**Module 3 Project 3**

****

**By**

**Sunil Raj Thota**

**Date: 03/21/2021**

**Title: Project 3 Report**

**ALY 6050 – Introduction to Enterprise Analytics**

**Prof. Tamir Hegazy**

**Introduction**

In this assignment, I am using MS Excel to work on 3 different problems as given. It helps us to identify and verify exponential smoothing, Adjusted exponential smoothing, R2 value, regression equation, and forecasting. This also gives an idea between the Actual values and Predicted values of each problem, and calculated the residuals accordingly. The average is calculated by AVERAGE() function in Excel.

In the Problem 1, It’s to predict the stock price of Honeywell International, an MNC company, and perform exponential smoothing forecasts on it. There are 2 sub-problems in it where in the 1st one we have to work on the smoothing parameters α and pick the best accurate one by using Mean Square Errors (MSEs). In the 2nd sub-problem, we need to use the β which is trend parameter with the α value as 0.75. Also, we pick the most accurate β by using MSE’s.

Exponential Smoothing is used to forecast the parameters to make necessary strategic decisions. It is called as smoothing out the data by removing the arbitrary values. In Excel, we can simply access it by clicking on Data Tab, and then click on Data Analysis in the Analysis group. If the Data Analysis is not available then we need to load the Analysis Toolpak from settings and options tabs.

In the Problem2, we need to work on the linear regression model on the helicopter data. Helicopter Number column is the independent variable and Labor Hours column is the dependent variable. Also, we were specified to plot a residual plot and produce a scatter plot with the best type of trendline is having maximum R2 value.

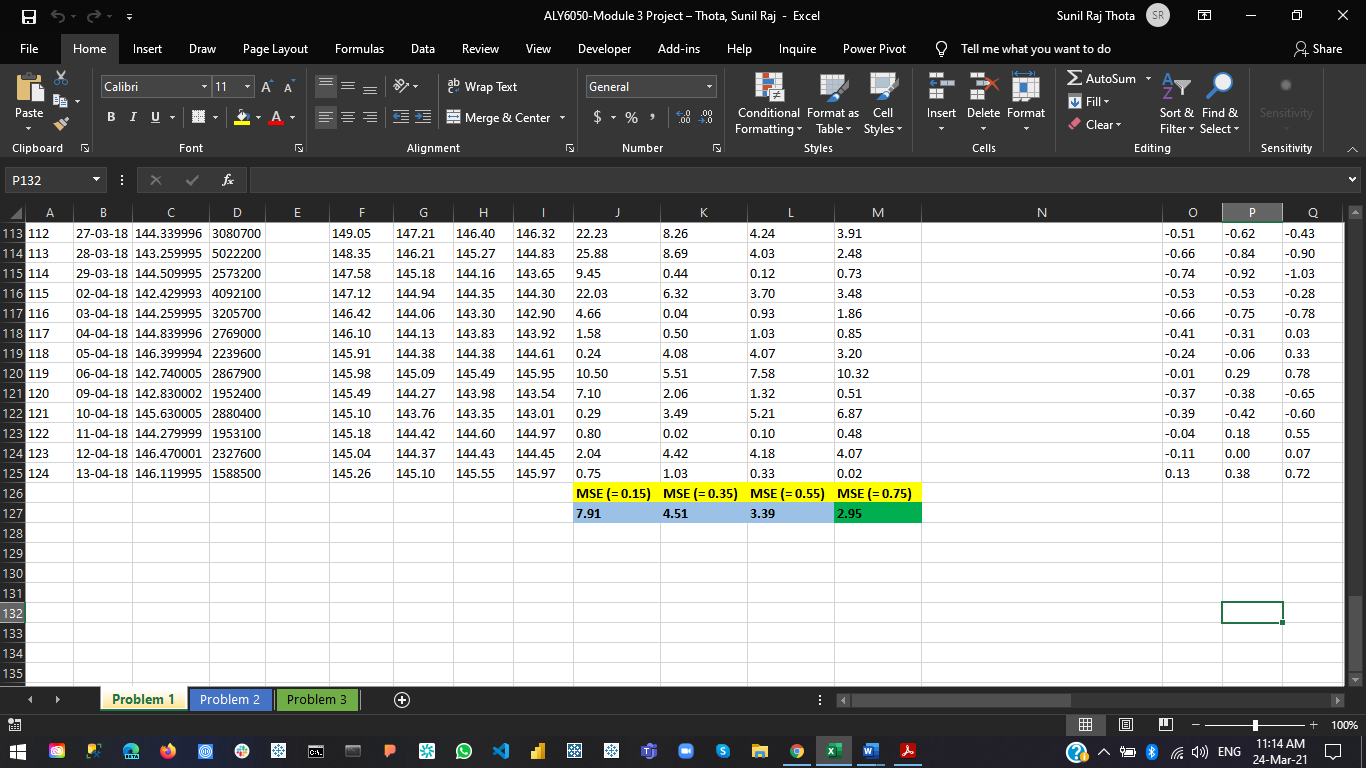
In the Problem 3, we need to work on the multiple regression model analysis on the categorical variables which shows some seasonality to forecast sales. We then need to compare the model accuracies by using MSE’s and other parameters like p-Value, standard errors, R2 value, and test statistic etc.

