Be Customer Wise or Otherwise

A Customer Segmentation Study based on Income and Expense in Philippines Household

Prepared by: Gabriella Pauline Djojosaputro

Date: 18 Oct 2020

Icons in this presentation are made by https://www.flaticon.com/authors/freepik unless otherwise specified

Introduction

For a hypermarket retail that is looking to develop their market into the Philippines region, customer segmentation study is helpful to inform their market development strategy. Market segmentation aims to divide the customers into homogenous groups that has similar characteristics. The aim of this analysis is to identify profitable segments and develop the profiles of these segments. The profiles may be used to inform the appropriate promotional strategy to attract each customer segments.

Profile of the customer is inferred from the Family Income and Expenditure Survey (FIES) conducted by the government of Philippines in 2018. The survey includes data regarding the income and expenses of households in the Philippines by categories, such as for bread, meat, fish, oil, sauce, vegetable, fruit, chocolate, bottled drinks, milk, alcohol, and tobacco. These data is useful to analyze the spending patterns of the households in the Philippines.

Mapa, C. D. S. (2020). *2018 Family Income and Expenditure Survey*. Philippines Statistics Authority. Retrieved October 5, 2020 from
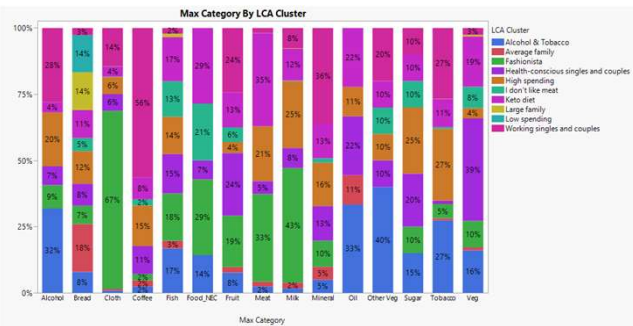
https://psa.gov.ph/sites/default/files/FIES%202018%20Final%20Report.pdf.

Approach
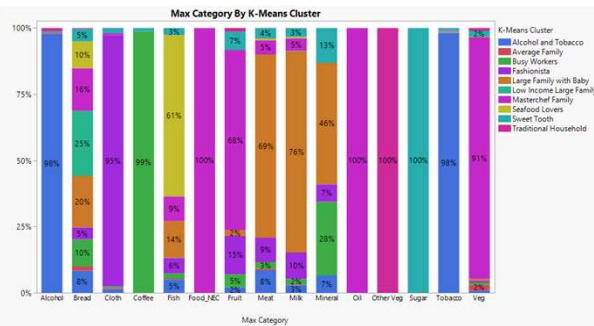
K-Means Clustering
• Shows clear pattern in the category that has maximum expense

Latent Class Analysis Clustering
• Pattern is not very clear

Two clustering techniques was used in the analysis, but k-means shows a clearer pattern when the categories which have the maximum expense for each record in the clusters are examined. More than 90% of the records that has alcohol, cloth, coffee, food that are not otherwise specified (food_NEC), oil, other vegetables, sugar, tobacco, and vegetable belong to a single cluster. This shows that the k-means technique has successfully identified a common pattern within each cluster that differentiates them from the other clusters.

Meanwhile, for latent class analysis (LCA) the distribution does not exhibit a clear pattern. This is probably due to the binning that was performed, as it would no longer retain the maximum expense but simply allocate it into different bins. LCA does not care about the actual value of the variables, but only whether the expenses belong to the lowest quartile, mid-low, mid-high, or the top quartile.

As such, the results from k-means clustering is chosen to describe the potential customer segments in the Philippines.
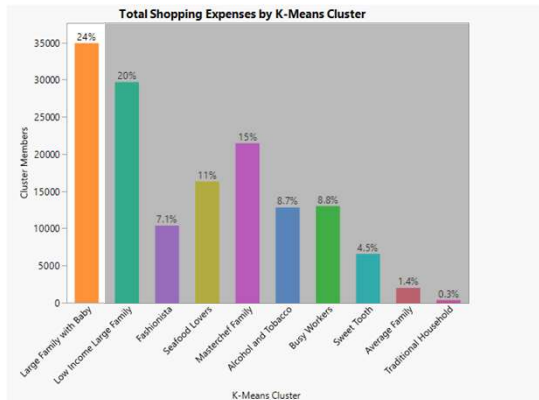
Customer Segments
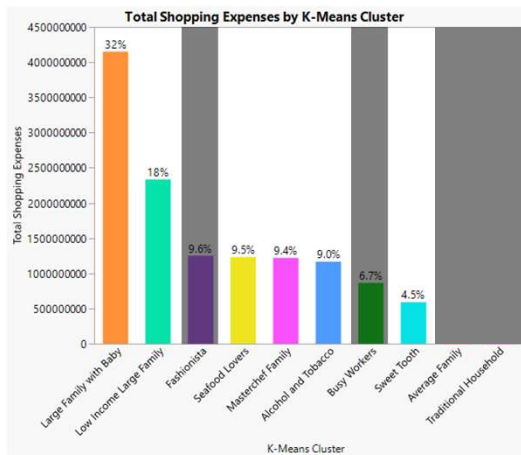
- 10 Clusters by K-means
  - Large Family with Baby
  - Low Income Large Family
  - ~~Fashionista~~
  - Seafood Lovers
  - Masterchef Family
  - Alcohol and Tobacco
  - ~~Busy Workers~~
  - Sweet Tooth
  - ~~Average Family~~
  - ~~Traditional Household~~

