

PRÉVENIR L'ATTRITION DES CLIENTS CHEZ STREAM



Christine Fouque
Louis Hiel
Rodrigue Bernard
NaoData
USID17 STATISTIQUES 2
MASTER MEDAS

Prédire la résiliation des clients chez Stream

Le directeur commercial d'un portefeuille de cartes de crédit à la consommation est confronté au problème de l'attrition des clients.



Démarche :

- Analyse des données
- Découvrir la raison de ce phénomène
- Prédire les clients susceptibles de résilier
- Proposer des actions proactives pour prévenir le départ du client

Méthodologie

- Analyse exploratoire
 - Statistiques bivariées
- Segmentation
 - Statistiques multivariées
- Modélisation
 - Evaluation des modèles

Exploration des
données
(stat descriptives)

Recherche de
groupes
(clustering)

Evaluation des
modèles par
itération



Recherche de
Corrélations

Réalisation des
Modèles
(algorithmes prédictifs)



STREAM

01

ANALYSE EXPLORATOIRE DES DONNÉES

Nettoyage des données
Analyse des variables



- 23 colonnes
- Pas de valeurs aberrantes ou nulles
- Anonymisation des données

Clientnum	Numéro de client. Identifiant unique du client titulaire du compte
-----------	--

- La variable d'intérêt : Attrition_Flag

Variable d'intérêt	
Attrition_Flag	Existing Customer (client), Attrited Customer (compte fermé)

- Existing Customer (84%)
- Attrited Customer (16%)

⇒ Prédiction par régression logistique

- Les 19 variables explicatives :
 - 5 qualitatives
 - 14 quantitatives
- 2 types d'informations
 - 6 variables socio-démographiques
Genre, Âge, Statut marital...
 - 13 variables services bancaires
 - Relation avec la banque
Ancienneté, Échanges
 - Produit
Carte, Crédit accordé
 - Les transactions
Nombre, Montants & variations sur l'année
 - Activités du compte
Utilisation de la carte, mois sans activité
 - Gestion financière
Montant impayé, à payer



STREAM

VARIABLES QUALITATIVES

Selon le type de clients, 2 populations , pour **chaque** variable :

- analyse visuelle des profils
- tableau de fréquence relatives de chaque modalités



Carte	1	0
Blue	0,934	0,931
Silver	0,05	0,056
Gold	0,013	0,011
Platinum	0,003	0,002

Profils
similaires

La variable
influence peu le
départ

Catégorie de la carte

Profils
différents

La variable
influence l'attrition

*Genre
Niveau d'études
Catégorie de revenus
et le statut marital*



Genre	1	0
M	0,43	0,48
F	0,57	0,52

VARIABLES QUANTITATIVES

La variable influence t-elle le départ des clients?

1- Sur les 2 populations (clients/anciens clients) :

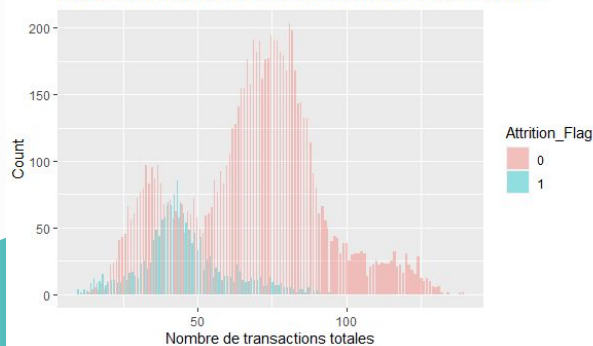
- Analyse visuelle
- Statistiques classiques

2- Comparaison des deux populations de clients

Tests statistiques de comparaison (moyenne ou médiane)

Les Populations sont différentes Les Populations sont semblables

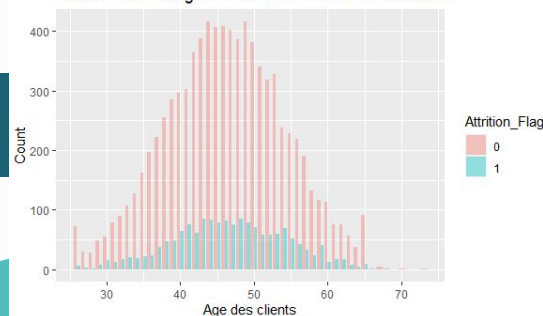
Distribution du nombre de transactions totales selon attrition



