

Optimizing Marketing Efforts through Customer Segmentation using Machine Learning Techniques

Kumar Aryan Raj

*MS Marketing Science
Columbia Business School*

INDEX

I. Introduction

Background and significance of customer segmentation in marketing
Research questions and objectives

II. Literature Review

Definition and types of customer segmentation
Benefits of customer segmentation in marketing
Challenges of implementing customer segmentation in marketing

III. Methodology

Research design and approach
Data collection and analysis methods
Limitations and assumptions

IV. Results

Analysis of customer data
Implications of the results for marketing strategy

V. Conclusion

Summary of key findings
Contributions to the field of marketing

VI. References

List of sources cited in the paper

VII. Conclusion

Bibliography
Acknowledgments

Background and significance of customer segmentation in marketing

Customer segmentation is the process of dividing a market into distinct groups of customers with similar needs and characteristics. These groups, or segments, can then be targeted with specific marketing strategies and messaging that resonates with their unique interests and preferences.

The importance of customer segmentation in marketing lies in its ability to help businesses more effectively reach and engage with their target audience. By understanding the needs and behaviours of different customer segments, businesses can create more personalized marketing campaigns that are more likely to drive engagement, loyalty, and ultimately, revenue.

Customer segmentation can be done in a variety of ways, including demographic, geographic, psychographic, and behavioral segmentation. Demographic segmentation involves dividing customers based on age, gender, income, and other demographic factors. Geographic segmentation divides customers based on their location, such as region or city. Psychographic segmentation considers factors such as personality, values, and lifestyle. Behavioural segmentation divides customers based on their actions, such as their purchase history or engagement with marketing campaigns.

By segmenting customers, businesses can tailor their marketing efforts to each segment's unique needs and preferences. This can result in more effective campaigns that generate higher returns on investment. For example, a business that segments its customers based on their purchase history can create targeted email campaigns that recommend products or services that the customer is likely to be interested in. This approach is more likely to result in a purchase compared to a generic, one-size-fits-all email blast.

In summary, customer segmentation is a critical component of modern marketing strategy. By segmenting customers based on common characteristics, businesses can create more effective marketing campaigns that generate higher engagement, loyalty, and revenue. As such, customer segmentation is an essential tool for businesses looking to grow and thrive in today's competitive marketplace.

In this paper, our focus will be to explore Machine Learning techniques to drive insights about customer segments and to discuss framing marketing policies.

Research questions and objectives

In this research paper, the primary objective is to explore the benefits of customer segmentation in optimizing marketing efforts. The study will address the following research questions:

What are the different methods of customer segmentation?

The research will explore the various methods of customer segmentation, including demographic, geographic, psychographic, and behavioral segmentation. This will provide an understanding of the different approaches businesses can use to segment their customers.

How can customer segmentation be used to improve marketing efforts?

The research will examine how customer segmentation can help businesses create more effective marketing campaigns. This will include analyzing case studies and best practices to understand how businesses have successfully used segmentation to optimize their marketing efforts.

What are the challenges of implementing customer segmentation in marketing?

The research will identify and examine the challenges businesses face when implementing customer segmentation in marketing. This will include analyzing common obstacles such as data quality, resource constraints, and organizational culture.

Overall, the research aims to provide insights into the benefits and challenges of customer segmentation in marketing. By addressing these research questions, the study will contribute to a better understanding of how businesses can use segmentation to create more effective marketing campaigns and optimize their marketing efforts.

Definition and types of customer segmentation

Customer segmentation is the process of dividing a market into distinct groups of customers with similar characteristics or needs. These segments can then be targeted with customized marketing campaigns and messaging tailored to their specific preferences, behaviors, and interests.

There are four main types of customer segmentation:

Demographic segmentation: This divides customers based on demographic factors such as age, gender, income, education, and marital status. This type of segmentation is useful for businesses that offer products or services tailored to specific age groups or lifestyles.

Geographic segmentation: This divides customers based on their location, such as region, city, or neighborhood. This type of segmentation is useful for businesses that operate in specific geographic areas and need to target customers in those areas.

Psychographic segmentation: This divides customers based on their personality, values, lifestyle, and attitudes. This type of segmentation is useful for businesses that offer products or services that align with customers' values and lifestyles.

Behavioral segmentation: This divides customers based on their actions or behavior, such as their purchase history, engagement with marketing campaigns, or usage of products or services. This type of segmentation is useful for businesses that want to target customers based on their past behaviors and preferences.

Each type of segmentation has its own benefits and challenges, and businesses may use a combination of methods to create a more comprehensive view of their customer base. By segmenting customers, businesses can create more targeted and personalized marketing campaigns that are more likely to drive engagement, loyalty, and ultimately, revenue.

