

# Logiciel R

PRÉSENTER PAR:

EL AMINE EL ALAMI NOHA  
EL HADRATI ABIR  
HACINI ALI

ENCADRER PAR:

DR. MOHAMED EL MEROUANI



## Plan:



- Section 1: Introduction
  - Définition
  - Description sommaire de R
  - Interface R
- Section2 : Les opérations sous R
  - Codes des opérations
  - Exercice d'application
- Section3 : La statistique sous R
  - Importer les données
  - Régression linéaire simple
  - Analyser la variance

# Définition:



- R est un logiciel libre de traitement statistiques des données disponible créé par R. Ihaka et R. Gentleman
- Vous pouvez installer R gratuitement d'après le lien:

<http://www.r-project.org/>

- Également disponible sous forme commercialisée S-Plus
- Distribué pour Windows, Linux, Unix, Macintosh

# Description sommaire de R

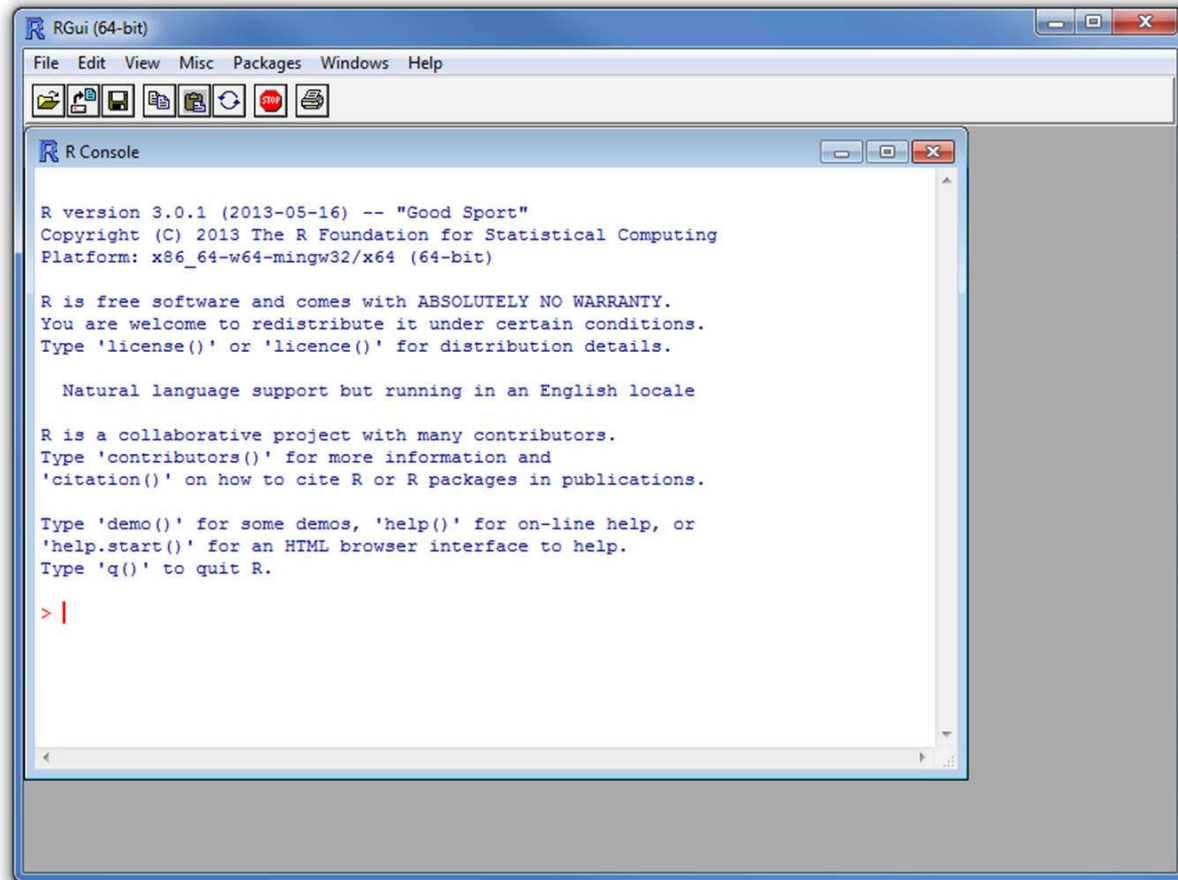


- R est un environnement intégré de manipulation de données, de calcul et de préparation de graphiques. Toutefois, ce n'est pas seulement un « autre » environnement statistique (comme SPSS ou SAS, par exemple), mais aussi un langage de programmation complet et autonome.
- Le R est un langage particulièrement puissant pour les applications mathématiques et statistiques puisque précisément développé dans ce but.



- Parmi ses caractéristiques particulièrement intéressantes, on note :
  - Langage basé sur la notion de vecteur, ce qui simplifie les calculs mathématiques et réduit considérablement le recours aux structures itératives (boucles for, while, etc.) ;
  - Pas de typage ni de déclaration obligatoire des variables ;
  - Programmes courts, en général quelques lignes de code seulement ;
  - Temps de développement très court.

