Indentifying Significant Variables of Poverty Within Indonesia Using Ridge and Principal Component Regression

Reyza Rahmatsyah   
*Statistics Department*  
*BINUS UNIVERSITY*DKI Jakarta, 11530  
reyza.rahmatsyah@binus.ac.idDi Raja Qusayyi Rabbani  
*Statistics Department*  
*BINUS UNIVERSITY*DKI Jakarta, 11530  
di.rabbani@binus.ac.id

Syarani Afa Natita Kusumah   
*Statistics Department*  
*BINUS UNIVERSITY*DKI Jakarta, 11530  
syarani.kusumah@binus.ac.id

*Abstract*—

Poverty has become a major issue that needs to be addressed around the world. It hinders the overall growth and prosperity of a country. For that reason, this study is aiming to identify and analyze the significant variables of poverty in Indonesia. To do that, multiple linear regression, ridge regression and Principal Component Regression (PCR) are being used to identify which variables have a significant impact on poverty. Since that the multiple regression model has violated the linear assumption which is multicollinearity, ridge regression and PCR are used to create a model to see the correlation between the independent variables and the dependent variable. Both ridge regression and PCR show that the average wage per hour has a significant impact toward poverty in a city. Both models also have a high R^2 score which is 0.87 and 0.78, meaning that the models show an adequate fit for the observed data. Further analysis using PCR shows that the rate of illiteracy also has a significant impact on poverty within a city.

Keywords—Principal Component Regression, Poverty, Ridge Regression

# Introduction

The welfare of its citizens is the dream of every country, including Indonesia. This can be seen from the number of citizens living in poverty. The challenge of poverty within the context of a culturally diverse and abundant natural resources poses an uphill battle for a country like Indonesia. Poverty is an important issue that hinders the development of a country. Poverty itself can include various things, not limited to financial shortages, but the lack of access to education, health, food, clothing, shelter, and others [8].

According to Ishartono and Rahorjo, poverty is a person's inability to meet basic or essential needs, which in turn causes more issues. These issues can be in the form of malnutrition, rapid exposure to disease, lack of access to education, which in the end poverty leaves behind a generation of people with social problems. Meanwhile, the World Bank defines poverty as it relates to welfare deprivation. Based on these definitions, people who live in poverty do not have enough resources to fulfill their lives.

Although categorized as a developing country, there exist a significant social divide within the Indonesian society [1]. We can categorize this divide as follows. The lower the poverty rate is, the smaller the number of people who live below the poverty line and more people are living above the poverty rate [2].

Poverty is and always will be a major thorn in the development of a country. Combating it will need full attention and a concentrated effort from all elements of society. Addressing poverty is not only the responsibility of the government, but also requires the participation of the private sector, community organizations, and individuals. To better understand this issue, we first need to know what factors are significantly influencing the poverty rate within one country. We identified this issue as a pressing matter that needs to be researched.

Based on *Badan Pusat Statistika* (BPS) data, Indonesia's poverty rate in March 2012 was 11.96% with a population of 29.25 million people. Meanwhile, in September 2012, the population level decreased with a value of 28.71 million people and a poverty rate of 11.66%. The data also shows that from 2012 to 2022, the poverty rate and population in Indonesia have decreased. The percentage of the poor people who live in the villages is 12.36%, while the percentage of the people who live in the urban areas is 7.53%.

# Previous Research

Hendriana and Kurniawan [2] in their article conducted research using multiple linear regression to analyze the effect of variables such as economic growth, human development index, and provincial minimum wage on poverty rates in Indonesia. The authors' objective was to identify the poverty rate at the provincial level in Indonesia and which variables affect it. The results showed that economic growth and unemployment rate have a significant effect on poverty in Indonesia. In addition, it was also concluded that economic growth has a negative effect, while the unemployment rate has a positive effect on the poverty rate.

