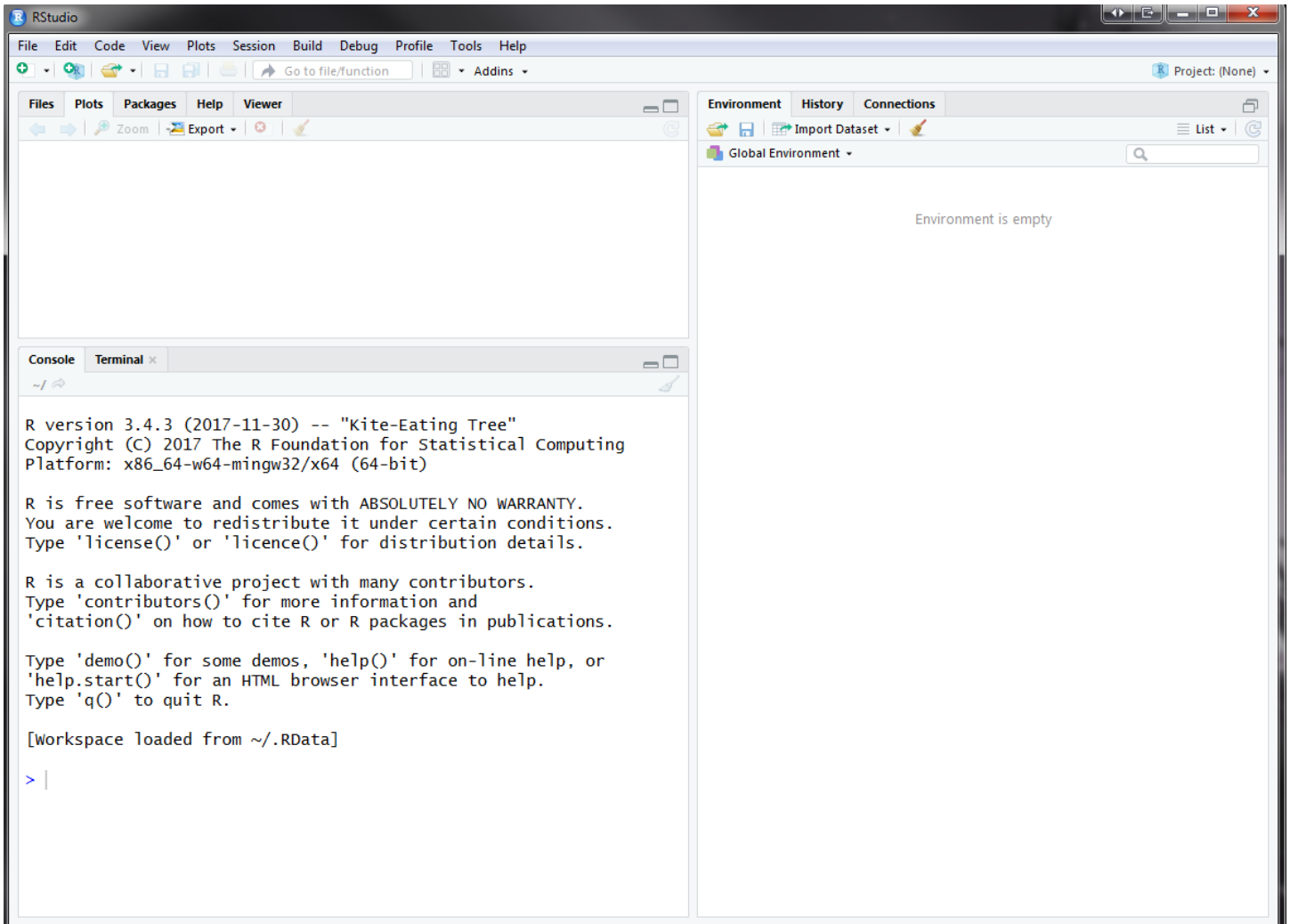


R Handout: Loading a Dataset into R

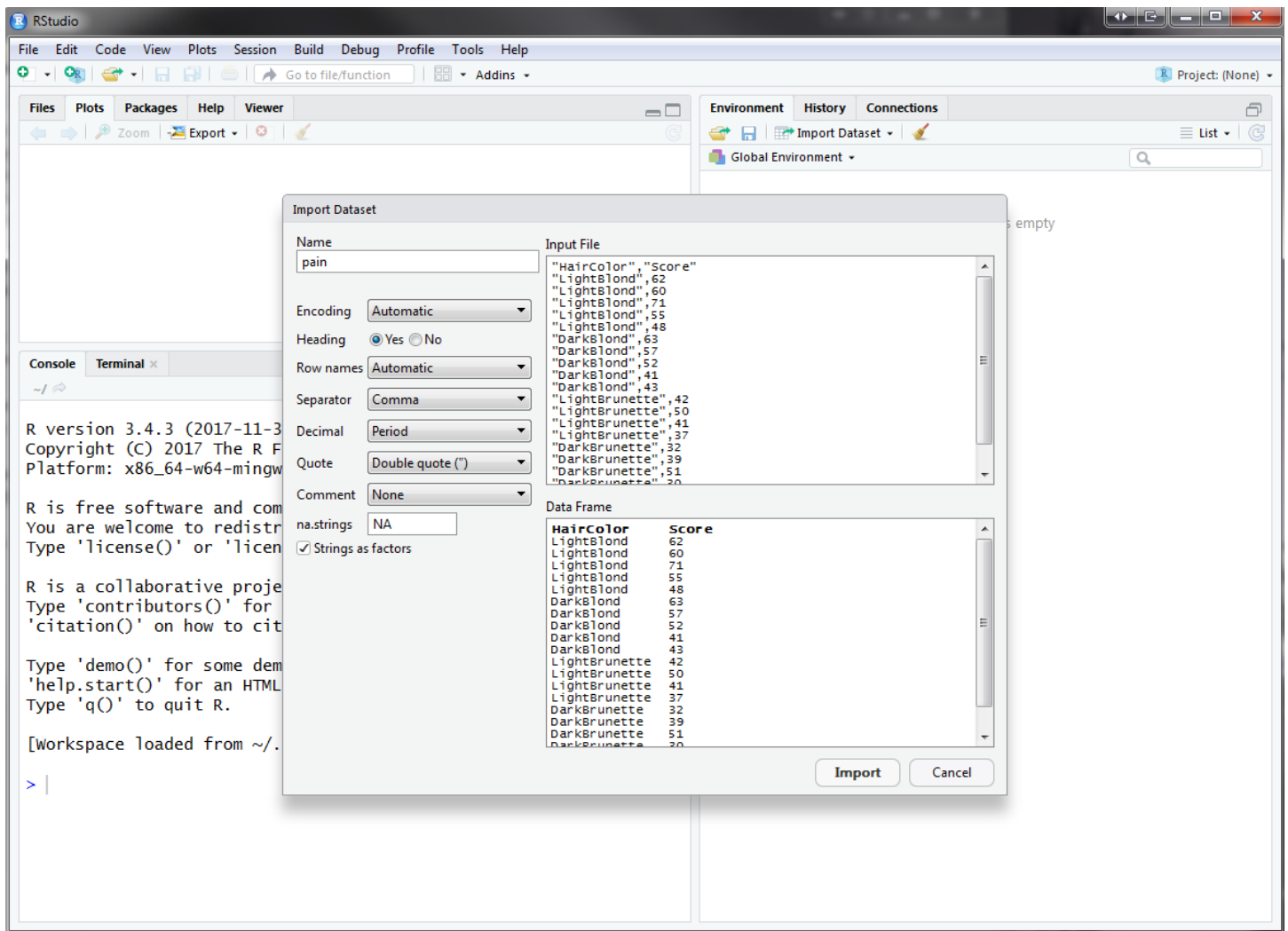
1 Loading a Dataset

Once you have installed R and R-studio, click on “File”, “New File”, “R Script”. This will open a .R file, where you can put your code that you want to run.

You should get a similar window to below (assuming you have reordered your windows to match mine):



Now, click “Import Dataset” (in the environment window), and select “From Text (base)”. Browse to the folder where you downloaded the dataset, select it, and then hit open. You should see something like this (the dataset will change with what file you open of course):



Next, hit “Import”.

