

# Analyse statistique du programme de fidélité

Identifier les facteurs différenciant les clients engagés et désengagés



# **SOMMAIRE**

## **1. Contexte**

1.1 Introduction

1.2 Objectifs

## **2. Méthodologie**

2.1 Description des données

2.2 Préparation et nettoyage des données

2.3 Méthodes statistiques utilisées

## **3. Résultats**

3.1 Test sur variables catégorielle (Test  $\chi^2$ )

3.2 Test sur variables numériques et catégorielles (Mann-Whitney U)

3.3 Corrélations (Test de Spearman)

## **4. Limites de l'étude**

## **5. Recommandations**

5.1 Actions pour les clients engagés

5.2 Actions pour les nouveaux clients

5.3 Actions pour anticiper le désengagement

## **6. Conclusion**

## 1. CONTEXTE

Canada Flight est une compagnie aérienne spécialisée dans les vols nationaux et internationaux, à destination et en provenance du Canada. Pour récompenser la fidélité de ses clients, la compagnie a mis en place un programme de fidélité qui permet aux voyageurs de cumuler des points à chaque vol. Ces points peuvent être ensuite utilisés pour obtenir des avantages divers comme des billets gratuits, des surclassements, et bien d'autres.

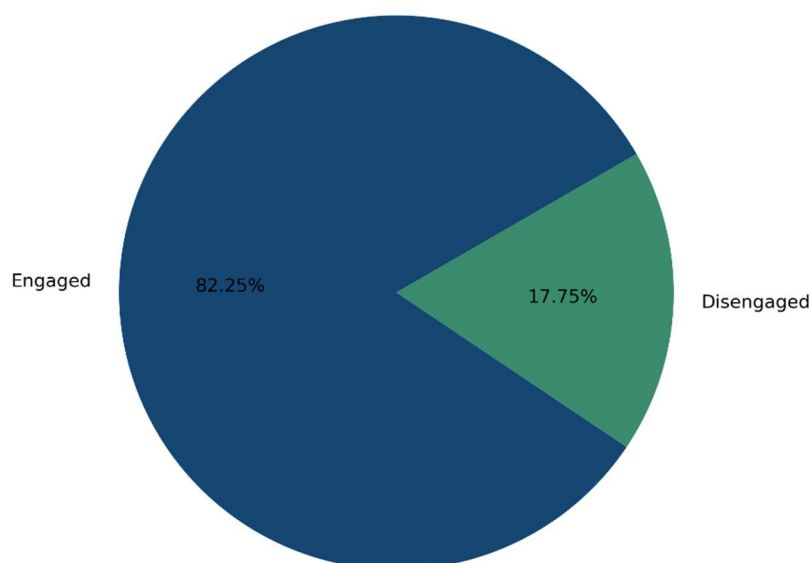
Dans une logique d'amélioration continue, Canada Fligth souhaite aujourd'hui développer davantage ce programme auprès de nouveaux adhérents. Cependant, l'entreprise a observé durant ces deux dernières années (2017-2018), que certains clients n'ont jamais effectué de réservation, tandis que d'autres ont effectué peu de vols.

Avant d'entreprendre toute campagne marketing, l'entreprise souhaite donc comprendre cette différence de comportement en analysant le profil des adhérents en fonction de leur niveau d'activité.

La base client a été segmentée en deux catégories d'activité :

- **Clients engagés** : plus de 20 vols
- **Clients désengagés** : clients inscrits ayant moins de 20 vols

Répartition des clients par statut d'engagement



L'objectif est d'identifier les facteurs clés qui différencient les clients actifs de ceux qui le sont moins, afin d'orienter efficacement les actions marketing à la fois pour :

- **promouvoir le programme** auprès de potentiels nouveaux adhérents
- **engager des actions concrètes** pour motiver les clients moins actifs et retenir les clients actifs

## Objectifs de l'analyse

Les objectifs principaux de cette étude sont :

- Comprendre le profil des adhérents** en analysant leurs comportements selon leur niveau d'activité (+ 20 vols vs - de 20 vols réalisés)
- Identifier les facteurs clés** qui différencient les clients engagés de ceux qui utilisent peu ou pas le programme
- Proposer des recommandations** pour aider l'entreprise à mieux cibler ses actions marketing et à favoriser une meilleure utilisation du programme de fidélité

## 2. MÉTHODOLOGIE

Pour cette étude, nous avons utilisé les données collectées par Canada flight en provenance de deux sources :

**La première table** contient les informations sociodémographique des clients (sexe, lieu de résidence, niveau d'études, statut marital, salaire).

