Understanding The Customer Lifetime Value of the Insurance Industry and Its Influencing Factors

Zhen Tian*

Corresponding author. Email:18332934983@163.com

School of Economics and Management, Southwest Petroleum University, Sichuan, China

Abstract. Customer lifetime value is the contribution of all profits that the enterprise get from the trading activities from their users. Companies and entities can use it as a method to increase their operating profit. Previous work predicted customer lifetime value based on customer-related cash flows and traditional financial metrics, but ignored the endogenous causes of customer lifetime value. This paper is designed to analyze several factors which affect the customer lifetime value. Taking the insurance industry as an example, through the establishment of a model and the implementation of analysis, the study figures out the connection between customer index and its lifetime value, and how can the region, income, marriage, purchase independent and comprehensive affect its lifetime value.

Keywords: customer lifetime value, insurance industry, factor analysis

1 Introduction

Customer value and firm profit are inextricably linked by a reinforcing relationship that often is endogenous to the firm through updating of customers' experiences and perceptions - this is the cornerstone of contemporary thinking in service marketing management [1]. For most organizations, understanding customers will be the key to success, while lack of this understanding can be a recipe for failure [2].

The Economist Intelligence Unit found two of the most important performance assessments that related to customers, which would be monitoring in the future by the companies in their survey, the first one was customer profitability and the second one was customer lifetime sales [3]. However, customer lifetime value measurement is of crucial importance for companies in managing demand more effectively than do their competitors [4]. If enterprises hope to obtain and retain high-value customers, they need to understand the influencing factors of the lifetime value of customers and pertinently improve it, so as to enhance the competitiveness of enterprises and occupy a dominant position in the market.

At present, the customer lifetime value is being used more and more widely in the sales field, especially in the insurance industry. People have tried to figure the customer lifetime value by combining FCM clustering algorithm and fuzzy AHP method [5], and the customer lifetime value assessment was based on the RFM analysis according to the regency, frequency and monetary shown in the RFM model [6], benchmarking the methods results, the stochastic model prediction approach and classic choice of RFM value variables [7].

ICBBEM 2024, March 29-31, Wuhan, People's Republic of China

DOI 10.4108/eai.29-3-2024.2347317

Based on those conclusion and the statistical data of an insurance company, this paper focuses on the analysis of customer lifetime value through the establishment of multiple linear regression model, non - parameter factors test and factor analysis, and provides a direction for the customer service of the insurance industry.

2 Analysis of Multiple Linear Regression Models

Regression analysis is a practical model, by fitting a set of collected data, it can indicates lots of information. Multiple linear regression analysis is a branch of the whole regression analysis, which can be taken to model a single response variable using several independent variables [8]. In this paper, this model can screen the variables and derive a reasonable correlation of the variables to the customer lifetime value.

2.1 Stepwise Regression

The interpreted variable is customer lifetime value. The explanatory variables are number of policies, claim amount, area, marital status, income, qualification, vintage, policy, type of policy. Screened by multiple linear regression, we can figure out that number of policies, claim amount, area, marital status, income were explanatory for the dependent variable. Variables like qualification, vintage, policy, type of policy are excluded. As shown in Table 1, Durbin - Watson test shows a result of 1.996, which means that there is no significant auto-correlation between the various independent variables.

Table 1. Model summary [Owner-draw]

Model summary [g]					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin - Watson
1	.386[e]	.149	.149	83605.014	1.996

a. Predictors: (Constant), number of policies, claim amount, Urban, married, income
b. Dependent Variable: customer lifetime value

2.2 Significance Test of The Regression Equation

As shown in Table 2, F statistics=2606.723. If the significant level is α=0.05, since P is less than α, the null hypothesis of the significance testing of the regression equation should be rejected. The regression coefficient is considered to be zero, and the linear relationship between the whole body of the explanatory variable is significant. The regression equations passed the significance test.

