

Business Intelligence Project Description 2024

IS Department

Submission:-

Please submit your code and report (including screenshot of code and the relevant figures) **note that** this project report should be in a presentation format (i.e., around 20 slides/PPT/PPTX) and code should be in a python file format (.py or .ipynb)

- Presentation should include Agenda, Introduction, Conclusion, etc.
- Each group member should spend equal time and effort.
- Project delivery date & discussion will be on the practical exam date.

Walmart Sales Problem

Dataset Description:-

The [Walmart dataset](#) consists of **21 columns with 9,995** number of **rows**.

Row ID => Unique ID for each row.

Order ID => Unique Order ID for each Customer.

Order Date => Order Date of the product.

Ship Date => Shipping Date of the Product.

Ship Mode=> Shipping Mode specified by the Customer.

Customer ID => Unique ID to identify each Customer.

Customer Name => Name of the Customer.

Segment => The segment where the Customer belongs.

Country => Country of residence of the Customer.

City => City of residence of the Customer.

State => State of residence of the Customer.

Postal Code => Postal Code of every Customer.

Region => Region where the Customer belongs.

Product ID => Unique ID of the Product.

Category => Category of the product ordered.
Sub-Category => Sub-Category of the product ordered.
Product Name => Name of the Product
Sales => Sales of the Product.
Quantity => Quantity of the Product.
Discount => Discount provided.
Profit => Profit/Loss incurred.

Project Description:-

Use the methods and techniques studied in lectures and labs in order to help the Walmart owners to understand which products, regions, categories and customer segments they should target or avoid.

- The questions of the owners of the Walmart:
 - Which product category has the highest sales?
 - Which customer segment that contribute to the highest sales?
 - Which region, state and city contribute to the highest sales and profit?
 - How the factors have influenced on each other?

Project implementation will include the following:-

1. Understand the topic of the dataset and define your project objective.
2. Apply data exploration methods (i.e., summary statistics and visualization methods)
3. Apply data cleaning or transformation method (i.e., Handle missing values, remove duplicates, handle outliers, etc.)
4. Visualize your dataset using **PowerBI**.
5. **Develop** and **evaluate** a **machine learning model** in **predicting sales** of the Walmart.
6. Making a **time series** for the **sales** of Walmart. It is between column sales and Order Date.