TOTAL STUDENT EXPENDITURE ANALYSIS BASED ON INCOME AND MONTHLY NEEDS USING LOGISTIC REGRESSION

F H Nasidin [1], L Louisa [2]
Actuarial Science. School of Business, President University, Indonesia
[1]*fauzan.nasidin@student.president.ac.id*
[2]*lovena.louisa@student.president.ac.id*

ABSTRACT
A binary logistic regression analysis was performed to analyse the most impacting variables on total student expenditure from a random sample of college student data gathered from a private survey done by the authors. The specific target of this research is to know whether students are prodigal based on Jim Rohn's budgeting strategy. The variables used as predictors are: a) gender, b) food, c) fashion, d) credit, e) hangout, f) assignment, g) electricity, h) laundry, i) training, j) others. The result of logistic regression analysis shows that a model with predictors of food, training, electricity, and others gives the best result with others expenses as the most significant factors. The data shows that females are more likely to shop more than males, while male students spend more money to hangout. Overall expenditure shows male are more prodigal than females. it is correctly classified in 76% of the cases. The key finding of this research is that the selected variables have significant effect on total student expenditure.

Keywords: *Expenditure, Jim Rohn's Budgeting Strategy, Logistic Regression, Prodigal*.

1. Introduction

Consumption is an activity of purchasing goods and services to meet the needs of everyday life. There are two types of consumption, food consumption and non-food consumption. Food consumption is an income expenditure for purchasing the food needs, while non-food consumption is an income expenditure for purchasing the needs outside food. According to Keynes theory, most people's expenditure is exposed by disposable income. Disposable income is an income remaining after decreased by tax. Otherwise, there are expenditures that do not depend on income. In other words, this expenditure will exist even if no income is received such as credit, expected standard of life and economic expectation. This type of expenditure is named autonomous consumption.

Most of the students are not having a permanent income since their income is depending on pocket money given by their parents. Instead, several students earn additional income by doing part-time jobs in their spare time. This will certainly affect the consumption pattern carried out by students. This pattern will influence the total expenditure they consume monthly. A proper expenditure planning is needed so that the income earned by a student can properly be distributed.

Various study has proven that there is a strong relationship between total income and the consumption pattern of the student. Sari (2019) analyse the relationship between the pocket money earned by student and their consumption pattern using simple linear regression. The result prove that 54,1% student consumption is affected by their pocket money, while the other 45.9% are effected by unexamined variables. Student expenditure are also affected by gender, class where they study and scholarship acceptance status. (Tama, 2014)

In general, students spent their income on food, laundry, fashion, internet package, travel, coursework and course training. To find out which variable has the most influence on total student expenditure, the researcher will create a logistic regression model based on the problem above. These models will later be compared to find out the best plan used for student expenditure.

2. Literature Review

2.1 Binomial Distribution and Logistic Regression

Define y_i variable is the variable of response with binary outcome (zero or one).

$$y_i = \begin{cases} 1 \\ 0 \end{cases} \quad (1)$$

y_i have value 1 if "success" on subject i and have value 0 if "fail" on subject i.

y_i is the realization of the random variable Y_i. The probability of $Y_i$is

$$P(Y_i = 1) = \pi_i \qquad P(Y_i = 0) = 1 - \pi_i \quad (2)$$

Y_i use Bernoulli distribution with parameter π_i and can be write with formula:

$$P(Y_i = y_i) = \pi_i^{y_i}(1-\pi_i)^{1-y_i} \quad (3)$$

Expectation value and variance respectively is

$$E(Y_i) = \mu_i = \pi_i \qquad Var(Y_i) = \pi_i(1-\pi_i) \quad (4)$$

In regression analysis, parameter π_i value is influenced by the variable X

$$\pi = \pi_i(X_i) \quad (5)$$

The variable X_i is called the independent variable (predictor) in the i-th subject. Since the mean and variance are dependent on the value of π_i, the linear model cannot be used. The linear model assumes that the predictors affect the fixed (same) variance. This condition is not met in the binary data response.

