

What defines a word?

Using the New York Times to understand the evolution of words throughout history.

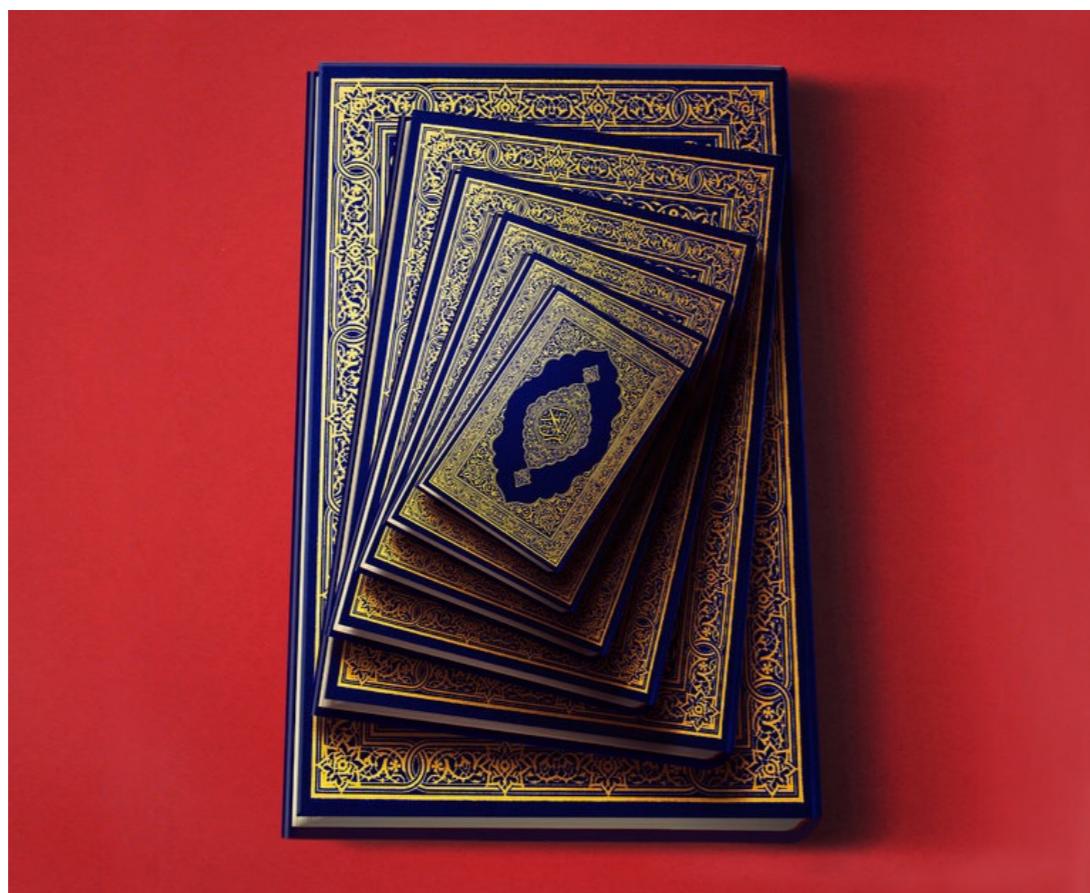
Madison Borgmann, Kate McManus

New York Times



Does this reveal more about our culture during time periods we're examining or how *The New York Times* is reporting these events.