# Interfaces: Fenêtre de démarrage R



The image shows a screenshot of the RGui (64-bit) window. The window title is "RGui (64-bit)" and it has a menu bar with "File", "Edit", "View", "Misc", "Packages", "Windows", and "Help". Below the menu bar is a toolbar with icons for file operations and execution. The main area of the window is the "R Console", which displays the following text:

```
R version 3.0.1 (2013-05-16) -- "Good Sport"
Copyright (C) 2013 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.

Natural language support but running in an English locale

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.

> |
```

Opération	Code
Pi	pi
Racine carré	sqrt
Multiplication,division,addition,soustraction	*, /, +, -
arrondir	round
Moyenne,médiane,variance	mean, median, var
Quartile et moyenne	summary
Ecart-type	sd
Etendue d'une série de valeur	range
Appliquer une même fonction à toutes les catégories de Y d'une variable X	tapply (X,Y,FUN)
Importer des données d'un tableau	read.delim, read.table
Créer une matrice	matrix
Effectuer un modèle lineaire	lm(x~y*z)
Représenter un nuage de points	plot(y~x)
Tracer une droite de regression correspondant à un modèle lineaire	abline(lm)
Fabriquer un graphe en boîte à moustache	boxplot(x~y)
Options graphiques	par

# Exercice d'application:



- **Exercice 1**

1. Calculez sous R la racine carrée de :  $\pi$  multiplié par l'âge de l'enseignant divisé par le nombre d'étudiants dans la salle. Supposons que l'âge de l'enseignant est 23 ans et le nombre d'étudiants égale à 18.
2. Arrondissez la valeur obtenue à 1 décimale.



# Solution:



Calculez sous R la racine carrée de : pi multiplié par l'âge de l'enseignant divisé par le nombre d'étudiants dans la salle.

Si l'enseignant a 23 ans et qu'il y a 18 étudiants dans la salle :

```
sqrt ( (pi*23) /18)
```

```
[1] 2.003561
```



- Arrondissez le résultat:

```
round(sqrt((pi*23)/18),1)
```

```
[1] 2.003561
```

OU

```
calcul=sqrt((pi*23)/18) #on crée ici un objet de  
type vecteur
```

```
round(calcul,1)
```

```
[1] 2.003561
```



## Différents types d'objets dans R :

- **vecteurs (vector)**

```
myvector=1
```

```
myvector
```

```
[1] 1
```

- **Matrices (matrix)**

```
> mymatrix
```

```
      [,1] [,2] [,3] [,4]
[1, ]    1    1    2    2
[2, ]    3    4    5    6
[3, ]    6    6    6    7
```



- **Tableau de données (data.frame):**

```
> mytab
  Echantillon Masse
1           53 0.0273
2          353 0.0586
```

# Logiciel R



- Installer un package sur R:

On lance le logiciel et on clique sur package:



Ensuite sur la première rubrique « Choisir un site miroir de CRAN », et on effectue le choix de l'endroit le plus proche où on se trouve:



R Console

```
R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"  
Copyright (C) 2016 The R Foundation for Statistical Computing  
Platform: i386-w64-mingw32/i386 (32-bit)
```

```
R est un logiciel libre livré sans AUCUNE GARANTIE.  
Vous pouvez le redistribuer sous certaines conditions.  
Tapez 'license()' ou 'licence()' pour plus de détails.
```

```
R est un projet collaboratif avec de nombreux contributeurs.  
Tapez 'contributors()' pour plus d'information et  
'citation()' pour la façon de le citer dans les publications.
```

```
Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide  
en ligne ou 'help.start()' pour obtenir l'aide au format HTML.  
Tapez 'q()' pour quitter R.
```

```
[Sauvegarde de la session précédente restaurée]
```

```
> chooseCRANmirror()
```

HTTPS CRAN mirror

0-Cloud [https]  
Algeria [https]  
Australia (Canberra) [https]  
Australia (Melbourne) [https]  
Australia (Perth) [https]  
Austria [https]  
Belgium (Ghent) [https]  
Brazil (RJ) [https]  
Brazil (SP 1) [https]  
Bulgaria [https]  
Canada (MB) [https]  
Chile 1 [https]  
Chile 2 [https]  
China (Beijing) [https]  
China (Hefei) [https]  
China (Lanzhou) [https]  
Colombia (Cali) [https]  
Czech Republic [https]  
Denmark [https]  
France (Lyon 1) [https]  
France (Lyon 2) [https]  
France (Marseille) [https]  
France (Montpellier) [https]  
France (Paris 2) [https]  
Germany (Falkenstein) [https]  
Germany (Münster) [https]  
Iceland [https]  
India [https]  
Ireland [https]  
Italy (Padua) [https]  
Japan (Tokyo) [https]  
Malaysia [https]

OK

Annuler



R Console

```
R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"  
Copyright (C) 2016 The R Foundation for Statistical Computing  
Platform: i386-w64-mingw32/i386 (32-bit)
```

```
R est un logiciel libre livré sans AUCUNE GARANTIE.  
Vous pouvez le redistribuer sous certaines conditions.  
Tapez 'license()' ou 'licence()' pour plus de détails.
```

```
R est un projet collaboratif avec de nombreux contributeurs.  
Tapez 'contributors()' pour plus d'information et  
'citation()' pour la façon de le citer dans les publications.
```

```
Tapez 'demo()' pour des démonstrations, 'help()' pour l'aide  
en ligne ou 'help.start()' pour obtenir l'aide au format HTML.  
Tapez 'q()' pour quitter R.
```

```
[Sauvegarde de la session précédente restaurée]
```

```
> chooseCRANmirror()  
> utils:::menuInstallPkgs()
```

