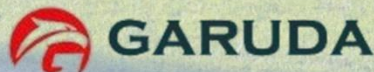


ISSN : 2089-0028
E-ISSN : 2654-7511

Jurnal Ilmiah Teori dan Aplikasi Statistika

Volume 15, No 1, JULI 2022



Penanggung Jawab	
Alfisyahrina Hapsery, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Pimpinan Redaksi	
Muhammad Athoillah, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Editor Pelaksana	
Sari Cahyaningtias, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Fenny Fitriani, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Nur Silviah Rahmi, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Sekar Utami Wijaya S.Stat., M.Si	(Universitas PGRI Adi Buana Surabaya)
Mitra Bestari (Reviewer)	
Dr.rer.pol. Dedy Dwi Prastyo, M.Si	(Institut Teknologi Sepuluh Nopember)
Dr. Drs. Agus Suharsono, M.S	(Institut Teknologi Sepuluh Nopember)
Dr. Bambang Widjanarko Otok	(Institut Teknologi Sepuluh Nopember)
Novri Suhermi, S.Si., M.Si., M.Sc	(Institut Teknologi Sepuluh Nopember)
Shofi Andari, S.Stat., M.Si	(Institut Teknologi Sepuluh Nopember)
Dr. RB Fajriya Hakim, S.Si., M.Si	(Universitas Islam Indonesia)
A'yunin Sofro, S.Si., M.Si., Ph.D.	(Universitas Negeri Surabaya)
Arief Rachman Hakim, S.Si., M.Si	(Universitas Diponegoro)
Dani Al Mahkya, S.Si., M.Si	(Sains Aktuaria Institut Teknologi Sumatra)
Dr. Sri Harini	(Universitas Islam Negeri Maulana Malik Ibrahim)
Dr. Faula Arina, M.Si	(Universitas Sultan Agung Tirtayasa)
Artanti Indrasetimingsih, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Gangga Anuraga, S.Si., M.Si	(Universitas PGRI Adi Buana Surabaya)
Winda Aprianti, S.Si., M.Si	(Politeknik Negeri Tanah Laut)

J Statistika: **Jurnal Ilmiah Teori dan Aplikasi Statistika** merupakan media penerbitan dan pembahasan karya ilmiah dalam bidang ilmu statistika beserta aplikasinya, baik berupa hasil penelitian, bahasan tentang teori, metodologi, komputasi maupun aplikasi statistika dalam bidang lain. Jurnal ini akan terbit dua kali dalam setahun setiap bulan Juli dan Desember.

Redaksi mengundang para pakar dan praktisi dari dalam dan luar lingkungan Universitas PGRI Adi Buana Surabaya untuk menuliskan karya ilmiahnya yang relevan dengan bidang ilmu statistika. Naskah dikirim pada website Jstatistika <http://jurnal.unipasby.ac.id/> dengan format A4 margin 4-3-3-3 times new roman 10, spasi 1, disertai biodata penulis. Email redaksi: jstat@unipasby.ac.id. Cp: 085645358346

Pengantar Redaksi

Syukur Alhamdulillah bahwa Jurnal Ilmiah: JStatistika, Program Studi Statistika Universitas PGRI Adi Buana Surabaya Volume 15 Nomor 1 dapat terbit bulan Juli 2022 ini. Jurnal Ilmiah JStatistika terbitan kali ini memuat artikel yang membahas tentang **“Pemodelan Persentase Kepesertaan Bpjs Non Penerima Bantuan Iuran Dengan Pendekatan Regresi Data Panel; Pemodelan Kejadian Balita Stunting di Kabupaten Bojonegoro dengan Metode Geographically Weighted Regression dan Multivariate Adaptive Regression Splines; Analisis Regulasi Uji Emisi Gas Buang Kendaraan Berdasarkan Pengaruhnya Terhadap Indeks Kualitas Udara di DKI Jakarta Menggunakan Metode Korelasi Pearson dan Regresi Linear; Estimasi Interval Kredibel Distribusi Normal Terpotong Kiri pada Data Waktu Sembuh Pasien Covid-19; Hotel Recommendation System with Content-Based Filtering Approach (Case Study: Hotel in Yogyakarta on Nusatrip Website); Spatial Autoregressive Moving Average Pada Pemodelan Persentase Penduduk Miskin Kabupaten/Kota Di Provinsi Jawa Timur Tahun 2020; Pemetaan Cakupan Pengguna Kb Aktif Dan Unmet Need Di Kabupaten Ponorogo Tahun 2021; Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan; Penerapan Text Mining pada Analisis Sentimen Pengguna Twitter Layanan Transportasi Online Menggunakan Metode Density Based Spatial Clustering of Applications with Noise (DBSCAN) dan K-Means; Peramalan Nilai Ekspor Provinsi Jawa Tengah dengan Metode Fuzzy Time Series Berbasis Algoritma Novel; Prediksi Pengeluaran Non Makanan Masyarakat Kabupaten Tulang Bawang Menggunakan Metode Rantai Markov”**.

Jurnal Ilmiah Statistika menerima artikel ilmiah dari hasil penelitian, laporan / studi kasus, kajian/ tinjauan pustaka, maupun penyegar ilmu statistika, yang berorientasi pada kemuthakhiran ilmu pengetahuan statistika agar dapat menjadi sumber informasi ilmiah yang mampu memberikan kontribusi dalam mengatasi permasalahan yang semakin kompleks. Redaksi mengundang berbagai ilmuwan dari berbagai lembaga pendidikan tinggi maupun penelitian untuk memberikan sumbangan ilmiahnya, baik berupa hasil penelitian maupun kajian ilmiah dalam bidang statistika dan aplikasinya.

Redaksi sangat mengharapkan masukan-masukan dari para pembaca demi semakin meningkatnya kualitas jurnal sebagaimana harapan kita bersama. Semoga artikel ilmiah yang termuat dalam Jurnal Ilmiah JStatistika bermanfaat bagi para akademisi dan professional yang berkecimpung dalam dunia statistika dan aplikasinya.

Surabaya, Juli 2022

Pimpinan Redaksi

DAFTAR ISI

HALAMAN JUDUL
PENGANTAR REDAKSI
DAFTAR ISI

- ❑ **Pemodelan Persentase Kepesertaan Bpjs Non Penerima Bantuan Iuran Dengan Pendekatan Regresi Data Panel**
Dhyana Venosia¹, Suliyanto², Sediono³, Nur Chamidah⁴
Universitas Airlangga 116 - 126

- ❑ **Pemodelan Kejadian Balita Stunting di Kabupaten Bojonegoro dengan Metode Geographically Weighted Regression dan Multivariate Adaptive Regression Splines**
Alif Yuanita Kartini¹, Laelatul Nur Ummah²
Universitas Nahdlatul Ulama Sunan Giri 127 - 136

- ❑ **Analisis Regulasi Uji Emisi Gas Buang Kendaraan Berdasarkan Pengaruhnya Terhadap Indeks Kualitas Udara di DKI Jakarta Menggunakan Metode Korelasi Pearson dan Regresi Linear**
M. Reza Akbar¹, M. Ihsan Akbar², dan Rizki Achmad D.³
Universitas Singaperbangsa Karawang dan Institut Teknologi Sains Bandung 137 - 146

- ❑ **Estimasi Interval Kredibel Distribusi Normal Terpotong Kiri pada Data Waktu Sembuh Pasien Covid-19**
Putri Fardha¹, Ardi Kurniawan², Sediono³, Elly Ana⁴, M. Fariz FM⁵
Universitas Airlangga 147 - 151

- ❑ **Hotel Recommendation System with Content-Based Filtering Approach (Case Study: Hotel in Yogyakarta on Nusatrip Website)**
Cheryl Ayu¹, Ayundyah K², RB Fajriya Hakim³, Arum Handini P⁴
Universitas Islam Indonesia 152 - 157

- ❑ **Spatial Autoregressive Moving Average Pada Pemodelan Persentase Penduduk Miskin Kabupaten/Kota Di Provinsi Jawa Timur Tahun 2020**
Kiki Laila Nirmala¹, Wara Pramesti², Fenny Fitriani³
Universitas PGRI Adi Buana Surabaya 158 - 166

- ❑ **Pemetaan Cakupan Pengguna Kb Aktif Dan Unmet Need Di Kabupaten Ponorogo Tahun 2021**
Bella Dwi Saputri¹, Diah Indriani²
Universitas Airlangga 167 - 177

- ❑ **Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan**
Weksi Budiaji¹, Juwarin Pancawati²
Universitas Sultan Ageng Tirtayasa 178 - 183

- ❑ **Penerapan Text Mining pada Analisis Sentimen Pengguna Twitter Layanan Transportasi Online Menggunakan Metode Density Based Spatial Clustering of Applications with Noise (DBSCAN) dan K-Means**
Dini Krisnawati Alfiki Astutik¹, Artanti Indrasetianingsih², Fenny Fitriani³
Universitas PGRI Adi Buana Surabaya 184 - 194

- ❑ **Peramalan Nilai Ekspor Provinsi Jawa Tengah dengan Metode Fuzzy Time Series Berbasis Algoritma Novel**
Septi Winda Utami¹, Indah Manfaati Nur², M. Al Haris³
Universitas Muhammadiyah Semarang 195 - 202

- ❑ **Prediksi Pengeluaran Non Makanan Masyarakat Kabupaten Tulang Bawang Menggunakan Metode Rantai Markov**
Ahmad Rizki Wiranto¹, Ahmad Faisol², Fitriani³
Universitas Lampung 203 - 209

Pemodelan Persentase Kepesertaan Bpjs Non Penerima Bantuan Iuran Dengan Pendekatan Regresi Data Panel

Dhyana Venosia⁽¹⁾, Suliyanto⁽²⁾, Sediono⁽³⁾, Nur Chamidah^{(4)*}

1,2,3,4Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas Airlangga

Jl. Dr. Ir. H. Soekarno, Mulyorejo, Kec. Mulyorejo, Kota Surabaya

e-mail: dhyanavenosia02@gmail.com, yanfit@yahoo.com, sediono101@gmail.com, dan nur-c@fst.unair.ac.id

ABSTRAK

Indonesia merupakan negara pengembang konsep *Universal Health Coverage* (UHC) pada Sistem Jaminan Sosial Nasional (SJSN) melalui program Jaminan Kesehatan Nasional (JKN) yang dikelola Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan. Peserta JKN, terbagi menjadi Penerima Bantuan Iuran (PBI) dan Non Penerima Bantuan Iuran (Non PBI). Tujuan dari penelitian ini yaitu, menganalisis faktor yang mempengaruhi persentase kepesertaan BPJS Non PBI khususnya Provinsi Jawa Timur tahun 2017 hingga 2020 yang mengalami fluktuasi. Maka, dalam mengestimasi fenomena tersebut menggunakan metode regresi data panel melalui pendekatan *Fixed Effect Model* (FEM) karena metode ini dapat mengakomodasi data *cross-section* dan *time series* dengan menerapkan adanya perbedaan *intercept* unit *cross-section* namun *slope* tetap diasumsikan konstan. Maka, secara statistik diperoleh kesimpulan bahwa yang berpengaruh signifikan adalah persentase penduduk miskin dan Tingkat Pengangguran Terbuka (TPT).

Kata kunci : BPJS Kesehatan Non PBI, Regresi Data Panel, *Fixed Effect Model*, Provinsi Jawa Timur.

ABSTRACT

Indonesia is a developing country in terms of the *Universal Health Coverage* (UHC) concept in the *National Social Security System* (NSSS), which is implemented through the *National Health Insurance* (NHI) program, which is administered by the *Health Social Security Administering Body* (HSSAB). *Contribution Assistance Recipients* (CAR) and *Non Contribution Assistance Recipients* (Non CAR) are two types of JKN participants. The goal of this study is to look at the factors that influence the percentage of HSSAB Non CAR participants, particularly in East Java Province, which has changed from 2017 to 2020. The panel data regression method uses the *Fixed Effect Model* (FEM) approach to estimate this phenomena since this method can accommodate *cross-section* and *time series* data by using the difference in *intercept* unit *cross-section* while the *slope* remains constant. As a result, statistically, the percentage of poor individuals and the *Open Unemployment Rate* (OUR) have a substantial effect.

Keywords : East Java Province, HSSAB Health Non CAR, Panel Data Regression, *Fixed Effect Model*.

1. PENDAHULUAN

Indonesia merupakan negara pengembang konsep *Universal Health Coverage* (UHC) di sektor kesehatan yang diterapkan pada Sistem Jaminan Sosial Nasional (SJSN) (Oldistra & Machdun, 2020). Kusumaningrum & Azinar (2018) mengemukakan bahwa, pada SJSN termuat Jaminan Kesehatan Nasional (JKN) yang berperan sebagai program khusus dalam perlindungan kesehatan dengan melibatkan Badan Penyelenggara Jaminan Sosial (BPJS) selaku pengelola. Menurut Kementerian Kesehatan, partisipan JKN dibedakan

menjadi Penerima Bantuan Iuran (PBI) dan Non Penerima Bantuan Non Iuran (Non PBI).

Menurut Badan Pusat Statistik (BPS) (2021a), rerata persentase kepesertaan BPJS Kesehatan Non PBI di Indonesia tahun 2017-2020 mengalami peningkatan sejumlah 2,52%. Namun, masih jauh dari target pemerintah. Hal ini, sebagaimana yang terjadi di Provinsi Jawa Timur dengan realita tahun 2017-2020 sebesar 17,165% dan rerata target pemerintah sebesar 82,835%. Sehingga, Provinsi Jawa Timur menempati 10 besar persentase kepesertaan BPJS Kesehatan Non PBI terendah.

Dalam meningkatkan kepesertaan BPJS kesehatan, sejauh ini upaya pemerintah yaitu melakukan sosialisasi rutin serta menerapkan prosedur pendaftaran dan pembayaran dengan mudah (Putro & Barida, 2017). Menurut Suprianto & Mutiarin (2017), pemerintah telah membuat model pembiayaan BPJS cukup ideal. Selain itu, pemerintah telah membatalkan kenaikan premi kelas III dan melakukan peningkatan kualitas pelayanan tingkat pertama dan tingkat lanjut. Upaya tersebut dirasa belum optimal, terbukti pada kenaikan persentase kepesertaan BPJS Kesehatan Non PBI Provinsi Jawa Timur tahun 2017-2020 secara berturut-turut hanya sejumlah 3.83%; 3.69%; dan 0.83% (BPS, 2021a).

Rendahnya kenaikan, diakibatkan karena terjadinya fluktuasi di Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020, seperti Kabupaten Pacitan dan Pamekasan dan Kota Kediri. Kabupaten Pacitan mengalami kenaikan sejumlah 0.82%; 4.65%; dan 0.56%. Kemudian, Kabupaten Pamekasan mengalami kemerosotan sejumlah 5.67%; 0.96%; dan 0.25%. Selanjutnya, Kota Kediri mengalami kemerosotan sejumlah 5.09%; kenaikan sejumlah 7.48%; dan kemerosotan kembali sejumlah 1.04% (BPS, 2021a). Fenomena tersebut membuktikan terjadinya dinamika persentase kepesertaan BPJS Kesehatan Non PBI di Kabupaten/Kota Provinsi Jawa Timur. Sehingga, pendugaan terkait faktor yang diasumsikan mempengaruhi diperlukan metode yang dapat mengakomodasi terjadinya dinamika. Maka, metode regresi data panel dirasa telah sesuai. Metode regresi data panel merupakan metode yang dapat mengakomodasi mengenai penggabungan antara data *cross section* dan *time series*.

Penelitian terkait telah dilakukan oleh Riza *et al.* (2020), bahwa penyebab ketidakpatuhan masyarakat disebabkan oleh kemiskinan. Selain itu, pemanfaatan JKN dilakukan masyarakat dengan mendaftar secara mandiri ketika menderita sakit. (Litawati, 2016). Menurut Putra & Arka (2018), Tingkat Pengangguran Terbuka (TPT) merupakan faktor dominan yang mempengaruhi persentase kepesertaan BPJS Kesehatan Non PBI. Maka, pendapatan dengan keikutsertaan menjadi peserta BPJS kesehatan memiliki keterkaitan (Abadi *et al.*, 2019). Penerapannya, aspek partisipasi masyarakat merupakan penentu keberhasilan JKN (Oldistra & Machdun, 2020).

Berdasarkan uraian di atas, fenomena tersebut diduga dipengaruhi oleh aspek ekonomi yaitu persentase penduduk miskin, persentase penduduk yang mempunyai keluhan kesehatan, TPT, dan

pendapatan yang melibatkan Kabupaten/Kota Provinsi Jawa Timur sebagai unit *cross section* dan rentang waktu pada tahun 2017-2020 sebagai unit *time series*. Penerapan metode regresi data panel akan menjadi sebuah kebaruan dalam menganalisis kepesertaan BPJS Non PBI. Selain itu, keunggulan yang diberikan dalam penelitian ini, secara khusus memberikan informasi terkait fluktuasi persentase kepesertaan BPJS Non PBI yang belum pernah ditemukan pada penelitian sebelumnya. Kemudian, diharapkan hasil dari penelitian ini mampu memberikan suatu prediksi yang berguna dalam mengoptimalkan program pemerintah.

2. METODE PENELITIAN

Persentase Kepesertaan BPJS Non Penerima Bantuan Iuran (Non PBI).

Menurut Kementerian Kesehatan, persentase kepesertaan BPJS Non PBI merupakan 1/100 peserta program JKN yang telah membayar iuran secara mandiri setiap bulannya (BPS, 2021b). Keterkaitan antara JKN dengan SJSN sebagaimana yang dikemukakan oleh Kusumaningrum & Azinar (2018), pada SJSN termuat Jaminan Kesehatan Nasional (JKN) yang berperan sebagai program khusus perlindungan kesehatan dengan melibatkan Badan Penyelenggara Jaminan Sosial (BPJS) selaku pengelola. Pengelolaan SJSN oleh BPJS sebagaimana memiliki tujuan yang telah sesuai dengan UU RI Nomor 40 Tahun 2004.

Faktor-Faktor yang Mempengaruhi Persentase Kepesertaan BPJS Non PBI

Berdasarkan latar belakang yang telah dipaparkan, peneliti menduga aspek ekonomi seperti persentase penduduk miskin, persentase penduduk yang mempunyai keluhan kesehatan, TPT, dan pendapatan mempengaruhi persentase kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur.

Persentase Penduduk Miskin

Menurut BPS (2021a), persentase penduduk miskin ialah 1/100 bagian dari penduduk dengan rerata pengeluaran per kapita untuk setiap bulan berkisar di bawah garis kemiskinan. Menurut Riza *et al.* (2020), kemiskinan merupakan penyebab ketidakpatuhan peserta Jaminan Kesehatan Berbasis Masyarakat (JKBM). Karena, apabila penduduk miskin menjadi peserta BPJS Non PBI, mereka harus mengurangi konsumsi untuk barang lain karena ketidakmampuannya dalam memenuhi kebutuhannya sekaligus. Masyarakat miskin

beranggapan bahwa masih banyak kebutuhan lain yang wajib terpenuhi dibanding menjadi partisipan keikutsertaan BPJS Non PBI yang harus membayar premi setiap bulannya.

Persentase Penduduk yang Mempunyai Keluhan Kesehatan

Menurut BPS (2021b), persentase dari penduduk yang mempunyai masalah mengenai keluhan kesehatan ialah 1/100 bagian dari penduduk yang mempunyai keluhan kesehatan secara fisik maupun psikis. Pemanfaatan program JKN dengan cara mendaftar secara mandiri ketika sudah menderita sakit atau mempunyai keluhan kesehatan mengindikasikan rendahnya pemahaman masyarakat mengenai pentingnya jaminan sosial. Maka, dapat dikatakan bahwa kenaikan persentase kepesertaan BPJS Non PBI memiliki hubungan linier dengan persentase penduduk yang terindikasi mempunyai keluhan kesehatan, namun belum menjadi anggota BPJS Non PBI (Litawati, 2016).

Tingkat Pengangguran Terbuka (TPT)

Menurut BPS (2021a), pendefinisian dari TPT ialah persentase antara jumlah pengangguran atas jumlah angkatan kerja. Menurut Litawati (2016), berkurangnya angka pengangguran mengakibatkan penurunan terhadap angka kemiskinan. Dalam hal ini, pengangguran diduga sebagai faktor yang dapat mempengaruhi fenomena terjadinya kemiskinan. Menurut Kementerian Pekerjaan Umum dan Perumahan Rakyat (2014), ditegaskan bahwa tidak semua pengangguran mengalami kondisi kemiskinan. Justru, saat ini semakin banyak ditemui pengangguran yang berkecukupan (*discourage worker*). Hal ini, diindikasikan dengan peningkatan investasi dari tahun 2016-2020. Dalam hal ini, seorang investor tetap hidup berkecukupan walaupun tidak memiliki pekerjaan. Maka, dapat dikatakan seorang investor merupakan bagian dari pengangguran, sehingga dapat mempengaruhi kenaikan kepesertaan BPJS Non PBI.

Pendapatan

Menurut Ham *et al.* (2020), pendapatan ialah arus masuk aktiva atas penyelesaian mengenai pemberian jasa atau aktivitas yang telah dilakukan. Riza *et al.* (2020), mengemukakan bahwa antara pendapatan dengan kepesertaan pada BPJS kesehatan mandiri memiliki keterkaitan. Menurut Nelisma *et al.* (2019), pendapatan ialah faktor pendukung terhadap kenaikan suatu layanan kesehatan. Hal ini terjadi karena, masyarakat pemilik pendapatan yang rendah atau terbatas berasumsi tidak memiliki kemampuan dalam

membayar premi BPJS kesehatan yang harus dibayarkan rutin setiap bulannya. Bagi mereka, dalam kondisi seperti ini kebutuhan konsumsi pangan dianggap lebih dominan (Abadi *et al.*, 2019). Dalam hal ini, pendapatan memegang pengaruh terjadinya perubahan perilaku masyarakat terhadap kepesertaan BPJS Non PBI (Pangestika *et al.*, 2017).

Regresi Data Panel

Sriyana (2014) menegaskan bahwa metode regresi data panel merupakan metode regresi yang menerapkan penggunaan konsep penggabungan antara pengamatan data pada unit *cross section* dan *time series* di sebuah persamaan seperti yang disajikan sebagai berikut:

$$Y_{it} = \alpha_{it} + X_{it}\beta + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (1)$$

Berdasarkan persamaan di atas, Hsiao (2003) mengelompokkan model regresi data panel menjadi 5 kelompok, sebagai berikut:

$$1. Y_{it} = \alpha^* + \sum_{k=1}^K \beta_k X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (2)$$

$$2. Y_{it} = \alpha_i^* + \sum_{k=1}^K \beta_k X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (3)$$

$$3. Y_{it} = \alpha_{it}^* + \sum_{k=1}^K \beta_k X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (4)$$

$$4. Y_{it} = \alpha_i^* + \sum_{k=1}^K \beta_{ki} X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (5)$$

$$5. Y_{it} = \alpha_{it}^* + \sum_{k=1}^K \beta_{kit} X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, N \quad (6)$$

Dalam penerapannya, metode regresi data panel memiliki keunggulan antara lain, data panel mampu menjelaskan dua macam informasi terkait dinamika yang berasal dari unit *cross section* dan *time series* sekaligus sehingga prediksi yang dihasilkan lebih akurat, efektif, sekaligus efisien (Sriyana, 2014). Menurut Rahmadeni & Wulandari (2017), pengestimasi model pada regresi data panel dikelompokkan menjadi, *Common Effect Model* (CEM), *Fixed Effect Model* (FEM), dan *Random Effect Model* (REM).

Common Effect Model (CEM)

Sriyana (2014) mengemukakan bahwa pengasumsian pendekatan CEM yaitu, dengan menganggap nilai *intercept* setiap variabel sama, begitu pula nilai pada koefisien *slope* bagi semua unit *cross section* dan *time series*. Dalam mengestimasi parameter CEM, dapat menggunakan *Ordinary Least Square* (OLS) (Rahmadeni & Wulandari, 2017). Maka, persamaan mengenai model CEM, dapat dinyatakan sebagai berikut:

$$Y_{it} = \alpha + X_{it}\beta + u_{it}; i = 1, \dots, N; t = 1, \dots, T \quad (7)$$

dengan penduga OLS sebagai berikut:

$$\hat{\alpha} = \frac{1}{NT} \mathbf{e}'_{NT} [\mathbf{Y} - \mathbf{X}\hat{\beta}] \quad (8)$$

$$\hat{\beta} = [\mathbf{X}'\mathbf{W}_{NT}\mathbf{X}]^{-1}\mathbf{X}'\mathbf{W}_{NT}\mathbf{Y} \quad (9)$$

Fixed Effect Model (FEM)

Rahmadeni & Wulandari (2017) mengemukakan bahwa berdasarkan pendekatan FEM terkait nilai *intercept* pada setiap variabel diasumsikan berbeda pada setiap unit *cross section* yang mana *slope* masih terasumsikan konstan. Dalam mengestimasi parameter FEM, dapat menggunakan OLS, dengan penduga OLS yang disebut *Least Square Dummy Variabel (LSDV)*. Maka, persamaan mengenai model FEM, dapat dinyatakan sebagai berikut:

$$Y_{it} = \alpha_i + \mathbf{X}_{it}\boldsymbol{\beta} + u_{it}; i = 1, \dots, N; t = 1, \dots, T \quad (10)$$

dengan penduga OLS sebagai berikut:

$$\hat{\alpha} = [\mathbf{D}'_N\mathbf{D}_N]^{-1}\mathbf{D}'_N[\mathbf{Y} - \mathbf{X}\hat{\beta}] \quad (11)$$

$$\hat{\beta} = [\mathbf{X}'\mathbf{W}_N\mathbf{X}]^{-1}\mathbf{X}'\mathbf{W}_N\mathbf{Y} \quad (12)$$

Random Effect Model (REM)

Rahmadeni & Wulandari (2017) mengemukakan bahwa REM mengasumsikan *effect* individu untuk seluruh unit *cross section* bersifat *random*. Parameter REM dapat diestimasi menggunakan *Generalized Least Square (GLS)*, dengan persamaan mengenai model REM yang dapat dituliskan sebagai berikut:

$$Y_{it} = \beta_0 + \sum_{k=1}^K \beta_k X_{kit} + u_{it}; i = 1, \dots, N; t = 1, \dots, T \quad (13)$$

Uji Spesifikasi Model

Menurut Ghazi & Hermansyah (2018), proses dalam pengestimasi regresi data panel didasarkan melalui model terbaik yang diperoleh melalui beberapa pengujian sebagai berikut:

Uji Chow

Uji *Chow* bertujuan untuk menentukan model estimasi yang sesuai antara model CEM atau model FEM melalui pengujian sebagai berikut:

$$F = \frac{\frac{SSE_{CEM} - SSE_{FEM}}{N-1}}{\frac{SSE_{FEM}}{NT - (N(K-1))}} \quad (14)$$

Uji Hausman

Uji *Hausman* diaplikasikan dalam menentukan estimasi berdasarkan model yang sesuai antara model FEM atau model REM melalui pengujian sebagai berikut:

$$W = [\hat{\beta}_{FEM} - \hat{\beta}_{REM}]' [\widehat{var}\hat{\beta}_{FEM} - \widehat{var}\hat{\beta}_{REM}]^{-1} [\hat{\beta}_{FEM} - \hat{\beta}_{REM}] \quad (15)$$

Uji Lagrange Multiplier

Uji *Lagrange Multiplier* bertujuan dalam menentukan suatu model estimasi yang sesuai antara model REM atau model CEM melalui pengujian sebagai berikut:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N [\sum_{t=1}^T e_{it}]^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} - 1 \right]^2 \quad (16)$$

Uji Asumsi Klasik

Menurut Ghazi & Hermansyah, (2018), estimasi model regresi yang bersifat linier tidak bias (*Best Linier Unbiased Estimator (BLUE)*) diperoleh melalui model regresi yang telah lolos pengujian asumsi klasik. Sehingga, model regresi dapat dikatakan baik. Berikut merupakan beberapa uji asumsi klasik yang digunakan:

Uji Normalitas

Menurut Yulianto *et al.* (2018), uji normalitas bertujuan untuk mendeteksi normalitas nilai residual dalam membuktikan data penelitian yang digunakan telah berasal dari populasi berdistribusi normal. Terkait pernyataan tersebut, normalitas residual dapat diuji melalui uji *Kolmogorov-Smirnov* melalui pengujian sebagai berikut:

$$D^+ = \sup_Y \left| \frac{1}{N} - F_0(Z_i) \right| \quad \text{dan} \quad D^- = \sup_Y \left| \frac{i-1}{N} - F_0(Z_i) \right| \quad (17)$$

Uji Multikolinieritas

Menurut Rahmadeni & Wulandari (2017), uji multikolinieritas digunakan sebagai estimator pendeteksi korelasi antar variabel independen yang ditinjau berdasarkan dari *Variance Inflation Factor (VIF)* yang diperoleh melalui pengujian sebagai berikut:

$$VIF = \frac{1}{tolerance} = \frac{1}{1-R_k^2} \quad (18)$$

Uji Heteroskedastisitas

Menurut Baum (2001), heteroskedastisitas bertujuan untuk mendeteksi masalah *heteroskedasticity* yang dikenal dengan nama *groupwise heteroskedasticity*. Dalam mendeteksi masalah heteroskedastisitas dapat dideteksi melalui *Wald Test* yang telah dimodifikasi untuk perhitungan *groupwise heteroskedasticity* melalui model FEM melalui pengujian sebagai berikut:

$$W = \sum_{i=1}^N \frac{(\hat{\sigma}_i^2 - \hat{\sigma}^2)^2}{V_i} \quad (19)$$

Uji Autokorelasi

Menurut Khosropour (2017), uji autokorelasi bertujuan untuk mengatasi masalah terjadinya

korelasi serial pada data panel yang bersifat membiaskan kesalahan standar sehingga menyebabkan hasil pengujian menjadi kurang efisien. Dalam mendeteksi masalah korelasi serial dapat dideteksi melalui *Wooldridge Test* melalui pengujian sebagai berikut:

$$(Y_{it} - Y_{it-1}) = (X_{it} - X_{it-1})'b_1 + (\varepsilon_{it} - \varepsilon_{it-1}) \quad (20)$$

$$\Delta Y_{it} = \Delta X_{it} b_1 + \Delta \varepsilon_{it} \quad (21)$$

Koefisien Determinasi

Menurut Rahmadeni & Wulandari (2017), perhitungan nilai koefisien determinasi bertujuan untuk menganalisis tingkat kemampuan yang dimiliki variabel independen dalam menjelaskan mengenai variabel dependen melalui sebuah model. Nilai koefisien determinasi yang diperoleh dalam model regresi, berkisar antara 0 hingga 1. Menurut Gujarati (2003), nilai koefisien determinasi dapat diperoleh melalui perhitungan sebagai berikut:

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (22)$$

Uji Signifikansi Parameter

Pengujian signifikansi parameter, terdiri dari 2 macam pengujian antara lain sebagai berikut:

Uji Serempak

Menurut Rahmadeni & Wulandari (2017), uji serempak digunakan menganalisis pengaruh pada seluruh variabel independen atas variabel dependen dalam model secara serempak. Menurut Gujarati (2003), nilai dari statistik uji *F* diperoleh dari:

$$F = \frac{\frac{R^2}{(K-1)}}{\frac{(1-R^2)}{(N-K)}} \quad (23)$$

Uji Parsial

Menurut Rahmadeni & Wulandari (2017), uji parsial digunakan menganalisis pengaruh secara parsial setiap variabel independen atas variabel dependen dalam model. Menurut Gujarati (2003), nilai dari statistik uji *t* diperoleh dari:

$$t = \frac{\hat{\beta}_i}{se(\hat{\beta}_i)} \quad (24)$$

Data dan Sumber Data

Data terkait penelitian ialah data sekunder tentang Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020, dengan rincian data beserta sumber datanya sebagai berikut:

Data	Sumber Data
- Persentase kepesertaan BPJS Non Penerima Bantuan Iuran di setiap Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020	Publikasi Badan Pusat Statistik Provinsi Jawa Timur, yang berjudul: - Provinsi Jawa Timur Dalam Angka 2018 - Provinsi Jawa Timur Dalam Angka 2019 - Provinsi Jawa Timur Dalam Angka 2020 - Provinsi Jawa Timur Dalam Angka 2021
- Persentase penduduk miskin di setiap Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020	
- Persentase penduduk yang mempunyai keluhan kesehatan di setiap Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020	
- TPT di setiap Kabupaten/Kota Provinsi Jawa Timur tahun 2017-2020	
Pendapatan di setiap Kabupaten/Kota Provinsi Jawa Timur	- Peraturan Gubernur Jawa Timur Nomor 121 Tahun 2016 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2017 - Peraturan Gubernur Jawa Timur Nomor 75 Tahun 2017 Tentang Upah minimum Kabupaten/Kota di Jawa Timur Tahun 2018 - Keputusan Gubernur Jawa Timur Nomor 188/665/KPTS/013/2018 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2019 - Keputusan Gubernur Jawa Timur Nomor 188/568/KPTS/013/2019 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2020

Gambar 1. Data dan Sumber Data

Variabel Penelitian

Penggunaan variabel pada penelitian ini terbagi menjadi variabel independen yang menyatakan variabel prediktor dan variabel dependen yang menyatakan variabel respon, dengan rincian sebagai berikut:

Variabel	Keterangan	Skala Pengukuran	Tipe Variabel
Y	Persentase kepesertaan BPJS Non PBI di setiap Kabupaten/Kota Provinsi Jawa Timur (%)	Rasio	Kontinu
X1	Persentase penduduk miskin (%)	Rasio	Kontinu
X2	Persentase penduduk yang mempunyai keluhan kesehatan (%)	Rasio	Kontinu
X3	TPT (%)	Rasio	Kontinu
X4	Pendapatan (Rupiah)	Rasio	Kontinu

Gambar 2. Variabel Penelitian

Berdasarkan data penelitian yang digunakan, maka diperoleh struktur data panel sebagai berikut:

Subjek Pengamatan	Tahun Pengamatan	Variabel Dependen	Variabel Independen			
			X1	X2	X3	X4
Kabupaten/Kota ke-1	2017	Y _{1,1}	X _{1,1,1}	X _{2,1,1}	X _{3,1,1}	X _{4,1,1}
	2020	Y _{1,4}	X _{1,1,4}	X _{2,1,4}	X _{3,1,4}	X _{4,1,4}
Kabupaten/Kota ke-2	2017	Y _{2,1}	X _{1,2,1}	X _{2,2,1}	X _{3,2,1}	X _{4,2,1}
	2020	Y _{2,4}	X _{1,2,4}	X _{2,2,4}	X _{3,2,4}	X _{4,2,4}
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Kabupaten/Kota ke-38	2017	Y _{38,1}	X _{1,38,1}	X _{2,38,1}	X _{3,38,1}	X _{4,38,1}
	2020	Y _{38,4}	X _{1,38,4}	X _{2,38,4}	X _{3,38,4}	X _{4,38,4}

Gambar 3. Struktur Data Panel

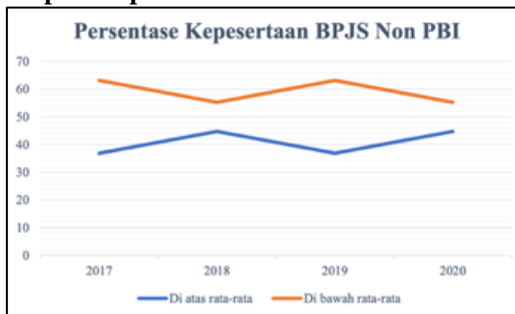
3. HASIL DAN PEMBAHASAN

Berdasarkan data penelitian, maka diperoleh hasil analisis sebagai berikut:

Deskripsi Variabel Penelitian

Dalam mendeskripsikan variabel penelitian mengenai persentase kepesertaan BPJS Non PBI di setiap Kabupaten/Kota Provinsi Jawa Timur beserta variabel-variabel yang diduga berpengaruh diselesaikan dengan menggunakan statistika deskriptif yang disajikan pada **Gambar 4.** hingga **Gambar 8.** sebagai berikut:

Persentase Kepesertaan BPJS Non PBI di Setiap Kabupaten/Kota Provinsi Jawa Timur



Gambar 4. Grafik Persentase Kepesertaan BPJS Non PBI

Berdasarkan **Gambar 4.**, diketahui bahwa rata-rata Kabupaten/Kota Provinsi Jawa Timur dengan persentase kepesertaan BPJS Non PBI di atas rata-rata pada tahun 2017 sebesar 36.84%; tahun 2018 sebesar 44.74%; tahun 2019 sebesar 36.84%; dan tahun 2020 sebesar 44.74%.

Persentase kepesertaan BPJS Non PBI terendah tahun 2017 sebesar 4.20% dan tahun 2018 sebesar 4.36% terdapat di Kabupaten Bangkalan. Kemudian, pada tahun 2019 sebesar 5.66% dan tahun 2020 sebesar 7.08% terdapat di Kabupaten Sampang. Kemudian, untuk persentase kepesertaan BPJS Non PBI tertinggi tahun 2017 sebesar 27.74% terdapat di Kota Madiun. Pada tahun 2018 sebesar 36.51%; tahun 2019 sebesar 43.76%; dan tahun 2020 sebesar 45.51% terdapat di Kabupaten Sidoarjo.

Selain itu, informasi yang dapat diperoleh berdasarkan **Gambar 4.** yaitu, rata-rata persentase kepesertaan BPJS Non PBI di setiap Kabupaten/Kota Provinsi Jawa Timur mengalami fluktuasi mulai tahun 2017 hingga tahun 2020.

Persentase Penduduk Miskin



Gambar 5. Grafik Persentase Penduduk Miskin

Berdasarkan **Gambar 5.**, diketahui bahwa rata-rata Kabupaten/Kota Provinsi Jawa Timur dengan persentase penduduk miskin di atas rata-rata pada tahun 2017 hingga tahun 2019 sebesar 44.73% dan tahun 2020 sebesar 47.37%.

Persentase penduduk miskin terendah tahun 2017 sebesar 4.31%; tahun 2018 sebesar 3.89%; tahun 2019 sebesar 3.81%; dan tahun 2020 sebesar 3.89% terdapat di Kota Batu. Kemudian, untuk persentase penduduk miskin tertinggi tahun 2017 sebesar 23.56%; tahun 2018 sebesar 21.21%; tahun 2019 sebesar 20.71%; dan tahun 2020 sebesar 22.78% terdapat di Kabupaten Sampang.

Selain itu, informasi yang dapat diperoleh berdasarkan **Gambar 5.** yaitu, rata-rata persentase penduduk miskin di atas rata-rata pada tahun 2020 mengalami peningkatan yang mengindikasikan bahwa persentase penduduk miskin di Provinsi Jawa Timur mengalami peningkatan dari 3 tahun sebelumnya.

Persentase Penduduk yang Mempunyai Keluhan Kesehatan



Gambar 6. Grafik Persentase Penduduk yang Mempunyai Keluhan Kesehatan

Berdasarkan **Gambar 6.**, diketahui bahwa rata-rata Kabupaten/Kota Provinsi Jawa Timur dengan persentase penduduk yang mengalami keluhan kesehatan di atas rata-rata pada tahun 2017 sebesar 47.37%; tahun 2018 sebesar 34.21%; tahun 2019 sebesar 52.63%; dan tahun 2020 sebesar 44.74%.

Persentase penduduk yang mempunyai keluhan kesehatan terendah tahun 2017 sebesar 16.54%; tahun 2018 sebesar 19.88%; dan tahun 2019 sebesar 18.64% terdapat di Kabupaten Bangkalan. Sedangkan, untuk persentase penduduk yang mempunyai keluhan kesehatan terendah tahun 2020 sebesar 17.18% terdapat di Kabupaten Sumenep. Kemudian, untuk persentase penduduk yang mempunyai keluhan kesehatan tertinggi tahun 2017 sebesar 47.37% terdapat di Kabupaten Trenggalek; tahun 2018 sebesar 52.15% dan tahun 2019 sebesar 49.49% terdapat di Kabupaten Blitar.

Sedangkan untuk persentase penduduk yang mengalami keluhan kesehatan tertinggi tahun 2020 sebesar 46.25% terdapat di Kabupaten Ponorogo.

Selain itu, informasi yang dapat diperoleh berdasarkan Gambar 6. yaitu, rata-rata persentase penduduk yang mempunyai keluhan kesehatan di setiap Kabupaten/Kota Provinsi Jawa Timur mengalami fluktuasi mulai tahun 2017 hingga tahun 2020.

Tingkat Pengangguran Terbuka (TPT)



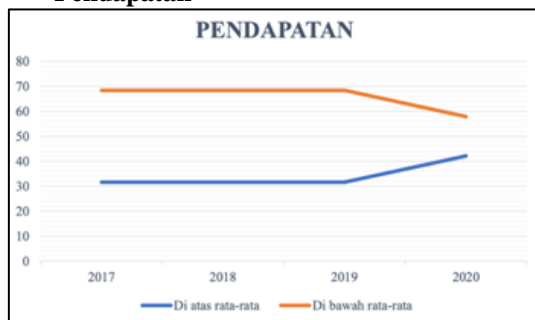
Gambar 7. Grafik Tingkat Pengangguran Terbuka

Berdasarkan Gambar 7., diketahui bahwa rata-rata Kabupaten/Kota Provinsi Jawa Timur dengan TPT di atas rata-rata pada tahun 2017 sebesar 47.37%; tahun 2018 sebesar 55.26%; tahun 2019 sebesar 44.74%; dan tahun 2020 sebesar 34.21%.

TPT terendah tahun 2017 sebesar 0.85%; tahun 2018 sebesar 1.39%; dan tahun 2019 sebesar 0.91%.; dan tahun 2020 sebesar 2.28% terdapat di Kabupaten Pacitan. Kemudian, untuk TPT tertinggi tahun 2017 sebesar 7.22%; tahun 2018 sebesar 6.65%; dan tahun 2019 sebesar 5.88% terdapat di Kota Malang. Sedangkan, untuk TPT tertinggi tahun 2020 sebesar 10.97% terdapat di Kabupaten Sidoarjo.

Selain itu, informasi yang dapat diperoleh berdasarkan Gambar 7. yaitu, rata-rata TPT di setiap Kabupaten/Kota Provinsi Jawa Timur mengalami fluktuasi mulai tahun 2017 hingga tahun 2020.

Pendapatan



Gambar 8. Grafik Tingkat Pengangguran Terbuka

Berdasarkan Gambar 8., diketahui bahwa rata-rata Kabupaten/Kota Provinsi Jawa Timur dengan pendapatan di atas rata-rata pada tahun 2017 hingga tahun 2019 sebesar 31.58%; dan tahun 2020 sebesar 42.11%.

Pendapatan terendah tahun 2017 sebesar Rp1388847.5 rupiah; tahun 2018 sebesar Rp1509816.12 terdapat di Kabupaten Pacitan, Ponorogo, Trenggalek, dan Magetan. Pada tahun 2019 sebesar Rp1763267,65 rupiah dan tahun 2020 sebesar Rp1913321.73 terdapat di Kabupaten Pacitan, Ponorogo, Trenggalek, Situbondo, Madiun, Magetan, Ngawi, Sampang, dan Pamekasan.

Estimasi Regresi Data Panel

Dalam penelitian ini, pengestimasi model regresi data panel terbagi menjadi 3, yaitu CEM, FEM, dan REM yang disajikan pada Tabel. 1 sebagai berikut:

Tabel 1. Estimasi Regresi Data Panel

	R-Sq	p-value	F
CEM	59.92%	0.0000	57.44
FEM	86.45%	0.0000	10.83
REM	59.82%	0.0000	111.29

Pemilihan Model Estimasi Terbaik

Berdasarkan model estimasi regresi data panel, dilakukan pemilihan model estimasi terbaik sebagaimana disajikan pada Tabel 2. sebagai berikut:

Tabel 2. Pemilihan Model Estimasi Terbaik

	p-value	Keputusan	Model Terpilih
Chow	0.0000	Tolak H_0	FEM
Hausman	0.0000	Tolak H_0	FEM
Lagrange Multiplier	0.0000	Tolak H_0	REM

Berdasarkan hasil pengujian dalam menentukan model estimasi terbaik melalui uji chow, hausman, dan lagrange multiplier, maka dapat disimpulkan bahwa model estimasi terbaik yang terpilih adalah FEM.

Uji Asumsi Klasik

Setelah menentukan model estimasi yang sesuai, akan dilanjutkan pada uji asumsi klasik sebagai berikut:

Uji Normalitas

Berdasarkan uji Kolmogorov-Smirnov, diperoleh nilai p-value sebesar 0.061. Sehingga,

keputusan yang terambil yaitu Terima H_0 . Maka, dapat diperoleh kesimpulan bahwa residual data telah berdistribusi normal.

Uji Multikolinieritas

Berdasarkan uji multikolinieritas, diperoleh nilai VIF < 10 pada setiap variabel independent penelitian, yang disajikan dalam **Tabel 3.**, sebagai berikut:

Memuat hasil dan pembahasan tentang penelitian yang dilakukan.

Tabel 3. Nilai VIF

	VIF	1/VIF
X1	5.05	0.197442
X2	9.36	0.106874
X3	9.36	0.106874
X4	3.52	0.283923

Maka, diperoleh kesimpulan bahwa pada model yang telah terpilih tidak terindikasi mengalami masalah multikolinieritas.

Uji Heteroskedastisitas

Berdasarkan *Walt test*, diperoleh nilai *p-value* sebesar 0.0000. Maka, keputusan yang terambil yaitu Tolak H_0 . Maka, dapat disimpulkan bahwa model yang telah terpilih terindikasi mengalami masalah heteroskedastisitas.

Uji Autokorelasi

Berdasarkan *Wooldridge test*, diperoleh nilai *p-value* sebesar 0.0000. Maka, keputusan yang terambil yaitu Tolak H_0 . Maka, dapat disimpulkan bahwa model yang telah terpilih terindikasi mengalami masalah autokorelasi.

Penanganan Pelanggaran Uji Asumsi Klasik

Berdasarkan hasil pengujian asumsi klasik, diketahui bahwa model yang telah sesuai terindikasi mengalami masalah heteroskedastisitas dan autokorelasi. Maka, dapat dilakukan penyelesaian masalah heteroskedastisitas dan autokorelasi sekaligus. Sehingga setelah dilakukan penanganan asumsi klasik, maka model estimasi telah terbebas dari masalah heteroskedastisitas dan autokorelasi sebagaimana disajikan dalam **Tabel 4.** sebagai berikut:

Tabel 4. Model Estimasi Setelah Penanganan

	Coef	p-value
X1	-2.832989	0.000
X2	-0.0082407	0.926

X3	1.565032	0.000
X4	-4.24e-09	0.989
Cons	50.06444	

Sehingga, bisa dilanjutkan pada pengujian selanjutnya.

Estimasi FEM

Pendekatan FEM yang digunakan, diasumsikan bahwa nilai *intercept* masing-masing variabel adalah berbeda pada setiap unit *cross section* namun *slope* masih diasumsikan konstan. Maka, nilai *intercept* pada pendekatan FEM untuk setiap unit *cross section* disajikan dalam **Tabel 5.** sebagai berikut:

Tabel 5. Intercept Unit Cross Section

Unit	Intercept
1	0
2	-14.67153
⋮	⋮
38	-15.54064

dengan, unit *cross section* pertama merupakan *benchmark* dari variabel *dummy*.

Koefisien Determinasi

Berdasarkan model terbaik yang telah terpilih yaitu FEM, diperoleh nilai R^2 sebesar 86.45%. Nilai ini, dapat diartikan dengan variabel persentase kepesertaan BPJS Non Non PBI di Kabupaten/Kota Provinsi Jawa Timur dapat dijelaskan oleh variabel persentase penduduk miskin, persentase penduduk yang mempunyai keluhan kesehatan, TPT, dan pendapatan sebesar 86.45%. Sedangkan, 13.65% mengenai variabel persentase kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur dapat dijelaskan melalui variabel lain yang berada di luar model.

Uji Signifikansi Parameter

Uji Serempak

Berdasarkan **Tabel 1.**, dapat disimpulkan bahwa uji serempak dalam model estimasi dapat terdapat pengaruh secara serempak antara persentase penduduk miskin, persentase penduduk yang mempunyai keluhan kesehatan, TPT, dan pendapatan terhadap persentase kepesertaan BPJS Non Non PBI di Kabupaten/Kota Provinsi Jawa Timur.

Uji Parsial

Setelah melakukan uji serempak, maka dilanjutkan dengan uji parsial sebagaimana disajikan pada **Tabel 6.** sebagai berikut:

Tabel 6. Uji Parsial

	<i>p-value</i>	Keputusan
$\hat{\beta}_1$	0.000	Tolak H_0
$\hat{\beta}_2$	0.926	Terima H_0
$\hat{\beta}_3$	0.000	Tolak H_0
$\hat{\beta}_4$	0.989	Terima H_0

Berdasarkan **Tabel 6.**, dapat diperoleh keputusan Tolak H_0 untuk $\hat{\beta}_1$ dan $\hat{\beta}_3$. Sehingga kesimpulan yang diperoleh yaitu, terdapat pengaruh secara parsial antara persentase penduduk miskin dan TPT terhadap persentase kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur.

Kemudian, untuk $\hat{\beta}_2$ dan $\hat{\beta}_4$ diperoleh keputusan Terima H_0 . Sehingga, kesimpulan yang diperoleh tidak terdapat pengaruh secara parsial antara persentase penduduk yang mempunyai keluhan kesehatan dan pendapatan terhadap persentase kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur.

Maka, diperoleh model estimasi yang sesuai pada regresi data panel dengan FEM sebagai berikut:

$$\hat{Y}_{it} = \alpha_i + 50.06444 - 2.832989X_{1it} + 1.565032X_{3it} \quad (25)$$

Intepretasi Model:

Tabel 7. Intepretasi Model

Intepretasi	
$\hat{\beta}_1$	Jika persentase penduduk miskin meningkat sebesar 1%, maka persentase kepesertaan BPJS Non PBI mengalami kemerosotan sebesar 2.832989% dengan menganggap variabel lain konstan.
$\hat{\beta}_3$	Jika TPT meningkat sebesar 1 satuan, maka persentase kepesertaan BPJS Non PBI mengalami peningkatan sebesar 1.565032% dengan menganggap variabel lain konstan.

Peningkatan persentase penduduk miskin mengindikasikan meningkatnya keadaan penduduk yang tidak mampu dalam memenuhi kebutuhan hidupnya, baik kebutuhan dasar makanan maupun kebutuhan dasar non makanan (Soleh, 2015). Hal ini, akan menyebabkan kemerosotan terhadap persentase kepesertaan BPJS Non PBI. Fenomena

ini terjadi karena, ketidakmampuannya dalam memenuhi kebutuhan sehari-hari membuat penduduk miskin enggan untuk mendaftarkan diri terhadap program BPJS Non PBI. Karena, selayaknya penduduk miskin telah terakomodasi dalam kepesertaan BPJS Kesehatan PBI. Namun, saat ini masih banyak ditemukan peserta BPJS Kesehatan PBI belum tepat sasaran (Nurgahayu & Ulfah, 2020). Hal ini, akan mengakibatkan terjadinya subsidi terbalik (Litawati, 2016). Hasil dari penelitian ini, sesuai dengan pemelitan Riza *et al.* (2020), bahwa ketidakmampuan ekonomi merupakan penyebab ketidakpatuhan sebagai peserta Jaminan Kesehatan Berbasis Masyarakat (JKBM).

Kemudian, meningkatnya TPT dapat menyebabkan peningkatan persentase kepesertaan BPJS Non PBI. Karena, menurut Kementerian Pekerjaan Umum dan Perumahan Rakyat (2014), tidak semua pengangguran tidak mampu memenuhi kebutuhan ekonomi. Hal ini, terjadi karena semakin banyak ditemui pengangguran yang berkecukupan (*discourage worker*). Fenomena ini terjadi karena semakin banyak pengangguran yang melakukan kegiatan investasi yang diindikasikan dengan peningkatan investor. Pernyataan ini sesuai dengan pendapat Astika (2019), bahwa investor dapat menghasilkan uang tanpa bekerja sehingga disebut pengangguran. Pengangguran yang berkecukupan dalam hal ini merupakan pengangguran yang mampu memenuhi kebutuhan perekonomiannya. Maka, tercukupinya kebutuhan ekonomi dapat memberikan pengaruh positif terhadap peningkatan kepesertaan BPJS Non PBI.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan, maka diperoleh kesimpulan sebagai berikut:

1. Persentase pada variabel persentase kepesertaan BPJS Non PBI, persentase penduduk yang mempunyai keluhan kesehatan, dan TPT di atas rerata pada tahun 2017 hingga tahun 2020 mengalami fluktuasi. Sedangkan, untuk persentase variabel persentase penduduk miskin dan pendapatan di atas rerata untuk tahun 2017 hingga tahun 2018 konstan dan mengalami peningkatan pada tahun 2020.

2. Model regresi yang sesuai untuk mengestimasi persentase kepesertaan BPJS Non PBI beserta variabel yang mempengaruhinya yaitu FEM dengan persamaan sebagai berikut:

$$\hat{Y}_{it} = \alpha_i + 50.06444 - 2.832989X_{1it} + 1.565032X_{3it}$$

dengan nilai R^2 sebesar 86.45%. Nilai ini, dapat diartikan dengan variabel persentase

kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur dapat dijelaskan oleh variabel persentase penduduk miskin, persentase penduduk yang mempunyai keluhan kesehatan, TPT, dan pendapatan sebesar 86.45%. Sedangkan, 13.65% mengenai variabel persentase kepesertaan BPJS Non PBI di Kabupaten/Kota Provinsi Jawa Timur dapat dijelaskan melalui variabel lain yang berada di luar model.

3. Berdasarkan hasil FEM, terdapat 2 variabel yang signifikan yaitu persentase penduduk miskin dan TPT. Maka, jika persentase penduduk miskin meningkat sebesar 1%, akan menyebabkan persentase kepesertaan BPJS Non PBI mengalami kemerosotan sebesar 2.832989% dengan menganggap variabel lain konstan. Kemudian, jika TPT meningkat sebesar 1 satuan, akan menyebabkan persentase kepesertaan BPJS Non PBI mengalami peningkatan sebesar 1.565032% dengan menganggap variabel lain konstan.

