BIS INFORMATION TECHNOLOGY AND COMPUTER SCIENCE

Volume 2 2025 | DOI: 10.31603/bistyc.246
6th Borobudur International Symposium on Science and Technology (BIS-STE) 2024

UNIMMA PRESS
UNIVERSITAS MUHAMMADIYAH MAGELANG

The influence of unhealthy lifestyles and habits on hypertension in data

Emilya Ully Artha[1*], Nova Aulia Burhan[1], Maimunah[1], Uky Yudatama[1], Endah Ratna Arumi[1]

[1] Universitas Muhammadiyah Magelang, Magelang, Indonesia
[*]Corresponding author's email: ully@ummgl.ac.id

Abstract

Serious illnesses like hypertension are frequently brought on by unhealthy eating habits and lifestyle choices. This disease, which manifests as fat, smoke, stress, and a lack of physical exercise, is not limited to metropolitan regions- it can also occur in suburban or rural areas. This study investigates the relationship between rural lifestyles and consumption habits with hypertension using data from BPS (BPS-Statistic Indonesia) Magelang Regency in Central Java. The population covered by the data includes 14 settlements, 10 columns, and 70 data points. The study's findings, which were derived by multivariate algorithms and linear regression, indicate that smoking and consuming processed foods and beverages both contribute to hypertension (74.76%). Particularly for rural areas, this study aids in the development of predictive models that identify and suggest healthy lifestyle modifications.

Keywords

Hypertension, Suburban, Rural, Linear regression, Predictive models

Introduction

Bad habits such as smoking, consuming fast food, and lack of vegetable intake have been shown to have a significant impact on health, especially in increasing the risk of hypertension. Smoking can damage blood vessels and increase blood pressure, while fast food high in salt and saturated fat can maintain this condition. In addition, lack of consumption of vegetables rich in fiber and essential nutrients can inhibit the body's ability to regulate blood pressure effectively [1].

Variables such as smoking habits, fast food consumption, lack of vegetables and hypertension will be used in this study, data taken from the Central Bureau of Statistics of Magelang Regency (BPS Kabupaten Magelang) from 2019 to 2023.

By using regression tests and multivariate tests to see the healthy lifestyle behavior of residents in Magelang Village, regression equations and relationships between these

Published:
April 28, 2025

Selection and Peer-review under the responsibility of the 6th BIS-STE 2024 Committee

variables will be obtained [2]. Our goal is to provide a comprehensive understanding of how data science can be harnessed to improve public health outcomes.

Method

From the variables that have been calculated for trend analysis, the next step is to calculate the statistical test (linear regression)[3] to see the influence of the variables. The next step is to use multivariate[4], [5] analysis to measure predictions between variables. The t-statistic test basically measures how much influence an individual of one independent variable has in explaining the variation of the dependent variable. This is done by comparing the calculated t value with the t table at a 95% confidence level (α = 0.05). The test criteria are to accept Ho if α <0.05, and reject Ho if α > 0.05

Z-score is a statistical measure that measures how far a value is from the population mean, measured in standard deviation units, and is used to standardize data and compare values from different distributions. Interpretation of the Z-score includes a value of 0 indicating a position exactly on the mean, a positive value indicating a position above the mean, and a negative value indicating a position below the mean. The Z-score is useful for detecting outliers, standardizing data, and hypothesis testing, but requires the assumption of a normal distribution and does not always provide a complete picture of the data distribution

Results and Discussion

The movement of hypertension data from 2019 to 2023 shows an increasing trend which will be explained in (Figure 1), there was a decrease in hypertension cases from 2019 to 2021. However, in 2022 and 2023, the number of hypertension cases increased again. Hypertension cases in 2019 were 1,812 people, in 2020 there were 1,169 people, in 2021 there were 891 people, in 2022 it increased to 3,999 people, and in 2023 it reached 7,383 people.

