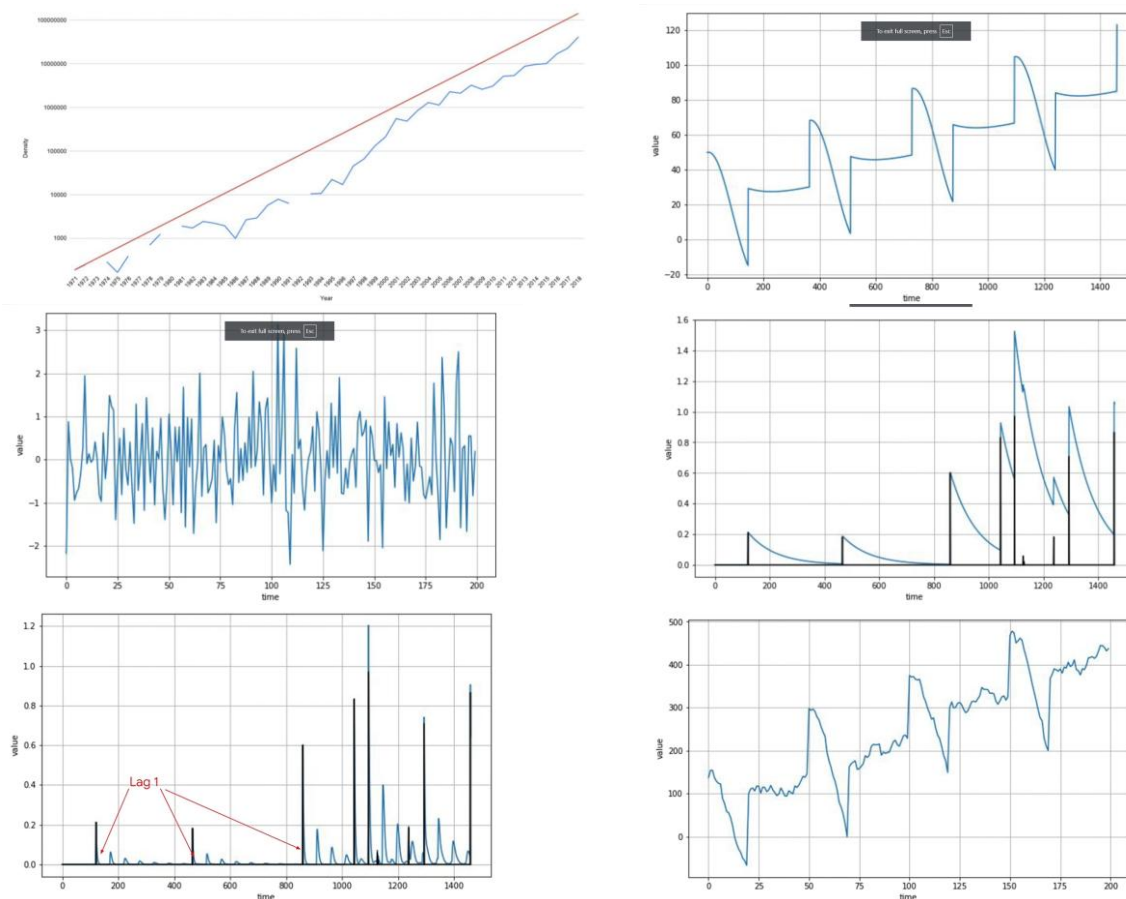


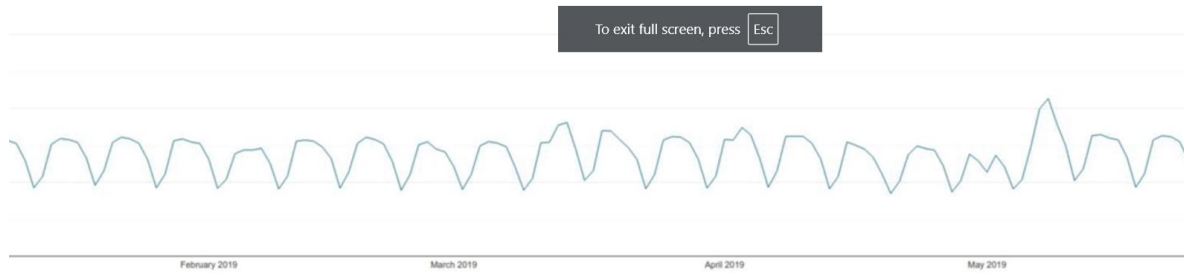
DeepLearning.AI Tensorflow Developer

Sequences, Time Series and Prediction: Week 1

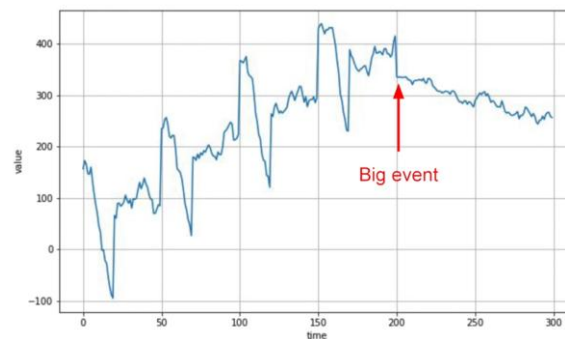
Time Series Data

Time series data have values progress over time, such as weather forecasting, stock price prediction. Time series data can be univariate, means single value for each time step or multivariate, means multiple values for each time step. Imputation also forecast missing values from time series data. Time series data also helps to detect anomaly and convert speech to text from sound waves. Time series data can have upward and downward trends or seasonality of specific intervals or both trends and seasonality. Time series data can even have autocorrelation and noise. Autocorrelation trends have deterministic decay with occasional spikes.





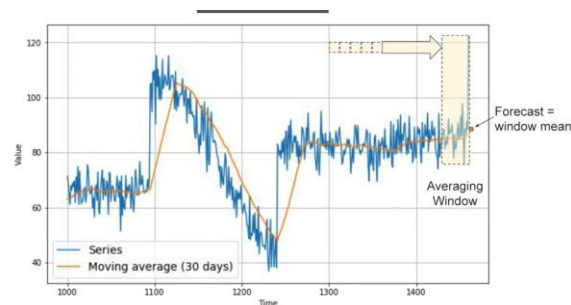
Machine learning models determine patterns on time series data for forecasting. Non stationary time series includes sudden change in pattern because of big events.



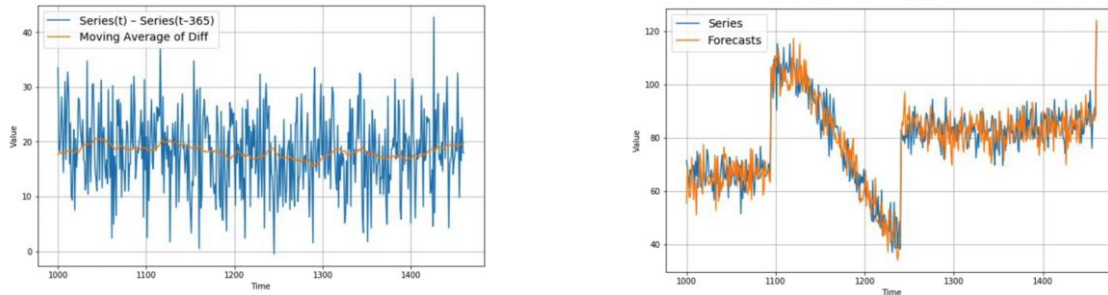
Stationary time series requires more data for accurate forecasting, whereas non stationary time series performs well with optimal training window which varies.

