Krishna Rao
12-16-2019

PART I: Experiment Report

Introduction

This report analyzes the results of an experiment on [redacted]s mobile home page. The original home page displays 10 categories of the highest popularity based on sales. The variant proposal displays the 10 categories of events happening closest to the users' location. In the experiment, a sample of data was collected on two conditions- the original (Control) and the variant versions of the home page. Conversion and bounce rate are the metrics of interest. Conversion measures the proportion of customers that make a purchase, and bounce measures the proportion that immediately leave the website. A successful variant of the mobile home page would result in a higher conversion rate and lower bounce rate than the control.

Exploratory Data Analysis

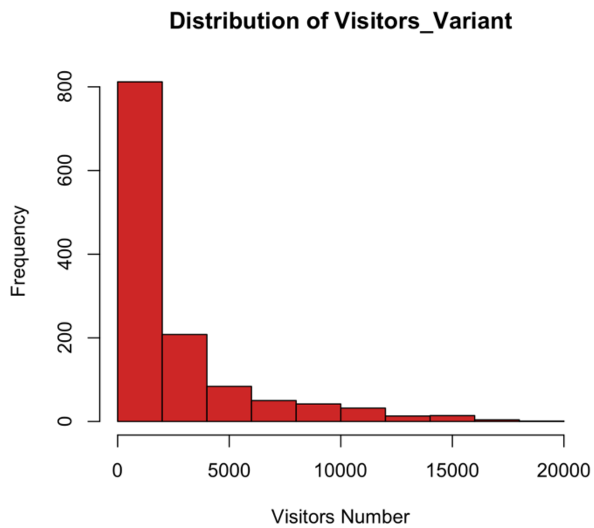
User behavior was tabulated for the following eight fields: 'Date', 'Channel', 'User Type', 'Land', 'Bounce', 'Purchase', 'Visitors Control' and 'Visitors Variant'. 'Date' spans 21 days starting on 10-10-2014.

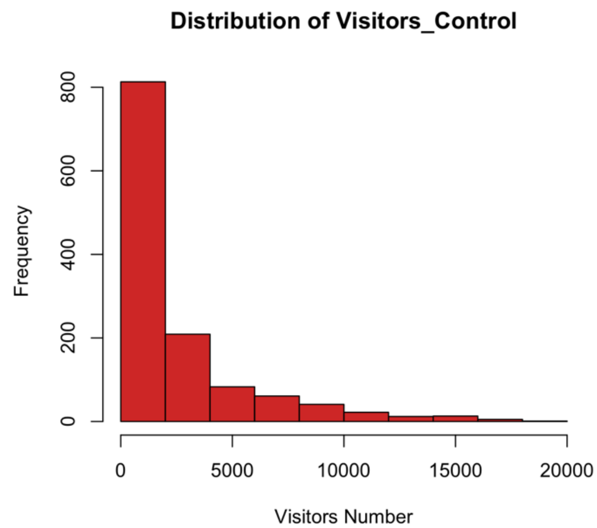
'Channel' and 'User Type' are categorical fields, with six and two unique values respectively. 'Channel' refers to how the user arrived on the website, and has values of 'Affiliate', 'Direct', 'Email', 'Paid Search', 'SEO', and 'Social Media'. 'User Type' has values 'Returning User' and 'New User'.

'Land', 'Bounce', 'Purchase' are indicator columns, which have value 1 or 0, indicating whether the event occurred or did not occur. 'Land' means that the user arrived directly on the home page. 'Bounce' means that the user left the website after landing. 'Purchase' indicates that the user made a purchase.

'Visitors Control' and 'Visitors Variant' are numerical columns- each containing the number of users that correspond to the unique set of conditions specified by the other columns- for control and variant groups respectively.

The data sample contains 1260 observations. The distribution of the two numeric columns, 'Visitors Control' and 'Visitors Variant' are highly skewed to the right, showing that there are a large number of observations for low values, and fewer observations for larger values in both fields. Normality is not a concern at this point, since the hypothesis tests are conducted on the conversion and bounce rate metrics.





