A Study of Customer Profitability in Automotive Retail

By Pierre-Paul Savoie

A research project submitted to

Saint Mary's University, Halifax Nova Scotia

In partial fulfillment of the requirements for

Degree of Master of Business Administration

December, 2014 Halifax Nova Scotia

Approved:

Dr. Ramesh Venkat

Associate Professor of Marketing

Approved:

Dr Margaret MacKee

Program Coordinator, MBA Program

Date: December 21st, 2014

Acknowledgement

I would like to acknowledge various people for helping me the course of this research project. First and foremost I would like to thank my supervisor, Ramesh Venkat, for his unwavering help and support throughout this project. This paper would never have been possible without his guidance, invaluable advice and most of all, extreme patience.

I also want to thank the dealer principal for allowing me to access and study his dealerships data for this project; this project would never have been possible without his support.

Finally I would like to thank Linda for her continued love and support in letting me complete my undergraduate degree and the MBA. I’m done with school, for now.

A Study of Customer Profitability in Automotive Retail

By Pierre-Paul Savoie

Abstract

The main purpose of this study was to explore Recency, Frequency & Monetary Value (RFM) analysis and Customer Lifetime Value (CLV), specifically looking at an automotive retail dealership. Though both these concepts are well known and have been used in many industries around the world, they are not used by automotive retailers.

This study attempted to identify the drivers of profitability in the dealership and to see if RFM analysis could be used to segment the customer database by profitability. The study also aimed to determine whether we could model projected CLV value from the RFM profile. In addition to using RFM and CLV analysis, the study used multiple regression, analysis of variance and cluster analysis.

The study identified RFM based customer segments and we also used CLV to model projected profitability within brands of vehicles. Several variables, including brand and customer longevity were found to be predictors or customer profitability and CLV.

The study shows that there is definitely a link between RFM and CLV. It also illustrates that more research needs to be done in this industry to develop accurate Customer Lifetime Value models, which encompass all the variables that affect profitability in both the sales and service areas of the auto dealership business.

December 21st, 2014

Table of Contents

List of Tables

1. Introduction

1.1. Automotive Dealership Industry

The vast majority of auto dealers are owned by private individuals or private corporations, very few are controlled by publicly traded corporations. Manufacturers will occasionally partner with an individual or group to open a new sales point that they will over time purchase from the manufacturer.

Whether a dealership is individually owned or is part of a group of dealers, the business itself operates essentially the same from store to store. Some details may change depending on what franchise or Original Equipment Manufacturer (OEM) they represent but the foundation of where revenue is generated is the same across all dealerships.

1.2. Dealership Customer Database

Dealership Management Systems (DMS) now allow the dealership marketing team to easily access the complete customer database with some measure of facility; from this data they can extract data which shows their customers' actual purchasing patterns and preferences rather than their intentions. Previously dealers were limited to sampling customer intentions through dealership and manufacturer surveys, as well as Customer Satisfaction Index (CSI) scores. CSI surveys are still performed today, but are more focused on new vehicle purchases and post warranty repair opinions and results. It is also important to note that a survey measuring both satisfaction and intent to purchase creates a strong method bias.

1.3 Sources of Revenue

Dealerships are essentially two separate businesses under one roof: the front-end, also known as variable operations and the back-end, also known as fixed operations.

The variable operations are responsible for sales of new and used vehicles as well as sales in the business office; these would include but are not limited to auto financing, extended warranties, service plans, protection products such as undercoating, protection packages (rust, paint & fabric), all of which are sources of revenue for the organization.

Fixed operations are responsible for ensuring all vehicles are ready to be sold, the long term servicing of vehicles for the end user, as well as the sales of parts and accessories to the service department and retail purchasers. This department also deals with other repair facilities such as corporate fleet and government facilities as well as small local service providers, which are collectively known as wholesale customers. For most dealerships this is a modest segment of their business and provides lower profit margins. This research project focuses more on retail customers where segmentation and individual attention will have greatest impact.

1.4 Importance of Measuring Customer Profitability

The ability to measure both long term and short term profitability is critical in determining the following: first, how to manage the customer base and target your marketing and communication strategies; second, how to identify, retain and nurture the right clients to maximize the return on their staff resources. Identifying which customers are most profitable, or have the potential to become more profitable through objective segmentation, gives us insight "into what customers want and how they behave, and marketing decisions made are evidence-based and result in more profitable outcomes from one-to-one customer interactions" (Vergara, 2009, p. 30).

Not all customers are equally profitable. It may even be desirable to fire some customers which are a drain on the resources of the firm, as many studies have shown, in order to allocate the finite resource on more potentially profitable customers. Increased focus on customer profitability has caused many companies to change their marketing strategy. Rather than focusing on acquiring new potentially unprofitable customers, which also typically have higher acquisition costs, they are "switching to defensive strategies" (Murby, 2007, p. 34), targeting retention of existing customers and increasing their spending with the firm. Utilizing tools such as RFM and (CLV), which are individualized (disaggregated) measures, allows the marketer to specifically identify customers in this regard (Gupta, et al., 2006), and will enable dealers to determine where and how to allocate resources.

1.5 Research Questions:

This Major Research Project (MRP) will address the following research questions using data from an automotive dealership:

Q1- Can we identify drivers of profitability?

Q2- Can customer database be segmented to identify which customers are the most profitable?

Q3- Can RFM segmentation be used to segment and predict profitability?

Q4- Can we model Projected CLV from the RFM profile?

1.6 Challenges Facing Automotive Dealerships

The financial crash experienced in 2008 triggered the closure of nearly 325 dealers in Canada between 2008 and 2010 with auto unit sales dropping from 1.7 million to 1.5 in 2009. Auto sales through 2013 have now recovered back to the 1.7 million unit sales level, but the number of dealerships has not increased proportionately. Even accounting for new dealers opening in zones that did not previously have auto dealers, there are still roughly 200 fewer dealerships in Canada in 2014 than there were in 2008. To survive, the dealer industry has implemented austerity measures and required that staff perform duties that had once been done by multiple personnel. As sales have recovered generally, dealerships have tried to not increase staffing levels and have required staff to do even more. Providing management with an easily understood way to segment their customers and determine appropriate actions is critical now more than ever.

1.6 Decreasing Margins and Increasing Market Pressures

The automotive retail industry has remained more or less stable in the service side of the business for more than 20 years; the target gross profit percentage for service labor for non-maintenance related items has remained at 70% for at least 10 years now. Competitive pressures from non OEM maintenance providers of basic services has forced dealerships to keep the prices lower for these services, therefore reducing the potential Gross Profit. The common thought process is to treat these services as a loss leader; hoping customers become habituated to return for maintenance, which will then provide service departments the opportunity to offer customers more profitable services.

1.7 Privacy Legislation

The latest challenge for the industry is the Canadian Anti-Spam Law (CASL) legislation enacted in July 2014, which requires implicit confirmation that dealerships have received permission to market to their customers by mail, phone and email.

According to an article by Li & Mee (2014), CASL generally requires express consent. Consent is implied in the following circumstances:

(a) The recipient and sender have an "existing business relationship" (EBR) or "existing non-business relationship" (ENBR) as defined by the legislation. For example, an EBR exists if the sender and recipient have engaged in specified types of business together in the previous two years (e.g. purchase of a product or service, written contract).

(b) The recipient has conspicuously published their electronic address, or has provided their electronic address to the sender, without indicating that they do not wish to receive unsolicited Commercial Electronic Messages(CEMs), and the CEM is relevant to the person's role in a business or official capacity.

This has made the use of automated communications much more complex for dealers, as they now have to get the express permission from the customer before communications can be sent to them. The use of automated voice, emails, text messages can no longer be used, unless explicit permission has been granted by the customer.

Considered one of the most stringent anti-spam regimes in the world, CASL will have a significant impact on the electronic communication practices of businesses operating in the Canadian marketplace (Li and Mee, 2014).

