Storyboarding as Part of the Process of Statistical Investigation

Sara Stoudt @sastoudt
Deborah Nolan @DebAtStats

Note: Data in sample figures comes from the Drug Abuse and Warning Network (DAWN)

What is a data storyboard?

- visual outline used in movie making → set of figures and tables with connecting text from an exploratory phase of analysis
- tactile way to group, filter, and order intermediate products of an analysis to find the bigger picture


https://oaklandliving.wordpress.com/2010/12/29/last-chance-to-see-pixar-at-oakland-museum-of-california/

Activity Learning Goals

- Create/choose visual elements that bolster a written argument.
- Incorporate iteration into a writing workflow to leave room for adapting the scope and honing statistical arguments.
- Break the habit of organizing writing (e.g. final report) based on the order of exploratory and analysis **processes** in favor of organizing it based on a **narrative**.
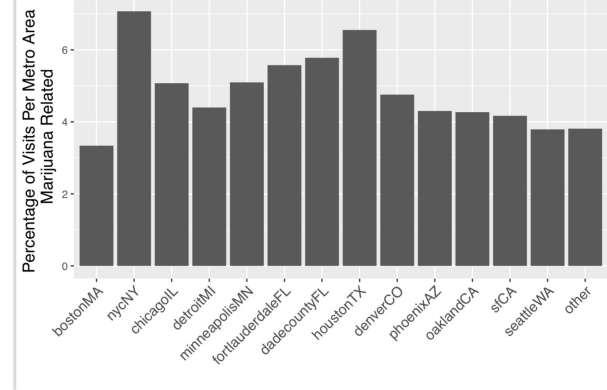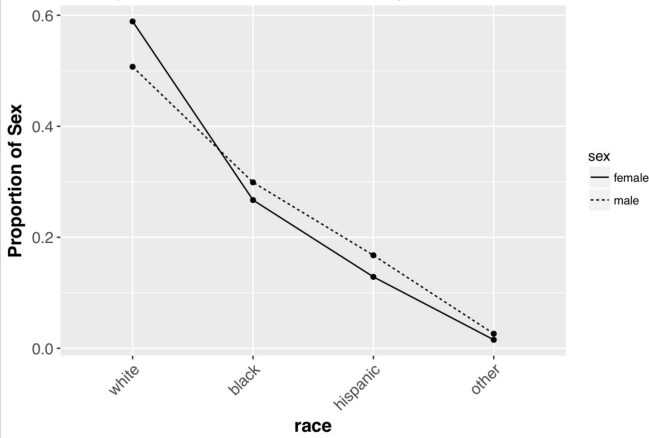
I did this, then this, then this,...

