Causal Impact for App Store Analysis

William Martin <w.martin@ucl.ac.uk>
University College London, UK

Design

App Store Analysis

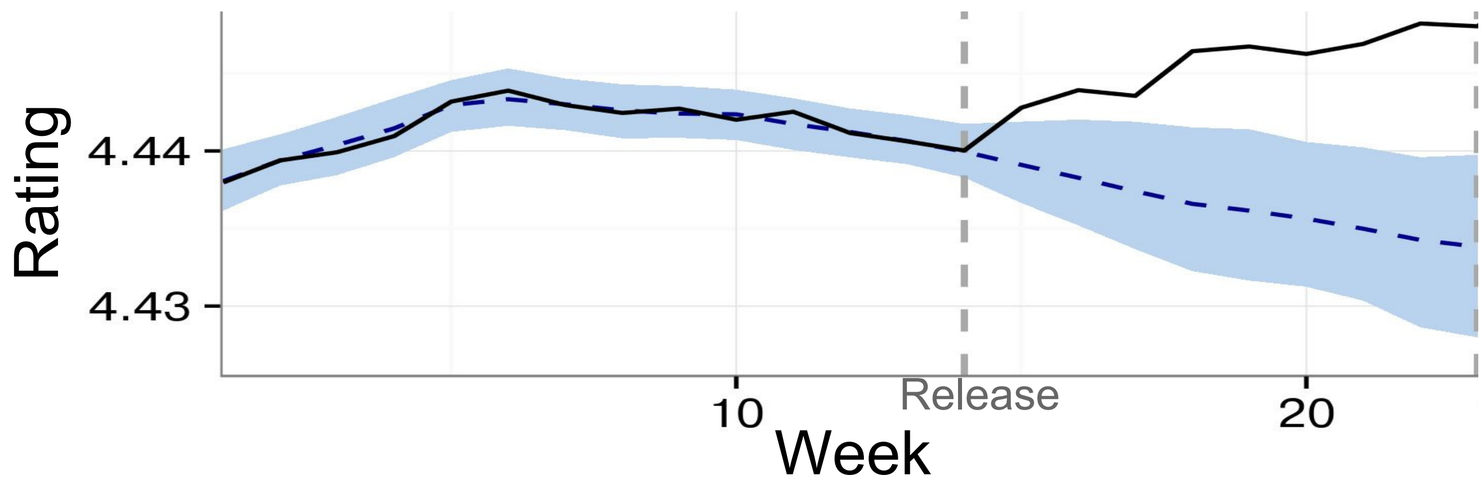
App developers naturally want to know which of their releases improve performance and which hinder it. App stores provide a method of measuring app performance, in the form of download ranks, ratings and reviews. Using recorded time series information, we can analyse the impact that releases had on these metrics.

Dataset

I collected information on the rating, download rank and number of ratings from Google Play and Windows Phone Store, over 52 weeks, for the apps which were consistently in the most popular store lists.

	Google	Windows
Apps	210	539
Releases	754	793
Non-releasing control set apps	97	397

Causal Impact Analysis



Causal Impact Analysis [2] is a form of causal inference, that works by training on the pre-event data vector in order to make a (counter-factual) prediction after the event. This prediction tells us the most likely course of the vector after the event, from which we can determine whether a significant change has occurred in the observed data vector. It accounts for global variance using a control set.

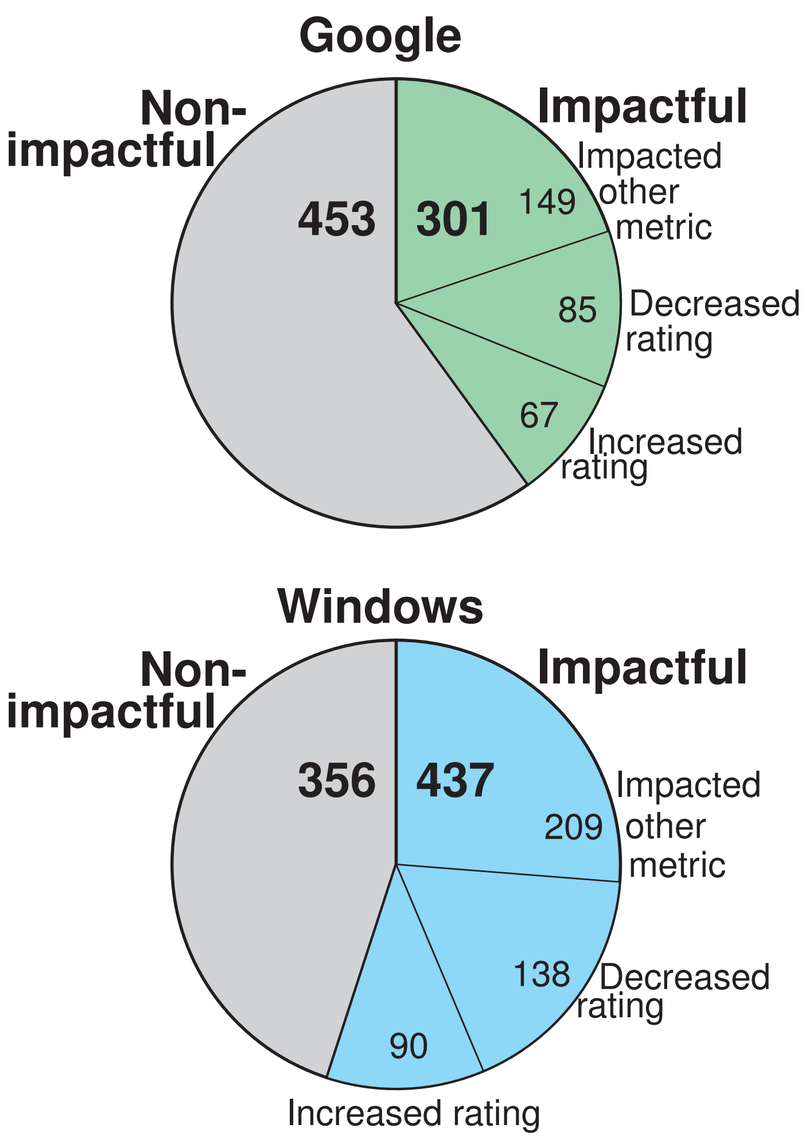
In the case of app release analysis, releases serve as the events and non-releasing apps serve as the control set.