Table 2. ANOVA [Owner-draw]

ANOVA [a]						
	model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	109187182036742.380	5	21837436407348.477	3124.187	.000[f]
	Residual	624790110959748.800	89386	6989798301.297		
	Total	733977292996491.100	89391			

a. Predictors: (Constant), number of policies, claim amount, Urban, married
b. Predictors: (Constant), number of policies, claim amount, Urban, married, income

2.3 Regression Equation Coefficient Test

Table 3. Coefficients [Owner-draw]

Coefficients a

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
5	(Constant)	-17076.087	1819.104		-9.387	.000	-20641.512	-13510.663		
	number of policies	64804.123	606.638	.335	106.825	.000	63615.120	65993.127	.967	1.034
	claim amount	2.853	.097	.103	29.264	.000	2.662	3.044	.773	1.293
	Urban	7748.414	696.319	.039	11.128	.000	6383.636	9113.192	.766	1.305
	married	-6565.596	570.786	-.036	-11.503	.000	-7684.331	-5446.861	.982	1.018
	income	-2612.623	429.726	-.019	-6.080	.000	-3454.882	-1770.364	.927	1.079

a. Dependent Variable: customer lifetime value

As shown in Table 3, by the T test, variables like number of policies, claim amount, area, marital status, income have significant effects on customer lifetime value respectively.

According to the regression model, we can conclude that:

(i) Customer lifetime value is significantly related to the number of policies purchased by the customer. The more policies they buy, the greater customer lifetime value they show. Customers who buy more than one insurance business have a significantly higher lifetime value than those who buy only one single business, showing a positive correlation.

(ii) A significant positive correlation is used to provide the relationship of customer lifetime value and the claim amount made by those clients. The higher the claim amount requires, the greater the lifetime value of the customer brings.

(iii) The lifetime value of customers living in urban is significantly higher than those living in rural areas, and the lifetime value of urban customers is higher than rural customers and the figure is 7660.013. It shows that the urbanization of residence can have a positive impact on the lifetime value of resident customers.

(iv) The lifetime value of married customers is significantly lower than single customers, and the figure is 6735.335. It shows that marriage has a negative impact on the lifetime value of customers.

(v) Customer lifetime value is significantly correlated with the income of customers. Customer income is divided into four groups in the order from low to high, the negative result has an ability to show that the income of customers is negatively correlated with its lifetime value.

After reaching the above conclusion, we can draw the factor determining formula for the customer lifetime value. It is surprising to find out that income and marriage match situation have negative influence on customer lifetime value and the area also need to be further analyzed. However, most coefficients of the variables cannot reflect the amplitude of fluctuation directly, especially the income, area and marital status. They have only directional significance no numerical significance. To more specifically demonstrate the actual impact of variables on customer lifetime value and make the results more accessible, the author conducted the normality

test and non-parametric test for non - continuous variables, and further analyzed the influence on the lifetime value of customers by comparing the percentiles of each group of data. It will be analyzed and elaborated in the following text.

3 Non - Parametric Test

The non - parametric methods is an extremely useful technique, and the main advantage in using non - parametric techniques is the avoidance of such assumptions [9]. Through non - parametric analysis and interquartile graph, we can directly derive different data on the lifetime value of customers under different groups.

3.1 Income Impact Factors

Table 4. Tests of Normality [Owner-draw]

Tests of Normality							
		Kolmogorov-Smirnova			Shapiro-Wilk		
	income	Statistic	df	Sig.	Statistic	df	Sig.
customer lifetime value	<=2L	.243	1776	.000	.715	1776	.000
	2L-5L	.235	20327	.000			
	5L-10L	.249	50538	.000			
	More than 10L	.243	13005	.000			

a. Lilliefors Significance Correction

Table 5. Test Statistics [Owner-draw]

Test Statistics a,b	
	customer lifetime value
Kruskal-Wallis H	614.212
df	3
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: income

As shown in Table 4, P is less than 0.05, reject the assumption that variables follow a normal distribution, using a non-parametric test whether there are significant differences between different groups. And Table 5 indicates that income can pass the non-parametric test.

