A Fully Customizable Textbook for Introductory Statistics/Data Science Courses

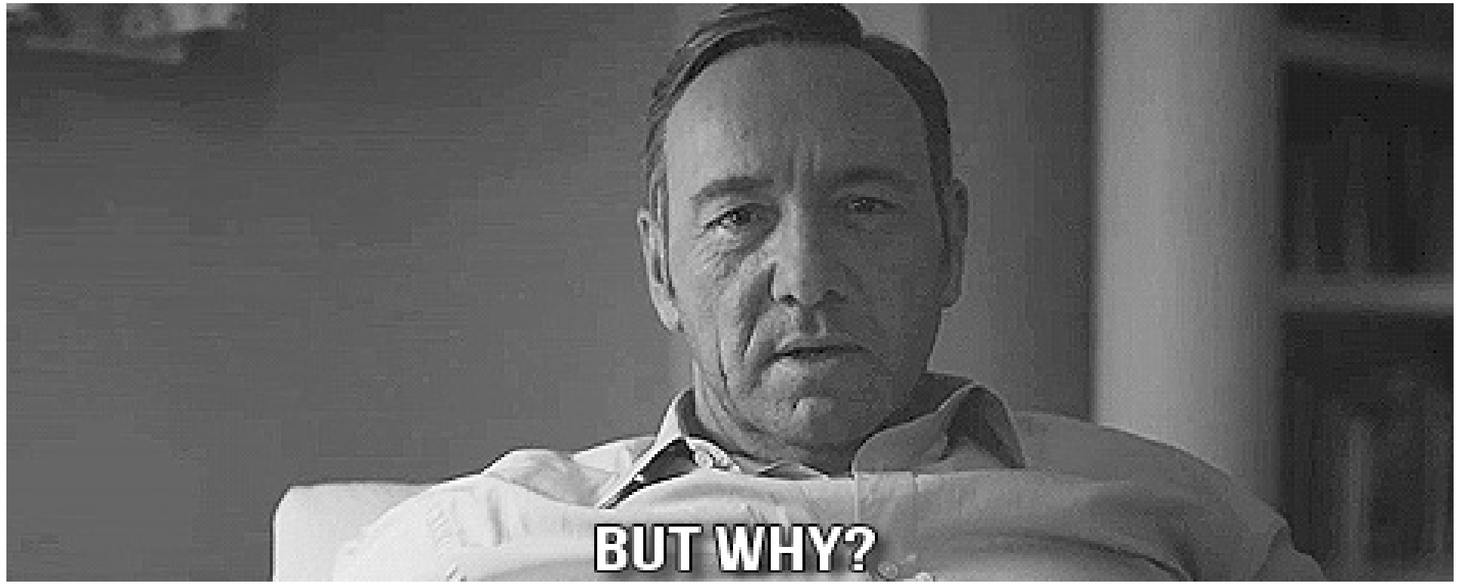
Chester Ismay - Reed College/Pacific University
Albert Y. Kim - Middlebury College

Email: chester@moderndive.com, albert@moderndive.com

2017/03/24

Slides available at http://bit.ly/moderndive-causeweb

Why is this needed?



Guiding Principles of ModernDive

---

1. Blur the lines between lecture and lab

While in lab section...

Then and Now

- Segregated lecture/lab is a legacy of when:
  - Desktops reigned
  - Proprietary statistical software was usually the best/only option

Then and Now

- Segregated lecture/lab is a legacy of when:
  - Desktops reigned
  - Proprietary statistical software was usually the best/only option
- Today
  - Almost all students have access to laptops
  - Open source software options are more palatable

