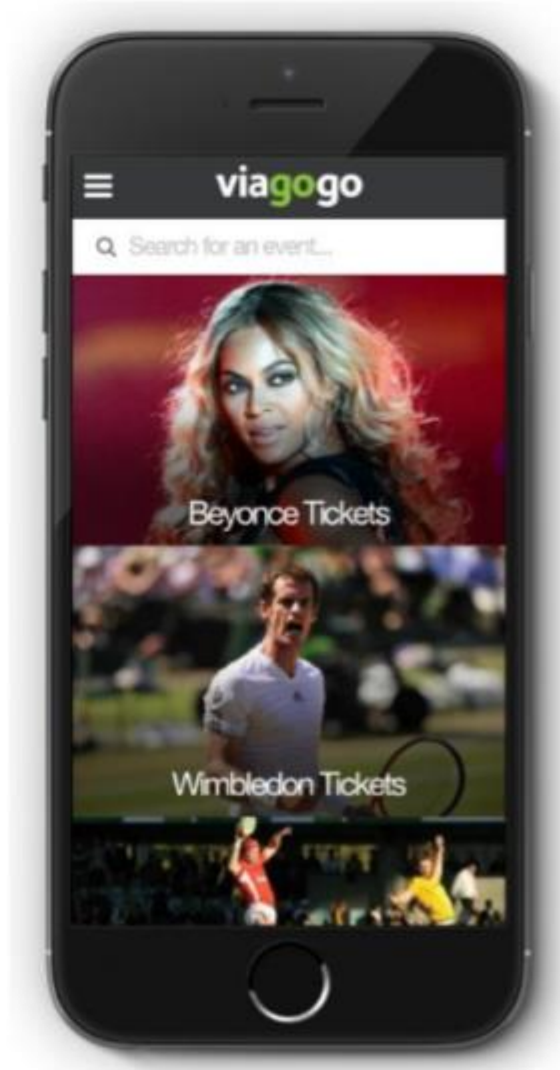


Viagogo Case Study Analysis

Prepared by:
Ujjwal Tamhankar
+1 (856)-952-6794
utamhank1@gmail.com
Feb16th, 2019

Contents

Executive Summary

Ticket sales are a critical part of Viagogo's business model. The ability to turn user website visits to our homepage into revenue is an important goal for the organization. In doing so, we utilize various tools and methods to determine what approaches are successful at increasing user engagement.

One of these is A/B testing. In this report we examine the data from our latest A/B test that was run from 10/10/2014 until 10/30/2014 which modified the mobile homepage and displayed the top 10 events nearest to a user's geographical location (hereby referred to as "the variant"). The default layout for the mobile homepage is to display the top 10 events based on popularity to every user regardless of that users location (hereby referred to as "the control).

In analyzing the data we observe two key metrics to judge our success, the conversion rate, which is a ratio describing the percentage of users that visit the page and make a purchase, and the bounce rate, which is the percentage of users who click away from the homepage after navigating on the page from an external site. Ideally, a change in website design should yield an increase in the conversion rate, and a decrease in the bounce rate to be deemed successful.

Key Findings:

- According to our metrics, the variant did not overall lead to a positive change in revenue or an increase in user engagement.
- Overall, the variant led to a 4.5% decrease in conversion rate and 4% increase in bounce rate when compared to the control.
- The control had an overall avg conversion rate of 5.5% and an overall avg bounce rate of 39% while the variant had an overall avg. conversion rate of 5.3% and an overall avg. bounce rate of 41%.
- Upon the performance of certain statistical tests, it was determined that these changes were not statistically significant.
- After further examining the slices of data categorized by the channel through which the user landed on the site (through an affiliate, direct, email link, paid search link, SEO, or social media link) as well as whether the user was new or returning, we were able to gather several insights as to where an A/B test like this might be most effective and/or places there is scope for additional investigation to be done.

		Channel						
		All Channels	Direct	Email	Paid Search	SEO	Social Media	Affiliate
All Users	conversion rate percent change (control -> variant)	-4.51%	-6.14%	-6.06%	-4.53%	-4.96%	0.38%	-1.05%
	bounce rate percent change (control -> variant)	4.00%	5.38%	4.05%	3.60%	3.61%	1.96%	2.34%
New Users	conversion rate percent change (control -> variant)	-2.99%	-5.46%	-6.00%	0.69%	-4.19%	0.34%	2.78%
	bounce rate percent change (control -> variant)	3.66%	4.57%	3.46%	4.28%	4.63%	-2.26%	2.04%
Returning Users	conversion rate percent change (control -> variant)	-6.35%	-7.07%	-5.22%	-10.2%	-5.94%	-0.80%	-4.77%
	bounce rate percent change (control -> variant)	4.38%	6.47%	4.75%	2.59%	1.27%	8.20%	2.92%

Background

This report describes the results and analysis of an A/B test performed on the Viagogo mobile homepage from 10/10/2014 to 10/30/2014 with the goal of increasing revenue through promoting user engagement. The change to the homepage to be tested was the 10 categories of events that a user is first displayed upon navigation to the home page from the most popular of the user base (in terms of ticket sales) to the 10 categories nearest to the users geographical location that are having events in that particular week.

