**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
* Variable Definition
* User\_ ID User ID
* Product\_ ID Product ID
* Gender Sex of User
* Age in bins
* Occupation (Masked)
* City\_ 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 Amount (Target Variable)