

ARJUN COLLEGE OF TECHNOLOGY  
**ASSIGNMENT – 3**  
NAAN MUDHALVAN

NAME : PERUVAILU KAMAL

REG NO: 723920243039



Post Attendee - Zoom | All Mail - peruvailukamal@g | My IBM | Home | Welcome to Colaboratory -

colab.research.google.com/?utm\_source=scs-index#scrollTo=KYcCregNQrIM

Welcome to Colaboratory

File Edit View Insert Runtime Tools Help Cannot save changes

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

```
[1] import pandas as pd
import numpy as np

[6] df=pd.read_csv('/content/House Price India.csv')
df.head()
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condition of the house	...	Built Year	Renovation Year	Pos
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4	5	...	1921	0	122
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0	5	...	1909	0	122
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0	3	...	1939	0	122
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0	3	...	2001	0	122
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0	4	...	1929	0	122

5 rows x 23 columns

15s completed at 11:13

25°C Mostly cloudy

Post Attendee - Zoom | All Mail - peruvailukamal@g | My IBM | Home | Welcome to Colaboratory -

colab.research.google.com/?utm\_source=scs-index#scrollTo=KYcCregNQrIM

Welcome to Colaboratory

File Edit View Insert Runtime Tools Help Cannot save changes

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

```
df.info
```

	number of views	condition of the house	...	Built Year	...
0	4	5	...	1921	
1	0	5	...	1909	
2	0	3	...	1939	
3	0	3	...	2001	
4	0	4	...	1929	
...	...	...	...	...	
14615	0	4	...	1957	
14616	0	4	...	1968	
14617	0	3	...	1962	
14618	0	4	...	1955	
14619	0	3	...	1969	

	Renovation Year	Postal Code	Latitude	Longitude	living_area_renov	...
0	0	122003	52.8645	-114.557	2880	
1	0	122004	52.8878	-114.470	2470	
2	0	122004	52.8852	-114.468	2940	
3	0	122005	52.9532	-114.321	3350	
4	0	122006	52.9047	-114.485	2060	

15s completed at 11:13

25°C Mostly cloudy

Post Attendee - Zoom | All Mail - peruvailukamal@g | My IBM | Home | Welcome to Colaboratory -

colab.research.google.com/?utm\_source=scs-index#scrollTo=KYcCregNQrIM

Welcome to Colaboratory

File Edit View Insert Runtime Tools Help Cannot save changes

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

```
df.info
```

	Renovation Year	Postal Code	Latitude	Longitude	living_area_renov	...
0	0	122003	52.8645	-114.557	2880	
1	0	122004	52.8878	-114.470	2470	
2	0	122004	52.8852	-114.468	2940	
3	0	122005	52.9532	-114.321	3350	
4	0	122006	52.9047	-114.485	2060	

# Welcome to Colaboratory

File Edit View Insert Runtime Tools Help [Cannot save changes](#)

Share Settings User

RAM Disk

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