First Words Articles



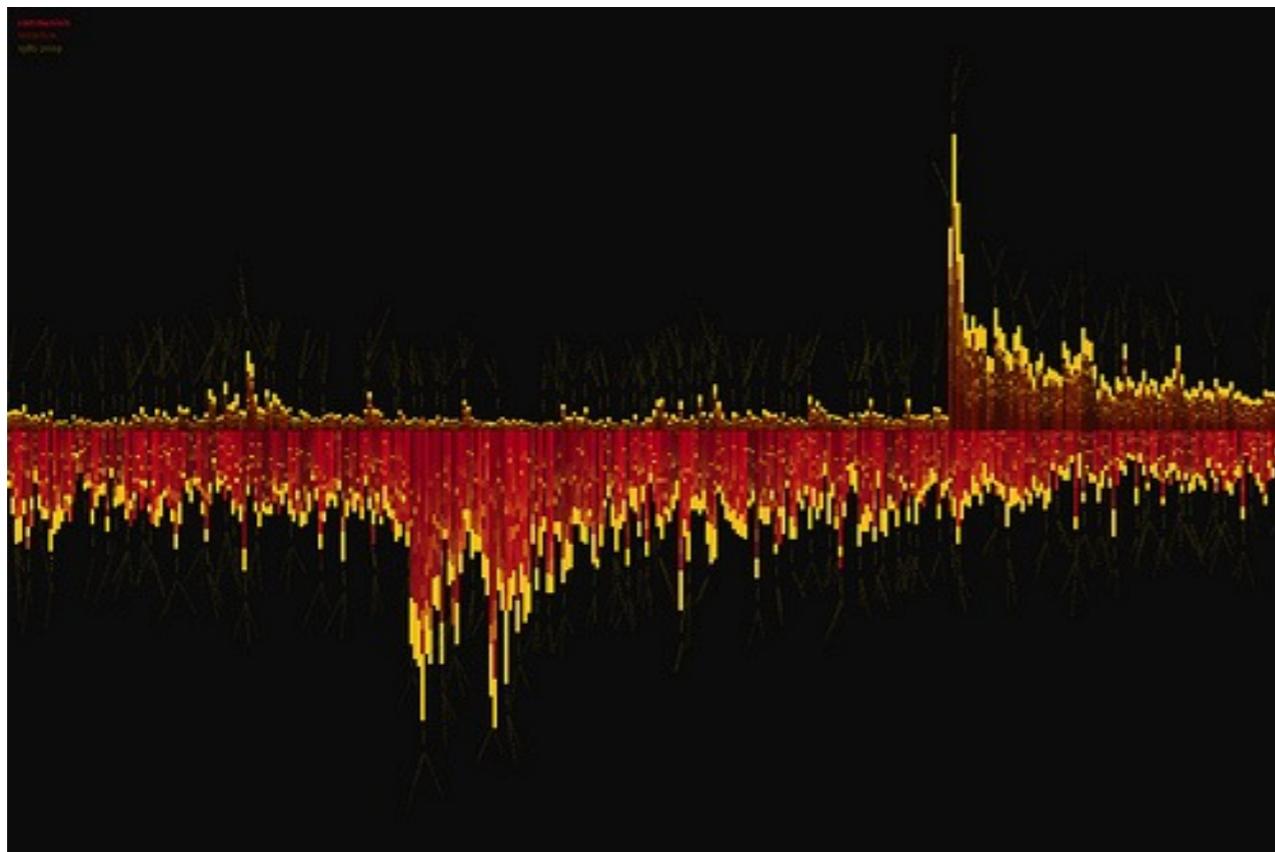
Radical



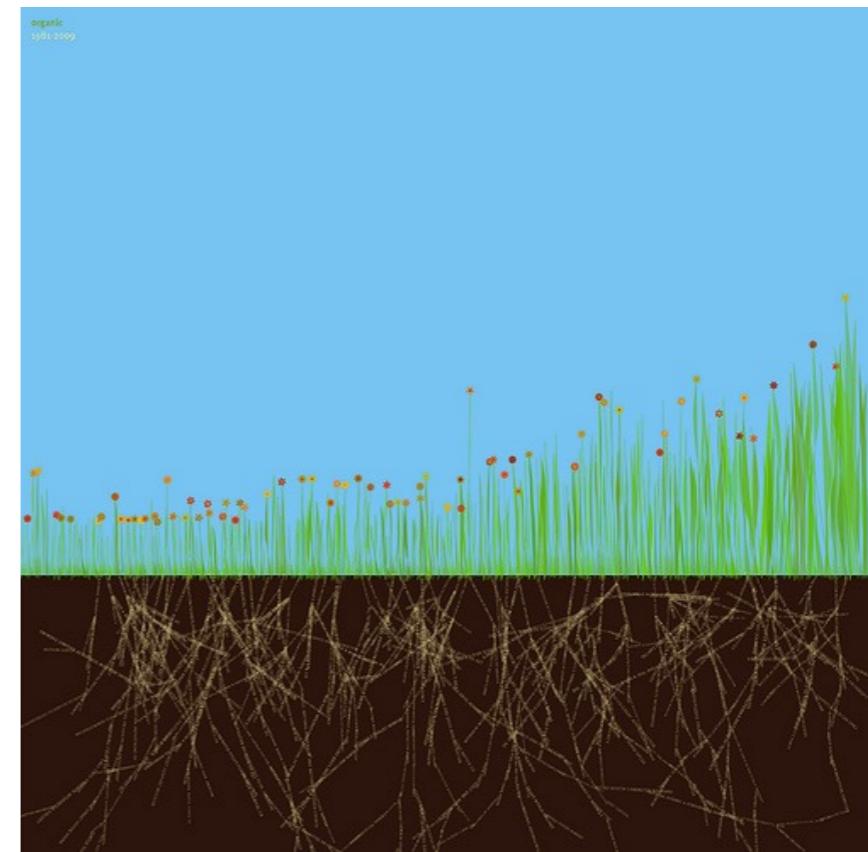
Progressive

- How do words change over time?
- Does this change align with how they're reported over time?

Past Work using the NYTimes API



Terrorism vs. Communism



Organic

Here's an Election Day diversion, courtesy of the research and development people at The Times. It's called Chronicle and it measures the number of articles in which a particular word or phrase has appeared since 1981.

The brainchild of Alexis Lloyd, a creative technologist, Chronicle has a serious purpose.

