IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Strategy to Increase Lifetime Value of a Customer using RFM

Author and Co-author

Ms. Shivangi Kiran Desai[1], Prof. Dr. Kavita. Kalyandurgmath[2]

Author: Ms. Shivangi Kiran Desai *LinkedIn*

Analytics professional with more than 2.5 years of experience in Marketing effectiveness measurement using real-time data. Passionate about converting data into a story. Experienced in end-to-end execution and project management for Digital Behaviour Analytics and Cross-Media Analytics.

Academic Qualifications: Full-time Post-Graduate Diploma in Management (PGDM) majoring in Research & Business Analytics from Welingkar College, Mumbai, India in 2018.

Currently a full-time student pursuing Master of Science in Business Analytics at the University of Washington, United States, started June 2021.

Co-author: Prof. Dr. Kavita. Kalyandurgmath *LinkedIn*

Professor in Operations and Analytics with over 22 years of experience in Teaching, Research, Consulting and Education Management. My Academic qualifications, Work Experience coupled with Teaching philosophy and passion in Teaching enabling me to impart Quality Education to graduate, PG Students, MDPs for over two decades. Recognition of my work through Excellent feedback by Students, Best Faculty and Outstanding Achiever's Awards, have been source of motivation for me.

Academic Qualifications: MSc (Statistics Gold medalist), MBA(Marketing).PG DCA, Diploma in Entrepreneurship, PhD in Management.

Subjects Taught: Quantitative Techniques in Management, Business Statistics, Operations Research, Research Methodology, Market Research, Decision Support Systems and Business Analytics for full-time MBA, MMS, PGDM, MDPs, Executive MBA, BBA, BCA

PhD. Guide to various Indian and International universities: successfully completed 5, ongoing 6

Education Management: Course Planning, Course Design, Academic delivery, Mentoring, Process owner of Assessment Centre/ University interface, Industry visits, Internships, Placements.

Consulting: Analytics Assignments undertaken for Many industries to name a few Honeywell Automation, Reliable Autotech.

Community Service: Member of Rotary International, Trustee Jag Kalyan, and Gyanam Ganga Trustees, volunteered during covid to help the society in many ways.

Abstract— *In this day and age, it is extremely easy and convenient to get data in most of the industries. In all business transactions, there is some basic information which is crucial and maintained by all the companies. Such data should be correctly stored and then it can be used to analyze, understand, interpret, and boost the profit margin of the company. Recency, Frequency and Monetary model (also called as RFM), has been widely applied in many practical fields in the past, especially, it helps to segment all customers and detect extremely valuable ones for direct marketing. This is done by studying their purchasing behavior i.e., RFM: Recency, Frequency, and Monetary segmentation: how recent was his or her last purchase, how frequently he or she purchased from the company, and the average or total amount of their purchase. This data helps in calculating RFM score which is later classified into segments. The decision makers can efficiently identify their valuable customers and then develop effective marketing strategy for each respective segment hereby to increase customer's lifetime value (CLV). It looks out for their best customers, churning customers or those have the possibility to be their profitable customers in near future. RFM helps in identifying the customers need most attention and the one who need to be retained. Further RFM would help in locating their loyal customers and the customers who respond to the company's promotional campaigns. This paper depicts the understanding and the scoring scheme of RFM thus summarizing how it has been effectively applied in a wide variety of areas.*

Key words: RFM, CLV, purchasing behavior, Recency Frequency and Monetary value, RFM score, Customer Lifetime value, segmentation, customer segmentation, Consumer Attitudes, Consumer Engagement, Recency, Frequency, Behavior, Monetary Value, Consumer Segmentation, Consumer Loyalty, Behavior Assessment

I. INTRODUCTION

Retail business is undergoing fast change in its marketing methods. Till a couple of years prior, we purchased the majority of the everyday use items from little shops around here and there. Generally, the shopkeepers sell goods—either one as a sole proprietor or with the help of a few assistants. Then, the concept of large departmental stores and malls provide same goods at the same or lower price. Today, supermarkets, departmental stores, online shopping is retailing like multilevel marketing that have replaced traditional retail businesspersons, such as hawkers and vendors. With such update in technology and ease in getting day- to day items, gathering data and maintaining a database has become convenient. This study enhances about the relationship between the customers and their consumption patterns in any specific store. RFM analysis is a marketing technique which helps us analyze their behaviour like when was the last date when the customer made any purchase or consumption (recency), how frequently does a customer purchase from a shop/store (frequency) and what is the monetary value of the goods/services of a customer purchase/s? There are wide range of shoppers, and this information helps us to segment every shopper by dividing them into various preformed groups. Thus, identifying the customers most likely to respond for promotions. Also, their transaction details and the number of years of their relationship with a retail sector has helped me achieve lifetime value of the customers, their trends, and a look at the possible measures.

This concept was originally introduced by Bult and Wansbeek in 1995. It was used effectively by catalogue marketers to minimize their printing and shipping costs while maximizing returns.

RFM is a scientifically proven process. It is based on the Pareto Principle: 80% of the results come from 20% of the causes. Similarly, 20% of customers contribute to 80% of your total revenue. People who spent once are more likely to spend again. People who make big-ticket purchases are more likely to repeat them. Pareto Principle is at the core of the RFM model. Focusing your efforts on critical segments of customers is likely to give you a much higher return on investment.

Rising popularity of computerization made it even easier to perform RFM studies because customer details and purchase records were digitized. An extensive study by Blattberg et al. in 2008 proved RFM's effectiveness when applied to marketing databases. Large retailers like Windsor circle reported significant success using RFM for his or her retail customers:

1. Eastwood increased their email marketing profits by 21%.
2. L'Occitane saw 25 times more revenue per email. 25 times, not 25%
3. Frederick's of Hollywood recorded conversion rates as high as 6-9% in their campaigns

RFM counts recency, frequency, and monetary values for every customer. Joins them, and afterward bunches them into various client sections for simple review and mission focusing on. RFM examination is valuable in understanding responsiveness of your clients and for division driven data set. Regardless of other development measurable strategies accessible, study shows that RFM is the second most basic strategy utilized by direct advertisers, after cross organizations. Regardless of that, RFM is famous for a couple of related reasons. As Kahan (1998) notes, RFM is easy to use and can generally be implemented very quickly.

