SESSIONS 5

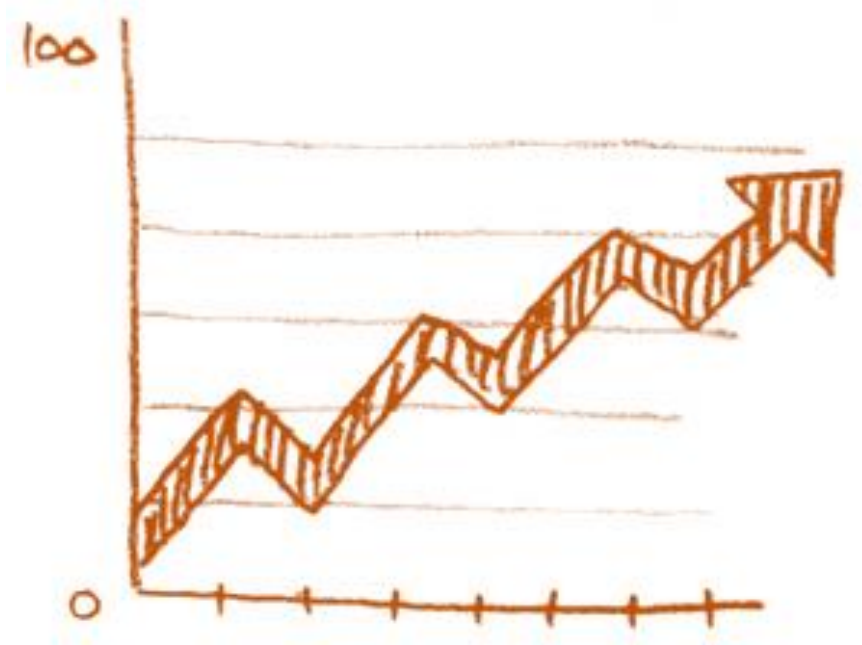
Line Plot

Data Science Program

Purwadhika
Startup and Coding School

Outline

- What is Line plot?
- When to use Line plot
- Create Line plot using Matplotlib, Seaborn, and Pandas



Purwadhika
Startup and Coding School

What is Line Plot?

Purwadhika
Startup and Coding School

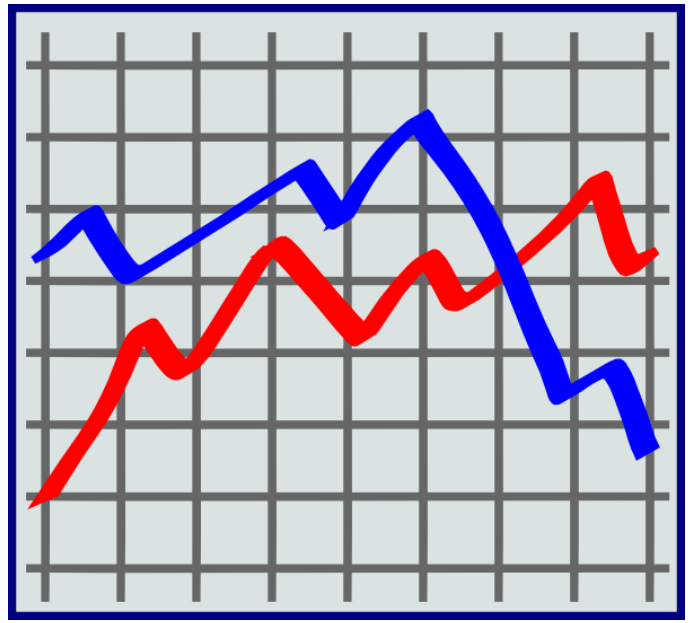
What is Line plot?

- A line chart or line plot is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments.
- It is a basic type of chart common in many fields. It is similar to a scatter plot except that the measurement points are ordered (typically by their x-axis value) and joined with straight line segments.
- A line chart is often used to visualize a trend in data over intervals of time – a time series – thus the line is often drawn chronologically.
- In these cases, they are known as run charts.

Purwadhika
Startup and Coding School

What is Line plot?

- A line chart is, as one can imagine, a line or multiple lines showing how single, or multiple variables develop over time.
- It is a great tool because we can easily highlight the magnitude of change of one or more variables over a period.



Purwadhika
Startup and Coding School

When to Use Line plot?

Purwadhika
Startup and Coding School

When to Use Line plot?

