

## Teknik Oversampling Pada Regresi Logistik Ordinal Dalam Menduga Faktor Yang Memengaruhi Risiko Penyebaran Zona Covid-19 di Kabupaten Garut

Ghina Fauziah<sup>(1)</sup>, Indahwati<sup>(2)</sup>, Erfiani<sup>(3)</sup>, Anwar Fitrianto<sup>(4)</sup>, Reni Amelia<sup>(5)</sup>

IPB University

Kampus IPB Baranangsiang, Jl. Raya Pajajaran, Bogor, Jawa Barat-Indonesia

e-mail: [ghinafauziah57@gmail.com](mailto:ghinafauziah57@gmail.com), [indah.stk@gmail.com](mailto:indah.stk@gmail.com), [erfiani@apps.ipb.ac.id](mailto:erfiani@apps.ipb.ac.id),  
[anwarstat@gmail.com](mailto:anwarstat@gmail.com), dan [reniamelia@apps.ipb.ac.id](mailto:reniamelia@apps.ipb.ac.id)

### ABSTRAK

Covid-19 merupakan penyakit yang disebabkan oleh infeksi virus SARS-CoV-2, pertama kali masuk ke Indonesia pada awal tahun 2020. Adanya wabah covid-19 menyebabkan setiap daerah yang ada di Indonesia khususnya kabupaten Garut harus terbagi ke dalam beberapa risiko zona covid-19 sesuai dengan kondisi dari suatu daerah tersebut. Beberapa faktor yang dapat memengaruhi suatu daerah masuk pada risiko zona tertentu dapat ditentukan berdasarkan jumlah kasus positif covid-19, kasus suspek, kasus kontak erat, jumlah desa, dan kepadatan penduduk daerah tersebut. Regresi logistik ordinal merupakan salah satu analisis regresi yang digunakan untuk menganalisa hubungan antara variabel prediktor dan variabel respon, dimana variabel respon tersebut bersifat kategorik dengan skala ordinal. Oleh karena itu, maka digunakan regresi logistik ordinal untuk mengetahui faktor apa saja yang memberikan pengaruh terhadap pembagian risiko zona covid-19 di kabupaten Garut pada bulan Juli tahun 2021. Sebelum melakukan pemodelan regresi logistik ordinal dilakukan terlebih dahulu proses teknik *resampling* dengan metode *oversampling* untuk menangani data yang tidak seimbang pada peubah respon. Berdasarkan pemodelan hasil dari pemodelan serta pengujian secara parsial, didapatkan bahwa peubah bebas yang memiliki pengaruh terhadap risiko zona covid-19 di kabupaten Garut yaitu jumlah desa, kepadatan penduduk, kasus suspek, dan kasus konfirmasi positif dengan nilai akurasi sebesar 85.71%.

**Kata kunci : Risiko Zona Covid-19, Oversampling, Regresi Logistik Ordinal.**

### ABSTRACT

Covid-19 is a disease caused by infection with the SARS-CoV-2 virus, which first entered Indonesia in early 2020. The Covid-19 outbreak has caused every region in Indonesia, especially Garut district, to be divided into several risk zones for COVID-19 according to the conditions of the area. Several factors that can affect an area entering a certain risk zone can be determined based on the number of positive cases of COVID-19, suspected cases, close contact cases, the number of villages, and the population density of the area. Ordinal logistic regression is one of the regression analyze the relationship between the predictor variable and the response variable, where the response variable is categorical with an ordinal scale. Therefore, so ordinal logistic regression was used to find out what factors had an influence on the distribution of the risk of the covid-19 zone in Garut district in July 2021. Before doing ordinal logistic regression modeling, the resampling technique process was carried out with the oversampling method to handle the data. unbalanced on the response variabel. Based on the modeling results from modeling and partial testing, it was found that the independent variabels that have an influence on the risk of the covid-19 zone in Garut district are the number of villages, population density, suspected cases, and positive confirmation cases with an accuracy value of 85.71%.

**Keywords : Covid-19 Zone Risk, Oversampling, Ordinal Logistiks Regression.**

### 1. PENDAHULUAN

Virus corona 2019 atau covid-19 merupakan penyakit menular yang disebabkan oleh virus SARS-CoV-2. Penyakit tersebut telah menyerang

sebagian besar manusia hampir diseluruh negara. Pandemi covid-19 telah dimulai pada tahun 2019 yang terjadi di negara China, diawali dengan menyebarnya virus yang ditemui pada manusia dan

hewan yang menyerang alat vital paru-paru dan memiliki risiko kematian tanpa memandang usia, sehingga semua manusia dapat terkena virus covid-19 tersebut tanpa memandang umur. Awal tahun 2020, *World Health Organization* (WHO) menetapkan bahwa wabah covid-19 telah ditetapkan menjadi masapandemi global dengan penyebaran virus yang sangat cepat dan berpotensi terserang infeksi penyakit covid-19 diseluruh dunia, termasuk Indonesia.

