

Questions

1. Display the top 10 rows of the dataset.
2. Check the last 10 rows of the dataset.
3. Find the shape of the dataset (number of rows and columns).
4. Retrieve information about the dataset, including data types and memory usage.
5. Fetch a random sample (50%) from the dataset.
6. Check for null values in the dataset.
7. Perform data cleaning by replacing '?' with NaN.
8. Drop all missing values.
9. Check for duplicate data and drop them.
10. Get overall statistics about the dataframe.
11. Drop the columns education-num, capital-gain, and capital-loss.
12. Analyze the distribution of the age column.
13. Find the total number of persons with age between 17 to 48 (inclusive).
14. Analyze the distribution of the workclass column.
15. Count the number of persons having Bachelor's and Master's degrees.
16. Perform bivariate analysis.
17. Replace salary values with 0 and 1.
18. Identify which workclass has the highest salary.
19. Determine whether males or females have a higher chance of earning more than 50K.
20. Convert the workclass column's data type to category.