BLACK FRIDAY SALES PREDICTION

DONE BY

Almasdivan.K(231008)

Ananda Dharshinee.M.S(231009)

Dheepika.R(231025)

Harini.K(231036)

Jayajanani.S(231041)

Keerthana.S(231050)

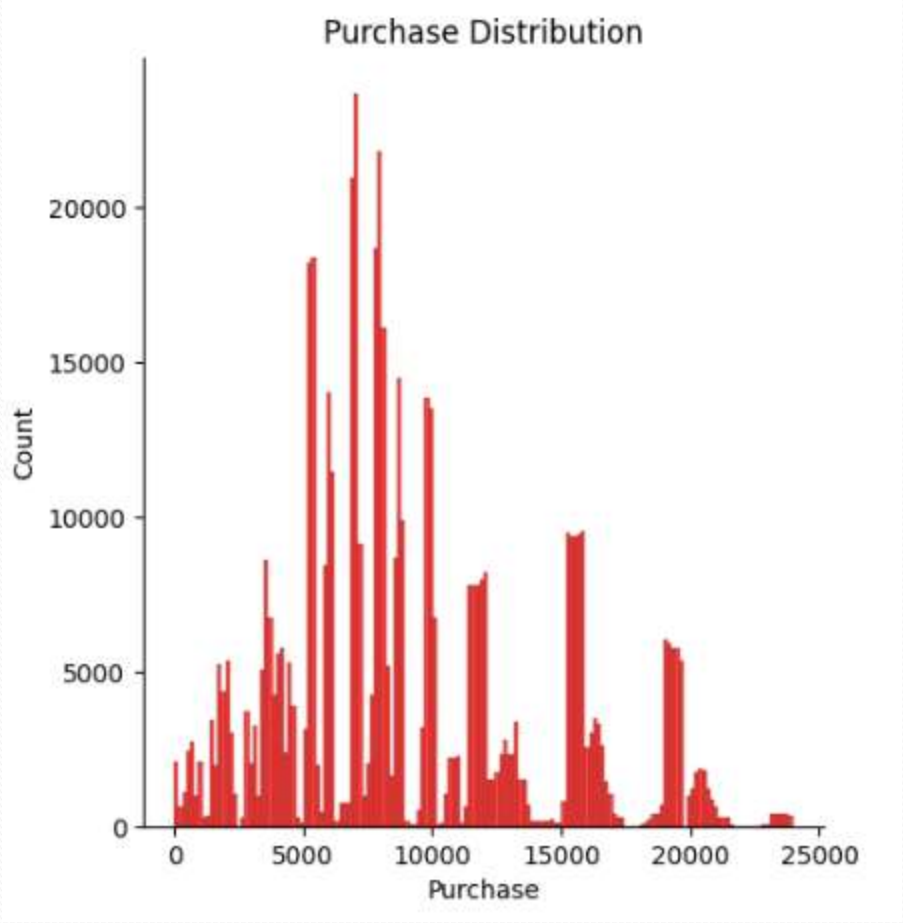
What is Black Friday?

- Black Friday is the biggest retail sales day in the United States, often regarded as the unofficial start of the holiday shopping season. This season is crucial for the economy, especially for retailers. It traditionally marks the start of the Christmas shopping season in the United States.

