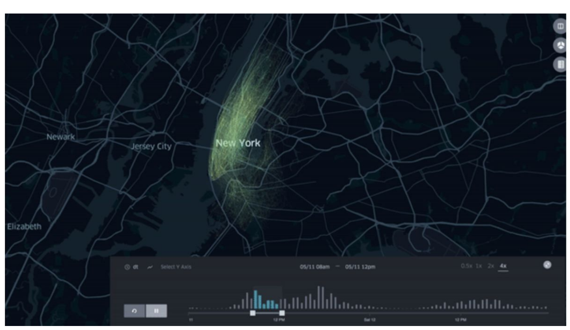
MSAN 631: Design and Analysis of Experiments

Bounce Rate Analysis using Google Analytics

Yue Lan, Bingyi Li, Guoqiang Liang, Beiming Liu, Zizhen Song

No discussion of transportation in New York would be complete without talking about one of the cheapest, easiest—and, increasingly, most popular—ways of getting around the city: biking. About 450,000 bike trips are taken every day in the five boroughs, with one in five of those trips being taken by someone who's commuting. So what are bikers' favorite route to ride in the city? And when is the most common time for biking? In order to clearly answer these questions, we created a dynamic visualization tool.

However, we come up with three different designs (Figure 1.) of the diagram and have troubles determining which one should be used on our website. We expect that users will be willing to interact with the webpage a little bit if the diagram is attractive enough. On the other hand, if they feel the diagram is boring, they will leave the webpage quickly without doing anything. So it will be useful for us to perform an A/B test on the bounce rate, which is the percentage of single-page sessions in which there was no interaction with the page.



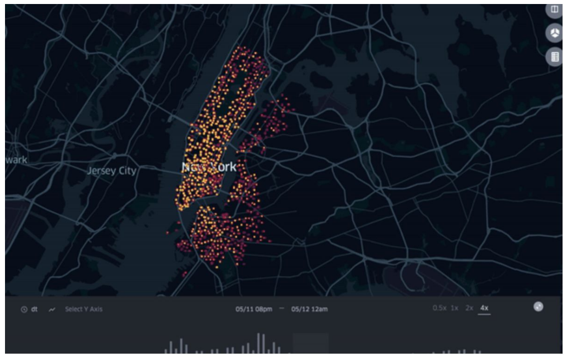


Figure 1.

So more formally, the optimization metric of our test is the bounce rate on the webpage, and the response variable is the number of bounce, which is calculated by the number of unique page view times the bounce rate. Our experimental factor has three levels, where each of those is a different visualization design. Our hypothesis is that webpage with different visualization diagrams will have different bounce rate, indicating one diagram might be more attractive than the other two.

We assume different groups have equal variance so that we can calculate effect size, which is defined as the difference of bounce rates divided by the standard deviation. For each pairwise comparison, we choose effect size of one, along with 0.01 significance level and 0.8 power. Based on these numbers, we calculate that the necessary sample size for our experiment is 21. For each pairwise comparison, our null hypothesis is that the two pages have the same bounce rate, and we expect to reject it. Here is how our data looks like.

	Unique Page View	Bounce Rate	Number of Bounce
V1	28	52.17%	15
V2	24	76.47%	18
V3	22	75.00%	17

And here is the result of our tests. One can see that for both tests, p-value is greater than 0.01, indicating that there is no significant difference in terms of the number of bounce for these three pages.

H0	HA	P-Value
V1 = V2	V1 != V2	0.1899
V1 = V3	V1 != V3	0.1509

So in conclusion, none of the design of visualization will give the webpage a different bounce rate, so we should feel free to use any of these three diagrams.

However, there is one potential problem with this conclusion. Although there is no significant difference between bounce rate, it might not suggest that these three visualization designs have the same level of attraction. For example, if a user plays with one diagram for 3 seconds and play with another for 30 seconds, there is definitely a preference, but bounce rate will not reflect this. So more ideally we should perform a test of average time on pages. But it is not feasible at this time, because with the experiment set up by Google Analytics, we do not have access to the sample variance. So a test on the bounce rate will be an acceptable substitute.