Retail Sales Data Analysis

📌 Executive Summary

This project analyzes a retail sales dataset to identify key business insights. Using Python for preprocessing and analysis, and BI tools (Power BI / Tableau / Looker Studio) for visualization, the project demonstrates how data-driven decision-making can improve sales strategies and operations.

🎯 Problem Statement

Retail businesses often struggle to understand which products perform best, how sales differ across regions, and how seasonality affects revenue. Without clear insights, it is difficult to make informed decisions to increase profits and optimize resources.

📊 Dataset Source

We used the Superstore Sales Dataset from Kaggle:

[Superstore Dataset on Kaggle](#)

🛠️ Methodology

1. Dataset Preparation

   ○ Download dataset from Kaggle.

   ○ Preprocess using Python (handle missing values, duplicates, format issues).

2. Data Analysis

   ○ Perform descriptive statistics (sales trends, top products, correlations).

3. Visualization with BI Tool

- Build an interactive dashboard with 5+ visualizations (bar chart, pie chart, time-series, heatmap, etc.).

- Add filters/slicers (region, product, month).

4. Insights & Recommendations

- Identify top-selling products.

- Compare sales across regions.

- Highlight seasonal trends.

📈 Results

- Top Products: Certain categories (e.g., office supplies) generate the most revenue.

- Regional Sales: Some regions outperform others in total sales.

- Seasonality: Sales tend to peak during specific months.

🖥️ BI Dashboard

The BI dashboard provides interactive visualizations for easy exploration of sales data.

Screenshots

🙌 Acknowledgment

This project was completed as part of BIT2053 – Fundamental of Modern Data at City University Malaysia.