Customer Segmentation Analysis:

Sam's Store

Key Focus

Who are our current and potential customers and how can we tailer our services to them, give them better service and increase company reveue?

Who is engaging most with the company?

What do our "best" customers look like?

Where We Are

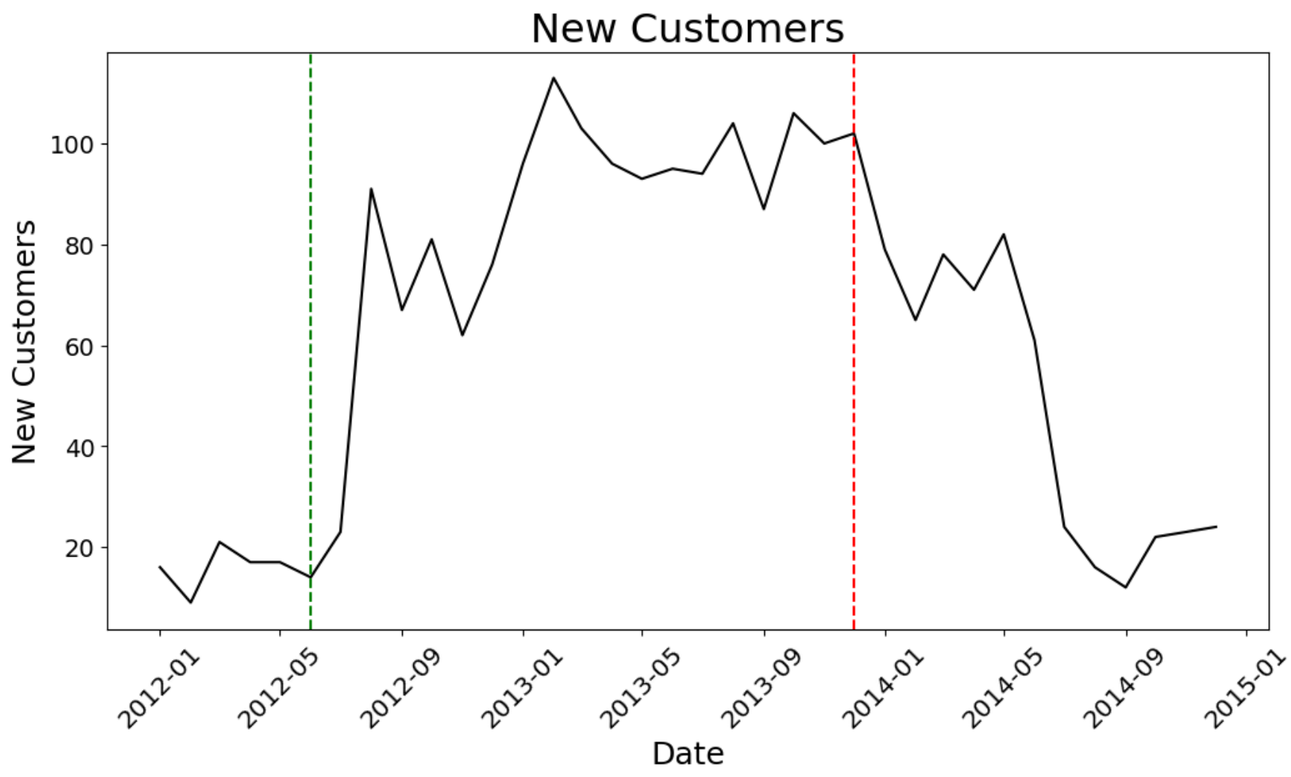
- The company has recently faced a drop in Enrollments after a 1.5yr high.
- We have netted $1.3 million over the past 3 years.
- Our campaigns aren't well received by our customers.

Our failure to bring in new customers and low campaign acceptance might indicate that we don't understand our customer base well. By Segmenting our customer base, we will be able to understand our customers and be able to identify and target potential customers.

