J STATISTIKA

Program Studi Statistika

EISSN : 2654-7511
PISSN : 2089-0028

J STATISTIKA

JURNAL ILMIAH TEORI DAN APLIKASI STATISTIKA

Volume 17 | Nomor 2 | 2024

EDITORIAL TEAM

INTRODUCTION

We are delighted to announce the current publication of Volume 18, Number 1 of JStatistika, affiliated with the Statistics Department at PGRI Adi Buana University Surabaya, has been released in July 2025. This particular issue of the JStatistika Scientific Journal features a diverse array of articles addressing a wide spectrum of topics. One of the highlighted articles delves into "Optimization of Raw Material Supply Scheduling to Maximize Storage Utilization at Company X; Optimization of Spare Part Inventory for Mask Production Machines at PT. XYZ to Enhance Productivity; Potential Productivity of Quail Farming in Ngaliyan, Pulutan, Wonosari, Gunungkidul; Analysis of Public Perception of Dynastic Politics in the 2024 Presidential and Vice Presidential Elections in Indonesia with a Chi-Square Approach; Three–Phase Traffic Light Petri Net Model Using The Modified Norwegian System; Weibull Regression Survival Analysis on the Rate of Recovery of Thyphoid Fever Patients: Case Study of RSUD Haji Makassar; Analysis of Environmental and Productivity Factors with the Number of Dengue Hemorrhagic Fever and Obesity Cases in Indonesia; Generalized Poisson Regression Modeling on the Number of Infant Deaths in East Nusa Tenggara Province in 2022"

The JStatistika Scientific Journal enthusiastically welcomes and invites contributions in a diverse range of formats, including but not limited to scholarly scientific articles that encompass various facets of statistical science. We eagerly seek research findings, comprehensive reports, insightful case studies, thorough literature reviews, and updates that pertain to the dynamic landscape of statistical science. Our overarching objective is to cultivate a repository of knowledge that is not only current but also invaluable in tackling the ever-evolving and intricate challenges confronting our field. We actively encourage authors to submit their work if it resonates with the most recent advancements and frontiers in statistical science. Our aspiration is to foster an environment where these contributions can flourish, ultimately serving as a wellspring of cutting-edge insights and understanding. We believe that these insights are instrumental in addressing the multifaceted issues that confront us in today's complex world.

Our editorial team extends a warm and inclusive invitation to scientists and scholars from diverse backgrounds and affiliations, including institutions of higher learning and esteemed research organizations. We seek your valuable contributions, whether they be grounded in empirical research results or rooted in rigorous scholarly studies within the expansive domain of statistics and its myriad practical applications. We hold a deep appreciation for the feedback and perspectives of our esteemed readership. Your input not only enriches the discourse but also plays a pivotal role in our continuous efforts to elevate the quality and relevance of the journal. We earnestly value your insights and ideas, recognizing that they are integral to our ongoing pursuit of excellence. Our ultimate vision is for the articles featured in the JStatistika Scientific Journal to transcend the confines of academia and serve as a wellspring of knowledge that benefits not only

scholars and researchers but also professionals actively engaged in the diverse realms of statistical science and its multifaceted real-world applications. Through collaborative efforts and a shared commitment to advancing our understanding of statistics, we aim to make a meaningful impact in the broader scientific community and beyond.

Jstatistika has been indexed by Sinta 4 Kemendikbud, Garuda, Google Scholar, Crossref, Worldcat, Scilit, ROAD, Onesearch, Journal Stories, Dimensions, Base, Open Alex, Wikidata, Internet Archieve, Root Indexing, Core, Harvard Library, Universiteit Leiden Library, Semantic Scholar, Open Air Explore, ASCI, Cite Factor, University of Saskatchewan Library, The University of Queensland Library, George University Library and Boston University Library.

Surabaya, December 2024

Editor in Chief

LIST OF CONTENTS

Optimization of Raw Material Supply Scheduling to Maximize Storage Utilization at Company X

Karina Rahmawati [(1)], Hery Murnawan [(2)]

[1,2] Universitas 17 Agustus 1945 Surabaya

Jalan Semolowaru 45, Surabaya 60118, East Java, Indonesia
e-mail: karinarhmwti@gmail.com[(1)], herymurnawan@untag-sby.ac.id[(2)]

ABSTRAK

Studi ini berfokus pada pengoptimalan penjadwalan pasokan bahan baku di Perusahaan X, sebuah bisnis di industri makanan dan minuman, untuk memaksimalkan pemanfaatan penyimpanan. Perusahaan X menghadapi tantangan dalam menyeimbangkan permintaan bahan baku dengan kapasitas penyimpanan yang terbatas, yang menyebabkan masalah seperti kekurangan stok atau kelebihan stok. Praktik manajemen inventaris saat ini terbukti tidak efektif dalam memenuhi permintaan konsumen secara efisien. Penelitian ini menggunakan metodologi kuantitatif, dengan fokus pada analisis data numerik untuk mengatasi tujuan penelitian. Melalui pendekatan ini, data yang dapat diukur dikumpulkan dan diperiksa untuk memberikan wawasan yang jelas tentang pola, hubungan, dan tren. Metode kuantitatif memastikan objektivitas dan ketepatan, menjadikannya ideal untuk mengevaluasi faktor-faktor seperti jadwal pasokan bahan baku, kapasitas penyimpanan, dan fluktuasi permintaan, sehingga mendukung pengambilan keputusan berdasarkan data. Melalui analisis data penjualan dan kebutuhan bahan baku, penelitian ini bertujuan untuk mengembangkan strategi penjadwalan yang lebih efektif yang menyelaraskan pasokan dengan permintaan sambil mengoptimalkan ruang penyimpanan. Temuan studi ini menawarkan solusi potensial untuk meningkatkan efisiensi operasional, mengurangi pemborosan, dan meningkatkan kinerja rantai pasokan secara keseluruhan. Dengan menerapkan penjadwalan pasokan yang dioptimalkan, Perusahaan X dapat memastikan operasi yang lebih lancar dan memenuhi kebutuhan konsumen dengan lebih baik, sehingga memaksimalkan nilai di seluruh rantai pasokan.

Kata kunci: Penjadwalan Pasokan; Bahan Baku; Optimasi; Pemanfaatan Penyimpanan; Manajemen Inventaris.

ABSTRACT

This study focuses on optimizing the raw material supply scheduling at Company X, a business in the food and beverage industry, to maximize storage utilization. Company X faces challenges in balancing raw material demand with limited storage capacity, leading to issues such as stock shortages or overstocking. The current inventory management practices have proven ineffective in meeting consumer demand efficiently. This research utilizes a quantitative methodology, focusing on numerical data analysis to address the research objectives. Through this approach, measurable data are collected and examined to provide clear insights into patterns, relationships, and trends. The quantitative method ensures objectivity and precision, making it ideal for evaluating factors such as raw material supply schedules, storage capacity, and demand fluctuations, thereby supporting data driven decision-makingThrough the analysis of sales data and raw material requirements, this research aims to develop a more effective scheduling strategy that aligns supply with demand while optimizing storage space. The findings of this study offer potential solutions for improving operational efficiency, reducing waste, and enhancing the overall supply chain performance. By implementing optimized supply scheduling, Company X can ensure smoother operations and better meet consumer needs, thus maximizing value across the entire supply chain.

Keywords: *Supply Scheduling; Raw Materials; Optimization; Storage Utilization; Inventory Management.*

INTRODUCTION

Company X operates in the food and beverage industry and is headquartered in Surabaya. Established in 2017, the company has gained significant recognition, as evidenced by its numerous branches widely distributed across Surabaya. The company specializes in producing healthy food products, primarily vegetable salads and cold-pressed juices. The raw materials used consist of fresh vegetables and fruits. Monitoring the inventory of raw materials is crucial to ensuring the smooth operation of the company. Effective monitoring guarantees sufficient stock availability, optimizes supply efficiency, ensures timely deliveries, and maintains the quality of raw materials, all of which significantly impact the production and distribution processes [1]. This, in turn, enables the company to maximize value creation across the entire supply chain, from suppliers and the company itself to the end consumers [2].

For Company X, inventory monitoring is conducted based on demand, which also dictates the production process. Higher demand results in increased raw material usage for production, and vice versa. This demand is often assessed based on the sales activity and foot traffic at the sales stands. The company faces storage limitations for both unprocessed and processed raw materials. The following outlines the chiller capacity and its utilization for storage purposes.

Based on the observations, several key improvements are required to enhance the efficiency of raw material management. First, evaluating raw material supply needs is essential to ensure that procurement aligns with production requirements and consumer demand [3]. By analyzing demand patterns and sales data, the company can estimate the necessary quantities of raw materials to avoid shortages or disruptions in production [4]. Additionally, scheduling procurement periods with suppliers is critical to maintaining a steady supply flow while minimizing storage constraints. Properly timed deliveries will help the company reduce the risk of overstocking and maintain the freshness of raw materials, ensuring high-quality products for consumers.

Another important aspect is optimizing the limited storage capacity of the company's chiller. By assessing the shelf life of raw materials, the company can implement a more efficient storage strategy, prioritizing perishable items and ensuring timely utilization [5]. This approach reduces waste due to spoilage and maximizes the effective use of the available space. Through this research, the company seeks to create a balance between supply chain efficiency and storage limitations, allowing it to meet consumer demand consistently while minimizing operational costs. These improvements will ultimately contribute to the company's ability to sustain growth and maintain its reputation for delivering fresh and healthy food products [6].

The primary objective of this research is to establish an optimal scheduling system for raw material procurement from suppliers. This scheduling aims to align with the company's production needs while addressing storage constraints [7]. By analyzing demand patterns and the available chiller capacity, the study seeks to determine the ideal frequency and volume of raw material deliveries to prevent stock shortages or overstock situations. Efficient scheduling is essential to maintain the freshness and quality of raw materials, which are critical for the company's products vegetable salads and cold-pressed juices [8]. Additionally, an accurate supply plan will enable the company to adapt quickly to fluctuating consumer demand, ensuring seamless production and sales operations [9,10].

Moreover, the research emphasizes maximizing the utilization of the company's limited chiller capacity. By integrating data on the shelf life of raw materials and their usage rates, the study

will propose strategies for efficient storage management. This includes prioritizing fast-moving or perishable items and minimizing waste caused by spoilage [11,12]. Ultimately, the research aims to enhance operational efficiency, reduce unnecessary costs, and support the company's commitment to delivering high-quality, healthy products to consumers. Through this approach, the company can achieve a balance between supply chain efficiency and its storage limitations, thereby sustaining its competitive advantage in the market.

METHOD

This research adopts a quantitative methodology as its core approach, emphasizing numerical data analysis to address the research objectives systematically and effectively. This method is particularly suited for studies that require precision and objectivity in evaluating measurable variables. By focusing on numerical data, this methodology provides the foundation for identifying clear insights into the patterns, relationships, and trends that influence the research problem. In this study, the quantitative approach is applied to assess critical factors such as raw material supply schedules, storage capacity, and fluctuations in consumer demand, all of which are integral to optimizing supply chain performance. The strength of a quantitative approach lies in its ability to support data-driven decision-making, enabling the development of solutions that are based on empirical evidence rather than assumptions. By collecting and analyzing measurable data, the research ensures objectivity, reliability, and replicability, making it particularly effective for evaluating operational and logistical challenges faced by businesses such as Company X.

To achieve the research objectives, the study follows a structured series of steps:

1. Identification Phase: This initial stage involves formulating the research problem based on observed challenges. In the context of Company X, the primary issues identified include inefficiencies in raw material supply scheduling, limited storage capacity, and misalignment between supply and demand.
2. Data Collection Phase: During this phase, relevant data are gathered to provide a comprehensive understanding of the operational context. Key data collected include:
3. Raw Material Supply Data: Information on the quantity and frequency of raw material deliveries from suppliers.
4. Bill of Materials (BOM) for Each Menu: Detailed breakdown of the raw materials required for producing specific menu items, ensuring accurate demand planning.
5. Chiller Capacity Data: Data on available storage space for both raw and processed materials, crucial for optimizing inventory management.
6. Data Processing Phase: This stage involves the application of advanced forecasting and error measurement techniques to analyze the collected data. The methods used include:
7. Moving Average: A technique to smooth data fluctuations and identify trends over time.
8. Exponential Smoothing: A forecasting method that assigns exponentially decreasing weights to past observations for predicting future demand.
9. Error Metrics: Methods such as Mean Absolute Deviation (MAD), Mean Square Error (MSE), and Mean Absolute Percentage Error (MAPE) are employed to evaluate the accuracy of the forecasting models and determine the best-fit parameters [13].
10. Conclusion Phase: The final step synthesizes the findings from the data analysis to develop actionable insights and recommendations. These conclusions are designed to address the

identified problems, offering practical solutions to optimize raw material supply scheduling, improve storage utilization, and enhance overall operational efficiency.

RESULT AND DISCUSSION

Raw Material Supply

The recapitulation of raw material supplies distributed to the Gwalk stand from each supplier over 14 days in July 2024 provides valuable insights into the contribution of each supplier to the availability of raw materials required by Company X. The raw material supplies provided by the suppliers are recorded as gross weight, which indicates that a portion of these materials is still in their raw state and must undergo kitchen processes, including waste removal and preparation, before being transformed into finished goods ready for sale.

Waste refers to the portion of raw materials that cannot be processed or must be discarded, such as inedible parts of vegetables. Additionally, waste arises from the reduction in material weight during processing. The table highlights that when raw materials enter the production process, all materials undergo kitchen preparation, resulting in a decrease in weight due to shrinkage or the removal of non-usable parts, such as vegetable cores. This emphasizes the importance of considering waste and processing loss in managing raw material supply effectively.

Bill of Materials for Each Menu

The bill of materials (BOM) is utilized to identify the specific raw materials and their required quantities for each menu item. This data provides detailed information on the components necessary for preparing each dish, ensuring accurate planning and efficient use of resources. Below is the data outlining the raw materials required for each menu item.

Table 1. Bill of Materials for Each Menu

Caesar Teaser (Rp. 56.000,-)		**Game of Truffle (Rp. 65.000,-)**		**The Hulk (Rp. 60.000,-)**	
BAHAN BAKU	QTY	BAHAN BAKU	QTY	BAHAN BAKU	QTY
Sayur	100 gram	Sayur	100 gram	Sayur	90 gram
Egg	0,5 butir	Aragula	5 gram	Kale	5 gram
Chicken Plain	40 gram (1 pac)	Egg	0,5 butir	Onion	8 gram
Smoked Beef	20 gram (1 pac)	Smoked Beef	20 gram (1 pac)	Kubis	15 gram
Parmesan	2,5 gram	Jamur	20 gram	Chicken Plain	40 gram (1 pac)
Cracker	15 gram	Mozarella	25 gram	Edamame	25 gram
		Crouton	15 gram	Wijen	2,5 gram
Waldorf Chicken (Rp. 60.000,-)		**Happy Garden (Rp. 60.000,-)**		**Konichiwa Salad (Rp. 60.000,-)**	
BAHAN BAKU	QTY	BAHAN BAKU	QTY	BAHAN BAKU	QTY
Sayur	90 gram	Sayur	90 gram	Sayur	90 gram
Aragula	5 gram	Aragula	5 gram	Onion	8 gram
Onion	8 gram	Onion	8 gram	Egg	0,5 butir
Celery	8 gram	Carrot	15 gram	Crabstick	30 gram (1 pac)
Chicken Plain	40 gram (1 pac)	Egg	0,5 butir	Cherry Tomato	10 gram (4-5 slice)
Cherry Tomato	10 gram (4-5 slice)	Chicken Plain	40 gram (1 pac)	Corn	30 gram
Grape	35 gram	Smoked Beef	20 gram (1 pac)	Edamame	25 gram
Feta	21 gram	Cherry Tomato	10 gram (4-5 slice)	Bonito	4 gram
Cracker	15 gram	Parmesan	2,5 gram		
		Fried Garlic	2,5 gram		
		Cracker	15 gram		
Wow Wow West (Rp. 60.000,-)		**Cajun Chicken (Rp. 60.000,-)**		**Tasty Thai (Rp. 60.000,-)**	
BAHAN BAKU	QTY	BAHAN BAKU	QTY	BAHAN BAKU	QTY
Sayur	90 gram	Sayur	90 gram	Sayur	90 gram
Kale	10 gram	Kale	5 gram	Carrot	15 gram
Carrot	15 gram	Onion	20 gram	Chicken Plain	40 gram (1 pac)
Onion	20 gram	Egg	0,5 butir	Grape	35 - 40 gram
Egg	0,5 butir	Chicken Plain	40 gram (1 pac)	Jicama	15 gram
Chicken Plain	40 gram (1 pac)	Smoked Beef	20 gram (1 pac)	Fried Shallot	15 gram
Cherry Tomato	20 gram (8-10 slice)	Cherry Tomato	20 gram (8-10 slice)		
Mozarella	25 gram	Corn	30 gram		
		Crouton	15 gram		

Mix Mix (Rp. 56.000,-)		The Yogi (Rp. 60.000,-)		Chop Suey (Rp. 60.000,-)	
BAHAN BAKU	QTY	BAHAN BAKU	QTY	BAHAN BAKU	QTY
Sayur	90 gram	Sayur	90 gram	Sayur	100 gram
Carrot	15 gram	Kale	5 gram	Carrot	15 gram
Cucumber	15 gram	Carrot	15 gram	Kubis	15 gram
Kubis	15 gram	Onion	8 gram	Chicken Plain	40 gram (1 pac)
Egg	0,5 butir	Cucumber	15 gram	Orange	35 gram (7 slice)
Tempe	2 pcs	Chicken Plain	40 gram (1 pac)	Wijen	2,5 gram
Potato	25 gram (4-5 slice)	Cherry Tomato	10 gram (4-5 slice)	Cracker	15 gram
Cherry Tomato	10 gram (4-5 slice)	Feta	21 gram		
Fried Shallot	15 gram	Corn	30 gram		
Ranch Chicken (Rp. 60.000,-)		**Autumn Crunch (Rp. 60.000,-)**		**Seoul Yummy (Rp. 60.000,-)**	
BAHAN BAKU	QTY	BAHAN BAKU	QTY	BAHAN BAKU	QTY
Sayur	90 gram	Sayur	90 gram	Sayur	90 gram
Celery	8 gram	Aragula	5 gram	Onion	8 gram
Chicken Plain	40 gram (1 pac)	Carrot	15 gram	Carrot	15 gram
Cherry Tomato	25 gram (10-12 slice)	Tofu	2 pcs	Egg	0,5 butir
Beetroot	15 gram	Orange	35 gram (7 slice)	Crabstick	30 gram (1 pac)
Potato	35 gram (8 slice)	Feta	21 gram	Mushroom	20 gram
Mozarella	25 gram			Jicama	15 gram
Crouton	15 gram				

Chiller Capacity Data

Company X has a total of 15 chillers. However, only a few shelves in each chiller are available for storing both raw and processed raw materials. This limitation is due to space constraints, which result in some chillers being used to store a mix of different types of materials. Below is the data detailing the available chiller capacity for storage purposes:

Table 2. Chiller Capacity Data

Letak	Nama	Kapasitas	Terpakai	Kegunaan
Ruang Mg	Chiller 1	568 kg	568 kg	Menyimpan sayur 80 kg, Kale
	Chiller 2	568 kg	568 kg	Menyimpan sayur 80 kg, Aragula
	Chiller 3	510 kg	510 kg	Menyimpan sayur 40 kg, Kubis, Celey
	Chiller 4	568 kg	568 kg	Menyimpan 50 box mg
	Chiller 5	180 kg	180 kg	Menyimpan 24 box mg
Lorong Mg	Chiller 6	180 kg	25 kg	Menyimpan Kale 10 pack
				Menyimpan Celery 15 pack
Kitchen	Chiller 7	180 kg	108 kg	Menyimpan Chicken Plain 60 pack
				Menyimpan Smoked Beef 100 pack
				Menyimpan Crabstick 20 pack
Condiment	Chiller 8	1081 kg	540,5 kg	Menyimpan Anggur 60 pack
				Menyimpan Onion 25 pack
				Menyimpan Edamame 30 pack
				Menyimpan Kubis 15 pack
				Menyimpan Jamur 20 pack
				Menyimpan Bengkuang 30 pack
				Menyimpan Wortel 30 pack
				Menyimpan Mozzarella 70 pack
				Menyimpan Tofu 10 pack
	Chiller 9	228 kg	228 kg	Menyimpan Ayam Mentah 80 kg
Gudang	Chiller 10	568 kg	189 kg	Menyimpan Parmesan 14 pack
	Chiller 11		284 kg	Menyimpan Feta 21 pail
	Chiller 12	370 kg	370 kg	Menyimpan Tortilla 244 pack
	Chiller 13	371 kg	371 kg	
	Chiller 14	372 kg	372 kg	
	Chiller 15	373 kg	373 kg	Menyimpan Beef Mentah 200 pack

After performing the forecasting calculations using the Exponential Smoothing method with alpha values ranging from 0.1 to 1, and calculating the error rate in the forecast, the next step is to select the optimal alpha (α) based on the error rate results. The purpose is to determine which alpha value provides the lowest error, as a smaller error indicates a more accurate forecast. Below is the recap of the error rate calculations from the forecasting trials using alpha (α) values ranging from 0.1 to 1.

Table 3. Exponential Smoothing

Menu	FE	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
Caesar Teaser	MAD	275,75	285,67	294,98	294,98	343,13	365,85	387,83	408,88	428,61	446,40
	MSE	120440,91	121800,06	128157,21	128157,21	150221,82	365,85	182206,72	201834,58	223789,16	247752,40
	MAPE	0,018	0,018	0,019	0,019	0,022	0,024	0,025	0,026	0,028	0,029
Waldorf Chicken	MAD	56,57	52,20	48,52	48,52	48,70	48,35	50,21	53,88	57,31	60,40
	MSE	4278,92	3778,94	3580,60	3580,60	3577,45	48,35	3822,10	4019,03	4267,36	4566,00
	MAPE	0,010	0,009	0,008	0,008	0,008	0,008	0,009	0,009	0,010	0,010
Happy Garden	MAD	97,09	93,14	90,57	90,57	98,29	104,74	110,85	116,58	121,83	126,40
	MSE	12239,06	12216,89	12676,20	12676,20	14335,39	104,74	16697,52	18149,16	19782,43	21574,80
	MAPE	0,031	0,030	0,029	0,029	0,031	0,033	0,035	0,037	0,038	0,040
Cajun Chicken	MAD	391,30	454,98	505,63	505,63	573,73	594,19	607,70	615,85	635,24	648,00
	MSE	449499,05	477289,12	511456,34	511456,34	589667,34	594,19	683114,29	739517,82	805354,19	883546,00
	MAPE	0,151	0,189	0,218	0,218	0,256	0,267	0,274	0,277	0,287	0,293
Game of Truffle	MAD	88,40	98,21	106,63	106,63	119,58	124,29	128,03	130,98	133,67	138,60
	MSE	15515,44	16937,81	18459,13	18459,13	21759,17	124,29	25653,42	27995,95	30725,42	33962,60
	MAPE	0,060	0,067	0,073	0,073	0,083	0,086	0,088	0,090	0,092	0,095
Konichiwa	MAD	26,34	27,30	28,25	28,25	31,45	33,61	35,90	38,31	40,82	43,40
	MSE	857,66	950,49	1055,44	1055,44	1308,56	33,61	1634,67	1833,36	2063,02	2332,60
	MAPE	0,076	0,071	0,067	0,067	0,065	0,066	0,072	0,077	0,083	0,089
Wow Wow West	MAD	48,72	45,41	43,53	43,53	42,19	43,03	47,10	50,72	54,65	58,60
	MSE	3536,22	3329,27	3321,52	3321,52	3560,73	43,03	3938,17	4153,66	4382,41	4623,00
	MAPE	0,046	0,047	0,049	0,049	0,054	0,058	0,062	0,067	0,071	0,076
The Hulk	MAD	23,99	23,27	22,77	22,77	23,70	25,70	27,60	29,32	30,81	32,00
	MSE	900,58	867,49	875,17	875,17	954,65	25,70	1075,73	1143,97	1214,00	1282,80
	MAPE	0,044	0,042	0,042	0,042	0,044	0,047	0,051	0,054	0,057	0,059

Menu	FE	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
The Yogi	MAD	12,91	10,49	8,61	8,61	7,28	7,58	8,08	8,50	9,09	9,80
	MSE	211,90	171,65	151,48	151,48	141,79	7,58	151,96	161,64	174,14	189,80
	MAPE	0,040	0,032	0,026	0,026	0,022	0,023	0,025	0,026	0,028	0,030
Ranch Yogurt	MAD	16,17	16,36	16,08	16,08	14,54	13,45	12,74	12,07	12,40	12,60
	MSE	288,53	289,43	292,67	292,67	291,00	13,45	279,81	274,07	269,99	268,60
	MAPE	0,036	0,037	0,037	0,037	0,033	0,031	0,029	0,028	0,028	0,029
Seoul Yummy	MAD	26,76	26,43	26,34	26,34	26,65	28,26	29,97	31,78	33,71	35,80
	MSE	858,28	920,48	999,65	999,65	1193,66	28,26	1433,93	1574,38	1731,74	1910,60
	MAPE	0,088	0,087	0,087	0,087	0,088	0,093	0,099	0,104	0,110	0,116
Tasty Thai	MAD	9,55	8,83	8,64	8,64	8,51	8,53	8,57	8,61	8,63	8,60
	MSE	160,70	139,77	128,78	128,78	122,04	8,53	123,57	125,60	128,13	131,00
	MAPE	0,018	0,017	0,016	0,016	0,016	0,016	0,016	0,016	0,016	0,016
Chop Suey	MAD	19,43	21,00	22,62	22,62	25,69	26,98	28,01	28,71	29,08	31,60
	MSE	888,45	943,87	1016,78	1016,78	1196,65	26,98	1421,61	1555,77	1708,53	1884,80
	MAPE	0,042	0,046	0,050	0,050	0,057	0,060	0,063	0,065	0,066	0,072
Autumn Crunch	MAD	26,74	27,08	29,76	29,76	32,71	33,23	33,27	32,94	33,30	34,00
	MSE	1290,41	1346,22	1406,54	1406,54	1500,38	33,23	1564,48	1595,93	1631,88	1674,80
	MAPE	0,059	0,060	0,067	0,067	0,074	0,075	0,075	0,074	0,075	0,077
Mix Mix	MAD	17,02	16,16	15,71	15,71	15,71	16,02	16,45	16,98	17,56	18,20
	MSE	342,27	303,69	286,40	286,40	290,35	16,02	326,30	353,97	388,67	431,80
	MAPE	0,146	0,138	0,133	0,133	0,131	0,133	0,135	0,139	0,142	0,147

