

Data Science

CS301

Let's Have A Group Project!

Activity 07

Week 10

Fall 2024

Oliver BONHAM-CARTER

Are you here today?!

ATTENDANCE

<https://forms.gle/iaY7zBmxj8KvsDMa8>

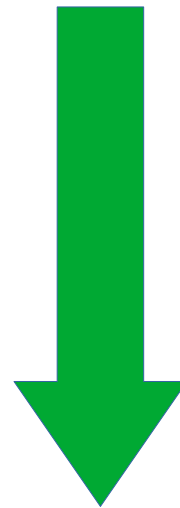


Group Work

- Break into groups of two or three people to complete the project.
- Use `data()` in R to get a list of available datasets to analyze
- Choose a dataset for the following super-fun activity that you will work on with your group members.

Which dataset
to use?

The scope of data?



Is this
numerical data?

Which is my group?



Group Work

- Inspect the data and discuss
 - What is the data *show*?
 - Is there a code book to explain the set?
 - What what kinds of questions can be answered using this data?
- Write down one of those questions for a research question.
- Define your hypothesis statements (**H_o** and **H_a**) to test

What can I do
with this data?

What are my hypotheses?



What questions can
this data answer?

Is there a codebook for my
data?

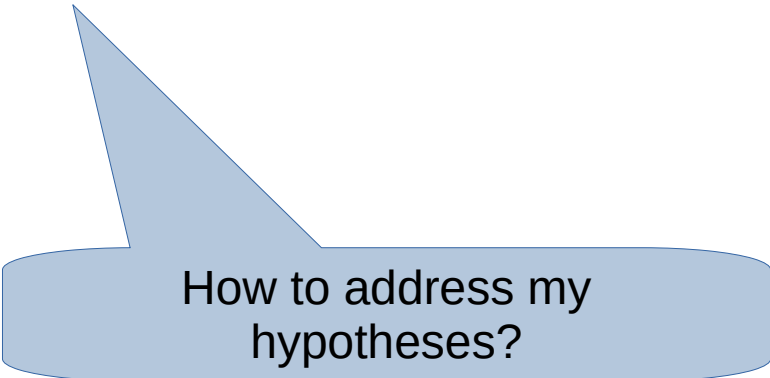


Group Work

- Use at least three different methods completed in class (or outside of class!) to form your study.
 - *Anova*
 - *Manova*
 - *T-tests*
 - *Linear Regression*
 - *Multivariate Regression*
 - *Other model*
- Note: you might need to adjust the question or choose a different dataset if the methods do not work.



Which model to use?



How to address my hypotheses?

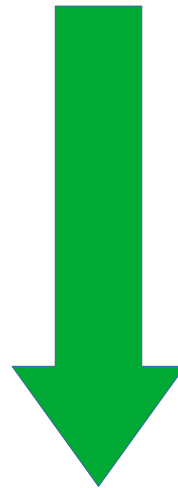


Group Work

- Perform the analysis. Compare methods: which *worked* best /worse for your data? Discuss with your group members!
- Create plots to show results to understand the models
- Write it down, save the model summaries to present
- Friday class to present your work as a group for your activity. Everyone must present something.

Which model
worked?

What plots to make?



Which slides to show

How do I know
my model worked?



Presentations on Friday

Lightning Talks of five minutes

- Come prepared on Friday to discuss the following using about five slides for a five minute lightning talk.
 - *The dataset*
 - *The research question and hypotheses*
 - *Statistical methods and summaries*
 - *Which “worked” best to address the hypotheses*
 - *Plots of the results*
 - *Lessons learned*
- **No submission, your presentation is the activity.**

