CS 293/EDUC 473

Discovery & Exploration in Educational Text Data

Parsing, Lexical Analyses

Stanford GRADUATE SCHOOL OF EDUCATION

Stanford | NLP

Reminders

- Teacher interview, getting to know you survey, discussant sign up due by **Wednesday's class**
- HW1 due **next Tuesday** at midnight
- Final project
  - Project rationale due the following Tuesday (Jan 28)
  - See Ed Forum announcement for finding project partners
- [Dan Meyer] from the Amplify/Desmos will join the beginning of Wednesday's class!

Today’s class

- Lecture by Dora on textbook paper
- Discussion!

Case Study

Lucy Li

Tricia Bromley

Dan Jurafsky

Content Analysis of Textbooks via Natural Language Processing: Findings on Gender, Race, and Ethnicity in Texas US History Textbooks

AERA Open
Volume 6, Issue 3, July-September 2020

https://doi.org/10.1177/2332858420940312

SAGE journals

Special Topic: Educational Data Science

Content Analysis of Textbooks via Natural Language Processing: Findings on Gender, Race, and Ethnicity in Texas U.S. History Textbooks

Li Lucy, Dorottya Demszky, Patricia Bromley, and Dan Jurafsky

Abstract
Cutting-edge data science techniques can shed new light on fundamental questions in educational research. We apply techniques from natural language processing (lexicons, word embeddings, topic models) to 15 U.S. history textbooks widely used in Texas between 2015 and 2017, studying their depiction of historically marginalized groups. We find that Latinx people are rarely discussed, and the most common famous figures are nearly all White men. Lexicon-based approaches show that Black people are described as performing actions associated with low agency and power. Word embeddings reveal that women tend to be discussed in the contexts of work and the home. Topic modeling highlights the higher prominence of political topics compared with social ones. We also find that more conservative counties tend to purchase textbooks with less representation of women and Black people. Building on a rich tradition of textbook analysis, we release our computational toolkit to support new research directions.

Keywords
artificial intelligence, case studies, content analysis, curriculum, data science, gender studies, history, natural language processing, race, textbooks, textual analysis

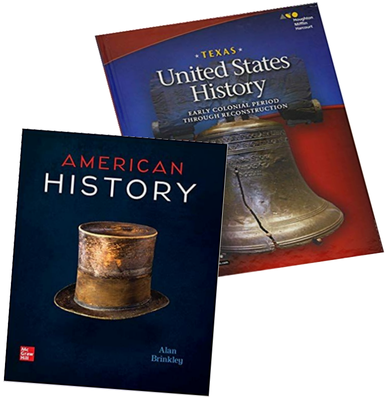
Why are we reading this old paper???

2020 is pre-ChatGPT… ancient times.

- Broad survey of lexical methods
- Illustrates how to apply NLP to **small** data w/ statistical rigor
- Great starting point for discussing the relevance of word-based analyses!

Motivation

Textbooks are the most widely used instructional tool around the world



Social & cultural values

Traditional Methods

Coding protocols, e.g:

36. Fill in each cell of the matrix below using the following codes:

	Groups/ Issues	Rights
Citizens / citizenship		
Children, youth		
Women		
Elderly / Old Age		
Ethnic minorities / racism		
Indigenous groups		
Immigrants / Immigration or Refugees		
Workers / Labor		
Disabled, handicapped		
Gays, lesbians		
The poor / Poverty (nationally or in an international development context)		
Health		
Environment		
Education		
Language and/or culture		
Other. List:		

1 = mentioned
0 = not mentioned

0 = no mention
1 = one or two sentences
2 = at least a paragraph
3 = at least one subheading
4 = at least one chapter heading
5 = over half the chapters

From Meyer, Bromley, & Ramirez (2010)

Texas

- 5.4M K-12 students (2017), 2nd largest in US
- Major textbook market
- Large influence on textbooks in U.S.

Texas Education Agency

Texas

The New York Times

How Texas Teaches History

THE TEXAS TRIBUNE MENU

Texas' Controversial Social Studies Textbooks Under Fire Again

The Washington Post

Education

What do students learn about slavery? It depends where they live.

The New York Times

Texas Mother Teaches Textbook Company a Lesson on Accuracy

Our Goal

Apply NLP to textbooks to answer questions that textbook researchers in education care about

Research Questions

RQ1 How much are different groups of people **mentioned**?

Research Questions

RQ1 How much are different groups of people **mentioned**?

RQ2 How are different groups and individuals **described**?

Research Questions

RQ1 How much are different groups of people **mentioned**?

RQ2 How are different groups and individuals **described**?

RQ3 Which **topics** are prominent and how do they relate to groups of people?

Research Questions

RQ1 How much are different groups of people **mentioned**?

RQ2 How are different groups and individuals **described**?

RQ3 Which **topics** are prominent and how do they relate to groups of people?

American History Textbook Data (2015-17)

American History Textbook Data (2015-17)

Messy purchase data from Texas districts

8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us history texas ed	B.ISD

American History Textbook Data (2015-17)

Messy purchase data from Texas districts

8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us history texas ed	B.ISD

manual clean-up & disambiguation

book	district	count
Am. Hist.	T.ISD	30
Give me lib.	B.ISD	100

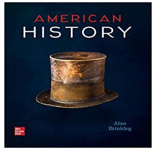
American History Textbook Data (2015-17)

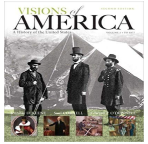
Messy purchase data from Texas districts

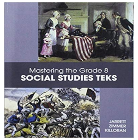
8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us history texas ed	B.ISD

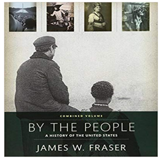
book	district	count
Am. Hist.	T.ISD	30
Give me lib.	B.ISD	100

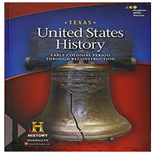
keep **15** most widely purchased textbooks









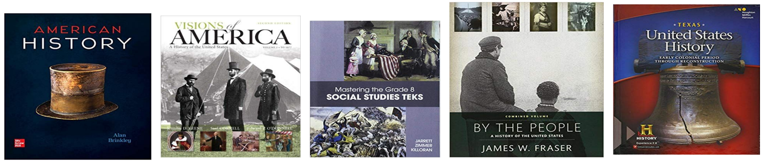


American History Textbook Data (2015-17)

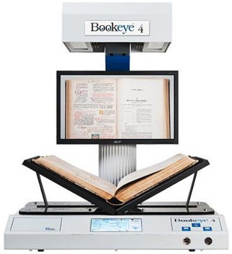
Messy purchase data from Texas districts

8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us history texas ed	B.ISD

book	district	count
Am. Hist.	T.ISD	30
Give me lib.	B.ISD	100



scan

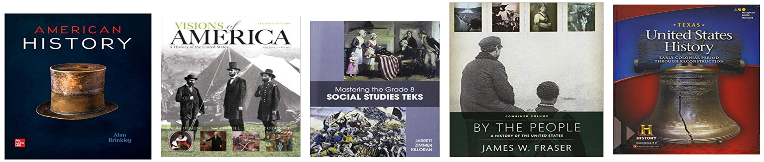


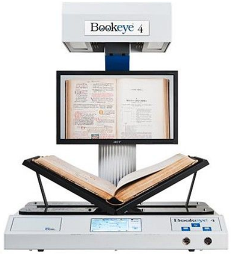
American History Textbook Data (2015-17)

Messy purchase data from Texas districts

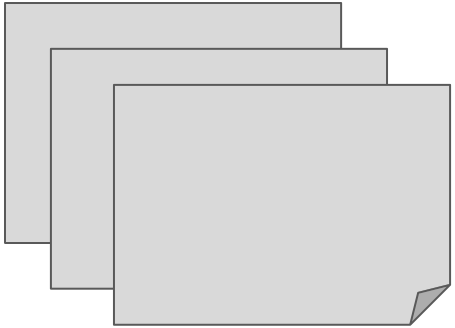
8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us history texas ed	B.ISD

book	district	count
Am. Hist.	T.ISD	30
Give me lib.	B.ISD	100





OCR w/ ABBYY FineReader



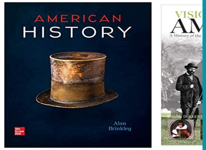
American History Textbook Data (2015-17)

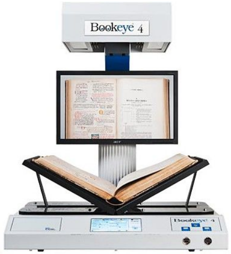
Messy purchase data from Texas districts

