**The Effect of Launching an Online Community in the game of Nicht-Soporific:**

A Business Report for KyngaCell

**Executive Summary:**

The gaming industry has noticed drastic growth and it attracts people across generations. As digital social tools built in the game system, companies are leveraging online communities for user engagements.

In 2020, KyngaCell launched an online community in its game Nicht-Soporific. In this report, we would analyze and quantify the effects of introducing this community to the short-term revenue, long-term customer lifetime value (CLV) and user retention rate. Also, users are classified into organic and campaign labels for segmentation analysis. By using 3 data sets of 199 users, a number of quantitative models were employed to calibrate these effects. Based on the model results, findings are listed as following. The online community would increase revenue both in the short term and long term. However, it reduced the retention rate of users. Managers should investigate the community’s environment for any potential toxicity or unattractive features. It can be further discovered this decrease only existed in campaign users. This could provide valuable insights for future community building. At last, the behavior of campaign users is the same with organic users. This proves the current marketing campaign effective reached its target audiences. Therefore, managers should continue and leverage campaigns in the future if the cost is reasonable.

**Introduction:**

The game industry recorded $43.4 billion in revenue in 2018 in US and is one of the fastest-growing sectors in the whole economy (ESA, 2019). Online gaming communities are the place where users share the same passion through a mix of the digital social tools built within the system and two-thirds of the young gamers embraced this interaction (GameTree, 2018; Wilber, 2018). KyngaCell, a leading mobile game company, launched an online community in its game Nicht-Soporific in 2020 which is believed to have increased revenue and retention. In this report, quantitative models would be used to analyze the effects of launching this online community in Nicht-Soporific.

**Problem Formation:**

Three data sets, containing 199 users, are provided for this report. Details can be found in Appendix A1. We define users who came from campaigns as campaign users while the others as organic users who naturally joined the game. In this report, 4 questions would be analyzed. (1) Has the online community increased user revenue in the short term? (2) Has the online community led to increased retention? (3) Has the online community led to an increase in customer lifetime value (CLV)? (4) Whether the marketing campaign of KyngaCell reached its target audiences of the game Nicht-Soporific?

[**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:**

For question 1, 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 (see Appendix E). Thus, we can conclude that the online community did increase user revenue in the short term (one month). However, the relatively small sample size might limit this extrapolation to millions of users.

For question 2, results show a negative effect of the community to the customer retention rate (see Appendix F). Joining the community is the only factor that affected the retention rate compared with existing user age and user spends. 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. More user details would be needed for better prediction.

For question 3, joining the community has a significant positive contribution to the customer lifetime value as illustrated in Appendix G. The users who joined the community contributed roughly 52 dollars more quarterly revenue than the users who did not. So, the online community did lead to an increase in customer lifetime value. Also, 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 used the existing churn rate to predict future churn rate (Appendix C). But it is still an acceptable approximation for analysis.

For question 4, we can find 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 confirmed by the results of formal statistical testing (see Appendix I to L). Also, in Appendix M and N, it can be found churn rate and priority to join the community is not depended on the type of users. But a special phenomenon was discovered. The effect of joining the community which increased churn rate only existed in the campaign users (Appendix O). However, sample size and the segmentation of only organic and campaign might be limitations.

**Recommendations and Implications:**

KyngaCell should turn more users into online community and develop more new product features to stimulate user spending. Becaue the online community did increase user spending in the following month among the users who joined. Possibly, other new features might also contribute to revenue increase. Moreover, it is seen that the community has a positive contribution to the CLV of joining the community in the long run. Therefore, more users in community would increase KyngaCell ‘s revenue in both short term and long run.

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:**

The results show that the online community increased Nicht-Soporific’s revenue both in the short term (one month) and long term (CLV). However, the launch of the online community reduced the retention rate of users. Managers should investigate the community’s environment for any potential toxicity or unattractive features. It can be further discovered this decrease only existed in campaign users. At last, the behavior of campaign users is the same with organic users. This proves the current marketing campaign effective reached its target audiences and should leverage more.

