Department of Economics
Data Analytics and Business Economics
DABN01: Master Thesis
May 2023

Segmenting Countries in the Food Packaging Market: A Cluster Analysis Approach

Authors:
Aksaraphak Nijsathorn
Ninda Atikah Mawaddah
Supervisor:
Joakim Westerlund

Acknowledgements

We would like to express our deepest appreciation to our supervisor, Joakim Westerlund, for the invaluable guidance and insightful suggestions, which greatly contributed to the improvement of this thesis. Our appreciation extends to Niklas Rystedt for the valuable advice and comments throughout the paper. Furthermore, we would like to thank our friends and family for supporting us in the process of writing this thesis.

Abstract

The food packaging market has been witnessing substantial growth due to changing consumer lifestyles, urbanization, increased purchasing power, and growing environmental sustainability awareness. This growth presents significant opportunities for companies operating in this market. To effectively capitalize on these opportunities, it is crucial to develop marketing strategies and forecasting analysis that cater to diverse consumer demands across different countries. In this master's thesis, clustering techniques are employed to segment countries in the food packaging market and identify distinct groups of countries based on product packed group, packaging size, and macroeconomic factors.

The primary objective of this study is to utilize unsupervised machine learning algorithms, specifically K-means and Hierarchical Clustering, to cluster countries or markets according to packaging product and size for the company in the food packaging industry. The findings indicate that K-Means with six clusters yields a higher Silhouette Score compared to Hierarchical Clustering. Moreover, an analysis of clustering trends from 2015 to 2019 reveals a consistent pattern in country clusters during the period of 2017 to 2019, signifying stability and similarity in country characteristics and packaging volumes. However, variations are observed in the clustering patterns of 2015 and 2016, suggesting distinct country characteristics and package volumes during those years. These findings emphasize the importance of considering temporal trends and dynamics when interpreting clustering results and understanding country characteristics and packaging volumes

Keywords: Unsupervised Machine Learning, Clustering, K-Means Clustering, Hierarchical Clustering, Principal Component Analysis

Table of Contents

List of Tables

List of Figures

1. Introduction

The food packaging market has been experiencing significant growth in recent years, driven by multiple factors. Changing consumer lifestyles have driven the demand for innovative and efficient packaging solutions. Increasing pace of urbanization has also contributed to the growth. According to the research by Versino (2023), by 2050 the world's population is estimated to reach 9.7 billion, more than 60% of the population living in urban areas tend to increase in food requirements and change in food consumption patterns. Moreover, rising income levels have led to higher purchasing power, enabling choices for consumers to select for their preferences packaged foods. In addition, the growing awareness of environmental sustainability, leading to the adoption of eco-friendly packaging materials and practices (Mahmoud et al., 2022). Based on a market forecast by Mordor Intelligence (2021), the food packaging market is projected to exhibit a Compound Annual Growth Rate (CAGR) of 5.5% from 2023 to 2028. This indicates a significant growth opportunity for industries operating within this market and the potential for increased value creation.

As the global market continues to expand, it becomes increasingly important for companies to gain insights to develop effective marketing strategies and forecasting analysis that meet diverse consumer demands across different countries. In line with this objective, the present master's thesis seeks to employ clustering techniques to segment packaging's markets and identify distinct groups of countries based on product characteristics, features, and macroeconomic factors including GDP per capita, inflation rate, unemployment rate, and urbanization ratio.

Clustering, a statistical technique used to group similar markets based on shared characteristics (Akay and Yüksel, 2017), will be employed in this study to segment the food packaging market by country. The primary objective is to identify and define the similarities and unique characteristics of each cluster, together with various aspects such as market movement and dynamics. The derived insights will facilitate the development of customized marketing strategies and predictions that effectively cater to the specific needs and preferences of each market segment.

This research goes beyond identifying country clusters by aiming to determine the importance of variables that influence cluster formation. The study involves analyzing the relative impact of macroeconomic factors, such as GDP per capita, inflation rate, and unemployment rate, on each cluster. By understanding the variable importance within each cluster, the company can gain insights into the key drivers of market demand and incorporate market characteristics to enhance its market competitiveness and profitability through the customization and localization of its products portfolio and marketing efforts.

2. Literature Review

The food packaging market displays diverse trends across various regions, reflecting the complexity of the global market. These regional variations contribute to the challenges faced in the food packaging industry. To address this complexity, market segmentation serves as a valuable tool for companies to effectively identify and target specific customer groups or countries. By customizing product portfolios and implementing tailored marketing strategies, companies can enhance customer satisfaction and improve profitability. Unsupervised machine learning, particularly clustering algorithms, offers a powerful approach for market segmentation. These algorithms enable the identification of patterns and similarities within intricate datasets, aiding in the understanding of customer behavior and preferences. To support this discussion, relevant literature reviews will be provided.

2.1 Global Food Packaging Market

Food packaging plays a critical component of the food industry as it serves various purposes such as safeguarding food from contamination, maintaining freshness, conveying essential information about ingredients, nutritional content, and storage instructions, as well as facilitating food transportation (Marsh and Bugusu, 2007). The global food packaging market is projected to grow with an estimated value of USD 440.3 billion by 2025 (Markets and Markets, 2020) with a CAGR of 5.5% (Mordor Intelligence, 2021). This growth can be attributed to several factors, including the increasing demand for convenience food, the growing popularity of sustainable packaging materials and rising personal income. Based on the increase of consumer preference for convenience and on-the-go food options, these significantly impact the food packaging industry. Changing lifestyles and eating habits as well as the need for portion control have altered the demand for single-serve packaging. Furthermore, sustainability has become a focus area in the food packaging market. With rising environmental awareness and concerns over plastic waste, sustainable design of packaging such as biodegradable, compostable, and recyclable packaging options are driven by consumer preferences for eco-friendly choices and regulatory initiatives promoting sustainable practices (Brennan et al., 2021).

