Customer Segmentation Using RFM Analysis: Realizing Through Python Implementation

Dr. Ahmad M. A. Zamil

Associate Professor,
Department of Marketing,
College of Business Administration,
Prince Sattam bin Abdulaziz University,
Saudi Arabia,

Dr. T. G. Vasista

Senior Consultant,
Vasista Consulting and Performing Services
OPC Pvt. Ltd Rajamahendravaram, India

Abstract

Purpose: The purpose of this research paper is to segment the best contributed retail management customers of a super store and corresponding least contributed customers for a given transaction period based on the customer purchase behavior.

Design/Methodology/Approach: Theoretically Market segmentation is determined based on geographical segmentation and RFM analysis is adopted to be implemented for obtaining customer segmentation. Data mining technique called clustering technique is used to implement RFM analysis in order to identify the highly profitable, high-valued customers and low-risk customers. The secondary file containing secondary data is obtained from the source web site called Github.

Findings:(i) Segmenting customers based on their geographical residence is easy to work using Python (ii) segmenting customers based on their purchase behavior patterns of grocery items is easy to find using Python programming kind of innovative information technologies with visualization. Though there are prior researches conducted using Python libraries, however our approach is an effort of reducing the gap between the theoretically coding the RFM model and its close implementation in Python based on the same RFM coding.

Research Limitations/Implications: The development of modern patterns of grocery stores data analysis is limited to the case of Britain and has been described and analyzed through Python programming implementation. Though it is not possible to obtain patterns for all countries globally, however it is possible to obtain such patterns of some of the leading countries like UAE, Saudi Arabia and some countries from Europe etc.

Practical Implications: Customer Segmentation through RFM analysis technique implementation in Python provides an opportunity to analyze not only the Britain region but also for different countries is possible. RFM analysis technique that is mentioned theoretically can be well programmed with possible visualization using Python coding and programming libraries. But one must be familiar with having awareness of how to code in Python programming in order to work towards obtaining geographical grouping and customer grouping using visualization techniques through graphs and charts.

Social Implications: It helps grocery stores management in identifying profitable customers and thus can campaign various customer

profitable offerings as a part of loyalty programs. Also, it is possible to focus on planning for customer retention programs.

Originality/Value: The value of this research paper lies in realizing through mapping the theoretical value of RFM analysis to practically visualizing the customer segmentation in the form of grouping best customers and low-risk customers using Python implementation. Such value can be found in following the research methodology steps while developing the content of the paper as well as showing practically obtaining results through dealing with data-based case study as a part of serving itself as a testing the proposed hypotheses.

Keywords: Behavioral Segmentation, Clustering, Customer Segmentation, Electronic Customer Relationship Management, Geographical Customer Segmentation, Market Segmentation, RFM Analysis.

Introduction

Globalization has major impact on the community cultures. It is slowly reflecting by merging the tastes and preferences across the globe. Therefore, future businesses require following a non-traditional i.e. innovative approach on orienting their strategic decision by adopting innovative information technologies in serving customers (Subramaniam, 2016). Customer service is an opportunity for segmenting markets (Sharma and Lambert, 1994). As the target marketing has been becoming an important strategy to achieve competitive advantage, customer segmentation problem for retail companies has been studied in many papers as cited in passing to Bilgic, Kantardzie & Cakir (2015).

As cited in Martin (2011) in passing, "market is a set of products or services"; "a market is a group of buyers and group of sellers that they serve" (Kotler & Armstrong, 1996).

Organizations should know their marketing size. Having knowledge on customer behavior can help marketing managers re-evaluate their strategies with the customers. It helps in improving and expanding the effective marketing strategies (Hosseini & Shabani, 2015). It is important to consider customer market segmentation before assessing customer loyalty.

Segmentation is "the classification of consumers within market that share related needs and similar purchasing behavioral habits" (Kotler, 2010). Customer segmentation is "the process of dividing customers into groups with similar characteristics or features" (Song & Kim, 2011).

Businesses put effort in understanding the rationale behind market segmentation. The rationale is to focus on the consumer behaviors and purchasing patterns. When effective marketing is done, organizations can achieve huge return on investment (ROI) for the expenses incurred on marketing and sales. Marketing segmentation can be done in multiple ways instead of following a single way

(Martin, 2011). As cited in Hosseini & Shabani (2015) current businesses are adopting innovative techniques for developing new strategies and for the integration of Customer Relationship Management and its analytics (Bose & Chen, 2009). These new innovative methods and techniques include: (i) association rules, clustering algorithms, decision trees and genetic algorithms (Berson et al, 1999; Turban et al, 2008).

Literature Review

Market Segmentation

Market segmentation is one of the most basic and important concepts of strategic marketing. The goal of customer segmentation is to identify customer groups, where in customers are greatly differentiated from customers in other segments. It is the process of dividing the customers into homogeneous sub-groups such as potential customers and repeated customers for the purpose of target marketing. Market segmentation is the process of categorical division of the population of possible customers into distinct groups. The customers within the same categorical segment share common characteristics that can help a firm in targeting customers and marketing to them effectively (marketbusinessnews.com, 2020) (Katti, 2015) (lovelock & Wirtz, 2011).

Customer Segmentation

The segmentation process helps in conducting analysis on not only on customers' needs and shopping habits but also helps in taking decisions on analyzing changing market conditions and competitions (Bilgic, Kantardzic & Cakir, 2015).

