

```

#Correction du TD4
#Répertoire de travail
getwd()
setwd(dirname(file.choose()))
#Vérifier le contenu des fichiers avant de les importer
file.show(file.choose())
#Importer DataTD4
Td4.1<-read.table(file.choose(),header=T,sep=";",dec=",")
#Importer DataTD4rbind
Td4.2<-read.table(file.choose(),header=T,sep=";",dec=",")
#Importer DataTD4merge
Td4.3<-read.table(file.choose(),header=T,sep=";",dec=",")
#Vérifier le contenu des fichiers
str(Td4.1)
str(Td4.2)
Td4.12<-rbind(Td4.1,Td4.2)
#vérifier l'importation de TD4.2 dans TD4.1
str(Td4.12)
str(Td4.3)
Td4.123<-merge(Td4.12,Td4.3,by="NUMIDENT")
str(Td4.123)
#sauvegarde du fichier
write.csv2(Td4.123,file=paste('DataTD4MAJ',as.character(Sys.Date(),"%Y%m%d"),".csv",sep=""))

#Découverte de l'outil graphique
set.seed(13) # permet que le tirage aléatoire suivant soit toujours identique
Echant1<-rnorm(20,mean=55,sd=10)
# construit un vecteur "a" contenant 20 points, extraits d'une population de moyenne 55 et
d'écart type de 10

#Analyse univariée des données – épidémiologie descriptive
mean(Echant1)
sd(Echant1)
max(Echant1)
summary(Echant1)

par(mfrow=c(3, 2))#partage la fenêtre graphique en 3 lignes et 2 colonnes
hist(Echant1)#histogramme
boxplot(Echant1)#boite à moustache
stripchart(Echant1, vertical=T) #nuage de point
stripchart(Echant1,pch=16,cex=1,method='jitter',jitter=0.1, vertical=T)
stripchart(Echant1,pch=16,cex=1,method='stack', vertical=T)
stripchart(Echant1,pch=12,cex=0.5,col=2,method='jitter', jitter=0.05,vertical=T)
stripchart(Echant1,pch=10,cex=1.2,col="violet",method='jitter', vertical=T)
par(new=T)
whiskerplot(Echant1)

#Découverte de l'outil graphique (2)

set.seed(13)
x1<-rnorm(10,mean=100,sd=10)
set.seed(9)
x2<-rnorm(10,mean=110,sd=10)
par(mfrow=c(2,1))#partage la fenêtre graphique en 2 lignes et une colonne
boxplot(x1,x2)
t.test(x1,x2)#test T de student
plot(x1,x2)
summary(lm(x2~x1))

#Découverte de l'outil graphique (3)

#On va utiliser le fichier créé lors de la 1° partie de la séance rm(list=ls())
Mydata<-read.table(file.choose(), header=T, sep=";", dec=",")
# que Mydata a bien été créé avec la fonction ls()
ls()

#Vérifier que l'objet créé contient bien l'information ?
names(Mydata)
str(Mydata)