The global food packaging market exhibits diverse trends and characteristics across different markets and continents, reflecting regional trends and market characteristics. In North America, advancements in technology, stringent regulations, and growing focus on sustainable packaging solutions drive the food packaging market (Grand View Research, 2021). Europe places a strong emphasis on sustainability and eco-friendly practices to reduce waste (European Economic and Social Committee, 2022). Additionally, the market is anticipated to grow from the change in consumer lifestyle and easy packaged food availability (Fortune Business Insights, 2019). While in the Asia Pacific region, the market tends to follow the growth of population expansion, urbanization, and evolving consumer preferences for convenient packaging products. Growing consumer demand for environment-friendly packaging with low cost may drive the demand in the region (Fortune Business Insights, 2019). Latin America focuses on the blooming food and beverage industry and differentiate products through innovative packaging designs and materials (Transparency Market Research, 2021). Africa can be attributed to the consumer's lifestyle and demand for specific products (Fortune Business Insights, 2019). These findings are beneficial in understanding the regional variations and dynamics, enabling companies to operate effectively and develop tailored strategies that align with specific consumer demands in different markets.

2.2 Market Segmentation

Market segmentation is crucial for companies as it allows them to identify and target specific customer groups by customizing and tailoring their product portfolios (Wedel and Kamakura, 2012). Products can no longer be produced and sold without considering customer needs. By focusing on the preferences and needs of consumers within distinct market segments, companies can enhance customer satisfaction, value, and profitability. Several studies highlight the importance of market segmentation in various industries. For example in the telecommunications industry, Bayer (2010) demonstrated the successful implementation of customer segmentation focusing on customer value, customer behavior, customer life cycle and migration. The result enabled precise targeting and enhanced business planning which can retain the number of consumers in the long term. Similarly, in the transportation sector, according to the study by Teichert et al. (2008), the authors emphasized the segmentation approach to identify groups of passengers along behavioral and socio demographic factors. Therefore, companies can tailor product packages and develop specific marketing strategies according to customers' preferences. Another study focused on the food industry (Verain et al., 2016), conducted a study focusing on the food-category attribute importance for consumer segmentation. They highlighted the importance of customers' perception of healthiness and sustainability of food products in creating distinct market segments. These segmentation efforts contributed to the development of effective policies and successful marketing strategies.

2.3 Unsupervised Machine Learning for Clustering Approach

Unsupervised machine learning is a powerful approach of machine learning in which an algorithm learns patterns in the data without the need for explicit supervision. One of the most popular applications of unsupervised learning is clustering (Popat, 2014). It is also a common technique for statistical analysis which is applied in various fields such as pattern recognition, image analysis, and bioinformatics. The advantage of clustering analysis is that it finds groups in data and produces reasonable grouping results and classifies the most similar series together (Akay and Yüksel, 2017). This technique is particularly valuable for market segmentation, as it can reveal patterns and similarities within large and complex datasets that would be challenging for humans to identify manually.

Several popular clustering algorithms have been employed in customer or market segmentation studies. K-means was implemented, partitioning the data into K distinct clusters by minimizing the within-cluster sum of squares (Kansal et al., 2018). Hierarchical clustering, on the other hand, creates a tree-like structure of clusters, allowing for both agglomerative and divisive approaches (Everitt et al., 2011). Another partitioning algorithm, i.e. K-Prototype, deriving from K-means practice, facilitates datasets consisting of continuous and categorical data (Rajagukguk and Fudholi, 2022). Density-based clustering, such as DBSCAN (Ester et al., 1996), identifies dense regions of data points separated by sparser regions. These algorithms have been extensively utilized in customer segmentation studies across various industries.

3. Methodology

In market or customer segment clustering, two commonly used algorithms, K-means and Hierarchical clustering (Abdulhafedh, 2021), have been selected in this study for the purpose of comparing their results and taking advantage of their distinct characteristics that are suitable for addressing different aspects of market or customer segmentation. Additionally, other preprocessing steps were performed prior to clustering. Firstly, standardization was applied to scale the numerical input variables. Secondly, Principal Component Analysis (PCA) was employed for dimension reduction and capturing variable importance through PCA loadings. Finally, the quality of clustering was determined and assessed using the Silhouette Score.

In the context of country segment clustering, the dataset used in this study consists of panel data with a combination of numerical and categorical variables. Prior to applying the clustering approach, the categorical variable, which is country, is transformed into dummy variables to ensure compatibility with the chosen methods. This preprocessing step enables the utilization of Principal Component Analysis (PCA) for dimension reduction. In a related analysis of customer segmentation by Abdulhafedh (2021), the PCA approach was applied in combination with K-means clustering.

To gain a deeper understanding of the dynamic nature and movement of clusters, the analysis is conducted year-by-year. Additionally, the analysis includes an examination of the differences in the yearly clustering results. This result highlights that countries are independent entities within the market, and their focus tends to shift from year to year in response to market changes (Ao and Wei, 2022). This approach facilitated the examination of changes and trends within clusters over time. Additionally, by performing the analysis on a yearly basis, the possibility of countries being assigned to multiple clusters, thereby avoiding duplication, is addressed.

The ultimate goal of the analysis is to determine the clustering of countries for each individual year. By considering the data, this process aims to reveal the specific clusters to which each country belongs during a particular year. This approach provides valuable insights into the segmentation of countries, allowing for a more understanding of their characteristics and behaviors over time. Grein (2010) conducted a comparative analysis of country clustering using data from 1995, 2000, and 2005 to examine how countries shifted their positions within clusters over the course of those years. The study focused on exploring the implications of these dynamic country clusters in relation to corruption and global firms. In another study, Scutariu et al. (2021) investigated the impact of the COVID-19 pandemic on the behavior of enterprises' e-commerce activity using Ward's method clustering. The researchers compared the movement of countries among clusters between 2018 (pre-pandemic) and 2020 (during the pandemic).

