

CODE R DU MEMOIRE

Table des matières

<i>Code used in Rstudio for the dtb database</i>	<i>1</i>
<i>Code used in RStudio for the dtblog database</i>	<i>5</i>
<i>Code used in RStudio for the dtb177 database</i>	<i>6</i>
<i>Code used in RStudio for the secondary database</i>	<i>7</i>
<i>Code used in RStudio for the tertiary database</i>	<i>9</i>

Code used in Rstudio for the *dtb* database

```
#importation des packages
```

```
Install.packages(corrplot)
```

```
Library(corrplot)
```

```
Install.packages(wooldridge)
```

```
Library(wooldridge)
```

```
Install.packages(car)
```

```
Library(car)
```

```
Install.packages(ggplot2)
```

```
Library(ggplot2)
```

```
#importation des données
```

```
#datatop <- read.csv("~/desktop/bdd", header=false, sep=";")
```

```
View(dtb)
```

```
Dtb<- dtb[2:4]
```

```
Colnames(dtb)=c('scope1','scope2','scope3')
```

```
View(dtb)
```

```
#analyse linéaire
```

```
#corrélation de pearson
```

```
Library(corrplot)
```

```
Cor(dtb)
```

```
Corrplot(cor(dtb))
```

```
#régression linéaire simple
```

```
Install.packages(wooldridge)
```

```
Library(wooldridge)
```

```
Modellinaire= lm(dtb$scope2~dtb$scope1)
```

```
Summary(modellinaire)
```

```
Check_model(modellinaire)
```

```
Modellinaire= lm(dtb$scope3~dtb$scope1) #pas de lien entre scope 1 et 3 car p-value > 5%
```

```
Summary(modellinaire)
```

```
Check_model(modellinaire)
```

```
Modellinaire= lm(dtb$scope3~dtb$scope2)
```

```
Summary(modellinaire)
```

```
Check_model(modellinaire)
```

```
#regression linéaire multiple
```

```
Modellinaire= lm(dtb$scope3~dtb$scope1+dtb$scope2)
Summary(modellinaire)
```

```
#visualisation des données en histogrammes
```

```
Hist(dtb$scope1)
Hist(dtb$scope2)
Hist(dtb$scope3)
```

```
#scatterplot 2d
```

```
Fig1 = ggplot(data = dtb,aes(x =scope1, y = scope2)) +
  Geom_point()
Fig1
```

```
Fig2 = ggplot(data = dtb,aes(x =scope1, y = scope3)) +
  Geom_point()
Fig2
```

```
Fig3 = ggplot(data = dtb,aes(x =scope2, y = scope3)) +
  Geom_point()
Fig3
```

```
#analyse non linéaire
```

```
#corrélation de spearman
```

```
Cor(dtb, method = "spearman")
Result = cor(dtb$scope2,dtb$scope1, method = "spearman")
Summary(result)
Print(result)
```

```
Result = cor(dtb$scope3,dtb$scope1, method = "spearman")
Summary(result)
Print(result)
```

```
Result = cor(dtb$scope3,dtb$scope2, method = "spearman")
Summary(result)
Print(result)
```

```
#régression non-linéaire simple
```

```
Modellinaire= lm(dtb$scope2~dtb$scope1+i(dtb$scope1^2))
Summary(modellinaire)
```

```
Modellinaire= lm(dtb$scope3~dtb$scope1+i(dtb$scope1^2))
Summary(modellinaire)
```

```
Modellinaire= lm(dtb$scope3~dtb$scope2+i(dtb$scope2^2))
```

```
Summary(modellineaire)
```

```
#régression non linéaire multiple
```

```
Modellineaire= lm(dtb$scope3~dtb$scope1+i(dtb$scope1^2)+ dtb$scope2+i(dtb$scope2^2))
```

```
Summary(modellineaire)
```

```
#trouver le nombre de cluster optimal
```

```
Install.packages("maptools")
```

```
Install.packages("purrr")
```

```
Install.packages()
```

```
Library(maptools)
```

```
Library(purrr)
```

```
Library(sp)
```

```
#elbow method
```

```
Library(factoextra)
```

```
Library(nbclust)
```

```
Fviz_nbclust(dtb, kmeans, method = "wss") +
```

```
  Geom_vline(xintercept = 3, linetype = 2)+
```

```
  Labs(subtitle = "elbow method")
```

```
#silhouette analysis
```

```
Install.packages("cluster")
```

```
Install.packages("factoextra")
```

```
Install.packages("ggplot2")
```

```
Install.packages("maptools")
```

```
Install.packages("purrr")
```

```
Install.packages()
```

```
Library(maptools)
```

```
Library(purrr)
```

```
Library(sp)
```

```
Library(cluster)
```

```
Library(ggplot2)
```

```
Library(factoextra)
```

```
#nouveau code pour silhouette methode
```

```
Fviz_nbclust(dtb, kmeans, method = "silhouette")+
```

```
  Labs(subtitle = "silhouette method")
```

```
#gap statistic
```

```
Library(factoextra)
```

```
Library(cluster)
```

```
Install.packages("clusgap")
```

```

# compute gap statistic qui nous en donne 1
Set.seed(123)
Gap_stat <- clusgap(dtb, fun = kmeans, nstart = 25,
                   K.max = 5, b = 50)

Fviz_gap_stat(gap_stat)

Fviz_nbclust(dtb, kmeans,
             Nstart = 25,
             Method = "gap_stat", k.max = 5,
             Nboot = 500 # reduce it for lower computation time (but less precise results)
) +
  Labs(subtitle = "gap statistic method")