- **A simple line chart** would have been much better. A line chart allows us to track the development of several variables at the same time. It is very easy to understand, and the reader doesn't feel overwhelmed.
- **Time series** is a line plot and it is basically connecting data points with a straight line. It is useful in understanding the trend over time. It can explain the correlation between points by the trend. An upward trend means positive correlation and downward trend means a negative correlation. It mostly used in forecasting and monitoring models.
- **When to use**: Time Series should be used when single or multiple variables are to be plotted over time.
- Ex: Stock Market Analysis of Companies, Weather Forecasting.

When to Use Line plot?

Time Series – AAPL vs IBM

Purwadhika
Startup and Coding School

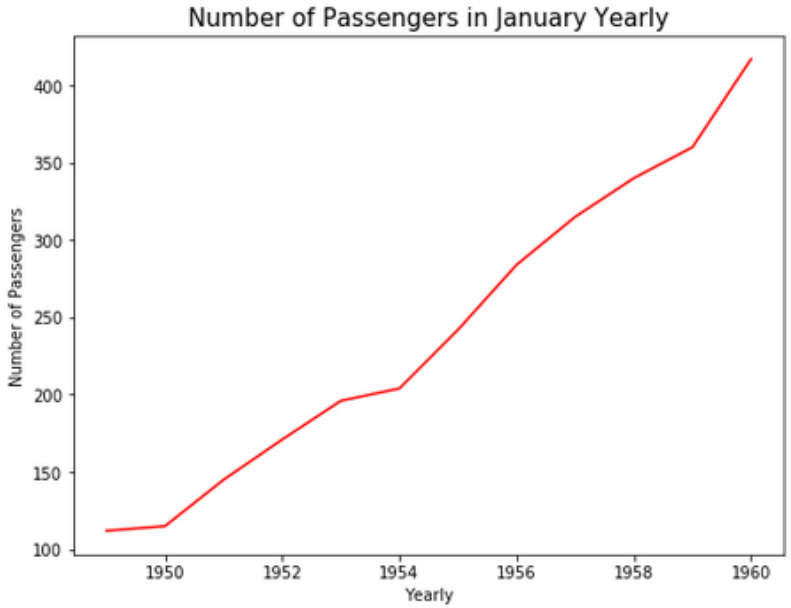
Create Line Plot using Matplotlib

Purwadhika
Startup and Coding School

Create Line Plot using Matplotlib

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python.

[10]:
```
plt.figure(figsize=(8,6))                                    # figure size
plt.plot(flights['year'][flights['month']=='January'],       # x axis = year name only in January
         flights['passengers'][flights['month']=='January'], # y axis = number of passengers only in January
         'red')                                              # line color
plt.title('Number of Passengers in January Yearly', size=15) # Title
plt.xlabel('Yearly')                                         # X label
plt.ylabel('Number of Passengers')                           # Y label
plt.show()
```



[3]:
```
# Import Matplotlib & Seaborn
import matplotlib.pyplot as plt
import seaborn as sns

# Import Flights Dataset from seaborn
flights = sns.load_dataset("flights")
flights.head(3)
```

[3]:

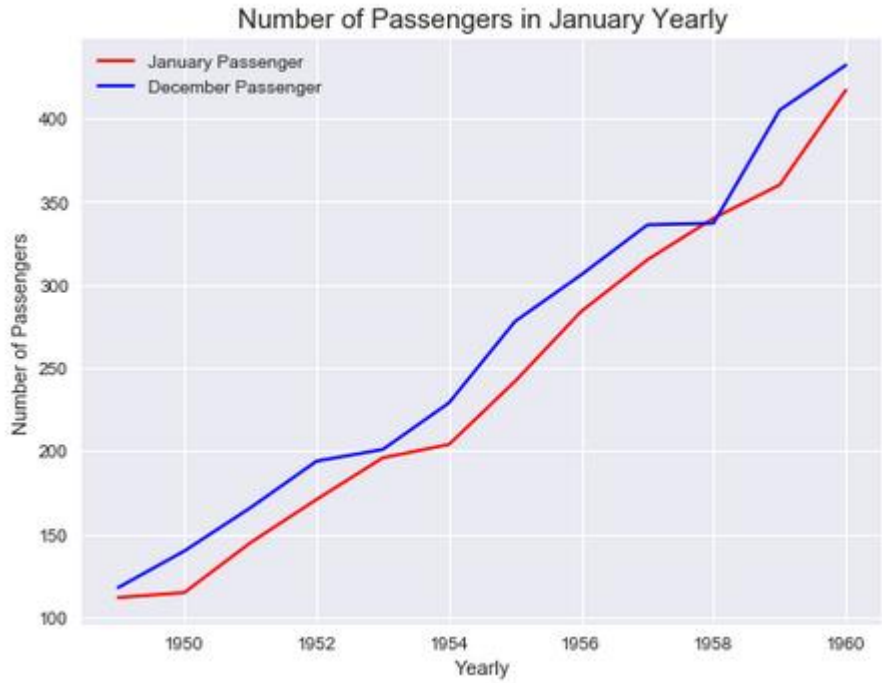
	year	month	passengers
0	1949	January	112
1	1949	February	118
2	1949	March	132

