**The Effect of Launching an Online Community:**

A Business Report for KyngaCell

**Executive Summary:**

The gaming industry has noticed drastic growth and it attracts people across generations. This combined with the mass consumption of internet-based content, gamers join online communities where they interact and share their gaming experience. Just like KyngaCell, companies love online communities because it helps promote their brands and keep track of what people’s perception of the brand and competitors is.

In this report, analyzes the introduction of a new online community feature in the game for the mobile gaming company KyngaCell. The online community provided multiple ways to connect and interact within the game – both during and outside of gameplay. The goal was to analyze if indeed the online community led to an increase in user revenue, retention, Customer Life Time (CLV), and finally to analyze which marketing venues efficiently reached its target audience.

The game team provided with data on user’s behavior for those play the games. Three datasets with 199 users was classified into whether they joined the community or not. Dataset 1 showed user monthly average spending before and after community launch. Dataset 2 recorded the users’ existing time, churn rate and average spending of the last 3 months. Dataset 3 showed whether the users were brought in from marketing campaigns.

Overall, the results showed that the online community increased revenue in the short term, and there is a possibility of long-term increase if KyngaCell offers features with added value. Unfortunately, the results show that the retention rate is decreased for those who joined. This could have been for any unknown variables such as the environment becoming unfriendly or toxic. The recommendation is to investigate and do more research to identify any unattractive feature. On the other hand, the online community showed positive to the CLV in the long term which means KyngaCell should find ways to increase its online user base. Lastly, results showed that campaign user and organic users portray similar behavior and offer similar values to KyngaCell. Therefore, they can continue and leverage campaigns in the future if the cost is reasonable.

**Introduction:**

The U.S. video game industry recorded $43.4 billion in revenue in 2018, and is considered one of the fastest-growing economic sectors. Online community is made of groups of individuals with similar interests who interact on the internet. It’s a popular means of communication among gamers and video game companies. KyngaCell is a gaming company that launched an online community in 2020. The gaming team believes this new feature improved user revenue and retention within the game. In this report, we will test this belief by using quantitative models to analyse the effects of launching this online community. The report will look into the impact of the online community on revenue, on the retention rate, on CLV, and finally on how the marketing campaign reached its target audience.

**Data Description: [will be in appendix later]**

KyngaCell’s game team provided data on user behavior of individuals who play the game which included those who decided to join and those who did not join the online community. Of those, three data sets containing 199 users was classified by whether they joined the community or not. Data set 1 provided their monthly average spending before and after the community launch. Data set 2 recorded the users’ existing time, churn rate and average spending of the last 3 months. Data set 3 provided additional information about whether the users were brought from marketing campaigns. [or all the data description above can be put in the Appendix]. We define users who came from campaigns as campaign users while the others as organic users who naturally joined the game.

[**Model Development**](https://docs.google.com/document/d/1njbd3EDm9jGE-Sh60mlGyGbRr0HFSwVM/edit#heading=h.obdip7v666i6)

Empowered by the data sets provided, different models would be employed to analyze the 4 questions. For question 1, we treat users into 2 groups. One is the users who joined the community and the other one is who did not. Their monthly spending before and after the launch of community would be compared by a model to test the difference (Appendix A). For question 2, a logistic regression model is used to depicture the churn rates of the users in 2 groups (Appendix B). For question 3, each user’s CLV is estimated first (Appendix C). And then a linear regression model is deployed to measure the relationship with their actions of joining the community (Appendix D). For question 4, a number of statistical tests would be employed to detect the difference of user behaviors among the 2 groups.

**Results and Limitations:**

As a preliminary analysis, the result shows that the online community has a significant positive effect on user revenue in the short term. If a user joined the community, it would lead to an increase of $29 more spending in the following month compared to the users who did not join (see Appendix E). Thus, the conclusion that the online community did increase user revenue in the short term (one month in this case). However, the relatively small sample size, 199 users, might limit this extrapolation to millions of users.

Next, it was important to determine whether there was an increase in retention rate. Results show a negative effect of the launch of community to the customer retention rate (see Appendix F). Joining the community is the only factor that affected the retention rate (compared to factors such as time and user spends) such as existing time or spending of users. The churn rate of users who joined the community is 1.5 times greater than (or 2.5 times of) those who did not join. Therefore, joining the online community actually lowered user retention. One limitation is that it only counts for 62.8% accuracy rate. Thus, more user details are needed.