1.8 Challenges in Measuring Customer Behavior and Profitability:

Long term profitability is not widely known or understood in the Auto industry, and although customer lifetime value is conceptually understood, it is not calculated precisely.

Dealership Management Systems (DMS) generally do not aggregate enough information to produce predictive analytics or calculate life time value and do not have any segmentation tools. Thankfully some DMS providers are now making it possible to extract large databases of customer information, which will allow marketers to identify groups of people with similar attributes (segments); “by matching marketing expenditure against the anticipated reaction of identified segments of the customer base they can manage marketing expenditure to optimize returns” (Murby, 2007).

1.9 Expected Contribution to Research

This research paper aims to broaden the understanding of analysis tools for automotive dealerships, by understanding of the importance of RFM analysis, customer life time value and customer segmentation. The ability to identify the profiles of customers to better manage customer communications and marketing efforts is important for dealerships, in order to maximize the return on their investment in these activities. Hopefully this research will provide a foundation for further study and development of models, which will enable the industry to manage their customers more effectively from both an effectiveness of communications and profitability standpoint.

II. Literature Review

The following is intended to be an overview of some of the literature pertinent to the research on automotive dealership and profitability measurement in general. It first looks at why identifying and measuring customer profitability is important and its impacts on marketing decision making and actions.

Second, we explore the first of two analysis concepts, Customer Lifetime value (CLV) which is widely used in many industries to profile and manage customer marketing and determining not only which are the most profitable customers, but also which can potentially be groomed to become more profitable (Yang, 2004). We will look at several models developed for specific uses to show how these models can be adapted to the needs of the analysis being performed. We begin with a basic structural model and build from there to more complex models developed for various industry needs, the last of which is an auto industry specific example for vehicle purchases.

Third, we look at RFM definition and models and how this concept and tool can be adapted for varying uses in analyzing and clustering customer segments together to determine managerial actions and develop plans to individualize communication and programs for varying segments. We also discuss strengths and weaknesses of RFM models, which leads us to suggestions on how utilizing RFM can be optimized in itself and why combining RFM with other methodologies results in a more comprehensive picture of individual customers and more accurate segmentation of the customer base.

2.1 Why Identifying Profitable Customers is Important

Traditionally dealerships have considered customer loyalty as a key measurement of success. Although important, this measurement can lead managers to make marketing and process decisions which are flawed, because a customer who is loyal to the firm may not be profitable (Jain & Singh, 2002) and may even be a drain on resources, in some cases. Research suggests that more comprehensive metrics such as Customer Lifetime Value (CLV) and RFM provide the ability to rank and group customers more accurately and therefore enable dealerships to customize the strategies to manage customers.

Ultimately, the goal for dealerships is to identify profiles for profitable customers and manage their marketing efforts to acquire and retain those types of customers, whilst also identifying the profiles of unprofitable customer and developing marketing strategies to manage both. Murby (2007, p. 34) suggests that "the goal of any business is not to improve customer or employee satisfaction at any cost, but rather to manage these relationships and the drivers of customer profitability to improve corporate performance".

Customer value is a marketing metric which, if well understood, will become a critical aid in decision making and optimizing marketing efforts of the dealership and measuring the effectiveness of marketing programs. Customer value can be expressed in many terms, including gross profit, net profit and contribution margin. Murby (2007, p.34) further suggests that, to achieve this, a company "… should identify the most and least profitable elements of its total customer base (and those in between) and manage these relationships accordingly".

As with most industries, automotive retail recognizes that the cost of acquiring new customers is much higher than retaining current customers, hence dealerships need to focus on these relationships.

2.2 CLV - Customer Lifetime Value

Customer life time value can be defined in various ways but at their core these definitions share a common thread. CLV is the present value of all incomes derived from the customer less the costs to serve these customers over time period in which the customer does business with the organization. One of the great advantages of CLV and RFM models is that they are built using transaction data for the entire customer base and not limited to a sample of customers as are attitudinal surveys (Gupta et al., 2006). Calculating the CLV for all of a firm's customers allows them to categorize customers based on their individual contribution to the organizations profits This helps to develop strategies to deal with each customer differently, instead of treating every customer the same way using the same marketing approaches.

Although firms are interested in knowing the current and predicted customer life time value of their customers, they also need to identify the factors they can control that can potentially increase the value of customer. It is not enough to know who are the most profitable customers, it is even more important to determine how to convert currently less profitable customers into more profitable ones.

Another important factor is that in most cases in most modeling approaches, competition is ignored due to the lack of competitive data (Gupta et al., 2006). Specifically in the automotive retail industry the competitors face the same pressures and relatively the same variable costs in regards to automotive technicians and support staff within their respective markets. Therefore, trying to factor this into the analysis is unlikely to add to the quality of the data. This is not the case in many other industries where the CLV is highly dependent on market dynamics; and most typical where margins and retention of customers fluctuate over short periods of time and would require a regular re-evaluation of CLV (Gupta et al., 2006).

2.3 Customer Lifetime Value Models

Researchers have developed a multitude of CLV models in an attempt to adapt to the myriad of business structures. What follows is a brief review of some of these models to help understand the basic ways one can start to build on basic formula and adapt to the various needs of the organizations being studied. It is not intended as an exhaustive list, but more of an overview of different ways these models can be viewed.

It was suggested by Jain & Singh (2002) that there are four basic types of models used to evaluate CLV, each of which can be adapted based on data or specific industry / user.

2.3.1 Basic Structural Model

$$\text{CLV} = \sum_{i=1}^{n} (R_i - C_i)/(1+d)^{i-0.5}$$

i= period of cash flow from customer transactions

R_i= revenue generated from customer in period i

C_i= total cost of generating Ri in period i

d= discount rate, which represents the cost of capital or time value of money

n= total number of periods of projected life time for the customer

This Particular Model assumes that all transactions take place at the end of a period

2.3.2 Customer Migration Model

In their article Jain & Singh refer to the work done by Dwyer (1997) where he proposed that there are 2 broad types of customers: "always-a-share" and "lost for good". The first assumes that customers will rely on several service providers and will adjust the share of business done with each as needed. In the second, customers have made a long term commitment, because the switching costs will be high and their transactions cannot be redeployed easily. Dwyer further defined the two types of customer categorization by suggesting that lost-for-good should be viewed as a customer retention problem and used a slight variation of the basic model of CLV.

For always-a-share Dwyer suggested a customer migration model in which purchase recency is factored into the model in order to predict future customer purchase behavior.

This model has certain advantages as it incorporates the purchasing probability into the model. In other words, even though a customer may not have purchased in a given period, they are still considered to have been retained. This model also has some weaknesses, because it assumes that time periods are fixed and the sales and cash flow occur at the same time in a period. It makes sense for businesses with regular cash flow, but would be less accurate for businesses where cash flow is more erratic.

2.3.3 Optimum Resource Allocation Models

Blattenberg & Deighton (1996) proposed models to find the optimal spending balance between customer acquisition and retention to maximize CLV. They also track customer equity gains and losses against the costs of marketing programs, which might highlight issues that might not be revealed by analyzing income statements and would help to put customers at the top-of-mind in strategic thinking. Their model is in two parts:

- Optimal level of acquisition spending

a= (ceiling rate) [1 – exp(- k1 *$A)]

Contribution margin from acquiring a prospect year 1 = a$m - $A

$A =acquisition expenditure

a= acquisition rate obtained as a result of $A

Ceiling rate =limit to attraction of new customers

k1= parameter controlling shape of exponent curve

- Optimal level of retention spending

This second formula describes how CE is dependent on retention spending:

***r*= (ceiling rate)[1 – exp(- *k2* *$*R*)]**

Year y contribution from retention = r[$m - $R/r]

$R= individual retention expenditure

r= retention rate as a result of expenditure

$m= margin earned from customer

*k*2= parameter controlling shape of exponent curve

In this model, the lifetime value is calculated by summing the annual values for the projected life of the customer, then adding in the first year and discounting to present value at the appropriate rate of return for marketing investments, which produces the customer equity value that can be attributed to that customer. This model also shows another use for CLV, as it can be used to determine optimal spending for marketing resources because it takes into account the expenses incurred to acquire individual customers.