**Reference:**

ESA (2019), *U.S. Video Game Sales Reach Record-Breaking $43.4 Billion in 2018*, The Entertainment Software Association, Available at: <https://www.theesa.com/press-releases/u-s-video-game-sales-reach-record-breaking-43-4-billion-in-2018/>

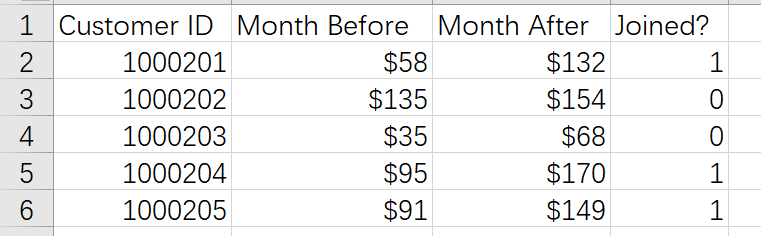
GameTreee (2018), *The Social Benefits of Online Gaming Communities*, Available at: <https://gametree.me/blog/the-social-benefits-of-online-gaming-communities/>

Wilber, J. (2018) *The Many Social Benefits of Playing Video Games*, Available at: <https://levelskip.com/community/The-Many-Social-Benefits-of-Playing-Video-Games>

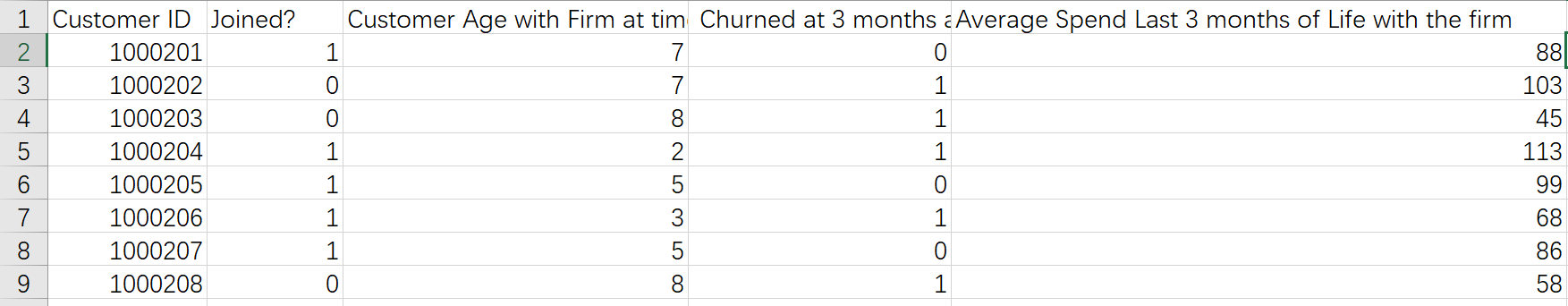
**Appendix**

**A1. Data description and examples**

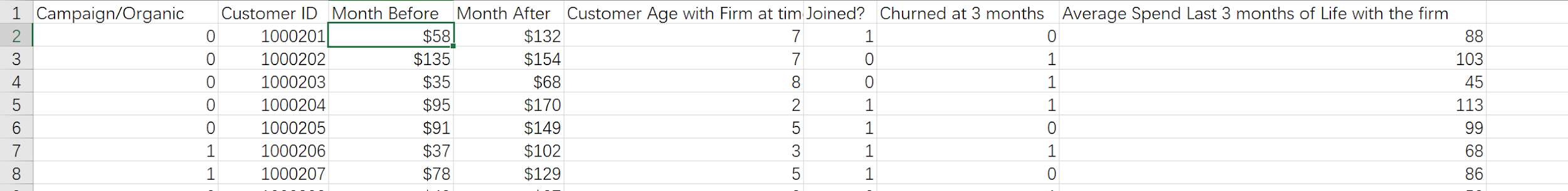
Data set 1 provided their spending in the month before the community launch and the spending in the following month.



