# Solution cas *ProFitness* (Thème 1)

## Mise en contexte

Un homme d’affaire de renommé vient d’acquérir le centre sportif **« ProFitness » situé au centre-ville de Sherbrooke.**

La seule préoccupation du nouveau propriétaire est de pouvoir préserver la clientèle du centre. Il croit fortement qu’acquérir de nouveaux clients peut être plus coûteux que d’entretenir ceux qui sont déjà abonnés.

Quelques semaines après l’acquisition, le nouveau propriétaire a envoyé un questionnaire aux 1833 membres du club. (Les réponses au questionnaire se trouvent dans le fichier « ***Pro\_Fitness\_données.xlsx*** »

Le sondage contient 8 questions pour évaluer les variables suivantes:

1. **Satisfaction vis-à-vis des équipements**
2. **Satisfaction vis-à-vis du personnel**
3. **Satisfaction vis-à-vis des programmes offerts**
4. **Niveau de satisfaction global**

Le niveau de satisfaction a été évalué sur une échelle ordinale codée de 1 à 5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 |
| Pas satisfait du tout | **Insatisfait** | **Neutre** | **Satisfait** | **Très satisfait** |

1. **Nombre d'années d'adhésion**
2. **Sexe (1 = Homme; 2= Femme)**
3. **Nombre de visites par semaine**
4. **Âge en années**
5. **Profession**

## **Partie 2 : Réponses aux Questions**

À l’aide d’Excel, tentons de répondre aux questions suivantes :

1. **Calculer le taux de réponse au sondage**.

Le nombre de répondant au sondage = 1214

Le nombre de questionnaires envoyés = 1833

Taux de réponse=1214/1833=66%

Il y a environ 2 personnes sur trois qui ont répondu au sondage.

Le taux de réponse n’est pas parfait, nos conclusions porteront donc sur les répondants et non tous les abonnés (cette section fera l’objet du cours: Statistique)

1. **Identifier le type de chaque variable dans la base de données.**

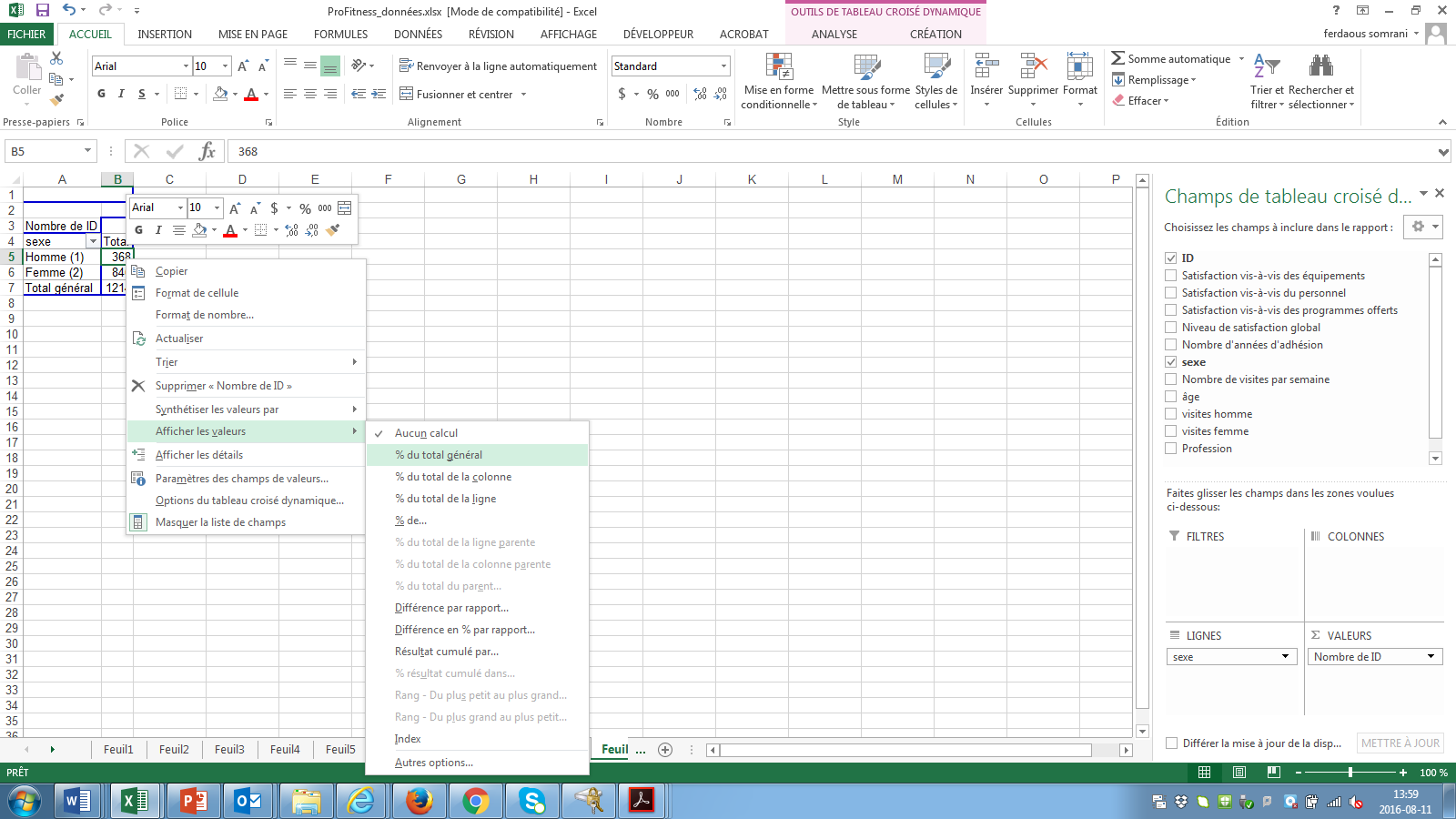
On dispose de 9 variables :