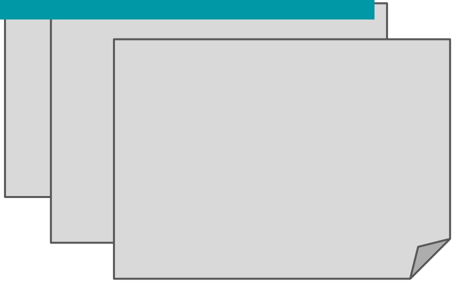
8th gr give me liberty	T.ISD
pearson US hsitory coloniz...	B.ISD
Pearson us his	

book	district	count
Am. Hist.	T.ISD	30
Give me lib	B.ISD	100

2 MONTHS OF WORK FOR TWO PEOPLE







Demographic Data

- district-level student demographic data
  - the National Center for Education Statistics (NCES), for AY 2016-17

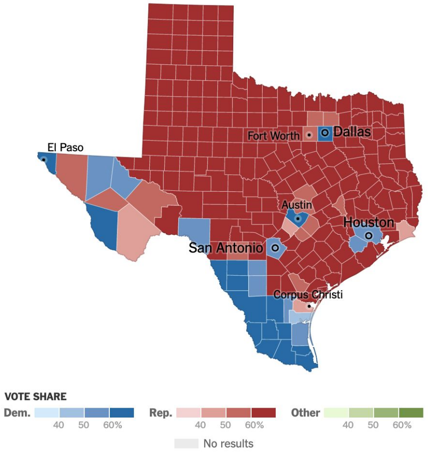
NCES

source: New York Times

Demographic Data

- district-level student demographic data
  - the National Center for Education Statistics (NCES), for AY 2016-17
- county-level political leaning
  - two party vote shares in 2016 elections

NCES



source: [New York Times]

Example

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

RQ1

How Much Are Different Groups of People **Mentioned**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women’s rights. (Brinkley, 2015: p. 330)

RQ1

How Much Are Different Groups of People **Mentioned**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

Coreference Resolution

RQ1

How Much Are Different Groups of People **Mentioned**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

Identifying people-related common nouns (WordNet, 95% accuracy)

RQ1

How Much Are Different Groups of People **Mentioned**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women’s rights. (Brinkley, 2015: p. 330)

Named Entity Recognition

RQ1

Race/Ethnicity & Gender

Common nouns referring to individuals or groups

446 marked

1665 unmarked
engineer, family

Women
wife, mother
Men
son, boy

Black
black, slaves, africans
Latinx
mexican, latina
White
colonist, white, european
Other
immigrants, asian-americans

RQ1

Race/Ethnicity & Gender

Common nouns referring to individuals or groups

446 marked

1665 unmarked
engineer, family

Women
wife, mother
Men
son, boy

Black
black, slaves, africans
Latinx
mexican, latina
White
colonist, white, european
Other
immigrants, asian-americans

Intersectionality
black women

RQ1

Race/Ethnicity & Gender

Common nouns referring to individuals or groups

446 marked

1665 unmarked
engineer, family

Women
wife, mother
Men
son, boy

Black
black, slaves, africans
Latinx
mexican, latina
White
colonist, white, european
Other
immigrants, asian-americans

Intersectionality
black women

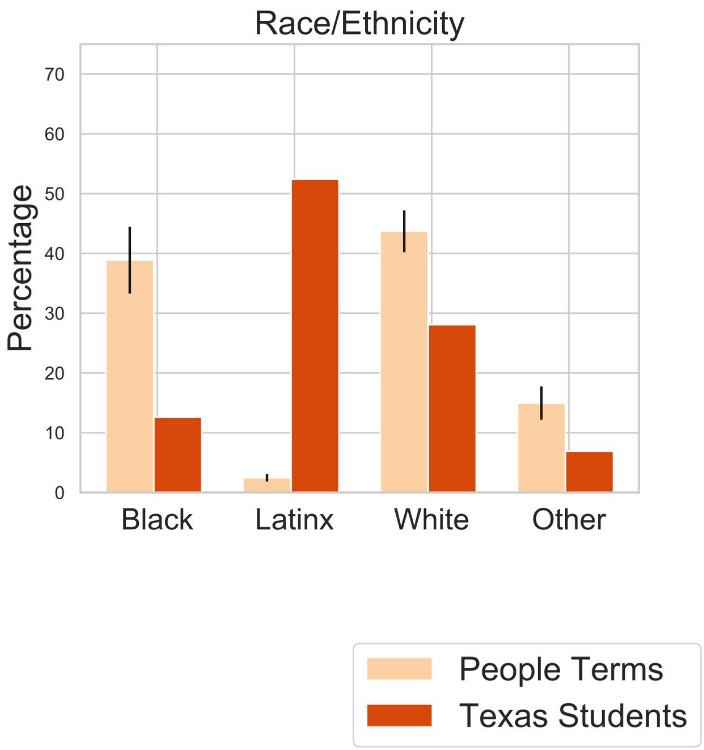
Named individuals

Gender
(Wikidata)

Race
(manual)

RQ1

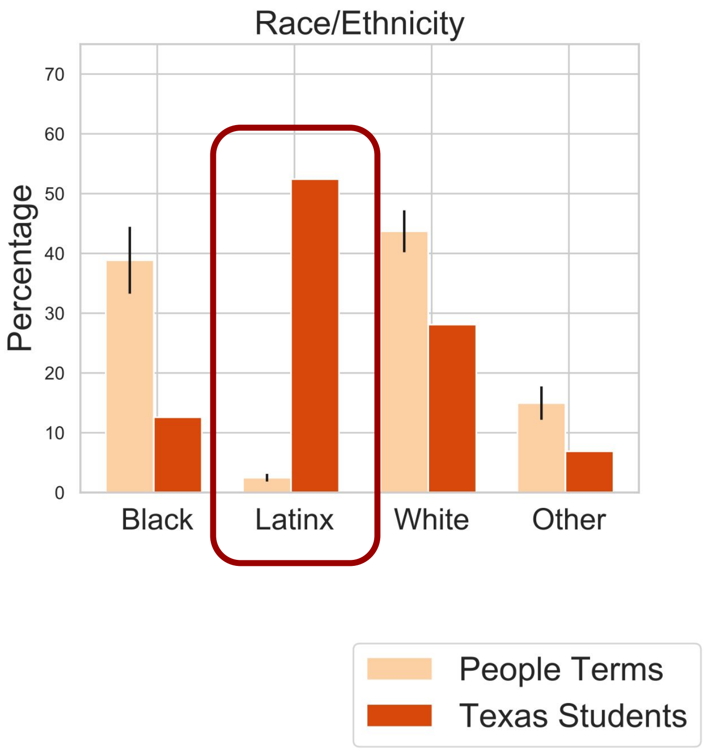
Comparing Student Demographics w/ Representation in Text