Your screen should now look something like this:

The screenshot shows the RStudio interface with the following components:

- Console:** Displays the R version (3.4.3), copyright information, and the commands used to load the 'pain' dataset from a CSV file. The output shows the first 14 rows of the dataset.
- Environment:** Shows the 'Global Environment' with the 'pain' object containing 19 observations of 2 variables.
- Data Viewer:** Displays a table with 19 rows and 2 columns: 'HairColor' and 'Score'. The first 14 rows are visible, showing a mix of 'LightBlond', 'DarkBlond', and 'LightBrunette' hair colors with corresponding scores.

Console Output:

```
R version 3.4.3 (2017-11-30) -- "Kite-Eating Tree"
Copyright (C) 2017 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

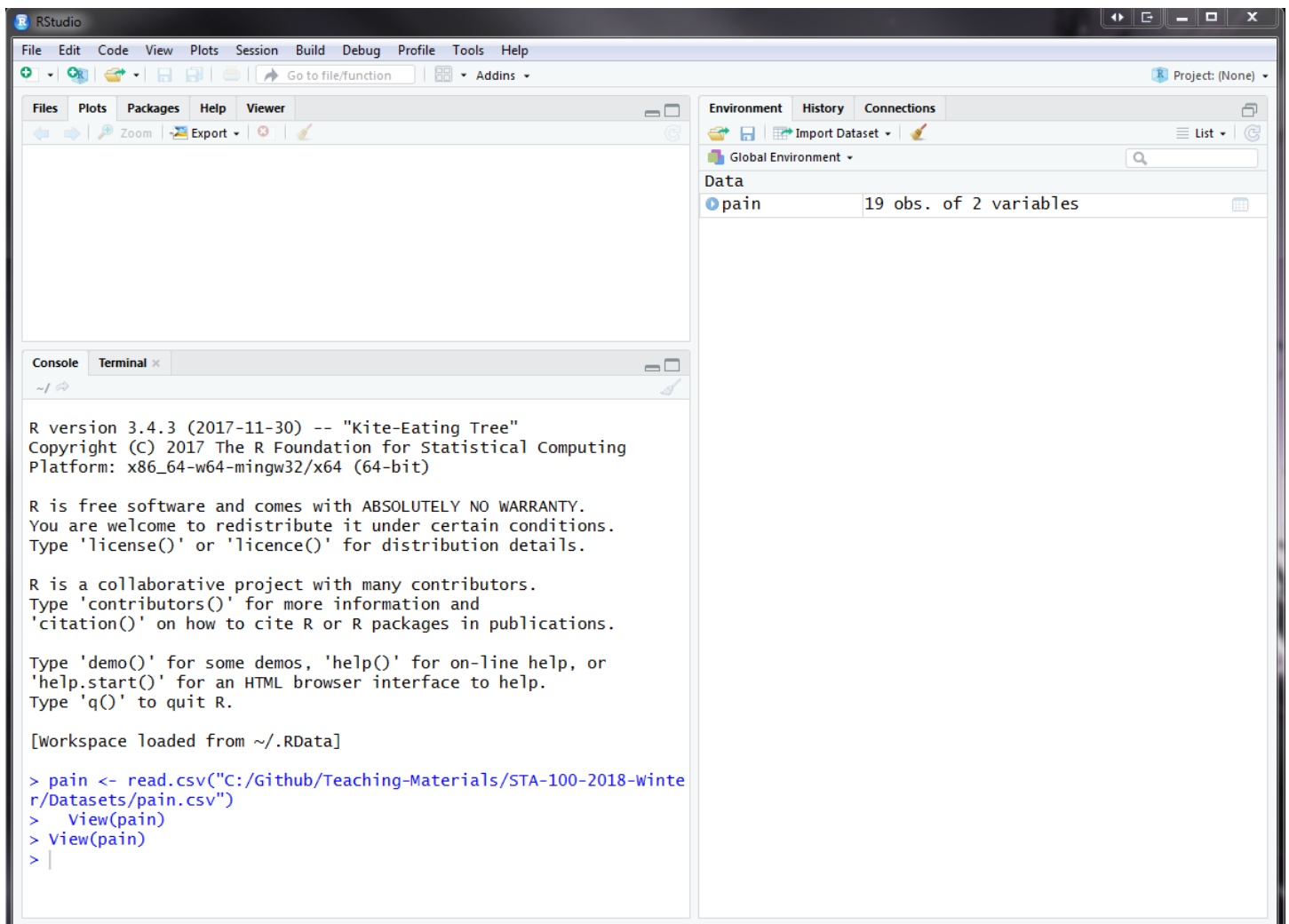
[Workspace loaded from ~/.RData]

> pain <- read.csv("C:/Github/Teaching-Materials/STA-100-2018-Winter/Datasets/pain.csv")
> View(pain)
>
```