In recent years, data mining applications based on RFM concepts have also been proposed for different areas such as for the computer security (Kim et al., 2010), for automobile industry (Chan, 2008) and for the electronics industry (Chiu et al., 2009). Research cases of data mining with RFM variables include different data mining techniques such as neural network and decision tree (Olson et al., 2009), rough set theory (Cheng & Chen, 2009), CHAID (McCarty and Hastak, 2007), genetic algorithm (Chan, 2008) and sequential pattern mining (Chen et al., 2009; Liu et al., 2009).

II. LITERATURE REVIEW

Numerous companies are currently operating in a competitive environment, which starts from shortening product life cycles, and hence decreasing customer brand loyalty (Cooil, Keiningham, Aksoy, & Hsu, 2007). To tighten the relationship that exists with a customer, many companies increasingly turn to the concepts of Customer Relationship Management (CRM) (Reinartz & Kumar, 2002; Winer, 2001) and, more specifically, database marketing. This way, companies determine current potential customers that can be converted to be the most profitable in upcoming days, i.e., the highest CLV customers. Then, build loyalty with those valuable customers and to increase their satisfaction. Therefore, it is not advisable to spend resources on unprofitable customers just to boost their levels of satisfaction.

According to the Vilfredo Pareto Principle or the eighty – twenty rules, in most companies, 20% of the shoppers generally account for the 80% or additional of firm's total value. Not every and each client square measure created equal. In fact, the highest 1% of consumers square measure value up to eighteen times over average customers.

II.1. Action plan for every CLV cohort

Due to the sudden rise in organized sector in India, majority of retail companies face difficulty in understanding customer the way they transact. The efforts of past researchers to name a few Professors V. Kumar & Prof. Bharat have showed the companies how to measure and predict the Customer Lifetime Value (CLV). There are several authors who have worked on prediction of Customer Life time Value. Also, recently the researchers introduced the RFM model, still lack a practical framework and combination of both models CLV and RFM. This study proposes step-by-step approach to measure and manage their customers. Most of the companies bring in consultants to understand their customers. Researchers in this study demonstrated calculation of RFM and CLV cohort scorecard and can be put to use for the practical tools that can support decision making

II.1.1. Managing high CLV customers

High CLV customers are every firm's pride and competitor's envy. These customers usually generate lopsidedly high profitability for the firm as compared to customers within the medium and low CLV segments. However, the high CLV customers could vary on the size of unused wallet dimension. Segment represents high CLV customers with low size of unused wallet. Customers in these segments typically represent true loyalists preferring to spend most of their entire share of wallet with the firm. The rewards are often tangible or intangible and will be directed at cultivating the attitudinal loyalty of these customers (Kumar and Shah 2004). For instance, Neiman Marcus's in Circle program rewards its best customers by offering a private six-day European golf tournament with 15 guests traveling in a private luxury jet. The objective is to the touch upon higher level goals and attitudinal aspects of the purchasers which will not be ordinarily met through tangible rewards. Customers in these segments represent extremely high net worth individuals that bear the potential of providing even greater revenue to the firm. Customers in these segments represent extremely high net worth individuals that bear the potential of providing even greater revenue to the firm.

II.1.2. Managing medium CLV customer

The medium CLV portion commonly is included the chief significant number of shoppers. Firms had the opportunity to break down this gigantic client portion cautiously to decide if there's any extension for increasing their CLV through up-sell and additionally strategically pitch advertising drives. Therefore, the firm should offer prizes coordinated at keeping up the client's present degree of paying. Promoting drives pointed toward relationship building or development of attitudinal devotion ought to be restricted to a small portion of CLV.

Thus, firm should focus their marketing initiatives towards buying more products from the firm or spending more on the prevailing products. Cox Communications, an American company that provides digital cable television, telecommunications and Home Automation services in the US offers its customers a steep discount if they buy all three services namely-. high-speed Internet, digital cable, and digital telephone line. Such marketing initiatives directly induce cross-sell as well as up-sell. If the purchasers don't respond favorably to the firm's effort of up-sell and cross-sell, then the firm should reduce its cost on such customers and instead focus in deriving profits from every transaction of these customers with the firm.

II.1.3. Managing low CLV customers:

Firms typically view these customers as a drain on the company's resources. Companies like Sprint Nextel view these customers as misfits and like to fireside them (as discussed within the earlier section of the paper). From the business standpoint, the action of Sprint Nextel might make commercial sense. Be that as it may, we aren't sure whether Sprint caused the choice to inclination to deter its clients upheld the previous worth or the more drawn-out term worth of its customers. Even if the selection was taken by watching the expected the longer-term value of the customer, Sprint may have omitted on evaluating the CLV (or the longer-term value of the customer) along the customer's overall potential to supply more business in future. Firms should aggressively roll out up-sell and cross-sell marketing initiatives to customers during this segment. Losing a customer could likewise be moderately more productive than over-spending on advertising programs for supporters during this fragment. By the by, the choice of Sprint to fireside its unbeneficial clients were an extraordinary measure and should welcome the fury of different clients or likely new clients through terrible informal exchange or negative exposure inside the press. A more practical methodology is diminishing transaction cost.

In the recent years, data mining has not only a great popularity in research area but also in commercialization. Data mining can help organizations discovering meaningful trends, patterns and correlations in their customer, product, or data, to drive improved customer relationships and then decrease the risk of business operations (Witten & Frank, 2005) In addition, firms might want to supply lavish rewards to the present segment of high CLV customers which will be financially infeasible for competition to emulate. For instance, American Express Card selectively chooses customers from its dataset and stows exclusive privileges such as 24-hour trip advisors, personal shopping assistants and complimentary stays at 5-star hotels (Kumar and Shah 2004) A customer having a high score of recency implies that he or she is more likely to make a repeat buy. The top 20% segment is coded as 5, while the next 20% segment is coded as 4 and so forth. Finally, the recency for every customer in the database is denoted by a number from 5 to 1 (Hughes, 1996; Kahan, 1998; Tsai and Chiu, 2004).

RFM (recency, frequency and monetary) model is a behavior-based model used to analyze the behavior of a customer and then make predictions based on the actual transactions made by them. According to Brand and Gerritsen (1998), Safe ways, like other large retail chains is an information. The supermarket gets demographic data directly from its customers by offering discounts in savings club card. To obtain the card, shoppers disclose their personal information which is later used in predictive modelling. An important discipline within database marketing is customer retention management, or the prevention of customer churn, defined as the propensity of customers to end the relationships with the company, and to switch to other suppliers. Several authors report the close link between customer retention and firm profitability (Gupta, Lehmann, & Stuart, 2004; Larivière & Van den Poel, 2005). Moreover, it is generally accepted that prolonging relationships with existing customers generates a higher return on investment than attracting new clients (Mozer, Wolniewicz, Grimes, Johnson, & Kaushansky, 2000; Rust & Zahorik, 1993).

