Lecture 9 - Basic Programming with R and Python

DSE 511

Drew Schmidt
2022-09-22

Announcements

- Nothing unresolved from last time
- Homework is live!
- Questions?

Content

- The Course So Far

The Course So Far

Where We've Been

Module 1: Introduction

- Lecture 1 - Course Introduction
- Lecture 2 - Introduction to VMs
- Lecture 3 - CANCELED 🏈

Where We've Been

Module 2: Version Control

- Lecture 4 - Introduction to Version Control
- Lecture 5 - Basic git
- Lecture 6 - Working with Remotes
- Lecture 7 - Collaborating on GitHub
- Lecture 8 - When Things Go Wrong

Where We're Headed

Module 3: Basic Programming with R and Python

- Lecture 9 - Introduction to R and Python
- Lecture 10 - Basic Programming
- Lecture 11 - Data Structures (Part 1)
- Lecture 12 - Data Structures (Part 2)
- Lecture 13 - Application

History of R and Python

What is R?

- *Lingua franca* for statistical computing
- Part programming language, part data analysis package
- Dialect of S (May 5, 1976, Bell Labs)
- Free software (GPL >= 2)

The R Language

- R is slow; if you don't know what you're doing, it's \emph{really} slow.
- High-level scripting language.
- Syntax designed for data: models are first-class constructs, missingness is built into the core of the language, \dots

Programming and Data Science

Why Do We Need This?

- Most of your problems *are not new problems*
- There are **known** good and bad ways to do almost anything you can think of

Thought Experiment

- I have two files
- How do I tell if they're "the same"?

Thought Experiment

- I have streaming data
- How do I process it?

What We Will and Won't Do

What's not the goal?

- Turn you into a computer scientist
- Make you into hardware experts
- Give you tons of algorithm puzzles

What's the goal?

-

Wrapup

Wrapup

-

Questions?