Conversion and bounce rate are the main metrics that are analyzed in this report. These metrics are defined as follows:

Figure 1:

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

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

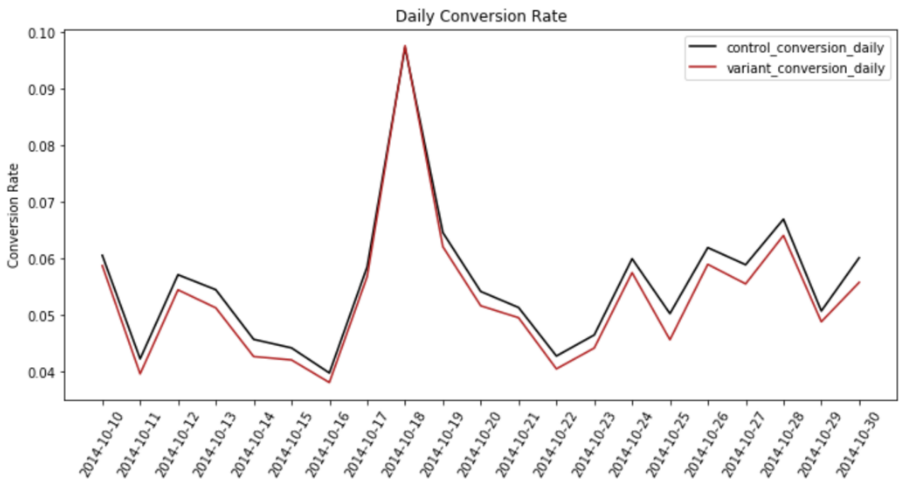
First, the conversion and bounce rates for the control and variant are calculated for the entire dataset. The results are tabulated below:

Table 1: Total Metrics

	Control	Variant	Aggregate Difference
Conversion	0.0556	0.0531	0.0025
Bounce	0.3966	0.4127	-0.016

From Table 1, we see that the total conversion rate from the control is higher than the conversion rate of the variant. The aggregate difference is therefore positive. The total bounce rate is lower for the control than the variant, and the aggregate difference is therefore negative. From this preliminary analysis, it appears that the control is outperforming the variant for both metrics.
Next, the daily metrics are examined. The experiment takes place over a 21-day period, from 10-10-2014 to 10-30-2014. Figure 2a shows the daily metrics for conversion rate, and Figure 2b shows the daily metrics for bounce rate.

Figure 2: Daily Metrics for Conversion and Bounce Rates
a.



b.

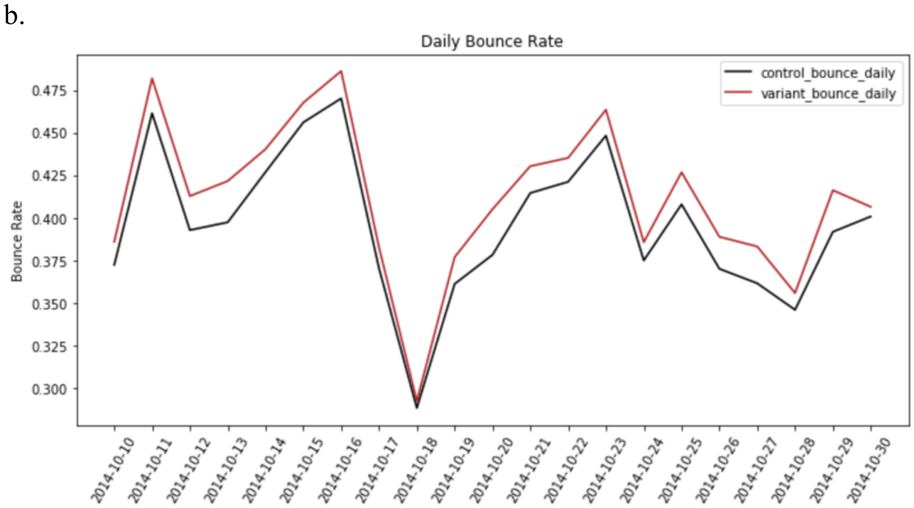


Figure 2a shows that the conversion rate of the control is higher than the variant on most days. The bounce rate for the control is lower than the variant on most days in the same period (Figure 2b). This follows the same trend as the total metrics, and suggests that the control is outperforming the variant. A clear trend in the data is a sharp increase in mean conversion rate around the middle of the month, 10-17, and a corresponding sharp decline in mean bounce rate around the same date. It is unclear without further information what may be causing this.

