

STID 1 - Programmation Statistique

TD3

Importer des données

Anthony SARDELLITTI

2023-01-01

Contents

1	Ressources documentaires	1
2	Exercice 1 : Importer les données	1
3	Exercice 2 : Compiler les données	2
4	Exercice 3 : Analyses statistiques	2

1 Ressources documentaires

Pour réaliser ce TD, vous aurez besoin des ressources suivantes :

- [Importer un fichier csv](#)
- [Modifier le répertoire par défaut](#)
- [Filtres et sélection](#)
- [Trier](#)
- [Agréger](#)
- [Les types de graphique](#)

2 Exercice 1 : Importer les données

Vous utiliserez dans ce TD les résultats de quiz Kahoot 2022. Ils ont été anonymisés. Il y a un fichier csv par quiz Attention, les questions ci-dessous vont vous contraindre d'utiliser les arguments des fonctions d'importations. Soyez vigilant. Nous vous conseillons de vérifier vos résultats avec la fonction `View()`.

- Importer le dataset `quiz1.csv` dans un objet appelé `quiz1` avec la fonction `read.csv()` uniquement.
- Importer le dataset `quiz2.csv` dans un objet appelé `quiz2` avec la fonction `read.csv()` uniquement.
- Importer le dataset `quiz3.csv` dans un objet appelé `quiz3` avec la fonction `read.csv()` uniquement.
- Importer le dataset `quiz4.csv` dans un objet appelé `quiz4` avec la fonction `read.csv()` uniquement.

3 Exercice 2 : Compiler les données

- a. Compiler l'ensemble des dataframe *quiz* dans un seul et même data frame appelé *quiz_final*. La fonction `rbind()` devrait vous aider (voir l'exemple ci-dessous).

```
df <- data.frame()
dim(df)
```

```
## [1] 0 0
```

```
df <- rbind(df, mtcars)
dim(df)
```

```
## [1] 32 11
```

```
df <- rbind(df, mtcars)
dim(df)
```

```
## [1] 64 11
```

- b. Modifier votre script de la question précédente afin d'ajouter dans le data frame, une colonne appelée *numero_quiz* qui renseigne le numéro du quiz (1,2,3 ou 4).

4 Exercice 3 : Analyses statistiques

- Déterminer la moyenne du nombre de points.
- Déterminer les quartiles du temps moyen de réponse.
- Constuire un boxplot de la distribution du temps moyen de réponse
- Constuire un histogramme de la distribution du nombre de points
- Calculer la répartition du nombre de participant par groupe
- Représenter cette répartition dans un diagramme en barre
- Calculer la moyenne agrégée pour groupe quiz.
- Représenter graphiquement cette différence avec un boxplot bivarié
- Calculer le nombre total de points de chaque joueur et afficher le podium.