

Meetup 1: Developing a Critical Eye

George Hagstrom



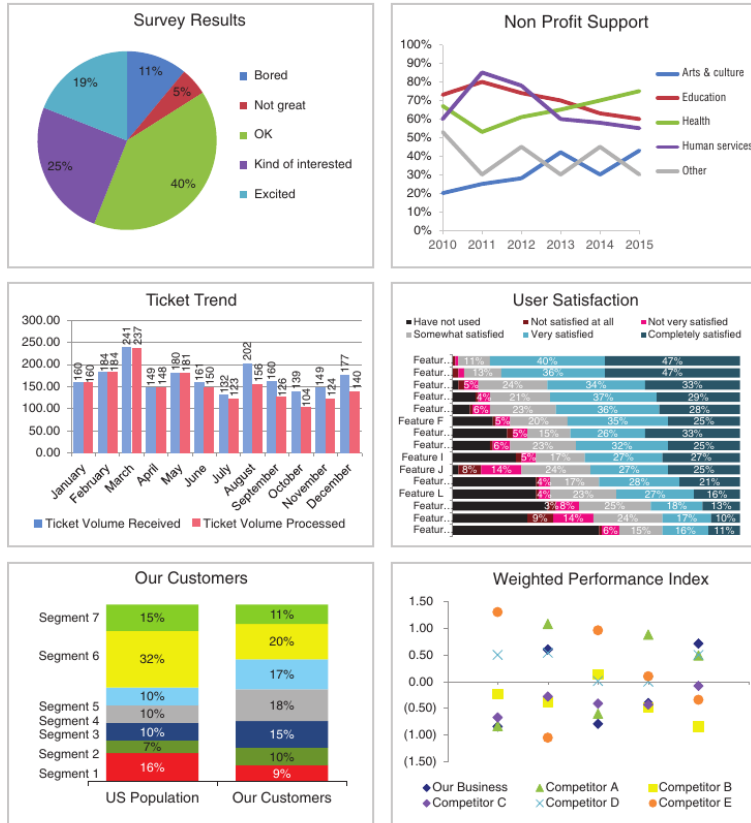
Why Visualizations Matter

- Graphs are the main way we communicate data
- Your visualizations are how people will remember your analysis, can make or break your recommendations to decision makers.
- Two goals of the class:
 1. Learn how to tell stories with graphs
 2. Learn to make good visualizations



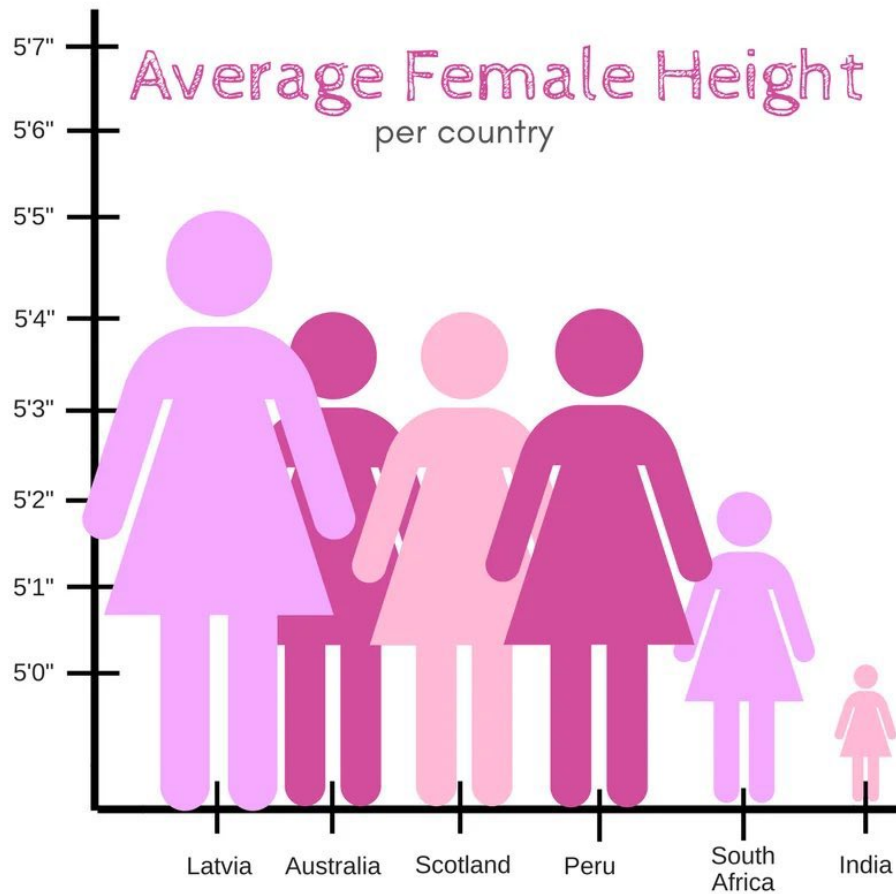
Why do we need a class on this?

- People are not naturally good at telling stories using data



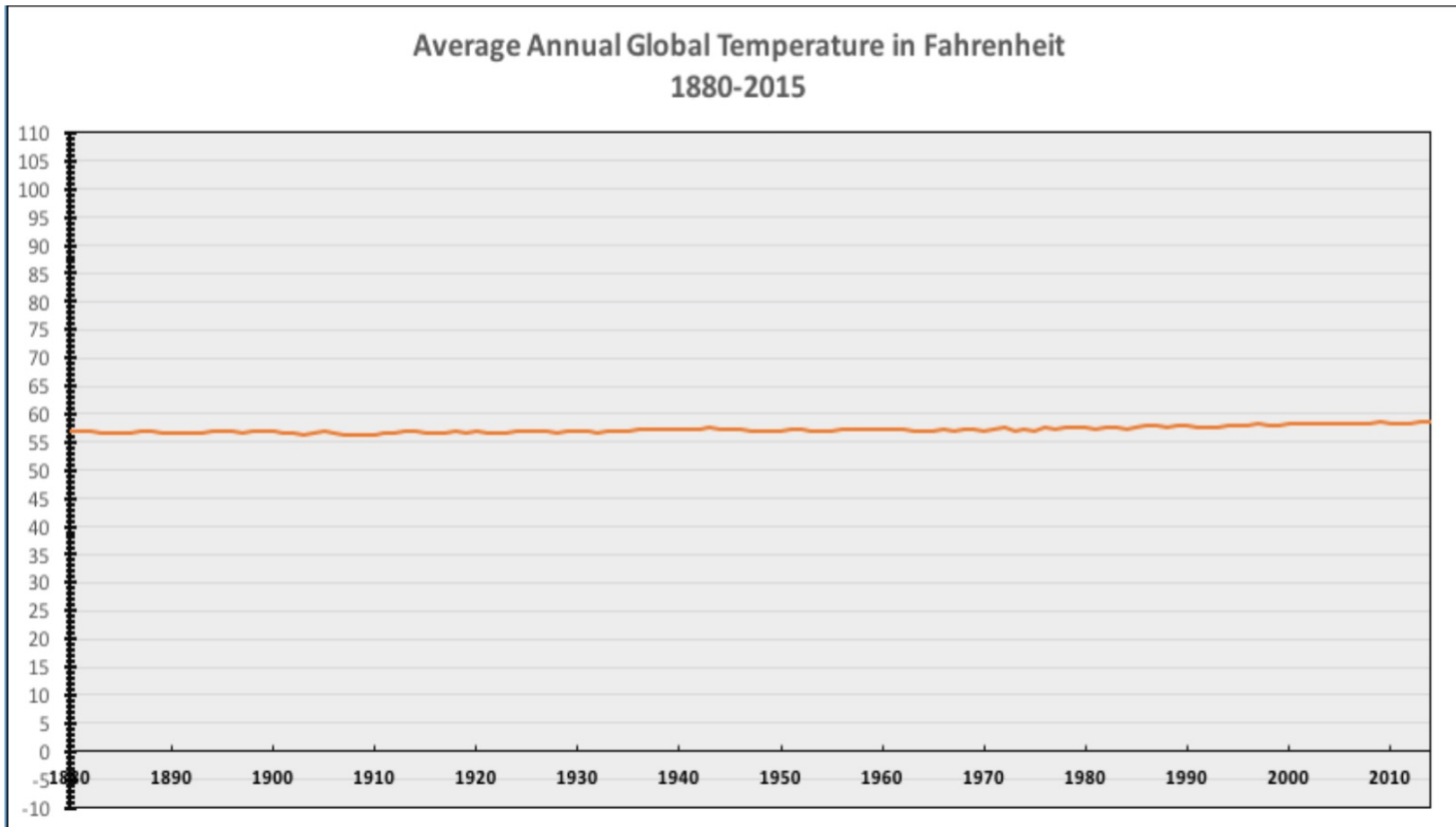
Why do we need a class on this?