Revenue
$1.3M

Avg. Campaign Acceptance
7.5%

New Customers
↓47%
between 2013 and 2014



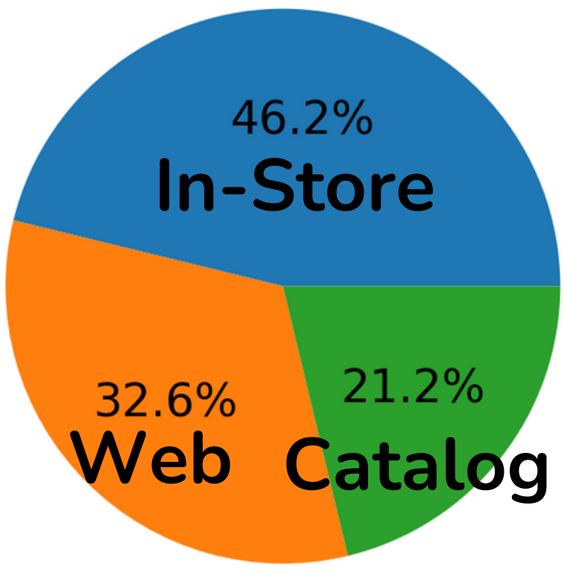
Understanding our customer base.

Unsupervised learning tools in Data Science give us the ability to identify hidden patterns in the data that reveal things about customer groups and preferences. This is called Customer Segmentation. Segmented marketing campaigns receive a higher return on investment when compared to non-segmented marketing campaigns.

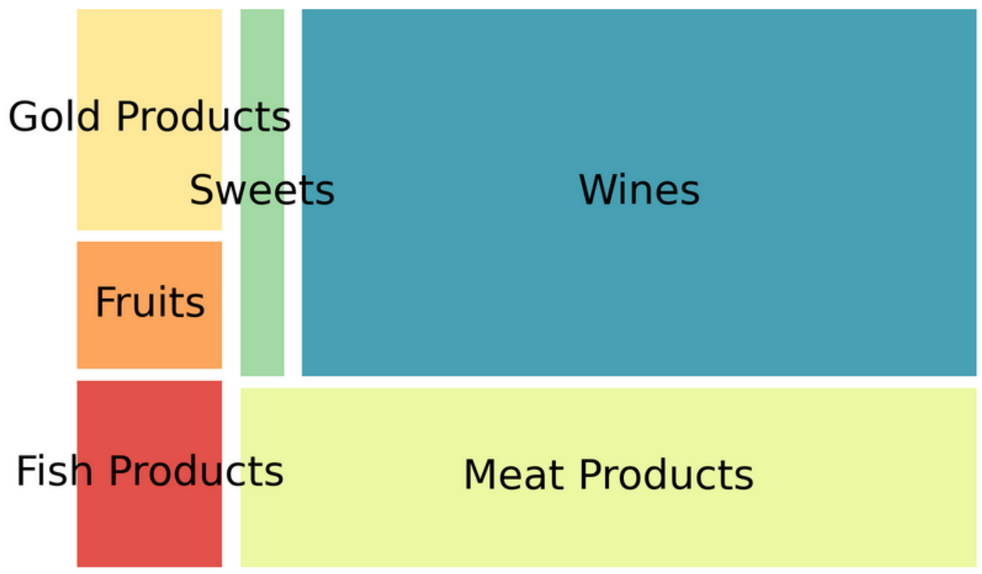
The Process:

Data Exploration	Simplifying	Modeling	Analysis	Recommendation
Understanding Data helps the analyst make decisions about the direction of the segmentation.	Preprocess the data for the segmentation models	Find the best possible model that helps us find the groups that make the most sense	Once we've found the best model, we use it to get an even deeper understanding of our Customers	Make a data driven decision based on the Customer segments and the original goals

