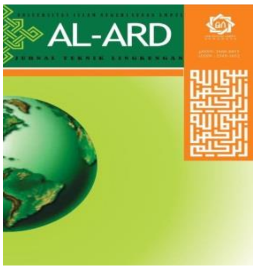

Al-Ard: Jurnal Teknik Lingkungan
Vol.7 No.2 – Maret 2022 (hal. 52-58)
http://jurnalsaintek.uinsby.ac.id/index.php/alard/index

Al-Ard:
Jurnal
Teknik Lingkungan

Determinants of the Amount of Waste in East Java

Joko Ade Nursiyono [1,*], Chusnul Chotimah [2], Warisna Endah Fitrianti [3]
[1,2,3] BPS Provinsi Jawa Timur, Surabaya, Indonesia
* Joko.ade@bps.go.id

Abstract
Listed as one of the largest waste contributor provinces in Indonesia. The population of East Java in 2020 reached 39 million people, it is the second highest in Indonesia. The increasing number of people accompanied by an increase in income will increase people's consumption in an area and this will cause the increasing amount of waste. If this waste problem is not handled properly, it will have a domino effect as well as degrading the environment. This study wanted to determine the effect of population, real expenditure per capita per year and the number of waste banks on the amount of waste in 2020 in East Java Province. This study uses a comparison of OLS Regression and Robust Regression models. The criteria for selecting the best model use the smallest MAPE, RMSE, and RSE values and the largest R-square value. The results of the partial test and the simultaneous test show that the variables of population, real expenditure per capita per year and the number of waste banks significantly affect the variable amount of waste in East Java with the selected model is the Robust Regression model. The R-square value of the Robust Regression model in this study is 0.8909, meaning that the model's ability to explain the variability of the East Java waste amount data is 89.09 percent, and the rest is explained by other variables not included in the model.
Keywords: Environment, Linear Regression, Robust Regression, Waste

Abstrak
Tercatat sebagai salah satu provinsi penyumbang sampah terbesar di Indonesia. Jumlah penduduk Jawa Timur pada tahun 2020 mencapai 39 juta jiwa, tertinggi kedua di Indonesia. Peningkatan jumlah penduduk yang disertai dengan peningkatan pendapatan akan meningkatkan konsumsi masyarakat di suatu daerah dan hal ini akan menyebabkan jumlah sampah yang semakin meningkat. Jika masalah sampah ini tidak ditangani dengan baik, maka akan berdampak domino sekaligus merusak lingkungan. Penelitian ini ingin mengetahui pengaruh jumlah penduduk, pengeluaran riil per kapita per tahun dan jumlah bank sampah terhadap jumlah sampah tahun 2020 di Provinsi Jawa Timur. Penelitian ini menggunakan perbandingan model Regresi OLS dan Regresi Robust. Kriteria pemilihan model terbaik menggunakan nilai MAPE, RMSE, dan RSE terkecil dan nilai R-square terbesar. Hasil uji parsial dan uji simultan menunjukkan bahwa variabel jumlah penduduk, pengeluaran riil per kapita per tahun dan jumlah bank sampah berpengaruh signifikan terhadap variabel jumlah sampah di Jawa Timur dengan model yang dipilih adalah model Robust Regression. Nilai R-square model Robust Regression dalam penelitian ini adalah 0,8909, artinya kemampuan model dalam menjelaskan variabilitas data jumlah sampah Jawa Timur sebesar 89,09 persen, dan sisanya dijelaskan oleh variabel lain yang tidak termasuk dalam model.
Kata Kunci: Lingkungan, Regresi Linier, Regresi Robust, Sampah

1. INTRODUCTION

One of the serious problems faced by big cities in Indonesia is solid waste. Currently, Indonesia is listed as the third largest contributor to waste in the world. In 2020 Indonesia produced 67.8 million tons of waste. Based on data from the Ministry of Environment and Forestry (KLHK), 37.3 percent of waste in Indonesia comes from household activities. The largest source of waste comes from traditional markets, amounting to 16.4 percent. Meanwhile, 15.9 percent of the waste originating from the commercial areas, and about 14.6 percent of the waste came from other sources. Waste is also generated by economic activities such as trades which is recorded at 7.29 percent, and waste generated from public facilities is 5.25 percent. Meanwhile, 3.22 percent of waste is generated in office areas.

