Product Demand Prediction

# INTRODUCTION

The problem is to implementing data science techniques for segment product demand prediction is a data-driven process that involves forecasting the future demand for a specific product or a range of products. The primary objective is to estimate the quantity of goods or services that customers are likely to purchase within a defined time frame. Demand prediction serves as a fundamental component of supply chain management, inventory control, production planning, and sales optimization for businesses. It allows organizations to make informed decisions related to inventory levels, procurement, pricing, and marketing strategies.

DATASET

The data is obtained from [**https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning**](https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ID | Store ID | Total Price | Base Price | Units Sold |
| 1 | 8091 | 99.0375 | 111.8625 | 20 |
| 2 | 8091 | 99.0375 | 99.0375 | 28 |
| 3 | 8091 | 133.95 | 133.95 | 19 |
| 4 | 8091 | 133.95 | 133.95 | 44 |
| 5 | 8091 | 141.075 | 141.075 | 52 |
| 9 | 8091 | 227.2875 | 227.2875 | 18 |
| 10 | 8091 | 327.0375 | 327.0375 | 47 |
| 13 | 8091 | 210.9 | 210.9 | 50 |
| 14 | 8091 | 190.2375 | 234.4125 | 82 |
| 17 | 8095 | 99.0375 | 99.0375 | 99 |
| 18 | 8095 | 97.6125 | 97.6125 | 120 |
| 19 | 8095 | 98.325 | 98.325 | 40 |
| 22 | 8095 | 133.2375 | 133.2375 | 68 |
| 23 | 8095 | 133.95 | 133.95 | 87 |
| 24 | 8095 | 139.65 | 139.65 | 186 |
| 27 | 8095 | 236.55 | 280.0125 | 54 |
| 28 | 8095 | 214.4625 | 214.4625 | 74 |
| 29 | 8095 | 266.475 | 296.4 | 102 |
| 30 | 8095 | 173.85 | 192.375 | 214 |

# COLUMNS USED

From Product demand .csv data the following columns are used

ID

Store ID

Total price

Base price

Unit sold

# LIBRARIES USED

The Python 3 environment comes with many helpful analytics libraries installed and several helpful packages to load.

The essential libraries used in this project are :

* Importing OS (for kaggle inputs)
* Numpy and Pandas libraries
* Matplotlib
* Seaborn

# TRAIN AND TEST

Training the dataset by describe(), isnull().sum(), drop(), show(), and by using Linear Regression algorithm we train the data

Testing the data by importing sklearn.linear\_modal from Linear Regression with ensuring the plot range and axis labels producing the values, scattering the data by mean\_absolute\_error and producing 3D plot.

# REST OF THE EXPLANATIONS

## Data Collection

The process involves gathering products data, which includes information about their purchase history, demographics, and interaction patterns.

## Data Preprocessing

The task involves preparing and cleaning data, handling missing values, and converting categorical features into numerical representations.

## Feature Engineering

Data preparation and cleaning, handling missing values, and the transformation of categorical features into numerical representations are all part of the task.

Modal Evaluation

Evaluate the model's performance on the test set using appropriate evaluation metrics. Common metrics for demand prediction include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

# ALGORITHMS USED

Apply clustering algorithms like K-Means, DBSCAN, or hierarchical clustering to segment customers.

Visualization: Visualize the customer segments using techniques like scatter plots, bar charts, and heatmaps. Interpretation: Analyze and interpret the characteristics of each customer segment to derive actionable insights for marketing strategies.

# DESIGN AND DATAFLOW

1.Physical data flow diagram:



