

TP1 - Naufrage du Titanic (tableur) Thème 4 : Données structurées

CORRECTION

Le numérique a permis de collecter, d'agréger et d'analyser de grandes quantités d'informations. Alors que notre monde est de plus en plus empli de données, il est nécessaire d'avoir la capacité d'extraire des informations de ces données et de les analyser de manière critique. Voici un exemple d'utilisation de logiciels pour étudier des données.

I. Le format CSV (comma-separated values)

Le RMS Titanic est un paquebot transatlantique britannique qui fait naufrage dans l'océan Atlantique Nord en 1912 à la suite d'une collision avec un iceberg, lors de son voyage inaugural de Southampton à New York. C'est l'une des plus grandes catastrophes maritimes survenues en temps de paix et la plus grande pour l'époque.



Afin d'étudier cet événement historique, on utilise le fichier « titanic.csv » qui contient des données sur les personnes (hors membres d'équipage) qui étaient présentes sur ce bateau lors de cette traversée.

1. Ouvrir le fichier « titanic.csv » à l'aide du logiciel Bloc-Notes (clic droit, « Ouvrir avec », « Bloc-Notes »). Quels sont les séparateurs qui ont été utilisés pour séparer les données de ce fichier CSV ?

Les séparateurs utilisés sont des points-virgules.

2. Quels sont les descripteurs de ce fichier ?

Classe, survie, nom, sexe, âge et tarif.

3. Fermer le fichier sans le modifier. Pourquoi ne pas utiliser Bloc-Notes pour étudier ces données ?

Car il n'y a aucune fonction qui permet de faire des opérations sur les données !

II. Traitement des données à l'aide d'un tableur

1. Pour étudier ces données, on va utiliser un tableur. Ouvrir le logiciel Excel (taper « Excel » dans le menu « Démarrer »). Dans le menu « Données », cliquer sur « Données externes » puis « Fichier texte » (ou directement sur « Fichier texte » selon la version d'Excel) et sélectionner le fichier « titanic.csv ».

2. Dans la fenêtre qui s'ouvre, cocher « Délimité » puis « Suivant ».

3. Pourquoi choisir de cocher « Point-virgule » dans la fenêtre ? Cocher cette option et cliquer sur « Suivant ».

Car le séparateur utilisé pour le fichier est le « point-virgule ».

4. Cliquer sur « Terminer » puis sur « OK ». Combien y avait-il de passagers à bord du Titanic (hors membres d'équipage) lors de cette traversée ?

Les objets vont de la ligne 2 à la ligne 1310 donc il y avait 1309 personnes à bord (hors membres d'équipage).

5. L'objet situé à la ligne 272 du tableur (si le fichier n'a pas été réorganisé) est Mme Smith qui était passagère de 1ère classe, âgée de 18 ans, qui avait payé son voyage 60€ et qui a survécu au naufrage. Compléter les phrases suivantes :

Pour le descripteur « sexe », la valeur 1 correspond à un homme et la valeur 2 correspond à une femme

Pour le descripteur « Survie », la valeur 1 correspond à la survie et la valeur 0 correspond à un décès.

6. Dans la cellule H1, saisir « =MOYENNE(E2:E2000) ». Interpréter le résultat.


On obtient la valeur 29,89770554. Cela signifie que la moyenne d'âge à bord du Titanic était environ de 30 ans.

7. Déterminer le tarif moyen payé par les passagers puis le tarif le plus élevé qui a été payé (pour cela, on utilisera la fonction MAX).

Le tarif moyen est environ de 33,36€. Le tarif le plus élevé est de 512€.

8. Dans une cellule vierge, saisir « =QUARTILE(F2:F2000 ;3)-QUARTILE(F2:F2000 ;1) ». Qu'obtient-on et comment interpréter ce résultat ?

Le nombre obtenu est 23€. Il s'agit de l'écart interquartile donc 50% des passagers ont payé un tarif qui diffère de moins de 23€.

9. Cliquer sur le B de la colonne B afin de sélectionner toute la colonne, puis, dans la partie « Données », cliquer sur l'icône  afin de trier les données par ordre décroissant selon le descripteur « survie ». Combien de personnes ont-elles survécu au naufrage du Titanic ? En déduire la proportion de personnes qui ont survécu.

Les personnes ayant survécu (c'est-à-dire celles ayant un descripteur « survie » égal à 1) vont de la ligne 2 à la ligne 501 ce qui donne 500 personnes.

500/1309 ≈ 0,38 donc il y a 38% des gens seulement qui ont survécu à ce naufrage !