1.   Energybills.xls

This dataset consists of energy bills over a period of time. There are 40 entries in this dataset. The first thing we will be doing is plotting the time series plot of energy bills and see what we can find out.
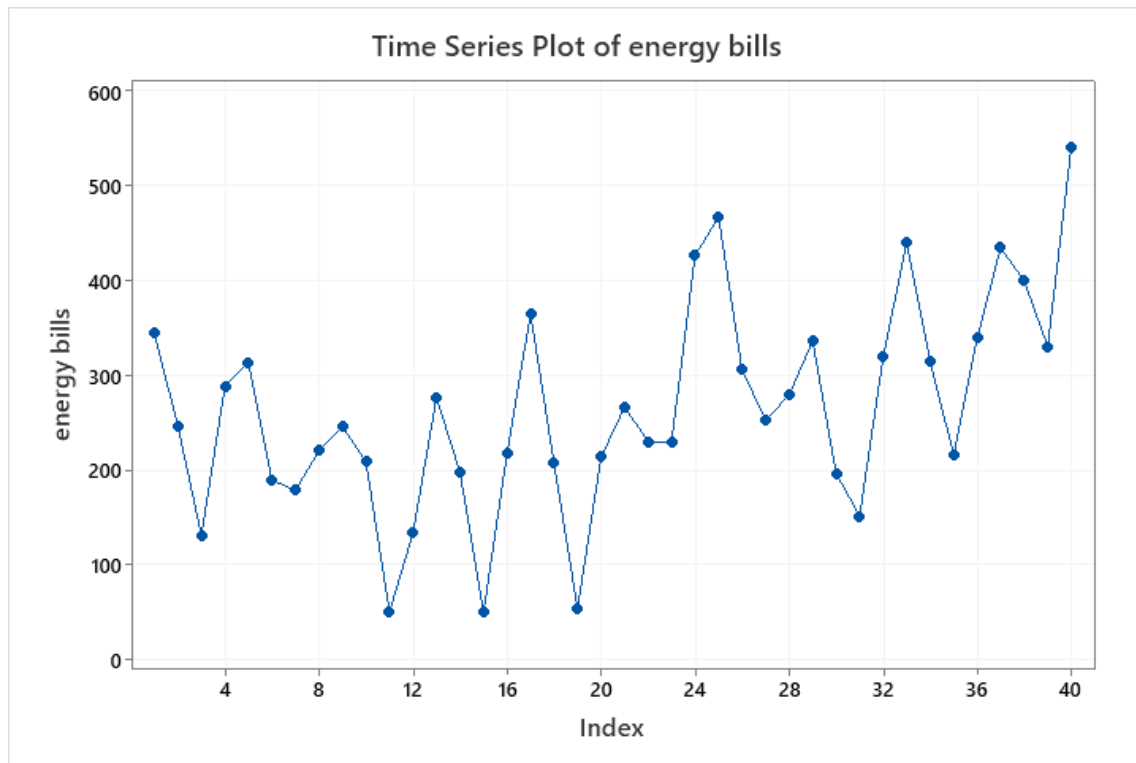

