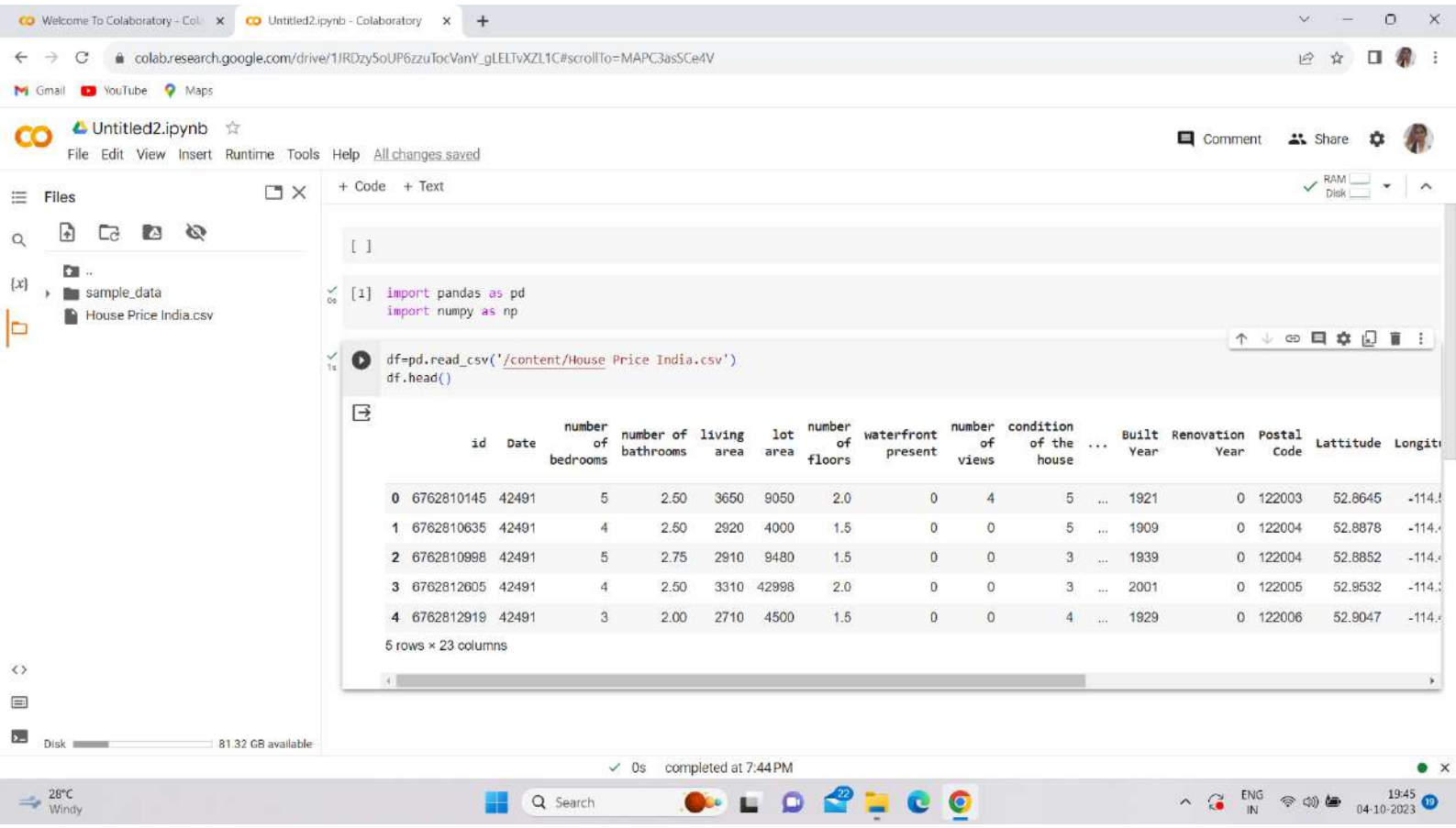
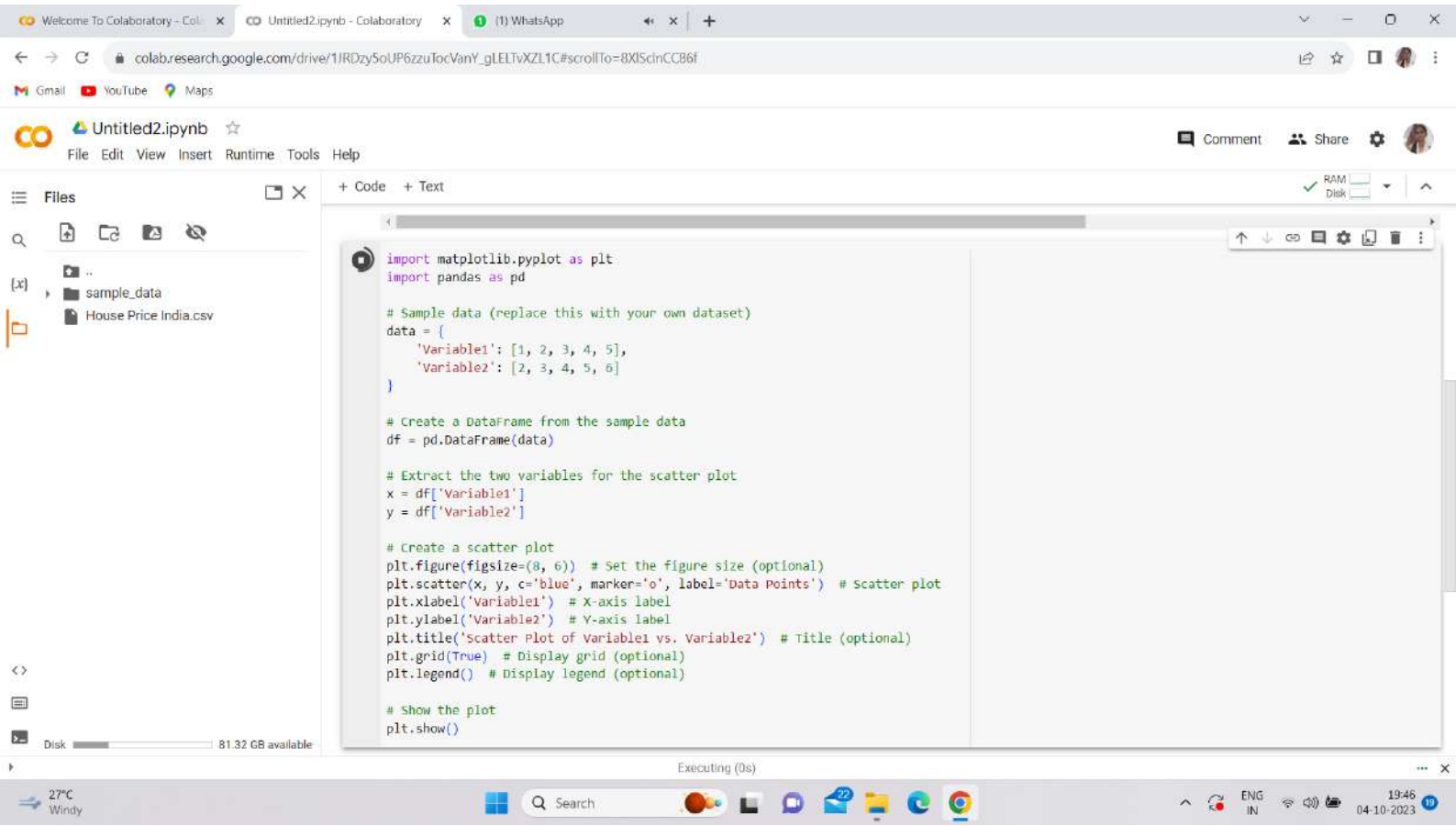

Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=MAPC3asSCe4V
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
..
sample_data
House Price India.csv
[ ]
[1] import pandas as pd
import numpy as np
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 Postal Code Lattitude Longitu
0 6762810145 42491 5 2.50 3650 9050 2.0 0 4 5 ... 1921 0 122003 52.8645 -114.5
1 6762810635 42491 4 2.50 2920 4000 1.5 0 0 5 ... 1909 0 122004 52.8878 -114.4
2 6762810998 42491 5 2.75 2910 9480 1.5 0 0 3 ... 1939 0 122004 52.8852 -114.4
3 6762812605 42491 4 2.50 3310 42998 2.0 0 0 3 ... 2001 0 122005 52.9532 -114.3
4 6762812919 42491 3 2.00 2710 4500 1.5 0 0 4 ... 1929 0 122006 52.9047 -114.4
5 rows × 23 columns
Disk 81.32 GB available
0s completed at 7:44 PM
28°C Windy
Search
ENG IN
19:45 04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=8XlSclnCC86f
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Files
+ Code + Text
RAM
Disk
..
sample_data
House Price India.csv
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
plt.title('Scatter Plot of Variable1 vs. Variable2') # Title (optional)
plt.grid(True) # Display grid (optional)
plt.legend() # Display legend (optional)
# Show the plot
plt.show()
Disk
81.32 GB available
Executing (0s)
27°C
Windy
Search
ENG
IN
19:46
04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=8XlSclnCC86f