Analysing CLV is an important way to determine the growth of a business. In this case, the results show that the online community led to an increase in CLV (See Appendix G for insights for problem three). Users who joined the community could contribute roughly $52 quarterly than those who did not. This result is supported by the statistical test performed in Appendix H. However, this correlation does not mean causality. For instance,, it is highly possible that the users who joined the community were the existing hardcore users. Even if there was no online community, they would probably spend more than ordinary users.**Also, the CLV calculation is based on the churn rate predicted by the logistic model (Appendix C). It used the existing churn rate to predict future churn rate. But it is still an acceptable approximation for analysis.**

Finally, an analysis of how users joined the game was conducted. Was it organically or through pain campaigns, and what the effects of both were. The results showed that campaign users and organic users generated the same data.  Their average spending, retention rate, percentage of joining the community and the age at the time of joining the community are basically the same.  This is also confirmed by the results of formal statistical testing (see Appendix I to L for raw data & more insight). When the two variables were added into the analysis it did not enhance the prediction of retention rate (See Appendix M for raw data).   In other words, churn rate is not depended on how users came in. In addition, as illustrated in Appendix N, user types have nothing to do with their priority to join the community. Interestingly, this analysis showed that churn rate increased only for campaign users (See raw data in Appendix O). However, two problems emerged. First, the smallness of the sample size questions the validity of the observation. Second, the segmentation of organic and campaign might be not accurate enough to gauge the different.

**Recommendations and Implications:**

KyngaCell should turn more users into online community and develop more new product features to stimulate user spending.[this is the topic sentence, our recommendation] Since the online community did increase user spending in the following month among the users who joined the community, it proved this new product feature is very effective in the short term. Possibly, other new features, as the launch of online community, might also contribute to revenue increase.[evidence one, as question 1 said, short term increase] Moreover, it is seen that the community has a positive contribution to the CLV of joining the community in the long run. [evidence two, as question 3 said, long term increase] Therefore, more users in community would increase KyngaCell ‘s revenue in both short term and long run. [reinforce, so more community user would increase both short term and long term]

Executive managers should investigate into the environment of the online community. The launch of this community decreased user retention rate of those who joined. This might reveal certain issues existed in the community. For example, they should check whether the current community’s environment became toxic. This would cause a negative effect for KyngaCell in the long run. However, we find this increase of churn rate only existed in the campaign users who joined the community. This can provide valuable insights for further research toward community building.

Managements should continue and expand their marketing campaign activities to attract more potential users. The users attracted from the current campaign have the similar behavior as the organic users, which also means similar value. This means the marketing campaign successfully reached KyngaCell’s target audience. KyngaCell can continue and leverage campaigns in the future if the cost is reasonable.

**Conclusion:**

Being one of the fastest growing industries in the market, it’s important that KyngaCell capitalizes on every opportunity available, and invest in the online community. The results of this report show that the online community increased revenue in the short term, and it could possibly increase in the long term if new features with added value are implemented. Results show that user retention rate decreased for those who joined. The recommendation is to investigate the community’s environment for any potential toxicity or unattractive behavior/feature. KyngaCell should find ways to increase the online community because the community has a positive contribution to the CLV in the long term. Lastly, they should expand their campaign because it proves to be more effective in reaching its target audience.

**Reference:**

**Appendix**

Model 1 Difference in difference approach

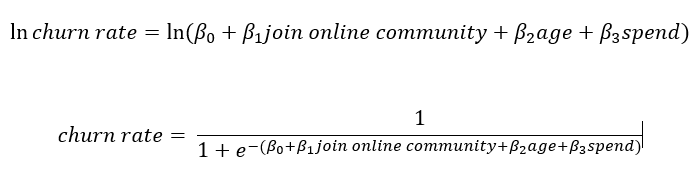
In this model, we assumed that users, whether joined the community or not, have parallel trends in spending over a short period of time. We divided users into two groups. Those who have joined the community went to the treatment group, and those who haven’t were put in the control group. Then we compared the difference in spending between two groups using the Difference in Difference approach.

The model we ran is as follows:

Y= β0 + β1\*[Time] + β2\*[Join Online Community] + β3\*[Time\*Join Online Community] +ε

1. A logistic regression model is used to detect whether joining the online community can lead customers to retain on the game within a three-month timeframe. Besides joining the online community, the ages of customers and the spendings of customers are two potential factors that impact the retention rate. In the logistic regression model, we use whether the customer joined the online community, the age of the customer, and the customer’s spend to predict the churn rate.