Based on the error rate calculations using the Exponential Smoothing method with alpha (α) values ranging from 0.1 to 1, the smallest forecast error was achieved with an alpha value of 0.1. After determining the optimal period for the Moving Average method and the appropriate alpha (α) for the Exponential Smoothing method, the next step is to compare the error rates from both forecasting methods. According to the calculations, the error rate for the Moving Average method was based on a 3-month period, while the Exponential Smoothing method used an alpha (α) of 0.1. The following results show the error rates for both methods:

Table 4. Exponential Smoothing and Moving Average

Menu	Moving Average			Exponential Smoothing		
	MAD	MSE	MAPE	MAD	MSE	MAPE
Caesar Teaser	322,11	115783,52	2,1%	275,75	120440,91	1,8%
Waldorf Chicken	30,00	1332,89	0,5%	56,57	4278,92	1,0%
Happy Garden	91,56	11420,22	2,9%	97,09	12239,06	3,1%
Cajun Chicken	480,56	234916,78	28,6%	391,30	449499,05	15,1%
Game of Truffle	87,44	8323,30	6,6%	88,40	15515,44	6,0%
Konichiwa	22,33	783,74	4,0%	26,34	857,66	4,6%
Wow Wow West	27,22	1513,30	3,8%	48,72	3536,22	7,6%
The Hulk	15,33	455,78	3,0%	23,99	900,58	4,4%
The Yogi	1,89	7,59	0,6%	12,91	211,90	4,0%
Ranch Yogurt	16,78	393,52	3,9%	16,17	288,53	3,6%
Seoul Yummy	32,22	1605,41	11,2%	26,76	858,28	8,8%
Tasty Thai	9,11	98,59	1,7%	9,55	160,70	1,8%
Chop Suey	20,44	526,15	4,9%	19,43	888,45	4,2%
Autumn Crunch	31,78	1236,37	7,8%	26,74	1290,41	5,9%
Mix Mix	10,67	142,44	9,3%	17,02	342,27	14,6%

Based on the comparison of the error rates between the Moving Average method with a 3-month period and the Exponential Smoothing method with an alpha (α) of 0.1, it can be concluded that the Exponential Smoothing method with an alpha (α) of 0.1 yields the smallest error rate. This comparison allows the researcher to more accurately determine future requirements based on the forecasting results. Therefore, the future demand can be forecasted by considering the calculations made using the Exponential Smoothing method with an alpha (α) of 0.1. Below is the data for the future demand forecast:

Table 5. Future needs

Menu	May-25	Jun-25	Jul-25	Aug-25	Sep-25	Oct-25	Total (Porsi)
Caesar Teaser	15827	15827	15781	15772	15774	15717	94698
Waldorf Chicken	5950	5950	5939	5933	5930	5924	35626
Happy Garden	3233	3233	3218	3213	3210	3195	19301
Cajun Chicken	1692	1692	1683	1831	1823	1803	10525
Game of Truffle	1355	1355	1348	1374	1373	1367	8172
Konichiwa	578	578	576	580	576	575	3463
Wow Wow West	720	720	709	703	706	703	4260
The Hulk	515	515	515	521	523	522	3111
The Yogi	306	306	309	309	310	311	1851
Ranch Yogurt	431	431	433	435	437	435	2603
Seoul Yummy	304	304	307	308	310	305	1839
Tasty Thai	552	552	552	550	548	548	3302
Chop Suey	423	423	423	430	428	427	2554
Autumn Crunch	415	415	416	424	425	422	2517
Mix Mix	138	138	135	136	135	133	816

Gross Raw Material Requirements

After performing the forecasting calculations, the results indicate the raw material requirements for the future. Raw material needs are divided into two types: gross requirements and net requirements. The gross raw material requirement refers to the future demand or the product demand for each period, without considering the initial inventory. The gross requirement is calculated as follows:

Gross Requirement = Sales Portion x Raw Material Quantity per Serving

For example, to calculate the gross requirement for each menu item, the calculation is as follows:

Table 6. Caesar teaser

Menu	May-24	Jun-24	Jul-24	Aug-24	Sep-24	Oct-24	Total (Porsi)	Rata-Rata/hari (ALL STAND)	Rata-rata Kebutuhan/stand
Caesar Teaser	15827	15827	15781	15772	15774	15717	94698	515	47

The forecast calculation for the Caesar Teaser menu using the Exponential Smoothing method with an alpha (α) of 0.1, spanning from May 2025 to October 2025, indicates a total of 94,698 servings, with sales occurring across 11 stands. The total forecast for each month is as follows:

1. May 2025 = 15,827 servings
2. June 2025 = 15,827 servings
3. July 2025 = 15,781 servings
4. August 2025 = 15,772 servings
5. September 2025 = 15,774 servings

Therefore, the total forecast for the Caesar Teaser menu over 6 months, across 11 stands, is 94,698 servings. To calculate the daily forecast for the Caesar Teaser menu, the following calculation is used:

$$\text{Daily Forecast for All Stands} = \frac{94.698 \text{ servings}}{184 \text{ hari (6 month)}} = 515 \text{ servings per day for all stands.}$$

To determine the daily forecast for each stand, the calculation is as follows:

$$\text{Daily Forecast per Stand} = \frac{515 \text{ servings}}{11 \text{ stand}} = 47 \text{ servings per stand per day.}$$

CONCLUSION

This study successfully developed optimized raw material supply scheduling for Company X to address challenges related to limited storage capacity and fluctuating demand. By applying advanced forecasting methods, such as Exponential Smoothing with an optimal alpha value, the research identified strategies to align raw material supply with production requirements effectively. The findings demonstrated the importance of balancing procurement schedules with storage constraints to minimize waste and ensure the freshness of raw materials. The integration of gross and net requirement calculations enabled precise forecasting of future needs, which significantly enhanced inventory management. Ultimately, the optimized scheduling system proposed in this study offers a robust solution for improving operational efficiency, reducing costs associated with waste and overstocking, and enhancing the overall performance of the supply chain. These strategies not only enable smoother operations but also help Company X sustain its competitive advantage in the food and beverage industry by consistently delivering high-quality, fresh products to consumers.

REFERENCE

[1] Amorim, P., Meyr, H., Almeder, C., & Almada-Lobo, B. (2013). Managing perishability in production-distribution planning: a discussion and review. *Flexible Services and Manufacturing Journal*, *25*, 389-413.

[2] Lee, Y. H., & Kim, S. H. (2002). Production–distribution planning in supply chain considering capacity constraints. *Computers & industrial engineering*, *43*(1-2), 169-190.

[3] Ben-Daya, M., As'Ad, R., & Seliaman, M. (2013). An integrated production inventory model with raw material replenishment considerations in a three layer supply chain. *International Journal of Production Economics*, *143*(1), 53-61.

[4] Wrigley, E. A. (2017). The supply of raw materials in the industrial revolution. In *The causes of the industrial revolution in England* (pp. 97-120). Routledge.

[5] Wei, X., Xu, G., & Kusiak, A. (2014). Modeling and optimization of a chiller plant. *Energy*, *73*, 898-907.

[6] Khan, R. S., Grigor, J., Winger, R., & Win, A. (2013). Functional food product development–Opportunities and challenges for food manufacturers. *Trends in food science & technology*, *30*(1), 27-37.

[7] Spinelli, S., Masi, C., Dinnella, C., Zoboli, G. P., & Monteleone, E. (2014). How does it make you feel? A new approach to measuring emotions in food product experience. *Food Quality and Preference*, *37*, 109-122.

[8] Gao, K., Huang, Y., Sadollah, A., & Wang, L. (2020). A review of energy-efficient scheduling in intelligent production systems. *Complex & Intelligent Systems*, *6*, 237-249.

[9] Du, B., Tan, T., Guo, J., Li, Y., & Guo, S. (2021). Energy-cost-aware resource-constrained project scheduling for complex product system with activity splitting and recombining. *Expert Systems with Applications*, *173*, 114754.

[10] M. Athoillah And M. I. Irawan, "Perancangan Sistem Informasi Mobile Berbasis Android Untuk Kontrol Persediaan Barang Di Gudang," Jurnal Sains Dan Seni Pomits, Vol. 1, No. 1, Pp. 1–6, 2013.

[11] K. Alim, B. Matsaany, and A. Rahmawati, "Diversification of Jakarta Islamic Index (JII) Stock Optimal Portfolio for the Period 2018-2023", J Statistika, vol. 16, no. 2, pp. 585–593, Dec. 2023.

[12] A. Amiseno, "Analisa Perbandingan Statistik Kelemahan SPI dan Biaya Politik sebagai Pemoderasi Kerugian Keuangan Daerah pada Belanja Daerah ", J Statistika, vol. 12, no. 2, pp. 17–22, Dec. 2019.

[13] G. Anuraga, A. Indrasetianingsih, and M. Athoillah, "Pelatihan pengujian hipotesis statistika dasar dengan software r," BUDIMAS: Jurnal Pengabdian Masyarakat, vol. 3, no. 2, pp. 327–334, 2021.

Optimization of Spare Part Inventory for Mask Production Machines at PT. XYZ to Enhance Productivity

Dony Mahardhika [(1)], Hery Murnawan [(2)]

[1,2] Universitas 17 Agustus 1945 Surabaya

Jalan Semolowaru 45, Surabaya 60118, East Java, Indonesia
e-mail: dhic.elc@gmail.com[(1)], herymurnawan@untag-sby.ac.id[(2)],

ABSTRAK

Penelitian ini berfokus pada optimalisasi persediaan suku cadang untuk mesin produksi masker di PT. XYZ dengan tujuan untuk meningkatkan produktivitas dan mengurangi biaya operasional. Penelitian ini mengidentifikasi titik-titik kegagalan utama pada mesin masker, termasuk masalah pada sistem konveyor, tarikan katrol, dan penyegel ultrasonik, yang merupakan kontributor utama terhadap waktu henti produksi. Penelitian ini mengadopsi pendekatan penelitian kuantitatif, dengan fokus pada pengumpulan dan analisis data numerik secara sistematis untuk menarik kesimpulan dan memberikan wawasan objektif terhadap masalah yang sedang dipelajari. Dengan menerapkan konsep manajemen inventaris seperti stok pengaman, tingkat layanan, dan tinjauan berkala, penelitian ini menentukan tingkat inventaris suku cadang optimal yang diperlukan untuk memastikan ketersediaan komponen dan meminimalkan risiko kehabisan stok. Penerapan metode PDCA (Plan-Do-Check-Act) semakin meningkatkan proses perawatan mesin, yang mengarah pada peningkatan tingkat efektivitas mesin. Selain itu, manajemen inventaris yang efektif mengurangi pemborosan karena pembelian berlebih dan kesalahan pengelolaan suku cadang. Temuan tersebut menyoroti bahwa dengan mengoptimalkan inventaris suku cadang dan meningkatkan perawatan mesin, PT. XYZ mampu mengurangi waktu henti, meminimalkan biaya operasional, dan meningkatkan produktivitas secara keseluruhan, yang berkontribusi pada efisiensi dan profitabilitas yang lebih besar.
Kata kunci: Persediaan Suku Cadang; Produksi Masker; Mesin; Produktivitas.

ABSTRACT

This study focuses on the optimization of spare part inventory for mask production machines at PT. XYZ with the aim of enhancing productivity and reducing operational costs. The research identifies key failure points in the mask machines, including issues with the conveyor system, pulley pulls, and ultrasonic sealers, which were the main contributors to production downtime. This study adopts a quantitative research approach, focusing on the systematic collection and analysis of numerical data to draw conclusions and provide objective insights into the issues being studied. By applying inventory management concepts such as safety stock, service levels, and periodic reviews, the study determines the optimal spare part inventory levels required to ensure component availability and minimize the risk of stockouts. The implementation of the PDCA (Plan-Do-Check-Act) method further improves machine maintenance processes, leading to an increase in machine effectiveness levels. Additionally, effective inventory management reduces waste due to over-purchasing and mismanagement of spare parts. The findings highlight that by optimizing spare part inventory and improving machine maintenance, PT. XYZ was able to reduce downtime, minimize operational costs, and enhance overall productivity, contributing to greater efficiency and profitability.
Keywords: *Spare part Inventory; Mask Production; Machines; Productivity.*

INTRODUCTION

The manufacturing industry has experienced rapid growth, with companies competing in both local and global markets to deliver high-quality products. The production of superior products is influenced by various factors, one of which is the smooth operation of production processes that adhere to established standards [1]. Machinery plays a critical role in the production process. A machine is considered efficient if it can achieve production targets without interruptions, operate at the designated speed, and produce high-quality outputs that meet the required standards. Therefore, proper maintenance is essential to ensure machinery remains in optimal condition and ready for use at any time [2]. Maintenance also extends the lifespan of the machines, enabling consistent performance over time [3].

PT. XYZ, an Indonesian pharmaceutical company based in Pandaan, East Java, was established in 2008. The company's primary vision is to foster Indonesia's independence in the healthcare sector, ensuring improved public health through the provision of high-quality, internationally standardized products. PT. XYZ's products cater to both domestic and international markets, including pharmaceuticals, medical devices, and household healthcare supplies. Its pharmaceutical portfolio spans categories such as antidotes and detoxifying agents, antiseptics, cardiovascular and hematopoietic drugs, dermatological treatments, herbal medicines, musculoskeletal remedies, vitamins, and minerals. The medical device range includes disposable devices, defibrillators, alcohol swabs, and ultrasound gel, among others. Additionally, its household healthcare supplies include hand sanitizers, underpads, medical plasters, and masks.

Despite its significant progress, PT. XYZ operates in an integrated and complex environment. Challenges occasionally arise in implementing systems on the ground, leading to discrepancies. For instance, in the production line for earloop masks, the high production costs result in narrow profit margins. Currently, the machines produce an average of 6,878 masks per shift. The total target for a single production batch is 250,000 masks, which requires 32 shifts to complete. With three shifts operating per day, it takes approximately 11 days to complete a single batch. This highlights the need for improvements in operational efficiency to optimize production and reduce costs [4], [5].

The initial stage involves the handover process by raw material warehouse personnel responsible for supplying materials, and production personnel responsible for counting and verifying whether the quantity and quality of raw materials meet the required standards. The second stage is the production process of masks. During the third stage, the packing and labeling processes occur simultaneously with production [6]. The final stage involves the handover of finished goods to the finished goods warehouse by production personnel and warehouse staff [7].

Observations indicate several issues during the mask production process, including production defects, breakdown losses, and setup and adjustment inefficiencies. Breakdown losses are primarily caused by machine failures, setup and adjustment times, and component replacement or repairs. In 2023, machine-related issues occurred an average of 2–3 times per month when production was limited to a single shift per day. However, since February 2024, with production operating on three shifts per day, machine issues have increased to an average of 15 occurrences per month, primarily due to problematic machine components. Data on these component issues will be used to calculate the Mean Time to Failure (MTTF), providing insights into the lifespan of

machine components. Understanding component lifespan is essential for preventing machine malfunctions and reducing breakdown time during production.

One of the key strategies to minimize breakdown time is the implementation of regular maintenance and timely replacement of machine components [8]. In this regard, the availability of machine components or spare parts plays a crucial role. A well-stocked inventory of spare parts ensures readiness when needed, thereby reducing breakdown time. This can be achieved through effective spare parts inventory management [9,10]. Efficient inventory management facilitates operational improvements by ensuring the availability of required spare parts within the designated time frame, while also avoiding wasteful over-purchasing or acquiring components that are unnecessary or unsuitable [11].

The study aims to address the issues described in the background by achieving several key objectives. First, it seeks to measure and evaluate the effectiveness of mask production machines. This involves analyzing their operational performance to determine their efficiency and reliability during the production process [12,13]. Second, the research focuses on identifying the types of failures that occur and uncovering the root causes of machine downtime, which is critical for minimizing disruptions and ensuring smooth operations. Lastly, the study aims to calculate the appropriate inventory levels for spare parts required for mask production machines. By determining the optimal stock levels, the research intends to support efficient inventory management, ensuring the availability of spare parts while avoiding overstocking or shortages.

METHOD

The research method refers to a series of strategies employed by the researcher to collect the necessary data to achieve the research objectives and address the identified problems. This study adopts a quantitative research approach, focusing on the systematic collection and analysis of numerical data to draw conclusions and provide objective insights into the issues being studied. The research steps undertaken are as follows:

1. Preliminary Phase: This phase involves determining the timeline and location of the research.
2. Identification Phase: formulating the problems identified during the research process.
3. Data Collection Phase: Gathering primary data, including production figures, machine downtime, rejection rates, and working hours. Secondary data collection includes an overview of the company and relevant documentation.
4. Data Processing Phase: Identifying machine problems, determining the causes of machine issues, analyzing the root causes of major problems using a fishbone diagram, implementing the PDCA method, and calculating productivity levels.
5. Conclusion Phase: The final stage involves drawing conclusions and providing recommendations based on the data analysis, aimed at addressing the research problems effectively.

RESULT AND DISCUSSION

Time Target

The time target is calculated by subtracting non-running time from the total working hours. This target corresponds to the productive working hours, which amount to 408 minutes per shift. This value serves as the basis for calculating the theoretical target.

Theoretical Target (Design Capacity)

The product target is determined by multiplying the machine speed with the time target. Machine speed refers to the number of products produced by the machine per minute, based on the design capacity of the Expert 9982 machine. The calculation is as follows:

$$\text{Theoretical Target} = \text{Machine Speed} \times \text{Time Target}$$
$$= 21 \text{ ppm} \times 408 \text{ minutes}$$
$$= 8{,}568 \text{ pcs}$$

Defective Product Percentage (%)

The percentage of defective products is calculated as the ratio of defective products to the total output. The calculation is illustrated below:

$$\text{Defective Product Percentage} = \frac{\text{Defective Products}}{\text{Total Products}} \times 100$$

$$\text{Defective Product Percentage} = \frac{393}{6093} \times 100 = 6{,}4\%$$

Average Actual Production Output

The average actual production output is calculated by dividing the total output by the number of shifts. The calculation is as follows:

$$\text{Average Actual Production Output} = \frac{\text{Total Output}}{\text{Total Shifts}}$$

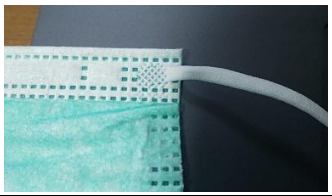
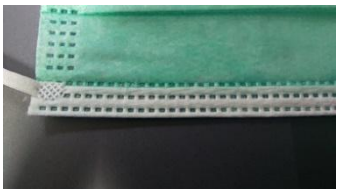
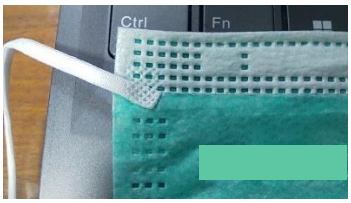
$$\text{Average Actual Production Output} = \frac{216{,}957}{32} = 6{,}779 \text{ pcs}$$

Defective Products

The percentage of defective products per shift remains above 5%, exceeding the company's acceptable threshold of 2%. Several types of defective products do not meet the company's quality standards. These defective products are detailed in the table below.

Table 1. Defective Products

No	Defective Products	No	Defective Products
1	*double welding*	2	*body - folded*
			

3	*body - wrinkle*	4	*body – shape failure*
			
5	*earloop – welding failure*	6	*earloop - not in position*
			

Calculation of Estimated Loss Potential

One critical factor contributing to decreased productivity is machine downtime. Downtime, spanning from the onset of a problem to its resolution, potentially reduces revenue due to suboptimal production capacity. This calculation estimates the loss potential based on the production output that could have been achieved if the machine operated without interruption. Data provided: Total machine downtime (January 2023–March 2024): 17,674 minutes; Machine speed (standard): 21 ppm; Cost of Goods Sold (COGS): IDR 340 per unit.

Estimated Loss = Machine Speed x Downtime x COGS
Estimated Loss = 21 ppm x 17,674 minutes x IDR 340
Estimated Loss = IDR 126,192,360

Analysis of Machine Problem Causes

The following table summarizes the frequency of machine issues and their respective percentages of occurrence based on data from 2023 to March 2024.

Table 2. Analysis of Machine Problem Causes

No	Problem Cause	Frequency (2023)	Frequency (February -March 2024)	Total Incidents	Percentage (%)
1	Roll brake worn	1	4	5	9,6
2	Pneumatic misalignment	8	6	14	26,9
3	Encoder zero-point shift	-	2	2	3,8
4	Guide roll detachment	-	2	2	3,8

5	Sensor roll bolt damage	-	1	1	1,9
6	Nose wire platform misalignment	2	3	5	9,6
7	Mask frequently pinched	-	1	1	1,9
8	Ultrasonic error	3	2	5	9,6
9	Sensor roll misalignment	2	2	4	7,7
10	Motor failure	1	-	1	1,9
11	Mould misalignment	3	-	3	5,8
12	Pneumatic failure	1	-	1	1,9
13	Conveyor detachment	1	-	1	1,9
14	Gripper bolt damage	1	-	1	1,9
15	Bolt housing damage	1	-	1	1,9
16	HMI failure	1	-	1	1,9
17	Pull roll damage	1	-	1	1,9
18	Ultrasonic unit damage	2	-	2	3,8
19	Clamp damage	1	-	1	1,9
Total		**29**	**23**	**52**	**100,00**

The five main causes (below the 80% threshold) of the majority of issues are primarily due to shear pneumatic pull, worn roll brakes, misalignment of the nose wire support, ultrasonic errors, and misaligned roll sensors, while the remaining causes are considered minor. According to the Pareto principle (80/20 rule), the issues contributing to 80% of the problems are the primary causes, while the remaining 20% are minor issues that can either be addressed later or disregarded. In addition to the five machine-related issues that are prioritized for resolution, the availability of spare parts also becomes a key priority to address. Therefore, a summary of the overall problems can be seen in the cause-and-effect diagram presented in Figure 1.

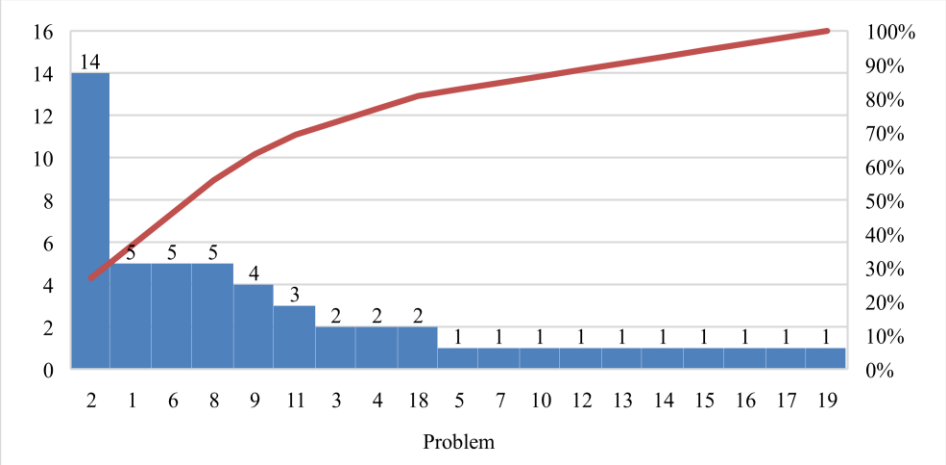


Figure 1. Diagram pareto

Safety Stock Calculation

The purpose of additional inventory is to account for uncertainties in demand and lead time. The calculation is as follows:

σD x safety factor

where:

σD = standard deviation

Standard Deviation Calculation:

The safety factor for a desired service level of 95% is 1.65 (refer to the attached Table of Safety Factors for Normal Distribution). The spare part usage data: 5, 1, 1, 1, 1, 1, 1, 2, 1

$$\text{Average} = \frac{\sum Xi}{n} = \frac{5+1+1+1+1+1+1+2+1}{9} = \frac{14}{9} = 1{,}56 \text{ units}$$

$$\text{Standard Deviation} = \sqrt{\frac{\sum(Xi-\mu)^2}{n}} = \sqrt{\frac{(5-2)^2+(1-2)^2+(1-2)^2}{9}} = 1{,}33 \text{ units}$$

σD x safety factor = 1,56 x 1,33 = 2,07 units = 3 units

Analysis of Reduced Potential Losses

The implementation of inventory management, such as the ordering calculation using the periodic review formula, can assist companies in consistently ensuring the availability of spare parts. More structured inventory management can significantly reduce machine breakdown time, as it enables quicker handling when machinery issues arise, thus lowering the potential for losses. Furthermore, other loss impacts, such as emergency costs to compensate for production shortfalls, can be eliminated or minimized. On the other hand, the application of the periodic review method also reduces the risk of losses due to excess stock. By ordering items based on actual needs and more accurate demand forecasts, the company can avoid waste from stockpile accumulation, whether due to expired items or increased storage costs. Below is a comparison of the production process before and after implementing machine maintenance and inventory management:

Table 3. Reduction of Potential Losses in a Single Production Batch

Periode	Time Target	Time Achieved	*Breakdown Time*	Machine Effectiveness	Estimated Loss
February – Marc 2024	13.056	11.605	1.451	88,89%	Rp 10.360.140
September – October 2024	13.056	12.361	695	94,67%	Rp 4.962.300

Analysis of PDCA Implementation

1. Analysis of the Plan Stage (Planning)
   The main issues have been clearly identified, including machine optimization, breakdown time, and insufficient operator training. The root causes of machine problems have been outlined, such as the lack of regular maintenance, insufficient or out-of-stock spare parts, and operator errors in handling the machines. The proposed solutions regular machine maintenance, spare parts availability, and operator training demonstrate a comprehensive approach to addressing these root causes.
2. Analysis of the Do Stage (Implementation)
   The proposed solutions have been implemented, including performing maintenance during non-operational hours, training operators to enhance their skills, and ensuring adequate spare parts are stocked in the warehouse. The execution of these actions involved the Production and Engineering Departments.
3. Analysis of the Check Stage (Inspection)
   An inspection is conducted by monitoring machine performance during operation and comparing effectiveness before and after maintenance. Data from this monitoring process are used to evaluate the success of the implemented solutions. Continuous monitoring is essential, as it serves as a foundation for further corrective actions.
4. Analysis of the Act Stage (Follow-up)
   The Engineering Department ensures consistency in follow-up actions, focusing on maintaining routine machine maintenance and conducting repairs promptly when machine performance deviates from the standard, without waiting for the next scheduled maintenance. Meanwhile, the Spare Parts Warehouse consistently monitors the availability of machine parts and evaluates the need to adjust the minimum stock levels, thus supporting smooth machine operation.