ACUANREFERENSI

- Abadi, M. Y., Marzuki, D. S., Arifin, M. A., Darmawansah., Rahmadani, S., & Fajrin, M. A. (2019). Analisis Kepesertaan Mandiri BPJS Kesehatan di Sektor Informal (Studi di Kota Makassar), *Jurnal Manajemen Kesehatan Yayasan RS. Dr. Soetomo*, 5(2):114-124.
- Astika, A. (2021). Bikin Heran, Orang Indonesia Ini 'Nganggur' tapi Punya Harta Triliunan dan Terus Bertambah. <https://www.google.com/amp/s/www.sonora.id/amp/422672782/bikin-heran-orang-indonesia-ini-nganggur-tapi-punya-harta-triliunan-dan-terus-bertambah>, diakses pada 22 Desember 2021.
- Badan Pusat Statistik. (2018). *Provinsi Jawa Timur Dalam Angka 2018*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik. (2019). *Provinsi Jawa Timur Dalam Angka 2019*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik. (2020). *Provinsi Jawa Timur Dalam Angka 2020*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik. (2021a). *Provinsi Jawa Timur Dalam Angka 2021*. Jawa Timur: Badan Pusat Statistik.
- Badan Pusat Statistik. (2021b). *Konsep Definisi dan Tata Cara Pengisian Kuesioner*. Jawa Timur: Badan Pusat Statistik.
- Baum, C. F. (2001). Residual Diagnostics for Cross Section Time Series Regression Models. *Stata Corporation*, 1(1):101-104.
- Gubernur Jawa Timur. Keputusan Gubernur Jawa Timur Nomor 188/665/KPTS/013/2018 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2019. Jawa Timur: Tim Penyusun Rancangan Peraturan Gubernur.
- Gubernur Jawa Timur. Keputusan Gubernur Jawa Timur Nomor 188/568/KPTS/013/2019 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2020. Jawa Timur: Tim Penyusun Rancangan Peraturan Gubernur.
- Gubernur Jawa Timur. Peraturan Gubernur Jawa Timur Nomor 121 Tahun 2016 Tentang Upah Minimum Kabupaten/Kota di Jawa Timur Tahun 2017. Jawa Timur: Tim Penyusun Rancangan Peraturan Gubernur.
- Gubernur Jawa Timur. Peraturan Gubernur Jawa Timur Nomor 75 Tahun 2017 Tentang Upah minimum Kabupaten/Kota di Jawa Timur Tahun 2018. Jawa Timur: Tim Penyusun Rancangan Peraturan Gubernur.
- Ghozi, S. & Hermansyah, H. (2018). Analisis Regresi Data Panel Profitabilitas Bank Pembangunan Daerah (PBD) di Indonesia. *Jurnal Matematika*, 8(1):1-12.
- Gujarati, D. N. (2003). *Basic Econometrics Fourth Edition*. McGraw-Hill Higher Education:New York.
- Ham, F. C., Karamoy, H., & Alexander, S. (2018). Analisis Pengakuan Pendapatan dan Beban pada PT. Bank Perkereditan Rakyat Prisma Dana Manado. *Jurnal Riset Akuntansi Going Concern*, 13(2):628-638.
- Hsiao, C. (2003). *Analysis of Panel Data*. Cambridge University Press:New York.
- Indonesia. Undang-Undang Republik Indonesia Nomor 40 Tahun 2004. Sekretariat Negara: Jakarta.
- Kementerian Kesehatan. (2021). *Jaminan Kesehatan Nasional untuk Indonesia Lebih Sehat*. <http://jkn.kemkes.go.id/faq/php>, diakses pada 31 Oktober 2021.
- Kementerian Pekerjaan Umum dan Perumahan Rakyat. (2014). Pengangguran tapi Tidak Miskin, Bekerja tapi Miskin. <http://kotaku.pu.go.id:8081/wartaarsipdetil.asp?mid=6438&catid=1&>, diakses pada 17 Desember 2021.
- Kusumaningrum, A. & Azinar, M. (2018). Kepesertaan Masyarakat Dalam Jaminan Kesehatan Nasional Secara Mandiri. *Higeia Journal of Public Health Research and Development*, 2(1):149-160.

- Khosropour, A. (2017). A Panel Data Analysis of The Relationship Between Corporate Social Responsibility and Earnings Management: Evidence From Iran. *Revista Quid*:2423-2431.
- Litawati, H. (2016). Strategi Perencanaan Jaminan Kesehatan Semesta Kabupaten Banyuwangi. *Jurnal Ilmiah Administrasi Publik*, 2(1):46-56.
- Nelisma., Afni, N., & Rosnawati. (2019). Faktor-Faktor yang Berhubungan dengan Minat Masyarakat Dalam Kepesertaan BPJS Kesehatan di Kelurahan Talise Valangguni, *Jurnal Kolaboratif Sains*, 2(1):471-479.
- Nurgahayu & Ulfa, N. (2020). Kesesuaian Anggota BPJS Penerima Bantuan Iuran (PBI) dengan Indikator Kemiskinan di Kecamatan Malua Kabupaten Enrekang. *Window of Public Health Journal*, 1(3):220-231.
- Oldistra, F. & Machdum, S. V. (2020). Analisis Perencanaan Kebijakan Jaminan Kesehatan Nasional dari Aspek Kepesertaan di Kementerian PPN/BAPPENAS. *Jurnal Ilmu Kesejahteraan Sosial*, 21(1):63-86.
- Pangestika, V. B., Jati, S. P., & Sriatmi, A. (2017). Faktor-Faktor yang Berhubungan dengan Kepesertaan Sektor Informal Dalam BPJS Kesehatan Mandiri di Kelurahan Poncol, Kecamatan Pekalongan Timur, Kota Pekalongan. *Jurnal Kesehatan Masyarakat*, 5(3):39-49.
- Putra, I. K. A. A., & Arka, S. (2018). Analisis Pengaruh Tingkat Pengangguran Terbuka, Kesempatan Kerja, dan Tingkat Pendidikan Terhadap Tingkat Kemiskinan pada Kabupaten/Kota di Provinsi Bali. *Jurnal Ekonomi Pembangunan Universitas Udayana*, 7(3):416-444.
- Putro, G. & Barida, I. (2017). Manajemen Peningkatan Kepesertaan Dalam Jaminan Kesehatan Nasional pada Kelompok Nelayan Non Penerima Bantuan Iuran (Non PBI). *Media Litbangkes*, 27(1):17-24.
- Rahmadeni. & Wulandari, N. (2017). Analisis Faktor-Faktor yang Mempengaruhi Inflasi pada Kota Metropolitan di Indonesia dengan Menggunakan Analisis Regresi Data Panel. *Jurnal Sains Matematika dan Statistika*, 3(2):34-42.
- Riza, Y., Budiarto, W., Anam, K., Qariati, N. I., Hayati, R., Irianty, H., Asrinawati., & Yani, A. (2020). Contribution Deposit Compliance: Income and Knowledge of BPJS Health Mandiri Participants. *Systematic Reviews in Pharmacy*, 11(10):854-858.
- Soleh, A. (2015). Pertumbuhan Ekonomi dan Kemiskinan di Indonesia. *Jurnal Ilmu Ekonomi dan Bisnis*, 2(2):197-209.
- Sriyana, J. (2014). *Metode Regresi Data Panel (Dilengkapi Analisis Kinerja Bank Syariah di Indonesia)*. Ekonisia:Yogyakarta.
- Suprianto, A. & Mutiarin, D. (2017). Evaluasi Pelaksanaan Jaminan Kesehatan Nasional (Studi Tentang Hubungan Stakeholder, Model Pembiayaan, dan Outcome JKN di Kabupaten Bantul Provinsi Daerah Istimewa Yogyakarta. *Journal of Governance and Public Policy*. 4(1):71-107.
- Yulianto, D. A., Sugiman., & Agoestanto, A. (2018). Estimasi Regresi Robust Model Seemingly Unrelated Regression (SUR) dengan Metode Generalized Least Square (GLS), *Unnes Journal of Mathematics*, 7(2):216-227.

Pemodelan Kejadian Balita Stunting di Kabupaten Bojonegoro dengan Metode *Geographically Weighted Regression* dan *Multivariate Adaptive Regression Splines*

Alif Yuanita Kartini⁽¹⁾, Laelatul Nur Ummah⁽²⁾

Universitas Nahdlatul Ulama Sunan Giri

Jl. A. Yani No. 10 Bojonegoro, 62115, Telp. (0353) 887341

e-mail: yuanitaalif5@gmail.com dan laelatul.nurummah9@gmail.com

ABSTRAK

Balita stunting adalah masalah kronik yang berkaitan dengan gizi yang disebabkan salah satunya karena asupan kandungan gizi yang kurang pada balita. Kabupaten Bojonegoro adalah satu dari beberapa kabupaten yang ada di provinsi Jawa Timur dengan kejadian balita stunting yang masih banyak dijumpai. Pada penelitian ini akan melakukan perbandingan pemodelan kejadian balita stunting di kabupaten Bojonegoro menggunakan metode GWR dan MARS. GWR mampu memodelkan kasus setiap wilayah secara spasial, sedangkan MARS mampu memodelkan kasus tanpa mempertimbangkan pola hubungan antara variabel prediktor dengan variabel respon. Model GWR terbaik dipilih dari nilai CV dan MSE yang paling kecil serta nilai *R-Square* yang paling besar. Diperoleh model terbaik dari GWR yaitu dengan pembobot fungsi kernel *Adaptive Bi-Square* dengan *bandwidth* 28, nilai CV=2,4635, MSE=0,8620, dan R-Square=0,8734. Model MARS terbaik dipilih berdasarkan nilai GCV terkecil dan *R-Square* terbesar. yaitu pada BF=24, MI=1 dan MO=1 dengan nilai GCV=1,29144 dan *R-Square*=0,841. Perbandingan kedua model didapatkan model MARS lebih baik karena nilai *R-Square* model GWR lebih besar dari MARS meskipun nilai angkanya tidak jauh berbeda, sedangkan nilai MSE model MARS lebih kecil dari nilai MSE model GWR dengan selisih jauh beda.

Kata kunci : *Balita Stunting, Bojonegoro, GWR, MARS*

ABSTRACT

Toddler stunting is a chronic problem related to nutrition, one of which is due to inadequate intake of nutritional content in toddlers. Bojonegoro Regency is one of several regencies in East Java province where the incidence of stunting under five is still common. In this study, we will compare the incidence of stunting under five in Bojonegoro district using the GWR and MARS methods. GWR is able to model cases for each region spatially, while MARS is able to model cases without considering the pattern of relationships between predictor variables and response variables. The best GWR model is selected from the smallest CV and MSE values and the largest R-Square value. The best model obtained from GWR is by weighting the Adaptive Bi-Square kernel function with a bandwidth of 28, CV = 2.4635, MSE = 0.8620, and R-Square = 0.8734. The best MARS model was chosen based on the smallest GCV value and the largest R-Square. namely at BF=24, MI=1 and MO=1 with GCV=1,29144 and R-Square=0,841 values. Comparing the two models, the MARS model is better because the R-Square value of the GWR model is greater than the MARS although the numerical value is not much different, while the MSE value of the MARS model is smaller than the MSE value of the GWR model with a much different difference.

Keywords : *Stunting Toddler, Bojonegoro, GWR, MARS*

1. PENDAHULUAN

Balita atau biasa disebut dengan bawah lima tahun merupakan seorang anak yang mempunyai usia di bawah lima tahun. Balita ini adalah sebutan untuk anak yang mempunyai usia antara 1 sampai 3 tahun atau dengan istilah batita dan anak usia prasekolah yaitu antara 3 sampai 5 tahun (Ningsih, 2017). Berdasarkan standard *World Health*

Organization (WHO), balita dikatakan mengalami stunting apabila tinggi badannya kurang dari minus dua standard deviasi median dari pertumbuhan anak. Balita stunting adalah masalah kronik yang menyangkut gizi yang disebabkan oleh beberapa faktor diantaranya status gizi ibu ketika hamil, kondisi sosial dan ekonomi dari keluarga, kondisi kesehatan pada bayi serta minimnya asupan gizi

pada bayi. Dalam kondisi seperti itu, balita stunting akan merasa sulit untuk mencapai perkembangan fisik dan kognitif yang maksimal (Kemenkes RI, 2018).

Bojonegoro adalah satu diantara beberapa kabupaten di provinsi Jawa Timur dengan kejadian balita stunting yang masih ada. Di kabupaten Bojonegoro terdapat 15 desa yang tersebar di sejumlah kecamatan yang menjadi lokasi kasus stunting tertinggi selama tiga tahun terakhir yaitu kecamatan Balen, Kapas, Kalitidu, Ngasem, Temayang, Gayam dan Bojonegoro (Kesehatan, 2020). Hal ini mengindikasikan bahwa kejadian stunting ini memiliki faktor penyebab yang berbeda di masing-masing kecamatan yang dipengaruhi oleh faktor geografis. Akan tetapi pola hubungan antara banyaknya kejadian stunting dengan faktor-faktor yang diduga mempengaruhinya juga tidak jelas. Untuk itu perlu dilakukan penelitian untuk mengetahui faktor penyebab kejadian balita stunting di kabupaten Bojonegoro.

Banyak penelitian yang sudah dilakukan tentang kejadian balita stunting diantaranya menggunakan metode *Geographically Weighted Regression* (GWR) yang dilakukan oleh (Al Azies et al., 2019), (Pramoedyo et al., 2020), (Fadliana & Darajat, 2021), dan (Ardianti et al., 2021). GWR disini digunakan untuk memodelkan kejadian stunting dengan memperhatikan efek spasial atau wilayah. Sementara itu pada penelitian yang lain untuk memodelkan kejadian balita stunting menggunakan metode *Multivariate Adaptive Regression Splines* (MARS) diantaranya yang dilakukan oleh (Azizah & Permatasari, 2020), (Kasse, 2017), dan (Ningrum, 2021). MARS digunakan karena tidak adanya pola yang jelas antara variabel respon dengan variabel prediktor yang digunakan dalam penelitian-penelitian tersebut. Olehkarena itu pada penelitian ini akan melakukan perbandingan pada metode *Geographically Weighted Regression* dan *Multivariate Adaptive Regression Splines* untuk memodelkan kejadian balita stunting di kabupaten Bojonegoro pada data Profil Kesehatan Kabupaten Bojonegoro Tahun 2019 (Dinas Kesehatan Kabupaten Bojonegoro, 2019) yang didapatkan dari website resmi Dinas Kesehatan kabupaten Bojonegoro. Dari beberapa penelitian terdahulu tersebut variabel yang digunakan sedikit oleh karena itu penelitian ini dibuat dengan variabel yang lebih banyak dan beda dari penelitian terdahulu tersebut.

Alasan penelitian ini menggunakan metode *Geographically Weighted Regression* (GWR) dan *Multivariate Adaptive Regression Splines* (MARS)

adalah ingin mengetahui model terbaik antara metode GWR yang mampu memodelkan kasus setiap wilayah secara spasial dengan metode MARS yang memodelkan kasus tanpa mempertimbangkan pola hubungan antara variabel prediktor dengan variabel respon untuk menanggulangi permasalahan pada data yang mempunyai dimensi tinggi dan menggunakan variabel prediktor lebih dari 3. Model terbaik yang didapatkan digunakan sebagai acuan dalam pemodelan kejadian balita stunting di kabupaten Bojonegoro. variabel prediktor yang mempengaruhi kejadian balita stunting di kabupaten Bojonegoro perlu diidentifikasi dan diselidiki pengaruhnya secara signifikan. Hasil penelitian ini akan memberikan masukan *knowledge* berupa model regresi untuk dinas kesehatan kabupaten Bojonegoro dalam memprediksi prediktor yang berpengaruh terhadap kejadian balita stunting di kabupaten Bojonegoro.

Dalam penelitian ini masalah yang muncul yaitu bagaimana karakteristik kejadian balita stunting di kabupaten Bojonegoro beserta faktor-faktor yang diduga mempengaruhinya, bagaimana memodelkan kejadian balita stunting di kabupaten Bojonegoro dengan metode *Geographically Weighted Regression* dan *Multivariate Adaptive Regression Splines*, serta bagaimana perbandingan model kejadian balita stunting dengan menggunakan *Geographically Weighted Regression* dan *Multivariate Adaptive Regression Splines*.

2. METODE PENELITIAN

Sumber data yang digunakan dalam penelitian ini merupakan data sekunder yang didapatkan dari data Pemantauan Status Gizi (PSG) tahun 2019 yang dilakukan oleh dinas kesehatan kabupaten Bojonegoro dan data hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2019 yang dilakukan oleh Badan Pusat Statistik (BPS). Unit pengamatan dalam penelitian adalah 28 kecamatan di kabupaten Bojonegoro yaitu kecamatan Margomulyo, Ngraho, Tambakrejo, Ngambon, Sekar, Bubulan, Gondang, Temayang, Sugihwaras, Kedungadem, Kepohbaru, Baureno, Kanor, Sumberejo, Balen, Sukosewu, Kapas, Bojonegoro, Trucuk, Dander, Ngasem, Gayam, Kalitidu, Malo, Purwosari, Padangan, Kasiman dan Kedewan (BPS, 2020).

Variabel yang digunakan dalam penelitian ini terdiri dari satu variabel respon dan delapan variabel prediktor dengan skala pengukuran rasio sebagaimana ditunjukkan pada Tabel 1 berikut.

Tabel 1. Variabel dan Indikator Penelitian

Atribut	Variabel Indikator	Skala Pengukuran
---------	--------------------	------------------

Y	Persentase Kejadian Balita Stunting Per Kecamatan	Rasio
X1	Persentase Bayi Baru Lahir Memperoleh Inisiasi Menyusui Dini (IMD)	Rasio
X2	Persentase Bayi Memperoleh Asi Eksklusif	Rasio
X3	Persentase Balita Memperoleh Vitamin A	Rasio
X4	Persentase Balita Memiliki KMS atau Buku KIA	Rasio
X5	Persentase Balita Ditimbang Empat Kali atau Lebih dalam Enam Bulan Terakhir	Rasio
X6	Persentase Balita Kurus Mendapatkan Makanan Tambahan	Rasio
X7	Persentase Rumah Tangga yang Menggunakan Air Bersih	Rasio
X8	Persentase Tenaga Kesehatan	Rasio

Adapun langkah-langkah analisis dalam penelitian ini yaitu sebagai berikut:

- a. Mendeskripsikan kejadian balita stunting di kabupaten Bojonegoro beserta faktor-faktor yang diduga mempengaruhinya
- b. Memodelkan kejadian balita stunting di kabupaten Bojonegoro menggunakan metode *Geographically Weighted Regression* (GWR) dengan langkah-langkah sebagai berikut (Lu et al., 2017)
 - 1) Melakukan perhitungan jarak *euclidean* antar wilayah pengamatan
 - 2) Melakukan penentuan bandwidth yang optimum dengan memperhatikan angka minimum pada *Cross Validation* (CV)
 - 3) Melakukan pemilihan fungsi pembobot kernel terbaik dengan memperhatikan angka *Akaike Information Criterion* (AIC)
 - 4) Melakukan perhitungan nilai estimasi parameter model GWR dengan fungsi kernel terbaik
 - 5) Melakukan uji kesesuaian model untuk mengetahui perbandingan antara model OLS dan metode GWR
 - 6) Melakukan pengujian signifikansi secara parsial terhadap parameter model GWR dengan fungsi pembobot terbaik

- 7) Memberikan interpretasi dari hasil model GWR
- c. Memodelkan kejadian balita stunting di kabupaten Bojonegoro menggunakan metode *Multivariate Adaptive Regression Splines* (MARS) dengan langkah-langkah sebagai berikut (Otok, 2014).
 - 1) Melakukan penentuan maksimum jumlah fungsi basis (BF), dimana basis fungsi yang disarankan untuk digunakan yaitu antara dua sampai empat kali jumlah variabel prediktor (Friedman, 1991), sehingga maksimum jumlah BF yang digunakan adalah 16, 24 dan 32.
 - 2) Melakukan penentuan jumlah maksimum (MI), dimana jumlah MI yang disarankan adalah 1, 2 dan 3. Apabila jumlah MI lebih dari 3, maka model yang dihasilkan akan sangat kompleks (Friedman & Roosen, 1995)
 - 3) Melakukan *trial and error* untuk penentuan minimum observasi diantara knot.
 - 4) Memilih model MARS terbaik dengan nilai GCV minimum
 - 5) Melakukan penaksiran parameter model MARS.
 - 6) Melakukan pengujian signifikansi parameter model MARS secara simultan (Uji F)
 - 7) Melakukan pengujian signifikansi parameter secara parsial (Uji t).
 - 8) Memberikan interpretasi dari model MARS yang didapatkan
- d. Membandingkan model MARS dan GWR dengan melihat nilai R-Square (R^2) dan nilai MSE pada masing-masing model.

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Kejadian Balita Stunting di kabupaten Bojonegoro

Karakteristik kejadian balita stunting di kabupaten Bojonegoro beserta faktor yang diduga mempengaruhinya dapat disajikan dalam bentuk statistik deskriptif sebagaimana ditunjukkan pada Tabel 2 berikut.

Tabel 2. Karakteristik Kejadian Balita Stunting di Kabupaten Bojonegoro beserta Faktor yang Diduga Mempengaruhinya

Variabel	N	Min	Max	Mean
Y	28	1.03	7.22	3.5721
X1	28	1.00	7.02	3.5714
X2	28	1.32	6.18	3.5704

X3	28	3.13	4.08	3.5711
X4	28	2.18	6.50	3.5707
X5	28	2.82	4.09	3.5711
X6	28	0.80	7.86	3.5718
X7	28	2.39	5.75	3.5718
X8	28	1.08	8.60	3.5729

Berdasarkan Tabel 2 jumlah data pada setiap variabel sebesar 28, berasal dari data Pemantauan Status Gizi (PSG) tahun 2019 yang dilakukan oleh Dinas Kesehatan Kabupaten Bojonegoro dan data hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2019 yang dilakukan oleh BPS.

Persentase kejadian balita stunting per kecamatan terendah sebesar 1.03% yaitu di kecamatan Padangan dan tertinggi sebesar 7.22% yaitu di kecamatan Kedungadem. Rata-rata persentase kejadian balita stunting per kecamatan sebesar 3.5721% maka dapat disimpulkan bahwa secara statistik persentase kejadian balita stunting per kecamatan berada di bawah standar *World Health Organization* (WHO) yaitu 20% pada tahun 2018 (Muktamar Umakaapa, 2020), itu artinya bahwa persentase kejadian balita stunting per kecamatan dikatakan baik.

Persentase bayi baru lahir yang memperoleh Inisiasi Menyusui Dini (IMD) terendah sebesar 1.00% yaitu di kecamatan Padangan dan tertinggi sebesar 7.02% yaitu di kecamatan Kapas. Rata-rata persentase bayi yang baru lahir memperoleh Inisiasi Menyusui Dini (IMD) sebesar 3.5714%. Maka dapat disimpulkan bahwa secara statistik persentase bayi yang baru lahir memperoleh Inisiasi Menyusui Dini (IMD) di bawah target renstra tahun 2019 yaitu sebesar 50.0% (K. Kesehatan & Indonesia, 2019), itu artinya persentase bayi baru lahir mendapat Inisiasi Menyusui Dini (IMD) dikatakan masih buruk, maka perlu dilakukan penyuluhan oleh bidan dan tenaga kesehatan lainnya.

Persentase bayi mendapat ASI eksklusif terendah sebesar 1.32% yaitu di kecamatan Sekar dan tertinggi sebesar 6.18% yaitu di kecamatan Kedungadem. Rata-rata persentase bayi mendapat ASI eksklusif sebesar 3.5704% maka dapat disimpulkan bahwa secara statistik persentase bayi mendapat ASI eksklusif berada di bawah target renstra tahun 2019 yaitu 50% (K. Kesehatan & Indonesia, 2019), itu artinya persentase bayi mendapat ASI eksklusif dikatakan masih buruk, maka perlu dilakukan penyuluhan tentang pemberian ASI eksklusif.

Persentase balita memperoleh vitamin A terendah sebesar 3.13% yaitu di kecamatan Kedungadem dan yang paling tinggi sebesar 4.08%

yaitu di kecamatan Baureno. Rata-rata persentase balita memperoleh vitamin A sebesar 3.5711% maka dapat disimpulkan bahwa secara statistik persentase balita memperoleh vitamin A berada di bawah standar nasional tahun 2019 yaitu 78.68% (K. Kesehatan & Indonesia, 2019), itu artinya bahwa persentase balita mendapat kapsul vitamin A dikatakan masih buruk dan perlu ditingkatkan lagi untuk memberikan vitamin A pada balita.

Persentase balita mempunyai KMS atau buku KIA terendah sebesar 2.18% yaitu di kecamatan Sugihwaras dan tertinggi sebesar 6.50% yaitu di kecamatan Baureno. Rata-rata persentase balita yang mempunyai KMS atau buku KIA sebesar 3.5707% artinya balita yang mempunyai KMS atau buku KIA di kabupaten Bojonegoro sebesar setengahnya dari angka maksimum.

Persentase balita yang ditimbang empat kali atau lebih dalam enam bulan terakhir terendah sebesar 2.82% yaitu di kecamatan Kedungadem dan tertinggi sebesar 4.09% yaitu di kecamatan Kasiman. Rata-rata persentase balita yang ditimbang empat kali atau lebih dalam enam bulan terakhir sebesar 3.5711% maka dapat disimpulkan bahwa secara statistik persentase balita yang ditimbang empat kali atau lebih dalam enam bulan terakhir berada di bawah standar nasional tahun 2019 yaitu 73.86% (K. Kesehatan & Indonesia, 2019), artinya balita yang ditimbang empat kali atau lebih dalam enam bulan terakhir dikatakan masih buruk, maka harus ditingkatkan lagi penimbangan bayi sebanyak empat kali atau lebih dalam enam bulan terakhir.

Persentase balita kurus mendapatkan makanan tambahan terendah sebesar 0.80% yaitu di kecamatan Ngambon dan tertinggi sebesar 7.86% yaitu di kecamatan Baureno. Rata-rata persentase balita kurus mendapatkan makanan tambahan sebesar 3.5718% maka dapat disimpulkan bahwa secara statistik persentase balita kurus mendapatkan makanan tambahan berada di bawah target renstra tahun 2019 yaitu 90% (K. Kesehatan & Indonesia, 2019), itu artinya bahwa persentase balita kurus mendapatkan makanan tambahan dikatakan masih buruk.

Persentase rumah tangga yang menggunakan air bersih terendah sebesar 2.39% yaitu di kecamatan Temayang dan tertinggi sebesar 5.75% yaitu di kecamatan Sumberejo. Rata-rata persentase rumah tangga yang menggunakan air bersih sebesar 3.5718% maka dapat disimpulkan bahwa secara statistik persentase rumah tangga yang menggunakan air bersih berada di bawah standar nasional tahun 2019 yaitu 87.81% (K. Kesehatan &

Indonesia, 2019), itu artinya bahwa persentase rumah tangga yang menggunakan air bersih dikatakan buruk, maka perlu ditingkatkan lagi rumah tangga yang menggunakan air bersih.

Persentase jumlah tenaga kesehatan paling rendah sebesar 1.08% yaitu di kecamatan Padangan dan paling tinggi sebesar 8.60% yaitu di kecamatan Kedungadem. Rata-rata persentase jumlah tenaga kesehatan sebesar 3.5729%. Kecamatan yang memiliki jumlah tenaga kesehatan di bawah rata-rata yaitu kecamatan Margomulyo, kecamatan Ngraho, kecamatan Ngambon, kecamatan Bubulan, kecamatan Gondang, kecamatan Sugihwaras, kecamatan Kanor, kecamatan Sumberejo, kecamatan Balen, kecamatan Sukosewu, kecamatan Trucuk, kecamatan Ngasem, kecamatan Gayam, kecamatan Malo, kecamatan Purwosari, kecamatan Padangan, kecamatan Kasiman dan kecamatan Kedewan. Sedangkan kecamatan yang mempunyai jumlah tenaga kesehatan di atas rata-rata yaitu di kecamatan Tambakrejo, Sekar, Temayang, Kedungadem, Kepohbaru, Baureno, Kapas, Bojonegoro, Dander dan Kalitidu.

3.2 Analisis Regresi Linier

Analisis menggunakan metode regresi linier atau regresi global digunakan untuk mengetahui variabel-variabel mana yang signifikan mempengaruhi persentase kejadian balita stunting per kecamatan di kabupaten Bojonegoro. Analisis menggunakan metode regresi linear pada penelitian ini tidak memperhitungkan efek spasial dan didapatkan hasil sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Hasil Estimasi Parameter dan Signifikansi Parameter

Variabel	Parameter	SE	Statistik t
$\hat{\beta}_0$	4.8836	3.3477	1.4588
$\hat{\beta}_1$	0.2747	0.2468	1.1132
$\hat{\beta}_2$	0.1698	0.4064	0.4180
$\hat{\beta}_3$	0.6413	1.2459	0.5147
$\hat{\beta}_4$	0.4596	0.9636	0.4770
$\hat{\beta}_5$	-1.8894	1.0556	-1.7898
$\hat{\beta}_6$	0.4650	0.1349	3.4464*
$\hat{\beta}_7$	-0.5860	1.0927	-0.5363

* Signifikan untuk α sebesar 5%. ($|t| \geq t\text{-tabel} = 2,093$)

Berdasarkan tabel 3, variabel prediktor yang signifikan mempengaruhi persentase kejadian balita stunting per kecamatan adalah persentase balita kurus mendapatkan makanan tambahan karena nilai

$|t| \geq t\text{-tabel}$ dimana $t\text{-tabel} = t_{(\alpha/2, df)} = t_{(0,025, 19)} = 2,093$ (lihat tabel t). Hasil pengujian ini disimpulkan bahwa ada pengaruh parsial/individu persentase balita kurus mendapatkan makanan tambahan terhadap persentase kejadian balita stunting per kecamatan. Untuk mengukur kebaikan model, diperoleh dari hasil output yaitu CV sebesar 1,836276, R-Square sebesar 0,757402, dan MSE = $SSE/(n-p) = 18,370346/(28-9) = 0,9667$.

3.3 Model Geographically Weighted Regression (GWR)

Dalam metode GWR, untuk menentukan fungsi kernel terbaik yang akan digunakan yaitu dengan melihat nilai CV, MSE dan nilai R-Square. Fungsi kernel yang dipilih yaitu yang mempunyai nilai CV dan MSE yang paling kecil serta nilai R-Square yang paling besar. Hasil dari pemilihan fungsi kernel yang akan digunakan sebagaimana ditunjukkan pada Tabel 4.

Tabel 4. Selection Criteria

Model	Fungsi Pembobot Kernel	Nilai CV	R-Square	MSE
OLS	-	1.8363	0.7574	0.9669
GWR	Fixed Gaussian	1.9766	0.8298	0.8920
GWR	Fixed Bi-Square	-	-	-
GWR	Adaptive Bi-Square	2.4635	0.8734	0.8620
GWR	Adaptive Gaussian	1.8504	0.7905	0.9320

Berdasarkan Tabel 4 didapatkan model GWR yang paling baik yaitu model GWR yang menggunakan fungsi pembobot kernel *Adaptive Bi-Square*. Selanjutnya untuk kesesuaian model dilakukan pengujian variabilitas spasial. Pada pengujian ini diwakili oleh *Geographical variability tests of local coefficients* yang dilakukan pada variabel independen untuk mengetahui apakah termasuk variabel global atau variabel lokal. Dikatakan sebagai variabel global atau dengan kata lain tidak ada variabilitas spasial apabila mempunyai nilai *DIFF of Criterion* yang positif, sedangkan dikatakan sebagai variabel lokal atau ada variabilitas spasial apabila nilai *DIFF of Criterion*nya adalah negatif. Hasil untuk pengujian variabilitas spasial ditunjukkan pada Tabel 5 berikut.

Tabel 5. Hasil Pengujian Variabilitas Spasial untuk Model GWR dengan kernel *Adaptive Bi-Square*

Variabel	<i>DIFF of Criterion</i>	Bernilai	Kesimpulan
Intercept	-7.074987	Negatif	Variabel Lokal
X1	0.069134	Positif	Variabel Global
X2	0.264507	Positif	Variabel Global
X3	0.158468	Positif	Variabel Global
X4	-0.475254	Negatif	Variabel Lokal
X5	-8586.732	Negatif	Variabel Lokal
X6	-0.688070	Negatif	Variabel Lokal
X7	0.221114	Positif	Variabel Global
X8	-0.240847	Negatif	Variabel Lokal

Berdasarkan Tabel 5, variabel penelitian dinilai sebagai variabel lokal dalam model regresi ini, karena sebagian besar ukuran *DIFF of Criterion* bernilai negatif. Hal ini menunjukkan bahwa model GWR sudah sesuai penerapannya. Berdasarkan model GWR dengan Kernel *Adaptive Bi-Square* dilakukan pengujian signifikansi parameter model secara serentak untuk menguji kesesuaian (*goodness of fit*) dari model GWR dengan tabel ANOVA (Uji-F). Pengujian ini digunakan untuk menguji hipotesis nol (H_0) bahwa model serupa dengan model Global (OLS) dengan hasil ditunjukkan pada Tabel 6.

Tabel 6. Hasil Output Tabel ANOVA untuk Model GWR

Sumber	Jumlah Kuadrat	Derajat Bebas	Mean Kuadrat	Statistik F
Residual Global	18.370	19.000		
Dugaan GWR	2.503	1.979	1.265	
Residual GWR	15.867	17.021	0.932	1.35654

Kesesuaian model GWR ditunjukkan ketika H_0 ditolak (Jika $F_{hitung} > F_{tabel}$) dan disimpulkan bahwa GWR tidak serupa dengan model Global (OLS). Dari hasil output tabel ANOVA, diperoleh statistik uji F (F_{hitung}) sebesar 1,356542. Disamping itu, dihasilkan juga derajat bebas df_1 dan df_2 masing-masing sebesar 19,000 (≈ 19) dan 1,979 (≈ 2), sehingga F-tabel untuk $\alpha = 5\%$ diperoleh $F_{tabel} = F_{\alpha(df_1, df_2)} = F_{0,05(19, 2)} = 19,4431$ (lihat tabel F). Dengan demikian diperoleh hasil pengujian $F_{hitung} < F_{tabel}$ sehingga didapatkan keputusan H_0 diterima dan disimpulkan bahwa model GWR serupa dengan model Global.

Selanjutnya untuk estimasi parameter model digunakan matriks pembobot spasial untuk melakukan pendugaan nilai parameter pada model GWR. Nilai parameter untuk model GWR tidak sama untuk setiap wilayah. Dengan demikian, model yang terbentuk dalam penelitian ini ada 28

persamaan yang berbeda untuk setiap kecamatan. Model persamaannya ditunjukkan pada Tabel 7.

Tabel 7. Model GWR di Setiap Kecamatan

Kecamatan	Model
Margomulyo	$\hat{y}_1 = 5,1863 + 0,3544X_1 + 0,2944X_2 + 1,2654X_3 + 0,6664X_4 - 2,6021X_5 + 0,3781X_6^* - 0,8712X_7 + 0,0529X_8 + e_1$
Ngraho	$\hat{y}_1 = 2,8126 - 0,0451X_1 + 0,0931X_2 - 0,1833X_3 + 0,4377X_4 - 0,7132X_5 + 0,8179X_6^* - 0,5464X_7 + 0,4002X_8 + e_2$
Tambakrejo	$\hat{y}_1 = 2,7562 - 0,0273X_1 + 0,1105X_2 - 0,1060X_3 + 0,432X_4 - 0,7689X_5 + 0,7684X_6^* - 0,5510X_7 + 0,4079X_8 + e_3$
Ngambon	$\hat{y}_1 = 2,9051 - 0,0145X_1 + 0,1195X_2 - 0,1002X_3 + 0,3929X_4 - 0,82X_5 + 0,7050X_6^* - 0,5126X_7 + 0,4389X_8 + e_4$
Sekar	$\hat{y}_1 = 3,0975 - 0,0095X_1 + 0,1818X_2 + 0,0104X_3 + 0,4043X_4 - 0,9905X_5 + 0,6594X_6^* - 0,5588X_7 + 0,4507X_8 + e_5$
Bubulan	$\hat{y}_1 = 3,1666 + 0,0989X_1 + 0,1124X_2 + 0,2846X_3 + 0,4405X_4 - 1,2266X_5 + 0,5914X_6^* - 0,5626X_7 + 0,3785X_8 + e_6$
Gondang	$\hat{y}_1 = 3,6682 + 0,1586X_1 + 0,2015X_2 + 0,7008X_3 + 0,5636X_4 - 1,7497X_5 + 0,5293X_6^* - 0,7364X_7 + 0,3026X_8 + e_7$
Temayang	$\hat{y}_1 = 4,7825 + 0,2918X_1 + 0,2589X_2 + 1,1077X_3 + 0,6378X_4 - 2,3717X_5 + 0,4224X_6^* - 0,8240X_7 + 0,1285X_8 + e_8$
Sugihwaras	$\hat{y}_1 = 5,1916 + 0,3557X_1 + 0,2960X_2 + 1,2759X_3 + 0,6704X_4 - 2,614X_5 + 0,3777X_6^* - 0,8771X_7 + 0,0520X_8 + e_9$
Kedungadem	$\hat{y}_1 = 5,4733 + 0,4244X_1 + 0,3491X_2 + 1,5100X_3 + 0,7486X_4 - 2,8987X_5 + 0,3441X_6^* - 1,0061X_7 - 0,0136X_8 + e_{10}$
Kepohbaru	$\hat{y}_1 = 5,5559 + 0,4403X_1 + 0,3348X_2 + 1,2607X_3 + 0,6709X_4 - 2,6489X_5 + 0,3214X_6^* - 0,8965X_7 - 0,0435X_8 + e_{11}$
Baureno	$\hat{y}_1 = 5,5428 + 0,4216X_1 + 0,2797X_2 + 1,0796X_3 + 0,5793X_4 - 2,4546X_5 + 0,3196X_6 - 0,7417X_7 - 0,0381X_8 + e_{12}$
Kanor	$\hat{y}_1 = 5,4742 + 0,384X_1 + 0,2062X_2 + 0,8946X_3 + 0,4589X_4 - 2,2459X_5 + 0,3291X_6^* - 0,5435X_7 - 0,0176X_8 + e_{13}$

Sumberejo	$\hat{y}_1 = 5,4507 + 0,4064X_1 + 0,3030X_2 + 1,1501X_3 + 0,6085X_4 - 2,5160X_5 + 0,3347X_6^* - 0,8002X_7 - 0,0175X_8 + e_{14}$
Balen	$\hat{y}_1 = 5,2554 + 0,3450X_1 + 0,2271X_2 + 0,8899X_3 + 0,4735X_4 - 2,2104X_5 + 0,3590X_6^* - 0,5848X_7 + 0,0253X_8 + e_{15}$
Sukosewu	$\hat{y}_1 = 5,0856 + 0,3269X_1 + 0,2498X_2 + 1,0049X_3 + 0,5517X_4 - 2,3087X_5 + 0,3844X_6^* - 0,7038X_7 + 0,0641X_8 + e_{16}$
Kapas	$\hat{y}_1 = 4,5921 + 0,2198X_1 + 0,0648X_2 + 0,4348X_3 + 0,2641X_4 - 1,6073X_5 + 0,4422X_6^* - 0,2355X_7 + 0,1279X_8 + e_{17}$
Bojonegoro	$\hat{y}_1 = 4,5363 + 0,1825X_1 - 0,0830X_2 + 0,2566X_3 + 0,0636X_4 - 1,3774X_5 + 0,4492X_6^* - 0,1317X_7 + 0,1028X_8 + e_{18}$
Trucuk	$\hat{y}_1 = 2,8108 + 0,0449X_1 - 0,3849X_2 - 0,4091X_3 + 0,4096X_4 - 0,3965X_5 + 0,7374X_6^* - 0,0222X_7 + 0,2307X_8 + e_{19}$
Dander	$\hat{y}_1 = 2,9635 + 0,0960X_1 - 0,0161X_2 + 0,0679X_3 + 0,3420X_4 - 0,9329X_5 + 0,6225X_6^* - 0,3788X_7 + 0,3760X_8 + e_{20}$
Ngasem	$\hat{y}_1 = 2,6870 - 0,0149X_1 + 0,0048X_2 - 0,2038X_3 + 0,3815X_4 - 0,6407X_5 + 0,7491X_6^* - 0,4245X_7 + 0,4216X_8 + e_{21}$
Gayam	$\hat{y}_1 = 2,8196 - 0,0262X_1 - 0,0463X_2 - 0,2966X_3 + 0,4304X_4 - 0,5854X_5 + 0,8052X_6^* - 0,4352X_7 + 0,4011X_8 + e_{22}$
Kalitidu	$\hat{y}_1 = 3,5269 + 0,0091X_1 - 0,3511X_2 - 0,6129X_3 + 0,4967X_4 - 0,4396X_5 + 0,8262X_6^* - 0,3037X_7 + 0,4238X_8 + e_{23}$
Malo	$\hat{y}_1 = 3,7155 - 0,0022X_1 - 0,1725X_2 - 0,5577X_3 + 0,5302X_4 - 0,5821X_5 + 0,8413X_6^* - 0,4667X_7 + 0,4188X_8 + e_{24}$
Purwosari	$\hat{y}_1 = 2,9309 - 0,0376X_1 + 0,0043X_2 - 0,2933X_3 + 0,4428X_4 - 0,6307X_5 + 0,8237X_6^* - 0,4868X_7 + 0,4021X_8 + e_{25}$
Padangan	$\hat{y}_1 = 2,9034 - 0,0446X_1 + 0,0504X_2 - 0,2522X_3 + 0,4404X_4 - 0,6692X_5 + 0,8275X_6^* - 0,5184X_7 + 0,4020X_8 + e_{26}$
Kasiman	$\hat{y}_1 = 3,2475 - 0,0365X_1 - 0,0247X_2 - 0,3904X_3 + 0,4705X_4 - 0,6304X_5 + 0,8457X_6^* - 0,5014X_7 + 0,4110X_8 + e_{27}$
Kedewan	$\hat{y}_1 = 3,9184 + 0,0216X_1 - 0,1350X_2 - 0,5244X_3 + 0,5092X_4 - 0,6636X_5 + 0,8222X_6^* - 0,4556X_7 + 0,3773X_8 + e_{28}$

Margomulyo	$\hat{y}_1 = 5,1863 + 0,3544X_1 + 0,2944X_2 + 1,2654X_3 + 0,6664X_4 - 2,6021X_5 + 0,3781X_6^* - 0,8712X_7 + 0,0529X_8 + e_1$
------------	--

* Signifikan untuk α sebesar 5%. ($|t| \geq t\text{-tabel} = 2,093$)

Dari Tabel 7 model GWR di atas sebagai contoh pada kecamatan Margomulyo dapat disimpulkan bahwa pada kecamatan Margomulyo terdapat pengaruh parsial/individu X_1 terhadap Y sebesar 0,3544 (pengaruh positif), artinya jika terjadi peningkatan persentase bayi yang baru lahir memperoleh Inisiasi Menyusui Dini (IMD) (X_1) sebesar 1 persen, maka persentase kejadian balita stunting per kecamatan (Y) akan mengalami peningkatan sebesar 0,3544 persen. Sebaliknya jika terjadi penurunan persentase bayi yang baru lahir memperoleh Inisiasi Menyusui Dini (IMD) (X_1) sebesar 1 persen, maka persentase kejadian balita stunting per kecamatan (Y) akan mengalami penurunan sebesar 0,3544 persen begitupun pada variabel yang lainnya.

Jika koefisien GWR dipaksakan untuk dilakukan pengujian t , maka digunakan nilai $t\text{-tabel} = t_{(\alpha/2, df)} = t_{(0,025, 16)} = 2,120$ (lihat tabel T) dengan *degree of freedom* (df) sebesar $16,231120 \approx 16$ (dibulatkan ke atas).

Selanjutnya yaitu melakukan pengujian parameter model GWR secara parsial. Pengujian ini dilakukan untuk mendapatkan variabel prediktor yang secara signifikan mempengaruhi persentase kejadian balita stunting per kecamatan. Model GWR pada setiap wilayah tidak sama didasarkan pada nilai koefisien GWR dan juga variabel prediktor yang signifikan berpengaruh terhadap variabel respon. Variabel prediktor dikatakan signifikan apabila nilai $t\text{-hitung} < t\text{-tabel}$. Variabel prediktor yang signifikan di masing-masing kecamatan hampir sama yaitu variabel X_6 (persentase balita kurus mendapatkan makanan tambahan) kecuali untuk kecamatan Baureno semua variabel prediktor tidak signifikan mempengaruhi persentase kejadian balita stunting.

3.4 Model *Multivariate Adaptive Regression Splines* (MARS)

Untuk mendapatkan model MARS terlebih dahulu harus menentukan jumlah maksimum Basis Fungsi (BF), Maksimum Interaksi (MI) dan Minimum Observasi (MO). Untuk mendapatkan model terbaik dilakukan *trial and error* terhadap kombinasi BF, MI dan MO hingga diperoleh model MARS dengan nilai GCV paling kecil dan nilai *R-Square* paling besar. Hasil *trial and error* model MARS diperlihatkan pada Tabel 8 berikut.

Tabel 8. Hasil Trial and Error Model MARS

BF	MI	MO	R-Square	GCV
16	1	1	0,841	1,29144
24	1	1	0,841	1,29144
24	2	1	0,682	1,39377
24	3	1	0,682	1,39377
24	1	3	0,558	1,59768
24	3	3	0,700	1,35655
32	1	1	0,841	1,29144
32	2	1	0,682	1,39377
32	3	1	0,682	1,39377
40	1	1	0,841	1,29144
40	2	1	0,682	1,39377
40	3	1	0,682	1,39377

Berdasarkan Tabel 8 didapatkan Model MARS terbaik dengan kombinasi BF=24, MI=1 dan MO=1. Nilai GCV untuk model MARS terbaik sebesar 1,29144 dan R-Square sebesar 0,841 dan bentuk modelnya sebagaimana ditunjukkan pada persamaan berikut:

$$\hat{Y} = -22,2586 + 0,454497*BF1 - 0,933528*BF2 - 17,0652*BF3 + 9,95784*BF4 + 11,6213*BF5;$$

dengan

$$BF1 = \max(0, X6 - 0,35);$$

$$BF2 = \max(0, 3,35 - X6);$$

$$BF3 = \max(0, X7 - 5,11);$$

$$BF4 = \max(0, 5,11 - X7);$$

$$BF5 = \max(0, X7 - 2,57);$$

Dari model MARS di atas dapat diinterpretasikan sebagai berikut.

- BF1 = max (0, X6 – 3.35) dengan koefisien 0.454497, artinya bahwa setiap terdapat kenaikan BF1 sebanyak satu satuan, maka persentase kejadian balita stunting per kecamatan akan meningkat sebesar 0.454497 persen.
- BF2 = max (0, 3.35 – X6) dengan koefisien - 0.933528, artinya bahwa setiap terdapat kenaikan BF2 sebanyak satu satuan, maka persentase kejadian balita stunting per kecamatan akan turun sebesar -0.933528 persen.
- BF3 = max (0, X7 – 5,11) dengan koefisien - 17,0652, artinya bahwa setiap terdapat kenaikan BF3 sebanyak satu satuan, maka persentase kejadian balita stunting per kecamatan akan turun sebesar -17,0652 persen.
- BF4 = max (0, 5,11 – X7) dengan koefisien 9,95784, artinya bahwa setiap terdapat kenaikan BF4 sebanyak satu satuan, maka akan meningkatkan persentase kejadian balita stunting per kecamatan sebesar 9,95784 persen.

- BF5 = max (0, X7 – 2,57) dengan koefisien 11,6213, artinya bahwa setiap terdapat kenaikan BF5 sebanyak satu satuan, maka akan meningkatkan persentase kejadian balita stunting per kecamatan sebesar 11,6213 persen.

Jadi, variabel yang signifikan mempengaruhi persentase kejadian balita stunting per kecamatan di kabupaten Bojonegoro adalah persentase balita kurus mendapatkan makanan tambahan (X6) dengan besar kepentingan pada model sebesar 100 persen dan persentase rumah tangga yang menggunakan air bersih (X7) dengan besar kepentingan pada model sebesar 51,08 persen.

Untuk pengujian signifikansi parameter model MARS secara simultan, didapatkan nilai F-hitung sebesar 23,23962 dan p-value sebesar 0,00000. Dengan menggunakan nilai α sebesar 5 persen, maka H_0 ditolak karena nilai p-value < 0,05. Sehingga kesimpulannya adalah model MARS yang didapatkan signifikan dan dapat digunakan untuk memprediksi kejadian balita stunting.

Edangkan untuk pengujian signifikansi secara parsial, dengan menggunakan nilai α sebesar 5 persen, didapatkan nilai p-value < 0,05 pada keseluruhan estimator parameter. Hal ini berarti H_0 ditolak yang artinya konstanta dan koefisien BF1, BF2, BF3, BF4 dan BF5 mempunyai pengaruh yang signifikan terhadap model.

3.5 Perbandingan Model GWR dan MARS

Pemilihan model terbaik dilakukan dengan membandingkan model GWR dan model MARS dengan melihat nilai *R-Square* (R^2) dan nilai MSE pada masing-masing model yang ditunjukkan pada Tabel 9.

Tabel 9. Perbandingan Model GWR dengan MARS

Model	R-Square	MSE
Model	R-Square	MSE
GWR	0,8734	0,8620
MARS	0,8408	0.4305

Nilai R-Square pada model GWR lebih besar MARS meskipun dari nilai angkanya tidak jauh berbeda, sedangkan nilai MSE model MARS diperoleh nilai yang lebih kecil dari nilai MSE model GWR dengan selisih jauh beda. Dengan demikian model terbaik jatuh pada model MARS. Dengan terbentuknya model MARS sumbangsih keilmuan yang diberikan adalah model regresi MARS yang menjelaskan pengaruh variabel-variabel prediktor terhadap kejadian balita stunting di kabupaten Bojonegoro yaitu variabel prediktor

X6 (persentase balita kurus mendapatkan makanan tambahan) dan X7 (persentase rumah tangga yang menggunakan air bersih).

4. KESIMPULAN DAN SARAN

Karakteristik kejadian balita stunting di kabupaten Bojonegoro beserta faktor yang diduga mempengaruhinya, untuk persentase kejadian balita stunting per kecamatan paling rendah 1,03% dan paling tinggi 7,22%, persentase bayi baru lahir memperoleh Inisiasi Menyusui Dini (IMD) paling rendah 1,00% dan tertinggi 7,02%, persentase bayi memperoleh ASI eksklusif paling rendah 1,32% dan tertinggi 6,18%, persentase balita memperoleh vitamin A terendah 3,13% dan paling tinggi 4,08%, persentase balita mempunyai KMS atau buku KIA paling rendah 2,18% dan tertinggi 6,50%, persentase balita ditimbang empat kali atau lebih dalam enam bulan terakhir paling rendah 2,82% dan paling tinggi 4,09%, persentase balita kurus memperoleh makanan tambahan paling rendah 0,80% dan tertinggi 7,86%, persentase rumah tangga yang menggunakan air bersih paling rendah 2,39% dan tertinggi 5,75% serta persentase tenaga kesehatan paling rendah 1,08% dan tertinggi 8,60%.

Pada model GWR, variabel prediktor yang signifikan di masing-masing kecamatan hampir sama yaitu variabel X6 (persentase balita kurus mendapatkan makanan tambahan) kecuali untuk kecamatan Baureno semua variabel prediktor tidak signifikan mempengaruhi persentase kejadian balita stunting. Sedangkan untuk model MARS variabel yang signifikan mempengaruhi persentase kejadian balita stunting per kecamatan di kabupaten Bojonegoro adalah persentase balita kurus mendapatkan makanan tambahan (X6) dan persentase rumah tangga yang menggunakan air bersih (X7).

Berdasarkan hasil analisis, model terbaik dalam pemodelan kejadian balita stunting di kabupaten Bojonegoro adalah model MARS karena mempunyai nilai MSE yang lebih kecil.

DAFTAR PUSTAKA

- Al Azies, H., Cholid, F., & Trishnanti, D. (2019). Pemetaan Faktor-Faktor yang Mempengaruhi Stunting pada Balita dengan Geographically Weighted Regression (GWR). *Semnaskes 2019*, 156–165.
- Ardianti, D., Pramoedyo, H., & Nurjannah, N. (2021). Distance weight of GWR-Kriging model for stunting cases in East Java. *Journal of Physics: Conference Series*, 1968(1), 12028.
- Azizah, D. M., & Permatasari, E. O. (2020). Modeling of toddler stunting in the province of east nusa tenggara using multivariate adaptive regression splines (mars) method. *Journal of Physics: Conference Series*, 1490(1), 12013.
- BPS, B. P. S. (2016). *Bojonegoro dalam angka 2016*. 1–180.
- Dinas Kesehatan Kabupaten Bojonegoro. (2019). Profil Kesehatan Kabupaten Bojonegoro Tahun 2019. *Kemntrian Kesehatan*, 141.
- Fadliana, A., & Darajat, P. P. (2021). Pemetaan Faktor Risiko Stunting Berbasis Sistem Informasi Geografis Menggunakan Metode Geographically Weighted Regression. *IKRA-ITH INFORMATIKA: Jurnal Komputer Dan Informatika*, 5(3), 91–102.
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1), 1–67. <https://doi.org/10.1214/aos/1176347963>
- Friedman, J. H., & Roosen, C. B. (1995). An introduction to multivariate adaptive regression splines. *Statistical Methods in Medical Research*, 4(3), 197–217. <https://doi.org/10.1177/096228029500400303>
- Kasse, I. (2017). MEMODELKAN FAKTOR-FAKTOR YANG MEMPENGARUHI GIZI BURUK BALITA DENGAN METODE MULTIVARIATE ADAPTIVE REGRESSION SPLINE (MARS). *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, 5(1), 69.
- Kemendes RI. (2018). Buletin Stunting. *Kementerian Kesehatan RI*, 301(5), 1163–1178.
- Kesehatan, D. (2020). *KABUPATEN BOJONEGORO TAHUN 2020*.
- Kesehatan, K., & Indonesia, R. (2019). *Profil Kesehatan Indonesia Tahun 2019*.
- Lu, B., Brunson, C., Charlton, M., & Harris, P. (2017). Geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, 31(5), 982–998. <https://doi.org/10.1080/13658816.2016.1263731>
- Muktamar Umakaapa. (2020). *Strategi 2020 Melawan Stunting*.
- Ningrum, C. A. K. (2021). *Pemodelan Kejadian Stunting Pada Balita Di Surabaya Dengan Pendekatan Multivariate Adaptive Regression*

- Spline*. UNIVERSITAS AIRLANGGA.
- Ningsih, W. A. (2017). *Gambaran Kelengkapan Imunisasi Dasar Pada Balita Di Puskesmas Garuda Kota Bandung Tahun 2017*.
- Otok, B. W. (2014). *Pendekatan Multivariate Adaptive Regression SPLINES (MARS) pada Pemodelan Penduduk Miskin di Indonesia Tahun 2008-2012*.
- Pramoedyo, H., Mudjiono, A. A. R. F., Ardianti, D., & Septiani, K. (2020). Determination of Stunting Risk Factors Using Spatial Interpolation Geographically Weighted Regression Kriging in Malang. *Mutiara Medika: Jurnal Kedokteran Dan Kesehatan*, 20(2), 98–103.

