

Data Visualization in Python

Why Data Visualization?

- Data visualization is a quick, easy way to convey concepts in a universal manner.

STATE	Rate 2004	Verbal	Math	Rate 2003	Verbal	Math
New York	87%	497	510	82%	496	510
Connecticut	85%	515	515	84%	512	514
Massachusetts	85%	518	523	82%	516	522
New Jersey	83%	501	514	85%	501	515
New Hampshire	80%	522	521	75%	522	521
D. C.	77%	489	476	77%	484	474
Maine	76%	505	501	70%	503	501
Pennsylvania	74%	501	502	73%	500	502
Delaware	73%	500	499	73%	501	501
Georgia	73%	491	493	66%	493	491
Rhode Island	72%	503	502	74%	502	504
Virginia	71%	515	509	71%	514	510
North Carolina	70%	499	507	68%	495	506
Maryland	68%	511	515	68%	509	515
Florida	67%	499	499	61%	498	498
Vermont	66%	516	512	70%	515	512
Indiana	64%	501	506	63%	500	504
South Carolina	62%	491	495	59%	493	496
Hawaii	60%	487	514	54%	486	516
Oregon	56%	527	528	57%	526	527
Alaska	53%	518	514	55%	518	518
Texas	52%	493	499	57%	493	500
Washington	52%	528	531	56%	530	532
California	49%	501	510	54%	499	519
Nevada	40%	507	514	36%	510	517
Arizona	32%	523	524	38%	524	525
Montana	29%	537	539	26%	538	543



Why Data Visualization?

- Data visualization is the graphical representation of information and data.
- visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

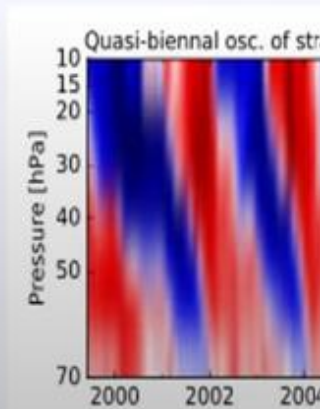
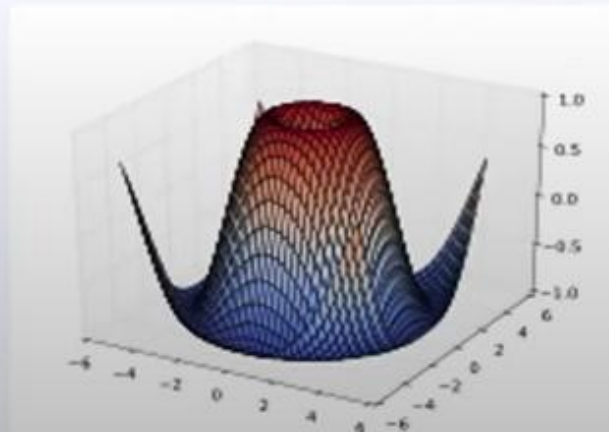
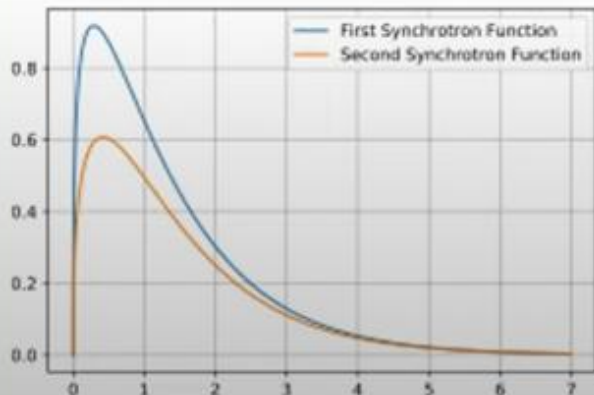
Types of Data Visualization in Python

Plotting Libraries:

- **Matplotlib:** low level, provides lots of freedom
- **Pandas Visualization:** easy to use interface, built on Matplotlib
- **Seaborn:** high-level interface, great default styles
- **Plotly:** can create interactive plots

What is Matplotlib?

