

# Chapter 4: EDA/Initial Results

*by* Fuyuan Cai

---

**Submission date:** 24-Jul-2024 06:20PM (UTC-0700)

**Submission ID:** 2422035794

**File name:** corypaste.txt (1.14K)

**Word count:** 177

**Character count:** 990

#### 4.1 Descriptive statistics

By loading the dataset, it shows that the dataset has 6 columns and 164628 rows, and the data includes Sales, Volume, Satisfaction, Date, and Product, which belongs to a large database and is sufficient to support the data analysis in this paper. image After simple data processing, the mean, median, standard deviation, and distribution of some key variables (e.g., sales, volume, customer satisfaction) were first calculated. The run was very smooth, confirming that the dataset was of good quality. image image

#### 4.2 Visualization

Initial testing used histograms, bar charts, and box plots to visualize other key metrics such as sales trends, seasonal patterns, product popularity, and customer demographics. image image image image image This study entails depicting line graphs of sales over time, highlighting peak sales periods and seasonal trends i.e. sales trends. A bar chart showing the frequency of sales identifies the best selling products i.e. product popularity. Customer satisfaction can be seen by showing the distribution of satisfaction ratings for different store locations and purchasing methods in a box and line plot.

# Chapter 4: EDA/Initial Results

## ORIGINALITY REPORT

0%

SIMILARITY INDEX

0%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

## PRIMARY SOURCES

Exclude quotes      On

Exclude bibliography      On

Exclude matches      Off