Analisis Regulasi Uji Emisi Gas Buang Kendaraan Berdasarkan Pengaruhnya Terhadap Indeks Kualitas Udara di DKI Jakarta Menggunakan Metode Korelasi Pearson dan Regresi Linear

Muhammad Reza Akbar⁽¹⁾, Muhammad Ihsan Akbar⁽²⁾, dan Rizki Achmad Darajatun⁽³⁾

^(1,3)Universitas Singaperbangsa Karawang

Jl. HS.Ronggo Waluyo, Puseurjaya, Kabupaten Karawang, Jawa Barat 41361

⁽²⁾Institut Teknologi Sains Bandung

Jl. Ganesha Boulevard Kota Deltamas Lot-A1 CBD, Pasirranji, Kec. Cikarang Pusat, Kabupaten Bekasi, Jawa Barat 17530

e-mail: rezakbar.m@gmail.com , ihsanakbar756@gmail.com , dan dosen@rizkidarajatun.org

ABSTRAK

Perkembangan zaman terus berlangsung pesat akhir-akhir ini. Salah satu dampak pengaruh dari kemajuan sektor transportasi adalah berupa fleksibilitas aktivitas masyarakat kota dalam berkegiatan menjadi lebih efisien dan efektif disegala sisi, seperti budaya masyarakat kota dalam beraktivitas mulai mengalami transisi, dari sebelumnya mengutamakan kendaraan bertenaga hewan beralih kepada penggunaan sektor kendaraan berbahan bakar fosil. Dampak perihal tersebut memiliki pengaruh langsung terhadap kualitas udara di DKI Jakarta. Penelitian dilakukan berdasarkan hasil identifikasi masalah dan studi literatur sebagai pedoman penelitian, data yang dibutuhkan yaitu data ISPU dan kendaraan lulus uji emisi selama periode Januari hingga September 2021, langkah Pengolahan data tahap 1 yaitu uji inferensial, dan uji normalitas data, serta uji korelasi Pearson dan uji regresi linear ditahap 2. Informasi yang diperoleh kemudian disimpulkan disertai saran. Variabel penelitian yaitu data uji emisi sebagai variabel X dan data ISPU sebagai variabel Y. Hasil identifikasi menunjukkan efektivitas regulasi dalam mereduksi persoalan lingkungan terbukti efektif dalam menangkal persoalan kualitas udara di DKI Jakarta, dimana seiring dengan meningkatnya jumlah kendaraan yang lolos uji emisi berhasil menurunkan skor ISPU sebesar 82% diperankan oleh upaya uji emisi dan 18% sisanya dipengaruhi oleh variabel lainnya.

Kata kunci : *Kendaraan, Emisi Gas Buang, ISPU, Kualitas Udara, Korelasi Pearson, Regresi Linear.*

ABSTRACT

The development of the times continues to take place rapidly lately. The impacts of the progress of the transportation sector are in the form of flexibility in the activities of urban communities in carrying out their activities to be more efficient and effective on all sides, such as the culture of urban communities in their activities starting to experience a transition, from prioritizing animal-powered vehicles to the use of the fossil fuel vehicle sector. The impact of this matter has a direct influence on the air quality in DKI Jakarta. The research was conducted based on the results of problem identification and literature studies as research guidelines, the data needed are ISPU and vehicles that pass the emission test during the period January to September 2021, step 1 data processing is an inferential test, and data normality test, then Pearson correlation test and linear regression test in stage 2. The information obtained is then concluded with suggestions. The research variables are emission test data as variable X and ISPU data as variable Y. The identification results show that the effectiveness of regulations in reducing environmental problems has proven to be effective in preventing air quality problems in DKI Jakarta, where along with the increasing number of vehicles that pass the emission test, the ISPU score has been reduced. of 82% is played by emission test efforts and the remaining 18% is influenced by other variables.

Keywords : *Vehicles, Exhaust Emissions, ISPU, Air Quality, Pearson Correlation, Linear Regression.*

1. PENDAHULUAN

Perkembangan zaman terus berkembang pesat akhir-akhir ini. Maju dan berkembangnya sektor tersebut telah membawa banyak perubahan di multisektor kehidupan. Salah satu dampak pengaruh dari kemajuan transportasi adalah berupa fleksibilitas aktivitas masyarakat kota dalam berkegiatan menjadi lebih efisien dan efektif disegala sisi, dimana budaya masyarakat kota dalam beraktivitas mulai mengalami transisi, dari sebelumnya mengutamakan penggunaan kendaraan bertenaga hewan yang kemudian mulai beralih kepada penggunaan mereka di sektor kendaraan berbahan bakar fosil. Walaupun demikian, tentunya akan ada dampak eksternalitas tersendiri yang diakibatkan olehnya terhadap aspek-aspek lain. Salah satu dampak positif adalah mampu menyalurkan komunikasi pelaku ekonomi disetiap lintas ruangannya dalam melakukan berbagai aktivitas yang dikehendaki secara lebih fleksibel dan produktif disegala sisi yang juga merupakan bagian dari representasi konsep penghematan aglomerasi dan urbanisasi dalam bidang ekonomika. Jika sebelumnya peran transportasi hanya berjasa dalam memacu aktivitas laju perekonomian wilayah kota, namun tidak peranannya terhadap kondisi lingkungan disekitarnya. Pasalnya, jenis kendaraan yang mendominasi di Indonesia, khususnya di DKI Jakarta saat ini masih didominasi oleh kendaraan berbahan bakar fosil yang umumnya dikenal sebagai sektor penyumbang gas polutan.

Dampak buruk gas polutan terhadap lingkungannya adalah berupa tercemarnya kualitas udara yang dapat mengganggu Kesehatan masyarakat. Pernyataan tersebut merujuk kepada hasil dari penelitian *Air Visual* di tahun 2019 yang menyebutkan bahwa kualitas udara disaat itu mengandung *toxic* polutan PM 2.5 sebesar 66.6 mg/m³. Angka tersebut telah melebihi batas minimal baku kenormalannya sebesar 65 mg/m³ yang berimplikasi kepada dinobatkannya Jakarta sebagai 10 kota di dunia yang tidak ramah lingkungan secara kualitas ambien udaranya (Ramdani, 2019). Dengan demikian, upaya penanggulangan dan pengendalian lingkungan perlu untuk dilakukan dalam upaya menyikapi isu lingkungan di DKI Jakarta yang semakin mengkhawatirkan. Salah satu upaya yang telah dilakukan oleh Pemprov DKI Jakarta bersama (Dinas Lingkungan Hidup dan Kebersihan) DLHK adalah pelaksanaan regulasi uji emisi disetiap kendaraan yang beroperasi di DKI Jakarta. Harapannya dengan terselenggarakannya upaya tersebut dapat menekan potensi gas buang emisi

kendaraan yang tak ramah lingkungan terhadap kualitas ambien udara di DKI Jakarta. Maka dari itu, penulis akan mengkaji secara matematis mengenai efektivitas regulasi uji emisi terhadap kualitas udara di DKI Jakarta selama periode januari hingga September 2021. Penelitian menggunakan data ISPU sebagai variabel terikat (Y) dan data kendaraan lolos uji emisi sebagai variabel bebas (X) melalui penggunaan dua metode hitung yakni analisis korelasi untuk mengetahui keterikatan hubungan sebab-akibat dan penggunaan analisis regresi sebagai prediksi pengaruh variabel X terhadap Y.

Penelitian ini dilakukan berdasarkan referensi dari penelitian terdahulu di antaranya:

1. Cindy Viane Bertan, A. K. T. Dundu, dan R. J. M. Mandagi yang berjudul “Pengaruh Pendayagunaan Sumber Daya Manusia (Tenaga Kerja) Terhadap Hasil Pekerjaan (Studi Kasus Perumahan Taman Mapanget Raya (Tamara))”. Penelitian dilakukan untuk menganalisis korelasi Pearson antara data pendayagunaan tenaga kerja sebagai variabel X dengan hasil pekerjaan sebagai variabel Y. Hasil penelitian menunjukkan bahwa korelasi bersifat positif yang signifikan sehingga dapat diketahui semakin besar nilai variabel X maka variabel Y semakin besar atau semakin cepat proyek tersebut diselesaikan. Penelitian juga melakukan uji regresi linear yang menghasilkan bahwa data berkaitan positif satu sama lain (Bertan, Dundu, & Mandagi, 2016), dan
2. Rahmana Dwi Saputra dengan Syarif Hidayat yang berjudul “Implementasi Regresi Linier untuk Prediksi Penjualan dan *Cash Flow* Pada Aplikasi *Point of Sales Restoran*”. Penelitian dilakukan untuk menganalisis uji regresi liner antara data penjualan (penjualan makanan dan minuman) dengan data *cash flow* (biaya produksi, pendapatan, biaya lain-lain, dan profit) restoran. Hasil penelitian menunjukkan bahwa peningkatan penjualan, modal, dan profit sangat berkaitan satu sama lain dimana jika penjualan bertambah maka modal yang dibutuhkan setiap bulannya pun juga bertambah sehingga mempengaruhi profit yang didapat setiap bulannya (Shaputra & Hidayat, 2021).

1.1. Geografis DKI Jakarta



Gambar 1. Administrasi DKI Jakarta

Kota DKI Jakarta merupakan sebuah wilayah dataran rendah dengan rata-rata ketinggian wilayah adalah 7 meter di atas permukaan laut. Berdasarkan SK Gubernur Nomor 171 tahun 2007, proporsi luas wilayah Provinsi DKI Jakarta adalah berupa daratan seluas 662,33 km² dan berupa lautan seluas 6.977,5 km². Pembagian wilayah administrasi diantaranya Jakarta Pusat dengan kisaran luas daratan 47,90 km², Jakarta Utara dengan kisaran luas daratan 154,01 km², Jakarta Barat dengan kisaran luas daratan 126,15 km², Jakarta Selatan dengan kisaran luas daratan 145,73 km², Jakarta Timur dengan kisaran luas daratan 187,73 km² (BPS Jakarta, 2019).

1.2. Aspek Kependudukan DKI Jakarta

Aspek kependudukan merupakan suatu komponen terpenting terhadap keberlangsungan sebuah wilayah kota dikarenakan mereka adalah pelaku berlangsungnya berbagai aktivitas setiap ruang kota. Selain itu keberadaan mereka dalam sebuah wilayah kota juga merupakan faktor penentu terhadap esensi keberlanjutan wilayah tersebut, khususnya mengenai kondisi kota di masa kini hingga dikedepannya. Keberlanjutan yang dimaksud disini adalah suatu kondisi dimana kehidupan masyarakat perkotaan atau pembangunan perkotaan dapat terpenuhi dengan baik tanpa sedikitpun mengurangi esensi ketersediaan sumber daya maupun kesempatan masyarakat dalam berkehidupan dikedepannya. DKI Jakarta secara demografis data kependudukan terus menunjukkan trend yang meningkat disetiap tahunnya. Pola yang ditunjukkan atasnya adalah secara tak menentu atau mengikuti kepada kaidah deret hitung yakni pola laju kependudukan disaat ini tidak menjamin pola yang sama kembali diwaktu-waktu berikutnya.

Tabel 1. Demografi Penduduk DKI Jakarta

Demografi Penduduk Jakarta				laju pertumbuhan penduduk (%)	
Kab/Kota	2018 (ribu)	2019 (ribu)	2020 (ribu)	2019-2018	2020-2019
Kep Seribu	24.13	24.3	27.75	0.7%	14%
Jakarta Selatan	2246.14	2264.7	2226.81	0.8%	-2%
Jakarta Timur	2916.02	2937.86	3037.14	0.7%	3%
Jakarta Pusat	924.69	928.11	1056.9	0.4%	14%
Jakarta Barat	2559.36	2589.93	2434.51	1.2%	-6%
Jakarta Utara	1747.31	1812.91	1778.98	3.8%	-2%
DKI Jakarta	10417.65	10557.81	10562.09	1.3%	0%
Total persentase				8.9%	22%

1.3. Transportasi DKI Jakarta

Definisi transportasi adalah kegiatan pemindahan barang dan penumpang dari suatu tempat ke tempat lain (Abbas, 2012). Transportasi juga didefinisikan sebagai usaha dan kegiatan mengangkut atau membawa barang dan penumpang dari suatu tempat ke tempat lainnya (Sugianto & Kurniawan, 2020). Sehingga dapat disimpulkan bahwa transportasi adalah kegiatan pemindahan suatu muatan baik barang maupun penumpang dari suatu tempat ke tempat lain.

Sektor transportasi merupakan salah satu sektor yang memiliki peranan terpenting dalam mejalin macam aktivitas masyarakat kota di berbagai aspeknya. Salah satu dampak positif atas penggunaan transportasi adalah berupa terciptanya iklim komunikasi ekonomi yang terkoneksi dengan baik disetiap antar pelaku ekonomi lintas ruang dan sektornya yang berdampak kepada meningkatnya utilitas aktivitas perekonomian sebuah wilayah kota secara produktif dan menguntungkan. Walaupun demikian, peranan transportasi terhadap lingkungannya justru tidak sebaik peranan mereka terhadap aspek perekonomian. Hal tersebut didasarkan kepada dampak pengaruh gas buang emisi yang dihasilkan disetiap jenis kategori kendaraannya. Berikut adalah komparasi kadar emisi gas buang disetiap jenis kendaraannya

Tabel 2. Komparasi Kadar Emisi Gas Buang Disetiap Jenis Kendaraannya

Jenis Teknologi Mesing Penggerak	Kadar Gas Emisi Yang Dihasilkan
Kendaraan listrik murni (EV)	0 gram/Km emisi CO ₂
Kendaraan listrik kombinasi listrik dan bensin (PHEV)	45 gram/Km emisi CO ₂
Kendaraan hybrid bensin dan baterai	70 gram/Km emisi CO ₂
Kendaraan konvensional	125 gram/Km emisi CO ₂

Berdasarkan tabel 2 diketahui bahwa di DKI Jakarta sendiri dominasi pengguna ruas jalan saat ini masih didominasi oleh jenis berbahan bakar fosil atau

kendaraan konvensional. Artinya potensi pencemaran lingkungan disana sangat mungkin terjadi atas kondisi kota yang demikian. Adapun, informasi statistik mengenai transportasi di ibu kota adalah sebagai berikut.

Tabel 3. Frekuensi Kendaraan DKI Jakarta

Jenis kendaraan	2016	2017	2019	2020	Jumlah Total	rasio prosentase
sepeda motor	13310672	14063540	15644530	16018716	59037458	75.4%
mobil penumpang	3525925	3711351	4064836	4061033	15363145	19.6%
truk	689561	708918	763374	772672	2934525	3.7%
bus	141516	144603	342036	342835	970990	1.2%
jumlah	17667674	18628412	20814776	21195256	78306118	

Berdasarkan tabel 3 diketahui bahwa proporsi kendaraan terbanyak di DKI Jakarta saat ini berasal dari sektor kendaraan sepeda bermotor dengan rasio prosentase sebesar 75.4% lebih banyak dibandingkan jenis kendaraan lainnya.

1.4. Kaidah ISPU Sebagai Indeks Kualitas Udara

Indeks Standar Pencemar Udara (ISPU) adalah laporan kualitas udara kepada masyarakat untuk menerangkan seberapa bersih atau tercemarnya udara dan bagaimana dampaknya terhadap kesehatan masyarakat setelah menghirup udara tersebut selama beberapa jam atau hari (Sianipar, 2017). ISPU juga didefinisikan sebagai angka yang tidak mempunyai satuan yang menggambarkan kondisi mutu udara ambien di lokasi tertentu, yang didasarkan kepada dampak terhadap kesehatan manusia, nilai estetika dan makhluk hidup lainnya (MENLH, 2020). Sehingga dapat disimpulkan bahwa ISPU adalah suatu alat ukur dan indikator kualitas udara pada suatu lingkungan kota dengan mengubah kadar pencemar udara yang terukur menjadi suatu angka yang tidak berdimensi. Maksud dari suatu angka yang tidak berdimensi adalah berubahnya konsentrasi kadar zat pencemar seperti mg/m3 yang telah didapatkan sebelumnya untuk kemudian dikonversikan kedalam bentuk bilangan yang dikehendaki sesuai dengan kaidah satuan ISPU dalam KEP 45/MENLH/1997 guna mempermudah informasi yang tersampaikan. Adapun, parameter zat pencemarnya terdiri dari Partikulat (PM10), Karbondioksida (CO), Sulfur dioksida (SO2), Nitrogen dioksida (NO2), dan Ozon (O3).

1.5. Regulasi Uji Emisi

Uji emisi merupakan suatu regulasi pengendalian lingkungan yang ditujukan untuk mereduksi emisi gas buang yang dihasilkan oleh setiap jenis kendaraan yang melintas di ruang wilayah kota. Di DKI Jakarta sendiri regulasi tersebut baru dilaksanakan pada awal Januari 2021. Dengan teralisasinya regulasi tersebut setidaknya dapat membantu pemerintah dalam

memantau tindak pelanggaran lingkungan yang dilakukan oleh setiap pengguna kendaraan yang tak ramah lingkungan secara emisi gas buang kendaraannya sehingga akan diberlakukan sanksi kepada pelanggar berupa disentif tilang, denda, hingga pada peneraan sanksi larangan untuk melintas di DKI Jakarta apabila telah diluar batas minimum yang telah ditentukan.

Secara teknis pelaksanaannya kendaraan yang telah tiba dan sudah saatnya untuk dilakukan pengujian akan diminta untuk mengondisikan kendaraannya dengan keadaan mesin menyala di putaran 1.900-2.00 rpm (rotasi per menit) dengan intensitas suhu kerja sekitar 60 hingga 70 derajat celsius. Kemudian kendaraan akan diinjeksikan dengan alat pengukur gas buang bernama *exhaust gas analyzer* untuk merekam gas buang yang dihasilkan oleh kendaraan tersebut selama 20 detik lamanya, lalu melakukan penilaian berdasarkan kriteria kendaraan tersebut apakah lulus atau tidaknya. Konsep penilaian regulasi tersebut adalah ketika gas buang kendaraan melebihi batas baku kenormalan yang telah ditetapkan disetiap kategori kendaraannya maka kendaraan tersebut tidak lulus sebagai kendaraan yang ramah lingkungan dan begitupun sebaliknya. Untuk lebih jelasnya, maka berikut adalah syarat kelulusan disetiap kategori kendaraannya,

1. Mobil berbahan bakar bensin. Kategori mobil dengan tahun produksi di bawah 2007 harus memiliki kadar CO2 di bawah 3% dan untuk kendaraan tahun produksi di atas 2007 harus memiliki kadar CO2 tidak lebih dari 1.5%,
2. Mobil berbahan bakar diesel. Kategori Mobil diesel dikisaran bobot 3.5 dan merupakan tahun produksi di bawah 2010 harus memiliki kadar opasitas tidak lebih dari 50%. Adapun, untuk mesin diesel dengan bobot yang sama namun tahun produksi di atas 2010 harus memiliki kadar opasitas 40%.

Adapun persyaratan yang harus dipenuhi lebih dulu bagi calon pesertanya adalah untuk rutin melakukan cek perawatan berkala di minggu-minggu sebelum pengujianya. Hal tersebut bertujuan untuk meminimalisir kinerja mesin kendaraan yang tidak efisien disaat kelangsungan pengujian emisi yang berimplikasi kepada tidak optimalnya output gas buang kendaraan yang dihasilkan disaat itu. Oleh karena itu tujuan dari dilaksanakannya regulasi tersebut tidak hanya untuk mereduksi persoalan lingkungan, melainkan juga dalam menumbuhkan rasa kesadaran diri atas perilaku masyarakat terhadap lingkungan sekitar.

1.6. Korelasi Pearson / KPM

Korelasi Pearson adalah alat uji statistik yang digunakan untuk menguji hipotesis asosiatif (uji

hubungan) dua variabel bila datanya berskala interval atau rasio. KPM dikembangkan oleh Karl Pearson (Oktavia, 2021). Korelasi Pearson juga didefinisikan sebagai metode yang mencari hubungan diantara variabel-variabel yang diteliti yang kemudian akan dijelaskan. Metode ini bertujuan untuk meneliti sejauh mana hubungan antara suatu variabel dengan variabel lainnya (Rufaidah, Erwina, & Yanto, 2019). Sehingga dapat disimpulkan bahwa korelasi Pearson adalah suatu metode uji statistik mengenai korelasi antara dua variabel berskala interval-rasio dalam mengetahui sifat, arah, dan tingkat hubungan diantara kedua variabel tersebut. Disebutnya sebagai KPM karena metode tersebut dikembangkan oleh Karl Pearson. Sehingga metode tersebut hanya diperuntukkan bagi data statistik berskala interval-rasio dalam menguji hipotesis asosiatif mengenai hubungan korelasi diantara kedua variabel uji tersebut. Berdasarkan jenisnya, uji korelasi terbagi menjadi dua yaitu,

1. Korelasi positif (searah): jika variasi terjadi searah atau skor tinggi pada variabel X diikuti oleh skor yang tinggi pula pada variabel Y, dan sebaliknya, skor rendah pada variabel X diikuti oleh skor yang rendah pada variabel Y,
2. Korelasi negatif (berlawanan): jika variasi terjadi sebaliknya, yakni skor tinggi pada variabel X diikuti oleh skor yang rendah pada variabel Y, dan sebaliknya, skor rendah pada variabel X diikuti oleh skor yang tinggi pada variabel Y

Koefisien korelasi dalam tingkat hubungannya sendiri terklasifikasikan kedalam 5 kelas interval sebagai berikut,

Tabel 4. Interval Koefisien Korelasi

Interval koefisien	Tingkat hubungan
0,00 – 0,199	Sangat rendah
0,20 – 0,399	Rendah
0,40 – 0,599	Sedang
0,90 – 0,799	Kuat
0,80 – 1,000	Sangat kuat

1.7. Regresi Linear

Regresi linear merupakan persamaan matematik yang memungkinkan kita meramalkan nilai-nilai suatu peubah tak bebas (y) dari satu nilai peubah bebas (x) yang hasilnya dapat berupa suatu persamaan linear serta garis regresi linear (Azizah, 2019). Regresi linear juga didefinisikan sebagai hubungan model secara linier antara variabel dependen dan variabel bebas untuk memprediksi nilai dari dependen jika variabel bebas mengalami

perubahan nilai (Shaputra & Hidayat, 2021). Sehingga dapat disimpulkan bahwa regresi linear adalah sebuah metode perhitungan yang bertujuan untuk mengetahui prediksi efektivitas variabel Y oleh X. Metode tersebut membantu peneliti untuk mengetahui seberapa besar prosentase Y yang diterangkan oleh X, serta dapat membantu peneliti dalam memprediksi suatu angka yang diharapkan terjadi dalam penelitiannya. Persamaan regresinya adalah sebagai berikut.

$$Y' = a + bX \tag{1}$$

Dari persamaan diatas perlu mengetahui lebih dulu maksud dari setiap variabel penyusunnya. Berikut adalah penjabarannya.

1. a (intersep) merupakan prediksi peningkatan nilai Y disaat nilai X = 0,
2. b (slop) merupakan koefisien arah regresi linier yang menunjukkan perubahan nilai Y disetiap peningkatan nilai X sebesar satu unit. Koefisien tersebut berfungsi dalam meramalkan nilai Y bila nilai X diketahui, sehingga b (slop) merupakan variabel indenpenden (bebas) sedangkan a (intersep) merupakan variabel dependen (terikat atau terpengaruh).

Pada umumnya data yang didapatkan tidak memiliki pola yang pasti, sehingga rumus disetiap variabel persamaannya adalah sebagai berikut, dengan lebih dulu untuk membuat tabel sebagai berikut.

Tabel 5. Perhitungan Koefisien Korelasi

X	x = X- Xrataan	Y	y =Y- Yrataan	xy	x ²	y ²

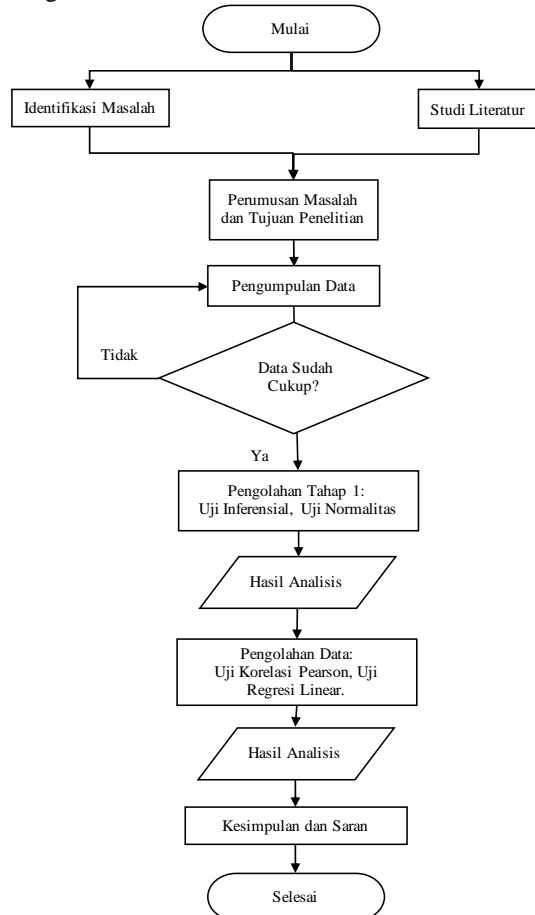
$$b = \frac{\sum xy}{\sum x^2} \tag{2}$$

$$a = \bar{Y} - b\bar{X} \tag{3}$$

2. METODE PENELITIAN

Penelitian menggunakan metode statistika kuantitatif deskriptif yaitu sebuah penelitian yang terfokuskan kepada makna informasi yang didapatkan dari sebuah angka bilangan yang terolah. Dimana nantinya, data yang telah terolah dijadikan sebagai bahan pendukung hipotesis di penelitian. Penelitian ini dilakukan pada Desember 2021 dengan format data yang digunakan adalah terhitung sejak Januari-September 2021. Penelitian dilakukan secara terintegrasi dengan media digital yakni memanfaatkan kondisional objek penelitian melalui sarana tersebut. Untuk lokasi penelitian yaitu DKI

Jakarta yang kemudian diperinci skala objek penelitiannya dengan memilih data lokasi uji sesuai dengan variabel datanya yaitu data kendaraan lulus uji emisi sebagai variabel X dan uji ISPU sebagai variabel Y. *Flow chart* proses penelitian adalah sebagai berikut.



Gambar 2. *Flow Chart* Proses Penelitian

Uraian flow chart proses penelitian adalah sebagai berikut:

1. Melakukan penelitian berdasarkan identifikasi masalah disertai studi literatur sebagai pedoman penelitian,
2. Menyusun rumusan masalah dan tujuan penelitian,
3. Pengumpulan data yaitu data ISPU dan kendaraan lulus uji emisi selama periode Januari hingga September 2021, serta memastikan data yang diperoleh tercukupi,
4. Pengolahan data tahap 1 yaitu uji inferensial dan uji normalitas data serta analisisnya,
5. Pengolahan data tahap 2 yaitu uji korelasi Pearson dan uji regresi linear, dan
6. Kesimpulan dan saran.

3. HASIL DAN PEMBAHASAN

3.1. Data Program Regulasi Uji Emisi Kendaraan DKI Jakarta

Upaya penanggulangan dan pengendalian lingkungan perlu untuk dilakukan dalam menyikapi isu lingkungan di DKI Jakarta yang semakin mengkhawatirkan. Salah satu upaya yang telah dilakukan oleh Pemprov DKI Jakarta bersama DLHK adalah pelaksanaan regulasi uji emisi disetiap kendaraan yang beroperasi di DKI Jakarta dengan harapan dapat menekan potensi emisi gas buang kendaraan yang tak ramah lingkungan terhadap kualitas ambien udara di kota tersebut.

Tabel 6. Tabel Uji Emisi Kendaraan Terhadap ISPU di DKI Jakarta

BULAN	BENSIN	keterangan	BULAN	SOLAR	keterangan
JANUARI	37525	lulus	JANUARI	5390	lulus
FEBRUARI	27335	lulus	FEBRUARI	4067	lulus
MARET	28553	lulus	MARET	4340	lulus
APRIL	22508	lulus	APRIL	3148	lulus
MEI	16060	lulus	MEI	1962	lulus
JUNI	15512	lulus	JUNI	2571	lulus
JULI	6213	lulus	JULI	1061	lulus
AGUSTUS	8935	lulus	AGUSTUS	1136	lulus
SEPTEMBER	10292	lulus	SEPTEMBER	1993	lulus
JANUARI	487	tidak lulus	JANUARI	290	tidak lulus
FEBRUARI	92	tidak lulus	FEBRUARI	15	tidak lulus
MARET	58	tidak lulus	MARET	16	tidak lulus
APRIL	32	tidak lulus	APRIL	11	tidak lulus
MEI	16	tidak lulus	MEI	3	tidak lulus
JUNI	23	tidak lulus	JUNI	15	tidak lulus
JULI	8	tidak lulus	JULI	1	tidak lulus
AGUSTUS	13	tidak lulus	AGUSTUS	6	tidak lulus
SEPTEMBER	41	tidak lulus	SEPTEMBER	20	tidak lulus

Berdasarkan tabel 6 untuk menguji efektivitas kebijakan tersebut, peneliti akan mengkaji efektivitas regulasi uji emisi terhadap kondisi kualitas udara di Jakarta khususnya dalam mereduksi persoalan pencemaran lingkungan secara matematis sesuai dengan urgensi peneliti dalam menyusun teks ilmiah ini.

3.2. Uji Inferensial Data Variabel X

Uji inferensial merupakan tahapan pengambilan sampel dengan menyeleksi data populasi yang ada sesuai dengan kaidah teoritik yang berlaku. pada tahap ini, uji inferensial akan dilakukan kepada variabel X yakni data uji emisi kendaraan yang terdata di Jakarta sebagaimana yang telah tersaji pada tabel 6. Diketahui total data variabel X adalah sebesar 32 secara terbagi, yakni 16 untuk data kendaraan lulus uji emisi dan sisanya adalah data kendaraan tidak lulus uji emisi. Dikarenakan data diatas masih berupa data populasinya, sehingga diperlukan proses penyeleksian data terlebih dahulu sebelum melakukan langkah pengujian di tahap selanjutnya.

$$s = \frac{n}{N} x S \tag{4}$$

s = Jumlah sampel setiap unit secara proporsional
 S = Jumlah seluruh sampel yang didapat
 N = Jumlah populasi
 n = Jumlah masing-masing unit populasi

Teknik di atas merupakan teknik *proportional random sampling* yaitu sebuah metode inferensial sampel yang didapatkan dari filterisasi data anggota kelompok populasi yang heterogen “bulan”. Maka dari itu, menurut Sugiyono penentuan sampel untuk masing-masing kelompok sampel yang heterogen menggunakan rumus sebagaimana yang tertera di atas (Lestari, 2014). Adapun, proses dan hasil inferensial data populasinya adalah sebagai berikut.

Tabel 7. Proses Inferensial dan Sampel Data Variabel X ISPU di DKI Jakarta

Variabel X (uji emisi dari kendaraan)				
Bulan	Total Populasi	s=n/N *30	Unit Sampel	
Januari	42915	Proses filterisasi data populasi untuk dijadikan sampel	6,482595757	
Februari	31402		4,743480647	
Maret	32893		4,968706099	
April	25656		3,875509187	
Mei	18022		2,722342788	
Juni	18083		2,731557243	
Juli	7274		1,098786008	
Agustus	10071		1,521291434	
September	12285		1,855730837	
Total Populasi	198601			

Berdasarkan tabel 7, data populasi yang digunakan di atas adalah data kendaraan berbahan bakar bensin dan solar yang telah *lulus uji emisi* saja untuk kemudian dikorelasikan dengan ISPU. Data subgroup perbulan sejumlah 30 mengikuti ketentuan sampel yang diambil yaitu sejumlah 30 karena metode yang akan digunakan adalah metode korelasi sehingga berdasarkan teoritik tertentu yaitu Gay berpendapat bahwa ukuran minimum sampel yang dapat diterima berdasarkan metode penelitiannya korelasi adalah sebesar 30 subjek (Lestari, 2014).

3.3. Uji Inferensial Data Variabel Y

ISPU merupakan satuan ukur indeks kualitas udara disebut kota. kedudukan ISPU pada penelitian ini diposisikan sebagai variabel Y yang diperankan sebagai variabel dependen atau variabel yang dipengaruhi atau bisa juga disebut sebagai variabel terikat. Sama halnya dengan jangka waktu

yang digunakan pada variabel X yaitu terhitung sejak Januari-September 2021, maka di variabel tersebut juga akan disesuaikan dengan ketentuan yang sama seperti yang ada pada variabel X. Total populasi variabel Y berkisar 156 data berupa penyajian tabel proses inferensial data variabel Y. Tabel proses inferensial data variabel Y di DKI Jakarta bulan Januari hingga September. Contoh analisis pengambilan data variabel Y diambil pada bulan Januari, sebagai berikut.

Tabel 8. Proses Inferensial Data Variabel Y DKI Jakarta Bulan Januari

DATA BULAN JANUARI											
tanggal	stasiun	pm10	pm25	so2	co	o3	no2	max	critical	kategori	KODE
01/01/21	DK11 (Bunderan HI)	38	53	29	6	21	13	33	PM25	SEDANG	2
02/01/21	DK11 (Bunderan HI)	27	46	27	7	47	7	47	O3	BAIK	1
03/01/21	DK11 (Bunderan HI)	44	58	25	7	40	13	58	PM25	SEDANG	2
04/01/21	DK11 (Bunderan HI)	30	48	24	4	32	7	48	PM25	BAIK	1
05/01/21	DK11 (Bunderan HI)	38	53	24	4	31	9	53	PM25	SEDANG	2
06/01/21	DK11 (Bunderan HI)	41	58	23	13	46	13	58	PM25	SEDANG	2
07/01/21	DK11 (Bunderan HI)	35	47	22	6	39	10	47	PM25	BAIK	1
08/01/21	DK11 (Bunderan HI)	37	54	26	56	27	10	54	PM25	SEDANG	2
09/01/21	DK11 (Bunderan HI)	47	61	16	27	22	12	61	PM25	SEDANG	2
10/01/21	DK11 (Bunderan HI)	23	25	16	11	33	8	33	O3	BAIK	1
11/01/21	DK11 (Bunderan HI)	38	54	17	14	27	10	54	PM25	SEDANG	2
12/01/21	DK11 (Bunderan HI)	29	50	20	12	26	15	50	PM25	BAIK	1
13/01/21	DK11 (Bunderan HI)	34	44	17	13	20	9	44	PM25	BAIK	1
14/01/21	DK11 (Bunderan HI)	36	57	15	15	26	13	57	PM25	SEDANG	2
15/01/21	DK11 (Bunderan HI)	47	62	14	15	32	12	62	PM25	SEDANG	2
16/01/21	DK11 (Bunderan HI)	52	78	15	17	25	14	78	PM25	SEDANG	2
17/01/21	DK11 (Bunderan HI)	52	71	17	15	26	14	71	PM25	SEDANG	2
18/01/21	DK11 (Bunderan HI)	40	52	16	11	26	12	52	PM25	SEDANG	2
19/01/21	DK11 (Bunderan HI)	31	41	13	12	25	12	41	PM25	BAIK	1
20/01/21	DK11 (Bunderan HI)	35	34	12	12	22	13	35	PM10	BAIK	1
21/01/21	DK11 (Bunderan HI)	43	51	13	14	28	15	51	PM25	SEDANG	2
22/01/21	DK11 (Bunderan HI)	49	63	14	14	28	15	63	PM25	SEDANG	2
23/01/21	DK11 (Bunderan HI)	72	108	14	43	44	20	108	PM25	TIDAK SEHAT	3
24/01/21	DK11 (Bunderan HI)	33	48	12	14	23	10	48	PM25	BAIK	1
25/01/21	DK11 (Bunderan HI)	56	82	13	29	62	15	82	PM25	SEDANG	2
26/01/21	DK11 (Bunderan HI)	53	70	14	19	28	13	70	PM25	SEDANG	2
27/01/21	DK11 (Bunderan HI)	20	20	13	10	19	8	20	PM10	BAIK	1
28/01/21	DK11 (Bunderan HI)	27	41	14	18	17	11	41	PM25	BAIK	1
29/01/21	DK11 (Bunderan HI)	30	53	16	17	24	11	53	PM25	SEDANG	2
30/01/21	DK11 (Bunderan HI)	27	37	13	16	24	12	37	PM25	BAIK	1

Pada tabel tersebut diketahui nilai maksimal ditandai dengan warna hijau perhariannya, misal diambil contoh yaitu pada periode 1 Januari 2021 dengan nilai maksimal sebesar 53 dihasilkan pada pengukuran pm25 yang dikategorikan sedang atau kode 2, begitupun data pada periode berikutnya sampai di bulan September 2021. Setelah nilai maksimal setiap periode diketahui kemudian di hitung nilai rata-rata perbulan untuk disajikan pada tabel 9. Berikut adalah data yang dihitung,

Tabel 9. Sampel Variabel Y

Variabel Y (ISPU)	
Bulan	Rataan
Januari	47,53333
Februari	54,13333
Maret	67,12903
April	72,83871
Mei	73,54839
Juni	80,6129
Juli	86,32258
Agustus	77,58065
September	75,74194

3.4. Uji Normalitas Variabel X dan Y

Sebagaimana yang telah dijelaskan pada bagian sebelumnya, bahwa syarat untuk dapat menggunakan analisis parametrik seperti korelasi R-Pearson dan regresi linear adalah data harus berdistribusi normal. Sehingga dilakukan uji normalitas untuk mengetahui sifat kenormalan dari kedua variabel tersebut melalui media alat bantu SPSS. Berikut adalah interpretasi hasil operasi hitung di SPSS,

Tabel 10. Interpretasi Hasil Uji Normalitas Variabel X dan Y

Data Variabel Uji	
Uji Emisi	ISPU
6,482595757	47,5333333
4,743480647	54,1333333
4,968706099	67,1290323
3,875509187	72,8387097
2,722342788	73,5483871
2,731557243	80,6129032
1,098786008	86,3225806
1,521291434	77,5806452
1,855730837	75,7419355

Adapun menurut kaidah teoritik yang berlaku mengenai sifat kenormalan data adalah sebagai berikut,

1. Jika $\text{sig.}(2\text{-tailed}) > 0.05$, maka data terdistribusi secara normal,
2. Jika $\text{sig.}(2\text{-tailed}) < 0.05$, maka data tidak terdistribusi normal.

Sebelum menginterpretasikan sifat kenormalannya, data lebih dulu harus disesuaikan kepada jenis kategori yang berlaku sebagai berikut,

1. Kolmogorov-smirnov, merupakan rujukan kenormalan data untuk jumlah data lebih dari 30 ($n > 30$),
2. Shapiro-wilk, merupakan rujukan kenormalan data untuk jumlah data kurang dari 30 ($n < 30$).

Setelah dilakukan kalkulasi pada media tersebut, hasil kalkulasi yang didapatkan adalah data terdistribusi normal dengan kriteria data sampel adalah ($n < 30$) serta *output* data $\text{sig.}(2\text{-tailed})$ yang dihasilkan pada tabel shapiro-wilk menunjukkan bilangan yang lebih dari 0.05. Dengan demikian, terbukti bahwa data sampel tersebut terdistribusi secara normal.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
UJI_EMISI	.238	9	.152	.911	9	.323
ISPU	.186	9	.200 [*]	.947	9	.659

^{*} This is a lower bound of the true significance.
^a Lilliefors Significance Correction

Gambar 3. Hasil Uji Normalitas

3.5. Uji Korelasi KPM-Pearson

Setelah didapatkannya data sampel melalui proses inferensial pada tahap sebelumnya serta telah teridentifikasinya sifat distribusi data sampelnya adalah normal, maka pengujian uji korelasi di tahap ini dapat dilakukan. dilakukannya uji korelasi tersebut, didasarkan kepada keingintahuan penulis mengenai apakah dari kedua data sampel yang telah didapatkan sebelumnya memiliki korelasi yang berarti atau tidak?. Hal tersebut sesuai dengan urgensi atas disusunnya teks penelitian ini. Untuk itu, berikut adalah komparasi antara data variabel X (uji emisi) dan variabel Y (ISPU) yang terhitung sejak Januari-September 2021 beserta interpretasinya.

Adapun, ketentuan dasar dalam pengambilan keputusan dari hasil uji korelasi tersebut adalah sebagai berikut,

1. Apabila skor $\text{sig.}(2\text{-tailed}) >$ probabilitas 0.05, maka tidak terdapat korelasi yang signifikan disetiap variabel ujinya atau H_0 diterima,
2. Apabila skor $\text{sig.}(2\text{-tailed}) <$ probabilitas 0.05, maka terdapat korelasi yang signifikan disetiap variabel ujinya atau H_1 diterima.

Setelah dilakukan uji korelasi R-Pearson, terbukti bahwasanya terdapat keterkaitan hubungan yang berarti antara variabel X dengan Y. Korelasi yang dihasilkan menunjukkan bilangan yang negatif sehingga hubungan disetiap variabel ujinya berlangsung secara berkebalikan yaitu tingginya skor disatu variabel diikuti oleh skor rendah divariabel selanjutnya. Hal tersebut sesuai dengan yang penulis perkirakan sebelumnya bahwa, seharusnya dengan telah diselenggarakannya uji emisi tersebut setidaknya kualitas pencemaran udara di ibu kota dapat terminimalisir walaupun tidak harus besar angkanya. Berikut adalah buktinya.

Correlations

		X	Y
X	Pearson Correlation	1	-.905**
	Sig. (2-tailed)		.001
	N	9	9
Y	Pearson Correlation	-.905**	1
	Sig. (2-tailed)	.001	
	N	9	9

** . Correlation is significant at the 0.01 level (2-tailed).

Gambar 4. Hasil Uji Korelasi

Dapat kita lihat darinya bahwa hasil sig.(2-tailed) menunjukkan angka berada di bawah 0.05, artinya ada korelasi yang berarti disetiap variabel ujinya dengan sifat ikatan yang berkebalikan (negatif). Namun, dikarenakan hasil yang ditampilkan hanya Sebatas informasi mengenai keterkaitan korelasi datanya saja dengan tidak dijelaskannya skor variabel mana yang menurun dan skor variabel mana yang meningkat. Untuk itu, diperlukan kembali metode pengujian regresi linear untuk mengetahui sifat keberadaan hubungan data serta seberapa besar prediksi Y yang diterangkan oleh X.

3.6. Uji Regresi Linear

Tahap berikutnya merupakan tahap analisis arah hubungan keberadaan data serta seberapa besar efektifitas variabel X sebagai variabel bebas terhadap Y sebagai variabel terikat. Untuk mengetahui prediksi Y yang diterangkan oleh X, maka lebih dulu kita harus mengetahui ketentuan pengambilan kesimpulan yang berlaku dipengujian tersebut,

1. Jika nilai skor sig.(2-tailed) > probabilitas 0.05, maka tidak terdapat pengaruh skor uji emisi (X) terhadap kualitas udara di ibu kota (Y) atau Ho diterima,
2. Jika nilai skor sig.(2-tailed) < probabilitas 0.05, maka terdapat pengaruh skor uji emisi (X) terhadap kualitas udara di ibu kota (Y) atau H1 diterima.

Setelah itu, berikut adalah hasil perhitungannya

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	91.514	4.163		21.992	.000
	UJI_EMISI_21	-6.273	1.112	-.905	-5.639	.001

a. Dependent Variable: ISPU_21

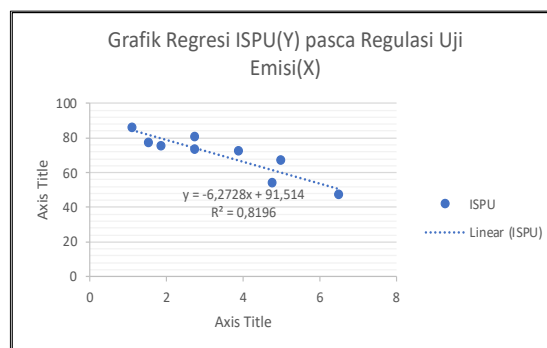
Gambar 5. Hasil Uji Regresi

Berdasarkan *output* di atas diketahui nilai signifikansi (sig.) sebesar 0.01 lebih kecil dari

probabilitas 0.05 [sig.(2-tailed) < 0.05], dengan demikian Ho ditolak dan H1 diterima sehingga ada pengaruh uji emisi (X) terhadap kualitas udara di ibu kota (Y). Jika dibuat persamaan regresinya maka konstanta (a) sebagai intersep ada pada tabel (constant) output di atas sedangkan koefisien arah regresi (b) ada pada tabel *Uji_Emisi. Ditunjukkan pada konstanta (a)/* (constant) bahwa disaat uji emisi/ variabel X tidak ada, maka nilai konsistensi ISPU sebesar 91.514. Sedangkan maksud dari koefisien regresi (b), yaitu suatu kondisi dimana disaat (X) mengalami penambahan sebesar 1% maka vairabel (Y) / ISPU mengalami penurunan sebesar (-6.273). Dalam artian, ketika jumlah kendaraan yang lolos uji emisi meningkat per 1% nya telah berkontribusi kepada menurunnya skor ISPU di Jakarta atau mampu mereduksi pencemaran udara disana. Sehingga susunan persamaan yang dihasilkan adalah sebagai berikut,

$$Y = 91.514 - 6.273x \quad (2)$$

Adapun jika diilustrasikan ke dalam bentuk grafik adalah sebagai berikut,



Gambar 6. Grafik Regresi ISPU (Y) Pasca Regulasi Uji Emisi (X)

Berdasarkan gambar 5, terbukti bahwa regulasi uji emisi di DKI Jakarta berdampak efektif, khususnya dalam mereduksi persoalan pencemaran lingkungan di ibu kota. dapat dilihat pada grafik diatas bahwa bentuk regresi ISPU berlangsung menurun, artinya kualitas udara di ibu kota berlangsung membaik pasca diberlakukannya regulasi uji emisi. Adapun, untuk mengetahui seberapa besar pengaruh uji emisi terhadap terjadinya penurunan ISPU disana adalah dengan mengkonversikan hasil dari R² sebesar 0.8196 kedalam bentuk persentase. Dengan demikian, kontribusi uji emisi terhadap menurunnya ISPU di Jakarta adalah sebesar 82% dikontribusikan olehnya dan 18% dikontribusikan oleh variabel lain yang tak terdefinisi.

4. KESIMPULAN DAN SARAN

4.1. Kesimpulan

Peranan regulasi pengendalian lingkungan menjadi sangat penting untuk segera dihadirkan terlebih dalam merespon perkembangan aktivitas perkotaan yang kian hari semakin meningkat pesat. Dampak eksternalitas terhadap lingkungannya adalah salah satunya berupa pencemaran kualitas udara di perkotaan akibat tingginya kadar gas buang emisi kendaraan dari aktivitas mobilisasi masyarakat kotanya. Oleh karenanya, salah satu regulasi yang menurut penulis cukup tepat dan inovatif adalah dengan menghadirkan upaya penyediaan regulasi uji emisi seperti yang telah dilakukan di Jakarta saat ini. Setelah penulis identifikasi mengenai efektivitas regulasi tersebut dalam mereduksi persoalan lingkungan di ibu kota, ternyata terbukti berhasil dan efektif dalam menangkal persoalan lingkungan 'udara' disana. dimana seiring dengan meningkatnya jumlah kendaraan yang lolos uji emisi di Jakarta telah berhasil menurunkan skor ISPU di ibu kota sebesar 82% diperankan oleh upaya uji emisi dan 18% sisanya dipengaruhi oleh variabel lainnya.

4.2. Saran

Saran untuk peneliti di masa yang akan datang, diketahui bahwa regulasi tersebut terbukti efektif dalam mereduksi persoalan lingkungan di DKI Jakarta, namun diharapkan model regulasi dapat diperhatikan mengikuti isu perkembangan zaman, dikarenakan aspek perkembangan zaman akan terus maju sehingga dinamika-dinamika eksternalitas atasnya tentu akan memiliki pola pengaruh yang berbeda pula dikedepannya. Salah satu contoh yang dapat diupayakan adalah program *redevelopment* ruang perkotaan melalui efisiensi intensitas lahan yang tak terpakai di perkotaan untuk dimanfaatkan kepada upaya penyediaan taman RTH, penyediaan fasilitas SPKLU bagi kendaraan listrik dalam memacu perkembangan moda transportasi berkelanjutan di perkotaan, serta upaya dalam menciptakan sistem *compact city* secara terpadu dan multifungsi diberbagai aspek ruang kotanya terkhusus bagi pengupayaan sistem kota terintegrasi disetiap multi-moda transportasi yang ada di kota tersebut layaknya konsep *Transit Oriented Development* (TOD).

DAFTAR PUSTAKA

- Abbas, S. (2012). Manajemen Transportasi. Cetakan Pertama. Edisi Kedua. In *Materials Science and Engineering A* (Vol. 27). Jakarta: Ghalia Indonesia.
- Azizah, F. N. (2019). *Modul Statistika Industri 2*. Karawang: Fakultas Teknik Universitas

- Singaperbangsa Karawang.
- Bertan, C. V., Dundu, A. K. T., & Mandagi, R. J. . (2016). Pengaruh Pendayagunaan Sumber Daya Manusia (Tenaga Kerja) Terhadap Hasil Pekerjaan (STUDI KASUS PERUMAHAN TAMAN MAPANGET RAYA(TAMARA)). *Jurnal Sipil Statik*, 4(1), 13–20.
- BPS Jakarta. (2019). *Pembagian Daerah Administrasi Menurut Kabupaten/Kota Administrasi di Provinsi DKI Jakarta*. Jakarta.
- Lestari, R. A. (2014). Pengaruh Kepemimpinan Partisipatif Dan Komitmen Organisasi Terhadap Efektifitas Implementasi Rencana Strategik Pada Madrasah Aliyah di Kabupaten Sukabumi. *Universitas Pendidikan Indonesia*, 1–25.
- MENLH. (2020). *Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia No 14 Tahun 2020 tentang Indeks Standar Pencemaran Udara*. 1–16.
- Oktavia, P. (2021). *Metode Analisis Perencanaan 1 Tentang Analisis Korelasi* (Vol. 13). Cikarang: Institut Teknologi Sains Bandung.
- Ramdani. (2019). *Polusi Udara Jakarta Terburuk di Dunia*. Jakarta.
- Rufaidah, V., Erwina, W., & Yanto, A. (2019). Hubungan Kualitas Web Kandaga Terhadap Kebutuhan Informasi Pengguna. *JUPI (Jurnal Ilmu Perpustakaan Dan Informasi)*, 4(1), 31. <https://doi.org/10.30829/jipi.v4i1.2917>
- Shaputra, R. D., & Hidayat, S. (2021). Implementasi regresi linear untuk prediksi penjualan pada aplikasi point of sales restoran. *Automata*. Retrieved from <https://103.220.113.195/AUTOMATA/article/view/17355>
- Sianipar, A. B. (2017). Optimalisasi Fungsi Papan Indeks Standar Pencemaran Udara (Ispu) Oleh Dinas Lingkungan Hidup Dan Kebersihan Kota Pekanbaru. *Ekp*, 13(3), 1576–1580.
- Sugianto, & Kurniawan, M. A. (2020). Tingkat Ketertarikan Masyarakat terhadap Transportasi Online, Angkutan Pribadi dan Angkutan Umum Berdasarkan Persepsi. In *Jurnal Teknologi Transportasi ...* (Vol. 1). Retrieved from <https://jurnal.poltradabali.ac.id/jttl/article/view/11>

Estimasi Interval Kredibel Distribusi Normal Terpotong Kiri pada Data Waktu Sembuh Pasien Covid-19

Putri Fardha Asa Oktavia Hans⁽¹⁾, Ardi Kurniawan⁽²⁾, Sediono⁽³⁾, Elly Ana⁽⁴⁾,
M. Fariz Fadillah Mardianto⁽⁵⁾

Universitas Airlangga

Kampus C Mulyorejo Telp./Fax. (031) 5915551 Surabaya 60115

e-mail: putri.fardha.asa-2018@fst.unair.ac.id, ardi-k@fst.unair.ac.id, sediono101@gmail.com,
elly-a@fst.unair.ac.id, dan m.fariz.fadillah.m@fst.unair.ac.id.

ABSTRAK

Pandemi Covid-19 dinyatakan sebagai *Public Health Emergency of International Concern* yaitu wabah global yang berisiko mengancam kesehatan masyarakat seluruh dunia. Tujuan dari penelitian ini adalah untuk mengestimasi interval kredibel dari distribusi normal terpotong kiri. Estimasi interval kredibel adalah estimasi interval dengan menggunakan metode Bayesian. Hasil estimasi interval kredibel yang didapatkan memiliki bentuk implisit yaitu bentuk persamaan yang tidak mudah untuk diselesaikan sehingga diselesaikan dengan pendekatan numerik yaitu integrasi komputasi dengan menggunakan bantuan software *Mathematica*. Hasil estimasi yang diperoleh diterapkan pada waktu sembuh pasien Covid-19 dari Puskesmas Jemursari Surabaya pada rentang bulan Desember 2020 sampai Februari 2021. Melalui pemotongan kiri, proses estimasi parameter hanya menggunakan data yang lebih dari 10 hari, sehingga didapat 37 data dari total 45 data. Kesimpulan yang diperoleh adalah waktu sembuh pasien Covid-19 di Puskesmas Jemursari mempunyai rata-rata antara 10.583 hari sampai dengan 11.087 hari dan variansi antara 1.706 hari sampai dengan 1.772 hari.