3.1 Standardization

Standardization is a preprocessing step in machine learning that aims to ensure that features are measured on the same scale. Standardization can improve the performance of machine learning models by making them less sensitive to the scale of the input features. Standardization is by using the mean and standard deviation in the training data (Lindholm, 2022):

$$x_{ij}^{new} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}, \qquad j = 1, \dots, p, \quad i = 1, \dots, n$$

where p is number of variable and observations i.

3.2 Principal Component Analysis (PCA)

According to Everitt (2011), the fundamental objective of principal components analysis (PCA) is to represent the variability in a set of correlated variables, denoted as $X_1, X_2, \dots, X_p$ using a new set of uncorrelated variables, denoted as $Z_1, Z_2, \dots, Z_p$. These new variables are obtained by creating linear combinations of the original X variables. The process involves deriving the new variables in a descending order of "importance," where Z_1 captures the maximum variation among all linear combinations of $X_1, X_2, \dots, X_p$. Subsequently, Z_2 is chosen to account for as much remaining variation as possible while being uncorrelated with Z_1, and so forth. The resulting new variables, $Z_1, Z_2, \dots, Z_p$, are referred to as the principal components.

It is often recommended to standardize the data before conducting PCA, particularly when the variables being analyzed differ significantly in their scales. Failing to standardize the data may introduce bias and favor variables that are measured in units with dominant scales. By standardizing the data, the influence of scale differences is minimized, ensuring that each variable contributes equally to the PCA process. (Lindholm, 2022)

Determining the number of components to retain in PCA can be a challenging task. One approach is to utilize the concept of Cumulative Percent Variance (CPV). CPV quantifies the percentage of variance captured by the first l principal components. The selection of a specific CPV value is subjective in nature. While there is a desire to account for a significant portion of the variance, it is also important to minimize the number of retained principal components. Therefore, the decision ultimately involves striking a balance between model parsimony and comprehensiveness, considering both the desire to capture variance and the need for a concise representation. (Sergio Valle, Weihua Li, and S. Joe Qin, 1999). A study by Suhr (1999) suggests that a cumulative proportion of variance explained between 70% and 80% is often considered acceptable.

3.3 K-Means Clustering

One of the most popular clustering algorithms is K-Means. Its primary objective is to classify observations into mutually exclusive groups, or clusters. The basic idea of K-Means is to maximize the similarity among observations within the same cluster while maximizing dissimilarity between observations belonging to different clusters. In K-Means clustering, each cluster is characterized by its centroid, which represents the mean of the observation values assigned to that particular cluster (Abdulhafedh, 2021).

The K-means algorithm is an iterative clustering method that operates based on a distance metric. Given a dataset with K classes, it calculates the mean distance to establish initial centroids, with each class represented by its respective centroid. In the case of a dataset X comprising n multidimensional data points and the objective of dividing it into K clusters, the Euclidean distance is commonly employed as a similarity measure. The clustering process aims to minimize the sum of squared distances between data points and their assigned centroids;

$$d = \sum_{k=1}^{k} \sum_{i=1}^{n} \|x_i - u_k\|^2$$

$$u_k = \frac{1}{n} \sum_{i=1}^{n} x_i$$

where *k* represents *K* cluster centers, u_k represents the *k*-th center, and x_i represents the *i*-th point in the data set (Yuan and Yang, 2019).

In the K-means clustering algorithm, it is necessary to specify the number of clusters before conducting the modeling process. One common approach for determining the optimal number of clusters is the Elbow Method. The Elbow method is a well-established approach for determining the optimal number of clusters in a given dataset. It involves computing the sum of squared distances between each data point in a cluster and its corresponding centroid, resulting in a series of K values. The sum of squared errors (SSE) is then calculated as a performance metric. Smaller SSE values indicate higher convergence within each cluster. When the number of clusters approaches the true number of underlying clusters, the SSE exhibits a rapid decrease. However, as the number of clusters exceeds, the SSE continues to decline at a slower rate (Yuan and Yang, 2019).

The optimal cluster number K is identified by the fact that the SSE rapidly reduces to the called cost peak value before reaching K, and after surpassing K, it continues to climb with the called cost peak value practically unaltered. Meanwhile, the best cluster number corresponding to the elbow point is determined by a deliberate choice (Shi et al., 2021).

3.4 Hierarchical Clustering

One of the conventional clustering algorithms is hierarchical algorithm (Akay and Yüksel, 2017). Two types of hierarchical algorithms exist: divisive hierarchical and agglomerative hierarchical algorithms. In the divisive approach, the algorithm starts from the top and proceeds downwards. Initially, there is a single large cluster encompassing all the data points, and the algorithm progressively splits clusters during the process. In the general case, one of the most widely used algorithms is agglomerative clustering (Popat et al., 2014). The agglomerative hierarchical algorithm operates in a bottom-up manner. Initially, each data point is treated as an individual cluster, and the algorithm progressively merges these clusters together (Akay and Yüksel, 2017).

Agglomerative hierarchical clustering starts with every single object in a single cluster. If there are N items, there will be N clusters, each containing one item (Popat et al., 2014). In each iteration, the algorithm merges the closest portions of clusters that meet specific similarity criteria, gradually combining the data until all of it resides within a single cluster (appendix B). The algorithm offers the benefit of generating an object ordering, which can be valuable for data presentation. To determine the merging of clusters in the agglomerative approach or the splitting of a cluster, it is necessary to employ a dissimilarity measure between observation sets, using an appropriate metric which is a distance measure between observation pairs, together with a linkage criterion that defines set dissimilarity based on pairwise distances among observations within the sets (Sasirekha and Baby, 2013). Some commonly used hierarchical agglomerative methods include single linkage, complete linkage, and average linkage (Šulc and Řezanková, 2019). Ward's method is also another practice, which is called the minimum variance method (Ward, 1963).

