Batch: D2	Roll No.: 16010122323
Experiment 05	

Title: Working with time-series data

Objective:

1. *Search/locate and download the time series Data*
2. *To learn how to visualize time series data in Tableau*
3. *Applying trend line over visualized time series data using analytics options in tableau*
4. *Perform forecast over the time series using analytics options in tableau*

Course Outcome:

CO1: Learn how to locate and download datasets, extract insights from that data and present their findings in a variety of different formats.

CO3 Apply data visualization best practices

Books/ Journals/ Websites referred:

1.   Data Visualization made simple New York: Routledge - Kristen Sosulski, First edition, 2019

2.   Sosulski, K. Data Visualization Made Simple: Insights into Becoming Visual, First edition, 2018

3.   https://www.tableau.com/learn/articles/time-series-analysis   4. https://www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm

Resources used:

DataSet: Total   female   births   in   California https://www.kaggle.com/dougcresswell/daily-total-female-births-in-california-1959

Theory:

Definition Time series:

Time series forecasting is a critical requirement for many organizations. It also helps to compare multiple dimensions over time, spot trends, and identify seasonal patterns in the data.
Time series analysis is a statistical technique used to record and analyze data points over a period of time, such as daily, monthly, yearly, etc.
A time series chart is the graphical representation of the time series data across the interval period.

Examples: stock market analysis, population trend analysis using a census, or sales and profit trends over time.

Time series analysis helps organizations understand the underlying causes of trends or systemic patterns over time. Using data visualizations, business users can see seasonal trends and dig deeper into why these trends occur. With modern analytics platforms, these visualizations can go far beyond line graphs.

In time series quantitative data are arranged in the order of their occurrence and resulting statistical series.

There are 4 components in a time series:

These four components are:
- Secular trend, which describe the movement along the term;
- Seasonal variations, which represent seasonal changes;
- Cyclical fluctuations, which correspond to periodical but not seasonal variations;
- Irregular variations, which are other non random sources of variations of series.

In the additive model, we represent a particular observation in a time series as the sum of these four components.
In this model, four components have a multiplicative relationship. So, we represent a particular observation in a time series as the product of these four components

Observations

The data set chosen for this experiment was "Daily Total Births of Female in California". The California Female Births dataset is a time series situation where we are trying to forecast future outcomes based on past data points. The problem is to forecast the daily number of female births in California. The dataset described a time-series of baby births over 12 months in 1959, and there are 365 observations.

The first graph gives a simple time series. It shows the Daily total female births in California Vs Day. As the chart shows data is highly volatile. Its been varying from 30 – 50 births in a day.
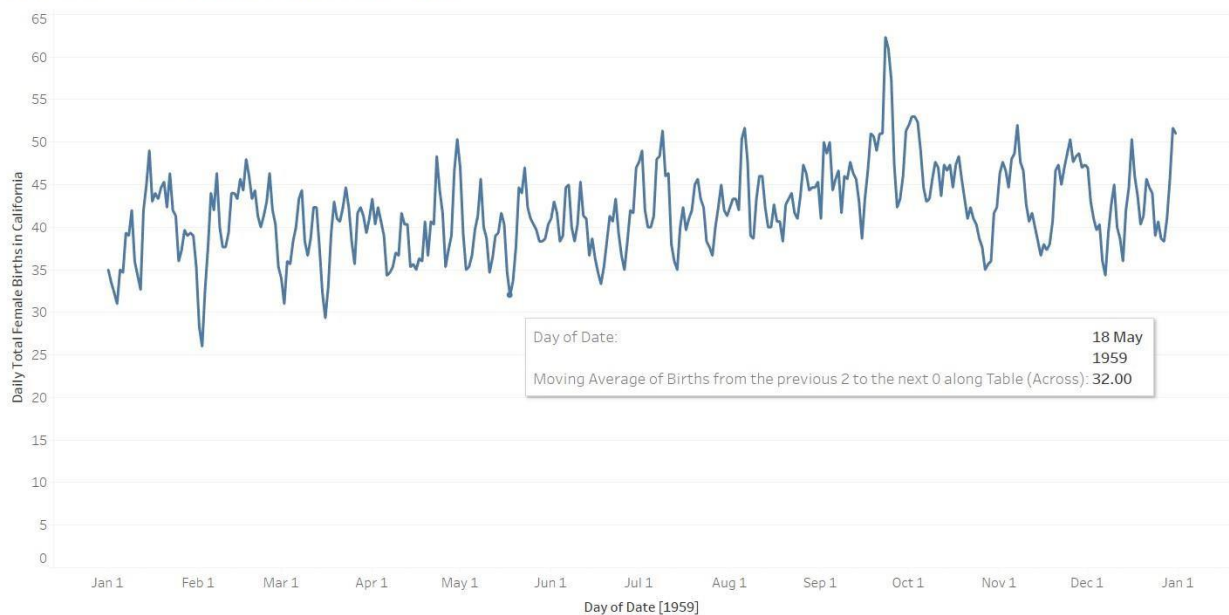
P.T.O