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

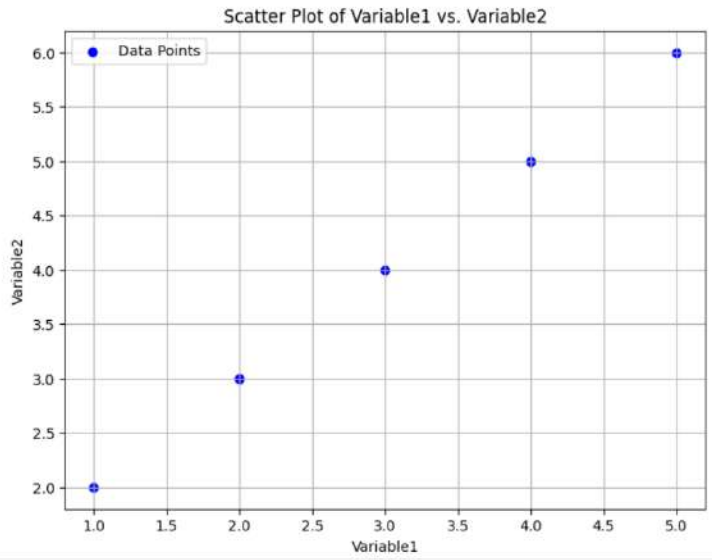
..
sample_data
House Price India.csv

Disk 81.32 GB available

+ Code + Text

RAM Disk

```
plt.show()
```



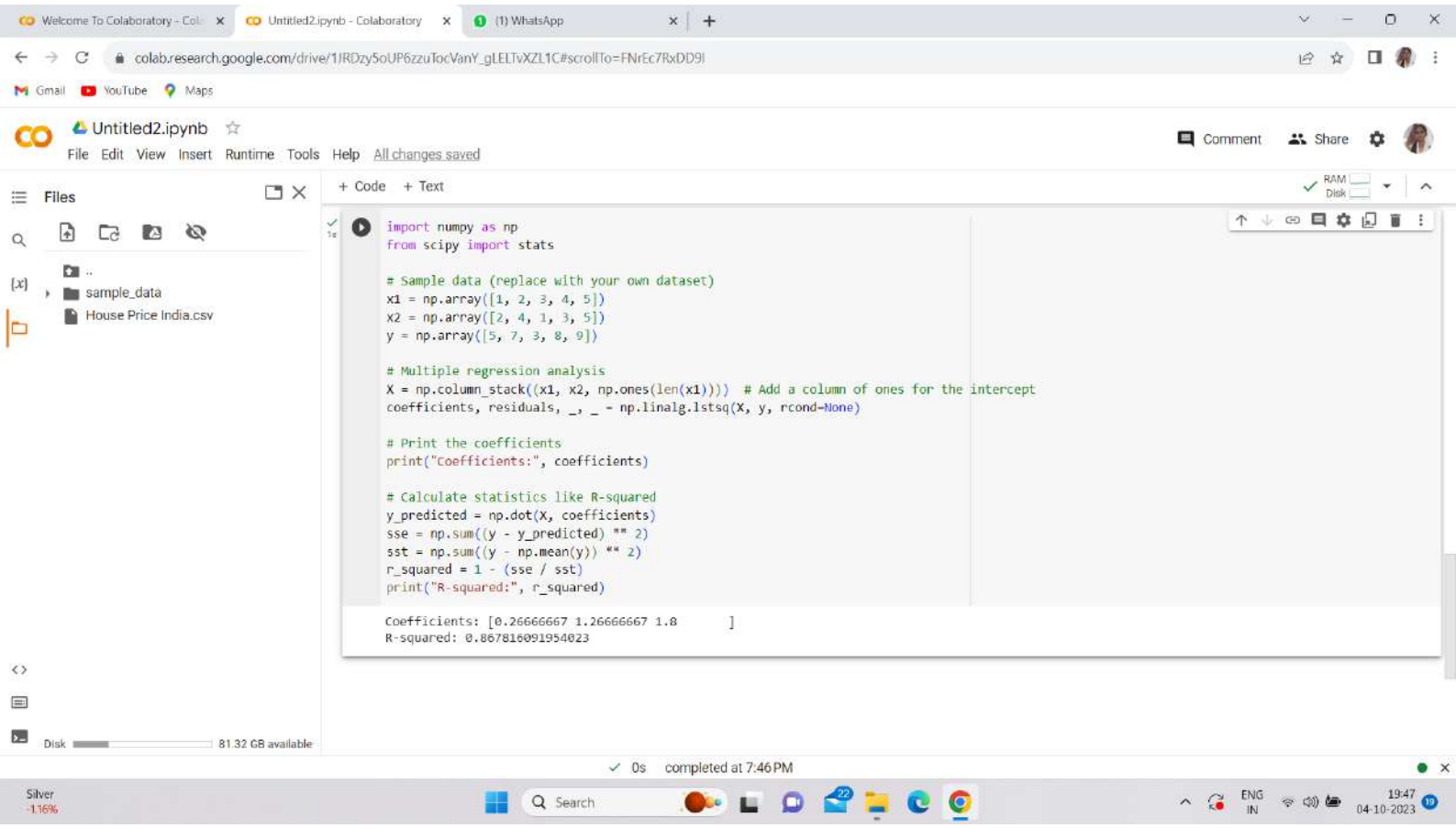
1s completed at 7:46 PM

27°C Windy

Search

ENG IN

19:46 04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=FNrEc7RxDD9I
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
House Price India.csv
+ Code + Text
RAM
Disk
import numpy as np
from scipy import stats
# Sample data (replace with your own dataset)
x1 = np.array([1, 2, 3, 4, 5])
x2 = np.array([2, 4, 1, 3, 5])
y = np.array([5, 7, 3, 8, 9])
# Multiple regression analysis
X = np.column_stack((x1, x2, np.ones(len(x1)))) # Add a column of ones for the intercept
coefficients, residuals, _, _ = np.linalg.lstsq(X, y, rcond=None)
# Print the coefficients
print("Coefficients:", coefficients)
# Calculate statistics like R-squared
y_predicted = np.dot(X, coefficients)
sse = np.sum((y - y_predicted) ** 2)
sst = np.sum((y - np.mean(y)) ** 2)
r_squared = 1 - (sse / sst)
print("R-squared:", r_squared)
Coefficients: [0.26666667 1.26666667 1.8]
R-squared: 0.867816091954023
Disk 81.32 GB available
0s completed at 7:46 PM
Silver -1.16%
Search
ENG IN
19:47 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=MUxMftmXDR85

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

Files

..
sample_data
House Price India.csv

Disk 81.32 GB available

+ Code + Text

RAM Disk

```
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"])
plt.ylabel("Value")
plt.title("Box Plot")
plt.show()
```

0s completed at 7:46 PM

Silver -1.16%

Search

ENG IN 19:47 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=MUxMftmXDR85

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

Files

..
sample_data
House Price India.csv

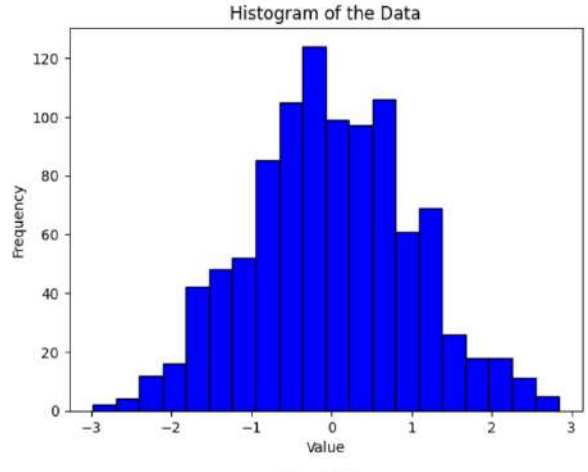
Disk 81.32 GB available

+ Code + Text

RAM Disk

```
plt.show()
```

```
Mean: -0.010086337268738753
Median: -0.041264079956700644
Standard Deviation: 1.0219189100143296
```



Box Plot

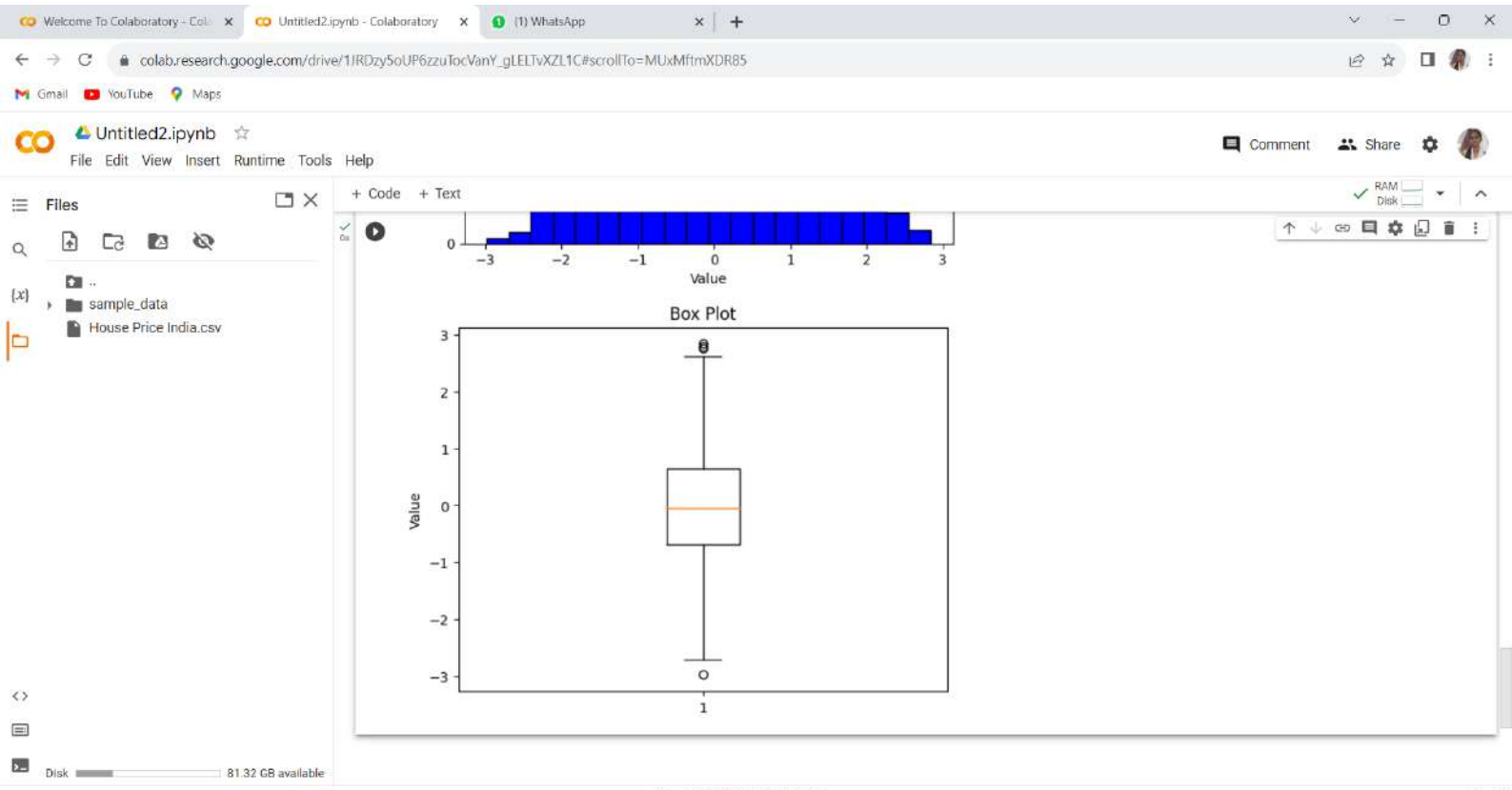
0s completed at 7:47 PM

GBP/INR +0.48%

Search

ENG IN

19:47 04-10-2023


Value
Box Plot
Value
1

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=tap-FaEgDjy7

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

RAM Disk

+ Code + Text

Files

..
sample_data
House Price India.csv

Disk 81.32 GB available

[5]

