Restaurant\_CaseStudy

Phanindra Panthagani

November 11, 2018

#import data from excel  
#library(readxl)  
#restaurant <- read\_excel("E:/FALL2018UTD/WebAnalyticsMIS6344.501/assignments/Sec501\_G3.xlsx",sheet = "ads\_analysis%202\_1fd2e6d3-bf38-")  
  
#Saving data to make knitting possible   
#save(restaurant, file="restaurant.RData")  
load(file="restaurant.RData")  
View(restaurant)  
restaurant$treatment[restaurant$treatment==0]<- "control"  
restaurant$treatment[restaurant$treatment==2]<- "ad group2"  
restaurant$treatment[restaurant$treatment==1]<- "ad group1"  
  
rest\_lm1<-lm(reservations~.,data=restaurant)  
summary(rest\_lm1)

##   
## Call:  
## lm(formula = reservations ~ ., data = restaurant)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -30.4488 -2.7384 0.0306 2.7995 28.7426   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.125e+01 4.210e-01 26.723 < 2e-16 \*\*\*  
## treatmentad group2 7.509e+00 1.072e-01 70.040 < 2e-16 \*\*\*  
## treatmentcontrol 3.658e+00 1.167e-01 31.350 < 2e-16 \*\*\*  
## pageviews 3.691e-02 5.715e-04 64.578 < 2e-16 \*\*\*  
## calls 1.041e-01 5.454e-03 19.093 < 2e-16 \*\*\*  
## business\_id 6.980e-05 1.563e-05 4.466 7.98e-06 \*\*\*  
## restaurant\_typeindependent -1.670e+00 2.829e-01 -5.902 3.62e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.643 on 29993 degrees of freedom  
## Multiple R-squared: 0.6626, Adjusted R-squared: 0.6625   
## F-statistic: 9817 on 6 and 29993 DF, p-value: < 2.2e-16

#Business id variable is removed from the linear model since it is not significant (p values=0.665)  
rest\_bid\_rem=restaurant[,-5]  
#Converting treatment to factor variable  
rest\_bid\_rem$treatment <- as.factor(rest\_bid\_rem$treatment)  
rest\_lm2<-lm(reservations~.,data=rest\_bid\_rem)  
summary(rest\_lm2)

##   
## Call:  
## lm(formula = reservations ~ ., data = rest\_bid\_rem)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -30.6262 -2.7306 0.0289 2.7968 28.7036   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 12.7993993 0.2385700 53.65 <2e-16 \*\*\*  
## treatmentad group2 7.8660841 0.0715751 109.90 <2e-16 \*\*\*  
## treatmentcontrol 3.2751252 0.0792176 41.34 <2e-16 \*\*\*  
## pageviews 0.0366507 0.0005688 64.44 <2e-16 \*\*\*  
## calls 0.1045278 0.0054553 19.16 <2e-16 \*\*\*  
## restaurant\_typeindependent -2.6382764 0.1818466 -14.51 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.644 on 29994 degrees of freedom  
## Multiple R-squared: 0.6624, Adjusted R-squared: 0.6623   
## F-statistic: 1.177e+04 on 5 and 29994 DF, p-value: < 2.2e-16

#Q1:Should RG stick to current ad design, or switch to the alternative?  
  
  
#Ans:  
#RG should switch to the alternative ad design (ad group 2)  
#Compared to adgroup 1(base group in rest\_lm2 linear model),the estimate value of ad group 2 is positive with a value of 7.866 and is significant. This means that reservations in adgroup 2 is 7.866 more than that of adgroup 1.  
#(Here, the outcome variable is chosen as reservations because the conversions are more important than the page views and call enquiries.)  
  
  
##Q2:Which outcome variable is most useful to consider? Are there other outcome variables that may be useful to measure  
  
#Ans:  
#The most useful outcome variable to consider is reservations  
#The other outcome variables that may be be useful to measure could be page views and calls; but these would be more important for brand awareness. For profits and conversions, reservations would be a good outcome variable.  
  
##Question 3): Which add design is better for independent restaurants?  
  
