



Cardiff Metropolitan University

**The Significance of Predictive Analysis in Forecasting Customer Retention: A
Study Conducted in the Insurance Industry of**

Sri Lanka

Submitted on Friday, 10th June 2022

By

Pethmi Senara Kalpage

Bachelor of Science (Honours)

BSc (Hons) Business Information Systems

DECLARATION

This work is being submitted in partial fulfillment of the requirements for the degree of BSc (Hons) Business Information Systems and has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed – Pethmi Senara Kalpage

Date - 10th of June 2022

SUPERVISOR'S DECLARATION STATEMENT

Student Name -

Supervisor's Name -

I acknowledge that the above-named student has regularly attended the meeting, and actively engaged in the dissertation supervision process.

Signed - (Supervisor)

Date -

ACKNOWLEDGEMENT

I'd like to express my deepest appreciation and gratitude to Mr. Induneth De Silva, our supervisor and lecturer for the Business Information Systems Dissertation Project. This endeavour would not have been possible without his knowledge, expertise as well as guidance throughout the completion of the dissertation.

I'd like to acknowledge my batchmates and seniors who supported me with advices and suggestions in order to improve the report in numerous ways. Lastly, I'd like to extend a sincere thanks to my family and friends who provided me with the moral support and motivation for the successful completion of my thesis.

ABSTRACT

This research is conducted with aims of achieving its primary objective; identifying the crucial elements that highly impacts customer retention of the insurance industry in Sri Lanka, while discovering how predictive analytics could be a significant technique to help companies with retaining their customers. The article dives into considering the variables; Assurance, Lifetime Value, Satisfaction, Financial Status and External Influences of specific demographics, identified through published articles, and analysing the perceptions of selected participants through quantitative data in order to uncover significant factors companies should consider.

The variables are then analysed to identify which has the highest impact on customer retention, and recommendations are made by the author based on the respective findings. Visualizations have been created using a Business Intelligence tool – Power BI, in order to portray customer segmentations and actual vs predicted results of Customer Retention. The importance of using predictive analytics for companies in insurance are highlighted and suggestions for future research is also mentioned by way of improving the built predictive model in order to improve accuracy and authenticity.

Table of Content

1.	CHAPTER 01 – INTRODUCTION	14
1.1.	Research Background.....	14
1.2.	Research Problem.....	15
1.3.	Research Questions	16
1.4.	Research Objectives	16
1.5.	Significance of the Research.....	16
1.6.	Scope of the Research	17
1.7.	Limitations of the Research.....	17
1.8.	Structure of Other Chapters.....	18
1.9.	Chapter Summary.....	19
2.	CHAPTER 02 – LITERATURE REVIEW	20
2.1.	Customer Retention.....	20
2.1.1.	Importance of Predictive Analysis.....	20
2.2.	Assurance	21
2.2.1.	Advertising.....	21
2.2.2.	Agent's communication	21
2.3.	Lifetime Value.....	22
2.3.1.	Perceived premium	22
2.3.2.	Risk attitude	23
2.3.3.	Product Type.....	23
2.4.	Satisfaction.....	23
2.4.1.	Discounts.....	23
2.4.2.	Benefits provided	24
2.5.	Financial Status	24
2.6.	External Influences.....	24

2.7. Socio Demographics	25
2.7.1. Age.....	25
2.7.2. Education level.....	25
2.7.3. Household dependencies.....	26
2.8. Chapter Summary.....	26
3. CHAPTER 03 – RESEARCH METHODOLOGY	27
3.1. Introduction	27
3.2. Research Design	27
3.3. Research Philosophy	28
3.4. Conceptual Framework	29
3.5. Research Process	30
3.6. Data Collection.....	31
3.7. Population And Sample Population	31
3.8. Hypothesis.....	33
3.8.1. Correlation based Hypothesis	33
3.8.2. Simple Linear based Hypothesis.....	34
3.8.3. Multiple Regression based Hypothesis	35
3.9. Operationalization	36
3.10. Pilot Study	40
3.11. Data coding.....	40
3.12. Reliability Analysis	41
3.12.1. Assurance	41
3.12.2. Lifetime Value.....	42
3.12.3. Satisfaction	42
3.12.4. Financial Status	43
3.12.5. External Influences.....	44
3.12.6. Retention	45

3.13.	Validity analysis	46
3.13.1.	Assurance	46
3.13.2.	Lifetime Value.....	47
3.13.3.	Satisfaction.....	48
3.13.4.	Financial Status	50
3.13.5.	External Influences.....	51
3.13.6.	Retention	52
3.14.	Chapter Summary	53
4.	CHAPTER 04 – ANALYSIS AND FINDINGS	54
4.1.	Introduction	54
4.2.	Data Preparation.....	55
4.3.	Reliability Analysis.....	56
4.4.	Validity Analysis.....	57
4.5.	Data Distribution and Analysis	58
4.5.1.	Frequency Measurement.....	58
4.5.1.1.	Age.....	59
4.5.1.2.	Education Level	60
4.5.1.3.	Employment Status.....	61
4.5.1.4.	No. of Household Dependencies	62
4.5.1.5.	No. of Insurance Policies Held	63
4.6.	Customer Segmentation based on Age Groups – Power BI Visualization ...	64
4.7.	Descriptive Statistics	66
4.8.	Inferential Statistics.....	67
4.9.	Correlation Analysis.....	67
4.9.1.	Hypothesis 1C.....	68
4.9.2.	Hypothesis 2C.....	68
4.9.3.	Hypothesis 3C.....	69
4.9.4.	Hypothesis 4C.....	70

4.9.5. Hypothesis 5C.....	71
4.10. Simple Linear Regression Analysis.....	72
4.10.1. Hypothesis 1R	73
4.10.2. Hypothesis 2R	74
4.10.3. Hypothesis 3R	75
4.10.4. Hypothesis 4R	76
4.10.5. Hypothesis 5R	77
4.11. Multiple Regression Analysis.....	78
4.11.1. Hypothesis 1MR.....	78
4.12. Discussion On Findings.....	81
4.12.1. Correlation Analysis.....	81
4.12.2. Regression Analysis	82
4.12.3. Multiple Regression Analysis	83
4.13. Multiple Regression – Predicted Model (Power Bi Visualization)	84
4.14. Chapter Summary	85
5. CHAPTER 05 – CONCLUSIONS AND RECOMMENDATIONS	86
5.1. Review Of Research Objectives.....	86
5.2. Research Implications and Contributions	91
5.3. Recommendations	91
5.4. Research Limitations	94
5.5. Future Research.....	95
6. REFERENCES	96
7. APPENDIX AND ANNEXURES	102
7.1. Appendix 01 – Supervisor Approval and Log Sheets	102
7.2. Appendix 02 – Final Survey Questionnaire	110
7.3. Appendix 03 - Individual Reliability Analysis	119
7.4. Appendix 04 - Individual Validity Analysis	122

7.5.	Appendix 05 - Individual Correlation Analysis Results	127
7.6.	Appendix 06 - Individual Regression Analysis Results.....	129
7.7.	Appendix 07 – Individual MR Hypothesis Testing	134
7.7.1.	Hypothesis 2MR	134
7.7.2.	Hypothesis 3MR	135
7.7.3.	Hypothesis 4MR	136
7.8.	Appendix 08 – Survey Results	137

Table of Figures

Figure 1 Research Onion	28
Figure 2 Conceptual Framework	29
Figure 3 Research Process	30
Figure 4 Calculating Sample Size.....	32
Figure 5 Assurance - Cronbach Alpha.....	41
Figure 6 Assurance - Item Statistics	41
Figure 7 Lifetime Value - Cronbach Alpha	42
Figure 8 Lifetime Value - Item Statistics.....	42
Figure 9 Satisfaction - Cronbach Alpha (Before)	42
Figure 10 Satisfaction - Item Statistics (Before).....	43
Figure 11 Satisfaction - Cronbach Alpha (After)	43
Figure 12 Satisfaction - Item Statistics (After).....	43
Figure 13 Financial Status - Cronbach Alpha.....	43
Figure 14 Financial Status - Item Statistics	44
Figure 15 External Influences - Cronbach Alpha (Before).....	44
Figure 16 External Influences - Item Statistics (Before)	44
Figure 17 External Influences - Cronbach Alpha (After)	45
Figure 18 External Influences - Item Statistics (After).....	45
Figure 19 Retention - Cronbach Alpha.....	45
Figure 20 Retention - Item Statistics	45
Figure 21 Validity % - Assurance.....	46
Figure 22 Component Matrix - Assurance.....	46
Figure 23 KMO & Bartlett's Test – Assurance.....	47
Figure 24 Validity % - Lifetime Value	47
Figure 25 Component Matrix - Lifetime Value	47
Figure 26 KMO & Bartlett's Test - Lifetime Value	48
Figure 27 Validity % and Component Matrix - Satisfaction (Before).....	48
Figure 28 Validity % - Satisfaction (After)	49
Figure 29 Component Matrix - Satisfaction (After)	49
Figure 30 KMO & Bartlett's Test – Satisfaction.....	49
Figure 31 Validity % - Finance.....	50
Figure 32 Component Matrix - Financial Status.....	50

Figure 33 KMO & Bartlett's Test - Financial Status	50
Figure 34 Validity % and Component Matrix – External Influences (Before).....	51
Figure 35 Validity % and Component Matrix – External Influences (After)	51
Figure 36 KMO & Bartlett's Test - External Influences	52
Figure 37 Validity % - Retention.....	52
Figure 38 Component Matrix -Retention.....	52
Figure 39 KMO & Bartlett's Test - Retention.....	53
Figure 40 Responses Collected.....	55
Figure 41 Age distribution.....	59
Figure 42 Education Level distribution	60
Figure 43 Employment Status distribution	61
Figure 44 Household dependency distribution	62
Figure 45 Number of policies held	63
Figure 46 Customer Segmentation based on Age - Below 25	64
Figure 47 Customer Segmentation based on Age - 25 – 35	65
Figure 48 Customer Segmentation based on Age - 35 - 50	65
Figure 49 Customer Segmentation based on Age - Above 50.....	66
Figure 50 Descriptive statistics.....	66
Figure 51 Scatter Plot - Assurance.....	68
Figure 52 Scatter Plot - Lifetime Value	69
Figure 53 Scatter Plot - Satisfaction	69
Figure 54 Scatter Plot - Financial Status.....	70
Figure 55 Scatter Plot - External Influences	71
Figure 56 Model Summary - Assurance	73
Figure 57 ANOVA Test - Assurance.....	73
Figure 58 Regression Coefficients - Assurance.....	73
Figure 59 Model Summary - Lifetime Value	74
Figure 60 ANOVA Test - Lifetime Value	74
Figure 61 Regression Coefficients - Lifetime Value	74
Figure 62 Model Summary - Satisfaction.....	75
Figure 63 ANOVA Test - Satisfaction.....	75
Figure 64Regression Coefficients - Satisfaction.....	75
Figure 65 Model Summary - Financial Status	76

Figure 66 ANOVA Test - Financial Status.....	76
Figure 67 Regression Coefficients - Financial Status.....	76
Figure 68 Model Summary - External Influences.....	77
Figure 69 ANOVA Test - External Influences	77
Figure 70 Regression Coefficients - External Influences	77
Figure 71 Model Summary - 1MR.....	78
Figure 72 ANOVA Test - 1MR	78
Figure 73 Regression Coefficients - 1MR	79
Figure 74 Predicted Model Visualization - Actual Vs Predicted Customer Retention	84
Figure 75 Improved Conceptual Framework.....	86
Figure 76 Supervisor Approval.....	102
Figure 77 Project Log Sheet - 01	103
Figure 78 Project Log Sheet - 02	104
Figure 79 Project Log Sheet - 03	105
Figure 80 Project Log Sheet - 04	106
Figure 81 Project Log Sheet - 05	107
Figure 82 Project Log Sheet - 06	108
Figure 83 Project Log Sheet - 07	109
Figure 84 Final Survey Questionnaire	110
Figure 85 Final Survey Questionnaire (cont'd).....	111
Figure 86 Final Survey Questionnaire (cont'd).....	112
Figure 87 Final Survey Questionnaire (cont'd).....	113
Figure 88 Final Survey Questionnaire (cont'd).....	114
Figure 89 Final Survey Questionnaire (cont'd).....	115
Figure 90 Final Survey Questionnaire (cont'd).....	116
Figure 91 Final Survey Questionnaire (cont'd).....	117
Figure 92 Final Survey Questionnaire (cont'd).....	118
Figure 93 Final Survey Questionnaire (cont'd).....	118
Figure 94 Assurance - Cronbach Alpha.....	119
Figure 95 Assurance - Item Statistics	119
Figure 96 Lifetime Value - Cronbach Alpha	119
Figure 97 Lifetime Value - Item Statistics.....	120
Figure 98 Satisfaction - Cronbach Alpha.....	120

Figure 99 Satisfaction - Item Statistics	120
Figure 100 Finance - Cronbach Alpha.....	120
Figure 101 Finance - Item Statistics	121
Figure 102 External Influences - Cronbach Alpha	121
Figure 103 External Influences - Item Statistics.....	121
Figure 104 Retention - Cronbach Alpha.....	121
Figure 105 Retention - Item Statistics	122
Figure 106 Validity Loading % - Assurance	122
Figure 107 Component Matrix - Assurance.....	122
Figure 108 KMO & Bartlett's Test – Assurance.....	123
Figure 109 Validity Loading % - Lifetime Value.....	123
Figure 110 Component Matrix - Lifetime Value.....	123
Figure 111 KMO & Bartlett's Test - Lifetime Value.....	123
Figure 112 Validity Loading % - Satisfaction	124
Figure 113 Component Matrix - Satisfaction	124
Figure 114 KMO & Bartlett's Test - Satisfaction	124
Figure 115 Validity Loading % - Finance	124
Figure 116 Component Matrix - Finance.....	125
Figure 117 KMO & Bartlett's Test - Finance	125
Figure 118 Validity Loading % - External Influences.....	125
Figure 119 Component Matrix - External Influences	125
Figure 120 KMO & Bartlett's Test - External Influences	126
Figure 121 Validity Loading % - Retention	126
Figure 122 Component Matrix - Retention.....	126
Figure 123 KMO & Bartlett's Test - Retention.....	126
Figure 124 Correlation Analysis - Assurance.....	127
Figure 125 Correlation Analysis - Lifetime Value	127
Figure 126 Correlation Analysis - Satisfaction.....	127
Figure 127 Correlation Analysis - Financial Status	128
Figure 128 Correlation Analysis - External Influences	128
Figure 129 RSR Histogram - Assurance.....	129
Figure 130 RSR Normal P-Plot – Assurance.....	129
Figure 131 RSR Normal P-Plot - Lifetime Value.....	130

Figure 132 RSR Normal P-Plot - Lifetime Value.....	130
Figure 133 RSR Histogram - Satisfaction	131
Figure 134 RSR Normal P-Plot - Satisfaction	131
Figure 135 RSR Histogram - Financial Status.....	132
Figure 136 RSR Normal P-Plot - Financial Status	132
Figure 137 RSR Histogram - External Influences	133
Figure 138 RSR Normal P-Plot - External Influences.....	133
Figure 139 Model Summary - 2MR.....	134
Figure 140 ANOVA Test - 2MR	134
Figure 141 Regression Coefficients - 2MR	134
Figure 142 Model Summary - 3MR.....	135
Figure 143 ANOVA Test - 3MR	135
Figure 144 Regression Coefficients - 3MR	135
Figure 145 Model Summary - 4MR.....	136
Figure 146 ANOVA Test - 4MR	136
Figure 147 Regression Coefficients - 4MR	136
Figure 148 Survey Results - Product Significance	137
Figure 149 Survey Results - Opinion on risk and perceived prices.....	137
Figure 150 Survey Results - Financial Perception.....	138
Figure 151 Survey Results - Discounts and other benefits	138
Figure 152 Survey Results - Assurance	138
Figure 153 Survey Results - Inflation and Competition	139
Figure 154 Survey Results - Overall Feedback	139

List of Tables

Table 1 Operationalization Breakdown – Dimensions & Elements	37
Table 2 Operationalization - Hypothesis and Questions.....	39
Table 3 Data coding	40
Table 4 Primary Data Coding	55
Table 5 Final Survey - Reliability Analysis.....	56
Table 6 Final Survey - Validity Analysis	57
Table 7 Measures of Frequency	58
Table 8 Correlation Analysis	67
Table 9 Regression Analysis.....	72
Table 10 MR Hypothesis Test Results.....	79
Table 11 Correlation Hypothesis - Summary	81
Table 12 Regression Hypothesis - Summary	82
Table 13 Multiple Regression Hypothesis - Summary	83
Table 14 Review of Research Questions	89
Table 15 Review of Research Objectives	90

Abbreviations

RSR – Regression Standardized Residual

1. CHAPTER 01 – INTRODUCTION

1.1. Research Background

The Sri Lankan insurance industry counts to being one of the major elements of the economy by its premium collection amounts; Rs.209 bn in 2020, as well as the role it plays in covering risks of both businesses and individuals. The overall health and performance of the industry lead to a 6% GDP growth and 14% asset growth in comparison to 2019 (KPMG, 2021). According to a press release by the (Insurance Regulatory Commission of Sri Lanka, 2019), the long-term insurance business, Sri Lanka records lapse rates at 11.69% in 2019, as a percentage of new policies issued, which are justified to be impacts of Covid-19 and higher churning rates.

Non-life insurance retention in 2017 was reported at 51.4%, a 79.75% decrease compared to the previous year, which results in an increase in business costs to acquire new customers. The importance of customer retention is therefore recognized, according to (Bitran, 1997), in saying that customer retention is more profitable than customer acquisition, resulting in companies adopting newer ways of relationship marketing. (Gunther, 2014) also states that insurance companies cannot depend on a stable customer base anymore, and agrees that a company would have to spend 12 times the value spent on retaining customers, in order to acquire new ones. A study shows that profits can be raised by 25% - 95% by just a 5% boost in retention. Therefore, the “most important customer retention strategy” would be to recognize existing clients most expected to churn.

1.2. Research Problem

Despite insurance being a pillar of support for individuals and companies in times of need, it operates in a competitive market, making it difficult for these companies to attract and be able to retain their customers. The matter in question here is the fact that insurance companies find it difficult to maintain a loyal customer base. This practically means that customer retention in these companies is unsteady, resulting in higher churning rates (customer turnover). In general, this could be assumed to be due to various reasons such as financial difficulties, lower levels of satisfactions, poor communication or having unsuccessful marketing strategies. (Swiss Re, 2014) states that “over 30% customers shop and compare” before making decisions, increasing risks of switching companies, while less than 20% purchase over a product.

Companies could make the mistake of attracting more low-value customers and miss out on converting first-time buyers into repeating customers. On the other hand, prioritizing the wrong set of customers could also lead to losses. These companies would thus benefit from identifying the main factors that contribute to, and affect customer behaviour in the insurance industry. This brings light onto the question, “what are the main factors that affect customer retention?”. For this, insurance companies should have proper time, resources and technology to provide a better experience.

Unpredictable customer retention tends to be a critical issue in the industry as the main source of income for such companies are through premiums paid by customers, and a change or break in this pattern could lead to inconsistent revenue. Therefore, if companies were able to pin down these issues, it would be an immense advantage to apply the concept of predictive analysis in order to forecast customer behaviour in the company in the upcoming future, that could stem into stronger decision-making attitudes. This can be though personalized recommendations or even having segmented campaigns to improve experience. There is a lack of literature to support companies with information on strategies they could follow to improve customer relations in this sector, and therefore this research would aid in finding the significant factors that the company needs to keep a close eye on.

1.3.Research Questions

The principal question that revolves around this research is to identify why customer retention in the insurance domain in Sri Lanka is low and uncertain with relate to variables that have the highest impact on retention. Sub-questions identified are as follows:

1. What are the main factors that influence customer retention?
2. What are reasons that cause a high churning rate?
3. What actions can the company take in order to retain customers?
4. What benefits will predictive analysis bring into companies in the insurance sector?
5. How much of an impact do external factors have on retention rates?

1.4.Research Objectives

- Identify key factors that impose a significant influence on customer retention in the insurance industry of Sri Lanka.
- Conduct proper customer segmentation
- Feed data gathered into a predictive model that would portray the anticipated change in customer retention in correlation with the variables identified.
- Utilize this information in order to make informed decisions and provide business information solutions to companies.
- Create a visual dashboard to display the final output using Power BI.

1.5.Significance of the Research

The importance of conducting this study would be to identify in the Sri Lankan context, peoples' perception on insurance, that results in retention and churning within the industry, and subsequently identifying methods in which companies could use the advantage of predicting these changes in order to build a strategic marketing approach and make effective decisions, resulting in higher revenue. It also aims to take into consideration external factors affecting churning/retention and aid in covering the academic gap in this area. Moreover, this research is also intended on covering the

research gap that exists, proved by the lack of literature and scarce resources available on how predictive analysis could play a major role in customer retention and company success in the insurance industry.

1.6.Scope of the Research

This research would be focused on analysing information that is gathered from the public, in relation to identified variables that have an impact on customer retention, while also analysing data gathered on customers from a leading insurance company in Sri Lanka. This information would be used to identify factors with a significant impact on retention, and a statistical model will be created to predict future changes in retention.

1.7.Limitations of the Research

- There are various limitations however, the main being the lack of literature available, in order to recognize appropriate variables. On the other hand, we do need to consider the fact that the secondary data which will be used for this analysis, could include numbers affected by Covid-19 in the past couple of years. This could result in identifying factors that may or may not be significant in a situation where the pandemic doesn't exist.
- Due to the prevailing situation in Sri Lanka, it is difficult to assume the current market situation, which has a huge impact on the day to day lives of citizens, which could cause them to make sudden and unpredictable decisions.

1.8.Structure of Other Chapters

This research article is broken down into five main chapters, summarized below:

Chapter 01 – Introduction

Contains the background to the identified problem, explaining the significance of the research towards the insurance industry in Sri Lanka. This chapter also contains the objectives and questions aimed to be prioritized, while also identifying limitations, scope and the research gap.

Chapter 02 – Literature Review

This chapter contains findings of past literature with regard to the selected variables of the research. Thereby, this secondary information can thereby be used to get a better idea of the behaviour of variables in similar past researches.

Chapter 03 – Research Design

The research design mainly explains the methodology in which the research was carried out. It contains the research philosophy, as well as the operationalization process. This chapter also goes on to explaining the conceptual framework and contains the hypothesis and pilot study analysis that would be benefiting the next few chapters.

Chapter 04 – Analysis and Findings

This chapter exhibits the quantitative data gathered during the research, as well as explains and shows how this information has been analysed with regard to the identified problem, correlated with the hypothesis created in the previous chapter. It contains analysis to test the reliability, validity, corelation, regression and multiple regression for the formulated hypothesis, as well as a comparison on the Actual Vs Predicted Customer Retention values.

Chapter 05 – Conclusion and Recommendations

This chapter will contain a review on the research questions and objectives covered in the study. The conclusions and recommendations will be identified and analysed in order to be of support for future research with given suggestions.

1.9. Chapter Summary

This section essentially identifies the research problem, and states the objectives and questions aimed to be covered by the end of this research. The research gap and significance of conducting this study has been stated, with clear limitations faced in the process.

2. CHAPTER 02 – LITERATURE REVIEW

Previous studies have been conducted around the world in order to identify the significance and usefulness of using predictive analytics when it comes to retaining customers in companies.

2.1.Customer Retention

(Gunther, 2014) identifies small increases in rates of retention to increase revenue gained by premiums by millions, increasing the importance of customer retention in CRM, in identifying customers that are likely to churn. This way companies can “refocus marketing efforts from acquisition” in order to protect customers with changing insurance demands, from competitors by providing enough offers. General marketing principles state that “keeping customers is more profitable than attracting new ones”, and its crucial for companies to “invest in customers who are potentially valuable for the company” (Peter C. Verhoef, 2010) . According to (Swiss Re, 2014), “over 30% customers shop and compare” before making decisions, increasing risks of switching companies, while less than 20% purchase over a product. Therefore, focusing on customer retention while reducing churning tends to be a crucial aspect companies must focus on.