Customer segmentation is virtually a potential tool to guide firms towards effective ways of marketing products and helps in developing new ways of realizing consumer behavior (Cooil, Aksoy & Keiningham, 2008). Further Kotler & Armstrong (1999) defined market segmentation and its types as "dividing a market into distinct groups of customers, with different needs, characteristics or

behavior, who might require separate products or who may respond differently to various combinations of marketing efforts" (Ahmad, 2003) (Aksoy, Keiningham & Bejou, 2014).It includes (i) geographical (ii) demographical (iii) behavioral and (iv) psycho graphical based segmentations. Customer segmentation improves profits through direct marketing where in catalogs, mailers etc are used (David Shepard Associates, 1998).

Geographical Customer Segmentation

Location is a transcendental decision to ensure the viability of retail stores. Geo-demographic segmentation can be considered as an analytical tool for optimizing retail location strategy. Geo-demographic segmentation seeks to capture the spatial heterogeneity of the urban market areas in terms of the characteristics of the residents (Gonzlez-Benito & Gonzlez-Benito, 2005). Geographical customer segmentation is basically segmenting customers based geographical areas to which customers are belonging to (Sulekha & Mor, 2014). It is believed that consumers who live in the same region share similar wants and needs and are different from the consumers who live in other regions of the world (Martin, 2011). In other words, the customers are divided based on geographical area units such as villages, towns, cities, states, countries and regions. Store wise segmentation divides a network of stores into meaningful groups. Geographical categorization of customers is possible with data base techniques by employing a set of pre-categorized examples to develop categorized population of records.

Previous research on geographical customer segmentation using historic customer and customer sale transaction data and filtering the data by customer location hierarchy (Oracle docs, undated; Campbell, 2015) was seen much. In this research database filtering approach is used to filter the customer transaction data in order to obtain the location hierarchy such as country. Filtering customers based on customer country has been applied.

Behavioral Customer Segmentation

Customer segmentation is a marketing strategy involving the customer division into various groups based on their underlying characteristics, needs and interests. It then opts for designing and implementing marketing strategies to target them. One of the most types of segmentation approaches is behavioral segmentation analysis (Huseynov & Ozkan, 2017).

In behavioral segmentation, customer purchase behavior needs are acknowledged not only in specific needs of products but also the interactions made among the whole range of products. This segmentation is done based on

purchase patterns. These purchasing patterns are interpreted by purchasing portfolios, which include list of categories of customer purchases and their consumption behaviors on these categories (Raje & Srivastava, 2014). The rationality for categorization is the consideration of factors such as percentage of frequency and percentage of monetary of all categories in range. It uses the clustering analysis by K-Mean technique (Do, 2011). Susilo (2016) conducted research on behavioral segmentation and price on customer value influenced for increasing the purchase that impacts the customer loyalty. The data is related to Jakarta. Huseynov & Ozkan (2017) have conducted the research study on behavioral segmentation analysis of online consumer audience in Turkey by using real-time e-commerce transaction data (Huseynov & Yildirim, 2017). Previous research in customer segmentation based on behavioral data in e-market place used Python programming language to conduct cluster analysis and visualization (Aziz, 2017). Further research on customer segmentation based on user behavior analysis with RFM model and data mining technique is conducted by Tavakoli, Molavi, Masoumi, Mobini, Etemad & Rahmani (2018), these authors used R programming for performing the exploratory analysis and the RFM analysis is done using Python. Kamthania, Pahwa & Madhavan (2018), Yoseph, Hashimah, Malim & Almalaily (2019); these people have visually shown the high valued and low valued customers. Though there are few researches that used Python for RFM analysis, our approach followed an approach that reduces the gap between theoretical RFM coding and practically implementing such RFM coding based filtering data using clustering technique of data mining on an existing e-commerce transaction data of a retail management.

Customer benefit from behavioral segmentation by (i) optimizing their market spending (ii) increasing the customer lifetime value (CLV) (iii) improving customer service and customer experience (iv) implementing optimal marketing channels selection for their each segment (v) improving product features and offerings (vi) identifying and cater to most profitable customers (Rivas, 2019).

Research Objective

Target marketing is a method of focusing and attracting customers who are likely to buy a product. Data mining techniques can be used to identify the highly profitable, high-valued and low-risk customers (Pandey, 2017).

Clustering is a data mining technique. It is used for geographical customer segmentation and consumer behavior-based customer segmentation, in view of implementing target marketing strategy. In this research

the goal is: (i) how to implement customer segmentation of stores, based on multiple data sources and (ii) how to create marketing strategies for each segment.

In this paper only (i) and (iii) consumer marketing techniques have been focused to be described and discussed to be implementation using innovative tools and languages like Python to realize the concept of consumer market segmentation.

Research Hypotheses

Null Hypothesis

H01 = Segmenting customers based on their geographical residence is easy to work using Python programming kind of innovative information technologies

H02 = Segmenting customers based on their purchase behavior patterns of grocery items is easy to find using Python programming kind of innovative information technologies.

Research Methodology

Contemporarily technology based online retailing has reshaped the retail landscape. Shopping in Grocery stores is emerging as one of the fastest growing categories of online retailing in the United Kingdom (Davies, Dolega & Arribas-Bel, 2019). The development of the modern patterns of grocery stores in Britain has been described and analyzed as a part of case study (Guy, 1996).

