Analyse marketing

2024-01-29

Le but de cet étude est de pouvoir créer un modèle pour que nous puissions optimiser nos investissements en terme de marketing, pour améliorer notre chiffre d'affaire.

Introduction

```r
# Importation du fichier CSV
ifood_df <- read.csv("C:/Users/belli/OneDrive/Bureau/Self learning/Data/Kaggle/Marketing Analytics/ifood_df.csv")
data_marketing <- ifood_df
head(data_marketing)
```

```
##   Income Kidhome Teenhome Recency MntWines MntFruits MntMeatProducts
## 1  58138       0        0      58      635        88             546
## 2  46344       1        1      38       11         1               6
## 3  71613       0        0      26      426        49             127
## 4  26646       1        0      26       11         4              20
## 5  58293       1        0      94      173        43             118
## 6  62513       0        1      16      520        42              98
##   MntFishProducts MntSweetProducts MntGoldProds NumDealsPurchases
## 1             172               88           88                 3
## 2               2                1            6                 2
## 3             111               21           42                 1
## 4              10                3            5                 2
## 5              46               27           15                 5
## 6               0               42           14                 2
##   NumWebPurchases NumCatalogPurchases NumStorePurchases NumWebVisitsMonth
## 1               8                  10                 4                 7
## 2               1                   1                 2                 5
## 3               8                   2                10                 4
## 4               2                   0                 4                 6
## 5               5                   3                 6                 5
## 6               6                   4                10                 6
##   AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1 AcceptedCmp2 Complain
## 1            0            0            0            0            0        0
## 2            0            0            0            0            0        0
## 3            0            0            0            0            0        0
## 4            0            0            0            0            0        0
## 5            0            0            0            0            0
```

```
0
## 6              0             0            0             0             0
0
##     Z_CostContact Z_Revenue Response Age Customer_Days marital_Divorced
## 1               3        11        1  63          2822                0
## 2               3        11        0  66          2272                0
## 3               3        11        0  55          2471                0
## 4               3        11        0  36          2298                0
## 5               3        11        0  39          2320                0
## 6               3        11        0  53          2452                0
##     marital_Married marital_Single marital_Together marital_Widow
## 1                 0              1                0             0
## 2                 0              1                0             0
## 3                 0              0                1             0
## 4                 0              0                1             0
## 5                 1              0                0             0
## 6                 0              0                1             0
##     education_2n.Cycle education_Basic education_Graduation education_Master
## 1                    0               0                    1                0
## 2                    0               0                    1                0
## 3                    0               0                    1                0
## 4                    0               0                    1                0
## 5                    0               0                    0                0
## 6                    0               0                    0                1
##     education_PhD MntTotal MntRegularProds AcceptedCmpOverall
## 1               0     1529            1441                  0
## 2               0       21              15                  0
## 3               0      734             692                  0
## 4               0       48              43                  0
## 5               1      407             392                  0
## 6               0      702             688                  0
```

Cet ensemble de données consiste en des données de la société XYZ sur les profils des clients, les préférences en matière de produits, les succès/échecs des campagnes et les performances des canaux de distribution.

Ces données nous seront utiles pour produire un modèle prédictif pour estimer la probabilité qu'un client réponde favorablement à une campagne future. Ainsi, nous pourrons étudier à optimiser les futurs efforts de marketing en fonction des enseignements tirés des données.

Le nombre de campagnes acceptés au total sera notre variable expliquée.

Etude sur les données

```
nombre_de_variables <- ncol(data_marketing)
nombre_d_observations <- nrow(data_marketing)
cat("Nous avons un nombre de variables de :", nombre_de_variables, "\n")

## Nous avons un nombre de variables de : 39
```

```r
cat("Nous avons un nombre d'observations de :", nombre_d_observations, "\n")
```

Nous avons un nombre d'observations de : 2205

Nous allons identifier si nous avons des valeurs nulles, c'est-à-dire des valeurs manquantes dans notre base de données, qui peuvent être suite à un manque d'information.

```r
nombre_valeurs_manquantes <- sum(is.na(data_marketing))
cat("Le nombre total de valeurs manquantes est :", nombre_valeurs_manquantes,
"\n")
```

Le nombre total de valeurs manquantes est : 0

```r
summary(data_marketing)
```

```
##      Income          Kidhome          Teenhome          Recency
##  Min.   :  1730   Min.   :0.0000   Min.   :0.0000   Min.   : 0.00
##  1st Qu.: 35196   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:24.00
##  Median : 51287   Median :0.0000   Median :0.0000   Median :49.00
##  Mean   : 51622   Mean   :0.4422   Mean   :0.5066   Mean   :49.01
##  3rd Qu.: 68281   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:74.00
##  Max.   :113734   Max.   :2.0000   Max.   :2.0000   Max.   :99.00
##     MntWines         MntFruits       MntMeatProducts  MntFishProducts
##  Min.   :   0.0   Min.   :  0.0    Min.   :   0.0   Min.   :  0.00
##  1st Qu.:  24.0   1st Qu.:  2.0    1st Qu.:  16.0   1st Qu.:  3.00
##  Median : 178.0   Median :  8.0    Median :  68.0   Median : 12.00
##  Mean   : 306.2   Mean   : 26.4    Mean   : 165.3   Mean   : 37.76
##  3rd Qu.: 507.0   3rd Qu.: 33.0    3rd Qu.: 232.0   3rd Qu.: 50.00
##  Max.   :1493.0   Max.   :199.0    Max.   :1725.0   Max.   :259.00
##  MntSweetProducts  MntGoldProds    NumDealsPurchases NumWebPurchases
##  Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   : 0.000
##  1st Qu.:  1.00   1st Qu.:  9.00   1st Qu.: 1.000   1st Qu.: 2.000
##  Median :  8.00   Median : 25.00   Median : 2.000   Median : 4.000
##  Mean   : 27.13   Mean   : 44.06   Mean   : 2.318   Mean   : 4.101
##  3rd Qu.: 34.00   3rd Qu.: 56.00   3rd Qu.: 3.000   3rd Qu.: 6.000
##  Max.   :262.00   Max.   :321.00   Max.   :15.000   Max.   :27.000
##  NumCatalogPurchases NumStorePurchases NumWebVisitsMonth  AcceptedCmp3
##  Min.   : 0.000      Min.   : 0.000    Min.   : 0.000    Min.   :0.00000
##  1st Qu.: 0.000      1st Qu.: 3.000    1st Qu.: 3.000    1st Qu.:0.00000
##  Median : 2.000      Median : 5.000    Median : 6.000    Median :0.00000
##  Mean   : 2.645      Mean   : 5.824    Mean   : 5.337    Mean   :0.07392
##  3rd Qu.: 4.000      3rd Qu.: 8.000    3rd Qu.: 7.000    3rd Qu.:0.00000
##  Max.   :28.000      Max.   :13.000    Max.   :20.000    Max.   :1.00000
##   AcceptedCmp4       AcceptedCmp5       AcceptedCmp1      AcceptedCmp2
##  Min.   :0.00000   Min.   :0.00000   Min.   :0.0000   Min.   :0.00000
##  1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000   1st Qu.:0.00000
##  Median :0.00000   Median :0.00000   Median :0.0000   Median :0.00000
##  Mean   :0.07438   Mean   :0.07302   Mean   :0.0644   Mean   :0.01361
##  3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.0000   3rd Qu.:0.00000
##  Max.   :1.00000   Max.   :1.00000   Max.   :1.0000   Max.   :1.00000
##     Complain        Z_CostContact    Z_Revenue      Response          Age
```

```
##    Min.   :0.00000    Min.   :3      Min.   :11    Min.    :0.000    Min.    :24.0
##    1st Qu.:0.00000    1st Qu.:3      1st Qu.:11    1st Qu.:0.000     1st Qu.:43.0
##    Median :0.00000    Median :3      Median :11    Median :0.000     Median :50.0
##    Mean   :0.00907    Mean   :3      Mean   :11    Mean    :0.151    Mean    :51.1
##    3rd Qu.:0.00000    3rd Qu.:3      3rd Qu.:11    3rd Qu.:0.000     3rd Qu.:61.0
##    Max.   :1.00000    Max.   :3      Max.   :11    Max.    :1.000    Max.    :80.0
##    Customer_Days   marital_Divorced marital_Married  marital_Single
##    Min.   :2159    Min.   :0.0000    Min.   :0.0000   Min.   :0.0000
##    1st Qu.:2339    1st Qu.:0.0000    1st Qu.:0.0000   1st Qu.:0.0000
##    Median :2515    Median :0.0000    Median :0.0000   Median :0.0000
##    Mean   :2513    Mean   :0.1043    Mean   :0.3873   Mean   :0.2163
##    3rd Qu.:2688    3rd Qu.:0.0000    3rd Qu.:1.0000   3rd Qu.:0.0000
##    Max.   :2858    Max.   :1.0000    Max.   :1.0000   Max.   :1.0000
##    marital_Together marital_Widow     education_2n.Cycle education_Basic
##    Min.   :0.0000   Min.   :0.00000   Min.   :0.0000     Min.   :0.00000
##    1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:0.0000     1st Qu.:0.00000
##    Median :0.0000   Median :0.00000   Median :0.0000     Median :0.00000
##    Mean   :0.2576   Mean   :0.03447   Mean   :0.0898     Mean   :0.02449
##    3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:0.0000     3rd Qu.:0.00000
##    Max.   :1.0000   Max.   :1.00000   Max.   :1.0000     Max.   :1.00000
##    education_Graduation education_Master education_PhD      MntTotal
##    Min.   :0.0000       Min.   :0.0000   Min.   :0.0000    Min.   :   4.0
##    1st Qu.:0.0000       1st Qu.:0.0000   1st Qu.:0.0000    1st Qu.:  56.0
##    Median :1.0000       Median :0.0000   Median :0.0000    Median : 343.0
##    Mean   :0.5048       Mean   :0.1651   Mean   :0.2159    Mean   : 562.8
##    3rd Qu.:1.0000       3rd Qu.:0.0000   3rd Qu.:0.0000    3rd Qu.: 964.0
##    Max.   :1.0000       Max.   :1.0000   Max.   :1.0000    Max.   :2491.0
##    MntRegularProds   AcceptedCmpOverall
##    Min.   :-283.0    Min.   :0.0000
##    1st Qu.:  42.0    1st Qu.:0.0000
##    Median : 288.0    Median :0.0000
##    Mean   : 518.7    Mean   :0.2993
##    3rd Qu.: 884.0    3rd Qu.:0.0000
##    Max.   :2458.0    Max.   :4.0000
```

