DATA ANALYTICS

ASSIGNMENT 6

ARMA AND ARIMA ANALYSIS

TEAM MEMBERS

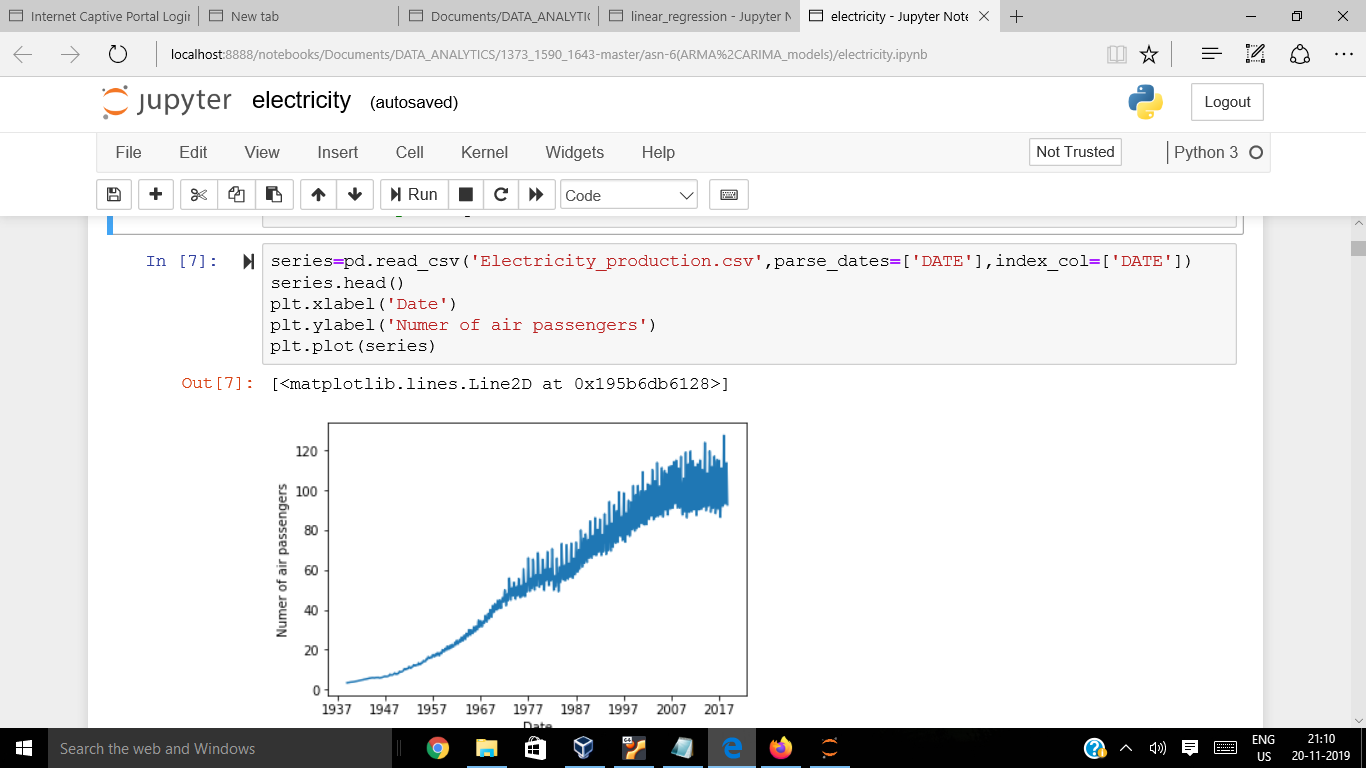
NAME SRN  
VIMARSHA R PES1201701590

VINUTHA M N PES1201701373

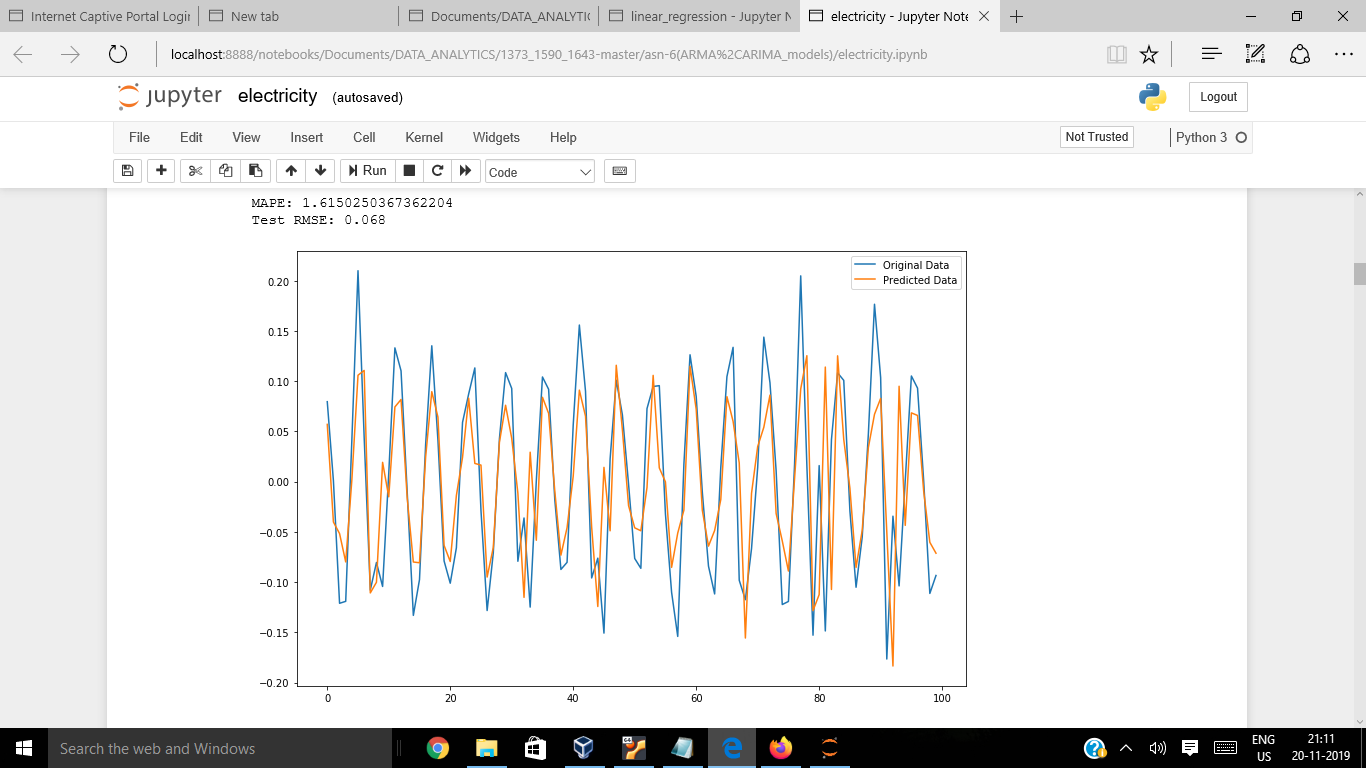
KEERTHI PRIYA P PES1201701643

***ARMA MODEL:*** An *ARMA model*, or Autoregressive Moving Average **model**, is used to describe weakly stationary stochastic time series in terms of two polynomials. The first of these polynomials is for autoregression, the second for the moving average.

***ARIMA MODEL:***  *ARIMA*, short for 'Auto Regressive Integrated Moving Average' is actually a class of **models** that 'explains' a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.



