

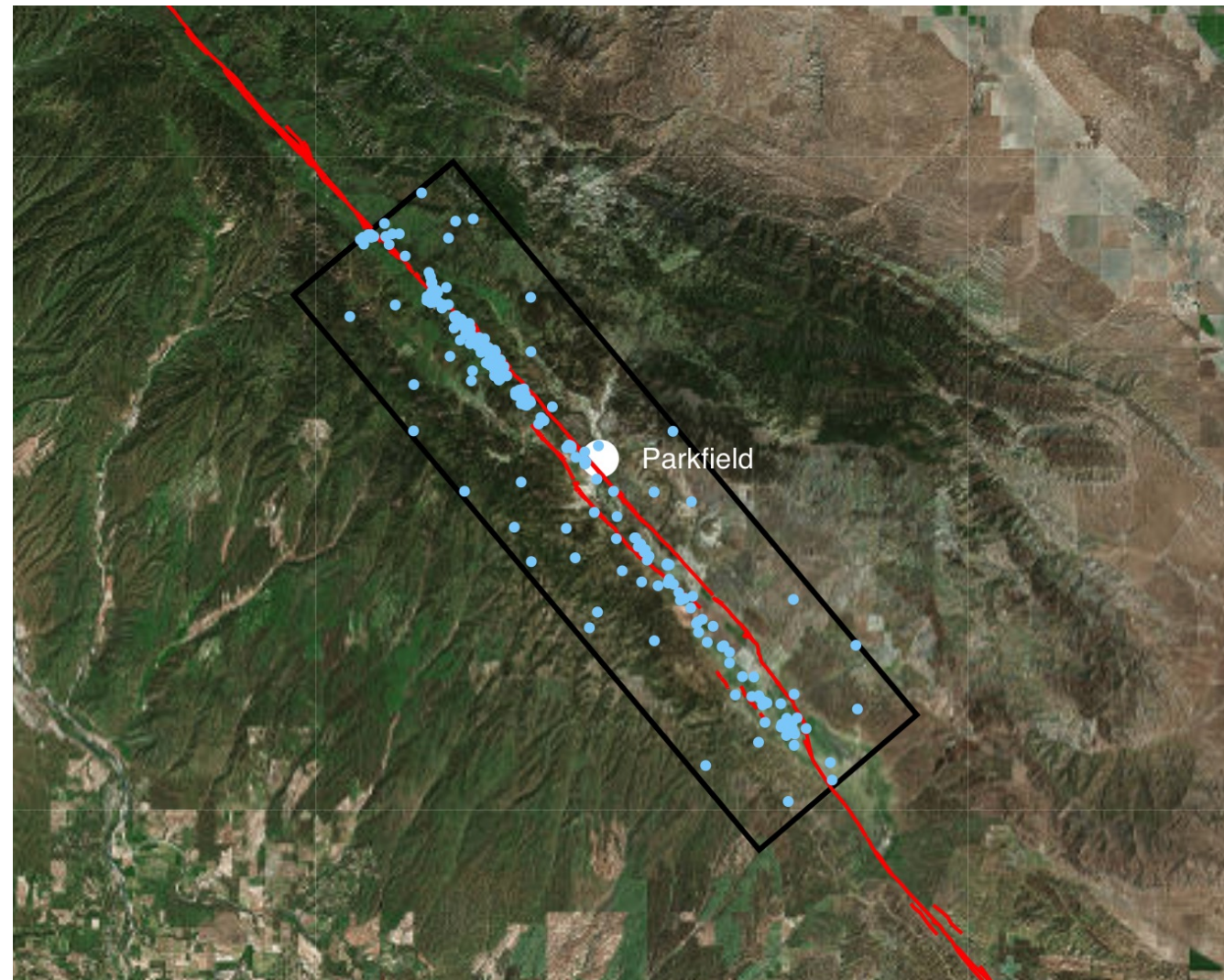


CASE STUDIES IN STATISTICAL THINKING

# **Variations in earthquake frequency and seismicity**

**Justin Bois**  
Lecturer, Caltech

# Movements along faults cause earthquakes



Fault data: USGS Quaternary Fault and Fold Database of the United States

Parkfield box proposed by Michael, and Jones, Bull. Seism. Soc. Am. **88**, 117-130, 1998

Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)