KLHK also mentioned that 55.87 percent of the waste that was successfully managed throughout 2020. The remaining 44.13 percent of waste is still left because it has not been managed. In response to this condition, the Government is targeting a 30 percent reduction in waste and 70 percent in waste management by 2025. The approaches taken include limiting plastic waste and recycling inorganic waste.

As a logical consequence of the activities of human life, it cannot be denied, garbage will always exist as long as life activities are still running. Each year, the volume of waste will always grow in line with the pattern of society's growing consumerism (Suryani, 2014).

The same problem also occurs in East Java. East Java is listed as one of the largest waste contributor provinces in Indonesia. Based on data from the National Waste Management Information System or Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) in 2020, waste generation in East Java from 38 regencies/cities is 5,719,360.64 tons per year. Furthermore, waste reduction is 14.81 percent or a total of 847,276.93 tons per year. Meanwhile, the managed waste is 54.91 percent or 3,140,310.48 tons per year. So that the unmanaged waste is 45.09 percent or a total of 2,579,050.16 tons per year. This indicates that the existing waste generation still needs to be handled properly and carried out in a sustainable manner.

According to Madasari, et al (2020) in their research, they carried out modeling with the Dynamic System Methodology to model the Surabaya waste formation process, landfills or TPS facilities, and formulate waste management policies that can increase the amount of processed waste and reduce the amount of waste production, and can be sustainable in the future. Policy formulation is carried out by paying attention to the details of the relationship between the variables of the waste management system. There are three proposed scenarios used in their study. The results showed that the scenario of developing a waste bank facility of 1,200 tons had the most significant impact compared to the existing condition, which was reviewed in terms of the amount of waste heaped, the amount of waste being used as raw materials and energy sources, and increased absorption of workers.

According to Wardhani and Harto (2018), the results of their reserach showed that the impact of the Waste Bank to reduce the amount of waste disposed of in landfills or TPA was felt in the city of Surabaya. The existence of a Waste Bank can reduce approximately 81.50% of the total volume of waste each month. Based on the results of further analysis, it can be seen that the average waste volume reduction with the existence of a waste bank is around 7.2% of the total volume of waste in East Java. Waste management strategies include community involvement in independent waste management, waste management using a sanitary landfill in accordance with environmental standards, and the development of high-tech waste processing for energy sources.

According to Prajati et al, (2015) The results of the analysis of variable relationships show that the consumer price index, population and GRDP affect the generation of waste. The test results of the Daskalopoulos model show that waste generation in Java and Sumatra can be explained by 33.7% by the consumer price index per category. The results of the trial of the Khajuria model show that waste generation in Java and Sumatra can be explained by 42.5% by population, GRDP, and years of schooling. The results of the Khajuria model development test show that waste generation in Java and Sumatra can be explained by population, GRDP, years of schooling, literacy rate, population density and economic growth of 65.6%. Projection of waste generation is carried out using the development of the Khajuria model and the discriminant equation. The results of the projected waste generation show that the cities of Pangkalpinang and Tanjungpinang are the cities with the highest waste generation in the next five years. The waste management costs required by the two cities are quite large, which is above 0.8% of GRDP.

The population of East Java in 2020 reached 39 million people, the second highest in Indonesia with a population growth rate of 0.53 percent. The increasing population and accompanied by an increase in income will increase the consumption of people in an area, this causes the generation of waste to also increase. If this waste is not handled properly, it will have a multiplier effect on the environment around us. One alternative in managing this waste generation is the creation of a waste bank.

The purpose of this study was to determine the effect of population number (number of residents), real expenditure per capita per year and the number of waste banks on the amount of waste in 2020 in East Java.