Kata kunci : *distribusi normal, metode bayesian, prior jeffrey's, terpotong kiri.*

ABSTRACT

The Covid-19 pandemic has been declared a *Public Health Emergency of International Concern*, which is a global pandemic that threatens public health throughout the world. The purpose of this study was to estimate the credible interval of the left truncated normal distribution. Credible interval estimation is interval estimation using Bayesian method. The results of the estimated credible intervals obtained have an implicit form or an equation that is not easy to solve so that they are solved by using a numerical computational integral approach using *Mathematica* software. The results of this study were applied to the recovery time of Covid-19 patients from the Jemursari Health Center Surabaya in the range of December 2020 to February 2021. Through left cutting, the parameter estimation process only uses data that is more than 10 days, so that 37 data is obtained from a total of 45 data. The conclusion is that the recovery time of Covid-19 patients at the Jemursari Health Center has an average of 10.583 days to 11.087 days and a variance between 1.706 days to 1.772 days.

Keywords : *normal distribution, bayesian method, jeffrey's prior, left truncated.*

1. PENDAHULUAN

Coronavirus Disease 2019 (Covid-19) menjadi wabah global yang menghampiri seluruh aspek kehidupan, khususnya aspek kesehatan manusia. Pemerintah di seluruh negara sedang berjuang untuk keluar dari wabah Covid-19 (Agustino, 2020). Ketua Perhimpunan Rumah Sakit Indonesia Jawa Timur, dr. Dodo Anondo menyampaikan bahwa terjadinya lonjakan kasus pada gelombang II Covid-19 pada bulan Januari

2021 membuat beberapa rumah sakit overload sehingga diberlakukan sistem tutup sementara. Sistem layanan rumah sakit yang overload bisa membahayakan keselamatan masyarakat yang tidak tertangani oleh tenaga medis. Beberapa upaya penanganan yang dilakukan pemerintah adalah dengan membuka rumah sakit lapangan, menambah tempat isolasi, hingga mempercepat vaksinasi. Selain itu, penting adanya estimasi terkait waktu

sembuh pasien Covid-19 untuk mencegah terjadinya Kembali overload kapasitas rumah sakit.

Inferensi statistik digunakan untuk mengambil suatu keputusan pada suatu populasi. Pada inferensi statistik, salah satu hal terpenting adalah melakukan estimasi. Estimasi parameter yaitu penaksiran parameter populasi dengan menggunakan sampel (Walpole, 2017). Secara teoritis terdapat dua metode dalam melakukan estimasi parameter yaitu metode klasik dan *Bayesian*. Metode *Bayesian* selain menggunakan informasi data sampel juga menggunakan informasi data sebelumnya atau yang disebut dengan distribusi *prior* (Prahutama, 2012). Distribusi *prior* yang tidak diketahui informasi sebelumnya disebut dengan *non-informative prior*. Salah satu metode yang digunakan untuk menentukan distribusi *non-informative prior* adalah metode *Jeffrey's*.

Estimasi parameter dibedakan menjadi dua yaitu estimasi titik dan estimasi interval. Estimasi interval merupakan pendugaan rentang nilai untuk parameter populasi. Estimasi interval dengan menggunakan metode *Bayesian* disebut dengan interval kredibel. Dalam statistik *Bayesian*, interval kredibel merupakan interval di dalam domain dari distribusi probabilitas *posterior* untuk penaksiran interval (Mutiarani, 2012). Interval kredibel disusun dengan menemukan kuantil $\alpha/2$ dan $1 - \alpha/2$ dengan tingkat signifikansi α .

Informasi data sampel mempunyai sebaran distribusi diskrit atau kontinu. Distribusi kontinu yang penting dan banyak digunakan dalam statistika adalah distribusi normal. Salah satu pemanfaatan distribusi normal pada bidang kedokteran adalah sebagai distribusi untuk analisis uji hidup. Analisis uji hidup adalah daya tahan pasien terhadap suatu jenis penyakit (Fernandes, 2016). Jika semua individu diuji sampai terjadi kegagalan disebut dengan sampel lengkap, maka ketika dilakukan pembatasan individu terhadap suatu nilai tertentu disebut dengan sampel terpotong. Distribusi akibat adanya pemotongan lebih besar dari suatu nilai tertentu disebut dengan distribusi terpotong kiri (Cha, 2015).

Penelitian terkait estimasi interval kepercayaan telah banyak dilakukan. Salah satunya dilakukan oleh Hazhiah (2012) dengan judul Estimasi Parameter Distribusi Weibull Dua Parameter Menggunakan Metode Bayes. Hasil dari penelitian ini adalah diketahuinya bentuk estimasi kredibel distribusi Weibull dua parameter pada sampel lengkap dan tersensor tipe II.

Dalam penelitian kali ini dilakukan estimasi interval kredibel dari distribusi normal terpotong kiri. Metode yang digunakan yaitu *Bayesian* dengan

prior Jeffrey's distribusi normal sebagai informasi *prior* dari parameter. Hasil estimasi diterapkan pada data waktu sembuh pasien Covid-19 di Puskesmas Jemursari. Dimisalkan X adalah variabel acak berdistribusi normal yang menyatakan waktu sembuh pasien Covid-19. Selanjutnya akan dilakukan pemotongan kiri yaitu $X > a$ dengan $a = 10$, merujuk pada himbauan WHO yaitu isolasi minimal 10 hari pertama sejak didiagnosis positif Covid-19. Estimasi waktu sembuh pasien Covid-19 diharapkan dapat menjadi salah satu upaya membantu pihak penyedia fasilitas kesehatan untuk menyiapkan sistem layanan perawatan pasien Covid-19.

2. METODE PENELITIAN

2.1 Sumber Data

Data dalam penelitian ini adalah data sekunder yang diperoleh dari Puskesmas Jemursari Kota Surabaya yaitu waktu sembuh pasien Covid-19 dalam rentang waktu bulan Desember 2020 sampai dengan Februari 2021. Data diukur sejak pasien dinyatakan positif hingga telah negatif dan didapatkan total 45 pasien. Lalu dilakukan pemotongan kiri yaitu pada $a = 10$, yang artinya data yang digunakan untuk proses estimasi adalah pasien yang memiliki rentang waktu sembuh lebih dari 10 hari.

2.2 Metode Analisis

Secara lengkap, prosedur analisis tersebut dapat dituliskan sebagai berikut.

1. Menentukan *Probability Density Function* (PDF) dari distribusi normal terpotong kiri berdasarkan persamaan berikut:

$$f(x|X > a; \theta, \sigma^2) = \frac{f(x)}{\text{Prob}(X > a)} \quad (1)$$

2. Menentukan bentuk fungsi *likelihood* distribusi normal terpotong kiri dengan parameter rata-rata θ dan variansi σ^2 berdasarkan persamaan berikut:

$$L(\theta, \sigma^2 | \underline{x}) = \prod_{i=1}^n f(x_i | X_i > a; \theta, \sigma^2) \quad (2)$$

3. Menentukan jenis *prior* yang digunakan yaitu *prior Jeffrey's* dari distribusi normal dengan θ dan σ^2 bersifat independent sebagai berikut:

$$p(\theta, \sigma^2) \propto p(\theta) p(\sigma^2) \propto \frac{1}{\sigma^2} \quad (3)$$

4. Menentukan distribusi *posterior* $g(\theta, \sigma^2 | \underline{x})$ berdasarkan persamaan berikut:

$$g(\theta, \sigma^2 | \underline{x}) = \frac{L(\theta, \sigma^2 | \underline{x}) p(\theta, \sigma^2)}{\int L(\theta, \sigma^2 | \underline{x}) p(\theta, \sigma^2)} \quad (4)$$

5. Menentukan distribusi marginal untuk parameter θ dan σ^2 yakni $p(\theta | \underline{x})$ dan $p(\sigma^2 | \underline{x})$ sebagai berikut:

- a. Distribusi marginal $p(\theta | \underline{x})$

$$p(\theta | \underline{x}) = \int_0^\infty g(\theta, \sigma^2 | \underline{x}) d\sigma^2 \quad (5)$$

- b. Distribusi marginal $p(\sigma^2 | \underline{x})$

$$p(\sigma^2 | \underline{x}) = \int_{-\infty}^\infty g(\theta, \sigma^2 | \underline{x}) d\theta \quad (6)$$

6. Menentukan estimasi interval kredibel untuk parameter θ berdasarkan persamaan berikut dengan tingkat signifikansi $\alpha = 0.05$:

- a. Batas bawah untuk θ atau θ_L

$$R(\theta_L) = \int_0^{\theta_L} p(\theta | \underline{x}) d\theta = \frac{\alpha}{2} \quad (7)$$

- b. Batas atas untuk θ atau θ_U

$$R(\theta_U) = \int_0^{\theta_U} p(\theta | \underline{x}) d\theta = 1 - \frac{\alpha}{2} \quad (8)$$

7. Menentukan estimasi interval kredibel untuk parameter σ^2 berdasarkan persamaan berikut dengan tingkat signifikansi $\alpha = 0.05$:

- a. Batas bawah untuk σ^2 atau σ^2_L

$$R(\sigma^2_L) = \int_0^{\sigma^2_L} p(\sigma^2 | \underline{x}) d\sigma^2 = \frac{\alpha}{2} \quad (9)$$

- b. Batas atas untuk σ^2 atau σ^2_U

$$R(\sigma^2_U) = \int_0^{\sigma^2_U} p(\sigma^2 | \underline{x}) d\sigma^2 = 1 - \frac{\alpha}{2} \quad (10)$$

8. Menerapkan hasil estimasi distribusi normal terpotong kiri pada data dengan langkah-langkah sebagai berikut:

- a. Membuat program untuk menentukan estimasi distribusi normal terpotong kiri yaitu pada persamaan (4) sampai dengan (10) menggunakan *software Mathematica*.

- b. Menguji kenormalan data waktu sembuh pasien Covid-19 menggunakan *software SPSS*.

3. HASIL DAN PEMBAHASAN

Diasumsikan X merupakan variabel *random* berdistribusi normal $X \sim N(\theta, \sigma^2)$. Bentuk PDF dari variabel *random* X adalah sebagai berikut:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}$$

atau

$$f(x) = \frac{1}{\sigma} \phi(y) \quad (11)$$

dengan $\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}$.

Selanjutnya, berikut merupakan bentuk *Cumulative Density Function* (CDF) variabel *random* X:

$$F(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2} dx$$

misalkan $z = \frac{x-\theta}{\sigma}$, maka:

$$F(x) = \int_{-\infty}^{\frac{x-\theta}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Bentuk CDF atau $F(x)$ dari distribusi normal adalah sebagai berikut:

$$F(x) = \int_{-\infty}^{\frac{x-\theta}{\sigma}} \phi(z) dz$$

atau

$$F(x) = \Phi\left(\frac{x-\theta}{\sigma}\right) \quad (12)$$

Tahap pertama untuk menentukan PDF distribusi normal terpotong kiri dengan terlebih dahulu menentukan bentuk $Prob(X > a)$ berdasarkan persamaan (12) adalah:

$$Prob(X > a) = 1 - \Phi\left(\frac{a-\theta}{\sigma}\right) \quad (13)$$

Berdasarkan persamaan (1) dan (13), bentuk PDF distribusi normal terpotong kiri dapat dituliskan sebagai berikut:

$$f(x|X > a; \theta, \sigma^2) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}}{1 - \Phi\left(\frac{a-\theta}{\sigma}\right)} \quad (14)$$

Dengan demikian bentuk fungsi *likelihood* sesuai persamaan (2) dan (14) adalah:

$$L(\theta, \sigma^2 | \underline{x}) = \frac{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} \quad (15)$$

Diasumsikan bahwa distribusi *prior* yang digunakan adalah *prior Jeffrey's* dari distribusi normal dengan θ dan σ^2 bersifat independent. Distribusi *prior* dari θ dan σ^2 masing-masing adalah $p(\theta) \propto c(\text{konstan})$ dan $p(\sigma^2) \propto \frac{1}{\sigma^2}$ (Subanar, 2019). Sesuai dengan persamaan (3), bentuk distribusi *prior Jeffrey's* dari distribusi normal adalah:

$$p(\theta, \sigma^2) \propto p(\theta) p(\sigma^2) \propto 1 \times \left(\frac{1}{\sigma^2}\right) \propto \frac{1}{\sigma^2} \quad (16)$$

Distribusi *posterior* berdasarkan persamaan (4) adalah pembagian antara fungsi *likelihood* (15) dengan distribusi *prior* (16), berikut merupakan bentuk distribusi *posterior* untuk distribusi normal terpotong kiri:

$$\begin{aligned} g(\theta, \sigma^2 | \underline{x}) &= \frac{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} \left(\frac{1}{\sigma^2}\right) \\ &= \frac{\int_{\sigma^2=0}^{\infty} \int_{\theta=-\infty}^{\infty} \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} \left(\frac{1}{\sigma^2}\right) d\theta d\sigma^2} \\ &= \frac{\left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} \\ &= \frac{\int_{\sigma^2=0}^{\infty} \int_{\theta=-\infty}^{\infty} \left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \int_{-\infty}^{\frac{a-\theta}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right)^n} d\theta d\sigma^2 \end{aligned}$$

Misalkan $w =$

$$\int_{\sigma^2=0}^{\infty} \int_{\theta=-\infty}^{\infty} \left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \left(1 - \int_{-\infty}^{\frac{a-\theta}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz\right)^n d\theta d\sigma^2,$$

maka distribusi *posterior* dapat dituliskan sebagai berikut:

$$g(\theta, \sigma^2 | \underline{x}) = \frac{1}{w} \left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n \quad (17)$$

Interval kepercayaan untuk parameter θ dengan metode estimasi *Bayesian* ditentukan dengan terlebih dahulu menentukan distribusi marginal dari parameter θ . Berdasarkan persamaan (5), bentuk $p(\theta | \underline{x})$ adalah:

$$p(\theta | \underline{x}) = \frac{1}{w} \int_0^{\infty} \left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n d\sigma^2 \quad (18)$$

Dengan demikian berdasarkan persamaan (7), dapat ditentukan nilai θ_L yaitu:

$$\int_{-\infty}^{\theta_L} \int_0^{\infty} \frac{\left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} d\sigma^2 d\theta = 0.025 \quad (19)$$

Sedangkan berdasarkan persamaan (8), dapat ditentukan nilai θ_U yaitu:

$$\int_{-\infty}^{\theta_U} \int_0^{\infty} \frac{\left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} d\sigma^2 d\theta = 0.975 \quad (20)$$

Interval kepercayaan untuk parameter σ^2 dengan metode estimasi *Bayesian* ditentukan dengan terlebih dahulu menentukan distribusi marginal dari parameter σ^2 . Berdasarkan persamaan (6), bentuk $p(\sigma^2 | \underline{x})$ adalah:

$$p(\sigma^2 | \underline{x}) = \frac{1}{w} \int_0^{\infty} \left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2} \left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n d\theta \quad (21)$$

Dengan demikian berdasarkan persamaan (9), dapat ditentukan nilai dari σ^2_L yaitu:

$$\int_0^{\sigma^2_L} \int_{-\infty}^{\infty} \frac{\left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} d\theta d\sigma^2 = 0.025 \quad (22)$$

Sedangkan berdasarkan persamaan (10), dapat ditentukan nilai dari σ^2_U yaitu:

$$\int_0^{\sigma^2_U} \int_{-\infty}^{\infty} \frac{\left(\frac{1}{\sigma^{n+2}}\right) e^{-\frac{1}{2\sigma^2}\sum(X_i-\theta)^2}}{\left(1 - \Phi\left(\frac{a-\theta}{\sigma}\right)\right)^n} d\theta d\sigma^2 = 0.975 \quad (23)$$

Persamaan (19), (20), (22), dan (23) masing-masing adalah estimasi interval kredibel dari parameter θ dan σ^2 . Bentuk estimasi yang didapatkan merupakan bentuk yang implisit, maka akan dihitung dengan pendekatan numerik menggunakan *software Mathematica*.

Selanjutnya, hasil estimasi interval akan diterapkan pada data waktu sembuh pasien Covid-19 yang berasal dari Puskesmas Jemursari Surabaya dalam rentang bulan Desember 2020 sampai Februari 2021. Data yang digunakan setelah dilakukan pemotongan kiri adalah sejumlah 37 data dari total keseluruhan 45 data. Pemotongan dilakukan pada titik $a = 10$, merujuk pada peraturan WHO bahwa setiap pasien Covid-19 wajib melakukan 10 hari masa isolasi awal.

Digunakan uji *Saphiro Wilk* untuk memastikan apakah data waktu sembuh pasien Covid-19 berdistribusi normal. Berikut merupakan hasil uji

Shapiro Wilk dengan tingkat signifikansi 5% pada *software SPSS*.

Tabel 1. Output uji *Shapiro Wilk*

	df	<i>P-value</i>
Data lengkap	45	0.756
Data terpotong kiri	37	0.087

Berdasarkan Tabel 1. diperoleh nilai *P-value* dari data lengkap dan data terpotong kiri waktu sembuh pasien Covid-19 lebih besar dari tingkat signifikan 0,05 sehingga dapat disimpulkan bahwa data berdistribusi normal. Selanjutnya, dengan menggunakan *software Mathematica* didapatkan interval kredibel dari parameter θ atau rata-rata waktu sembuh pasien Covid-19 terpotong kiri antara 10.583 hari sampai dengan 11.087 hari. Sedangkan, interval kredibel dari parameter σ^2 atau variansi waktu sembuh pasien Covid-19 terpotong kiri antara 1.706 hari sampai dengan 1.772 hari.

4. KESIMPULAN DAN SARAN

Berdasarkan pembahasan didapatkan estimasi interval kredibel dari distribusi normal terpotong kiri yang berbentuk implisit, sehingga diselesaikan dengan pendekatan numerik. Hasil estimasi interval kredibel distribusi normal terpotong kiri diterapkan pada data waktu sembuh pasien Covid-19 dengan total data yang digunakan yaitu 37 data. Dengan tingkat signifikansi $\alpha = 0.05$, rata-rata waktu sembuh pasien Covid-19 antara 10.583 hari sampai dengan 11.087 hari. Sedangkan, variansi waktu sembuh pasien Covid-19 antara 1.706 hari sampai dengan 1.772 hari. Estimasi interval kredibel yang diperoleh memiliki rentang yang pendek, sehingga dapat disimpulkan bahwa hasil estimasi interval yang didapatkan sudah cukup baik.

Penelitian ini hanya terbatas pada pembahasan estimasi interval kredibel dengan menggunakan *prior Jeffrey's*, sehingga saran untuk penelitian selanjutnya adalah dapat dilakukan pengembangan dengan menentukan estimasi titik, reliabilitas atau dengan menggunakan jenis distribusi prior lainnya.

DAFTAR PUSTAKA

- Agustino, L. (2020). Analisis Kebijakan Penanganan Wabah Covid-19: Pengalaman Indonesia. *Jurnal Borneo Administrator*, 16(2), 253-270.
- Cha, J. (2015). *Re-Estabilishing the Theoretical Foundations of a Truncated Normal Distribution: Standardization Statistical Inference, and Convolution*. Clomson University.
- Fernandes, A. A. R., Solimun. (2016). *Pemodelan Statistika pada Analisis Reliabilitas dan Survival*. Universitas Brawijaya Press.
- Hazhiah, I. T., Sugito, Rahmawati, R. (2012). Estimasi Parameter Distribusi Weibull Dua Parameter Menggunakan Metode Bayes. *Jurnal Gaussian*, 1(1), 103-112.
- Mutiarani, V. S. (2012). Estimasi Parameter Dan Interval Kredibel Dengan Model Regresi Linier Berganda Bayesian. *Seminar Nasional Pendidikan Matematika Ahmad Dahlan*. Yogyakarta.
- Prahatama, A. S. (2012). Inferensi Statistik dari Distribusi Normal dengan Metode Bayes untuk Non-Informatif Prior. *Media Statistika*, 5(2), 95-104.
- Subanar. (2019). *Inferensi Bayesian Dengan R*. Yogyakarta: Gadjah Mada University Press.
- Walpole, R. E. (2017). *Probability & Statistics for Engineers & Scientist*. United States: Ninth Edition, Pearson Education.

Hotel Recommendation System with Content-Based Filtering Approach (Case Study: Hotel in Yogyakarta on Nusatrip Website)

Cheryl Ayu Melyani⁽¹⁾, Ayundyah Kesumawati⁽²⁾, Raden Bagus Fajriya Hakim⁽³⁾, Arum Handini Primandari⁽⁴⁾

Department of Statistics, Universitas Islam Indonesia
Jl.Kaliurang km 14,5, Yogyakarta 55584, Indonesia
e-mail: cherylayu25@gmail.com

ABSTRAK

Meningkatnya pandemi Covid-19 membuat aktivitas masyarakat menjadi terhambat seringkali memunculkan stress jika harus berada dirumah secara terus menerus. Hal ini menyebabkan meningkatnya tren staycation atau kegiatan berlibur di kota sendiri dengan menyewa sebuah hotel. Teknologi penyewaan hotel mulai dialihkan dengan adanya OTA (Online Travel Agent). Adanya berbagai macam hotel dengan berbagai macam fasilitas yang disuguhkan membuat masyarakat sering merasa kebingungan dalam memilih hotel yang akan ditematinya. Untuk membantu mengatasi hal tersebut, peneliti membuat sebuah sistem rekomendasi untuk membantu calon penghuni hotel dalam memilih hotel sesuai dengan pilihannya. Selain itu juga dapat membantu perusahaan dalam meningkatkan pemesanan kamar hotel melalui website-nya. Dalam penelitian ini, peneliti membangun sebuah sistem rekomendasi hotel di Yogyakarta pada salah satu OTA di Indonesia menggunakan Content-Based Filtering Methods, pembobotan data teks menggunakan Term Frequency-Invers Document Frequency (TF-IDF) Methods, dan mengukur kemiripan dokumen menggunakan Cosine Similarity Methods. Data yang digunakan adalah deskripsi dari masing-masing hotel. Berdasarkan hasil rekomendasi hotel Good Karma Yogyakarta sebagai contoh pengujian, didapatkan 10 hotel yang mirip yaitu Happy Buddha Yogyakarta – Hostel, Nextdoor Homestay, Hotel Puspita, OYO 426 Hotel Gading Resto, Omah Jegog Homestay, Prawirotaman Homestay, RedDoorz near Prawirotaman, Ayodhya Garden Hostel Yogyakarta by HOM, Bringin House Yogyakarta, dan House 24 Yogyakarta dengan nilai cosine similarity secara berturut-turut sebesar 0.956666513, 0.946570717, 0.917459394, 0.912534877, 0.886439718, 0.88221982, 0.881775275, 0.875845794, 0.872030219, dan 0.871514859.

Kata Kunci : *Staycation, Hotel, Recommendation System, Content-Based Filtering*

ABSTRACT

The increasing of Covid-19 pandemic has hampered people's activities, often causing stress if they are only stay at home continuously. This has led to an increasing trend of staycations or holiday activities in the city itself by renting a hotel. Hotel rental technology has begun to be transferred with the existence of OTA (Online Travel Agent). The existence of various kinds of hotels with various kinds of facilities that makes people feel confused in choosing which hotel to occupy. To help overcome this, the researchers create a recommendation system to help prospective hotel residents choose the hotel according to their choice. In addition, it can also assist companies in increasing hotel room reservations through its website. In this study, researchers build a hotel recommendation system in Yogyakarta at one of the OTAs in Indonesia using Content-Based Filtering Methods, weighting text data using Term Frequency-Inverse Document Frequency (TF-IDF) Methods and measuring document similarity using Cosine Similarity Methods. The data used is a description of each hotel. Based on the results of the Good Karma Yogyakarta hotel recommendations as a test example, 10 similar hotels were obtained, namely Happy Buddha Yogyakarta – Hostel, Nextdoor Homestay, Hotel Puspita, OYO 426 Hotel Gading Resto, Omah Jegog Homestay, Prawirotaman Homestay, RedDoorz near Prawirotaman, Ayodhya Garden Hostel Yogyakarta by HOM, Bringin House Yogyakarta, and House 24 Yogyakarta with cosine similarity values 0.956666513, 0.946570717, 0.917459394, 0.912534877, 0.886439718, 0.88221982, 0.881775275, 0.875845794, 0.872030219, and 0.871514859.

Keywords : *Staycation, Hotel, Recommendation System, Content-Based Filtering*

1. INTRODUCTION

The COVID-19 pandemic, which has lasted for approximately 2 years has an impact on various sectors. Activities that people do are hampered. On the one hand, people do not want to be infected with the Corona Virus and decide to carry out all activities in their respective homes. On the other hand, they realized that if they stay at home all time, it will create stress. Based on Google Trends data, the graph of the staycation trend tends to rise from the beginning of 2020 until now. According to HowStuffWorks, a staycation is a combination of taking a vacation and staying at home (Layton, 2009). A survey conducted by Wego.co.id regarding the staycation trend, shows that Yogyakarta is one of the most popular destinations as a destination city for staycation. This is shown by the number of hotel searches in May-August 2020 on their platform (Ahsan, 2020).

In this digital era, the development of information technology has grown rapidly where various innovations in offering a product or service have emerged. One of the innovations that has evolved from conventional to modern is in the tourism sector (Sandi Wachyuni & Wiweka, 2020). The existence of an Online Travel Agent (OTA) is able to replace hotel booking activities where previously book a hotel directly, now it can be done online using website-based OTA (Hendriyati, 2019). There are various kinds of star and non-star hotels in Yogyakarta, people who want to do a staycation will be confused in choosing a hotel according to the criteria and facilities of interest. OTA provides a variety of hotel options that can be tailored to the needs and desires of the user. All information regarding the location, facilities, and advantages of the hotel which includes explanations about rooms, restaurants, swimming pools, parking areas, and others are explained in the description column on an OTA website. In this study, researchers create a hotel recommendation system in Yogyakarta at one of the OTAs in Indonesia which is Nusatrip.com using Content-Based Filtering Methods which is recommendation method based on content or features of the items, weighting text data using Term Frequency-Inverse Document Frequency (TF-IDF) , and measuring document similarity using Cosine Similarity to make it easier for someone in choosing a hotel. One of studies about recommendation system is "Research Supervisor Recommendation System Based on Topic Conformity". the purpose of this research is to make it easier for students to determine their supervisor according to the research topic. This study applies word weighting using TF_IDF method

and calculating the similarity between two objects using Cosine Similarity Method. Based on the results of the analysis, the accuracy value of comparison of the recommendations from the system with the actual data is 75% .

2. METHOD AND MATERIALS

2.1. Data and Data Source

This study used secondary data that obtained from Nusatrip.com website on November 16, 2021. The variables used are name, description, and hotel star data in Yogyakarta.

2.2. Recommendation System

Recommendation system is a software and a technique that suggests an item of interest to users and intended to support its use in various decision-making processes (Larasati & Februariyanti, 2021). For example, in the decision-making process such as what items will users buy, what music will be heard, what hotels will be ordered, what news will be read, and much more. In how its works, the recommendation system will not only recommend items that are most liked by user but also provide recommendations personally to each user which will provide item recommendations according to user (Mondi et al., 2019). The recommendation system is one of the most powerful and popular information discovery tools on a website to overcome information overload and can help them to make good choices (Ricci et al., 2015).

2.3. Content-Based Filtering

The Content-Based Filtering method is a recommendation system method based on content or features of the items then compared with the items that user liked before (Raghuwanshi & Pateriya, 2019). The recommendation system using Content-Based Filtering method is only based on the item that user is looking for or likes own and not involve other users in making the recommendations. Thus, if the user changes, the technique with Content-Based Filtering is still possible to adjust the recommendation or suggestion of the appropriate item in a short time (Purnaramadhan, 2021). The advantages and disadvantages of the Content-Based Filtering are (Aamir & Bhusry, 2015):

Advantages:

1. The Content-Based Filtering method only requires the content of the item and the user profile itself for recommendations.

2. The Content-Based Filtering method can explain the features of the item on which the recommendation is based to the user.
3. New item can be recommended to users even though they don't have ratings from other users because they are based on the content of the item.

Disadvantages:

1. If the content on an item does not include complete and sufficient information to accurately distinguish it from other items, the recommendations will be less precise.
2. Serendipity constraints (unexpected events), where the system with this method will be difficult to provide recommendations or suggestions that are not unexpected items that are selected only based on content.

2.4. Text Preprocessing

The first step that must be passed before the analysis process is prepare the data to be analyzed. A text usually has a messy data structure, so it cannot be processed immediately. At this stage, the text will be cleaned of unnecessary parts until its prepared into a structured text and ready to be processed further.

1. Case Folding: the process of converting capital letter to lowercase. It its intended that the same word but containing capital letters is not detected differently from words but containing capital letters is not detected differently from words that do not contain capital letters.
2. Remove Punctuation: The stage for remove punctuation characters.
3. Remove Whitespace: Process of removing excess whitespace in text or documents.
4. Remove Stopword: Step to remove words that have no meaning or meaningless.
5. Lemmatization: A stage to find the basic form of a word (lemma) according to morphological analysis and dictionaries (Supriyati & Iqbal, 2018).

Tokenizing: Process to separating words in a text or document based on hyphehs (-) and spaces.

2.5. Term Frequency-Inverse Document Frequency (TF-IDF)

Term Frequency-Inverse Document Frequency (TF-IDF) weighting method is a method used to gibe weight to the relationship of a term in a document by combining to concept for weight calculation, namely Term Frequency (TF) which is the frequency of occurrence of words in the

document and Inverse Document Frequency (IDF) which is the inverse frequency of document containing words (Nurjannah & Fitri Astuti, 2013). The equation used to calculate TF-IDF:

$$W_{ij} = TF_{ij} \times \left(\ln \left(\frac{D}{df_j} \right) + 1 \right) \quad (1)$$

with:

- W_{ij} : Weight of term j to document i
- TF_{ij} : Frequency of occurrence of term j in document i
- IDF_j : Inverse document frequency of term j
- D : Total documents
- df_j : Total documents from term j

2.6. Cosine Similarity

Cosine similarity is a method for measuring the similarity between two n-dimensional vectors which is usually used in the information search field to compare two texts or documents (Jannach et al., 2011). If two texts or documents are increasingly similar, the value of cosine similarity will be closer to 1, whereas if the value of cosine similarity is close to 0 then the two texts or documents are increasingly dissimilar. The equation used to calculate cosine similarity:

$$Cos \propto = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \quad (2)$$

With:

- A : Vector A
- B : Vector B
- A_i : Term i in document A
- B_i : Term I in document B
- $A \cdot B$: Dot product between vector A and vector B
- $|A|$: Vector length A
- $|B|$: Vector length B
- $|A||B|$: Cross product between $|A|$ and $|B|$

3. RESULT AND DISCUSSION

3.1. Descriptive Analysis

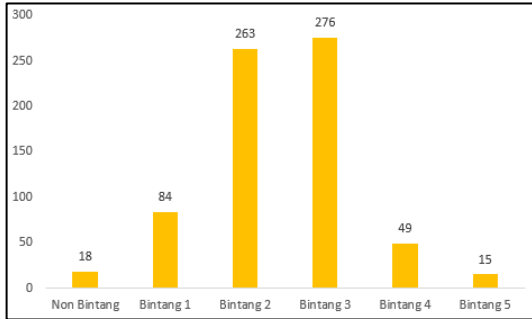


Figure 1. Number of Hotels in Yogyakarta by Star

The graph shows the number of hotels in Yogyakarta on Nusatrip.com website classified by star. From the graph, the largest distribution of hotels is in two-star and three-star hotels. One of the factors that the hotel has a high of frequency is adapting to the pattern of consumer tendencies. The two-star and three-star hotels has quite complete facilities and the price is not too expensive. This is in line with the opinion expressed by Co-Founder and Chief Marketing Office of Tiket.com in Republika.co.id article where consumers tend to choose accommodation with low process, get discounts, and promotions. This sees the tendency of consumers to choose hotel accommodation (Asrianti & Azizah, 2020).

3.2. Preprocessing Data

The first stage in the analysis process is data preprocessing. Data preprocessing carried out in this research is Case Folding, Remove Punctuation, Stopword Removal, Lemmatization, and Tokenizing text. This is an example of the data preprocessing stage using 132nd document that present in Table 1.

Table 1. Original Data

Hotel Descriptions
This property located near Malioboro Street. This smoke-free hostel features a 24-hour front desk, tour/ticket assistance, and a garden. WiFi in public areas is free. Rooms offer free WiFi and showers.
Take in the views from a terrace and a garden and make use of amenities such as complimentary wireless Internet access.
Nearby Attractions Batik Plentong 0.5 km / 0.3 mi Lana Gallery 1 km / 0.6 mi Masjid Jogokariyan 1.3 km.

This is the result of a document that has gone through the data preprocessing stage document that present in Table 2.

Table 2. Data after preprocessing

Hotel Descriptions
property located near malioboro street smoke free hostel feature 24 hour front desk tourticket assistance garden wifi public area free room offer free wifi shower view terrace garden use amenity complimentary wireless internet access... nearby attraction batik plentong 0 5 km 0 3 mi lana gallery 1 km 0 6 mi masjid jogokariyan 1 3 km

After that, the data that has been processed is separated for each word namely tokenizing.

Table 3. Word tokenizing

Hotel Descriptions
'property', 'located', 'near', 'malioboro', 'street', 'smoke', 'free', 'hostel', 'feature', '24', 'hour', 'front', 'desk', 'tourticket', 'assistance', 'garden', 'wifi', 'public', 'area', 'free', 'room', 'offer', 'free', 'wifi', 'shower', 'view', 'terrace', 'garden', 'use', 'amenity', 'complimentary', 'wireless', 'internet', 'access', '...', 'nearby', 'attraction', 'batik', 'plentong', '0', '5', 'km', '0', '3', 'mi', 'lana', 'gallery', '1', 'km', '0', '6', 'mi', 'masjid', 'jogokariyan', '1', '3', 'km', '.....

3.3. Term Weighting using TF-IDF

The first step to calculate term weighting is calculate the value of Term Frequency (TF). This is a table of term frequency:

Table 4. TF Table

Doc	laundry	nearby	...	wifi
1	1	0	...	0
2	1	0	...	0
3	0	1	...	1
4	1	2	...	2
5	0	0	...	0
⋮	⋮	⋮	...	⋮
705	0	0	...	0
df	164	461	...	456

The example used in this calculation is the word "laundry" in 4th document. A value of 0 in TF

table means that in a document, a certain word does not appear in the document, while a value of 1 means that in a document, a certain word appears 1 time in the document. The TF value for the word “laundry” in 4th document is 1, so it means that the word “laundry” appears 1 time in the 4th document. Next, look for the df value which is the document frequency value. For example, in word “laundry”, there are 164 documents containing that word. D is total document, which is 705 documents.

$$W_{ij} = TF_{ij} \times \left(\ln \left(\frac{D}{df_j} \right) + 1 \right)$$

$$W_{4,laundry} = 1 \times \left(\ln \left(\frac{705}{164} \right) + 1 \right)$$

$$W_{4,laundry} = 1 \times (1,458331 + 1) = 2,458831$$

Based on the result, the weight of the word “laundry” in 4th document is 2.458831.

3.4. Cosine Similarity

Previously, the value of word weighting has been calculated using TF-IDF, then calculate the document similarity. After getting the similarity value of a document with other documents, it will be sorted from the largest and the top 10 documents with the highest similarity value will be taken as recommendations. For example, the 132nd and 170th documents as a sample for manual calculations.

$$Cos \alpha = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

$$Cos \alpha = \frac{1638,29225}{\sqrt{1708,006} \times \sqrt{1812,564}} = 0,956666513$$

From the results of cosine similarity calculation, the 132nd document has the highest similarity value to the 170th document by 96% and it followed by nine other documents. According to the data obtained, the 132nd document is a document from Good Karma Yogyakarta-Hostel. Thus, 10 sequences hotel recommendations that are similar with Good Karma Yogyakarta Hostel from the recommendation system are shown in Table 5. The table contains the document number, hotel’s name, and cosine similarity value of the hotel document to other hotel documents. Then 10 order of hotel documents which have largest cosine similarity value are displayed in Table 5.

Table 5. Recommendation Result

No	Doc	Hotel’s Name	Cosim Value
1	170	Happy Buddha Yogyak arta - Hostel	0,9566
2	303	Nextdoor Homestay	0,9465
3	210	Hotel Puspita	0,9174

No	Doc	Hotel’s Name	Cosim Value
4	360	OYO 426 Hotel Gading Resto	0,9125
5	315	Omah Jegog Homestay	0,8864
6	422	Prawirotaman Homestay	0,8822
7	501	RedDoorz near Prawirot aman	0,8817
8	45	Ayodhya Garden Hostel Yogyakarta by HOM	0,8758
9	63	Bringin House Yogyakarta	0,8720
10	223	House 24 Yogyakarta	0,8715

When the user clicks on Good Karma Yogyakarta-Hostel, the system will provide 10 other hotel recommendations that have the highest similarity to the hotel that user has chosen. The similarity is measured based on the similarity of the description of each hotel.

3.5. Application of the Recommendation System on Website

Deploying a recommendation system that has been create into a simple website using Flask Python and Heroku to make it easier for users to find hotels that are similar with the previous selected hotel. The appearance pf the recommendation system application website presented in Figure 2 and Figure 3.

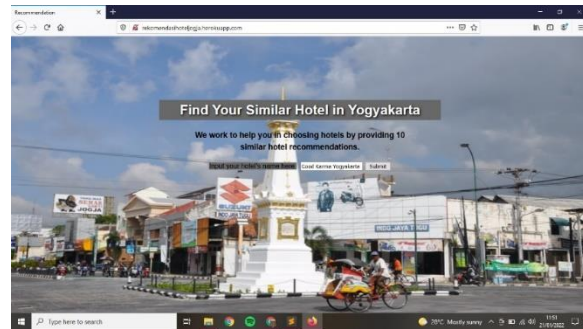


Figure 2. Search Page



Figure 3. Recommendation Result Page

Accessible via <https://rekomendasihoteljogja.herokuapp.com/>

4. Conclusion

Based on the result and discussion of this study, it can be concluded that document similarity calculation using cosine similarity in this study as an example of using the 132nd document. There were 10 documents with the highest similarity values which are 170th, 303rd, 210th, 360th, 315th, 422nd, 510th, 45th, 63rd, and 223rd documents. These documents have a cosine similarity value of 0.956666513, 0.946570717, 0.917459394, 0.912534877, 0.886439718, 0.88221982, 0.881775275, 0.875845794, 0.872030219, and 0.871514859. The result of hotel recommendations in Yogyakarta as an example using the 132nd document, namely Good Karma Yogyakarta-Hostel with Content-Based Filtering generated in this study based on the similarity of hotel descriptions are Happy Buddha Yogyakarta – Hostel, Nextdoor Homestay, Hotel Puspita, OYO 426 Hotel Gading Resto, Omah Jegog Homestay, Prawirotaman Homestay, RedDoorz near Prawirotaman, Ayodhya Garden Hostel Yogyakarta by HOM, Bringin House Yogyakarta, and House 24 Yogyakarta.

REFERENCES

- Aamir, M., & Bhusry, M. (2015). Recommendation System: State of the Art Approach. *International Journal of Computer Applications*, 120(12), 25–32.
- Ahsan. (2020.) Staycation Kala Pandemi, Bandung dan Yogyakarta Jadi Incaran. Diambil kembali dari <https://travel.wego.com/berita/staycationkala-pandemi-bandung-dan-yogyakarta-jadi-incaran/>
- Asrianti, S, &Azizah, N. (2020). Pandemi Corona Ubah Kecenderungan Wisatawan Pilih Akomodasi. Diambil kembali pada <https://www.republika.co.id/berita/qc6j66463/pandemicoronaubah-kecenderungan-wisatawan-pilih-akomodasi>
- Hendriyati, L. (2019). Online travel agent. *Jurnal Media Wisata*, 17(1), 1–10.
- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2011). *Recommender systems: an introduction*. Cambridge University Press.
- Larasati, F. B. A., & Februriyanti, H. (2021). Sistem Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode Content - Based Filtering. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 45.
- Layton, Julia. 2009. What is a Staycation. Diambil kembali pada <https://money.howstuffworks.com/personal-finance/budgeting/staycation.htm>
- Mondi, R. H., Wijayanto, A., & Winarno. (2019). Recommendation System With Content-Based Filtering Method for Culinary Tourism in Mangan Application. *Itsmart*, 8(2), 65–72.
- Nurjannah, M., & Fitri Astuti, I. (2013). Penerapan Algoritma Term Frequency-Inverse Document Frequency (Tf-Idf) Untuk Text Mining. Mahasiswa S1 Program Studi Ilmu Komputer FMIPA Universitas Mulawarman Dosen Program Studi Ilmu Komputer FMIPA Universitas Mulawarman. *Jurnal Informatika Mulawarman*, 8(3), 110–113.
- Purnaramadhan, R. (2021). Recommendation System Model Untuk Merekomendasikan Produk Pada Website Menggunakan Metode Content-Based. *Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Islam Indonesia*.
- Raghuwanshi, S. K., & Pateriya, R. K. (2019). Recommendations System : Techniques, Challenges, Application, and Evaluation. *Advances in Intelligent Systems and Computing*, 2(January), 107–119.
- Ricci, F., Rokch, L., & Shapira, B. (2015). *Recommender Systems Handbook* (F. Ricci, L. Rokach, & B. Shapira (eds.); Second Edi). Springer Science+Business Media.
- Rismanto, R., Syulistyo, A. R., & Agusta, B. P. C. (2020). Research supervisor recommendation system based on topic conformity. *International Journal of Modern Education and Computer Science*, 12(1), 26–34
- Sandi Wachyuni, S., & Wiweka, K. (2020). Kepuasan Wisatawan Dalam Penggunaan E-Commerce Agoda Dalam Pemesanan Hotel. *Journal of Tourism Destination and Attraction*, 8(1), 275.
- Supriyati, E., & Iqbal, M. (2018). Pengukuran Similarity Tema Pada Juz 30 Al Qur'an Menggunakan Teks Klasifikasi. *Simetris: Jurnal Teknik Mesin, Elektro Dan Ilmu Komputer*, 9(1), 361–370.

Spatial Autoregressive Moving Average Pada Pemodelan Persentase Penduduk Miskin Kabupaten/Kota Di Provinsi Jawa Timur Tahun 2020

Kiki Laila Nirmala⁽¹⁾, Wara Pramesti⁽²⁾, Fenny Fitriani⁽³⁾

^(1,2,3)Program Studi Statistika Universitas PGRI Adi Buana Surabaya,

Jalan Dukuh Menanggal gang XII/4 Surabaya

e-mail: kiki.lailanirmala@gmail.com⁽¹⁾, warapra@unipasby.ac.id⁽²⁾, fenny_f@unipasby.ac.id⁽³⁾

ABSTRAK

Persentase penduduk miskin yang tinggi menunjukkan bahwa tingkat kemiskinan di suatu wilayah juga tinggi. Persentase penduduk miskin sering menjadi pembicaraan publik. Kenaikan ataupun penurunan persentase penduduk miskin tetap menjadi permasalahan di setiap daerah. Sehingga pemerintah daerah tetap berusaha untuk mengendalikan kesejahteraan daerahnya dengan melakukan perlindungan serta menyejahterakan kehidupan sosial dan ekonomi daerahnya. Naik turunnya persentase penduduk miskin ini dipengaruhi oleh beberapa variabel. Untuk mengetahui variabel-variabel tersebut dapat digunakan salah satunya dengan pendekatan regresi berganda. Akan tetapi, dalam permasalahan kemiskinan ini tentu dipengaruhi juga oleh keadaan sekitarnya. Hal ini menyebabkan kemungkinan asumsi residual independen tidak dipenuhi, sehingga analisis regresi dapat dikembangkan ke metode spasial area. Metode regresi berbasis spasial area ini biasa disebut dengan metode Spatial Autoregressive Moving Average (SARMA). Hasil analisis yang dilakukan menunjukkan bahwa selain dipengaruhi oleh variabel-variabel yang diduga berpengaruh, juga dipengaruhi oleh wilayah sekitar.

Kata kunci : *Metode spasial area, Penduduk miskin, Regresi, SARMA.*

ABSTRACT

A high percentage of poor people indicates that the poverty rate in an area is also high. The percentage of poor people is often a public talk. The increase or decrease in the percentage of the poor remains a problem in every region. So that the local government must try to control the welfare of the region by protecting and prospering the social and economic life of the region. The increase and decrease of the percentage of the poor is influenced by several variables. To find out these variables, one of them can be used with a multiple regression approach. However, the problem of poverty is of course also influenced by the surrounding conditions. This causes the possibility of independent residual assumptions not being met, so that the regression analysis can be extended to the spatial area method. This spatial area-based regression method is commonly referred to as the Spatial Autoregressive Moving Average method (SARMA). The results of the analysis carried out show that apart from being influenced by variables that are thought to have an effect, it is also influenced by the surrounding area.

Keywords : *spatial area method, Poor people, Regression, SARMA.*

1. PENDAHULUAN

Kenaikan dan penurunan persentase kemiskinan disebabkan oleh banyak faktor, antara lain kenaikan harga barang yang disebabkan oleh naiknya harga BBM pada September 2013 dan Maret 2015. Didasarkan pada (BPS, Persentase Penduduk Miskin Maret 2021 turun menjadi 10,14 persen, 2021) diketahui bahwa prosentase penduduk miskin pada Maret 2021 meningkat sebesar 0,36 jika dibandingkan dengan periode Maret 2020, akan tetapi prosentase tersebut menurun sebesar 0,05 jika dibandingkan dengan periode September 2020. Hal

ini juga berlaku untuk prosentase kemiskinan di Prov Jawa Timur. Dari (BPS, 2022) didapatkan bahwa prosentase penduduk miskin Prov Jawa Timur pada periode Maret 2021 mengalami peningkatan 0,31 jika dibandingkan dengan periode Maret 2020 dan mengalami penurunan 0,06 jika dibandingkan dengan periode September 2020. Dari data tersebut, prosentase kemiskinan meningkat pada saat terjadi pandemi covid-19 mulai bulan Maret 2020 (Maaruf, 2021). Selain adanya covid-19 tentu ada beberapa faktor yang diduga sebagai penyebab kenaikan persentase penduduk miskin.

Untuk mengetahui beberapa faktor yang diduga berpengaruh dari penyebab naik turunnya persentase tersebut dapat digunakan analisis regresi. Analisis regresi ini merupakan salah satu metode statistika yang dapat digunakan untuk mengetahui seberapa besar pengaruh variabel prediktor terhadap variabel respon, sehingga dihasilkan suatu model persamaan yang dapat digunakan untuk memprediksi variabel respon jika diketahui variabel prediktornya. Pada analisis regresi diasumsikan bahwa residual identik, independen dan berdistribusi normal. Apabila asumsi residual identik tidak dipenuhi, maka dilakukan deteksi awal apakah ada efek spasial. Jika didapatkan ada efek spasial, maka digunakan metode Regresi Spasial Titik. Demikian pula apabila asumsi residual independen tidak dipenuhi dan ternyata ada efek spasial, maka digunakan metode Regresi Spasial Area. Apabila asumsi residual berdistribusi normal tidak terpenuhi, ini diduga ada data outlier, dan dilakukan transformasi yang sesuai.

Metode regresi yang berbasis spasial area ini biasa disebut dengan *Spatial Autoregressive Moving Average* (SARMA). Seperti yang dijabarkan di atas, metode SARMA ini digunakan jika diasumsikan variabel dependen dan sisaan terpengaruh juga oleh efek spasial. Selain itu, menurut Huang pada tahun 1984 menyatakan bahwa untuk model SARMA ini biasanya digunakan dalam melakukan analisis data cross section dengan matriks pembobot spasial sebagai bentuk hubungan antar daerah (Lispani, Sumarjaya, & Sukarsa, 2018). Metode SARMA ini telah digunakan dalam banyak penelitian, diantaranya adalah penelitian yang menerapkan SARMA pada 1) pemodelan pada masalah DBD di Kalimantan Timur dan Kalimantan Tengah (Sari, Hayati, & Wahyuningsih, 2021), 2) pemodelan tindakan kriminalitas di Jatim (Lispani, Sumarjaya, & Sukarsa, 2018), 3) pemodelan kasus pneumonia (Widyastuti, Srinadi, & Susilawati, 2019), dan 4) pemodelan tingkat pengangguran terbuka di Jawa Tengah (Wardani, Handajani, & Zukhronah, 2019)

Penelitian terdahulu yang menyangkut masalah penduduk miskin ini antara lain 1) penelitian yang dilakukan oleh (Purba & Soleman, 2020) dengan menggunakan Spatial Autoregressive (SAR) dan Regresi Spatial Error Model (SEM). Hasil dari pemodelan dengan SAR menunjukkan bahwa variabel IPM signifikan berpengaruh terhadap kemiskinan. Pemodelan dengan SEM menunjukkan bahwa variabel IPM signifikan berpengaruh terhadap kemiskinan. 2) penelitian yang dilakukan (Septiana & Rusdiansyah, 2019) didapatkan bahwa hasil pemilihan model terbaik

melalui uji diagnostik regresi spasial diperoleh model terbaik adalah SAR dimana PDRB dan Pengangguran berpengaruh signifikan terhadap kemiskinan. 3) penelitian yang dilakukan oleh (Laswinia & Chamid, 2016) mendapatkan bahwa faktor IKLH, IPM, kepadatan penduduk, dan laju pertumbuhan ekonomi adalah hal yang berhubungan dengan prosentase penduduk miskin. 4) penelitian yang dilakukan oleh (Ristika, Primandhana, & Wahed, 2021) mendapatkan bahwa jumlah penduduk dan indeks pembangunan manusia berdampak signifikan terhadap tingkat kemiskinan.

2. METODE PENELITIAN

Data yang digunakan dalam penelitian ini berupa data sekunder yang diperoleh dari Badan Pusat Statistik Jawa Timur pada tahun 2020, terdiri dari persentase penduduk miskin sebagai variabel respon. Wilayah yang digunakan dalam penelitian adalah kabupaten/kota di Jawa Timur terdiri dari 38 kabupaten/kota.

Variabel yang digunakan sebanyak delapan, terdiri dari satu variabel respon, yaitu Presentase Penduduk Miskin (Y) dan tujuh variabel prediktor. Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan < SD (X_1), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SD/SLTP (X_2), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SLTA (X_3), Presentase Sumber Air Minum Kemasan Bermerk (X_4), Presentase Sumber Air Minum Isi Ulang (X_5), Presentase Penduduk di Jawa Timur dan Angka Kesakitan (X_6), serta Presentase Penduduk di Jawa Timur yang Memiliki Jaminan Kesehatan (X_7).

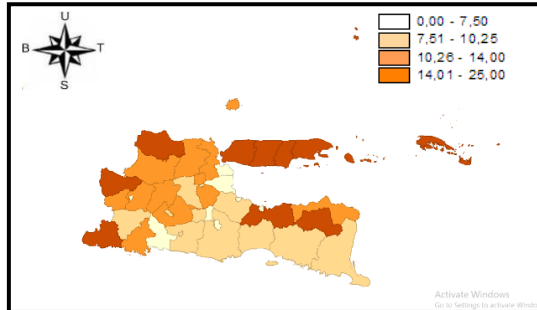
Langkah-langkah penyelesaian penelitian dan analisis data sebagai berikut:

1. Mendeskripsikan presentase penduduk miskin
2. Deteksi multikolinieritas.
3. Estimasi parameter dan pengujian parameter regresi OLS
4. Uji Asumsi residual IIDN
5. Menentukan matriks pembobot spasial dan menstandarisasi
6. Menguji efek ketergantungan spasial dengan uji dependensi Moran's I
7. Penentuan model spasial dengan uji Lagrange Multiplier (LM)
8. Pemodelan Regresi Spasial (SAR, SEM dan SARMA)
9. Pemilihan model terbaik berdasarkan R^2 dan AIC
10. Interpretasi model terbaik

11. Menarik kesimpulan

3. **HASIL DAN PEMBAHASAN**

Persentase penduduk miskin di Jawa Timur dapat digambarkan secara visual sebagai berikut :



Gambar 1. Sebaran persentase penduduk miskin kabupaten/kota di Jawa Timur 2020

Gambar 1 merupakan pola sebaran persentase kemiskinan kabupaten dan kota di Jawa Timur yang terbagi menjadi 4 kelompok. Kelompok pertama 0% sampai 7,5% terdiri dari sembilan kabupaten/kota, kelompok kedua terdiri dari sepuluh kabupaten/kota dengan persentase 7,51% sampai 10,25 % , kelompok ketiga terdapat sepuluh kabupaten/kota dengan persentase 10,26 % sampai 14%, dan kelompok keempat sembilan kabupaten/kota dengan presentase sebesar 14,01% sampai 25%.

Analisis Regresi

Analisis regresi dilakukan untuk mengetahui pola hubungan antara respon persentase kemiskinan dengan semua variabel prediktornya. Analisis regresi dapat dilakukan apabila tidak terjadi kasus multikolinieritas antar variabel prediktor. Berikut hasil pengujian kasus multikolinieritas :

Tabel 1. Nilai VIF

Variabel	VIF
X1	9,64
X2	4,08
X3	5,64
X4	4,44
X5	2,98
X6	1,28
X7	1,86

Kasus multikolinieritas dapat dideteksi salah satunya dengan melihat besaran Varians Inflasi Faktor (VIF). Apabila nilai VIF ≥ 10 , maka terjadi kasus multikolinieritas dan harus diselesaikan

terlebih dahulu. Pada Tabel 1 nilai VIF semua kurang dari sepuluh, maka tidak terjadi kasus multikolinieritas, sehingga dapat dilanjutkan untuk melakukan analisis regresi klasik.