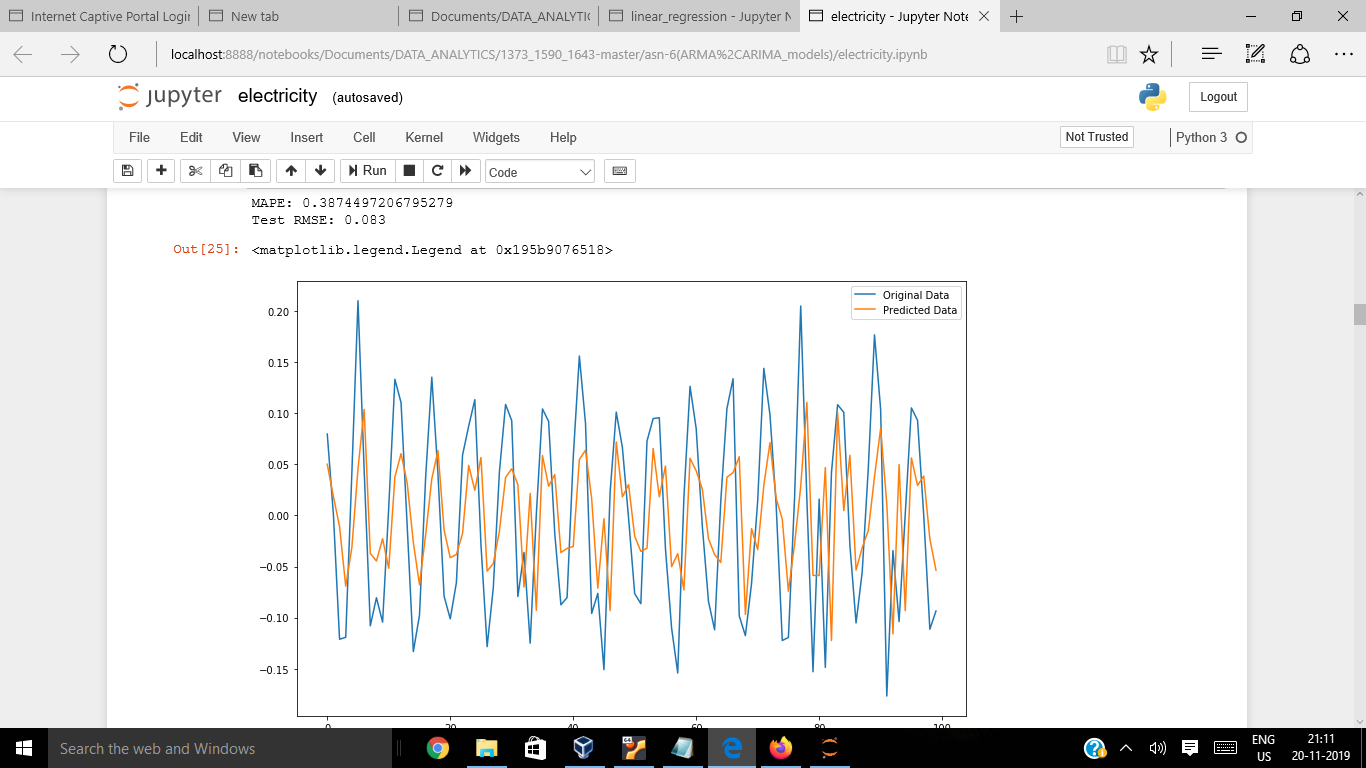
We are considering *Electricity\_Production* dataset, as per the above graph mean and variance are varying with time indicating that it is a ***non-stationary time series*** dataset.