The specific model we use is illustrated as below:



1. Customer Lifetime Value (CLV) calculation assumptions

Here, we use the formula below to calculate the value of CLV

CLV = 3\* m \* L - AC

The term m represent the margin for each user, which is 50% of customer spend. The expected lifetime of customer (L) is the prediction output from the logistic regression model between online community and churn rate. Since the raw margin value is monthly and the expected lifetime of customer is quarterly, to balance the equation, the margin value has been multiplied by 3 to become a quarterly margin value. As a result, CLV is quarterly value. Furthermore, since the linear regression is looking for the relationship between CLV and online community, the term AC, which means the acquisition value of users, is a constant and it would affect the value of the coefficients. At last, since the time period is relatively short, we would ignore the time value of money or discounting rate in this model also.

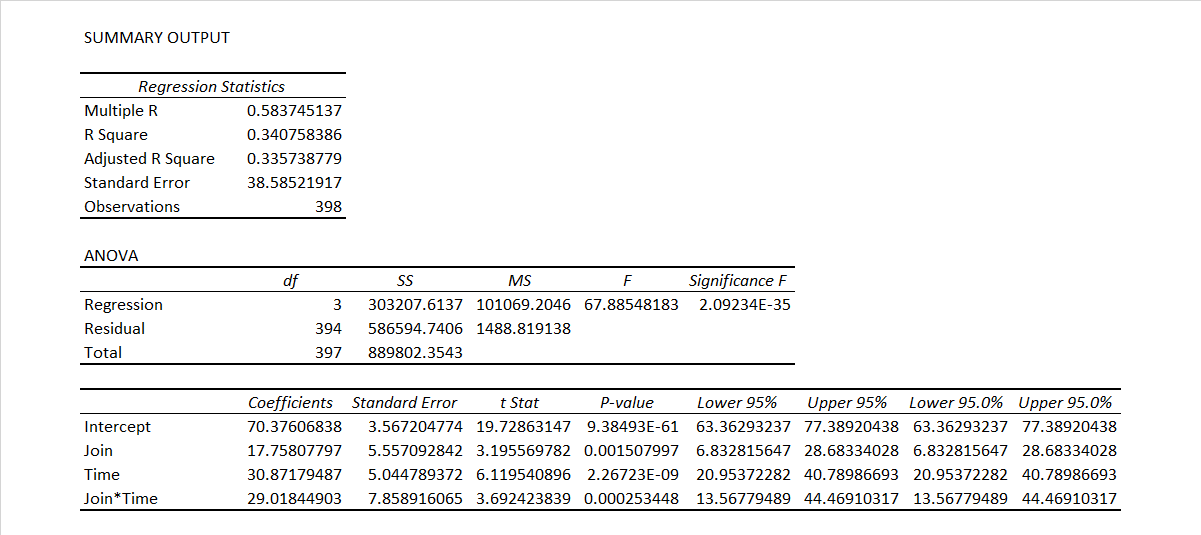
C.

Linear regression is a linear approach to study the relationship between two or more variables. In this case, linear regression is used to predict the relationship between online community and CLV.

The specific model is shown below:

CLV = beta0 + beta1\*I[join] + beta2\*[age] + e

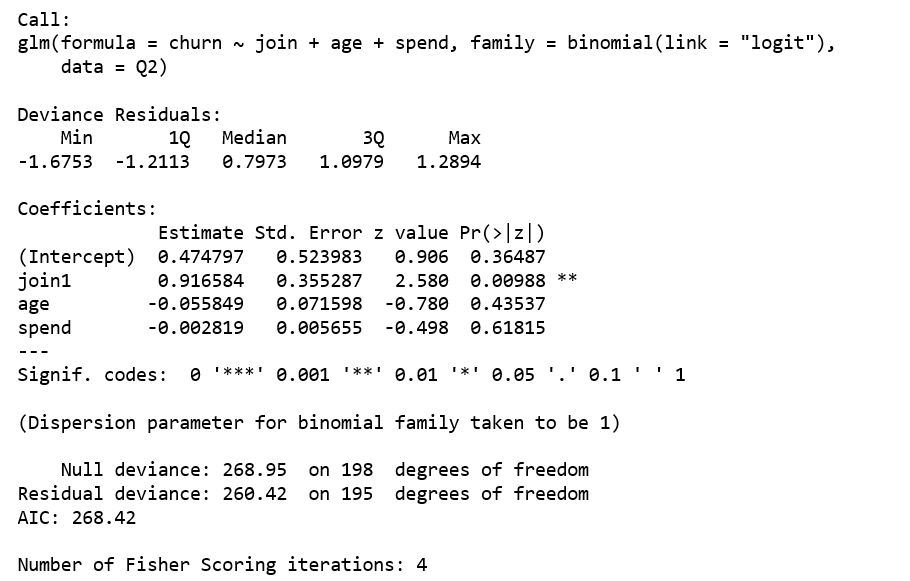
1. Result of difference in difference model