Pada awal tahun 2020 merupakan awal baru bagi seluruh manusia di dunia, tidak terkecuali Indonesia. Pandemi wabah covid-19 telah menyerang hampir ke seluruh wilayah di Indonesia, salah satunya Jawa Barat. Berdasarkan data dari dinas Kesehatan Jawa Barat bahwa dari awal tahun 2020 data konfirmasi positif covid-19 semakin hari semakin meningkat, terutama pada tahun 2021 dimana terjadi lonjakan yang sangat tinggi terhadap konfirmasi positif kasus covid-19. Menurut Pusat Informasi dan Koordinasi Jawa Barat, kabupaten Garut merupakan salah satu dari 15 kabupaten/kota di Jawa Barat dengan kasus konfirmasi positif paling banyak dengan jumlah kasus total di Jawa Barat hingga 28 Mei 2022 yaitu 1.106.499 orang terkonfirmasi positif covid-19. Bulan Juli merupakan salah satu bulan dengan kasus konfirmasi positif covid-19 paling tinggi di kabupaten Garut dibandingkan bulan-bulan pada tahun sebelumnya maupun setelahnya yaitu pada tahun 2021 terjadi sebanyak 4.141 konfirmasi positif covid-19, kemudian pada tahun 2021 terjadi sebanyak 20.633 kasus konfirmasi positif covid-19, dan 2022 hingga 27 Mei sebanyak 6.134 kasus konfirmasi positif covid-19.

Dengan adanya wabah pandemi covid-19, maka dari itu seluruh daerah di Indonesia khususnya kecamatan-kecamatan yang ada di kabupaten Garut menjadi terbagi ke dalam beberapa risiko zona yang salah satunya ditentukan berdasarkan kenaikan atau penurunan jumlah kasus covid-19 di kecamatan tersebut. Tujuan adanya pembagian risiko zona covid-19 untuk mengendalikan kasusnya yang berbasis pada pendataan wilayah masing-masing yang diawasi secara ketat serta membedakan antar wilayah yang memiliki risiko tinggi, risiko sedang, risiko rendah, dan wilayah yang tidak berisiko. Risiko zona ditentukan berdasarkan hasil skoring dan pembobotan dari berbagai indikator kesehatan masyarakat. Maka dari itu, penulis bermaksud melakukan penelitian mengenai pengaruh kasus harian covid-19 meliputi jumlah desa, kepadatan penduduk, suspek, konfirmasi positif, dan kontak erat terhadap risiko zona covid-19 di kabupaten Garut periode bulan Juli tahun 2021.

Dalam menganalisis kasus risiko sebaran zona covid-19 di Kabupaten Garut, digunakan model regresi logistik ordinal. Regresi logistik ordinal digunakan untuk memodelkan hubungan antara peubah respon yang terdiri lebih dari dua kategori dengan satu atau lebih peubah penjelas (Hosmer dan Lemeshow 2000). Adapun alasan mengapa metode yang digunakan adalah regresi logistik ordinal adalah karena kasus penyebaran covid-19 merupakan variabel kategorik (lebih dari 2 kategori) dengan skala ordinal sehingga tepat jika menganalisis sebaran risiko zona covid di kabupaten Garut menggunakan regresi logistik ordinal, berbeda dengan regresi logistik lainnya seperti regresi logistik biner dan regresi logistik multinomial. Selain itu, regresi logistik ordinal dapat menjawab variabel apa saja yang dapat memengaruhi sebaran zona covid-19 di kabupaten Garut. Sebelum melakukan proses pemodelan, dilakukan terlebih dahulu pre-processing data dengan menggunakan teknik resampling dengan tujuan untuk menyeimbangkan distribusi data sehingga mengurangi efek distribusi kelas atau kategori yang tidak seimbang (Jian et al. 2016), adapun teknik resampling yang digunakan adalah random oversampling. Random oversampling digunakan untuk menyeimbangkan data dengan cara menambahkan data secara acak pada kelas minoritas atau paling sedikit (risiko zona covid-19) yaitu pada kelas kategori zona risiko tinggi dan zona risiko rendah.

Berdasarkan penelitian yang telah dilakukan oleh Wara Pramesti dkk mengenai risiko penyebaran covid-19 di Jawa Timur menggunakan regresi logistik ordinal dengan metode *backward elimination*, dari hasil penelitian tersebut dapat disimpulkan bahwa yang memengaruhi risiko penyebaran covid-19 di Jawa Timur adalah jumlah kasus covid-19 yang meninggal dan rata-rata anggota rumah tangga. Dari hasil penelitian tersebut, Pemerintah Jawa Timur dapat membuat kebijakan dengan cara memprioritaskan hal-hal yang dapat menyebabkan penyebaran covid-19 semakin parah.

