



Customer Segmentation

What is customer segmentation?

Customer segmentation is the process of separating your customers into groups based on certain traits (e.g. personality, interests, habits) and factors (e.g. demographics, industry, income) they share.

Segmentation offers a simple way of organizing and managing your company's relationships with your customers. This process also makes it easy to tailor and personalize your marketing, service, and sales efforts to the needs of specific groups. This helps boost customer loyalty and conversions.

References

<https://blog.hubspot.com/service/customer-segmentation>

Why segment customer?

There are a number of other reasons why customer segmentation is so important. Here are some of the things this process can help your business accomplish:

- Learning about your customers on a deeper level so you can tailor your content to their unique needs and challenges
- Creating targeted campaigns and ads to resonate with and convert segments of customers
- Improving your customer service and customer support efforts by understanding and preparing for challenges different groups are likely to experience
- Increasing customer loyalty with customized content and interactions
- Understanding who your most valuable customers are and why
- Communicating with segments of customers via preferred channel or platform.

K-means clustering for customer segmentation with supermarket data

K-means clustering

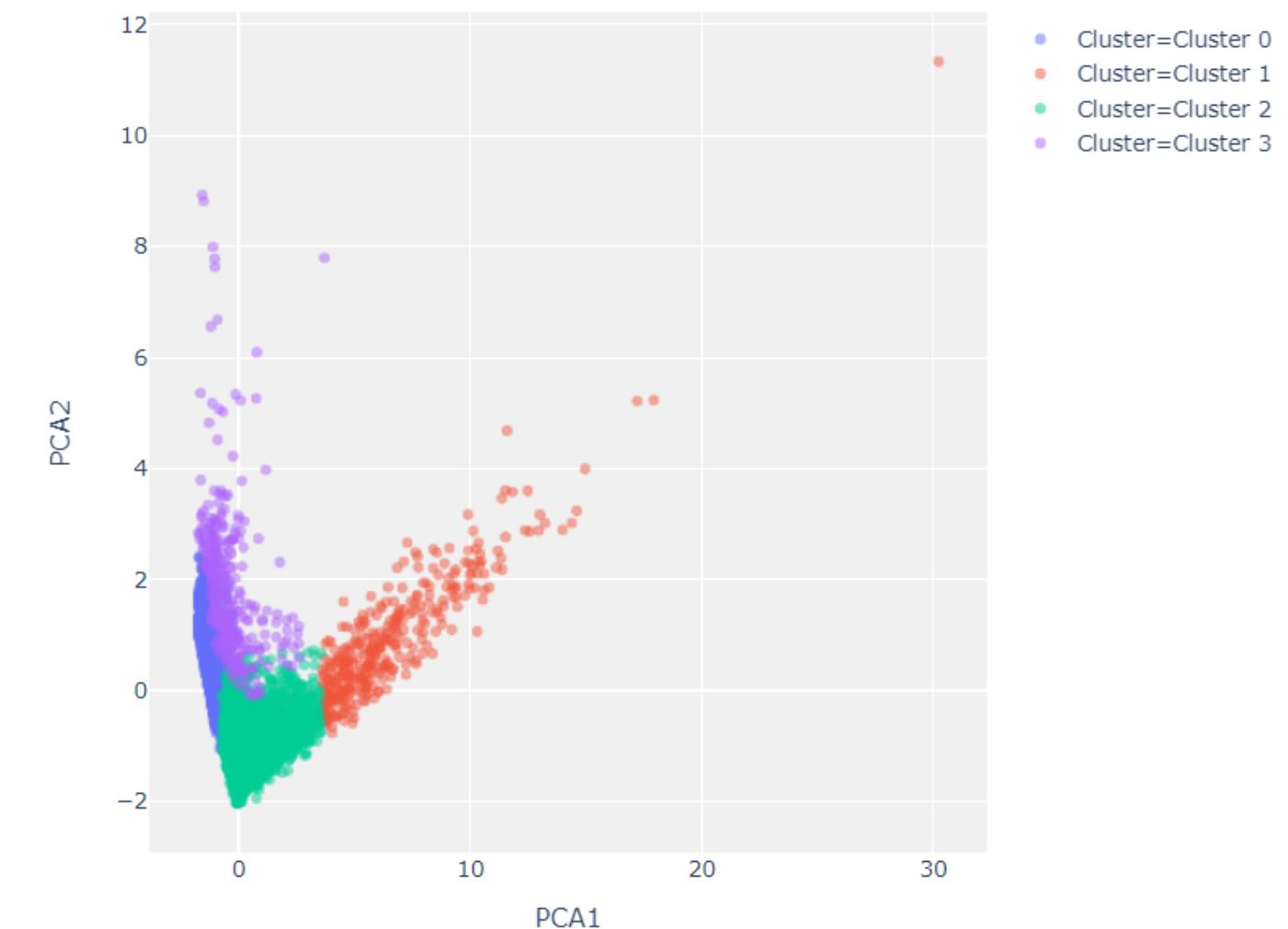
K-Means clustering is an unsupervised machine learning algorithm that divides the given data into the given number of clusters. Here, the "K" is the given number of predefined clusters, that need to be created.