A well-documented approach to enhance customer retention is the practice of customer churn prediction, in which a classification model is built to identify those customers that are most likely to demonstrate churning behavior (Xie, Li, Ngai, & Ying, 2009). With this simplified approach to Customer Lifetime Value, one can easily be able to take a snapshot of their customers' purchasing history and flip it into a widescreen forecast of their future actions.

According to Letitiah Obiri, a marketing consultant, in an article on encouraging customer loyally post festive season, "Customer Lifetime Value is a clear look at the benefit of acquiring and keeping any given customer."

one of the primaries uses for CLV is that Cost Per Acquisition will be as low as possible.

III. RESEARCH METHODOLOGY

Research design is very well crafted to meet the study objectives. The fundamental objective of this study is to predict lifetime value of a customer based on the type of transactions they do for retail business. This study is designed to help the companies to segment their customers based on their transactions with their date of transaction, the monetary value of the transaction and how frequently they visit the retail store. Secondary dataset is adopted from Kaggle. The study unit is individual customer, and the data set has 9994 transaction details for 793 customers. The sample data set is covering one of the USA's superstore as real data of large Indian retail companies is not quite easy to obtain as secondary data. The research is intended to be completed in various stages as follows:

Acquiring secondary dataset of US superstore from 2014-2018 from Kaggle for the study.

1. Taking literature summary of peer reviewed international journals on RFM and CLV models.
2. Using MS Excel for data cleaning, tabulating and present the analysis for practical purposes.
3. Demonstrated RFM analysis on collected data, assigned scores to each customer by studying their transaction over the period of time and tabulate that data in excel charts
4. Created clusters and further segmented RFM segments in high, medium, and low CLV
5. Drawing conclusions and giving suggestions for each customer segment

The detailed methodology of analysis is explained as below:

III.1. Customer Segments with RFM Model

There is no rule of the number of segments to be used for RFM analysis. Here, we have used eleven customer segments. In the data, recency signifies the length of a time period since the last order, while frequency denotes the number of orders within a specified time period which can be one year, 2 years or from the company incorporation. Monetary value means the total amount of money that the customer has spent in this specified time period (Wang, 2010). In this case, it is very important to take the same analysis period for all three variables: Recency, frequency and monetary value. RFM basically studies the purchase behaviour which is used in segmenting. Customers attitude toward any product, brand, benefit, or even loyalty from the database is observed. Evaluate the actual percentage of the existing customers would be in every of these segments. And evaluate the effectiveness of the recommended marketing action used in the for business.

III.2. RFM Score Calculations

Following step is to organize every customer detail from the database namely: Customer ID / Email / Name etc. – to identify them. Recency (R) as days since last purchase: Exact how long prior was their last purchase? Deduct latest purchase date from the present date to ascertain the recency esteem. 1 day prior? 14 days prior? 500 days prior?

Frequency (F) as complete number of transactions: what number occasions has the client bought from our store? For instance, in the event that a person ordered 10 times over a time then their frequency number is 10. Monetary (M) as complete cash spent: What is the aggregate sum this client spent? Once more, breaking point to most recent two years – or take all time. Simply total up the total amount to get the M worth. If the amount is extremely high for you to use, then you can also take the average of the total amount.

After this step, we will assign scores from 1 to 5 to each customer for each factor on a 1 to 5 scale (5 is highest).

In Recency, maximum score 5 will be given to recent customers (or customers with the smallest recent date) while minimum score 1 will be given to customers with the highest number of days. This is because those who have purchased recently are likely to revisit the store again in future as compared to those who purchased 500 days ago.

Scoring approach for frequency and monetary value is the opposite of recency. Here, higher the transactions and total money spent, higher is the score. Lower the amount and frequency, lower the score.

III.3. Why is RFM used as a method to analyze Customer Lifetime Value?

Customer lifetime value (CLV) is a metric that represents the total net profit that a company makes from any given customer. CLV is a projection to estimate a customer's monetary worth to a business after factoring in the value of the relationship with a customer over time. Out of all the assets that an organization has, customers are the key driver of future revenues. For an organization, customers are the key drivers who will generate revenue in future. Hence, it is important do make decisions keeping customer in mind at the center.

III.4. Customer Analytics:

It is about analyzing customer data and observe their behavior to manage business decisions. Using this metric-based approach will help to manage company’s marketing retention and to manage customer relationship. The analytics is worked to achieve following objectives.

1. **Customer segmentation**

    Segmentation into small groups and addressing customers individually based on actual behaviours instead of only looking at average data customers.

2. **Tracking customers**

    Tracking their move among different segments over time, instead of just determining in what segments customers are now without regard for how they arrived there

3. **Prediction**

    Accurately predicting the customers future behaviours of customers (e.g., convert, churn, spend more, spend less) using predictive customer behaviour modelling techniques

4. **Determine customer lifetime value**

    Using advanced calculations to determine the customer lifetime value (LTV) of every customer and basing decisions on it.

CLV is a projection to estimate a customer's monetary worth to a business and this is done only after factoring in the value of the relationship with a customer from the beginning to the current time. CLV is an important metric for determining how much money a company wants to spend on acquiring new clients and how much repeat business a company can expect from certain consumers.

Here, we are just segregating the customers in 3 groups to increase their lifetime value, hence, calculation is not required. If a company needs to calculate their customer lifetime value, then it can be calculated using any one of the methods illustrated below:

- **Average revenue per user**
    - Customer Lifetime Value = Average Purchase Value Per Customer * Gross Margin * Average Number of Purchases OR
    - Customer Lifetime Value = (Annual revenue per customer * Customer relationship in years) – Customer acquisition cost
- **Cohort analysis:** A cohort is a group of customers that share a characteristic or set of characteristics. By examining cohorts instead of individual users, companies can get a picture of the variations that exist over the course of an entire relationship with groups of customers.

- **Individualized CLV:** Companies interested in broadly calculating CLV can focus on determining the total value of customers by source, channel, campaign, or other mediums such as coupons or landing pages on a company website.