I found that…
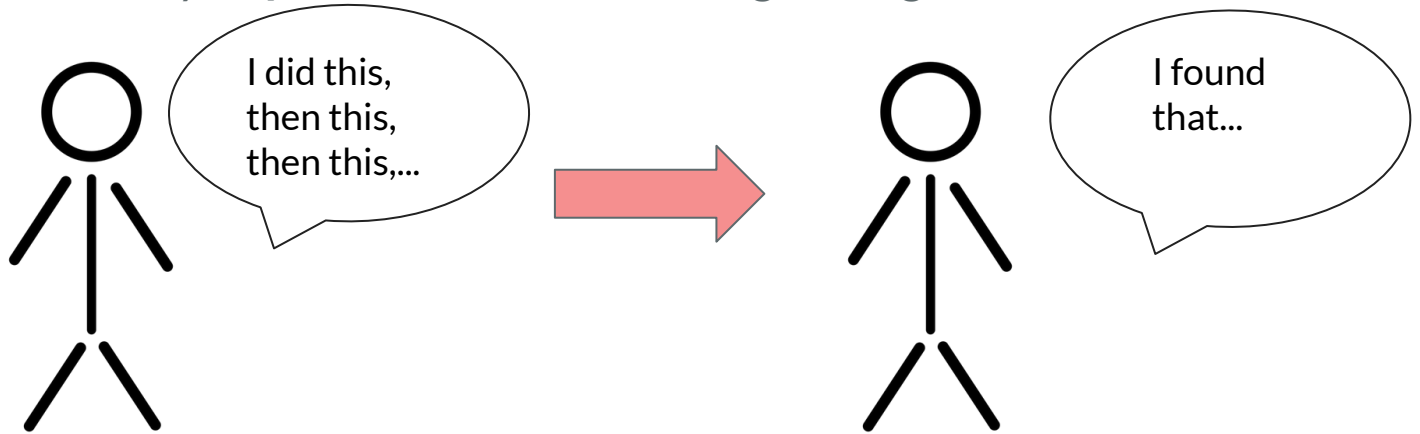
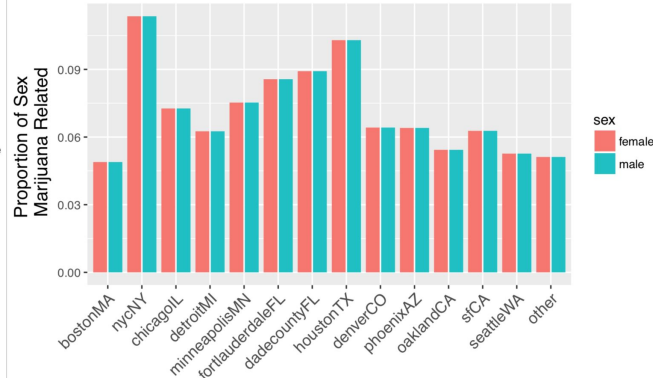# Marijuana Related Visits By Age and Sex
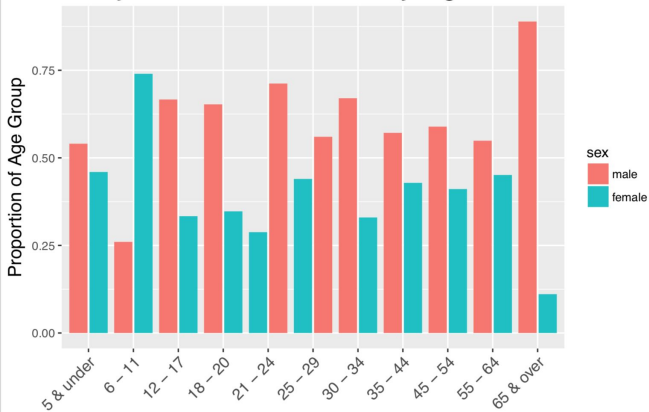
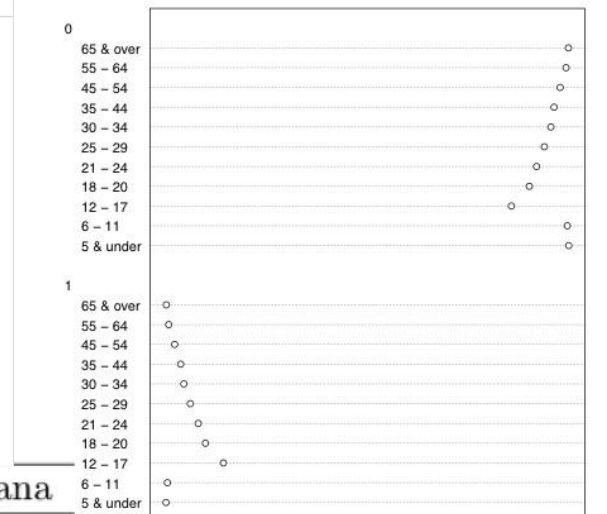# Marijuana Related Visits by Race