This research uses a case study approach. It uses qualitative case study methodology to enable researchers in conducting deep exploration of complex phenomenon within the specific domain context (Rashid, Rashid, Warraich, Sabir and Waseem, 2009). A case study approach is adopted for the investigation, when the realization through theoretical aspects becomes limited with operational activities and synergies. According to Miles & Huberman (1994) a case study is a "phenomenon of some sort of occurring in a bounded context" as cited in Knoll (2008) in passing. Case studies are an established research approach for exploring complex phenomenon from such domains where the extracting insights become hazy (Lovas & Goshal, 2000).

Due to the sensitive nature of primary data gathering, companies generally do not provide the primary data. As a result, the required data is gathered from web sites in the form of a secondary data, in order to process them in python programming and to realize the result of data mining techniques (Schrantz, 2013). The secondary data is collected during the period of May-June 2020 and the research is conducted during May-July 2020.

The methodology proposed in this paper contributes to the electronic marketing literature by exploring analytical Customer Relationship Management perspectives in the retail management of grocery stores. This methodology aims to make an attempt to identify customers and also segmenting them using clustering technique of data mining through a case study as an exploratory method. It uses secondary data available from web sites sources such as Kaggle.com and Github.com etc. Clustering analysis is mostly used as a data mining technique. It maps data items to more unknown similar item groups (Oliverira, 2012).

Realizing Geographical Segmentation in Python

Geographic segmentation variables can include location hierarchy, city, state, country, population density, economic status, zip code and regional climate (Oracle docs, undated). For example, if ecommerce ships internationally, user interface may want drill down behaviors, conversations and purchase patterns from a particular geographic region or country or the location hierarchy (Campbell, 2015).

For Python Implementation of Geographical Segmentation see Figure-1 to Figure-4.

Figure-1: Declaring Python Libraries and Loading Data file in Python

```
In [178]: import pandas as pd
          import numpy as np
          import datetime as dt
          import warnings
          warnings.filterwarnings('ignore')
          import matplotlib.pyplot as plt
          import seaborn as sns
          color = sns.color_palette()
```

Loading Datafile in Python

```
In [179]: df = pd.read_excel("OnlineRetail4.xlsx")
          df.head()
```

Out[179]:

	InvoiceNo	StockCode	lower	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country
0	536365	85123A	white hanging heart t-light holder	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0	United Kingdom
1	536365	71053	white metal lantern	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
2	536365	84406B	cream cupid hearts coat hanger	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0	United Kingdom
3	536365	84029G	knitted union flag hot water bottle	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom
4	536365	84029E	red woolly hottie white heart.	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0	United Kingdom

Figure-2: Listing unique geographical nations and its corresponding Customer IDs

```
In [180]: df1=df
          df1.Country.nunique()
Out[180]: 38

In [181]: df1.Country.unique()
Out[181]: array(['United Kingdom', 'France', 'Australia', 'Netherlands', 'Germany',
                 'Norway', 'EIRE', 'Switzerland', 'Spain', 'Poland', 'Portugal',
                 'Italy', 'Belgium', 'Lithuania', 'Japan', 'Iceland',
                 'Channel Islands', 'Denmark', 'Cyprus', 'Sweden', 'Austria',
                 'Israel', 'Finland', 'Bahrain', 'Greece', 'Hong Kong', 'Singapore',
                 'Lebanon', 'United Arab Emirates', 'Saudi Arabia',
                 'Czech Republic', 'Canada', 'Unspecified', 'Brazil', 'USA',
                 'European Community', 'Malta', 'RSA'], dtype=object)

In [182]: customer_country = df1[['Country','CustomerID']].drop_duplicates()

In [183]: customer_country.head()
```

Out[183]:

	Country	CustomerID
0	United Kingdom	17850.0
9	United Kingdom	13047.0
26	France	12583.0
46	United Kingdom	13748.0
65	United Kingdom	15100.0

