

Visualizing RFM Segmentation

Ron Kohavi
Amazon.com^(*)
1200 12th Ave South, Suite 1200
Seattle, WA 98144-2734
ronnyk@cs.stanford.edu

Rajesh Parekh
Blue Martini Software
2600 Campus Dr.
San Mateo, CA 94025
rparekh@bluemartini.com

ABSTRACT

Segmentation based on RFM (Recency, Frequency, and Monetary) has been used for over 50 years by direct marketers to target a subset of their customers, save mailing costs, and improve profits. RFM analysis is commonly performed using the Arthur Hughes method, which bins each of the three RFM attributes independently into five equal frequency bins. The resulting 125 cells are depicted in a tabular format or as bar graphs and analyzed by marketers, who determine the best cells (customer segments) to target. We propose an interactive visualization of RFM that helps marketers visualize and quickly identify important customer segments. Additionally, we show an integrated filtering approach that allows marketers to interactively explore the RFM segments in relation to other customer attributes, such as behavioral or demographic, to identify interesting sub-segments in the customer base. We depict these RFM visualizations on two large real-world data sets and discuss how customers have used these visualizations in practice to glean interesting insights from their data. Given the widespread use of RFM as a critical, and many times the only, segmentation tool, we believe that the proposed intuitive and interactive visualization will provide significant business value.

Keywords

Visualization, RFM (Recency, Frequency, Monetary) analysis, customer segmentation, database marketing.

1. INTRODUCTION

Direct marketers, especially catalog mailers, recognized long ago that customer segmentation improves profits [1]. One of the earliest segmentation techniques, called *RFM analysis*, has been in use for over 50 years. It is based on three simple customer attributes: Recency of purchase, Frequency of purchase, and Monetary value of purchase, hence the name **RFM**.

RFM analysis is based on the following simple observations that have been made over and over again across multiple industries:

1. Customers who purchased recently are likely to respond better to messages. They are also more likely to purchase again, compared to someone who has not purchased for a long time. For certain product families, where a purchase provides a rush of enthusiasm, this effect is very strong.

2. Frequent buyers, i.e., customers who purchase often, are more likely to buy again than infrequent buyers.
3. Big spenders often respond better than low spenders.

The three behavioral attributes, Recency, Frequency, and Monetary are extremely simple, in that they can be easily computed for any database that has purchase history, and are easy to comprehend, yet very powerful in their predictive ability. Arthur Hughes points out very nicely in the article "RFM: Is it 'Kudzu' or is it Gold" [3] that RFM analysis is easy to perform and results in a good segmentation. Unless very large mailings are done, the cost of developing specific models may be higher than the profits generated by a slightly better response rate. Some researchers claim that RFM based segmentation provides 75% to 85% of segmentation capabilities [2][4].

Our goal in this paper is to show how using some intuitive and interactive visualization methods can add significant value to the RFM analysis and simplify the marketers' task of identifying interesting customer segments based on RFM. We do not claim that RFM segmentation is better than segmentation built using richer models with more attributes; surely it is not. With additional attributes and good statistical software, better models can definitely be built. Further, when these richer models are being built, the RFM attributes can be thrown into the mix with the other attributes used for modeling. We have also found it useful to create RFM attributes for specific product categories and families, creating hundreds of attributes available for more complex modeling. For example, in addition to computing the customer's overall RFM, one can compute the Recency, Frequency, and Monetary spending for *Cosmetic* products.

The cost of the modeling effort increases dramatically when richer models are built using a variety of attributes. As an example, the following simple calculation shows that demographic data overlays need to improve the model significantly to provide a positive return-on-investment (ROI). A demographic overlay for a list with 1 million names will cost \$50,000 at a typical rate of 5 cents per person. If a company makes a 20% profit margin, the overlay must generate \$250,000 in incremental sales to recover the cost of the demographic overlay (20% of \$250,000 = \$50,000). Assuming the average purchase is \$100 and the targeted list is 10% of the size of the customer base (100,000 people from the one million), the demographic overlay needs to identify an additional 2,500 people, or provide an increase of 2.5% in the percentage of responders. This would be extremely hard when response rates are around 1-5% for targeted lists. Demographic overlays are therefore useful when very large lists are mailed to or when the same demographic information is used for many campaigns and the cost can be

^(*)The work on this paper was done while the author was at Blue Martini Software

amortized over time. This simple calculation ignores the additional cost of building more complex models, which would make it even harder to recoup the investment cost. In several applications RFM analysis is a critical, and at times the only, segmentation tool used. It is these application scenarios that will benefit most from the proposed visualization methods.

We propose a visualization that represents a RFM segmentation in a scatter plot using X, Y, and optionally Z axes to show the RF or RFM attributes, size to show segment size, and color to show response. The visualization summarizes the different RFM segments for marketers. Further, a visual filtering technique based on filter charts enables business users to interactively view the effect of other attributes on the RFM analysis and reveals interesting sub-segments in the customer base. We depict these visualizations on two real-world data sets BMSRFM1 and BMSRFM2 that will be made publicly available to enable fellow researchers to further experiment with this data and possibly generate exciting new insights.

The paper is organized as follows. Section 2 describes the two datasets we use in the paper and describes their characteristics. Section 3 presents an overview of the Arthur Hughes RFM analysis and identifies its limitations. Section 4 addresses our proposals for visualizing RFM segments. Section 5 discusses how these visualizations have been used in practice on customer datasets to glean interesting insights. Section 6 concludes with a summary and ideas for future work.

2. Real-World Data

We describe the two data sets (BMSRFM1 and BMSRFM2) and review their characteristics in detail, both to show the complexity of real data and to aid other researchers who might want to use the data in their experiments.