#Choosing the interaction terms for better understanding  
rest\_lm3\_interactions<-lm(reservations~pageviews+calls+restaurant\_type\*treatment,data=rest\_bid\_rem)  
summary(rest\_lm3\_interactions)

##   
## Call:  
## lm(formula = reservations ~ pageviews + calls + restaurant\_type \*   
## treatment, data = rest\_bid\_rem)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -30.3114 -2.7486 0.0114 2.7764 29.0272   
##   
## Coefficients:  
## Estimate Std. Error  
## (Intercept) 12.4795681 0.2428857  
## pageviews 0.0376214 0.0005791  
## calls 0.1024794 0.0054581  
## restaurant\_typeindependent -2.8635123 0.1978522  
## treatmentad group2 8.1682416 0.0913491  
## treatmentcontrol 3.1480999 0.0943000  
## restaurant\_typeindependent:treatmentad group2 -0.6896075 0.1351644  
## restaurant\_typeindependent:treatmentcontrol 0.4978785 0.1342049  
## t value Pr(>|t|)   
## (Intercept) 51.380 < 2e-16 \*\*\*  
## pageviews 64.968 < 2e-16 \*\*\*  
## calls 18.776 < 2e-16 \*\*\*  
## restaurant\_typeindependent -14.473 < 2e-16 \*\*\*  
## treatmentad group2 89.418 < 2e-16 \*\*\*  
## treatmentcontrol 33.384 < 2e-16 \*\*\*  
## restaurant\_typeindependent:treatmentad group2 -5.102 3.38e-07 \*\*\*  
## restaurant\_typeindependent:treatmentcontrol 3.710 0.000208 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.639 on 29992 degrees of freedom  
## Multiple R-squared: 0.6632, Adjusted R-squared: 0.6632   
## F-statistic: 8438 on 7 and 29992 DF, p-value: < 2.2e-16

#Question3ans:)From the summary of the terms  
#restaurant\_typeindependent:treatmentad group2 -0.6896075 0.1351644 -5.102 3.38e-07 \*\*\*   
#restaurant\_typeindependent:treatmentcontrol 0.4978785 0.1342049 3.710 0.000208 \*\*\*  
#Adgroup 1 is better for independent restaurants.   
#Compared to ad group 1 ,ad group 2 for chain restaurants has 0.689 less number of reservations and control group has 0.497 more reservations   
  
  
##Question 4:)   
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

rest\_bid\_rem$restaurant\_type <- as.factor(rest\_bid\_rem$restaurant\_type)  
levels(rest\_bid\_rem$restaurant\_type)

## [1] "chain" "independent"

library(tidyverse)

## -- Attaching packages --------------------------------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.0.0 v readr 1.1.1  
## v tibble 1.4.2 v purrr 0.2.5  
## v tidyr 0.8.1 v stringr 1.3.1  
## v ggplot2 3.0.0 v forcats 0.3.0

## -- Conflicts ------------------------------------------------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

#Reversing the factor levels to get chain in the linear model instead of independent. Now we can easily interpret the interaction with chain and treatment  
rest\_bid\_rem$restaurant\_type<-fct\_rev(rest\_bid\_rem$restaurant\_type)  
levels(rest\_bid\_rem$restaurant\_type)

## [1] "independent" "chain"

rest\_lm4\_interactions<-lm(reservations~pageviews+calls+restaurant\_type\*treatment,data=rest\_bid\_rem)  
  
summary(rest\_lm4\_interactions)

##   
## Call:  
## lm(formula = reservations ~ pageviews + calls + restaurant\_type \*   
## treatment, data = rest\_bid\_rem)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -30.3114 -2.7486 0.0114 2.7764 29.0272   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 9.6160558 0.4044872 23.773  
## pageviews 0.0376214 0.0005791 64.968  
## calls 0.1024794 0.0054581 18.776  
## restaurant\_typechain 2.8635123 0.1978522 14.473  
## treatmentad group2 7.4786341 0.1060524 70.518  
## treatmentcontrol 3.6459784 0.1152769 31.628  
## restaurant\_typechain:treatmentad group2 0.6896075 0.1351644 5.102  
## restaurant\_typechain:treatmentcontrol -0.4978785 0.1342049 -3.710  
## Pr(>|t|)   
## (Intercept) < 2e-16 \*\*\*  
## pageviews < 2e-16 \*\*\*  
## calls < 2e-16 \*\*\*  
## restaurant\_typechain < 2e-16 \*\*\*  
## treatmentad group2 < 2e-16 \*\*\*  
## treatmentcontrol < 2e-16 \*\*\*  
## restaurant\_typechain:treatmentad group2 3.38e-07 \*\*\*  
## restaurant\_typechain:treatmentcontrol 0.000208 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.639 on 29992 degrees of freedom  
## Multiple R-squared: 0.6632, Adjusted R-squared: 0.6632   
## F-statistic: 8438 on 7 and 29992 DF, p-value: < 2.2e-16

#  
##Q4ans:  
#From the summary of the terms  
#restaurant\_typechain:treatmentad group2 0.6896075 0.1351644 5.102 3.38e-07 \*\*\*  
#restaurant\_typechain:treatmentcontrol -0.4978785 0.1342049 -3.710 0.000208 \*\*\*  
#Adgroup 2 is better for chain restaurants.   
#Compared to ad group 1 ,ad group 2 for chain restaurants has 0.689 more number of reservations and control group has 0.497 less reservations.