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.



**A2.**

**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.

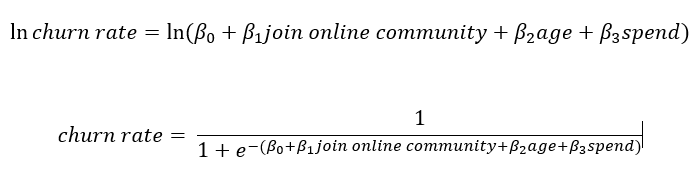
The model we ran is as follows:

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

1. **Logistic model**

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.

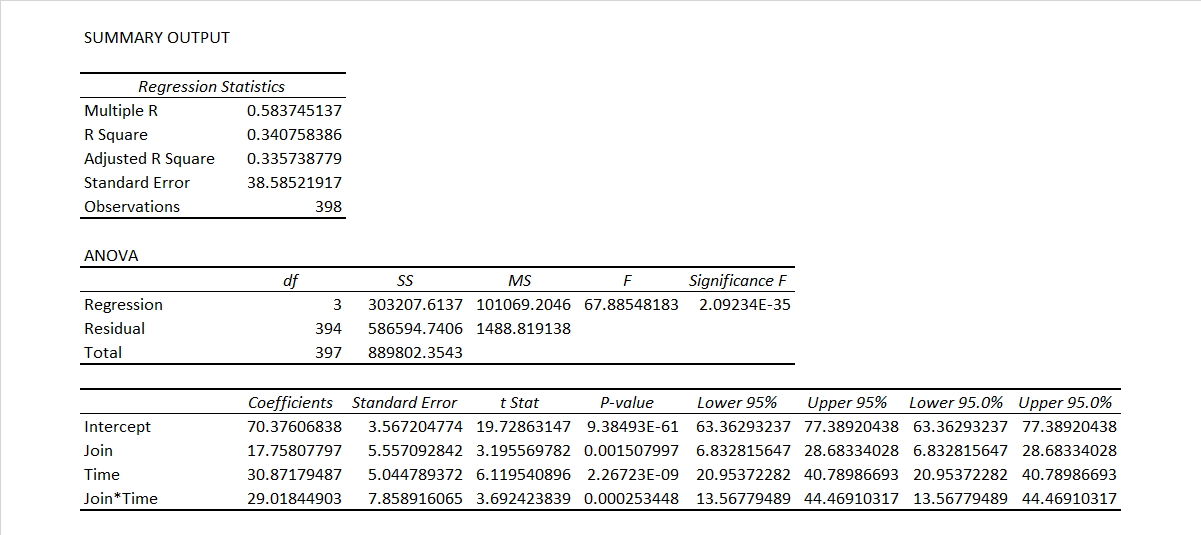
1. **Linear regression model**

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

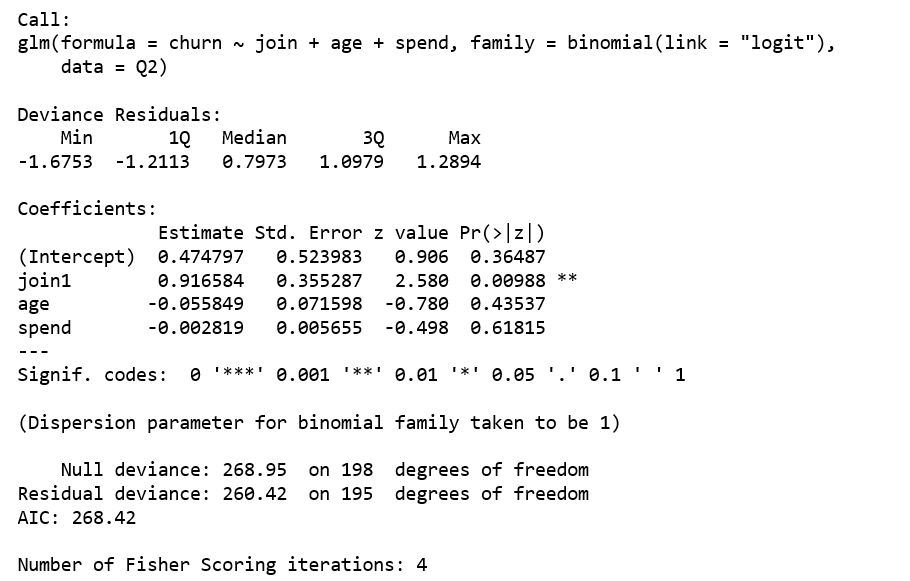
1. **Result of difference in difference model**



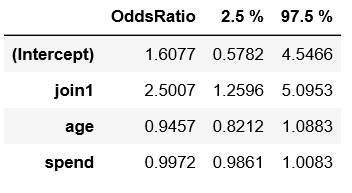
Regression Model using Difference in difference approach