Analisis regresi klasik digunakan untuk melihat pola hubungan secara matematis antara variabel respon dan prediktor. Model regresi dikatakan baik apabila pengujian parameternya signifikan dan asumsi-asumsi residual dipenuhi. Pengujian parameter regresi dilakukan terhadap parameter β , dan asumsi residual identik independen dan berdistribusi normal harus dipenuhi. Menurut (Drapper & Smith, 1992), estimasi parameter regresi dilakukan dengan metode kuadrat terkecil (*Ordinary Least Square*), yang sering ditulis dengan OLS. Hasil estimasi regresi klasik sebagai berikut:

Tabel 2. Hasil Estimasi Parameter Regresi

Parameter	Estimasi	P-Value	Keputusan
β_0	$5,944 \times 10^{-16}$	1	Ho tidak ditolak
β_1	0,0122	0,76	Ho tidak ditolak
β_2	-0,0126	0,62	Ho tidak ditolak
β_3	-0,0247	0,41	Ho tidak ditolak
β_4	-0,0424	0,13	Ho tidak ditolak
β_5	0,0489	0,02	Tolak Ho
β_6	0,0096	0,51	Ho tidak ditolak
β_7	0,0150	0,05	Tolak Ho
R² = 0, 5329 (53,29 %)			
Keterangan $t_{(0,10/30)}=1,31042$			

Berdasarkan Tabel 2, parameter yang berbeda secara signifikan adalah β_5 dan β_7 , dan memperoleh nilai koefisien determinasi sebesar 53.29%.

Model regresi yang terbentuk:

$$\hat{y} = 5,944 \times 10^{-16} + 0,0122x_1 - 0,0126x_2 - 0,0247x_3 - 0,0424x_4 + 0,0489x_5 + 0,0096x_6 + 0,0150x_7 \tag{1}$$

Selanjutnya dilakukan pengujian terhadap parameter regresi. Uji yang pertama adalah uji simultan sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \text{Paling tidak ada satu } \beta_j \neq 0$$

P-Value yang diperoleh 0.0009012, maka dapat disimpulkan bahwa H_0 ditolak, yang artinya

minimal ada satu parameter yang nilainya tidak nol, dan model signifikan.

Uji signifikansi yang kedua, yaitu uji parsial atau individu digunakan untuk mengetahui parameter-parameter yang tidak bernilai nol, hasilnya dapat dilihat pada Tabel 2. Parameter yang signifikan β_5 dan β_7 .

Pengujian Asumsi Residual

Asumsi-asumsi residual yang harus dipenuhi sebagai berikut:

1. Uji Identik

Deteksi adanya heteroskedastisitas dalam model regresi diuji dengan uji *Breusch-Pagan*. Apabila uji identik tidak dipenuhi, kemungkinan terdapat efek spasial. Hipotesis yang diuji adalah sebagai berikut:

$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (Varians residual homogen)

H_1 : minimal terdapat satu $\sigma_n^2 \neq \sigma^2$ (Varians residual tidak homogen)

Hasil perhitungan nilai statistik uji *Breusch-Pagan* 8.5805, dengan derajat bebas 7 dan p value 0.2842 lebih besar dari nilai alpha 0.10, artinya H_0 tidak ditolak. Kesimpulan asumsi varians residual homogen dipenuhi.

2. Uji Independen

Uji residual independen digunakan untuk mengetahui apakah antara residual satu dengan yang lain independen. Apabila tidak independen, diduga ada efek spasial pada persentase penduduk miskin. Pengujian dilakukan dengan menggunakan uji *Durbin Watson*. Langkah pengujian dan hasil uji *Durbin Watson*, sebagai berikut :

$H_0: \rho = 0$ (tidak ada autokorelasi antar residual)

$H_1: \rho \neq 0$ (ada autokorelasi antar residual)

Statistik uji *Durbin Watson* 1.3897, dengan p value 0.0162 lebih kecil dari nilai alpha 0.10, artinya H_0 ditolak, jadi residual tidak independen, sehingga diindikasikan adanya efek spasial pada variabel respon.

3. Uji Normalitas

Uji normalitas untuk mengetahui apakah asumsi residual berasal dari distribusi normal dipenuhi. Pengujian asumsi residual berdistribusi normal $(0, \sigma^2)$ dilakukan dengan menggunakan uji *Kolmogorov-Sminorv*.

Hipotesis yang digunakan adalah :

$H_0: F_0(\epsilon) = F(\epsilon)$ (Residual berdistribusi normal)

$H_1: F_0(\epsilon) \neq F(\epsilon)$ (Residual tidak berdistribusi normal)

Statistik uji Kolmogorov-Smirnow 0.1166 dengan p value 0.64 lebih dari $\alpha = 0.10$, artinya H_0 tidak ditolak dan dapat disimpulkan bahwa residual berasal dari distribusi normal.

Asumsi residual independen pada analisis regresi klasik tidak dipenuhi, hal ini mengindikasikan bahwa terdapat efek spasial pada variabel respon (persentase penduduk miskin). Dugaan adanya efek spasial dapat diuji dengan menggunakan uji *Moran's I*. Untuk melakukan uji *Moran's I* diperlukan matriks pembobot. Matriks pembobot yang digunakan adalah pembobot *Queen contiguity*. *Queen contiguity* merupakan matriks yang menunjukkan keterkaitan antar daerah atau matriks yang menggambarkan hubungan kedekatan antar daerah. Pembentukan matriks pembobot spasial ini dilambangkan dengan huruf **W**.

Langkah pengujian *Moran's I*:

$H_0 : I = 0$ (tidak ada autokorelasi antar lokasi)

$H_1 : I \neq 0$ (ada autokorelasi antar lokasi)

Statistik Uji *Moran's I* 0.22826437 dengan p value 0.03056 kurang dari 0.10, maka H_0 ditolak, yang dapat disimpulkan terdapat autokorelasi spasial antar lokasi pengamatan.

Untuk mengetahui apakah persentase penduduk miskin dipengaruhi oleh variabel prediktor dan wilayah sekitar dapat digunakan metode spasial, tetapi sebelum ditentukan metode spasial yang sebaiknya digunakan, maka dilakukan pengujian dengan *Lagrange Multiplier*. Hasil uji *Lagrange Multiplier* sebagai berikut :

Tabel 4 Nilai *Lagrange Multiplier* (LM)

	Statistic	df	P-Value	Kriteria
LM lag	5,4408	1	0,0197	Tolak H_0
LM error	1,4926	1	0,2218	H_0 tidak ditolak
LM SARMA	9,489	2	0,008	Tolak H_0

Ket : Siginifikan pada $\alpha = 10\%$

Berdasarkan Tabel 4 didapatkan bahwa:

- a) P-Value LM lag sebesar 0,0197, lebih kecil dari $\alpha=10\%$, berarti bahwa terdapat dependensi spasial lag sehingga perlu dilanjutkan ke Pemodelan SAR.
- b) P-Value LM error sebesar 0.2218, lebih besar dari $\alpha = 10\%$, maka dapat disimpulkan bahwa tidak terdapat dependensi spasial error sehingga tidak perlu dilanjutkan ke pemodelan SEM.
- c) P-Value LM lag dan error sebesar 0.0087, lebih

kecil dari $\alpha = 10\%$, artinya bahwa terdapat dependensi spasial lag dan error sehingga dapat dilakukan pemodelan campuran atau SARMA.

Pemodelan Regresi Spasial (*Spatial Autoregressive Model*)

Hasil uji *Lagrange Multiplier* menunjukkan bahwa terdapat autokorelasi spasial lag pada variabel respon, sehingga dapat dilakukan analisis pemodelan *Spatial Autoregressive Model* (SAR). Estimasi parameter menggunakan uji wald. Pengujian parameter menghasilkan *p value* lebih kecil dari $\alpha = 10\%$, untuk β_5 , dan β_7 , sehingga dapat disimpulkan bahwa parameter-parameter tersebut signifikan secara statistik.

Tabel 5. Nilai Estimasi Parameter SAR

Parameter	ρ	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
Estimasi	0,312	-0,059	0,155	0,067	-0,143
Std. Error	0,152	0,104	0,319	0,209	0,247
Z-hitung	2,052	0,563	0,486	0,321	-0,579
P-value	0,040	0,573	0,627	0,748	0,562

Berdasarkan hasil estimasi parameter pada Tabel 5, dapat ditentukan model SAR sebagai berikut:

$$\hat{y}_i = \rho \sum_{i=1, i \neq j}^{38} w_{ij} y_j + \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} \quad (2)$$

$$\hat{y}_i = 0,312 \sum_{i=1, i \neq j}^{38} w_{ij} y_j - 0,059 + 0,155 X_{1i} - 0,067 X_{2i} - 0,143 X_{3i} - 0,369 X_{4i} + 0,420 X_{5i} - 0,062 X_{6i} - 0,346 X_{7i} \quad (3)$$

Interpretasi dari model SAR pada persamaan diatas. Koefisien Rho (ρ) menunjukkan bahwa suatu wilayah dikelilingi oleh wilayah lain, maka pengaruh dari masing-masing wilayah yang mengelilingi wilayah lain sebesar 0,312. Pada variabel (X_1) Presentase Penduduk Usia 15 Tahun ke atas Tamatan < SD naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,155. Jika variabel (X_2) Presentase Penduduk Usia 15 Tahun ke atas Tamatan SD/SLTP naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,067. Jika variabel Presentase Penduduk

Usia 15 Tahun ke atas Tamatan SLTA (X_3) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,143.

Jika variabel Presentase Sumber Air Minum Kemasan Bermerk naik satu satuan (X_4), maka kasus presentase penduduk miskin turun sebesar 0,369. Jika variabel Presentase Sumber Air Minum Isi Ulang naik satu satuan (X_5), maka kasus presentase penduduk miskin naik sebesar 0,420. Jika variabel Presentase Penduduk di Jawa Timur dengan Angka kesakitan (X_6) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,062. Jika variabel Presentase penduduk di Jawa Timur yang memiliki Jaminan Kesehatan (X_7) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,346.

Dari persamaan SAR dapat digambarkan dalam suatu wilayah. Misalkan diambil wilayah yang diamati yaitu Kabupaten Jombang. Kabupaten Jombang dengan kode wilayah 17 yang berbatasan dengan Kota Kediri dengan kode 06, Kabupaten Malang dengan kode 07, Kabupaten Mojokerto dengan kode 16, Kabupaten Nganjuk dengan kode 18, Kabupaten Bojonegoro dengan kode 22 dan Kabupaten Lamongan dengan kode 24, maka persamaan regresi dugaan adalah sebagai berikut:

Parameter	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$
Estimasi	-0,369	0,420	-0,062	-0,346
Std. Error	0,229	0,165	0,119	0,141
z-hitung	-1,607	2,554	-0,519	-2,459
P-value	0,108	0,010	0,603	0,013

$$\hat{y}_{jombang} = 0,312 \sum_{i=1, i \neq j}^{38} w_{ij} y_j - 0,059 - 0,155 X_{jombang1} - 0,067 X_{jombang2} - 0,143 X_{jombang3} - 0,369 X_{jombang4} + 0,420 X_{jombang5} - 0,062 X_{jombang6} - 0,346 X_{jombang7} \quad (4)$$

Interpretasi dari persamaan diatas, nilai koefisien Rho (ρ) sebesar 0,312 artinya terdapat pengaruh wilayah atau lokasi yang mengelilingi wilayah lain (kabupaten Jombang bertetangga dengan Kota Kediri, Kabupaten Malang, Kabupaten Mojokerto, Kabupaten Nganjuk, Kabupaten Bojonegoro dan Kabupaten Lamongan). Pada

variabel (X_1), yaitu Presentase Penduduk Usia 15 Tahun ke atas Tamatan < SD naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,155. Jika variabel Presentase Penduduk Miskin Usia 15 Tahun ke atas Tamatan SD/SLTP (X_2) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,067. Jika variabel Presentase Penduduk Usia 15 Tahun ke atas Tamatan SLTA (X_3) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,143.

Jika variabel Presentase Sumber Air Minum Kemasan Bermerk (X_4) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,369. Jika variabel Presentase Sumber Air Minum Isi Ulang (X_5) naik satu satuan, maka kasus presentase penduduk miskin naik sebesar 0,420. Jika variabel Presentase Penduduk di Jawa Timur dengan Angka kesakitan (X_6) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,062. Jika variabel Presentase penduduk di Jawa Timur yang Memiliki Jaminan Kesehatan (X_7) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,346.

Hasil uji *Lagrange Multiplier lag* menunjukkan terdapat efek ketergantungan lokasi pada suatu wilayah dengan wilayah lainnya yang berdekatan, ditunjukkan dari nilai ρ . *p-value* yang diperoleh kurang dari $\alpha = 10\%$. Koefisien determinasi (R^2) diperoleh sebesar 0,5850.

Hal tersebut menunjukkan bahwa variabel Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan < SD (X_1), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SD/SLTP (X_2), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SLTA (X_3), Presentase Sumber Air Minum Kemasan Bermerk (X_4), Presentase Sumber Air Minum Isi Ulang (X_5), Presentase Penduduk di Jawa Timur dan Angka (X_6), Presentase penduduk di Jawa Timur yang Memiliki Jaminan Kesehatan (X_7) mampu menjelaskan sebesar 58,50% terhadap persentase penduduk miskin dan sisanya sebesar 41,5 % dijelaskan oleh faktor lain yang tidak terdapat dalam model.

Spatial Autoregressive Moving Average Model (SARMA)

Hasil uji *Lagrange Multiplier* menunjukkan bahwa terdapat autokorelasi spasial *lag* dan error, sehingga dapat dilakukan analisis pemodelan SARMA. Estimasi parameter menggunakan uji wald. Pengujian parameter menghasilkan *p value* lebih kecil dari $\alpha = 10\%$,

Tabel 6 Nilai estimasi parameter SARMA

Parameter	Rho (ρ)	Lamda (λ)	$\hat{\beta}_0$	$\hat{\beta}_1$
Estimasi	0,690	-0,734	-0,088	0,079
Std. Error	0,098	0,154	0,046	0,222
z-hitung	7,029	-4,752	-1,888	0,356
P-value	2,084x10 ⁻¹²	2,013x10 ⁻⁶	0,059	0,722

Parameter	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
Estimasi	-0,018	-0,117	-0,401	0,482
Std. Error	0,129	0,175	0,194	0,118
z-hitung	-0,140	-0,668	-2,061	4,089
P-value	0,888	0,504	0,039	1,332x10⁻⁰⁵

Parameter	$\hat{\beta}_6$	$\hat{\beta}_7$
Estimasi	-0,091	-0,296
Std. Error	0,109	0,113
z-hitung	-0,830	-2,607
P-value	0,406	0,009

Model yang terbentuk sebagai berikut:

$$\hat{y}_i = \rho \sum_{j=1, i \neq j}^{38} w_{ij} y_j + \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \lambda \sum_{j=1, i \neq j}^{38} w_{ij} U_j \tag{5}$$

$$\hat{y}_i = 0,690 \sum_{j=1, i \neq j}^{38} w_{ij} y_j - 0,088 + 0,079 X_{i1} - 0,018 X_{i2} - 0,117 X_{i3} - 0,401 X_{i4} + 0,482 X_{i5} - 0,091 X_{i6} - 0,296 X_{i7} + u_i \tag{6}$$

$$u_i = -0,734 \sum_{j=1, i \neq j}^{38} w_{ij} u_j \tag{7}$$

Berdasarkan Persamaan (6) dapat dijelaskan bahwa, jika tidak terdapat pengaruh dari variabel apapun maka persentase penduduk miskin akan menurun sebesar 1.135. Nilai koefisien spasial lag (ρ) sebesar 0,690 dan nilai koefisien spasial error (λ) sebesar -0.734 signifikan pada $\alpha = 0,1$ sehingga dapat disimpulkan bahwa persentase penduduk miskin pada suatu wilayah atau lokasi ke-i akan berpengaruh terhadap persentase penduduk miskin pada lokasi yang menjadi tetangganya.

Jika variabel Presentase Penduduk Usia 15 Tahun ke atas Tamatan < SD (X_1) naik satu satuan, maka kasus presentase penduduk miskin naik sebesar 0,079. Jika variabel Presentase Penduduk Miskin Usia 15 Tahun ke atas Tamatan SD/SLTP (X_2) naik satu satuan, maka kasus presentase

penduduk miskin turun sebesar 0,018 . Jika variabel Presentase Penduduk Usia 15 Tahun ke atas (X_3) Tamatan SLTA naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,117.

Jika variabel Presentase Sumber Air Minum Kemasan Bermerk (X_4) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,401 . Jika variabel Presentase Sumber Air Minum Isi Ulang (X_5) naik satu satuan, maka kasus presentase penduduk miskin naik sebesar 0,482 . Jika variabel Presentase Penduduk di Jawa Timur dengan Angka kesakitan (X_6) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,091. Jika variabel Presentase penduduk di Jawa Timur yang Memiliki Jaminan Kesehatan (X_7) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,296 .

Dari persamaan SARMA dapat digambarkan dalam suatu wilayah. Misalkan diambil wilayah yang diamati yaitu Kabupaten Jombang. Kabupaten Jombang dengan kode wilayah 17 yang berbatasan dengan Kota Kediri dengan kode 06, Kabupaten Malang dengan kode 07, Kabupaten Mojokerto dengan kode 16, Kabupaten Nganjuk dengan kode 18, Kabupaten Bojonegoro dengan kode 22 dan Kabupaten Lamongan dengan kode 24, maka model persamaan regresi dugaan adalah sebagai berikut :

$$\hat{y}_{Jombang} = 0,690 - 0,088 + 0,079 X_{jombang1} - 0,018X_{jombang2} - 0,117X_{jombang3} - 0,401X_{jombang4} + 0,482X_{jombang5} - 0,091X_{jombang6} - 0,296 X_{jombang7} - 0,734 \tag{8}$$

$$u_{jombang} = -0,734 \sum_{j=1, i \neq j}^n w_{ij} u_j \tag{9}$$

Interpretasi dari model diatas, apabila nilai koefisien Rho (ρ) bermakna, maka suatu wilayah dikelilingi wilayah lain dan pengaruh masing-masing wilayah sebesar 0,690. Jika variabel Presentase Penduduk Usia 15 Tahun ke atas Tamatan < SD (X_1) naik satu satuan, maka kasus presentase penduduk miskin naik sebesar 0,079 . Jika variabel Presentase Penduduk Usia 15 Tahun ke atas Tamatan SD/SLTP (X_2) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,018. Jika variabel Presentase Penduduk Usia 15 Tahun keatas Tamatan SLTA

(X_3) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,117 .

Jika variabel Presentase Sumber Air Minum Kemasan Bermerk (X_4) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,401 . Jika variabel Presentase Sumber Air Minum Isi Ulang (X_5) naik satu satuan, maka kasus presentase penduduk miskin naik sebesar 0,482 . Jika variabel Presentase Penduduk di Jawa Timur dengan Angka kesakitan (X_6) naik satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,091 . Jika variabel Presentase penduduk di Jawa Timur yang Memiliki Jaminan Kesehatan naik (X_7) satu satuan, maka kasus presentase penduduk miskin turun sebesar 0,296 . Nilai koefisien spasial error (λ) sebesar 0,734 artinya presentasi penduduk miskin pada suatu wilayah atau lokasi ke-i akan berpengaruh terhadap presentase penduduk miskin pada lokasi yang menjadi tetangganya .

Hasil uji *Lagrange Multiplier lag dan error* menunjukkan terdapat efek ketergantungan lokasi pada suatu wilayah dengan wilayah lainnya yang berdekatan, ditunjukkan dari nilai λ dan ρ pada Tabel 5 dan *p-value* yang diperoleh kurang dari $\alpha = 10\%$. Koefisien determinasi (R^2) diperoleh sebesar 0,6748.

Hal ini dapat disimpulkan bahwa variabel variabel Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan < SD (X_1), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SD/SLTP (X_2), Presentase Penduduk Usia 15 Tahun ke atas dan Tamatan SLTA (X_3), Presentase Sumber Air Minum Kemasan Bermerk (X_4), Presentase Sumber Air Minum Isi Ulang (X_5), Presentase Penduduk di Jawa Timur dan Angka (X_6), Penduduk Presentase di Jawa Timur yang Memiliki Jaminan Kesehatan (X_7) mampu menjelaskan sebesar 67,49 % terhadap presentase penduduk miskin dan sisanya sebesar 32,52 % dijelaskan oleh variabel lain yang tidak terdapat pada model .

Pemilihan Model Terbaik

Kriteria pemilihan model terbaik yang digunakan nilai *R-Square* (R^2) dan *Akaike's Information Criterion* (AIC). Pemilihan didasarkan pada nilai (R^2) terbesar dan nilai AIC terkecil.

Tabel 7 Nilai R^2 dan nilai AIC

Model	R^2	AIC
Regresi Linier Berganda (OLS)	53,29%	95,89
<i>Spatial Autoregressive Model</i> (SAR)	58,50%	93,40

<i>Spatial Autoregressive Moving Average (SARMA)</i>	67,49 %	86,134
--	----------------	---------------

Berdasarkan Tabel 7 diketahui bahwa nilai R² terbesar dan AIC terkecil diperoleh pada model SARMA berturut-turut dengan nilai R² sebesar 67,49% dan AIC sebesar 86,134. Hal ini menunjukkan bahwa pemodelan Persentase SARMA merupakan model terbaik yang digunakan untuk memodelkan persentase penduduk miskin di Jawa Timur tahun 2020. .

Persamaan model SARMA sebagai berikut

$$\hat{y}_i = 0,690 \sum_{j=1, i \neq j}^{38} w_{ij} y_j - 0,088 + 0,079 X_{i1} - 0,018 X_{i2} - 0,117 X_{i3} - 0,401 X_{i4} + 0,482 X_{i5} - 0,091 X_{i6} - 0,296 X_{i7} + u_i \tag{10}$$

$$u_i = -0,734 \sum_{j=1, i \neq j}^{38} w_{ij} u_j \tag{11}$$

4. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan didapatkan kesimpulan bahwa pola persebaran persentase penduduk miskin di kabupaten dan kota di Jawa Timur terbagi menjadi empat kelompok. Kelompok pertama mencakup 9 kabupaten/kota, kelompok kedua mencakup 10 kabupaten/kota, kelompok ketiga mencakup 10 kabupaten/kota dan kelompok keempat mencakup 9 kabupaten/kota.

Model terbaik yaitu SARMA dengan pembobot *Queen Contiguity*, nilai R² sebesar 67,49% dan AIC sebesar 86,134. Faktor yang berpengaruh signifikan yaitu presentase sumber air minum kemasan bermerk, persentase sumber air minum isi ulang dan variabel persentase rumah tangga memiliki jaminan kesehatan.

DAFTAR PUSTAKA

BPS. (2021). *Persentase Penduduk Miskin Maret 2021 turun menjadi 10,14 persen*. Jakarta: Badan Pusat Statistik.
 BPS. (2022, Juni 20). *Persentase Penduduk Miskin Menurut Provinsi (Persen), 2019-2021*. Diambil kembali dari Badan Pusat Statistik Provinsi Jawa Timur: <https://jatim.bps.go.id/indicator/23/344/1/>

persentase-penduduk-miskin-menurut-provinsi-.html
 Drapper, N., & Smith, H. (1992). *Analisis Regresi Terapan Edisi Kedua*. Penerjemah: Bambang Sumantri. Jakarta: Gramedia Pustaka Umum.
 Laswinia, V. D., & Chamid, M. S. (2016). Analisis Pola Hubungan Persentase Penduduk Miskin dengan Faktor Lingkungan, Ekonomi, dan Sosial di Indonesia Menggunakan Regresi Spasial. *JURNAL SAINS DAN SENI ITS*, 5(2), D-235 - D-240.
 Lispani, N. M., Sumarjaya, I. W., & Sukarsa, I. K. (2018). Pemodelan Jumlah Tindak Kriminalitas Di Provinsi Jawa Timur Dengan Analisis Regresi Spasial Autoregressive And Moving Average. *E-Jurnal Matematika*, 7(4), 346-356.
 Maaruf, E. (2021, Agustus 03). *Angka Kemiskinan di Indonesia Naik Akibat Pandemi Covid-19*. Diambil kembali dari Sindonews.com: <https://nasional.sindonews.com/read/500774/15/angka-kemiskinan-di-indonesia-naik-akibat-pandemi-covid-19-1627999725>
 Purba, N. S., & Soleman, L. A. (2020). Analisis Spasial Mengenai Faktor-Faktor Yang Memengaruhi Kemiskinan Di Provinsi Papua Tahun 2019. *Seminar Nasional Hasil Penelitian & Pengabdian Kepada Masyarakat* (hal. 71-80). Banten: Universitas Pamulang.
 Ristika, E. D., Primandhana, W. P., & Wahed, M. (2021). Analisis Pengaruh Jumlah Penduduk, Tingkat Pengangguran Terbuka Dan Indeks Pembangunan Manusia Terhadap Tingkat Kemiskinan Di Provinsi Jawa Timur. *Eksis, Jurnal Ilmiah Ekonomi dan Bisnis*, 12(2), 118-122.
 Sari, D. N., Hayati, M. N., & Wahyuningsih, S. (2021). Model Spatial Autoregressive Moving Average (SARMA) pada Data Jumlah Kejadian Demam Berdarah Dengue (DBD) di Provinsi Kalimantan Timur dan Tengah Tahun 2016. *EKSPONENSIAL*, 11(1), 57-64.
 Septiana, S., & Rusdiansyah, R. (2019). PENGARUH PDRB, Jumlah Penduduk Dan Pengangguran Terhadap Kemiskinan Di Kota Banjarmasin. *JIEP: Jurnal Ilmu Ekonomi dan Pembangunan*, 2(1), 197-210.

- Wardani, I. K., Handajani, S. S., & Zukhronah, E. (2019). Pemodelan Tingkat Pengangguran Terbuka Provinsi Jawa Tengah Menggunakan Model Spatial Autoregressive Moving Average. *Prosiding Sendika* (hal. 199-205). Purworejo: Department of Mathematics Education Universitas Muhammadiyah Purworejo.
- Widyastuti, M. N., Srinadi, I. G., & Susilawati, M. (2019). Pemodelan Jumlah Kasus Pneumonia Balita Di Jawa Timur Menggunakan Regresi Spatial Autoregressive Moving Average. *E-Jurnal Matematika*, 8(3), 236-245.

Pemetaan Cakupan Pengguna Kb Aktif Dan *Unmet Need* Di Kabupaten Ponorogo Tahun 2021

Bella Dwi Saputri⁽¹⁾, Diah Indriani⁽²⁾

Fakultas Kesehatan Masyarakat, Universitas Airlangga, Mulyorejo, Kecamatan Mulyorejo, Surabaya, Jawa Timur, 60115 Indonesia

¹ bella.dwi.saputri-2018@fkm.unair.ac.id

² diah-i@fkm.unair.ac.id

ABSTRAK

Keluarga Berencana bertujuan untuk mengontrol jumlah penduduk dengan membatasi jumlah anak setiap keluarga untuk mencapai kesejahteraan keluarga. Program tersebut dilakukan dengan menghimbau Pasangan usia subur untuk menggunakan metode atau alat kontrasepsi sesuai kebutuhan dan keinginan. Pada tahun 2021, Kabupaten Ponorogo mengalami penurunan cakupan pengguna keluarga berencana aktif (Pasangan Usia Subur yang sedang menggunakan alat kontrasepsi) sekaligus peningkatan prevalensi *unmet need* (Pasangan Usia Subur yang kebutuhan berkeluarga berencananya belum terpenuhi). Sistem Informasi Geografi merupakan ilmu yang mempelajari pembuatan peta secara komputasi atau biasa disebut dengan pemetaan. Pemetaan dilakukan untuk memantau capaian cakupan pengguna keluarga berencana aktif dan *unmet need* setiap wilayah kecamatan di Kabupaten Ponorogo. Penelitian ini merupakan penelitian deskriptif yang menggambarkan pengguna keluarga berencana aktif dan *unmet need* di Kabupaten Ponorogo. Penelitian ini menggunakan data sekunder yang selanjutnya akan dianalisis secara spasial. Hasil penelitian menunjukkan bahwa terdapat 4 kecamatan yang memiliki cakupan pengguna keluarga berencana aktif dibawah capaian Kabupaten Ponorogo dan terdapat 9 kecamatan dengan prevalensi *unmet need* diatas capaian Kabupaten Ponorogo. Metode kontrasepsi yang paling diminati oleh masyarakat Kabupaten Ponorogo adalah suntik. Analisis kluster berdasarkan cakupan KB Aktif dan *unmet need* membagi Kabupaten Ponorogo menjadi 4 kluster. Hasil pemetaan yang telah dilakukan, dapat menjadi bahan pertimbangan untuk meningkatkan keberhasilan Program Keluarga Berencana di Kabupaten Ponorogo.

Kata Kunci: Keluarga Berencana, Pengguna Keluarga Berencana Aktif, *Unmet Need*, Pemetaan

ABSTRACT

Family Planning aims to control the population with limiting number of children in family to achieve family welfare. This program encouraging couples of childbearing age to use contraceptive methods or devices according to their needs and desires. In 2021, Ponorogo had a decrease of active family planning participants and an increase of prevalence unmet need (couple of childbearing age who need contraception but do not use contraception). Geographic Information System is a science that studies map making computationally or commonly called mapping. The mapping can used to monitor coverage of active family planning participants and prevalence unmet need of each sub-district. This study is a descriptive study that describes active family planning participants and unmet need in Ponorogo Regency. This study uses secondary data which will be analyzed spatially. The results showed that there were 4 sub-districts that had active family planning participant below the district's achievement and 9 sub-districts with the prevalence of unmet need above the district's achievement. Injection is the method of contraception that the people of Ponorogo Regency are most interested. Cluster analysis based on active family planning coverage and unmet need divided Ponorogo Regency into 4 clusters.. The results of this study can be taken into consideration to increase the success of Family Planning Program in Ponorogo Regency.

Keywords: Family Planning, Active Family Planning Participants, *Unmet Need*, Mapping

1. PENDAHULUAN

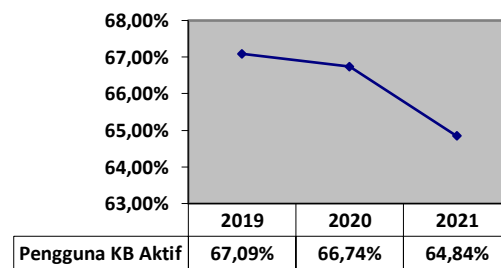
Indonesia mengalami peningkatan jumlah penduduk pada setiap tahunnya. Penduduk Indonesia pada tahun 2021 mencapai 272,30 juta jiwa (Dirjendukcapil, 2021). Jumlah ini mengalami kenaikan dibandingkan tahun sebelumnya yaitu pada tahun 2020 yaitu sebanyak 270,20 juta jiwa (BPS, 2020). Pulau Jawa menjadi pulau dengan jumlah penduduk terbanyak se-Indonesia dengan jumlah penduduk yang berada di provinsi Jawa Timur mencapai 40,67 juta jiwa pada tahun 2020. Jumlah penduduk pada tahun 2020 ini mengalami kenaikan dibandingkan tahun-tahun sebelumnya yaitu 39,5 juta pada tahun 2018 dan 39,70 juta jiwa pada tahun 2019 (BPS Provinsi Jawa Timur, 2021). Kabupaten Ponorogo juga mengalami kenaikan jumlah penduduk pada setiap tahunnya. Sebanyak 870.705 jiwa terdapat di Provinsi Jawa Timur pada tahun 2018, lalu meningkat menjadi 871.370 jiwa pada tahun 2019 dan kembali meningkat menjadi 949.318 pada tahun 2020 (BPS Provinsi Jawa Timur, 2020).

Kenaikan penduduk yang terjadi setiap tahun di Kabupaten Ponorogo dapat menyebabkan tingginya jumlah penduduk sehingga akan menimbulkan masalah kepadatan penduduk. Selain itu, banyaknya jumlah penduduk juga dapat menyebabkan munculnya berbagai masalah kesehatan yang terjadi di dalam keluarga dan tentunya akan mengancam kesejahteraan keluarga. Kenaikan jumlah penduduk pada setiap tahun ini tentu perlu menjadi perhatian khusus utamanya bagi Pemerintah Kabupaten Ponorogo sehingga harus dilakukan upaya untuk mengontrol jumlah penduduk di Kabupaten Ponorogo. Program KB merupakan upaya yang dilakukan pemerintah dalam usaha mengontrol jumlah penduduk.

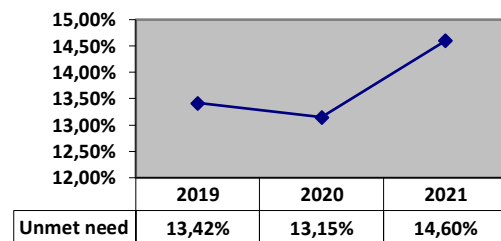
Program Keluarga Berencana merupakan upaya dalam membatasi jumlah anak setiap keluarga agar tercapai kesejahteraan dalam keluarga (Matahari et al., 2018). Beberapa indikator yang berperan untuk menunjukkan keberhasilan pelaksanaan Program Keluarga Berencana di suatu wilayah diantaranya adalah *Total Fertility Rate* (TFR) dan Laju pertumbuhan penduduk (LPP). TFR merupakan rata-rata jumlah anak yang dilahirkan oleh setiap wanita selama masa subur berlangsung (BPS, 2022). LPP merupakan angka yang menunjukkan rata-rata tingkat kenaikan jumlah penduduk per tahun pada suatu wilayah dan dalam jangka waktu tertentu (BPS, 2021). Laju pertumbuhan penduduk dan *Total Fertility Rate* yang ada di suatu wilayah dipengaruhi oleh banyak aspek, salah satunya dipengaruhi oleh jumlah Pasangan Usia Subur (PUS) yang berpartisipasi

dalam Program Keluarga Berencana (KB). Keterlibatan masyarakat dalam Program Keluarga Berencana dapat membantu menurunkan jumlah penduduk di suatu wilayah. Cakupan pengguna KB Aktif pada setiap wilayah dapat mengindikasikan keberhasilan program KB di wilayah tersebut. Pengguna KB Aktif merupakan pengguna KB yang saat ini sedang menggunakan kontrasepsi dengan tujuan menunda kehamilan, mengatur atau menjarangkan kehamilan, atau mengakhiri kesuburan. Pengguna KB Aktif dapat menggunakan jenis alat atau obat kontrasepsi yang tersedia diantaranya adalah suntik, pil, kondom, implan, Metode Operasi Wanita (MOW), Metode Operasi Pria (MOP), dan *Intrauterine Device* (IUD). Pemilihan jenis alat atau obat kontrasepsi ini disesuaikan dengan tujuan lama penggunaan alat atau obat kontrasepsi. PUS dapat memilih jenis kontrasepsi sesuai keinginan baik jenis kontrasepsi metode jangka panjang maupun jangka pendek.

Berikut terdapat data yang menunjukkan bahwa cakupan pengguna KB Aktif di Kabupaten Ponorogo mengalami penurunan dan prevalensi *unmet need* (PUS yang membutuhkan kontrasepsi namun tidak menggunakan kontrasepsi) mengalami peningkatan.



Gambar 1. Grafik presentase pengguna kb aktif



Gambar 2. Presentase *unmet need* di kab.ponorogo

Berdasarkan grafik diatas, pada tahun 2019 cakupan pengguna KB Aktif di Kabupaten Ponorogo sebanyak 67,09% atau 114.704 pengguna KB Aktif. Pada tahun 2020, cakupan pengguna KB Aktif di Kabupaten Ponorogo menurun menjadi 66,74% atau 114.760 pengguna. Pada tahun 2021, cakupan

pengguna KB Aktif di Kabupaten Ponorogo kembali mengalami penurunan menjadi 64,84% atau 110.959 pengguna.

Selain presentase kepesertaan KB aktif, persentase *unmet need* juga digunakan sebagai indikator keberhasilan pelaksanaan Program KB.. *Unmet need* merupakan kebutuhan KB yang belum terpenuhi atau dengan kata lain *unmet need* terjadi ketika keinginan atau kebutuhan PUS untuk menggunakan alat atau obat kontrasepsi namun belum terpenuhi sehingga PUS tersebut memutuskan untuk tidak menggunakan alat atau metode kontrasepsi. Data Laporan Statistika Rutin menunjukkan bahwa prevalensi kejadian *unmet need* di Kabupaten Ponorogo mengalami peningkatan dalam 2 tahun terakhir yaitu sebagai berikut. Prevalensi *unmet need* di Kabupaten Ponorogo pada tahun 2021 mencapai 14,60%. Prevalensi ini mengalami peningkatan dibandingkan tahun 2020 yaitu sebanyak 13,15% kejadian *unmet need*.

Penurunan cakupan pengguna KB Aktif dan peningkatan prevalensi *unmet need* di Kabupaten Ponorogo pada tahun 2021 menjadi salah satu indikator bahwa Program Keluarga Berencana belum berjalan secara maksimal. Oleh karenanya dapat dilakukan sebuah pemetaan yang bertujuan untuk memantau perkembangan pelaksanaan Program Keluarga Berencana pada setiap wilayah kecamatan yang ada di Kabupaten Ponorogo. Pemetaan ini merupakan pengaplikasian Sistem Informasi Geografis (SIG) yaitu ilmu yang mempelajari mengenai pembuatan peta secara komputasi dengan tahapan meliputi input, proses, serta output (Sugaadi et al., 2009). Artikel ini bertujuan untuk membuat pemetaan mengenai capaian cakupan pengguna KB Aktif dan pemetaan prevalensi *unmet need* pada setiap wilayah kecamatan di Kabupaten Ponorogo pada tahun 2021. Selain itu, akan dilakukan pemetaan hasil pengklusteran berdasarkan cakupan KB Aktif dan prevalensi *unmet need*. Hasil pemetaan ini dapat digunakan sebagai alternatif solusi untuk mengatasi permasalahan cakupan KB Aktif yang rendah dan prevalensi *unmet need* yang tinggi pada suatu wilayah di kecamatan Kabupaten Ponorogo.

2. METODE PENELITIAN

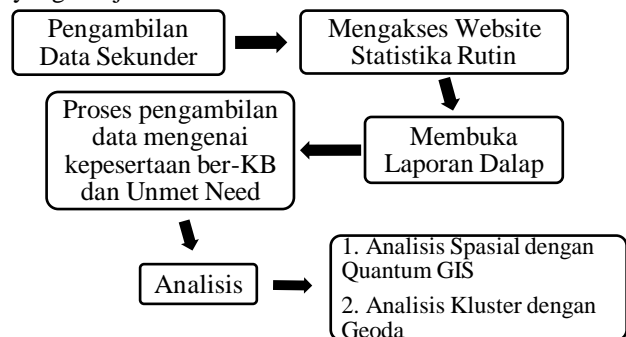
Penelitian ini merupakan jenis penelitian deskriptif dimana penelitian ini akan menggambarkan objek penelitian. Objek dalam penelitian ini adalah cakupan pengguna KB Aktif dan *unmet need* di Kabupaten Ponorogo pada tahun 2021. Data yang digunakan dalam penelitian merupakan jenis data sekunder dari Laporan Statistika Rutin. Laporan ini menyediakan informasi mengenai pencapaian

program KB secara nasional. Pencatatan dan pelaporan data dalam laporan ini rutin dilakukan setiap tahun dan dipublikasikan setiap bulan dan setiap tahun. Laporan Statistika Rutin tersedia dalam aplikasi berbasis web yang dapat diakses melalui situs <https://aplikasi.bkkbn.go.id>.

Laporan yang terdapat di dalam Laporan Statistika Rutin terbagi menjadi 2 yaitu Laporan Pengendalian Lapangan (Laporan Dalap) dan Laporan Pelayanan Kontrasepsi (Laporan Pelkon). Laporan Dalap merupakan laporan hasil pencatatan dan pelaporan kegiatan koordinasi serta pengendalian pelaksanaan Program KB di lapangan. Laporan Dalap merupakan laporan yang alur penginputan datanya dilakukan oleh penyuluh keluarga berencana (PKB) pada setiap kecamatan di wilayah Kabupaten Ponorogo. Sedangkan Laporan Pelkon merupakan laporan hasil pencatatan dan pelaporan yang berkaitan dengan pelayanan kontrasepsi dan dilakukan oleh fasilitas kesehatan pada setiap wilayah seperti klinik KB, dan dokter/bidan praktek swasta. Laporan Pelkon merupakan laporan yang penginputan datanya dilakukan oleh setiap fasilitas kesehatan yang terdapat di setiap wilayah di Kabupaten Ponorogo. Data yang akan digunakan adalah data dalam Laporan Dalap yaitu mengenai pengguna KB Aktif berdasarkan jenis metode kontrasepsi dan data mengenai pasangan usia subur bukan pengguna KB (BKKBN, 2021).

Analisis data dilakukan dengan menggunakan bantuan *software* Quantum GIS dan Geoda. Metode pengukuran pada penelitian ini berupa pengklasifikasian warna pada peta. Pengklasifikasian warna untuk pemetaan cakupan pengguna KB Aktif dan prevalensi *unmet need* terbagi menjadi 2 yaitu warna merah untuk wilayah yang tidak berhasil memenuhi standar capaian kabupaten dan warna kuning untuk wilayah yang berhasil memenuhi standar capaian kabupaten. Sedangkan untuk pemetaan hasil pengklusteran akan terbagi menjadi 4 yaitu warna kuning untuk kluster 1, warna biru untuk kluster 2, warna hijau untuk kluster 3, dan warna merah untuk kluster 4.

Berikut merupakan ringkasan alur penelitian yang disajikan dalam bentuk flowchart:



Gambar 3. Flowchart Alur Penelitian

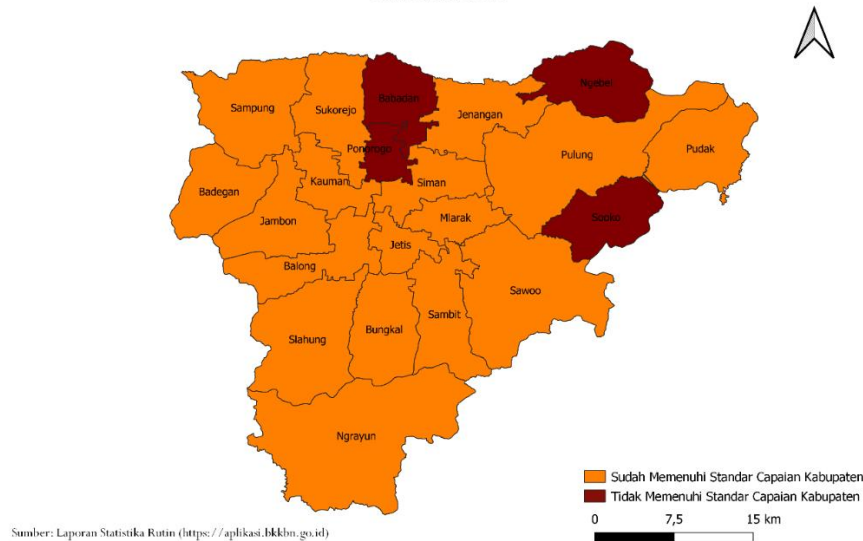
3. HASIL DAN PEMBAHASAN

Hasil

Data sekunder yang bersumber dari Laporan Statistika Rutin selanjutnya akan dilakukan pengolahan data sekaligus analisis data. Pengolahan data dilakukan dengan menggunakan software

Quantum GIS dan Geoda *filtering* dan *cleaning* data sehingga menghasilkan data yang sesuai dengan kebutuhan peneliti. Berikut merupakan data yang telah dilakukan proses *filter* dan *Cleaning* (Tabel 1):

PEMETAAN CAPAIAN CAKUPAN KB AKTIF TIAP KECAMATAN DI KABUPATEN PONOROGO
Bulan Desember 2021



Gambar 4. Pemetaan Cakupan KB Aktif

1. Pemetaan Capaian Cakupan KB Aktif

Capaian Cakupan KB Aktif Kabupaten Ponorogo pada bulan Desember 2021 mencapai 64,84%. Dari hasil pemetaan yang telah dilakukan, didapatkan hasil bahwa terdapat 4 kecamatan yang tidak memenuhi standar capaian kabupaten yaitu memiliki cakupan KB Aktif dibawah 64,84%. 4 kecamatan tersebut yaitu Kecamatan Ngebel, Babadan, Ponorogo, dan Kecamatan Sooko. Sedangkan 17 kecamatan yang lain yaitu Kec. Badegan, Kec. Bungal, Balong, Jambon, Jetis, Jenangan, Mlarak, Kauman, Ngrayun, Pudak, Pulung, Sambit, Sampung, Sawoo, Siman, Slahung, dan Sukorejo sudah memenuhi standar capaian kabupaten yaitu memiliki capaian cakupan KB Aktif diatas 64,84%.

Pengguna KB Aktif yang terdapat di Kabupaten Ponorogo menggunakan jenis metode kontrasepsi yang beragam. Jenis metode kontrasepsi jangka pendek (non MKJP) yang dapat digunakan diantaranya adalah pil, kondom, dan suntik. Sedangkan jenis kontrasepsi yang tergolong MKJP

yang dapat digunakan diantaranya adalah IUD, MOW, MOP dan implan.

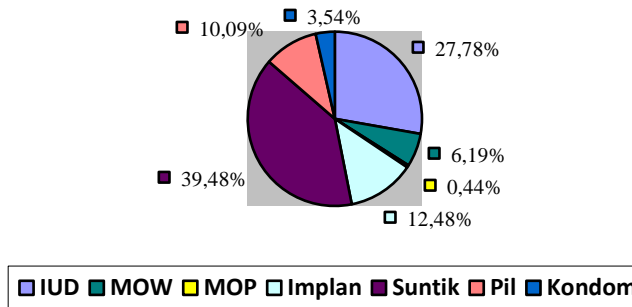
Berdasarkan grafik pengguna KB Aktif per metode kontrasepsi (Gambar 5), jenis alkon yang paling banyak dipakai dan diminati Kabupaten Ponorogo adalah jenis alkon suntik dengan prevalensi pengguna jenis metode ini sebanyak 39,48%. Sedangkan jenis metode kontrasepsi yang masih jarang dipakai adalah MOP dengan prevalensi 0,44% dan kondom dengan prevalensi 3,54%. Sedangkan untuk metode kontrasepsi yang lain yaitu MOW dengan prevalensi 6,19%, pil dengan prevalensi 10,09%, implan dengan prevalensi 12,48%, dan IUD dengan prevalensi 27,78%.

Tabel 1. Data Hasil *Filtering* dan *Cleaning*

Kecamatan	Jumlah PUS	KB Aktif	% KB Aktif	Unmet Need			% Unmet Need	IUD	MOW	MOP	IMPLAN	SUNTIK	PIL	KONDOM	%MKJP	%Non MKJP
				Ingin anak ditunda	Tidak ingin anak lagi	Total										
Babadan	10.797	6163	57,08	538	354	892	8,26	1854	352	8	258	2846	556	289	40,11	59,89
Badegan	5.921	3932	66,41	582	632	1.214	20,50	649	218	6	708	2028	240	83	40,21	59,79
Balong	8.363	5570	66,60	468	849	1.317	15,75	1994	143	36	641	2317	346	93	50,52	49,48
Bungkal	7.107	4767	67,07	313	388	701	9,86	1850	315	19	821	1188	314	260	63,04	36,96
Jambon	9.113	6157	67,56	808	701	1.509	16,56	1032	196	39	958	3262	494	176	36,14	63,86
Jenangan	11.102	7271	65,49	878	1.126	2.004	18,05	812	542	13	632	3362	1572	338	27,49	72,51
Jetis	5.792	3878	66,95	373	359	732	12,64	440	286	5	314	1814	840	179	26,95	73,05
Kauman	7.515	4938	65,71	661	751	1.412	18,79	1923	317	9	552	1583	376	178	56,72	43,28
Mlarak	5.579	3758	67,36	265	254	519	9,30	998	177	7	403	1803	246	124	42,18	57,82
Ngebel	3.761	2366	62,91	437	408	845	22,47	454	146	20	592	585	448	121	51,23	48,77
Ngrayun	10.426	6937	66,54	531	522	1.053	10,10	1349	515	69	265	3679	1020	40	31,69	68,31
Ponorogo	10.720	4773	44,52	788	3.546	4.334	40,43	2148	599	10	247	1297	300	172	62,94	37,06
Pudak	2.115	1450	68,56	76	321	397	18,77	317	109	18	497	434	75	0	64,90	35,10
Pulung	11.451	7841	68,47	506	579	1.085	9,48	1873	660	81	1492	2903	550	282	52,37	47,63
Sambit	8.860	5974	67,43	635	496	1.131	12,77	1535	337	3	766	1896	1061	376	44,21	55,79
Sampung	7.085	4744	66,96	348	257	605	8,54	1454	272	3	384	2161	380	90	44,54	55,46
Sawoo	11.245	7455	66,30	597	680	1.277	11,36	2466	498	20	813	2729	665	264	50,93	49,07
Siman	8.691	6032	69,41	217	244	461	5,30	2988	338	19	580	1543	357	207	65,07	34,93
Slahung	10.484	7105	67,77	560	485	1.045	9,97	2295	375	22	1152	2581	544	136	54,10	45,90
Sooko	4.735	3068	64,79	250	156	406	8,57	375	172	71	812	1418	162	58	46,61	53,39
Sukorejo	10.099	6780	67,14	994	1.087	2.081	20,61	2022	296	10	961	2382	647	462	48,51	51,49

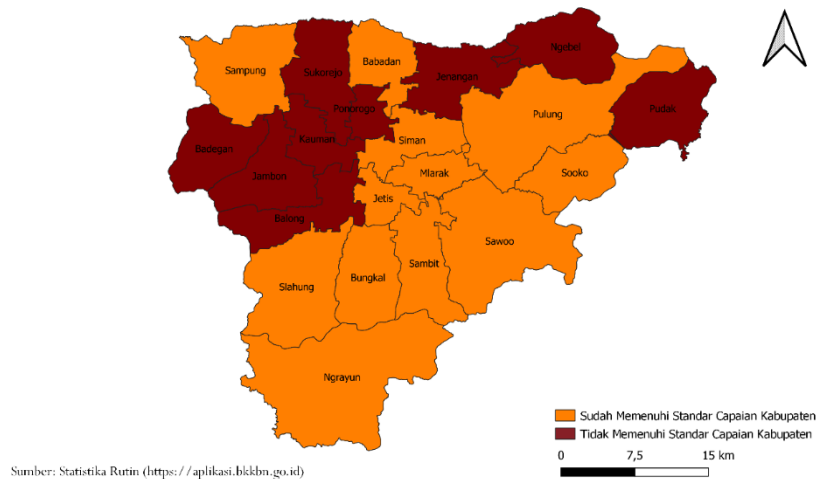
Data yang telah diolah selanjutnya dilakukan analisis spasial dan analisis kluster menggunakan *software* Quantum GIS dan Geoda dengan hasil sebagai berikut.

Presentase Penggunaan Alat/Obat Kontrasepsi



Gambar 5. Presentase Pengguna KB Aktif berdasar Alat/Obat kontrasepsi

2. Pemetaan Prevalensi *Unmet Need*



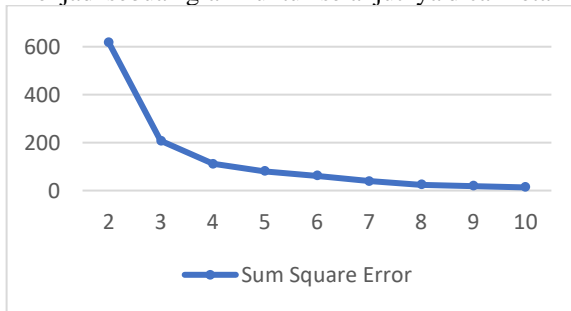
Gambar 6. Pemetaan Prevalensi *Unmet Need* Tiap Kecamatan di Kabupaten Ponorogo

Capaian prevalensi *unmet need* di Kabupaten Ponorogo pada bulan Desember 2021 mencapai 14,60%. Dari hasil pemetaan yang telah dilakukan, didapatkan hasil bahwa terdapat 9 kecamatan yang tidak memenuhi standar capaian kabupaten yaitu memiliki prevalensi *unmet need* diatas 14,60% yaitu Kecamatan Pudak, Ponorogo, Ngebel , Kauman Jenangan, Sukorejo, , Jambon, Balong, dan Badegan. Sedangkan 12 kecamatan yang lain yaitu Kec. Babadan, Sawoo Pulung Siman, Jetis, , Sambit, Bungkal, Slahung, Ngrayun, ,Sooko, Mlarak dan Kecamatan Sampung sudah memenuhi standar capaian kabupaten yaitu memiliki capaian prevalensi *unmet need* dibawah capaian Kabupaten Ponorogo.

3. Analisis Kluster

Analisis kluster dilakukan menggunakan bantuan software Geoda. Analisis kluster menggunakan metode K-Means yang termasuk kedalam metode pengklusteran *non-hirarki*. Metode ini dipilih dikarenakan peneliti ingin membuat pengelompokan dengan membentuk kluster sesuai dengan jumlah kluster yang telah ditentukan dan jumlah kluster yang ideal. Sehingga dalam proses pengklusterannya nanti, jumlah kluster sudah ditentukan diawal. Penentuan jumlah kluster dilakukan menggunakan metode Elbow. Penentuan jumlah kluster dengan metode ini dilakukan dengan menghitung nilai *Sum Square*

Error (SSE) setiap kluster sesuai dengan data yang ada. Nilai SSE setiap kluster selanjutnya dibuat menjadi sebuah grafik untuk selanjutnya dicari letak



Gambar 7. Grafik *Sum of Square Error*

Berdasarkan grafik disamping, siku terbentuk antara kluster 3 dan kluster 4. Oleh karenanya, jumlah kluster yang ideal adalah 4 kluster. Setelah menemukan jumlah kluster yang ideal, selanjutnya

siku pada grafik tersebut berada (Fadilla et al., 2020). Berikut merupakan grafik SSE yang telah dibuat.

dilakukan analisis kluster menggunakan Geoda dengan hasil sebagai berikut.

Within-cluster sum of square	111,60
Between-cluster sum of square	1656,84

Tabel 2. Analisis Kluster dengan Geoda

Hasil analisis kluster dengan geoda menunjukkan bahwa nilai within-cluster sum of square sebesar 111,60 dan nilai between-cluster sum of square sebesar 1656,84. Dari hasil tersebut, dapat diketahui bahwa kluster yang terbentuk sudah ideal, dikarenakan nilai perbedaan antar kecamatan dalam 1 kluster kecil dan perbedaan antar kluster sangat besar. Berikut merupakan pemetaan hasil analisis kluster yang telah dilakukan.



Gambar 8. Pemetaan Hasil Pengklusteran Wilayah Berdasarkan Cakupan KB Aktif dan Prevalensi *Unmet Need*

Dari pemetaan diatas, terdapat 4 kluster yang terbentuk yaitu sebagai berikut.

1. Kluster 1

Jumlah wilayah yang masuk pada kluster 1 sebanyak 11 kecamatan yaitu Kec. Sampung, Kec. Jatis, Kec. Siman, Kec. Pulung, Kec. Sooko, Kec.