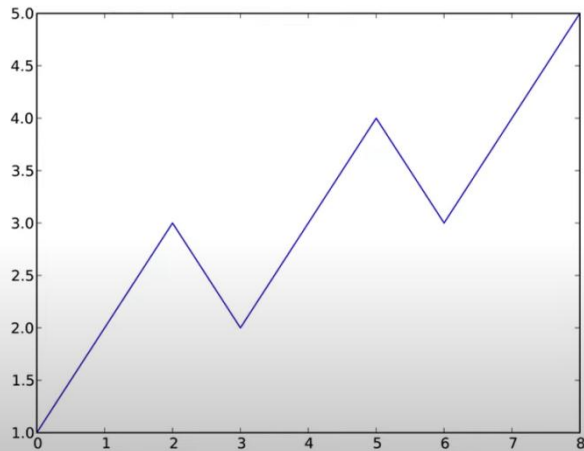
- Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.
- Matplotlib is a python library used for Data Visualization.
- Matplotlib is 2D and 3D plotting python library.
- It was introduced by John Hunter in the year 2002.



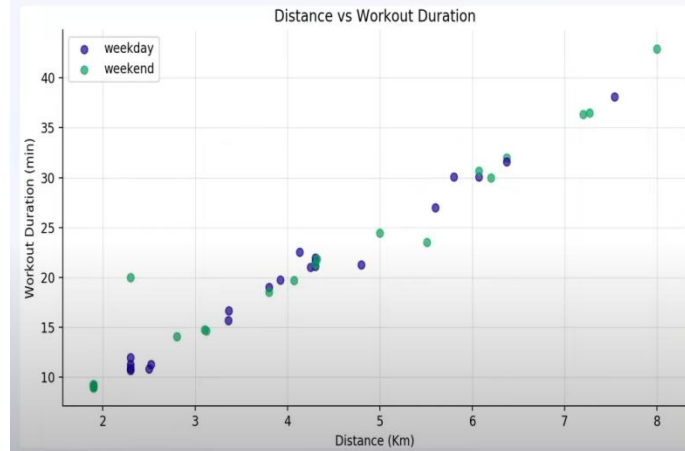
Matplotlib Graphs

- **Linear Plot**
- **Scatter Plot**
- **Bar Plot**
- **Stem Plot**
- **Step Plot**
- **Hist Plot**
- **Box Plot**
- **Pie Plot**
- **Fill_Between Plot**