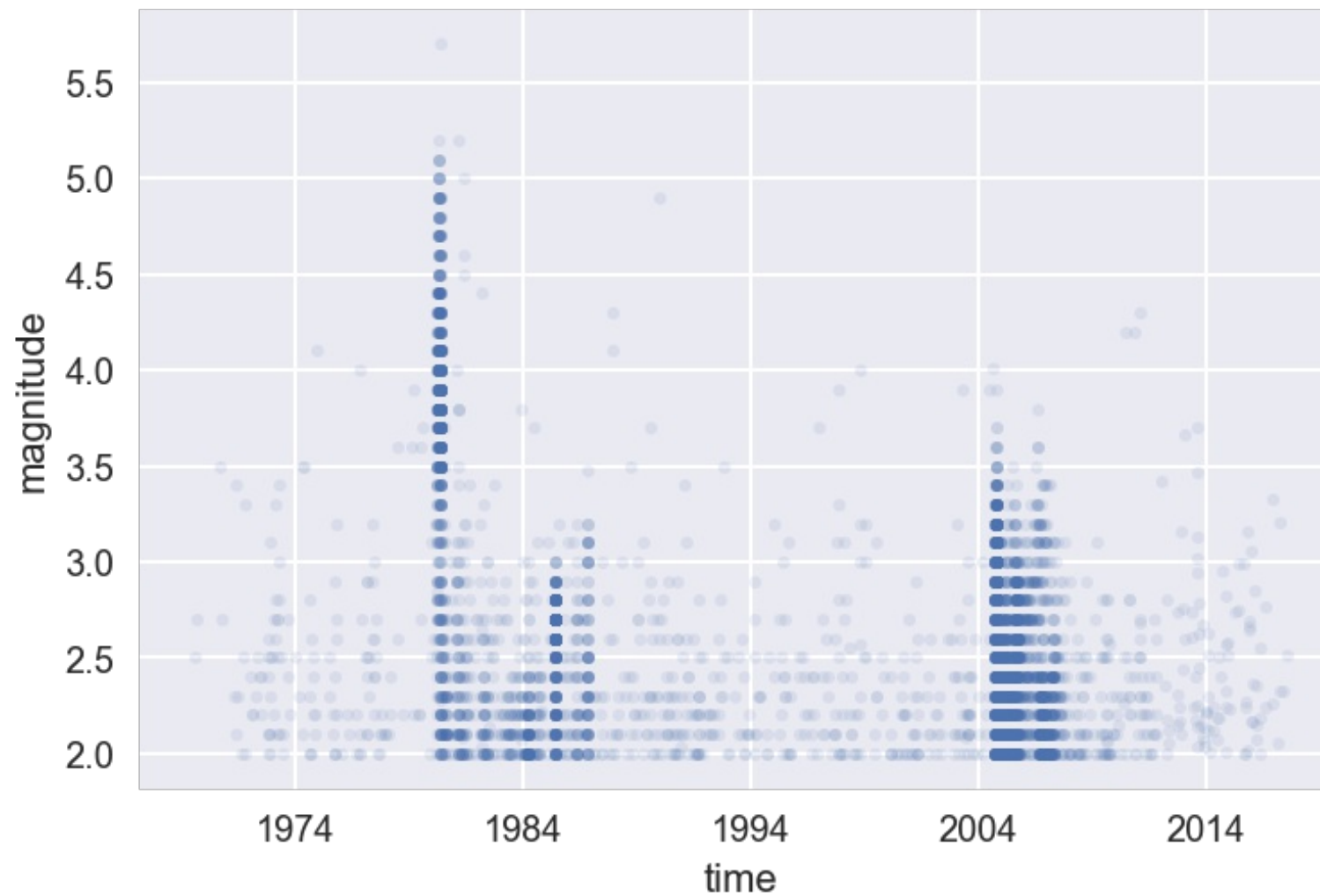
# Mt. St. Helens



*Photo: Austin Post, USGS, public domain*

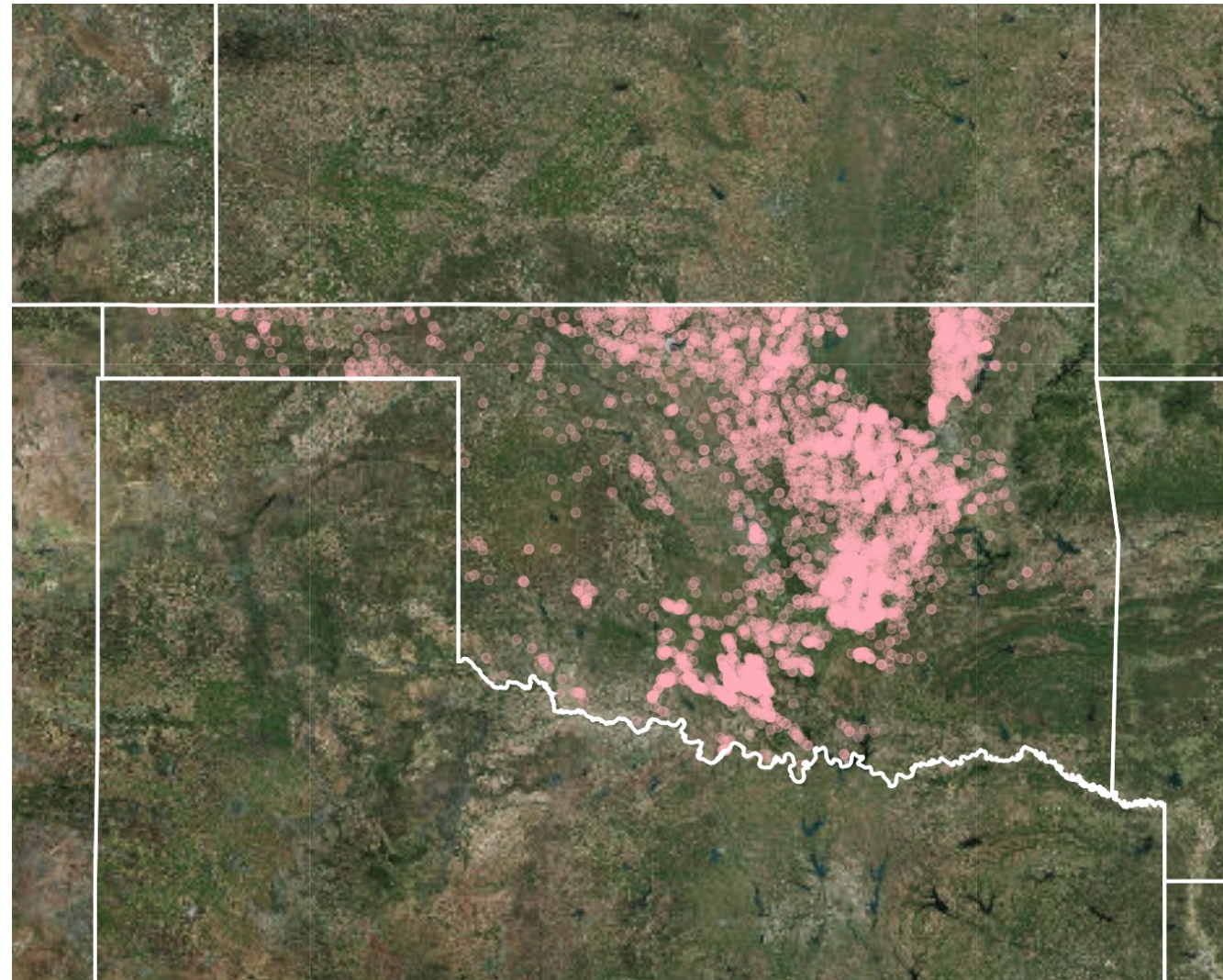


# Mt. St. Helens earthquakes



*Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)*

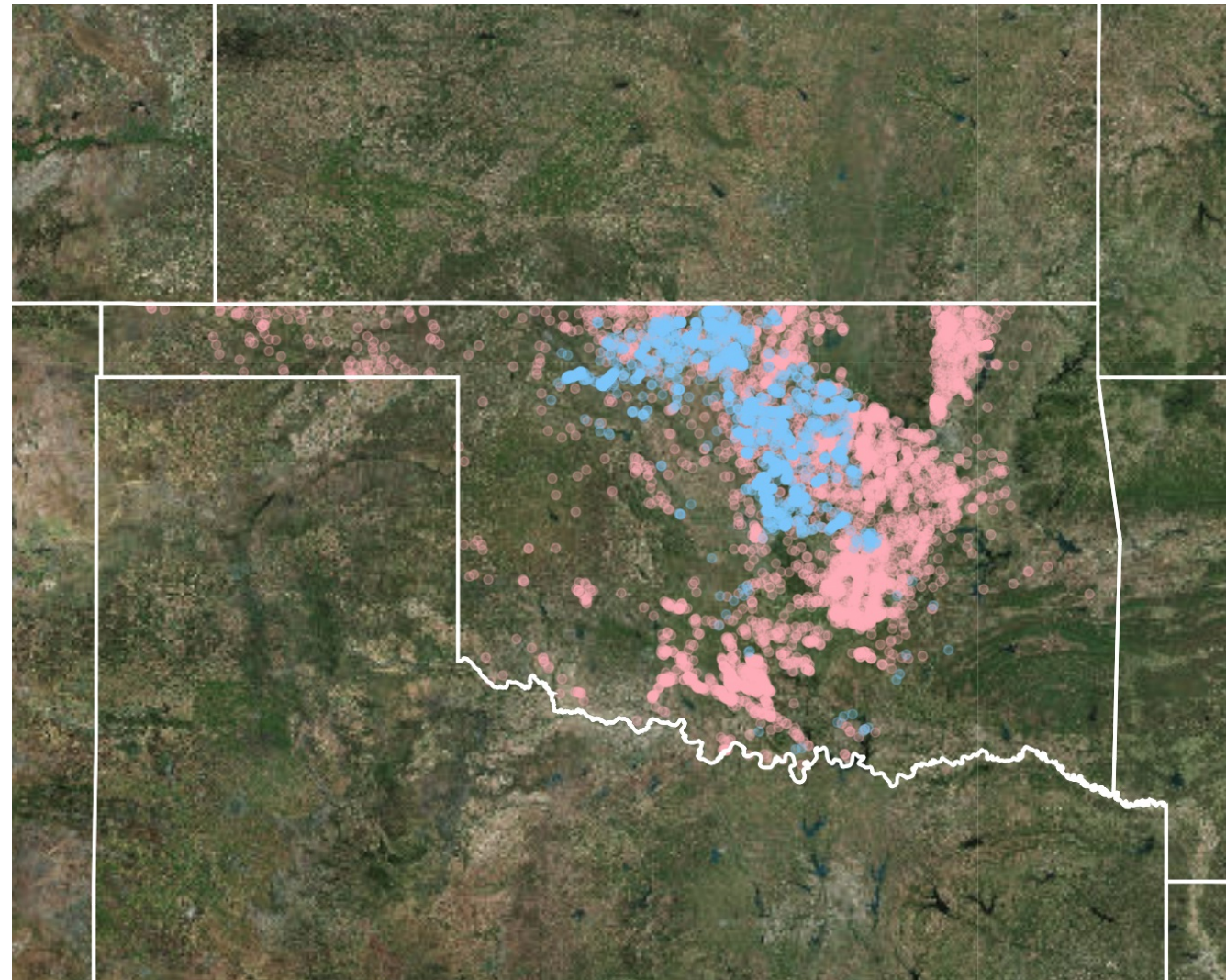
# Wastewater injection in Oklahoma



*Data source: Oklahoma Corporation Commission Oil and Gas Division*



