**FGP Customer**

**Loyalty Program**



**MARK3054**

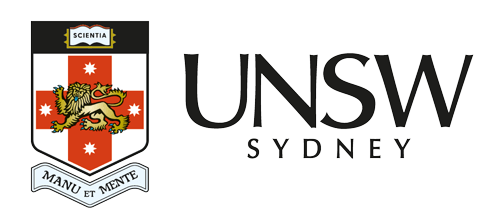
**Final Report**

Created by:

**Ashleigh Smith (z5160534)**

**Ethan Whitehouse (z5160339)**

**Olivya Precious (Z5286497)**



**TABLE OF CONTENTS**

[**EXECUTIVE SUMMARY**](#_fvzkztrupq6e) **3**

[**INTRODUCTION**](#_6iua4oyq143v) **4**

[**III MANAGERIAL PROBLEM**](#_225459vo5rpb) **4**

[**IV RESEARCH QUESTIONS**](#_1l1x7142d3m6) **5**

[**V METHODOLOGY**](#_hro5kzbwnaxj) **5**

[**VI RESULTS AND FINDINGS**](#_c69bo6myjtj) **7**

[Research Question 1](#_oxrz0ivs0l3t)

[Research Question 2](#_o3la2jr24i5)

[Research Question](#_j7r3lkmqkt83) 3

[**VII DEFICIENCIES**](#_cvq0occ1caru) **13**

[**VIII APPENDIX**](#_p7ksz66o13x7) **16**

# EXECUTIVE SUMMARY

The FGP customer loyalty program consists of three merchant members; a fast food chain, grocery store chain and petrol store chain which operate in Asia. To enable the expansion of the FGP loyalty program, a thorough analysis of its current customer base is required to better understand the market and inform its future marketing strategies. A survey was conducted by the FGP loyalty program manager, consisting of 1,995 customers within the FGP loyalty program. These responses accompanied a dataset of 43,805 purchase records and 649 redeem records gathered from the 2015 fiscal year. This dataset was analysed and segmented through a number of methods and statistical tests, particularly regression and RFM analysis to uncover several key findings.

Six key segments of customers were identified and profiled by the team which found the most valuable customer to be recognised as segment 6, aptly named FGP Champions. Descriptive analysis of this segment found this cluster to mainly consist of males who own a credit card, are middle aged and have a high level of customer satisfaction across the program and merchants. A series of regression further revealed customer value, as determined by total RFM score, to be influenced by gender, credit card ownership and satisfaction for petrol merchants. The results of this regression were aligned with the FGP Champion profile, confirming this demographic should be targeted by the FGP loyalty program in its expansion.

Our findings into customer purchases and the redemption of points found points were unused by a significant number of customers along with an imbalance in redemption and sales quantities in each of the chains. Further promotion into the value of the FPG loyalty program to its customers is therefore recommended. In determining customer churn, customers from City A were found to be insignificant in determining the likelihood that a customer would continue with business. A customer’s RFM score was also deemed insignificant in determining customer churn.

Despite the large dataset, the data provided by the FGP loyalty program manager only included data for 2015, meaning deeper analysis over a longer period of time was unable to be conducted. This limits our ability to recommend sustainable marketing strategies as the team isn’t able to understand the longitudinal dynamics of customers with the FGP loyalty program.

# INTRODUCTION

This report will focus on the FGP customer loyalty program which consists of three merchant members; a fast food chain, grocery store chain and petrol store chain operating in Asia. To assist with the potential expansion of the FGP loyalty program, our team has conducted a study into customer value and predictors of customer behaviour to best manage current and future customers. Thus, the team has undertaken extensive analysis and testing of variables within the FGP loyalty program dataset, consisting of 1,995 customers and 43,805 purchase records and 649 redeem records gathered from the 2015 fiscal year. Based on the results of our analysis, the team has correspondingly recommended targeted marketing strategies.

# III MANAGERIAL PROBLEM

The FGP loyalty program manager seeks to expand the FGP loyalty program to attract more customers and additional merchant members. However, the manager currently lacks a comprehensive understanding of the program’s current customers and associated behaviours. This information is key to achieving their core marketing and overall business goals as these insights allow the business to understand customer needs and optimise the customer experience accordingly. Specifically, these insights can be utilised to promote the value of the FGP loyalty program to potential and existing customers, in a targeted and effective manner. To achieve this, segmentation of customers on an individual level is vital to understand customer heterogeneity, driving factors relating to customer behaviour. and ultimately predict potential future behaviour of customers within each customer segment.

This research project will explore the overarching managerial question:

*How can we estimate the value of a customer in the loyalty program, predict if a customer will churn and manage the customers accordingly?*

# IV RESEARCH QUESTIONS

This research project will address three key research questions to best address and provide meaningful insights into the managerial problem.

1. Who are the most valuable customers within the FGP loyalty program? What are the characteristics of a valuable customer?
2. What is the churn rate per segment and is there an indication that the customer will churn?
3. What is the relationship between customers and purchases made within the loyalty program or redemption of points?

# V METHODOLOGY

**Recency, frequency, monetary value (RFM) Analysis:** This marketing analysis tool was used to segment the customers based on historical purchase behaviours, and identify the FGP loyalty program’s best customers, and therefore who are considered most valuable.

