Retail Sales Intelligence Toolkit

Overview

This project integrates a Power BI dashboard with a machine learning prediction model to analyze and forecast retail sales. The Power BI component provides visual insights into regional performance, discount strategies, and profitability, while the Python-based model uses Random Forest to predict future sales outcomes from transactional data.

Power BI Dashboard

- Cleaned the Superstore dataset: removed blanks, converted dates
- Built line, bar, and map visuals with slicers for interactivity
- KPIs included total sales, profit by region, and discount averages
- Slicers allow filtering by region and year to support business decisions

Machine Learning Model

- Used a Random Forest Regressor with one-hot encoded inputs
- Dataset split into 80% train, 10% validation, and 10% test
- Focused on prediction of sales using features like region, discount, quantity, segment, and category
- Achieved R^2 of 0.51 on the test set with clear signs of overfitting (train R^2 was 0.98)
- Included evaluation visual: Actual vs Predicted Sales

Included Downloads

- Power BI Dashboard (.pbix)
- Python script with Random Forest model
- Feature importance plot