Figure 7 : Time series for energy bills

From figure 7, we can see that there may be a linear trend. Let us apply time series analysis to check which trend model works best for this dataset.
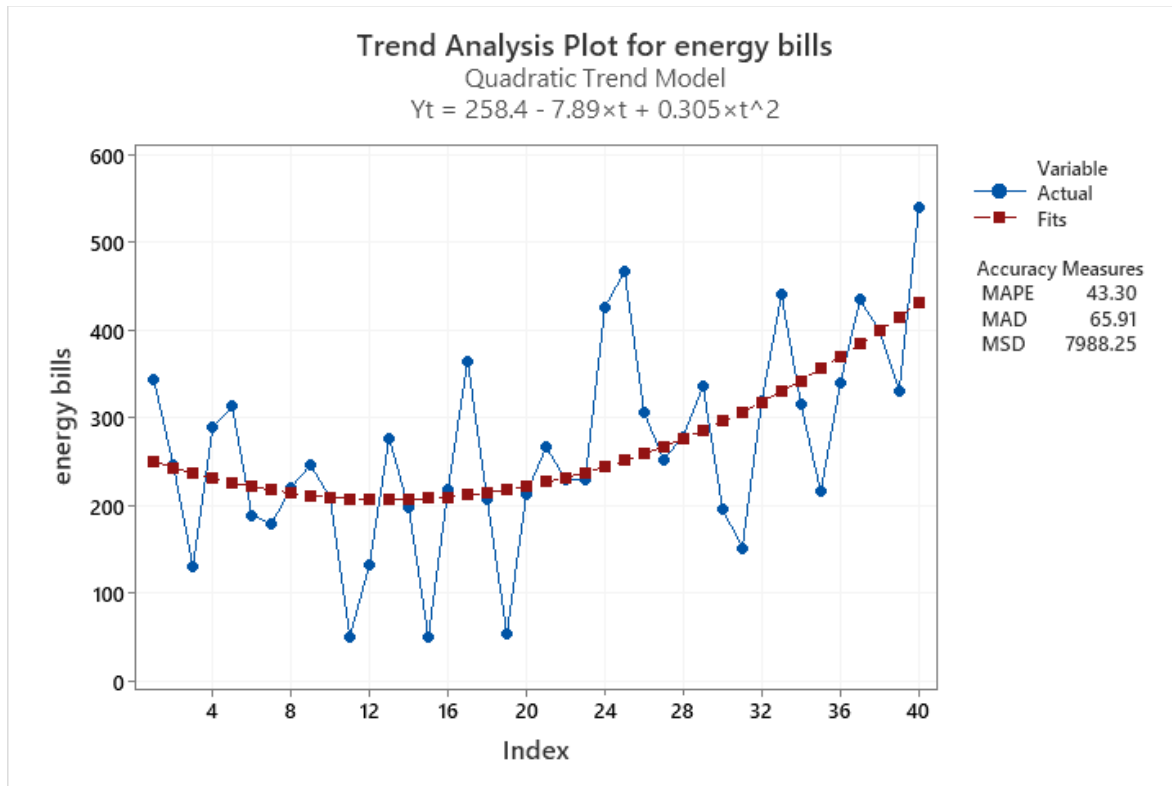
Figure 8 : Quadratic Trend model for energy bills

After checking with Linear, Quadratic, Exponential trend models, the quadratic model seems to be better than the other two.

Let us check for the ACF and partial ACF for energy bills with default lag to gain some insights.
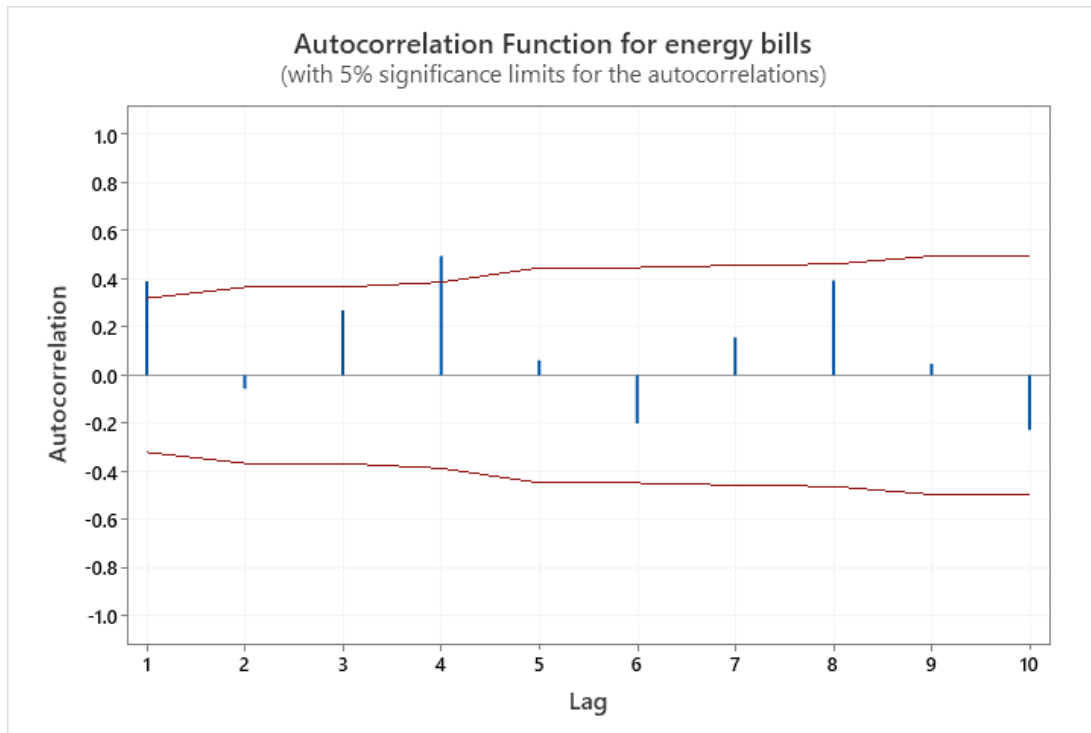
Figure 9 : ACF for energy bills with default lag

