Collecting data/sampling and introduction to R

Overview

Review

Intro to R and some DataCamp exercises

Sampling

- Parameters and statistics
- Proportions

Announcements

If you haven’t done so yet, please remember to fill out the class survey:

https://goo.gl/GxE8qk

Also join the class Slack group – has anyone not joined?

https://https://goo.gl/DvMsNB

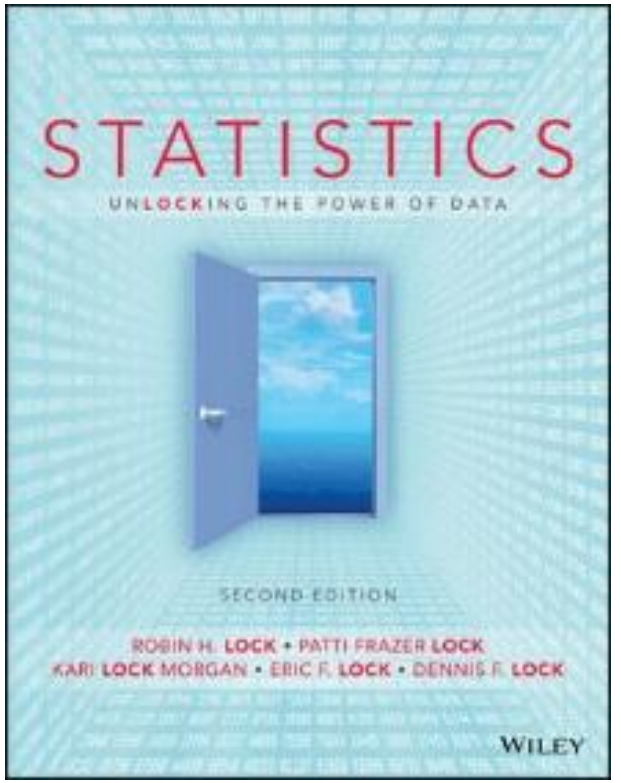
Any questions about the Lock5 practice problems?

Practice problems from Lock 5, first edition:

1.1, 1.3, 1.5, 1.11, 1.25, 1.26

Have people ordered the book?

- First edition or second edition?



Quiz time!

1. What is a population?

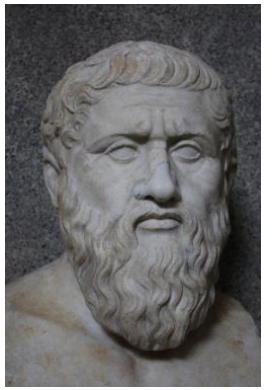
2. What is a sample?

3. What is statistical inference?

4. What are the rows of a data set called?

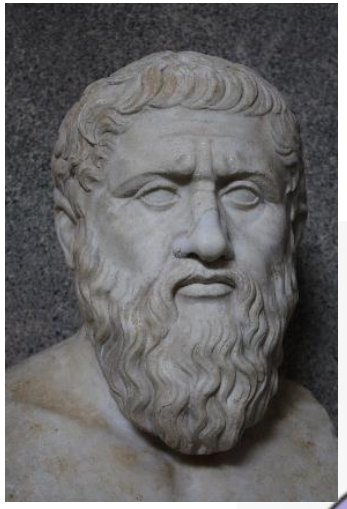
5 . What are the columns of a data set called?

