Poonam Ahuja, Sammi Chang, Isha Chaudhari, Sujaya Darke, Song Han

BANA277 Star Digital Assignment

Table 1
Data Description

Variable Name	Description
purchase	A dummy variable indicating whether the consumer eventually purchased at Star Digital or not =0 if there was no purchase =1 if there was a purchase
imp_1	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 1
imp_2	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 2
imp_3	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 3
imp_4	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 4
imp_5	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 5
imp_6	The number of ad impressions for either Star Digital or the charity that the consumer saw at website # 6
test	A dummy variable indicating whether the consumer was in the test or control group =0 if the consumer was in the control group =1 if the consumer was in the test group

The cost of advertising at Sites 1 through 5 is $25 per thousand impressions, while the cost of advertising at Site 6 is $20 per thousand impressions. While the advertiser cannot control which of Sites 1 through 5 it can advertise on (these sites are part of a single ad network and the ad serving software automatically decides which site the advertisement appears on within the network), it does have the ability to specify if the advertising should appear on Site 6 or Sites 1 through 5, or both these options. A purchase results in a lifetime contribution of $1,200 for Star Digital.

1. Is online advertising effective for Star Digital? In other words, is there a difference in conversion rate between the treatment and control groups?

First, we ran a two sample t-test to test for means. The null hypothesis for the test was that the means are equal. Since the p-value is .06, we can say that this is not significant under an alpha level of .05. However, in this case, we will choose to use a significance level of 0.1 since it is also used as an accepted level of significance in the industry.

```
data:  star_test$purchase and star_control$purchase
t = 1.8713, df = 3309.2, p-value = 0.06139
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.000916332  0.039289257
sample estimates:
mean of x mean of y
0.5048792 0.4856928
```

Keeping this significance level in mind, we ran a regression to see the effect of advertising on the test group as shown below:

```
Call:
glm(formula = purchase ~ test, family = binomial(), data = star)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.186  -1.186   1.169   1.169   1.202

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.05724    0.03882  -1.474   0.1404
test         0.07676    0.04104   1.871   0.0614 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 35077  on 25302  degrees of freedom
Residual deviance: 35073  on 25301  degrees of freedom
AIC: 35077

Number of Fisher Scoring iterations: 3
```

```
> exp(coef(regression))
(Intercept)        test
  0.9443631   1.0797852
```

In the regression results, we can see that at a significance level of 0.1, the variable test is significant. We can interpret the results to mean that online advertising is effective for Star Digital and that if Star Digital’s advertising is shown, then the odds of purchase increases by (exp(0.07676)-1)*100% = 7.98%.

Because purchases are considered as conversions, we can say that there is an increase in conversion rate with advertising for the treatment group in comparison to the control group, and thus advertising is effective for Star Digital.

2. Is there a frequency effect of advertising on purchase? In particular, the question is whether increasing the frequency of advertising (number of impressions) increases the probability of purchase?

To test whether or not there is a frequency effect, we ran a regression with the total impressions, test, and interaction variable of both total impressions and treatment/advertising. We created a variable called Timp which contains the total impressions on sites 1-6. Then, we created an interaction variable which measures the combined interaction of advertising and increased impressions. The result of the regression can be seen below:

```
Call:
glm(formula = purchase ~ Timp_t + Timp + test, family = binomial(),
    data = star)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.9145  -1.1266   0.1299   1.2156   1.2433

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.169577   0.042895  -3.953 7.71e-05 ***
Timp_t       0.015466   0.003207   4.823 1.42e-06 ***
Timp         0.015889   0.002876   5.524 3.32e-08 ***
test        -0.013903   0.045613  -0.305    0.761
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 35077  on 25302  degrees of freedom
Residual deviance: 34190  on 25299  degrees of freedom
AIC: 34198

Number of Fisher Scoring iterations: 5


> exp(coef(regression2))
(Intercept)      Timp_t        Timp        test
  0.8440214   1.0155865   1.0160156   0.9861932
```

Interpretation: For every unit increase in total impressions, the odds of purchase increases by (exp(0.015889)-1)*100% = 1.60%. For every unit increase in total impressions pertaining to the test group, the odds of purchase increases by (exp(0.015466)-1)*100% = 1.56%. Here, the test variable is not significant even at the alpha = 0.1 level, so the estimate of the coefficient, -.013903, does not explain much in variation.