In order to measure the success of the change (referred to as "the variant") against the original (referred to as "the control"), we examined the change in the webpage's conversion rate. The conversion rate is defined by the number of users that visit the homepage and subsequently make a ticket purchase divided by the total visitors to the homepage:

$$\text{Conversion Rate} = \frac{\#\ \textit{visitors to the home page that subsequently make a purchase}}{\#\ \textit{total visitors to the home page}}$$

Secondly, the goal of the test was to determine if a change in website design would yield a decrease in the bounce rate, defined as the number of visitors that "bounce" from the homepage after only viewing one page divided by the total number of visitors that that land directly on the homepage (from an external site):

$$\text{Bounce Rate} = \frac{\#\ \textit{visitors that bounce from the home page}}{\#\ \textit{total visitors that land on the home page}}$$

Thirdly, we delve into some supplementary statistical analysis and determine the overall success or failure of the variant over the control, as well as some key considerations for evaluating the results.

Lastly, we discuss recommendations for further potential improvements to the webpage, set criteria for future A/B tests, and discuss other metrics to judge success or failure of those tests.

Experimental Design

Hypothesis

Our hypothesis for this A/B test was that a change in the homepage design according to the variant will lead to a statistically significant change in user engagement by increasing the conversion rate and decreasing the bounce rate.

Methodology

This experiment was performed by showing half of the visitors to our site the control version of the homepage, and the other half of our users the variant. We then collected data cataloging certain metrics of the users including whether the user was new or returning, and what channel through which the user landed on our site (through an affiliate, direct search bar, email link, paid search link, SEO, or social media link).

Description of Collected Data

This test was run from 10/10/2014 until 10/30/2014. For each day, the following metrics were cataloged in the .csv that we analyzed:

- Date: Date of the of the user's visit.
- Channel: How the user arrived to the Viagogo site (from an affiliated website, directly, email, paid search, SEO, or social media link).
- User Type: Describes if the user that visited the page was a returning or new user.
- Land: Whether the user navigated to our homepage from another page on our site (0) or if they navigated directly to our homepage from another site (1).
- Bounce: Describes if the user navigated to another page on our website after landing on our homepage (0) or if they left our site altogether (1).
- Purchase: Describes if the user made a purchase (1) or if they did not (0)
- Visitors_Control: Describes the number of users shown the "control" webpage that met the criteria in the previous columns that visited the webpage for that sample.
- Visitors_Variant: Describes the number of users shown the "control" webpage that met the criteria in the previous columns that visited the webpage for that sample.x``

There were thus 60 samples taken per day of the study on users that satisfied relevant combinations of the 3 metrics (Land, Bounce and Purchase). A sample of the collected data is shown in the figure below:

Date	Channel	User Type	Land	Bounce	Purchase	Visitors_Control	Visitors_Variant
10/10/2014	Affiliate	Returning User	0	0	0	1211	1175
10/10/2014	Affiliate	Returning User	1	0	0	4076	4810
10/10/2014	Affiliate	Returning User	1	1	0	2766	3386
10/10/2014	Affiliate	Returning User	0	0	1	196	159
10/10/2014	Affiliate	Returning User	1	0	1	358	332
10/10/2014	Affiliate	New User	0	0	0	1589	1574
10/10/2014	Affiliate	New User	1	0	0	7165	6501
10/10/2014	Affiliate	New User	1	1	0	4709	4211
10/10/2014	Affiliate	New User	0	0	1	132	150
10/10/2014	Affiliate	New User	1	0	1	640	694
10/10/2014	Direct	Returning User	0	0	0	2336	2311
10/10/2014	Direct	Returning User	1	0	0	9099	8452

Data Analysis

Description of Data Structures extracted for analysis

In order to extract the relevant metrics, we wrote code in python (available on Github at https://github.com/utamhank1/viagogo_case_study) to restructure the data into a format that that would present the data that we were interested in (the conversion and bounce rates on a day by day basis, throughout the time period the A/B test was run). In addition, we parsed out a small table of the mean values for Bounce Rate, Conversion Rate, and their aggregate difference.

A sample of the results output tables is shown below:

	Date	Users Purchased Control	Users Purchased Variant	Users Landed Control	Users Bounced Control	Users Landed Variant	Users Bounced Variant	Users Visited Control	Users Visited Variant	Conversion Rate Control	Conversion Rate Variant	Bounce Rate Control	Bounce Rate Variant
0	10/10/2014	8164	8034	116714	43480	118595	45786	134823	136799	0.06055347	0.0587285	0.372535	0.38607
1	10/11/2014	6185	5948	128451	59294	132081	63668	146452	150276	0.04223227	0.03958051	0.461608	0.482038
2	10/12/2014	7989	7589	120823	47482	120355	49703	139861	139387	0.057121	0.05444554	0.392988	0.41297
3	10/13/2014	7794	7383	124564	49522	126101	53187	143092	144040	0.05446845	0.0512566	0.397563	0.421781
4	10/14/2014	6583	6291	125204	53472	128482	56601	144092	147529	0.04568609	0.04264246	0.427079	0.440536
5	10/15/2014	6401	6179	126881	57887	128842	60269	144917	146976	0.04417011	0.04204088	0.456231	0.467774
6	10/16/2014	6091	5943	135773	63856	138574	67398	153276	156195	0.03973877	0.03804859	0.470314	0.486368
7	10/17/2014	8344	8147	124010	45956	124898	47830	142540	143404	0.05853795	0.05681153	0.370583	0.382952