Clustering using k-means find 10 as the optimum number of clusters. Out of the 10 clusters, 5 segments are found to be key customer segments. The clusters are ranked by the sum of shopping expenses of its members. Large Family with Baby seems to be the most profitable segment, contributing to 32% of the total shopping expenses. Some segments are not considered as key customer segments for the following reasons:

- Fashionista
  - Although hypermarket may sell clothing, they are likely to purchase it from a clothing line instead.
- Busy Workers
  - Their expenses are mostly on coffee, which are probably not bought from hypermarket but coffee shop.
  - Considering that they have little time, they are more likely to purchase other items from convenience shop as well.
- Average Family
  - Excluded because they have no strong preference on a particular item.
  - Only contribute to 1.2% of total shopping expenses.
- Traditional Household
  - More likely to purchase from a traditional market.
  - Only contribute to 0.2% of total shopping expenses

4

Large Family with Baby

Total Shopping Expenses by K-Means Cluster

- Family Size
- Characteristics:
  - Total Income: ₱350,440
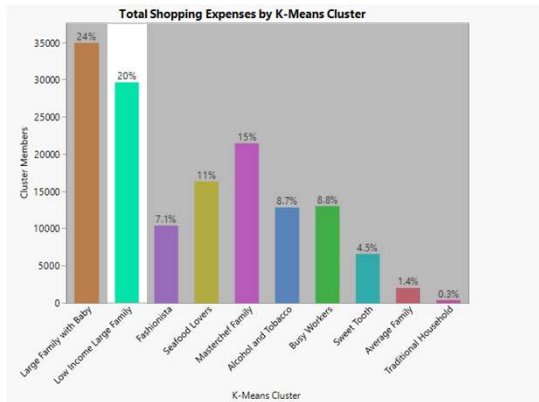 - Shopping Expenses: ₱109,028
  - 24% of the sample
- Preferred Items:

The first key customer segment identified is large family with baby, which is also likely to be the most profitable segment because it contributes to 32% of the total shopping expenses of the whole dataset, despite only making up 24% of the records.

The characteristic of this segment is that is has a larger family (median family size: 5). It also has the highest median total income (₱350,440) and shopping expenses (₱109,028) compared to the other segments. They spend a higher percentage of their expenses on milk, meat, and bottled drinks compared to the other segments. They are likely to have a baby in the family due to the high proportion of expenditure on milk.

Low Income Large Family

Total Shopping Expenses by K-Means Cluster

- Family Size
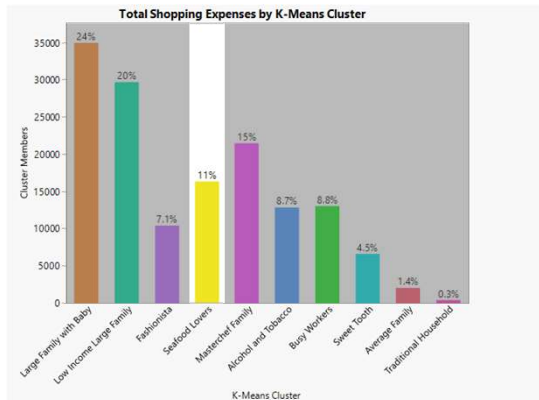
- Characteristics:
  - Total Income: ₱153,319
  - Shopping Expenses: ₱74,061
  - 20% of the sample
- Preferred Items:

The second identified segment is low income large family which has a slightly higher median family size (6) compared to large family with baby, and lower median for total income and shopping expenses. They make up 20% of the sample but only contributes to 18% of the overall total expenses. Their most preferred items is bread, making up a very large percentage in their shopping expenses.

Seafood Lovers

- Family Size

- Characteristics:
  - Total Income: ₱143,852
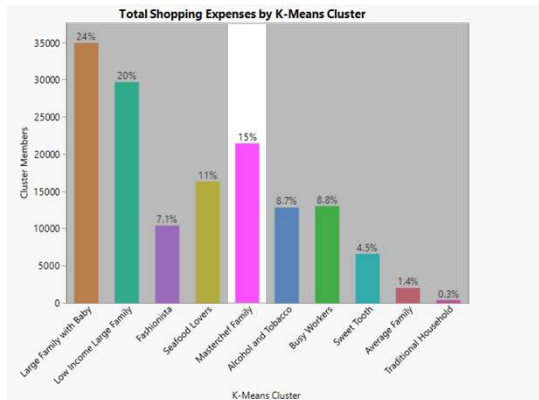 - Shopping Expenses: ₱71,262.5
  - 11% of the sample
- Preferred Items:

Seafood lovers make up 11% of the sample and contributes to 9.5% of the overall total shopping expenses. They display a higher tendency to spend on fish and seafood compared to the other segments. The median family size, total income, and shopping expenses are 4, ₱143,852, and ₱71,262.5 respectively, all slightly lower than low income large family segment.

