Capstone Project: Customer Segmentation
Final Submission

Executive Summary

- Customers can be segmented into 3 groups
- The key attributes for the 3 groups are income and family size, which customer profiles can be based on
- The more meat products and wine products a customer purchases, the more likely they are to be in the high income group
- Customers with the least amount of kids (family size) have the highest incomes
- Higher income groups tend to make more in store and catalog purchases, while lower income tend to make more website visits/purchases.
- The company marketing team can use these three groups to target their ads to the correct channels and the correct market.
- The marketing team can also personalize their communication efforts a lot better based of the group characteristics and behaviors, leading to higher ROI
- The final proposed model is K-Means at K=3 as it had the most clearly distinguishable clusters, can handle a large amount of data, and is cost-effective

Problem and Solution Summary

- The problem is we do not have segmented customers. This problem is crucial to solve as customer segmentation is a critical part of marketing operations with direct consequences on sales and marketing strategy.
    - Customer segmentations help marketers develop personalized communications and offerings that cater to each segment's interest.
    - In today's world, personalized customer interaction is highly preferred by customers. Overall, customer segmentation plays a vital role in optimizing ROI for marketing efforts.

- The goal is to figure out the best possible customer segments using the given customer dataset. This can be achieved using Unsupervised Learning techniques like Dimensionality Reduction and Clustering.

- After visualizing the data by applying T-SNE and PCA, and running clustering algorithms like Gaussian Mixture Model, DBSCAN, Hierarchical Clustering, K-Medoids, K-Means to cluster the data, we found K-Means to be the best