Partition time series dataset into training, validation and test period based on fixed partitioning. Train on both validation and test dataset to achieve optimal performance. In roll forward partitioning training period is increased as each week or month.

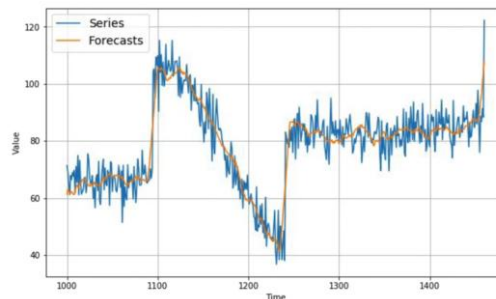
Metrics for evaluating performance of time series forecasting are usually RMSE or MAD or MAPE which can be computed using Keras metrics APIs. Common forecasting method is calculating the moving average with averaging window which removes noise but does not anticipate trend or seasonality.



Moving average technique can perform worse than naïve forecast resulting higher error values. Performance can be improved by removing trend and seasonality from time series data called differencing. Differencing technique works on the difference of time series data at time t and time $t-k$. Then apply moving average technique on difference time series data to get difference forecast. To get time series forecasting we add back time $t-k$ data.



Differencing slightly improves the MAE error values, although noise comes from past data. Further noise can be removed by smoothing both past and present data, hence by taking moving average of time $t-k$ data which also improves MAE error value.



Trailing window is used to calculate moving average of present values and centered window is used for past values. Centered window average gives accurate prediction than trailing window.