Regression Model using Difference in difference approach

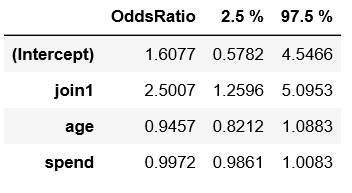
As the result shown above, the intercept β0 ≈ 70.38 means the baseline average revenue of users before joining the community. The coefficient of Join β1 ≈ 17.76, which means the difference between the two groups (Join or Not Join) before the joining decision happens. The coefficient of Time β2 ≈ 30.87, which means the time trend in the control group. The coefficient of Join\*Time β3 ≈ 29.02, which means the difference brought only by joining the community. Here we focused on β3 because it excluded the time effect and revealed the true effect joining the community had on user revenue. The p-value shows it has a significant effect on the revenue. The coefficient is positive, which can be interpreted as joining the community will lead to an average increase of $29 more spending per user in the following month.

F.

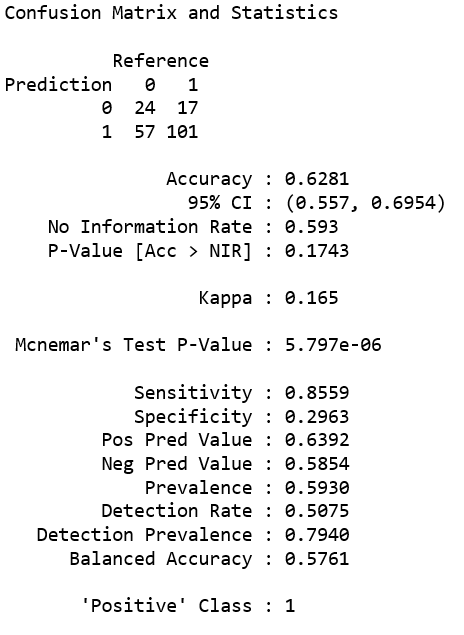


The result of logistic regression model

The model shows that only the action of joining the online community has significant effect (with p-value 0.00988) on the churn rate. The result of the logistic regression model shows that, based on the given data, the age of the customer and the customer spending do not have proved impact on the churn rate. In contrast, the online community does impact on the churn rate. However, the result shows the online community has a negative impact on retaining customers.



