NumPy, Pandas, and MatplotLIb

Key concepts

- Load and explore a DataFrame
- Perform simple transformations
- Run basic statistical analyses
- Create a plot

Reading and Exploring a DataFrame

Run the following code and register its output

```python
import pandas

# Example data
data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'Diana', 'Ethan'],
    'Age': [25, 30, 35, 40, 45],
    'Score': [88, 92, 85, 90, 87]
}

df = pandas.DataFrame(data)
df
```

Run the following code and register its output

```python
import pandas as pd

# Example data
data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'Diana', 'Ethan'],
    'Age': [25, 30, 35, 40, 45],
    'Score': [88, 92, 85, 90, 87]
}

df = pd.DataFrame(data)
print(df)
```

Challenge: What is the average score? What is the youngest age?

Use Pandas methods to find:

- The average score

- The youngest age

Modifying a Dataframe

Run the following code and register its output

```python
import pandas as pd

data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'Diana', 'Ethan'],
    'Age': [25, 30, 35, 40, 45],
    'Score': [88, 92, 85, 90, 87]
}

df = pd.DataFrame(data)

new_data = ['Fail', 'Pass', 'Fail', 'Pass', 'Fail']

df["Result"] = new_data

print(df)
```

Now add this and register the new output

```python
# Add a new column with Pass/Fail based on Score >= 90
# print(df['Score'].apply(lambda x: 'Pass' if x >= 90 else 'Fail'))
df['Result'] = df['Score'].apply(lambda x: 'Pass' if x >= 90 else 'Fail')
print(df)
```

Challenge: Add a new column that classifies people into 'Young' (<35) and 'Old' (>=35)

Add a new column that classifies people into 'Young' (<35) and 'Old' (>=35).

Basic statistics with NumPy

Run the following code and register the output

```python
import numpy as np

ages = df['Age'].to_numpy()
scores = df['Score'].to_numpy()

print(f"ages = {ages}")
print(f"scores = {scores}")

print("Standard Deviation of Scores:", np.std(scores))
print("Mean Age:", np.mean(ages))
```

Challenge: What is the correlation between Age and Score?

Use `np.corrcoef` to find the correlation.

Plotting with Matplotlib

Run the following code and register its output

```python
import matplotlib.pyplot as plt

plt.plot(df['Age'], df['Score'], marker='o')
plt.title('Score by Age')
plt.xlabel('Age')
plt.ylabel('Score')
plt.grid(True)
plt.show()
```

Challenge: Create a bar plot showing how many students passed vs failed

Create a bar plot showing how many students passed vs failed.