Project Title: *E-Commerce Customer Churn Analysis*

Problem Statement:

Inthe realmof e-commerce, businesses face the challenge of understanding customer churn patterns to ensure customer satisfaction and sustained profitability. This project aims to delve into the dynamics of customer churn within an e-commerce domain, utilizing historical transactional data to uncover underlying patterns and drivers of churn. By analyzing customer attributes such as tenure, preferred payment modes, satisfaction scores, and purchase behavior, the project seeks to investigate and understand the dynamics of customer attrition and their propensity to churn. The ultimate objective is to equip e-commerce enterprises with actionable insights to implement targeted retention strategies and mitigate churn, thereby fostering long-term customer relationships and ensuring business viability in a competitive landscape.

Dataset Download:

https://drive.google.com/uc?export=download&id=1iKKCze_Fpk2n_g3BIZBiSjcDFdFcEn3D

Project Stepsand Objectives:

Data Cleaning:

Handling Missing Values and Outliers:

➢ Impute mean for the following columns, and round off to the nearest integer if required: WarehouseToHome, HourSpendOnApp, OrderAmountHikeFromlastYear, DaySinceLastOrder.

➢ Impute mode for the following columns: Tenure, CouponUsed, OrderCount.

➢ Handle outliers in the 'WarehouseToHome' column by deleting rows where the values are greater than 100.

Dealing with Inconsistencies:
- ➢ Replace occurrences of "Phone" in the 'PreferredLoginDevice' column and "Mobile" in the 'PreferedOrderCat' column with "Mobile Phone" to ensure uniformity.

- ➢ Standardize payment mode values: Replace "COD" with "Cash on Delivery" and "CC" with "Credit Card" in the PreferredPaymentMode column.

Data Transformation:

Column Renaming:
- ➢ Rename the column "PreferedOrderCat" to "PreferredOrderCat".
- ➢ Rename the column "HourSpendOnApp" to "HoursSpentOnApp".

Creating New Columns:
- ➢ Create a new column named 'ComplaintReceived' with values "Yes" if the corresponding value in the 'Complain' is 1, and "No" otherwise.

- ➢ Create a new column named 'ChurnStatus'. Set its value to "Churned" if the corresponding value in the 'Churn' column is 1, else assign "Active".

Column Dropping:
- ➢ Drop the columns "Churn" and "Complain" from the table.

Data Exploration and Analysis:

- ➢ Retrieve the count of churned and active customers from the dataset.

- ➢ Display the average tenure and total cashback amount of customers who churned.

- ➢ Determine the percentage of churned customers who complained.

- ➢ Identify the city tier with the highest number of churned customers whose preferred order category is Laptop & Accessory.

- ➤ Identify the most preferred payment mode among active customers. Calculate the
- ➤ total order amount hike from last year for customers who are single
  and prefer mobile phones for ordering.
- ➤ Find the average number of devices registered among customers who used UPI as
  their preferred payment mode.
  Determine the city tier with the highest number of customers.
- ➤
  Identify the gender that utilized the highest number of coupons.
- ➤
  List the number of customers and the maximum hours spent on the app in each
- ➤ preferred order category.
  Calculate the total order count for customers who prefer using credit cards and
  have the maximum satisfaction score.
- ➤
  What is the average satisfaction score of customers who have complained?

- ➤ List the preferred order category among customers who used more than 5
  coupons.
- ➤
  List the top 3 preferred order categories with the highest average cashback
  amount.
- ➤ Find the preferred payment modes of customers whose average tenure is 10
  months and have placed more than 500 orders.
  Categorize customers based on their distance from the warehouse to home such
- ➤ as 'Very Close Distance' for distances <=5km, 'Close Distance' for <=10km,
  'Moderate Distance' for <=15km, and 'Far Distance' for >15km. Then, display the
- ➤ churn status breakdown for each distance category.
  List the customer's order details who are married, live in City Tier-1, and their
  order counts are more than the average number of orders placed by all
  customers.

➢ a) Create a '**customer_returns**' table in the 'ecomm' database and insert the following data:

ReturnID	CustomerID	ReturnDate	RefundAmount
١٠٠١	٥٠٠٢٢	٢٠٢٣-٠١-٠١	٢١٣٠
١٠٠٢	٥٠٣١٦	٢٠٢٣-٠١-٢٣	٢٠٠٠
١٠٠٣	٥١٠٩٩	٢٠٢٣-٠٢-١٤	٢٢٩٠
١٠٠٤	٥٢٣٢١	٢٠٢٣-٠٣-٠٨	٢٥١٠
١٠٠٥	٥٢٩٢٨	٢٠٢٣-٠٣-٢٠	٣٠٠٠
١٠٠٦	٥٣٧٤٩	٢٠٢٣-٠٤-١٧	١٧٤٠
١٠٠٧	٥٤٢٠٦	٢٠٢٣-٠٤-٢١	٣٢٥٠
١٠٠٨	٥٤٨٣٨	٢٠٢٣-٠٤-٣٠	١٩٩٠

b) Display the return details along with the customer details of those who have churned and have made complaints.