For the purposes of the RFM model, recency, frequency, and monetary value measures were defined as:

* Recency: the number of days since the last purchase made by the member
* Frequency: the number of purchases made by the member in 2015
* Monetary Value: the total sales amount made by the member during 2015

The customers were sorted based on each RFM attribute, and assigned a score from 0-9, with 0 being the least and 9 being the best score. The scores were allocated based on customer percentile with the top 10% receiving a score of 9, the next 10% a score of 8, and so on, with the bottom 10% receiving a score 0.

The three separate RFM scores as well as the aggregated RFM score (calculated by combining the R, F, and M ranking) were used as a primary data point to segment and analyse customers within all four research questions.

**Cluster Analysis:**  This method was used to categorise individuals R, F and M scores to identify relatively homogeneous groups, based on the degree of (dis)similarity in the data set. Specifically the team engaged a PDA Hierarchical Cluster Analysis utilising euclidean distance and Ward's minimum variance method. Further analysis was conducted into the optimal N of clusters provided it satisfied the 2 < N < 10 to ensure a manageable and meaningful number of clusters were identified and profiled.

**Descriptive analysis:** This method was utilised to understand the demographics of the FGP customers within the dataset and profiled clustered and to determine any trends or patterns that may exist amongst the categorical variables. Specifically measures of central tendency, mainly mean, and proportionality were to identify the ‘majority’ within groups.

**Inferential analysis:** T-Tests and ANOVA was used to infer relationships and observed differences between demographic variables and customer segments. Independent T- Tests were used to compare groups with one other based upon the aggregated mean of linear demographic data and customer purchase behaviour from the dataset and further analysis. Various types of ANOVA were also applied to determine the mean of scale attribute within several groups, mainly relating to customer segments and indicators of churn.

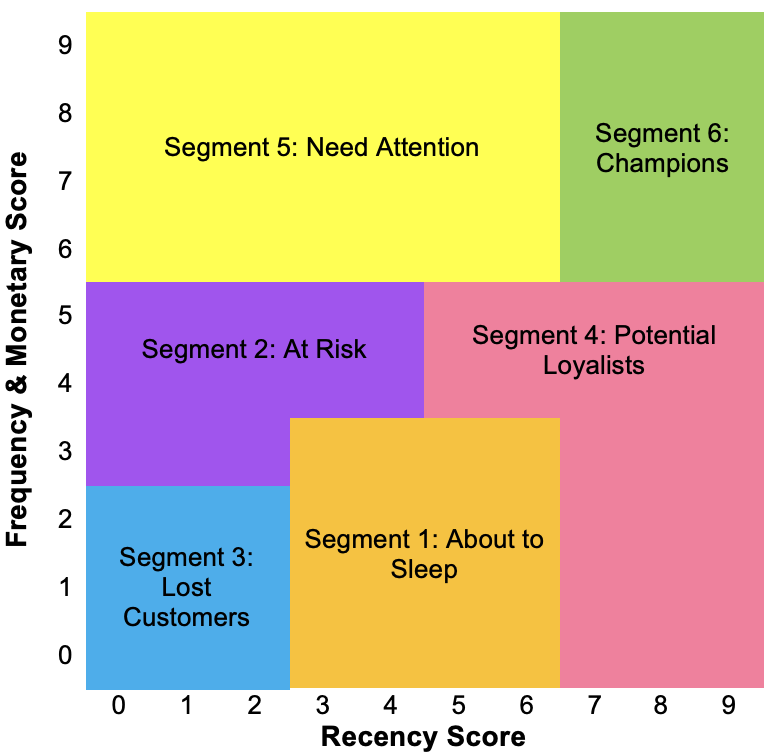
**Regression:** A series of regression analysis were conducted to make meaningful predictions about future customers and purchase behaviour to inform segmentation, targeting and positioning of the FGP loyalty program and cater to the needs of their target, and most valuable, consumers.

* *Multiple Linear Regression:* Used to analyse any significantly correlated variables and to then determine whether the dependent variable can be explained by the independent variable and whether this relationship was significant and strong
* *Binary Regression:* Multiple variables were dummy coded to utilise binary regression and analyse the relationship between categorical variables on a greater extent. This method was particularly utilised in the team’s approach to research question 1 and 2.

