

Retail Sales Analysis Report

Team Leader: Yassin Islam (Data Preprocessing & Report Manager)

Team Members:

- Mariam Mostafa (ML Engineer)
- Mariam Melad (EDA Analyst)
- Nada Etman (EDA Analyst)

1. Introduction

This report documents the analysis of synthetic retail sales data covering:

- Preprocessing of complete and missing datasets
- Exploratory visualizations of sales patterns
- KNN model performance metrics

2. Data Overview

Cleaned Dataset (First 3 Rows)


ProductCategory	UnitPrice	QuantitySold	Returned
Clothing	1603.67	4	0
Clothing	3698.71	4	0
Electronics	3088.44	9	0

Key Statistics:


- Mean UnitPrice: \$2,791
- Avg. Quantity Sold: 5.85 units

3. Visualizations

3.1 Data Quality


 Missing Values

22% missing UnitPrice values

 Price Distribution

Outlier removal impact

3.2 Product Analysis

 Return Rates

Electronics show highest return rate (12%)

4. Model Results

Dataset	Accuracy
Full Data	17.5%
Missing Data	12.5%

5. Conclusion

- Completed preprocessing of both datasets
 - Generated 3 key visualizations
 - Established baseline KNN accuracy metrics
-

Appendix: Code Files

Repository: github.com/k0rn3lDev/DS_Project