**Column definitions (df\_cleaned)**

1. **Customer Demographic**

* **ID**: Unique identifier for each customer.
* **Year\_Birth**: Year of birth of the customer.
* **Age**: Calculated age of the customer based on the year of the latest customer enrollment and Year\_Birth
* **Age\_Group**: Categorical grouping of customers based on their age.
  + **Values**: Gen X (892), Millennials (716), Boomers+ (567), Gen Z- (61)
  + Customers were grouped into segments:  
    - Gen Z (0-24): Young adults, likely early-stage earners.  
    - Millennials (25-39): Early to mid-career, potentially family-oriented.  
    - Gen X (40-54): Established professionals.  
    - Boomers+ (55+): Retirees or late-career individuals.
* **Education**: Level of education attained by the customer (e.g., Bachelor, Master).
  + **Values**: Bachelor (1126), Master (571), Doctoral (485), High School (54)
* **Marital\_Status**: Marital status of the customer (e.g., Married, Single, Divorced).
  + **Values**: Married (1442), Single (482), Divorced (308), Other (4)
* **Income**: Annual income of the customer.
* **Income\_Group**: Categorical grouping of customers based on their income levels (e.g., Low, Medium, High).
  + **Values**: Medium (1290), Low (731), High (215)

1. **Household Composition**

* **Kidhome**: Number of small children in the customer’s household.
  + **Values**: 0 (1291), 1 (897), 2 (48)
* **Teenhome**: Number of teenagers in the customer’s household.
  + **Values**: 0 (1155), 1 (1029), 2 (52)
* **Total\_Kids**: Total number of children ()KidHome and TeenHome) in the customer’s household.
  + **Values**: 1 (1125), 0 (637), 2 (421), 3 (53)
* **Has\_Kids**: Binary indicator for whether the customer has children (1: Yes, 0: No).
  + **Values**: 1 (1599), 0 (637)
* **No\_Kids**: Binary indicator for whether the customer has no children (1: No, 0: Yes).
  + **Values**: 1 (637), 0 (1599)

1. **Customer Lifecycle**

* **Dt\_Customer**: Date when the customer was enrolled in the company.
* **Customer\_Tenure**: Number of days since the customer joined the company.
* **Tenure\_Group**: Categorical grouping of customers based on their tenure duration (e.g., 0-1 Years, 1-2 Years).
  + **Values**: 0-1 Years (1156), 1-2 Years (1080)

1. **Purchase Behaviour**

* **Recency**: Number of days since the customer’s last purchase.
* **Recency\_Group**: Categorical grouping based on Recency (e.g., <30 Days, 30–90 Days).
  + **Values**: 30–90 Days (1317), <30 Days (722), >90 Days (197)
* **MntWines**: Amount spent on wines in the last two years.
* **MntFruits**: Amount spent on fruits in the last two years.
* **MntMeatProducts**: Amount spent on meat products in the last two years.
* **MntFishProducts**: Amount spent on fish products in the last two years.
* **MntSweetProducts**: Amount spent on sweet products in the last two years.
* **MntGoldProds**: Amount spent on gold products in the last two years.
* **Total\_Spending**: Total amount spent by the customer across all product categories.
* **Dominant\_Category**: Product category on which the customer spent the most.
  + **Values**: MntWines (1537), MntMeatProducts (442), MntGoldProds (167), MntFishProducts (59), MntSweetProducts (16), MntFruits (15)
* **Spending\_Group**: Categorical grouping of customers based on their total spending (e.g., Low, Medium, High).
  + **Values**: Low (1635), Medium (356), High (245)

1. **Purchase Channels**

* **NumDealsPurchases**: Number of purchases made using deals.
* **Deal\_Purchase\_Category**: Categorical grouping of customers based on deal purchase activity (e.g., Low, Medium, High).
  + **Values**: No/Low (1510), Medium (579), High (147)
* **NumWebPurchases**: Number of purchases made through the company’s website.
* **NumCatalogPurchases**: Number of purchases made through catalogs.
* **NumStorePurchases**: Number of purchases made in physical stores.
* **Total\_Purchases**: Total number of purchases across all channels (web, catalog, store).
* **Dominant\_Channel**: Channel (web, catalog, store) where the customer made the most purchases.
  + **Values**: NumStorePurchases (1477), NumWebPurchases (600), NumCatalogPurchases (159)

1. **Engagement & Interaction**

* **NumWebVisitsMonth**: Number of visits to the company’s website in the last month.
* **Visit\_Frequency\_Category**: Categorical grouping of customers based on web visit frequency (e.g., Low, Medium, High).
  + **Values**: High (1169), Medium (904), Low (163)

1. **Marketing Campaigns**

* **AcceptedCmp1**: Binary indicator of whether the customer accepted Campaign 1 (1: Yes, 0: No).
  + **Values**: 0 (2092), 1 (144)
* **AcceptedCmp2**: Binary indicator of whether the customer accepted Campaign 2 (1: Yes, 0: No).
  + **Values**: 0 (2206), 1 (30)
* **AcceptedCmp3**: Binary indicator of whether the customer accepted Campaign 3 (1: Yes, 0: No).
  + **Values**: 0 (2073), 1 (163)
* **AcceptedCmp4**: Binary indicator of whether the customer accepted Campaign 4 (1: Yes, 0: No).
  + **Values**: 0 (2069), 1 (167)
* **AcceptedCmp5**: Binary indicator of whether the customer accepted Campaign 5 (1: Yes, 0: No).
  + **Values**: 0 (2074), 1 (162)
* **Response**: Binary indicator of whether the customer responded to the last campaign (1: Yes, 0: No).
  + **Values**: 0 (1902), 1 (334)
* **Total\_Campaign\_Acceptances**: Total number of campaigns accepted by the customer.
  + **Values**: 0 (1774), 1 (324), 2 (83), 3 (44), 4 (11)
* **Is\_Campaign\_Responder**: Binary indicator of whether the customer accepted any campaign (1: Yes, 0: No).
  + **Values**: 0 (1774), 1 (462)
* **Complain**: Binary indicator of whether the customer filed a complaint (1: Yes, 0: No).
  + **Values**: 0 (2216), 1 (20)

**Total Columns**: 44