Purwadhika
Startup and Coding School

Create Line Plot using Matplotlib

[13]:
```
x1 = flights['year'][flights['month']=='January']
y1 = flights['passengers'][flights['month']=='January']
x2 = flights['year'][flights['month']=='December']
y2 = flights['passengers'][flights['month']=='December']

plt.style.use('seaborn')                                    # change style
plt.figure(figsize=(8,6))                                   # figure size
plt.plot(x1, y1, 'red',                                     # January Passenger Yearly
         x2, y2, 'blue')                                    # December Passenger Yearly
plt.title('Number of Passengers in January Yearly', size=15) # Title
plt.xlabel('Yearly')                                        # X label
plt.ylabel('Number of Passengers')                          # Y label
plt.legend(['January Passenger', 'December Passenger'])     # add legend
plt.grid(True)                                              # add grid
plt.savefig('JanDes_Passengers_Lineplot.png')               # saving plot
plt.show()
```



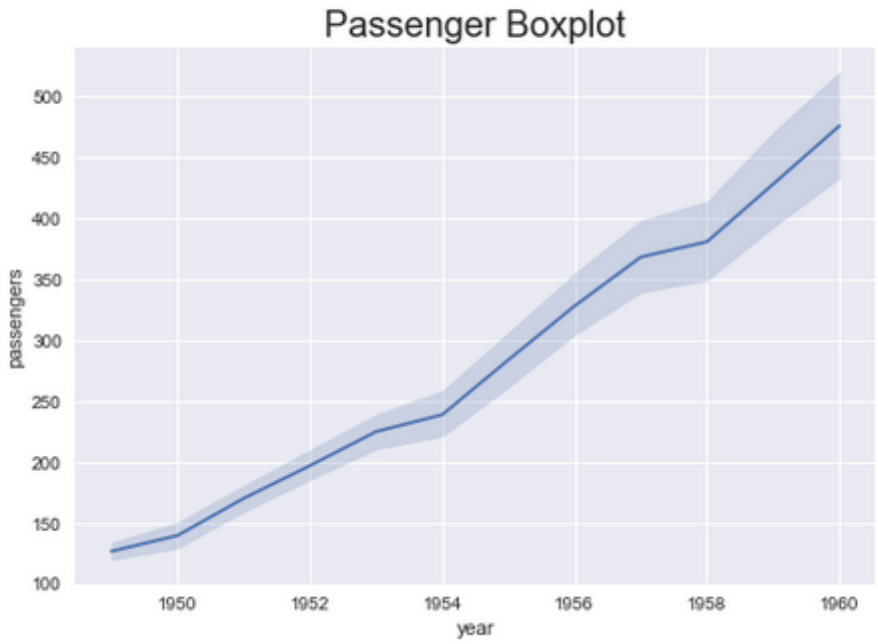
Purwadhika
Startup and Coding School

Create Line Plot using Seaborn

Purwadhika
Startup and Coding School

Create Line Plot using Seaborn

```
[21]: sns.lineplot(data=flights, x="year", y="passengers") # create line plot using seaborn
      plt.title('Passenger Boxplot', size=20)              # add title
      plt.show()
```



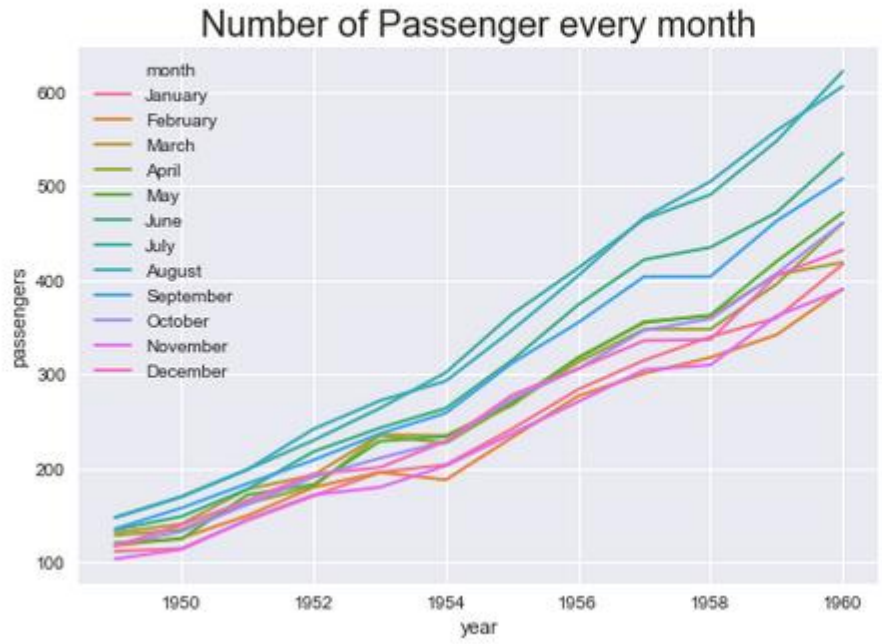
Seaborn is a Python data visualization library based on matplotlib.