The above ACF shows us that the first and fourth entries are outside of the marked red region. This tells us that we probably need to take a difference of the energy bills.
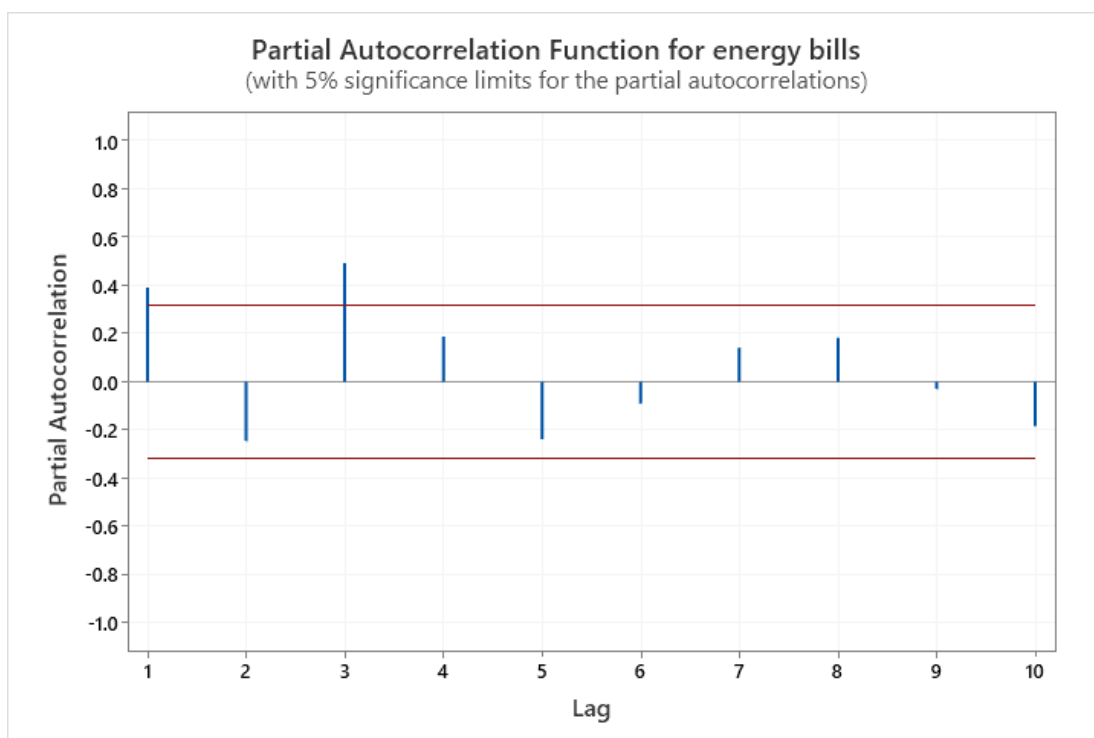