	Type	Conversion Rate Avg	Bounce Rate Avg
0	Control	0.055613329	0.3960114
1	Variant	0.053105799	0.4118704
2	Aggregate Difference	-0.00250753	0.015859

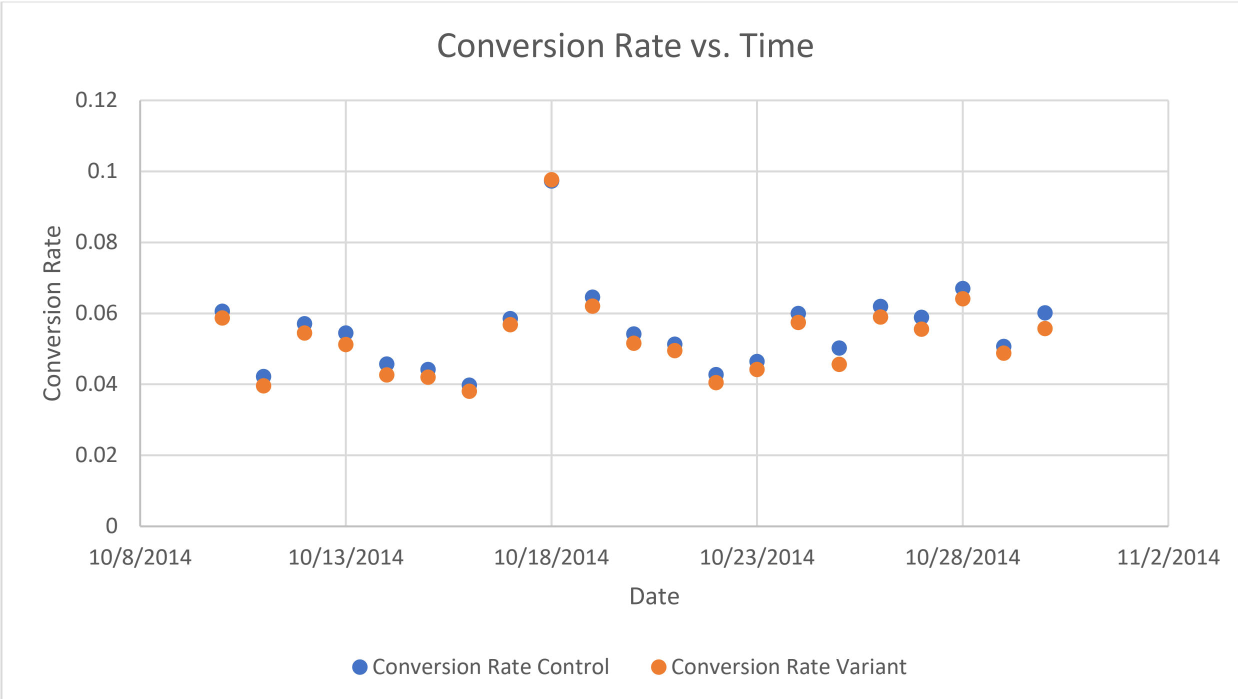
In our primary analysis, we look at the conversion and bounce rates for a combination of all users. In our subsequent secondary analysis, we stratify the data by looking at the Conversion and Bounce rates for each of the 20 other subgroups of users given by:

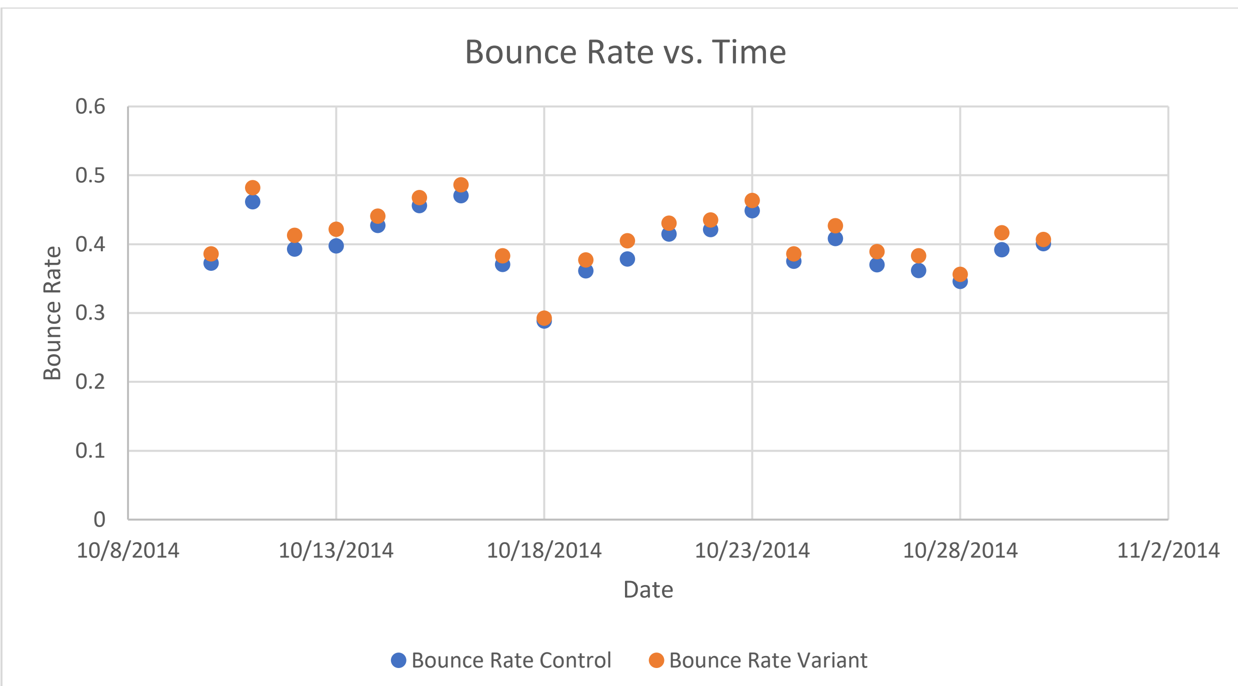
1. New users – All Channels
2. Returning Users – All Channels
3. All Users – Direct Channel
4. All Users – Email Channel
5. All Users – Paid Search Channel
6. All Users – SEO Channel
7. All Users – Social Media Channel
8. All Users – Affiliate Channel
9. New Users – Direct Channel
10. New Users – Email Channel
11. New Users – Paid Search Channel
12. New Users – SEO Channel
13. New Users – Social Media Channel
14. New Users – Affiliate Channel
15. Returning Users – Direct Channel
16. Returning Users – Email Channel
17. Returning Users – Paid Search Channel
18. Returning Users – SEO Channel
19. Returning Users – Social Media Channel
20. Returning Users – Affiliate Channel

