Amirtha Raja Rajeswari N A

231501012

LAB EXPERIMENTS – 1 to 5

FRAMEWORK FOR DATA AND VISUAL ANALYTICS

AD23632

## Ex 1 Setting up the Python environment and libraries - Juypter Notebook

**AIM:**  
To understand the working of Jupyter Notebook, write and execute Python code, create new code and Markdown cells, and demonstrate the use of Jupyter Widgets and Jupyter AI.

1. Create a new notebook for Python

Open Anaconda Navigator or Jupyter Lab / Notebook.  
Click **New → Python 3** to open a fresh notebook.

1. Write and execute Python code

a = 5

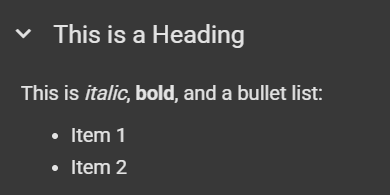
b = 7

print(a + b)

OUTPUT:



1. Create new cells for code and Markdown



1. Demonstrate the application of Jupyter Widgets, Jupyter AI

!jupyter labextension install @jupyter-widgets/jupyterlab-manager

import ipywidgets as widgets

widgets.IntSlider(

value=10,

min=0,

max=100,

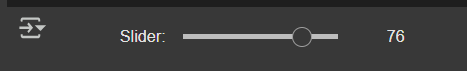
step=1,

description='Slider:',

continuous\_update=True

)

OUTPUT:



## Ex 2 EDA-Data Import and Export

## 

**AIM:**  
To perform exploratory data analysis by importing data from various sources such as CSV, Excel, SQL, and web scraping, and export DataFrames into Excel and CSV formats using Python.

1. Importing data from CSV, Excel, SQL databases, and web scraping

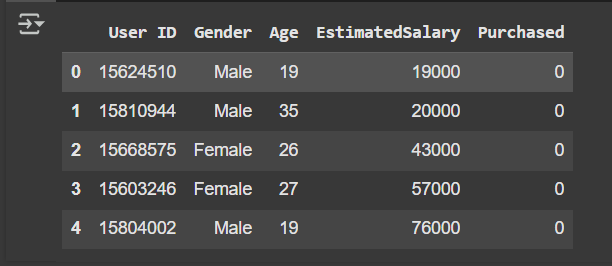
* **CSV**

import pandas as pd

df = pd.read\_csv('/content/suv\_data.csv')

df.head()

OUTPUT:



* **EXCEL**

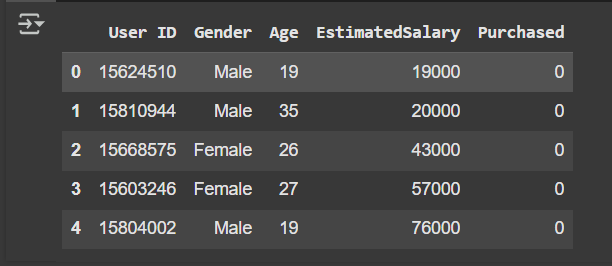
!pip install openpyxl

df.to\_excel('suv\_data.xlsx', index=False)

df2 = pd.read\_excel('suv\_data.xlsx')

df2.head()

OUTPUT:



* **SQL DB**

import sqlite3

import pandas as pd

conn = sqlite3.connect('[mydata.db](http://mydata.db)')

df = pd.read\_csv('suv\_data.csv') # Your existing data

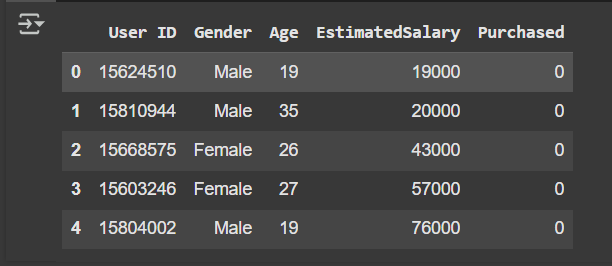
df.to\_sql('suv\_table', conn, if\_exists='replace', index=False) # Store to SQL

df\_sql = pd.read\_sql\_query("SELECT \* FROM suv\_table", conn)

df\_sql.head()

conn.close()

OUTPUT:



* **WEB SCRAPING**

import pandas as pd

url = 'https://en.wikipedia.org/wiki/List\_of\_countries\_by\_GDP\_(nominal)'

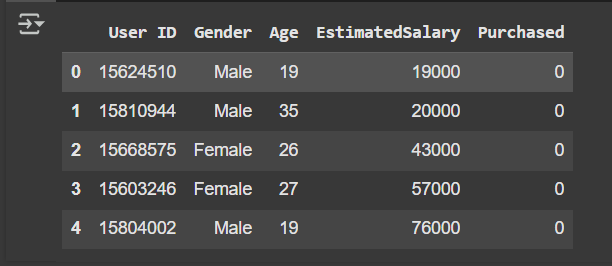
tables = pd.read\_html(url) # This will return a list of tables

print(len(tables)) # See how many tables were found

df\_web = tables[1] # You can try 0, 1, 2, etc.