2.1.1.Importance of Predictive Analysis

Once predictive models are in place retention can thereby be improved by contacting “the right customer, with the right offer, at the right time” (Verhoef & Venkatesan, 2010). Author (Abby Borden, 2021) identifies predictive analytics to thereby be able to help companies measure responses of customers in relation to the promotional efforts undertaken, in order to understand how certain segments would react to changes. (Stefan Mau, 2017) states that only 9% of insurers use predictive models, and the usage of this would help companies in order to avoid negative consequences such as over-touching customers. Therefore, analytical models such as ones built in this study would help in order to identify customers with changing demands.

The following variables have been recognized to have certain gravity over the chosen topic, and are split into the following subtopics:

- Assurance
- Lifetime Value
- Satisfaction
- Financial Status
- External Influences
- Socio Demographics

2.2.Assurance

A customer is usually considered to have churned if they've decided to switch to different providers, or even cancelled their policies, for main reasons being a lack of need or dissatisfaction in general (Gunther, 2014). Assurance tends to be one of the most decisive factors as it brings up the risk of "over-touching" if used incorrectly (Stefan Mau, 2017). This variable can be measured with the following indicators:

2.2.1.Advertising

One of the most important research aspects to be considered is targeted advertising as it provides businesses with proper insights on promoting their products. The author, (Hsiao-Fan Wang, 2005), as well as (D. Kalaivani, 2018) suggest incorporating data mining in order to achieve effective profitability management. The author further observes the importance of business intelligence in order to analyze the efficiency measures. This way, patterns and suggestions can be made accordingly, with the help of knowledge management. According to research by (Negi, 2010), one to one marketing and loyalty programs can be given more attention to, in order to segment customers according to their preferences, and identify purchase habits for valid promotions.

2.2.2.Agent's communication

In the insurance sector, sales agents personally contacting customers would be an effective way to use human resources. According to author (Swiss Re, 2014), although people have access to information online, personal interactions remain crucial for a

majority of customers. The article also highlights benefits the predictive model could bring to the company, such as having the upper hand on correct timing when it comes to market interventions. According to the results of the article, 25% of customers find receiving calls to be disturbing, and thus, in order to reduce over-touching of customers, it would be important to understand the customers preference on personal contact, which could directly impact retention.

2.3.Lifetime Value

(Werner J. Reinartz, V. Kumar, 2003) emphasizes on the importance of lifetime value over time, in order to improve business performance, which could increase customer loyalty. The objective of maximizing customer's lifetime value is one of the main aspects of customer relationship managements (Kumar.V, 2006). Expanding businesses with existing customers, rather than acquiring new ones is considered to be less costly and more profitable for companies. Predictive analytics thus aids in improving the lifetime value by considering on the equity growth (Tala Mirzaei, 2014). The article by (Gunther, 2014) concludes that there is an increased risk of churn when there is a shorter lifetime. That is, even if a customer is to re-join, when compared to a customer who retained, the churning probabilities are considered high. Once the lifetime value is predicted, measured can be taken to save costs of reaching out to customers (Milliman, 2016). This variable is broken down into the following sub-categories:

2.3.1.Perceived premium

(Gunther, 2014) associates perceived premium to be a decisive variable and identifies customers who yearly pay a large premium, and have at least 3 main insurances to be of high risk when it comes to churn. With the analysis, the results also conclude that clients that pay a high instalment are bound to retain if their partners are also customers in the same company. According to (Noorhannah Boodhun; Manoj Jayabalan, 2018), analytics would thereby be helpful in order to provide the correct amount of premiums according to the specific risks, as it also depends on the customer's lifestyle (Milliman, 2016).

2.3.2.Risk attitude

(Peter C. Verhoef, 2010) implies unobserved attitudes towards risk of customers could lead to interdependencies, when it comes to purchase decisions across various policies. This is important when it comes to predicting the potential value in areas such as consumer research, as it very much affects the acquisition of products and services.

2.3.3.Product Type

(D. Kalaivani, 2018) analyses the importance of knowing how consumers would behave when it comes to changes in factors such as delivery of product, cost and also type of product. The article also acknowledges the fact that there could be an increase in ROI, customer happiness and conversions in a predictive model when it comes to delivering a specific product to the right customer. Better strategic decisions can also be made once customer profiles are identified and customers are provided with the required products (Mark Xu, 2005).

2.4.Satisfaction

(Fatemeh Safara, 2020) reveals that improving customer satisfaction remains challenging in any industry and thus machine learning and predictive models can be proposed to forecast the behaviour of customers. Profitability can thereby be identified, according to (Tala Mirzaei, 2014) once similar patterns are identified and profiles are segmented. The author, with relate to prior research thus identifies satisfaction to be one of the most prevalent variables when it comes to retaining customers. The following are factors that have an obvious impact to customer satisfaction.

2.4.1.Discounts

(Gunther, 2014) groups the variable discounts into few categories, and identifies that the customers offered with discount programs substantially pertains to a lower churn level. Concluding results of the analysis shows that the most important factor that could affect customer retention is thereby discounts.

2.4.2.Benefits provided

When considering a rational customer, (Maresova Petra, 2012) identifies that consumer mainly focus on maximizing their benefits, and thus providing more value to a service could be effective. The article also focuses on customer's perceptions on the rate of future benefits they'll receive as well.

2.5.Financial Status

Financial status of a customer could be an important factor when it comes to them deciding whether or not to continue paying premiums. This can be broken down into the following categories.

- Job status
- Income group

(Wagner A. Kamakura, 1991) reports that income is an important determinant when it comes to financial services, and shows a strong interdependency between the variable with purchased services, and breaks it down into low, middle, high, very high, and unknown-income levels. Potential behaviours can thus be analysed once the income bracket and job status of an individual is known. (Maresova Petra, 2012) thereby identifies that customer who are financially sound, have a lower chance of churning as they would not question the price of products as much.

2.6.External Influences

Various articles find the need to identify the impact of external factors on customer retention. These influences can be broken down into 2 main categories.

- Inflation rates
- Competitor influences

(Maresova Petra, 2012) reveals that the insurance industry is no exception when it comes to increasing competition and (Gunther, 2014) emphasizes on the need to focus on factors such as media and competitor campaigns as they are bound to have an

adverse effect on customer retention. Identifying customer strategies would thus be beneficial. (Elsayed Elashkar, 2020) uses probit linear regression in order to analyze the competitor product behaviour in order to gain competitive advantage, and analyze more dynamic factors in the process as well.

2.7.Socio Demographics

Socio demographics are considered to be the main factors/variables that have significant levels of impacts on customer retention, in most articles. In this research, the following socio demographic factors will be analyzed in order to identify correlations and also analyze the sample target population;

2.7.1.Age

According to the correlation analysis conducted by (Fatemeh Safara, 2020), age has the highest correlation with the features and is considered a more influential variable, when it comes to predicting customer behaviour, which is also agreed to by (Tala Mirzaei, 2014) to be of higher prevalence. On the other hand, (Gunther, 2014) analyzes a specific age group of customers and identifies linear relationships between age and customer churn. According to a study conducted in the health insurance industry, age is used as a variable that provides more useful information of payments, which could determine the probability of customer retention accordingly (Jamal, 2019).

2.7.2.Education level

The article by (Fatemeh Safara, 2020) also takes into consideration the education level, which is split into 5 main stages:

- No
- Diploma
- Bachelor
- Masters
- PhD

And identifies that the level of education correlates with the customers decision to retain in the company as it also is an important determinant when it comes to acquiring

financial services (Peter C. Verhoef, 2010). It is also assumed that the rational behaviour of customers can be challenged when factors such as education levels are taken into consideration.

2.7.3.Household dependencies

The article by (Stefan Mau, 2017) takes into consideration the household dependencies, which covers for common attributes such as family status and the value of the household, and states that including data mining to this variable could account to improved prediction performance. To this, (Peter C. Verhoef, 2010) also agrees with regard to the fact that a person's needs could be affected by such factors, and providing unique/specific products catered to their requirements could increase retention within the company.

When taking all the factors discussed in literature into consideration, the following independent variables will be taken into account going forth with this research. The conceptual framework of the study would thereby consist of the following notable variables as the main independent variables which will be operationalized in chapter 03;

- Assurance
- Lifetime Value
- Satisfaction
- Financial Status
- External Influences

2.8.Chapter Summary

This chapter consists of the literature that was used in order to identify variables, as well as determine relationships between the chosen topic and previous researches. The main variables; Assurance, Lifetime Value, Satisfaction, Financial Status, and External Influences, as well as supporting factors such as Socio Demographics have been taken into consideration and explained in this chapter.

3. CHAPTER 03 – RESEARCH METHODOLOGY

3.1. Introduction

Chapter 03 contains the research design, philosophy and process. It also dives into the conceptual framework that is based on the literature examined in the earlier chapter, as well as formulating the hypothesis, variable operationalization and consists of the pilot study and analysis done in order to test the validity and reliability of the variables.

3.2. Research Design

This research conducted for the purpose of identifying factors that affect customer retention in insurance will follow a mono method of study, where only one method; quantitative data will be considered. The strategy chosen based on the purpose of the study would be mainly analyzing the distributed survey and in this way, variables can be analyzed, and accumulated data through questionnaires can be used to predict changes in customer retention. As a result, primary data will have been collected and analyzed, while theories in previous literature can be tested. The cross-sectional time horizon will be followed not only for the fact that observations will be carried out for a single point in time, but also for the added advantage that it would aid in comparing various variables parallelly without having to manipulate the environment (Writers, 2022).

A pilot study will be conducted before collecting the primary data, intending to test the accuracy, reliability and significance of the variables identified through literature. A successful pilot study will thereby result in a successful data collection and analysis process, which is the main objective of this chapter. Thereby, the relationships between the dependent and independent variables can be identified, and the hypothesis can be tested.

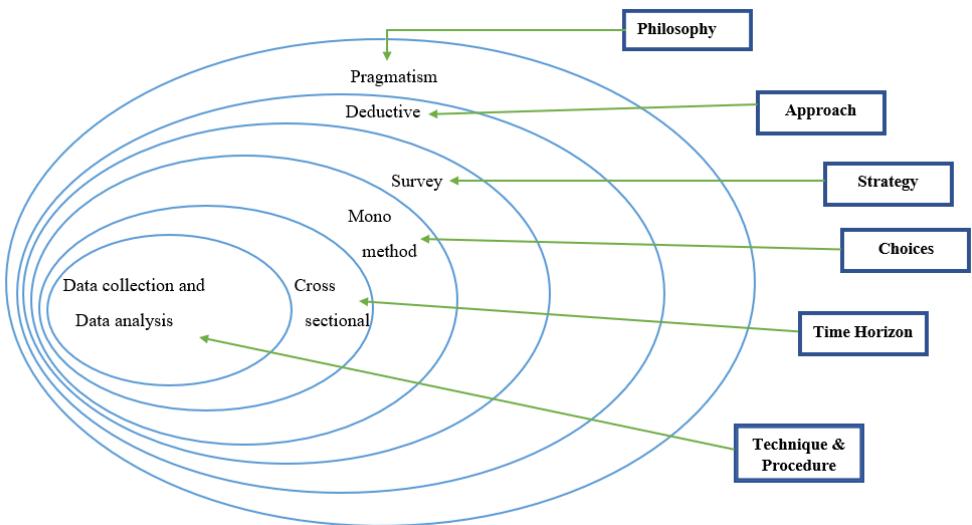


Figure 1 Research Onion

3.3. Research Philosophy

This research will be done based on the Pragmatism research philosophy, which recognizes that there can be different points of view in order to interpret actions. This makes use of both positivism and interpretivism, involving statistical tools, and structured data, but at the same time, understanding that human nature will vary based on different social settings. Here, research questions will be one of the most important determinants of the research philosophy, and there can also be more than approaches to the research and strategies considered in the study (John Dudovskiy, 2022).

A deductive approach will also be used in order to conduct this research. This involves studying previous literature and existing theories, and testing the hypothesis. In this study, it would help in identifying relationships between variables and also in order to measure findings and predictions. The research mainly focuses on subjective variables, which analyses the perceptions and experiences of an individual, and not just the final outcome. This is considered important in this research as it mainly focuses on identifying reasons for churn, which depends, and is subjective to every person. Other data such as socio demographic factors would thereby be considered in order to help understand and segment participants more effectively.

3.4. Conceptual Framework

According to the factors identified and examined in Chapter 02 literature review, the variables that would affect customer retention rates in the Sri Lankan insurance domain are included in a conceptual framework as follows;

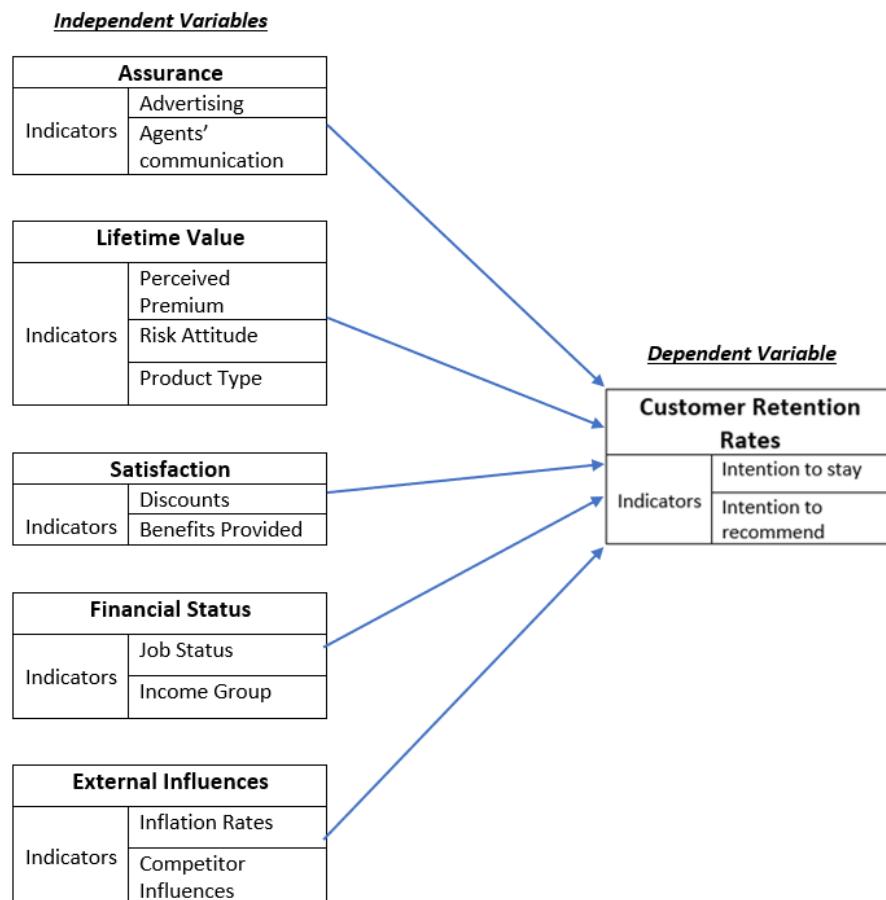


Figure 2 Conceptual Framework

3.5.Research Process

The structure intended to be followed in the research is shown below;

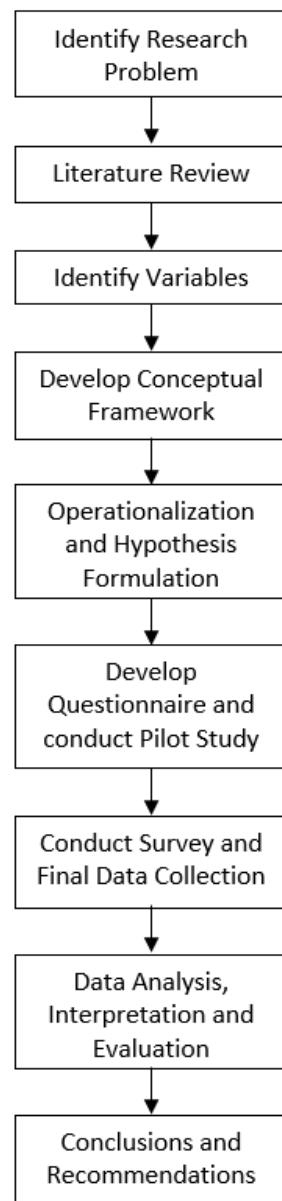


Figure 3 Research Process

3.6.Data Collection

When it comes to data collection, it can be of two methods; primary and secondary. Primary data is data which is usually collected through first-hand experiences and has not been used in the past, and can be collected through interviews, questionnaires, etc. on the other hand secondary data is data that can be obtained through sources, both internal and external of an organization. This could be data from reports such as executive summaries or financial statements (QuestionPro, 2022).

Once the variables have been identified from the past researches, the researcher will proceed to collect the quantitative data through google forms. The forms will be distributed, specific to customers of the insurance domain, in order to get insight on the selected variables. In order to conduct the pilot study, a sample of 35 respondents were collected. This was done in order to identify whether the variables collected through literature can be practically applied especially in the insurance domain in Sri Lanka. Valid and reliable variables can thus be identified and any inconsistencies.

3.7.Population And Sample Population

The target participants of this research are existing customers who currently hold at least one policy in the insurance domain. According to (Insurance Regulatory Commission of Sri Lanka , 2019) the industry expanded by issuing 736,961 new insurance policies. Considering this, a sample of 384 respondents will be targeted, with a confidence level of 95% and a margin of error of 5%.

Samples were collected randomly by distributing the questionnaire (Refer Appendix 02) over a selected number of social media platforms, which turned out to be the most convenient method for participants. Thereby, the questionnaire was distributed in order to receive 384 responds for the purpose of going ahead with analysis.

Population Size ⓘ

Confidence Level (%) ⓘ

▼

Margin of Error (%) ⓘ

Sample size
384

Doing market research? SurveyMonkey Audience gets you the right survey respondents fast and easy and helps you target them by demographics, consumer behavior, geography, or even designated marketing area.

[Choose your audience](#)

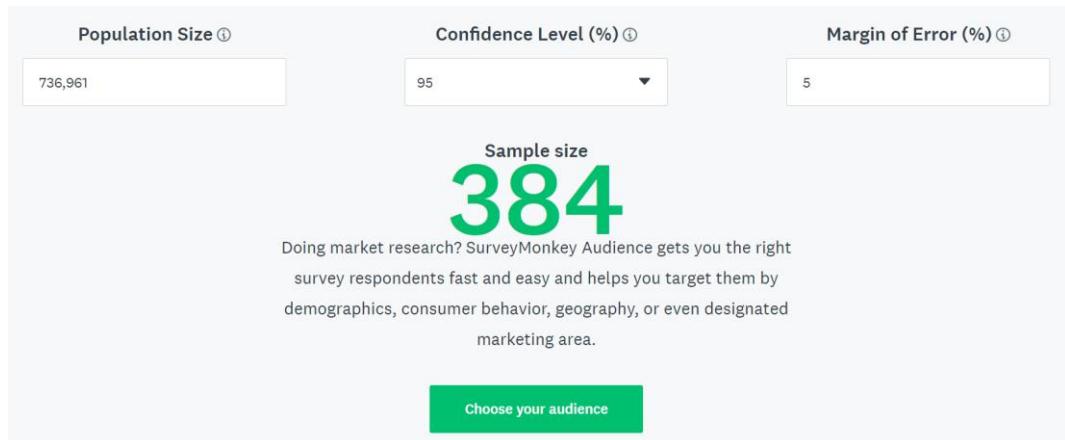


Figure 4 Calculating Sample Size

3.8.Hypothesis

The stated hypothesis is developed based on the conceptual framework of the research, named with the variable numbers, followed by letters ‘C’, ‘R’ and ‘MR’, respectively indicating Correlation, Simple Linear Regression and Multiple Regression.

3.8.1.Correlation based Hypothesis

Hypothesis 1C

H0 – Assurance does not have a relationship with Customer retention in Insurance

H1 – Assurance has a positive relationship with Customer retention in Insurance

Hypothesis 2C

H0 – Lifetime Value does not have a relationship with Customer retention in Insurance

H1 – Lifetime Value has a positive relationship with Customer retention in Insurance

Hypothesis 3C

H0 – Satisfaction does not have a relationship with Customer retention in Insurance

H1 – Satisfaction has a positive relationship with Customer retention in Insurance

Hypothesis 4C

H0 – Financial Status does not have a relationship with Customer retention in Insurance

H1 – Financial Status has a positive relationship with Customer retention in Insurance

Hypothesis 5C

H0 – External Influences does not have a relationship with Customer retention in Insurance

H1 – External Influences has a positive relationship with Customer retention in Insurance

3.8.2.Simple Linear based Hypothesis

Hypothesis 1R

H0 – Customer retention in Insurance is not affected by Assurance

H1 – Customer retention in Insurance is affected by Assurance

Hypothesis 2R

H0 – Customer retention in Insurance is not affected by Lifetime Value

H1 – Customer retention in Insurance is affected by Lifetime Value

Hypothesis 3R

H0 – Customer retention in Insurance is not affected by Satisfaction

H1 – Customer retention in Insurance is affected by Satisfaction

Hypothesis 4R

H0 – Customer retention in Insurance is not affected by Financial Status

H1 – Customer retention in Insurance is affected by Financial Status

Hypothesis 5R

H0 – Customer retention in Insurance is not affected by External Influences

H1 – Customer retention in Insurance is affected by External Influences

3.8.3.Multiple Regression based Hypothesis

Hypothesis 1MR

H0 – Assurance, Lifetime Value, Satisfaction, Financial Status and External Influences have no impact on Customer Retention in Insurance

H1 – Assurance, Lifetime Value, Satisfaction, Financial Status and External Influences have an impact on Customer Retention in Insurance

When conducting the analysis, the author was able to test and identify different combinations of variables in order to get a better and more accurate R Square value in order to get a more accurate statistical model. As a result, the following revised hypothesis were also evaluated;

Hypothesis 2MR

H0 –Lifetime Value, Satisfaction, Financial Status and External Influences have no impact on Customer Retention in Insurance

H1 –Lifetime Value, Satisfaction, Financial Status and External Influences have an impact on Customer Retention in Insurance

Hypothesis 3MR

H0 – Assurance, Lifetime Value, Financial Status and External Influences have no impact on Customer Retention in Insurance

H1 – Assurance, Lifetime Value, Financial Status and External Influences have an impact on Customer Retention in Insurance

Hypothesis 4MR

H0 – Lifetime Value, Financial Status and External Influences have no impact on Customer Retention in Insurance

H1 – Lifetime Value, Financial Status and External Influences have an impact on Customer Retention in Insurance

3.9. Operationalization

Variable	Dimensions	Elements
Dependent Variable:		
Customer Retention (Gunther, 2014) (Peter C. Verhoef, 2010) (Verhoef & Venkatesan, 2010)	<ul style="list-style-type: none"> • Intention to stay • Intention to recommend 	<ul style="list-style-type: none"> • Intention to stay with the insurance company in the foreseeable future. • Intention to recommend having an insurance to others.
Independent Variables:		
Assurance (Gunther, 2014) (Stefan Mau, 2017) (Hsiao-Fan Wang, 2005) (D. Kalaivani, 2018) (Negi, 2010) (Swiss Re, 2014)	<ul style="list-style-type: none"> • Advertising • Agent's communication 	<ul style="list-style-type: none"> • Social media advertising • Agent contact
Lifetime Value (Werner J. Reinartz, V. Kumar, 2003) (Kumar.V, 2006) (Tala Mirzaei, 2014) (Gunther, 2014) (Milliman, 2016)	<ul style="list-style-type: none"> • Perceived premium • Risk attitude • Product type 	<ul style="list-style-type: none"> • Insurance cost/price • Risk taking ability • Product preference • Needs/requirement

(Peter C. Verhoef, 2010) (D. Kalaivani, 2018) (Mark Xu, 2005)		
Satisfaction (Fatemeh Safara, 2020) (Tala Mirzaei, 2014) (Gunther, 2014) (Maresova Petra, 2012)	<ul style="list-style-type: none"> • Discounts • Benefits provided 	<ul style="list-style-type: none"> • Attractive price reductions • Additional benefits
Financial status (Wagner A. Kamakura, 1991) (Maresova Petra, 2012)	<ul style="list-style-type: none"> • Job status • Income group 	<ul style="list-style-type: none"> • Employment status • Likelihood of paying
External influences (Maresova Petra, 2012) (Gunther, 2014) (Elsayed Elashkar, 2020)	<ul style="list-style-type: none"> • Inflation rates • Competitor influences 	<ul style="list-style-type: none"> • Perception on price hikes • Competitive advantage

Table 1 Operationalization Breakdown – Dimensions & Elements

Once the variables have been broken down into dimensions and elements, the hypothesis and scale of the variables have been determined. This is shown in the table below along with the respective questions and scale of measurement.

