

# **TP – Formation Statistique de base**

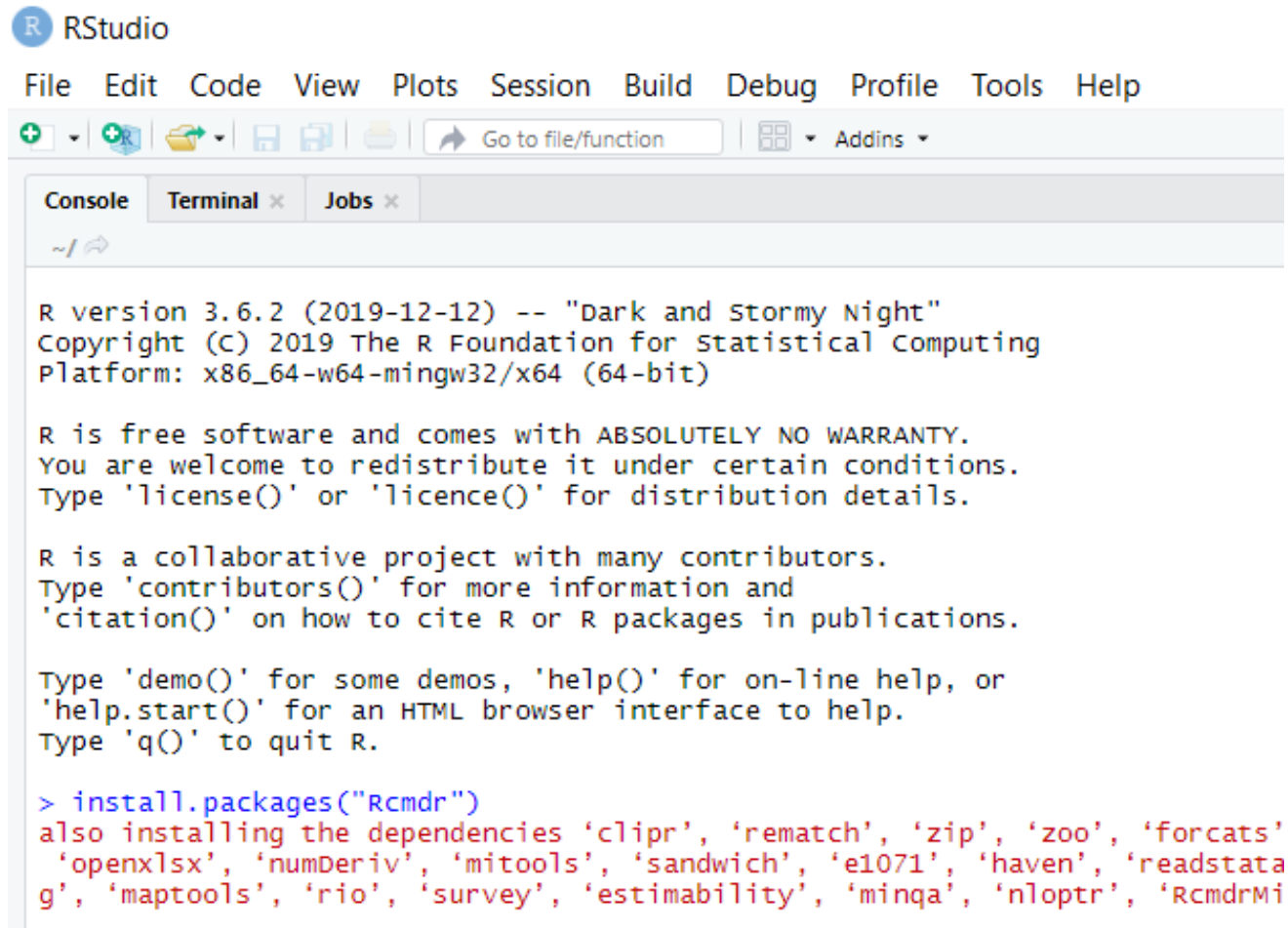
**Parties 1, 2 et 3**

09-10/03/2020

Nathalie Vialaneix, Sandrine Laguerre

# **Partie 1 : Installation, importation des données, manipulations de base**

# Installation du package Rcmdr



The image shows a screenshot of the RStudio interface. At the top, the RStudio logo and name are visible. Below that is a menu bar with options: File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help. A toolbar contains various icons for file operations and a search bar labeled 'Go to file/function'. Below the toolbar, there are tabs for 'Console', 'Terminal', and 'Jobs'. The 'Terminal' tab is active, displaying the following text:

```
~/ ↵  
  
R version 3.6.2 (2019-12-12) -- "Dark and Stormy Night"  
Copyright (C) 2019 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> install.packages("Rcmdr")  
also installing the dependencies 'clipr', 'rematch', 'zip', 'zoo', 'forcats',  
'openxlsx', 'numberj', 'mitools', 'sandwich', 'e1071', 'haven', 'readstata',  
'g', 'maptools', 'rio', 'survey', 'estimability', 'minqa', 'nloptr', 'RcmdrMi'
```

## Les données utilisées

 2020-03-09	05/03/2020 09:21	Dossier de fichiers	
 diet.csv	09/09/2019 07:32	Fichier CSV Micros...	1 Ko
 gene.csv	09/09/2019 07:32	Fichier CSV Micros...	29 Ko
 genotype.csv	09/09/2019 07:32	Fichier CSV Micros...	1 Ko
 Histo1.png	09/09/2019 11:38	Fichier PNG	4 Ko
 lipid.csv	09/09/2019 07:32	Fichier CSV Micros...	4 Ko
 logfilename	10/09/2019 13:43	Fichier	2 Ko

## Nutrimouse

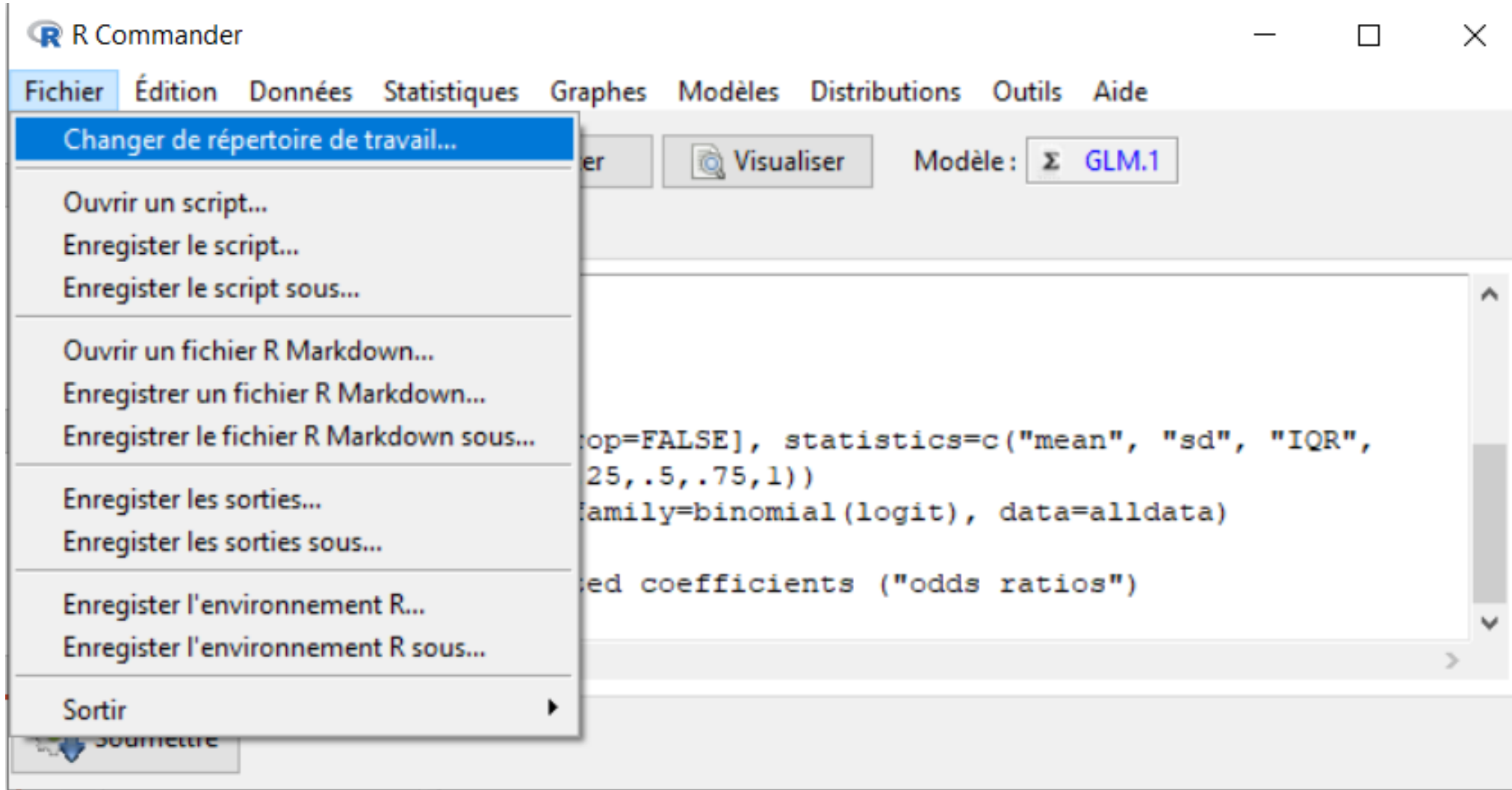
- Gene : expression de 120 gènes pertinents dans le cadre d'une étude sur la nutrition
- Lipid : concentration de 21 acides gras hépatiques
- Diet : 5 régimes testés REF (maïs, colza), COC (huile de coco), SUN (tournesol), LIN (lin), FISH (huiles de poissons)
- Genotype : 2 génotypes (WT / PPAR)

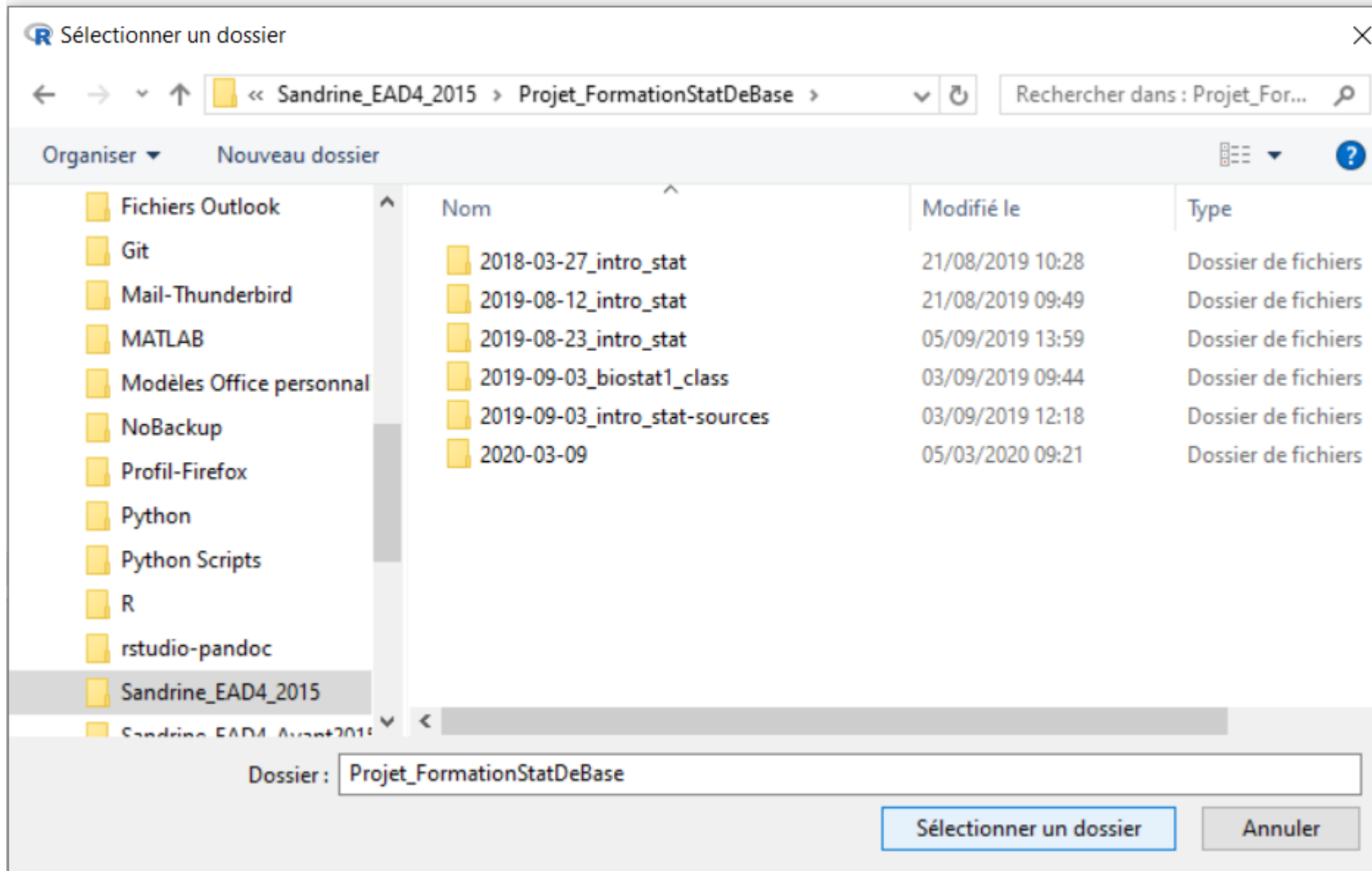
# Lancement de Rcmdr

```
> library("Rcmdr")
```

The screenshot shows the R Commander application window. The title bar reads "R Commander". The menu bar includes "Fichier", "Édition", "Données", "Statistiques", "Graphes", "Modèles", "Distributions", "Outils", and "Aide". Below the menu bar, there are two status fields: "Données: <Pas de données>" and "Modèle: <Pas de modèle>". Between these fields are three buttons: "Éditer" (with a pencil icon), "Visualiser" (with a magnifying glass icon), and "Soumettre" (with a gear icon). The main workspace is divided into three sections: "Script R" and "R Markdown" tabs at the top, a large empty text area in the middle, and a "Sortie" (Output) section at the bottom. The "Messages" section at the very bottom contains the following text: "[2] AVIS: The Windows version of the R Commander works best under RGui with the single-document interface (SDI); see ?Commander."


## Définition du répertoire de travail








Données :

 Éditer

 Visualiser

Modèle :

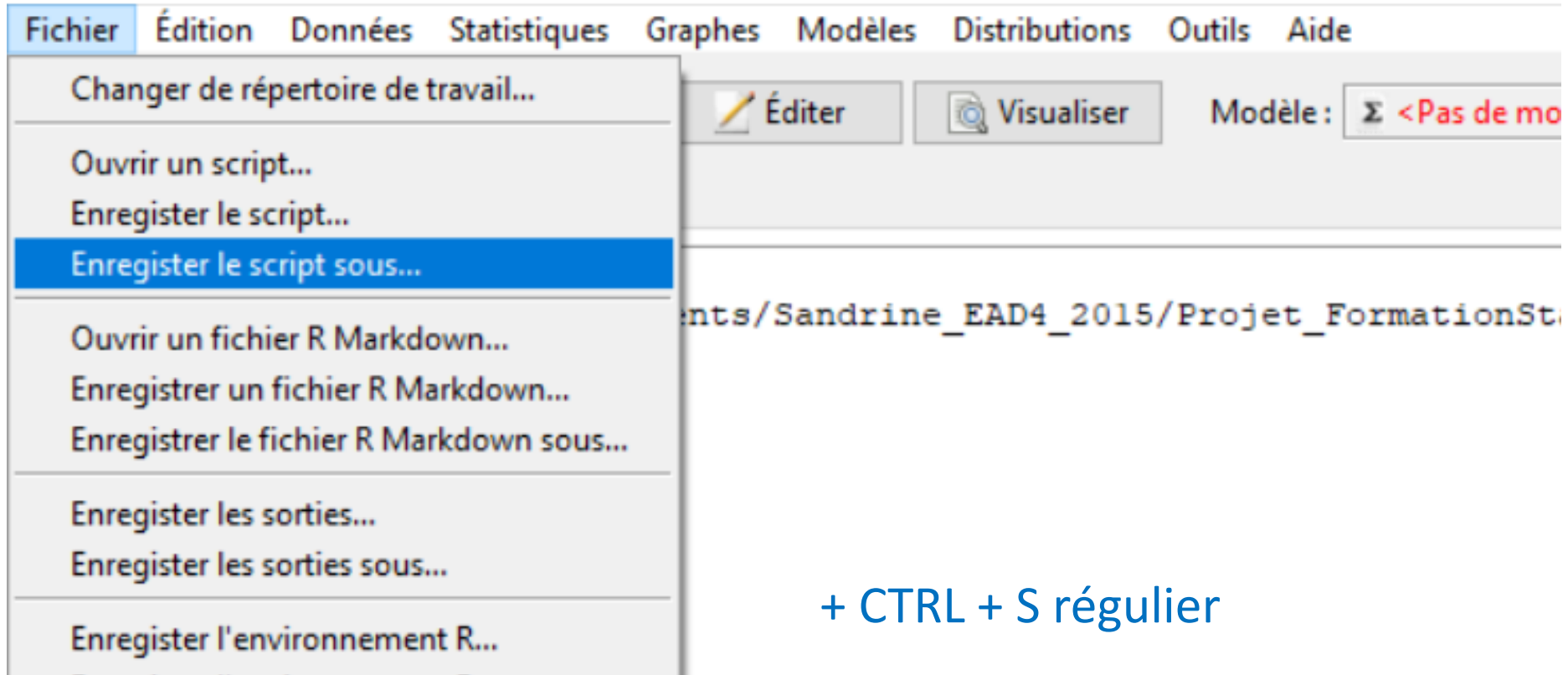
Script R

```
setwd("C:/Users/laguerre/Documents/Sandrine_EAD4_2015/Projet_FormationStatDeBase")
```

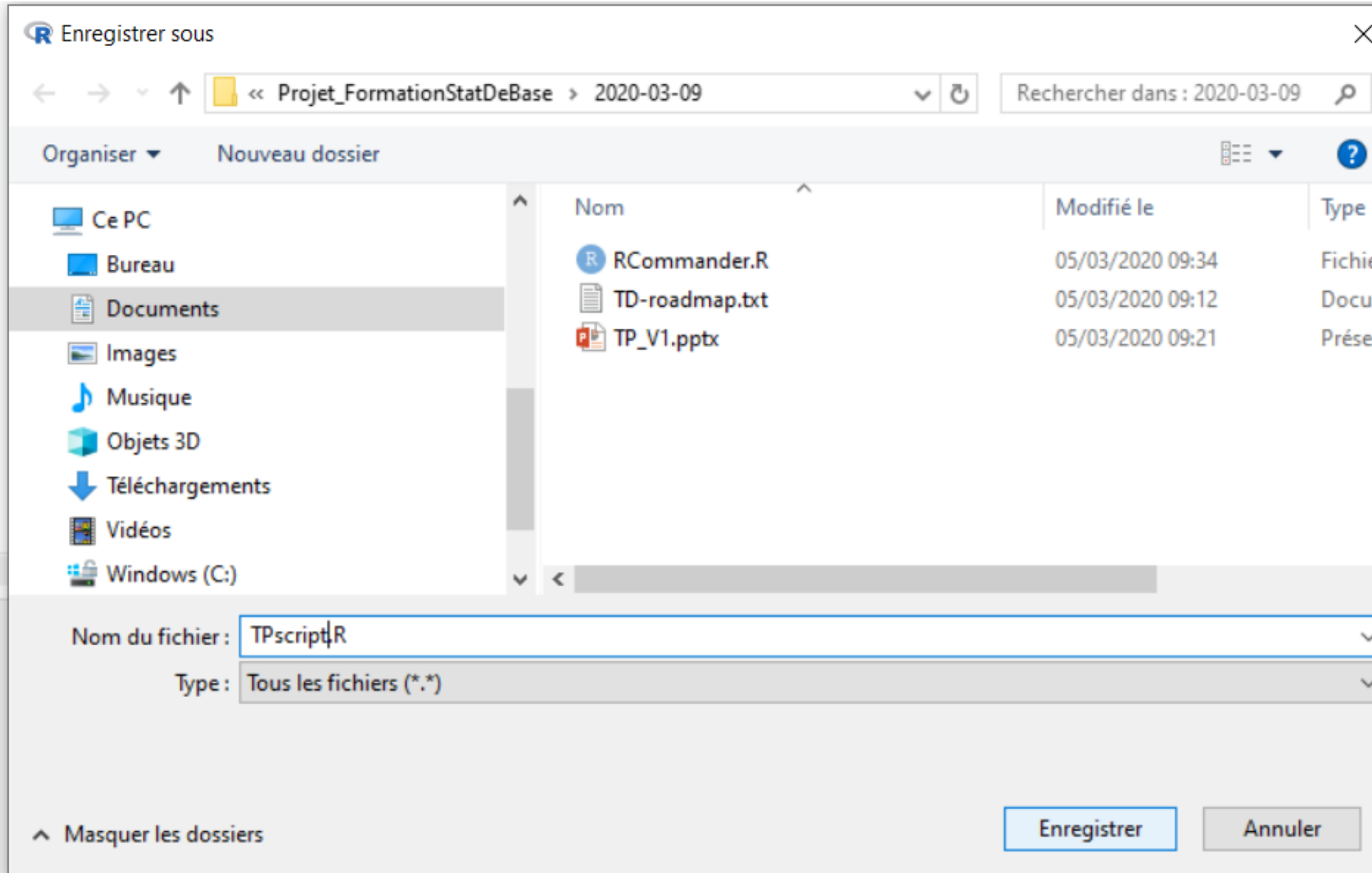


## Sauvegarde du script courant


R Commander




+ CTRL + S régulier



## Importation des données

 C:\Users\laguerre\Documents\Sandrine\_EAD4\_2015\Projet\_FormationStatDeBase\gene.csv - Notepad

```
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?  
  
TD-roadmap.txt x gene.csv x  
1 "X36b4", "ACAT1", "ACAT2", "ACBP", "ACC1", "ACC2", "ACOTH", "ADISP", "AD  
2 -0.42, -0.65, -0.84, -0.34, -1.29, -1.13, -0.93, -0.98, -1.19, -0.68, -0.5  
3 -0.44, -0.68, -0.91, -0.32, -1.23, -1.06, -0.99, -0.97, -1, -0.62, -0.58, -  
4 -0.48, -0.74, -1.1, -0.46, -1.3, -1.09, -1.06, -1.08, -1.18, -0.75, -0.66,  
5 -0.45, -0.69, -0.65, -0.41, -1.26, -1.09, -0.93, -1.02, -1.07, -0.71, -0.6  
6 -0.42, -0.71, -0.54, -0.38, -1.21, -0.89, -1, -0.95, -1.08, -0.76, -0.59, -  
7 -0.43, -0.69, -0.8, -0.32, -1.13, -0.79, -0.93, -0.97, -1.07, -0.75, -0.55  
8 -0.53, -0.62, -1, -0.44, -1.22, -1, -0.94, -0.94, -1.05, -0.67, -0.66, -0.0  
9 -0.49, -0.69, -0.91, -0.37, -1.29, -1.06, -1.05, -1.02, -1.16, -0.75, -0.6  
10 -0.36, -0.66, -0.74, -0.39, -1.15, -1.08, -0.88, -0.98, -1.05, -0.66, -0.5  
11 -0.5, -0.62, -0.79, -0.36, -1.21, -0.82, -0.92, -0.99, -1, -0.69, -0.62, -0  
12 -0.4, -0.6, -0.55, -0.25, -1.22, -1.13, -0.81, -0.99, -1.1, -0.77, -0.52, -  
13 -0.52, -0.66, -0.66, -0.41, -1.28, -1.1, -0.95, -1.04, -1.14, -0.79, -0.72  
14 -0.52, -0.63, -0.99, -0.43, -1.24, -0.96, -0.96, -0.96, -1.06, -0.78, -0.6  
15 -0.47, -0.71, -0.44, -0.45, -1.44, -1.17, -1.02, -1.02, -1.19, -0.88, -0.6  
16 -0.42, -0.66, -0.88, -0.33, -1.24, -1.03, -0.92, -0.94, -1.02, -0.68, -0.5
```

Fichier Édition **Données** Statistiques Graphes Modèles Distributions Outils Aide

Données Visualiser Modèle:  $\Sigma$  <Pas de modèle>

Script R R Markdown

```
setwd("C:/Us
```

- Nouveau jeu de données...
- Charger un jeu de données...
- Fusionner des jeux de données...
- Importer des données**
  - depuis un fichier texte, le presse-papier ou une URL...
  - depuis des données SPSS..
  - depuis un fichier SAS xport...
  - depuis un fichier SAS b7dat...
  - depuis des données Minitab...
  - depuis des données STATA...
  - depuis un fichier Excel...
- Données dans les paquets
- Jeu de données actif
- Gérer les variables du jeu de données actif

R Lire des données depuis un fichier, le presse-papier... ✕

Nom du tableau de données :

Noms de variables dans le fichier :

Indicateur de données manquantes :

Emplacement du fichier de données

Système de fichiers local

Presse-papier

Lien internet (URL)

Séparateur de champs

Espaces  Virgules [,]

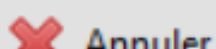
Semicolons [:]  Tabulations

Autre Spécifiez :

Séparateur décimal

Point [.]

Virgule [,]



```
setwd("C:/Users/laguerre/Documents/Sandrine_EAD4_2015/Projet_FormationStatDeBase")
gene <-
  read.table("C:/Users/laguerre/Documents/Sandrine_EAD4_2015/Projet_FormationStatDeBase/gene.csv",
            header=TRUE, sep=",", na.strings="NA", dec=".", strip.white=TRUE)
```

R Ouvrir ✕

« Sandrine\_EAD4\_2015 » Projet\_FormationStatDeBase > Rechercher dans : Projet\_For...

Organiser Nouveau dossier

	Nom	Modifié le	Type
AMAP			
OneDrive			
Ce PC			
Bureau			
Documents			
Images			
Musique			
Objets 3D			
Téléchargements			
Vidéos			
Windows (C:)			
	2018-03-27_intro_stat	21/08/2019 10:28	Dossier de fichiers
	2019-08-12_intro_stat	21/08/2019 09:49	Dossier de fichiers
	2019-08-23_intro_stat	05/09/2019 13:59	Dossier de fichiers
	2019-09-03_biostat1_class	03/09/2019 09:44	Dossier de fichiers
	2019-09-03_intro_stat-sources	03/09/2019 12:18	Dossier de fichiers
	2020-03-09	05/03/2020 09:36	Dossier de fichiers
	diet.csv	09/09/2019 07:32	Fichier CSV Micro.
	gene.csv	09/09/2019 07:32	Fichier CSV Micro.
	genotype.csv	09/09/2019 07:32	Fichier CSV Micro.
	Histo1.png	09/09/2019 11:38	Fichier PNG
	lipid.csv	09/09/2019 07:32	Fichier CSV Micro.
	logfilename	10/09/2019 13:43	Fichier

Nom du fichier :  Tous les fichiers (\*.\*)

Ouvrir Annuler

Ajouter une ligne

Ajouter une colonne

		1	2	3	4	5	6	7
	rowname	X36b4	ACAT1	ACAT2	ACBP	ACC1	ACC2	ACOTH
1	1	-0.42	-0.65	-0.84	-0.34	-1.29	-1.13	-0.93
2	2	-0.44	-0.68	-0.91	-0.32	-1.23	-1.06	-0.99
3	3	-0.48	-0.74	-1.1	-0.46	-1.3	-1.09	-1.06
4	4	-0.45	-0.69	-0.65	-0.41	-1.26	-1.09	-0.93
5	5	-0.42	-0.71	-0.54	-0.38	-1.21	-0.89	-1
6	6	-0.43	-0.69	-0.8	-0.32	-1.13	-0.79	-0.93
7	7	-0.53	-0.62	-1	-0.44	-1.22	-1	-0.94
8	8	-0.49	-0.69	-0.91	-0.37	-1.29	-1.06	-1.05
9	9	-0.36	-0.66	-0.74	-0.39	-1.15	-1.08	-0.88
10	10	-0.5	-0.62	-0.79	-0.36	-1.21	-0.82	-0.92
11	11	-0.4	-0.6	-0.55	-0.25	-1.22	-1.13	-0.81
12	12	-0.52	-0.66	-0.66	-0.41	-1.28	-1.1	-0.95
13	13	-0.52	-0.63	-0.99	-0.43	-1.24	-0.96	-0.96
14	14	-0.47	-0.71	-0.44	-0.45	-1.44	-1.17	-1.02
15	15	-0.42	-0.66	-0.88	-0.33	-1.24	-1.03	-0.92

## **TO DO**

- **Ouvrir successivement les 3 autres fichiers de données**
- **Les importer et les nommer lipid, genotype, diet**
- **Les regarder dans Rcmdr pour vérifier l'import**

R Lire des données depuis un fichier, le presse-papier... ✕

Nom du tableau de données : diet

Noms de variables dans le fichier :

Indicateur de données manquantes : NA

Emplacement du fichier de données

Système de fichiers local

Presse-papier

Lien internet (URL)

Séparateur de champs

Espaces  Virgules [,]

Semicolons [;]  Tabulations

Autre Spécifiez :

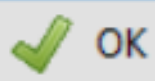
Séparateur décimal

Point [.]

Virgule [,]



Aide



OK

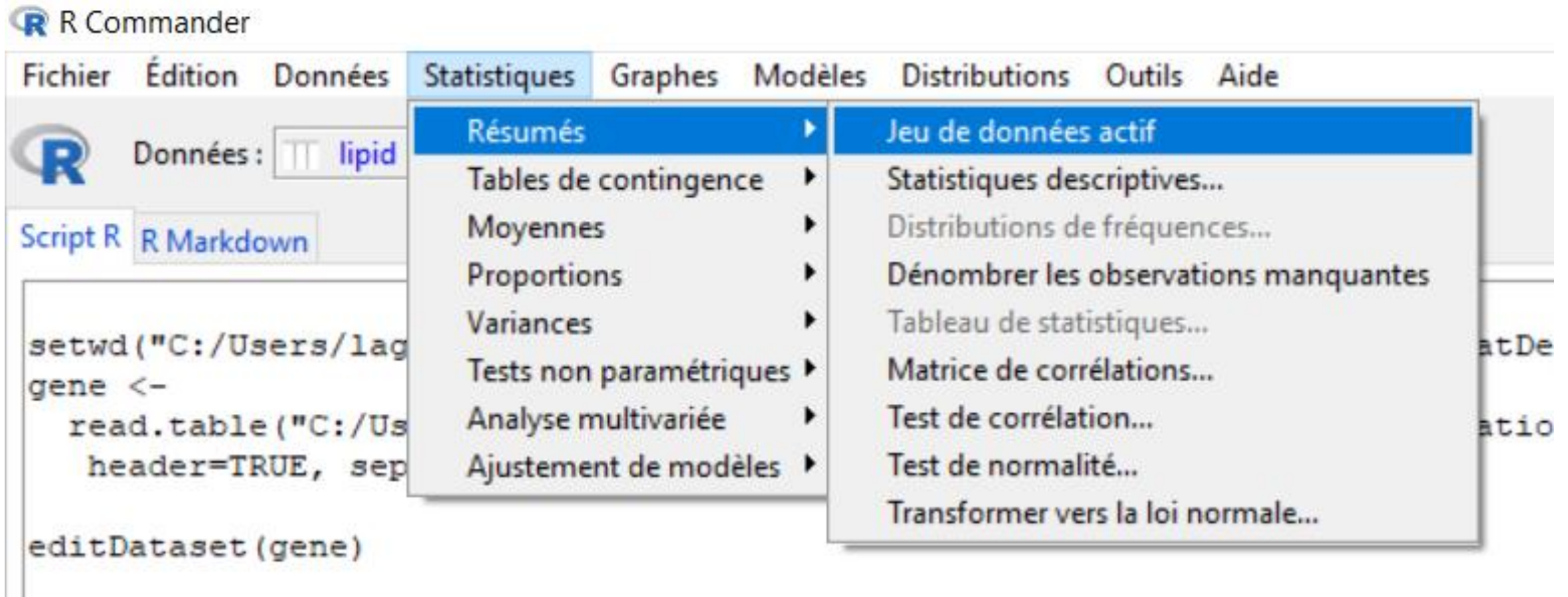


Annuler

Attention pour diet et genotype, il n'y a pas de nom de variables  
Si oubli de décocher, l'effectif n'est pas le bon



## Statistiques de base sur le jeu de données actif



R Commander

Fichier Édition Données **Statistiques** Graphes Modèles Distributions Outils Aide

Données : lipid

Script R R Markdown

```
setwd("C:/Users/lag  
gene <-  
  read.table("C:/Us  
  header=TRUE, sep  
  
editDataset(gene)
```

**Résumés**

- Tables de contingence
- Moyennes
- Proportions
- Variances
- Tests non paramétriques
- Analyse multivariée
- Ajustement de modèles

**Jeu de données actif**

- Statistiques descriptives...
- Distributions de fréquences...
- Dénombrer les observations manquantes
- Tableau de statistiques...
- Matrice de corrélations...
- Test de corrélation...
- Test de normalité...
- Transformer vers la loi normale...

```

Rcmdr> summary(lipid)
  C14.0      C16.0      C18.0      C16.1n.9
Min.   :0.2200  Min.   :14.65  Min.   : 1.680  Min.   :0.2900
1st Qu.:0.3500  1st Qu.:20.65  1st Qu.: 4.540  1st Qu.:0.4575
Median :0.4200  Median :23.64  Median : 6.500  Median :0.6450
Mean   :0.7630  Mean   :23.03  Mean   : 6.747  Mean   :0.6870
3rd Qu.:0.6325  3rd Qu.:25.54  3rd Qu.: 9.065  3rd Qu.:0.8125
Max.   :3.2400  Max.   :29.72  Max.   :10.970  Max.   :1.5000
  C16.1n.7      C18.1n.9      C18.1n.7      C20.1n.9
Min.   : 1.590  Min.   :14.69  Min.   : 1.530  Min.   :0.0000
1st Qu.: 2.632  1st Qu.:20.05  1st Qu.: 2.140  1st Qu.:0.2450
Median : 3.300  Median :23.02  Median : 3.075  Median :0.2850
Mean   : 4.419  Mean   :25.27  Mean   : 4.426  Mean   :0.2840
3rd Qu.: 4.952  3rd Qu.:32.41  3rd Qu.: 4.622  3rd Qu.:0.3725
Max.   :13.900  Max.   :41.23  Max.   :15.030  Max.   :0.6500
  C20.3n.9      C18.2n.6      C18.3n.6      C20.2n.6
Min.   :0.0000  Min.   : 2.31  Min.   :0.0000  Min.   :0.0000
1st Qu.:0.0000  1st Qu.: 8.88  1st Qu.:0.0000  1st Qu.:0.0000
Median :0.0000  Median :14.87  Median :0.1800  Median :0.1250
Mean   :0.3068  Mean   :15.28  Mean   :0.3745  Mean   :0.1852
3rd Qu.:0.2125  3rd Qu.:21.20  3rd Qu.:0.3025  3rd Qu.:0.3025
Max.   :2.8900  Max.   :40.02  Max.   :5.0700  Max.   :0.8300
  C20.3n.6      C20.4n.6      C22.4n.6      C22.5n.6
Min.   :0.110  Min.   : 0.750  Min.   :0.000  Min.   :0.0000
1st Qu.:0.400  1st Qu.: 1.893  1st Qu.:0.000  1st Qu.:0.0000
Median :0.725  Median : 3.710  Median :0.025  Median :0.1700
Mean   :0.776  Mean   : 5.279  Mean   :0.184  Mean   :0.4370
3rd Qu.:1.055  3rd Qu.: 6.270  3rd Qu.:0.310  3rd Qu.:0.4425
Max.   :1.640  Max.   :15.760  Max.   :0.730  Max.   :2.5200
  C18.3n.3      C20.3n.3      C20.5n.3      C22.5n.3
Min.   : 0.000  Min.   :0.000  Min.   :0.00  Min.   :0.0000
1st Qu.: 0.000  1st Qu.:0.000  1st Qu.:0.07  1st Qu.:0.1275
Median : 0.460  Median :0.000  Median :0.25  Median :0.4400
Mean   : 2.888  Mean   :0.091  Mean   :1.79  Mean   :0.8718
3rd Qu.: 1.310  3rd Qu.:0.070  3rd Qu.:2.84  3rd Qu.:1.6750
Max.   :21.620  Max.   :0.640  Max.   :9.48  Max.   :2.5800
  C22.6n.3
Min.   : 0.280
1st Qu.: 1.853
Median : 3.355
Mean   : 5.914
3rd Qu.:10.393
Max.   :17.350

```

RcmdrMsg: [19] NOTE: Script enregistré dans

RcmdrMsg+ C:/Users/laguerr/Document/Sandrine\_EAD4\_2015/Projet\_FormationStatDeBase/2020-03-09/TPscript.R

## **TO DO**

- **Statistiques descriptives sur diet avant de revenir à lipid**

R Commander

Fichier Édition Données **Statistiques** Graphes Modèles Distributions Outils Aide

Données :  diet

Script R R Markdown

```
read.table("C:/Us  
header=TRUE, sep
```

```
editDataset(gene)
```

```
lipid <-
```

```
read.table("C:/Users/laguerre/Documents/Sandrine EAD4 2015/Projet Formati
```

Résumés ▶

Tables de contingence ▶

Moyennes ▶

Proportions ▶

Variances ▶

Tests non paramétriques ▶

Analyse multivariée ▶

Ajustement de modèles ▶

Jeu de données actif

Statistiques descriptives...