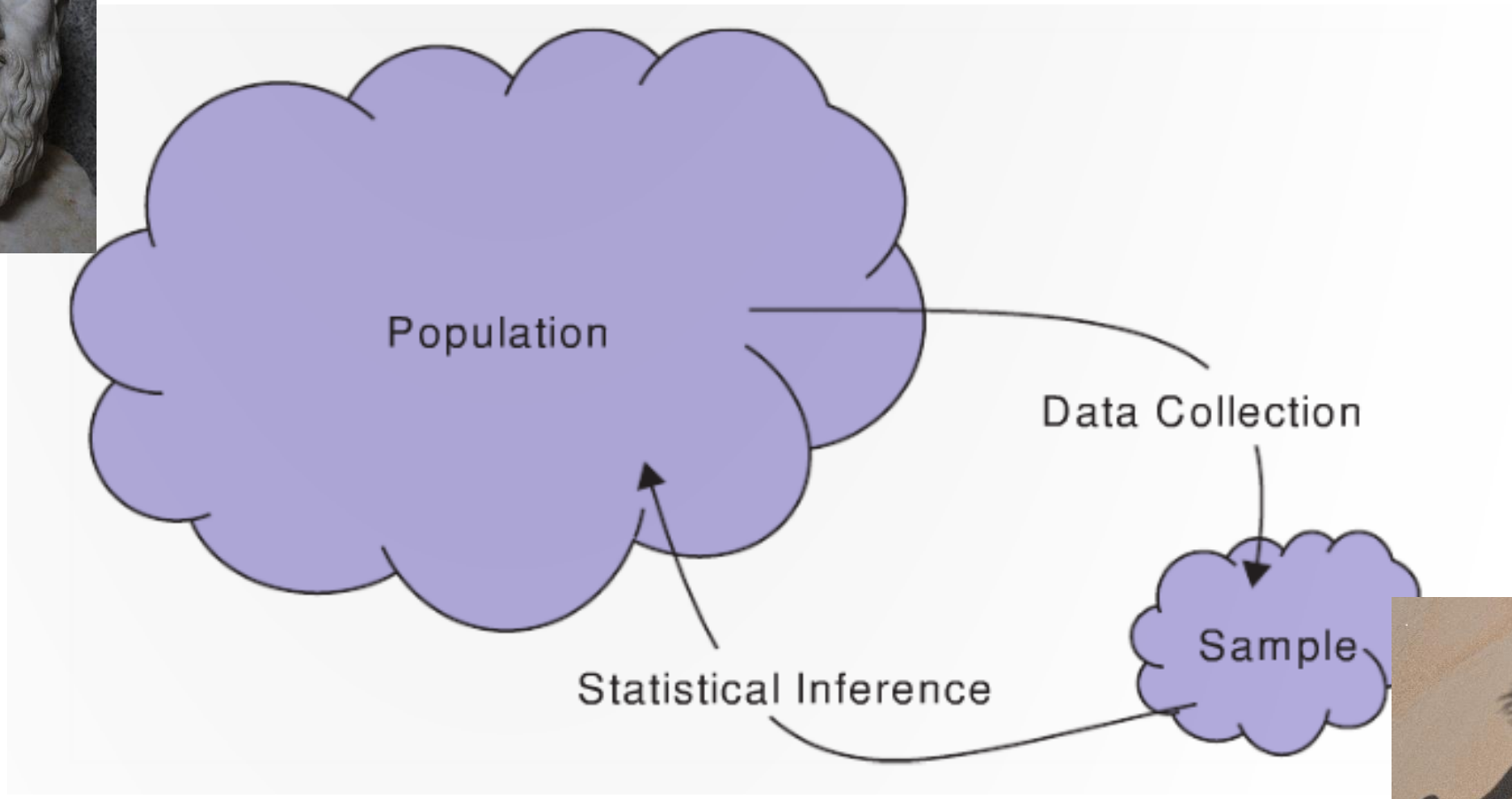
6. What is the difference between categorical and quantitative variables?

