**IE 360**

**STATISTICAL FORECASTING AND TIME SERIES**

**HOMEWORK 1**

by

İrem Beyza Hüner

**Course Instructor** : Assoc. Prof. Mustafa Baydoğan

**Date of Submission** : 05.04.2024

**Department of Industrial Engineering**

**Boğaziçi University**

**Bebek, Istanbul**

**TABLE OF CONTENTS**

[TABLE OF FIGURES iii](#_Toc163246878)

[1. INTRODUCTION 1](#_Toc163246879)

[2. DATA MANIPULATION AND VISUALIZATION 2](#_Toc163246880)

[3. TIME SERIES REGRESSION ANALYSIS 5](#_Toc163246881)

[4. CONCLUSIONS 10](#_Toc163246882)

[REFERENCES 11](#_Toc163246883)

# TABLE OF FIGURES

[**Figure 2.1** House Sales Time Series Plot. 2](#_Toc163246885)

[**Figure 2.2** US Dollar Selling Exchange Rate Time Series Plot. 2](#_Toc163246886)

[**Figure 2.3** Employment Rate Time Series Plot. 3](#_Toc163246887)

[**Figure 2.4** Search Volume of the Word “Zam” Time Series Plot. 3](#_Toc163246888)

[**Figure 2.5** Scatter Plot. 4](#_Toc163246889)

[**Figure 3.1** ACF Plot of residuals for regression model 1. 5](#_Toc163246890)

[**Figure 3.2** Regression between house sales and the search volume of the word “Zam”. 6](#_Toc163246891)

[**Figure 3.3** ACF Plot of residuals for regression model 2. 7](#_Toc163246892)

[**Figure 3.4** Regression between US Dollar Selling Exchange Rate and the search volume of the word “Zam”. 7](#_Toc163246893)

[**Figure 3.5** ACF Plot of residuals for regression model 3. 8](#_Toc163246894)

[**Figure 3.6** Regression between Employment Rate and the search volume of the word “Zam”. 9](#_Toc163246895)

# INTRODUCTION

The Electronic Data Delivery System of the Central Bank of the Republic of Turkey (CBRT) is used to obtain data related to economic measures. Three time series from different data categories are selected. Under the category “House and Construction statistics" House Sales Statistics - Mortgaged sales data is selected. Under the "Labor Force Statistics" category, Employment rate (%) is selected. Finally, under the “Exchange Rates” category, US Dollar (Selling) data is selected.

All three time series that are selected are sampled at a monthly level and the length of the observation period in each of the time series is selected as 8 years, from the beginning of 2016 to the end of 2023. In the data, the column name TP\_AKONUTSAT2\_TOPLAM represents House Sales Statistics - Mortgaged sales; TP.DK.USD.S.YTL represents the US Dollar selling exchange rate and TP.TIG07 represents Employment rate (%).

In the R code, House Sales Statistics time series is referred as dataset\_1, US Dollar selling exchange rate time series is referred as dataset\_2 and the employment rate time series is referred as dataset\_3. The correlation coefficient between datasets are calculated. Consequently, the absolute correlation coefficient between each pair is less than 0.5.

In this homework, it is investigated whether price raise (zam) is related to employment rate, house sales and US Dollar (selling) exchange rate. The analysis is done via comparing the search volume data obtained from Google Trends for the word “zam” meaning price raise in Turkish with the other datasets in hand, obtained from CBRT.

# DATA MANIPULATION AND VISUALIZATION

Firstly, all time series are plotted to observe their trend and seasonality.

A graph showing a number of numbers and a line graph

Description automatically generated

**Figure 2.1** House Sales Time Series Plot.

A graph showing the growth of the stock market

Description automatically generated

**Figure 2.2** US Dollar Selling Exchange Rate Time Series Plot.

A graph showing the number of stock prices

Description automatically generated

**Figure 2.3** Employment Rate Time Series Plot.

A graph showing a number of different types of data

Description automatically generated with medium confidence

**Figure 2.4** Search Volume of the Word “Zam” Time Series Plot.

Then, a scatterplot matrix is constructed for these 4 datasets.

