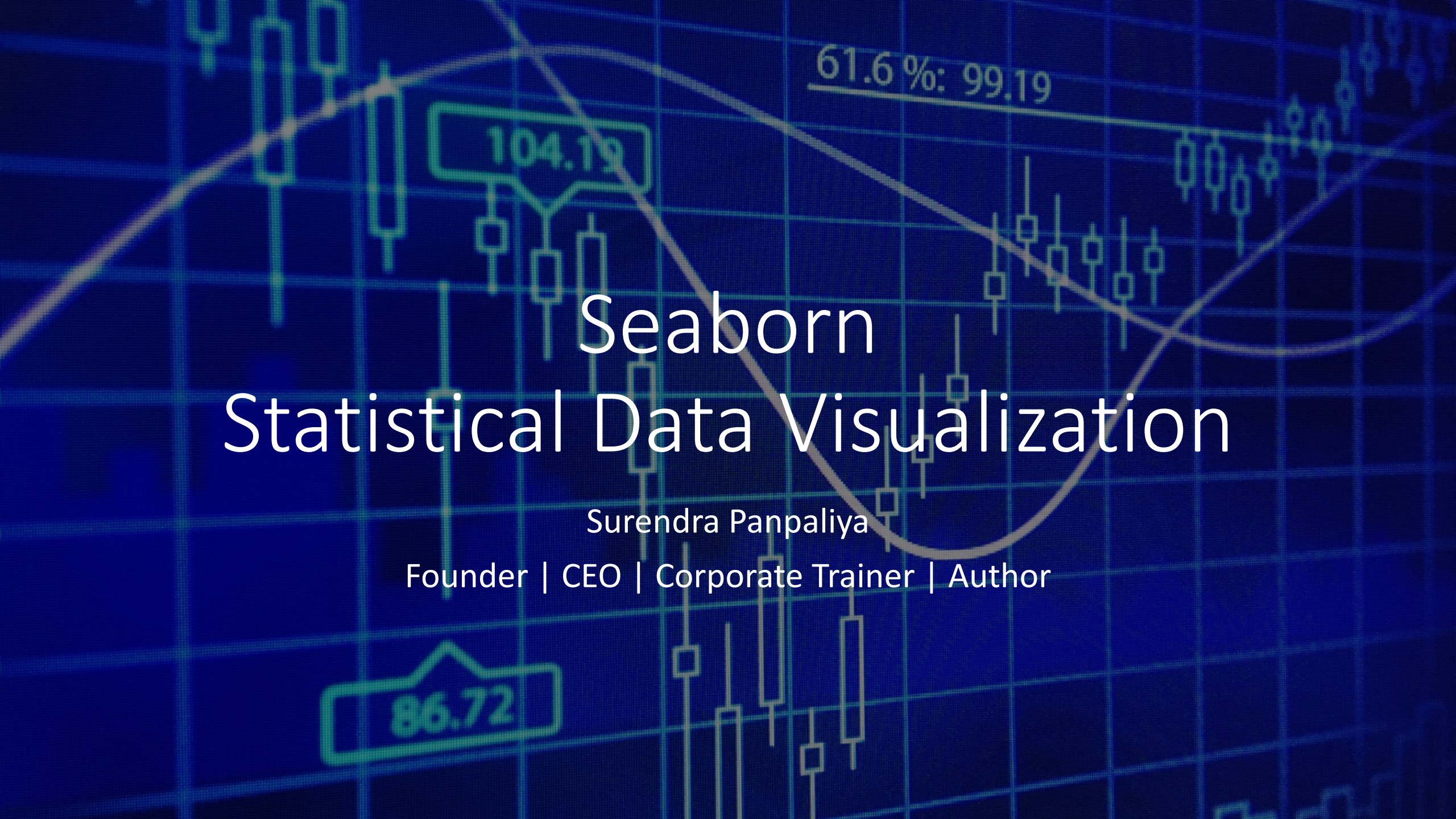

Seaborn
Statistical Data Visualization

Surendra Panpaliya

Founder | CEO | Corporate Trainer | Author

Agenda

- What is Seaborn?
- Seaborn Vs Matplotlib
- Important Features of Seaborn
- Importing Libraries

Agenda

- Seaborn Datasets
- DataFrames
- Visualization
- Seaborn Figure Styles
- Seaborn Hands On

What is Seaborn?

Seaborn is a Python data visualization library based on matplotlib.

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

An open source, BSD-licensed Python library.

Providing high level API for visualizing the data.

Seaborn Vs Matplotlib

Matplotlib tries to make easy things easy and hard things possible.

Seaborn tries to make a well-defined set of hard things easy too.

Seaborn Vs Matplotlib

Seaborn helps resolve the two major problems faced by Matplotlib

Default Matplotlib parameters

Working with data frames

Seaborn compliments and extends Matplotlib.

Data Visualization

- Data Visualization is the graphic representation of data.
- It converts a huge dataset into small graphs, thus aids in data analysis and predictions.
- It is an indispensable element of data science which makes complex data more understandable and accessible.
- Matplotlib and Seaborn act as the backbone of data visualization through Python.

Matplotlib

It is a Python library used for plotting graphs with the help of other libraries like Numpy and Pandas.

It is a powerful tool for visualizing data in Python. It is used for creating statical interferences and plotting 2D graphs of arrays.

