LAB Assignment No. 1

Lab Assignment – Dataset Creation & Analysis Objective

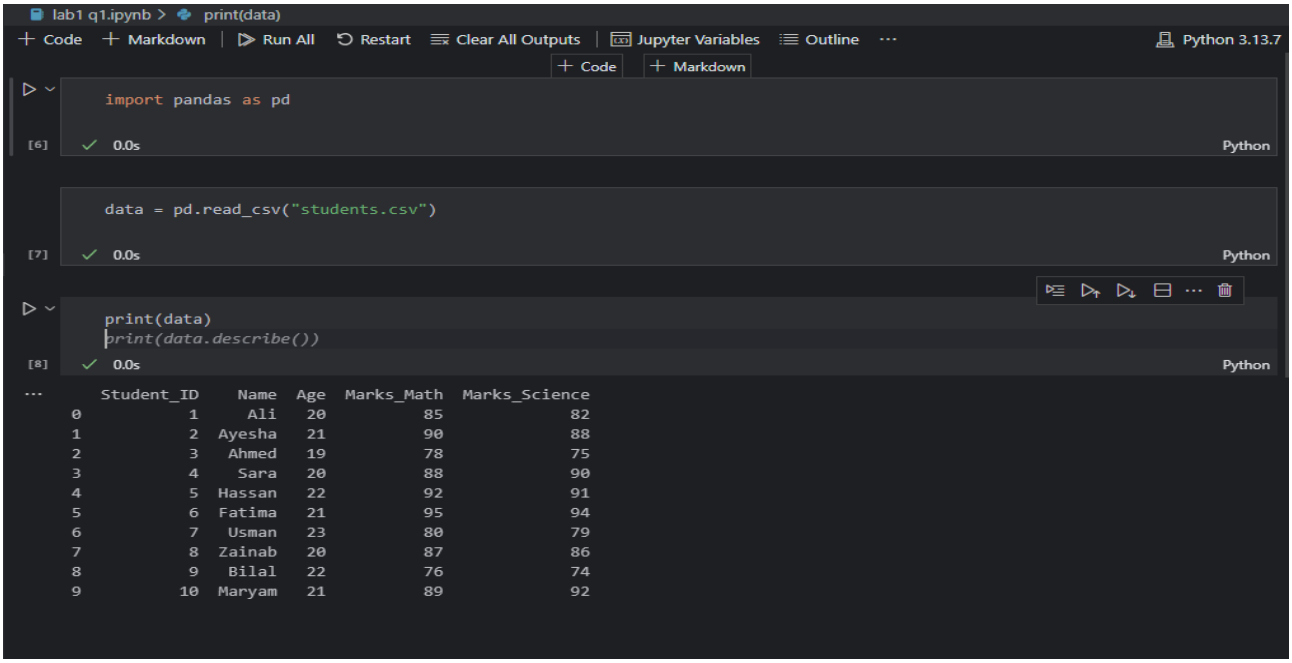
To learn how to create and upload a dataset in Python, perform basic statistical analysis, and visualize data using graphs.

Tasks Q1: Create a Dataset Manually

- Create a dataset of at least **10 students** with the following columns:
    - Student_ID, o
      Name, o
      Age, o
      Marks_Math, o
      Marks_Science.
- Store the dataset in a **CSV file** named students.csv.

Q2: Upload Dataset in Python

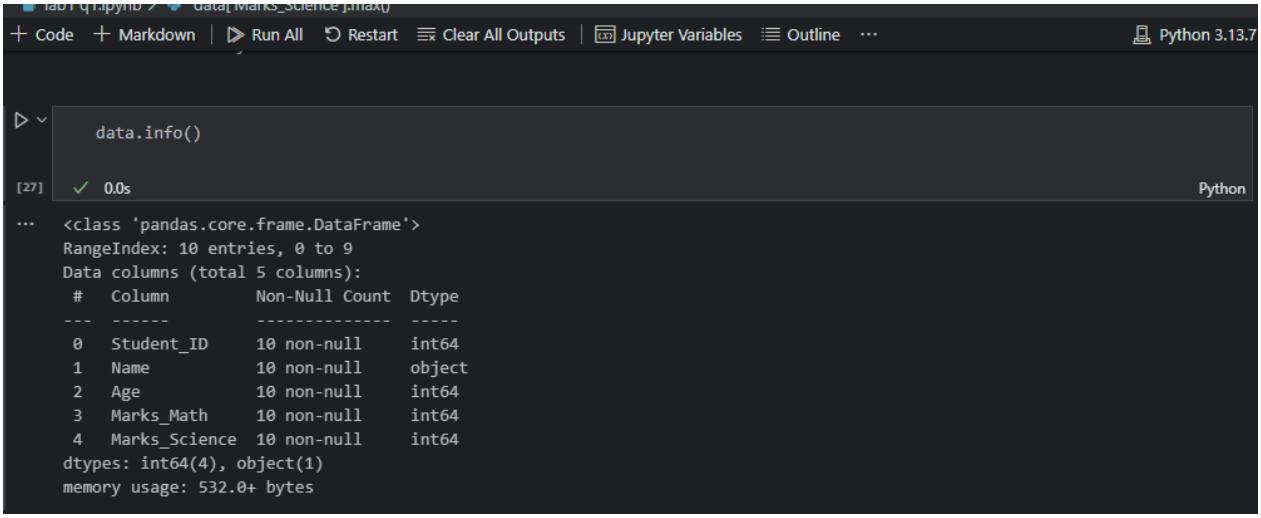
- Use **Pandas** to load the dataset.



