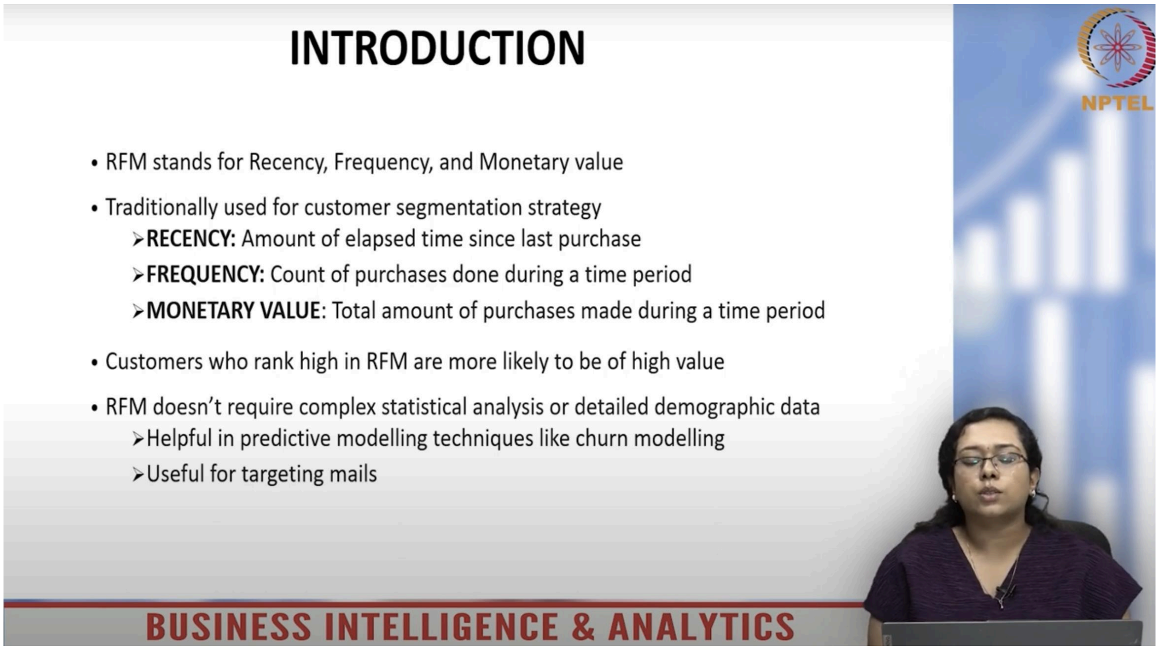
Course Name:Business Intelligence and Analytics
Professor Name:Prof. Saji.K.Mathew
Department Name:Department of Management Studies
Institute Name:Indian Institute of Technology Madras
Week:10
Lecture:37

RFM ANALYSIS

Hello all and welcome to today's Business Intelligence and Analytics course. Today we are diving into a new topic which is known as RFM Analysis. I do not know if you might have heard about this, but RFM stands for Recency, Frequency and Monetary Value. So we are going to see what RFM Analysis is and where it is used and how it can bring more insights to the data that companies have. So RFM Analysis is a marketing analytics tool or mostly it is used by businesses for their marketing purposes in order to segment customers into different buckets.

So why do we need to segment our customers? So most of the businesses, they segment customers in order to, maybe to increase profit or for marketing purposes or for direct emailing. So not every customer needs to be treated similarly because there might be high net worth customers, there might be customers who purchase frequently, but the monetary value of those purchases would be very less. So that is why we need a recency, frequency and monetary analysis in order to bucket these customers into different groups. So these groups might vary and it depends on how the business wants to segment their customers.

So this RFM Analysis was born from the idea from a paper that was published in Marketing Science Journal and by Jan and Tom. So it was broadly based on direct mail marketing and how the businesses get their profits or the 80% of the profits come from 20% of the customers. So that is how the importance of segmenting the customers into different buckets or different segments, so that we will know who the top 20 percentile customers are, so that we can you know give personalized offers or loyalty programs or, you know even give them more offers during their purchases so that they still remain loyal as long as the business exists. So that is the overview of what RFM is and we will see how it is done. So as I already explained, RFM stands for recency, frequency and monetary value.



So traditionally it has been used as a customer segmentation strategy wherein you bucket these customers, say there are millions of customers, so not every customer would be alike. The customers may come frequently to the store, if it is an in-store kind of business model then the customers may come to the store frequently but for very low value purchases like toothbrush or daily items but some customers may come very very less frequently but they might be purchasing a, say you know high value purchase like some video games or some high quality TVs or even, you know high quality purchases or high monetary value purchases. So not, these two customers should not be treated alike. If you want to market some product to them or you know you want to send out marketing mails to them, then these two customers fall in different buckets because one customer continuously buys low value purchases, so he needs to be marketed with items that are of his or her interest and the second customer who is frequently making high value purchases probably uses some other store to, you know purchase these daily items but uses this store to make high value purchases. So he has to be targeted in a different way. So that is why this RFM analysis is so important.

