Mean: The average of a numerical set of values.

Median: The middle value when data is sorted.

Minimum and Maximum: The smallest and largest values respectively.

Standard Deviation (std): Measures how much the data deviates from the mean.

Percentile: The value below which a given percentage of observations fall.

Categorical Variable: A variable that takes on a limited number of distinct categories (e.g., Species, Loan_Status).

Grouping: Process of dividing data based on a categorical variable for individual statistical analysis.

🔹 Algorithm

For Summary Statistics Grouped by Categorical Variable:

1. **Load dataset using `pandas`.**
2. **Identify numerical and categorical columns.**
3. **Use `groupby()` on the categorical column.**
4. **Apply `agg()` to calculate statistics like mean, median, min, max, std.**

For General Statistical Summary:

1. **Load dataset using `pandas`.**
2. **Use `describe()` for quick summary.**
3. **Use `.mean()`, `.median()`, `.std()`, etc., for specific stats.**

4. **Use `numpy` or `seaborn` for more insights (e.g., correlation, plots).**

COnclusion

This program efficiently analyzes numerical data by grouping it with a categorical variable and extracting key statistical summaries. It also provides deep insights using plots and correlation matrices, useful for decision-making and data understanding.