Project Report: Sales Data Analysis

Sales Data Analysis Project

1. Introduction

This report presents an analysis of sales data to understand patterns, trends, and key insights. The analysis includes data cleaning, manipulation, exploratory data analysis (EDA), and statistical modeling using techniques such as ANOVA and multiple regression.

2. Data Cleaning and Preprocessing

2.1 Handling Missing Values

- Checked for missing values in the dataset.
- Imputed missing values using appropriate methods (e.g., replacing null values with zero or using statistical imputation techniques like mean or median).

2.2 Dealing with Data Types

- Identified and corrected inconsistencies in variable data types.

2.3 Removing Duplicates

- Detected and removed duplicate records to ensure data integrity.

3. Data Manipulation

3.1 Merging and Concatenating Data

- Combined multiple datasets to ensure a comprehensive dataset for analysis.

4. Exploratory Data Analysis (EDA)

4.1 Frequency Analysis

- Conducted frequency analysis to gain insights into sales trends.

4.2 Detecting Normality

Several transformation techniques were applied to check data normality:

- **Square root method**
- **Log method**
- **Box-Cox method**
- **Yeo-Johnson method**

5. Statistical Analysis

5.1 One-Way ANOVA

- Performed ANOVA to determine whether there are statistically significant differences among groups.

5.2 Multiple Regression Analysis

- Developed a multiple regression model to identify key factors influencing sales performance.

6. Results and Insights

- The analysis revealed key trends in sales performance and factors affecting sales outcomes.
- Statistical analysis provided a basis for making data-driven business decisions.

This professional report outlines the structured approach to analyzing sales data, ensuring accurate insights and actionable recommendations.