

Business Intelligence

- SEMINAR WINTER SEMESTER 2013/2014 -

Introduction to the Statistics GUI

"R Commander"

– SEMINAR PAPER –

Submitted by:Maximilian RappStudent-ID:Advisor:Prof. Dr. Dirk Neumann

Table of contents

| 1. | Int | roduction |
|-----|-----|--------------------------------------|
| 2. | Ob | jectives |
| 3. | Sta | rting "R Commander" 4 |
| | 3.1 | Installing the package "R Commander" |
| | 3.2 | Loading the package "R Commander" |
| 4. | "R | Commander" window 6 |
| Z | 4.1 | Toolbar |
| ۷ | 4.2 | Script window |
| Z | 4.3 | Output window7 |
| Z | 1.4 | Message window7 |
| Z | 4.5 | Menu bar |
| 5. | Cre | eating a working directory |
| 6. | Cre | eating data sets |
| e | 5.1 | Windows |
| e | 5.2 | OS X |
| 7. | Im | porting data files |
| 7 | 7.1 | Created data set |
| 7 | 7.2 | CRAN data sets |
| 8. | Sav | ving data file |
| 9. | Sta | tistics |
| ç | 9.1 | Numerical summaries |
| ç | 9.2 | Correlation matrix |
| 10. | G | raphs |
| 1 | 0.1 | Histogram12 |
| 1 | 0.2 | Scatterplot |
| 1 | 0.3 | Scatterplot matrix |
| 11 | N | Iultiple linear regression |
| 11. | 10. | |

| 11 | 1.2 Durbin-Watson test | 5 |
|-----|--|---|
| 11 | 1.3 Variance Inflation | 6 |
| 12. | Getting help | 6 |
| 13. | Closing "R Commander" | 7 |
| 14. | The Comprehensive R Archive Network (CRAN) 1 | 7 |
| 14 | 1.1 CRAN packages | 7 |
| | 14.1.1 Contributed "R Commander" plugins 1 | 7 |
| 14 | 1.2 CRAN task views | 9 |
| 15. | Tutorials & additional reading material | 9 |
| 16. | Conclusion | 0 |
| 17. | References | 1 |

1. Introduction

The Package "R Commander", a menu driven free statistical software for "R" developed by Prof. John Fox, allows to analyse and graph data by running "R" from a graphical user interface (GUI). Instead of typing commands, it provides a selection of the most commonly used commands, which can be easily executed by drop down buttons from the menu bar. The purpose of this seminar paper is to help one to work with "R Commander" correctly. Among other things, one will learn to get started with basic operations. Special focus is placed on how to perform multiple linear regressions and how to check requirements for the OLS estimator. This guidance is based on OS X but there are only small differences in application for Windows and Linux. Commands will be indicated in red courier font.

2. Objectives

The aim of "R Commander" was to develop a friendlier user interface to facilitate students to use statistical software by removing the hindrance of knowing and especially learning commands in order to indicate how statistical software can be used meaningfully. Moreover, it should make it relatively difficult to do unreasonable things and therefore get useless results. It is a very good alternative for users who are not familiar with statistical software and with commands (Eitel 2012). In addition, "R Commander" displays the relationship between choices made in the GUI and "R" for each operation via the corresponding underlying code, which is highly useful for novices, eventually hoping to wean users from the GUI to writing commands (Scott 2012).

3. Starting "R Commander"

Users can already get started after successfully installed "R Commander" within a few minutes.

3.1 Installing the package "R Commander"

The package "R Commander" can be downloaded in "R" by typing at the prompt (>) the following command *winstall.packages(Rcmdr,dependencies=TRUE)* «. A CRAN mirrow window will appear in which one selects a mirror (e.g. *»Germany(Goettingen)*«). "R" installs "R Commander" by pressing *»OK*«.