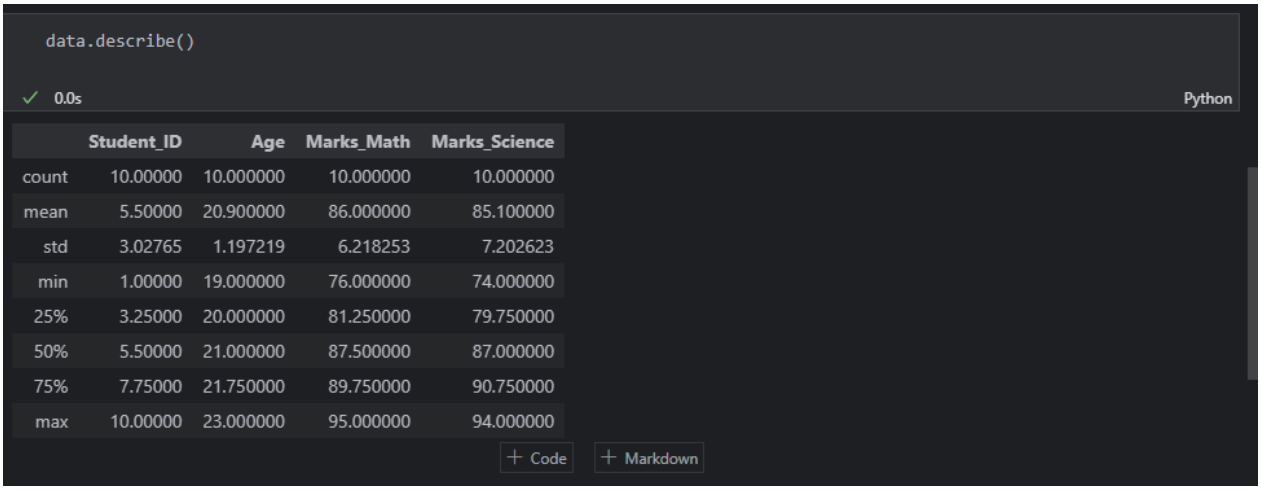
Q3: Observe Dataset Information

Run the following commands and explain the output:

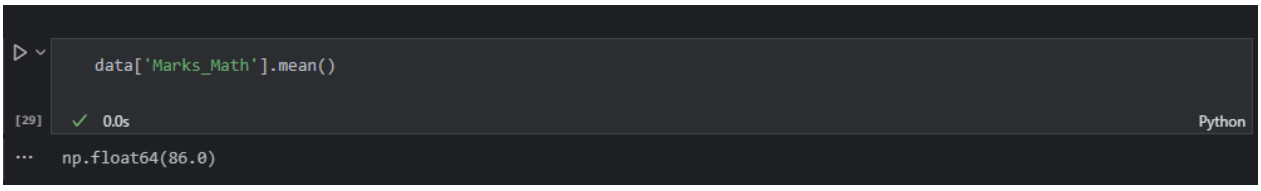
1. data.info() → Dataset structure



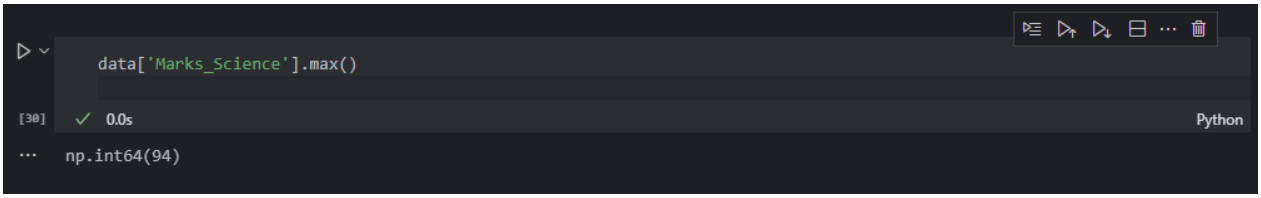
2. data.describe() → Summary statistics (mean, std, min, max, etc.)