2.1 Data Overview

The two data sets (BMSRFM1 and BMSRFM2) are based on samples from two real-world datasets with differing characteristics. The datasets represent mostly Business to Consumer (B2C) orders, but a small percentage of the orders come from Businesses (B2B). The customer type (Business or Consumer) was not marked in the initial databases and hence cannot be differentiated reliably.

The sizes and duration of these datasets are shown in Table 1. Each original dataset was split into training, test, and future sets. For BMSRFM1 we reserved the two most recent years for future research since it spans a longer duration; for BMSRFM2 we reserved the most recent year for future use. This was done to avoid human overfitting of these dataset should they be used in future research. For each dataset, the last year not marked for “future use” is the test set, and the earlier years comprise the training set. The actual dates in these data sets have been converted to relative days and the dollar amounts have been perturbed using a simple linear transformation to further protect the identity of the original data sources.

Table 1: Dataset Sizes

	BMSRFM1	BMSRFM2
Approximate years	7	5
Number of customers	478,139	301,070
Number of purchases	1,099,005	1,628,884
Customers in training set	268,893	242,298
Customers in test set	41,668	54,604

2.2 Data characteristics

Figure 1 shows the number of purchases per customer, trimmed at 12. The graph accounts for 98.2% of the customers for BMSRFM1 and 91.4% of the customers for BMSRFM2. There are about a dozen customers in BMSRFM1 that purchased more than 200 times and a similar number of customers in BMSRFM2 that purchased more than 400 times. In both datasets customer identification is not perfect and some percentage of “new customers” are probably old customers that are not correctly recognized. Data warehouses are rarely perfect and this is one of the challenges in working with real data.

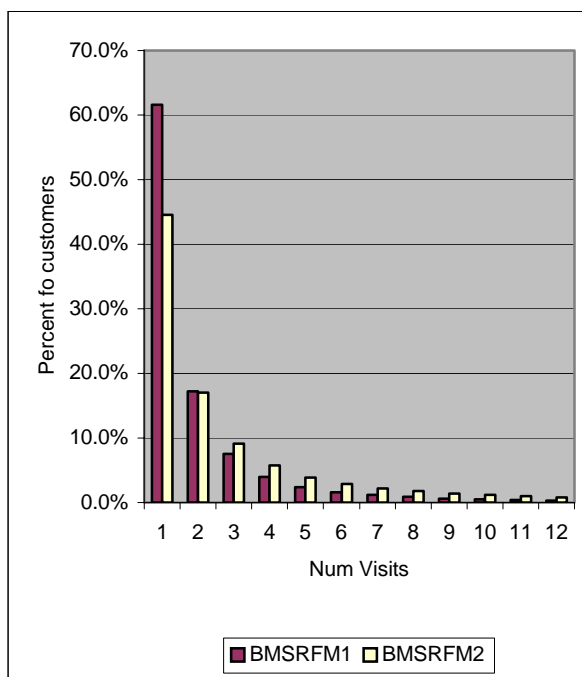


Figure 1: Number of purchases per customer

The two datasets differ in customer frequency and monetary spending over the years. Figure 2 and Figure 3 show customer frequency (average orders per customer per year) over time and the average order amount over time.

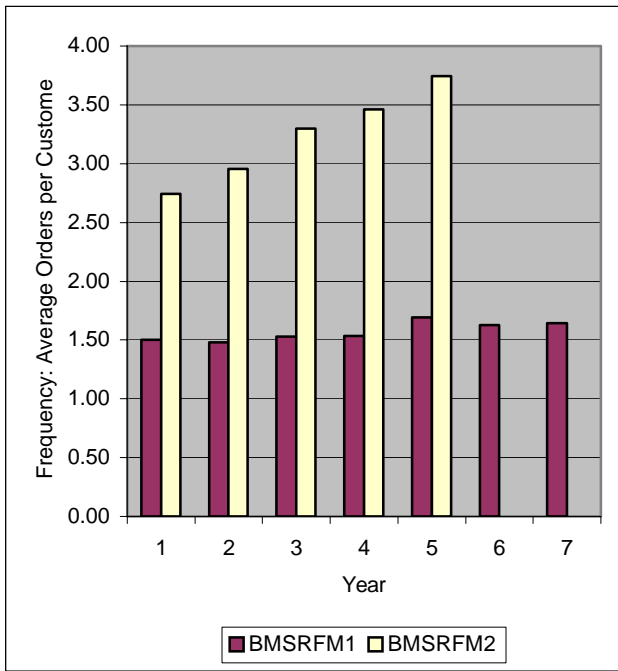


Figure 2: Average Orders Per Customer Per Year

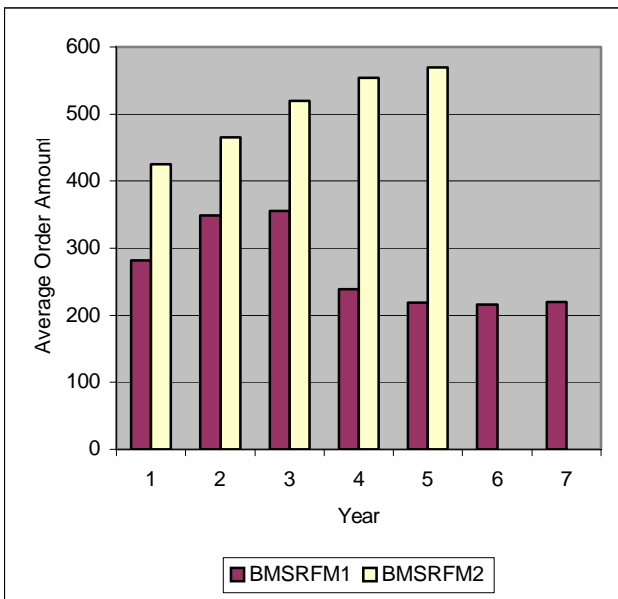


Figure 3: Average Order Amount Across Years

For BMSRFM1, the frequencies stay fairly constant and average order amounts vary over the years and decline slightly. For BMSRFM2, both frequency and average order amount increase. As with frequency per customer, the distribution of the average order amount is extremely skewed with some values that are over 50 times the mean, representing purchases by businesses and extremely wealthy individuals.