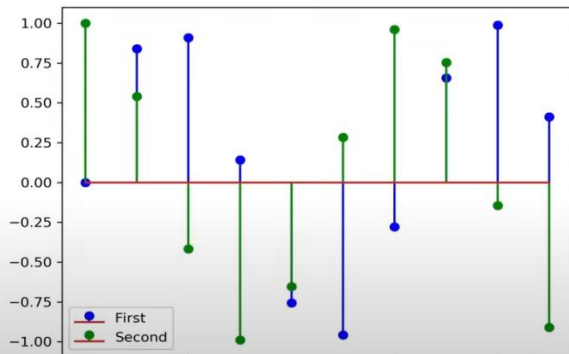
Linear Plot



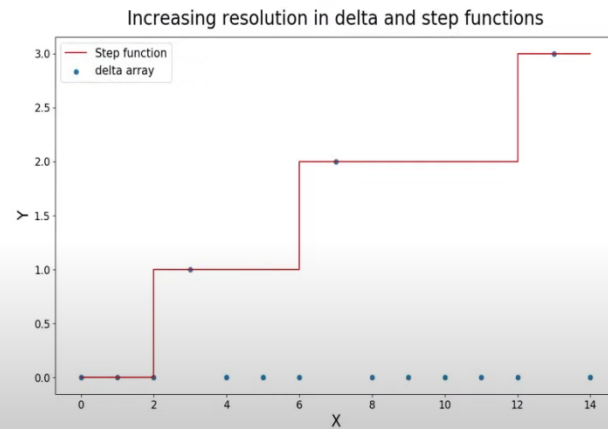
Scatter Plot



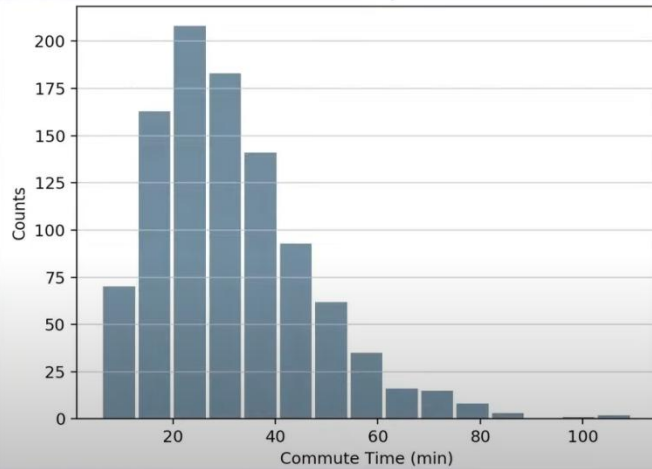
Stem Plot



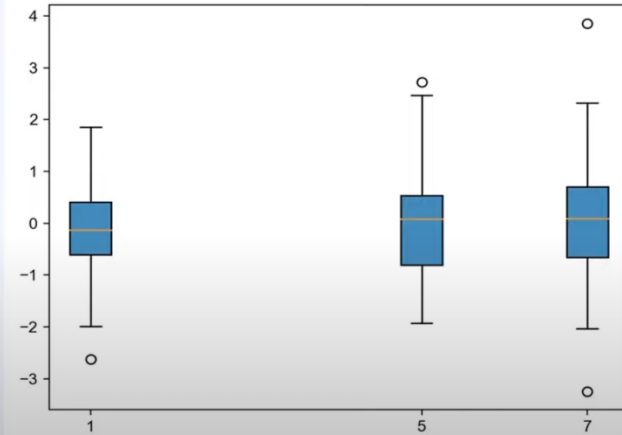
Step Plot



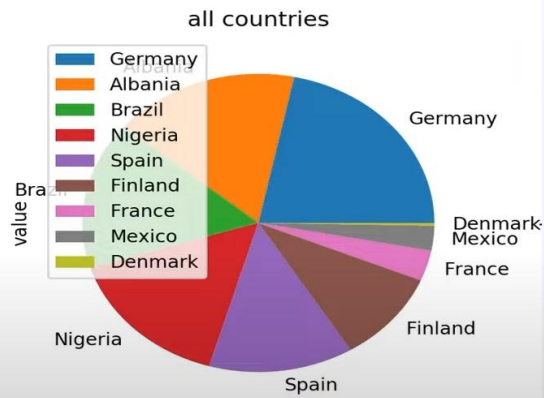
Hist Plot



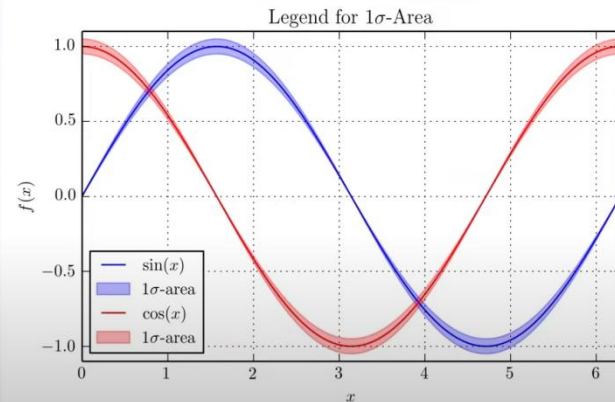
Box Plot



Pie Plot



Fill_Between Plot



Installation of matplotlib

pip install matplotlib

Importing Matplotlib

```
import matplotlib.pyplot as plt  
or  
from matplotlib import pyplot as plt
```

Matplotlib Bar Plot

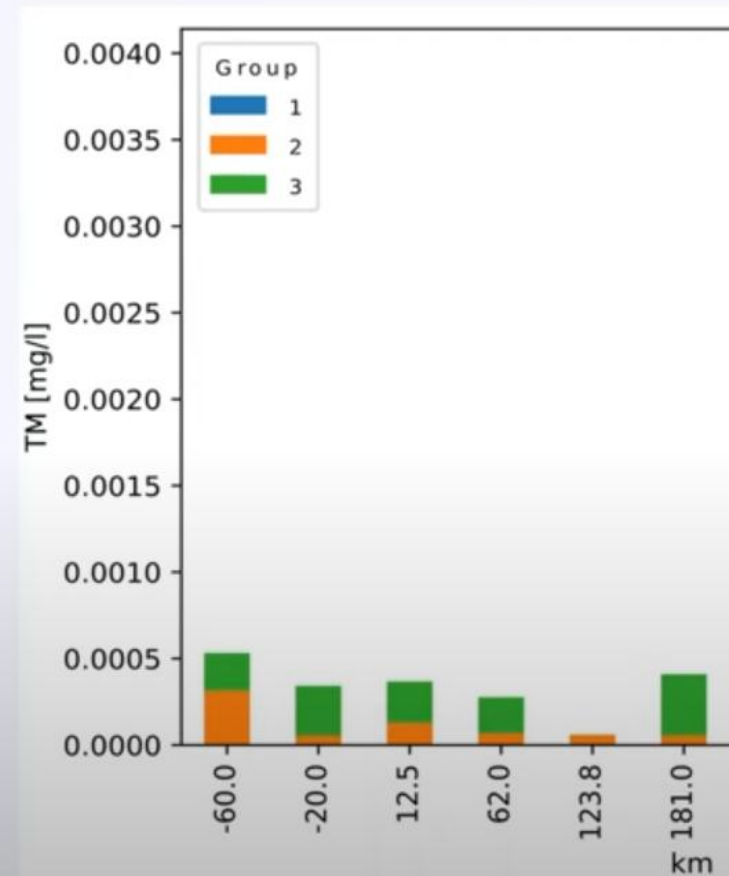
```
import matplotlib.pyplot as plt
```

```
x = [ ]
```

```
y = [ ]
```

```
plt.bar(x, y)
```

```
plt.show()
```