3. Kluster 3

Jumlah wilayah yang masuk pada kluster 3 sebanyak 8 yaitu Kecamatan Jenangan, Kauman, Jambon, Sukorejo, Badegan, Ngebel, Balong, Pudak

Bungkal, Kec. Ngrayun, Kec. Sambit, Kec. Mlarak, Kec. Slahung, Kec. Sawoo

2. Kluster 2

Jumlah wilayah yang masuk pada kluster 2 sebanyak 1 kecamatan yaitu Kecamatan Babadan

4. Kluster 4

Jumlah wilayah yang masuk pada kluster 4 sebanyak 1 kecamatan yaitu Kecamatan Ponorogo.

Pembahasan

Berdasarkan pemetaan yang telah dilakukan, cakupan KB Aktif di wilayah Kabupaten Ponorogo sudah merata (Gambar 4). Hal ini ditunjukkan dengan mayoritas kecamatan sudah memiliki capaian cakupan KB Aktif melebihi angka capaian Kabupaten Ponorogo. Meskipun mayoritas wilayah kecamatan sudah melaksanakan program KB dengan baik, namun masih terdapat beberapa wilayah kecamatan dimana program KB belum berjalan secara maksimal yang ditunjukkan dengan cakupan KB Aktif yang masih rendah. Terdapat 4 kecamatan yang perlu mendapat perhatian khusus dalam peningkatan pencapaian keberhasilan Program KB, dikarenakan pada 4 kecamatan ini masih memiliki capaian cakupan KB Aktif dibawah angka capaian Kabupaten Ponorogo. 4 Kecamatan ini adalah Kecamatan Ngebel, Kecamatan Babadan, Kecamatan Ponorogo, dan Kecamatan Sooko. Penyebab rendahnya cakupan pengguna KB Aktif dilatarbelakangi oleh banyak faktor, seperti tingkat ekonomi keluarga, tingkat pendidikan, dan budaya.

Tingkat ekonomi dapat mempengaruhi keinginan PUS untuk mengikuti program KB (Rahmayanti, 2015). Keluarga dengan tingkat ekonomi menengah kebawah akan lebih mengutamakan kebutuhan primer setiap harinya dibandingkan kebutuhan lainnya seperti kebutuhan untuk ber-KB. Keluarga dengan tingkat ekonomi menengah kebawah akan lebih cenderung untuk mengesampingkan penggunaan penghasilan dalam pemenuhan kebutuhan ber-KB. Hal ini akan menyebabkan banyaknya keluarga yang tidak berpartisipasi menjadi pengguna KB Aktif dikarenakan keterbatasan ekonomi.

Selain tingkat ekonomi, tingkat pendidikan juga dapat mempengaruhi keinginan PUS untuk ber-KB (Rahmayanti, 2015). Tingkat pendidikan ini berhubungan dengan bagaimana seorang individu lebih mudah memahami dan menerima adanya informasi. Tingkat pendidikan juga akan mempengaruhi pola pikir atau pandangan seorang individu, tidak terkecuali pandangan mengenai keikutsertaan dalam Program KB. PUS dengan tingkat pendidikan yang tinggi lebih cenderung mudah memahami makna dibalik pelaksanaan Program Keluarga Berencana yaitu untuk mengontrol jumlah penduduk dan mencapai kesejahteraan keluarga, namun PUS dengan tingkat pendidikan rendah cenderung tidak menerima dan memahami makna pelaksanaan Program KB sehingga lebih berisiko untuk tidak berpartisipasi dalam penggunaan alat kontrasepsi.

Faktor lain yang berpengaruh terhadap keinginan PUS untuk ber-KB adalah budaya yang terdapat di wilayah lingkungan mereka (Wilisandi & Feriani, 2020). Seringkali anggapan yang bertentangan dengan tujuan Program Keluarga Berencana seperti misalnya “banyak anak banyak rezeki” menjadi penyebab PUS tidak ingin ber-KB. Selain itu, pengaruh dari lingkungan sekitar juga akan mempengaruhi keikutsertaan PUS dalam ber-KB. Pengaruh tetangga yang memiliki dukungan negatif terhadap program KB dapat menyebabkan PUS menjadi tidak ikut serta dalam program KB dan begitupun sebaliknya. Apabila tetangga memiliki dukungan positif terhadap pelaksanaan program KB, maka akan semakin besar pula keinginan PUS untuk ikut serta dalam program KB.

Jenis metode kontrasepsi yang digunakan oleh masyarakat Kabupaten Ponorogo diantaranya adalah MKJP yang meliputi IUD, MOW, MOP, implan dan metode kontrasepsi jangka pendek meliputi kondom, pil, dan suntik. Jenis kontrasepsi yang paling diminati adalah jenis metode kontrasepsi suntik dengan persentase sebesar 39,48% yang artinya sebanyak 39,48% pengguna KB Aktif di Kabupaten Ponorogo menggunakan jenis KB suntik (Gambar 5). Jenis KB suntik merupakan jenis metode kontrasepsi jangka pendek yaitu jenis metode kontrasepsi yang kegunaannya hanya bersifat sementara atau tidak permanen. Jenis metode kontrasepsi ini memiliki tingkat efektivitas yang lebih rendah dibandingkan dengan MKJP sehingga jenis ini memiliki resiko yang tinggi terhadap terjadinya *drop out*. Hal ini tentunya harus dihindari, dikarenakan apabila angka *drop out* tinggi maka akan berisiko terjadinya kelahiran yang akan meningkatkan kepadatan penduduk di suatu wilayah. Selain itu, angka *drop out* yang tinggi juga dapat menyebabkan capaian KB Aktif suatu wilayah menjadi menurun.

Sedangkan untuk jenis metode kontrasepsi MOP dan kondom menjadi jenis KB yang kurang diminati oleh masyarakat Kabupaten Ponorogo. Metode kontrasepsi yang ditujukan untuk PUS berjenis kelamin laki-laki ini masih sangat rendah penggunaannya. Prevalensi pengguna jenis metode kontrasepsi MOP dan Kondom yaitu 0,44% dan 3,54% (Gambar 5). Rendahnya pengguna jenis metode kontrasepsi ini dapat mengindikasikan bahwa masih sedikit PUS laki-laki yang berkontribusi dalam pelaksanaan Program KB dan menggunakan alat kontrasepsi. Faktor yang mempengaruhi keikutsertaan laki-laki dalam ber-KB diantaranya adalah tingkat pengetahuan tentang KB yang masih kurang, tingkat pendapatan rendah,

persepsi atau pandangan yang positif terhadap nilai anak seperti menganggap bahwa anak itu adalah berkah dsb, serta sikap negatif terhadap alat kontrasepsi misalnya takut akan resiko atau efek samping yang akan timbul akibat pemakaian (Tamtomo, 2015).

Selain cakupan KB Aktif, keberhasilan program KB juga ditunjukkan dengan prevalensi *unmet need* yang rendah. Berdasarkan hasil pemetaan yang telah dilakukan, terdapat 14 kecamatan yang memiliki prevalensi *unmet need* dibawah capaian Kabupaten Ponorogo dan terdapat 9 kecamatan yang memiliki prevalensi *unmet need* diatas capaian Kabupaten Ponorogo yaitu Kecamatan Ngebel, Kecamatan Jenangan, Kecamatan Balong, Kecamatan Ponorogo, Kecamatan Sukorejo, Kecamatan Kauman, Kecamatan Puduk, Kecamatan Jambon, dan Kecamatan Badegan (Gambar 6). *Unmet need* menjadi salah satu penyebab terjadinya kehamilan tidak diinginkan (Ratnaningsih, 2018). Kehamilan tidak diinginkan akan berpotensi untuk mengalami berbagai masalah kesehatan dan berdampak pada kesejahteraan keluarga. *Unmet need* dapat terjadi karena dilatarbelakangi oleh usia, dukungan suami dan pendidikan. PUS dengan kategori usia tidak berisiko (usia 20-35 tahun) apabila mengalami kehamilan cenderung tidak menggunakan KB dikarenakan pada usia ini mereka masih menginginkan anak lagi. PUS pada usia ini termasuk dalam kategori ingin memiliki anak namun ditunda, oleh karenanya mereka cenderung tidak menggunakan KB dikarenakan beberapa hal seperti misalnya takut apabila kesuburannya terganggu ketika menggunakan KB dan lain sebagainya. Sedangkan pada PUS dengan usia berisiko yaitu PUS usia <20 tahun atau >35 tahun cenderung belum ingin memiliki kehamilan atau tidak menginginkan adanya kehamilan. PUS pada usia muda memilih untuk tidak menggunakan KB dikarenakan mereka beranggapan bahwa menggunakan KB pada usia muda dapat mempengaruhi tingkat kesuburannya kelak. PUS pada usia >35 tahun memilih tidak menggunakan KB dikarenakan mereka beranggapan bahwa pada usia ini mereka tidak mungkin mengalami kehamilan. Berbagai anggapan yang muncul ini terjadi akibat masih minimnya informasi yang diterima oleh masyarakat mengenai Program Keluarga Berencana. Oleh karenanya, tingkat pendidikan dan tingkat pengetahuan yang rendah juga menjadi faktor penyebab terjadinya *unmet need*. *Unmet need* juga dapat dipengaruhi oleh dukungan suami terhadap pelaksanaan Program Keluarga Berencana. Dukungan suami dapat berupa

pemberian motivasi, edukasi dan memfasilitasi istri untuk ber-KB.

Pemetaan hasil pengklusteran berdasarkan cakupan KB Aktif dan prevalensi *unmet need* menunjukkan bahwa Kabupaten Ponorogo terbagi menjadi 4 kluster (Gambar 8). Kluster 1 memiliki rata-rata cakupan KB Aktif sebesar 67,19% dan *unmet need* sebesar 9,80%. Sehingga kecamatan yang termasuk dalam kluster 1 merupakan kecamatan yang sudah memenuhi standar capaian kabupaten baik untuk cakupan KB Aktif maupun untuk prevalensi *unmet need*. Kluster 2 memiliki rata-rata cakupan KB Aktif sebesar 57,08% dan prevalensi *unmet need* sebesar 8,26%. Kecamatan pada kluster ini merupakan kecamatan yang sudah memenuhi mencapai standar capaian *unmet need* kabupaten namun tidak memenuhi standar capaian cakupan KB Aktif kabupaten. Kluster 3 memiliki rata-rata cakupan KB Aktif sebesar 66,30% dan prevalensi *unmet need* sebesar 18,93%. Pada kluster 3 terdapat kecamatan yang rata-rata memiliki cakupan KB Aktif diatas capaian kabupaten namun tidak memenuhi standar capaian *unmet need* kabupaten. Sedangkan kluster 4 memiliki rata-rata cakupan KB Aktif sebesar 44,52% dan prevalensi *unmet need* sebesar 40,43. Kecamatan pada kluster ini merupakan kecamatan yang tidak memenuhi standar capaian kabupaten baik dalam cakupan KB Aktif maupun dalam prevalensi *unmet need*.

Berdasarkan hasil pemetaan mengenai cakupan KB Aktif dan *unmet need* yang telah dilakukan, terdapat 2 kecamatan yang berpotensi akan mengalami peningkatan jumlah penduduk pada tahun yang akan mendatang dan akan mempengaruhi kepadatan penduduk di wilayah tersebut. 2 kecamatan tersebut yaitu Kecamatan Ngebel dan Kecamatan Ponorogo. Kedua kecamatan ini memiliki capaian cakupan KB Aktif dibawah capaian kabupaten dan memiliki prevalensi *unmet need* diatas prevalensi kabupaten. Beberapa upaya dapat dilakukan untuk meningkatkan cakupan pengguna KB Aktif pada wilayah kecamatan yang masih memiliki cakupan dibawah capaian cakupan Kabupaten. Upaya yang dapat dilakukan diantaranya adalah dengan pemberian bantuan dana bagi keluarga yang merasa kesulitan mendapat alat kontrasepsi dikarenakan keterbatasan ekonomi, pemberian wawasan mengenai alat kontrasepsi pada PUS, menjalin hubungan berupa sebuah kerjasama dengan pemangku kepentingan di wilayah setempat, melakukan pendekatan dan advokasi kepada tokoh masyarakat (Sukmaniar et al., 2018). Sedangkan upaya yang dapat dilakukan untuk menurunkan prevalensi *unmet need* pada suatu wilayah adalah

dengan memaksimalkan peran aktif petugas Penyuluh Keluarga Berencana (PKB) maupun tenaga KB lain yang ada di suatu wilayah dalam memperkenalkan program KB kepada masyarakat. Selain itu, dapat pula dilakukan penyuluhan kepada masyarakat mengenai pentingnya untuk ikut serta dalam Program Keluarga Berencana.

4. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian dan pembahasan diatas, dapat disimpulkan bahwa:

1. Hasil pemetaan mengenai cakupan pengguna KB Aktif menunjukkan bahwa terdapat 4 kecamatan yang memiliki capaian cakupan pengguna KB Aktif dibawah capaian Kabupaten Ponorogo yaitu Kecamatan Ngebel, Kecamatan Babadan, Kecamatan Ponorogo, dan Kecamatan Sooko
2. Jenis metode kontrasepsi yang paling diminati PUS di Kabupaten Ponorogo adalah suntik, sedangkan jenis metode kontrasepsi kondom, MOP, dan MOW masih kurang diminati oleh PUS di Kabupaten Ponorogo
3. Hasil pemetaan mengenai prevalensi *unmet need* menunjukkan bahwa terdapat 9 kecamatan yang memiliki prevalensi *unmet need* diatas capaian Kabupaten Ponorogo yaitu Kec. Badegan, Kec. Pudak, Kec. Jenangan, Kec. Ngebel, Kec. Jambon, Kec. Ponorogo, Kec. Kauman, , Kec. Balong, dan Kec. Sukorejo
4. Pemetaan pengklusteran wilayah menghasilkan Kabupaten Ponorogo terbagi menjadi 4 kluster. Kluster 1 berisi Kec. Jetis, Kec. Siman, Kec. Pulung, Kec. Bungkal, Kec. Sawoo, Kec. Ngrayun, Kec. Sambit, Kec. Mlarak, Kec. Slahung, Kec. Sooko, Kec. Sampung. Kluster 2 terdiri dari 1 kecamatan yaitu Kecamatan Babadan. Kluster 3 terdiri dari Kec. Jambon, Kec. Jenangan, Kec. Kauman, Kec. Balong, Kec. Ngebel, Kec. Badegan, Kec. Pudak, dan Kec. Sukorejo. Kluster 4 terdiri dari Kecamatan Ponorogo

Saran

Penelitian ini mengolah data menggunakan aplikasi Quantum GIS dan Geoda, diharapkan peneliti selanjutnya dapat mencoba menggunakan jenis software lainnya untuk mengolah data penelitian. Selain itu, metode pengklusteran terdiri dari banyak metode. Diharapkan peneliti selanjutnya dapat menggunakan jenis metode pengklusteran

yang lain agar dapat membandingkan hasilnya dan diperoleh metode pengklusteran terbaik.

DAFTAR PUSTAKA

- BPS. (2020). *Hasil Sensus Penduduk 2020*. Retrieved March 21, 2022, from <https://www.bps.go.id/pressrelease/2021/01/21/1854/hasil-sensus-penduduk-2020.html>
- BPS. (2021). *Laju Pertumbuhan Penduduk (Persen)*. <https://www.bps.go.id/indicator/12/1976/1/laju-pertumbuhan-penduduk.html>
- BPS. (2022). *Sistem Informasi Rujukan Statistik*. <https://sirusa.bps.go.id/sirusa/index.php/indikator/49>
- Dirjendukcapil. (2021). *Distribusi Penduduk Indonesia Per Juni 2021: Jabar Terbanyak, Kaltara Paling Sedikit*. <https://dukcapil.kemendagri.go.id/berita/baca/809/distribusi-penduduk-indonesia-per-juni-2021-jabar-terbanyak-kaltara-paling-sedikit>
- Fadilla, R., Andarsyah, R., & Awangga, R. M. (2020). *Data analytics: Peningkatan Performa Algoritma Rekomendasi Collaborative Filtering Menggunakan K-Means Clustering* (R. M. Awanga (ed.)). Kreatif Industri Nusantara. https://www.google.co.id/books/edition/Data_analytics_peningkatan_performa_algo/JgkCEAAAQBAJ?hl=en&gbpv=1&dq=metode+elbow&pg=PA91&printsec=frontcover
- Rahmayanti, O. M. (2015). *Hubungan Antara Tingkat Pendidikan dan Pendapatan dengan Tingkat Partisipasi PUS dalam Pelaksanaan Program Keluarga Berencana di Desa Menganti Kecamatan Kesugihan Kabupaten Cilacap Tahun 2015* [Universitas Negeri Semarang]. <http://lib.unnes.ac.id/23279/1/3201411034.pdf>
- Ratnaningsih, E. (2018). Analisis Dampak Unmet Need Keluarga Berencana Terhadap Kehamilan Tidak Diinginkan Di Rumah Sakit Panti Wilasa Citarum Semarang. *Jurnal Kebidanan*, 7(2), 80–94.
- Sukmaniar, Saputra, W., & Saputra, D. (2018). *Upaya Peningkatan Peserta KB Aktif dalam Rangka Pencapaian Target Renstra BKKBN 2015-2019 di Kabupaten Musi Rawas, Provinsi Sumatera Selatan*. 26, 39–50.
- Tamtomo, A. D. (2015). *KAJIAN RENDAHNYA PARTISIPASI SUAMI (PUS) SEBAGAI AKSEPTOR KB DI DESA TANJUNG SARI (JURNAL) Oleh Agung Dwi Tamtomo 2015*

Mahasiswa pendidikan Geografi Dosen Pembimbing 1 Dosen Pembimbing 2. Universitas Lampung.

- Wilisandi, W., & Feriani, P. (2020). Hubungan Faktor Budaya dengan Perilaku Penggunaan Alat Kontrasepsi (KB) di Puskesmas Samarinda Kota. *Journal Borneo Student Research*, Vol.2 No.1(1), 8. <https://journals.umkt.ac.id/index.php/bsr/article/download/1491/669>

Medoid-based Clustering pada Kecamatan di Kabupaten Lebak dan Pandeglang Provinsi Banten Berdasarkan Trilogi Ketahanan Pangan

Weksi Budiaji^(1*,2), Juwarin Pancawati⁽¹⁾

¹Fakultas Pertanian, Universitas Sultan Ageng Tirtayasa

Jl. Palka Km 3, Sindang Sari, Pabuaran, Serang, Banten 42163

²Pusat Unggulan Ipteks Inovasi Pangan Lokal, Universitas Sultan Ageng Tirtayasa

Jl. Palka Km 3, Sindang Sari, Pabuaran, Serang, Banten 42163

e-mail: budiaji@untirta.ac.id (*penulis korespondensi), juwarinpancawati@untirta.ac.id

ABSTRAK

Kabupaten Lebak dan Pandeglang Provinsi Banten memiliki prevalensi status gizi buruk pada balita yang tinggi dan berada pada ketahanan pangan yang paling rendah diantara seluruh kabupaten di Provinsi Banten. Analisis kluster untuk mengelompokkan kecamatan di kedua kabupaten sangat penting untuk mengeksplorasi karakteristik kecamatan di Kabupaten Lebak dan Pandeglang. Data indikator trilogi ketahanan pangan yaitu ketersediaan bahan pangan, akses ke bahan pangan, dan pemanfaatan bahan pangan diambil dari data sekunder Badan Pusat Statistik Kabupaten Lebak dan Pandeglang tahun 2019. Ketiga variabel tersebut digunakan untuk menghitung jarak pada algoritma *simple k-medoids clustering*. Perhitungan jarak divariasikan dengan konsep jarak Euclidean, squared Euclidean, dan Manhattan. Hasil kluster kemudian divalidasi dengan consensus clustering, dan validasi internal. Jumlah kluster yang paling sesuai dengan mempertimbangkan hasil validasi adalah empat kluster yaitu kluster ketersediaan dan akses (kluster 1), kluster akses (kluster 2), kluster rawan (kluster 3), dan kluster ketersediaan (kluster 4). Kluster 3 sebagai kluster rawan pangan perlu diperhatikan karena tersusun sebanyak 38% dari total kecamatan di Kabupaten Lebak dan Pandeglang.

Kata kunci : Ketahanan pangan, Kluster, Jarak, Medoid, Consensus, Validasi.

ABSTRACT

Lebak and Pandeglang Regions in Banten Province have a high stunting prevalence of children under 5 years old and have the lowest value of food security index among regions in Banten Province. Cluster analysis to group districts in Lebak and Pandeglang Regions is indispensable to characterize the district members in those two regions. The trilogy data of food security indicator, i.e., the availability, access, and utility of the food, were obtained from Bureau of Statistics of Lebak and Pandeglang Regions 2019 secondary data. The trilogy data were supplied for distance calculation in the simple k-medoids clustering. The distances were varied among Euclidean, squared Euclidean, and Manhattan distances. The clustering result was then validated via consensus clustering and internal validation. The suitable number of clusters w.r.t validation results was four clusters defined as the available and access cluster (cluster 1), the access cluster (cluster 2), the vulnerable cluster (cluster 3), and the available cluster (cluster 4). The cluster 3 as the vulnerable cluster should be focused on because it consists of 38% from overall districts in Lebak and Banten Regions.

Keywords : Food security, Clustering, Distance, Medoid, Consensus, Validation.

1. PENDAHULUAN

Analisis kluster yang bertujuan mengelompokkan objek sejenis pada satu kelompok yang sama merupakan analisis eksploratif. Analisis eksploratif sangat bermanfaat untuk mengidentifikasi struktur data. Salah satu analisis kluster yang populer dan juga salah satu dari top 10 algoritma pada *data mining* adalah k-means (Wu et

al., 2008). Berbeda dengan k-means yang menggunakan rata-rata sebagai pusat dari kluster, k-medoids menggunakan salah satu objek sebagai pusat kluster. Simple k-medoids (SKM) telah dikembangkan dengan mengadopsi k-means dan mempertimbangkan kelemahan yang dimiliki k-means yaitu adanya kluster kosong dan *local optima* (Budiaji & Leisch, 2019).

K-means dan SKM merupakan analisis kluster dengan metode algoritma partisi yang menggunakan konsep jarak sebagai ukuran kemiripan objek. Proses partisi memungkinkan setiap objek berpindah dari kluster satu ke kluster yang lain. Hasil analisis kluster dengan metode algoritma partisi sangat besar dipengaruhi oleh pemilihan konsep jarak yang digunakan (Leisch, 2006), sehingga penggunaan jarak yang beragam sangat penting, seperti Pamulang et al. (2021) dan Aditya et al. (2021) yang menggunakan variasi jarak pada algoritma partisi.

Artikel ini membahas penggunaan berbagai variasi jarak dengan algoritma SKM pada variabel trilogi ketahanan pangan. Trilogi ketahanan pangan yang terdiri atas ketersediaan bahan pangan, akses ke bahan pangan, dan pemanfaatan bahan pangan (Simatupang, 2007). Ketiga variabel ini digunakan untuk mengukur ketahanan pangan di Jawa Tengah (Pujiati et al. 2020) dan bahkan kabupaten-kabupaten di Indonesia (Mun'im, 2012; FSC & WFP, 2015).

Studi FSC & WFP (2015) menyebutkan bahwa Kabupaten Lebak dan Pandeglang memiliki ketahanan pangan yang paling rendah diantara seluruh kabupaten di Provinsi Banten, sehingga kedua kabupaten ini menjadi fokus perhatian di artikel ini. Kedua kabupaten memiliki persentase prevalensi status gizi (berat badan/ umur) buruk pada balita yang tinggi (4.8 % dan 6.98%) melebihi rata-rata Provinsi Banten (3.65%) (Balitbangkes, 2019). Belum ada analisis ketahanan pangan pada level kecamatan di kedua kabupaten juga menjadi alasan pentingnya penyusunan gerombol pada tingkat kecamatan.

2. METODE PENELITIAN

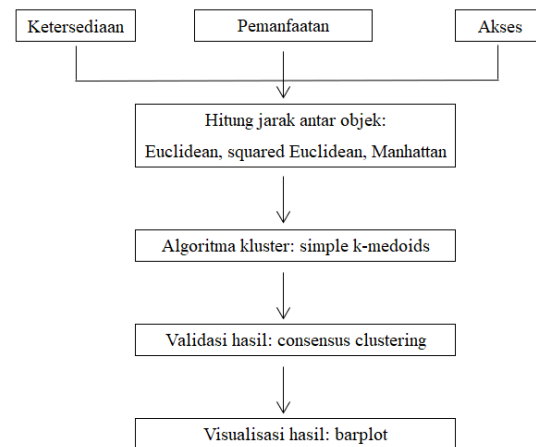
Medoid merupakan objek yang menjadi pusat kluster. Pengembangan simple k-medoid (SKM) yang mengadopsi k-means dengan mempertimbangkan kluster kosong dan local optima telah dilakukan (Budiaji & Leisch, 2019). Algoritma tersebut juga diimplementasikan pada paket R dengan nama paket *kmed* (Budiaji, 2021).

Algoritma SKM diterapkan pada data ketersediaan, akses, dan pemanfaatan bahan pangan. Data yang digunakan adalah data sekunder dari BPS Lebak (2019) dan BPS Pandeglang (2019). Data ketersediaan pangan menggunakan indikator rasio produksi sereal/ia tiap kecamatan dengan konsumsi normatif (FSC & WFP, 2015). Akses bahan pangan merupakan jumlah rumah tangga pengguna/pelanggan air bersih per 10000 penduduk, sedangkan pemanfaatan bahan pangan merasiokan jumlah fasilitas kesehatan tiap kecamatan per kapita.

Variabel trilogi ketahanan pangan yaitu ketersediaan, akses, dan pemanfaatan digunakan sebagai dasar perhitungan jarak tiap kecamatan. SKM sebagai metode algoritma partisi sangat dipengaruhi oleh jarak yang digunakan untuk mengukur kemiripan (Leisch, 2006). Jarak yang digunakan adalah *Euclidean*, *squared Euclidean*, dan *Manhattan*.

Hasil clustering kemudian divalidasi dengan salah satu metode kriteria relatif yaitu *consensus clustering* (Monti et al., 2003). Metode ini menggunakan dasar teknik *re-sampling* yang hasilnya dapat menggambarkan kestabilan anggota kluster. Validasi dengan metode kriteria relatif dapat divisualkan dalam *heatmap* dengan gambar blok diagonal mengindikasikan kluster yang dihasilkan stabil.

Hasil kluster yang paling sesuai dengan kondisi kecamatan di Kabupaten Lebak dan Pandeglang dideskripsikan dengan barplot. Barplot ditandai nilai tengah per kluster untuk menginterpretasikan hasil kluster (Dolnicar & Leisch, 2014). Gambar 1 menunjukkan proses analisis data trilogi ketahanan pangan sampai dengan visualisasi. Proses ini dibantu dengan software R (R Core Team, 2020), paket *kmed* (Budiaji, 2021), dan *rgl* (Murdoch & Adler, 2021).



Gambar 1. Langkah-langkah analisis data sampai hasil akhir

3. HASIL DAN PEMBAHASAN

3.1 Deskripsi variabel trilogi ketahanan pangan

Ketersediaan bahan pangan diukur dengan produksi sereal/ia tiap kecamatan tahun 2018 (BPS Lebak, 2019; BPS Pandeglang, 2019) yang terdiri atas produksi padi dan jagung dan konsumsi normatif. Konsumsi normatif sereal/ia per hari per kapita tiap kecamatan diperoleh dari perkalian

jumlah pendudukan dengan 0.3 (FSC & WFP, 2015). Indeks ketersediaan bahan pangan kemudian merupakan rasio produksi sereal dengan konsumsi normatif. Berbeda dengan indeks ketersediaan FSC & WFP (2015), indeks ketersediaan yang dihitung ini jika nilai lebih dari 1 mengindikasikan kelebihan bahan pangan, sedangkan nilai kurang dari 1 menunjukkan kekurangan bahan pangan.

Tabel 1 menunjukkan bahwa kecamatan di Kabupaten Lebak mayoritas kekurangan bahan pangan. Kondisi ini bertolak belakang dengan kecamatan di Kabupaten Pandeglang. Rataan nilai indeks kelebihan pangan di Kabupaten Pandeglang walaupun lebih kecil yaitu sebesar 2.79 dibandingkan Kabupaten Lebak (3.52), rentang/jangkauan nilai kelebihan pangan di Kabupaten Lebak lebih tinggi.

Tabel 1. Ringkasan ketersediaan bahan pangan

Kabupaten	Status Ketersediaan Bahan Pangan *	Jumlah Kecamatan	Nilai Indeks		
			Min	Rataan	Mak
Lebak	Kekurangan (indeks < 1)	24	0	0.20	0.77
	Kelebihan (indeks > 1)	4	1.44	3.52	6.05
Pandeglang	Kekurangan (indeks < 1)	3	0.50	0.68	0.82
	Kelebihan (indeks > 1)	32	1.18	2.79	5.34
Total		63			

* Kelebihan bahan pangan ditandai dengan nilai indeks yang lebih dari 1

Jumlah rumah tangga pengguna/ pelanggan air bersih per 10000 jumlah penduduk merupakan indikator dari akses bahan pangan. Semakin tinggi indeks rumah tangga pengguna/ pelanggan air bersih mengindikasikan semakin mudah akses ke bahan pangan. Tabel 2 menunjukkan bahwa kedua kabupaten memiliki karakter yang mirip dalam hal akses ke bahan pangan. Setengah dari keseluruhan kecamatan di kedua kabupaten bukan merupakan pengguna/ pelanggan air bersih.

Tabel 2. Ringkasan akses ke bahan pangan

Kabupaten	Status Akses Bahan Pangan *	Jumlah Kecamatan	Nilai Indeks		
			Min	Rataan	Mak
Lebak	Kurang (indeks = 0)	14	0	0	0
	Baik (indeks > 0)	14	0.01	2.07	7.23
Pandeglang	Kurang (indeks = 0)	17	0	0	0
	Baik (indeks > 0)	18	0.02	2.52	8.87
Total		63			

* Akses ke bahan pangan yang baik diindikasikan dengan nilai indeks yang tidak 0

Pemanfaatan bahan pangan menggunakan jumlah fasilitas Kesehatan per kecamatan per kapita sebagai indikator. Semakin tinggi indeks pemanfaatan bahan pangan mengindikasikan semakin mudah dalam pemanfaatan bahan pangan. Tabel 3 menunjukkan bahwa kedua kabupaten memiliki karakter yang berbeda pada pemanfaatan bahan pangan. Kabupaten Pandeglang mayoritas dengan pemanfaatan bahan pangan yang baik,

sedangkan Kabupaten Lebak sebaliknya, lebih dari 50% kecamatan di Kabupaten Lebak memiliki indeks pemanfaatan bahan pangan yang kurang.

Tabel 3. Ringkasan pemanfaatan bahan pangan

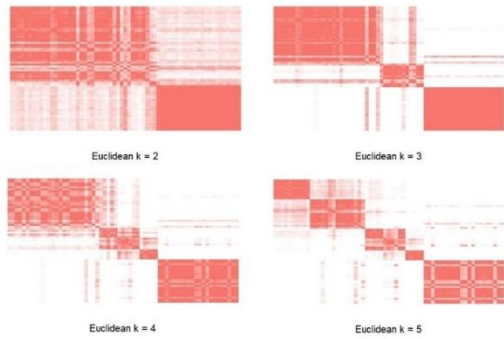
Kabupaten	Status Pemanfaatan Bahan Pangan *	Jumlah Kecamatan	Nilai Indeks		
			Min	Rataan	Mak
Lebak	Kurang (indeks < 1)	18	0.52	0.74	0.97
	Baik (indeks > 1)	10	1.01	1.28	1.66
Pandeglang	Kurang (indeks < 1)	13	0.39	0.7	0.98
	Baik (indeks > 1)	22	1.03	1.38	2.84
Total		63			

* Pemanfaatan bahan pangan yang baik diindikasikan dengan nilai indeks yang lebih dari 1

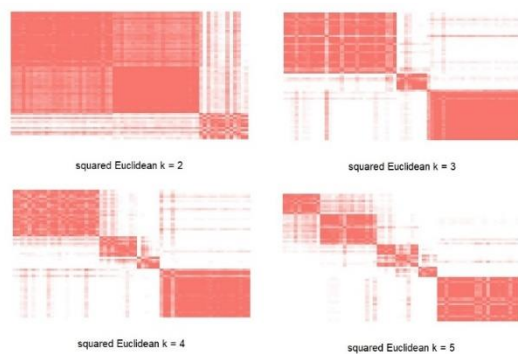
3.2 Analisis kluster

Analisis kluster diterapkan pada data kecamatan di Kabupaten Lebak dan Pandeglang. Data yang digunakan adalah data asli (bukan data terstandarisasi) karena semua variabel yang digunakan berupa indeks hasil dari proporsi, sehingga setiap variabel memiliki kontribusi yang sama. Penggunaan data tanpa distandarisasi juga memudahkan dalam interpretasi. Data asli ini kemudian dihitung jarak antar objeknya. Jarak yang digunakan adalah *Euclidean*, *squared Euclidean*, dan *Manhattan*.

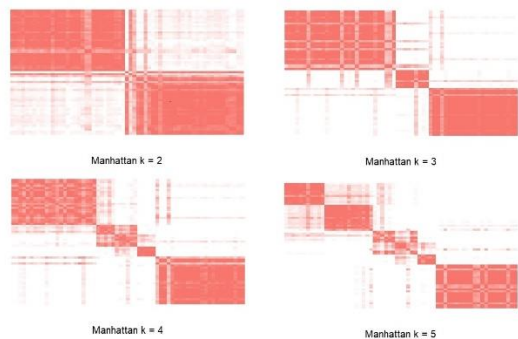
Hasil kluster dengan jarak *Euclidean*, *squared Euclidean*, dan *Manhattan* divalidasi dengan consensus clustering (Monti et al., 2003; Rupji et al., 2019) yang mendasarkan pada teknik bootstrapping. Heatmap dengan blok warna diagonal yang bersih dan penuh (tanpa titik-titik putih) menggambarkan hasil yang paling stabil, sedangkan Heatmap dengan warna yang tidak penuh (memiliki titik-titik putih yang banyak) menunjukkan separasi didalam cluster kurang baik. Gambar 2, 3, dan 4 menunjukkan *heatmap* hasil kluster dengan jumlah kluster 2, 3, 4, dan 5 pada jarak *Euclidean*, *squared Euclidean*, dan *Manhattan*. Gambar 2 memperlihatkan jumlah kluster 3 menghasilkan *heatmap* yang paling stabil karena memiliki blok warna diagonal paling penuh, sedangkan Gambar 3 dan 4 jumlah kluster 3 dan 4 menghasilkan *heatmap* yang paling baik.



Gambar 2. Heatmap hasil validasi kluster pada jumlah kluster (k) = 2, 3, 4, dan 5 dengan jarak *Euclidean*



Gambar 3. Heatmap hasil validasi kluster pada jumlah kluster (k) = 2, 3, 4, dan 5 dengan jarak *squared Euclidean*



Gambar 4. Heatmap hasil validasi kluster pada jumlah kluster (k) = 2, 3, 4, dan 5 dengan jarak *Manhattan*

Teknik validasi kriteria internal juga diaplikasikan selain teknik kriteria relatif dengan *consensus clustering* untuk memperkuat hasil validasi. Validasi internal digunakan untuk mengetahui kualitas hasil dari partisi. Ukuran yang digunakan pada validasi internal adalah kepadatan dan pemisahan (Arbelaitz et al., 2013). Kriteria internal yang digunakan adalah silhouette

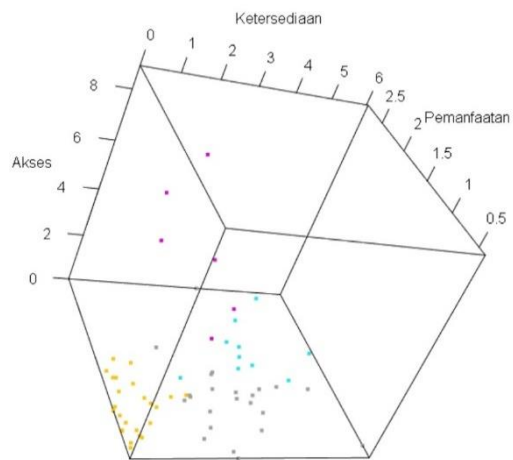
(Rousseeuw, 1987) dan *medoid-based shadow value* (Budiaji, 2019). Nilai silhouette berkisar [-1,1] sedangkan nilai *medoid-based shadow value* berada pada [0,1]. Semakin tinggi nilai silhouette/ *medoid-based shadow value*, menunjukkan cluster terpartisi dengan baik (nilai 1 artinya terpisah sempurna). Tabel 4 menunjukkan bahwa jumlah kluster 2, 3, atau 4 dengan jarak *squared Euclidean* menghasilkan validasi internal yang paling baik karena nilainya paling tinggi. Berdasarkan pertimbangan kriteria relatif, internal, dan kemudahan analisis lanjutan terkait ketahanan pangan di tingkat kecamatan, jumlah kluster 4 dengan jarak *squared Euclidean* dipilih sebagai jumlah kluster yang paling cocok/ sesuai.

Tabel 4. Nilai validasi internal dengan silhouette dan *medoid-based shadow value*

Kriteria	Jarak	Jumlah Kluster			
		2	3	4	5
Silhouette	Euclidean	0.369	0.470	0.480	0.438
	squared Euclidean	0.659*	0.628*	0.645*	0.618
	Manhattan	0.386	0.461	0.464	0.427
Medoid-based shadow value	Euclidean	0.547	0.620	0.615	0.615
	squared Euclidean	0.790*	0.788*	0.797*	0.771
	Manhattan	0.580	0.630	0.607	0.597

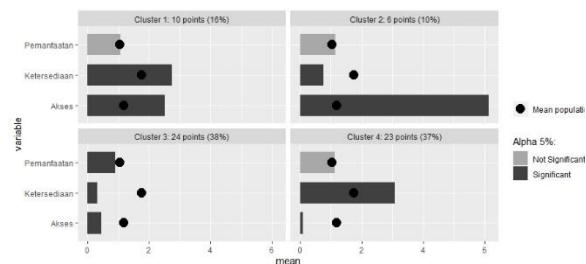
* Jumlah kluster yang cocok berdasarkan nilai yang paling tinggi

Plot 3 dimensi mudah dibuat karena variabel yang digunakan dalam analisis kluster ada 3. Gambar 5 menunjukkan plot objek dalam tiga dimensi dengan jumlah kluster sebanyak 4 berdasarkan hasil analisis kluster dengan jarak *squared Euclidean*. Untuk mempermudah interpretasi hasil analisis, barplot (Gambar 6) dibuat.



Gambar 5. Plot 3 dimensi hasil analisis kluster dengan $k = 4$ dengan jarak squared *Euclidean*

Gambar 6 menunjukkan bahwa analisis kluster menghasilkan kluster 3 (38%) dengan karakteristik bernilai rendah pada ketiga variabel (kluster rawan). Kluster 1 (16%) adalah kluster ketersediaan dan akses karena anggota kluster memiliki nilai yang tinggi pada kedua variabel tersebut. Kluster 2 (10%) merupakan kluster dengan akses ke bahan pangan yang sangat baik, sedangkan kluster 4 (37%) sangat baik pada ketersediaan bahan pangan. Analisis lanjutan perlu pada kluster 1 karena kluster tersebut merupakan kluster dengan kategori yang rawan pada ketiga aspek (ketersediaan, akses, dan pemanfaatan bahan pangan).



Gambar 6. Barplot 4 kluster berdasarkan variabel pemanfaatan, ketersediaan dan akses bahan pangan

Anggota kluster 3 berjumlah 24 kecamatan (38%) dari 63 kecamatan di Kabupaten Lebak dan Pandeglang. Hanya ada satu kecamatan di Kabupaten Pandeglang yang masuk dalam kluster 3 yaitu Kecamatan Cadasari, sedangkan 23 kecamatan lainnya berasal dari Kabupaten Lebak. Dengan jumlah kecamatan sebanyak 28 kecamatan di Kabupaten Lebak, berarti sebanyak 82% kecamatan di Kabupaten Lebak masih berkategori rawan.

4. KESIMPULAN

Kecamatan di Kabupaten Pandeglang memiliki ketersediaan dan pemanfaatan bahan pangan yang lebih baik daripada kecamatan di Kabupaten Lebak. Pada variabel akses ke bahan pangan kecamatan di kedua kabupaten berkarakter mirip yaitu setengahnya masih kurang akses. Analisis kluster yang telah divalidasi menghasilkan 4 kluster pada jarak squared *Euclidean* sebagai jumlah kluster yang paling sesuai. Kluster ketiga perlu diperhatikan karena merupakan kluster yang rawan pangan dengan anggota 38% dari total populasi kecamatan di Kabupaten Lebak dan Pandeglang.

UCAPAN TERIMA KASIH

Ucapan terima kasih kami sampaikan kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Sultan Ageng Tirtayasa (Untirta) yang telah mendanai penelitian ini melalui skema Hibah Penelitian Madya Fakultas Pertanian Untirta.

DAFTAR PUSTAKA

Aditya A, Sari BN, dan Padilah TN. 2021. Perbandingan pengukuran jarak Euclidean dan Gower pada kluster k-medoids. *Jurnal Teknologi dan Sistem Komputer* Vol. 9, No. 1, Hal. 1-7

Arbelaitz O, Gurrutxaga I, Muguerza J, Perez J, and Perona I. 2013. An extensive comparative study of cluster validity indices. *Pattern Recognition* 46, 243–256

Balitbangkes. 2019. *Laporan Provinsi Banten Riskesdas 2018/* Badan Penelitian dan Pengembangan Kesehatan. Jakarta : Lembaga Penerbit Badan Penelitian dan Pengembangan Kesehatan.

BPS Lebak. 2019. *Kabupaten Lebak dalam Angka*. Lebak: Badan Pusat Statistik Kabupaten Lebak.

BPS Pandeglang. 2019. *Kabupaten Pandeglang dalam Angka*. Pandeglang: Badan Pusat Statistik Kabupaten Lebak.

Budiaji W. 2019. Medoid-based shadow value validation and visualization. *International Journal of Advances in Intelligent Informatics* Vol. 5 (2), pp 76-88

Budiaji, Weksi, and Friedrich Leisch. 2019. Simple K-Medoids Partitioning Algorithm for Mixed Variable Data. *Algorithms* 12, no. 9: 177. <https://doi.org/10.3390/a12090177>

Budiaji, W. 2021. kmed: distance-based k-medoids. R package version 0.4.0. <https://cran.r-project.org/package=kmed>

Dolnicar, S. and F. Leisch. 2014. Using graphical statistics to better understand market segmentation solutions. *International Journal of Market Research* 56, 207–230.

FSC and WFP. 2015. *Food Security and Vulnerability Atlas of Indonesia*. Jakarta, Indonesia: FSC and WFP.

Leisch, F. 2006. A toolbox for K-centroids cluster analysis. *Comput. Stat. Data Anal.* 51, 526–544. [<http://dx.doi.org/10.1016/j.csda.2005.10.006>]

Monti S, Tamayo P, Mesirov J, and Golub T. 2003. Consensus Clustering: A Resampling-Based

- Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Mach. Learn.*, vol. 52, pp. 91–118, doi: 10.1023/A:1023949509487.
- Mun'im A. 2012. Analisis Pengaruh Faktor Ketersediaan, Akses, dan Penyerapan, Pangan terhadap Ketahanan Pangan di Kabupaten Surplus Pangan: Pendekatan Partial Least Square Path Modelling. *Jurnal Agro Ekonomi* Vol. 30 No. 1, Hal: 41-58
- Murdoch, D & Adler D. 2021. rgl: 3D Visualization Using OpenGL. R package version 0.108.3. <https://CRAN.R-project.org/package=rgl>
- Pamulang MNP, Aini MN, dan Enri U. 2021. Komparasi Distance Measure Pada K-Medoids Clustering untuk Pengelompokan Penyakit Ispa. *Edumatic: Jurnal Pendidikan Informatika* Vol. 5 No. 1, Hal. 99-107
- Pujiati S, Pertiwi A, Silfia CCS, Ibrahim DM, Hafida SHN. 2020. Analisis Ketersediaan, Keterjangkauan, dan Pemanfaatan Pangan dalam Mendukung Tercapainya Ketahanan Pangan Masyarakat di Provinsi Jawa Tengah. *Jurnal Sosial Ekonomi Pertanian* Vol. 16 No. 2, Hal: 123-133
- R Core Team. 2020. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2015, available at: <https://www.r-project.org/>
- Rousseeuw, PJ. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, vol. 20, pp. 53–65
- Rupji M, Dwivedi B, Kowalski J. 2019. NOJAH: NOt Just Another Heatmap for genome-wide cluster analysis. *PLoS ONE* 14(3): e0204542
- Simatupang P. 2007. Analisis Kritis Terhadap Paradigma dan Kerangka Dasar Kebijakan Ketahanan Pangan Nasional. *Forum Penelitian Agro Ekonomi* Vol. 25 No.1 Hal: 1–18
- Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Yu, P.S.; Steinbach M; Hand D; and Steinberg D. 2008. Top 10 algorithms in data mining. *Knowl. Inf. Syst.*, 14, 1–37. [<http://dx.doi.org/10.1007/s10115-007-0114-2>]

LAMPIRAN

Data yang digunakan dalam artikel ini dapat diunduh pada DOI: 10.5281/zenodo.6525997. Kode software R juga tersedia untuk reproduksi ulang pada DOI: 10.5281/zenodo.6526018 (<https://zenodo.org/record/6526018/files/jstat.R?download=1>).

Penerapan Text Mining pada Analisis Sentimen Pengguna Twitter Layanan Transportasi Online Menggunakan Metode *Density Based Spatial Clustering of Applications with Noise* (DBSCAN) dan K-Means

Dini Krisnawati Alfiki Astutik⁽¹⁾, Artanti Indrasetianingsih⁽²⁾, Fenny Fitriani⁽³⁾
Program Studi Statistika, Fakultas Sains dan Teknologi, Universitas PGRI Adi Buana Surabaya
Jl. Dukuh Menanggal XII/4, Surabaya 60234
dinialfiki98@gmail.com⁽¹⁾, artanti.indra@unipasby.ac.id⁽²⁾, fenny_f@unipasby.ac.id

ABSTRAK

Transportasi *online* saat ini menjadi populer dan diminati masyarakat di Indonesia dengan transportasi *online* yang banyak digunakan adalah Grab dan Gojek. Meskipun layanan transportasi *online* mendapat respon positif namun terdapat masalah yang dihadapi yaitu banyaknya konsumen yang kecewa dan merasa tidak puas dengan pelayanan yang diberikan. Penelitian ini bertujuan untuk mengetahui pengelompokan tanggapan masyarakat terhadap kedua transportasi *online* tersebut. Tanggapan masyarakat mengenai layanan transportasi *online* didapat dari salah satu media sosial yang banyak digunakan oleh masyarakat Indonesia yaitu twitter. Datapada twitter berupa kumpulan *text* sehingga diperlukan *text mining* untuk menganalisisnya. Salah satu analisis dalam *text mining* adalah *text clustering* sehingga pada penelitian ini menggunakan *text clustering* untuk mengelompokkan pendapat menjadi beberapa kategori. Metode yang digunakan pada *text clustering* adalah metode *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) dan *K-Means*. DBSCAN adalah sebuah metode yang membentuk *cluster* dari data-data yang saling berdekatan, sedangkan data yang saling berjauhan tidak akan menjadi anggota *cluster* dan biasa disebut sebagai *noise*. *K-Means* adalah teknik *clustering* yang sederhana dan cepat dalam proses *clustering* obyek serta mampu mengelompokkan data dalam jumlah cukup besar. Hasil penelitian menunjukkan bahwa metode DBSCAN dan *K-Means* kurang tepat digunakan pada penelitian ini dalam mengelompokkan *tweet* yang ditujukan kepada layanan transportasi online Gojek dan Grab karena memiliki nilai *silhouette coefficient* kurang dari 0.5 artinya struktur lemah atau *tweet* tanggapan masyarakat kepada layanan transportasi belum berada pada kelompok yang tepat.

Kata kunci : DBSCAN, K-Means, Text Clustering, Transportasi Online

ABSTRACT

Online transportation is now becoming popular and in demand by people in Indonesia with online transportation that is widely used, namely Grab and Gojek. Although online transportation services received a positive response, there were problems faced, namely the number of consumers who were disappointed and dissatisfied with the services provided. This study aims to determine the grouping of public responses to the two online transportation. The public's response to online transportation services is obtained from one of the social media that is widely used by the Indonesian people, namely Twitter. The data on Twitter is in the form of a collection of text, so text mining is needed to analyze it. One of the analyzes in text mining is text clustering, so in this study using text clustering to group opinions into several categories. The methods used in text clustering are *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) and *K-Means* methods. DBSCAN is a method that forms clusters of data that are close to each other, while data that are far from each other will not be a member of the cluster and are commonly referred to as noise. *K-Means* is a clustering technique that is simple and fast in the object clustering process and is able to group large amounts of data. The results show that the DBSCAN and *K-Means* methods are not appropriate to be used in this study in classifying tweets aimed at online transportation services Gojek and Grab because they have a *silhouette coefficient* value of less than 0.5 meaning that the structure is weak or the public response tweets to transportation services are not yet in the cluster right.

Keywords : DBSCAN, K-Means, Online Transportation, Text Clustering

1. PENDAHULUAN

Berdasarkan data google temasek, Brain dan Company yang bertajuk economy SEA pada tahun 2019, Indonesia menjadi negara tercepat dan terbesar dalam pertumbuhan ekonomi digital di kawasan Asia Tenggara. Selain *e-commerce*, ekonomi digital Indonesia juga disumbang oleh bisnis transportasi *online*. Fenomena transportasi *online* ini menjadi cepat populer dan diminati oleh masyarakat karena menawarkan inovasi terbaru mengenai transportasi yang digabungkan dengan teknologi komunikasi secara *online*, dan menyediakan berbagai layanan, sehingga memberikan kemudahan bagi masyarakat dalam efisiensi waktu untuk memesan layanan transportasi online dimana saja dan kapan saja dengan hanya menggunakan *smartphone*. Selain itu, transportasi online menjadi salah satu pilihan utama masyarakat karena harga dan kenyamanan.

Hasil survei Asosiasi Penyelenggara Jasa Internet Indonesia (APJI) (2020) periode 2019-kuartal II/2020 mencatat sebanyak 41,9% masyarakat Indonesia menggunakan transportasi online dengan Grab dan Gojek menjadi layanan aplikasi transportasi *online* yang paling sering digunakan oleh masyarakat. Sebanyak 21,3% masyarakat menggunakan Grab dan 19,4% masyarakat menggunakan Gojek. Namun, meskipun transportasi online dapat diterima dengan baik oleh masyarakat di Indonesia, terdapat permasalahan yang dirasakan oleh kedua perusahaan tersebut, yaitu masih banyaknya konsumen yang tidak puas dan kecewa dengan pelayanan yang diberikan. Dominannya pendapat konsumen yang menilai positif pelayanan jasa transportasi online, nampaknya tidak serta merta menghapus kekecewaan konsumen.

Berdasarkan survei yang dilakukan oleh Yayasan Lembaga Konsumen Indonesia (YLKI) pada 5-16 April 2017 diketahui bahwa 59% konsumen transportasi *online* pernah dikecewakan. Banyaknya keluhan konsumen terhadap transportasi *online* menunjukkan bahwa transportasi *online* belum mempunyai standar pelayanan minimal yang jelas dan terukur yang berakibat antara pengemudi satu dengan lainnya tidak sama dalam memberikan pelayanan terhadap konsumennya serta belum mempunyai mekanisme penanganan pengaduan. Tantangan lain yang ada pada konsumen layanan transportasi online saat ini adalah harga. Sebuah penelitian mengenai kesan konsumen tentang ojek *online* juga pernah dilakukan oleh *Research Institute of Economic Development (RISED)*. Hasil penelitian menyatakan bahwa sejumlah 71,12 persen konsumen transportasi *online*

di Indonesia bisa jadi berpaling menggunakan kendaraan pribadi jika tarif transportasi *online* naik (Lingga, Murti Ali. 2019).

Konsumen yang setia merupakan hal yang penting untuk kelangsungan bisnis, jadi bagaimana cara menarik dan mempertahankan loyalitas konsumen. Kepuasan konsumen tentunya merupakan hal penting bagi perusahaan apabila ingin terus bersaing. Konsumen dapat dengan mudah mengekspresikan ulasan, pengaduan, kepuasan dan lainnya dengan mudah melalui berbagai media sosial, salah satunya adalah Twitter. Twitter merupakan media sosial yang populer di Indonesia maupun di berbagai negara lainnya. Menurut data dari Pos dan Informatika Kementerian Komunikasi dan Informatika pada tahun 2020 tercatat sebanyak 19,5 juta dari 245 juta penduduk Indonesia menggunakan twitter. Perusahaan transportasi *online* memiliki akun resmi di Twitter yang menampung *tweet* komentar seperti pelayanan, driver, dan aplikasi, maupun memberikan informasi yang terkini yaitu @gojekindonesia dan @GrabID.

Data dari pendapat masyarakat pada twitter sangat berperan sebagai umpan balik layanan suatu produk atau layanan. Pendapat masyarakat yang ditujukan pada akun resmi transportasi *online* di twitter dapat dikelompokkan menjadi beberapa kategori atau *cluster*. Sehingga perlu dilakukan *text clustering* untuk mempermudah pihak transportasi *online* dalam menanggapi pendapat terkait keluhan, memberikan jawaban dan memberikan informasi kepada masyarakat terkait promosi atau info penting di media sosial twitter. *Text clustering* dapat mengelompokkan kata pada sebuah pendapat yang memiliki kemiripan untuk menjadi beberapa kategori atau *cluster*. Penentuan kategori atau kelompok dari *tweet* yang ditujukan terhadap akun resmi transportasi *online* Gojek dan Grab dalam mengantisipasi banyaknya *tweet*, seperti membuat template tanggapan untuk setiap kategori, dan mengelompokkan *tweet* yang masuk berdasarkan kategori. Sehingga perlu dilakukan penelitian untuk mengetahui tanggapan masyarakat berdasarkan sentimen di media sosial twitter terhadap kedua layanan transportasi tersebut.