Aulina and Mirtawati [3] in their article conducted research that aims to analyze the factors that affect the poverty rate in Indonesia in the period 2015-2019. The variables used by the researchers in panel data regression analysis to determine the factors affecting poverty in Indonesia are poverty rates, economic growth, literacy rates, average years of school education, open unemployment rates, and human development index. The results showed that the economic growth variable (GRDP) had a negative and significant effect, and the open unemployment rate variable had a positive and significant effect on the poverty rate in Indonesia in 2015-2019. It can be concluded that economic growth and open unemployment rate are factors that affect the poverty rate in Indonesia.

Suprijati and Damayanti [4] in their article conducted research aimed at examining poverty alleviation in 31 provinces in Indonesia through economic growth formed from domestic investment and workers. The variables used by researchers in panel data regression analysis to analyze the effect of PMDN and workers on economic growth, as well as the relationship between economic growth and poverty in cities and villages are economic growth, domestic investment (PMDN), the number of workers, urban poverty, and rural poverty. The results show that domestic investment has a negative and significant effect on economic growth, while the number of workers has a positive and significant effect on economic growth. In addition, economic growth also has a positive and significant effect on urban and rural poverty. It can be concluded that the economic growth formed by foreign direct investment and labor can have an impact on the poverty rate in Indonesia.

Aini and Islamy [14] in their article examined the influence of variables such as unemployment rate, education, health, GRDP, and HDI on poverty rate in Indonesia. This research was conducted using quantitative methods with secondary data from Badan Pusat Statistik Indonesia. The analysis method used is panel data analysis using Eviews 8 software. The results showed that the variables of unemployment and education had no significant effect on poverty in Indonesia. However, the variables of health, GRDP and HDI show an influence on poverty. GRDP and HDI have a significant effect on the poverty rate, while health is also shown to influence poverty.

Farida et al [15] in their article conducted research using regression analysis to see the effect of village development on poverty in Indonesia. The unit of analysis was 434 districts/cities in Indonesia. The dependent variable is the percentage of poor people as a measure of poverty in a region (Pov). The independent variables that are the focus of the analysis are the Village Development Index (HDI) which is broken down into five sub-indices, namely basic services, infrastructure, accessibility, public services, economic growth, and the Human Development Index (HDI). The results show that basic services, public services, economic growth, and HDI have a negative and significant influence on poverty in Indonesia. This means that an increase in these variables will be able to reduce the poverty rate in the community. In addition, this study also conducted hypothesis testing which includes a partial test (t test) and simultaneous test (F test). The partial test results show that four independent variables in the study proved to have a significant effect on poverty, namely basic services, public services, economic growth, and HDI.

# metodology

The results show within chapter and after III above was created using the R and RStudio Software.

## Dataset The dataset used for this research comes from Badan Pusata Statistika (BPS) website. The selection of the dataset indicators is based on the desire to analyze the factors that influence the poverty rate in Indonesia. The data collected is in provinces of Indonesia in 2022. This research consists of independent variables and dependent variables. These variables are explained as follows:

1. Dataset Variables

|  |  |
| --- | --- |
| Variable Name | Definition |
| Tingkat\_kemiskinan\_kota (Y) | Percentage of urban poverty |
| Gini\_ratio | A ratio that describes all forms of inequality. 0-1, where 0 is equal and 1 is inequality. |
| Tingkat\_penyelasaian\_SD | Percentage of primary school graduates |
| Tingkat\_penyelasaian\_SMP | Percentage of middle school graduates |
| Tingkat\_penyelasaian\_SMA | Percentage of high school graduates |
| Buta\_huruf\_15\_minus | Percentage of illiteracy among individuals below 15 years old |
| Buta\_huruf\_15 | Percentage of illiteracy among individuals above 15 years old |
| Buta\_huruf\_45 | Percentage of illiteracy among individuals above 45 years old |
| Anak\_bekerja | Percentage of child labor |
| Tingkat\_setengah\_pengangguran | Percentage of temporarily unemployed population |
| Tenaga\_kerja\_formal | Percentage of the population with formal employment |
| Kerja\_informal\_pertanian | Percentage of the population with informal employment in agricultural work |
| Upah\_rata2\_perjam | Average hourly income |
| Konsumsi\_kalori\_perhari | Daily calories consumption |
| Konsumsi\_protein\_perhari | Daily protein consumption |
| GK\_perkotaan | Minimum expenditure level for basic needs (in Rupiah/capita/month) in urban areas |
| GK\_non\_makanan\_kota | Minimum expenditure limit for non-food needs (in Rupiah/capita/month) in urban areas |
| Tingkat\_kerentanan\_penduduk | Percentage of economic instability |
| Kepemilikan\_akta\_40kebawah | Percentage of the population under 40 years old with legal documents |
| Median\_penduduk\_dibawah40 | Median income of the population under 40 percentiles. |