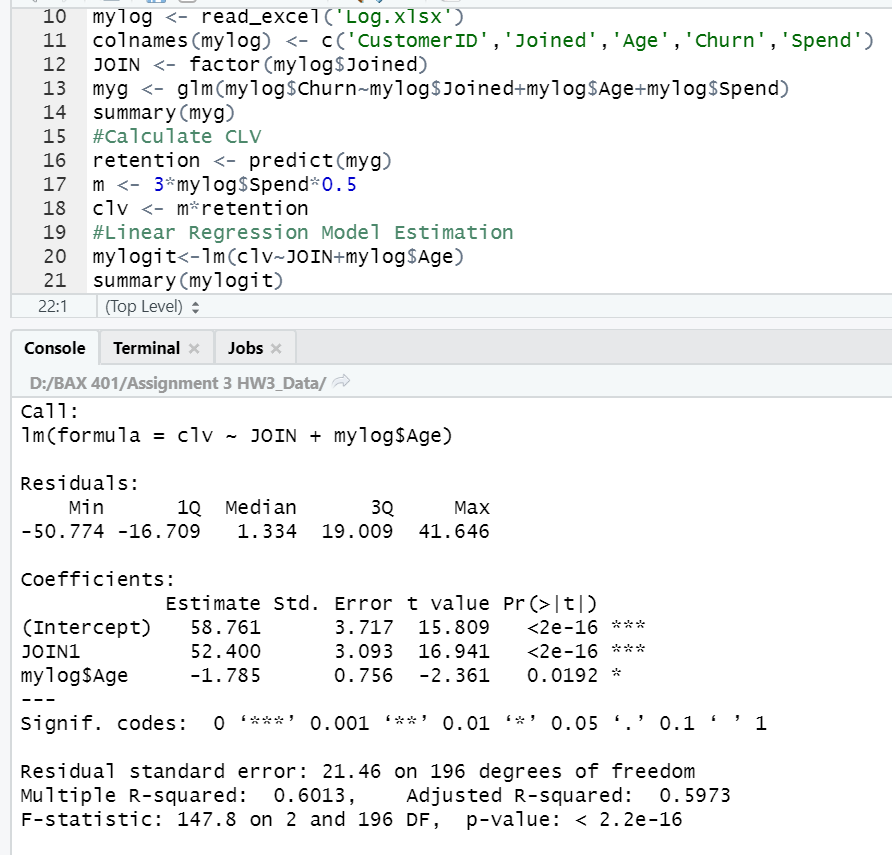
Exponent Coefficients: After transformation the coefficients of variables, we can find the online community’s coefficient with the churn rate is 2.5007, a positive number that means people joined the online community would have 1.5007 times greater probability to churn within three months than the users who did not join the online community.



Confusion Matrix and Statistics: From the confusion matrix, the accuracy of this model is 0.6281, which is an acceptable number.  The F-1 Score is 0.7319, which means the model is fine to predict the churn rate, and the result of the model is acceptable.

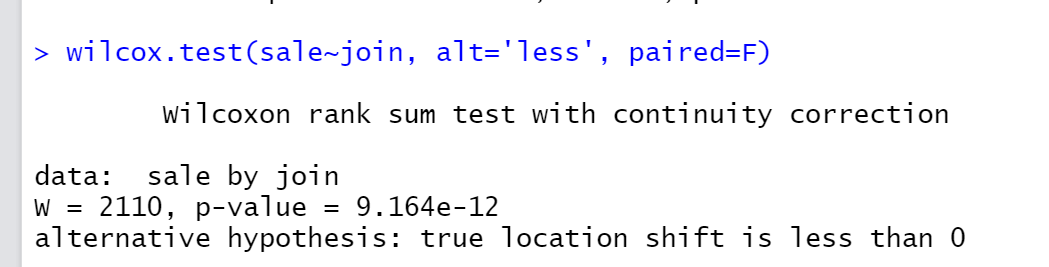
G.

Below is the result of the linear regression model for question 3. From this model, since estimation of Join is a positive value and p-value is extremely significance, the model concludes that there is a strong positive linear regression between CLV and the online community. Specifically, the users who joined the community have $52.4 more CLV in every quarter compared with the users who did not join the community.



H.

Anova Analysis:It is not normal but has constant variance.  So, the wilcox test is used.

