Engagement Test

Mitul Shah

8/8/2017

Loading the data

```
user_table <- read.csv("Engagement_Test/user_table.csv")
test_table <- read.csv("Engagement_Test/test_table.csv")
```

Checking Data Quality

```
## Are there any duplicates?
length(user_table$user_id) == length(unique(user_table$user_id)) ## looks good!
```

```
## [1] TRUE
```

```
length(test_table$user_id) == length(unique(test_table$user_id)) ## looks good!
```

```
## [1] TRUE
```

```
## Merge two datasets
data <- merge(user_table, test_table, by = "user_id")

## Converting the mode of date columns to date
data$signup_date <- as.Date(data$signup_date, format = "%Y-%m-%d")
```

```
## Warning in strptime(x, format, tz = "GMT"): unknown timezone 'zone/tz/
## 2018c.1.0/zoneinfo/America/Los_Angeles'
```

```
data$date <- as.Date(data$date, format = "%Y-%m-%d")
```

Overall Impact of the feature: Recommended Friends

```
## Doing a t-test overall
t.test(data$pages_visited[data$test == 0], data$pages_visited[data$test == 1])
```

```
##
##   Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0] and data$pages_visited[data$test == 1]
## t = 0.55711, df = 95835, p-value = 0.5775
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.02190997  0.03931178
## sample estimates:
## mean of x mean of y
##   4.608394  4.599693
```

The results show that the test was not significant overall. Moreover, the mean of the test group is also a bit less than the control group. Let's try to see whether this was actually true by investigating into different user segments.

Impact of the feature across different user segments

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(rpart)
```

```
## Checking to see whether the split was ~ 50/50 or not across all browser
table(data$browser, data$test) ## looks good!
```

```
##
##               0     1
##   Chrome  21453 21974
##   Firefox 10972 10786
##   IE      10906 10974
##   Opera    1109  1018
##   Safari   5406  5402
```

```
## Chrome Users
t.test(data$pages_visited[data$test == 0 & data$browser == "Chrome"], data$pages_visited[data$test == 1
```

```
##
##  Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0 & data$browser == "Chrome"] and data$pages_visited[data$test
## t = -3.3071, df = 40793, p-value = 0.0009434
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.12317542 -0.03150278
## sample estimates:
## mean of x mean of y
##  4.613341  4.690680
```

```
## Safari Users
t.test(data$pages_visited[data$test == 0 & data$browser == "Safari"], data$pages_visited[data$test == 1
```

```
##
##  Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0 & data$browser == "Safari"] and data$pages_visited[data$tes
## t = -1.1721, df = 10290, p-value = 0.2412
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.14472434  0.03641159
## sample estimates:
## mean of x mean of y
##  4.638180  4.692336
```

```
## Opera Users
t.test(data$pages_visited[data$test == 0 & data$browser == "Opera"], data$pages_visited[data$test == 1 &
```

```
##
##  Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0 & data$browser == "Opera"] and data$pages_visited[data$test
## t = 55.369, df = 1108, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  4.385326 4.707551
## sample estimates:
## mean of x mean of y
##  4.546438  0.000000
```

```
## Firebox Users
t.test(data$pages_visited[data$test == 0 & data$browser == "Firefox"], data$pages_visited[data$test == 1
```

```
##
##  Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0 & data$browser == "Firefox"] and data$pages_visited[data$tes
## t = -3.4405, df = 20716, p-value = 0.0005817
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.17909547 -0.04909487
## sample estimates:
## mean of x mean of y
##  4.600164  4.714259
```

```
## Internet Explorer Users
t.test(data$pages_visited[data$test == 0 & data$browser == "IE"], data$pages_visited[data$test == 1 & da
```

```
##
##  Welch Two Sample t-test
##
## data:  data$pages_visited[data$test == 0 & data$browser == "IE"] and data$pages_visited[data$test ==
## t = -2.6596, df = 20497, p-value = 0.00783
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1519985 -0.0230158
## sample estimates:
## mean of x mean of y
##  4.598478  4.685985
```

```
## Subsetting test and control users
data_test <- filter(data, test == 1)
data_control <- filter(data, test == 0)