## Base Model

### Simple Linear Regression

Simple linear regression [12] is a method in statistics to model the relationship between the dependent variable and an independent variable. A simple linear model is a simple straight line that relates one independent variable to the dependent variable. The equation for this model is as follows:

Where Y is the dependent variable and X is the independent variable. One the other hand, is called the intercept and is called the slope. If the slope is 0, then there is no connection between X and Y. But if the slope is negative, then there is a negative connection between X and Y. Likewise, when the slope is positive, then the connection between X and Y is also positive [12].

### Multiple Linear Regression

### Multiple linear regression [13] is an extension of simple linear regression that models the relationship between a singular dependent variable with two or more independent variables. A multiple linear regression is the same as simple linear regression where the model is also a straight line, but the difference is that there are more independent variables. The equation for this model is as follows:

Where Y is the dependent variable and X1, X2, X3,…,Xn is the independent variables. are the slopes and is the intercept of the model. The slopes interpret the connection between the independent variables and the dependent variable. If the value of the slope is negative, then the connection is also negative. On the other hand, if the value of the slope is positive, then the connection is positive as well. The larger the slope, the stronger the connection is between the independent variables and the dependent variable [13].

### The model generated from these two methods are used as a base for further analysis. The generated model is then tested for linear assumption violations. Violations of the linear assumptions will result in a development of the model that tries to mitigate the violation. Below is the generated model using Multiple Linear Regression where “Tingkat\_kemiskinan\_kota” is the dependent variable. (1)

## Residual Analysis

After defining the base model, we then test the model for linear assumption violations. The results of each are shown below:

### Normality Test

The normality test is performed to test whether the errors in the data are normally distributed. The normality test must show normally distributed errors.

To test the normality of a model, we use the (Kolmogorov-Smirnov) *Lillie.test* function. We can assume that the errors are normally distributed if the *p-value* from the test is greater than . Table II shows that our *p-value* 0.4668 Which is greater than . Therefore, we can assume that the errors are normally distributed.

1. Kolomgorov- Smirnov Normality Test

|  |  |
| --- | --- |
| D | 0.10369 |
| p-value | 0.4668 |

### Homoscedasticitas Test

The homoscedasticity test is performed to test whether the error variation for each independent variable is constant. The homoscedasticity test must show constant error variation.

To test the homoscedasticity of a model, we use the (Breusch-Pagan) *bptest* function. We can assume homoscedasticity if the *p-value* resulting from the test is greater than . Table III shows that our *p-value* is which is greater than . Therefore, we can assume that there is homoscedasticity in the data.

1. Breusch-pagan Heteroscedasticity Test

|  |  |
| --- | --- |
| BP | 18.285 |
| p-value | 0.5035 |

### Autocorrelation Test

The autocorrelation test is performed to test whether there is a correlation between the times of the variables used. The autocorrelation test must show that the residuals are independent, in other words, No Autocorrelation.

To test the autocorrelation of a model, we use the (Durbin-Watson) *dwtest* function. We can assume the absence of autocorrelation if the resulting from the test is greater than . It can be seen in Table IV. our result is 0.2482 which is greater than . Therefore, we can assume that there is no autocorrelation in the data.