- The world is full of awful visualizations



Why do we need a class on this?

- The world is full of awful visualizations

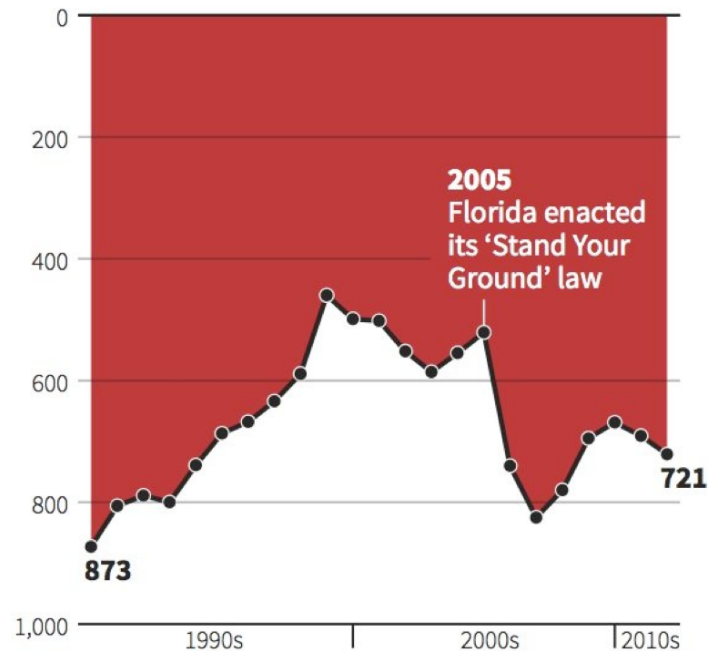


Why do we need a class on this?

- The world is full of awful visualizations

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

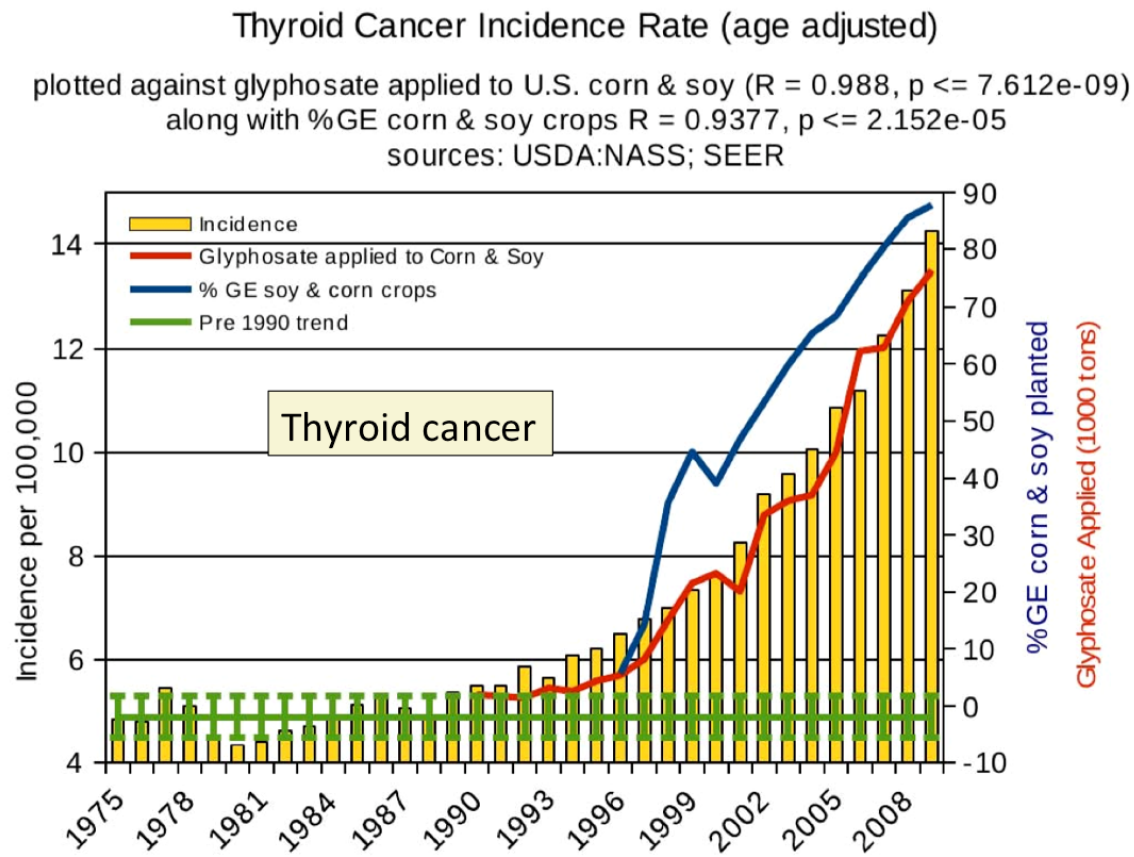
C. Chan 16/02/2014

REUTERS



Why do we need a class on this?

- The world is full of awful visualizations

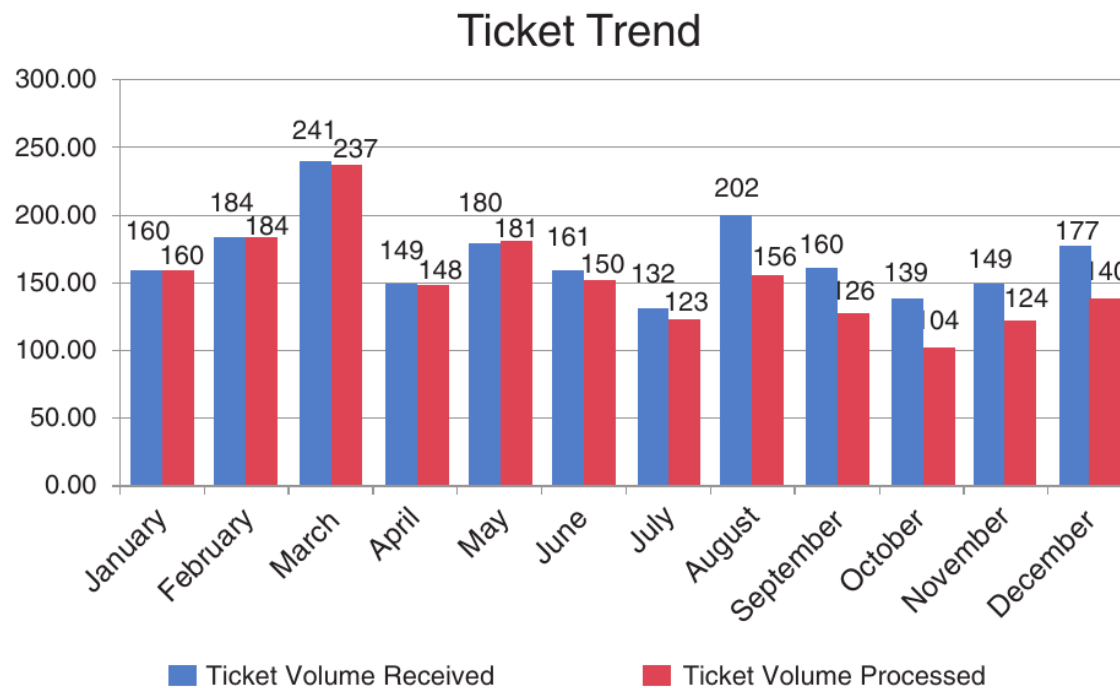


*Figure 10, Swanson et al. Journal of Organic Systems 2014; 9(2):6-37.



