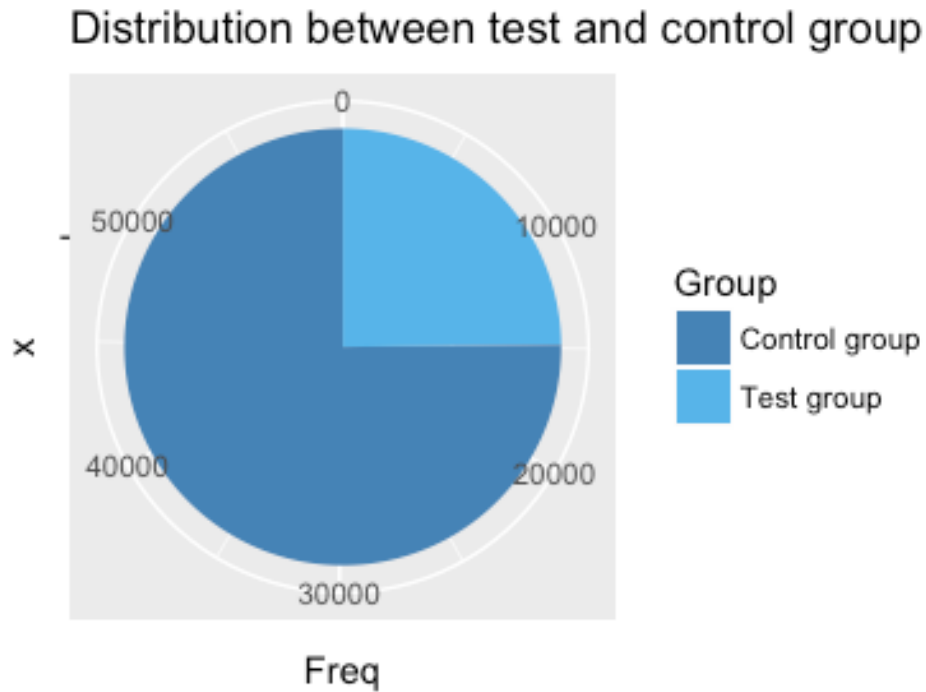


SeedboxTest

Olivier Froment

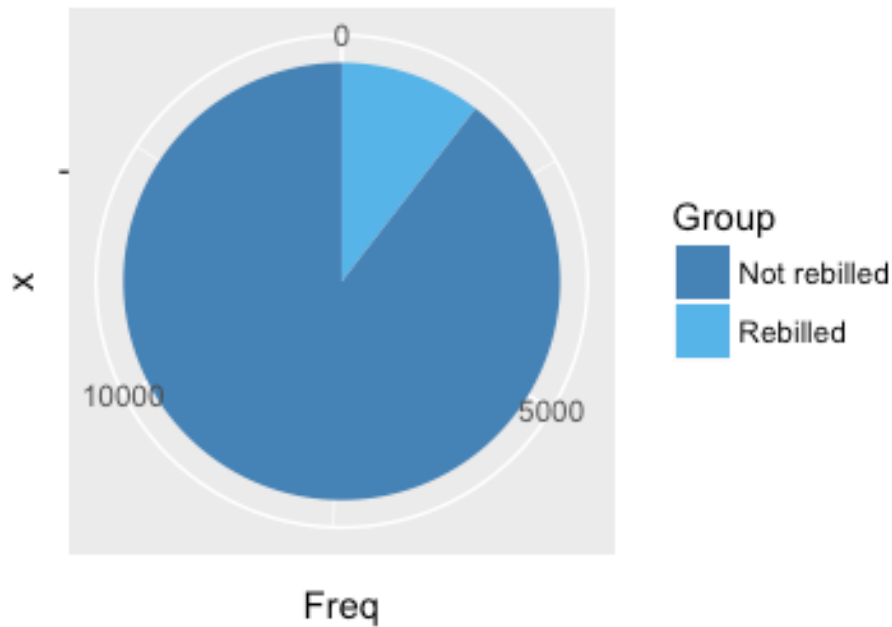
2018-07-15

Exploratory Analysis

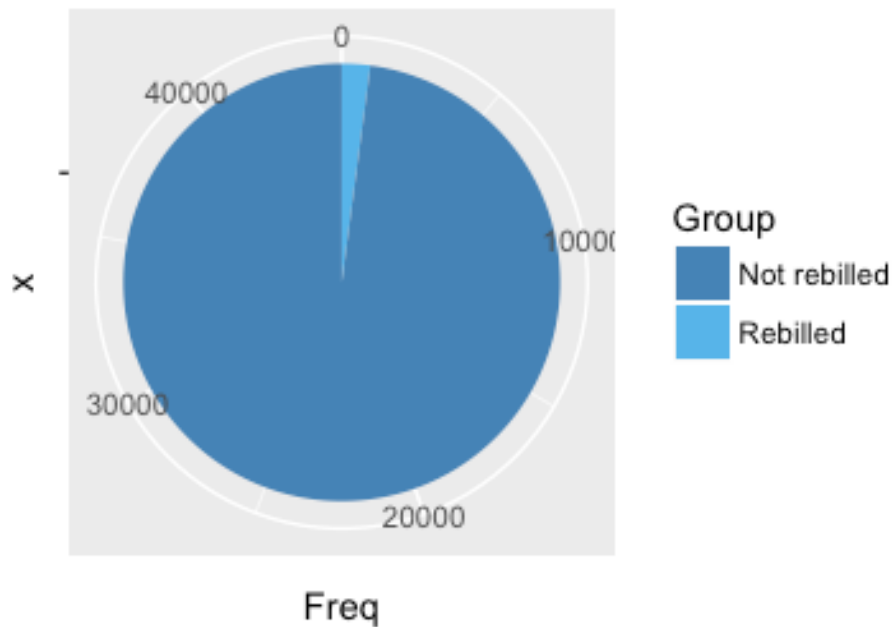