**La deuxième table** contient des informations sur l'activité des clients regroupant des variables comme le nombre de vols, la distance total parcourue, les points cumulés ou échangés.

## 2.1 Nettoyage et préparation des données

Le nettoyage des données comprend la vérification des doublons et des types de données. De nombreux salaires manquants étaient présents, notamment pour les clients issus d'un niveau d'études "College" (BTS/Bac+2). Ces valeurs ont été imputées via la médiane des salaires moyens entre les clients d'un niveau "Bachelor" (license) et "High School" (lycée). De plus, de nouvelles variables ont été créées comme la variable ancienneté du client (calculée à partir de l'année d'inscription), ainsi qu'une variable de date obtenue par concaténation de l'année et du mois d'inscription.

Une partie porte sur la visualisation exploratoire, plusieurs graphiques ont été créés pour comprendre la diversité des profils et identifier les relations pertinentes à tester.

Avant d'appliquer les tests statistiques, une version du jeu de données sans valeurs extrêmes (outliers) a été construite afin d'évaluer la présence d'une différence significative de résultats entre les données avec et sans.

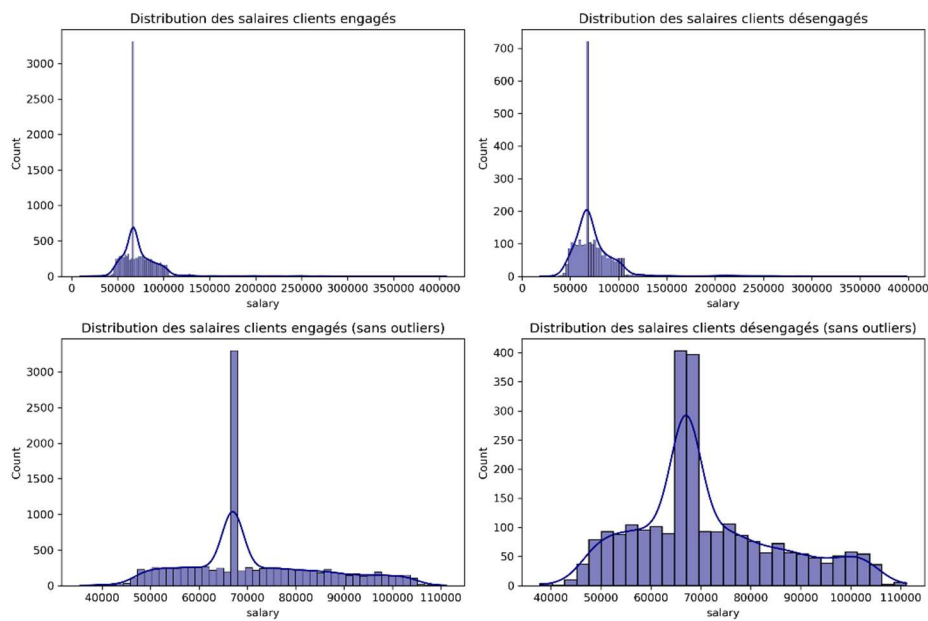
## 2.2 Test statistiques

Pour la partie analytique, plusieurs tests statistiques ont été réalisés :

**Le test de chi-carré** a été utilisé pour vérifier l'association entre les critères sociodémographiques des clients et les statuts d'engagement (Engagés vs désengagés). Ce test est adapté pour examiner les associations entre variables catégoriques dont les effectifs sont supérieurs à 5 pour chaque croisement de variable.