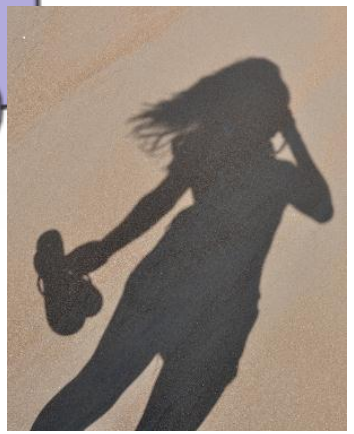


Plato

7. Who is this?







We will come back to this soon, but first…

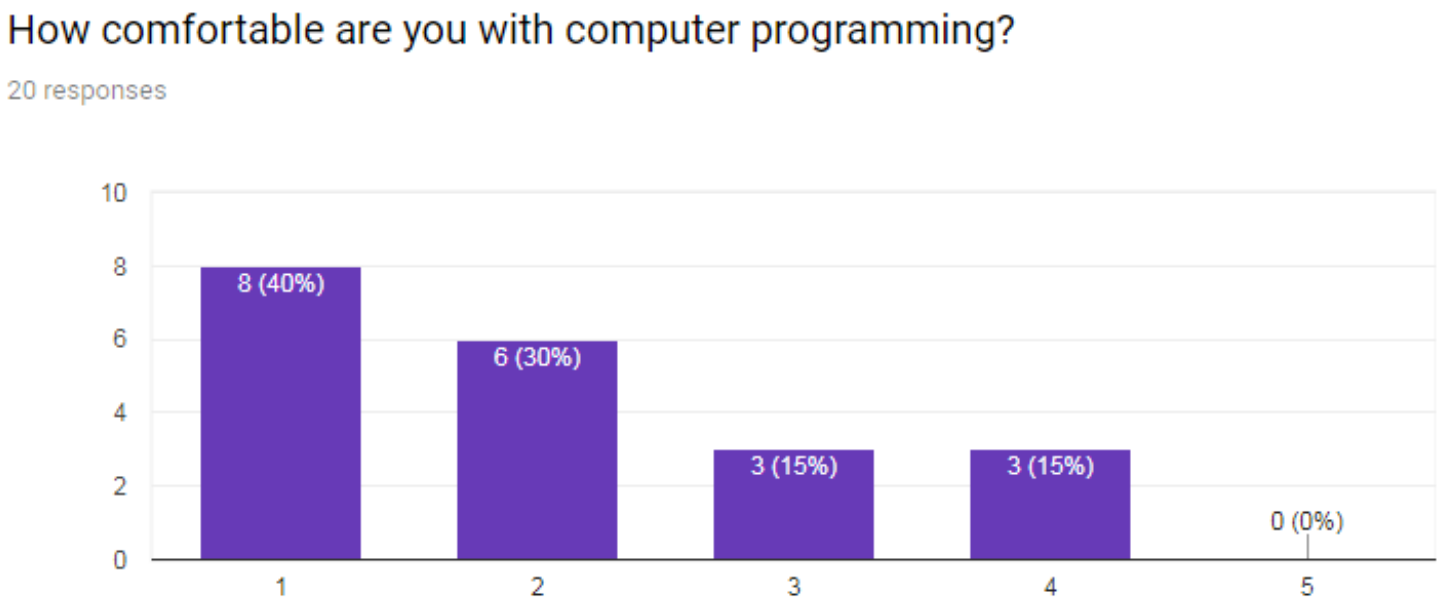
Question