Case Study

- Your firm had two employees leave in May
- This caused a backlog of work



Case Study

- Highlight the key features to tell a story

Please approve the hire of 2 FTEs

to backfill those who quit in the past year

Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

Storytelling 0.3



Course Website and Meetup

- Brightspace to submit homework
- Course Meetup: Monday 8:00-9:00PM
- [Zoom Link](#)



Slack Channel

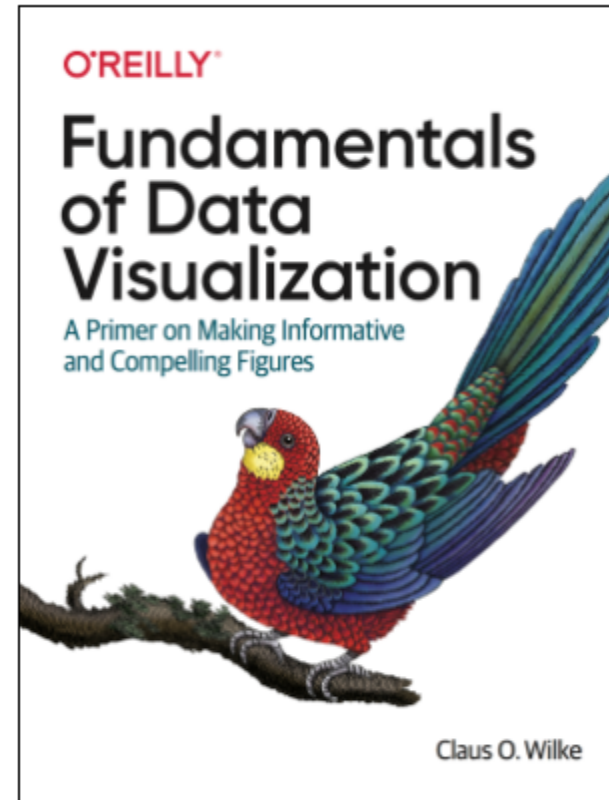
- Course slack channel for messaging and rapid communications
- <https://data608-spring-2025.slack.com>
- [Invite Link](#)



Textbooks

Two most important textbooks:

1. [Fundamentals of Data Visualization](#)
2. [Storytelling with data](#)

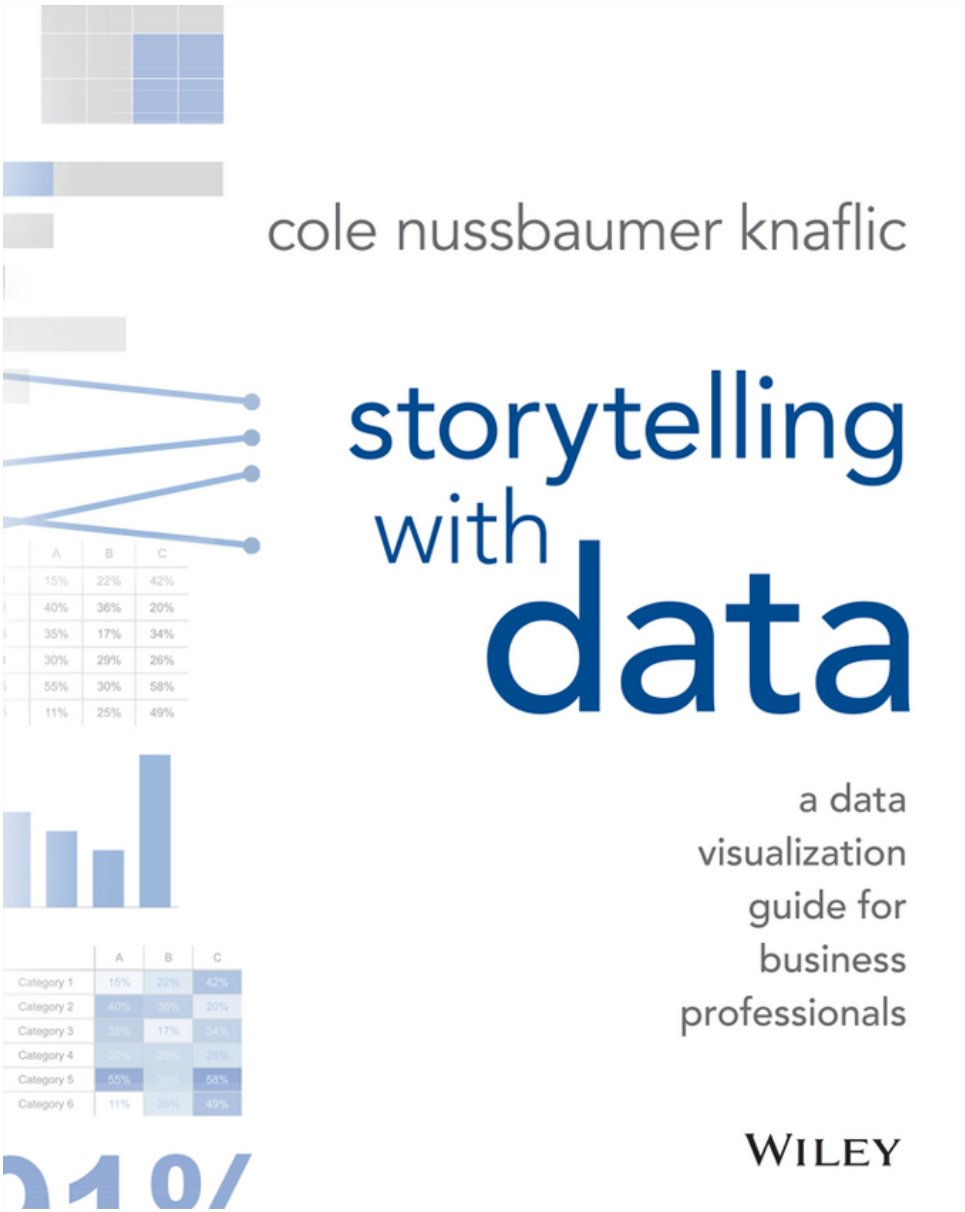


Textbooks



Two most important textbooks:

- 1. Fundamentals of Data Visualization
- 2. Storytelling with data



Assignments

- 8 Lab Assignments (80%)
 - Submit a quarto slideshow and rendered slides to brightspace
 - Language freedom, R or python recommended
- Participation/Discussions (20%)
 - Usually one per week



Week Plan: Readings

- Chapter 1 of Storytelling
- Chapters 1 and 2 of Fundamentals of Dataviz
- First 11 slides of Data Practitioner's Guide to Storytelling with Data
- For fun: Two sets of History of Dataviz slides



Week Plan: Assignments

- Elevator Pitch Intro
- Bad Figure (Critical Eye) Discussion
 - By Friday to give classmates a chance to vote on the worst figure
- Due in two weeks: Data Story 1



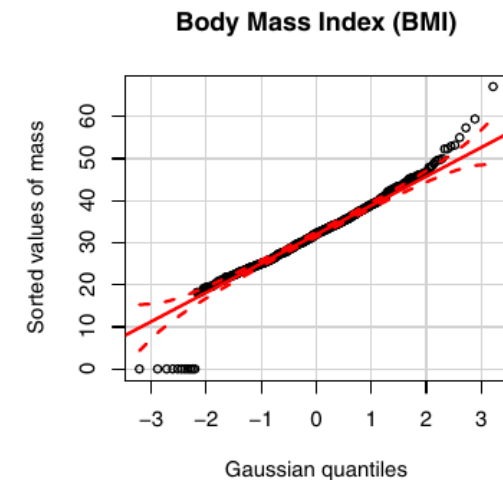
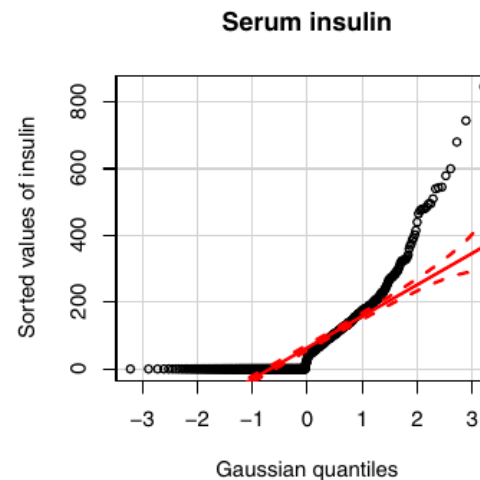
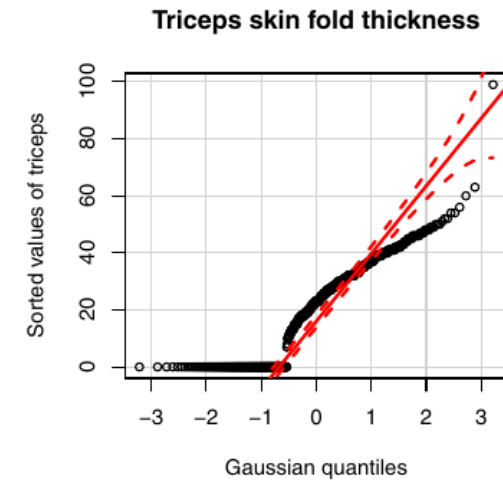
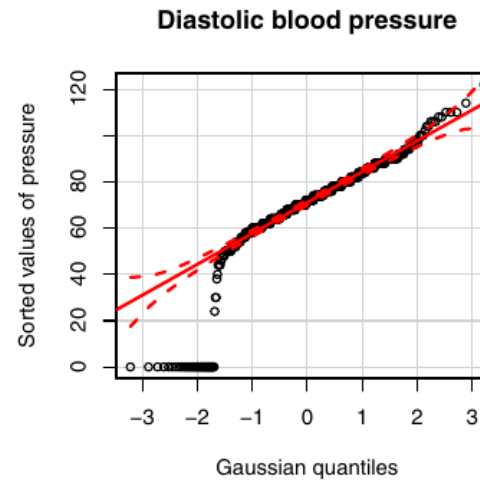
Week Plan: Key Concepts