I will also tend to determine the type of the plots for each problem and showcase the results, and findings. In this report, I have implemented 3 Problems and their related questions.

**Analysis**

**Problem 1(A):**

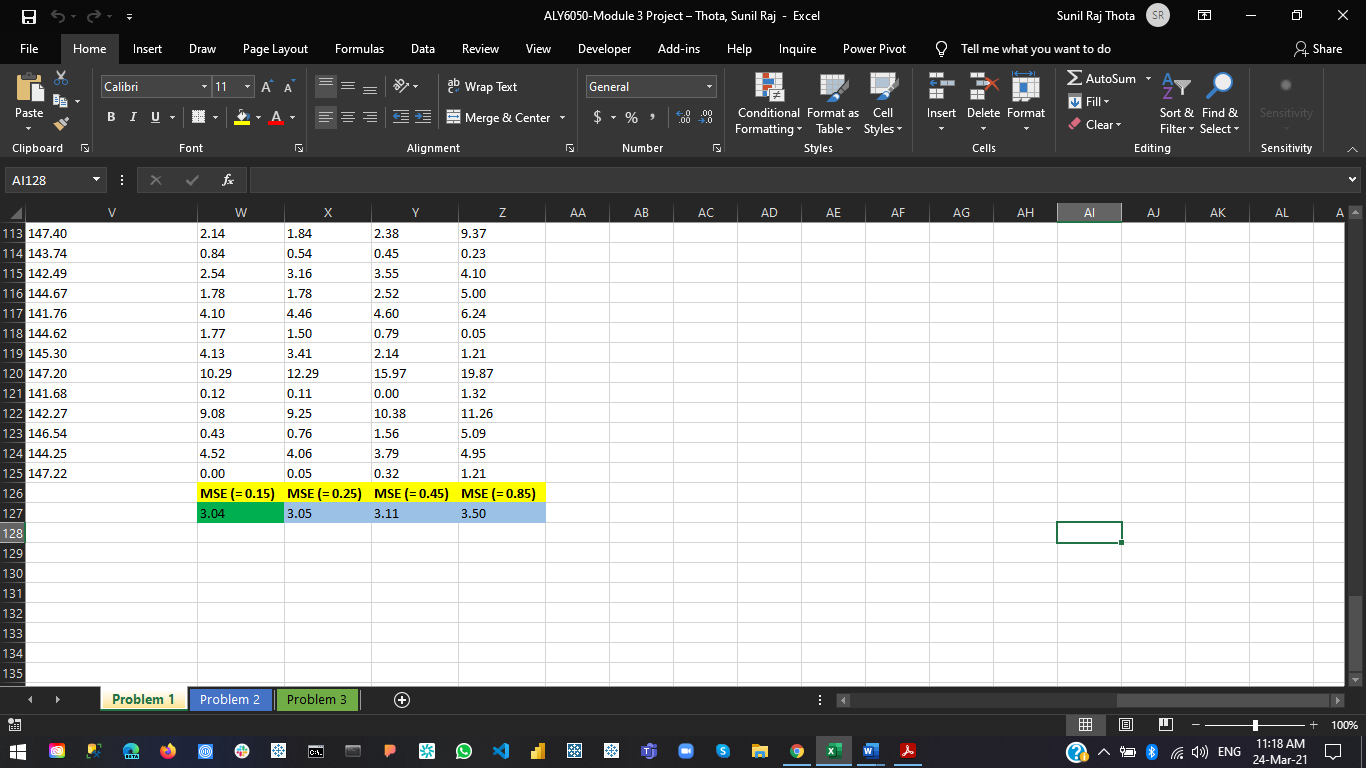
In this problem, I have taken the α values as 0.15, 0.35, 0.55, and 0.75 smoothing parameters. After the analysis and calculations I observed that 0.75 smoothing parameter as the MSE value of 2.95 which is a great way to check the predictiveness. This value must be positive and the values which are near to 0 are good. So, if the less MSE is there, then the model is more accurate. There are 2 types of forecasting methods which are exponential smoothing and adjusted exponential smoothing. Here, we have measured several α values. In this, we have found out it to be the 0.75 value that can give the best prediction.