Figure 1. hypertension data visualization

The average movement of vegetable consumption from 2019 to 2023 shows instability with an increase followed by a decrease. This can be seen in Figure 2, the number of people consuming vegetables shows an increasing trend from 2019 to 2021, with a peak in 2021 of 2,873 people. However, in the following years there was a decrease, namely

to 2,736 people in 2022 and 2,442 people in 2023. Despite a significant initial increase, the amount of vegetable consumption tends to decline after reaching its peak in 2021.

Figure 2. Visualization of average vegetable consumption

The number of people consuming processed food and beverages has experienced a gradual decline as seen in Figure 3, from 10,685 people in 2019 to 9,223 people in 2021. However, this trend reversed in 2022 with an increase to 9,729 people, and continued to increase to reach 10,920 people in 2023, making it the highest number during that period.

Figure 3. Visualization of average fast food consumption

The number of people who consume cigarettes shows a consistent increase from 2019 to 2023 as seen in Figure 4. In 2019, the number was recorded at 3,169 people, and continues to increase every year to reach 4,662 people in 2023. This trend reflects 18 continuous growths in cigarette consumption over the five-year period.

Figure 4. Visualization of average cigarette consumption

The normal distribution graph of hypertension sufferers from 2019 to 2023 shows that the number of hypertension sufferers has an average of around 200 people, with the curve being seen in . This means that most of the data on hypertension sufferers is around the number 200. This distribution reflects that there are more cases around the average, while the number of cases that are much higher or lower than 200 is getting smaller, showing a regular and predictive distribution in the analysis of hypertension sufferer data during the period.

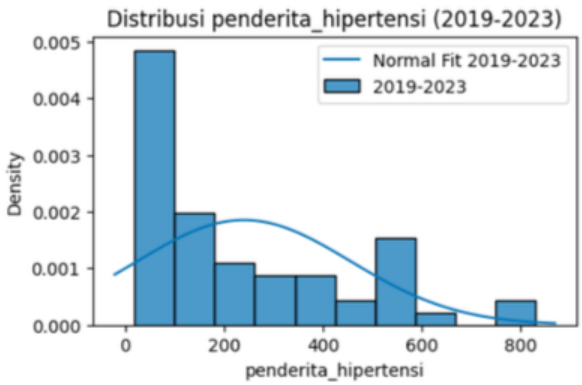

Figure 5. Distribution curve of hypertension sufferers

This study utilizes the statsmodels library to apply the Ordinary Least Squares (OLS) model. The results of the multiple linear regression model using OLS can be seen in Figure 6.

OLS Regression Results

Dep. Variable:	y	R-squared:	0.746
Model:	OLS	Adj. R-squared:	0.731
Method:	Least Squares	F-statistic:	50.94
Date:	Sat, 24 Aug 2024	Prob (F-statistic):	1.68e-15
Time:	05:00:25	Log-Likelihood:	-340.76
No. Observations:	56	AIC:	689.5
Df Residuals:	52	BIC:	697.6
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1429.3701	3796.784	0.376	0.708	-6189.434	9048.174
x1	-389.1018	390.820	-0.996	0.324	-1173.340	395.136
x2	-80.2811	177.620	-0.452	0.653	-436.702	276.140
x3	270.1006	106.215	2.543	0.014	56.965	483.236

Omnibus:	13.147	Durbin-Watson:	2.284
Prob(Omnibus):	0.001	Jarque-Bera (JB):	17.250
Skew:	0.851	Prob(JB):	0.000180
Kurtosis:	5.120	Cond. No.	4.61e+03

Figure 6. Results of multiple linear regression calculations

The R-squared value or coefficient of determination obtained is 0.746 and the remaining 0.254 is influenced by other factors. Based on the results, the resulting regression model equation is: $Y = 1429.37 - 389.1X1 - 80.28X2 + 270.1X3$. The results of the previous multiple linear regression analysis can be concluded as follows:

The constant (const) of 1429.37 indicates that if all independent variables have a value of zero, then the value of the dependent variable will be at 1429.37. This constant represents the intercept in the regression model.

The coefficient of the vegetable consumption variable of -389.10 indicates that every one-unit increase in vegetable consumption, assuming other variables remain constant, will cause a decrease in the dependent variable of 389.10. This negative coefficient indicates a negative relationship between vegetable consumption and the dependent variable.

