

Rpaillassse

Pauline et Thomas

15 septembre 2016

Le Cake Report est un projet visant à se familiariser avec quelques différents outils libres de droit pouvant être utilisés en science: pour collecter, interpréter et publier des données. Le but de cette base de données est de tester différents biscuits du commerce lors du goûter!

Importer les données

```
getwd()

## [1] "/Users/Ofix/ateliers-open-geek/Atelier_04"

setwd("/Users/Ofix/ateliers-open-geek/Atelier_04/")
data<-read.csv("grades.csv")
```

Data dictionary

subject: numero du sujet (données anonymisées) cake: marque du biscuit price: prix de la boîte de gâteaux nb.of.cookies: nb de cookies par boîtes aspect: note entre 1 et 5 taste: goût, note entre 1 et 5 texture: note entre 1 et 5 packaging: note entre 1 et 5 quality by price: rapport qualité prix; note entre 1 et 5 total: note sur 20 comprenant l'aspect, le goût, la texture, le packaging et le rapport qualité prix

Decrire les données

```
data$price <- as.numeric(data$price)
data$qualityByPrice <- as.numeric(data$qualityByPrice)
str(data)

## 'data.frame':    58 obs. of  10 variables:
## $ subject      : int  1 2 3 4 5 1 2 3 4 5 ...
## $ cake         : Factor w/ 11 levels "carrefour stick",...: 4 4 4 4 4 5 5 5 5 5 ...
## $ price        : num  9 9 9 9 9 8 8 8 8 8 ...
## $ nb.of.cookies: int   9 9 9 9 9 12 12 12 12 12 ...
## $ aspect       : Factor w/ 9 levels "", "1", "1,5", "2",...: 6 7 9 9 5 5 8 9 6 9 ...
## $ taste        : Factor w/ 9 levels "", "1", "1,5", "2",...: 8 7 9 7 7 8 8 9 4 9 ...
## $ texture      : Factor w/ 11 levels "", "0,5", "1", "1,5",...: 10 8 8 11 11 8 9 11 5 11 ...
## $ packaging    : Factor w/ 10 levels "", "0,5", "1", "1,25",...: 6 5 6 3 7 6 8 3 10 9 ...
## $ qualityByPrice: num   7 4 4 6 6 6 6 4 6 5 ...
## $ total        : Factor w/ 28 levels "", "0", "10", "10,5",...: 16 10 15 15 15 13 19 15 13 22 ...

head(data)

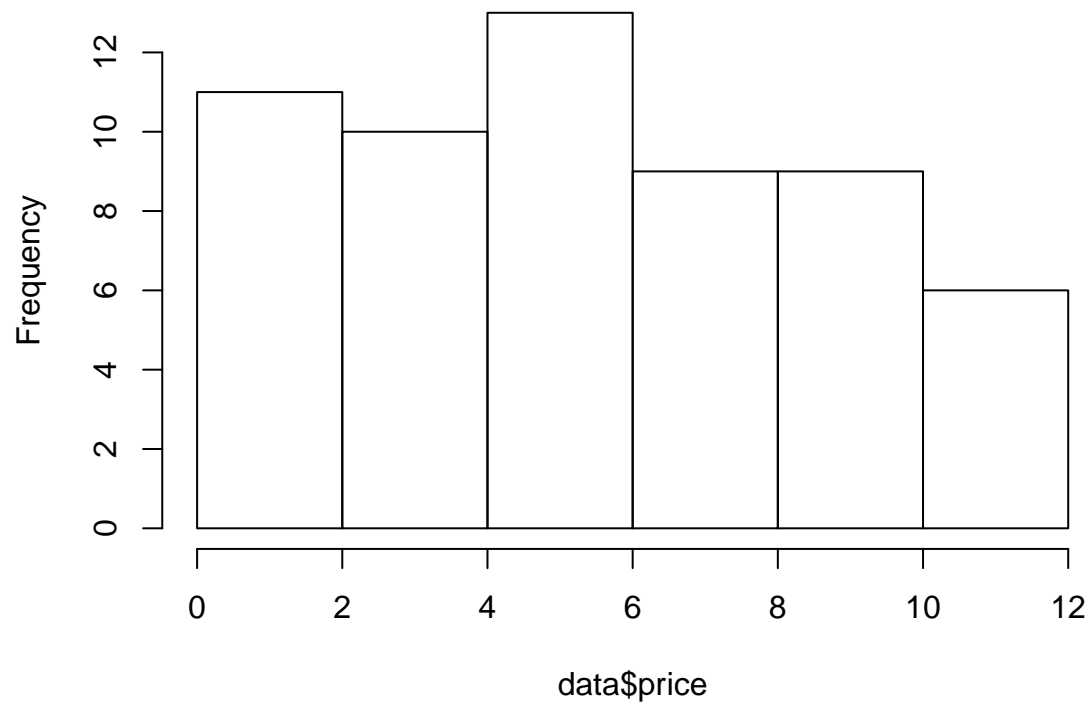
##   subject      cake price nb.of.cookies aspect taste texture
## 1      1      cookie bm      9           9      3    3,5    3,5
## 2      2      cookie bm      9           9    3,5      3      3
```

```
## 3      3      cookie bm      9      9      4      4      3
## 4      4      cookie bm      9      9      4      3      4
## 5      5      cookie bm      9      9      2,5      3      4
## 6      1 cookie granola goc      8      12      2,5      3,5      3
## packaging qualityByPrice total
## 1      2      7      15,5
## 2      1,5      4      13
## 3      2      4      15
## 4      1      6      15
## 5      2,5      6      15
## 6      2      6      14
```