# VI RESULTS AND FINDINGS

## Research Question 1: Who are the most valuable customers?

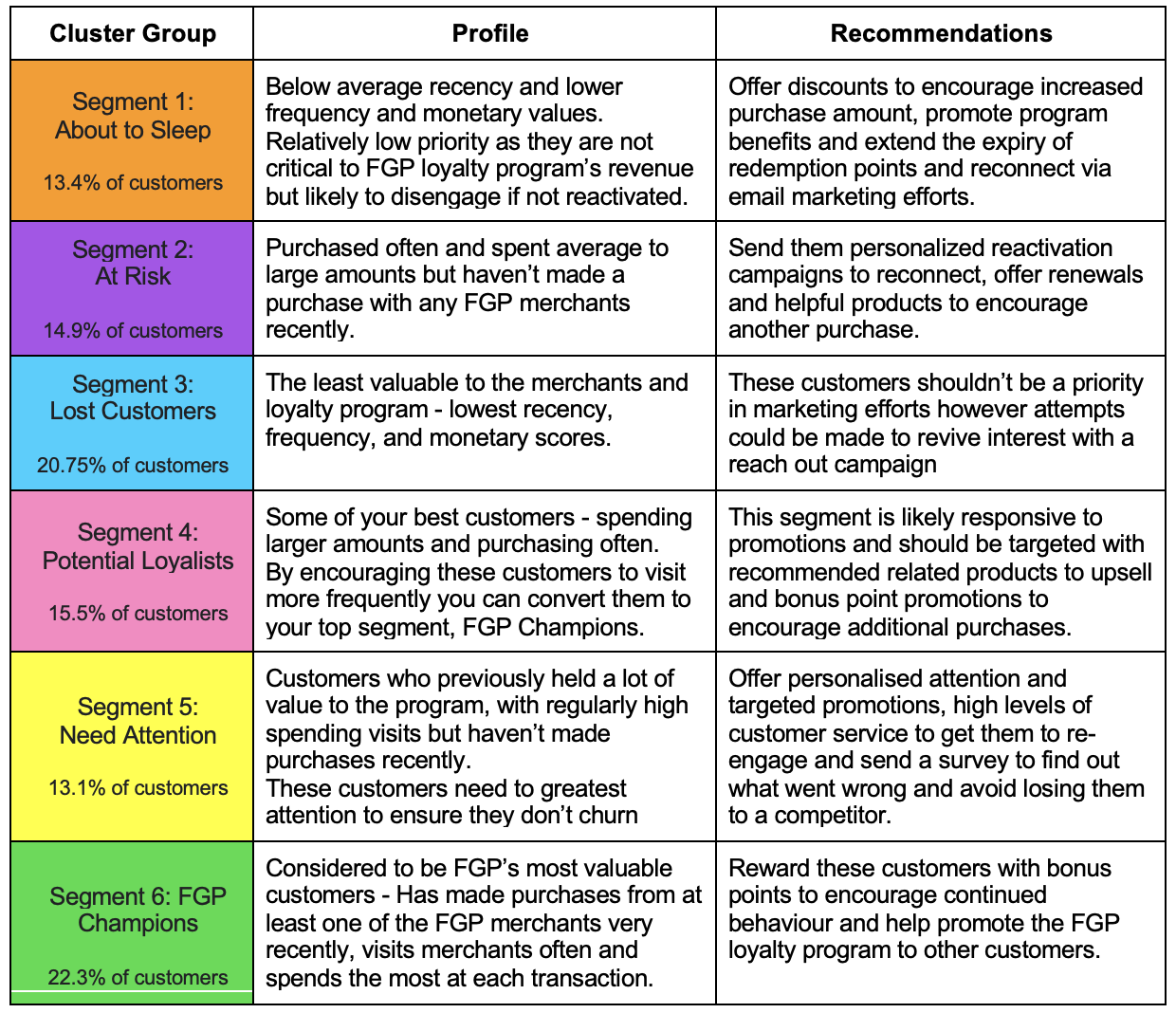
The data obtained from the FGP customer data and purchase records were used to assess each customer’s value within the FGP loyalty program. Furthermore, to address the relationship between recency, frequency, and monetary rankings, the team proceeded to utilise PDA hierarchical cluster analysis to segment the customers into homogeneous groups. In doing so, the most valuable customers can be differentiated as well as identifying other critical customer segments, based on their purchase behaviour. Specifically, an optimal number of clusters were identified with the results of the PDA hierarchical cluster generating six distinct segments of customers within the FGP loyalty program. Based on RFM scores, a visualisation of the predicted segments was generated, as shown in Figure 1.



***Figure 1: RFM Predictive Segments for the FGP Customer Loyalty Program***

The central tendencies of the RFM scores of each segment were analysed and subsequently profiled (shown in Appendix A and B). These findings further allowed the team to generate personalised marketing strategies to maximise value to the program, summarised in Figure 2.

***Figure 2: Profiles of FGP Customer Loyalty Program Segments + Relevant Recommendations***



To gain an understanding of the differences between these groups, proportional analysis of categorical attributes was conducted and the central tendency, specifically mean, of scale variables was determined (Appendix B). Based on this, the most valuable segment of customers, i.e. Segment 6: FGP Champions were majority males (58%), own a credit card (55%) and the ages of customers are skewed older, when compared to other clusters, with the majority falling between 30-49 age groups. FGP Champions further have an almost equal split between those who do and don’t own a car. This split is significant as all other segments show a larger majority of customers who don’t own a car. These customers have an extremely low churn rate based on 2015 to 2016 data with only 1% of customers failing to engage. The level of satisfaction for the program (8.57), petrol merchant (9.33) and grocery merchant (7.71) are the highest of all segments, with fast food satisfaction for this group ranking average (7.05). The NPS score for this cluster is furthermore the highest at 7.82, suggesting this group is likely to advocate for the brand and recommend the FGP loyalty program to friends and family.