# Seismic Oklahoma



*Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)*



## CASE STUDIES IN STATISTICAL THINKING

**Let's practice!**



CASE STUDIES IN STATISTICAL THINKING

# Earthquake magnitudes in Oklahoma

Justin Bois  
Lecturer, Caltech



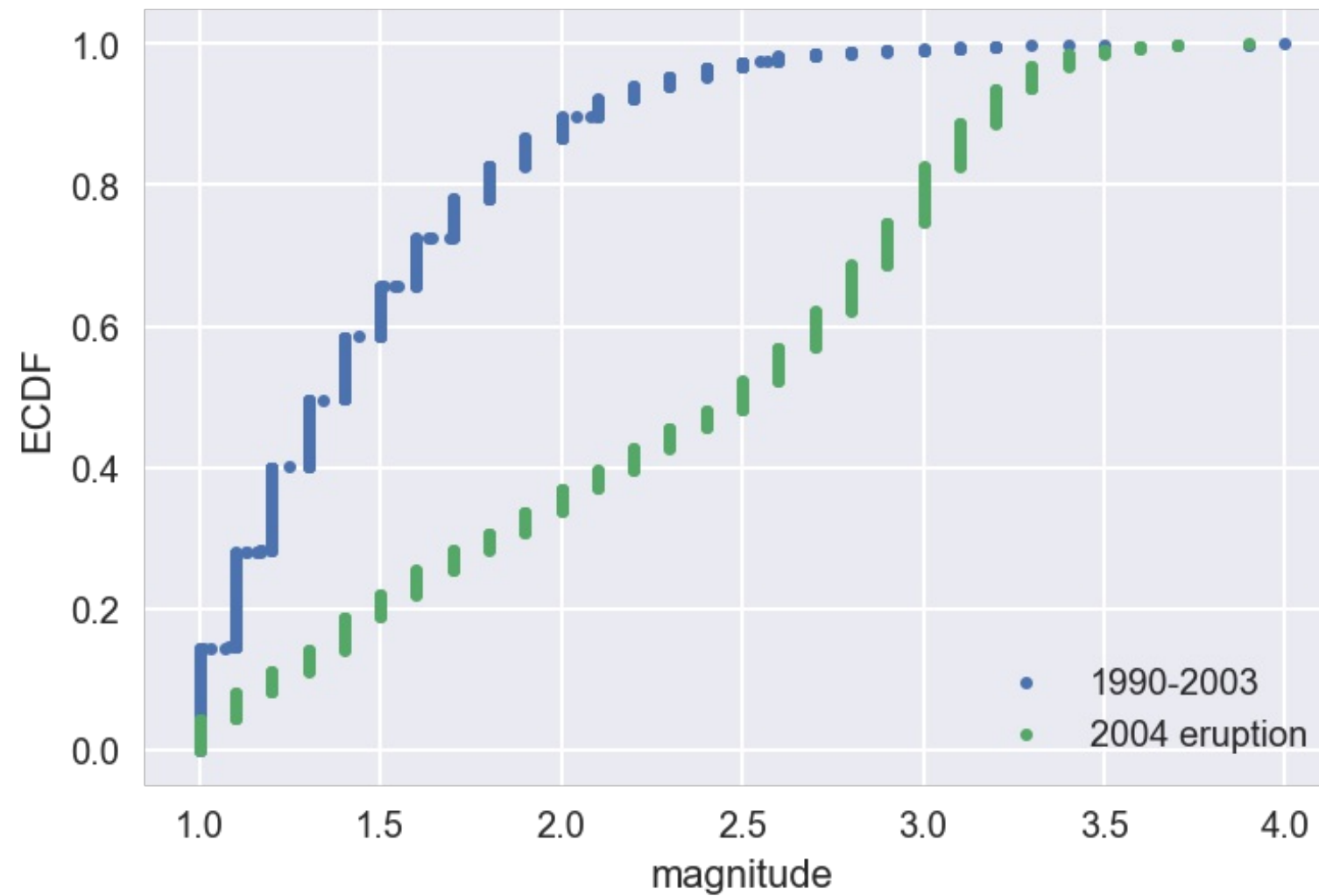
# Seismicity around Mt. St. Helens



*Photo: Steve Schilling, USGS, public domain*



# ECDF of Mt. St. Helens earthquakes



Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)



# Magnitudes in Oklahoma

- Verify that the Gutenberg-Richter Law holds before and after 2010
- Compute  $b$ -values
- Perform hypothesis test