Nous pouvons étudier la clientèle des six dernières campagnes de l'entreprise :

Nous pouvons voir ici que la plage de revenu varie de 1 730 à 113 734, avec une moyenne d'environ 51 622. Il y a une variation importante dans les niveaux de revenu des individus dans l'échantillon.

La plupart des clients n'ont pas d'enfants à charge, les valeurs médianes et du 1er quartile étant toutes deux égales à zéro.

L'âge des clients varie de 24 à 80 ans, avec une moyenne d'environ 51 ans.

```r
#Création d'un modèle avec le nombre d'acceptation total de campagne en tant
que variable expliquée.

library(leaps)
```

UONG Bellinna

```
Modèle1 <- regsubsets(
  AcceptedCmpOverall~Income + Kidhome + Teenhome + Recency + MntWines + MntFr
uits
  + MntMeatProducts + MntSweetProducts + MntGoldProds + NumDealsPurchases +
    NumCatalogPurchases + NumStorePurchases + NumWebVisitsMonth + AcceptedCmp
3
  + AcceptedCmp4 + AcceptedCmp5 + AcceptedCmp1 + AcceptedCmp2 + Complain +
    Z_CostContact + Z_Revenue + Response + Age + Customer_Days + marital_Divo
rced
  + marital_Married + marital_Single + marital_Together + marital_Widow +
    education_2n.Cycle + education_Basic + education_Graduation + education_M
aster
  + education_PhD + MntTotal + MntRegularProds, data = data_marketing,
  nvmax = ncol(data_marketing), method = "exhaustive", really.big = TRUE
  )

## Warning in leaps.setup(x, y, wt = wt, nbest = nbest, nvmax = nvmax, force.
in =
## force.in, : 5 linear dependencies found

## Reordering variables and trying again:
```

La sortie indique que le modèle considéré a des dépendances linéaires entre certaines
variables. Les dépendances linéaires peuvent poser des problèmes lors de la modélisation.
Nous allons tout d'abord vérifier les corrélations.

```
cor(data_marketing)

## Warning in cor(data_marketing): l'écart type est nul

##                          Income       Kidhome      Teenhome         Recency
## Income              1.000000000  -0.531698920   0.042482894    0.0067162140
## Kidhome            -0.531698920   1.000000000  -0.040206742    0.0118290564
## Teenhome            0.042482894  -0.040206742   1.000000000    0.0138814421
## Recency             0.006716214   0.011829056   0.013881442    1.0000000000
## MntWines            0.730495161  -0.499287609   0.002783478    0.0164704563
## MntFruits           0.537920258  -0.374388312  -0.176925289   -0.0049087482
## MntMeatProducts     0.702499620  -0.445665370  -0.267176819    0.0261384879
## MntFishProducts     0.551758002  -0.389894860  -0.206371293    0.0011773193
## MntSweetProducts    0.555601037  -0.379101220  -0.164246040    0.0255352851
## MntGoldProds        0.417652875  -0.356549651  -0.019619039    0.0185575295
## NumDealsPurchases  -0.132426606   0.226433801   0.394340657    0.0004791932
## NumWebPurchases     0.503184202  -0.375590214   0.161228959   -0.0051044692
## NumCatalogPurchases 0.710056534  -0.519813290  -0.114018504    0.0297495560
## NumStorePurchases   0.687205716  -0.506543168   0.047321148    0.0004621200
## NumWebVisitsMonth  -0.648306273   0.448497177   0.129364940   -0.0179057586
## AcceptedCmp3       -0.011180865   0.015897071  -0.043222979   -0.0323274025
## AcceptedCmp4        0.233266915  -0.162596703   0.037860408    0.0176583761
## AcceptedCmp5        0.416385928  -0.205123678  -0.190760447    0.0003338488
## AcceptedCmp1        0.345242062  -0.174740838  -0.145747695   -0.0210968376
## AcceptedCmp2        0.110209533  -0.082124037  -0.015804880   -0.0013902334
```

```
## Complain                -0.027487682  0.037025481  0.007633360  0.0057582284
## Z_CostContact                     NA           NA           NA            NA
## Z_Revenue                         NA           NA           NA            NA
## Response                 0.174901766 -0.078409289 -0.155196040 -0.2004131359
## Age                      0.212625315 -0.238082811  0.362918641  0.0142277160
## Customer_Days           -0.024892175 -0.055743362  0.019394380  0.0283384487
## marital_Divorced         0.013892287 -0.018513519  0.055851956  0.0014831209
## marital_Married         -0.010426970  0.019730566  0.007499336 -0.0211055870
## marital_Single          -0.015539170  0.014525125 -0.100454277 -0.0009263338
## marital_Together        -0.001959733  0.007422499  0.027180977  0.0239081937
## marital_Widow            0.044336453 -0.072243537  0.047962337 -0.0013484303
## education_2n.Cycle      -0.060620749  0.019050320 -0.056259159 -0.0067891971
## education_Basic         -0.239604426  0.055307759 -0.120519094 -0.0030928283
## education_Graduation     0.017643915 -0.001930415 -0.024697766  0.0314185337
## education_Master         0.021632987  0.011482238  0.023805901 -0.0255630459
## education_PhD            0.091175907 -0.042030614  0.092901462 -0.0092342602
## MntTotal                 0.823066002 -0.551152041 -0.142995165  0.0211321193
## MntRegularProds          0.816879229 -0.539828043 -0.146865727  0.0202414498
## AcceptedCmpOverall       0.388247306 -0.212079586 -0.130254745 -0.0133440232
##                             MntWines      MntFruits MntMeatProducts MntFishProd
ucts
## Income                   0.730495161  0.537920258     0.702499620    0.551758
0018
## Kidhome                 -0.499287609 -0.374388312    -0.445665370   -0.389894
8601
## Teenhome                 0.002783478 -0.176925289    -0.267176819   -0.206371
2925
## Recency                  0.016470456 -0.004908748     0.026138488    0.001177
3193
## MntWines                 1.000000000  0.384946529     0.593118931    0.395966
6904
## MntFruits                0.384946529  1.000000000     0.568100098    0.592556
4348
## MntMeatProducts          0.593118931  0.568100098     1.000000000    0.595673
0220
## MntFishProducts          0.395966690  0.592556435     0.595673022    1.000000
0000
## MntSweetProducts         0.388612745  0.570986145     0.556511024    0.582974
2402
## MntGoldProds             0.390193557  0.392596096     0.375581109    0.425420
3988
## NumDealsPurchases        0.011858138 -0.136350120    -0.165521696   -0.145029
9761
## NumWebPurchases          0.552342087  0.300813200     0.329453160    0.297775
8808
## NumCatalogPurchases      0.673233794  0.513686252     0.714382319    0.563173
5086
## NumStorePurchases        0.639372823  0.459055955     0.517244926    0.456896
1632
## NumWebVisitsMonth       -0.329394738 -0.424463422    -0.543387420   -0.453353
```