About a quarter of users were put into the test group. Therefore, a larger amount of data will come from users that were placed into the control group.

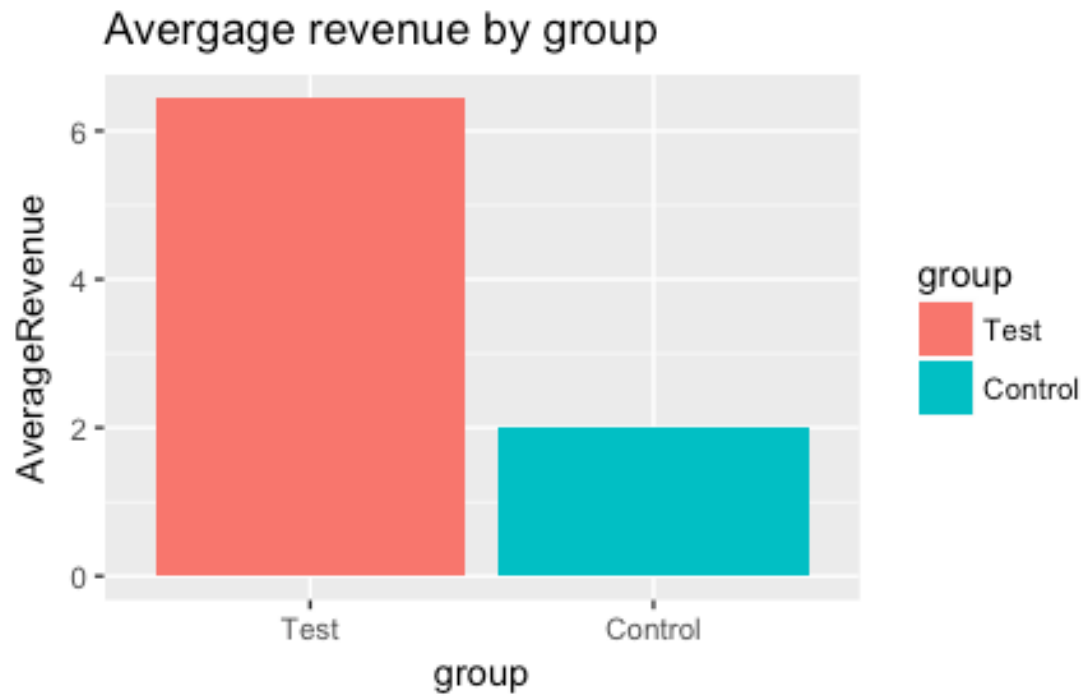
Distribution of rebilled users for test group



