Name: Sanika Hande

PRN: 22060641035

TSA Assignment

Define

The dataset is a time series dataset of sales quantity and beverage manufacturers from January 2015-April 2020. The question asks for suggestions for a suitable smoothing technique on the dataset, comparing the average cost and revenue, and whether the same smoothing technique applies to both average cost and revenue.

Analysis

1. Overview

In the given dataset there are 64 observations and 3 variables namely Revenue, Average Cost, and Sales quantity which are measured along with time at regular intervals so it is a time series data. Also, using R we can check whether the given dataset is a time series or not.

There are three types of smoothing techniques namely Exponential Smoothing, Holt's Smoothing (double exponential smoothing), and Holt's Winter smoothing (triple exponential smoothing) for time series forecasting. Here we are interested in smoothing out the series so that longer trend patterns can be revealed and short-term fluctuations and noise can be smoothened. When there is a trend and a seasonal component in the data Holt's winter smoothing is used; and when there is a trend but no seasonal component Holt's smoothing is used and when there is neither a trend nor a seasonal component, Single exponential smoothing is used. ACF graph has been used to detect Seasonality whereas Mann Kendall has been used to check whether there is a trend or not.

2. Software Used

Excel is used to carry out smoothing Whereas R is used to detect Seasonality and trends

3. Results/ Tables/ Graphs

```
> is.ts(dataset)
[1] TRUE
```

Fig 1

```
> MannKendall(dataset$Average_cost)
tau = 0.252, 2-sided pvalue =0.0033101
> MannKendall(dataset$Revenue)
tau = 0.617, 2-sided pvalue =< 2.22e-16
> MannKendall(dataset$Sales_quantity)
tau = 0.494, 2-sided pvalue =< 2.22e-16
```