So what is recency? Recency is nothing but how recent is a customer. For example, if the customer visits a store has visited a store one month back, then he is supposed to be recent. If he has visited the store probably, you know 5 years back then he is not a recent customer. So recency is nothing but a measure of how much time has elapsed between the last purchase and the present date. So that is what recency stands for. This is one of the most important matrix which the businesses use to know if the customers are still their own customers or if they have churned or, you know moved to some other businesses.

For example, you might not use Amazon always to make your purchases. Sometimes you might use Flipkart or Tata Neu or so many other, you know online retails but how does Amazon, you know come to know that you are a recent customer or not? By analyzing how much time you know you spend visiting their website or how recently you have made a purchase. So that is how they come to know about your recency values. So if you were a good customer, you had made high net worth purchases from Amazon, say you know 10 lakhs worth of purchase over last, you know 5 years then you are a good customer and Amazon should not, you know have the chance of, you know you churning off to some other website. So what Amazon will do is, see your recency metrics like how recently you have come to Amazon, what have you searched, have you actually made that purchases which you searched, did you add some products to the market, to the cart.

So all there are lots of metrics but RFM analysis, it does not take your psychographic data or any other data. It just deals with 3 very very simple things that is how much recently have you made the purchase, how frequently are you coming to make the purchase and monetary value is nothing but the total amount of purchases that you have made over a time horizon. It might be the entire customer lifetime like, you know from the day you became a customer to a particular business to till date or it may be like you know previous 5 years. So it can be a time period or your entire customer time period. So that is what recency is and frequency is nothing but the total count of purchases that you have done during the time period. Say if it is 5 years, how many times have you visited the store. If we are talking about a brick and mortar kind of stores or if it is a online retail website, then how many times have you frequented that website is something that is known as the frequency matrix .

And the monetary value is, it is direct. So it is nothing but how many what is the total amount that you have spent to purchase all the things over a particular time period. So high RFM is good or low RFM is good. Obviously high RFM means you are a more loyal and high valued customer. So if it is, if your recency value is very high, that means you are very recent. If your frequency value is very high, that means you frequent this store the most, that means you are a very loyal customer and if your monetary value is also high, then you make high value purchases in the store.

That means high RFM means the customer is likely to be high value and what should the business do if a customer is of high value? The business should concentrate on not losing the customer at any cost. So if RFM value is calculated to be very high, then at no cost should the company lose the customer. There are, you know multiple brands or multiple companies working in the same segment. For example, if it is online shopping then there are so many business that work in the same segment.

So it is very important that, you know you might churn off to some other business if you do not like the way that, you know the customer service is treating you or if your order is very delayed. So all these metrics help you, help the business to identify if the customers are staying loyal and it also gives you the prediction of how the customer will act in the future as well.

So as I already told, RFM does not require any complex statistical analysis or any detailed demographic data. So it just takes the R value, F value and the M value and no other values. Are there other values? There are millions of values that you can analyze to know how your customer is.

For example, if you are visiting an online website like Amazon, then you might be looking at some product you might be scrolling and, you know looking at the product but you will not click on it or so, or adding it to the cart. So there will be look to click value, click to cart value and cart to buy value. So there are so many metrics that the online business can use, to know how the customers are acting or will act in the future. But our RFM analysis is very very simple. It does not need any complex algorithms or complex analysis techniques to know the customer performance.