La variable influence
l'attrition

La variable n'influence
pas l'attrition

Distribution de l'âge des client en fonction de l'attrition

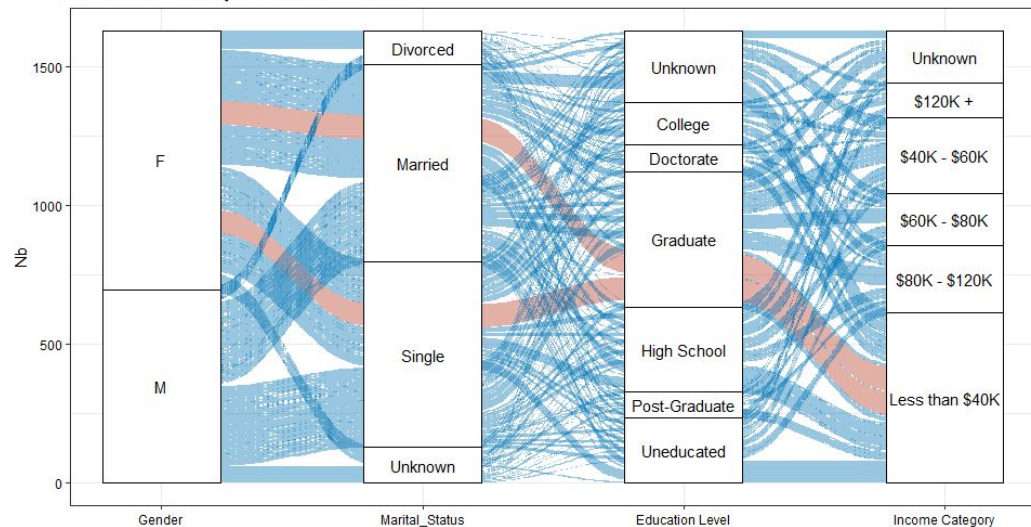


4 Variables qualitatives influentes :

- genre
- niveau d'étude
- le revenu
- niveau d'éducation

9 Variables quantitatives influentes:

- Activité du compte
 - utilisation de la carte
 - nombre de moins sans activité
- Les transactions :
 - leur nombres et la variation
 - les montants et la variation
- Nombre de produits
- Montant impayé (Revolving balance)
- Nombre d'échanges annuels avec la banque



Des pistes à explorer ...