IV. DATA ANALYSIS

IV.1. Data Analysis-1

Researchers imported a relevant dataset from Kaggle. This data is about superstore of USA supermarket data, and it contains 9994 transaction details of 794 clients. With this detail of customer's first buy, last buy, total amount spent by every customer along with their frequency of buy. I have used RFM method to segment customers according to their buy history as in how often they buy, when do they buy and how much have they spent. used 5 pointer scales to segment them and have segmented them into:

- Champions
  - These are those customers who buy very often, who spends the most and who purchased most recently.
- Loyal Customers
  - These customers respond to promotions offered by the seller and those who spend more
- Potential Loyalist
  - These are customers who have shopped recently, more than once and spend a higher amount
- Recent Customers
  - These are recent shoppers but don't purchase often
- Promising
  - These are customers who have shopped recently but have spent less
- Customers Needing Attention
  - These are customers whose recency, frequency and monetary value is a bit above average
- About to Sleep
  - These are customers whose recency, frequency and monetary value is a bit below average

- At Risk
    - These customers are those who used to spend high and buy often in the past. As they have not purchased in the recent times, it is important to bring them back
- Can't Lose Them
    - These are customers who made biggest buy but that was a while ago
- Hibernating
    - These customers Last order was long back, they spend less and most probably purchased only once
- Lost
    - These are customers with the lowest recency, frequency, and monetary scores

Table 1 and 2 are first 10 rows of the dataset. We can sort the data as per the Order Date but as we only require frequency, recent purchase and total amount, sorting is not required. I created a pivot table and that helped me get the right data required for the analysis.

Table 1: First 11 columns of the dataset

Table 2: Remaining columns of the dataset

Every type of customer is different, and it is important to treat them differently. As everyone is not your target customer and it is necessary to understand and implement marketing efforts for different customers to retain them. Some customers are deal seekers, some are loyalists, and some are waited to be treated differently. Understanding demands of every segment of customers will help to retain the existing ones and acquire old and lost customers. This is how customer lifetime value can be increased.

Now, we have a lot of rows, but all the rows are not necessary for current analysis. I have taken a copy of the dataset and in that copy, I just kept required rows: Customer ID, Order Date and Sales. I created a pivot table by keeping Customer ID in row. On the column side, I placed minimum and maximum of Order Date, count of Customer ID and Sum of Sales. This helped me get Recency, Frequency and Monetary value of each customer.

For recency, I have calculated by taking max of all the Order Dates – minimum order date for each customer. For example, if max of all Order Dates is 05-01-2018 so I will reduce max of all customers all Order Dates with 05-01-2018. This will be my recency variable. Count of Customer ID is my frequency variable whereas Sum of Sales will be my monetary variable. These will show values for each customer as Customer ID was kept on row side of the pivot table.

Then, minimum and maximum (min and max) of each variable is picked and they are divided into 5 equal parts i.e., 20% in each. Then, points are assigned points to each range which is from 1-5 as shown in Table 3. These points help us to segment based on their score.

More the frequency and monetary value, higher the score and in recency, the lesser number of days, more the score. For the same reason, Min of recency has score 5 and max or frequency and monetary value has score 5.

Table 3: Scores assigned as per minimum and maximum of the dataset- Recent days (When was the last purchase done), Frequency (of purchases over the time), Monetary Value- (Total of all transactions)

Chart 1: Count of each score for Recency, Frequency and Monetary Value

Also, segmentation can be shown as 1, 1, 1 or 111. It depends on the company. However, both meaning the same grade. Also, the sequence taken is of RFM - recency score, frequency score and then monetary score so it is important to follow the same sequence in whole study. You can also take score as MFR or FRM but then the segment names will change so to avoid any confusion, it is better to follow the sequence of RFM.

Segmentation:

Table 4 explains how I have combines MFR points of every customer and segmented them on the basis.

1, 1, 1 means the customer falls into lost segment while 5,5,5 falls under champions. In the dataset I have used 5, 5, 5 being the highest comes into champions' segment. The highest score need not be 5, 5, 5. The lowest score can be 1, 1, 1 as it is possible that a person might be travelling and would have stopped for some purchases. Hence recency, frequency and monetary value can be 1. Segmentation depends on the nature of product, Industry, Location, and the response from customers. There is no specific metric to segment customers. It depends on industry to industry and company to company.

Table 4: Counts of each score in each RFM segment

Image 1: RFM scores further segmented into high, medium, and low CLV

Chart 2: RFM Customer segmentation

In the chart 2 most of the customers i.e., 157 belong to the segment: Potential Loyalist. Hibernating being second highest, occupy only 102 customers. While third highest, Champions, being the highest in terms of scores occupies 93 customers. The numbers are quite high for a the highest as it falls amongst the top 3 segments. Hence, it can be said that the supermarket is doing well. However, there is always a room for the improvement and so strategies can be designed to retain top segments and focus on the lower ones. The lowest scores belong to recent Customers. Promising, about to sleep and cannot afford to lose them are around the same range.

Chart 3 depicts top 20 customer scores:

Chart 3: Top 20 customer scores with their segments

Below are the strategies for each CLV segment:

- ***Low CLV score range is from 1, 1, 1 to 2, 5, 5***
- ***Medium CLV score range is from 3, 1, 1 to 4, 5, 5***
- ***High CLV score range is from 5, 1, 1 to 5, 5, 5***

CLV	Purchase Pattern	Strategies for each CLV segment
High CLV	• Bought as of late, purchase regularly, and spend the most! • Responsive to advancements. Ongoing clients yet spent a decent sum and purchased more than once. • Recent buyers but have not spent a lot.	• Can be early adopters for new items, offer free preliminaries • Upsell higher worth items. • Ask for reviews and engage them. Offer participation/unwaveringness program, suggest different items.
Medium CLV	• Bought most as of late, however not frequently • Recent customers yet have not spent a lot. Bought more than once. • Above normal recency, recurrence, and financial qualities. Might not have purchased as of late however.	• Provide on-boarding support, begin building relationship, recommend items. • Create brand mindfulness, offer participation/loyalty program. • Make restricted time offers, recommend based on past purchases, • Reactivate them with offers.

Low CLV	• Lowest recency, recurrence, and money related scores. • Last purchase was long back, low spenders and low number of requests. • Spent large cash and bought frequently. However, quite a while past. Need to bring them back!	• Revive interest with connect crusade, overlook in any case. Offer other significant items and extraordinary limits. • Reproduce brand esteem. Win them back through recharges or more up to date items, don't lose them to rivalry, converse with them. • Send customized messages to reconnect, offer recharges, give accommodating assets.

In the above demonstration, we have used RFM score to segment customers and have mentioned strategies to increase customer lifetime value for each segment.

Now, we will use the same data and check the Active probability of the customers i.e., estimating the chance of a customer is still Active or not. To predict company's future profits, it is important to predict your active customers. Method given by Schmittlein, Morrison, and Colombo estimates the probability that the customer is still active. You need to have a data with 3 variables: N = Number of buys, t = Time of last buy, T = Time elapsed between acquisition of customer and present time.