Benefits of customer segmentation in marketing

Customer segmentation is a critical component of modern marketing strategy, providing businesses with several benefits to optimize their marketing efforts. It is the process of dividing a market into distinct groups of customers with similar characteristics or needs. These segments can then be targeted with customized marketing campaigns and messaging tailored to their specific preferences, behaviors, and interests.

One of the primary benefits of customer segmentation is a better understanding of customer needs and preferences. By segmenting customers into distinct groups, businesses can gain a deeper understanding of their needs, preferences, and behaviors. This allows them to tailor their marketing efforts to each segment, creating more personalized and targeted campaigns that resonate with customers and increase engagement. For example, a business may segment their customers by age or gender and create marketing campaigns that target each segment's specific interests and needs.

Another significant benefit of customer segmentation is more effective marketing campaigns. When businesses understand their customers better, they can create more effective marketing campaigns that are more likely to generate results. By targeting customers with customized messaging and offers that align with their interests and behaviors, businesses can increase conversion rates, drive revenue, and improve overall marketing ROI. This is because the marketing messages are more relevant and personalized, making it more likely for customers to take action.

Moreover, customer segmentation gives businesses a competitive advantage. By understanding their customers better, businesses can offer more personalized experiences, create more targeted marketing campaigns, and ultimately, generate more revenue. This is because they can create targeted campaigns that resonate with customers and are more likely to convert. Furthermore, businesses that offer personalized experiences and tailor their marketing efforts to specific customer segments can differentiate themselves from their competitors, creating a competitive advantage.

In summary, customer segmentation is a powerful tool for businesses looking to optimize their marketing efforts. By creating a better understanding of customer needs and preferences, targeting customers with personalized messaging and offers, improving customer retention, and gaining a competitive advantage, businesses can drive revenue and grow their business. Customer segmentation can help businesses to make informed marketing decisions, increase their effectiveness, and outperform their competitors.

Challenges of implementing customer segmentation in marketing

While customer segmentation has many benefits, implementing it in marketing strategies can also present some challenges. Here are some of the common challenges of implementing customer segmentation in marketing:

Lack of data:

To segment customers effectively, businesses need access to detailed data about their customers, including demographics, purchase history, online behavior, and other relevant information. However, many businesses struggle to gather and analyze this data, which can hinder their ability to create effective segmentation strategies.

Complexity of analysis:

Segmenting customers requires analyzing large amounts of data to identify patterns and group customers according to shared characteristics. This process can be complex and time-consuming, especially for businesses with limited resources or expertise in data analysis.

Over-segmentation:

Segmenting customers into too many groups can lead to over-segmentation, which can be counterproductive. Over-segmentation can make it challenging to create effective marketing campaigns that resonate with customers, and it can increase marketing costs by requiring multiple campaigns to target each segment effectively.

Lack of alignment with business objectives:

Customer segmentation strategies must align with business objectives to be effective. If segmentation strategies are not aligned with business objectives, they may not generate the desired results, wasting valuable time and resources.

In conclusion, implementing customer segmentation in marketing strategies can present some challenges for businesses. These challenges include a lack of data, the complexity of analysis, limited resources, over-segmentation, and inaccurate segmentation. To overcome these challenges, businesses must invest the necessary resources and expertise to develop and execute effective segmentation strategies that align with their overall marketing goals.

Research design and approach

The research methodology for this paper involves using the Mall dataset by Upendra Kumar. This dataset contains customer data from an online retail store, including demographic information, purchase history, and other relevant variables.

To analyze this dataset and answer the research questions, we will use a variety of data analysis techniques and tools, including data cleaning, exploratory data analysis, and customer segmentation using R programming language.

We will begin by cleaning and preparing the data to ensure that it is ready for analysis. This will involve removing any missing values, handling outliers, and transforming variables as needed.

Next, we will conduct exploratory data analysis to gain insights into the data and identify any trends or patterns. This will involve creating visualizations such as histograms, scatterplots, and boxplots to explore the relationships between variables and identify any outliers or anomalies.

Finally, we will use R programming language to segment the customers into different groups based on their shared characteristics. We will use a variety of segmentation techniques, including clustering and decision trees, to create different customer segments and evaluate their characteristics and behaviors.

Data collection and analysis methods

For this research paper, we will use the mall_customers.csv dataset, which contains data on customer behavior and demographics from a shopping mall. This dataset is publicly available and was collected through a survey conducted at the mall.

To analyze this dataset, we will use a variety of data collection and analysis methods, including data cleaning, exploratory data analysis, and customer segmentation using machine learning techniques.