Figure 4 shows customer attrition over the first four years of the customers' tenure. The data is restricted to customers who bought for the first time over 4 years "ago" (prior to the end of the dataset) so that we could characterize attrition rates.

For both datasets, 60-70% of the customers who shop once never shop again. Customers that shopped for two or three years abandon at a lower rate of about 28%.

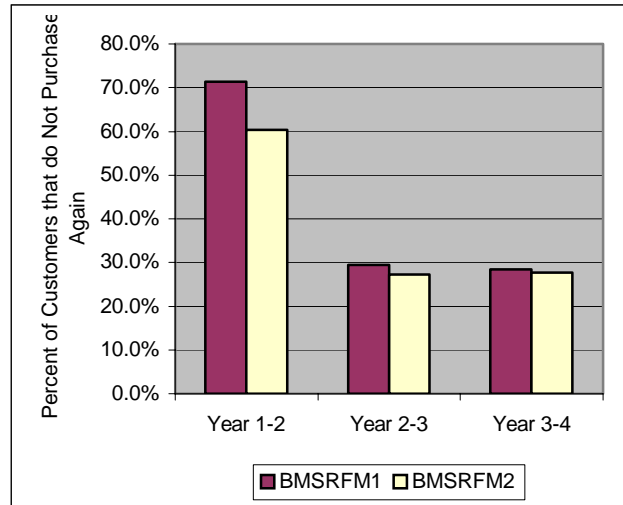


Figure 4: Customer attrition over first four years of the customer's tenure

3. Arthur Hughes' RFM Analysis

Arthur Hughes popularized RFM analysis by binning, or discretizing, the three RFM attributes into five ranges each, thus creating $5 \times 5 \times 5 = 125$ cells [4]. There is nothing magical about the number 5, but 125 cells create a reasonable set of segments to review by humans, yet a large enough set to provide significant lift for targeted marketing. Binning each range into 10 bins would create $10 \times 10 \times 10 = 1,000$ cells, which is likely to be too high for human analysis; binning each range into three bins only gives 27 segments.

There are several ways to bin attributes [5], and in this context equal-frequency binning is the most appropriate since the attributes have a very large range that is sparsely populated at the high end. Equal frequency binning creates cells that contain approximately the same number of customers. The bins cannot always contain the exact same number of customers because there are duplicate values for attributes. For example, if 19% of customers spent less than \$50 and 5% spent exactly \$50, then a right threshold of \$50.00 (right thresholds are defined here to be non-inclusive) will create a cell with 19% and a threshold of \$50.01 will create a bin with 24% of customers. Threshold-based binning will therefore usually create bins that are similar in the number of customers they contain, but not exactly the same.

This is usually a minor problem, but one exception to note is the case when we bin customer frequency. In this case, the bins vary significantly in the number of customers because a large number of customers have a frequency of 1 (i.e., the customers who have

purchased just once). As shown in Figure 1 and Figure 5 below, most customers have short tenure and therefore the most common frequency is one.

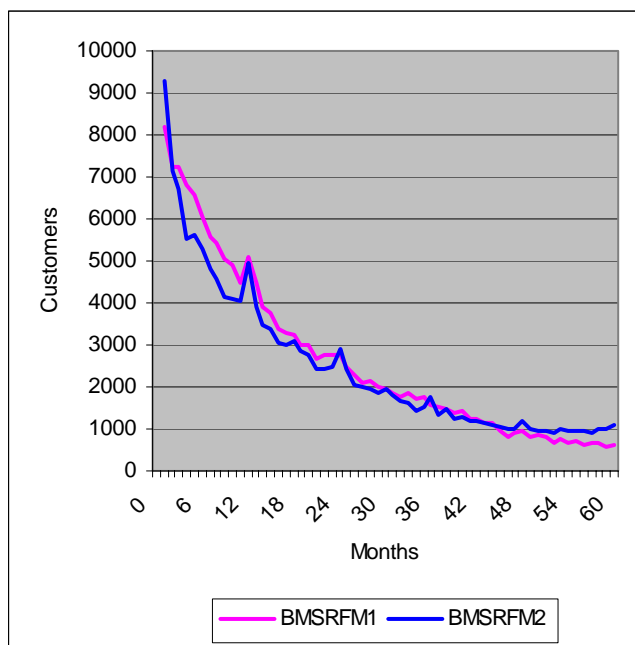


Figure 5: Customer tenure (first to last order); Customers with less than 1 month removed

Arthur Hughes suggests creating bins with exactly the same size (exact quintiles) by labeling some people with the same threshold in one bin and some in the other (based on record ordering). As we concentrate on visual representations of RFM, it is important to us to be able to say that a given bin represents frequency of, say, 2.0-2.5 times a year, something one cannot do with exact quintiles. Moreover, splitting the customers with the same frequency value (say 1) over 2-3 bins seems very misleading. Finally, we prefer the more stable approach of creating bins based on thresholds, so that the binning process is independent of the initial record ordering. The rest of this paper therefore uses threshold-based binning.

We begin with viewing the RFM attributes independently, as suggested by Arthur Hughes. From this point on, all graphs represent data from the two training sets and the response is estimated through the test set. Specifically, the business problem we are looking at is whether there was a purchase in the test set year.