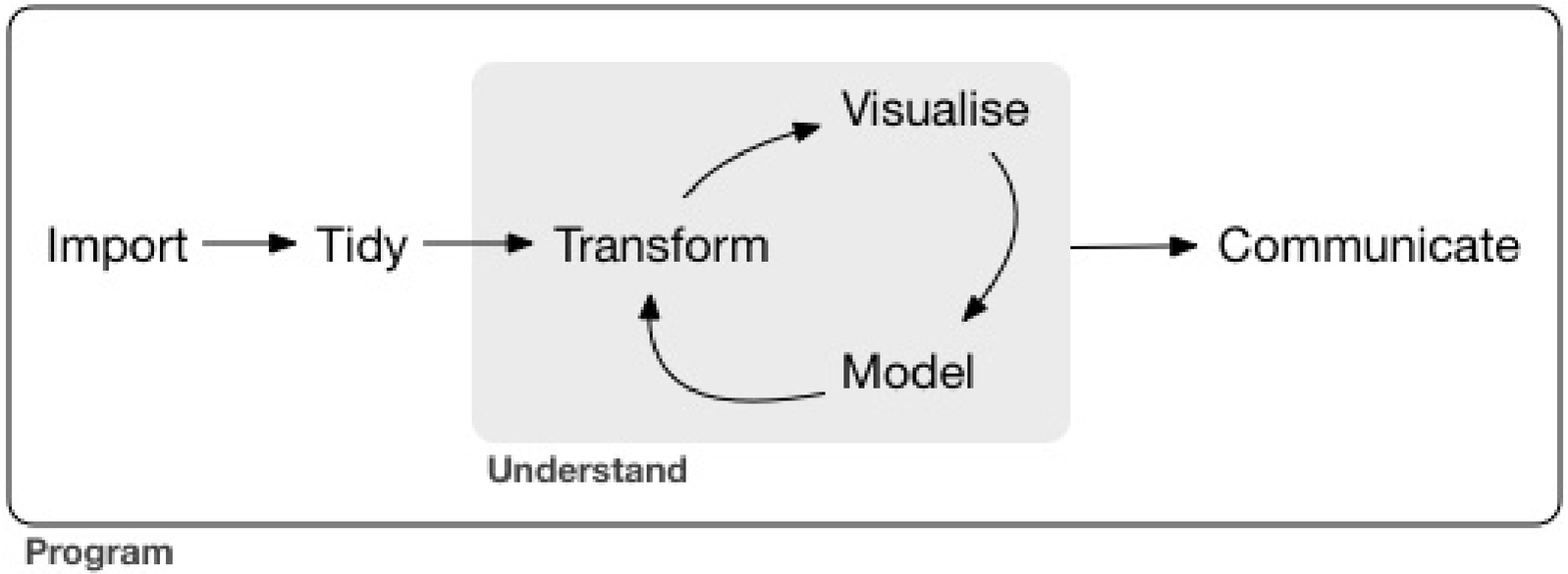
A new classroom environment

Working like data scientists/statisticians work

Guiding Principles of ModernDive

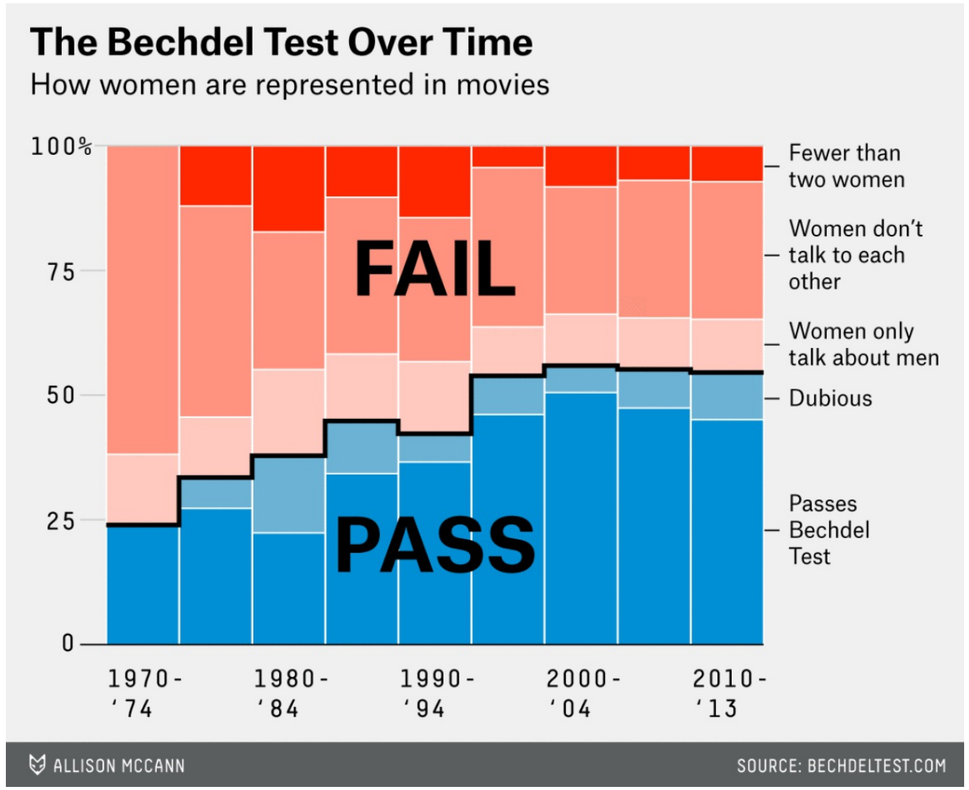
---

2. Focus on the data/science research pipeline



Creating effective data stories is the key

- Each topic builds on previous topics towards improving communication using data