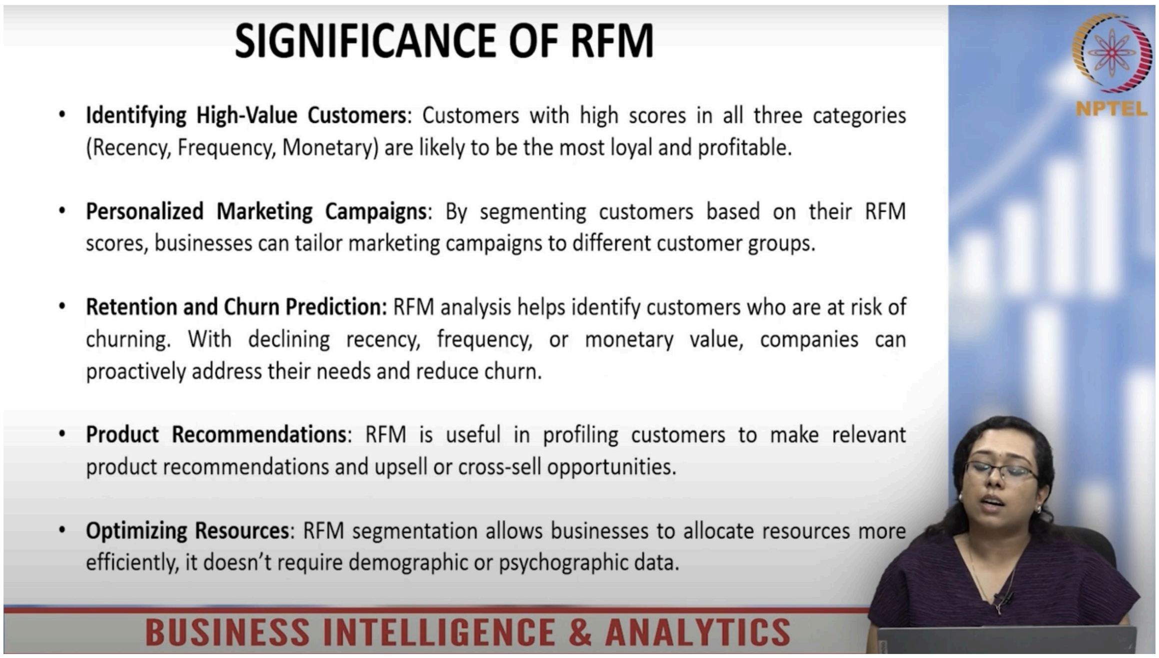
It just depends on the three values that is R, F and M. So churn modeling, it is helpful in predictive modeling techniques like churn modeling. So RFM analysis is based on the data that is already existing now with the business, like how much recent the customer has been, how much frequent the customer has been and how much monetary value he or she has purchased during his or her lifetime. So all these values are pre-existing. So it is not a futuristic data that is, based on which RFM is analyzed.

So it is a past value that is being analyzed. But can RFM analysis also predict the future? Yes. Because how the customer has been acting till now, is a good indicator of how he or she will be acting in the future as well. So this RFM analysis can also act for predictive modeling like churn modeling. Churning is nothing but when you leave one business to move to the other business.

For example, you are moving completely from Amazon to Flipkart because you did not like something from Amazon. So that is known as churning off. So it is useful to know the predictive possibilities of, if a customer would be churning or not. So it is also, RFM is also used for targeting mails.

So for example, in US various zip codes are segmented into different types. For example, people from a particular zip code might not be you know purchasing high value items.

But people from you know a particular other zip code might be very rich and would be always purchasing high value items. So marketing mails or direct marketing mails are not sent in a similar way to both the zip codes. To the zip code where people make very low value purchases are not, you know marketed with items that are, you know very costly. So that is what is known as targeting mails and RFM analysis is a very good method for targeting mails.



Next we will look into the significance of RFM. So most points we have already discussed. So RFM is a very simple technique but it is also highly effective in customer segmentation. And the first point that is, identifying high value customers. So as we already saw high R, high F and high M means the customers are most likely to be loyal as well as profitable.

Loyal because the recency and frequency is high, profitable because monetary value is high. So for example, say you have a Netflix account and you are, you know you have not been frequently visiting the Netflix account for a long time. Then Netflix will come to know that you are not visiting. So it can send you prompts or messages to, you know attract you with messages like, there is a new movie release, this might interest you and so on. Similarly, if you are using a BookMyShow app or something like that, then it might, if you are not frequently visiting or recently you have not visited the app, then you might be seeing notifications trying to attract you into visiting that website. So that is what these businesses do in order to prevent churn. So that is done in order to protect the high value customers that the businesses have.

Second point is personalized marketing campaigns. So by segmenting customers into various different segments, we can, you know tailor the marketing campaigns to different customer groups. I have already told while discussing the targeting mail example that you know not every customer is treated alike.

So high net worth customers are treated in a different manner. They are shown with things that are of their interest. For example, a person who likes violent or action movies are shown with movies of his or her interest and not rom-com movies.

Third point is retention and churn prediction. So RFM analysis helps to identify customers who are at the risk of churning, who are dissatisfied with the business, so that they are looking into another business which they can churn off to.