CONCLUSION

Based on the research conducted on the optimization of spare part inventory for mask machines at PT. XYZ, the following conclusions can be drawn:

1. Identification of Masking Machine Failures: The study successfully identified the types of failures in the mask machines, such as disruptions in the conveyor, pulley pulls, and ultrasonic sealers, which were the primary causes of downtime during the production process.
2. Optimal Inventory Levels: By applying the concepts of safety stock, service level, and periodic review, this research determined the optimal quantity of spare parts required to maintain component availability when needed. This strategy effectively reduces the risk of stockouts, which can lead to production downtimes.
3. Effectiveness of Masking Machine Operations: The implementation of the PDCA (Plan-Do-Check-Act) method in production and inventory management significantly improved the machines' Overall Equipment Effectiveness (OEE), which in turn supported the company’s productivity.
4. Impact of Effective Inventory Management: Effective inventory management not only helped reduce downtime but also minimized cost waste due to over-purchasing or improper

use of spare parts. This holistic approach to managing inventory proved to be a key factor in optimizing operational efficiency and reducing unnecessary expenses.

REFERENCE

[1] Luckyardi, S., Soeryanto Soegoto, E., Supatmi, S., Warlina, L., & Hassan, F. (2022). Marketing strategy for local superior commodities and regional economic contributions of Indonesia. *Journal of Eastern European and Central Asian Research*, *9*(1).

[2] Erhueh, O. V., Nwakile, C., Akano, O. A., Aderamo, A. T., & Hanson, E. (2024). Advanced maintenance strategies for energy infrastructure: Lessons for optimizing rotating machinery. *Global Journal of Research in Science and Technology*, *2*(02), 065-093.

[3] Cao, X. G., Zhang, M. Y., Gong, Y. R., Jia, X. L., & Zhang, R. Y. (2021). Maintenance decision method considering inspection of mining equipment. *International Journal of Metrology and Quality Engineering*, *12*, 21.

[4] Raval, S. J., Kant, R., & Shankar, R. (2020). Analyzing the Lean Six Sigma enabled organizational performance to enhance operational efficiency. *Benchmarking: An International Journal*, *27*(8), 2401-2434.

[5] Ajiga, D., Okeleke, P. A., Folorunsho, S. O., & Ezeigweneme, C. (2024). The role of software automation in improving industrial operations and efficiency. *International Journal of Engineering Research Updates*, *7*(1), 22-35.

[6] Knop, K. (2021). Management of packaging labeling technology in the context of improving the final product quality and work safety. *System Safety: Human-Technical Facility-Environment*, *3*(1), 116-128.

[7] Wruck, S., Vis, I. F., & Boter, J. (2017). Risk control for staff planning in e-commerce warehouses. *International Journal of Production Research*, *55*(21), 6453-6469.

[8] Kolte, T. S., & Dabade, U. (2017). Machine operational availability improvement by implementing effective preventive maintenance strategies-A review and case study. *International Journal of Engineering Research and Technology*, *10*(1), 700-708.

[9] Yazdi, M. (2024). Maintenance strategies and optimization techniques. In *Advances in Computational Mathematics for Industrial System Reliability and Maintainability* (pp. 43-58). Cham: Springer Nature Switzerland.

[10] M. Athoillah and M. I. Irawan, "Perancangan sistem informasi mobile berbasis Android untuk kontrol persediaan barang di gudang," Jurnal Sains dan Seni Pomits, vol. 1, no. 1, pp. 1–6, 2013.

[11] Madu, C. N. (2000). Competing through maintenance strategies. *International Journal of Quality & Reliability Management*, *17*(9), 937-949.

[12] D. . Rizqi Aulia and T. N. . Wijayaningrum, "Statistical Quality Control (SQC) Method Analysis Regarding Qualty Control of Shoe Products (Case Study of PT-X)", J Statistika, vol. 17, no. 1, pp. 691–702, Jul. 2024.

[13] Zurqoni and F. . Rahman, "Aplikasi Statistik Proses Kontrol Untuk Evaluasi Proses Pembelajaran Menggunakan Diagram Kontrol $\bar{X}$ dan R ", J Statistika, vol. 12, no. 2, pp. 30–35, Dec. 2019.

Potential Productivity of Quail Farming in Ngaliyan, Pulutan, Wonosari, Gunungkidul

Annida Purnamawati[(1)] , Wawan Nugroho [(2)] , Raffi Fa`Iq Firjatullah [(3)] , Mochamad Akbar Garudea Syahroni [(4)]

[(1,3,4)] Bina Sarana Informatika University Yogyakarta, Jl. West Ringroad, Gamping, Yogyakarta,
[(2)] Bina Sarana Informatika University Surakarta, Jl. Letjen Sutoyo No.43, Surakarta City
email: annida.npr@bsi.ac.id

ABSTRAK

Burung puyuh merupakan jenis hewan unggas yang mempunyai potensi produktivitas untuk dapat dikembangkan dan ditingkatkan dalam segi pemeliharaan demi terwujudnya produksi yang maksimal. Penelitian bertujuan menganalisis apasaja yang menjadi faktor pengaruh dari produksi usaha peternakan burung puyuh petelur yang ada di Ngaliyan, Pulutan, Wonosari, Gunungkidul. Jika rendahnya produksi telur burung puyuh menjadi kendala bagi peternak, hal ini dapat menghambat upaya untuk meningkatkan produktivitas dan memenuhi kebutuhan konsumen. Tujuan dari penelitian untuk mengidentifikasi berbagai faktor apasaja yang memengaruhi produksi telur burung puyuh. Penggunaan pendekatan gabungan antara metode kualitatif dan kuantitatif dengan menerapkan metode regresi linier berganda. Hasil penelitian mendapatkan pendapatan dari produksi telur burung puyuh Rp193.011.950 per periode usaha (15 bulan). Analisis regresi linier berganda menunjukkan bahwa kedua variabel independen, yaitu biaya pakan ($\boldsymbol{X_1}$) dan biaya tenaga kerja ($\boldsymbol{X_2}$), berpengaruh signifikan terhadap produksi telur burung puyuh (Y), dengan nilai signifikansi < 0.001. Variabel biaya pakan ($\boldsymbol{X_1}$) memiliki pengaruh positif dominan, sedangkan biaya tenaga kerja ($\boldsymbol{X_2}$) memiliki pengaruh negatif. Hubungan antara kedua variabel independen terhadap produksi telur mencapai 88.1%, dengan 11.9% dipengaruhi oleh variabel lain di luar penelitian. Model regresi yang digunakan kuat dan valid, menunjukkan pengaruh signifikan kedua variabel independen terhadap variabel dependen.

Kata kunci: pendapatan, faktor produksi, burung puyuh, regresi linier berganda, spss.

ABSTRACT

Quails are a type of poultry with significant productivity potential that can be developed and improved through enhanced care to achieve optimal production. This study aims to analyze the factors influencing the production of laying quail farms in Ngaliyan, Pulutan, Wonosari, Gunungkidul. If low egg production becomes an obstacle for farmers, it can hinder efforts to increase productivity and meet consumer demand. The purpose of this research is to identify the various factors influencing quail egg production. The study employs a mixed-method approach, combining qualitative and quantitative techniques, with multiple linear regression analysis. The results indicate that income from quail egg production is IDR 193.011.950 per business period (15 months). Multiple linear regression analysis shows that both independent variables, feed cost (X_1) and labor cost (X_2), have a significant effect on quail egg production (Y), with a significance level of <0.001. The feed cost variable (X_1) has a dominant positive influence, while labor cost (X_2) has a negative influence. The relationship between the two independent variables and egg production is 88.1%, with 11.9% influenced by other factors outside of the study. The regression model used is strong and valid, demonstrating the significant impact of both independent variables on the dependent variable.

Keywords: *income, production factors, quail, multiple linear regression, spss.*

INTRODUCTION

Based on data from the Coordinating Ministry for Economic Affairs, the livestock sub-sector contributes positively to the national economy by providing growth of 0.34% and contributing 1.58% to GDP [1] . One of the direct benefits obtained from the livestock sub-sector is animal food that has high value, especially as a source of animal protein. Along with population growth, education and public awareness also increase where the importance of nutrition and the role of nutrients in life, especially in children. So the higher the needs of the community also have an impact on the livestock sub-sector so that it has a positive impact on improving nutritional quality and also provides benefits for livestock businesses [2] .

Quail farming is one of the livestock sub-sectors that has the potential to produce eggs and meat, and is a source of animal food that is rich in nutrition [3] . Quail, in addition to producing eggs and meat, quail droppings can also increase income for livestock business actors. Quail can be cultivated by the community as a source of daily income, although the marketing of quail eggs is not as large as chicken eggs due to different selling prices, but quail farming can also have great potential to be developed [4] .

Various studies have examined analytical methods for processing poultry farm data. Research conducted by Naurah examined the financial feasibility of a quail egg farming business in Pringsewu District, Pringsewu Regency, by applying various approaches, such as NPV, IRR, and sensitivity analysis. The results of the analysis showed that this business is vulnerable to decreased production, fluctuations in selling prices, and increased feed costs [5] . Another study by Boni reviewed the feasibility of a quail egg farming business in Tebing Tinggi Okura Village, Rumbai Pesisir District, Pekanbaru, this study reviewed the financial aspects using investment criteria such as NPV, IRR, Net B/C Ratio, PBP, and BEP. The findings showed that the business was classified as feasible to run with an NPV of IDR 64.518.459, an IRR of 18.94%, and a Net B/C Ratio of 1.2 [6] . However, this study specifically explores the effectiveness of multiple linear regression method which can help in determining the potential factors in quail farming.

This study aims to reveal the effectiveness of analysis techniques on quail farming. The background of the research is to see today's society who are enthusiastic about the nutritional value of children in order to support maximum growth and development. As well as seeing the closest local livestock environment, namely quail farming. In addition, the author hopes that the results of the study can add to the contribution of research developers to always be able to improve the quality of studies in the future. Through this study, the author also hopes that readers can understand more deeply regarding the implementation of the multiple regression linear method effectively.

METHOD

This research was conducted in Ngaliyan, Pulutan, Wonosari, Gunungkidul. The location was chosen intentionally because it is one of the quail egg production farms in Wonosari District, Gunungkidul Regency. The stages carried out in the research are described in Figure 1.

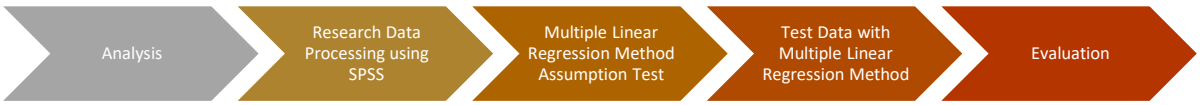


Figure 1. Research Stage

Explanation of the research stages in Figure 1 as follows:

1. Analysis

Writer adopt technique data collection based quantitative through survey literature academic in the field science related, aimed at For to obtain relevant concepts so that can support innovation study use development policy public [9]. In the research this is also done approach qualitative, so the data obtained is results from interview. Interview can give outlook deep about perspectives, experiences and information that are not can obtained through other methods such as survey. Data Preparation. At the stage This collect data with ensure all the necessary data for analysis Already collected and in a suitable format (example: Excel, CSV, etc.). Identification variable dependent *(Y)* and variables independent (X_1 and X_2). Turning off the complete and clean data, doing cleaning data from outliers, missing values, or data input error.

2. Research Data Processing using SPSS

Study This use SPSS tools for help in apply method multiple linear regression. Research data processing using SPSS is very easy for analyzing quantitative data with method statistics, such as multiple linear regression.

3. Multiple Linear Regression Method Assumption Test

The process of evaluating whether a multiple linear regression model meets the necessary statistical assumptions to ensure accurate and reliable results. These assumptions are critical for the validity of the model's coefficients, predictions, and associated statistical inferences

a. Normality Test
   To ensure that the residual data is normally distributed by visualizing it using a residual histogram or P-Plot.
b. Heteroscedasticity Test
   To test heteroscedasticity that the residual variance is homogeneous so that there is no heteroscedasticity pattern. The test results can be seen with the residual plot against the prediction or scatterplot.
c. Multicollinearity Test
   To ensure that there is no strong linear relationship between the independent variables depicted by looking at the VIF Value <10 indicating no multicollinearity.
d. Autocorrelation Test
   Autocorrelation test is a process to identify whether there is a relationship between the residuals (prediction errors) of one observation with the residuals of other observations in a regression model.

4. Test Data with Multiple Linear Regression Method

Multiple linear regression refers to a statistical technique that uses two or more independent variables to predict the outcome of a dependent variable. The technique enables analysts to

determine the variation of the model and the relative contribution of each independent variable in the total variance.

a. Determination Coefficient Test (R^2))
   Measures how much variation in the dependent variable can be explained by the independent variables. Ensures the R² value ranges from 0 to 1; the closer to 1, the better the model is at explaining the data.
b. ANOVA Test/Simultaneous Test (F Test)
   Testing whether independent variables simultaneously (together) have a significant effect on the dependent variable. By looking at the calculated F with the F table or p-value with a significance level.
c. Partial Test (t Test)
   Testing the effect of each independent variable on the dependent variable. To see the calculated t value with the t table or p-value with a significance level.

The variable limitations that the author uses in the study using the SPSS multiple linear regression method, namely using the feed cost variables (X_1) and labor costs (X_2) to carry out calculations based on related variables which are things that can affect the production of quail eggs.

The calculation of income from quail egg farming in the study was carried out using the formula for cost of receipt and profit. With the ultimate goal of achieving maximum profit, so that the farm tries to maximize how to produce quail egg production. By considering several things to obtain input (total cost) and the amount received, namely egg production (total income). The profit formula with total income minus production costs is as follows [7] :

$$\boldsymbol{\pi = TR - TC} \quad \textbf{(1)}$$

Information:

π = Business income (Rp/period)
TR(Total Revenue) = Income (RP/period)
TC((Total Cost) = Production Cost (RP/period)

The multiple linear regression model formula used in nature to carry out analysis in looking at production factors is as follows [8] :

$$\boldsymbol{Y = a + b_1X_1 + b_2X_2 + e} \quad \textbf{(2)}$$

Information:

A = Constant
b_1 = Product variable regression coefficient
b2 = Regression coefficient of price variable
Y = Egg production
X_1 = Amount of feed
X_2 = Labor costs
e = Error (independent variable)

5. Evaluation

Evaluation in the research stage of multiple linear regression models is a process to assess whether the regression model built has met statistical criteria and can be relied on to explain the relationship between independent variables (predictors) and dependent variables (responses). This

evaluation involves several important steps aimed at ensuring the validity, accuracy, and usefulness of the model.

RESULTS AND DISCUSSION

Livestock business bird quail started in 2009, background behind from election business This Because with see opportunity big in cultivation farm bird quail in society. With supplies information and knowledge obtained from society and online media then decide for try manage business farm bird quail. The beginning cultivation bird quail as many as 1000 tails the longer the more develop business farm bird quail. Until now cultivation bird quail can develop with Good from in terms of production volume and marketing. Until Now 2024 cultivation egg bird quail with total of 8000 tails. This business is managed directly by the owner namely Mr. Marino and his wife. If there needs urge so There is power Work addition with payment system daily. According to Mr. Marino business this is one of them prospective business, on the other hand he also has a livelihood as a farmer.

A number of assumptions obtained in discussion cost production and implementation linear regression method multiple done in analysis:

1. Seeds egg bird quail purchased 1 month old
2. Bird quail Already Ready lay eggs age 40-45 days and effectiveness lay eggs up to 15 months.

Cost production in business cultivation egg quail reaching Rp475.238.050 per period business. Details cost includes: seeds amounting to Rp64.000.000, feed Rp257.394.300, vitamins Rp30.768.750, energy work Rp92.306.250, and depreciation Rp30.768.750 per period business. From the percentage cost production, feed become component the biggest with 54.16%, followed by power Work by 19.43%.

Reception in context business or business often refers to total revenue or revenue earned from sale egg bird quail before reduced cost others. Acceptance is results from the total production multiplied with price per unit of product such as, for example reception from results sale egg in farm bird quail. Acceptance This reflect flow enter funds into in organization or company, which will later use for cover cost operational and productive profit clean after reduced total cost [10] .

Following This is the data obtained in study cultivation egg bird quail with data calculation per period business which is 15 months.

Table 1. Income Breeder Bird Quail Egg Layers Per Period (15 months) in Ngaliyan, Pulutan, Wonosari, Gunungkidul 2024

Description	Unit	Amount
1) Production	Kg/ period business	24,750
2) Price	Rp/kg	27,000
3) Acceptance (AxB)	Rp/ period business	668,250,000
4) Cost Production	Rp/ period business	475,238,050
5) Income (CD)	Rp/ period business	193,011,950

In table 1, it is stated that the average production egg bird quail per day an average of 55 kg, with price sell Rp27.000 per kg. The total production egg during One period business reached

Rp668.250.000. Profit business egg bird quail counted with reduce total revenue of the total cost production so Rp668.250.000–Rp475.238.050, so that obtained profit amounting to Rp193.011.950 per period business. Based on cost data analysis production, it is known that component the biggest in cost business cultivation egg quail in Ngaliyan, Pulutan, Wonosari, Gunungkidul is feed and energy work. For know how far has it changed cost feed and energy Work influence production egg quail, done analysis multiple linear regression. Analysis results using SPSS software is presented as follows:

1. Multiple Linear Regression Method Assumption Test

Before do multiple linear regression, must Conduct Assumption Test Classic For ensure results valid regression, with:

a. Normality Test

Normality test in multiple linear regression an assumption test classic aiming ensured that the residual (difference) between predicted value and actual value actually). This is important done due to the statistical tests used in Linear regression assumes that the residuals are normally distributed. If the assumption test violated, result analysis can invalid.

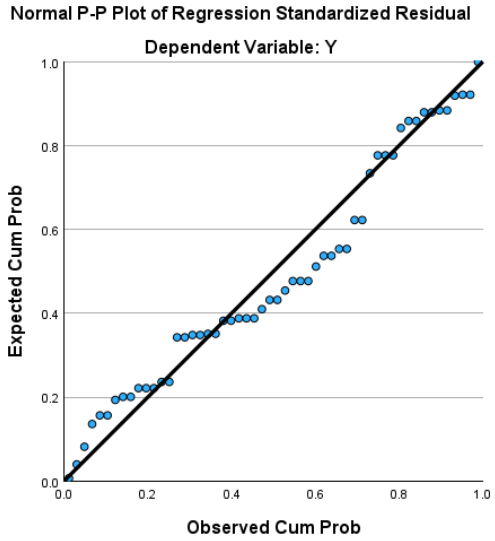


Figure 2. Normality Test

In Figure 2 the data is said to be normal if all data in motion following (approaching) the diagonal line.

b. Heteroscedasticity Test

Heteroscedasticity test in multiple linear regression aiming For see is residual (error) prediction) has constant variance or no. Assumption base in linear regression is that the residual must be own the same variance (homoscedasticity). If the residual variance is not constant (heteroscedasticity), results regression Can So less valid or biased.

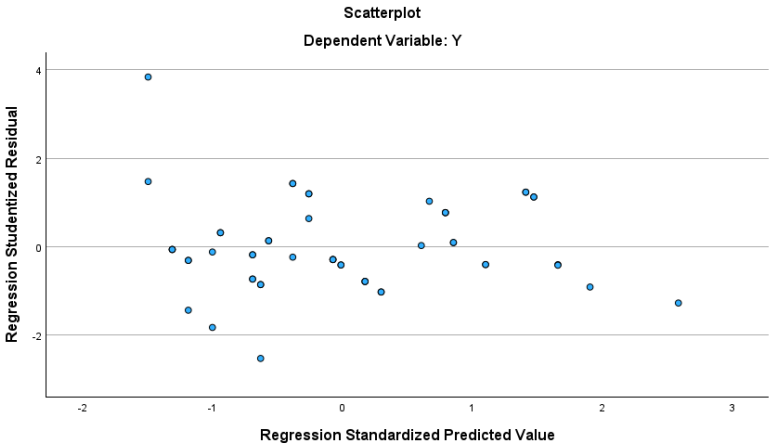


Figure 3. Heteroscedasticity Test

In figure 3 the data is stated safe from symptom heteroscedasticity Because all data is spread to all over part graph .

c. Multicollinearity Test

Multicollinearity test was conducted for identify whether there is very strong linear relationship between the independent variables in a regression model multiple.

Table 2. Multicollinearity Test

Coefficients [a]

Model		VIF
1	(Constant)	
	X1	1.026
	X2	1.026

a. Dependent Variable: Y

In table 2 is multicollinearity test results it is no happen Because VIF value of each variable more low of 10.

d. Autocorrelation Test

Autocorrelation test in multiple linear regression used to test whether there is correlation between residuals (error prediction) in a regression model.

Table 3. Autocorrelation Test

Runs Test

	Unstandardized Residual
Test Value [a]	-5184.82711
Cases < Test Value	27
Cases >= Test Value	27
Total Cases	54

Number of Runs	32
Z	1.099
Asymp . Sig. (2-tailed)	.272

a. Median

In table 3 the results of the autocorrelation test it is said no occurs because results from auto correlation test show mark asymp sig. more tall from 0.05 namely of 0.272.

1. Multiple Linear Regression Method Test

a. Coefficient Test Determination (R^2)

To know magnitude the influence exerted by all independent variables (variable X) on the dependent variable (variable Y).

Table 4. Coefficient Test Determination

Model Summary [b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.938 [a]	.881	.876	70328.755	2.418

a. Predictors: (Constant), X2, X1

b. Dependent Variable: Y

In table 4 the results of the coefficient test determination ability from all $X_1$and X_2(independent variables) together in influencing the Y variable (dependent variable), from the coefficient test determination produce by 88.1%. Then the rest by 11.9% means can explained by other factors that are not included in study This .

b. ANOVA test/ simultaneous test (F test)

To know significant the influence exerted by all independent variables (variable X) on the dependent variable (variable Y).

Table 5. ANOVA test

ANOVA [a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1861080511344.988	2	930540255672.494	188.135	<.001 [b]
	Residual	252252821988.345	51	4946133764.477		
	Total	2113333333333.334	53			

a. Dependent Variable: Y

b. Predictors: (Constant), X2, X1

In table 5 we can see seen Sig. value < 0.05 and F count Far more big than F table value, we can conclude that the regression model in a way overall significant. This means that, X_1 and X_2 together have a significant influence on the dependent variable Y.

c. Partial Test (t Test)

To know significance the influence exerted by each independent variable (variable X) on the dependent variable (variable Y).

Table 6. Partial Test

Coefficients [a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	11900121.319	2683811.153		4.434	<.001
	X1	1.063	.062	.841	17.168	<.001
	X2	-6.677	1.080	-.303	-6.180	<.001

a. Dependent Variable: Y

Table 6 shows result partial test analysis, to know the value of the independent variable (X) individually against the dependent variable *(Y)*. Variable it is said influential in a way significant if mark significant <0.05. So it can be seen in the table mark variable X_1 and X_2 influential significant on the dependent variable *(Y)*. From the analysis the obtained the equation model multiple linear regression as following:

$$Y = 11{,}900{,}121{,}319 + 1{,}063\ X_1 - 6{,}677\ X_2 + e$$

a. Constant (Intercept)

B (constant)= 11.900.121.319 shows intercept value. This is mark from Y (variable dependent) when all variable independent ($X_1$and X_2) are zero. With a t value of 4.434 and sig <0.001, this constant is statistically significant. That is, there is a basis for stating that this value is significantly different from zero in the model.

b. Coefficient X_1(first independent variable)

B (X_1) = 1.063 indicates that every one unit increase in X_1(the first independent variable) will increase Y (the dependent variable) by 1.063 units, assuming the other variables remain constant. The value of t = 17.168 and Sig < 0.001 indicate that it $X_1$is very significant. In other words, the variable X_1has a strong influence on the Y variable.

c. Coefficient X_2(Second Independent Variable)

B (X_2) = -6.677 shows that every increase one unit in X_2 (the second independent variable) will decrease Y by 6.677 units, assuming the other variables remain constant. The t value = -6.180 and sig < 0.001 indicate that X_2 is also statistically significant. That is, X_2 it has a significant effect on Y, even though the effect is negative.

CONCLUSION

Reception from business egg bird quail in Ngaliyan, Pulutan, Wonosari, Gunungkidul reaching Rp668.250.000 per period business, with cost production amounting to Rp475.238.050. This resulted in an income of Rp193.011.950 per period business. Based on results analysis from method multiple linear regression the can with conclusion that in a way overall, both variable independent (X_1 and X_2) have a significant effect on the dependent variable Y because the Sig value for both variables is < 0.001 (meaning it is smaller than 0.05, the general significance limit). X_1 has a significant positive influence on Y, while X_2 is also statistically significant for Y, although its

influence is negative. This model Can it is said Enough strong Because second variable independent significant. Two variables were studied, namely cost feed (X_1), indicating the most dominant influence to production egg bird quail. Index correlation between variable amount feed (X_1) and costs power work (X_2) against production egg bird quail (Y) is 88.1%, while the remaining 11.9% influenced by other variables that are not including in study this. In overall, regression model This can considered strong and valid, with second variable independent own significant contribution to prediction variable dependen.

ACKNOWLEDGEMENT

This research was funded by the Institute for Research and Community Service, Bina Sarana Informatika University (LPPM-UBSI).

REFERENCES

[1] H. Limanseto, "Support for Integrated Farming Corporations with Cluster Models to Drive the Rural Economy," *https://www.ekon.go.id/* , 2022. https://www.ekon.go.id/publikasi/detail/4020/dukungan-bagi-korporasi-peternakan terintegrasi-model-klaster-untuk-gerakkan-ekonomi-pedesaan#:~:text=The livestock subsector itself also contributes,1%2C58%25 to GDP.

