

Etude sur le risque de récurrence d'anévrisme.

N. Glade

Données

Une étude (réelle) a été réalisée en France et au Vietnam pour déterminer les récurrences d'anévrismes chez des personnes ayant déjà été prises en charge pour un anévrisme.

Cette étude comporte 5 variables :

- country : France ou Vietnam
- gender : M ou F
- aneurysm : âge auquel a été pris en charge le patient pour la première fois
- bmi : indice de masse corporelle
- risk : nombre de récurrences (0 = pas de récurrences)

Livrables.

L'objectif est de poser une (idéalement) ou deux questions scientifiques pouvant être faites sur cette étude et de présenter vos résultats dans un article de 4 pages exactement (min et max), police type Roman taille 10, double colonne, interligne 1. Le programme R fera partie des livrables.

L'article au format PDF devra comporter :

- un titre
- les auteurs
- un résumé (les questions, les données/contexte, ce qui a été fait, les résultats)
- une intro présentant
 - le jeu de données (effectifs par pays / genre, histogrammes de l'âge, du bmi)
 - les questions
- la méthode
- les résultats
- une discussion

Soyez précis-e sur ce que vous dites : tout doit avoir du sens, être utile à votre étude et ne pas être scolaire. Ne pas ajouter de superflu.

Le rendu du code se fera au format texte (.R) ou jupyter (noyau R).

L'ensemble (code R + PDF) seront déposés sous Chamilo dans un seul fichier Zip portant le nom du binôme : TIS4_Rendu2_NOM1_NOM2.zip

Pistes d'études.

Vous pouvez étudier par exemple :

- le lien entre le risque de récurrence et les autres facteurs (anova)
- faire une classification gaussienne (mclust) sur le bmi puis expliquer les clusters avec les autres facteurs (mclust + anova)
- prédiction du risque, ... ou d'autres variables

Pour ouvrir le fichier.

```
d <- read.table("aneurysm.csv", header=T, sep=";", stringsAsFactors=T)
head(d)
```