For recency, we created five bins of approximately the same number of customers. The bins are numbered 1 to 5, with the best customers in bin 1 and the worst customers in bin 5. We note that this follows the numbering scheme described in [1], but it is the reverse of Arthur Hughes recommended numbering where 5 is the best cell and 1 is the worst cell [4]. We found Arthur Hughes' graphs, which range from 5 on the left side of the x-axis to 1 on the right side unconventional. Moreover, when plotting two variables in a scatter plot (discussed later in this paper), we felt

that most people would find (0,0) as the origin more intuitive than (5,5). Obviously, this is just a labeling standard and none of the results would change if either labeling scheme were used.

Figure 6 shows the distribution of responses. As expected, recency is an excellent predictor of response.

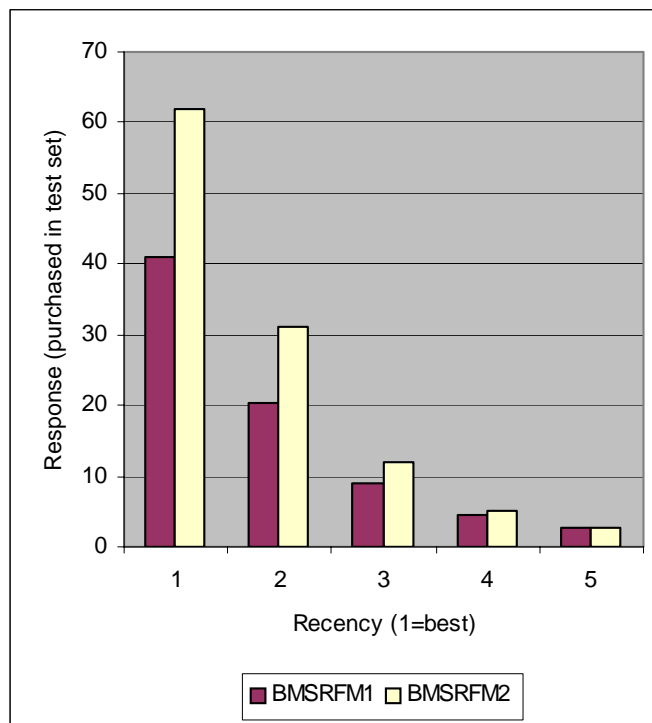


Figure 6: Recency Bins

Frequency, unlike recency, can be measured in multiple ways. Some advocate using the number of purchases over the history of the customer, but this creates a huge bias towards preferring high-tenured customers. Why should a customer who purchased a large number of items four years ago but nothing since be ranked high? Also, since most customers have purchased once, this leads to a very large bin at Frequency = 5. 67% of customers in the BMSRFM1 training set and 46% of customers in the BMSRFM2 training set have purchased only once. Similarly, defining frequency to be the number of purchases in the last year of the training set would create a very large bin of people with zero frequency. We therefore chose to use the average yearly frequency over the lifetime of the customer, i.e., how many purchases does the customer make in a year on average? Figure 7 shows the distribution.

While both datasets show a very similar effect for Recency and Frequency, Recency is the better of the two. Indeed, Arthur Hughes [4] wrote "frequency is also a good predictor of behavior, but much less so than recency... That is why RFM is RFM instead of FRM or FMR."

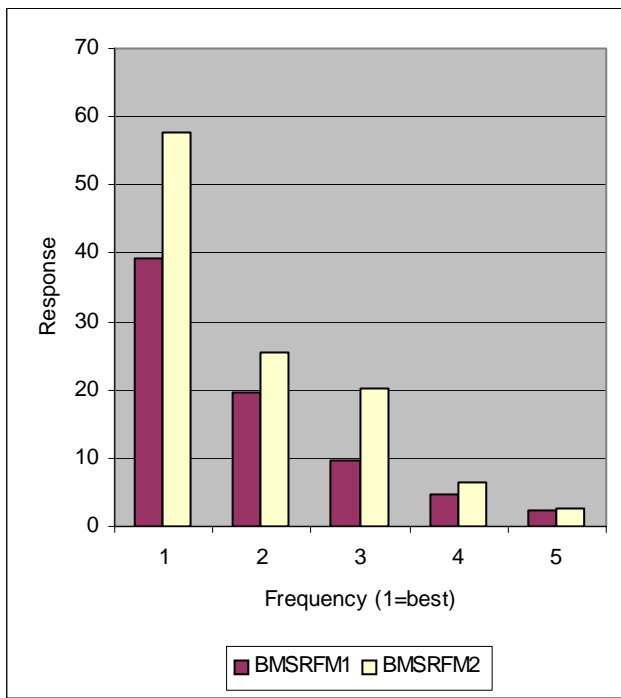


Figure 7: Frequency Bins

Monetary spending, like Frequency, can be defined based on average or total. We chose average with similar reasoning to the above discussion. The distribution is very similar to the distribution of Frequency and hence omitted.

In addition to viewing each attribute alone, Arthur Hughes recommends looking at the combined score created by joining Recency, Frequency, and Monetary as a 3-digit number ranging from 111 (best customers) to 555 (worse customers). Figure 8 and Figure 9 show the resulting graphs. This simple segmentation scheme works well on the BMSRFM1 and BMSRFM2 datasets. However, note the limitation 1 listed below.

Limitations of Arthur Hughes' RFM Analysis

The Arthur Hughes RFM analysis has the following limitations:

1. Reviewing 125 bins as in Figure 8 and Figure 9 shows that the ordering is reasonably good, but it gives very little insight about relationships, such as the response for Frequency=3 (every 25th bar).
2. If threshold-based binning is used (as we have done), then attribute interactions will create different-sized bins (or segments). Figure 8 only shows response, but it is possible (and often happens) that several segments are very small. Even if the response rate is high for a segment, it may not be worth targeting the segment if it contains a very small set of customers.
3. Knowing that someone is in a given segment does not provide his or her expected frequency. Specifically, if someone purchased twice last year, how likely is it that they will visit at least once this year? The Arthur

Hughes method provides a rank order, but not an expected value.