Statistical hypothesis tests (AB tests) are necessary to assess whether the control and variant differ significantly for the conversion and bounce rates. First, conversion rate is analyzed. A one-tailed, two-sample t-test performed to assess the validity of the following hypothesis:

Figure 3: Conversion Rate Hypothesis Test

$$H_0: \mu_{Control} - \mu_{Variant} = 0$$
$$H_1: \mu_{Control} - \mu_{Variant} < 0$$

The null hypothesis (H_0) states there is no difference between the mean value of the conversion rate for the control and the variant. The alternate hypothesis states that the mean value of conversion rate for the variant is greater than that of the control. Note that the mean values of the daily metrics are equivalent to the total metrics.
The resulting test statistic is t = 0.646, with p-value of 0.739. There is insufficient evidence to reject the null hypothesis at the $\alpha = 0.05$ significance level. The variant's conversion rate is not significantly different from that of the control. The interpretation is that the variant likely *does not* significantly increase conversion rate.

Figure 4: Bounce Rate Hypothesis Test

H_0: $\mu_{Control} - \mu_{Variant} = 0$
H_1: $\mu_{Control} - \mu_{Variant} > 0$

A similar test is performed for the bounce rate. The null hypothesis states that there is no difference between the mean value of the bounce rate for the control and the variant. The alternate hypothesis states that the mean value of bounce rate for the control is greater than that of the variant. If the bounce rate of the variant is lower, this is considered a successful outcome.
The resulting test statistic is t = -1.165, with a p-value of 0.875. There is insufficient evidence to reject the null hypothesis at the $\alpha = 0.05$ significance level. This means that the variant likely *does not* significantly decrease bounce rate.

Discussion and Supplementary Analysis

The results above consider the entire dataset. It is important to consider the results of different categories as well. First, the 'Channel' variable is examined. Recall that 'Channel' refers to how the user arrived at the website, and it comprises six different categories. Metrics are calculated for each of the different categories and summarized below.

Table 2: Data Split by Channel p-values

Channel	**Conversion p-value**	**Bounce p-value**
Affiliate	0.567	0.739
Direct	0.783	0.917
Email	0.764	0.893
Paid Search	0.743	0.858
SEO	0.755	0.812
Social Media	0.476	0.716

The data is split according to the channel, and the original hypotheses tests (see Figures 3 and 4) are conducted on each of the new data sets. As evident from the p-value columns, there are no significant results (p-value < 0.05) where the conversion or bounce rate of the variant outperformed the control.

Table 3: Data Split by User Type p-values

User Type	**Conversion p-value**	**Bounce p-value**
Returning User	0.788	0.882
New User	0.675	0.832

The data is split according to the User Type, and the original hypotheses tests are conducted. Similar to the Channel-split data, there are no significant results from splitting the data according to User Type.

Conclusion

Based on the results of the statistical tests, *there is not* strong evidence to choose the Variant over the Control version of the mobile home page. The hypothesis test showed that there is likely not a significant difference in the mean conversion or bounce rate across the entire dataset. Splitting the dataset and testing each of the six Channels and each of the two User Types also did not show statistically significant differences for the metrics. The alpha level used was $\alpha = 0.05$, meaning that the results are true with a 95% confidence level. As a result of these findings, the Variant should not be deployed.

For this situation, there is not any additional information that needs to be considered, since the change is fairly straightforward. It would be interesting for further study to see how often users pick categories in the top-10 list instead of searching for events manually. If users mostly searched for events, it might explain the negligible change by changing the top-10 list.

PART II: Next Experiment

The goal of this section is to design a new experiment to run on [redacted]s mobile home page. The current home page displays 10 categories of the highest popularity based on sales. Below I list five potential improvements I would make to the current home page. With each suggested change, I provide a brief discussion of the additional data I need, and how I would measure success or failure.

1. **I would use colored frames surrounding a smaller picture for categories. The current large-format pictures are aesthetically**

I would simply conduct an AB test on the original compared to this version, and measure the conversion and bounce metrics. Variant data should be recorded to test this potential improvement.

2. **Increase the size of the font of the category titles. This makes it easier to read and may increase chance of a click.**

Again, a simple AB test should be sufficient to assess this change. The variant version of the home page would have titles with larger fonts. Variant data should be recorded.