Common nouns referring to individuals or groups

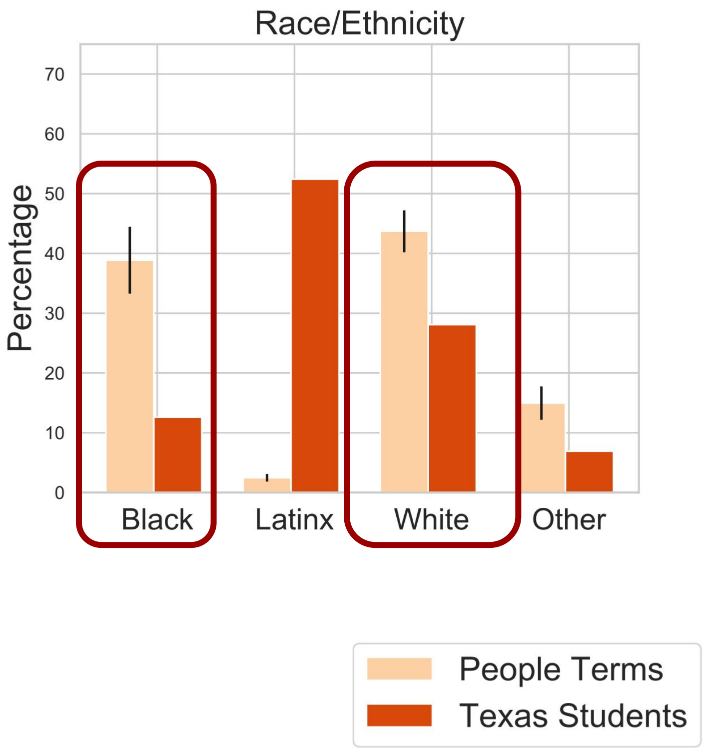
RQ1

Hispanic / Latinx Students are Disproportionately Underrepresented



RQ1

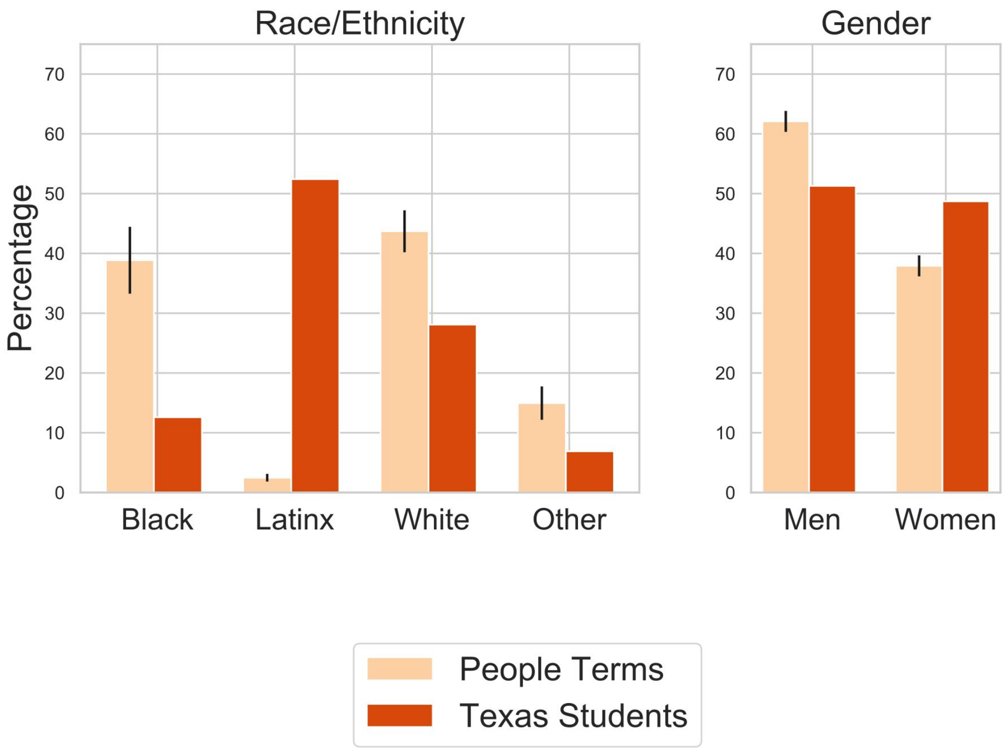
African Americans and White People are Mentioned Disproportionately More



white people are mentioned even more often than the plot shows (since this ethnicity is often unmarked)

RQ1

Men Are Mentioned Disproportionately More Often Than Women



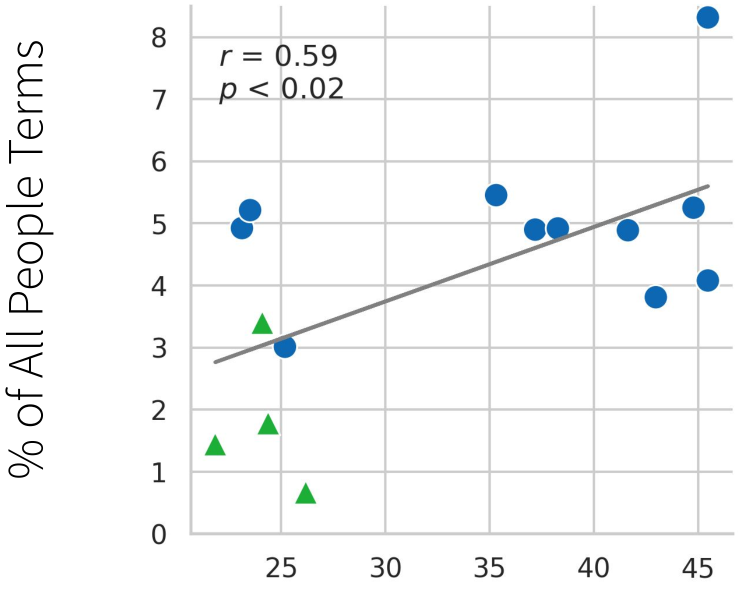
RQ1

Top 50 Named People

RQ1

Books in More Democratic Counties Mention Black People and Women More

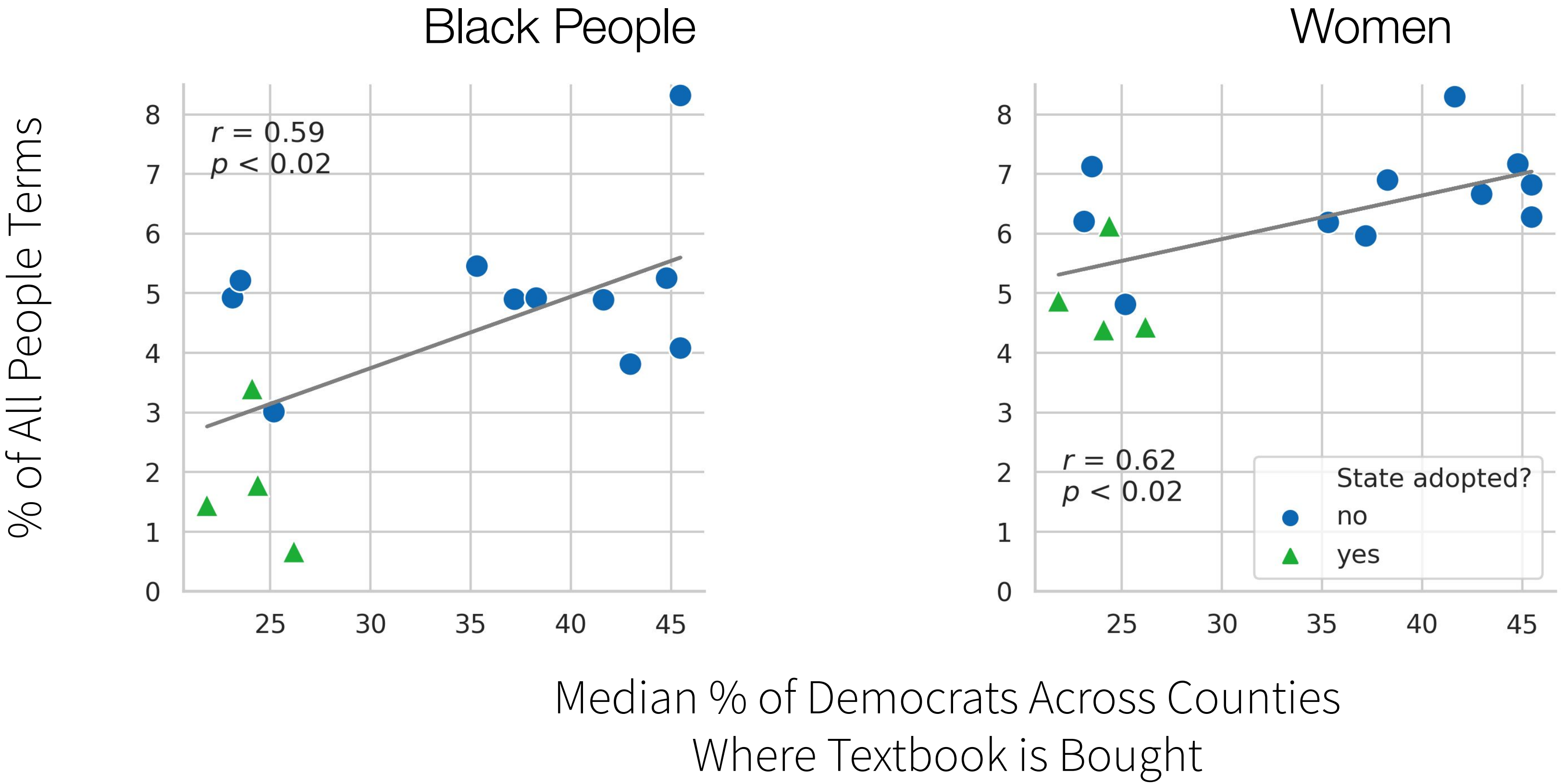
Black People



Median % of Democrats Across Counties Where Textbook is Bought

RQ1

Books in More Democratic Counties Mention Black People and Women More



RQ2

How Are Different Groups and Individuals **Described**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

RQ2

How Are Different Groups and Individuals **Described**?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

Dependency Parsing

RQ2

How Are Different Groups and Individuals **Described**?

Progress toward feminist goals was limited in the antebellum years, but in [illegible]omen did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women’s rights. (Brinkley, 2015: p. 330)

subject

Dependency Parsing

RQ2

How Are Different Groups and Individuals **Described**?

