Creating Interactive Visualization Software

IMT 589

Kevin McCraney

Motivation and Dataset

Motivation

Pitchfork is an interesting anomaly in the world of critique. Founded by Ryan Schreiber in 1995, it was one of the first music journalism outlets to come out of the blogosphere and turn into a cultural institution. Pitchfork started off with an occasionally experimental style and was also the launching point for several now-acclaimed authors or journalists. However, I don't want to do an oral history of this publication. I want to think about critique, particularly how the character of it has changed in the age of the internet.

What is critique? For my purposes, it's the process of applying an expert's eye to a cultural work. The expert provides context to guide the reader through the intricacies of the work they're discussing. In the case of music criticism, they might be conversant in the field by providing historical or musical context to a particular work. However, as far as I know, no one has produced music critique that looks *at a distance*, rather than at a particular work.

As part of this project, I hoped to use my analytical skills to produce a new lens, through which we can view the work Pitchfork is doing. I achieved this through several means: by plotting the rolling mean of review by genre in Tableau; and by performing sentiment analysis on the writings Pitchfork's reviewers using the VADER method, plotting those data against the average score for each reviewer. I believe that looking at the average sentiment and the average score together might allow us to better understand whether an author was writing using more even, neutral-evaluatory language or more decisive, critical-evaluatory language.

Dataset

Initially, I sought to scrape Pitchfork's reviews into a NoSQL database, clean the reviews, and then perform analysis. However, Pitchfork severely rate-limited me and encrypted all the results. I ended up using a pre-existing dataset that contains all the reviews until 2017. There are 18,393 reviews in total. I used these reviews to produce a series of smaller datasets (CSVs) in Python.

Storyboard

I began this project wanting to address four key points: a network of influence; a mapping between the average sentiment of a reviewer and the average score that reviewer gave; a boxplot detailing genre distribution statistics; and a plot of average scores per genre. I outline the four different visualizations in this sketch below.

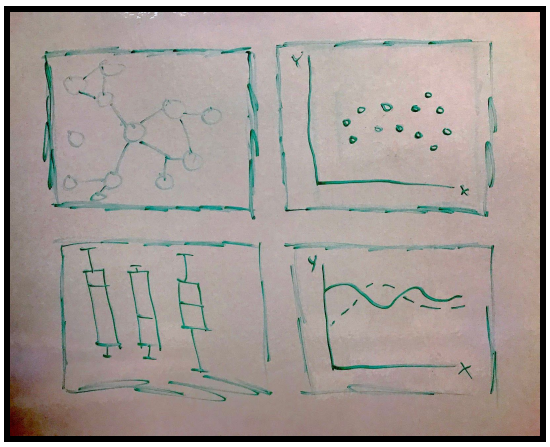

Figure 1; Storyboard/Sketch

I imagined that users would have fairly minimal interaction with the visualizations themselves, hovering over particular data points to get some key information.

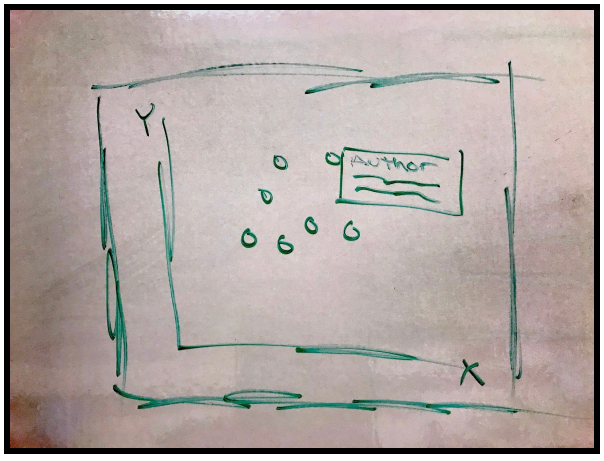

Figure 2; Interaction

Initial Prototype and Reorientation

As above, what I wanted to do first was produce a network visualization of all of the hyperlinks cited in every review, in order to show whether or not a particular author had more outsize influence to others (through more frequent references to their writing). Then, users would be able to scale the results based on most or least cited pages in order to see which things Pitchfork reviewers were referencing.