3.5 Silhouette Score

The Silhouette Score is a commonly employed metric for assessing the quality of clustering results. It enables the comparison of outputs from different clustering algorithms. (Rousseeuw, 1987). The Silhouette Score ranges from -1 to 1, with values closer to 1 indicating a favorable clustering outcome, while values closer to -1 suggest that a data point might be assigned to an incorrect cluster. The Silhouette Coefficient is measured as follows:

$$S_i = \frac{b_i - a_i}{max\{a_i, b_i\}}$$

where a_i is the average distance of observation i from all other observations in its cluster and b_i is the smallest average distance of i to all observations in any other cluster. To clarify, b_i is found by measuring the average distance of i from every observation in cluster A, the average distance of i from every point in cluster B and taking the smallest resulting value (Rousseeuw, 1987).

4. Data

4.1 Data Sources

The data in this thesis is from diverse sources. Specifically, the internal dataset comprises beverage packaging volume data from 114 countries between 2015 and 2021, including various features such as date, market/country, packed product group, process, distribution, category, sub-category, package size group, package type group, and total volume which is forecasted by a research company. The study aims to cluster countries based on their packaging volume and macroeconomic factors as well as analyze the trend of these clusters over time. To provide a comprehensive overview of country characteristics, supplementary data from open data sources (external data) which are the World Bank and the International Monetary Fund (IMF) is incorporated into the analysis.

Unemployment rate, inflation, urbanization rate, GDP per capita and consumption per capita (total volume over total population) are included in this study. Unemployment rate defines the labor market, assuming that the unemployment rate implies the overall purchasing power of consumers. Urbanization ratio indicates a view of infrastructure needs, and socio-economic patterns. The factor is linked to the population density. Regions with high urbanization ratios, may have a larger number of potential customers or consumption. GDP per capita is included in this study to provide a more detailed understanding of economic conditions in different countries, as it takes into account the economic well-being of each individual, rather than just the overall economic activity of the country. The consumption per capita metric, which is calculated by dividing the total packaging volume by the population size, offers valuable insights into the consumption patterns of individuals in various countries. By examining both GDP per capita and consumption per capita, this study is able to provide a more comprehensive evaluation of economic and consumption behavior in different countries.

4.2 Data Processing

This study is centered around three primary objectives, namely, the comparison of algorithms to identify the optimal model and the appropriate number of clusters, the analysis of cluster characteristics to determine the differences between clusters, and the tracking of cluster movements over time. The insights gained from this analysis can be leveraged to develop targeted marketing strategies that cater to the unique needs and preferences of each market segment, in this case the country segment.

To ensure the analysis reflects the current market situation and provides relevant insights, the selected period of data for analysis spans from 2015 to 2019. This five-year period was chosen to focus on the most recent data, excluding the years 2020 and 2021, which were significantly impacted by the Covid-19 pandemic. By excluding the pandemic years, the analysis can avoid distortions caused by exceptional circumstances and provide a clearer picture of the market's normal trends and patterns. This approach allows for a more accurate understanding of the market dynamics during a period closer to the present, enabling researchers and industry professionals to make informed decisions based on up-to-date information.

The internal variables that included in this analysis are year, market/country, packed product group, package size group, and total volume. These variables are essential for examining the characteristics and trends within the dataset. However, certain variables such as process, distribution, category, sub-category, and package type group have been deemed unnecessary for the current analysis. These variables primarily provide more detailed descriptions of the packed product itself, rather than capturing broader market characteristics or patterns. The variable "process" specifies the specific production process employed for a particular product, while "category" and "sub-category" further classifies the type or nature of the product. Similarly, the variable "distribution" indicates the distribution channels utilized for the product, and "package type group" describes the specific packaging format or design. These variables inclusion in the current analysis may introduce unnecessary complexity without significantly enhancing the understanding of the market dynamics or clustering patterns. Thus, to streamline the analysis and focus on the key variables of interest, they have been removed from the dataset.

To enhance the analysis and obtain country clusters based on total volume per product and country characteristics, the analysis focuses on internal variables, including year, market/country, packed product group, package size group, and total volume. In order to capture a more detailed view of the data, the packed product group and package size group variables are combined to create a comprehensive set of new variables.

Specifically, the combinations of packed product and size groups resulted in a total of 102 distinct categories (detailed in Appendix A). These combinations represented different combinations of product types and their corresponding sizes. By transposing these combinations, each cell becomes a new variable representing the total volume associated with a specific packed product and size group.

This transformation expands the scope of the analysis, enabling a more comprehensive exploration of the data. The resulting 102 new variables provide a granular view of the total volume per packed product and size group, facilitating a more detailed examination of market dynamics and clustering patterns. By considering the total volume across different packed product and size groups, this approach aims to capture the variations in identifying distinct clusters based on both product characteristics and country-specific factors. This enrich perspective on the data allowed for a more nuanced understanding of the relationships between different variables and their impact on market segmentation. Lastly, country characteristics or macroeconomic factors (external data) are added to the data set.

Figure 1 shows the overview of how this analysis is conducted from data processing to interpret the variable importance for clustering. The data is splitted by year to analyze the cluster movement. Categorical variable, “country”, is changed to dummy variables while numerical variables are standardized to ensure uniformity in their measurement scales.

