**SMWA - Restaurant Grades Case Assignment**

By Balaji Venktesh(119100)

Gireesh Sundaram(119100)

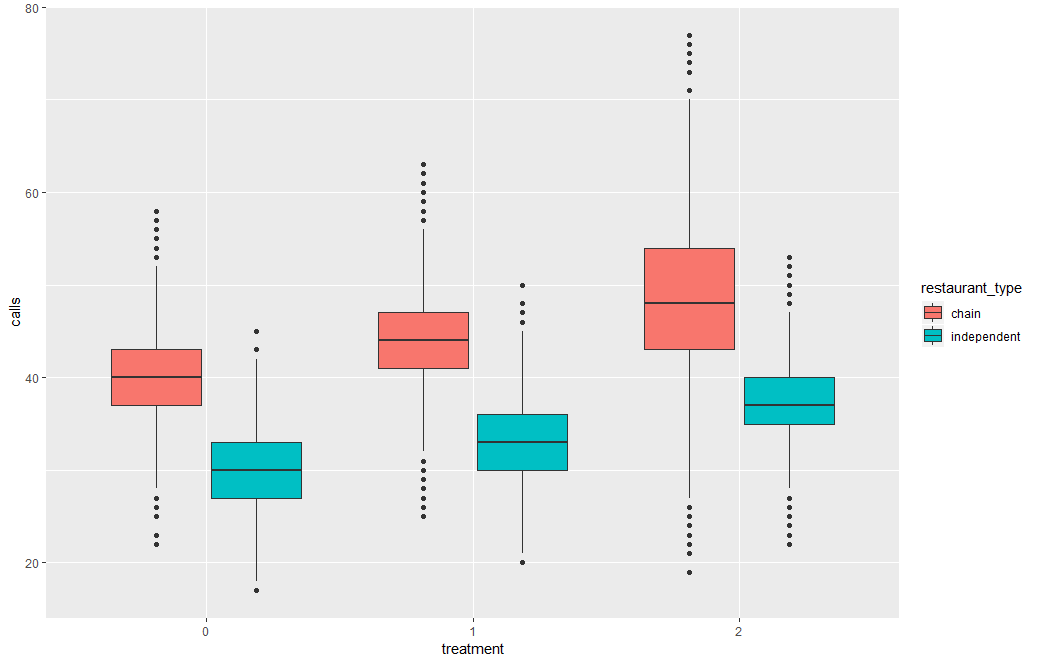
Vineet Kapoor (11910076)

**Q: Given that there is a set of restaurants on the platform that have bought ads, and another one that have not bought ads, why is an A/B testing required. Can we not compare the outcomes for restaurants that do, and those that do not advertise to find the effectiveness of advertising?**

Ans : Yes, A/B testing is required. In this experiment, there are 3 different groups. We can compare number of pageviews, number of reservation and calls using A/B test. We can compare two groups at a particular point of time and analyze which group is more effective in ROI. No, we cannot compare the outcomes of the two groups because there are two more sub-groups in advertising. The ads display algorithm is different.

**Q: Which outcome variable is most useful to measure advertising effectiveness? How will you make this choice? Please explain your answer.**

Ans. **Calls** is more effective outcome variable. By visualizing it, we can see that means of calls is different for all the treatments. Reservations and pageviews are less important as their visualizations for different treatments is not having that much variation in means of calls for each treatment as compared to calls



**Q: We want to compare the chosen outcome variable in the treatment and control groups to ascertain the effectiveness of each type of advertising. Can you say why conducting a regression analysis to do this analysis would be a better approach than using t-tests of means?**

Ans: Regression analysis is used to test the relationship between two variables.

T-test is used to measure the difference in means in case of two different groups.

Using regression analysis, we can estimate the confidence intervals, p-values and differences between treatment groups. Using Regression analysis, we can also find out the amount of variance of data, the model is capturing. It can be verified using Adjusted R-square value.

**Q: Should RG stick to the current ad display algorithm, or should it switch to the alternative? Is the answer same for both types of restaurants? Design and execute a regression model to answer this question and append the screenshot of your regression result to your submission.**

Ans: The firm should not stick to current ad display algorithm and switch to alternative advertisement algorithm. Increment in the mean in calls by applying alternative ad display algorithm is almost 4.00

Using current ad display algorithm, the difference in means in 1.1421.

