Assignment 6

Define seasonality, trends and stationarity in time series.

Seasonality : Seasonal variation (or seasonality) are repeated cycles over time. If a cycle structure repeates at the same frequency, then it is seasonal . the seasonality is a periodic behavior that time series follow.

Trends:

Trends show variations of low frequency in a time series. Trends are long-term movement of a time series. A trend happens when a variable (in a long period – yearly or so) gets any changes (increase or decrease).

Stationarity:

if a time series has : 1- constant mean and 2-constant variance in time periods of our series and also 3- non seasonal . then it is called stationary. There are multiple statistical tests to check stationarity in time series , for example ADF test and KPSS test. We have used ADF technique for checking the stationarity of our data in the code.

I have used this data set for the time series assignment, this data set consisted of some feature, which the wind speed was selected from, the wind speed records were recorded over many cities, though I have used the Chicago wind speed over years 2014 to 2017

The data was recorded hourly, for visualization purposes we aggregate them to daily by the mean function

The index is set to the date, and the format is set to the datetime index of pandas, also null values are dealt with

Here is how the data looks before and after cleaning

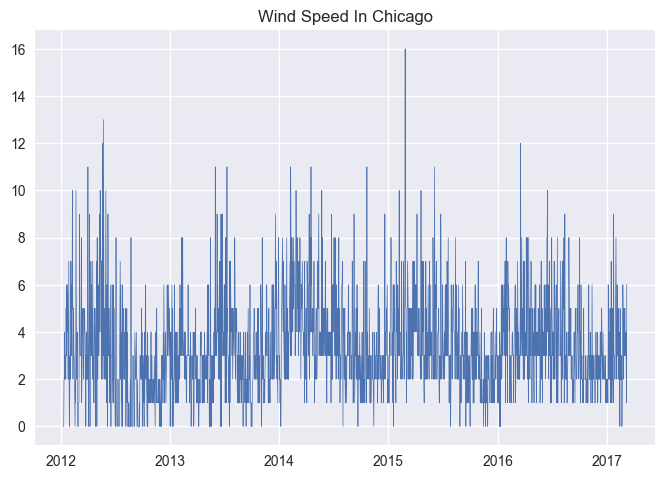


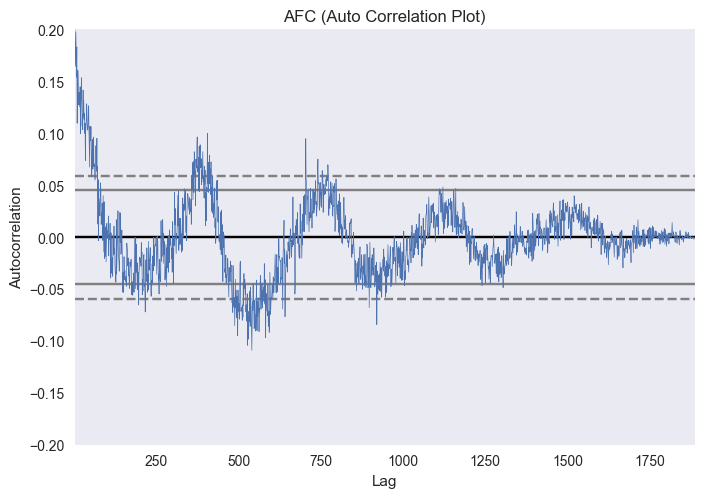
After :

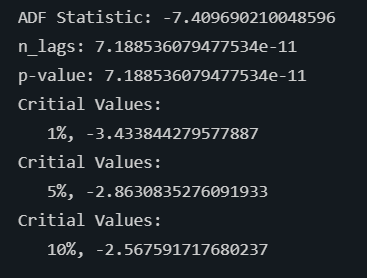


Now to check for the seasonality and the stationarity :

Lets first take a look at the time series itself:



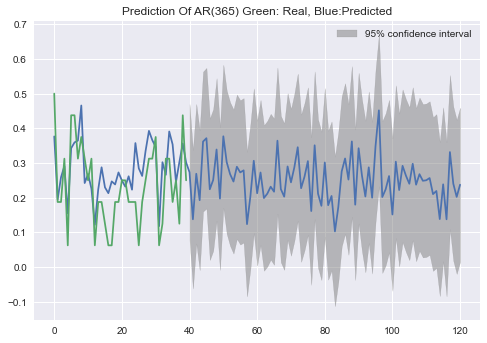
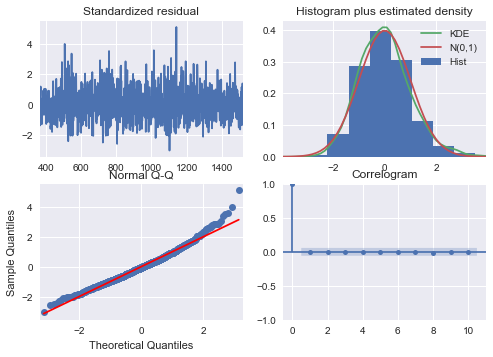
And now the auto correlation plot which reveals if our data has seasonality, seasonality here could be shown with high correlations of lags (above and beaneath the threshold) which is alsomt not seasonal, below we’ll see the ADF test result which infact proves that our series is stationary and thus we’re able to fit it to the Auto regression model



I have used AR and ARIMA model, to forecast the series,

The ar model : 365 lags are selected for the represtation of a year

Here are the diagnostics and the prediction chart:



Arima model :