UONG Bellinna

```
1126
## AcceptedCmp3           0.060699843  0.014131448    0.021224280   -0.000831
7887
## AcceptedCmp4           0.373062550  0.006078480    0.096798404    0.015512
8497
## AcceptedCmp5           0.472728958  0.208615442    0.389276144    0.194387
1733
## AcceptedCmp1           0.351345596  0.192061430    0.325306235    0.261389
0433
## AcceptedCmp2           0.206230787 -0.010146894    0.045841628    0.002093
1042
## Complain              -0.036709042 -0.003134561   -0.020920757   -0.019298
7079
## Z_CostContact                  NA           NA             NA
NA
## Z_Revenue                      NA           NA             NA
NA
## Response               0.245559120  0.122331423    0.248821266    0.107404
6263
## Age                    0.164437717  0.013148505    0.041540408    0.040854
9404
## Customer_Days          0.168102391  0.067978065    0.089203013    0.081610
6853
## marital_Divorced       0.021679242  0.010566658   -0.021688450   -0.015212
5174
## marital_Married       -0.012597170 -0.013723228   -0.027768869   -0.031728
3469
## marital_Single        -0.022597803  0.011981916    0.045574836    0.013808
7450
## marital_Together       0.005914940 -0.014209713   -0.004063846    0.015501
8206
## marital_Widow          0.034138632  0.025960526    0.017369525    0.041886
0288
## education_2n.Cycle    -0.096259081  0.025451644   -0.041737589    0.061303
8319
## education_Basic       -0.140368593 -0.060915462   -0.111968356   -0.059839
6169
## education_Graduation  -0.060919622  0.114919364    0.064916634    0.106227
0438
## education_Master       0.036403273 -0.055581096   -0.004020271   -0.050152
8593
## education_PhD          0.160803896 -0.084300533   -0.004194439   -0.103952
4467
## MntTotal               0.902309602  0.606657736    0.861391527    0.635038
0686
## MntRegularProds        0.901848425  0.594180343    0.860663357    0.620626
4002
## AcceptedCmpOverall     0.509912967  0.155132826    0.319553355    0.174674
5809
##                       MntSweetProducts MntGoldProds NumDealsPurchases
```

```
## Income                0.5556010366    0.41765287    -0.1324266057
## Kidhome              -0.3791012196   -0.35654965     0.2264338014
## Teenhome             -0.1642460404   -0.01961904     0.3943406572
## Recency               0.0255352851    0.01855753     0.0004791932
## MntWines              0.3886127454    0.39019356     0.0118581384
## MntFruits             0.5709861452    0.39259610    -0.1363501196
## MntMeatProducts       0.5565110236    0.37558111    -0.1655216958
## MntFishProducts       0.5829742402    0.42542040    -0.1450299761
## MntSweetProducts      1.0000000000    0.35574718    -0.1222791081
## MntGoldProds          0.3557471824    1.00000000     0.0569260865
## NumDealsPurchases    -0.1222791081    0.05692609     1.0000000000
## NumWebPurchases       0.3320569860    0.40596098     0.2525018096
## NumCatalogPurchases   0.5243692500    0.47103151    -0.0586248572
## NumStorePurchases     0.4541333507    0.38857519     0.0727099347
## NumWebVisitsMonth    -0.4293754169   -0.25302187     0.3635450133
## AcceptedCmp3          0.0010993039    0.12498401    -0.0228824818
## AcceptedCmp4          0.0286645349    0.02361303     0.0172212120
## AcceptedCmp5          0.2580529217    0.17611781    -0.1869746513
## AcceptedCmp1          0.2447711537    0.17037979    -0.1295378645
## AcceptedCmp2          0.0099152222    0.05073102    -0.0385132167
## Complain             -0.0207728019   -0.03044021     0.0041417648
## Z_CostContact                   NA           NA              NA
## Z_Revenue                       NA           NA              NA
## Response              0.1153256618    0.14020969     0.0046907375
## Age                   0.0210751922    0.05929469     0.0720051178
## Customer_Days         0.0808433945    0.16140682     0.2210926616
## marital_Divorced     -0.0008125344    0.01663326     0.0242153105
## marital_Married      -0.0056058313   -0.01641145     0.0158552796
## marital_Single       -0.0027112645   -0.00100633    -0.0454802055
## marital_Together     -0.0112201466   -0.01037480     0.0066915948
## marital_Widow         0.0493470493    0.04309587     0.0036968197
## education_2n.Cycle    0.0605500038    0.01918907    -0.0076021686
## education_Basic      -0.0578630690   -0.06501355    -0.0438670606
## education_Graduation  0.1040753586    0.13175936    -0.0020890503
## education_Master     -0.0677233283   -0.03249239     0.0266346756
## education_PhD        -0.0857016184   -0.11970789     0.0002672288
## MntTotal              0.6045142898    0.46369396    -0.0875985595
## MntRegularProds       0.5953935536    0.38877553    -0.0964099469
## AcceptedCmpOverall    0.2001735293    0.19464651    -0.1269623221
##                    NumWebPurchases NumCatalogPurchases NumStorePurchases
## Income               0.5031842017         0.710056534       0.687205716
## Kidhome             -0.3755902140        -0.519813290      -0.506543168
## Teenhome             0.1612289589        -0.114018504       0.047321148
## Recency             -0.0051044692         0.029749556       0.000462120
## MntWines             0.5523420869         0.673233794       0.639372823
## MntFruits            0.3008132004         0.513686252       0.459055955
## MntMeatProducts      0.3294531598         0.714382319       0.517244926
## MntFishProducts      0.2977758808         0.563173509       0.456896163
## MntSweetProducts     0.3320569860         0.524369250       0.454133351
## MntGoldProds         0.4059609804         0.471031515       0.388575189
```

