# Installing R and RStudio

* Download and install the R language from [r-project.org](https://cloud.r-project.org/).
* Download and install RStudio Desktop from [RStudio.com](https://www.rstudio.com/products/rstudio/download/#download): Use RStudio to program, write scripts, plot graphs, and look up documentation in R; this is the standard integrated development environment (IDE).

# Importing Data

Once you have RStudio installed and open, there are a few ways to import your data:

* Use the data() command for built in datasets; for example, typing data(iris) will import the classic “iris” dataset into the current environment.

> data(iris)

Then, type

> iris

It will display the entire dataset of “iris” including the variable names.

* Use ***read.csv*** statement to import a spreadsheet as a two-dimensional data frame (array-like structure) from a \*.csv file used to store data. For example, Excel can output CSV files using the "Save as" menu and you have a saved “iris.csv” in the default path after installing R.

>my.data<- read.csv("iris.csv", sep = ',', header = TRUE)

Or if you know a specific path to your data, you can specify like this, e.g.

> my.data<- read.csv("/users/hfang/iris.csv", sep = ',', header = TRUE)

* + Assign the output of read.csv to a variable using “<-”, e.g. in this example “my.data”
  + To learn more options for read.csv, type ***?read.csv*** or ***help(read.csv****)*
* R can store data in a variety of data structures, including data frames, matrices, and lists. ***read.csv*** reads in data as a dataframe: A dataframe stores data such that each row represents a single observation and each column represents a single variable. Create a dataframe with the ***data.frame()*** command, e.g.,

**> data.frame(iris)**

* Use c( ) to manually enter a vector of values, as in ***c(1, 2, 3)***

# Using Data

Once you have your data imported, using data(iris) for example, you will want to get a sense of what it looks like.

* ***View(iris)*** Invoke a spreadsheet-style data viewer on a matrix-like R object in RStudio.
* ***summary(iris)*** shows descriptive statistics and counts for a data frame.
* ***head(iris)*** will show you the first 6 rows
* ***names(iris)*** will show the names of the columns.
* The dollar sign allows access to a specific column, like ***iris$Species***
* To access a specific row, use ***iris[1,]*** (don't forget the comma).
* To access a specific value, use ***iris[1,1]*** to access the value of Row 1 and Col 1 in the data frame.
* Most operations in R will operate on more than one value at a time, for instance if you run my.variable <- c(1, 2, 3) + 5

> my.variable

[1] 6 7 8

my.variable will be equal to c(6, 7, 8)

* ***rm()*** or **remove()**: e.g., ***rm*(iris)** will remove the loaded iris dataset.

# Using Libraries

Several thousand libraries are available for R.

* To install the library “e1071”, type ***install.packages("e1071")***; or click on the “tools” menu on the top bar.

Graphical user interface, text, application, chat or text message

Description automatically generated

Or

In the bottom right window of RStudio, look for the menu “packages”.

Graphical user interface, table

Description automatically generated

* To use the library, add ***library("e1071")*** to your code after installation.

# Exporting Data

* Use the ***save()*** command to save a dataframe, like so: ***save***(iris, file="iris.rda").
* Use the load command to restore the data that you saved ***load***(file = "iris.rda")
* Output to a CSV file with ***write.csv*** like so: ***write.csv***(iris, file="iris.csv")

# Install R packages

* If you need to review packages available in R, go to <https://cran.rstudio.com/>
* For Windows users, if you want to install a package after your package review at the above link, go to RGui and click on the menu called “Packages” 🡪 click on “Install package(s)…” [🡪 pick any place (e.g. server at “USA (OH)”] 🡪pick and click on the package you want from the window. Once you use this procedure for the first time, the step in the bracket won’t be shown the next time.
* For Mac users, if you want to install a package after your package review at the above link, go to R Console and click on the menu called “Packages & Data” 🡪 click on “Package Installer” 🡪 choose “CRAN (sources)” and click on “At User Level” at the bottom of the interface and then check “Get List”[🡪 pick any place (e.g. server at “USA (OH)”] 🡪pick and click on the package you want from the window. Once you use this procedure for the first time, the step in the bracket won’t be shown the next time.

# Getting Help

* To view the documentation for a command in R, simply type ***?[command]***

For example, to remind yourself of the parameters for creating a dataframe, you can type and run the command ***?data.frame,*** which will then show the documentation on the bottom right window of RStudio.

* Do not be afraid to read sites such as StackOverflow or to search the internet for documentation on specific packages. You can even search for packages for a specific function that you want. For example, if you needed to process strings, an internet search would probably yield documentation for the package “stringr” (<https://www.rdocumentation.org/packages/stringr/versions/1.4.0>) which provides the information you need.