L'objectif était d'identifier si ces facteurs sociodémographiques avaient une influence sur le niveau d'activité des clients. Ce test permet d'identifier les caractéristiques des clients engagés, permettant ainsi à la compagnie de cibler de potentiels adhérents engagés.

**Le test de Mann-Whitney U** a été réalisé pour observer la différence de salaire, du nombre total de vols, de l'ancienneté et de la valeur vie client entre les clients engagés et désengagés. Ce test est adapté pour comparer des variables quantitatives non normalement distribuées entre deux groupes indépendants et évaluer s'il existe une différence statistiquement significative.



*les salaires ne suivent pas une distribution normale. Confirmé par les tests de Kolmogorov-Smirnov et Shapiro pour les deux distributions avec et sans valeurs extrêmes*

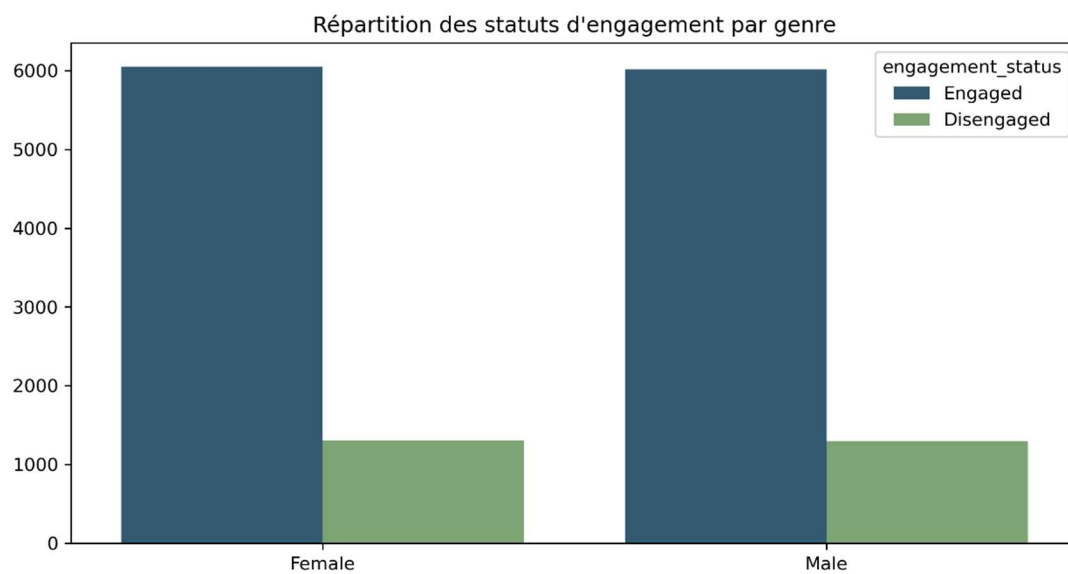
L'objectif était de déterminer si le salaire, le nombre total de vols, l'ancienneté ou la valeur vie client varient selon le niveau d'activité, ce qui permettrait à la compagnie de cibler les profils à forte valeur ajoutée.

**2.3 Des analyses complémentaires** ont été réalisées afin d'analyser la corrélation entre variables quantitatives, pour une vue plus globale (Test de Spearman).