Alternatively, it is also possible to download and install the package "R Commander" as follows: Click in the menu bar »Packages« \Rightarrow »Install package(s)«. As in Figure 1 one searches in the left column the package "R Commander" labelled "Rcmdr". After pressing the button »Install« "R Commander" is installed.

| Paket-Repository | | |
|---|----------------------------------|-------------------------|
| CRAN (binaries) | * | |
| Liste holen 🗹 Pa | kete im Binärformat | • Suche |
| Paket | Installierte Version | Version im Repository |
| rcdklibs | | 1.5.4 |
| RCEIM | | 0.1 |
| Rcell | | 1.2-6 |
| Rcgmin | | 2013-02.20 |
| Rchemcpp | | 1.1.1 |
| rChoiceDialogs | | 1.0.4 |
| RCircos | | 1.1.2 |
| RClimMAWGEN | | 1.1 |
| Rclusterpp | | 0.2.3 |
| rCMA | | 1.0.0 |
| Rcmdr | 2.0-1 | 2.0-2 |
| RcmdrPlugin.BCA | | 0.9-7 |
| RcmdrPlugin.coin | | 1.0-21 |
| RcmdrPlugin.depthTools | | 1.3 |
| RcmdrPlugin.doBy | | 0.1-3 |
| RcmdrPlugin.DoE | | 0.12-2 |
| RcmdrPlugin.doex | | 0.2.0 |
| RcmdrPlugin.EACSPIR | | 0.2 |
| Installationsziel | | |
| Systemweit (in R framework) | | Installieren |
| 🔘 Für diesen Benutzer | | inkl. abhängiger Pakete |
| 🔵 In ein anderes Verzeichnis (wir | rd bei der Installation erfragt) | a man uphangiger Pakete |
| | | Allo Aktualisioren |

Figure 1. Installing package "Rcmdr".

Windows users should select the »*SDI* (single-document interface) option«, before loading "R Commander" allowing "R Commander" to work properly with "R". Therefore, one has to make a copy of the "R" desktop icon (Fox et al. 2013). Right-click the new icon and select »properties«. Add » - - sdi« (preceded by a single space) to the target field on the shortcut tab of the properties dialog box (Fox et al. 2013). Alternatively, one can select the »*SDI*« by clicking in "R" on »Edit« \Rightarrow »GUI preferences...« (Chang et al. 2010). For detailed information, use the following link: http://people.ysu.edu/~gchang/r/R_ChangeToSDI.pdf.

OS X users must install the application "XQUARTZ", which can be downloaded at the following link: http://xquartz.macosforge.org/landing/ (Fox et al. 2013).

For some important hints of how to install "R Commander" properly, use the following link: http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/nstallation-notes.html (Fox et al. 2013).

3.2 Loading the package "R Commander"

One can load "R Commander" with "R" by typing the following command »library (Rcmdr) « at the prompt (>) and pressing enter. Alternatively, one can load "R Commander" by clicking in the menu bar »Packages« ▷ »Load packages« ▷ »Rcmdr«. Loading "R Commander" for the first time "R" will ask to install some additional packages. Click »*Yes*« (Gutermuth 2010).

4. "R Commander" window

The following sectors provide details for each window illustrated in Figure 2.

| | ● O O X R Commander | |
|---------|---|---------------|
| | File Edit Data Statistics Graphs Models Distributions Tools Help | |
| / | 🕼 Data set: 🔟 < <u>No active dataset></u> Z Edit data set 🗟 View data set Model: 🗴 < <u>No</u> | active model> |
| | R Script R Markdown | |
| | | Toolbar |
| Manulan | | |
| menu ba | Script Window | |
| | | |
| | Tal. | X |
| | | F |
| | Output | Submit |
| | | |
| | | |
| | | |
| | Output Window | |
| | Output window | |
| | | |
| | | |
| | | |
| | | - |
| | | |
| | Messages | |
| | Message Window | |
| | | × |

Figure 2. "R Commander" Window.

4.1 Toolbar

The toolbar in Figure 2 is explained in Table 1.

| FUNCTION | DESCRIPTION | | |
|--------------|---|--|--|
| Dataset | Displays the current name of the active data set. | | |
| | One can choose among the data sets currently in | | |
| | memory by clicking the button »No active | | |
| | dataset«. | | |
| | | | |
| Edit dataset | One can edit the active data set by clicking the | | |
| | button »Edit dataset«. | | |
| | | | |
| View dataset | One can view the active data set by clicking the | | |
| | button »View dataset«. | | |
| | | | |
| Model | Displays the current name of the active model. | | |
| | One can choose among the models currently in | | |
| | memory by clicking the button »No active | | |
| | model«. | | |
| | | | |

4.2 Script window

This window displays commands generated by "R Commander". It provides the ability to edit, enter and re-execute them (Fox 2005). Commands can also be directly entered and executed by highlighting and clicking the button *»Submit«* (Gutermuth 2010). Several lines can easily be executed by left clicking and dragging over them. Commands must not be preceded by *»>«* (Gutermuth 2010).

