

R - the simple way

Manchester R

Graphical User Interfaces for R

Graeme Hutcheson

University of Manchester

I teach statistics to social scientists who need to use a full range of techniques.

I teach statistics to social scientists who need to use a full range of techniques.

The package I use is R...

I teach statistics to social scientists who need to use a full range of techniques.

The package I use is R...

It includes all the techniques I need, is open-source, available for all platforms, is easy to install and use, and runs from a CD/USB drive...

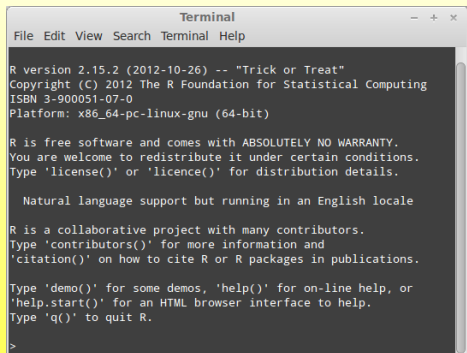
R is easy to use...

R is easy to use...

First, start R...

R is easy to use...

First, start R...



```
Terminal
File Edit View Search Terminal Help

R version 2.15.2 (2012-10-26) -- "Trick or Treat"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: x86_64-pc-linux-gnu (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

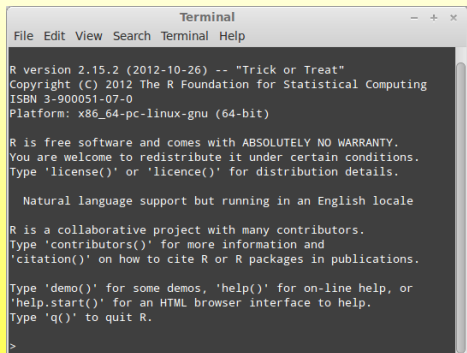
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

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.

>
```

R is easy to use...

First, start R...



```
Terminal
File Edit View Search Terminal Help

R version 2.15.2 (2012-10-26) -- "Trick or Treat"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: x86_64-pc-linux-gnu (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.

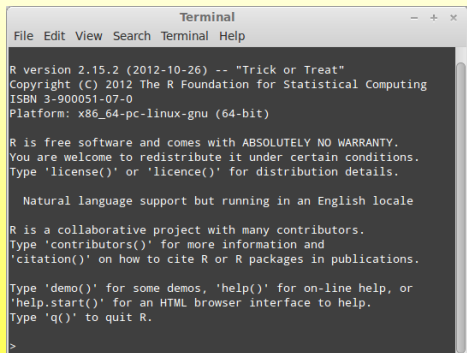
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Then, import
your data using
a command such as...

R is easy to use...

First, start R...



```
Terminal
File Edit View Search Terminal Help

R version 2.15.2 (2012-10-26) -- "Trick or Treat"
Copyright (C) 2012 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: x86_64-pc-linux-gnu (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.

>
```

Then, import
your data using
a command such as...

```
Dataset <- read.table("/Documents/Data/ExampleData.csv",  
                      header=TRUE, sep=";", na.strings="NA",  
                      dec=".", strip.white=TRUE)
```

R is easy to use...

It is easy to run regression models...

R is easy to use...

It is easy to run regression models...

```
GLM.2 <- glm(wage ~ Company + Age,  
             family=gaussian(identity), data=Dataset)  
summary(GLM.2)
```

It is easy to run regression models...

```
GLM.2 <- glm(wage ~ Company + Age,  
             family=gaussian(identity), data=Dataset)  
summary(GLM.2)
```

```
glm(formula = wage ~ Company + Age,  
     family = gaussian(identity), data = Dataset)
```

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	10.8925	10.5343	1.034
Company[T.company02]	0.4746	7.5364	0.063
Company[T.company03]	6.8441	8.7338	0.784
Age	0.7210	0.2240	3.219

R is easy to use...

It is easy to run regression models...