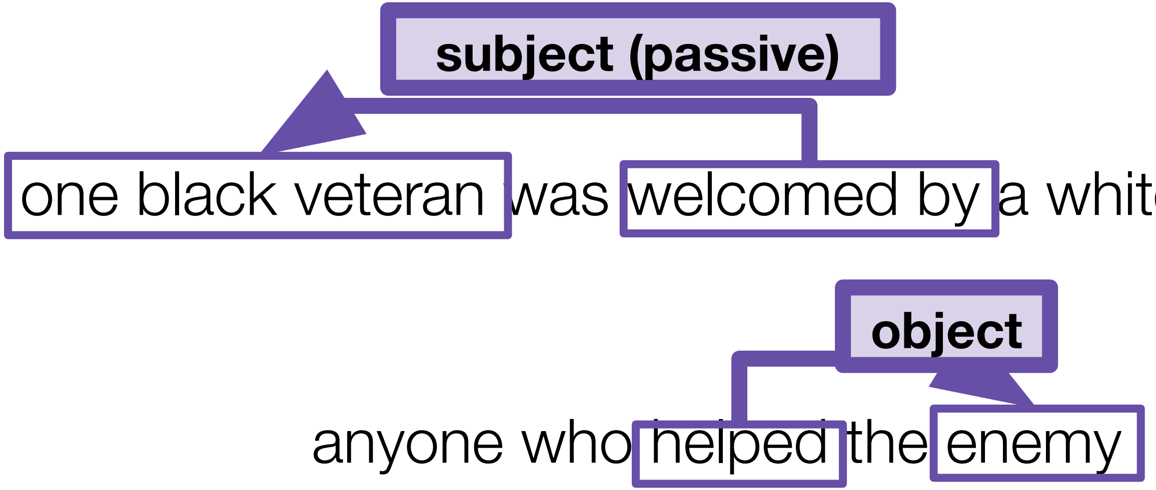
adj modifier

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

Dependency Parsing

RQ2

How Are Different Groups and Individuals **Described**?



Dependency Parsing

RQ2

Lexicons

adjectives

NRC Valence, Arousal, Dominance lexicons (Mohammad, 2018)

amazing (↑ valence) *asleep* (↓ arousal) *competitive* (↑ dominance)

RQ2

Lexicons

adjectives

NRC Valence, Arousal, Dominance lexicons (Mohammad, 2018)

amazing (↑ valence) *asleep* (↓ arousal) *competitive* (↑ dominance)

verbs

Connotation frames (Rashkin et al, 2016; Sap et al., 2017)

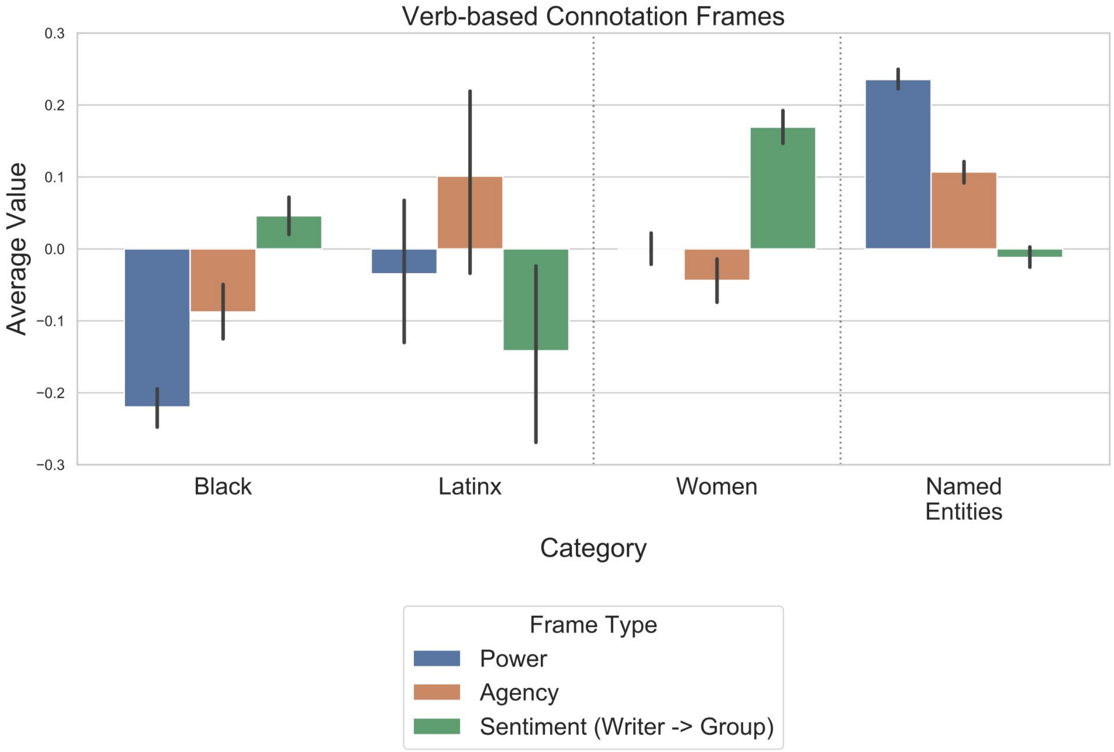
X (-1 agency) *obeys*

X (-1 power) *applauds Y* (+1 power)