However, this proved too arduous for my computer to do. I could produce a network visualization, but it was unable to scale the data such that it was interpretable. It took so long for the browser to load and plot the dataset that I eventually gave up.

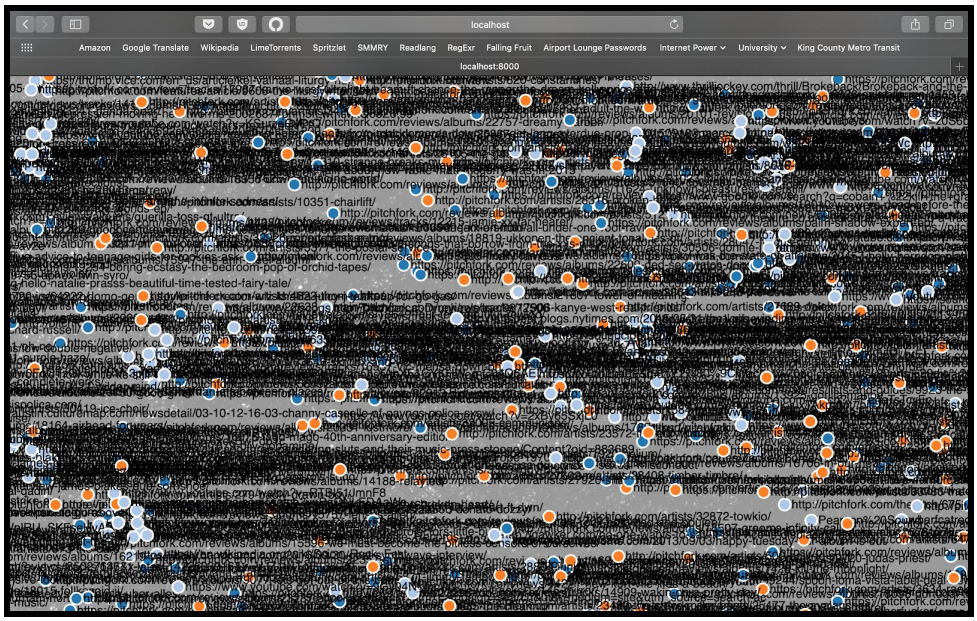

Figure 2; Aborted Network Visualization

What I ended up doing instead was spending most of my time writing code to structure, analyze, and aggregate data in order to produce a CSV that contains metrics related to reviewers.

One of my primary metrics was the VADER sentiment metric. The VADER sentiment metric is a score based on the presence of a word in a dictionary of positive, negative, or neutral words. The score comes from the ratio of a particular type of word to all words present in the document. It's a somewhat coarse metric, because it looks at individual words rather than entire sentence context, but it is useful here to gauge simple

relationships across a corpus (body of texts). If a reviewer has a score closer to 0, we can assume their review uses more neutral language. If the score is further from 0, then their language is more critical.

Actual Visualizations

The visualization I have produced is a simple scatterplot. There is minimal user interaction--no clicks, menus, or scrolling is possible. A user can hover over the data points and get some key information about the reviewer's average score, their sentiment score, and the number of reviews they completed. In that way, the visualization is not at all dissimilar from the first prototype.