Preferred Channels



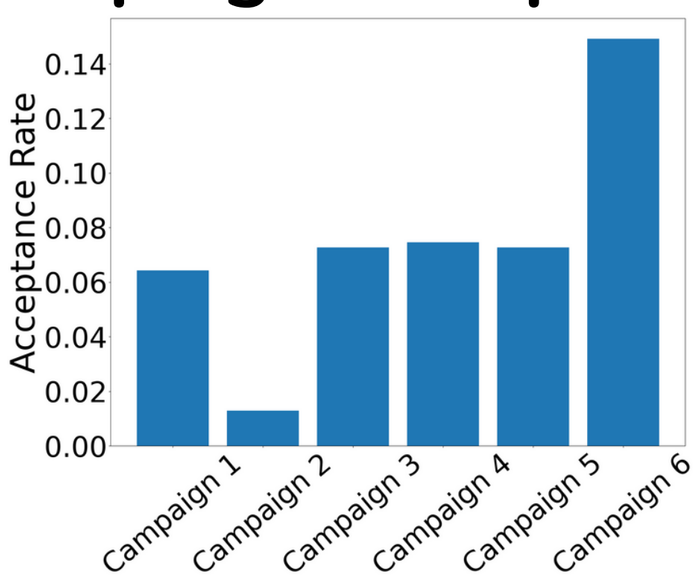
Total Product Sales Distribution



Analytics

Avg Purchases	12.6
Cost per Cart	$38.20
Customer Income	$52K

Campaign Acceptance



It was apparent in the analysis That Wine and Meat are our most popular products. This claim is evidenced by the fact that the Total expenditure is most strongly correlated to the expenditure on Meat and Wine products especially through the catalog and in-store. I recommend that we prioritize selling these products on our web platform. Web purchases account for ~33% of our total sales, yet according to my analysis, the amount spent on Meat and Web purchases have a low correlation compared to the correlation between the amount spent on meat and catalog and store purchases. This analysis is substantiated when examining the variable loadings for our principal components. The amount spent on Wine and Meat appears to be what drives purchasing in the catalog, and in-store and increases the cost of our cart.
This suggests that we approach our web channel as a weak link in our business model. We should analyze our web platform to see why we aren't selling as much meat online as we are in-store and in our catalogs. It is possible that an incentive program to increase meat expenditure online will be what customers need to get excited about making purchases.

Dimensionality Reduction

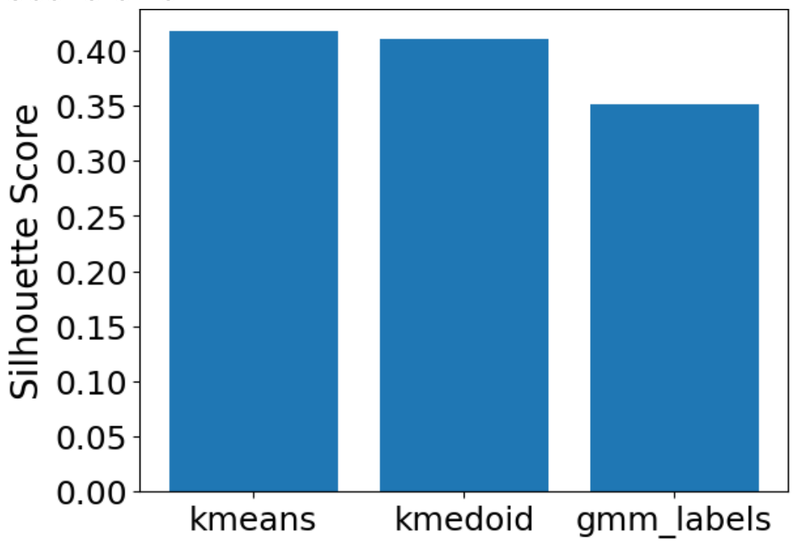
The dataset after Feature Engineering and analysis contained 32 variables. In order to ensure that our unsupervised learning techniques can produce the most distinct and usable cluster groups, we need to reduce the complexity of the data while retaining as much informaiton. This is where Dimensionality Reduction comes into play.
I chose to use Principal Component Analysis, a Feature Extraction method. Feature Extraction transforms the variables instead of just selecting variables. Principal Components transforms the data into components that are correlate with the most important factors of the dataset.

PC1: Customers with a high PC1 score may be our higher spenders.
PC2: Customers with a high PC2 score are NOT very engaged with the store.
PC3:This may underpin a weak pattern of customers that don't shop very often, but may like to purchase meat from us.

