***Hamna Jamila***

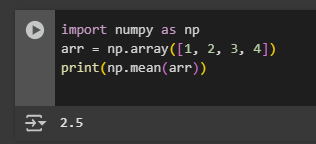
**AKSA- AI Internship**

**Data Handling, Cleaning, and Visualization**

**1. Data Handling with NumPy and Pandas**

**NumPy (Numerical Python)**

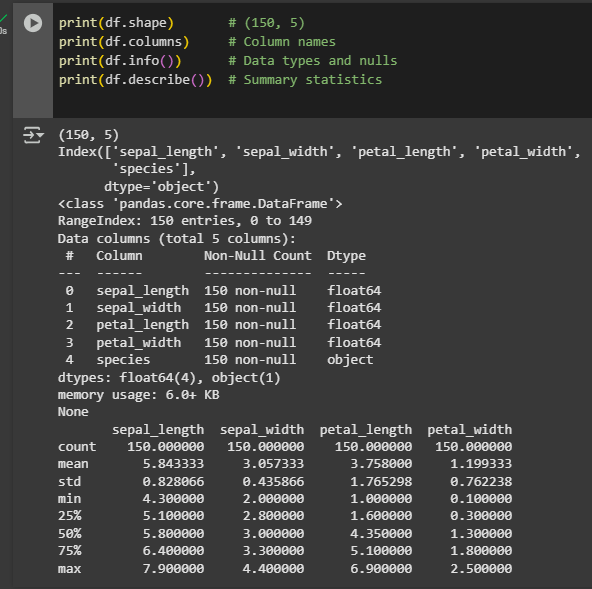
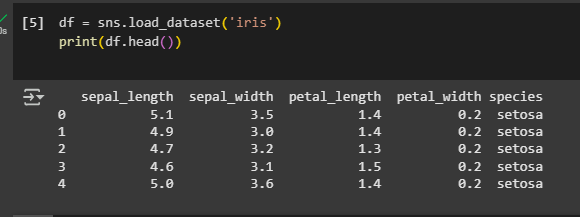
* **What is numpy?**  
  Numpy is a python library used for working with arrays. It helps perform mathematical and logical operations on large datasets easily.
* **Concepts:**
  + **ndarray**: n-dimensional array like a list but faster and more powerful
  + **Vectorized operations**: Perform operations on entire arrays without using loops
  + **Functions**: np.array(), np.mean(), np.sum(), np.arange(), np.linspace()
* **Example:**



I’ll be taking the example of iris data set for all these concepts

**Pandas**

* **What is Pandas?**  
  Pandas is a powerful Python library for data analysis and manipulation. It uses two main structures: **Series** (1D) and **DataFrame** (2D).
* **Concepts:**
  + **Reading data**: pd.read\_csv(), pd.read\_excel()
  + **Accessing data**: .head(), .tail(), .shape, .info(), .describe()
  + **Selecting data**: df['column'], df.loc[], df.iloc[]
* **Example:**

****

**2. Data Cleaning & Manipulation**

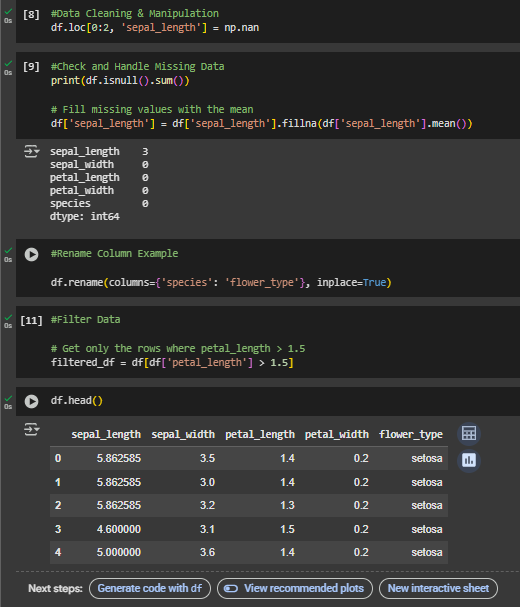
**Why Clean Data?**

Real-world data is messy. Cleaning helps remove incorrect, missing, or duplicate values to make data reliable for analysis.

**Common Cleaning Tasks:**

* **Handling Missing Values:**
  + df.isnull(), df.dropna(), df.fillna(value)
* **Removing Duplicates:**
  + df.duplicated(), df.drop\_duplicates()
* **Fixing Data Types:**
  + df['col'] = df['col'].astype(int)
* **Renaming Columns:**
  + df.rename(columns={'old': 'new'})

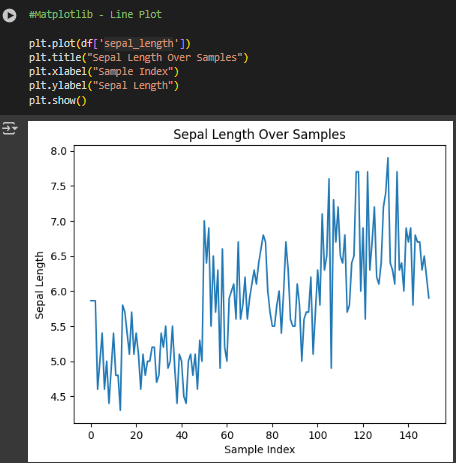
**Example:**

****

**3. Visualization with Matplotlib, Seaborn, and Plotly**

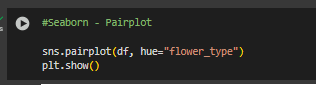
**Matplotlib**

* **What is it?**  
  A basic plotting library to create static graphs.
* **Common Plots:** Line, Bar, Histogram, Pie
* **Example:**



**Seaborn**

* **What is it?**  
  Built on top of Matplotlib, Seaborn makes it easier to make attractive and informative statistical graphics.
* **Common Plots:** Heatmap, Boxplot, Barplot, Pairplot
* **Example:**



A screenshot of a computer screen

AI-generated content may be incorrect.

**Plotly**

* **What is it?**  
  A library for creating interactive plots for web applications.
* **Example:**

A screen shot of a computer

AI-generated content may be incorrect.