2. RESEARCH METHODS

Research Scope

This study covers socio-economic conditions that are closely related to the amount of waste during 2020. This research data comes from several pages, data on the amount of waste and the number of Waste Banks from the Ministry of Environment's National Waste Management Information System or Sistem Informasi Pengelolaan Sampah Nasional (SIPSN), population, real expenditure per capita per year according to regencies/cities throughout East Java from the Central Statistics Agency or Badan Pusat Statistik (BPS) of East Java Province (jatim.bps.go.id). Given the limitations of data, this study only covers 2020 conditions in the East Java region.

Data Analysis

The analysis used in this study includes descriptive analysis and inferential analysis. Descriptive analysis aims to provide a description of the data seen from the mean (mean), standard deviation, variance, maximum, minimum, sum, range, kurtosis, and skewness descriptively (Anderha and Sugama, 2021). Descriptive analysis is generally applied to see the pattern of the dependent variable. While the inferential analysis used in this study is a multiple linear regression model with the Ordinary Least Square (OLS) parameter estimation method followed by Robust regression. According to Nursiyono and Pray (2016), linear regression is a regression model that is linear in parameters and describes a causal relationship (cause and effect) of the independent variable on the dependent variable. This causal relationship is the reason for choosing this research model. The following is the form of the equation of multiple linear regression used in this study:

$$yt = \beta_0 + \beta_1 Pddk_{1i} + \beta_2 Inc_{2i} + \beta_3 Jbs_{3i} + \varepsilon_i \qquad (1)$$

With β_0 is intercept mdoel, $\beta_1, \beta_2, \beta_3$ are multiple linear regression coefficents, $Pddk_{1i}$ is the amount of resident observations *i-th* (person), Inc_{2i} real spending per capita per year observation *i-th* (thousand rupiahs), Jbs_{3i} is the amount of waste bank observation *i-th* (unit), ε_i is the error component (residual) on the *i-th* observation, while *i* = 1, 2, 3, …, 38.

The selection of this research model uses the method *enter*. Which is, all independent variables are directly entered into the model to see whether each independent variable has a partial and significant effect on the dependent variable. After the regression model is formed, the next step is to test the classical assumptions. This assumption test is used to ensure that the estimator generated in the modeling is BLUE (Best Linear Unbiased Estimation).

The classical assumption tests performed were the residual normality test (Kolmogorov-Smirnov test), non-multicollinearity test (Variation Inflation Factor value), non-autocorrelation test (Durbin-Watson test), and homoscedasticity test (Breush-Pagan test). The measure used for the normality and homoscedasticity test is the p-value, if it is greater than the study 0.05 (p-value > 0.05), it is concluded that the residual model has met the assumption test. While the non-multicollinearity test uses the Variation Inflation Factor (VIF) measure, if the value is less than 10, it can be concluded that all independent variables in the model are free from multicollinearity disorders (Nursiyono and Pray, 2016).

Although the regression model that has been formed has been significant both partially and simultaneously, often the OLS estimation method has not produced a good model estimate. According to Sari and Bustami (2021), the OLS method is not appropriate for data containing outliers. According to Nursiyono and Pray (2016), outliers cause the residual variance to be higher. The following is a box plot to see the presence of outliers:

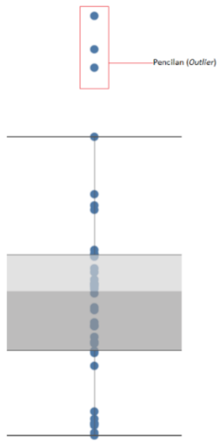

Figure 1. Detection of Data Outlier
(Source: SIPSN, BPS Jatim, Proceeded Using Tableau)

Due to the outliers in the data used in this study and causing the heteroscedasticity of the

OLS model to be disturbed, the modeling of this study also uses robust regression (robust regression).