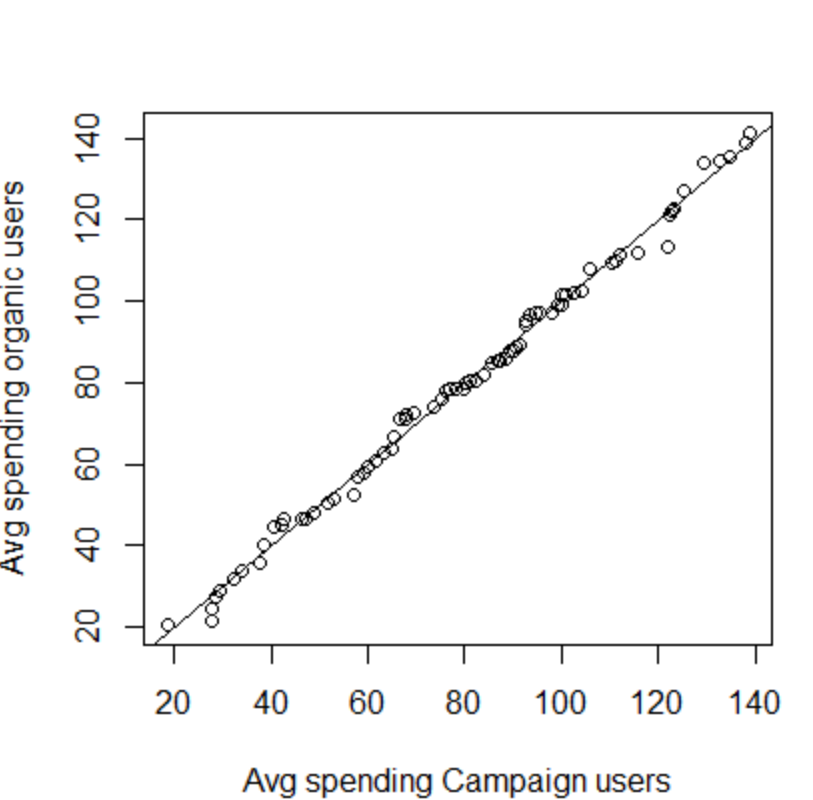


The users who joined community spend more than the users who did not join in the 90 days.

The mean average spending in last 3 months of campaign users is $80.38 with a standard deviation of 30.25.

The mean average spending in last 3 months of organic users is $80.30 with a standard deviation of 30.76.

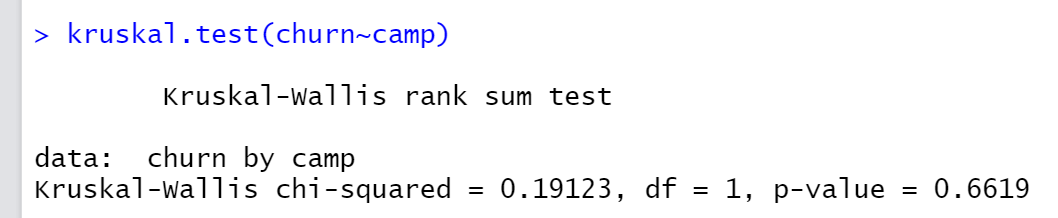
According to the QQ-plot below, we can clearly see there is no difference between the spending of organic users and the users from campaigns.



1. Retention rate of the campaign users is 60.48% with a standard deviation of 0.49

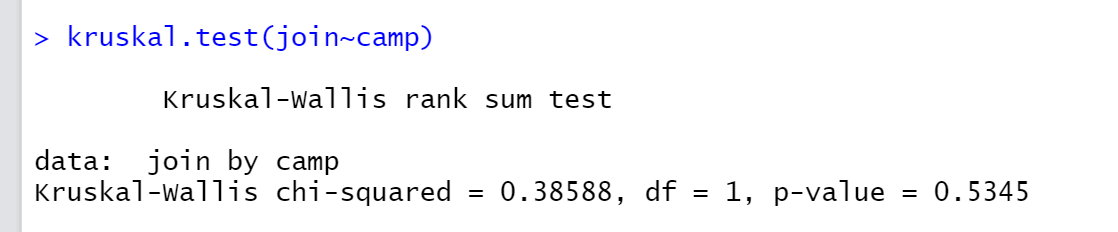
Retention rate of the organic users is 56% with a standard deviation of 0.5

According to the Kruskal test below, we can conclude that there is no statistical difference between the retention rate of organic users and campaign users.



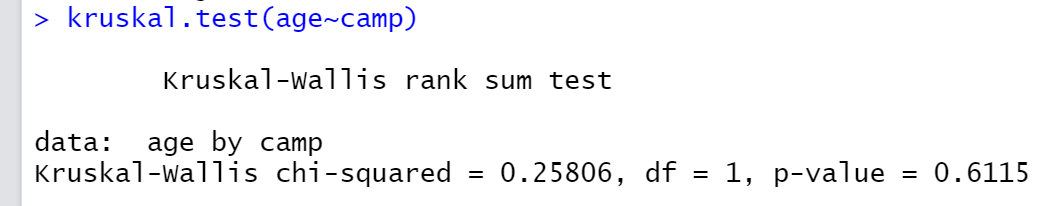
1. In our dataset, 39.52% of the campaign users joined the community and 44% of the organic users joined the community.

However, according to the Kruskal test below, we can conclude that there is no statistical difference between the percentages.