Variable	Hypothesis	Dimensions	Scale	Questions
Assurance (Gunther, 2014) (Stefan Mau, 2017) (Hsiao-Fan Wang, 2005) (D. Kalaivani, 2018) (Negi, 2010) (Swiss Re, 2014)	H1C, H1R	Advertising	Likert	11.1, 11.2,
		Agent's communication	Likert	11.3, 11.4, 11.5, 11.6
Lifetime Value (Werner J. Reinartz, V. Kumar, 2003) (Kumar.V, 2006) (Tala Mirzaei, 2014) (Gunther, 2014) (Milliman, 2016) (Peter C. Verhoef, 2010) (D. Kalaivani, 2018) (Mark Xu, 2005)	H2C, H2R	Perceived premium	Likert	7.2, 8.3
		Risk attitude	Likert	8.1, 8.2, 8.4
		Product type	Likert	7.1, 7.3, 7.4,
Satisfaction (Fatemeh Safara, 2020) (Tala Mirzaei, 2014) (Gunther, 2014) (Maresova Petra, 2012)	H3C, H3R	Discounts	Likert	10.1, 10.3
		Benefits provided	Likert	10.2, 10.4

Financial status (Wagner A. Kamakura, 1991) (Maresova Petra, 2012)	H4C, H4R	Job status	Likert	9.1, 9.4
		Income group	Likert	9.2, 9.3
External influences (Maresova Petra, 2012) (Gunther, 2014) (Elsayed Elashkar, 2020)	H5C, H5R	Inflation rates	Likert	12.1, 12.2
		Competitor influences	Likert	12.3, 12.4
Customer Retention (Gunther, 2014) (Peter C. Verhoef, 2010) (Verhoef & Venkatesan, 2010)	-	Intention to stay	Likert	13.1, 13.2
		Intention to recommend	Likert	13.3

Table 2 Operationalization - Hypothesis and Questions

3.10. Pilot Study

The pilot study for this research is done by gathering data from 35 participants, and will cover tests conducted such as the reliability and validity analysis. The data gathered will be analyzed using the IBM SPSS Statistics – Version 23.

3.11. Data coding

In order to analyze the variables, the following data coding method was used through a Likert scale, with values ranging 1 through 5 in order to give a measurable value to responses as shown below;

5	4	3	2	1
Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

Table 3 Data coding

3.12. Reliability Analysis

A reliability analysis provides insight on the extent to which certain variables provide consistent results a repeated number of times (Statistic Solutions , 2022). The reliability in this research has been conducted through the Cronbach alpha analysis in SPSS. According to the standards, only variables with a value greater than 0.7 were considered reliable, with the capability of defining the selected variable. The results for each variable are as follows;

3.12.1. Assurance

An alpha value of 0.809 was identified for the 6 questions in the variable ‘Assurance’, and this was thereby considered reliable

Reliability Statistics	
Cronbach's Alpha	N of Items
.809	6

Figure 5 Assurance - Cronbach Alpha

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Assurance_1	18.4667	10.395	.554	.784
Assurance_2	18.8333	10.282	.468	.800
Assurance_3	18.6667	9.540	.638	.764
Assurance_4	18.8667	9.361	.670	.757
Assurance_5	19.0333	9.551	.508	.794
Assurance_6	18.4667	8.671	.610	.771

Figure 6 Assurance - Item Statistics

3.12.2. Lifetime Value

An alpha value of 0.863 was identified for the 8 questions in the variable ‘Lifetime Value’, and thereby this is considered reliable.

Reliability Statistics	
Cronbach's Alpha	N of Items
.863	8

Figure 7 Lifetime Value - Cronbach Alpha

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Lifetime_Value_1	29.0000	19.448	.633	.844
Lifetime_Value_2	29.2000	18.372	.615	.847
Lifetime_Value_3	29.1333	19.154	.576	.850
Lifetime_Value_4	28.9333	19.030	.726	.834
Lifetime_Value_5	29.1667	17.937	.805	.823
Lifetime_Value_6	29.1333	20.395	.580	.850
Lifetime_Value_7	29.1000	19.886	.561	.851
Lifetime_Value_8	29.2000	20.372	.435	.866

Figure 8 Lifetime Value - Item Statistics

3.12.3. Satisfaction

The variable ‘Satisfaction’ initially gives a Cronbach value of 0.621, due to this being unreliable, removing question 3 would increase the reliability of this variable.

Reliability Statistics	
Cronbach's Alpha	N of Items
.621	4

Figure 9 Satisfaction - Cronbach Alpha (Before)

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Satisfaction_1	12.0667	4.340	.276	.628
Satisfaction_2	12.0000	3.172	.564	.425
Satisfaction_3	12.6000	3.766	.199	.726
Satisfaction_4	12.0333	2.999	.654	.353

Figure 10 Satisfaction - Item Statistics (Before)

Once question 3 was removed, this increased the cronbach alpha value to 0.726, which is considered reliable for the author to go ahead with.

Reliability Statistics

Cronbach's Alpha	N of Items
.726	3

*Figure 11 Satisfaction - Cronbach Alpha (After)***Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Satisfaction_1	8.4333	2.461	.372	.818
Satisfaction_2	8.3667	1.757	.559	.628
Satisfaction_4	8.4000	1.490	.754	.355

Figure 12 Satisfaction - Item Statistics (After)

3.12.4. Financial Status

The variable ‘Financial Status’ has a Cronbach value of 0.758, which is considered reliable and has no issue going forth.

Reliability Statistics

Cronbach's Alpha	N of Items
.758	4

Figure 13 Financial Status - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Finance_1	12.5075	4.314	.628	.660
Finance_2	12.4179	4.762	.566	.696
Finance_3	12.2836	4.964	.608	.681
Finance_4	12.4030	4.699	.451	.766

Figure 14 Financial Status - Item Statistics

3.12.5. External Influences

Initially the Cronbach alpha value for the variable ‘External Influences’ was 0.594. as this is considered unreliable, the best that can be done is to remove question 3.

Reliability Statistics

Cronbach's Alpha	N of Items
.594	4

*Figure 15 External Influences - Cronbach Alpha (Before)***Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
External_1	11.8636	3.997	.363	.533
External_2	12.0455	3.583	.440	.472
External_3	12.6061	3.812	.250	.639
External_4	12.0758	3.794	.491	.447

Figure 16 External Influences - Item Statistics (Before)

However, even if question 3 is removed, a Cronbach value of just 0.639 can be obtained, which is considered unsatisfactory. The author however hopes to make improvements in order to increase the alpha value when collecting data from the sample respondents.

Reliability Statistics

Cronbach's Alpha	N of Items
.639	3

Figure 17 External Influences - Cronbach Alpha (After)

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
External_1	8.2727	2.017	.447	.543
External_2	8.4545	1.759	.500	.466
External_4	8.4848	2.223	.404	.600

Figure 18 External Influences - Item Statistics (After)

3.12.6. Retention

The dependent variable ‘Retention’ has a Cronbach alpha of 0.924, which is considered a high and reliable value, and therefore, no changes are to be made.

Reliability Statistics

Cronbach's Alpha	N of Items
.924	3

Figure 19 Retention - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Retention_1	8.1333	4.395	.781	.941
Retention_2	7.9667	4.378	.856	.882
Retention_3	8.1000	3.955	.903	.842

Figure 20 Retention - Item Statistics

3.13. Validity analysis

The validity analysis is done in order to identify how accurately a method is measured the way it's actually supposed to be measured (WordPress, 2011). Here, a value >50% can be considered valid. The validity analysis for the identified variables are as follows;

3.13.1. Assurance

The variable ‘Assurance’ gives a validity of 51.964% which is considered valid and no changes are to be made.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.118	51.964	51.964	3.118	51.964	51.964
2	1.260	20.996	72.960			
3	.719	11.983	84.943			
4	.546	9.100	94.043			
5	.197	3.282	97.325			
6	.160	2.675	100.000			

Figure 21 Validity % - Assurance

If later on any changes are to be made, according to the component matrix below, removing the 2nd question could increase the loading % by a relative amount.

Component Matrix ^a	
	Component
	1
Assurance_1	.689
Assurance_2	.633
Assurance_3	.776
Assurance_4	.785
Assurance_5	.661
Assurance_6	.765

Figure 22 Component Matrix - Assurance

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.607
Bartlett's Test of Sphericity	79.062
df	15
Sig.	.000

Figure 23 KMO & Bartlett's Test – Assurance

3.13.2. Lifetime Value

The variable ‘Lifetime Value’ gives a validity of 52.311% which is considered valid and no changes are to be made.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.185	52.311	52.311	4.185	52.311	52.311
2	1.446	18.081	70.393			
3	.776	9.696	80.089			
4	.672	8.398	88.486			
5	.356	4.448	92.935			
6	.249	3.111	96.046			
7	.199	2.488	98.534			
8	.117	1.466	100.000			

Figure 24 Validity % - Lifetime Value

If later on any changes are to be made, according to the component matrix below, removing the 8th question could increase the loading % by a relative amount.

Component Matrix^a

	Component
	1
Lifetime_Value_1	.733
Lifetime_Value_2	.707
Lifetime_Value_3	.687
Lifetime_Value_4	.825
Lifetime_Value_5	.883
Lifetime_Value_6	.700
Lifetime_Value_7	.654
Lifetime_Value_8	.545

Figure 25 Component Matrix - Lifetime Value

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.734
Bartlett's Test of Sphericity	128.327
df	28
Sig.	.000

Figure 26 KMO & Bartlett's Test - Lifetime Value

3.13.3. Satisfaction

The variable ‘Satisfaction’ gives a validity of 50.649% which is considered valid. However, as the initial reliability of the variable was low, question 3 could be removed.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.026	50.649	50.649	2.026	50.649	50.649
2	1.031	25.765	76.413			
3	.688	17.200	93.614			
4	.255	6.386	100.000			

Extraction Method: Principal Component Analysis.

Component Matrix^a

	Component
	1
Satisfaction_1	.590
Satisfaction_2	.837
Satisfaction_3	.385
Satisfaction_4	.910

Figure 27 Validity % and Component Matrix - Satisfaction (Before)

Once question 3 is removed, the cumulative loading % increases to 64.880%, which is considered valid to go forth.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.946	64.880	64.880	1.946	64.880	64.880
2	.798	26.605	91.485			
3	.255	8.515	100.000			

Figure 28 Validity % - Satisfaction (After)

Component Matrix^a

	Component
	1
Satisfaction_1	.646
Satisfaction_2	.825
Satisfaction_4	.922

Figure 29 Component Matrix - Satisfaction (After)

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.519
Bartlett's Test of Sphericity	25.107
Approx. Chi-Square	
df	3
Sig.	.000

Figure 30 KMO & Bartlett's Test – Satisfaction

3.13.4. Financial Status

The variable ‘Financial Status’ gives a validity of 59.025% which is considered valid and no changes are to be made.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.361	59.025	59.025	2.361	59.025	59.025
2	.710	17.761	76.786			
3	.470	11.743	88.529			
4	.459	11.471	100.000			

Figure 31 Validity % - Finance

If later on any changes are to be made, according to the component matrix below, removing the 4th question could increase the loading % by a relative amount.

Component Matrix^a

	Component
	1
Finance_1	.818
Finance_2	.780
Finance_3	.806
Finance_4	.659

Figure 32 Component Matrix - Financial Status

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.766
Bartlett's Test of Sphericity	64.948
Approx. Chi-Square	6
df	.000
Sig.	

Figure 33 KMO & Bartlett's Test - Financial Status

3.13.5. External Influences

The variable ‘External Influences’ gives a validity of 46.826% which is not considered valid.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.873	46.826	46.826	1.873	46.826	46.826
2	.985	24.623	71.449			
3	.588	14.699	86.148			
4	.554	13.852	100.000			

Extraction Method: Principal Component Analysis.

Component Matrix^a

	Component
	1
External_1	.687
External_2	.758
External_3	.500
External_4	.759

Figure 34 Validity % and Component Matrix – External Influences (Before)

Considering the above component matrix, removing question 3 brings up the value to 58.091%, which is considered valid, and this value can be taken forth.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.743	58.091	58.091	1.743	58.091	58.091
2	.701	23.375	81.467			
3	.556	18.533	100.000			

Extraction Method: Principal Component Analysis.

Component Matrix^a

	Component
	1
External_1	.760
External_2	.804
External_4	.719

Figure 35 Validity % and Component Matrix – External Influences (After)

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.638
Bartlett's Test of Sphericity	24.407
df	3
Sig.	.000

Figure 36 KMO & Bartlett's Test - External Influences

3.13.6. Retention

The variable ‘Retention’ gives a validity of 86.911% which is considered valid and no changes are to be made.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.607	86.911	86.911	2.607	86.911	86.911
2	.291	9.715	96.625			
3	.101	3.375	100.000			

Figure 37 Validity % - Retention

Component Matrix^a

	Component
	1
Retention_1	.896
Retention_2	.939
Retention_3	.961

Figure 38 Component Matrix -Retention

The KMO and Bartlett test is also considered appropriate as the KMO value is over 0.5 and the Bartlett test gives a significance value below 0.05.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.714
Bartlett's Test of Sphericity	69.678
df	3
Sig.	.000

Figure 39 KMO & Bartlett's Test - Retention

3.14. Chapter Summary

This chapter has covered the research methodologies, philosophies and processes carried out by the researcher. It also contains the hypothesis and operationalization formulation, as well as goes in depth of the pilot study and the analysis derived in order to improve the structuring and formulation of questions for the final questionnaire, in detail.

4. CHAPTER 04 – ANALYSIS AND FINDINGS

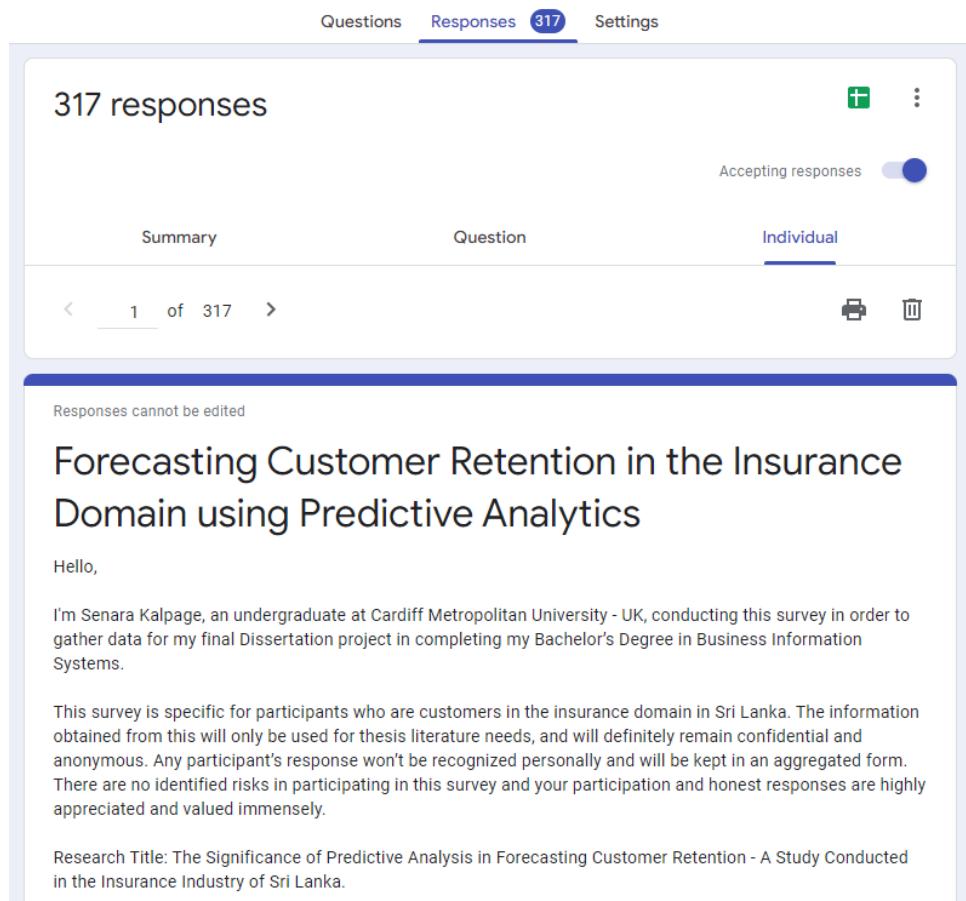
4.1. Introduction

This chapter comprises the findings of the research, gathered using primary data collection methods; distributing the final survey to the target audience of the study (Appendix 02). The upcoming sections will thereby discuss the process of which the data collected was analyzed to examine the reliability and validity of variables, while also diving into the descriptive and statistical analysis that were conducted. Findings and conclusions are thereby drawn and presented against the formulated hypothesis testing.

The final questionnaire was structured following the findings and results of the pilot study already discussed in the previous chapter. Once necessary updates were made to improve the reliability and validity of the variables, the questionnaire was distributed among the target audience; current insurance holders of Sri Lanka, initially apprehending the respondent demographics, followed by questions based on a Likert scale to measure the attitudes and opinions of respondents.

4.2.Data Preparation

The data that was gathered through the questionnaire was initially prepared by removing records of participants who didn't give consent, as well as participants who mentioned having no current insurance policies, which reduced the collected amount of responds from 317 to 311.



The screenshot shows the Google Forms interface for managing survey responses. At the top, there are tabs for 'Questions', 'Responses' (which is selected, showing the number 317), and 'Settings'. Below the tabs, it displays '317 responses'. There is a toggle switch for 'Accepting responses' which is turned off. A navigation bar below includes 'Summary', 'Question', and 'Individual' (which is underlined). At the bottom of the main area, it says 'Responses cannot be edited'. The survey title is 'Forecasting Customer Retention in the Insurance Domain using Predictive Analytics'. The survey begins with a greeting: 'Hello,' followed by a statement: 'I'm Senara Kalpage, an undergraduate at Cardiff Metropolitan University - UK, conducting this survey in order to gather data for my final Dissertation project in completing my Bachelor's Degree in Business Information Systems.' It continues with a note about the survey's purpose: 'This survey is specific for participants who are customers in the insurance domain in Sri Lanka. The information obtained from this will only be used for thesis literature needs, and will definitely remain confidential and anonymous. Any participant's response won't be recognized personally and will be kept in an aggregated form. There are no identified risks in participating in this survey and your participation and honest responses are highly appreciated and valued immensely.' At the bottom, it specifies the 'Research Title: The Significance of Predictive Analysis in Forecasting Customer Retention - A Study Conducted in the Insurance Industry of Sri Lanka.'

Figure 40 Responses Collected

The Likert scale used was coded by the following mapping in order to convert received variables to be measurable:

5	4	3	2	1
Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree

Table 4 Primary Data Coding

4.3.Reliability Analysis

Reliability analysis makes sure that the scale in consideration consistently reflects its construct, showing that the researcher has coded the variables consistently, providing a similar meaning across items (Statistic Solutions , 2022)

The following table lists the results of the reliability that was tested of the gathered data for internal consistencies, using the Cronbach Alpha of each variable;

Variable	No. of Items/Questions	Cronbach Alpha
Assurance	6	0.867
Lifetime Value	7	0.844
Satisfaction	3	0.721
Financial Status	4	0.756
External Influences	3	0.703
Customer Retention	3	0.834

Table 5 Final Survey - Reliability Analysis

According to theories, the Cronbach Alpha values range between 0 and 1, and an alpha greater than 0.70 is considered to be better and acceptable (Statistics Solutions, 2022). Taking this into consideration, the reliability of the variables in this research are considered reliable as shown in the table above. (Refer Appendix 03 for individual reliability results).

4.4. Validity Analysis

Validity analysis assesses the extent to which variables are measured against the way it was initially designed to be measured (WordPress, 2011). The validity of the data gathered in this study was measured by the Validity Loading %, as well as the KMO measure of Sampling Adequacy and Bartlett's Test of Sphericity, and is listed in the table below;

Variable	Validity Loading %	KMO measure of Sampling Adequacy	Bartlett's Test of Sphericity
Assurance	60.706%	0.863	<0.001
Lifetime Value	51.79%	0.866	<0.001
Satisfaction	65.28%	0.619	<0.001
Financial Status	59.602%	0.767	<0.001
External Influences	62.977%	0.653	<0.001
Customer Retention	75.179%	0.720	<0.001

Table 6 Final Survey - Validity Analysis

According to (Michaela Weber, Jeanine Van Ancum, 2018), a percentage between 50 and 75 is considered to be acceptable, whereas (Vincent Nijs, 2019) states that a KMO value > 0.5 and the Bartlett significant level < 0.05 is also considered acceptable. Therefore, the author can come to the conclusion that the variables in this study can be considered valid according to the results of the data gathered, listed in the table above. (Refer Appendix 04 for individual validity results).

4.5.Data Distribution and Analysis

4.5.1.Frequency Measurement

The following chart represents the data distribution of the 311 responses that were collected through the final survey for the study.

Variable	Characteristic	Frequency	Percentage (%)
Age	Below 25	40	13
	25 – 35	126	41
	35 – 50	91	29
	Above 50	54	17
Education Level	A Levels	68	22
	Diploma/HND	65	21
	Bachelors	120	39
	Masters	47	15
	PhD	5	2
	Do not wish to specify	6	2
Employment Status	Yes	255	82
	No	56	18
No. of Household Dependencies	None	31	10
	1 or 2	165	53
	3 or More	115	37
No. of Insurance Policies Held	1 or 2	203	65
	3 or More	108	35

Table 7 Measures of Frequency

4.5.1.1. Age

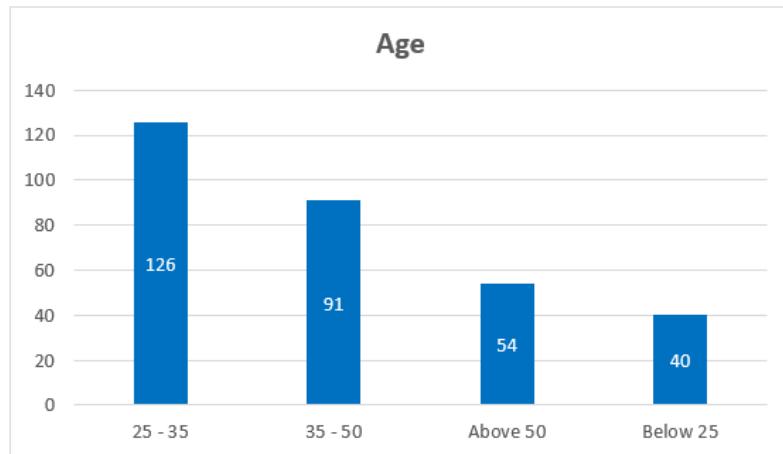


Figure 41 Age distribution

The survey was distributed among the target population, of whom a majority was between ages 25 and 35, and the minority was below 25. The majority accounted for 41% of the respondents, whereas 29% of the respondents aged between 35 and 50. The reason for the demographics to align in this manner is due to the fact that the questionnaire was distributed mainly aiming existing customers who have a proper idea of the insurance domain as well as find insurance to be of importance, namely ages 25 through 50.

4.5.1.2. Education Level

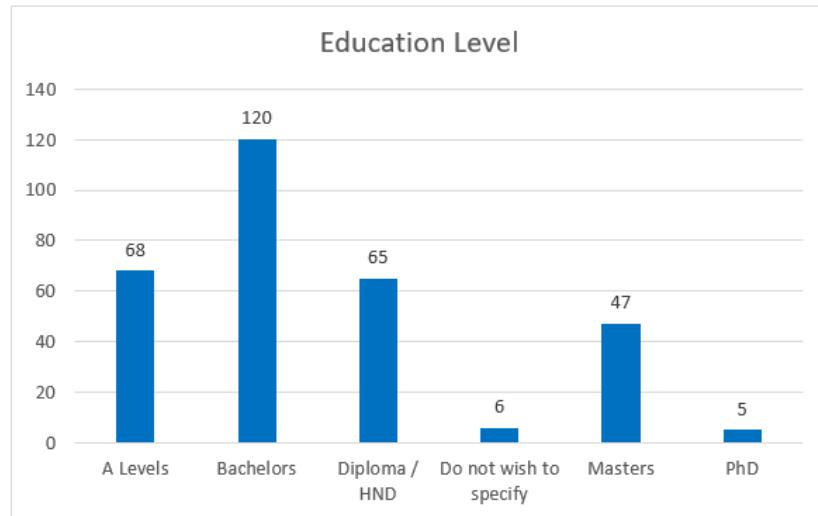


Figure 42 Education Level distribution

Most of the respondents, 120 to be exact, accounting to 39% of the participants held a Bachelor's Degree, whereas 15% to 22% of the respondents each had completed a Diploma, Masters or A Levels, thereby possessing a proper academic background. With this, it can be concluded that the respondents targeted does have a good idea of the variables, and are able to better understand the questions asked.