Q: What was the movie, ‘Pirates of the Caribbean’ rated?
A: PG-13

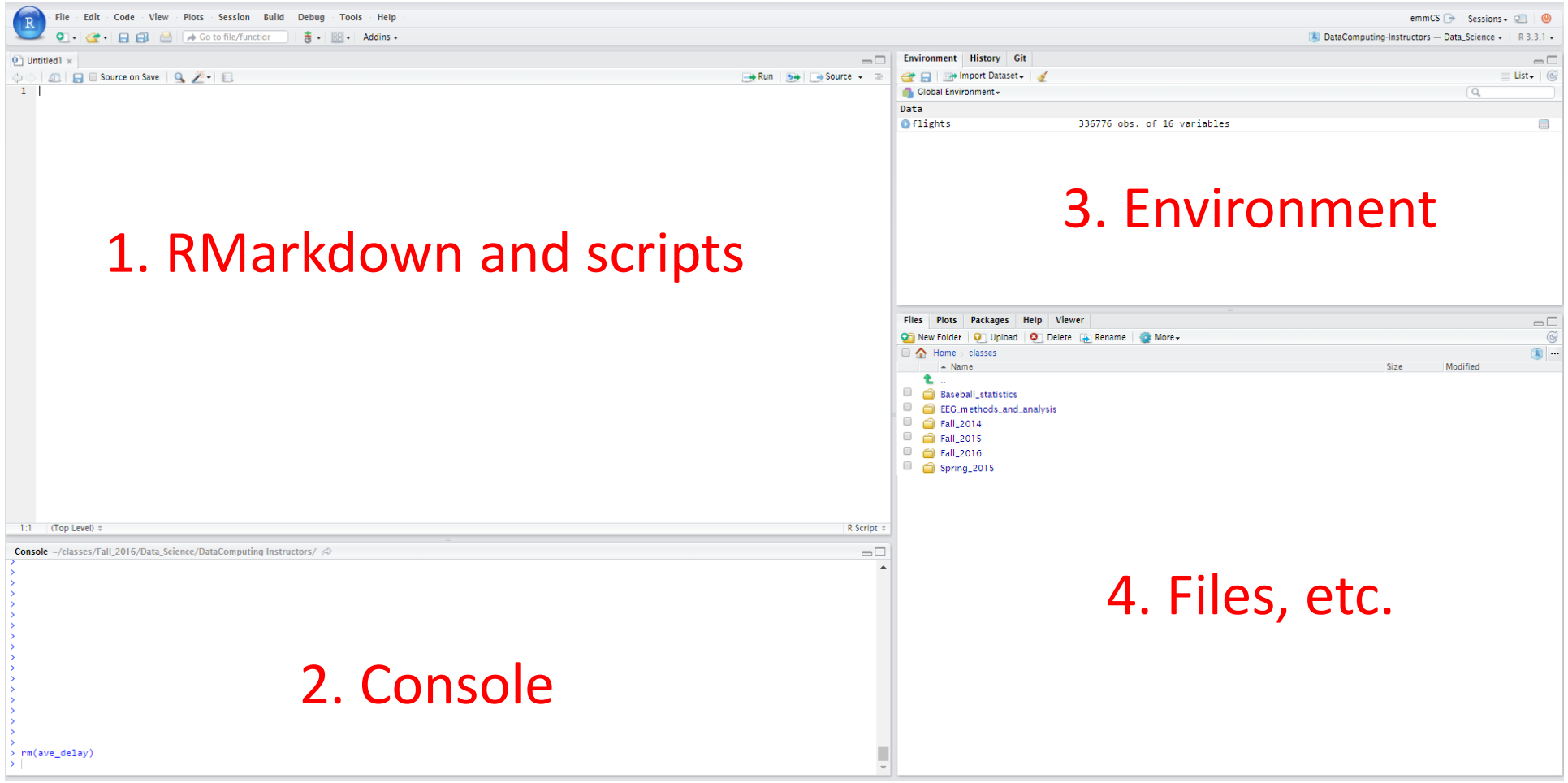
Q: Worst joke of the semester?
A: We are just getting started!