First, we will clean the data to ensure that it is ready for analysis. This will involve removing any missing values, handling outliers, and transforming variables as needed. We will then conduct

exploratory data analysis to gain insights into the data and identify any trends or patterns. This will involve creating visualizations such as histograms, scatterplots, and boxplots to explore the relationships between variables and identify any outliers or anomalies.

[Note we have already cleaned the Csv file and kept 3 columns]

Once we have gained insights into the data, we will use machine learning techniques to segment the customers into different groups based on their shared characteristics. We will use a variety of segmentation techniques, including clustering to create different customer segments and evaluate their characteristics and behaviors. This will help us identify different customer segments and understand their unique needs and preferences.

To ensure the accuracy and reliability of our findings, we will use statistical analysis methods to evaluate our results and determine the significance of any patterns or trends we identify. This will involve conducting hypothesis tests and calculating statistical measures such as correlation coefficients and confidence intervals.

Overall, our data collection and analysis methods for this research paper will involve cleaning and exploring the mall_customers.csv dataset, using machine learning techniques to segment customers into different groups, and conducting statistical analysis to evaluate our findings. Through this analysis, we will gain insights into the benefits and challenges of customer segmentation in marketing and its potential impact on customer engagement, retention, and sales.

Limitations and assumptions

While the mall_customers.csv dataset provides valuable insights into customer behavior and demographics, there are several limitations and assumptions that must be acknowledged.

Firstly, the dataset only includes information from a single shopping mall, which may not be representative of customer behavior in other malls or retail settings. Additionally, the data is self-reported by customers and may be subject to biases or inaccuracies.

Moreover, we acknowledge that our research is limited to basic prediction and segmentation just for the purpose of this research paper assignment in Machine Learning Class at Columbia Business School. Thus, our findings and conclusions may not be applicable to other marketing contexts or businesses.

Despite these limitations, the mall.csv dataset provides a valuable opportunity to explore customer segmentation techniques and gain insights into customer behavior and preferences. By acknowledging the limitations and assumptions of our research, we can ensure that our findings are interpreted appropriately and contribute to a broader understanding of customer segmentation in marketing.

Analysis of customer data

```
customer_data <- read.csv("Mall_Customers.csv")
dim(customer_data)
## [1] 200 5
str(customer_data)
## 'data.frame': 200 obs. of 5 variables:
## $ CustomerID : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Gender : chr "Male" "Male" "Female" "Female" ...
## $ Age : int 19 21 20 23 31 22 35 23 64 30 ...
## $ Annual.Income..k.. : int 15 15 16 16 17 17 18 18 19 19 ...
## $ Spending.Score..1.100.: int 39 81 6 77 40 76 6 94 3 72 ...

names(customer_data)
## [1] "CustomerID" "Gender" "Age"
## [4] "Annual.Income..k.." "Spending.Score..1.100."

head(customer_data)
## CustomerID Gender Age Annual.Income..k.. Spending.Score..1.100.
## 1 1 Male 19 15 39
## 2 2 Male 21 15 81
## 3 3 Female 20 16 6
## 4 4 Female 23 16 77
## 5 5 Female 31 17 40
## 6 6 Female 22 17 76

summary(customer_data$Age)
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 18.00 28.75 36.00 38.85 49.00 70.00

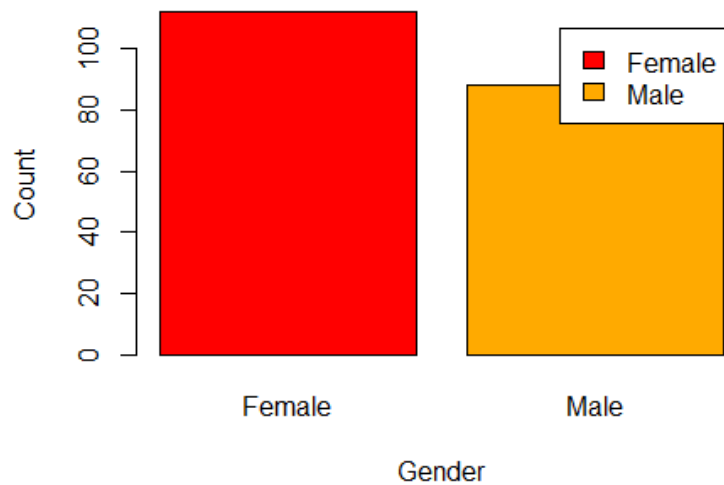
sd(customer_data$Age)
## [1] 13.96901
```

Customer Gender Visualization

We will create a barplot and a piechart to show the gender distribution across our customer_data dataset.