According to Ryan (1997) in Edriani (dkk), robust regression is an alternative to the OLS method that is used when the distribution of the residuals is not normal and/or there are several indices that affect the model. To see the significance as well as the effect of outliers in the model, this study uses a difference in fits (DFFITS) measure, namely by using a threshold $2\sqrt{p/n}$, with p is the number of independent variables in the model, and n is the number of observations in the model. A model is said to contain significant and influential outliers in the model if the value $DFFITS > 2\sqrt{p/n}$ (Zach, 2020).

The estimation method for the robust regression parameter of this research is MM-regression. According to Utomo, et al. (2014), MM regression is a method of merger between the M- regression and S- regression that are resistant and have higher efficiency.

Model Selection Criteria

In this study, two models were obtained, namely the Ordinary Least Square (OLS) regression model and the robust regression model. To get the best model, this study uses several statistical measures, namely the smallest Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) as an indicator of the goodness of the model as a predictive tool, the smallest RSE (specifically the robust model seen from the standard robust residual size). error) to see how much accuracy the model has with the actual observed value, and the largest adjusted R square value $\left(R^2_{Adj}\right)$ as a measure of the strength of the independent variables included in the model to explain the large proportion of the diversity of the dependent variables.

3. RESULTS

Descriptive Analysis

Tabel 1. Descriptive Analysis of Research Variables

Description	Total Residents	Real Expenditure per Capita per Year	Amount Waste Bank
Average (mean)	1.070.150	11.569	50.74
Maximum	2.874.314	17.862	236.00
Minimum	132.434	8.673	0.00

(Source: Data Proceeded by R)

Descriptively, this research variable can be seen in Table 1. The population variable has an average value of 1,070,150 people with a minimum value of 132,434 people and a maximum value of 2,874,314 people. The real expenditure variable per capita per year is 11,569 thousands of rupiah with a minimum value of 8,673 thousands of rupiah and a maximum value of 17,862 thousands of rupiah. The variable number of garbage bins has an average value of 50.74 units with a minimum value of 0 units and a maximum value of 236 units.

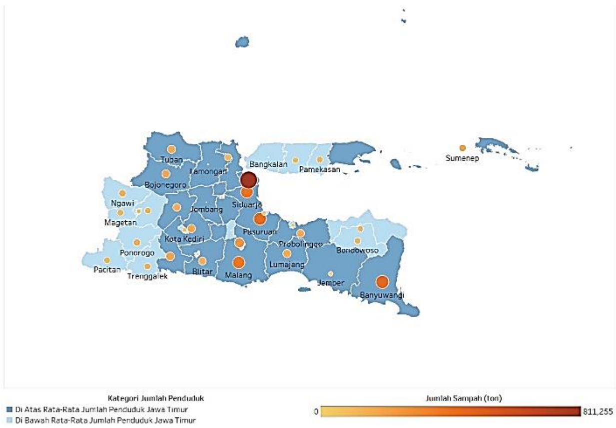

Figure 2. Map of Population Distribution and The Amount of Waste in 2020 in East Java
(Source: SIPSN, BPS Jatim Proceeded Using Tableau)

Figure 2 shows a map of the distribution of the population and the amount of waste in East Java in 2020. Districts/cities whose population is above the average population of East Java have a tendency of increasement in the amount of waste produced. The city of Surabaya has a population above the average population of East Java and has the largest amount of waste.

Figure 3 shows that districts/cities in quadrant I are districts/cities with a high level of real expenditure per capita per year and also have many waste banks. This means that in addition to high expenditure or consumption, the district/city also pays attention to good waste management, for example, Surabaya City is the best city in terms of waste handling, as evidenced by the determination of Surabaya City as a role model in urban waste management in 2018 by the Ministry of Environment and Forestry.