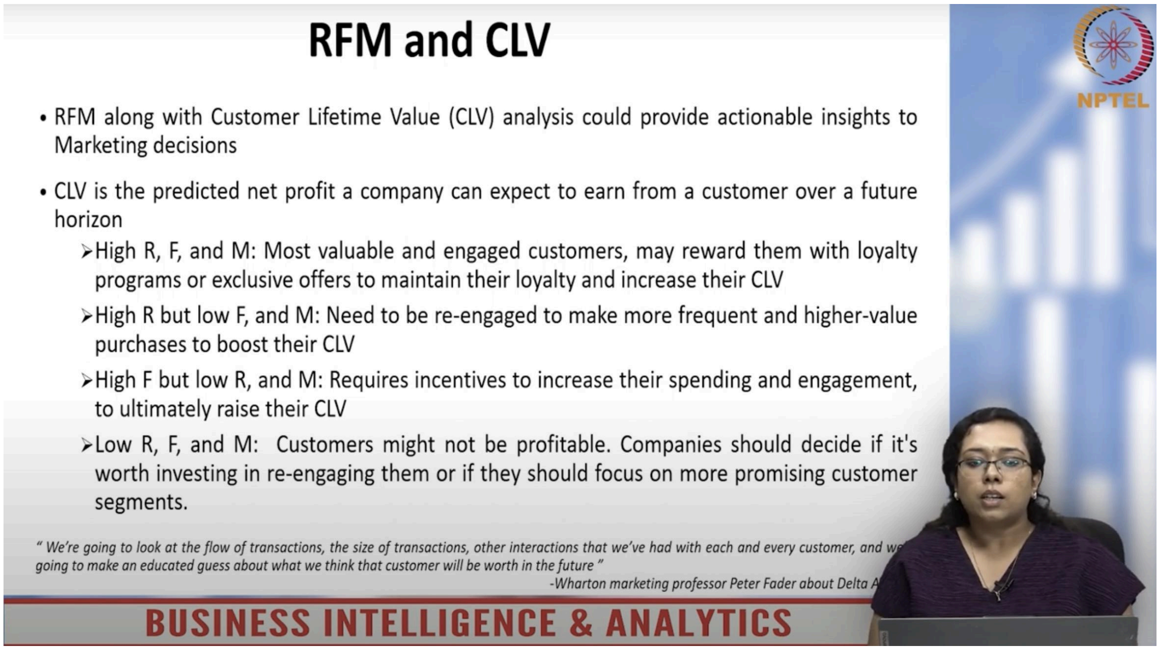
So I have already said what churning is and so to retain these customers who are at the risk of churning or who are at the brim of churning is something that businesses focus on. So with declining recency, frequency or monetary values the companies can, you know know which customers are likely to churn off especially if the customers were of high net worth or high RFM previously and now their RFM is decreasing, that is an indicator that the customer may churn. So business might proactively give out strategies for retaining those customers in their business itself.

The fourth point is product recommendations. So as I already told, product recommendations or recommendation systems that are modeled by various businesses work on RFM values.

So high RFM value customers are given with more attractive offers, while recommending products, they are given more preference during some sale or offers that are going on. So they are given relevant product recommendations and to upsell or cross-sell opportunities.

The fifth and the final point is optimizing resources. This also I have told that RFM is a technique which does not use many complex statistical analysis tools. It just depends on three matrices that is R, F and M.

So this kind of segmentation allows business to allocate resources more efficiently and since it does not require any demographic or psychographic data. So we need not give out complex analysis techniques to do RFM. It can be done very simply using three matrices. So this does not overuse the business resources like other models. So this is quite simple and easy and uses, it optimizes resource usage.



RFM and CLV. What CLV? CLV is a customer lifetime value. So what is RFM? RFM is the three matrices that would determine how the customer has acted in the past. CLV is something that is a predicted possibility. So CLV is the predicted net profit a company can expect from a customer throughout a time period, throughout a future time horizon. Say it can be the entire lifetime of the customer, in future how much profit he will bring to the company. So that is what is known as CLV. So

CLV is a predicted value and RFM is a value that is already available with us. So high R, high F and high M value customers. They are the most valuable, profitable and engaged customers. So what businesses can do is give them rewards, offers, loyalty programs, etc. So high R, low F and M customers that means they have high recency, that is they have visited recently but they have low frequency.

They are not common visitors, they do not come often and their monetary value is also low, that means they have come recently. They are quite new to the business. Say you have just installed your Amazon app today and you are just going and seeing the Amazon, if it is better than a brick and mortar store that is present. So that means that you have come recently but you have not come frequently or you have not purchased anything. So the strategy for these kinds of customers need to be different from the first bucket. That is these customers need to be attracted or pulled into the business so that they remain as high RFM customers.

So they need to be re-engaged to make them more frequent or to make them do high value purchases. So there is a re-engagement strategy that need to be focused by the businesses. And the third category is high F, low R and M. So high F means you are a frequent customer, you come to the store often but you have not come to the store recently. You used to come frequently but you have not come for a past particular time and your monetary value is also less. That means earlier you used to come frequently. That means there needs to be incentives or offers that need to be given to such customers to increase their spending and engagement, so that it can lead to high CLV.

So why we are focusing on this R, F and M values is that in future the customer lifetime value can be increased or the predicted CLV values can be increased, so that the customer also remains as a loyal customer in the future.