- Methods for creating a story (Storytelling)
- What makes a quality figure (Wilke, Practitioner's Guide)



Context: Purposes of Visualizations

- Exploratory
- Explanatory
- Inferential



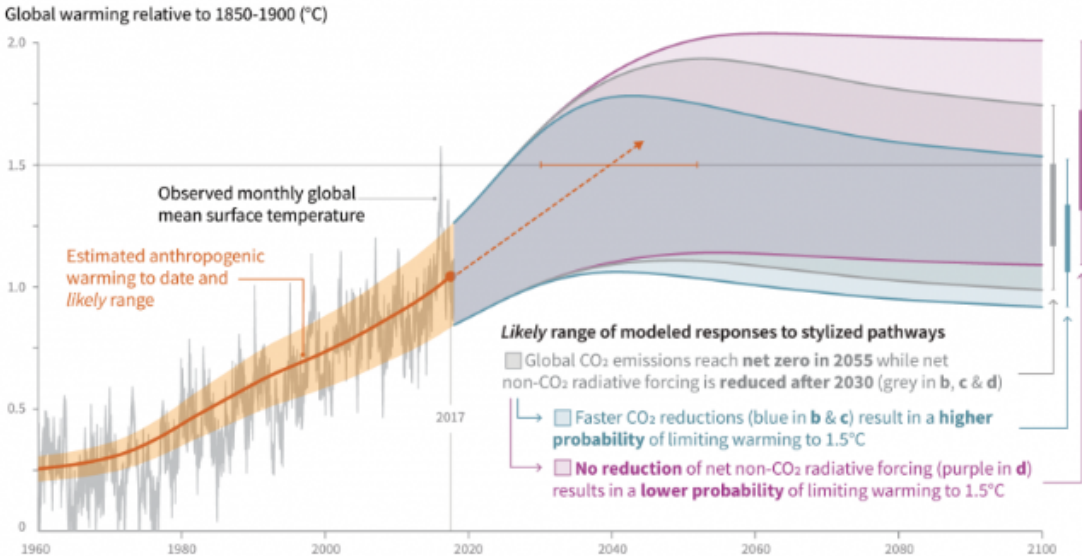
Context: Purposes of Visualizations



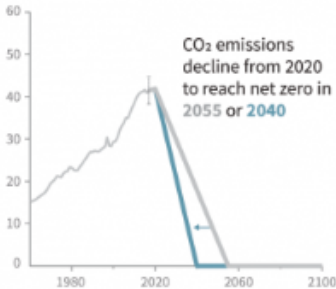
- Exploratory
- Explanatory
- Inferential

Cumulative emissions of CO₂ and future non-CO₂ radiative forcing determine the probability of limiting warming to 1.5°C

a) Observed global temperature change and modeled responses to stylized anthropogenic emission and forcing pathways

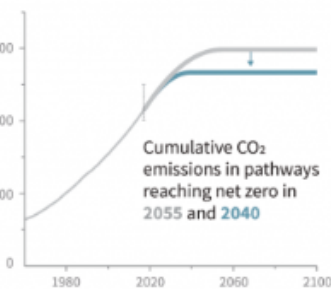


b) Stylized net global CO₂ emission pathways
Billion tonnes CO₂ per year (GtCO₂/yr)



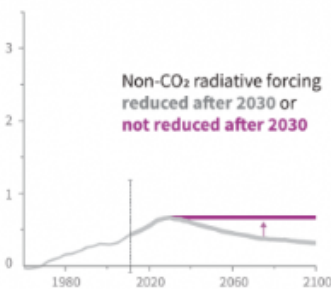
Faster immediate CO₂ emission reductions limit cumulative CO₂ emissions shown in panel (c).

c) Cumulative net CO₂ emissions
Billion tonnes CO₂ (GtCO₂)



Maximum temperature rise is determined by cumulative net CO₂ emissions and net non-CO₂ radiative forcing due to methane, nitrous oxide, aerosols and other anthropogenic forcing agents.

d) Non-CO₂ radiative forcing pathways
Watts per square metre (W/m²)



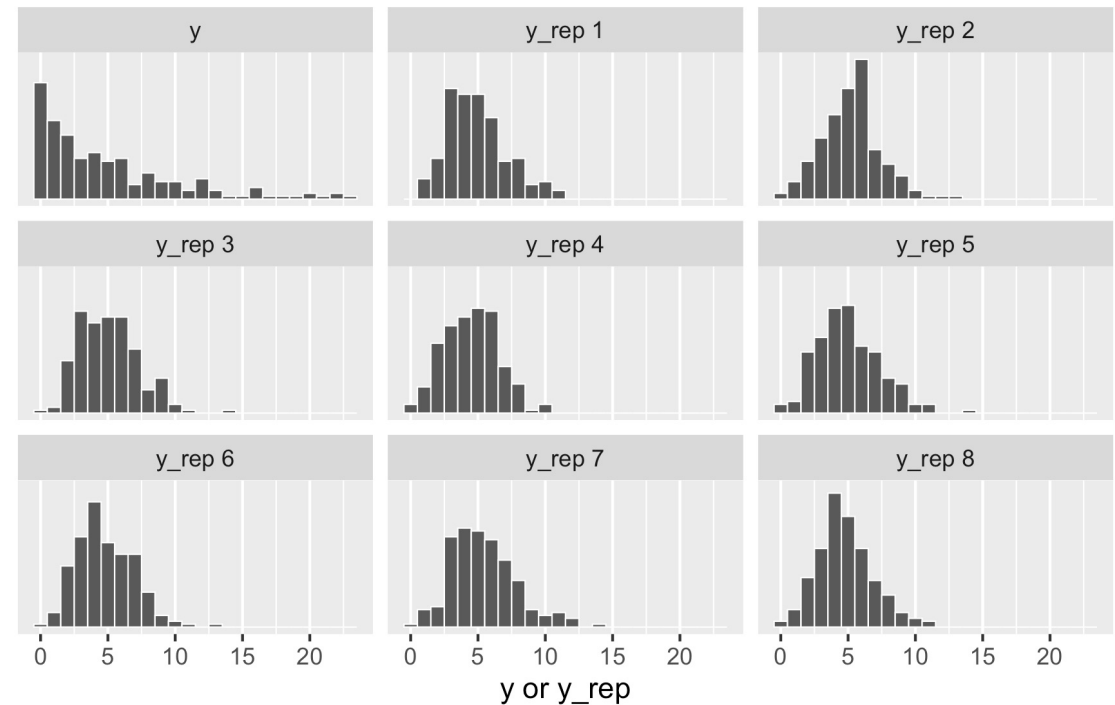
IPCC SPM1 Special Report on 1.5C



Context: Purposes of Visualizations

- Exploratory
- Explanatory
- Inferential

Negative binomial data, Poisson model



Context: Who, What, and How

Before you begin, start with three questions

- Who will see your figure?
 - You would make different choices presenting to experts, decision makers/stakeholders, your supervisor, or the general public



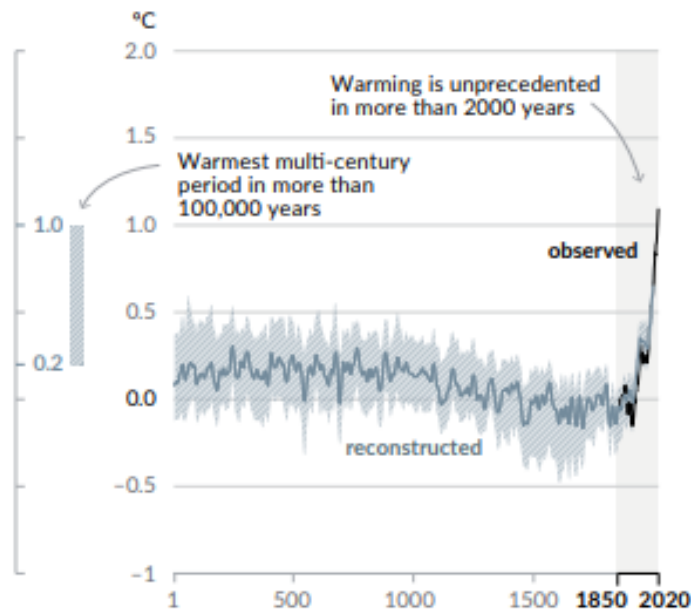
Context: Who, What, and How

- Figure for Decision Makers (IPCC SPM):