Suppose the variable X is a factor that can be classified into k groups, $i = 1, 2, ..., k$The effect of factor X on the value of π_i will be analysed. Individuals or subjects who are located in one group have the same X value. n_i represents the total observations in group *i* and y_i represents the total "successes" in group i, so $y_i = 0, 1, 2, ..., n_i$. If each group observations are independent the probability of "success" is π_i, then Yi has a binomial distribution.

$$P(Y_i = y_i) = \binom{n_i}{y_i}\pi_i^{y_i}(1-\pi_i)^{n_i - y_i} \quad (6)$$

Mean and variance Y_i is

$$E(Y_i) = \mu_i = n_i\pi_i \qquad Var(Y_i) = n_i\pi_i(1-\pi_i) \quad (7)$$

2.2 Regression Logistic Model with Single Independent Variable

Based on bivariate data (X, Y) where X is the variable of predict and Y is the binary variable of response, $\pi(x)$, $\pi(x)$expresses the probability of "success" at the value x so that $\pi(x)$ is a parameter in the binomial distribution. Logistic regression can be defined as a function:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} \quad (8)$$

The logit of this probability is a linear function:

$$\text{Logit } [\pi(x)] = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x \quad (9)$$

The logit transformation is the natural logarithm of the odds value. From this equation, logistic regression indicates that:

A. For $\beta_1 > 0$, an increase of one unit x has the effect on an increase in the logit value. If $x \to \infty$ then $\pi(x) \to 1$ and if $x \to -\infty$ then $\pi(x) \to 0$.

B. For $\beta_1 < 0$, an increase of one unit x the effect on a decrease in the logit value. When $x \to \infty$, then $\pi(x) \to 0$. and when $x \to -\infty$, then $\pi(x) \to 1$.

C. There is a linear relationship between $\log \frac{\pi(x)}{1-\pi(x)}$ and variable X. If for each value of X there are enough observations (measurement repetition), then a scatter diagram can be made between the value $\log \frac{\pi(x)}{1-\pi(x)}$ against the variable X to see the linear relationship pattern which itself applies x = 0 for $\log \frac{\pi(x)}{1-\pi(x)} = \beta_0$ and x = 1 for $\log \frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1$.

2.3 Logistic Regression Model with Many Independent Variables

Logistic regression is a type of regression that uses two different values to describe the response variable (Y). It commonly used 0 as fail, and 1 as success. The distribution function used is the logistic distribution with the notation $\pi(x)$ to express the conditional mean of Y given the covariate vector $X = (x_1, x_2, \ldots, x_p)^T$. The formula of logistic regression is as follows

$$\pi(x) = \frac{\exp(X^T\beta)}{1+\exp(X^T\beta)} \text{ with } X^T\beta = \beta_0 + x_1\beta_1 + \ldots + x_p\beta_p \tag{10}$$

$\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ is a parameter vector. A logit $\pi(x)$ transformation is defined as:

$$g(X) = \log \frac{\pi(X)}{1-\pi(X)} = X^T\beta \tag{11}$$

Therefore, g (X) is linear in the β parameter

2.4 Maximum Likelihood Estimator for Logistic Regression

Suppose a sample consists of n observations of the pair (X_i, Y_i) with $i = 1, 2, \ldots, n$. The formula of logistic regression model is as follow:

$$\pi(X_i) = \frac{\exp(X_i^T\beta)}{1+\exp(X_i^T\beta)} \tag{12}$$

To determine the regression model, β is estimated using the maximum likelihood method at first. The log likelihood function is:

$$\log L(\beta) = \sum_{i=1}^{n} \{ y_i \log(\pi_i) + (n_i - y_i)\log(1-\pi_i) \} \tag{13}$$

2.5 Logistic Regression Inference

2.5.1 Confidence Interval

If the sample used in the observation is large, the confidence interval of β_j in the logistic regression model logit is $[\pi(x)] = \beta_0 + x_1\beta_1 + \ldots + x_p\beta_p$ is $\hat{\beta}_j \pm Z_{\alpha/2}\sqrt{Var(\hat{\beta}_j)}$ for j = 0, 1, …, p. (14)

2.5.2. Significance Test

To test the hypothesis $H_0: \beta_j = 0$ on large samples, statistics test can be used:

$$Z = \frac{\hat{\beta}_j}{\sqrt{Var(\hat{\beta}_j)}} \tag{15}$$

The Z statistic is standard normal distribution.

$$Z^2 = \left(\frac{\hat{\beta}_j}{\sqrt{Var(\hat{\beta}_j)}} \right)^2 \tag{16}$$

The Z^2 statistic has a Chi-Square distribution with df = 1. This Z statistic is also called the Wald test statistic.