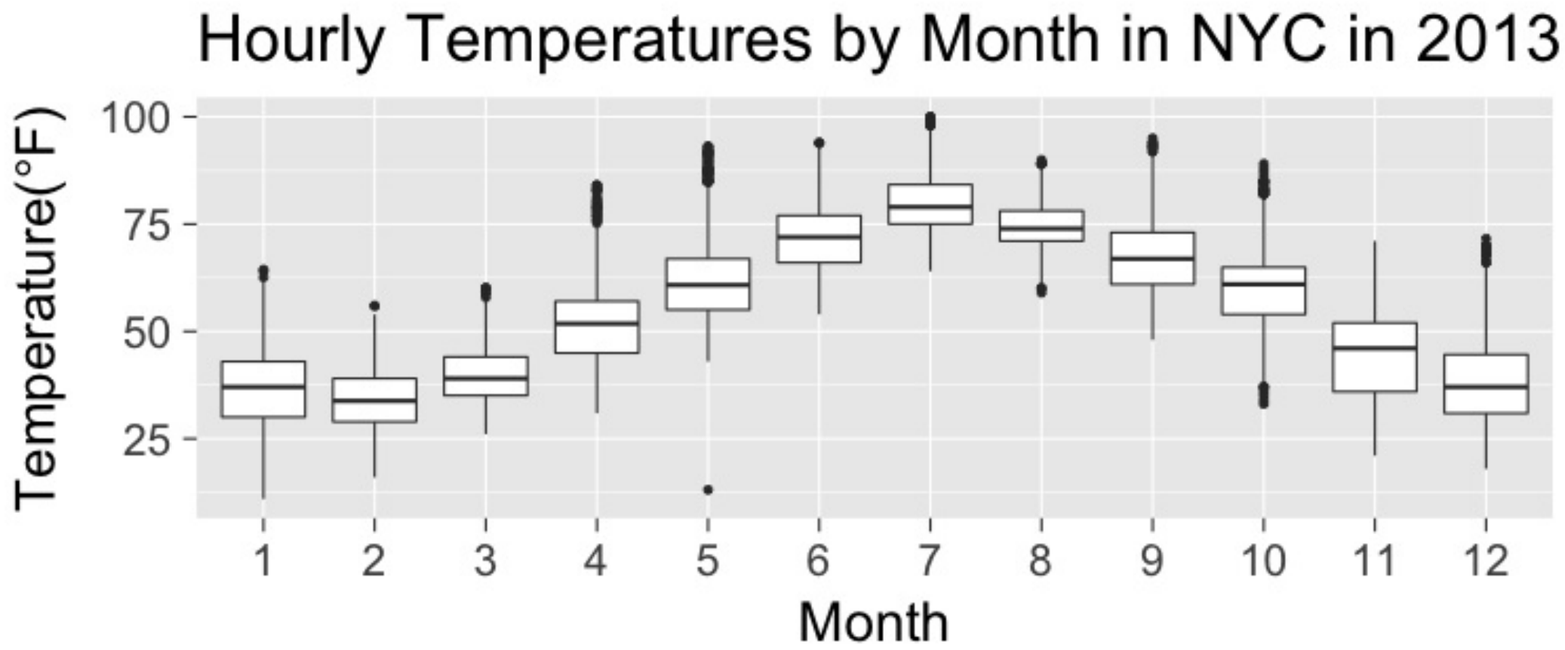


Guiding Principles of ModernDive

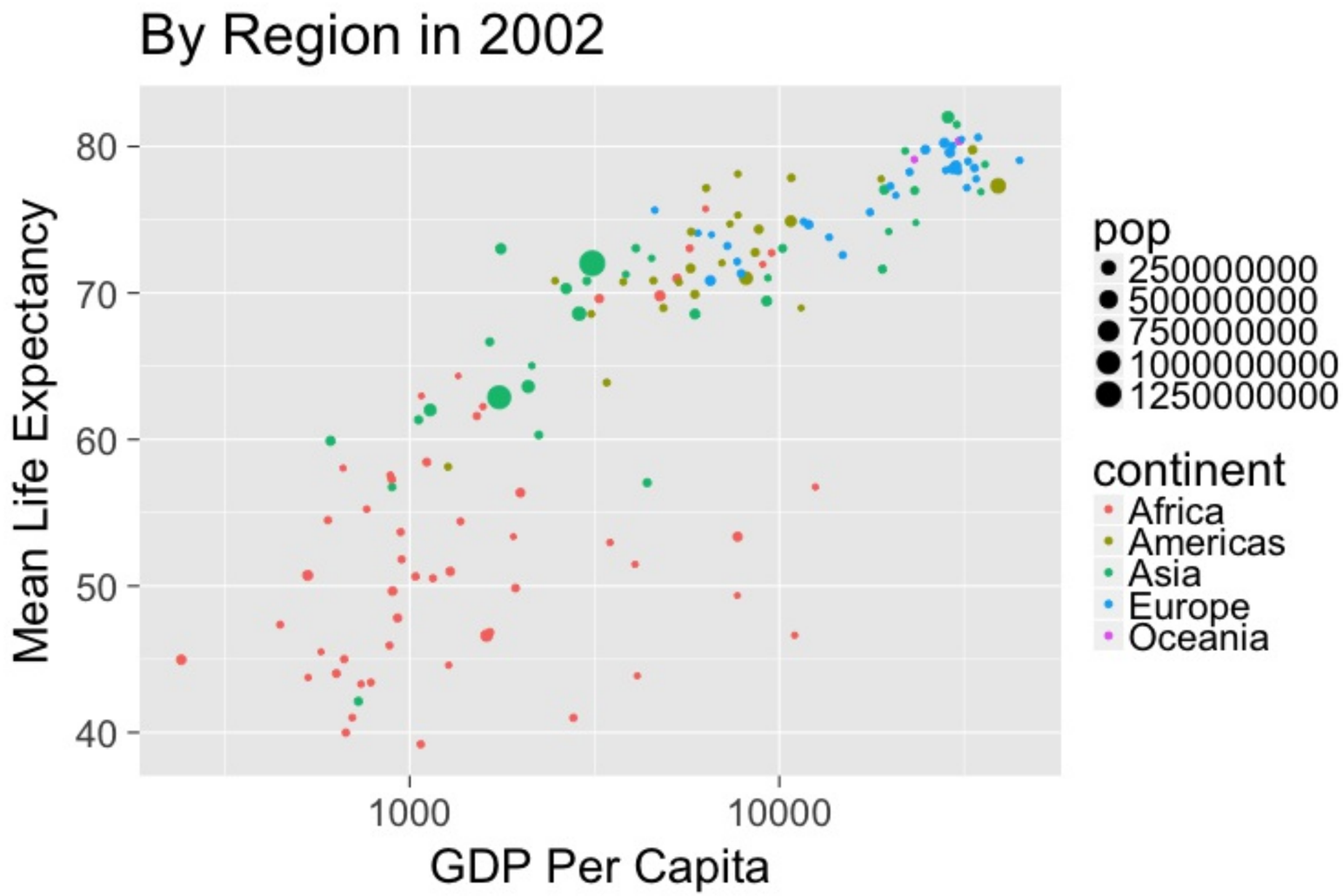
---

3. It's all about the data

- Use modern R packages with rich, interesting, open data



Have data visualization be the driver

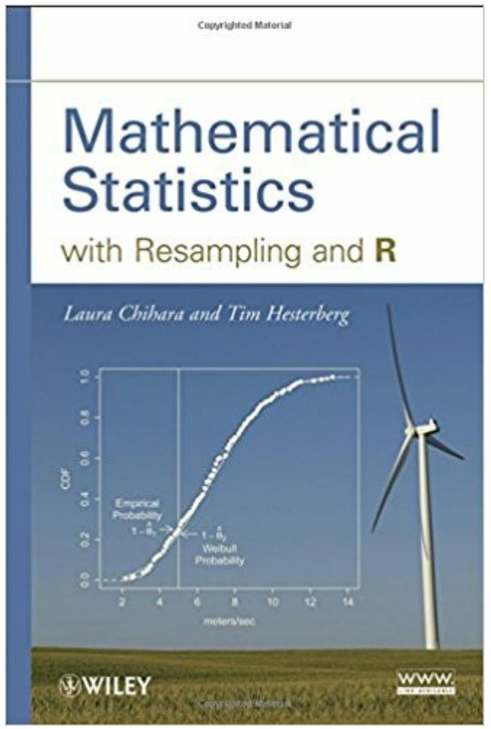


Guiding Principles of ModernDive

---