iPhone Sales Data Analysis Project


- 1) What are the top 10 highest-rated iPhones on Flipkart in India?
- 2) How many ratings do the highest-rated iPhones on Flipkart have?
- 3) Which iPhone has the highest number of reviews on Flipkart?
- 4) What is the relationship between the sale price of iPhones and the number of ratings on Flipkart?
- 5) What is the relationship between the discount percentage and the number of ratings of iPhones on Flipkart?

Python SEABORN



Seaborn Graphs

- Scatter Plot
- Customizing with Matplotlib
- Box Plot
- Violin Plot
- Swarm Plot
- Overlaying plots
- Heatmap
- Histogram
- Bar Plot
- Factor Plot
- Density Plot
- Joint Distribution Plot

Differences between Seaborn and Matplotlib

Feature	Matplotlib	Seaborn	
Level	Low-level, gives full control over plots.	High-level, built on top of Matplotlib.	
Ease of Use	Requires more code for styling/complex plots.	Fewer lines of code for clean, statistical plots.	
Default Style	Basic (plain plots).	Prettier by default (color palettes, themes).	
Focus	General-purpose plotting (line, bar, scatter, etc.).	Statistical visualization (distribution, regression, categorical data).	
Customization	Extremely customizable, but verbose.	Easier for statistical plots, but internally still uses Matplotlib.	

When to Use Which?

-  Use Matplotlib when:
 - You need **full control** over figure design (research publications, dashboards).
 - You want to **customize every aspect** (fonts, ticks, axes, annotations).
 - You're creating **non-standard plots** (e.g., radar charts, custom animations).
-  Use Seaborn when:
 - You want **beautiful plots quickly** with less code.
 - You're doing **Exploratory Data Analysis (EDA)** or reports.
 - You're working with **statistical data** (distributions, categorical comparisons, correlations).

Dependency of Seaborn



Installation Seaborn

`pip install seaborn`

`Conda install seaborn`

`pip install matplotlib`

Importing Matplotlib

```
import matplotlib.pyplot as plt
or
from matplotlib import pyplot as plt
and
Import seaborn as sns
```

Seaborn line plot

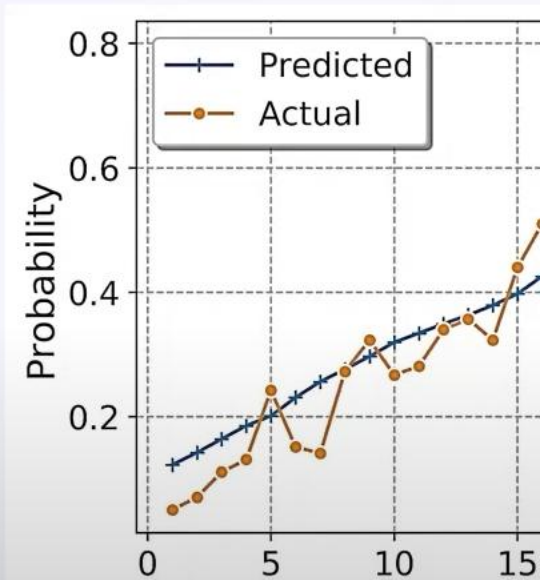
```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