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.

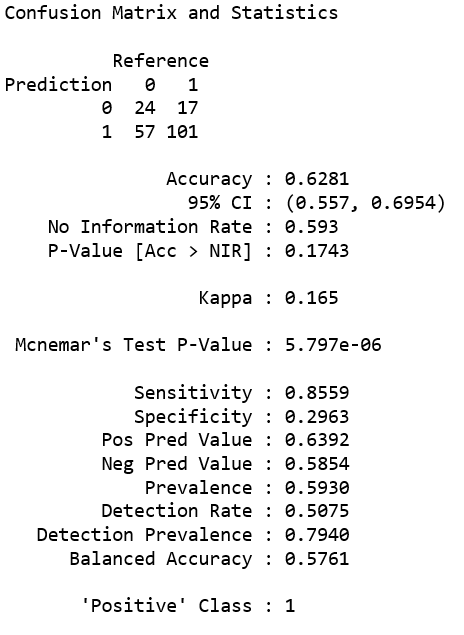
1. **The result of logistic regression model**



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.



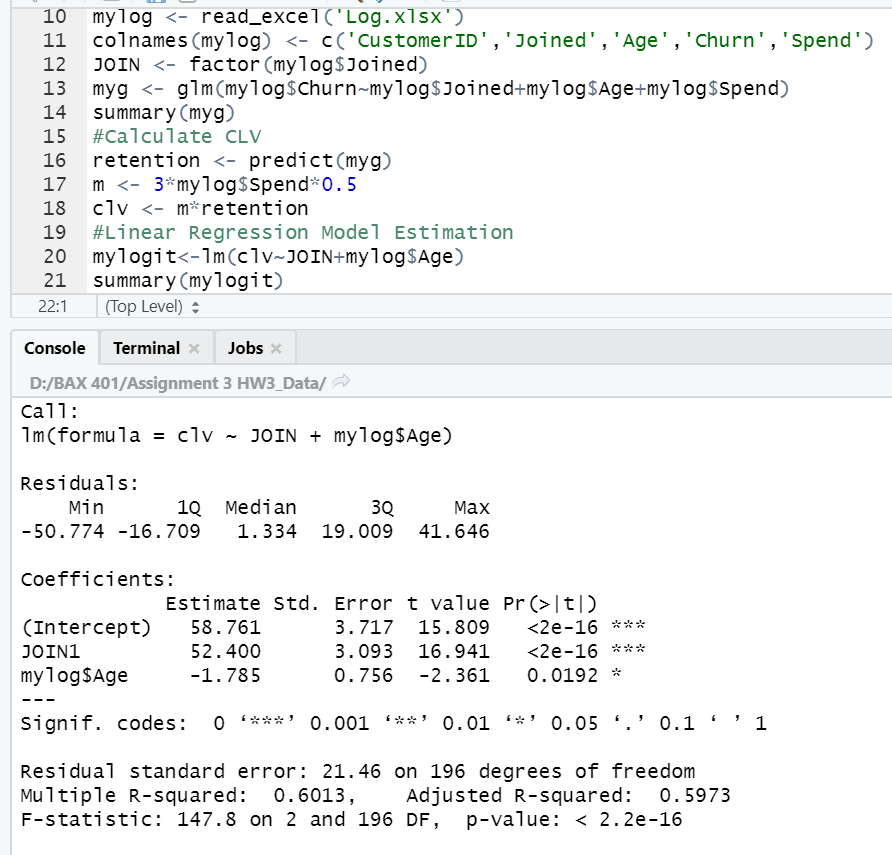
Exponent Coefficients: After transformation the coefficients of variables, we can find the online community’s coefficient with the churn rate is 2.5007, a positive number that means people joined the online community would have 1.5007 times greater probability to churn within three months than the users 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.