Keep in mind that it also suffers from many of the same weaknesses as other models, in that it assumes cash flows remain constant and that they occur at the same time in each period (Jain & Singh, 2002).

2.3.4 Automotive Retail Specific Example

To show a more concrete example, which relates to the auto industry, we refer to a case written in 2008 for the Ivey School of business by Mike Moffat & Kyle B. Murray, in which they featured Conroy's Acura. This was a fairly comprehensive CLV model which took into account all the major factors and revenue streams related to vehicle purchases that are common in automotive dealerships: retention rate, acquisition cost, discount rate, gross profit per vehicle, and cost related to the retention of each customer.

Their formula is as follows:

$$CLV= ((1-d)^{y} * ((p-c-(m * y))/1-r)))$$

d= discount rate or time value of money (5% was used in their example)

y= length of time customer keeps their vehicle before trading it in for a new one

r= retention rate, the number of customer returning to dealership to purchase their next vehicle

m= yearly maintenance cost for retaining customer relationship

p= sales price of the vehicle

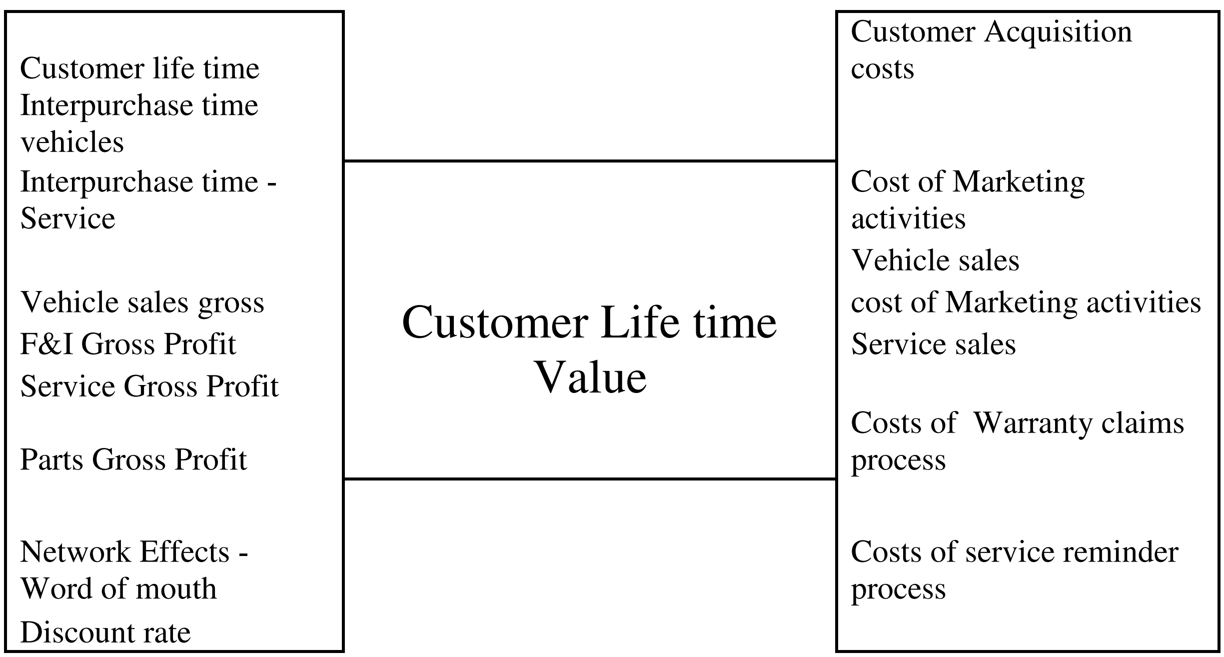
c= dealer cost of vehicle

This model does not factor in revenues derived from service and parts sales generated throughout the lifetime cycle of the vehicle, which can be significantly more than revenues generated by the sale of a new vehicle.

2.4 Factors which Affect Customer Lifetime Value

As discussed earlier there are many models for CLV, which have been developed for various industries and are modified to fit the individual Customer-Firm relationship. These models have to incorporate various industry-specific factors that affect CLV to achieve the most accurate measure possible. Jain & Singh (2013) proposed a basic template for identifying these factors, I have adapted this to the automotive industry to show what factors are in play for them.

Figure 1 – Factors Impact CLV



2.5 RFM Analysis - Recency, Frequency Monetary Value (RFM)

RFM analysis has been used to evaluate customer behavior for more than 30 years in many industries. In both B2B and B2C applications, RFM has proven to be a tried and true analysis method which most direct marketers will use as a matter of course (Yang, Oct 2004).

- *Recency* is a measure of how much time has passed since the customer last purchased from the firm. Many marketers feel that the most recent purchasers are more likely re-purchase than their less recent customers (Birant, 2011).
- *Frequency* measures how often a customer purchases from the firm over a set period of time. This measure assumes that customers with more purchases are more likely to re-purchase than less recent purchasers (Birant, 2011).
- *Monetary* takes into consideration the value of how much the customer has purchased over the same period of time or the average transaction value depending on the type of analysis being performed.

Although RFM was developed decades ago, it has not advanced appreciably and many researchers have suggested enhancing the model and/or supplementing it with other methodologies to more accurately segment the customer database (Neal 2004). RFM has a lot of appeal because it uses a common sense approach that most marketers and business owners can understand, hence its popularity. Although it is commonly used, RFM also has a number of drawbacks.

RFM is not forward looking and does not consider whether the customer will continue to do business with the firm. As such, this tool can only look at past purchase behavior and needs to assume that past behavior will reflect future behavior. RFM does not take into account externalities such as marketing actions, legislation and alterations to product lines and services that could help to predict the customer's future value to the firm. Also, if Weighted RFM is used, the weighting values given to R, F & M will greatly influence the calculation of customers' individual worth to the firm (Kumar, Venkatesan and Beckman, 2008).

Though counter intuitive, coding RFM based mostly on the experience of the marketer may skew the segments being developed and result in *overlapping syndrome*. This may happen if you eliminate cells with fewer categories in order to cut down on the complexity; whereby increasing the risk of losing significant segments in favor of insignificant ones. The hard coding process can also produce uneven segment sizes, which may also threaten segmentation efficiency (Yang, 2004).

As suggested by Birant (2011), there are several studies that have looked at other versions of RFM analysis, including Weighted RFM (WRFM), mentioned above. In WRFM, R, F & M are each calculated using a weighted value to highlight the relative importance of each value and enable the marketer to make 'intuitive judgments' about customer ranking wR , wF & wM.

Other examples include TRFM (Timely, Recency, Frequency and Monetary), RML (Recency, Monetary and Loyalty), and FRAT (Frequency, Recency, Amount and Type of goods).

TRFM can be used to analyze products with demands at different times, whereas RML is an adaptation of RFM which can be used for environments in which transaction periods are annual in nature, and where loyalty is the frequency normalized over an annual period. Interestingly, in her Doctoral thesis on a specific dealership analysis Bonicalzi, M. (2004), the current Dean of the Canadian of Automotive Business School of Canada, found that through a correlation analysis “between customer satisfaction, loyalty and profitability shows that there is a low correlation between tenure (the Loyalty measure) and NPV total net profit (the profitability measure)”. She also determined that “there is no correlation between CSI (measure of satisfaction) and tenure or between CSI and NPV total Net Profit”.

FRAT is another extension of the RFM model that provides an improvement on the segmentation potential of RFM by taking into account the category or type of goods purchased by the customer. For example: “0- no buy, 1– buy compact car, 2– buy an economy car, 3- buy a mid-size car, 4- buy a luxury car where the order is defined in increasing order of size“ (Birant, 2011).