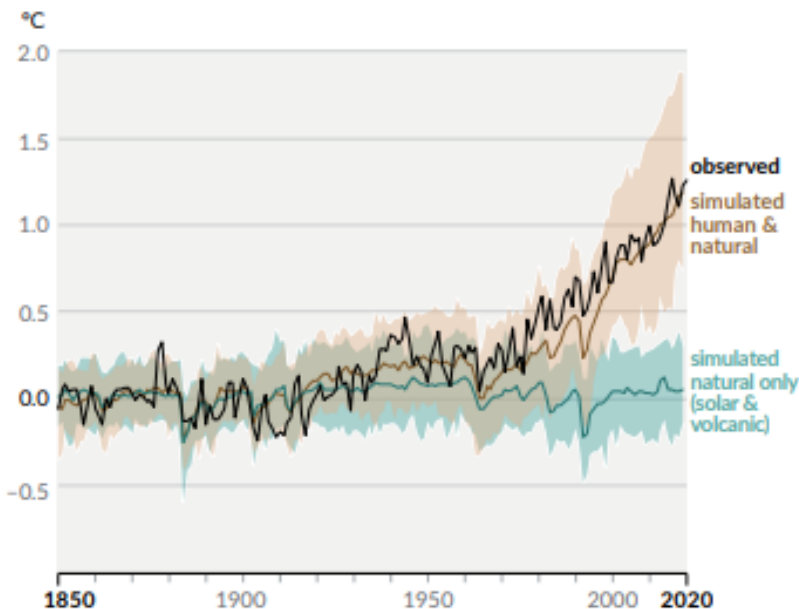
Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

Changes in global surface temperature relative to 1850–1900

(a) Change in global surface temperature (decadal average) as **reconstructed** (1–2000) and **observed** (1850–2020)



(b) Change in global surface temperature (annual average) as **observed** and simulated using **human & natural** and **only natural** factors (both 1850–2020)

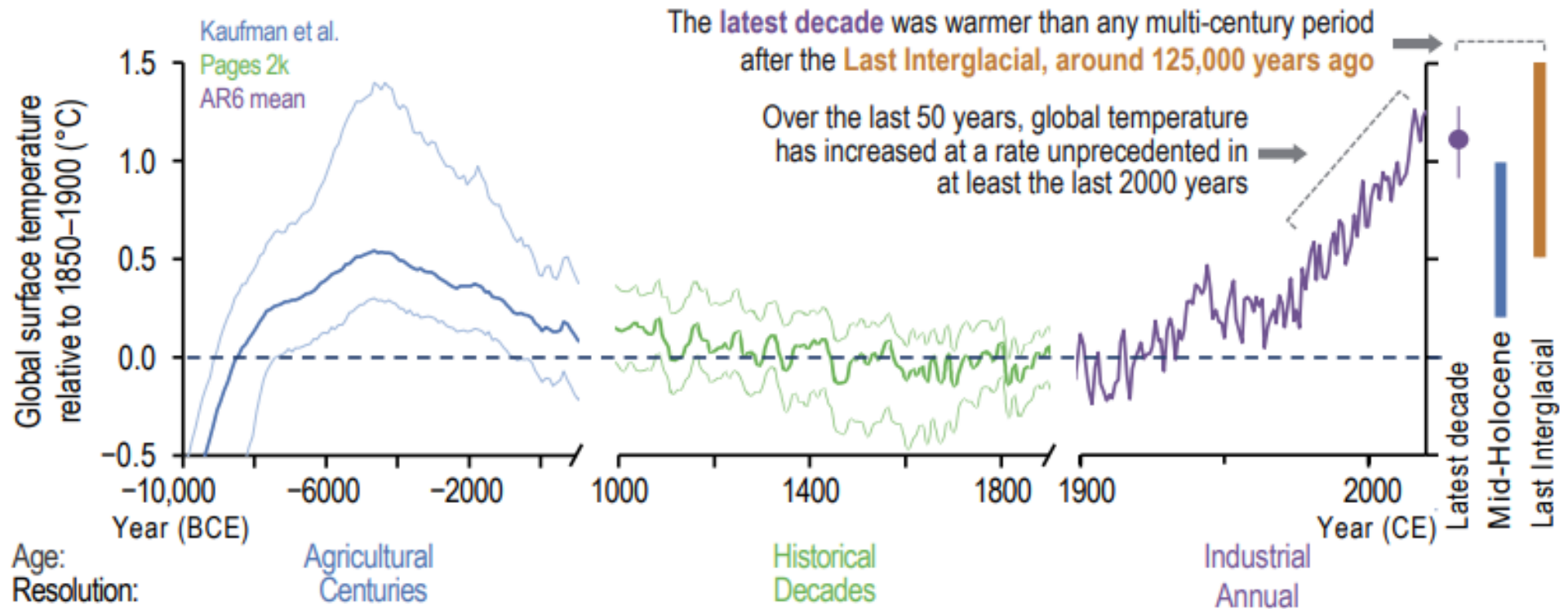


Context: Who, What, and How

- Figure for other scientists (IPCC)

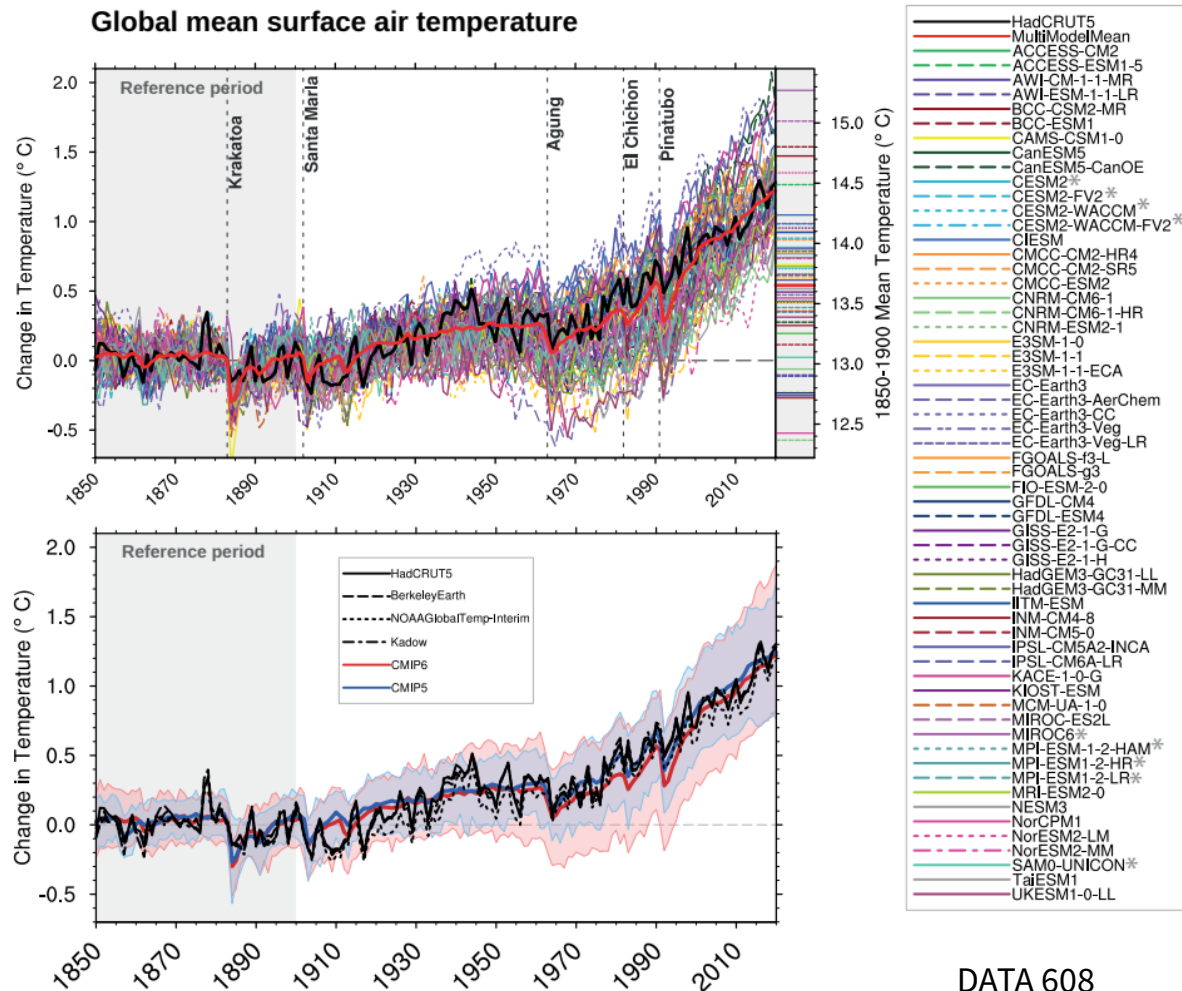
Changes in surface temperature

(a) Global surface temperatures are more likely than not unprecedented in the past 125,000 years



Context: Who, What, and How

- Figure for other scientists (IPCC)