### 3. RÉSULTATS

#### 3.1 Test Chi-carré

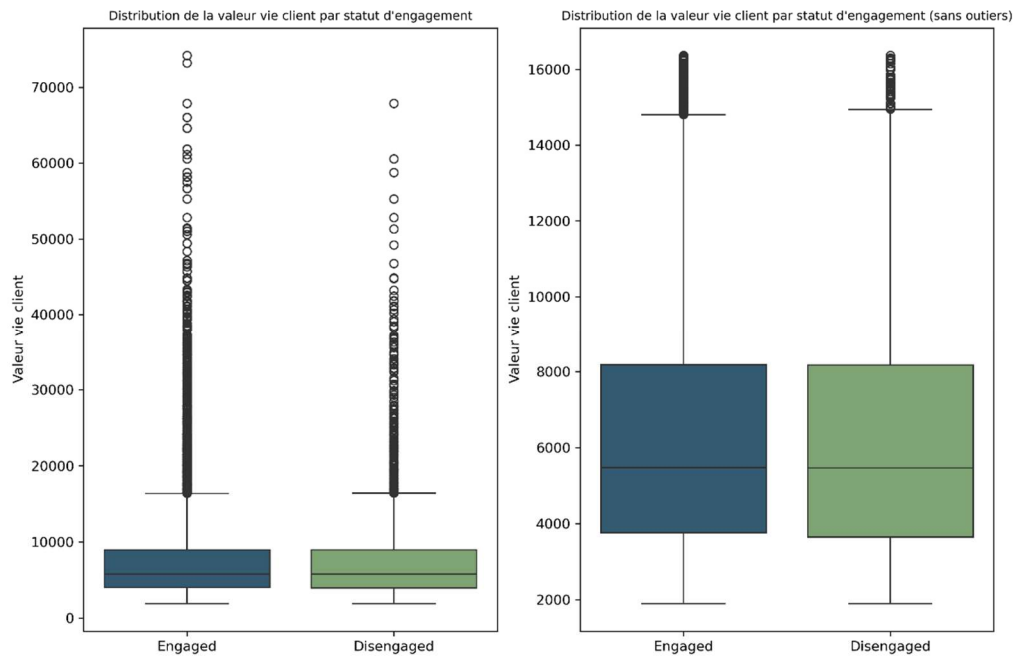
Les test de Chi-carré ont démontré que les facteurs sociodémographiques (genre, statut marital, niveau d'études et lieu de résidence) n'ont aucune influence sur le statut d'engagement. Tous les résultats ont une p-valeur supérieure au seuil de 0.05, ce qui signifie qu'il n'y pas de différence statistiquement significative entre ces variables et le statut d'engagement.



*Aucune différence significative entre les hommes et les femmes selon les statuts d'engagements*

### 3.2 Test de Man-Whitney U

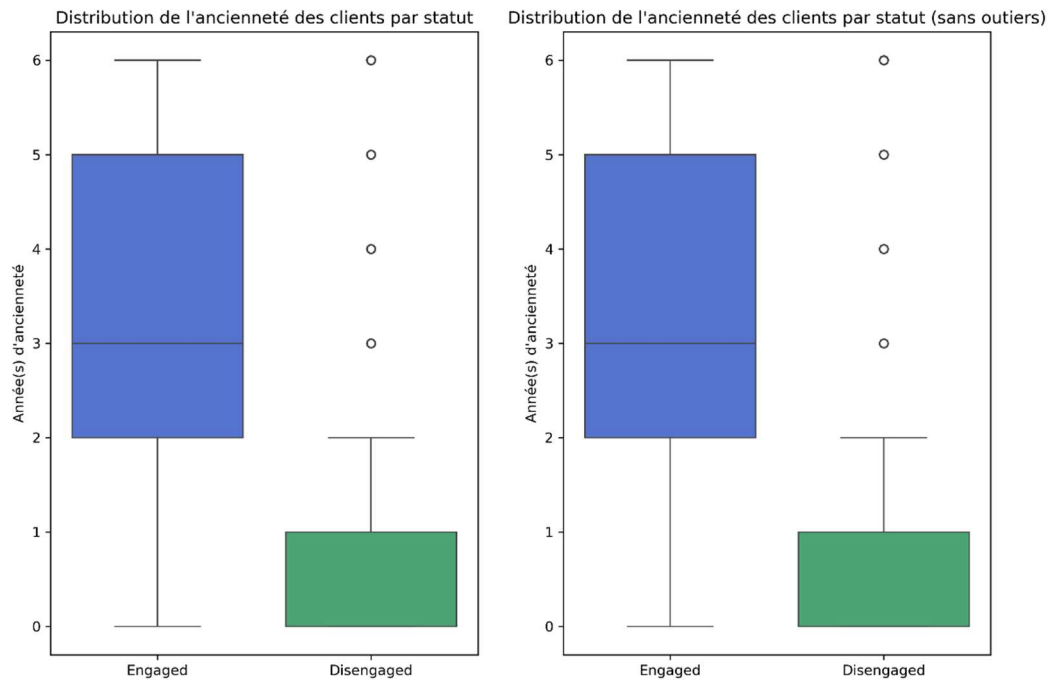
Le test de Mann-Whitney U appliqués sur la distribution des salaires et la valeur vie client (CLV) entre les deux groupes, ne présentent pas non plus de différence statistiquement significative entre les clients engagés et désengagés.