X (↑ sentiment) *suffered*

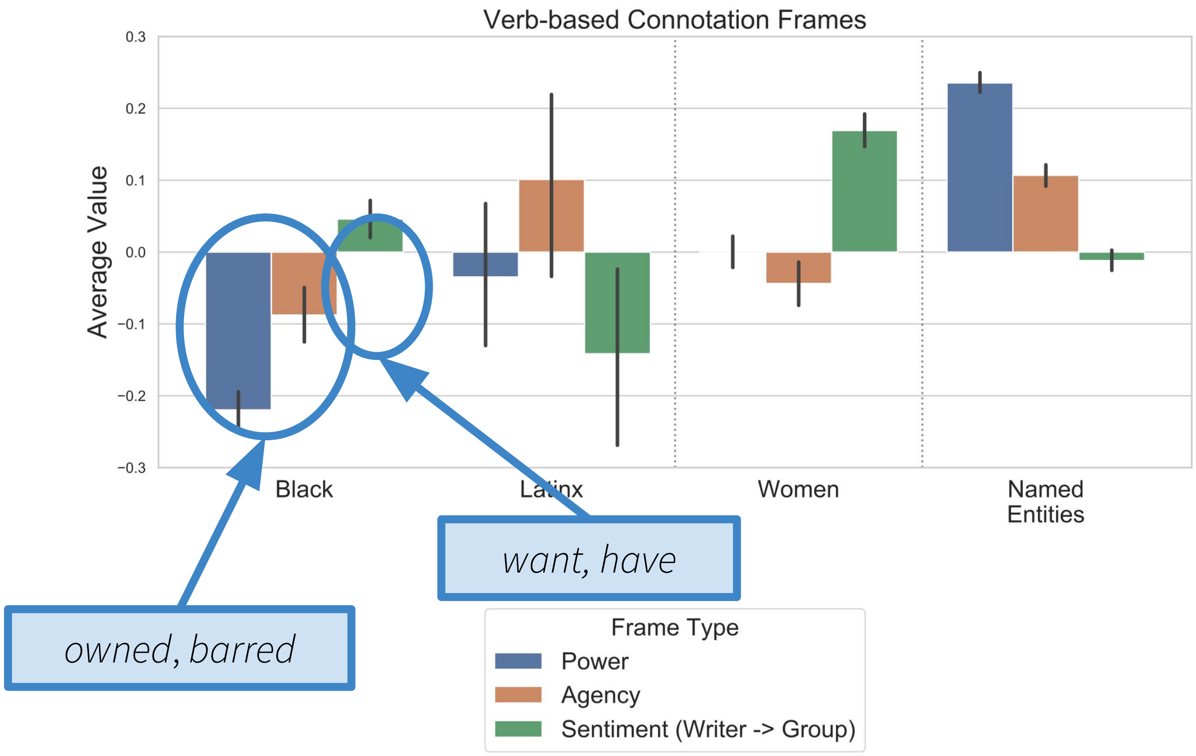
RQ2

Power & Agency



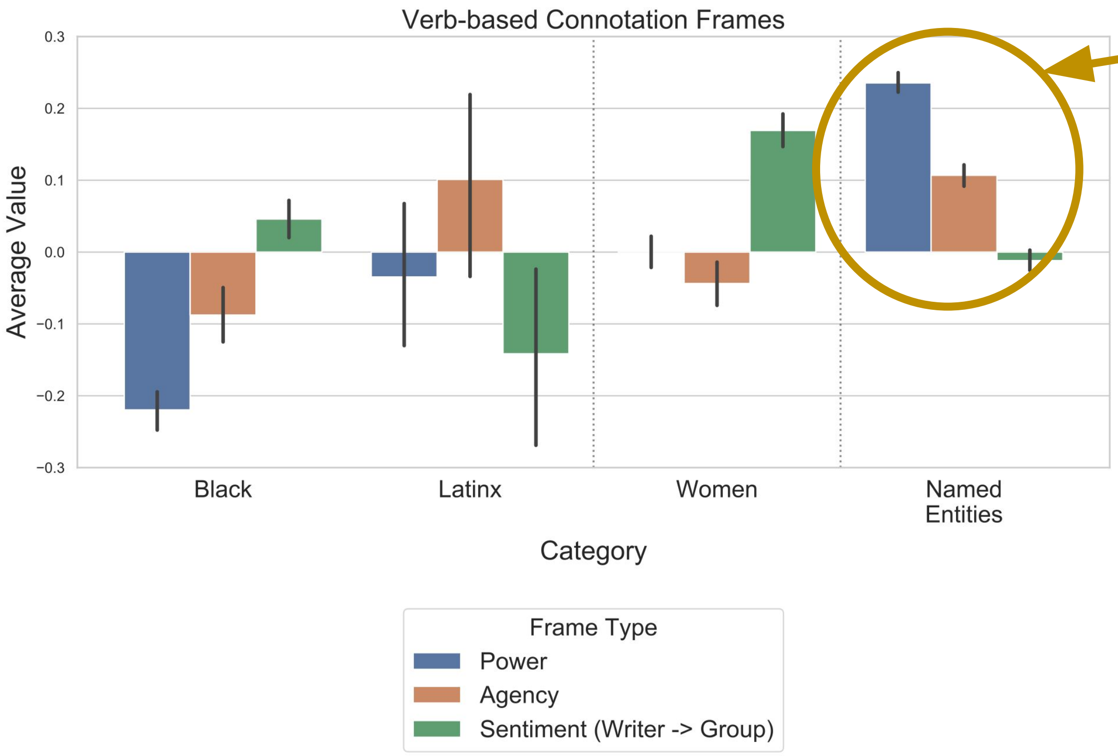
RQ2

Power & Agency



Power & Agency

veto, initiate



RQ2

Other Lexicon Findings

African Americans (↓ adjective dominance)

Ex: *slave*, *inferior*

Famous people (↑ adjective arousal)

Ex: *worried*, *victorious*, *furious*

Women (↑ verb sentiment)

Ex: women *marry* or *help*

RQ2

GloVe Embeddings w/ Bootstrapping

- unigrams & bigrams (skip stopwords)
- GloVe training w/ bootstrapping (Antoniak & Mimno, 2018)
- mean cosine similarity across 50 runs, between:

Bootstrapping helps mitigate data sparsity!
Create samples of the data (e.g. sample 50 times with replacement), train model on each and aggregate results.

RQ2

GloVe Embeddings w/ Bootstrapping

- unigrams & bigrams (skip stopwords)
- GloVe training w/ bootstrapping (Antoniak & Mimno, 2018)
- mean cosine similarity across 50 runs, between:

man-related terms

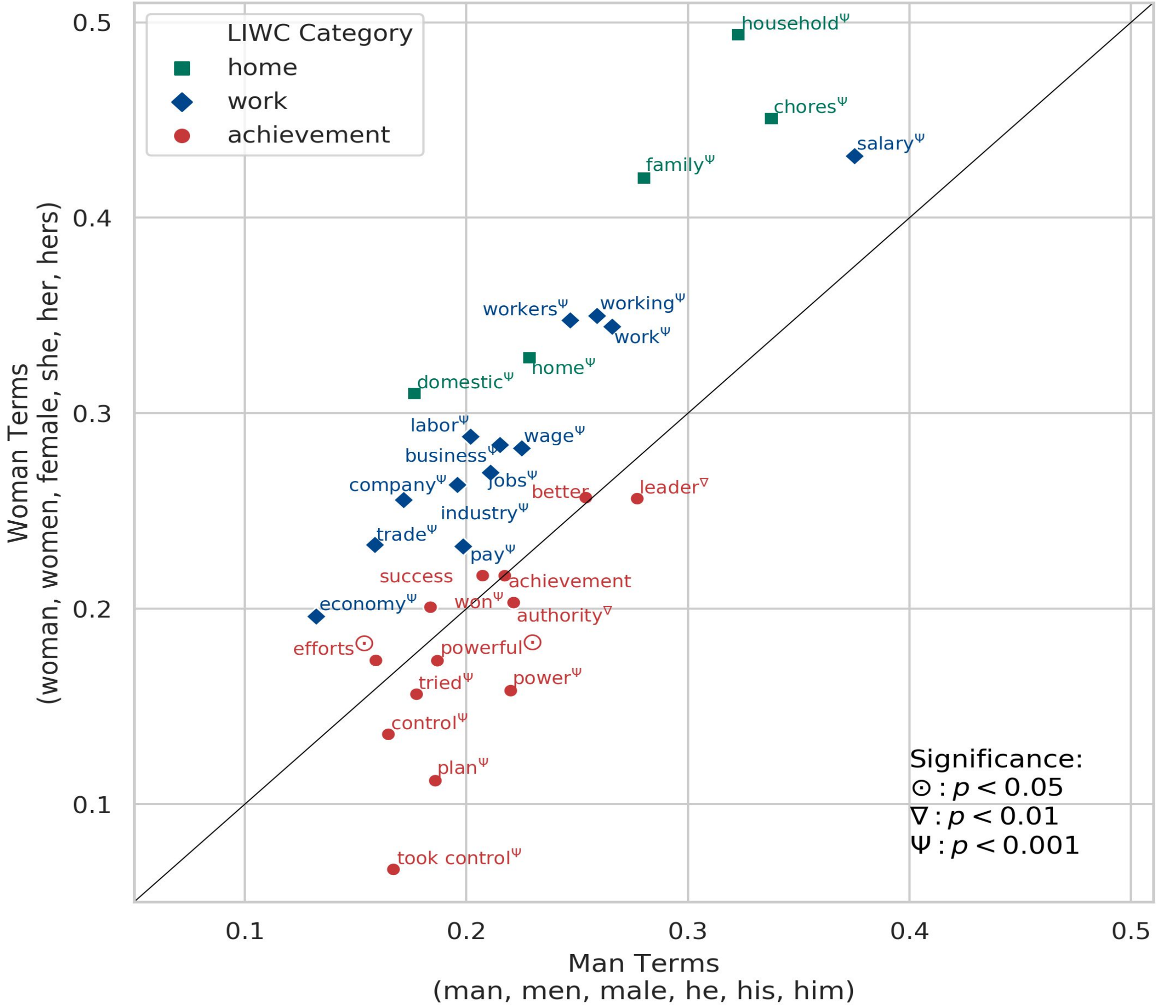
(*man, men, male, he, his, him*)

woman-related terms

(*woman*, *women*, *female*, *she*, *her*, *hers*)