This is for AR model with,

*MAPE* - 1.6150250367362204

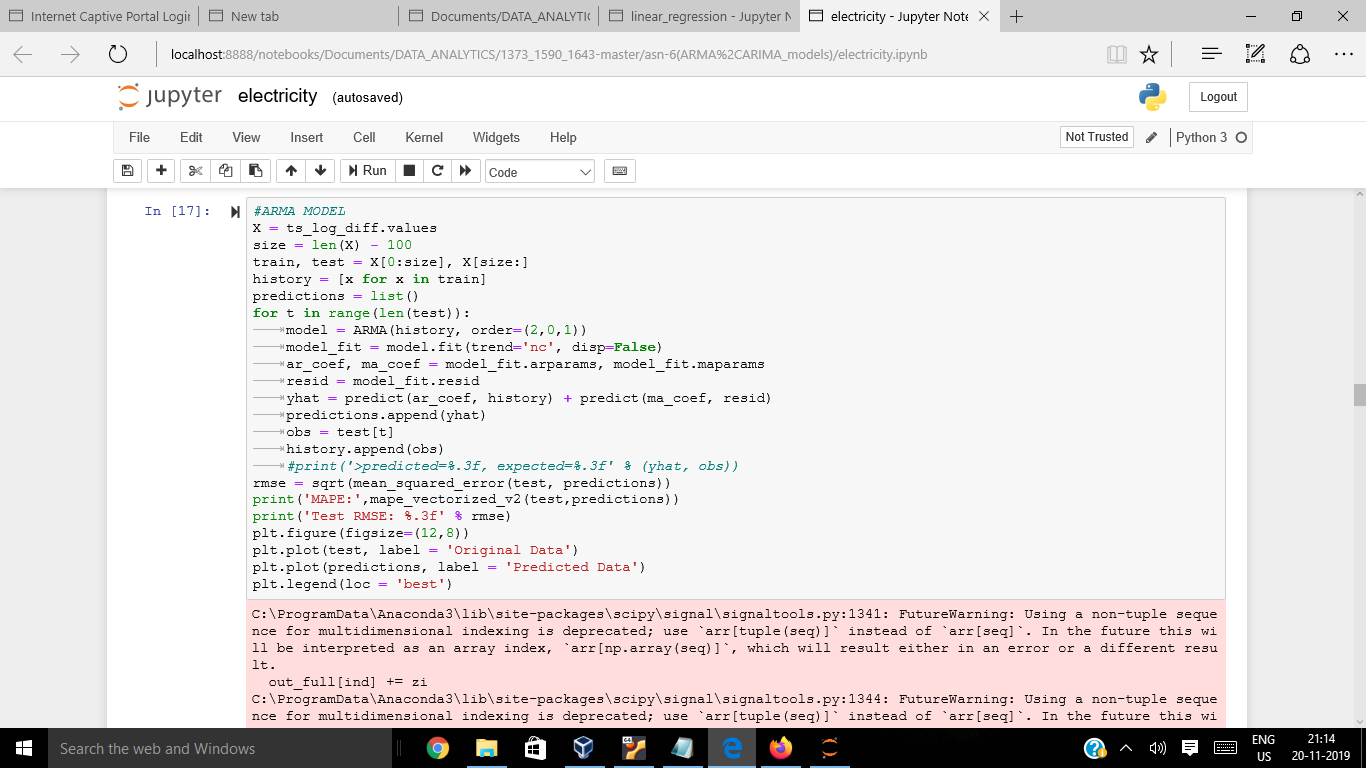
*Test RMSE* - 0.068

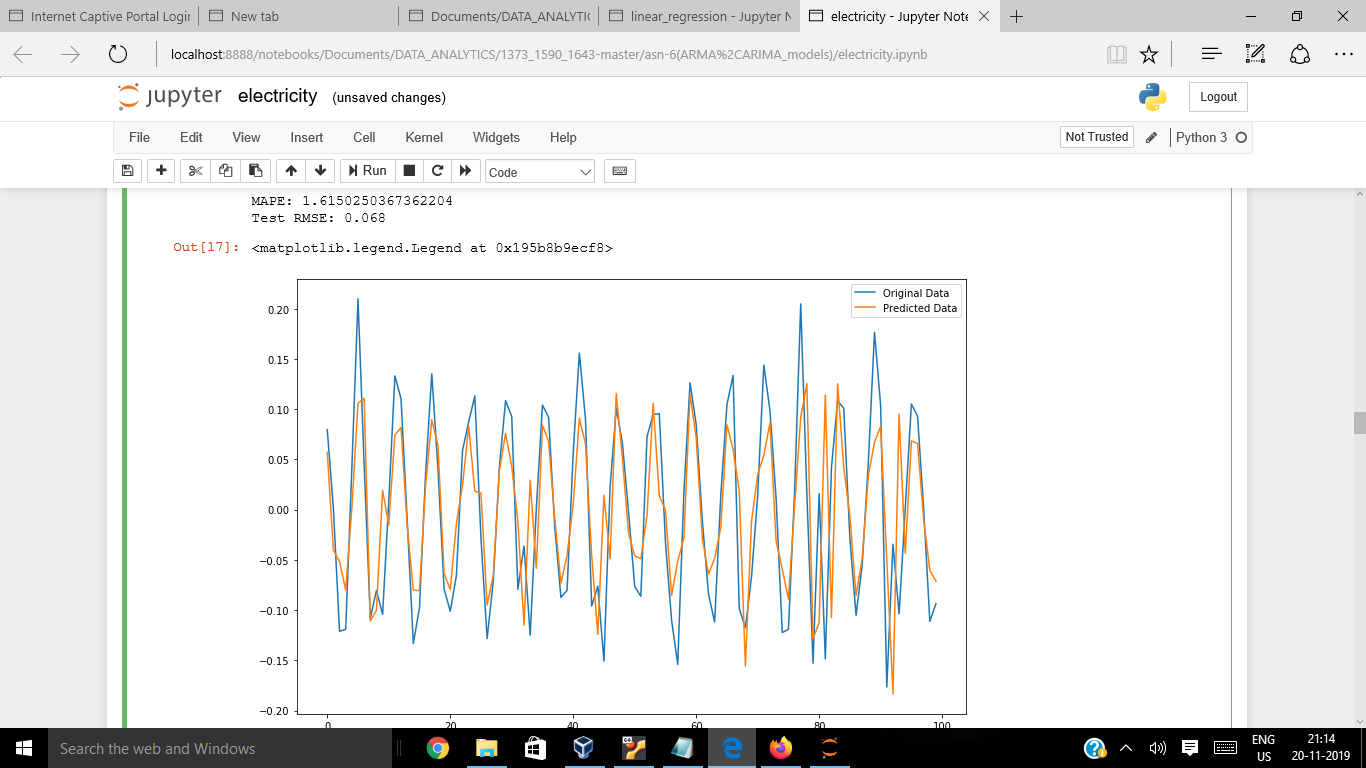


This is for MA model with,