A screenshot of a graph

Description automatically generated

**Figure 2.5** Scatter Plot.

# TIME SERIES REGRESSION ANALYSIS

**Regression Model 1**

In this model, the forecasting variable is house sales, and the predictor variable is search volume of the word “Zam”.

**Summary:**

Residuals:

Min 1Q Median 3Q Max

-26228 -12740 -1819 7037 100583

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 32966.2 2663.9 12.375 <2e-16 \*\*\*

Zam -202.0 117.2 -1.724 0.088 .

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 18740 on 94 degrees of freedom

Multiple R-squared: 0.03066, Adjusted R-squared: 0.02035

F-statistic: 2.973 on 1 and 94 DF, p-value: 0.08795

A graph of a plot of residuals

Description automatically generated

**Figure 3.1** ACF Plot of residuals for regression model 1.

A graph with black dots and a blue line

Description automatically generated

**Figure 3.2** Regression between house sales and the search volume of the word “Zam”.

**Regression Model 2**

In this model, the forecasting variable is US Dollar Selling Exchange Rate and the predictor variable is search volume of the word “Zam”.

**Summary:**

Residuals:

Min 1Q Median 3Q Max

-9.600 -2.975 -1.752 1.150 18.875

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.84581 0.78076 6.207 1.45e-08 \*\*\*

Zam 0.27327 0.03434 7.958 3.92e-12 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.493 on 94 degrees of freedom

Multiple R-squared: 0.4025, Adjusted R-squared: 0.3961

F-statistic: 63.32 on 1 and 94 DF, p-value: 3.919e-12

A graph of a plot

Description automatically generated with medium confidence

**Figure 3.3** ACF Plot of residuals for regression model 2.

A graph with a red line and black dots

Description automatically generated

**Figure 3.4** Regression between US Dollar Selling Exchange Rate and the search volume of the word “Zam”.

**Regression Model 3**

In this model, the forecasting variable is Employment Rate and the predictor variable is search volume of the word “Zam”.

**Summary:**

Residuals:

Min 1Q Median 3Q Max

-5.4274 -0.7333 0.2923 1.2821 2.3521

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 45.56721 0.23648 192.688 < 2e-16 \*\*\*

Zam 0.04006 0.01040 3.851 0.000215 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.664 on 94 degrees of freedom

Multiple R-squared: 0.1363, Adjusted R-squared: 0.1271

F-statistic: 14.83 on 1 and 94 DF, p-value: 0.0002146

A graph of a plot

Description automatically generated with medium confidence

**Figure 3.5** ACF Plot of residuals for regression model 3.

A graph with a green line and a green line

Description automatically generated

**Figure 3.6** Regression between Employment Rate and the search volume of the word “Zam”.

# CONCLUSIONS

**Regression Model 1**

As it is seen from the Figure 3.2, the relationship between house sales and the search volume of the word “Zam” is negative. However, the relationship is not that strong, which can be observed by the R-squared value equal to 0.03066. When the ACF of residuals is observed, it is seen that residuals are correlated with lag 1. Additionally, the P-value is higher than 0.05, therefore it can be said that the regression is not significant.

**Regression Model 2**

Figure 3.4 shows that there is a positive correlation between US Dollar Selling Exchange Rate and the predictor variable is search volume of the word “Zam”. It is also physically meaningful since as the US Dollar exchange rate increases, the overall prices of foreign goods increase in the country. Therefore, in most goods, a price increase is observed, this is reflected by an increase in the search volume of the word “Zam”. The R-squared value is 0.4025 and the p-value is smaller than 0.05, therefore, it can be said that the regression is significant. Also, ACF plot of the residuals are correlated. Correlation of residuals indicate that the model is not perfect and the assumption of uncorrelated residuals is violated. Therefore, it may indicate that might be more predictor variables other than search volume of the word “Zam” to forecast US Dollar Selling Exchange Rate.

**Regression Model 3**

Figure 3.4 shows that there is too little positive correlation between Employment Rate and the search volume of the word “Zam. R-squared value is 0.1363, which also shows they are not very correlated. ACF plot shows that the residuals are correlated, therefore the uncorrelated residual assumption is violated.

# REFERENCES

https://evds2.tcmb.gov.tr/

http://trends.google.com/