[2] S. Maryam *et al.* , "ANALYSIS OF PURCHASE DECISION MAKING OF QUAIL (Coturnix coturnix) CARCASSES ON CONSUMER BEHAVIOR IN SUKABUMI REGENCY," vol. 9, no. 1, pp. 1371–1383, 2023.

[3] T. Estining, MDS T, FR Industri, U. Telkom, P. Puyuh, and P. Eka, "Feasibility Analysis of the Development of Cultivating Quail Eka Farm in Bogor," *e-Proceeding Eng.* , vol. 4, no. 2, pp. 2–7, 2017.

[4] Riyanti, K. Nova, and MMP Sirat, *Production of Various Poultry Livestock* . 2020. [Online]. Available: http://repository.lppm.unila.ac.id/30686/7/BUKU PRODUKSI ANEKA TERNAK UNGGAS.pdf

[5] N. Nisrina, MI Affandi, and L. Marlina, "Financial Feasibility Analysis of Quail Egg-Laying Business in Pringsewu District, Pringsewu Regency," *J. Animal Science, Padjadjaran University* , vol. 22, no. 2, p. 137, 2022, doi: 10.24198/jit.v22i2.40491.

[6] B. Sanjaya, H. Yasid, P. Studi, A. Faculty, P. University, and L. Kuning, "Sanjaya Previous Research and Bep," vol. 13, no. 1, pp. 47–58, 2016.

[7] AF Syahputra and I. Wicaksana, "Financial Analysis of Tilapia Fish (Oreochromis Niloticus) Seeding Business at the Pendem Fish Seed Center, East Java," *J. Agrimanex Agribusiness, Rural Manag. Dev. Ext.* , vol. 3, no. 1, pp. 37–46, 2022, doi: 10.35706/agrimanex.v3i1.7000.

[8] AA-F. Nur Wahyudin, A. Primajaya, and ASY Irawan, "Application of Multiple Linear Regression Algorithm on Astra Isuzu Car Sales Estimation," *Techno.Com* , vol. 19, no. 4, pp. 364–374, 2020, doi: 10.33633/tc.v19i4.3834.

[9] A. Purnamawati, MN Winnarto, and M. Mailasari, "Cart Analysis (Classification and Regression Trees) for Predicting Bicycle Users Based on Weather," *J. Teknoinfo* , vol. 16, no. 1, p. 14, 2022, doi: 10.33365/jti.v16i1.1478.

[10] D. Saadudin, Y. Rusman, and C. Perdani, "ANALYSIS OF COST, INCOME AND R/C OF GINGER FARMING (Zingiber officinale)," *J. Ilm. Mhs. Agroinfo Galuh* , vol. 3, no. 2, p. 85, 2017, doi: 10.25157/jimag.v3i2.216.

Analysis of Public Perception of Dynastic Politics in the 2024 Presidential and Vice Presidential Elections in Indonesia with a *Chi-Square Approach*

Nur Chamidah [(1)], Firda Aulia Pratiwi [(2)], Naura Qathrunnada Bi Arafah [(3)], Bintang Alyaa Sabila [(4)], Sheila Sevira Asteriska Naura [(5)]

[1,2,3,4,5]Departement of Mathematics, Faculty of Science and Technology, Universitas Airlangga

Jl. Dr. Ir. H. Soekarno, Mulyorejo, Surabaya, Jawa Timur

e-mail: nur-c@fst.unair.ac.id[(1)], firda.aulia.pratiwi-2021@fst.unair.ac.id[(2)], naura.qathrunnada.bi-2021@fst.unair.ac.id[(3)], bintang.alyaa.sabila-2021@fst.unair.ac.id[(4)], sheila.sevira.asteriska-2022@fst.unair.ac.id[(5)]

ABSTRAK

Fenomena politik dinasti dalam Pemilihan Presiden dan Wakil Presiden 2024 menjadi isu hangat di Indonesia. Politik dinasti merupakan pewarisan kekuasaan dalam keluarga untuk mempertahankan pengaruh politik, yang berpotensi mengancam demokrasi dan pembangunan negara. Penelitian ini bertujuan mengidentifikasi hubungan antara persepsi masyarakat terhadap politik dinasti berdasarkan karakteristik responden, sejalan dengan SDGs poin ke-16 yaitu, perdamaian, keadilan, dan kelembagaan yang tangguh, jika politik dinasti tidak transparan maka dapat menghambat pemerintahan yang baik. Data diperoleh melalui kuesioner yang didistribusikan kepada 210 responden, kemudian dianalisis menggunakan uji *Chi-Square* untuk mengukur hubungan persepsi masyarakat terhadap politik dinasti dengan karakteristik jenis kelamin, profesi, dan wilayah kependudukan. Hasil analisis menunjukkan bahwa persepsi masyarakat tidak memiliki hubungan signifikan dengan jenis kelamin, namun signifikan terkait dengan profesi dan wilayah kependudukan pada salah satu pernyataan, yaitu setuju bahwa keberadaan dinasti politik berdampak negatif terhadap demokrasi Indonesia.

Kata kunci: Politik Dinasti; Politik Demokrasi; *Chi-square*

ABSTRACT

The phenomenon of dynastic politics in the 2024 Presidential and Vice Presidential Elections has become a hot issue in Indonesia. Dynastic politics is the inheritance of power in the family to maintain political influence, which has the potential to threaten democracy and state development. This study aims to identify the relationship between public perceptions of dynastic politics based on the characteristics of respondents, in line with SDGs point 16, namely, peace, justice and resilient institutions, if dynastic politics is not transparent it can hinder good governance. Data was obtained through questionnaires distributed to 210 respondents, then analyzed using the Chi-Square test to measure the relationship between public perceptions of dynastic politics and the characteristics of gender, profession, and region of residence. The results of the analysis show that public perception does not have a significant relationship with gender, but is significantly related to profession and region of residence on one of the statements, namely agreeing that the existence of political dynasties has a negative impact on Indonesian democracy.

Keywords: *Dynasty Politics; Democratic Politics; Chi-Square.*

INTRODUCTION

Indonesia is a developing country that adheres to a democratic system characterized by public participation through general elections to determine leaders and representatives of the people. Democracy, derived from the Greek demos (people) and kratos (power), means "government by the people" or "the people in power," namely government of the people, by the people, and for the people [1]. Indonesia as a developing country always faces challenges regarding political dynamics, including the issue of dynasty politics in the 2024 Presidential and Vice Presidential Elections. According to Dedi (2022), dynasty politics refers to formal political power that is inherited in families or relatives for more than one generation to maintain the dominance of power [2]. Dynasty politics is a double-edged phenomenon. On the one hand, this practice can violate the principles of democracy and human rights by depriving others of their rights. On the other hand, although it is considered normal if filled by competent individuals, dynasty politics is still prone to causing imbalances due to the mixture of family and community interests [3].

The issue of dynasty politics has drawn pros and cons because it is considered to be contrary to democracy in Indonesia and benefits certain families or groups in power [4]. Dynasty politics has been a topic of previous research, in studies *"Politik Dinasti di Indonesia: Tinjauan Kritis terhadap Penerapan Demokrasi di Era Kepemimpinan Presiden Jokowi"* by Syanur et al. (2022). This study states that dynasty politics are developing in Indonesia, especially in the Jokowi era, with members of the president's family involved in formal politics. This practice is considered to violate democratic ethics because it prioritizes family ties, which can hinder public participation and damage the democratic system [4]. Another study entitled *"Dampak Dinasti Politik terhadap Culture Demokrasi Pancasila di Indonesia"* by Gadafi and Nelwati (2022) discusses the impact of political dynasties on Pancasila democracy. This study shows that political dynasties bring positive and negative dynamics, with the risk of exacerbating inequality in the implementation of democracy in Indonesia [5]. Although the study relied on qualitative analysis to assess the impact of dynasty politics, this study will continue the previous study by using the chi-square method to test the relationship between public perceptions of dynasty politics based on demographic characteristics. The aim is to identify significant differences in perceptions of dynasty politics across different community groups.

The chi-square approach is one of the statistical methods that is often used in various disciplines. The chi-square test tests the match between the observed frequency and the expected frequency based on certain calculations [6]. This method has the ability to analyze data that is categorical or discrete [7]-[8] Several studies related to the use of the chi-square method, namely by Kleden et.al, (2023) regarding the relationship between demographic factors and the incidence of malaria in East Wewewa sub-district [9], research by Widayanti & Kusumawati related to perceptions of vaccine effectiveness with attitudes of willingness to participate in COVID-19 vaccination [10]. Previous research that is relevant to discussing public perception of leadership using chi-square test analysis is research by Nurhalimah (2022) entitled "Public Perception of Female Village Head Leadership" using the Chi-square test to analyze public perception. The results showed no significant difference based on gender, but there was a significant difference based on age, indicating that perception of leadership is influenced by age [11]. Therefore, based

on the description above, the author is interested in conducting research on "Public Perception of Dynasty Politics in the 2024 Presidential and Vice President Elections in Indonesia using the Chi-Square approach". This is also in line with the Sustainable Development Goals (SDGs) point 16, namely peace, justice, and resilient development because the existence of non-transparent dynasty politics will be an obstacle for the government in Indonesia in implementing the SDGs [12].

METHOD

This research is a quantitative study that uses primary data obtained directly from 210 sample respondents with an Indonesian population and collected online through Google Form. The sampling method in this study uses purposive random sampling because the researcher sets criteria considerations for the research sample. This research was conducted over a period of one month from November 2 to November 30, 2023. The analysis technique used in this study is the Chi-Square Test. According to Nugraha (2013), For a two-way contingency table, the chi-square test statistic calculation formula is used as follows [13]:

$$\chi^2 = \sum_{i=1}^{b}\sum_{j=1}^{k}\frac{(n_{ij}-e_{ij})^2}{e_{ij}} \sim \chi^2(v)\,; v = (b-1)(k-1) \quad (1)$$

with n_{ij} being the observation frequency in the i-th row and j-th column and e_{ij} being the expected frequency in the i-th row and j-th column. If there is a relationship between variables, the Cramer's V contingency coefficient test is conducted on a b×k contingency table where b and k are greater than two and asymmetric (b≠k). According to Everitt (1993), the calculation statistics for the Cramér's V contingency coefficient test are as follows [14]:

$$V = \sqrt{\frac{\chi^2}{n.\min{(b-1,k-1)}}} \quad (2)$$

The steps used to collect data to answer the research questions can be outlined through the following flow chart:

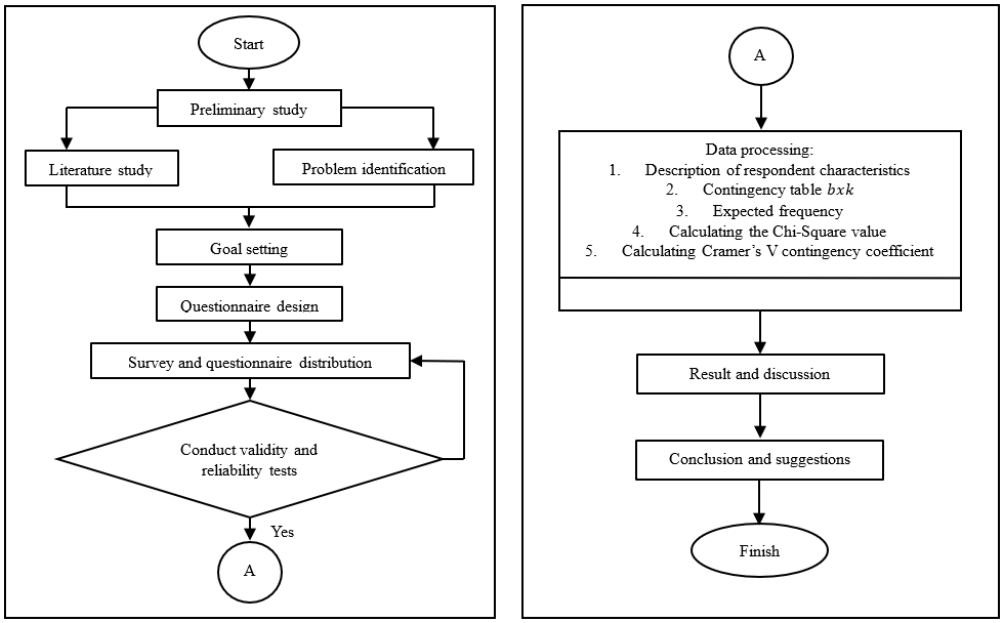


Figure 1. Research Flow Chart

Next is the most crucial stage in the research procedure, namely the data analysis process with the following steps:

1. Conduct validity and reliability tests based on answers from respondents
2. Interpret descriptive statistics of the data obtained
3. Create a contingency table in the form of rows and columns for each of the community characteristic variables to be analyzed, namely gender, profession, and residency area and calculate the total number of row values and the total number
4. Calculating the expected frequency value of each cell using the expected frequency formula
5. Calculating the *Chi – Square* value
6. Comparing the calculated *Chi – Square* value with the Chi – Square table $\chi^2{}_{\alpha(k-1)(s-1)}$
7. Making a decision with the criteria: Reject if the value $\chi^2{}_{value} > \chi^2{}_{\alpha(k-1)(s-1)}$
8. Calculating the value of *Cramers's V* contingency coefficient
9. Interpreting the research data

RESULTS AND DISCUSSION

Based on the research that has been done, the results of the research questionnaire contain four main statements that are indicators of public perception of dynastic politics, namely

1. I agree that dynastic politics has an effect on the election of the President and Vice President of Indonesia in 2024.
2. I agree that the existence of political dynasties has a negative impact on Indonesian democracy.
3. I agree that members of political dynasties get political positions because of family relationships, not because of their own abilities.
4. I agree that dynastic politics produces policies that do not benefit society, because it is only concerned with their own group.

Each component of the questionnaire statement consists of two answer choices with detailed scores: Disagree (1), agree (2).

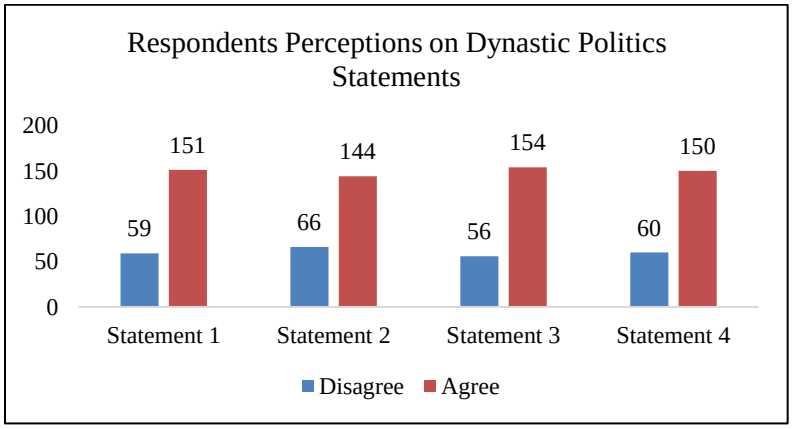


Figure 2. Descriptive Statistics of the Number of Respondents for Each Statments

It can be seen from Figure 2**.** As many as 210 respondents who had filled out this research questionnaire, in statement 1 received a disagreeing answer by 59 (28%) respondents and agreed by 151 (72%) respondents. In statement 2 received a disagreeing answer by 66 (31.43%) respondents and agreed by 144 (68.57%) respondents. In statement 3 received a disapproving answer by 56 (27%) respondents and agreed by 154 (73%) respondents. In statement 4, they received a disapproving answer by 60 (28.57%) respondents and agreed by 150 (71.43%) respondents. It can be concluded overall that the dominant respondents answered in agreement to the statements given.

1. Validity Test

Validity test is a test used to determine and test the accuracy and determination of a measuring instrument. The hypotheses used are [15] :

$H_0 : \rho = 0$ (The measuring instruments/questions used in the questionnaire are invalid)

$H_1 : \rho \neq 0$ (The measuring instruments/questions used in the questionnaire are valid)

Test statistics are used with the rejection area if the H_0 *p-value is* $\leq \alpha$ (0.05).

Table 1. Validity Test on All Questions

Variable	*P-Value*	Decision	Conclusion
Statement 1	0.000	**Reject H_0**	Valid
Statement 2	0.000	**Reject H_0**	Valid
Statement 3	0.000	**Reject H_0**	Valid
Statement 4	0.000	**Reject H_0**	Valid

Based on the results of the validity test, the *p-value* for the four statements in the questionnaire is 0.000 so that it can be concluded that all category statement items on the questionnaire are valid. Therefore, such statements can be used in questionnaires.

2. Reliability Test

After performing the validity test, the next step before entering the test *Chi-square* is to conduct reliability tests. A variable is said to be good if it has a value *Cronbach's Alpha* > of 0.6 [16].

Table 2. Reliability Test

Category	Cronbach's Alpha value	Decision
Respondents' Perceptions of Dynastic Politics in the 2024 Presidential and Vice Presidential Elections in Indonesia	0.915	Reliable

Based on the results of the Reliability test, the reliability value of 0.915 which is classified as very high value so that it can be concluded that the questionnaire is considered capable of revealing information.

3. Descriptive Statistics

a) Results of analysis of community characteristics based on gender

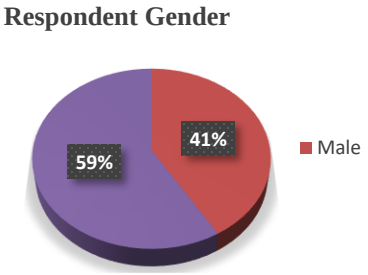


Figure 3. Percentage of Respondents by Gender

Based on Figure 1. From 210 respondents, it can be seen that the gender of people who have filled out questionnaires related to public perceptions of dynastic politics in the 2024 Presidential and Vice Presidential Elections in Indonesia is 41% or 87 male respondents and 59% or 123 respondents are female.

b) Results of analysis of community characteristics based on profession (Job type)

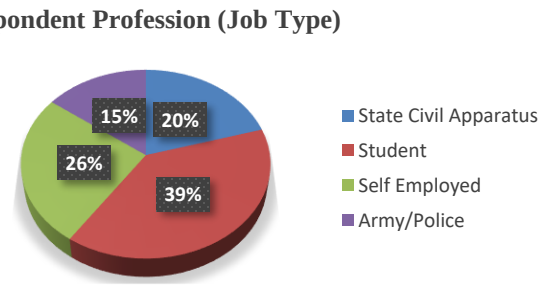


Figure 4. Percentage of Respondents by Profession (Job Type)

Based on Figure 2. From 210 respondents, it can be seen that the type of community work that has filled out questionnaires related to public perception of dynastic politics in the 2024 Presidential and Vice Presidential Elections in Indonesia is 20% or 43 respondents of the State Civil Apparatus, of 39% or 82 respondents of Students / Students, of 26% or 54 respondents of Self-Employed, and of 15% or 31 respondents of Members of the TNI / Polri.

c) Results of analysis of community characteristics based on residency area

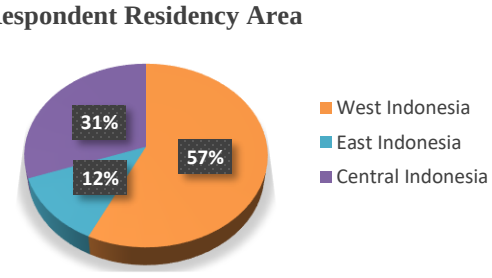


Figure 5. Percentage of Respondents by Residency Area

Based on Figure 3. From 210 respondents, it can be seen that the population area of the people who have filled out questionnaires related to public perceptions of dynastic politics in the 2024 Presidential and Vice Presidential Elections in Indonesia is 57% or 120 respondents of western Indonesia, of 12% or 26 respondents of eastern Indonesia, of 31% or 64 respondents of central Indonesia.

4. *Chi-Square* Independence Test

This research uses the *Chi-Square* test to examine the relationship between public perceptions of dynastic politics in the 2024 elections and three demographic characteristics: gender, profession, and region of residence. The following are the results of the analysis for each statement based on the Chi-Square test:

a. The Independence Test was conducted on statement 1, which agreed that dynasty politics influenced the election of the President and Vice Presidential of Indonesia in 2024. The results showed that 75% of men and 70% of women agreed with this statement. Based on profession, students had the highest agreement rate of 75%, followed by TNI/Polri Members (74%), Civil Servants (70%), and Private Workers (67%). In addition, based on the area of residence, respondents from central Indonesia showed the highest agreement rate (80%), followed by western Indonesia (70%) and eastern Indonesia (65%). *Chi-Square* Test results are obtained from statements 1 to 3 characteristics in the following table 3.

Table 3. Chi-Square Test Result Statement 1

Characteristics	Chi-square Value	Chi-square Table	Decision
Gender	0.57969	3.84145	Accept H_0
Profession	3.06721	11.07049	Accept H_0
Residency Area	2.91138	5.99146	Accept H_0

Based on Table 3, the results of the Chi-Square test show that there is no significant relationship between statement 4 and gender, profession, or area of residence. The Chi-Square value for all characteristics is smaller than the table value, so the null hypothesis is accepted. This means that the agreed perception that dynastic politics influences the 2024 Indonesian Presidential and Vice Presidential elections is not influenced by these three demographic characteristics.

b. The Independence Test was conducted on statement 2, which agreed that the existence of political dynasties has a negative impact on Indonesian democracy. The results showed that 71% of men and 70% of women agreed with this statement. Based on profession, respondents from the Private Worker profession showed the highest level of agreement, which was 81%. This was followed by Students (69%), Civil Servants (58%) and TNI/Polri Members (58%). In addition, based on area of residence, respondents from western Indonesia had the highest level of agreement, which was 79%. Meanwhile, respondents from central and eastern Indonesia had lower levels of agreement, at 54% and 53% respectively. *Chi-Square* Test results are obtained from statements 2 to 3 characteristics in the following table.

Table 4. Chi-Square Test Result Statement 2

Characteristics	Chi-square Test Result	Chi-square Table	Decision	Cramer's V
Gender	0.49982	3.84145	Accept H_0	-
Profession	11.24943	11.07049	**Reject H_0**	0.23145
Residency Area	14.59124	5.99146	**Reject H_0**	0.26359

Based on Table 4, the follow-up test is because the results obtained are **Reject H0** on the characteristics of the profession and recidency area**,** where there is at least one unequal proportion. Chisquare's advanced Test results are shown in the following table 5 and table 6.

Table 5. Chi-Square Further Test Results Statement 2 on Profession

Profession Characteristics	*Chi-square* Result Test	*Chi-square* Table	Decision	*Cramer's V*
State Civil Apparatus – Student	1.61671	3.84145	Accept H_0	-
State Civil Apparatus – Self Employed	6.35186	3.84145	**Reject H_0**	0.25589
State Civil Apparatus – Army/Police	4.1646E-05	3.84145	Accept H_0	-
Student – Self Employed	2.44055	3.84145	Accept H_0	-
Student – Army/Police	1.32082	3.84145	Accept H_0	-
Self Employed – Army/Police	5.47162	3.84145	**Reject H_0**	0.253716

Based on Table 5, the results of the Chi-Square test show that there is no significant relationship between the profession of Civil Servants and students or TNI/Polri towards statement 2. However, there is a significant relationship between Private Workers and Civil Servants and TNI/Polri, although the relationship is weak (Cramer's V around 0.25). Although there is a relationship between the profession of Private Workers and other professions, the strength of the relationship is very small and not enough to influence the perception of dynasty politics.

Table 6. Chi-Square Further Test Results Statement 2 on Residency Area

Residency Area Characteristics	*Chi-square* Test Result	*Chi-square* Table	Decision	*Cramer's V*
West Indonesia – East	7.24142	3.84145	**Reject H_0**	0.2227
West Indonesia – Central	12.0624	3.84145	**Reject H_0**	0.25604
East Indonesia – Central	0.005276	3.84145	Accept H_0	-

Based on Table 6, the results of the Chi-Square test show that there is a significant relationship between West Indonesia and East Indonesia, and between West Indonesia and Central Indonesia towards statement 2, with Cramer's V values of 0.22 and 0.26 respectively, indicating a weak relationship. However, there is no significant relationship between East Indonesia and Central Indonesia, because the null hypothesis is accepted with a very small Chi-Square value.

c. The Independence Test was conducted on statement 3, which agreed that members of political dynasties obtain political positions because of family ties, not because of their own abilities. The results showed that the level of agreement of female respondents (74%) was slightly higher than that of males (72%). Based on profession, students had the highest level of agreement, which was 78%, followed by civil servants (76%) , TNI/Polri (74%), and Private Workers (63%). In addition, based on the area of residence, respondents from western Indonesia showed the highest level of agreement (77%), followed by central Indonesia (70%), and eastern Indonesia had the lowest level of agreement, which was 61%. *Chi-Square* Test results are obtained from statements 3 to 3 characteristics as follows.

Table 7. Chi-Square Test Result Statement 3

Characteristics	*Chi-square* Test Result	*Chi-square* Table	Decision
Gender	0.064225	3.84145	Accept H_0
Profession	6.69805	11.07049	Accept H_0
Residency Area	3.24931	5.99146	Accept H_0

Based on Table 7, the results of the Chi-Square test show no significant relationship between statement 3 and gender, profession, or area of residence. The Chi-Square value for all characteristics is smaller than the table value, so the null hypothesis is accepted. This means that the agreed perception that members of political dynasties obtain political positions because of family ties, not because of their own abilities, is not influenced by the three demographic characteristics.

d. The Independence Test was conducted on statement 4, which agreed that dynasty politics resulted in policies that were not beneficial to the community, because they only prioritized their own group. The results showed that 75% of men and 70% of women agreed with this statement. Based on profession, respondents who worked as Private Workers had the highest level of agreement (76%), followed by Students (73%), TNI/Polri Members (68%), and Civil Servants (65%). Based on the area of residence, respondents from western Indonesia showed the highest level of agreement (75%), eastern Indonesia (70%), and central Indonesia (65%). *Chi-Square Test results are obtained* from statements 4 to 3 characteristics as follows.

