

Il faut s'entraîner... Importation

juin 2015

1 Fichiers texte

Basique mais indispensable, le changement de répertoire de travail.

1. utilisez l'interface graphique pour changer le répertoire courant
2. vérifiez le répertoire courant dans le terminal
3. utilisez le terminal pour changer le répertoire courant et vérifiez

Sur le fichier texte, on va travailler sur la data.frame des Iris de Fisher. Ce fichier est très bien connu de tous les statisticiens car pris en exemple pour les analyses factorielles et les analyses discriminantes.

Le fichier est assez simple et le data.frame existe en mémoire. Vous pourrez comparer l'importation avec le data.frame stocké dans la collection d'exemples de R.

Pour récupérer une data.frame *exemple*, il faut taper :

1. Regarder le fichier avec un éditeur de texte (utilisez NotePad++, clic-droit sur le fichier)
2. Une fois repérer les caractéristiques du fichier : séparateur de décimales, séparateurs de champ, ...
3. Choisir la bonne fonction et les bonnes options pour R.

Lors de l'importation, essayer d'importer en transformant les variables caractère tantôt en *character* tantôt en *factor*.

1. Iris.csv
2. Iris.txt

Maintenant même exercice sur les deux fichiers :

1. patient.dat
2. patient.csv
3. pour des raisons de traitement de données, importer les *ACP*, *peridurale* et *periACP* sous forme *character* et non *numeric/integer*.

1.1 Exportation

Le but est d'exporter le data.frame mtcars dans les data.frame exemples :

1. pour Excel
2. pour SAS (sans les row.names)

2 Format SPSS

Il faudra :

1. charger le paquet foreign
2. utiliser la bonne commande pour charger le paquet SPSS

3 Importation de fichier Excel

Dans cette partie, on utilise *XLConnect* pour importer un fichier Excel.

Les **données** sont les appréciations d'enseignants par des élèves d'un collège américain.

Il est au format Excel 2010.

1. charger le fichier
2. Transformer le fichier "xlsx" en fichier Excel 2003
3. Charger correctement le patient.xlsx en évitant d'importer les descriptions des variables

4 En supplément...

1. Il faut charger le fichier texte avec un format fixe *Iris-fixed.txt*
2. Importation des données ESLC depuis une base SQLite. Pour des questions de taille, faites une requête SQL pour n'importer que les données pour la France (condition 1). Puis importer les données pour les élèves français qui ont passé l'anglais (condition 1 et 2)

```
> # 1) country_id = 'FR'
>
> # 2) targetLanguage_id='EN'
```

3. insérer une nouvelle table dans la base de données SQLite avec seulement les données pour les élèves français. Le nom de la table sera "fr"