```
[7] 14619 146000
[14620 rows x 23 columns]>

[8] import matplotlib.pyplot as plt
import pandas as pd

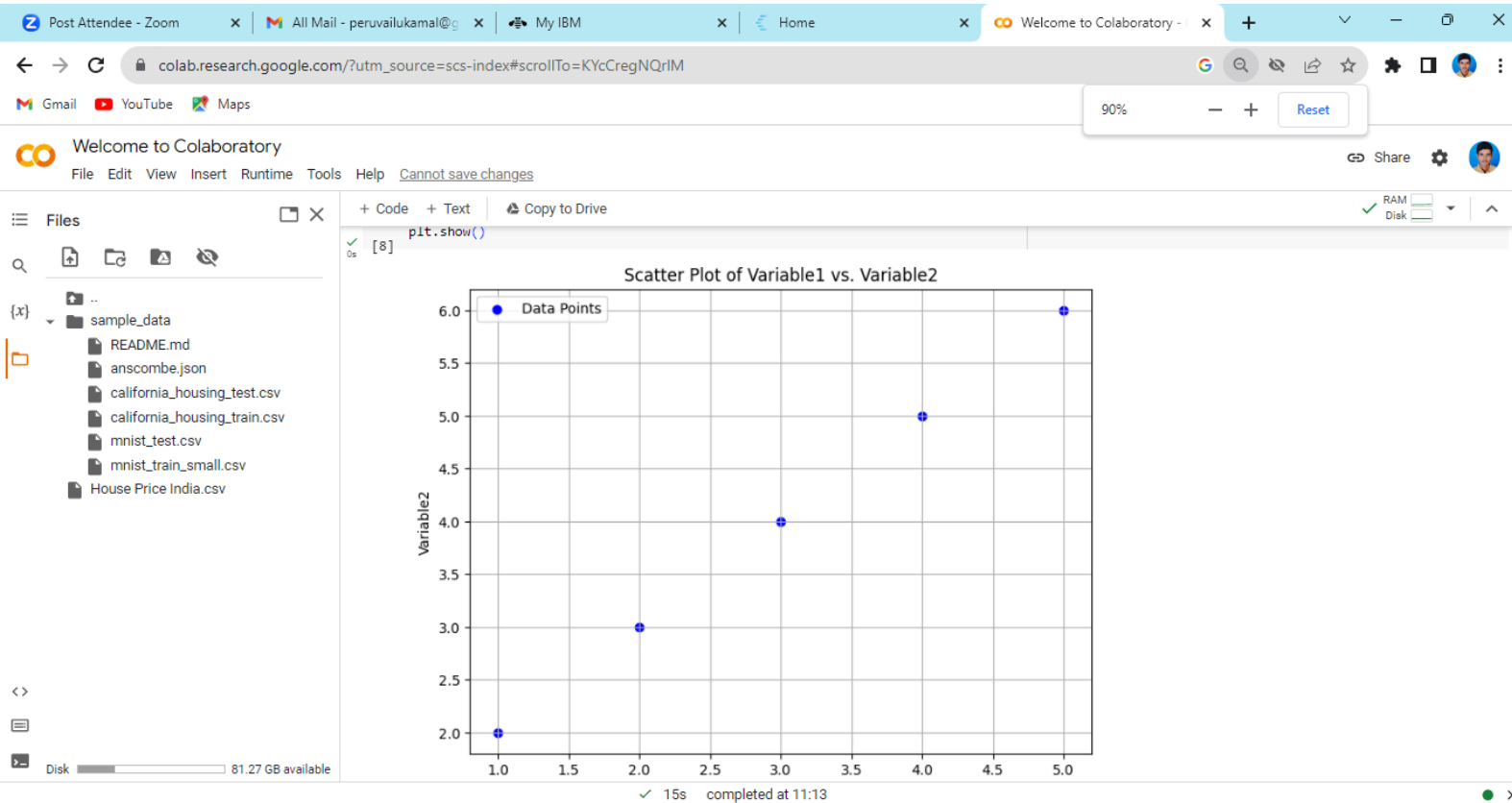
# Sample data (replace this with your own dataset)
data = {
    'Variable1': [1, 2, 3, 4, 5],
    'Variable2': [2, 3, 4, 5, 6]
}

# Create a DataFrame from the sample data
df = pd.DataFrame(data)

# Extract the two variables for the scatter plot
x = df['Variable1']
y = df['Variable2']

# Create a scatter plot
plt.figure(figsize=(8, 6)) # Set the figure size (optional)
plt.scatter(x, y, c='blue', marker='o', label='Data Points') # Scatter plot
plt.xlabel('Variable1') # X-axis label
plt.ylabel('Variable2') # Y-axis label
```

✓ 15s completed at 11:13



colab.research.google.com/?utm\_source=scs-index#scrollTo=KYcCregNQrIM

Welcome to Colaboratory

File Edit View Insert Runtime Tools Help Cannot save changes

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

Code

```
2s import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# Generate a sample dataset (you can replace this with your own data)
data = np.random.normal(0, 1, 1000) # Generating 1000 random data points with a mean of 0 and standard deviation of 1

# Create a pandas DataFrame
df = pd.DataFrame(data, columns=["Value"])

# Summary statistics
mean = df["Value"].mean()
median = df["Value"].median()
std_dev = df["Value"].std()

print("Mean:", mean)
print("Median:", median)
print("Standard Deviation:", std_dev)

# Create a histogram to visualize the distribution
plt.hist(df["Value"], bins=20, color='blue', edgecolor='black')
plt.xlabel("Value")
plt.ylabel("Frequency")
plt.title("Histogram of the Data")
plt.show()