Marijuana Related Visits By Metro Area
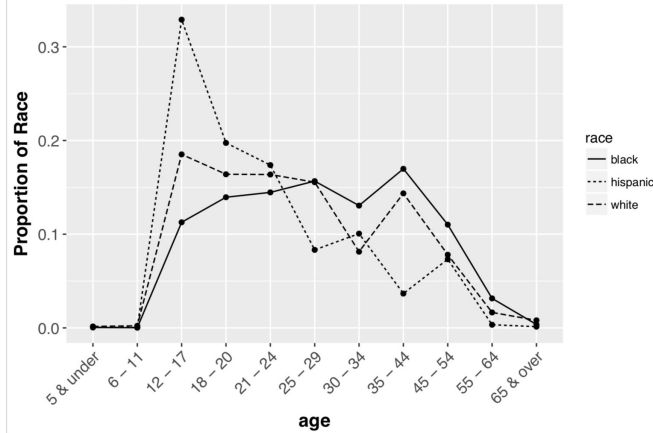
Marijuana Use Normalized by Number of People Per Age Group

Marijuana Related Visits by Sex and Race

Marijuana Related Visits By Metro Area and Sex

	isMarijuana
0	95.96
1	4.04

1. Collect tables and plots.

Marijuana Related Visits By Age and Sex

Marijuana Related Visits By Metro Area

Marijuana Related Visits by Race

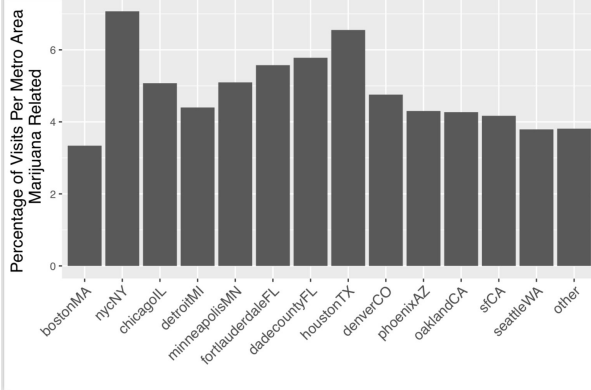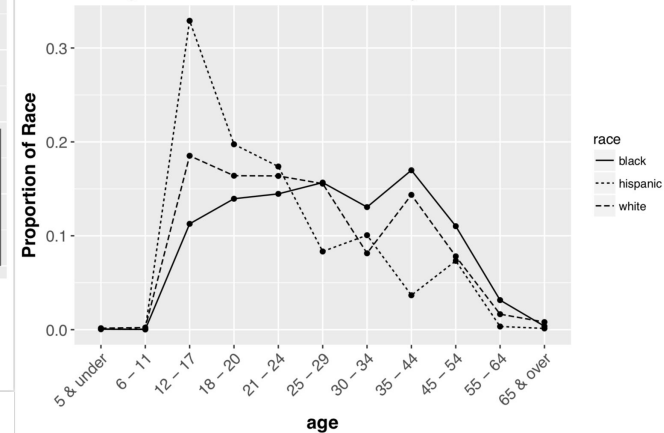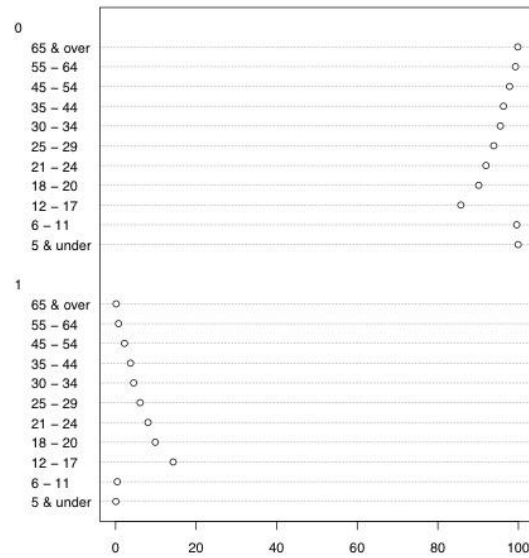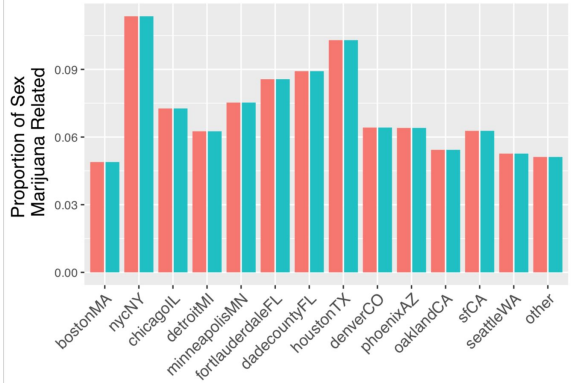## Marijuana Use Normalized by Number of People Per Age Group