Figure 10 : Partial ACF for energy bills with default lag

The partial ACF shows us the first and third entries are out of bounds.
The second order difference is taken and the time series plot is checked to see if there is any change.
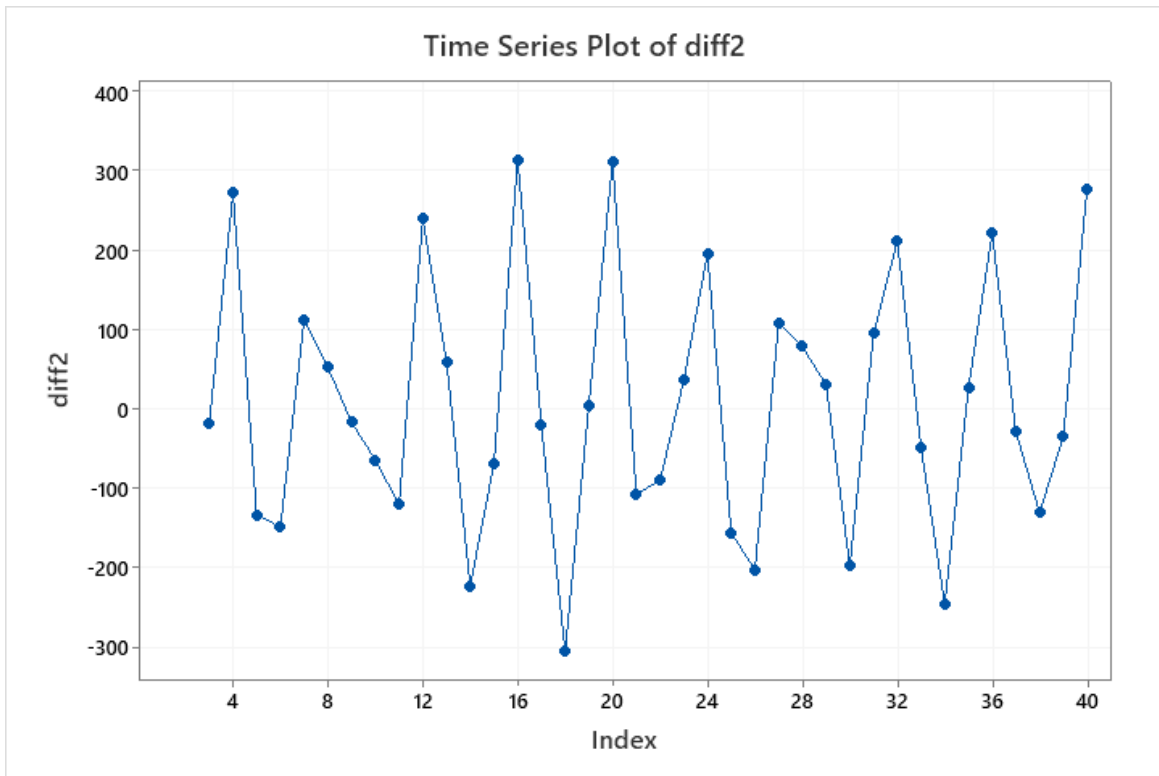

