

Examples of AB Testing

Sam Pritchard

December 26, 2017

Headline:

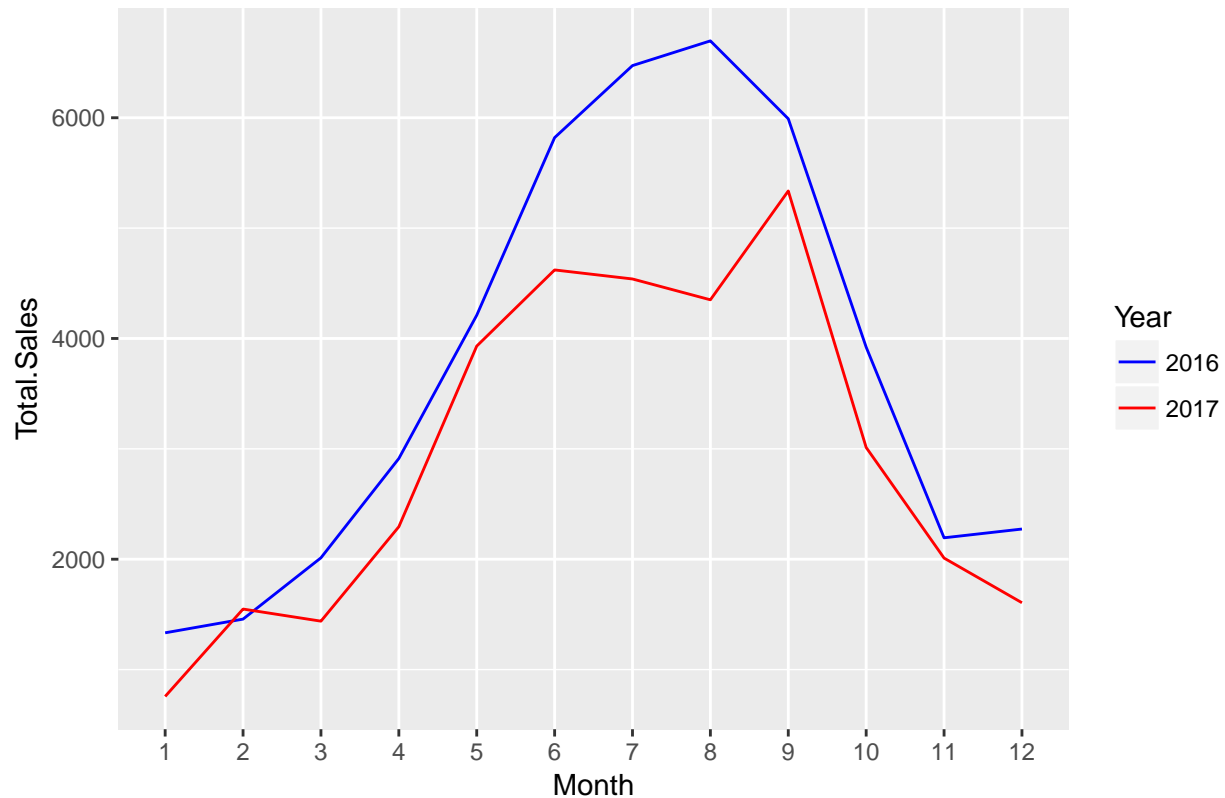
In our last conversation, Colin asked if I could help make it easier to do AB tests; comparing the results of 2 options. This document lays out a general direction for how I can help you do AB tests. Essentially, I plan to set up a website that would allow you to compare almost anything with anything. Below I show the two main components I'm planning to use for that site: the AB Arrow and AB Trails.

Demonstrating AB Test Visualizations

Scenario 1: Sellers Publishing in June-Sept 2016 vs. 2017

As shown below 2017 lagged behind 2016 significantly. If we were to do an AB test using something called a T-Test, how soon could we determine 2017 was lagging? This same T-Test approach can be used in comparing a new supplier with an existing one, comparing a pricing change to the status quo, etc. Scroll down for more.

Trend Year Over Year for SELLERS PUBLISHING, INC.