There are also versions of RFM that are focused on electronic and social media such as RFD (Recency, Frequency, and Duration), which was suggested to be used for analyzing how sticky a website is i.e. how much time is spent on a website; and RFR (Recency, Frequency, and Reach), which has been proposed for “social graph”: Recency – for the last post, Frequency – total number of posts, and Reach – number of networks or friends.

According to the review performed by Wei et al (2010), RFM has several advantages and disadvantages for decision makers as it is a very cost effective way to acquire data and quantify customer behavior. RFM is valuable in predicting customer responses and affecting company profits in the short term. Because RFM summarizes purchase behavior by using a small number of variables, it is a very easily understood and effective model. RFM utilizes internal databases of customer specific transactions; because this history is not obtained through aggregate level demographic databases, the RFM analysis becomes more meaningful for targeting particular customers.

Wei et al. (2010) outline some further disadvantages of this analysis. RFM is typically focused on identifying the most valuable customers, which means that there is little meaningful scoring possible on R, F & M when most customers have spent little or do not purchase often. They suggest that this is particularly true for most firms sales because of the 80-20 rule (i.e. when 80% of sales come from 20% of their customers) although this is not usually the case with Auto dealerships.

2.6 Combining Methodologies to Produce More Accurate Profiling

Several researchers have determined that utilizing both RFM and CLV will enable us to cluster customer data more accurately: Chuang & Shen (2008) first assessed the Weights of R, F & M to determine the relative importance of each value by utilizing the Analytical Hierarchy Process Method They then evaluated the customer life time value (CLV) through clustering analysis and then finally sorted the customers using a self-organizing map methodology to identify high-value customers.

As stated above, the likelihood of a repeat purchase declines as the period of time since the last purchase grows. The chart below provides an example of how automotive dealerships currently segment their customers from an activity point of view, which I have used in various iterations over my 20 plus years in the automotive dealer management system industry. It shows what are typically considered active, non-active and lost customers. Most vehicle brands now recommend maintenance services every 6 months or 8000 KM, though some have even longer intervals or active reminders built into the vehicle software, which will inform customer when they are due for an interval service. Determining how and when to change customer marketing strategies is critical to keeping retention rates high and thus continuing the relationship with the customer over the long term.

Table 1: CRM Activity Matrix Example

Period	Status	Actions
1-3 months (R=5)	Recent & active	Send post service surveys
4-6 months (R=5)	Active	Send interval service reminder
7-9 Months (R=4)	Semi active	Send second service reminder
10-12 Months (R=3)	Semi active	Send personalized message or phone call
13 -18 Months (R=2)	Potential lost customer	Contact customer to determine status
19 – 24 Months (R=1)	Last chance to retain	Special offer for last chance retention
25 + Months (R=1)	Lost customer	Customer has defected remove from activity list

2.7 Suggestions for Optimizing RFM Analysis

As mentioned above and in Table 1, auto dealer service customers normally purchase maintenance service on an interval basis, however the services performed at each interval will fluctuate (See Table 2), therefore their spending patterns will be highly variable. If you Combine F & M together as a single variable, we can then extrapolate the average spend to make the comparison of customer revenue more representative of the cash flow created by customer activity. Below is an example of interval services performed at automotive dealerships. This helps to understand that cash flows from customers can vary widely depending on what services they select and where each vehicle is in its life cycle.

Table 2: Example of Typical Basic Maintenance Schedule

Service 1(6000K, 30000K)	Lube oil Filter (LOF)	$49.95
Service 2(12000K, 36000K)	Service1 + Tire rotation	$89.95
Service 3(18000K, 42000K)	Lube oil Filter (LOF)	$49.95
Service 4 (24000K,48000K)	Service 2 + brake service	159.95

2.8 Clustering Customer Segments

Once RFM values have been identified, we can then cluster customers with similar RFM values and assign them to an appropriate segment. This would then allow us to target customers with varying RFM values appropriately; developing and adopting different marketing strategies for each identified segment to nurture high RFM customers and find ways to increase support customer with growth potential, etc.

By clustering customers into different groups, we can improve the quality of recommendation, this helps decision-makers identify market segments more clearly and therefore develop more effective strategies (Birant, 2011).

Vergara (2009) suggested that we can obtain a clearer picture of our customer databases by clustering customers using both RFM & CLV. Specifically, he suggests that once the organization has identified customers based on purchasing patterns, segmentation analysis can be used to target the core audience you want to reach. What follows are some of the suggested steps and activities to perform to get the best results:

- Identify customer segments that enable differential marketing programs
- Use past purchase data to identify customer groups
- Identify key factors which drive customer value to enable segmentation
- Identify clusters of customers which can be used to develop marketing strategies to increase or maintain current value to the firm.
- Align the marketing budget priorities against each subgroup

Sohrabi & Khanlari (2007) suggested clustering the customer database using K-means cluster analysis. This approach attempts to identify relatively homogeneous groups of customers based on selected characteristics. The K-Means algorithm is designed to handle large numbers of records, which is perfect for dealerships as they do not purge their customer files for 10 years or more, and sometimes never, so the data continually grows. The dealership we used for this study is medium in size and their entire database includes well over 24,000 individual customer records added to the system since 1997. They suggest setting the number of clusters to 8, which is the total number of permutations of RFM if you look at each variable as a High / Low variable (2 X 2 X 2) according to the average R, F & M score. From this data, they built an 8 cluster comparative chart (see Table 8). By comparing the results in each cluster to the average they assigned the Low / High score to each RFM variable.

2.8 Summary

This review has focused on three main areas: first, why measuring customer profitability is important for automotive dealerships to understand and the advantages of analyzing their customer base from this perspective.

Now that dealership management systems have become more comprehensive, many dealerships are now able to extract the transaction data from these systems, which allows them to perform both RFM and CLV analyses over their entire customer database and not limit themselves to a sample of customers through surveys or manually extracted data from paper files. This allows them to categorize each and every customer's contribution to the firm and determine the best actions to enhance relationships, retain valuable customers and even divest unprofitable ones.

Secondly, Customer Lifetime Value (CLV) at its core is the present value of all revenues that can be derived from each of the dealerships customers over the entire period that this customer does business with them. There are a multitude of CLV models that have been developed in an attempt to adapt to various industries. By way of example, we explored some of the basic models that have been developed to adapt to various industry needs. They encompass many variables, such as average lifetime, acquisition costs, discount rates, retention/defection rates, and gross margins to name a few.

Each model of course has its drawbacks because certain assumptions need to be made. For example, some models assume that cash flows always occur at the same time in every period. This makes sense for businesses with regular cash flow, but would be less accurate for business where cash flows are more erratic. Another important factor is that in most cases in most modeling approaches, competition is ignored due to lack of competitive data (Gupta et al., 2006)

Finally, we looked at RFM analysis, which has been in use for more than 30 years and is a conceptually easy to understand method of analyzing the firm's customers. This method has proven to be effective for many industries in both B2B and B2C transactions. RFM is a very adaptable analysis tool as it can be tailored to the specific needs of any business and each variable can be altered or enhanced to suit the requirements of the analysis being performed.

As with any method of analysis, there are drawbacks as well as advantages. For example, the grouping or weighting of data, which is often based on experience of the marketer, may reduce the complexity of the analysis to the detriment of its accuracy or significance. This is called the overlapping syndrome (Yang, 2004). How each value is weighted can also lead researchers to inappropriate conclusions, as does the assumption that a particular industry can count on customer loyalty for repeat purchases. Since RFM analyses past behavior, there is no guarantee that this predicts future purchasing behavior, hence why utilizing both RFM and CLV to cluster customers into actionable segments will yield better results.

III. Methodology

3.1 Data Gathering & Cleaning

The customer data for this research paper were gathered from a single automotive retail dealer that specializes in North American brands of cars and trucks. The data was extracted and exported to Microsoft Excel utilizing a dynamic reporting tool resident on the Dealership Management System (DMS) utilized by the dealership. We did not differentiate between different vehicle types, i.e. cars, trucks and crossovers. We did, however, gather the vehicle carline when the data were available in the database.