UONG Bellinna

```
## NumDealsPurchases        0.2525018096      -0.058624857       0.072709935
## NumWebPurchases          1.0000000000       0.420889600       0.512108636
## NumCatalogPurchases      0.4208896004       1.000000000       0.561261336
## NumStorePurchases        0.5121086364       0.561261336       1.000000000
## NumWebVisitsMonth       -0.0614499203      -0.530622915      -0.449355732
## AcceptedCmp3             0.0415400083       0.112007125      -0.071258836
## AcceptedCmp4             0.1619872516       0.149593798       0.177018648
## AcceptedCmp5             0.1406317650       0.339603307       0.213742416
## AcceptedCmp1             0.1584462891       0.326437463       0.177887283
## AcceptedCmp2             0.0343010965       0.105829931       0.084904675
## Complain                -0.0140069543      -0.018647308      -0.012503751
## Z_CostContact                     NA                NA                NA
## Z_Revenue                         NA                NA                NA
## Response                 0.1496994148       0.234523961       0.033508605
## Age                      0.1618386420       0.140724065       0.138613576
## Customer_Days            0.1929023597       0.096862709       0.112057171
## marital_Divorced         0.0297327080       0.005604585       0.003925981
## marital_Married         -0.0003336935      -0.018343534       0.011105077
## marital_Single          -0.0430733433      -0.001115097      -0.023398026
## marital_Together         0.0052343772      -0.000948703      -0.006013909
## marital_Widow            0.0357427377       0.044383325       0.030994339
## education_2n.Cycle      -0.0358991519      -0.030490293      -0.022059118
## education_Basic         -0.1280493303      -0.122534453      -0.145277689
## education_Graduation     0.0085983551       0.026817729       0.009614146
## education_Master        -0.0092161358      -0.014803832       0.010634784
## education_PhD            0.0709179438       0.047989030       0.048626240
## MntTotal                 0.5210859914       0.791186893       0.677892837
## MntRegularProds          0.5039472057       0.778742425       0.668632189
## AcceptedCmpOverall       0.1952484979       0.366459493       0.201254206
##                        NumWebVisitsMonth AcceptedCmp3  AcceptedCmp4
## Income                      -0.648306273 -0.0111808646  0.2332669151
## Kidhome                      0.448497177  0.0158970709 -0.1625967027
## Teenhome                     0.129364940 -0.0432229787  0.0378604081
## Recency                     -0.017905759 -0.0323274025  0.0176583761
## MntWines                    -0.329394738  0.0606998426  0.3730625502
## MntFruits                   -0.424463422  0.0141314481  0.0060784801
## MntMeatProducts             -0.543387420  0.0212242802  0.0967984039
## MntFishProducts             -0.453353113 -0.0008317887  0.0155128497
## MntSweetProducts            -0.429375417  0.0010993039  0.0286645349
## MntGoldProds                -0.253021875  0.1249840073  0.0236130261
## NumDealsPurchases            0.363545013 -0.0228824818  0.0172212120
## NumWebPurchases             -0.061449920  0.0415400083  0.1619872516
## NumCatalogPurchases         -0.530622915  0.1120071254  0.1495937979
## NumStorePurchases           -0.449355732 -0.0712588363  0.1770186478
## NumWebVisitsMonth            1.000000000  0.0596753516 -0.0309888001
## AcceptedCmp3                 0.059675352  1.0000000000 -0.0800877984
## AcceptedCmp4                -0.030988800 -0.0800877984  1.0000000000
## AcceptedCmp5                -0.279758802  0.0806049910  0.3124239024
## AcceptedCmp1                -0.198196845  0.0953502260  0.2425025017
## AcceptedCmp2                -0.008288523  0.0715541217  0.2949532491
```

```
## Complain                     0.020341553   0.0095353239 -0.0271200009
## Z_CostContact                          NA             NA            NA
## Z_Revenue                              NA             NA            NA
## Response                    -0.005358181   0.2535758176  0.1797273256
## Age                         -0.122712674  -0.0614048421  0.0699031070
## Customer_Days                0.277656346  -0.0079519735  0.0157227120
## marital_Divorced             0.024902344   0.0169993502  0.0050521395
## marital_Married              0.012436977  -0.0004631307 -0.0053842585
## marital_Single              -0.011288822   0.0157392942 -0.0145984793
## marital_Together            -0.007475640  -0.0197711997 -0.0009715325
## marital_Widow               -0.031535545  -0.0153747672  0.0411907189
## education_2n.Cycle           0.017277897   0.0022024424 -0.0346218661
## education_Basic              0.100688411   0.0225200244 -0.0449135203
## education_Graduation        -0.018433997  -0.0148246710 -0.0130709596
## education_Master            -0.022099934  -0.0135767011  0.0182829692
## education_PhD               -0.007488926   0.0202756412  0.0403173323
## MntTotal                    -0.501639049   0.0445707968  0.2591582573
## MntRegularProds             -0.498011053   0.0346734015  0.2672887843
## AcceptedCmpOverall          -0.168914313   0.4311369779  0.6121006925
##                        AcceptedCmp5 AcceptedCmp1  AcceptedCmp2       Complai
n
## Income                 0.4163859276   0.345242062  0.1102095332 -0.027487682
3
## Kidhome               -0.2051236779  -0.174740838 -0.0821240370  0.037025480
9
## Teenhome              -0.1907604472  -0.145747695 -0.0158048805  0.007633360
1
## Recency                0.0003338488  -0.021096838 -0.0013902334  0.005758228
4
## MntWines               0.4727289575   0.351345596  0.2062307875 -0.036709042
0
## MntFruits              0.2086154417   0.192061430 -0.0101468939 -0.003134560
5
## MntMeatProducts        0.3892761445   0.325306235  0.0458416282 -0.020920756
7
## MntFishProducts        0.1943871733   0.261389043  0.0020931042 -0.019298707
9
## MntSweetProducts       0.2580529217   0.244771154  0.0099152222 -0.020772801
9
## MntGoldProds           0.1761178114   0.170379785  0.0507310189 -0.030440206
6
## NumDealsPurchases     -0.1869746513  -0.129537865 -0.0385132167  0.004141764
8
## NumWebPurchases        0.1406317650   0.158446289  0.0343010965 -0.014006954
3
## NumCatalogPurchases    0.3396033073   0.326437463  0.1058299315 -0.018647308
1
## NumStorePurchases      0.2137424156   0.177887283  0.0849046745 -0.012503751
2
## NumWebVisitsMonth     -0.2797588016  -0.198196845 -0.0082885234  0.020341552
```

```
8
## AcceptedCmp3           0.0806049910   0.095350226  0.0715541217  0.009535323
9
## AcceptedCmp4           0.3124239024   0.242502502  0.2949532491 -0.027120000
9
## AcceptedCmp5           1.0000000000   0.409282524  0.2228466174 -0.008463927
6
## AcceptedCmp1           0.4092825241   1.000000000  0.1765203451 -0.025100584
5
## AcceptedCmp2           0.2228466174   0.176520345  1.0000000000 -0.011236221
0
## Complain              -0.0084639276  -0.025100585 -0.0112362210  1.000000000
0
## Z_CostContact                   NA            NA            NA            N
A
## Z_Revenue                       NA            NA            NA            N
A
## Response                0.3246481844   0.296977534  0.1691283541 -0.000272644
5
## Age                    -0.0192752300   0.011747497  0.0077369304  0.004531459
4
## Customer_Days         -0.0045504853  -0.040196525  0.0060205151  0.035895077
4
## marital_Divorced      -0.0216354082  -0.016996145  0.0239599676 -0.016998763
5
## marital_Married        0.0130415717   0.026561534 -0.0371205298  0.002493967
0
## marital_Single        -0.0119754351   0.001263766 -0.0141649178  0.019442611
1
## marital_Together       0.0060868307  -0.019344479  0.0382438968 -0.001661902
7
## marital_Widow          0.0138632875   0.001070206 -0.0007299219 -0.018076246
0
## education_2n.Cycle    -0.0271772483   0.008071679 -0.0095015595  0.020147331
2
## education_Basic       -0.0444681608  -0.041569158 -0.0186083417 -0.015158839
4
## education_Graduation   0.0165030190   0.030759413  0.0067113926  0.037359713
0
## education_Master       0.0019825252  -0.027079303 -0.0311324389 -0.016771178
9
## education_PhD          0.0137465773  -0.002936774  0.0335297093 -0.038572042
7
## MntTotal               0.4755592486   0.384526158  0.1383903298 -0.032959232
8
## MntRegularProds        0.4780748792   0.383947025  0.1391709872 -0.031430286
6
## AcceptedCmpOverall     0.7195601412   0.677609509  0.4604888865 -0.020999797
9
##                        Z_CostContact Z_Revenue      Response          Age
```

```
## Income                        NA        NA  0.1749017657  0.212625315
## Kidhome                       NA        NA -0.0784092890 -0.238082811
## Teenhome                      NA        NA -0.1551960399  0.362918641
## Recency                       NA        NA -0.2004131359  0.014227716
## MntWines                      NA        NA  0.2455591201  0.164437717
## MntFruits                     NA        NA  0.1223314226  0.013148505
## MntMeatProducts               NA        NA  0.2488212658  0.041540408
## MntFishProducts               NA        NA  0.1074046263  0.040854940
## MntSweetProducts              NA        NA  0.1153256618  0.021075192
## MntGoldProds                  NA        NA  0.1402096860  0.059294693
## NumDealsPurchases             NA        NA  0.0046907375  0.072005118
## NumWebPurchases               NA        NA  0.1496994148  0.161838642
## NumCatalogPurchases           NA        NA  0.2345239607  0.140724065
## NumStorePurchases             NA        NA  0.0335086052  0.138613576
## NumWebVisitsMonth             NA        NA -0.0053581812 -0.122712674
## AcceptedCmp3                  NA        NA  0.2535758176 -0.061404842
## AcceptedCmp4                  NA        NA  0.1797273256  0.069903107
## AcceptedCmp5                  NA        NA  0.3246481844 -0.019275230
## AcceptedCmp1                  NA        NA  0.2969775341  0.011747497
## AcceptedCmp2                  NA        NA  0.1691283541  0.007736930
## Complain                      NA        NA -0.0002726445  0.004531459
## Z_CostContact                  1        NA           NA           NA
## Z_Revenue                     NA         1           NA           NA
## Response                      NA        NA  1.0000000000 -0.021305489
## Age                           NA        NA -0.0213054894  1.000000000
## Customer_Days                 NA        NA  0.1965199350 -0.013462073
## marital_Divorced              NA        NA  0.0549671682  0.069347375
## marital_Married               NA        NA -0.0805263812 -0.045480437
## marital_Single                NA        NA  0.1137032362 -0.128435761
## marital_Together              NA        NA -0.0746639710  0.054820361
## marital_Widow                 NA        NA  0.0452845043  0.163720688
## education_2n.Cycle            NA        NA -0.0350079389 -0.104364280
## education_Basic               NA        NA -0.0504372880 -0.115872277
## education_Graduation          NA        NA -0.0407488169 -0.061579332
## education_Master              NA        NA  0.0035090570  0.074754071
## education_PhD                 NA        NA  0.0896269714  0.123428972
## MntTotal                      NA        NA  0.2648950158  0.118369510
## MntRegularProds               NA        NA  0.2623627717  0.117551703
## AcceptedCmpOverall            NA        NA  0.4269607089  0.001529105
##                     Customer_Days marital_Divorced marital_Married
## Income               -0.024892175      0.0138922870    -0.0104269699
## Kidhome              -0.055743362     -0.0185135191     0.0197305661
## Teenhome              0.019394380      0.0558519558     0.0074993358
## Recency               0.028338449      0.0014831209    -0.0211055870
## MntWines              0.168102391      0.0216792418    -0.0125971702
## MntFruits             0.067978065      0.0105666578    -0.0137232283
## MntMeatProducts       0.089203013     -0.0216884499    -0.0277688690
## MntFishProducts       0.081610685     -0.0152125174    -0.0317283469
## MntSweetProducts      0.080843394     -0.0008125344    -0.0056058313
## MntGoldProds          0.161406816      0.0166332618    -0.0164114463
```

