

The web-based estimation of motorcycles sales using linear regression method

by Dedy Prehanto

Submission date: 05-Oct-2021 08:17PM (UTC+0700)

Submission ID: 1665899876

File name: Indriyanti_2021_IOP_Conf._Ser._Mater._Sci._Eng._1098_052115.pdf (764.81K)

Word count: 3106

Character count: 14787

PAPER · OPEN ACCESS

The web-based estimation of motorcycles sales using linear regression method

To cite this article: A D Indriyanti *et al* 2021 *IOP Conf. Ser.: Mater. Sci. Eng.* **1098** 052115

⁶ View the [article online](#) for updates and enhancements.



ECS 240th ECS Meeting
Digital Meeting, Oct 10-14, 2021
We are going fully digital!
Attendees register for free!
REGISTER NOW

The banner features a group of diverse people in professional attire, smiling and engaged in conversation. The background is a bright, modern setting. The text is overlaid on the left side of the image.

This content was downloaded from IP address 110.139.82.196 on 05/10/2021 at 01:01

The web-based estimation of motorcycles sales using linear regression method

A D Indriyanti*, D R Prehanto, I G L P E Prisma and I K D Nuryana

Information Technology, Informatics Engineering, Universitas Negeri Surabaya, 60231, Indonesia

*ariesdwi@unesa.ac.id

Abstract. The purpose of this research is to use the linear regression method to predict motorcycle sales results, the variable used is the period as an independent variable (X) and sales as the dependent variable (Y). The data used in the calculation of linear regression is the last four years data, from January 2014 to December 2019. The implementation of the motorcycle sales forecasting system is to predict sales in the coming months. To find out the level of accuracy of the prediction error calculation is needed so that it is known how many error levels are obtained. Calculation of forecasting errors using Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE). The results of this study are web based motorcycle sales prediction systems using linear regression method. From this system, motorcycle sales forecasting is obtained the following month. In January 2015 with forecasting results of 12.63. To find out how accurate the forecasting level is, the error calculation result using Mean Absolute Deviation (MAD) is 3.40 and Mean Absolute Percentage Error (MAPE) is 44.33%. The results show that the error rate is small and the forecasting results are close to accurate.

1. Introduction

The development of motorcycles in Indonesia has increased from year to year, making many companies compete in offering their products. Each company provides the best advantage of the products offered to consumers, so that the company can win a competitive market with other companies. Many ways are taken by the company to encourage the public's attraction towards the motorbikes offered. One way that is often used is through the variety of promotional media used. Promotion media evaluation system must always be run at the end of each year to find out the role of promotional media towards marketing targets [1].

Likewise, Agung Motor dealers who are the object of this research want motorcycle sales to reach high sales targets to sell motorcycle units. Besides having to think about sales targets that must be achieved every year, the manager must determine the number of motorcycle products that will be in stock each year. Determination of the number of stock of motorcycle products must be in accordance with the sales target to be achieved to avoid unsold products. Products that are not sold at the end of this year will become a new problem in stock inventory management, resulting in product build up. In overcoming this problem an information system is needed that is able to forecast product sales which can be used to prepare strategies for increasing motorcycle sales.

Forecasting is the process of estimating products that will be sold in the future in certain circumstances and made based on data that has occurred within a certain time period [2]. In forecasting



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI

Published under licence by IOP Publishing Ltd

there must be a difference between forecasting results with sales reality, therefore forecasting results can be said to be good if the level of distrust shows a small value [3]. In forecasting, sales data with monthly periods in 2014-2019 are needed to produce the number of forecasting every month going forward. The forecasting process will be more effective if more and more sales data are used and the error result shows a small number [4]. There are many forecasting methods that can be used in this case, including the fuzzy time series method, linear regression, moving average, and other forecasting methods [5]. The method used for forecasting motorcycle sales at Agung Motor dealers is linear regression. Linear regression is a measuring tool used to determine whether there is a relationship between variables [6]. Linear regression is more accurate in knowing the rate of change of a variable against other variables. The variables used are period as independent variable (X) and sale as dependent variable (Y) [7].

The results of this study are in the form of information systems forecasts for motorcycle sales in the coming months based on 2014-2019 sales data. To find out the results of forecasting that are obtained will be calculated using Mean Absolute Deviation (MAD), Mean Absolute Percentage Error (MAPE) [8]. So that the level of distrust will be known from the results of forecasting.