We gathered 10941 lines of data from the selected database, limited on a period basis to records where the last repair order date was between December 1st 2009 and the day the data was retrieved on December 1st 2014. We then did a manual purge of 285 records, which were obviously business clients or employees. This provided a set of 10656 records to analyze. Customer records are repeated where the data included reference to multiple vehicles. For example, if a customer has owned four vehicles and the last repair order for one of these vehicles met our selection criteria, then the customer data is repeated four times. However, the data in each line remained the same except for the vehicle specific information. To eliminate some of the noise in the data caused by these duplicates, we elected to discard the duplicate records while retaining the one with the most recent service history. This left us with 9693 individual customer records in the final data set for analysis.

In the initial pull of data, we also tried to exclude any customers who were flagged in the database as either Employees (former or current) or Businesses, so that we could focus our attention solely on privately-owned vehicles being serviced at this dealer.

3.2 Analysis and Creation of New Data Fields

With the data set cleaned, we were able to perform three different RFM analyses, which we thought would improve the predictability of the customer's past behavior, i.e. RFM for Service (RFMsvc), RFM for Total transactions (RFMtot) and RFM for Monthly average service revenue (RFMave). To calculate these different RFM scores we first had to identify and rank the individual building blocks of R, F & M the description of each of these follows:

Recency (R) - calculated the difference, in days, between the day the data was extracted and the last repair order date

Frequency (F) - for all transactions was extracted by total all transactions for each customer record [Ftot]

Frequency (F) - for service was extracted by total all service and Truck shop transactions for each customer record. [Fsvc]

Monetary value (M) – for service was extracted by total all service and truck shop gross profit for each customer record. [Msvc]

Monetary value (M) – for all was extracted by total all gross for each customer record.

Monetary value (M) – for average service was extracted by totaling all service and truck shop transactions and dividing by the number of service transaction gross for each record.

Once all the baseline RFM values had been calculated, we then started building our scoring of the individual records for each value. Ranking followed the usual pattern for RFM where 5 is best score and 1 is lowest score, e.g., the most recent service transactions would be ranked as a 5 and the least recent would be ranked as 1.

The Recency Scoring (R) score followed the same logic that is used in Chart 1: a score of 5 is assigned for vehicles that had been serviced at the dealership in the last 6 months, score of 4 for 7 to 9 months, score of 3 for 10 – 12 months, score of 2 for 13 to 18 months and finally score of 1 for vehicles that had not been serviced in over 19 months.

The Frequency Scoring (Ftot) for total transactions was sorted and ranked to identify customers in segments of 20% of the database from the sorted list of records, going from highest to lowest number of total of transactions.

The Frequency scoring (Fsvc) for service transactions was sorted and ranked to identify customers in segments of 20% of the database from the sorted list of records, going from highest to lowest number of total service and truck shop transactions.

The Monetary scoring (Msvc) for service transactions was sorted and ranked to identify customers in segments of 20% of the database from the sorted list of records, highest to lowest number of total service and truck shop transactions.

The Monetary Scoring (Mtot) for all transactions was sorted and ranked to identify customers in segments of 20% of the database from the sorted list of records, going from highest to lowest number of total service and truck shop transactions.

The Monetary Scoring (Mave) for all service transactions was sorted and ranked to identify customers in segments of 20% of the database from the sorted list of records, going from highest to lowest number of total service and truck shop transactions. This score was developed to see if a monthly M may be more accurate predictor than an annual indicator, due to the non-cyclical nature of vehicle maintenance.

Table 3: Field Descriptions in the Data Set

Field	Variable Name	Description
1	Cust #	Customer number / record ID
2	CREATE-DATE4	Date customer record was created
3	LAST-RO-DATE	Last repair order date
4	PRIOR RO DATE	Previous repair order date
5	CARLINE	Vehicle carline
6	VEH YEAR	Vehicle year (when entered)
7	SVC-GP%-PY-C	Service gross profit percent prior years customer pay
8	SVC-GP PY W	Service gross profit dollars prior years warranty pay
9	SERV-GROSS-TOT-C	Service gross profit dollars customer pay
10	SERV-GROSS-TOT-W	Service gross profit dollars warranty pay
11	PY-SERV-GROSS-C	Service gross profit dollars customer pay - prior year
12	PY-SERV-GROSS-W	Service gross profit dollars prior year warranty pay
13	SVC-GP%-PY-C	Service gross profit percent prior year customer pay
14	SVC-GP%-PY-W	Service gross profit percent prior year warranty pay
15	SERV-UNITS-TOT-C	# of individual job performed for vehicle customer pay current year
16	SERV-UNITS-TOT-W	# of individual job performed for vehicle warranty pay current year
17	SVC-TRN-PY-C	# of individual job performed for vehicle customer pay Prior year
18	SVC-TRN-PY-W	# of individual job performed for vehicle warranty pay Prior year

Field	Variable Name	Description
19	ALL-GP-CY	All gross profit generated current
20	ALL-GP-PY	All gross profit generated prior
21	BS-GROSS-TOT-C	Truck shop gross profit customer pay
22	BS-GROSS-TOT-W	Truck shop gross profit warranty pay
23	BS-GP%-PY-C	Truck shop gross profit percent customer pay
24	BS-GP%-PY-W	Truck shop gross profit percent warranty pay
25	BS-UNITS-TOT-C	# of individual truck shop job performed for vehicle customer pay prior year
26	BS-UNITS-TOT-W	# of individual truck shop job performed for vehicle warranty pay prior year
27	ALL-TRAN-CY	All transactions current year
28	ALL-TRAN-PY	All transaction prior year
29	ALL-GROSS-CY	All gross generated in current year
30	ALL-GROSS-PY	All gross generated in current year
31	ALL-GP%	Gross profit % for all transaction
32	ALL-GP%-CY	Gross profit % for all transaction current year
33	ALL-GP%-PY	Gross profit % for all transaction current year
34	ALL-TRAN	Total of all transactions
35	ALL-TRAN-CY	All transactions current year
36	ALL-TRAN-PY	All transaction prior year
37	TOT-USED-FI-GP	Total gross profit from Used vehicle F&I sales
38	TOT-NEW-FI-GP	Total gross profit from new F&I vehicle sales
39	TOT-USED-VEH-GROSS	Total gross profit from used vehicle sales
40	Date Today	Used to calculate offset from last repair order
41	Receny Days - Raw	# of days since last Repair Order
42	Frequency all Raw	Total of all transactions (line 34)
43	Frequency all SVC	Total frequency of all Service & truck shop transactions (Sum 15, 16, 17, 18, 25, 26)
44	Monetary Raw SVC	Total $ of all Service & truck shop transactions (sum 9, 10, 11, 12, 20, 22)
45	Monetary Raw ALL	Total $ of all transactions (sum 29, 30)
46	Monetary service Ave	Average $ per service transaction (div 44 / 43)
47	R	Recency score
48	Ftot	Frequency score for all transactions
49	Fsvc	Frequency score for all service transactions
50	Msvc	Monetary score for service transactions
51	Mtot	Monetary score for all transactions
52	Mave	Monetary score for average service transactions
53	RFMsvc	RFM score for service (47,49, 50)

Field	Variable Name	Description
54	RFMtot	RFM score for all transactions (47,48, 51)
55	RFMave	RFM score for service average (47,48, 51)
56	CLVsvc: m(r/1+i-r)	CLV calculation for service transactions
57	CLV tot: m(r/1+i-r)	CLV calculation for all transactions

Once the data was imported into SPSS we also added CLV calculation, which used the average service and truck shop spend from current year and prior year to predict a CLV amount for the next three years using the formula suggested by Jain & Singh (2002).

$$\text{CLV} = \sum_{i=1}^{n} (R_i - C_i) / (1+d)^{i-0.5}$$

We also converted the carlines for the brands into 11 dummy variables, either 0 or 1, which enabled us to then perform regression analyses utilizing brands (carlines) as variables for further analysis. Furthermore, we added customer age as a variable, by calculating the difference in days between when the customer record was created and when the data was extracted.