Context: Who, What, and How

Before you begin, start with three questions

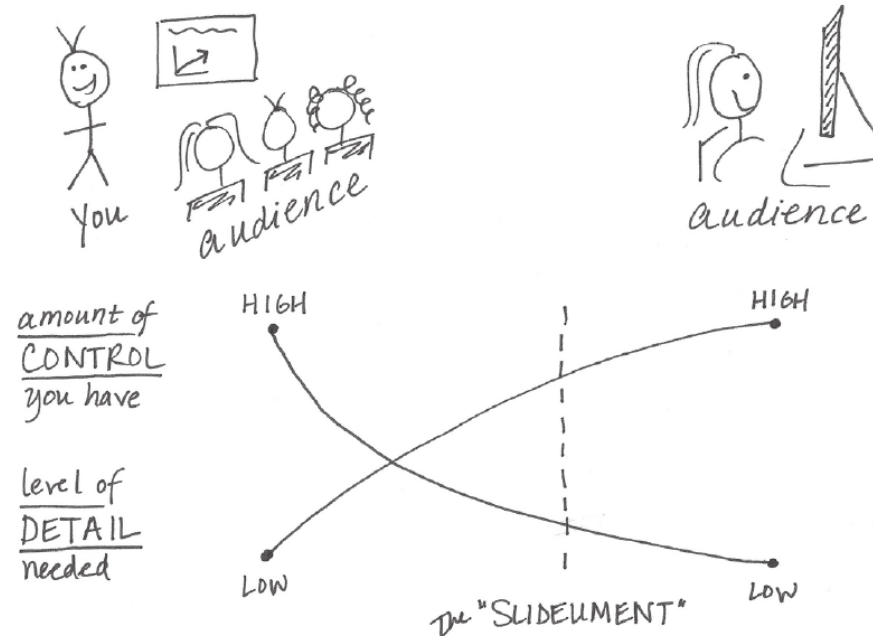
- What do you want your audience to know or do?
 - There needs to be a point or story beyond “this is interesting”



Context: Who, What, and How

- Consider the Mechanism
- More details in written comms than presentations

LIVE PRESENTATION WRITTEN DOC or EMAIL



Storytelling



Context: Who, What, and How

Before you begin, start with three questions

- How?
 - What data will you use to accomplish this?
- Don't hide non-supporting data
 - Trust is built over time but destroyed in an instant
 - Building trust is key for convincing people to take action on your recommendations, in all sorts of settings



The Elevator Pitch and Big Idea

- If you had just 3 minutes to tell your audience what they need to know, what would you say?
 - What is the overall topic/problem (get into it fast)
 - Why is it important (bring emotion into it)
 - Try to find a good punchline to complete the narrative



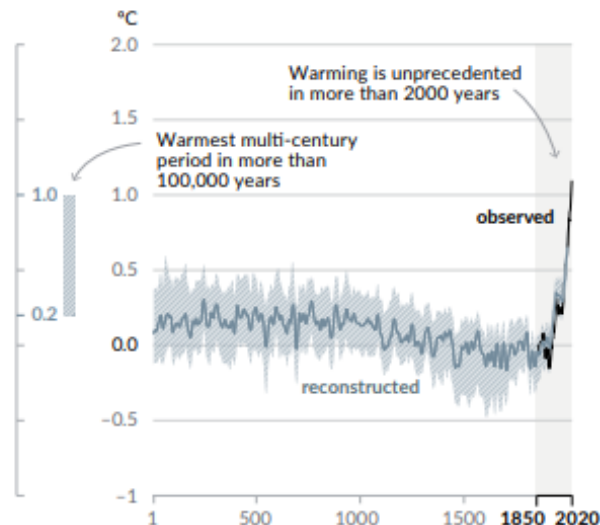
Big Idea: Single Sentence Summary

- Convey your unique point of view
- Must convey what is at stake

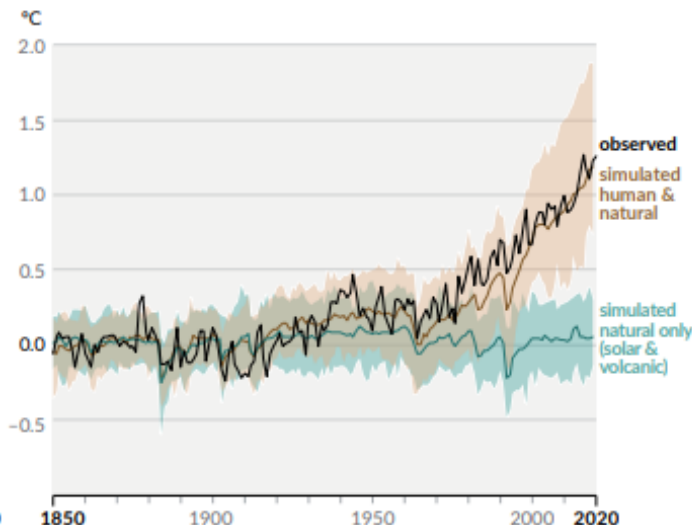
Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

Changes in global surface temperature relative to 1850–1900

(a) Change in global surface temperature (decadal average) as **reconstructed** (1–2000) and **observed** (1850–2020)



(b) Change in global surface temperature (annual average) as **observed** and simulated using **human & natural** and **only natural** factors (both 1850–2020)



The Storyboard

- Storyboard is like a visual outline of the story you want to tell
- Arose in TV, film, and other narrative media
- Sketch out the narrative and visual elements of your data story.
- 2-6 elements
- Problem, Research Question, Data, Insights, Recommendation/Conclusion



The Storyboard

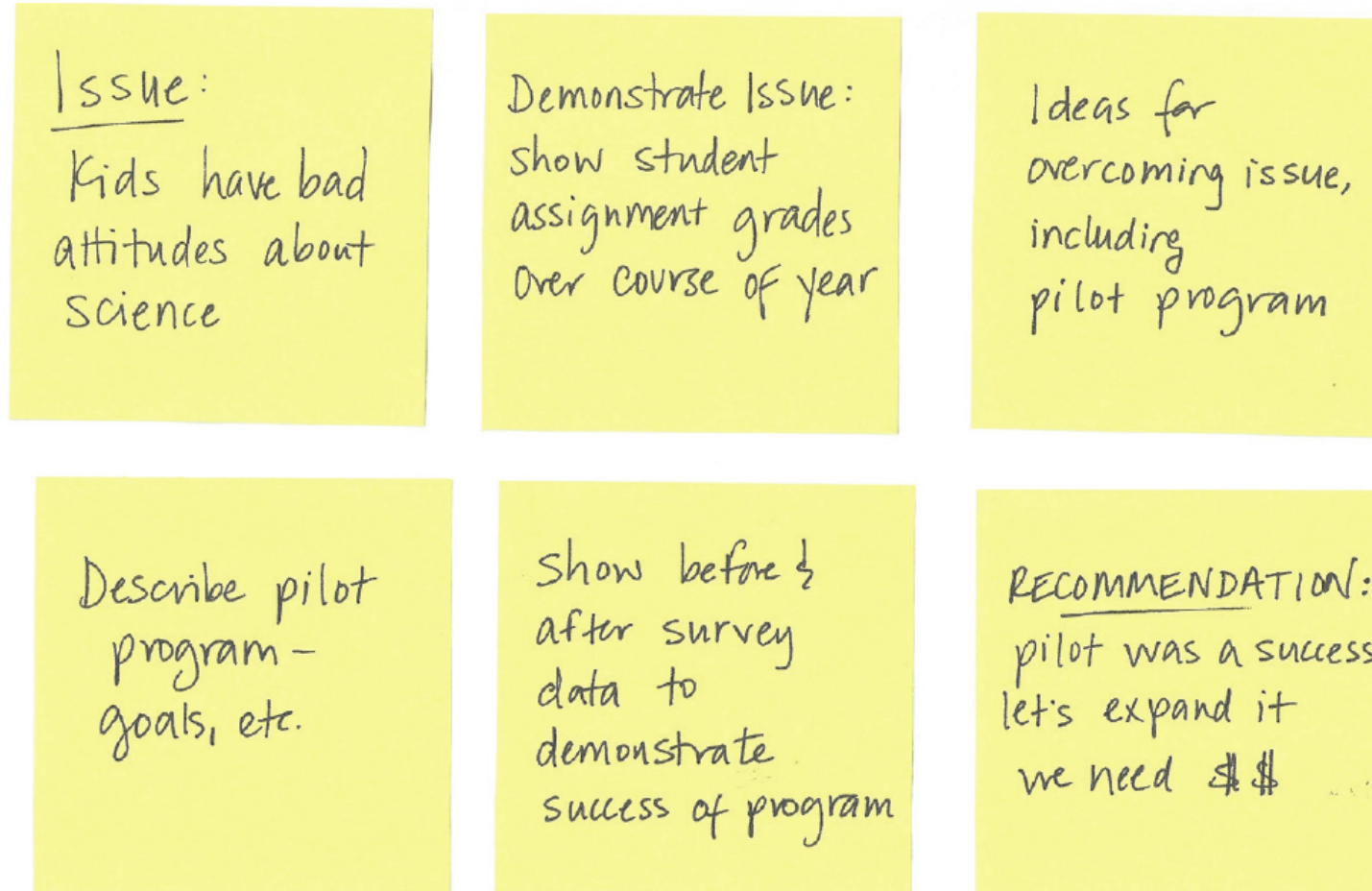


FIGURE 1.2 Example storyboard



Why Stories?

