Computer Science

E-ISSN No : 2581-6195 | Volume : 3 | Issue : 2| Mar-Apr 2020

ANALYSIS OF CHANGE PRICE ON CUMULATIVE INFLATION USING LINEAR REGRESSION

Yunita Sartika Sari

ABSTRACT

In this digital era, all data is easily accessed, stored, traced, and even analyzed. Big Data is not just a trend popularized by exclusive circles, but is a marker of a paradigm shift to understanding social processes. Referring to the big benefits that Big Data technology can offer, it is interesting to see the extent to which Big Data technology has been utilized in Indonesia, for example in government agencies, and what challenges arise in its application. There are 7 groups of expenditure of goods and services based on the Classification of individual consumption by purpose, one of which is the clothing group. This study aims to determine the correlation between changes in clothing prices to inflation. The method used in this study is linear regression.

KEYWORDS: Big Data, Regression, Classification, Clothing

1. INTRODUCTION

In the digital era, the data is one of the components that are important in decision making. Big data can be interpreted as a collection of data that is very large (volume), very quickly changing / growing (velocity), present in various forms / formats (variety), and has a certain value (value), with a note if it comes from accurate source (veracity).

Fluctuations in commodity prices provide sentiment for the economy and can have an impact on a country's economic stability. One factor that can illustrate a country's economic stability is the inflation rate. Inflation becomes a quantitative measure of fluctuations in the average price of a group of goods and services. One of the main indicators to measure inflation is the price movement of the group of goods and services consumed by the public. There are 7 groups of expenditure of goods and services based on the Classification of individual consumption by purpose, one of which is the clothing group.

2. PLATFORM THEORY

2.1 Definition of Linear Regression

Linear regression is an approach to model the relationship between Y dependent variables and one or more independent variables called X. One of the uses of linear regression is to make predictions based on data that has been previously owned. The relationship between these variables is called the linear regression model. Based on the use of independent variables, the linear regression can be divided into two, namely univariate linear regression and multivariate linear regression.

2.2. Definition of Inflation

Inflation is a process of increasing prices in general and continuously related to market mechanisms that can be caused by various factors, among others, increased public consumption, excess liquidity in the market that triggers consumption or even speculation, to include also due to the uneven distribution of goods. In other words, inflation is also a process of continuously decreasing the value of a currency. Inflation is a process of an event, not a high or low price level. That is, the price level that is considered high does not necessarily indicate inflation. Inflation is an indicator to see the level of change, and is considered to occur if the price increase process takes place continuously and influence each other. The term inflation is also used to mean an increase in the money supply which is sometimes seen as a cause of rising prices.

3. RESULTS AND DISCUSSION

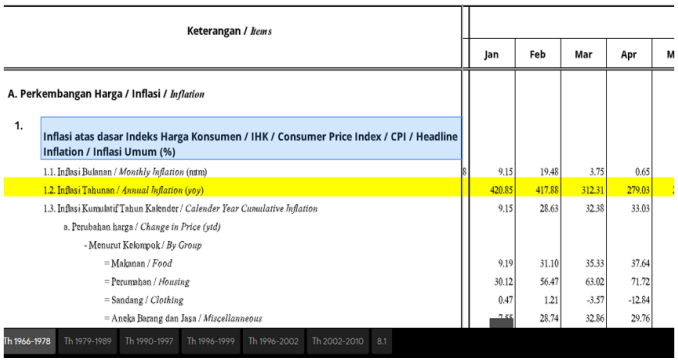
a. Data

The data source used is the website bi.go.id, from the website we get Inflation data to be processed into Big Data. From these data there are cumulative inflation figures with changes in clothing prices.

Faculty of Computer Science, Mercu Buana University, Indonesia

HOW TO CITE THIS ARTICLE:
Yunita Sartika Sari (2020). Analysis Of Change Price On Cumulative Inflation Using Linear Regression , International Educational Journal of Science and Engineering (IEJSE), Vol: 3, Issue: 2, 01-03

Research Paper



Keterangan / *Items*	Jan	Feb	Mar	Apr
A. Perkembangan Harga / Inflasi / *Inflation*				
1. Inflasi atas dasar Indeks Harga Konsumen / IHK / Consumer Price Index / CPI / Headline Inflation / Inflasi Umum (%)				
1.1. Inflasi Bulanan / *Monthly Inflation* (mtm)	9.15	19.48	3.75	0.65
1.2. Inflasi Tahunan / *Annual Inflation (yoy)*	420.85	417.88	312.31	279.03
1.3. Inflasi Kumulatif Tahun Kalender / *Calender Year Cumulative Inflation*	9.15	28.63	32.38	33.03
a. Perubahan harga / *Change in Price (ytd)*				
- Menurut Kelompok / *By Group*				
= Makanan / *Food*	9.19	31.10	35.33	37.64
= Perumahan / *Housing*	30.12	56.47	63.02	71.72
= Sandang / *Clothing*	0.47	1.21	-3.57	-12.84
= Aneka Barang dan Jasa / *Miscellaneous*	[illegible]	28.74	32.86	29.76

Th 1966-1978 | Th 1979-1989 | Th 1990-1997 | Th 1996-1999 | Th 1996-2002 | Th 2002-2010 | 8.1

Figure 1: Dataset

b. Method

Regression is used in building models to predict values (one target or multiple targets) of input data (with certain dimensions) that already exist. Data processing to visualization:
Import libraries and data used

```
In [1]: import pandas as pd
        df = pd.read_excel("D:/FGA/code/inflasi.xlsx")
        dfa = pd.read_excel("D:/FGA/code/ratainflasi.xlsx")
```

```
In [2]: df.head()
```