4.5.1.3. Employment Status

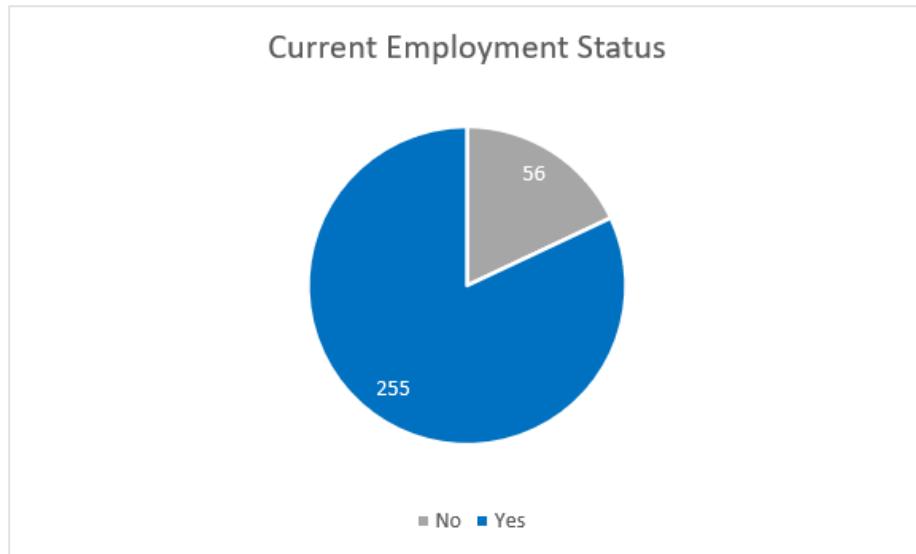


Figure 43 Employment Status distribution

Of the target respondents, 255 out of a 311 were employed at the time, accounting to 82% of the population, whereas 18% of the respondents stated to be unemployed. The questionnaire was distributed mainly targeting those employed as it would give a better understanding on how people perceive insurance considering the financial status, which is one of the variables in this study. This would thereby give a better understanding on peoples' necessities when they are considered to work hard for their income, as opposed to not.

4.5.1.4. No. of Household Dependencies

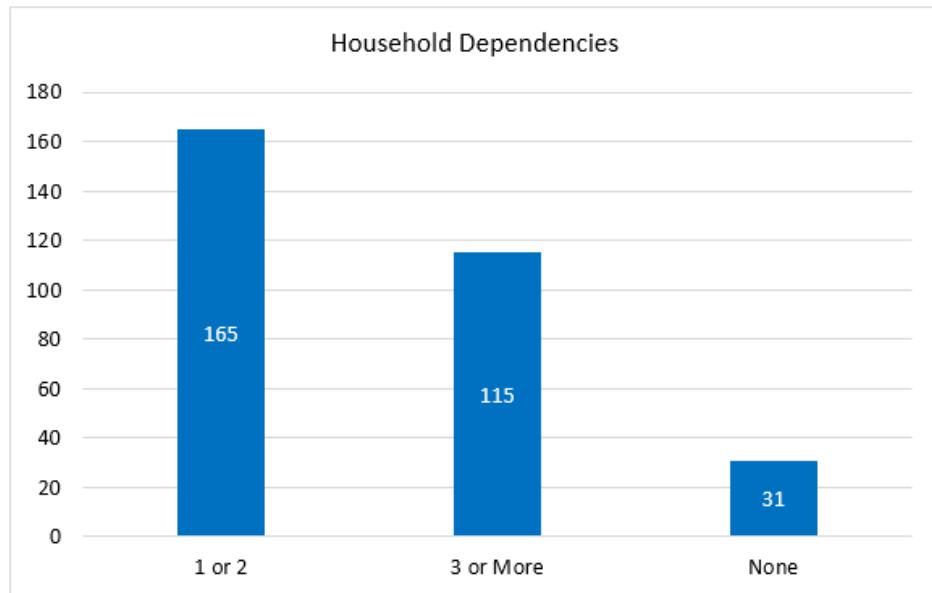


Figure 44 Household dependency distribution

Of the total respondents, 165 stated to have at least 1 dependent within their household, which accounted to 53% of the participants, whereas 37% states to have over 3 dependents. The reason for the inclusion of gathering household dependency data is due to the fact that once people have more dependents, they tend to not want a risk, and therefore look into insurance in order to mitigate or reduce the aspect of risk, which is also considered a dimension in the study.

4.5.1.5. No. of Insurance Policies Held

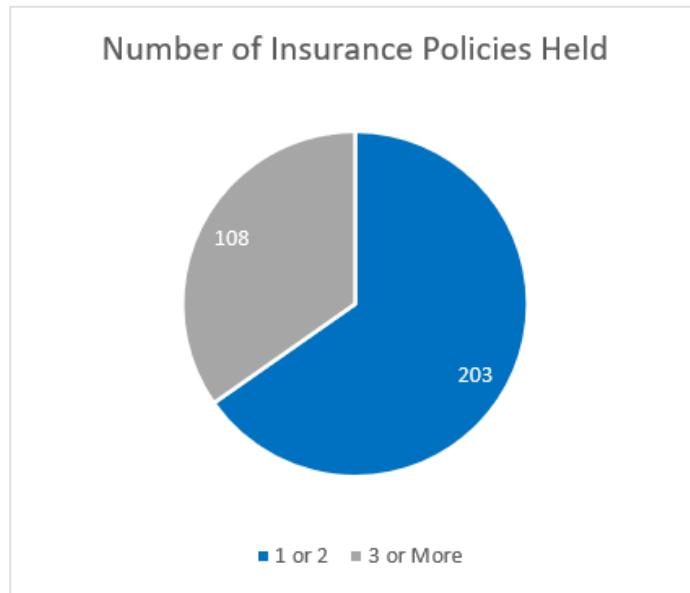


Figure 45 Number of policies held

The target population of this study is existing customers in the insurance domain of Sri Lanka. Through this survey the author was able to reach 203 individuals who held a minimum of 1 insurance policy, which amounted to 65% of the respondents, whereas the rest of the 35% had over 3 policies. According to literature, when people hold more insurance policies, there is a less likely chance of churn.

4.6.Customer Segmentation based on Age Groups – Power BI Visualization

Predictive analytics is an important tool to help companies measure responses of customers in relation to the promotional efforts undertaken, in order to understand how certain segments would react to changes (Abby Borden, 2021). Following are screenshots of a dashboard created in order to visualize the segmentation of customers using the Power BI software. This would thereby give a clear visualization of the demographic factors considered in the study, against the age group of customers.



Figure 46 Customer Segmentation based on Age - Below 25

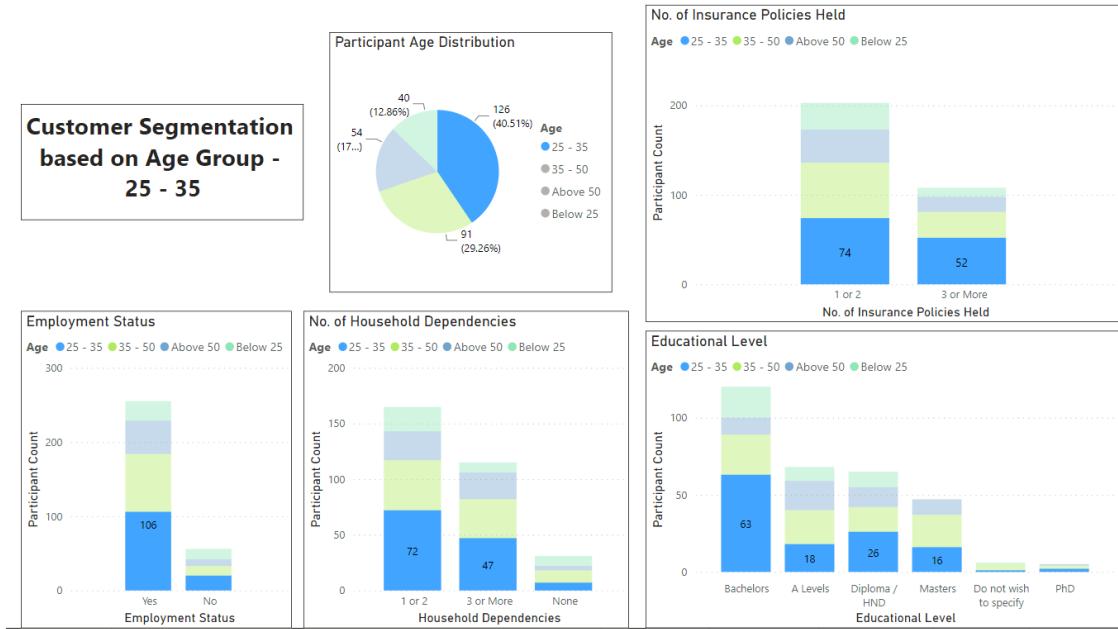


Figure 47 Customer Segmentation based on Age - 25 – 35

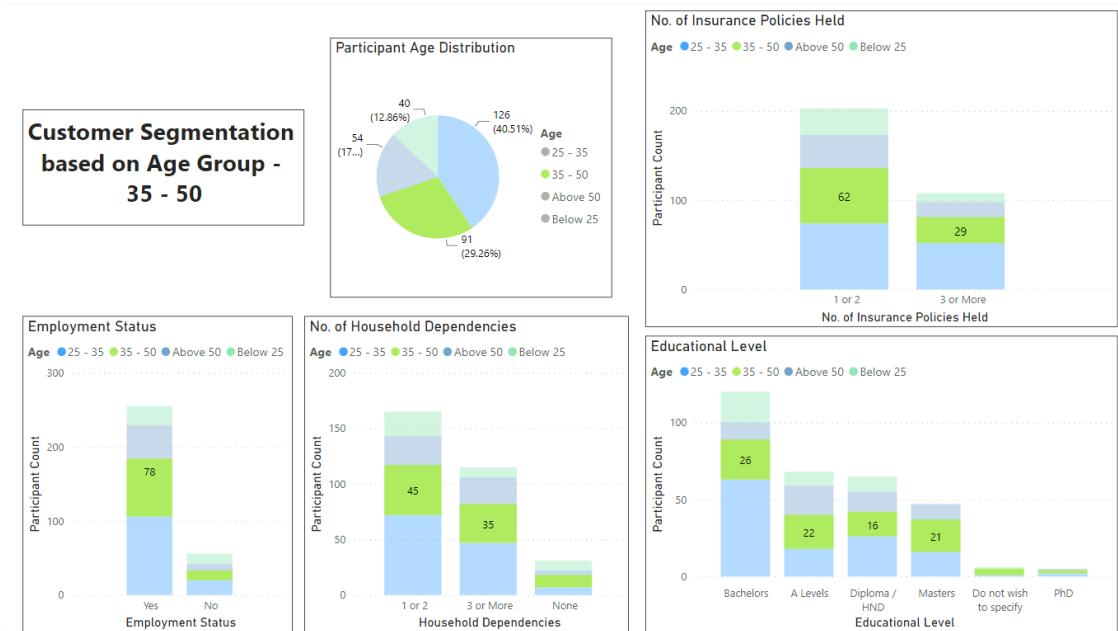


Figure 48 Customer Segmentation based on Age - 35 - 50

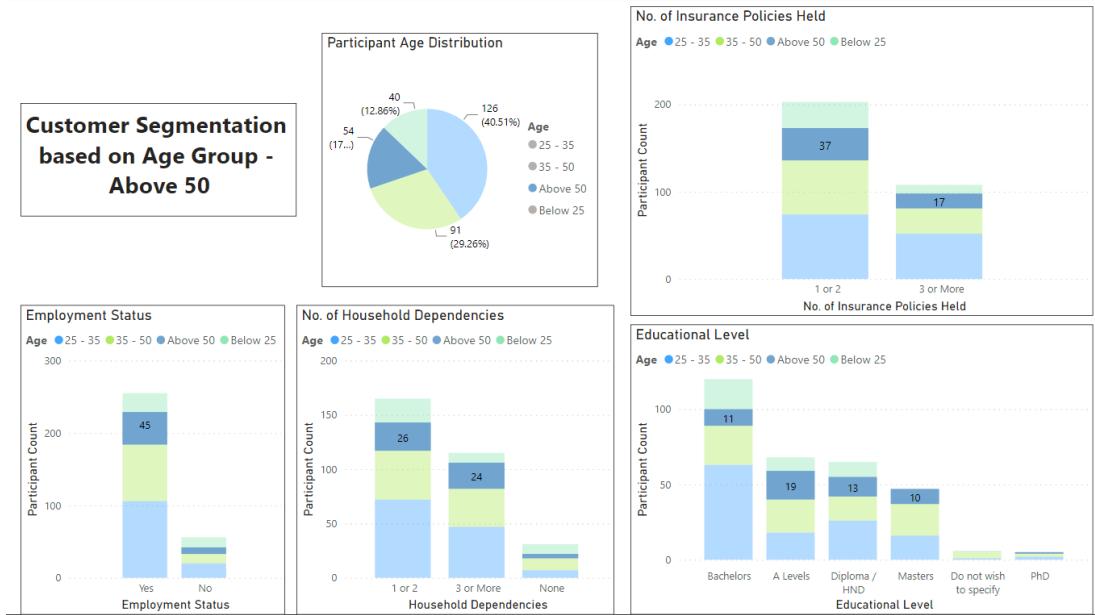


Figure 49 Customer Segmentation based on Age - Above 50

4.7.Descriptive Statistics

Descriptive Statistics

	N	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Assurance	311	3.8869	.67321	.453	-.205	.138	.088	.276
Lifetime_Value	311	4.3064	.52285	.273	-.750	.138	.889	.276
Satisfaction	311	4.2969	.57557	.331	-.500	.138	-.410	.276
Financial_Status	311	4.2267	.62157	.386	-.821	.138	1.584	.276
External_Influences	311	4.2529	.63610	.405	-.576	.138	-.454	.276
Customer_Retention	311	4.2144	.67025	.449	-.813	.138	1.162	.276
Valid N (listwise)	311							

Figure 50 Descriptive statistics

The mean statistics of the data analyzed all ranged in between 3.887 and 4.306, while the standard deviation ranged from 0.523 and 0.673. This indicates that the responses on the Likert scale was more in favour to the positive. Since the lowest mean statistic falls over 3, it can be concluded that most responds ranged from ‘Neutral’ to ‘Strongly Agree’. The only disagreeing variable to be identified here would be Assurance, as it has a mean statistic of 3.89 while having a standard deviation of 0.67. On the other hand, the skewness ranged from -1 to +1 and is thereby considered to be acceptable (Diva Dugar, 2018).

4.8.Inferential Statistics

4.9.Correlation Analysis

The correlation analysis is done through Pearson's correlation coefficient measuring the strength of linear associations between 2 variables, usually denoted by 'r', which ranges between values +1 and -1 representing a perfect positive to a perfect negative correlation. A value of 0 would thereby mean no correlation between two variables (Bianca Williams, Caroline Halloin, 2020).

Independent Variable	Dependent Variable: Customer Retention		
	Pearson Correlation	Significance (2-tailed)	N
Assurance	0.404	<0.001	311
Lifetime Value	0.572	<0.001	311
Satisfaction	0.570	<0.001	311
Financial Status	0.565	<0.001	311
External Influences	0.588	<0.001	311

Table 8 Correlation Analysis

The above table illustrates the correlation between the dependent and independent variables of the study. The hypothesis created in Chapter 03 were tested against the correlations, with scatter plots generated to support the understanding of the relationships. (Refer Appendix 05 for individual correlation results).

4.9.1.Hypothesis 1C

For the variable Assurance, with the data gathered through the questionnaire, a correlation coefficient of 0.404 has been identified with a significance level <0.001. With this, the null hypothesis can be rejected with a 99% confidence level in favour of the alternate hypothesis. Thereby Assurance can be concluded to have a weak positive relationship with the dependent variable; Customer Retention.

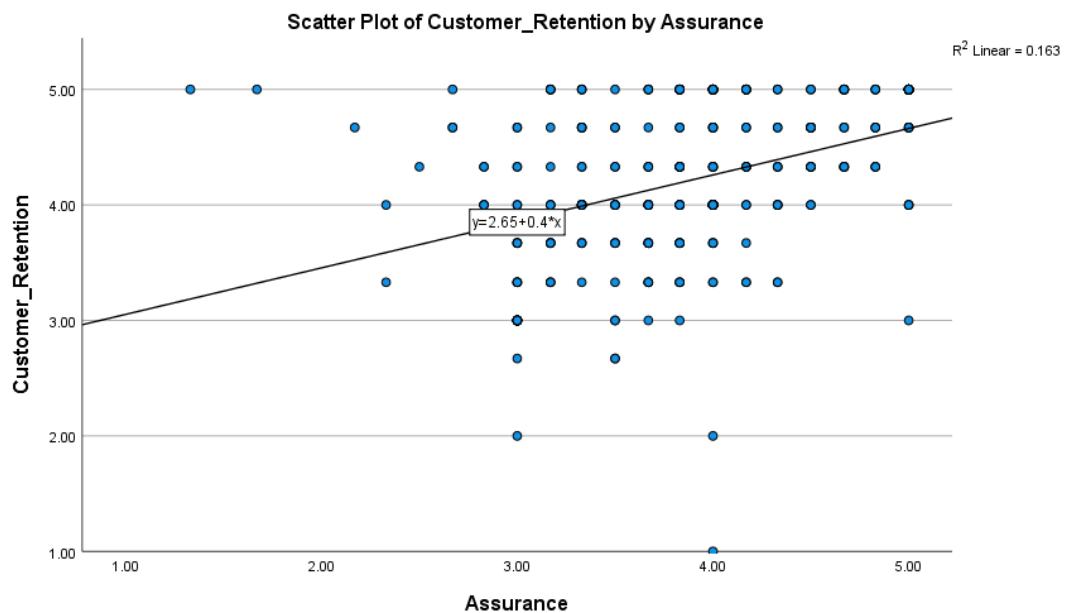


Figure 51 Scatter Plot - Assurance

4.9.2.Hypothesis 2C

The variable Lifetime Value, with the data gathered through the questionnaire, shows a correlation coefficient of 0.572 with a significance level <0.001. With this, the null hypothesis can be rejected with a 99% confidence level in favour of the alternate hypothesis. Thereby Lifetime Value can be concluded to have a strong positive relationship with the dependent variable; Customer Retention.

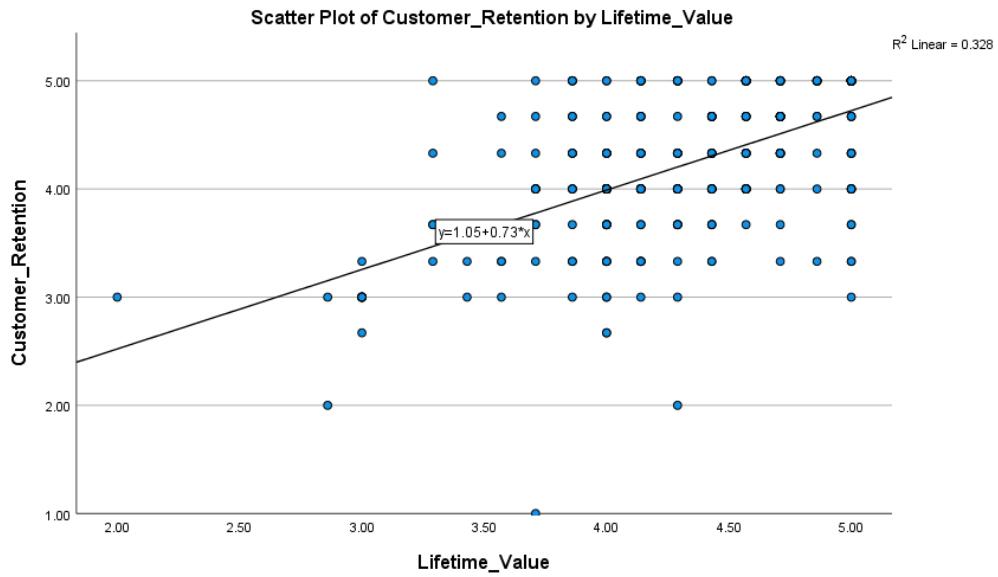


Figure 52 Scatter Plot - Lifetime Value

4.9.3.Hypothesis 3C

When considering the correlation of Satisfaction, a coefficient of 0.570 has been identified with a significance level <0.001. This results in rejecting the null hypothesis, in favour of the alternate hypothesis. As a result, Satisfaction can be concluded to have a strong positive relationship with Customer Retention, meaning, the higher the Satisfaction of the customer, the higher the levels of retention within the company.

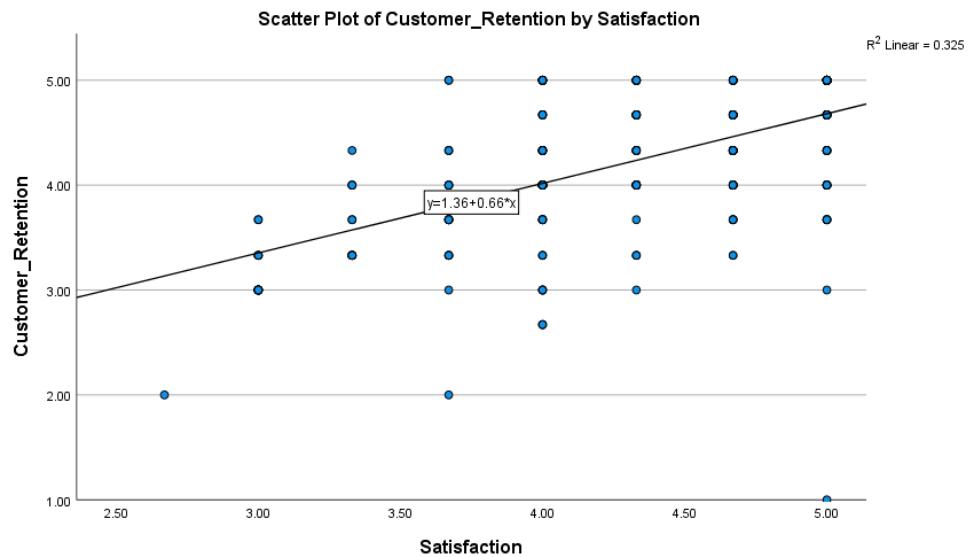


Figure 53 Scatter Plot - Satisfaction

4.9.4.Hypothesis 4C

With the data gathered, it has been identified that the Financial Status of a customer has a positive correlation with Customer Retention. This is proven with a correlation coefficient of 0.565 and a significant <0.001, with a 99% confidence level. With this, the null hypothesis can thereby be rejected.

