**Report on Data Cleaning and Exploratory Data Analysis (EDA)**

**Introduction**

This report describes the tasks performed for data cleaning and exploratory data analysis (EDA). The main goal of this process is to prepare and analyze the dataset to extract useful insights.

**Tasks Performed**

**1. Data Cleaning**

* **Loading the Dataset:** The dataset was loaded and its structure was checked.
* **Handling Missing Values:** Missing values were treated using imputation techniques or removed if necessary.
* **Removing Duplicates:** Duplicate records were identified and removed to avoid redundancy.
* **Detecting Outliers:** Statistical methods were used to find and treat outliers.
* **Standardizing Categorical Values:** Formatting issues and typos in categorical data were corrected.

**2. Exploratory Data Analysis (EDA)**

**Univariate Analysis (Single-Variable Exploration)**

* Summary statistics like mean, median, mode, variance, and skewness were calculated.
* Frequency distributions were analyzed for categorical variables.
* Histograms and box plots were created to visualize distributions.

**Bivariate Analysis (Two-Variable Exploration)**

* A correlation matrix was generated to find relationships between numerical variables.
* Scatter plots were used to check relationships between continuous variables.
* Bar plots, violin plots, and box plots were used to compare categorical and numerical variables.

**Multivariate Analysis (Multiple Variables Exploration)**

* Pair plots were created to analyze multiple relationships simultaneously.
* Heatmaps were used to visualize correlations among multiple variables.
* Grouped comparisons were performed to examine the combined effects of multiple features.

**Expected Outcome**

By completing these tasks, the following outcomes were achieved:

* The dataset was cleaned and prepared for analysis.
* The distributions of individual variables were visualized and understood.
* Relationships between variables were identified and interpreted.
* Multivariate techniques were applied to explore complex interactions within the data.

**Conclusion**

The data cleaning and EDA steps helped in preparing the dataset for further analysis. These methods provided valuable insights into the structure and relationships within the data, which will be useful for future decision-making and predictive modeling.