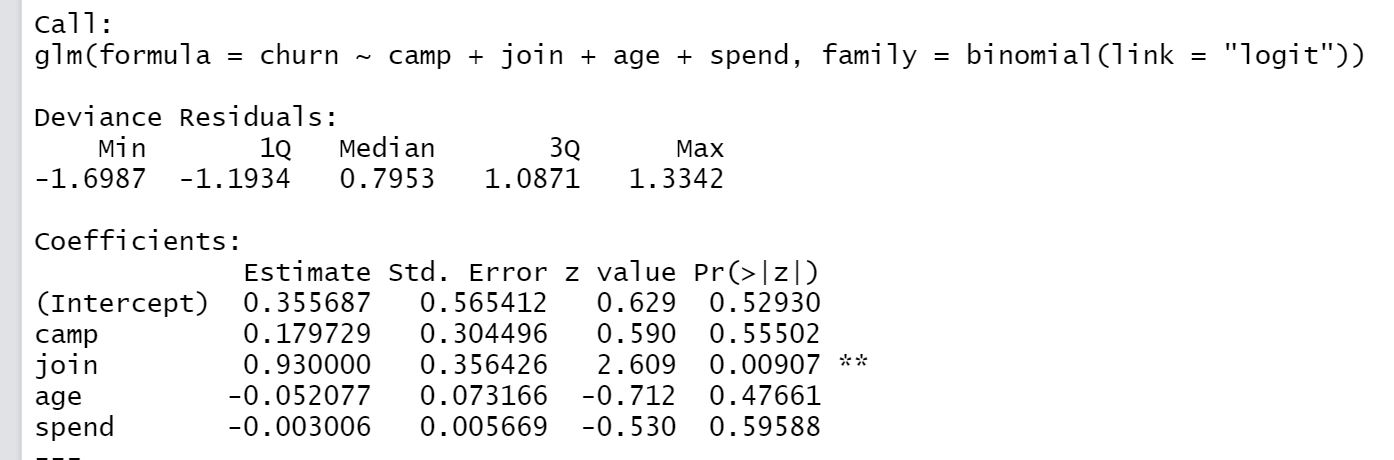


1. The mean customer age at the time of joining the community of campaign users is 4.12 months with a standard deviation of 1.86 and that of the organic users is 4.07 months with a standard deviation of 2.26.

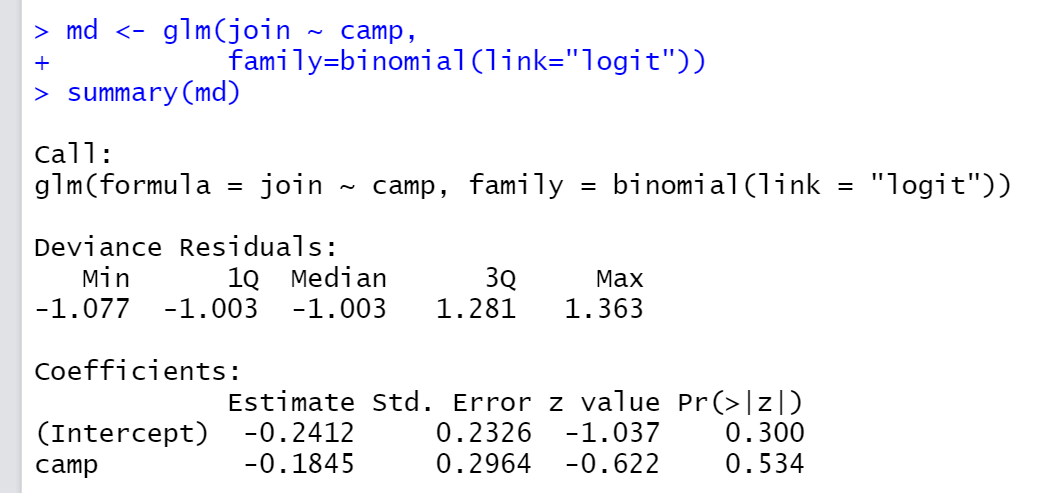
According to the Kruskal test below, we can conclude that there are no differences between these 2 numbers.