*Call:*

*lm(formula = calls ~ reservations + pageviews + restaurant\_type +*

*treatment, data = df)*

*Residuals:*

*Min 1Q Median 3Q Max*

*-28.4243 -2.9841 -0.0075 3.0583 25.4825*

*Coefficients:*

*Estimate Std. Error t value Pr(>|t|)*

*(Intercept) 19.0579561 0.3660572 52.063 < 2e-16 \*\*\**

*reservations 0.1155969 0.0060376 19.146 < 2e-16 \*\*\**

*pageviews 0.0272564 0.0006187 44.051 < 2e-16 \*\*\**

*restaurant\_typeindependent -0.6751090 0.1919361 -3.517 0.000436 \*\*\**

*treatment1 1.1421866 0.0854253 13.371 < 2e-16 \*\*\**

*treatment2 5.0728608 0.0840230 60.375 < 2e-16 \*\*\**

*---*

*Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

*Residual standard error: 4.886 on 29994 degrees of freedom*

*Multiple R-squared: 0.6238, Adjusted R-squared: 0.6237*

*F-statistic: 9947 on 5 and 29994 DF, p-value: < 2.2e-16*

**For Chain type of restaurants: The answer is same for both type of restaurants. The alternative ad display algorithm has more calls as compared to current ad display algorithm.**

*Call:*

*lm(formula = calls ~ reservations + pageviews + treatment, data = df\_chain)*

*Residuals:*

*Min 1Q Median 3Q Max*

*-27.9464 -3.7711 -0.0399 3.8093 25.7230*

*Coefficients:*

*Estimate Std. Error t value Pr(>|t|)*

*(Intercept) 17.905948 0.599194 29.883 < 2e-16 \*\*\**

*reservations 0.112598 0.009296 12.113 < 2e-16 \*\*\**

*pageviews 0.029470 0.001044 28.215 < 2e-16 \*\*\**

*treatment1 1.253775 0.164854 7.605 3.05e-14 \*\*\**

*treatment2 4.416500 0.165645 26.662 < 2e-16 \*\*\**

*---*

*Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

*Residual standard error: 6.049 on 11995 degrees of freedom*

*Multiple R-squared: 0.2911, Adjusted R-squared: 0.2909*

*F-statistic: 1231 on 4 and 11995 DF, p-value: < 2.2e-16*

**For Independent type restaurants:**

*Call:*

*lm(formula = calls ~ reservations + pageviews + treatment, data = df\_independent)*

*Residuals:*

*Min 1Q Median 3Q Max*

*-14.7054 -2.5899 -0.0274 2.6299 17.2621*

*Coefficients:*

*Estimate Std. Error t value Pr(>|t|)*

*(Intercept) 1.865e+01 2.634e-01 70.81 <2e-16 \*\*\**

*reservations 1.132e-01 8.168e-03 13.86 <2e-16 \*\*\**

*pageviews 2.647e-02 7.619e-04 34.75 <2e-16 \*\*\**

*treatment1 9.929e-01 9.151e-02 10.85 <2e-16 \*\*\**

*treatment2 5.445e+00 9.112e-02 59.76 <2e-16 \*\*\**

*---*

*Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

*Residual standard error: 3.908 on 17995 degrees of freedom*

*Multiple R-squared: 0.425, Adjusted R-squared: 0.4249*

*F-statistic: 3325 on 4 and 17995 DF, p-value: < 2.2e-16*

**Q: We are interested in knowing whether the current ad display algorithm and the alternative algorithmare better than the control of not advertising for the restaurants. Can you use the analysis you already conducted to answer this question? If yes, how? If not, what can you do to answer this question? Explain. Finally, report your findings to answer the question.**

Ans: No , the above regression analysis can’t be used to answer the question of current ad display algorithm and the alternative algorithm are better than the control of not advertising for the restaurants.

By seeing the coefficients of the result of below regression output. We can conclude that treatment 1 and treatment 2 are both better than no ads group. The difference in means is 3.17 between control group and combination of current advertisement groups and alternative ad group.

*Call:*

*lm(formula = calls ~ reservations + pageviews + restaurant\_type +*

*ad\_type, data = df)*

*Residuals:*

*Min 1Q Median 3Q Max*

*-29.808 -3.172 -0.060 3.103 26.903*

*Coefficients:*

*Estimate Std. Error t value Pr(>|t|)*

*(Intercept) 21.5948825 0.4191144 51.52 <2e-16 \*\*\**

*reservations 0.2769908 0.0050564 54.78 <2e-16 \*\*\**

*pageviews 0.0177195 0.0006021 29.43 <2e-16 \*\*\**

*restaurant\_typeindependent -2.0820916 0.1958837 -10.63 <2e-16 \*\*\**

*ad\_typecontrol group -3.1703111 0.0754011 -42.05 <2e-16 \*\*\**

*---*

*Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1*

*Residual standard error: 5.052 on 29995 degrees of freedom*

*Multiple R-squared: 0.5978, Adjusted R-squared: 0.5977*

*F-statistic: 1.114e+04 on 4 and 29995 DF, p-value: < 2.2e-16*