UONG Bellinna

```
## NumDealsPurchases       0.221092662      0.0242153105     0.0158552796
## NumWebPurchases         0.192902360      0.0297327080    -0.0003336935
## NumCatalogPurchases     0.096862709      0.0056045845    -0.0183435341
## NumStorePurchases       0.112057171      0.0039259813     0.0111050769
## NumWebVisitsMonth       0.277656346      0.0249023437     0.0124369767
## AcceptedCmp3           -0.007951973      0.0169993502    -0.0004631307
## AcceptedCmp4            0.015722712      0.0050521395    -0.0053842585
## AcceptedCmp5           -0.004550485     -0.0216354082     0.0130415717
## AcceptedCmp1           -0.040196525     -0.0169961453     0.0265615339
## AcceptedCmp2            0.006020515      0.0239599676    -0.0371205298
## Complain                0.035895077     -0.0169987635     0.0024939670
## Z_CostContact                    NA               NA               NA
## Z_Revenue                        NA               NA               NA
## Response                0.196519935      0.0549671682    -0.0805263812
## Age                    -0.013462073      0.0693473747    -0.0454804366
## Customer_Days           1.000000000      0.0097204635    -0.0158253211
## marital_Divorced        0.009720464      1.0000000000    -0.2713198878
## marital_Married        -0.015825321     -0.2713198878     1.0000000000
## marital_Single          0.002079478     -0.1792947977    -0.4177231856
## marital_Together        0.003425653     -0.2010158007    -0.4683290408
## marital_Widow           0.013066188     -0.0644761851    -0.1502173950
## education_2n.Cycle      0.011465726      0.0069904108     0.0107928377
## education_Basic         0.058275291     -0.0444706849    -0.0055070223
## education_Graduation    0.029693136      0.0056525306    -0.0038482521
## education_Master       -0.033257199     -0.0038696602    -0.0074673530
## education_PhD          -0.035934184      0.0084719097     0.0059839156
## MntTotal                0.150475562      0.0037263640    -0.0222509333
## MntRegularProds         0.141399733      0.0023212332    -0.0216053574
## AcceptedCmpOverall     -0.012213083     -0.0018397785     0.0059954332
##                       marital_Single marital_Together marital_Widow
## Income                  -0.0155391701     -0.0019597334  0.0443364527
## Kidhome                  0.0145251253      0.0074224991 -0.0722435373
## Teenhome                -0.1004542768      0.0271809772  0.0479623371
## Recency                 -0.0009263338      0.0239081937 -0.0013484303
## MntWines                -0.0225978032      0.0059149404  0.0341386316
## MntFruits                0.0119819160     -0.0142097127  0.0259605261
## MntMeatProducts          0.0455748362     -0.0040638461  0.0173695249
## MntFishProducts          0.0138087450      0.0155018206  0.0418860288
## MntSweetProducts        -0.0027112645     -0.0112201466  0.0493470493
## MntGoldProds            -0.0010063297     -0.0103748006  0.0430958712
## NumDealsPurchases       -0.0454802055      0.0066915948  0.0036968197
## NumWebPurchases         -0.0430733433      0.0052343772  0.0357427377
## NumCatalogPurchases     -0.0011150971     -0.0009487030  0.0443833245
## NumStorePurchases       -0.0233980257     -0.0060139090  0.0309943387
## NumWebVisitsMonth       -0.0112888223     -0.0074756402 -0.0315355448
## AcceptedCmp3             0.0157392942     -0.0197711997 -0.0153747672
## AcceptedCmp4            -0.0145984793     -0.0009715325  0.0411907189
## AcceptedCmp5            -0.0119754351      0.0060868307  0.0138632875
## AcceptedCmp1             0.0012637663     -0.0193444795  0.0010702063
## AcceptedCmp2            -0.0141649178      0.0382438968 -0.0007299219
```

```
## Complain                 0.0194426111  -0.0016619027 -0.0180762460
## Z_CostContact                       NA             NA            NA
## Z_Revenue                           NA             NA            NA
## Response                  0.1137032362  -0.0746639710  0.0452845043
## Age                      -0.1284357614   0.0548203605  0.1637206879
## Customer_Days             0.0020794778   0.0034256533  0.0130661879
## marital_Divorced         -0.1792947977  -0.2010158007 -0.0644761851
## marital_Married          -0.4177231856  -0.4683290408 -0.1502173950
## marital_Single            1.0000000000  -0.3094832498 -0.0992673174
## marital_Together         -0.3094832498   1.0000000000 -0.1112932420
## marital_Widow            -0.0992673174  -0.1112932420  1.0000000000
## education_2n.Cycle       -0.0301772836   0.0181225430 -0.0158653305
## education_Basic           0.0450262027   0.0006024897 -0.0138520246
## education_Graduation      0.0159246945  -0.0076844270 -0.0167163121
## education_Master         -0.0051708476   0.0202099980 -0.0103527276
## education_PhD            -0.0106322044  -0.0217174286  0.0458842477
## MntTotal                  0.0059400812   0.0016221897  0.0358776638
## MntRegularProds           0.0062709993   0.0026560244  0.0332829202
## AcceptedCmpOverall       -0.0061129837  -0.0061184979  0.0155373413
##                        education_2n.Cycle education_Basic education_Graduati
## on
## Income                       -0.060620749    -0.2396044262        0.0176439
## 15
## Kidhome                       0.019050320     0.0553077590       -0.0019304
## 15
## Teenhome                     -0.056259159    -0.1205190938       -0.0246977
## 66
## Recency                      -0.006789197    -0.0030928283        0.0314185
## 34
## MntWines                     -0.096259081    -0.1403685926       -0.0609196
## 22
## MntFruits                     0.025451644    -0.0609154617        0.1149193
## 64
## MntMeatProducts              -0.041737589    -0.1119683560        0.0649166
## 34
## MntFishProducts               0.061303832    -0.0598396169        0.1062270
## 44
## MntSweetProducts              0.060550004    -0.0578630690        0.1040753
## 59
## MntGoldProds                  0.019189066    -0.0650135518        0.1317593
## 64
## NumDealsPurchases            -0.007602169    -0.0438670606       -0.0020890
## 50
## NumWebPurchases              -0.035899152    -0.1280493303        0.0085983
## 55
## NumCatalogPurchases          -0.030490293    -0.1225344530        0.0268177
## 29
## NumStorePurchases            -0.022059118    -0.1452776887        0.0096141
## 46
## NumWebVisitsMonth             0.017277897     0.1006884107       -0.0184339
```

```
97
## AcceptedCmp3              0.002202442   0.0225200244    -0.0148246
71
## AcceptedCmp4             -0.034621866  -0.0449135203    -0.0130709
60
## AcceptedCmp5             -0.027177248  -0.0444681608     0.0165030
19
## AcceptedCmp1              0.008071679  -0.0415691585     0.0307594
13
