BUBT | BANGLADESH UNIVERSITY OF BUSINESS AND TECHNOLOGY

Committed to Academic Excellence

Assignment on

Time Series Analysis and its Applications in Real-world

Course Code: CSE 475

Course Title: Data Mining

Submitted to:	Submitted by:
Name: Badhan Chandra Das Lecturer Dept. of CSE at Bangladesh University of Business and Technology.	Name: Syeda Nowshin Ibnat ID: 17183103020 Intake: 39 Section: 02 Program: B.Sc. in CSE Semester: Spring 2022

Date of Submission: 27-07-2022

1) Introduction: The Series of data points recorded over a specified period of time is called Time-series data. Time-series analysis is a technique for analyzing time series data and extract meaningful statistical information and characteristics of the data. One of the major objectives of the analysis is to forecast future value.

➔ **What is Time Series Analysis?**

Time series data is everywhere, since time is a constituent of everything that is observable. As our world gets increasingly instrumented, sensors and systems are constantly emitting a relentless stream of time series data. Such data has numerous applications across various industries. Time series analysis is a statistical method to analyze the past data within a given duration of time to forecast the future. It comprises of ordered sequence of data at equally spaced interval.

Example:

To understand the time series data & the analysis let us consider an example of Airline Passenger data. It has the count of passenger over a period of time.

Exploratory Analysis:

The first step is to perform the exploratory analysis which is carried out by plotting a line chart of the count of passengers against time. Figure 1 shows the count of passenger on y-axis and time on x-axis where each interval can be considered as a year.

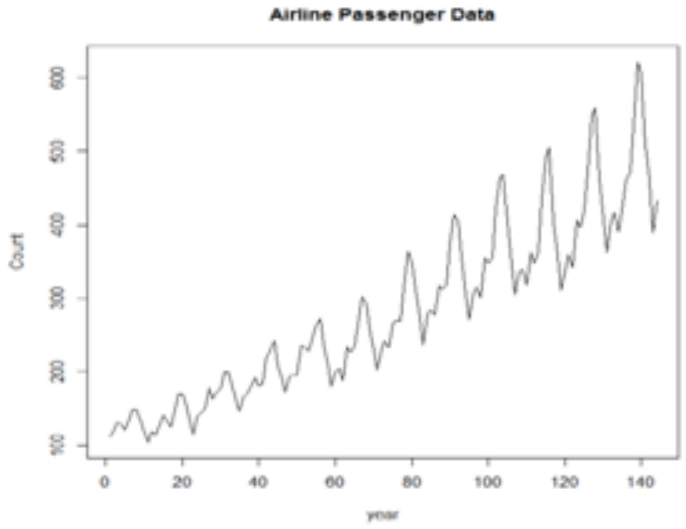

Figure 1: Exploratory Analysis of Airline passenger data

On careful observation of the given figure following observations can be derived:

a) **Trend:** Increasing or decreasing pattern has been observed over a period of time. In this case, the gradually increasing underlying trend is observed. i.e. the count of passengers has increased over a period of time.
b) **Seasonality:** Refers to cyclic pattern. A similar pattern that repeats after a certain interval of time. In the airline passenger example, we can observe a cyclic pattern that has a certain high & a low point which is visible in all the interval.
c) **Heteroscedasticity:** Refers to Non-constant variance or varying deflection from the mean over a period of time.

How Does Time Series Analysis Work?

➔ **Method and models of time series analysis.**

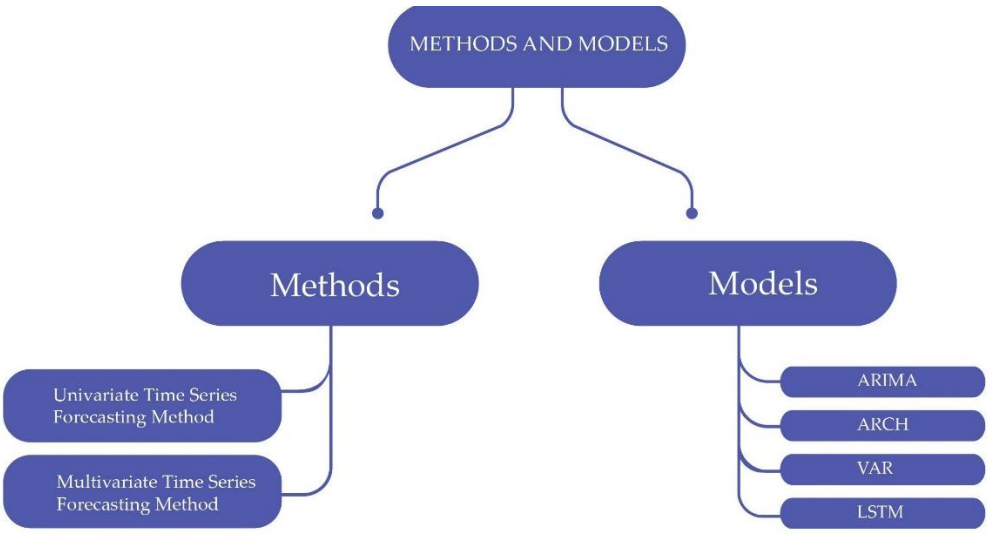

Figure 2: Method and models of time series analysis

Univariate time series: Only one variable varies over time. for example, data collected from a sensor measuring the temperature of a room every second.