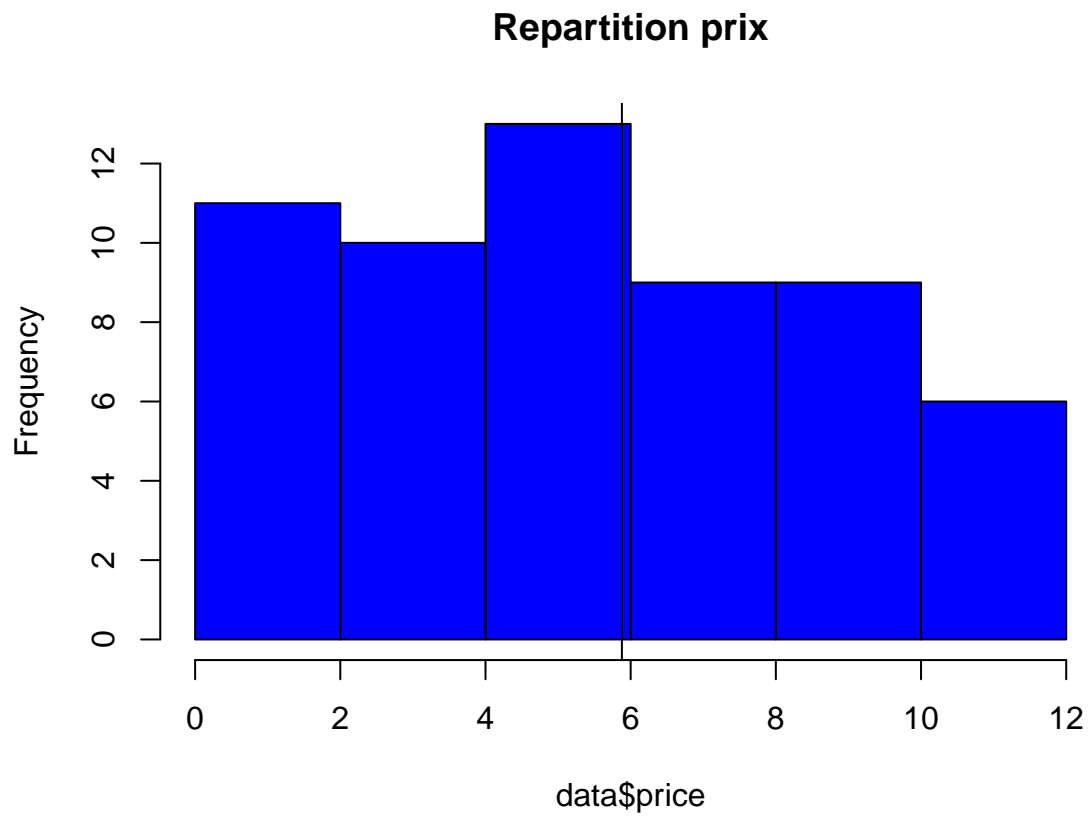
Si on veut ?tudier les prix

```
#tracer l'histogramme de la r?partition des prix
par(mfrow=c(1,1),mar=c(4,4,4,4))
help(hist) #bespon d'aide?
hist(data$price)
```