## AcceptedCmp2             -0.009501559  -0.0186083417     0.0067113
93
## Complain                 0.020147331  -0.0151588394     0.0373597
13
## Z_CostContact                     NA            NA
NA
## Z_Revenue                         NA            NA
NA
## Response                -0.035007939  -0.0504372880    -0.0407488
17
## Age                     -0.104364280  -0.1158722769    -0.0615793
32
## Customer_Days            0.011465726   0.0582752913     0.0296931
36
## marital_Divorced         0.006990411  -0.0444706849     0.0056525
31
## marital_Married          0.010792838  -0.0055070223    -0.0038482
52
## marital_Single          -0.030177284   0.0450262027     0.0159246
95
## marital_Together         0.018122543   0.0006024897    -0.0076844
27
## marital_Widow           -0.015865331  -0.0138520246    -0.0167163
12
## education_2n.Cycle       1.000000000  -0.0497663264    -0.3170992
19
## education_Basic         -0.049766326   1.0000000000    -0.1599605
68
## education_Graduation    -0.317099219  -0.1599605680     1.0000000
00
## education_Master        -0.139663431  -0.0704531592    -0.4489108
07
## education_PhD           -0.164802915  -0.0831347614    -0.5297149
65
## MntTotal                -0.060271630  -0.1386308418     0.0143320
81
## MntRegularProds         -0.064468002  -0.1380869260     0.0025957
41
## AcceptedCmpOverall      -0.021605392  -0.0438353851     0.0078097
22
##                     education_Master education_PhD     MntTotal
```

```
## Income                    0.021632987   0.0911759074   0.823066002
## Kidhome                   0.011482238  -0.0420306144  -0.551152041
## Teenhome                  0.023805901   0.0929014623  -0.142995165
## Recency                  -0.025563046  -0.0092342602   0.021132119
## MntWines                  0.036403273   0.1608038957   0.902309602
## MntFruits                -0.055581096  -0.0843005327   0.606657736
## MntMeatProducts          -0.004020271  -0.0041944389   0.861391527
## MntFishProducts          -0.050152859  -0.1039524467   0.635038069
## MntSweetProducts         -0.067723328  -0.0857016184   0.604514290
## MntGoldProds             -0.032492386  -0.1197078948   0.463693960
## NumDealsPurchases         0.026634676   0.0002672288  -0.087598559
## NumWebPurchases          -0.009216136   0.0709179438   0.521085991
## NumCatalogPurchases      -0.014803832   0.0479890296   0.791186893
## NumStorePurchases         0.010634784   0.0486262403   0.677892837
## NumWebVisitsMonth        -0.022099934  -0.0074889259  -0.501639049
## AcceptedCmp3             -0.013576701   0.0202756412   0.044570797
## AcceptedCmp4              0.018282969   0.0403173323   0.259158257
## AcceptedCmp5              0.001982525   0.0137465773   0.475559249
## AcceptedCmp1             -0.027079303  -0.0029367738   0.384526158
## AcceptedCmp2             -0.031132439   0.0335297093   0.138390330
## Complain                 -0.016771179  -0.0385720427  -0.032959233
## Z_CostContact                      NA             NA             NA
## Z_Revenue                          NA             NA             NA
## Response                  0.003509057   0.0896269714   0.264895016
## Age                       0.074754071   0.1234289722   0.118369510
## Customer_Days            -0.033257199  -0.0359341837   0.150475562
## marital_Divorced         -0.003869660   0.0084719097   0.003726364
## marital_Married          -0.007467353   0.0059839156  -0.022250933
## marital_Single           -0.005170848  -0.0106322044   0.005940081
## marital_Together          0.020209998  -0.0217174286   0.001622190
## marital_Widow            -0.010352728   0.0458842477   0.035877664
## education_2n.Cycle       -0.139663431  -0.1648029148  -0.060271630
## education_Basic          -0.070453159  -0.0831347614  -0.138630842
## education_Graduation     -0.448910807  -0.5297149654   0.014332081
## education_Master          1.000000000  -0.2333080787   0.006361735
## education_PhD            -0.233308079   1.0000000000   0.070804430
## MntTotal                  0.006361735   0.0708044297   1.000000000
## MntRegularProds           0.009650659   0.0848106001   0.996569260
## AcceptedCmpOverall       -0.012484312   0.0332556661   0.461279404
##                       MntRegularProds AcceptedCmpOverall
## Income                     0.816879229        0.388247306
## Kidhome                   -0.539828043       -0.212079586
## Teenhome                  -0.146865727       -0.130254745
## Recency                    0.020241450       -0.013344023
## MntWines                   0.901848425        0.509912967
## MntFruits                  0.594180343        0.155132826
## MntMeatProducts            0.860663357        0.319553355
## MntFishProducts            0.620626400        0.174674581
## MntSweetProducts           0.595393554        0.200173529
## MntGoldProds               0.388775527        0.194646508
```

```
## NumDealsPurchases        -0.096409947        -0.126962322
## NumWebPurchases           0.503947206         0.195248498
## NumCatalogPurchases       0.778742425         0.366459493
## NumStorePurchases         0.668632189         0.201254206
## NumWebVisitsMonth        -0.498011053        -0.168914313
## AcceptedCmp3              0.034673402         0.431136978
## AcceptedCmp4              0.267288784         0.612100693
## AcceptedCmp5              0.478074879         0.719560141
## AcceptedCmp1              0.383947025         0.677609509
## AcceptedCmp2              0.139170987         0.460488886
## Complain                 -0.031430287        -0.020999798
## Z_CostContact                     NA                  NA
## Z_Revenue                         NA                  NA
## Response                  0.262362772         0.426960709
## Age                       0.117551703         0.001529105
## Customer_Days             0.141399733        -0.012213083
## marital_Divorced          0.002321233        -0.001839778
## marital_Married          -0.021605357         0.005995433
## marital_Single            0.006270999        -0.006112984
## marital_Together          0.002656024        -0.006118498
## marital_Widow             0.033282920         0.015537341
## education_2n.Cycle       -0.064468002        -0.021605392
## education_Basic          -0.138086926        -0.043835385
## education_Graduation      0.002595741         0.007809722
## education_Master          0.009650659        -0.012484312
## education_PhD             0.084810600         0.033255666
## MntTotal                  0.996569260         0.461279404
## MntRegularProds           1.000000000         0.461494686
## AcceptedCmpOverall        0.461494686         1.000000000
```