4. RFM is used to evaluate response, while monetary customer value may be much more important. Specifically, if two people purchase at the same frequency, yet one purchased last week and the other 3 months ago, the Arthur Hughes ranking will assign the latter person a lower Recency value and hence order it relatively low. However, that person could be spending 10 times as much per visit and their expected value could be much higher.

We address issue 1, 2, and 4 in the next section and leave issue 3 for future work.

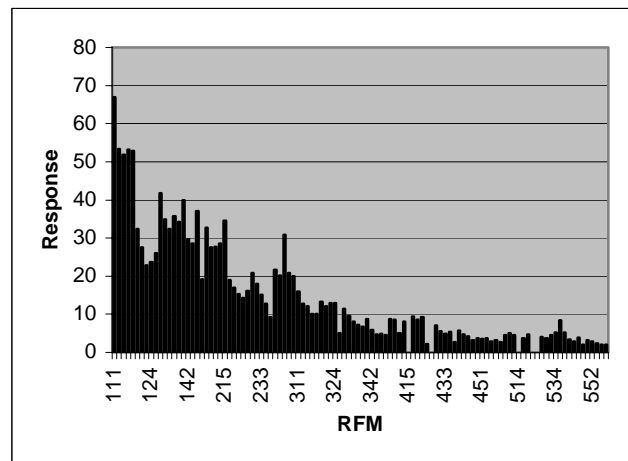


Figure 8: RFM Order for BMSRFM1

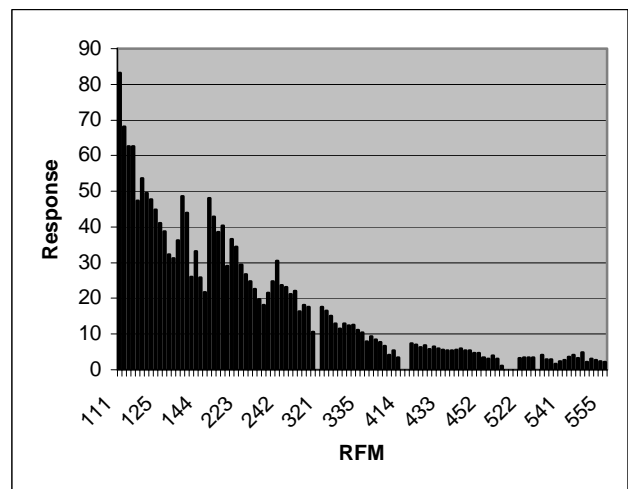


Figure 9: RFM Order for BMSRFM2

4. Visualizing RFM Segments

In this section we describe interactive visualizations for the Arthur Hughes RFM based on *scatter plots* and *filter charts* that alleviate three of the limitations mentioned above.

4.1 Computing RFM Scores and Response Spending

The RFM scores are computed for each customer based on their purchase history in the training set. We have two options for computing the *response spending*.

- The data is split into a training set containing data up to a year ago and a test set containing the last year. The total amount spent by the customer in the test set is assumed to be the *response spending*.
- A sample of the data from the full date range is used for modeling and a real campaign is sent out to the sample. The actual response is used to evaluate the segments. This is the better experiment, as it measures the response to the specific campaign, but it also costs more, both in terms of the money spent and in the time it takes to get the results.

In this paper, we compute *response spending* for the customer as the total amount spent by the customer in the test set. Note that we use monetary value instead of a Boolean (purchase/no purchase) to address issue 4 mentioned above. Looking at customer value for the next year is more important to the business than just whether a customer purchased or not. Segments within the customer base that have a high average response spending are interesting segments that should be targeted.

4.2 Segments based on RFM

The following charts are plotted using the BMSRFM2 data. The charts for BMSRFM1 are similar. Figure 10 depicts a bar chart for customer recency. The height of the bars represents the number of customers in the corresponding recency bin. As mentioned in section 3, the bins are of approximately the same size, which can be seen in the graph by the similar bar heights. Further, color is mapped to the average response spending¹. It is readily apparent that the more recent customers (Recency = 1) are the more attractive customers with average response spending of over \$2000. The less recent customers (Recency = 5) on the other hand have very low average response spending.

Figure 11 shows 2-d scatter plot of Recency versus Frequency. Recency and Frequency are mapped to the X and Y-axes respectively; the size of each square is mapped to the number of customers in that segment; and color is mapped to the average response spending. The more recent and more frequent purchasers (for example, Recency = 1 and 2, and Frequency = 1 and 2) tend to have a significantly better response than those who have shopped less recently or infrequently (for example, Recency = 4 and 5, and Frequency = 4 and 5).

One important aspect of this visualization is that it emphasizes the segment sizes due to attribute interactions, something that is

¹ We recommend that the charts in this paper be viewed in color in order to clearly understand the significance of the attribute that is mapped to color.

impossible to discern in bar graphs that are commonly used in RFM analysis. The squares are of different sizes indicating that is a different number of customers in each segment. For instance, there are no squares corresponding to Recency = 1 and Frequency = 5 and Recency = 2 and Frequency = 5 indicating that no customers fall in these segments. The square corresponding to Recency = 1 and Frequency = 1 is larger than its neighbors since it represents a larger segment of the customer base as compared to its neighbors. Marketers are mostly interested in large segments of the customer base with high response spending like the segment with Recency = 1 and Frequency = 1.

