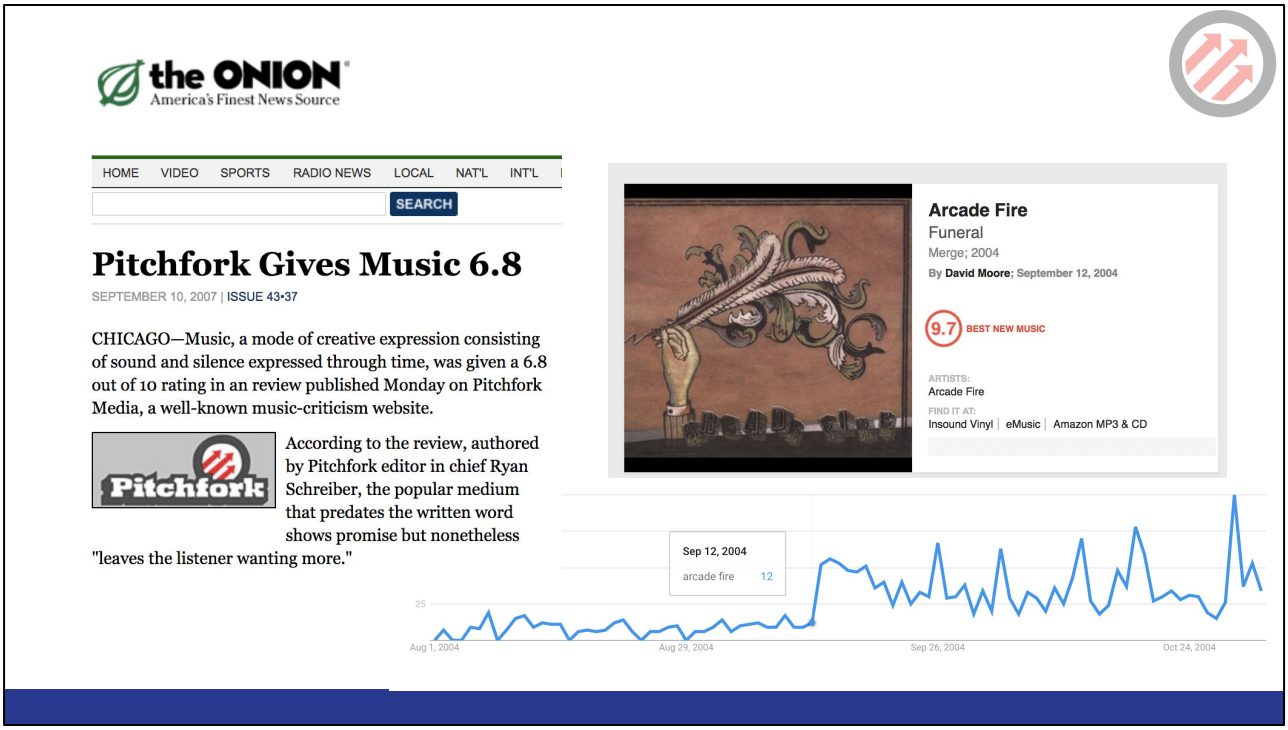
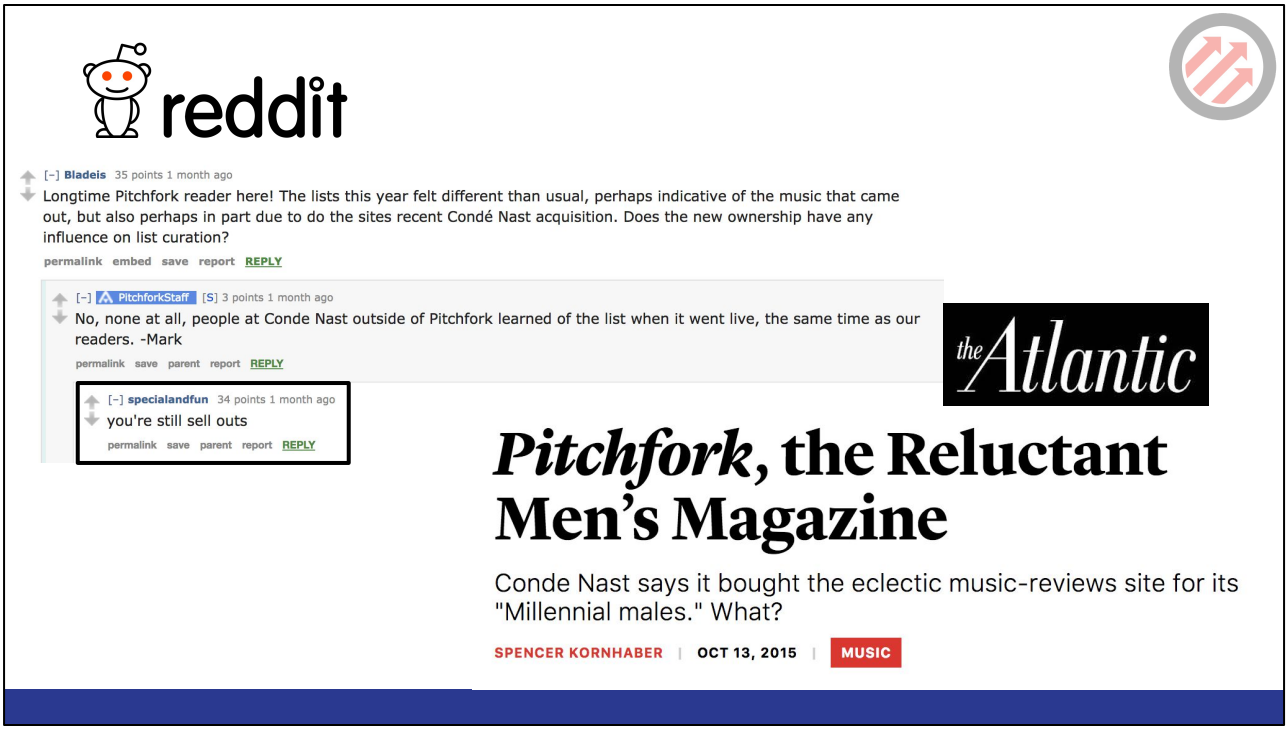
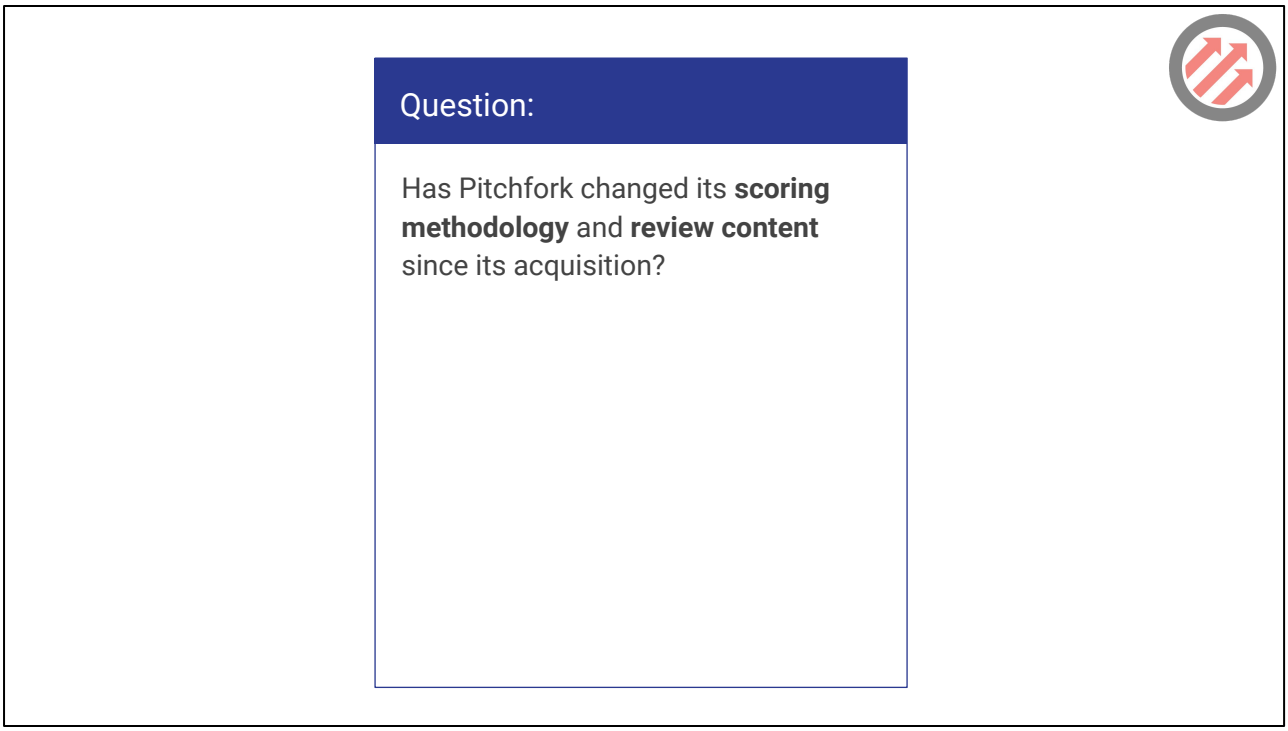
Hello! My name is Andy Bashford. I'm a data scientist with an interest in music and music journalism.