Primary Analysis

Upon running code to extract the relevant criteria, the following graphs and table on the next page summarize the results of the experiment for all users on all channels:

Type	Conversion Rate Avg	Bounce Rate Avg
Control	0.055613329	0.396011432
Variant	0.053105799	0.411870393
Aggregate Difference	-0.00250753	0.015858961
Percent Change	-4.5%	4.0%





Here we see, based on just looking at the average change in Conversion Rates and Bounce Rates, that the variant resulted in an average *lower* conversion rate and a *higher* bounce rate for the redesigned homepage. Another trend we noticed was that there was an overall negative correlation between conversion rates and bounce rates. As conversion rates increased, bounce rates decreased and vice versa. This makes sense since if more people are buying tickets, one can infer an increase in user engagement in general and thus there will be a decrease in users navigating away from the page after visiting the site once. Further, it appears that conversion rates peaked (and bounce rates valleyed) during the middle of the month while remaining relatively stable during the rest of the month.

Independent samples t-test for significance

Looking at the data we collected and analyzed, it was important to determine if the resulting changes in conversion and bounce rates could be deemed statistically significant beyond a 95% threshold. Due to the following assumptions of there being two groups, the independent variable (homepage) being assigned randomly to two homogeneous groups, the sample sizes being equal, no extreme outliers, and the dependent variable being qualitative, we determined that the independent samples t-test was a good statistical measure to determine if the sample means differed in a significant way. Thus, we set our null hypothesis H_0: $\mu_A - \mu_B = 0$ and setting our alpha value at alpha = .05. We calculated the t statistics comparing the differences between the conversion and bounce rates (results available in primary_results_analysis.xlsx).

We determined that neither the means for the bounce and conversion rates differed statistically significantly, all assumptions holding true. Since a key assumption of the independent samples t-test was that the underlying data is normally distributed, we applied an additional test to determine if the samples for conversion rates, and bounce rates, were in fact normally distributed.

Kolmogorov – Smirnov Test for normality

The Kolmogorov-Smirnov test attempts to determine if two datasets differ significantly, in this case, we compare each of the variant and control conversion and bounce rates to the normal distribution and try to determine if there is a significant difference. A key advantage of the K-S test over other, more traditional tests such as Chi-squared is that the KS-test makes no assumption about the underlying distribution of the data.

Upon running the K-S test on the Bounce Rates and Conversion rates we determine that the key statistic “D” is less than the D critical for 21 samples at an alpha of .05, and therefore conclude that both the bounce rates and the conversion rates are normally distributed. This proving of a key underlying assumption of the t-test gives credibility to the fact that the results of the t-test are valid and that there is no statistically significant difference exists among the means of the control and variant.