Packages

- rcircr
- RCircos
- rclimatega
- RClimMAWGEN
- rClinicalCodes
- rclinicaltrials
- RClone
- Rclusterpp
- rCMA
- rcmdcheck
- Rcmdr**
- RcmdrMisc
- RcmdrPlugin.BCA
- RcmdrPlugin.BiclustGUI
- RcmdrPlugin.coin
- RcmdrPlugin.depthTools
- RcmdrPlugin.DoE
- RcmdrPlugin.doex
- RcmdrPlugin.EACSPIR
- RcmdrPlugin.EBM
- RcmdrPlugin.EcoVirtual
- RcmdrPlugin.epack
- RcmdrPlugin.Export
- RcmdrPlugin.EZR
- RcmdrPlugin.FactoMineR
- RcmdrPlugin.FuzzyClust
- RcmdrPlugin.GWRM
- RcmdrPlugin.HH
- RcmdrPlugin.IPSUR
- RcmdrPlugin.KMggplot2
- RcmdrPlugin.lfst
- RcmdrPlugin.MA

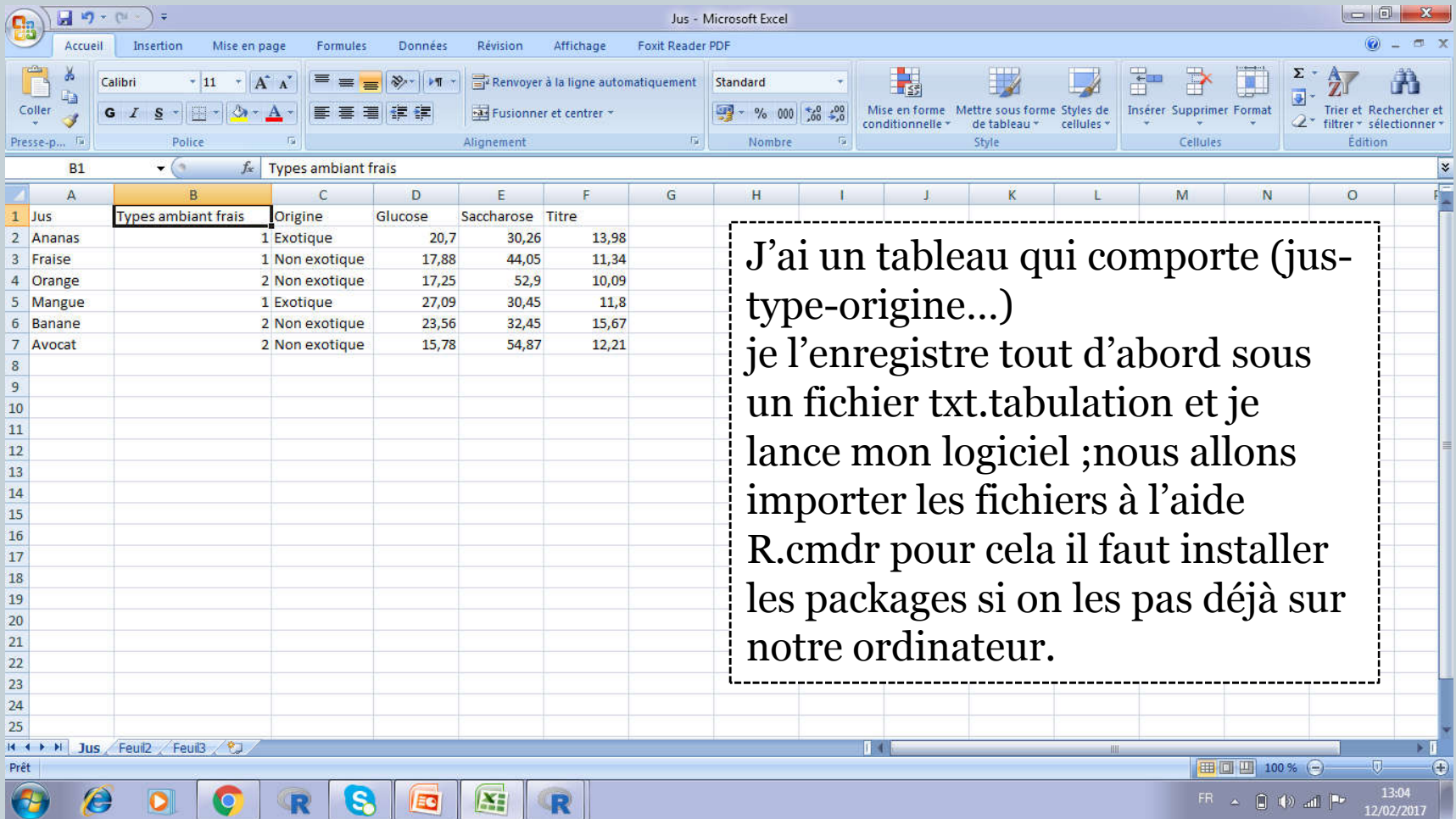
OK

Annuler

Enfin on clique sur installer les packages pour faire le choix ici on a choisi Rcmdr:

# Importer les données sur R

Nous allons voir un petit exemple, je prends alors un fichier Excel



The screenshot shows a Microsoft Excel spreadsheet with the following data:

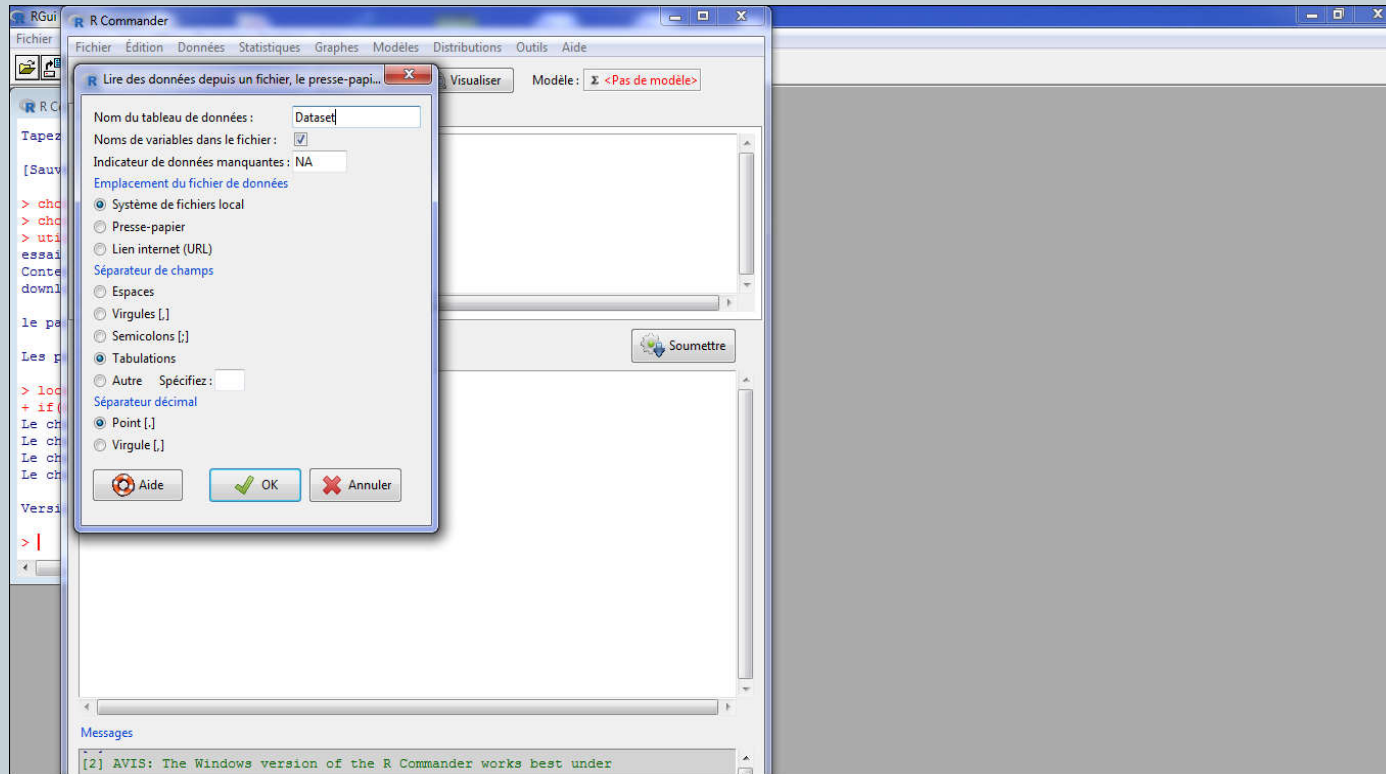
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Jus	Types ambient frais	Origine	Glucose	Saccharose	Titre									
2	Ananas		1 Exotique	20,7	30,26	13,98									
3	Fraise		1 Non exotique	17,88	44,05	11,34									
4	Orange		2 Non exotique	17,25	52,9	10,09									
5	Mangue		1 Exotique	27,09	30,45	11,8									
6	Banane		2 Non exotique	23,56	32,45	15,67									
7	Avocat		2 Non exotique	15,78	54,87	12,21									

A text box on the right contains the following text:

J'ai un tableau qui comporte (juste-type-origine...)  
je l'enregistre tout d'abord sous un fichier txt.tabulation et je lance mon logiciel ; nous allons importer les fichiers à l'aide R.cmdr pour cela il faut installer les packages si on les pas déjà sur notre ordinateur.