```
data = sns.load_dataset(" ")
```

```
sns.lineplot(x=" ", y=" ", data = data)
```

```
plt.show()
```



Seaborn Heatmap Plot

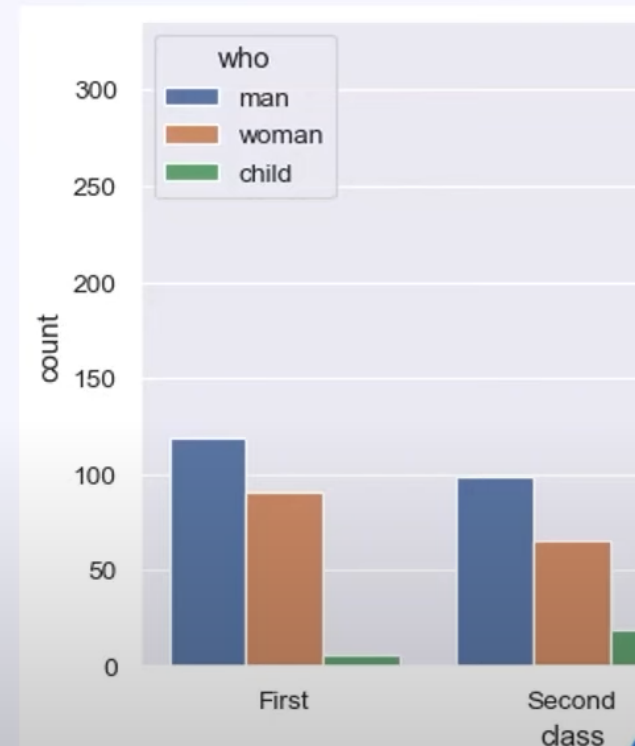
```
import seaborn as sns
import matplotlib.pyplot as plt

df = sns.load_dataset(' ')
sns.heatmap(data = data)
plt.show()
```



Count Plot in Seaborn

```
import seaborn as sns
import matplotlib.pyplot as plt
df = sns.load_dataset("")
sns.countplot(x="", data = df)
plt.show()
```



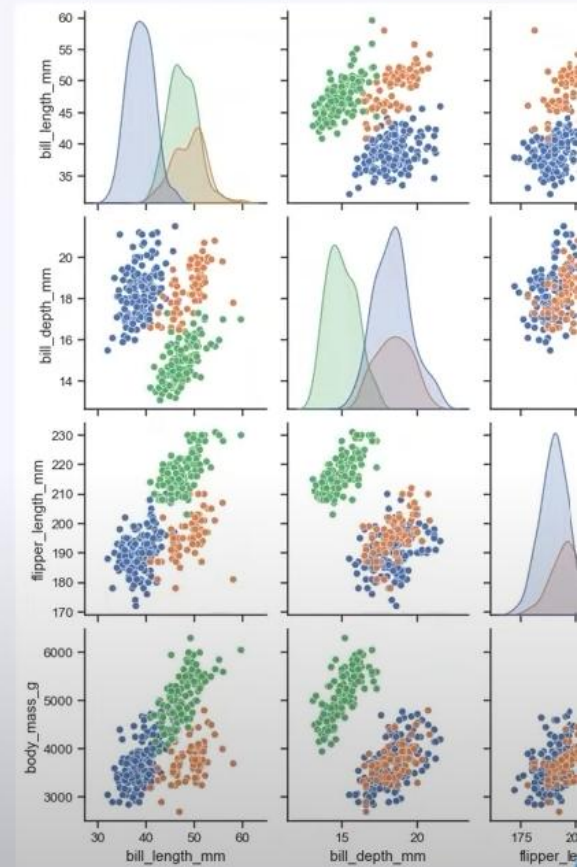
Count Plot in Seaborn

What's the difference between countplot and barplot ?

- Countplot plots the count of the number of records by category.
- Barplot plots a value or metric for each category (by default, barplot plots the mean of a variable, by category).

Pair Plot in Seaborn

```
import seaborn  
  
import matplotlib.pyplot as plt  
  
df = seaborn.load_dataset('penguins')  
seaborn.pairplot(df)  
plt.show()
```



Cat Plot in Seaborn

```
import seaborn as sns  
exercise = sns.load_dataset("")  
g = sns.catplot( x="", y="", data=exercise, kind = "" )
```

Cat Plot in Seaborn

Categorical scatterplots with catplot

- `stripplot()` – with `kind="strip"`
- `swarmplot()` – with `kind="swarm"`

Categorical estimate plots with catplot

- `pointplot()` – with `kind="point"`
- `barplot()` – with `kind="bar"`
- `countplot()` – with `kind="count"`

Categorical distribution plots with catplot

- `boxplot()` – with `kind="box"`
- `violinplot()` – with `kind="violin"`
- `boxenplot()` – with `kind="boxen"`

Styling Plots in Seaborn

- **Seaborn Figure Styles**
- **Removing Axes Spines**
- **Scale and Context**

Multiple Plots (Facet - Grid) in Seaborn

```
import seaborn
import matplotlib.pyplot as plt
df = seaborn.load_dataset("")
graph = seaborn.FacetGrid(df, col = "", hue = "")
graph.map(plt.scatter).add_legend()
plt.show()
```