1. Durbin-Watson Autocorrelation Test

|  |  |
| --- | --- |
| DW | 2.0866 |
| p-value | 0.2482 |

### Multicolinearitas Test

The multicollinearity test is performed to test whether there is a linear relationship between independent variables. The multicollinearity test must show that the independent variables do not have a perfect linear relationship.

To test the multicollinearity of a model, we use the *Varian Inflation Factor* (VIF) function. We can assume the presence of multicollinearity if the VIF resulting from the test is greater than 10. It is shown in Table V. the variables "buta\_huruf" and "Tingkat\_Penyelesaian" have VIF values greater than 10. These results indicate that there is multicollinearity in the dataset used for modeling.

1. VIF Test

|  |  |
| --- | --- |
| Variable Name | VIF |
| Gini\_ratio | 5.08753 |
| Tingkat\_penyelasaian\_SD | 16.82353 |
| Tingkat\_penyelasaian\_SMP | 11.1087 |
| Tingkat\_penyelasaian\_SMA | 7.860765 |
| Buta\_huruf\_15\_minus | 84.26995 |
| Buta\_huruf\_15 | 584.0985 |
| Buta\_huruf\_45 | 307.3356 |
| Anak\_bekerja | 3.793655 |
| Tingkat\_setengah\_pengangguran | 8.627038 |
| Tenaga\_kerja\_formal | 24.46599 |
| Kerja\_informal\_pertanian | 9.583717 |
| Upah\_rata2\_perjam | 6.217335 |
| Konsumsi\_kalori\_perhari | 12.34225 |
| Konsumsi\_protein\_perhari | 20.16863 |
| GK\_perkotaan | 18.4239 |
| GK\_non\_makanan\_kota | 18.0415 |
| Tingkat\_kerentanan\_penduduk | 11.00783 |
| Kepemilikan\_akta\_40kebawah | 10.42408 |
| Median\_pendapatan\_penduduk\_dibawah40 | 9.377495 |

The results from these tests performed on the model showed that there is a fundamental problem within the dataset. We can assume that a normal multiple linear regression is not adequate for this dataset. Because the problem is multicollinearity, we can circumvent this issue by using the Ridge Regression or the Principal Component Regression.

## Model Development

### Ridge Regression

Ridge regression is often used to overcome multicollinearity problems in regression analysis. The problem [10] occurs when the independent variables in the regression model have a strong relationship with each other, causing the regression coefficients to be unstable and difficult to interpret. To overcome multicollinearity, the ridge [9] method adds a bias constant C to the diagonal of . The value of C for the ridge regression coefficient is between 0 and 1. In simple form is as follows:

(2)

### Principal Component Regression

Principal Component regression (PCR) is a method used to handle multicollinearity and dimensionality reduction in regression models. To manage multicollinearity in a model, the PCR [11] method initially performs a Principal Component Analysis (PCA) on the independent variables. Then, those principal components will be used as independent variables in a new regression model. The number of principal components used in PCR is chosen by considering the proportion of variance explained [8]. The cutoff used in this paper will be 80% .

# Results and Analysis

## Ridge Regresssion

To make a ridge regression, we use the lmridge library from the R software. Based on probability shown in Table VI, the significant variables are *Gini\_ratio*, and *Upah\_rat-perjam*