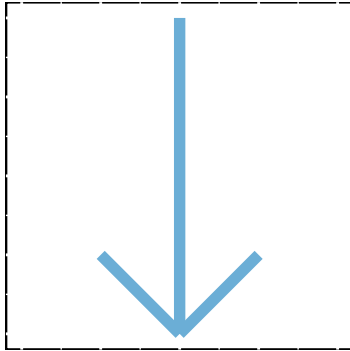
AB Test Arrow

The AB Arrow answers the question “Is B better or worse than A, or is it too close to call?” It can answer this question in five possible ways

1. B is better with 95% certainty -> Medium Blue Arrow Up
2. B is worse with 95% certainty -> Medium Blue Arrow Down
3. B is better with 75% certainty -> Light Blue Arrow Up
4. B is worse with 75% certainty -> Light Blue Arrow Down
5. B is too similar to A to decide -> Question Mark

I plan to make the AB Arrow available for any AB test you'd like to do by selecting a couple pieces of information

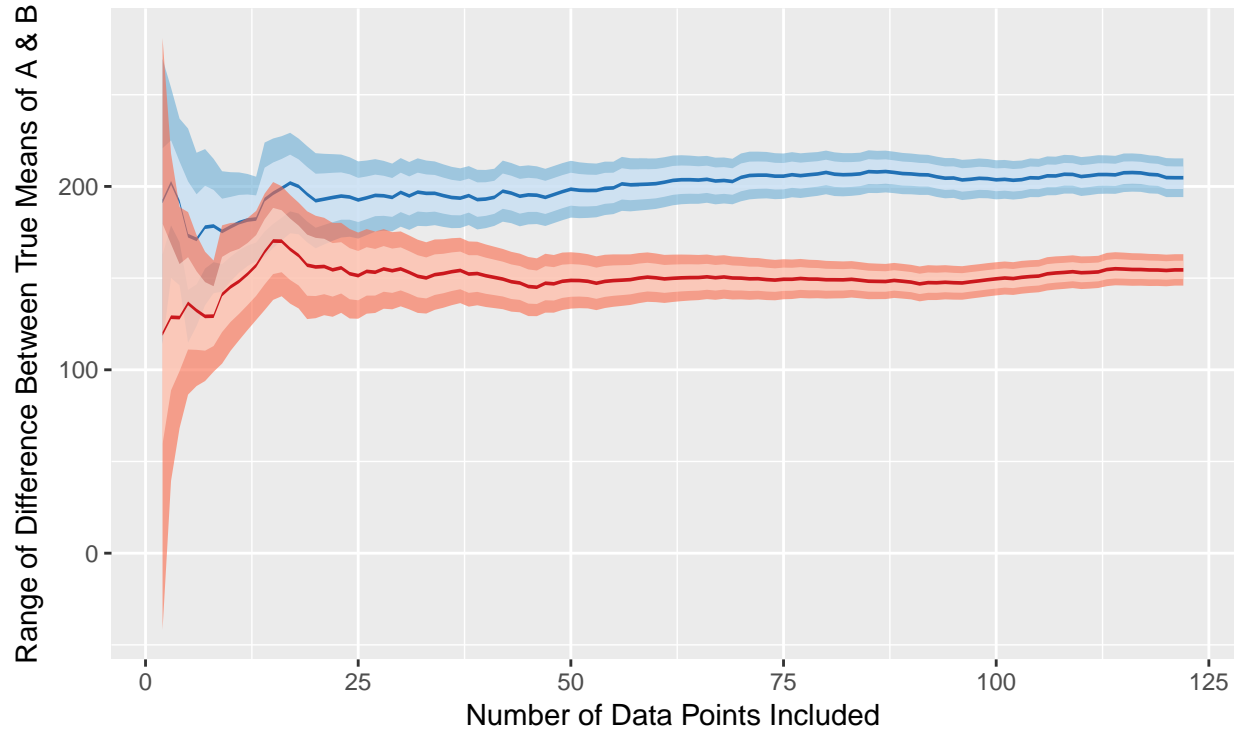
Here's the AB Arrow for Sellers Publishing June-Sep 2016 (A) vs. 2017 (B) (as we saw above, 2017 was significantly lower, so the medium blue down arrow makes sense)



Visualization of how the AB Test unfolds with each new day of data

AB Test Progress with 95% & 75% Confidence Intervals

Lines show average of A (blue) and B (red) and
ribbons show possible variation of that average at 75% and 95% confidence