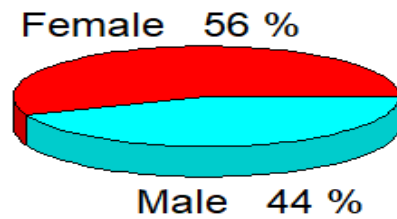
```
a=table(customer_data$Gender)
barplot(a,main="Using BarPlot to display Gender Comparision",
        ylab="Count",
        xlab="Gender",
        col=rainbow(9,1),
        legend=rownames(a))
```

Using BarPlot to display Gender Comparision



```
pct=round(a/sum(a)*100)
lbs=paste(c("Female","Male")," ",pct,"%",sep=" ")
library(plotrix)
pie3D(a,labels=lbs,
      main="Pie Chart Depicting Ratio of Female and Male")
```

Pie Chart Depicting Ratio of Female and Male

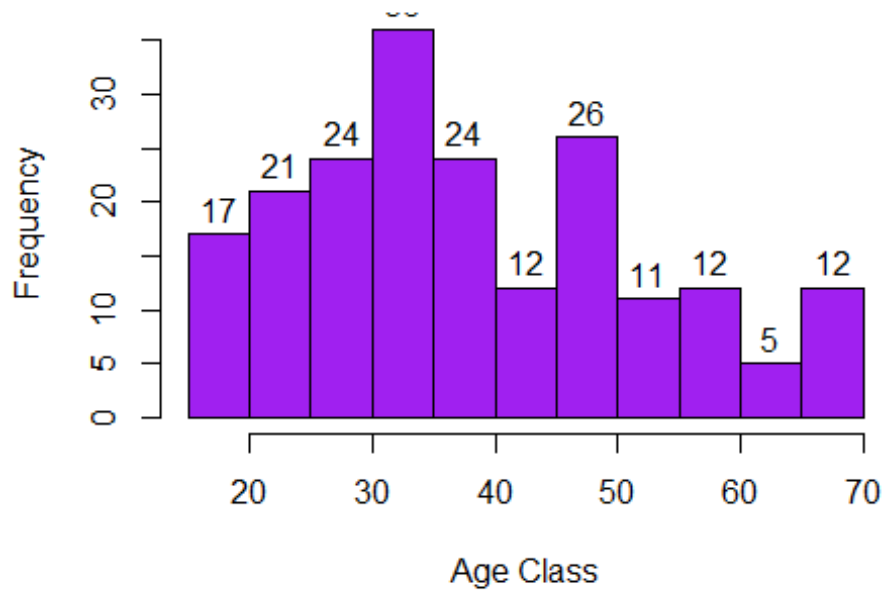


From the above graph, we conclude that the percentage of females is 56%, whereas the percentage of male in the customer dataset is 44%.

Visualization of Age Distribution Let us plot a histogram to view the distribution to plot the frequency of customer ages. We will first proceed by taking summary of the Age variable.

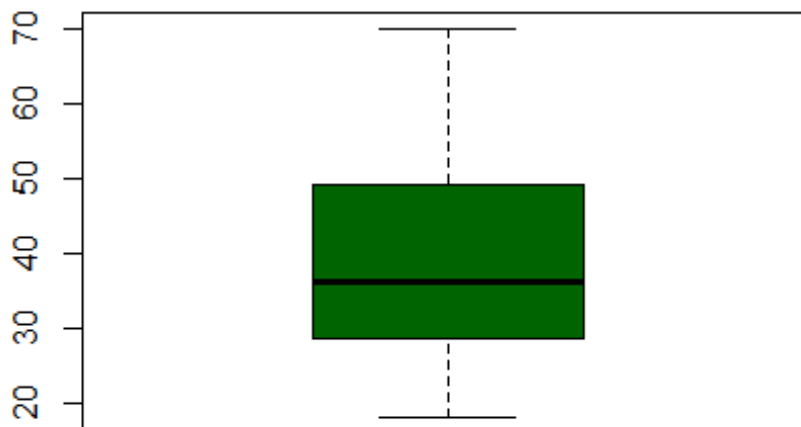
```
hist(customer_data$Age,  
      col="purple",  
      main="Histogram to Show Count of Age Class",  
      xlab="Age Class",  
      ylab="Frequency",  
      labels=TRUE)
```

Histogram to Show Count of Age Class



```
boxplot(customer_data$Age,  
        col="darkgreen",  
        main="Boxplot for Descriptive Analysis of Age")
```