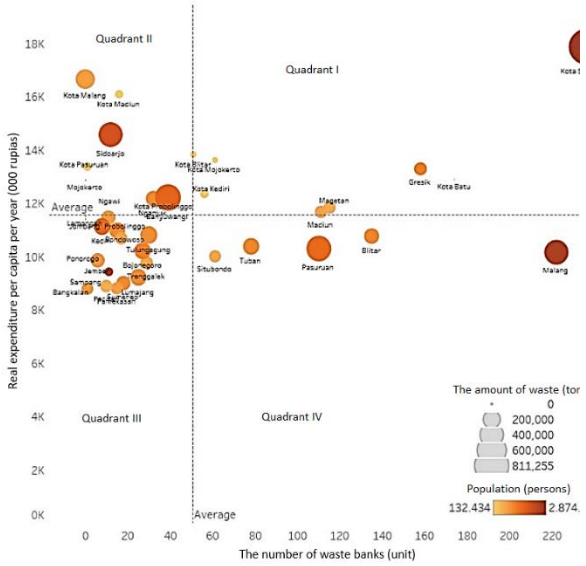

Figure 3. Quadrant Analysis of Real Expenditure Per Capita Per Year and The Number of Waste Banks In 2020
(Source: SIPSN, Proceeded Using Tableau)

Inferential Analysis

The inferential analysis used in this study is a multiple linear regression model with the Ordinary Least Square (OLS) parameter estimation method. Based on Table 2, the regression model obtained is as follows:

$$yt = -0,097 + 0,519\ Pddk_{1i} + 0,231\ Inc_{2i} + 0\ ,155\ Jbs_{3i}$$

Tabel 2. OLS Regression Estimation Results

	Estimate	Std.Error	T-value	Pr(>|t|)
intercept	-0,097	0,044	-2,192	0,035
Pddk	0,519	0,089	5,825	1,45e-06
Inc	0,231	0,088	2,601	0,0137
Jbs	0,155	0,084	1,833	0,0756
F-statistic	20,05	DF	3 dan 34	1,164e-07

(Source: Data Proceeded by R)

The results of the partial test of the population and real expenditure per capita per year were significant in the 0.05 study. For the variable the number of waste banks is not significant and the sign should reverse direction with the variable amount of waste. While the results of the simultaneous test p-value = 1.164e-07 <0.05, this means that the independent variable significantly affects the dependent variable. With an R-Square value of 63.89 percent, it means that the model's ability to explain the variability of data on the amount of waste in East Java is 63.89 percent, and the rest is explained by other variables not included in the model.

Assumption Tests

1. Normality Test

The normality test in this study used the Kolmogorov Smirnov test. The normality test results p-value = 1 is greater than the 0.05 study, this means that the residual regression results have followed a normal distribution.

2. Multicolinearity Test

The VIF value of the three independent variables is less than 10, meaning that there is no multicollinearity between the independent variables in the model.

VIF Pddk =1.13; VIF Inc=1.09; VIF Jbs=1.22

3. Homoskedasticity Test

The results of the Breush-Pagan test p-value = 0.0093 <0.05, thus rejecting H0 at a significance level of 5 percent. This means that there is heteroscedasticity in the resulting OLS regression model. Due to the presence of heteroscedasticity disorders, this study will continue the analysis using Robust Regression.

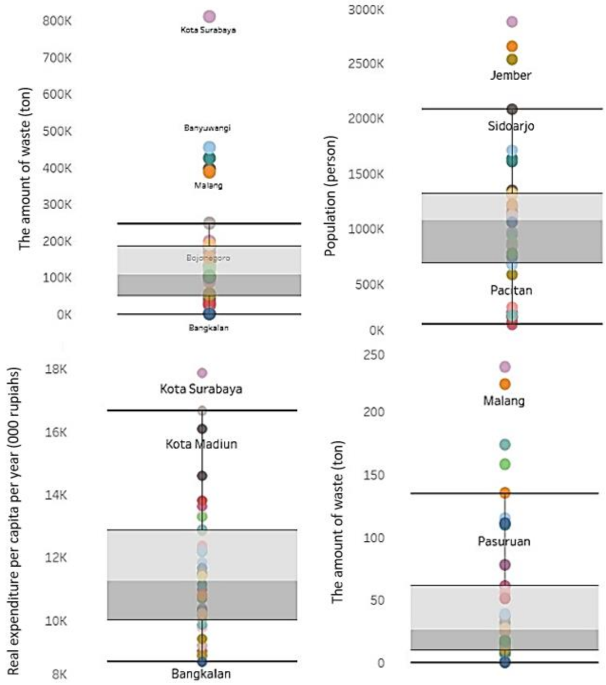

Figure 4. Box Plot of Research Variables
(Source: SIPSN, BPS Jatim, Proceeded Using Tableau)