Out[2]:

	Tahun	Inflasi_Kumulatif	ChangeInPrice_Sandang	Unnamed: 3	Unnamed: 4
0	1967-01-01	9.15	0.47	NaN	NaN
1	1967-02-01	28.63	1.21	44.195833	9.170000
2	1967-03-01	32.38	-3.57	56.250000	62.269167
3	1967-04-01	33.03	-12.84	NaN	NaN
4	1967-05-01	32.82	0.51	NaN	NaN

```
In [3]: dfa.head()
```

Out[3]:

	RataInflasi	RataSandang
0	44.195833	9.170000
1	56.250000	62.269167
2	4.170833	1.310833
3	6.499167	8.952500
4	2.062500	0.933333

The data that is ready to be processed consists of 3 columns, namely Year, Cumulative Inflation and Clothing Change in Price. The main variables taken are cumulative inflation and changes in clothing prices.

Description of the data

```
df.describe()
```

	Inflasi_Kumulatif	ChangeInPrice_Sandang	Unnamed: 3	Unnamed: 4
count	230.000000	230.000000	14.000000	14.000000
mean	11.345217	7.903435	13.394286	9.924464
std	15.356095	15.568204	16.366810	15.997618
min	-1.650000	-12.840000	2.062500	-1.540833
25%	3.115000	0.897500	5.808750	2.699167
50%	6.385000	2.955000	6.764167	4.685000
75%	11.207500	7.607500	9.458125	9.115625
max	81.830000	100.410000	56.250000	62.269167

The data type is float. The amount of data is 230, where the average for cumulative inflation is 11,345 while for Change in Price Clothing is 7.9. The standard deviations of the two variables are spaced 15 so that they have fluctuations that are almost as large.

Import model Regresi Linear

```
In [4]: import numpy as np
        from sklearn.linear_model import LinearRegression
        y=np.array(df["Inflasi_Kumulatif"])
        x=np.array(df["ChangeInPrice_Sandang"]).reshape((-1,1))
        model = LinearRegression()
        model.fit(x,y)
        model = LinearRegression().fit(x,y)
        print('intercept:', model.intercept_)
        print('slope:', model.coef_)

        intercept: 5.098110736332823
        slope: [0.79042933]
```

Intercept is an intersection point between a line with the Y axis on the diagram axis when the value of X = 0.
Slope is a measure of the slope of a line.

c. Data visualization

```
In [16]: import matplotlib.pyplot as plt
         plt.scatter(df["ChangeInPrice_Sandang"], df["Inflasi_Kumulatif"],  color='red')
         plt.title('Scatter Plot')
         plt.xlabel('x')
         plt.ylabel('y')
         plt.plot(x,y_prediksi, '-b')
         plt.show()
```

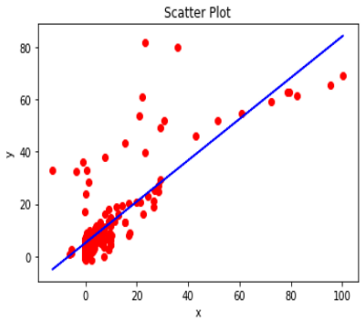


Figure 2: Data Visualization

With scatter plots in matplotlib.pyplot we can see the spread of data points and regression lines in blue.
The more data spread around the regression line, the better regression model obtained. In this case, there is still data that does not spread around the regression line.

4. CONCLUSION

Cumulative inflation and changes in clothing prices have a close relationship with a correlation of 0.8013. By placing changes in clothing prices as independent variables x and cumulative inflation as dependent variables y, with Mean Square Error obtained from predictions amounting to 84.0158. From the results obtained, it can be concluded that the regression model above is good enough to predict cumulative inflation from data on changes in clothing prices.

REFERENCE

1. Triana, Yaya Sudarya, Astari Retnowardhani. Enhance Interval Width of Crime Forcasting with ARIMA model-fuzzy alpha cut. Telkomnika. Vol 17 No 3. 2019
2. Nia Rahma Kurnianda & Yunita Sartika Sari. Analysis and Design

of Information System for Self-Journal on Food Based Dietary Assessment Record for Diabetes Patients. International Research Journal of Computer Science (IRJCS). Volume 06 Issue 5. 2018
3. Ranggadara, Indra & Suhendra "Naive Bayes Algorithm with Chi Square and NGram Feature for Reviewing Laptop Product on Amazon Site.".IRJCS, Vol 7, issue 2.2018
4. William H. Kruskal and Judith M. Tanur, ed. "Linear Hypotheses," International Encyclopedia of Statistics. Free Press.1982
5. Lindley, D.V. (1987). "Regression and correlation analysis," New Palgrave: A Dictionary of Economics, v. 4, pp. 120–23.
6. Birkes, David and Yadolah Dodge, Alternative Methods of Regression. ISBN 0-471-56881-3
7. Corder, G.W. and Foreman, D.I. (2009).Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach Wiley, ISBN 978-0-470-45461-9
8. Draper, N.R. and Smith, H. (1998).Applied Regression Analysis Wiley Series in Probability and Statistics
9. Fox, J. (2017). Applied Regression Analysis, Linear Models and Related Methods. Sage
10. Hardle, W., Applied Nonparametric Regression (1990), ISBN 0-521-42950-1
11. Meade, N. and T. Islam (1995) "Prediction Intervals for Growth Curve Forecasts," Journal of Forecasting, 14, pp. 413–430.
12. Witten, H.Ian, Eibe, F. and Mark, H., 2011. Data Mining Practical Mechine Learning Tools and Techniques Third Edition (3rd Ed.). Elsevier Inc.