Icons made by https://www.flaticon.com/authors/iconixar

Masterchef Family

Total Shopping Expenses by K-Means Cluster

- Family Size

- Characteristics:
  - Total Income: ₱141,364
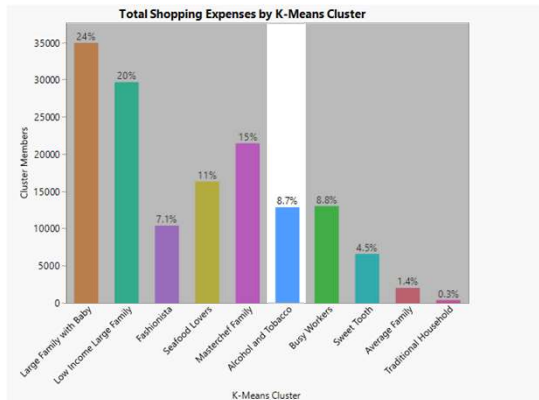 - Shopping Expenses: ₱53,487
  - 15% of the sample
- Preferred Items:

Masterchef family has the lowest median shopping expenses (₱53,487) but it still contributes to 9.4% of the overall total shopping expenses, which is higher than Alcohol and Tobacco segment and Sweet Tooth Segment. This is due to the fact that they make up a higher percentage of the sample (15%) compared to the latter two segments.

The lower shopping expenses is probably due to the smaller family size (3). However, they seem to be cooking a lot more than the other segments because their proportion of expenditure on cooking ingredients such as oil, sauce, vegetable, and fruits are higher.

Alcohol and Tobacco
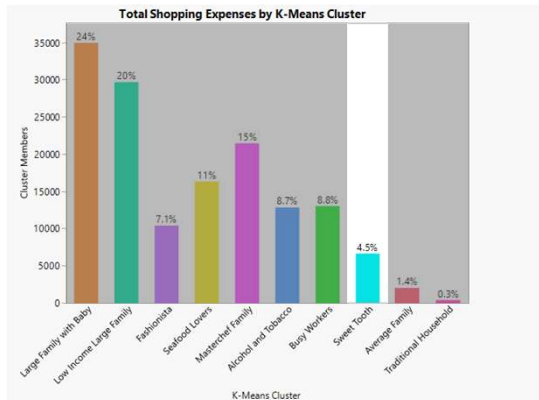

- Family Size

- Characteristics:
  - Total Income: ₱210,435
  - Shopping Expenses: ₱84,388
  - 8.7% of the sample
- Preferred Items:

The alcohol and tobacco segment has the 3rd highest median total income with ₱210,435, and the 2nd highest median shopping expense (₱84,388) but they are only 8.7% of the sample, hence not contributing a lot to the overall total of shopping expenses. Their purchase on alcohol and tobacco has a larger proportion compared to the other segments, often being the category of item that they spend the most on.

Sweet Tooth

- Family Size
- Characteristics:
  - Total Income: ₱220,009
  - Shopping Expenses: ₱81,669
  - 4.5% of the sample
- Preferred Items:

Sweet Tooth also has a median total income of ₱220,009, ranking 2nd out of the other customer segments, but a low contribution to overall total of shopping expenses due to the low prevalence in the sample. They allocate more spending on sugar, chocolate, jam, and sweets compared to the other segments.
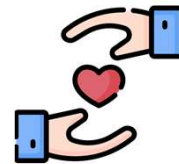
Recommendations

Act on the profiles to inform **market development strategy**

Collect **quantity** of the items purchased

Create **loyalty programs**

Recommendations based on the results of this analysis are:
1. Collect the quantity of the items purchased if further analysis is required for the market development strategy. This data allows us to calculate the average price of items the customers bought in each category. A high shopping expense may mean they purchase a bulk of inexpensive items, or a few of expensive items. This can help to inform pricing.
2. Act on the profiles to inform market development strategy if the current analysis is deemed sufficient. The basic characteristics of family size, income, shopping expenses, and distribution of expenses between the categories are included in this analysis.
3. After the retail is established in the Philippines, loyalty programs should be launched to keep track of the customer's purchase. Loyalty programs provide data on the items that customers purchase which can be further analyzed in market basket analysis to encourage cross-selling and up-selling, or RFM analysis to determine potential loyalty.

Basket icon created by https://www.flaticon.com/authors/photo3idea-studio