Figure-3: Listing country-wise customers sorting by Customer ID

```
In [184]: ccgc=customer_country.groupby(['Country']).sum()
          ccgc ['CustomerID'] = ccgc['CustomerID'].astype(int)
          ccgc.head()

Out[184]:              CustomerID
            Country
            Australia    115614
            Austria      137631
            Bahrain       24708
            Belgium      311123
            Brazil        12769

In [185]: customer_country.groupby(['Country']).sum().reset_index().sort_values('CustomerID', ascending=False)
          customer_country.head()

Out[185]:      Country        CustomerID
            0   United Kingdom    17850.0
            9   United Kingdom    13047.0
           26          France    12583.0
           46   United Kingdom    13748.0
           65   United Kingdom    15100.0

          Plotting the count of customers
```
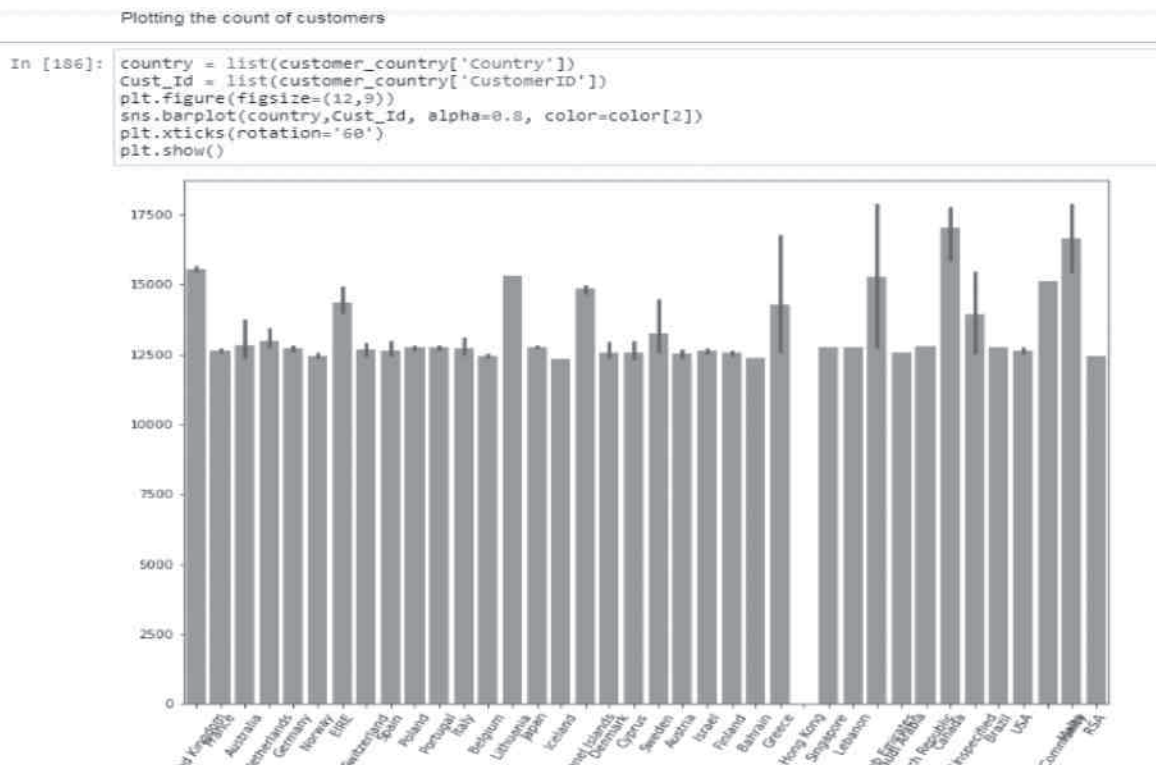
Figure-4. Listing country wise customers sorting by Customer ID (Geographical Segmentation)

```
          Plotting the count of customers

In [186]: country = list(customer_country['Country'])
          Cust_Id = list(customer_country['CustomerID'])
          plt.figure(figsize=(12,9))
          sns.barplot(country,Cust_Id, alpha=0.8, color=color[2])
          plt.xticks(rotation='60')
          plt.show()
```

Behavioral Segmentation

For companies it is becoming mandatory to better understand the customers' data in order to find similarities and differences among customers as well as to predict their behavior. Thus, customer segmentation according to their data is becoming vital (Dogan, Aycin & Bulut, 2018).

In behavioral segmentation, the variables normally include a sub-segment of consumer segmentation. Knowledge on buyer behavior analysis helps in understanding, customer commitment and pricing plans as well as to develop market positions (Martin, 2011).

Recency, Frequency and Monetary (RFM) analysis is widely applied in many practical areas like direct marketing (Wei, Lin & Wu, 2010). Further, RFM model allows decision makers in identifying valuable customers (Wei, Lin & Wu, 2010). RFM model serves as an effective marketing strategy as well as considered as an effective method of segmenting. It is a behavioral analysis performed for doing market segmentation. RFM method obtains homogeneous cluster groups through conducting behavioral analysis of customers. It improves the target marketing segmentation by examining the recency, frequency and money spent in grocery items of a retail stores. This paper summarizes case study example for the customers who had bought most recently, most frequently and had spent more money as well as who contributed least in a specified period for the geographical area of UK (Ait Daoud, Bouikhalene, Amine & Lbibb, 2015).

Clustering is used for market segmentation. Clustering is the most important data mining technique used in customer relationship management and marketing domains. Clustering would use customer purchase transaction data to track buying behavior and patterns; and creates strategic business initiatives (Rajagopal, 2011). While tools of data mining categorize automatically by looking for the best-fit simplification of data set, clustering analysis helps in revealing a statistically significant correlation between data items for e.g. location and buying habits in a customer database (Katsaras, Wolfsons, Kinsey & Senauer, 2001).

For Python Implementation of behavioral segmentation see Figure 5 – Figure 11.

Finding Best Contributing Customers Using RFM Matrix Principle

The RFM analysis is the most frequently adopted direct marketing segmentation technique. It comprises of three measures viz. Recency, Frequency and Monetary (Wei, Lin and Wu, 2010). It is fast and simple as its original purpose is to quantify customer behavior (Sokol& Holy, 2020).