Marijuana Related Visits By Metro Area and Sex

Marijuana Related Visits by Sex and Race

isMarijuana	
0	95.96
1	4.04

2. Group related findings.

3. Make an argument (find the story).

Although marijuana-related visits to the emergency room do not make up a large proportion of drug related visits, we are still interested in who is affected. Are there demographic discrepancies that may inform future prevention policies?

Marijuana Related Visits By Age and Sex

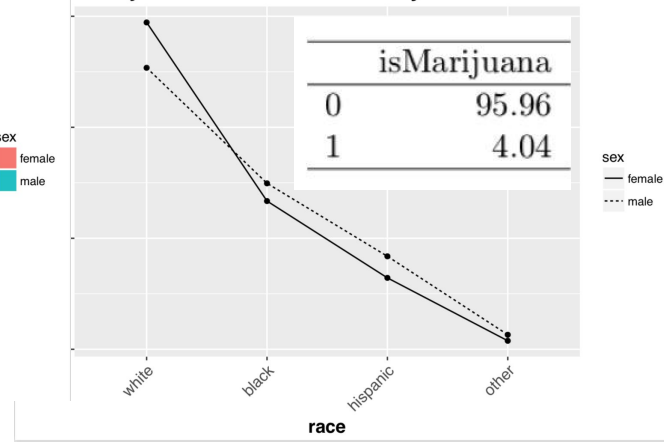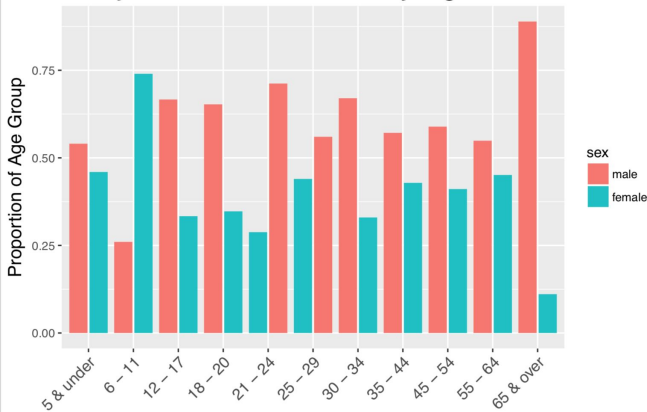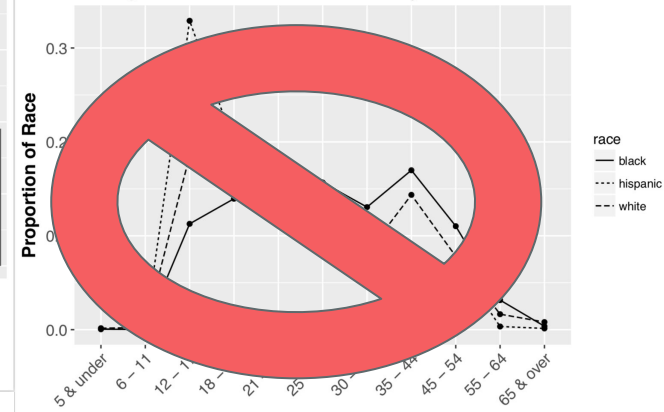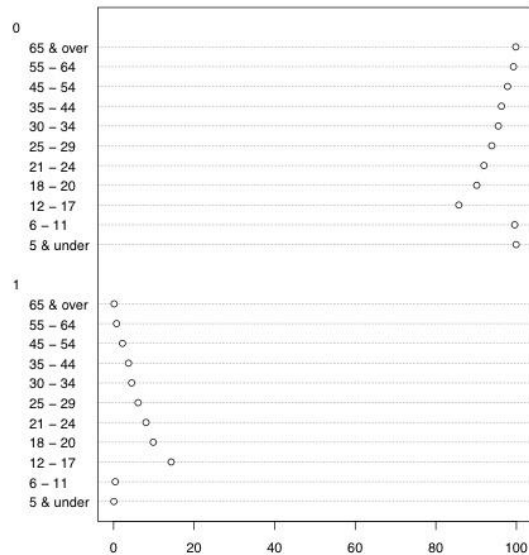# Marijuana Related Visits By Metro Area