β_0 + timp_t + timp + test

Control Group: Yi = β_0 + timp_t(0) + timp + test(0) = 84% + 101%

Yi = -.017 + .015 * total impressions

One unit increase in total impressions increases the odds of purchase by (exp(.015)-1)*100% = around 1.5%

Treatment Group: β_0 + timp_t(1) + timp + test(1) = 84% + 101% + 101% + 98%

Yi = -.017 + .016 * total impressions + -.014 + .015 * total impressions * test

One unit increase in total impressions increases the odds of making a purchase by (exp(.015889 + .015466) - 1) * 100%, which is around 3%.

Frequency of ads does have a higher effect on odds of making a purchase. The frequency effect is higher for the treatment group than for the control group.

3. How does the conversion effectiveness of Sites 1-5 compare with that of Site 6?

For this question, we ran a regression to compare the effectiveness of advertising on sites 1-5 measured by the variable “Imp15” and site 6 measured by the variable “Imp_6”. Additionally, we also added interaction terms for both the sites to measure the true effect of advertising.

```
Call:
glm(formula = purchase ~ Imp15 + imp_6 + test * Imp15 + test *
    imp_6, family = binomial(), data = star)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.1280  -1.1195   0.1185   1.2217   1.2472

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.166556   0.042533  -3.916 9.01e-05 ***
Imp15        0.019452   0.003443   5.650 1.61e-08 ***
imp_6        0.003978   0.004294   0.927 0.354179
test        -0.006087   0.045314  -0.134 0.893139
Imp15:test   0.014617   0.003794   3.852 0.000117 ***
imp_6:test   0.013483   0.005405   2.494 0.012616 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 35077  on 25302  degrees of freedom
Residual deviance: 34166  on 25297  degrees of freedom
AIC: 34178

Number of Fisher Scoring iterations: 5
```

```
> exp(coef(regression3))
(Intercept)       Imp15       imp_6        test  Imp15:test  imp_6:test
  0.8465755   1.0196422   1.0039860   0.9939313   1.0147240   1.0135741
```

Control group: $yi = \beta_0 + \beta_1 * impressions1\_5 + \beta_2 * impressions6$