1. Ridge Regression

|  |  |  |
| --- | --- | --- |
| Variable | Estimate | Pr > t |
| Intercept | -3.31E+01 | 0.0398 \* |
| Gini\_ratio | 4.38E+01 | 0.0025 \*\* |
| Tingkat\_penyelasaian\_SD | 2.23E-01 | 0.5298 |
| Tingkat\_penyelasaian\_SMP | 3.12E-01 | 0.0512 . |
| Tingkat\_penyelasaian\_SMA | -1.16E-01 | 0.1066 |
| Buta\_huruf\_15\_minus | 1.08E+00 | 0.2658 |
| Buta\_huruf\_15 | -1.85E+00 | 0.2985 |
| Buta\_huruf\_45 | 6.63E-01 | 0.3424 |
| Anak\_bekerja | -8.89E-01 | 0.0251 \* |
| Tingkat\_setengah\_pengangguran | -2.23E-01 | 0.4352 |
| Tenaga\_kerja\_formal | -2.82E-01 | 0.0370 \* |
| Kerja\_informal\_pertanian | -3.34E-01 | 0.0163 \* |
| Upah\_rata2\_perjam | -5.00E-04 | 0.0097 \*\* |
| Konsumsi\_kalori\_perhari | 6.30E-03 | 0.3855 |
| Konsumsi\_protein\_perhari | 5.70E-03 | 0.9792 |
| GK\_perkotaan | 0.00E+00 | 0.0253 \* |
| GK\_non\_makanan\_kota | 0.00E+00 | 0.7009 |
| Tingkat\_kerentanan\_penduduk | 2.98E-02 | 0.9278 |
| kepemi1ikan\_akta\_40kebawah | -6.59E-02 | 0.4048 |
| Median\_pendapatan\_penduduk\_dibawah40 | 1.66E-01 | 0.1727 |

An of indicates that the ridge regression can explain of the variance within can be explained by the model. Table VI shows that the significant variables impacting the percentage of poverty within a city are the gini ratio, average wage, percentage of people in formal and informal workforces, and the expenses in a city.

## Principal Component Regresssion

Before using the principal components in a regression, we first need to define the number of components used. This is achieved by creating a principal component for all variables in the data and looking at the variances explained. Based on Figure 1, most of the variances are explained between Component 1 through 5.

A graph with numbers and lines

Description automatically generated

Figure 1. Principal Component Analysis Results

After deciding the number of components, we then pass these components into the linear regression. Results in Table VII, show that components 1,2,4, and 5 are significant towards predicting the overall poverty percentage of poverty within a city.

1. Summary Table From PCR

|  |  |  |  |
| --- | --- | --- | --- |
|  | Estimate | Pr>t | Significance |
| (Intercept) | 7.26706 | < 2e-16 | \*\*\* |
| Comp.1 | -0.27942 | 0.00338 | \*\*\* |
| Comp.2 | -0.66581 | 2.26e-06 | \*\*\* |
| Comp.3 | 0.21506 | 0.20486 |  |
| Comp.4 | 1.33488 | 4.05e-07 | \*\*\* |
| Comp.5 | 0.88453 | 0.00149 | \*\*\* |

From the output in Table VII, using component 1,2,4, and 5, the model created is:

(3)

Each PCA uses different combinations of components. Inside each PCA, the combinations are weighted and dropped accordingly. Examining the combinations inside the PCA can give insight on the significant variables that the model uses. Using the loadings function in R, loadings represent the relationships between the original variables and the principal components.

1. Loadings

|  |  |  |  |
| --- | --- | --- | --- |
| PCA1 | PCA2 | PC4 | PC5 |
| >0.1 | -0.31541 | -0.09172 | 0.201924 |
| 0.229735 | -0.29593 | >0.1 | >0.1 |
| 0.235291 | -0.22203 | 0.219583 | >0.1 |
| 0.276732 | -0.11694 | 0.154477 | 0.234508 |
| -0.26126 | >0.1 | 0.177469 | 0.152524 |
| -0.23626 | 0.239817 | 0.125479 | >0.1 |
| -0.2561 | >0.1 | 0.204537 | 0.112601 |
| -0.23775 | >0.1 | -0.32103 | -0.29731 |
| -0.19667 | -0.10222 | 0.289299 | -0.12983 |
| 0.325866 | >0.1 | >0.1 | >0.1 |
| -0.22422 | -0.26995 | >0.1 | >0.1 |
| -0.32587 | >0.1 | >0.1 | >0.1 |
| 0.159481 | 0.286389 | >0.1 | -0.13357 |
| >0.1 | -0.24011 | 0.328028 | -0.50974 |
| 0.185254 | -0.2259 | 0.200267 | -0.49644 |
| 0.1406 | 0.334706 | 0.307614 | >0.1 |
| 0.13866 | 0.358547 | 0.265185 | >0.1 |
| -0.18007 | >0.1 | 0.138717 | -0.26286 |
| 0.23519 | -0.24621 | -0.13745 | 0.168583 |
| -0.26265 | -0.19235 | -0.20329 | >0.1 |

