**Problem Statement**

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**

|  |  |
| --- | --- |
| Column | Description |
| User\_ID | User ID |
| Product\_ID | Product ID |
| Gender | Gender |
| Age | Age in bins |
| Occupation | Occupation (masked) |
| City\_Category | Category of a 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 |
| Product\_Category\_2 | Product may belong to another category also |
| Product\_Category\_3 | Product may belong to another category also |
| **Purchase** | **Purchase amount (target variable)** |