4.3 Output window

This window displays the output in dark blue generated by the Script Window and therefore "R" Commands used in red.

4.4 Message window

This window displays error messages in red, warnings in green and notes in blue.

4.5 Menu bar

The menu bar in Figure 2 is explained in Table 2.

| MENU | SUBMENU |
|------------|--|
| File | Menu items for loading and saving script files, R |
| | Markdown files, for saving output and the R |
| | workspace and for exiting. |
| Edit | Menu items for editing the contents of the script |
| | and output windows. |
| Data | Submenus for loading data sets and manipulating |
| | data files. |
| Statistics | Submenus for a variety of basic statistical |
| | analyses. |
| Graphs | Menu items for creating a variety of basic |
| | statistical graphs. |
| Models | Menu items and submenus for obtaining |
| | numerical summaries, confidence intervals, and |
| | hypothesis tests, numerical diagnostics and |
| | graphs for a statistical model, and for adding |
| | diagnostic quantities (e.g. residuals) to the data |

| | set. | | | |
|---------------|---|--|--|--|
| Distributions | Submenus for continuous and discrete | | | |
| | distributions. | | | |
| Tools | Menu items for loading "R Commander" plugins and for setting some options. | | | |
| Help | Menu items to obtain information about the "R Commander" including an introductory manual. | | | |

Table 2. "R Commander" Menu bar.

Menu items and submenus, which are not applicable to the current content, are grayed out (inactive) (Scott 2012).

5. Creating a working directory

At all times "R" and therefore "R Commander" looks in a certain folder called directory. If one needs to work with this directory, it is of significant importance to create an own folder, which is easy to find. By typing the command »getwd() « and clicking »*Submit*« one can find out in which folder "R Commander" is currently looking in. One can change this directory by typing the command »setwd(»required directory«) «. By re-execution of »getwd() « the output window should now display the new directory as in Figure 3.



Figure 3. Changing the working directory.

Alternatively, one can change the directory with the menu bar. Click on »File« \Rightarrow »Change working directory« and select the required directory.

6. Creating data sets

"R Commander" is able to easily create data, distinguished by the operating system.

6.1 Windows

One can create a new data set by clicking on »Data« \Rightarrow »New Dataset«. One will then be asked to enter a name for the data set. As in Figure 4 a new window called »Data Editor« will appear in which one can type in the data. Name and type of the variables of each column can be defined by clicking on the »column header« (e.g. »Var1«) (Eitel 2012). If one wants to enter numeric values, select »numeric« under »variable type« (Eitel 2012). One can close the »Data editor« by clicking on »X« in the right hand corner. This data set is now the active data set displayed in the Toolbar. By clicking on »View data set«, the data set will appear in a new window. Clicking on »Edit data set« will bring up the data editor.

| R Da | R Data Editor | | | | | | | |
|----------|----------------|------|------|------|------|------|------|--|
| File i | File Edit Help | | | | | | | |
| | var1 | var2 | var3 | var4 | var5 | varб | var7 | |
| 1 | WT | 5.4 | | | | | | |
| 2 | WT | 5.3 | | | | | | |
| 3 | WT | 5.2 | | | | | | |
| 4 | WT | 5.4 | | | | | | |
| 5 | WΤ | 5.5 | | | | | | |
| 6 | HOM | 6.1 | | | | | | |
| - 7 | HOM | 6.2 | | | | | | |
| В | HOM | 6.5 | | | | | | |
| 9 | HOM | 6.2 | | | | | | |
| 10 | HOM | 6.4 | | | | | | |
| 11 | | | | | | | | |
| 12 | | | | | | | | |
| 13 | | | | | | | | |
| 14 | | | | | | | | |
| 15 | | | | | | | | |
| 16 | | | | | | | | |
| 17 | | | | | | | | |
| 18 | | | | | | | | |
| 19 | | | | | | | | |
| 20 | | | | | | | | |
| 19 20 | | | | | | | | |

Figure 4. Creating a data set.

6.2 OS X

The OS X version of "R Commander" does not simply have an option for creating a new data set. One can either use "R" codes to create a new data set and save it or use other software such as *»Excel«* or *»Text Editor«.* One needs to save the data set created in the working directory as a comma-delimited (*».csv«*) file from *»Excel«* or as a tab-delimited (*».txt«*) file from a *»Text Editor«.*

7. Importing data files

Importing data sets can be done by creating ones own or choosing data sets of the CRAN website.