En examinant la matrice de corrélation fournie, voici quelques paires de variables avec des corrélations importantes (en valeur absolue) :

Incom et MntTotal (corr = 0,823) MntWines et MntTotal (corr = 0.9617) MntMeatProducts et MntTotal (corr = 0.9435) MntFruits et MntTotal (corr = 0.9288) MntSweetProducts et MntTotal (corr = 0.9079) MntFishProducts et MntTotal (corr = 0.8875)

Pour éviter des erreurs de modèles, nous allons omettre les variables suivantes : MntWines,MntMeatProducts, MntFruits,MntSweetProducts,MntFishProducts, MntRegularProds, AcceptedCmp1, AcceptedCmp2, AcceptedCmp3, AcceptedCmp4, AcceptedCmp5, Income.

De plus, certaines variables telles que Z_CostContact et Z_Revenue ne semblent pas contribuer à la variance des autres variables et pourraient être retirées.

```
# Nous utilisons la méthode de régularisation ridge pour aider à traiter les
dépendances linéaires.
library(glmnet)

## Warning: le package 'glmnet' a été compilé avec la version R 4.2.3
```

UONG Bellinna

```
## Le chargement a nécessité le package : Matrix

## Loaded glmnet 4.1-8

# Création de la matrice de conception X et du vecteur de réponse y
X <- model.matrix(AcceptedCmpOverall~Income + Kidhome + Teenhome + Recency +
                  MntWines + MntFruits + MntMeatProducts + MntSweetProducts
+
                  MntGoldProds + NumDealsPurchases + NumCatalogPurchases +
                  NumStorePurchases + NumWebVisitsMonth + Complain + Respon
se
                + Age + Customer_Days + marital_Divorced + marital_Married
+
                  marital_Single + marital_Together + marital_Widow +
                  education_2n.Cycle + education_Basic + education_Graduati
on +
                  education_Master + education_PhD, data = data_marketing)[
,-1]

y <- data_marketing$AcceptedCmpOverall

# Ajustement du modèle ridge
ridge_model <- cv.glmnet(X, y, alpha = 0)  # alpha = 0 pour la régression rid
ge
# Obtention du meilleur modèle
meilleur_model <- glmnet(X, y, alpha = 0, lambda = ridge_model$lambda.min)

# Affichage des coefficients
coef(meilleur_model)

## 28 x 1 sparse Matrix of class "dgCMatrix"
##                                s0
## (Intercept)          9.537217e-01
## Income               5.240863e-06
## Kidhome              2.745105e-02
## Teenhome            -5.649917e-02
## Recency              1.062665e-03
## MntWines             9.183677e-04
## MntFruits           -1.426082e-03
## MntMeatProducts     -1.740945e-04
## MntSweetProducts     7.283173e-05
## MntGoldProds         1.210976e-04
## NumDealsPurchases   -3.466732e-02
## NumCatalogPurchases  1.418888e-02
## NumStorePurchases   -2.054277e-02
## NumWebVisitsMonth    1.885507e-02
## Complain            -9.096639e-03
## Response             6.257670e-01
## Age                 -2.285600e-03
## Customer_Days       -4.425738e-04
```

```
## marital_Divorced       -4.033282e-02
## marital_Married         3.478218e-02
## marital_Single         -5.210948e-02
## marital_Together        2.245180e-02
## marital_Widow           1.692716e-03
## education_2n.Cycle      8.001846e-02
## education_Basic         1.888396e-01
## education_Graduation    4.781983e-02
## education_Master       -3.716857e-02
## education_PhD          -9.251309e-02
```

```r
# Création de modèle linéaire
modele_lm <- lm(AcceptedCmpOverall~Income + Kidhome + Teenhome + Recency +
                MntWines + MntFruits + MntMeatProducts + MntSweetProducts +
                MntGoldProds + NumDealsPurchases + NumCatalogPurchases +
                NumStorePurchases + NumWebVisitsMonth + Complain + Response
+
                Age + Customer_Days + marital_Divorced + marital_Married +
                marital_Single + marital_Together + marital_Widow + educati
on_2n.Cycle
                + education_Basic + education_Graduation + education_Master
                + education_PhD, data = data_marketing)
summary(modele_lm)
```

```
##
## Call:
## lm(formula = AcceptedCmpOverall ~ Income + Kidhome + Teenhome +
##     Recency + MntWines + MntFruits + MntMeatProducts + MntSweetProducts +
##     MntGoldProds + NumDealsPurchases + NumCatalogPurchases +
##     NumStorePurchases + NumWebVisitsMonth + Complain + Response +
##     Age + Customer_Days + marital_Divorced + marital_Married +
##     marital_Single + marital_Together + marital_Widow + education_2n.Cycle
+
##     education_Basic + education_Graduation + education_Master +
##     education_PhD, data = data_marketing)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7916 -0.2387 -0.0487  0.1414  3.4241
##
## Coefficients: (2 not defined because of singularities)
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        9.504e-01  1.832e-01   5.187 2.33e-07 ***
## Income             5.583e-06  1.221e-06   4.572 5.09e-06 ***
## Kidhome            3.604e-02  2.799e-02   1.287  0.19811
## Teenhome          -5.893e-02  2.665e-02  -2.211  0.02711 *
## Recency            1.193e-03  3.925e-04   3.039  0.00240 **
## MntWines           1.034e-03  5.861e-05  17.646  < 2e-16 ***
## MntFruits         -1.504e-03  3.783e-04  -3.977 7.21e-05 ***
## MntMeatProducts   -2.523e-04  8.867e-05  -2.845  0.00448 **
```

```
## MntSweetProducts        1.433e-04  3.669e-04   0.391  0.69612
## MntGoldProds            9.166e-05  2.576e-04   0.356  0.72199
## NumDealsPurchases      -3.662e-02  7.511e-03  -4.876 1.16e-06 ***
## NumCatalogPurchases     1.221e-02  6.914e-03   1.766  0.07759 .
## NumStorePurchases      -2.528e-02  5.367e-03  -4.711 2.62e-06 ***
## NumWebVisitsMonth       1.956e-02  7.426e-03   2.634  0.00849 **
## Complain               -2.135e-03  1.165e-01  -0.018  0.98539
## Response                6.566e-01  3.515e-02  18.679  < 2e-16 ***
## Age                    -2.421e-03  1.079e-03  -2.244  0.02493 *
## Customer_Days          -4.884e-04  6.284e-05  -7.773 1.18e-14 ***
## marital_Divorced       -4.841e-02  6.888e-02  -0.703  0.48224
## marital_Married         3.414e-02  6.296e-02   0.542  0.58767
## marital_Single         -5.799e-02  6.518e-02  -0.890  0.37375
## marital_Together        2.107e-02  6.400e-02   0.329  0.74196
## marital_Widow                  NA         NA      NA       NA
## education_2n.Cycle      2.023e-01  4.585e-02   4.412 1.07e-05 ***
## education_Basic         3.226e-01  7.889e-02   4.090 4.47e-05 ***
## education_Graduation    1.627e-01  3.000e-02   5.423 6.51e-08 ***
## education_Master        6.696e-02  3.636e-02   1.841  0.06569 .
## education_PhD                  NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5162 on 2179 degrees of freedom
## Multiple R-squared:  0.431,  Adjusted R-squared:  0.4244
## F-statistic: 66.01 on 25 and 2179 DF,  p-value: < 2.2e-16
```

La valeur de R^2 est 0,431, ce qui signifie que le modèle explique environ 43,1% de la variance totale dans la variable dépendante. Cela suggère que le modèle explique une proportion significative de la variabilité observée.

Le test F évalue la significativité globale du modèle. Avec une statistique F de 66,01 et une p-valeur très proche de zéro (p-value : < 2.2e-16), cela suggère que le modèle dans son ensemble est statistiquement significatif. En d'autres termes, au moins une des variables indépendantes a un effet significatif sur la variable dépendante.

L'analyse révèle des tendances significatives : chaque adolescent supplémentaire dans le ménage est associé à une diminution de 0,059 dans l'acceptation de la campagne marketing. De même, chaque achat de promotions est lié à une réduction de 0,0366 dans cette acceptain, et chaque achat en magasin entraîne une diminution de 0,02528. En revanche, les clients ayant accepté la dernière campagne présentent une augmentation significative de 0,6566 dans la variable dépendante. En termes d'éducation, des niveaux tels que 2nd Cycle, Basic et Graduation sont associés à des réponses plus élevées, soulignant l'influence significative de l'éducation sur les attitudes liées à la variable dépendante.

```
shapiro.test(data_marketing$AcceptedCmpOverall)

##
##  Shapiro-Wilk normality test
##
```

```
## data:  data_marketing$AcceptedCmpOverall
## W = 0.49952, p-value < 2.2e-16
```

```r
library(lmtest)
```

```
## Warning: le package 'lmtest' a été compilé avec la version R 4.2.3
```

```
## Le chargement a nécessité le package : zoo
```

```
##
## Attachement du package : 'zoo'
```

```
## Les objets suivants sont masqués depuis 'package:base':
##
##     as.Date, as.Date.numeric
```

```r
bptest(modele_lm)
```

```
##
##   studentized Breusch-Pagan test
##
## data:  modele_lm
## BP = 474.13, df = 25, p-value < 2.2e-16
```

```r
library(car)
```

```
## Le chargement a nécessité le package : carData
```

```r
durbinWatsonTest(modele_lm)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1      0.04190782      1.914008   0.042
##  Alternative hypothesis: rho != 0
```

Pour un seuil de 5%, nous pouvons voir que ces trois tests sont validés, ainsi les résidus suivent une distribution normale, une absence d'hétéroscédasticité, puis les erreurs ne sont pas corrélées dans le temps, renforçant la validité des estimations et des intervalles de prédiction.

Analyse des données

```r
library(ggplot2)
```

```
## Warning: le package 'ggplot2' a été compilé avec la version R 4.2.3
```

```r
# Diagramme en barres des réponses acceptées et non acceptées
ggplot(data_marketing, aes(x = AcceptedCmpOverall)) +
  geom_bar(fill = "grey") +
  labs(title = "Distribution des Réponses",
       x = "Nombre de réponse positive envers les campagnes publicitaires",
```

```
        y = "Fréquence") +
  theme_minimal()
```

Distribution des Réponses
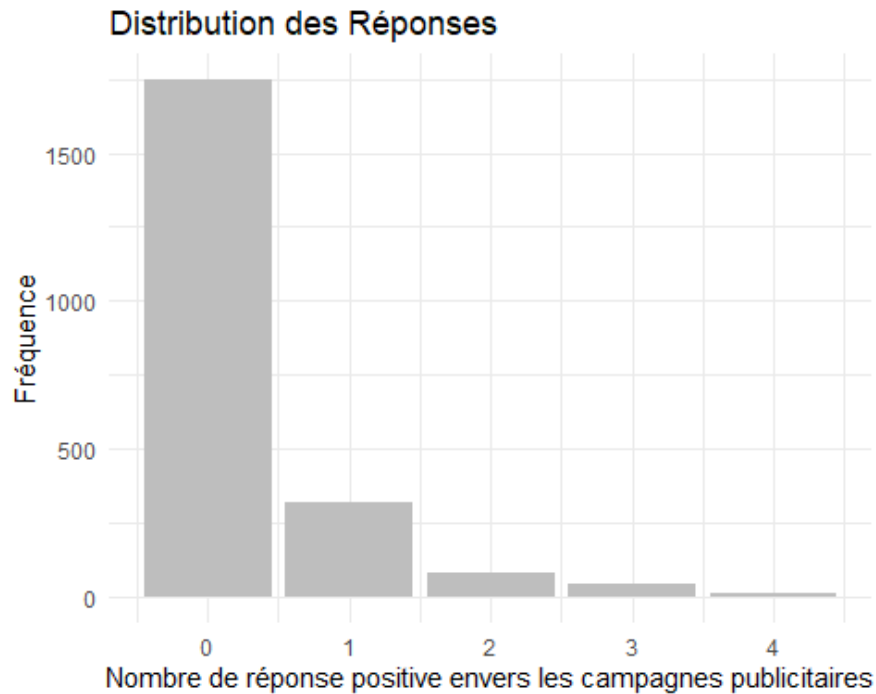
Sans aucun doute, la quasi-totalité des clients n'ont pas répondu de façon positives à aucun des campagnes publicitaires que nous avons mené.