Distributions de fréquences...

Dénombrer les observations manquantes

Tableau de statistiques...

Matrice de corrélations...

Test de corrélation...

Test de normalité...

Transformer vers la loi normale...

```
RcmdrMsg: [19] NOTE: Script enregistré dans  
RcmdrMsg+ C:/Users/laguerr/Document/Sandrine_EAD4_2015/Projet_FormationStatDeBase/2020-03-09/TPscript.R  
RcmdrMsg: [20] NOTE: Le jeu de données diet a 40 lignes et 1 colonnes.
```

```
Rcmdr> summary(diet)
```

```
  v1
```

```
 coc :8
```

```
 fish:8
```

```
 lin :8
```

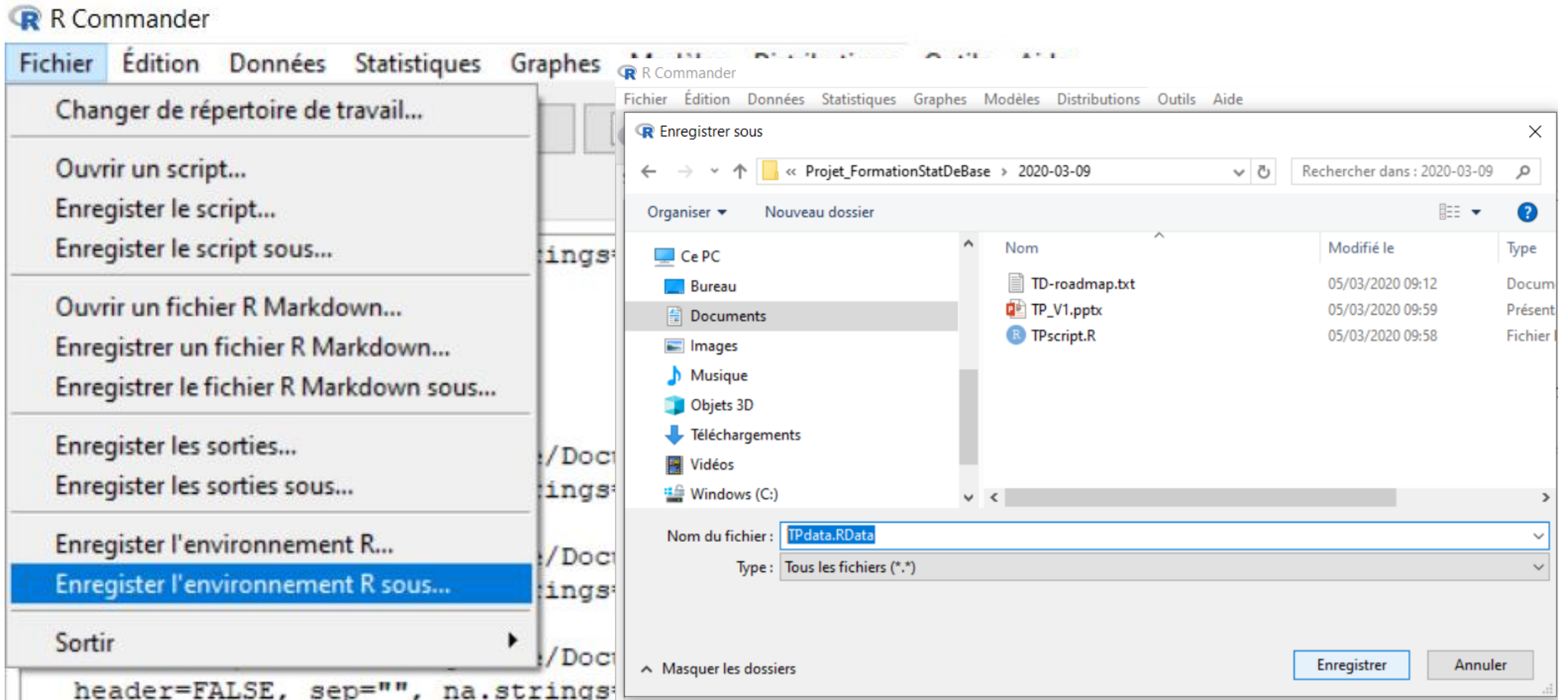
```
 ref :8
```

```
 sun :8
```

```
>
```



# Sauvegarde des données sous la forme Rdata



```
header=FALSE, sep="", na.strings  
editDataset(diet)  
editDataset(genotype)
```

.Rdata, .Rdata, .rdatt, .rdata = différentes extensions pour le même fichier binaire que l'on ne peut ouvrir qu'avec R

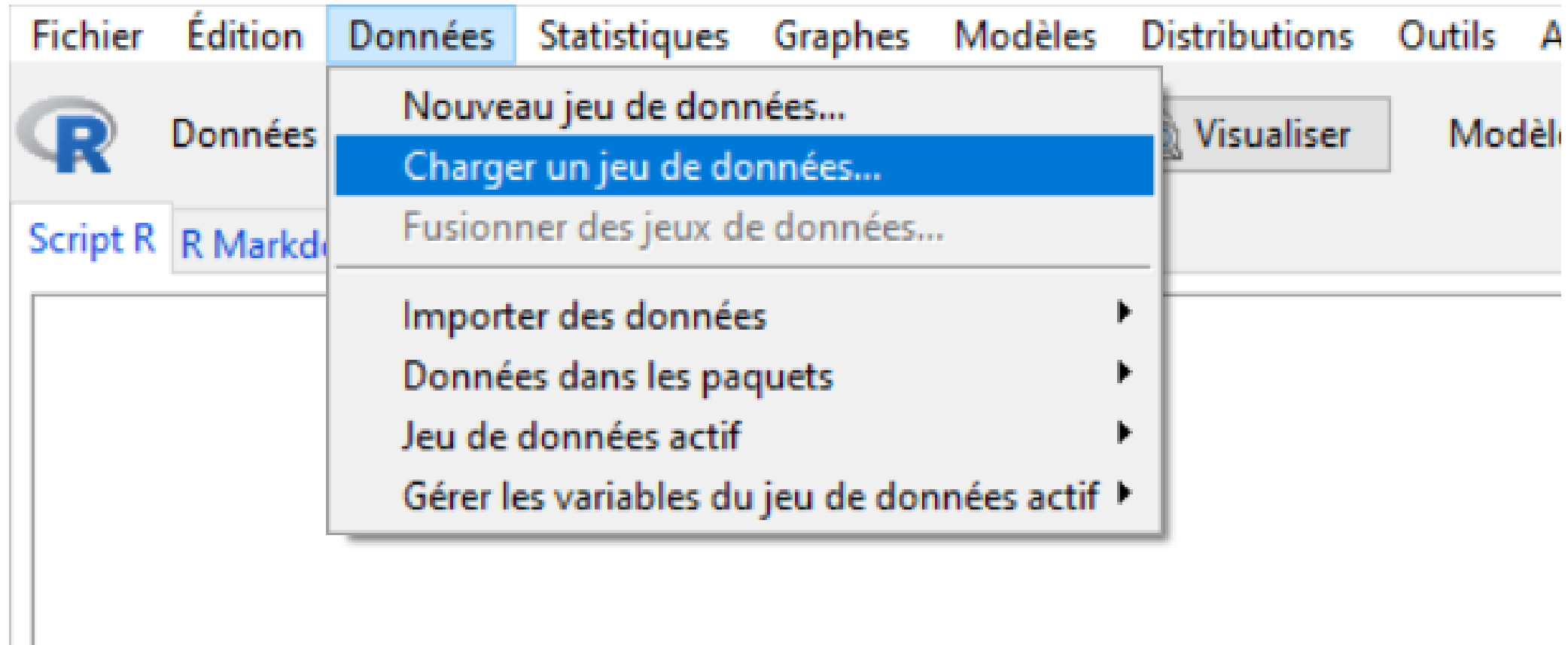
# **Partie 2 : Statistiques univariées**

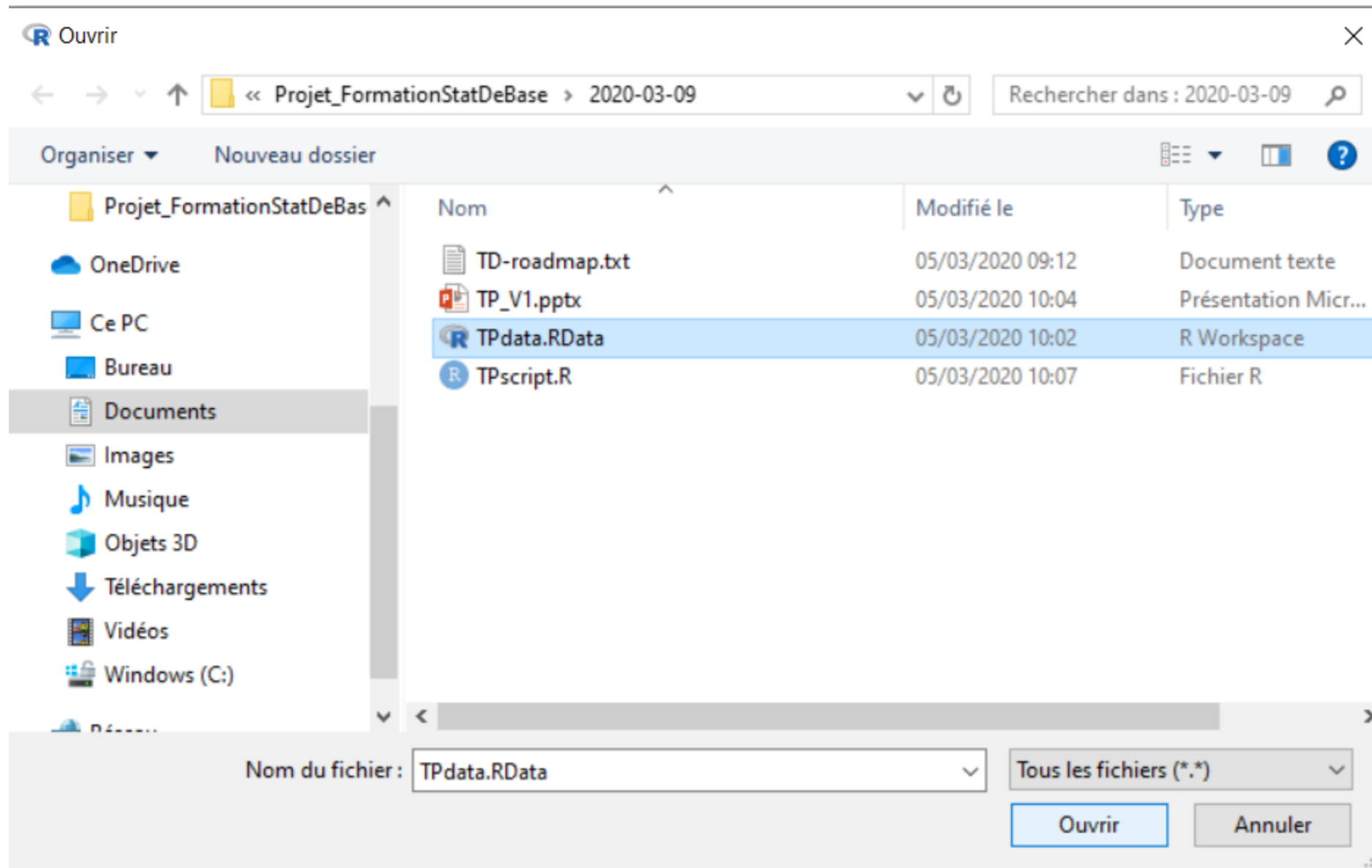
# **Partie 2.1 : Variables quantitatives (numériques)**

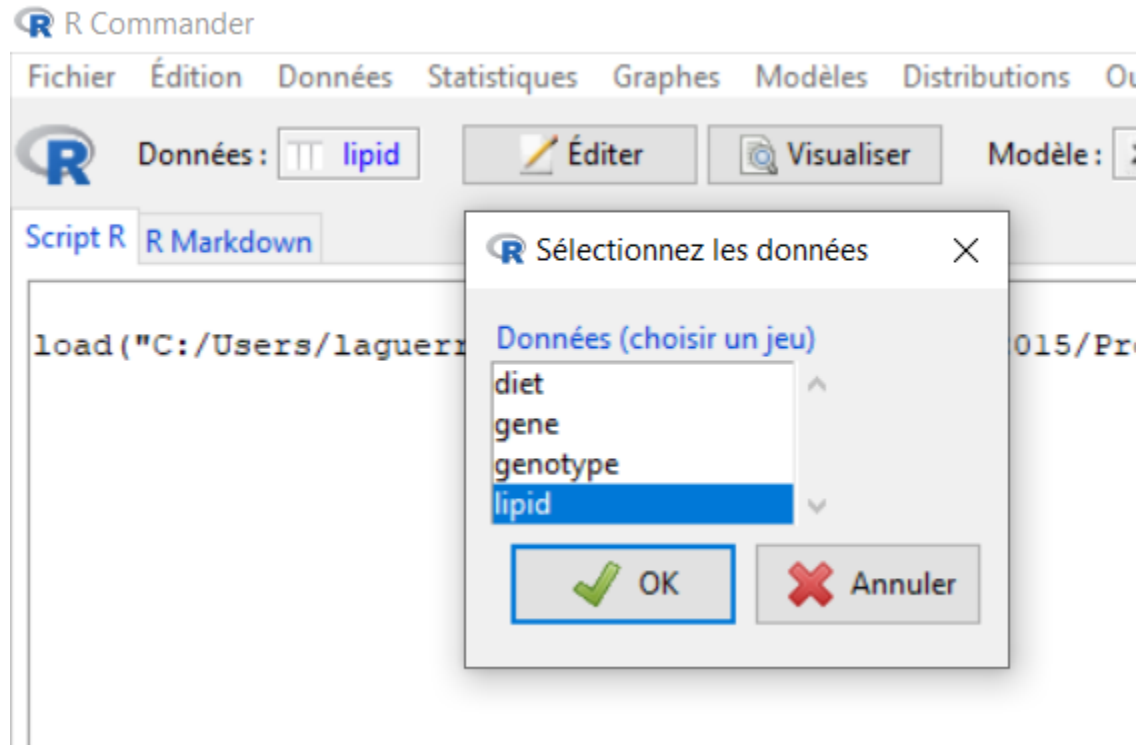


## Charger un jeu de données Rdata

 R Commander

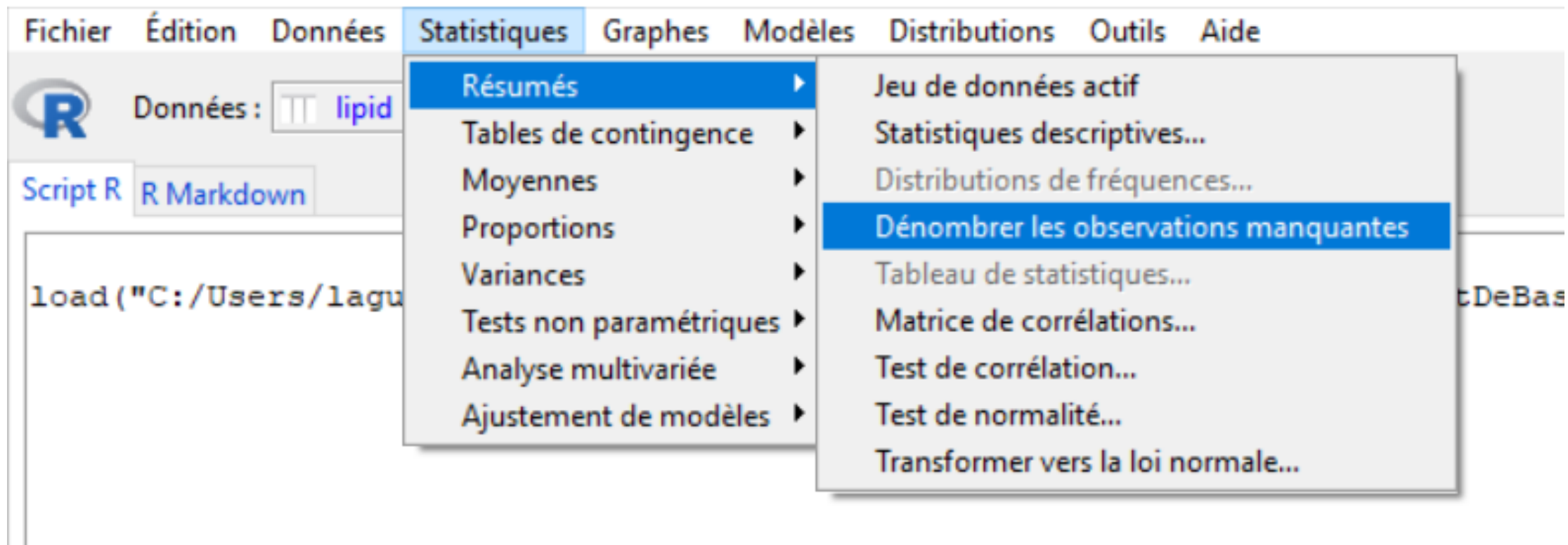






## Dénombrer les observations manquantes

R Commander

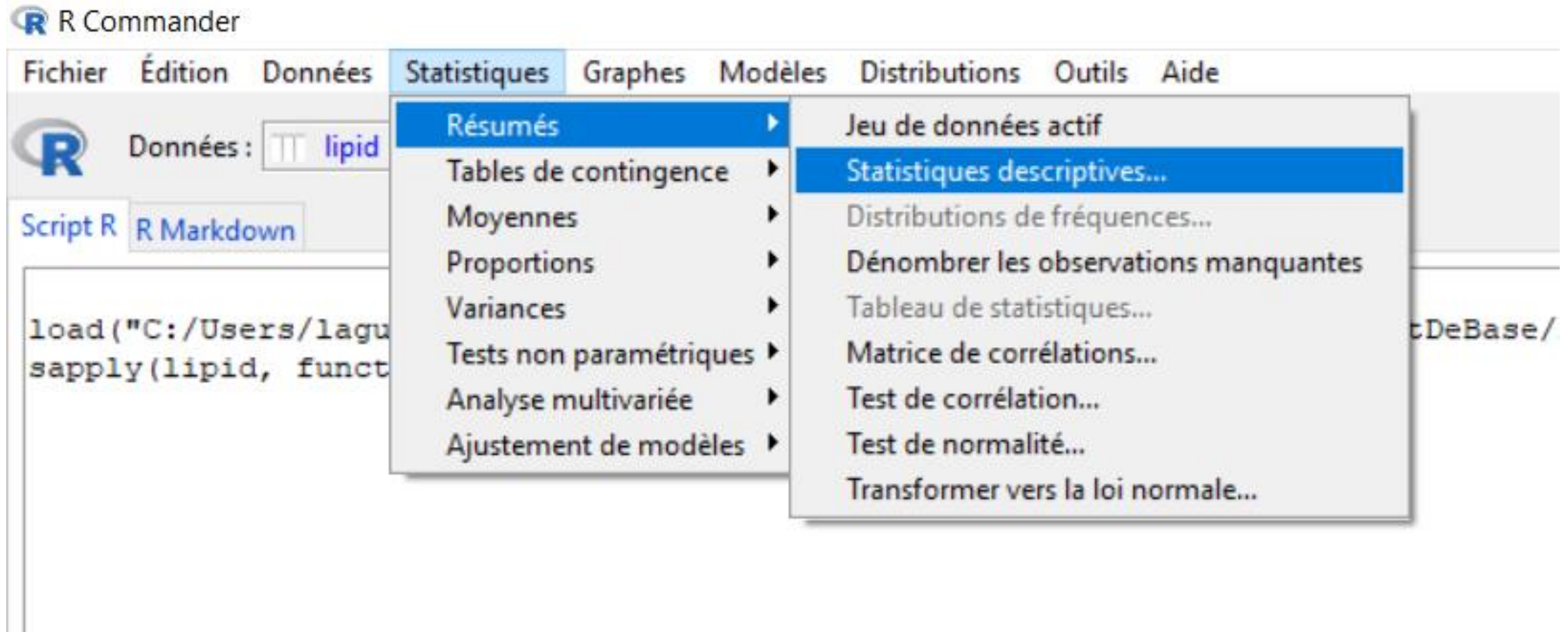


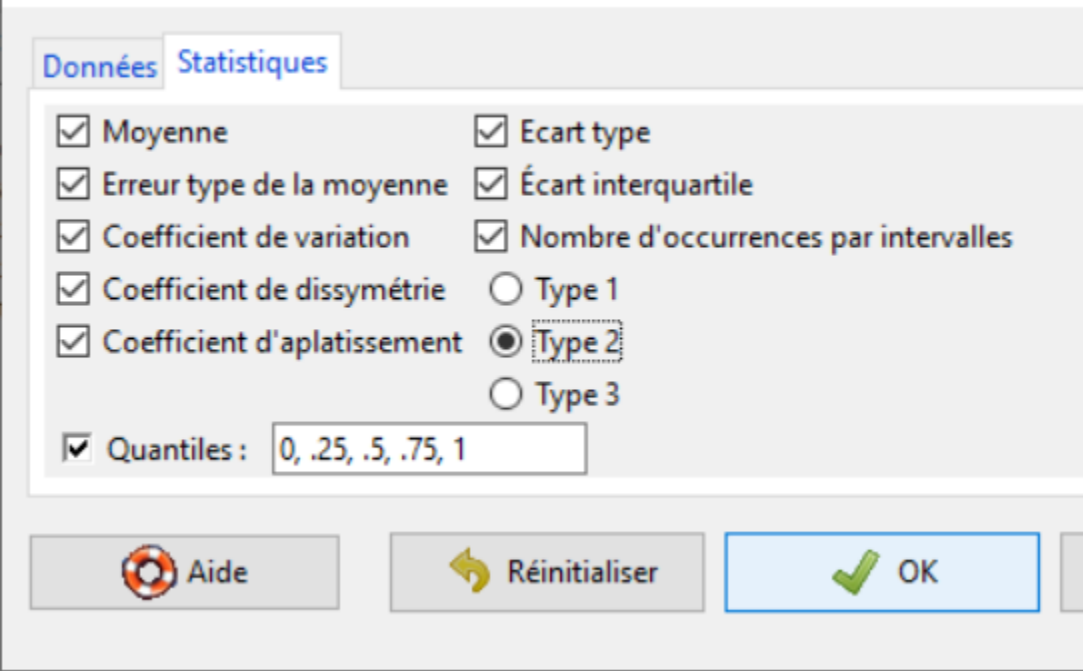
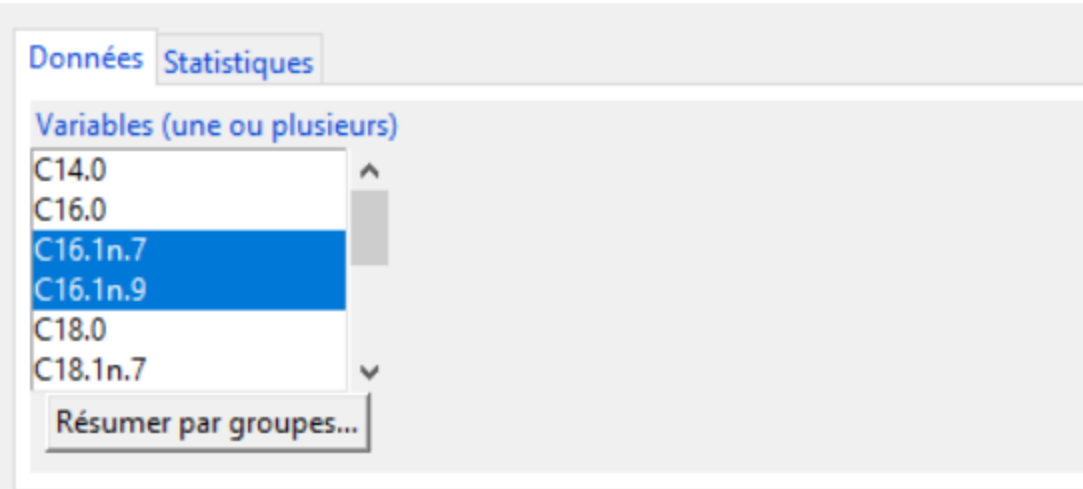
The screenshot shows the R Commander interface. The menu bar includes 'Fichier', 'Édition', 'Données', 'Statistiques', 'Graphes', 'Modèles', 'Distributions', 'Outils', and 'Aide'. The 'Statistiques' menu is open, showing options like 'Résumés', 'Tables de contingence', 'Moyennes', 'Proportions', 'Variances', 'Tests non paramétriques', 'Analyse multivariée', and 'Ajustement de modèles'. The 'Résumés' sub-menu is also open, with 'Dénombrer les observations manquantes' highlighted. Other options in this sub-menu include 'Jeu de données actif', 'Statistiques descriptives...', 'Distributions de fréquences...', 'Tableau de statistiques...', 'Matrice de corrélations...', 'Test de corrélation...', 'Test de normalité...', and 'Transformer vers la loi normale...'. The 'Données' pane shows a variable named 'lipid'. The 'Script R' pane contains the command `load("C:/Users/lagu`.

```
Rcmdr> sapply(lipid, function(x)(sum(is.na(x)))) # NA counts
  c14.0   c16.0   c18.0 c16.1n.9 c16.1n.7 c18.1n.9 c18.1n.7 c20.1n.9
    0      0      0      0      0      0      0      0
c20.3n.9 c18.2n.6 c18.3n.6 c20.2n.6 c20.3n.6 c20.4n.6 c22.4n.6 c22.5n.6
    0      0      0      0      0      0      0      0
c18.3n.3 c20.3n.3 c20.5n.3 c22.5n.3 c22.6n.3
    0      0      0      0      0
```

>

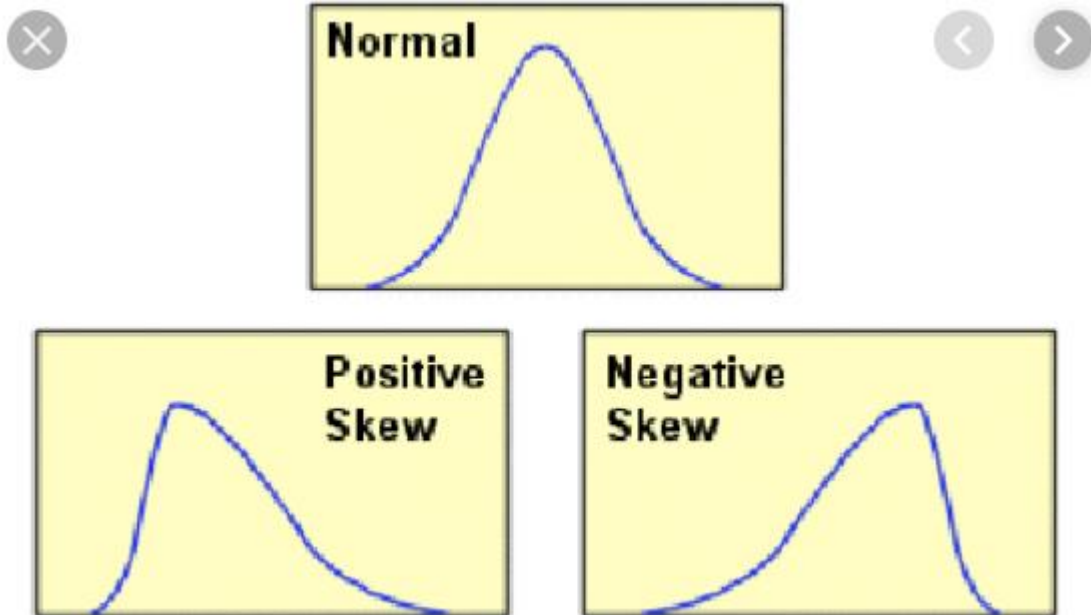
# Statistiques descriptives





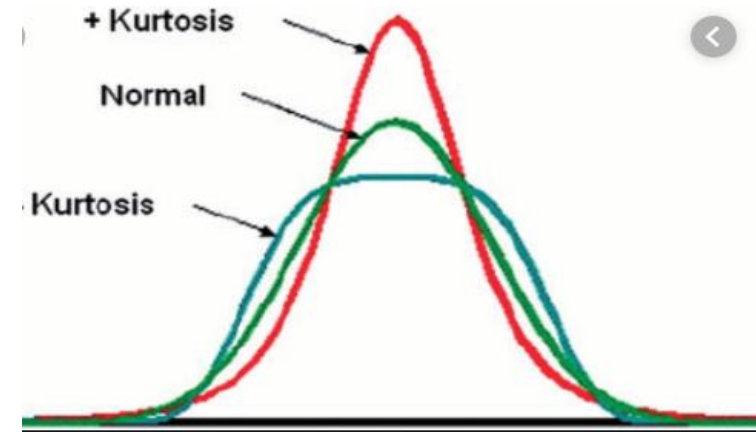
- Ecart-type = SD : mesure de la dispersion des données autour de la moyenne
- Erreur type de la moyenne = SEM : mesure de la précision de l'estimation de la moyenne
- $SEM = SD/\sqrt{n}$
- Ecart interquartile =  $Q_{75} - Q_{25}$
- Coefficient de variation (CV) =  $SD / \text{moyenne}$

## Coefficient de dissymétrie (skewness)



- Positive skew : moyenne  $>$  médiane
- Negative skew : moyenne  $<$  médiane

## Coefficient d'aplatissement (kurtosis)



- Dist. normale : kurtosis = 0
- Kurtosis  $> 0$  : distribution plus pointue (leptokurtic)
- Kurtosis  $< 0$  : distribution plus plate (platykurtic)

```
Rcmdr> numSummary(lipid[,c("c16.1n.7", "c16.1n.9"), drop=FALSE],
Rcmdr+   statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "cv", "skewness",
Rcmdr+   "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")
      mean      sd  se(mean)  IQR      cv skewness kurtosis  0%
c16.1n.7 4.41875 2.9849702 0.47196523 2.320 0.6755237 1.9309189 3.1882201 1.59
c16.1n.9 0.68700 0.2849849 0.04506007 0.355 0.4148252 0.9524791 0.5377586 0.29
      25%  50%  75% 100%  n
c16.1n.7 2.6325 3.300 4.9525 13.9 40
c16.1n.9 0.4575 0.645 0.8125  1.5 40
```

```
Rcmdr> binnedCounts(lipid[,c("c16.1n.7", "c16.1n.9"), drop=FALSE])
```

Binned distribution of c16.1n.7

	Count	Percent
[0, 2]	1	2.5
(2, 4]	26	65.0
(4, 6]	5	12.5
(6, 8]	4	10.0
(8, 10]	0	0.0
(10, 12]	2	5.0
(12, 14]	2	5.0
Total	40	100.0