1. **Satisfaction vis-à-vis des équipements : Ordinale**
2. **Satisfaction vis-à-vis du personnel : Ordinale**
3. **Satisfaction vis-à-vis des programmes offerts : Ordinale**
4. **Niveau de satisfaction global : Ordinale**
5. **Nombre d'années d'adhésion : Discrète**
6. **Sexe (1 = Homme; 2= Femme) : Nominale**
7. **Nombre de visites par semaine : Discrète**
8. **Âge en années : Discrète (mais peut être considérée comme continue.**
9. **Profession** (1=Étudiant; 2=Professionnel; 3=Cadre; 4=Retraité; 5=Chômeur) : **Nominale**

1. **La clientèle du centre** (on parle ici des répondants au sondage) **est-elle majoritairement masculine ?**

Il s’agit ici de retrouver le % homme /% femmes (répondants).

On se sert ici du TCD pour retrouver la répartition des répondants Homme/femme (69,69% Femmes vs 30.31% Hommes).



1. **Quelle est la proportion de Femmes globalement très satisfaites des services offerts par le centre?**





**2.25%** (19 femmes parmi 846) des clientes sont globalement très satisfaites des services offerts.

1. **Commenter le profil des répondants (âge et profession)**

**Répartition de la variable Profession**

Les répondants sont majoritairement cadres ou professionnels (531+173)/1214= 58%.

Les étudiants représentent également une proportion considérable des répondants, soit, (36%).

Attention : On voit apparaitre une catégorie 33 (voir codification de la variable profession). Il existe une valeur aberrante (erronée dans ce cas-ci). Il peut s’agir d’une mauvaise transcription des données (ID= 816).

Les traitements possibles : revenir aux questionnaires pour rectifier le tir, imputer la valeur la plus probable (ici c’est 3) ou éliminer l’observation des données (cette pratique peut être dangereuse dans le cas où les valeurs manquantes traduisent un pattern dans les données).

