Produced by Group Tue_09_4:

Shaantam Dar (z5207798)
Jiaying Guan (z5274301)
Faiyam Islam (z5258151)
Xichen Song (z5221864)

Table of Contents

Executive Summary

The FGP customer loyalty program was developed in Asia, currently including three merchant members – the fast food, grocery store and petrol station chain. The purpose of this report is to assist the manager of the loyalty program to understand the optimal way to expand the program and also meet the interests of the three merchant members related to the program through in-depth analysis. 4x Marketing Consultants has conducted a rigorous review through a quantitative analysis of customer characteristics. The following insights will be scrutinised in order to provide comprehensive solutions for the research questions in relation to the managerial problem. Firstly, utilised multiple linear regression to determine which merchant has the greatest impact on customer satisfaction, consequently the manager can target this merchant. According to purchase behaviours of customers, RFM and CLV were used to identify and investigate the most valuable customers. Our team has defined a customer to be 'valuable' if they remain loyal to the program despite other variables impacting their satisfaction. Through binary logistic regression, the report explains in detail there are some possible factors that may lead to customer churn. Although there lacks a holistic model to predict churn rates, we can determine which independent variables may influence customer satisfaction and activeness across the loyalty program in 2016. Regression analysis proves that car/credit card ownership, gender, and city are significant factors affecting the likelihood of promotion. Indeed, we understand that the factors that influence the decision on whether to recommend the loyalty program to others are related to the customer satisfaction with the program. In essence, by analysing how points accumulate within the program, the sales and redemption rates of the three merchant stores are respectively compared to help the manager understand which chain should be targeted. This would allow the manager to determine how to achieve the best interests for three merchant members. There are some limitations to the validity of the findings of this research. These limitations are mainly attributed to possible factors such as relatively small sample size, time constraints and a lack of variables in the data, which will be further scrutinised and explored in the report. As a result, the recommendations provided for the manager of the loyalty program are to adopt a customer-centric approach to understand customer needs and preferences. Moreover, through extensive analysis, the FGP program manager can justify and facilitate expansion across certain merchant chains. The main objective for Jennifer is to fortify and enlarge the existing customer base, using marketing strategies on customer segments and increase sales. Through a combination of optimising customer satisfaction, increasing sales through encouraging customer activeness, reducing attrition rates and targeting the most influential merchant across the program is essential for future investments in the FGP program.

Research Background

Customer loyalty programs (CLP) are an effective marketing and promotional tool to create value for customers to create a relationship between brands and promote repeat business with them, as a result this maximises the profitability from each customer. CLP's are essentially based on the theory of cost-effectiveness; that retaining a customer is five times cheaper than acquiring a new one and that an increase in 5% of retention rate corresponds to an increase of profits by 25% (Saleh, 2020). FGP utilises this theory with their CLP that consists of three merchants including a Fast Food, Grocery Store and Petrol station chain, where accumulated points that are earned through expenditure can be redeemed at any one of these chains, creating more value and increasing customer loyalty. Jennifer, the manager has noted that she would like to expand the loyalty program to increase its effectiveness, however she must keep in mind the importance of "value perception" in this expansion process and how it affects customer loyalty (Yi, Y. and Jeon, H., 2003). Customer's can display either high or low-involvement and are uniquely categorised by their satisfaction of the program, as each customer has different needs it is important for Jennifer to understand the different criteria such as their race, gender etc. rather than just spending habits to justify and predict the impact of this expansion. This level of personalisation is now the most important factor that needs to be considered for FGP as it can potentially increase overall "consumer spending by 500%" (Morgan, 2019). This study will explore the importance of understanding customers at an individual level and how heterogeneity impacts their value perception and engagement within the program, providing personalised solutions and strategies that are customer centric based on empirical evidence.

Managerial Problem

The objective of the managerial problem is to estimate the value of a customer in the loyalty program, predict if a customer will churn, and manage the customers accordingly. As a means to achieve a holistic solution, the relationship between a customer's satisfaction with the 3 merchant chains and their satisfaction with the loyalty program must be heavily scrutinised. Through the accumulation of qualitative and quantitative research, the manager can determine the value of a customer and ultimately produce marketing strategies to improve the current loyalty program. To successfully achieve a solution for the managerial problem, the following research questions have been developed:

1. ***What are the performances of the 3 merchants and which one should the manager target for expansion?***
2. ***How to identify valuable customers in the FGP loyalty program ?***

3. *How do we determine customer churn rate from the loyalty program?*

Research Objective

In order to determine customer value, it is critical that effective marketing strategies are utilised to deliver satisfaction towards each merchant. A major objective for the manager is to scrutinise if it is economically viable to focus on a particular merchant based on the interaction of customers. Understanding the strengths and weaknesses of the FGP program will allow Jennifer to emphasise on existing customer growth and potentially expand the program. It is highly desirable for Jennifer to not only determine which loyalty program provides the most customer value, but also investigate variables which may influence customer churn. Thus, it is paramount to provide Jennifer the necessary solutions for the managerial problem and necessitate an incentive to expand the loyalty program.

Methodology

Data Exploration

The FGP loyalty program consisted of 1,995 customers who have been surveyed for initial data findings and exploration. All customers are situated in a designated Asian country, with ages between 19 and 84. Each individual was asked to provide their satisfaction on the loyalty program with one prerequisite of this survey which was if the customer had been a previous member of the program. In addition, to maintain fluidity in the surveying process, all respondents were asked to provide information regarding their redeem and purchase records. Data regarding all of this information will be invaluable, especially when exploring the ceteris paribus effect of independent variables on a customer's willingness to remain loyal to this program (dependent variable).

Data analysis and interpretation

With the aim of precisely investigating the relationship between various factors which contribute to customer churn and loyalty of the FGP program, a survey has been conducted for the manager to determine the customers' opinions about the program. Various analytical techniques have been utilised to test the overall attractiveness of this loyalty program to customers. This report integrates descriptive statistics, multiple and binary logistic regression analysis, ANOVA (Analysis of Variance), RFM and CLV scores and cluster analysis. The aforementioned techniques will aid the data analysis we conducted to test the significance of variables, determine customer value and churn rates.