```
GLM.2 <- glm(wage ~ Company + Age,  
             family=gaussian(identity), data=Dataset)  
summary(GLM.2)
```

```
glm(formula = wage ~ Company + Age,  
     family = gaussian(identity), data = Dataset)
```

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	10.8925	10.5343	1.034
Company[T.company02]	0.4746	7.5364	0.063
Company[T.company03]	6.8441	8.7338	0.784
Age	0.7210	0.2240	3.219

it is easy to draw graphics...

R is easy to use...

It is easy to run regression models...

```
GLM.2 <- glm(wage ~ Company + Age,  
             family=gaussian(identity), data=Dataset)  
summary(GLM.2)
```

```
glm(formula = wage ~ Company + Age,  
     family = gaussian(identity), data = Dataset)
```

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	10.8925	10.5343	1.034
Company[T.company02]	0.4746	7.5364	0.063
Company[T.company03]	6.8441	8.7338	0.784
Age	0.7210	0.2240	3.219

it is easy to draw graphics...

```
scatterplot(wage~Age, reg.line=FALSE,  
            smooth=TRUE, spread=FALSE, xlab="Age",  
            ylab="Wage", span=0.5, data=Dataset)
```

R is easy to use...

It is easy to run regression models...

```
GLM.2 <- glm(wage ~ Company + Age,  
             family=gaussian(identity), data=Dataset)  
summary(GLM.2)
```

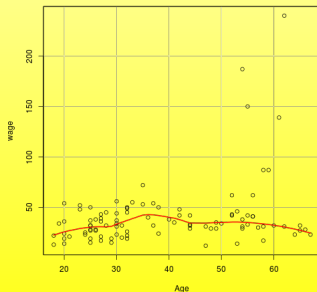
```
glm(formula = wage ~ Company + Age,  
     family = gaussian(identity), data = Dataset)
```

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	10.8925	10.5343	1.034
Company[T.company02]	0.4746	7.5364	0.063
Company[T.company03]	6.8441	8.7338	0.784
Age	0.7210	0.2240	3.219

it is easy to draw graphics...