3. **Resize and reformat the categories so that they can all be displayed on a single page.**

This suggestion is based off the hypothesis that pages at the bottom of the top-10 list are less frequently purchased as a result of less optimal placement. For testing this hypothesis, it would help to have data on purchase by category. A test could compare the purchase proportions of the original top-ten categories compared to a version with randomly shuffled top-ten categories. If the proportions purchased by category differ significantly, this would indicate that placement of the category affects purchase.

4. **Place countdown clocks on the categories. This gives people an incentive to purchase now rather than later.**

If an event in a category is approaching, a banner announcing this may increase last-minute purchases. This can be tested with an AB test, and variant data should be recorded.

5. **Use a recommendation engine based off individual User's behavior. This may result in better targeted advertisement and higher conversion.**

With the rise big data, it is important to make use of collected data to improve user experience. A recommendation engine can use the user's previous selections to suggest categories the user is most likely to purchase. I would test this using an AB test comparing original and variant experimental data.

Success and Failure would be measured by using statistical hypothesis testing to compare the mean conversion and mean bounce rates of the variant with the original.

[redacted]_Case_Study

Krishna Rao

12/15/2019

```r
#note, I created a csv from 'Spring 2018 - Product Case Data.xlsx
dat <- read.csv('product_case_data.csv')
```

EDA

```r
colnames(dat)
```

```
## [1] "Date"             "Channel"          "User.Type"
## [4] "Land"             "Bounce"           "Purchase"
## [7] "Visitors_Control" "Visitors_Variant"
```

```r
head(dat, 5)
```

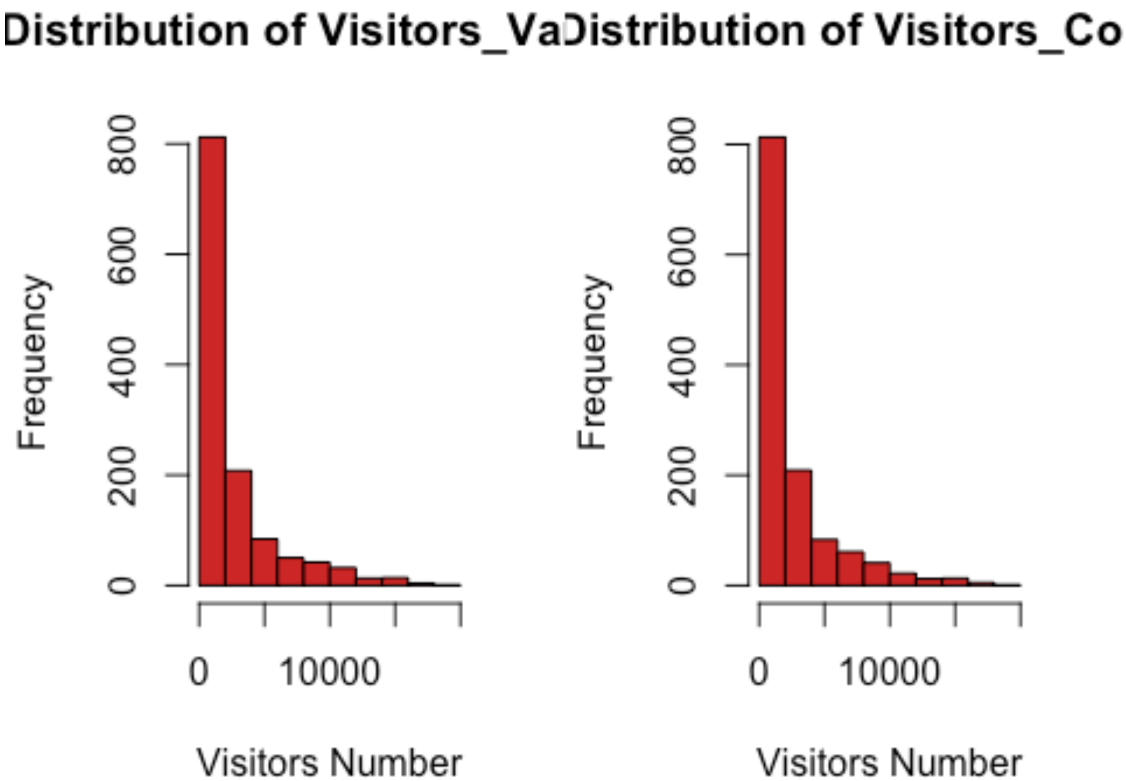
```
##       Date   Channel      User.Type Land Bounce Purchase Visitors_Control
## 1 10/10/14 Affiliate Returning User    0      0        0             1211
## 2 10/10/14 Affiliate Returning User    1      0        0             4076
## 3 10/10/14 Affiliate Returning User    1      1        0             2766
## 4 10/10/14 Affiliate Returning User    0      0        1              196
## 5 10/10/14 Affiliate Returning User    1      0        1              358
##   Visitors_Variant
## 1             1175
## 2             4810
## 3             3386
## 4              159
## 5              332
```

```r
str(dat)
```

```
## 'data.frame':    1260 obs. of  8 variables:
##  $ Date            : Factor w/ 21 levels "10/10/14","10/11/14",..: 1 1 1 1
1 1 1 1 1 1 ...
##  $ Channel         : Factor w/ 6 levels "Affiliate","Direct",..: 1 1 1 1 1
1 1 1 1 1 ...
##  $ User.Type       : Factor w/ 2 levels "New User","Returning User": 2 2 2
2 2 1 1 1 1 1 ...
##  $ Land            : int  0 1 1 0 1 0 1 1 0 1 ...
##  $ Bounce          : int  0 0 1 0 0 0 0 1 0 0 ...
##  $ Purchase        : int  0 0 0 1 1 0 0 0 1 1 ...
##  $ Visitors_Control: int  1211 4076 2766 196 358 1589 7165 4709 132 640
...
##  $ Visitors_Variant: int  1175 4810 3386 159 332 1574 6501 4211 150 694
...
```