```

```

head(Mydata)
summary(Mydata)#permet de relever quelles variables sont à factoriser et lesquelles sont à
garder en format numérique

#Créer un vecteur contenant les numéros des colonnes à transformer
Transvar<- c(7,13,15,16) # à vous de définir ces x
#Programmer une boucle qui va transformer chacune des variables choisies en variable
qualitative
for (i in 1:length(Transvar)) { Mydata[, Transvar[i]] <- factor(Mydata[, Transvar[i]))}# à
vous de définir X
for (i in c(7,13,15,16)){Mydata[,i]<-factor(Mydata[,i])}
#9) Créer une table de données qui ne contienne que les données pour les classe d'âge sont
: (20,22] et (25,28.7]
Mydata.1<-subset(Mydata, CLASSAGE=="(18,22]"|CLASSAGE=="(25,28.7]")
dim(Mydata.1)

#9bis) diagramme en bâton données catégorielles
plot(Mydata$CLASSAGE)
plot(Mydata$CLASSAGE,col=rainbow(4))
plot(Mydata$CLASSAGE,col=rainbow(4), xlab="Classe d'âge", ylab="Effectifs", cex.lab=1.5,
cex.axis=0.5)
plot(Mydata$CLASSAGE,col=rainbow(4), xlab="Classe d'âge",
ylab="Effectifs",cex.lab=1.5,cex.axis=0.5,
xaxt="n")
axis(side=1,labels=paste(1:4,"°GROUPE",sep=""), tck=0, at= plot(Mydata$CLASSAGE,plot=F))

#10) camembert
pie(summary(Mydata$SEXE)) #ou pie(table(Mydata$SEXE)
nom<-c("Hommes", "Femmes")
titre<- "répartition par sexe"
pie(table(Mydata$SEXE),label=nom, main = titre,cex.main=2, col=colors()[c(131,213)])

#la fonction palette() permet de connaître les numéros des couleurs.
palette()
#[1] "black" "red" "green3" "blue" "cyan" "magenta" "yellow" "gray"
#11) données quantitatives
plot(Mydata$AGE)#nuage de point
hist(Mydata$AGE)#histogramme
hist(Mydata$AGE,nclass=5)#on choisit le nombre de classe
hist(Mydata$AGE,nclass=5,plot=F)#pour avoir les données du graphiques
hist(Mydata$AGE,nclass=5,plot=F)$breaks; hist(Mydata$AGE,nclass=5,plot=F)$counts
hist(Mydata$AGE,freq=T,breaks=c(19:31))

#12) boîtes à moustache simple
par(mfrow=c(2,2))
boxplot(Mydata$AGE, main="AGE")
boxplot(Mydata$POIDS, main="POIDS")
boxplot(Mydata$TAILLE, main='TAILLE')
boxplot(Mydata$IMC, main='IMC')

#13) boîtes à moustache par catégorie
boxplot(split(Mydata$AGE,Mydata$SEXE))
boxplot(Mydata$AGE~Mydata$SEXE, main="Age/Sexe")
boxplot(Mydata$POIDS~Mydata$SURPOIDS, main="Poids/Surpoids")
boxplot(Mydata$IMC~Mydata$SEXE, main='IMC/Sexe')
boxplot(Mydata$IMC~Mydata$CLASSAGE, main="IMC/Classe d'âge")

#14) ajouter un titre
boxplot(Mydata$AGE~Mydata$SEXE, main="Moyenne d'âge par sexe")
#ajouter un nom à l'axe des abscisses
boxplot(Mydata$AGE~Mydata$SEXE, main="Moyenne d'âge par sexe", xlab="SEXE")

#15) partager la fenêtre graphique pour présenter plusieurs graphiques
par(mfrow= c(3,2))
par(mar=c(4,3,1.5,2))
boxplot(Mydata$AGE,main="Age median")
boxplot(Mydata$POIDS, main='Poids')
boxplot(Mydata$AGE~Mydata$SEXE, main="Age moyen par sexe",xlab="SEXE")
pie(summary(as.factor(Mydata$MALADIE)),label=c("Male","Female"), main = "Répartition par
sexe")
plot(Mydata$CLASSAGE)

```

```

plot(Mydata$SEXE)
#supprimer le partage
par(mfrow=c(1,1))
16) autre méthode fonction layout() pour avoir des fenêtres de taille différente
zones<-matrix(c(1,1,1,2,3,4),ncol=2)
layout(zones)
layout.show(max(zones)) #affiche les régions de figure
boxplot(Mydata$AGE~Mydata$SEXE, main="Age moyen par sexe",xlab="SEXE")
pie(summary(as.factor(Mydata$MALADIE)),label=c("non malade","malade"), main = "Repartition
\n en fonction du statut \n en lien avec la maladie")
plot(Mydata$CLASSAGE, main ="répartition par age",col="blue")
plot(Mydata$SEXE, main ="répartition par sexe", col ="red")
layout(matrix(1))
17) Exporter un graphique
#Exemple1
graphics.off()
pdf(file="exempleA.pdf")
hist(Mydata$AGE,nclass=5, main ="répartition par âge",col="blue")
plot(Mydata$SEXE, main ="répartition par sexe", col =c("red","white"))
dev.off()
#Exemple2
graphics.off()
pdf(file="exempleB%03d.pdf",onefile=FALSE)
hist(Mydata$AGE,nclass=5, main ="répartition par age",col="blue")
plot(Mydata$SEXE, main ="répartition par sexe", col ="red")
dev.off()

#17) Exercice pour la découverte de l'outil graphique et des tests statistiques
Utiliser ici le fichier DataTD4 2° partie.
Mydata2<-read.table(file.choose(), header=T, sep=";", dec=",")
names(Mydata2)
by(Mydata2$AGE,Mydata2$SEXE,mean)
par(mfrow=c(1,1))
Titre<-"Age moyen en fonction du sexe"
boxplot(Mydata2$AGE~Mydata2$SEXE,col=c(2,4), main=Titre, ylab="Age en année", xlab="Sexe")
jpeg(file.choose())
boxplot(Mydata2$AGE~Mydata2$SEXE,col=c(2,4), main=Titre, ylab="Age en année", xlab="Sexe")
dev.off() # ferme la fenêtre graphique active
wilcox.test(Mydata2$AGE[Mydata2$SEXE==2],Mydata2$AGE[Mydata2$SEXE==1],paired=FALSE)
wilcox.test(Mydata2$AGE~Mydata2$SEXE,paired=FALSE)

```