**Name:** Pooja Patil

**Selected dataset**: Customer Personality Analysis

**Source**: Kaggle

**Software Used**: Excel

**Information About Dataset**:

**Total Observation**: 2240 ---------------------🡪 **After Cleaning**: 2209

**DATA CLEANING STEPS:**

* **Ckecked For Missing values by using Filter: In Income column there are total 23 missing values which is deleted and I left with 2217 observation.**
* **In Education column, I replace Basic(54 cells) with SSC and 2n cycle(200), Master and Materss cells with Masters.**
* **After that I Notice in Marital status Column there are some cells of absurd, alone and polo which is irrelevant so I deleted those columns and left with 2209 observations. Also Replace the Together cell (580 cells) with Married by using Replace function.**
* **Replace Mnt in columns names to Amt (like NmtWines to AmtWines)**
* **Rearrange the columns in order to AccecptedCmp1 to AccecptedCmp5 for better understanding.**
* **Convert income to ₹ English(India)**
* **Added some columns with their relevant column totals which may useful for further analysis:**

**They are as follows**

**Total\_kids :** kidhome + teenhome

**Total\_Amt\_Purchase:** AmtWines+AmtFruits+AmtMeatProducts+AmtFishProducts+AmtSweetProducts+AmtGoldProdsAmtWines

**Total\_NumPurchase:**

NumDealsPurchases+ NumWebPurchase+ NumCatalogPurchases+ NumStorePurchases+ NumWebVisitsMonth

**Total\_Accecpted\_Cam:** AcceptedCmp1+AcceptedCmp2+AcceptedCmp3+ AcceptedCmp4+AcceptedCmp5

**Column info:**

* ID: Customer's unique identifier
* Year\_Birth: Customer's birth year
* Education: Customer's education level
* Marital\_Status: Customer's marital status
* Income: Customer's yearly household income
* Kidhome: Number of children in customer's household
* Teenhome: Number of teenagers in customer's household
* Dt\_Customer: Date of customer's enrollment with the company
* Recency: Number of days since customer's last purchase
* Complain: 1 if the customer complained in the last 2 years, 0 otherwise

**Products**

* MntWines: Amount spent on wine in last 2 years
* MntFruits: Amount spent on fruits in last 2 years
* MntMeatProducts: Amount spent on meat in last 2 years
* MntFishProducts: Amount spent on fish in last 2 years
* MntSweetProducts: Amount spent on sweets in last 2 years
* MntGoldProds: Amount spent on gold in last 2 years

**Promotion**

* NumDealsPurchases: Number of purchases made with a discount
* AcceptedCmp1: 1 if customer accepted the offer in the 1st campaign, 0 otherwise
* AcceptedCmp2: 1 if customer accepted the offer in the 2nd campaign, 0 otherwise
* AcceptedCmp3: 1 if customer accepted the offer in the 3rd campaign, 0 otherwise
* AcceptedCmp4: 1 if customer accepted the offer in the 4th campaign, 0 otherwise
* AcceptedCmp5: 1 if customer accepted the offer in the 5th campaign, 0 otherwise
* Response: 1 if customer accepted the offer in the last campaign, 0 otherwise

**Place**

* NumWebPurchases: Number of purchases made through the company’s website
* NumCatalogPurchases: Number of purchases made using a catalogue
* NumStorePurchases: Number of purchases made directly in stores
* NumWebVisitsMonth: Number of visits to company’s website in the last month