Knowledge Streams

Numpy Mini Project - 01

Total Points: 10

Learning Outcomes

By the end of this Numpy Mini Project, students will be able to:

1. **Data Handling**: Load and manipulate datasets using NumPy.
2. **Data Analysis**: Perform basic statistical analyses such as finding maximum, minimum, mean, and standard deviation.
3. **Subsetting Data**: Create subsets of data based on specific criteria.
4. **Comparative Analysis**: Compare different subsets of data to draw meaningful insights.
5. **Policy Evaluation**: Evaluate real-world policies using data-driven approaches.
6. **Data Interpretation**: Interpret data to understand demographic and socio-economic trends.

Most importantly, IF YOU DO NOT UNDERSTAND ANYTHING Syntax or Logic wise, PLEASE VISIT THE NUMPY DOCUMENTATION (It is your best friend 🙂)

This project evaluates your understanding of the first week so Instructors will not help you in this.

Well, if you are ready now, move to the next page!

CONGRATULATIONS!

You graduated from KS and started working as a data analyst for NADRA. Your first task is to load and inspect the census data of Lahore.

Step 1: Load the data

- You receive a CSV file named makeSenseOfCensus.csv containing the census data. Your job is to load this data into a NumPy array for analysis.
- To understand what the data looks like, print the entire dataset.
- Check the type of the data structure you've loaded to ensure it's a NumPy array.

Step 2: **Append the Data**
Scenario: A new record has come in, and you need to update your dataset to include this new information.

- You receive a new record to append to your dataset using only the given command.

Step 3: **Check if it's a Young City or Old City**
Scenario: Your next task is to analyze the age distribution to determine if the population is generally young or old.

- Create an Age array and examine the statistics shared in the .ipynb file.

Step 4: **Check the Country's Race Distribution**
Scenario: Understanding racial demographics is important for policy making. Your task is to analyze the race distribution.

- Create Race Arrays and make numerical identifications as shared in the .ipynb file.

Step 5: **Check if Senior Citizens Follow Work Hour Policy**
Scenario: The government has a policy that senior citizens (age > 60) should not work more than 25 hours a week. Your job is to check if this policy is being followed.

- Create a Senior Citizen Array and find out using the steps in the .ipynb file.

Step 6: Check Education and Income Relationship
Scenario: You need to determine if higher education correlates with higher income.

- Create relative Arrays and find the relations as shared in the .ipynb file.

Voila! That's it. You have done your work. Now take a cup of Chai and relax 🙂

The .ipynb file that I shared with you guys in the group contains more detailed descriptions and approaches for this Project. So, do give it a view first