3. data['Marks_Math'].mean() → Mean of Math marks

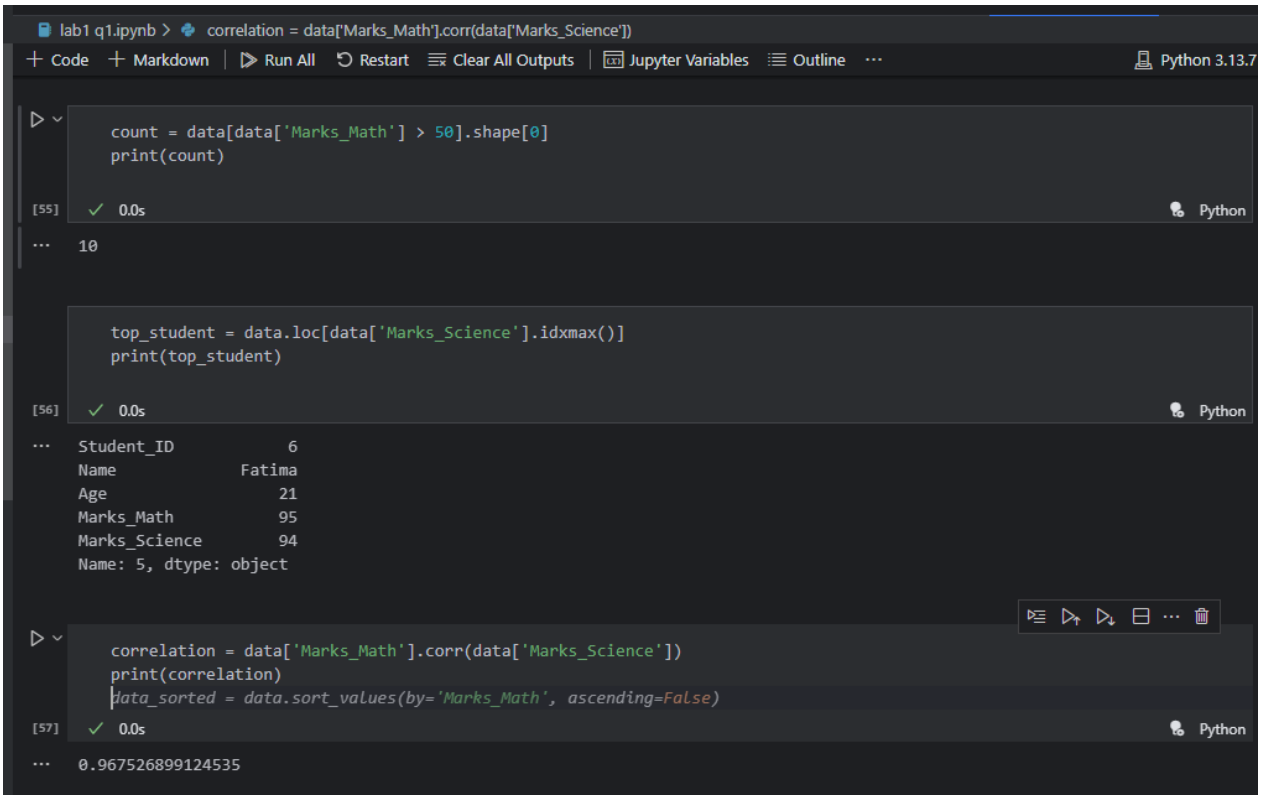


4. data['Marks_Science'].max() → Maximum Science marks



Q4: Perform Some Data Analysis

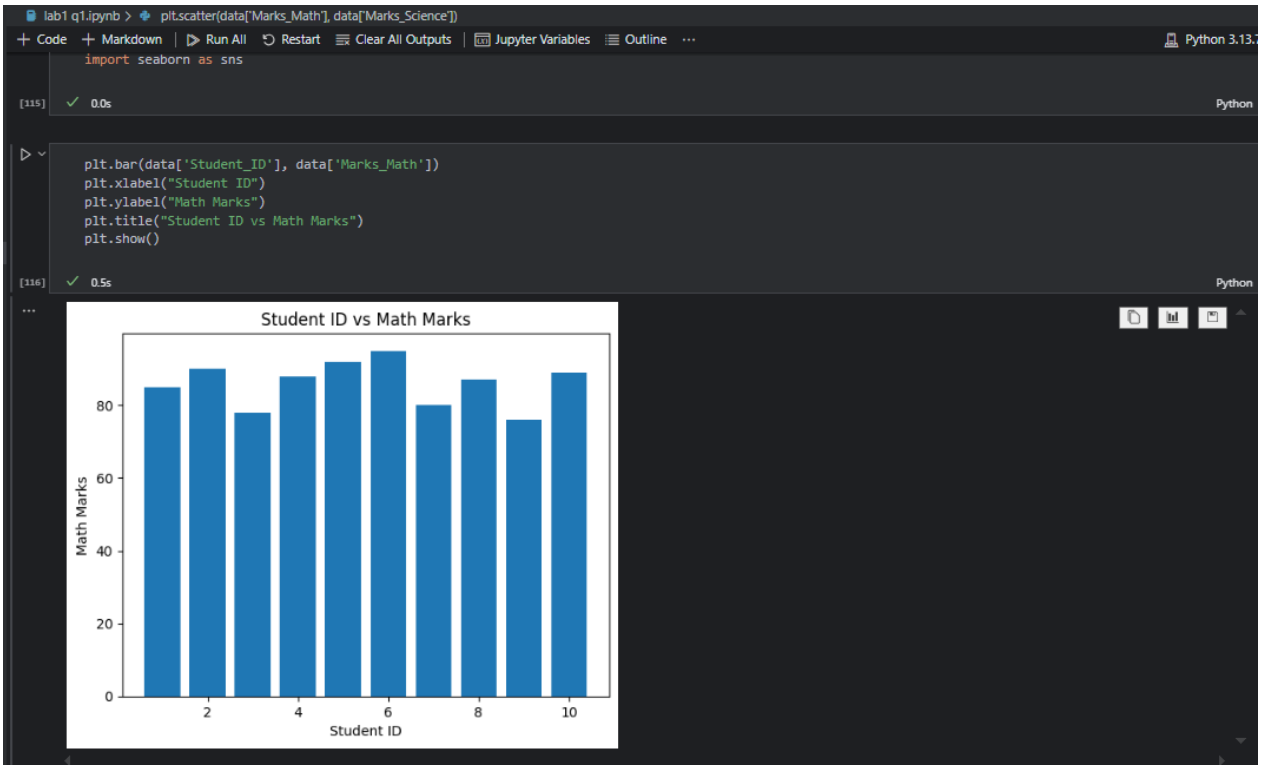
- Find how many students have Marks_Math > 50.
- Find the student with the **highest Science marks**.
- Calculate the **correlation** between Marks_Math and Marks_Science.