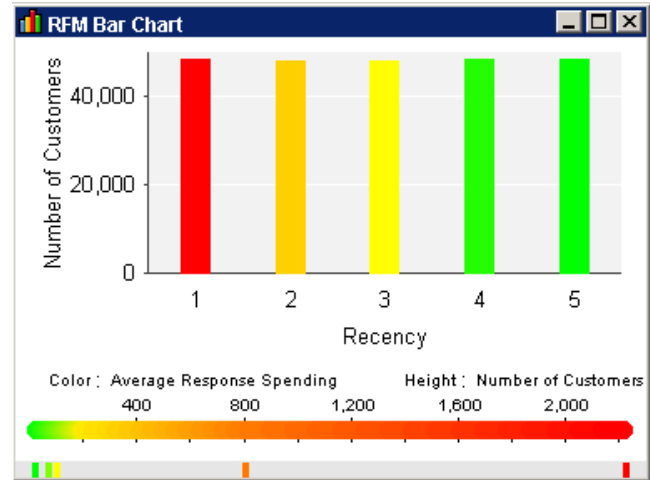


Figure 10: Recency by Average Response Spending

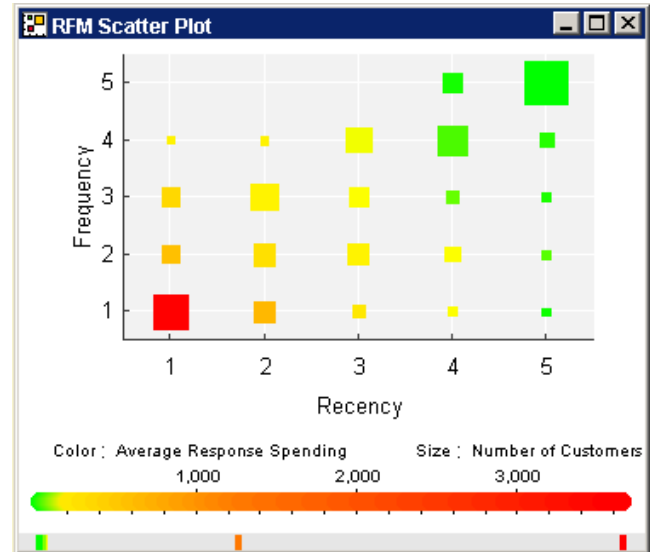


Figure 11: RFM Scatter Plot

A 3D *scatter plot* is used to depict the complete RFM (see Figure 12). Recency, Frequency, and Monetary are mapped to the X, Y, and Z-axes respectively; the size of each cube is mapped to the number of customers in that segment; and color is mapped to the average response spending. On screen rotation of this 3D chart facilitates viewing the individual cubes from different angles.

Figure 12 depicts segments of vastly different sizes and a full range of average response spending.

4.3 Sub-segments using other attributes

Marketers are often interested in visualizing the impact of demographic and behavioural attributes on RFM. A *Filter chart* is a valuable tool that allows the user to interactively pick different attributes to filter the data on and immediately see the impact of the selected filter settings. Figure 13 shows a sample filter chart with 4 attributes: *state*, *city*, *repeat purchaser*, and *tenure*. *State* and *city* are part of the dataset whereas *repeat purchaser* and *tenure* are derived from the *number of orders* placed by the customer and *order recency* respectively.

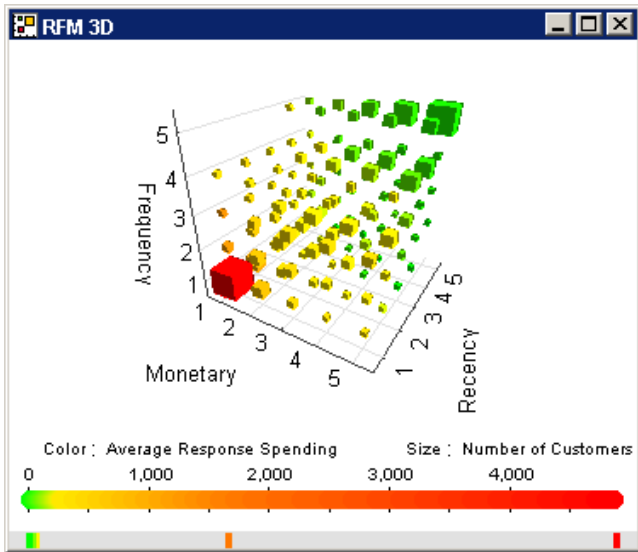


Figure 12: Complete RFM in 3D

Each attribute is depicted as a histogram with the height corresponding to the number of customers. The user can select one or more attribute values by clicking on the corresponding bars and apply these settings to see the effect on the RFM charts as depicted in Figure 14 and Figure 15. Values from multiple attributes can be selected by checking the ‘Fixed’ checkbox in one attribute and going on to the next to select other values.

Behind the scenes, a multidimensional cube is created for every combination of the attribute values in the filters and the 125 values of Recency, Frequency, and Monetary values. Suppose the marketers want to see the average response spending of high tenure customers (those customers whose first purchase was several years ago). Select the high value (3) in the *tenure* bar, and the scatter plot will change to show the RFM values only for those customers as shown in Figure 14. It is clear that a majority of the customers in the high tenure segment fall under the Recency=5 and Frequency=5 bin. This is due to the fact that there is a large attrition after the first year. The marketer might then ask, “How about the high tenure customers who have purchased more than once?” In this case, the settings *tenure=3* and *repeat purchaser=true* are selected from the filter chart. The resulting RFM chart is shown in Figure 15. This sub-segment of the customer base behaves much differently. There is a sizeable

group of customers who have a high frequency (Frequency=1) who also have a very high average response spending. This is the set of “loyal customers”—high tenure and also frequent purchasers.

Marketers can thus easily interact with the filter charts to select combinations of one or more attributes and instantly observe their impact on the RFM segments.