Table 6. Percentile Table [Owner-draw]

	customer lifetime value	P
≤2L	67464 (58899~126783)	0
2L-5L	68280 (58380~113238)	
5L-10L	65856 (50520~101676)	
More than 10L	65604 (50091~99048)	

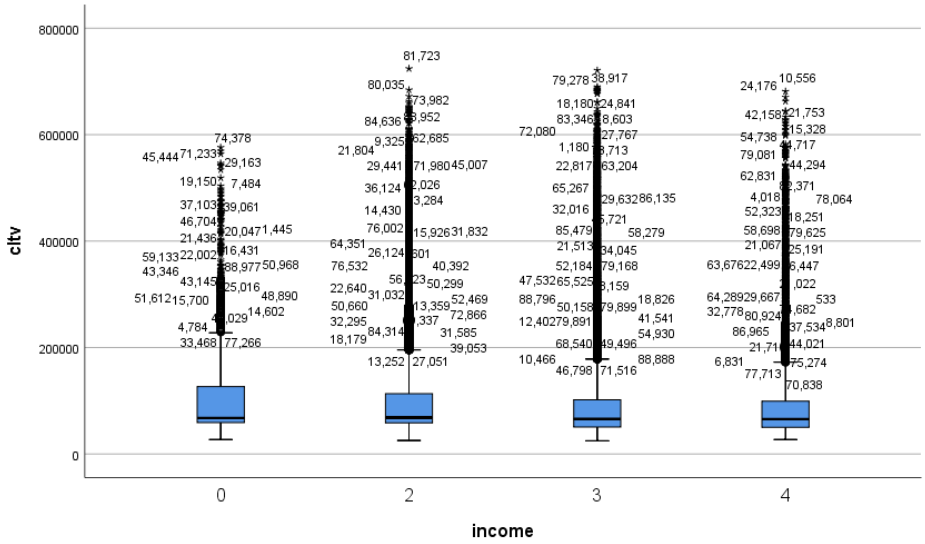

Fig. 1. Interquartile of income [Owner-draw]

According to the difference shown in Table 6 and Figure 1 between the different groups, the mean and interquartile of customer lifetime value produced show a gradual decline as the customer income changes from group 0 to group 4. It can further be concluded that customers' income has a negative impact on their customer lifetime value, but the customer lifetime value does not change a lot, as we can indicate in Table 6, the number of customer lifetime value only reduced by 1860.

3.2 Marital Status Impact Factors

Table 7. Tests of Normality [Owner-draw]

Tests of Normality				
		Kolmogorov-Smirnova		
	Marital status	Statistic	df	Sig.
customer lifetime value	single	.246	36386	.000
	married	.247	49260	.000

a. Lilliefors Significance Correction

Table 8. Test Statistics [Owner-draw]

Test Statistics a	
	customer lifetime value
Mann-Whitney U	873051514.500
Wilcoxon W	2196319804.500
Z	-27.022
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable:marital_status

As shown in Table 7, P is less than 0.05, reject the assumption that variables follow a normal distribution, using a non-parametric test whether there are significant differences between different groups. And Table 8 indicates that income can pass the non-parametric test.