# Create a box plot to visualize the summary statistics
plt.boxplot(df["Value"], showfliers=False)
```

15s completed at 11:13

colab.research.google.com/?utm\_source=scs-index#scrollTo=KYcCregNQrIM

Welcome to Colaboratory

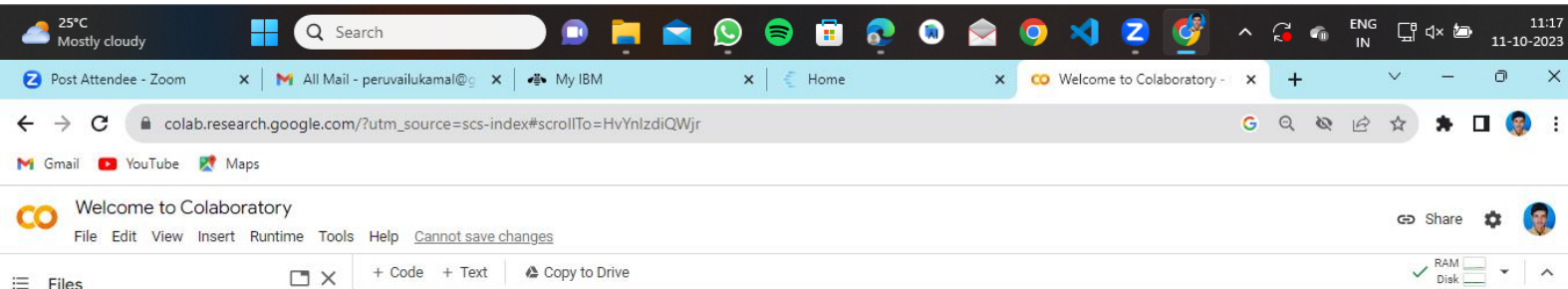
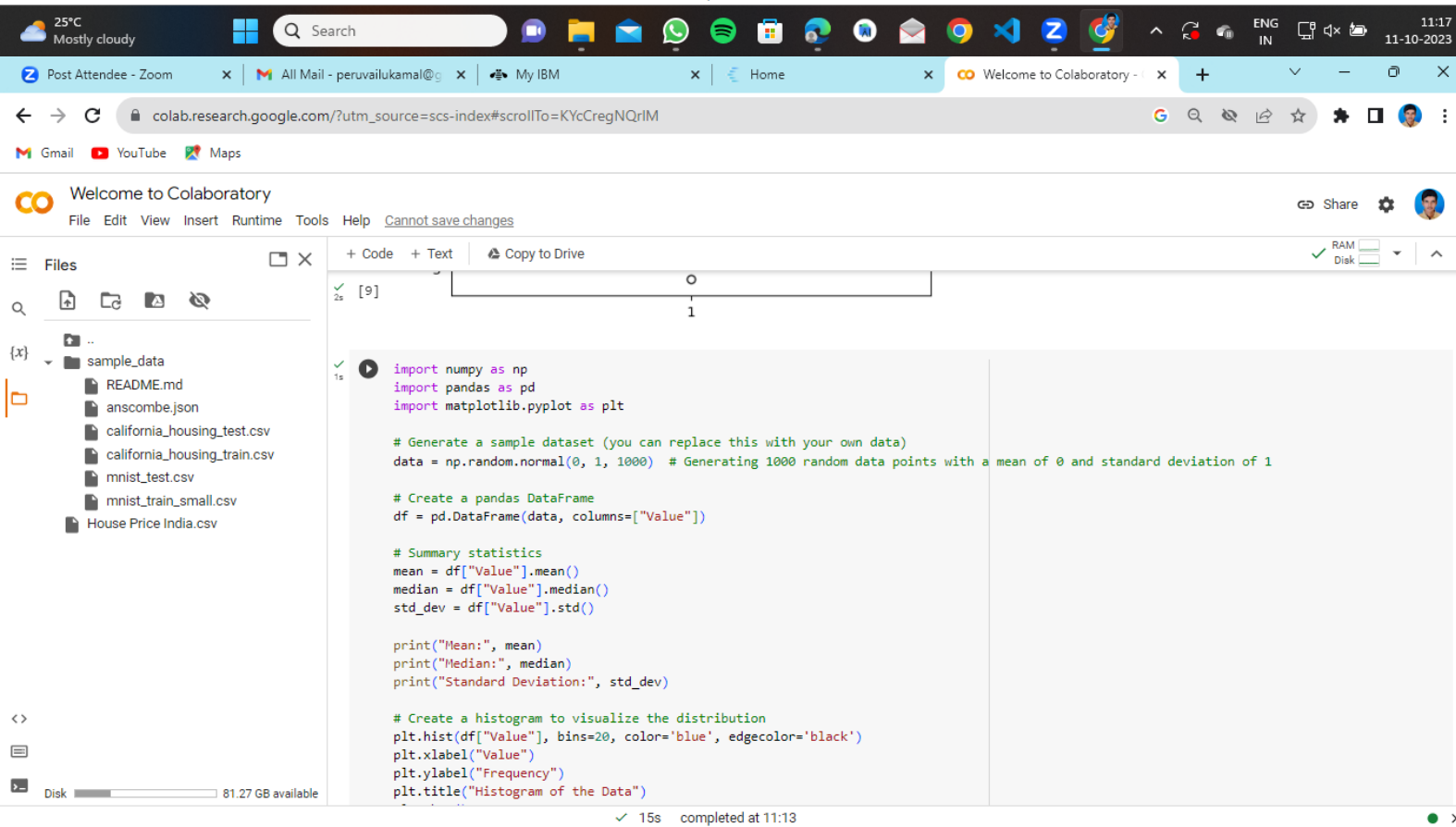
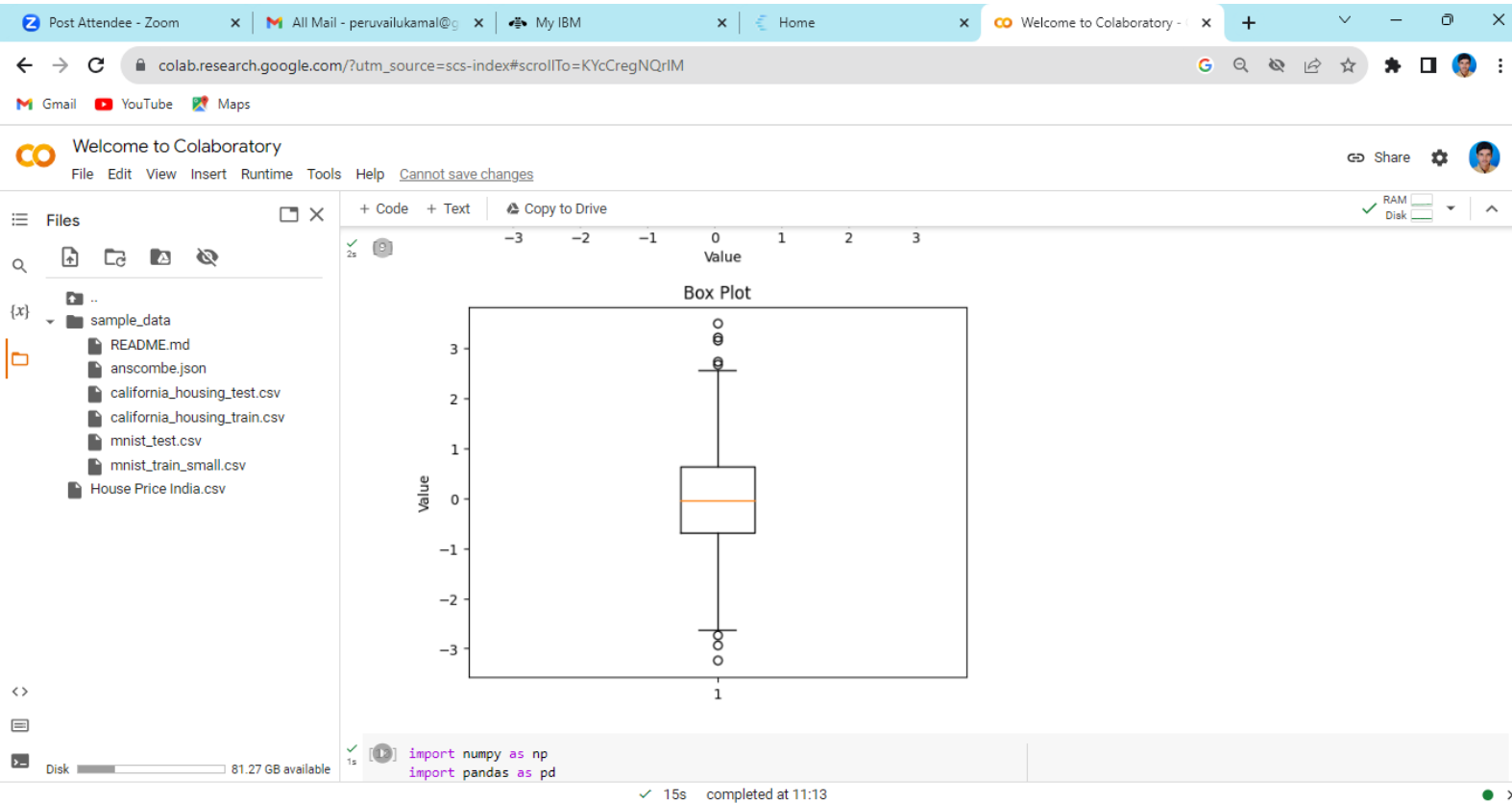
File Edit View Insert Runtime Tools Help Cannot save changes

Files

- sample\_data
  - README.md
  - anscombe.json
  - california\_housing\_test.csv
  - california\_housing\_train.csv
  - mnist\_test.csv
  - mnist\_train\_small.csv
  - House Price India.csv

Code

```
plt.ylabel("Value")
```



Files

sample\_data

README.md

anscombe.json

california\_housing\_test.csv

california\_housing\_train.csv

mnist\_test.csv

mnist\_train\_small.csv

House Price India.csv

Disk 81.27 GB available