7.1 Created data set

One can import a file saved in the working directory by clicking on $Data \ll P Theoremath{\mathbb{S}}$ and $P = Theoremath{\mathbb{S}}$ and

| 0 0 | 🗙 Read Text Data I | From File, Clip | board, or UR | L |
|--------------------------------|--------------------|-----------------|--------------|----|
| Enter name | for data set: | Dataset | | |
| Variable nan | nes in file: | | | |
| Missing data | indicator: | NA | | |
| Location o | f Data File | | | |
| Local file | system | | | |
| \diamond Clipboard | | | | |
| ♦ Internet U | RL | | | |
| Field Separ | ator | | | |
| White spa | ce | | | |
| Commas | | | | |
| | | | | |
| ♦ Other S | pecify: | | | |
| Decimal-Po | oint Character | | | |
| Period [.] | | | | |
| Comma [,] | l | | | |
| 🔞 Hel | p 🔰 | 🕻 Cancel | | OK |
| | | | | |

Figure 5. Characterising the data set.

Note:

The name of this data set cannot have spaces in it as well as "R" is case-sensitive ($>mydata \neq MyData \ll$) (Karp 2010).

7.2 CRAN data sets

Once a package is installed with "R" one can load a data set by clicking on »Data« ⇒ »Data in packages« ⇒ »Read data set from an attached package...«. A dialog box will appear as follows. The data set used is called »Anscombe« from the package »car«.

| \varTheta 🔿 🔿 🕅 🕅 🕅 🕅 🕅 | | | | | | |
|----------------------------------|----------------|---------------------|--|--|--|--|
| Package (Double-click to select) | Data set (Doub | le-click to select) | | | | |
| car | AMSsurvey | | | | | |
| datasets | Adler | _ | | | | |
| Imtest | Angell | | | | | |
| MASS | Anscombe | - | | | | |
| OR | | | | | | |
| Enter name of data set: | | | | | | |
| Help on selected data set | | | | | | |
| 🔞 Help | 💢 Cancel | OK | | | | |

Figure 6. Loading data set from package.

One can select the package on the left side by double-clicking. The corresponding data sets of this package will appear on the right side. By selecting a data set and clicking >OK this data set is then activated.

8. Saving data file

One is able to save the script as well as the output by clicking on \gg File« $\implies \gg$ Save script as...« or \gg Save output as...« respectively. When saving the script, always add the \gg .r« extension at the end of the name (Gutermuth 2010).

9. Statistics

"R Commander" is able to provide profound information about the data set in few steps.

9.1 Numerical summaries

If one wants to get some summary statistics of the active data set, open as in Figure 7 »Statistics≪ ⇒ »Summaries≪ ⇒ »Numerical summaries≪.

| 000 | | X R Commander | | |
|----------------|---------------------------|--------------------------------|--------|---|
| File Edit D | Data Statistics Graphs Mo | dels Distributions Tools Help | | |
| 📿 Data s | et: Summaries | ► Active data set | Model: | Solution Section 2 - Sectio |
| P. Script P. I | - Contingency tables | Numerical summaries | | , , |
| K SCIPT KI | Means | Frequency distributions | | |
| | Proportions | Count missing observations | | <u></u> |
| | Variances | Table of statistics | | |
| | Nonparametric tests | Correlation matrix | | |
| | Dimensional analysis | Correlation test | | |
| | Fit models | Shapiro-Wilk test of normality | | |
| | | | - | |
| | | | | |

Figure 7. Open Numerical summaries.

A dialog box will appear with the menu bar *»Data«* and *»Statistics«* as in Figure 8. In *»Data«*, one can select (greyed out) the variable(s) one would like to have summarized. Numerical variables but no factors are shown to choose from. By holding down the *»Ctrl key«* one is able to select more variables (Mecklin 2011). It is also possible to summarize by groups. Therefore, click on the button *»Summarized by groups...«* (Chang et al. 2010).

| 0 | X Numerical Summaries | | |
|---|-----------------------|------|-------|
| Data Statistics | | | |
| Variables (pick one or more) education income urban young | | | |
| 🔞 Help 😽 Reset | 💥 Cancel | 🚽 ОК | Apply |

Figure 8. Selecting variable(s) for Numerical summaries.