2. Methods

This research focuses on forecasting motorcycle sales using linear regression methods. To do this forecasting there are two variables used, namely the period as an independent variable (X) and sales as the dependent variable (Y) [9]. The data that will be used are sales data for the past four years that come from Agung Motor dealers. The system can predict motorcycle sales for the following year through calculations using linear regression. In a forecasting there will definitely be a difference with the actual data, so to measure the level of accuracy in the forecasting system is done by using some calculation errors in forecasting accuracy, namely the Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) [10]. The user is the admin, admin as the main operator in using the system and can add or change data according to the provisions. Research also requires steps to make it easier to conduct research, and know which needs to be done first. Figure 1 below is a further explanation of the research procedure.

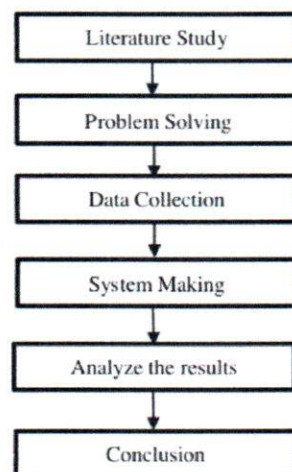


Figure 1. Research procedure.

The picture above explains that the first thing to do in a research procedure is to conduct a literature study to find information contained in online and printed books. Next is the formulation of the problem, namely the stage of identifying the problem, in this thesis the problem to be identified is how to design and implement a linear regression method to determine the forecasting of web-based motorcycle sales. Then to complete the data in this study, it requires the data collection stage where the data obtained from the Agung Motor dealer sales data in Maumere, the data obtained are four-year sales data, namely in January 2014 - December 2019. To complete this study the variables used are the independent variable (X) is the period and for the dependent variable (Y) is the sale [11]. Next is the stage of making a Honda motorcycle sales system, the programming language used in the system is PHP, and for the database used to store data is MySQL. In this system for forecasting the linear regression method is used.

Linear regression is a method that functions to test the extent of the causal relationship between the independent variable (X) and the dependent variable (Y), where the independent variable is the period and the dependent variable is sales [12].

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

$$Y = X\beta + \varepsilon \quad (2)$$

where :

Y = dependent variable (predicted value)

X = the matrix of independent variable ($n \times (p + 1)$)

β = vector of regression model parameters ($(p + 1) \times 1$)

ε = vector of error ($n \times 1$)

The results obtained from this study in the form of a forecasting system for Honda motorcycle sales using linear regression method. The user of this system is the admin as the main user. Admin entered the data for the last four years, from January 2014 to December 2019, then the data was processed into the system and produced a linear regression equation to calculate sales forecasting. After the results of the forecasting come out to determine the level of accuracy of the forecasting results, it is necessary to know the level of error using the Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) [13].

The MAD calculation equation is as follows [14]:

$$\text{MAD} = \frac{\sum |X_t - F_t|}{n} \quad (3)$$

when:

X_t = sales

F_t = result of forecasting

n = amount of Period

The MAPE calculation equation is as follows [15]:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left(\frac{|X_t - F_t|}{X_t} \times 100\% \right) \quad (4)$$

when:

X_t = sales

F_t = result of forecasting

n = amount of period

2.1. Flowchart system

The flow of the application flowchart is first started from the user logging into the system by entering the username and password then the system will process by matching the username and password in the

database, if the username and password match then the system will proceed to the dashboard data page if incorrect then the system will ask to enter the correct username and password to the admin. After successfully logging in the system will directly lead to the dashboard data page and several menus, if there is no menu to be opened the user can log out to exit the application system, aimed at Figure 3 System Flowchart.

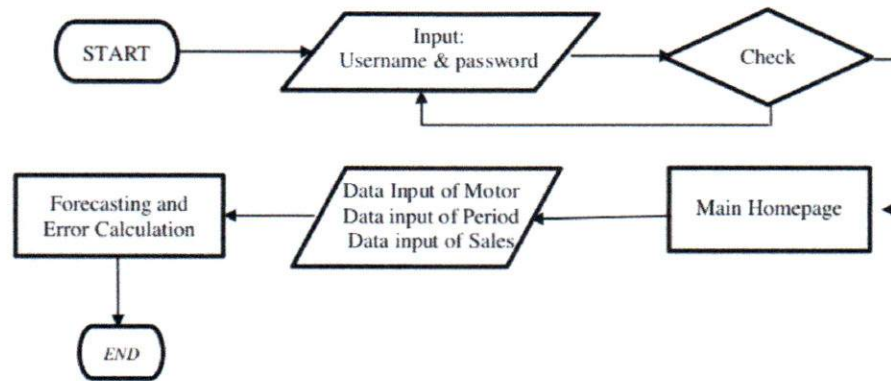


Figure 2. Flowchart of system.