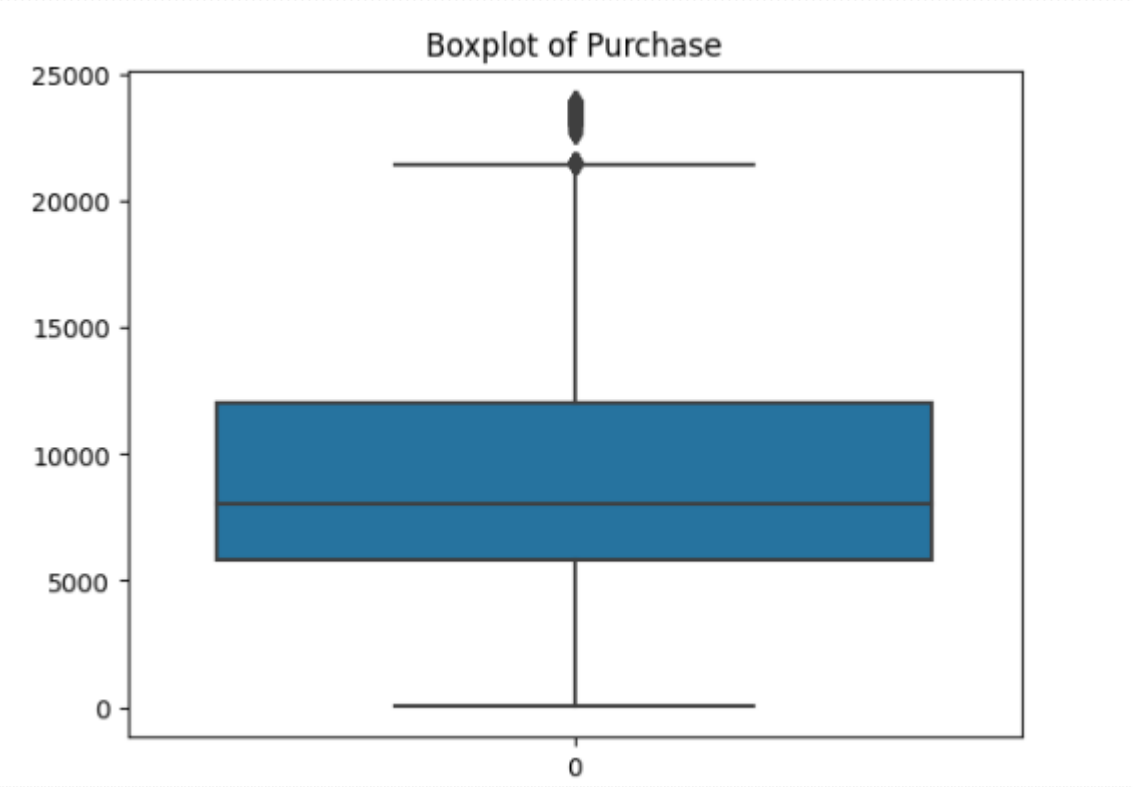
PROBLEM STATEMENT:

- A retail company wants to understand the customer purchase behaviour against various products of different categories. They have shared purchase summary of various customers for selected high volume products from last month. The data set also contains customer demographics product details and Total purchase amount from last month.
- Now, they want to build a model to predict the purchase amount of customer against various products which will help them to create personalized offer for customers against different products.

