

R Commander Worksheet for Practice

Open R by clicking on the icon on your desktop. At the prompt, type:
`library(Rcmdr)`
This should open R Commander in a separate window.

IMPORTING DATA:

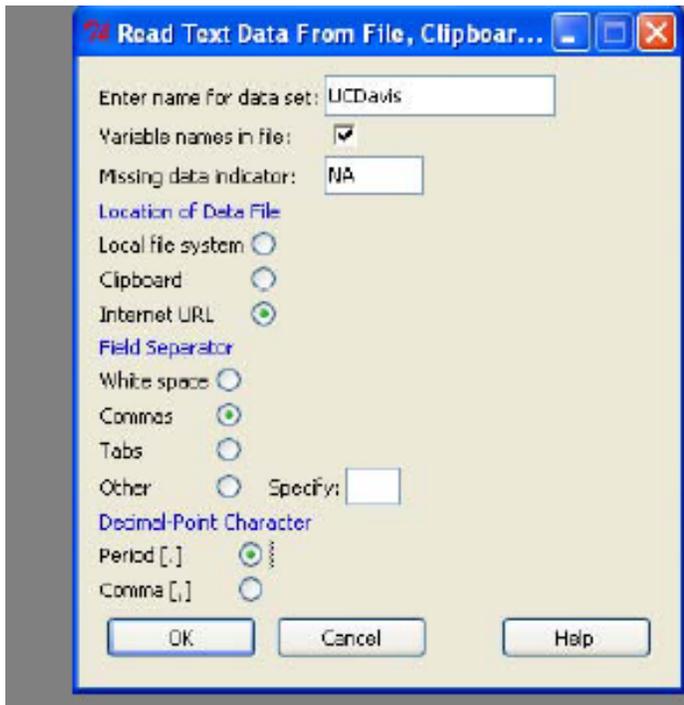
Data -> Import data -> from url textfile, clipboard or url...

Fill in the popup box as shown, except you can choose your own name for the data set. Things to notice:

Click the buttons for “Internet URL”

Click the “Field Separator” choice “Commas”

Otherwise, you will leave things as they were.



In the box asking for the url, type <http://www.ics.uci.edu/~jutt/7/ucdavis1.txt>

NOTE: If you prefer, you can save the data set to your computer first, then in the screenshot above you would click “Local file system” instead of “Internet URL.”

The data set should now be there. To check, click on “View Data Set.” It will pop up in a new window. Be sure to close it before you proceed.

TO SAVE YOUR DATA SET SO YOU CAN WORK ON IT AGAIN LATER:

Data → *Active data set* → *Save active data set*

Save it where you like. Then the next time you open R Commander, go to:

Data → *Load data set*

It should bring up your files, and just find the one you saved and open it.

CREATE SOME GRAPHS:

Create some histograms [*Graphs* -> *Histogram*, then choose variable]

Create some boxplots and stem-and-leaf displays.

TO SAVE A GRAPH:

Once you close the graph window, the graph is gone. (This is not true for stemplots, which appear in the regular output window.) So, do this first:

Graphs → *Save graph to file* → *as bitmap* (or pdf, etc, if you prefer), then I usually choose JPEG. You can then insert it as a picture into a Word document, or you can just print it.

NOTE: If your graph seems to have disappeared but you did not close the graph window, click on “R Gui” and it should reappear.

FINDING SUMMARY STATISTICS:

Try finding some summaries for the variables in the data set. For example, to find various summary values for Exercise:

Statistics -> *Summaries* -> *Numerical summaries*

Then click on “Exercise” and leave everything else as is. You should get these results:

mean	sd	0%	25%	50%	75%	100%	n	NA
4.52907	4.463861	0	2	3	6	30	172	1

CREATING A SUBSET OF THE DATA SET:

For example, to get a data set with the females only:

Data -> *Active data set* -> *Subset Data Set*, then in the “Subset Expression” box type

Sex = = “Female”

TO FIND PROPORTIONS FOR NORMAL DISTRIBUTIONS:

Distributions → *Continuous distributions* → *Normal distribution* → *Normal probabilities*

A pop-up box will ask you for the value, the mean, and the standard deviation. You do not need to compute the z-score first, just enter the appropriate mean and standard deviation. Alternatively (and more work) you can compute the z-score first, and then use the default mean of 0 and standard deviation of 1. In either case, the answer will appear as a decimal number in the Output Window.

Example: The probability of having a UCI Verbal SAT score of at least 674 was shown in lecture to be about .08. Do the above, then enter these in the popup box:

Variable value: 674

Mean: 569

Standard deviation 75

Upper tail [gives you the proportion above the value instead of below it]

Result:

[1] 0.08075666

This means that the proportion of Verbal SAT scores above 674 is about .0807... or about 8%.

EXITING R AND R COMMANDER

You can exit both at the same time using *File* → *Exit* → *From Commander and R*

You will be asked if you want to save your script file and your output. If you need the output for your homework, you should save it. It will be saved as a text file (.txt).