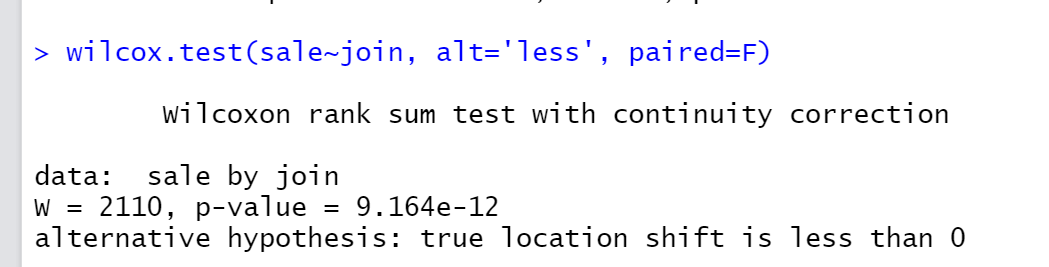
**G. the result of the linear regression model**

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**

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



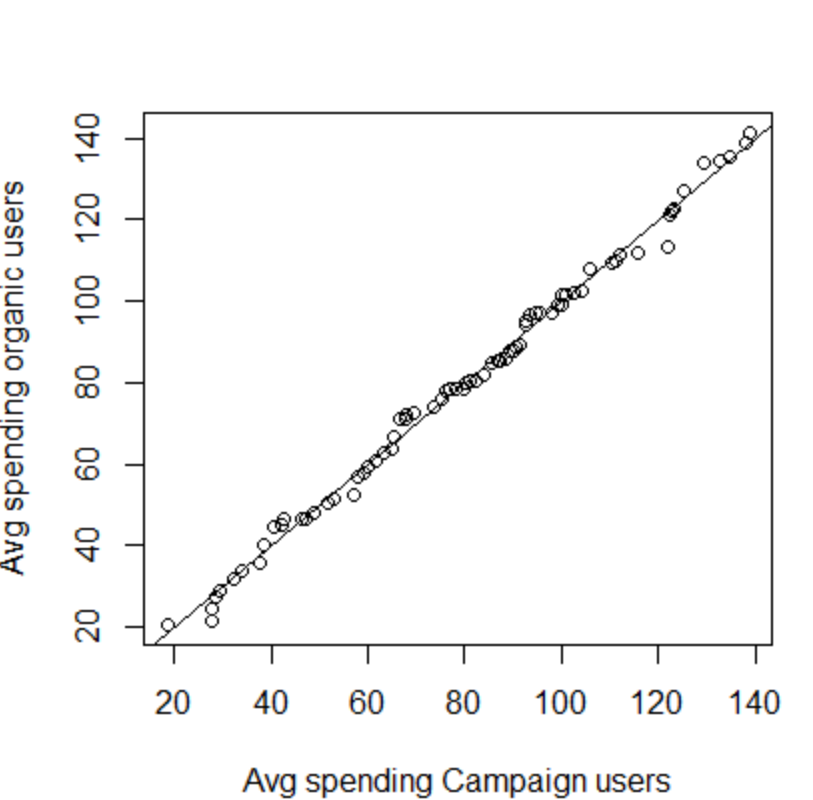
The users who joined community spend more than the users who did not join in the 90 days.