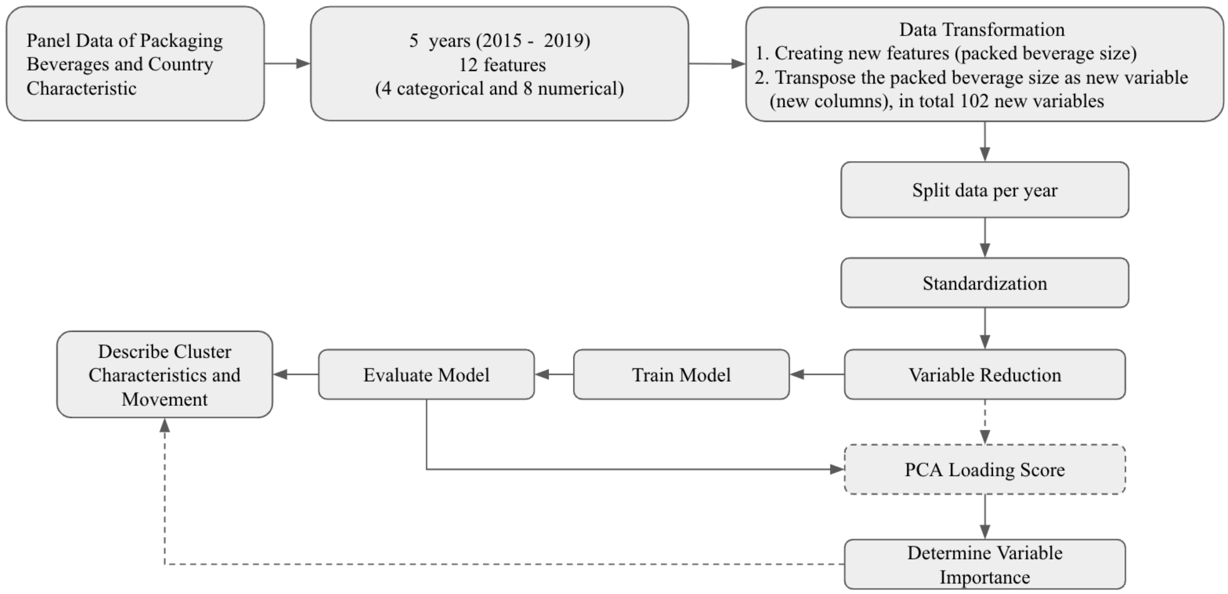


Figure 1. Study Flow Chart

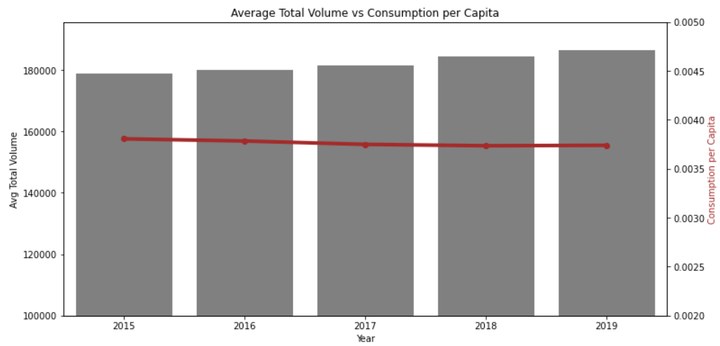
4.3 Descriptive Statistics Data

In this section, a descriptive analysis of the data is conducted to gain an overview and deeper understanding of the dataset. The analysis focuses on the volume variable and certain economic factors. Specifically, data from the period between 2015 and 2019 was exclusively utilized to examine the trends.

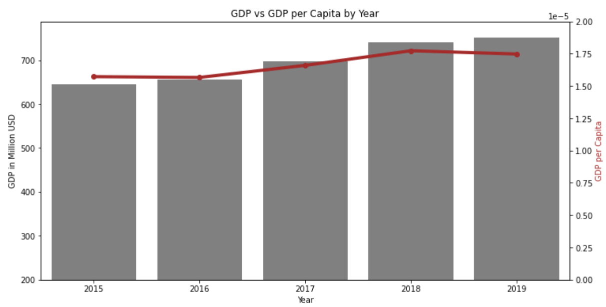
Figure 2.a. presents the trend of total volume and consumption per capita. The total volume demonstrates a gradual increase since 2015, indicating a growing market. On the other hand, consumption per capita experienced a slight decline year-by-year, suggesting potential changes in individual consumption patterns.

Figure 2.b. illustrates the trends of GDP and GDP per Capita. These two indicators display a similar pattern, reflecting the overall economic performance during the selected period. Both show a consistent trend, which can be further analyzed to understand the relationship between economic growth and the market under consideration.

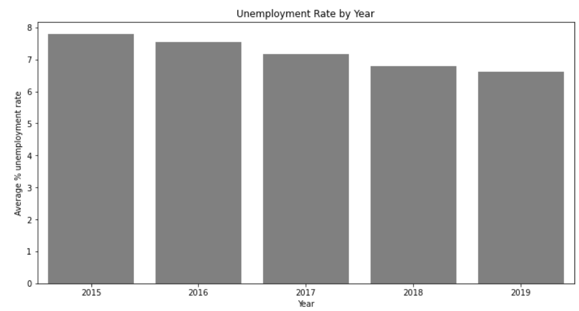
Figure 2.c. focuses on the Unemployment Rate. It demonstrates a positive impact by showcasing a decline from 2015 onwards. This decrease in the Unemployment Rate indicates improving economic conditions, which can potentially influence market dynamics and consumer behavior.



(a) Trend of Total Volume and Consumption per Capita



(b) Trend of GDP and GDP per Capita



(c) Trend of unemployment rate

Figure 2. Trend of Country Characteristic and Economics Variables

By analyzing these variables and their trends, a comprehensive overview of the dataset can be obtained. These insights provide information for understanding the food packaging market dynamics and country characteristics in the next section.

