Maryna Pavlenko

Click-through-rate (CTR) Analysis

1. Analysis

Data Processing

```
getwd()
install.packages("dplyr")
library("dplyr")
data <- read.csv("Geo-Fence Analytics.csv")
data$imp_large<- ifelse(data$imp_size == "728x90",1,0)
data$cat_entertainment<-ifelse(data$app_topcat %in% c("IAB1", "IAB1-6"),1,0)
data$cat_social<- ifelse(data$app_topcat  == "IAB14",1,0)
data$cat_tech<- ifelse(data$app_topcat  == "IAB19-6",1,0)
data$os_ios<-ifelse(data$device_os == "iOS",1,0)
install.packages("aspace")
library("aspace")
data$distance <- 6371*acos(cos(as_radians(data$device_lat)) * cos(as_radians(data$geofence_lat)) *
                        cos(as_radians(data$device_lon) - as_radians(data$geofence_lon)) +
                        sin(as_radians(data$device_lat)) * sin(as_radians(data$geofence_lat)))
data$distance_squared<-data$distance^2

data$ln_app_review_vol<-log(data$app_review_vol)
```

Descriptive Statistics

```
install.packages("psych")
library(psych)
attach(data)
variables<-cbind(didclick,distance, imp_large, cat_entertainment, cat_social, cat_tech, os_ios,
ln_app_review_vol, app_review_val)
describe(variables)

cor(variables)
data$distance_group<-as.list(distance)
newdf <- data %>%
  mutate(distance_group = case_when(
    distance_group > 0 & distance_group<= 0.5 ~ "1",
    distance_group > 0.5 & distance_group <=1 ~ "2",
    distance_group > 1 & distance_group <=2   ~ "3",
    distance_group > 2 & distance_group <= 4  ~ "4",
    distance_group > 4 & distance_group <= 7  ~ "5",
    distance_group > 7 & distance_group <= 10 ~ "6",
    distance_group > 10  ~ "7",
    TRUE                              ~ "NA"
```

```
  ))
newdf$impressions<-c(1)
plot(newdf %>%
 group_by(distance_group) %>%
 summarise(ctr = sum(didclick)/sum(impressions)))

plot(newdf %>%
    group_by(app_review_val) %>%
    summarise(ctr = sum(didclick)/sum(impressions)))
plot(newdf %>%
    group_by(ln_app_review_vol) %>%
    summarise(ctr = sum(didclick)/sum(impressions)))
```

Logistics Regression

```
newdf$norm_dist<-(distance-mean(distance))/sd(distance)
newdf$norm_dist_squared<-norm_dist^2
summary(glm(didclick ~ norm_dist + norm_dist_squared + imp_large + cat_entertainment + cat_social +
cat_tech +  os_ios + ln_app_review_vol + app_review_val, family="binomial", data=newdf))
```

The regression output demonstrates the following:
Distance (variable “norm_dist”) has inverse relationships with clicks, because its estimated coefficient is negative. Its p-value is less than 0.05 which indicates that it is significant in the model. Next variable “norm_dist_squared” has linear relationships with the dependent variable “didclick”, its coefficient is positive and p-value is less significant than for the distance.
Imp_large variable is significant in the model and its relationship with the dependent variable is not linear, which means for a one-unit increase in the variable “imp_large”, we expect a 0.35216 decrease in the log-odds of the dependent variable didclick.
Cat_entertainment, cat_social, app_review_val and ln_app_review_vol have very high p-values, therefore they are not significant in this model.
Cat_tech and os_ios are significant due to their low p-values and they have linear relationships with the dependent variable “didclick” which means that for a one-unit increase in these independent variables we expect a 0.68766 and 0.38589 increase in the log-odds of the dependent variable didclick.
With all said, we can conclude that number of clicks depend on the following factors:

1) How far the user is from the AMC Theaters. The farther he is from the advertiser, the less likely he will click on the ad.
2) The size of the impression. The larger the impression size is the less chances the user will click on it.
3) Device type. iOS users click on an ad more often than Android users.
4) Pinger app users demonstrate higher click rate than other app users.

Based on this analysis I suggest some recommendations to AMC Theaters: target iOS users in the close proximity to the Theaters, use small to medium impression size. Pinger app proved to be a good advertising platform for their campaigns. Having all this as focus, it is also recommended to customize and target the ads to other segments of their customers as well.