

Practicals based on Data Science (Python) (230701105)

Practical ASSIGNMENT- 4

1. Consider following dataset in CSV file with the following fields.
Rollno, Name, Gender, English, Maths, Science, Category
Insert category values as A, B or C. At least insert 30 records with some missing values in subject marks data.
2. Load the dataset (student_marks.csv) into a pandas DataFrame.
Display information: number of columns, number of rows/records and the name of all columns.
3. Display the total number of missing values per column. Fill missing marks with the mean of the column/subject. Also handle missing values appropriately if exists in any column.
4. Create the following charts to demonstrate your finding using above dataset. Consider appropriate data for visualization.
 - A) Histogram
 - B) Bar Graph
 - C) Scatter Plot
 - D) Line chart
 - E) Area Graph
 - F) Pie chart
 - G) Donut chart
 - H) BoxPlot
5. Plot a histogram showing the distribution of students' marks in English.
6. Create a bar graph to compare the average marks in each subject (English, Maths, Science) for students in each Category (A, B, and C).
7. Plot a scatter plot to visualize the relationship between Maths and Science marks.
8. Plot a line chart to show the progress of English marks for the first 10 students based on their roll numbers.
9. Create a pie chart to show the distribution of students by Category (A, B, C).
10. Create a donut chart showing the distribution of students by Gender (Males vs. Females).
11. Generate a box plot to visualize the spread and outliers of Maths marks for students.