In the above example, the release version `6.0` of the Google Play app `Flightradar24 - Flight Tracker` causes a significant increase in the rating (a small increase in this case corresponds to thousands of users rating higher than the mean). The observed rating, shown as the solid line, clearly deviates significantly from the prediction, shown as the dotted line. The 95% confidence interval is plotted as the shaded blue region.

Causal impact analysis is applied using `CausalImpact` [1] in these experiments.

Results

App Release Analysis



Candidate Causes of Impacts

To identify potential causes for the impacts observed, I compared impactful against non-impactful releases in each store. I also compared releases that increased rating against those that decreased rating.

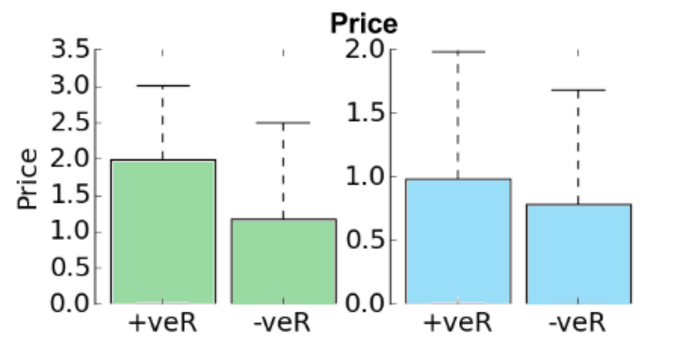
Price - Rating
Expensive apps more likely to positively impact rating

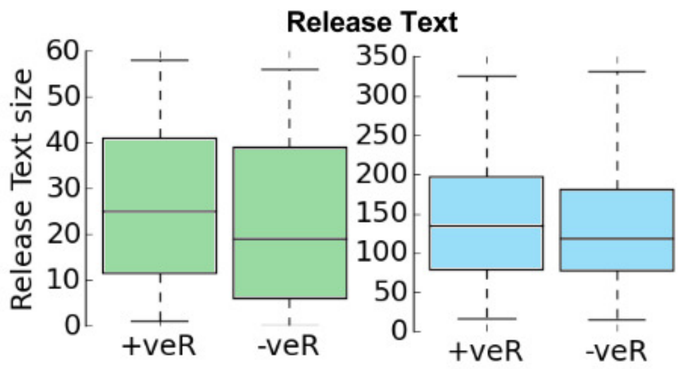
New feature - Rating
2% more positive releases introduce features

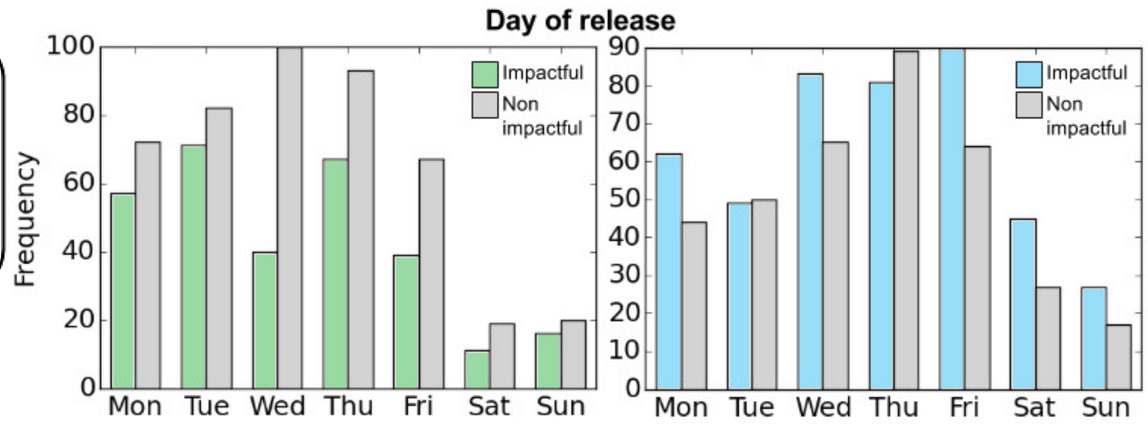
Release day - Impact
A greater proportion of releases achieve impact if released at the weekend

Bug fix - Rating
4% more positive releases fix bugs

Release text - Impact + Rating
Impactful releases and positive releases have longer, more descriptive release text







Future Work

Since this preliminary study I have developed an alternative tool for performing causal impact analysis on large sets of app store data, and continue to investigate the properties of impactful releases in larger datasets.

References

[1] K. H. Brodersen. CausalImpact. `https://google.github.io/CausalImpact/CausalImpact.html`. Retrieved May 28 2015.

[2] K. H. Brodersen, F. Gallusser, J. Koehler, N. Remy, and S. L. Scott. Inferring causal impact using bayesian structural time-series models. *Annals of Applied Statistics*, 9:247–274, 2015.

William Martin