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

# Generate sample data (you can replace this with your own data)
np.random.seed(0)
x = np.random.rand(100)  # Numerical variable 1
y = 2 * x + 1 + np.random.randn(100)  # Numerical variable 2 with a linear relationship to x

# Create a pandas DataFrame
df = pd.DataFrame({'X': x, 'Y': y})

# Scatter plot
plt.scatter(df['X'], df['Y'], color='blue', alpha=0.7)
plt.xlabel('X')
plt.ylabel('Y')
plt.title('Scatter Plot of X vs. Y')
plt.grid(True)
plt.show()
```

0s completed at 7:48 PM

GBP/INR +0.48%

Search

ENG IN

19:48 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=tap-FaEgDjy7

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help Saving...

Comment Share

+ Code + Text

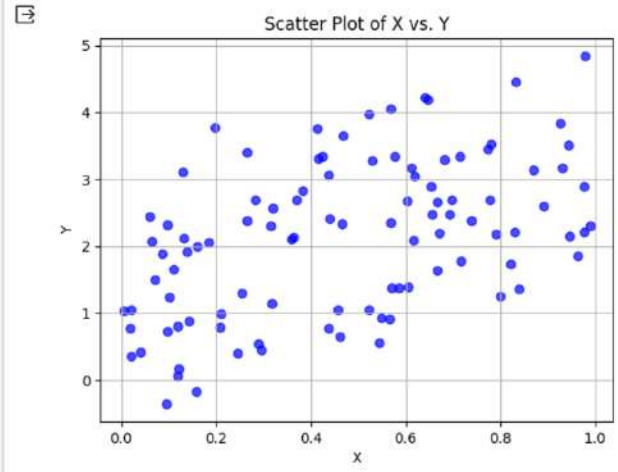
RAM Disk

Files

..
sample_data
House Price India.csv

Disk 81.32 GB available

```
plt.title('Scatter Plot of X vs. Y')
plt.grid(True)
plt.show()
```



0s completed at 7:48 PM

27°C Windy

Search

ENG IN

19:48 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help

Comment Share

Files

..
sample_data
House Price India.csv

Disk 81.32 GB available

+ Code + Text

RAM Disk

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

# Load the Iris dataset from Seaborn
iris = sns.load_dataset("iris")

# Explore the dataset
print(iris.head())

# Pairplot to visualize pairwise relationships between numerical features
sns.pairplot(iris, hue="species", markers=["o", "s", "D"])
plt.title("Pairplot of Iris Dataset")
plt.show()

# Correlation heatmap for numerical features
correlation_matrix = iris.corr()
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title("Correlation Heatmap of Iris Dataset")
plt.show()

# Box plots to compare the distribution of features by species
plt.figure(figsize=(12, 6))
plt.subplot(1, 2, 1)
sns.boxplot(x="species", y="sepal_length", data=iris)
plt.title("Sepal Length by Species")
plt.subplot(1, 2, 2)
sns.boxplot(x="species", y="sepal_width", data=iris)
plt.title("Sepal Width by Species")
```

0s completed at 7:48 PM

27°C Windy

Search

ENG IN

19:49 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help Saving...