It is a centroid-based algorithm in which each cluster is associated with a centroid. The main idea is to reduce the distance between the data points and their respective cluster centroid.

In K-Mean clustering, we can apply the **elbow method** for selecting the K number of clusters. Another method is calculating the **silhouette score**.

K-means clustering results

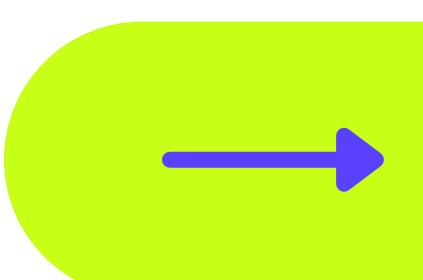
Applying K-Means for segmentation supermarket data **can be clustered into 4 groups** of customers according to the variables that are analyzed as follows: Total Spend Total visits Total SKUs Total quantity Ticket size Recency Customer price sensitive.



References

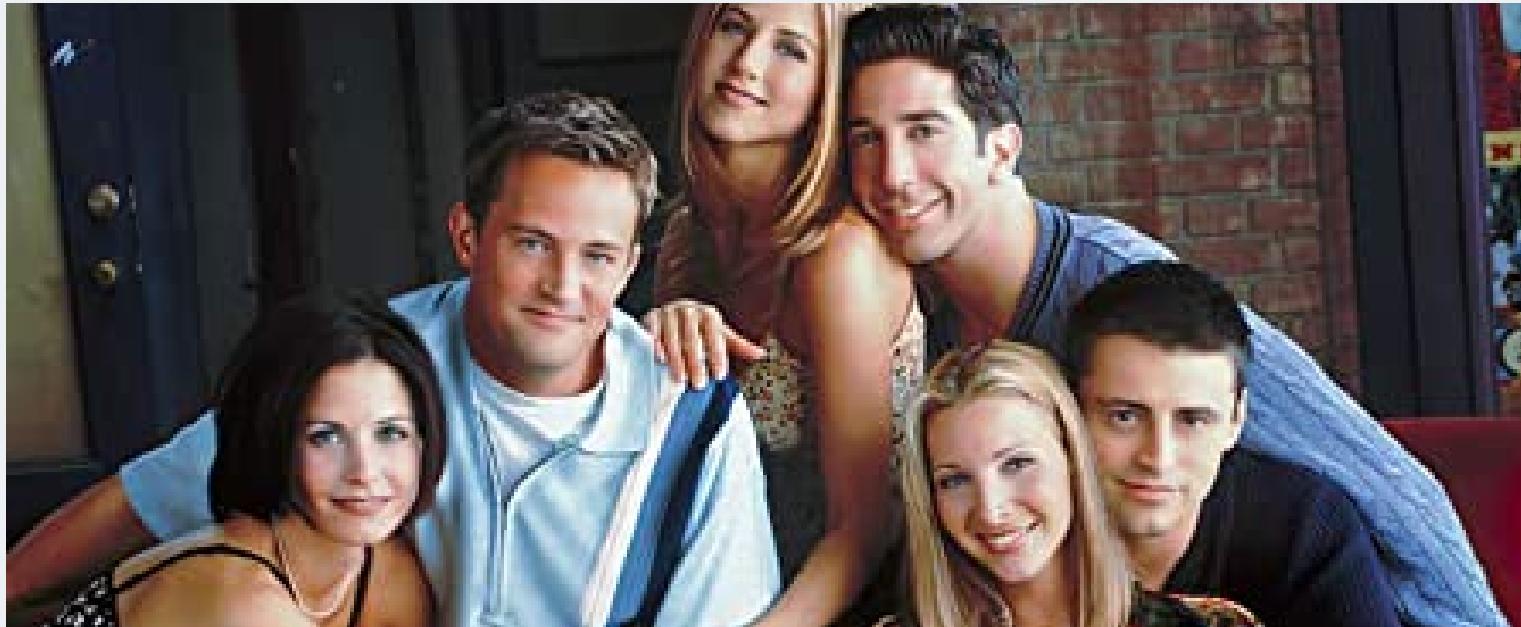
<https://medium.com/analytics-vidhya/customer-segmentation-with-k-mean-clustering-852203e18ab7>

<https://www.analyticsvidhya.com/blog/2021/05/k-means-clustering-with-mall-customer-segmentation-data-full-detailed-code-and-explanation/>



Interpret results and plan for actions

Name each cluster of customers like our friend's group.