1. **User comparison**

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.

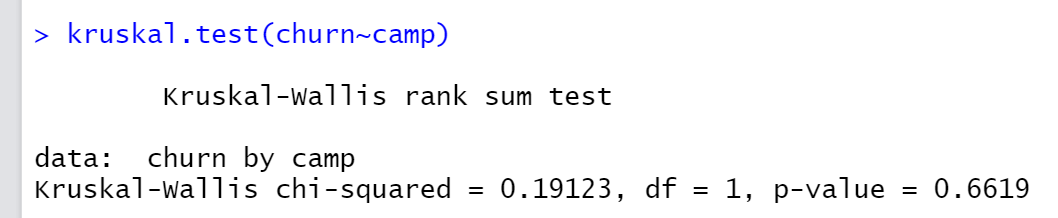
According to the QQ-plot below, we can clearly see there is no difference between the spending of organic users and the users from campaigns.



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

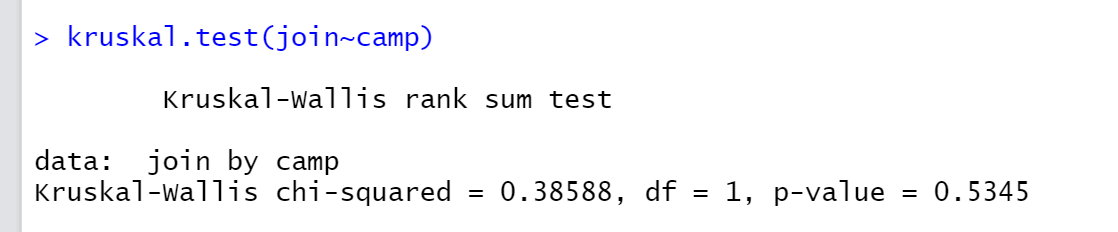
Retention rate of the organic users is 56% with a standard deviation of 0.5

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.



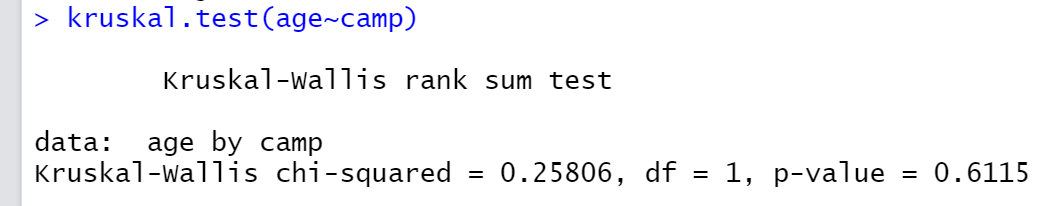
1. In our dataset, 39.52% of the campaign users joined the community and 44% of the organic users joined the community.