The coefficient of the processed food and beverage consumption variable of -80.28 indicates that every one-unit increase in processed food and beverage consumption, assuming other variables remain constant, will cause a decrease in the dependent variable of 80.28. This also shows a negative relationship between consumption of processed food and beverages and the dependent variable.

The coefficient of the cigarette consumption variable of 270.10 shows that every one unit increase in cigarette consumption, assuming other variables remain constant, will cause an increase in the dependent variable of 270.10. This positive coefficient indicates a positive relationship between cigarette consumption and the dependent variable.

After the model was estimated using multivariate analysis, the results obtained were an R-squared value of 0.7476. The R-Squared value of 74.76% indicates that there is a significant relationship between the behavior of residents in Bandongan District who suffer from hypertension with healthy eating habits, such as vegetable consumption, and unhealthy lifestyles, such as consumption of instant foods and drinks, and smoking habits. Thus, it can be concluded that these habits significantly influence the behavior of residents who suffer from hypertension. The results of the F test on the multivariate model with an F-statistic of 65.1777 and a p-value of 0.0000. A high F-statistic value indicates that this model significantly explains the variability of the data, while a very small p-value indicates that the results are statistically significant. Thus, the variables in this model have a significant influence on the dependent variable of the r-squared results of the multivariate model.

Conclusion

Overall, this model has a good ability to explain the relationship between consumption patterns and hypertension, with an R-squared of 74.76%. The behavior of people who suffer from hypertension is significantly influenced by their eating habits and lifestyle.

Based on the results of the model estimation using multivariate analysis, the R-Squared value of 74.76% shows that healthy eating habits, such as vegetable consumption, as well as unhealthy lifestyles, such as instant food and beverage consumption, and smoking habits, have a significant influence on the behavior of people who suffer from hypertension. This shows the importance of diet and lifestyle in influencing health, especially in the context of hypertension.

Acknowledgement

I would like to express my deepest gratitude to all those who have supported me throughout this research project. First of all, I am very grateful to my fellow supervisors and examiners and students who are willing to contribute and are enthusiastic in examining the data used in more depth. I would also like to thank my colleagues and fellow researchers at the University of Muhammadiyah Magelang for their collaborative spirit and for providing a stimulating research environment. Furthermore, I express my appreciation to the participants of this study, whose cooperation and willingness to share their experiences have made this research possible.

Lastly, I am very grateful to my family and friends for their unwavering support and understanding throughout this journey. Their love and encouragement have been a constant source of strength.

References

[1] P. M. Hulls, R. C. Richmond, R. M. Martin, and F. De Vocht, "A systematic review protocol examining workplace interventions that aim to improve employee health and wellbeing in male-dominated industries," *Syst Rev*, vol. 9, no. 1, 2020, doi: 10.1186/s13643-019-1260-9.

[2] A. Sameh and I. Abunadi, "Piecewise multi-regression: superior predictor for housing index/prices at the districts of Riyadh-KSA," *... Conference on Data Science, E-Learning and ...*, 2019, doi: 10.1145/3368691.3368719.

[3] M. Khashei, A. Zeinal Hamadani, and M. Bijari, "A novel hybrid classification model of artificial neural networks and multiple linear regression models," *Expert Syst Appl*, vol. 39, no. 3, pp. 2606–2620, 2012, doi: 10.1016/j.eswa.2011.08.116.

[4] C.-C. M. Yeh, Z. Zhuang, W. Zhang, and L. Wang, "Multi-future Merchant Transaction Prediction," 2021, *Springer Science and Business Media Deutschland GmbH, Visa Research, Palo Alto, CA 94306, United States*. doi: 10.1007/978-3-030-67670-4_17.

[5] W. Chango, R. Cerezo, and C. Romero, "Predicting academic performance of university students from multi-sources data in blended learning," *... on Data Science, E-Learning and ...*, 2019, doi: 10.1145/3368691.3368694.