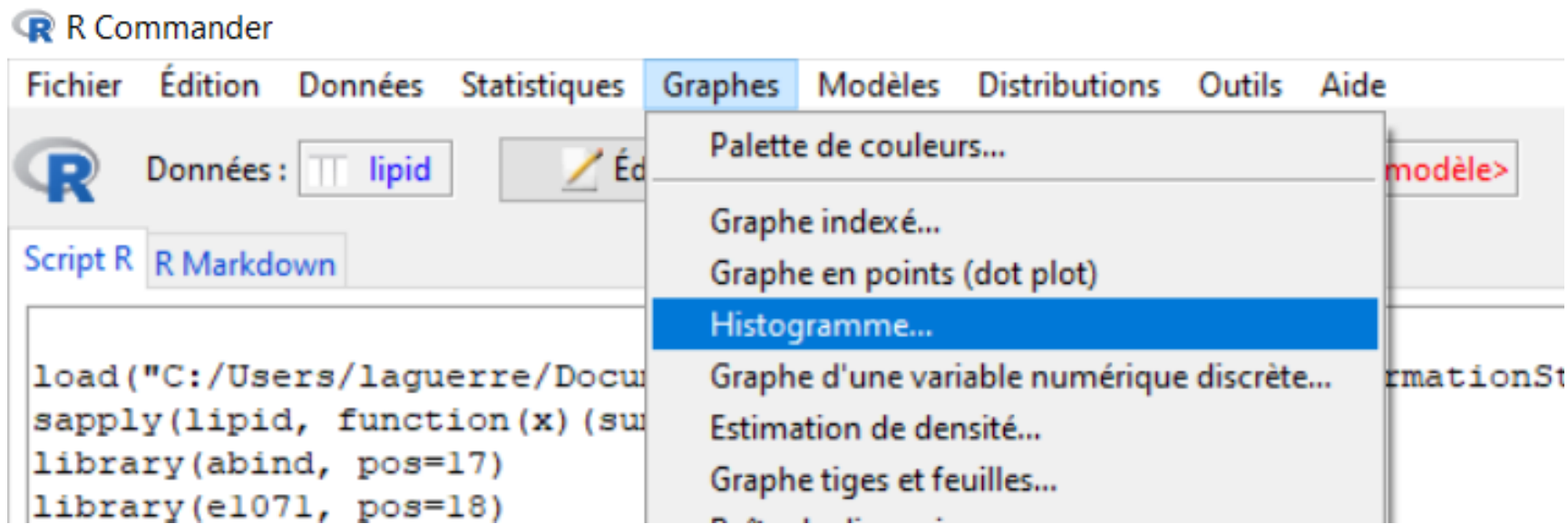
Binned distribution of c16.1n.9

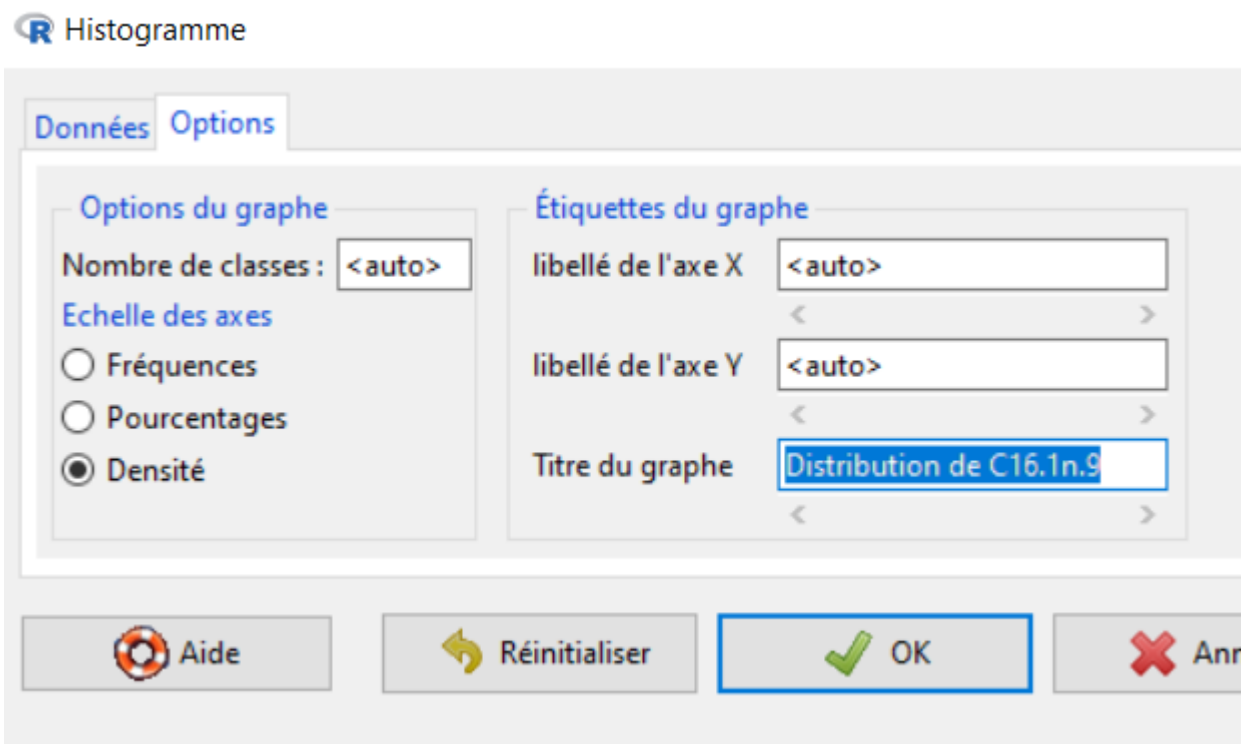
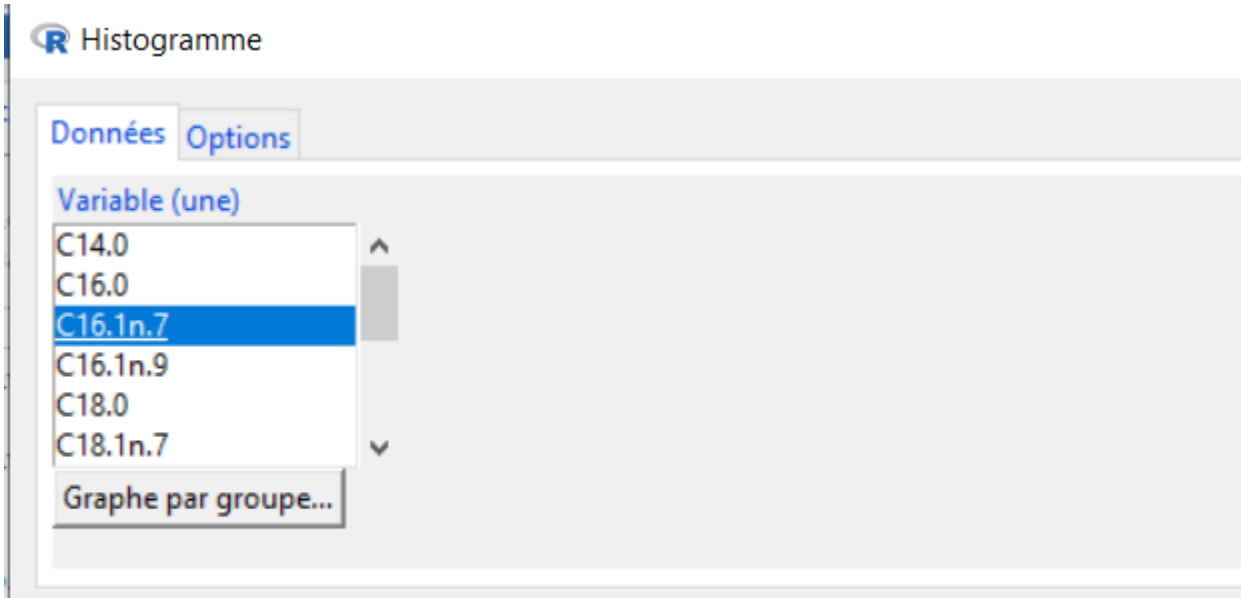
	Count	Percent
[0.2, 0.4]	5	12.5
(0.4, 0.6]	13	32.5
(0.6, 0.8]	11	27.5
(0.8, 1]	5	12.5
(1, 1.2]	3	7.5
(1.2, 1.4]	2	5.0
(1.4, 1.6]	1	2.5
Total	40	100.0

- Pour les 2 variables positive skew : moyenne > médiane
- Pour les 2 variables : positive kurtosis : distribution plus pointue que la loi normale



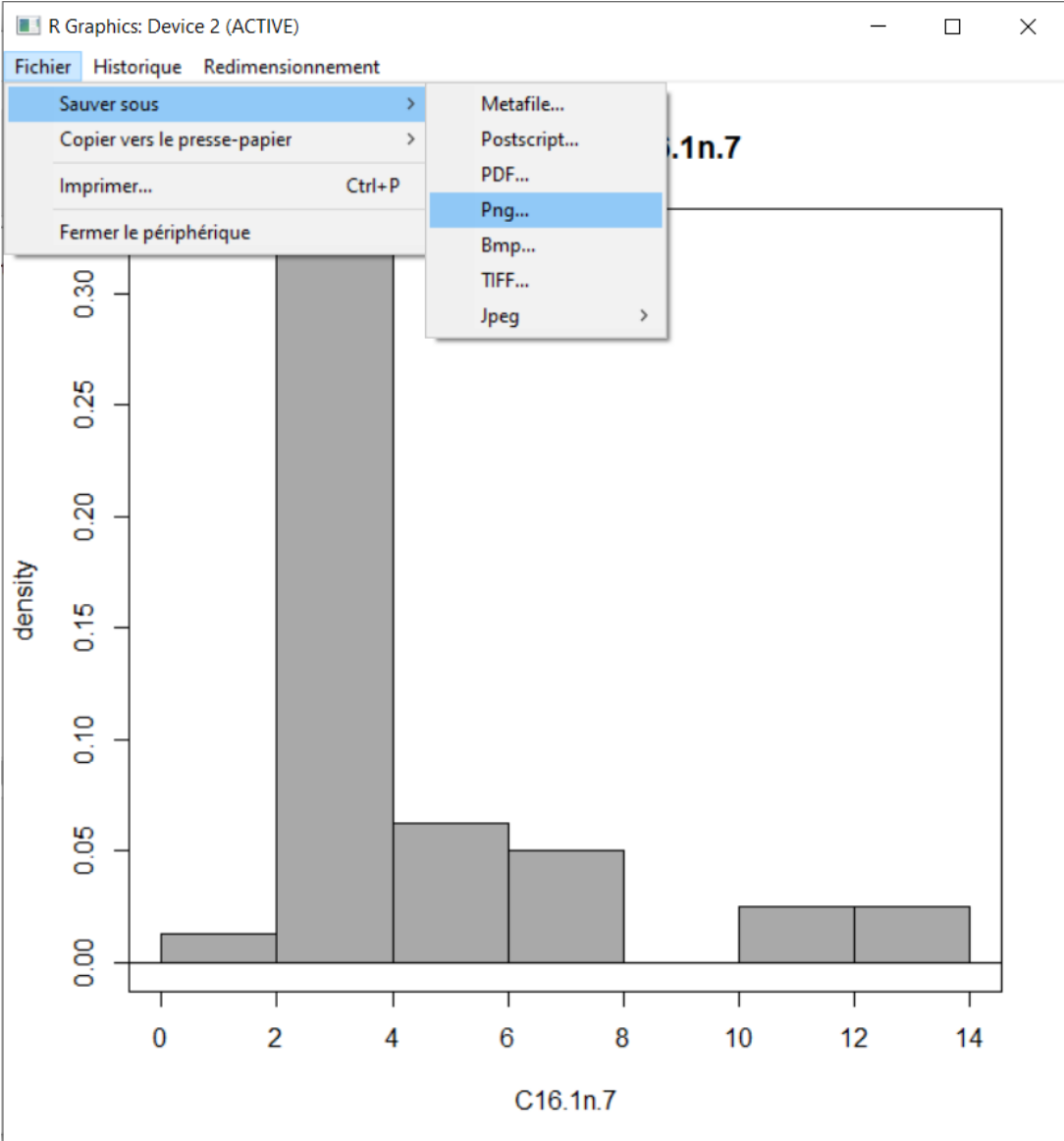
# Histogramme





- Sur un histogramme en fréquence, l'axe vertical affiche le nombre d'occurrence dans une classe
- Sur un histogramme en densité, l'axe vertical permet que l'aire de l'ensemble des barres fasse 1

# Enregistrer un graphique



## **TO DO**

- **Faire l'histogramme de C16.1n.9**

Données: lipid Éd

Script R R Markdown

```
load("C:/Users/laguerr/Docu
sapply(lipid, function(x) su
library(abind, pos=17)
library(el071, pos=18)
numSummary(lipid[,c("C16.1n.
  statistics=c("mean", "sd",
numSummary(lipid[,c("C16.1n.
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
binnedCounts(lipid[,c("C16.1
numSummary(lipid[,c("C16.1n.
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
binnedCounts(lipid[,c("C16.1
numSummary(lipid[,c("C16.1n.
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
binnedCounts(lipid[,c("C16.1
```

Palette de couleurs...

Graphe indexé...

Graphe en points (dot plot)

**Histogramme...**

Graphe d'une variable numérique discrète...

Estimation de densité...

Graphe tiges et feuilles...

Boîte de dispersion...

Graphe quantile-quantile...

Boîte à moustaches pour symétrie...

Nuage de points...

Matrice de nuages de points...

Graphe en lignes...

Graphe XY conditionnel...

Graphe des moyennes...

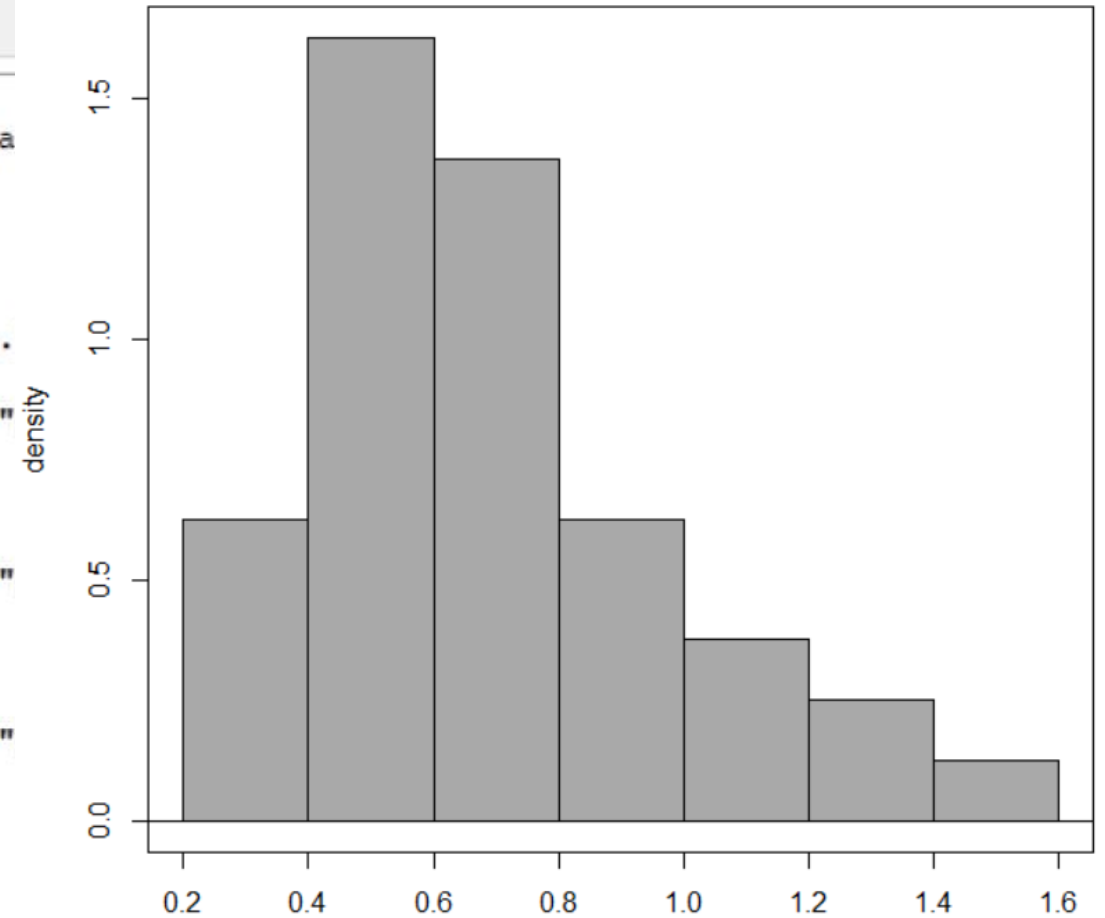
Graphe en bande...

Graphe en barres...

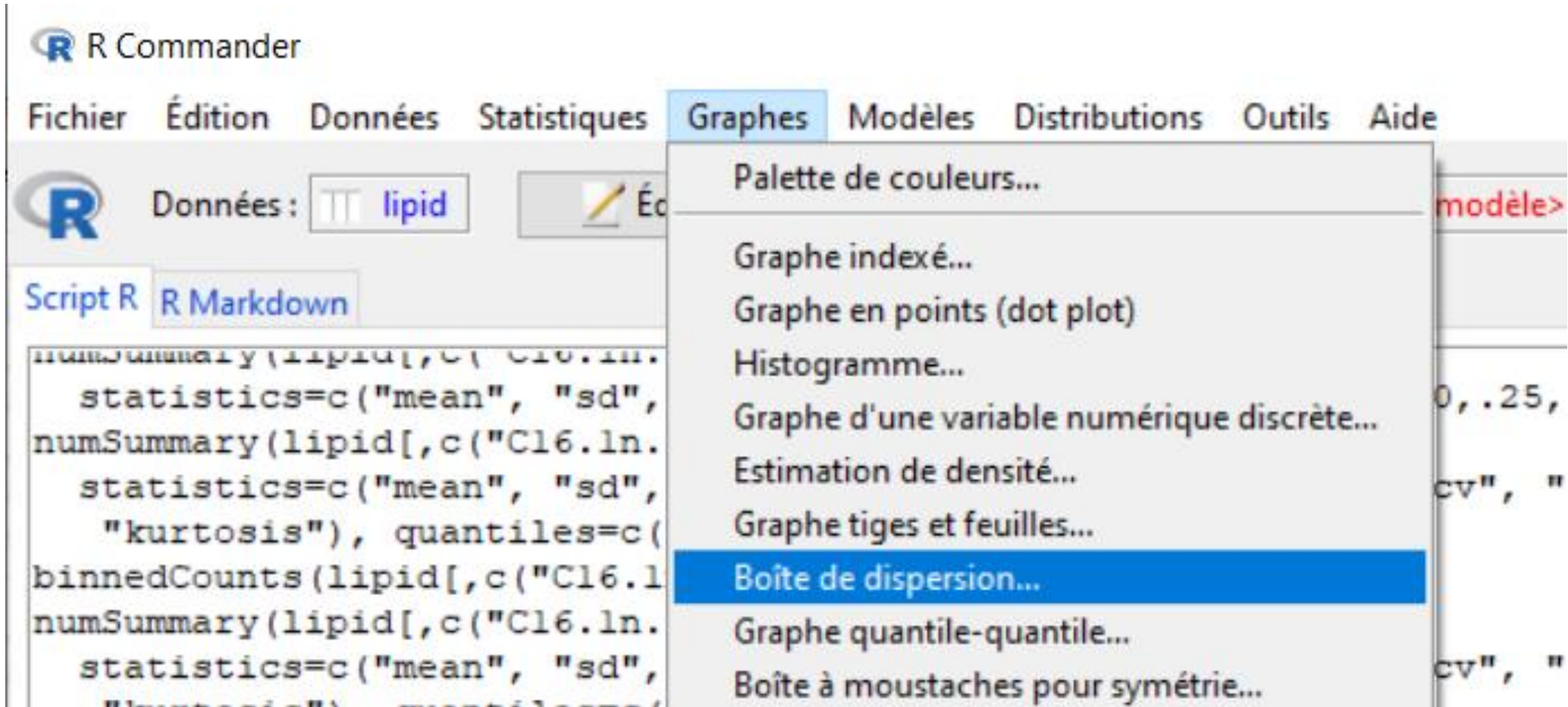
Graphe en camembert...

Graphe 3D ▶

Enregistrer le graphe dans un fichier... ▶

**Distribution de C16.1n.9**

## Boxplot : boite de dispersion



The image shows the R Commander interface. The menu bar includes 'Fichier', 'Édition', 'Données', 'Statistiques', 'Graphes', 'Modèles', 'Distributions', 'Outils', and 'Aide'. The 'Graphes' menu is open, displaying options such as 'Palette de couleurs...', 'Graphe indexé...', 'Graphe en points (dot plot)', 'Histogramme...', 'Graphe d'une variable numérique discrète...', 'Estimation de densité...', 'Graphe tiges et feuilles...', 'Boîte de dispersion...' (highlighted), 'Graphe quantile-quantile...', and 'Boîte à moustaches pour symétrie...'. The 'Données' pane shows 'lipid' as the active dataset. The 'Script R' pane contains R code for summarizing data, including `statistics=c("mean", "sd", "kurtosis")` and `quantiles=c(...)`.

Données Options

Variable (une)

C14.0  
C16.0  
**C16.1n.7**  
C16.1n.9  
C18.0  
C18.1n.7

Graphe par groupe...

Données Options

Identifier les valeurs aberrantes

- Automatiquement**  
 À la souris  
 Non

Étiquettes du graphe

libellé de l'axe X &lt; auto &gt;

&lt; &gt;

libellé de l'axe Y &lt; auto &gt;

&lt; &gt;

Titre du graphe &lt; auto &gt;

&lt; &gt;



Aide



Réinitialiser



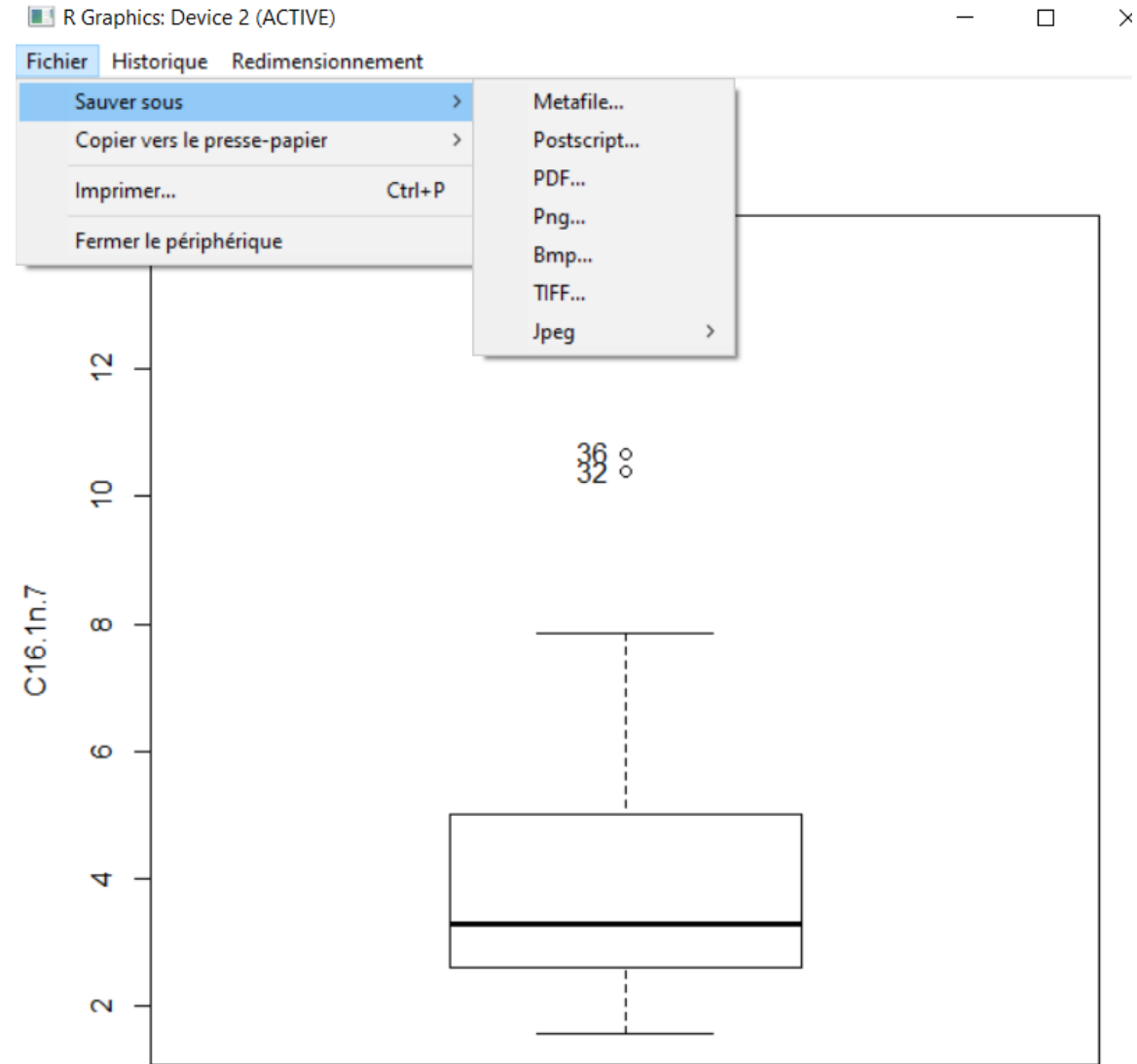
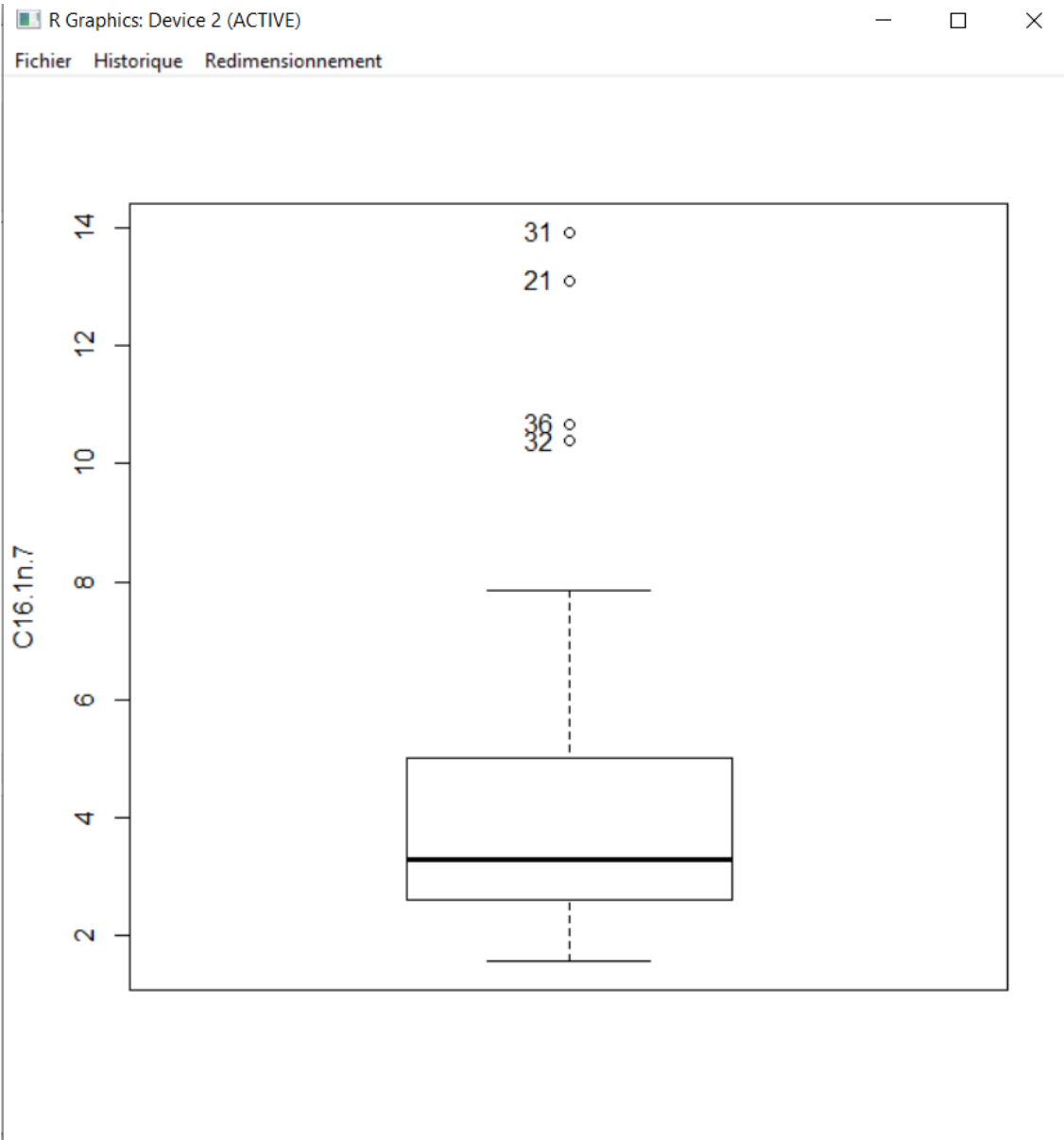
OK



Annuler



Appliquer





# Ajout de variables standardisées

The image shows the R Commander interface. The 'Données' menu is open, with 'Gérer les variables du jeu de données actif' selected. A sub-menu is open, showing 'Standardiser des variables...' as the selected option. Below, the 'Standardiser des variables' dialog box is open, showing a list of variables with 'C16.1n.7' and 'C16.1n.9' selected. The 'OK' button is highlighted.

R Commander

Fichier Édition **Données** Statistiques Graphes Modèles Distributions Outils Aide

Données

Modèle:  $\Sigma$  <Pas de modèle>

Script R R Markd

statistics  
numSummary(  
statistics  
"kurtosis  
binnedCounts  
numSummary(lipid[,c("C16.1n.7", "C16.1n.9"),  
se(mean)", "I  
.25, .5, .75, 1),  
7", "C16.1n.9"  
, "C16.1n.9"),

Gérer les variables du jeu de données actif

- Recoder des variables...
- Calculer une nouvelle variable...
- Ajouter les numéros d'observation au jeu de données...
- Standardiser des variables...**
- Convertir des variables numériques en facteurs...

Standardiser des variables

Variables (une ou plusieurs)

- C14.0
- C16.0
- C16.1n.7**
- C16.1n.9**
- C18.0
- C18.1n.7

Aide OK Annuler

## TO DO

- **Comment s'appellent les nouvelles variables et où sont elles situées?**
- **Calculez les statistiques de base sur ces variables. Que sont-elles?**

R lipid

— □ ×

	n.6	C22.5n.6	C18.3n.3	C20.3n.3	C20.5n.3	C22.5n.3	C22.6n.3	Z.C16.1n.7	Z.C16.1n.9
1	.00	0.00	5.97	0.37	8.62	1.75	10.39	-0.4417967	-1.18251867
2	.58	2.10	0.00	0.00	0.00	0.48	2.61	-0.6294033	-0.48072717
3	.54	1.77	0.00	0.00	0.00	0.22	2.51	-0.5925520	-0.48072717
4	.00	0.00	0.49	0.00	2.99	1.04	14.99	-0.5356000	-0.69126462
5	.32	0.44	0.42	0.00	0.30	0.35	6.69	-0.5255496	-0.79653335
6	.00	0.56	0.00	0.00	0.00	2.13	2.56	0.9518520	-0.09474185
7	.00	0.00	8.40	0.42	7.37	2.05	9.84	-0.2742908	-1.14742909
8	.00	0.00	6.01	0.39	7.96	2.33	10.40	-0.3848447	-1.39305612
9	.00	0.00	0.55	0.00	3.13	1.65	16.36	-0.6897054	-0.86671250
10	.00	0.39	0.00	0.00	0.00	0.00	1.86	0.8680991	0.74740794

Données : lipid

Script R R Markdown

```
binCounts(lipid[,c(
numSummary(lipid[,c
statistics=c("mea
"kurtosis"), qua
binCounts(lipid[
```

**Statistiques**

- Résumés
- Tables de contingence
- Moyennes
- Proportions
- Variances
- Tests non paramétriques
- Analyse multivariée
- Ajustement de modèles

Jeu de données actif

**Statistiques descriptives...**

- Distributions de fréquences...
- Dénombrer les observations manquantes
- Tableau de statistiques...
- Matrice de corrélations...
- Test de corrélation...
- Test de normalité...
- Transformer vers la loi normale...

Statistiques générales

Données Statistiques

Variables (une ou plusieurs)

- C22.4n.6
- C22.5n.3
- C22.5n.6
- C22.6n.3
- Z.C16.1n.7
- Z.C16.1n.9**

Résumer par groupes...

Aide Réinitialiser OK Annuler Appliquer

```
Rcmdr> numSummary(lipid[,c("Z.C16.1n.7", "Z.C16.1n.9"), drop=FALSE],
Rcmdr+   statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "cv", "skewness",
Rcmdr+   "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")
      mean sd  se(mean)      IQR cv  skewness  kurtosis
Z.C16.1n.7 -4.857226e-17  1 0.1581139 0.7772272 NA 1.9309189 3.1882201
Z.C16.1n.9 -1.724315e-16  1 0.1581139 1.2456799 NA 0.9524791 0.5377586
      0%      25%      50%      75%     100%  n
Z.C16.1n.7 -0.9476644 -0.5984147 -0.3747944 0.1788125 3.176330 40
Z.C16.1n.9 -1.3930561 -0.8053057 -0.1473762 0.4403742 2.852782 40
RcmdrMsg: [6] AVIS: warning in cv(x) : not all values are positive
```

```
Rcmdr> binnedCounts(lipid[,c("Z.C16.1n.7", "Z.C16.1n.9"), drop=FALSE])
Binned distribution of Z.C16.1n.7
```

	Count	Percent
[-1, -0.5]	18	45.0
(-0.5, 0]	11	27.5
(0, 0.5]	3	7.5
(0.5, 1]	3	7.5
(1, 1.5]	1	2.5
(1.5, 2]	1	2.5
(2, 2.5]	1	2.5
(2.5, 3]	1	2.5
(3, 3.5]	1	2.5
Total	40	100.0

```
Binned distribution of Z.C16.1n.9
```

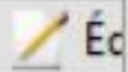
	Count	Percent
[-1.5, -1]	5	12.5
(-1, -0.5]	9	22.5
(-0.5, 0]	10	25.0
(0, 0.5]	7	17.5
(0.5, 1]	2	5.0
(1, 1.5]	3	7.5
(1.5, 2]	3	7.5
(2, 2.5]	0	0.0
(2.5, 3]	1	2.5
Total	40	100.0

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide



Données : lipid



Éd

Script R R Markdown

```
binCounts(lipid[,c("C16.1  
with(lipid, Hist(C16.ln.7, s  
col="darkgray", main="Dist  
with(lipid, Hist(C16.ln.7, s  
col="darkgray", main="Dist
```

Palette de couleurs...

Graphe indexé...

Graphe en points (dot plot)

**Histogramme...**

Graphe d'une variable numérique discrète...

Estimation de densité...

Graphe tiges et feuilles...

Boîtes de dispersion

modèle>



Données Options

Variable (une)

C22.4n.6

C22.5n.3

C22.5n.6

C22.6n.3

Z.C16.1n.7

Z.C16.1n.9

Graphe par groupe...

