Introducing an Online Community at KyngaCell

Section 1: Group E

DATE: 11-11-2021

Table of Contents

Executive Summary:

The report elucidates the impact on the retention, Customer Lifetime Value and revenue for KyngaCell for a novel gaming feature.Diff-in-Diff method for multiple regression models captures the distinctions in outcomes across the groups before and after the online features was introduced in the game. It yields the conclusion that the online community increased the revenue by 29.02$. Logistic regression model results in 21.35% reduction in retention probability. The multiple regression coefficients, we obtain CLV reduced by approximately 20.88$ by the online community. With the inclusion of 'Campaign/Organic' variable in the model, the retention probability and CLV indicates, respectively, 4.1% and 4.02$ less, than for an organically acquired customer. Quantifying CLV, retention and user revenue changes are imperative in forecasting future growth and better decision making leading to increased profitability. It's the one of the KPI which is the highest indicator of the customer experience strategy effectiveness.

Introduction:

In this report we determine the revenue, retention, churn and Customer Lifetime Value change for mobile gaming powerhouse KyngaCell after introducing a new feature in the game. We use Diff-in-Diff and logistic regression to evaluate the user behavior and generate insights. Diff and Diff captures the differences in outcomes across the groups before and after the online features have been introduced in the game. Customer churn is the number of customers that leave in given amount of time, whereas customer retention is the number of customers that stay with the company. CLV is computed by multiplying the client's worth to the firm by their average lifespan. It aids in determining the revenue a company may expect from a client over the course of their lifetime. Using CLV, we forecast the growth in the number of customers in coming years.

Problem Formulation:

To calculate the increased user revenue, we use Diff and Diff, and we use logistic regression for calculating the increased retention and CLV. We take "Churned at 3 months" as the dependent variable and Joined and Customer Age with the firm as the independent variables.We calculated the CLV using the formula:

$$CLV = \frac{Revenue\ generated * Gross\ Margin * Average\ Lifetime(months)}{Number\ of\ customers}$$

We further use AIC, R^2 and Adjusted R^2 to identify and gauge the performance of our best models.Post calculating retention and CLV we used multiple regression to understand their magnitude. For retention, we used calculated retention percentage as a dependent variable and Joined? , Average Spend Last 3 months, Customer Age with Firm at time of launching the online community as independent variables. We used the indicator variable joined to get the magnitude of retention. Similarly, for CLV we changed the dependent variable to calculated CLV and used indicator variables to understand the magnitude of CLV.

Data Description:

The data in this report consists of mainly three sets.

	Campaign/Organic	Month Before	Month After	Customer Age	Joined?	Churned at 3 months	Average Spend
Min.	0	$ 11	$ 28	1	0	0%	$ 19
1st Qu.	0	$ 42	$ 87	2	0	0%	$ 58
Median	1	$ 81	$ 123	4	0	100%	$ 82
Mean	1	$ 78	$ 121	4	0	59%	$ 80
3rd Qu.	1	$ 110	$ 152	6	1	100%	$ 101
Max.	1	$ 145	$ 212	8	1	100%	$ 141

Figure 1: Descriptive statistics of the variables of interest that are considered for modelling.

For the first stage, information on the players is considered which is provided by the gaming team. This contains the information on who joined to engage in the community and who was not. Furthermore, the data shows how much money each of them spent in the game one month before and after the online community was added. This provides a short-term statistic for sales growth that can be traced back to only those who joined the community. This is the first data set.

The data on 90-day churn (i.e., did the player churn within 90 days after the online community was introduced), customer's age with the firm at the time of joining the community in months, and average spend within the game for these 90 days, in order to determine long-term effects such as retention and CLV is considered. This is the second data set.

Finally, the gaming team has an extra variable added to the dataset. The 'Campaign/Organic' variable keeps track of how the player got into the game. A value of '1' indicates that the user joined the game as a result of the company's marketing campaign. A value of '0' denotes that the person joined through organically. This is the third dataset that is evaluated.

Model Development:

For understanding whether the online community increased the user revenue, we identified the focal groups from the first dataset as two different groups 1 as Joined and 0 as not joined.Then we stacked the revenue data and created a treatment column according to the pre- and post-months data. We then applied multiple regression on the stacked data to understand the impact of the online community on the user revenue.

Equation 1:

$$\text{User Revenue} = 70.38 + 30.87 * (Treatment) + 17.76 * (Group) - 29.02 * (Treatment{:}\,Group)$$

Using the second data set to understand the impact of online community on retention and CLV, we built a model based on logistic regression to understand the odds ratio for churn and to quantify the impact of online community on churn and CLV. In logistic regression, the variables of interest were (‘Joined?’, ‘Customer Age with Firm at time of launching the online community’, ‘Average Spend Last 3 months of Life with the firm’) to develop our initial model. The AIC we achieved with the model is 268.54 and then we predicted the probability of churn (probchurn) for each customer.