```

```

#consensus-based algorithm
Install.packages("parameters")
Install.packages("easystats")
Install.packages("nbclust")
Install.packages("mclust")
Library(parameters)
Library(easystats)
Library(nbclust)
Library(mclust)

```

```

N_clust <- n_clusters(dtb,
                    Package = c("easystats", "nbclust", "mclust"),
                    Standardize = false)

N_clust
Plot(n_clust)

```

```

#cluster analysis
Kc<-kmeans(dtb,3)
Kc

```

```

Kc<-kmeans(dtb,2)
Kc

```

```

Kc<-kmeans(dtb,4)
Kc

```

Code used in RStudio for the *dtblog* database

```
#analyse lineaire sur dtblog
```

```
#transformation de la dtb en log
```

```
Dtblog<- data.frame(scope1log= log(dtb$scope1),  
                    Scope2log=log(dtb$scope2),  
                    Scope3log = log(dtb$scope3))
```

```
View(dtblog)
```

```
#corrélation linéaire
```

```
Cor(dtblog)
```

```
#régression linéaire simple
```

```
Modellog= lm(dtblog$scope2log~dtblog$scope1log)
```

```
Summary(modellog)
```

```
Check_model(modellog)
```

```
Modellog= lm(dtblog$scope3log~dtblog$scope1log) #fonctionne
```

```
Summary(modellog)#p-valeur hyper faible, on a qqc de significatif (on l'a vu grâce au  
graphique)
```

```
Check_model(modellog)
```

```
Modellog= lm(dtblog$scope3log~dtblog$scope2log)
```

```
Summary(modellog)
```

```
Check_model(modellog)
```

```
#régression linéaire multiple
```

```
Modellog=lm(dtblog$scope3log~dtblog$scope1log+dtblog$scope2log)
```

```
Summary(modellog)
```

```
Check_model(modellog)
```

```
#analyse non lineaire
```

```
#corrélation de spearman
```

```
Cor(dtblog, method = "spearman")
```

```
#régression non linéaire
```

```
Modellinaire= lm(dtblog$scope2~dtblog$scope1+i(dtblog$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire= lm(dtblog$scope3~dtblog$scope1+i(dtblog$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire= lm(dtblog$scope3~dtblog$scope2+i(dtblog$scope2^2))
Summary(modellinaire)
```

```
#régression multiple non-linéaire
```

```
Modellinaire= lm(dtblog$scope3~dtblog$scope1+i(dtblog$scope1^2)+
dtblog$scope2+i(dtblog$scope2^2))
Summary(modellinaire)
```

Code used in RStudio for the *dtb177* database

```
#importation des données
```

```
Colnames(dtb177)=c('scope1','scope2','scope3')
View(dtb177)
```

```
#analyse lineaire
```

```
#corrélation de pearson des scopes 1, 2 et 3 ensemble
```

```
Library(corrplot)
Corrplot(cor(dtb177))
Cor(dtb177)
```

```
#régression linéaire simple
```

```
Library(wooldridge)
```

```
Modellinaire= lm(dtb177$scope2~dtb177$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(dtb177$scope3~dtb177$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(dtb177$scope3~dtb177$scope2)
Summary(modellinaire)
```

```
#régression linéaire multiple scope 1-2-3
```

```
Library(wooldridge)
Library(car)
Reglinmult=lm(dtb177$scope3~dtb177$scope1 +dtb177$scope2)
Summary(reglinmult)
```

```
#non-parametric analysis
```

```
#corrélation de spearman
```

```
Cor(dtb177, method = "spearman")
```

```
Result = cor(dtb177$scope3,dtb177$scope2, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
Result = cor(dtb177$scope2,dtb177$scope1, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
Result = cor(dtb177$scope3,dtb177$scope1, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
#régression non linéaire
```

```
Modellinaire= lm(dtb177$scope2~dtb177$scope1+i(dtb177$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire= lm(dtb177$scope3~dtb177$scope1+i(dtb177$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire= lm(dtb177$scope3~dtb177$scope2+i(dtb177$scope2^2))
```

```
Summary(modellinaire)
```

```
#régression non linéaire multiple
```

```
Modellinaire= lm(dtb177$scope3~dtb177$scope1+i(dtb177$scope1^2)+  
dtb177$scope2+i(dtb177$scope2^2))
```

```
Summary(modellinaire)
```