Tujuan dari penelitian ini : (1) mengetahui gambaran sebaran risiko zona covid-19 di kabupaten Garut; (2) mengaplikasikan metode regresi logistik ordinal untuk mengidentifikasi faktor apa saja yang mempengaruhi sebaran risiko zona covid-19 di kabupaten Garut.

**2. METODE PENELITIAN**

**2.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder mengenai kasus risiko sebaran zona covid-19 kabupaten Garut bulan Juli tahun 2021 yang diperoleh dari Dinas Kesehatan kabupaten Garut.

**2.2 Variabel dan Definisi Operasional Variabel**

Variabel yang digunakan dalam penelitian ini yaitu risiko penyebaran zona covid-19 sebagai variabel respon berbentuk data ordinal (rendah, sedang, tinggi) yang dinotasikan dengan Y. Variabel bebas terdiri dari jumlah desa ( $X_1$ ), kepadatan penduduk ( $X_2$ ), kasus suspek ( $X_3$ ), kasus konfirmasi positif ( $X_4$ ), dan kasus kontak erat ( $X_5$ ).

**2.3 Teknik Analisis**

Tahapan analisis yang dilakukan dengan rincian sebagai berikut:

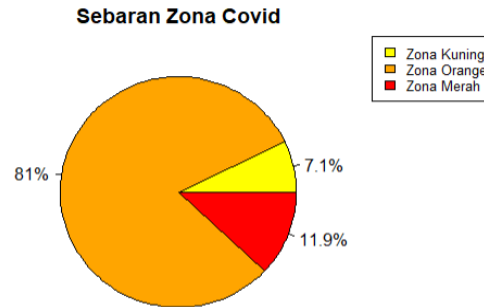
1. Melakukan eksplorasi terhadap data awal
2. Melakukan pengujian korelasi antar variabel bebas.
3. Membagi data menjadi 80% data latih dan 20% data uji.
4. Melakukan proses oversampling.
5. Menyusun model regresi logistik ordinal menggunakan teknik oversampling sebagai berikut:
  - a. Pemodelan regresi logistik ordinal
  - b. Pengujian parameter secara simultan
  - c. Pengujian parameter secara parsial
6. Menyusun model regresi logistik ordinal menggunakan teknik oversampling dengan metode stepwise sebagai berikut:
  - a. Pemodelan regresi logistik ordinal
  - b. Pengujian parameter secara simultan
  - c. Pengujian parameter secara parsial
7. Menentukan akurasi model
8. Menghitung oddsratio

**3. HASIL DAN PEMBAHASAN**

**3.1 Hasil Eksplorasi Terhadap Data Awal**

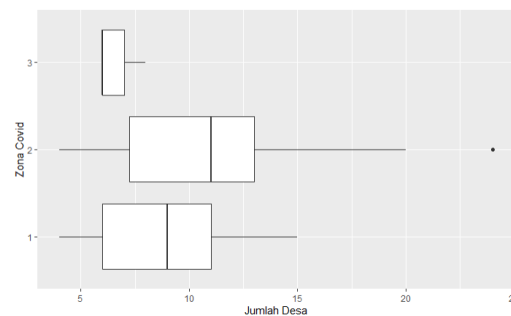
Dari data sampel 42 kecamatan di kabupaten Garut terdapat 7.1% atau 3 kecamatan termasuk pada risiko zona kuning, 81% atau 34 kecamatan termasuk pada risiko zona oranye, dan sisanya sebesar 11.9% atau 5 kecamatan termasuk pada risiko zona merah (Gambar 1). Median jumlah desa yang termasuk pada risiko zona oranye lebih besar dibandingkan median risiko zona lainnya (Gambar 2). Median persentase kepadatan penduduk yang termasuk pada risiko zona oranye lebih besar dibandingkan median risiko zona lainnya (Gambar

3). Median kasus suspek yang termasuk pada risiko zona oranye lebih besar dibandingkan median risiko zona lainnya (Gambar 4). Median kasus konfirmasi positif yang termasuk pada risiko zona oranye lebih besar dibandingkan median risiko zona lainnya (Gambar 5). Median kasus kontak erat yang termasuk pada risiko zona rendah lebih besar dibandingkan median risiko zona lainnya (Gambar 6).