3. Results and discussion

Linear regression is a measuring tool used to determine whether there is a relationship between variables. The following are sales data for all types of Honda motorcycles from January 2014 to December 2019.

Table 1. Data pf sales years 2014-2019.

No.	Type of Motor	Month												Years
		Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agu	Sep	Okt	Nov	Des	
1	VARIO	25	13	24	26	8	5	-	25	8	7	18	15	2014
2	VARIO	8	5	5	20	15	12	13	16	19	24	17	28	2015
3	VARIO	3	1	2	12	2	9	2	5	2	-	4	6	2016
4	VARIO	-	-	-	-	5	1	5	4	3	5	6	13	2019
5	BEAT	21	24	22	24	20	19	19	24	22	26	24	29	2014
6	BEAT	18	12	14	12	7	9	8	13	15	15	16	24	2015
7	BEAT	13	7	4	7	9	2	7	6	5	8	6	10	2016
8	BEAT	-	-	-	-	-	-	-	4	5	5	11	11	2019
9	SUPRA X	23	23	26	18	13	9	8	5	11	17	6	14	2014
10	SUPRA X	10	11	9	10	4	9	6	10	13	6	4	12	2015
11	SUPRA X	10	9	8	6	5	7	13	6	10	6	11	11	2018
12	SUPRA X	7	7	9	9	2	-	3	2	5	6	5	6	2019
13	SCOOPY	8	9	18	9	1	2	6	8	9	12	3	9	2014
14	SCOOPY	11	16	4	8	4	9	3	5	2	5	8	4	2015
15	SCOOPY	8	7	9	12	10	11	16	9	8	5	12	7	2016
16	SCOOPY	18	10	12	15	9	4	7	4	-	-	-	1	2019
17	REVO	9	19	26	6	-	9	10	7	10	2	14	7	2014

Table 1. Cont.

18	REVO	11	6	4	3	5	5	3	10	7	7	5	15	2015
19	REVO	3	1	2	4	4	6	5	2	3	1	2	3	2016
20	REVO	15	9	15	10	11	12	12	7	16	12	9	6	2014
21	MEGAPRO	3	11	1	6	1	2	3	2	3	4	7	1	2016
22	MEGAPRO	1	1	1	-	3	-	2	-	3	1	3	2	2017
23	MEGAPRO	-	1	2	-	-	-	-	-	-	1	-	-	2018
24	VERZA	2	-	1	2	1	3	-	3	-	1	3	2	2018
25	VERZA	2	1	2	4	3	4	-	5	1	3	-	2	2019
26	CB 150 R	2	-	1	2	2	-	4	3	5	1	2	2	2018
27	CB 150 R	4	1	1	3	3	4	2	3	3	5	4	3	2019
Total		235	204	222	228	147	153	157	188	188	185	200	243	

Calculation of forecasting sales of Honda motorcycles with type SUPRA X motorcycles, with sales data from the period 1 - 48 or in January 2011 - December 2014:

$$b = \frac{(n)(\Sigma XY) - (\Sigma X)(\Sigma Y)}{(n)(\Sigma X^2) - (\Sigma X)^2}$$

$$b = \frac{(48)(8537) - (1176)(440)}{(48)(36260) - 1382976}$$

$$b = -0,301$$

To calculate the value of a, it can be calculated by the following formula:

$$a = \bar{Y} - b\bar{X}$$

$$a = 9,17 - (-0,301)(24,5)$$

$$a = 9,17 - (-7,37)$$

$$a = 16,54$$

If the results of the a and b values are known then are entered into the linear regression equation as follows:

$$Y = a + bX$$

$$Y = 16,54 + (-0,301)(X)$$

So, the linear regression equation above is used for forecasting sales of SUPRA X type motorcycles in the following month by entering the X value, i.e. period. This forecast menu serves as a forecast process for sales using the linear regression method.