5. Result & Discussion

5.1 Principal Component Analysis

A large number of variables, specifically 108 including country, are initially considered for the clustering analysis. However, utilizing a high number of variables in clustering can introduce challenges in terms of interpretability. The resulting clusters may become complex and difficult to comprehend, making it challenging to discern the characteristics or factors that differentiate one cluster from another. To address this issue, Principal Component Analysis (PCA) is employed. Abdulhafedh (2021) implemented PCA as a means to reduce the dimensionality of the dataset. Specifically, they applied PCA with 5 components, which allowed for the consolidation of 17 variables. By reducing the number of variables through PCA, the resulting clusters become more manageable and interpretable.

In this study, the determination of the number of components for PCA is based on the cumulative percent variance (CPV). The CPV represents the proportion of variance accounted for by the principal components in the model. Following the recommendation by Suhr (1999), a CPV value of at least 70% was considered desirable. Table 1 displays the CPV values for each principal component, and it can be observed that the cumulative percent variance reaches 70% at the 9th principal component in the PCA analysis. Consequently, PCA 9 is selected as the input for the clustering model, aiming to capture a significant portion of the data's variability. As a result, rather than employing all 108 variables as input for clustering, only 9 variables (PCA 1 - PCA 9) are included in the clustering model, ensuring a more concise and meaningful representation of the data.

Table 1. Cumulative Percent Variance (CPV)

PC	2015	2016	2017	2018	2019
1	26,58%	26,75%	26,95%	27,22%	27,51%
2	42,26%	42,36%	42,62%	43,17%	43,64%
3	48,63%	48,84%	49,16%	49,81%	50,37%
4	54,17%	54,29%	54,59%	55,20%	55,74%
5	58,99%	59,04%	59,28%	59,93%	60,52%
6	62,53%	62,62%	62,86%	63,47%	64,08%
7	65,50%	65,64%	65,80%	66,38%	67,03%
8	68,11%	68,16%	68,25%	68,85%	69,56%
9	70,42%	70,46%	70,53%	71,17%	71,88%
10	72,50%	72,49%	72,57%	73,22%	73,95%
…	…	…	…	…	…
108	99,97%	99,97%	99,96%	99,96%	99,96%

5.2 K-Means and Elbow Method

The primary clustering algorithm utilized in this study is K-Means, a widely used method for grouping observations into distinct clusters based on their similarity. In this analysis, a total of 9 principal component analysis (PCA) components are included as variables, representing a dimensionality reduction technique applied to the original dataset. Determining the appropriate number of clusters, denoted as K, is a crucial step in the K-Means algorithm. To address this challenge, the Elbow method is utilized, followed by an evaluation using the Silhouette Score.

In Figure 3, the Elbow graph for each year is presented, depicting the relationship between the number of clusters and the sum of squared errors (SSE). Although the graph exhibits a relatively smooth trend, a distinct point of inflection can be observed around K = 7. This suggests that beyond this point, the reduction in SSE becomes less significant. As a result, K = 6 clusters are selected as the initial configuration for this study, taking into account both the Elbow method and the goal of keeping the model simple.

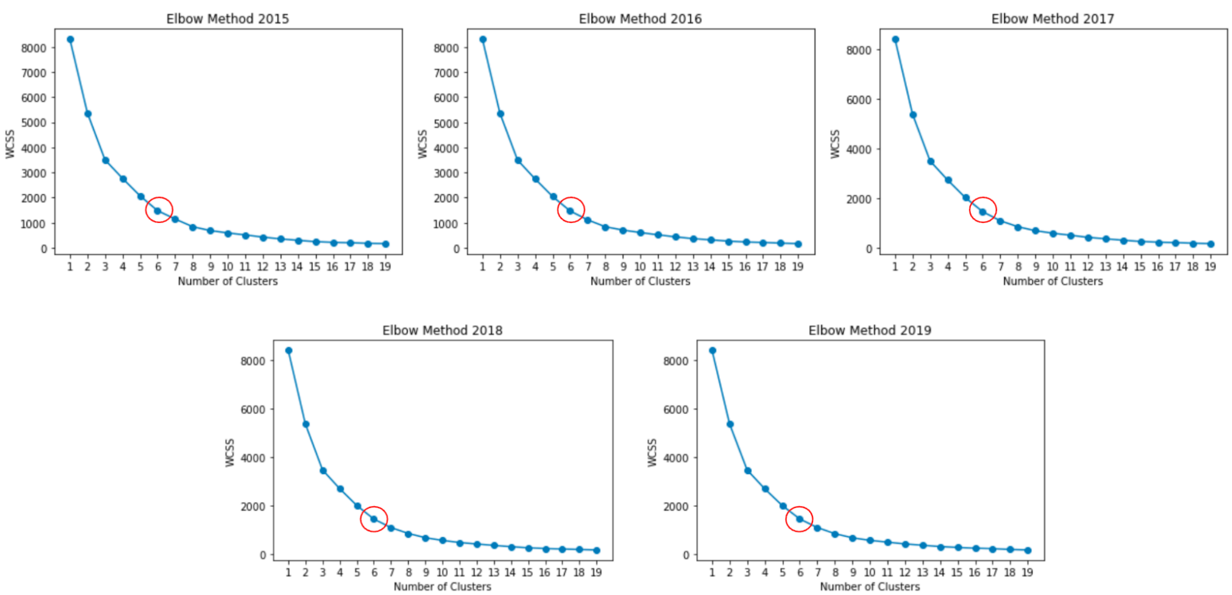


Figure 3. SEE of Elbow Method from 2015 - 2019

5.3 Modeling Comparison

In the previous section, the number of clusters was determined to be 6 using the K-Means algorithm. To further assess the suitability of this choice, a comparison was made between K = 6 and K = 7 clusters and also using another algorithm, Hierarchical Clustering. Table 2 presents the Silhouette Scores for both models across the period of 2015-2019.