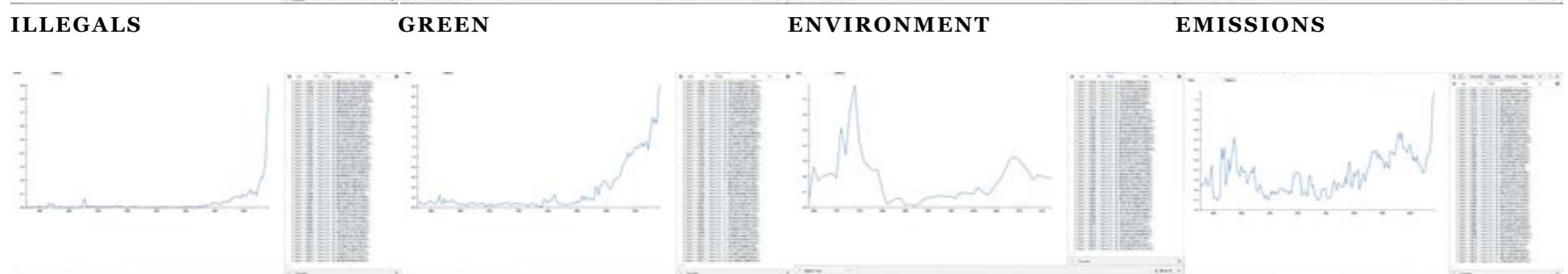
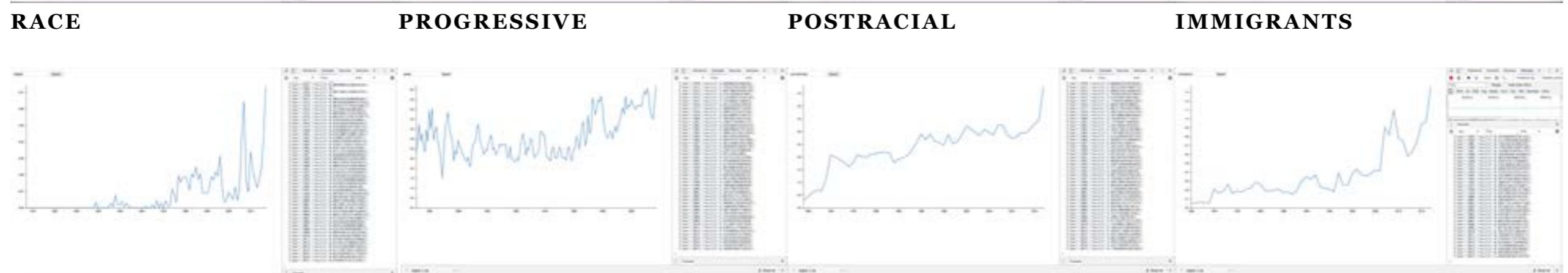
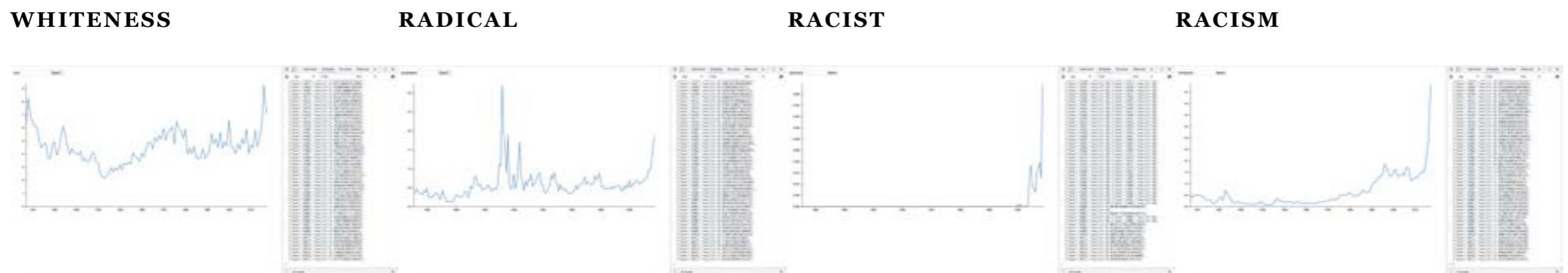
"You can see how topics have risen and fallen over time, but you can also see the way language and terminology change over time," she said. For example, she said, the term "greenhouse effect" has given way to the broader discussion of "climate change."

She built a rudimentary version of the tool about six months ago and has been refining it ever since. Eventually, she intends to take its reach back to include more than another 100 years of Times usage. It is intended for internal use, at least for now.