```
scatterplot(wage~Age, reg.line=FALSE,  
           smooth=TRUE, spread=FALSE, xlab="Age",  
           ylab="Wage", span=0.5, data=Dataset)
```



I think that R is simple and I like using it...

I think that R is simple and I like using it...

My students, on the other hand, who are typically used to point-and-click software are horrified by R...

I think that R is simple and I like using it...

My students, on the other hand, who are typically used to point-and-click software are horrified by R...

...they don't like the way it looks, they don't know the syntax or having to search for commands and they don't like analyses failing due to 'trivial' spelling mistakes.

I think that R is simple and I like using it...

My students, on the other hand, who are typically used to point-and-click software are horrified by R...

...they don't like the way it looks, they don't know the syntax or having to search for commands and they don't like analyses failing due to 'trivial' spelling mistakes.

In practice, many take one look at R, resort to and stay with SPSS.

I think that R is simple and I like using it...

My students, on the other hand, who are typically used to point-and-click software are horrified by R...

...they don't like the way it looks, they don't know the syntax or having to search for commands and they don't like analyses failing due to 'trivial' spelling mistakes.

In practice, many take one look at R, resort to and stay with SPSS.

Despite trying, I have not been able to solve this over the last 10 years, but have managed to teach with R by using a number of graphical interfaces.

Using R - the simple way

There are many GUI interfaces that simplify R; for example, JGR, R-commander, RKward, Rattle...

Using R - the simple way

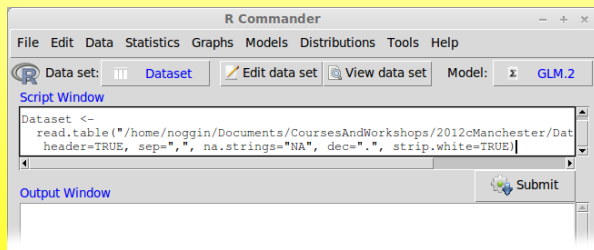
There are many GUI interfaces that simplify R; for example, JGR, R-commander, RKward, Rattle...

I teach using the R-commander (Rcmdr), a GUI developed by John Fox.

Using R - the simple way

There are many GUI interfaces that simplify R; for example, JGR, R-commander, RKward, Rattle...

I teach using the R-commander (Rcmdr), a GUI developed by John Fox.



The Rcmdr offers a range of menu-driven options for data management, statistical analysis, graphics and model manipulation.

The Rcmdr offers a range of menu-driven options for data management, statistical analysis, graphics and model manipulation.

It is particularly valuable as it integrates with editors (eg., Rstudio, Rgedit, Tinn-R, Komodo...) and teaches users about coding...

The Rcmdr offers a range of menu-driven options for data management, statistical analysis, graphics and model manipulation.

It is particularly valuable as it integrates with editors (eg., Rstudio, Rgedit, Tinn-R, Komodo...) and teaches users about coding...

I use R and the Rcmdr interchangeably with an editor (Rstudio)...

Rcmdr Plugins...

The Rcmdr only provides a subset of available techniques...

Rcmdr Plugins...

The Rcmdr only provides a subset of available techniques...

It is, however, simple to extend this using Rcmdr extensions; a list of which are available on CRAN.

Rcmdr Plugins...

The Rcmdr only provides a subset of available techniques...

It is, however, simple to extend this using Rcmdr extensions; a list of which are available on CRAN.

Rcmdr	R Commander
RcmdrPlugin.BCA	Rcmdr Plug-In for Business and Customer Analytics
RcmdrPlugin.coin	Rcmdr Coin Plug-In
RcmdrPlugin.doBy	Rcmdr doBy Plug-In
RcmdrPlugin.DoE	R Commander Plugin for (industrial) Design of Experiments
RcmdrPlugin.doex	Rcmdr plugin for Stat 4309 course
RcmdrPlugin.EACSPIR	Plugin de R-Commander para el manual EACSPIR
RcmdrPlugin.EBM	Rcmdr Evidence Based Medicine Plug-In package
RcmdrPlugin.epack	Rcmdr plugin for time series
RcmdrPlugin.EZR	R Commander Plug-in for the EZR (Easy R) Package
RcmdrPlugin.HH	Rcmdr support for the HH package
RcmdrPlugin.IPSUR	An IPSUR Plugin for the R Commander
RcmdrPlugin.KMggplot2	An Rcmdr Plug-In for Kaplan-Meier Plots and Other Plots by Using the ggplot2 Package
RcmdrPlugin.mosaic	Adds menu items to produce mosaic plots and assoc plots to Rcmdr
RcmdrPlugin.MPAStats	R Commander Plug-in for MPA Statistics
RcmdrPlugin.orloca	orloca Rcmdr Plug-in
RcmdrPlugin.plotByGroup	Rcmdr plots by group using lattice
RcmdrPlugin.qcc	Rcmdr qcc Plug-In
RcmdrPlugin.qual	Rcmdr plugin for quality control course

Rcmdr Plugins...

The plugins provide additional functionality and menu options for the Rcmdr...

Rcmdr Plugins...

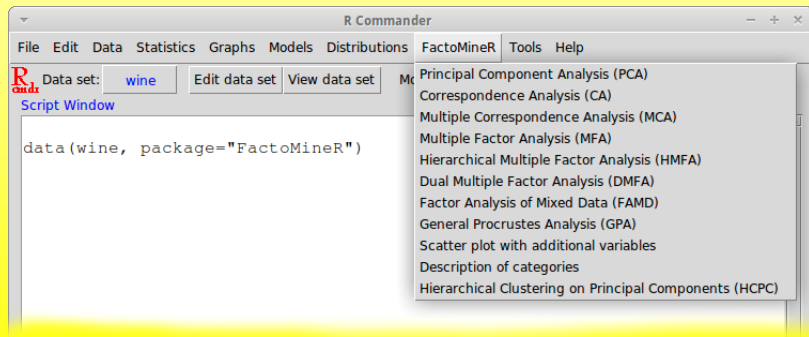
The plugins provide additional functionality and menu options for the Rcmdr...

For example, the FactoMineR plugin provides easy access to the FactoMineR package...

Rcmdr Plugins...

The plugins provide additional functionality and menu options for the Rcmdr...

For example, the FactoMineR plugin provides easy access to the FactoMineR package...



Simple extensions to R

There are nearly 4,500 libraries for R.

Simple extensions to R

There are nearly 4,500 libraries for R.

Many of these libraries are command-line, but there are an increasing number that provide simple graphical interfaces.

Simple extensions to R

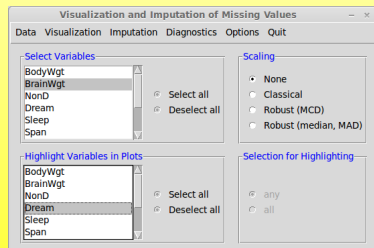
There are nearly 4,500 libraries for R.

Many of these libraries are command-line, but there are an increasing number that provide simple graphical interfaces.

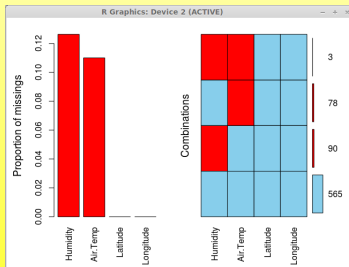
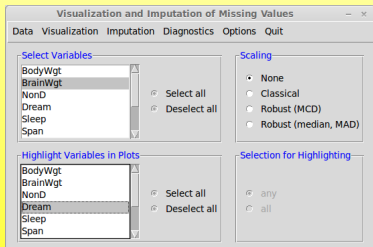
As an example, there are many libraries dedicated to missing data analysis and imputation. When teaching about this, it is useful to have a simple interface. Two that I have found useful are [VIM](#) and [Amelia](#), both of which have easy to use interfaces...

Install VIM. Call using... `library(VIM), vmGUImenu()`

Install VIM. Call using... `library(VIM), vmGUIMenu()`

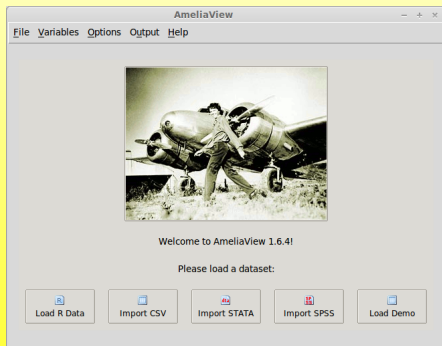


Install VIM. Call using... `library(VIM), vmGUImenu()`



Install Amelia. Call using... `library(Amelia), AmeliaView()`


Install Amelia. Call using... `library(Amelia), AmeliaView()`



Install Amelia. Call using... `library(Amelia), AmeliaView()`

AmeliaView

File Variables Options Output Help



AmeliaView

File Variables Options Output Help

Load Session Save Session Plot Histogram Edit Data Missingness Map Impute! Output Log

Variable	Transformation	Lag	Lea	Bounds	Min	Max	Mean	SD	Missing
Company	ID			(factor)	0/98
Gender	ID			(factor)	0/98
Education	ID			(factor)	0/98
Attitude	ID			(factor)	0/98
wage					12	240	40.92	32.94	0/98
Age					18	67	39.02	14.35	0/98

Time-Series Variable Cross-Section Variable Unhandled Factor Variable

Data Loaded: /home/noggin/Documents/CoursesAndWork Obs: 98 Vars: 6 No im

Install Amelia. Call using... `library(Amelia), AmeliaView()`

The screenshot displays the AmeliaView application window. The top window shows a menu bar (File, Variables, Options, Output, Help) and a central image of Amelia Earhart. Below it, a second window shows a data table with the following structure:

Variable	Transformation	Lag	Lea	Bounds	Min	Max	Mean	St
Company	ID				(factor)	
Gender	ID				(factor)	
Education	ID				(factor)	
Attitude	ID				(factor)	
wage					12	240	40.92	35
Age					18	67	39.02	14

Below the table, there are icons for Time-Series Variable, Cross-Section Variable, and Unhandled. At the bottom, it shows "Data Loaded: /home/noggin/Documents/CoursesAndWork Obs: 98 Vars: 6 No im".

To the right, an R Graphics window titled "R Graphics: Device 3 (ACTIVE)" displays a "Missingness Map". The map shows a grid of red and white cells. The y-axis represents observations (6 to 171) and the x-axis represents variables (tariff, fivecp, intresmi, signed, polly, ushng, gdp.pc, pop, country, year). A legend indicates that white cells represent "Missing" data and red cells represent "Observed" data.

Graphics - the simple way

R has a huge number of graphics that can be drawn.

Graphics - the simple way

R has a huge number of graphics that can be drawn.

the Rcmdr allows a number of these to be drawn using a simple interface (including graphical plugins such as mosaic, .HH, plotbygroup, Kaplan-Meier).

Graphics - the simple way

R has a huge number of graphics that can be drawn.

the Rcmdr allows a number of these to be drawn using a simple interface (including graphical plugins such as mosaic, .HH, plotbygroup, Kaplan-Meier).

There are also some stand-alone interfaces that provide access to graphical capabilities.

Graphics - the simple way

R has a huge number of graphics that can be drawn.

the Rcmdr allows a number of these to be drawn using a simple interface (including graphical plugins such as mosaic, .HH, plotbygroup, Kaplan-Meier).

There are also some stand-alone interfaces that provide access to graphical capabilities.

One of these is [latticist](#)...

Easy graphics using Latticist

Latticist: Dataset

Latticist v. 0.9-44 reset

Hypervariate
marginals splom (pairs) parallel

Variables on axes
switch x/y Aspect:
x= Disposition
y= Course
z=
 Segments (x - z) (x +/- z)

Groups / Color
LPN
explode superpose Tile

Conditioning
Ethnicity

Scales: Sep. strata

Subset

Style and Theme
 Lines Style...

Plot call