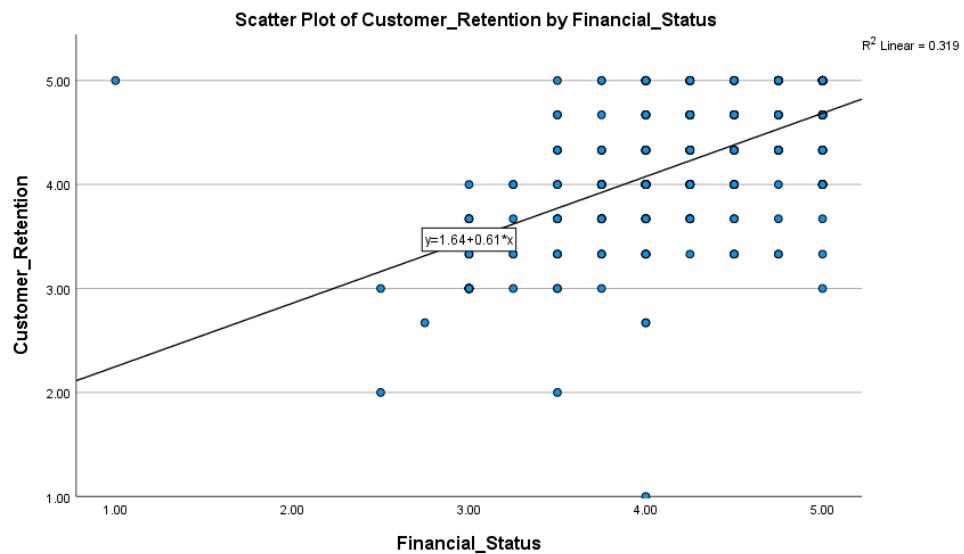


Figure 54 Scatter Plot - Financial Status

4.9.5.Hypothesis 5C

External Influences shows a correlation coefficient of 0.588 with its significance falling <0.001 at a confidence level of 99%. This shows a strong positive correlation between the variable and Customer Retention, meaning the null hypothesis can be rejected as it more favours the alternate hypothesis.

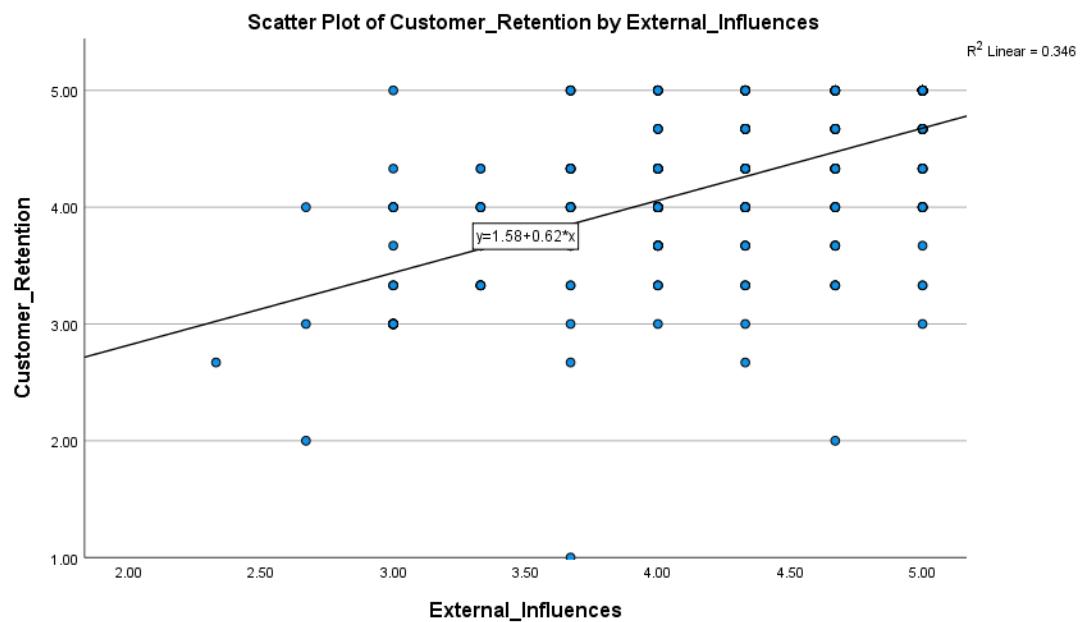


Figure 55 Scatter Plot - External Influences

4.10. Simple Linear Regression Analysis

Simple linear regression estimated the association between two variables, which are usually quantitative. It shows the strength of the relationship, as well as depicts the value changes of the dependent variable, with a change in value of the independent variable (Rebecca Bevans, 2022).

Independent Variable	Dependent Variable: Customer Retention					
	Model Summary		ANOVA		Coefficients	
	R	R Square	F	Sig	Constant	Variable
Assurance	0.404	0.163	60.341	<0.001	2.650	0.402
Lifetime Value	0.572	0.328	150.646	<0.001	1.054	0.734
Satisfaction	0.570	0.325	148.968	<0.001	1.361	0.664
Financial Status	0.565	0.319	144.903	<0.001	1.639	0.609
External Influences	0.588	0.346	163.145	<0.001	1.580	0.619

Table 9 Regression Analysis

(Refer Appendix 06 for individual regression results – RSR Histograms and Normal P-Plots).

4.10.1. Hypothesis 1R

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.404 ^a	.163	.161	.61405	.163	60.341	1	309	.000

a. Predictors: (Constant), Assurance

b. Dependent Variable: Customer_Retention

Figure 56 Model Summary - Assurance

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	22.752	1	22.752	60.341	.000 ^b
Residual	116.513	309	.377		
Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), Assurance

Figure 57 ANOVA Test - Assurance

Coefficients^a

Model	Unstandardized Coefficients		Beta	t	Sig.
	B	Std. Error			
1 (Constant)	2.650	.204	.404	12.969	.000
	.402	.052		7.768	.000

a. Dependent Variable: Customer_Retention

Figure 58 Regression Coefficients - Assurance

According to the model summary the variable Assurance shows a R^2 value of 16.3%, whereas the ANOVA analysis shows a F statistic of 60.341 and a significance alpha of 0.000. Due to the alpha being <0.001 at a confidence level of 99%, the null hypothesis can be rejected. Moreover, a slope of +0.402 can also be identified in the coefficients, signifying a proving that Assurance does have an effect on Customer Retention. The following formula can thereby be derived:

$$\text{Customer Retention} = 2.650 + (0.402 * \text{Assurance})$$

4.10.2. Hypothesis 2R

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.572 ^a	.328	.326	.55044	.328	150.646	1	309	.000

a. Predictors: (Constant), Lifetime_Value
b. Dependent Variable: Customer_Retention

Figure 59 Model Summary - Lifetime Value

ANOVA ^a					
Model		Sum of Squares	df	Mean Square	F
1	Regression	45.643	1	45.643	150.646
	Residual	93.622	309	.303	
	Total	139.265	310		

- a. Dependent Variable: Customer_Retention
b. Predictors: (Constant), Lifetime_Value

Figure 60 ANOVA Test - Lifetime Value

Model	Unstandardized Coefficients			t	Sig.
	B	Std. Error	Beta		
1	(Constant)	1.054	.259	4.063	.000
	Lifetime_Value	.734	.060	.572	12.274

- a. Dependent Variable: Customer_Retention

Figure 61 Regression Coefficients - Lifetime Value

According to the ANOVA test, the variable Lifetime Value with Customer Retention has a F statistic of 150.646 and significance of 0.000, which is <0.001. At a 99% confidence level, the null hypothesis can thereby be rejected in favour of the alternate hypothesis. Moreover, the model summary of the regression test also shows that the Lifetime Value has a 32.8% effect on the dependent variable, proved through the R², with a coefficient slope of +0.734 to prove the alternate hypothesis. The following formula can thereby be derived:

$$\text{Customer Retention} = 1.054 + (0.734 * \text{Lifetime Value})$$

4.10.3. Hypothesis 3R

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.570 ^a	.325	.323	.55145	.325	148.968	1	309	.000

a. Predictors: (Constant), Satisfaction

b. Dependent Variable: Customer_Retention

Figure 62 Model Summary - Satisfaction

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	45.300	1	45.300	148.968	.000 ^b
Residual	93.965	309	.304		
Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), Satisfaction

Figure 63 ANOVA Test - Satisfaction

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	1.361	.236		5.768	.000
Satisfaction	.664	.054	.570	12.205	.000

a. Dependent Variable: Customer_Retention

Figure 64 Regression Coefficients - Satisfaction

The model summary of the regression test reveals an R^2 of 32.5%, showing the percentage by which, the variable Satisfaction has an effect on the dependent variable Customer Retention. With a F statistic of 148.968, and a significance alpha <0.001, the null hypothesis can thereby be rejected at a 99% confidence level in favour of the alternate hypothesis. Moreover, a slope of +0.664 can also be seen, proving that Satisfaction does have an impact on the dependent variable.

$$\text{Customer Retention} = 1.361 + (0.664 * \text{Satisfaction})$$

4.10.4. Hypothesis 4R

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.565 ^a	.319	.317	.55391	.319	144.903	1	309	.000

a. Predictors: (Constant), Financial_Status
b. Dependent Variable: Customer_Retention

Figure 65 Model Summary - Financial Status

ANOVA ^a					
Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.459	1	44.459	144.903
	Residual	94.806	309	.307	
	Total	139.265	310		

- a. Dependent Variable: Customer_Retention
b. Predictors: (Constant), Financial_Status

Figure 66 ANOVA Test - Financial Status

Model	Unstandardized Coefficients			Standardized Coefficients	t	Sig.
	B	Std. Error	Beta			
1	(Constant)	1.639	.216	.565	7.581	.000
	Financial_Status	.609	.051		12.038	.000

- a. Dependent Variable: Customer_Retention

Figure 67 Regression Coefficients - Financial Status

According to the ANOVA test, the variable Financial Status with Customer Retention has a F statistic of 144.903 and significance of 0.000, which is <0.001. Therefore, at a 99% confidence level, the null hypothesis can be rejected in favour of the alternate hypothesis. Moreover, the model summary of the regression test also shows that the Financial Status has a 32% effect on the dependent variable, proved through the R², with a coefficient slope of +0.609 to prove the alternate hypothesis.

$$\text{Customer Retention} = 1.639 + (0.609 * \text{Financial Status})$$

4.10.5. Hypothesis 5R

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.588 ^a	.346	.343	.54310	.346	163.145	1	309	.000

a. Predictors: (Constant), External_Influences

b. Dependent Variable: Customer_Retention

Figure 68 Model Summary - External Influences

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	48.121	1	48.121	163.145	.000 ^b
Residual	91.143	309	.295		
Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), External_Influences

Figure 69 ANOVA Test - External Influences

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	1.580	.209		7.578	.000
External_Influences	.619	.048	.588	12.773	.000

a. Dependent Variable: Customer_Retention

Figure 70 Regression Coefficients - External Influences

The model summary of the regression test reveals an R^2 of 34.6%, showing the percentage by which, the variable External Influences has an effect on the dependent variable Customer Retention. With a F statistic of 163.145, and a significance alpha <0.001, the null hypothesis can be rejected at a 99% confidence level in favour of the alternate hypothesis. Moreover, a slope of +0.619 can also be seen, proving that External Influences does have an impact on the dependent variable.

$$\text{Customer Retention} = 1.580 + (0.619 * \text{External Influences})$$

4.11. Multiple Regression Analysis

Multiple regression is a statistical technique that can be taken for the purpose of analyzing and determining the relationship between a number of independent variables, to a dependent variable. The purpose of using this type of an analysis is to therefore identify the independent variables that tend to predict/have a particular effect on the dependent variable (Andrew W. Moore, 2006).

According to Chapter 03, the multiple regression hypothesis is tested below;

4.11.1. Hypothesis 1MR

The initial hypothesis consisted of testing the relationship of all independent variables against the dependent variable Customer Retention in order to get the following results;

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.689 ^a	.474	.466	.48997	.474	55.019	5	305	.000

a. Predictors: (Constant), External_Influences, Assurance, Financial_Status, Lifetime_Value, Satisfaction

b. Dependent Variable: Customer_Retention

Figure 71 Model Summary - 1MR

ANOVA ^a					
Model		Sum of Squares	df	Mean Square	F
1	Regression	66.043	5	13.209	55.019
	Residual	73.222	305	.240	
	Total	139.265	310		

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), External_Influences, Assurance, Financial_Status, Lifetime_Value, Satisfaction

Figure 72 ANOVA Test - 1MR

Model	Coefficients ^a				
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	.212	.252		.840	.401
Assurance	.074	.049	.074	1.514	.131
Lifetime_Value	.218	.079	.170	2.759	.006
Satisfaction	.152	.074	.131	2.045	.042
Financial_Status	.200	.068	.185	2.940	.004
External_Influences	.301	.059	.285	5.121	.000

a. Dependent Variable: Customer_Retention

Figure 73 Regression Coefficients - IMR

According to the results, the model summary for the multiple regression gives a R^2 of 47.4%. the rest of the 52.6% is considered to be variables other than the ones selected, that could have an effect on Customer Retention. However, according to the coefficient test, the variables Assurance and Satisfaction have the highest significant values of 0.131 and 0.042 respectively. Due to this not achieving a 99% confidence level, the author, as mentioned in Chapter 03 will be trying different combinations of variables in order to get the most accurate and significant statistical model. The following changes were identified once different combinations were tested;

Variable Deleted	Hypothesis	Improved R Squared	Improved F statistic
None deleted	1MR	47.4%	55.019
Assurance only (Appendix 07 – Section 7.7.1)	2MR	47.0%	67.914
Satisfaction only (Appendix 07 – Section 7.7.2)	3MR	46.7%	67.030
Both Assurance and Satisfaction deleted (Appendix 07 – Section 7.7.3)	4MR	46.3%	88.066

Table 10 MR Hypothesis Test Results

After taking into consideration the results of the different multiple regression tests, the author has decided to go ahead with the initial test. This is due to the test having the highest R^2 , and although a few variables had a significance slightly >0.001 , the overall

significance of the model according to the ANOVA and model summary is 0.000, which is acceptable and considered significant. The following formula for multiple regression can thus be derived;

$$\text{Customer Retention} = 0.212 + (0.074 * \text{Assurance}) + (0.218 * \text{Lifetime Value}) + \\ (0.152 * \text{Satisfaction}) + (0.200 * \text{Financial Status}) + (0.301 * \text{External Influences})$$

4.12. Discussion On Findings

4.12.1. Correlation Analysis

The following table shows a summary of the correlation analysis done on the study.

Variable	Hypothesis	Coefficient	Significance	Result
			Alpha	
Assurance	1C	0.404	<0.001	Accepted
Lifetime Value	2C	0.572	<0.001	Accepted
Satisfaction	3C	0.570	<0.001	Accepted
Financial Status	4C	0.565	<0.001	Accepted
External Influences	5C	0.588	<0.001	Accepted

Table 11 Correlation Hypothesis - Summary

According to the data gathered in the study and the correlation analysis conducted, all 5 variables are found to have a positive correlation with the dependent variable Customer Retention, with a significance level <0.001. As a result, the correlations of all 5 variables are accepted, whereas the null hypothesis of the variables have been rejected.

The variable External Influences has been identified with the highest correlation with a coefficient of 0.588, denoting a strong positive relationship with the dependent variable. The main dimensions captured in this variable was the competitor influences as well as inflation rates within the country. According to the data gathered its thereby understood that people are more affected by these factors, which could lead to a low retention rate within the company. On the other hand, the variable with the lowest correlation identified was Assurance, with a coefficient of 0.404. The variable is thereby considered to have a weak positive relationship with the dependent variable, meaning an increase in this variable will only increase retention by a small amount.

The other 3 variables, namely Satisfaction, Financial Status and External Influences have a correlation coefficient > +0.5, implying a positive correlation. This therefore shows that the variables have a considerable amount of impact on the dependent variable, and as a result are all accepted.

4.12.2. Regression Analysis

The following table shows a summary of the regression analysis done on the study.

Variable	Hypothesis	R Square	F Statistics	Significance Alpha	Result
Assurance	1R	0.163	60.341	<0.001	Accepted
Lifetime Value	2R	0.328	150.646	<0.001	Accepted
Satisfaction	3R	0.325	148.968	<0.001	Accepted
Financial Status	4R	0.319	144.903	<0.001	Accepted
External Influences	5R	0.346	163.145	<0.001	Accepted

Table 12 Regression Hypothesis - Summary

With light to the regression analysis conducted with the data gathered from the study, all 5 variables are found to have a significant amount of impact on Customer Retention, and thereby are accepted with a significance alpha <0.001. As a result, the null hypothesis of all 5 variables have been rejected in favour of the alternate hypothesis.

When considering the results of the analysis, External Influences yet again has the highest impact on the dependent variable. This is implied by the R² value, signifying that the variable has a 34.6% impact on Customer Retention. The variable also has the highest F statistic value of 163.145.

On the other hand, Assurance has the lowest influence on the dependent variable, with a low R² value of 16.3%, and a lower F statistic of 60.341. Lifetime Value, Satisfaction and the Financial Status of customers thereby are shown to have an impact of over 30% on the dependent variable.

4.12.3. Multiple Regression Analysis

The table below illustrates in summary the Multiple Regression Analysis done on the study.

Variable Deleted	Hypothesis	Improved R Squared	Improved F statistic	Result
None deleted	1MR	47.4%	55.019	Accepted
Assurance only (Appendix 07 – Section 7.7.1)	2MR	47.0%	67.914	Rejected
Satisfaction only (Appendix 07 – Section 7.7.2)	3MR	46.7%	67.030	Rejected
Both Assurance and Satisfaction deleted (Appendix 07 – Section 7.7.3)	4MR	46.3%	88.066	Rejected

Table 13 Multiple Regression Hypothesis - Summary

The multiple regression analysis was carried out several times intending to find the best combination of variables that have the highest significance on the dependent variable in order to create a more accurate statistical model.

Throughout testing, the highest R^2 value recorded was at 47.4%, with all 5 variables in consideration. Variables Assurance and Satisfaction had a significance value slightly >0.001 , resulting in the different testing combinations in order to increase significance. During this it was identified that removing either variables, or both variables would result in a decrease in the R^2 value by over 1% although an increase in the F statistic has been recorded.

Therefore, in order to maintain the significance, the test with the highest R^2 value (test 01) was accepted and the other tests were rejected. The 2 variables in question, although having a higher significant value does not impose an issue in finalizing the test result as the overall model significance resulted as 0.000, which is <0.001 , and thereby accepted.

$$\text{Customer Retention} = 0.212 + (0.074 \times \text{Assurance}) + (0.218 \times \text{Lifetime Value}) + (0.152 \times \text{Satisfaction}) + (0.200 \times \text{Financial Status}) + (0.301 \times \text{External Influences})$$

4.13. Multiple Regression – Predicted Model (Power Bi Visualization)

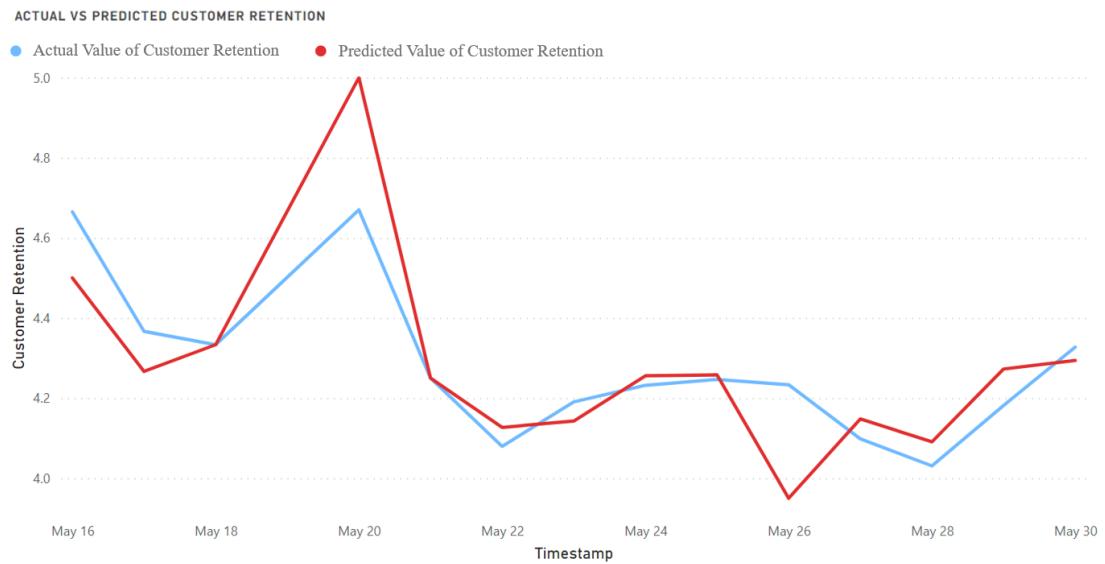


Figure 74 Predicted Model Visualization - Actual Vs Predicted Customer Retention

In regression analysis, several multiple regression hypotheses were tested, and considering the outcome of the analysis, the first statistical hypothesis (1MR) was accepted, and the predictive model for the hypothesis was deemed to be the most suitable statistical model in predicting Customer Retention as it contains all variables that have a higher significance on the dependent variable. From the predictive model, the author was able to derive the predictive value of Customer Retention. The above graph portrays the comparison of the actual values vs predictive values. The actual values are taken from the data gathered from the survey. Since the 2 patterns are pretty identical, it again proves the fact that the selected predictive model is suitable in order to forecast Customer Retention.

4.14. Chapter Summary

This summary consists of the data analytics of the research from the point of data preparation, to conducting various analysis such as testing the validity, reliability, correlation, regression and multiple regression of each individual variable with all the data gathered throughout the research. The findings have been analyzed and discussed in this chapter, while also taking into consideration the demographic attributes of the respondents.

5. CHAPTER 05 – CONCLUSIONS AND RECOMMENDATIONS

5.1. Review Of Research Objectives

The main research question and purpose of the study was to recognize the main factors that have an impact on customer retention. After the data gathering and conducting analysis of the variables (validity, reliability, correlation and regression), the variables with a higher significance value was identified. The below variables were identified to have a strong positive correlation with customer retention rates of the company, with a correlation coefficient $> +0.5$, and therefore the conceptual framework can be updated as follows;

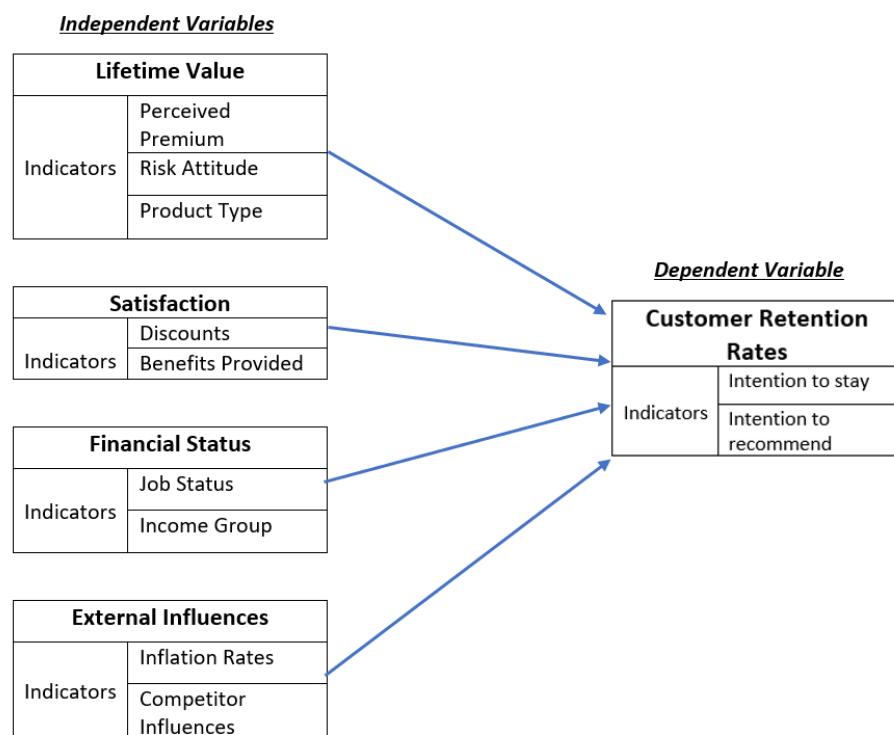


Figure 75 Improved Conceptual Framework

The following table discusses the areas of the research problems that have been discovered throughout the research.