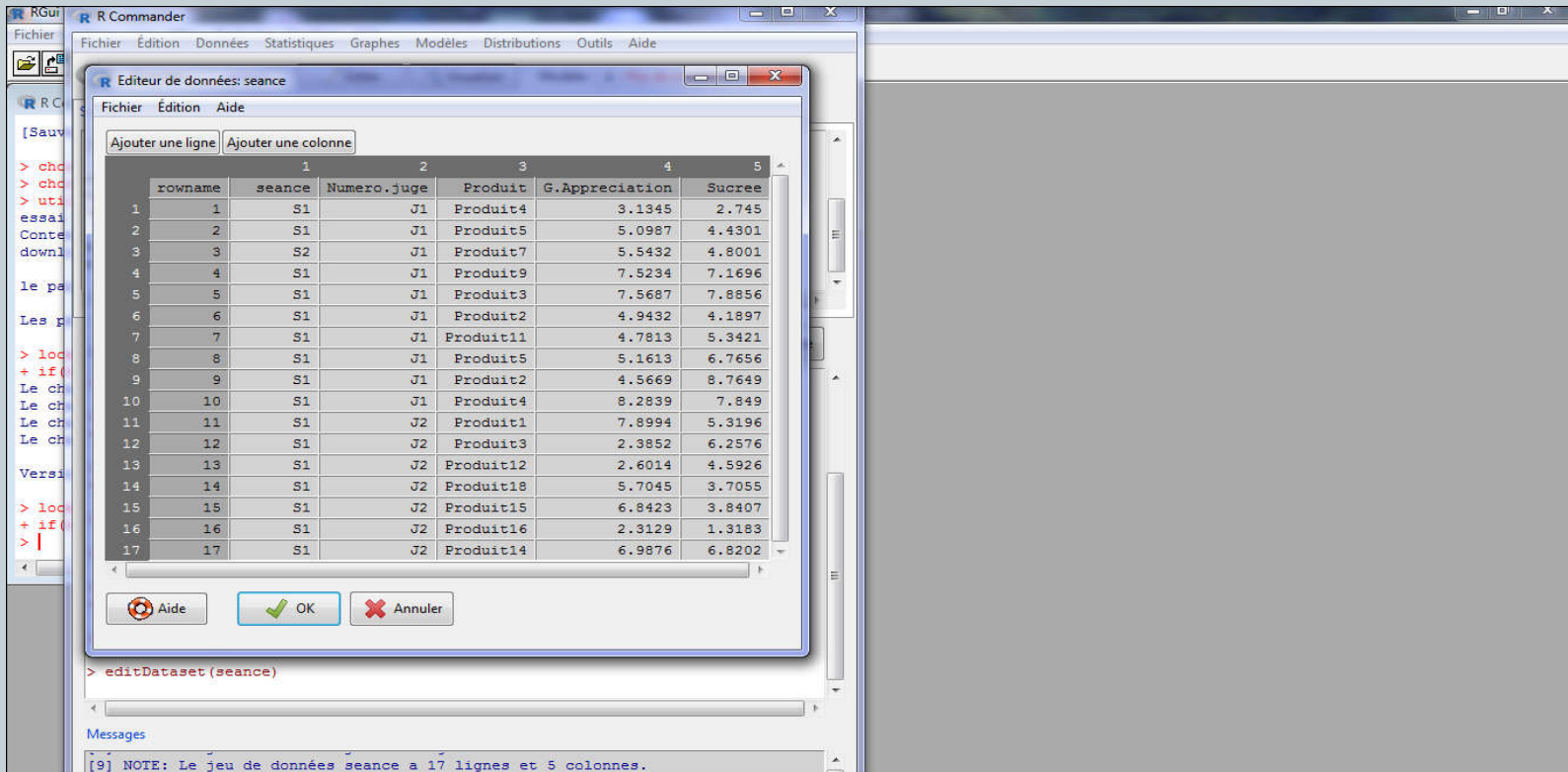


Sinon, on clique sur charger les packages une fois ce dernier est chargé la fenêtre r-cmdr s'ouvre et on clique sur donnée qui nous permet de les importées.



# Régression linéaire sous R

Pour faire une régression linéaire je vais éditer un autre jeu de donnée



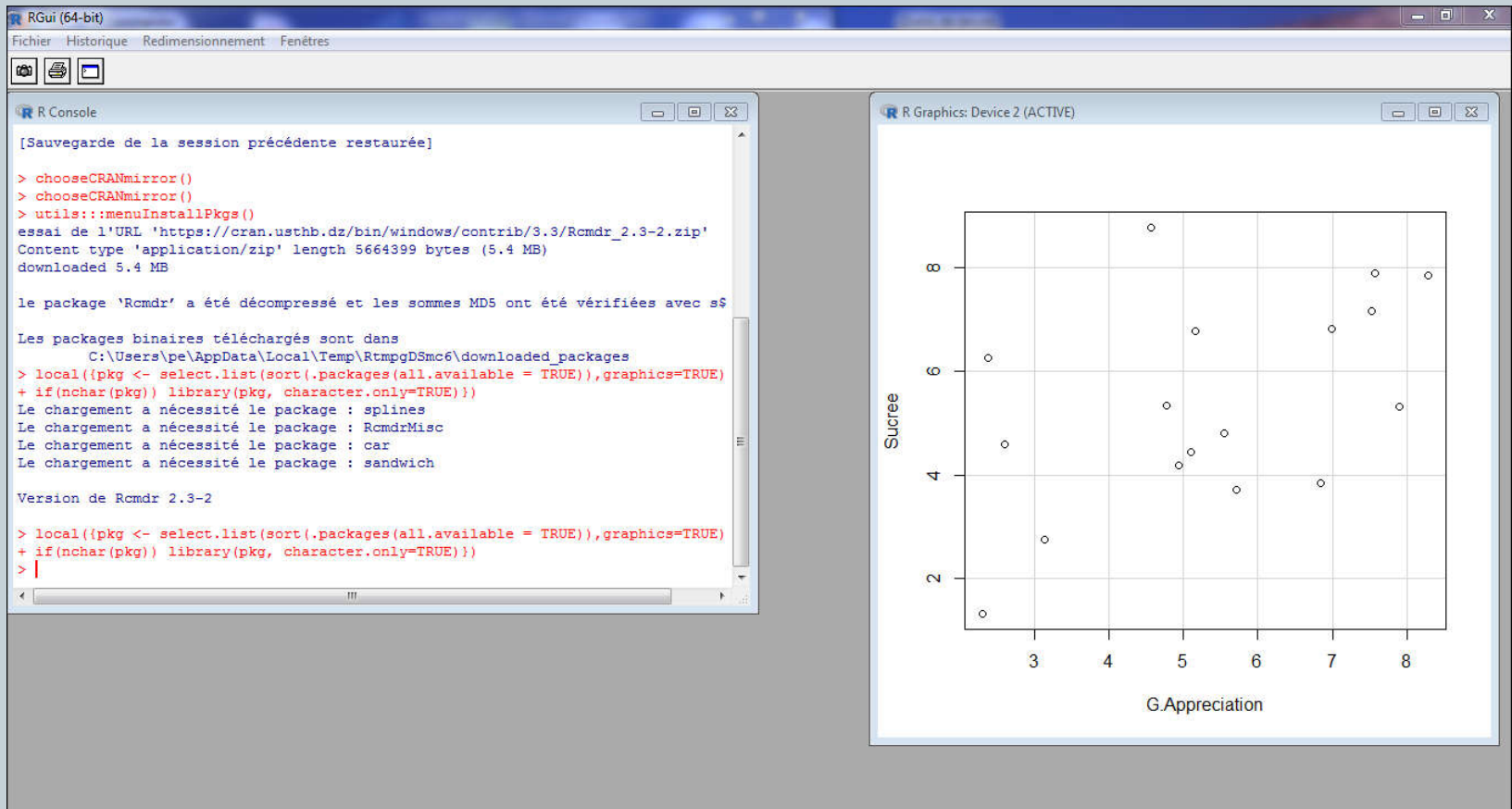
The screenshot shows the R Commander interface. The main window displays a data editor for the dataset 'seance'. The data is organized into 17 rows and 5 columns. The columns are labeled: 'rowname', 'seance', 'Numero.juge', 'Produit', 'G.Appreciation', and 'Sucree'. The rows contain numerical identifiers, session codes (S1, S2), judge names (J1, J2), product names (Produit1-18), and two numerical values representing appreciation and sucrose levels.