Code used in RStudio for the *secondary* database

```
#importation des données
```

```
View(secondary)
```

```
Colnames(secondary)=c('scope1','scope2','scope3')
```

```
View(secondary)
```

```
#analyse lineaire
```

```
#corrélation de pearson des scopes 1, 2 et 3 ensemble
```

```
Cor(secondary)
```

```
#régression linéaire simple
Library(wooldridge)
```

```
Modellinaire= lm(secondary$scope2~secondary$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(secondary$scope3~secondary$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(secondary$scope3~secondary$scope2)
Summary(modellinaire)
```

```
#régression linéaire multiple scope 1-2-3
```

```
Library(wooldridge)
```

```
Library(car)
```

```
Reglinmult=lm(secondary$scope3~secondary$scope1 +secondary$scope2)
```

```
Summary(reglinmult)
```

```
#non-parametric analysis
```

```
#corrélacion de spearman
```

```
Result = cor(secondary, method = "spearman")
```

```
Print(result)
```

```
Result = cor(secondary$scope2,secondary$scope1, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
Result = cor(secondary$scope3,secondary$scope1, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
Result = cor(secondary$scope3,secondary$scope2, method = "spearman")
```

```
Summary(result)
```

```
Print(result)
```

```
#régression non linéaire simple
```

```
Modellinaire=lm(secondary$scope2~secondary$scope1+i(secondary$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire=lm(secondary$scope3~secondary$scope1+i(secondary$scope1^2))
```

```
Summary(modellinaire)
```

```
Modellinaire=lm(secondary$scope3~secondary$scope2+i(secondary$scope2^2))
Summary(modellinaire)
```

```
#régression non linéaire multiple
```

```
Modellinaire=lm(secondary$scope3~secondary$scope1+i(secondary$scope1^2)+
secondary$scope2+i(secondary$scope2^2))
Summary(modellinaire)
```

Code used in RStudio for the *tertiary* database

```
#importation des données
```

```
View(tertiary)
Colnames(tertiary)=c('scope1','scope2','scope3')
View(tertiary)
```

```
#corrélation de pearson des scopes 1, 2 et 3 ensemble
```

```
Cor(tertiary)
```

```
#régression linéaire simple
```

```
Library(wooldridge)
```

```
Modellinaire= lm(tertiary$scope2~tertiary$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(tertiary$scope3~tertiary$scope1)
Summary(modellinaire)
```

```
Modellinaire= lm(tertiary$scope3~tertiary$scope2)
Summary(modellinaire)
```

```
#régression linéaire multiple scope 1-2-3
```

```
Library(wooldridge)
```

```
Library(car)
```

```
Reglinmult=lm(tertiary$scope3~tertiary$scope1 +tertiary$scope2)
Summary(reglinmult)
```

```
#non-parametric analysis
```

```
#corrélation de spearman
```

```
Cor(tertiary, method = "spearman")
```

```
Result = cor(tertiary$scope2,tertiary$scope1, method = "spearman")
Summary(result)
Print(result)
```

```
Result = cor(tertiary$scope3,tertiary$scope1, method = "spearman")
Summary(result)
Print(result)
```

```
Result = cor(tertiary$scope3,tertiary$scope2, method = "spearman")
Summary(result)
Print(result)
```

#régression non linéaire simple

```
Modellinaire= lm(tertiary$scope2~tertiary$scope1+i(tertiary$scope1^2))
Summary(modellinaire)
```

```
Modellinaire= lm(tertiary$scope3~tertiary$scope1+i(tertiary$scope1^2))
Summary(modellinaire)
```

```
Modellinaire= lm(tertiary$scope3~tertiary$scope2+i(tertiary$scope2^2))
Summary(modellinaire)
```

#régression non linéaire multiple

```
Modellinaire= lm(tertiary$scope3~tertiary$scope1+i(tertiary$scope1^2)+
tertiary$scope2+i(tertiary$scope2^2))
Summary(modellinaire)
```