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

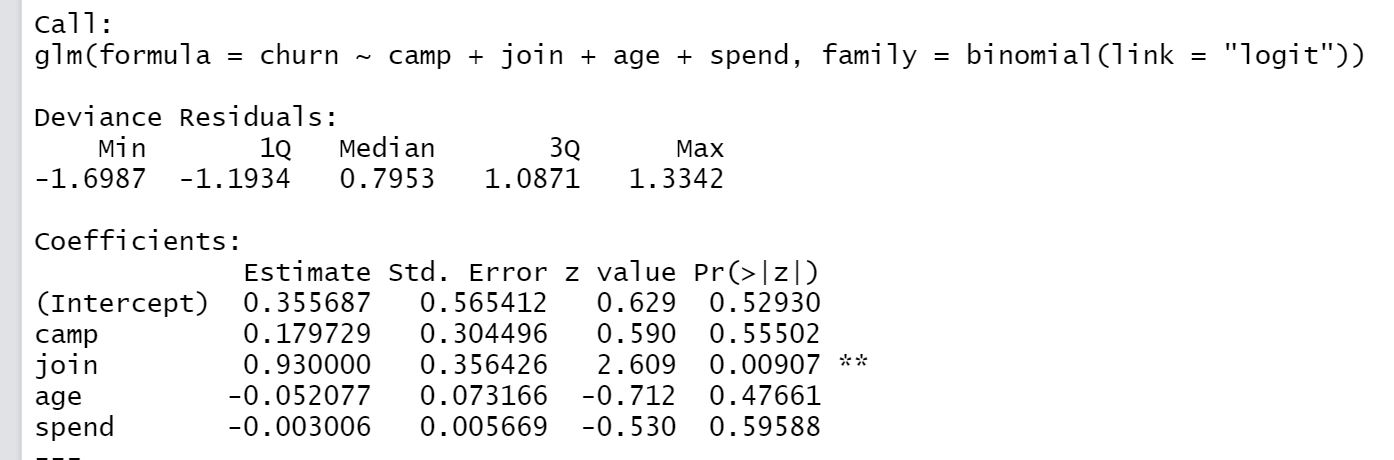


1. The mean customer age at the time of joining the community of campaign users is 4.12 months with a standard deviation of 1.86 and that of the organic users is 4.07 months with a standard deviation of 2.26.

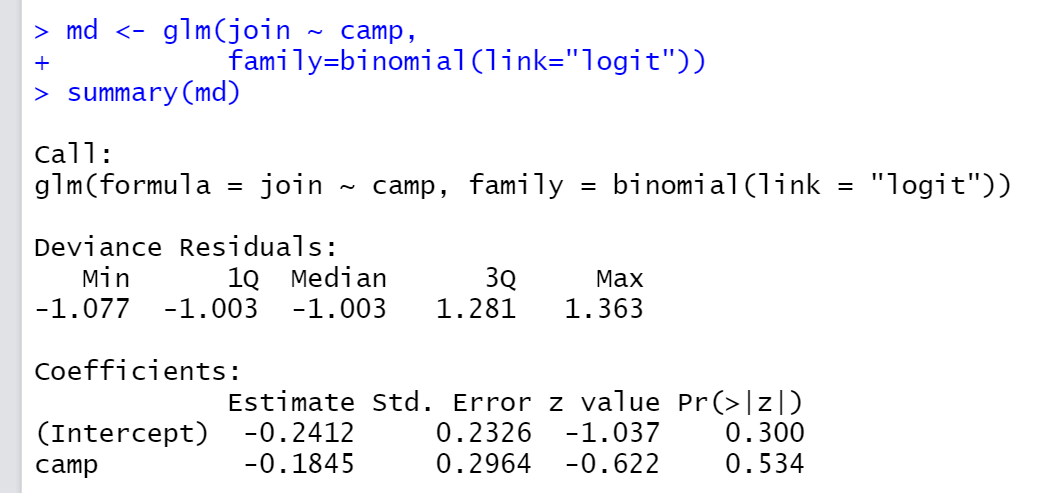
According to the Kruskal test below, we can conclude that there are no differences between these 2 numbers.