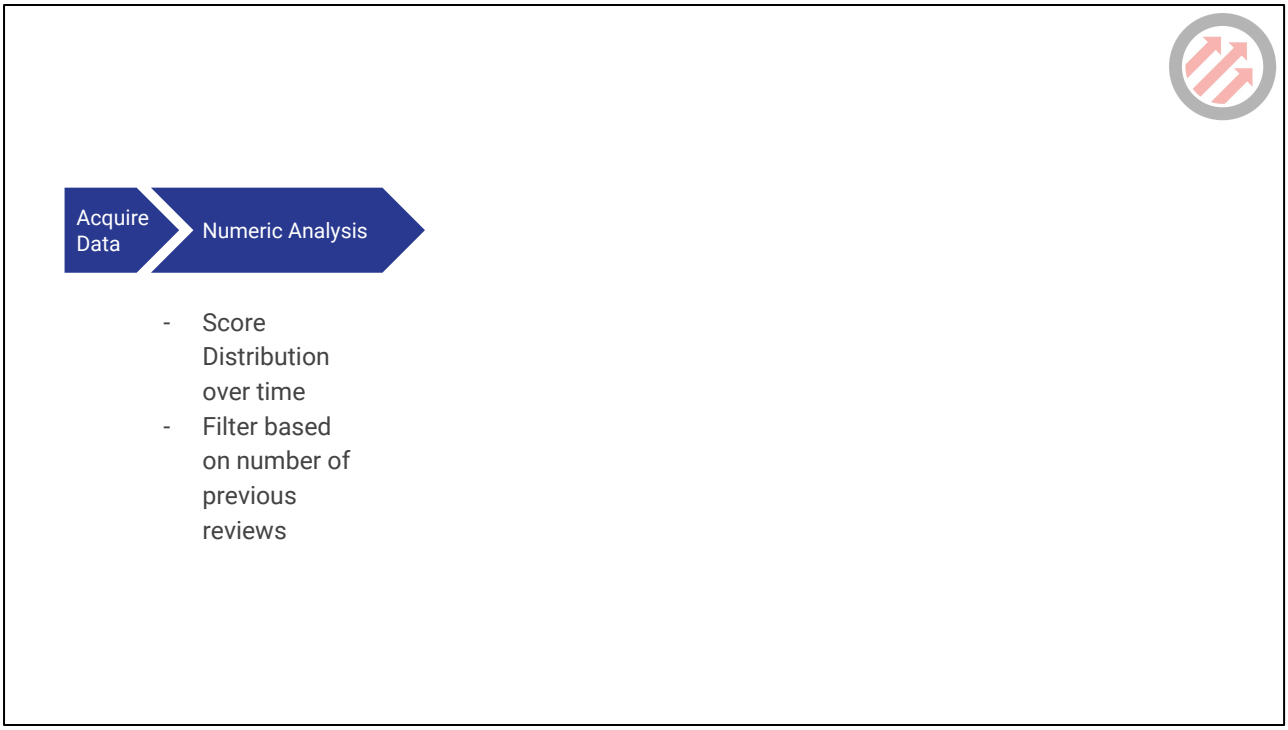
My capstone for Galvanize focuses on Pitchfork. Pitchfork is an online music publication with a reputation for being opinionated, snobby, but usually right.