The larger the absolute value of a variable in a principal component, the larger the contribution of said variable toward the component. A value close to 0 indicates a weak contribution to the component. Loadings are useful to find patterns of contributing significant variables. But this is not a good baseline, choosing significant variables on loadings can be difficult [16].

based on the equation in (3), the model uses the PCA in predicting poverty within a city. We can transform the PCA back into its original variable of by using the formula:

(4)

is the coefficients transformed back into the variable for . are the coefficients from Table VII or equation (3), and is the eigenvectors or the loadings of the data [16]. Using the equation in (4), we can transform the model back into its original variable . The coefficients found:

(5)

An of indicates that the ridge regression can explain of the variance within can be explained by the model. Using The values in (5), interpreting the model will be based on the value of coefficient from each variable. If the positive and negative values indicate the direction of the relationship, we can interpret the result as:

1. When all predictor variables are 0, the base value for the percentage of poverty within one city is 7.267.
2. The significant contributing factors for the increase of the percentage of poverty within a city, are gini ratio (0.213), Percentage of graduates from primary, middle, and high school (0.274, 0.488, 0.394), illiteracy particularly from the age of 15-45 (0.314 and 0.378), the half unemployment rate (0.477), those who do farm work (0.314), and expenses within a city for its citizens (0.227).
3. The significant contributing factors for the decrease of the percentage of poverty within a city are the percentage of people working formally (-0.363) and the average wage per hour (-0.312)

# Conclusion

Poverty is an ongoing issue where all aspects of society have a role in fighting. But, going into a fight without properly checking and preparing oneself is detrimental to the efforts in reducing poverty. Knowing what programs to support is crucial in developing a framework where poverty is tackled in a clear and concise manner. Using a linear regression to check the importance of certain variables is a step in the right direction. Despite this, using just one model is not sufficient in interpreting the results. Violations towards regression assumptions will skew the meaning of each significant variable.

To counteract the issues of violations, particularly the violation of multicollinearity, the paper used the ridge and principal component regression. Both models shown an adequate fit for the data producing an of and respectively. The significance of each variable can also be concluded from the respective output summary from each model. Both the Ridge and the Principal Component Regression agree that the average wage per hour (*Upah\_rata2\_perjam*) is a contributing variable towards the decreasing percentage of poverty in a city. Further analysis in the principal component regression shows that the rate of illiteracy is a contributing variable in increasing the percentage of poverty within a city.

The research conducted in this study is not a rigid conclusion. Poverty is a multilayered issue where fixing one problem may lead to another. The results found in this study should be treated as a starting point for a more deliberate approach in handling poverty. Results in this study can be used as a guideline for programs to invest into and what factors of society need help. This study is limited to only finding the significant variables within a given dataset. Predicting the poverty percentage may also need more intuitive model selection and variable selection. Future research on this issue needs to use more varied data to further narrow the significance of each variable.