Marijuana Related Visits by Race

Marijuana Use Normalized by Number of People Per Age Group

Marijuana Related Visits By Metro Area

Marijuana Related Visits by Sex and Race

isMarijuana	
0	95.96
1	4.04

4. Choose the tables and plots needed to tell your story.

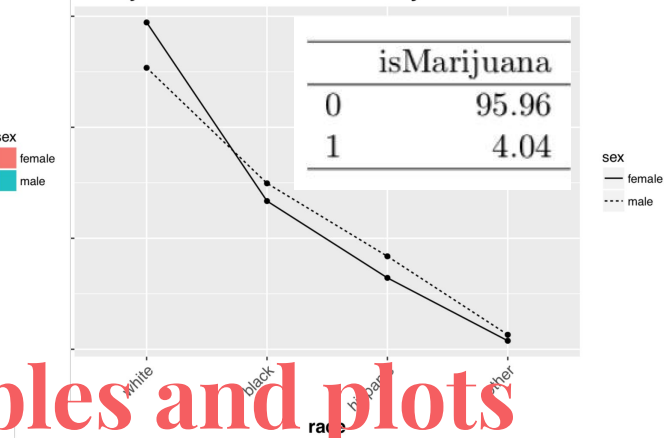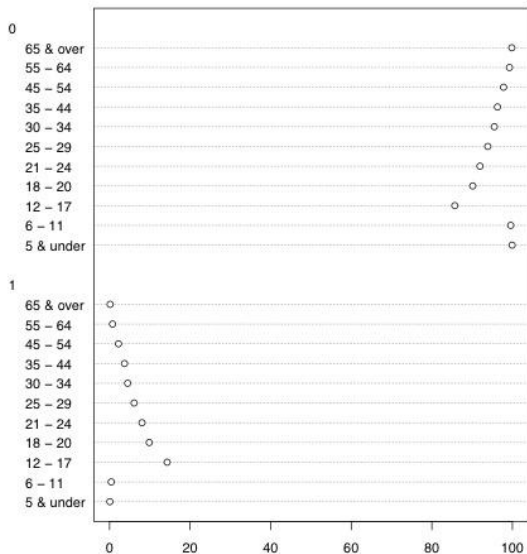| isMarijuana	
0	95.96
1	4.04

#1

Marijuana Use Normalized by Number of People Per Age Group

#2

Marijuana Related Visits By Age and Sex

#3

Marijuana Related Visits by Sex and Race
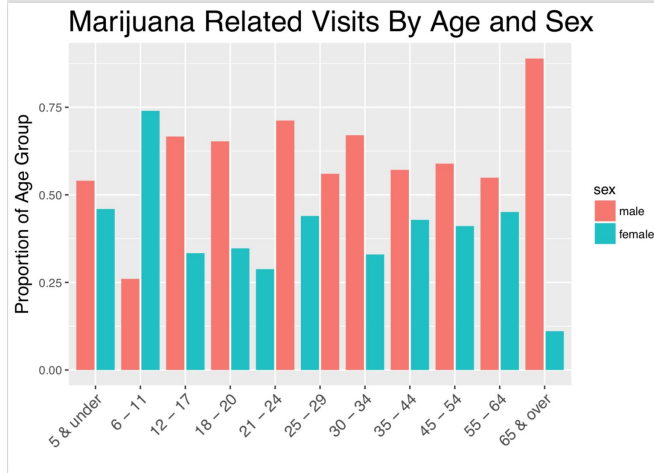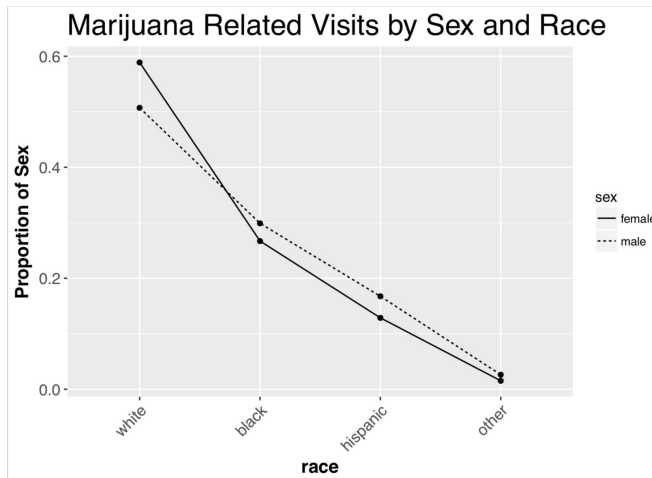
#4