Recency is commonly defined as "the number of periods since the last purchase". It measures the time gap between the most recent transaction time and the analyzing time. Frequency is defined as the "two states of purchase such as single purchase and repeated purchase" (Wei, Lin & Wu, 2010). If the frequency score is high, it indicates greater customer loyalty. Monetary is defined as "the purchase value that the customer spends in this period" (Wei, Lin & Wu, 2010). RFM cell covers five equal quintiles with 20% groups. All customers are presented in the form of codes as 555, 554, 553….111, which create 5x5x5 = 125 cells. Thus, the best customer segment is denoted by 555 and the worst customer segment is denoted by 111. Based on RFM scores, customers can be grouped into different segments and correspondingly their profitability is analyzed further. RFM scoring method sorts customers in descending order from best contributing to least contributing. RFM model is applied to categorize customers into high-contribution loyal customers, low-contribution loyal customers, uncertain customers, high-spending lost customers and low-spending lost customers. RFM analysis can also use to calculate lifetime value (CLV) in addition to analyzing customer profitability.

Python Implementation of RFM Matrix Analysis

Figure-5: Listing UK Customers IDs

```
from datetime import date date_format = "%m/%d/%Y" today = date.today() print(today) print(maxdate) print(mindate)

RFM Segmentation and Creating RFM Table
```

```
In [ ]:   Recency is calculated for a point in time and the last invoice data is
          2011-12-09, so we use 2011-12-10 for calculating the recency
```

```
In [196]:  df1 = df1.loc[df1['Country']=='United Kingdom']
           Cust_Id_UK = df1[['CustomerID']].drop_duplicates()
```

```
In [197]:  Cust_Id_UK.head()
```

Out[197]:

	CustomerID
0	17850.0
9	13047.0
46	13748.0
65	15100.0
82	15291.0

```
Create RFM table

import datetime NOW = datetime.datetime.now() rfmTable = df1.groupby('CustomerID').agg({'InvoiceDate': lambda x: (NOW - x.max()).days, 'InvoiceNo':
lambda x: len(x), 'TotalPrice': lambda x: x.sum()})
```

Figure-6: Creating RFM table

```
Create RFM table

import datetime NOW = datetime.datetime.now() rfmTable = df1.groupby('CustomerID').agg({'InvoiceDate': lambda x: (NOW - x.max()).days, 'InvoiceNo':
lambda x: len(x), 'TotalPrice': lambda x: x.sum()})
```

```
In [219]:  import datetime
           NOW = datetime.datetime.now()
           rfmTable = df1.groupby('CustomerID').agg({'InvoiceDate': lambda x:
            (x.max() - x.min()).days, 'InvoiceNo': lambda x: len(x), 'TotalPrice': lambda x: x.sum()})
```

```
In [220]:  rfmTable

           segmented_rfm = rfmTable.rename(columns={'InvoiceDate': 'recency',
                                 'InvoiceNo': 'frequency',
                                 'TotalPrice': 'monetary_value'})
           segmented_rfm
```

Out[220]:

CustomerID	recency	frequency	monetary_value
12346.0	0	1	77183.60
12747.0	366	103	4196.01
12748.0	372	4596	33719.73
12749.0	209	199	4090.88
12820.0	323	59	942.34

Figure-7: Coding RFM (RFM Scores)

```
In [ ]:  Segmenting gives a starting point for the detailied analysis
         5 segments are easy to understand and explain
```

Best Customers will have values with low recency, highest frequeny and higest monetary_amount

Add segment numbers rfm Table

```
In [221]:  def RScore_compare(r):
               if (r <= 74.6):
                   RScore = 1
               elif (r > 74.6 and r <= 149.2):
                   RScore = 2
               elif (r > 149.2 and r <= 223.8):
                   RScore = 3
               elif (r > 223.8 and r <= 298.4):
                   RScore = 4
               else:
                   RScore = 5
               return RScore

           def FMScore_compare(r):
               if (r <= 74.6):
                   FMScore = 5
               elif (r > 74.6 and r <= 149.2):
                   FMScore = 4
               elif (r > 149.2 and r <= 223.8):
                   FMScore = 3
               elif (r > 223.8 and r <= 298.4):
                   FMScore = 2
               else:
```

Figure-8: Listing RFM Customers IDs and their corresponding RFM Scores

```
In [231]:  segmented_rfm['RFMScore'] = segmented_rfm.recency.map(str)+ segmented_rfm.frequency.map(str)+segmented_rfm.monetary_value.map(str
           segmented_rfm.head()
```

Out[231]:

CustomerID	recency	frequency	monetary_value	RFMScore
12346.0	1	5	1	151
12747.0	5	4	1	541
12748.0	5	1	1	511
12749.0	3	3	1	331
12820.0	5	5	1	551

```
In [232]:  segmented_rfm.head()
```

Out[232]:

CustomerID	recency	frequency	monetary_value	RFMScore
12346.0	1	5	1	151
12747.0	5	4	1	541
12748.0	5	1	1	511
12749.0	3	3	1	331
12820.0	5	5	1	551

```
In [233]:  segmented_rfm[segmented_rfm['RFMScore']=='111'].sort_values('monetary_value', ascending=False)
```

Best way contributing Customers

These are the customers with RFM score 5-5-5. It means these are the customers who bought recently, buy often and spend a lot. It is likely that they will continue to do so. Since they already like you, the recommendation is that the store has to keep informing customers about new products, new incentives, any loyalty programs or social media incentives that is run by the stores.