To further clarify the relationship between RFM scores and customer variables and attributes, we conducted a series of multiple linear regression tests, with the dependent variable being total RFM score. Several attempts to remove insignificant results of individual variables within the regression were made by the team to try and generate a completely statistically significant result and greater model fit. The most significant (p-value = 0) and strongest linear relationship (Adj R Square = 0.58) was defined with the following equation:

**Total RFM Score** = -26.92 - 1.17 x *Gender* + 0.52 x *OwnCreditCard* + 0.06 x *Age* + 4.88 x *Active2016* + 0.77 x *Sat\_Program* + 0.41 x *Sat\_FastFood* + 3.25 x *Sat\_Petrol + ɛ*

The coefficients of this regression (see Appendix C) predict the variable that impact RFM scores the most to be whether the customer is a male as well as customer activity in 2016 and overall satisfaction with the Petrol merchant. A positive relationship with credit card ownership and age is also demonstrated. These results are in alignment with the analysis of the FGP Champions segment. The adjusted R square figure in this regression forecasts that approx 60% of RFM scores can be explained by the regression line equation. Although the results indicate a relatively weak to medium strength, its relative significance in this data set of combining these specific demographic variables should be capitalised on in future marketing efforts.

Furthermore, results from an initial regression involving every demographic variable (see Appendix D) show an insignificant improvement in the adjusted R square value along with multiple variables with a p-value > 0.05 despite significance of the overall ANOVA test. The lack of difference in model fit between the two regressions may suggest the irrelevance of the variables; Days registered, grocery satisfaction, location, and race, in predicting customer value. Thus, based on regression and segmentation analysis, it’s recommended that middle-aged males who own a credit card be targeted by the FGP merchants and loyalty program.

## Research Question 2: What is the churn rate per segment and is there an indication that the customer will churn?

The FGP customer data was used in order to extrapolate information regarding the likelihood that a customer continued their business from 2015 to 2016, as represented by the variables ‘Active2015’ and ‘Active2016’. These figures revealed a fall in customer activity from 2015 to 2016, a phenomenon which was reflected by the fact that 1996 customers were active in 2015, compared to 1499 in 2016, a drop of 497. When computed this a reveals a churn rate of 24.9%

.

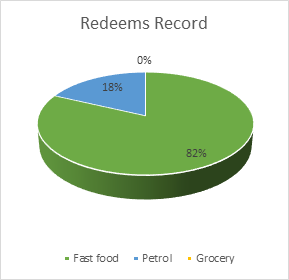
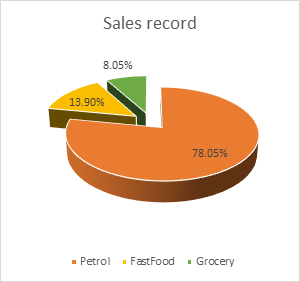
In order to understand what factors were significant in affecting a customers churn rate several statistical tests were undertaken. Chiefly, a binary logistic regression was run in order to determine the significance of RFM scores in affecting the likelihood that a customer continued their business in 2016 *(appendix H)*. The r-square value produced by this test, 0.237, was of low value, reflecting the fact that a customer's RFM score was not a useful predictor of continued business in 2016.

Furthermore ANOVA tests were run in order to determine the significance that certain customer attributes or demographic segments had upon the likelihood that they would churn in the following period *(appendix G)*. This included an ANOVA test which segmented customers activity in 2016 based upon the city which they were from. This test produced a p-value of 0.09 between groups, reflecting the fact that a customer's city was not of statistical relevance in determining the likelihood that they would continue their business.

## Research Question 3: What is the relationship between customers and purchases made within the loyalty program or redemption of points?

This question uses multiple linear regression to determine whether engagement and customer satisfaction have a significant relationship with customer loyalty (net-promoter) which has been proven using 95% confidence level. According to Bansal (2016), engagement helps to keep the customer remained associated and involved with the brand for a period time which can be done through online such as social media or offline such as billboard.

As shown in Appendix E, the resulting 0.462 R Square value indicates that the independent factors explain 46.2 percent of the variation in the dependent variable (net promoter). As the p is less than 0.05, it shows that all independent variables have a significant positive relationship with the customer loyalty. However, the coefficient of both Satisfaction Grocery and Purchase Number are less than 0.1 which means insignificant effect on customer loyalty. Furthermore, Fast Food Satisfaction is the biggest contributor from all independent variables.

In this report, our team presumes that the relationship between customer and the chains are strong even though there is no loyalty program in the first place. The Appendix E shows the sales amount and redeem amount that was being summarized.

The pie chart showed that 78.05% of total sales came from Petrol, followed by FastFood and Grocery which were 13.90% and 8.05% respectively. In the redeems record, 82% of the records was FastFood followed by 18% at Petrol chain, but no records showed that customers used redeem for Grocery. These statistics demonstrate that customers continue to buy at Grocery despite the fact that they never use the redemption points. This result also proves that one chain is more significant than the other two chains.

Descriptive analysis of the point redemption data shows that only 20% of members in the FGP loyalty redeemed earned points at any merchant in 2015. This demonstrates the majority of loyalty members are not actively utilising the points systems and may suggest the program’s lack of marketing efforts to promote benefit usage and the value proposition.

To address this issue, the loyalty program manager might incorporate cost sharing among retailers to operate the loyalty program more effectively in the long-term

**VII DEFICIENCIES**

A key deficiency in the dataset was it’s limited scope as purchase and redeem records and the survey only included data for the 2015 fiscal year, i.e. 1 year’s worth of data. The lack of data collected over a greater period limits any longitudinal analysis and the ability to detect any changes or patterns in customer behaviour over an extended period of time. Therefore this data cannot fully represent the entirety of the customer’s relationship with the FGP loyalty program and its merchants. Particularly in determining the association between redemption and purchase, further data would be useful to gain a better interpretation.