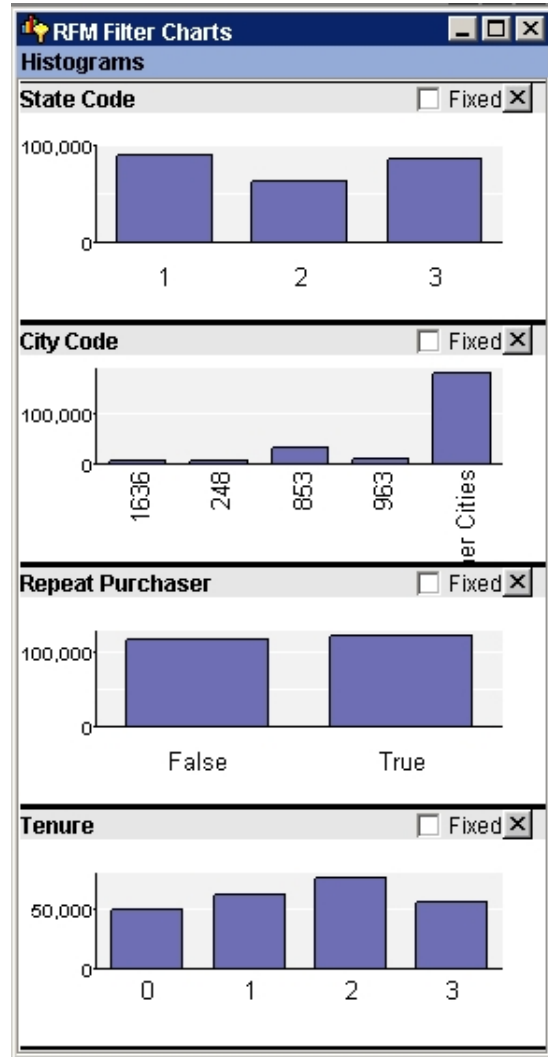


Figure 13: Filter Chart with 4 Attributes

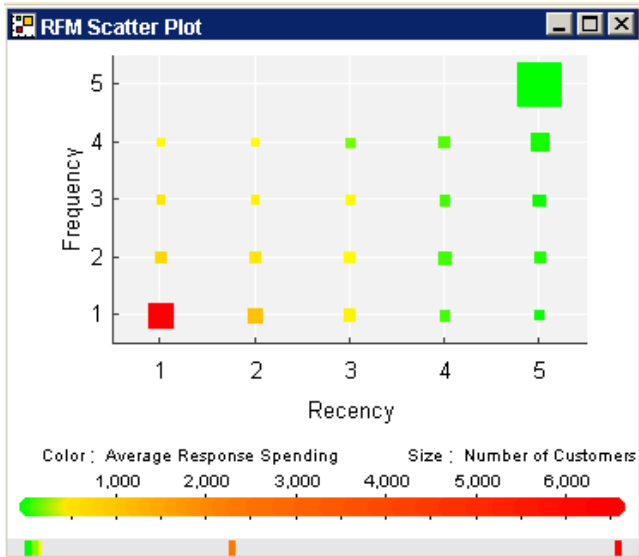


Figure 14: RFM for customers with high values of tenure

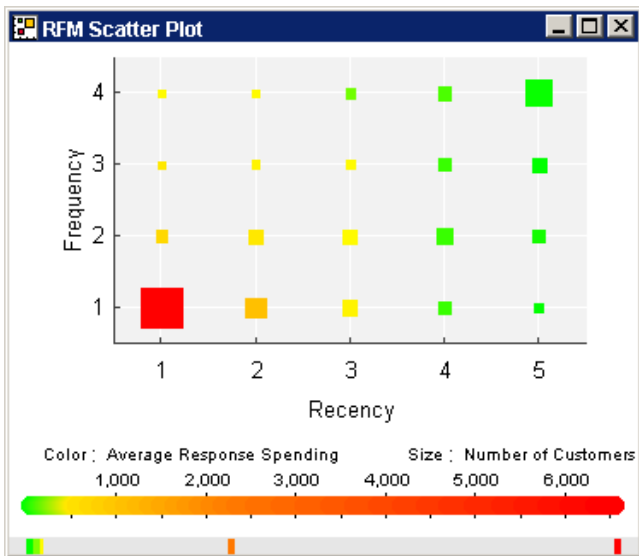


Figure 15: RFM for high tenure, repeat purchasers

5. Applying RFM Visualization

In this section we discuss how the RFM visualization methods described above were used to glean interesting insights from two Blue Martini retail e-commerce customer data sets, Debenhams (United Kingdom's fashion retail leader for over 200 years) and MEC (Canada's leading supplier of quality outdoor gear and clothing). Note that although the publication of these insights was approved (see [8] and [9]), the public release of these datasets was not. The two datasets BMSRFM1 and BMSRFM2 described earlier are not from these clients.

The RFM visualization on the Debenhams' data shown in Figure 16 and Figure 17 revealed a good-sized segment of Debenhams' customers who had high average spending but had not purchased

in the recent past. Using filter charts, the visualizations in Figure 16 and Figure 17 were restricted to customers who owned the Debenhams' loyalty card. The segment (Recency = 5 and Frequency = 5) denotes customers who have not purchased recently and not purchased very frequently. However, this is a good-sized with high average spending that can be targeted by a suitably designed marketing promotion.