Distribution of rebilled users for control group

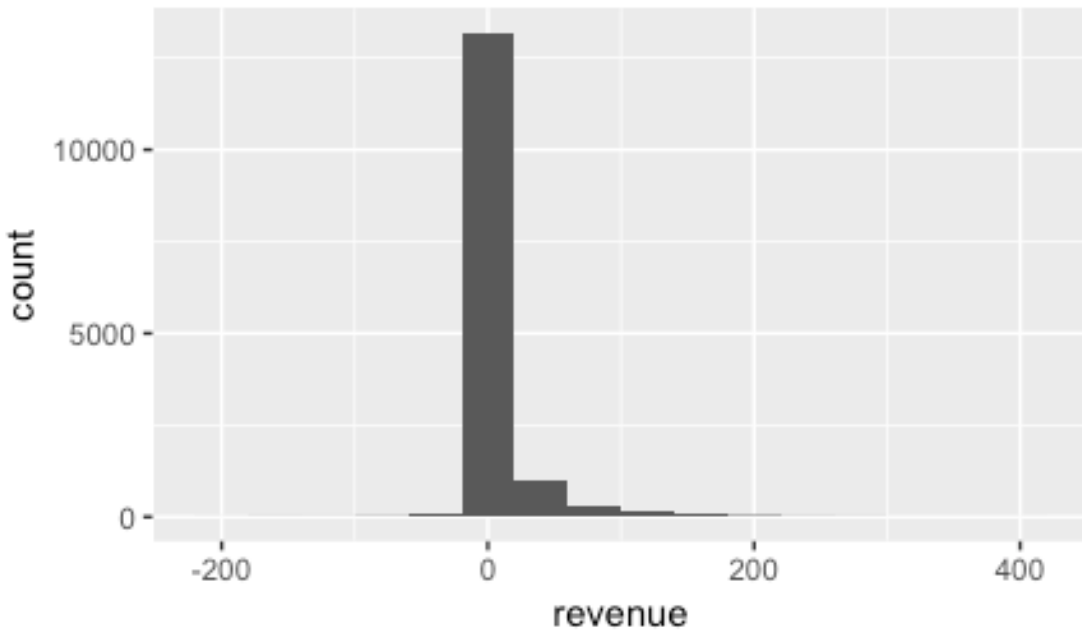


Within the sample, the proportion of of users that were rebilled is much higher when considering users from the test group rather than the control group. Empirically, a user that must call-in to cancel is much more likely to generate at least 1 additionnal rebill (about 5 times as likely).

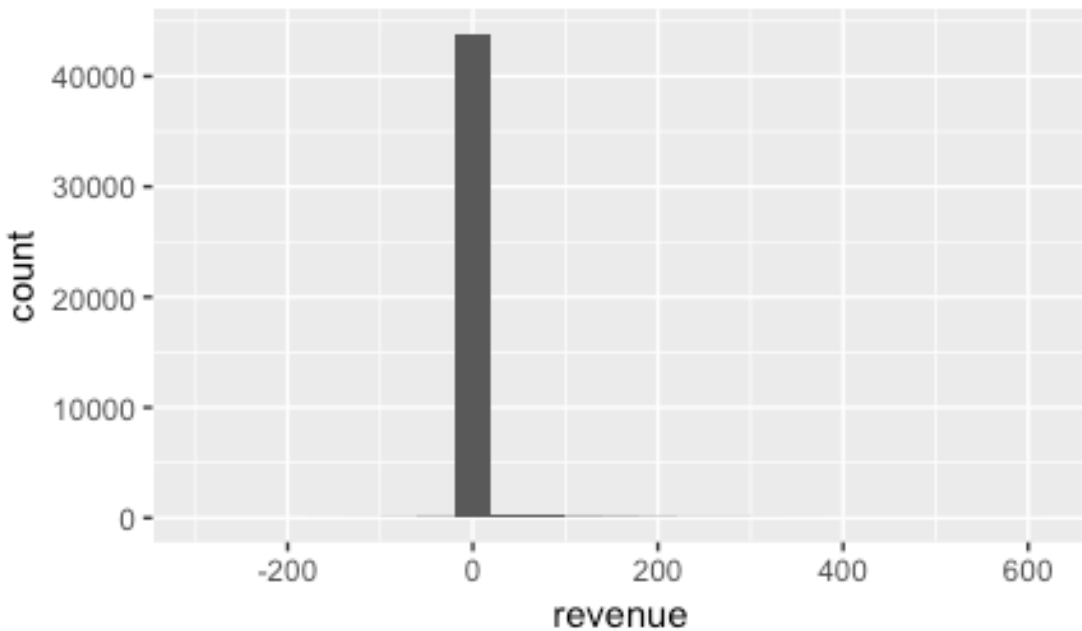


The average revenue generated by users in the test group is approximately three times higher than that generated by users in the control group. Given this fact, there are some good chances that a user that must call-in to cancel will generate more revenues.