4. Use simulation/resampling instead of probability

From Albert's 300-level ~~Mathematical Statistics~~ Theory of Statistics:

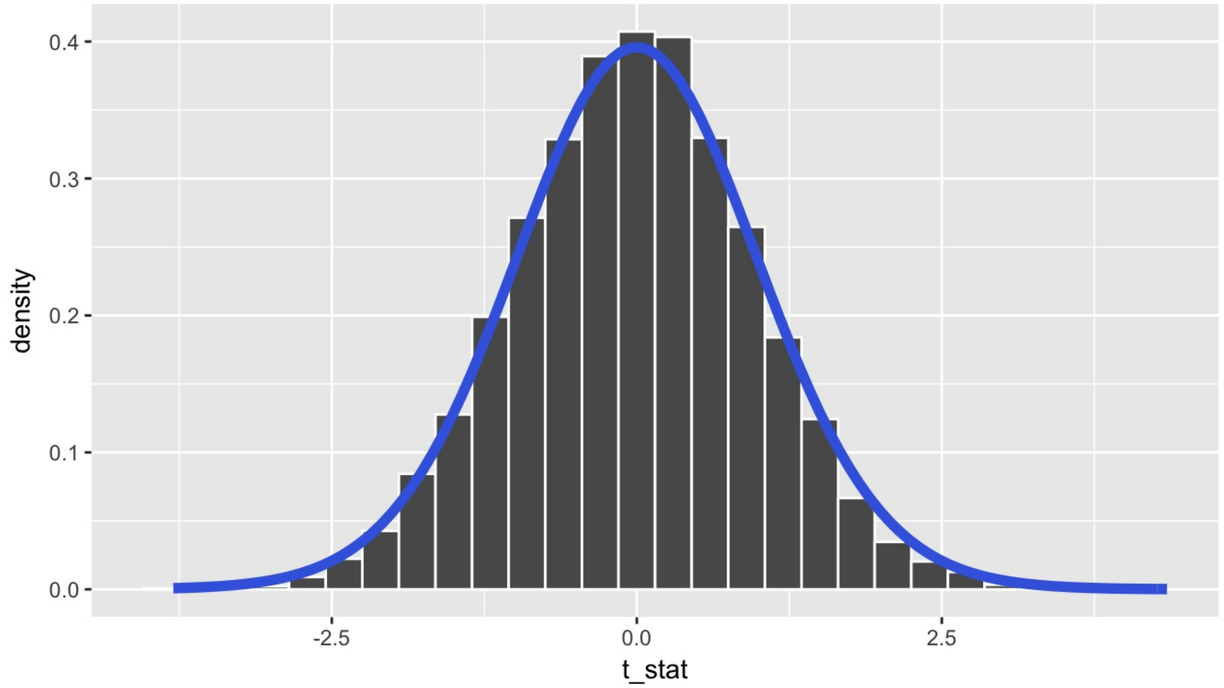


Reinforce concepts instead of equations, formulas, and probability tables

- Build the Central Limit Theorem using computation
- Check for assumptions needed for traditional inference using algorithmic thinking