It was first introduced by John D. Hunter in 2002. It uses Pyplot for providing MATLAB like interface free and open-source.

It is capable of dealing with various operating systems and their graphical backends.

Seaborn

It is also a Python library used for plotting graphs with the help of Matplotlib, Pandas, and Numpy.

It is built on the roof of Matplotlib and is considered as a superset of the Matplotlib library.

It helps in visualizing univariate and bivariate data. It uses beautiful themes for decorating Matplotlib graphics.

Seaborn

- It acts as an important tool in picturing Linear Regression Models.
- It serves in making graphs of statical Time-Series data.
- It eliminates the overlapping of graphs and also aids in their beautification.

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Functionality	It is utilized for making basic graphs. Datasets are visualised with the help of bargraphs, histograms, piecharts, scatter plots, lines and so on.	Seaborn contains a number of patterns and plots for data visualization. It uses fascinating themes. It helps in compiling whole data into a single plot. It also provides distribution of data.

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Syntax	It uses comparatively complex and lengthy syntax. Example: Syntax for bargraph- matplotlib.pyplot.bar(x_axis, y_axis).	It uses comparatively simple syntax which is easier to learn and understand. Example: Syntax for bargraph- seaborn.barplot(x_axis, y_axis).

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Dealing Multiple Figures	We can open and use multiple figures simultaneously. However they are closed distinctly. Syntax to close one figure at a time: matplotlib.pyplot.close(). Syntax to close all the figures: matplotlib.pyplot.close("all")	Seaborn sets time for the creation of each figure. However, it may lead to (OOM) out of memory issues

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Visualization	Matplotlib is well connected with Numpy and Pandas and acts as a graphics package for data visualization in python. Pyplot provides similar features and syntax as in MATLAB. Therefore, MATLAB users can easily study it.	Seaborn is more comfortable in handling Pandas data frames. It uses basic sets of methods to provide beautiful graphics in python.

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Pliability	Matplotlib is a highly customized and robust	Seaborn avoids overlapping of plots with the help of its default themes

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Data Frames and Arrays	Matplotlib works efficiently with data frames and arrays.It treats figures and aces as objects. It contains various stateful APIs for plotting. Therefore plot() like methods can work without parameters.	Seaborn is much more functional and organized than Matplotlib and treats the whole dataset as a single unit. Seaborn is not so stateful and therefore, parameters are required while calling methods like plot()

Matplotlib vs Seaborn

FEATURES	MATPLOTLIB	SEABORN
Use Cases	Matplotlib plots various graphs using Pandas and Numpy	Seaborn is the extended version of Matplotlib which uses Matplotlib along with Numpy and Pandas for plotting graphs

Important Features of Seaborn

Built in themes for styling matplotlib graphics

Visualizing univariate and bivariate data

Fitting in and visualizing linear regression models

Important Features of Seaborn

Plotting statistical time series data

Seaborn works well with NumPy and Pandas data structures

It comes with built in themes for styling Matplotlib graphics

Importing Libraries

```
import pandas as pd
```

```
from matplotlib import pyplot as plt
```

```
import seaborn as sb
```

```
df = sb.load_dataset('tips')
```

```
print(df.head())
```

df.head()

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Seaborn Datasets

To view all the available data sets in the Seaborn library

print(sb.get_dataset_names())

['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes', 'diamonds', 'dots', 'exercise', 'flights', 'fmri', 'gammas', 'geyser', 'iris', 'mpg', 'penguins', 'planets', 'tips', 'titanic']

DataFrames

It store data in the form of rectangular grids by which the data can be over viewed easily.

Each row of the rectangular grid contains values of an instance

Each column of the grid is a vector which holds data for a specific variable.

DataFrames

Rows of a DataFrame do not need to contain values of same data type.

Can be numeric, character or logical.

DataFrames for Python come with the Pandas library

Defined as two-dimensional labeled data structures with potentially different types of columns

Visualization

Visualizing data is one step and further making the visualized data more pleasing is another step.

Visualization is an art of representing data in effective and easiest possible way.

Unlike Matplotlib, Seaborn comes packed with customized themes

A high-level interface for customizing and controlling the look of Matplotlib figures

Matplotlib

```
import numpy as np
from matplotlib import pyplot as plt


def sinplot(flip = 1):
    x = np.linspace(0, 14, 100)
    print(x)
    for i in range(1, 5):
        plt.plot(x, np.sin(x + i * .5) * (7 - i) * flip)


sinplot()
plt.show()
```

Seaborn

- To change the same plot to Seaborn defaults,
- use the **set()** function −

```
import seaborn as sb
sb.set()
sinplot()
plt.show()
```

Seaborn Figure Styles

The interface for manipulating the styles is **set_style()**.

Using this function you can set the theme of the plot.

As per the latest updated version, below are the five themes available.

Darkgrid

Whitegrid

Seaborn Figure Styles

- Darkgrid
- Whitegrid
- Dark
- White
- Ticks

Thank You

Surendra Panpaliya
Founder | CEO | International Corporate Trainer | Author
Email: Surendra@gktcs.com
Mobile: +91-9975072320 / +91-7620379390

Web: https://www.gktcs.com

Profile link: https://www.gktcs.com/surendra

Linkedin: https://www.linkedin.com/in/surendrarp