

Introduction to R

The R interface is a just that - a basic interface. In order to obtain R, you must first download the software from the link provided here: <https://www.r-project.org/>. Then, in order to use the software, you must load a dataset and use packages and commands in order to tell the software how to evaluate and handle the data. Each package is open-sourced and contains different tools in order to implement all sorts of data manipulation techniques.

Note: The dataset used in this tutorial and the R Script are on Moodle:

Packages, Libraries, Loading Data, and Viewing Data

1. Install packages for use.

Example: The foreign package and library allows you to load SPSS and STATA dataset file formats. You will want to load dependency packages also just in case the package/commands needs the support of others to work.

Note: You need to select a mirror site for downloading packages. The mirror site you choose does not matter.

Command:

```
install.packages("foreign", dependencies=TRUE)
```

2. Loading a package for use.

Command:

```
library(foreign)
```

3. Loading the dataset.

Example: Loading the 2016 Cooperative Congressional Election Study dataset in STATA format (.dta format) that is provided on Canvas or found at the study's webpage. When loading the dataset, the code here specified that variables should remain numerical rather than convert them to factors. Therefore, if a variable was say gender (i.e. men and women), the variable will be coded as a number, such as 1 for men 2 for women. Note: I named the dataset "dat" here.

Command:

```
dat <- read.dta(file.choose(), convert.factors=FALSE)
```

Commands: If the dataset is a STATA version 13 dataset.

```
install.packages("readstata13")  
library(readstata13)
```

```
dat <- read.dta13(file.choose(), convert.factors=FALSE)
```

4. Additional examples for different data formats:

Example: Reading an SPSS file (.sav format) into R

Command:

```
dat <- read.spss(file.choose(), convert.factors=FALSE)
```

Example: If you want to read an excel file (.xls or .xlsx format) into R you will need a different package. The code is below.

Command:

```
install.packages(gdata, dependencies=TRUE)
library(gdata)
dat <- read.xls(file.choose())
```

5. Exploring the Dataset.

Example: Looking at the names of the variables in the dataset.

Command:

```
names(dat)
```

Example: If we wanted to know the number of rows we have in the dataset (i.e. observations) we could use the “nrow” function. Here, we have 64,600 American Survey Respondents.

Command:

```
nrow(dat)
```

Example: If we wanted to get a quick idea of the structure of the variables. We could simply summarize the entire dataset.

Command:

```
summary(dat)
```

6. Saving the Workspace.

Example: If you want to save the workspace and all of the work you have done the command is simple (note: this does not save your code though. You must save the code separately).

Command:

```
save.image("TestWorkspace.Data")
```

Additional Note: Help with a Library, Command, or Function

Example: If you would like to find out what function a command performs. A help file can be accessed by typing ?? and the code. If we wanted to know what function the “names” command performed, we could type the command below.

```
??names
```