2.5.3. Model Fit Test

Suppose we want to test whether the model fits the data and we want to test how much the fit is then a measure of deviation can be used. The deviation statistic (D) measures the discrepancy between the observed value and the value predicted by the model.

$$D = -2\sum_{i=1}^{k}\left\{ y_i \log\left(\frac{n_i\hat{\pi}_i}{y_i}\right) + (n_i - y_i)\log\left(\frac{n_i - n_i\hat{\pi}_i}{n_i - y_i}\right)\right\} \tag{17}$$

or

$$D = 2\sum_{i=1}^{k}\left\{ y_i \log\left(\frac{y_i}{n_i\hat{\pi}_i}\right) + (n_i - y_i)\log\left(\frac{n_i - y_i}{n_i - n_i\hat{\pi}_i}\right)\right\} \tag{18}$$

3. Method

This research used primary data gained from the author's private survey on 63 respondents of student college from various universities around the world. The questionnaire is self-administered, asking about their total income and total expenditure based on these variables; a) gender, b) food, c) fashion, d) credit, e) hangout, f) assignment, g) electricity, h) laundry, i) training, j) others. The result of this questionnaire was analyzed using logistic regression in R to get which variable gives the most significant impact to their expenditure.

Since most logistic regression produces a binary result, it is necessary to define a binary outcome on this data. The binary result is determined based on the ratio of total expenditure compared to the income. using 70-10-10-10 budgeting strategy by Jim Rohn, it can be concluded as 0 for total expenditure less than 70% of income and 1 for total expenditure more than 70% of income. Table below shows head of data afterward

Table 1. Head of The Analyzed Data

Income	Expenditure	Binary	Gender	Fashion	Food	Credit	Hangout	Assignment	Training	Laundry	Electricity	Others
2500000	1950000	1	M	100000	600000	100000	300000	100000	400000	50000	200000	100000
2000000	1210000	0	F	100000	300000	70000	200000	70000	200000	20000	200000	50000
2500000	1510000	0	M	50000	300000	150000	300000	10000	250000	50000	250000	150000
2000000	1615000	1	M	100000	750000	50000	100000	20000	250000	75000	250000	20000
1800000	1394000	1	F	250000	450000	25000	150000	15000	200000	54000	50000	200000

This binary logistic regression procedure in R was used to perform the analysis to determine whether likelihood of total expenditure could be predicted from the independent variables on this data. The research used data from 63 student colleges around the world.

4. Results and Discussion

The binary logistics regression analysis was performed in R with procedures as below:

4.1 Data Preparation

Data is named by dataset, and *str* command will show this data in a string type.

```
> str(dataset)
tibble [63 x 15] (S3: tbl_df/tbl/data.frame)
 $ Name          : chr [1:63] "Jason Nathanael" "Veldelen Yaphira" "Mika Daniel Nainggolan" "Didik Wibowo" ...
 $ Gender        : chr [1:63] "Male" "Female" "Male" "Male" ...
 $ Income Total  : num [1:63] 2500000 2000000 2500000 2000000 1800000 2000000 3000000 2800000 4000000 2500000 ...
 $ Expense Total : num [1:63] 1950000 1210000 1510000 1615000 1394000 ...
 $ Food          : num [1:63] 20000 10000 10000 25000 15000 15000 20000 20000 30000 25000 ...
 $ Fashion       : num [1:63] 100000 100000 50000 100000 250000 100000 200000 100000 200000 250000 ...
 $ Credit        : num [1:63] 100000 70000 150000 50000 25000 75000 100000 50000 150000 150000 ...
 $ Hangout       : num [1:63] 300000 200000 300000 100000 150000 200000 200000 300000 1000000 25000 ...
 $ Assignment    : num [1:63] 100000 70000 10000 20000 15000 100000 35000 30000 100000 100000 ...
 $ Training      : num [1:63] 400000 200000 250000 250000 200000 90000 200000 300000 350000 90000 ...
 $ Electricity   : num [1:63] 200000 200000 250000 250000 50000 200000 100000 50000 200000 200000 ...
 $ Laundry       : num [1:63] 50000 20000 50000 75000 54000 24000 53000 40000 32000 12000 ...
 $ Others        : num [1:63] 100000 50000 150000 20000 200000 50000 300000 100000 100000 200000 ...
 $ Food (Monthly): num [1:63] 600000 300000 300000 750000 450000 450000 600000 600000 900000 750000 ...
 $ Binary        : num [1:63] 1 0 0 1 1 0 0 0 1 1 ...
```