```
stripplot(Course ~ Disposition | Ethnicity,  
data = dat, groups = LPN, main = "Course  
vs Disposition by Ethnicity and LPN", xlab  
= "Disposition", jitter.data = TRUE, type =  
c("g", "p", "a"), fun = function(x) {
```

Easy graphics using Latticist

Latticist: Dataset

Latticist v. 0.9-44 reset

Hypervariate

marginals splom (pairs) parallel

Variables on axes

switch x/y Aspect: []

x= Disposition []

y= Course []

z= []

Segments (x - z) [] (x +/- z)

Groups / Color

LPN []

explode superpose Tile

Conditioning

Ethnicity []

Scales: [] Sep. strata

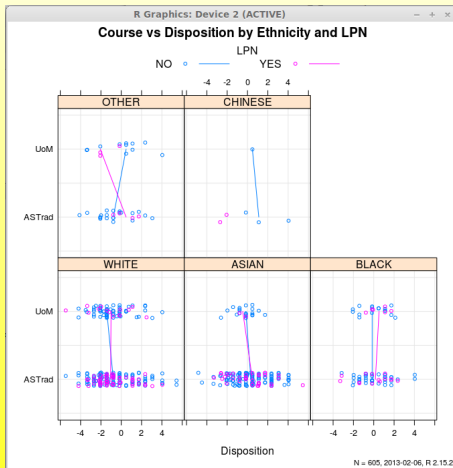
Subset

Style and Theme

Lines Style... []

Plot call