After defining T* = t/T, the authors show that (T*) estimates the probability that the customer is still active. For instance, if T = 10 and customer has made buy, 1, 5, 6 and 9. One can estimate that a customer is still active to equal 0.9 ^ 4= 0.6561.

Chart 4: Range of Active probability along with the count of customers in each range

With Chart 4 as reference, most of the clients fall under active probability range above 50%. Highest being the range 67% to 87% (3 ranges in the chart).

RFM – Variations and Limitations:

Variation of RFM include RFD (Recency Frequency Duration) and RFE (Recency Frequency Engagement), replacing Monetary Value by Duration and Engagement, respectively.

RFM segmentation focuses on just three variables while there could be others that are critical for a business. RFM model has its weakness, but some can be resolved with minor modification or extension like sub-segmentation of customers.

V. SUMMARY

Customer retention is a serious concern in India, especially in retail sector which is growing leaps and bounds in India. Calculating Customer lifetime value is a challenge for nearly every organization. Following are the benefits of the study. Identifying a profitable customer base. Clear understanding of how to grow business and profits, based on a better understanding of their customers. An ability to respond quickly to the customer needs and changes in the business environment and better protection in downturns, due to more loyal customers. Maintaining Healthy Customer base and helps acquire in-depth knowledge to grow business and profits with the fastest response time for changes and needs will be fast. More loyal customers make company profitable. The main objective of this research is to study about the relationship between the customers and their transactional details. RFM strategy can help to segment customers according to their buy pattern. Also, the relation between number of years since they became a customer, understand their trends, and a look at the possible measures. The key to achieving better retention, at its considerably basic, is, understanding the core issues that drive turnover.

VI. REFERENCES

- Kim, H. K.; Im, K. H. & Park, S. C. (2010). DSS for computer security incident response applying CBR and collaborative response, *Expert Systems with Applications*, Vol. 37, No. 1, (January 2010) 852-870, ISSN:0957-4174.
- Chan, C.C.H. (2008). Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer, *Expert Systems with Applications*, Vol. 34, No. 4, (May 2008) 2754-2762, ISSN:0957-4174.
- Chiu, C-Y.; Kuo, I-T. & Chen, P-C. (2009). A market segmentation system for consumer electronics industry using particle swarm optimization and honey bee mating optimization, *Global Perspective for Competitive Enterprise*, *Economy and Ecology*, Springer London, pp. 681- 689.

- Olson, D.L.; Cao, Q.; Gu, C. & Lee, D. (2009). Comparison of customer response models,
    - *Service Business*, Vol. 3, No. 2, (June 2009) 117-130, ISSN: 1862-8516.
- Quinlan, J. R. (1993). *C4.5 Programs for Machine Learning*, Morgan Kaufmann Publishers. 302 pages.
- Chen, Y-L.; Kuo, M-H.; Wu, S-Y. & Tang, K. (2009). Discovering recency, frequency, and monetary (RFM) sequential patterns from customers' purchasing data, *Electronic Commerce Research and Applications*, Vol. 8, No. 5, (October 2009) 241-251, ISSN: 1567-4223.
- Winer, R. S. (2001). A framework for customer relationship management. California Management Review, 43(4), 89-108.
- Zeithaml, V.A., Rust, R.T. & Lemon, K.N. 2001. 'The customer pyramid: Creating and serving profitable customers', California Management Review, 42(4): 118– 142
- Cheng, C-H. & Chen, Y-S. (2009). Classifying the segmentation of customer value via RFM model and RS theory, *Expert Systems with Applications*, Vol. 36, No. 3, (April 2009) 4176-4184, ISSN: 0957-4174.
- Venkatesan, R. & Kumar, V. 2004. 'A customer lifetime value framework for customer selection and resource allocation', Journal of Marketing, 68(4): 106-125
- Cooil, B., Keiningham, T. L., Aksoy, L., & Hsu, M. (2007). A longitudinal analysis of customer satisfaction and share of wallet: Investigating the moderating effect of customer characteristics. Journal of Marketing, 71(1), 67-83.
- A marketing view of the customer value: Customer lifetime value and customer equity by A.M. Estrella-Ramón and M. Sánchez-Pérez
- Reinartz, W., & Kumar, V. (2002). The mismanagement of customer loyalty. Harvard Business Review, 80(7), 86-94.
- Rust, R. T., & Zahorik, A. J. (1993). Customer satisfaction, customer retention, and market share. Journal of Retailing, 69(2), 193-215.
- McCarty, J. A. & Hastak, M. (2007). Segmentation approaches in data mining: A comparison of RFM, CHAID, and logistic regression, *Journal of Business Research*, Vol. 60, No. 6, (June 2007) 656-662, ISSN:0148-2963.
- Chiu, C-Y.; Kuo, I-T. & Chen, P-C. (2009). A market segmentation system for consumer electronics industry using particle swarm optimization and honey bee mating optimization, *Global Perspective for*

Competitive Enterprise, Economy and Ecology, Springer London, pp. 681- 689.

- Chuang, H. & Shen, C. (2008). A study on the applications of data mining techniques to enhance customer lifetime value – based on the department store industry, *Proceedings of the 7th International Conference on Machine Learning and Cybernetics*, pp. 168-173, ISBN: 978-1424420964, Kunming, China, July 2008, IEEE.
- Blattberg, R.C.; Kim, B-D. & Neslin, S.A. (2008). *Database Marketing: Analyzing and Managing Customers*, Chapter 12, pp. 323-337, Springer, ISBN: 978-0387725789, New York, USA.
- Bult, J. R. & Wansbeek, T. (1995). Optimal selection for direct mail, *Marketing Science*, Vol. 14, No. 4, (Fall 1995) 378-394, ISSN:0732-2399.
- Kahan R (1998). Using database marketing techniques to enhance your one-to-one marketing initiatives. J. Consum. Mark., 15(5): 491-493.
- Hosseini, S.M.; Maleki, A. & Gholamian, M.R. (2010). Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty, Expert Systems with Applications, Vol. 37, No. 7, (July 2010) 5259–5264, ISSN:0957-4174
- Wang CH (2010). Apply robust segmentation to the service industry using kernel induced fuzzy clustering techniques. Expert Syst. Appl., 37: 8395-8400.
- Hughes AM (1994). Strategic Database Marketing. Probus Publishing, Hughes AM (1996). Boosting reponse with RFM. Mark. Tools, 5: 4-10.
- Hung C, Tsai C (2008). Market segmentation based on hierarchical self-organizing map for markets of multimedia on demand. Expert Syst. Appl., 34(1): 780-787
- Book by Derya B and Prof. Kimito F named as Data Mining Using RFM
- A review of the application of RFM model by Jo-Ting Wei, Shih-Yen Lin and Hsin-Hung Wu
- Hill, N., Roche, G. & Allen, R. 2007. Customer Satisfaction: The customer experience through the customer's eyes. London: Cogent Publishing Ltd.
- Reversing the Logic: The Path to Profitability through Relationship Marketing by Author: V. Kumar Ilaria Dalla Pozza, J. Andrew Petersen, Denish Shah
- Reconciling Performance and Interpretability in Customer Churn Prediction using Ensemble Learning based on Generalized Additive Models by Koen W. De Bock and Dirk Van den Poel
- Classifying the segmentation of customer value via RFM model and RS theory by Author: Ching-Hsue

