

ST 516: Foundations of Data Analytics

Introduction to R

What is R?

How we will use R

What is R?

R is a programming language.

Since R is a complete language, you can write R code to do almost anything. For example, R can say “hello world,” make webpages, plan your next chess move, keep track of the weather, and even check your email.

R is a tool for data analysis.

R was built to do data analysis. From the start, it was built to read and manipulate data, perform statistical modelling, and make graphical displays. There is a whole community of people developing new packages to work with data.

Why should I learn a programming language for data analysis?

- Documentation: your R script becomes a recipe for how you conducted your analysis, from raw data to plots and confidence intervals.
- Communication: Your R script tells you, your future self, and your peers how you went about an analysis.
- Automation: You can reuse your R script next time you have a similar analysis to do.

Why R?

R isn't the only tool for data analysis.

Some good reasons to learn it:

- It's free \$0
- It's free - "to run, copy, distribute, study, change and improve the software."
- It has a large (and still growing) community of users and developers

RStudio and Datacamp

RStudio is our environment for running R on your own computer. You will have to install both R and RStudio.

Datacamp is an online interactive environment for learning R.

The first time you visit you will need to create an account.

You'll complete some lab exercises in datacamp, but when it comes time to doing your homework or your own analyses you will use RStudio.

A quick look

Watch this video for a quick look at both RStudio and Datacamp.

R for simulation

In this class, we will also use R to simulate random processes. For example, we can ask R to flip a coin for us:

```
sample(c("Heads", "Tails"), size = 1)
```

```
## [1] "Tails"
```

Of course, R can flip a coin thousands of times for us really quickly,

```
x <- replicate(10000, sample(c("Heads", "Tails"), size = 1))
```

And make us a table (or chart!) summarizing the flips.

```
table(x)
```

```
## x
```

```
## Heads Tails
```

```
## 4917 5083
```


Getting started

Get R and RStudio installed on your computer

Head to [datacamp](#), register, and complete first two chapters of Introduction to R