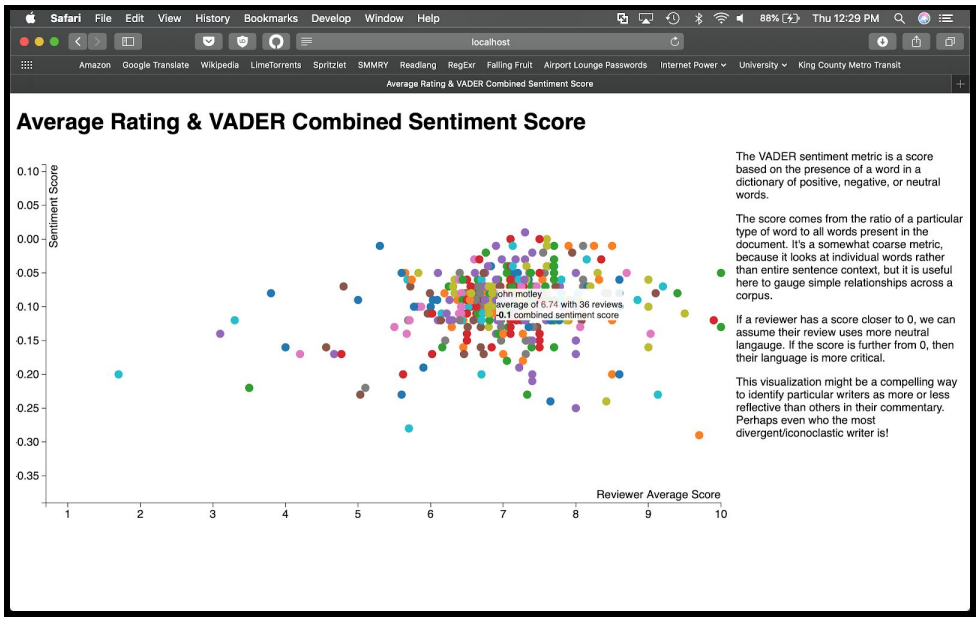

Figure 3; Rating/Sentiment Visualization

You can view this project at the following URL:
https://kevinrmccraney.github.io/pitchfork-visualization/

Development Process and Challenges

Practically speaking, the hardest task I encountered during this process was dealing with the repercussions of scraping the Pitchfork website, and figuring out how to pivot when I was unable to get the data I needed. That's what caused my pivot toward the end of the process. I would say that I spent about a solid week of work (8hr days) trying

to get data, get it in the proper form, and use it. Developing the visualization itself was a mere afterthought in comparison.

As part of this project, I relied heavily on Mike Bostock's ObservableHQ visualization resource. Many different developers have provided extensible visualization tools and datasets. I have cobbled together my work from several examples and pieces on there. Mike's work is obviously informed by the venerable Edward Tufte, which I had previously encountered in my education and reviewed in order to supplement my knowledge.

Like many of my peers, I also used some of the guiding principles and ideas from Udacity's course on d3. In addition to the curriculum available there, I used Michael Freeman's book "Interactive Information Visualization" to supplement my knowledge.

One of the ways I might expand this project in the future is to allow the user to filter by number of reviews. That way the critical evaluation can be focused on people who have stronger engagement with the site over time, rather than a one-off review. Additionally, I'd like to use a LSTM neural network model to figure out the sentiment on a sentence-by-sentence level.

References

Conaway, Nolan. “18,393 Pitchfork Reviews.” Kaggle, 13 Jan. 2017, https://www.kaggle.com/nolanbconaway/pitchfork-data

“Data Visualization and D3.Js.” Udacity, Udacity, www.udacity.com/course/data-visualization-and-d3js--ud507

Freeman, Michael. *Interactive Information Visualization*. GitHub Pages, 2017. https://info474-s17.github.io/book/index.html

Pandey, Parul. “Simplifying Sentiment Analysis Using VADER in Python (on Social Media Text).” Medium, Analytics Vidhya, 23 Sept. 2018, https://medium.com/analytics-vidhya/simplifying-social-media-sentiment-analysis-using-vader-in-python-f9e6ec6fc52f

Tufte, Edward R., 1942-. *Visual Explanations : Images and Quantities, Evidence and Narrative*. Cheshire, Conn. :Graphics Press, 1997. Print.