Comment Share

Files

..
sample_data
House Price India.csv

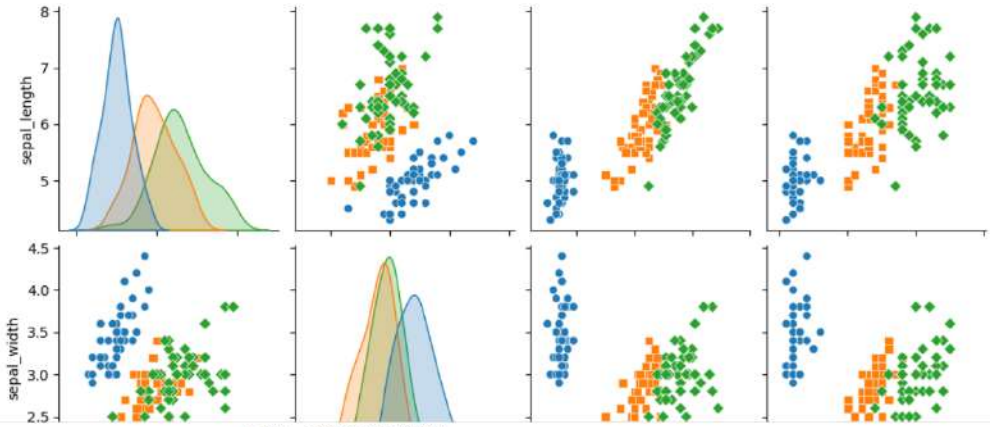
Disk 81.32 GB available

+ Code + Text

RAM Disk

```
plt.title("Petal Width by Species")
plt.show()
```

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa



✓ 20s completed at 7:50PM

EUR/INR +0.31%

Search

ENG IN

19:50 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

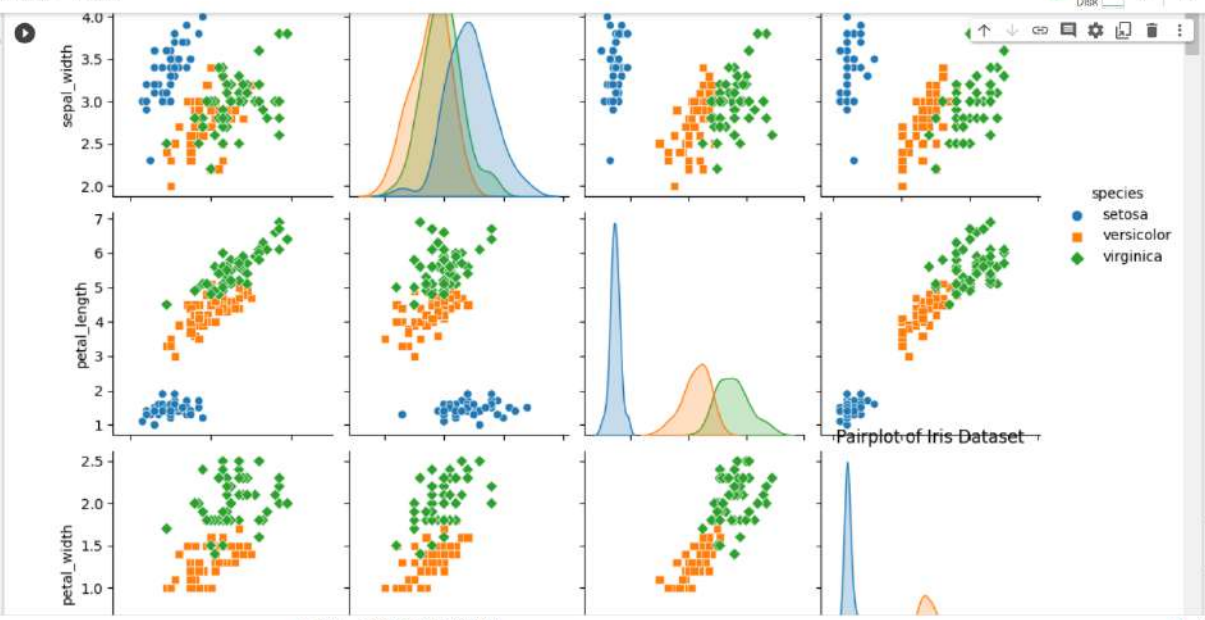
+ Code + Text

RAM Disk

Files

..
sample_data
House Price India.csv

Disk 81.32 GB available



20s completed at 7:50PM

EUR/INR +0.31%

Search

ENG IN

19:50 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

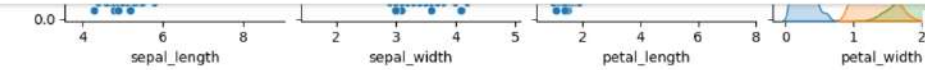
Files

..
sample_data
House Price India.csv

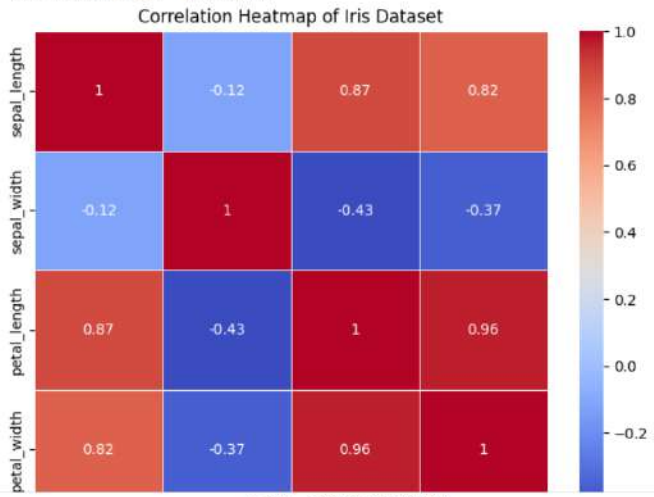
Disk 81.32 GB available

+ Code + Text

RAM Disk




```
<ipython-input-7-0081365d979a>:17: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it wi
  correlation_matrix = iris.corr()
```