Based on the results, K-Means with 6 clusters is higher than K-Means with 7 Clusters. Other than that, K-Means with 6 clusters yields the highest Silhouette Score compared to

Hierarchical Clustering. Therefore, the subsequent in-depth analysis in this study will focus on K-Means with 6 clusters.

Table 2. Silhouette Score Comparison of 6 and 7 Clusters

year	K-Means Clustering		Hierarchical Clustering	
	6	7	6	7
2015	0.745	0.645	0.659	0.655
2016	0.631	0.614	0.653	0.649
2017	0.594	0.595	0.623	0.620
2018	0.607	0.604	0.616	0.613
2019	0.623	0.608	0.625	0.622
Average	**0.640**	**0.613**	**0.635**	**0.632**

5.4 K-Means Result

In order to analyze the clusters derived from the K-Means algorithm with 6 clusters, which were discussed in the previous section, it is useful to visualize the data by creating a scatter plot of PCA 1 and PCA 2. This scatter plot, as depicted in Figure 4, provides insights into the underlying patterns and relationships between data points from 2015 to 2019 in lower-dimensional space.

The PCA 1 axis (x-axis), representing the direction of maximum variance in the dataset, captures the most significant source of variation in the data. On the other hand, the PCA 2 axis (y-axis), orthogonal to PCA 1, captures the second highest source of variation. By plotting the data points on the scatter plot based on their respective PCA 1 and PCA 2 scores, it can show the arrangement of the data points in the reduced-dimensions. The position of a data point on the scatter plot indicates its projection onto the lower-dimensional space defined by PCA1 and PCA2. The relative positions of data points provide insights into their similarities and dissimilarities based on the underlying patterns in the original variables.

The analysis of the scatter plots of PCA 1 and PCA 2 for each year reveals interesting trends and shifts in the clustering patterns. Overall, the plots exhibit a consistent trend across the years, indicating a stable clustering structure. However, in 2015, a notable deviation is observed, suggesting a distinct pattern or behavior in the data points for that specific year. Additionally, two data points stand out in the scatter plot, displaying significantly higher PCA 1 and PCA 2 scores compared to the other data points. These data points may represent unique cases or outliers.
In 2015, most of the data points fell into Cluster 2. A majority of the data points are concentrated within this cluster, indicating a higher similarity among the corresponding countries. However, as we move forward to the subsequent years, 2016 to 2019, the clustering patterns noticeable changes.

During this period, there is a discernible shift in the cluster assignments for certain countries. Specifically, some of the countries that were initially classified within Cluster 2 in 2015 transition to Cluster 5 in the subsequent years. This shift suggests a change in the similarity patterns and potentially reflects evolving dynamics or factors affecting those countries.

Additionally, in the year 2017, an intriguing shift is observed within Cluster 2. One data point, represented by the 'yellow' marker, changes in its coordinate position. This change indicates a shift in the characteristics or features of the corresponding country, resulting in its repositioning within the clustering structure. These findings highlight the dynamic nature of the clustering patterns over time and the potential influences that drive the shifts in cluster assignments.