In conducting the analysis based on the data points in the survey, there was a lack of strength in the results among most tests, particularly in regression testing. This limits the recommendation potential and further customer research by the FGP loyalty program manager should be conducted to gain additional data that allows for stronger relationships to emerge.

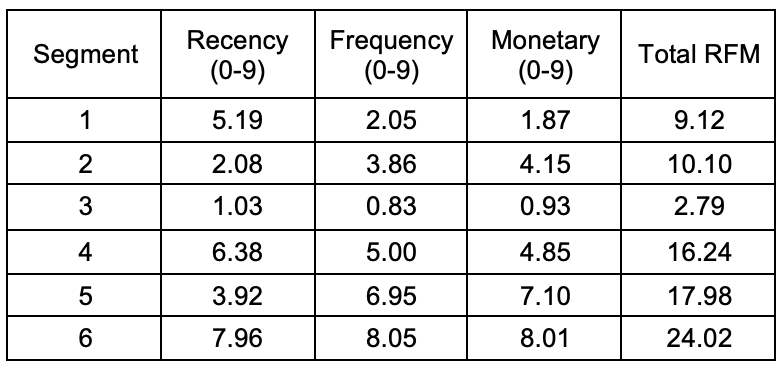
Another issue was the labeling of variables which meant insights into demographic variables involving city and race were unable to be turned into appropriate marketing recommendations beyond the superficial and non-meaningful labels. Correspondingly, there was also a lack of insights beyond the data provided into promotional activities of the FGP program. This is necessary to understand their relative impact on sales and redemption behaviour.

**VIII. REFERENCE**

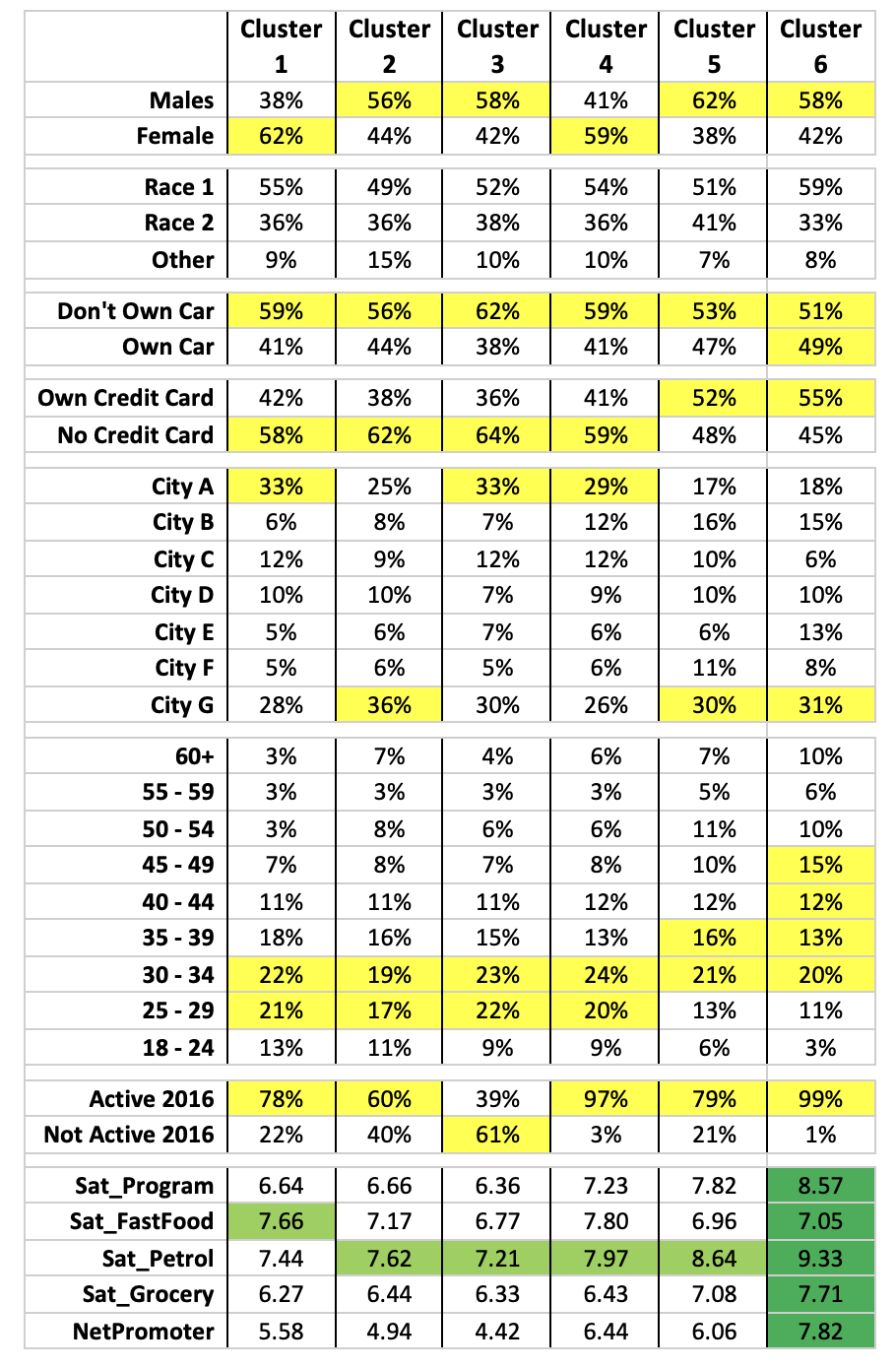
Bansal, Rohit. (2016). CUSTOMER ENGAGEMENT – A LITERATURE REVIEW. Global International Research Thoughts ISSN: 2347-8861. 2. 15-20.

