

Project Title	Black Friday Sales Prediction
Technologies	Business Intelligence
Domain	Retail and Sales
Project Difficulties level	Advanced

Problem Statement:

Retail is **the sale of goods and services from individuals or businesses** to the end-user. The retail industry provides consumers with goods and services for their everyday needs. In retail one of crucial part is to understand the consumer behaviour and make various arrangements for the sales of the company.

A retail company “ABC Private Limited” wants to understand the customer purchase behaviour (specifically, purchase amount) 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 (age, gender, marital status, city_type, stay_in_current_city), product details (product_id and product category) 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.

Data

Variable	Definition
User_ID	User ID
Product_ID	Product ID
Gender	Sex of User
Age	Age in bins

Occupation

Occupation (Masked)

City_Category

Category of the City (A,B,C)

Stay_In_Current_City_Years

Number of years stay in current city

Marital_Status

Marital Status

Product_Category_1	Product Category (Masked)
Product_Category_2	Product may belongs to other category also (Masked)
Product_Category_3	Product may belongs to other category also (Masked)
Purchase	Purchase Amount (Target Variable)

Your model performance will be evaluated on the basis of your prediction of the purchase amount for the test data (test.csv), which contains similar data-points as train except for their purchase amount. Your submission needs to be in the format as shown in "SampleSubmission.csv".

We at our end, have the actual purchase amount for the test dataset, against which your predictions will be evaluated. Submissions are scored on the root mean squared error (RMSE). RMSE is very common and is a suitable general-purpose error metric. Compared to the Mean Absolute Error, RMSE punishes large errors:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2},$$

Where \hat{y} is the predicted value and y is the original value.

Find key metrics and factors and show the meaningful relationships between attributes. Do your own research and come up with your findings.

Dataset:

You can find the dataset on the given link

<https://drive.google.com/drive/folders/1IOQh5zJ4QDLfU55-gOzI39kCZx3ikJB5?usp=sharing>

Approaches:

Python, R, Tableau, Power BI or you can use any tools and techniques as per your convenience. We would appreciate your valid imagination in finding solutions

Project Evaluation metrics:

Code: As per the requirements

- You are supposed to write a code in a modular fashion
- Safe: It can be used without causing harm.
- Testable: It can be tested at the code level.
- Maintainable: It can be maintained, even as your codebase grows.
- Portable: It works the same in every environment (operating system)
- You have to maintain your code on GitHub.

- You have to keep your GitHub repo public so that anyone can check your code.
- Proper readme file you have to maintain for any project development.
- You should include basic workflow and execution of the entire project in the readme file on GitHub
- Follow the coding standards: <https://www.python.org/dev/peps/pep-0008/>

Database:

- You are supposed to use a given dataset/resource for this project.

<https://drive.google.com/drive/folders/1IOQh5zJ4QDLfU55-gOzl39kCZx3ikJB5?usp=sharing>

Submission requirements:

High-level Document:

You have to create a high-level document design for your project. You can reference the HLD form below the link.

Demo link:

[HLD Document Link](#)

Low-level document:

You have to create a Low-level document design for your project; you can refer to the LLD from the below link.

Demo link:

[Low Level Design Sample document link](#)

Architecture:

You have to create an Architecture document design for your project; you can refer to the Architecture from the below link.

Demo Link:

[Architecture Document Link](#)

Wireframe:

You have to create a Wireframe document design for your project; refer to the Wireframe from the below link.

Demo link

[Wire-frame link](#)

Project work:

You will have to share the Tableau Public Link of your work

You have to submit your code GitHub repo in your dashboard when the final submission of your project.

Demo link

[Project code sample link :](#)

Detail project report:

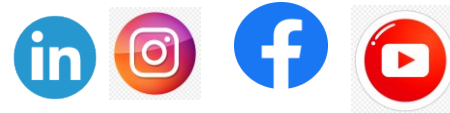
You have to create a detailed project report and submit that document as per the given sample.

Demo link

[DPR sample link](#)

Project demo video:

You have to record a project demo video for at least 5 Minutes and submit that link as per the given demo.



The project LinkedIn a post:

You have to post your project detail on LinkedIn and submit that post link in your dashboard in your respective field.

Demo link

[Linkedin post sample link :](#)