EDA plot

```
par(mfrow=c(1,2))
hist(dat$Visitors_Variant, main = 'Distribution of Visitors_Variant', xlab=
'Visitors Number', ylab = 'Frequency', col = 'firebrick3')
hist(dat$Visitors_Control, main = 'Distribution of Visitors_Control', xlab=
'Visitors Number', ylab = 'Frequency', col = 'firebrick3')
```



function to calculate metrics

```
metrics_calculator <- function(df, control_or_variant, conversion_or_bounce){

  #CONTROL conversion rate:
  control_conversion_num <- df$Visitors_Control[which(df$Purchase == 1)]
  control_conversion_denom <- df$Visitors_Control
  control_conversion_rate <-
sum(control_conversion_num)/sum(control_conversion_denom)

  #VARIANT converison rate:
  variant_conversion_num <-df$Visitors_Variant[which(df$Purchase == 1)]
  variant_conversion_denom <- df$Visitors_Variant
  variant_conversion_rate <-
sum(variant_conversion_num)/sum(variant_conversion_denom)
```

```r
  #CONTROL bounce rate:
  control_bounce_num <- df$Visitors_Control[which(df$Bounce == 1 & df$Land ==
1)]
  control_bounce_denom <-df$Visitors_Control[which(df$Land == 1)]
  control_bounce_rate <- sum(control_bounce_num)/sum(control_bounce_denom)

  #VARIANT bounce rate:
  variant_bounce_num <- df$Visitors_Variant[which(df$Bounce == 1 & df$Land ==
1)]
  variant_bounce_denom <-df$Visitors_Variant[which(df$Land == 1)]
  variant_bounce_rate <-sum(variant_bounce_num)/sum(variant_bounce_denom)

  if (control_or_variant == 'control' & (conversion_or_bounce ==
'conversion')) {
     return (control_conversion_rate)
  }
  if (control_or_variant == 'variant' & (conversion_or_bounce ==
'conversion')) {
     return (variant_conversion_rate)
  }
  if (control_or_variant == 'control' & (conversion_or_bounce == 'bounce')) {
     return (control_bounce_rate)
  }
  if (control_or_variant == 'variant' & (conversion_or_bounce == 'bounce')) {
     return (variant_bounce_rate)
  }
}

print(metrics_calculator(dat, 'control', 'conversion'))

## [1] 0.0555822

print(metrics_calculator(dat, 'variant', 'conversion'))

## [1] 0.05305115

print(metrics_calculator(dat, 'control', 'bounce'))

## [1] 0.3966413

print(metrics_calculator(dat, 'control', 'bounce'))

## [1] 0.3966413
```