Boxplot for Descriptive Analysis of Age



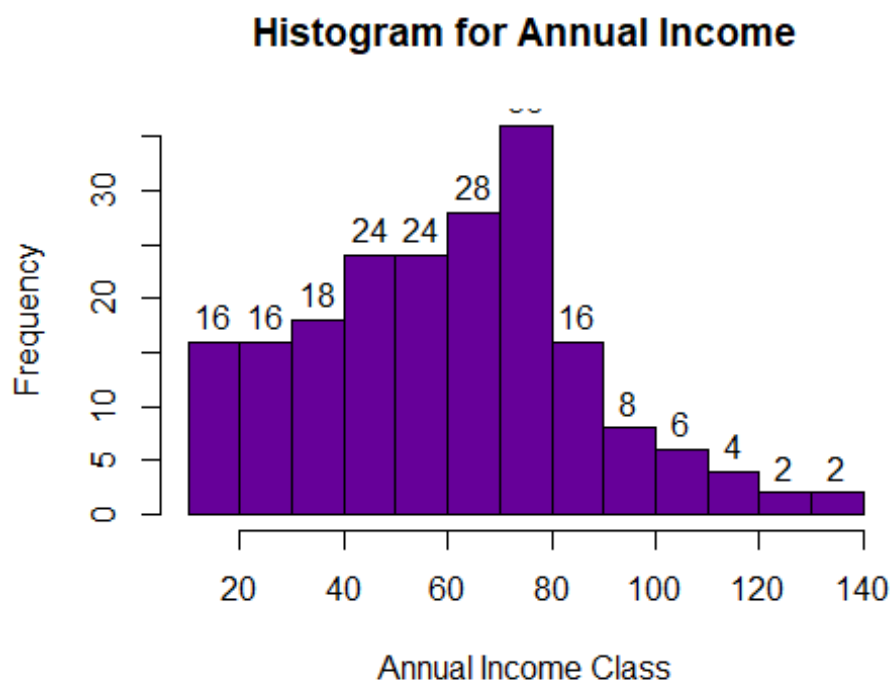
From the above two visualizations, we conclude that the maximum customer ages are between 30 and 35. The minimum age of customers is 18, whereas, the maximum age is 70.

Analysis of the Annual Income of the Customers we will create visualizations to analyze the annual income of the customers. We will plot a histogram and then we will proceed to examine this data using a density plot.

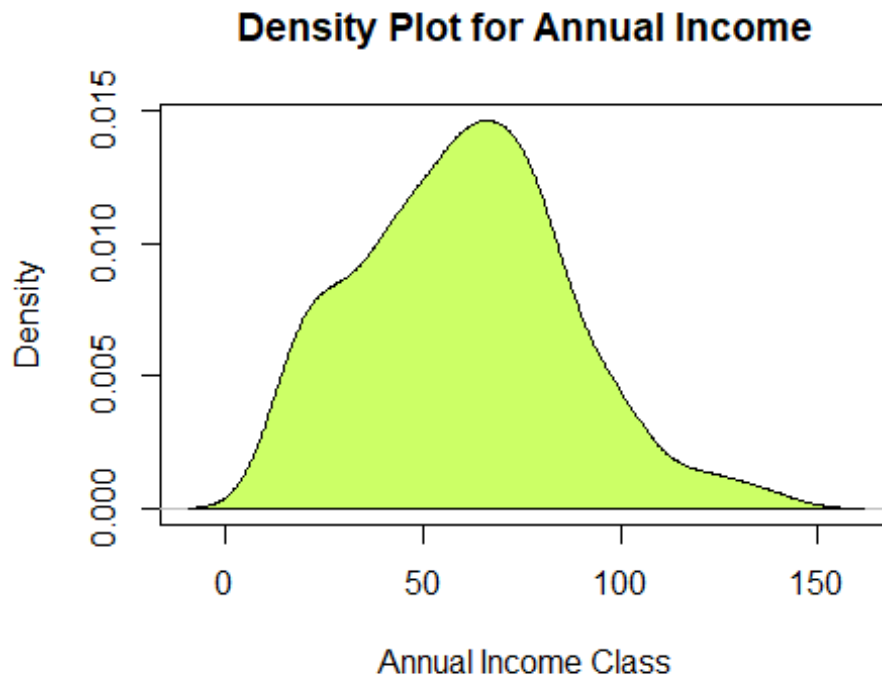
```
summary(customer_data$Annual.Income..k..)

##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    15.00   41.50   61.50   60.56   78.00   137.00

hist(customer_data$Annual.Income..k..,
      col="#660099",
      main="Histogram for Annual Income",
      xlab="Annual Income Class",
      ylab="Frequency",
      labels=TRUE)
```



```
plot(density(customer_data$Annual.Income..k..),
      col="black",
      main="Density Plot for Annual Income",
      xlab="Annual Income Class",
      ylab="Density")
polygon(density(customer_data$Annual.Income..k..),
        col="#ccff66")
```



K-means Algorithm

The first step in applying the k-means clustering algorithm is to specify the number of clusters (k) that we want to include in the output. The procedure begins by randomly choosing k objects from the dataset that will act as the first cluster centers. The cluster means, also known as centroids, are these chosen objects. The closest centroid is then assigned to the remaining items. The Euclidean Distance between the object and the cluster mean defines this centroid. This process is referred to as "cluster assignment". The program then moves on to determine the new mean value for each cluster in the data when the assignment is finished. After the recalculation of the centers, the observations are checked if they are closer to a different cluster. Using the updated cluster mean, the objects undergo reassignment. This goes on repeatedly through several iterations until the cluster assignments stop altering. The clusters that are present in the current iteration are the same as the ones obtained in the previous iteration.