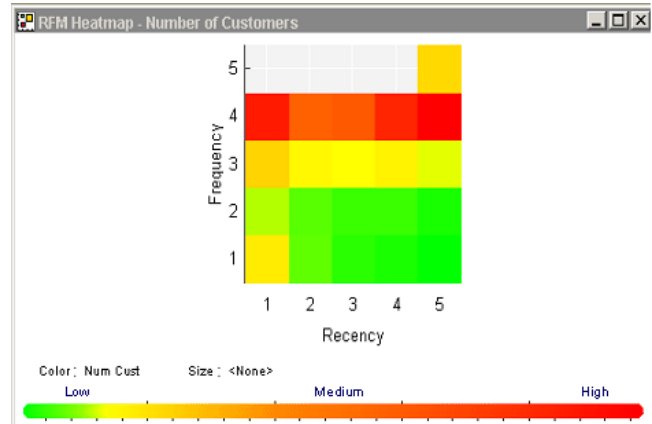


Figure 16: Recency, Frequency, and Number of Customers

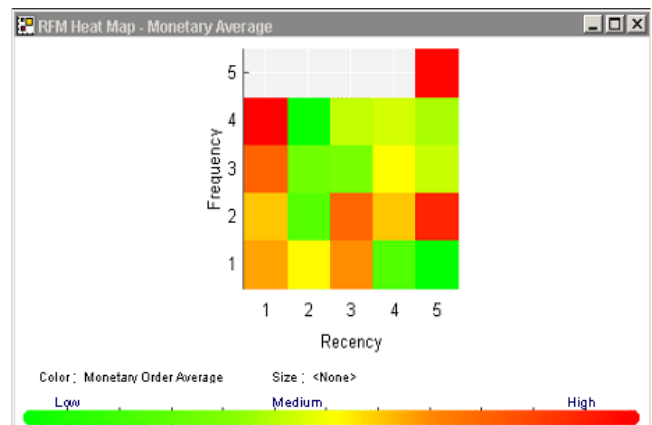


Figure 17: Recency, Frequency, and Average Spending

In the case of MEC, the RFM visualization was able to identify a segment of holiday buyers who have not purchased since the last holiday season. In Figure 18, the group of customers with recency = 4 corresponds to December holiday shoppers who have not shopped at MEC since then. Among these, the group of customers with frequency = 3 and 4 also happen to have a high average spending which makes them an attractive target for a marketing campaign to encourage them to shop again MEC.

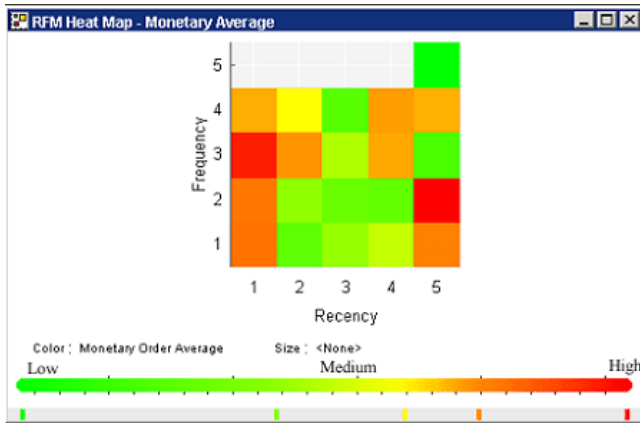


Figure 18: Holiday Shoppers at MEC (Recency = 4)

6. Summary and Future Work

In this paper, we have presented an interactive visualization of RFM that helps marketers visualize and quickly identify important customer segments. We showed how visualizations based on *filter charts* allow interactive exploration of RFM segments in relation to other customer attributes (such as behavioral or demographic) to identify interesting sub-segments in the customer base. While the visualizations may seem simple (which was our goal), they are not trivial (e.g., they cannot be created using commonly available tools such as Excel) and much of their power is derived from the interactivity. The proposed visualizations were shown on two large real-world data sets that will be made publicly available to enable fellow researchers to further experiment with this data and possibly generate exciting new insights. Applications of these visualization methods on two retail e-commerce customer data sets revealed interesting customer segments that could be targeted using appropriately designed marketing campaigns. Given the widespread use of RFM based segmentation in database marketing, we believe that the proposed interactive visualization methods will provide significant business value.

Several avenues are available for future work. Our visualization can be extended to help migration analysis [7]. It would be interesting to see how the proposed visualization methods can complement more sophisticated models built using RFM and other available attributes to predict customer response (both binary and expected value).

7. ACKNOWLEDGMENTS

The authors performed this work at Blue Martini Software, San Mateo, CA. We are grateful to our clients for sharing their data with us.

8. REFERENCES

- [1] David Shepard Associates. The New Direct Marketing: How to Implement a Profit-Driven Database Marketing Strategy, 3rd edition, 1998.
- [2] Renee Retka, Beyond Traditional Segmentation. Direct Marketing News, Nov 2001.
<http://www.customerman.com/Segmentation.htm>
- [3] Arthur M. Hughes. RFM: IS it “kudzu” or is it Gold?
<http://www.dbmarketing.com/articles/Art123.htm>
- [4] Arthur M. Hughes, Strategic Database Marketing, 2nd edition, 2000.
- [5] James Dougherty, Ron Kohavi, and Mehran Sahami. Supervised and unsupervised discretization of continuous features. In Armand Prieditis and Stuart Russell, editors, Machine Learning: Proceedings of the Twelfth International Conference, pages 194-202. Morgan Kaufmann Publishers, San Francisco, CA, July 1995.
- [6] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. Numerical Recipes in C (Second Edition). Cambridge University Press, 1992.
- [7] Jim Sellers and Arthur M. Hughes. RFM Migration Analysis: A New Approach to a Proven Technique.
<http://www.dbmarketing.com/articles/Art123.htm>
- [8] Blue Martini Case Study. Blue Martini Business Intelligence Delivers Unparalleled Insight into User Behavior at the Debenhams Website.
http://www.bluemartini.com/redirect/redirect_login_pdf.jsp?CONTENT%3C%3Ecnt_id=169201&FOLDER%3C%3Efolder_id=160845&bmUID=1063554761907
- [9] Blue Martini Case Study. Blue Martini Business Intelligence at Work: Charting the Terrains of MEC Website Data.
http://www.bluemartini.com/redirect/redirect_login_pdf.jsp?CONTENT%3C%3Ecnt_id=167181&FOLDER%3C%3Efolder_id=160845&bmUID=1063554761908