Secondary Analysis

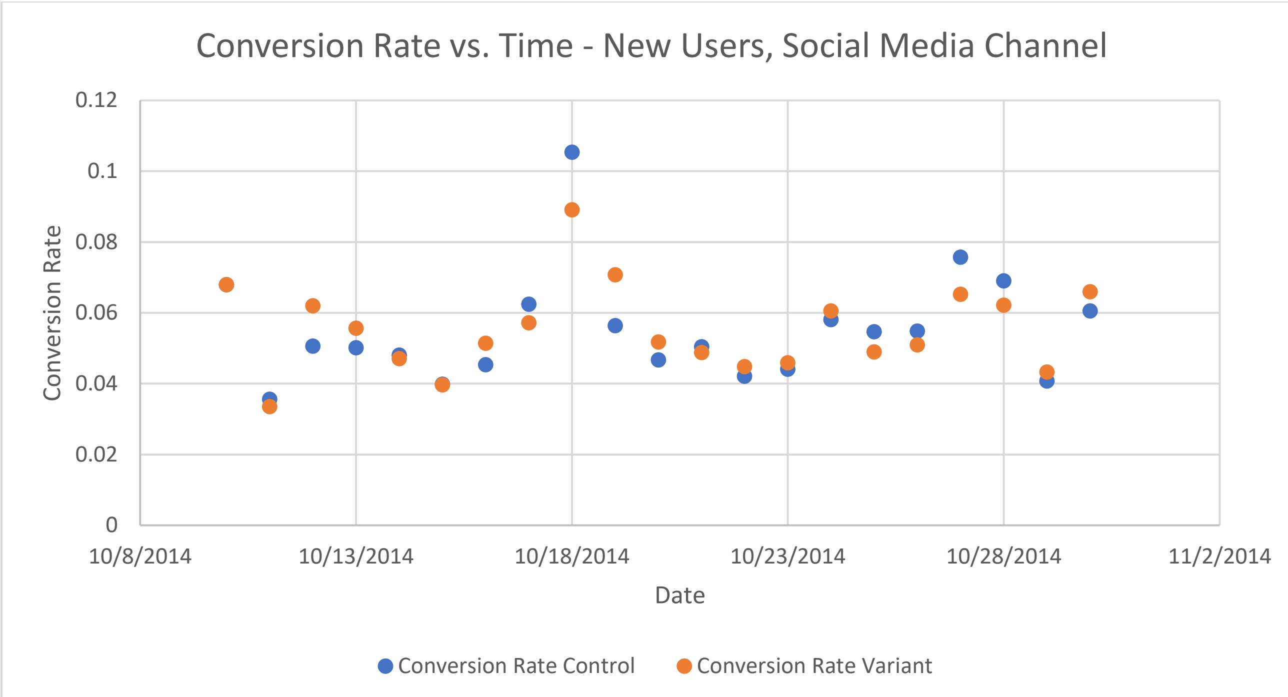
Although the changes to the conversion rate and bounce rate among the data gathered from the experiment did not show a strong discrepancy between the control and variant homepages, it was worth investigating whether there was a strong change in bounce and conversion rate amongst the 20 population subgroups (see section

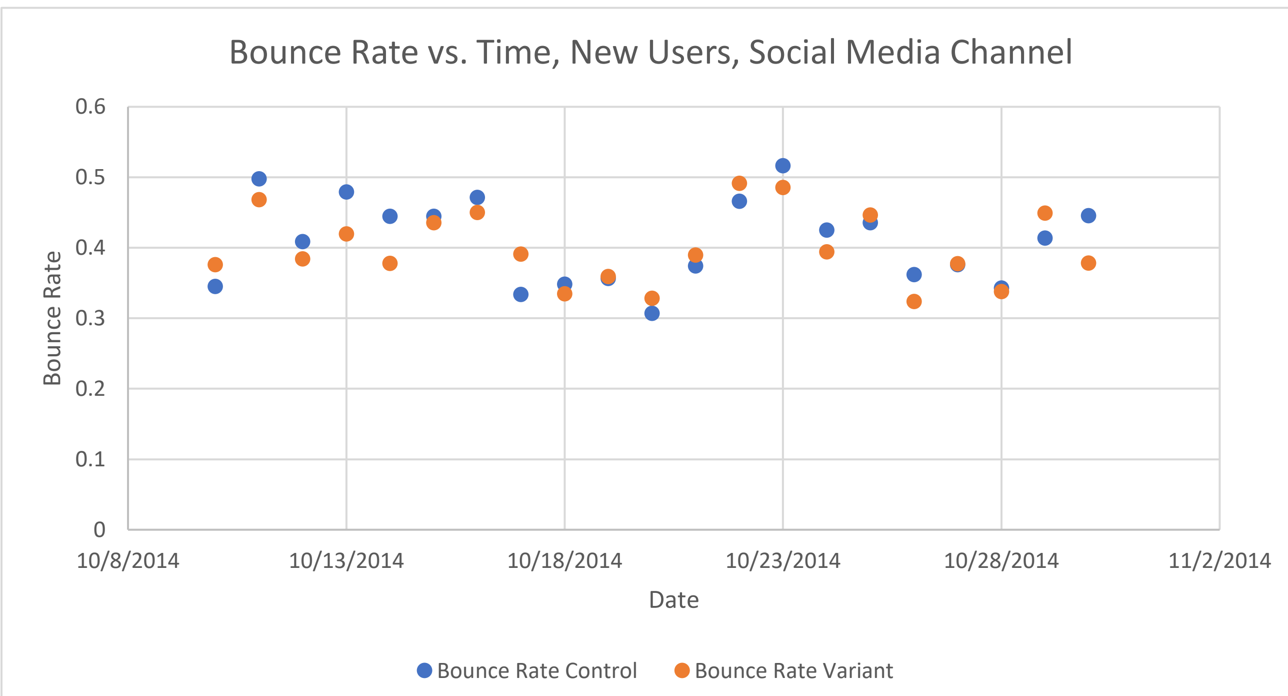
Description of Data Structures extracted for analysis) that could be extracted from the data. Thus, the original data extraction script was modified, and the table below summarizes the means of the conversion and bounce rates for each subgroup, as well as their respective percent differences between the control and variant groups:

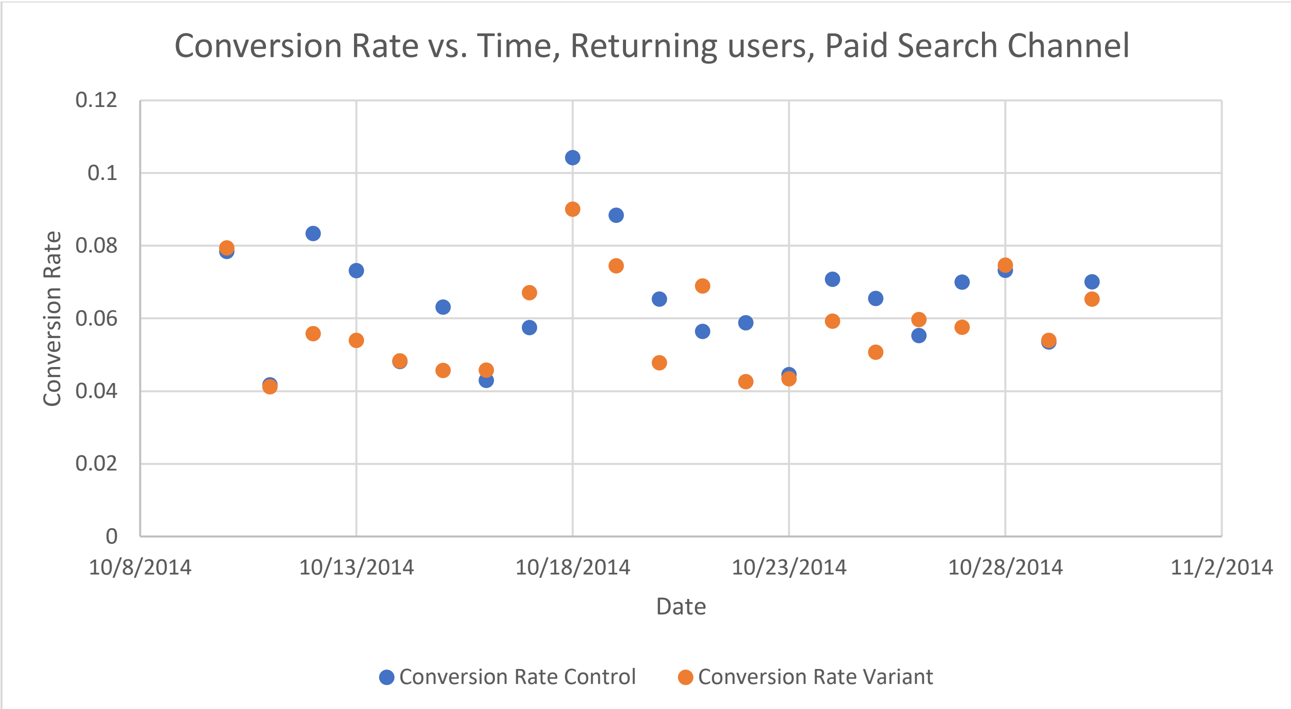
		Channel						
		All Channels	Direct	Email	Paid Search	SEO	Social Media	Affiliate
All Users	conversion rate percent change (control -> variant)	-4.51%	-6.14%	-6.06%	-4.53%	-4.96%	0.38%	-1.05%
	bounce rate percent change (control -> variant)	4.00%	5.38%	4.05%	3.60%	3.61%	1.96%	2.34%
New Users	conversion rate percent change (control -> variant)	-2.99%	-5.46%	-6.00%	0.69%	-4.19%	0.34%	2.78%
	bounce rate percent change (control -> variant)	3.66%	4.57%	3.46%	4.28%	4.63%	-2.26%	2.04%
Returning Users	conversion rate percent change (control -> variant)	-6.35%	-7.07%	-5.22%	-10.2%	-5.94%	-0.80%	-4.77%
	bounce rate percent change (control -> variant)	4.38%	6.47%	4.75%	2.59%	1.27%	8.20%	2.92%