Nous devrions nous concentrer sur quel critère nous devrions nous concentrer par rapport aux habitudes que suivent nos clients, et ainsi placer nos investissements de marketing correctement.

```
library(dplyr)

## Warning: le package 'dplyr' a été compilé avec la version R 4.2.3

##
## Attachement du package : 'dplyr'

## L'objet suivant est masqué depuis 'package:car':
##
##     recode

## Les objets suivants sont masqués depuis 'package:stats':
##
##     filter, lag

## Les objets suivants sont masqués depuis 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)

##
## Attachement du package : 'tidyr'

## Les objets suivants sont masqués depuis 'package:Matrix':
##
##      expand, pack, unpack

lieu<-data_marketing %>%
  select(c('NumWebPurchases','NumCatalogPurchases','NumStorePurchases','NumWe
bVisitsMonth'))
lieu_key_value <- pivot_longer(lieu, cols = everything(), names_to = "key", v
alues_to = "value")

ggplot(lieu_key_value, aes(value)) +
geom_histogram(fill = 'brown', binwidth = 1) +
facet_wrap(~key, scales = 'free_x', ncol = 2) +
labs(title = "Histograms of Place Variables") +
theme(plot.title = element_text(hjust = 0.5))
```


Histograms of Place Variables
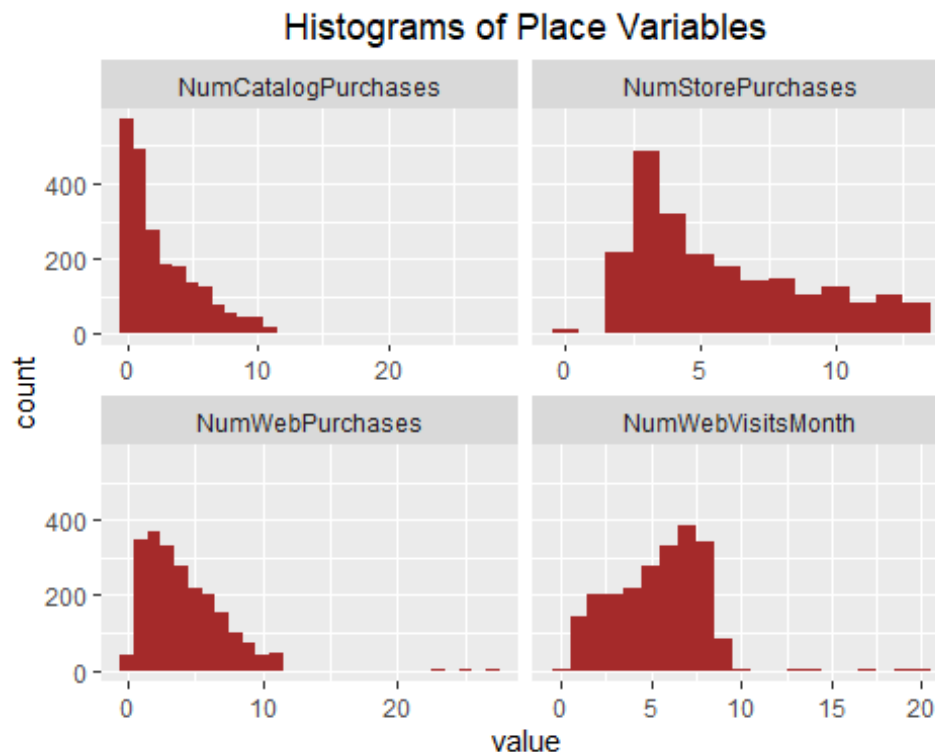
Nous pouvons voir que la quasi-totalité des clients vont leur course en magasin, mais dont beaucoup n'utilisent pas le catalogue. Nous pouvons améliorer ce cas, soit en revoyant la composition des catalogues, et ses distributions à plus de monde.

De plus, le site du magasin a une importance signifiante sur les achats. Nous pouvons penser à faire plus de marketing digital pour ainsi attirer plus de clients, surtout dans notre ère actuelle qui devient de plus en plus digital.

```r
Age_plot <- ggplot(data_marketing, aes(x = Age)) +
  geom_histogram(fill = 'dodgerblue4', binwidth = 5) +
  labs(title = "Histogramme des âges") +
  theme(plot.title = element_text(hjust = 0.5))

print(Age_plot)
```


Histogramme des âges
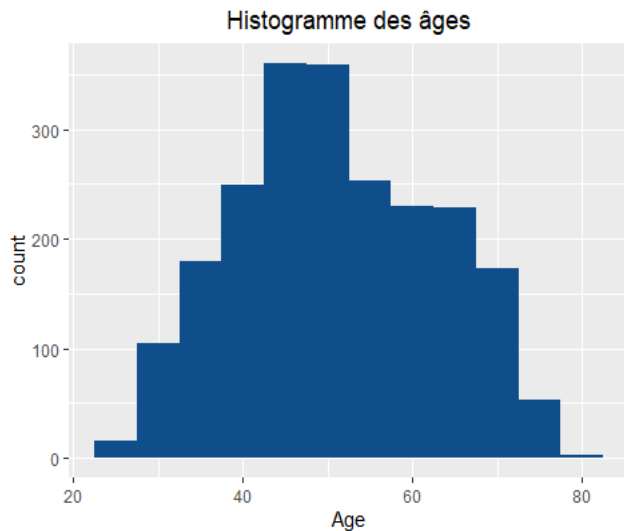
```r
library(ggplot2)

# Créer un nuage de points pour visualiser la corrélation entre l'âge et AcceptedCmpOverall
ggplot(modele_lm, aes(x = Age, y = AcceptedCmpOverall)) +
  geom_point() +  # Nuage de points
  labs(x = "Âge", y = "AcceptedCmpOverall") +
  ggtitle("Corrélation entre l'âge et AcceptedCmpOverall")
```

Corrélation entre l'âge et AcceptedCmpOverall
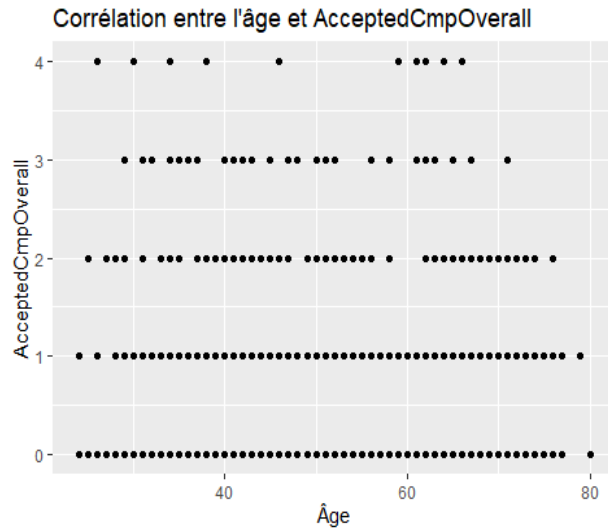
Nous voyons que la plus part des clients sont âgés entre 30 ans et 70 ans, nous pouvons identifiés cela à des adultes, des parents de famille, ou bien des grands-parents. On pourrait croire qu'il faudrait se concentrer alors sur ce point-ci pour orienter nos démarches marketing.

Cependant, la graphique de corrélation ne montre pas une corrélation directe de l'âge sur l'acceptation de campagne publicitaire. Donc au lieu de faire des publicitaires ciblés, comme nous aurions pu faire, il faut viser un public large, pour aussi toucher de nouveaux clients.