

Tutoriel sur R

François Husson

UP de mathématiques appliquées - l'institut Agro



Journées d'études en statistique – SFdS 2021

- ① Visualisation
- ② Imputation par ACP
- ③ Imputation par ACM
- ④ Imputation par AFDM

Visualiser le dispositif de données avec les fonctions du package VIM

- utiliser les fonction `aggr`, `matrixplot` et `marginplot` du package VIM pour visualiser le dispositif de données manquantes
- Construire une matrice de présence-absence et faire une ACM pour visualiser les associations de données manquantes entre variables

- ① Visualisation
- ② Imputation par ACP
- ③ Imputation par ACM
- ④ Imputation par AFDM

Imputation simple par ACP et ACP sur tableau incomplet

- ① Récupérer ou importer le jeu de données ozone depuis
<http://factominer.free.fr/missMDA/ozoneNA.csv>
- ② Sélectionner les variables quanti uniquement
- ③ Estimer le nombre de composantes pour imputer
- ④ Imputer
- ⑤ Faire l'ACP à partir du jeu de données complété (les variables quali du jeu initiales seront utilisées après les avoir concaténées au jeu complété)

Imputer par ACP et faire une ACP avec Factoshiny

- ① Utiliser la fonction `Factoshiny` du package `Factoshiny` sur le jeu de donnée

Imputation multiple ACP avec mice et missMDA

- ① Utiliser la fonction `mice` (du package `mice`) pour faire de l'imputation multiple
- ② Utiliser la fonction `MIPCA` (du package `missMDA`) pour faire de l'imputation multiple
- ③ Utiliser les tableaux imputés obtenu par `mice` pour faire la régression du maximum d'ozone en fonction des autres variables sur le tableau incomplet
- ④ mêmes question avec les tableaux imputés par `MIPCA`

- ① Visualisation
- ② Imputation par ACP
- ③ Imputation par ACM**
- ④ Imputation par AFDM

ACM sur tableau incomplet

- ① Charger le jeu de données `vnf` : `data(vnf)`
- ② Faire l'ACM en ajoutant une modalité NA pour chaque variable ayant au moins un NA (ce qui est fait par défaut par la fonction MCA)
- ③ Imputer le tableau de données par la fonction `imputeMCA` après avoir estimé le nombre de composantes pour imputer. Faire l'ACM en utilisant le tableau disjonctif imputé

- ① Visualisation
- ② Imputation par ACP
- ③ Imputation par ACM
- ④ Imputation par AFDM

Imputation par AFDM et AFDM sur tableau incomplet

- ① Reprendre le jeu de données ozone
- ② Imputer par AFDM
- ③ Faire l'AFDM sur le tableau en utilisant le tableau disjonctif complété
- ④ Imputer par forêts aléatoire avec la fonction `missForest` package `missForest`