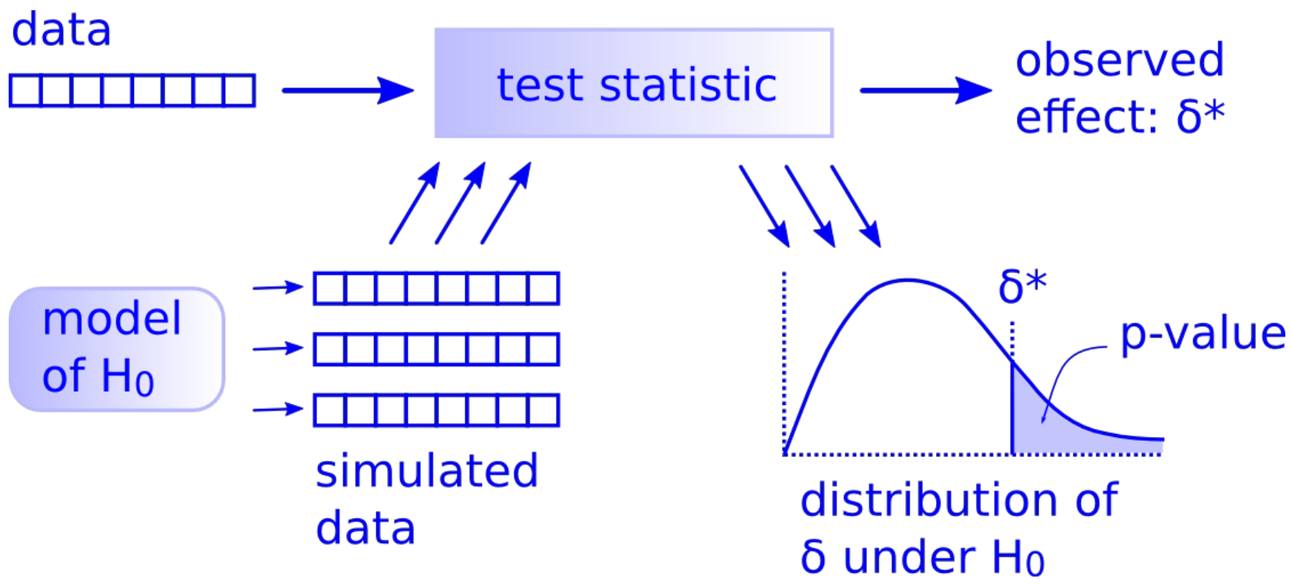
Reinforce concepts instead of equations, formulas, and probability tables

- Build the Central Limit Theorem using computation
- Check for assumptions needed for traditional inference using algorithmic thinking



Inspiration and common theme

There is only ONE (hypothesis) test!