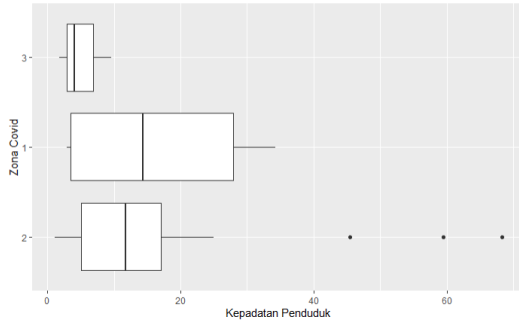
**Gambar 1.** Risiko sebaran zona covid-19

Pada gambar 1 diatas dapat dilihat bahwa dari 42 kecamatan yang ada di kabupaten Garut terdapat 34 kecamatan (81%) masuk pada zona oranye (tingkat keparahan sedang) terdiri dari kecamatan Wanaraja, Tarogong Kidul, Tarogong Kaler, Talegong, Sukawning, Sukaresmi, Sucinaraja, Singajaya, Selaawi, Samarang, PAsirwangi, Pangatikan, Pamulihan, Pakenjeng, Mekarmukti, Malangbong, Limbangan, Leuwigoong, Leles, Karangpawitan, Kadungora, Garut Kota, Cisurupan, Cilawu, Cikelet, Cikajang, Cihurip, Cigedug, Cibiuk, Cibatu, Cibalong, Bungbulang, Bayongbong, dan Banjarwangi. Sedangkan terdapat 3 kecamatan (7.1%) masuk pada zona kuning (zona keparahan rendah) terdiri dari kecamatan Peundeuy, Pameungpeuk, dan Caringin. Kemudian sisanya sebanyak 5 kecamatan (11.9%) masuk pada zona merah (tingkat keparahan tinggi) terdiri dari kecamatan Kersamanah, Karangtengah, Cisompet, Cisewu, dan Banyuresmi.



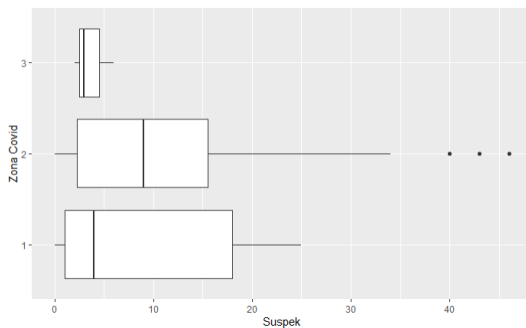
**Gambar 2.** Jumlah desa menurut risiko sebaran zona covid-19

Berdasarkan gambar 2 diatas, dapat disimpulkan bahwa nilai tengah jumlah desa dari zona kuning, zona oranye, dan zona merah berbeda. Daerah dengan zona oranye memiliki nilai tengah jumlah desa yang paling besar dibandingkan dengan zona lainnya.



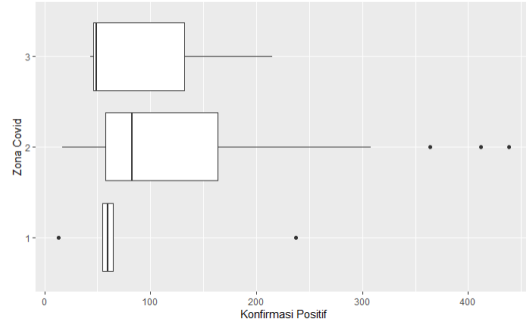
**Gambar 3.** Persentase laju pertumbuhan penduduk menurut risiko sebaran zona covid-19

Berdasarkan gambar 3 diatas, dapat disimpulkan bahwa nilai tengah kepadatan penduduk dari zona kuning, zona oranye, dan zona merah berbeda. Daerah dengan zona oranye memiliki nilai tengah kepadatan penduduk yang paling besar dibandingkan dengan zona lainnya.



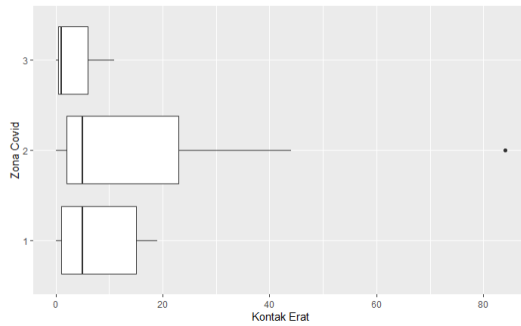
**Gambar 4.** Kasus suspek menurut risiko sebaran zona covid-19

Berdasarkan gambar 4 diatas, dapat disimpulkan bahwa nilai tengah kasus suspek dari zona kuning, zona oranye, dan zona merah berbeda. Daerah dengan zona oranye memiliki nilai tengah kasus suspek yang paling besar dibandingkan dengan zona lainnya.



**Gambar 5.** Kasus konfirmasi positif menurut risiko sebaran zona covid-19

Berdasarkan gambar 5 diatas, dapat disimpulkan bahwa nilai tengah kasus konfirmasi positif dari zona kuning, zona oranye, dan zona merah berbeda. Daerah dengan zona oranye memiliki nilai tengah kasus konfirmasi positif yang paling besar dibandingkan dengan zona lainnya.

