PromotionalCampaign_Analysis

Shivani Sharma

4/8/2020

Objective: Measure the effectivess of Promotional Campaign

Glimpse of the Dataset:

```
options(repos = c(CRAN = "http://cran.rstudio.com"))
#Importing Needed packages
install.packages("dplyr")
```

```
## Installing package into 'C:/Users/shiva/Documents/R/win-library/3.6'
## (as 'lib' is unspecified)
```

```
## package 'dplyr' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'dplyr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\shiva\Documents\R\win-library\3.6\00LOCK\dplyr\libs\x64\dplyr.dll to C:
## \Users\shiva\Documents\R\win-library\3.6\dplyr\libs\x64\dplyr.dll: Permission
## denied
```

```
## Warning: restored 'dplyr'
```

```
##
## The downloaded binary packages are in
##  C:\Users\shiva\AppData\Local\Temp\RtmpILFJEz\downloaded_packages
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
##
## 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
```

```
install.packages("tidyr")
```

```
## Installing package into 'C:/Users/shiva/Documents/R/win-library/3.6'
## (as 'lib' is unspecified)
```

```
## package 'tidyr' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'tidyr'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying C:
## \Users\shiva\Documents\R\win-library\3.6\00LOCK\tidyr\libs\x64\tidyr.dll to C:
## \Users\shiva\Documents\R\win-library\3.6\tidyr\libs\x64\tidyr.dll: Permission
## denied
```

```
## Warning: restored 'tidyr'
```

```
##
## The downloaded binary packages are in
##  C:\Users\shiva\AppData\Local\Temp\RtmpILFJEz\downloaded_packages
```

```
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 3.6.3
```

```
#Loading Data
data <- read.csv("GameFun.csv")
head(data)
```

```
##      id test purchase  site impressions income gender gamer
## 1  1956    0        0 site1           0    100      1     0
## 2 45821    1        0 site1          20     70      1     0
## 3 59690    1        0 site1          22    100      1     0
## 4 18851    0        0 site1          13     90      1     0
## 5 60647    1        0 site1          12     60      1     0
## 6 30908    1        0 site1           0     40      1     0
```

```
control<-data%>%select(test,purchase,site,impressions,gender,gamer,income)%>%filter(test==0)
treatment<-data%>%select(test,purchase,site,impressions,gender,gamer,income)%>%filter(test==1)
```

Checking whether the test and control groups are probabilistically equivalent:

```
#averages, HO: Income Means of control and treatment are equal
#          H1:Income Means of control and treatment are unequal
var.test(control$income, treatment$income, ratio = 1, alternative = "two.sided")
```

```
##
##  F test to compare two variances
##
## data:  control$income and treatment$income
## F = 0.98973, num df = 11956, denom df = 28090, p-value = 0.5055
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.9603114 1.0202414
## sample estimates:
## ratio of variances
##          0.9897303
```

Variance is equal as p=0.5 (we consider significance at 0.05 for consumer research) and null hypothesis is not rejected

```
#Since population deviance is not known
t.test(control$income, treatment$income, alternative = "two.sided", mu = 0, paired = FALSE, var.
equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  control$income and treatment$income
## t = 1.5206, df = 40046, p-value = 0.1284
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.06581471  0.52136490
## sample estimates:
## mean of x mean of y
##  55.16601  54.93824
```

At 5% significance we can conclude that there is no difference between the average incomes of the control and treatment groups

```
#gender
female_control<-control %>% filter(gender==0) %>%summarise(female=n())
all_control<-control %>% select(gender) %>%summarise(all=n())
female_prop_control<-female_control/all_control


female_treat<-treatment %>% filter(gender==0) %>%summarise(female=n())
all_treat<-treatment %>% select(gender) %>%summarise(all=n())
female_prop_treat<-female_treat/all_treat

prop.test(c(4210,9908), c(11957, 28091), alternative = "two.sided", correct = FALSE)
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(4210, 9908) out of c(11957, 28091)
## X-squared = 0.013935, df = 1, p-value = 0.906
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.010838889  0.009607238
## sample estimates:
##    prop 1    prop 2
## 0.3520950 0.3527108
```

At 5 % significance there is no difference in the proportions of gender between the control & test

```
#gamers
gamer_control<-control %>% filter(gamer==1) %>%summarise(gamer=n())
all_control<-control %>% select(gamer) %>%summarise(all=n())
gamer_prop_control<-gamer_control/all_control


gamer_treat<-treatment %>% filter(gamer==1) %>%summarise(gamer=n())
all_treat<-treatment %>% select(gamer) %>%summarise(all=n())
gamer_prop_treat<-gamer_treat/all_treat

prop.test(c(7196,16892), c(11957, 28091), alternative = "two.sided", correct = FALSE)
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(7196, 16892) out of c(11957, 28091)
## X-squared = 0.0084632, df = 1, p-value = 0.9267
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.00998534  0.01096896
## sample estimates:
##    prop 1    prop 2
## 0.6018232 0.6013314
```

At 5 % significance there is no difference in the proportions of gamer between the control & test

Conclusion: There is no statistically significant difference between the control and treatment groups in terms of the observables. The test and control group are probabilistic equivalent

If there was a large difference between test and control groups, I would check for statistical significance. If there is a statistically significant difference between the control and treatment groups, then this experiment would be invalid and hence I would not conduct the experiment as the results would be inaccurate and no inferences can be drawn. I would need to redesign and ensure test and control are probabilistic equivalent.

```
control_purchaser<-control %>% filter(purchase==1) %>%summarise(gamer=n())
control_nonpurchaser<-control %>% filter(purchase==0) %>%summarise(gamer=n())
control_purchaserate<-control_purchaser/(control_purchaser+control_nonpurchaser)

treat_purchaser<-treatment %>% filter(purchase==1) %>%summarise(gamer=n())
treat_nonpurchaser<-treatment %>% filter(purchase==0) %>%summarise(gamer=n())
treat_purchaserate<-treat_purchaser/(treat_purchaser+treat_nonpurchaser)

prop.test(c(433,2158), c(11957, 28091), alternative = "two.sided", correct = FALSE)
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(433, 2158) out of c(11957, 28091)
## X-squared = 228.56, df = 1, p-value < 2.2e-16
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.04518155 -0.03603577
## sample estimates:
##     prop 1     prop 2
## 0.03621310 0.07682176
```

At 5% significance there is a difference in the purchase rates of control and treatment groups The difference between Treatment & Control purchase rate for all customers is 0.041.

Conclusion: The Promotional campaign is effective.