Données Options

Options du graphe

Nombre de classes : &lt;auto&gt;

Echelle des axes

 Fréquences Pourcentages Densité

Étiquettes du graphe

libellé de l'axe X &lt;auto&gt;

&lt; &gt;

libellé de l'axe Y &lt;auto&gt;

&lt; &gt;

Titre du graphe &lt;auto&gt;

&lt; &gt;



Aide



Réinitialiser



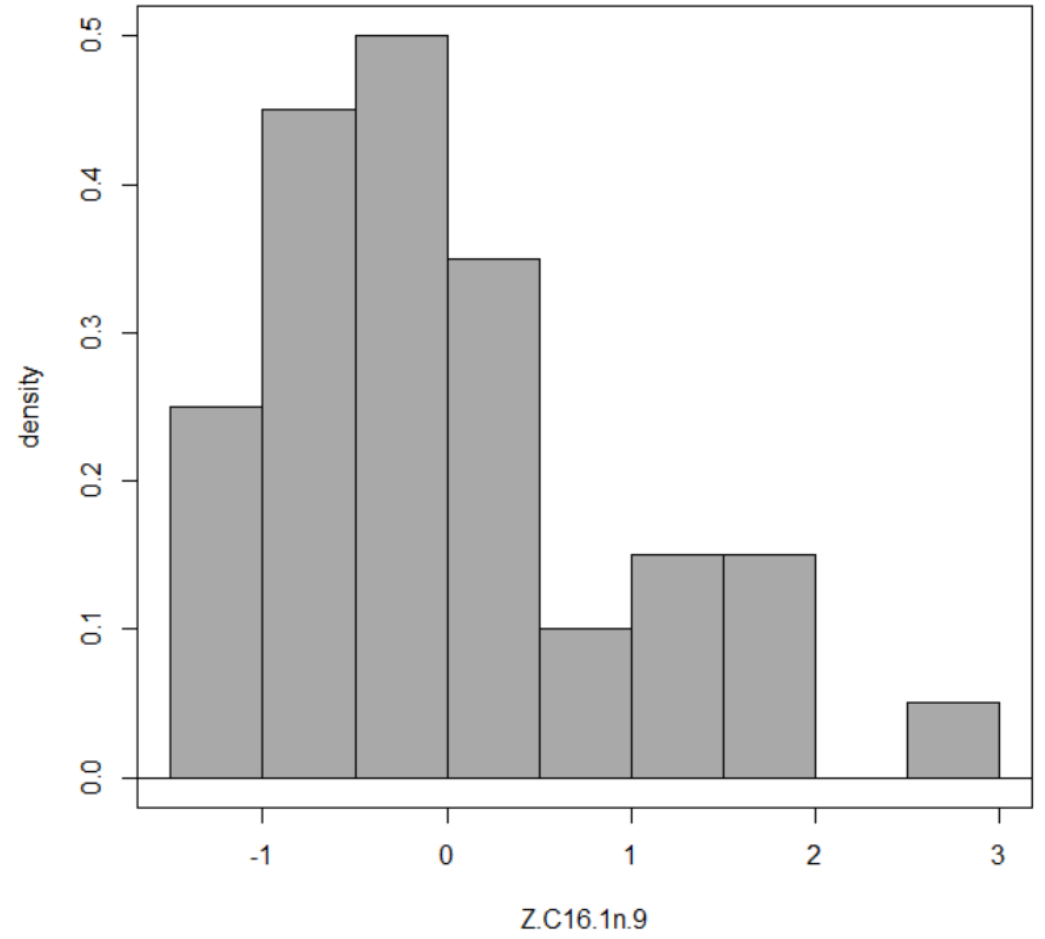
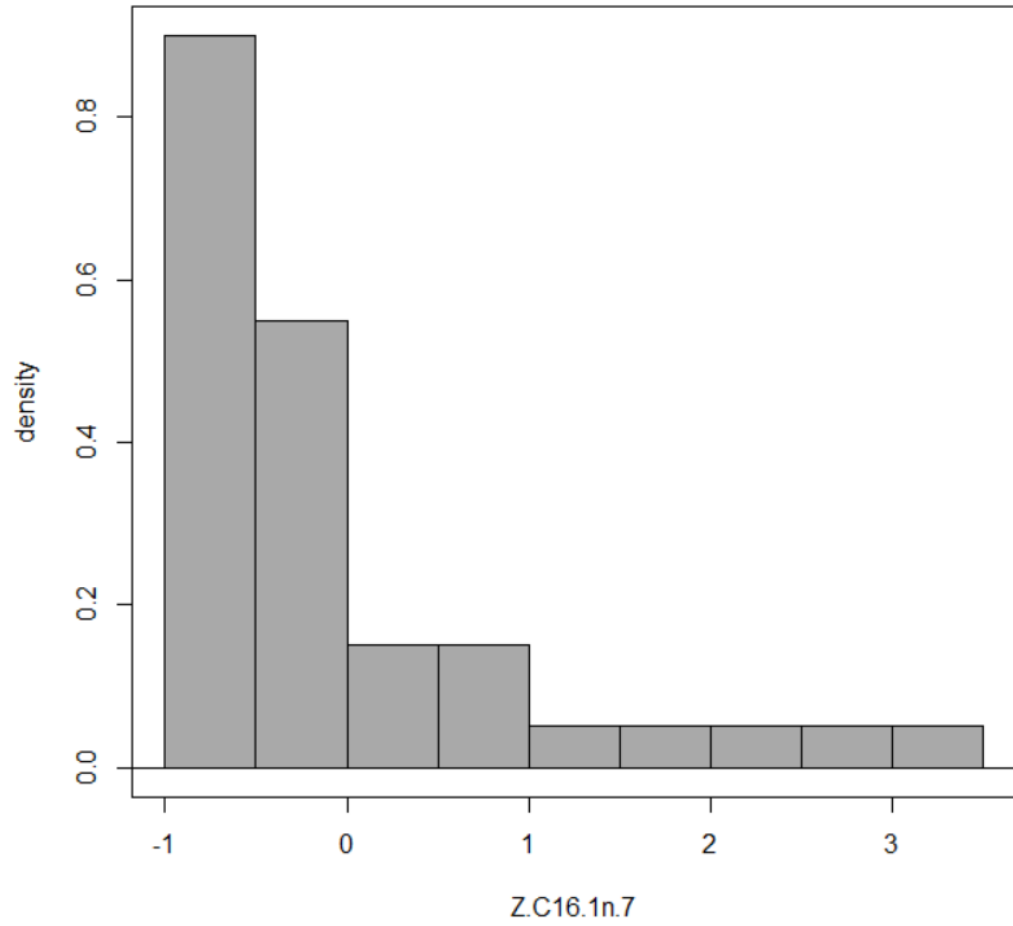
OK



Annuler

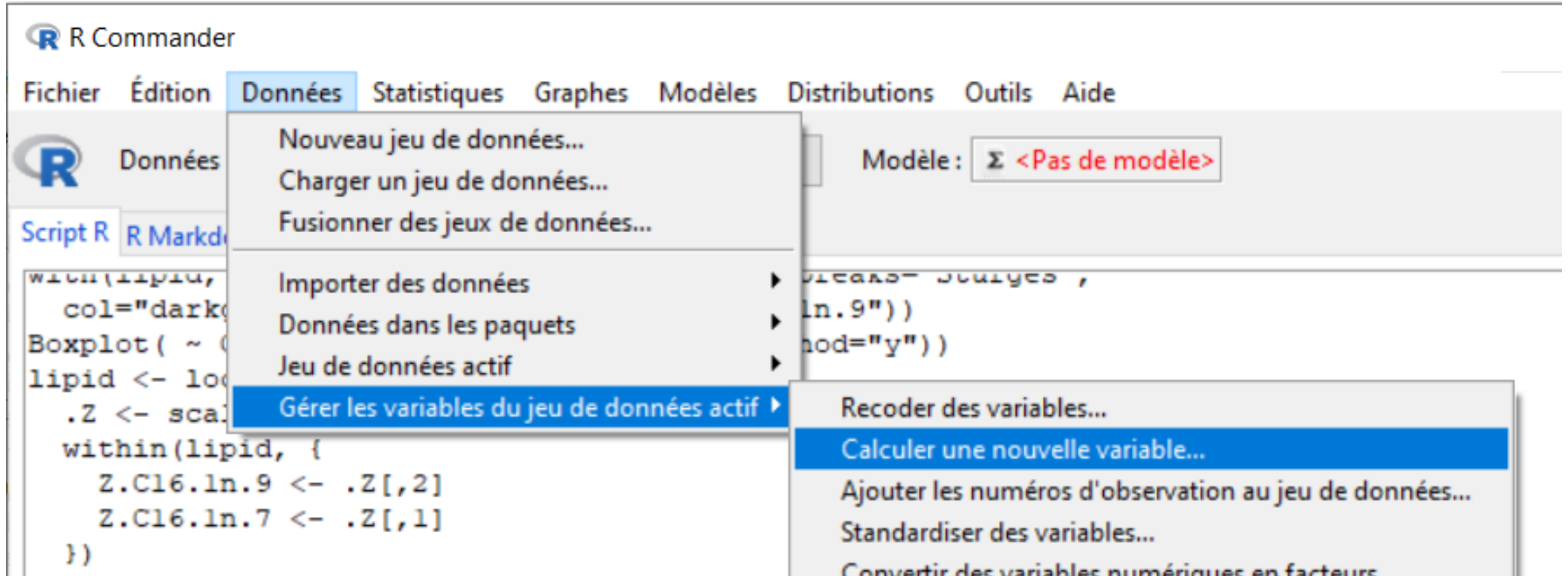


Appliquer





## Ajouter log2(C16.1n.7)



The image shows the R Commander interface. The 'Données' menu is open, and the submenu 'Gérer les variables du jeu de données actif' is also open. The submenu contains the following options:

- Recoder des variables...
- Calculer une nouvelle variable...**
- Ajouter les numéros d'observation au jeu de données...
- Standardiser des variables...
- Convertir des variables numériques en facteurs...

The main menu items are: Fichier, Édition, Données, Statistiques, Graphes, Modèles, Distributions, Outils, Aide.

The 'Données' menu items are: Nouveau jeu de données..., Charger un jeu de données..., Fusionner des jeux de données..., Importer des données, Données dans les paquets, Jeu de données actif, Gérer les variables du jeu de données actif.

The submenu items are: Recoder des variables..., Calculer une nouvelle variable..., Ajouter les numéros d'observation au jeu de données..., Standardiser des variables..., Convertir des variables numériques en facteurs...

The code editor shows the following R code:

```
with(lipid, {  
  col="darkred"  
  Boxplot( ~ C16.1n.9 )  
  lipid <- log2(C16.1n.7)  
  .Z <- scale(lipid) # standardize  
  within(lipid, {  
    Z.C16.1n.9 <- .Z[,2]  
    Z.C16.1n.7 <- .Z[,1]  
  })  
})
```

Variables existantes (double-clic envoie vers l'expression)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C18.0
- C18.1n.7

Nom de la nouvelle variable

log.C16.1n.7

Expression à calculer

log2 (C16.1n.7)

	C18.3n.3	C20.3n.3	C20.5n.3	C22.5n.3	C22.6n.3	Z.C16.1n.7	Z.C16.1n.9	log.C16.1n.7
1	5.97	0.37	8.62	1.75	10.39	-0.4417967	-1.18251867	1.6322682
2	0.00	0.00	0.00	0.48	2.61	-0.6294033	-0.48072717	1.3448285
3	0.00	0.00	0.00	0.22	2.51	-0.5925520	-0.48072717	1.4059924
4	0.49	0.00	2.99	1.04	14.99	-0.5356000	-0.69126462	1.4956952
5	0.42	0.00	0.30	0.35	6.69	-0.5255496	-0.79653335	1.5109619
6	0.00	0.00	0.00	2.13	2.56	0.9518520	-0.09474185	2.8599695
7	8.40	0.42	7.37	2.05	9.84	-0.2742908	-1.14742909	1.8479969
8	6.01	0.39	7.96	2.33	10.40	-0.3848447	-1.39305612	1.7092906
9	0.55	0.00	3.13	1.65	16.36	-0.6897054	-0.86671250	1.2387869
10	0.00	0.00	0.00	0.00	1.86	0.8680991	0.74740794	2.8094144

## TO DO

- **Faire la boîte à moustache de la nouvelle variable et commenter en particulier les variables atypiques**

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données : lipid

Script R R Markdown

```
with(lipid, hist(C16.ln.7, s  
  col="darkgray", main="Dist  
with(lipid, Hist(C16.ln.7, s  
  col="darkgray", main="Dist  
with(lipid, Hist(C16.ln.9, s  
  col="darkgray", main="Dist  
Boxplot( ~ C16.ln.7, data=li  
lipid <- local({  
  ? <- scale(lipid["C16
```

Palette de couleurs...

Graphe indexé...

Graphe en points (dot plot)

Histogramme...

Graphe d'une variable numérique discrète...

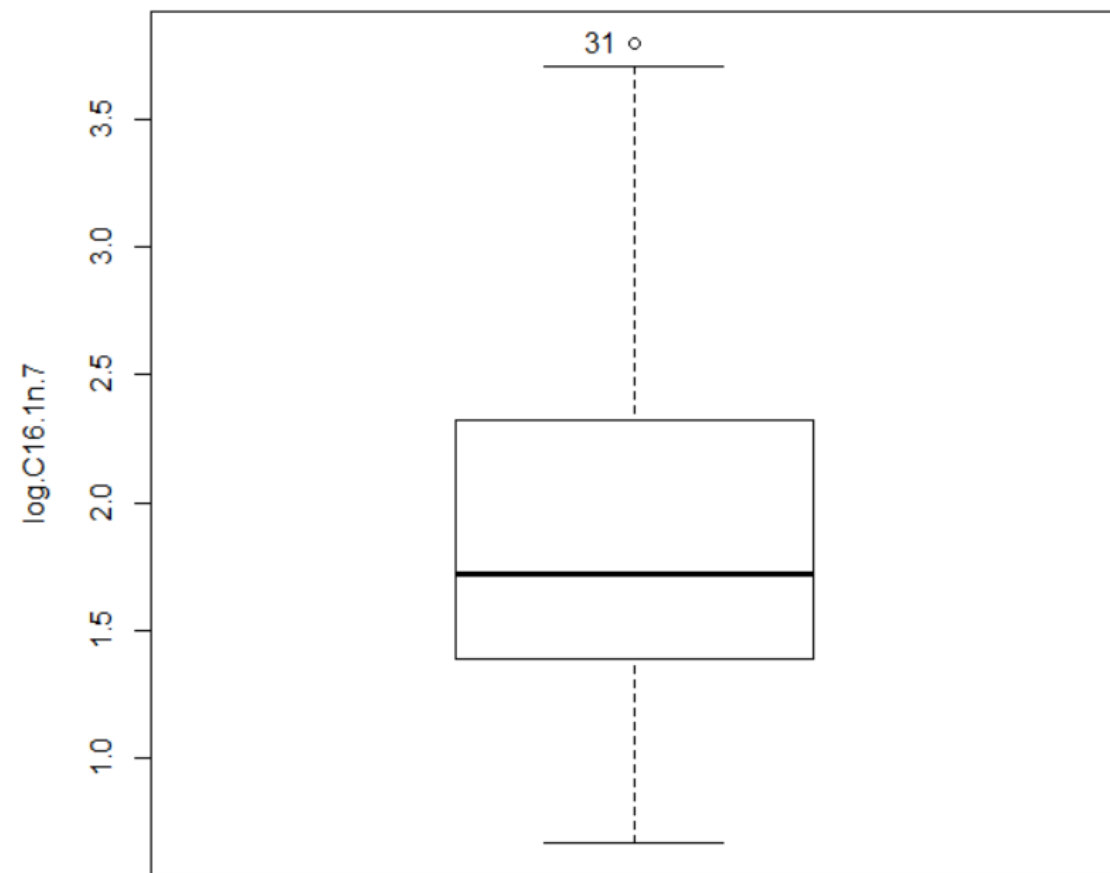
Estimation de densité...

Graphe tiges et feuilles...

**Boîte de dispersion...**

Graphe quantile-quantile...

Boîte à moustaches pour symétrie...



## **Partie 2.2 : Variables qualitatives**

# Switch sur diet

R Commander

Fichier Édition Données Statistiques Graphes Modèles Distributions O

Données : lipid Éditer Visualiser Modèle :

Script R

Sélectionnez les données

Données (choisir un jeu)

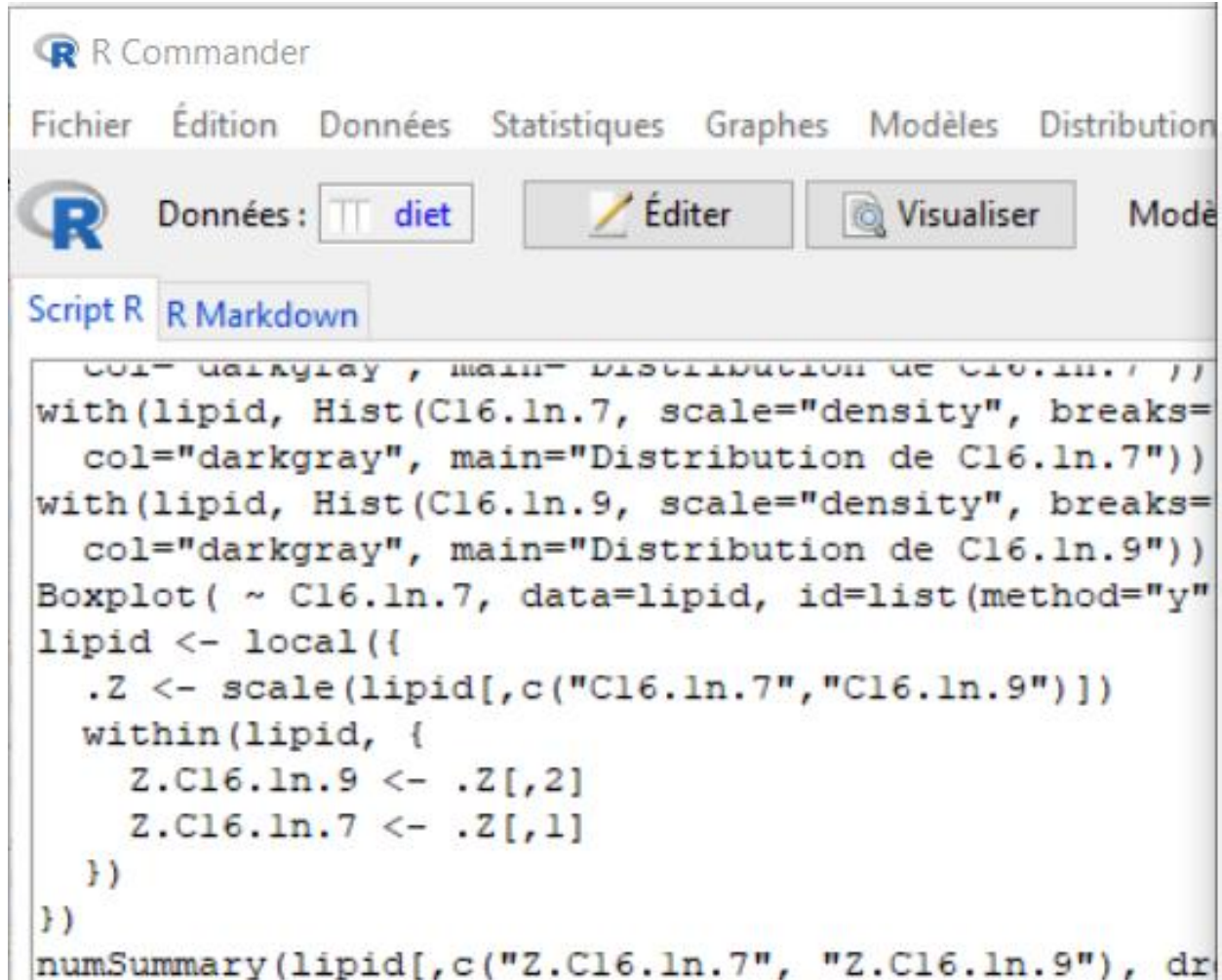
- diet
- gene
- genotype
- lipid

OK Annuler

```
with(lipid, colMeans(density ~ genotype, breaks="Sturges", plot=TRUE))
with(lipid, colMeans(density ~ genotype, breaks="Sturges", plot=TRUE))
Boxplot(density ~ genotype, data=lipid, id=list(method="y"))
plot(density ~ genotype, data=lipid, id=list(method="y"))
numSummary(lipid[,c("Z.C16.ln.7", "Z.C16.ln.9"), drop=FALSE],
  statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "kurtosis"),
  quantiles=c(0, .25, .5, .75, 1), type="2")
```



## Changement du nom de la colonne via l'éditeur



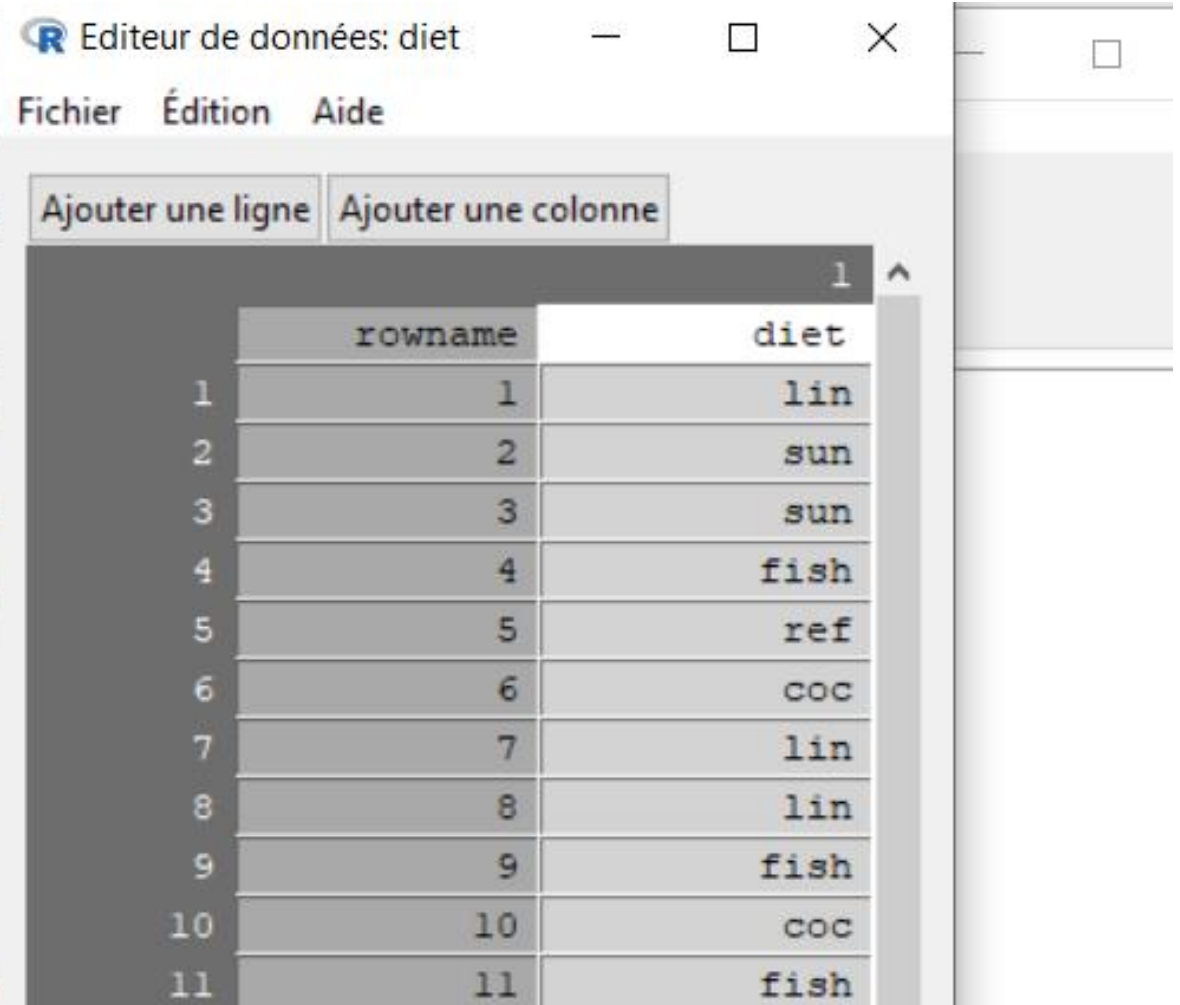
R Commander

Fichier Édition Données Statistiques Graphes Modèles Distribution

Données:  Éditer Visualiser Modèles

Script R R Markdown

```
col = darkgray, main = "Distribution de C16.ln.7" ))
with(lipid, Hist(C16.ln.7, scale="density", breaks=
  col="darkgray", main="Distribution de C16.ln.7"))
with(lipid, Hist(C16.ln.9, scale="density", breaks=
  col="darkgray", main="Distribution de C16.ln.9"))
Boxplot( ~ C16.ln.7, data=lipid, id=list(method="y")
lipid <- local({
  .Z <- scale(lipid[,c("C16.ln.7", "C16.ln.9")])
  within(lipid, {
    Z.C16.ln.9 <- .Z[,2]
    Z.C16.ln.7 <- .Z[,1]
  })
})
numSummary(lipid[,c("Z.C16.ln.7", "Z.C16.ln.9"), dr
```



R Editeur de données: diet

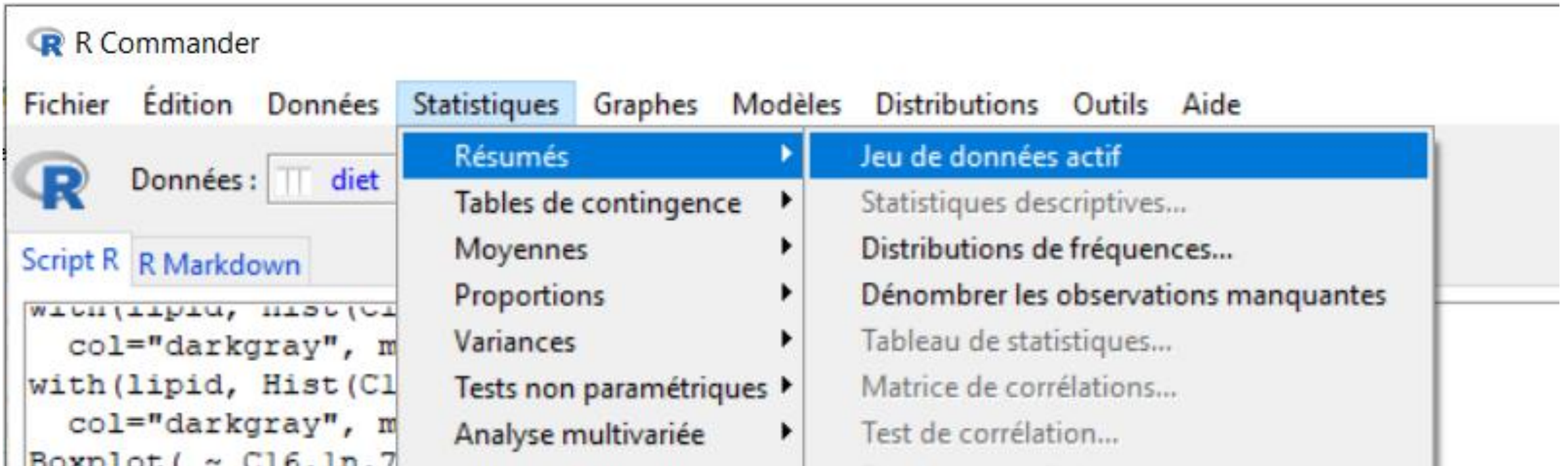
Fichier Édition Aide

Ajouter une ligne Ajouter une colonne

	rowname	diet
1	1	lin
2	2	sun
3	3	sun
4	4	fish
5	5	ref
6	6	coc
7	7	lin
8	8	lin
9	9	fish
10	10	coc
11	11	fish

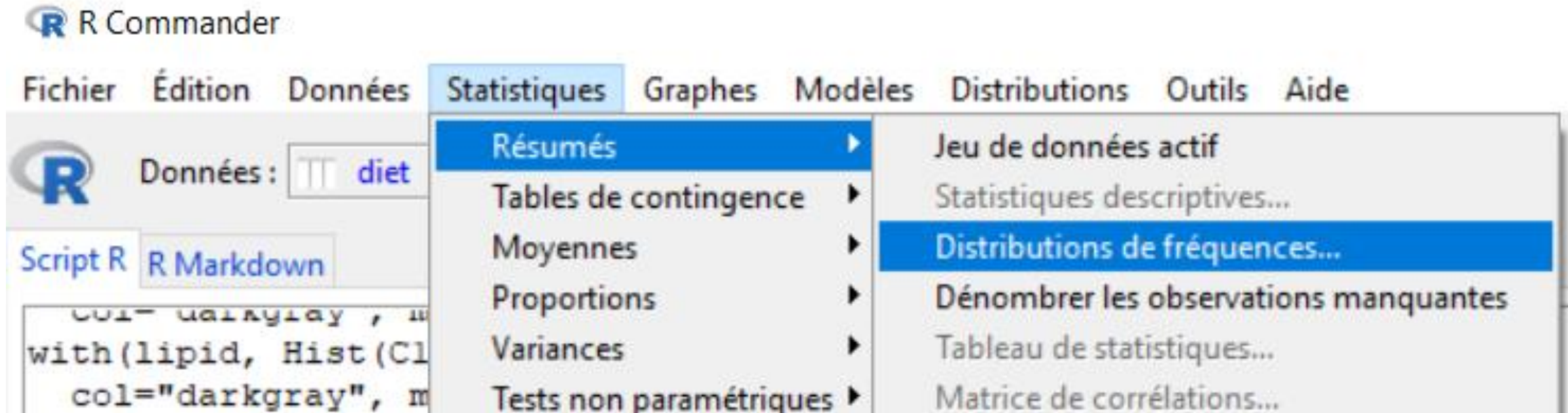


# Résumé statistique



```
Rcmdr> summary(diet)
  diet
coc  :8
fish:8
lin  :8
ref  :8
sun  :8
> |
```

## Distribution de fréquences



R Commander

Fichier Édition Données **Statistiques** Graphes Modèles Distributions Outils Aide

Données :

Script R R Markdown

```
col="darkgray", m  
with(lipid, Hist(Cl  
col="darkgray", m
```

- Résumés ▶
  - Jeu de données actif
  - Statistiques descriptives...
  - Distributions de fréquences...**
  - Dénombrer les observations manquantes
  - Tableau de statistiques...
  - Matrice de corrélations...
- Tables de contingence ▶
- Moyennes ▶
- Proportions ▶
- Variations ▶
- Tests non paramétriques ▶

Variables (une ou plusieurs)

diet

 Test d'ajustement du Chi-deux (une variable)

Aide



Réinitialiser



OK



Annuler



Appliquer

```
Rcmdr> local({  
Rcmdr+   .Table <- with(diet, table(diet))  
Rcmdr+   cat("\ncounts:\n")  
Rcmdr+   print(.Table)  
Rcmdr+   cat("\npercentages:\n")  
Rcmdr+   print(round(100*.Table/sum(.Table), 2))  
Rcmdr+ })
```

counts:

diet

	coc	fish	lin	ref	sun
	8	8	8	8	8

percentages:

diet

	coc	fish	lin	ref	sun
	20	20	20	20	20

# Graphe en barres (barchart)

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide



Données:  Éd

Script R R Markdown

```
lipid$log.C16.ln.7 <- with(1
Z.C16.ln.7 <- .Z[,1]
})
})
numSummary(lipid[,c("Z.C16.1
statistics=c("mean", "sd",
"kurtosis"), quantiles=c(
binnedCounts(lipid[,c("Z.C16

lipid$log.C16.ln.7 <- with(1
Boxplot( ~ log.C16.ln.7, dat
editDataset(diet)
summary(diet)
local({
.Table <- with(diet, table
cat("\ncounts:\n")
print(.Table)
cat("\npercentages:\n")
print(round(100*.Table/sum
```

Palette de couleurs...

Graphe indexé...

Graphe en points (dot plot)

Histogramme...

Graphe d'une variable numérique discrète...

Estimation de densité...

Graphe tiges et feuilles...

Boîte de dispersion...

Graphe quantile-quantile...

Boîte à moustaches pour symétrie...

Nuage de points...

Matrice de nuages de points...

Graphe en lignes...

Graphe XY conditionnel...

Graphe des moyennes...

Graphe en bande...

**Graphe en barres...**

Graphe en camembert...