Analisis sentimen merupakan metode analisis berbasis komputasi mengenai pendapat, sentimen, evaluasi, sikap, penilaian dan emosi berkenaan dengan suatu topik, produk, layanan, organisasi, individu, ataupun kegiatan tertentu lainnya (Liu, 2010). Sebelum melakukan analisis sentimen, diperlukan praproses teks dengan metode *text mining* untuk mengolah data teks agar siap untuk dianalisis. *Text mining* dapat memberikan solusi dari permasalahan seperti pemrosesan, pengelompokkan, dan menganalisa *unstructured text* dalam jumlah besar.

Terdapat banyak metode *text clustering* yang sering digunakan untuk analisis sentimen namun metode yang sering digunakan dalam *text clustering* adalah metode *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)* dan *K-Means*. Metode DBSCAN adalah algoritma pengelompokan yang didasarkan pada kepadatan (density) data (Tan, dkk., 2007). Metode ini membentuk *cluster* dari data-data yang saling berdekatan, sedangkan data yang saling berjauhan tidak akan menjadi anggota *cluster* (Adinugroho dan Sari, 2018). Sedangkan *K-Means* merupakan metode *clustering* yang memiliki kemampuan untuk mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi (Irwansyah dan Faisal, 2015).

Penelitian mengenai analisis sentimen pernah dilakukan oleh Isnawarty, et al (2019) mengenai *text clustering* pada akun twitter layanan ekspedisi JNE, J&T dan POS Indonesia menggunakan DBSCAN dan *K-Means* menghasilkan kesimpulan bahwa DBSCAN menjadi metode lebih baik dari *K-Means* dalam mengelompokkan *tweet* layanan ekspedisi karena memiliki nilai *silhouette coefficient* yang lebih tinggi. Penelitian lain dilakukan oleh Arsih, et al (2016) mengenai metode pengclusteran berbasis densitas menggunakan algoritma DBSCAN terhadap fasilitas kredit nasabah yang menghasilkan kesimpulan bahwa metode DBSCAN bekerja lebih baik dari *K-Means* karena memiliki run time lebih cepat dalam mengelompokkan fasilitas kredit nasabah.

Penelitian ini bertujuan melakukan analisis sentimen mengenai tanggapan masyarakat terhadap layanan transportasi *online* berbasis *text mining* data twitter. Pendapat perlu diidentifikasi dan dikelompokkan menjadi beberapa kategori. Kategori pendapat dapat berupa kepuasan, keluhan, saran, opini, pemberian informasi, dan lain-lain. Pengidentifikasi informasi dari sebuah *text* diperlukan analisis menggunakan *text mining*. Penentuan kategori pengelompokan pada *text mining* dapat dilakukan dengan menggunakan metode *text clustering*. Metode *text clustering* yang digunakan pada penelitian ini adalah metode DBSCAN dan *K-Means*. Dari kedua metode a metode tersebut, akan dipilih metode terbaik yang diharapkan dapat memberikan saran kepada jasa layanan transportasi *online* yaitu Grab dan Gojek dapat mengambil keputusan yang tepat untuk keberlangsungan perusahaan kedepannya dengan mengetahui kategori pendapat yang paling sering muncul dan pengelompokan tanggapan yang diberikan oleh masyarakat dalam mengukur kepuasan dan meningkatkan performa layanan.

2. METODE PENELITIAN

Data yang digunakan pada penelitian ini adalah data yang diambil dari kumpulan *tweet* dari pengguna Twitter di Indonesia dengan akun Twitter yang digunakan dalam analisis adalah akun resmi dari layanan transportasi *online* Grab Indonesia (@GrabID) dan Gojek Indonesia (@gojekindonesia). Data *tweet* diambil pada tanggal 01 Januari hingga 10 Februari 2021 dengan menggunakan Twitter API (*Application Programming Interface*).

Variabel penelitian yang digunakan adalah frekuensi kemunculan kata dasar dari setiap *tweet* yang ditujukan kepada @GrabID dan @gojekindonesia yang telah dilakukan preprocessing dengan skala data rasio.

Langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut:

1. Mengambil data *tweet* dengan menggunakan Twitter API
 - a. Memasukkan *keyword* yang berhubungan dengan akun Twitter @GrabID dan @gojekindonesia
 - b. Menyimpan hasil crawling ke database.
2. Melakukan *text preprocessing* pada data *tweet* layanan transportasi *online* Gojek dan Grab.
 - a. Melakukan *cleaning* data, yaitu menghapus karakter HTML, link URL, *retweet* (RT), username (@username), baris baru, angka, hashtag (#), emoticons, tanda baca, dan *double space*.
 - b. Melakukan *case folding*, yaitu mengubah seluruh teks dengan huruf kecil (non kapital).
 - c. Melakukan *stemming*, yaitu menghilangkan kata imbuhan sehingga didapatkan kata dasar.
 - d. Menghapus kata pada *tweet* yang terdapat pada *stopwords*.
 - e. Melakukan *tokenizing*, yaitu memecah *tweet* menjadi kata per kata.
 - f. Mengubah kata-kata informal menjadi bentuk formal.
 - g. Menghubungkan kumpulan kata yang memiliki makna bila digabungkan.
3. Mengubah data *tweet* ke dalam bentuk frekuensi kemunculan kata dengan menggunakan TF-IDF.
4. Melakukan visualisasi *tweet* dengan *Word Cloud*.
5. Melakukan clustering data

- 5.1. Melakukan clustering data menggunakan metode DBSCAN
 - a. Menentukan parameter *MinPts* dan *Eps*
 - b. Memilih *tweet* *p* secara acak.
 - c. Menandai *tweet* sebagai *core point* dengan berdasarkan nilai *Eps* dan melanjutkan perhitungan jarak antara *core point* dengan *point* yang lain dengan menggunakan rumus jarak *Euclidean*.
 - d. Membuat *cluster* baru dengan menambahkan menambahkan *tweet p* ke dalam *cluster*.
 - e. Identifikasi pada data yang ditandai sebagai *core point* dan melanjutkan sampai semua print telah diproses.
 - f. Jika ada *tweet* yang tidak masuk ke dalam *cluster* manapun akan ditandai sebagai *noise*.
- 5.2. Melakukan clustering data menggunakan metode K-Means.
 - a. Memilih secara acak *k centroid* awal dalam data
 - b. Menentukan jarak setiap kata terhadap pusat *cluster (centroid)*
 - c. Mengelompokkan setiap kata berdasarkan kedekatannya dengan *centroid* terdekat dengan menggunakan *Euclidean distance*.
 - d. Menghitung ulang pusat *cluster (centroid)* baru
 - e. Ulangi langkah (b) hingga (d), sampai anggota yang ada pada tiap *cluster* tidak berubah. Jika jumlah *cluster k* belum diketahui, dapat menggunakan nilai *Variance Ratio Criterion (VRC)* dan metode *Elbow*.
 - f. Membandingkan hasil *clustering* metode *DBSCAN* dan *K-Means* dengan melihat nilai *silhouette coefficient* terbesar untuk menentukan metode *clustering* terbaik yang terpilih
 - g. Melakukan interpretasi dan menarik kesimpulan

3. HASIL DAN PEMBAHASAN

Hasil proses *crawling* data yang telah dilakukan menggunakan *Twitter API* mulai tanggal 01 Januari 2021 hingga 10 Februari 2021 pada akun resmi layanan transportasi online Gojek yaitu @GojekIndonesia diperoleh tanggapan pelanggan

sebanyak 2.291 *tweet* dan 2.416 *tweet* yang diperoleh dari akun resmi layanan transportasi online Grab yaitu @GrabID.

Text preprocessing merupakan tahapan awal dalam analisis *clustering* yang digunakan untuk mengubah data pada kumpulan *tweet* yang didapatkan dari proses *crawling* menjadi data yang terstruktur sehingga dapat mempermudah analisis selanjutnya. Tahapan yang dilakukan pada *Text preprocessing* yaitu menghapus link, hastag, simbol *retweet (RT)*, *username*, angka, *emoticon*, baris kosong, *punctuation*, spasi berlebih, *case folding*, *stemming*, *stopword*, dan *tokenizing*.

A. Layanan Transportasi Online Gojek

Langkah selanjutnya yang dilakukan setelah melakukan *text preprocessing* adalah melakukan perhitungan frekuensi kemunculan pada masing-masing kata pada setiap *tweet* dan selanjutnya kata tersebut akan dihitung bobot menggunakan TF-IDF. Struktur data berdasarkan frekuensi kemunculan kata yang diperoleh setelah dilakukannya *text preprocessing* pada akun layanan transportasi online Gojek ditunjukkan pada Tabel 1 sebagai berikut.

Tabel 1. Struktur Data Layanan Transportasi Online Gojek Setelah *Text Preprocessing*

Tweet ke-	Kata						
	akun	...	diskon	...	promo	...	voucher
1	0		1		1		0
2	0		0		0		0
...							
1059	2		0		0		0
...							
2291	0		0		0		0

Jumlah kata yang diperoleh dari proses *text preprocessing* sejumlah 104 kata dasar dan kata-kata tersebut merupakan variabel penelitian yang akan digunakan. Namun, sebelum melakukan proses *text clustering*, perlu dilakukan pembobotan pada kata dasar dari setiap *tweet* dengan menggunakan *Term Frequency – Inverse Document Frequency (TF-IDF)*. Proses perhitungan TF-IDF dilakukan dengan menggunakan persamaan (1) dan (2) yang ditunjukkan pada Tabel 2 sebagai berikut.

Tabel 2. Hasil perhitungan TF-IDF Layanan Transportasi Online Gojek

Twee t ke-	Kata						
	akun	...	diskon	...	promo	...	vouche r
1	0		0.670 3		0.632 2		0
2	0		0		0		0
...							

1059	0,6589		0		0		0
...							
2291	0		0		0		0

Tabel 2 menunjukkan bahwa kata “diskon” yang terdapat pada *tweet* ke-1 memiliki bobot sebesar 0,6703. Hal ini menunjukkan bahwa kata “diskon” pada *tweet* ke-1 memiliki frekuensi kemunculan yaitu $1 \times 0,6703 = 0,6703$. Begitu pula dengan perhitungan TF-IDF pada kata-kata dasar yang lainnya.

Visualisasi data text menggunakan *Word Cloud* digunakan untuk mengetahui frekuensi kata-kata yang paling sering muncul pada data dengan menggunakan data pada Tabel 2 yang telah boboti menggunakan TF-IDF yang ditunjukkan pada Gambar 1 sebagai berikut.



Gambar 1. Visualisasi *Word Cloud* dari Layanan Transportasi Online Gojek

Gambar 1 menunjukkan bahwa kata yang sering muncul atau memiliki frekuensi tinggi pada tanggapan masyarakat yang ditujukan kepada akun resmi gojek (@GojekIndonesia) adalah kata “dm”, “gojek”, “admin”, “tolong”, “cek”, “order”, “gopay”, “akun”, “gofood”, “bayar”. Sedangkan kata-kata lain yang berukuran kecil artinya kemunculan kata tersebut memiliki frekuensi rendah. Namun, untuk dapat melihat lebih detail apa saja kata-kata yang dan hal yang paling sering dikeluhkan pelanggan kepada akun resmi layanan transportasi online Gojek, data yang telah didapatkan tersebut akan diolah dan dikelompokkan menggunakan metode DBSCAN dan *K-Means*.

Clustering dengan metode DBSCAN pada data *tweet* yang berasal dari akun resmi layanan transportasi online Gojek dilakukan dengan cara melakukan percobaan manual menggunakan

kombinasi nilai *MinPts* dan *Eps* yang berbeda-beda sesuai dengan keinginan peneliti hingga mendapatkan hasil kombinasi *MinPts* dan *Eps* yang terbaik, kombinasi yang digunakan adalah nilai *MinPts* sebesar (5, 15, 25,...95) serta nilai *Eps* antara 0.1 sampai 1. Berikut merupakan hasil 5 kombinasi *MinPts* dan *Eps* yang menghasilkan nilai *silhouette coefficient* tertinggi yang ditunjukkan pada Tabel 3.

Tabel 3. Nilai *Silhouette Coefficient* dan *Noise* metode DBSCAN pada Akun Layanan Transportasi Online Gojek

<i>MinPts</i>	<i>Eps</i>	<i>Silhouette Coefficient</i>	<i>Noise</i>
5	0.6	0.4253	752
5	0.7	0.4082	627
5	0.6	0.4062	858
5	0.5	0.4009	972
5	0.4	0.3958	991

Tabel 4 menunjukkan bahwa hasil *clustering* dari 2.291 *tweet* yang ditujukan kepada akun layanan transportasi gojek menggunakan metode DBSCAN dengan parameter *Eps* sebesar 0,6 dan *MinPts* sebesar 5 menghasilkan nilai *silhouette coefficient* sebesar 0,4253.

Proses *Clustering* menghasilkan jumlah *cluster* sebanyak 69 *cluster* dengan *tweet* yang masuk ke dalam *cluster* sebanyak 1.539 *tweet* dan sisanya sebanyak 752 *tweet* tidak masuk ke dalam *cluster* atau disebut sebagai noise.

Hasil *clustering* *tweet* tanggapan pelanggan layanan transportasi gojek menggunakan metode DBSCAN pada *cluster* ke-2 dengan anggota *cluster* sebanyak 19 *tweet*, diperoleh hasil kata yang paling sering muncul yaitu "hapus", "akun", "gojek", "admin", "aplikasi", "gopay", "tolong", "login", "nomor" sehingga diperoleh kesimpulan bahwa kata menunjukkan mengenai keinginan pelanggan untuk menghapus aplikasi gojek dengan penjabarannya ditunjukkan pada Gambar 2 dan Tabel 4 sebagai berikut.



Gambar 2. Visualisasi *Word Cloud* pada *Cluster 2 metode DBSCAN* Layanan Transportasi *Online Gojek*

Clustering menggunakan metode *K-Means*. Dalam metode *K-Means*, perlu menentukan jumlah *cluster (K)* optimum terlebih dahulu dengan menggunakan nilai *Variance Ratio Criterion (VRC)* dan metode *elbow*. Jumlah *cluster* adalah yang memiliki nilai yang paling tinggi. Index yang digunakan untuk menentukan jumlah *cluster* optimum dilakukan dengan *K* mulai 2 sampai dengan 20. Berikut adalah nilai *VRC* pada data layanan transportasi online Gojek yang ditunjukkan pada Tabel 5 dan Gambar 3. Berdasarkan hasil nilai *VRC* pada Tabel 5 dan Gambar 3 menunjukkan jumlah *cluster (K)* yang terbentuk sebanyak 20 dan yang memiliki nilai *VRC* paling tinggi yaitu sebesar 150,386 pada 2 *cluster*. Dimana memberikan nilai *silhouette coefficient* sebesar 0.1771.

Tabel 4. Hasil Clustering pada Cluster ke-2 Metode DBSCAN pada Akun Layanan Transportasi *Online Gojek*

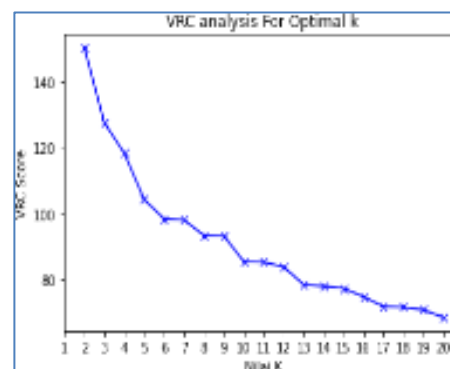
Responden	Tweet	Kata yang paling sering muncul
3	Hallo @gojekindonesia Bagaimana lcara hapus akun gojek y?	"hapus", "akun", "gojek", "admin", "aplikasi", "gopay", "tolong", "login", "nomor"
267	@bapakayammm @BuruhYogyakarta GOJEK sejak dipegang KEVIN memang sadis. driver bonus dihapus, sekarang meres mercha... https://t.co/YgMNORU1Nh	
876	@gojekindonesia Saya mau data2 pribadi sy dihapus/dihanguskan dr database gojek	
880	@gojekindonesia halo, saya ingin mengajukan penghapusan akun gojek saya, karena saya sudah tidak sudi menggunakan a... https://t.co/gecbMvJk0f	
1018	@gojekindonesia alloo gimana cara hapus akun gojek costumer.. please..	
1095	@gojekindonesia hapus permanen	
1096	@gojekindonesia hapus akun gojek bersama gopaynya?	
1332	@gojekindonesia hapus akun min:(klo gitu kan masih ada akunnya bisa login lagi	
	@gojekindonesia kak saya mau	

1499	menghapus akun yg ada ktp saya di gopay plus..karena akunnya hilang
1675	min @gojekindonesia slow respond ah w kan buru mau hapus akun
1715	@bucinnyathehun @gojekindonesia Ya selama tidak di gunakan ^^ Karena kalo dah di hapus maka kamu tidak akan bisa m...

Berikut adalah nilai Metode *Elbow* pada data akun layanan transportasi *online Gojek* yang ditunjukkan pada Tabel 6 dan divisualisaikan pada Gambar 4 sebagai berikut.

Tabel 5. Nilai *VRC* pada layanan transportasi online Gojek.

Cluster (k)	Nilai VRC	Cluster (k)	Nilai VRC
2	150,386	12	83,874
3	127,620	13	78,450
4	118,427	14	78,035
5	104,331	15	77,364
6	98,454	16	74,825
7	98,127	17	71,798
8	93,404	18	71,650
9	93,247	19	70,344
10	85,427	20	68,773
11	85,312		

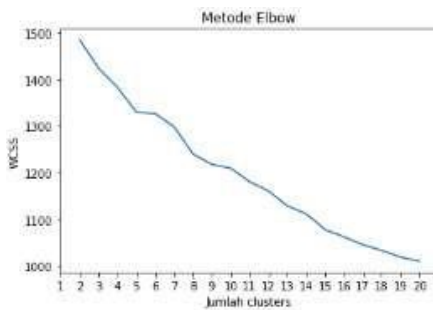


Gambar 3. Nilai *VRC* pada Layanan Transportasi *Online Gojek*

Tabel 6 Nilai Metode *Elbow* pada Layanan Transportasi *Online Gojek*

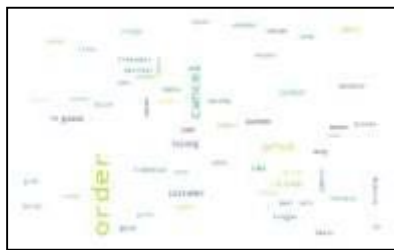
Cluster (k)	Metode Elbow	Cluster (k)	Metode Elbow
2	1485.220	12	1161.421
3	1423.951	13	1129.239

4	1383.177	14	1112.434
5	1329.829	15	1078.576
6	1327.121	16	1063.132
7	1298.649	17	1046.238
8	1240.454	18	1033.773
9	1218.157	19	1019.511
10	1209.940	20	1010.881
11	1180.991		



Gambar 4. Nilai Metode Elbow pada Layanan Transportasi Online Gojek

Berdasarkan hasil nilai metode elbow pada Tabel 6 dan Gambar 4 menunjukkan jumlah cluster (K) yang terbentuk sebanyak 5 karena membentuk siku atau mengalami penurunan nilai SSE. Dimana memberikan nilai *silhouette coefficient* sebesar 0.2138. sehingga dipilih K Optimum yang digunakan adalah 5 karena memiliki nilai *silhouette coefficient* lebih tinggi dibandingkan dengan $K = 2$. Hasil *clustering* pada layanan transportasi online Gojek menggunakan metode *K-Means* dengan visualisasi *word cloud* yang ditunjukkan pada Gambar 5. Hasil *clustering* pada cluster ke-1 layanan transportasi online Gojek menggunakan metode *K-Means* dengan visualisasi *word cloud* yang ditunjukkan pada Gambar 5 sebagai berikut.



Gambar 5 Visualisasi Word Cloud pada Cluster 1 metode *K-Means* Layanan Transportasi Online Gojek

Gambar 5 menunjukkan bahwa hasil *clustering* dari 2.291 tweet yang ditunjukkan

kepada layanan transportasi online Gojek menggunakan metode *K-Means* menghasilkan cluster optimum sebanyak 5 cluster dengan nilai *silhouette coefficient* sebesar 0,2138, dimana cluster 1 dengan anggota cluster sebanyak 112 tweet menunjukkan kata pembatalan pemesanan layanan transportasi online Gojek.

B. Layanan Transportasi Online Grab

Langkah selanjutnya yang dilakukan setelah melakukan *text preprocessing* adalah melakukan perhitungan frekuensi kemunculan pada masing-masing kata pada setiap *tweet* dan selanjutnya kata tersebut akan dihitung bobot menggunakan TF-IDF. Struktur data berdasarkan frekuensi kemunculan kata yang diperoleh setelah dilakukannya *text preprocessing* pada akun layanan transportasi online Grab ditunjukkan pada Tabel 7 sebagai berikut.

Tabel 7. Struktur Data Layanan Transportasi Online Grab Setelah *Text Preprocessing*

Tweet ke-	Kata						
	aku	...	pesan	...	sesuai	...	update
1	1		0		0		0
2	0		1		1		0
...							
1168	0		0		0		1
...							
2416	0		0		0		0

Tabel 7 menunjukkan bahwa jumlah kata yang diperoleh telah berkurang dikarenakan ada beberapa kata yang tidak memiliki makna dalam kalimat sehingga dihapus dan tidak digunakan pada analisis. Jumlah kata yang diperoleh dari proses *text preprocessing* sejumlah 100 kata dasar dan kata-kata tersebut merupakan variabel penelitian yang akan digunakan. Namun, sebelum melakukan proses *text clustering*, perlu dilakukan pembobotan pada kata dasar dari setiap *tweet* dengan menggunakan *Term Frequency – Inverse Document Frequency* (TF-IDF). Hasil perhitungan TF-IDF dapat dilihat pada Tabel 8.

Tabel 8. Hasil perhitungan TF-IDF Layanan Transportasi Online Grab

Tweet ke-	Kata						
	Aku n	...	pesan	...	sesuai	...	update
1	1		0		0		0
2	0		0.6111		0.7915		0

...						
1168	0		0		0	0.6076
...						
2416	0		0		0	0

Tabel 8 menunjukkan bahwa kata “akun” yang terdapat pada *tweet* ke-1 memiliki bobot sebesar 1. Hal ini menunjukkan bahwa kata “akun” pada *tweet* ke-1 memiliki frekuensi kemunculan yaitu $1 \times 1 = 1$. Begitupula dengan perhitungan TF-IDF pada kata-kata dasar yang lainnya.

Visualisasi data text menggunakan *Word Cloud* digunakan untuk mengetahui frekuensi kata-kata yang paling sering muncul pada data dengan menggunakan data pada Tabel 8 yang telah diboboti menggunakan TF-IDF yang ditunjukkan pada Gambar 6 sebagai berikut.



Gambar 6. Visualisasi *Word Cloud* dari Layanan Transportasi *Online Grab*

Gambar 6 menunjukkan bahwa kata yang sering muncul atau memiliki frekuensi tinggi pada tanggapan masyarakat yang ditujukan kepada akun resmi Grab (@GrabID) adalah kata “dm”, “grab”, “kirim”, “admin”, “cek”, “order”, “tolong”, “aplikasi”, “grabfood”, “ovo”. Sedangkan kata-katalain yang berukuran kecil artinya kemunculan kata tersebut memiliki frekuensi rendah.

Namun, untuk dapat melihat lebih detail apa saja kata-kata yang dan hal yang paling sering dikeluhkan pelanggan kepada akun resmi layanan transportasi *online Grab*, data yang telah didapatkan tersebut akan diolah dan dikelompokkan menggunakan metode DBSCAN dan *K-Means*.

Hasil *clustering* menggunakan metode metode *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) dengan kombinasi yang digunakan adalah nilai *MinPts* sebesar (10, 20, 30,...,60) serta nilai *Eps* antara 0.1 sampai 1. Tabel 9 merupakan hasil kombinasi *MinPts* dan *Eps* yang menghasilkan nilai *silhouette coefficient* tertinggi, yaitu sebesar 0.3419.. Proses *clustering* menghasilkan jumlah *cluster* sebanyak 27 *cluster* dan terdapat 1043 *noise* pada *tweet*.

Tabel 9. Nilai *Silhouette Coefficient* dan *Noise* metode DBSCAN pada Akun Layanan Transportasi *Online Grab*

<i>MinPts</i>	<i>Eps</i>	<i>Silhouette Coefficient</i>	<i>Noise</i>
10	0.6	0.3419	1043
10	0.7	0.3353	936
10	0.5	0.3312	1196
10	0.5	0.3255	1217
10	0.4	0.3192	1232

Berikut adalah visualisasi hasil *clustering* pada cluster ke-1 akun resmi layanan transportasi *online Grab* menggunakan metode DBSCAN yang ditampilkan secara visual menggunakan *word cloud*.

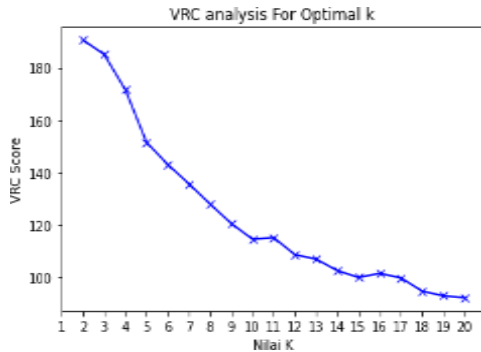


Gambar 7. Visualisasi *Word Cloud* pada Cluster 1 metode DBSCAN Layanan Transportasi *Online Grab*

Hasil *clustering* menggunakan metode *K-Means* pada layanan transportasi *online Grab*. Penentuan jumlah *cluster* optimum berdasarkan dari nilai *Variance Ratio Criterion* (VRC) dan *elbow* yang nilai tertinggi. *Index* yang digunakan untuk penentuan jumlah *cluster* optimum dilakukan dengan *K* mulai dari 2 hingga 20. Berikut adalah nilai VRC pada data akun layanan transportasi *online Grab* yang ditunjukkan pada Tabel 10 dan divisualisaikan pada Gambar 8.

Tabel 10. Nilai VRC pada Layanan Transportasi *Online Grab*

Cluster (k)	Nilai VRC	Cluster (k)	Nilai VRC
2	190.640	12	108.693
3	185.101	13	107.007
4	171.618	14	102.608
5	151.555	15	100,021
6	143.134	16	101.556
7	135.731	17	99.838
8	127.994	18	94.779
9	120.579	19	93.000
10	114,604	20	92.269
11	115,191		



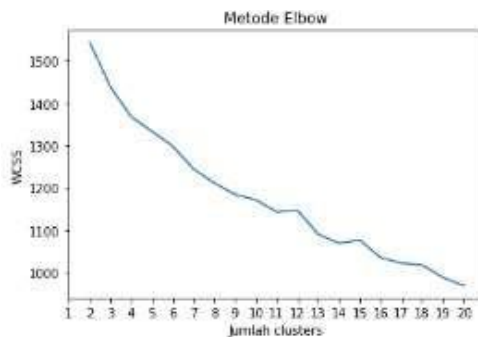
Gambar 9 Nilai VRC pada Layanan Transportasi Online Grab

Berdasarkan hasil nilai VRC pada Tabel 10 dan Gambar 9 menunjukkan jumlah *cluster* (*K*) yang terbentuk sebanyak 20 dan yang memiliki nilai VRC paling tinggi yaitu sebesar 190,640 pada 2 *cluster*. Dimana memberikan nilai *silhouette* sebesar 0.1845.

Berikut adalah nilai Metode Elbow pada data akun layanan transportasi online Grab yang ditunjukkan pada Tabel 11 dan divisualisasikan pada Gambar 10 sebagai berikut.

Tabel 11 Nilai Metode Elbow pada Layanan Transportasi Online Grab

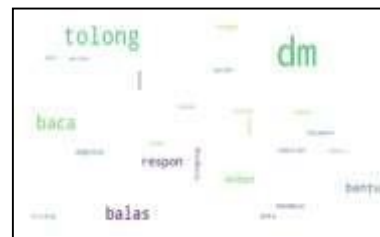
Cluster (k)	Metode Elbow	Cluster (k)	Metode Elbow
2	1544.283	12	1147.376
3	1439.247	13	1090.477
4	1368.063	14	1069.779
5	1333.848	15	1077.132
6	1298.827	16	1035.677
7	1245.134	17	1022.728
8	1211.793	18	1018.164
9	1184.853	19	988.834
10	1172.101	20	969.441
11	1143.816		



Gambar 10 Nilai Metode Elbow pada Layanan Transportasi Online Grab

Berdasarkan hasil nilai metode elbow pada Tabel 11 dan Gambar 10 menunjukkan jumlah *cluster* (*K*) yang terbentuk sebanyak 20 dan yang membentuk siku dan mengalami penurunan nilai SSE pada saat terbentuk 11 *cluster*. Dimana memberikan nilai *silhouette coefficient* sebesar 0.2897. Sehingga dipilih *K* Optimum yang digunakan adalah 11 karena memiliki nilai *silhouette coefficient* lebih tinggi dibandingkan dengan *K* = 2. Hasil *clustering* pada cluster ke-2 layanan transportasi online Grab menggunakan metode *K-Means* dengan visualisasi *word cloud* yang ditunjukkan pada Gambar 11 sebagai berikut.

Berdasarkan hasil nilai VRC pada Tabel 10 dan Gambar 8 menunjukkan jumlah *cluster* (*K*) yang terbentuk sebanyak 20 dan yang memiliki nilai VRC paling tinggi yaitu sebesar 190,640 pada 2 *cluster*. Dimana memberikan nilai *silhouette* sebesar 0.1845.



Gambar 11. Visualisasi *Word Cloud* pada Cluster ke-2 metode *K-Means* Layanan Transportasi Online Grab

Hasil *clustering* dari 2.416 tweet yang ditujukan kepada layanan transportasi online Grab menggunakan metode *K-Means* menghasilkan sebanyak 11 *cluster* dengan nilai *Silhouette coefficient* sebesar 0,2897. Gambar 11 merupakan visualisasi *word cloud* pada cluster 2 dengan anggota cluster sebanyak 94 tweet menunjukkan kata permohonan untuk merespon dan cek dan membaca *dm* (*direct message*).

Perbandingan metode DBSCAN dan *K-Means* digunakan untuk mengetahui metode terbaik yang digunakan untuk mengelompokkan tweet pada akun resmi layanan transportasi online Gojek dan Grab. Perbandingan metode *clustering* ini didasarkan pada hasil nilai *silhouette coefficient* yang diperoleh dari metode DBSCAN dan *K-Means*. Perbandingan dari kedua metode yang digunakan ditunjukkan pada Tabel 12.

Tabel 12. Perbandingan antara Metode DBSCAN dan *K-Means* pada Layanan Transportasi Online Gojek dan Grab

Layanan Transportasi Online	Metode Clustering	Silhouette Coefficient	Jumlah Cluster
Gojek	DBSCAN	0.4253	69
	K-Means	0.1771	2
Grab	DBSCAN	0.3419	27
	K-Means	0.1845	2

Tabel 12 menunjukkan bahwa nilai *silhouette coefficient* pada metode DBSCAN di semua layanan transportasi online Gojek maupun Grab lebih baik dalam mengelompokkan tweet dibandingkan dengan metode K-Means. Hal ini dapat dilihat dari nilai *silhouette coefficient* pada metode DBSCAN lebih besar dibandingkan dengan K-Means. Namun nilai *silhouette coefficient* berada diantara $0,25 < SC < 0,5$ yang berarti termasuk struktur lemah atau pengelompokkan mengenai tanggapan konsumen terhadap layanan transportasi online yang terbentuk memiliki akurasi kurang baik dan belum bisa berpartisipasi tweet ke dalam cluster yang homogen secara tepat.

Metode DBSCAN maupun K-Means kurang tepat diterapkan pada data penelitian ini karena pengelompokkan tweet layanan transportasi grab masih cenderung heterogen atau menyebar secara acak tidak sesuai dengan tujuan dari analisis cluster yaitu mengelompokkan data ke dalam beberapa cluster sedemikian hingga data yang berada di dalam cluster yang sama cenderung mempunyai sifat yang lebih homogen daripada data yang berada di cluster yang berbeda. Namun persaingan diantara kedua perusahaan tersebut tidak jauh berbeda dan sangat kompetitif sehingga perusahaan harus berinovasi secara cepat dan membuat diferensiasi dibandingkan kompetitornya agar dapat meraih pasar yang lebih tinggi.

4. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil dan pembahasan, maka kesimpulan yang diperoleh sebagai berikut :

1. Berdasarkan hasil crawling data, diketahui bahwa layanan transportasi online Grab memiliki jumlah tweet lebih banyak dibandingkan dengan Gojek. Kata yang paling sering muncul pada tweet yang ditujukan pada layanan transportasi online Gojek adalah “dm”, “Gojek”, “admin”, sedangkan kata yang paling sering muncul pada tweet yang ditujukan pada Grab adalah “dm”, “Grab”, “kirim”.
2. Berdasarkan hasil analisis clustering menggunakan metode DBSCAN pada layanan

transportasi online Gojek memiliki nilai *silhouette coefficient* sebesar 0.4253 dan noise sebanyak 752 tweet dengan 69 cluster yang terbentuk. Sedangkan untuk Grab memiliki nilai *silhouette coefficient* sebesar 0.3419 dan noise sebanyak 1043 tweet dengan 27 cluster yang terbentuk.

3. Berdasarkan hasil analisis clustering menggunakan metode K-Means diperoleh nilai *silhouette coefficient* sebesar 0.1771 pada layanan transportasi online Gojek dan sebesar 0.1845 pada Grab dengan masing-masing terbentuk 2 kluster.
4. Berdasarkan hasil analisis clustering, diketahui bahwa metode DBSCAN merupakan metode terbaik untuk mengelompokkan tweet yang ditujukan kepada akun media sosial twitter layanan transportasi Gojek dan Grab karena memiliki nilai *silhouette coefficient* tertinggi dengan menghasilkan 69 cluster untuk layanan transportasi online Gojek dan 27 cluster pada Grab.

B. Saran

Saran yang dapat diberikan kepada layanan transportasi online Gojek dan Grab yaitu sebagai bahan pertimbangan untuk evaluasi dan meningkatkan performa layanan dalam memberikan respon terhadap pertanyaan dan keluhan pelanggan dengan mengacu pada hasil clustering. Sedangkan saran untuk penelitian selanjutnya, diharapkan agar mendapatkan lebih banyak data dan lebih teliti dalam proses *text preprocessing* karena akan dapat mempengaruhi hasil cluster.

DAFTAR PUSTAKA

- Adinugroho, S. & Sari, Y. A. (2018). Implementasi Data Mining Menggunakan Weka. Malang: UB Press.
- Arsih, Nur., Nusar Hajarisman & Sutawanir Darwis (2016). Metode Pengclusteran Berbasis Densitas Menggunakan Algoritma DBSCAN. Prosiding Statistika, Vol. 2, Hal. 153-163.
- Irwansyah, E., & Faisal, M. (2015). Advanced Clustering : Teori dan Aplikasi. Yogyakarta: DeePublish.
- Isnawarty, Devi Putri & Irhamah (2019). Text Clustering pada Akun Twitter Layanan Ekspedisi JNE, J&T dan Pos Indonesia Menggunakan Metode Density Based Spatial Clustering of Applications with Noise (DBSCAN) dan K-Means. Jurnal Sains dan Seni ITS, Vol. 8, No. 2.

- Lingga, Murti Ali. (2019).
<https://ekonomi.kompas.com/read/2019/02/11/175518026/bila-tarif-ojek-online-naik-konsumen-kembali-ke-kendaraan-pribadi>.
tanggal akses: 20-Des-2020.
- Liu, Bing (2010). Sentiment Analysis and Subjectivity. Chicago : University of Illinois
- Tan, H. P., Plowman, D., & Hancock, P., (2007). Intellectual Capital and Financial Returns of Companies. Journal of Intellectual Capital, Vol. 8, Hal 76-91.

Peramalan Nilai Ekspor Provinsi Jawa Tengah dengan Metode Fuzzy Time Series Berbasis Algoritma Novel

Septi Winda Utami⁽¹⁾, Indah Manfaati Nur⁽²⁾, M. Al Haris^{(3)*}

Universitas Muhammadiyah Semarang

Jl. Kedungmundu No.18, Kedungmundu, Kec. Tembalang, Kota Semarang, Jawa Tengah 50273

e-mail: septiutami23@gmail.com, indahmnur@unimus.ac.id, alharis@unimus.ac.id*

ABSTRAK

Ekspor berperan penting dalam perekonomian suatu daerah. Ekspor dapat digambarkan melalui nilai ekspor, yaitu nilai berupa uang yang merupakan hasil dari penjualan barang/jasa ke luar negeri. Nilai ekspor perlu diramalkan untuk merumuskan target pertumbuhan ekonomi dan menetapkan kebijakan yang tepat dalam perekonomian di suatu daerah. Nilai ekspor dapat diramalkan menggunakan metode *Fuzzy time series* berbasis algoritma Novel yang dikembangkan oleh Haneen Talal Jasim, Abdul Ghafoor Jasim Salim, dan Kais Ismail Ibraheem. Tahapan metode *Fuzzy time series* berbasis algoritma Novel memiliki tahapan yang hampir sama dengan *Fuzzy time series* klasik dari penentuan semesta pembicaraan hingga penentuan *Fuzzy Logical Relationship Group* (FLRG). Perbedaannya terdapat pada tahap perhitungan hasil prediksinya. Tujuan dari penelitian ini adalah meramalkan nilai ekspor Provinsi Jawa Tengah di masa yang akan datang dengan metode *Fuzzy time series* berbasis algoritma Novel serta mendapatkan tingkat akurasi peramalan menggunakan *Mean Absolute Percentage Error* (MAPE). Berdasarkan hasil analisis penerapan metode *Fuzzy time series* berbasis algoritma Novel pada data nilai ekspor di Provinsi Jawa Tengah diperoleh hasil ramalan pada bulan Januari 2022 sebesar 1054,178 juta US \$ serta nilai MAPE sebesar 11,653%.

Kata Kunci: *Ekspor, Fuzzy time series berbasis algoritma Novel, MAPE.*

ABSTRACT

Exports play an important role in the economy of a region. Exports can be described by the value of exports, namely the value in the form of money resulting from the sale of goods/services abroad. The value of exports needs to be forecasted to formulate targets for economic growth and establish appropriate policies in the economy of a region. The export value can be predicted using the Fuzzy time series method based on the Novel algorithm developed by Haneen Talal Jasim, Abdul Ghafoor Jasim Salim, and Kais Ismail Ibraheem. The stages of the Novel algorithm-based Fuzzy time series method have almost the same stages as the classic fuzzy time series from determining the universe of conversation to determining the Fuzzy Logical Relationship Group (FLRG). The difference is in the stage of calculating the prediction results. The purpose of this study is to predict the export value of Central Java Province in the future using the Fuzzy time series method based on the Novel algorithm and to get the level of forecasting accuracy using Mean Absolute Percentage Error (MAPE). Based on the results of the analysis of the application of the Fuzzy time series method based on the Novel algorithm on export value data in Central Java Province, the forecast results in January 2022 of 1054.178 million US \$ and the MAPE value of 11.653%.

Keyword: *Export, a Novel algorithm Fuzzy time series, MAPE.*

1. PENDAHULUAN

Perekonomian suatu negara adalah hal yang paling penting dan erat kaitannya dengan pertumbuhan ekonomi negara tersebut. Salah satu cara untuk meningkatkan pertumbuhan ekonomi suatu negara yaitu melalui perdagangan internasional. Perdagangan internasional adalah hubungan antar negara dalam kegiatan ekonomi yang diwujudkan dengan proses pertukaran barang

atau jasa secara sukarela dan saling menguntungkan (Aprita & Adhitya, 2020). Pelaksanaan dari kegiatan perdagangan internasional dapat terjadi dalam bentuk ekspor dan impor.

Menurut Kementerian Perdagangan Republik Indonesia, ekspor merupakan kegiatan mengeluarkan barang dari daerah pabean Indonesia ke daerah pabean negara lain, sedangkan impor adalah kegiatan memasukkan barang kedalam

daerah pabean Indonesia (Benny, 2013). Ekspor dan impor mengambil peran penting dalam kestabilan perekonomian suatu negara karena mempengaruhi jumlah devisa suatu negara. Ekspor dan impor merupakan upaya memenuhi kebutuhan masyarakat dan menambah pendapatan devisa bagi negara dalam mensejahterakan kehidupan masyarakat, dimana keduanya saling berkaitan satu sama lain dalam mempengaruhi perekonomian (Sumantri & Latifah, 2021).

Sedyaningrum, dkk (2016) menyatakan bahwa Nilai Tukar Rupiah terhadap Dollar AS akan menguat sejalan dengan ekspor yang mengalami peningkatan. Sedangkan Nilai Tukar Rupiah akan melemah ketika ekspor mengalami penurunan. Nilai ekspor yang tinggi akan meningkatkan produktifitas dalam negeri, yang mengakibatkan tenaga kerja dapat terserap secara penuh dan mengurangi jumlah pengangguran. Jika jumlah pengangguran suatu negara berkurang, maka akan meningkatkan pendapatan perkapita sehingga daya beli masyarakat akan meningkat (Sedyaningrum, Suhadak, & Nuzula, 2016).

Provinsi Jawa Tengah adalah provinsi dengan jumlah penduduk terbanyak ke-3 di Indonesia yakni berjumlah 36.516.035 penduduk menurut hasil sensus penduduk 2020. Jumlah penduduk di Provinsi Jawa Tengah belum seimbang dengan posisi Jawa Tengah yang menduduki peringkat ke-9 sebagai daerah penyumbang ekspor terbesar di Indonesia. Masalah perekonomian di Provinsi Jawa Tengah tentunya harus mendapat perhatian lebih dari pemerintah. Perekonomian di Provinsi Jawa Tengah dapat digambarkan salah satunya menggunakan nilai ekspor Provinsi Jawa Tengah yang dapat dilihat melalui website resmi Badan Pusat Statistik (BPS) Provinsi Jawa Tengah (BPS Jawa Tengah, 2020).

Pemerintah tentunya harus merumuskan target pertumbuhan ekonomi dan menetapkan kebijakan yang tepat dalam perekonomian di Provinsi Jawa Tengah baik dalam jangka pendek maupun dalam jangka panjang. Salah satunya adalah dengan melakukan peramalan. Peramalan adalah proses perkiraan besarnya atau jumlah sesuatu pada waktu yang akan datang berdasarkan data pada masa lampau yang dianalisis secara alamiah khususnya menggunakan metode statistika (Muhammad Wahdeni Pramana, Purnamasari, & Prangga, 2021).

Fuzzy Time Series (FTS) adalah peramalan data yang menggunakan himpunan *fuzzy* sebagai dasar peramalan. Peramalan dengan FTS dapat merekam pola dari data masa lalu untuk memprediksi data yang akan datang. Seiring

berkembangnya waktu, banyak model pengembangan dari Metode FTS yang memiliki hasil prediksi lebih akurat. Algoritma Novel pada metode FTS yang dikembangkan oleh Jasmin, dkk merupakan salah satu pengembangan metode FTS. Perbedaan metode FTS berbasis Algoritma Novel dengan metode FTS klasik terletak pada langkah perhitungan hasil prediksi setelah diperoleh *Fuzzy Logical Relations Group* (Jasim, Salim, & Ibraheem, 2012; Meirissa, 2016). Meirissa (2016) mengatakan bahwa metode FTS berbasis algoritma Novel yang dikembangkan oleh Jasmin, dkk (2012) lebih akurat dibandingkan dengan Metode FTS klasik.

Beberapa penelitian terkait dengan Metode FTS berbasis algoritma Novel adalah Meirissa (2016) dengan judul Penerapan algoritma Novel berdasarkan metode *Fuzzy time series* untuk memprediksi nilai tukar petani. Penelitian ini menghasilkan kesimpulan bahwa peramalan FTS berbasis algoritma Novel lebih akurat dibandingkan metode FTS klasik dengan nilai MSE sebesar 0,01 dan MAPE sebesar 0,08. Penelitian lainnya juga telah dilakukan oleh Riyadli (2016) dengan judul Analisis Perbandingan Logika *Fuzzy Time Series* sebagai Metode Peramalan. Hasil penelitian menunjukkan bahwa metode FTS berbasis algoritma Novel lebih baik dibandingkan metode FTS Ruy Chyn Tsaur dengan kesalahan prediksi sebesar 0,33%.

2. METODE PENELITIAN

2.1 Peramalan

Peramalan merupakan dugaan atau perkiraan mengenai terjadinya suatu kejadian atau peristiwa di waktu yang akan datang (Afdianti Nurkhasanah, Suparti, & Sudarno, 2015; Haris, 2020). Peramalan ini sangat berguna dalam berbagai aspek kehidupan, terutama dalam rangka perencanaan untuk menghadapi bahkan mengantisipasi berbagai keadaan yang terjadi pada masa yang akan datang. Peramalan dibedakan menjadi 3 berdasarkan jangka waktu peralamannya, yaitu peramalan jangka pendek, peramalan jangka Menengah, dan peramalan jangka panjang. FTS termasuk FTS berbasis algoritma Haneen Talal Jasim termasuk kategori peramalan jangka pendek karena hanya dapat meramalkan sebanyak 1 periode kedepan (Muhammad Wahdeni Pramana et al., 2021).

2.2 Fuzzy Time Series berbasis Algoritma Novel

Fuzzy Time Series (FTS) berbasis algoritma Novel merupakan pengembangan dari Metode FTS

klasik yang dikemukakan oleh Haneen Talal Jasim, Abdul Ghafoor Jasim Salim, dan Kais Ismail Ibraheem pada Tahun 2012 untuk meramalkan pendaftaran Universitas Alabama, Mosul, Irak (Jasim et al., 2012). Tahapan metode FTS berbasis algoritma Novel adalah sebagai berikut:

Langkah 1: Mengumpulkan data

Langkah 2: Menentukan nilai maksimum dan nilai minimum dari interval :

$$[D_{\min} - D_1, D_{\max} + D_2] \quad (1)$$

dengan D_1 dan D_2 merupakan nilai konstanta untuk menentukan *Universe of discourse* U (Tsaur, 2012).

Langkah 3: Penentuan Interval I menggunakan metode *average based length* (Riyadli, 2016) sebagai berikut :

- a. Menemukan perbedaan D_{vt} , D_{vt-1} kemudian temukan sisa perbedaan pertama, lalu temukan rata-rata perbedaan pertama.

$$av = \frac{\sum_{i=1}^n (D_i - D_{i-1})}{n-1} \quad (2)$$

dengan n adalah banyaknya data.

- b. Menghitung nilai tengah menggunakan rumus:

$$B = \frac{av}{2} \quad (3)$$

- c. Menentukan range dari hasil B berdasarkan tabel 1 di bawah ini:

Tabel 1. Tabel basis

Jangkauan	Basis
0,1 - 1	0,1
1 - 10	1
10 - 100	10
100 - 1000	100
1000 - 10.000	1000

Langkah 4. Menentukan nilai dari interval fuzzy menggunakan rumus sebagai berikut:

$$m = \frac{(D_{\max} + D_1 - D_{\min} + D_2)}{I} \quad (4)$$

Langkah 5. Menentukan himpunan *Fuzzy logical* dengan cara, $A_i = (d_{i-1}, d_i, d_{i+1}, d_{i+2})$ dimulai dari $A_1 = (d_0, d_1, d_2, d_3)$ dan berakhir pada $A_m = (d_{m-1}, d_m, d_{m+1}, d_{m+2})$, di mana $d_0 = D_{\min} - l$ dan $d_{m+2} = D_{\max}$ dan fuzzifikasi data historis menandakan nilai linguistik data diwakili oleh satu set himpunan fuzzy $1 \leq l \leq m$.

Langkah 6. Fuzzifikasi data historis menandakan nilai linguistik data diwakili oleh satu set himpunan fuzzy $1 \leq I \leq m$.

Langkah 7. Menentukan *Fuzzy Logical Relations* sebagai berikut : $A_i \rightarrow A_j$

Langkah 8. Menentukan *Fuzzy Logical Relations Group*

Langkah 9. Menghitung nilai peramalan pada waktu t berdasarkan algoritma Novel sebagai berikut (Meirissa, 2016):

- a. Jika *Fuzzy Logical Relationship Group* (FLRG) A_j adalah kosong $A_j \rightarrow \emptyset$, maka nilai F_{vt} adalah nilai tengah dari interval A_j dengan $A_j = (d_{j-1}, d_j, d_{j+1}, d_{j+2})$.
- b. Jika FLRG A_j adalah relasi *one to one* $A_j \rightarrow A_k$ maka interval yang mengandung nilai peramalan adalah A_k , untuk memperoleh hasil prediksi langkah pertama adalah membandingkan nilai j dan i yang terdapat pada *current state* dan *next state*. Kemudian menentukan nilai Y dengan rumus sebagai berikut (Jasim et al., 2012):

$$Y = (D_t - D_{t-1}) - (D_{t-1} - D_{t-2}) \quad (5)$$

dengan,

D_t : data pada periode t

D_{t-1} : data pada periode $t-1$

D_{t-2} : data pada periode $t-2$

menggunakan ketentuan sebagai berikut:

1. Jika $j > i$, dan $Y > 0$, maka kecenderungan peramalan akan menaik dan menggunakan aturan 2 untuk meramalkan data tersebut.
2. Jika $j > i$, dan $Y < 0$, maka kecenderungan peramalan akan menurun dan menggunakan aturan 3 untuk meramalkan data tersebut.
3. Jika $j < i$, dan $Y > 0$, maka kecenderungan peramalan akan menaik dan menggunakan aturan 2 untuk meramalkan data tersebut.
4. Jika $j < i$, dan $Y < 0$, maka kecenderungan peramalan akan menurun dan menggunakan aturan 3 untuk meramalkan data tersebut.
5. Jika $j = i$, dan $Y > 0$, maka kecenderungan peramalan akan menaik dan menggunakan aturan 2 untuk meramalkan data tersebut.
6. Jika $j = i$, dan $Y < 0$, maka kecenderungan peramalan akan menurun dan menggunakan aturan 3 untuk meramalkan data tersebut.

Aturan 1,2, dan 3 yang digunakan untuk memprediksi adalah sebagai berikut (Jasim et al., 2012):

Aturan 1: Digunakan jika Y tidak terpenuhi, ketentuannya sebagai berikut:

- 1) Jika nilai $\frac{|D_t - D_{(t-1)}|}{2} > \frac{A_j}{2}$, maka kecenderungan peramalan pada interval ini menaik dan $F_n = 0,75$ dari A_j .
- 2) Jika nilai $\frac{|D_t - D_{(t-1)}|}{2} = \frac{A_j}{2}$, maka peramalan adalah nilai tengah interval.
- 3) Jika nilai $\frac{|D_t - D_{(t-1)}|}{2} < \frac{A_j}{2}$, maka kecenderungan peramalan pada interval ini menurun dan $F_n = 0,25$ dari A_j .

Aturan 2:

- 1) Jika $x = |Y| * 2 + D_{(t-1)} \in A_j$ atau $x = D_{(t-1)} - |Y| * 2 \in A_j$, maka kecenderungan peramalan pada interval ini menaik dan $F_n = 0,75$ dari A_j .
- 2) Jika $x = \frac{|Y|}{2} + D_{(t-1)} \in A_j$ atau $x = D_{(t-1)} - \frac{|Y|}{2} \in A_j$, maka kecenderungan peramalan menurun dan $F_n = 0,25$ dari A_j .
- 3) Jika bukan keduanya, maka hasil peramalan adalah nilai tengah dari interval A_j .

Aturan 3:

- 1) Jika $x = \frac{|Y|}{2} + D_{(t-1)} \in A_j$ atau $x = D_{(t-1)} - \frac{|Y|}{2} \in A_j$, maka kecenderungan peramalan pada interval ini menurun dan $F_n = 0,25$ dari A_j .
 - 2) Jika $x = |Y| * 2 + D_{(t-1)} \in A_j$ atau $x = D_{(t-1)} - |Y| * 2 \in A_j$, maka kecenderungan peramalan pada interval ini menaik dan $F_n = 0,75$ dari A_j .
 - 3) Jika bukan keduanya, maka hasil peramalan adalah nilai tengah dari interval A_j .
- c. Jika FLRG A_j adalah relasi *one to many* $A_j \rightarrow A_{k1}, A_{k2}, \dots, A_{kp}$, maka peramalan mengikuti ketentuan sebagai berikut :
- 1) Jika perbedaan antara dua $k1, k2, \dots, kp \leq 2$ maka hasil peramalannya adalah nilai tengah interval tersebut.
 - 2) Jika perbedaan antara dua $k1, k2, \dots, kp > 2$ maka hasil prediksi dihitung menggunakan one to one fuzzy logical relationship dengan mengaplikasikan ketentuan B aturan 2.

2.3 Ukuran Akurasi Peramalan

Tujuan dari analisis runtun waktu (*time series*) adalah untuk meramalkan atau memprediksi nilai masa depan. Metode peramalan yang bertujuan untuk menghasilkan ramalan optimum yang memiliki tingkat kesalahan kecil atau akurasi yang tinggi. Menghitung tingkat akurasi setiap model peramalan digunakan metode uji yaitu MAPE (*Mean Absolute Percentage Error*). Semakin tinggi nilai MAPE yang dihasilkan maka akurasi peramalannya semakin rendah dan berlaku sebaliknya, semakin rendah nilai MAPE maka akurasi dari peramalannya semakin tinggi (Dewi & Ibnu Hadi, 2019).

$$MAPE = \frac{1}{N} \sum_{t=1}^n \left[\frac{X_t - F_t}{X_t} \right] \times 100\% \quad (6)$$

dengan N merupakan jumlah data yang digunakan, X_t merupakan data aktual pada periode ke t dan F_t merupakan data hasil peramalan pada periode ke t .

2.4 Sumber Data

Data penelitian diperoleh dari Badan Pusat Statistika (BPS) Provinsi Jawa Tengah. Data tersebut merupakan data nilai ekspor (juta US \$) yang diukur bulanan dari bulan Januari 2010 hingga Desember 2021.