Table 8. Chi-Square Test Result Statement 4

Characteristics	*Chi-square* Test Result	*Chi-square* Table	Decision
Gender	0.78497	3.84145	Accept H_0
Profession	3.35444	11.07049	Accept H_0
Residency Area	1.86777	5.99146	Accept H_0

Based on Table 8, the results of the Chi-Square test show no significant relationship between statement 4 and gender, profession, or area of residence. The Chi-Square value for all characteristics is smaller than the table value, so the null hypothesis is accepted. This means that

the perception that dynastic politics produces policies that are not beneficial to the community is not influenced by the three demographic characteristics.

CONCLUSION

Based on research conducted on 210 respondents with questionnaire filling media, it was concluded that people's perception of dynastic politics was not significantly related to gender. However, significant differences arise when involving certain professions (types of work) and population areas, especially in the second statement which agrees that the existence of dynastic politics has a negative impact on Indonesian democracy. These results provide a deeper picture of the complexity and dynamics of public perceptions of dynastic politics in Indonesia and strengthen understanding of how certain factors can influence their views. This research also has the potential to support the government's efforts in building public trust in political institutions in the electoral process.

ACKNOWLEDGMENTS

The author would like to thank the Statistics Study Program, Faculty of Science and Technology and Universitas Airlangga, as well as all parties involved in supporting this research and publication.

REFERENCE

[1] A. S. Purba, "Potret Pandangan Akademisi di Jurnal Ilmu Sosial dan Ilmu Politik UGM (JSP) Mengenai Permasalahan Demokrasi di Indonesia," *Jurnal Politik Muda*, vol. 4, pp. 1–2, 2015.

[2] A. Dedi, "Politik Dinasti dalam Prespektif Demokrasi," *Moderat: Jurnal Ilmiah Ilmu Pemerintahan*, pp. 92–101, 2022.

[3] A. A. Rahma, A. A. Oktaviani, A. Hofifah, T. Z. Ahda, and R. G. Nugraha, "Pengaruh Dinasti Politik Terhadap Perkembangan Demokrasi Pancasila di Indonesia," *Jurnal Kewarganegaraan*, vol. 6, 2022.

[4] A. D. M. Syanur, I. B. S. AW, and L. Afrilia, "Politik Dinasti Di Indonesia: Tinjauan Kritis Terhadap Penerapan Demokrasi Di Era Kepemimpinan Presiden Jokowi," *Kultura: Jurnal Ilmu Hukum, Sosial, Dan Humaniora*, vol. 1, no. 3, pp. 17–26, 2023.

[5] K. Gadafi and S. Nelwati, "Dampak Dinasti Politik Terhadap Culture Demokrasi Pancasila di Indonesia," *Journal Innovation In Education*, vol. 2, no. 3, pp. 133–148, 2022.

[6] W. Pramesti, "Tabel Kontingensi Untuk Mengetahui Hubungan Antara Jenis Penyakit, Jenis Kelamin, Usia, Lama Rawat dan Keadaan Keluar Pasien," *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, pp. 4–1, 2012, doi: 10.36456/jstat.vol4.no1.a1165.

[7] X. Ji, W. Gu, X. Qian, H. Wei, and C. Zhang, "Combined Neyman–Pearson chi-square: An improved approximation to the Poisson-likelihood chi-square," *Nucl Instrum Methods Phys Res A*, vol. 961, 2020, doi: 10.1016/j.nima.2020.163677.

[8] S. D. Bolboacă, L. Jäntschi, A. F. Sestraş, R. E. Sestraş, and D. C. Pamfil, "Pearson-fisher chi-square statistic revisited," *Information (Switzerland)*, vol. 2, no. 3, pp. 528–545, 2011, doi: 10.3390/info2030528.

[9] M. A. Kleden, J. U. Moto, and R. D. Guntur, "Hubungan Faktor Demografis dengan Kejadian Malaria di Kecamatan Wewewa Timur: Pendekatan Analisis Chi-Square," *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 16, no. 2, pp. 499–513, 2023.

[10] L. P. Widayanti and E. Kusumawati, "Hubungan persepsi tentang efektifitas vaksin dengan sikap kesediaan mengikuti vaksinasi Covid-19," *Jurnal Hearty: Jurnal Kesehatan Masyarakat*, vol. 9, no. 2, pp. 78–85, 2021.

[11] S. Nurhalimah, "Persepsi Masyarakat terhadap Kepemimpinan Perempuan Kepala Desa (Studi di Desa Ciasihan Kecamatan Pamijahan Kabupaten Bogor)," 2022.

[12] United Nations, "Sustainable Development Goal 16: Peace, Justice, and Strong Institutions." Accessed: Dec. 14, 2024. [Online]. Available: https://indonesia.un.org/id/sdgs/16/progress

[13] J. Nugraha, *Pengantar Analisis Data Kategorik : Metode dan Aplikasi menggunakan Program R*. Yogyakarta: Group Penerbitan CV BUDI UTAMA, 2013.

[14] B. S. Everitt and Edward Arnold, *The Analysis of Contingency Tables (2nd Edition)*. London: Edward Arnold, 1993.

[15] I. Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: CV. Alfabeta, 2019.

[16] H. Retnawati, "Reliabilitas Instrumen Penelitian," *Jurnal Pendidikan Teknik Mesin Unnes*, vol. 12, no. 1, 2017.

Three–Phase Traffic Light Petri Net Model Using The Modified Norwegian System

Tomi Tristono[(1)], Setiyo Daru Cahyono[(2)], Pradityo Utomo[(3)], Mochamad Sidqon[(4)], Sudarno[(5)], Daniel Wahyu Suprayoga Prabowo[(6)], Hendro Susilo[(7)]

[1,2,3,5,6,7]University of Merdeka Madiun – Serayu Road 79, Madiun City

[4]University 17 August 1945 Surabaya – Semolowaru Road 45, Surabaya City

e-mail correspondence: tomitristono@unmer-madiun.ac.id[(1)]

ABSTRAK

Petri net dapat dipakai untuk memodelkan perilaku sinyal lampu lalu lintas. Model Petri net juga memungkinkan untuk menyajikan singkronisasi beberapa fase lampu lalu lintas. Model Petri net dapat mereprentasikan sinyal modifikasi sistem Norwegia yang menyala dengan urutan hijau, kuning, merah, kuning, dan kemudian hijau lagi. Pada setiap siklus lampu lalu lintas modifikasi sistem Norwegia, sinyal kuning menyala dua kali. Studi ini bertujuan untuk mengkaji model Petri net lampu lalu lintas dengan tiga fase yang menggunakan modifikasi sistem Norwegia. Metode uji validasi dan verifikasi kebenaran model Petri net diantaranya menggunakan sejumlah *Place-Invariant*, properti keterbatasan (*boundedness*) pada Petri net, konservasi (*conservation*), *coverability tree* untuk berbagai keadaan sinyal, dan simulasi. Berdasarkan hasil kajian menyatakan bahwa *Place-Invariant*, dan properti Petri net dapat merepresentasikan bahwa model adalah benar dan *feasible*. Simulasinya juga menyajikan urutan yang benar sinyal modifikasi sistem Norwegia.

Kata kunci: model Petri net, lampu lalu lintas, modifikasi sistem Norwegia

ABSTRACT

Petri nets can be used to model the behavior of traffic light signals. The Petri net model also makes it possible to provide synchronization of several traffic light phases. The Petri net model can represent the modified Norwegian system signal, which lights up in the sequence green, yellow, red, yellow, and then goes back to green again. The yellow signal flashes twice in each traffic light cycle modified Norwegian system. This study aims to examine the Petri net model of traffic lights with three phases using the modified Norwegian system. Methods for validating and verifying the correctness of the Petri net model used several Place-Invariants, boundedness properties on the Petri net, conservation, coverability trees for various signal conditions, and simulation. Based on the study results, the Place-Invariant and the Petri net properties can represent that the model is correct and feasible. The simulation also presents the correct sequence of modified Norwegian system signals.

Keywords: *Petri net model, traffic lights, modified Norwegian system*

INTRODUCTION

This study aims to study the Petri net model of the modified Norwegian traffic light system. Traffic lights have three phases and a fixed order: west, north, and east. They then return to the initial state, the west phase.

Generally, a standard traffic light system has three discrete state sequences: green, yellow, and red. All states light up in specific and at permanent time intervals, alternating with other phases and repeatedly forming a traffic light cycle [1]. The Norwegian traffic light system is slightly different from the standard system. On the Norwegian system, the yellow and red signals flash simultaneously before returning to the green signal. These two signals aim to reduce travel delays and not to trigger conflict. If the signal lights up red–yellow, vehicles stopped for a long time can prepare to start their travel again. The Modified Norwegian system changes the red–yellow signal to a single-yellow signal. In each traffic light cycle of the modified Norwegian system, the yellow signal flashes twice [2].

Petri nets are models that can graphically represent the behavioral structure of a distributed system. It can also model control systems, sensor networks, and manufacturing [3]. Petri nets can also represent traffic light scheduling [2]. Petri net is a directed graph with four essential elements: places (P/roundabouts), transitions (T/squares), directed arcs, and tokens/dots. Place (P) is a representation of the state that occurred. The enabled transition (T) is ready to fire. It can trigger and transform an initial state into the next state according to the arc direction. A token indicates that a condition related to a particular place is occurring [4].

Several Petri net properties, including boundedness, conservation, and state tree coverability, are used for validation tests and model verification. Model validation and verification tests also use Place-Invariant and simulation. The analysis uses the coverability tree method, a sequence of state occurred, and fire transitions, which can be represented using the multiplication of the connection/ incidence matrix and the enable transition matrix. [3].

Previous research has studied the structure of traffic light behavior using Petri nets. It is the study of traffic light structure modifications using Petri nets, which can reduce delays but do not trigger conflicts [2]. The study of invariants on Petri nets was conducted to prove that a traffic light model is correct [4]. Research on modified binary Petri nets (MBPNs) to design traffic light models [5]. Research on the validity of traffic light models and analysis using Petri net properties [6]. Business process system studies must be deadlock-free. This study was analyzed using the Petri net model [7]. It is the research on the Petri net model of a network system with multichannel queues [8].

METHOD

The following describes the definition of a Petri net, the three-phase system using the modified Norwegian system, and the connectivity/incidence matrix and invariants.

1. Petri Net

Four Petri net elements must be known, including N(P, T, A, w). The elements of P are a finite set of places. P = {P_1, P_2, . . . , Pm}. The finite set of transitions T is T = {t_1, t_2, . . . , tn}, which m and n are positive integers. Element A is a set of directed arcs that relate place to transition and the reverse, namely from transition to place. It is written as A ⊆ (P×T) ∪ (T×P). The element

w (weight) is the weight function of an arc with w, i.e., A→{1, 2, 3, . . }. The model's arc weight is written near the arc line using an integer number that is greater than one. It is presented with a directed line only if it equals one [3].

The concept of 'marking' in a Petri net model refers to the use of tokens in a place, which is mathematically represented as M(Pi) ={0, 1}, where i=1,2,3,..15. There are 15 places in a traffic light model with three phases of the modified Norwegian system. A 'marking place' in a traffic light signal model represents the states marked with tokens. For a signal to be on, its corresponding place will have a token, which will be empty if the signal is off. The symbol '1' represents 'on ', while '0' represents 'off '.

2. Three–Phase Traffic Light Using Modified Norwegian System

Figure 1., shows a road junction with a three-phase traffic light schedule. Phase scenario is a traffic light scheduling design that aims to ensure that the space at a road intersection can be used alternately by a group of vehicles coming from different directions [9].

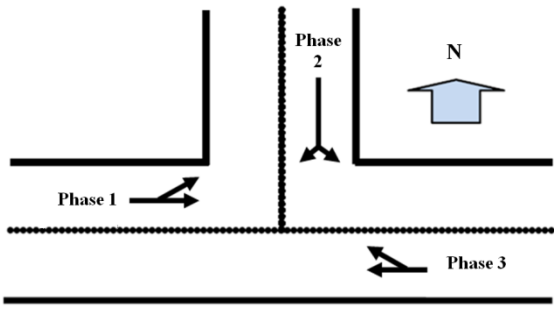


Figure 1. Road junction with three arms and its traffic flow

Figure 2. presents a three-phase traffic light Petri net model with a modified Norwegian system. The west arm features signals P_1, P_2, and P_3, representing green, yellow, and red signals. The north arm has P_6, P_7, and P_8, and the east has P_{11}, P_{12}, and P_{13} signals. Control places P_4, P_9, and P_{14} are crucial in the system. They are responsible for changing the yellow signal to red. Similarly, control places P_5, P_{10}, and P_{15} facilitate transforming the red signal to yellow. These three places also function to synchronize the three traffic light phases.

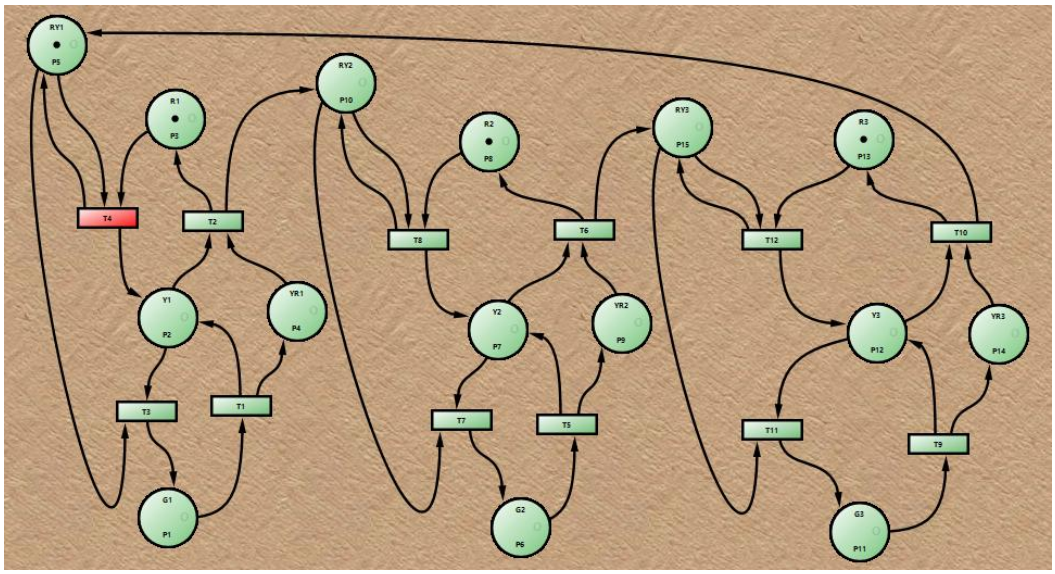


Figure 2. Petri net model of three-phase traffic lights with a modified Norwegian system when the All Red Signal is on.

As depicted in Figure 2., a token on one of P_5, P_{10}, or P_{15} triggers the activation of All Red signals, causing all traffic light signals for the three arms to turn red. There is always an All Red Signal at the end of each traffic light phase of an arm. The aim is to provide a moment's pause to clear the junction of other vehicles that have not yet finished passing, thereby ensuring travel safety and preventing traffic flow conflicts [2]. The synchronous three-phase traffic lights are preceded by the T_2, T_6, or T_{10} fire transition and then progress with T_8, T_{12}, or T_4, respectively.

The green signal for the west arm is on if there is a token on P1. The signals on the north and east arms must be red. There should be a token each in places P_8 and P_{13}. It is illustrated in Figure 2. The signal will light up yellow after the red light elapses and then transform to a green signal in phase 1, namely phase west.

3. The Connectivity/ Incidence Matrix and Invariants

Connections in the Petri net model can be represented using a Forward Incidence Matrix and a Backward Incidence Matrix. The connectivity matrix of the Petri net model is presented in Figure 3. The connectivity matrix A subtracts the Backward Incidence Matrix from the Forward Incidence Matrix. The matrix size is 15 x 12, meaning the model has 15 places and 12 transitions.

The value of the Forward Incidence Matrix elements is the weight of the arcs that connect the transition to place, or place is the output of the transition. The Backward Incidence Matrix elements are the weight of the arcs that connect the place to the transition or mean that the places are the transitions's input. If there is no arc connecting the place to the transition or vice versa, then the arc weight is given a value of zero [3].

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12		
A =	-1	0	1	0	0	0	0	0	0	0	0	0	G1	= P1
	1	-1	-1	1	0	0	0	0	0	0	0	0	Y1	= P2
	0	1	0	-1	0	0	0	0	0	0	0	0	R1	= P3
	1	-1	0	0	0	0	0	0	0	0	0	0	YR1	= P4
	0	0	-1	0	0	0	0	0	0	1	0	0	RY1	= P5
	0	0	0	0	-1	0	1	0	0	0	0	0	G2	= P6
	0	0	0	0	1	-1	-1	1	0	0	0	0	Y2	= P7
	0	0	0	0	0	1	0	-1	0	0	0	0	R2	= P8
	0	0	0	0	1	-1	0	0	0	0	0	0	YR2	= P9
	0	1	0	0	0	0	-1	0	0	0	0	0	RY2	= P10
	0	0	0	0	0	0	0	0	-1	0	1	0	G3	= P11
	0	0	0	0	0	0	0	0	1	-1	-1	1	Y3	= P12
	0	0	0	0	0	0	0	0	0	1	0	-1	R3	= P13
	0	0	0	0	0	0	0	0	1	-1	0	0	YR3	= P14
	0	0	0	0	0	1	0	0	0	0	-1	0	RY3	= P15

Figure 3. Connectivity/ incidence matrix of the Petri net model

Invariant, or Place-Invariant, is a term that describes a Linear Time-Invariant (LTI) system in dynamic marking places. Traffic lights are a Linear Time-Invariant (LTI) system. The system has dynamic behavior, but the marking uses tokens; it has permanent elements unaffected by time [10]. Invariants are a property of Petri net used to validate and verify the correctness of a model. While this study does not discuss invariants in transitions, they are also a part of this concept. The manifestations of Invariants or Place-Invariants are written in Invariant (1) to Invariant (3). Extra, Invariant (4), and Invariant (5) are used to present a conservation number of Petri net models.

Traffic lights regulate travel schedules at intersections/ junctions to avoid conflicts in the flow of vehicles coming from different directions. It must serve all signal phases of the road arm and can return to the initial state to create cycles and repetition of traffic light schedules [3].

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
T =	1	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0
	0	0	0	1	0	0	0	0	0	0	0	0
	0	0	0	0	1	0	0	0	0	0	0	0
	0	0	0	0	0	1	0	0	0	0	0	0
	0	0	0	0	0	0	1	0	0	0	0	0
	0	0	0	0	0	0	0	1	0	0	0	0
	0	0	0	0	0	0	0	0	1	0	0	0
	0	0	0	0	0	0	0	0	0	1	0	0
	0	0	0	0	0	0	0	0	0	0	1	0
	0	0	0	0	0	0	0	0	0	0	0	1

Figure 4. A set of transition matrices in a modified Norwegian traffic light system for junctions with three arms.

In Figure 4, the transition matrix is presented in order from T_1 to T_{12}. However, Figure 4 does not reflect the fire sequence in the traffic light Petri net model with a modified Norwegian system.

RESULT AND DISCUSSION

Figure 5 shows the sequence of events in the Petri net model of three-phase traffic lights, which was designed to implement the modified Norwegian system. Twelve events are combined in a row into a matrix O (Occurrence), which is 15 x 12.

	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10	O11	O12	
O =	1	0	0	0	0	0	0	0	0	0	0	0	P1
	0	1	0	0	0	0	0	0	0	0	0	1	P2
	0	0	1	1	1	1	1	1	1	1	1	0	P3
	0	1	0	0	0	0	0	0	0	0	0	0	P4
	0	0	0	0	0	0	0	0	0	0	1	1	P5
	0	0	0	0	1	0	0	0	0	0	0	0	P6
	0	0	0	1	0	1	0	0	0	0	0	0	P7
	1	1	1	0	0	0	1	1	1	1	1	1	P8
	0	0	0	0	0	1	0	0	0	0	0	0	P9
	0	0	1	1	0	0	0	0	0	0	0	0	P10
	0	0	0	0	0	0	0	0	1	0	0	0	P11
	0	0	0	0	0	0	0	1	0	1	0	0	P12
	1	1	1	1	1	1	1	0	0	0	1	1	P13
	0	0	0	0	0	0	0	0	0	1	0	0	P14
	0	0	0	0	0	0	1	1	0	0	0	0	P15

Figure 5. The Occurrence Matrix presents all states/ events in the modified Norwegian traffic light system.

Equations (1) to Equation (7) are determined based on the connection matrix/ Incidence Matrix A shown in Figure 3. It is the formula after fire with the enabled transition. A transition is enabled if all input places contain tokens greater than the arc weight to the next destination place [3].

$$O_{i+1}=O_i+A.T_i, \text{ which } i=1, 2, 5, 6, 9,10 \quad (1)$$

$$O_4=O_3+A.T_8 \quad (2)$$

$$O_5=O_4+A.T_7 \quad (3)$$

$$O_8=O_7+A.T_{12} \quad (4)$$

$$O_9=O_8+A.T_{11} \quad (5)$$

$$O_{12}=O_{11}+A.T_4 \quad (6)$$

$$O_1=O_{12}+A.T_3 \quad (7)$$

Figure 6 shows the sequence of fire transitions and Occurrences. One traffic light cycle starts from [O1] → T1, with the event [O1] and the transition T1 enabled. All transitions in the model fire once in one traffic light cycle without missing anything, and events can return to their initial state, which indicates coverability.

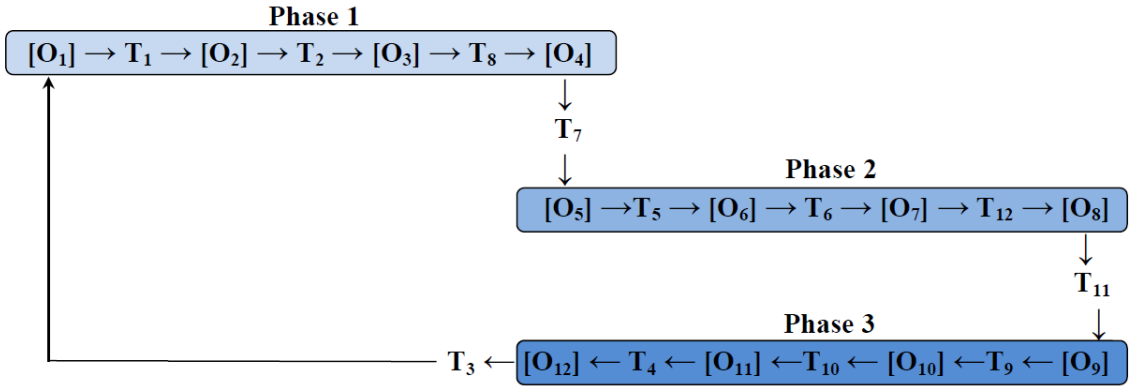


Figure 6. The sequence of all occurrences and fire transitions.

The following are three permanent invariants with parts that do not change over time. Invariant (1) states that only one signal can light at a phase in one junction arm: green, yellow, or red. The *Marking*: $M(G_i), M(Y_i), M(R_i), M(G_j), M(Y_j), M(R_j), M(YR_i), M(RY_i) = \{0,1\}$.

Invariant (2) states that a traffic light signal on a phase may turn on a green or yellow signal if the signals of the other two arms are light red. Invariant (2) guarantees travel safety for vehicles coming from an arm so there is no conflict with other traffic from the different road arms.

$$M(G_i) + M(Y_i) + M(R_i) = 1, \ for\ i,j = 1,2,3 \qquad \text{Invariant (1)}$$

$$M(G_i) + M(Y_i) + M(R_i) = M(R_j) \ \ for\ i \neq j, and\ M(R_j) = 1 \qquad \text{Invariant (2)}$$

$$M(G_1) + M(YR_1) + M(RY_1) + M(G_2) + M(YR_2) + M(RY_2) + M(G_3) + M(YR_3) + M(RY_3) = 1 \qquad \text{Invariant (3)}$$

Invariant (3) describes a synchronized travel schedule for the three road arms of the junction. *M(YRi)* is the marking symbol on the control place for the yellow signal, which will transform into a red signal for i = 1, 2, 3. *M(RYi)* is the marking on the synchronized control place for the three arms and the control place for the red signal, which will transform into a yellow signal at a traffic light phase.

Figure 7. Three-phase traffic light simulation results with a modified Norwegian system

After the verification, the modified Norwegian traffic light system model with three phases was declared valid. The fire sequence of transitions in the traffic light Petri net model, which has three phases, constructs a traffic light cycle. The model has a qualified coverability tree that can return to its initial state.

Table 1. The Traffic Light Schedule

Phase	Green	InterGreen			Red	Cycle
		Yellow I	Yellow II	All Red		
	Seconds					
1. West	27	3	3	3	66	99
2. North	18	3	3	3	75	99
3. East	27	3	3	3	66	99

Figure 7 shows the simulation. Table 1 presents the complete schedule. Both present the correct sequence and time interval for each signal of the modified Norwegian system. The traffic light Petri net model meets the boundedness property; the number of tokens in a place does not grow to infinity or is bounded. The model also fulfills the conservation property. There are no deadlocks or situations when no transitions can fire at all.

$$2M(P_1) + M(P_2) + M(P_3) + M(P_5) + 2M(P_6) \\ +M(P_7) + M(P_8) + M(P_9) + M(P_{10}) + 2M(P_{11}) \\ +M(P_{12}) + M(P_{13}) + M(P_{14}) + M(P_{15}) = 4 \qquad \text{Invariant (4)}$$

Petri net is conservation if the number of tokens in all places after being multiplied by their respective weights is constant for all situations [3]. Multiplication of the matrix B(1x15) on the Occurrence matrix O in Figure 5., will always give a constant value. Multiplication B(1x15) x O(15 x 12) = C(1x12). The row matrix B = [2 1 1 1 2 1 1 1 2 1 1 1 1 1 1] and the row matrix C = [4 4 4 4 4 4 4 4 4 4 4 4]. There are 12 occurrence states in one modified Norwegian traffic light cycle. The conservation constant value is 4. This conservation property can be written in Invariant (4). $M(P_i) = \{0,1\},\ which\ i = 1, 2, \ldots 15$.

$$2M(G_1) + M(K_1) + M(R_1) + M(YR_1) + M(RY_1) + 2M(G_2) \\ +M(Y_2) + M(R_2) + M(YR_2) + M(RY_2) + 2M(G_3) \\ +M(Y_3) + M(R_3) + M(YR_3) + M(RY_3) = 4 \qquad \text{Invariant (5)}$$

Invariant (4), which applies to all places in the model, can also be written as Invariant (5). The marking applies: $M(G_i), M(Y_i), M(R_i), M(G_j), M(Y_j), M(R_j), M(YR_i), M(RY_i) = \{0,1\}$. Solving conservation problems is sufficient if the coverability tree is met.