# VIII APPENDIX

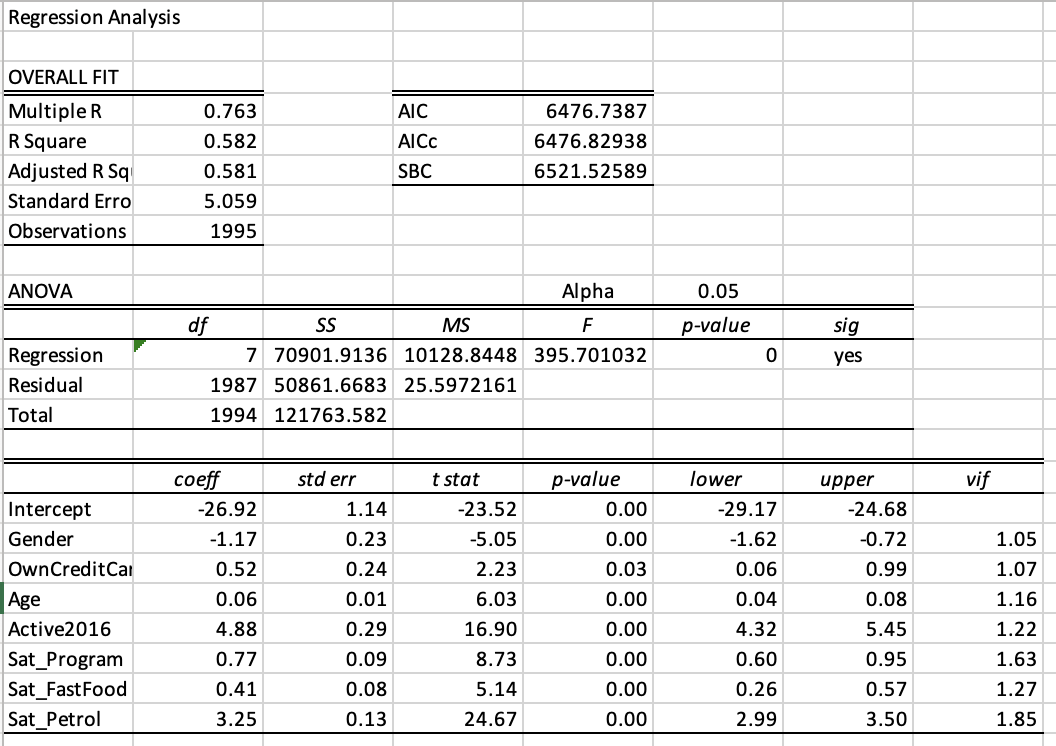
**Appendix A Mean RFM scores of FGP Loyalty Program customer segments**



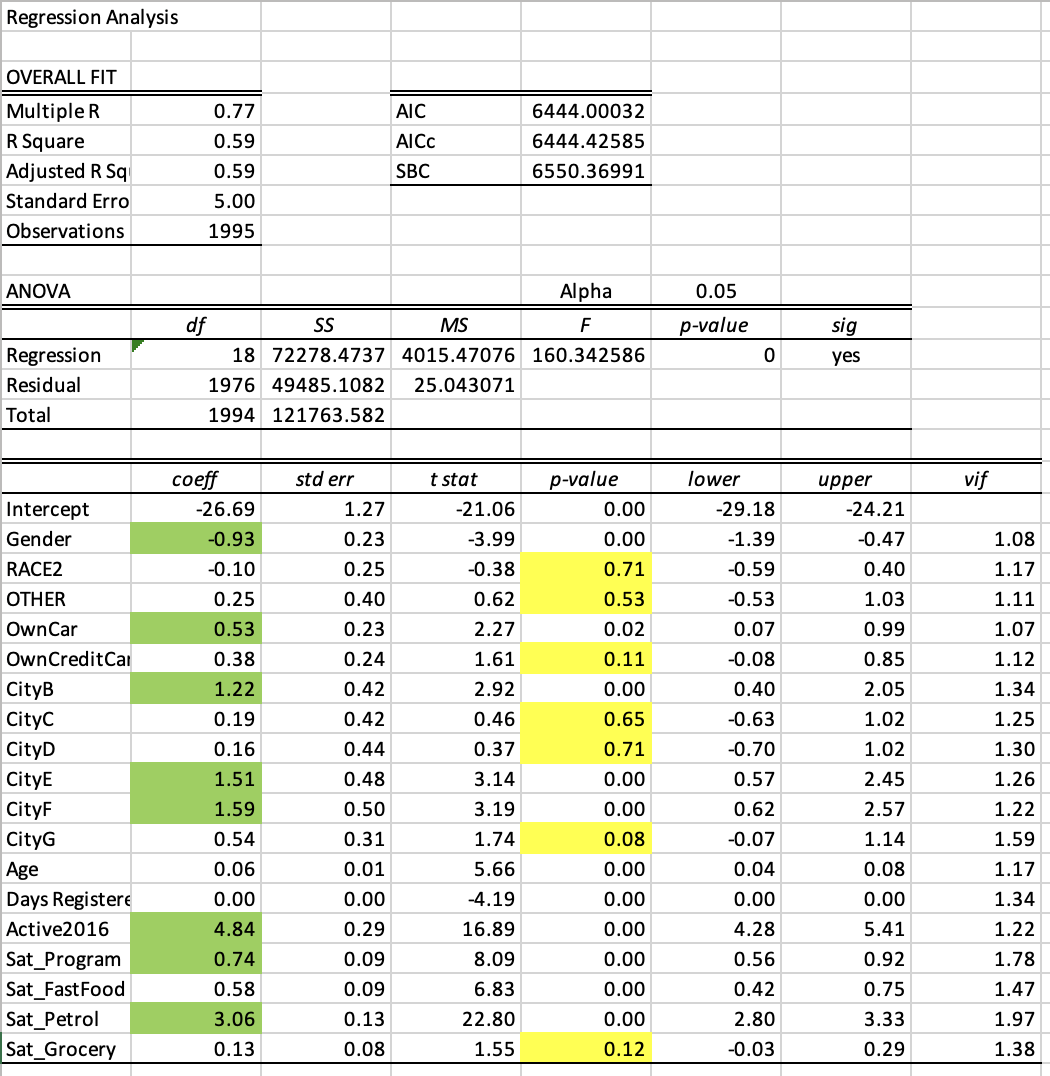
**Appendix B: Summary of Clusters - Descriptive Analysis of Customer Data and Variable**