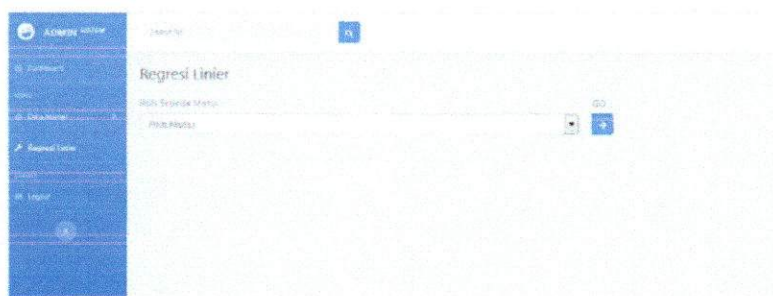


Figure 3. Forecasting page features.

Then after knowing the results of forecasting, error calculation is needed to find out how accurate the results of forecasting are obtained using the following equation.

The MAD calculation equation is as follows:

$$MAD = \frac{\sum |X_t - F_t|}{n}$$

$$MAD = \frac{163,19}{48}$$

$$MAD = 3,40$$

Equation calculation MAPE as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left(\frac{|X_t - F_t|}{X_t} \times 100\% \right)$$

$$MAPE = \frac{1}{48} \sum_{t=1}^n (21,28 \times 100\%)$$

$$MAPE = \frac{2128\%}{48}$$

$$MAPE = 44,33 \%$$

The results of the calculation of errors using Mean Absolute Deviation (MAD) of 3.40 and Mean Absolute Percentage Error (MAPE) of 44.33% of these results indicate that the error rate is small and the results of forecasting close to accurate. This forecast menu serves as a forecast process for sales using the linear regression method.



Figure 4. Sales features.

4. Conclusion

The design of a forecasting system for Honda motorbike sales in the Great Web-based motorcycle using the linear regression method. The design of sales forecasting uses flowcharts, Data Flow Diagrams (DFD), Entity Relationship Diagrams (ERD), Conceptual Data Models (CMD), Physical Data Models (PDM).

Implementation of a forecasting system for Honda motorbike sales in the Great Web-based motorcycle using the linear regression method. From this system we will get the results of motorcycle sales forecasting in the following month. As one example on a beat motorcycle, motor beat data will be processed by the system then a linear regression equation is obtained which is used to calculate motorcycle sales forecasting. In January 2020, the results of forecasting for supra x motorbikes were 12.63 with actual data of 10. To find out how accurate the forecasting results were, the error calculation

using the Mean Absolute Deviation (MAD) was 3.40 and the Mean Absolute Percentage Error (MAPE) of 44.33%. These results indicate that the error rate is small and the forecasting results close to accurate.

References

- [1] Permadi G S, Vitadiar T Z, Kistofer T, & Mujiyanto A H 2019 The Decision Making Trial and Evaluation Laboratory (Dematel) and Analytic Network Process (ANP) for Learning Material Evaluation System *In E3S Web of Conferences* **125** p 23011 EDP Sciences
- [2] Prehanto D R, Indriyanti A D, Mashuri C, & Permadi G S 2019 Soil Moisture Prediction using Fuzzy Time Series and Moisture sensor Technology on Shallot Farming *In E3S Web of Conferences* **125** p 23002 EDP Sciences
- [3] Indriyanti A D, Prehanto D R, Permadi G S, Mashuri C, & Vitadiar T Z 2019 Using Fuzzy Time Series (FTS) and Linear Programming for Production Planning and Planting Pattern Scheduling Red Onion *In E3S Web of Conferences* **125** p 23007 EDP Sciences
- [4] Vitadiar T Z, Farikhin F, & Surarso B 2018 Production Planning and Planting Pattern Scheduling Information System for Horticulture *In E3S Web of Conferences* **31** p 10004 EDP Sciences
- [5] Land K C 2015 Forecasting *In International Encyclopedia of the Social & Behavioral Sciences: Second Edition*
- [6] Permai S D & Tanty H 2018 Linear regression model using bayesian approach for energy performance of residential building *Procedia Computer Science* **135** pp 671-677
- [7] Gang Su X 2009 *Linear regression analysis: Theory and computing*
- [8] Abolghasemi M, Beh E, Tarr G, & Gerlach R 2020 Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion *Computers & Industrial Engineering* 106380
- [9] Huber J & Stuckenschmidt H 2020 Daily retail demand forecasting using machine learning with emphasis on calendric special days *International Journal of Forecasting*
- [10] Lourenço N, Gouveia C M, & Rua A 2020 Forecasting tourism with targeted predictors in a data-rich environment *Economic Modelling*
- [11] Schuld M, Sinayskiy I, & Petruccione F 2016 Prediction by linear regression on a quantum computer *Physical Review A* **94**(2) 022342
- [12] Feng C & Zhang J 2020 Assessment of aggregation strategies for machine-learning based short-term load forecasting *Electric Power Systems Research* **184** 106304
- [13] Yang Y 2018 Prediction and analysis of aero-material consumption based on multivariate linear regression model *In 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* pp 628-632 IEEE
- [14] Lu X & Martin G 2014 An improved rate control algorithm for SVC with optimised MAD prediction *In 2014 IEEE 16th International Workshop on Multimedia Signal Processing (MMSP)* pp 1-6 IEEE
- [15] Eswaran C & Logeswaran R 2010 A comparison of ARIMA, neural network and linear regression models for the prediction of infant mortality rate *In 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation* pp 34-39 IEEE