3.3 Confidentiality and Privacy

The dealer principal of this store gave us permission to access this data for the purposes of this academic study. To ensure privacy for the dealership's customers, we specifically made sure that all data we gathered for this study was anonymized and could not be used to identify customers individually. Therefore, we used only customer numbers and vehicle data to differentiate records.

IV. Results

4.1 *Customer Lifetime Value (CLV)*

The data we studied allowed us to compare and study RFM and CLV as tools to understand the customer database more accurately. We were trying to look at trends and correlations between the data collected and how the concepts of RFM and CLV could be applied to understand the data.

4.1.1 CLV by Brand

The dataset had a large number of records (7373) which could be identified as Ford Brands (carline), this allowed us to perform a regression analysis for CLV by brand. This analysis uses an averaged CLV by brand, which yielded the following results:

Table 4: Customer Lifetime Value by Brand

Brand	CLV ($)	N	Standard Deviation
Focus	$888.54	1103	1638.05
Fiesta	$300.54	193	628.03
Fusion	$807.65	589	1512.77
F-Series	$708.52	2033	1423.33
Ranger	$576.68	1049	961.00
Taurus	$1,308.77	195	2091.43
Edge	$1,075.75	304	1938.11
Mustang	$526.49	256	859.74
Explorer	$968.28	204	1571.13
Escape	$817.47	1401	1487.83
MKX	$891.15	46	1716.10

Note: N = Sample Size; Standard Deviation is quite large which means there is big range for CLV for each brand.

4.1.2 CLV Percentile Analysis

We were also able to look at CLV for service from a percentile point of view to look at the spread of CLV values within the customer database. We found that there were some large dollar value outliers, as evidenced by the mean CLV.

Table 5: CLV Percentile Analysis

N	9633
Overall Mean	762.23
Percentiles	**Mean CLV**
0 to 10	< 0
11 to 20	< 0
21 to 30	$ 37.94
31 to 40	$ 108.64
41 to 50	$ 223.07
51 to 60	$ 379.58
61 to 70	$ 630.85
71 to 80	$ 1,081.39
81 to 90	$ 2,106.30
91 Plus	> $2106.30

It should be noted that the bottom 20% customer segment showed a negative CLV and can therefore be considered a drain on the organization; whereas the top 20% are contributing substantially to total CLV. This being said, the customers in the 21% to 80% actually contribute the majority of the firm's gross profit.

4.2 RFM

In this next section, we further explore RFM scores and what relationships these scores have with each other and other variables through several analyses. First, we look at a quartile and frequency analysis. Secondly, we look at RFM as an indicator of profitability. Thirdly, we consider R, F & M and their correlation to customer tenure and finally, we investigate a cluster analysis for RFM.

4.2.1 Quartile & Frequency Analysis

We performed a quartile analysis of the RFM score for the entire database and determined that most of the 125 possible RFM scores were found within the data set. The vast majority of these scores were between 224 and 532 (50%). From a predictive perspective these scores simply tell us that the full range of customer types is present in the customer base, through all ranges from low to high recency, frequency, and monetary value.

Table 6: RFM Quartile and Frequency Analysis

Total records (N)		9633			Max # RFM
	Percentiles	LOW	HIGH	FREQ	Groups
	1 to 25	111	123	2465	8
	26 to 50	224	531	2398	78
	51 to 75	**532**	**532**	**2339**	1
	76 +	533	555	2431	13
			Totals	9633	100

It should be noted that the 51% to 75% quartile, 2339 customers, is comprised of one single RFM score. There were a maximum of 100 RFM groups present in the data.

4.2.2 RFM Profitability

We compared the RFM for service to CLV and the monetary value for service to confirm which components of RFM would drive the CLV and customer profitability. What follows are some of the final data records of this analysis. Table 7 is divided into two groups of RFM Scores to show the RFMsvc score, the mean and standard deviation for CLV, raw service revenue and the number (#) of records which received that score. The left side shows a sampling of RFM scores from 111 to 445 and the right shows scores from 534 to 555.

Table 7: RFM Score Comparison to Monetary Value and CLV

RFM svc		CLV	Raw SVC $	#
111	Mean	10.37	7.43	975
	Std. Dev.	18.75	13.44	
112	Mean	166.75	119.52	332
	Std. Dev.	71.61	51.32	
113	Mean	493.89	353.98	128
	Std. Dev.	134.53	96.42	
114	Mean	1073.00	769.04	40
	Std. Dev.	209.42	150.10	
115	Mean	2372.10	1700.13	4
343	Mean	326.03	431.41	26
	Std. Dev.	243.27	91.15	
344	Mean	811.75	902.53	35
	Std. Dev.	519.03	196.70	
345	Mean	1814.97	1966.80	16
	Std. Dev.	1764.37	948.01	
443	Mean	216.20	384.87	36
	Std. Dev.	173.05	87.48	
444	Mean	844.76	932.97	62
	Std. Dev.	482.21	233.52	
445	Mean	2528.39	2169.97	28
	Std. Dev.	1338.95	737.00	

RFM svc		CLV	Raw SVC $	#
534	Mean	714.82	840.20	102
	Std. Dev.	533.90	194.71	
535	Mean	1948.31	1916.56	19
	Std. Dev.	1136.00	464.20	
543	Mean	215.36	411.70	313
	Std. Dev.	215.05	95.92	
544	Mean	670.04	886.00	465
	Std. Dev.	476.45	208.21	
545	Mean	1862.15	1959.42	171
	Std. Dev.	1166.44	676.85	
553	Mean	189.98	485.99	12
	Std. Dev.	265.91	72.36	
554	Mean	546.78	1023.69	273
	Std. Dev.	481.73	199.14	
555	Mean	3157.53	3145.31	686
	Std. Dev.	2888.10	2025.10	

4.2.3 Correlations of RFM & Customer Age

We also looked at the relationship between R, F & M for service and customer age, e.g., how long this customer has been coming to the dealership. You will observe that customer age is positively correlated to all other factors, and recency is negatively correlated to both frequency and monetary value.

Table 8: Correlation Analysis

	Recency Days Raw	**Frequency all SVC**	**Monetary Raw SVC**	**Customer Age**
Recency Days - Raw	1			
Frequency all SVC	-.233**	1		
Monetary Raw SVC	-.186**	.809**	1	
Customer Age	.205**	.488**	.401**	1
** indicates p<0.001				

4.3 Cluster Analysis

Based on the K-means cluster analysis suggested by Sohrabi & Khanlari (2007), we developed the following eight (8) category cluster analysis for the customer database. It is interesting to note that 91.7% of the customers fall within 2 categories LHL and HLL. Each of R, F & M were scored as either high (H) or low (L) by whether they were above or below the average score for that variable. This was to replicate the methodology applied in the study with the exception of the monetary value, which factored in the author's experience to mitigate the high average values indicated in the data.

Table 9: RFM Cluster Analysis

	1	2	3	4	5	6	7	8	AVE
Recency Days - Raw	120	155	424	458	339	763	349	278	361
Monetary Raw SVC	18821	16033	13552	3052	1264	222	5654	8610	8401
Frequency all SVC	93	82	92	29	15	4	53	66	54
Cluster Type	L,H,H	L,H,H	H,H,L	H,H,L	L,H,L	H,L,L	L,L,L	L,H,H	Total
Number of Cases	1	2	5	619	2196	6637	135	38	9633
Percentile Share	0.01 %	0.02 %	0.05 %	6.43 %	22.80 %	68.90 %	1.40 %	0.39 %	100.00 %

4.4 *Regression Analysis*

Two regression analyses were performed. The first one used CLV as the dependent to understand how brands could predict CLV. The second used a measure of profit as the dependent variable, alongside demographics (gender and lifespan as a customer) and a few other independent variables. This second regression was done using a different dataset to the one utilized for all other analyses in this project.