	PC1	PC2	PC3
Recency	0.014815	-0.994746	-0.099663
MntWines	0.335678	0.018088	-0.162794
MntMeatProducts	0.513621	-0.020158	0.352904
NumWebPurchases	0.232037	0.060846	-0.570088
NumCatalogPurchases	0.421343	-0.012760	0.055300
NumStorePurchases	0.335257	0.046453	-0.325362
NumWebVisitsMonth	-0.186125	0.031584	-0.356704
TotalNumPurchases	0.211934	0.047668	-0.450908
AvgCostperCart	0.447177	-0.021293	0.273037

Clustering Methods

3 types of Clustering methods were tested: K-Means, K-Medoids, and Gaussian Mixture Modeling. Silhouette Scoring and useful cluster profiling were used to determine the best clustering method.
K-Means produced the best clusters of the clustering methods used with a cluster score of 0.41.



While K-Medoids was close behind at 0.40, a comparison between the cluster profiles revealed that the clusters themselves didn't reveal any information that could be used to identify potential customers. K-Medoids also featured clusters that had PC score distributions with a large spread. While interpretation of Principal Component Scores are of themselves difficult to interpret, the K-Medoids clustering rendered them incomprehensible. The Gaussian Mixture model produced clusters that were very similar in profile and PC score distribution to those of the K-Means Clustering algorithm, the Silhouette score favored K-Means over the Gaussian clustering method.

Cluster Analysis

Cluster 0 - Wealth Wizards
has the highest income and the smallest household sizes. On average, they are big spenders, and have responded well to our campaigns, but do not shop deals very often. They prefer to shop in-store and order through the catalog.

Cluster 1 - Aloof Amigos
has the lowest income and highest family sizes. Our campaigns have resonated with them the least of all the clusters. While they do not shop often, their preferred channel is in store. Even though they are lower income, they do not shop deals as often as cluster 2. This suggests low engagement.

Cluster 2 - Engaged Edgars
has a moderate income and mode rately sized households that are more likely to have more teens than younger children. Those in Cluster 2 favor our web and in-store channels, though appear to be familiar enough with the catalog. They appear to love deals and accept a moderate amount of campaigns. These customers are Highly engaged.

Iterate

Many Iterations of this analysis were executed. A variety of changes were made across iterations including changing the way outliers were treated, which variables had their outliers treated, and a process of eliminating or including variables into the dimensionality reduction and clustering analysis. The success of an an iteration was determined by the end result, the silhouette scores of the clustering methods used.

Final Recommendation

We must use the information revealed about our clusters to make our campaigns targeted. Our enrollments have dropped significantly and it can be seen that the campaigns aren't successfully engaging our customers to a high enough degree. Targeted campaigns will increase engagement and profit by engaging untapped customer potential. Campaigns should be targeted at our Aloof Amigo and Engaged Edgar groups, the clusters with the most potential, and our largest groups.

Those in Aloof Amigos are our largest-sized group and comprise 50% of our customers, yet aren't engaged. They rarely accept campaigns and even though they have less income than those in Engaged Edgars, they don't accept deals as often. In addition to making less money, those in Aloof Amigos and Engaged Edgars have similar-sized families, which should necessitate the same thriftiness that those in Engaged Edgars demonstrate. Though it isn't clear why Aloof Amigos are less engaged, the data provides some feedback as to how we can improve our engagement with the Aloof Amigos: I believe that we can convert customers in Aloof Amigos to Engaged Edgars customers. Those in Aloof Amigos more often live in lower-income homes with younger children. While they will never be able to spend as much as the Wealth Wizards, we can increase the Aloof Amigos' engagement by appealing to their need for budget-friendly options using our success with the Engaged Edgars as a template(with a few modifications).

Aloof Amigos families will likely respond well to a marketing/branding approach that helps them see our business as family and budget-friendly. This approach will pair nicely with my earlier proposal to analyze our online platform and optimize our branding approach to those in Aloof Amigos. The targeted marketing may have the best ROI if we focus on drawing them into the store and to our website to shop for items that they need for their younger children (Aloof Amigos customers are more likely to have younger children than teens).