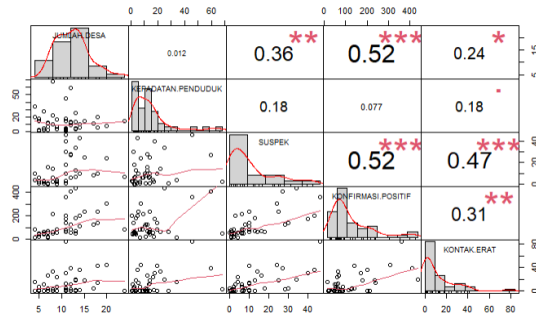


**Gambar 6.** Kasus kontak erat menurut risiko sebaran zona covid-19

Berdasarkan gambar 6 diatas, dapat disimpulkan bahwa nilai tengah kasus kontak erat dari zona kuning, zona oranye, dan zona merah berbeda. Daerah yang memiliki zona oranye dan zona merah memiliki nilai tengah kasus kontak erat yang paling besar dibandingkan dengan zona lainnya.

### 3.2 Korelasi Antar Variabel Bebas

Berikut adalah gambar dari korelasi antar variabel bebas.



Gambar 7. Korelasi antar variabel bebas

Berdasarkan gambar 7 diatas ditunjukkan bahwa hasil korelasi antar variabel bebas tidak ada yang lebih besar dari 0.8, maka dapat disimpulkan bahwa tidak terdapat korelasi antar variabel bebas.

**3.3 Membagi Data Menjadi 80% Data Latih dan 20% Data Uji**

Terdapat 33 kecamatan yang menjadi data latih dan 9 kecamatan menjadi data uji. Dari 33 kecamatan yang menjadi data latih, 3 kecamatan dikategorikan sebagai kecamatan dengan risiko zona tinggi (merah), 28 kecamatan dikategorikan sebagai kecamatan dengan risiko zona sedang (oranye), dan 2 kecamatan dikategorikan sebagai kecamatan dengan risiko zona rendah (kuning).

**3.4 Proses Oversampling**

Jumlah data pada masing-masing kategori rendah, sedang, dan tinggi tidak sama banyak atau tidak seimbang, sehingga dilakukan resampling menggunakan metode oversampling dengan cara menambah data secara acak pada kelas kategori yang memiliki data paling sedikit atau minoritas, yaitu menambah data secara acak pada risiko zona merah sebanyak 25 data dan pada risiko zona kuning sebanyak 26 data, sehingga setelah dilakukan oversampling maka jumlah data yang dimiliki yaitu 84 data.

**3.5 Regresi Logistik Ordinal Menggunakan Teknik Oversampling**

**1) Pemodelan Regresi Logistik Ordinal**

Tabel 1 menyajikan hasil pemodelan regresi logistik ordinal dengan teknik oversampling.

Tabel 1. Model Regresi Logistik Ordinal

Variabel	Koefisien regresi	Standard error
Tinggi   Sedang	-4.5116	0.9825
Sedang   Rendah	-1.7076	0.7400
Jumlah Desa	-0.2571	0.0892
Kepadatan Penduduk	-0.0878	0.0308

Suspek	-0.1880	0.0455
Konfirmasi Positif	0.0205	0.0043
Kontak Erat	0.0223	0.032

Pada tabel 1 diatas diperoleh dua persamaan regresi logistik ordinal, karena terdapat tiga kategori pada variabel respon sehingga didapatkan model persamaan sebanyak dua model sebagai berikut:

$$\log \text{it}[P(Y \leq 1|X)] = -4.5116 - 0.2571 * \text{JumlahDesa} - 0.0878 * \text{KepadatanPenduduk} - 0.1880 * \text{Suspek} + 0.0205 * \text{KonfirmasiPositif} + 0.0223 * \text{KontakErat} \quad (1)$$

$$\log \text{it}[P(Y \leq 3|X)] = -1.7076 - 0.2571 * \text{JumlahDesa} - 0.0878 * \text{KepadatanPenduduk} - 0.1880 * \text{Suspek} + 0.0205 * \text{KonfirmasiPositif} + 0.0223 * \text{KontakErat} \quad (2)$$

**2) Uji Parameter Secara Simultan**

Pengujian parameter secara serentak atau simultan digunakan untuk mengetahui apakah variabel bebas berpengaruh signifikan terhadap variabel respon. Hasil uji menggunakan Likelihood Ratio Test dan didapatkan  $X^2 = 59.823$  dengan nilai p-value =  $1.322e-11$ . Karena nilai p-value < 0.05 sehingga dapat disimpulkan bahwa  $H_0$  ditolak, artinya minimal ada satu variabel bebas yang memiliki pengaruh terhadap variabel respon.