Three-phase traffic lights can be implemented at all road junctions. These lights, designed for safety, are particularly effective at most Trans Java Toll exits – in Indonesia, which are junctions with three arms. When the schedule is appropriate, these lights ensure safe and smooth traffic flow without causing jams.

CONCLUSION

A study has reviewed the traffic light model with three phases of the modified Norwegian system using Petri nets, along with model correctness verification tests using Invariants and several Petri net properties. Based on the study results, the model was declared valid and feasible. The Petri net property can show that the model can return to the initial. Analysis of the traffic light Petri net model has fulfilled the boundedness, conservation, and tree coverability properties for all

conditions. The simulation has presented the correct sequence of each signal in the modified Norwegian system.

ACKNOWLEDGMENTS

Thanks for financing research support in the 2024 budget year is conveyed to LPPM Merdeka Madiun University.

REFERENCE

[1] *MKJI*. Jakarta: Direktorat Jenderal Bina Marga Direktorat Bina Jalan Kota, Jakarta, 1997.
[2] S. D. Cahyono, T. Tristono, Sutomo, and P. Utomo, "Model of demand order method of traffic lights phases," *J. Phys. Conf. Ser.*, vol. 1211, no. 1, pp. 0–9, 2019, doi: 10.1088/1742-6596/1211/1/012036.
[3] D. Adzkiya, "Membangun Model Petri Net Lampu lalu Lintas dan Simulasinya," ITS Surabaya, 2008.
[4] T. Tristono, S. D. Cahyono, and P. Utomo, "' Place-Invariant' Pada Model Petri Net Lampu Lalu Lintas," in *Prosiding Sendika, UMPwr*, 2021, pp. 173–178.
[5] O. Yaqub and L. Li, "Modeling and analysis of connected traffic intersections based on modified binary petri nets," *Int. J. Veh. Technol.*, vol. 2013, 2013, doi: 10.1155/2013/192516.
[6] M. D. S. Soares, "Architecture-Driven Integration of Modeling Languages for the Design of Software-Intensive Systems.," TU Delft, 2010.
[7] R. Anggrainingsih, S. P. Yohanes, and U. Salamah, "Analisis Dan Verifikasi Workflow Mengggunakan Petri Net," in *UDINUS, Semarang, Semantik*, 2014, pp. 150–156.
[8] S. R. Puri and W. Pramesthi, "Model Petri Net Sistem Jaringan Antrean Multichannel Tak-Siklik 5 Server," *J. Transform. - Univ. PGRI Banyuwangi*, vol. 2, no. 2, pp. 40–50, 2018.
[9] M. Z. Irawan, "Analisis Dampak Lalu Lintas (Simpang Bersinyal)." UGM, Yogyakarta, 2019.
[10] N. D. Safitri, "Karakteristik Domain Waktu dan Frekuensi pada Sistem Linear Time Invariant (LTI)," UIN Sunan Kalijaga - Yogyakarta, 2019.

Weibull Regression Survival Analysis on the Rate of Recovery of Thyphoid Fever Patients: Case Study of RSUD Haji Makassar

Nurul Maghfira [(1)], Bobby Poerwanto [(2)], Hardianti Hafid [(3)]

[1,2,3] Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,

Universitas Negeri Makassar

Kampus UNM Parangtambung, Jalan Daeng Tata Makassar

e-mail: nurulmaghfira90@gmail.com[(1)], bobby_poerwanto@unm.ac.id[(2)], hardiantihf@unm.ac.id[(3)]

ABSTRAK

Analisis *survival* adalah metode statistika yang digunakan dalam mempelajari ketahanan hidup yang berhubungan dengan waktu, mulai waktu awal penelitian sampai waktu akhir penelitian. Tujuan analisis *survival* adalah untuk mengetahui hubungan antara waktu *survival* dengan variabel bebas pada penelitian yang diduga mempengaruhi waktu *survival*-nya. Pada penelitian ini menggunakan analisis *survival* regresi weibull untuk melihat faktor-faktor yang berpengaruh secara signifikan terhadap laju kesembuhan pasien penderita demam tifoid di RSUD Haji Makassar Tahun 2022. weibull untuk melihat faktor-faktor yang berpengaruh secara signifikan terhadap laju kesembuhan pasien penderita demam tifoid di RSUD Haji Makassar Tahun 2022. Diperoleh hasil analisis faktor yang berpengaruh signifikan terhadap laju kesembuhan pasien penderita demam tifoid adalah nyeri ulu hati. Pasien dengan riwayat mengalami nyeri ulu hati memiliki nilai *hazard ratio* sebesar 1,779, yang berarti bahwa pasien penderita demam tifoid yang mengalami nyeri ulu hati memiliki laju kesembuhan lebih cepat sebesar 1,779 kali dibandingkan dengan pasien yang tidak mengalami nyeri ulu hati.

Kata kunci: Anslisis Survival, Regresi Weibull, Demam Tifoid

ABSTRACT

Survival analysis is a statistical method used in studying survival related to time, from the beginning of the study to the end of the study. The purpose of survival analysis is to determine the relationship between survival time and independent variables in research that are thought to affect survival time. This study uses weibull regression survival analysis to see the factors that significantly affect the recovery rate of patients with typhoid fever at the Makassar Hajj Hospital in 2022. The results of the analysis of factors that have a significant effect on the rate of recovery of patients with typhoid fever are heartburn. Patients with a history of having heartburn have a hazard ratio value of 1.779, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of 1.779 times compared to patients who do not experience heartburn.

Keywords: *Survival Analysis, Weibull Regerssion, Thyphoid Fever*

INTRODUCTION

Survival analysis is one of the statistical methods that is often used to analyze data related to the timing of events [1]. Survival time is defined as the survival time of an object at the beginning of observation until the occurrence of an event [2] Events in survival analysis can be death, disease relapse, treatment or others [3].

The purpose of survival analysis is to determine the relationship between the time of the event and the independent variables measured at the time of the study. In addition, it is also used to see the factors that significantly affect an event or event [4].

One of the modeling of survival analysis is a parametric survival model, which is a model that has the assumption of following a certain known distribution on its dependent variable. In some circumstances, parametric models can provide more accurate estimates ([5]. Distributions that are often used in survival analysis include the Weibull, exponential, log normal, and log-logistic distributions.

The Weibull distribution is a generalization of the exponential distribution, it is known that in the exponential distribution the hazard value is constant, so it often does not match the actual situation. However, in the Weibull distribution, the hazard value is not constant so that it is closer to the actual situation. [6]. The Weibull distribution is flexible due to the shape parameter that determines the change in shape of the hazard curve [7].

Typhoid fever is an acute infectious disease of the digestive system caused by salmonella typhi or salmonella paratyphoid bacteria. Typhoid fever is transmitted through food or drink contaminated with salmonella typhi bacteria, in addition the disease can be transmitted through direct contact with feces, urine or secretions of people with typhoid fever. [8].

Based on data from the World Health Organization (WHO) an estimated 11-20 million people fall ill with typhoid and about 128,000 to 161,000 people die from it each year. The disease is mostly found in areas that lack clean water and environmental sanitation. It can also be caused by unhealthy food [9].

Typhoid fever is one of the endemic diseases in Indonesia that must be given serious attention because it can be a public health threat. In Indonesia alone, typhoid fever cases range from 350-810 per 100,000 population, the prevalence of this disease in Indonesia is 1.6% and ranks 5th in infectious diseases that occur at all ages in Indonesia, which is 6.0% and ranks 15th in the cause of death of all ages in Indonesia, which is 1.6%. [10], So it is necessary to reduce the mortality rate and one of the things that can be done is to analyze the factors that affect the recovery rate of patients with typhoid fever.

METHOD

1. Data Source

The data used in this study are secondary data and identification of variables set as criteria for typhoid fever patients obtained from the medical records of the RSUD Haji Makassar in 2022.

2. Variable Operational Definition

Tabel 1. Variable Operational Definition

Variable	Name Variable	Descriptive
Y	Survival Time	Time during which typhoid fever patients were hospitalized (days).
S	Status	0 = If the patient was involuntarily discharged, or changed treatment, died during the study period;1 = If the patient recovers
X_1	Gender	0 = Male; 1 = Female
X_2	Age	-
X_3	Heartburn	0 = Not heartburn; 1 = Heartburn
X_4	Dirty Tongue	0 = Not dirty tongue; 1 = Dirty tongue
X_5	Vomit	0 = Not Vomit; 1 = Vomit
X_6	Diarrhea	0 = Not diarrhea; 1 = Diarrhea
X_7	Dizzy	0 = Not Dizzy; 1 = Dizzy

3. Research Produce

The procedures carried out in this study are:

1. Collecting sources of information needed in the research.
2. Recapitulation of typhoid fever patient data at this stage the data is taken from the Makassar Hajj Hospital to be processed using the Weibull regression method.
3. Performing data processing using statistical software.
4. Compile a research report.
5. Make conclusions based on the problems that have been discussed

4. Data Analysis Technique

The data analysis techniques in this study are:

1. Describe the characteristics of data in patients with typhoid fever.
2. Testing the assumption of proportional hazard

There are three types of checking the Proportional Hazard assumption[11]:

1. The survival lines on the Kaplan-Meier curve do not intersect each other.
2. Survival lines on the ln-ln survival do not intersect each other.
3. Global test or Goodness of Fit (GOF) test
4. Weibull distribution test on dependent variable [12]

$$D = sup|S(t) - F_0(t)|$$

where:

D = kolmogorov Smirnov

$S(t)$: empirical value of sample cumulative distribution

$F_0(t)$: cumulative distribution function

sup: Supremum or upper limit

5. Weibull regression modeling [13]:

$$h(t|X) = \lambda\gamma t^{-1}\exp(\beta^t X)$$

where:
$h(t|X)$ = Cox model
β^t = regression coefficient
λ = scale parameter value
γ = shape parameter value
t = time variable

6. Parameter estimation with the method Maximum Likelihood Estimation [13], [14]:

$$L(\theta|t) = \frac{n!}{(n-r)!}S(t|X)^{n-r}\prod_{i=1}^{r} f(t|X)$$

Where:
$L(\theta|t) =$ *likelihood* θ function
X_i = a vector of variables of failed individuals at time t_i
$r =$ the desired number of deaths or failures
n = the number of observed data.

7. Parameter significance testing [15], [16]:

a. Simultaneous Test

$$G^2 = -2\left[\ln L_R - \ln L_f\right]$$

L_R = *partial likelihood* starting model.
L_f = *partial likelihood* final model.

b. Partial Test

$$Z = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$$

$\hat{\beta}_i$ = the parameters of the regression model to be estimated,
$SE(\hat{\beta}_i)$ = *standard error* the parameters of the regression model to be estimated,

c. Calculate the hazard ratio value [17]:
The hazard ratio for categorical data is as follows:

$$\widehat{HR} = \frac{h(t|x=1)}{h(t|x=0)} = \frac{h(t)e^{\hat{\beta}}}{h(t)} = e^{\hat{\beta}}$$

$\widehat{HR}$ = hazard ratio estimation
$e^{\hat{\beta}}$ = exponential coefficient

d. Interpretation of the Weibull regression results obtained
e. Conclusion.

RESULT AND DISCUSSION

1. Descriptive Analysis

Data are presented in the form of histograms, tables and diagrams with a total of 75 patients with typhoid fever at the Makassar Hajj Hospital in 2022. The longest typhoid fever patient was 8 days as many as 1 patient, while the fastest typhoid fever patient was 1 day as many as 4 patients and the most patients were 3 days, namely 19 patients. The percentage of typhoid fever patients

who recovered was 86.67% (65 people). The oldest age of the patient was 69 years, while the youngest age of the patient was 1 year. the average age of the patient was 25 years. Inpatients with typhoid fever were more female than male. More typhoid fever inpatients experienced heartburn, vomiting, dizziness and compared to those who did not experience heartburn, vomiting, and dizziness, and more typhoid fever patients did not experience dirty tongue and diarrhea than those who experienced dirty tongue and diarrhea.

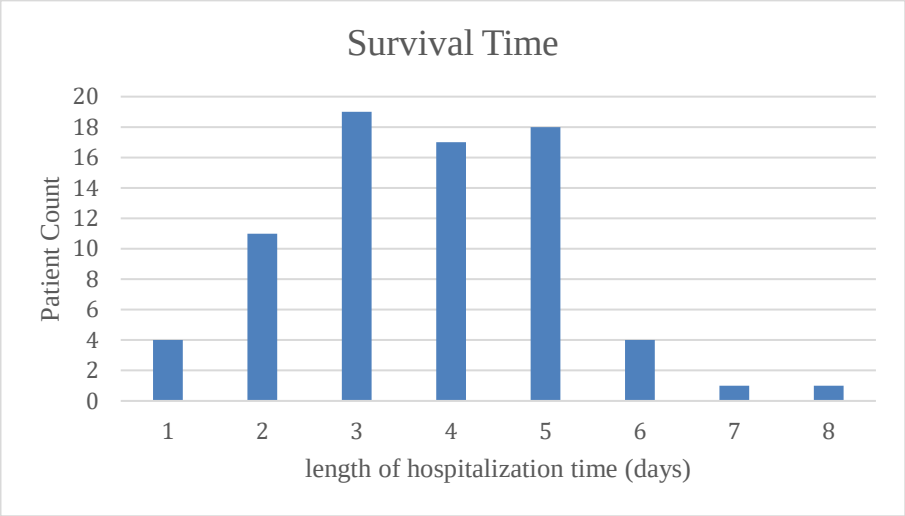
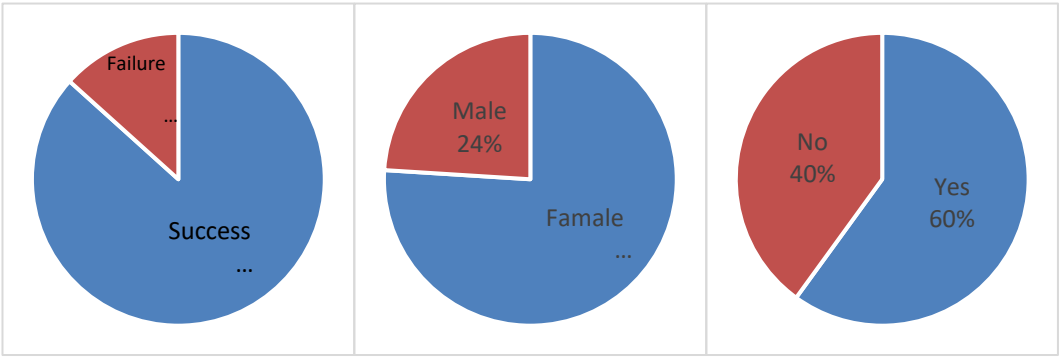


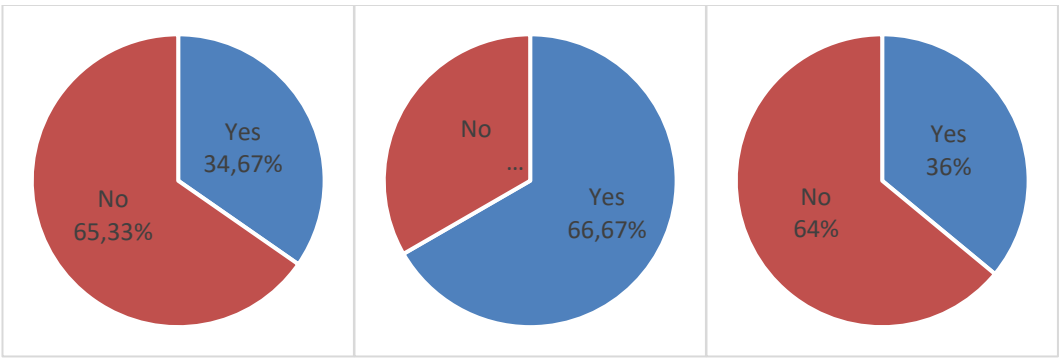
Figure 1. Survival Time Histogram

Table 2. Age of Typhoid Fever Patient

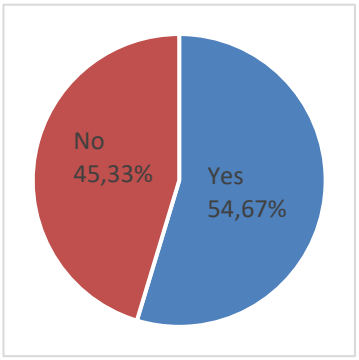
Variable	Average (Year)	Maximum (Year)	Minimum (Year)
Age	25	69	1



(a) (b) (c)



(d) (e) (f)



(g)

Figure 2. Pie Chart for (a) Event Censorship Status, (b) Gender, (c)Heartburn, (d)Dirty Tongue, (e)Vomit, (f)Diarrhea, (g)Dizzy.

2. Testing the Proportional Hazard Assumption

In testing the proportional hazard assumption, this study uses a global test or goodness of fit test. With hypothesis:

H_0 = the data meets the proportional hazard assumption

H_1 = the data does not fulfill the proportional hazard assumption

Reject H_0 if $p - value < \alpha(0{,}05)$

Table 3. Global Test p-value

Variable	p-value	Description
Gender (X_1)	0,52	Accept H_0
Age (X_2)	0,46	Accept H_0
Heartburn (X_3)	0,97	Accept H_0
Dirty Tongue (X_4)	0,89	Accept H_0
Vomit (X_5)	0,65	Accept H_0
Diarrhea (X_6)	0,70	Accept H_0
Dizzy (X_7)	0,96	Accept H_0

Based on table 3, it shows that the results of the global test for each variable each have a p-value> α = 0.05, so it can be concluded that it Accept H_0, which means that the proportional hazard assumption on typhoid fever patient data is met.

3. Testing Data Distribution

Data distribution testing is used to determine the distribution used is in accordance with the length of hospitalization time (Y) data. With Hypothesis:

$H_0 =$ survival data follows the Weibull distribution

$H_1 =$ survival data does not follow the Weibull distribution

Reject H_0 if $D_{hit} > D_{(\alpha,n-1)}$ or $p - value < \alpha(0{,}05)$

Table 4. Distribution Test Values

Distribution	D_{hit}	*p-value*	Description
Weibull	0,132	0,146	Accept H_0

Based on table 4, it shows that the data on the length of hospitalization of patients with typhoid fever obtained a p-value from the Kolmogorov Smirnov test of 0.146 where the p-value = 0.146> α = 0. and the value of $D_{hit} = 0{,}132 < D_{(0{,}05,74)} = 0{,}157$ so that it Accept H_0, which means that the data on the length of hospitalization of patients with typhoid fever follows the Weibull distribution.

4. Parameter Estimation and Parameter Testing

Parameter estimates in the Weibull regression model can be estimated using the Maximum Likelihood Estimation (MLE) method. In parameter testing including simultaneous testing and partial testing, the results of parameter estimation and parameter testing are obtained as follows:

Table 5. Parameter Estimation and Partial Test

Variable	$\|Z\|$	Coefficient	*P-value*	Description
Gender (X_1)	1,860	0,625	0,063	Not significant
Age (X_2)	0,444	-0,004	0,645	Not significant
Heartburn (X_3)	**2**	**0,576**	**0,045**	Significant
Dirty Tongue (X_4)	1,063	-0,300	0,287	Not significant
Vomit (X_5)	0,056	-0,016	0,955	Not significant
Diarrhea (X_6)	0,418	-0,134	0,675	Not significant
Dizzy (X_7)	1,436	-0,405	0,151	Not significant

a. Simultaneous test

The simultaneous test is carried out to see the effect of the independent variables together on the dependent variable in a model. With hypothesis:

$H_0 =$ The independent variable has no effect on the dependent variable

$H_1 =$ At least one of the independent variables affects the variable.

Reject H_0 if p-value $\leq \alpha(0{,}05)$ or $G_{hit} \geq = x^2_{\alpha;db=p}$

Table 6. Simultaneous test for all variables

Likelihood ratio test (G)	$x^2_{(0,05;6)}$	***p-value***
16,4	12,59	0,021

Based on the results of the simultaneous test, the likelihood ratio test value is 16.4 and the p-value = 0.021 < α = 0.05, because the value of $G_{hit} = 16{,}4 > x^2_{(0,05;6)} = 12{,}59$, it can be concluded that rejecting H_0 means that there are one or more variables that have a significant effect or it can be interpreted that the model as a whole can contribute to the recovery rate of patients with typhoid fever.

b. Partial Test

This test is a test conducted to determine the effect of each independent variable on the dependent variable. With hypothesis:

$H_0 =$ Independent variable i has no effect on survival time

$H_1 =$ Independent variable i affects survival time

Reject H_0 if *p-value* $\leq \leq \alpha(0{,}05)$ or $|Z| \geq Z_{\frac{0.05}{2}}$

Based on the results in table 5, the test results for each variable are shown. Based on the H_0 rejection criteria, it can be concluded that only the heartburn variable has a p-value < α = 0.05 and a value of $|Z| = 2 > Z_{\frac{0.05}{2}} = 1{,}96$ it can be concluded that reject H_0, which means that only the heartburn variable has a significant effect on the recovery rate of patients with typhoid fever.

From the parameter estimation results in the table above, a Weibull regression model is obtained which has a significance value of less than α = 0.05 so that a Weibull regression model is obtained as follows:

$$h(t|X) = \lambda\gamma t^{\gamma-1} \exp(\beta^t X)$$

$$h(t|x) = (4{,}190)(2{,}836)t^{2{,}836-1}exp(0{,}576X_3)$$

$$\boldsymbol{h(t|x) = (4,190)(2,836)t^{1{,}836}exp(0{,}576X_3)}$$

5. Hazard Ratio

To find out the survival rate of patients with typhoid fever can be seen by looking for the hazard ratio value of significant variables. The results of the analysis for the hazard ratio value can be seen in the following table:

Table 7. Hazard Ratio values for significant variables

Variable	***Hazard Ratio***
Heartburn (X_3)	1,779

Based on table 7, the hazard ratio value for the heartburn variable is 1.779. This value is more than one, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of 1.779 times compared to patients who do not experience heartburn.

CONCLUSION

Of the several factors that are considered to have a significant effect on the recovery rate of inpatients with typhoid fever at the Makassar Hajj Hospital, there is only one variable that has a significant effect, namely the variable of heartburn. The Weibull regression model for data on typhoid fever patients at the RSUD Haji Makassar in 2022 is:

$$h(t|x) = (4{,}190)(2{,}836)t^{1{,}836}exp(0{,}576X_3)$$

The hazard ratio value of 1.779 was obtained, which means that patients with typhoid fever who experience heartburn have a faster recovery rate of about 1.779 times compared to patients who do not experience heartburn.

For future researchers, the authors suggest using other methods in determining the factors that influence the recovery rate of typhoid fever patients. before designing a study, it is better to find out in advance whether or not the data to be used in the study so that it will be easier when doing a study and can conduct research with different case studies.

REFERENCE

[1] Collet David, *Modelling Survival Data In Medical Research third Edition*. 2013.

[2] R. Y. Fa'rifah and Purhadi, "Analisis Survival Faktor-Faktor yang Mempengaruhi Laju Kesembuhan Pasien Penderita Demam Berdarah Dengue (DBD) di RSU Haji Surabaya dengan Regresi Cox," *Jurnal Sains dan Seni ITS*, vol. 1, no. 1, pp. 271–276, 2012.

[3] M. K. Yusuf, "Analisis Survival Lama Waktu Sembuh Dengan Perawatan Standar Pada Pasien Rawat Inap Penyakit Tifus Dengan Menggunakan Metode Regresi Cox Proportional Hazard," 2018, [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true&db=sph&AN=119374333&site=ehost-live&scope=site%0Ahttps://doi.org/10.1016/j.neuron.2018.07.032%0Ahttp://dx.doi.org/10.1016/j.tics.2017.03.010%0Ahttps://doi.org/10.1016/j.neuron.2018.08.006

[4] A. R. Faisal, M. N. Bustan, and S. Annas, "ANALISIS SURVIVAL DENGAN PEMODELAN REGRESI COX PROPORTIONAL HAZARD MENGGUNAKAN PENDEKATAN BAYESIAN (Studi Kasus: Pasien Rawat Inap Penderita Demam Tifoid di RSUD Haji Makassar)," *VARIANSI: Journal of Statistics and Its application on Teaching and Research*, vol. 2, no. 2, p. 62, 2020, doi: 10.35580/variansiunm14629.

[5] H. P. Zhu, X. Xia, C. H. Yu, A. Adnan, S. F. Liu, and Y. K. Du, "Application of Weibull model for survival of patients with gastric cancer," *BMC Gastroenterol*, vol. 11, pp. 1–6, 2011, doi: 10.1186/1471-230X-11-1.

[6] L. E. Widyawati and R. D. Bekti, "Analisis Faktor yang Mempengaruhi Laju Kesembuhan Pasien Demam Berdarah Dengue Menggunakan Regresi Cox dan Regresi Weibull (Studi

Kasus : Kabupaten Bantul, Yogyakarta)," *Jurnal Statistika Industri dan Komputasi*, vol. 05, no. 2, p. 26, 2020.

[7] A. Solehah, "Analisis Ketahanan Hidup Pasien Kanker Paru Menggunakan Regresi Weibull," 2018.

[8] Y. Levani and A. D. Prastya, "Demam Tifoid: Manifestasi Klinis, Pilihan Terapi Dan Pandangan Dalam Islam," *Al-Iqra Medical Journal : Jurnal Berkala Ilmiah Kedokteran*, vol. 1, no. 2, pp. 10–16, 2020, doi: 10.26618/aimj.v3i1.4038.

[9] Juniah and M. Arianti, "Pentingnya Pengetahuan Kesehatan Demam Typhoid Anak," *Jurnal Pengabdian Masyarakat*, vol. 2, no. 1, pp. 7–14, 2023, doi: 10.59030/jpmbd.v2i1.18.