Time Series: Total Female Births in California
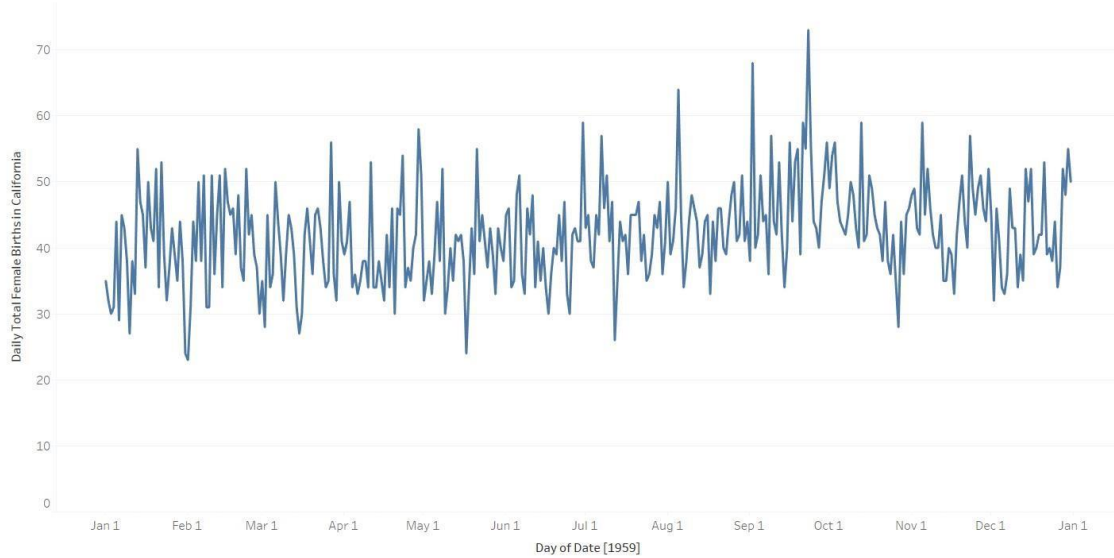
The next graph shows the moving average of Daily births vs Day graph. Moving Average gives a good idea of the temperature variations over the last years if we want to consider the reasons for change. It smooths out the noise by filtering short term variations.

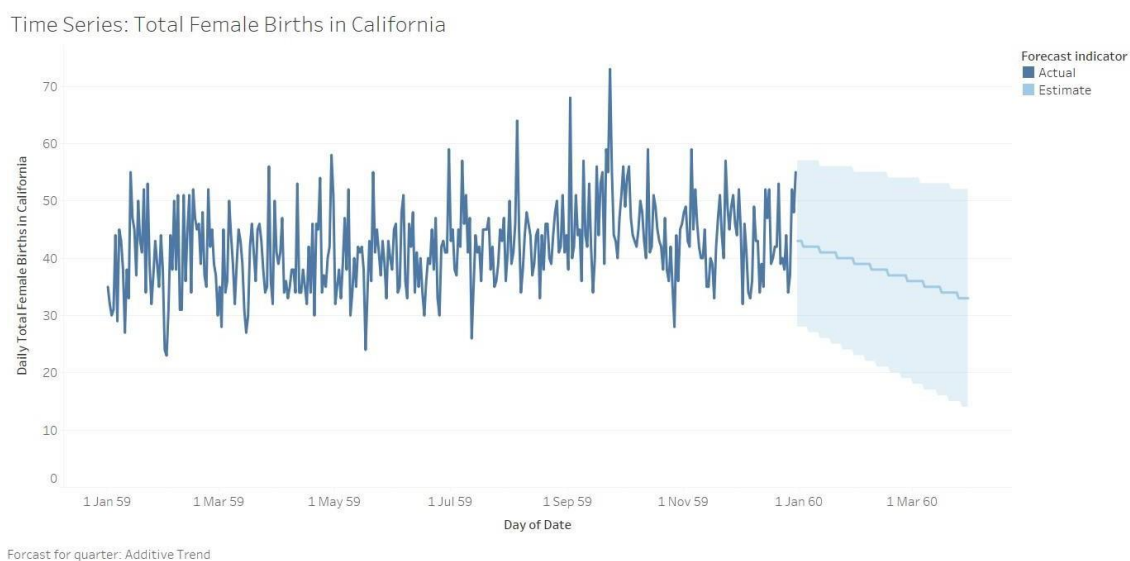Time Series: Total Female Births in California