Figure 1. Data String

4.2 Binary Logistic Regression Model

A binary logistic regression model was created by using *glm* command in R. the first model will include all the independent variables to measure its correlation with the dependent variable Binary. A variable that gives significant effect to the model has *p-value* lower than 0.05. First model is named model_logit1.

model_logit1=glm(Binary~Gender+Food+Fashion+Credit+Hangout+Assignment+Training+Electricity+Laundry+Others,data=dataset,family=binomial(link="logit"))

In the first model, there are two variables that give significant effect to the model, which is "Training" and "Others". The other independent variables are still not significant. Then, we need to build a new model by removing those insignificant variables. In model 2, "Credit" is erased because it has the biggest *p-value*. The equation will be:

model_logit2=glm(Binary~Gender+Food+Fashion+Hangout+Assignment+Training+Electricity+Laundry+Others,data=dataset,family=binomial(link="logit"))

The summary of model 1 is:

Table 2. Regression Analysis Summary from Model 1

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.463e+00	1.949e+00	-2.803	0.00506
GenderMale	1.384e+00	9.111e-01	1.519	0.12885
Food	4.076e-05	3.362e-05	1.213	0.22532
Fashion	1.828e-06	2.114e-06	0.864	0.38736
Credit	-1.838e-06	6.587e-06	-0.279	0.78021
Hangout	-1.566e-06	1.597e-06	-0.981	0.32671
Assignment	8.748e-06	7.701e-06	1.136	0.25599
Training*	6.223e-06	3.062e-06	2.033	0.04209
Electricity'	9.851e-06	5.318e-06	1.852	0.06396
Laundry	-7.539e-06	9.980e-06	-0.755	0.44999

The summary of model 2 is:

Table 3. Regression Analysis Summary from Model 2

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.557e+00	1.931e+00	-2.879	0.00399
GenderMale	1.333e+00	8.969e-01	1.487	0.13710
Food	3.919e-05	3.304e-05	1.186	0.23558
Fashion	1.806e-06	2.111e-06	0.855	0.39229
Hangout	-1.644e-06	1.573e-06	-1.045	0.29596
Assignment	8.387e-06	7.575e-06	1.107	0.26822
Training*	6.365e-06	3.041e-06	2.093	0.03634
Electricity'	9.773e-06	5.304e-06	1.843	0.06539
Laundry	-7.322e-06	9.982e-06	-0.734	0.46324
Others*	6.507e-06	2.686e-06	2.423	0.01539

The second model still gives us the same output with the first model with a little decrease in *p-value* of all variables. This model is still not efficient. In the next model, we will remove laundry from the model for its big *p-value.* The equation for model 3 as follow:

model_logit3=glm(Binary~Gender+Food+Fashion+Hangout+Assignment+Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The significance of Gender is increasing in the third model, while "Fashion" holds the biggest *p-value*. Otherwise, "Gender" will be removed in the next chapter. If we still conclude "Gender" to the next model, it will decrease the significance of the other variables.

model_logit4=glm(Binary~Food+Fashion+Hangout+Assignment+Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The summary of model 3 is:

Table 4. Regression Analysis Summary from Model 3

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.450e+00	1.901e+00	-2.867	0.00414
GenderMale'	1.509e+00	8.663e-01	1.742	0.08152
Food	3.571e-05	3.291e-05	1.085	0.27787
Fashion	1.769e-06	2.113e-06	0.837	0.40235
Hangout	1.655e-06	1.556e-06	-1.064	0.28755
Assignment	8.486e-06	7.489e-06	1.133	0.25716
Training*	5.332e-06	2.622e-06	2.034	0.04199
Electricity'	8.795e-06	5.073e-06	1.734	0.08299
Others*	6.084e-06	2.638e-06	2.306	0.02110

The summary of this model is:

Table 5. Regression Analysis Summary from Model 4

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.246e+00	1.866e+00	-2.811	0.00493
Food	-1.458e-06	1.445e-06	-1.009	0.31316
Fashion	6.493e-06	7.221e-06	0.899	0.36855
Hangout	3.910e-05	3.268e-05	1.196	0.23154
Assignment	7.628e-07	1.978e-06	0.386	0.69979
Training*	5.698e-06	2.521e-06	2.260	0.02383
Electricity*	1.102e-05	5.052e-06	2.182	0.02912
Others*	5.771e-06	2.327e-06	2.480	0.01316

From the picture below, “Electricity” gains its significance in model four and we have a total of 3 significant factors to the model. But, it stills not the most efficient model since it has several insignificant factors. We will remove “Hangout” in the next model.

model_logit5=glm(Binary~Food+Fashion+Assignment+Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The summary of model 5 is:

Table 6. Regression Analysis Summary from Model 5

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.411e+00	1.844e+00	-2.934	0.00335
Food	3.863e-05	3.178e-05	1.216	0.22417
Fashion	5.394e-07	1.932e-06	0.279	0.78014
Assignment	5.735e-06	7.154e-06	0.802	0.42280
Training*	5.191e-06	2.402e-06	2.161	0.03068
Electricity*	1.056e-05	4.945e-06	2.135	0.03273
Others*	5.962e-06	2.367e-06	2.519	0.01178

There is no significant change in model 5. We still need to remove several variables that do not give a significant effect to the model. As for that, “Fashion” will be erased from the model.

model_logit6=glm(Binary~Food+Assignment+Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The summary of model 6 is:

Table 7. Regression Analysis Summary from Model 6

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-5.265e+00	1.762e+00	-2.989	0.0028
Food	5.384e-06	7.037e-06	0.765	0.4442
Assignment	4.096e-05	3.069e-05	1.335	0.1820
Training*	5.093e-06	2.387e-06	2.133	0.0329
Electricity*	1.044e-05	4.915e-06	2.125	0.0336
Others*	6.015e-06	2.357e-06	2.552	0.0107

The picture above shows that “Assignment” and “Food” are the only variables that are insignificant, while other variables such as “Training”, “Electricity” and “Others” are significant factors. Then, we will remove “Assignment” from the next model since it has higher *p-value* than “Food”.

model_logit7=glm(Binary~Food+Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The summary of model 7 is:

Table 8. Regression Analysis Summary from Model 7

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-4.751e+00	1.583e+00	-3.001	0.00269
Food	4.392e-05	3.079e-05	1.427	0.15368
Training*	4.947e-06	2.393e-06	2.067	0.03870
Electricity*	9.542e-06	4.606e-06	2.072	0.03831
Others*	5.647e-06	2.274e-06	2.483	0.01303

This model is nearly efficient because most of the variables are significant. Then, we will erase the “Food” factor from the model.

model_logit8=glm(Binary~Training+Electricity+Others,data=dataset,family=binomial(link="logit"))

The summary of model 8 is:

Table 9. Regression Analysis Summary from Model 8

Coefficients	Estimate	Std. Error	z value	Pr (>|z|)
(intercept)**	-3.645e+00	1.300e+00	-2.804	0.00505
Training*	5.242e-06	2.480e-06	2.114	0.03455
Electricity*	9.648e-06	4.467e-06	2.160	0.03077
Others*	5.172e-06	2.114e-06	2.447	0.01440

Since all the variables in this model are significant, we can conclude that this is the most efficient model. But, it is necessary to check its error by comparing the AIC value from all these models. The AIC value of each model can be seen in the table below:

Table 10. AIC Table from Model 1 to Model 8

No	Model	AIC
1	model_logit1	83.23
2	model_logit2	81.31
3	model_logit3	79.85
4	model_logit4	81.32
5	model_logit5	79.47
6	model_logit6	78.19
7	model_logit7	77.02
8	model_logit8	77.24

