Assignment

Pandas, Data Visualization & EDA

Adult Census Income Data Analysis

Learning Outcome:

- Pandas.
- Exploratory Data Analysis & Data Visualization.

Objective:

To do data analysis using python and explore the adult census income dataset using numpy and pandas libraries and drive meaningful insights by performing Exploratory data analysis using visualization libraries.

Data Description:

The dataset contains information about the individual's age, gender, and several attributes affecting the income of a person.

SL.No	Column Name	Description
1	Age	Age of an individual.
2	Workclass	Working class of an individual
3	Fnlwgt	Weights based on demographic characteristics.
4	Education	Highest education of an individual.
5	Education.num	Education encoded in unique numbers.
6	Marital Status	Marital status of an individual.
7	Occupation	Occupation of an individual.
8	Relationship	Relation of an individual in the family.
9	Race	Race of the individual.
10	Sex	Gender (0=Male, 1=Female)
11	Capital.gain	Gain amount of the individual
12	Capital.loss	Loss amount of the individual

13	Hours.per.week	Working hours per week
14	Native.country	The native country of the individual
15.	Income	Income of the Individual.

Questions:-

1. Import necessary libraries and load the dataset and display random 5 samples. Check the info of the data and write your findings. (2 points)
2. What is the average age of males and females in the data? (1 point)
3. Identify the qualifications of most of the individuals and display it using an appropriate plot? (1 point)
4. Find the middle most observation of age for different working classes. (1 point)
5. What is the percentage of the Asian-Pac-Islander race people in the data? (1 point)
6. Which occupation has more variability in the working hours? (1 point)
7. What is the range of the middle 50% of working hours per week? (2 points)
8. Are there any negatively skewed features? Support your answer with a metric and the plot. (2 points)
9. Identify the presence of extreme values in age using visualization techniques. (1 point)
10. Is there any effect of age on the working hours of a person? Support your answer with a metric and the plot. (2 points)
11. Is there any individual with age less than 18 and capital loss less than 100? (2 points)
12. Which occupation has the highest number of males? (1 point)
13. Analyze the trend of the capital gain amount for different educational qualifications and write your observations. (1 point)
14. Compare the working hours for each working class and write your observations. (1 point)
15. Is there an all male profession according to the data? (1 point)