Données Options

## Echelle des axes

- Fréquences  
 Pourcentages

## Sélection des couleurs

- Par défaut  
 À partir de la palette de couleurs

## Style des barres de groupes

- Divisées (empilement)  
 Côte-à-côte (en parallèle)

## Pourcentages pour les barres de groupes

- Conditionnel  
 Total

## Position de la légende

- Au dessus de graphe  
 A droite  
 Centrer  
 A gauche

## Étiquettes du graphe

libellé de l'axe X   
< >  
libellé de l'axe Y   
< >  
Titre du graphe   
< >



Aide



Réinitialiser



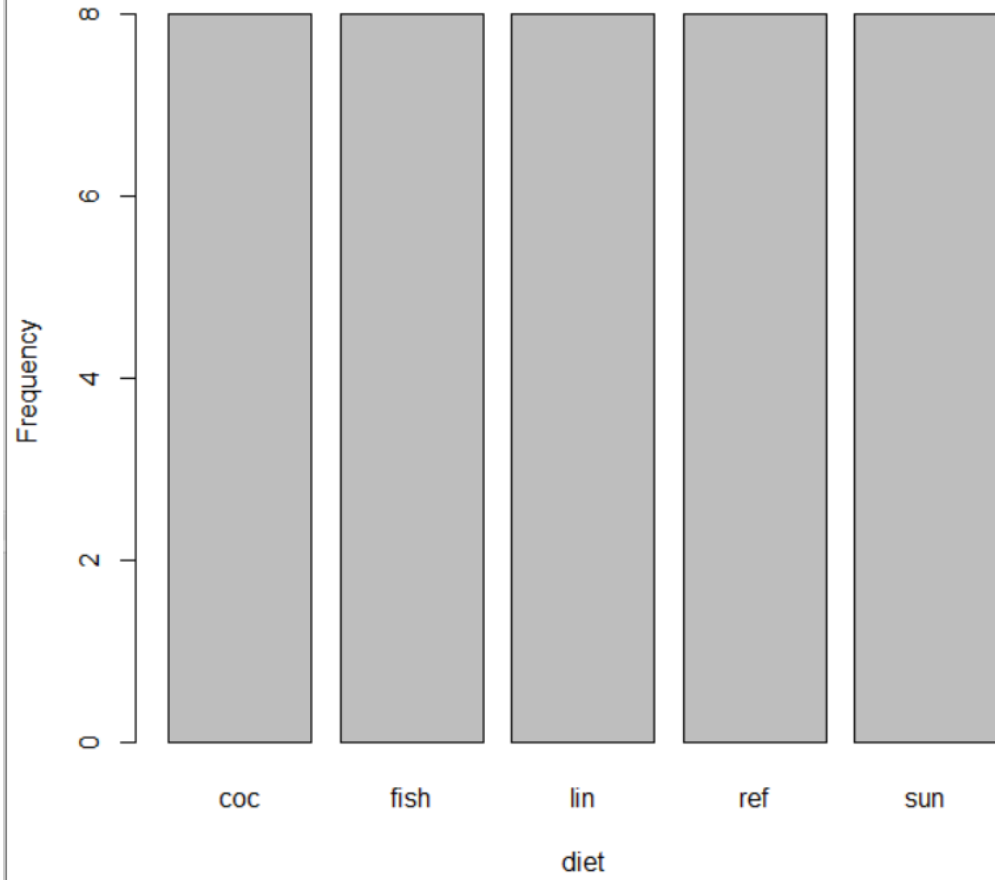
OK



Annuler



Appliquer





# Graphe en camembert (pie-chart)

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données:  Éd

Script R R Markdown

```
lipid$Z.C16.ln.7 <- with(1, lipid[,4])
})
})
numSummary(lipid[,c("Z.C16.ln.7", "Z.C16.ln.8", "Z.C16.ln.9", "Z.C16.ln.10", "Z.C16.ln.11", "Z.C16.ln.12", "Z.C16.ln.13", "Z.C16.ln.14", "Z.C16.ln.15", "Z.C16.ln.16", "Z.C16.ln.17", "Z.C16.ln.18", "Z.C16.ln.19", "Z.C16.ln.20", "Z.C16.ln.21", "Z.C16.ln.22", "Z.C16.ln.23", "Z.C16.ln.24", "Z.C16.ln.25", "Z.C16.ln.26", "Z.C16.ln.27", "Z.C16.ln.28", "Z.C16.ln.29", "Z.C16.ln.30", "Z.C16.ln.31", "Z.C16.ln.32", "Z.C16.ln.33", "Z.C16.ln.34", "Z.C16.ln.35", "Z.C16.ln.36", "Z.C16.ln.37", "Z.C16.ln.38", "Z.C16.ln.39", "Z.C16.ln.40", "Z.C16.ln.41", "Z.C16.ln.42", "Z.C16.ln.43", "Z.C16.ln.44", "Z.C16.ln.45", "Z.C16.ln.46", "Z.C16.ln.47", "Z.C16.ln.48", "Z.C16.ln.49", "Z.C16.ln.50", "Z.C16.ln.51", "Z.C16.ln.52", "Z.C16.ln.53", "Z.C16.ln.54", "Z.C16.ln.55", "Z.C16.ln.56", "Z.C16.ln.57", "Z.C16.ln.58", "Z.C16.ln.59", "Z.C16.ln.60", "Z.C16.ln.61", "Z.C16.ln.62", "Z.C16.ln.63", "Z.C16.ln.64", "Z.C16.ln.65", "Z.C16.ln.66", "Z.C16.ln.67", "Z.C16.ln.68", "Z.C16.ln.69", "Z.C16.ln.70", "Z.C16.ln.71", "Z.C16.ln.72", "Z.C16.ln.73", "Z.C16.ln.74", "Z.C16.ln.75", "Z.C16.ln.76", "Z.C16.ln.77", "Z.C16.ln.78", "Z.C16.ln.79", "Z.C16.ln.80", "Z.C16.ln.81", "Z.C16.ln.82", "Z.C16.ln.83", "Z.C16.ln.84", "Z.C16.ln.85", "Z.C16.ln.86", "Z.C16.ln.87", "Z.C16.ln.88", "Z.C16.ln.89", "Z.C16.ln.90", "Z.C16.ln.91", "Z.C16.ln.92", "Z.C16.ln.93", "Z.C16.ln.94", "Z.C16.ln.95", "Z.C16.ln.96", "Z.C16.ln.97", "Z.C16.ln.98", "Z.C16.ln.99", "Z.C16.ln.100"),
  statistics=c("mean", "sd", "kurtosis"), quantiles=c("q0", "q1", "q2", "q3", "q4"),
  binnedCounts(lipid[,c("Z.C16.ln.7", "Z.C16.ln.8", "Z.C16.ln.9", "Z.C16.ln.10", "Z.C16.ln.11", "Z.C16.ln.12", "Z.C16.ln.13", "Z.C16.ln.14", "Z.C16.ln.15", "Z.C16.ln.16", "Z.C16.ln.17", "Z.C16.ln.18", "Z.C16.ln.19", "Z.C16.ln.20", "Z.C16.ln.21", "Z.C16.ln.22", "Z.C16.ln.23", "Z.C16.ln.24", "Z.C16.ln.25", "Z.C16.ln.26", "Z.C16.ln.27", "Z.C16.ln.28", "Z.C16.ln.29", "Z.C16.ln.30", "Z.C16.ln.31", "Z.C16.ln.32", "Z.C16.ln.33", "Z.C16.ln.34", "Z.C16.ln.35", "Z.C16.ln.36", "Z.C16.ln.37", "Z.C16.ln.38", "Z.C16.ln.39", "Z.C16.ln.40", "Z.C16.ln.41", "Z.C16.ln.42", "Z.C16.ln.43", "Z.C16.ln.44", "Z.C16.ln.45", "Z.C16.ln.46", "Z.C16.ln.47", "Z.C16.ln.48", "Z.C16.ln.49", "Z.C16.ln.50", "Z.C16.ln.51", "Z.C16.ln.52", "Z.C16.ln.53", "Z.C16.ln.54", "Z.C16.ln.55", "Z.C16.ln.56", "Z.C16.ln.57", "Z.C16.ln.58", "Z.C16.ln.59", "Z.C16.ln.60", "Z.C16.ln.61", "Z.C16.ln.62", "Z.C16.ln.63", "Z.C16.ln.64", "Z.C16.ln.65", "Z.C16.ln.66", "Z.C16.ln.67", "Z.C16.ln.68", "Z.C16.ln.69", "Z.C16.ln.70", "Z.C16.ln.71", "Z.C16.ln.72", "Z.C16.ln.73", "Z.C16.ln.74", "Z.C16.ln.75", "Z.C16.ln.76", "Z.C16.ln.77", "Z.C16.ln.78", "Z.C16.ln.79", "Z.C16.ln.80", "Z.C16.ln.81", "Z.C16.ln.82", "Z.C16.ln.83", "Z.C16.ln.84", "Z.C16.ln.85", "Z.C16.ln.86", "Z.C16.ln.87", "Z.C16.ln.88", "Z.C16.ln.89", "Z.C16.ln.90", "Z.C16.ln.91", "Z.C16.ln.92", "Z.C16.ln.93", "Z.C16.ln.94", "Z.C16.ln.95", "Z.C16.ln.96", "Z.C16.ln.97", "Z.C16.ln.98", "Z.C16.ln.99", "Z.C16.ln.100"),
  plot=FALSE)

lipid$log.C16.ln.7 <- with(1, lipid[,4])
Boxplot( ~ log.C16.ln.7, data=lipid, editDataset(diet))
summary(diet)
local({
  .Table <- with(diet, table(x=lipid$log.C16.ln.7))
  cat("\ncounts:\n")
  print(.Table)
  cat("\npercentages:\n")
  print(round(100*.Table/sum(.Table)))
})
with(diet, Barplot(diet, xlab="log.C16.ln.7", ylab="count", main="Barplot of log.C16.ln.7"))
```

modèle>

cv", "

)

- Palette de couleurs...
- Graphe indexé...
- Graphe en points (dot plot)
- Histogramme...
- Graphe d'une variable numérique discrète...
- Estimation de densité...
- Graphe tiges et feuilles...
- Boîte de dispersion...
- Graphe quantile-quantile...
- Boîte à moustaches pour symétrie...
- Nuage de points...
- Matrice de nuages de points...
- Graphe en lignes...
- Graphe XY conditionnel...
- Graphe des moyennes...
- Graphe en bande...
- Graphe en barres...
- Graphe en camembert...**
- Graphe 3D

Graphique en camembert

Variable (une)

diet

Sélection des couleurs

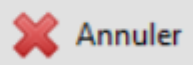
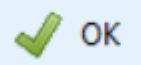
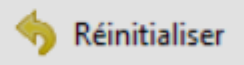
- Par défaut
- À partir de la palette de couleurs

Étiquettes du graphe

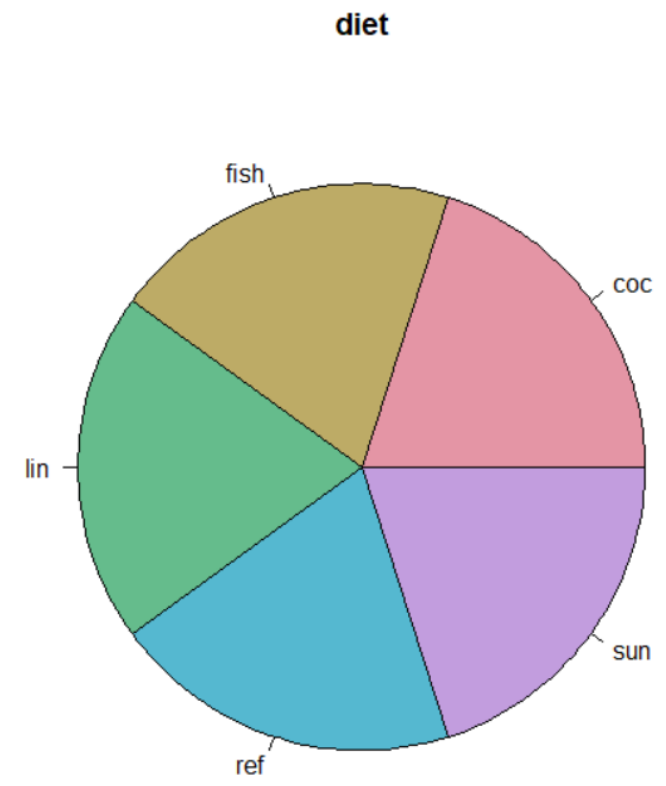
libellé de l'axe X < auto >

libellé de l'axe Y < auto >

Titre du graphe < auto >



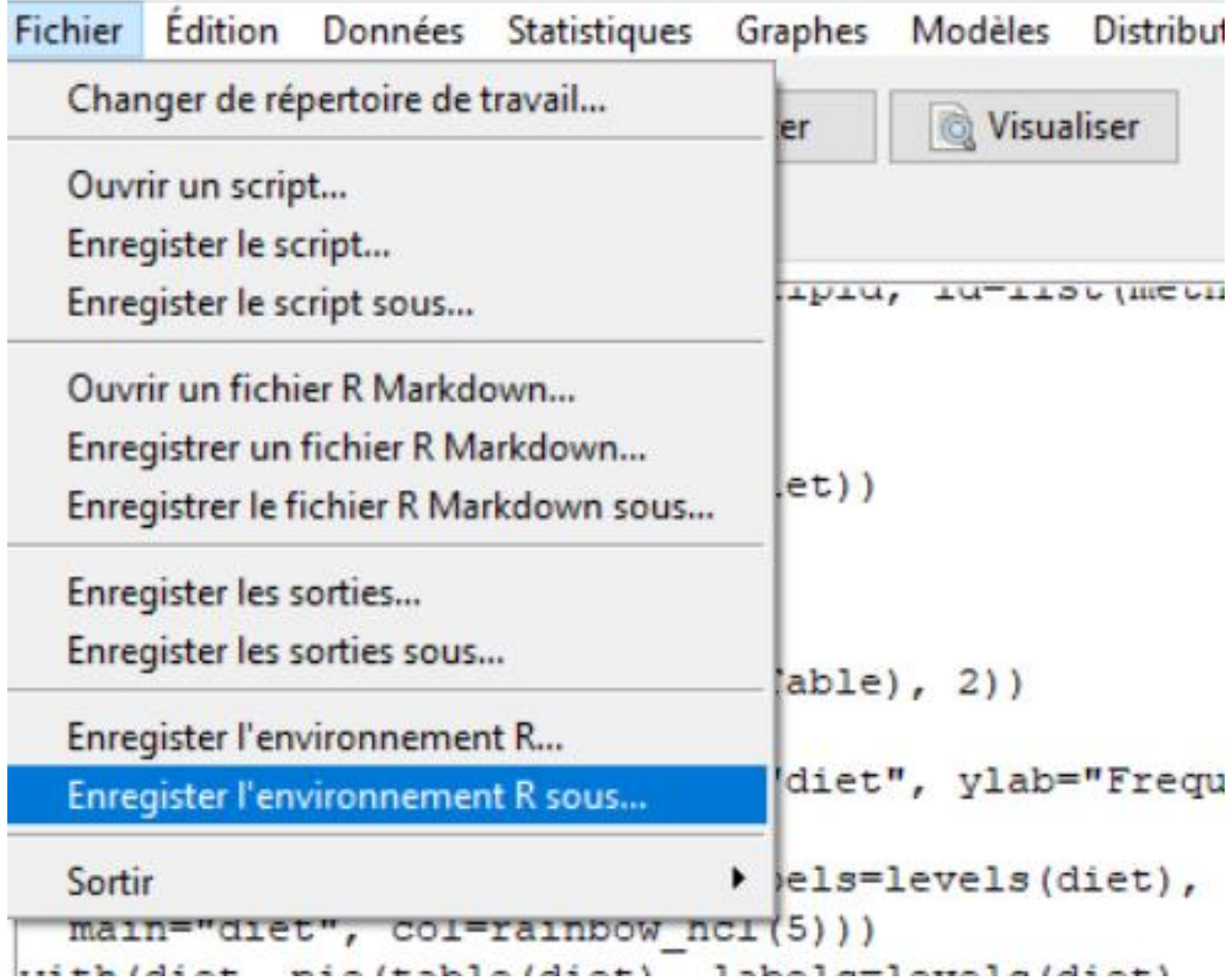
Redimensionnement





# Sauvegarde de l'environnement Rdata

R Commander

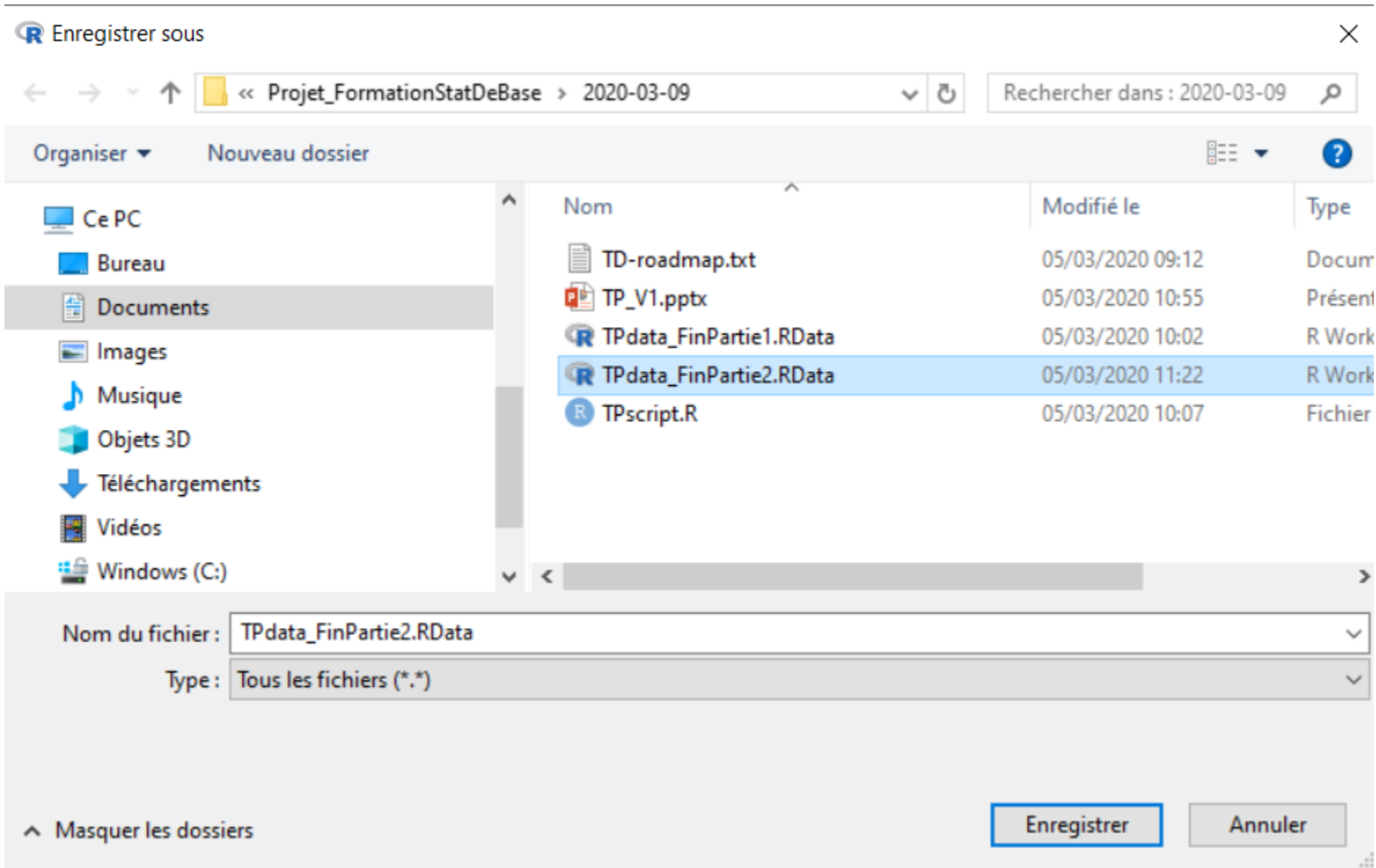


The image shows the R Commander application window. The 'Fichier' (File) menu is open, displaying various options. The option 'Enregistrer l'environnement R sous...' (Save R environment as...) is highlighted in blue. The menu items are:

- Changer de répertoire de travail...
- Ouvrir un script...
- Enregistrer le script...
- Enregistrer le script sous...
- Ouvrir un fichier R Markdown...
- Enregistrer un fichier R Markdown...
- Enregistrer le fichier R Markdown sous...
- Enregistrer les sorties...
- Enregistrer les sorties sous...
- Enregistrer l'environnement R...
- Enregistrer l'environnement R sous...**
- Sortir

The background shows a code editor with R code, including a `plot` function call:

```
plot(table, las=1, col=rainbow(nci))  
  
plot(diet, ylab="Freque  
      xlab="Niveau", col=rainbow(nci),  
      main="table", col=rainbow(nci(5)))  
with(diet, pie(table(diet), labels=levels(diet))
```

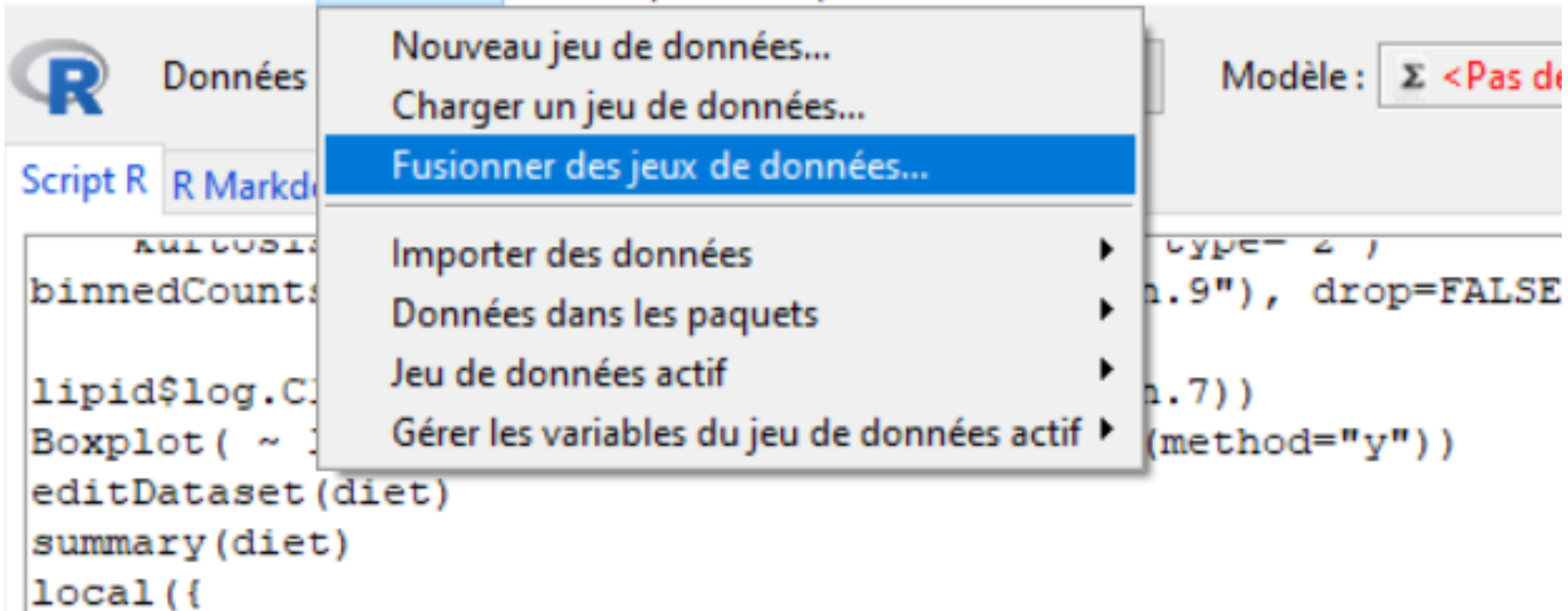


# **Partie 3 : Statistiques bivariées**

## Fusion des jeux de données

R Commander

Fichier Édition **Données** Statistiques Graphes Modèles Distributions Outils Aide



The image shows the 'Données' menu in R Commander. The menu items are:

- Nouveau jeu de données...
- Charger un jeu de données...
- Fusionner des jeux de données...** (highlighted)
- Importer des données ▶
- Données dans les paquets ▶
- Jeu de données actif ▶
- Gérer les variables du jeu de données actif ▶

The background shows a script editor with the following code:

```
lipid$log.C  
Boxplot( ~  
editDataset(diet)  
summary(diet)  
local({
```

R Fusionner des jeux de données

Nom du tableau de données fusionné :

Premier jeu de données (en choisir un)  Second jeu de données (en choisir un)

Sens de la fusion

Fusionner des lignes

Fusionner des colonnes

Fusionner uniquement les lignes et les colonnes communes

Go to file/function | Addins

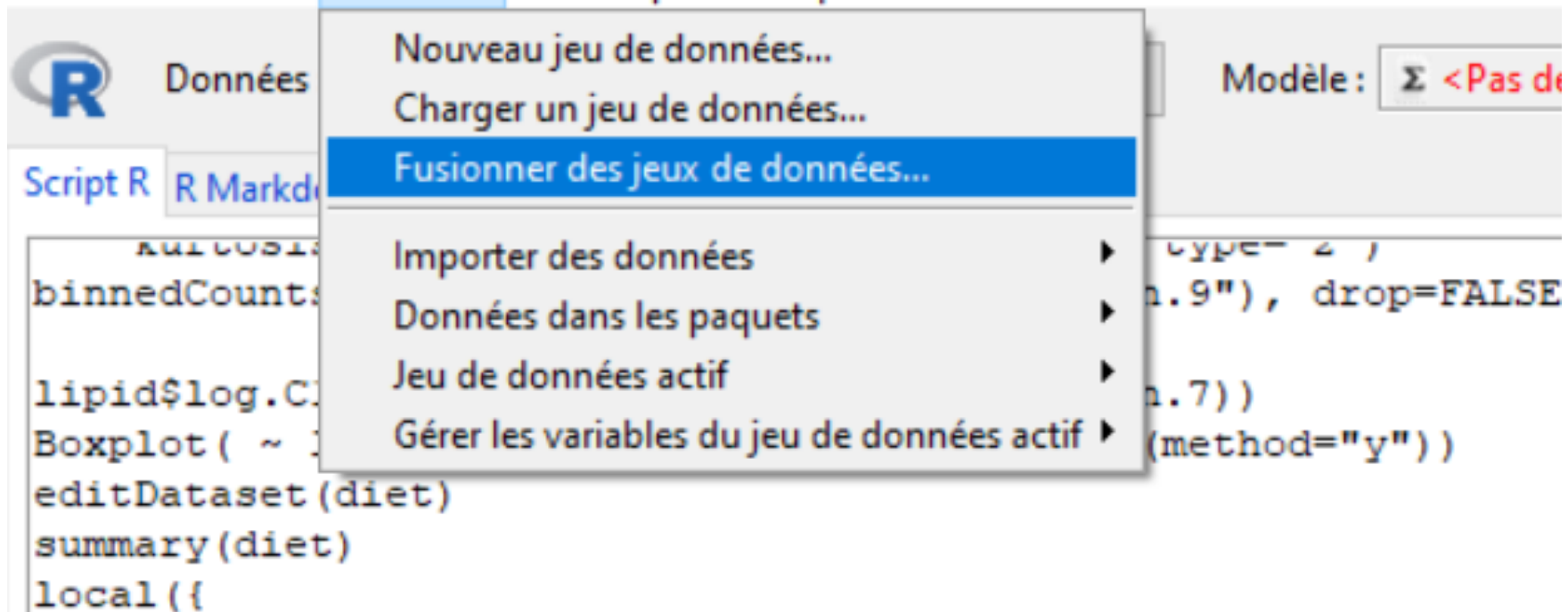
as.data.frame(alldata) x

Filter | Cols: 1 - 50

	diet	X36b4	ACAT1	ACAT2	ACBP	ACC1	ACC2	ACOTH	ADI
1	lin	-0.42	-0.65	-0.84	-0.34	-1.29	-1.13	-0.93	-0.91
10	coc	-0.50	-0.62	-0.79	-0.36	-1.21	-0.82	-0.92	-0.91
11	fish	-0.40	-0.60	-0.55	-0.25	-1.22	-1.13	-0.81	-0.91
12	ref	-0.52	-0.66	-0.66	-0.41	-1.28	-1.10	-0.95	-1.04
13	sun	-0.52	-0.63	-0.99	-0.43	-1.24	-0.96	-0.96	-0.91
14	ref	-0.47	-0.71	-0.44	-0.45	-1.44	-1.17	-1.02	-1.01
15	sun	-0.42	-0.66	-0.88	-0.33	-1.24	-1.03	-0.92	-0.94
16	lin	-0.58	-0.62	-0.85	-0.42	-1.33	-1.19	-0.91	-0.91
17	coc	-0.46	-0.69	-0.45	-0.32	-1.31	-0.93	-1.06	-1.01
18	fish	-0.36	-0.58	-0.71	-0.34	-1.20	-0.98	-0.92	-0.94
19	coc	-0.35	-0.62	-0.56	-0.24	-1.24	-0.88	-0.95	-0.91
2	sun	-0.44	-0.68	-0.91	-0.32	-1.23	-1.06	-0.99	-0.91

## TO DO

- **Fusionner alldata avec avec genotype, puis avec lipid en conservant le meme nom alldata**
- **Résumé numérique de ce fichier**



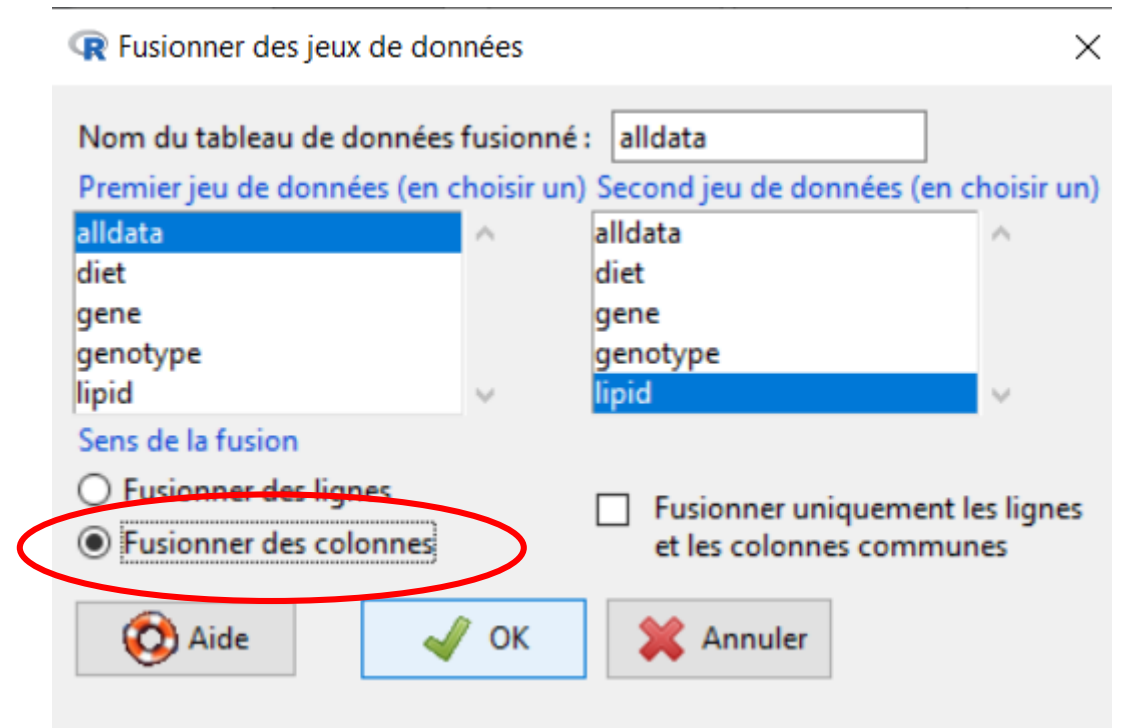
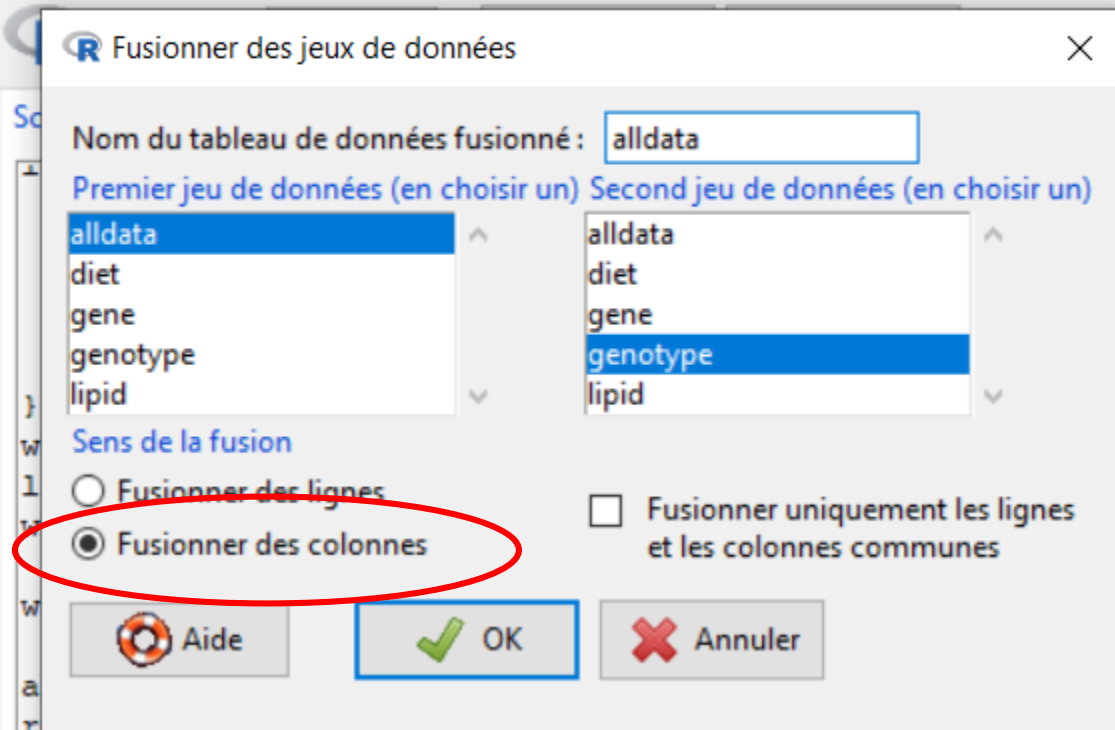
The image shows the 'Données' menu in R Commander. The menu items are:

- Nouveau jeu de données...
- Charger un jeu de données...
- Fusionner des jeux de données...** (highlighted)
- Importer des données ▶
- Données dans les paquets ▶
- Jeu de données actif ▶
- Gérer les variables du jeu de données actif ▶

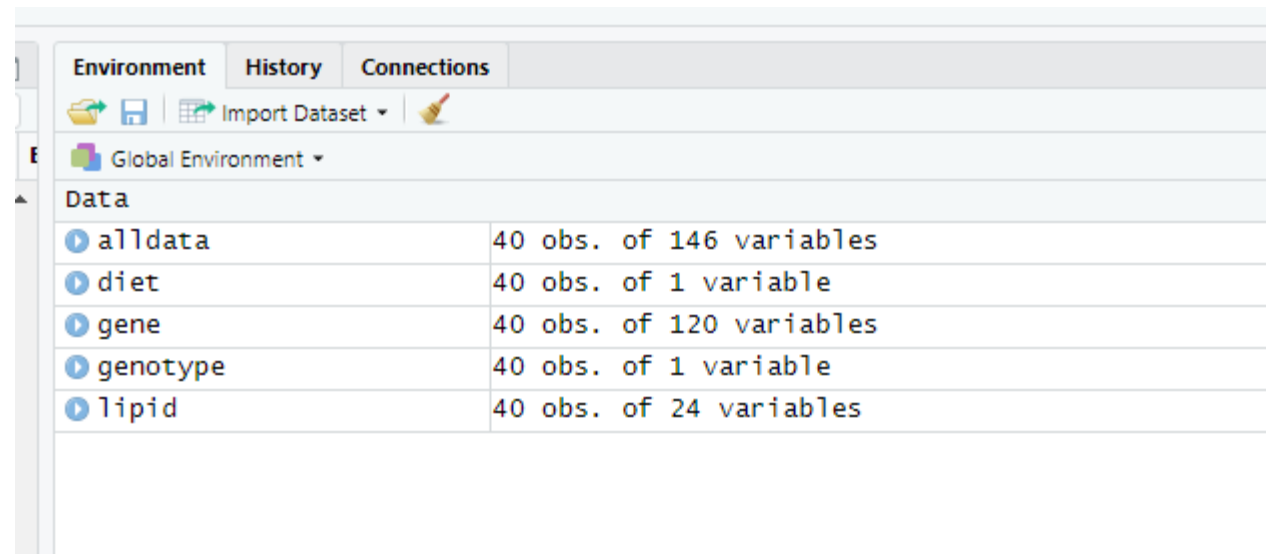
The background shows a script editor with R code:

```
lipid$log.C...  
Boxplot( ~ ...  
editDataset(diet)  
summary(diet)  
local({
```





```
Rcmdr> rownames(alldata) <- alldata$Row.names  
Rcmdr> alldata$Row.names <- NULL  
RcmdrMsg: [18] NOTE: Le jeu de données alldata a 40 lignes et 146 colonnes.  
> |
```



Environment	History	Connections
Global Environment		
Data		
alldata		40 obs. of 146 variables
diet		40 obs. of 1 variable
gene		40 obs. of 120 variables
genotype		40 obs. of 1 variable
lipid		40 obs. of 24 variables

Données : 

Script R R Markdown

```
##  
with(diet, Barplot(  
library(colorspace,  
with(diet, pie(tabl  
main="diet", col=  
with(diet, pie(tabl  
main="diet", col=rainbow_hcl(5)))  
alldata <- merge(diet, gene, all=TRUE, by="row.names")  
rownames(alldata) <- alldata$Row.names
```

Résumés ▶

Tables de contingence ▶

Moyennes ▶

Proportions ▶

Variances ▶

Tests non paramétriques ▶

Analyse multivariée ▶

Ajustement de modèles ▶

Jeu de données actif

Statistiques descriptives...