```
stripplot(Course ~ Disposition | Ethnicity,
data = dat, groups = LPN, main = "Course
vs Disposition by Ethnicity and LPN", xlab
= "Disposition", jitter.data = TRUE, type =
c("g", "p", "a"), fun = function(x) {
```



Saving graphics...

Graphics can be saved using a variety of formats (pdf, png, eps, jpg).

Saving graphics...

Graphics can be saved using a variety of formats (pdf, png, eps, jpg).

The Rcmdr allows graphics to be saved and adjustments made for width, height, resolution and text size.

Saving graphics...

Graphics can be saved using a variety of formats (pdf, png, eps, jpg).

The Rcmdr allows graphics to be saved and adjustments made for width, height, resolution and text size.

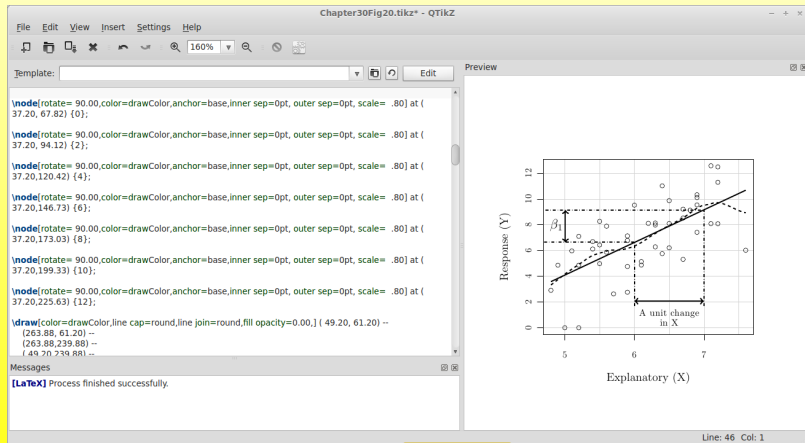
Graphics can also be saved in other formats. The one I use most frequently is the tikZ language (a \LaTeX package).

Easy editing using tikzDevice

Save graphic using tikzDevice library. Then open and edit in Qtikz...

Easy editing using tikzDevice

Save graphic using tikzDevice library. Then open and edit in Qtikz...

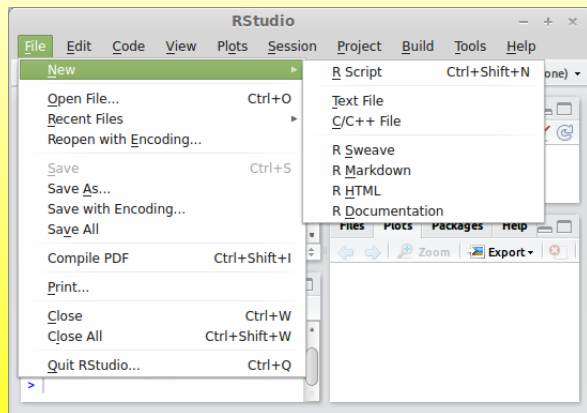


The screenshot shows the Qtikz editor interface. The left pane displays the TikZ source code for a scatter plot with a regression line. The right pane shows the rendered preview of the plot.