As in Figure 9 one can select *»descriptive statistics*« in the menu bar *»Statistics*« such as mean, standard deviation and quantiles (Mecklin 2011).

| • • • | Numerical Summ | aries | |
|--|----------------|-------|---------|
| Data Statistics | | | |
| Mean Standard Deviation Interquartile Range Coefficient of Variation Skewness Type 1 Kurtosis | | | |
| 🔞 Help 🥎 Reset | 💢 Cancel | 🖋 ОК | i Apply |

Figure 9. Selecting descriptive statistics.

By clicking »*OK*« a summary table will appear in the output window as follows:

| <pre>> numSummary(Anscombe[,c("education", "income", "urban", "young")], statistics=c("mean", "sd", + "quantiles"), quantiles=c(0,.25,.5,.75,1))</pre> | | | | | | | "sd", | | | |
|---|-----------|-----------|--------|---------|--------|---------|--------|----|--|--|
| | mean | sd | 0% | 25% | 50% | 75% | 100% | n | | |
| education | 196.3137 | 46.45449 | 112.0 | 165.00 | 192.0 | 228.50 | 372.0 | 51 | | |
| income | 3225.2941 | 560.02597 | 2081.0 | 2785.50 | 3257.0 | 3612.00 | 4425.0 | 51 | | |
| urban | 664.5098 | 151.34482 | 322.0 | 552.50 | 664.0 | 790.50 | 1000.0 | 51 | | |
| young | 358.8863 | 23.95998 | 326.2 | 342.05 | 354.1 | 369.15 | 439.7 | 51 | | |

Figure 10. Summary table of the data set »Anscombe«.

9.2 Correlation matrix

One can examine the strength of the linear relationship between each pair of variables. Therefore open »Statistics« \Rightarrow »Summaries« \Rightarrow »Correlation matrix...«. By selecting the variables in the dialog box and clicking »OK« one will get the matrix as follows. Closer to 1 means strong correlation.

| > cor(Anso | combe[,c("e | education",' | 'income","u | rban","young")], | use="complete") |
|------------|-------------|--------------|-------------|------------------|-----------------|
| | education | income | urban | young | |
| education | 1.0000000 | 0.6675773 | 0.2633238 | 0.3114855 | |
| income | 0.6675773 | 1.0000000 | 0.6854580 | -0.1623600 | |
| urban | 0.2633238 | 0.6854580 | 1.0000000 | -0.1386334 | |
| young | 0.3114855 | -0.1623600 | -0.1386334 | 1.0000000 | |

Figure 11. Correlation matrix of the data set »Anscombe«.

10. Graphs

"R Commander" is able to generate basic statistic graphs but the variety is strictly limited by the choice offered in the menu bar *»Graphs«.* In the following are three common graphs namely, Histogram, Scatterplot and Scatterplot matrix explained.

10.1 Histogram

If one wants to see the proportion of cases falling into each of several categories, use *»Histogram«.* Open *»Graphs« ⇒ »Histogram...«.* In *»Data«* one can select the variable one wants to display. In *»Options*« one can label the axes and choose axis scaling. By clicking *»OK*«, the histogram will appear in a new window as follows:



Figure 12. Histogram of the data set »Anscombe«.

10.2 Scatterplot

If one wants to examine the relationship between two variables, use Scatterplot. Open Graphs Scatterplots. In Data one can select the x and y variables. In Options one can modify the Scatterplot such as least square line, plot labels and points. By clicking OK the Scatterplot will appear in a new window as follows:



Figure 13. Scatterplot of the data set »Anscombe«.

10.3 Scatterplot matrix

If one wants to examine the relationship between more than two variables, use »Scatterplot matrix«. Open »Graphs« ▷ »Scatterplots matrix«. In »Data« one can select the variables (at least three). In »Options« one can modify the Scatterplot matrix such as the type of graph

on the diagonal and least squares lines. By clicking *»OK*« the Scatterplot matrix will appear in a new window as follows:



Figure 14. Scatterplot matrix of the data set »Anscombe«.