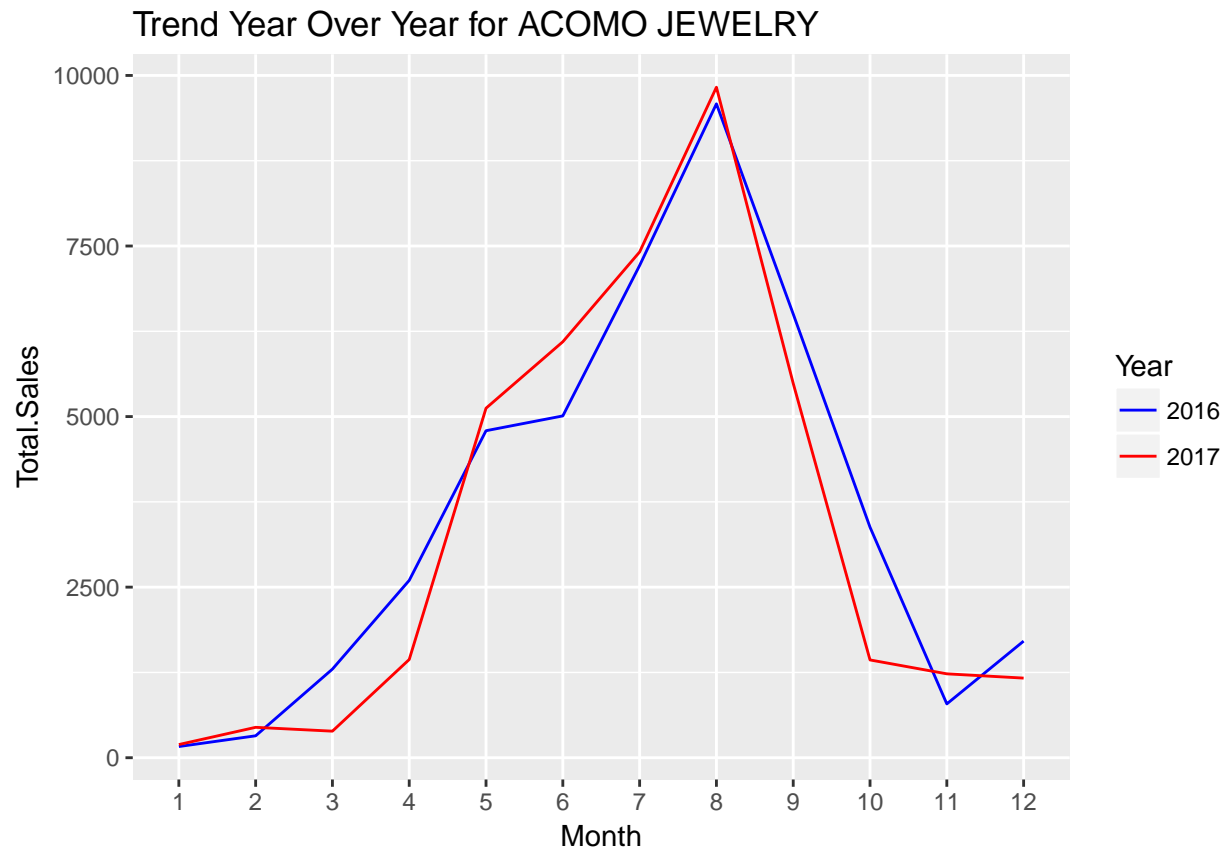


Acomo Jewelry

Demonstrating AB Test Visualizations

Scenario 2: Acomo Jewelry in June-Sept 2016 vs. 2017

The daily sales rate of Acomo Jewelry through the Summer was very similar in 2017 to its rate in 2016. This will mean the AB test will be unable to pick a winner, but this can be a useful finding as well.



AB Arrow for Acoma Jewelry June-Sept 2016 (A) vs. 2017 (B)

The AB Arrow is a question mark because the means of the daily sales rate are too similar to pick a winner statistically. This fits with the trend shown in the previous line graph where the lines followed each other in the Summer months very closely.



Visualization of how the AB Test unfolded with each new day for Acoma Jewelry

The range of possible difference between the means narrowed over time but even after 4 months of time the means weren't statistically significant.

AB Test Progress with 95% & 75% Confidence Intervals

Lines show average of A (blue) and B (red) and ribbons show possible variation of that average at 75% and 95% confidence