​​

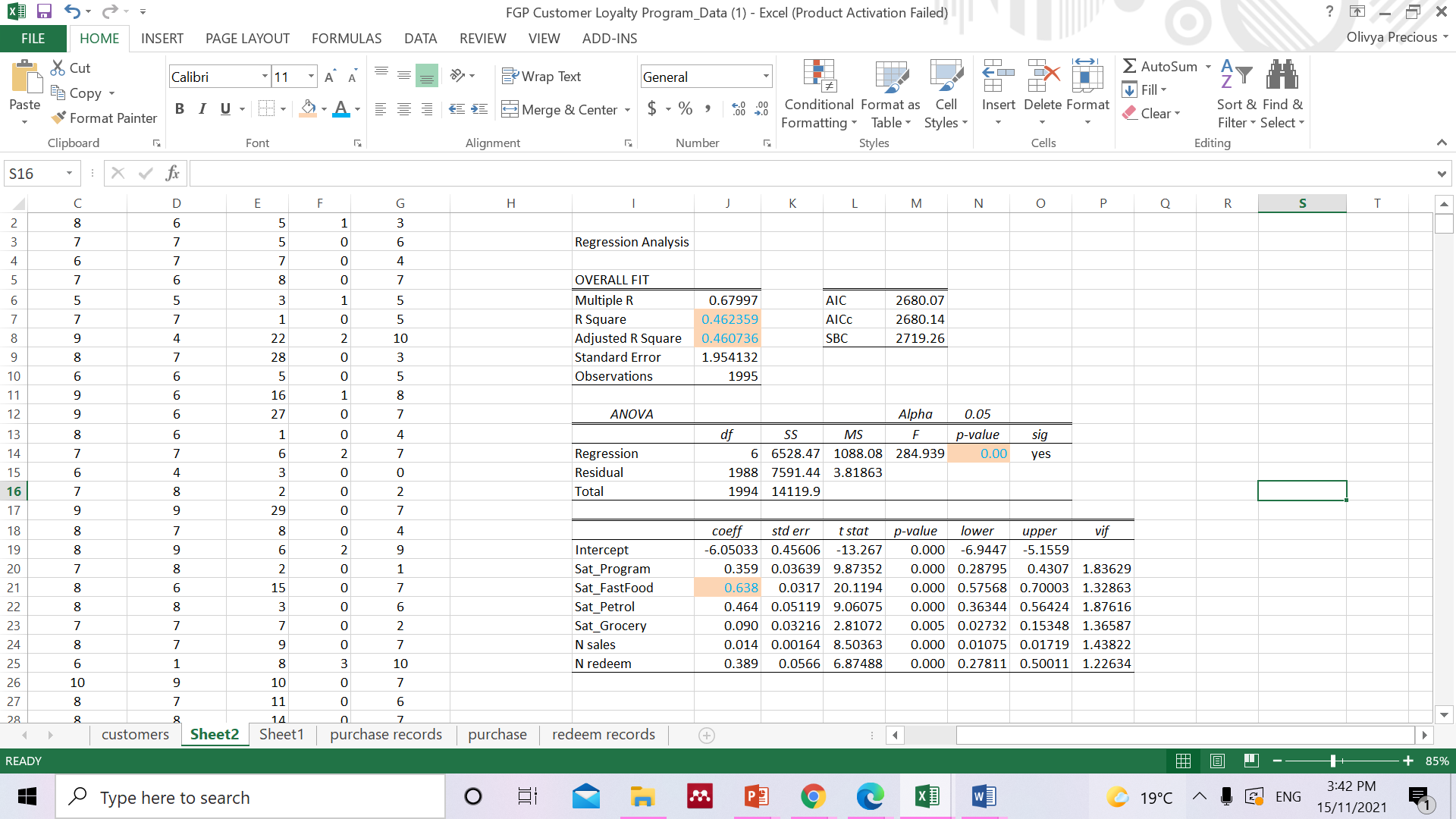
**Appendix C: Total RFM Regression Results by Gender, Credit Card, Age, Active 2016 and Customer Satisfaction (excl. Grocery)**