2.5 Langkah-Langkah Penelitian

Langkah-langkah penelitian untuk melakukan peramalan nilai ekspor provinsi jawa tengah dengan metode FTS berbasis algoritma Novel adalah sebagai berikut:

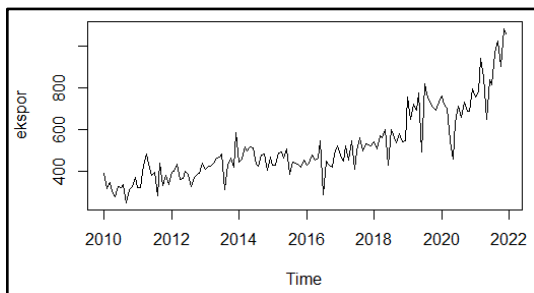
1. Melakukan analisis deskriptif yang bertujuan untuk melihat keadaan atau gambaran data yang sesungguhnya.
2. Melakukan peramalan *Fuzzy time series* berbasis algoritma Novel dengan langkah sebagai berikut:
 - a. Menentukan *universe of discourse* (Semesta pembicaraan U);
 - b. Menentukan panjang interval;
 - c. Menentukan jumlah interval;
 - d. Menentukan himpunan fuzzy;
 - e. Melakukan fuzzifikasi data;
 - f. Menentukan *Fuzzy Logical Relationship* (FLR);
 - g. Menentukan *Fuzzy Logical Relationship Group* (FLRG);
 - h. Menghitung hasil peramalan nilai ekspor menggunakan *Fuzzy time series* berbasis algoritma Novel.

3. Melakukan evaluasi hasil peramalan menggunakan nilai MAPE.

3. HASIL DAN PEMBAHASAN

3.1 Statistika Deskriptif

Nilai ekspor Provinsi Jawa Tengah pada periode Januari 2010 hingga Desember 2021 secara umum menunjukkan bahwa nilai ekspor tertinggi sebesar 1083,13 Juta US \$ pada November 2021 sedangkan nilai ekspor terendah sebesar 249,2 Juta US \$ pada September 2010 dengan rata-rata nilai ekspor sebesar 518,4097 Juta US \$. Plot deret waktu nilai ekspor Provinsi Jawa Tengah ditunjukkan pada gambar 1.



Gambar 1. Plot deret waktu nilai ekspor Provinsi Jawa Tengah 2010-2021

Kondisi nilai ekspor Provinsi Jawa Tengah pada periode Januari 2010 hingga Desember 2021 membentuk pola siklis yang terjadi akibat data yang dipengaruhi fluktuatif jangka panjang. Berbagai faktor yang mempengaruhi fluktuasi tersebut diantaranya seperti naik turunnya produksi dalam negeri maupun naik turunnya permintaan pasar internasional.

3.2 Mendefinisikan Universe of Discourse

Nilai minimal (D_{min}) dan maksimal (D_{max}) dari data Nilai Ekspor Provinsi Jawa Tengah adalah sebesar 249,2 dan 1083,13. D_1 dan D_2 merupakan suatu konstanta atau bilangan positif yang telah sesuai untuk membulatkan himpunan semesta U , sehingga diambil nilai $D_1 = 0,20$ dan $D_2 = 0,87$ untuk membulatkan himpunan semesta U . Sehingga himpunan semesta diperoleh seperti berikut:

$$\begin{aligned}
 U &= [D_{min} - D_1 ; D_{max} + D_2] \\
 &= [249,2 - 0,20 ; 1083,13 + 0,87] \\
 &= [249 ; 1084]
 \end{aligned}$$

3.3 Menentukan Panjang Interval

Metode yang digunakan untuk menentukan panjang interval (I) adalah metode *average based*

(av), dengan adalah menghitung seluruh nilai selisih mutlak antara D_{t+1} dan D_t untuk $t = 1, \dots, 144$. Kemudian menghitung rata-rata nilai selisih mutlak dari data nilai ekspor sesuai persamaan 2, sehingga diperoleh nilai av adalah 61,13371. Nilai B diperoleh dengan membagi nilai av dengan 2 sehingga didapatkan nilai 30,56685. Berdasarkan hasil tersebut, maka basis interval yang digunakan adalah 10. Kemudian 30,56685 dibulatkan berdasarkan basis, sehingga diperoleh panjang interval (I) sebesar 30.

3.4 Menentukan Jumlah Interval

Penentuan jumlah interval menggunakan persamaan (4) berikut:

$$\begin{aligned}
 m &= \frac{D_{max} + D_1 - D_{min} + D_2}{I} \\
 m &= \frac{1083,13 + 0,20 - 249,2 + 0,87}{30} \\
 m &= 27,83333 \approx 28
 \end{aligned}$$

Hasil di atas menunjukkan jumlah interval (m) sebanyak 28 interval. Kemudian himpunan semesta U dibagi menjadi 28 interval dengan panjang masing-masing interval sebesar 30 yang disajikan pada tabel berikut:

Tabel 2. Interval himpunan semesta

Interval	
$U_1 = [249,0000 ; 278,8214]$	$U_{15} = [666,5000 ; 696,3214]$
$U_2 = [278,8214 ; 308,6429]$	$U_{16} = [696,3214 ; 726,1429]$
$U_3 = [308,6429 ; 338,4643]$	$U_{17} = [726,1429 ; 755,9643]$
$U_4 = [338,4643 ; 368,2857]$	$U_{18} = [755,9643 ; 785,7857]$
$U_5 = [368,2857 ; 398,1071]$	$U_{19} = [785,7857 ; 815,6071]$
$U_6 = [398,1071 ; 427,9286]$	$U_{20} = [815,6071 ; 845,4286]$
$U_7 = [427,9286 ; 457,7500]$	$U_{21} = [845,4286 ; 875,2500]$
$U_8 = [457,7500 ; 487,5714]$	$U_{22} = [875,2500 ; 905,0714]$
$U_9 = [487,5714 ; 517,3929]$	$U_{23} = [905,0714 ; 934,8929]$
$U_{10} = [517,3929 ; 547,2143]$	$U_{24} = [934,8929 ; 964,7143]$
$U_{11} = [547,2143 ; 577,0357]$	$U_{25} = [964,7143 ; 994,5357]$
$U_{12} = [577,0357 ; 606,8571]$	$U_{26} = [994,5357 ; 1024,3571]$
$U_{13} = [606,8571 ; 636,6786]$	$U_{27} = [1024,3571 ; 1054,1786]$
$U_{14} = [636,6786 ; 666,5000]$	$U_{28} = [1054,1786 ; 1084,0000]$

3.5 Fuzzifikasi Data

Proses fuzzifikasi merupakan proses menentukan nilai linguistik dari setiap data nilai ekspor. Nilai linguistik dari suatu data ditentukan berdasarkan interval yang telah dibentuk, kemudian membandingkan derajat keanggotaan yang sesuai dan memilih derajat keanggotaan 1. Sebagai contoh data pada bulan Januari 2010 sebesar 387,36 nilai

tersebut termasuk dalam interval $U_5 = [368,2857; 398,1071]$, maka fuzzifikasi dari data tersebut adalah A_5 .

3.6 Menentukan Fuzzy Logical Relationship (FLR)

Berdasarkan hasil fuzzifikasi, dapat ditentukan FLR dengan menghubungkan data pada bulan sebelumnya dengan data pada bulan selanjutnya dalam bentuk $A_i \rightarrow A_j$ yang mana A_i disebut sisi kiri (*current state*) dan A_j sisi kanan (*next state*). Misalnya data pada periode Januari 2010 fuzzifikasi 5 dan data pada periode Februari 2010 fuzzifikasi 3 sehingga terbentuk FLR ($A_5 \rightarrow A_3$) dan seterusnya.

3.7 Menentukan Fuzzy Logical Relationship Group (FLRG)

FLRG dibentuk untuk mempermudah proses perhitungan prediksi yaitu dengan mengelompokkan setiap perpindahan *state*. Berikut adalah FLRG untuk seluruh data nilai ekspor di Jawa Tengah periode Januari 2010 hingga Desember 2021.

Tabel 3. Fuzzy logical relationship group

Current State	Next State
$A_1 \rightarrow$	$2(A_3)$
$A_2 \rightarrow$	$A_1, 2(A_7)$
$A_3 \rightarrow$	$A_1, 4(A_3), A_4, 4(A_5), 2(A_7)$
$A_4 \rightarrow$	A_2, A_4, A_6
$A_5 \rightarrow$	$A_2, 4(A_3), 3(A_5), A_6, 2(A_7)$
$A_6 \rightarrow$	$A_5, 2(A_6), 4(A_7), 3(A_8), A_9, 2(A_{12})$
$A_7 \rightarrow$	$A_3, A_4, A_5, 4(A_6), 5(A_7), 6(A_8), A_9, A_{10}, A_{11}, A_{13}$
$A_8 \rightarrow$	$A_3, 3(A_6), 3(A_7), 3(A_8), A_9, 3(A_{10})$
$A_9 \rightarrow$	$A_5, A_7, A_8, 3(A_9), A_{10}, 2(A_{11}), A_{20}$
$A_{10} \rightarrow$	$A_2, 2(A_7), A_8, 2(A_9), 4(A_{10}), A_{12}, A_{18}$
$A_{11} \rightarrow$	$A_6, A_9, A_{10}, A_{11}, A_{12}$
$A_{12} \rightarrow$	A_6, A_7, A_{10}, A_{11}
$A_{13} \rightarrow$	A_{16}
$A_{14} \rightarrow$	A_{16}, A_{17}, A_{20}
$A_{15} \rightarrow$	$A_{15}, A_{17}, A_{18}, A_{19}$
$A_{16} \rightarrow$	$A_{10}, A_{14}, 2(A_{15}), 2(A_{17})$
$A_{17} \rightarrow$	A_{15}, A_{16}, A_{18}
$A_{18} \rightarrow$	$A_{10}, A_{14}, A_{16}, A_{18}, A_{24}$
$A_{19} \rightarrow$	A_{18}
$A_{20} \rightarrow$	A_{17}, A_{20}, A_{25}
$A_{21} \rightarrow$	A_{14}
$A_{22} \rightarrow$	A_{28}

$A_{23} \rightarrow$	-
$A_{24} \rightarrow$	A_{21}
$A_{25} \rightarrow$	A_{27}
$A_{26} \rightarrow$	-
$A_{27} \rightarrow$	A_{22}
$A_{28} \rightarrow$	A_{28}

3.8 Menghitung hasil prediksi

Berdasarkan ketentuan dan aturan pada metode FTS berbasis algoritma Novel, hasil prediksi dihitung menggunakan data histori sebelumnya dalam bentuk FLR dan FLRG. Sebagai contoh, untuk memprediksi data pada bulan Februari 2010, dapat dilihat bahwa FLR dari bulan Januari 2010 menuju bulan Februari 2010 yang mana *state* bertransisi dari $A_5 \rightarrow A_3$, sehingga FLR A_5 adalah satu ke satu, maka untuk memprediksi menggunakan ketentuan 2. Langkah berikutnya adalah menghitung nilai Y berdasarkan persamaan (5), tetapi karena data pada bulan Desember 2009 tidak diketahui maka nilai Y tidak dapat dihitung, sehingga hasil prediksi pada bulan Februari 2010 menggunakan aturan 1. Pada aturan 1 langkah pertama adalah menghitung nilai x dengan rumus sebagai berikut:

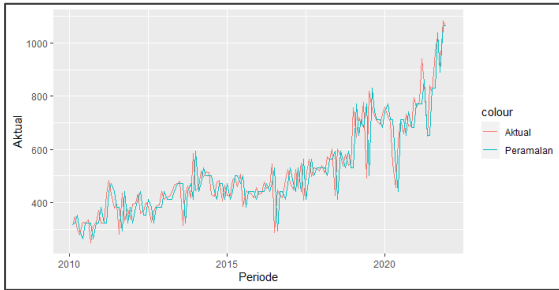
$$x = \frac{|D_t - D_{t-1}|}{2}$$

$$x = \frac{|314,47 - 387,36|}{2}$$

$$x = 36,445$$

Diperoleh nilai x sebesar 36,445. Kemudian menghitung setengah dari interval A_3 , berdasarkan himpunan fuzzy yang dibentuk dapat diperoleh interval A_3 adalah $U_3 = [308,6429 ; 338,4643]$, sehingga setengah dari interval A_3 sebesar 323,5536 atau sama dengan nilai tengah U_3 . Dari perhitungan tersebut diperoleh $x < \frac{A_3}{2}$. Berdasarkan aturan 1, maka hasil prediksi bulan Februari 2010 cenderung turun dan jatuh pada 0,25 poin dari interval A_3 yaitu 316,1.

Menggunakan cara yang sama berdasarkan aturan peramalan berdasarkan metode FTS berbasis algoritma Novel didapatkan plot antara data aktual dan data peramalan nilai ekspor Provinsi Jawa Tengah periode Januari 2010 hingga Desember 2021 sebagai berikut:



Gambar 2. Plot data aktual dan peramalan nilai ekspor di Provinsi Jawa Tengah periode Januari 2010-Desember 2021

Peramalan pada bulan selanjutnya, yaitu bulan Januari 2022 diketahui *current state* adalah A_{28} dan *next state* adalah himpunan kosong ($A_{28} \rightarrow \emptyset$). Sehingga untuk menghitung hasil prediksi digunakan ketentuan 1 yaitu nilai peramalan berada ditengah interval *Fuzzy* dengan $A_i = (d_{i-1}, d_i, d_{i+1}, d_{i+2})$ yaitu $A_{28} = (d_{27}, d_{28}, d_{29}, d_{30})$. Karena tidak terdapat nilai d_{29} dan d_{30} maka perhitungannya sebagai berikut :

$$\begin{aligned} \text{Januari 2022} &= \frac{d_{27} + d_{28}}{2} \\ &= \frac{1039,2679 + 1069,0893}{2} \\ &= 1054,178 \end{aligned}$$

Didapatkan hasil prediksi nilai ekspor Provinsi Jawa Tengah pada bulan Januari 2022 adalah sebesar 1054,178 juta US \$.

3.9 Ketepatan Metode Peramalan

Pengujian tingkat akurasi dilakukan dengan melihat kriteria pengukuran *Mean Absolute Percentage Error* (MAPE). Dari hasil tersebut dapat diketahui apakah metode yang digunakan tidak bias dan tingkat akurasi tinggi untuk dapat digunakan untuk memprediksi suatu nilai.

Tabel 3. Akurasi Peramalan

Metode	MAPE (%)
FTS berbasis algoritma Novel	11,653

Tabel 3 di atas menunjukkan bahwa hasil prediksi data nilai ekspor di Provinsi Jawa Tengah pada Januari 2010 hingga Desember 2021 dengan menggunakan FTS berbasis algoritma Novel diperoleh nilai MAPE sebesar 11,653%. Maka dapat dikatakan bahwa hasil peramalan baik dan dapat

digunakan untuk memprediksi nilai ekspor pada bulan selanjutnya yaitu bulan Januari 2022.

4. KESIMPULAN DAN SARAN

Hasil penelitian prediksi data nilai ekspor di Provinsi Jawa Tengah dengan menggunakan metode FTS berbasis algoritma Novel menghasilkan kesimpulan sebagai berikut:

1. Hasil prediksi nilai ekspor di Provinsi Jawa Tengah pada bulan Januari 2022 dengan menggunakan metode FTS berbasis Algoritma Novel yaitu sebesar 1054,178 juta US \$.
2. Metode Fuzzy Time Series berbasis Algoritma Novel mempunyai hasil peramalan yang akurat pada kasus Nilai Ekspor di Provinsi Jawa Tengah. Karena memiliki nilai MAPE (Mean Absolute Percentage Error) sebesar 11,653% sehingga dapat dikatakan baik untuk meramalkan nilai ekspor Provinsi Jawa Tengah.

DAFTAR PUSTAKA

Afdianti Nurkhasanah, L., Suparti, & Sudarno. (2015). Perbandingan Metode Runtun Waktu Fuzzy-Chen dan Fuzzy-Markov Chain untuk Meramalkan Data Inflasi di Indonesia. *Jurnal Gaussian*, 4(4).

Aprita, S., & Adhitya, R. (2020). *Hukum Perdagangan Internasional* (R. Mirsawati, ed.). Depok: Rajawali Pers.

Benny, J. (2013). Ekspor dan Impor Pengaruhnya Terhadap Posisi Cadangan Devisa di Indonesia. *Jurnal EMBA*, 1(4), 1406–1415. https://doi.org/10.1007/springerreference_67401

BPS Jawa Tengah. (2020). Provinsi Jawa Tengah Dalam Angka. In *Badan Pusat Statistik*. Badan Pusat Statistik Provinsi Jawa Tengah.

Dewi, E. R., & Ibnu Hadi. (2019). Peramalan Produk Domestik Bruto (PDB) Industri Pengolahan Non Migas di Indonesia dengan menggunakan Metode Fuzzy Time Series. *Jurnal Statistika Dan Aplikasinya*, 3(2). <https://doi.org/10.21009/jsa.03203>

Haris, M. Al. (2020). Peramalan Harga Emas dengan Model Generalized Autoregressive Conditional Heteroscedasticity (GARCH). *Jurnal Sainika Unpam: Jurnal Sains Dan Matematika Unpam*, 3(1). <https://doi.org/10.32493/jsmu.v3i1.5263>

Jasim, H. T., Salim, A. G. J., & Ibraheem, K. I. (2012). A Novel Algorithm to Forecast Enrollment Based on Fuzzy Time Series. *Applications and Applied Mathematics: An International Journal (AAM)*, 7(1), 385–397.

- Meirissa, I. R. (2016). Penerapan Algoritma Novel Berdasarkan Metode Fuzzy Time Series Untuk Memprediksi Nilai Tukar Petani. *Revista CENIC. Ciencias Biológicas*, 152(3).
- Muhammad Wahdeni Pramana, Purnamasari, I., & Prangga, S. (2021). Peramalan Data Ekspor Nonmigas Provinsi Kalimantan Timur Menggunakan Metode Weighted Fuzzy Time Series Lee. *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, 14(1). <https://doi.org/10.36456/jstat.vol14.no1.a3747>
- Riyadli, H. (2016). Analisis Perbandingan Logika Fuzzy Time Series Sebagai Metode Peramalan. *Speed-Sentra Penelitian Engineering Dan Edukasi*, 8(1).
- Sedyaningrum, M., Suhadak, & Nuzula, N. F. (2016). Daya Beli Masyarakat di Indonesia Studi pada Bank Indonesia Periode Tahun 2006 : IV-2015 : III. *Jurnal Administrasi Bisnis*, 34(1), 114–121.
- Sumantri, F., & Latifah, U. (2021). Analisa Pengaruh Marko Ekonomi Terhadap Indeks Harga Saham Gabungan Periode 2015-2019. *DIALEKTIKA: Jurnal Ekonomi Dan Ilmu Sosial*, 6(2). <https://doi.org/10.36636/dialektika.v6i2.617>
- Tsaur, R. C. (2012). A fuzzy time series-Markov chain model with an application to forecast the exchange rate between the Taiwan and us Dollar. *International Journal of Innovative Computing, Information and Control*, 8(7 B).

Prediksi Pengeluaran Non Makanan Masyarakat Kabupaten Tulang Bawang Menggunakan Metode Rantai Markov

Ahmad Rizki Wiranto⁽¹⁾, Ahmad Faisol⁽²⁾, Fitriani⁽³⁾

Jurusan Matematika FMIPA Universitas Lampung

Jl. Prof. Dr. Soemantri Brodjonegoro No. 1 Bandar Lampung 35145

e-mail: ahmad.rizki100819@students.unila.ac.id, ahmadfaisol@fmipa.unila.ac.id dan fitriani.1984@fmipa.unila.ac.id

ABSTRAK

Rantai Markov (*Markov Chains*) merupakan suatu metode yang mempelajari sifat-sifat suatu variabel pada masa sekarang, yang didasarkan pada sifat-sifatnya di masa lalu, dalam usaha menaksir sifat-sifat variabel tersebut di masa yang akan datang. Dengan menggunakan metode *Markov Chains* diperoleh prediksi untuk pengeluaran perumahan dan fasilitas rumah tangga pada tahun 2022-2025 secara berturut-turut 52.5%, 50.3705%, 50.31882%, dan 50.31764%. Untuk prediksi pengeluaran aneka barang dan jasa pada tahun 2022-2025 adalah 24.2%, 23.131, %23.03519%, dan 23.03152%. Prediksi pengeluaran pakaian, alas kaki dan tutup kepala pada periode 2022-2025 sebesar 8.2%, 7.9442%, 7.9319%, dan 7.9913%. Sedangkan prediksi pengeluaran keperluan barang tahan lama tahun 2022-2025 sebesar 7.4%, 10.1995%, 10.34567%, dan 10.35161%. Prediksi pada tahun 2022-2025 untuk pengeluaran keperluan pajak, pungutan dan asuransi sebesar 4.6%, 5.3499%, 5.3642%, dan 5.36373%. Serta pengeluaran untuk keperluan pesta pada periode 2022-2025 sebesar 3.1%, 3.0049%, 3.00422%, dan 3.00417%. Prediksi pengeluaran barang tahan lama setiap tahunnya selalu mengalami peningkatan. Faktor keperluan jangka panjang dapat menjadi alasan masyarakat mengeluarkan pengeluaran untuk barang tahan lama menjadi keperluan yang penting dan mengalami peningkatan setiap tahunnya. Selain itu, pengeluaran pajak, pungutan dan asuransi ikut meningkat. Sedangkan pengeluaran pakaian, alas kaki dan tutup kepala akan diprediksi meningkat di tahun 2025. Untuk keperluan Perumahan, dan fasilitas rumah tangga, pengeluaran aneka barang dan jasa dan keperluan pesta mengalami penurunan tiap tahunnya.

Kata kunci : COVID-19, Kesejahteraan, Pengeluaran Non Makanan, Prediksi, dan Markov Chains

ABSTRACT

Markov Chains is a method that studies the properties of a variable in the present, which is based on its properties in the past, in an effort to estimate the properties of these variables in the future. By using the Markov Chains method, predictions for housing and household facility expenditure in the 2022–2025 period are 52.5%, 50.3705%, 50.31882%, and 50.31764%, respectively. For predictions of expenditure for various goods and services in 2022-2025 are 24.2%, 23.131, %23.03519%, and 23.03152%. Expenditure predictions for clothing, footwear, and headgear in the 2022–2025 period are 8.2%, 7.9442%, 7.9319%, and 7.9913%. Meanwhile, the predicted expenditure for durable goods in 2022–2025 is 7.4%, 10.1995%, 10.34567%, and 10.35161%. Predictions for 2022-2025 for expenditures on taxes, levies, and insurance are 4.6%, 5.3499%, 5.3642%, and 5.36373%. And expenses for party purposes in the period 2022-2025 are 3.1%, 3.0049%, 3.00422%, and 3.00417%. The prediction of durable goods expenditures every year always increases. The factor of long-term needs can be the reason people spend on durable goods which are important needs and increase every year. In addition, spending on taxes, levies and insurance also increased. Meanwhile, expenditure on clothing, footwear and headgear is predicted to increase in 2025. For housing purposes, and household facilities, expenditure on various goods and services and party needs has decreased every year.

Keywords : COVID-19, Welfare, Non Food Expenditure, Prediction, Markov Chains

1. PENDAHULUAN

Dampak dari pandemi COVID-19 tidak hanya berpengaruh pada kesehatan masyarakat, perekonomian negarapun ikut terkena dampak. Bahkan saat ini perekonomian dunia mengalami tekanan berat yang diakibatkan oleh COVID-19 (Nugroho, 2020). Masalah perekonomian dipandang sebagai masalah utama yang harus segera dicarikan solusi untuk mengatasi krisis ini. Pemerintah telah berupaya dengan berbagai kebijakan strategis agar dapat memulihkan perekonomian di Indonesia sebagai implementasi dari amanat undang-undang dasar negara 1945 untuk mensejahterakan rakyat (Susanto, dkk., 2021).

Keberhasilan suatu negara dalam mencapai tujuannya dapat dilihat dari tingkat kesejahteraan penduduknya (Susanto, dkk., 2021). Banyak indikator yang dapat digunakan untuk melihat kesejahteraan penduduk, diantaranya dapat dilihat dari tingkat konsumsi dan pengeluaran. Menurut Badan Pusat Statistik (2012), Pengeluaran rumah tangga dapat dibagi menjadi dua, yaitu pengeluaran makanan dan pengeluaran non makanan. Semakin banyak presentase pengeluaran rumah tangga untuk membeli kelompok non makanan maka semakin tinggi kesejahteraan rumah tangga tersebut, dikarenakan kondisi perekonomiannya mengalami peningkatan atau pendapatan yang mulai meningkat (Praza & Shamdiyah, 2020).

Kabupaten Tulang Bawang merupakan kabupaten di Provinsi Lampung yang daerahnya cukup strategis. Hal ini dikarenakan Kabupaten Tulang Bawang dilalui oleh Lintas Sumatera yang membuat mobilitasnya cukup tinggi. Namun, selama Periode 2016 - 2020, Garis kemiskinan Kabupaten Tulang Bawang selalu mengalami peningkatan setiap tahunnya. Pada Maret 2016, Garis kemiskinan Kabupaten Tulang Bawang berada diangka Rp. 362.185,- perkapita perbulan, meningkat menjadi Rp 373.681,- perkapita perbulan, dan pada tahun 2020 menjadi Rp. 438.744,- perbulan. Peningkatan Garis kemiskinan tidak menjadikan persentase penduduk miskinnya meningkat, hal ini menunjukkan bahwa tingkat pendapatan sebgaiian penduduk miskin khususnya mereka yang berada di sekitar garis kemiskinan mampu mengimbangi kenaikan barang.

Peningkatan pendapatan idealnya diikuti dengan pemerataan pendapatan, karena pemerataan merupakan salah satu strategi dan tujuan pembangunan nasional. Ketimpangan dalam menikmati hasil pembangunan diantara kelompok penduduk dikhawatirkan akan menimbulkan

masalah-masalah sosial. Menurut BPS Kabupaten Tulang Bawang (2020), Secara umum tingkat konsumsi yang tergambar dalam pengeluaran per kapita penduduk Kabupaten Tulang Bawang di tahun 2020 mengalami penurunan sebesar 5,13 persen dibanding tahun 2019 lalu. Pengeluaran per kapita di tahun 2020 mencapai Rp 842,3 ribu, sedikit lebih rendah dibanding konsumsi per kapita tahun sebelumnya yang tercatat sebesar Rp 887,87 ribu. Penurunan pengeluaran per kapita tahun 2020 dibanding tahun 2019 terjadi pula untuk konsumsi makanan dan bukan makanan. Jika dilihat menurut jenis pengeluaran, terjadi penurunan pengeluaran per kapita makanan sebesar 1,88 persen sedangkan untuk bukan makanan sebesar 8,62 persen dibandingkan tahun sebelumnya.

Setiap tahun Pemerintah Kabupaten Tulang Bawang telah berusaha untuk meningkatkan taraf hidup masyarakatnya melalui pembangunan di berbagai bidang. Hal ini perlu dilakukan untuk melihat kualitas manusia dalam suatu daerah yang berguna untuk memajukan daerah tersebut. Menurut Suryono (2014), perlu adanya sistem kebijakan pemerintah yang terorganisasi dengan pelayanan dan lembaga sosial untuk mendorong individu dan kelompok agar mencapai tingkat hidup dan kesehatan yang sejahtera. Oleh karena itu, dalam merancang suatu kebijakan, perlu dilakukannya evaluasi dari kebijakan-kebijakan sebelumnya. Selain itu, diperlukan juga prediksi, agar nantinya dapat ditentukan target dan kebijakan tersebut dapat berjalan dengan baik dan memberikan dampak jangka panjang.

Markov Chains merupakan suatu proses random (proses stokastik) di mana keadaan (state) yang akan datang bersifat independen terhadap keadaan (state) yang lampau dan hanya tergantung pada keadaan yang terdekat sebelumnya (Pramuditya, dkk., 2014). Selain itu, *markov chains* dapat menganalisis pergerakan dari satu kondisi ke kondisi lainnya dengan menggunakan teknik probabilitas. Metode ini mempelajari sifat-sifat suatu variabel dari masa sekarang yang didasarkan pada sifat-sifat dari masa lalu untuk memprediksi sifat-sifat di masa yang akan datang (Putri & Muliawati, 2021). Oleh karena itu, penulis tertarik untuk memprediksi pengeluaran non makanan masyarakat Tulang Bawang dengan menggunakan metode Markov Chains.

2. METODE PENELITIAN

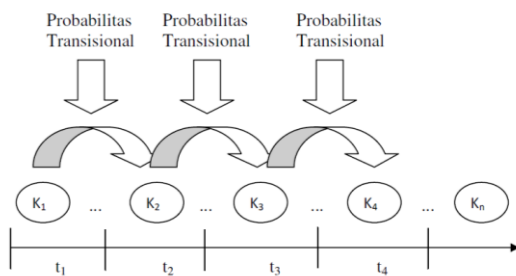
Sumber data yang diperoleh adalah data sekunder yang bersumber dari publikasi Statistik Kesejahteraan Rakyat Badan Pusat Statistik

Kabupaten Tulang Bawang tahun 2016-2021. Adapun metode dalam menyelesaikan penelitian ini adalah sebagai berikut:

1. Mengumpulkan data pengeluaran non makanan dari tahun 2016-2021 yang diperoleh dari BPS Kabupaten Tulang Bawang
2. Menentukan Matriks peluang transisi
3. Menentukan initial state
4. Melakukan prediksi dari tahun 2022-2025

Proses stokastik $\{X_n\}$ adalah kumpulan dari variabel acak. Indeks n diartikan sebagai waktu dan X_n adalah state pada waktu n . Himpunan n disebut ruang parameter atau himpunan indeks proses. Jika dilihat dari parameter waktunya, proses stokastik dibedakan menjadi dua yaitu proses stokastik waktu diskrit dan proses stokastik waktu kontinu. Bagian dari proses stokastik dengan himpunan variabel acak yang berkaitan dengan waktu merupakan rantai markov. Proses stokastik adalah barisan peristiwa yang memenuhi hukum-hukum peluang. Suatu nilai dikatakan mengikuti proses stokastik ketika nilai tersebut berubah terhadap waktu secara acak (Nurman dan Syata, 2021).

Metode ini termasuk ke dalam metode kuantitatif yang digunakan dalam perhitungan probabilitas perubahan-perubahan yang terjadi berdasarkan probabilitas perubahan selama periode waktu tertentu (Noeryanti, dkk., 2019). Rantai markov merupakan bagian dari proses stokastik yang memenuhi sifat markov, yaitu peluang kejadian atau peubah acak K pada waktu $t + 1$ hanya akan dipengaruhi oleh kejadian K pada waktu t dan tidak akan dipengaruhi oleh kejadian sebelum waktu t (Wahyudi, 2021).



Gambar 1. Proses rantai markov

Menurut Pramuditya, dkk. (2014), dalam penerapan Metode *markov chains*, ada beberapa asumsi yang harus terpenuhi yaitu:

- a. Jumlah probabilitas transisi untuk suatu keadaan awal dari sistem sama dengan 1.

- b. Probabilitas-probabilitas tersebut berlaku untuk semua partisipan dalam sistem.
- c. Probabilitas transisi konstan sepanjang waktu, artinya peluang untuk setiap keadaan dari periode $n \geq 0$ adalah sama.
- d. State independen sepanjang waktu.

Menurut Noeryanti, dkk. (2019), jika sebuah rantai markov $\{X_t, t = 0, 1, \dots\}$ dengan ruang state $\{0, 1, \dots, m\}$, maka peluang sistem itu dalam state i pada suatu state j pada pengamatan sebelumnya diberikan notasi P_{ij} dan disebut peluang transisi dari state i ke state j . Matriks $P = [P_{ij}] = P\{X_{n+1} = j | X_n = i\}$ disebut matriks transisi rantai markov dengan elemen-elemennya bernilai positif dan jumlah elemen-elemen pada suatu baris sama dengan 1. Berikut merupakan matriks P berukuran $m \times m$ untuk peluang transisi satu langkah:

$$P = \begin{matrix} \text{State} & 0 & 1 & \dots & m \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} p_{00} & p_{01} & \dots & p_{0m} \\ p_{10} & p_{11} & \dots & p_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m0} & p_{m1} & \dots & p_{mm} \end{bmatrix} \end{matrix}$$

dimana:

$i = \text{state ke-}i (0, 1, \dots, m)$

$j = \text{state ke-}j (0, 1, \dots, m)$

$$P_{ij} \geq 0 \text{ dan } \sum_{j=0}^m P_{ij} = 1; (i, j = 0, 1, 2, \dots, m)$$

Sedangkan untuk matriks P berukuran $m \times m$ dalam peluang transisi n -langkah

$$P^n = \begin{matrix} \text{state} & 0 & 1 & \dots & m \\ \begin{matrix} 0 \\ 1 \\ \vdots \\ m \end{matrix} & \begin{bmatrix} P_{00}^n & P_{01}^n & \dots & P_{0m}^n \\ P_{10}^n & P_{11}^n & \dots & P_{1m}^n \\ \vdots & \vdots & \ddots & \vdots \\ P_{m0}^n & P_{m1}^n & \dots & P_{mm}^n \end{bmatrix} \end{matrix}$$

Maka P_{ij}^n merupakan peluang bahwa proses pada state ke- i akan berada pada state ke- j setelah proses mengalami n tambahan transisi. Jadi

$$P_{ij}^n = P\{X_{n+t} = j | X_t = i\}, n, i, j \geq 0 \quad (1)$$

dimana:

$i = \text{state ke-}i (0, 1, \dots, m)$

$j = \text{state ke-}j (0, 1, \dots, m)$

$$P_{ij}^n \geq 0 \text{ dan } \sum_{j=0}^m P_{ij}^n = 1; (i, j = 0, 1, 2, \dots, m)$$

3. HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data pengeluaran non makanan masyarakat Kabupaten Tulang Bawang dari tahun 2016 sampai dengan tahun 2021.

Tabel 1. Pengeluaran Non Makanan Masyarakat Kabupaten Tulang Bawang 2016-2021

Tahun	Pengeluaran menurut Kelompok Non Makanan (Rupiah)						Jumlah
	Perumahan dan Fasilitas Rumah Tangga	Aneka Barang dan Jasa	Pakaian, Alas Kaki dan Tutup Kepala	Barang Tahan Lama	Pajak, Pungutan dan Asuransi	Keperluan Pesta	
2016	177219	81719	27742	25020	15362	10297	337359
2017	182196	83766	30071	47058	25371	11314	379776
2018	217357	113280	34535	95953	24243	15576	500944
2019	205320	86834	32194	67650	23370	12854	428222
2020	208305	86998	32065	26512	25556	11859	391295
2021	206466	91455	28728	32623	27079	4763	391114

Dari data pengeluaran non makanan masyarakat Kabupaten Tulang Bawang akan dibentuk matriks peluang transisi P

$$P = P_{ij} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} & P_{15} & P_{16} \\ P_{21} & P_{22} & P_{23} & P_{24} & P_{25} & P_{26} \\ P_{31} & P_{32} & P_{33} & P_{34} & P_{35} & P_{36} \\ P_{41} & P_{42} & P_{43} & P_{44} & P_{45} & P_{46} \\ P_{51} & P_{52} & P_{53} & P_{54} & P_{55} & P_{56} \\ P_{61} & P_{62} & P_{63} & P_{64} & P_{65} & P_{66} \end{bmatrix}$$

dimana:

- i=1, mewakilkan tahun 2016
- i=2, mewakilkan tahun 2017
- i=3, mewakilkan tahun 2018
- i=4, mewakilkan tahun 2019
- i=5, mewakilkan tahun 2020
- i=6, mewakilkan tahun 2021
- j=1, mewakilkan pengeluaran untuk perumahan dan fasilitas rumah tangga
- j=2, mewakilkan pengeluaran untuk aneka barang dan jasa
- j=3, mewakilkan pengeluaran untuk pakaian, alas kaki dan tutup kepala
- j=4, mewakilkan pengeluaran untuk barang tahan lama
- j=5, mewakilkan pengeluaran untuk pajak, pungutan dan asuransi
- j=6, mewakilkan pengeluaran untuk keperluan pesta

Sehingga matriks peluang transisi pengeluaran non makanan masyarakat Kabupaten Tulang Bawang sebagai berikut:

$$P = \begin{bmatrix} 0.525 & 0.242 & 0.082 & 0.074 & 0.046 & 0.031 \\ 0.480 & 0.220 & 0.079 & 0.124 & 0.067 & 0.030 \\ 0.434 & 0.226 & 0.069 & 0.192 & 0.048 & 0.031 \\ 0.479 & 0.203 & 0.075 & 0.158 & 0.055 & 0.030 \\ 0.533 & 0.222 & 0.082 & 0.065 & 0.065 & 0.030 \\ 0.528 & 0.234 & 0.074 & 0.083 & 0.069 & 0.012 \end{bmatrix}$$

Sebelum melakukan prediksi terlebih dahulu menentukan initial state. Dalam penelitian ini, peneliti menggunakan initial state sebagai berikut

$$\pi(0) = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$$

Prediksi Pengeluaran Non-Makanan tahun 2022

Dalam memprediksi pengeluaran non-makanan pada tahun 2022, maka akan digunakan initial state yang akan dikalikan dengan matriks peluang transisi P

$$\pi(1) = \pi(0)P$$

$$\pi(1) = [1 \ 0 \ 0 \ 0 \ 0 \ 0] \times \begin{bmatrix} 0.525 & 0.242 & 0.082 & 0.074 & 0.046 & 0.031 \\ 0.480 & 0.220 & 0.079 & 0.124 & 0.067 & 0.030 \\ 0.434 & 0.226 & 0.069 & 0.192 & 0.048 & 0.031 \\ 0.479 & 0.203 & 0.075 & 0.158 & 0.055 & 0.030 \\ 0.533 & 0.222 & 0.082 & 0.065 & 0.065 & 0.030 \\ 0.528 & 0.234 & 0.074 & 0.083 & 0.069 & 0.012 \end{bmatrix}$$

$$\pi(1) = [0.525 \ 0.242 \ 0.082 \ 0.074 \ 0.046 \ 0.031]$$

Kemudian untuk mengubah bentuk peluang menjadi persentase, maka hasil dari $\pi(1)$ dikalikan dengan 100%.

$$\pi(1) \times 100\% = [52.5\% \ 24.2\% \ 8.2\% \ 7.4\% \ 4.6\% \ 3.1\%]$$

Dari hasil prediksi pada tahun 2022, diperoleh bahwa prediksi pengeluaran untuk keperluan perumahan dan fasilitas rumah tangga sebesar 52.5% dari pengeluaran makanan. Sedangkan sebesar 24.2% digunakan untuk pengeluaran barang dan jasa serta 8.2% digunakan untuk pengeluaran pakaian, alas kaki, dan tutup kepala. Untuk pengeluaran barang tahan lama, pengeluaran pajak, pungutan dan asuransi, serta pengeluaran untuk keperluan pesta berturut-turut sebesar 7.4%, 4.6%, dan 3.1% dari pengeluaran non-makanan.

Prediksi Pengeluaran Non-Makanan tahun 2023

Dalam melakukan prediksi pengeluaran non makanan pada tahun 2023, maka akan digunakan $\pi(1)$ yang dikalikan dengan matriks P

$$\pi(2) = \pi(1) \times P$$

$$\pi(2) = [0.525 \ 0.242 \ 0.082 \ 0.074 \ 0.046 \ 0.031] \times \begin{bmatrix} 0.525 & 0.242 & 0.082 & 0.074 & 0.046 & 0.031 \\ 0.480 & 0.220 & 0.079 & 0.124 & 0.067 & 0.030 \\ 0.434 & 0.226 & 0.069 & 0.192 & 0.048 & 0.031 \\ 0.479 & 0.203 & 0.075 & 0.158 & 0.055 & 0.030 \\ 0.533 & 0.222 & 0.082 & 0.065 & 0.065 & 0.030 \\ 0.528 & 0.234 & 0.074 & 0.083 & 0.069 & 0.012 \end{bmatrix}$$

$$\pi(2) = [0.503705 \quad 0.23131 \quad 0.079442 \quad 0.101995 \quad 0.053499 \quad 0.030049]$$

$$\pi(2) \times 100\% = [50.3705\% \quad 23.131\% \quad 7.9442\% \quad 10.1995\% \quad 5.3499\% \quad 3.0049\%]$$

Sehingga diperoleh bahwa prediksi peluang pengeluaran untuk keperluan perumahan dan fasilitas rumah tangga adalah sebesar 50.3705% dan turun sebesar 2.1295% dari pengeluaran keperluan perumahan dan fasilitas rumah tangga pada tahun 2022. Pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, serta pengeluaran untuk keperluan pesta juga mengalami penurunan sebesar 1.069%, 0.2558%, dan 0.0951% dari tahun 2022. Dengan prediksi pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, serta pengeluaran untuk keperluan pesta pada tahun 2023 adalah sebesar 23.131%, 7.9442%, dan 3.0049%. Sedangkan prediksi pengeluaran barang tahan lama, serta pengeluaran pajak, pungutan dan asuransi sebesar 10.1995% dan 5.3499%, yang artinya mengalami kenaikan dari prediksi pada tahun 2022 sebesar 2.7995% dan 0.7499%.

Prediksi Pengeluaran Non-Makanan tahun 2024

Dalam melakukan prediksi pengeluaran non makanan pada tahun 2024, maka akan digunakan $\pi(2)$ yang dikalikan dengan matriks P

$$\pi(3) = \pi(2) \times P$$

$$\pi(3) = [0.503705 \quad 0.23131 \quad 0.079442 \quad 0.101995 \quad 0.053499 \quad 0.030049] \times \begin{bmatrix} 0.525 & 0.242 & 0.082 & 0.074 & 0.046 & 0.031 \\ 0.480 & 0.220 & 0.079 & 0.124 & 0.067 & 0.030 \\ 0.434 & 0.226 & 0.069 & 0.192 & 0.048 & 0.031 \\ 0.479 & 0.203 & 0.075 & 0.158 & 0.055 & 0.030 \\ 0.533 & 0.222 & 0.082 & 0.065 & 0.065 & 0.030 \\ 0.528 & 0.234 & 0.074 & 0.083 & 0.069 & 0.012 \end{bmatrix}$$

$$\pi(3) = [0.5031882 \quad 0.2303519 \quad 0.0793190 \quad 0.1034567 \quad 0.053642 \quad 0.0300422]$$

$$\pi(3) \times 100\% = [50.31882\% \quad 23.03519\% \quad 7.93190\% \quad 10.34567\% \quad 5.3642\% \quad 3.00422\%]$$

Pada prediksi tahun 2024, peluang pengeluaran yang mengalami peningkatan hanya pada pengeluaran barang tahan lama dan pengeluaran pajak, pungutan dan asuransi yang mengalami kenaikan sebesar 0.14617% dan 0.0143% dari pengeluaran barang tahan lama dan pengeluaran pajak, pungutan dan asuransi pada tahun 2023. Dengan prediksinya sebesar 10.34567% dan 5.3642%. Sedangkan prediksi pengeluaran untuk keperluan perumahan dan fasilitas rumah tangga, pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, serta pengeluaran untuk keperluan pesta kembali mengalami penurunan sebesar 0.05168%, 0.09581%, 0.0123%, dan 0.00068 dari prediksi pengeluaran untuk keperluan

perumahan dan fasilitas rumah tangga, pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, serta pengeluaran untuk keperluan pesta pada tahun 2023, dengan prediksinya sebesar 50.31882%, 23.03519%, 7.93190%, dan 3.00422%.

Prediksi Pengeluaran Non-Makanan tahun 2025

Dalam melakukan prediksi pengeluaran non makanan pada tahun 2025, maka akan digunakan $\pi(3)$ yang dikalikan dengan matriks P

$$\pi(4) = \pi(3) \times P$$

$$\pi(4) = [0.5031882 \quad 0.2303519 \quad 0.0793190 \quad 0.1034567 \quad 0.0536420 \quad 0.0300422] \times \begin{bmatrix} 0.525 & 0.242 & 0.082 & 0.074 & 0.046 & 0.031 \\ 0.480 & 0.220 & 0.079 & 0.124 & 0.067 & 0.030 \\ 0.434 & 0.226 & 0.069 & 0.192 & 0.048 & 0.031 \\ 0.479 & 0.203 & 0.075 & 0.158 & 0.055 & 0.030 \\ 0.533 & 0.222 & 0.082 & 0.065 & 0.065 & 0.030 \\ 0.528 & 0.234 & 0.074 & 0.083 & 0.069 & 0.012 \end{bmatrix}$$

$$\pi(4) = [0.5031764 \quad 0.2303152 \quad 0.0793133 \quad 0.1035161 \quad 0.0536373 \quad 0.0300417]$$

$$\pi(4) \times 100\% = [50.31764\% \quad 23.03152\% \quad 7.93133\% \quad 10.35161\% \quad 5.36373\% \quad 3.00417\%]$$

Berbeda dengan prediksi tahun-tahun 2022,2023, dan 2024, pada prediksi 2025, hanya prediksi pengeluaran barang tahan lama yang mengalami kenaikan sebesar 0.00594%. Dengan prediksi pengeluaran barang tahan lama pada tahun 2025 sebesar 10.35161%. Sedangkan prediksi pengeluaran untuk keperluan perumahan dan fasilitas rumah tangga, pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, pengeluaran pajak, pungutan, dan asuransi, serta pengeluaran untuk keperluan pesta sebesar 50.31764%, 23.03152%, 7.93133%, 5.36373%, dan 3.00417. Dengan penurunan untuk pengeluaran untuk keperluan perumahan dan fasilitas rumah tangga, pengeluaran barang dan jasa, pengeluaran pakaian, alas kaki, dan tutup kepala, pengeluaran pajak, pungutan, dan asuransi, serta pengeluaran untuk keperluan pesta masing-masing sebesar 0.00118%, 0.00367%, 0.00057%, 0.00047%, dan 0.00005%.

Dari prediksi yang telah dilakukan, maka diperoleh

Tabel 2. Prediksi pengeluaran menurut kelompok non makanan

Prediksi Tahun	Prediksi pengeluaran perbulan menurut kelompok non makanan					
	Perumahan dan Fasilitas Rumah tangga	Aneka Barang dan Jasa	Pakaian, Alas Kaki dan Tutup Kepala	Barang Tahan Lama	Pajak, Pungutan dan Asuransi	Keperluan Pesta
2022	52.5%	24.2%	8.2%	7.4%	4.6%	3.1%
2023	50.3705%	23.131%	7.9442%	10.1995%	5.3499%	3.0049%
2024	50.31882%	23.03519%	7.9319%	10.34567%	5.3642%	3.00422%
2025	50.31764%	23.03152%	7.9913%	10.35161%	5.36373%	3.00417%

Prediksi pengeluaran perumahan dan fasilitas rumah tangga selalu mengalami penurunan setiap tahunnya, yaitu pada tahun 2022-2025 berturut-turut mengalami penurunan menjadi Rp. 205.847, Rp. 201.463, Rp. 201.359 dan Rp. 201.357. Penurunan pengeluaran perumahan dan fasilitas rumah tangga setiap tahunnya dapat diakibatkan oleh barang yang ada dianggap masih layak atau keperluan tersebut baru saja dipenuhi beberapa tahun belakangan. Sedangkan prediksi pengeluaran untuk aneka barang dan jasa pada prediksi tahun 2022 mengalami kenaikan dibandingkan tahun sebelumnya sebesar 0.8% yaitu menjadi Rp. 92.187, namun untuk tahun 2023-2025 mengalami penurunan secara berturut-turut menjadi Rp. 91.202, Rp. 91.115 dan Rp. 91.112.

Prediksi pengeluaran untuk pakaian, alas kaki dan tutup kepala juga mengalami kenaikan pada tahun 2022 sebesar 0.8% sehingga menjadi Rp. 29.012 dan prediksinya mengalami penurunan pada tahun 2023-2025 secara berturut-turut menjadi Rp. 28.938, Rp. 28.902 dan Rp. 28.901. Untuk prediksi pengeluaran pakaian, alas kaki dan tutup kepala pada tahun 2022 mengalami kenaikan salah satu faktornya dikarenakan pada tahun ini perkembangan dari aplikasi belanja *online* sedang mengalami peningkatan. Pada tahun 2022 prediksi pengeluaran untuk keperluan barang tahan lama mengalami penurunan sebesar 0.9% yaitu menjadi Rp. 32.329, namun mengalami peningkatan terus menerus pada tahun 2023-2025 berturut-turut menjadi Rp. 33.234, Rp. 33.283, dan Rp. 33.285. Faktor keperluan jangka panjang dapat menjadi alasan masyarakat mengeluarkan pengeluaran untuk barang tahan lama menjadi keperluan yang penting dan mengalami peningkatan setiap tahunnya. Untuk pengeluaran keperluan pajak, pungutan dan asuransi mengalami penurunan pada tahun 2022 sebesar 2.3%, sehingga menjadi Rp. 26.456. Sedangkan mengalami kenaikan pada prediksi 2023 dan 2024 menjadi Rp. 26.654 dan Rp. 26.658. Pada tahun 2025,

pengeluaran keperluan pajak, pungutan dan asuransi mengalami penurunan menjadi Rp. 26.657.

Sedangkan prediksi pengeluaran untuk keperluan pesta pada prediksi 2022 mengalami kenaikan sebesar 1.9% sehingga menjadi Rp. 4.854, dan pada 2023-2025 selalu mengalami penurunan menjadi Rp. 4.853, Rp. 4.852, dan Rp. 4.851. Keperluan pesta sendiri mengalami penurunan yang dapat diakibatkan beberapa tahun terakhir ini diberlakukan pengurangan mobalitas masyarakat oleh pemerintah, sehingga untuk acara yang mengumpulkan banyak orang seperti pesta mengalami penurunan yang berdampak pada pengeluaran untuk kebutuhan pesta ikut mengalami penurunan

4. KESIMPULAN DAN SARAN

Pengeluaran non makanan masyarakat dapat menjadi salah satu indikator untuk melihat kesejahteraan rakyat di suatu daerah. Hal ini dikarenakan semakin tinggi pengeluaran untuk keperluan non makanan, maka dapat dikatakan bahwa pengeluaran untuk makanan sudah tercukupi, yang dapat diakibatkan oleh meningkatnya pendapatan. Dengan menggunakan metode *Markov Chains* untuk memprediksi pengeluaran barang tahan lama setiap tahunnya selalu mengalami peningkatan. Faktor keperluan jangka panjang dapat menjadi alasan masyarakat mengeluarkan pengeluaran untuk barang tahan lama menjadi keperluan yang penting dan mengalami peningkatan setiap tahunnya. Selain itu, pengeluaran pajak, pungutan dan asuransi ikut meningkat. Sedangkan pengeluaran pakaian, alas kaki dan tutup kepala akan diprediksi meningkat di tahun 2025. Untuk keperluan Perumahan, dan fasilitas rumah tangga, pengeluaran aneka barang dan jasa dan keperluan pesta mengalami penurunan tiap tahunnya.

Oleh karena itu, pemerintah dapat meningkatkan lagi upaya dalam meningkatkan kesejahteraan rakyat pada saat pandemi Covid-19. Dikarenakan hasil dari prediksi pengeluaran masyarakat di Kabupaten Tulang Bawang pada tahun 2022-2025 belum terjadinya kekonsistenan peningkatan untuk pengeluaran non makanan. hal tersebut dapat diakibatkan pendapatan dari masyarakat belum stabil.

DAFTAR PUSTAKA

Badan Pusat Statistik. 2012. SIRUSA Pengeluaran per Kapita. <https://sirusa.bps.go.id/sirusa/index.php/indikator/197> . Diakses pada 3 Februari 2022.

Badan Pusat Statistik Kabupaten Tulang Bawang. 2020. Indikator Kesejahteraan Rakyat

Kabupaten Tulang Bawang 2020. Badan Pusat Statistik Kabupaten Tulang Bawang.

Noeryanti, N., Setyawan, Y., dan Hadinagara, H. 2019. *Memprediksi Kemiskinan Di Provinsi Daerah Istimewa Yogyakarta Menggunakan Metode Analisis Rantai Markov*. Jurnal Teknologi. 12, 45-53.

Nugroho, R. B. 2020. *Dampak Covid-19 Bagi Kegiatan Ekonomi Masyarakat*. Orphanet Journal of Rare Diseases. 21, 1-9.

Nurman, T. A., dan Syata, I. 2021. *Prediksi Hasil Panen Kopi di Sulawesi Menggunakan Analisis Rantai Markov*. Jurnal MSA (Matematika dan Statistika serta Aplikasinya). 9, 120-127.

Putri, N. N., dan Muliawati, T. 2021. *Analisis Rantai Markov dalam Memprediksi Status Pasien COVID-19 di Indonesia*. Indonesian Journal of Applied Mathematics. 1, 44-50.

Pramuditya, S. A., Marwati, R., dan Puspita, E. 2014. *Peramalan Pangsa Pasar Kartu GSM dengan Pendekatan Rantai Markov*. Euclid. 1, 116-124.

Praza, R., dan Shamadiyah, N. 2020. *Analisis Hubungan Pengeluaran dengan Ketahanan Pangan Rumah Tangga Petani di Kabupaten Aceh Utara*. Agrifo: Jurnal Agribisnis Universitas Malikussaleh. 5, 23-34.

Suryono, A. 2014. *Kebijakan Publik Untuk Kesejahteraan Rakyat. Transparansi: Jurnal Ilmiah Ilmu Administrasi*. 6, 98-102.

Susanto, E. R., Puspaningrum, A. S., dan Neneng, N. 2021. *Model Rekomendasi Penerima Bantuan Sosial Berdasarkan Data Kesejahteraan Rakyat*. Jurnal Tekno Kompak. 15, 1-12.

Wahyudi, S. R., Sari, R. F., dan Widyasari, R. 2021. *Penentuan Pola Penyebaran Curah Hujan Harian Kabupaten Karo Dengan Menggunakan Rantai Markov Orde-N*. Jurnal Riset dan Aplikasi Matematika (JRAM). 5, 144-157.

JStatistika : Jurnal Ilmiah Teori Dan Aplikasi Statistika ber ISSN online 2654-7511 dan ISSN cetak 2089-0028,

JStatistika merupakan media penerbitan karya ilmiah berupa literasi hasil penelitian ataupun kajian tentang teori, metodologi serta aplikasi dalam bidang ilmu statistika dan data sains. **JStatistika** diterbitkan oleh Program Studi Statistika Fakultas Sains dan Teknologi Universitas PGRI Adi Buana Surabaya dan terbit dua kali dalam setahun setiap bulan Juli dan Desember.

PUBLISHED BY :

**PRODI STATISTIKA
FAKULTAS SAINS DAN TEKNOLOGI**