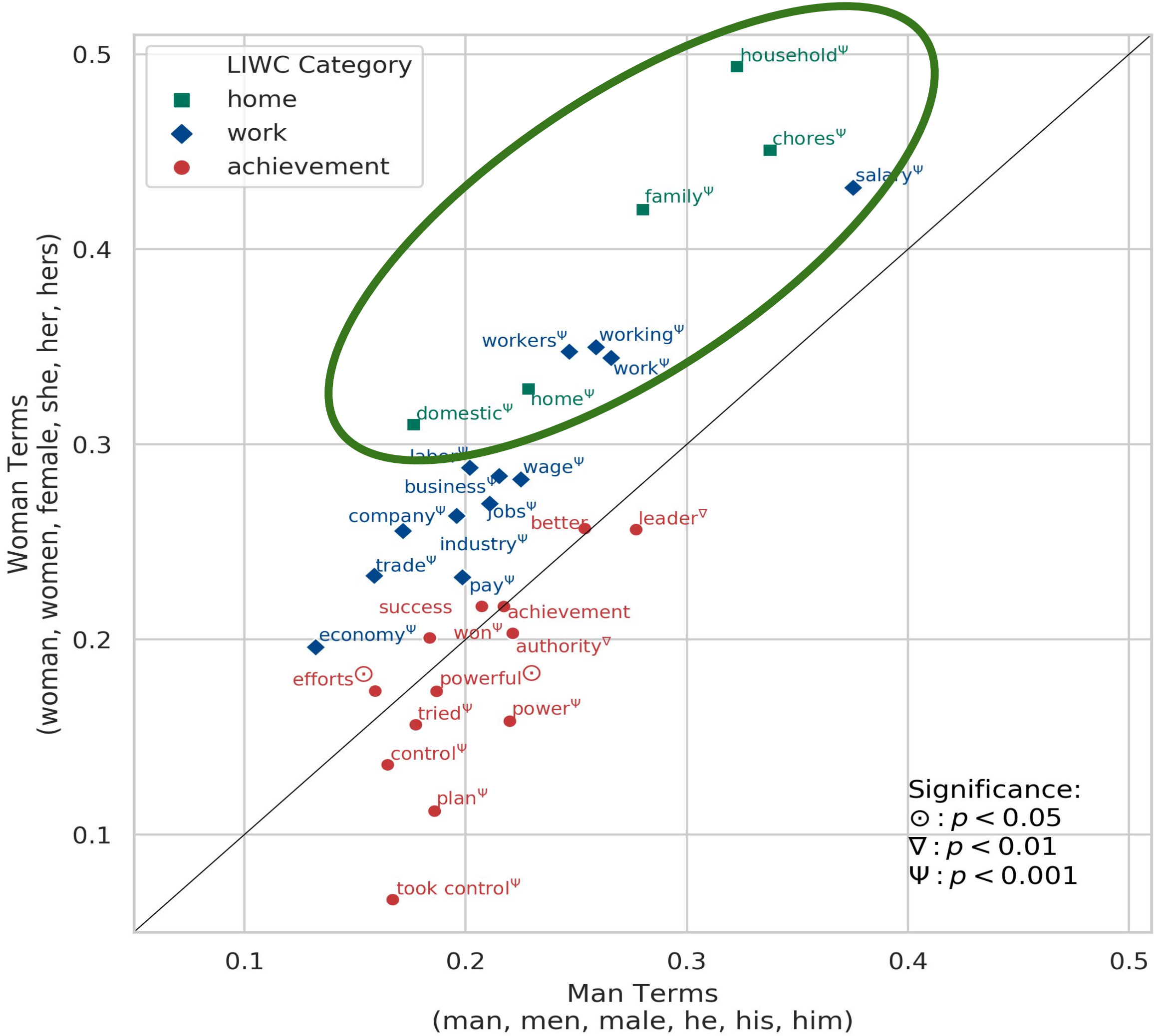
most frequent words in ***home***, ***work*** and ***achievement*** LIWC categories

RQ2


LIWC Category
home
work
achievement
household$^{\Psi}$
chores$^{\Psi}$
salary$^{\Psi}$
family$^{\Psi}$
workers$^{\Psi}$
working$^{\Psi}$
work$^{\Psi}$
home$^{\Psi}$
domestic$^{\Psi}$
labor$^{\Psi}$
wage$^{\Psi}$
business$^{\Psi}$
company$^{\Psi}$
jobs$^{\Psi}$
better
leader$^{\nabla}$
industry$^{\Psi}$
trade$^{\Psi}$
pay$^{\Psi}$
success
achievement
economy$^{\Psi}$
won$^{\Psi}$
authority$^{\nabla}$
efforts$^{\odot}$
powerful$^{\odot}$
tried$^{\Psi}$
power$^{\Psi}$
control$^{\Psi}$
plan$^{\Psi}$
took control$^{\Psi}$
Significance:
$\odot : p < 0.05$
$\nabla : p < 0.01$
$\Psi : p < 0.001$
0.1
0.2
0.3
0.4
0.5
0.1
0.2
0.3
0.4
0.5
Woman Terms
(woman, women, female, she, her, hers)
Man Terms
(man, men, male, he, his, him)

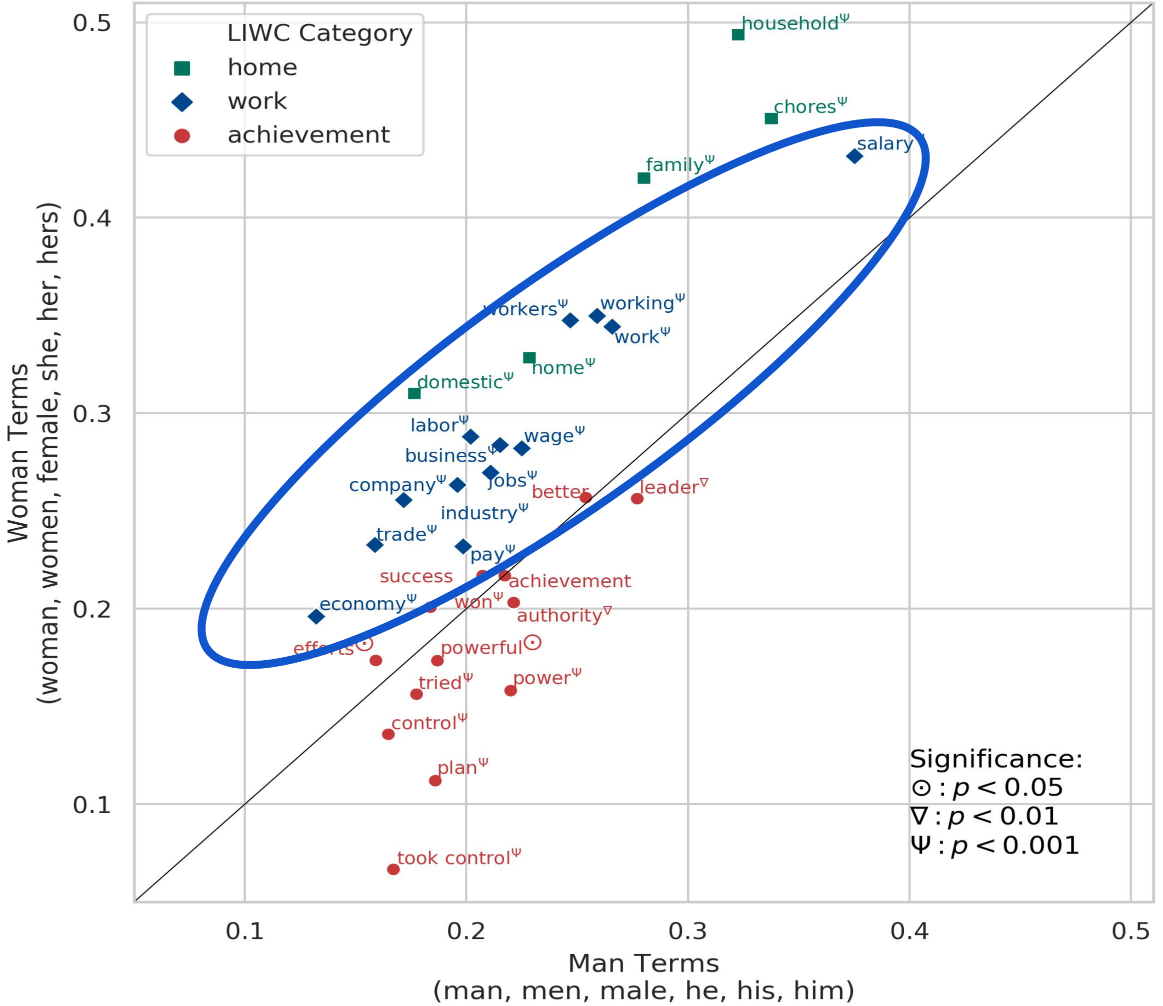
RQ2

Home related terms are more closely related to women, with very high significance