Figure-9: Listing Best RFM Customers IDs and their corresponding RFM Scores

```
In [33]: segmented_rfm[segmented_rfm['RFMScore']=='555'].sort_values('monetary_value', ascending=False)
```

Out[33]:

CustomerID	recency	frequency	monetary_value	RFMScore
14865.0	5	5	5	555
17816.0	5	5	5	555

F=5 indicates that the customers are loyal with high frequency of purchases at stores. But it does not mean they bought it recently and spent high money at stores in purchases. M=5 indicates that customers are big spenders. They spent a lot of money over the given transaction period; so, it is likely that their lifetime value is having potential at stores in terms of investing a lot in your products. The recommendation here is that the stores that it should keep informing about innovative products, expensive products and top line products to this cluster group of customers.

Least contributing Customers

Figure-10: Listing least contributing RFM Customers IDs and their corresponding RFM Scores

```
In [233]: segmented_rfm[segmented_rfm['RFMScore']=='111'].sort_values('monetary_value', ascending=False)
```

Out[233]:

CustomerID	recency	frequency	monetary_value	RFMScore
12856.0	1	1	1	111
13521.0	1	1	1	111
13755.0	1	1	1	111
14547.0	1	1	1	111
15427.0	1	1	1	111
15993.0	1	1	1	111
16049.0	1	1	1	111
16984.0	1	1	1	111
17509.0	1	1	1	111
17813.0	1	1	1	111

These are the customers with RFM score 1-1-1. It means these customers last ordered a lot of time ago, bought very few times and customers spent very, very little. They are unlikely to be considered worth.

Count of RFM Score

Figure-11: Chart listing differently contributing RFM Customers Count
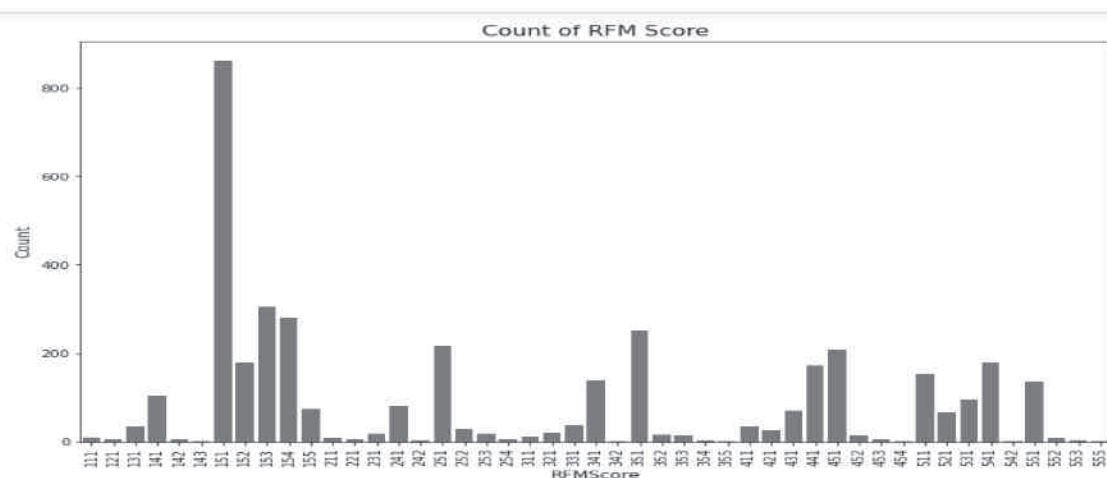
Conclusion

Customer market segmentation is an increasingly becoming significant issue in today's competitive innovative marketing domain like retail groceries store management. In this paper, geographical segmentation is done by filtering the transaction data using data base technique. Further, customer behavioral segmentation is done based on RFM analysis using Python programming statement constructs. Contemporarily, use of data mining techniques is playing important roles in marketing. Though data mining techniques cannot fully replace the significant role of domain experts and their domain expertise knowledge, application of such techniques actively support business experts in business oriented big data analytics and in making better decision in the area of target marketing and electronic customer relationship management. Data mining techniques can help organizations in extracting patterns from the transaction data to gain a competitive advantage in their efforts to optimize customer relationship management. Use of clustering algorithm is one of the ways to segment data set according to the gathered similarities in view of developing insights on behavior-based customer segmentation.

Further, effective loyalty programs rely on sophisticated CRM analytics to understand factors that influence the consumer behavior. Understanding consumer behavior allow to build loyalty programs by adding value in term of conducting customer segmentation and product personalization (Zamil, Al Adwan & Vasista, 2020; Banerjee, 2018).

Further the research study is limited in considering online customers, who order to have grocery items to be delivered online. The research value can be observed on bridging the gap between theoretical understanding and the visible realization through implementation (Bagla & Khan, 2017). As cited in Kasliwal & Agarwal (2014), based on the modified theory of planned behavior, it is found that customers were ready to pay prices (Kim & Han, 2010) online when purchasing grocery items similar to those who make payment at brick and mortar stores without hesitating much.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ahmed, R. (2003). Benefit Segmentation: A Potentially useful technique of segmenting and targeting older consumers. International Journal of Market Research 45 (3), Retrieved on 04-11-2020 https://www.warc.com/fulltext/JMRS/78268.htm

Ait Daoud, R. Bouikhalene, B. Amine, A. & Lbibb, R. (2015). Combining RFM model and clustering techniques for customer value analysis of a company selling online. 978-1-5090-0478-2/15, IEEE. [In text citation: Aid Daoud, Bouikhalene, Amine & Lbibb, 2015]

Aksoy, L., Keiningham, T. & Bejou, D. (2014). Profit Maximization through customer relationship management, Measurement, Prediction and Implementation, New York: Routledge

Aziz, A. (2017). Customer segmentation based on behavioral data in e-marketplace, Thesis submitted to Department of Information Technology, Uppsala University, Sweden.