## CASE STUDIES IN STATISTICAL THINKING

**Let's practice!**



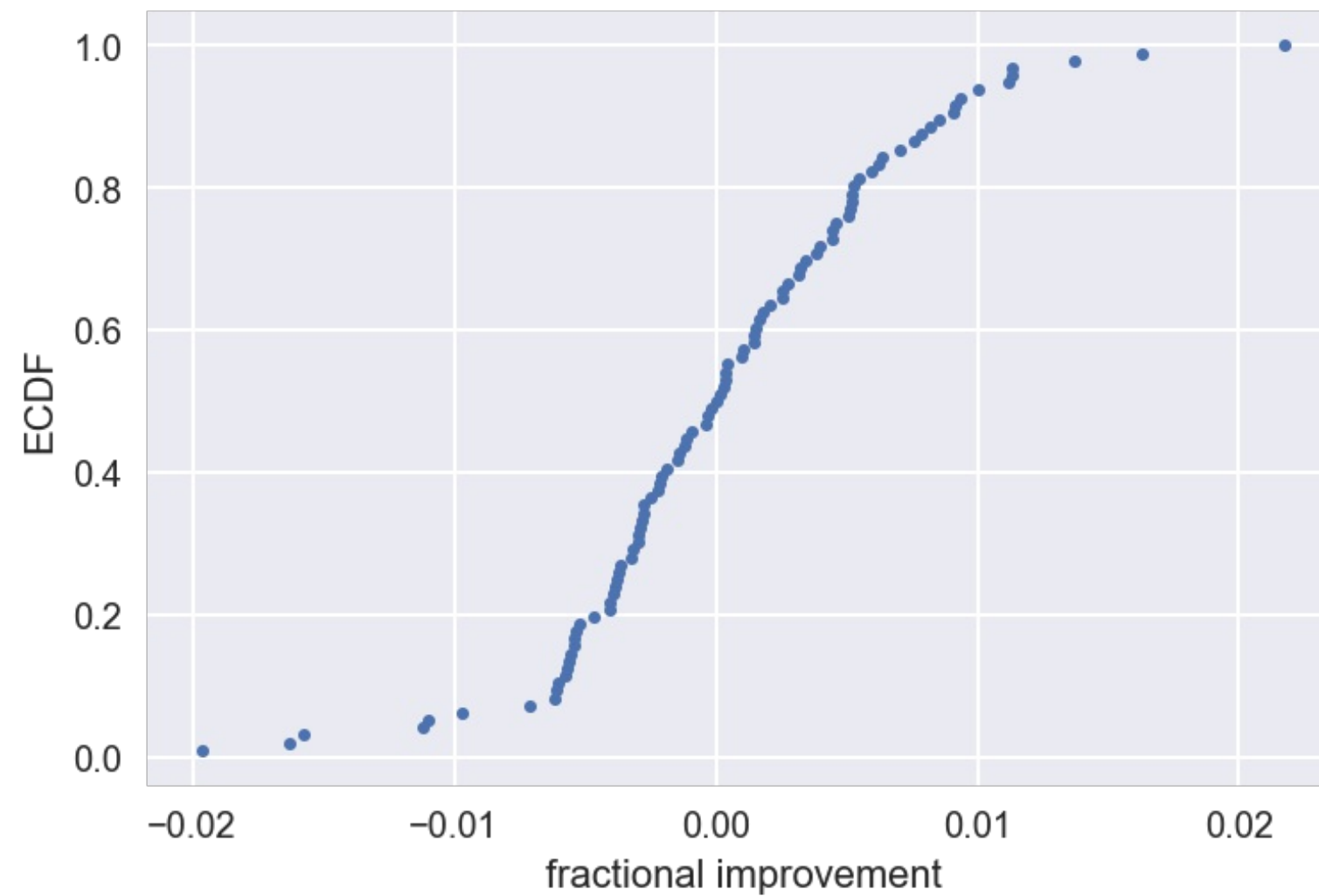
## CASE STUDIES IN STATISTICAL THINKING

# Closing comments

Justin Bois  
Lecturer, Caltech

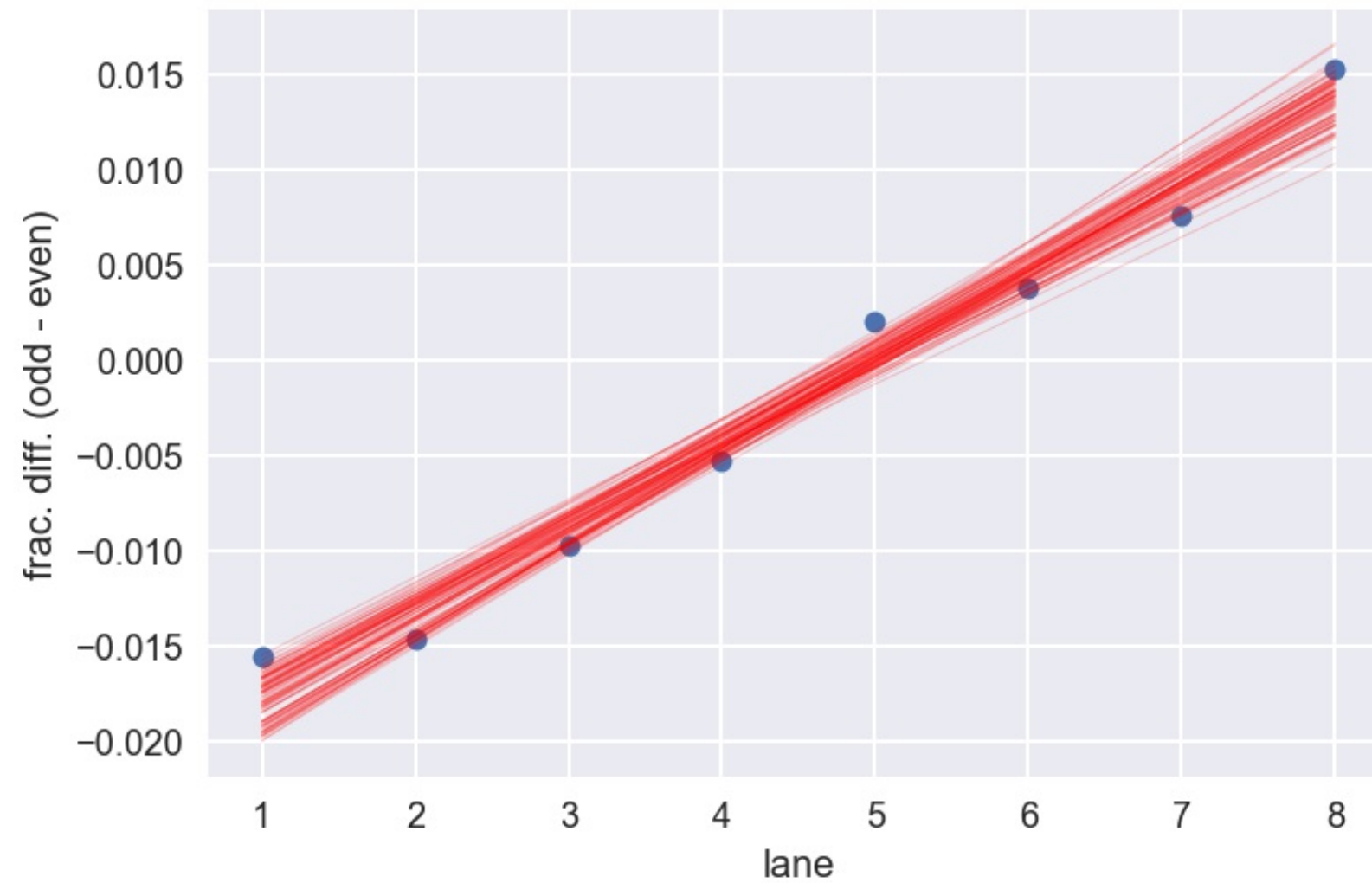


# Fractional improvment in finals



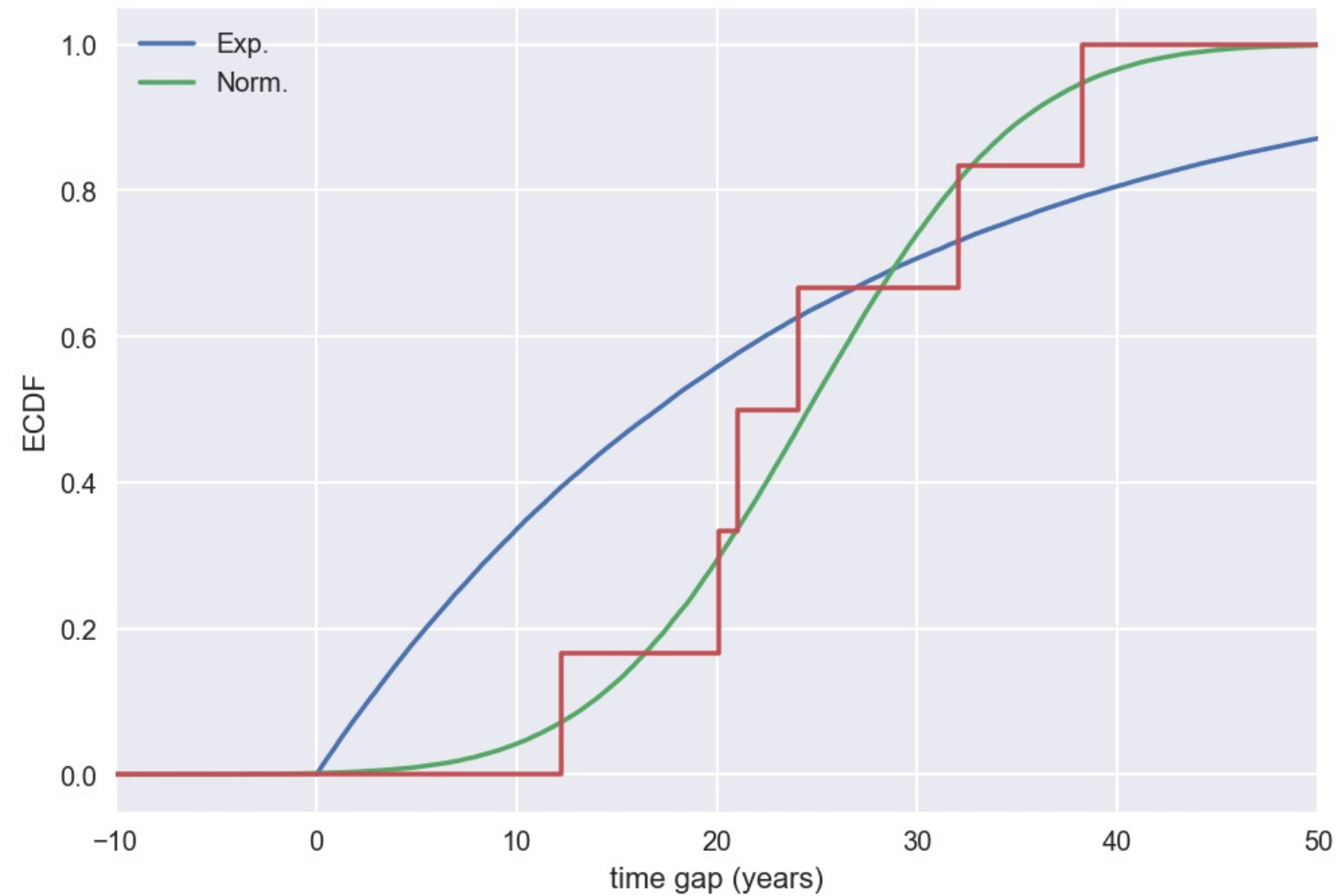


# Quantifying the Current Controversy



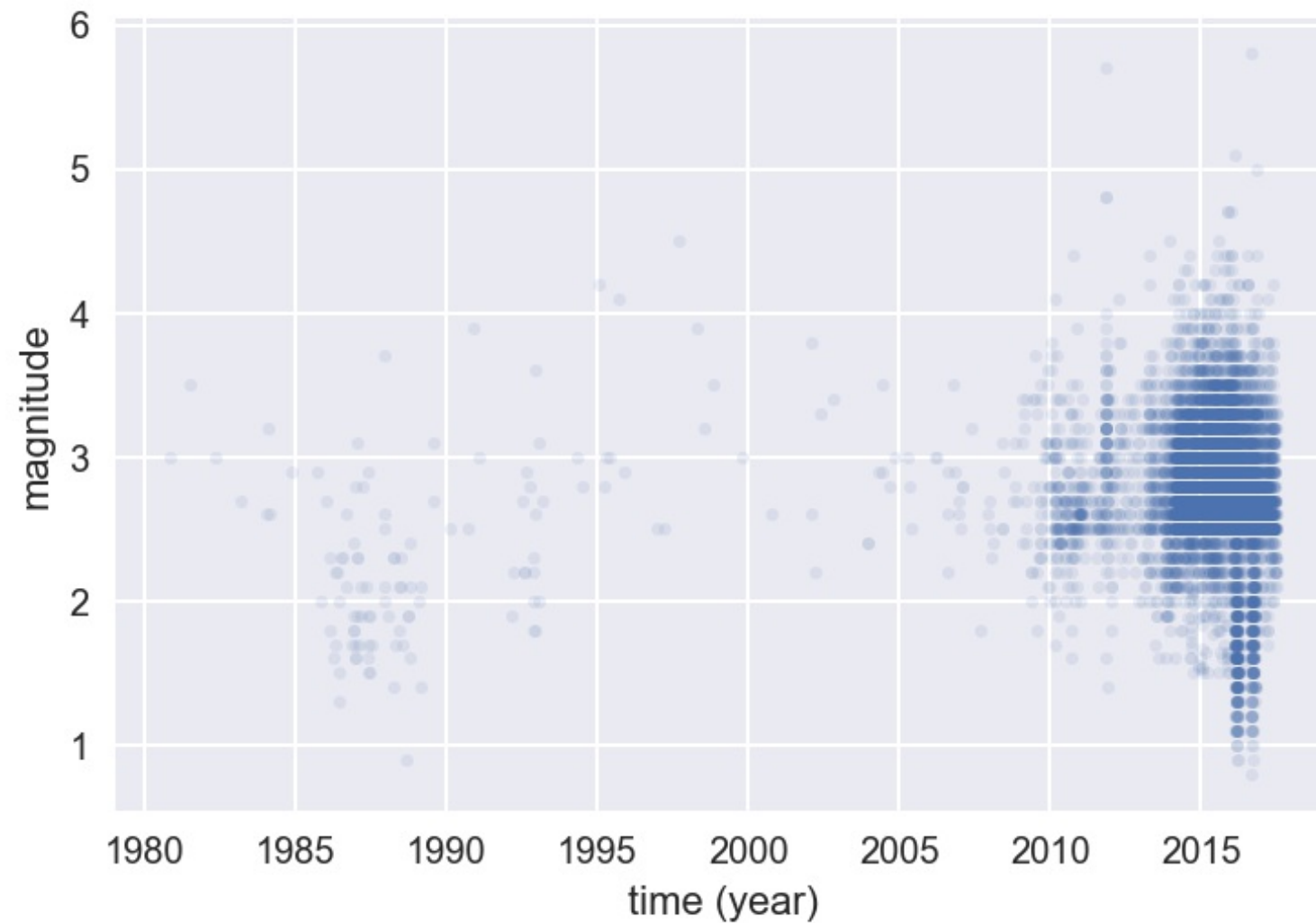