Results

1. *According to the performances of the 3 merchants, which one should the manager target for expansion?*

Linear Regression and ANOVA tests

The incentive to implement regression analysis on the FGP program is to determine which merchant performs well with customer satisfaction and allows the manager to target that particular program for expansion. MLR analysis of satisfaction program (dependent variable) against the petrol, grocery and fast food firms (independent variables) have been executed in figure 1. For every 1 unit increase of satisfaction in the petrol firm equates to 0.73 unit of increase on average for the satisfaction program, which, compared to other firms, has the highest contribution. In conjunction, the ANOVA test results in figure 2 showcases a similar pattern with the petrol firm comprising the highest mean squared value compared to the other firms. However, the R^2value indicates 0.43, meaning 43% of the variation of the dependent variable is explained by the independent variables which are the petrol, grocery and fast food firm and 57% is left unexplained. In spite of the low goodness of fit within the MLR model, we deduce that the petrol chain is the most impactful merchant.

Further investigation of Petrol Chain

Figure 3 exemplifies an MLR analysis of the satisfaction program against race and gender, to determine if there is any significance of other lurking variables. The estimated coefficients are relatively low and the p-values for the independent variables stipulate no significance. Furthermore, a binary logistic regression model was conducted to determine any significance of independent variables which may influence customer activeness in 2016. From figure 6, we interpret that the petrol firm contains the highest estimated coefficient of 0.71. The p-value of Sat_Petrol also indicates significance alongside the Sat_FastFood, NetPromoter and Gender_F. Other than the Gender_F variable, the other significant variables share positive linear relationships. McFadden's $R^2$is 0.23, meaning only 23% of the variation in the dependent variable is explained by the independent variables. There is also evidence of the demographics affecting the satisfaction program for petrol chains, with CityA, CityB, CityC and CityE comprising negative coefficients. Petrol satisfaction based on the location of each city is determined by the ANOVA test results in figure 5, where the mean squared value of CityA has the largest coefficient with 50.78 and the p-value indicates the variable is statistically significant. Overall, there seems to be no significant relationships between the city variables and the petrol chain in order for the manager to target a certain

location for customer value and expansion. This analysis is further confirmed through an MLR regression with the petrol firm against the city variables in figure 8, where there are no noteworthy independent variables which contain a large coefficient. In addition, correlation tests between the satisfaction program and the 3 merchants in figure 7 signify a similar pattern of previous results, where the petrol firm has the highest correlation value of 0.58. From the results of MLR, binary regression and correlation tests, we can make a sensible judgement that for Jennifer it will be fruitful to expand upon the petrol chain of the FGP loyalty program to customers.

2. *How to identify valuable customers in the FGP loyalty program?*

Customer RFM Scores

For the purpose of identifying the most valuable customers of the FGP Loyalty Program based on their purchasing behavior, an RFM model will be used in this report. According to Troy (2019), recency, frequency, and monetary are three quantifiable factors that determine the RFM model. Recency represents the time interval since the consumer's last purchase, which can show whether they are an active customer. Whereas frequency reflects the number of purchases made by the consumer in a particular time period. Obviously, customers with high purchase frequency are more engaged and likely to be more loyal than customers with low purchase frequency. In addition, monetary means how much money the customer spent on the FGP Program in a certain period. Monetary is also an important factor because it shows the customer's purchasing power and directly affects company revenue. Figure 9 exhibits how to assign RFM scores. In order to show the equal importance of recency, frequency, and Monetary scores, we divide each variable into 5 levels, and the corresponding scores are all from 1 to 5. Furthermore, customers with low recency are more valuable than customers with high recency; therefore, the more recently a customer has made a purchase, the higher R-score he can get. The frequency scoring standard is similar to recency scoring, the higher the frequency, the higher the score. Since higher monetary value is more beneficial to the FGP program, we are here to score customers, those who spend the most get 5, and those who spend the least get 1. For the RFM score calculation, we select 3 variables (MemberID, SalesDate, SalesAmt) from the 'purchase records' table. MemberID is used to identify customers, while SalesAmt is used to calculate customers' total spending amount. Furthermore, we find the last purchase date and purchase time from SalesDate. Then we calculate each RFM score and analyse the results. Figure 10 shows the descriptive statistics of RFM scores. The mean of RFM score is 7.76 and the maximum score is 15 while the minimum score is 3. When discussing valuable customers, we should pay attention to the customers whose RFM score is above the mean,

especially the customers whose RFM score reaches 15 points . Besides, we can ignore the customers who scored 3 points, because they did not make a significant contribution to the FGP program. Therefore, for indicating the most valuable customers , we list 23 customers with the highest RFM scores in Figure 11.

Customer Lifetime Value (CLV)

According to Kumar (2006), compared to RFM analysis, CLV analysis method is more forward-looking and considers whether the customer will be active in the future. Furthermore, Monica (2011) indicates that the RFM explains the value of customers based on recency, frequency, and monetary value of purchases, but it cannot consider other factors to estimate future purchase behavior and the value of customers. Therefore, in an attempt to find the most profitable customers more accurately, we use the CLV method to predict the net profit contributed by the entire future relationship with the customer. In the process of calculating the customer's CLV, the customer's information of revenue(R), cost(C), acquisition cost(A), discount rate(r), and retention rate(p) should be provided. We are able to calculate revenue(R) by using total sales amount from purchase records. Whereas, retention rate can be extracted through calculating the proportion of the number of active customers in 2016 divided by the number of customers in 2015. From the data, we can know that there are 1499 customers still active in 2016, while the number of active customers in 2015 was 1995; therefore, retention rate(p) is 75%. However, cost, acquisition cost, and discount rate are not given from the FGP database. In this case, we simply assume cost(C) and acquisition cost(A) both are 0 and assume discount rate(r) is 15%. Then using the CLV formula:

$$CLV = \frac{(1+r)}{(1+p+r)}(R-C) - A$$

to drive each customer's lifetime value. After calculating each customer's CLV score, we do descriptive statistics to summarize crucial features on the results. Figure 12 shows that the maximum CLV score is 81362.5 while the minimum score is 3.54. And the mean of CLV is 1650.01. Besides, from Figure 13, we can see that the majority of customers had CLV scores between 0 and 2000, and only 31 customers had CLV scores over 10,000. Because there are 1995 customers' CLV scores, it is hard to show them all in the report; thus, we select 31 customers with a CLV score of more than 10,000 and consider them as valuable customers. Figure 14 shows their Customer ID and we will analyze them in the next part. Both RFM model and CLV model can identify valuable customers. After finding valuable customers, FGP needs to focus more on retaining these customers. First, we can analyse the

characteristics of valuable customers. These common characteristics will help FGP to know which types of customers are more likely to become loyal customers. Furthermore, FGP can provide them with personalised services, for example, when they are identified as valuable customers, the points they consume can be doubled. Personalised services can motivate them to spend and increase their loyalty to FGP.

