

Introducing R

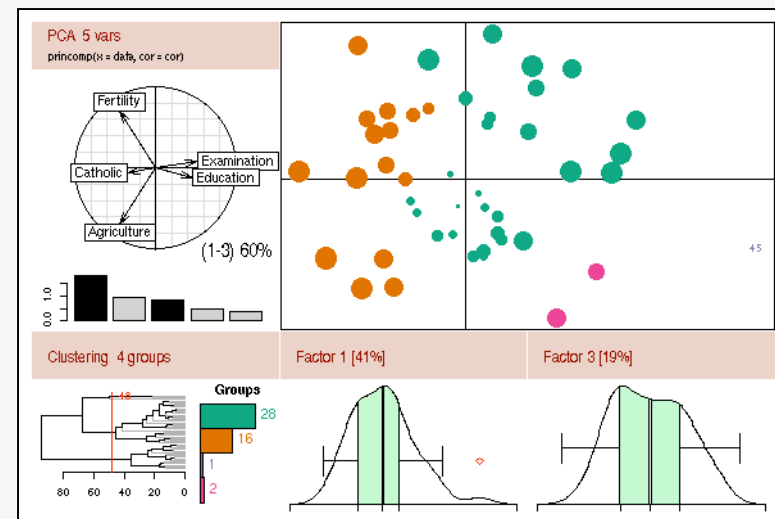
Cameron Wellock

cwellock@rockefeller.edu

Part I: preliminaries

What is R?

- A tool for statistics
- A programming language
- A plotting system



What isn't R?

- Excel
- A statistician



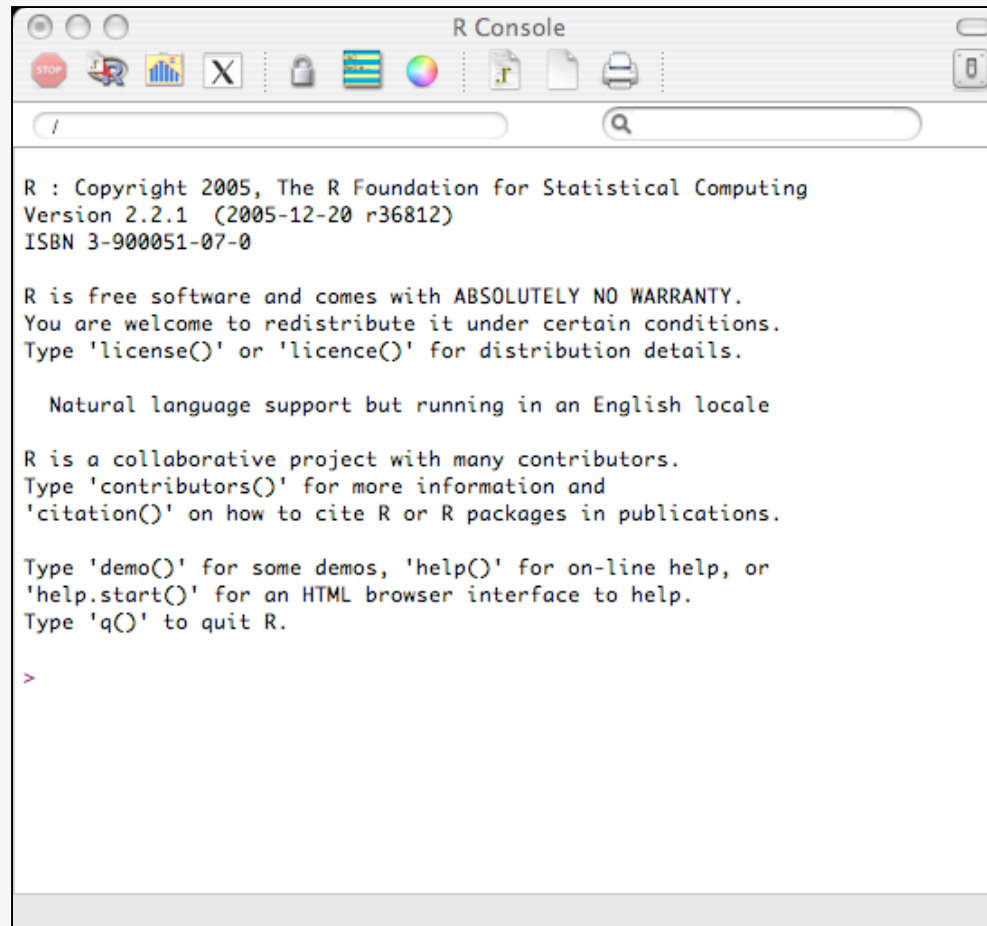
The Most Important Thing

- www.r-project.org

How to start R on OS X

- Click on the smiley icon on the dock
- Go to 'applications'
- Double click the 'R' icon

Do you get this?



```
R Console  
/   
R : Copyright 2005, The R Foundation for Statistical Computing  
Version 2.2.1 (2005-12-20 r36812)  
ISBN 3-900051-07-0  
  
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.  
  
Natural language support but running in an English locale  
  
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.  
  
>
```

Part II: get to work!

