Introduction

Animals and exam stress

Question: Have you ever been to an event designed to help students reduce exam stress by interacting with animals (dog petting, goat yoga, etc.)?

Do dogs help exam stress?

- Data collected on 284 students at a mid-size Canadian university
- Students randomly assigned to one of three treatment groups: handler-only contact (control), indirect contact, and direct contact
- Well-being and ill-being measures recorded before and after treatment for each student
- Approach: compare pre/post measures of well-being and ill-being

Recording well-being and ill-being measures

- Likert items for each well-being / ill-being measure
- Average the likert items to get a score for each measure
- E.g.:
  - Positive affect score is the average of 5 Likert items
  - Social connectedness is the average of 20 Likert items
  - Happiness is the average of 4 Likert items

Example Likert item for social connectedness

“I am able to relate to my peers.”

- Strongly disagree (1)
- Disagree
- Somewhat disagree
- Somewhat agree
- Agree
- Strongly agree (6)

The raw data

- 284 rows (one per student)
- 200+ columns
- Example: social connectedness

	SC1_1	SC1_2	SC1_3	SC1_4	SC1_5	SC1_6	SC1_7	SC1_8	SC1_9	SC1_10	SC1_11	SC1_12
1	5	4	5	5	5	2	1	6	5	6	5	2
2	5	6	2	4	6	3	1	5	1	4	2	6
3	3	4	2	3	4	4	3	3	2	5	3	5
4	4	6	2	5	5	2	2	4	4	5	5	5
5	1	5	2	4	3	2	5	2	5	4	4	5
6	3	6	4	5	4	1	3	5	2	5	2	

Activity

- Need to process the raw data before it can be used to answer the research question
- Work with your neighbor to brainstorm data processing steps, then we will discuss as a class
- I will collect the handout at the end of class

Activity

What do you want the *final* data to look like, to make it easy to answer the research question (do students who interact more with dogs see a greater reduction in stress and a greater improvement in well-being?)

Activity

What are some of the steps that need to be done to turn the raw data into the data you *want*?

Activity

Are you familiar with any tools (e.g. R functions) that would allow you to carry out the data processing steps?

Activity

The original data from the Binfet *et al.* paper actually includes a *lot* more demographic information on the students (their ethnicity, gender identity, etc.). Why do you think I removed these variables before sharing the data with you?

Class plan

Unit 1: data wrangling

- tools for cleaning and processing data (modifying columns, creating new columns, subsetting rows, etc)
- working with different variable types
- reshaping data
- wrangling with multiple tables
- iteration across columns

Course goals

- Develop computing skills to work with data and answer statistical questions
- Emphasize reproducibility and good coding practices
- Introduce other important computing tools for statistics and data science (Python, SQL, Git)

What this course isn't:

- An exhaustive list of R or Python functions
- A computer science course
- A deep dive into how R actually works

Expectations

- Complete any assigned reading ahead of class
- Become comfortable finding and using unfamiliar functions on your own
- Bring laptop each day
- Submit class activities and occasional concept quizzes (graded for effort, not completeness or correctness)
- Attend department seminars (more info to follow)

Course components

- Class participation (graded for effort)
- HW assignments (roughly one per week)
- Exams (2 midterms, 1 final)
- Project