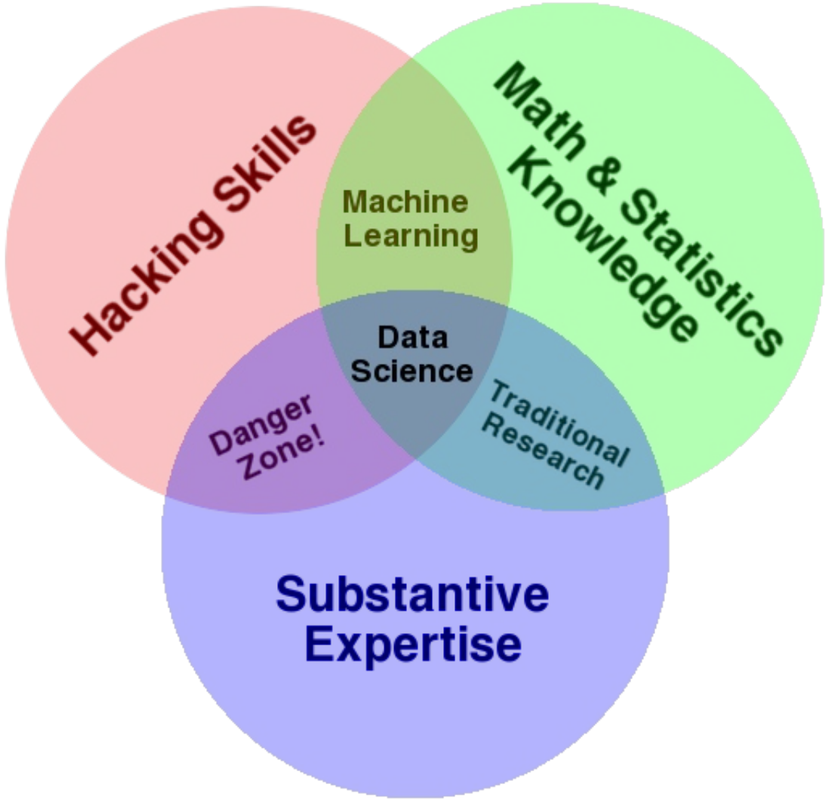


5. Don’t fence off students from the computation pool, throw them in!

- Scaffold & support as a good foreign language professor would

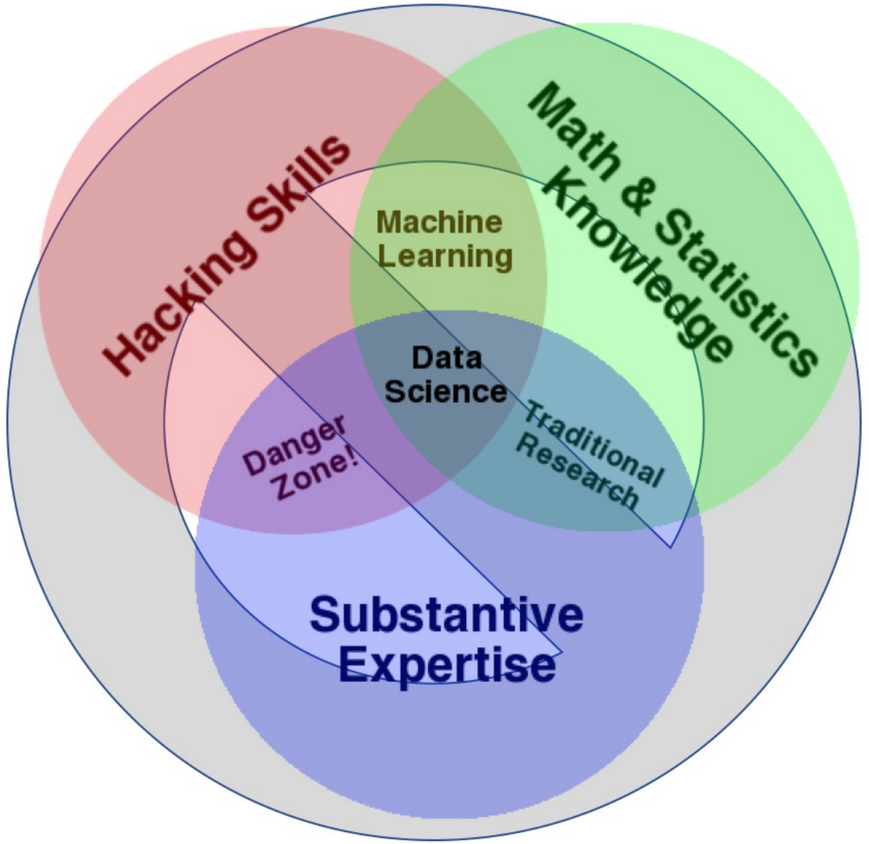
Developing a needed skill

- Coding will be a basic skill on par with reading and writing.



Developing a needed skill

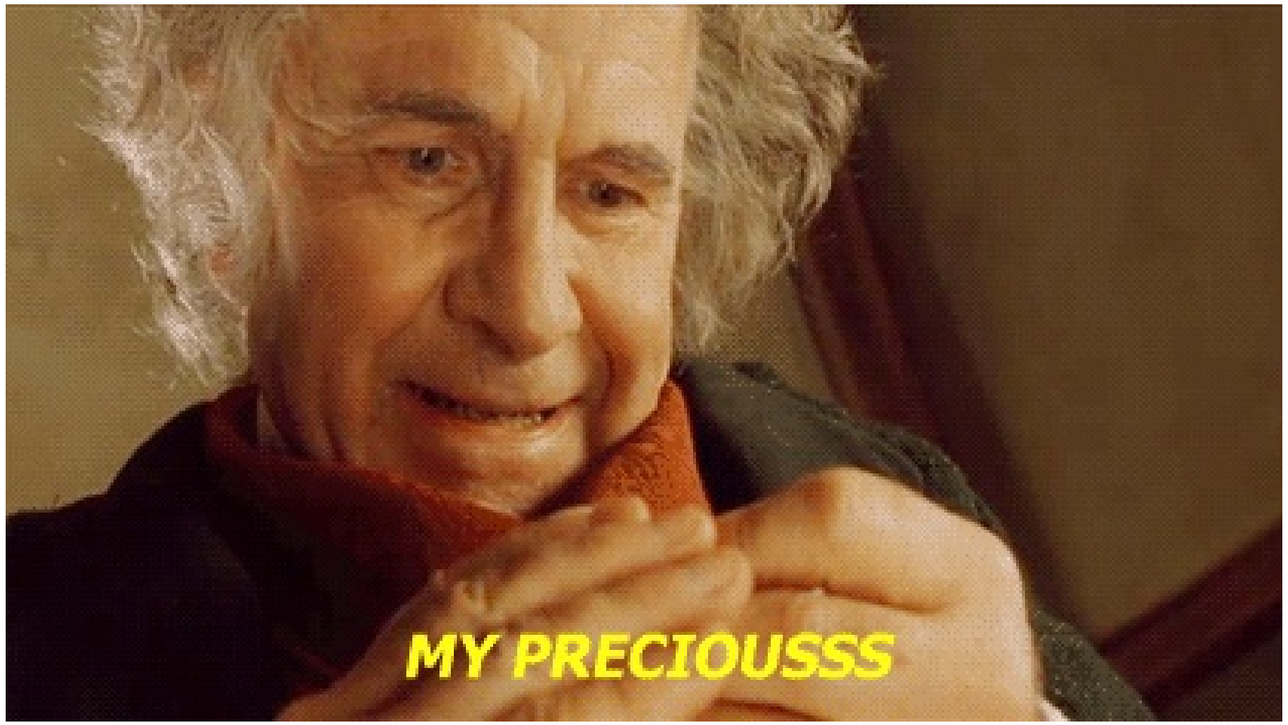
- Coding will be a basic skill on par with reading and writing.