Table 9. Percentile Table [Owner-draw]

	customer lifetime value	Z	P
single	68196 (56676~108912)	-27.022	0
married	65292 (49824~99420)		

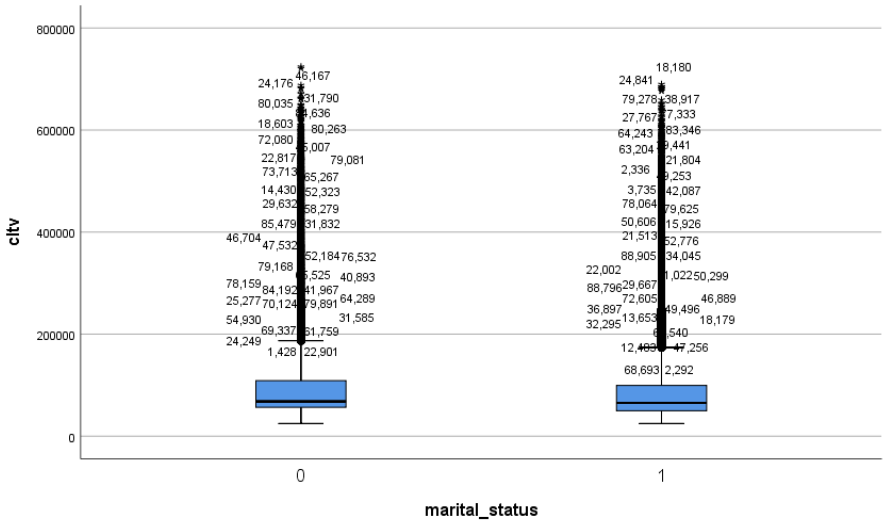

Fig. 2. Interquartile of marital status [Owner-draw]

Different marriage conditions have significantly different value to customers. According to the difference shown in Table 9 and Figure 2 in different groups, single customers have a larger customer lifetime value than married customers. It can be further concluded that the marriage situation of customers has a negative impact on their customer lifetime value and the number of customer lifetime value reduced by 2904 from single client to married client. We can figure out that marital status has a more significant effect on customer lifetime value than income.

3.3 Area Impact Factors

Table 10. Tests of Normality [Owner-draw]

Tests of Normality				
		Kolmogorov-Smirnova		
	area	Statistic	df	Sig.
customer lifetime value	Urban	.246	60176	.000
	Rural	.247	25470	.000

a. Lilliefors Significance Correction

Table 11. Test Statistics [Owner-draw]

Test Statistics a	
	customer lifetime value
Mann-Whitney U	701743471.500
Wilcoxon W	1064557924.500
Z	-39.386
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable:area

As shown in Table 10, P is less than 0.05, reject the assumption that variables follow a normal distribution, using a non-parametric test whether there are significant differences between different groups. And Table 11 indicates that income can pass the non-parametric test.

Table 12. Percentile Table [Owner-draw]

	customer lifetime value	Z	P
rural	63924(45624~90480)	-39.386	0
urban	67596(56340~108420)		

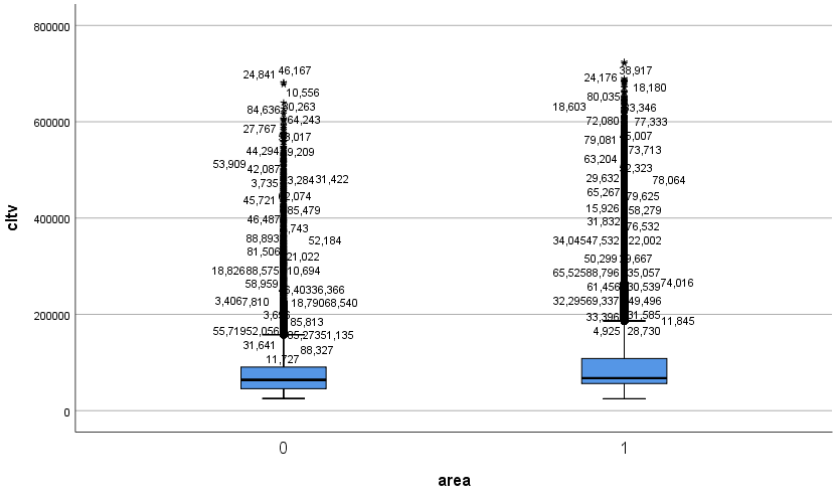

Fig. 3. Interquartile of area [Owner-draw]

Different regional conditions have significantly different customer lifetime value. According to the difference in percentiles of different groups shown in Table 12 and Figure 3, customers living in cities have greater customer lifetime value than customers living in towns. It can be further concluded that the urbanization of the customer's residence has a positive impact on their customer lifetime value. It is easy to find out that the number of customer lifetime value increases by 3672 from rural clients to urban clients and the area has a most significant effect on customer lifetime value in terms of the three variables.