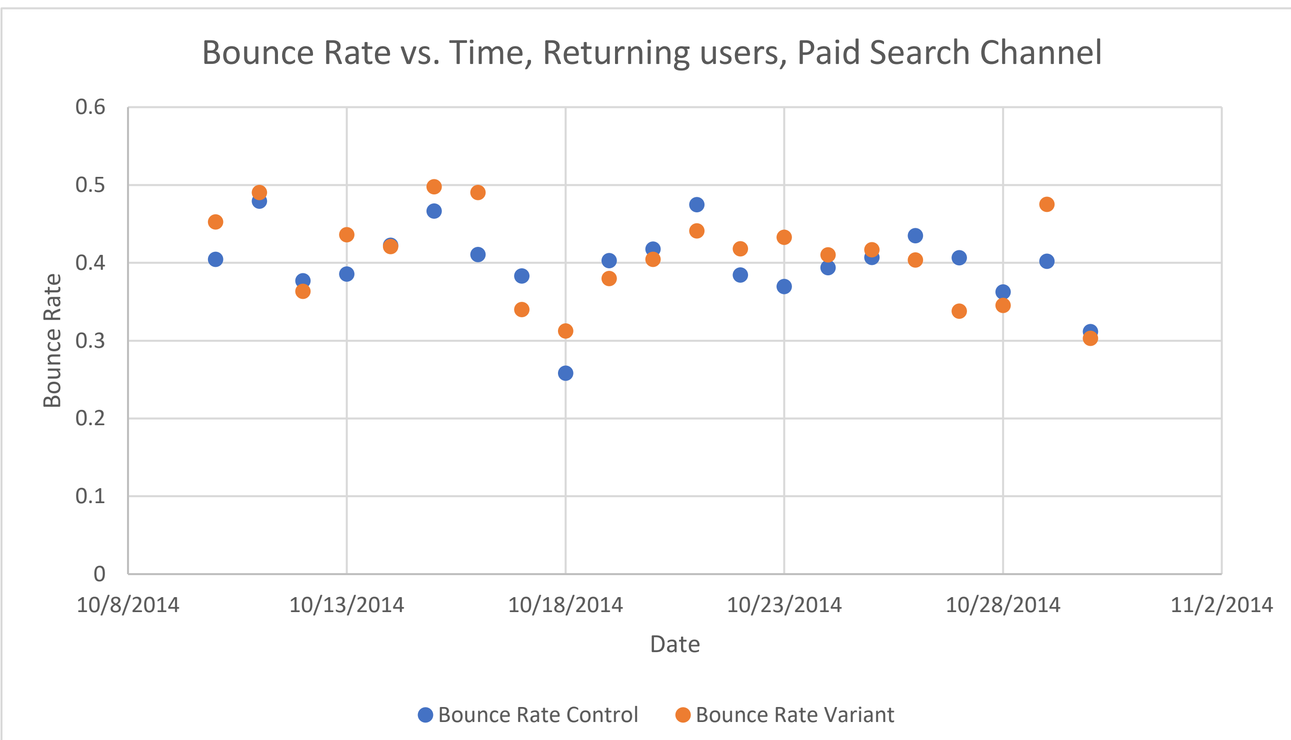
Here we have highlighted interesting findings amongst the New Users – Social Media channel subgroup and the Returning Users – Paid Search channel subgroup. The subgroup highlighted in green denotes a possible area where the variant homepage may have been effective in driving revenue and user engagement (indicated by a positive conversion rate difference between the variant and control, and a decrease in bounce rate between the variant and control). The subgroup highlighted in red indicates a group of users where the variant may have performed especially poorly as indicated by an over 10% decrease in conversion rate and a 2.5% increase in the bounce rate. Upon performing the t-test on both areas, we still saw that the results were not statistically significant for a $p>.05$ confidence threshold (data was still normally distributed according to Kolmogorov – Smirnov test, results of the two tests are available in the files secondary_results_land_paid_search_returning_users_analysis.xlsx and secondary_results_social_media_new_users_analysis.xlsx)

Day-by- Day graphs of the two categories of users are presented below:


Conversion Rate vs. Time - New Users, Social Media Channel
Conversion Rate
0.12
0.1
0.08
0.06
0.04
0.02
0
10/8/2014
10/13/2014
10/18/2014
10/23/2014
10/28/2014
11/2/2014
Date
Conversion Rate Control
Conversion Rate Variant


Bounce Rate vs. Time, New Users, Social Media Channel
Bounce Rate
0.6
0.5
0.4
0.3
0.2
0.1
0
10/8/2014
10/13/2014
10/18/2014
10/23/2014
10/28/2014
11/2/2014
Date
Bounce Rate Control
Bounce Rate Variant





Observing that the social media category was the only category where the variant was somewhat successful, this may be the most promising of the categories to explore with further testing.

It was also interesting to explore the raw means of the conversion and bounce rates for all the categories for both the variant and control groups, as shown below:

Control		Channel						
		All Channels	Direct	Email	Paid Search	SEO	Social Media	Affiliate
All Users	Conversion Rate	0.0556	0.0550	0.0592	0.0573	0.0535	0.0630	0.0555
	Bounce Rate	0.3960	0.3921	0.3894	0.4082	0.3954	0.4004	0.4001
New Users	Conversion Rate	0.0530	0.0537	0.0536	0.0522	0.0523	0.0552	0.0528
	Bounce Rate	0.3985	0.3934	0.4032	0.4152	0.3925	0.4092	0.4019
Returning Users	Conversion Rate	0.0596	0.0573	0.0678	0.0650	0.0559	0.0761	0.0599
	Bounce Rate	0.3927	0.3898	0.3705	0.3980	0.4021	0.3897	0.3972

Variant		Channel						
		All Channels	Direct	Email	Paid Search	SEO	Social Media	Affiliate
All Users	Conversion Rate	0.0531	0.0516	0.0556	0.0547	0.0509	0.0632	0.0549
	Bounce Rate	0.4119	0.4132	0.4052	0.4229	0.4096	0.4083	0.4095
New Users	Conversion Rate	0.0514	0.0508	0.0504	0.0526	0.0501	0.0554	0.0542
	Bounce Rate	0.4131	0.4114	0.4171	0.4329	0.4107	0.4000	0.4101
Returning Users	Conversion Rate	0.0558	0.0533	0.0642	0.0584	0.0526	0.0755	0.0570
	Bounce Rate	0.4099	0.4150	0.3881	0.4083	0.4072	0.4217	0.4088

Here we see certain groups of users where the tests performed relatively strongly in terms of either high conversion rates or low bounce rate (highlighted in green) or areas where they performed poorly, as highlighted in red. It seems as if, for both tests, users that were routed from social media links engaged the most and were more likely to purchase tickets.

