Name: Ei Tanaka

Class: DATS 6313 Time Series Analysis Section 10

Date: September 13, 2023

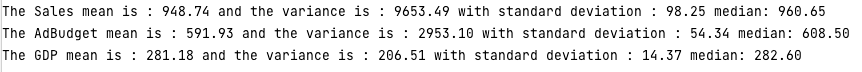
Lab #1 Stationary & Non-stationary

1. Load the time series data called ‘tute1.csv’ [ the dataset can be found on the course GitHub]. This dataset relates to the quarterly sales for a small company in 1981. Sales contains the quarterly sales, AdBudget is the advertisement budget and GPD is the gross domestic product for a small company. Plot Sales, AdBudget and GPD versus time step in one graph. Add grid and appropriate title, legend to each plot. The x-axis is the time, and it should show the time (year). The y-axis is the USD ($).

A graph of sales and sales

Description automatically generated with medium confidence

1. Find the time series statistics (average, variance, standard deviation, median) of Sales, AdBudget and GPD and display the Average, variance, and standard deviation as follow on the console: [5pts]



1. Prove that the Sales, AdBudget and GDP in this time series dataset is stationary. Plot all means and variances and show that the means and variances are almost constant.

A graph of sales and sales

Description automatically generated

A graph of a graph showing a number of samples

Description automatically generated with medium confidence

A graph of a graph showing the value of a product

Description automatically generated with medium confidence

1. Write down your observation about the plot of the mean and variance in the previous step. Is Sales, GDP and AdBudget stationary or not? Explain why. [5pts]

**For all “Sales”, “AdBudget”, and “GDP”, The rolling mean values seem to fluctuate around a central value, but without a clear increasing or decreasing trend. This suggests that the mean might be relatively constant over time. Similarly, the rolling variance values also fluctuate, but they don't show a clear increasing or decreasing trend either. This suggests that the variance might also be relatively constant over time.**

1. Perform an ADF-test to check if the Sales, AdBudget and GDP stationary or not (confidence interval 95% or above). Does your answer for this question reinforce your observations in the previous step? Hint: You can use the following code to calculate the ADF-test. [5pts]

The result of ADF test for Sales

A screenshot of a computer

Description automatically generated

The result of ADF test for AdBudget

A white background with numbers and symbols

Description automatically generated

The result of ADF test for GDP

A white background with black text

Description automatically generated

All three time series datasets are considered stationary at the 95% confidence interval. This conclusion is drawn from the fact that the p-values for all datasets are less than 0.05. Specifically:

* The first dataset has a p-value of 0.016628.
* The second dataset has a p-value of 0.064434, but it's worth noting that this dataset is not stationary at the 95% confidence level; however, it would be considered stationary at the 90% confidence level.
* The third dataset has a p-value of 0.018443.

Thus, based on the ADF test, the majority of the datasets exhibit stationary behavior at the 95% confidence level.

1. Perform an KPSS-test to check if the Sales, AdBudget and GDP stationary or not (confidence interval 95% or above). Does your answer for this question reinforce your observations in the previous steps? Hint: You can use the following code to calculate the KPSS-test. [5pts]
2. Repeat step 1-6 with ‘AirPassengers.csv’ on the GitHub. This timeseries dataset is univariate with #passengers as an attribute. [40pts]
3. If the #passengers is not stationary, it needs to become stationary by transformation. [20pts] a. Perform a 1st order non-seasonal differencing. Is the dataset become stationary? Explain why. b. Perform a 2nd order non-seasonal differencing. Is the dataset become stationary? Explain why. c. Perform a 3rd order non-seasonal differencing. Is the dataset become stationary? Explain why. d. If the procedures in steps a, b and c does not make the dataset stationary then perform a log transformation of the original raw dataset followed by a 1st order differencing then plot the rolling mean and variance. Perform ADF-test and KPSS-test on the transformed dataset and display the results on the console. This step should make the dataset stationary which means the rolling mean and variance is stabilize and the ADF-test confirms stationarity.