It provides a high-level interface for drawing attractive and informative statistical graphics.

Purwadhika
Startup and Coding School

Create Line Plot using Seaborn

```
[22]: sns.lineplot(data=flights, x="year", y="passengers", hue="month") # create line plot using seaborn
      plt.title('Number of Passenger every month', size=20)             # add title
      plt.show()
```



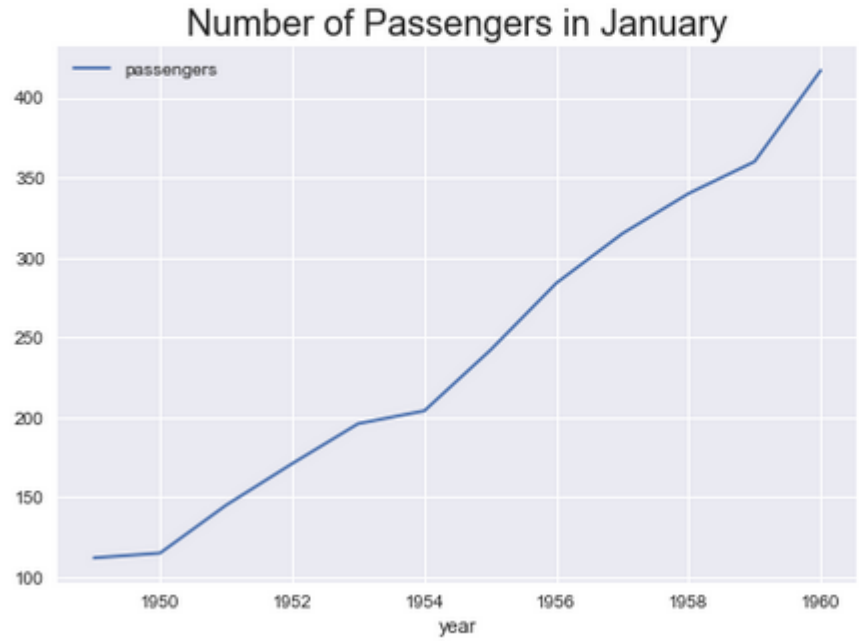
Purwadhika
Startup and Coding School

Create Line Plot using Pandas

Purwadhika
Startup and Coding School

Create Line Plot using Pandas

```
[29]: df_january = flights[flights['month']=='January']     # passenger in January
      df_january.plot.line(x='year', y='passengers')       # create line plot using pandas
      plt.title('Number of Passengers in January', size=20) # add title
      plt.show()
```



Pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.

Purwadhika
Startup and Coding School

Reference

- Wikipedia, "Line chart", https://en.wikipedia.org/wiki/Line_chart
- Veekesh Dhununjoy, "Advanced Visualization for Data Scientists with Matplotlib", https://medium.com/sfu-cspmp/advanced-visualization-for-data-scientists-with-matplotlib-15c28863c41c
- 365datascience, "Choosing the right chart: Selecting among 14 chart types", https://365datascience.com/chart-types-and-how-to-select-the-right-one/
- Badreesh Shetty, "Data Visualization using Matplotlib", https://towardsdatascience.com/data-visualization-using-matplotlib-16f1aae5ce70
- Daftar kode saham, https://www.sahamok.com/emiten/sektor-industri-barang-konsumsi/sub-sektor-makanan-minuman/
- Pandas.Series.diff, https://pandas.pydata.org/docs/reference/api/pandas.Series.diff.html
- Amanda Iglesias Moreno, "Simple and multiple linear regression with Python", https://towardsdatascience.com/simple-and-multiple-linear-regression-with-python-c9ab422ec29c
- Seaborn, “seaborn.lineplot”, https://seaborn.pydata.org/generated/seaborn.lineplot.html

Purwadhika
Startup and Coding School