Introduction to EDA:

- EDA aims to summarize main characteristics of data, often using visual methods.
- It helps to understand the data distribution, identify outliers, and assess the underlying assumptions.

Key Techniques:

- **Summary Statistics**: Mean, median, mode, standard deviation, range, percentiles.
- **Visualization**: Histograms, box plots, scatter plots, heatmaps, pair plots.
- **Correlation Analysis**: Pearson correlation, Spearman rank correlation.

Performing Exploratory Data Analysis (EDA) on the Titanic dataset is a common exercise in data analysis and machine learning. This dataset provides information about the passengers aboard the Titanic, including whether they survived or not, their age, class, fare, and other attributes. Here's a step-by-step guide and documentation outline for performing EDA on the Titanic dataset:

1. Loading the Dataset

First, load the dataset into your Python environment. You can download it from various sources such as Kaggle or use a pre-existing library that includes it (like Seaborn).

2. Understanding the Dataset

Explore the structure and contents of the dataset to understand what variables are available and their types.

3. Data Cleaning and Preprocessing

Handle missing values, convert data types if necessary, and preprocess the data for analysis.

4. Exploratory Data Analysis (EDA)

Perform visual and statistical analysis to explore relationships, distributions, and anomalies in the data.

5. Conclusion

Summarize findings and insights from EDA, which can guide further analysis or modeling tasks.