The third graph shows a trendline (polynomial) which can be used to understand the trend over the years and also calculate the approximate births in the coming years. As we can see the trend line is gentle and shows low overall variations.
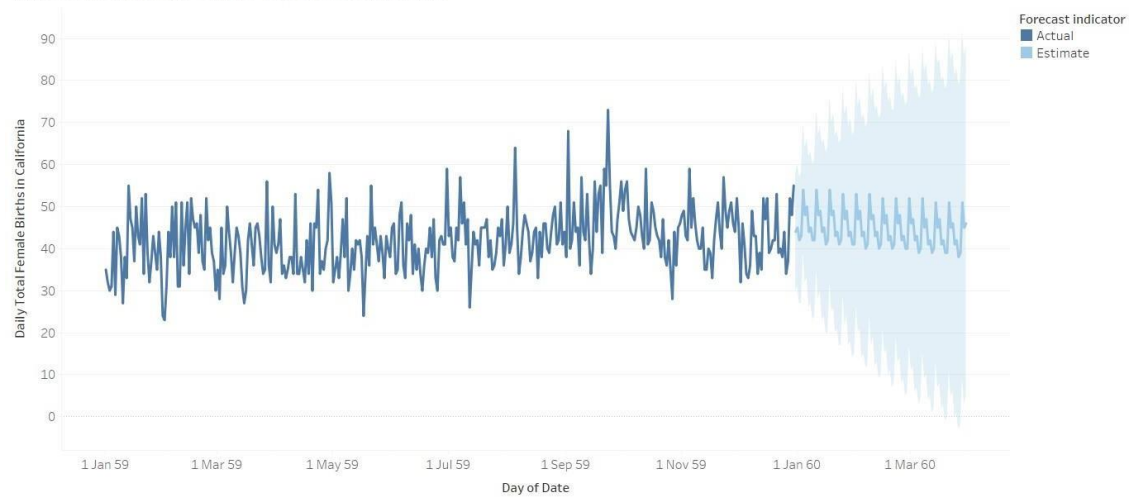
Time Series: Total Female Births in California
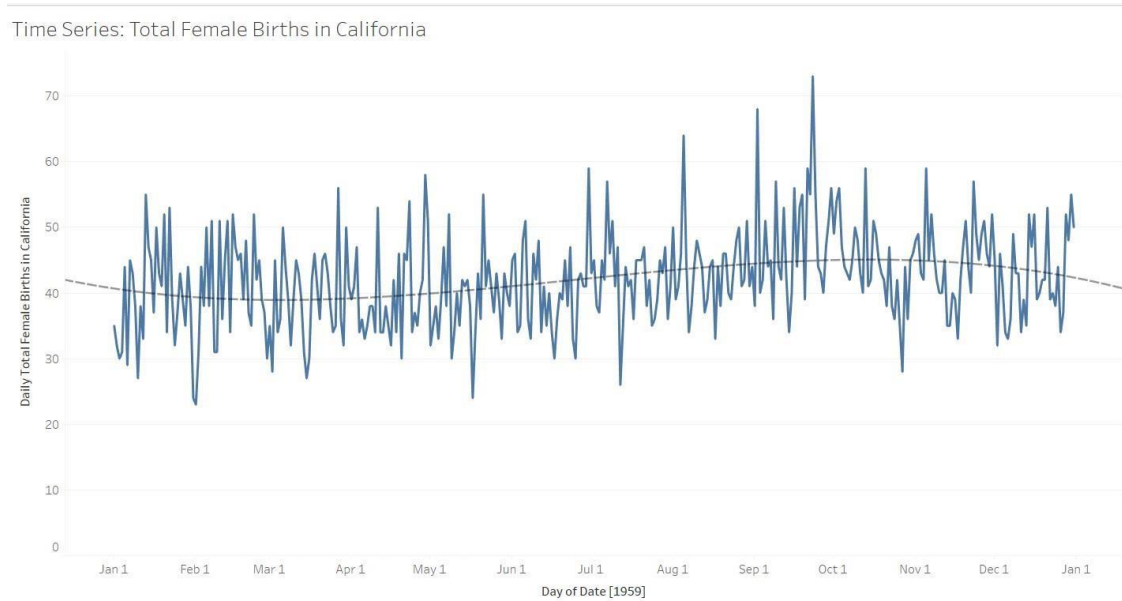
The next 4 graphs give a forecast of Births vs Day for the next quarter. The four forecasts take into consideration 4 models:

- Additive Trend and Seasonality,
- Additive Trend and Additive Seasonality
- Additive Trend and Multiplicative Seasonality,
- Multiplicative Trend and Multiplicative Seasonality

The additive trend shows a decrease in no of births over the next quarter but rest of the model show a rather volatile yet average forecast for the next quarter, which shows that time is not affecting the average births.