Cluster Analysis

Through cluster analysis, customers of the loyalty program are segmented for further research. The purpose of this analysis is to reveal the inherent heterogeneity between customers and allow our team to divide customer data into homogeneous groups. The segmentation and cluster group are based on customer satisfaction with the program and the three merchant stores as well as the likelihood of customers recommending the program. To help understand the large data set, a dendrogram shown in figure 15 was constructed, as the hierarchical clustering algorithm allows us to flexibly choose the suitable number of customer segments. Therefore, our team strategically selected four market segments for analysis. According to figure 16, we observe that segment 1 is the largest group, followed by segment 2 and 3, and segment 4 is the smallest. It is found that the manager might need to pay more attention to the customers aged 43-48 years old which account for around 67% of all segments. In addition, as shown in figure 17, the customers of the four market segments mainly originate from CityA. The phenomenon indicates that CityA could be focused on taking more marketing strategies to make it become the largest market of selling the loyalty program as it is an advantageous geographic area. Through comparing customer satisfaction with the three merchant stores in the program, it is significantly shown that customers have the highest satisfaction with petrol station chains among all market segments, at about 8.08 shown in figure 18. Furthermore, according to the average satisfaction levels of the four segments, segment 4 has the highest satisfaction with the program (9.31) and is the most likely to recommend the program (8.63), while segment 2 is the opposite. It can be seen that there is a positive correlation between satisfaction and net promoter score. Therefore, increasing the satisfaction of existing customers with the loyalty program is essential because this factor can produce a great influence on whether customers are willing to recommend the program.

3. How do we determine customer churn rate from the loyalty program?

Customer Churn for FGP Program

Customer attrition revolves around customers willingly halting purchasing or using products from a firm. Determining or predicting which customers will churn is

advantageous for firms as a potential revenue source. Although there is no definitive model to facilitate an exact churn analysis, a binary logistic regression will allow the manager to identify several indicators for potential customer attrition. It is necessary for Jennifer to increase the customer lifetime value and expansion of the loyalty program by increasing the retention rate. Referring to figure 6, customer activeness in 2016 was regressed with the satisfaction rate of the 3 merchant variables and personal factors of the customers such as owning a credit card, car, demographic variables, race and gender. The satisfaction rate for fast food and petrol as well as net promoter and female customers are highly significant as the p-value is less than the 5% significance level. The manager can target fast food and petrol chains for expansion of the program and focus on male customers since the female gender variable is negative, but significant. In terms of the descriptor variables, customers who own a car, credit card and are situated in CityD are significant. So as to achieve an improved retention rate, the manager is required to maximise the satisfaction rate by considering customers with the aforementioned personal factors and demographic, as well as fast food and petrol chains.

Customer Net Promoter Scores (NPS)

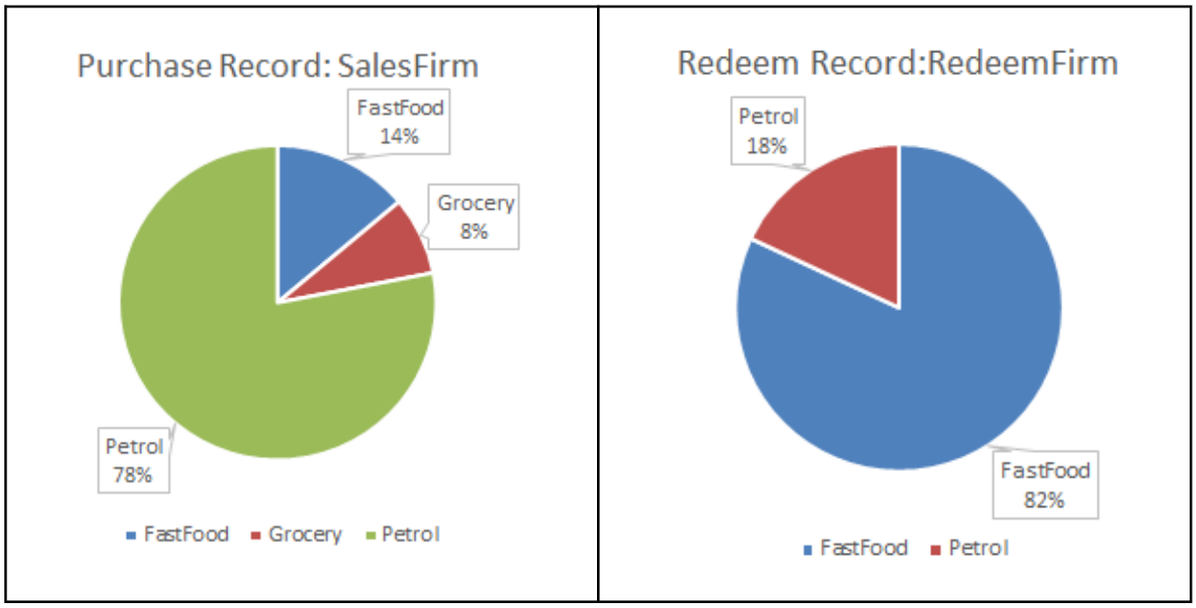
The NPS is calculated below:

$$NPS = \frac{397}{1995} - \frac{1071}{1995} = 19.9\% - 53.7\% = -33.8\%$$

Applying NPS assists our team in understanding whether the participating customers are willing to recommend the loyalty program to others, so as to measure whether the program is truly recognised by customers. Through analysing data from NetPromoter, it is indicated that the NPS of the loyalty program is relatively low. Since NPS could be closely related to retaining existing customers and attracting potential new customers, more efforts are required to increase NPS. According to figure 19, there is a positive correlation between whether customers are willing to recommend the loyalty program and their satisfaction with this program and the three merchant stores. By using a regression model for comparison in figure 20, it is found that the increase in satisfaction with grocery store chains can produce less positive impact on the willingness to promote the FGP loyalty program. While the increase in the satisfaction with fast food and petrol station chains will be more likely to enhance positive attitudes of customers in recommending this program. Therefore, in order to increase the likelihood of existing members promoting the program to potential customers, marketing efforts should focus on improving the satisfaction of fast food and petrol chains.