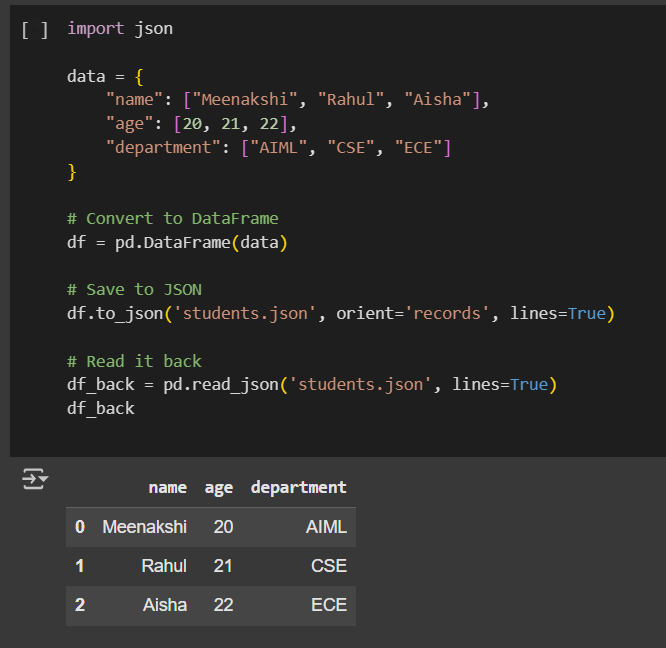
df\_web.head()

OUTPUT:

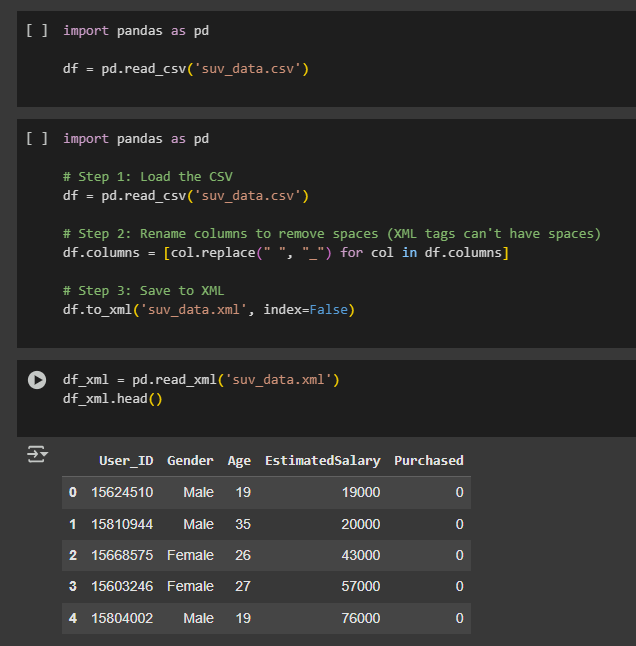


1. Handling different data formats

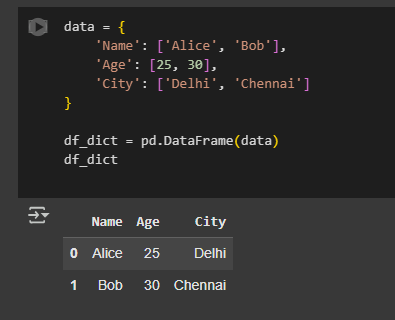
* **JSON**



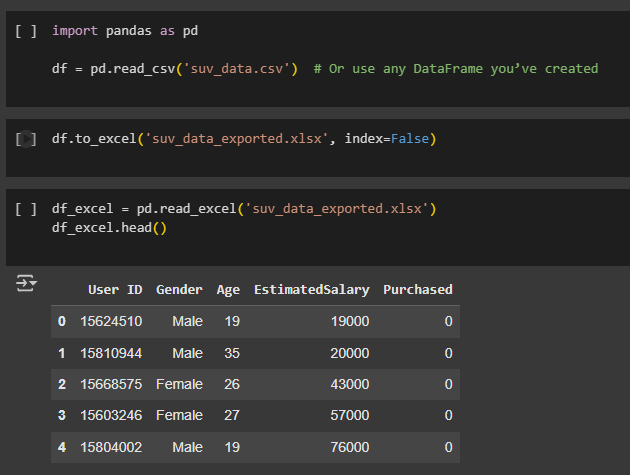
* **XML**



* **PYTHON DICTIONARY**



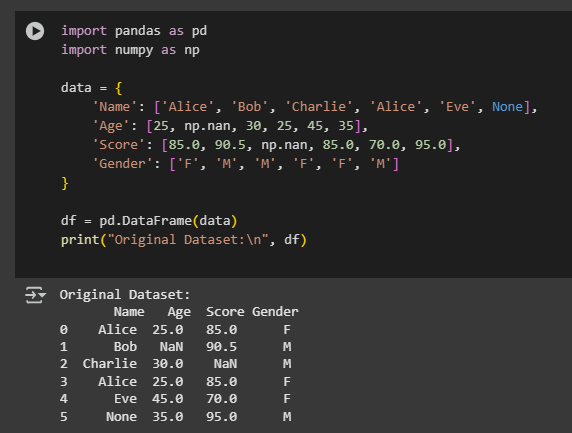
1. Export a DataFrame to an Excel file.

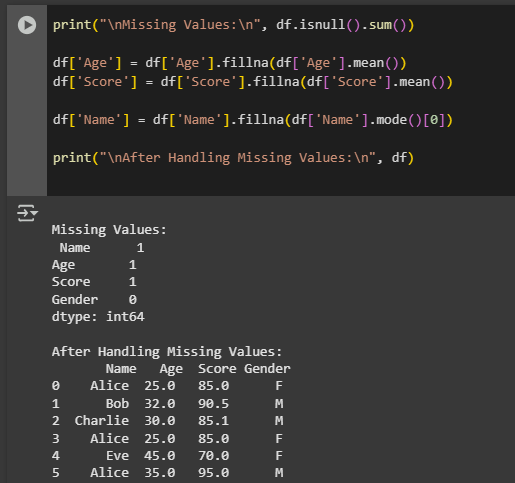


## Ex 3 EDA-Data Cleaning

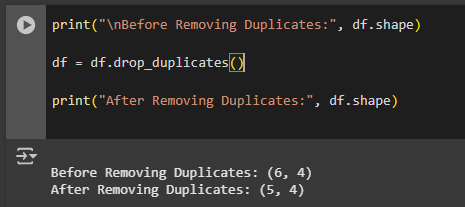
**AIM:**  
To clean the dataset by handling missing values, removing duplicates, performing data type conversion, and normalizing data using standardization and min-max scaling.

1. Handlingmissing values: detection, filling, and dropping

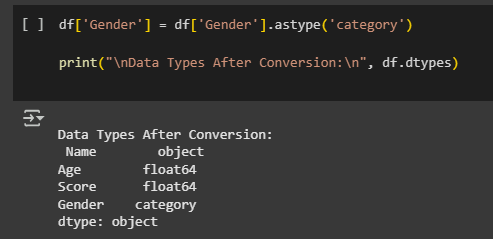




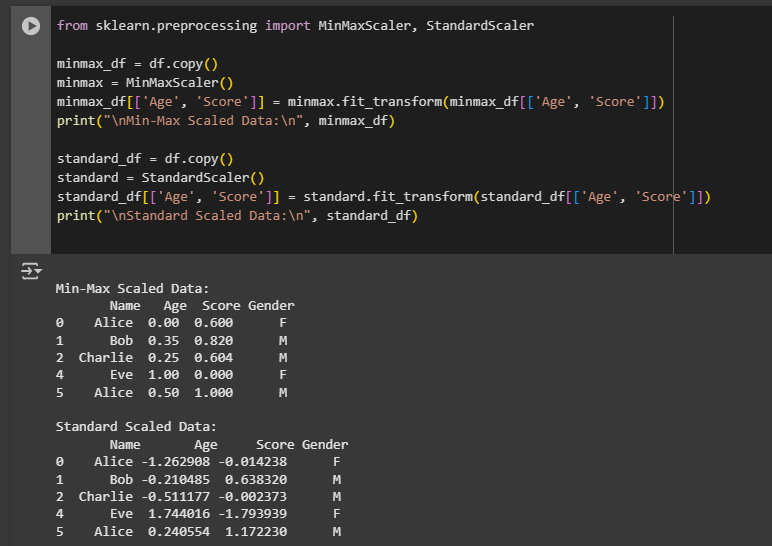
2. Removing duplicates and unnecessary data



1. Data type conversion and ensuring consistency



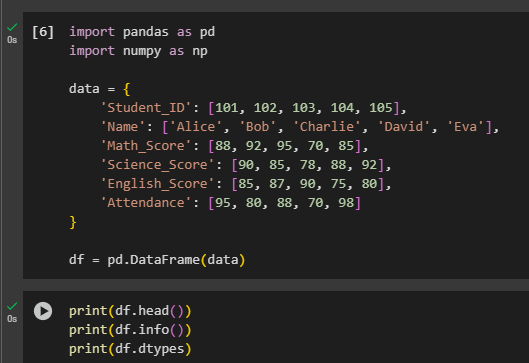
1. Normalize data (e.g., standardization, min-max scaling).