02

RECHERCHE DE CORRÉLATIONS

- Corrélations
- Segmentation

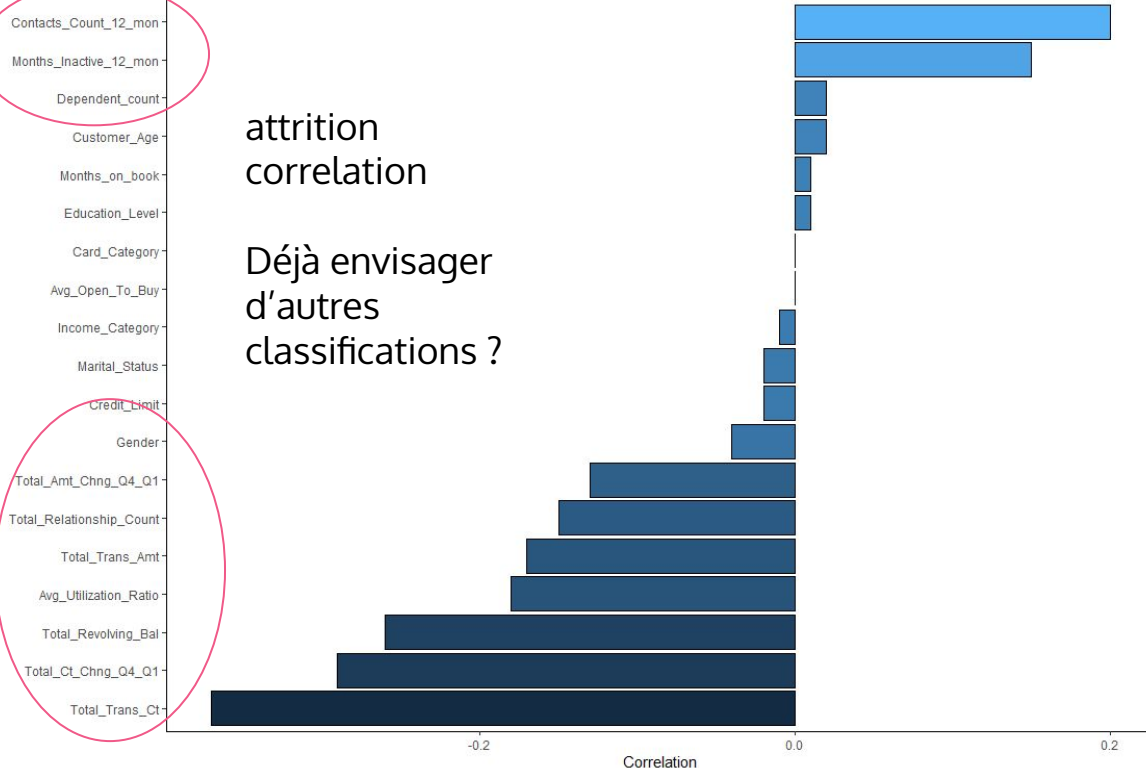




Attrition_Flag Correlation

attrition
correlation

Déjà envisager
d'autres
classifications ?

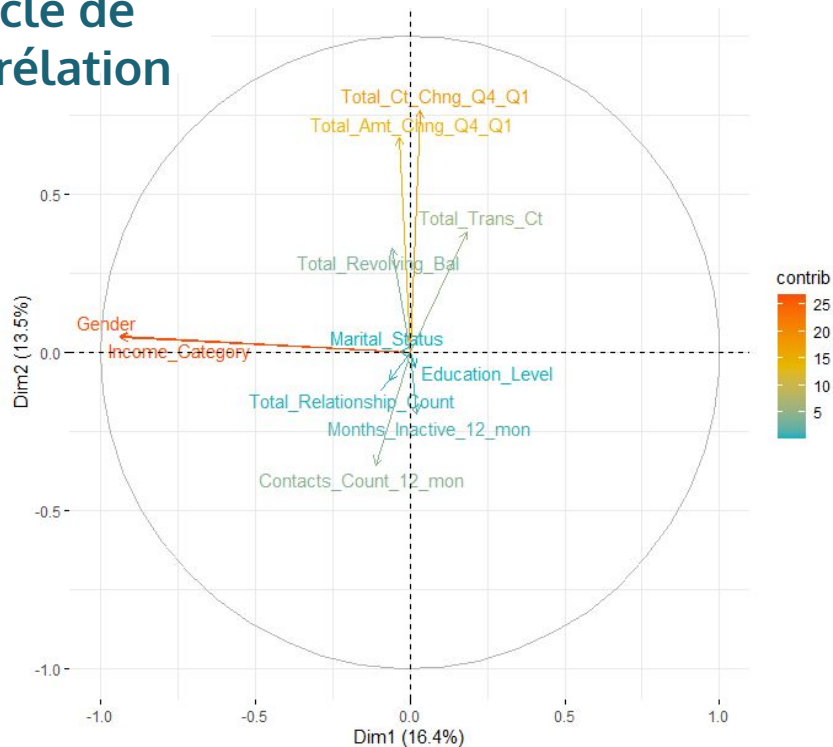


Identifier les variables explicatives
les plus corrélées avec la variable
d'intérêt, taux d'attrition

- **Corrélations positives :**
 - Activité bancaire : crédit consommé
 - Nombre et montant total de transactions, usage de la carte
 - Genre qui influe un peu
- **Corrélations négatives**
 - Echanges du client avec la banque
 - Inactivité du compte