Bagla, R. K. & Khan, J. (2017). Customers' expectations and satisfaction with online food ordering portals. Prabandhan: Indian Journal of Management, 10(11). [Bagla & Khan, 2017]

Banerjee, K. (2018). Enhancing customer loyalty using market basket analysis, big data and analytics customer experience, customer sentiment analysis, enterprise analytics, market basket analysis, April 2, Retrieved from Aspiresys.com blog. [In text citation: Banerjee, 2018]

Berson, A., Smith, S. & Thearling, K. (1999). Building Data Mining Applications for CRM, 1st Edition, New York: McGraw-Hill Professional. [In text citation: Berson, Smith and Thearling, 1999]

Bilgic, E., Kantardzic, M. & Cakr, O. (2015). Retail Store segmentation for target marketing. In proceedings of ICDM, Industrial conference on data mining: Advances in data mining: Applications and theoretical aspects, 32-44. [In text citation: Bilgic, Kantardzic, Cakr, 2015]

Bose, I. & Chen, X. (2009). Quantitative models for direct marketing: A review from systems perspective, European Journal of Operational Research, 195(1), 1-16. [Bose & Chen, 2009]

Campbell, K. (2015). Google Analytics: How to use Advanced Segments. Retrieved on 24-10-2020 https://www.practicalecommerce.com/Google-

Analytics-How-to-Use-Advanced-Segments

Cooil, B., Aksoy, L.& Keiningham, T. L. (2008). Approaches to customer segmentation. Journal of Relationship Marketing, 6(3-4), 9-39. [Cooil, Aksoy &Keiningha, 2008]

David Shepard Associates (1998). The new direct marketing: How to implement a profit-driven database marketing strategy, 3rd Edition.

Davies, A., Dolega, L. & Arrisbas-Bel, D. (2019). Buy online collect in-store: exploring grocery click & collect using a national case study. International Journal of Retail & Distribution Management, doi: 10.1108/IJRDM-01-2018-0025. [In text citation: Davies, Dolega & Arrisbas-Bel]

Do, T. N. (2011). Market Segmentation: Applying data mining to segment retail market based on purchasing portfolios. In Master Thesis in Business and Informatics submitted to University of Technology in Hochiminh in association with University of Boras, Sweden.

Dogan, O., Aycin, E. & Bulut, Z. A. (2018). Customer Segmentation by using RFM model and clustering methods: A case study in retail industry. International Journal of Contemporary Economics and Administrative Sciences, 8(11), 1-19.[Dogan, Aycin & Bulut, 2018]

Gonzlez-Benito, s. & Gonzalez-Benito, J. (2005). The role of geodemographic segmentation in retail location strategy. International Journal of Market Research, 47 (3), 295-316.

Guy, C. M. (1996). Corporate strategies in good retailing and their local impacts: a case study of Cardiff. Environment and Planning, 28, 1575-1602. [In text citation: Guy, 1996]

Hosseini, M. & Shabani, M. (2015). New approach to customer segmentation based on changes in customer value. Journal of Marketing Analytics, 3(3), 110-121. [In text citation: Hosseini & Shabani]

Huseynov, F. & Yirdirim, S. O. (2017). Behavioral segmentation analysis of online consumer audience in Turkey by using real e-commerce transaction data. International Journal of Economics and Business Research, 14(1), DOI:10.1504/IJBER.2017.085549

Huseynov F. & Ozkan, S. (2017). Behavioral segmentation analysis of online consumer audience in Turkey by suing real e-commerce transaction data. International Journal of Economics and Business Research 14(1), 12-28.

Kamthania, D., Pahwa, A. & Madhvan, S. (2018). Market segmentation analysis and visualization using K-Mode clustering algorithm for E-Commerce Business. Journal of Computing and Information Technology, 26(1), 57-68.

Kasliwal, N. & Agarwal, S. (2014). A Study on Indian consumer's attitude and choice of preferences for green attributes of the hotel industry. Prabandhan: Indian Journal of Management, 8(1), 21-33. [Kasliwal & Agarwal, 2014]

Katsaras, N., Wolfsons, P., Kinsey, J. & Senauer, B. (2001). Data mining a segmentation analysis of U. S. Grocery shoppers. Working paper 01-01, The retail food industry Centre, University of Minnesota. Retrieved on May 20, 2020https://core.ac.uk/download/pdf/6546203.pdf [In text citation: Katsaras, Wolfsons, Kinsey & Senauer, 2001].