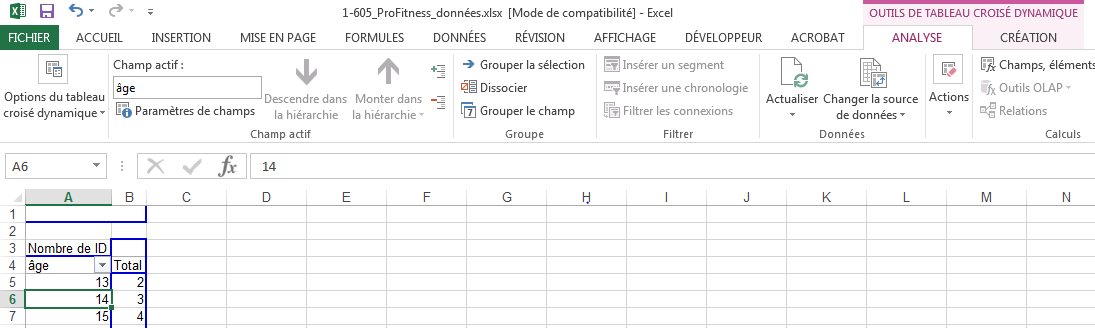
|  |  |
| --- | --- |
|  |  |
| Nombre de ID |  |
| Profession | Total |
| 1-Étudiant | 440 |
| 2-Professionnel | 531 |
| 3-Cadre | 173 |
| 4-Retraité | 28 |
| 5-Chômeur | 41 |
| **33** | 1 |
| Total général | 1214 |

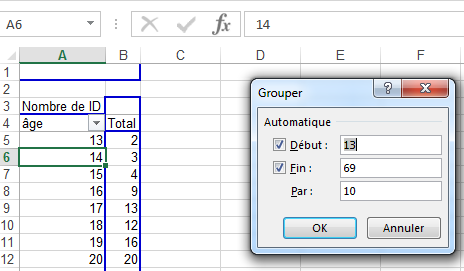
****

**Âge**

Plusieurs mesures peuvent décrire la variable âge. On peut parler d’âge moyen, âge médian, min, max… (on va revenir à ces mesures à la séance 2)

On peut aussi regrouper l’âge par tranche de 10 pour décrire la distribution de l’âge.





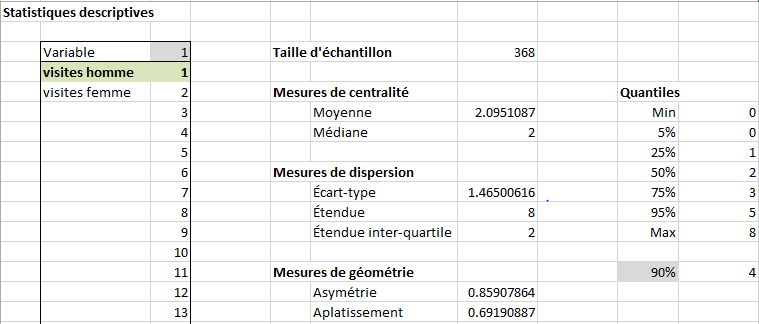


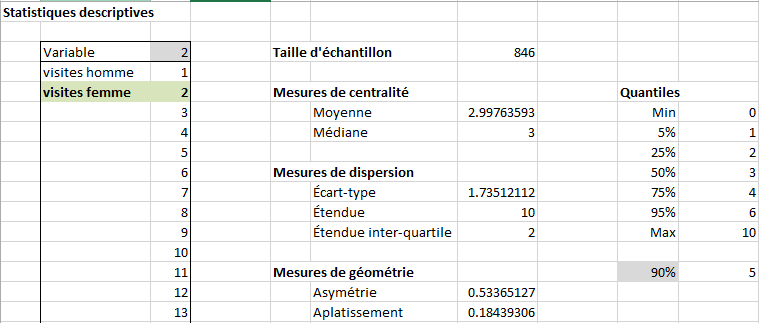
La clientèle est relativement jeune : près de 92% des clients sont âgés de moins de 52 ans.

## **Partie 2 : réponses aux questions**

On s’intéresse à la comparaison du nombre de visites des hommes et des femmes.