Identifier des groupes d'observations ayant des caractéristiques similaires

Cercle de corrélation



- Genre et revenu bien représenté
- activité bancaire par trimestre
- une dizaine de facteurs influent sur le taux d'attrition

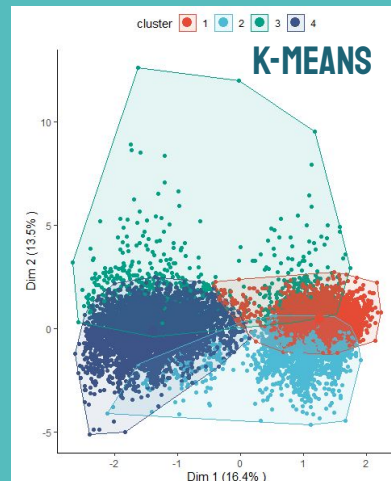
les 6 premières dimensions ne parviennent pas à expliquer plus de 70% de la variance ou masse totale d'inertie

Quels individus dans un même groupe se ressemblent ?

Quels individus dans des groupes différents se démarquent ?

Première itération en projetant les combinaisons des couples de dimensions sur un nuage de points, sur 4 clusters

DIM	VARIANCE	CUMUL
Dim1	16.4	16.4
Dim2	13.5	29.9
Dim3	11.8	41.8
Dim4	9.3	51.1
Dim5	9.1	60.2
Dim6	8.9	69.1



⇒ Les cumuls de variance ne nous permettent pas ici d'obtenir pour le moment une qualité globale explicative de l'analyse



STREAM

03

EVALUATION DES MODÈLES



Taux de vraisemblance = 996

Variables influentes :

- Nombre total des transactions (12 derniers mois)
- Montant total des transactions (12 derniers mois)
- Variation du nombre total des transactions (Q4 par rapport à Q1)
- Solde de crédit impayé
- Contact de la banque sur les 12 derniers mois
- Nombre total de produits détenus par le client
- Inactivité du client au cours des 12 derniers mois
- Le genre

ATTRITION	0	1
0	244	44
1	56	256

CRITÈRE	%
bien classé	83%
sensibilité	85%
spécificité	81%



Affinage de la classification des indicateurs

Taux de vraisemblance = 705

Variables influentes :

- Montant total des transactions (12 derniers mois) *inférieur à 3000 \$*
- Ratio créé: Montant total des transactions / Nombre total des transactions
- Nombre total des transactions (12 derniers mois) *très faible*
- Solde de crédit impayé *supérieur à 2200 \$ ou inférieur à 500 \$*
- Variation du nombre total des transactions (Q4 par rapport à Q1) *très faible*
- Variation du montant total des transactions (Q4 par rapport à Q1) *très faible ou très fort*
- Nombre total de produits détenus par le client
- Inactivité du client au cours des 12 derniers mois
- Contact de la banque sur les 12 derniers mois
- Les hommes

ATTRITION	0	1
0	265	31
1	35	269

CRITÈRE	%
bien classé	89%
sensibilité	90%
spécificité	88%



Taux de vraisemblance = 743

Restructuration :

- Le statut marital et le nombre de personnes à charge écartés
- Restructuration de la catégorie du revenu du client



EVALUATION DES MODÈLES



~90%
CATÉGORISÉ

10 % de faux positif

~ 90% de Sensibilité, Spécificité,
Taux de bien classé

SIMPLIFIÉ

GLOBAL



1



2



3



STREAM

CONCLUSION

Proposition de solution pour réduire l'attrition des clients chez STREAM :

- Optimisation des flux d'appels afin de cibler les clients avec une inactivité significative
- Affinage des campagnes de communication centrée sur une approche produit
- Enquête consommateur sur les raisons des crédits impayés...

NEXT STEP

- Prediction segmentée par produit (type de carte...)
- Analyse du type de produit par rapport au profil de consommation
- vers un outils de supervision en temps réel pour prioriser les appels aux clients ?



STREAM

NOTRE ÉQUIPE

MERCI !



CHRISTINE FOUQUE
cheffe de projet



LOUIS HIEL
data manager



RODRIGUE BERNARD
data analyst



Questions ?



NaoData