Super best friends (Cluster 1)

With a group of 366 persons. The cluster of less but quality. They come to our supermarket a lot. The average total spend of them is 2526 which is the highest spend of all clusters. They buy plenty of products about 2074 quantity. The recency is 5.63 (shortest). It means the more recently a customer has purchased with our store, the more likely they will continue to keep the business for subsequent purchases. The low recency is the sign of they might switch to brand competitors although they have loyal to us. The average ticket size is 18.58 which is a value that tells us about what amount of money that each buyer spends per visit.

plan to action:

- They are pure gold customers. Should keep them happy and gain loyalty from them.
- Understand the reasons why they become customers with our store.
- Give them a high level of service
- Apply personalization such as a surprise coupon discount. Send promotion coupons to a direct email or another contact.
- Feedback is key. Survey about pain points of customers and plan solutions to fix them immediately.



(Almost) best friends (Cluster 2)

They are a huge cluster with 2518 buyers. They spend quite much about 198 which is the second-highest spend. The average total visit is 21.08. They sometimes come but sometimes don't. However, they buy about 157 quantities not as high as the quantity of our super best friend but it's higher than the rest of the group. The recency is 53.56 and the average ticket size is 9.69

plan to action:

- They will be the super best friends if you treat them well.
- Make them buy more by upselling which persuading them to buy higher-priced items
- Cross-selling is an option to buy related products to increase basket size.
- Improve their experience and apply personalization as a cluster of super best friends.
- Social media analysis for collecting feedback from customers. How they talk about us. For a deeper understanding of customers' behavior and demographics.

Interpret results and plan for actions

Name each cluster of customers like our friend's group.



Long lost friends (Cluster 3)

Even if they are the smallest group but they influence us in some way. The average total spend is similar to the (almost) best friend cluster with around 120.41. Although they spend quite a lot, they still don't often come to our store. The average total visit is low(2.8). The average total quantity is 73.03. When considering the ticket size of this cluster found that they have the highest ticket size with 40.41. The recency is 317.75. (second longest)

plan to action:

- They are the type of win-back customers.
- Analyze buying pattern which is purchase decision and customers' behavior in the past.
- Make them feel personalized by directing text with a message like "Haven't seen you in a while we have a special offer for you" "Come back and receive"
- Calculate the customer's lifetime value.



Strangers (Cluster 0)

aka those barely bring in any revenue and are not necessarily loyal to ours. (it's sad but it is true) The biggest cluster of our supermarket (2856). They hardly come to our store. Their average total visit is 1.78 (lowest). They spend less about 11.14. It is a group of customers who rarely come and when they come they rarely buy anyway. The average ticket size is 5.57(lowest). The recency is 431.34 (longest).

plan to action:

- Don't just eliminate our strangers, they can turn into best friends someday.
- Churn analysis to predict the trend of the number of customers who will leave our store.
- Risk or fraud pattern to analyze if the customer is at risk of discontinuing purchase.

References

- <https://www.tractionwise.com/en/magazine/customer-retention-strategies/> <https://www.commbox.io/8-best-customer-retention-strategies-your-business-should-use/>
<https://www.shopify.com/retail/the-art-of-the-upsell-how-to-get-customers-to-buy-more-without-being-overbearing> <https://blog.fivestars.com/3-marketing-ideas-ways-to-win-back-lost-customers/>

Results from python

	TotalSpend	TotalVisits	TotalSKUs	TotalQuantity	CustPrice	TicketSize	total_days	recency	SpendPerQuantity
Cluster									
Cluster 0	11.140823	1.782213	6.157913	9.614496	1.881653	5.568577	51.249300	431.343838	1.144871
Cluster 1	2526.647186	172.131148	367.245902	2074.076503	1.852459	18.580841	792.866120	5.631148	1.253190
Cluster 2	198.933507	21.079031	55.832010	157.539317	1.881255	9.690359	608.814138	53.359412	1.285401
Cluster 3	120.411278	2.808333	27.647222	73.030556	1.983333	40.410560	148.077778	317.747222	3.723818