- Humans are wired to hear and tell stories, it is how we relate
- Neuroscience research backs this up
- Good stories create synchrony across multiple brain regions between storyteller and listener

Speaker–listener neural coupling underlies successful communication

[Greg J Stephens](#)^{a,b,1}, [Lauren J Silbert](#)^{c,1}, [Uri Hasson](#)^{c,d,2}

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#)

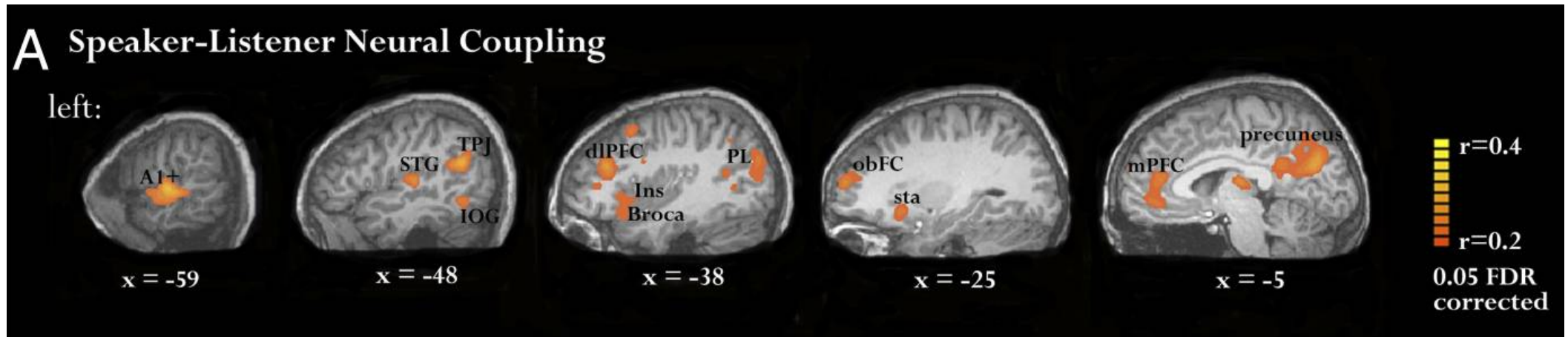
PMCID: PMC2922522 PMID: [20660768](#)

Abstract

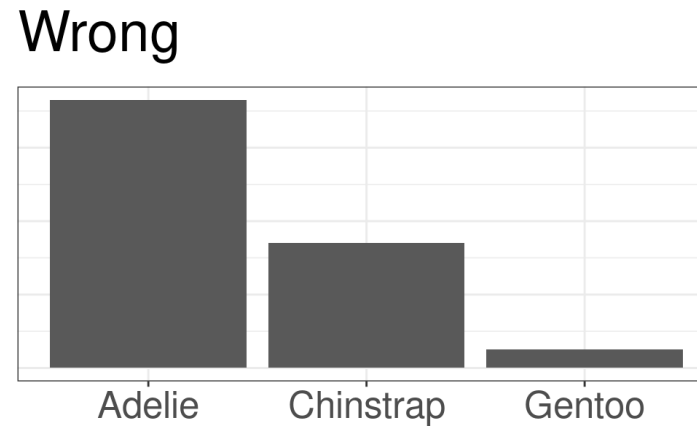
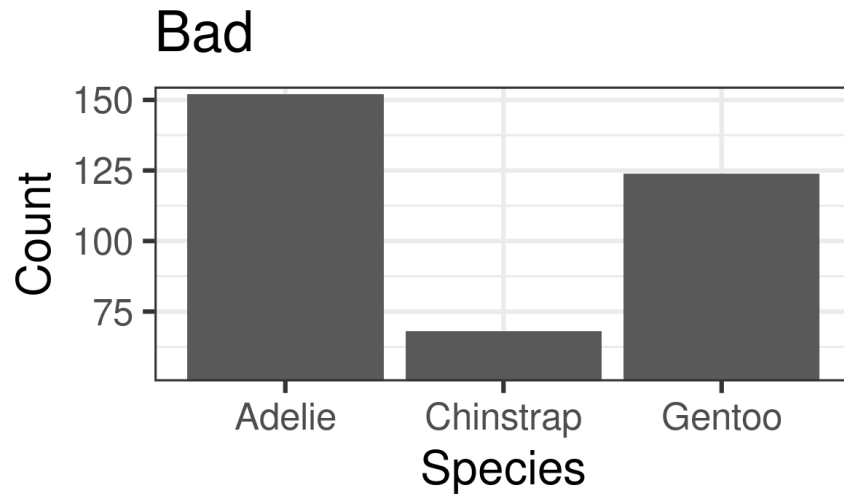
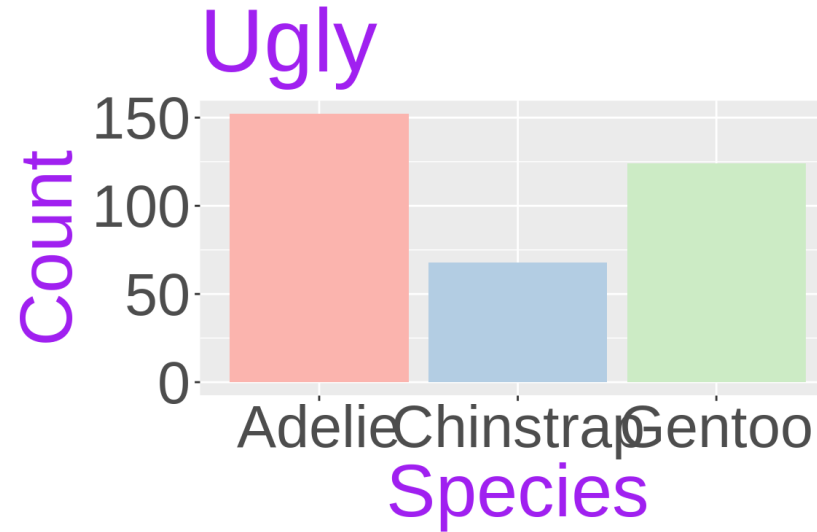
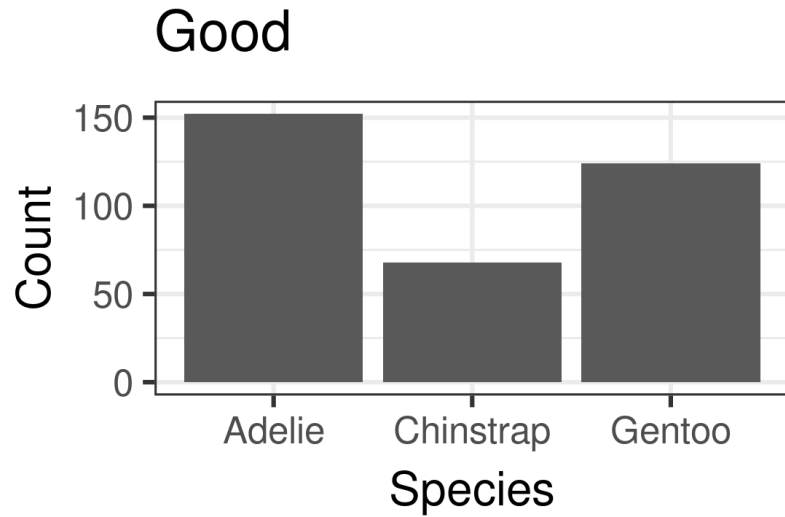


Why Stories?