*MAPE* - 0.387449

*Test RMSE* - 0.083

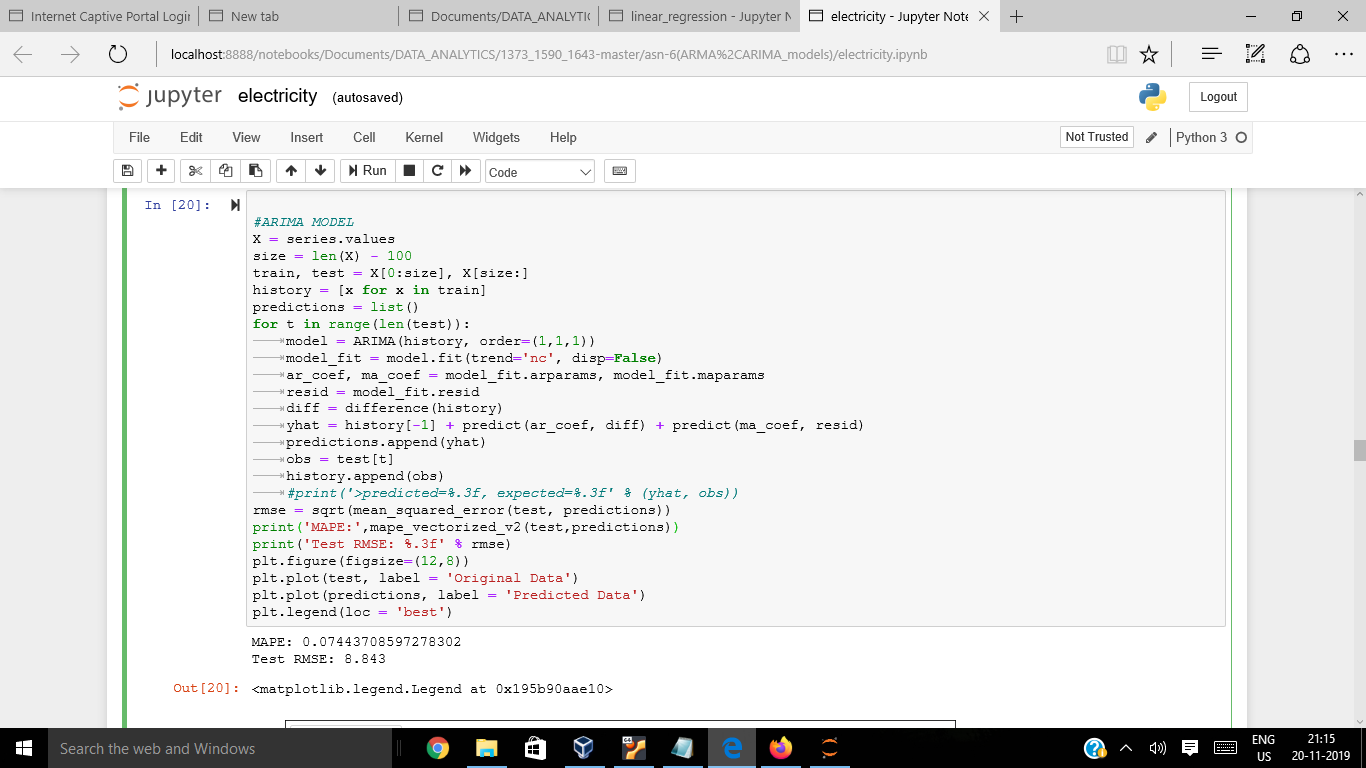


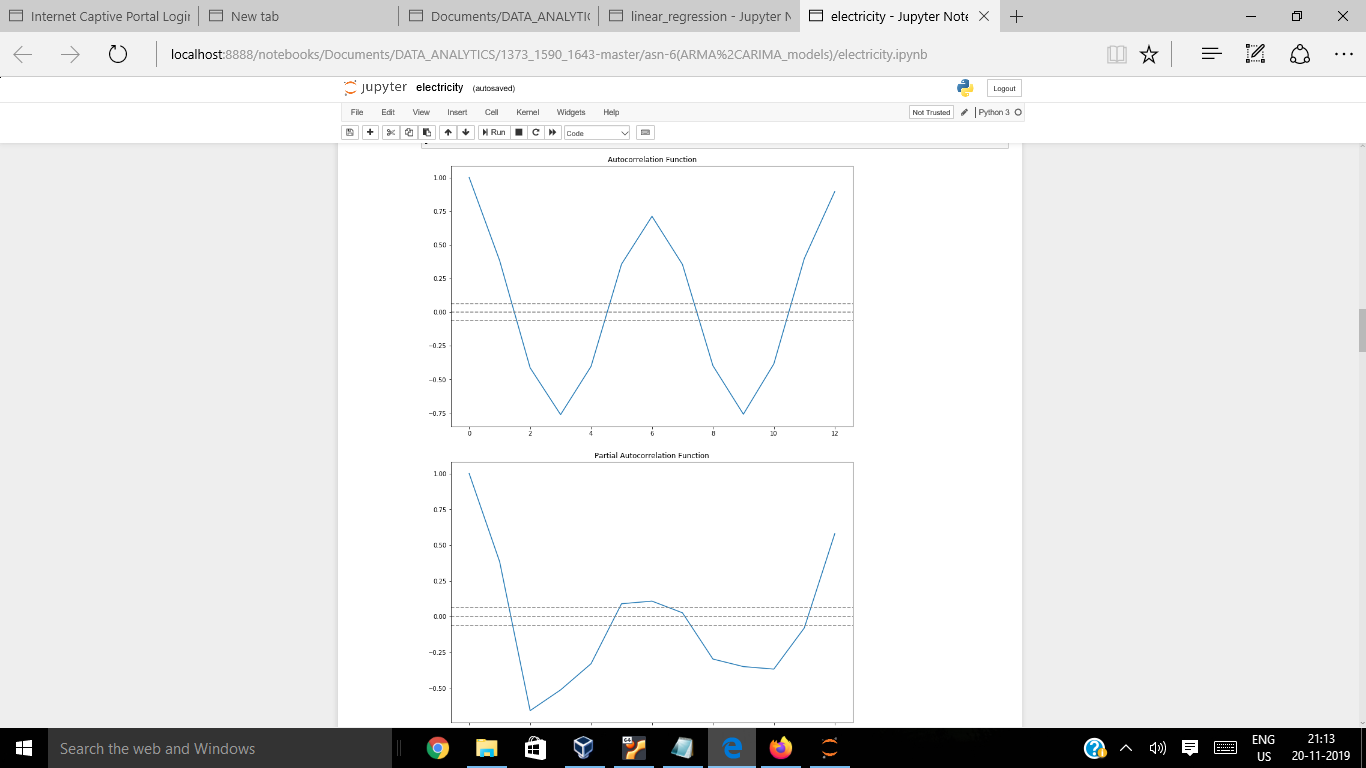


This is for ARMA model with,

*MAPE* : 1.6150250367362204

*Test RMSE* - 0.068





***Auto Correlation Function:***

**ACF** is an (complete) auto-correlation function which gives us values of auto-correlation of any series with its lagged values.