Distribution of revenue users for test group



Distribution of revenue for control group



The empirical distribution of revenues generated by users from the test group has a larger tail than that of the empirical distribution of revenues generated by users from the control group. By the shape of these two distributions, it would seem like a user that must call-in to cancel is more likely to generate more revenues, although this must be verified with a formal test.

Methodology, Statistical Analysis and Conclusions

1. We wish to find a confidence interval for a proportion. The `prop.test()` function is most appropriate to obtain this result.

At a 95% confidence level, it is estimated that the probability distribution between the test group and the control group is as follows: **between 24.49% and 25.19% of users are randomized into the test group and conversely, between 74.81% and 75.51% of users are randomized into the control group.**

2. We want to know if a user that must call-in to cancel is more likely to generate at least one additional rebill. Equivalently, we are asking ourselves if the probability that a user from the test group generates at least one additional rebill is larger than that of a user from the control group. To answer this question, a one-tailed test for difference of proportions will be used.

The null hypothesis is: the probability that a user from the test group generates at least one additional rebill is smaller or equal than that of a user from the control group.

The alternate hypothesis is: the probability that a user from the test group generates at least one additional rebill is larger than that of a user from the control group.

After running the test, we obtain the following result: $p\text{-value} < 2.2e-16$. Therefore, we reject the null hypothesis in favor of the alternate hypothesis and **we conclude that, at a 95% confidence level, the probability that a user that must call-in to cancel generates at least one additional rebill is larger than that of a user that may cancel using the web form.**

Furthermore, the following 95% confidence interval was calculated: $[0.0796, 1]$. **This means that the probability that a user that must call-in to cancel generates at least one additional rebill is at least 7.96% larger than that of a user that may cancel using the web form.** This is congruent with the intuition that if it is harder for a user to cancel his subscription, it is more likely that he will stay subscribed.

3. Let X and Y be 2 random variables that describe the revenue generated by a user that must call-in to cancel and the revenue generated by a user that may cancel using the web form, respectively. We wish to know if a user that must call-in to cancel is more likely to generate more revenues than it is to generate less revenue. That is, $P(X > Y) > 0.5$, which is the alternate hypothesis for the Wilcoxon test. The null hypothesis is $P(X > Y) = 0.5$. That is, a user that must call-in to cancel is just as likely to generate more revenues than it is to generate less revenue.

We may also run a t-test to test whether or not the average revenue generated by a user that must call-in to cancel is greater than the revenue generated by a user that may cancel using the web form. In this case, the null hypothesis is that both averages are equal, while the null hypothesis is that the average revenue generated by the user that must call-in to cancel is greater.

The Wilcoxon test, being more robust, is more appropriate given the asymmetric nature of the revenue distribution. On the other hand, the t-test will allow for a better conclusion. Should the two tests contradict each other, the Wilcoxon test will be favored over the t-test.

After running the Wilcoxon test, we obtain the following result: $p\text{-value} < 2.2e-16$. We reject the null hypothesis and **we conclude that, at a 95% confidence level, a user that must call-in to cancel is more likely to generate more revenues.**

After running the t-test, we obtain the following result: $p\text{-value} < 2.2e-16$. The two tests are congruent. We reject the null hypothesis and **conclude that, at the 95% confidence level, the average revenue generated by a user that must call-in to cancel is larger.**

The following 95% confidence interval was also calculated: $[4.0498, \text{inf}]$. This means that **the average revenue generated by a user that must call-in to cancel is at least 4.0498\$ higher than that of a user that may cancel using the web form.**

These results are mostly due to the fact that 10.48% of users from the test group were rebilled while only 2.09% of users from the control group were rebilled.

4. We wish to know if user that must call-in more likely to produce a higher chargeback rate. But, this metric is undefined for most users, since most do not have more than one rebill. No notion of order can be attributed to the chargeback rate of users that have 0 rebills. On the other hand, it is still possible to get an idea about the chargeback rate within a group as a whole. Therefore, the chargeback rate will be computed for each group. **The results are as follow: 1.78% for the test group and 2.82% for the control group. We can see that, within the sample, the group of users that had to call-in to cancel had a much smaller chargeback rate as a whole.** This is congruent with our previous results. Since a larger percentage of users from the test group were rebilled, it only makes sense that the chargeback rate would be smaller.

While the result is still interesting, it is important to note that it does not allow us to make any inference about the population whatsoever, nor does it tell us anything about individual chargeback rates.