

PHASE 3

PRODUCT DEMAND PREDICTION WITH MACHINE LEARNING

Dataset and its detail Explanation :

Product ID :

This is a unique identifier assigned to each product in a store or inventory system. It ensures that every product can be specifically identified and differentiated from all other products.

Store ID :

Just as products have a unique identifier, so do stores. This is especially relevant for businesses with multiple branches or franchises.

Total Price at Which Product Was Sold :

This is the final selling price of the product to the customer. It may include any discounts, promotions, or additional charges applied.

Base Price at Which Product Was Sold :

This is the initial or listed price of the product before any discounts, promotions, or additional charges. It's the standard price at which the product is typically sold when no promotions are in effect.

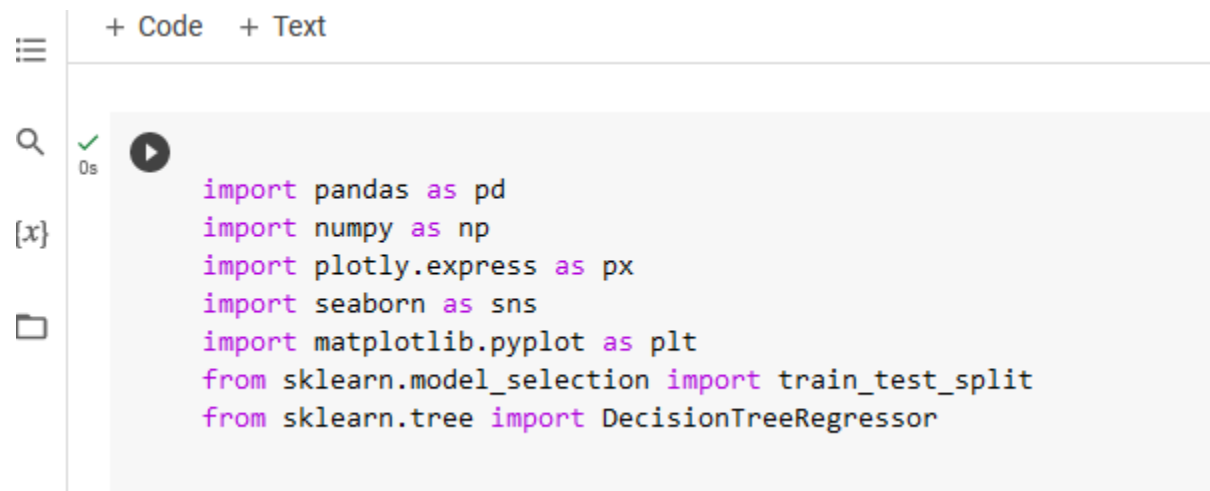
PHASE 3

Units Sold (Quantity Demanded) :

This refers to the number of individual units of a product that were sold during a specific period.

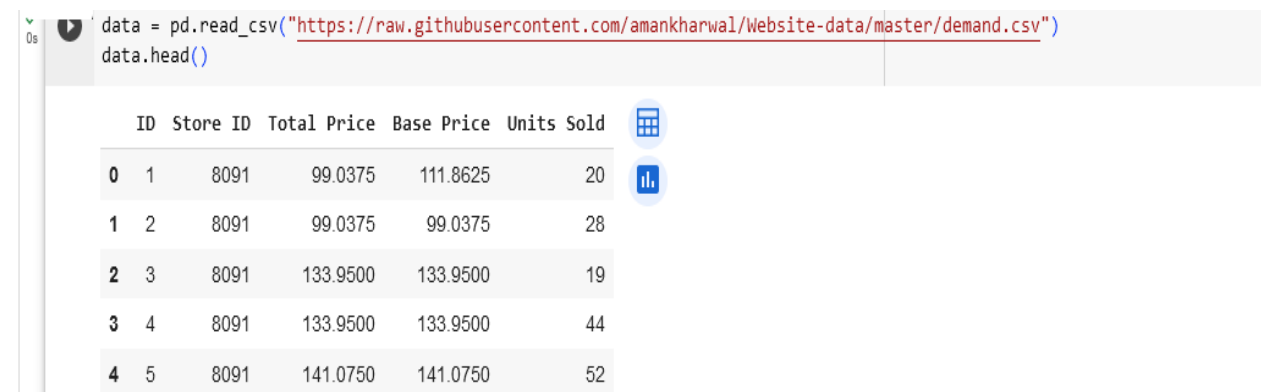
Begin building the project by load the dataset :

Importing the packages:



A Jupyter Notebook interface with a sidebar on the left containing icons for a menu, search, a variable [x], and a file folder. The main area has a toolbar with '+ Code' and '+ Text' buttons. Below the toolbar is a code cell with a play button icon and a '0s' execution time indicator. The code cell contains the following Python code:

```
import pandas as pd
import numpy as np
import plotly.express as px
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeRegressor
```



A Jupyter Notebook interface showing the execution of code to load a dataset. The code cell contains:

```
data = pd.read_csv("https://raw.githubusercontent.com/amankharwal/Website-data/master/demand.csv")
data.head()
```

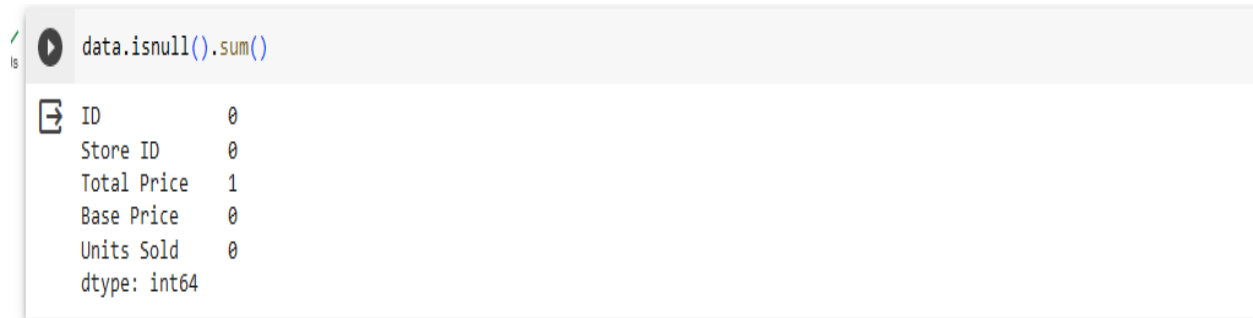
Below the code cell, the first five rows of the dataset are displayed as a table. To the right of the table are icons for a grid and a bar chart.

	ID	Store ID	Total Price	Base Price	Units Sold
0	1	8091	99.0375	111.8625	20
1	2	8091	99.0375	99.0375	28
2	3	8091	133.9500	133.9500	19
3	4	8091	133.9500	133.9500	44
4	5	8091	141.0750	141.0750	52

PHASE 3

Preprocess Dataset :

Now lets have a look at whether this dataset contains any null values or not.



```
data.isnull().sum()
```

ID	0
Store ID	0
Total Price	1
Base Price	0
Units Sold	0
dtype: int64	