$Yi = -0.166+0.0194*impressions1\_5 + 0.00397 * impressions6$

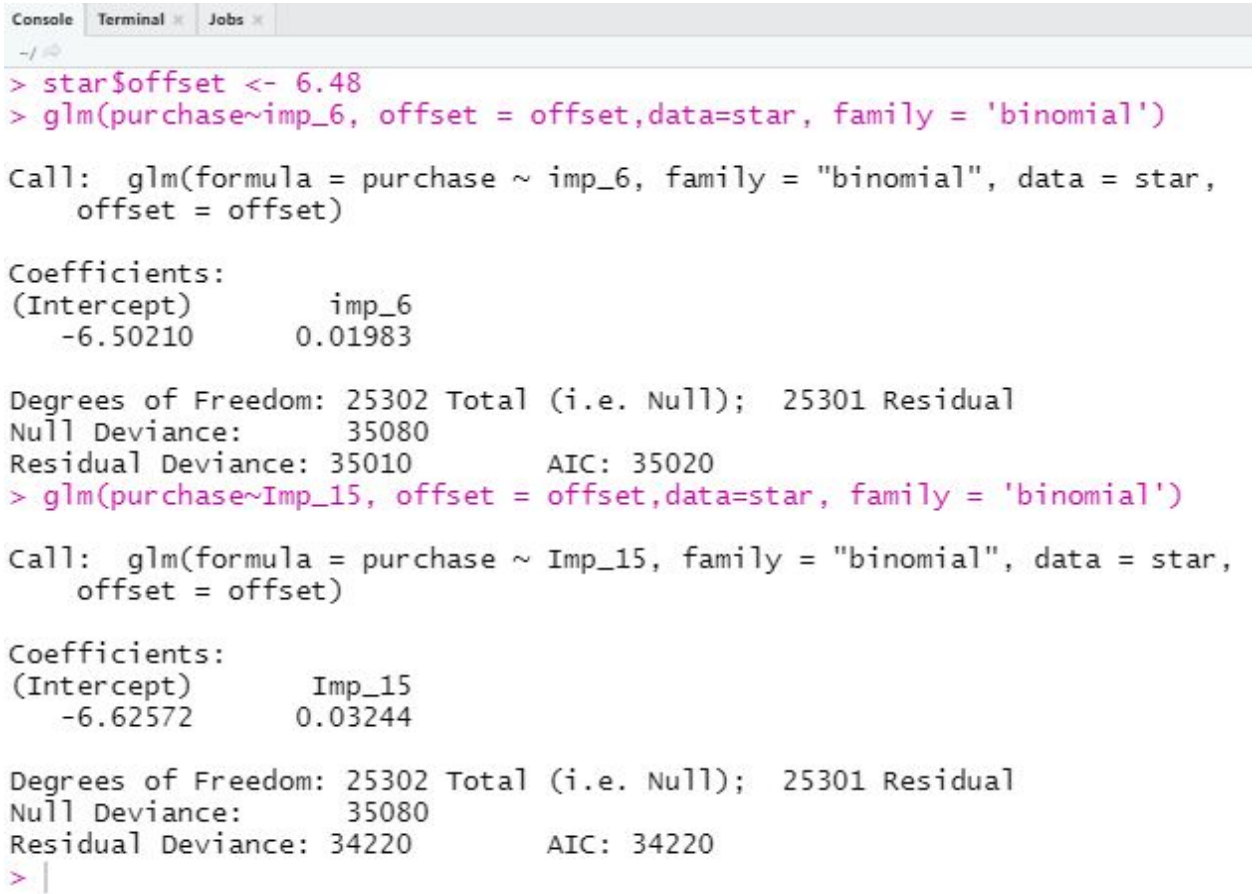
For sites 1-5, the odds of making a purchase increases by (exp(.019)-1)*100% which is around 1.92%. For site 6, the odds of making a purchase increases by (exp(.0039)-1)*100% which is around 0.39%

Treatment group: $yi = \beta_0 + \beta_1 * impressions15 + \beta_2 * impressions6 + \beta_3 * test(1) + \beta_4 * impressions15 * test(1) + \beta_5 * impressions6 * test(1)$

For sites 1-5, the odds of making a purchase increases by (exp(.019+.014) -1) * 100% which is around 3.4% and for site 6, the odds of making a purchase increases by (exp(.0039+.013) -1) * 100%, which is around 1.7%.

In conclusion, the conversion effectiveness of sites 1 to 5 is slightly higher than site 6, based on the calculation of the results for the control and treatment group.

4. Optional Challenge Question -- Which sites should Star Digital advertise on? In particular, should it put its advertising dollars in Site 6 or in Sites 1 through 5?



```
Console  Terminal  Jobs
~/
> star$offset <- 6.48
> glm(purchase~imp_6, offset = offset,data=star, family = 'binomial')

Call:  glm(formula = purchase ~ imp_6, family = "binomial", data = star,
    offset = offset)

Coefficients:
(Intercept)        imp_6
   -6.50210      0.01983

Degrees of Freedom: 25302 Total (i.e. Null);  25301 Residual
Null Deviance:      35080
Residual Deviance: 35010        AIC: 35020
> glm(purchase~Imp_15, offset = offset,data=star, family = 'binomial')

Call:  glm(formula = purchase ~ Imp_15, family = "binomial", data = star,
    offset = offset)

Coefficients:
(Intercept)       Imp_15
   -6.62572      0.03244

Degrees of Freedom: 25302 Total (i.e. Null);  25301 Residual
Null Deviance:      35080
Residual Deviance: 34220        AIC: 34220
>
```

ROI = CLV per customer * probability of buying – cost of each impression *average impressions per person

Probability of buying for sites 1-5: exp(.03244) = p/(1-p) ⇒ p= .508

Probability of buying for site 6: exp(.01983) = p/(1-p) ⇒ p = .505

Sites 1-5 : $1200 * .508 - ($25/1000) * 6.09 = $609.45

Sites 6: $1200 * .505 - ($20/1000) * 1.78 = $605.96

The ROI for sites 1 through 5 is around $4 higher per customer than on site 6. We would recommend spending advertising dollars on sites 1 through 5.