United States International University-Africa

Education to take you places

SCHOOL OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF DATA SCIENCE AND ANALYTICS

SUMMER 2024 – ASSIGNMENT 4

COURSE CODE: STA 3050A **UNIT NAME:** TIME SERIES AND FORECASTING

DATE: 12TH JULY 2024 **TOTAL MARKS**: 100 MARKS

INSTRUCTIONS:

For this exercise:

1. ANSWER ALL QUESTIONS
2. Do all your work in the Rmarkdown (.rmd).
3. Submissions should be in either a **`.rmd` file**
4. NO SUBMISSIONS SHOULD BE DONE VIA EMAIL

Dataset Description: You have been provided with monthly sales data for a new retail store over a period of approximately 6.5 years (80 months). The data is measured in thousands of dollars.

Dataset: timeseries_data.csv

1. Load the dataset into RStudio and plot a time series graph to visualize the monthly sales data.
    - Provide an interpretation of any observed trends or patterns in the data.
2. Perform the Augmented Dickey-Fuller (ADF) test on the sales data to determine if it is stationary.
    - Interpret the results of the test and discuss the implications for further analysis.
3. Apply the Box-Cox transformation to the sales data to stabilize variance.
    - Plot the transformed data and interpret how the transformation impacts the distribution and stationarity of the data.
4. Perform the ADF test on the transformed sales data to confirm stationarity.
    - Interpret the results and compare them with those from Question 2 to assess the effectiveness of the Box-Cox transformation.
5. Compute and plot the Autocorrelation Function (ACF) for up to 20 lags of the transformed sales data.
    - Interpret the correlogram to identify any significant autocorrelation patterns and their implications for forecasting.