

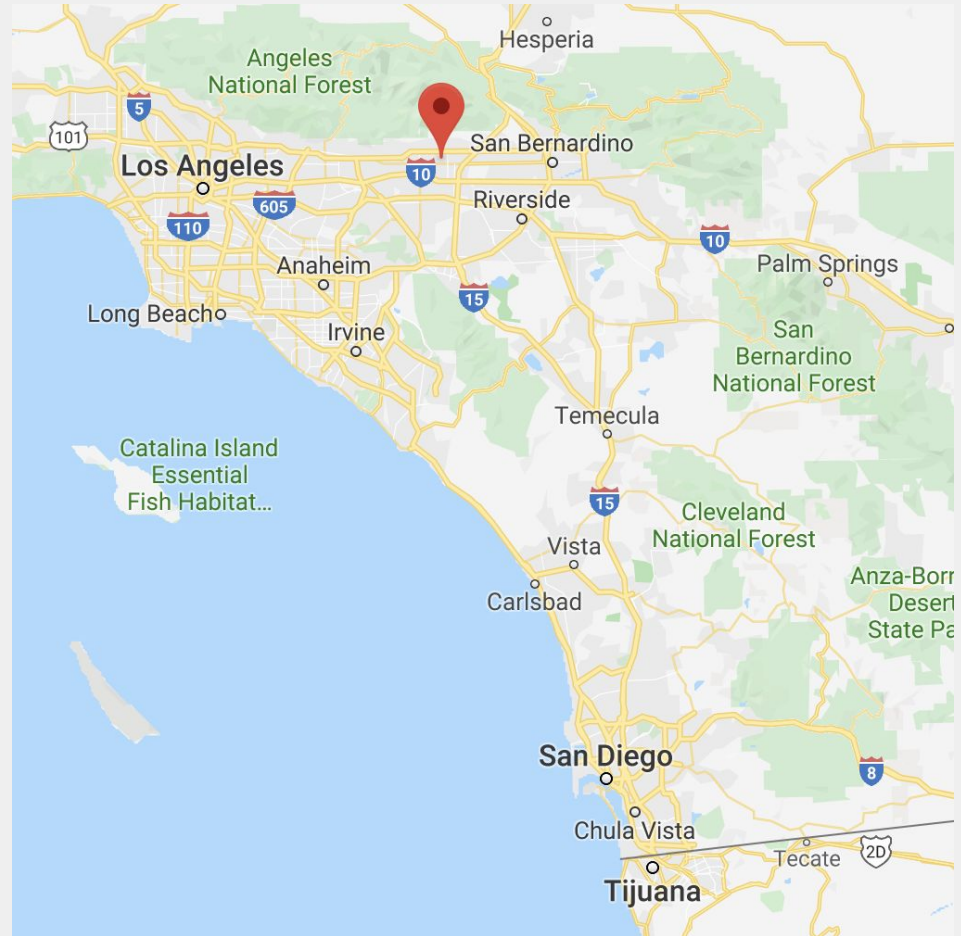


How I Became a Harvard Data Scientist

Heather Mattie

July 5, 2022

→ I was born and raised in Rancho Cucamonga, California



- I was born and raised in Rancho Cucamonga, California
- I took honors math and AP Calculus in high school along with AP Biology, Econ/History and Spanish



- I was born and raised in Rancho Cucamonga, California
- I took honors math and AP Calculus in high school along with AP Biology, Econ/History and Spanish
- Became a Math major in undergrad at the University of La Verne





Claremont
Graduate
University

Harvard
University



Undergraduate Math Major

Goal: become a high
school math teacher



**Undergraduate
Math Major**

Goal: become a high
school math teacher

**Master's Degree
Applied Math Major**

Goal: become a college
math teacher



**Undergraduate
Math Major**

Goal: become a high
school math teacher

**Master's Degree
Applied Math Major**

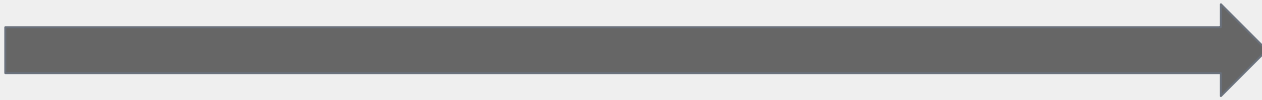
Goal: become a college
math teacher

**PhD
Biostatistics Major**

Goal: become a
Biostatistician and
make bank working in
industry

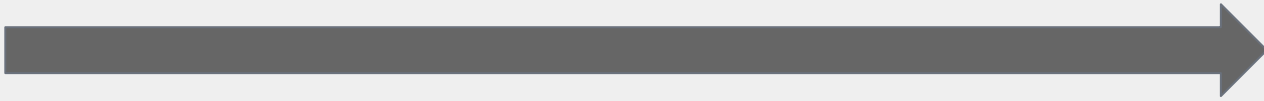


How I thought life/academia would go:





How I thought life/academia would go:



How life/academia actually go:



Main Lesson: Plans change!

How I thought life/academia would go:



How life/academia actually go:



Main Lesson: Plans change!

- Rediscovered my passion for teaching
- Loved my department and collaborators
- Created a job just for me
 - Teach 4 data science and biostatistics courses a year
 - Work on several research projects
 - Help run a Master's degree program in Health Data Science

Research



What is a health disparity?

A disproportionate number of health conditions and deaths compared with the general population

African Americans make up 13 percent of the U.S. population ...



... but represent almost half of all new HIV cases.



Algorithmic Bias

Algorithmic bias occurs when the application of an algorithm **compounds existing inequities** in socioeconomic status, race, ethnic background, religion, gender, disability or sexual orientation to amplify them and **adversely impact equity in health systems**.



Network Science

Developed methods that predict tie strength in a network, which assists in modeling the spread of disease and information. Now working on how to use mechanistic models to generate real-world networks



Health Disparities

How to improve health outcomes for marginalized communities like BIPOC and those suffering from HIV and the homeless population, especially with regard to COVID-19

AI and Healthcare

Implementing algorithms in healthcare, generalizability of the outcomes of these algorithms, infrastructure needed to implement AI in healthcare



Resources

- My email: hemattie@hsph.harvard.edu
- Harvard Biostatistics [Undergraduate Summer Program](#)
- My [Intro to Data Science course website](#)
- My [GitHub page](#)
- Free data science, statistics, math and computer science courses:
 - ◆ edx.org
 - ◆ coursera.org
 - ◆ datacamp.com