4.4.1 Regression Analysis - CLV by Brand

A regression analysis was conducted with CLV as the dependent variable and the following independent variables: Focus, Fiesta, Fusion, Ranger, Taurus, Edge, Mustang, Explorer, Escape, MKX, Recency days, Frequency all SVC and customer age. The R-square was: 0.525. The overall model was statistically significant [$F(13,7372)=624.70$, $p<0.0001$]. See Table 10 for regression coefficients.

Table 10: Regression Analysis - CLV by Brands

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-127.912	27.720		-4.614	.000
Focus	-.863	37.621	.000	-.023	.982
Fiesta	-215.572	75.923	-.024	-2.839	.005
Fusion	-59.689	47.022	-.011	-1.269	.204
Ranger	-126.058	38.583	-.030	-3.267	.001
Taurus	290.809	75.487	.032	3.852	.000
Edge	86.387	61.858	.012	1.397	.163
Mustang	4.315	66.809	.001	.065	.949
Explorer	157.606	73.777	.018	2.136	.033
Escape	-24.563	34.987	-.007	-.702	.483
MKX	-523.704	149.832	-.028	-3.495	.000
Recency Days - Raw	.102	.024	.039	4.286	.000
Frequency all SVC	74.669	1.111	.680	67.203	.000
Customer Age	.072	.009	.084	8.372	.000

4.4.2 Regression Analysis - Profit vs. Miscellaneous Variables

A regression analysis was conducted on a previous iteration of the dataset, which had included gender and overall profit. The analysis was performed with profit as the dependent variable together with the following independent variables: gender, transactions, recency days, years, new vehicle gross current year and new vehicle gross prior year. The R-square was: 0.152. The overall model was statistically significant [F(6,5583)=166.04, p<0.0001]. See Table 11 for regression coefficients.

Table 11: Regression Analysis - Profit

Model	Un standardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	3053.015	170.921		17.862	.000
Gender	-29.367	89.467	-.004	-.328	.743
Transactions	2.344	.966	.034	2.427	.015
Recency Days	-1.997	.126	-.200	-15.858	.000
Years	-179.524	10.317	-.236	-17.401	.000
NV Current	1382.678	85.389	.201	16.193	.000
NV Prior	26.917	15.747	.022	1.709	.087

4.4.3 Vehicle Brand Anova Analysis

A one way Anova analysis was performed comparing each of the 11 Ford brands identified in the database. The mean differences were compared using a Tukey test, which revealed that many of the means showed significant differences. Two examples follow.

Table 12: Edge - CLV Comparison

(I) Brand	(J) Brand	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
	Focus	187.20	93.60	0.649	-114.15	488.55
	Fiesta	775.21	132.99	**0**	347.03	1203.38
	Fusion	268.09	102.04	0.235	-60.45	596.63
	F-Series	367.22	88.85	**0.002**	81.14	653.29
	Ranger	499.07	94.12	**0**	196.03	802.09
	Taurus	-233.02	132.57	0.805	-659.85	193.80
	Mustang	549.25	122.57	**0**	154.61	943.86
	Explorer	107.47	130.78	0.999	-313.58	528.52
	Escape	258.28	91.42	**0.149**	-36.07	552.63
Edge	MKX	184.59	228.59	0.999	-551.40	920.59

Table 13: Explorer - CLV Comparison

(I) Brand	(J) Brand	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
	Focus	79.73	110.12	1	-274.83283	434.29388
	Fiesta	667.73	145.09	**0**	200.58	1134.89
	Fusion	160.62	117.38	0.956	-217.31	538.56
	F-Series	259.75	106.12	0.336	-81.91	601.42
	Ranger	391.59	110.57	**0.017**	35.61	747.58
	Taurus	-340.49	144.71	0.398	-806.41	125.42
	Edge	-107.47	130.78	0.999	-528.52	313.58
	Mustang	441.78	135.61	**0.044**	5.16	878.40
	Escape	150.81	108.28	**0.95**	-197.81	499.43
Explorer	MKX	77.12	235.84	1	-682.21	836.46

V. Conclusion

In closing, this section will evaluate our initial research questions, new information acquired during the course of the analysis, and applicability of the concepts for research and management. This will be followed with a discussion of the limitations of this research project and suggested future research that would address some of these limitations.

5.1 Discussion of Initial Research Questions

In this first part we will address the research questions that were posited at the outset of this study and address the extent to which the results answer these questions.

Q1- Can we identify drivers of profitability?

Overall the answer is yes. The drivers of profitability we were able to identify in this study included the brand of vehicle, as well as Recency, Frequency and Monetary value. Obviously, as expected, the largest driver of profitability was monetary value, but all the others were also proven to have had an effect.

Q2- Can customer database be segmented to identify which customers are the most profitable?

RFM is an effective way to classify customers and segment the database for analysis and market research use. Dealership management can use these scores to identify their most profitable customers and those who can be groomed to become more profitable. Cluster analysis can be used to further refine customer profiles, but cleaner data would be required to prove this more accurately. As evidenced by the results in Table 9, the vast majority of the customers are relatively new customers with high recency and low frequency and monetary value; or are likely to be defectors who have low recency and monetary value, but a high frequency.

Q3- Can RFM segmentation be used to segment and predict profitability?

RFM seems to be so strongly correlated to monetary value and CLV in this study that it makes clear that more work is needed to identify a methodology which accurately predicts profitability. As the reader can see from the sample RFM scores, CLV dollars and raw service dollars, whenever an RFM score ends in a “5” the dollar values will be high. It should also be noted that there is a large cluster of customers who fall into the 543 and 544 category, which together account for a significant contribution to the bottom line. RFM definitely identifies past trends but more variables including vehicle age and vehicle brand service plans costs, need to be factored in the model to make it more effective. This was beyond the scope and availability of data for this study.

Q4- Can we model Projected CLV from the RFM profile?

In this study, we were able to validate that RFM scores and CLV are interrelated, but more work needs to be done to study this area. The auto dealership business is very complex and many more factors need to be worked on in order to build a predictive model.

5.2 Discussion of Results

If we look back at the percentile breakdown of CLV, the top 30% are the most profitable customers. If you consider that the very top CLV customers are outliers and exceptions then the 80 / 20 rule still applies and the top 20% of customers are the most profitable individually. However, Dealership marketers should not be fooled into thinking they should concentrate their efforts on these customers alone. Collectively the mid-range customers are the ones that supply the vast majority of the revenue stream to the dealership.

The Reader should also note that there is a large variation in the CLV to Brand regression (Table 9), at the same time the standard deviation is also large and therefore, there is a lot of variation around the mean. This does indicate that there is some correlation between CLV and the Brand, but from the data acquired it cannot be quantified accurately.

When comparing the mean CLV scores from brand to brand several of these means were significantly different, whether negative or positive. The implication is that brands have a correlation to service gross profit. This can be explained somewhat by the fact that both service intervals and pricing varies significantly from brand to brand, so the amount of time and the cost of parts vary depending on brand as well.

5.3 Managerial Implications

There is no doubt that RFM can be extracted from the basic data this research paper used and that this level of analysis can be performed by dealership personnel. The marketer though would need to temper the absolute results of a score with their experience and the actual service history records for the vehicle(s) owned by the customer (see limitations section).

The CLV figures that we have calculated in this study are beyond the scope of the average dealership marketer from both a technical and data acquisition point of view. Although the DMS system they use does provide total lifetime revenue for the customer, these totals are for every transaction and vehicle they have owned from the very first day the customer was added to the system.

5.4 Limitations of Data and Research

The results of this study should be taken with some note of caution as there was a lot of noise in the data due to multiple vehicles being owned by the same customer. Much more data would be required to calculate the true CLV properly. For example, we predicted only three years of CLV because we could not identify where each vehicle was in its life cycle. We also only performed CLV calculations on customers added after November 2011, so that we would capture mostly new vehicle purchasers and new service clients who would have been in the beginning or middle of their vehicles' life cycle.