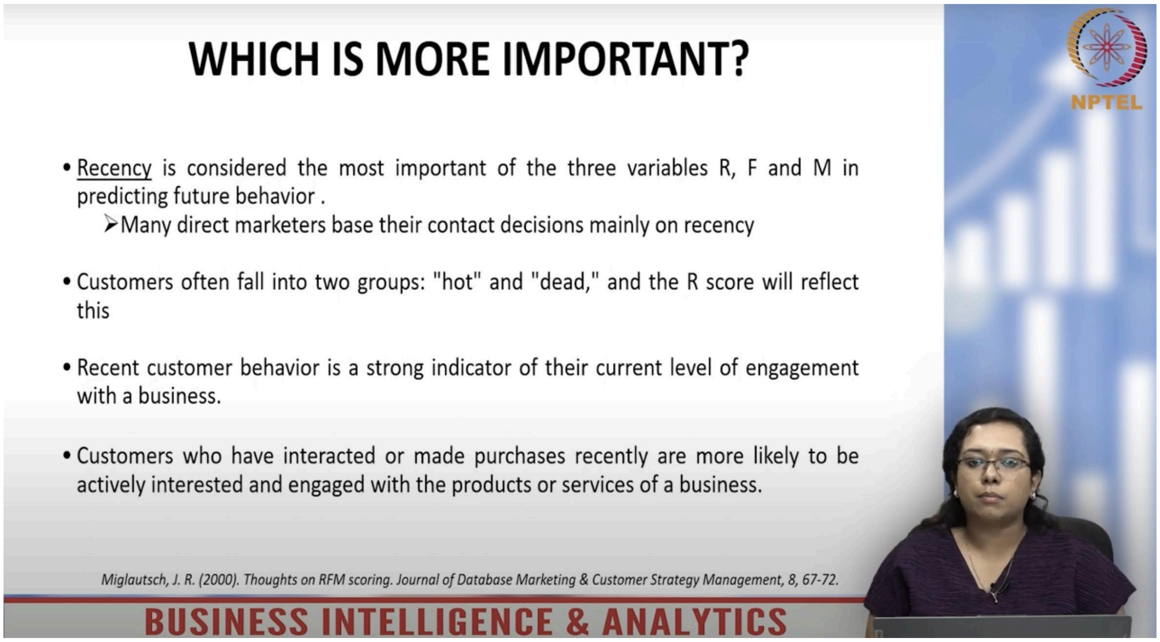
And the last segment is low R, low F and low M. Everything is low, that means the customers are not profitable and they are not recent or frequent. So companies should have a strategy to know if it is worth investing in them or engaging in them or if they should focus on more promising segments like the first three segments.

So the last line that you can see in the slide is given out by a consultant who is a Wharton marketing professor. This is a line that he is talking about Delta Airlines. So Delta Airlines is a huge company and it has said that it uses RFM strategy to market its customer segments, so that it can give out good offers for high value or high RFM customers. So this is what he tells. “We are going to look at the flow of the transaction, the size of the transaction and other interaction that we have had with each and every customer and we are going to make an educated guess about what we think that the customer will be worth in the future.”

That means that it is doing RFM analysis to calculate the CLV values of all the customers and then bucket the customers into various segments, so that they can target the customers who are in the high value segment and give them more promotional tactics or offers and you know retain them in the business.

So which is more important? We have three values with us which is recency, frequency and monetary value. If you can take a guess which is important among the three matrices, it is pretty difficult to decide because we do not have any data with us. So from many analysis that scholars have done around the world, they have you know tried to prove that recency is considered to be the most important of the three variables.

However, this might differ from business to business. Some business might prefer frequency over recency. Some might prefer monetary value over recency and frequency. So it depends on business. But generally speaking, recency is considered to be the most important of the three variables. So many direct mail marketers base their decisions mainly on recency.