Time Series: Total Female Births in California

Forcast for quarter: Additive Trend

Time Series: Total Female Births in California

Forecast indicator
- Actual
- Estimate

Forcast for quarter: Additive Trend, Additive Seasonality

Time Series: Total Female Births in California

Forecast indicator
- Actual
- Estimate

Forcast for quarter: Additive Trend, Multiplicative Seasonality

Time Series: Total Female Births in California

Forecast indicator
- Actual
- Estimate

Forcast for quarter: Multiplicative Trend, Multiplicative Seasonality
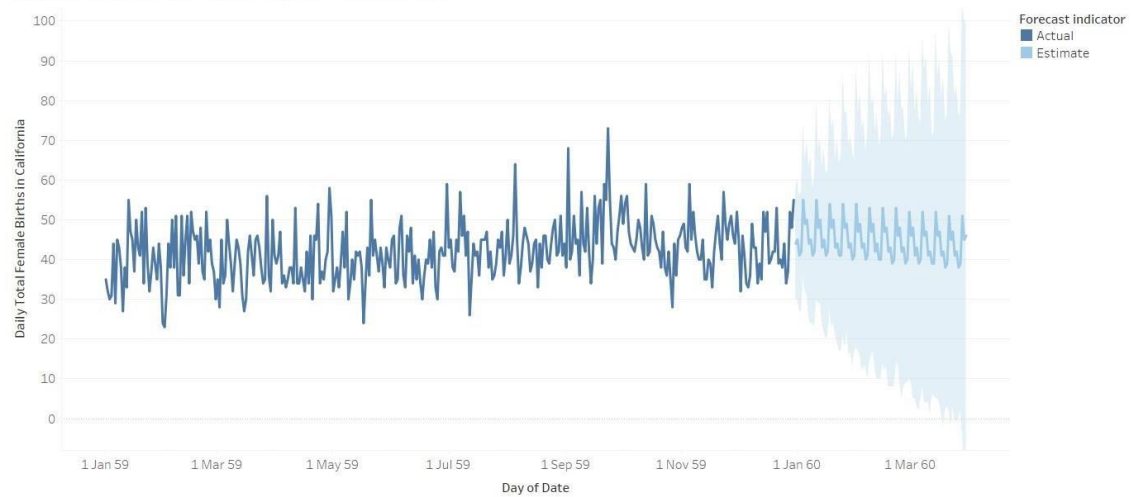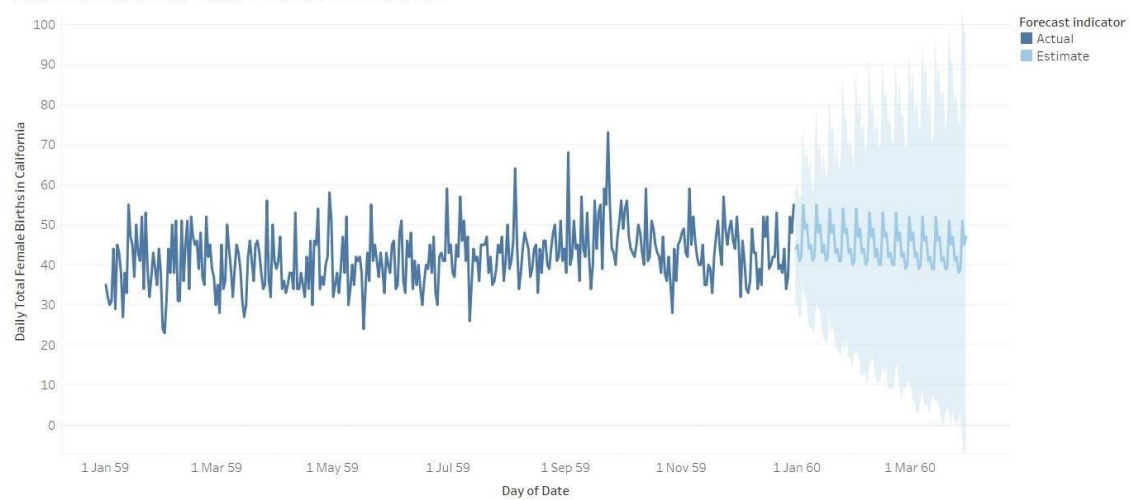
Conclusion

Through this experiment we learnt the concept of time series and implemented the same on tableau. Tableau gives many options for visualization of a time series data. This makes working with the time series data easier and it also reduces the calculations and is user friendly.

Date: 14-11-2021   **Signature of faculty in-charge**

Post Lab Question:

1. Compare the additive and multiplicative model of time series

- In the additive model, we represent a particular observation in a time series as the sum of components (seasonality, trend, cyclical and residual). It is used where change is measured in absolute quantity. Here the data is modelled as it is.
- In the multiplicative model, we represent a particular observation in a time series as the product of components. It is Used where change is measured in percent(%) change. Here the data is additive but after taking logarithmic value.