Figure 8 : Time series plot for energy bills with second order difference

As we can see there is a small linear trend and now we will check for the ACF and partial ACF for this second order difference.
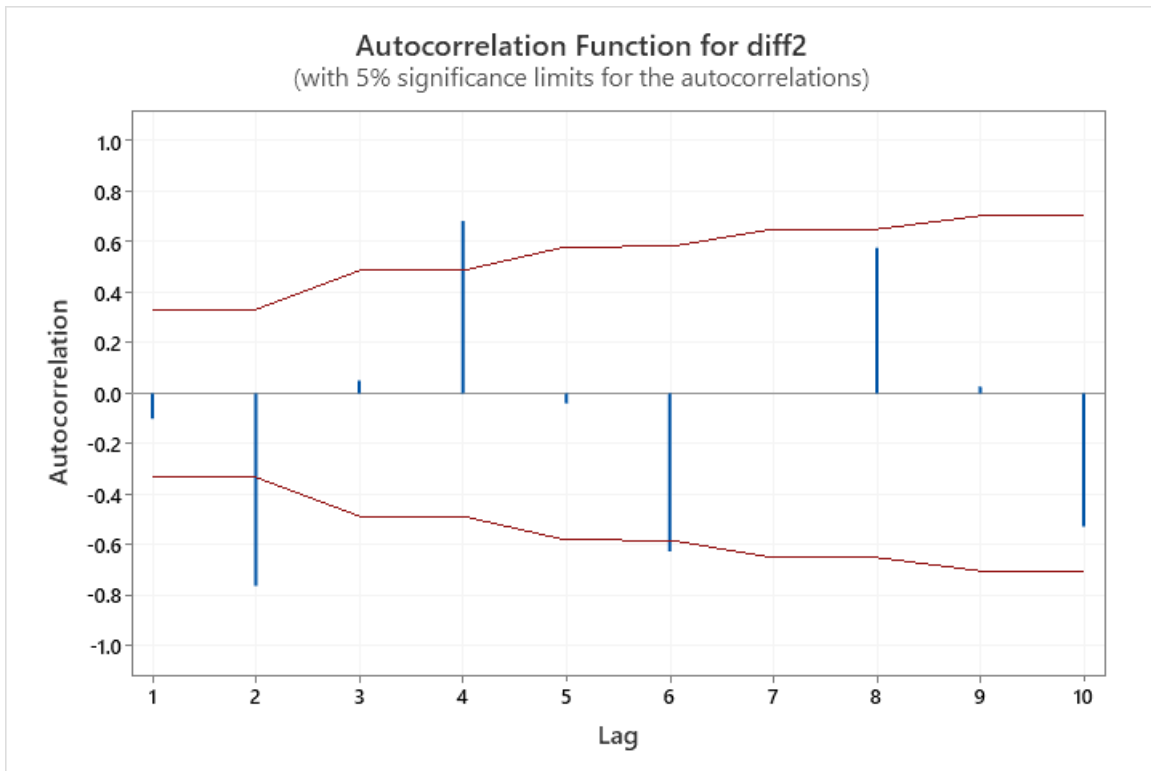
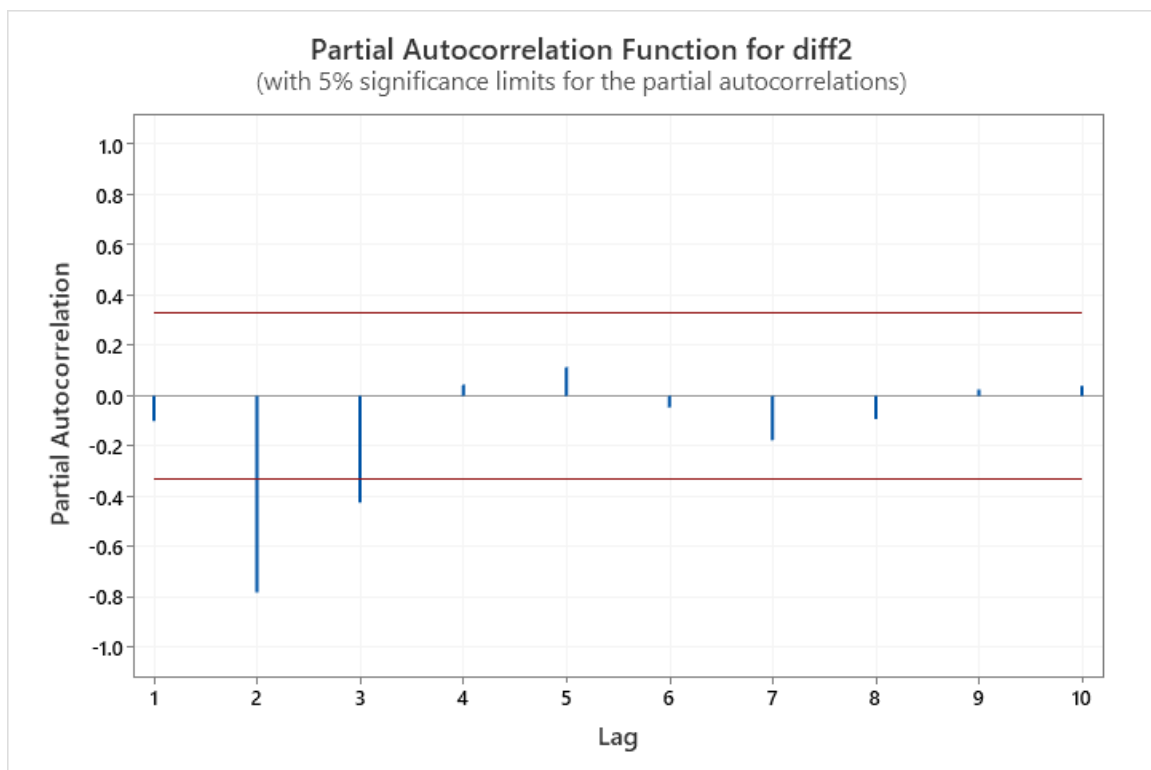**Figure 9 : ACF of second order difference**

Figure 10 : Partial ACF of second order difference

The second entry in both of these ACF and partial ACF tells us that we could run ARIMA with an autoregressive value of 2 and seasonal difference of 2. Moving average is taken as zero.