11. Multiple linear regression

A multiple linear regression consists of one response variable *Y« and two or more explanatory variables *X«. Rather than modelling the mean response as a straight line, as in simple regression, it is now modelled as a function of several variables (Lindquist 2012). One can do a regression modelling by clicking on $*Statistics \ll *$ Fit models \implies *Linear regression...« (Robison-Cox 2011). As in Figure 15 a dialog box will appear in which one can name the regression model (e.g. *Anscombe.regression*) and select the response variable (e.g. *income*) and the explanatory variables (e.g. *education, urban and young*).

| 0 0 | X Linear Regression | | | |
|--------------------------------|--|--|--|--|
| Enter name for model: Anscombe | e.regression | | | |
| Response variable (pick one) | Explanatory variables (pick one or more) | | | |
| education | education | | | |
| income | income | | | |
| urban | urban | | | |
| young | young | | | |
| Subset expression | | | | |
| <all cases="" valid=""></all> | | | | |
| | | | | |
| 🔞 Help 😽 Re | eset 🔰 🛠 Cancel 🖌 🗸 OK 🔗 Apply | | | |

Figure 15. Fitting a multiple linear regression.

By clicking »*OK*« "R Commander" will display the summary of the model as follows:

| Residuals: | | | | | | | |
|--------------------------------|----------------------------|-----------------------|----------------------|------------------------|-----------------|-----|--|
| Min | 10 Median | 30 | Max | | | | |
| -438.54 -145. | 21 - 41.65 | 119.20 | 739.69 | | | | |
| Coefficients: | | | | | | | |
| | Estimate Sto | d. Error | t value | Pr(> t) | | | |
| (Intercept) 3 | 011.4633 (| 509.4769 | 4.941 | 1.03e-05 | *** | | |
| education | 7.6313 | 0.8798 | 8.674 | 2.56e-11 | *** | | |
| urban | 1.7692 | 0.2591 | 6.828 | 1.49e-08 | *** | | |
| young | -6.8544 | 1.6617 | -4.125 | 0.00015 | *** | | |
| | | | | | | | |
| Signif. codes | : 0 '***' (| 0.001 '* | *' 0.01 | '*' 0.05 | | ' 1 | |
| Residual stan Multiple R-sq | dard error: uared: 0.79 | 259.7 or 979, Adju | n 47 deg usted R- | rees of fi squared: | reedom 0.785 | | |
| F-Statistic: | 61.86 ON 3 a | and 47 Di | F, p-va | tue: 2.384 | te-16 | | |

Figure 16. Output of the multiple linear regression.

By typing the regression model (*»Anscombe.regression«*) in the script window and clicking *»Submit«* one is able to see only the parameter estimates.

11.1 Breusch-Pagan test

An assumption of regression is that the variance of the errors is constant, or homogeneous, across observations (Oklahoma State University 1999). If this assumption is violated, the errors are said to be heteroscedastic. With the *Breusch-Pagan test* one can test the regression model of homoscedasticity. Therefore open *Models* \Rightarrow *Numerical diagnostics* \Rightarrow *Breusch-Pagan test for heteroscedasticity...* A dialog box will appear in which one can select the *Variance Formula*. By choosing *Explanatory variables* and clicking *OK* regression for homoscedasticity has to be rejected or not.

```
> bptest(income ~ education + urban + young, studentize=FALSE, data=Anscombe)
    Breusch-Pagan test
data: income ~ education + urban + young
BP = 5.6728, df = 3, p-value = 0.1287
```



11.2 Durbin-Watson test

The Durbin–Watson test is used to detect the presence of autocorrelation, a condition in which a relationship exists between consecutive residuals (Sullivan 2013). It tests the null hypothesis of no autocorrelation against the alternative hypothesis of autocorrelation. Open »Models« \Rightarrow »Numerical diagnostics« \Rightarrow »Durbin–Watson test for autocorrelation...«. By selecting »*rho!* = 0« in the dialog box, one gets the following result. Looking at the p-value tells one if the null hypothesis has to be rejected or not.

Figure 18. Durbin-Watson test for autocorrelation.

11.3 Variance Inflation

An important assumption for multiple regressions is that independent variables are not perfectly multicolinear. The variance inflation factor (VIF) tests the variables of multicollinearity. Open »Models« ⇒ »Numerical diagnostics« ⇒ »Variance-inflation factors«. "R Commander" will display results as follows:

```
> vif(Anscombe.regression)
education urban young
1.238866 1.140588 1.175557
```

Figure 19. Variance-inflation test.