- Humans are wired to hear and tell stories, it is how we relate
- Neuroscience research backs this up
- Good stories create synchrony across multiple brain regions between storyteller and listener



Bad, Ugly, and Wrong Figures



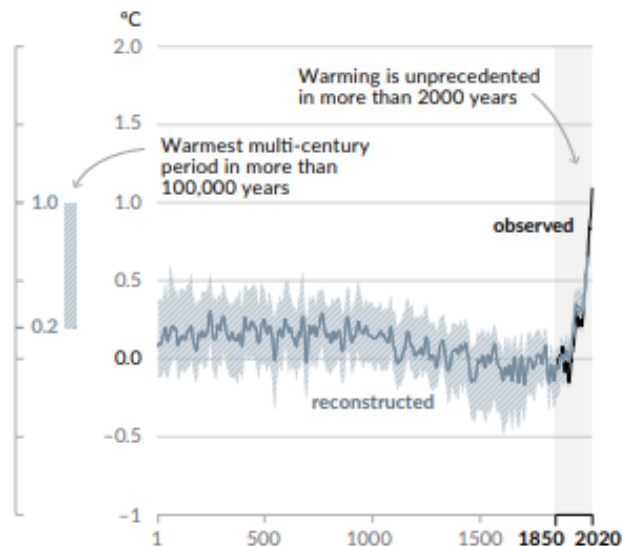
Visualization Quality Features

1. **Fidelity:** Should be a faithful representation of the underlying data, nothing added or hidden
2. **Simplicity:** As simple as possible but no simpler. No irrelevant embellishments

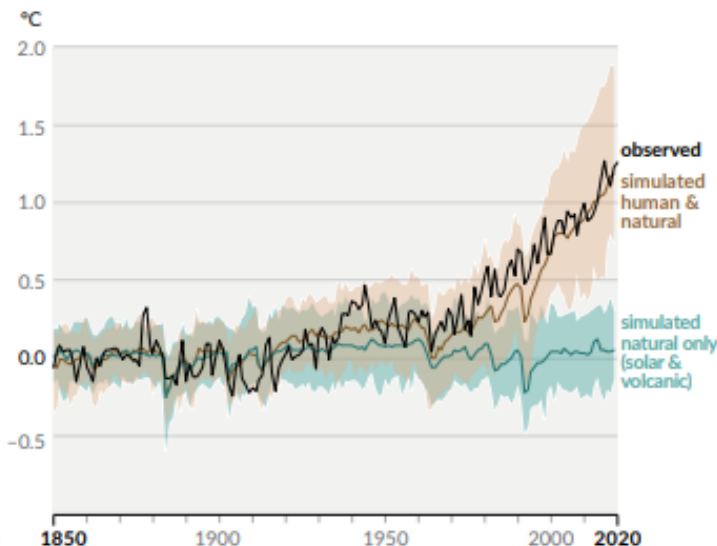
Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

Changes in global surface temperature relative to 1850–1900

(a) Change in global surface temperature (decadal average) as reconstructed (1–2000) and observed (1850–2020)



(b) Change in global surface temperature (annual average) as observed and simulated using human & natural and only natural factors (both 1850–2020)



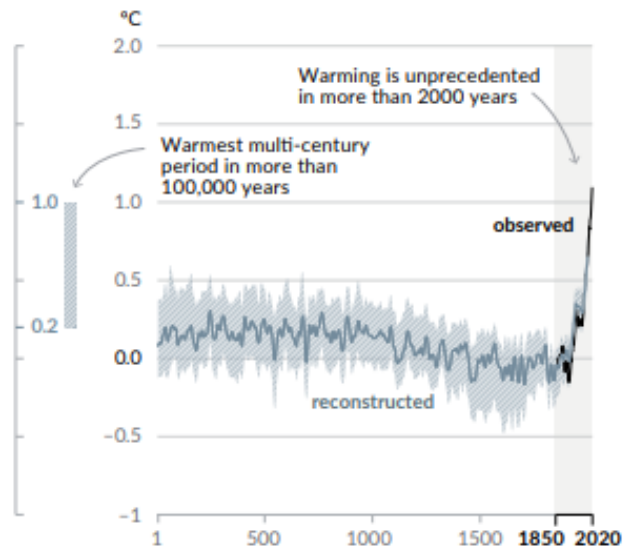
Visualization Quality Features

3. **Utility:** Fit for purpose as defined by the viewers needs. No extraneous information
4. **Saliency:** Focuses attention on the most important facts and conclusions

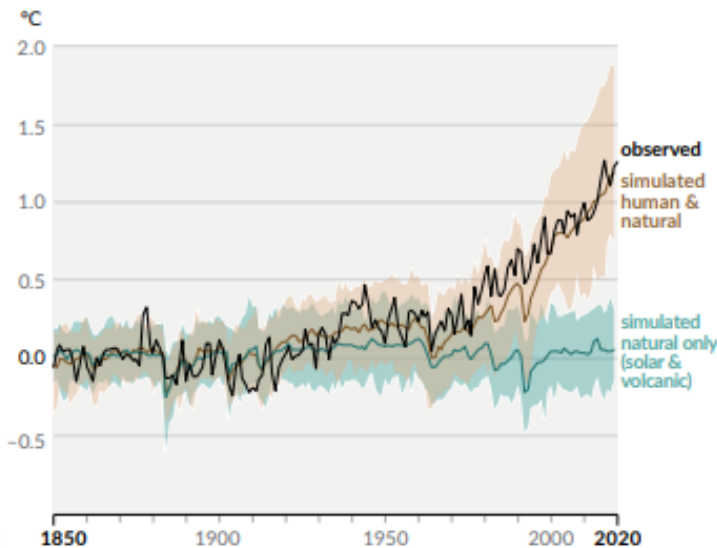
Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

Changes in global surface temperature relative to 1850–1900

(a) Change in global surface temperature (decadal average) as **reconstructed** (1–2000) and **observed** (1850–2020)



(b) Change in global surface temperature (annual average) as **observed** and simulated using **human & natural** and **only natural** factors (both 1850–2020)



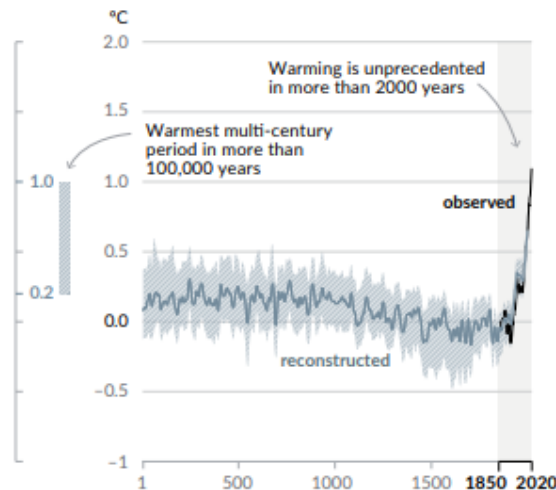
Visualization Quality Features

- 5. **Efficacy:** Accomplishes the purpose for which it was designed, in context
- 6. **Uniformity:** Adheres to a predetermined presentation theme
- 7. **Amity:** User friendly, takes minimal effort to grasp the main message

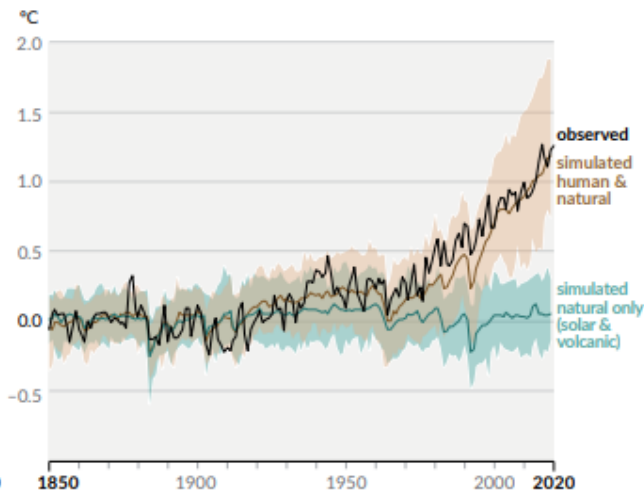
Human influence has warmed the climate at a rate that is unprecedented in at least the last 2000 years

Changes in global surface temperature relative to 1850–1900

(a) Change in global surface temperature (decadal average) as **reconstructed** (1–2000) and **observed** (1850–2020)



(b) Change in global surface temperature (annual average) as **observed** and simulated using **human & natural** and **only natural** factors (both 1850–2020)



Advert



- Once per month (Tuesday, Wednesday, Thursday) “field trips” to New York Open Statistical Computing Meetup
- Send me a DM I’ll add you to the email list
- nyhackr.org



About nyhackr

Founded by [Josh Reich](#) and [Drew Conway](#), the [New York Open Statistical Programming Meetup](#) started as the New York R Meetup with a handful of people in an office at [Union Square Ventures](#). Since then it has grown to over 12,000 members and has been hosted at [NYU](#), [Columbia](#), [AOL](#), [iHeartRadio](#), [eBay](#), [Work-Bench](#) and other locations.

Our mission is to spread knowledge of statistical programming techniques in open-source languages such as R, Python, Julia and Go, and data science in general. Another important aspect is community building and socializing. The meetups start with pizza, followed by a 45-90 minute talk, ending with a trip to the local bar.

Jobs

Job openings and other announcements are on our [Slack channel](#) and on the [Meetup message board](#).



Thanks!