Question No.	Research Question	Conclusion
01	What are the main factors that influence customer retention?	According to the analysis, the main factors that affect and influence customer retention have been identified as Lifetime Value, Satisfaction, Financial Status and External Influences, with a coefficient greater than 0.5, also achieving objective 01 (Refer table 15 – Review on Research Objectives). This indicates a strong positive correlation between the variables and Customer Retention. The variable with the greatest influence on the dependent variable was thus identified as External Influences, indicating that the inflation rates and competition does have a higher impact on people's perceptions toward insurance, answering the first research question.
02	What are reasons that cause a high churning rate?	When considering the individual questions, reasons for high churning rates have also been identified. According to a question asked in the variable Financial Status, individuals would most find it difficult to pay insurance premiums if at all they lose their sources of income, whereas 77% of the respondents also stated that choosing of an insurance policy would most likely depend on the financial status (Appendix 08). Therefore, a negative change in the job status of an individual could lead to a higher churn rate in the company. On the other hand, 89% of the respondents also admitted to wanting a change in the insurance policy if they find the product unsatisfactory (Appendix 08). Therefore, the type of product is also

		<p>considered significant, which could highly impact churn rates within the company.</p> <p>Moreover, 80% of the respondents also reconsiders necessities with regards to increases in inflation within the economy. Due to this, it can be understood that there is a high chance policy that require higher premiums will lead to higher churns. This is backed up by 88% of the respondents who state that affordable premiums and benefits would increase their satisfaction, leading to lower churn rates (Appendix 08).</p>
03	What actions can the company take in order to retain customers?	Answered in recommendations (Section 5.3)
04	What benefits will predictive analysis bring into companies in the insurance sector?	Answered in recommendations (Section 5.3)
05	How much of an impact do external factors have on retention rates?	A particular crucial research questions attempted in the study happened to be, to find how much of an effect the external influences would have on customer retention. This is an important aspect to be considered, as authors such as (Stefan Mau, 2017) and (Gunther, 2014) signifies the importance of having to consider external aforementioned external influences in future research as it has an increasing effect on churn. According to the analysis, it was identified that this

		variable was one that had the highest coefficients; 0.588, with the strongest positive correlation with the variables considered, and the highest R^2 recorded of 34.6%. It was thereby identified that an increase in factors such as economic inflation and competitiveness, there is a high chance for churning rates of the company to increase. With this analysis the author was thereby able to address the existing research gap on considering external influences when formulating a predictive model.
--	--	--

Table 14 Review of Research Questions

The following section uncovers the areas of the research objectives that have been discovered throughout the research.

Objective No.	Research Objective	Conclusion
01	Identify key factors that have a significant impact on customer retention in the insurance industry of Sri Lanka.	Refer Research Question 01 above (Review of Research Questions)
02	Conduct proper customer segmentation	When conducting the research, the author was able to segment the target population into the following segments; <ul style="list-style-type: none"> - Age - Education Level - Employment Status - Number of household dependencies - Number of insurance policies held

		With this segmentation, the author was able to identify the impact of the variables on Customer retention according to the demographics of the respondents. This was visualized through a Power BI dashboard in Section 4.6, where the demographics were segmented according to the participant age groups.
03	Feed data gathered into a predictive model that would portray the anticipated change in customer retention in correlation with the variables identified.	The data gathered was analyzed in the previous chapter (Chapter 04), and variables with the highest correlation with the dependent variable was identified. All independent variables were identified to have a strong positive correlation, with coefficients $> +0.5$, except the variable Assurance which had a weak positive relationship with a coefficient of 0.404 with the dependent variable.
04	Utilize this information in order to make informed decisions and provide business information solutions to companies.	Will be discussed in recommendations (Section 5.3)
05	Create a visual dashboard to display the final output using Power BI	Dashboards were created using Power BI to visualize; <ul style="list-style-type: none"> - Customer segmentation based on age groups (Section 4.6) - Actual Vs Predictive Customer Retention values (Section 4.13)

Table 15 Review of Research Objectives

5.2. Research Implications and Contributions

This research would be of use for insurance companies in an effort to get a comprehensive understanding of the significant factors that would affect their customer retention. Companies can thereby use the statistical model created in order to feed data as required in order to get predictions and results according to the data they possess. The findings of the research would moreover aid in recognizing the behaviour of clients, specific to the insurance domain in Sri Lanka, which lacks literature in the current context. It would thereby also aid in future research and give a basic idea of the conception to make further improvements as required necessary.

5.3. Recommendations

In accordance to the analysis conducted in previous chapters, it has been identified that the variables with the highest significance are Lifetime Value, Satisfaction, Financial Status and External Influences, with a strong positive correlation. This thereby confirms that these variables have a higher impact on the dependent variable; Customer Retention. Insurance companies in Sri Lanka can use this information in order to make a positive impact in various areas of the business, mainly making use of the technology of predictive analytics.

With relation to the data gathered, it can be understood that the type of insurance products is highly significant to a customer, with 89% of the respondents agreeing that they would most likely change policies if they are unsatisfied with the type of policy, whereas 88.7% of the respondents preferred being more aware of the type of policy required (Appendix 08). The author (Stefan Mau, 2017) also identifies from his research that 90% of customers who already churned cannot be brought back, and therefore, (Verhoef & Venkatesan, 2010) reveals PA modelling could help companies contact “the right customer, with the right offer, at the right time”.

An important step insurance companies can be recommended to take is to acknowledge the needs and interests of the customer, and to thereby offer a personalized service. This is supported through the customer segmentation, in order to analyze the customer lifestyle. In this way the companies would be able to recognize the risk-taking attitudes

of the customers, and predict how changes in the product and premium amounts would affect customer retention, and then revenue. Lifestyle details such as whether or not a customer consumes alcohol, or even partake in various sports would be indicators of determining the best product for the customer (Yogesh Dhavale, 2020). Once these aspects are thought of and fed into a predictive model, it would help companies improve the lifetime period and value, taking into consideration in this study. This would be a step further that can be taken, rather than making assumptions based on basic demographic factors, in order to cater to specific needs. **Buyer-focused marketing strategies** can thus be implemented in order to accommodate to unmet needs. (Abby Borden, 2021) identifies predictive analytics to thereby be able to help companies measure responses of customers in relation to the promotional efforts undertaken, in order to understand how certain segments would react to changes. Insurance companies can in this way segment the customers based on how likely they are to repurchase within the upcoming month, and thus tailor marketing strategies to best fit each segment.

(Yogesh Dhavale, 2020) also states that companies can adopt **cross-selling strategies** in order to reduce acquisition costs low, by selling additional policies to existing customers according to their needs, saving companies a lot of time that could otherwise be invested in other crucial functions. According to the study conducted around 91% of the respondents agreed to be conscious of the insurance coverage, and therefore, it would be best if companies are able to keep a close consideration on this aspect when adopting cross-selling strategies. Predictive analysis would aid in taking into account the past behaviour of customers, as well as the behaviour of similar customers in order to assist agents with suggesting certain products, or even identifying dissatisfaction of customers which could help reduce over-touching of customers. This can also be done with the aid of **collaborative filtering**, to provide personalized recommendations, and this combines with machine learning would help companies make more accurate churn predictions. Once customer journeys are mapped along with their lifecycle, insurers would be able to make connections on when customers are led to abandon policies, and work their way around the issue to not have it be a problem in the future again, thereby increasing satisfaction of customers while saving costs and maintaining a proper customer base.

With intentions to analyze the success of the business information decisions taken by the company, various **KPIs** such as the “**Net Promoter Index (NPI)**”, **Overall Satisfaction Index and Brand Equity**. Aforementioned KPIs would relatively aid in measuring customer referrals, as to how many customers would promote or recommend the brand to others, or even how the satisfaction levels of customers have changed comparatively, against previous products and competitors (Shashank Singh, 2022). Companies can thereby use this information in order to improve or change the strategies taken up, in order to improve their retention.

An issue identified when analyzing data is the factor of trust individuals have on the companies, for not misusing their personal information and data. With just over 60% of customers having a satisfying level of trust, insurers must assure their clients of maintaining confidentiality without being victims of data breaches. It can be recommended for the insurers to therefore be transparent on how and what personal data would be used for, and letting customers have control over what they decide to share, and opt out any details they are unwilling to provide (Arnie Gordon, 2021). Companies should also be clear on their **privacy policies**, as well as be open to external audits to verify **compliance**, which would give customers a better impression (Harold Li, 2021). User agreements, and badges of trust provided from security companies can also aid in this aspect. By giving customers the right to request for their data to be updated or deleted, would thereby provide a certain assurance to clients, which could lead to higher satisfaction and an increase in lifetime value and retention.

One of the main variables with the highest impact on customer retention was revealed to be External influences, and this insurers in Sri Lanka, especially considering the current economic situation in the country should pay close attention to the effects of inflation rates and competition, on its customer base. Predictive analytics here would be of use in order to look at **price optimization**, where companies can also get effective results on their ROI. This can be done by diving deep into customer segments in order to identify product ad purchasing patterns, while also taking into consideration the external influences. Once all this information is taken into account, agents can also get an idea on how consumers would behave in different purchasing situations, as well as environments by analyzing data to identify trends and patterns to optimize pricing that would be personalized and suitable for customers in different situations (KeenFolks,

2022). Now that significant variables have been identified, insurers can utilize and understand current and future trends in order to create strategies to gain competitive advantages, increase market share and opportunities. Market gaps can thereby be filled as companies make use of the technology revealing customer perceptions, in order to improve customer retention in the company by understanding target audiences better.

5.4. Research Limitations

With the ongoing crisis in Sri Lanka, the researcher did find it difficult to gather data from the target population, mainly due to the fact that people in general were not in a mindset to provide required data with the current economic downfall in the country. A population of 384 respondents was defined to be collected, however the author was only able to collect responses from 317 individuals, of which 6 responses were considered invalid. The research was thereby conducted with the responses of 311 participants, which still accounts to 81% of the initial target sample.

With the analysis conducted, a R^2 value of 47.4% was identified, meaning there can be variables amounting to 52.6% that are not included in this study. As a result of this, the built model would only aid insurers with the variables that are considered, while there can be various other factors that affect customer retention in insurance.

On the other hand, one of the main limitations identified by the author is the scarcity of resources and literature in order to conduct this research. (Stefan Mau, 2017), along with other researchers such as (Tala Mirzaei, 2014) and (Noorhannah Boodhun; Manoj Jayabalan, 2018) agree that there are a very few academic articles and comprehensive schemes for the adoption of predictive analytics techniques on how this could enrich the insurance domain and business efficacy. With this reason, the author found it challenging to identify variables and dimensions that would give a proper outcome, and is also a reason to dive into the pragmatism approach.

The author did find it challenging to analyze the current customer base from a leading insurance company in the country. This was mainly due to the time constraints faced while conducting the research, and also the variables that were taken into account for

the study as well. Therefore, the model was not directed towards the company secondary data, but with the primary data gathered from the questionnaire.

5.5.Future Research

Taking all aspects covered in the study into account, a number of improvements can be made by future researchers to come up with a more significant model. A larger sample population can be analyzed in order to get the perceptions of a larger audience. As recommended, companies can take better advantage of predictive models by understanding the needs of the different segments. Therefore, future research can also be done where data collected from participants are more personalized, to give a better view on the different customer groups. This way, predictions can be more accurate and also catered to the different types of customers an insurer could interact with.

The present study revealed an R^2 value of 47.4%, meaning there are numerous variables that amount to the residual and thereby future research can look into other variables and factors that have a wider and significant impact on customer retention. In this way, once a proper model is built, companies would be able to rely on the output on grounds of more factors being looked into. Moreover, analyzed data can also be used to incorporate current customer bases of companies to the built predictive model, in order to test accuracy and observe how current trends can be improved with the aid of predictive analytics.

6. REFERENCES

- Abby Borden, 2021. *5 ways predictive analytics improves customer retention.* [Online] Available at: <https://lexer.io/5-ways-predictive-analytics-improves-customer-retention/#:~:text=Predictive%20analytics%20enables%20you%20to,experiences%20at%20every%20lifecycle%20stage> [Accessed 4 6 2021].
- Andrew W. Moore, 2006. Combining Multiple Signals for Biosurveillance. *Handbook of Biosurveillance*.
- Arnie Gordon, 2021. *12 Ways To Assure Your Customers That Their Data Is Protected.* [Online] Available at: <https://www.forbes.com/sites/forbestechcouncil/2021/06/08/12-ways-to-assure-your-customers-that-their-data-is-protected/?sh=249e1e9a5925> [Accessed 5 6 2022].
- Bianca Williams, Caroline Halloin, 2020. Data-Driven Model Development for Cardiomyocyte Production Experimental Failure Prediction. *Computer Aided Chemical Engineering*, Issue 30.
- Bitran, G., 1997. A comparative analysis of decision making procedures in the catalog sales industry. *European Management Journal*, 15(2).
- Carmichael, A. & Anderson, D. J., 2015. *Essential Kanban Condensed.* illustrated ed. s.l.:Lean-Kanban University, 2015.
- D. Kalaivani, 2018. Factor based prediction model for customer behavior analysis. *Int J Syst Assur Eng Manag*.
- Diva Dugar, 2018. *Skew and Kurtosis: 2 Important Statistics terms you need to know in Data Science.* [Online] Available at: <https://codeburst.io/2-important-statistics-terms-you-need-to-know-in-data-science-skewness-and-kurtosis-388fef94eeaa> [Accessed 1 6 2022].

Do, D., 2017. *The Five Principles of Lean*. [Online] Available at: <https://theleanway.net/The-Five-Principles-of-Lean> [Accessed 05 August 2017].

Doll, J., 2020. *Effective Agile sprint cycles for data consulting projects*. [Online] Available at: <https://nealanalytics.com/blog/effective-agile-sprint-cycles-for-data-consulting-projects/> [Accessed 1 Effective Agile sprint cycles for data consulting projects 4 07 2020].

Elsayed Elashkar, 2020. Business predictive analysis from business insurance data using business strategic planning techniques. *KNOWLEDGE MANAGEMENT RESEARCH & PRACTICE*.

Fatemeh Safara, 2020. A Computational Model to Predict Consumer Behaviour During COVID-19 Pandemic. *Computational Economics*.

Gunther, C.-C., 2014. Modelling and predicting customer churn from an insurance company. *Scandinavian Actuarial Journal*, 2014(1), pp. 58-71.

Harold Li, 2021. *12 Ways To Assure Your Customers That Their Data Is Protected*. [Online]

Available at: <https://www.forbes.com/sites/forbestechcouncil/2021/06/08/12-ways-to-assure-your-customers-that-their-data-is-protected/?sh=249e1e9a5925> [Accessed 5 6 2022].

Hsiao-Fan Wang, 2005. Managing customer profitability in a competitive market by continuous data mining. *Industrial Marketing Management*.

Insurance Regulatory Commission of Sri Lanka , 2019. *Performance of the Insurance Industry for the year 2019*, Colombo: Insurance Regulatory Commission of Sri Lanka .

Insurance Regulatory Commission of Sri Lanka, 2019. *Performance of the Insurance Industry for the year 2019*, Colombo: Insurance Regulatory Commission of Sri Lanka.

Jaaron, A. . A. & Backhouse, C., 2011. A methodology for the implementation of lean thinking In manufacturing support services. *International Journal of Services and Operations Management (Int J Serv Oper Manag)*, pp. 4-6.

Jamal, S., 2019. Predictive Analysis for Claims in Insurance Industry using Machine Learning. *International Research Journal of Engineering and Technology (IRJET)*, 6(8).

John Dudovskiy, 2022. *Pragmatism Research Philosophy*. [Online] Available at: <https://research-methodology.net/research-philosophy/pragmatism-research-philosophy/> [Accessed 15 5 2022].

KeenFolks, 2022. *How Predictive Analytics Can Improve Company Performance And Drive Revenue*. [Online] Available at: <https://thekeenfolks.com/predictive-analytics/> [Accessed 5 6 2022].

KPMG, 2021. *Sri Lanka Insurance Report*, s.l.: KPMG Sri Lanka.

Kumar.V, 2006. Customer Relationship Management: A Databased Approach, Wiley.

MANTEC, 2018. *WHAT IS SIX SIGMA & ITS BENEFITS?*. [Online] Available at: <https://mantec.org/six-sigma-benefit-manufacturers/> [Accessed 28 June 2018].

Maresova Petra, 2012. Research of the Behavior of Consumers in the Insurance Market in the Czech Republic. *Journal of Competitiveness*, 4(2), pp. 20-37.

Mark Xu, 2005. Gaining customer knowledge through analytical CRM. *Industrial Management & Data Systems*, 105(7), pp. 995-971.

Michaela Weber, Jeanine Van Ancum, 2018. Concurrent validity and reliability of the Community Balance and Mobility scale in young-older adults. *BMC Geriatrics*, p. 18:156.

Milliman, 2016. *The use of AI and Data Analytics in Life Insurance*, s.l.: Milliman.

Negi, R., 2010. Determining satisfaction through customer-perceived service quality and value: evidence from Ethiopian Telecommunications. *International Journal of Services and Operations Management*, pp. 333-350.

Noorhannah Boodhun; Manoj Jayabalan, 2018. Risk prediction in life insurance industry using supervised learning algorithms. *Complex & Intelligent Systems*.

Peter C. Verhoef, R. V. L. M. E. C. M. M. K. & S. G., 2010. CRM in Data-Rich Multichannel Retailing Environments: A Review and Future Research Directions. *Journal of Interactive Marketing*, 24(2), pp. 121-137.

QuestionPro, 2022. *Data Collection Methods: Definition, Examples and Sources*. [Online]

Available at: <https://www.questionpro.com/blog/data-collection-methods/> [Accessed 24 5 2022].

Rebecca Bevans, 2022. *Simple Linear Regression / An Easy Introduction & Examples*. [Online]

Available at: <https://www.scribbr.com/statistics/simple-linear-regression/#:~:text=What%20is%20simple%20linear%20regression,Both%20variables%20should%20be%20quantitative.> [Accessed 2 6 2022].

Shashank Singh, 2022. *Top Strategies For Increasing Insurance Customer Engagement*. [Online]

Available at: <https://www.reutersevents.com/insurance/blog/top-strategies-increasing-insurance-customer-engagement> [Accessed 7 2022].

Sherman, R., 2015. *Business Intelligence Guidebook : From Data Integration to Analytics*. illustrated ed. s.l.:Elsevier Science, 2014.

Singh, V., 2021. *HOME : AGILE : KANBAN GUIDE 2019*. [Online] Available at: <https://www.toolsqa.com/agile/kanban/> [Accessed 07 07 2021].

Statistic Solutions , 2022. *Reliability Analysis*. [Online] Available at: <https://www.statisticssolutions.com/reliability-analysis/> [Accessed 29 5 2022].

Statistics Solutions, 2022. *Cronbach's Alpha*. [Online] Available at: <https://www.statisticssolutions.com/cronbachs-alpha/> [Accessed 29 5 2022].

Stefan Mau, 2017. Forecasting the next likely purchase events of insurance customers. *International Journal of Bank Marketing*.

Swiss Re, 2014. Digital distribution in insurance: a quiet revolution.

Tala Mirzaei, 2014. Application of predictive analytics in CRM. *SAIS 2014 Proceedings*, Volume 23.

Tennant, G., 2001. *Six Sigma: SPC and TQM in Manufacturing and Services*. Hampshire: Gower Publishing LTD.

Verhoef, P. & Venkatesan, R. M. L. M. E. K. M. a. G. S., 2010. CRM in data-rich multichannel retailing environments: a review and future research directions. *Journal of Interactive Marketing*, 24(2), pp. 121-137.

Vincent Nijs, 2019. *Multivariate > Factor > Pre-factor*. [Online] Available at: https://rstantstats.github.io/docs/multivariate/pre_factor.html#:~:text=The%20KMO%20and%20Bartlett%20test,is%20correlated%20with%20other%20variables [Accessed 30 5 2022].

Wagner A. Kamakura, 1991. Applying latent trait analysis in the evaluation of prospects for cross-selling of financial services. *International Journal of Research in Marketing* 8, pp. 329-349.

Werner J. Reinartz, V. Kumar, 2003. The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration. *Journal of Marketing*, Volume 67, pp. 77-99.

WordPress, 2011. *What is validity and why is it important in research?*. [Online] Available at: <https://psuqd8.wordpress.com/2011/11/20/why-is-validity-important-in-research/#:~:text=So%20why%20is%20validity%20important,main%20aim%20of%20the%20study> [Accessed 29 5 2022].

Writers, 1., 2022. *Understanding the Research Onion.* [Online] Available at: <https://15writers.com/research-onion/> [Accessed 30 3 2022].

Yogesh Dhavale, 2020. *3 Ways to Target the Right Customers in the Insurance Industry.* [Online] Available at: <https://blog.aureusanalytics.com/blog/different-ways-to-target-the-right-customers-in-the-insurance-industry> [Accessed 4 6 2022].

7. APPENDIX AND ANNEXURES

7.1. Appendix 01 – Supervisor Approval and Log Sheets

Approval Request - Supervisor Sessions Inbox x ▼ □

 **Senara Kalpage** Sat, Jun 4, 8:04 PM (4 days ago) ★
Dear Mr. Induneth, Please find my supervisor session log sheets attached herewith for your approval. Thank you, Senara Kalpage, CL/CBISM/...

 **Induneth De Silva** Sun, Jun 5, 7:55 AM (3 days ago) ★ ↗ ⋮
to me, Induneth ▼
Approved!
...

On Sat, Jun 4, 2022 at 8:04 PM Senara Kalpage <senara286@gmail.com> wrote:
Dear Mr. Induneth,

Please find my supervisor session log sheets attached herewith for your approval.

Thank you,
Senara Kalpage,
CL/CBISM/12/07 - BIS_Batch_12.

Figure 76 Supervisor Approval



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project – CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should remain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 13/03/2022

Meeting No: 1

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):

Initial Research Title discussion

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Explain research problem
2. Feedback on research title

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Research on BIS related research problems
2. Better to have a different topic or specify scope of current topic

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 77 Project Log Sheet - 01



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project – CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at least ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should remain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 20/03/2022

Meeting No: 2

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study
Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):
Studying different research areas

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Explain research problem
2. Feedback on research title
3. Discussion on way forward

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Received confirmation, and therefore to identify variables.
2. Prepare research objectives and methodology.
3. Work on proposal presentation.

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 78 Project Log Sheet - 02



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project – CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should retain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 01/05/2022

Meeting No: 3

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):

Identified variables, and working on literature review

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Need an explanation of the Interim Report 1 structure
2. Clarifications on how variables should be broken down into dimensions and elements
3. Clarifications on how the hypothesis needs to be created

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Completion of Interim Report 1
2. Bring a proper breakdown of the variables, dimensions and elements
3. Have an idea on how to structure the questionnaire

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 79 Project Log Sheet - 03



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project - CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should retain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 08/05/2022

Meeting No: 4

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):
Submitted Interim Report 1 and working on Interim Report 2

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Walkthrough SPSS
2. Clarification on the inclusion of dimensions in questionnaires
3. Understand the different types of analysis that will be needed

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Prepare a draft of the questionnaire
2. Understand the process of pilot testing and align questions accordingly.

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 80 Project Log Sheet - 04



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project - CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should retain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 15/05/2022

Meeting No: 5

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):
Started first rounds of pilot testing

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Clarifications on how validity and reliability tests need to be done
2. Clarifications on how the results of respondents affect the analysis results

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Remove or add questions to increase Cronbach Alpha
2. Complete pilot testing and start analysis

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 81 Project Log Sheet - 05



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project – CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should retain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 22/05/2022

Meeting No: 6

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):

Completed operationalization and pilot testing

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Clarify issues regarding to low Cronbach
2. Understand the process of analysis in chapter 4
3. Discuss how the statistical model needs to be built

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Make improvements to questions if necessary
2. Try to find questions through literature
3. Get financial information for statistical model

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 82 Project Log Sheet - 06



ICBT
International College of
Business & Technology

Project Log Sheet – Supervisory Sessions for Information Systems Project – CIS6000

Notes on use of the project log sheet:

1. This log sheet is designed for meetings of more than 15 minutes duration, of which there must be at ten (10) during the course of the project (TEN mandatory supervisory sessions).
2. The student should prepare for the supervisory sessions by deciding which question(s) he or she needs to ask the supervisor and what progress has been made (if any), since the last session.
3. A log sheet is to be brought by the student to each supervisory session.
4. The actions by the student (and, perhaps the supervisor), which should be carried out before the next supervisory meeting should be noted briefly in the relevant section in the form.
5. The student should leave a copy (after the session) of the Project Log sheet with the supervisor and to the coordinator (for Assistant Manager to include his signature). One copy should retain with the student.
6. It is compulsory that students bring their previous supervisory session log sheets together with the project file during each supervisory session.
7. The log sheet is a deliverable for the project and it is an important record of a student's organization and learning experience. The student MUST hand in the log sheets as an appendix of the final year documentation, with sheets dated and numbered consecutively.