Basics of R

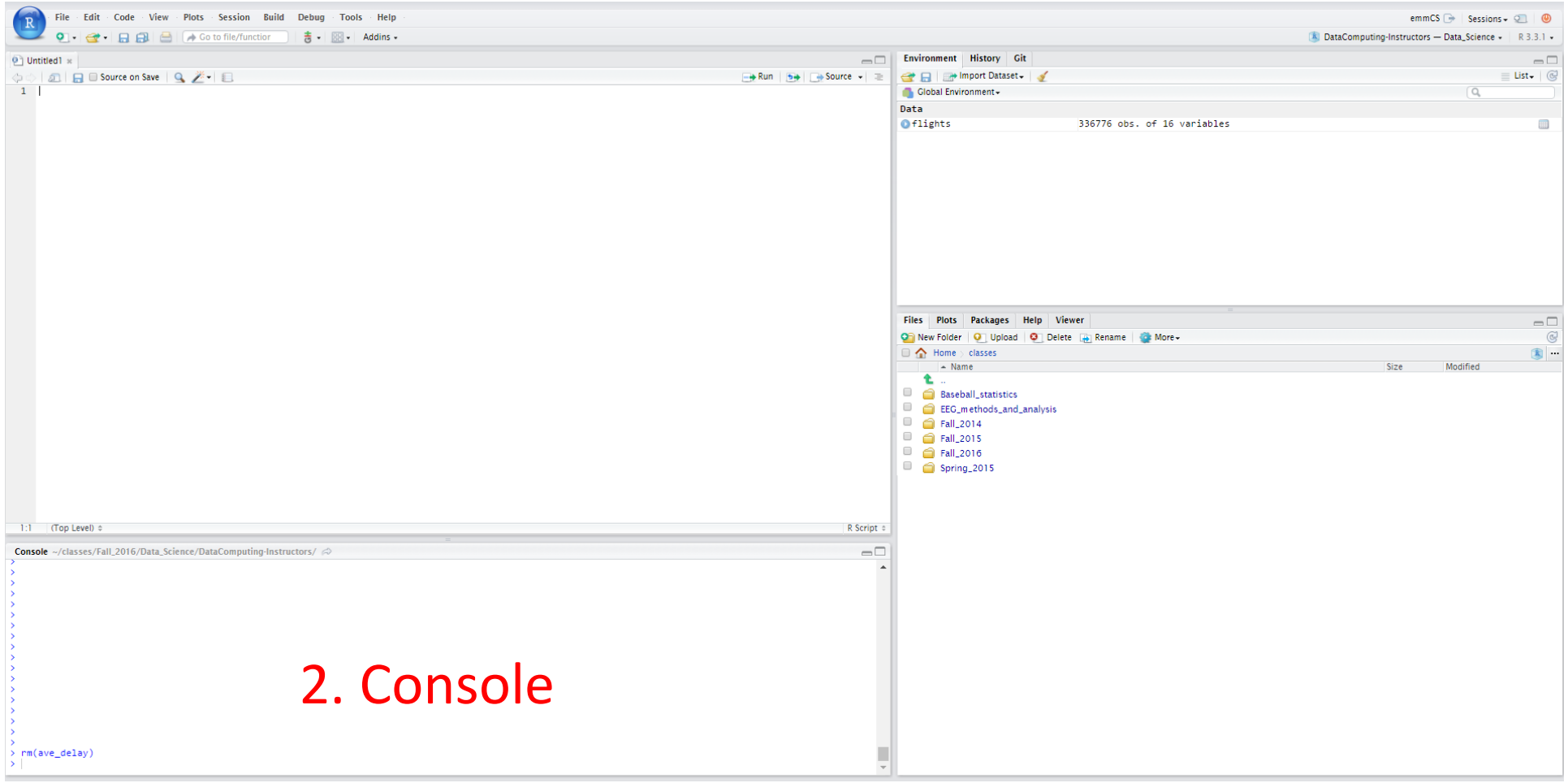
Everyone log on to: https://asterius.hampshire.edu/