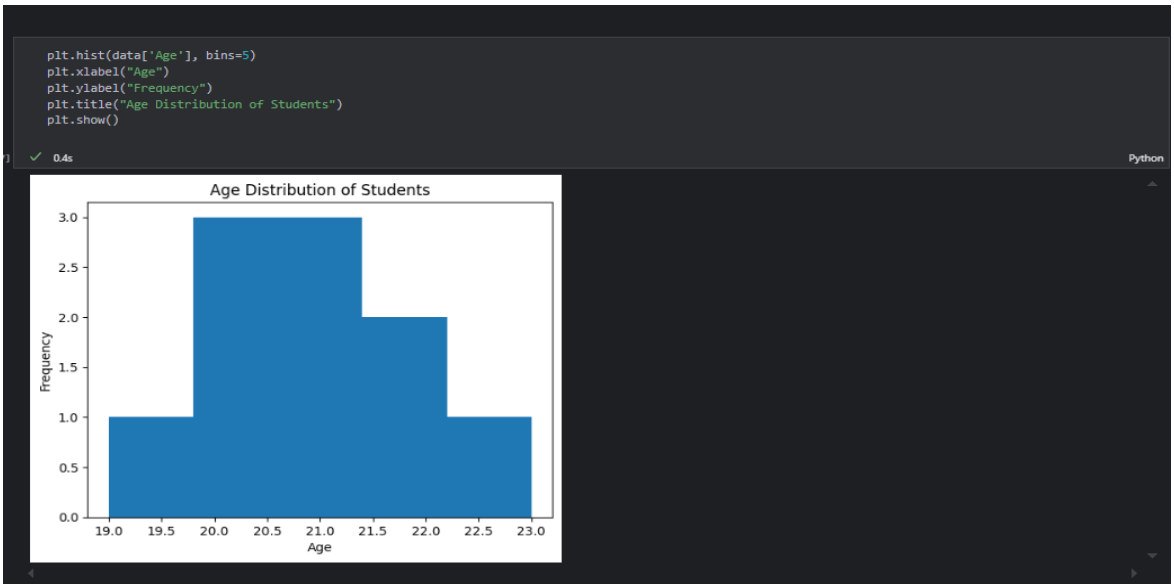
Q5: Data Visualization

Use **Matplotlib/Seaborn** to create graphs:

1. A bar chart of Student_ID vs Marks_Math.



A histogram of Age.



2. A scatter plot of Marks_Math vs Marks_Science.

```
plt.scatter(data['Marks_Math'], data['Marks_Science'])
plt.xlabel("Math Marks")
plt.ylabel("Science Marks")
plt.title("Math vs Science Marks")
plt.show()
```

0.3s Python