6. Complete reproducibility with bookdown

- Put it all out there

- Ultimately the best textbook is one you've written yourself



The bookdown R package

- Write an entire book using R and Markdown
- Rapid iteration and easily-updateable
- Exports book to multiple formats
- Slick cross-references
- Textbook has versions not editions
- Wikipedia model for intro stats/data science

- A bookdown book about writing with bookdown

ModernDive

An Introduction to Statistical and Data Sciences via R

Authors: Chester Ismay, Albert Y. Kim and you?

ModernDive.com

OR

ModernDive.org

Supplementary materials to the textbook

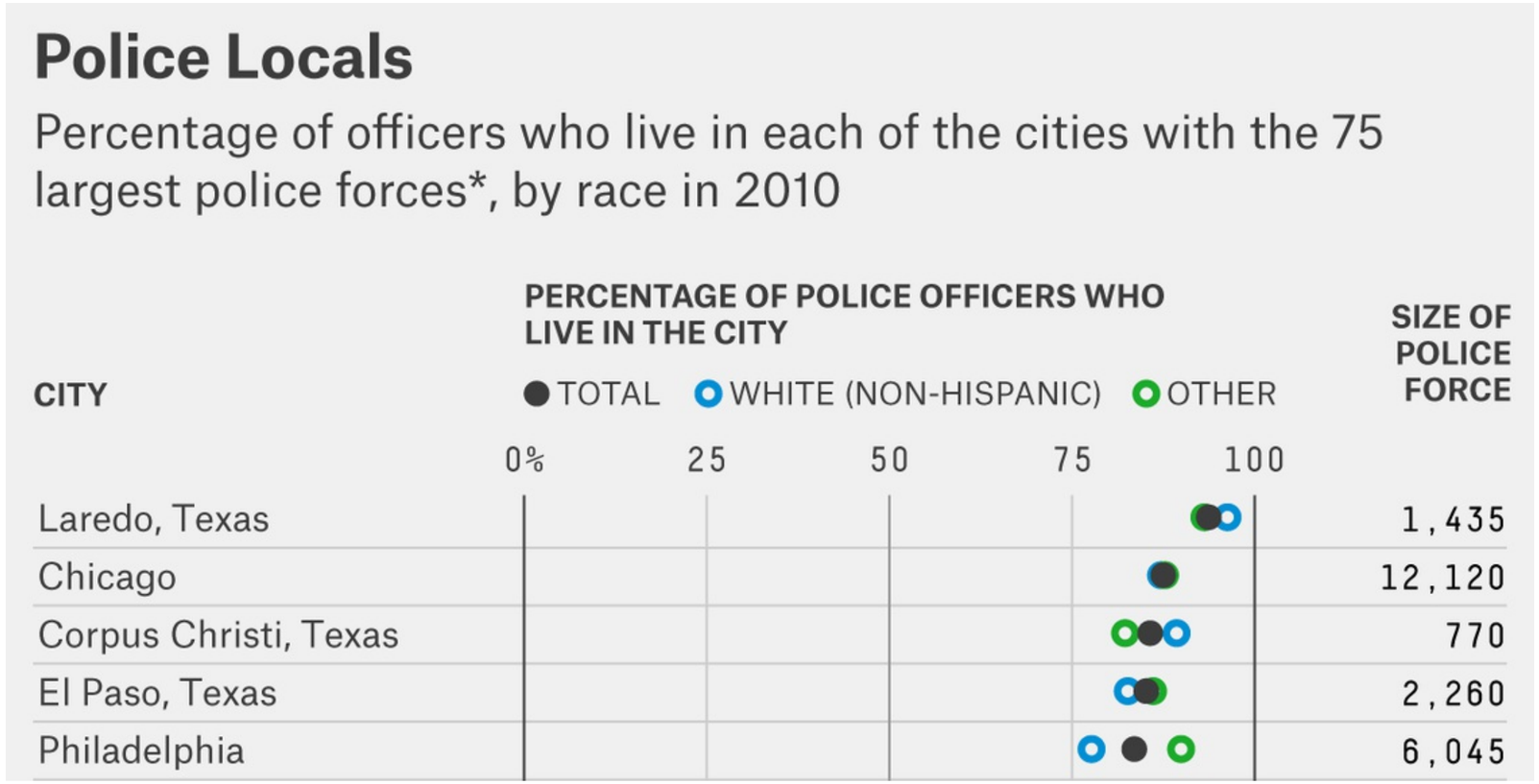
The `fivethirtyeight` R package

- Data sets that balance being
  - rich enough to answer meaningful questions with,
  - real enough to ensure that there is context, and
  - realistic enough to convey to students that data as it exists "in the wild" often needs processing.

The `fivethirtyeight` R package

- Data sets that balance being
  - rich enough to answer meaningful questions with,
  - real enough to ensure that there is context, and
  - realistic enough to convey to students that data as it exists "in the wild" often needs processing.
- Easily and quickly accessible to novices, so that we minimize the prerequisites to research.

The `fivethirtyeight` R package



DataCamp course

FREE COURSE

Effective Data Storytelling using the tidyverse

Continue Course

0 Videos · 30 Exercises · 15 hours · 30 Participants · 2550 XP

Course Description

This course is designed to supplement and build on the content covered at http://moderndive.com . It assumes that you have completed the Introduction to R course on DataCamp at https://www.datacamp.com/courses/free-introduction-to-r .

1 Applying R Basics

20%

Exploring the basics of R on a data set based on the fivethirtyeight.com "Most Police Don't Live In The Cities They Serve" article.

