Visualizing Market Data for Mobile Applications

Jeffrey Xiao, Dengning He

Abstract

In an era where convenience is king, 77% of US adults own a smartphone which on average has enough computational power to rival that of a midrange desktop from just a few years in the past. Because of this, the mobile app market has grown larger than ever, with 2.1 million apps available on the Google Play and 2 million apps on the Apple App Store in 2018. With so much competition, development teams often have trouble ensuring that their apps stand out in today's oversaturated market. The app stores themselves have layouts more tailored to the consumer than the developer, with navigation designed to lead them to trying out more applications and little information about the application beyond a description, rating, pictures and a couple reviews. There plenty of tools available for professional market analysis, but almost all of them feature plain tables and line graphs with relatively bare-bones data.

In this project, we will aim to create an interactive visualization given the app data that would allow a market research team to better navigate, visualize, and draw comparisons between applications for mobile devices. In addition, we aim to use advanced tools such as sentiment analysis to construct graphs over time that would better illustrate how an app is performing at a certain time due to what factors. It would also have the ability to display two different apps side-by-side and comparison by digital distribution platform.

Summary

Given data from the Apple App Store and the Google Play Store, create an interactive application featuring a circular network diagram that allows users to explore apps and draw novel conclusions based on key performance indicators (KPI) and sentiment analysis of reviews over time.

Project Type

Interactive Visualization

Major Projects Referenced

“Why People Hate Your App — Making Sense of User Feedback in a Mobile App Store” is a paper that was published by Carnegie Mellon detailing much of the sentiment analysis process for reviews of a mobile application. It discusses the method of collecting, parsing, and analysing the data.

Federica Fragapane and Giorgia Lupi work on a data visualization team whose main goal is to represent data accessibly, often in such a way that represents the complex and nuanced human nature behind the raw numbers. During her TED Talk, Giorgia Lupi discussed how part of the reason why we were so bad at predicting outcome of the 2016 US presidential election was because we failed to consider the story behind the data. In our project, we will aim to borrow some of the philosophy behind their designs as well as some of their visualization techniques.

Target Audience

Market research team interested in sentiment analysis;
The project manager who is monitoring the app performance over time;
Random users interested in the mobile app market.

The common ground of the target audience is the mixture of novices and experts, which requires the design to balance easily understandable visualization and in-depth comparison.

Key Features

1. Identify the desired app using filtering and sorting
Using multiple attributes such as star rating, category, and publication date, the user can define their targeted subsection of the market.

2. Gain the insights on the contribution of each review category to final rate
After exploring the 42matter's dataset and API, we found that each app has sentiment analysis based on user reviews and sorted by category. Each category adds a weighted score base on the user favorability to the final sentiment rating. We would like to visualize these categories in such a way that the user can intuitively view them.

3. Compare the two app performance based on sentiment analysis
Our visualization should also allow the user to compare at least two selected apps across the two most popular mobile application distribution platforms (Google and IOS), giving sentiment analysis data over time on each app in the side panel.

Milestones

Gain approval on project prospectus

Explore data using R and Python and modify the ideation based on the results

Create a high-fidelity design prototype and style-guide.

Chart the original circular network diagram, line charts, and pie chart using d3.js

Link dataset and API to the demo and test functionalities

Connect our reviews dataset to an external sentiment analysis software such as Google NLP

Layout the page and finalize the visualization

Design usability testings and gain feedbacks

Iterate the original visualization based on the results of the UX research.

Definition of Success

After deploying the visualization, we are going to conduct usability tests to our target users. The use should be able navigate through visualization as quickly as they could through a mobile application distribution platform with the same if not more grouping and filtering features. They should also be able to gain insight beyond that of the application distribution platform.

The evaluation process would be designed in the future. The iteration of this project will be conducted based on further requirements.

References

Fu, B., Lin, J., Li, L., Faloutsos, C., Hong, J., & Sadeh, N. (2013). Why people hate your app. Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 13. doi:10.1145/2487575.2488202

Oelke, D., Hao, M., Rohrdantz, C., Keim, D. A., Dayal, U., Haug, L. E., & Janetzko, H. (2009, October). Visual opinion analysis of customer feedback data. In *2009 IEEE Symposium on Visual Analytics Science and Technology* (pp. 187-194). IEEE.

Gregory, M. L., Chinchor, N., Whitney, P., Carter, R., Hetzler, E., & Turner, A. (2006, July). User-directed sentiment analysis: Visualizing the affective content of documents. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text* (pp. 23-30). Association for Computational Linguistics.

Wang, C., Xiao, Z., Liu, Y., Xu, Y., Zhou, A., & Zhang, K. (2013). SentiView: Sentiment analysis and visualization for internet popular topics. *IEEE transactions on human-machine systems*, *43*(6), 620-630.

Lupi, G. (2017, March). Newton Aduaka: How we can Find Ourselves in Data [Video file]: https://www.ted.com/talks/giorgia_lupi_how_we_can_find_ourselves_in_data