Documentation for Exploratory Data Analysis

Overview

This project performs a comprehensive exploratory data analysis (EDA) to understand customer behavior, sales trends, and product performance. The analysis is based on three datasets: `Customers`, `Products`, and `Transactions`.

Exploratory Data Analysis (EDA)

Univariate Analysis

- **Numerical Features**: Distributions of `Price`, `Quantity`, and `TotalValue` are analyzed using histograms and box plots.
- **Categorical Features**: Frequency counts for `Region` and `Category` provide insights into data distribution.

Bivariate and Multivariate Analysis

- **Correlation Analysis**:
    - Heatmaps identify relationships between numerical variables.
    - Key Insights: Positive correlation between `Price` and `TotalValue`.
- **Scatter Plots**:
    - Analyzed `Price` vs. `Quantity` to understand purchasing patterns.

Temporal Analysis

- Daily total sales trends are visualized to identify peak sales periods and seasonal patterns.
- Insights:
    - Sales exhibit periodic spikes, potentially aligned with promotions or holidays.

Key Business Insights

1. **High-Value Customers**:
    - Customers with frequent and high-value transactions were identified, providing opportunities for loyalty programs.
2. **Seasonal Trends**:
    - Peak sales periods highlight the effectiveness of promotions and holidays.
3. **Product Preferences**:
    - Insights into regional and category preferences guide inventory planning.

Conclusion

The EDA provides a structured framework for understanding customer behavior, sales trends, and product performance. These insights can guide strategic decision-making in marketing, inventory management, and customer relationship management.