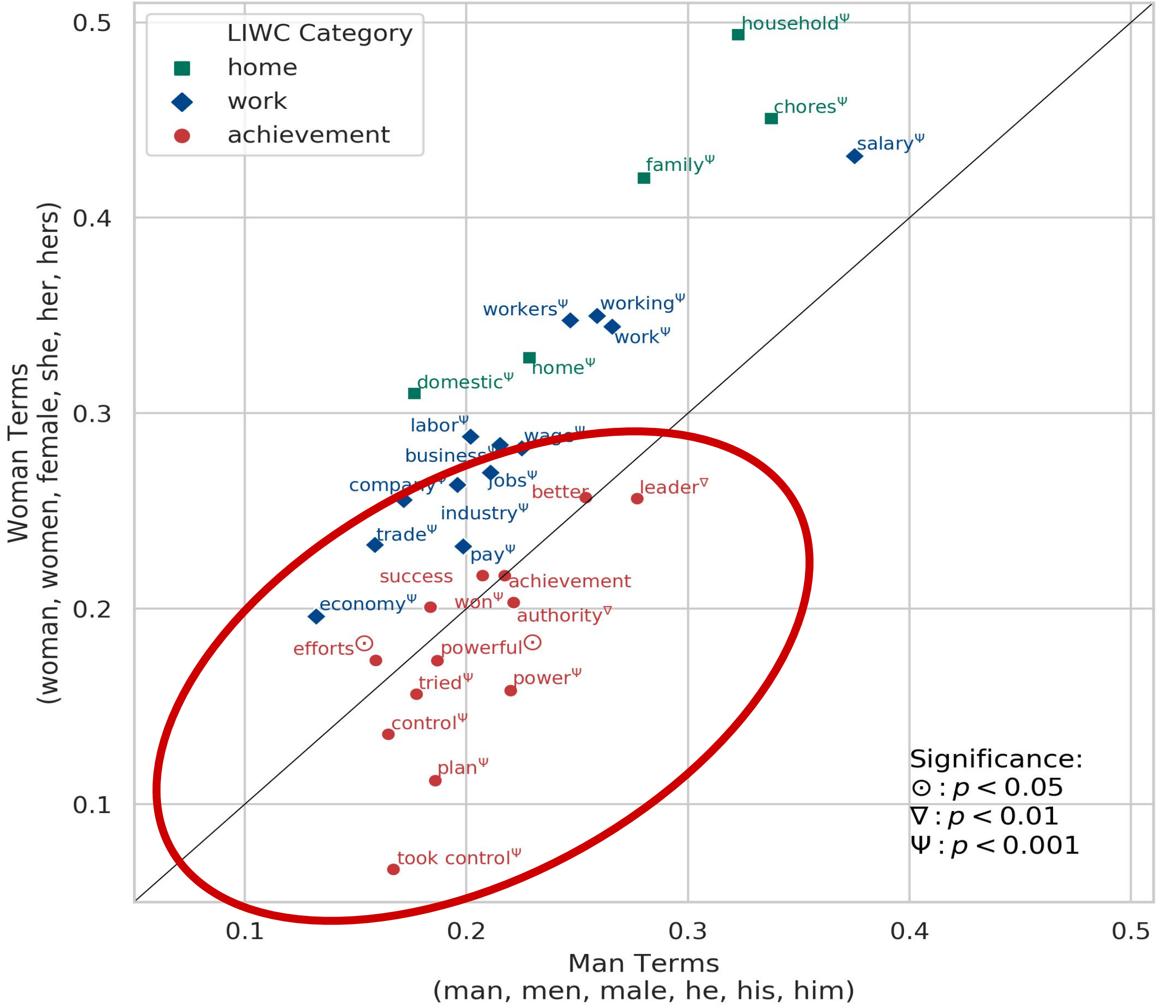
RQ2

Work related terms are more closely related to women, with very high significance



RQ2

Most **achievement** related terms are more closely related to men but not all



RQ3

Which **Topics** Are Prominent and How Do They Relate to Groups of People?

Progress toward feminist goals was limited in the antebellum years, but individual women did manage to break the social barriers to advancement. Elizabeth Blackwell, born in England, gained acceptance and fame as a physician. Her sister-in-law Antoinette Brown Blackwell became the first ordained woman minister in the United States; and another sister-in-law, Lucy Stone, took the revolutionary step of retaining her maiden name after marriage. Stone became a successful and influential lecturer on women's rights. (Brinkley, 2015: p. 330)

Topics

reform (reform,progress, social,...)

women's rights (women,right, movement,...)

religion (church,religion, christian,...)

marriage (women,men, young,...)

RQ3

Topic Modeling

- LDA (Blei & Jordan, 2003)
  - 50 topics, induced at the sentence-level
  - run together on all books

RQ3

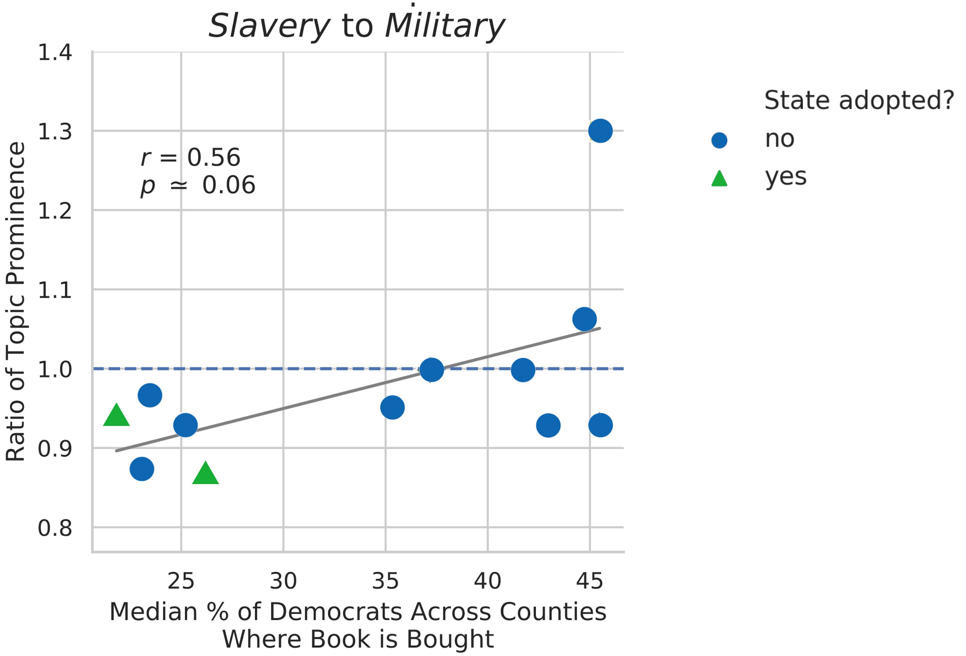
Comparing Topic Prominence Across Books

Using **ratio of relative frequencies** of topics helps control for noise

$$\frac{\text{Mean freq. of topic(s) related to } \mathbf{X} \text{ in Book A}}{\text{Mean freq. of topic(s) related to } \mathbf{Y} \text{ in Book A}}$$