In calculating RFM, we had to use the transaction count as a proxy for frequency. So, transactions equated to the number of individual service operations and not to the actual number of repair orders, which would have been a true reflection of frequency. It is not uncommon for one repair order to have two or three and even more operations performed in one service visit. There is also the fact that service intervals can be erratic, since most are based on kilometers driven, one customer may have services performed at a very different time interval than the next.

The data were gathered from one dealership selling vehicles from one original equipment manufacturer (OEM). The implication of this is that the study was not cross sectional from an OEM point of view, and we could easily have found very different results in an import or luxury brand dealership.

To predict future behavior from the data, we would have needed information about the customers before and after vehicles were purchased. This would have necessitated more of a longitudinal study, which was well beyond the scope and resources of this research project.

5.4 Further Research

An important area for further research would be to develop a more accurate CLV model to incorporate the life cycle of a vehicle going forward. This model would need to be able to identify: 1) where the vehicle is in the life cycle, e.g. 3rd year of an average 7 year cycle, 2) what carline to determine revenue stream and 3) also account for some growth in the revenue.

More research should also be done in identifying practical uses of the RFM score for the management team to use at the dealership level. This could be done by further studying potential clustering methodologies, which would break down the customer base into smaller homogeneous groups that could be acted upon, i.e. segments of customers who are similar enough in profile to be targeted communicated with in the same way or would behave similarly to a service offering or special promotion.

Finally, we suggest that a great deal more could be learned from conducting a survey of customers at the dealership and identifying specific attitudes and correlating these to the RFM score. Some of research questions that come to mind are intent to repurchase, satisfaction of service department, and brand satisfaction.

Very little work has been done in studying RFM, and CLV at the automotive retailer level. The data are stored on many of the Dealer Management Systems (DMS) available in the marketplace As yet the author knows of no tools or methods that have been developed to easily retrieve the data in a useful way, though he intends to continue looking for new and better ways to find and manipulate the data.

VI. References

Anonymous, (2005). *Can customer profitability be accurately measured?* .Business Credit, 107(10), 20-21.

Asllani, A., & Halstead, D. (2011). *Using RFM data to optimize direct marketing campaigns: A linear programming approach*. Academy of Marketing Studies Journal*, 15*, 59-76.

Berger, P. D., Bolton, R. N., Bowman, D., Briggs, E., & et al. (2002). *Marketing actions and the value of customer assets: A framework for customer asset management*. Journal of Service Research: *JSR, 5*(1), 39.

Birant, Derya (2011). *Data Mining using RFM Analysis knowledge-Oriented applications in data Mining*, 91-108

Blattberg, R.C. & Deighton, J. (1996). *Manage Marketing by the Customer Equity Test*, Harvard Business Review, July – August 136-144

Bonicalzi, M. (2004). Analysis *of customer relationship profitability in a car dealership*. (Order No. 3136557, Argosy University/Sarasota). ProQuest Dissertations and Theses, 195-195 p.

Calculating value, (2003). *Business Mexico, 13*(9), 52-55.

Chuang H.M & Shen C.C, *A study on the application of data mining techniques to enhance customer life- time value based on the department store industry*, Proceedings of 7th International Conference on Machine Learning and Cybernetics, California, USA, 2008, pp. 168-173.

Drèze, X., & Bonfrer, A. (2009). *Moving from customer lifetime value to customer equity*. Quantitative Marketing and Economics, *7*(3), 289-320.

Dwyer R.F (1997). *Customer Lifetime valuation to support Marketing Decision Making*. Journal of Direct Marketing*, 11* (4), 6 - 13

Fader, P. S., Hardie, B. G. S., & Lee, K. L. (2005*). RFM and CLV: Using iso-value curves for customer base analysis*. *JMR,* Journal of Marketing Research*, 42*(4), 415-430.

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., & et al. (2006*). Modeling customer lifetime value*. Journal of Service Research *: JSR, 9*(2), 139-155.

Ishiguro, H., & Amasaka, K. (2012*). Establishment of A strategic total direct mail model to bring customers into auto dealerships*. Journal of Business & Economics *Research (Online), 10*(8), 493-n/a.

Jain, D., & Singh, S. S. (2002). *Customer lifetime value research in marketing: A review and future directions*. Journal of Interactive Marketing, 16(2), 34-46.

Jain, D., & Singh, S. S. (2013). *Measuring Customer Lifetime Value: Models and Analysis*. INSEAD, Faculty & Research Paper.

Kumar, V., & Rajan, B. (2009). *Profitable customer management: Measuring and maximizing customer lifetime value*. Management Accounting Quarterly*, 10*(3), 1-18.

Kumar, V., Shah, D., & Venkatesan, R. (2006). *Managing retailer profitability--one customer at a time!* Journal of Retailing*, 82*(4), 277-294.

Kumar, V., Venkatesan, R., Bohling, T., & Beckmann, D. (2008). *The power of CLV: Managing customer lifetime value at IBM*. Marketing Science, *27*(4), 585-599,749,751,753.

Kumar V. & Bharath R.(2009 September) Strategic finance – Nurturing the right customers

Li, P., & Mee, W. (2014). *CANADA'S NEW ANTI-SPAM LAW*. AgriMarketing, 52(5), 32.

Mccrary, M. (2009*). Enhanced customer targeting with multi-stage models: Predicting customer sales and profit in the retail industry*. Journal of Targeting, Measurement and Analysis for Marketing, *17*(4), 273-295.

Moffat, Mike, Murray Kyle B., (2008*). Conroy's Acura: customer Lifetime Value and return on Marketing,* Richard Ivey school of Business cases, Case 9B08A001 Version (A) 2008-02-04

Murby, L. (2007). *Customer Profitability*. Financial Management, , 33-35.

Neal, W. D. (2004) *'Principles of marketing segmentation'*, Marketing Research, April.

Neslin, S. A., Taylor, G. A., Grantham, K. D., & Mcneil, K. R. (2013). *Overcoming the "recency trap" in customer relationship management*. Academy of Marketing Science.Journal*, 41*(3), 320-337.

Niraj, R., Foster, G., Gupta, M. R., & Narasimhan, C. (2008). *Understanding customer level profitability implications of satisfaction programs.* The Journal of Business & Industrial Marketing, 23(7), 454-463.

Pepe, M. (2012). *Customer lifetime value: A vital Marketing/Financial concept for businesses*. *Journal* of Business & Economics Research (Online)*, 10*(1), 1.

Roberts, M. L., Liu, R. R., & Hazard, K. (2005). *Strategy, technology and organisational alignment: Key components of CRM success*. Journal of Database Marketing & Customer Strategy Management, *12*(4), 315-326.

Schweidel, D. A., Fader, P., & Meyer, R. J. (2012). *What matters in managerial judgments of customer value?*. Rochester, Rochester: Social Science Research Network.

Sohrabi, Babak & Khanlari, Amir (2007). Customer *lifetime Value (CLV) Measurement based on RFM Model*. Iranian Accouting & Auditing review, spring 2007, vol. 14 No. 47, pp 7 - 20

Talaba, M., PhD. (2013). *Comparison between customer lifetime value (clv) and traditional measurement tools of customer value*. Knowledge Horizons.Economics*,* *5*(2), 189-192.

Vergara, D. (2009). *Get a little closer.* Target Marketing*, 32*(5), 30-31.

Yang, A. X. (2004*). How to develop new approaches to RFM segmentation*. Journal of Targeting, Measurement and Analysis for Marketing, 13(1), 50-60.

Wei, Jo-Ting, Lin, Shih-Yen and Wu Hsin-Hung (2010*). A Review of the Application of RFM Model*. African Journal of Business Management, Vol. 4(19), pp 4190-4206, December Special Review, 2010

Zhang, J. Q., Dixit, A., & Friedmann, R. (2010). Customer loyalty and lifetime value: An empirical investigation of consumer packaged goods. *Journal of Marketing Theory and Practice, 18*(2), 127-139.

Yang, A. X. (2010*). How to develop new approaches to RFM segmentation*. Journal of Targeting, Measurement and Analysis for Marketing, 13(1), 50-60.