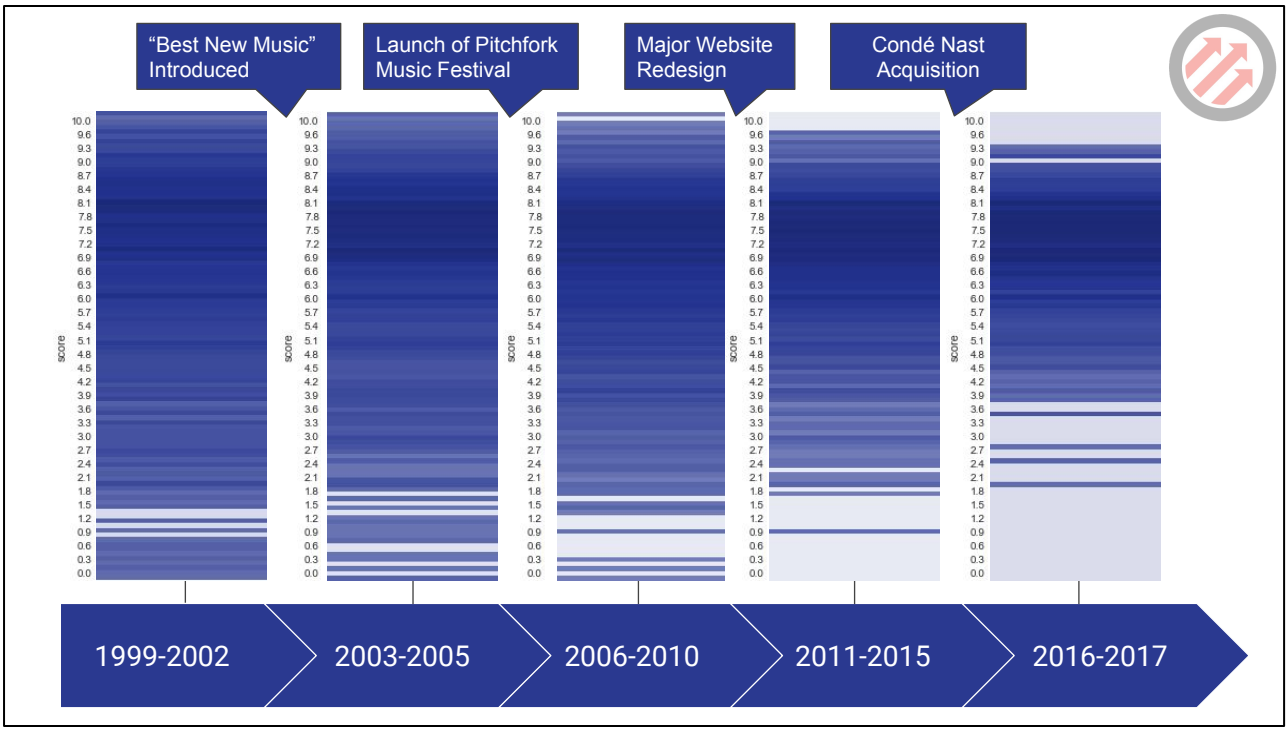
In 2015, they were acquired by the large media company Conde Nast, leaving fans skeptical of the future of Pitchfork's reputation.