Fig 2

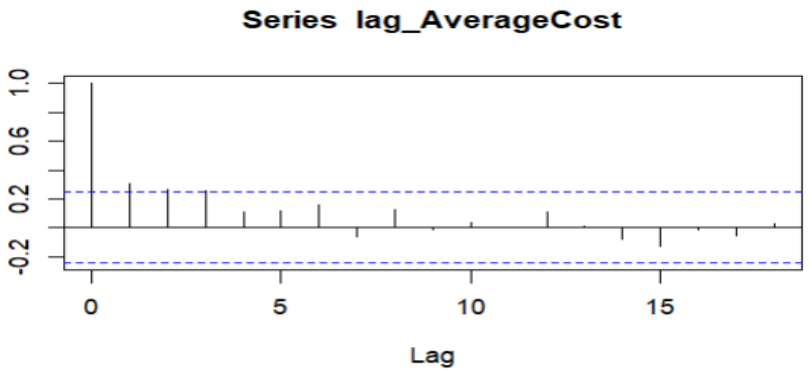

Fig 3

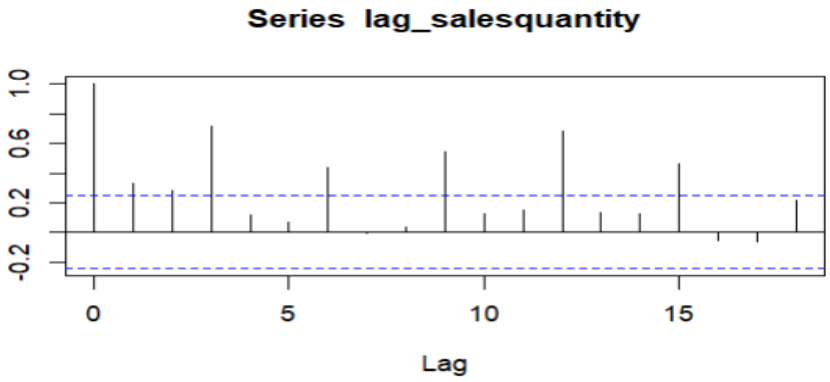

Fig 4

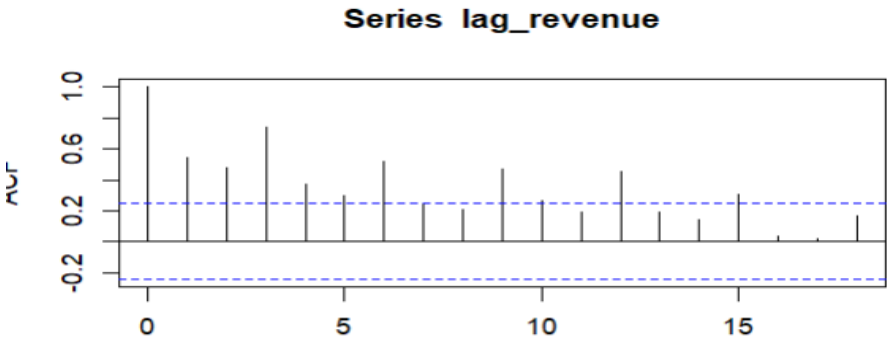

Fig 5

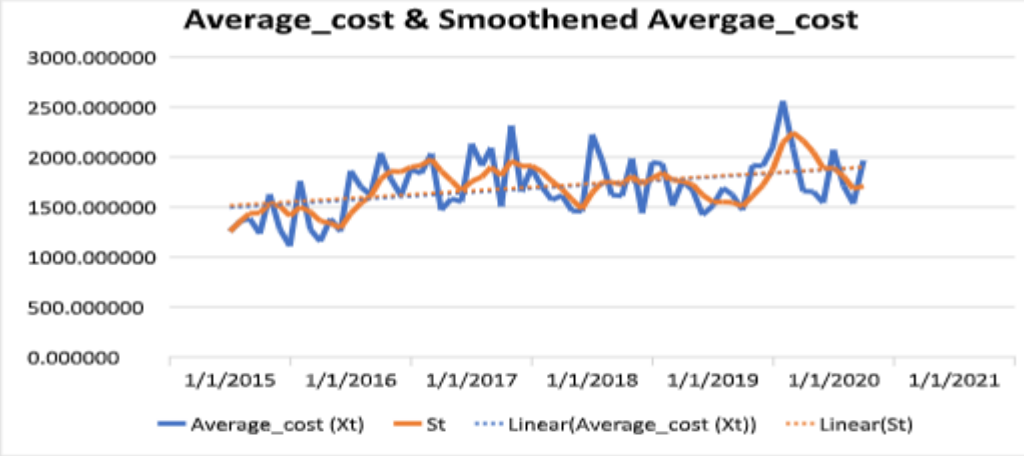

Fig 6

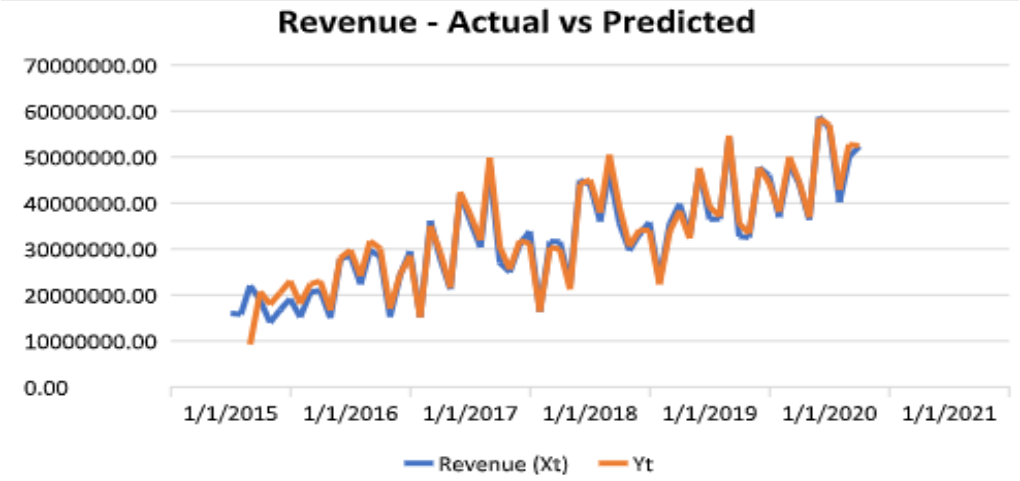

Fig 7

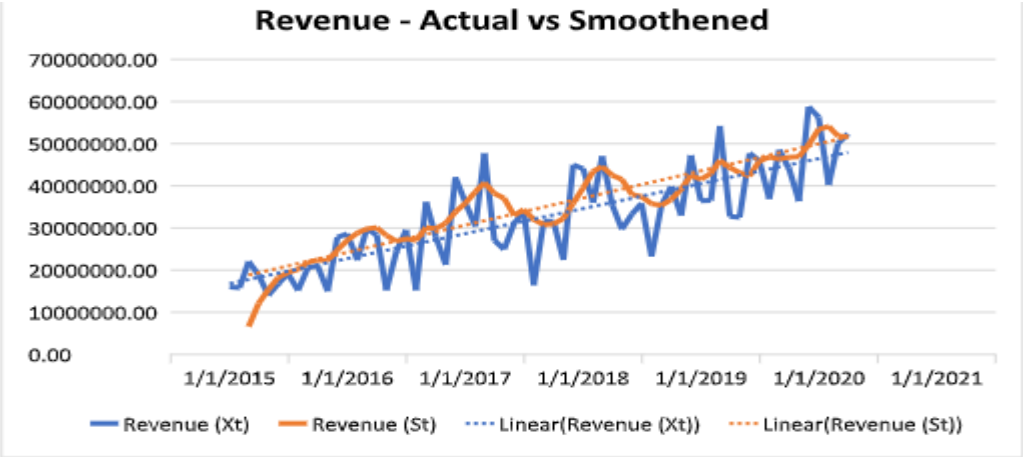

Fig 8

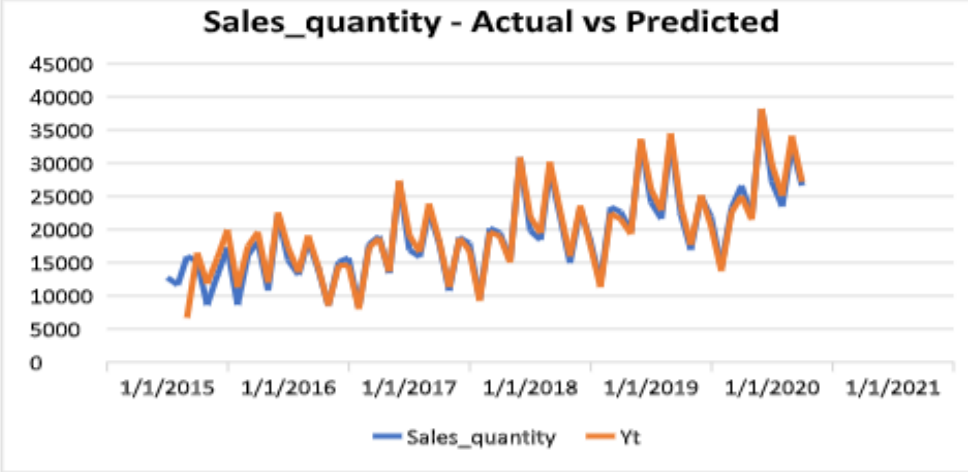

Fig 9

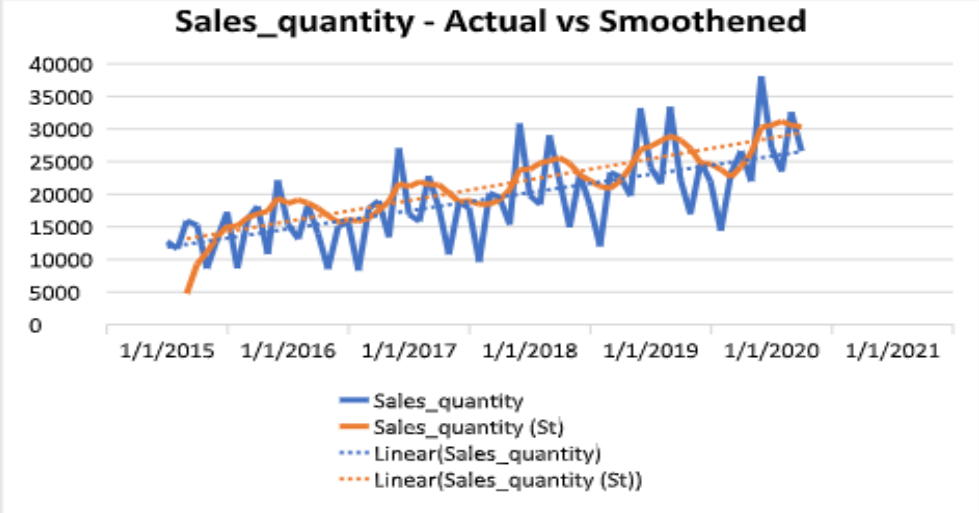

Fig 10

4. Exploratory Data Analysis

In the given time series data set,

- There are no missing values present. So there is no need for data cleaning.
- There is no need for data integration(as only one dataset is available so need to merge the dataset), data transformation (as the values of the variable lie in the required range)

- Through data mining we can see revenue has both Seasonality and trend (therefore Holt's winter Smoothing is used).
  Sales Quantity also has both seasonality and trend therefore(we have used Holt's Winter Smoothing
  In Average cost-only trend is present and Seasonality is absent thus, (we have used Holt's Smoothing).ACF plots and the Mannkandall test has been used to detect seasonality in Revenue and Average cost and trend in sales Quantity, average cost, and revenue respectively.

Interpretation

- From **Fig 1** using the is.ts() command we can see that the given dataset is of time series.