rowname	seance	Numero.juge	Produit	G.Appreciation	Sucree
1	S1	J1	Produit4	3.1345	2.745
2	S1	J1	Produit5	5.0987	4.4301
3	S2	J1	Produit7	5.5432	4.8001
4	S1	J1	Produit9	7.5234	7.1696
5	S1	J1	Produit3	7.5687	7.8856
6	S1	J1	Produit2	4.9432	4.1897
7	S1	J1	Produit11	4.7813	5.3421
8	S1	J1	Produit5	5.1613	6.7656
9	S1	J1	Produit2	4.5669	8.7649
10	S1	J1	Produit4	8.2839	7.849
11	S1	J2	Produit1	7.8994	5.3196
12	S1	J2	Produit3	2.3852	6.2576
13	S1	J2	Produit12	2.6014	4.5926
14	S1	J2	Produit18	5.7045	3.7055
15	S1	J2	Produit15	6.8423	3.8407
16	S1	J2	Produit16	2.3129	1.3183
17	S1	J2	Produit14	6.9876	6.8202

The R Commander console shows the command `editDataset(seance)` and a message: `[9] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes.`

- Etudions par exemple le degré d'appréciation en fonction de la variable sucre

- Dans un premier temps il est préférable de décrire les données avec un graphe nuage des points de la variable sucre avec degré d'appréciation puis je clique sur statistique puis ajustement puis régression et on obtient le «  $\hat{a}_1$  ».  $A_1$  estimé.



- je clique ensuite sur statistique puis ajustement et régression et on obtient le «  $\hat{a}$  » A estimé.



The screenshot shows the R Commander window with the following content:

**Script R**

```
editDataset(jus)
editDataset(seance)
editDataset(seance)
scatterplot(Sucree~G.Appreciation, reg.line=FALSE, smooth=FALSE,
  spread=FALSE, boxplots=FALSE, span=0.5, ellipse=FALSE, levels=c(.5, .9),
  data=seance)
RegModel.1 <- lm(Sucree~G.Appreciation, data=seance)
summary(RegModel.1)
```