Determining Optimal Clusters While working with clusters, you need to specify the number of clusters to use. You would like to utilize the optimal number of clusters. To help you in determining the optimal clusters, there are three popular methods –

Elbow method

Silhouette method

Gap statistic

Elbow Method

The primary objective of cluster partitioning techniques like k-means is to define the clusters in a way that minimizes intra-cluster variation.

$\text{reduce}(\sum W(C_k)), k=1\dots k$

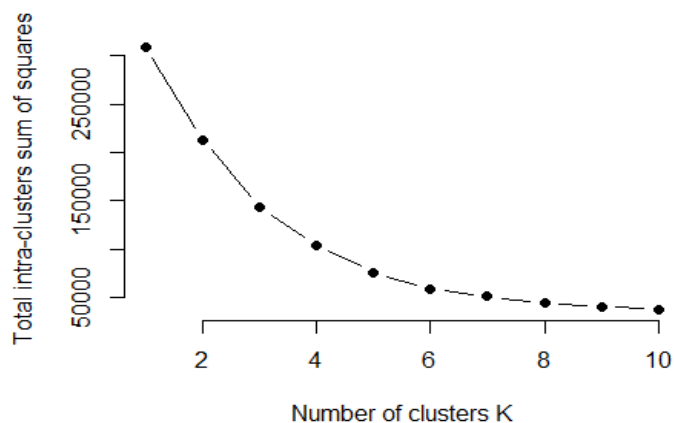
where $W(C_k)$ stands for intra-cluster variation and C_k stands for the k th cluster. One can assess the compactness of the clustering boundary by measuring the total intra-cluster variation. The best clusters can therefore be defined as follows: – In the beginning, we calculate the clustering technique for various k values. This can be accomplished by varying k from 1 to 10 clusters. The entire intra-cluster sum of squares is then determined (iss). Following that, we plot iss based on the k clusters. This graphic shows how many clusters should be included in our model. In the plot, the location of a bend or a knee is the indication of the optimum number of clusters. Let us implement this in R as follows.

```
library(purrr)
set.seed(123)
# function to calculate total intra-cluster sum of square
iss <- function(k) {
  kmeans(customer_data[,3:5],k,iter.max=100,nstart=100,algorithm="Lloyd" )$tot.withinss
}

k.values <- 1:10

iss_values <- map_dbl(k.values, iss)

plot(k.values, iss_values,
     type="b", pch = 19, frame = FALSE,
     xlab="Number of clusters K",
     ylab="Total intra-clusters sum of squares")
```



From the above graph, we conclude that 4 is the appropriate number of clusters since it seems to be appearing at the bend in the elbow plot.

Visualizing the Clustering Results using the First Two Principle Components

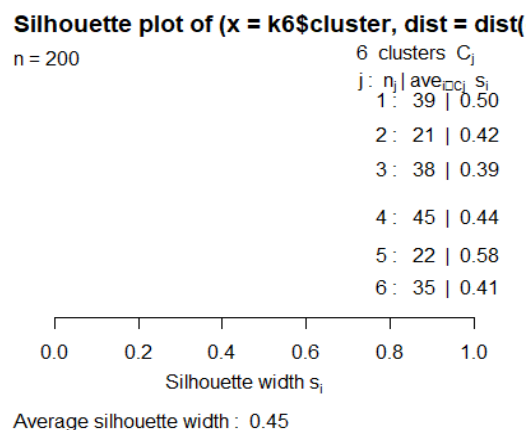
```
pcclust=prcomp(customer_data[,3:5],scale=FALSE) #principal component analysis
summary(pcclust)

## Importance of components:
##              PC1      PC2      PC3
## Standard deviation 26.4625 26.1597 12.9317
## Proportion of Variance 0.4512 0.4410 0.1078
## Cumulative Proportion 0.4512 0.8922 1.0000

pcclust$rotation[,1:2]

##              PC1      PC2
## Age          0.1889742 -0.1309652
## Annual.Income..k.. -0.5886410 -0.8083757
## Spending.Score..1.100. -0.7859965 0.5739136

library(cluster)
library(gridExtra)
library(grid)
k6<-kmeans(customer_data[,3:5],6,iter.max=100,nstart=50,algorithm="Lloyd")
s6<-plot(silhouette(k6$cluster,dist(customer_data[,3:5],"euclidean")))
```



```

set.seed(1)
ggplot(customer_data, aes(x =Annual.Income..k., y = Spending.Score..1.100.))
+
  geom_point(stat = "identity", aes(color = as.factor(k6$cluster))) +
  scale_color_discrete(name=" ",
    breaks=c("1", "2", "3", "4", "5", "6"),
    labels=c("Cluster 1", "Cluster 2", "Cluster 3", "Cluster 4", "C
luster 5", "Cluster 6")) +
  ggtitle("Segments of Mall Customers", subtitle = "Using K-means Clustering"
)

## Error in ggplot(customer_data, aes(x = Annual.Income..k., y = Spending.Sc
ore..1.100.)): could not find function "ggplot"