My question: Has Pitchfork changed its scoring methodology and review content since its acquisition?

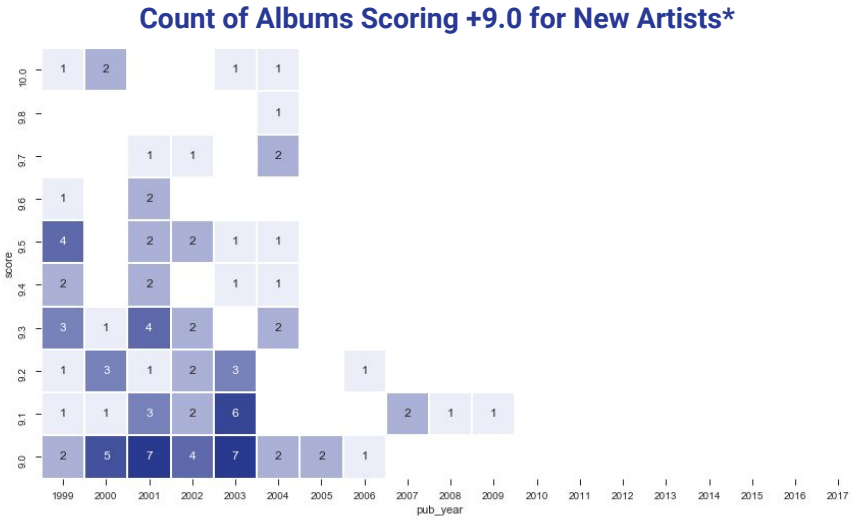


First, I scraped Pitchfork's archives to gather over 19,000 reviews since 1999 and performed some numeric analysis.