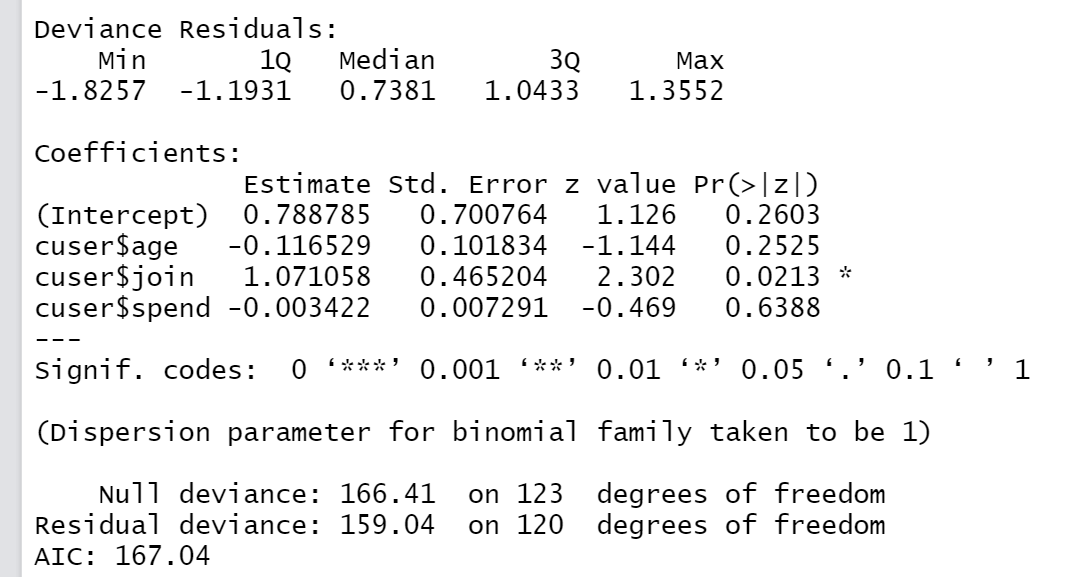
1. By adding the variable ‘Campaign’ into our logistic regression model, there is no improvement to the overall model. In other words, the ‘Campaign’ variable did not provide any useful information for depicture the churn rate.



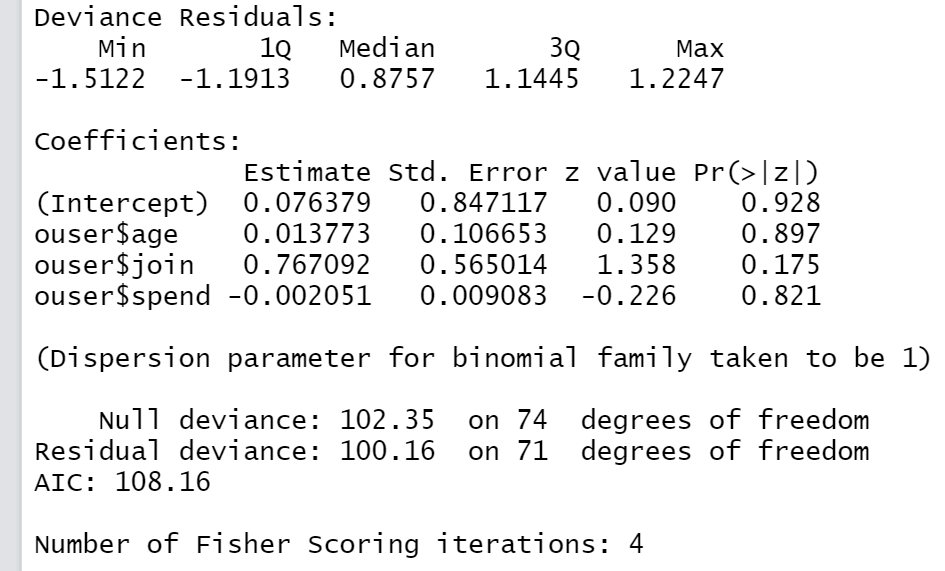
1. Using logistic regression to measure whether users are more likely to join the community based on user types, there is no evidence to say which type of users made a different. In other words, organic users and campaign users have the same chance to join the community.



**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.