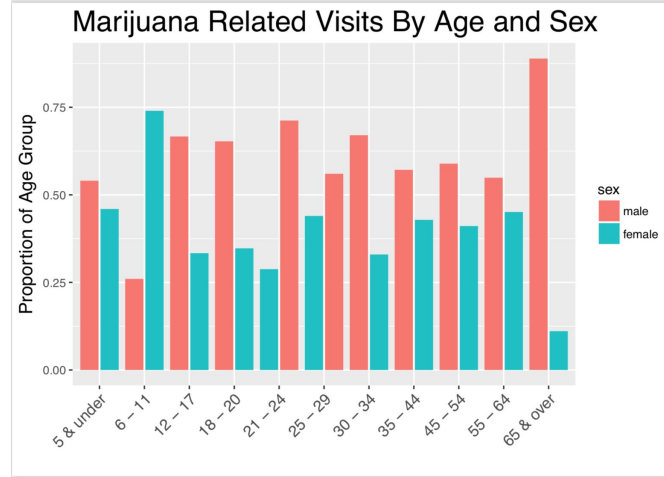5. Sequence the chosen tables and plots.

isMarijuana	
0	95.96
1	4.04

Of drug-related ER visits, those dealing with marijuana are few.


Marijuana Use Normalized by Number of People Per Age Group
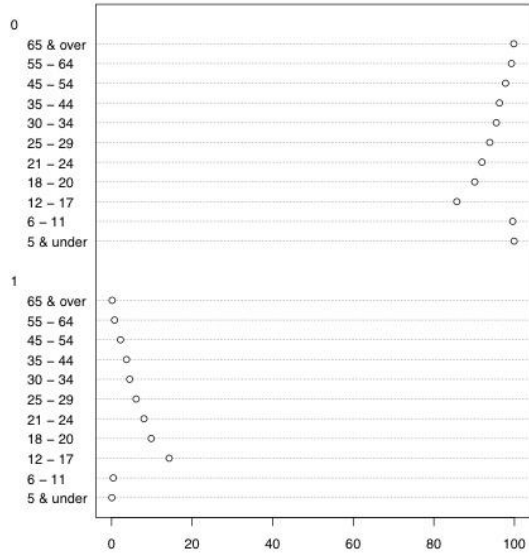
Marijuana is most prevalent in the 12-17 year old ER patients and steadily declines with age.

6. Add captions and transitions.

Marijuana Related Visits By Age and Sex

Do the age trends differ between gender? Yes! Females have their greatest percentage of marijuana incidents in their late 20s while males have their peak much earlier (12-17).

There are noticeable differences between genders across all age groups. More men than women come into the ER having used marijuana.