[10] S. Khairunnisa, E. M. Hidayat, and R. Herardi, "Hubungan Jumlah Leukosit dan Persentase Limfosit terhadap Tingkat Demam pada Pasien Anak dengan Demam Tifoid di RSUD Budhi Asih Tahun 2018 – Oktober 2019," *Seminar Nasional Riset Kedokteran (SENSORIK)*, pp. 60–69, 2020.

[11] M. Sopiyudin Dahlan, *Analisis Survival Dasar-Dasar Teori Dan Aplikasi Program Stata*. Jakarta: Sagung Seto, 2013.

[12] N. H. Hasa, M. N. Bustan, and Aswi, "ANALISIS BAYESIAN SURVIVAL WEIBULL UNTUK MENENTUKAN FAKTOR YANG MEMPENGARUHI LAJU KESEMBUHAN PASIEN RAWAT INAP KANKER SERVIKS DI RSDU KOTA MAKASSAR," *VARIANSI: Journal of Statistics and Its Application on Teaching and Research ISSN 2684-7590 (Online) Vol. 4 No. 1 (2022), 1-8*, vol. 4, no. 1, pp. 1–8, 2022.

[13] I. Sari, "Analisis Regresi Weibull (Kasus: Faktor-faktor yang mempengaruhi laju kesembuhan penderita penyakit demam berdarah dengue di RSAL JALA AMMARI Makassar)," 2017.

[14] Yogi Adam Firdaus, N. Ngatini, and Sekarsari Utami Wijaya, "Pemodelan Regresi Data Panel Harga Beras di Wilayah Indonesia Bagian Barat," *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, vol. 16, no. 2, pp. 486–498, Dec. 2023, doi: 10.36456/jstat.vol16.no2.a8061.

[15] C. H. Imanina, "ANALISIS SURVIVAL TERHADAP PASIEN PENYAKIT GINJAL KRONIS DENGAN MENGGUNAKAN COX REGRESSION (Studi Kasus: Pasien Penyakit Ginjal Kronis di RSUD Arifin Achmad, Pekanbaru, Riau)," no. 21, pp. 1–9, 2018.

[16] G. Anuraga, A. Indrasetianingsih, and M. Athoillah, "Pelatihan pengujian hipotesis statistika dasar dengan software r," *BUDIMAS: Jurnal Pengabdian Masyarakat*, vol. 3, no. 2, pp. 327–334, 2021.

[17] Riangkaryaman, "MODEL SEMIPARAMETRIK COX PH DAN PARAMETRIK (WEIBULL DAN LOGNORMAL) DALAM ANALISIS SURVIVAL (Kasus: Rawat Inap Penderita DBD di RSUD Haji Kota Makassar," 2017.

Analysis of Environmental and Productivity Factors with the Number of Dengue Hemorrhagic Fever and Obesity Cases in Indonesia

Abu Hanifah Al Faruqy [(1)], A'yunin Sofro [(2)].

[1,2] Universitas Negeri Surabaya

Jl. Ketintang, Ketintang, Kec. Gayungan, Kota Surabaya, Jawa Timur
e-mail: ayuninsofro@unesa.ac.id[(2)]

ABSTRAK

Demam Berdarah Dengue (DBD) adalah penyakit yang sering terjadi di daerah tropis. Hal ini terjadi karena vektor pembawa penyakit ini adalah nyamuk Ae. aegypti yang habitatnya berada di lingkungan tropis. Obesitas telah menjadi masalah global bagi seluruh belahan dunia. Pasien dengan gizi obesitas mempunyai respon imunitas yang lebih kuat karena terjadi peningkatan inflamasi di dalam sirkulasi. Sehingga pembuluh darah dapat menjadi lebih lebar dari biasanya. Hal ini memicu kebocoran plasma dan memperparah penyakit DBD sehingga dapat menimbulkan SSD. Kedua penyakit ini memiliki faktor pemicu bermacam-macam seperti lingkungan dan produktivitas. Dalam penelitian ini, akan menggunakan analisis regresi linear multivariat untuk mengetahui faktor mana saja yang signifikan terhadap kedua penyakit tersebut. Analisis regresi linear menggunakan uji wilks lambda secara serentak serta parsial. Berdasarkan hasil penelitian dinyatakan bahwa ada 4 variabel prediktor yang signifikan terhadap tingkat kasus DBD dan obesitas yaitu, Indeks Kualitas Udara (x_1), prduktivitas (x_2), tingkat kemiskinan (x_3) dan jumlah PUSKESMAS (x_4).
Kata kunci: DBD, Obesitas, Analisis Regresi Linear Multivariat

ABSTRACT

Dengue Hemorrhagic Fever (DHF) is a frequently occurring disease in tropical regions. This is due to the disease vector being the Aedes aegypti mosquito, whose habitat is in tropical environments. Obesity has become a global issue worldwide. Patients with obesity have a stronger immune response due to increased inflammation in circulation. As a result, blood vessels may become wider than usual. This triggers plasma leakage and exacerbates DHF, potentially leading to Severe Dengue Syndrome (SDS). Both of these diseases have various triggering factors such as environmental and productivity-related aspects. In this study, multivariate linear regression analysis will be employed to identify which factors are significant for both diseases. The linear regression analysis will use simultaneous and partial Wilks' Lambda tests. Based on the research findings, it is stated that there are four predictor variables significant for the levels of DHF and obesity, namely, Air Quality Index (x_1), productivity (x_2), poverty rate (x_3), and the number of health centers (x_4).
Keywords: *Regression Linear Multivariat; Analysis; DHS; Obesit*

INTRODUCTION

Dengue Hemorrhagic Fever (DHF) is a disease that frequently occurs in tropical areas and become a global concern [1], [2]. This happens because the vector carrying this disease is the Aedes aegypti mosquito, which habitat is in tropical environments. According to the World Health Organization, DHF is estimated to have 50 to 100 million cases of infection, with 22,000 deaths occurring. The most severe cases of DHF occurred in 2019, with America reporting 3.1 million cases and in Asia itself, reports from Bangladesh (101,000), the Philippines (420,000), Vietnam (320,000), and Malaysia (131,000) [3]. In Indonesia, DHF cases increased from 65,602 cases in 2018 to 138,127 cases in 2019. This indicates that the possibility of DHF occurring in Indonesia is still very high [4].

Obesity has become a global issue for the entire world [5]. Data shows that in the United States in 2007, 2.2 percent of men and 35.5 percent of women had a Body Mass Index (BMI) above normal [6]. In Indonesia, according to the Ministry of Health (KEMENKES), obesity is one of the diseases that will be targeted for eradication.

Given this global trend, obesity is often associated with an increased risk of dengue shock syndrome (DSS) or critical phase dengue hemorrhagic fever (DHF) cases. The prevalence of obesity in individuals over 18 years old reaches 12 percent [7]. Patients with obesity have a stronger immune response due to increased inflammation in the circulation. This can cause blood vessels to become wider than usual. This triggers plasma leakage and worsens the condition of DHF, potentially leading to DSS [3].

So, people with obesity are more susceptible to dengue fever than other people. This could be a serious condition because if obesity individuals infected with dengue fever, the obesity patient will increase the severity of dengue fever [8]. With the high number of obesity and dengue fever cases in the whole world, it is important, especially for people with obesity to take preventive steps.

By deeper understanding of the pattern of both diseases is key to control and prevention strategy [2]. To prevent this from happening, preventive measures can be taken by controlling the causes of both dengue fever and obesity. Environmental factors, including the cleanliness of the environment, weather conditions, and pollution influence the spread of dengue hemorrhagic fever (DHF) itself [9], [10]. Unlike DHF, obesity is not a contagious disease but can occur when a person does not maintain a healthy lifestyle, one aspect of which is the frequency of physical activity [11], [12], [13], [14].

Multivariate linear regression is a regression analysis that involves more than one response variable (y) and more than one predictor variable (x) hat are correlated with each other [14], [15]. Therefore, to get to know which causes that possibly to be controlled, this multivariate linear regression analysis can be used as a tool to determine the extent to which environmental factors and productivity affect DHF and obesity. Hence, this study is titled "Analysis of Environmental and Productivity Factors with the Number of Dengue Hemorrhagic Fever and Obesity Cases in Indonesia."

METHOD

The research method used is multivariate regression analysis, with data collection stages from the official websites of the Ministry of Health, the Central Statistics Agency, and the Directorate of Air Pollution Control, Directorate General of Pollution and Environmental Damage

Control. The collected data can then be processed using multivariate regression analysis with the help of R Studio, and the research steps are as follows:

1. Determine the descriptive statistical analysis of the research data.
2. Determine the estimation of multivariate linear regression parameters.
3. Perform simultaneous and partial significance testing of the parameters using the Wilks' Test.
4. Determine the multivariate linear regression model.
5. Determine the simple linear regression model.
6. Conduct the Akaike Information Criterion (AIC) test on the simple and multivariate linear regression models.
7. Draw conclusions and make recommendations.

The response variables used in this study are as follows :

- y_1 : Dengue Hamorrhaguc Fever (DHF) case rate
- y_2 : Obesity prevalence rate

There are 6 predictor variables in this study are as follows:

- x_1 : Air Quality Index (AQI)
- x_2 : Work productivity level
- x_3 : Poverty rate
- x_4 : Number of community health centers
- x_5 : Unemployment rate
- x_6 : Level of heavy activity in the past week

In this study, environmental factors are derived from variables x_1 dan x_4. The chosen variables representing environmental factors because of the dengue fever can be controlled by the air quality and treatments of dengue disease case in helath center. While productivity factors are derived from variables x_2, x_3, x_5, dan x_6. Therefore, these variables represent productivity because obesity is a disease that related to lifestyle, so these variables were chosen.

RESULT AND DISCUSSION

There are 2 response variables consisting of y_1 and y_2 dan 6 predictor variables, namely, x_1, x_2, x_3, x_4, x_5, and x_6 which will be used in this study. Descriptive statistics are an initial statistical analysis to understand basic information in the data [17]. Here is the descriptive statistics for the entire research dataset

Table 1. Descriptive Statistics of Study.

Variable	Minimum	Maximum	Mean (Average)
y_1	2,9	80,90	28,31
y_2	13	42	29,76
x_1	68,06	95,79	89,18
x_2	24,94	400	94,5
x_3	4,6	27,4	10,43
x_4	0,29	7,16	1,59

x_5	2,27	7,53	4,61
x_6	40,66	75,01	53,09

Multivariate Linear Regression Analysis

Multivariate linear regression analysis is a linear regression analysis that involves more than one or more response variables (y) and predictor variables (x) [18]. Subsequently, from the response variables and predictor variables, a multivariate linear regression model with i responses can be formed as follows.

$$\begin{aligned} y_1 &= \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \cdots + \beta_{n1}x_n + \epsilon_1 \\ y_2 &= \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \cdots + \beta_{n2}x_n + \epsilon_2 \\ &\vdots \\ y_i &= \beta_{0i} + \beta_{1i}x_1 + \beta_{2i}x_2 + \cdots + \beta_{ni}x_n + \epsilon_i \end{aligned} \quad (1)$$

In determining regression analysis, there are several steps to follow. Firstly, estimating the parameters.

Estimated Parameter Results

One of the main objectives of regression analysis is to express the predicted value of the response variable through the values of the predictor variables from the regression equation. Before determining the regression equation, the regression coefficients, denoted as β. $\bar{\beta}$ can be obtained from the following equation [19] :

$$\bar{\beta} = (X^T X)^{-1} X^T Y \quad (2)$$

Where,

X : Matrix of predictor variables $n\ x\ (n+1)$

Y : Matrix of response variable n

$\bar{\beta}$: Matrix of regression parameters

Thus, from the costruction of parameter estimation for each variable, the following is obtained :

Table 2. Estimated parameter results each predictor variables regrading the resnponse variables

Parameter	y_1	y_2
x_0	-26,7373	-72,3188
x_1	0,1101	0,9716
x_2	0,0089	0,0284
x_3	3,5636	0,4109
x_4	2,5688	-1,4411
x_5	-0,8525	0,9694
x_6	0,1328	0,1182

Based on the estimated parameter results above, a temporary multivariate regression model for y_1

and y_2 can be formulated as follows:

$$\begin{aligned} y_1 &= -26{,}7373 + 0{,}1101x_1 + 0{,}0089x_2 + 3{,}5636x_3 + \\ &\quad 2{,}5688x_4 - 0{,}8525x_5 + 0{,}1328x_6 \\ y2 &= -72{,}3188 + 0{,}9716x_1 + 0{,}0284x_2 + 0{,}4109x_3 - \\ &\quad 1{,}4411x_4 + 0{,}9694x_5 + 0{,}1182x_6 \end{aligned} \tag{3}$$

Parameter Significance Testing

Significance testing is used to determine whether all parameters have a significant effect or not. The significance level can be observed; if all parameters have a value of zero, then there is no influence from the predictor variables on the response variable. Conversely, if the parameter values are not zero, there is an influence from the predictor variables on the response variable [19]. In multivariate regression analysis, there are several tests: simultaneous test and partial test.

Simultaneous Testing

Simultaneous testing is used to determine whether all parameters collectively have a significant effect on the model. Simultaneous testing is performed using Wilk's Lambda test with the following hypotheses:

$H_0 : \bar{\beta}_{11} = \bar{\beta}_{12} = \ ... = \bar{\beta}_{in} = 0$ (model is not significant)

H_1 : paling sedikit ada satu $\bar{\beta}_{in} \neq 0$ (model significant)

The computed Wilk's Lambda test result is $\lambda_{calculate} = 0.0156$ with a $p - value = 2.22\ x\ 10^{-16}$. Thus, it can be concluded that the $p - value < \alpha = 0.05$, Therefore, the decision is to reject H_0 and the model is significant collectively.

Partial Testing

Using the same hypothesis, partial testing will be conducted for each parameter. The results of partial testing are as follows:

Table 3. Variable analysis with Wilks value and p-value result

Variable	Nilai Wilks	p-value
x_1	0,4912	9,636 x 10^{-5}
x_2	0,7214	0,0143
x_3	0,0272	2,2 x 10^{-16}
x_4	0,7694	0,0331
x_5	0,9203	0,3399
x_6	0,9201	0,3389

Based ont Table 3., the predictor variables $x_1, x_2,\ x_3$ and x_4 have p-values for each variables less than α(0,05), thus it can be concluded that $x_1, x_2,\ x_3$ and x_4 significantly influence the response variables y_1 and y_2. Conversely, the predictor variables x_5 and x_6 have p-values greater than α(0,005) indicating that x_5 and x_6 are not significant for the response variables y_1 and y_2.

Parameter estimation will be performed with significant variables, resulting in the following:

Table 4. Significant estimated parameters

Variable	y_1	y_2
β_0	-30,1027	-52,0933
β_1	0,1951	0,8573
β_2	0,0074	0,0308
β_3	3,5318	0,4865
β_4	2,1728	-1,6222

Multivariate Regression Model

Based on the partial Wilks' test results, there are 4 predictor variables that have a significant influence on the response variable. These predictor variables are $x_1, x_2,\ x_3$ and x_4. Therefore, the multivariate linear regression model can be written as follows:

$$\begin{aligned} y_1 &= -30{,}1027 + 0{,}1951x_1 + 0{,}0074x_2 + 3{,}5318x_3 + 2{,}1728x_4 \\ y_2 &= -52{,}0933 + 0{,}8573x_1 + 0{,}0308x_2 + 0{,}4865x_3 - 1{,}6222x_4 \end{aligned} \quad (4)$$

From this multivariate linear regression model, the interpretations are as follows:

1. For every one percent increase in the Air Quality Index (AQI), the number of DHF cases will increase by 0.1951. For every one percent increase in work productivity level, the number of DHF cases will increase by 0.0074. For every one percent increase in the poverty rate, the number of DHF cases will increase by 3.5318, and for every one percent increase in the number of community health centers, the number of DHF cases will increase by 2.1728.
2. For every one percent increase in the Air Quality Index (AQI), the obesity prevalence rate will increase by 0.8573. For every one percent increase in work productivity level, the obesity prevalence rate will increase by 0.0308. For every one percent increase in the poverty rate, the obesity prevalence rate will increase by 0.4865, and for every one percent increase in the number of community health centers, the obesity prevalence rate will decrease by 1.6222.

Simple Linear Regression Analysis

Linear regression analysis is a method used to establish a functional relationship between two variables. Typically, this analysis involves one response variable and one predictor variable. The regression model can be written as follows [20].

$$y = a + bx. \quad (5)$$

In this study, the regression of the response variable on each predictor variable can be observed as follows:

Table 5. Estimated parameters variables that significant to y_1

Variable	y_1
a	-30,1027
x_1	0,1951

x_2	0,0074
x_3	3,5318
x_4	2,1728

According to simple linear regression result, can constructed regression model as follows:

$$y_1 = -30{,}1027 + 0{,}1951x_1 + 0{,}0074x_2 + 3{,}5318x_3 + 2{,}1728x_4 \quad (6)$$

Table 6. Estimated parameters variables that significant to y_2

Variable	y_2
a	-52,0933
x_1	0,8573
x_2	0,0308
x_3	0,4865
x_4	-1,6222

According to simple linear regression result, can constructed regression model as follows:

$$y_2 = -52{,}0933 + 0{,}8573x_1 + 0{,}0308x_2 + 0{,}4865x_3 - 1{,}6222x_4 \quad (7)$$

Akaike Information Criterion (AIC)

AIC (Akaike Information Criterion) is used to evaluate how well a model fits the data. The AIC value chosen is the smallest among other AIC values [21,22,23]. In this study, the AIC values of the simple linear regression model and the multivariate linear regression model are compared. The AIC values of both models can be seen as follows:

Table 7. Comparison of AIC values between the simple linear regression model and the multivariate linear regression model

Model	Nilai AIC
Regresi Linear Sederhana	219,55
Regresi Linear Multivariat	145,86

According AIC values, the model with the smallest AIC value is considered the best model. Therefore, the model from multivariate linear regression is selected as the chosen regression model in this study.

CONCLUSION

Based on the results of the Wilks' Lambda test, there are 4 predictor variables that significantly influence the response variable, namely x_1 (Air Quality Index), x_2 (Work productivity), x_3 (Poverty rate), and x_4 (Number of community health centers). From these four variables, a linear regression model can be derived. In this study, simple linear regression analysis was also conducted on each predictor variable against each response variable. Based on the results

of both simple linear regression and multivariate linear regression analyses, two different AIC values were obtained. Therefore, the model with the smallest AIC value will be chosen as the best model in this study. Thus, the multivariate linear regression model is selected as the best model, with the following model:

$$\begin{aligned} y_1 &= -30{,}1027 + 0{,}1951x_1 + 0{,}0074x_2 + 3{,}5318x_3 + 2{,}1728x_4 \\ y_2 &= -52{,}0933 + 0{,}8573x_1 + 0{,}0308x_2 + 0{,}4865x_3 - 1{,}6222x_4 \end{aligned} \quad (8)$$

REFERENCE

[1] "World Mosquito Program," Dengue. [Online]. Available: https://www.worldmosquitoprogram.org/en/learn/mosquito-borne-diseases/dengue#

[2] L. P. S. Pratiwi, "Binary Logistic Regression Analysis on the Spread of Dengue Fever in Bali Province," vol. 17, no. 1, 2024.

[3] R. A. A. Naiem, R. Rompies, and S. N. N. Tatura, "Hubungan antara Status Nutrisi dengan Tingkat Keparahan Infeksi Dengue pada Pasien Anak di RSUP Prof. Dr. R. D. Kandou, Manado, Indonesia," *eCL*, vol. 11, no. 1, Oct. 2022, doi: 10.35790/ecl.v11i1.37887.

[4] M. Arisanti and N. H. Suryaningtyas, "KEJADIAN DEMAM BERDARAH DENGUE (DBD) DI INDONESIA TAHUN 2010-2019," *J.SPIRAKEL*, vol. 13, no. 1, pp. 34–41, Dec. 2021, doi: 10.22435/spirakel.v13i1.5439.

[5] S. Mugianti, T. Setijaningsih, and K. Fransiska, "Perkembangan Motorik Balita Obesitas Usia 3-5 Tahun," *JNK JOURNAL*, vol. 5, no. 1, pp. 046–052, Apr. 2018, doi: 10.26699/jnk.v5i1.ART.p046-052.

[6] Ni Putu Indah Kartika Putri and I Made Gede Dwi Lingga Utama, "HUBUNGAN OBESITAS DENGAN KEJADIAN SINDROM SYOK DENGUE PADA ANAK," *JURNAL MEDIKA UDAYA*, 10 2020.

[7] KEMENKES RI, *Laporan Kinerja Kementrian Kesehatan Tahun 2022*. 2022.

[8] Naseem S. Miller, "Dengue: A primer and research roundup," Aug. 2024, [Online]. Available: https://journalistsresource.org/home/dengue-research-roundup/

[9] L. L. Ayun and E. T. Pawenang, "Hubungan antara Faktor Lingkungan Fisik dan Perilaku dengan Kejadian Demam Berdarah Dengue (DBD) Di Wilayah Kerja Puskesmas Sekaran, Kecamatan Gunungpati, Kota Semarang," 2017.

[10] H. Husin, A. Ramon, W. Angraini, N. Wati, and P. Anugrah, "ANALISIS FAKTOR RISIKO YANG BERHUBUNGAN DENGAN KEJADIAN DEMAM BERDARAH DENGUE (DBD)".

[11] D. D. Amanda and S. Martini, "Relationship Of Characteristic and Status of Central Obesity with The Prevalence Of Hypertension," *JBE*, vol. 6, no. 1, p. 57, Mar. 2018, doi: 10.20473/jbe.V6I12018.57-66.

[12] W. Widiantini and Z. Tafal, "Aktivitas Fisik, Stres, dan Obesitas pada Pegawai Negeri Sipil," *Kesmas: National Public Health Journal*, p. 325, Feb. 2014, doi: 10.21109/kesmas.v0i0.374.

[13] I. Trisna and S. Hamid, "FAKTOR-FAKTOR YANG BERHUBUNGAN DENGAN OBESITAS SENTRAL PADA WANITADEWASA (30-50 TAH UN) DI KECAMATAN LUBUK SIKAPING TAHUN 2008," *JKMA*, vol. 3, no. 2, pp. 68–71, Mar. 2009, doi: 10.24893/jkma.v3i2.62.

[14] M. Wirastuti and A. Sofro, "Analisis Longitudinal pada Data Pasien Diabetes Melitus," *J Statistika*, vol. 12, no. 1, pp. 13–19, Jul. 2019, doi: 10.36456/jstat.vol12.no1.a1993.

[15] S. Sukirman, "Analisis Cloud Computing Untuk Penyimpanan Dokumen Terhadap Proses Pembelajaran Menggunakan Algoritma Regresi Linear Berganda (Studi Kasus : SMA Chandra

Kusuma Jakarta Utara)," *Jurnal.ilmiah.informatika*, vol. 7, no. 1, pp. 1–12, Jun. 2022, doi: 10.35316/jimi.v7i1.1-12.
[16] P. R. Arum, R. P. Gautama, I. Fitriani, and F. Naza, "Identifying Factors that Influence Life Expectancy in Central Java Using Spatial Regression Models," vol. 16, no. 2, 2023.
[17] L. D. Martias, "STATISTIKA DESKRIPTIF SEBAGAI KUMPULAN INFORMASI," *FHRS*, vol. 16, no. 1, p. 40, Jun. 2021, doi: 10.14421/fhrs.2021.161.40-59.
[18] Richard A J et al and Dean W, *Applied multivariate statistical analysis*, 5th ed. 2002.
[19] Alvin C. Rencher, *Multivariate Statistical Inference and Applications*. A Wiley-Interscience Publication, 1998.
[20] N. Suhandi, E. A. K. Putri, and S. Agnisa, "Analisis Pengaruh Jumlah Penduduk terhadap Jumlah Kemiskinan Menggunakan Metode Regresi Linear di Kota Palembang," vol. 09, no. 2, 2018.
[21] R. Bevans, "Akaike Information Criterion | When & How to Use It (Example)," *Scibbr*, Jun. 2023, [Online]. Available: https://www.scribbr.com/statistics/akaike-information-criterion/
[22] A. Salsabila, Erfiani, Indahwati, A. . Fitrianto, and M. A. . Aliu, "Association of Poverty Categories, Educational Characteristics, and Area of Residence in Indonesia Using a Three-Way Log-Linear Model", *J Statistika*, vol. 17, no. 1, pp. 624–634, Jul. 2024.
[23] G. Anuraga, A. Indrasetianingsih, and M. Athoillah, "Pelatihan pengujian hipotesis statistika dasar dengan software r," *BUDIMAS: Jurnal Pengabdian Masyarakat*, vol. 3, no. 2, pp. 327–334, 2021.

Generalized Poisson Regression Modeling on the Number of Infant Deaths in East Nusa Tenggara Province in 2022

Maria Ririanti da Rato [(1)], Robertus Dole Guntur [(2)]

[1,2]Mathematics Department, Science and Engineering Faculty, Universitas Nusa Cendana

Jln. Adisucipto, Penfui Timur, Kabupaten Kupang, Nusa Tenggara Timur

e-mail: ririndarato17@gmail.com[(1)] , robertus_guntur@staf.undana.ac.id[(2)]

ABSTRAK

Jumlah kematian bayi di Provinsi Nusa Tenggara Timur (NTT) masih berada di atas rata-rata jumlah kematian bayi secara nasional. Penelitian ini dilakukan untuk menentukan model *Generalized Poisson Regression* (GPR) dalam mengatasi overdispersi pada model regresi poisson pada kasus jumlah kematian bayi serta untuk melihat faktor-faktor yang berpengaruh signifikan terhadap jumlah kematian bayi di Provinsi NTT. Variabel penelitian yang digunakan yakni jumlah kematian bayi sebagai variabel repson, dan sejumlah variabel prediktor yang diduga dapat mempengaruhi variabel respon. Data yang digunakan merupakan data sekunder yang diperoleh dari publikasi Badan Pusat Statistik (BPS) Provinsi Nusa Tenggara Timur (NTT) dari setiap 22 Kota/Kabupaten. Hasil penelitian menunjukkan bahwa data jumlah kasus kematian bayi mengalami overdispersi dengan rasio antara deviance dan derajat kebebasan sebesar 3,578. Pemodelan dengan GPR menunjukan bahwa model dengan 5 variabel bebas menghasilkan model yang optimal dengan nilai AIC sebesar 184,145. Kelima variabel yang berpengaruh signifikan terhadap jumlah kematian bayi di Provinsi NTT yakni persentase rumah tangga dengan akses sanitasi layak, persentase persalinan ditolong oleh pihak diluar tenaga medis, jumlah remaja yang mendapat penyuluhan kesehatan reproduksi, persentase penduduk usia 0-59 bulan menurut pemberian imunisasi tidak lengkap, dan jumlah fasilitas kesehatan.

