Welcome to the Workshop!

- Find a seat with a nametag, and plug in!
  - WiFi instructions are on the white board
- Make sure you've completed the setup
  - https://uw-madison-aci.github.io/2018-08-27-uwmadison-dc/#setup
- And the pre-workshop survey
  - https://www.surveymonkey.com/r/dcpreworkshopassessment?workshop_id=2018-08-27-uwmadison-dc
- Say 'Hi' to the other people at your table

Welcome to

@

https://uw-madison-aci.github.io/2018-08-27-uwmadison-dc/

Instructor/Helper Introductions

Workshop Content

Data Problems*

- "Where did this data file come from?!"
- "Can you repeat the analysis, omitting subject X?"
- "You used the wrong data- the full dataset has more subjects."
- "The color-coding in this spreadsheet was off."
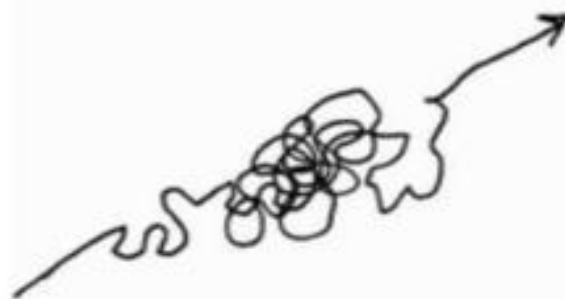
Our Goal: Reproducibility

- Learn data analysis tools **and** best practices for organizing your data
- Two themes:
  - Plan ahead: take time to organize and document
  - Avoid doing things "by hand": use tools that parse and analyze data, rather than editing data directly
- Take away something immediately useful
  - Pick one thing you learned to implement right away

Success
what people think
it looks like

Success
what it really
looks like

What We'll Learn

- Data Organization and Formatting in Spreadsheets
- Data Cleaning/Reformatting
- (Meta: Tips for Data Entry/Creation)
- Data Management using a Database
- Automated Data Analysis
- Automated Data Visualization and Reporting

Tools We'll Use

- Spreadsheet Program (Excel, or other)
- Open Refine
- SQL
- R w/packages:
  - dplyr
  - ggplot2
  - knitr

Workshop Logistics

Workshop Format

- Hands-on, working with real data.
- Be prepared to:
  - follow along with the instructor on your own computer
  - work on exercises by yourself or with people near you
- Ask questions!
- Share notes in the etherpad:
  - https://pad.carpentries.org/2018-08-27-uwmadison-dc

Nametags

- Your nametags are a visual feedback tool for instructors

- If you are lost, confused, or need help, put the red side up.

- If all is well, put up the green side up.

Code of Conduct

- In order to facilitate the best learning for everyone, we expect everyone to abide by the Data Carpentry Code of Conduct: http://www.datacarpentry.org/code-of-conduct/

  If you have any concerns, email Christina at ckoch5@wisc.edu or Sarah at sarah.stevens@wisc.edu, or contact the Carpentry staff listed on the CoC above.

Other Logistics

- We will have breaks during the morning, afternoon and for lunch.

- Coffee, tea, and morning snacks are provided, lunch is on your own.

- We can provide a refrigerator / microwave for food.

- Restrooms (and drinking fountains) are across the hall or on the first floor.

Meet your neighbors!

- Talk to 1-2 people around you by:
  - summarizing your research/work in three words (and then briefly describe it)
  - sharing one thing you hope to learn at this workshop

Important Links

- Workshop website (schedule, materials):

https://uw-madison-aci.github.io/2018-08-27-uwmadison-dc/

- Etherpad (links, online notes + live chat):

https://pad.carpentries.org/2018-08-27-uwmadison-dc

- Twitter:
  @thecarpentries, #dcuwm

Portal Dataset

Creatures!

Intro to the Portal Dataset

- Tabular dataset

- Observations of small mammals

- Collected over more than 40 years

- 24 plots of a desert ecosystem in Arizona

- Different experimental manipulations controlling which rodents are allowed to access which plot

- Real dataset, but simplified

Questions

- Is there a difference between the weight of male and female kangaroo rats?
- Do different species of kangaroo rats have different hindfoot lengths?
- How does hindfoot length relate to weight the weight of a kangaroo rat?
- How does the number of Kangaroo rats that get into the enclosure change over time?
- Is there a difference in the number of kangaroo rats that get into the different enclosures?