A high variance inflation factor (VIF) indicates a multicollinearity problem. A generally approved convention, whether the values are acceptable or not does not exist. As a general rule, values of 5 or above are high (Robison-Cox 2011).

12. Getting help

If one needs some help on a function (e.g. *»Breusch-Pagan test (»bptest«)«*), type as in Figure 20 the following command *»help(bptest)«* or *»?bptest«* in the script window and click *»Submit«*. A new window will appear with detailed information about this asked function (Gutermuth 2010).

| help(bptest) ?bptest |
|-------------------------|
| |
| Output |
| > help(bptest) |
| > ?(bptest) |

Figure 20. Getting help function.

If one needs some general help about "R Commander" click in the menu bar *»Help«* as follows:

| \varTheta 🖸 🖸 📉 🛛 🕅 🕅 | |
|---|--|
| File Edit Data Statistics Graphs Models Distributions Tools | Help |
| R Script R Markdown | Commander help Introduction to the R Commander R Commander website |
| | About Rcmdr |
| | Help on active data set (if available) |
| | Start R help system |
| | R website |
| | Using R Markdown |
| 4 | |

Figure 21. Menu bar Help.

13. Closing "R Commander"

Closing "R Commander" can be done by clicking on "File" \Rightarrow "Exit" or by clicking the "X" of the "R Commander" window. If one has not saved previously the data set, one will always be asked to save the script file, markdown and output generated. "R" can get confused by objects (data sets / parameters) carried over, therefore avoid exiting from "R" (Karp 2010). Accidentally closed "R Commander", one can easily load it again by typing the command "Commander()" (Gutermuth 2010).

14. The Comprehensive R Archive Network (CRAN)

CRAN provides, with increasing tendency, packages for every type of science.

14.1 CRAN packages

The official website of "R" provides packages which contain plugins, tools and functions. All these packages sorted alphabetically are available and can be downloaded under the following link: http://cran.r-project.org/web/packages/available_packages_by_name.html#available-packages-R.

These packages will be installed exactly as the "R Commander" package "Rcmdr" described above (cf. Chapter 3). "R Commander" already contains some additional packages. One can import these data sets by clicking on $Tools \iff Load package(s) \ll (Gutermuth 2010).$

14.1.1 Contributed "R Commander" plugins

The following list contains all plugins for "R Commander" available in the *»CRAN Archive«*. These packages extend the range of applications and are therefore highly useful (Gutermuth 2010). Clicking on the package on the left side will directly lead one to the *»CRAN Package«* website.

PACKAGE

DESCRIPTION

RcmdrPlugin.BCA

Rcmdr Plug-In for Business and Customer Analytics

| RcmdrPlugin.coin | Remdr Coin Plug-In |
|-------------------------|---|
| RcmdrPlugin.depthTools | R commander Depth Tools 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) |
| RcmdrPlugin.HH | Remdr 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 |
| RcmdrPlugin.lfstat | Rcmdr Plug-In for low flow analysis |
| RcmdrPlugin.MA | Graphical User Interface for Conducting Meta- Analyses in R |
| RcmdrPlugin.mosaic | Adds menu items to produce mosaic plots and assoc plots to Remdr |
| RcmdrPlugin.MPAStats | R Commander Plug-in for MPA Statistics |
| RcmdrPlugin.orloca | orloca Rcmdr Plug-in |
| RcmdrPlugin.plotByGroup | Rcmdr plots by group using lattice |
| RcmdrPlugin.qual | Rcmdr plugin for quality control course |

| Tools for sampling in Official Statistical Surveys |
|--|
| Rcmdr Plugin for seeg |
| SLC Rcmdr Plug-in |
| Rcmdr Sport Management Plug-In |
| Efficiently search the R help pages |
| Statistical URV Rcmdr Plug-In |
| R Commander Plug-in for the survival Package |
| Rcmdr Teaching Demos Plug-In |
| Graphical Integrated Text Mining Solution |
| UCA Rcmdr Plug-in |
| |

Table 3. "R Commander" Plugins.

14.2 CRAN task views

For the sake of simplicity, all packages are sorted by *»Task Views«* under the following link: http://cran.r-project.org/web/views/. *»CRAN Task Views«* will help to choose the appropriate package(s) corresponding to the empirical work. If one wants to install *»CRAN Task Views«*, the *»ctv package«* initially needs to be installed. Once the *»ctv package«* is installed, one can use the command *»install.views(»required Task View«)«*.