# Inter-earthquake times of the Parkfield sequence





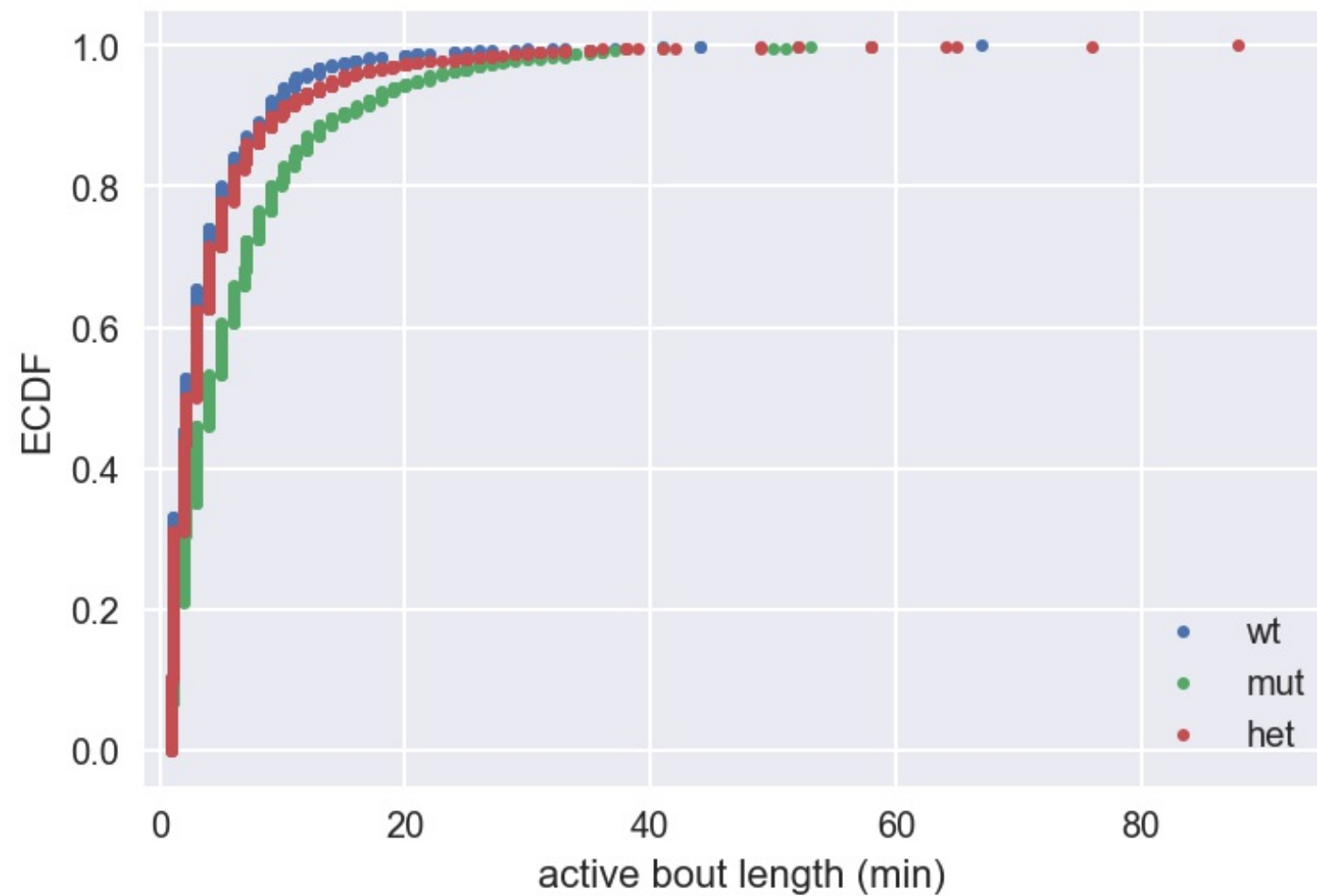
# Oklahoma earthquakes



*Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)*



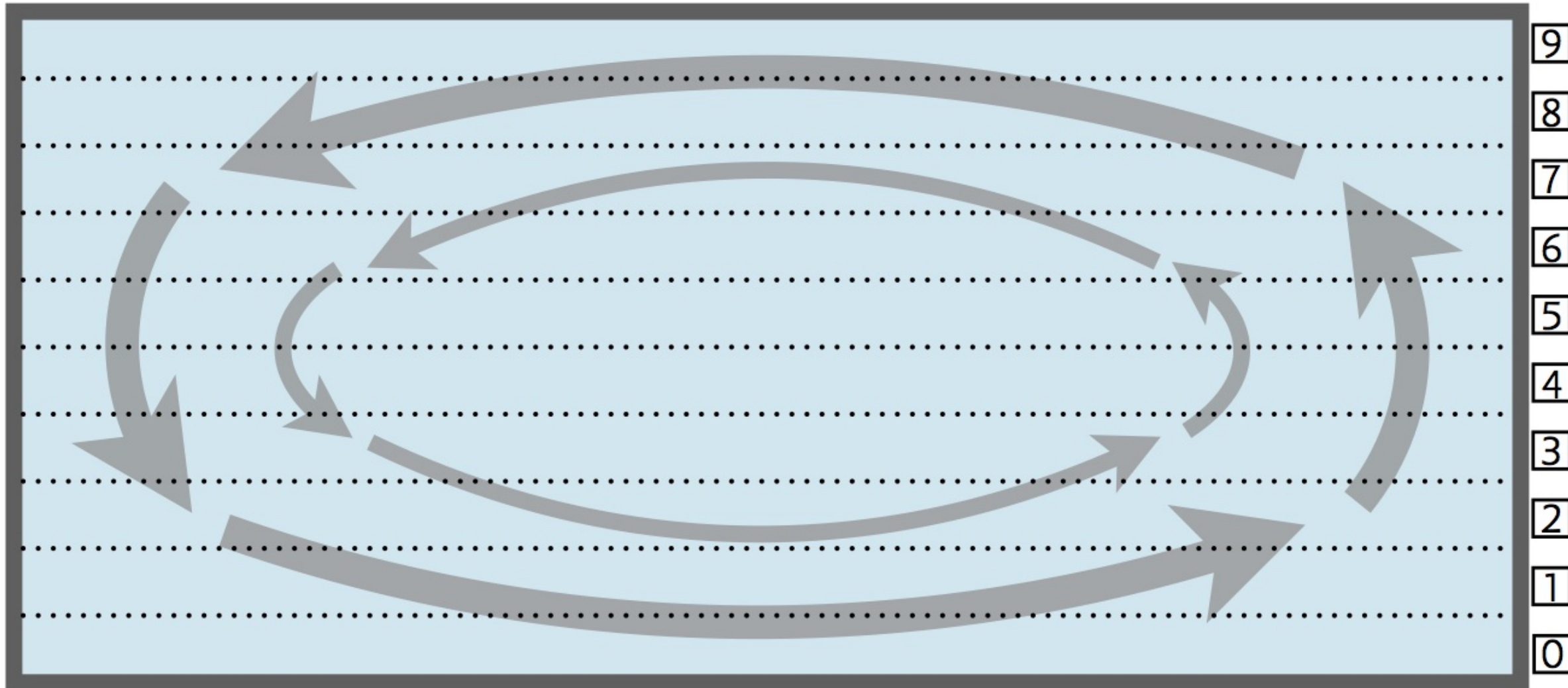
# Active bout lengths



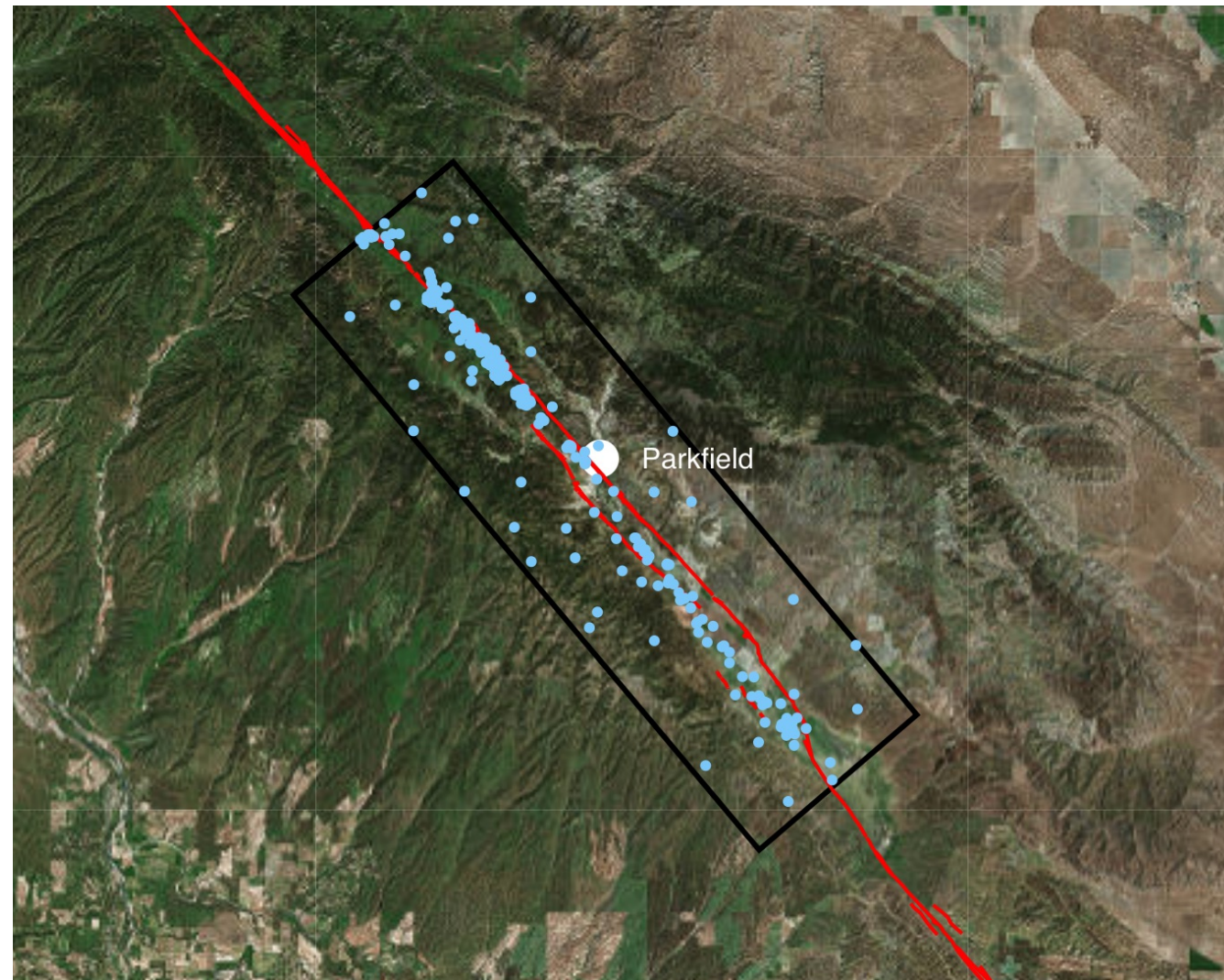
Data courtesy of Avni Gandhi, Grigorios Oikonomou, and David Prober, Caltech



# Each data set offers unique challenges



# Each data set offers unique challenges



Fault data: USGS Quaternary Fault and Fold Database of the United States  
 Parkfield box proposed by Michael, and Jones, Bull. Seism. Soc. Am. **88**, 117-130, 1998  
 Data source: USGS ANSS Comprehensive Earthquake Catalog (ComCat)





# My goals for you

- Recognize your opportunity to learn and contribute



# My goals for you

- Sharpen hacker stats skills
  - Bootstrap confidence intervals
  - Permutation and bootstrap hypothesis tests
  - Kolmogorov-Smirnov test
  - Location parameters
  - Log-linear regression



# My goals for you

- Statistical inference pipeline
  - EDA
  - Parameter estimation/confidence intervals
  - Hypothesis test

# My goals for you

- Ability to frame precise questions





## CASE STUDIES IN STATISTICAL THINKING

**Go forth and data!**