The web-based estimation of motorcycles sales using linear regression method

ORIGINALITY REPORT

15%

SIMILARITY INDEX

14%

INTERNET SOURCES

13%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Universitas Negeri Surabaya The State University of Surabaya Student Paper	7%
2	scholar.lib.vt.edu Internet Source	1%
3	www.efsa.unsa.ba Internet Source	1%
4	Syarifah Diana Permai, Heruna Tanty. "Linear regression model using bayesian approach for energy performance of residential building", Procedia Computer Science, 2018 Publication	1%
5	repository.radenintan.ac.id Internet Source	1%
6	repository.unair.ac.id Internet Source	1%
7	insightsociety.org Internet Source	1%

The web-based estimation of motorcycles sales using linear regression method

ORIGINAL REPORT

15%

SIMILARITY INDEX

14%

INTERNET SOURCES

13%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES



Submitted to Universitas Negeri Surabaya The State University of Surabaya

Internet source

7%



scholarlib.widya.ac.id

Internet source

1%



www.ets.uns.ac.id

Internet source

1%



Statistik Dalam Rerangka Herings Tancy, Linear regression model using Bayesian approach for energy performance of residential building, Procedia (Computer Science), 2018

Publication

1%



repository.radenintan.ac.id

Internet source

1%



repository.unair.ac.id

Internet source

1%



insightsociety.org

Internet source

1%

8	www.scribd.com Internet Source	<1 %
9	real.mtak.hu Internet Source	<1 %
10	www.hindawi.com Internet Source	<1 %
11	spatial.usc.edu Internet Source	<1 %
12	www.icee.usm.edu Internet Source	<1 %
13	www.inderscience.com Internet Source	<1 %
14	A D Indriyanti, D R Prehanto, I G L E P Prisma, Soeryanto, B Sujatmiko, J Fikandda. "Simple Additive Weighting algorithm to aid administrator decision making of the underprivileged scholarship", Journal of Physics: Conference Series, 2019 Publication	<1 %
15	Lee, Park, Kim. "Incremental Capacity Curve Peak Points-Based Regression Analysis for the State-of-Health Prediction of a Retired LiNiCoAlO ₂ Series/Parallel Configured Battery Pack", Electronics, 2019 Publication	<1 %

1%

1%

1%

1%

1%

1%

1%

1%

www.scribd.com



real.mtak.hu



www.hindawi.com



spatial.usc.edu



www.icee.basim.edu



www.inderscience.com



A.D. Indriyanti, D.R. Pribandito, I.G.P.E.P. Prishana, P. Suryanto, S. Supatmiko, J. Fikanda, "Simple Additive Weights algorithm to aid administrator decision making of the underprivileged scholarship", Journal of Physics: Conference Series, 2019



Lee, Park, Kim, "Incremental Capacity Curve Peak Points Based Regression Analysis for the State-of-Health Prediction of a Retired LiNiCoAlO2 Series Parallel Configured Battery Pack", Electronics, 2019



Exclude quotes On

Exclude matches Off

Exclude bibliography On