library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, last
```
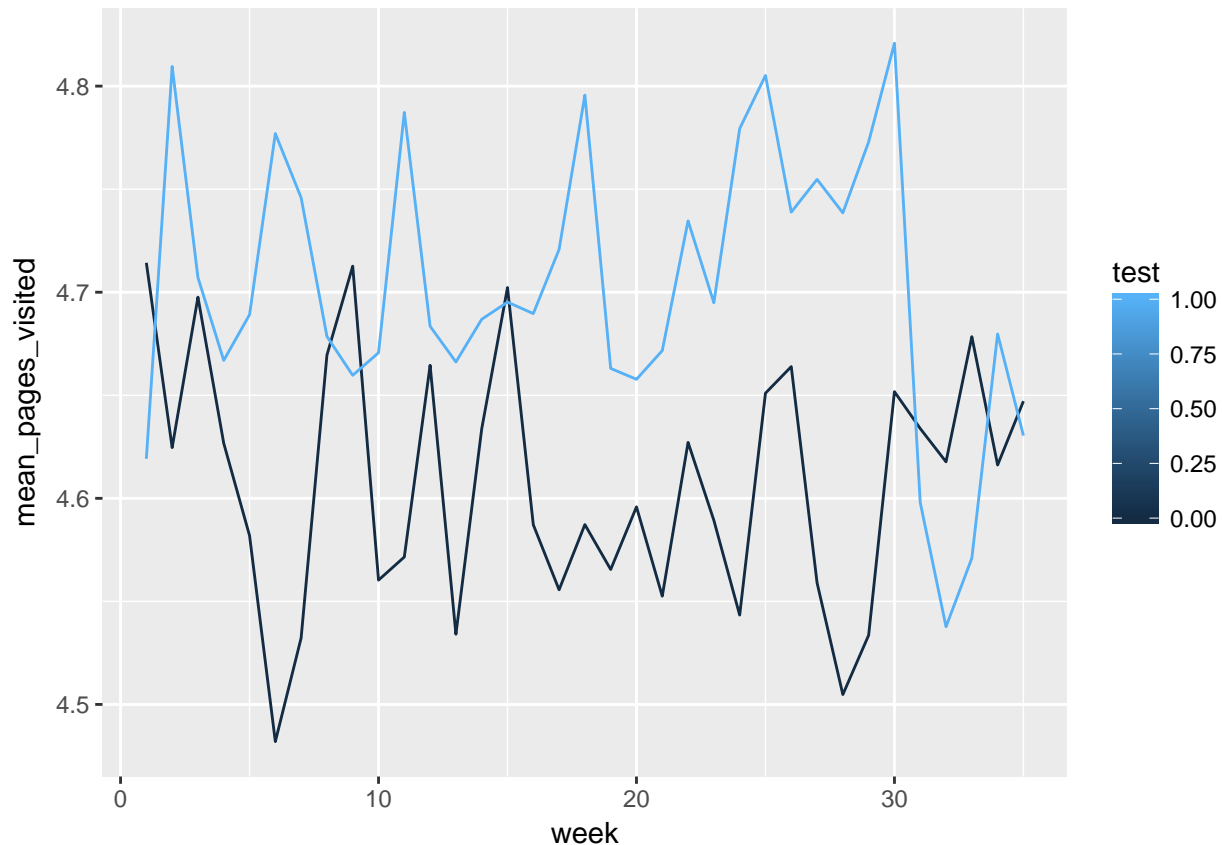
```
## Adding a week variable
data$week <- week(data$signup_date)

## Mean pages visited for control and test by week
dat <- subset(data, browser!="Opera") %>% group_by(week, test) %>% summarize(mean_pages_visited = mean(

library(ggplot2)

## Visualizing it!
ggplot(dat, aes(week, mean_pages_visited, group = test, color = test)) + geom_line()
```

```
## Building a tree
tree = rpart(test ~ ., data = data, control = rpart.control(minbucket = nrow(data)/100, maxdepth = 2))
tree
```

```
## n= 100000
##
## node), split, n, deviance, yval
##       * denotes terminal node
##
## 1) root 100000 24999.760 0.5015400
##   2) pages_visited< 1.5 13083  3039.202 0.3669648
##     4) pages_visited>=0.5 12065  2596.834 0.3135516 *
##     5) pages_visited< 0.5 1018     0.000 1.0000000 *
##   3) pages_visited>=1.5 86917 21687.960 0.5217967
##     6) pages_visited>=7.5 12269  2855.846 0.3687342 *
##     7) pages_visited< 7.5 74648 18497.430 0.5469537 *
```

Conclusion

We can clearly see that the conversion rate for the Opera Users in the test group was 0 and they responded differently. There is something wrong with this and it needs to be investigated further to find out the reason behind it.

If you look at the plot sign-up date vs pages visited for test and control, you will see that the test is performing proportionally better for old users than new users.

The reason of this is novelty effect. You put a new feature on the site and, just because it is new, old users click a lot on it just out of curiosity. But not because the new feature is actually better, it is simply curiosity. Then they don't click on it anymore as they get used to it. Obviously, new users don't show this behavior since for them everything is new. If you suspect your new feature might be affected by novelty effect, you should compare test and control only for users for which that was the first overall experience on the site (or at least very recent). If that's actually flat while overall test is winning (like here), it is a big warning sign of novelty effect issues. So there is no point in making the change on the site.