STAT 135
1. Overview

Spring 2022

Lecturer: Dr Rebecca Barter (*she/her*)
Office hours: Tu 9:30-10:30, Th 1:00-2:00
Office: Evans 339

Email: rebeccabarter@berkeley.edu
Twitter: @rlbarter
GitHub: rlbarter

About me

Hi, my name is Rebecca!

Zoom class conduct

- Keep your microphone muted during class
- Feel free to “raise your hand” and to type questions in the chat (I am more likely to get to a raised hand faster than a chat question)
- Please only unmute yourself if I call on you to do so
- I know zoom is tiring. I will do my best to plan for breaks during class
- Class will be recorded and posted on BCourses (both when remote and in-person)

Photo by Chris Montgomery on Unsplash

Getting in touch

Piazza (https://piazza.com/berkeley/spring2022/stat135/home)

- Use Piazza to ask questions about course content, the homeworks and projects
- You are encouraged to answer one another's questions

Office Hours

- If you have questions for me or your GSI, feel free to attend our office hours
- My office hours will be right after class and will take place in the same zoom link while remote, or in Evans 339 when in person.

Email

- Try to utilize piazza and office hours when asking ask questions about the homework, projects, etc. Of course if you have a sensitive or personal question, feel free to email us.

What to expect from STAT 135

What to expect from STAT 135

1. **Critical thinking with data**
   - Data and reality
   - Exploratory Data Analysis

2. **Inference**
   - Sampling
   - Likelihoods
   - Maximum likelihood estimation
   - Sufficient statistics
   - Confidence intervals
   - The delta method
   - Bootstrapping

What to expect from STAT 135

3. **Hypothesis testing**
    - Null and alternative hypotheses
    - p-values
    - Significance levels and power
    - Confidence intervals
    - Tests for proportions
    - Multiple testing issues
    - Tests of comparison
    - Non-parametric tests

4. **Causal inference**
    - A/B testing
    - Neuman-Rubin potential outcomes
    - Confounding

What to expect from STAT 135

5. **Linear models**
    - Prediction problems
    - Least squares
    - Linear regression (+ inference)
    - Logistic regression

Assessments

Homeworks (20%)

There will be ~8 homeworks.
Homeworks will involve mostly technical and theoretical questions.

Data projects (20%)

There will be ~3 data projects.
These projects will involve conducting some relevant data analysis and drawing conclusions.

Midterm (20%)

There will be one midterm, currently schedules for Thursday March 10

Final exam (40%)

There will be a final exam, currently scheduled for May 12

How fun! A survey!