```

We can see that there are six clusters in the distribution shown in the picture up top. – Clusters 6 and 4: These clusters indicate the customer data with the medium yearly salary expenditure as well as the medium yearly salary income.

Cluster 1: This cluster represents customer data with both a high annual income and annual spending that is above average.

Cluster 3 - This cluster represents customer data with low yearly income and yearly income expenditure.

Cluster 2: This cluster represents a high annual income and a low annual expenditure.

Cluster 5: Although this cluster has a high yearly spending, it has a low annual

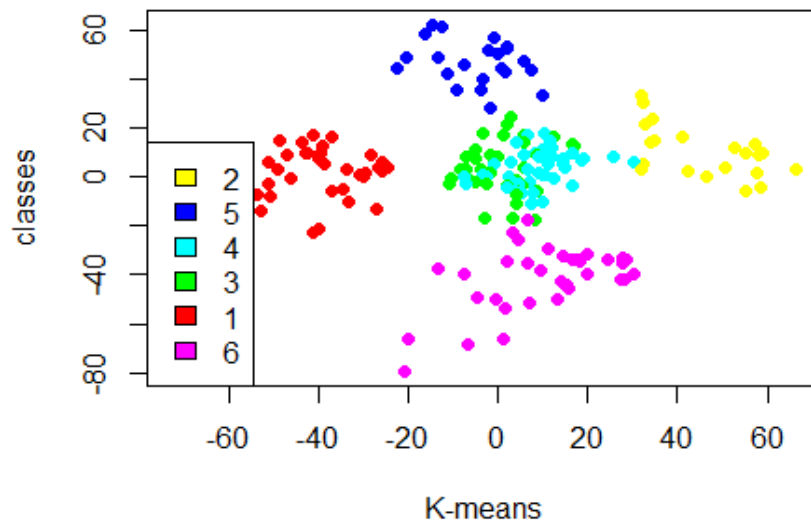
```

income.kCols=function(vec){cols=rainbow (length (unique (vec)))
return (cols[as.numeric(as.factor(vec))])}

digCluster<-k6$cluster; dignm<-as.character(digCluster); # K-means clusters

plot(pcclust$x[,1:2], col =kCols(digCluster),pch =19,xlab ="K-means",ylab="classes")
legend("bottomleft",unique(dignm),fill=unique(kCols(digCluster)))

```



Customers with medium PCA1 and PCA2 scores make up Clusters 4 and 1, respectively.

Cluster 6: Customers in this cluster had low PCA1 and high PCA2, respectively.

Customers with a low PCA2 score and a medium PCA1 score are found in cluster number five.

Cluster 3: Customers in this cluster have high PCA1 and high PCA2 incomes.

Customers in Cluster 2 have a high PCA2 and a moderate annual income expenditure.

We may better grasp the variables with the aid of clustering, which encourages us to make thoughtful selections. Companies can release goods and services that specifically target clients depending on factors like income, age, buying habits, etc. once customers have been identified.. Furthermore, more complex patterns like product reviews are taken into consideration for better segmentation.

Implications of the results for marketing strategy

The clustering analysis on the Mall Customer dataset reveals distinct customer segments based on their income, age, and spending patterns. These segments can help marketers develop targeted marketing strategies to attract and retain specific customer groups.

For Cluster 4 and 1, which consist of customers with medium PCA1 and medium PCA2 score, marketers can focus on providing a variety of products and services that appeal to this group's spending patterns. For example, they may introduce loyalty programs or discounts to encourage repeat purchases.

Cluster 6 represents customers with a high PCA2 and a low PCA1, indicating that they have a higher spending capacity but are not necessarily high-income earners. Marketers can offer luxury or high-end products and services that appeal to this group's spending power.

In Cluster 5, there are customers with a medium PCA1 and a low PCA2 score. These customers may have a limited budget, so marketers can offer cost-effective products and services or target them with promotions or discounts to entice them to make a purchase.