Conclusion

Based on the data collected and analyzed from this experiment, it is certain that overall, the change did not result in a positive effect on user engagement and/or revenue. While the differences between the control and variant were not statistically significant, the overall tendency was that bounce rates increased, and conversion rates decreased among users that were shown the variant homepage as compared to the control.

On the question of whether to revert to the control or stick with the variant, there are several external considerations that must be considered when making this decision. Are there any major advantages to showing the variant to users that can be quantified and reported? Quantitative data such as load times, software development and maintenance speed, amount of time that a user stays on the homepage, and qualitative metrics such as popularity of the app amongst key influencers and artists, paid promotions from local artists to promote sales of their tickets, and user experience would be key factors to consider when making the decision. These extra sources of data may give insights that show that the control has other advantages that don’t currently manifest in the given data. Going off only the data we are given however, it would be advisable to revert to the control if the only goals are to keep as high a conversion rate and low a bounce rate as possible.

Recommendations for New Experiments

Based on the results of this experiment, we recommend five potential changes to explore with subsequent A/B testing:

1. On the homepage, list the top ten *most popular* categories that are having events nearest to the users geographical location for that week. This way, we combine aspects of the current control (only top ten most popular events) and the current variant (events nearest to the geographical location of the user). In order to judge success or failure of this approach, we could use the same metrics as the current experiment, conversion rate and bounce rate and a similar data collection format.

2. Display the top 10 most popular events that are happening soonest and indicate to the user what day ticket sales will end, regardless of location. This may create a sense of urgency and increase the conversion rate. In order to determine success of this experiment, it would be important to look at the conversion rate and bounce rate, but also look at how many people clicked on events that were expiring soon, which would indicate a level of interest.

3. Add an easy slider/radio buttons for users to switch between types of events (concerts, sports, theater, festivals) that would display the ten most popular events in the category the user selected without having to navigate away from the home page. In order to judge the success of this feature, it would be important to catalog the number of users that clicked/engaged with the slider/radio button, and given that they engaged with the feature, the amount of users that converted (made a purchase) and the number of users that bounced (navigated away from the site). Another variant of this would be to display the top ten most popular events nearest the users geographical location once they select a category.

4. Change the design of the homepage based on what channel the user arrived to the Viagogo homepage from. As described from this experiment, users that viewed the variant who arrived from a social media link had a slightly higher conversion rate and slightly lower bounce rate when compared to users who arrived from another channel. It may be worth investigating the profile of the average social media user who navigates to the Viagogo site and optimizing a webpage that caters to them. Similarly, it may be useful to do this for the users in all channels and see what each group responds best too. Again, the conversion and bounce rates might be useful to judge the success of these experiments. A further factor that may be considered is the amount of time the user actively spends on the homepage.

5. Show the user the top tickets for popular events that are the most discounted from the “face value”. A key advantage Viagogo has as a secondhand ticket

marketplace is that they do not have to sell tickets above a price minimum that firsthand ticket resellers do. Why not take advantage of this and show users the most discounted tickets that ticket resellers are offering? This would incentivize some friendly competition on the seller side (since sellers would know their tickets would be prioritized to users if they sell at a discount under the face value) and benefit both users and Viagogo through being able to buy discounted tickets and increased sales. Further, this may increase user engagement by forcing users to check the site frequently for tickets for events they are interested in showing up at a steep markdown (since it is likely guaranteed the most discounted tickets would sell the most rapidly).

In order to determine the success of this approach, it would be advantageous to collect data on conversion rates, bounce rates, the number of returning users to the website in a given random time period, the amount of time each user spends on the homepage, the average amount of time that the most mark downed tickets stay on the homepage before selling out (less time the better), and the average percentage markdown that results in the highest conversion rates (for seller feedback).

Prioritizing the Recommendations

Additional data that would be useful to prioritize implementation of these recommendations would be the categories that most users engaged with the most (concerts, sports, theater, festivals). An even spread among all four would indicate that the suggestion (3) might be a promising first step. The average price of the ticket that a user purchased might give some insight on the behavior of the average consumer (do consumers, on average, purchase cheaper tickets from the mobile homepage, and buy expensive ones from the desktop site?). If so, it may be useful to implement suggestion number (5) and only display sharply discounted tickets on the mobile page as opposed to expensive ones for events such as large festivals. Recommendations (1), and (2) could be quite easily implemented based on data that we already have. Recommendation (4) would take the most work as it would involve substantial market research to create user profiles for visitors to the homepage, and several redesigns of the page to suit the interests of every group. However, a very targeted, comprehensive marketing strategy such as the one outlined in (4) may be the most promising as it would have the goal of giving each user segment what they desire the most and would ideally result in higher ticket sales and revenue.