- When the p-value in the Mannkandall test is less than 0.05 we say that trend is present in the dataset and from the tau value we can interpret that trend is increasing if the tau value is positive and decreasing when the tau value is negative. So from **Fig 2,** we can see that the p-value for Average cost, Revenue, and Sales quantity are less than 0.05 Therefore trend is present in Revenue, Sales Quantity, and Average cost. And also the tau value is positive therefore the trend is increasing.
- **Fig 3** shows the ACF plot for the Average cost it can be seen from the graph that seasonality is absent as the patterns are not repeated after the regular interval and also some of the plot lines also line below the x-axis. **Fig 4** shows the ACF plot for Sales Quantity Seasonality can be clearly seen from the graph since the pattern is repeated after regular intervals of 3 time periods. Also as the time period is 3 lag can also be considered as 3. Similarly, **fig 5** shows the ACF plot for Revenue there is seasonality present as the same patterns are repeated after regular intervals of 3 periods as here also the periods are 3 lag can be considered as 3.
- **Fig 6** is the graph of average cost since the only trend is present and seasonality is absent we have used Holt's smoothing for Average cost. Here the α (data smoothing factor) is taken as 0.3 and the β (trend smoothing factor) is taken as 0.4. The orange line shows the smoothened curve and the blue line depicts the plot of the actual values.

- Since seasonality is present for Revenue and sales Quantity Holt's Winter Smoothing is used, the value of α, β,γ are 0.3,0.4 and 0.5 respectively. Here the γ is the seasonality smoothing factor. **Fig 7** and **Fig 9** represent the smoothened graph and actual graph for Revenue and Sales quantity respectively. **Fig 8 and Fig 10** show the actual vs predicted plots for Revenue and Sales Quantity respectively. From the actual vs predicted graph, we can interpret that predicted values are closer to the actual values which suggests that the series is smoothened out.

Conclusion

From the above-given analysis and interpretation, we can conclude that smoothing is used to reveal long-term trends in the time series and get rid of short-term noise and fluctuation. The appropriate technique for sales and revenue is Holt's winter smoothing whereas for Average cost Holt's Smoothing is the most suitable technique. In revenue seasonality and the trend is present so we have used Holt's Winter smoothing and in Average cost, there is only a trend therefore, we have used Holt's smoothing. So we can conclude by saying that average cost and revenue use different smoothing techniques because of the different patterns present in the data.