PRODUCT DEMAND PREDICTION WITH MACHINE LEARNING

DATA COLLECTION

- *Gather historical sales data , which typically includes information such as date, product ID, quantity sold, price, and other relevant features.*
- *You may also need external factors like economic indicators, seasonality, and marketing efforts data.*

DATA PREPROCESSING

- ***Loading Data****:*

Load your historical sales data and external factors data into data structures like Pandas DataFrames in Python.

- ***Data Integration****:*

Merge or join the historical sales data and external factors data, aligning them by the date or another common key.

- ***Handling Missing Values****:*

Check for missing values in both datasets. You can impute missing values using methods like mean, median, or interpolation, or drop rows with missing data if appropriate.

DATASET

- **Dataset Link:** https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning

PREPROCESSED DATASET

CODING:

```
import pandas as pd
import numpy as np
import plotly.express as px
import seaborn as sns
I mport matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeRegressor
data = pd.read_csv(“productdemand.csv")
data.head()
```

OUTPUT

ID	STORE ID	TOTAL PRICE	BASE PRICE	UNIT SOLD
1	8091	99.0375	111.8625	20
2	8091	99.0375	99.0375	28
3	8091	133.9500	133.95	19
4	8091	133.9500	133.95	44
5	8091	141.0750	141.075	52

ANALYZE THE PRICE AND DEMAND PRODUCT

```
fig = px.scatter(data,x="Units Sold",y="Total Price",size='Units Sold')
fig.show()
```

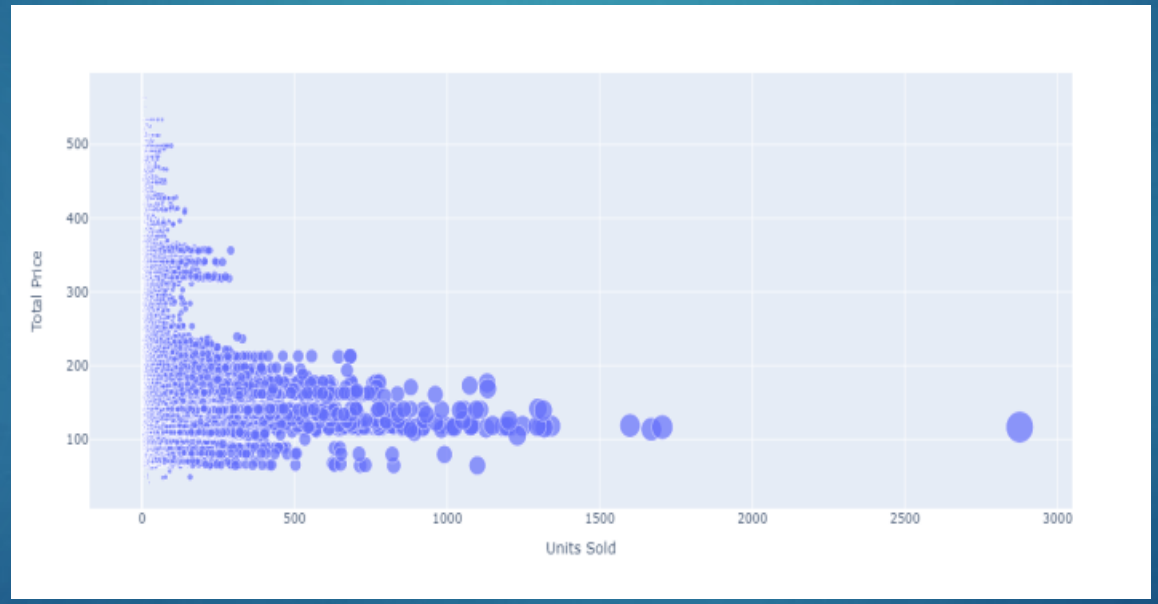

Fig: Scatter Plot

EXPLORATORY DATA ANALYSIS

- *Analyze the data to identify pattern, trends and correlations.*
- *EDA can help you make informed decision about feature selection and model choice.*

CODING:

```
missing_values = data.isnull().sum()
# Summary statistics
summary_stats = data.describe()
# Data visualization (e.g., histograms, scatter plots)
import matplotlib.pyplot as plt
import seaborn as sns
# perform EDA here
```

ANALYZE THE CORRELATION

- **CODING:**

```
print(data.corr())
correlations = data.corr(method='pearson’)
plt.figure(figsize=(15, 12))
sns.heatmap(correlations, cmap="coolwarm", annot=True)
plt.show()
```

OUTPUT OF THE CORRELATION

	ID	STORE ID	TOTAL PRICE	BASE PRICE	UNITS SOLD
ID	1.000000	0.007464	0.008473	0.018932	-0.010616
STORE ID	0.007464	1.000000	-0.038315	-0.038848	-0.007432
TOTAL PRICE	0.008473	-0.038315	1.000000	0.958885	-0.235625
BASE PRICE	0.018932	-0.038848	0.958885	1.000000	-0.140032
UNITS SOLD	-0.010616	-0.004372	-0.235625	-0.140032	1.000000

```
Correlations=data.corr(method='pearson')
plt.figure(figsize=(15,12))
sns.heatmap(correlations,cmap="coolwarm",annot=true)
plt.show()
```

Fig: Heatmap

CONCLUSION

- *Product demand prediction with machine learning were classified and various processing were done using the given dataset.*