**Sortie**

```
Call:
lm(formula = Sucree ~ G.Appreciation, data = seance)

Residuals:
    Min       1Q   Median       3Q      Max
-2.5163 -1.3725  0.2449  0.9602  3.7774

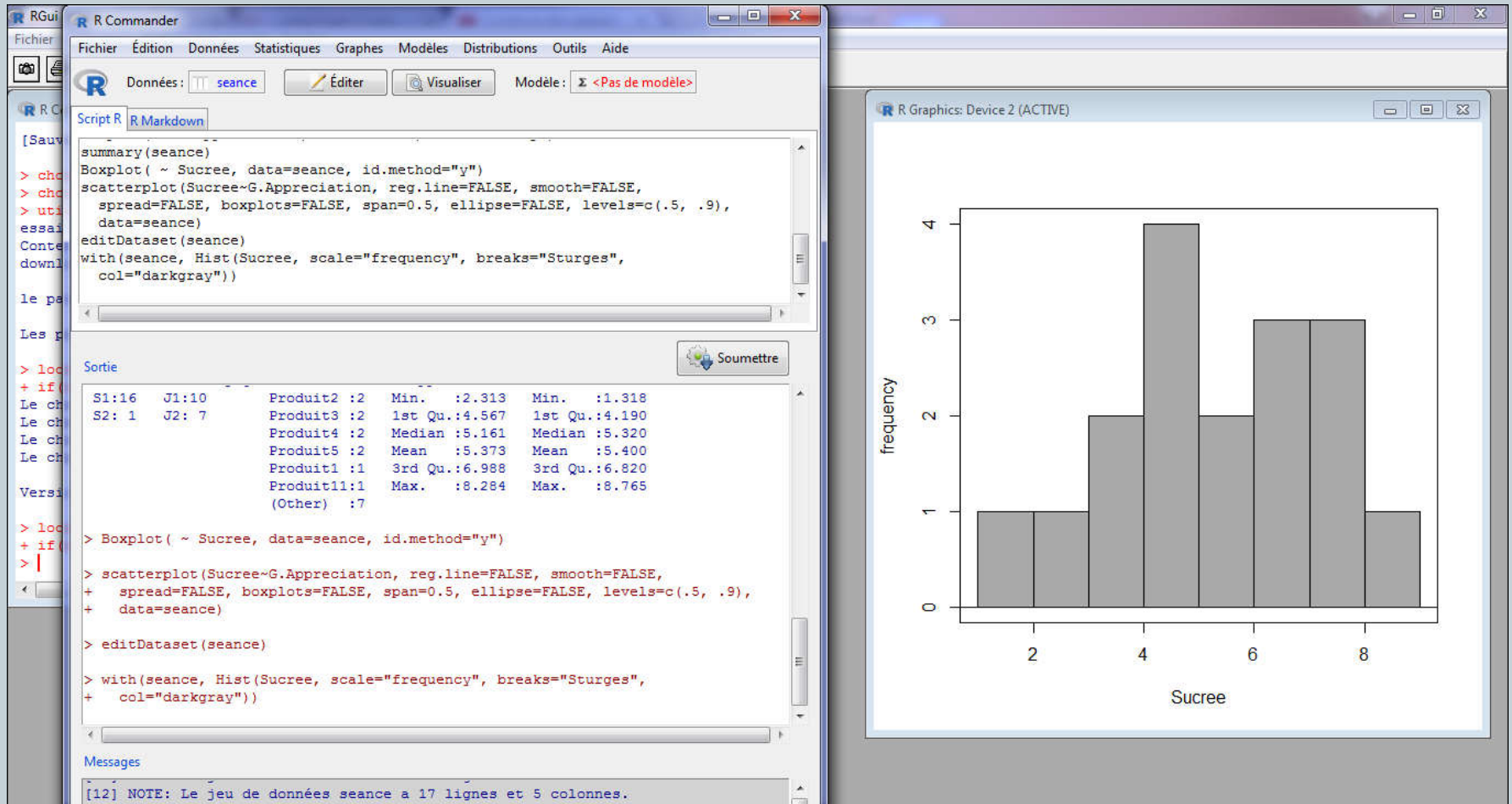
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.6515     1.2947   2.048  0.0585 .
G.Appreciation  0.5115     0.2272   2.252  0.0398 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.781 on 15 degrees of freedom
Multiple R-squared:  0.2526, Adjusted R-squared:  0.2028
F-statistic:  5.07 on 1 and 15 DF, p-value: 0.03975
```

**Messages**

```
[10] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes.
```

# Insérer les graphes sous R



The image displays the R Commander interface, which is used for running R scripts and viewing the results. The main window is titled "R Commander" and contains a menu bar (Fichier, Édition, Données, Statistiques, Graphes, Modèles, Distributions, Outils, Aide) and a toolbar with buttons for "Données", "Éditer", "Visualiser", and "Modèle". The "Données" field shows "seance" and the "Modèle" field shows "<Pas de modèle>".

The "Script R" pane contains the following R code:

```
summary(seance)
Boxplot( ~ Sucree, data=seance, id.method="y")
scatterplot(Sucree~G.Appreciation, reg.line=FALSE, smooth=FALSE,
  spread=FALSE, boxplots=FALSE, span=0.5, ellipse=FALSE, levels=c(.5, .9),
  data=seance)
editDataset(seance)
with(seance, Hist(Sucree, scale="frequency", breaks="Sturges",
  col="darkgray"))
```

The "Sortie" pane shows the output of the script, including a summary of the 'seance' dataset and the execution of the Boxplot and Hist functions. The summary output is as follows:

Variable	Unit	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Produit2	:2	2.313	4.567	5.161	5.373	6.988	8.284
Produit3	:2	1.318	4.190	5.320	5.400	6.820	8.765
Produit4	:2						
Produit5	:2						
Produit1	:1						
Produit11	:1						
(Other)	:7						

The "Messages" pane shows a note: "[12] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes."

The "R Graphics: Device 2 (ACTIVE)" window displays a histogram of the 'Sucree' variable. The x-axis is labeled "Sucree" and has major ticks at 2, 4, 6, and 8. The y-axis is labeled "frequency" and has major ticks at 0, 1, 2, 3, and 4. The histogram consists of seven bars with the following approximate frequencies: 1, 1, 2, 4, 2, 3, and 1.

