What is Time Series Analysis

Edris Safari

Bellevue University, Nebraska U.S.A.

Abstract

Time series analysis deals with data that are ordered in time. In other words, the data set has an attribute whose value is time based. This could be in nanoseconds, days, years or decades. In this paper, we will describe time series analysis and explain why time is important in data analysis.

What is Time Series Analysis

## Introduction

A time series is simply a set of data points that are ordered in time. For example, the frequency of google searches for the word 'diet' in a period of one year or the average annual temperature in Boston since 1830. Time series analysis is used in stocks and bonds and even crypto currency such as Bitcoin. There is also the concept of multiple timer series analysis such as Heating oil and natural gas prices in the course of the year. There is also data that are ordered in time but taken at a point in time. These are called cross-sectional data. Surveys and government records are common sources. In a time, series, time is typically the independent variable. The dependent variable can be any type and depends on the use case. In the financial sector, we are looking for future prices, trends, earnings, etc.

Essentially analysis of time-based data is like data set with discrete or continuous values whose frequency is not based on passage of time. As such, analysis of time-based data involves a model (i.e. mean squared), fitting the data into the model and using the model to make a prediction.

The simplest way to start the analysis is using the linear regression model. As shown below, in the linear regression model in a time-based data set, the dependent and independent variables as well as the error(noise) are all indexed by time.

$$Y_t = \alpha + \beta X_t + \epsilon_t$$

Other common models are moving average and exponentially moving average. More complex models for forecasting involve Fourier repetitive transforms to smooth out the data set to where reliable predictions could be made (see ref. 3).

The important aspects to consider in time series analysis are:

**Stationarity**

Time series that is stationary if its statistical properties do not change over time. In other words, it has constant mean and variance, and that the covariance is independent of time. Not all time series data are stationary. For example, the stock prices are affected by numerous attributes which make the mean and variance to change over time. However, there are transformational techniques (see ref 2) that can be used to make the data stationary

**Seasonality**

Periodic fluctuations in the data manifests seasonality of data. For example, the rush of google search for the word 'diet', ticks high after the holidays, but low during the holidays. Energy consumption is another example.

**Autocorrelation**

Autocorrelation deals with variations in time and the related pattern. The idea is that if certain value changes in time, there are occasions when there are high and low correlations between values.

The auto correlation plot shows the similarity between observations as a function of time lag between them. In the picture below, we see that there is high correlation between 1st, and 24th, 12th and 36th and 24th and 48th. This means that data has similar value every 24 unit of time (seconds, minutes, days, etc.) This is a valuable tool for making predictions or to detect anomaly (if the 24th value starts drifting or peeking).
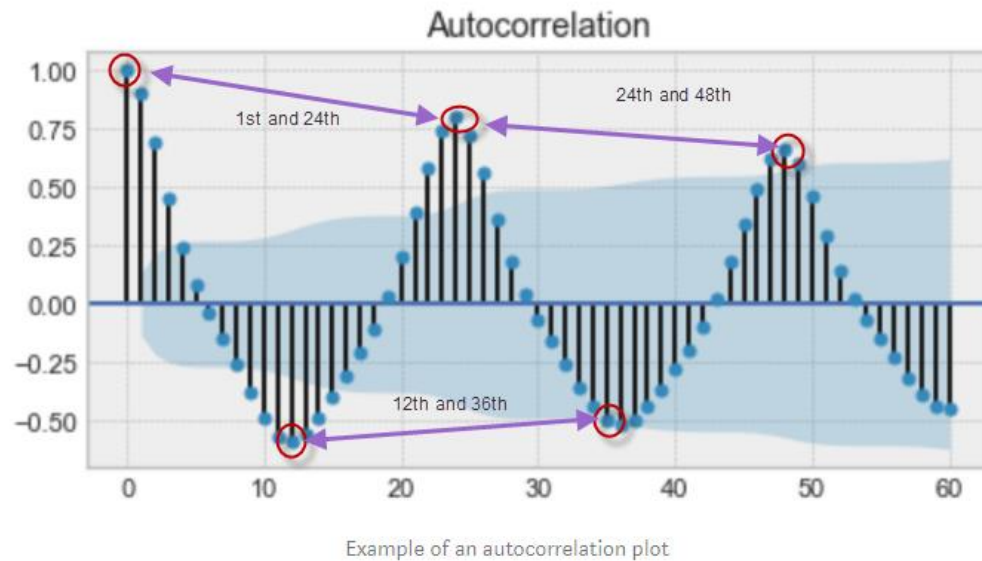
Example of an autocorrelation plot

*Figure 1- Autocorrelation graph*

## Conclusion

Time series analysis is concerned with predicting future values based on past values that are themselves based on passage of time. For example, the pressure in a chamber will fluctuate over time. These fluctuations are recorded and analyzed, and any future "abnormal" fluctuations can be predicted and thus prevented. Not all data sets are time-based, nor do they need be. The algorithms used to model non-time-based data may or may not be appropriate for time series data analysis and vice versa. It would depend on the use case and the dataset itself. However, in most cases, the elementary tools such as plotting, and simple modeling may put us on the right path.

## References

1. Downey, Allen B.. Think Stats: Exploratory Data Analysis . O'Reilly Media. Kindle Edition.

2. Two Effective Algorithms for Time Series Forecasting - https://www.youtube.com/watch?v=VYpAodcdFfA

3. https://www.datacamp.com/courses/introduction-to-time-series-analysis-in-python

4. The Complete Guide to Time Series Analysis and Forecasting - https://towardsdatascience.com/the-complete-guide-to-time-series-analysis-and-forecasting-70d476bfe775