Which merchant should the manager target?

In order to address and provide a concrete solution for which chain to target for expansion, Jennifer needs to look at whether customers are valuable or not. Jennifer's team defines a 'valuable' customer if they remain loyal to a certain chain despite other independent variables which can impact their satisfaction. Customer purchase rates and redeem rates among these 3 merchants is summarised by the pie charts below:



Interpreting the pie charts, we observe that the petrol firm contains the highest amount of sales, resulting in 78%, compared to 14% in fast food and 8% in grocery. Incongruous to these results, the redemption record indicates the fast food chain with 82% redemption rate compared to 18% for petrol and 0% for grocery. The contrasting results of the petrol chain having the greatest sales and fast food chain having the greatest redemption rate indicates an imbalance. To counteract this imbalance, it necessitates for the manager to focus primarily on these two chains. In essence, we interpret that the churn rate is much higher for the petrol chain compared to the fast food chain. A definite conclusion we can deduce is the omission of targeting the grocery chain in order to target the merchant chains for expansion more efficiently. Furthemore, in establishing a confident approach towards each individual chain expansion, Jennifer is also required to adhere to previous analysis on linear regression, RFM, CLV and Cluster Analysis and combine these results to determine a conclusion.

Recommendations

1. ***Improving Customer Satisfaction through Personalisation:***

   Personalisation is now wanted more than anything else by customers, a study by Wunderman 56% of consumers are more loyal to brands that prioritise their needs and preferences which indicates a link between personalisation and improved loyalty (Abramovich, 2017). As discussed before, retaining a customer is five times cheaper than acquiring a new one therefore an effort must be made to improve satisfaction through personalisation to reduce churn of dissatisfied consumers and increase revenue for FGP. This can be practically achieved through tailoring rewards based on customers' frequency visiting a specific chain, rather than just the basic 0.01 cents per dollar spent. For example, if a customer favours petrol pumps offer cents per liter off while fast food (The most popular redemption chain) offers meal discounts based on upsizing or purchasing extras. This will increase the value of the customer loyalty program for both customers and FGP through increasing CLV scores, improving loyalty and increasing the likelihood of recommending FGP's services through greater Net Promoter Scores (NPS).

2. ***Customer Incentivisation through Referral Programs and Word-of-Mouth Marketing:***

   FGP should also focus on improving Net Promoter Score, this is the indication of how likely a customer is to positively recommend using FGP's services to their networks. As NPS increases this encourages positive behaviour that not only prompts additional purchases at different chains but also provides Word-of-Mouth (WOM) marketing to acquire new customers organically. WOM is incredibly important as a study showed that 74% of customers indicated that this is a key influencer in their decision making process to conduct business with an organisation (LoyaltyLion, 2018). This can be achieved through strategic marketing and other product perks to encourage consumers who actively refer others and increase the NPS of customers who currently sit low on this scoring system to create a feedback loop. FGP should utilise these strategies specifically on Petrol and Fast food chains as this is where the majority of their revenue is stemming from to improve NPS, reduce churn, improve customer company relationships and increase market share.

3. ***Build partnerships with other businesses (B2B partnerships):***

   A vital aspect in strategic marketing is the inclusion of business partners. In today's digitally and data-driven world, customer preferences and expectations are continuously changing (Nicita, 2019). It's crucial for FGP to attain access to customer bases from other businesses and potentially target and attract customers

who own a car and credit card as well as CityD which was the only statistically significant variable. By building partnerships with other businesses with a similar goal, Jennifer can expand the loyalty program whilst simultaneously minimising churn rates. It is also recommended that the manager is able to grasp unique insights from other businesses through a B2B partnership on different analytical tools to use and identify more about their customers.

4. *Redistribution of Costs to promote Healthy Competition:*

It is also important for FGP to consider the satisfaction of its merchant managers who are looking after the day-to-day operations of chains, to address the concerns of compensation for more important chains that SongTing rose. Through an analysis of the chains that are bringing in the most value to FGP we have deduced that the Petrol chain is the most effective, as this chain has the highest influence on the satisfaction program in terms of regression and the greatest purchase and redemption rate . With this we recommend a periodic redistribution of operating costs for the customer loyalty program to Grocery, Fast Food and Petrol to 70%, 20%, 10% respectively, which can be adjusted in future years to meet changing performances of chains. This strategy will also promote healthy competition between chains to improve motivation, place more effort to please customers, innovate and to achieve higher sales targets, inadvertently reducing the churn rate across low-performing merchants (Wardley, Cable and Steinhage, 2017). However, this strategy should be used with consideration and watched closely as merchants should not be using unethical means to redistribute costs.

Limitations

Through an analysis of the survey data provided by FGP's customer loyalty program we have noticed that there may be a level of sampling bias (Jovancic, 2021) which may have skewed our conclusions, as the survey has only captured a small sample of 1995 customers using the CLP. This relatively small sample size can be a causation for the low adjusted R-squared values calculated through the discussion of our multiple linear regression modelling and may not be an accurate representation of current users of the customer loyalty program. For future reference and more accurate analysis of the CLP it is important to have a larger survey sampling size to mitigate this sampling bias that may have interfered with our strategic recommendations.

Furthermore, the FGP customer loyalty program has been effectively utilised for the past 8 years as indicated by Jennifer the manager. However 4x Marketing Consultants have only been provided with the selective data of customers, purchasing and redeeming information between 2015 to 2016. This has created a time period bias and made the analysis of correlation within customer lifetime value, heterogeneity of behaviour and which customers are likely to churn challenging to extrapolate. If a larger sample size had been provided across the 8 year period the accuracy and relevance of our analysis would be in better support of our recommendations for Jennifers managerial needs. The survey itself is also quite small and doesn't ask customers enough questions to assist with our analysis, we believe that there may be other variables such as personal income and data on each merchant's operations which would have allowed us to further reinforce our solutions.