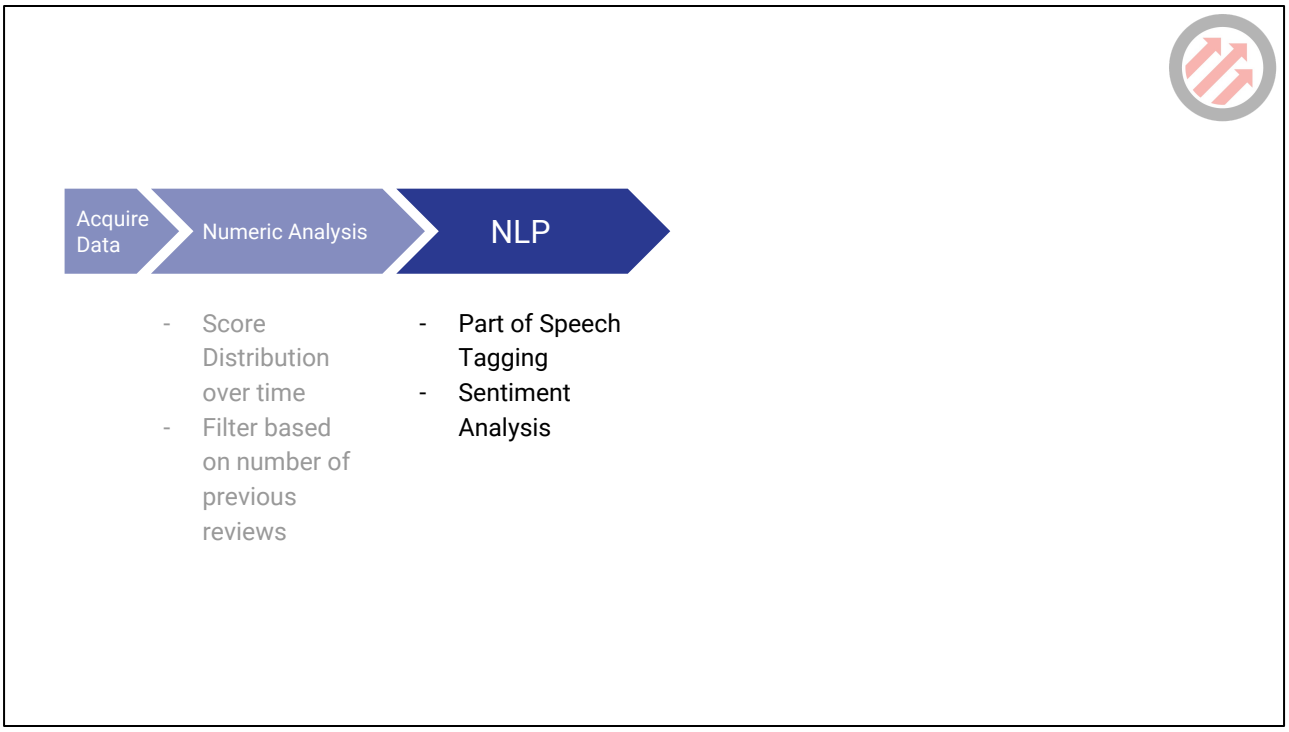
I broke down Pitchfork reviews into five major buckets. One before BNM, one before the launch of Pitchfork Music Festival (the first time Pitchfork had an incentive to promote specific artists), one before a major website redesign in 2011, one before the Conde Nast acquisition, up to present day. It's pretty plain to see how the spread of scores given to albums has tightened with each major event.

Does Pitchfork Still Gamble on New Artists?



** Artists with 0 previous reviews on Pitchfork at time of publication*

Next, I wanted to evaluate if Pitchfork is still willing to gamble on new artists. In 2004, they gave Arcade Fire a 9.7 for their first release, putting them on the map. But, since then, they have only given an album a 9.0 or above 8 times when the artist is being reviewed for the first time… zero times since 2010.



After finding that the distribution of scores has changed, I used Natural Language Processing to evaluate the content of the reviews.

POS Tagging

```
abstract = """Sufjan Stevens has always written
              personally, weaving his life story
              into larger narratives, but here his
              autobiography is front and center.
              Carrie & Lowell is a return to the
              stripped-back folk of Seven Swans but
              with a decade's worth of refinement
              and exploration packed into it."""
import spacy
nlp = spacy.load('en')
```

```
make_POS_df(abs_desc, nlp)
```

	word	POS	tag
0	has	VERB	VBZ
1	always	ADV	RB
2	personally	ADV	RB
3	weaving	VERB	VBG
4	larger	ADJ	JJR
5	here	ADV	RB
6	is	VERB	VBZ
7	front	ADJ	JJ

spaCy

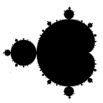
Sentiment Analysis

```
abstract = """Sufjan Stevens has always written
              personally, weaving his life story
              into larger narratives, but here his
              autobiography is front and center.
              Carrie & Lowell is a return to the
              stripped-back folk of Seven Swans but
              with a decade's worth of refinement
              and exploration packed into it."""
from textblob import TextBlob
TextBlob(abstract).sentiment
```

```
Sentiment(polarity=0.049999999999999996, subjectivity=0.25)
```

```
parse_sentiment(abstract)
```