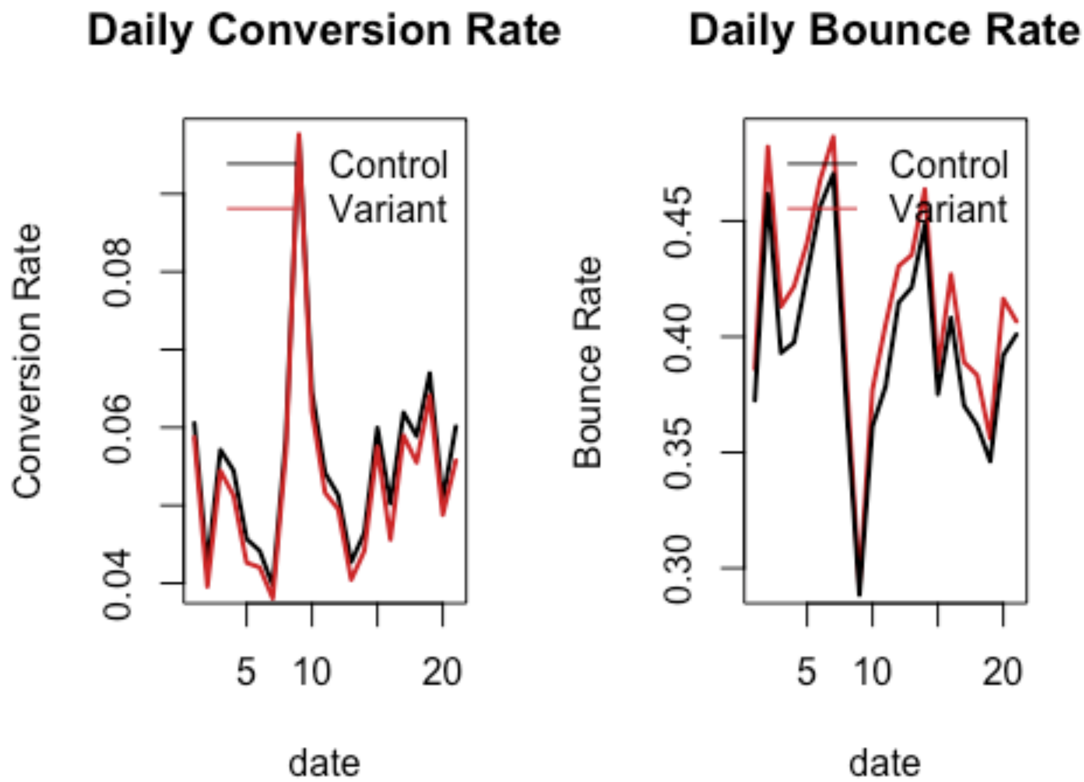
getting daily values

```r
unique_dates_list <- unique(dat$Date)
N = length(unique_dates_list)

daily_control_conversion = rep(NA, N)
daily_variant_conversion = rep(NA, N)
```

```
daily_control_bounce = rep(NA, N)
daily_variant_bounce = rep(NA, N)


for (i in 1:N){
  date = toString(unique_dates_list[i])
  temp_df = dat[which(dat$Date == date), ]

  daily_control_conversion[i] = metrics_calculator(temp_df, 'control',
'conversion')
  daily_variant_conversion[i] = metrics_calculator(temp_df, 'variant',
'conversion')
  daily_control_bounce[i] = metrics_calculator(temp_df, 'control', 'bounce')
  daily_variant_bounce[i] = metrics_calculator(temp_df, 'variant', 'bounce')

}
```

df holding daily values

```
daily_df<- data.frame( 'dates' = unique_dates_list,
                 'daily_control_conversion' = daily_control_conversion,
                'daily_variant_conversion' = daily_variant_conversion,
                'daily_control_bounce' = daily_control_bounce,
                'daily_variant_bounce' = daily_variant_bounce)

par(mfrow = c(1, 2))
plot.default(daily_df$dates, daily_df$daily_control_conversion, type = 'l',
lwd = 2, col = 'black', main = 'Daily Conversion Rate', xlab = 'date', ylab =
'Conversion Rate')
lines(daily_df$dates, daily_df$daily_variant_conversion, type = 'l', lwd = 2,
col = 'firebrick3')
legend('topright',  legend=c("Control", "Variant"),
       col=c("black", "firebrick3"), lty=1, bty = "n")

plot.default(daily_df$dates, daily_df$daily_variant_bounce, type = 'l', lwd =
2, col = 'firebrick3', main = 'Daily Bounce Rate', xlab = 'date', ylab =
'Bounce Rate')
lines(daily_df$dates, daily_df$daily_control_bounce, type = 'l', lwd = 2, col
= 'black')
legend('topright',  legend=c("Control", "Variant"),
       col=c("black", "firebrick3"), lty=1, bty = "n")
```



T-test for conversion

```
t.test(daily_df$daily_control_conversion, daily_df$daily_variant_conversion,
alternative= 'less')

##
##  Welch Two Sample t-test
##
## data:  daily_df$daily_control_conversion and
daily_df$daily_variant_conversion
## t = 0.64584, df = 39.954, p-value = 0.739
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##         -Inf 0.009045375
## sample estimates:
##  mean of x  mean of y
## 0.05561333 0.05310580
```

T-test for bounce

```
t.test(daily_df$daily_control_bounce, daily_df$daily_variant_bounce,
alternative = 'greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  daily_df$daily_control_bounce and daily_df$daily_variant_bounce
## t = -1.1648, df = 39.936, p-value = 0.8745
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.03878625         Inf
## sample estimates:
## mean of x mean of y
## 0.3960114 0.4118704
```

Conducting t-tests for each of the Channels and User Type:

```
#convert earlier analysis into a function


customize_df <- function(df){

  unique_dates_list <- unique(df$Date)
  N = length(unique_dates_list)

  daily_control_conversion = rep(NA, N)
  daily_variant_conversion = rep(NA, N)
  daily_control_bounce = rep(NA, N)
  daily_variant_bounce = rep(NA, N)

  for (i in 1:N){
    date = toString(unique_dates_list[i])
    temp_df = df[which(df$Date == date), ]

    daily_control_conversion[i] = metrics_calculator(temp_df, 'control',
'conversion')
    daily_variant_conversion[i] = metrics_calculator(temp_df, 'variant',
'conversion')
    daily_control_bounce[i] = metrics_calculator(temp_df, 'control',
'bounce')
    daily_variant_bounce[i] = metrics_calculator(temp_df, 'variant',
'bounce')

  }
    daily_df<- data.frame( 'dates' = unique_dates_list,
                 'daily_control_conversion' = daily_control_conversion,
                'daily_variant_conversion' = daily_variant_conversion,
                'daily_control_bounce' = daily_control_bounce,
                'daily_variant_bounce' = daily_variant_bounce)

  conversion_pval = t.test(daily_df$daily_control_conversion,
daily_df$daily_variant_conversion, alternative= 'less')$p.value
```

```r
  bounce_pval = t.test(daily_df$daily_control_bounce,
daily_df$daily_variant_bounce, alternative = 'greater')$p.value

  return (c(conversion_pval, bounce_pval))
}
```

Channel

```r
unique_channel_list <- unique(dat$Channel)

for (channel in unique_channel_list) {
  print(toString(channel))
  print(customize_df(dat[which(dat$Channel == toString(channel)), ]))
}
```

```
## [1] "Affiliate"
## [1] 0.5673098 0.7394567
## [1] "Direct"
## [1] 0.7832885 0.9167125
## [1] "Email"
## [1] 0.7635217 0.8931854
## [1] "Paid Search"
## [1] 0.7431651 0.8581569
## [1] "SEO"
## [1] 0.7554400 0.8117665
## [1] "Social Media"
## [1] 0.4764297 0.7164641
```

User Type

```r
unique_user_type_list <- unique(dat$User.Type)

for (user_type in unique_user_type_list) {
  print(toString(user_type))
  print(customize_df(dat[which(dat$User.Type == toString(user_type)), ]))
}
```

```
## [1] "Returning User"
## [1] 0.7880284 0.8820692
## [1] "New User"
## [1] 0.6752223 0.8324742
```