Alexis Lloyd's Chronicle Tool

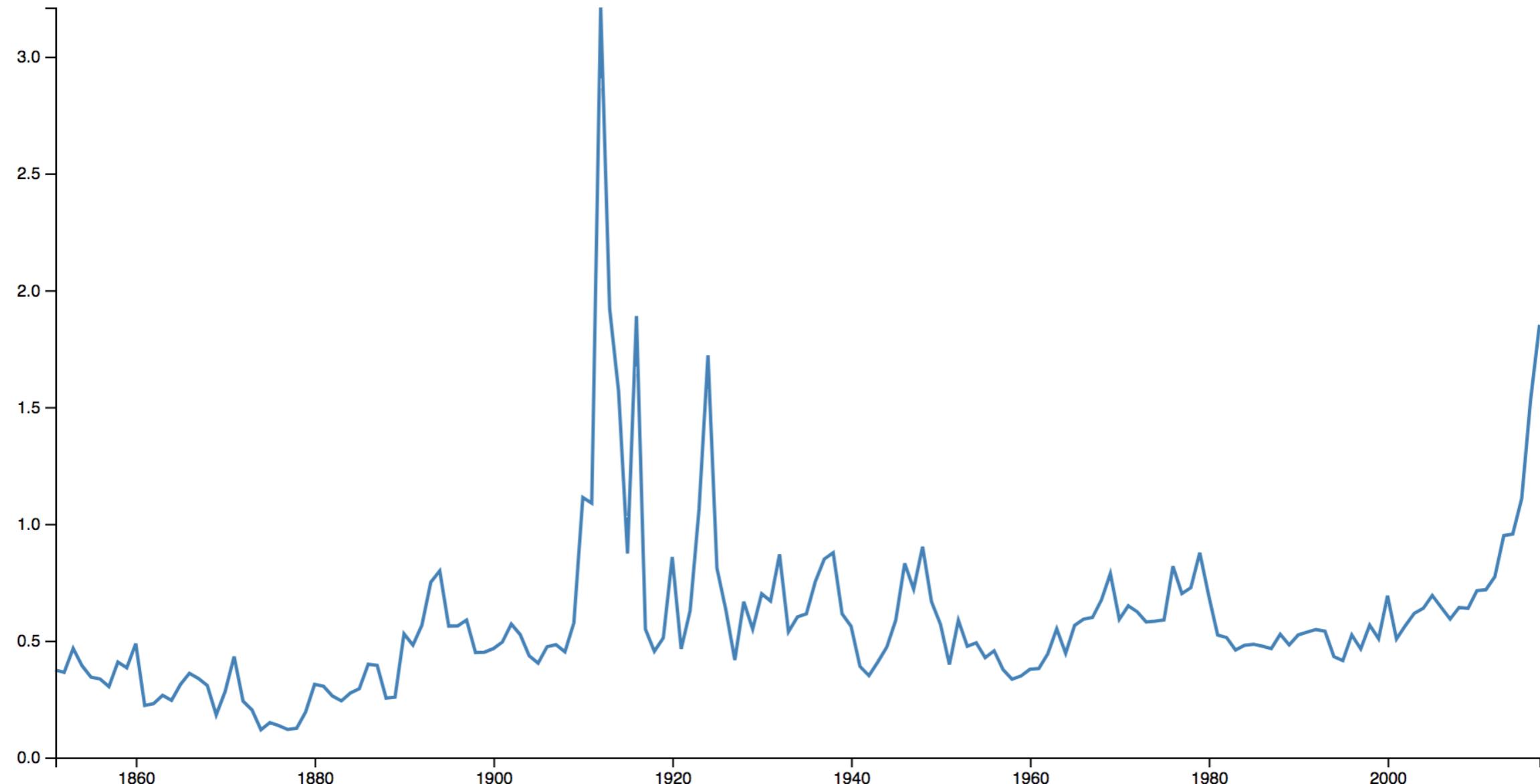
Currently in use by *The New York Times* to track word usage

Previous examples lack historical context or emotions of the times.