Cheng, You-Shyang Chen

- Customer Satisfaction and Customer Loyalty by Kabu Khadka & Soniya Maharjan
- https://www.putler.com/rfm-analysis/
- https://dma.org.uk/article/rfm-modelling-easily-generate-successful-customer-segments
- https://searchcustomerexperience.techtarget.com/definition/customer-lifetime-value-CLV
- https://alexgenovese.it/blog/applying-rfm-customer-segmentation-to-your-business-right-now/
- https://is.cuni.cz/webapps/zzp/download/120167358
- https://wps-feb.ugent.be/Papers/wp_12_805.pdf
- https://postfunnel.com/rfm-analysis-important-e-commerce/
- https://www.unc.edu/
- https://www.optimove.com/learning-center/deep-customer-analytics
- https://www.coursehero.com/file/p2al2ue/Thats-the-catch-Instead-of-reaching-out-to-100-of-your-audience-you-need-to/
- http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.615.7927&rep=rep1&type=pdf
- https://custora.com/tour/feature_predictive_customer_lifetime_value_clv_retail/
- https://link.springer.com/content/pdf/10.1057%2Fpalgrave.im.4340243.pdf
- http://shodhganga.inflibnet.ac.in/bitstream/10603/11075/6/06_chapter2.pdf
- https://www.researchgate.net/publication/228399859_A_review_of_the_application_of_RFM_model
- http://iaiest.com/dl/journals/5-%20IAJ%20of%20Accounting%20and%20Financial%20Management/v3-i6-jun2016/paper3.pdf
- https://www.itproportal.com/features/emerging-technology-and-the-race-to-uptime/
- https://www.retaildoc.com/lifetime-value-of-customer-service-retail
- https://www.shopify.in/blog/customer-lifetime-value
- https://internetretailing.net/2018/01/guest-comment-encouraging-customer-loyalty-post-festive-season/

This research includes secondary data collection from Kaggle https://www.kaggle.com/juhi1994/superstore?select=US+Superstore+data.xls and application of statistical analysis to this data. The insights gathered from this research and analysis served as the outcome of this research.

Tables & Figures

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Na	Segment	Country	City	State
1	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky
2	CA-2016-152156	08-11-2016	11-11-2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky
3	CA-2016-138688	12-06-2016	16-06-2016	Second Class	DV-13045	Darrin Van H	Corporate	United States	Los Angeles	California
4	US-2015-108966	11-10-2015	18-10-2015	Standard Clas	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderda	Florida
5	US-2015-108966	11-10-2015	18-10-2015	Standard Clas	SO-20335	Sean O'Donn	Consumer	United States	Fort Lauderda	Florida
6	CA-2014-115812	09-06-2014	14-06-2014	Standard Clas	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
7	CA-2014-115812	09-06-2014	14-06-2014	Standard Clas	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
8	CA-2014-115812	09-06-2014	14-06-2014	Standard Clas	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
9	CA-2014-115812	09-06-2014	14-06-2014	Standard Clas	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California
10	CA-2014-115812	09-06-2014	14-06-2014	Standard Clas	BH-11710	Brosina Hoffr	Consumer	United States	Los Angeles	California

Table 1: First 11 columns of the dataset

Postal Code	Region	Product ID	Category	Sub-Categor	Product Nan	Sales	Quantity	Discount	Profit
42420	South	FUR-BO-1000	Furniture	Bookcases	Bush Somers	261.96	2	0	41.9136
42420	South	FUR-CH-1000	Furniture	Chairs	Hon Deluxe F	731.94	3	0	219.582
90036	West	OFF-LA-1000	Office Suppli	Labels	Self-Adhesive	14.62	2	0	6.8714
33311	South	FUR-TA-1000	Furniture	Tables	Bretford CR4!	957.5775	5	0.45	-383.031
33311	South	OFF-ST-1000	Office Suppli	Storage	Eldon Fold 'N	22.368	2	0.2	2.5164
90032	West	FUR-FU-1000	Furniture	Furnishings	Eldon Expres	48.86	7	0	14.1694
90032	West	OFF-AR-1000	Office Suppli	Art	Newell 322	7.28	4	0	1.9656
90032	West	TEC-PH-1000	Technology	Phones	Mitel 5320 IP	907.152	6	0.2	90.7152
90032	West	OFF-BI-1000	Office Suppli	Binders	DXL Angle-Vi	18.504	3	0.2	5.7825
90032	West	OFF-AP-1000	Office Suppli	Appliances	Belkin F5C20	114.9	5	0	34.47

Table 2: Remaining columns of the dataset

Recency (Days)			Frequency			Monetary value (Total)		
Score	Min	Max	Score	Min	Max	Score	Min	Max
5	0	28	5	19	37	5	4282.64	25043.05
4	29	54	4	14	18	4	2750.76	4282.63
3	55	104	3	11	13	3	1739.44	2750.75
2	105	226	2	8	10	2	961.55	1739.43
1	227	1167	1	1	7	1	4.83	961.54

Table 3: Scores assigned as per minimum and maximum of the dataset- Recent days (When was the last purchase done), Frequency (of purchases over the time), Monetary Value- (Total of all transactions)