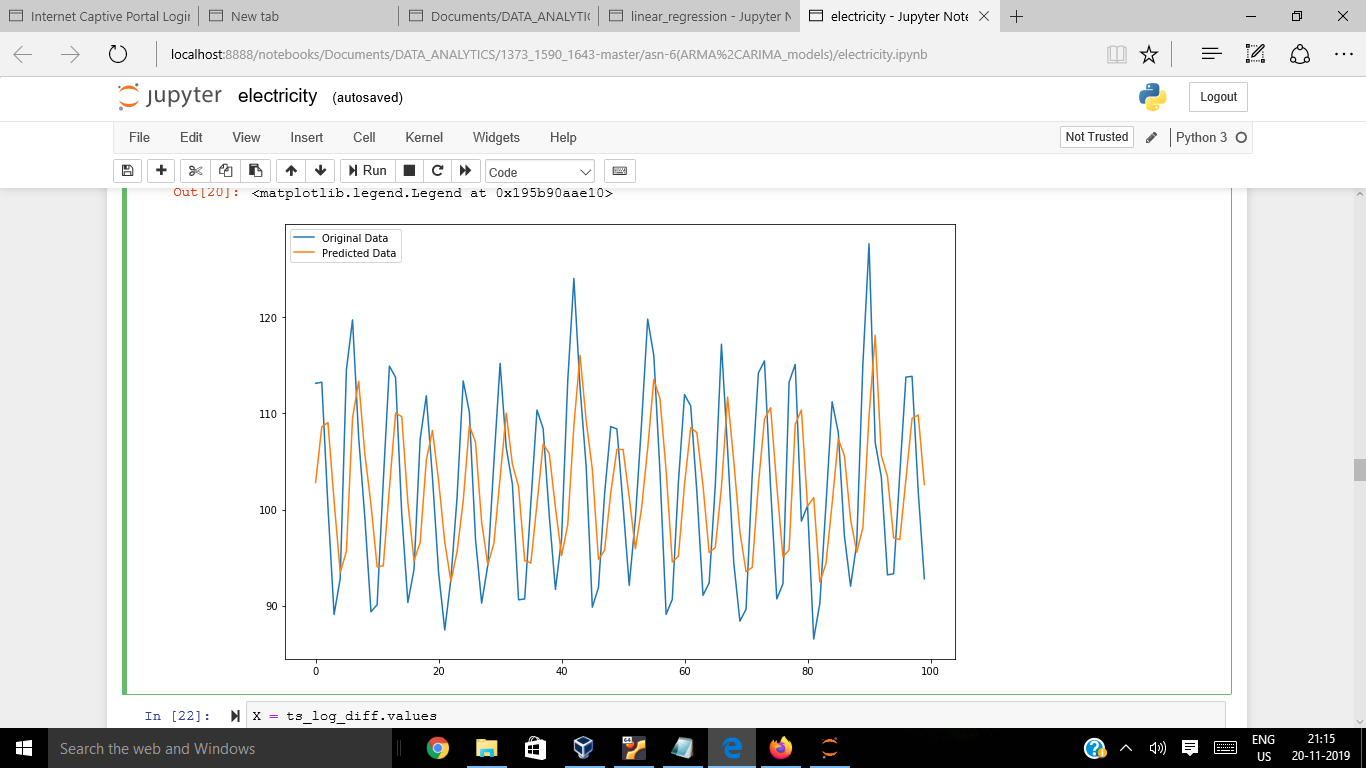
***Partial Autocorrelation Function:***

**PACF** is a partial auto-correlation function. Basically instead of finding correlations of present with lags like ACF, it finds correlation of the residuals (which remains after removing the effects which are already explained by the earlier lag(s)) with the next lag value hence ‘partial’.

*Graph shown below is for ARIMA model:*

*MAPE -*  0.07443708597278302

*Test RMSE -*  8.843



In most of the time series models , we use MAPE as an accuracy indicator, comparing MAPE values of all the models , we have got the ***least MAPE for ARIMA model*** , so ARIMA model is the best suited model to analyze the Electricity\_production non-stationary time series dataset.