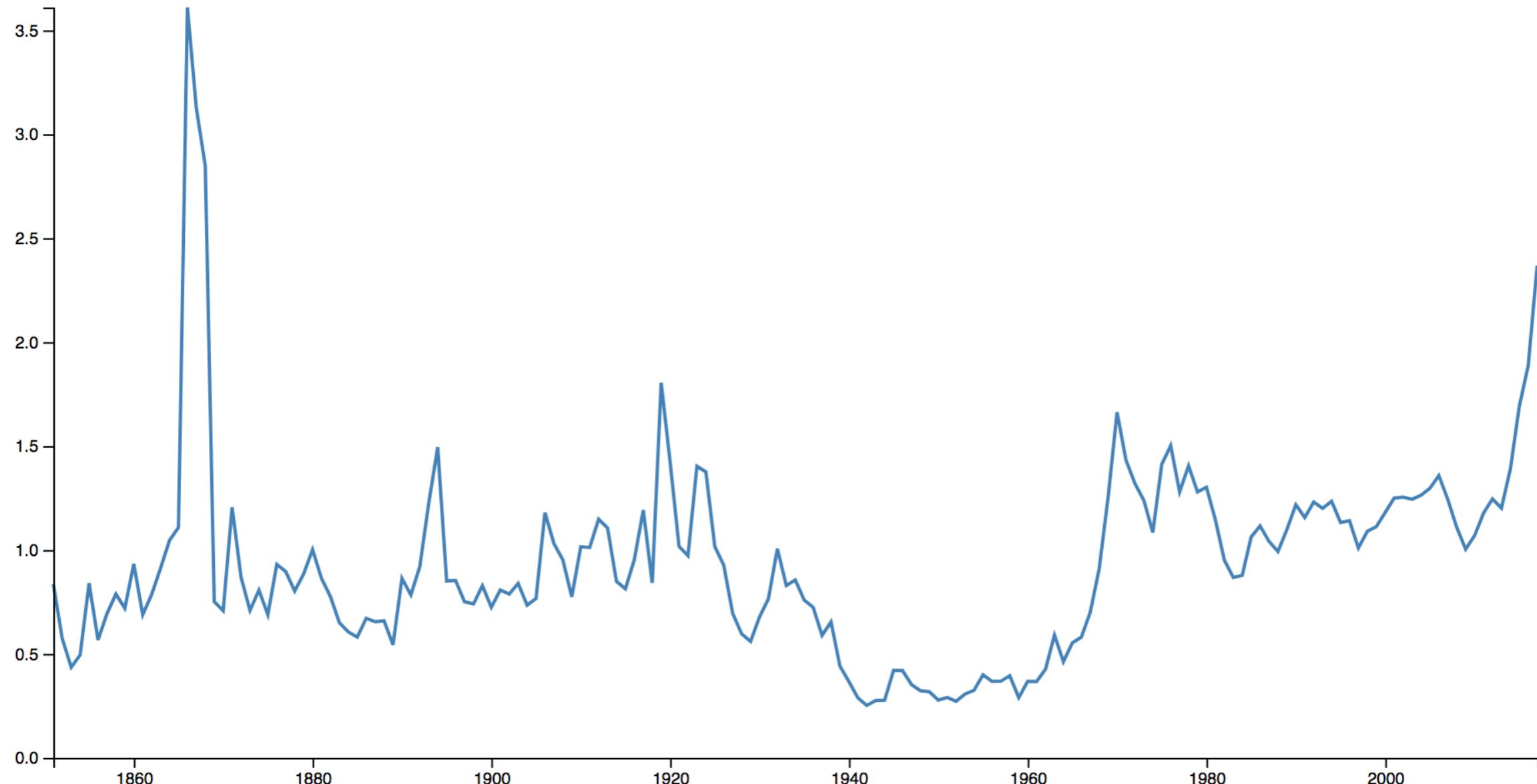


PROGRESSIVE

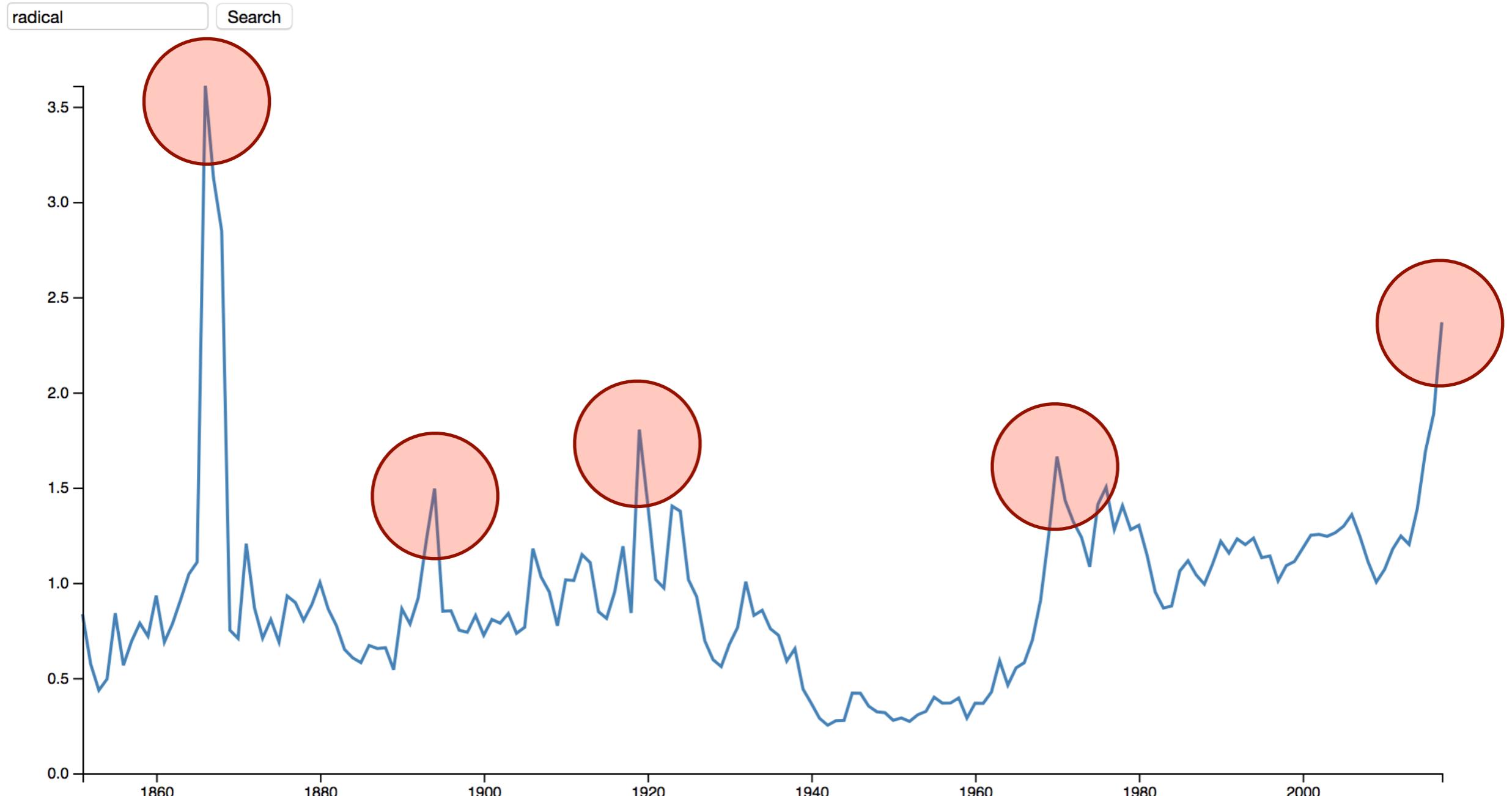
progressive



RADICAL



RADICAL



*What was happening in society
during these peaks and valleys?*

Data Collection Difficulties

- NYTimes API limited amount of calls
- Only chose two words: radical and progressive
- Lengthy, monotonous process