Multivariate time series: There will be more than one variable to work with.

ARIMA Model: It is a combination of three different models itself, AR, MA and I, where "AR" reflects the evolving variable of interest is regressed on its own prior values, "MA" infers that the

regression error is the linear combination of error terms values happened at various stages of time priory, and "I" shows the data values are replaced by the difference between their values and the previous values. Combinedly "ARIMA" tries to fit the data into the model, and also ARIMA depends on the accuracy over a broad width of time series.

ARCH/GARCH Model: Being the extended model of its common version GARCH, Autoregressive Conditional Heteroscedasticity (ARCH) is the most volatile model for time series forecasting, and are well trained for catching dynamic variations of volatility from time series.

Vector Autoregressive Model or VAR model: It gives the independencies between various time-series data which as a generalization of the Univariate Autoregression Model.

LSTM: Long-short term memory(LSTM)is a deep learning model, it is a kind of Recurrent Neural Network(RNN) to read the sequence dependencies.

2) Real-world applications of Time series analysis:

Weather records, economic indicators and patient health evolution metrics all are time series data. Time series data could also be server metrics, application performance monitoring, network data, sensor data, events, clicks and many other types of analytics data.

We can notice how time depicted at the bottom of the below chart is the axis:

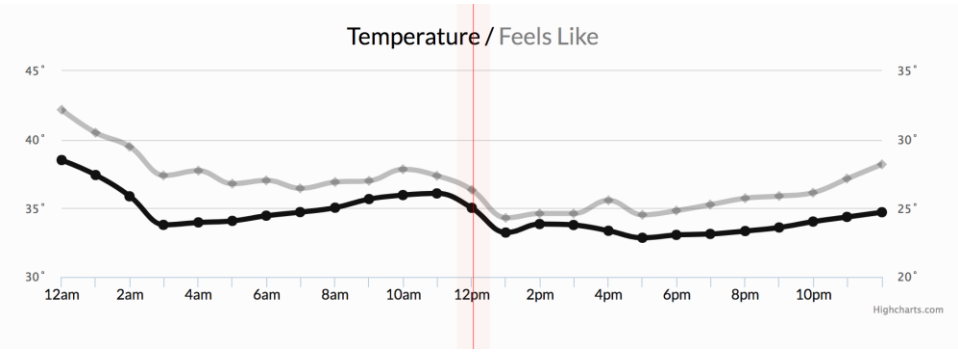

Figure 3: Weather conditions

In the next chart below, note time as the axis over which stock price changes are measured. In investing, a time series tracks the movement of data points, such as a security's price over a specified period of time with data points recorded at regular intervals. This can be tracked over the short term (such as a security's price on the hour over the course of a business day) or the long term (last day of every month over the course of five years).

Figure 4: Stock exchange

The cluster monitoring example below, depicting disk ops write and usage data, would be familiar to Network Operation Center teams.

Figure 5: Cluster monitoring

Another familiar example of time series data is patient health monitoring, such as in an electrocardiogram (ECG), which monitors the heart's activity to show whether it is working normally.

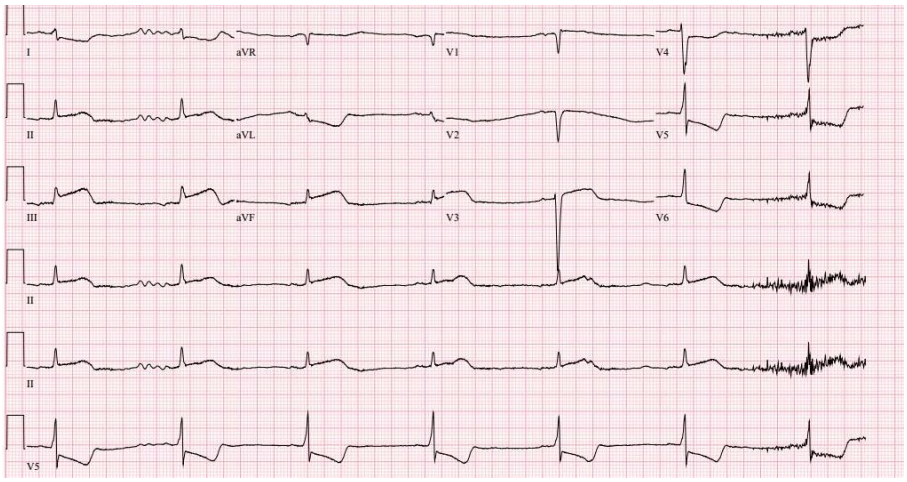

Figure 6: Health monitoring

In addition to being captured at regular time intervals, time series data can be captured whenever it happens regardless of the time interval, such as in logs. Logs are a registry of events, processes, messages and communication between software applications and the operating system. Every executable file produces a log file where all activities are noted. Log data is an important contextual source to triage and resolve issues. For example, in networking, an event log helps provide information about network traffic, usage and other conditions.