4 Factor Analysis

Factor analysis, which can be used to analyze some underlying structure between the variables, is an effective method [10]. To provide users with clearer clarity of influencing factors, this paper will further compound the influence variables of customer lifetime value by means of factor analysis.

4.1 KMO and Bartlett Tests

The Bartlett sphere test and KMO test are used to analyze whether the variables are correlated. Table 13 indicates that the result of KMO test measure is 0.605, indicating suitability for factor analysis. The probability value of the Bartlett test is 0.000, which is 0.05 below the significance level, which shows that the data base is suitable for the following factor analysis.

Table 13. KMO and Bartlett's Test [Owner-draw]

KMO and Bartlett's Test		
Kaiser - Meyer - Olkin Measure of Sampling Adequacy.		.605
Bartlett's Test of Sphericity	Approx. Chi-Square	29685.113
	df	10
	Sig.	.000

4.2 Total Variance Interpretation

According to the standardized data, the correlation coefficient matrix, eigenvalues and eigenvectors of the variables are established, so as to obtain the factor feature root and variance contribution rate of the influencing factors.

Table 14 together with Figure 4 can show that there was a second significant decrease between 4 and 5, so four common factors were extracted accordingly using principal component analysis. 89.214% is the result of final cumulative variance rate the four factors can contribute, reflecting that the information is primitive and usable for research.

Table 14. Total Variance Explained [Owner-draw]

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.700	34.000	34.000	1.700	34.000	34.000	1.452	29.049	29.049
2	1.020	20.399	54.399	1.020	20.399	54.399	1.005	20.090	49.139
3	.908	18.151	72.550	.908	18.151	72.550	1.003	20.069	69.208
4	.833	16.664	89.214	.833	16.664	89.214	1.000	20.006	89.214
5	.539	10.786	100.000						

Extraction Method: Principal Component Analysis.

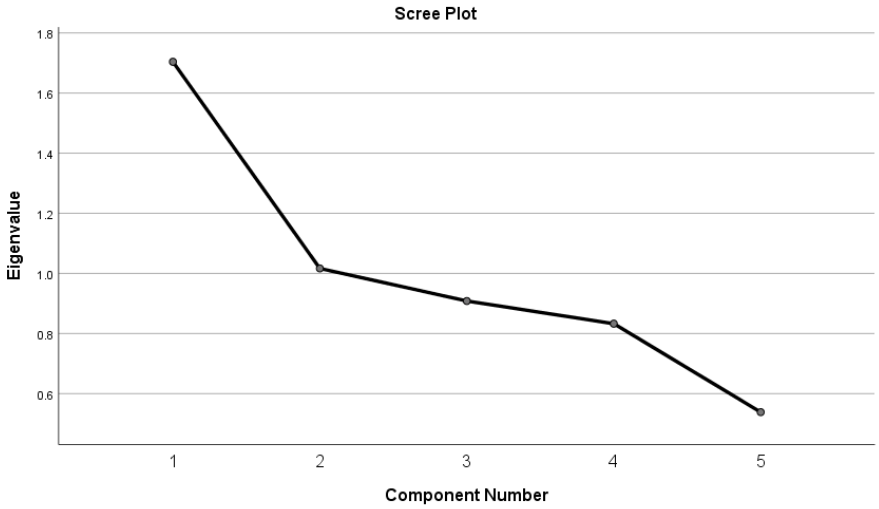

Fig. 4. Scree Plot [Owner-draw]

4.3 The Explanatory Factors of The Factor

Table 15. Rotated Component Matrix [Owner-draw]

Rotated Component Matrix a				
	Component			
	1	2	3	4
area	.832			
income			.985	
marital_status		.995		
claim amount	-.853			
number of policies				.995

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 4 iterations.