Even though model_logit8 is the most efficient model, with all variable in the model is significant, but model_logit7 appears to have the lowest AIC value. In this case, we need to do Hosmer-Lemeshow Goodness of Fit test to check whether model_logit8 is the proper model to analyse this data. The result of this test is:

$$X\text{-squared} = 9.0999,\ df = 8,\ p\text{-value} = 0.3339$$

Because its *p-value* is higher than 0.05, we can conclude that model_logit8 is the best model to analyse our data. This model will be taken to the further test.

7.2 Best Model Selection

7.2.1 Goodness of Fit Test

In goodness of fit tests, deviances are used to measure the significance of each variable in the model and to test whether the model is proper to use. In this goodness of fit test, we will use McFadden Pseudo R^2 with the formula $R^2 = 1 - \frac{I(\hat{\beta})}{I(\bar{y})}$ where $I(\hat{\beta})$ is log likelihood value form fitted model and $I(\bar{y})$ is logit model that only has constant component.

The value of this test is -0.21, which means that 21% of the variation of the dependent variable can be explained by the formed model.

7.2.2 Odds Ratio

Odds ratio measures the change of one-value in the independent variable will affect the increase on the dependent variable at some certain point based on the odds ratio value. The following table shows the result of odds ratio of each variable

Table 11. Odds Ratio

Intercept	Training	Electricity	Others
0.03	1.00	1.00	1.00

Based on this result, we can interpret that, if the value of training increase by 1, then, the value of total student expenditure will also increase 1.

7.2.3 Predicted Table

After having a fitted model to our data, we could create a prediction. To test whether our model perform well in predicting the data, we could create a confusion matrix to display the success rate of the model prediction against the data. Produces an output as:

Table 12. Predicted Table

True	Pred	
	0	1
0	14	10
1	7	32

The value on (0,0) and (1,1) represent the number of correct prediction. Otherwise, the value on (0,1) and (1,0) represent the number of incorrect prediction. If the correctly-predicted value more than 50% of the total prediction, than we could say that our data is fit to predict our data.

$$\frac{(14+32)}{(14+32+10+7)} = 0.73 \sim 73\%$$

The results show that the model can give correct prediction about 73%.

5. Conclusion and Implications

Model_logit8 is the best model to draw total student expenditure. In creating a budget planning for a month, we recommend student to take on training expense, electricity expense and other expense on their consideration because those expense will have a significant impact to their expenditure. The binary logistic regression model used to plan student expenditure is

Expenditure = -3.64 + 0.0000056*Training + 0.00000965*Electricity + 0.00000517*Others.

References

Mubarok, W. (2018, August 09). Regresi Logistik DENGAN R. Retrieved May 06, 2021, from https://muhammadilhammubarok.wordpress.com/2018/08/09/regresi-logistik-dengan-r/

Nugraha, J. (september 2013). *Pengantar Analisis Data Kategorik* (1st ed.). Yogyakarta, Indonesia: Deepublish.

Sari, N. P. (2019). *Pengaruh Uang Saku Terhadap Pola Konsumsi Mahasiswa Fakultas Ekonomi dan Bisnis Islam IAIN Bengkulu* [IAIN Bengkulu]. http://repository.iainbengkulu.ac.id/3611/1/NONI PURNAMA SARI.pdf

Tama, R. T. (2014). *Pengeluaran Konsimsi Mahasiswa Program Studi Pendidikan Ekonomi Fakultas Ekonomi Universitas Negeri Yogyakarta*. Universitas Negeri Yogyakarta.

TheStatsGeek. (2014, August 25). Hosmer-Lemeshow goodness of fit test in R. Retrieved May 06, 2021, from https://www.youtube.com/watch?v=MYW8gA1EQCQ

Trent Hamm – Founder & Columnist Last Updated: May 9, Hamm, T., & Columnist, F. (2019, May 09). The 70-10-10-10 budgeting model. Retrieved May 06, 2021, from https://www.thesimpledollar.com/save-money/the-70-10-10-10-budgeting-model#