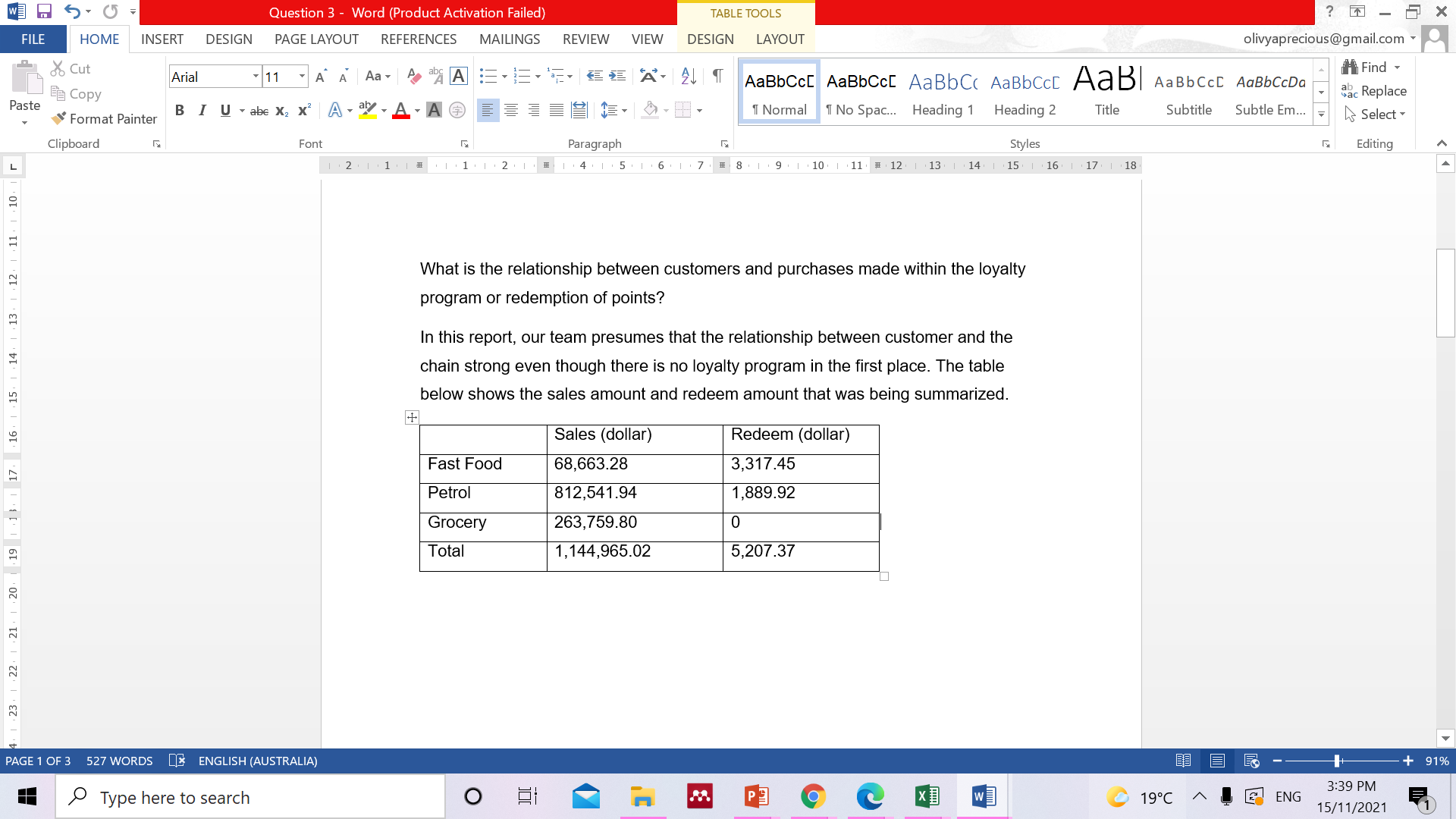


**Appendix D: Total RFM Regression Results with all Demographic Customer Variables in the Dataset**

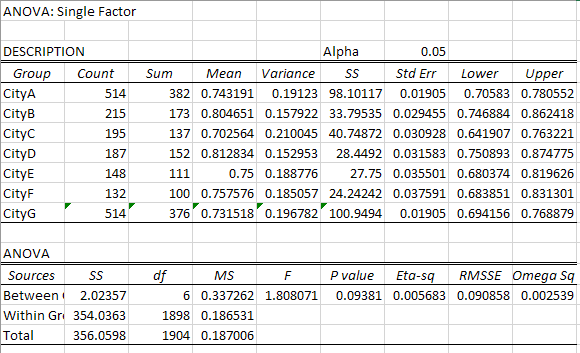


**Appendix E: Multiple Regression Analysis Between Satisfaction and Customer Loyalty**



**Appendix F: Imbalances in redemption and sales at each chain**

**Appendix G: ANOVA - City/Activity2016**



**Appendix H: Binary logistic regression - RFM/Activity2016**