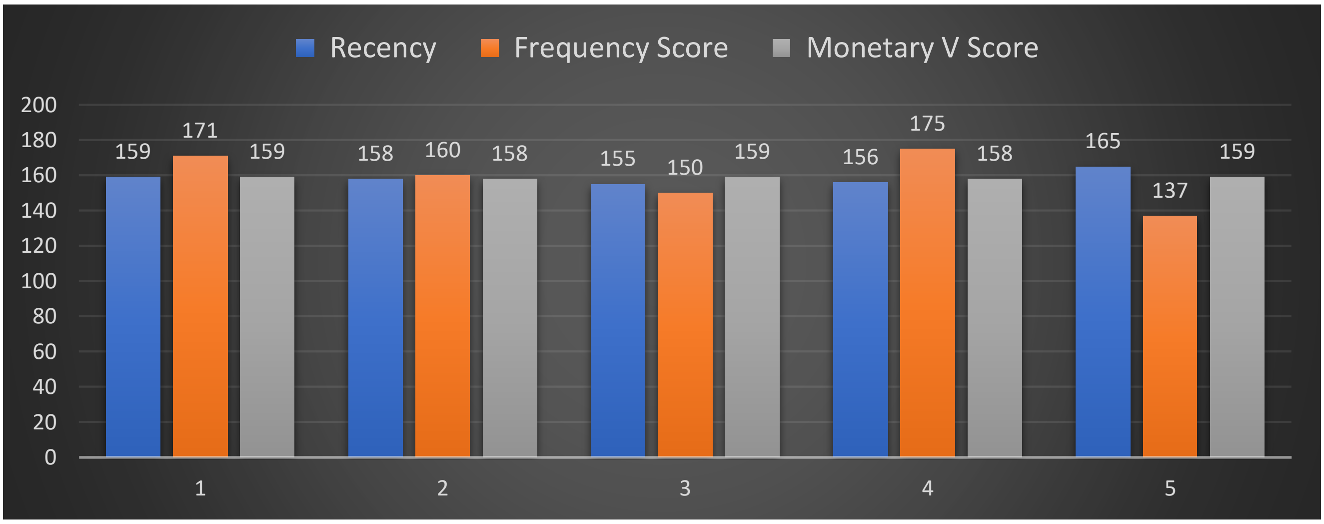

Chart 1: Count of each score for Recency, Frequency and Monetary Value

Count (M, F, R)	No.	Customer Segment
1, 1, 1	42	Lost
1, 1, 2	15	Lost
1, 1, 3	17	Can't Lose Them
1, 1, 4	11	Can't Lose Them
1, 1, 5	13	Can't Lose Them
1, 2, 1	13	Lost
1, 2, 2	8	Hibernating
1, 2, 3	5	Hibernating
1, 2, 4	6	At Risk
1, 2, 5	8	At Risk
1, 3, 1	1	Lost
1, 3, 2	5	Hibernating
1, 3, 3	5	At Risk
1, 3, 4	1	At Risk
1, 3, 5	5	At Risk
1, 4, 2	3	At Risk
1, 4, 3	1	At Risk
2, 1, 1	18	Hibernating
2, 1, 2	11	Hibernating

Count (M, F, R)	No.	Customer Segment
3, 4, 1	6	Potential Loyalist
3, 4, 2	6	Potential Loyalist
3, 4, 3	7	Customers Needing Attention
3, 4, 4	9	Loyal Customers
3, 4, 5	19	Loyal Customers
3, 5, 1	2	Potential Loyalist
3, 5, 2	1	Potential Loyalist
3, 5, 3	6	Potential Loyalist
3, 5, 4	5	Loyal Customers
3, 5, 5	6	Loyal Customers
4, 1, 1	2	Recent Customers
4, 1, 2	1	Recent Customers
4, 1, 3	1	Promising
4, 1, 4	1	Promising
4, 2, 1	2	Recent Customers
4, 2, 2	6	Recent Customers
4, 2, 3	8	Potential Loyalist
4, 2, 4	7	Promising
4, 2, 5	4	Promising

2, 1, 3	9	About To Sleep
2, 1, 4	3	Can't Lose Them
2, 1, 5	5	Can't Lose Them
2, 2, 1	10	About To Sleep
2, 2, 2	10	Hibernating
2, 2, 3	12	Hibernating
2, 2, 4	7	At Risk
2, 2, 5	6	At Risk
2, 3, 1	5	About To Sleep
2, 3, 2	9	Hibernating
2, 3, 3	4	Hibernating
2, 3, 4	9	At Risk
2, 3, 5	9	At Risk
2, 4, 1	3	Hibernating
2, 4, 2	7	At Risk
2, 4, 3	4	At Risk
2, 4, 4	3	At Risk
2, 4, 5	6	At Risk
2, 5, 1	2	Hibernating
2, 5, 3	1	At Risk
2, 5, 4	2	At Risk
2, 5, 5	3	At Risk
3, 1, 1	7	Recent Customers
3, 1, 2	5	About To Sleep
3, 1, 3	2	Promising
3, 1, 4	2	Promising
3, 1, 5	4	Promising
3, 2, 1	10	About To Sleep
3, 2, 2	6	Hibernating
3, 2, 3	7	Potential Loyalist
3, 2, 4	9	Customers Needing Attention
3, 2, 5	3	Customers Needing Attention
3, 3, 1	7	About To Sleep

4, 3, 1	5	Potential Loyalist
4, 3, 2	5	Potential Loyalist
4, 3, 3	10	Potential Loyalist
4, 3, 4	7	Customers Needing Attention
4, 3, 5	6	Loyal Customers
4, 4, 1	6	Potential Loyalist
4, 4, 2	11	Potential Loyalist
4, 4, 3	15	Customers Needing Attention
4, 4, 4	14	Loyal Customers
4, 4, 5	11	Champions
4, 5, 1	1	Potential Loyalist
4, 5, 2	6	Potential Loyalist
4, 5, 3	9	Potential Loyalist
4, 5, 4	11	Champions
4, 5, 5	9	Champions
5, 1, 3	1	Promising
5, 1, 4	1	Promising
5, 2, 1	2	Promising
5, 2, 2	1	Promising
5, 2, 3	4	Promising
5, 2, 4	3	Promising
5, 2, 5	3	Promising
5, 3, 1	3	Potential Loyalist
5, 3, 2	9	Potential Loyalist
5, 3, 3	4	Potential Loyalist
5, 3, 4	5	Customers Needing Attention
5, 3, 5	6	Customers Needing Attention
5, 4, 1	4	Potential Loyalist
5, 4, 2	9	Potential Loyalist
5, 4, 3	7	Loyal Customers
5, 4, 4	12	Champions
5, 4, 5	12	Champions
5, 5, 1	8	Potential Loyalist

3, 3, 2	9	Hibernating
3, 3, 3	4	Potential Loyalist
3, 3, 4	9	Customers Needing Attention
3, 3, 5	8	Loyal Customers