Marijuana Related Visits by Sex and Race
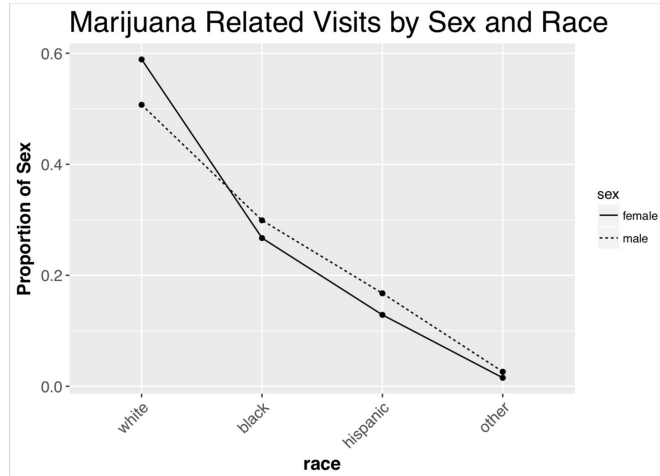
Are the differences between gender consistent across race? Yes! The gap between males and females does narrow from white to black to Hispanic but not dramatically.

7. Iterate.

- Are there any gaps in the argument?
  - Acknowledge in text but don't explicitly fill.
  - Patch the gap by updating a plot or piece of the analysis.
  - Rework the analysis more thoroughly. Is there a plot or finding missing?
- Can a peer understand the main takeaways from the storyboard?
  - What pieces are confusing?
  - What questions do they still have after browsing?
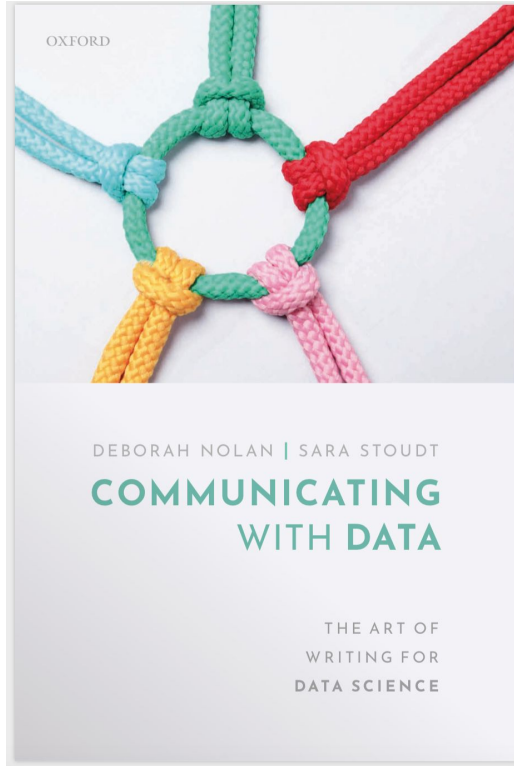
Want more details and examples? Join us on gather.town

**Thursday, July 1st
2:45pm-3:30pm ET**
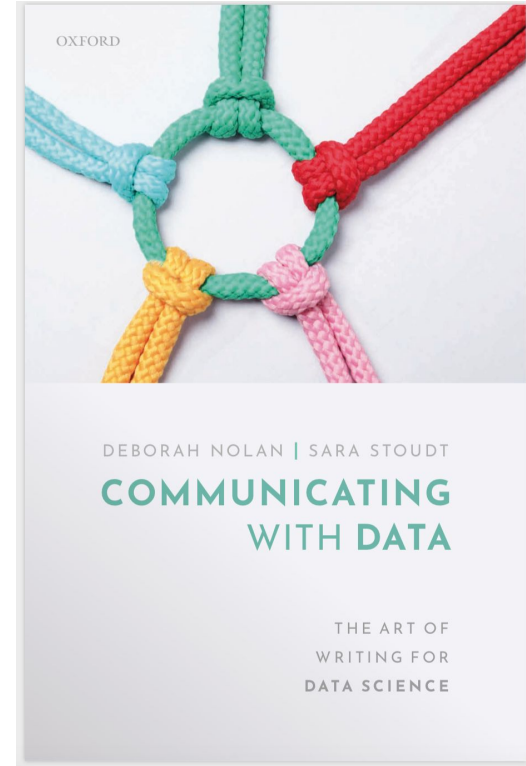
Note: The idea for storyboarding our research was inspired by an activity led by Sara ElShafie in the Data Science for the 21st Century NSF Training Program Science Communication Short Course at UC Berkeley taken by Sara Stoudt.