https://4h2azvklbdb.typeform.com/to/RpG7bG9G

Quick breakout room

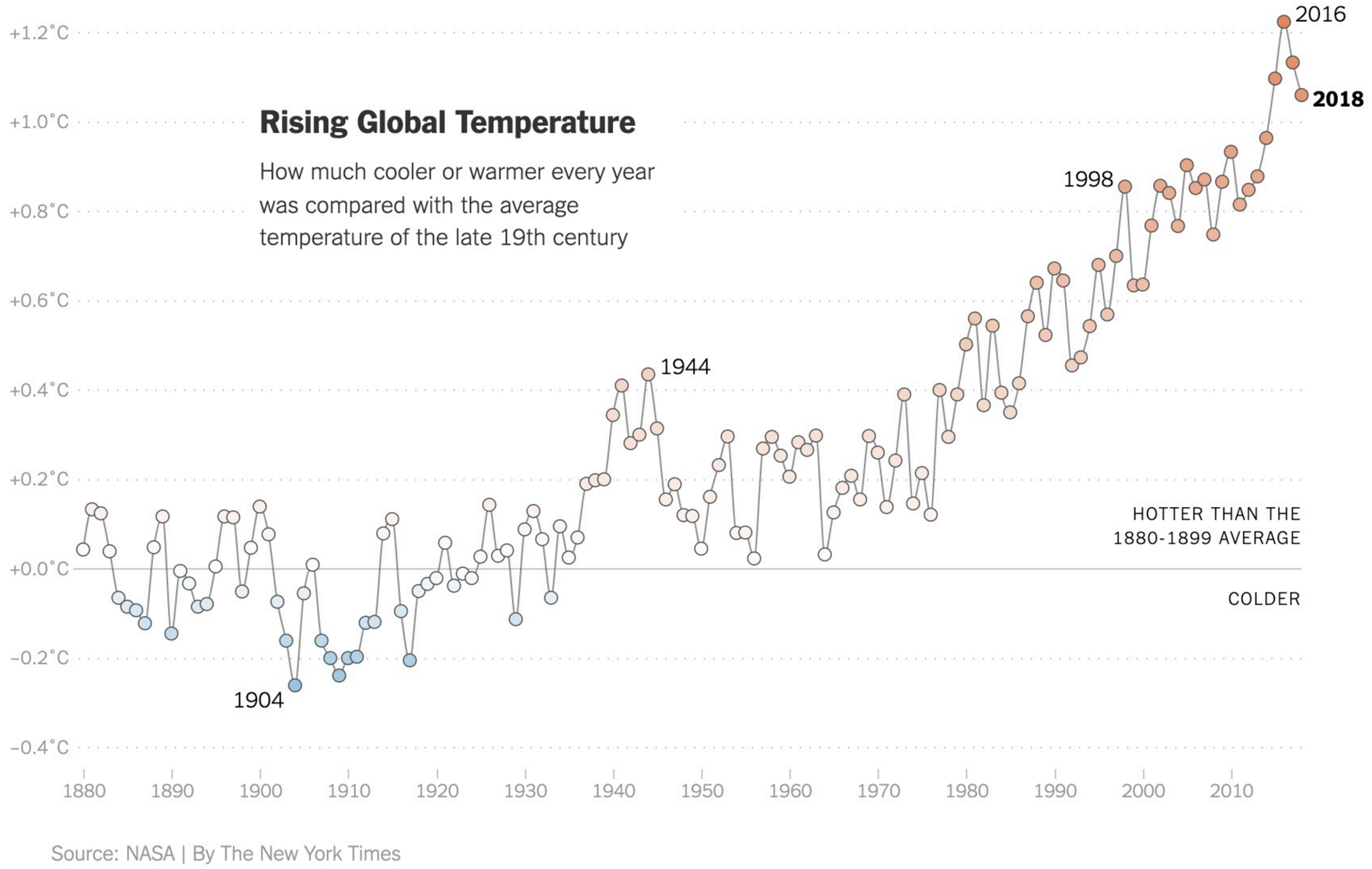
Introduce yourself to a few of your fellow classmates.
1. What is your major?
2. What do you do for fun?

A whirlwind tour of STAT 135

Critical thinking

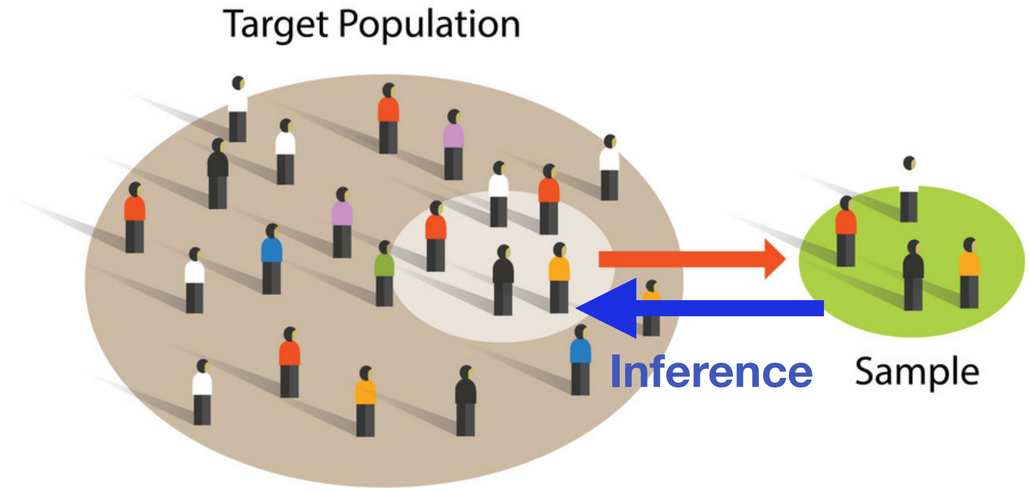
Real data is a usually a very **imperfect** reflection of reality.