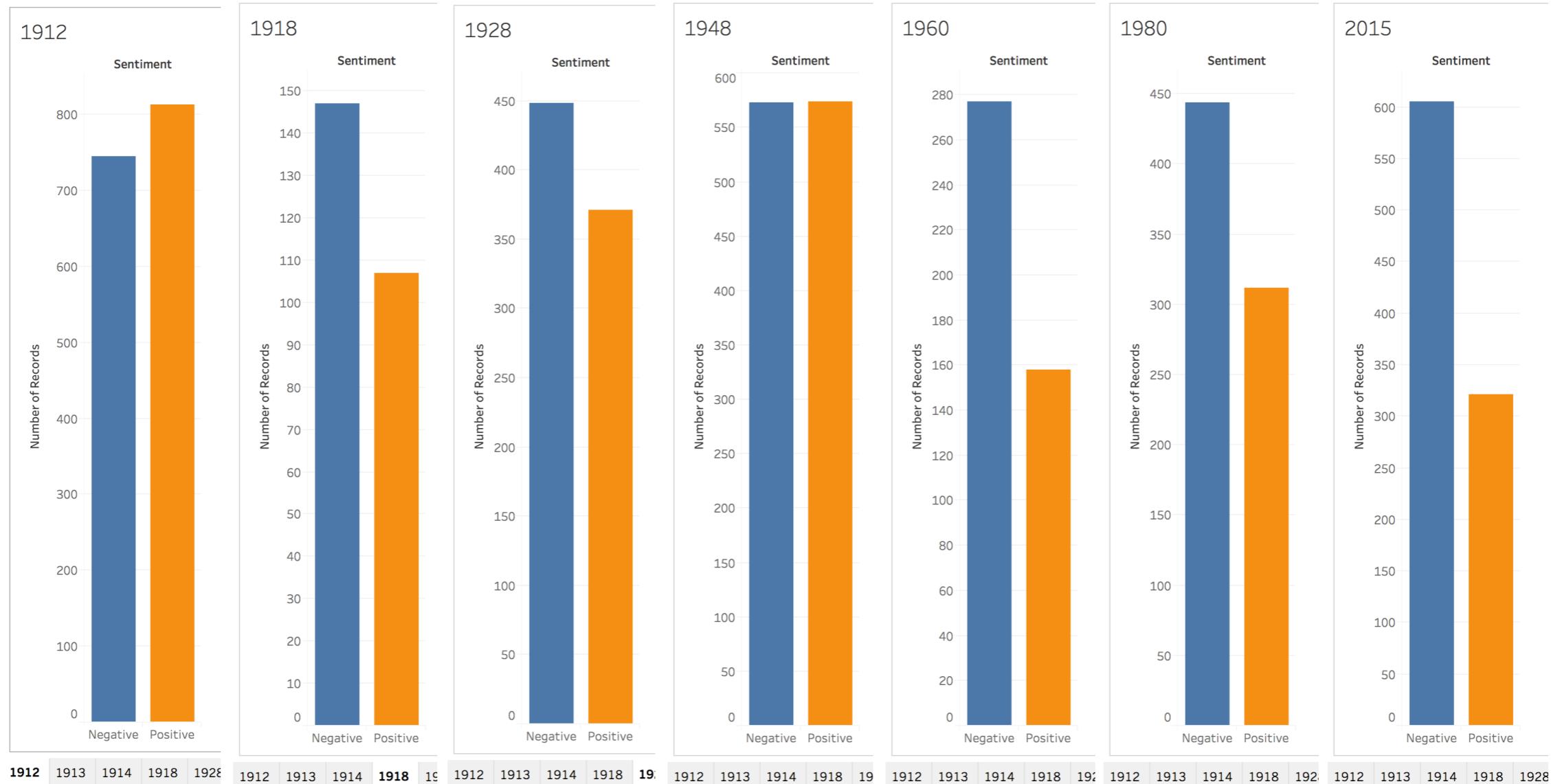


Positive



Negative

PROGRESSIVE



What was the sentiment of the articles being reported?

Sentiment Analysis Difficulties

- Either unsupervised or use someone else's tool
 - LIWC2015, Lexalytics, LightSide
- Analysis on lead paragraphs - low word count
- Commentary on *New York Times* articles rather than the sentiment of the time.

Our Tool

Looking Ahead

- Scalable tool allowing for word search from all First Word articles
- Create our own sentiment analysis model from our labeled data.
- Allow a user to explore more of the NYTimes API through these words-reader driven approach.

“We have to be really careful that people feel like they can see themselves in The New York Times. I want us to be perceived as fair and honest to the world, not just a segment of it. It's a really difficult goal.”

—New York Times Executive Editor