✓ 20s completed at 7:50PM

EUR/INR +0.31%

Search

ENG IN

19:50 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

Files

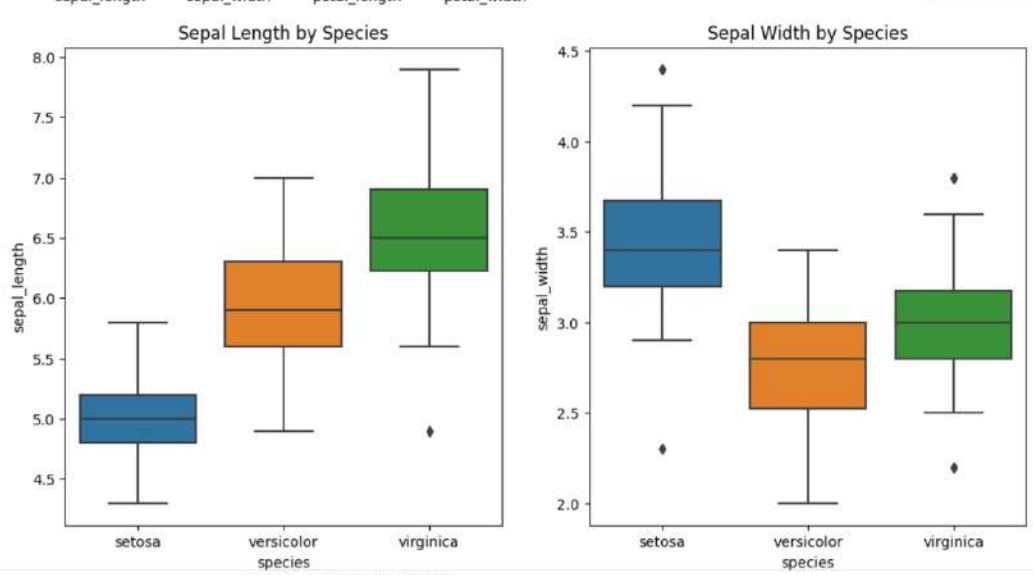
..
sample_data
House Price India.csv

Disk 81.32 GB available

+ Code + Text

RAM Disk

sepal_length sepal_width petal_length petal_width



20s completed at 7:50PM

EUR/INR +0.31%

Search

ENG IN

19:50 04-10-2023

Welcome To Colaboratory - Col | Untitled2.ipynb - Colaboratory | (1) WhatsApp

colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=3t9wr4IXDwJM

Gmail YouTube Maps

Untitled2.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

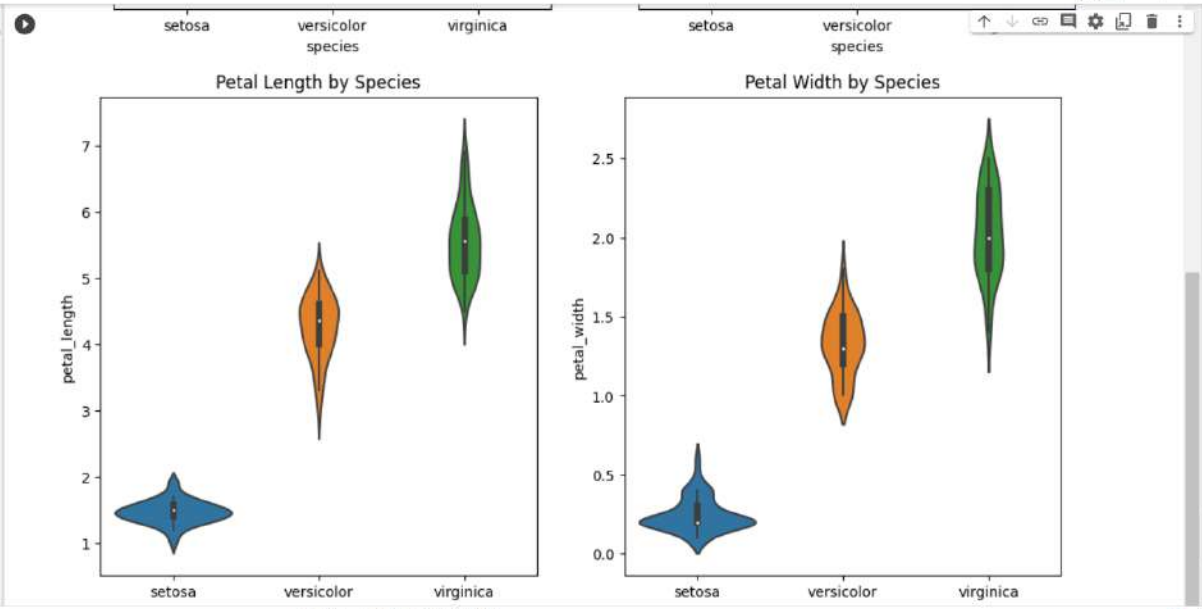
Files

..
sample_data
House Price India.csv

Disk 81.32 GB available

+ Code + Text

RAM Disk



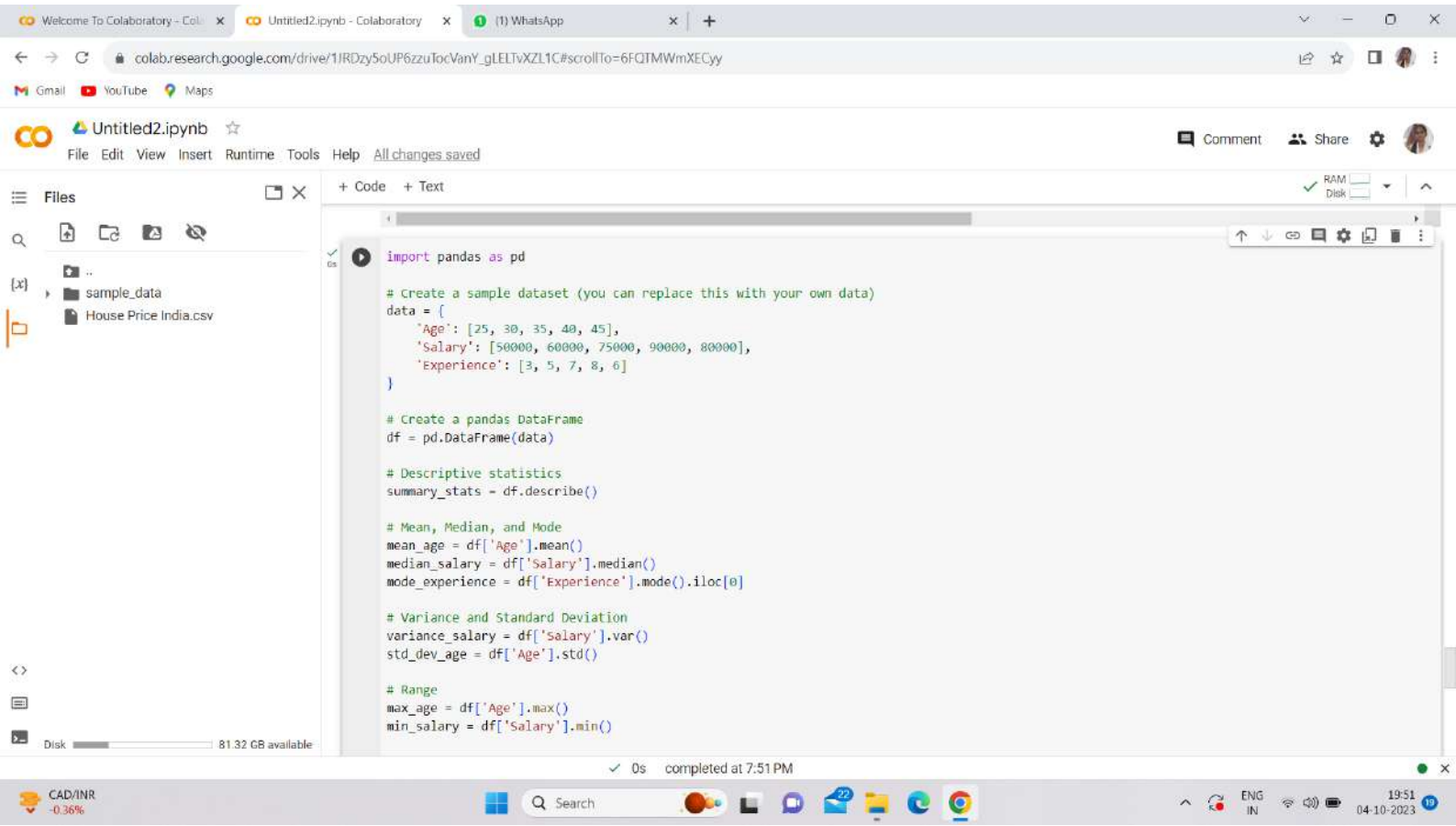
20s completed at 7:50PM

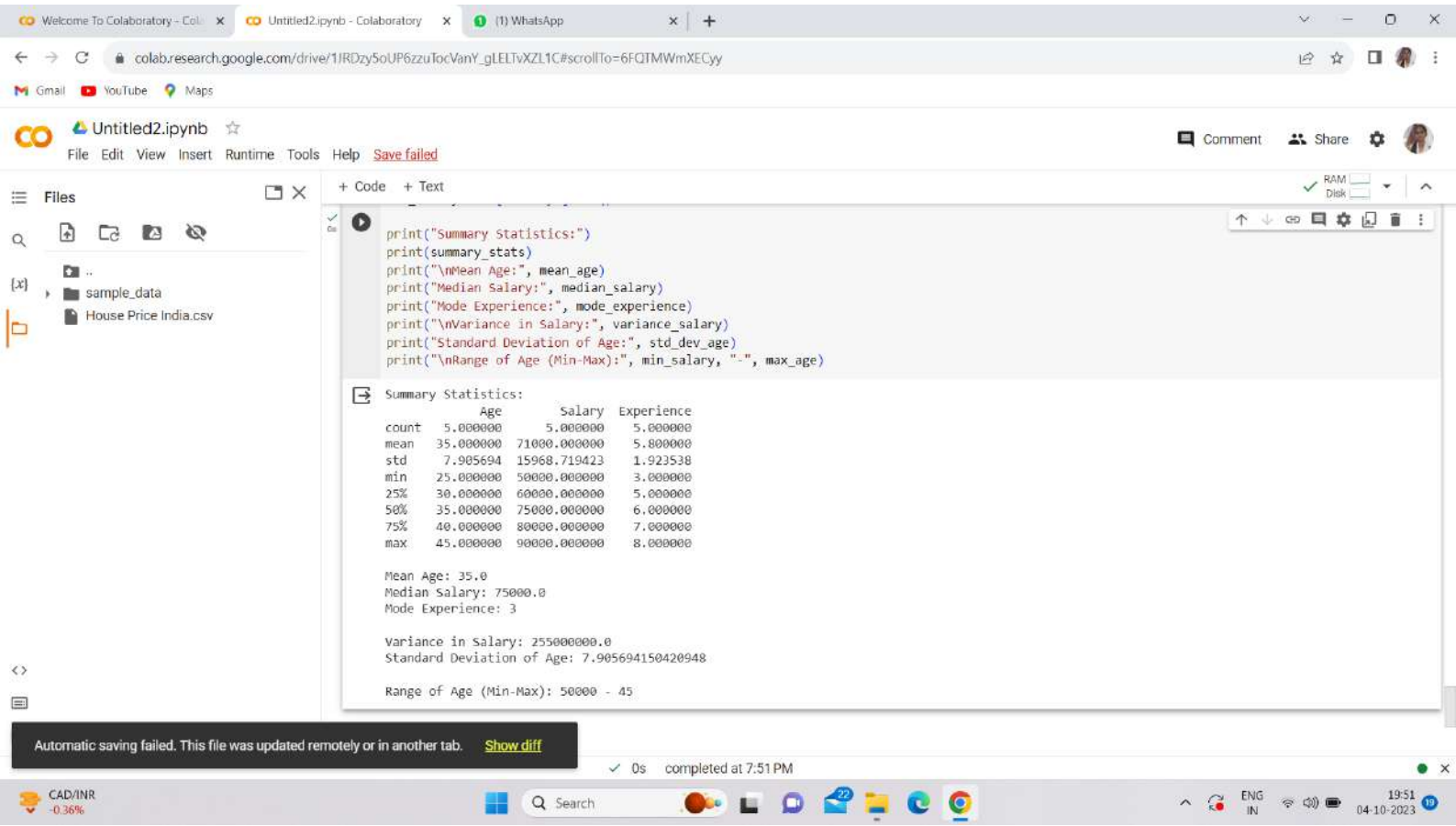
EUR/INR +0.31%

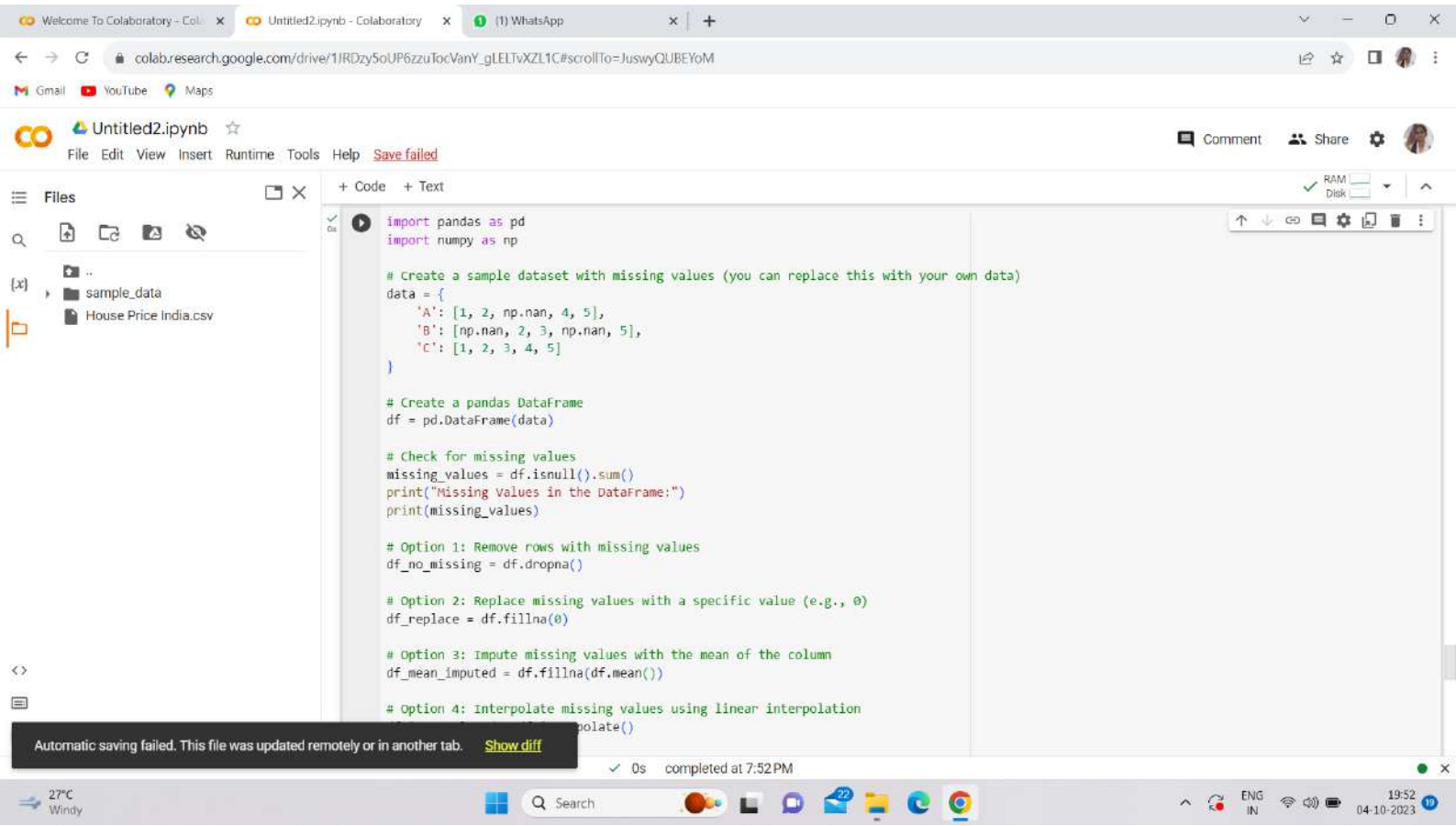
Search

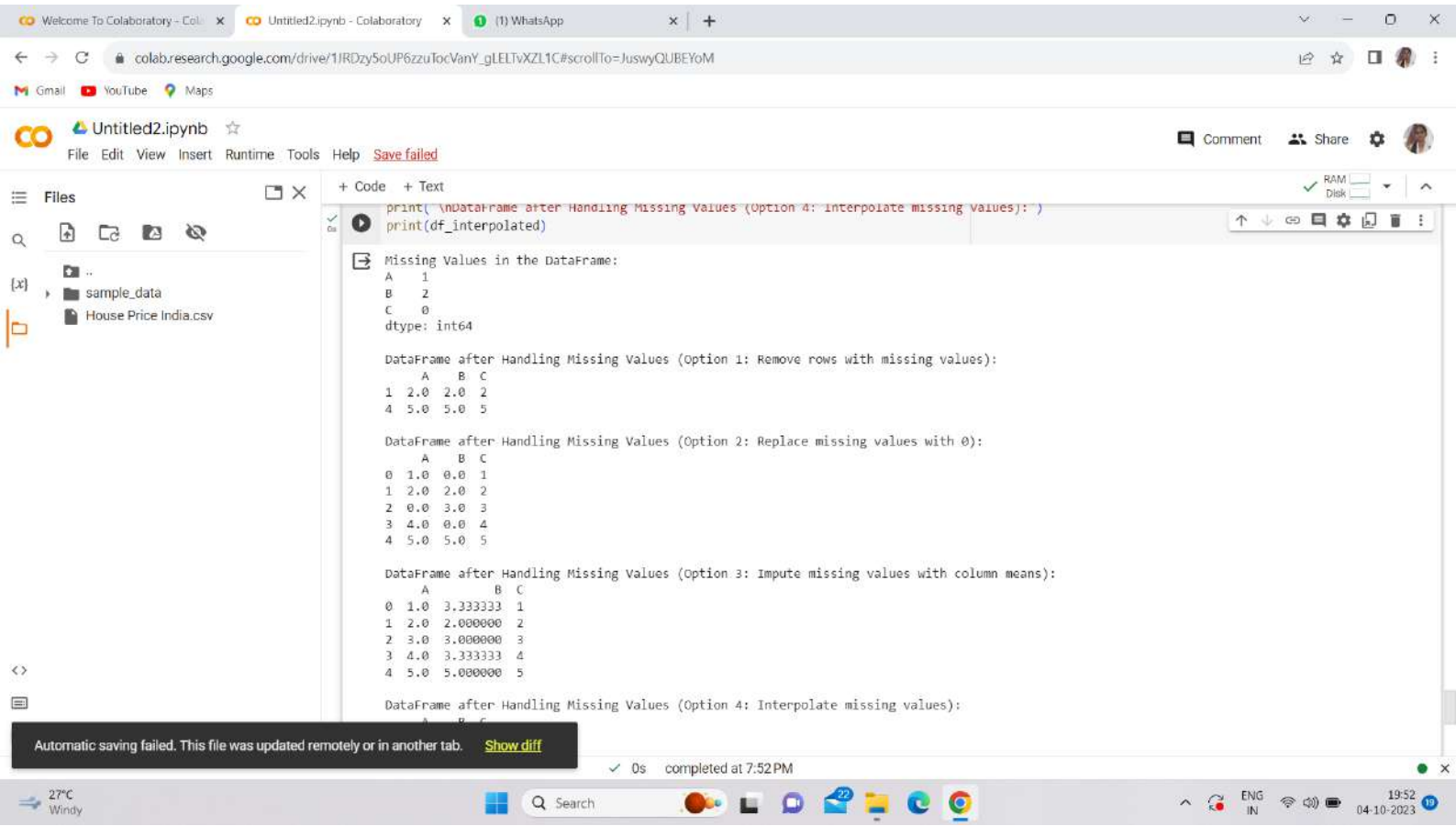
ENG IN

19:50 04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=6FQTMWmXECyy
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
..
sample_data
House Price India.csv
import pandas as pd
# Create a sample dataset (you can replace this with your own data)
data = {
'Age': [25, 30, 35, 40, 45],
'Salary': [50000, 60000, 75000, 90000, 80000],
'Experience': [3, 5, 7, 8, 6]
}
# Create a pandas DataFrame
df = pd.DataFrame(data)
# Descriptive statistics
summary_stats = df.describe()
# Mean, Median, and Mode
mean_age = df['Age'].mean()
median_salary = df['Salary'].median()
mode_experience = df['Experience'].mode().iloc[0]
# Variance and Standard Deviation
variance_salary = df['Salary'].var()
std_dev_age = df['Age'].std()
# Range
max_age = df['Age'].max()
min_salary = df['Salary'].min()