According to Table 15, we can conclude from the factor analysis that:

(i) Factor 1 has a high correlation with the variable region and the claim amount, so factor 1 can be named as a regional influence factor. The customer is located in different urban and rural areas, and the living situation and its attitude towards consumer products claims affect the customer's lifetime value.

(ii) Factor 2 has a high correlation with variable marriage status and can be named as family influence factor. Family is an important consumption unit, and the marriage situation will significantly affect the choice of customers' purchase decision. Marriage is the first step to establish a family, and the family life cycle will lead to the loss of lifelong value of customers.

(iii) Factor 3 has a high correlation with variable income and can be named as an income influence factor. The income of customers significantly affects the lifetime value that customers

can bring to the enterprise. The potential consumption habits of high-income customers make them have higher demands, reduce their returns, and erode the profitability of the enterprise, which is most likely to lead to net loss.

(iv) Factor 4 has a high correlation with the amount of insurance purchased by the variable, and it can be named as the consumption influence factor. The amount of insurance purchased has a significant impact on the lifetime value of customers. Customers who participate in more than one insurance obviously have a higher lifetime value. It can be concluded that the lifetime value of customers is significantly related to the customers' re-purchase of enterprise products.

5 Conclusion

Based on the data provided by the insurance industry, this paper makes a unified and grouped analysis of several factors that may affect the lifetime value of customers. It is concluded that the lifetime value of customers is affected by four factors: region, family, income and consumption, among which income has unexpected side effects on the lifetime value of customers, and at the same time, marriage will significantly reduce the lifetime value brought by customers. Accordingly, the insurance industry can target its target customers and provide targeted services to achieve the purpose of profit.

References

[1]Liu, B.S.C., Petruzzi, N.C., Sudharshan, D. (2007) A service effort allocation model for assessing customer lifetime value in service marketing. Journal of Services Marketing, 21: 24–35.

[2]Hajipour, B., Esfahani, M. (2019) Delta model application for developing customer lifetime value. Marketing Intelligence & Planning, 37: 298-309.

[3]Ryals, L.J., Knox, S. (2005) Measuring risk-adjusted customer lifetime value and its impact on relationship marketing strategies and shareholder value. European Journal of Marketing, 39: 456-472.

[4]Ekinci, Y., Uray, N., U¨lengin, F. (2014) A customer lifetime value model for the banking industry: a guide to marketing actions. European Journal of Marketing, 48: 761-784.

[5]Safari, F., Safari, N., Montazer, G.A. (2016) Customer lifetime value determination based on RFM model. Marketing Intelligence & Planning, 34: 446-461.

[6]Tsai, C.F., Hu, Y.H., Hung, C.S., Hsu, Y.F. (2013) A comparative study of hybrid machine learning techniques for customer lifetime value prediction. Kybernetes, 42: 357-370.

[7]Sant'Anna, A.P., Ribeiro, R.O.A. (2009) Statistical modeling and probabilistic composition in the prediction of the customer lifetime value. Benchmarking: An International Journal, 16: 335-350.

[8]Blyth, K., Kaka, A. (2006) A novel multiple linear regression model for forecasting S-curves. Engineering, Construction and Architectural Management, 13: 82-95.

[9]Bower, J.A. (1998) Statistics for food science – V part C: non-parametric ANOVA. Nutrition & Food Science, 2: 102–108.

[10] Saeed, B., Tasmin, R., Mahmood, A., Hafeez, A. (2022) Development of a multi-item Operational Excellence scale: Exploratory and confirmatory factor analysis. The TQM Journal, 34: 576-602.