Student's Name: Senara Kalpage

Cardiff Number: 20216562

Date: 03/06/2022

Meeting No: 7

Intake: 2021/2022

Project Title: The Significance of Predictive Analysis in Forecasting Customer Retention: A Study Conducted in the Insurance Industry of Sri Lanka

Supervisor's Name: Mr. Induneth De Silva

Supervisor Signature:

Coordinator's Name: Ms. Kalpani Wickramasinghe

Program Manager Signature:

Work progression as to date (noted by student BEFORE mandatory supervisor meeting):
Started correlation and regression analysis in chapter 4

Items for Discussion (noted by student BEFORE mandatory supervisor meeting):

1. Issues in the statistical model
2. Way forward if objective variables cannot be taken forth

Action List (to be attempted by student by the NEXT mandatory supervisory meeting – TO BE FILLED SUPERVISOR):

1. Identify and explain correlations
2. Make improvements to questionnaires if needed
3. Try different analysis methods and dive deep into explaining relationships between existing variables.

Note: A student should make an appointment to meet his or her supervisor (via phone call or e-mail) at least 3 days prior to supervisory session. In the event a supervisor could not be booked for consultation, the Assistant Manager should be informed 2 days prior to supervisory meeting so that a meeting can be subsequently arranged.

Figure 83 Project Log Sheet - 07

7.2.Appendix 02 – Final Survey Questionnaire

Below is the questionnaire distributed among participants

Forecasting Customer Retention in the Insurance Domain using Predictive Analytics

Hello,

I'm Senara Kalpage, an undergraduate at Cardiff Metropolitan University - UK, conducting this survey in order to gather data for my final Dissertation project in completing my Bachelor's Degree in Business Information Systems.

This survey is specific for participants who are customers in the insurance domain in Sri Lanka. The information obtained from this will only be used for thesis literature needs, and will definitely remain confidential and anonymous. Any participant's response won't be recognized personally and will be kept in an aggregated form. There are no identified risks in participating in this survey and your participation and honest responses are highly appreciated and valued immensely.

Research Title: The Significance of Predictive Analysis in Forecasting Customer Retention - A Study Conducted in the Insurance Industry of Sri Lanka.

I kindly invite you to participate in this survey conducted to identify factors influencing customer retention in the insurance industry.

If you have any recommendations or clarifications, please feel free to reach out:

Email - senara286@gmail.com

Thank you in advance, and Best Regards,
Sincerely,
Senara Kalpage.

Figure 84 Final Survey Questionnaire

Participant Consent: I have read and understood the above information and agree to voluntarily participate in the survey.

*

- Yes
- No

Please indicate your age

- Below 25
- 25 - 35
- 35 - 50
- Above 50

Please indicate your highest Educational Level

- A Levels
- Diploma / HND
- Bachelors
- Masters
- PhD
- Do not wish to specify

Figure 85 Final Survey Questionnaire (cont'd)

Are you currently employed?

- Yes
- No

Number of household dependencies

- None
- 1 - 2
- 3 or more

How many insurance policies do you have?

- None
- 1 - 2
- 3 or more

[Next](#)



Page 1 of 2

[Clear form](#)

Figure 86 Final Survey Questionnaire (cont'd)

Analyzing the significant components

Product significance

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I prefer being well aware of the type of policy I require	<input type="radio"/>				
I consider the insurance coverage before I purchase a policy	<input type="radio"/>				
I'm willing to reach out to my agent if I have an issue with my policy	<input type="radio"/>				
It's more likely that I'll change my policy if I find the product unsatisfactory	<input type="radio"/>				

Figure 87 Final Survey Questionnaire (cont'd)

What is your opinion on risk and perceived prices?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I tend to think into the future before making a purchase	<input type="radio"/>				
I try to minimize taking risks and take extra precaution to reduce risks	<input type="radio"/>				
I'm conscious about premium amounts when choosing an insurance policy	<input type="radio"/>				
I trust that the company will not misuse my personal data	<input type="radio"/>				

Figure 88 Final Survey Questionnaire (cont'd)

Financial perception					
	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I consider having an insurance policy to be beneficial	<input type="radio"/>				
I would prefer paying premiums if they are affordable	<input type="radio"/>				
I would choose an insurance policy considering my financial status	<input type="radio"/>				
I would face difficulties paying my insurance if I lose my source of income	<input type="radio"/>				

Figure 89 Final Survey Questionnaire (cont'd)

Discounts and other benefits

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I'm more interested in a product if there is a discount provided	<input type="radio"/>				
I prefer being aware of the benefits my insurance policy will bring me	<input type="radio"/>				
I wouldn't necessarily discontinue payments if no discounts are received	<input type="radio"/>				
Affordable policies with benefits increase my satisfaction	<input type="radio"/>				

Figure 90 Final Survey Questionnaire (cont'd)

Do insurance companies do a good job assuring customers?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I've come across insurance adverts on various social platforms	<input type="radio"/>				
Most adverts I've come across give me reassurance that I've purchased the right insurance	<input type="radio"/>				
I have been contacted by my agents occasionally	<input type="radio"/>				
My agents are able to provide all necessary information I need	<input type="radio"/>				
There have been times where my agent assured the best when I was in doubt	<input type="radio"/>				
Being contacted by an agent frequently can cause inconvenience	<input type="radio"/>				

Figure 91 Final Survey Questionnaire (cont'd)

Inflation and competition

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I'm well aware of the current price hikes in Sri Lanka	<input type="radio"/>				
I've re-considered my necessities due to the price hikes	<input type="radio"/>				
Other insurance providers have reached out to me	<input type="radio"/>				
I'm aware of other insurance providers	<input type="radio"/>				

Figure 92 Final Survey Questionnaire (cont'd)

Overall feedback

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
I consider insurance to be of importance	<input type="radio"/>				
I'm confident that affordable premiums and proper discounts are attractive offers	<input type="radio"/>				
I would recommend having an insurance to a friend/family.	<input type="radio"/>				

Back Submit Page 2 of 2 Clear form

Figure 93 Final Survey Questionnaire (cont'd)

7.3.Appendix 03 - Individual Reliability Analysis

1. Assurance

Reliability Statistics	
Cronbach's Alpha	N of Items
.867	6

Figure 94 Assurance - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Assurance_1	19.2476	12.716	.557	.862
Assurance_2	19.5080	11.741	.684	.841
Assurance_3	19.5273	10.869	.723	.834
Assurance_4	19.4566	11.178	.754	.828
Assurance_5	19.5016	11.431	.733	.833
Assurance_6	19.3666	11.852	.551	.866

Figure 95 Assurance - Item Statistics

2. Lifetime Value

Reliability Statistics

Cronbach's Alpha	N of Items
.844	7

Figure 96 Lifetime Value - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Lifetime_Value_1	25.7910	10.140	.576	.826
Lifetime_Value_2	25.7235	10.227	.615	.820
Lifetime_Value_3	25.7621	9.788	.636	.817
Lifetime_Value_4	25.8006	9.921	.632	.817
Lifetime_Value_5	25.8842	10.051	.637	.817
Lifetime_Value_6	25.9550	10.372	.542	.831
Lifetime_Value_7	25.9518	10.182	.558	.829

Figure 97 Lifetime Value - Item Statistics

3. Satisfaction

Reliability Statistics

Cronbach's Alpha	N of Items
.721	3

*Figure 98 Satisfaction - Cronbach Alpha***Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Satisfaction_1	8.7331	1.590	.413	.797
Satisfaction_2	8.5241	1.508	.594	.574
Satisfaction_4	8.5241	1.431	.640	.514

Figure 99 Satisfaction - Item Statistics

4. Financial Status

Reliability Statistics

Cronbach's Alpha	N of Items
.756	4

Figure 100 Finance - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Finance_1	12.6656	3.836	.589	.682
Finance_2	12.7074	3.737	.617	.666
Finance_3	12.5466	3.771	.631	.661
Finance_4	12.8006	3.696	.421	.790

Figure 101 Finance - Item Statistics

5. External Influences

Reliability Statistics

Cronbach's Alpha	N of Items
.703	3

*Figure 102 External Influences - Cronbach Alpha***Item-Total Statistics**

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
External_1	8.4373	1.821	.591	.526
External_2	8.5595	1.834	.502	.635
External_4	8.5209	1.921	.471	.672

Figure 103 External Influences - Item Statistics

6. Retention

Reliability Statistics

Cronbach's Alpha	N of Items
.834	3

Figure 104 Retention - Cronbach Alpha

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Retention_1	8.4084	1.881	.729	.736
Retention_2	8.4116	2.088	.674	.792
Retention_3	8.4662	1.869	.686	.781

Figure 105 Retention - Item Statistics

7.4. Appendix 04 - Individual Validity Analysis

1. Assurance

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.642	60.706	60.706	3.642	60.706	60.706
2	.725	12.077	72.783			
3	.621	10.343	83.126			
4	.409	6.819	89.945			
5	.369	6.145	96.090			
6	.235	3.910	100.000			

Extraction Method: Principal Component Analysis.

*Figure 106 Validity Loading % - Assurance***Component Matrix^a**

	Component
	1
Assurance_1	.675
Assurance_2	.795
Assurance_3	.826
Assurance_4	.850
Assurance_5	.836
Assurance_6	.672

Extraction Method: Principal Component Analysis.

Figure 107 Component Matrix - Assurance

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.863
Bartlett's Test of Sphericity	Approx. Chi-Square df Sig.	874.788 15 .000

Figure 108 KMO & Bartlett's Test – Assurance

2. Lifetime Value

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.625	51.790	51.790	3.625	51.790	51.790
2	.967	13.818	65.609			
3	.592	8.458	74.067			
4	.536	7.659	81.725			
5	.457	6.522	88.248			
6	.418	5.971	94.219			
7	.405	5.781	100.000			

Extraction Method: Principal Component Analysis.

Figure 109 Validity Loading % - Lifetime Value

Component Matrix^a

	Component
	1
Lifetime_Value_1	.702
Lifetime_Value_2	.733
Lifetime_Value_3	.754
Lifetime_Value_4	.748
Lifetime_Value_5	.749
Lifetime_Value_6	.666
Lifetime_Value_7	.681

Extraction Method: Principal Component Analysis.

Figure 110 Component Matrix - Lifetime Value

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.866
Bartlett's Test of Sphericity	Approx. Chi-Square df Sig.	752.932 21 .000

Figure 111 KMO & Bartlett's Test - Lifetime Value

3. Satisfaction

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.958	65.280	65.280	1.958	65.280	65.280
2	.708	23.591	88.871			
3	.334	11.129	100.000			

Extraction Method: Principal Component Analysis.

Figure 112 Validity Loading % - Satisfaction

Component Matrix^a

	Component
	1
Satisfaction_1	.680
Satisfaction_2	.853
Satisfaction_4	.877

Extraction Method: Principal Component Analysis.

Figure 113 Component Matrix - Satisfaction

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.619
Bartlett's Test of Sphericity	237.455
df	3
Sig.	.000

Figure 114 KMO & Bartlett's Test - Satisfaction

4. Financial Status

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.384	59.602	59.602	2.384	59.602	59.602
2	.733	18.325	77.927			
3	.468	11.707	89.634			
4	.415	10.366	100.000			

Extraction Method: Principal Component Analysis.

Figure 115 Validity Loading % - Finance

Component Matrix^a

	Component
	1
Finance_1	.799
Finance_2	.822
Finance_3	.825
Finance_4	.625

Extraction Method:
Principal Component
Analysis.

*Figure 116 Component Matrix - Finance***KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.767
Bartlett's Test of Sphericity	332.718
Approx. Chi-Square	
df	6
Sig.	.000

Figure 117 KMO & Bartlett's Test - Finance

5. External Influences

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.889	62.977	62.977	1.889	62.977	62.977
2	.646	21.528	84.505			
3	.465	15.495	100.000			

Extraction Method: Principal Component Analysis.

*Figure 118 Validity Loading % - External Influences***Component Matrix^a**

	Component
	1
External_1	.841
External_2	.783
External_4	.754

Extraction Method:
Principal Component
Analysis.

Figure 119 Component Matrix - External Influences

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.653
Bartlett's Test of Sphericity	174.734
df	3
Sig.	.000

Figure 120 KMO & Bartlett's Test - External Influences

6. Retention

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.255	75.179	75.179	2.255	75.179	75.179
2	.416	13.852	89.032			
3	.329	10.968	100.000			

Extraction Method: Principal Component Analysis.

Figure 121 Validity Loading % - Retention

Component Matrix^a

	Component
	1
Retention_1	.886
Retention_2	.854
Retention_3	.861

Extraction Method: Principal Component Analysis.

Figure 122 Component Matrix - Retention

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	.720
Bartlett's Test of Sphericity	362.502
df	3
Sig.	.000

Figure 123 KMO & Bartlett's Test - Retention

7.5.Appendix 05 - Individual Correlation Analysis Results

1. Assurance

Correlations		
	Assurance	Customer_Retention
Assurance	Pearson Correlation	1
	Sig. (2-tailed)	.404**
	N	311
Customer_Retention	Pearson Correlation	.404**
	Sig. (2-tailed)	.000
	N	311

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 124 Correlation Analysis - Assurance

2. Lifetime Value

Correlations		
	Lifetime_Value	Customer_Retention
Lifetime_Value	Pearson Correlation	1
	Sig. (2-tailed)	.572**
	N	311
Customer_Retention	Pearson Correlation	.572**
	Sig. (2-tailed)	.000
	N	311

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 125 Correlation Analysis - Lifetime Value

3. Satisfaction

Correlations		
	Satisfaction	Customer_Retention
Satisfaction	Pearson Correlation	1
	Sig. (2-tailed)	.570**
	N	311
Customer_Retention	Pearson Correlation	.570**
	Sig. (2-tailed)	.000
	N	311

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 126 Correlation Analysis - Satisfaction

4. Financial Status

		Correlations	
		Financial_Status	Customer_Retention
Financial_Status	Pearson Correlation	1	.565**
	Sig. (2-tailed)		.000
	N	311	311
Customer_Retention	Pearson Correlation	.565**	1
	Sig. (2-tailed)	.000	
	N	311	311

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 127 Correlation Analysis - Financial Status

5. External Influences

		Correlations	
		External_Influences	Customer_Retention
External_Influences	Pearson Correlation	1	.588**
	Sig. (2-tailed)		.000
	N	311	311
Customer_Retention	Pearson Correlation	.588**	1
	Sig. (2-tailed)	.000	
	N	311	311

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 128 Correlation Analysis - External Influences