ARIMA is utilized to come with a model with autoregressive value and the seasonal difference of 2. The given picture below is for forecasts from the last row(40). These values are the next 5 forecasted values for energy bills.

Forecasts from period 40

		95% Limits		
Period	Forecast	Lower	Upper	Actual
41	729.93	557.608	902.26	
42	672.48	316.667	1028.28	
43	684.09	219.657	1148.52	
44	913.55	334.579	1492.51	
45	1041.28	262.646	1819.90	

We will be storing the forecasted values and upper and lower bound values. A time series plot is utilized to check and verify forecasted values lie in the range of the lower and upper bounds.
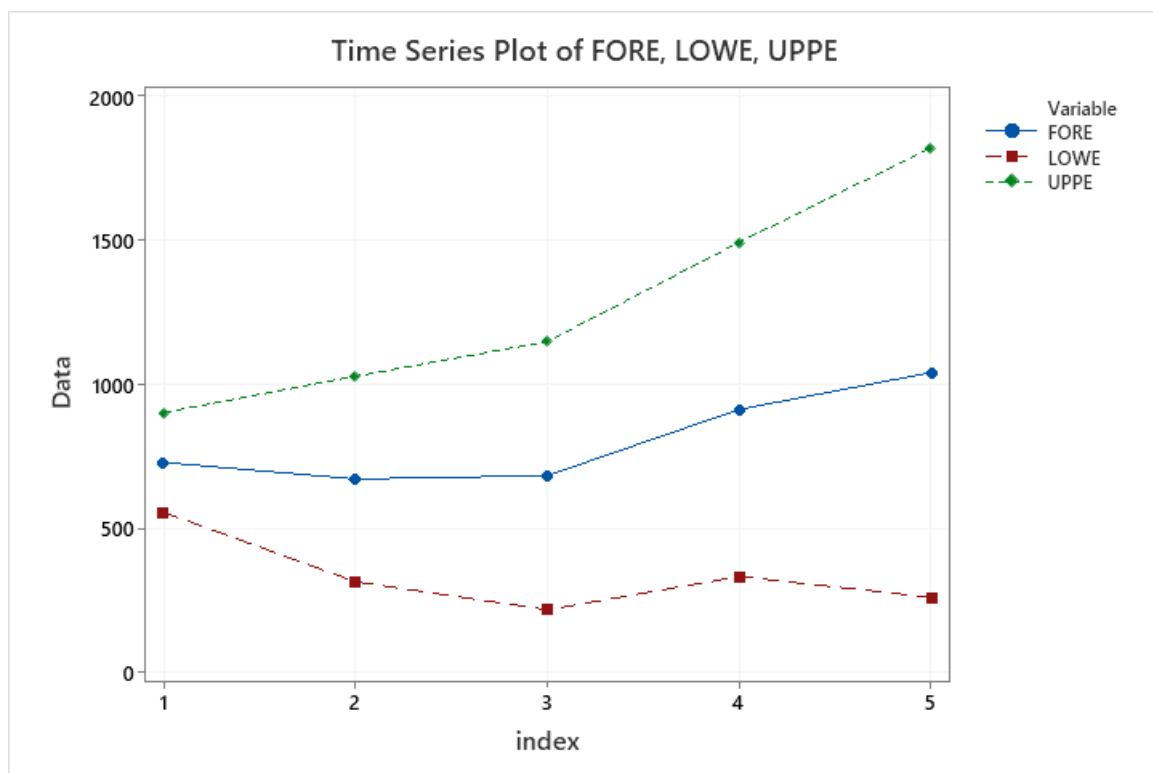

Figure 11 : Plot for forecast, lower and upper bound values