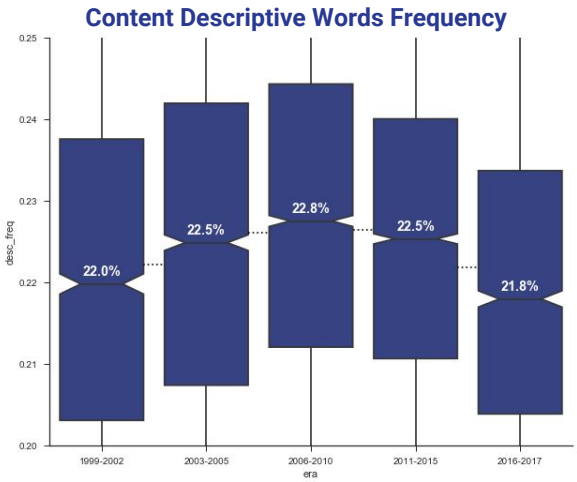
```
personally,
Subjectivity: 0.3

larger
Subjectivity: 0.5

center.
Subjectivity: 0.1

worth
Subjectivity: 0.1
```



TextBlob

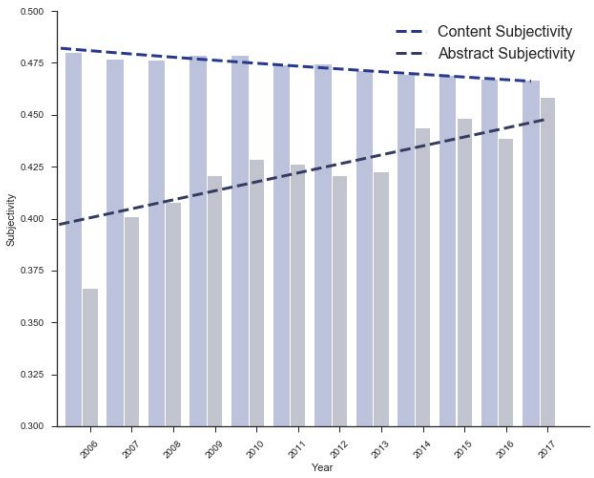
I used the tool SpaCy to tag all 14 million words from reviews with their part of speech, counting up the words that are most likely to be descriptive (adjectives, adverbs and certain tenses of verbs). I then used TextBlob to get a subjectivity score for each review, measuring how opinionated or objective the author was.

Part of Speech Tagging



** Notches indicate 95% confidence intervals for the median*

Sentiment Analysis



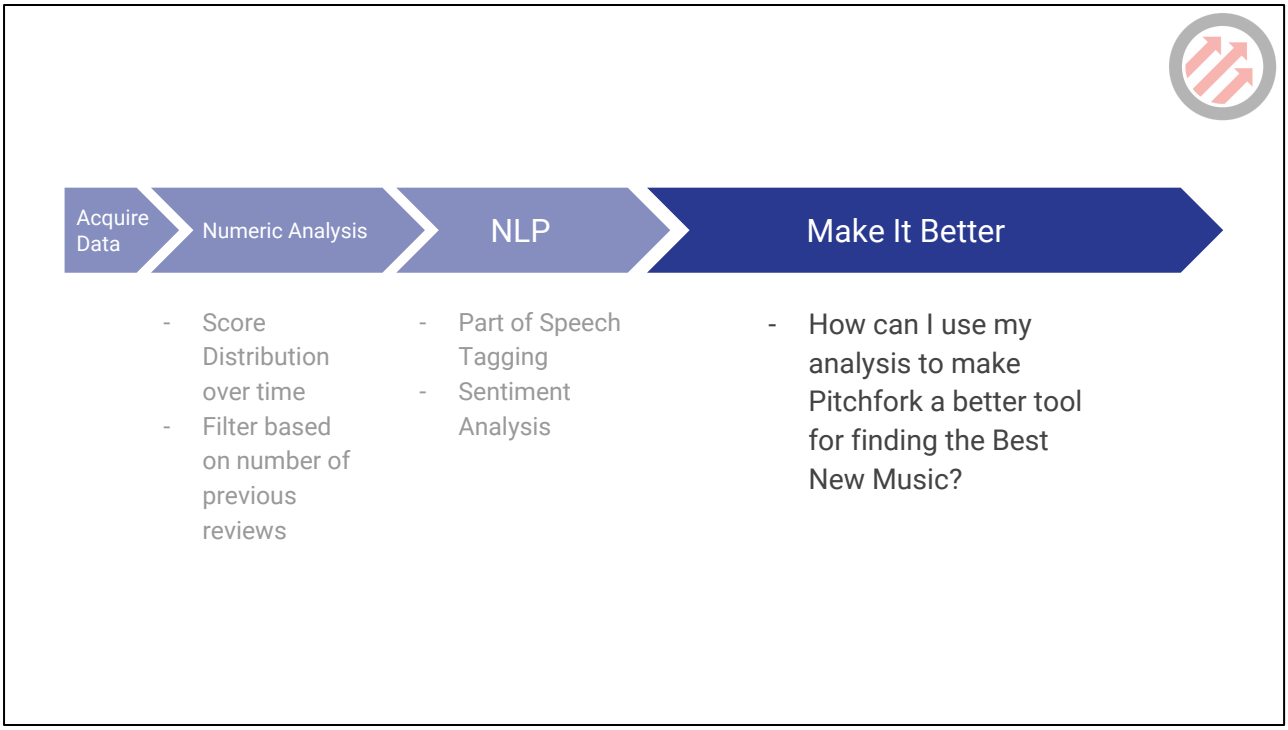
I found that since 2003, Pitchfork reviews are getting less descriptive. The boxplots have notches indicating a 95% confidence interval of the median, showing that these changes are statistically significant. I also found that, while the content of the reviews has become more objective, the one-to-two sentence abstract at the top of the review has become more subjective. The subjectivity of the content and the abstract are converging together, giving the reader a more consistent account of Pitchfork's opinion.