*Il n'y a pas de différence significative de la valeur vie client entre les deux groupes*



En revanche, le test de Mann-Whitey U appliqués sur le nombre de vols effectués et l'ancienneté dans le programme, révèlent une différence significative entre les deux groupes, avec une p-valeur inférieure au seuil de 0.05. Cela suggère que les clients engagés volent davantage et son présents dans le programme depuis plus longtemps.



*Les clients engagés sont majoritairement plus anciens dans le programme que les clients désengagés*

### 3.3 Test de Spearman

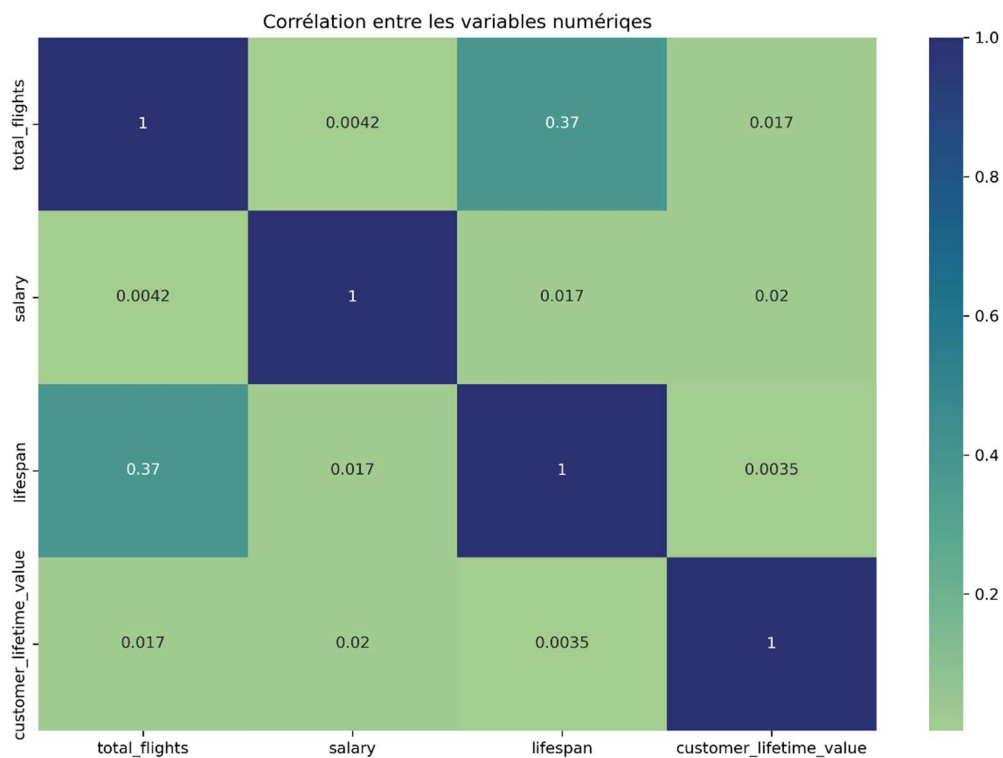
Les tests de corrélation de Spearman ont été réalisés pour observer les relations entre certaines variables quantitatives, pour une analyse plus globale et hors comparaison de groupes). Bien que le test de Spearman est plus robuste face aux valeurs extrêmes, les tests ont été effectués sur les données sans outliers pour limiter leur influence.

