RStudio Lab Week 9 - TS 2

1. Load the following packages:

- forecast
- fpp2 (or fpp3)
- ggplot2

2. Assign the `wineind` dataset in the `forecast` package to an object called `df`. This time series shows monthly Australian wine sales. Now use a function to extract only the wine sales from the start of 1980 to the end of 1985 from the `df` data and save as an object `dt`. This dataset is what we will refer to as our 'time series' or 'original time series' in this lab.

3. Print the dataset to the console and plot the dataset. Give the plot an appropriate title and y-axis label.

4. What components are present in this time series? Describe them fully.

   Random variation, Increasing linear, trend

5. Apply a variance stabilising transformation and then differencing at the appropriate lag value. Then plot the transformed data.

6. Does the data in (5) above look stationary? Justify your answer by using a function on the transformed data to see if the data needs further transforming, and apply the appropriate transformation if one is needed.

7. Now plot the ACF of the transformed data. Comment on what you see.

   All lag values are within the dashed blue lines which represent a 95% confidence interval on the hypothesis that there is no autocorrelation. Therefore, our model has captured all possible information.

8. Split the original time series data `dt` into a training set and a test set. The training set should end at the end of 1984 and the test set should start in January of 1985 and end in December of 1985.

9. Fit the most appropriate simple forecasting model to forecast for the original time series in the short-term to the training data. Specify a forecast horizon of 12 time periods for which to make predictions. Plot the original time series `dt` and the forecasts of your fitted model on one plot. Then plot the original time series `dt` and the predictions without the prediction intervals. What do you see?

10. Assess whether you would be happy with the model you have fitted by analysing the residuals. Comment on each element of the output.

11. Evaluate the accuracy of the forecasts of your fitted model in (9). Do the values you see make sense, taking into account the components in the time series and the scale of the original data?