Additionally, our analysis was further limited due to missing data in the Birth Year category of customers utilising the loyalty program as 44 respondents had their data categorised as "NULL" and the presence of outliers skewing our results. We are also missing major data to calculate the CLV of average gross margin per customer lifespan, the rate of discount and acquisition costs. Hence, through the expansion of the customer loyalty program data sampling size across a larger time period and addition of questions to the survey would enable 4x Marketing Consultants to develop stronger data driven recommendations. Through adopting these changes there may be more evidence to support our strategies to reduce churn, better tailor rewards to customers behaviours and meet the needs of the manager.

Conclusion

To increase the strength of the customer base, through statistical analysis of FGP's survey data Jennifer can predict the ideal customers to target for the customer loyalty program. We set ourselves 3 research questions to extrapolate FGP's data and were able to synthesis key recommendations based on what we uncovered. These questions were focused on, which merchant should we be targeting for expansion, what makes a valuable customer and how we can determine churn rate from the program in itself. Through our analysis we have deduced that the petrol merchant should be prioritised due to the firm having 78% of purchases which make up the majority of FGP's sales while the fast food chain should also be targeted as it contributes to 82% in redemption rate. Furthermore, customers that are 43-48 years old account for 67% of segments and males from cityA should be a focus on the types of customers that should be targeted by the strategies provided. Further predictors of customers who are likely to be valuable customers are ones with car/credit card ownership as well. Finally, adopting a customer-centric and personalised marketing approach, incentivised referral programs to the customer loyalty program, paired with an internal strategy of redistributing costs this will poise FGP to maximise customer satisfaction and reduce churn.

References

Abramovich, G., 2017. *Study: Majority Of People More Loyal To Brands That Care About Them*. [online] Adobe Blog. Available at: <https://blog.adobe.com/en/publish/2017/01/28/wunderman-tlp-wantedness-study.html#gs.62x3kr> [Accessed 11 July 2021].

Effects of loyalty programs on value perception, program loyalty, and brand loyalty. *Journal of the academy of marketing science*, 31(3), pp.229-240.

Jovancic, N., 2021. *5 Types of Bias in Research and How to Make Your Surveys Bias-Free*. [online] LeadQuizzes. Available at: <https://www.leadquizzes.com/blog/types-of-bias-in-research/> [Accessed 10 July 2021].

Kumar, V. (2006), *Customer lifetime value*. http://www.drvkumar.com. Accessed on 8.09.2010

LoyaltyLion, 2018. *A guide to loyalty and NPS: everything you need to know*. [online] LoyaltyLion. Available at: <https://loyaltylion.com/blog/a-guide-to-loyalty-and-nps> [Accessed 11 July 2021].

Monica, T., 2011. *Comparison between Customer Lifetime Value (CLV) and Traditional Measurement Tools of Customer Value*. [online] Poseidon01.ssrn.com. Available at: <https://poseidon01.ssrn.com/delivery.php?ID=299118127070015118110125073107114107103081020021005061018084066068096112022116003030107045042043006097108094118000120076029124042090023016077064099082127116079115074027079048119094006001029097087000098125100120017005103111073067071017023066117090100070&EXT=pdf&INDEX=TRUE> [Accessed 21 July 2021].

Morgan, B., 2019. *Does It Still Cost 5x More To Create A New Customer Than Retain An Old One?*. [online] Forbes. Available at: <https://www.forbes.com/sites/blakemorgan/2019/04/29/does-it-still-cost-5x-more-to-create-a-new-customer-than-retain-an-old-one/?sh=3a088cc03516> [Accessed 3 July 2021].

Nicita, C., 2019. *Council Post: Purpose In Partnerships: Adding Value With A Customer-First Mindset*. [online] Forbes. Available at:

<https://www.forbes.com/sites/forbesagencycouncil/2019/09/18/purpose-in-partnerships-adding-value-with-a-customer-first-mindset/?sh=2e28dea5678a> [Accessed 23 July 2021].

Saleh, K., 2020. *Customer Acquisition Vs.Retention Costs – Statistics And Trends*. [online] Invesp. Available at: <https://www.invespcro.com/blog/customer-acquisition-retention/> [Accessed 3 July 2021].

Tory, S., 2019. *Inside Recency, Frequency, Monetary Value (RFM)*. [online] Investopedia. Available at: <https://www.investopedia.com/terms/r/rfm-recency-frequency-monetary-value.asp> [Accessed 21 July 2021].

Wardley, D., Cable, D. and Steinhage, A., 2017. *The Pros and Cons of Competition Among Employees*. [online] Harvard Business Review. Available at: <https://hbr.org/2017/03/the-pros-and-cons-of-competition-among-employees#:~:text=Some%20research%20studies%20suggest%20such,effort%20and%20enables%20higher%20performance.> [Accessed 11 July 2021].

Appendix

Figure 1: Multiple Linear Regression of Sat_Program against Sat_Petrol, Sat_Grocery and Sat_FastFood

lm(formula = Sat_Program ~ Sat_Petrol + Sat_Grocery + Sat_FastFood, data = customer)
Residuals:

Min	1Q	Median	3Q	Max
-6.22	-0.74	0.05	0.81	4.73

Coefficients:

	Estimate	Std. Error	t-value	p-value	significance
(Intercept)	**-2.07**	**0.26**	**-8.05**	**1.37e-15**	***
Sat_Petrol	**0.73**	**0.03**	**27.89**	**< 2e-16**	***
Sat_Grocery	**0.25**	**0.02**	**12.84**	**< 2e-16**	***
Sat_FastFood	**0.25**	**0.02**	**13.98**	**< 2e-16**	***

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Goodness of fit:

Multiple R^2	Adjusted R^2	F-statistic	DF	p-value
0.43	0.43	502.8	1991	< 2.2e-16

Figure 2: ANOVA test results of Sat_Program vs 3 merchants

Analysis of Variance Table
Response: Sat_Program

	Df	Sum Sq	Mean Sq	F-value	p-value	Significance
Sat_Petrol	1	1806.06	1806.06	1193.70	< 2.2e-16	***
Sat_Grocery	1	180.08	180.08	119.02	< 2.2e-16	***
Sat_FastFood	1	295.86	295.86	195.55	< 2.2e-16	***
Residuals	1991	3012.37	1.51			