**3) Uji Parameter Secara Parsial**

Tabel 2 menyajikan hasil pengujian parameter secara parsial menggunakan uji Wald.

Tabel 2. Hasil uji parameter secara parsial

Variabel	Nilai wald	p-value
Tinggi   Sedang	-4.5116	4.39e-06
Sedang   Rendah	-1.7076	2.10e-02
Jumlah Desa	-0.2571	3.97e-03
Kepadatan Penduduk	-0.0878	4.42e-03
Suspek	-0.1880	3.67e-05
Konfirmasi Positif	0.0205	2.27e-06
Kontak Erat	0.0223	4.89e-01

Berdasarkan tabel 2 diatas menunjukkan bahwa variabel jumlah desa, kepadatan penduduk, suspek, dan konfirmasi positif berpengaruh terhadap risiko sebaran zona covid di kabupaten Garut Juli tahun 2021. Hal tersebut dapat dilihat pada nilai p-value yang kurang atau lebih kecil dari  $\alpha = 0.05$ . Pada variabel kontak erat tidak berpengaruh signifikan terhadap variabel risiko sebaran zona covid-19, hal ini ditunjukkan karena nilai p-value lebih besar dari  $\alpha = 0.05$ .



**3.6 Regresi Logistik Ordinal Menggunakan Teknik Oversampling Disertai Stepwise**

**1) Pemodelan Regresi Logistik Ordinal**

Tabel 3 menyajikan hasil pemodelan regresi logistik ordinal dengan teknik oversampling yang disertai metode stepwise.

**Tabel 3.** Model Regresi Logistik Ordinal

Variabel	Koefisien regresi	Standard error
Tinggi   Sedang	-4.3489	0.9275
Sedang   Rendah	-1.5719	0.6975
Jumlah Desa	-0.2472	0.0866
Kepadatan Penduduk	-0.0755	0.0232
Suspek	-0.1759	0.0419
Konfirmasi Positif	0.0203	0.0043

Pada tabel 3 diatas diperoleh dua persamaan regresi logistik ordinal menggunakan teknik oversampling yang disertai dengan metode stepwise, karena terdapat tiga kategori pada variabel respon sehingga didapatkan model persamaan sebanyak dua model sebagai berikut:

$$\log \text{it}[P(Y \leq 1|X)] = -4.3489 - 0.2472 * \text{JumlahDesa} - 0.0755 * \text{KepadatanPenduduk} - 0.1759 * \text{Suspek} + 0.0235 * \text{KonfirmasiPositif} \quad (3)$$

$$\log \text{it}[P(Y \leq 3|X)] = -1.5719 - 0.2472 * \text{JumlahDesa} - 0.0755 * \text{KepadatanPenduduk} - 0.1759 * \text{Suspek} + 0.0235 * \text{KonfirmasiPositif} \quad (4)$$

Hasil estimasi  $\alpha_1 = -4.3489$  lebih kecil dari  $\alpha_3 = -1.5719$  berarti model ini adalah model kumulatif dengan kemiringan yang sama yaitu model garis regresi yang berdasarkan pada peluang kumulatif kategori respon (Albana 2013). Hal tersebut menunjukkan bahwa peluang kumulatif risiko tinggi lebih kecil daripada peluang kumulatif risiko rendah, berarti setiap kenaikan risiko Covid-19 akan menurunkan peluang logit kumulatif.

**2) Uji Parameter Secara Simultan**

Uji parameter secara simultan digunakan untuk mengetahui apakah variabel bebas berpengaruh terhadap variabel respon. Hasil uji dengan menggunakan Likelihood Ratio Test dan didapatkan  $X^2 = 59.356$  dengan nilai p-value = 3.96e-12. Karena nilai p-value < 0.05 maka dapat disimpulkan bahwa  $H_0$  ditolak, artinya minimal ada satu variabel bebas yang memiliki pengaruh terhadap variabel respon.

**3) Uji Parameter Secara Parsial**

Tabel 4 menyajikan hasil pengujian parameter secara parsial menggunakan uji Wald.

**Tabel 4.** Hasil uji parameter secara parsial

Variabel	Nilai wald	p-value
Tinggi   Sedang	-4.3489	2.75e-06
Sedang   Rendah	-1.5719	2.42e-02
Jumlah Desa	-0.2472	4.31e-03
Kepadatan Penduduk	-0.0755	1.15e-03
Suspek	-0.1759	2.77e-05
Konfirmasi Positif	0.0203	2.41e-06

Berdasarkan tabel 4 diatas menunjukkan bahwa variabel jumlah desa, kepadatan penduduk, suspek, dan konfirmasi positif berpengaruh terhadap risiko sebaran zona covid di kabupaten Garut Juli tahun 2021. Hal ini dapat dilihat pada nilai p-value yang kurang atau lebih kecil dari  $\alpha = 0.05$ .

