

Data science without coding

Welcome! This is the landing page for our breakout session.

This page will persist and has links to all of the things we will do.

DO NOT USE INTERNET EXPLORER.

- [CODAP](#) (Also see <https://codap.concord.org> for more info and resources)
- [NHANES data](#): Health data with height, weight, etc. 1000 kids. Static file.
- [American Community Survey data](#): Census data with income. California, 2013. Data portal.
- [BART data](#): Mass transit in the SF Bay Area. 40M+ cases!
 - [BART suggested topics and a link](#) to a 4-lesson sequence

Underlying idea

Although coding is [currently] essential for doing real data science, starting with coding for *learning* has problems for all sorts of reasons. Maybe it's better to start learning data science with a more drag-and-drop kind of system. CODAP is one such system; what can we do that "smells like" data science?

Basic plan

We will spend most of our time hands-on, using CODAP to do various tasks. We'll begin by following me directly, then branching out to more independent work.

We will address questions that come up. If there's time, you'll see some student work.

As you work, think about

- What are you doing that "smells like" data science?
- What part of all this (if any) appears in a traditional stats course?
- Data science often is attached to "computational thinking." Are we doing that here?

CODAP tips

- Drag categorical labels on graphs to re-order them
- Column titles are also menus where you can find **Edit Formula**
- Drag attribute names leftwards in the table to group by that attribute
- Click the gray, circular "plus" to add a new attribute in that area
- Use the "eyeball" menus to hide/show cases (in graphs) or set aside (in tables)

Data Moves

Actions that alter a dataset's values, contents, or structure.

Examples: filtering, grouping, summarizing, recoding.

Contact!

Tim Erickson • Lick-Wilmerding High School • Epistemological Engineering • eepsmedia at gmail dot com

High-school assignment, March 2019

Students—mostly juniors and seniors—had two class sessions of 75 minutes each, starting from nothing. The ending mini-investigation was homework over the following weekend.

- [Class notes and assignment.](#)
 - This is a running document in reverse chronological order, that is, the most recent text is at the top.
 - It's in a class called Applied Math, which has few prerequisites. As such, although students are well-behaved and pretty motivated, some have not had much success in high-school math. You will find the first homework near the bottom
- [Student work.](#)
 - These are from the final (top-most in the previous link) assignment, where students picked their own topic.
 - The first two (with asterisks) are both pretty good and interesting.
 - They were mostly google docs. I have "printed" them to PDF and redacted names.
 - In much of this, students make mistakes and erroneous assumptions. And sometimes they're pretty far off-base. This has not been filtered to show only great work!

Other things to explore

- A game about [diagnosis and classification trees.](#)
- [Playing Galileo.](#)