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 3: Multiple Linear Regression of Sat_Program against Race and Gender

lm(formula = Sat_Program ~ factor(Gender) + factor(Race), data = customer)
Residuals:

Min	1Q	Median	3Q	Max
-6.37	-1.20	-0.12	0.88	2.89

Coefficients:

	Estimate	Std. Error	t-value	p-value	significance
(Intercept)	**7.20**	**0.12**	**59.24**	**< 2e-16**	***
Gender	**0.01**	**0.07**	**0.20**	**0.85**	
Race1	**0.16**	**0.13**	**1.25**	**0.21**	
Race2	**-0.09**	**0.13**	**-0.66**	**0.51**	

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Goodness of fit:

Multiple R^2	Adjusted R^2	F-statistic	DF	p-value
0.00496	0.00346	3.30	1991	0.01941

Figure 4: Multiple Linear Regression of Sat_Petrol against Sat_Grocery and Sat_FastFood

lm(formula = Sat_Petrol ~ Sat_Grocery + Sat_FastFood, data = customer)
Residuals:

Min	1Q	Median	3Q	Max
-5.46	-0.74	0.01	0.82	3.82

Coefficients:

	Estimate	Std. Error	t-value	p-value	significance
(Intercept)	**6.51**	**0.17**	**39.20**	**< 2e-16**	***
Sat_Grocery	**0.30**	**0.02**	**19.62**	**< 2e-16**	***
Sast_FastFood	**-0.06**	**0.02**	**-4.13**	**3.87e-05**	***

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Goodness of fit:

Multiple R^2	Adjusted R^2	F-statistic	DF	p-value
0.19	0.18	226.4	1992	< 2.2e-16

Figure 5: ANOVA results for Sat_Petrol against city variables

Analysis of Variance Table
Response: Sat_Petrol

	Df	Sum Sq	Mean Sq	F-value	p-value	Significance
CityA	1	50.78	50.78	38.35	7.16e-10	***
CityB	1	14.14	14.14	10.68	0.001101	**

CityC	1	20.90	20.90	15.79	7.34e-05	***
CityD	1	1.89	1.89	1.43	0.23	
CityE	1	13.13	13.13	9.92	0.001662	**
CityF	1	0.94	0.94	0.71	0.40	

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 6: Binary Logistic Regression for customers active in 2016

lm(formula = Act_16 ~ Sat_Program + Sat_FastFood + Sat_Petrol + Sat_Grocery + NetPromoter + Gender_F + Race1 + Race2 + Car + CCard + CityA + CityB + CityC + CityD + CityE + CityF, family = "binomial", data = customer)

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.01	0.00015	0.37	0.70	2.16

Coefficients:

	Estimate	Std. Error	z-value	p-value	significance
(Intercept)	**6.12**	**1694.55**	**0.004**	**0.9971**	
BirthYear	**0.0028**	**0.0084**	**0.33**	**0.74**	
Sat_Program	**-0.06**	**0.05**	**-1.13**	**0.26**	
Sast_FastFood	**0.38**	**0.05**	**7.20**	**6.18e-13**	***
Sat_Petrol	**0.71**	**0.08**	**9.30**	**< 2e-16**	***
Sat_Grocery	**-0.02**	**0.05**	**-0.34**	**0.73**	
NetPromoter	**0.27**	**0.04**	**7.44**	**9.96e-14**	***
Gender_F	**-0.63**	**0.16**	**-4.00**	**6.22e-05**	***
Race1	**0.46**	**0.20**	**2.29**	**0.02**	*
Race2	**0.25**	**0.21**	**1.23**	**0.22**	
Car	**0.33**	**0.13**	**2.48**	**0.01**	*
CCard	**0.32**	**0.13**	**2.41**	**0.02**	*
CityA	**-0.10**	**0.16**	**-0.61**	**0.54**	
CityB	**-0.06**	**0.23**	**-0.27**	**0.78**	
CityC	**-0.36**	**0.22**	**-1.66**	**0.10**	
CityD	**0.53**	**0.24**	**2.20**	**0.03**	*
CityE	**-0.06**	**0.26**	**-0.25**	**0.80**	
CityF	**0.18**	**0.27**	**0.69**	**0.49**	

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

Goodness of fit:

llh	llhNull	G2	McFadden	r2ML	r2CU
-857.72	-1118.83	522.22	0.23	0.23	0.34

Figure 7: Correlation of Sat_Program against 3 merchants

Independent Variables	Sat_Program
Sat_Petrol	0.5840614
Sat_FastFood	0.1065004
Sat_Grocery	0.4137724

Figure 8: Multiple Linear Regression of Sat_Petrol against city variables

lm(formula = Sat_Petrol ~ CityA + CityB + CityC + CityD + CityE + CityF, data = customer)
Residuals:

Min	1Q	Median	3Q	Max
-5.19	-0.81	-0.10	0.90	2.19

Coefficients:

	Estimate	Std. Error	z-value	p-value	significance
(Intercept)	**8.10**	**0.05**	**172.92**	**< 2e-16**	***
CityA	**-0.29**	**0.07**	**-4.18**	**3.04e-05**	***
CityB	**0.31**	**0.09**	**3.43**	**0.00062**	***
CityC	**-0.27**	**0.09**	**-2.80**	**0.00518**	**
CityD	**0.18**	**0.10**	**1.89**	**0.05882**	
CityE	**0.34**	**0.11**	**3.25**	**0.00117**	**
CityF	**0.09**	**0.11**	**0.85**	**0.39846**	

Signif. Codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ‘ 1
Goodness of fit:

Multiple R^2	Adjusted R^2	F-statistic	DF	p-value
0.037	0.034	12.81	1988	3.02e-14

Figure 9: Recency, Frequency, Monetary (RFM) Scores of Customers

R(Recency)	R-Score
January-March, 2015(1-3)	1
April-June, 2015(4-6)	2
July-September, 2015(7-9)	3
October-November, 2015(10-11)	4
December, 2015(12)	5