VISUALISATION



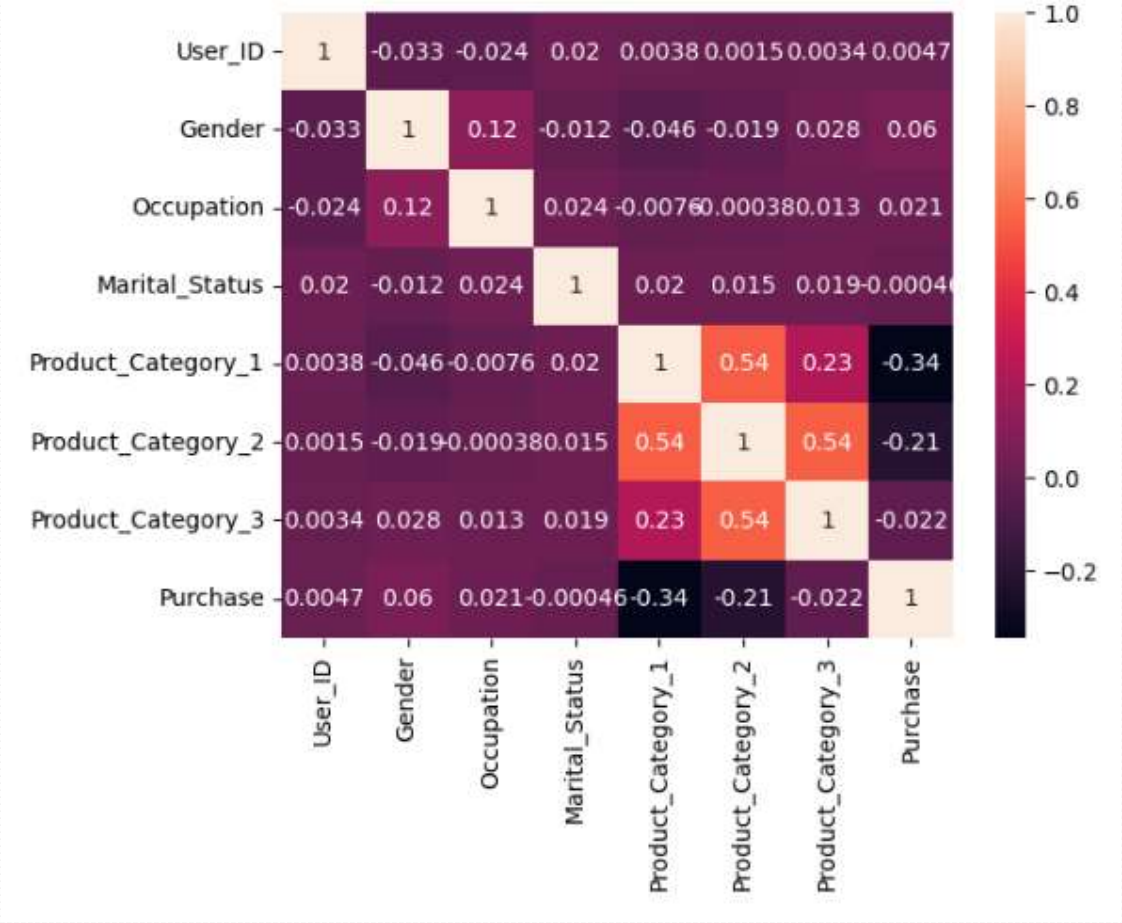
OUTLIERS



Transforming

	User_ID	Product_ID	Gender	Age	Occupation	City_Category	Stay_In_Current_City_Years	Marital_Status	Product_Category_1	Product_Category_2	Product_Category_3	Purchase
0	1000001	P00069042	0	0-17	10	A	2	0	3	NaN	NaN	8370
1	1000001	P00248942	0	0-17	10	A	2	0	1	6.0	14.0	15200
2	1000001	P00087842	0	0-17	10	A	2	0	12	NaN	NaN	1422
3	1000001	P00085442	0	0-17	10	A	2	0	12	14.0	NaN	1057
4	1000002	P00285442	1	55+	16	C	4+	0	8	NaN	NaN	7969

HEAT MAP


User_ID
Gender
Occupation
Marital_Status
Product_Category_1
Product_Category_2
Product_Category_3
Purchase
1 -0.033 -0.024 0.02 0.0038 0.0015 0.0034 0.0047
-0.033 1 0.12 -0.012 -0.046 -0.019 0.028 0.06
-0.024 0.12 1 0.024 -0.0076 0.00038 0.013 0.021
0.02 -0.012 0.024 1 0.02 0.015 0.019 -0.0004
0.0038 -0.046 -0.0076 0.02 1 0.54 0.23 -0.34
0.0015 -0.019 0.00038 0.015 0.54 1 0.54 -0.21
0.0034 0.028 0.013 0.019 0.23 0.54 1 -0.022
0.0047 0.06 0.021 -0.00046 -0.34 -0.21 -0.022 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

Predicting NULL Values

```
df.isnull().sum()
```

```
User_ID                       0
Product_ID                    0
Gender                        0
Age                           0
Occupation                    0
City_Category                 0
Marital_Status                0
Product_Category_1            0
Product_Category_2            0
Product_Category_3            0
Purchase                      0
Stay_In_Current_City_Years_0  0
Stay_In_Current_City_Years_1  0
Stay_In_Current_City_Years_2  0
Stay_In_Current_City_Years_3  0
Stay_In_Current_City_Years_4+ 0
dtype: int64
```

Spliting the Data Set

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=123)
```

Linear Regression

- Linear Regression is an algorithm that belongs to supervised Machine Learning. It tries to apply relations that will predict the outcome of an event based on the independent variable data points. The relation is usually a straight line that best fits the different data points as close as possible.

```
r2_score(y_test, y_pred)
```

```
0.15192944521481688
```

DECISION TREE REGRESSOR

- Decision tree builds regression or classification models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes.

```
r2_score(y_test, dt_y_pred)
```

```
0.5521191505924365
```

XGBOOST REGRESSOR

- XGBoost, which stands for Extreme Gradient Boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems.

```
r2_score(y_test, xgb_y_pred)
```

```
0.67227891659979
```

LINK

GITHUB LINK:

1. https://github.com/Almasdivan/_Black_Friday_Sale_Prediction_Analysis-ML
2. https://github.com/ANANDADHARSHINEE/Black-Friday-Sales-Prediction
3. https://github.com/dheepika-r/Inlustro
4. https://github.com/Harinivaradhan
5. https://github.com/Jayajanani27/inlustro-blackfriday
6. https://github.com/KeerthanaS01/Inlustro

THANK YOU