Figure 4 shows that in this study the variables used were outliers. Outlier data contained in this research variable such as the City of Surabaya has a real per capita annual expenditure and the amount of waste exceeds twice the standard deviation of the data distribution. Malang Regency has the number of waste banks exceedingly twice the standard deviation of the data distribution of the number of waste banks. Jember Regency has a population exceedingly twice the standard deviation of the population data distribution.

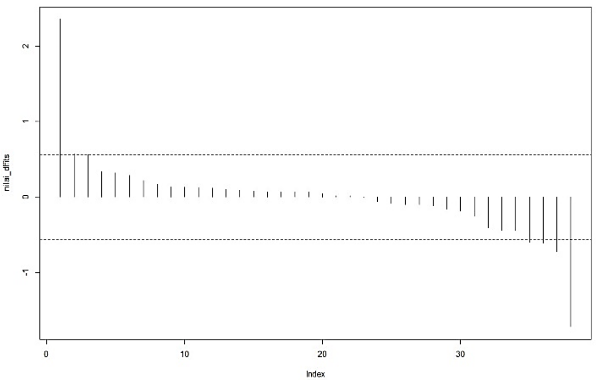
Figure 5. Visualization Of DFFITS Test Results
(Source: Data Proceeded by R)

To see the significance as well as the effect of outliers in the model, this study used the difference in fits (DFFITS) measure . The visualization of the DFFITS test results can be seen in Figure 5. Because in this study there are outliers that have an important role so that they need to be accommodated in the model using Robust regression.

The horizontal line is a boundary line (threshold) as an indicator of the existence of outliers and how important the outliers are in the model, if they cross that line the outliers are important and need to be accommodated.

Tabel 3. Robust Regression Estimation Result

	Estimate	Std.Error	T-value	Pr(>|t|)
intercept	-0,017	0,023	-0,723	0,474
Pddk	0,561	0,038	14,767	2,36e-16
Inc	0,124	0,055	2,252	0,0309
Jbs	-0,067	0,030	-2,225	0,0328

(Source: Data Proceeded by R)

Based on Table 3, the resulting Robust Regression model is as follows:

$$yt = -0{,}017 + 0{,}561\, Pddk_{1i} + 0{,}124\, Inc_{2i} - 0\,{,}067\, Jbs_{3i}$$

The results of the partial and simultaneous tests were significant in the 0.05 study. The R-square value of the robust regression model is 89.09 percent, meaning that the model's ability to explain the variability of data on the amount of waste in East Java is 89.09 percent, and the rest is explained by other variables not included in the model.

Best Model Selection

To find out the best model among the OLS Regression and Robust Regression models, the size of the model selection can be seen from the MAPE, RMSE, RSE and R 2 values. Table 4 shows that both the MAPE, RMSE, RSE values of the Robust Regression model are smaller than OLS, while the R 2 value is greater than the OLS Regression. Therefore, it can be concluded that the best model is the Robust Regression model.

Tabel 4. Comparison of OLS Regression and Robust Regression Model

	MAPE	RMSE	RSE	R^2
OLS	4,299	0,246	0,124	63,89
Robust	2,048	0,159	0,045	89,09

(Source: Data Proceeded by R)

Model Interpretation

1. If the population increases by 1 unit, the amount of waste will increase by 0.561 units assuming other factors are considered constant.
2. If real expenditure per capita per year increases by 1 unit, the amount of waste will increase by 0.124 units assuming other factors are considered constant.
3. If the waste bank increases by 1 unit, the amount of waste will decrease by 0.067 units assuming other factors are considered constant.