Kim, Y. & Han, H. (2010). Intention to pay conventional-Hotel prices at a green hotel-a modification of the theory of planned behavior. Journal of Sustainable Tourism, 18(8), 997-1014. [In text citation: Kim & Han, 2010]

Katti, S. (2015). Market Segmentation. Retrieved on 04-11-2020   https://www.linkedin.com/pulse/market-segmentation-shashank-katti/

Knoll, S. (2008). Cross-Business Synergies. 1st Edition, Gabler Edition Wissenschaft, GmbH, Wiesbaden: Gabler-Verlag. [In text citation: Knoll, 2008]

Kotler, P. & Armstrong, G. (1996). Principles of Marketing. Seventh Edition, New Delhi: Prentice Hall of India. [In text citation: Kotler & Armstrong]

Kotler, P. & Armstrong, G. (1999). Principles of marketing. London: Prentice Hall. [In text citation: Kotler & Armstrong]

Lovas, B. & Ghoshal, S. (2000). Strategy as guided evolution. Strategic Management Journal, Wiley Online Library [In text citation: Lovas & Ghoshal, 2000]

marketbusinessnews.com (2020). What is market segmentation? Retrieved on 04-11-2020 https://marketbusinessnews.com/financial-glossary/market-segmentation/

Martin, G. (2011). The importance of marketing segmentation. American Journal of Business Education, 4(6) 15-18. [In text citation: Martin,

2011]

Miles, M. B. & Huberman, A. M. (1994). Qualitative Data Analysis. Second Edition, USA: Sage Publications. [In text citation: Miles & Huberman, 1994]

Oliverira, V. L. M. (2012). Analytical Customer Relationship Management in Retailing Supported by Data Mining Techniques. A thesis submitted to Faculdad de Engenharia da Universidad do Porto for the doctoral degree in Industrial Engineering and Management.

Oracle docs (n. d.). Customer segmentation. Retrieved on 24-10-2020 https://docs.oracle.com/cd/E76441_01/orase/pdf/orase/170/html/implementation_guide/customer-seg.htm

Pandey, R. (2017). Target marketing using data mining. In proceedings of SARC international conference,March 5th, New Delhi, India. [In text citation: Pandey, 2017]

Raje, A. & Srivastava, R. K. (2014). The connection between data mining and segmentation in marketing area. International Journal of Innovative Research in Computer and Communication Engineering, 2 (12), 7415-7418.

Rajagopal, S. (2011). Customer data clustering using data mining technique. International Journal of Database Management Systems (IJDMS), 3(4), 1-11. [In text citation: Rajagopal, 2011]

Rashid, Y., Rashid, A., Warraich, M. A., Sameen, S. and Waseem, A. (2019). Case Study Method: A step-by-step guide business researcher. doi:org/10.1177/1609406919862424 [In text citation: Rashid, Rashid, Warraich, Sameen and Waseem, 2019]

Rivas, A. (2019). How to Ignite organic growth: Customer segmentation. Retrieved on 24-10-2020 https://www.mktr.ai/how-to-ignite-growth-with-customer-segmentation/

Sharma, A. & Lambert, D. M. (1990). Segmentation of markets based on customer service, International journal of Physical distribution & Logistics Management, 20(7), 19-27. [In text citation: Sharma & Lambert, 1990]

Schrantz, K. P. (2013). Location and Competitive Strategy in Retail: The case of GameStop in Michigan. Master thesis submitted to Western Michigan University as part of Scholar Works at WMU. USA. [Schrantz, 2013]

Sokol, O. & Holy, V. (2020). The role of shopping mission in Retail Customer segmentation. Retrieved onJune 16, 2020 https://arxiv.org/pdf/1909.02996.pdf [In text citation: Sokol & Holy, 2020]

Song, H. S. & Kim, S. H. (2001). Mining the change of customer behavior in an internet shopping mall. Expert Systems with Applications, 21(3), 157-168. [In text citation: Song & Kim, 2001]

Subramaniam, D. (2016). Evolving global business trend and India's challenges. Prabandhan: Indian Journal of Management, 9(3). [In text citation: Subramaniam, 2016]

Susilo, W. H. (2016). An Impact of Behavioral Segmentation to increase consumer Loyalty: Empirical study in higher education of postgraduate institutes at Jakarta. Procedia – Social and Behavioral Sciences, 229, 183-195.

Tavakoli, M., Molavi, M. & Masoumi, V., Mobini, M., Etemad, S. & Rahmani, R. (2018). Customer segmentation and strategy development based on user behavior analysis, RFM model and data mining techniques: A case study. 15th IEEE International Conference on e-Business Engineering. October 12-14, Xi'an China

Turban, E., Sharda, R. Aronson, J. E. & King D. N. (2008). Business Intelligence: A managerial Approach, Upper Saddle River, NJ: Pearson Prentice Hall. [In text citation: Turban, Sharda, Aronson & King, 2008]

Wei, J-T., Lin, S-Y., & Wu, H-H. (2010). A review of the application of RFM model. African Journal of Business Management, 4(19), 4199-4206. [In text citation: Wei, Lin & Wu, 2010]

Yoseph, F., Malim. N. H. A. H. & AlMalaily, M. (2019). New Behavioral segmentation methods to understand consumers in retail industry. International Journal of Computer Science & Information Technology, 11 (1), 43-61.

Zamil, A. M. A., Al-Adwan, A. & Vasista, T. G. (2020). Enhancing customer loyalty with market basket analysis using innovative methods: A Python Implementation approach. International Journal of Innovation, Creativity and Change, September issue (Accepted). [In text citation: [Zamil, Al-Adwan & Vasista, 2020]