**Les tests ont montrés une absence de corrélation** (P-value > 0.05) entre :

- Le salaire et le nombre total de vols effectué
- La valeur vie client et l'ancienneté

En revanche, **les tests ont montré une corrélation positive faible** malgré la significativité (P-valeur < 0.05) entre :

- Le nombre total de vols effectués et la valeur vie client (Coefficient de Spearman : 0.01) : plus un client réalise de vols, plus il contribue à la rentabilité
- Le nombre total de vols effectués et l'ancienneté (Coefficient de Spearman : 0.35), plus le nombre de vols réalisé est élevé, plus le client possède de l'ancienneté dans le programme
- Le salaire et la valeur vie client (Coefficient de Spearman : 0.02) : plus le salaire d'un client est élevé, plus il génère de la valeur



## 4. LIMITES DE L'ÉTUDE

Cette étude statistique présente plusieurs limites qui pourraient influencer l'interprétation des résultats.

Premièrement, certaines **valeurs manquantes ont été imputées** (notamment les salaires), ce qui ne reflète pas les données réelles. L'imputation via la médiane permet de limiter l'impact des valeurs extrêmes, mais peuvent également introduire un biais dans les résultats.

D'autre part, **l'absence de variables clés comme l'âge ou le montant total dépensé** par client limite la capacité à affiner les profils et à évaluer la valeur réelle des clients. Ces informations auraient été utiles pour enrichir l'analyse comportementale et affiner davantage la segmentation client.

Enfin, **les tests statistiques utilisés dans cette étude sont principalement non paramétriques** (comme le test de Mann-Whitney U ou le test de Spearman), car les données ne suivaient pas une condition normale. Bien que ces tests soient adaptés aux données disponibles, ils sont généralement moins puissants que les tests paramétriques, ce qui signifie que certaines relations plus fines ou effets faibles ont pu ne pas être détectés.

## 5. CONCLUSION

L'étude statistique met en évidence que les clients engagés sont généralement ceux ayant plus d'ancienneté et ayant effectués le plus de vols. À l'inverse, les clients qualifiés de désengagés sont majoritairement de nouveaux adhérents (moins de 2 ans d'ancienneté), ce qui suggère qu'ils ne sont peut-être pas encore familiers avec le programme, plutôt que réellement désengagés.

## **6. RECOMMANDATIONS**

D'après les résultats de cette étude, les actions à mettre en place doivent s'adresser à la fois aux clients engagés, qu'il est essentiel de fidéliser, et aux nouveaux inscrits, qui peuvent devenir des clients rentables à long terme s'ils sont bien accompagnés.

### **1. Cibler les clients à fort potentiel déjà engagés**

Il est essentiel pour l'entreprise de concentrer ses efforts sur les clients ayant plus de 20 vols à leur actif et ceux ayant une ancienneté supérieure à 2 ans, qui n'ont pas encore adhéré au programme. Ces profils montrent une forte activité avec la compagnie et représentent une cible idéale pour rejoindre le programme.

### **2. Accompagner les nouveaux clients récents**

Les clients ayant moins de 2 ans d'ancienneté ne sont pas nécessairement désengagés, mais simplement en phase de découverte. Ils peuvent devenir des clients fidèles s'ils sont bien orientés dès le départ. Il serait pertinent d'identifier ces clients afin de leur envoyer des mails de présentation, leur expliquant le fonctionnement du programme, les bénéfices concrets qu'ils peuvent en tirer et les avantages exclusifs disponibles.

### **3. Retenir les adhérents engagés et relancer les profils récents**

L'entreprise peut également relancer les membres peu actifs, ou les récents inscrits au programme via des offres promotionnelles ciblées, des rappels personnalisés sur les avantages ou les nouveautés du programme, ainsi qu'une communication régulière pour maintenir le programme dans leur esprit.

### **4. Comprendre les comportements de désengagement**

Il pourrait être intéressant pour l'entreprise de comparer les profils engagés avec ceux des clients ayant abandonnés le programme, afin d'identifier les signaux de désengagement en amont et adapter les campagnes de fidélisation pour prolonger l'engagement des adhérents.

## **5. Une compréhension des attentes des adhérents**

Réaliser des enquêtes auprès des adhérents afin de mieux comprendre les préférences (destination préférée, budget disponible, services préférés, etc.) , ce qui permettrait d'adapter les offres, les communications et de personnaliser davantage l'expérience utilisateur.