4. CONCLUSION

Regression models were chosen in this study is a robust regression, with an excess of the value of R2 is greater and the value of MAPE, RMSE, and RSE is smaller than the OLS regression. Robust regression can also overcome the presence of heteroscedasticity disorders. In addition, with Robust Regression the sign/direction of the correlation between the independent variable and the dependent variable is in accordance with the proper theory. The partial and simultaneous test results show that the variables of population, real expenditure per capita per year and the number of waste banks significantly affect the variable amount of waste in East Java in 2020.

The recommendation for further research based on the results of this study is the use of spatial models as an alternative to see the influence of population, real per capita expenditure per year, and number of waste banks on the amount of waste by region.

5. REFERENCES

AdminDLH. (2019, Mei 23). berita. Retrieved from dlh.jatimprov.go.id: http://dlh.jatimprov.go.id/berita-penanganan-dan-pengendalian-sampah-regional-di-jawa-timur.html

Anderha, R., & Maskar, S. (2021). Pengaruh

Kemampuan Numerasi dalam Menyelesaikan Masalah Matematika terhadap Prestasi Belajar Mahasiswa Pendidikan Matematika. Jurnal Ilmiah Matematika Realistik, Vol. 2 (1), 1-10.

Edriani, T., Rahmadania, A., & Noora, D. (2021). Analisis Hubungan Kepadatan Penduduk dengan Pola Penyebaran COVID-19 Provinsi DKI Jakarta menggunakan Regresi Robust. Indonesian Journal Of Applied Mathematics, Vol. 1 (2), 51-60.

Mandasari, D., Wirjodirdjo, B., & Anityasari, M. (2020). Peningkatan Fasilitas Bank Sampah sebagai. Jurnal Teknik ITS, Vol. 9(2), F322-F327.

Nursiyono, J., & Pray, H. N. (2016). Setetes Ilmu Regresi Linier. Malang: Media Nusa Creative.

Prajati, G., Padmi, T., & Rahardyan, B. (2015). Pengaruh Faktor-faktor Ekonomi dan Kependudukan terhadap Timbulan Sampah di Ibu Kota Provinsi Jawa dan Sumatera. Jurnal Teknik Lingkungan, Vol. 21 (1), 39-47.

Rizaty, M. A. (2021, Juli 29). datapublish. Retrieved from databoks.katadata.co.id: https://databoks.katadata.co.id/datapublish/2021/07/29/mayoritas-sampah-nasional-dari-aktivitas-rumah-tangga-pada-2020

Sari, R., & Bustami. (2021). Penerapan Regresi Komponen Utama Robust S-estimator untuk Analisis Pengangguran di Kota Dumai. Karya Ilmiah Universitas Riau, 1-12.

Suryani, A. S. (2014). Peran Bank Sampah dalam Eektivitas Pengelolaan Sampah (Studi Kasus Bank Sampah Malang). Aspirasi: Jurnal Masalah-masalah Sosial, 71-84.

Utomo, A., Sumartini, N. P., Siregar, A., Nagari, N., & Triyaningsih, S. (2014). Regresi Robust untuk Memodelkan Pendapatan Usaha Industri Makanan dan Non-Makloon Berskala Mikro dan Kecil di Jawa Barat. Jurnal Matematika, Sains, dan teknologi, Vol. 15 (2), 63-74.

Wardhani, M., & Harto, A. D. (2018). Komparasi Pengurangan Timbulan Sampah Berbasis Masyarakat menggunakan Prinsip Bank Sampah di Surabaya, Gresik, dan Sidoarjo. Journal Ilmiah Universitas Trunojoyo, Vol. 11 (1), 52-63.

Zach. (2020, Desember 13). How to Calculate DFFITS in R. Retrieved from statology.org: https://www.statology.org/dffits-in-r/