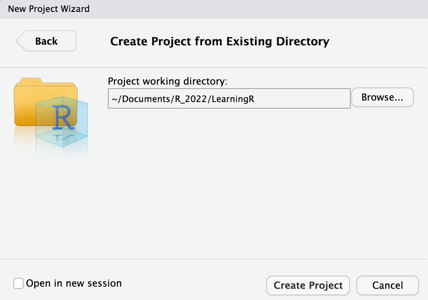
# Ch02: The basic data types and structures

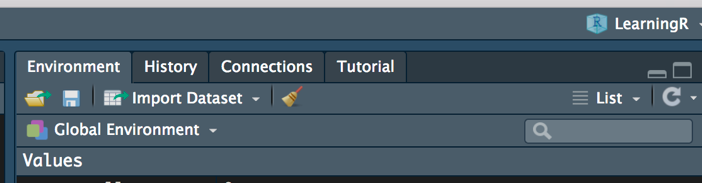
## Good practices

### Create a project

Now that you have set up R and R Studio, here are a few pointers to have a clean work environment. Create a project for this course. Go to File>New Project> Existing or new directory (depending on if you’ve already made a folder with the documents of this course). It should look like this:



Create your project. RStudio should look like this now:



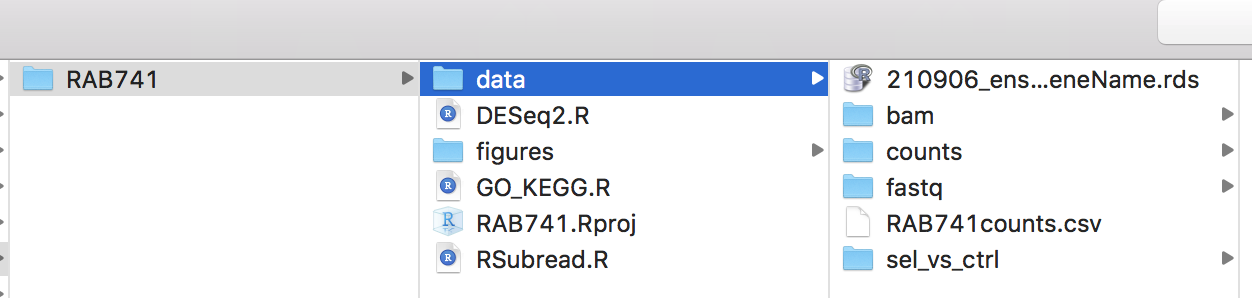
Now everytime you open this project, all your scripts previously opened in it will be available and you working directory will have been set to your project path.

### Comment your code

Use the # sign to write comment in your code. It’s useful and crucial to reproducibility. In theory, any person with some notions in R should be able to understand every step of your script just by reading your comments.

### Create your usual subfolders

In the spirit of clarity and reproducibility, your files should also be stored in an organized manner. It’s a good reflex to create subfolders like “data” and “figures” in each of your project. A typical RNAseq project might look like this:



### Naming your variables

Choose short variable names that do not start with numbers. R is case-sensitive so Gene\_name and gene\_name are considered different. Avoid using dots in the name (data.set > dataset).

## Get started with coding

### The different data types

| Type | Abbreviation | Example |
| --- | --- | --- |
| Characters | chr | “CD44” “10” “TRUE” |
| Numeric | num | 10 0.001 |
| Logical | logi | TRUE FALSE |

### The different data structures

| Type | Characteristics | Example |
| --- | --- | --- |
| Vector | Contains only a single datatype at the time | c(“CD44” ,“10” ,“TRUE”)  c(1, 3, 0.001)  ! be mindful of class coercion |
| List | Can contain multiple datatypes | list(1,"PRKAA2", TRUE) |
| Matrix | Assembly of same size vectors | matrix(1:6, nrow = 2, ncol = 3) |
| Dataframe | Assembly of same length lists | data.frame(ID = 1:3,  name = c("PRKAA1", "CFTR", "EGFR")  expr = round(runif(3,0,1), 2)) |

### Operations on the various data structures

### Creating



### Subsetting



### Modifying



### Other useful things