F (Frequency)	F-Score
0-25	1
26-50	2
51-75	3
76-100	4
>100	5

M (Money)	M-Score
0-100	1
101-200	2
200-1000	3
1000-2000	4
>2000	5

Figure 10: Descriptive Statistics of RFM scores

	R-Score	F-Score	M-Score	FRM-Score
MIN	1	1	1	3
1st Qu.	3	1	1	6
MEAN	3.939348371	1.463157895	2.362406015	7.764912281
MEDIAN	4	1	3	8
3rd Qu.	5	2	3	10
MAX	5	5	5	15

Figure 11: Customers with the highest RFM scores (15 points)

	A	B	C	D	E	F	G	H	I	
1	MemberID	Total SalesAmt	Last day	Purchase Times	Last Purchase MONTH	R(Recency)	F (Frequency)	M (Money)	Total RFMScore	R
2	10172519455	28243.06	2015/12/30	243	12	5	5	5	15	
3	10173159205	10808.55	2015/12/28	139	12	5	5	5	15	
4	10173322855	10705.75	2015/12/23	139	12	5	5	5	15	
5	10173998455	23642.95	2015/12/28	378	12	5	5	5	15	
6	10176477805	3975.32	2015/12/31	107	12	5	5	5	15	
7	11249539555	2350.16	2015/12/24	108	12	5	5	5	15	
8	11250732655	2041.87	2015/12/31	117	12	5	5	5	15	
9	11250866155	7394.81	2015/12/31	220	12	5	5	5	15	
10	11253707905	5130.53	2015/12/29	149	12	5	5	5	15	
11	11254145605	3188.02	2015/12/30	142	12	5	5	5	15	
12	11263879405	4397.96	2015/12/28	114	12	5	5	5	15	
13	11267456755	2658.84	2015/12/31	230	12	5	5	5	15	
14	11277927355	2546.66	2015/12/31	216	12	5	5	5	15	
15	11278394155	28300	2015/12/31	283	12	5	5	5	15	
16	11282767105	3778.15	2015/12/31	275	12	5	5	5	15	
17	11286420355	4402.7	2015/12/31	193	12	5	5	5	15	
18	11299664605	3219.52	2015/12/25	132	12	5	5	5	15	
19	11313296605	20874.18	2015/12/30	506	12	5	5	5	15	
20	11752046455	2376.62	2015/12/31	224	12	5	5	5	15	
21	11752369255	12663.93	2015/12/30	136	12	5	5	5	15	
22	11755510705	3766.28	2015/12/26	134	12	5	5	5	15	
23	11761623205	3897.31	2015/12/29	144	12	5	5	5	15	
24	10008415255	2270.04	2015/11/12	113	11	4	5	5	14	

Figure 12:Descriptive Statistics of Customer Lifetime Value (CLV)

	MIN	1st Qu.	MEAN	MEDIAN	3rd Qu.	MAX
CLV	3.54	175.66	1650.01	589.23	1856.62	81362.5

Figure 13:CLV customer score histogram

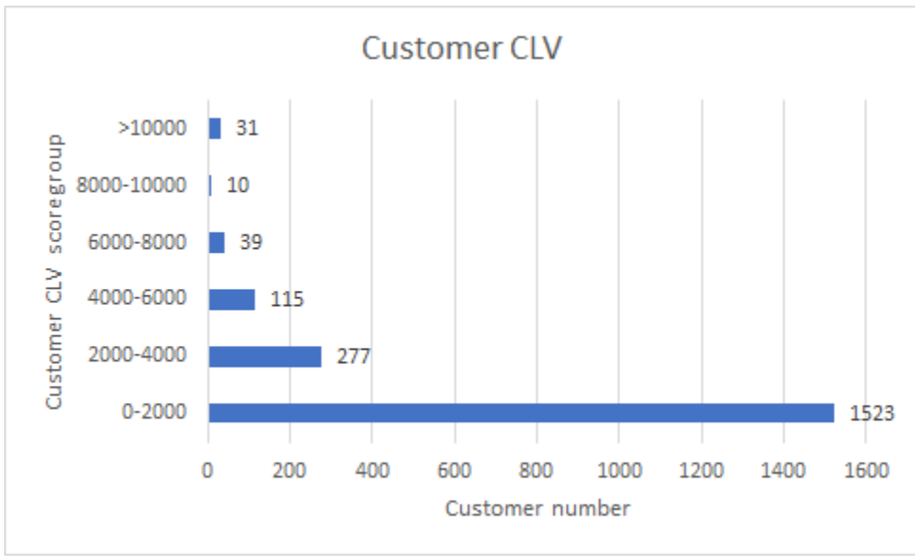


Figure 14 :The list of 31 customer with score more than 10000

	A	B	C	D	E	F	G
1	Curtomer ID	Revenue	Cost	Acquisition Co	Discount rate	Retention rate	CLV
2	11278394155	28300	0	0	0.15	0.75	81362.5
3	10172519455	28243.06	0	0	0.15	0.75	81198.7975
4	10173998455	23642.95	0	0	0.15	0.75	67973.48125
5	11313296605	20874.18	0	0	0.15	0.75	60013.2675
6	11311168555	15509.9	0	0	0.15	0.75	44590.9625
7	11752369255	12663.93	0	0	0.15	0.75	36408.79875
8	11805149155	11486.32	0	0	0.15	0.75	33023.17
9	10173159205	10808.55	0	0	0.15	0.75	31074.58125
10	10173322855	10705.75	0	0	0.15	0.75	30779.03125
11	10117037005	9727.12	0	0	0.15	0.75	27965.47
12	11732353405	9125	0	0	0.15	0.75	26234.375
13	10172617555	8828.21	0	0	0.15	0.75	25381.10375
14	10172621005	7873.14	0	0	0.15	0.75	22635.2775
15	11250866155	7394.81	0	0	0.15	0.75	21260.07875
16	11264032855	6600.48	0	0	0.15	0.75	18976.38
17	11253707905	5130.53	0	0	0.15	0.75	14750.27375
18	11754332005	4673.99	0	0	0.15	0.75	13437.72125
19	10054307455	4548.11	0	0	0.15	0.75	13075.81625
20	11286420355	4402.7	0	0	0.15	0.75	12657.7625
21	11263879405	4397.96	0	0	0.15	0.75	12644.135
22	10153390405	4138.2	0	0	0.15	0.75	11897.325
23	10173325555	3977.03	0	0	0.15	0.75	11433.96125
24	10176477805	3975.32	0	0	0.15	0.75	11429.045
25	11761623205	3897.31	0	0	0.15	0.75	11204.76625
26	11250036355	3871.71	0	0	0.15	0.75	11131.16625
27	11282767105	3778.15	0	0	0.15	0.75	10862.18125
28	11755510705	3766.28	0	0	0.15	0.75	10828.055
29	11311486105	3732.45	0	0	0.15	0.75	10730.79375
30	11755996855	3732.19	0	0	0.15	0.75	10730.04625
31	10141493455	3713.75	0	0	0.15	0.75	10677.03125
32	11263394455	3497.17	0	0	0.15	0.75	10054.36375