```
Chapter30Fig20.tikz* - QTikZ
File Edit View Insert Settings Help
160%
Template:
Edit
Preview
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20, 67.82) {0};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20, 94.12) {2};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20,120.42) {4};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20,146.73) {6};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20,173.03) {8};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20,199.33) {10};
\draw[rotate=90.00,color=drawColor,anchor=base,inner sep=0pt, outer sep=0pt, scale= .80] at (37.20,225.63) {12};
\draw[color=drawColor,line cap=round,line join=round,fill opacity=0.00,] ( 49.20, 61.20) -- (263.88, 61.20) -- (263.88,239.88) -- ( 49.20,239.88) --
Messages
[LaTeX] Process finished successfully.
Line: 46 Col: 1
```

The rendered plot shows a scatter plot of data points with a solid regression line. The x-axis is labeled "Explanatory (X)" and the y-axis is labeled "Response (Y)". A vertical dashed line is drawn at X=6, and a horizontal dashed line is drawn at Y=8. A vertical double-headed arrow between these lines is labeled β_1 . A horizontal double-headed arrow at the bottom of the plot is labeled "A unit change in X".

Easy publishing using R-studio, Sweave, Knitr and \LaTeX



Many of the capabilities of R can be accessed through simple user interfaces.

Many of the capabilities of R can be accessed through simple user interfaces.

R is a truly general package that should be taught to all analysts, whether they are used to graphical interfaces or not...