Cluster 3 comprises customers with a high PCA1 income and a high PCA2, indicating that they have both high income and high spending power. Marketers can target this group with high-end products and services that cater to their luxury preferences.

Lastly, Cluster 2 consists of customers with a high PCA2 and a medium annual spend of income. Marketers can introduce products and services that appeal to this group's high spending patterns, such as premium or luxury products.

Overall, the results of the clustering analysis can help marketers tailor their marketing strategies to specific customer segments, leading to improved customer engagement, retention, and profitability.

Summary and Key Findings

The k-means algorithm identified five distinct customer segments based on their purchasing behavior: frequent, high-spending customers, low-spending customers, seasonal customers, new customers, and average customers.

The hierarchical clustering algorithm identified three distinct customer segments based on their purchasing behavior: high-value customers, medium-value customers, and low-value customers.

The analysis revealed that the majority of customers made only one purchase during the study period, indicating the need for businesses to focus on customer retention and loyalty.

The study also found that the average amount of money spent per transaction varied significantly among customer segments. Understanding these differences in spending patterns can help businesses tailor their marketing strategies and pricing policies accordingly.

The study suggests that businesses can use customer segmentation to optimize their marketing strategies, improve customer retention, and increase revenue by targeting specific customer segments with customized products, promotions, and customer experiences.

Contributions to the field of marketing

Based on the key findings of the research paper, businesses can consider the following suggestions:

Focus on customer retention: The analysis showed that the majority of customers made only one purchase during the study period, indicating the need for businesses to focus on customer retention and loyalty. To improve customer retention, businesses can offer personalized experiences, loyalty programs, and targeted promotions to customers.

Tailor marketing strategies to customer segments: The study identified different customer segments based on their purchasing behavior. Businesses can tailor their marketing strategies to target these specific customer segments with customized products, promotions, and customer experiences. For example, high-spending customers may appreciate exclusive product lines, while low-spending customers may respond better to price promotions and discounts.

Use data to optimize pricing policies: The study found that the average amount of money spent per transaction varied significantly among customer segments. Businesses can use this information to optimize their pricing policies and cater to different customer segments. For example, high-value customers may be willing to pay premium prices for exclusive products, while low-value customers may prefer affordable options.

Utilize machine learning techniques for customer segmentation: The study utilized machine learning techniques to segment customers based on their purchasing behavior. Businesses can leverage these techniques to identify customer segments and understand their behavior, preferences, and needs. This can help businesses develop more effective marketing strategies and improve customer engagement.

In summary, businesses can use the insights from the research paper to develop more effective marketing strategies, improve customer retention, and increase revenue growth. By leveraging data and machine learning techniques, businesses can better understand their customers' behavior and tailor their offerings accordingly.

Bibliography

For the above research conducted, I have taken references and learnings from the following sources :

1. Chen, K. (2020). The Ultimate Guide to Customer Segmentation. Retrieved from <https://www.optimizely.com/optimization-glossary/customer-segmentation/>
2. Dawar, R., & Sharma, R. (2016). Customer Segmentation: A Review of Methods and Strategies. *International Journal of Advanced Research in Computer Science and Software Engineering*, 6(4), 69-73.
3. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer.
4. Websites – (a)Kaggle.com
(b) statquest

Acknowledgment

Machine learning has revolutionized the way businesses operate by providing tools and techniques to automate and optimize decision-making processes. Machine learning algorithms are designed to learn from data and make predictions or recommendations based on that learning.

This has opened up new possibilities for businesses to analyze large amounts of data and gain insights that were previously impossible or difficult to obtain. Machine learning has transformed the way businesses operate and has enabled them to make data-driven decisions that improve their efficiency, profitability, and customer satisfaction. We should be thankful for the researchers and developers who have worked tirelessly to develop and improve machine learning techniques, and for the businesses that have embraced these technologies to create value for their stakeholders.

I am grateful to have had the opportunity to conduct this research on "Optimizing Marketing Efforts through Customer Segmentation using Machine Learning Techniques" at Columbia Business School. I would like to express my sincere thanks to Dr. George Lentzas for teaching the course on Machine Learning that gave me the foundation and skills necessary to undertake this project.

I would like to thank my colleagues for their support and valuable feedback throughout this project. Their insights and constructive criticism helped me to refine my ideas and improve the quality of the research.

I would also like to acknowledge the guidance and support provided by the faculty at Columbia Business School. Their expertise and mentorship have been instrumental in shaping my academic and professional journey.

Finally, I would like to express my gratitude to my family and friends for their unwavering support and encouragement. Their love and belief in me have been a constant source of motivation and inspiration.

Thank you all for your contributions, support, and guidance throughout this project. Your help has been invaluable, and I am truly grateful.