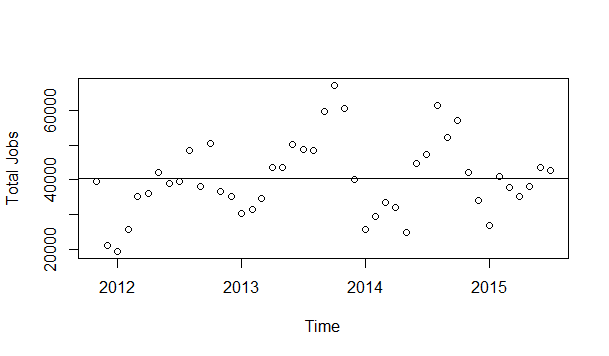
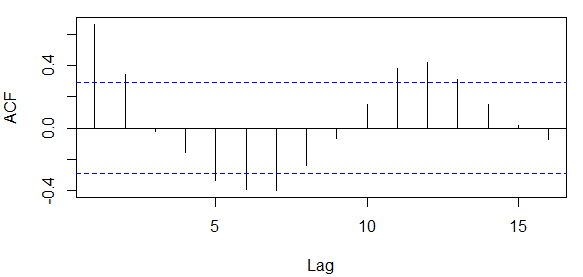
1. I used the useful variables, import the file “Synthetic Model.xlsx” into SAS, used SQL to group the data by month and year. Out put the data into “structured.csv”.

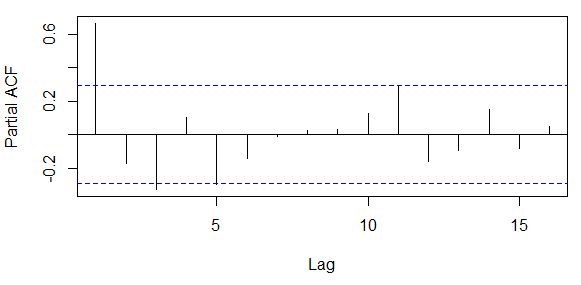
2. Values for all variables from and after August 2015 are extremely small compared to the majority of the observations, so I deleted them.

3. Import the data into R, graph below shows the Total Jobs over Time, no obvious significant trend, but seasonality was probably there.

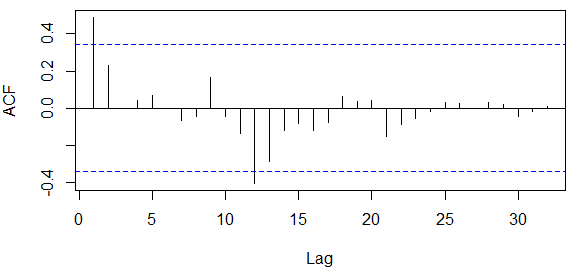


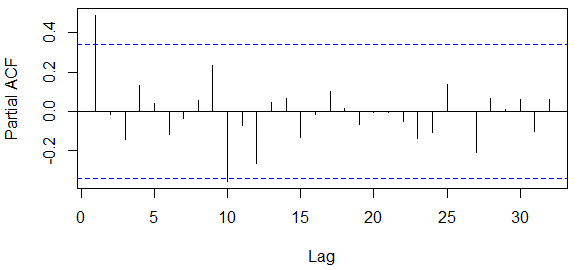
4. Looking at the ACF and PACF plots for the original data below, significant seasonal lags are shown, we want to take the seasonal difference for the data.



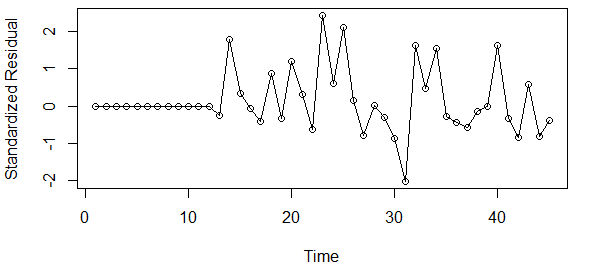


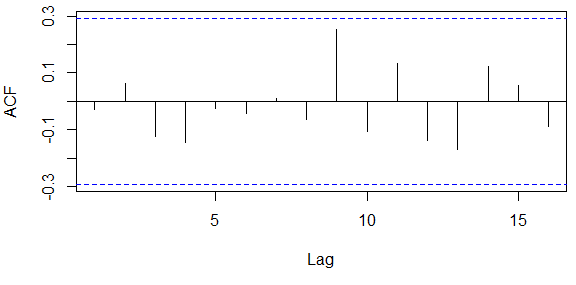
5. Took the first seasonal difference and computed ACF and PACF for the data. The regular orders chosen for the ARIMA model are (1,0,0), and the seasonal orders are (1,1,0)

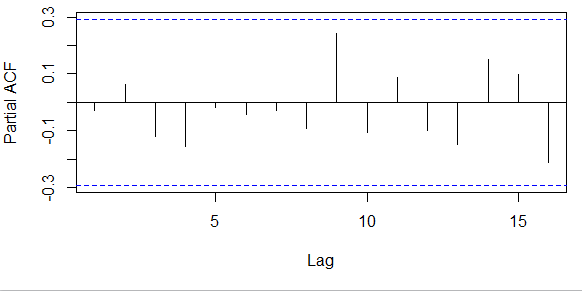




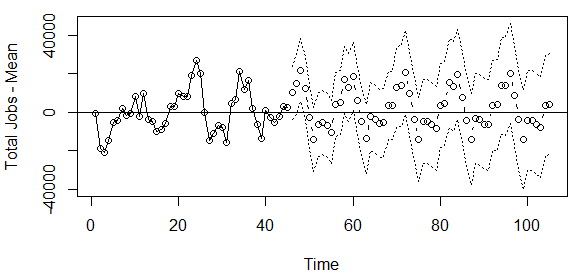
6. The standardized residuals from the ARIMA(1,0,0)x(1,1,0) are shown below, along with the corresponding ACF and PACF. No significant lags shown from both plots, this means we have a good model.



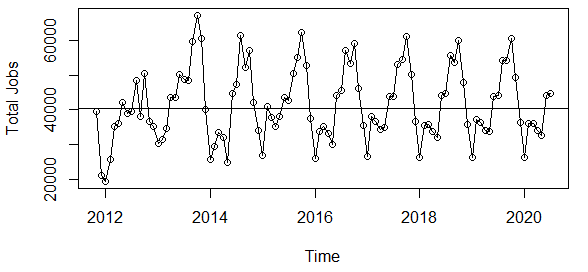




7. Since we deducted mean of Total Jobs before the model building, the graph below shows the predicted values and the 95% upper and lower confidence limit for the “Total Jobs - Mean”



8. Add the mean back gives the final plot below from 2011/7 all the way up to the 5 years predicted values of 2020/7 for Total Jobs.



9. Saved the predicted values as “Predicted Total Jobs.csv” file.