##### References

1. Wiganepdo, S. S., Soegoto, H., Komputer Indonesia, U., Dipati Ukur, J., & Barat Indonesia, J. (2022). PERAN PMDN DAN PMA TERHADAP PENURUNAN ANGKA KEMISKINAN DI INDONESIA PROGRAM STUDI MANAJEMEN FAKULTAS EKONOMI DAN BISNIS UNIVERSITAS KOMPUTER INDONESIA BANDUNG. In *JURISMA: Jurnal Riset Bisnis dan Manajemen* (Vol. 12, Issue 1).
2. Herdiana, A., & Kurniawan, F. (n.d.). (2022). Studi Kasus Kemiskinan Di Indonesia Level Provinsi Dan Faktor-Faktor Yang Mempengaruhinya Menggunakan Regresi Linear Berganda. In *Jurnal Matematika dan Statistika serta Aplikasinya* (Vol. 10, Issue 1).
3. Aulina, N., & Mirtawati. (2021). ANALISIS REGRESI DATA PANEL PADA FAKTOR-FAKTOR YANG MEMPENGARUHI KEMISKINAN DI INDONESIA TAHUN 2015 – 2019. In *Kinerja Jurnal Ekonomi dan Bisnis* (Vol. 4, Issue 1).
4. Suprijati, J., & Damayanti, S. R. (2022). *PENGENTASAN KEMISKINAN KOTA DAN DESA 31 PROPINSI DI INDONESIA MELALUI PERTUMBUHAN EKONOMI YANG DIBENTUK DARI PMDN DAN PEKERJA* (Vol. 6, Issue 1).
5. Wicaksono, S. P., Hutajulu, D. M. (2023). Analisis Faktor-Faktor yang Mempengaruhi Kemiskinan di Indonesia Tahun 1999-2020. Transekonomika. 3(2).
6. Fadhillah, A., Arintoko, A., & Kamio, K. (2021). Dampak Investasi, Proyek dan Utang Luar Negeri Terhadap Kemiskinan Indonesia Tahun 2010-2020. Eksis: Jurnal Ilmiah Ekonomi Dan Bisnis, 12(1), 1. <https://doi.org/10.33087/eksis.v12i1.216>
7. Samputra, P.L., & Munandar, A.I. (2019). Korupsi, Indikator Makro Ekonomi, dan IPM terhadap Tingkat Kemiskinan di Indonesia. Jurnal Ekonomi Kuantitatif Terapan.
8. Lee H, Park YM, Lee S.  **Principal Component Regression by Principal Component Selection**.  CSAM 2015;22:173-180.  <https://doi.org/10.5351/CSAM.2015.22.2.173>
9. Wailaine, T. L., Talakua, M. W., & Lesnussa, Y. A. (2014). Model regresi RIdge Untuk Mengatasi Model Regesi Linear Bergana yang Mengandung Multikolineritas. Jurnal Barekeng. 8(1).
10. Mcdonald, G. C. (2009). Ridge Regression. John Wiley & Sons, Inc. 1. [Sci-Hub | Ridge regression. Wiley Interdisciplinary Reviews: Computational Statistics, 1(1), 93–100 | 10.1002/wics.14](https://sci-hub.se/https://doi.org/10.1002/wics.14)
11. Massy, W. F. (2016). Principal Components Regression In Exploratory Statistical Research. Journal of hte American Statistical Association. 60(309). <https://www.jstor.org/stable/2283149>
12. Bangdiwala S.I. (2018). Regression: Simple Linear. INTERNATIONAL JOURNAL OF INJURY CONTROL AND SAFETY PROMOTION. 25(1). 113-115. <https://www.tandfonline.com/doi/full/10.1080/17457300.2018.1426702>
13. Tranmer, M., Murphy, J., Elliot, M., & Pampaka, M. (2020). Multiple Linear Regression (2nd Edition). Cathie Marsh Institute Working Paper. <https://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/multiple-linear-regression.pdf>
14. Aini, L. N., & Islamy, S. N. (2021). Dampak penganguran, pendidikan, kesehatan, PDRB dan indeks pembangunan manusia terhadap kemiskinan di Indonesia. *Journal of Economics Research and Policy Studies*, *1*(3), 132–141. https://doi.org/10.53088/jerps.v1i3.325
15. Farida, A., Arifuddin, M. H., Rahimi, N., & Kamarni, N. (2022). Pengaruh Pembangunan Desa terhadap Penanggulangan Kemiskinan di Indonesia. *Jurnal Samudra Ekonomi Dan Bisnis*, *13*(2), 191–205. <https://doi.org/10.33059/jseb.v13i2.5253>
16. Jolliffe, I. (2013). *Principal component analysis*. Springer Science & Business Media.