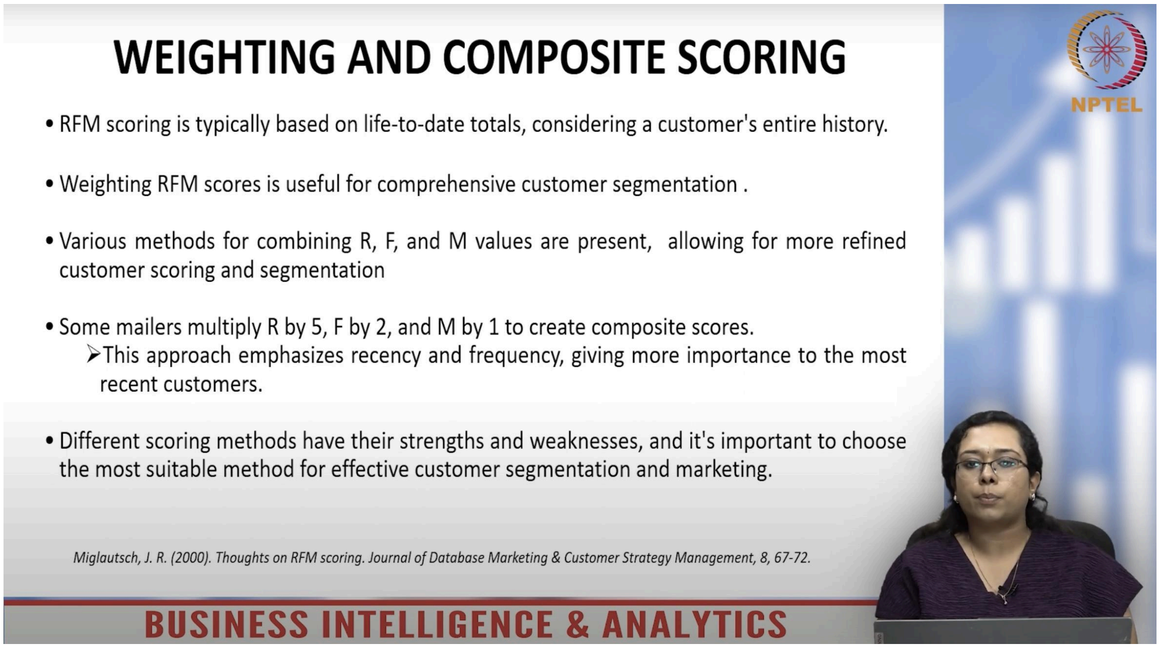


So customers can be divided into two groups, which is hot and dead. So you guessed it right, hot means customers who have high recency values or customers who have come recently to the business to purchase something and the dead group customers are those who have very less R score. So those are hot and dead customers. So recent customer behaviour is a strong indicator of the current level of engagement with the business. So the recency score or the R score is a very good indicator of how currently they are engaging with the business.

So if they have been a frequent customer before 5 years and they have also purchased high monetary value worth of items, but they are not at all recent. So they have come before 5 years and made purchases. They are not coming now. That means that the customer has turned off. So different strategies has to be formulated in order to bring back the customer into the system.

That strategy might not be the same as what we are doing with high RF and M customers. And customers who have interacted or made purchases recently are more likely to be actively interested and engaged with the products or services of a business. So that is what we discussed. So if the customer is recent, that means that even in future they might stay loyal.

That is because they have come recently and made a purchase. That means that they have not churned off. So this churn modelling is a very important topic that very big businesses do in order to retain the customers who have not yet churned and who stay loyal. And for those who have churned off they follow a different marketing strategy or you know customer segmentation strategy in order to engage those customers who have churned off. So this is taken from a paper which is 'Thoughts on RFM' scoring that was published in the journal of Database marketing and Customer strategy management. So they have told that recency is important among all the three matrices.

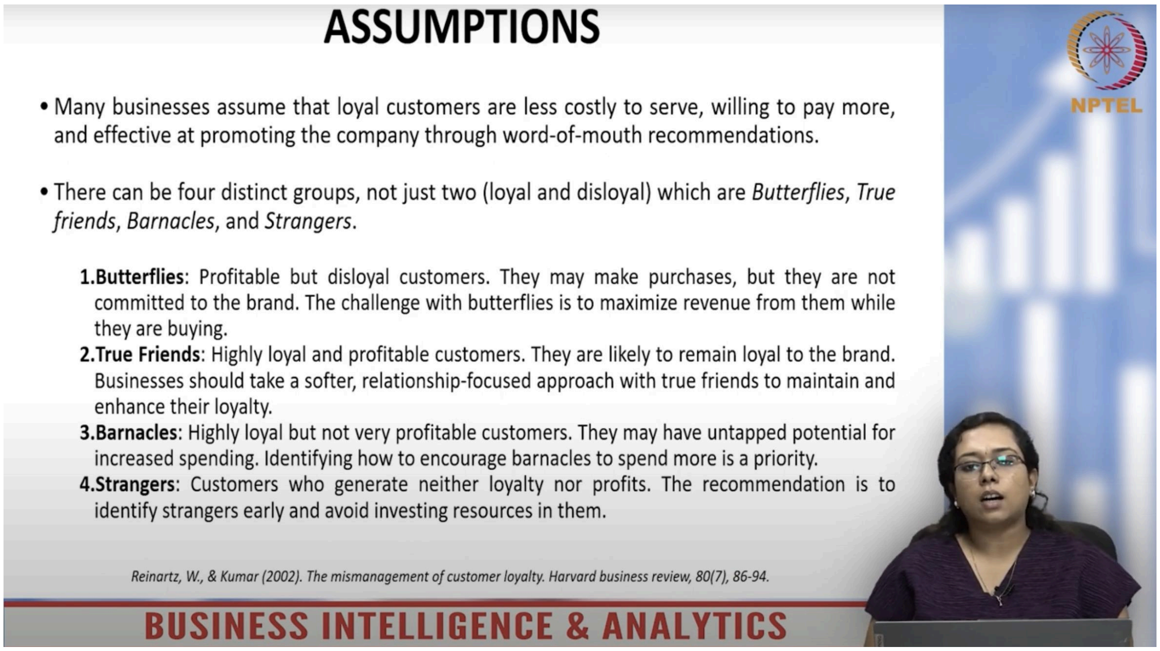


So we have discussed that recency is the most important matrix . So we will see how we weight and score these three matrices in order to come to a RFM total score. So this is weighting and composite scoring. So RFM scoring is typically done based on life to date totals, that is the over the customer's entire history of purchase with the particular business. So the customer's entire history is taken in order to come to a RFM score.