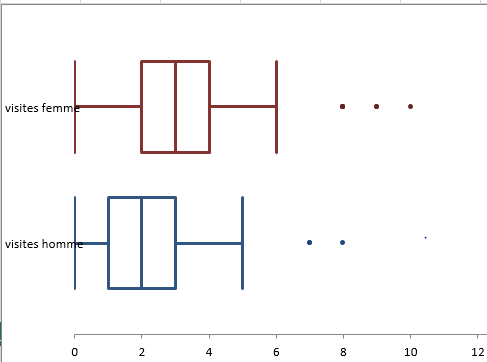
**Question 1 :** Calculer la moyenne, la médiane, le coefficient d’asymétrie, l’étendue, l’écart-type et l’IQR pour chaque variable.

****

****

**Question 2 :** Comparer la distribution de la variable nombre de visites chez les Hommes et chez les Femmes et commenter (utilisez les Boxplots).

On peut comparer le nombre de visites des hommes et des femmes à l’aide des boxplots des 2 groupes.

****

**Mesures de tendance centrale :**

Les femmes visitent le centre en moyenne une fois de plus que les hommes. D’autre part, le nombre de visites médian des femmes est de 3 alors que celui des hommes est de 1. Cela signifie, qu’au plus 50% des femmes visitent 3 fois et plus le centre sportif et qu’au plus 50% des hommes y vont plus d’une fois.

**Mesures de dispersion :**

L’étendue des visites des femmes est supérieure de 2 visites à celui des hommes. Chez les Hommes on observe un écart de 8 visites entre le nombre de visites le plus élevé et le nombre de visites le plus bas. Ce même écart s’élève à 10 chez les femmes. Finalement, l'écart type du nombre de visites des femmes est légèrement supérieur à celui des hommes. Cela signifie que la dispersion moyenne des visites des femmes autour du nombre de visites moyen est légèrement supérieure à la dispersion moyenne des visites moyenne des hommes.

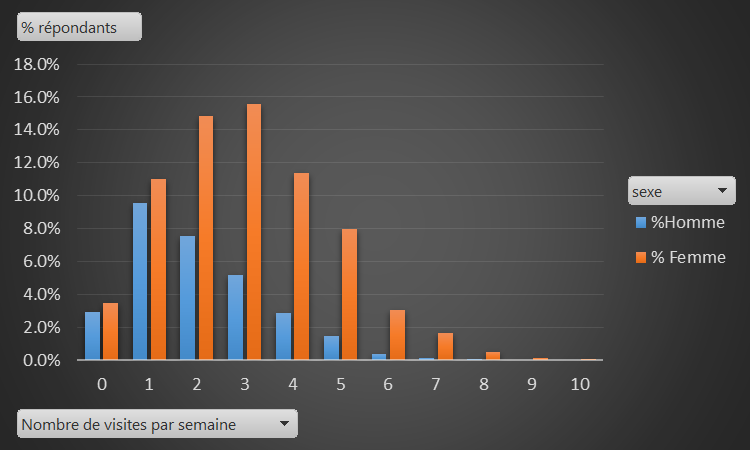
**Mesures de géométrie :**

Le coefficient d’asymétrie des visites des hommes est légèrement supérieur à 0. Cela signifie que la distribution des visites des hommes est légèrement asymétrique à droite. On observe la même chose chez les femmes (moins prononcé).