Question:

Has Pitchfork changed its **scoring methodology** and **review content** since its acquisition? **Yes!**

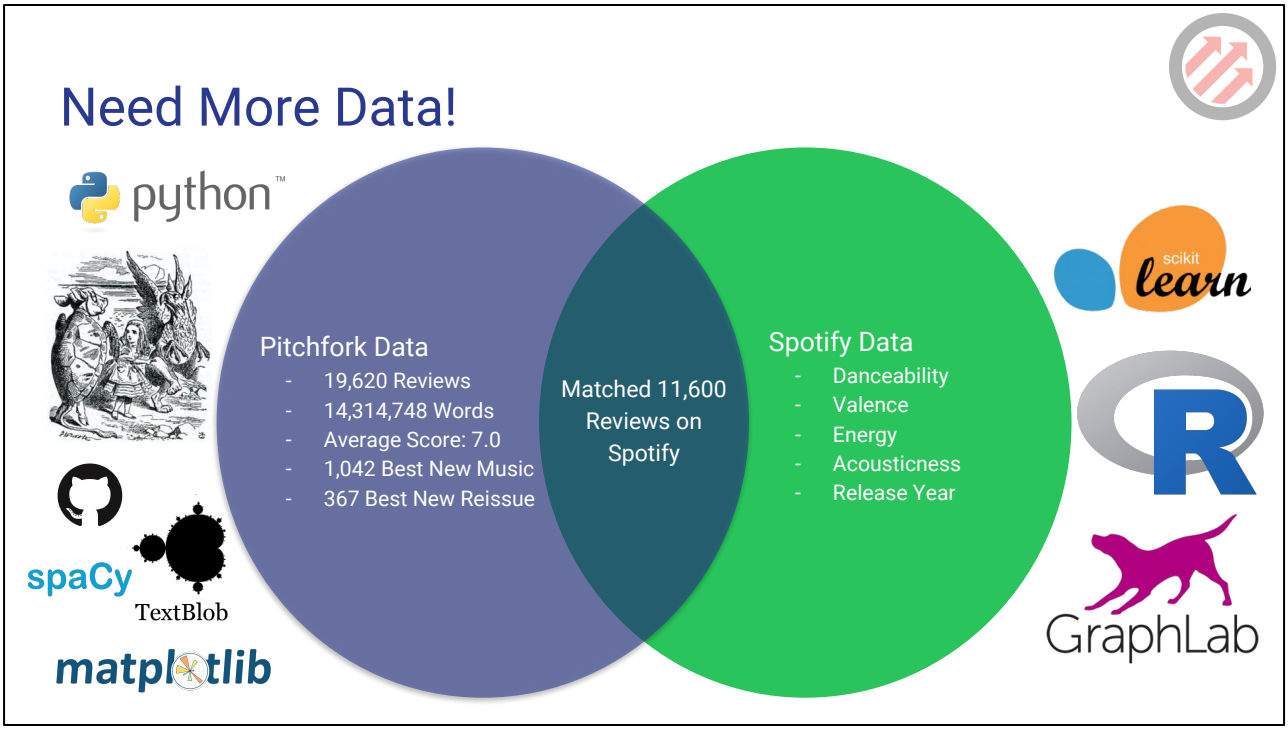
- **The spread of the score distribution has tightened**
- **Pitchfork doesn't gamble big on new artists**
- **Pitchfork's language has become less descriptive and more objective**
- **Pitchfork has become more consistent in the Abstract**