15. Tutorials & additional reading material

One can find a lot of information about "R Commander" in the Internet. Nevertheless, most papers are quite similar to each other and therefore not helpful in getting profound information. The first listed website contains the basic document about "R Commander" written by the developer Prof. John Fox that should once be read. The second and third contain written tutorials on specific topics explained understandably and very often further accompanied by an appropriate video.

- 1. http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/Getting-Started-with-the-Rcmdr.pdf
- 2. http://people.ysu.edu/~gchang/r/R_Instructions.htm

 http://instruction.bus.wisc.edu/jfrees/UWCAELearn/Lists/Rcmdr%20Teaching%20Videos/All Items.aspx

16. Conclusion

Concluding, this paper gives a short overview about the basic operations of "R Commander" and additionally some hints to get profound information. It will benefit students and novices to get more easily familiar with the program "R" and eventually with the commands. Therefore, "R Commander" is a highly useful and a recommended plugin, helping to get started with data analysing.

17. References

Chang, G.-A., Kerns, G.-J. 2010. "Introduction to probability and statistics using R/RCommander with IPSUR Plug-In". http://people.ysu.edu/~gchang/r/IPSURRcmdr.pdf. Reviewed at 19.03.2014

Diez, D. "R Basics". http://www.ddiez.com/teac/r/basics.php. Reviewed at 19.03.2014

- Eitel, J.-U.-H. 2012. "R Commander Tutorial". http://ecosensing.org/wp-content/uploads/2012/04/R-commander-tutorial1.pdf. Reviewed at 19.03.2014
- Fox, J., Bouchet-Valat, M. 2013. "Getting started with the R Commander". http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/Getting-Started-with-the-Rcmdr.pdf. Reviewed at 19.03.2014
- Fox, J. 2013. "R Commander Installation Notes". http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/installation-notes.html. Reviewed at 19.03.2014
- Fox, J. 2005. "The R Commander: A Basic-Statistics Graphical User Interface to R". http://www.jstatsoft.org/v14/i09/paper. Reviewed at 19.03.2014
- Graham, J. "Introduction to R/R-Commander for Windows". http://www.math.umt.edu/graham/stat216/Rchapter1.pdf. Reviewed at 19.03.2014
- Gutermuth, D. 2010. "Introduction to R and R Commander". http://www.wlu.ca/documents/42689/Introduction_to_R_and_R_Commander.pdf. Reviewed at 19.03.2014
- Hutcheson, G. 2012. "R and Rcmdr". http://www.methods.manchester.ac.uk/events/2012-0416/Rhandout.pdf. Reviewed at 19.03.2014
- Karp, N.-A. 2010. "R Commander an Introduction". http://cran.r-project.org/doc/contrib/Karp-Rcommander-intro.pdf. Reviewed at 19.03.2014
- Lindquist, M.-A. 2012. "Multiple Linear Regression". http://www.stat.columbia.edu/~martin/W2024/R6.pdf. Reviewed at 19.03.2014
- Mecklin, C.-J. 2011. "An Introduction to the R Commander". http://campus.murraystate.edu/academic/faculty/cmecklin/Presentations/IntroductiontoRCommand erHandout.pdf. Reviewed at 19.03.2014

Oklahoma State University. 1999 "Heteroscedasticity". http://www.okstate.edu/sas/v8/sashtml/ets/chap14/sect39.htm. Reviewed at 19.03.2014

Robison-Cox, J. 2011. "Using R-Commander". http://www.math.montana.edu/~jimrc/classes/stat410/notes/RcmdrInstructions.pdf. Reviewed at 19.03.2014

Scott, T.-A. 2012. "An Introduction to the "R Commander: A Basic-Statistics Graphical User Interface (GUI) for R".

http://biostat.mc.vanderbilt.edu/wiki/pub/Main/TheresaScott/RCommander.TAScott.handout.pdf. Reviewed at 19.03.2014

- Sullivan, J.-H. 2013. "Regression diagnostics III (Time Series)". http://www.jhs14.business.msstate.edu/bqa9333/tips/11.65_durbinWatsonTest_Tables_kw.pdf. Reviewed at 19.03.2014
- Torres-Reyna, O. "Linear Regression". http://dss.princeton.edu/training/Regression101.pdf. Reviewed at 19.03.2014