**Assignment 2**

**Author:** Hamna Ashraf, Student # 8826836

**Course:** PROG8431

**Question 1**

**Data Transformation (2 points):**

1.1 Read in the Ayr data and transform it into an appropriate time series datatype.

**A screenshot of a computer

Description automatically generated**

**Question 2:**

**Descriptive Time Series Data (5 points)**

2.1. Summarize the information (mean, std dev, etc.) (1 point)

**A screenshot of a computer

Description automatically generated**

2.2) Plot the time series data. (1 point)

A graph showing the temperature of the year

Description automatically generated

2.3) Decompose the times series data into the constituent components. Comment on each (any trends you observe, etc.) (1 point)

We decompose our dataset to observe the trends in data and to understand patterns over period of time.

A graph of different types of curves

Description automatically generated with medium confidence

In the trend plot we can see that there is a gradual increase in the temperature over time. The line increases and then smooths down.

2.4) Deseasonalize the information and plot the result. (1 point)

A graph showing the growth of the number of data

Description automatically generated

Figure 1: Plot after de-seasonalizing

2.5) Add any comments about what you observe: seasonality of temperature, trends, etc. (1 point)

Now that we have de-seasonalized the data, we can see a *slight* difference in the plot (figure 1) than the one observed earlier (figure 2). It can be due to irregularities or any other random factor that could be a reason for this change in the plot.

A graph with black text

Description automatically generated

Figure 2: Shows the Decomposed Monthly Data Trend 1988-1990

**Question 3**

3.1) Smooth the temperature chart using a moving average. Try 3 different values for the moving average and choose the one you think best shows the trend (if any). (2 points) **A graph showing different moving average

Description automatically generated**

In time series analysis, after three moving averages, we can see that the short-term fluctuations are smooth out in all three compared to the original data.

However, when moving the averages by 12 period, we can see the upward trend, which is more effective than 4 or 8 period averages.

3.2) Determine if the time series is stationary. (2 points)

A computer screen shot of a computer code

Description automatically generated

The data is stationary which means that the statistical properties of mean, and variance remains constant throughout the time. It also shows that there is no seasonality in data.

3.3) Create an autocorrelation chart (using acf) and comment on which lags are significant. Do previous values seem to influence current values? (2 points)

A graph with lines and text

Description automatically generated

ACF is when the data is compared with itself after different lags.

Initially at lag 0, the autocorrelation is always 1, but as the lag increases, we can see a gradual decrease in the ACF which shows a decreasing correlation. Decreasing correlation represents that data is stationary and that the statistical measures don’t change over time.

**Question 4**

4.1) Create a simple moving average forecast of temperature in Ayr for five years beyond the data provided. Graph your results along with a 75% prediction interval. (2 points)

A graph of a graph showing the growth of a number of years

Description automatically generated with medium confidence

4.2) Create an exponentially smoothing forecast of temperature in Ayr for five years beyond the data provided. Graph your results along with a 75% prediction interval.

A graph with lines and numbers

Description automatically generated

4.3) Compare the two forecasts you created in steps 1 and 2 above. Which forecast seems superior? Why?

A white background with numbers and letters

Description automatically generated

For the forecast of single moving average (step 1), we can see that just like the original trend, with increase in year, there is an increase in temperature. However, for exponential average forecast (step 2), we can notice that all the predicted temperatures for next 5 years are same. It doesn’t provide us any useful information with this forecast.

Therefore, the single moving average forecast seems to be the most suitable for our dataset.