7.6.Appendix 06 - Individual Regression Analysis Results

1. Assurance

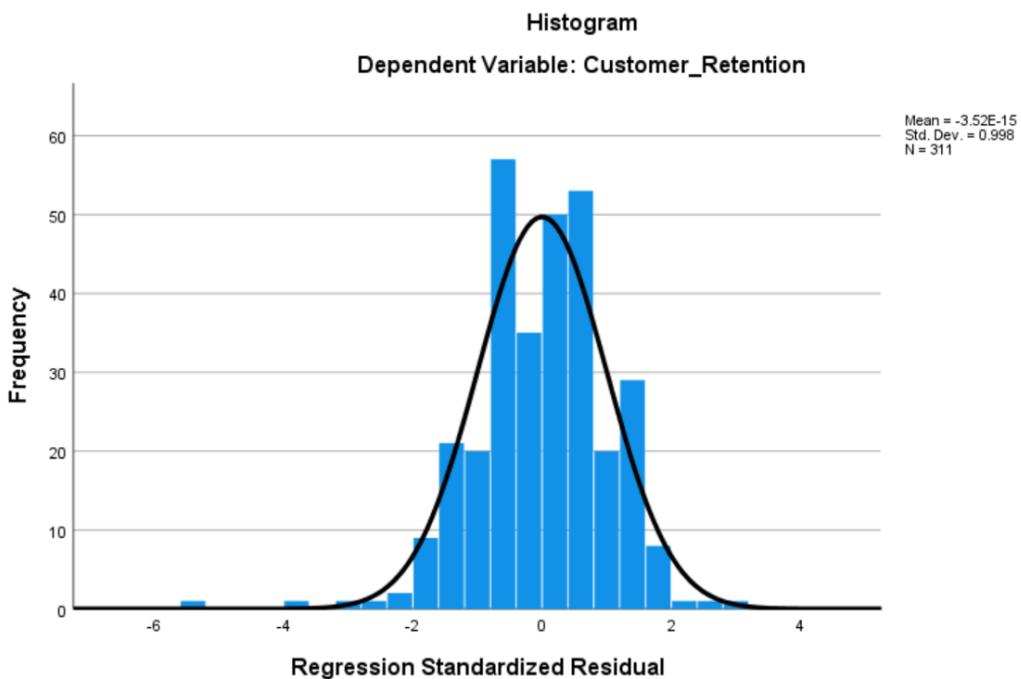


Figure 129 RSR Histogram - Assurance

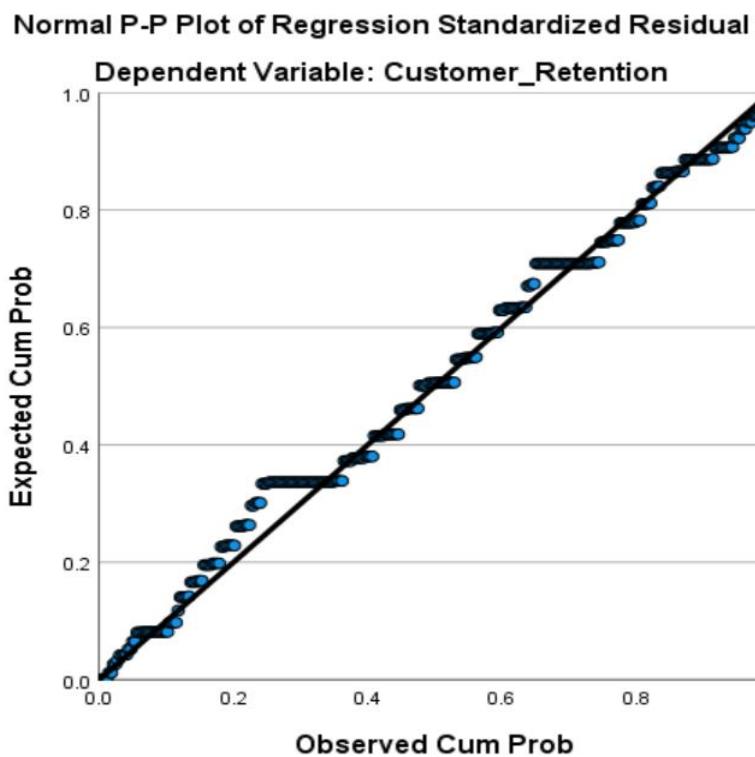


Figure 130 RSR Normal P-Plot – Assurance

2. Lifetime Value

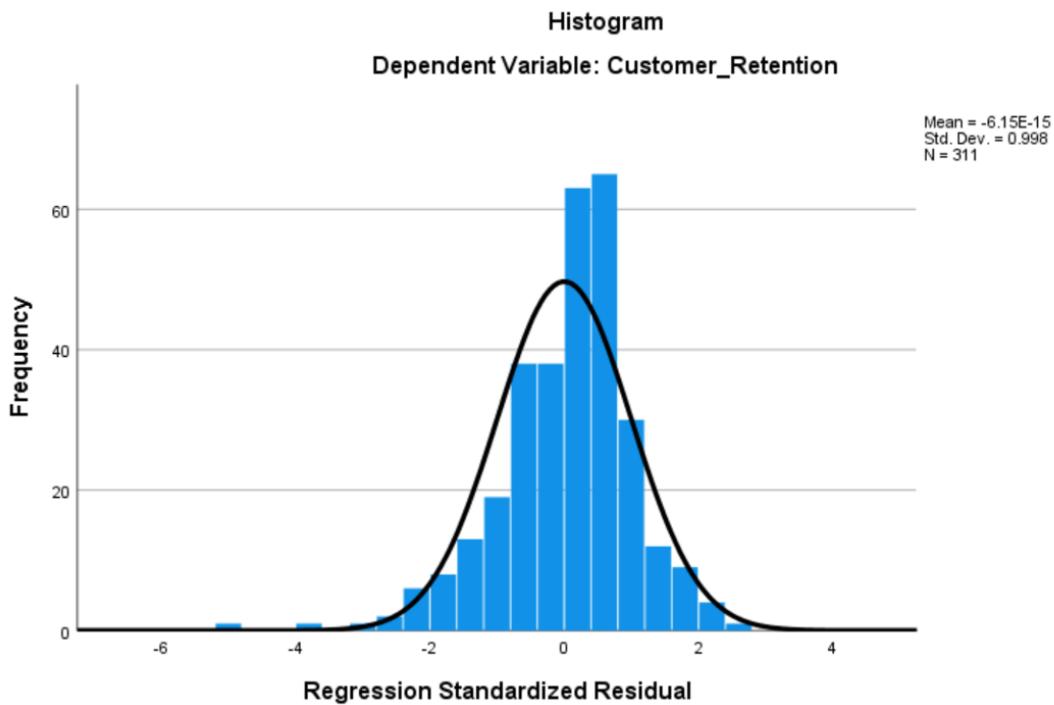


Figure 131 RSR Normal P-Plot - Lifetime Value

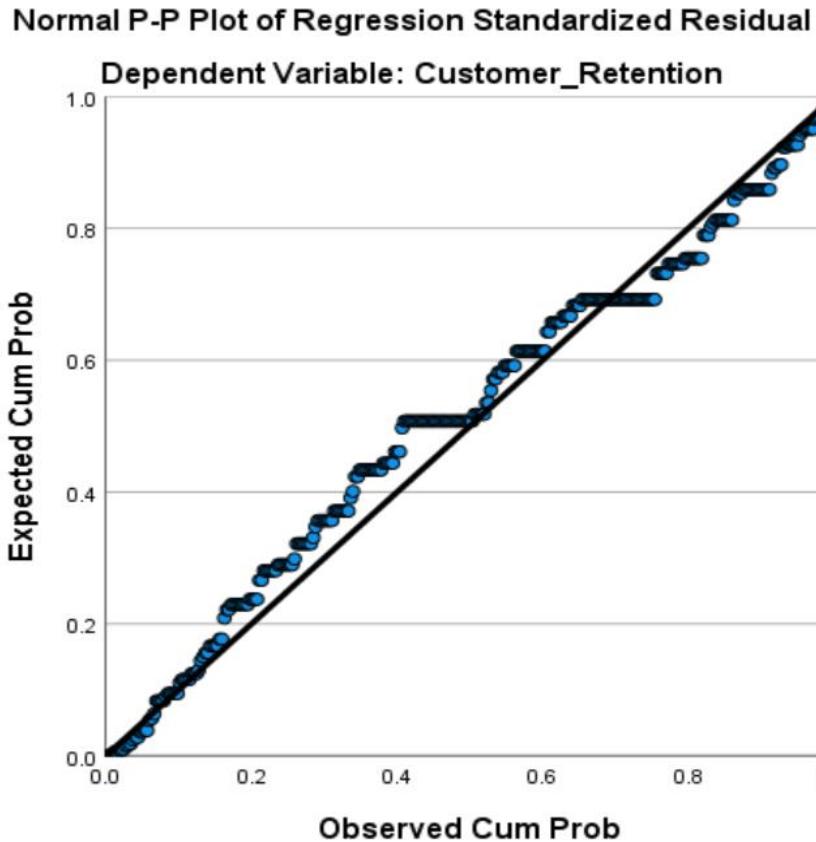


Figure 132 RSR Normal P-Plot - Lifetime Value

3. Satisfaction

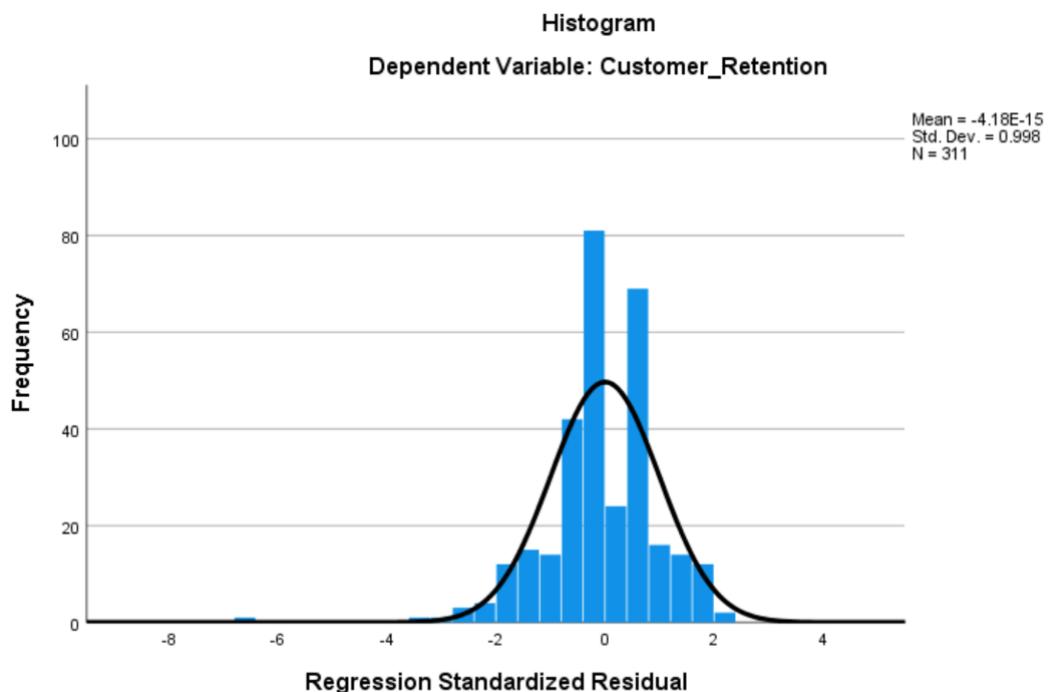


Figure 133 RSR Histogram - Satisfaction

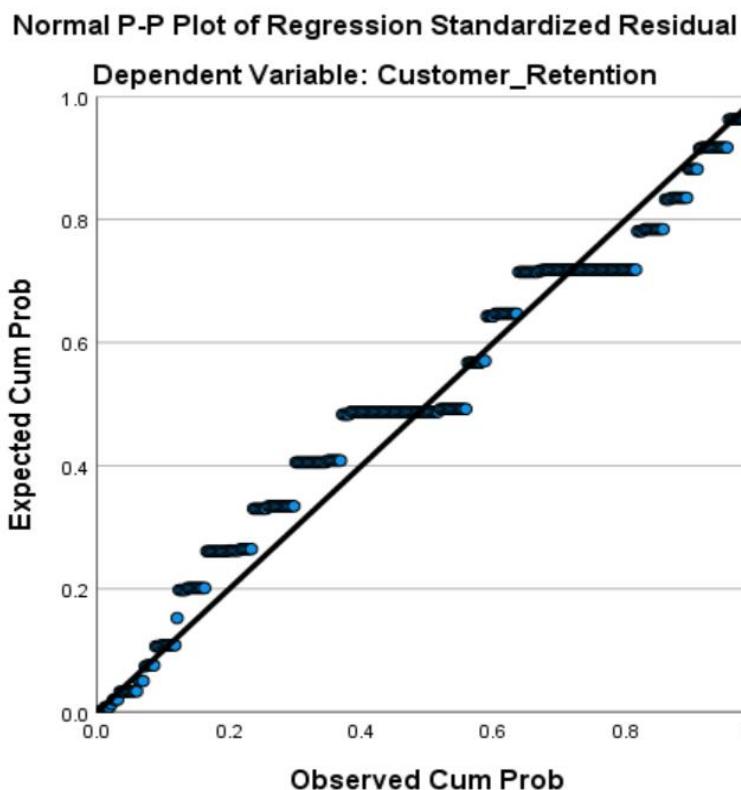


Figure 134 RSR Normal P-Plot - Satisfaction

4. Financial Status

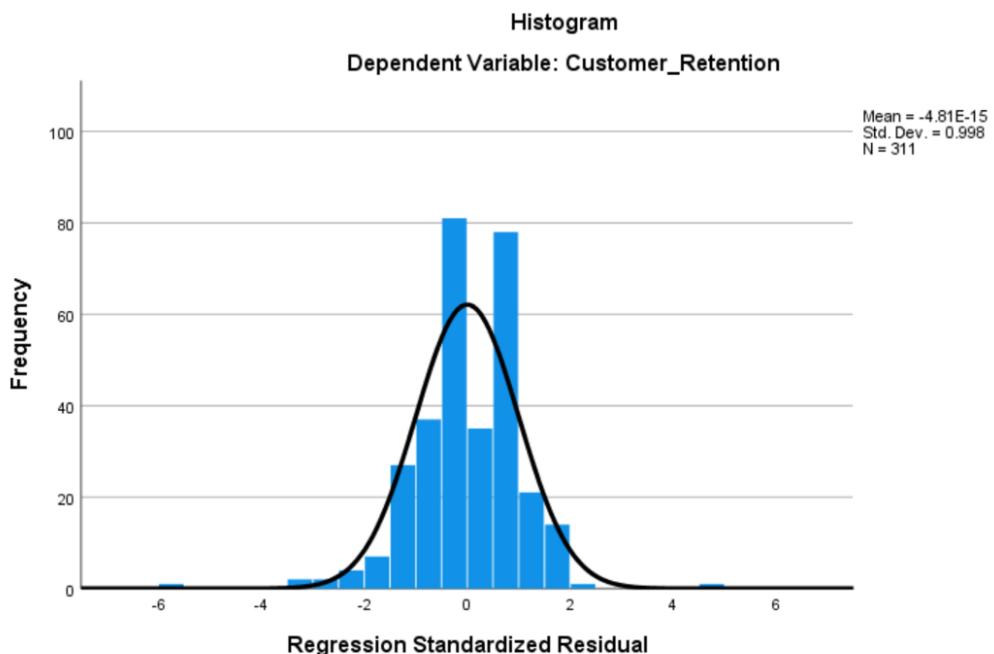


Figure 135 RSR Histogram - Financial Status

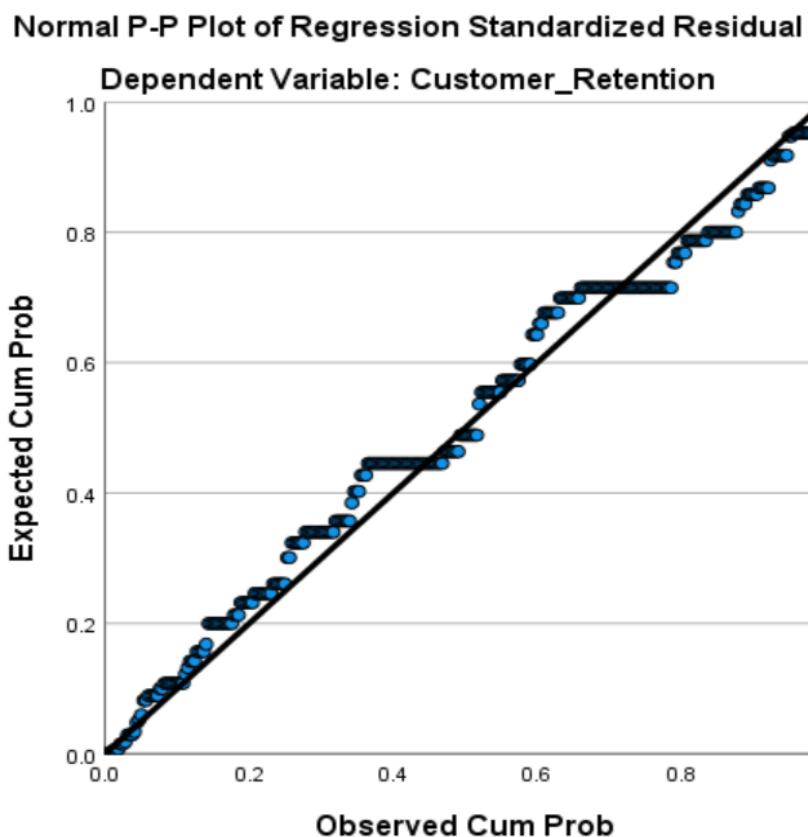


Figure 136 RSR Normal P-Plot - Financial Status

5. External Influences

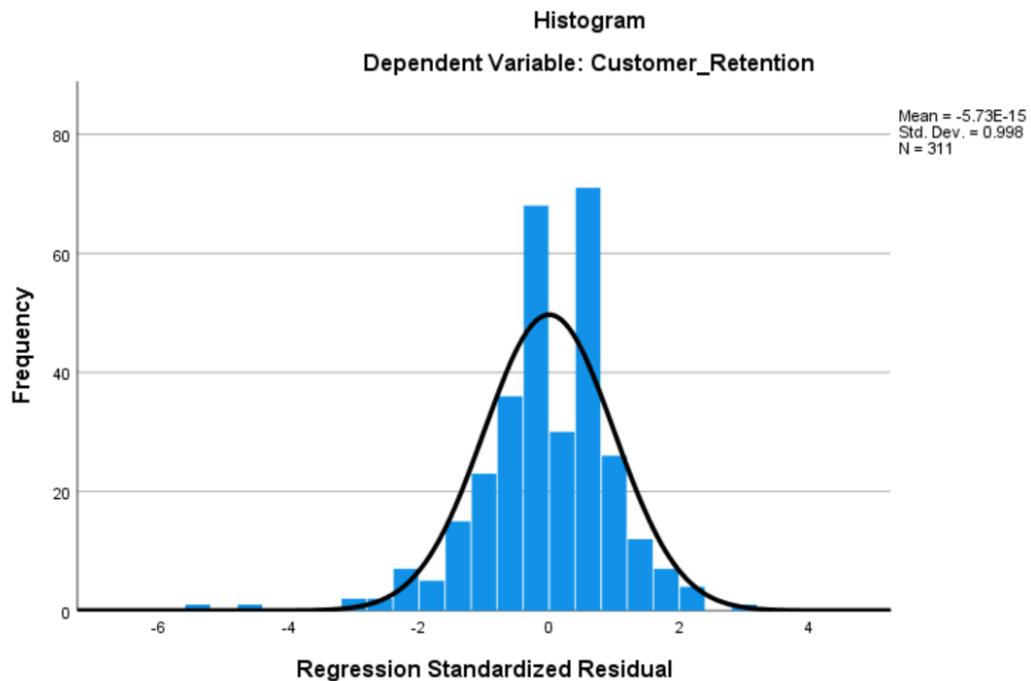


Figure 137 RSR Histogram - External Influences

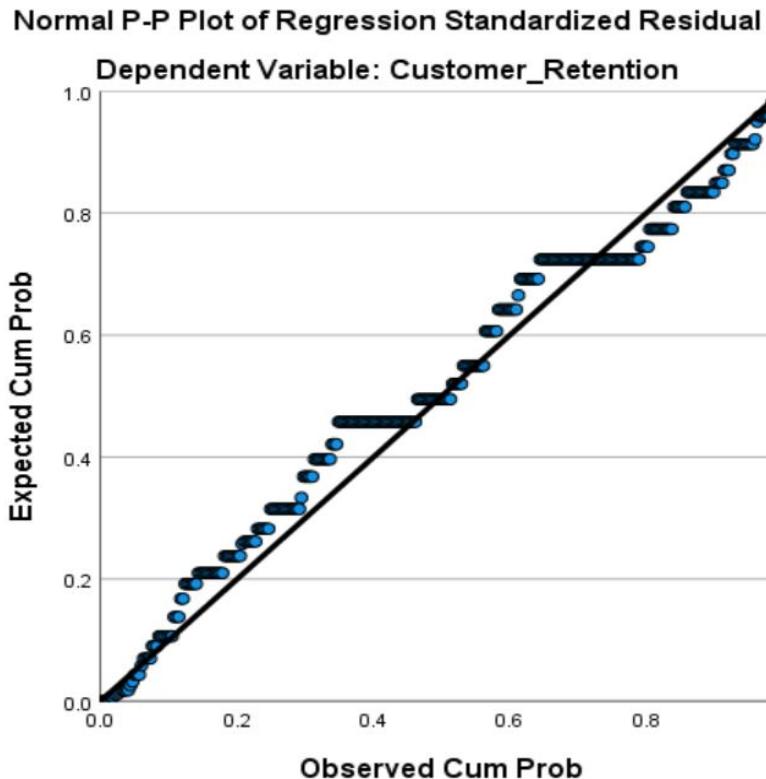


Figure 138 RSR Normal P-Plot - External Influences

7.7.Appendix 07 – Individual MR Hypothesis Testing

7.7.1.Hypothesis 2MR

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.686 ^a	.470	.463	.49101	.470	67.914	4	306	.000

a. Predictors: (Constant), External_Influences, Financial_Status, Lifetime_Value, Satisfaction

b. Dependent Variable: Customer_Retention

Figure 139 Model Summary - 2MR

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	65.492	4	16.373	67.914	.000 ^b
Residual	73.773	306	.241		
Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), External_Influences, Financial_Status, Lifetime_Value, Satisfaction

Figure 140 ANOVA Test - 2MR

Coefficients^a

Model	Unstandardized Coefficients			t	Sig.
	B	Std. Error	Beta		
1 (Constant)	.288	.248		1.165	.245
Lifetime_Value	.220	.079	.172	2.789	.006
Satisfaction	.158	.075	.135	2.114	.035
Financial_Status	.229	.065	.213	3.522	.000
External_Influences	.313	.058	.297	5.364	.000

a. Dependent Variable: Customer_Retention

Figure 141 Regression Coefficients - 2MR

7.7.2.Hypothesis 3MR

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.683 ^a	.467	.460	.49251	.467	67.030	4	306	.000

a. Predictors: (Constant), External_Influences, Assurance, Financial_Status, Lifetime_Value

b. Dependent Variable: Customer_Retention

Figure 142 Model Summary - 3MR

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	65.038	4	16.260	67.030	.000 ^b
	Residual	74.227	306	.243		
	Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), External_Influences, Assurance, Financial_Status, Lifetime_Value

Figure 143 ANOVA Test - 3MR

Coefficients^a

Model	Unstandardized Coefficients		Beta	t	Sig.
	B	Std. Error			
1	(Constant)	.298	.250		.1.194
	Assurance	.079	.049	.079	1.603
	Lifetime_Value	.261	.076	.204	3.419
	Financial_Status	.249	.064	.231	3.905
	External_Influences	.337	.056	.320	5.986

a. Dependent Variable: Customer_Retention

Figure 144 Regression Coefficients - 3MR

7.7.3.Hypothesis 4MR

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.680 ^a	.463	.457	.49377	.463	88.066	3	307	.000

a. Predictors: (Constant), External_Influences, Financial_Status, Lifetime_Value

b. Dependent Variable: Customer_Retention

Figure 145 Model Summary - 4MR

ANOVA^a

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	64.414	3	21.471	88.066	.000 ^b
Residual	74.850	307	.244		
Total	139.265	310			

a. Dependent Variable: Customer_Retention

b. Predictors: (Constant), External_Influences, Financial_Status, Lifetime_Value

Figure 146 ANOVA Test - 4MR

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	.383	.245		1.564	.119
Lifetime_Value	.266	.077	.207	3.472	.001
Financial_Status	.283	.060	.262	4.678	.000
External_Influences	.351	.056	.333	6.300	.000

a. Dependent Variable: Customer_Retention

Figure 147 Regression Coefficients - 4MR

7.8.Appendix 08 – Survey Results

Product significance

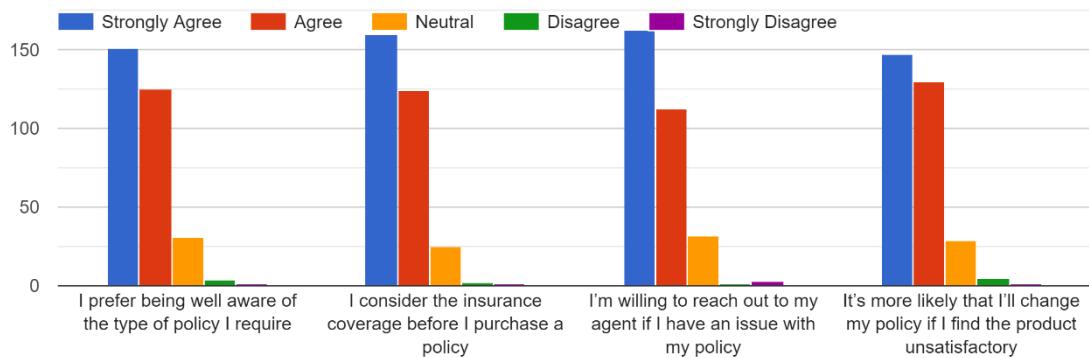


Figure 148 Survey Results - Product Significance

What is your opinion on risk and perceived prices?

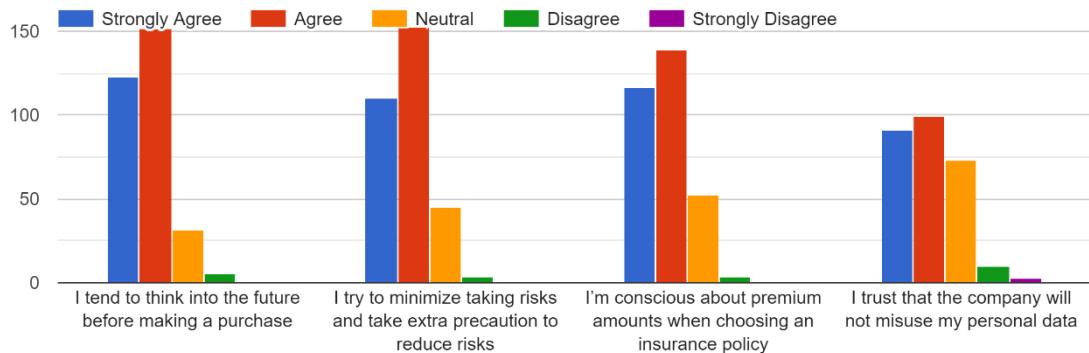


Figure 149 Survey Results - Opinion on risk and perceived prices

Financial perception

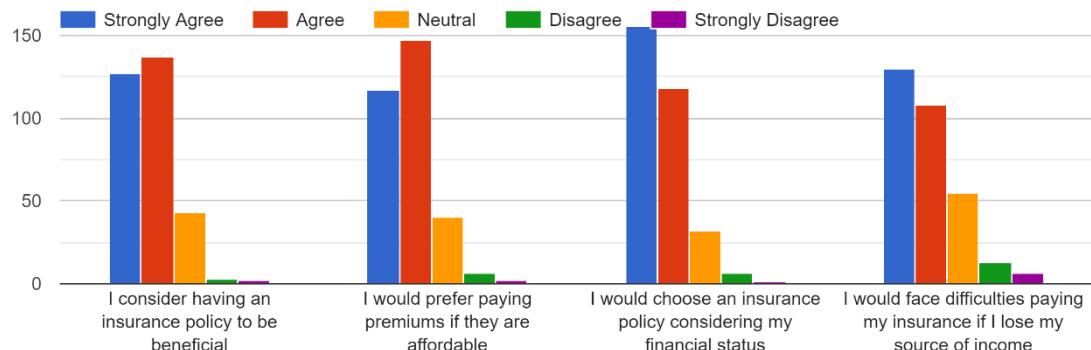


Figure 150 Survey Results - Financial Perception

Discounts and other benefits

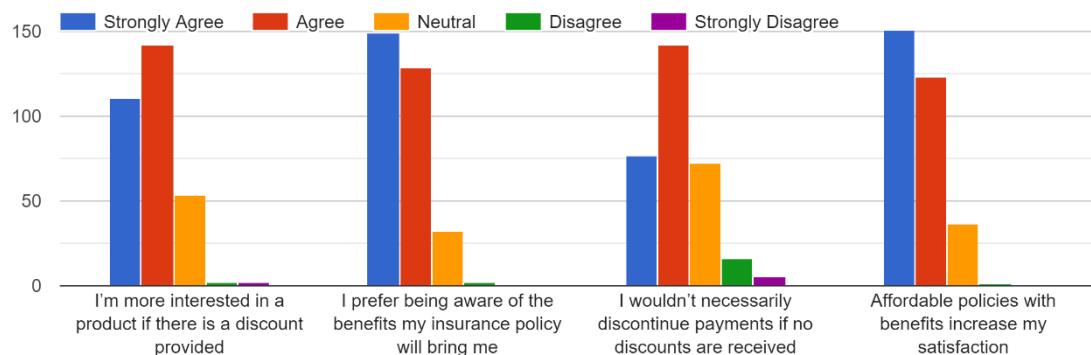


Figure 151 Survey Results - Discounts and other benefits

Do insurance companies do a good job assuring customers?

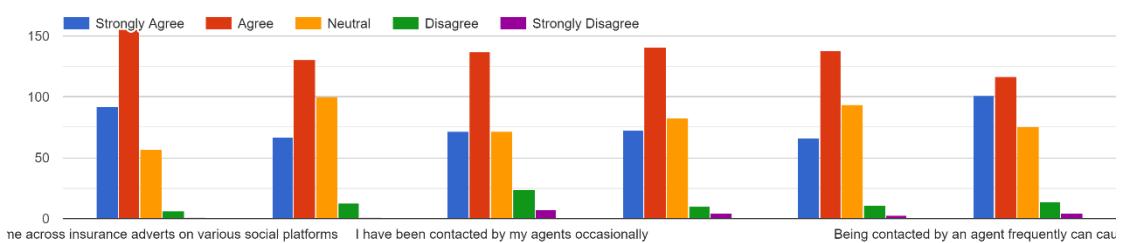


Figure 152 Survey Results - Assurance

Inflation and competition

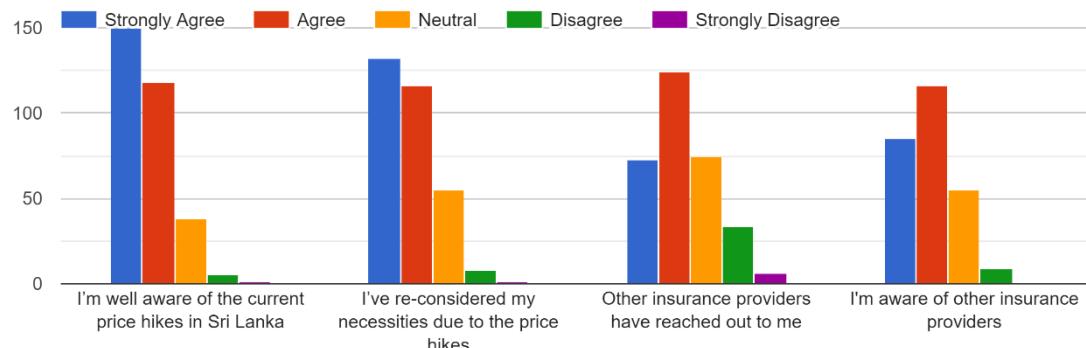


Figure 153 Survey Results - Inflation and Competition

Overall feedback

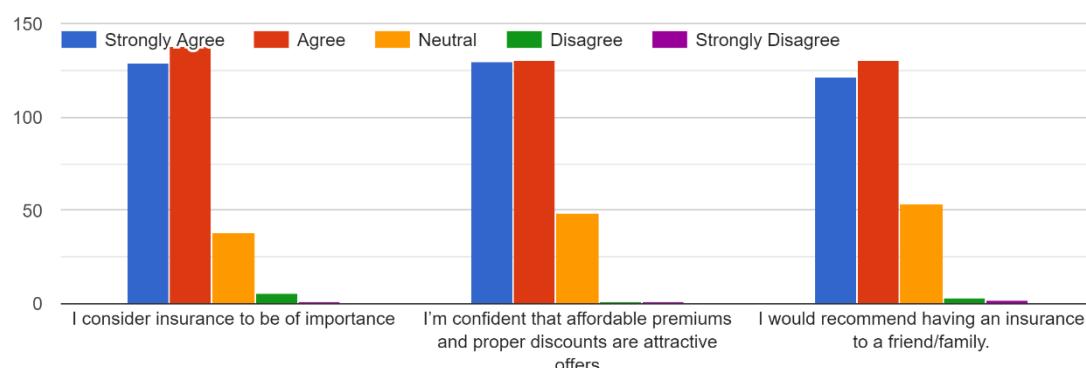


Figure 154 Survey Results - Overall Feedback