Investigate the properties of a data frame — 70 xp

Instructor(s):

Chester Ismay

Albert Y. Kim

Chester's Social Statistics course webpage

Schedule

The references to **Chapters/Sections** here correspond to the **MODERN DIVE into Data with R** book. Be sure to check the **DataCamp** link above for more details on the DataCamp (DC) assignments. More details on the Problem Sets (PS) are available in the link above for Problem Sets.

Search:

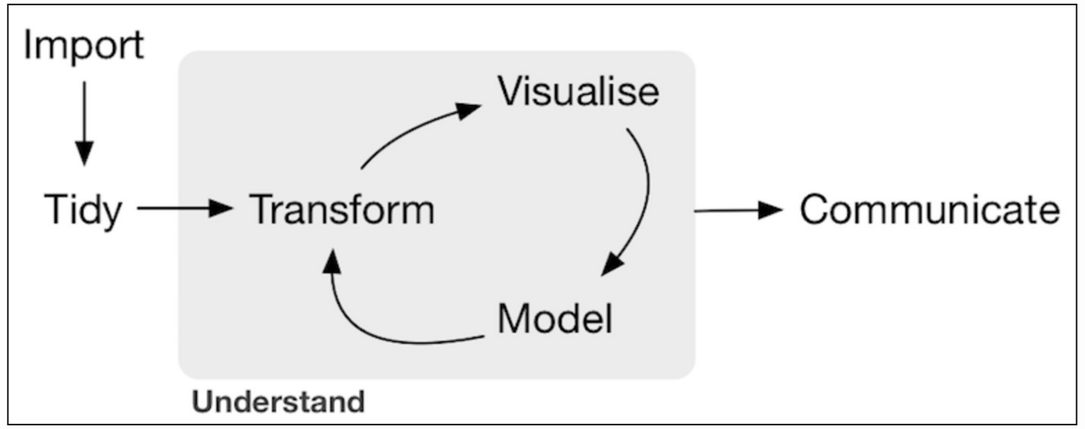
Weekday	Date	Content / Assessment	Material Due
All	All	All	All
Tuesday	January 31	Introduction (Chapters 1 and 2)	-
Thursday	February 2	Introduction to R on DataCamp	-
Tuesday	February 7	Review of Introduction to R on DataCamp	DC course: Introduction to R

Albert's Intro to Stat & Data Sciences course webpage

Intro to Stat & Data Sciences

Slack RStudio Server Problem Sets Syllabus

Fork me on GitHub



Topics

- Slides. Also in HTML format.
- Learning checks.
- ModernDive textbook. Feedback form here.

What's to come

- Add more interactive shiny apps into the book
- Create more Review Questions at chapter ends using `fivethirtyeight` and other open data sources
- Design and share instructor resources
- Finish DataCamp course to supplement and assist with more immediate feedback

ModernDive.com

- Join us for a workshop with many more details at USCOTS at Penn State on May 17-18
- Slides created via the R package xaringan by Yihui Xie.
- Source code for ModernDive
  - Feel free to modify the book as you wish for your own needs! Just please list the authors as "Chester Ismay, Albert Y. Kim, and YOU!"
- Fill out form to receive updates using MailChimp
- Email us at chester@moderndive.com or albert@moderndive.com
- Follow us on Twitter: @old_man_chester, @rudeboybert

These slides available at http://bit.ly/moderndive-causeweb
Source code for slides at https://github.com/ismayc/causeweb2017