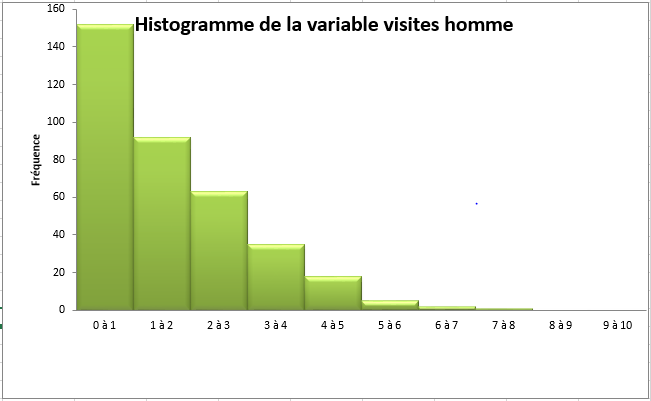
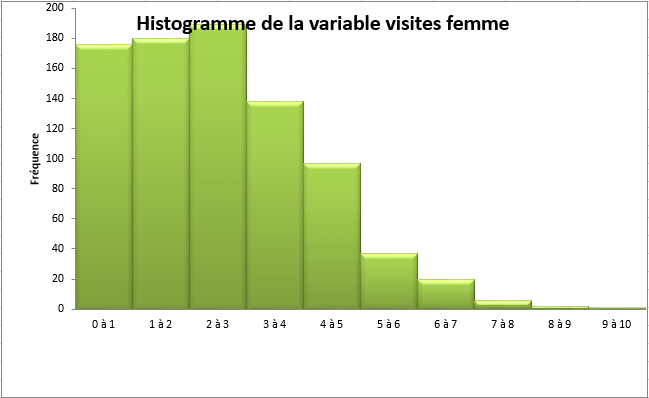
**Mesures de position :**

* Le plus petit nombre de visites observé chez les Hommes et les femmes est de 0. le client qui visite 10 fois par semaine le centre sportif s’avère une femme.
* Le nombre de visites le plus élevé chez les hommes est de 8 visites par semaine.
* Comme pour les mesures de tendance centrale, les quantiles des visites des femmes sont supérieurs aux quantiles des visites des hommes d’une visite. Par exemple, le premier quartile des visites des hommes est égal à 1 alors que celui des femmes est égal à 2. Cela signifie qu’au plus 75% des hommes vont plus d’une fois au centre sportif et au plus 25% des filles y vont moins de deux fois.

Comme la variable Nombre de visites est discrète et comporte peu de modalités un diagramme à bâtons mettant en lumière le % d’hommes et de femmes effectuant un nombre de visites en particulier peut être construit.



**Question3 :** Tracer les histogrammes pour chaque variable et commenter.

****

* Dans les 2 cas, on note une asymétrie à droite.
* L’intervalle modale est de (2-3 visites) chez les femmes alors qu’il est de (0 à 1) visite chez les hommes.
* On note une plus grande dispersion chez les femmes.

**Question 4** : Le nouveau propriétaire désire augmenter le nombre de visites au centre sportif. Quelles recommandations pouvez-vous lui suggérer à la lumière de ces résultats? Utilisez les variables disponibles dans la base de données?

Si l’échantillon est représentatif (on en fait l’hypothèse ici), on peut présumer que la clientèle est majoritairement féminine et que la grande majorité des membres sont des cardes et professionnels.

On peut proposer :

1. des accommodements des horaires d’ouverture et de fermeture (le cas échéant). Faire un test sur une période bien précise pour ne pas supporter des coûts de fonctionnement élevé sans contrepartie.
2. Faire des ententes avec les entreprises (à proximité du centre) pour attirer plus de clientèle.
3. Faire des offres promotionnelles pour couple Homme/femme ou ami(e)s…

**Question 5 :** Le nouveau propriétaire désire améliorer le niveau de satisfaction global. Fiez-vous aux données pour lui communiquer des pistes d’amélioration.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Niveau de satisfaction | Satisfaction globale | Équipements | Personnel | Programmes offerts |
| Pas du tout satisfait | 0% | 3% | 3% | 13% |
| Insatisfait | 7% | 11% | 8% | 21% |
| Neutre | 48% | 31% | 12% | 41% |
| Satisfait | 44% | **46%** | **48%** | **20%** |
| Très satisfait | 2% | **9%** | **29%** | **5%** |
| Total | 100% | 100% | 100% | 100% |

* Le personnel est un atout important pour le centre.
* Le niveau d’insatisfaction concerne principalement les programmes offerts suivi des équipements.
* Pour pallier à cela, le propriétaire doit mettre vers l’avant une gamme de programme attrayante et renouveler le parc sportif (si le coût le permet).