This might or might not be true because, you know some business do not want the customer's entire life history but just the past 5 years. So it depends from business to business on how they want to score the customers based on RFM values. So weighting the RFM values is useful for comprehensive customer segmentation.

So there are various methods on how we combine the R, F and M values. Some mailers may multiply R by 5, F by 2 and M by 1 to create a composite score. That means that the recency value is given high importance by the mailers. This approach emphasizes recency and frequency giving more importance to the most recent customers. That means that the monetary value of purchase by a particular customer is not given that much importance. Even if you have purchased for 5 rupees or 500 rupees, it does not make that much difference. They are giving more preference to, if you have been recent and frequent in the past.

So different, as I already told different scoring methods, they have different strengths and different weaknesses and ultimately it is the business's choice to choose which strategy works best for the customer segmentation and marketing. So this is also taken from the same paper.



So we have already seen how the weights are given to various matrices by the business. For example, some business might give more privilege or points to R, some might give more points to F and some may give more weights to monetary value.

For example, say a business which is running a fitness app. So a person who is dedicated to fitness would usually use the fitness app on a daily basis. So frequency might be given more weightage. But say an online retail platform, like Flipkart that might give more weightage to recency because if the customers are not recent that means that the customer has moved on to some other business. So it ultimately depends on business, on which R or F or M they should give more privilege.

Next we are going to see different customer buckets that we can segment the customers into. So many business assume that loyal customers are less costly to serve, willing to pay more and effective at promoting the company through word of mouth recommendation. For example, if you have liked some business so much that you are using it on a daily basis and also making almost every purchase through that app. For example, the daily purchases like, daily purchase businesses like Zepto or Blinkit or Swiggy Instamart, all these business you will be purchasing on a daily basis or alternate day basis. So it is very difficult if you have got accustomed to one business platform, then it is very difficult to move into another platform unless you have some dissatisfaction with it. For example, if I am a user of Zepto, then it is very difficult for me to change to Swiggy Instamart, unless Zepto starts pricing its products very high or does not give me any offer or I am dissatisfied with the customer service, that is when I churn off to a different company like Swiggy Instamart. So that is what happens.

So most business assume that loyal customers are the most high net worth customers because they need not spend a lot to retain those customers, because they are already there and they are already using the platform and they pay more on a daily basis and they can also promote the business by word of mouth communications to friends or family. So there can be, broadly classifying, the customers can be classified into four groups, not just loyal and disloyal customers but we can club these customers into four groups which we will be seeing on a live problem that we are going to solve at a later stage today.

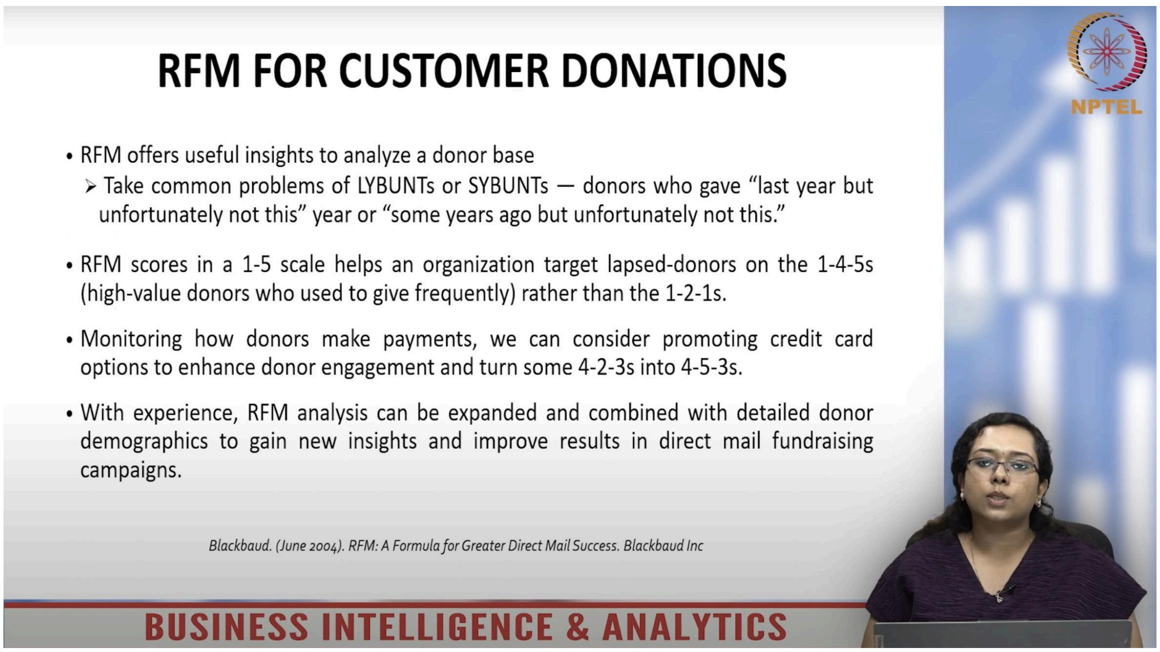
So the first customer segment is called butterflies. So how do butterflies behave? Butterflies are not very loyal, that means that we do not know when the butterflies would come but when they come it is very beautiful. So when they come, they make high net worth purchases but they are not committed to the brand. So they might come next or might not come next. So the challenge with butterflies is to maximize the revenue when they are coming, that is all. So they are very profitable but they are disloyal, they are not loyal to the brand at all. So whenever a butterfly is coming and making a purchase, you need to ensure that they are given with the right promotional strategies in order to retain them in the business.