The world's ugliest calculator

- Type this:

1 + 1

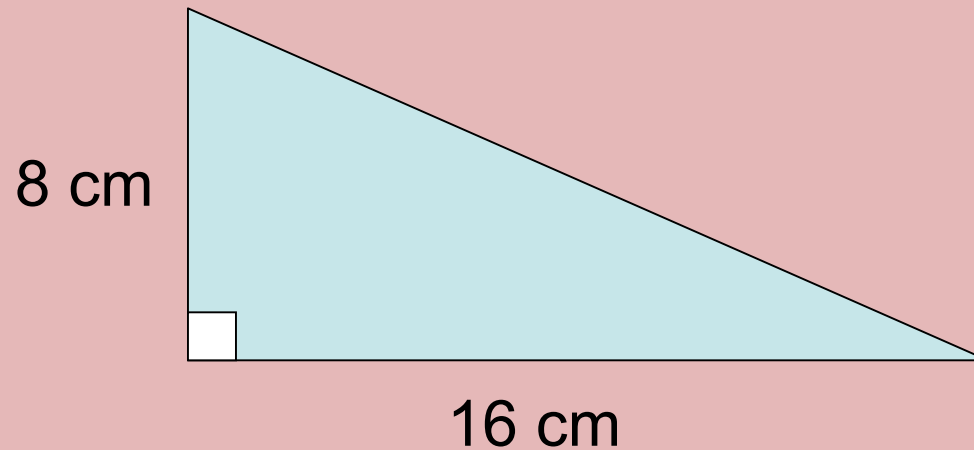
2 * 3

4e-3 / 2.5e-1

((-1 - -1) * -1 - -1) ^ -1 - 1

ACTIVITY #1

- Find the length of the hypotenuse for the following right-angle triangle:



Finding the truth

- Type this:

```
3 < 4
```

```
3 >= 4
```

```
3 == 4
```

```
3 != 4
```

Holding on to your results

- Type this:

```
X <- 3.14159 * 6^2
```

```
3.14159 * 6^2 -> X
```

```
X
```

Beyond numbers

- Strings: “like this” or ‘like this’
- NA: when you don’t have a value
- NaN: when math goes bad
- Inf: infinity

Goodbye, world

- Type this:

```
print('goodbye, world!')
```

```
log(10)
```

```
log(10, base=10)
```

```
log(10, 10)
```

Finally, statistics

- Type this:

```
X
```

```
c(1, 2, 3, 4, 5)
```

```
Y <- c(1, 2, 3, 4, 5)
```

```
mean(Y)
```

ACTIVITY #2

- Find the mean and median ages of five people, including yourself
- Find the standard deviation too

Vector arithmetic

- Type this:

```
Z <- c(6, 1, 3, 1, 1.5)
```

```
Z * 2
```

```
Z + Y
```

```
Z * Y
```

Getting parts of a vector

- Type this:

```
Z[2]
```

```
Z[2:4]
```

```
Z[-2]
```

```
Z[c(4, 2, 3, 1)]
```

A good trick

- Type this:

```
a.vector <- c(11, 45, 28, 37, 51, 19)
```

```
a.vector > 35
```

```
which(a.vector > 35)
```

```
a.vector[which(a.vector > 35)]
```

ACTIVITY #3

- Type this:

```
A <- 101:200
```

- Find a way to get every *other* item from A, starting with the first item, but you can't use:

```
A[c(1, 3, 5, 7, 9, ..., 199)]
```

Things we covered in part II

- Concepts: *variables, functions, arguments*
- Data types: *numbers, booleans, strings, vectors*
- Special values: *NA, NaN, Inf*
- Basic and vector arithmetic
- Indexing vectors

Part III: enter the data

Lists

- Type the following:

```
A <- list(cam=c(85,75,85), sindy=c(95,95,100))
```

```
A
```

```
A[[“cam”]]
```

```
A$cam
```

Data frames

- Type the following:

```
v1 <- c("billy", "suzie", "edgardo", "ishmael",  
"monica", "phoebe", "rachel")
```

```
v2 <- c(15, 19, 11, 45, 38, 37, 39)
```

```
v3 <- c(10, -20, 45, 1000, 10000, 10000, 10000)
```

```
D <- data.frame(name=v1, age=v2, net.worth=v3)
```

Fun with columns

- Type the following:

```
D$name
```

```
D$assistants <- c(0, 0, 1, 1, 3, 2, 4)
```

```
D$assistants <- c()
```

ACTIVITY #4

- Find a way to get a subset of the people in D--only those over 30 years of age.