So, has Pitchfork changed its scoring methodology and review content since its acquisition? Yes! The spread of the score distribution has tightened, Pitchfork doesn't gamble big on new artists, Pitchfork's language has become less descriptive and more objective, and Pitchfork has become more consistent in the Abstract

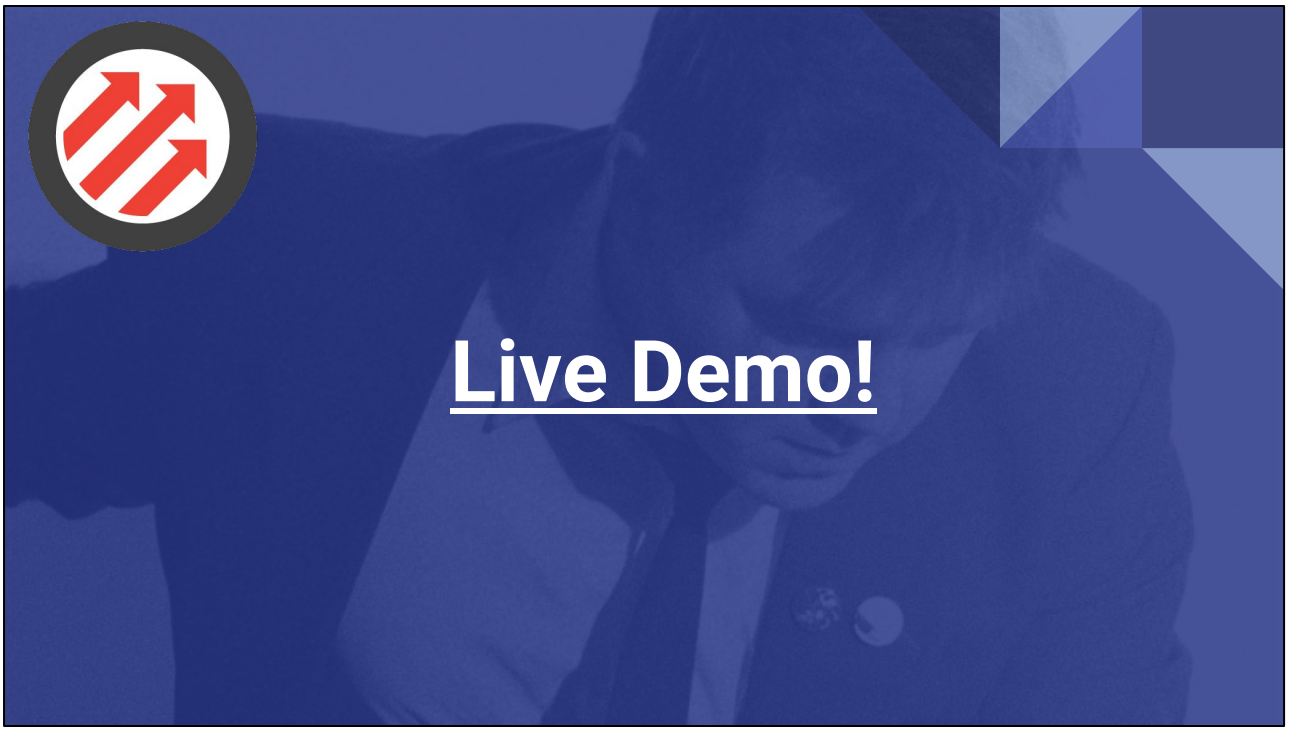


I didn't stop there. My next step was to use the data I collected to make Pitchfork a better tool for finding new music, since at the end of the day that's what most people are after when they are visiting the site.

Pitchfork Review
Recommender System



I took the review data I had collected and merged it with data from Spotify, where they have done audio analysis on 3-second samples of tracks to get some key metrics. I was able to match 11,600 albums from Pitchfork with data from Spotify.


Live Demo!

Github: https://github.com/andrewjbashford
LinkedIn: https://www.linkedin.com/in/andybashford/