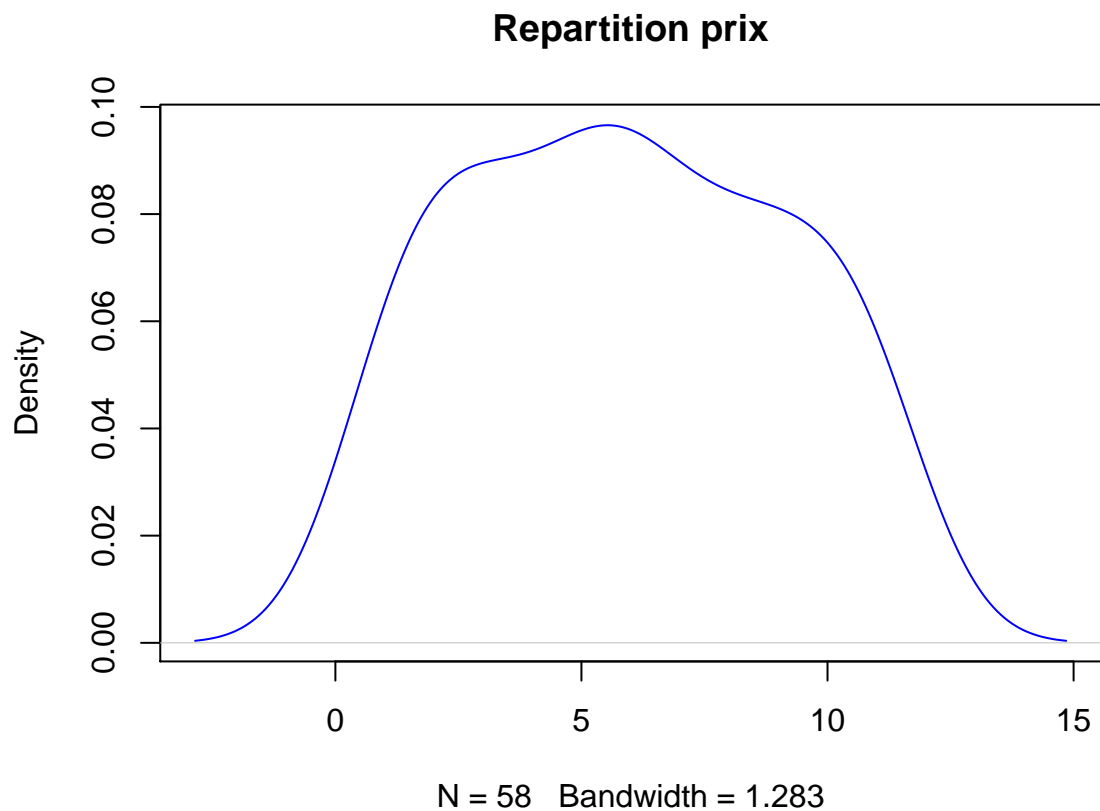
Histogram of data\$price



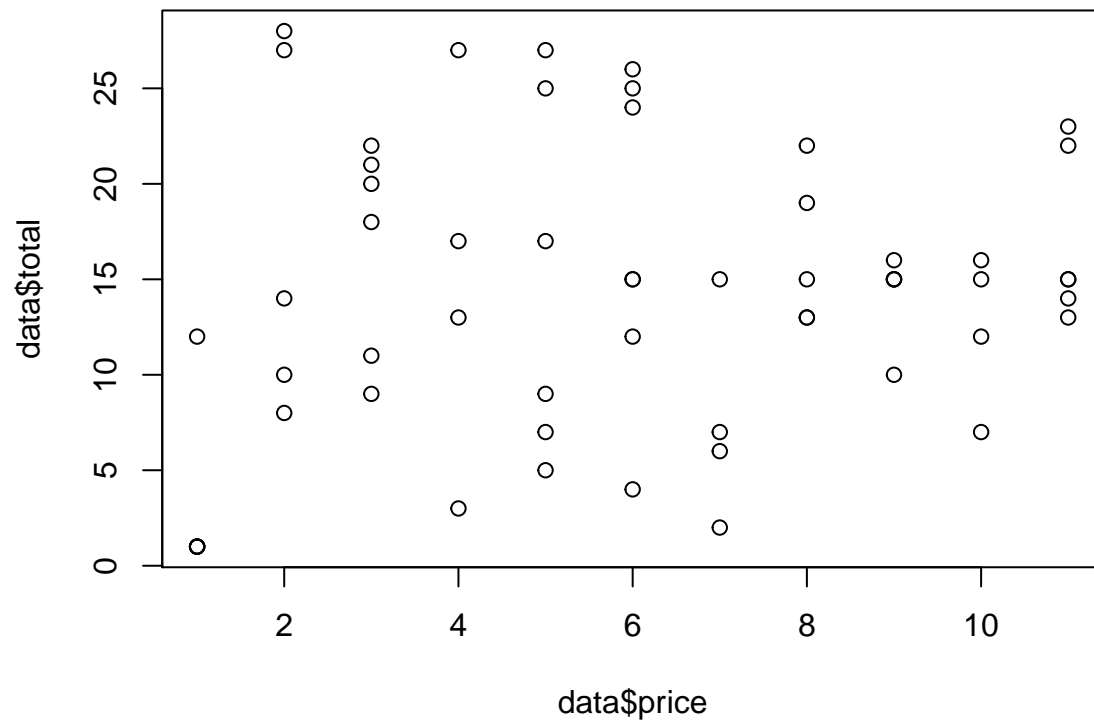
```
hist(data$price,main="Repartition prix",col="blue")
#rajouter la moyenne des prix
abline(v=mean(data$price, na.rm=TRUE))
```



```
#tracer la densit?,c'est ? dire la distribution de la variable  
plot(density(data$price,na="TRUE"),main="Repartition prix",col="blue")
```

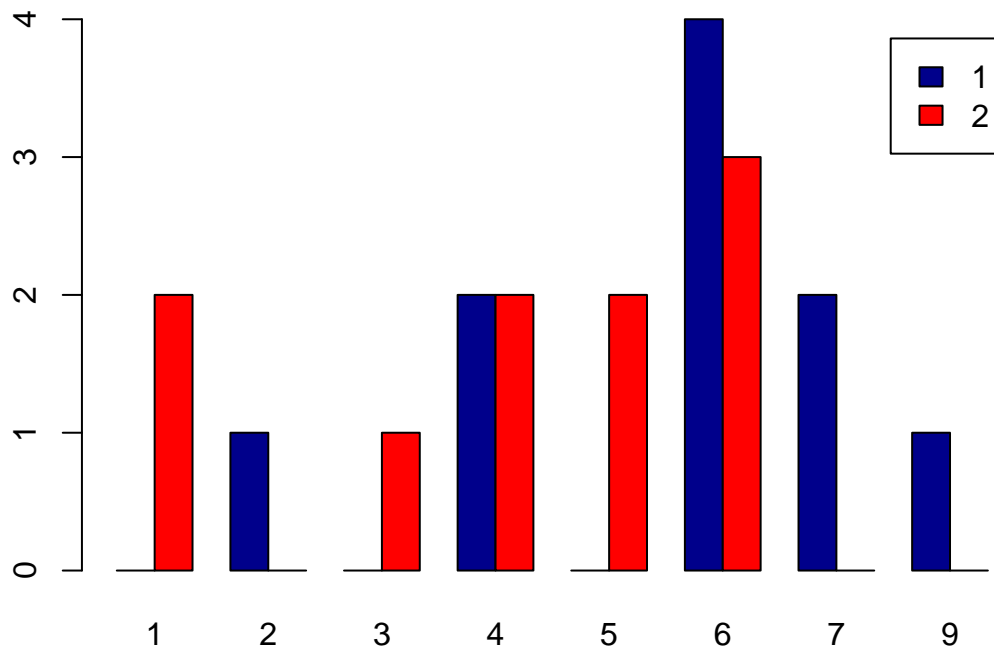


```
#tracer des points repr?sentant la note toal en fonction du prix
plot(data$price, data$total)
```



```
#tracer le rapport qualit? prix suivant les sujets
# Grouped Bar Plot
counts <- table(data[which(data$subject==1 | data$subject==2),]$subject, data[which(data$subject==1 | d
barplot(counts, main="ddistribution du rapport qualit?-prix selon le sujet 1 et 2 ",
  xlab="Rapport qualit? prix", col=c("darkblue","red"),
  legend = rownames(counts), beside=TRUE)
```