Distributions de fréquences...

Dénombrer les observations manquantes

Tableau de statistiques...

Matrice de corrélations...

Test de corrélation...

Test de normalité...

Transformer vers la loi normale...

Rcmdr> summary(alldata)

diet	X36b4	ACAT1	ACAT2
coc :8	Min. :-0.5800	Min. :-0.7500	Min. :-1.1000
fish:8	1st Qu.:-0.5025	1st Qu.:-0.6900	1st Qu.:-0.8800
lin :8	Median :-0.4600	Median :-0.6600	Median :-0.7950
ref :8	Mean :-0.4552	Mean :-0.6552	Mean :-0.7668
sun :8	3rd Qu.:-0.4200	3rd Qu.:-0.6200	3rd Qu.:-0.6450
	Max. :-0.3000	Max. :-0.5200	Max. :-0.3900

ACBP	ACC1	ACC2	ACOTH
Min. :-0.6600	Min. :-1.440	Min. :-1.2000	Min. :-1.0600
1st Qu.:-0.5025	1st Qu.:-1.300	1st Qu.:-1.0900	1st Qu.:-0.9500
Median :-0.4250	Median :-1.260	Median :-1.0450	Median :-0.9200
Mean :-0.4338	Mean :-1.258	Mean :-1.0280	Mean :-0.9107
3rd Qu.:-0.3550	3rd Qu.:-1.220	3rd Qu.:-0.9875	3rd Qu.:-0.8800
Max. :-0.2400	Max. :-1.070	Max. :-0.7900	Max. :-0.7300

ADISP	ADSS1	ALDH3	AM2R
Min. :-1.0800	Min. :-1.190	Min. :-0.9900	Min. :-0.780
1st Qu.:-1.0200	1st Qu.:-1.140	1st Qu.:-0.9100	1st Qu.:-0.670
Median :-0.9700	Median :-1.075	Median :-0.7850	Median :-0.630
Mean :-0.9782	Mean :-1.076	Mean :-0.8100	Mean :-0.628
3rd Qu.:-0.9400	3rd Qu.:-1.035	3rd Qu.:-0.7475	3rd Qu.:-0.590
Max. :-0.8700	Max. :-0.910	Max. :-0.6200	Max. :-0.460

AOX	BACT	BIEN	BSEP
Min. :-0.4800	Min. :-0.4400	Min. :-1.1600	Min. :-0.9000
1st Qu.:-0.3175	1st Qu.:-0.3225	1st Qu.:-0.9900	1st Qu.:-0.7600
Median :-0.2300	Median :-0.3000	Median :-0.9200	Median :-0.7000
Mean :-0.2505	Mean :-0.2828	Mean :-0.9213	Mean :-0.6910
3rd Qu.:-0.1675	3rd Qu.:-0.2350	3rd Qu.:-0.8550	3rd Qu.:-0.6275
Max. :-0.0400	Max. :-0.1100	Max. :-0.6400	Max. :-0.5100

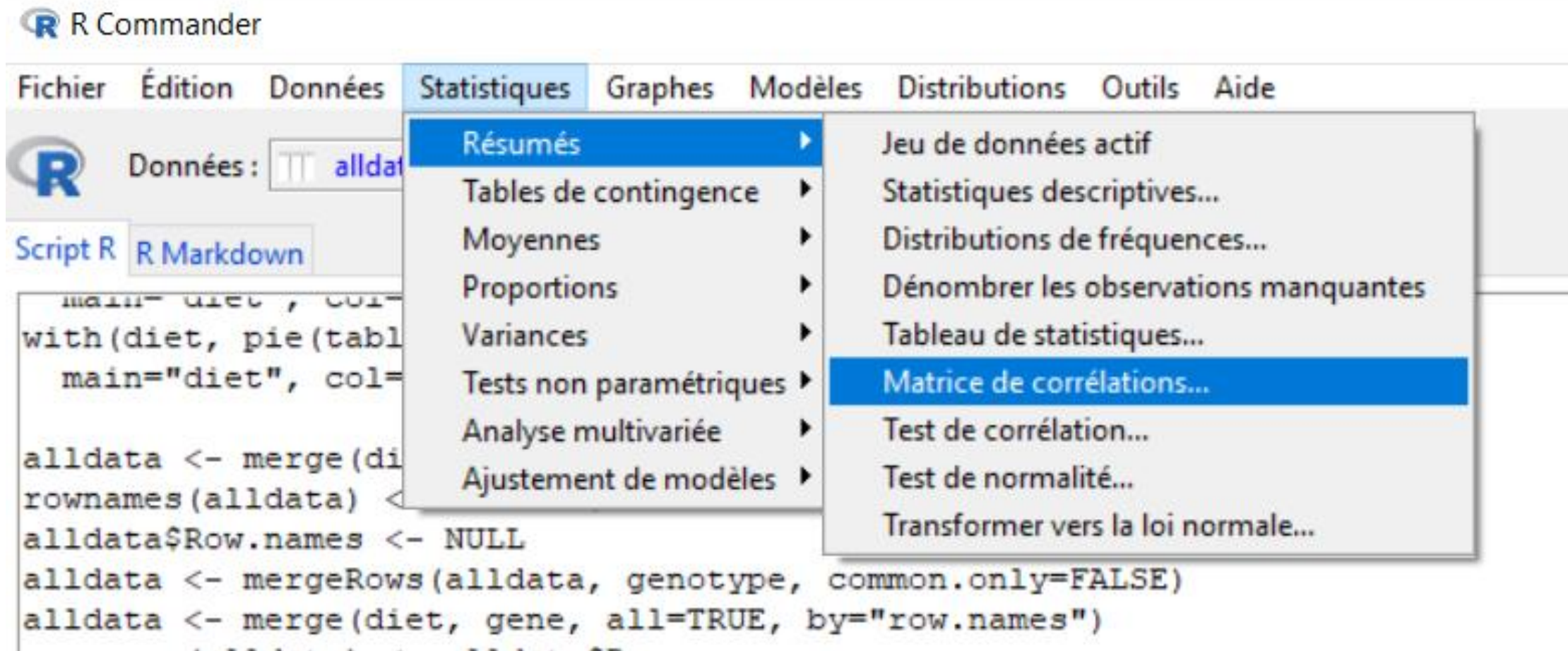
  

Bcl.3	C16SR	CACP	CAR1
-------	-------	------	------

# **Partie 3.1**

## **Deux variables quantitatives**

# Matrice de corrélation



R Commander

Fichier Édition Données **Statistiques** Graphes Modèles Distributions Outils Aide

Données: alldata

Script R R Markdown

- Résumés
  - Jeu de données actif
  - Statistiques descriptives...
  - Distributions de fréquences...
  - Dénombrer les observations manquantes
  - Tableau de statistiques...
  - Matrice de corrélations...**
  - Test de corrélation...
  - Test de normalité...
  - Transformer vers la loi normale...
- Tables de contingence
- Moyennes
- Proportions
- Variances
- Tests non paramétriques
- Analyse multivariée
- Ajustement de modèles

```
with(diet, pie(tabl
  main="diet", col=

alldata <- merge(di
rownames(alldata) <
alldata$Row.names <- NULL
alldata <- mergeRows(alldata, genotype, common.only=FALSE)
alldata <- merge(diet, gene, all=TRUE, by="row.names")
```

## R Matrice de corrélations

Variables (deux ou plus)

- C16SR
- C18.0
- C18.1n.7
- C18.1n.9
- C18.2n.6
- C18.3n.3

Type de corrélations

Coefficient de Pearson

Coefficient de Spearman


Coefficients partiels


Observations à utiliser

Observations complètes

Observations avec paires complètes

p-values par paires

 Aide

 Réinitialiser

- Coefficient de corrélation : mesure à quel point 2 variables changent ensemble (importance et sens de la relation)
- Pearson : relation linéaire (modification de l'une entraîne modification proportionnelle de l'autre)
- Spearman : relation monotone (les variables changent ensemble mais pas forcément de façon proportionnelle)



```

Rcmdr> cor(alldata[,c("C14.0", "C16.0", "C16.1n.7", "C16.1n.9", "C16SR", "C18.0",
Rcmdr+ "C18.1n.7", "C18.1n.9", "C18.2n.6", "C18.3n.3", "C18.3n.6", "C20.1n.9",
Rcmdr+ "C20.2n.6", "C20.3n.3", "C20.3n.6", "C20.3n.9", "C20.4n.6", "C20.5n.3",
Rcmdr+ "C22.4n.6", "C22.5n.3", "C22.5n.6", "C22.6n.3")], use="complete")

```

	C14.0	C16.0	C16.1n.7	C16.1n.9	C16SR
C14.0	1.000000000	0.009348349	0.97196895	0.65743596	0.09478165
C16.0	0.009348349	1.000000000	0.01841975	-0.37214015	0.21390606
C16.1n.7	0.971968954	0.018419753	1.000000000	0.64766480	0.09075282
C16.1n.9	0.657435963	-0.372140152	0.64766480	1.000000000	0.03900575
C16SR	0.094781651	0.213906059	0.09075282	0.03900575	1.000000000
C18.0	-0.590873033	0.555890766	-0.61195072	-0.89638629	-0.04547242
C18.1n.7	0.893836709	0.113774805	0.89299985	0.71481640	0.12219683
C18.1n.9	0.694888252	-0.204173703	0.70711689	0.91413764	0.06168371
C18.2n.6	-0.507051393	-0.662911221	-0.56976030	-0.03849041	-0.27834522
C18.3n.3	-0.157757498	-0.368367808	-0.04769143	-0.15461666	0.07045364
C18.3n.6	0.153286083	-0.052227022	0.13946596	0.02786462	-0.05821699
C20.1n.9	0.124795106	-0.504826290	0.04432566	0.18831207	-0.40068740
C20.2n.6	-0.320990267	-0.561811334	-0.39626192	0.04487598	-0.41655417
C20.3n.3	-0.190551705	-0.147447123	-0.07459729	-0.24958985	0.01241970
C20.3n.6	-0.449002513	0.397253433	-0.46817202	-0.58375462	-0.15078663
C20.3n.9	0.326138010	0.302601367	0.34643288	0.10692196	0.04370407
C20.4n.6	-0.310621238	0.175685522	-0.34817392	-0.38547994	-0.27629007
C20.5n.3	-0.331766866	0.262608682	-0.27643931	-0.52477362	0.10049333
C22.4n.6	-0.281655529	-0.303996540	-0.33117994	-0.12433592	-0.42311856
C22.5n.3	-0.272727372	0.182657062	-0.27191149	-0.45282778	0.18217094
C22.5n.6	-0.142763084	-0.060888607	-0.21063969	-0.16473036	-0.32573750
C22.6n.3	-0.479336339	0.454584010	-0.50198134	-0.59055129	0.28618519

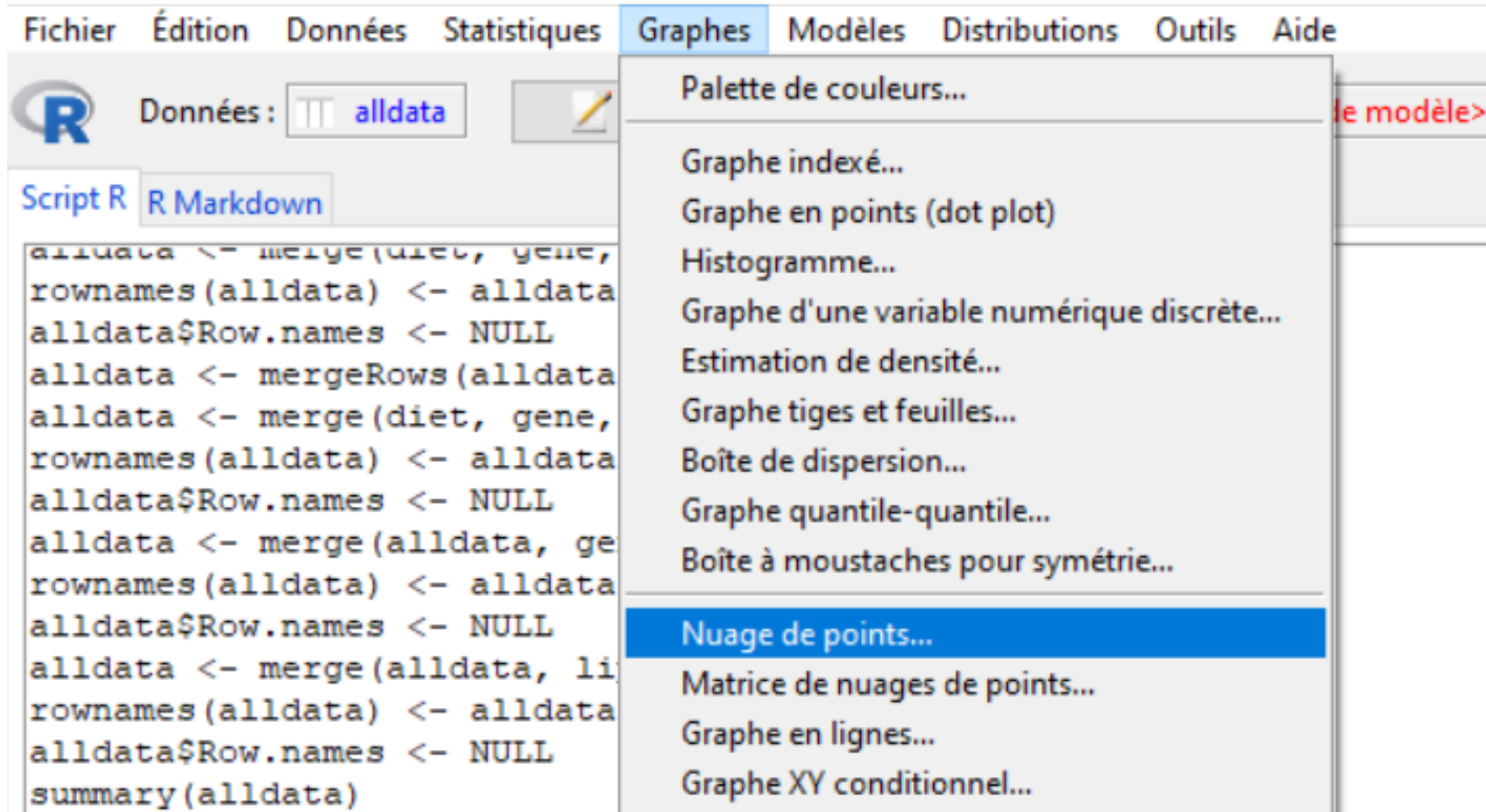
  

	C18.0	C18.1n.7	C18.1n.9	C18.2n.6	C18.3n.3
C14.0	-0.59087303	0.893836709	0.69488825	-0.50705139	-0.15775750
C16.0	0.55589077	0.113774805	-0.20417370	-0.66291122	-0.36836781
C16.1n.7	-0.61195072	0.892999848	0.70711689	-0.56976030	-0.04769143
C16.1n.9	-0.89638629	0.714816404	0.91413764	-0.03849041	-0.15461666
C16SR	-0.04547242	0.122196828	0.06168371	-0.27834522	0.07045364
C18.0	1.000000000	-0.572752923	-0.83994297	-0.08426218	-0.07850200
C18.1n.7	-0.57275292	1.000000000	0.80031523	-0.55637932	-0.31106295
C18.1n.9	0.83994297	0.80031523	1.000000000	0.72477087	0.72467607



# Nuage de points (scatterplot)

R Commander



The screenshot shows the R Commander interface. The menu bar includes 'Fichier', 'Édition', 'Données', 'Statistiques', 'Graphes', 'Modèles', 'Distributions', 'Outils', and 'Aide'. The 'Graphes' menu is open, displaying the following options: 'Palette de couleurs...', 'Graphe indexé...', 'Graphe en points (dot plot)', 'Histogramme...', 'Graphe d'une variable numérique discrète...', 'Estimation de densité...', 'Graphe tiges et feuilles...', 'Boîte de dispersion...', 'Graphe quantile-quantile...', 'Boîte à moustaches pour symétrie...', 'Nuage de points...' (highlighted in blue), 'Matrice de nuages de points...', 'Graphe en lignes...', and 'Graphe XY conditionnel...'. The 'Données' field shows 'alldata'. The script editor contains the following R code:

```
alldata <- merge(diet, gene,  
rownames(alldata) <- alldata  
alldata$Row.names <- NULL  
alldata <- mergeRows(alldata  
alldata <- merge(diet, gene,  
rownames(alldata) <- alldata  
alldata$Row.names <- NULL  
alldata <- merge(alldata, ge  
rownames(alldata) <- alldata  
alldata$Row.names <- NULL  
alldata <- merge(alldata, li  
rownames(alldata) <- alldata  
alldata$Row.names <- NULL  
summary(alldata)
```

## Nuage de points

Données Options

variable x (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

variable y (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

Graphe par groupe...

Expression de sélection

<tous les cas valides>

## Nuage de points

Données Options

Options du graphe

- Décalages aléatoires x
- Décalages aléatoires y
- Axe X logarithmique
- Axe Y logarithmique
- Boîte à dispersion marginales
- Ligne des moindres carrés
- Courbe de lissage
- Afficher l'étendue

Fenêtre de lissage: 50

Graphe d'ellipses de concentration

Niveaux de concentration: .5, .9

Identifier des points

- Automatiquement
- Avec la souris de manière interactive
- Ne pas identifier

Nombre de points à identifier: 2

Étiquettes et points du graphe

libellé de l'axe X: <auto>

libellé de l'axe Y: <auto>

Titre du graphe: <auto>

Caractères à utiliser: <auto>

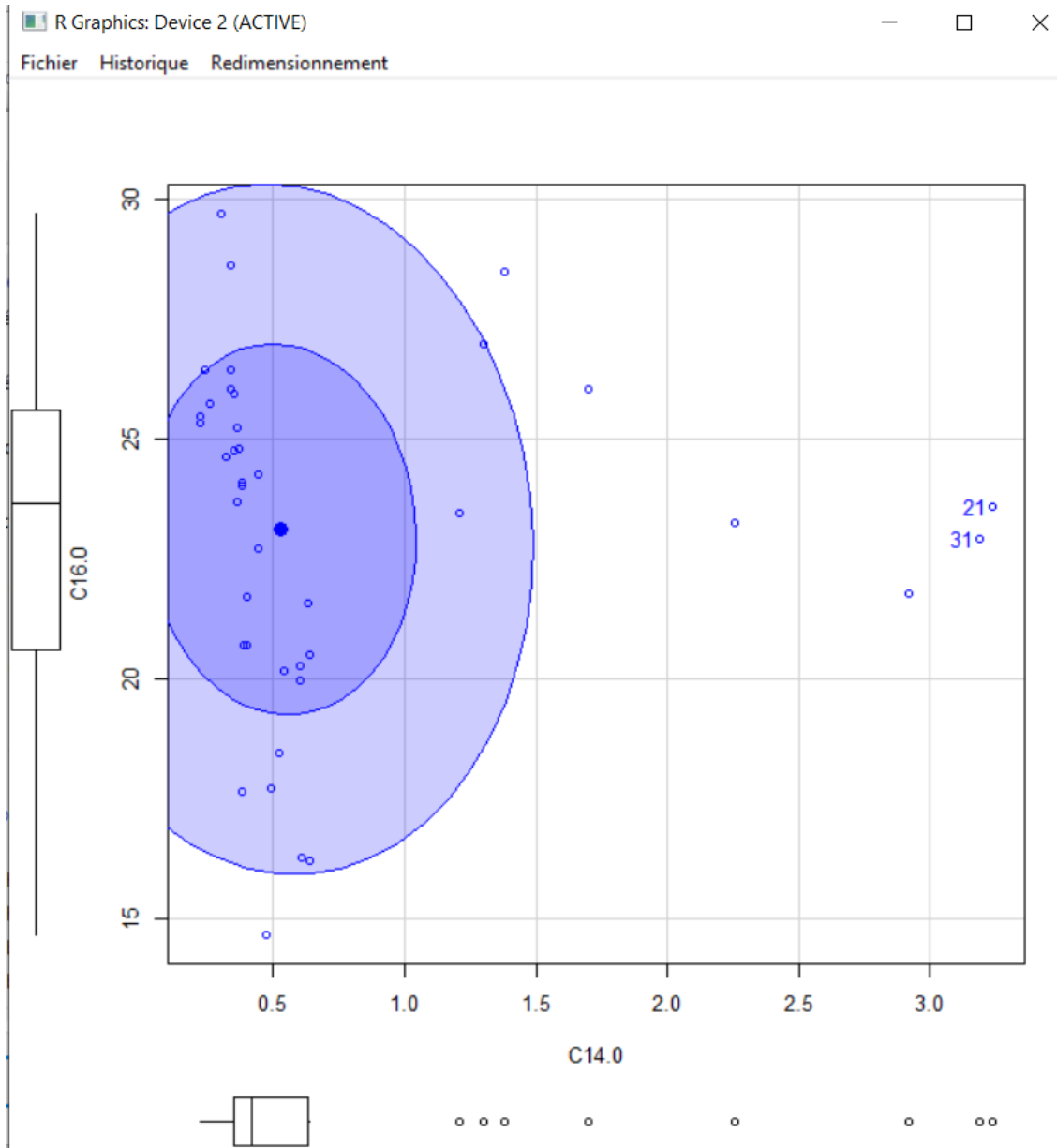
Taille de point: 1.0

Taille du texte des axes: 1.0

Taille du texte des libellés d'axes: 1.0

Position de la légende


- Au dessus de graphe
- En haut à gauche
- En haut à droite
- En bas à gauche
- En bas à droite



Corrélation de Pearson = 0,009  
Cohérent avec le graphe

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données:  

Script R R Markdown

```
alldata <- merge(diet, gene,
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- mergeRows(alldata
alldata <- merge(diet, gene,
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- merge(alldata, ge
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- merge(alldata, li
rownames(alldata) <- alldata
```

- Palette de couleurs...
- Graphe indexé...
- Graphe en points (dot plot)
- Histogramme...
- Graphe d'une variable numérique discrète...
- Estimation de densité...
- Graphe tiges et feuilles...
- Boîte de dispersion...
- Graphe quantile-quantile...
- Boîte à moustaches pour symétrie...
- Nuage de points...**
- Matrice de nuages de points...
- Graphe en lianes...

## R Nuage de points

Données Options

variable x (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

variable y (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

Graphe par groupe...

Expression de sélection

<tous les cas valides>

## R Nuage de points

Données Options

Options du graphe

- Décalages aléatoires x
- Décalages aléatoires y
- Axe X logarithmique
- Axe Y logarithmique
- Boîte à dispersion marginales
- Ligne des moindres carrés
- Courbe de lissage
- Afficher l'étendue

Fenêtre de lissage

Graphe d'ellipses de concentration

Niveaux de concentration :

Identifier des points

- Automatiquement
- Avec la souris de manière interactive
- Ne pas identifier

Nombre de points à identifier :

Étiquettes et points du graphe

libellé de l'axe X

libellé de l'axe Y

Titre du graphe

Caractères à utiliser

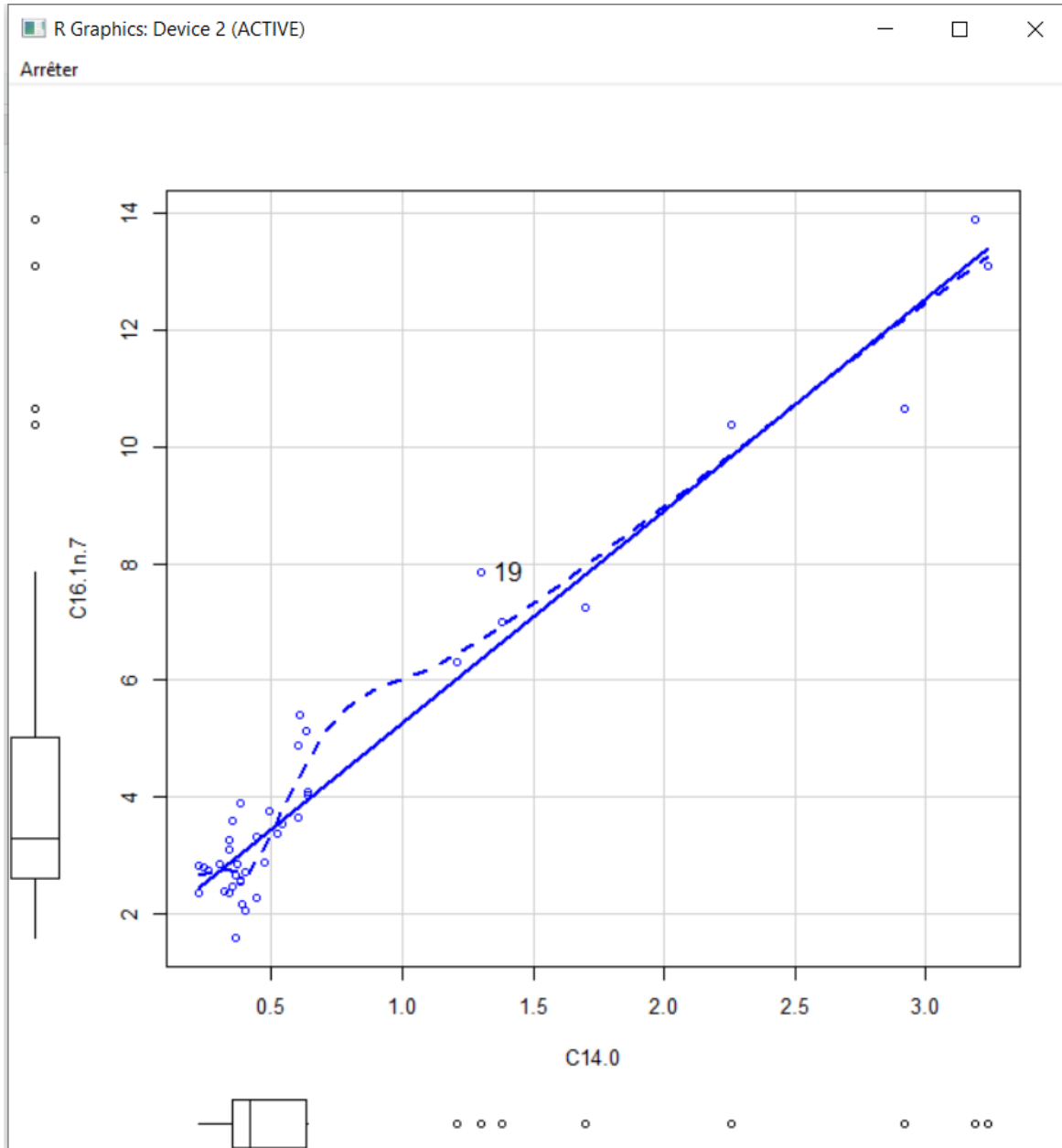
Taille de point

Taille du texte des axes

Taille du texte des libellés d'axes

Position de la légende

- Au dessus de graphe
- En haut à gauche
- En haut à droite
- En bas à gauche
- En bas à droite




Corrélation de Pearson = 0,97  
Moins net si on se concentre sur les valeurs faibles qui sont la grand majorité  
La corrélation de Pearson est tirée par les points extrêmes  
Corrélation Spearman = 0,78

## TO DO

- **Refaire le graphe de C16.1n.7 en fonction de C14 mais avec les échelles log**
- **Conclusion?**

R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données:  

Script R R Markdown

```
alldata <- merge(diet, gene,
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- mergeRows(alldata
alldata <- merge(diet, gene,
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- merge(alldata, ge
rownames(alldata) <- alldata
alldata$Row.names <- NULL
alldata <- merge(alldata, li
rownames(alldata) <- alldata
```

- Palette de couleurs...
- Graphe indexé...
- Graphe en points (dot plot)
- Histogramme...
- Graphe d'une variable numérique discrète...
- Estimation de densité...
- Graphe tiges et feuilles...
- Boîte de dispersion...
- Graphe quantile-quantile...
- Boîte à moustaches pour symétrie...
- Nuage de points...**
- Matrice de nuages de points...
- Graphe en lianes...



## Nuage de points

Données Options

variable x (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

variable y (une)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

Graphe par groupe...

Expression de sélection

<tous les cas valides>

## Nuage de points

Données Options

Options du graphe

- Décalages aléatoires x
- Décalages aléatoires y
- Axe X logarithmique
- Axe Y logarithmique
- Boîte à dispersion marginales
- Ligne des moindres carrés
- Courbe de lissage
- Afficher l'étendue

Fenêtre de lissage 50

Graphe d'ellipses de concentration

Niveaux de concentration : .5, .9

Identifier des points

- Automatiquement
- Avec la souris de manière interactive
- Ne pas identifier

Nombre de points à identifier : 2

Étiquettes et points du graphe

libellé de l'axe X <auto>

libellé de l'axe Y <auto>

Titre du graphe <auto>

Caractères à utiliser <auto>

Taille de point 1.0

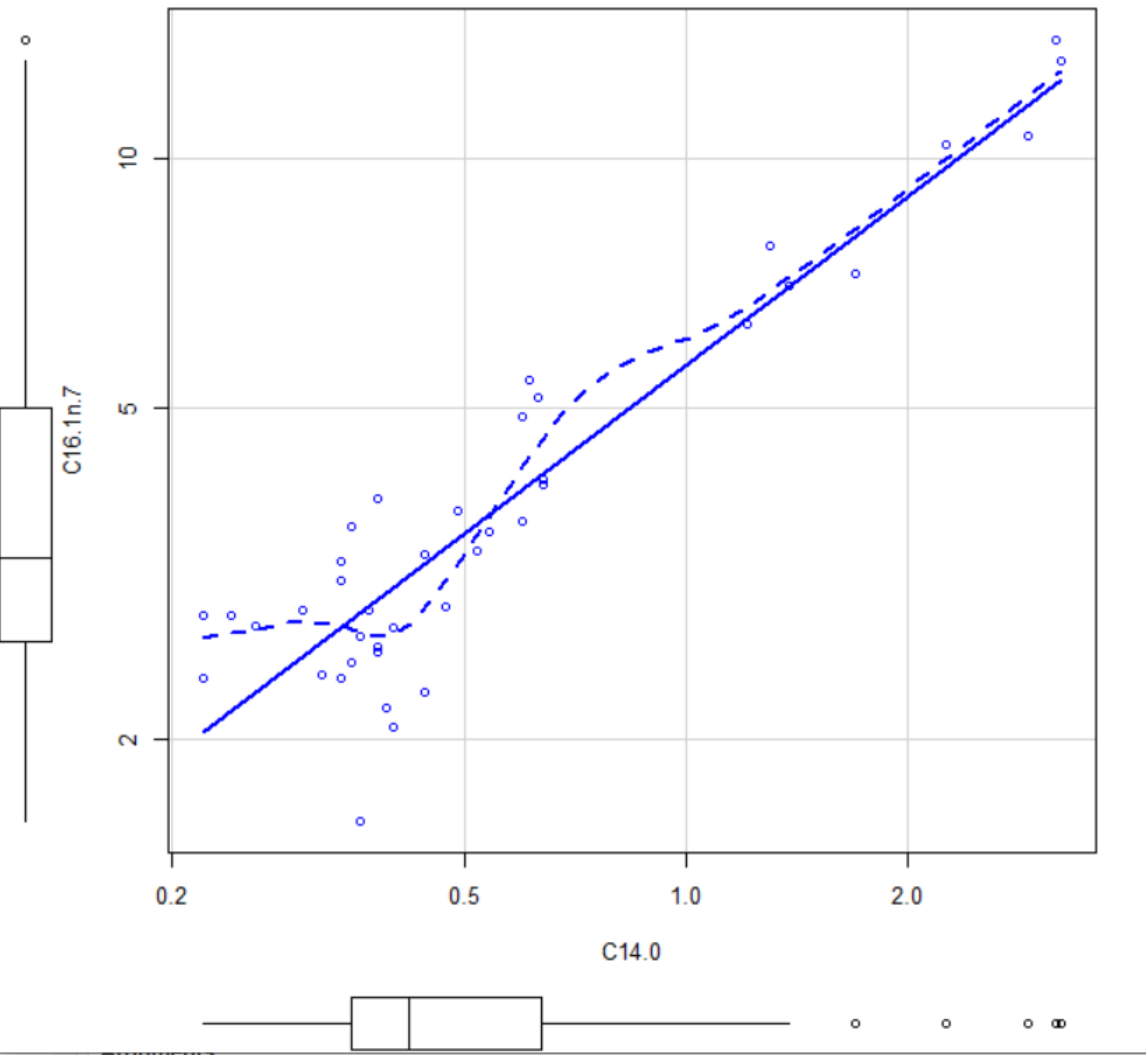
Taille du texte des axes 1.0

Taille du texte des libellés d'axes 1.0

Position de la légende

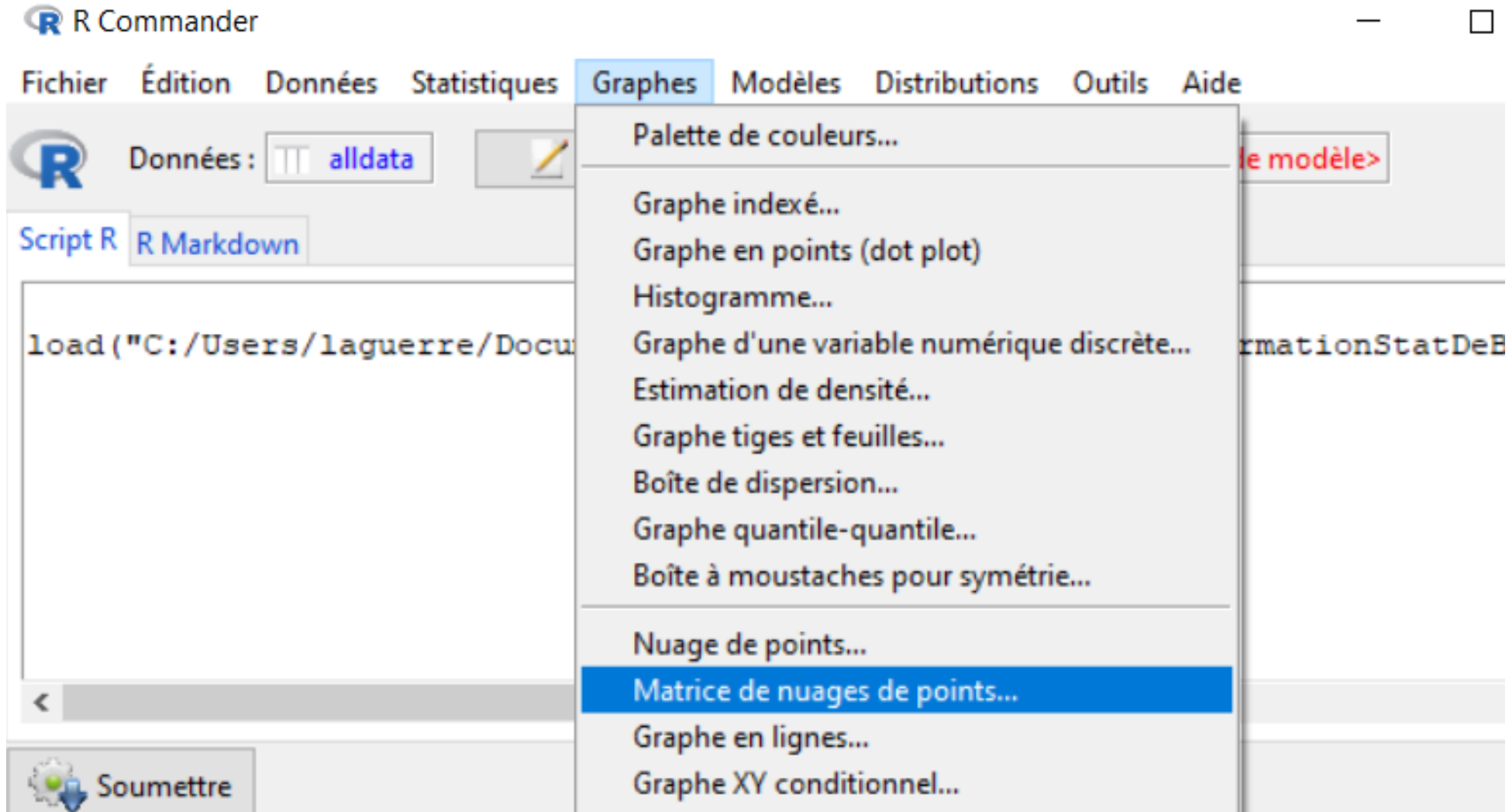
- Au dessus de graphe
- En haut à gauche
- En haut à droite
- En bas à gauche
- En bas à droite

Arrêter




Corrélation de Pearson = 0,92  
Corrélation de Pearson = 0,78

# Matrice de nuages de points



R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données :  


Script R R Markdown

```
load("C:/Users/laguerre/Docu
```

le modèle>

rmationStatDeB

- Palette de couleurs...
- Graphe indexé...
- Graphe en points (dot plot)
- Histogramme...
- Graphe d'une variable numérique discrète...
- Estimation de densité...
- Graphe tiges et feuilles...
- Boîte de dispersion...
- Graphe quantile-quantile...
- Boîte à moustaches pour symétrie...
- Nuage de points...
- Matrice de nuages de points...**
- Graphe en lignes...
- Graphe XY conditionnel...