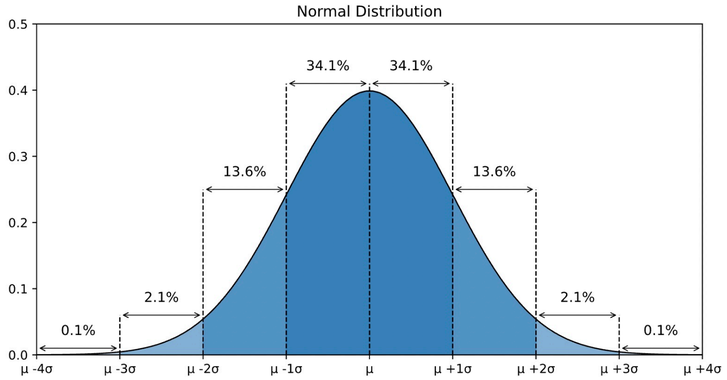
Exploratory Data Analysis



Inference

How do we learn about a population from a data sample?



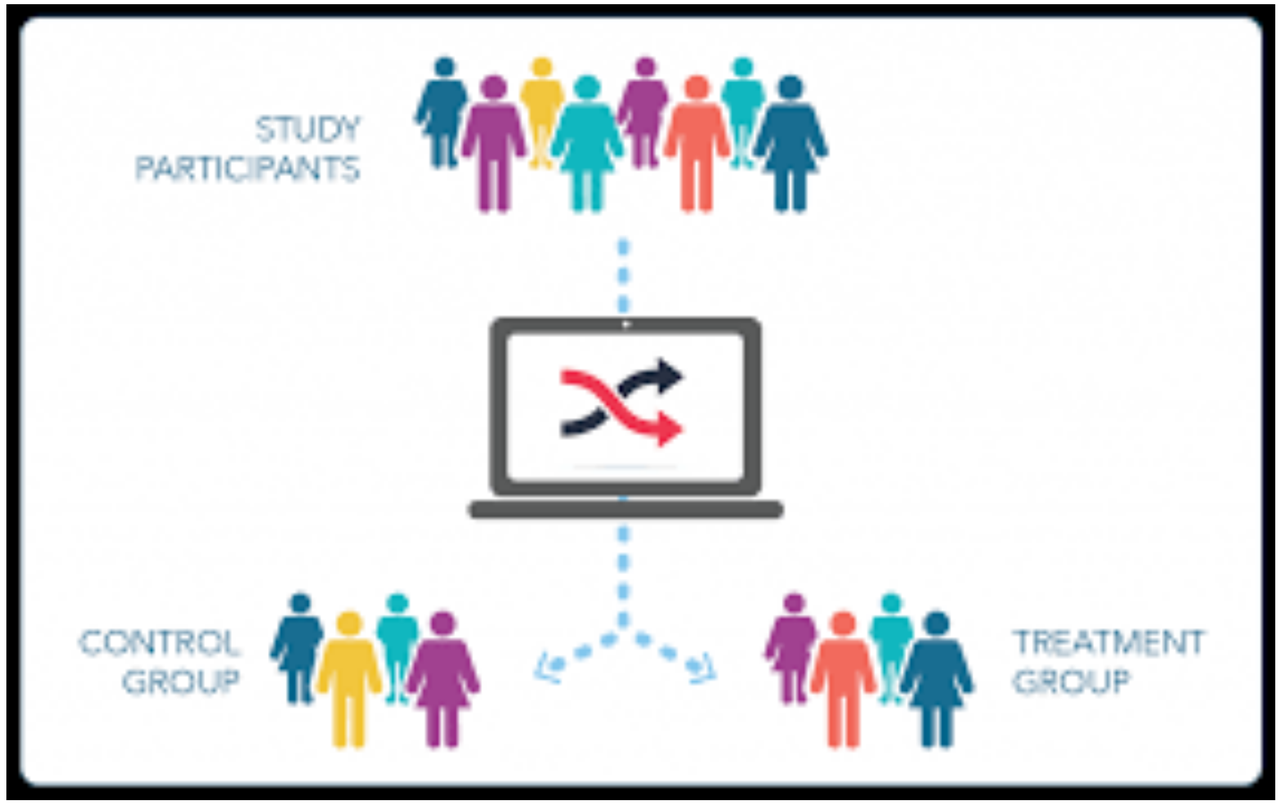


$$\hat{\theta} = \arg\max_{\theta \in \Theta} \widehat{L}_n(\theta\,; \mathbf{y})$$

$$\sqrt{n}\left(\widehat{\theta}_{\text{mle}} - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0,\, \mathcal{I}^{-1}\right)$$

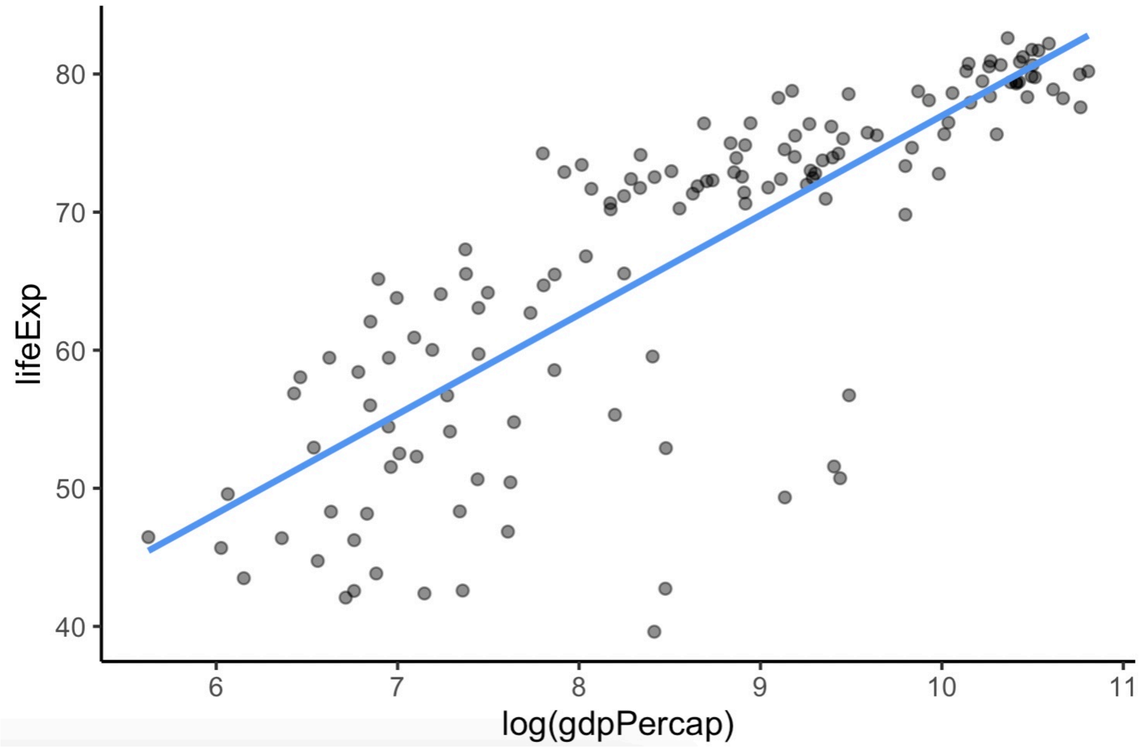
$$\mathcal{I}_{jk} = \mathbb{E}\left[-\frac{\partial^2 \ln f_{\theta_0}(X_t)}{\partial\theta_j\, \partial\theta_k} \right]$$

Hypothesis testing
causal inference & A/B testing



Is there a difference between the treatment and control response?

Linear models



Your goal might be:
Prediction:

- What would the life expectancy of a country whose GDP per capita is 9?

Inference

- What is the global relationship between GDP per cap and life expectancy?

Questions?

I love questions. They make me feel less lonely.