

The Big Data Camp Challenge

1. Use the techniques you are practicing here to collaboratively demonstrate **one plausibly true thing** about a topic that interests you.
2. Reflect on the process of demonstrating that thing
3. Present your finding, your process and the fruits of your reflection to the group **on FRIDAY 06/09.**

Here are our expectations

- Form groups
- Articulate a topic or question of shared interest
- Identify and gather relevant data
- Create a dataset through scraping, API utilization, and/or merging existing datasets
- Run queries, statistical analyses and/or visualizations using your data to demonstrate your one true thing
- Prepare a presentation that describes your question, your process, your findings, and what doing this taught you about working with “big data”
- Have Fun



Examples of Language

6+ / 0- = Failed



Perfect Joy Oil Paintings

by Janelle Wingert

Perfect Joy - a series of oil paintings - expressing perfect joy in those precious moments of childhood delight. Check it out!

📍 Cheyenne, WY

Funding Unsuccessful

Project ended on Jun 28 2012

5+ / 0- = Success



NOMADIC: "Major Love Journey"

My first project, "Major Love Journey" is a compilation of pieces that celebrate love in family, love in friendship, love of education



by Nomadic and 14 backers

3+ / 4- = Failure



Help Mark Get Dressed

by Alex Hoekstra

Poor Mark is poor. Poor and sexy. Poor, sexy Mark. He wants to dress well, but he needs help. A lot of help.

📍 Boston, MA

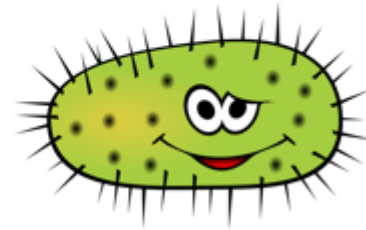
Funding Unsuccessful

Project ended on Aug 3 2014

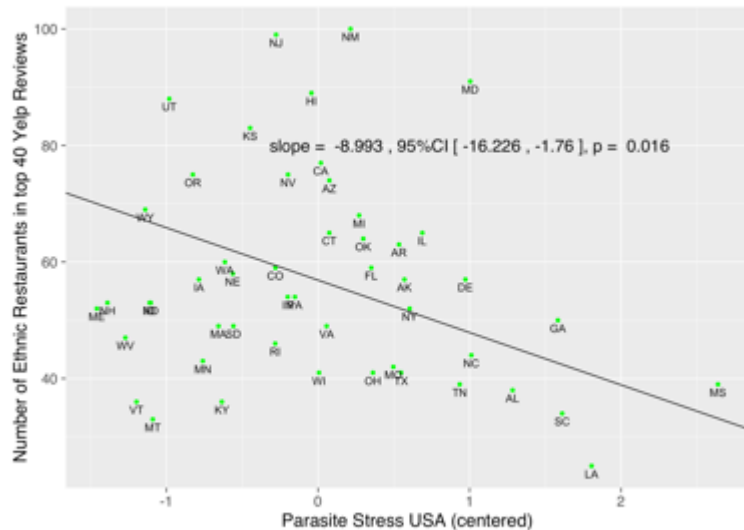
Finding success on Kickstarter: Being too positive or too negative results in not getting funded

- Threat of disease can lead people to increase their preference for conformity, familiarity, conservatism, to be less open to novelty, and to be more ethnocentric

(Murray, Schaller, & Suedfeld, 2013)



One true thing



$r(43) = -.36, 95\% \text{ CI } [-.58, -.09]$ (slope); adjusted $R^2 = .20$ (m)



Higher pathogen prevalence in a region is associated with fewer non-American restaurants in the top 40 rated restaurants on Yelp.

Presentation Parts

- Group Members
- Our motivating question was...
- Why this question was interesting to us
- Process of dataset construction
- Here is our fact: _____
- Our preliminary findings or results
- How you did what you did
- What you would do with more time

Presentation Parts (2)

- Slides of some sort – Powerpoint, pdf
- Presentation should be about 7 minutes
- At the very end of the presentation provide one to two tweets that summarize your project
- PLEASE KEEP AND SHARE YOUR CODE WITH TEDDY - this is helpful for future camps

How to come up with ideas?

THE SEVEN WHATS!

- What is a question floating around your field that people are having trouble answering?
- What is an old fundamental question or idea in your field that is incomplete?
- What is an interesting dataset that is under – explored?
- What are several massive datasets that I can connect?
- What resources are on the Github site?
- What do the other campers have to say?

Go forth and do interesting stuff.

WAIT! You need teams.

Team Formation (1)

- Objectives
 - Self select into teams around the topics that enflame your passions
 - Cross disciplinary teams less than or equal to six
 - Go through a series of steps to make teams

Team Formation (2)

- Education/Education Policy
- Networks
- Family Relationships
- Rankings and Ratings
- Social Movements
- Elites
- Emotions
- Race
- Politics
- Healthcare
- Firm Performance

Three Questions, Three Data Sources

Cross-Disciplinary Conversations

1. Send one person to get a sheet.
2. Go some place and chat.
3. Come up with a preliminary idea.
4. Return the sheet filled out in this room at 4:45.