 Soumettre

## R Matrice de nuages de points

Données Options

Sélectionnez les variables (trois au plus)

- C14.0
- C16.0
- C16.1n.7
- C16.1n.9
- C16SR
- C18.0

Graphe par groupe...

Expression de sélection

<tous les cas valides>

## R Matrice de nuages de points

Données Options

Sur la diagonale

- Graphes de densité
- Histogrammes
- Boîtes de dispersion
- Nuage de points à une dimension
- Graphes quantiles-quantiles normaux
- Rien (vide)

Titre du graphe <auto>

Autres Options

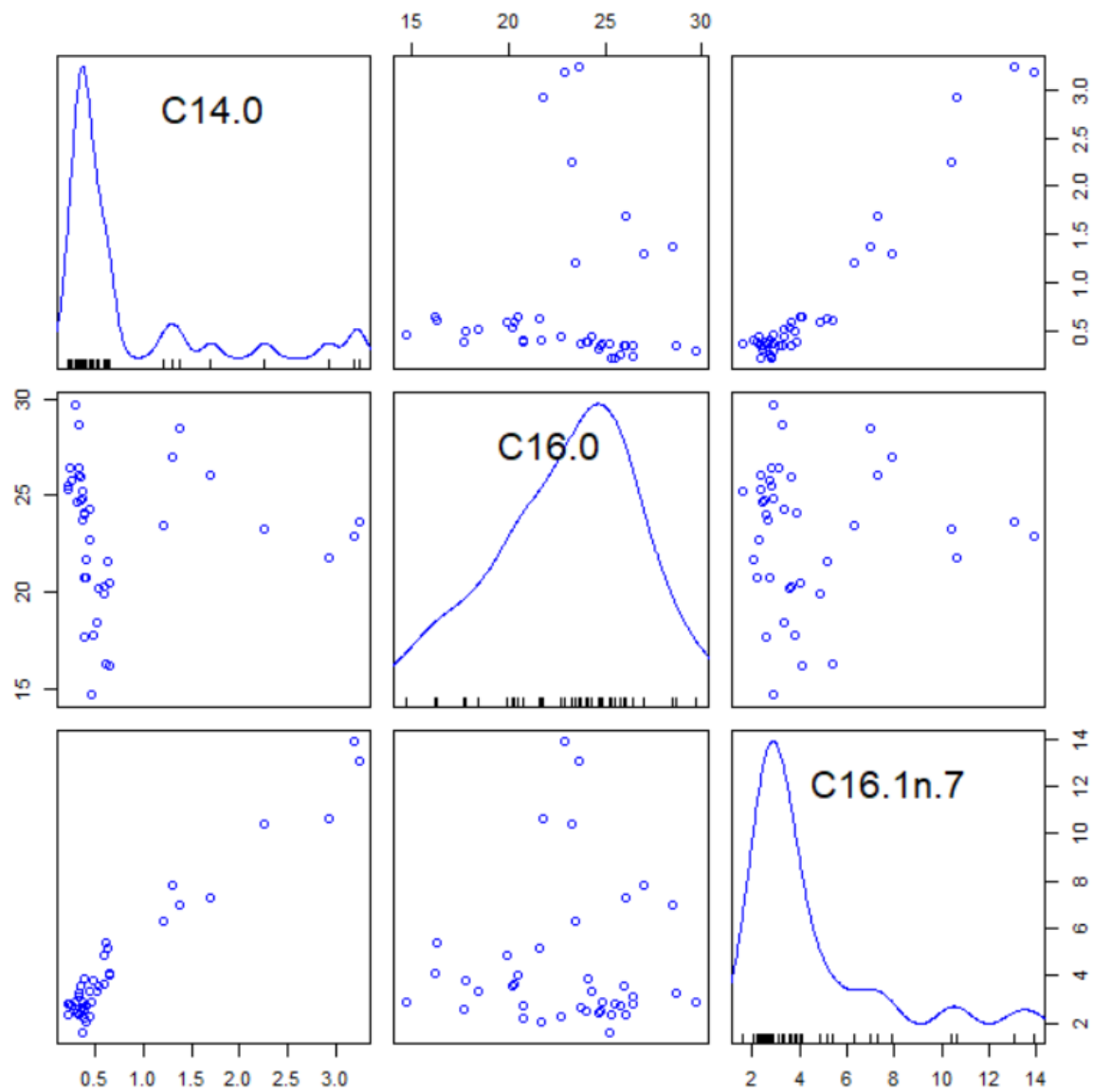
- Lignes des moindres carrés
- Courbes de lissage
- Afficher l'étendue

Fenêtre de lissage 50

Nombre de points à identifier 0 dans chaque panneau et groupe

- Graphe d'ellipses de concentration

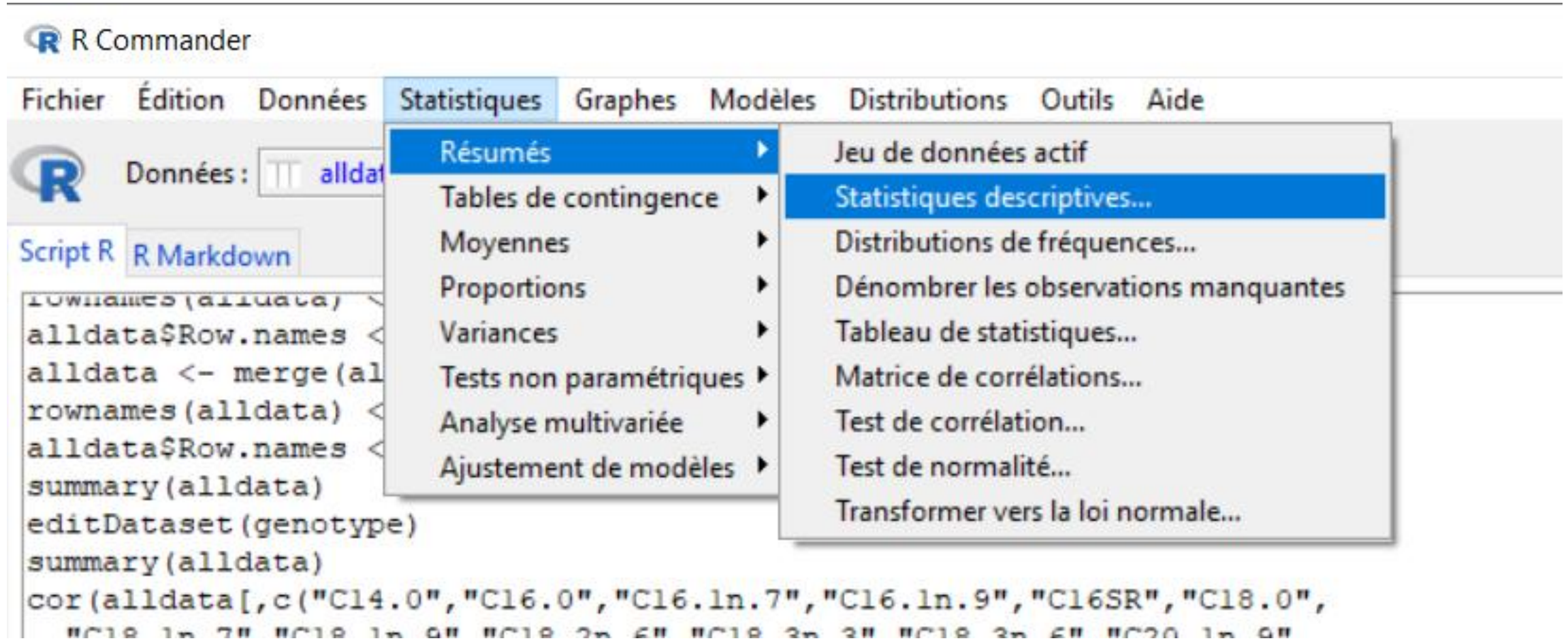
Niveaux de concentration : .5, .9



## **Partie 3.2**

# **Une variable quantitative et une variable qualitative**

# Statistiques descriptive par groupe



R Commander

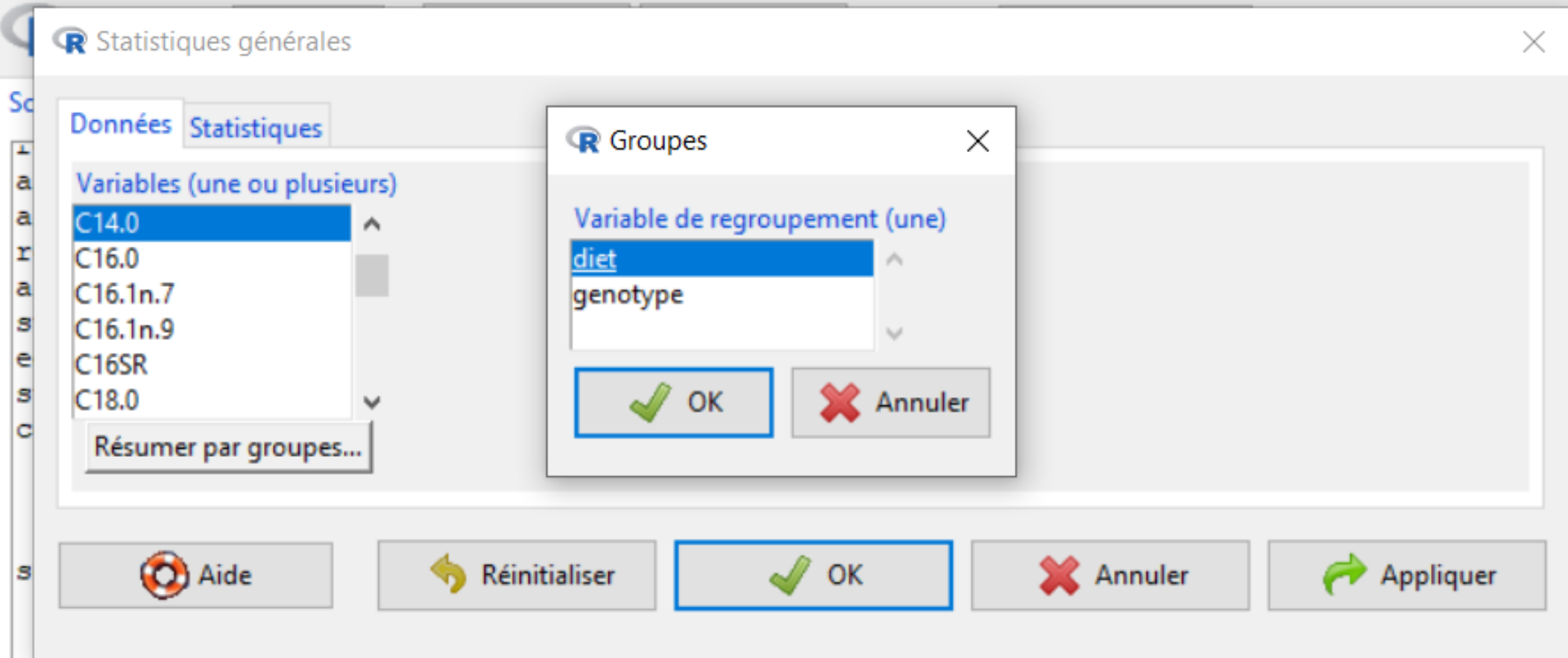
Fichier Édition Données **Statistiques** Graphes Modèles Distributions Outils Aide

Données :

Script R R Markdown

```
rownames(alldata) <-  
alldata$Row.names <-  
alldata <- merge(alldata, rownames(alldata)) <-  
alldata$Row.names <-  
summary(alldata)  
editDataset(genotype)  
summary(alldata)  
cor(alldata[,c("C14.0", "C16.0", "C16.1n.7", "C16.1n.9", "C16SR", "C18.0",  
"C18.1n.7", "C18.1n.9", "C18.2n.5", "C18.2n.3", "C18.2n.6", "C20.1n.9")])
```

- Résumés
  - Jeu de données actif
  - Statistiques descriptives...**
  - Distributions de fréquences...
  - Dénombrer les observations manquantes
  - Tableau de statistiques...
  - Matrice de corrélations...
  - Test de corrélation...
  - Test de normalité...
  - Transformer vers la loi normale...
- Tables de contingence
- Moyennes
- Proportions
- Variances
- Tests non paramétriques
- Analyse multivariée
- Ajustement de modèles



```
scatterplot(C16.ln.7~C14.0, regLine=TRUE, smooth=list(span=0.5,  
spread=FALSE), id=list(method=identif), boxplots=TRUE, data=alldata)
```



Données Statistiques

- Moyenne
- Erreur type de la moyenne
- Coefficient de variation
- Coefficient de dissymétrie
- Coefficient d'aplatissement
- Quantiles :
- Ecart type
- Écart interquartile
- Nombre d'occurrences par intervalles
- Type 1
- Type 2
- Type 3



Aide



Réinitialiser



OK



Annuler



Appliquer

```

Rcmdr> numSummary(alldata[, "c14.0", drop=FALSE], groups=alldata$diet,
Rcmdr+   statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "cv", "skewness",
Rcmdr+   "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")

```

	mean	sd	se(mean)	IQR	cv	skewness	kurtosis	0%
coc	2.15000	0.86818365	0.30694927	1.6275	0.4038064	0.25771547	-2.106823	1.21
fish	0.32375	0.08551316	0.03023347	0.1425	0.2641333	-0.04669076	-1.783927	0.22
lin	0.45500	0.15757085	0.05570971	0.2750	0.3463096	0.06426685	-1.954484	0.24
ref	0.48625	0.11891624	0.04204324	0.2225	0.2445578	0.07011804	-2.041715	0.34
sun	0.40000	0.07151423	0.02528410	0.0600	0.1787856	1.23427395	1.106966	0.32

	25%	50%	75%	100%	C14.0:n
coc	1.3600	1.98	2.9875	3.24	8
fish	0.2500	0.33	0.3925	0.44	8
lin	0.3400	0.42	0.6150	0.64	8
ref	0.3775	0.48	0.6000	0.64	8
sun	0.3575	0.38	0.4175	0.54	8

## TO DO

- **Faire de même avec C16 et C22.4n.6**
- **Commenter**



Données

Statistiques

Variables (une ou plusieurs)

- BIEN
- BSEP
- C14.0
- C16.0
- C16.1n.7
- C16.1n.9

Résumer par : diet



Aide



Réinitialiser



OK



Annuler



Appliquer

```

Rcmdr> numSummary(alldata[, "C16.0", drop=FALSE], groups=alldata$diet,
Rcmdr+   statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "cv", "skewness", "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")
Rcmdr+

```

	mean	sd	se(mean)	IQR	cv	skewness	kurtosis	0%
coc	24.5675	2.324434	0.8218115	3.1000	0.09461418	0.71955056	-0.8033068	21.79
fish	24.5800	2.834633	1.0021940	3.0700	0.11532273	0.42299005	0.4408166	20.71
lin	22.4025	5.114187	1.8081381	9.1000	0.22828643	-0.22195189	-2.0962394	16.20
ref	22.2925	2.809737	0.9933920	4.2350	0.12603956	-0.02022623	-1.9230624	18.44
sun	21.2875	3.694771	1.3062990	4.6425	0.17356530	-0.86253865	-0.3353724	14.65

	25%	50%	75%	100%	C16.0:n
coc	23.1800	23.520	26.2800	28.49	8
fish	22.4725	25.285	25.5425	29.72	8
lin	17.3525	23.755	26.4525	28.63	8
ref	20.1750	22.290	24.4100	26.04	8
sun	19.5450	22.200	24.1875	24.76	8

Données Statistiques

Variables (une ou plusieurs)

- C20.4n.6
- C20.5n.3
- C22.4n.6
- C22.5n.3
- C22.5n.6
- C22.6n.3

Résumer par : diet

Aide Réinitialiser OK Annuler Appliquer

```

Rcmdr> numSummary(alldata[, "c22.4n.6", drop=FALSE], groups=alldata$diet,
Rcmdr+   statistics=c("mean", "sd", "se(mean)", "IQR", "quantiles", "cv", "ske
Rcmdr+   "kurtosis"), quantiles=c(0,.25,.5,.75,1), type="2")

```

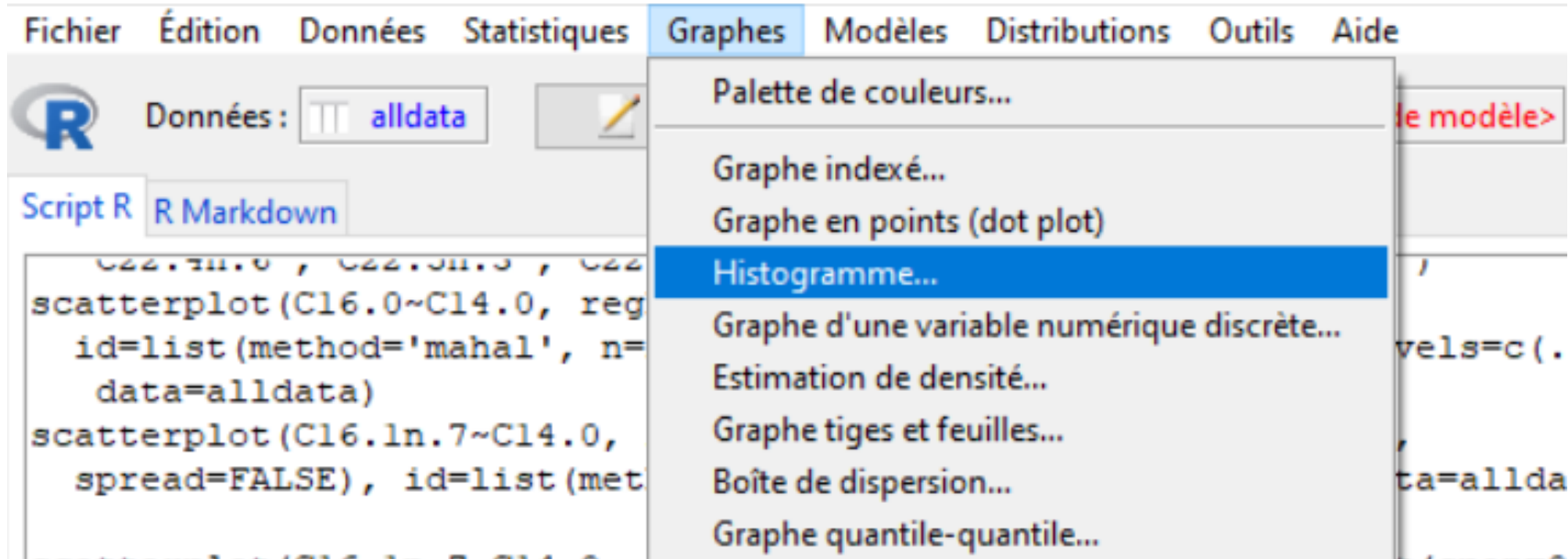
	mean	sd	se(mean)	IQR	cv	skewness	kurtosis	0%
coc	0.01500	0.02828427	0.01000000	0.0125	1.8856181	1.59099026	0.968750	0.00
fish	0.00000	0.00000000	0.00000000	0.0000	NA	NaN	NaN	0.00
lin	0.02125	0.04155461	0.01469177	0.0150	1.9555111	1.85723878	2.633546	0.00
ref	0.24500	0.07151423	0.02528410	0.1325	0.2918948	-0.44996316	-1.769433	0.14
sun	0.63875	0.06577831	0.02325614	0.0800	0.1029797	0.05527641	-0.826992	0.54

	25%	50%	75%	100%	c22.4n.6:n
coc	0.0000	0.000	0.0125	0.07	8
fish	0.0000	0.000	0.0000	0.00	8
lin	0.0000	0.000	0.0150	0.11	8
ref	0.1775	0.265	0.3100	0.32	8
sun	0.5950	0.640	0.6750	0.73	8

# Histogramme par groupe

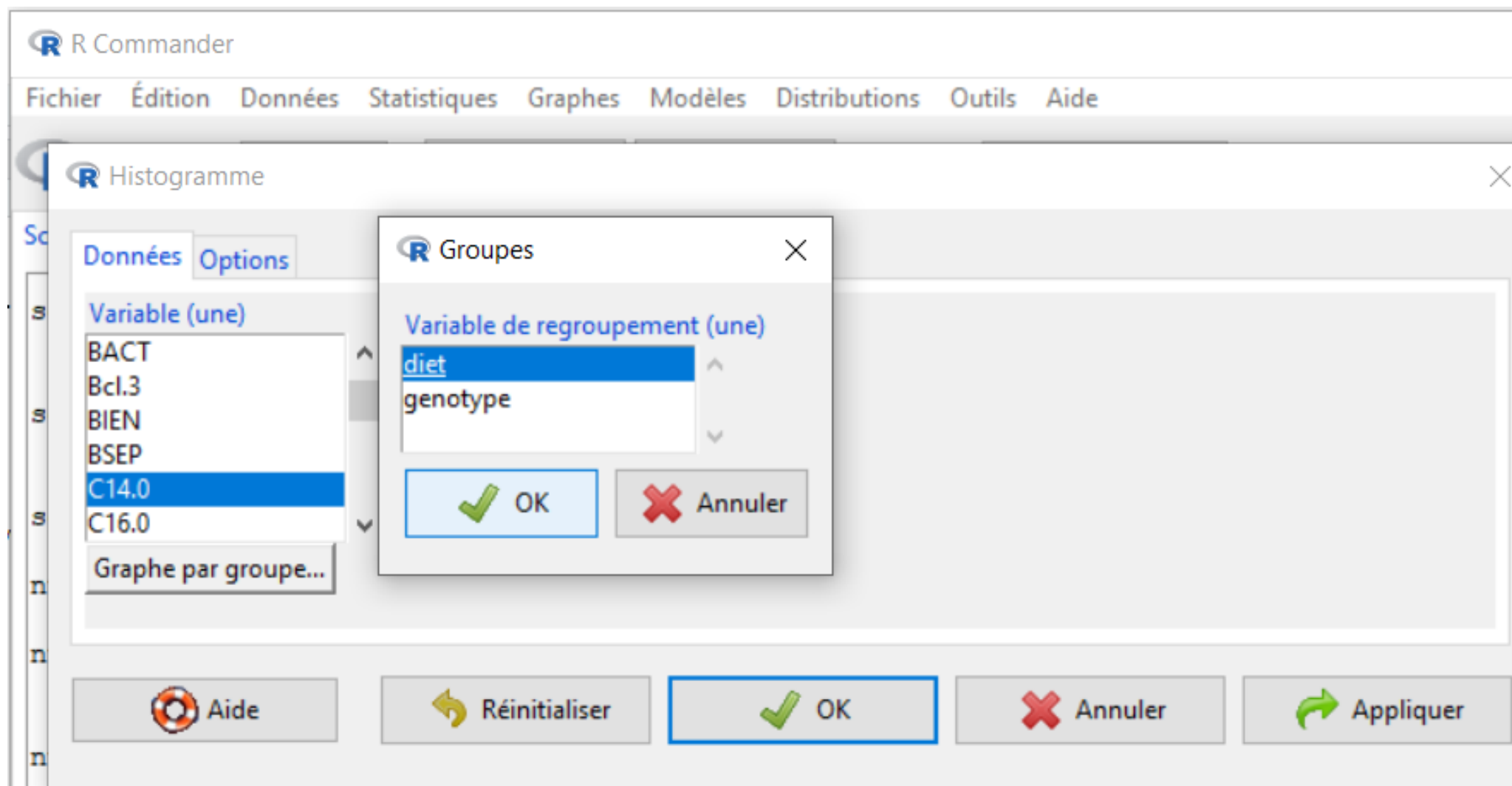
R Commander

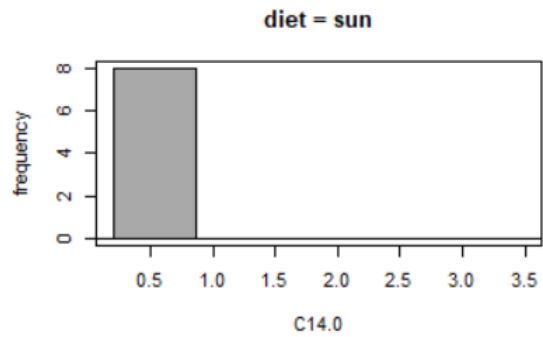
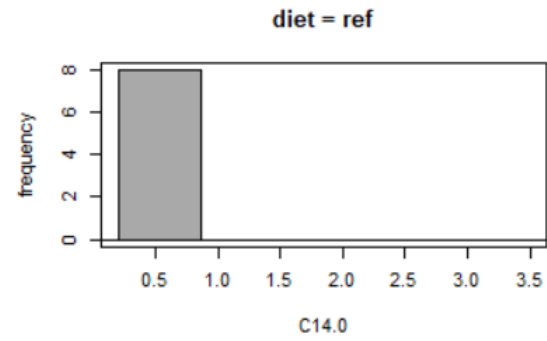
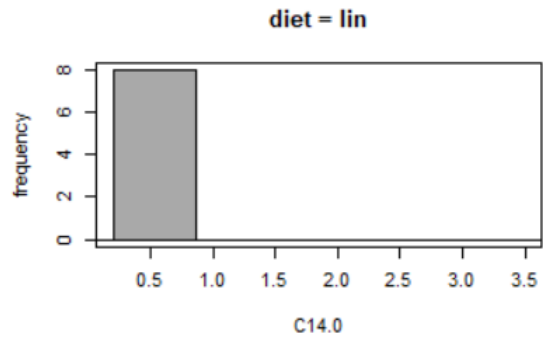
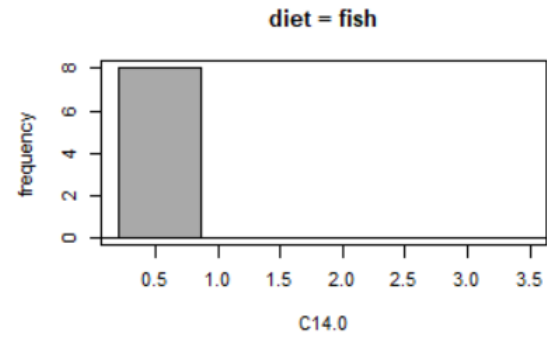
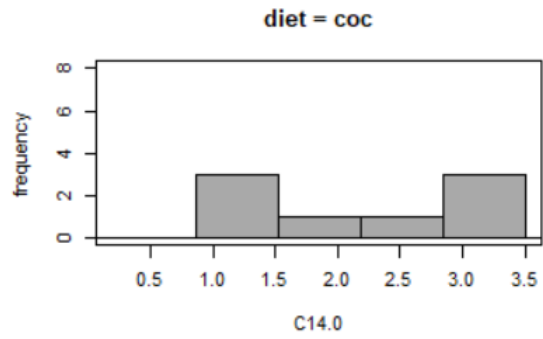


The screenshot shows the R Commander interface. The menu bar includes 'Fichier', 'Édition', 'Données', 'Statistiques', 'Graphes', 'Modèles', 'Distributions', 'Outils', and 'Aide'. The 'Graphes' menu is open, displaying options such as 'Palette de couleurs...', 'Graphe indexé...', 'Graphe en points (dot plot)', 'Histogramme...', 'Graphe d'une variable numérique discrète...', 'Estimation de densité...', 'Graphe tiges et feuilles...', 'Boîte de dispersion...', and 'Graphe quantile-quantile...'. The 'Histogramme...' option is highlighted in blue. In the background, the 'Données' field shows 'alldata' and the script editor contains R code for scatterplots.