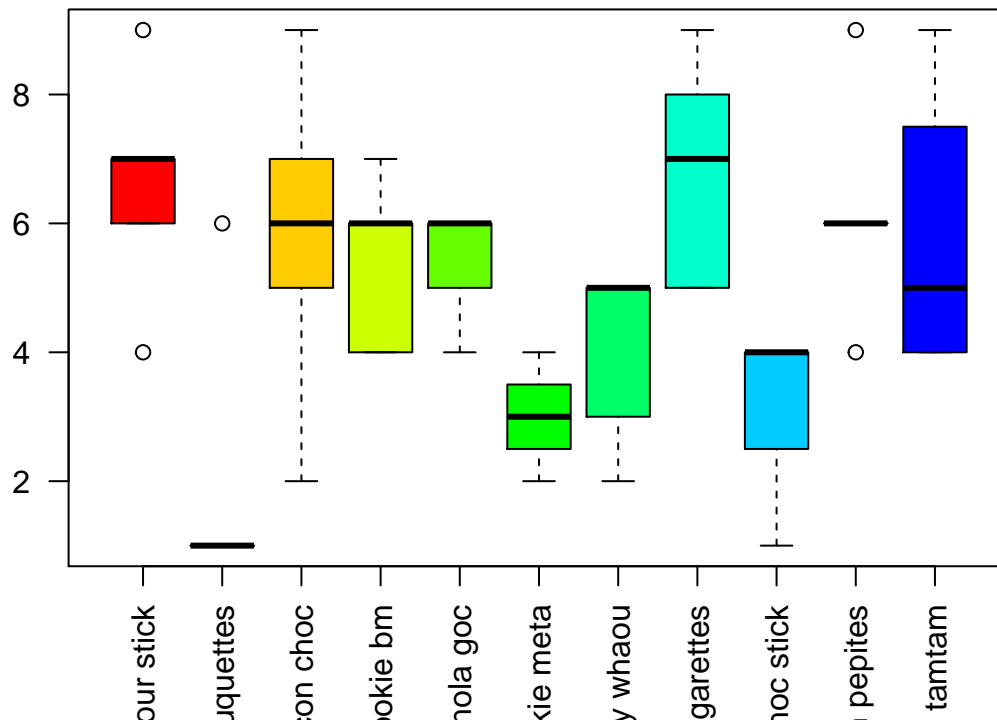
ddistribution du rapport qualit?–prix selon le sujet 1 et 2



Rapport qualit? prix

#tracer le note du rapport qualit? pris en fonction du gateau

```
boxplot(data$qualityByPrice ~ data$cake, horizontal=FALSE,col=rainbow(15),las=2)
```



En resume, pour les representations graphiques (non exhaustives): * description d'1 variable quantitative continue: ** histogramme (hist()) ** densite (plot(density())) * description d'1 variable quantitative discontinue:

** barplot (barplot()) * description d'1 variable qualitative: ** camembert (pie()) *2 variables quantitatives:
** nuage de point(plot) *1 variable quantitative en fonction d'une variable qualitative ** boite ? moustache
(boxplot())