Kata kunci: Jumlah Kematian Bayi, Generalized Poisson Regression (GPR), Overdispersi

ABSTRACT

The number of infant deaths in East Nusa Tenggara (NTT) Province is still above the national average. This research was conducted to investigate the Generalized Poisson Regression (GPR) model for addressing the overdispersion in the Poisson regression model of the number of cases of infant deaths and to explore the potential factors influencing the number of infant deaths in the province. The variable used is the number infant mortality as a response variable, and the number of predictor variables that are thought to influence the response variable. The data used is secondary data obtained from the publication by the Central Statistics Agency of ENT Province from each of the 22 cities/regencies. The study shows data on the number of cases of infant mortality experienced overdispersion with a ratio between deviance and degrees of freedom of 3.578. Modeling with GPR shows that the model with 5 independent variables produces an optimal model with an AIC value of 184.145. Those variables are the percentage of households with access to adequate sanitation, the percentage of births assisted by parties other than medical personnel, the number of teenagers who received reproductive health counseling, the percentage of the population aged 0-59 months who received incomplete immunization, and the number of health facilities.

Keywords: *Number Of Infant Deaths; Generalized Poisson Regression (GPR), Overdispersion*

INTRODUCTION

The number of babies who die before the age of one year in a certain period is known as the Infant Mortality Rate (IMR). Infant mortality remains a major problem throughout the world, especially for developing countries, even though IMR worldwide has made progress in reducing infant mortality [1].

Based on the results of the Indonesian Demographic and Health Survey (IDHS), there was a decrease in IMR in 2022 compared to the previous year, where infant mortality in 2022 was 16.85/1,000 Live Births while in 2017 it was 24/1,000 Live Births. Based on data on infant mortality in Indonesia, there has been a decrease in IMR, but it is still quite high when compared to other ASEAN countries, which is 4.2 times higher than Malaysia and 1.2 times higher than the Philippines [2].

Based on the results of the 2020 Population Census Long Form, it explains that IMR in East Nusa Tenggara Province (NTT) is 25.67/1,000 live births. The high IMR in NTT Province is caused by many factors including factors directly related to infants, one of which is related to low birth weight (LBW) which is the leading cause of infant mortality. The distribution of infant mortality cases in NTT Province is spread across all regions, however, the distribution pattern and mapping of the distribution of death cases have not been optimally investigated.

Data shows that the number of infant deaths in NTT Province continues to increase. In 2022, There are about 1244 cases, of which 114 cases were in Kupang district which is the highest infant mortality of 22 district in the province. Therefore, due to the characteristics of the district that vary from one another, it is necessary to investigate the potential factor contributing to this infant death this by developing a mathematical model. The distribution of infant mortality cases in NTT Province is spread across all regions, however, the distribution pattern and mapping of the distribution of death cases have not been optimally investigated [3].

The number of infant deaths is one example of count data where the data distribution follows the Poisson distribution [4]. In modeling with the Poisson regression model, the assumption of equal dispersion should be satisfied namely the mean and variance of the response variable must be the same. However, this condition is often violated when using real data where the mean of response variable is lower than variance, or vice versa so that the use of the Poisson regression model is no longer effective. For this, other methods are needed, for example using the Generalized Poisson Regression (GPR) model. The use of the GPR model is very effective in overcoming the problem of overdispersion [5–8]. For this reason, the GPR model will be applied to investigate the factors that influence the number of infant deaths in NTT Province. The results of this study are expected to contribute to the local government in designing health policy to reduce the IMR in NTT Province.

METHOD

Data Source

The data sources used in this study are secondary data regarding cases of infant mortality and factors suspected of influencing it in NTT Province in 2022. Data were obtained from the publication of the Central Statistics Agency (BPS) of NTT Province in figures 2023 and People's Welfare Statistics [9]. The unit analysis of this study is 22 cities/regencies in NTT Province.

Research Variables

The research variables consist of response variables (Y) and predictor variables (X). The response variables used in this study are the number of infant deaths (Y), as well as predictor variables which are factors suspected to influence the number of infant deaths in 22 districts/cities in NTT Province. The following are details of the predictor variables of this study.

X_1 = Percentage of poor population
X_2 = Percentage of households with access to proper sanitation
X_3 = Percentage of deliveries assisted by parties other than medical personnel
X_4 = Percentage of deliveries assisted by midwives/nurses
X_5 = Percentage of pregnant women under 19 years old
X_6 = Number of adolescents who received reproductive health education
X_7 = Percentage of population aged 0-59 months according to incomplete immunization
X_8 = Percentage of population aged 0-23 who have ever been breastfed
X_9 = Number of health facilities
X_{10} = Number of health workers

Analysis Methods and Analysis Steps

Data analysis was supported by SPSS software. The data processing steps are as follows: firstly, we conducted a descriptive statistic to find out the general picture of data on infant mortality and its predictors in NTT Province in 2022. Then we applied Poisson distribution test to assess the distribution of data.

The Poisson distribution test is used to determine whether the response variable in this study is distributed Poisson or not. For this reason, the Kolmogorov-Smirnov test will be used in this Poisson distribution test. The hypotheses used are as follows:

H_0 = The number of infant deaths is distributed Poisson
H_1 = The number of infant deaths is not distributed Poisson

The test statistics used are,

$$D_{count} = \max |F_n(x) - F_0(x)| \quad (1)$$

Then, we conducted a multicollinearity test**.** In the Poisson regression model, there are several assumptions that must be met. One of them is that there is no multicollinearity between predictor variables. Multicollinearity is a situation that indicates a strong correlation or relationship between two or more predictor variables in a model. Detection of multicollinearity cases can be done by looking at the Variance Inflation Factor (VIF) value. If the VIF value is greater than 10, then there is multicollinearity between the predictor variables [10]. The test statistics in multicollinearity are,

$$VIF = \frac{1}{(1 - r_{i,j}^2)} \quad (2)$$

Where $r_{i,j}$ is the correlation coefficient X_i and Y_j.

Furthermore, we conducted a Poisson regression modeling. The Poisson regression is a non-linear regression model that is often used to model the relationship between response variables in the form of discrete data (counts) with predictor variables in the form of discrete or continuous data [11]. The following is the general form of the Poisson regression equation for ten predictors.

$$\mu_i = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \cdots + \beta_{10} X_{10}) \quad (3)$$

In evaluating the effect of each parameter to the model in equation 3, we conducted a test for testing the significance of model parameters. Firstly, we did a simultaneous test. This test used the results of the omnibus test. The omnibus test is a likelihood ratio test to determine whether all predictor variables collectively improve the model compared to the intercept model alone. In this test, all predictor variables are tested together to see whether the predictor variables jointly affect the response variable [12]. In the omnibus test, there are the results of the likelihood ratio chi-square which is usually symbolized as G^2. The following are the criteria for the hypothesis being tested

$H_0: \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = \beta_{10} = 0$

$H_1: at\ least\ there\ is\ \beta_k \neq 0, k = 1,2,3,\ldots,10$

Test Statistics used,

$$G^2 = 2\sum_{i=1}^{n}(y_i x_i^T \hat{\beta} - \exp(x_i^T \hat{\beta}) - (y_i \hat{\beta}_0 \exp(\hat{\beta}_0))) \quad (4)$$

The omnibus test is expected to have a value of $G^2 > \chi^2_{table}$ that can reject the null hypothesis, or can be explained in the following table. Secondly, we conducted the partial test aimed to determine the predictor variables having a significant effect on the response variable [13]. The Wald test will be used to obtain predictor variables that have a significant effect on the response variable. The following are the test hypothesis criteria used

$H_0: \beta_k = 0, for\ k = 1,2,\ldots,10$

$H_1: \beta_k \neq 0, for\ k = 1,2,3,\ldots,10$

The test statistics in the partial test are,

$$W = \frac{(\hat{\beta}_k)}{SE\ (\hat{\beta}_k)} \quad (5)$$

Where $\hat{\beta}_k$ is the estimated value for parameter $\hat{\beta}_k$ and $SE\ (\hat{\beta}_k)$ is the estimated standard error of $\hat{\beta}_k$. The decision to be taken is if the $p - value < \alpha$ with the significance level used is 5% or the value of the Wald test is greater than X^2_{table}.

On the next stage, we did a test to investigate the possibility of the overdispersion of the data. When using Poisson regression, the equidispersion assumption must be met, namely the assumption of equality between the mean and variance values. However, Poisson regression modeling sometimes contains overdispersion, namely when the variance value is greater than the mean value. The phenomenon of overdispersion can be written as $var\ (Y) > E(Y)$. Conversely, data whose variance is smaller than the mean is called underdispersion. The phenomenon of underdispersion can be written as $var\ (Y) < E(Y)$. Overdispersion can be indicated by the deviance and Pearson chi-squares values divided by the degrees of freedom [14]. If both values are more than 1, then overdispersion is said to occur. Finally, we conducted modeling of Generalized Poisson Regression (GPR) and interpreting the results.

RESULT AND DISCUSSION

Descriptive Statistics.

Descriptive statistics aims to determine the characteristics of each variable. Variables are described using the amount of data, minimum value, maximum value, average value, standard

deviation and variance in Table 1. To see the number of infant mortality cases in each city/district in NTT, it can be seen in the following Figure 1.

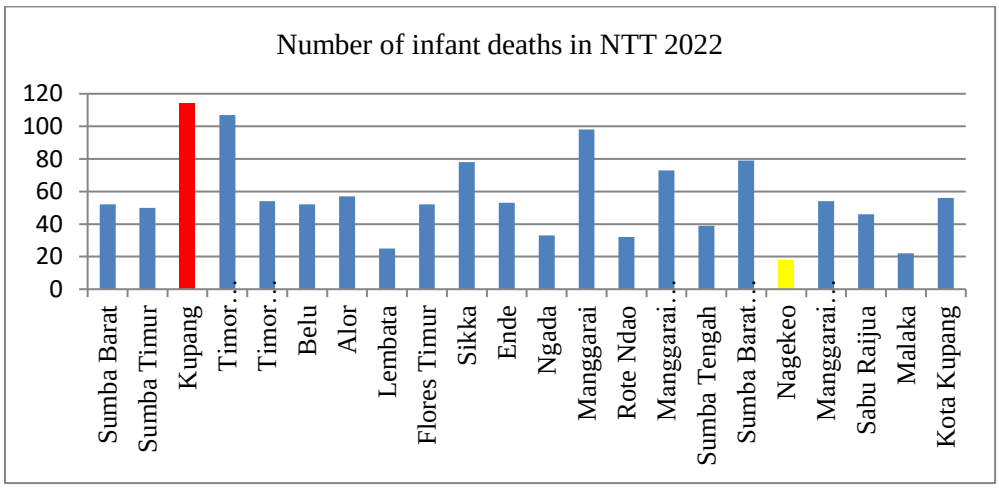


Figure 1. Graph of Infant Mortality Rate in East Nusa Tenggara Province in 2022

From Figure 1, the lowest number of infant mortality cases was 18 cases in Nagekeo Regency. Meanwhile, the highest number of infant mortality cases was 114 cases in Kupang Regency in 2022.

Table 1. Characteristics of Infant Mortality Data in NTT Province in 2022

Variables	N	Min	Max	Mean	Std. Deviation	Variance
Y	22	18	114	56,5455	25,98668	675,307
X_1	22	8,61	32,51	20,7732	6,91628	47,835
X_2	22	45,83	88,82	73,7091	14,17349	200,888
X_3	22	0	29,22	9,26	8,21708	67,520
X_4	22	658	8065	6389,2273	1650,08865	2722792,565
X_5	22	12	2764	1373,6091	822,31369	676199,8
X_6	22	200	1200	596,5909	239,20696	57219,968
X_7	22	22,20	49,19	33,4759	7,99320	63,891
X_8	22	83,75	100	95,6568	4,42380	19,570
X_9	22	224	927	520,7727	187,83628	35282,0170
X_{10}	22	126	1985	1108,4091	493,9801	244016,348

From Table 1, the average number of infant deaths in NTT Province in 2022 was 56.5455 or 56 cases with a variance of 675.307.

Poisson Distribution.

Following the equation (1), the Kolmogorov-Smirnov test had been applied to test Poisson distribution of the dependent variable. The value of Kolmogorov-Smirnov test is presented in Table 2.

Table 2. *Kolmogorov-Smirnov* Test Result

$One - Sample\ Kolmogorov - Smirnov$		
		Y
N		22
$Most\ Extreme\ Differences$	$absolute$	0,264
$Asymp.Sig.(2-tailed)$		0,093

From Table 2, obtained the Kolmogorov-Smirnov test value for $D_{count} = 0{,}264$ which is in the absolute value. While the $D_{table} = 0{,}281$ for a significance level of 5%. The $D_{count}(0{,}264) < D_{table}(0{,}281)$ is obtained or can be seen from the $p - value$ $(0{,}093) > \alpha$ $(0{,}05)$ then accept H_0 which means the number of infant mortality cases in NTT Province in 2022 is following the Poisson distribution.

Multicollinearity Test.

The strong correlation or relationship between two or more predictor variables was evaluated by looking at the Variance Inflation Factor (VIF) as indicated in equation (2). The results of the multicollinearity test on the number of infant mortality cases in NTT Province in 2022 are shown in the Table 3 below.

Table 3. VIF Values on Predictor Variables

Variables	VIF
Poor population (X_1)	2,843
Households with access to proper sanitation (X_2)	1,972
Delivery assisted by parties other than medical personnel (X_3)	3,140
Delivery assisted by midwives/nurses (X_4)	2,857
Pregnant women <19 years old (X_5)	1,816
Adolescents who receive reproductive health counseling (X_6)	2,029
Age 0−59 months who are given incomplete immunization (X_7)	2,444
Age 0-23 months who are given breast milk (X_8)	1,846
Health facilities (X_9)	2,549
Health workers (X_{10})	2,924

From Table 3, it can be seen that the VIF value of each predictor variable is less than 10, so there is no multicollinearity between the predictor variables of the number of infant mortality cases in NTT Province.

Poisson Regression Modeling.

In developing Poisson regression model, firstly, we did a simultaneous test by determining the value of the omnibus test as indicated in equation (4). The omnibus test is a likelihood ratio test to determine whether all predictor variables collectively improve the model compared to the intercept model alone. The value of Omnibus Test was presented in Table 4 below.

Table 4. Simultaneous Parameter Testing Using Omnibus Test

Omnibus Test	
Likelihood Ratio Chi-Square	206,149
df	10
Sig	0,000

Based on Table 4, if the model value (sig) $> \alpha$ or the $G^2 > \chi^2_{table}$, it means that H_0 is accepted. Because the model value (sig) in the omnibus test is 0,00 $> \alpha$ (0,05) or the $G^2(206{,}149) > \chi^2_{tabel}(18{,}307)$, then H_0 is rejected. So, it can be concluded that there is one or more predictor variables that have a significant relationship to the number of infant mortality cases in NTT Province in 2022.

Secondly, we conducted the partial test by calculating the Wald test value as shown in equation (5). The estimation of each parameter model and the value of Wald test for each predictor variable was presented in Table 5.

Tabel 5. Parameter Estimation Value

Parameter Estimation				
Parameter	Estimation	SE	Sig	*Wald Chi-Square*
β_0	3,286	0,9548	0,001	11,844
β_1	-0,003	0,0074	0,716	0,133
β_2	-0,013	0,0031	0,000	16,338
β_3	0,019	0,0065	0,003	8,729
β_4	-1,837e-5	3,2925e-5	0,577	0,311
β_5	-6,500e-6	4,2133e-5	0,877	0,024
β_6	0,001	0,0002	0,000	18,702
β_7	-0,019	0,0059	0,002	9,809
β_8	0,012	0,0092	0,185	1,757
β_9	0,001	0,0002	0,000	18,144
β_{10}	8,311e-5	9,6760e-5	0,390	0,738

Based on Table 5, the values (sig) of the parameters $\beta_0 = 0{,}001, \beta_2 = 0{,}000, \beta_3 = 0{,}003, \beta_6 = 0{,}000, \beta_7 = 0{,}002,$ dan $\beta_9 = 0{,}000$ are obtained, where the value of each parameter is smaller than the significance level used, which is 0,05. As well as the Wald Chi-Square test value, from each parameter $\beta_0 = 11{,}844$; $\beta_2 = 16{,}338$; $\beta_3 = 8{,}729$; $\beta_6 = 18{,}702$; $\beta_7 = 9{,}809$; dan $\beta_9 = 18{,}144$ where this value is greater than $X^2_{table} = 3{,}8415$. Then H_0 is accepted, which means that the predictor variables, namely the percentage of households with access to proper sanitation, the percentage of deliveries assisted by parties other than medical personnel, the number of adolescents who receive reproductive health education, the percentage of the population aged 0-59 months according to incomplete immunization, and the number of health facilities have a significant effect on the number of infant deaths in NTT province in 2022. Therefore the initial model of the number of infant mortality cases in NTT Province is

$$\mu_i = \exp(3{,}431 - 0{,}015X_2 + 0{,}024X_3 + 0{,}001X_6 - 0{,}019X_7 + 0{,}001X_9) \quad (6)$$

Testing the overdispersion of the data

Overdispersion can be indicated by the deviance and Pearson chi-squares values divided by the degrees of freedom [14,15]. If both values are more than 1, then overdispersion is said to occur. The results of the overdispersion test are presented in the Table 6 below.

Table 6. Overdispersion Test Results

Infant and toddler mortality	***Value***	***df***	***Value/df***
Deviance	39,359	11	3,578
Pearson cji-square	39,166	11	3,561

Table 6 shows that the Deviance value and Pearson Chi Square value are 3,578 and 3,561 respectively, which are more than 1. This shows that the Poisson regression model does not meet the equidispersion assumption or experiences overdispersion so that the Poisson regression model is not suitable for modeling the number of infant mortality cases in NTT Province in 2022. Therefore, further analysis was carried out using the Generalized Poisson Regression (GPR) model.

Generalized Poisson Regression (GPR) Modeling on the Number of Infant Mortality Cases in East Nusa Tenggara Province in 2022.

To determine the GPR model, namely by estimating and testing parameters where previously a simultaneous test had been carried out to see which variables had a significant effect on the model, but only involving one or several predictor variables selected based on the smallest Akaike Information Criteria (AIC) value [16,17]. AIC is a statistical measuring tool used to measure model suitability. The following are the results of the AIC test.

Tabel 7. AIC Values of All Selected Models

Variables	**AIC Value**	**Model**
X_2, X_3, X_6, X_7, X_9	184,145	$\exp(4{,}283 - 0{,}013X_2 + 0{,}019X_3 + 0{,}001X_6 - 0{,}019X_7 + 0{,}001X_9)$
X_2, X_3, X_6, X_9	196,242	$\exp(3{,}592 - 0{,}012X_2 + 0{,}011X_3 + 0{,}001X_6 + 0{,}001X_9)$
X_2, X_6, X_9	201,794	$\exp(3{,}862 - 0{,}015X_2 + 0{,}001X_6 + 0{,}001X_9)$
X_3, X_9	240,849	$\exp(2{,}985 + 0{,}017X_3 + 0{,}002X_9)$
X_9	260,159	$\exp(3{,}136 + 0{,}002X_9)$

Based on Table 7, the best model selected based on the smallest AIC value is the model involving $X_2, X_3, X_6, X_7,$ dan X_9 as predictor variables that have a significant effect with an AIC value of 184.145. So, the Generalized Poisson Regression model for infant mortality cases in NTT Province in 2022 is

$$\mu = \exp(4{,}283 - 0{,}013X_2 + 0{,}019X_3 + 0{,}001X_6 - 0{,}019X_7 + 0{,}001X_9) \quad (7)$$

Based on the model above, it can be interpreted as follows, if the variable has a constant value, then the number of infant deaths in NTT Province in 2022 that occurred was an average of $e^{4{,}283}$ or 74,457 cases. Each additional 1 coefficient of household population with Access to Proper Sanitation will reduce the average number of infant deaths by $e^{-0{,}013}$ or 0,987 cases. The more households that have inadequate access to proper sanitation, the more the number of infant deaths in each district/city will increase by 1 case.

Each additional 1 coefficient of childbirth assisted by parties other than medical personnel will increase the average number of infant deaths by $e^{0{,}019}$ or 1,019 cases. This means that if each delivery assisted by parties other than medical personnel increases, the number of infant deaths in each district/city will also increase by 1 case.

Each additional 1 coefficient of adolescents who receive reproductive health counseling will increase the average number of infant deaths by $e^{0{,}001}$ or 1,001 cases. This means that every teenager who receives reproductive health counseling is likely to increase the number of infant deaths in each district/city by 1 case.

Every additional 1 coefficient of the population aged 0-59 months according to incomplete immunization will reduce the average number of infant deaths by $e^{-0{,}019}$ or 0,981 cases. This means that the population at that age according to incomplete immunization can reduce the number of infant deaths in each district/city by 1 case.

Every additional 1 coefficient of health facilities will cause an increase in the average number of infant deaths by $e^{0{,}001}$ or around 1,001 cases. This means that if the health facilities obtained by the community increase, it is likely to increase the number of infant deaths in each district/city by 1 case.

CONCLUSION

The distribution of infant death is uneven in NTT Province with the highest was in Kupang district. The GPR model for representing the cases in this province in 2022 is $\mu = \exp(4{,}283 - 0{,}013X_2 + 0{,}019X_3 + 0{,}001X_6 - 0{,}019X_7 + 0{,}001X_9)$. Of the ten factors suspected of influencing the response variable, there are five factors that have an impact on the number of toddler deaths in NTT Province in 2022, which is the percentage of households with access to proper sanitation (X_2), the percentage of deliveries assisted by parties other than medical personnel (X_3), the number of adolescents receiving reproductive health education (X_6), the percentage of the population aged 0-59 months according to incomplete immunization (X_7), and the number of health facilities (X_9).

REFERENCES

[1] Nwanze LD, Siuliman A, Ibrahim N. Factors associated with infant mortality in Nigeria: A scoping review. *PLoS ONE* 2023;**18**(11):e0294434.

[2] Databox. Provinsi dengan Angka Tertinggi di Indonesia. https://databoks.katadata.co.id/datapublish/2023/02/27/papua-provinsidengan-angka-kematian-bayi-tertinggi-di-indonesia.

[3] Elisabeth Brielin Sinu, Ambrosius Dedi A. Sinu. Analisis Angka Kematian Bayi Di Provinsi Nusa Tenggara Timur Dengan Model Regresi Spasial. *jkpu-nalanda* 2023;**1**(6):287–299.

[4] Maxwell O, Mayowa BA, Chinedu IU, Peace AE. Modelling Count Data; A Generalized Linear Model Framework. *American Journal of Mathematics and Statistics* 2018;**8**(6):179–183.
[5] Chaniago AD, Wulandari SP. Pemodelan Generalized Poisson Regression (GPR) dan Negative Binomial Regression (NBR) untuk Mengatasi Overdispersi pada Jumlah Kematian Bayi di Kabupaten Probolinggo. *JSSITS* 2023;**11**(6):D448–D455.
[6] Saputro DRS, Susanti A, Pratiwi NBI. The handling of overdispersion on Poisson regression model with the generalized Poisson regression model. Surakarta, Indonesia, 2021: 020026.
[7] Guntur RD, Dappa JSB. PEMODELAN GENERALIZED POISSON REGRESSION PADA KASUS AIDS DI PROVINSI NUSA TENGGARA TIMUR TAHUN 2023. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research* 2024;**6**(2):95–104.
[8] Ibeji JU, Zewotir T, North D, Amusa L. Modelling fertility levels in Nigeria using Generalized Poisson regression-based approach. *Scientific African* 2020;**9**:e00494.
[9] BPS NTT. Statistik Kesejahteraan Rakyat Provinsi Nusa Tenggara Timur 2022. Badan Pusat Statistik Provinsi NTT: Kupang, 2022 https://ntt.bps.go.id/id/publication/2022/12/28/6a49f879c5d9063b1c4b5396/statistik-kesejahteraan-rakyat-provinsi-nusa-tenggara-timur-2022.html.
[10] Ihsan H, Sanusi W, Ulfadwiyanti R. Model Generalized Poisson Regression (GPR) dan Penerapannya pada Angka Pengangguran bagi Penduduk Usia Kerja di Provinsi Sulawesi Selatan. *JMathCoS* 2021;**3**(2):109.
[11] Maneking FDG, Salaki DT, Hatidja D. Model Regresi Poisson Tergeneralisasi untuk Anak Gizi Buruk di Sulawesi Utara. *JIS* 2020;**20**(2):141.
[12] Darnah. Mengatasi Overdispersi pada Model Regresi Poisson dengan Generalized Poisson Regresion I. *Jurnal Eksponensial* 2011;**2**(2):5–10.
[13] Maulana AD, Anggraini D, Sukmawaty Y. PEMODELAN KEPADATAN PENDUDUK DENGAN PENDEKATAN REGRESI POISSON TERGENERALISASI. *ragam* 2023;**2**(1):33.
[14] Dwi Putri A, Devianto D, Yanuar F. Pemodelan Jumlah Kematian Bayi di Kota Bandung dengan Menggunakan Regresi Zero-Inflated Poisson. *Jurnal Matematika UNAND* 2022;**11**(1):12.
[15] R. K. Putri, M. Athoillah, and A. Haqiqiyah, "Analisis Faktor Yang Mempengaruhi Ketepatan Kelulusan Mahasiswa Dengan Algoritma Regresi Linear," Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika, vol. 5, no. 2, pp. 671–680, 2024.
[16] W. Y. . Rusyida and A. Y. . Nugroho, "Multiperiod Logit on Survival Analysis of Financial Distress in Manufacturing Company", J Statistika, vol. 16, no. 1, pp. 354–370, Jul. 2023.
[17] Lais MF, Atti A, Pangaribuan RM, Guntur RD. Model Generalized Poisson Regression (GPR) Pada Kasus Stunting Di Provinsi Nusa Tenggara Timur. *JD* 2023;**5**(2):68–75.