**Problem 1(B):**

In this problem, I need to find the best trend value with the best accurate smoothing parameter 0.75. We have already performed smoothing parameter and calculated the MSE’s with the 0.15, 0.35, 0.55, and 0.75 values and found out the 0.75 value as the best one. For this, I have also done some calculations on the adjusted exponential smoothing. Now, let’s perform similar analysis on the trend parameters i.e., β value. We are given with the 0.15, 0.25, 0.45, and 0.85 values to find out the best one.

After some calculations and analysis, I have found out 0.15 trend parameter has the best MSE value of 3.04. From this, we can infer that the 0.15 trend param is the best one to predict the accurate value.



**Problem 2:**

In this problem, I have worked on the Helicopter Data and applied simple linear regression model on top of it. Also, I have plotted the Scatter and Residuals Plot. The plots showcase that its not a linear relationship.

The R squared value is also approximated to 0.92 for the best value when the polynomial trendline is taken. When a linear trendline is selected the best value, we got is 0.7285. Also, found out the slope and intercept values as -119.88 and 1767.7 respectively. I also calculated the residuals and found out to be the sum of them as 0.

**Problem 3:**

In this problem, I have worked on developing a multiple linear regression model with categorical variables that incorporate seasonality for forecasting sales using the last three years of data in the Excel file New Car Sales. Also, I had made sure that the p-values are acceptably small by consecutively removing the months with high p-value and repeated the analysis.

The Regression analysis is used to find and measure the properties of 1/ additional independent variables on 1 dependent variable. It is used to estimate and calculate the potential by choosing the each and every parameter. The R2 and Adjusted R2 values provides the goodness of fit. After some analysis and calculations, I have noticed that the months of November and December have large p-values and removed. The Regression Model Equation is = 43390.33 + 3406.33 Feb + 10802.67 Mar + 7134 Apr + 10693 May + 11866 Jun + 8610.67 Jul + 10155 Aug + 7092.67 Sep + 6689.67 Oct.

**Conclusion**

From this project I have gained a lot of knowledge and got familiar with the line plots, scatter plots, residual plots, regression, exponential smoothing, adjusted exponential smoothing, p-values, and forecasting calculations. Now, I am clear and work on these concepts very well with deep understanding. Also, I have worked on the calculations of MSE’s to find the best smoothing parameters and trend parameters. The less the MSE is the more accurate the forecasts will be. So, I used these analysis and models to define and predict the sales and forecasted them. Also, working on MS Excel gave me an opportunity to work on these techniques. The r squared values have also been calculated to predict the sales using the regression techniques. I previously did not work on these aspects, but from this Assignment I had a great time in learning them. In coming days and assignments, I will make sure to use this knowledge and apply on the problems.

**References**

[1] Svetlana Cheusheva (March 2, 2021) How to forecast in Excel: linear and non-linear forecasting methods was retrieved from https://www.ablebits.com/office-addins-blog/2019/03/20/forecast-excel-linear-exponential-smoothing-forecasting-models/

[2] Exponential Smoothing was retrieved from https://www.excel-easy.com/examples/exponential-smoothing.html

[3] Aishwarya Singh (September 23, 2020) Time Series in Excel! Learn Exponential Smoothing Models for Time Series Forecasting in Excel was retrieved from https://www.analyticsvidhya.com/blog/2020/09/time-series-forecasting-ms-excel-exponential-smoothing/