Sales Forecasting & Demand Prediction

Overview

With the increasing demands and intense competition in the market, businesses need to understand which products, regions, categories, and customer segments to target for maximizing sales and profitability. This project utilizes machine learning techniques to forecast sales and predict demand based on historical data.

Dataset

The dataset used in this project comes from a Superstore and contains detailed transactional data. The dataset includes various attributes such as order details, customer information, product details, and financial metrics.

Metadata

The dataset consists of the following features:

- Row ID: Unique ID for each row.
- Order ID:Unique Order ID for each customer.
- Order Date: Date when the order was placed.
- Ship Date: Date when the order was shipped.
- Ship Mode: Shipping mode selected by the customer.
- Customer ID: Unique ID assigned to each customer.
- Customer Name: Name of the customer.
- Segment: The segment to which the customer belongs.
- Country: Country of residence of the customer.
- City: City of residence of the customer.
- State: State of residence of the customer.
- Postal Code: Postal code of the customer.
- Region: Region where the customer belongs.
- Product ID: Unique ID of the product.

- Category: Category of the ordered product.

- Sub-Category: Sub-category of the ordered product.

- Product Name: Name of the product.

- Sales: Sales revenue from the product.

- Quantity: Quantity of the product sold.

- Discount: Discount applied to the product.

- Profit: Profit or loss incurred from the sale.

Project Objectives

The Sales Forecasting and Demand Prediction project aims to build a machine learning

model that predicts future sales and demand for products based on historical data. Accurate forecasting

helps businesses optimize inventory management, staffing, and marketing strategies. This project will apply

data science techniques, from data collection and analysis to model deployment and monitoring, enabling

businesses to make data-driven decisions.

Technologies Used

- Python

- Pandas, NumPy, Matplotlib, Seaborn (Data Analysis & Visualization)

- Scikit-learn (Machine Learning Models)