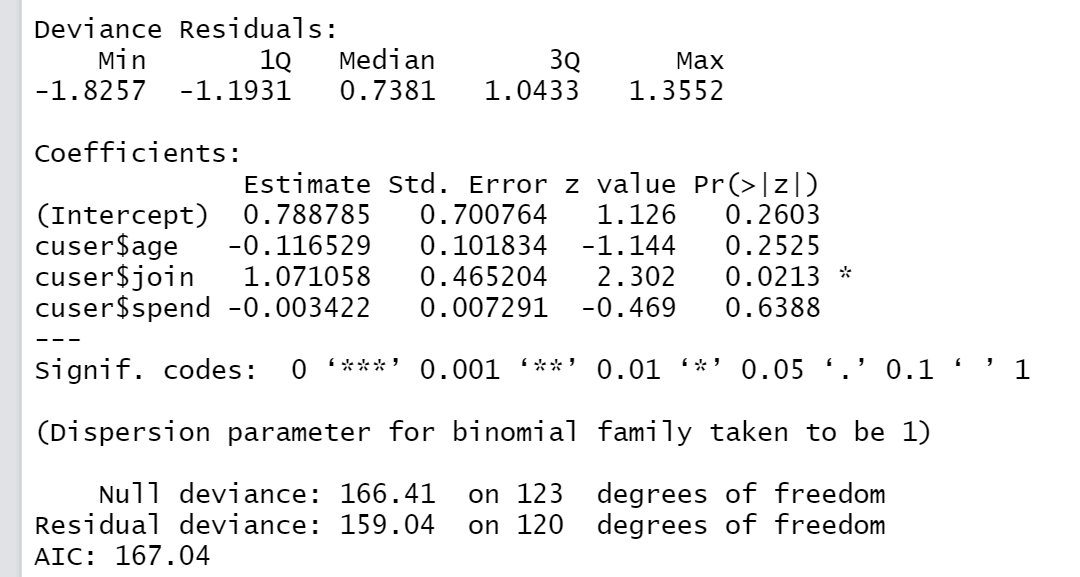
1. By adding the variable ‘Campaign’ into our logistic regression model, there is no improvement to the overall model.  In other words, the ‘Campaign’ variable did not provide any useful information for depicture the churn rate.



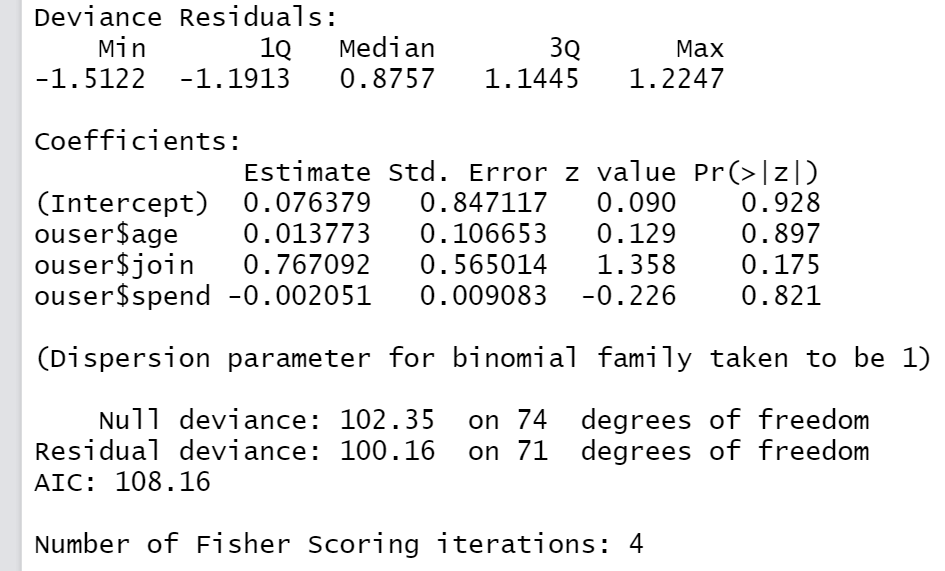
1. Using logistic regression to measure whether users are more likely to join the community based on user types, there is no evidence to say which type of users made a different.  In other words, organic users and campaign users have the same chance to join the community.



O.



The table above shows the result of a logistic regression of churn rate within the campaign users. As we can see, if the campaign users joined the community, it will increase their churn rate. Because the coefficient of the join term is 1.07 with a significant p-value, which proved positive relationships between churn rate and the action of joining the community.



The table above shows the result of a logistic regression of churn rate within the organic users. As we can see, if the organic users joined the community, it will not affect their churn rate.

By comparing these 2 groups, we can find the phenomenon of joining the community will increase churn rate only existed in the campaign users.