RStudio layout



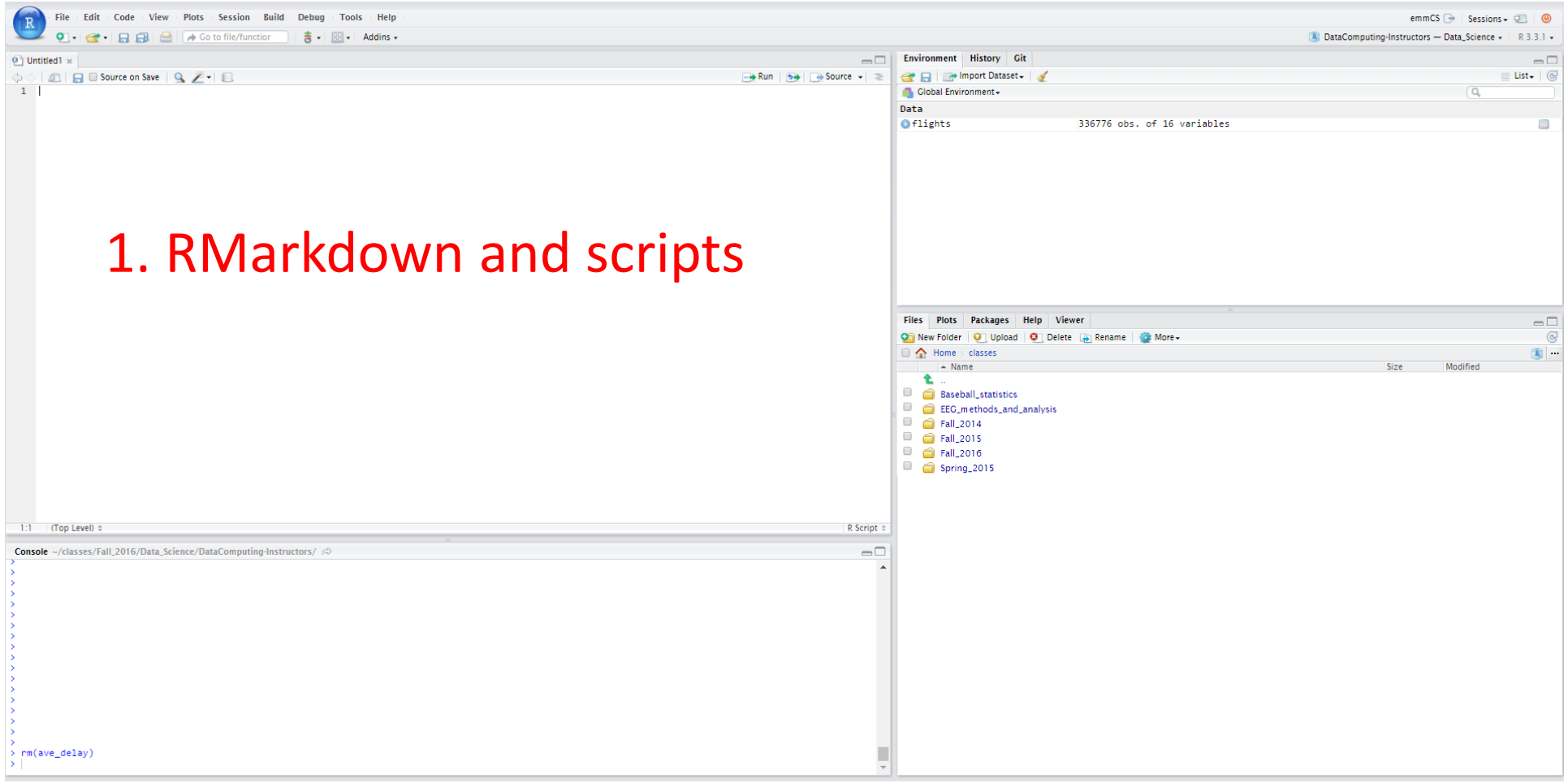
RStudio layout



R as a calculator

```
> 2 + 2
> 7 * 5
```

RStudio layout



1. RMarkdown and scripts

Create a new script

File -> New File -> R Script

Save the script with a reasonable name, e.g., notes.R

R Basics

Arithmetic:

```
> 2 + 2
> 7 * 5
```

Assignment:

```
> a <- 4
> b <- 7
> z <- a + b
> z
[1] 11
```

Number journey…

Number journey

```
> a <- 7
> b <- 52
> d <- a * b
> d
[1]  364
```

Character strings and booleans

```
> a <- 7
> s <- "hello everyone"
> b <- TRUE

> class(a)
[1] numeric

> class(s)
[1] character
```

Functions

Functions use parenthesis: functionName(x)

```
> sqrt(49)
> tolower("HELLO everyone")
```

To get help

```
> ? sqrt
```

One can add comments to your code

```
> sqrt(49)    # this takes the square root of 49
```

Question

Q: What kind of grades the pirate get in Introduction to Statistics?

A: High Seas

Q: Worst joke of the semester?

A: Not likely

Vectors

Vectors are ordered sequences of numbers or letters

The c() function is used to create vectors

```
> v <- c(5, 232, 5, 543)
> s <- c("a", "b", "c", "d")
```

One can access elements of a vector using square brackets []

```
> s[3]        # what will the answer be?
```

We can get multiple elements from a vector too

```
> s[c(1, 2)]
```

Vectors continued

One can assign a sequence of numbers to a vector

```
> z <- 2:10
> z[3]
```

One can test which elements are greater than a value

```
> z > 3
```

Can add names to vector elements

```
> names(v) <- c("first", "second", "third", "fourth")
```

Announcement

International talk like a pirate day: **September 19th**

DataCamp Intro to R chapters 1 and 2

Please signup for the CS206 DataCamp group (link also on Moodle):

https://www.datacamp.com/groups/shared_links/c00a3d77b6a18dec7c5b6844526b34dd70450542

Please then login and complete the first two chapters for the Introduction to R class:

https://www.datacamp.com/courses/free-introduction-to-r

Try to complete the first two chapters by 11:59pm on Wednesday September 12th