# Analyse de la variance sous R

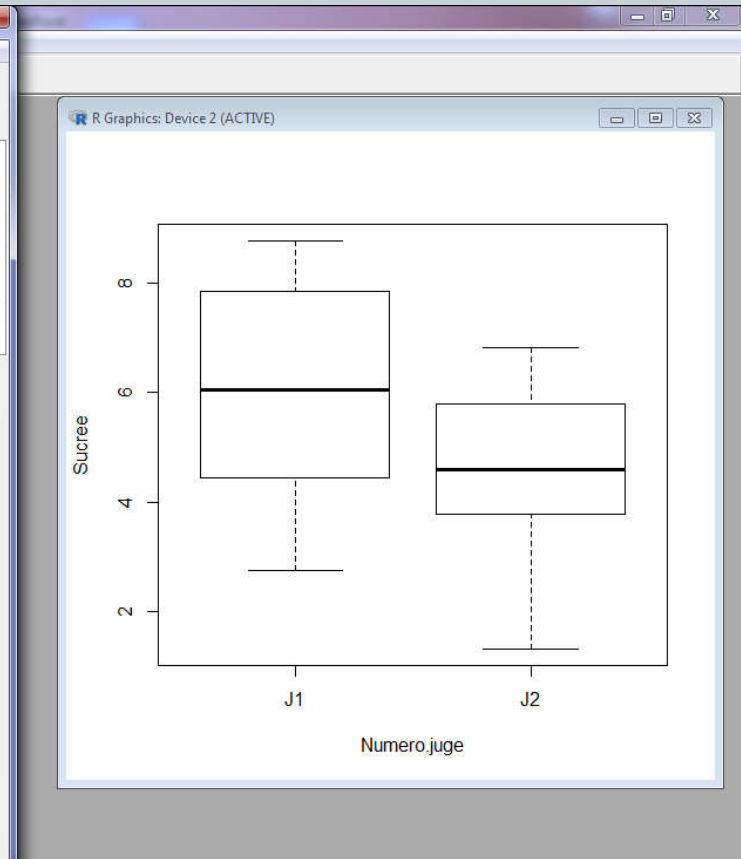
Je vais aussi éditer un jeu de donnée et analyser le degré d'appréciation avec les variables jus et production. je visualise les données en utilisant un graphe (la boîte à moustache)

```
R Commander
Fichier Édition Données Statistiques Graphes Modèles Distributions Outils Aide
Données: seance Éditer Visualiser Modèle: Σ <Pas de modèle>
Script R R Markdown
[Sauv]
> chd
> chd
> utl
essai
Conte
downl
le pa
Les g
> loc
+ if (
Le ch
Le ch
Le ch
Le ch
Versi
> loc
+ if (
> |
Messages
[12] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes.
```

```
scatterplot(Sucree~G.Appreciation, reg.line=FALSE, smooth=FALSE,
  spread=FALSE, boxplots=FALSE, span=0.5, ellipse=FALSE, levels=c(.5, .9),
  data=seance)
editDataset(seance)
with(seance, Hist(Sucree, scale="frequency", breaks="Sturges",
  col="darkgray"))
Boxplot(G.Appreciation~seance, data=seance, id.method="y")
Boxplot(Sucree~Numero.juge, data=seance, id.method="y")
```

```
Sortie
Produit1 :1 3rd Qu.:6.988 3rd Qu.:6.820
Produit11:1 Max. :8.284 Max. :8.765
(Other) :7
```

```
> Boxplot(~ Sucree, data=seance, id.method="y")
> scatterplot(Sucree~G.Appreciation, reg.line=FALSE, smooth=FALSE,
+ spread=FALSE, boxplots=FALSE, span=0.5, ellipse=FALSE, levels=c(.5, .9),
+ data=seance)
> editDataset(seance)
> with(seance, Hist(Sucree, scale="frequency", breaks="Sturges",
+ col="darkgray"))
> Boxplot(G.Appreciation~seance, data=seance, id.method="y")
> Boxplot(Sucree~Numero.juge, data=seance, id.method="y")
```



- je clique sur statistique puis je choisis ajustement du modèle ensuite modèle linéaire et je clique sur appréciation en fonction de produit et de juge.

The screenshot shows the R Commander interface. The 'Modèle linéaire' dialog box is open, allowing the user to define a linear model. The model name is 'LinearModel.2'. The response variable is 'G.Appreciation'. The predictor variables are 'Numero.juge [facteur]' and 'Produit [facteur]'. The 'Formule du modèle' section shows the formula 'G.Appreciation ~ Numero.juge + Produit'. The 'Expression de sélection' is set to 'Hauteur'. The console window below shows the following R commands:

```
+ data=seance)
> editDataset(seance)
> with(seance, Hist(Sucree, scale="frequency", breaks="Sturges",
+ col="darkgray"))
> Boxplot(G.Appreciation~seance, data=seance, id.method="y")
> Boxplot(Sucree~Numero.juge, data=seance, id.method="y")
```

The console also displays a message: [12] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes.

On trouve les coefficients fournis et je clique sur modèle, test hypothese et finalement table de l ANOVA et on trouve la table d'analyse de ola variance avec les sommes des carrées.



```
editDataset(seance)
with(seance, Hist(Sucree, scale="frequency", breaks="Sturges",
  col="darkgray"))
Boxplot(G.Appreciation~seance, data=seance, id.method="y")
Boxplot(Sucree~Numero.juge, data=seance, id.method="y")
LinearModel.2 <- lm(G.Appreciation ~ Produit + Numero.juge, data=seance)
summary(LinearModel.2)
anova(LinearModel.2)
```

Sortie

```
Produit[T.Produit9]    -5.5595    4.2160   -1.319    0.2789
Numero.juge[T.J2]     -5.1835    2.9812   -1.739    0.1805
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.108 on 3 degrees of freedom
Multiple R-squared:  0.7832, Adjusted R-squared:  -0.1562
F-statistic: 0.8338 on 13 and 3 DF,  p-value: 0.6513

> anova(LinearModel.2)
Analysis of Variance Table

Response: G.Appreciation
      Df Sum Sq Mean Sq F value Pr(>F)
Produit    12  34.729   2.8941  0.6513 0.7440
Numero.juge  1  13.434  13.4343  3.0233 0.1805
Residuals    3  13.331   4.4436

[13] NOTE: Le jeu de données seance a 17 lignes et 5 colonnes.
```



**MERCI POUR VOTRE  
ATTENTION**