Figure 15: Cluster Dendrogram of FGP Program customer data

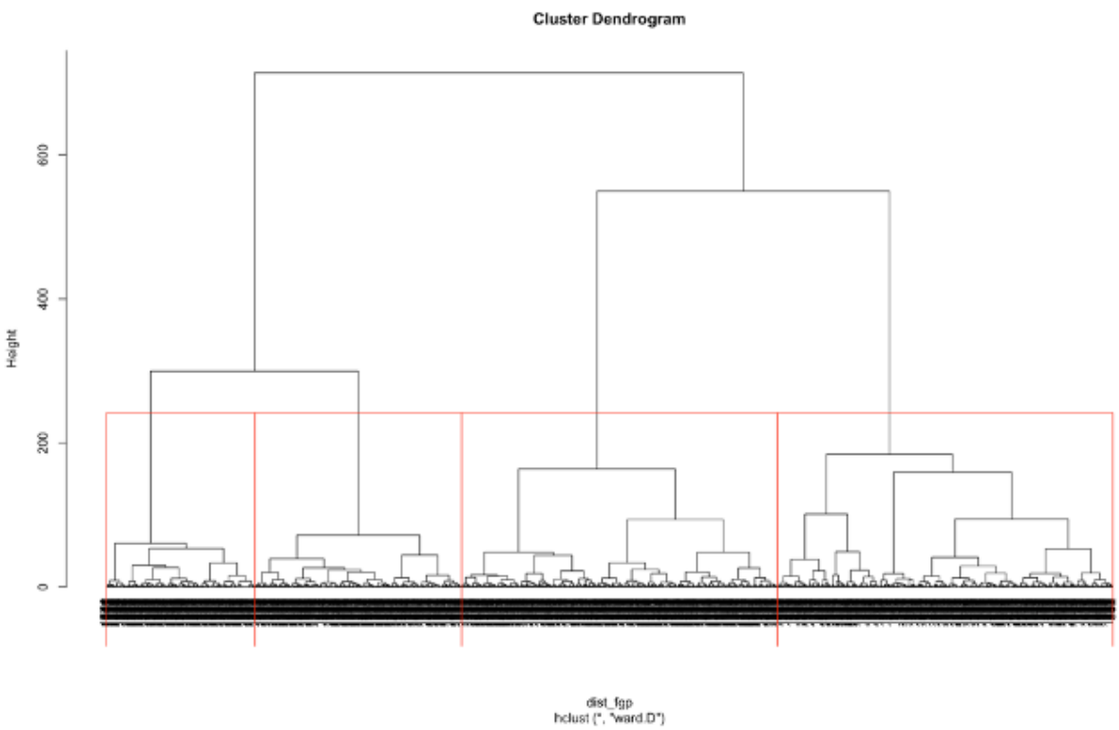


Figure 16: Cluster Analysis of 4 segments containing member ID and average birth year

Row Labels	Count of MemberID	Average of BirthYear
Segment 1	665	1983.892802
Segment 2	626	1976.42404
Segment 3	410	1978.345588
Segment 4	294	1973.453608
Grand Total	**1995**	**1978.882624**

Figure 17: Cluster Analysis of 4 segments containing sum of city variables

Row Labels	Sum of CityA	Sum of CityB	Sum of CityC	Sum of CityD	Sum of CityE	Sum of CityF
Segment 1	236	55	84	47	28	29
Segment 2	138	60	58	57	51	46
Segment 3	104	60	37	50	31	26
Segment 4	36	40	16	33	38	31
Grand Total	**514**	**215**	**195**	**187**	**148**	**132**

Figure 18: Cluster Analysis of 4 segments containing average of satisfaction programs, the 3 merchant chains and net promoter

Row Labels	Average of Sat_Program	Average of Sat_FastFood	Average of Sat_Petrol	Average of Sat_Grocery	Average of NetPromoter
Segment 1	7.257142857	8.70075188	7.479699248	6.102255639	7.142857143
Segment 2	5.924920128	6.07028754	7.53514377	6.276357827	3.591054313
Segment 3	7.843902439	6.987804878	8.631707317	7.317073171	5.648780488
Segment 4	9.31292517	6.476190476	9.816326531	8.5	8.629251701
Grand Total	**7.262656642**	**7.195488722**	**8.078195489**	**6.759899749**	**5.940350877**

Figure 19: Correlation test between NetPromoter and Sat_Program, Sat_FastFood, Sat_Petrol and Sat_Grocery

	NetPromoter	*Sat_Program*	*Sat_FastFood*	*Sat_Petrol*	*Sat_Grocery*
NetPromoter	1				
Sat_Program	0.5027966	1			
Sat_FastFood	0.40110266	0.10650043	1		
Sat_Petrol	0.38601585	0.5840614	-0.1667654	1	
Sat_Grocery	0.20473074	0.41377237	-0.201464	0.42219268	1

Figure 20: Multiple Linear Regression of NetPromoter against Sat_FastFood, Sat_Petrol and Sat_Grocery as the independent variables

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-9.101739	0.43657124	-20.848233	1.7945E-87	-9.9579233	-8.2455546	-9.9579233	-8.2455546
Sat_FastFood	0.83008416	0.03012665	27.5531508	7.47E-142	0.77100109	0.88916723	0.77100109	0.88916723
Sat_Petrol	0.94067542	0.04423322	21.2662649	1.3194E-90	0.85392716	1.02742368	0.85392716	1.02742368
Sat_Grocery	0.21749859	0.03278716	6.63365171	4.2069E-11	0.15319785	0.28179933	0.15319785	0.28179933