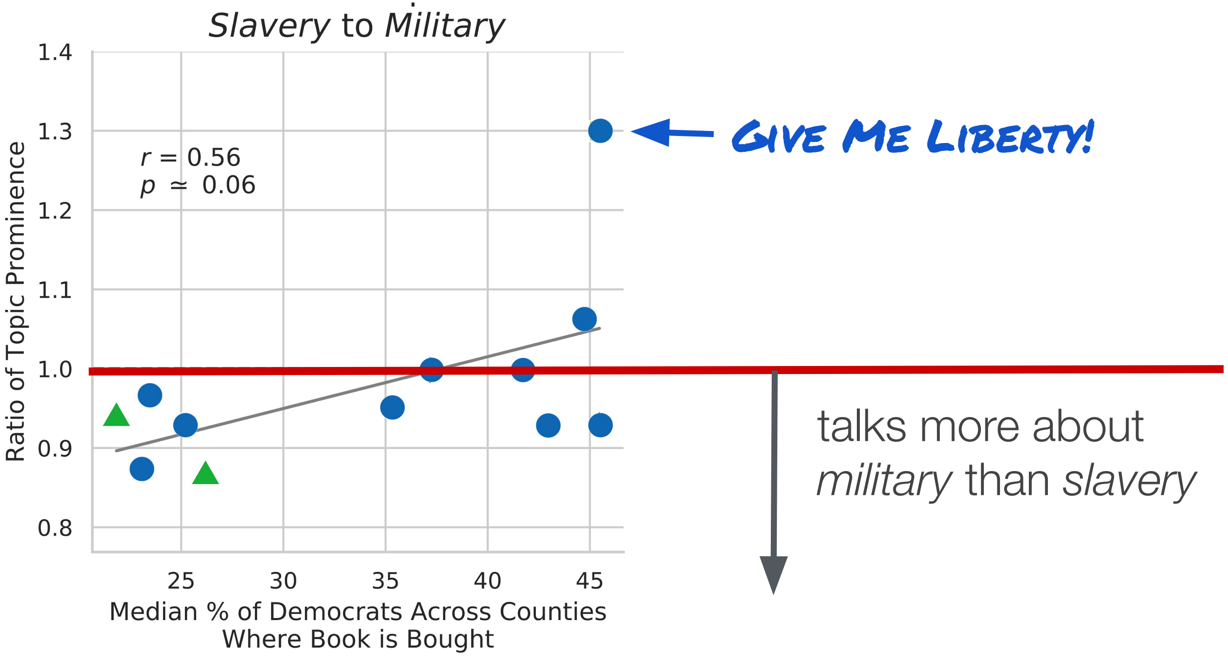
RQ3

Comparing Groups of Topics



RQ3

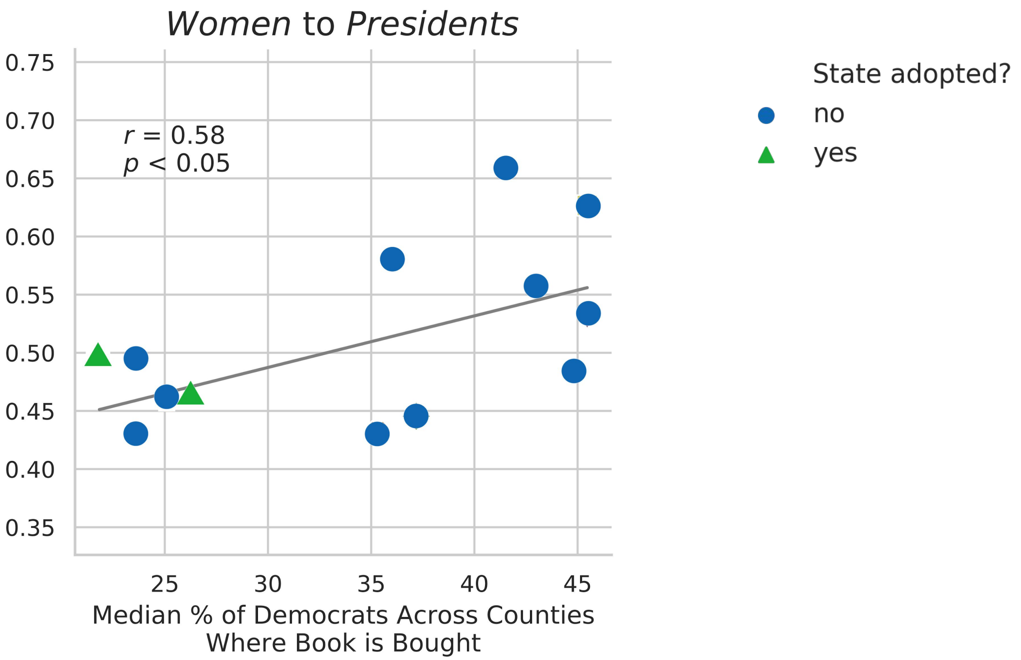
Besides One Outlier, The Ratio of These Topics is Similar Across Books



talks more about *military* than *slavery*

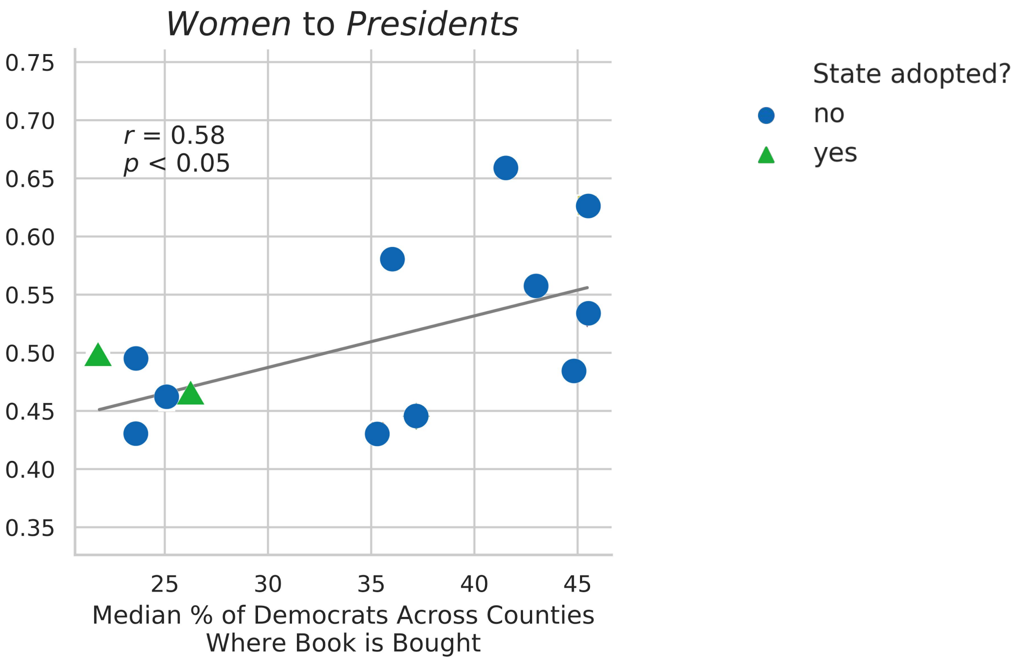
RQ3

Comparing Groups of Topics



RQ3

All Books Talk More about *Presidents* than about *Women*, but the Ratio is Closer to 1 in Books in Democratic Counties



Summary

Methods

Results

RQ1

How much are different groups of people **mentioned**?

Methods

Coref
NER
Wordnet

Results

Latinx people virtually absent;
Named people mainly white men;
More diverse representation in books bought in more Democratic counties;

RQ2

How are different groups and individuals **described**?

Methods	Results
Coref NER Wordnet	Latinx people virtually absent; Named people mainly white men; More diverse representation in books bought in more Democratic counties;
VAD Lexicon Connotation Frames Semantic similarity	Women discussed in context of marriage, home & work; Black people have low agency, power, and dominance

RQ3

Which **topics** are prominent and how do they relate to groups of people?

Methods	Results
Coref NER Wordnet	Latinx people virtually absent; Named people mainly white men; More diverse representation in books bought in more Democratic counties;
VAD Lexicon Connotation Frames Semantic similarity	Women discussed in context of marriage, home & work; Black people have low agency, power, and dominance
Topic modeling	Social history topics tend be more prominent in books bought in more Democratic counties, but similarities across books are greater than their differences

What questions do you have?

Discussion

Theme #1: Methodological Tensions

- “But are we being clear about what’s qualitative here in a quantitative paper?”
  - “While the paper’s NLP methods excelled at scaling pattern detection, interpreting the deeper social and historical implications of these patterns still remained firmly in the realm of human judgment”
  - “Obviously, we need some form of qualitative decision making (e.g. the labels used given for the topic modeling analysis in the paper). However, *what should our boundaries be for these types of decisions*? When do we choose between a numerical explanation or an informed interpretation of the data?”

Theme #1: Methodological Tensions

- "Could LLMs do the same, or would it be just as difficult to get the nuances right?"
  - "While I usually default to a "sledgehammer" approach with LLMs when trying to extract information, the authors employed an intelligent combination of focused NLP techniques."
  - "What I found most enlightening was seeing how "classical" NLP methods, when thoughtfully combined, can provide efficient and interpretable insights from educational texts."
  - "Named Entity Recognition had an F1 score of 0.71… do you not think that the NER would work worse for people of color than for white history figures? My point is, the paper uses methodologies that seem objective, and they are convincing, but at the same time relying on lexicons and NER and other things and not vetting the bias in those areas as well could lead to another host of problems."
  - "You could train an LLM on only U.S. history textbooks, then surface what facts the model learned about social groups via fill in the blank questions like: Black people are ____."

(try to)

Wait… let's answer that real quick.

Theme #2: Interpreting Results

- “Does quantifying mentions of different groups actually capture meaningful inclusion?”
  - “The textbook does have its problems, but it’s also a problem with history as a whole, as a practice that has been dominated by a white nationalist American narrative within the US.”
  - “Yet, from a historical perspective, such representations may be factually accurate, reflecting systemic inequities. The issue, then, may not lie solely in the content but in how it is contextualized and presented to students.”
  - “For example, it made sense to me that women should be spoken about in the context of the home for most of pre-modern US history, since their roles were largely in the home back then.”
  - “I wonder how the methods and findings could be further informed by textbook research insights into what representation in history could look like (e.g. it's true that almost all of our presidents were white men - how does that influence what we 'expect' in curricula?)”
  - “Is this phenomenon a result of shifts in how historical topics are framed or is this a reflection of the inclusion of more contemporary history in textbooks?”

Theme #2: Interpreting Results

- "How computational tools like those used in the paper can be used for systemic change instead of just for critiquing?"
  - "Can we introduce these methods of analysis into the workflow of textbook authorship?"
  - "Should it be required that before adoption in schools, a textbook must meet some threshold of word frequencies representing racial groups in America?"

California doesn’t want to be like Texas!

You work at a research think tank advising the state on which textbooks to keep, toss, and buy.

What kinds of quantitative metrics (inspired by the paper) can help the state make this decision?

Theme #3: Extensions Galore!

- Can we use/extend these methods to analyze…
    - Movies, worksheets, other supplementary materials?
    - Standardized tests?
    - Textbooks in other states? Other countries?
    - Specific chapters and events in history textbooks?
    - Change over time… As the population changes? Within the same publisher?
    - Textbooks in other subjects, like math?

Please watch Dan Meyer's ASU GSV keynote by Wednesday!