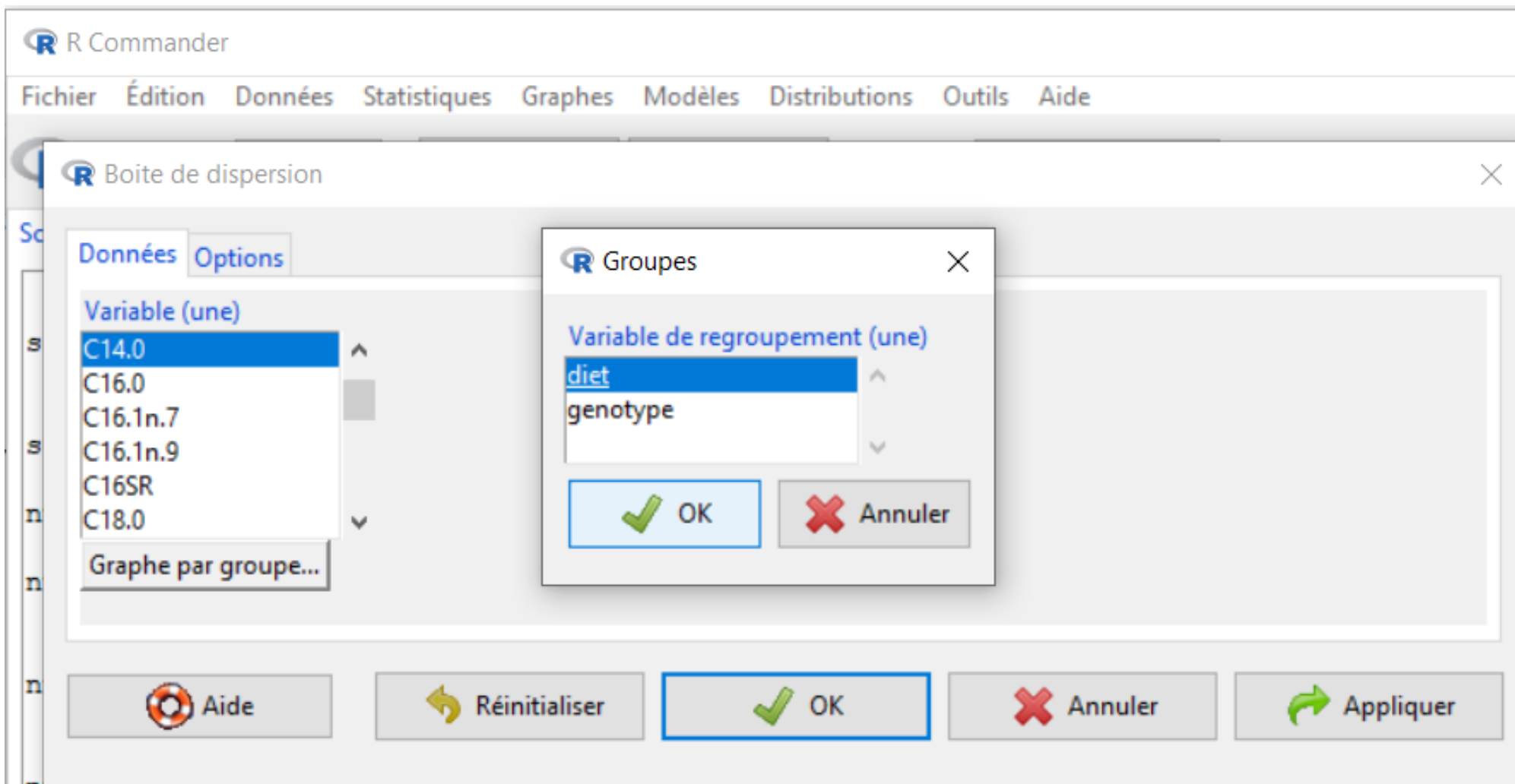
```
scatterplot(C16.0~C14.0, reg  
id=list(method='mahal', n=  
data=alldata)  
scatterplot(C16.ln.7~C14.0,  
spread=FALSE), id=list(met
```

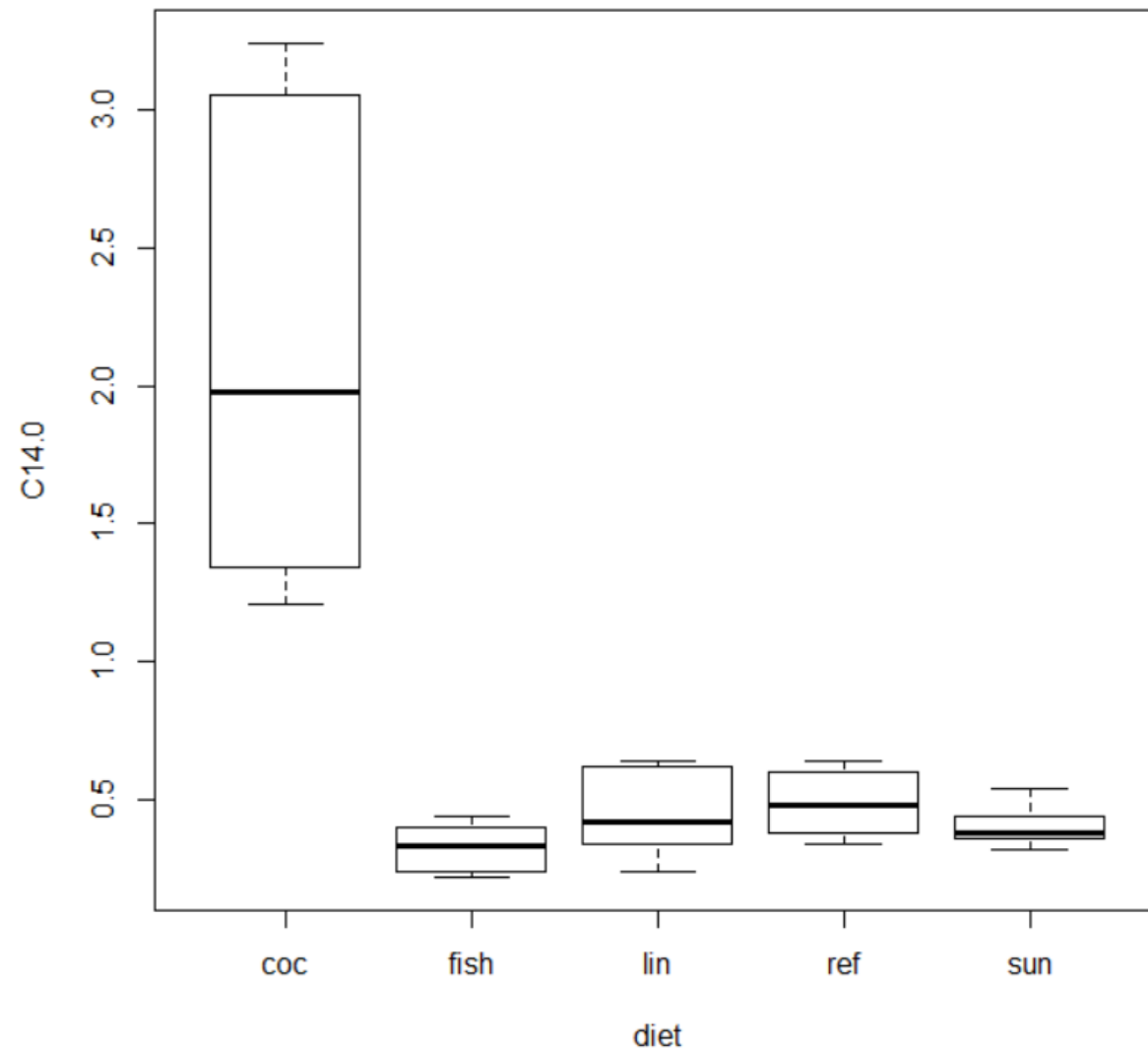




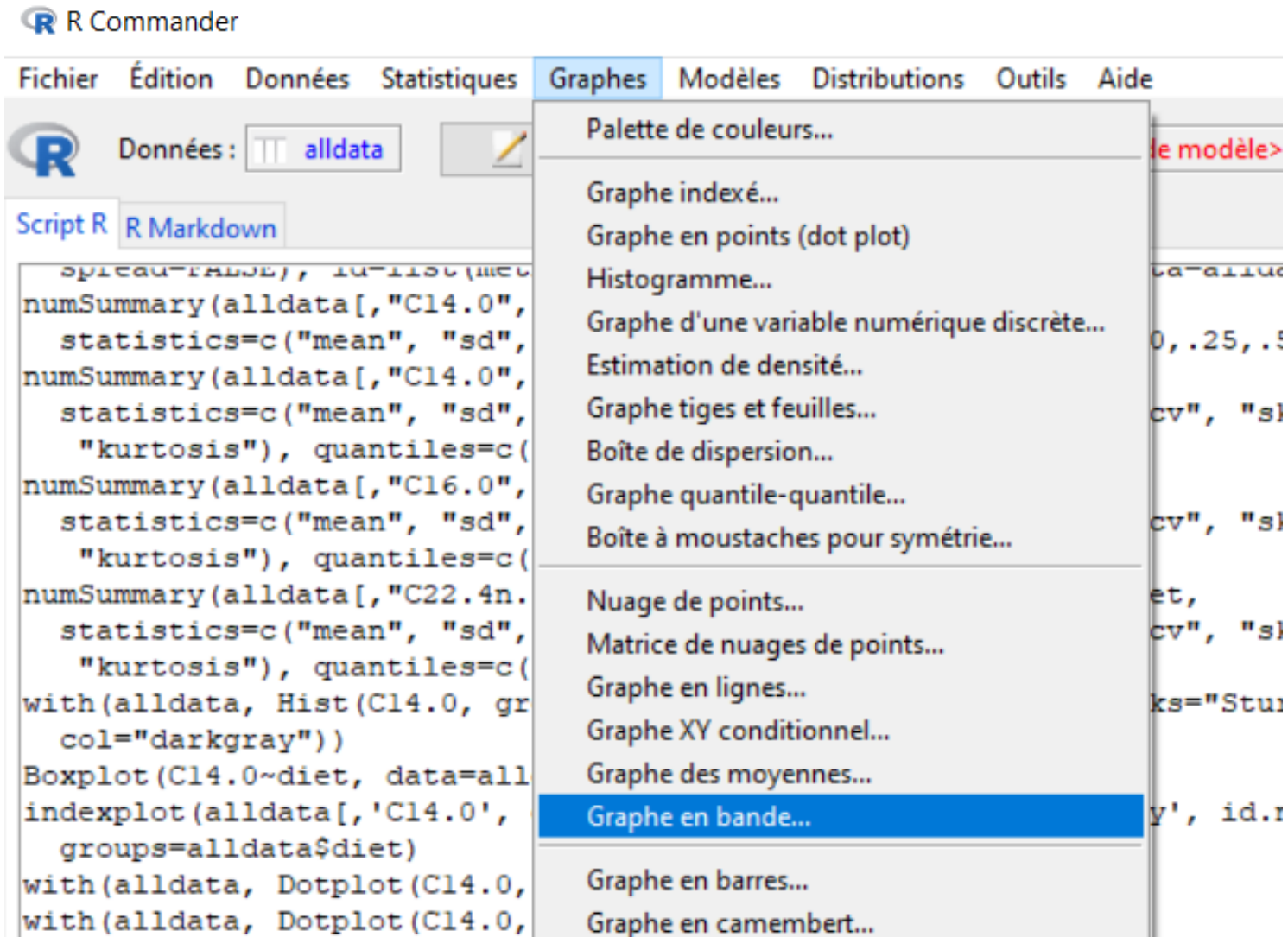









# Graphe en bandes (strip-chart)



R Commander

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données:  

Script R R Markdown

```
opieau=FALSE), id=1150 (me...
numSummary(alldata[, "C14.0",
  statistics=c("mean", "sd",
numSummary(alldata[, "C14.0",
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
numSummary(alldata[, "C16.0",
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
numSummary(alldata[, "C22.4n.
  statistics=c("mean", "sd",
    "kurtosis"), quantiles=c(
with(alldata, Hist(C14.0, gr
  col="darkgray"))
Boxplot(C14.0~diet, data=all
indexplot(alldata[, 'C14.0',
  groups=alldata$diet)
with(alldata, Dotplot(C14.0,
with(alldata, Dotplot(C14.0,
```

le modèle>

ca=aiiuc

0, .25, .5

cv", "s)

cv", "s)

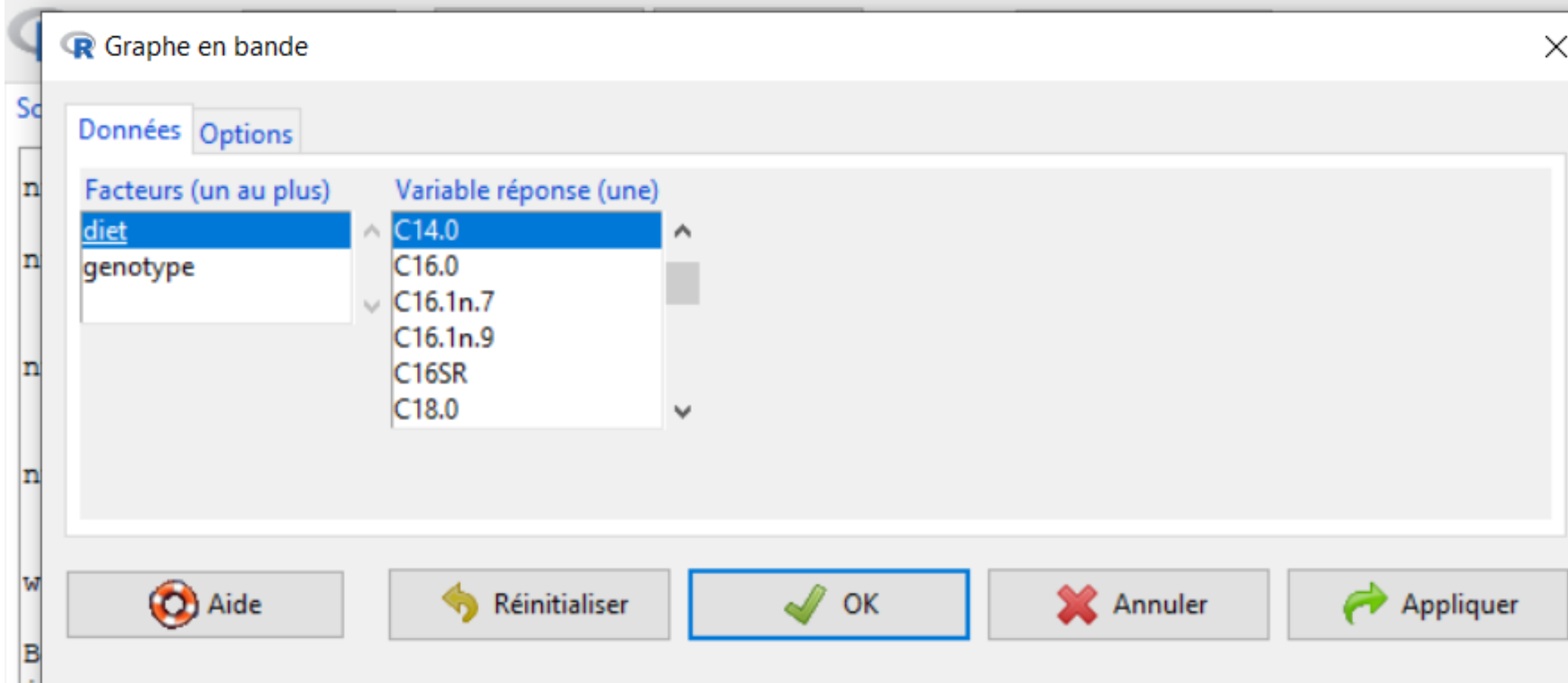
et,

cv", "s)

ks="Stu

y', id.r

- Palette de couleurs...
- Graphe indexé...
- Graphe en points (dot plot)
- Histogramme...
- Graphe d'une variable numérique discrète...
- Estimation de densité...
- Graphe tiges et feuilles...
- Boîte de dispersion...
- Graphe quantile-quantile...
- Boîte à moustaches pour symétrie...
- Nuage de points...
- Matrice de nuages de points...
- Graphe en lignes...
- Graphe XY conditionnel...
- Graphe des moyennes...
- Graphe en bande...**
- Graphe en barres...
- Graphe en camembert...



Données Options

Valeurs dupliquées

- Données empilées
- Décalage

Étiquettes du graphe

libellé de l'axe X   
< >

libellé de l'axe Y   
< >

Titre du graphe   
< >



Aide



Réinitialiser



OK

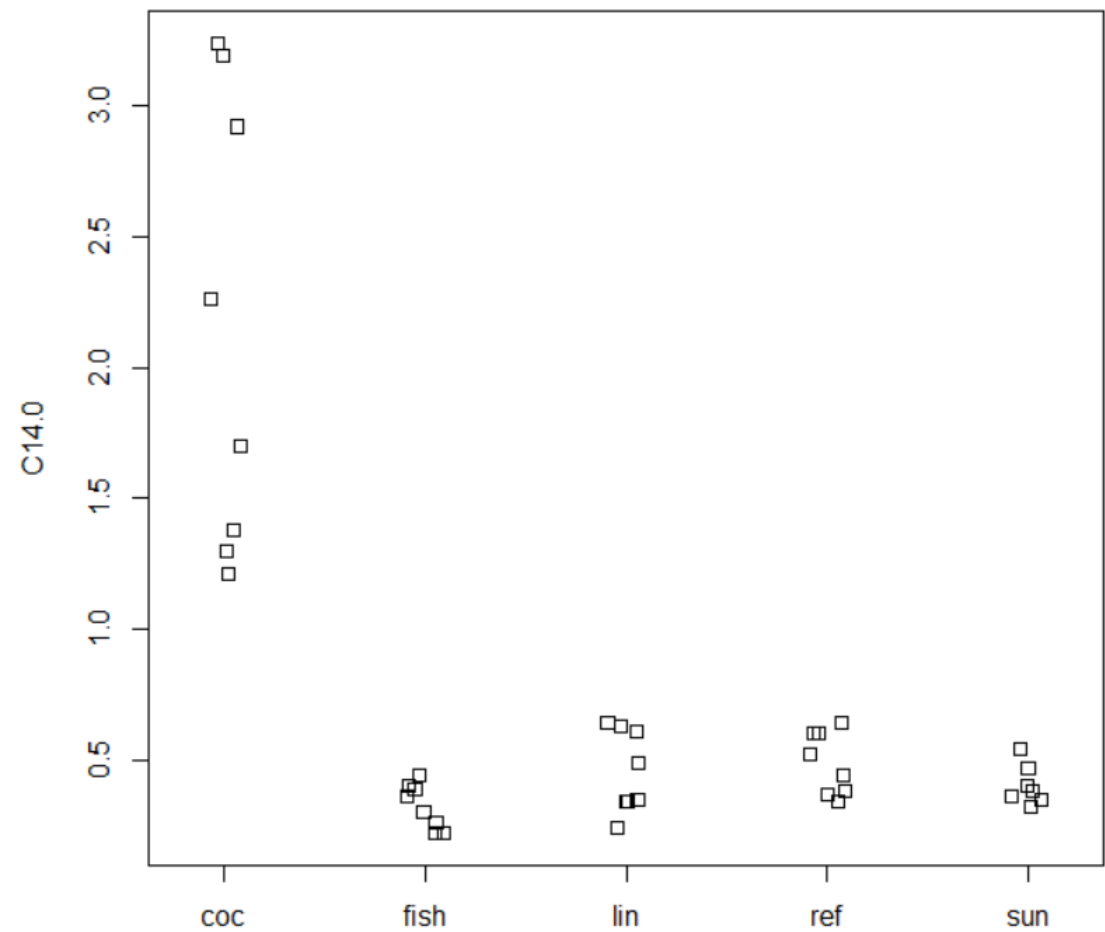


Annuler



Appliquer





## TO DO

- **Regarder les distributions conditionnelles à diet de C16 et C22.4n.6**
- **Commenter**

## R Graphe en bande

Données Options

Facteurs (un au plus)

diet

genotype

Variable réponse (une)

BACT

Bcl.3

BIEN

BSEP

C14.0

C16.0



Aide



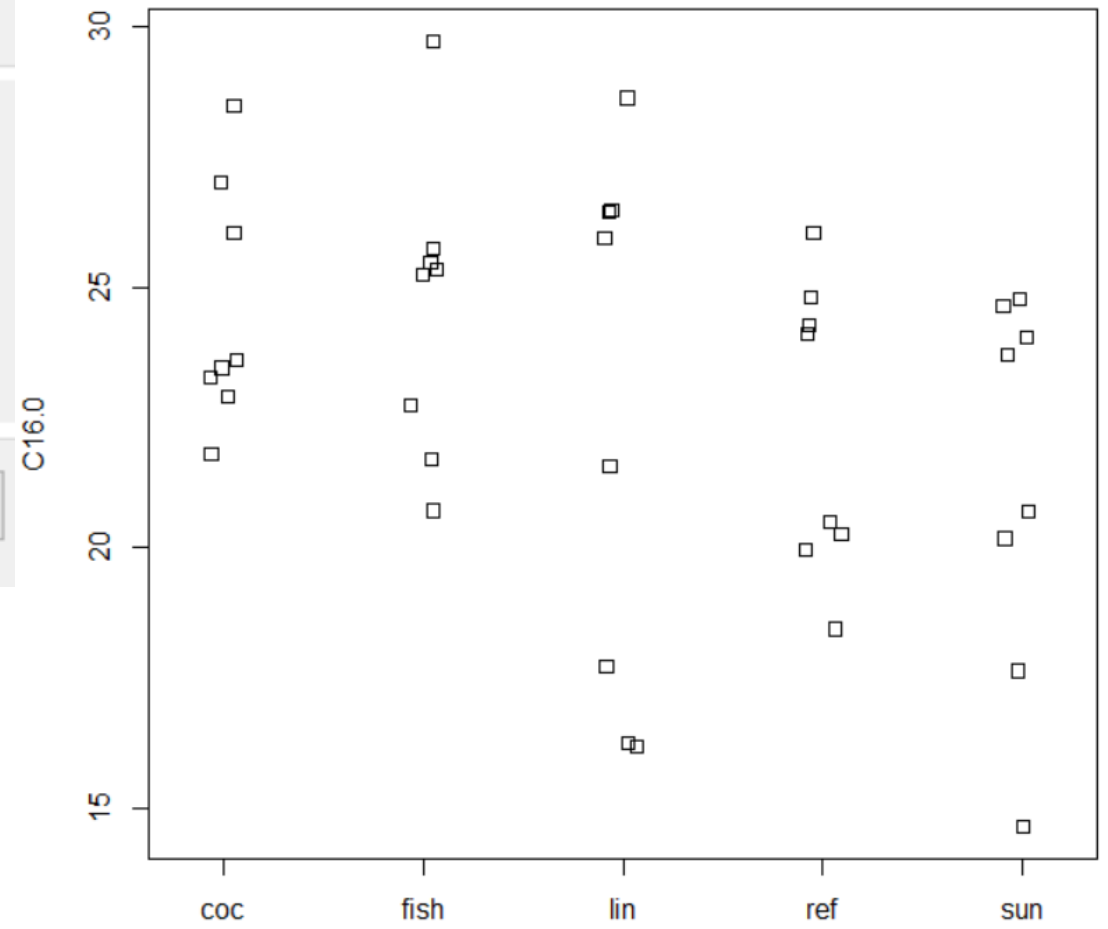
Réinitialiser



OK



Annuler



## R Graphe en bande

Données Options

Facteurs (un au plus)

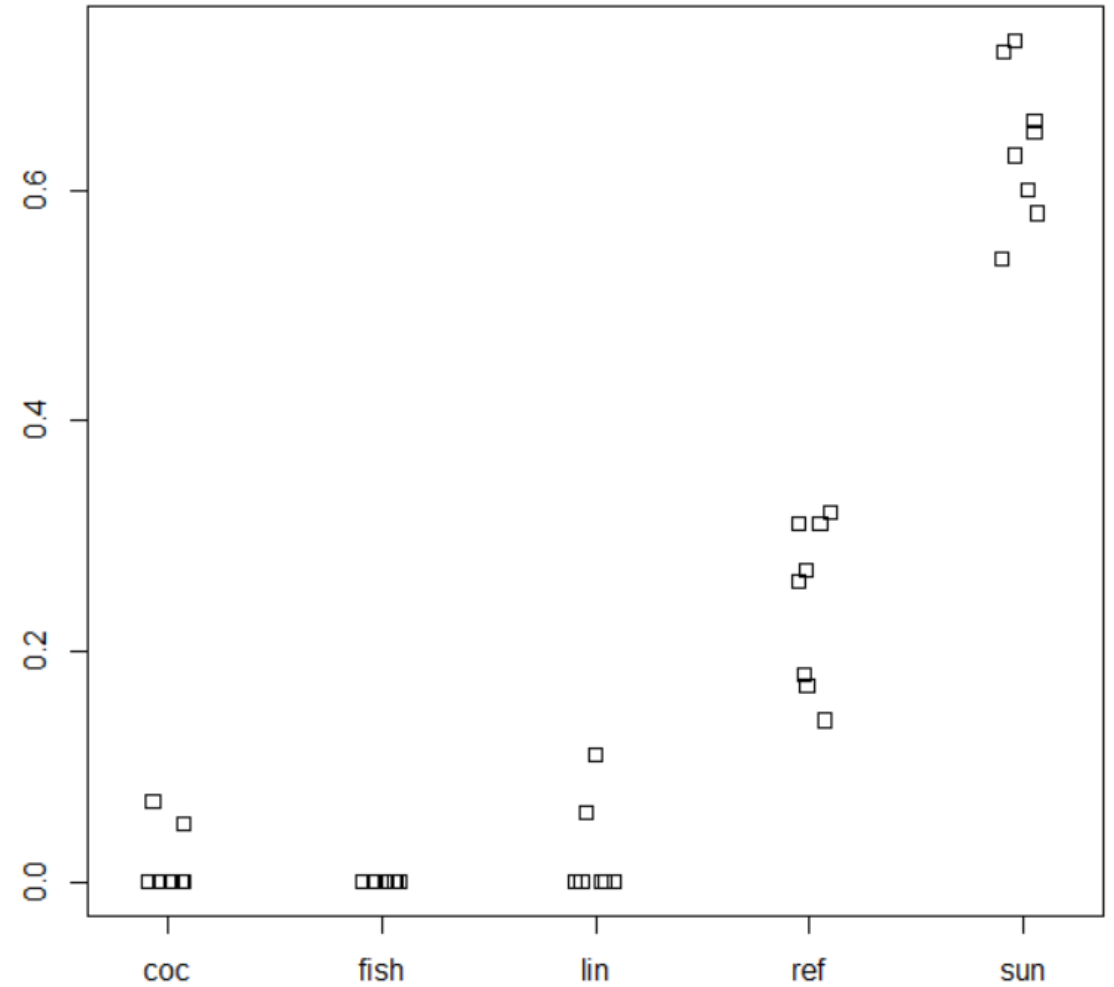
- diet
- genotype

Variable réponse (une)

- C20.4n.6
- C20.5n.3
- C22.4n.6
- C22.5n.3
- C22.5n.6
- C22.6n.3

Aide Réinitialiser OK

C22.4n.6



# **Partie 3.3**

## **Deux variables qualitatives**

## **TO DO**

- **Editer le jeu de données genotype et modifier le nom de la colonne en « genotype »**



Données: genotype



Éditer



Visualiser

Modèle:  $\Sigma$  <Pas de modèle>

Script R R Markdown

```
library(cowplot, pos=15)
with(diet, pie(table(diet), labels=levels(diet), xlab="", ylab="",
  main="diet", col=rainbow_hcl(5)))
with(diet, pie(table(diet), labels=levels(diet), xlab="", ylab="",
  main="diet", col=rainbow_hcl(5)))
alldata <- merge(diet, gene, all=TRUE, by="row.names")
rownames(alldata) <- alldata$row.names
```



Editeur de données: genoty...

Fichier Édition Aide

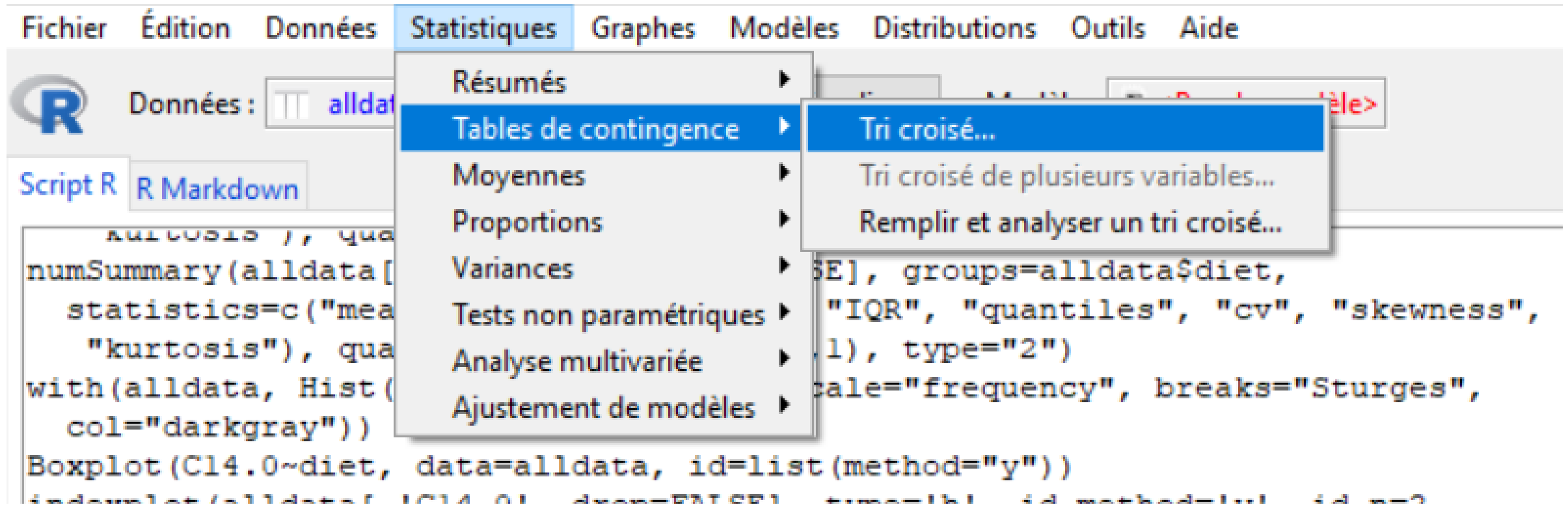
Ajouter une ligne

Ajouter une colonne

	rowname	genotype
1	1	wt
2	2	wt
3	3	wt
4	4	wt
5	5	wt
6	6	wt
7	7	wt
8	8	wt

## Tableau croisé (table de contingence)

R Commander



The screenshot shows the R Commander interface. The menu bar includes 'Fichier', 'Édition', 'Données', 'Statistiques', 'Graphes', 'Modèles', 'Distributions', 'Outils', and 'Aide'. The 'Statistiques' menu is open, showing options like 'Résumés', 'Tables de contingence', 'Moyennes', 'Proportions', 'Variances', 'Tests non paramétriques', 'Analyse multivariée', and 'Ajustement de modèles'. The 'Tables de contingence' option is selected, and a sub-menu is displayed with 'Tri croisé...' highlighted. Other sub-menu options include 'Tri croisé de plusieurs variables...' and 'Remplir et analyser un tri croisé...'. The background shows a script editor with R code for data analysis.

```
numSummary(alldata[, qua
  statistics=c("mea
    "kurtosis"), qua
with(alldata, Hist(
  col="darkgray"))
Boxplot(C14.0~diet, data=alldata, id=list(method="y"))
id=plot(alldata[, C14.0], data=EMSE1, type="h", id=method="y", id=2
```



R Table de contingence à double entrée



Données Statistiques

Variable en ligne (une)

diet  
genotype

Variable en colonne (une)

diet  
genotype

Expression de sélection

<tous les cas valides>



Aide



Réinitialiser



OK



Annuler



Appliquer

R Table de contingence à double entrée



Données Statistiques

Calculer les pourcentages

- Pourcentages des lignes  Pourcentage des colonnes
- Pourcentages du total  Pas de pourcentages

Test d'hypothèses

- Test Chi-deux d'indépendance  Composants de la statistique du Chi-deux
- Imprimer les fréquences attendues  Test exact de Fisher



Aide



Réinitialiser



OK



Annuler



Appliquer

```

Rcmdr> local({
Rcmdr+   .Table <- xtabs(~diet+genotype, data=alldata)
Rcmdr+   cat("\nFrequency table:\n")
Rcmdr+   print(.Table)
Rcmdr+   cat("\nRow percentages:\n")
Rcmdr+   print(rowPercents(.Table))
Rcmdr+ })

```

Frequency table:

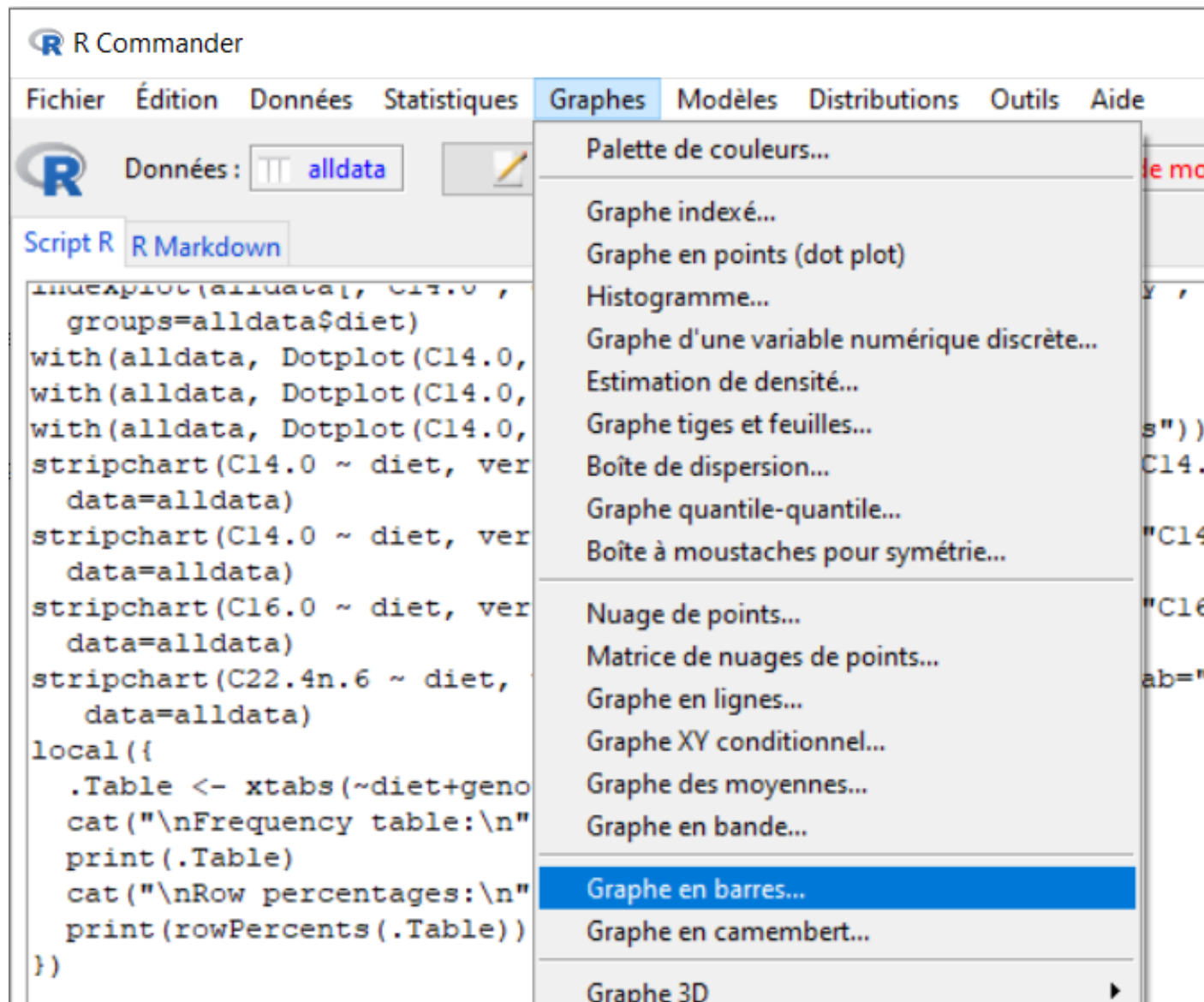
	genotype	
diet	ppar	wt
coc	4	4
fish	4	4
lin	4	4
ref	4	4
sun	4	4

Row percentages:

	genotype		Total	Count
diet	ppar	wt		
coc	50	50	100	8
fish	50	50	100	8
lin	50	50	100	8
ref	50	50	100	8
sun	50	50	100	8

Plan d'expérience équilibré  
Pas de lien entre Diet et Genotype (par construction)


# Graphe en barres (barplot)



The image shows the R Commander interface. The 'Graphes' menu is open, displaying various plotting options. The option 'Graphe en barres...' is highlighted in blue. The background shows a script editor with R code for data analysis and plotting.

**R Commander**

Fichier Édition Données Statistiques **Graphes** Modèles Distributions Outils Aide

Données :  

Script R

```
indexplot(alldata[, C14.0 ,
  groups=alldata$diet)
with(alldata, Dotplot(C14.0,
with(alldata, Dotplot(C14.0,
with(alldata, Dotplot(C14.0,
stripchart(C14.0 ~ diet, ver
  data=alldata)
stripchart(C14.0 ~ diet, ver
  data=alldata)
stripchart(C16.0 ~ diet, ver
  data=alldata)
stripchart(C22.4n.6 ~ diet,
  data=alldata)
local({
  .Table <- xtabs(~diet+geno
  cat("\nFrequency table:\n"
  print(.Table)
  cat("\nRow percentages:\n"
  print(rowPercents(.Table))
})
```

Palette de couleurs...

Graphe indexé...

Graphe en points (dot plot)

Histogramme...

Graphe d'une variable numérique discrète...

Estimation de densité...

Graphe tiges et feuilles...

Boîte de dispersion...

Graphe quantile-quantile...

Boîte à moustaches pour symétrie...

Nuage de points...

Matrice de nuages de points...

Graphe en lignes...

Graphe XY conditionnel...

Graphe des moyennes...

Graphe en bande...

**Graphe en barres...**

Graphe en camembert...

Graphe 3D ▶

## R Graphe en barres

Données Options

Variable (une)

diet  
genotype

Graphe par groupe...

### R Groupes

Variable de regroupement (une)

diet  
genotype



OK



Annuler

## R Graphe en barres

Données Options

Echelle des axes

- Fréquences
- Pourcentages

Sélection des couleurs

- Par défaut
- À partir de la palette de couleurs

Style des barres de groupes

- Divisées (empilement)
- Côte-à-côte (en parallèle)

Pourcentages pour les barres de groupes

- Conditionnel
- Total

Position de la légende

- Au dessus de graphe
- À droite
- Centrer
- A gauche

Étiquettes du graphe

libellé de l'axe X < auto >



libellé de l'axe Y < auto >



Titre du graphe < auto >



