**TP1. GROUPE (Ghislain, Fule, et Eveline)**

**1. Structure de la base de données :**

* **Nombre de colonnes :** 28
* **Nombre de lignes :** 2240
* **Types de données :**
  + int64 : 22 colonnes
  + nvarchar : 5 colonnes
  + date : 1 colonne

**Colonnes et types de données :**

| **Column Name** | **Description** | **Data Type** | **Data Type** |
| --- | --- | --- | --- |
| ID | Unique identifier for each customer | smallint | smallint |
| Year\_Birth | Year of birth of the customer | smallint | smallint |
| Education | Customer's level of education | nvarchar(50) | nvarchar(50) |
| Marital\_Status | Customer's marital status | nvarchar(50) | nvarchar(50) |
| Income | Customer's yearly household income | nvarchar(50) | int |
| Kidhome | Number of small children in customer's household | tinyint | tinyint |
| Teenhome | Number of teenagers in customer's household | tinyint | tinyint |
| Dt\_Customer | Date of customer's enrollment with the company | date | date |
| Recency | Number of days since the last purchase | tinyint | tinyint |
| MntWines | Amount spent on wines in the last 2 years | smallint | smallint |
| MntFruits | Amount spent on fruits in the last 2 years | tinyint | tinyint |
| MntMeatProducts | Amount spent on meat products in the last 2 years | smallint | smallint |
| MntFishProducts | Amount spent on fish products in the last 2 years | smallint | smallint |
| MntSweetProducts | Amount spent on sweet products in the last 2 years | smallint | smallint |
| MntGoldProds | Amount spent on gold products in the last 2 years | smallint | smallint |
| NumDealsPurchases | Number of purchases made with discount | tinyint | tinyint |
| NumWebPurchases | Number of purchases made through company's website | tinyint | tinyint |
| NumCatalogPurchases | Number of purchases made using catalog | tinyint | tinyint |
| NumStorePurchases | Number of purchases made directly in stores | tinyint | tinyint |
| NumWebVisitsMonth | Number of visits to company's website in the last month | tinyint | tinyint |
| AcceptedCmp3 | 1 if customer accepted the offer in the 3rd campaign, 0 otherwise | bit | bit |
| AcceptedCmp4 | 1 if customer accepted the offer in the 4th campaign, 0 otherwise | bit | bit |
| AcceptedCmp5 | 1 if customer accepted the offer in the 5th campaign, 0 otherwise | bit | bit |
| AcceptedCmp1 | 1 if customer accepted the offer in the 1st campaign, 0 otherwise | bit | bit |
| AcceptedCmp2 | 1 if customer accepted the offer in the 2nd campaign, 0 otherwise | bit | bit |
| Response | 1 if customer accepted the offer in the last campaign, 0 otherwise | bit | bit |
| Complain | 1 if customer complained in the last 2 years | bit | bit |
| Country | Customer's country | nvarchar(50) | nvarchar(50) |

**2.** Lien de la base de données : <https://www.kaggle.com/datasets/sahilnbajaj/marketing-campaigns-data-set>

**Qualité des données :**

* **Valeurs nulles :**
  + Income : 24 valeurs nulles
* **Forme et contenu :**
  + Income : contient des symboles de monnaie (ex: "$")
  + Dt\_Customer : contient des dates sous forme de chaînes de caractères (ex: "6/16/14")

**Charger les données dans une table dans SQL Server :**

**Étapes nécessaires :**

1. **Conversion de type :**
   * Income : supprimer les symboles de monnaie et convertir en numérique.
   * Dt\_Customer : convertir en type date.
2. **Mise en forme (casse) :**
   * Uniformiser la casse des valeurs de type object si nécessaire.