Disk
81.32 GB available
0s completed at 7:51 PM
CAD/INR
-0.36%
Search
ENG
IN
19:51
04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=6FQTMWmXECyy
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Save failed
Comment
Share
Files
+ Code + Text
RAM
Disk
..
sample_data
House Price India.csv
print("Summary Statistics:")
print(summary_stats)
print("\nMean Age:", mean_age)
print("Median Salary:", median_salary)
print("Mode Experience:", mode_experience)
print("\nVariance in Salary:", variance_salary)
print("Standard Deviation of Age:", std_dev_age)
print("\nRange of Age (Min-Max):", min_salary, "-", max_age)
Summary Statistics:
Age Salary Experience
count 5.000000 5.000000 5.000000
mean 35.000000 71000.000000 5.800000
std 7.905694 15968.719423 1.923538
min 25.000000 50000.000000 3.000000
25% 30.000000 60000.000000 5.000000
50% 35.000000 75000.000000 6.000000
75% 40.000000 80000.000000 7.000000
max 45.000000 90000.000000 8.000000
Mean Age: 35.0
Median Salary: 75000.0
Mode Experience: 3
Variance in Salary: 255000000.0
Standard Deviation of Age: 7.905694150420948
Range of Age (Min-Max): 50000 - 45
Automatic saving failed. This file was updated remotely or in another tab.
Show diff
0s completed at 7:51 PM
CAD/INR
-0.36%
Search
ENG
IN
19:51
04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=JuswyQUBEYoM
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Save failed
Comment
Share
Files
+ Code + Text
RAM
Disk
..
sample_data
House Price India.csv
import pandas as pd
import numpy as np
# Create a sample dataset with missing values (you can replace this with your own data)
data = {
'A': [1, 2, np.nan, 4, 5],
'B': [np.nan, 2, 3, np.nan, 5],
'C': [1, 2, 3, 4, 5]
}
# Create a pandas DataFrame
df = pd.DataFrame(data)