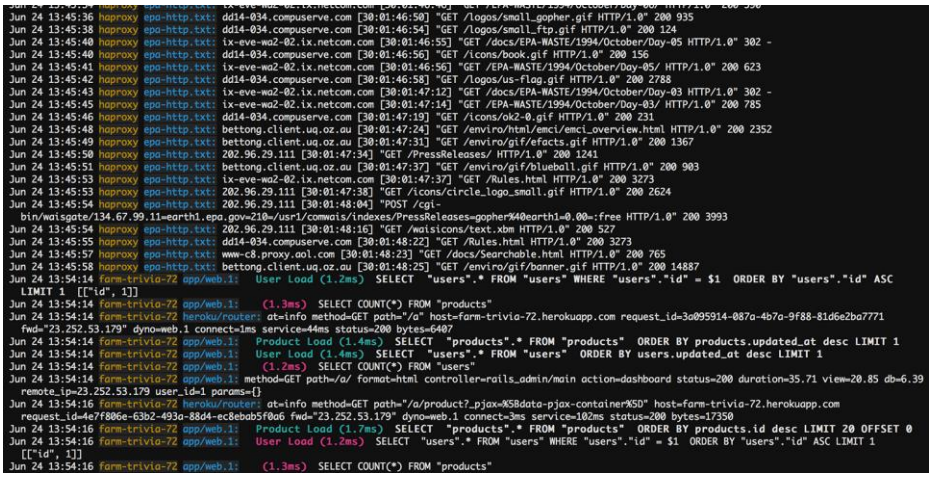

Figure 7: Logs

Traces (a list of the subroutine calls that an application performs during execution) are also time series data. Over the colored bands in the traces chart below, you can see examples of time series

data. The goal of tracing is to follow a program's flow and data progression. Tracing encompasses a wide, continuous view of an application to find bugs in a program or application.

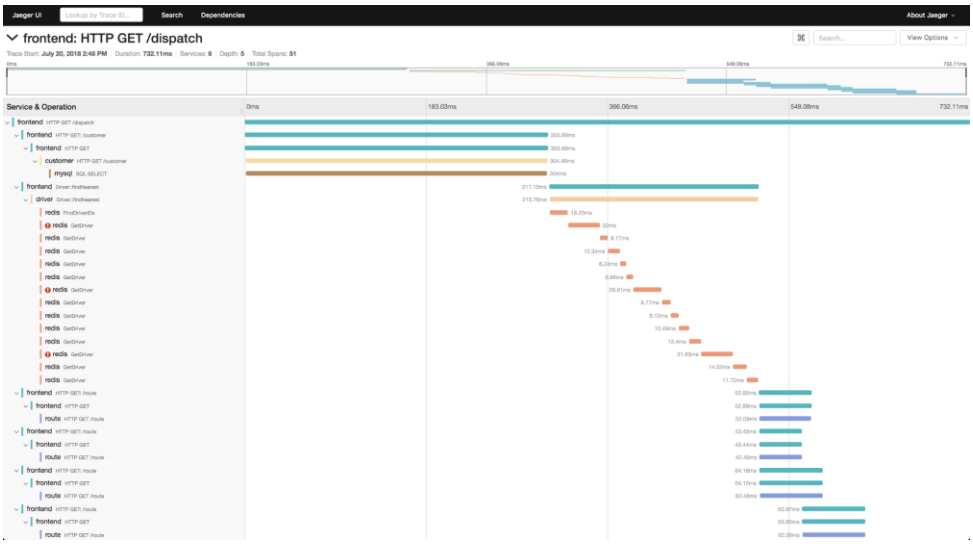

Figure 8: Traces

3) Conclusion:

The primitive decisions were made on the basis of gut feelings and common sense. A systematic and collaborative approach to make a decision supported by the data is a real game-changer. With a wide range of applications. It has become one of the most important areas of study. It plays a crucial role in understanding the underlying structure of the time series data with aid in extracting meaningful statistical & characteristic information and henceforth the decision making backed by the data.

References

[1] https://www.influxdata.com/what-is-time-series-data/

[2] https://www.influxdata.com/what-is-time-series-data/

[3] https://www.educba.com/time-series-analysis/

[4] https://www.investopedia.com/terms/t/timeseries.asp