Next is the true friends category. As we all know, true friends are loyal, they can be counted on with anything. So similarly, here also they are highly loyal and profitable customers, that means R, F and M value everything is high. That is why they are called true friends. So they are likely to remain loyal to the brand. So what businesses have to do is, they have to take a softer and relationship focused approach with true friends to maintain and enhance the loyalty.

And the third category is what is known as barnacles. So what are barnacles? Do you know what is a barnacle? It is an organism, it is an arthropod which sticks itself to boats or ships and then you know remains itself, remains there during the entire journey of the ship or boat. So does it do any good to the ship? No, it is actually dragging the ship with its weight and it is not allowing the ship to travel efficiently. So what it is doing is, it is pulling the ship down. That is the same connotation with which it has been used here also. So barnacles are highly loyal, because once it sticks to something, it does not tend to go anywhere but it sticks there itself. So it is highly loyal but it is not profitable, like the actual barnacles which does not bring any good to the ship. So it is not very profitable, but they may have untapped potential for increased spending. So we need to identify how to increase these barnacles to spend more which should be the priority of the business.

And the fourth and final customer segment is what is known as strangers. So customers who generate neither loyalty nor profits. So strangers are someone who has low R, low F and low M values. So the recommendation is to identify these strangers early and avoid investing resources in them. So mostly they are disloyal customers who do not make any good net worth purchases. So it is recommended that we do not spend our company resources on investing in this strangers category but instead focus our marketing strategy more on the first three customer segments.

So this has been taken, this classification has been taken from a paper which is a Harvard Business Review paper which was titled The Mismanagement of Customer Loyalty.



Next we are going to see RFM for customer donations. So not all businesses are alike. There are business which depend on donations to run themselves. So RFM analysis also offers useful insights to analyze a donor base, to determine how donors can be segmented in various categories. For example, there are two categories like Lybunts and Sybunts as you can see from the slide. So Lybunts are nothing but donors who gave last year but unfortunately not this year. And Sybunts, are some years ago but unfortunately not this. So Sybunts are not recent because they have given the donation some years ago but this year they are not given.

So there are various strategies on how to deal with such donors and RFM analysis gives useful insights on how to analyze these donor bases. For example, let us think that we are clubbing the customers based on RFM scores which are from 1 to 5, 1 meaning low R, low F and low M, 5 meaning it is the highest that a person can have. So if the recency value is 5, that means the customer has been most recent. If the frequency is 5, that means he is very frequent and if the monetary value is 5, that means he has made high net worth purchases. So that is what we have scaled from 1 to 5 and this RFM scores help an organization target the lapsed donors.

Lapsed donors are some donors that are not at all recent. So they are not coming recently to the business to donate. So we can target the donors on 1, 4, 5 rather than 1, 2, 1s. So what is 1, 4, 5? That means that recency is 1 in both but frequency is 4 in the first one. So the first donor is more frequent and is making a high net worth purchase because his monetary value is scaled to 5. So we can, you know prioritize the 1, 4, 5s rather than the 1, 2, 1s whose RFM values are just 1 for R, 2 for F and 1 for M.

So monitoring how the donors make payments we can consider promoting credit card options to engage donor engagements and turn some 4,2, 3s into 4, 5, 3s. So what are we changing here? The frequency value, frequency values being changed from 2 to 5, that is by giving them you know credit card options to pay we are making the customers or the donors highly frequent to the business. And with experience RFM analysis can be expanded and combined with detailed donor demographics to gain new insights. So RFM analysis need not be done in isolation, but it can also be combined with various demographic or psychographic data to obtain new insights and improve results so that it can be used for direct mail fundraising campaigns. So this has been taken from a paper that is ‘RFM-A formula for greater direct mail success’.