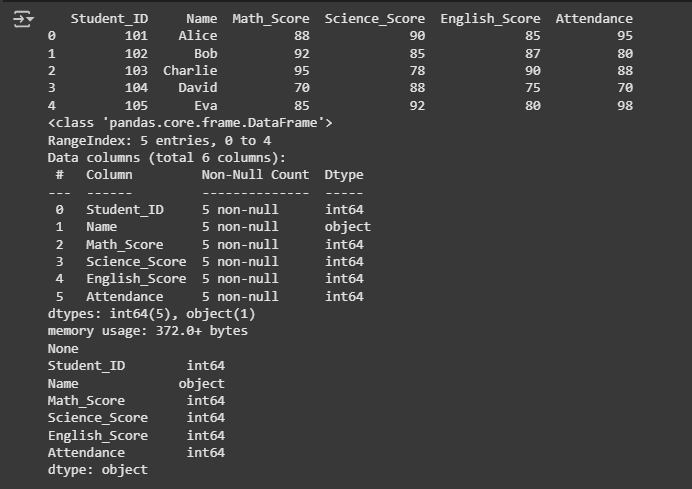
## Ex 4 EDA-Data Inspection and Analysis

**AIM:**  
To inspect and analyze datasets by viewing DataFrames, filtering and subsetting data using conditions, and calculating descriptive statistics including measures of central tendency and dispersion.

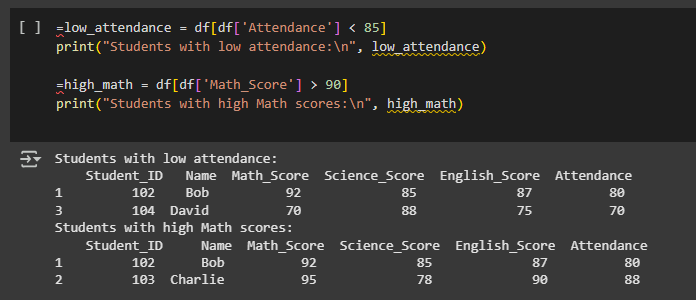
1. Viewing and inspecting DataFrames



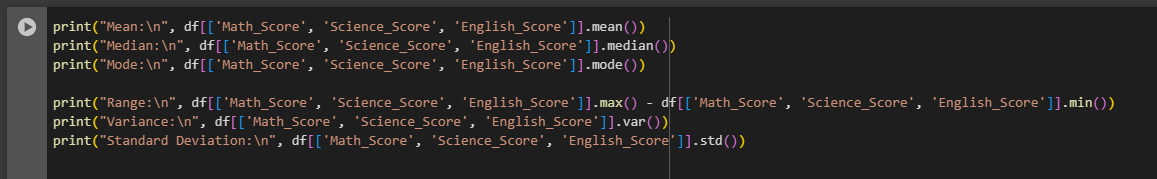
OUTPUT:

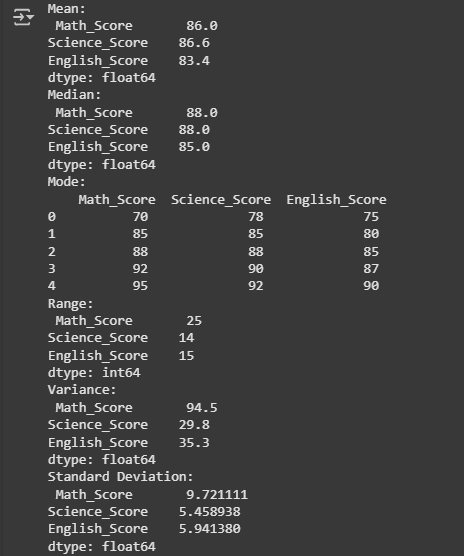


1. Filtering and subsetting data using conditions



1. Descriptive statistics: measures of central tendency (mean, median, mode) and measures of dispersion (range, variance, standard deviation)

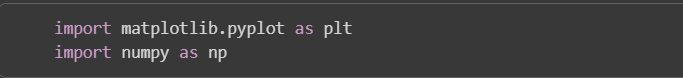


****

## Ex 5 EDA-Data Visualization with Matplotlib

AIM:

To visualize and understand datasets using basic plots- **line charts, bar charts** and **histograms** with the Matplotlib library in Python.

 A screen shot of a computer code

AI-generated content may be incorrect.A graph with a line going up

AI-generated content may be incorrect.

A computer screen shot of text

AI-generated content may be incorrect.

A graph of a bar chart

AI-generated content may be incorrect.

A screen shot of a computer code

AI-generated content may be incorrect.

A graph of a distribution

AI-generated content may be incorrect.