

Tasks (use NumPy only):

1. Create Arrays

- A 1D array of 20 random integers from 1 to 100
- A 5×5 matrix with values from 1 to 25

2. Reshape and Slicing

- Reshape the 1D array into a 4×5 matrix
- Extract the 2nd and 4th columns

3. Mathematical Operations

- Find the mean, median, max, and standard deviation of your matrix
- Replace all values divisible by 5 with -1

4. Boolean Masking

- Get all values greater than the matrix mean
- Count how many values are less than 10

Section 2: Pandas

Use iris.csv

◆ Tasks:

1. Load & Explore

- Load the dataset using `pd.read_csv()`
- Show: `.head()`, `.info()`, `.describe()`

2. Selection & Filtering

- Select only columns 'sepal_length' and 'species'
- Filter rows where `petal_width > 1.0` and `species == 'setosa'`

3. Group & Aggregate

- Group by species and calculate the average of `sepal_length`
- Count how many records per species

4. Missing Values

- Inject missing values in some rows

- Fill missing sepal_length with column mean
- Drop rows where species is missing

5. **Sorting & Value Counts**

- Sort the DataFrame by petal_length in descending order
- Find top 2 most frequent species

Section 3: Matplotlib

Tasks (use Matplotlib & optionally Pandas plots):

1. **Line Plot**

- Plot the average sepal length for each species

2. **Bar Plot**

- Count of each species as a bar plot (.value_counts()).plot(kind='bar')

3. **Scatter Plot**

- Plot sepal_length vs petal_length, color by species

4. **Histogram**

- Plot histogram of sepal_width with 10 bins

5. **Customization**

- Add title, xlabel, ylabel, grid, and legend to each plot

-
1. Plot a **boxplot** comparing sepal_length across different species.
 2. Use df.corr() and .heatmap() from seaborn (if allowed).
 3. Export your cleaned DataFrame to CSV.

Submission Guidelines

- Submit as a .ipynb (Jupyter Notebook) or .py script
- Include markdown/text for short explanations
- Use comments to explain steps