5, 5, 2	15	Potential Loyalist
5, 5, 3	12	Potential Loyalist
5, 5, 4	19	Champions
5, 5, 5	19	Champions

Table 4: Counts of each score in each RFM segment

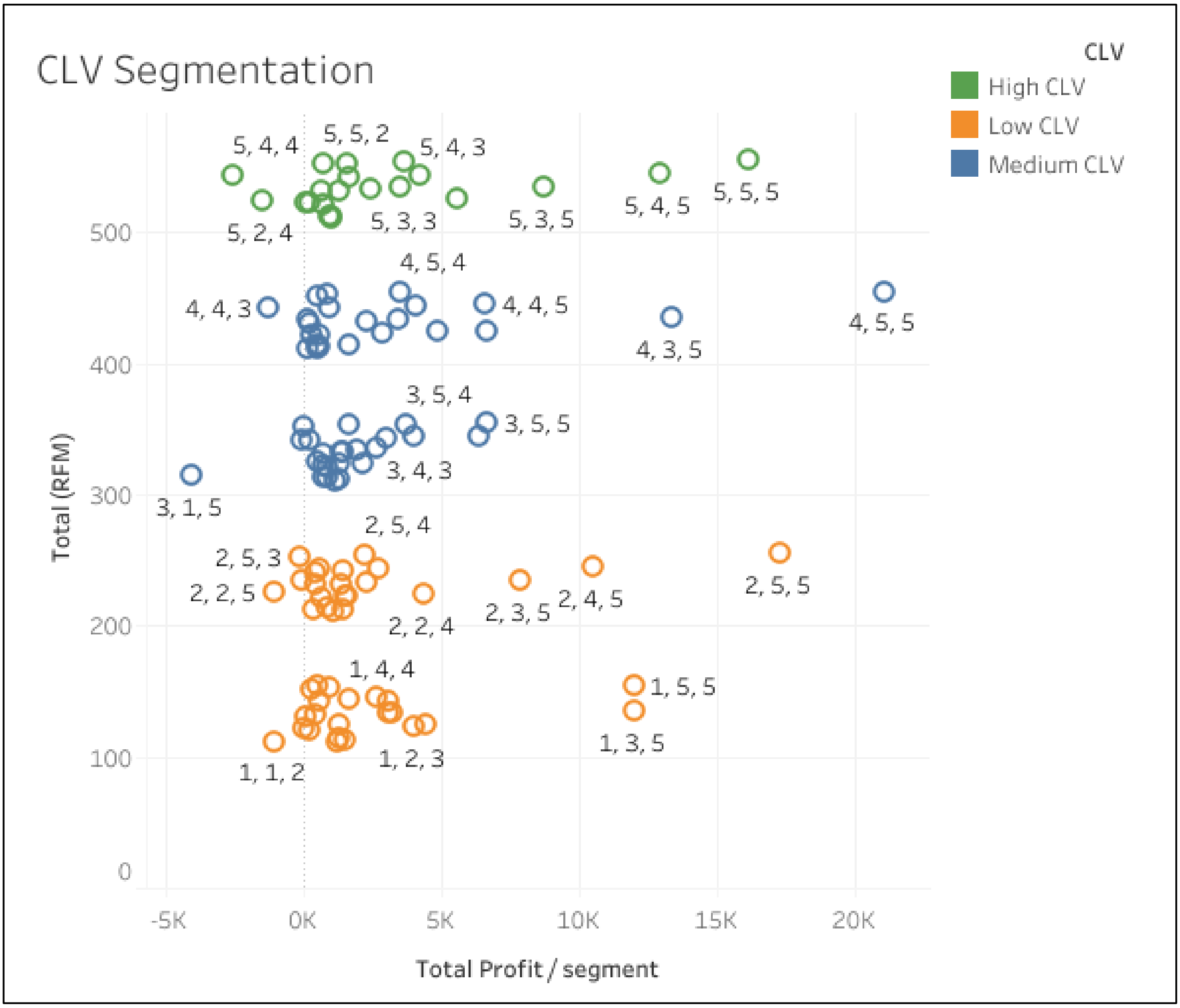

Image 1: RFM scores further segmented into high, medium, and low CLV

Chart 2: RFM Customer segmentation

Chart 3: Top 20 customer scores with their segments

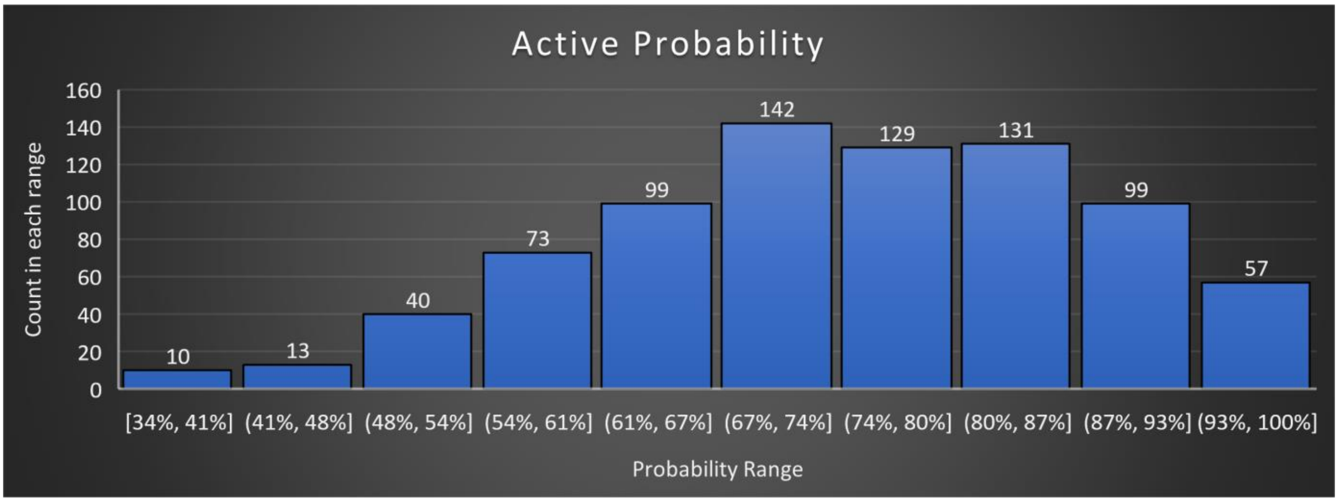

Chart 4: Range of Active probability along with the count of customers in each range