**3.7 Akurasi Model**

Model akurasi digunakan untuk menilai seberapa besar keakurasian model yang dihasilkan. Nilai akurasi didapatkan dari perhitungan *confussion matrix* antara data hasil prediksi dengan data yang sebenarnya pada data uji, kemudian dilakukan pengururan ketepatan antara data prediksi dan data sebenarnya yang dapat menghasilkan tingkat ketepatan dari *confussion matrix* tersebut. Baik model dengan menggunakan teknik oversampling yang disertai stepwise maupun tidak memiliki nilai akurasi yang sama besar yaitu 85,71%. Artinya model regresi logistik ordinal yang didapatkan baik menggunakan metode stepwise ataupun tidak itu baik karena memiliki nilai akurasi yang tinggi.

**3.8 Interpretasi Odds Ratio**

Interpretasi terhadap model dapat dilihat berdasarkan nilai odds ratio. Nilai odds ratio didapatkan dari nilai koefisien variabel bebas yang di eksponensialkan. Adapun hasil dari nilai odds ratio sebagai berikut:

**Tabel 5.** Hasil odds ratio

Variabel	Koefisien regresi	Odds ratio
Tinggi   Sedang	-4.3489	0.0129
Sedang   Rendah	-1.5719	0.2076
Jumlah Desa	-0.2472	0.7809
Kepadatan Penduduk	-0.0755	0.9271
Suspek	-0.1759	0.8386
Konfirmasi Positif	0.0203	1.0205

Interpretasi odds ratio dari output diatas adalah sebagai berikut:

1) Interpretasi Model Risiko Zona Tinggi

- $\text{Exp}(\text{intercept risiko zona tinggi}) = 0.01292 \rightarrow$  tanpa memperhatikan peubah bebasnya, peluang sebaran covid zona tinggi (merah) adalah 0.01292 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya.

- $\text{Exp}(\text{jumlah desa}) = 0.780917 \rightarrow$  Odds kecamatan dengan kategori zona tinggi (merah) akan meningkat sebesar 0.780917 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika jumlah desa naik sebesar satu desa dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{kepadatan penduduk}) = 0.92719 \rightarrow$  Odds kecamatan dengan kategori zona tinggi (merah) akan meningkat sebesar 0.92719 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika kepadatan penduduk naik sebesar satu desa dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{suspek}) = 0.83868 \rightarrow$  Odds kecamatan dengan kategori zona tinggi (merah) akan meningkat sebesar 0.83868 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika jumlah kasus suspek naik sebesar satu kasus dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{konfirmasi positif}) = 1.02058 \rightarrow$  Odds kecamatan dengan kategori zona tinggi (merah) akan meningkat sebesar 1.02058 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika jumlah konfirmasi positif naik sebesar satu kasus dengan peubah bebas lainnya konstan.

2) Interpretasi Model Risiko Zona Rendah

- $\text{Exp}(\text{intercept risiko zona rendah}) = 0.20764 \rightarrow$  tanpa memperhatikan peubah bebasnya, peluang sebaran covid zona rendah (kuning) adalah 0.20764 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya.

- $\text{Exp}(\text{jumlah desa}) = 0.780917 \rightarrow$  Odds kecamatan dengan kategori zona rendah (kuning) akan meningkat sebesar 0.780917 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika jumlah desa naik sebesar satu desa dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{kepadatan penduduk}) = 0.92719 \rightarrow$  Odds kecamatan dengan kategori zona rendah (kuning) akan meningkat sebesar 0.92719 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika kepadatan penduduk naik sebesar satu desa dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{suspek}) = 0.83868 \rightarrow$  Odds kecamatan dengan kategori zona rendah (kuning) akan meningkat sebesar 0.83868 kali dibandingkan

peluang untuk masuk pada kategori zona covid lainnya jika jumlah kasus suspek naik sebesar satu kasus dengan peubah bebas lainnya konstan.

- $\text{Exp}(\text{konfirmasi positif}) = 1.02058 \rightarrow$  Odds kecamatan dengan kategori zona rendah (kuning) akan meningkat sebesar 1.02058 kali dibandingkan peluang untuk masuk pada kategori zona covid lainnya jika jumlah konfirmasi positif naik sebesar satu kasus dengan peubah bebas lainnya konstan.

