Data Visualization Final Project Process Book

Jefrey Alexander, Janavi Kumar, and Peter Silverstein

IDEA CONCEPTION

We wanted to use this final project as a way to explore current events and public policy. Gaining inspiration from in-class examples of pairing data visualizations with lexical storytelling and text analysis, Jefrey floated the idea of somehow tracking specific words targeted by the current administration. We then decided to expand our exploration by including other forms of censorship pre-dating the past year.

ROUGH DRAFT & MOCK UPS (on next page)
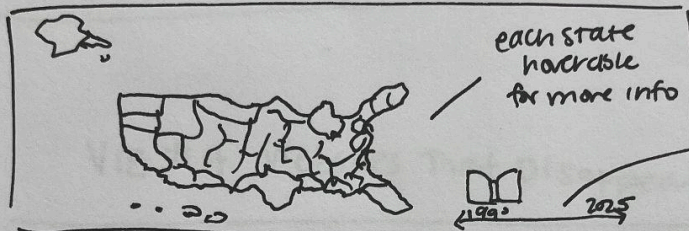
America, Censored
↓

~~America, Censored~~

) time-based animation on load

Intro paragraph on redaction/ censorship, state of things, etc.

↓ ← include cues for users to know to scroll?
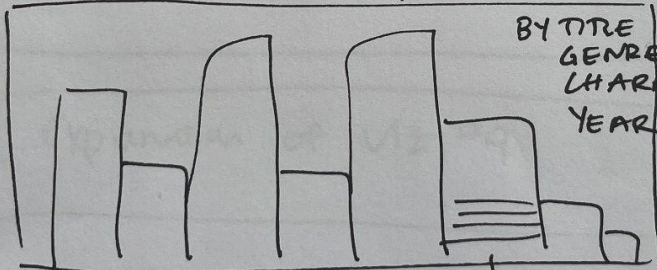
Viz #1 - PEN America data - book bans over time

each state hoverable for more info

bar w/ years (stylize if possible)

∙∙ ২০   ⟨199°   2025

explainer about viz #1

↓

Viz #2 - Most Commonly Banned Books

BY TITLE
GENRE
CHARACTER TYPE } toggleable
YEAR

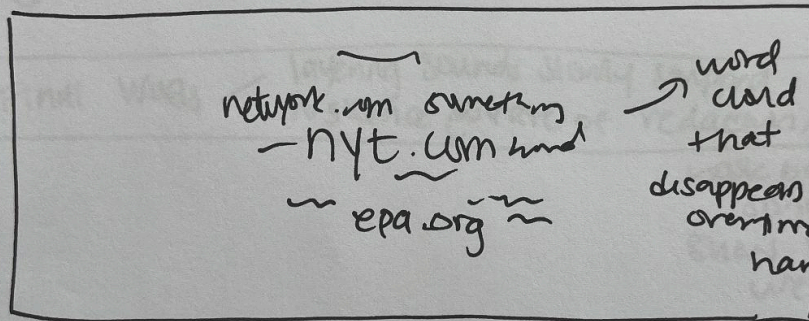→ bars made of stacked books

explainer about viz #2

vertical scrollable interface

Viz #3 - Gizmodo/Archive of NYT Words Not Allowed in
Research - Bubble Chart _or_ bar graph? ← don't
want too
many

← non-graphed bubbles
showing frequency/size
of removed
words
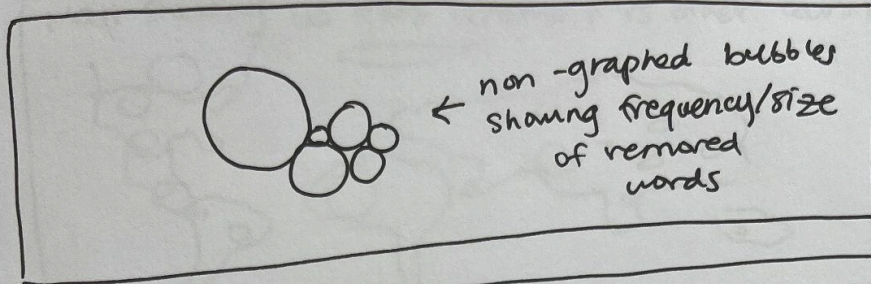
expanation of Viz #3

- idea: have
text
disappear
or look
redacted
at
some
parts?

VIZ #4 Websites That Disappear

netyork.com ownership
— NYt.com kind

epa.org

word
cloud
that
disappears
over time? Sites that
have been a redacted
from most the
7) largest
or removed

explanation of Viz #4

Other ideas
↳ Pet store?
Art more?
Entertainment?

Viz #5 — transparencyreport.google.com — gov't request to
remove content

Map showing US gov't censorship vs other countries
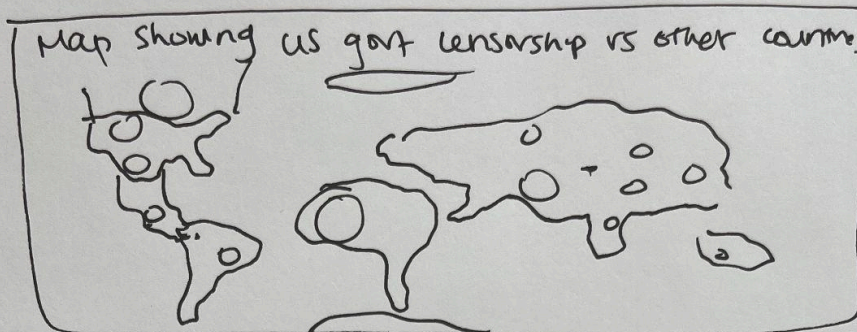
chloropleth or bubble map w/ info on how many
request per year

expl. viz #5

Viz #6

Final Words / layering sounds slowly stripped
to silence @ rate of redaction /

book ban /
something to
show how quickly
we're losing
info

Data & Methods Section   / maybe
incorporate
throughout...?

other Ideas
  ↳ Red scare?
   Art work?
   Entertainment ?

EXECUTION & CHALLENGES

During our brainstorming, there were several other areas of censorship we were excited to explore, such as book bans, general government oversight on the internet and freedom of access, censorship in Hollywood in the time of the Red Scare and art censorship. We also were interested in exploring more unconventional data analysis techniques, deviating from visualizations to auditory analysis, having our website incorporate a playable track sound to silence showing the loss of data over time using the current rate of website removal.

As you can see from our drawings, not everything that we wanted to implement was feasible, and we decided to use different visualizations from some topics after getting a better picture of what we could do with the data available to us in terms of size and features. Some of data we found needed to be aggregated from various sources, had to be manipulated and run through our own re-labelling and classification, and the amount of data available was often not enough to drive a meaningful visualization. Ultimately, however, given the time constraints and the purpose of this project, we decided to limit ourselves to three topical visualizations that dealt with the direct education of the public: books, research papers, and websites. We also decided that we didn't want too many examples of the same type of visualization, so our final execution differed from our mockups in this way. Other challenges that prevented us from expanding our visualizations involved the time it took us to get used to the streamlit platform and its limitations on user interactivity, figuring out how to incorporate html/css and python together to create a seamless website, and general design struggles (navigation sidebar, caching issues, slow loading maps, etc.) In the end, we decided a simpler, specific, color-thematic website would serve us best.