# Check for missing values
missing_values = df.isnull().sum()
print("Missing Values in the DataFrame:")
print(missing_values)
# Option 1: Remove rows with missing values
df_no_missing = df.dropna()
# Option 2: Replace missing values with a specific value (e.g., 0)
df_replace = df.fillna(0)
# Option 3: Impute missing values with the mean of the column
df_mean_imputed = df.fillna(df.mean())
# Option 4: Interpolate missing values using linear interpolation
Automatic saving failed. This file was updated remotely or in another tab.
Show diff
0s completed at 7:52 PM
27°C
Windy
Search
ENG
IN
19:52
04-10-2023


Welcome To Colaboratory - Col
Untitled2.ipynb - Colaboratory
(1) WhatsApp
colab.research.google.com/drive/1JRDzy5oUP6zzuTocVanY_gLELTvXZL1C#scrollTo=JuswyQUBEYoM
Gmail
YouTube
Maps
Untitled2.ipynb
File Edit View Insert Runtime Tools Help Save failed
Comment
Share
Files
sample_data
House Price India.csv
+ Code + Text
RAM
Disk
print(\nDataFrame after Handling Missing Values (Option 4: Interpolate missing values):)
print(df_interpolated)
Missing Values in the DataFrame:
A 1
B 2
C 0
dtype: int64
DataFrame after Handling Missing Values (Option 1: Remove rows with missing values):
A B C
1 2.0 2.0 2
4 5.0 5.0 5
DataFrame after Handling Missing Values (Option 2: Replace missing values with 0):
A B C
0 1.0 0.0 1
1 2.0 2.0 2
2 0.0 3.0 3
3 4.0 0.0 4
4 5.0 5.0 5
DataFrame after Handling Missing Values (Option 3: Impute missing values with column means):
A B C
0 1.0 3.333333 1
1 2.0 2.000000 2
2 3.0 3.000000 3
3 4.0 3.333333 4
4 5.0 5.000000 5
DataFrame after Handling Missing Values (Option 4: Interpolate missing values):
Automatic saving failed. This file was updated remotely or in another tab.
Show diff
0s completed at 7:52 PM
27°C
Windy
Search
ENG
IN
19:52
04-10-2023