Equation 2:

$$\log\left(\frac{p(Churn|X)}{1-p(Churn|X)}\right) = 0.46 + 0.92 * (Joined) - 0.052 * (Customer\ Age) - 0.0029 * (Average\ Spend\ Last\ 3\ months)$$

Post that, using the third dataset which now includes the ‘Campaign/Organic’ variable we again applied the logistic regression to understand the difference, which new variable would bring to the model and provide value to the team. We included all the variables of interest used in the above model along with ‘Campaign/Organic’ variable. The AIC we achieved with this model is at 270.18.

Then again using the new model we predicted the new probability of churn(probchurn) for each customer.

Equation 3:

$$\text{Log}\left(\frac{p(Churn|X)}{1-p(Churn|X)}\right) = 0.35 + 0.17 * (Campaign\ or\ Organic) + \ 0.93 * (Joined) - 0.052 * (Customer\ Age) - 0.003 * \ (Average\ Spend\ Last\ 3\ months)$$

We calculated the respective Customer Lifetime Value(CLV) at each customer level for both the models using the predicted probability of churn. To quantify the impact of online community alone on the user revenue, retention and CLV, we created the multiple regression models as needed. In the appendix section, we will discuss these multiple regression models in detail.

Results:

From the multiple regression model created using Difference-in-Difference methodology on the first dataset, we understand that the online community increased the revenue. Since the coefficient of Treatment:Group is significant. And on an average the online community helped revenue increase by 29.02$.

And, from the coefficient of "Joined?" variable obtained from the logistic regression model based on the second dataset, we can infer that the online community has not led to increased retention. Since, odds ratio above 1 indicates that the probability of churn increases. Instead, it increased the churn according to the data. Then, from the output of multiple regression model we also deduced that joining in the online community resulted in 21.35% reduction in retention probability. Furthermore, the online community has also led to a decrease in CLV. We deduced this from the coefficients of multiple regression model used. On an average keeping all other variables used in the model constant when a customer joined the online community the CLV reduced by approximately 20.88$. However, we still need to examine the negative coefficients we are obtaining for "Joined?" variable in detail. Since, the negative coefficients might be arising due to a variety of reasons like insufficient data or lurking variables etc. But, with the present data we can conclude that the online community is not helping retention and CLV.

From the 'Campaign/Organic' variable included model we deduce that while keeping all other variables constant, when the customer is acquired using Campaign, the probability of retention is approximately 4.1% less when compared to an organically acquired customer. Similarly, the CLV of the customer when acquired using Campaign is on an average 4.02$ less than the customer acquired organically. This might be happening because the organically acquired customer might be more attached to the game comparatively.

Recommendations and Managerial Implications:

Evaluation of whether the online community increased user revenue and led to increased retention and CLV is extremely crucial in managerial decisions. The possible avenues that can be an outcome of the results are channelizing budget to maximize profits by optimizing revenue sources, estimating the user acquisition costs, determining the loyal customer base. It aids in evaluating the impact of content revisions and also in assessing the potential of a new game in the beta stage. The spend on campaigns to acquire customers is not bringing in additional revenue as compared to customers who joined organically. The odds of a customer acquired from Campaign getting churned 1.2 times greater than the odds of a customer joined Organically. On the other side, Age is helping in retention and a month increase in age of the customer decreases the odds of churn by 1.1%. The Average CLV for the customers who have not joined the community and acquired through organic channels is higher as compared to people who have joined the community and acquired through campaigns. Customers with lesser age, acquired organically and joined the community tend to spend more over their lifetime.

Conclusion:

In conclusion, with the available data in mobile gaming powerhouse KyngaCell, campaigns to acquire customers and the community created to improve customer engagement are not giving the desired outcomes. As discussed in the earlier sections gaming KyngaCell is performing well in terms of retention and revenue without the online community and marketing campaign.

Appendix:

Model Equations:

Multiple regression model for calculating user retention percentage:

$$\text{Retention} = 0.392 - 0.213 \ * (Joined) + 0.000663 * (Average\ Spend\ Last\ 3\ months) + 0.019 * (Customer\ Age)$$

Multiple regression model for calculating customer lifetime value

$$\text{CLV} = -2.12 - 20.88 \ * (Joined) + 0.87 * (Average\ Spend\ Last\ 3\ months) + 1.04 * (Customer\ Age)$$

After adding Campaign/Organic variable

Multiple regression model for calculating user retention percentage

$$\text{Retention} = 0.417 - 0.0418 \ * (Campaign\ or\ Organic) - 0.216 \ * (Joined) + 0.000685 * (Average\ Spend\ Last\ 3\ months) + 0.012 \ * (Customer\ Age)$$

Multiple regression model for calculating customer lifetime value

$$\text{CLV} = 0.18 - 4.02 \ * (Campaign\ or\ Organic) - 21.22 \ * (Joined) + 0.88 * (Average\ Spend\ Last\ 3\ months) + 1.06 * (Customer\ Age)$$