#### 4. KESIMPULAN DAN SARAN

Hasil penelitian dalam mengidentifikasi faktor yang memengaruhi risiko zona covid-19 menggunakan regresi logistik ordinal dengan teknik oversampling diperoleh kesimpulan yaitu:

1. Kecamatan dengan risiko zona sedang memiliki proporsi yang lebih tinggi dibandingkan dengan risiko zona tinggi dan rendah yaitu sebanyak 33 kecamatan termasuk pada risiko zona sedang, 5 kecamatan termasuk pada daerah dengan risiko zona tinggi, dan 3 kecamatan termasuk pada daerah dengan risiko zona rendah.

2. Terdapat dua model regresi logistik ordinal menggunakan teknik oversampling, yaitu model untuk kecamatan dengan risiko zona tinggi dan kecamatan dengan risiko zona rendah. Bentuk umum model dapat dituliskan:

$$\log \text{it}[P(Y \leq 1|X)] = -4.3489 - 0.2472 * \text{JumlahDesa} - 0.0755 * \text{KepadatanPenduduk} - 0.1759 * \text{Suspek} + 0.0235 * \text{KonfirmasiPositif}$$

$$\log \text{it}[P(Y \leq 3|X)] = -1.5719 - 0.2472 * \text{JumlahDesa} - 0.0755 * \text{KepadatanPenduduk} - 0.1759 * \text{Suspek} + 0.0235 * \text{KonfirmasiPositif}$$

Hasil estimasi  $\alpha_1 = -4.3489$  lebih kecil dari  $\alpha_3 = -1.5719$  berarti model ini adalah model kumulatif dengan kemiringan yang sama yaitu model garis regresi yang berdasarkan pada peluang kumulatif kategori respon (Albana 2013). Hal tersebut menunjukkan bahwa peluang kumulatif risiko tinggi lebih kecil daripada peluang kumulatif risiko rendah, berarti setiap kenaikan risiko Covid-19 akan menurunkan peluang logit kumulatif.

3. Variabel prediktor yang memengaruhi terhadap sebaran zona covid-19 di Kabupaten Garut periode Juli 2021 adalah jumlah desa, kepadatan penduduk, kasus suspek, dan konfirmasi positif dengan nilai akurasi sebesar 85.71%.

**DAFTAR PUSTAKA**

- Agresti, A. (2002). *Categorical Data Analysis*. Inc. New York.
- Ananth, Cande V. (1997). *Regression Models for Ordinal Responses: A Review of Methods and Applications*. *International Journal of Epidemiology*, 2-6.
- Arofah, Irvana. (2018). Analisis Persepsi Biaya Kuliah Perguruan Tinggi dengan Menggunakan Metode Regresi Logistik Ordinal. *Jurnal Statistika dan Aplikasinya*, 1-6.
- B. R. A. Febrilia, S. Rahayu, and B. D. Korida. (2019). "Ordinal Logistic Regression Analysis of Factors Affecting the Length of Student Study". *J. Mat. "MANTIK"*, 1-7.
- Choi, Seung W. (2011). *lordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations*. *J Stat Softw*, 2-3.
- Hosmer, D.W. and Lemeshow, S. (2000) *Applied logistik regression*. 2nd Edition. A John Wiley and Sons, Inc. New York.
- Liang, Jiaqi. (2020). *Multinomial and ordinal Logistic regression analyses with multi-categorical variables using*. *Journal of Annals Of Translational Medicine*, 1-7.
- Liu, Xing. (2012). *Ordinal Regression Analysis: Using Generalized Ordinal Logistic Regression Models to Estimate Educational Data*. *Journal Of Modern Applied Statistical Methods*. 2-6.
- Mugirahayu, A. S. (2021). *Penentuan Status Kewaspadaan COVID-19 Pada Suatu Wilayah Menggunakan Metode Fuzzy Inference System (FIS) Mamdani*. *Jurnal Sains dan Edukasi Sains*, 1-2.
- Nugraha, Jaka. (2017). *Pemodelan Data Ordinal, Nominal dan Cacah*. Universitas Islam Indonesia.
- P. Thomas, N.A. Salmon, W. Riana. (2016). *Analisis Regresi Logistik Ordinal*. *Barekeng Jurnal Matematika dan Ilmu Terapan*, 1-6.
- Pramesti, W. (2021). *Regresi Logistik Backward Elimination pada Risiko Penyebaran Covid-19 di Jawa Timur*. *Journal of Statistics and Its Applivation on Teaching and Research*, 1-8.
- S. Imaslihkah, M. Ratna, and V. Ratnasari. (2013). "Analisis Regresi Logistik Ordinal terhadap Faktor-faktor yang mempengaruhi Predikat Kelulusan Mahasiswa S1 di ITS Surabaya". *J. Sains dan Seni Pomits*, 1-6.
- Shen, C., Bar-yam, Y., & New England Complex Systems Institute. (2020). *Color Zone Pandemic Response Version 2*. 2020. Retrieved from <https://necsi.edu/color-zone-pandemic-response-version-2>, diakses tanggal 4 April 2022.