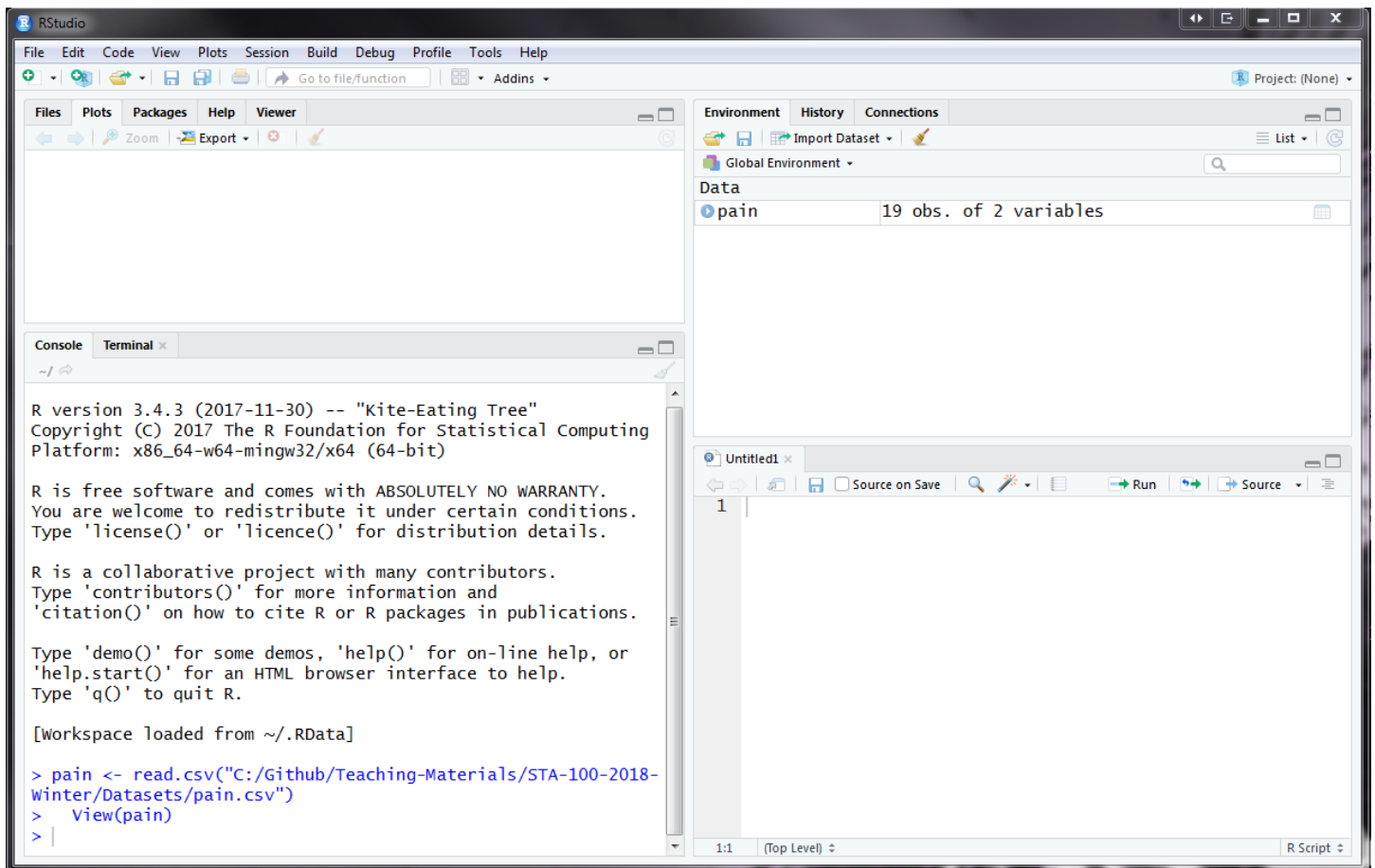
Data Viewer Table:

| | HairColor | Score |
|----|---------------|-------|
| 1 | LightBlond | 62 |
| 2 | LightBlond | 60 |
| 3 | LightBlond | 71 |
| 4 | LightBlond | 55 |
| 5 | LightBlond | 48 |
| 6 | DarkBlond | 63 |
| 7 | DarkBlond | 57 |
| 8 | DarkBlond | 52 |
| 9 | DarkBlond | 41 |
| 10 | DarkBlond | 43 |
| 11 | LightBrunette | 42 |
| 12 | LightBrunette | 50 |
| 13 | LightBrunette | 41 |
| 14 | LightBrunette | 37 |

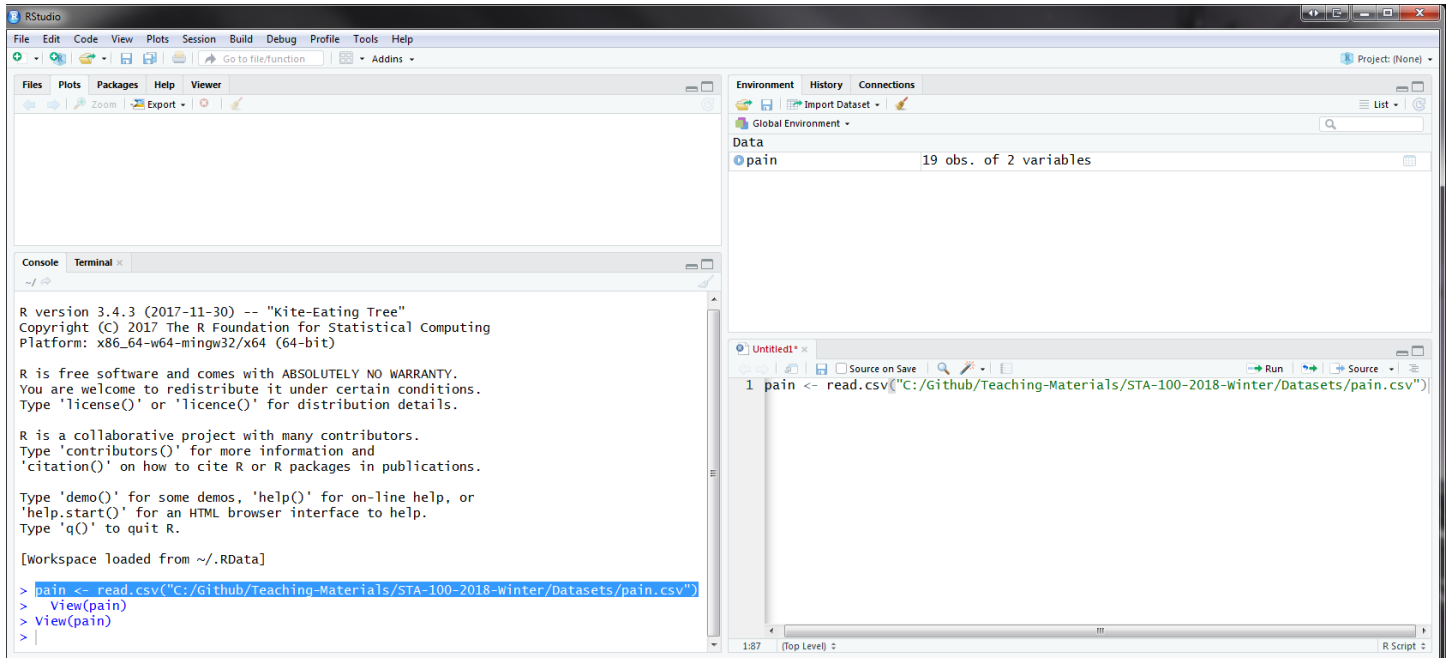
Notice this opened a tab that says “pain”, or in general the name of your dataset. This simply displays the format of the data, so we may go ahead and close that by pressing the “x” button next to the name. Now your window should look like this:



Next, we want to open a new R script to copy and paste code into. To do that, click “File”, “New File”, “ R Script” Your window should now look like this:



I highly recommend copying and pasting the command `R` used to load the data into the `R` script window. This allows you to load in the data without going to the dropdown menu, and is useful if you use this code again. You want to select everything starting from the name of the dataset (in this case, `pain`) to the end of that line (be sure to remove the “>” sign):

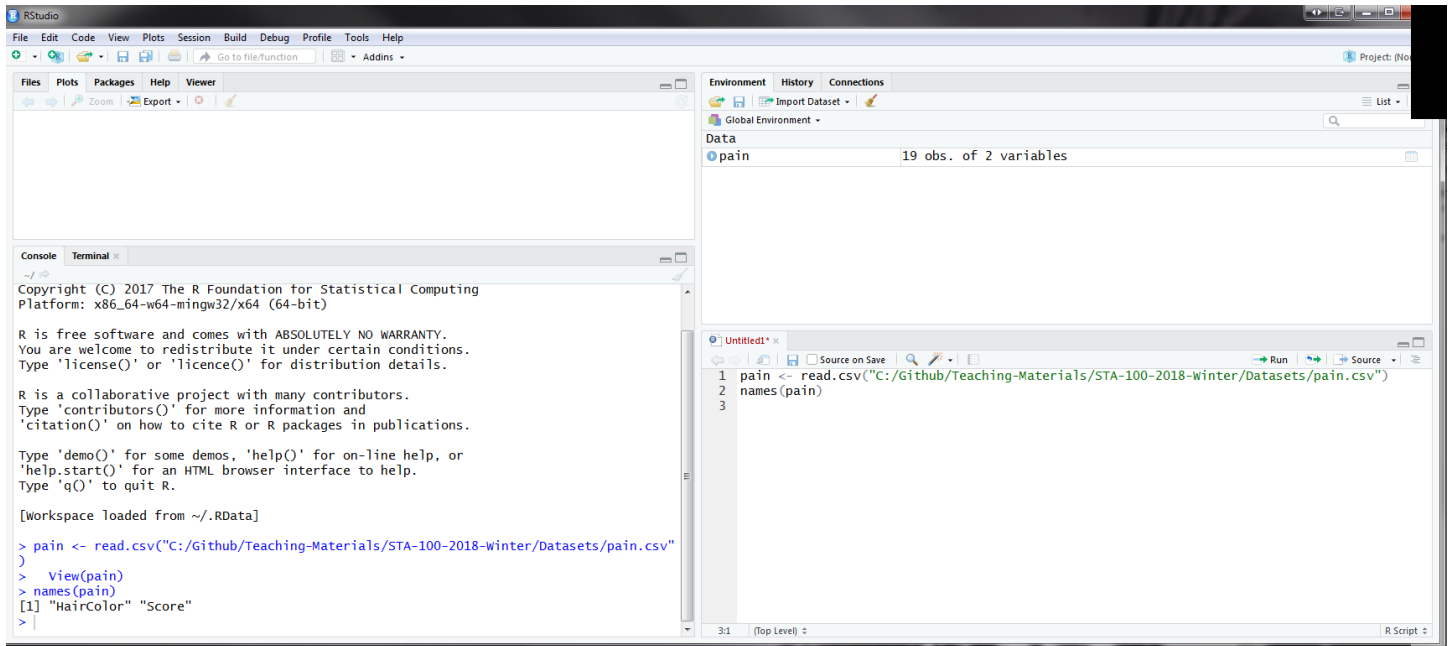


This allows you to simply copy and paste those commands next time you run `R`, without having to use the dropdown menu. In other words, with these two commands, the first thing that your `.R` file will do is **load the data**. Then we can manipulate the data within `R`, and put all of our code on different lines on the `.R` file.

To save the `.R` file (and all your code), either hit “Ctrl”, “s” on your keyboard (and save it with whatever name you like), or click “File” and “Save” from the drop down menu.

Now, we want to put all of our commands in the `.R` file, on separate lines. To run a command and get the result in `R`, we can put our mouse on that line, and hit “Run” in the upper right hand corner, or “Ctrl” and “Enter” on our keyboard.

For example, if I wanted the names of the columns (See week 1 handout for further details), my .R file and console (after I have run the command) would look like:



Please see the weekly handouts to learn new commands.