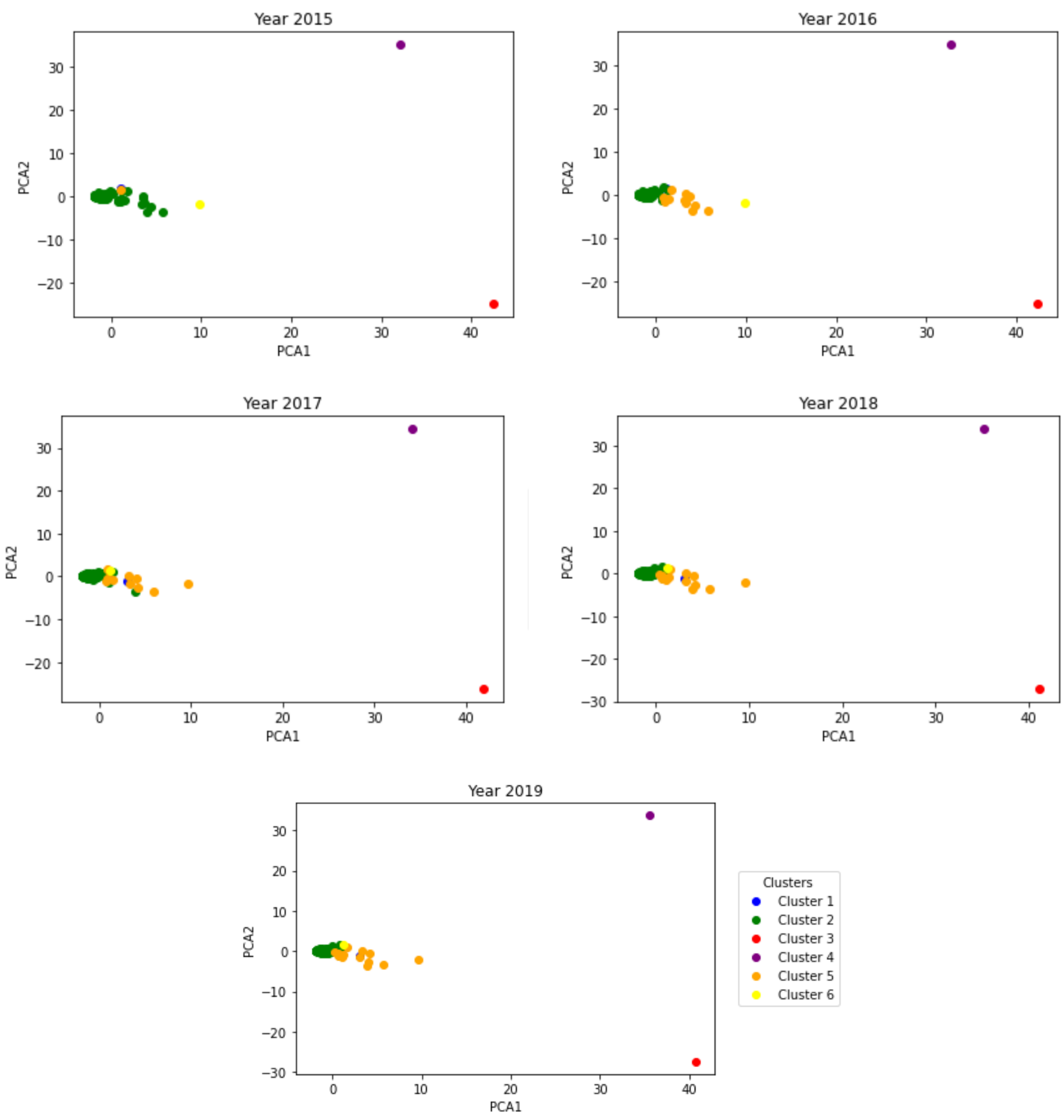
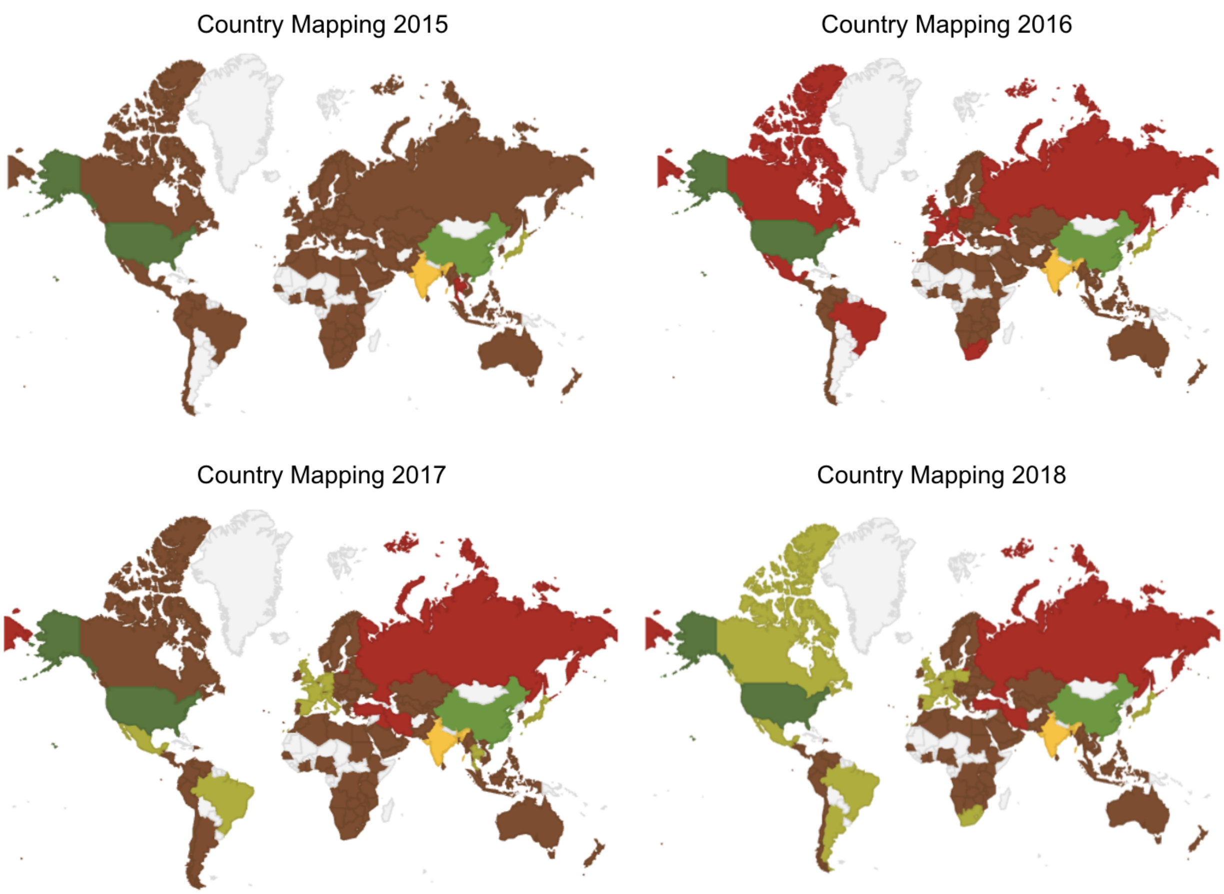


Figure 4. Plot of PCA 1 and PCA 2 from 2015 to 2019

To provide a visual representation of the country groupings, Figure 5 displays a country map illustrating the clustering results. As depicted in Figure 4, in 2015, a significant number of countries fell within Cluster 2. However, starting from 2016, the country clusters undergo changes, with certain countries transitioning to Cluster 1. The years 2018 and 2019 exhibit relatively similar groups of countries, indicating the similarity. For detailed information on individual countries and their specific characteristics is described in Appendix E.

Since in the last 3 years the trend showed a similarity, it becomes increasingly relevant for the company to utilize the clustering results from this specific time period for their strategic decision-making and forecasting. By leveraging the clustering outcomes, the company can effectively segment their target markets and tailor their strategies accordingly. The clustering results serve as a valuable input for feature vetting.

Applying the clustering results from the last three years enables the company to capture the most recent trends and dynamics in the market, allowing for more accurate needs within each cluster.



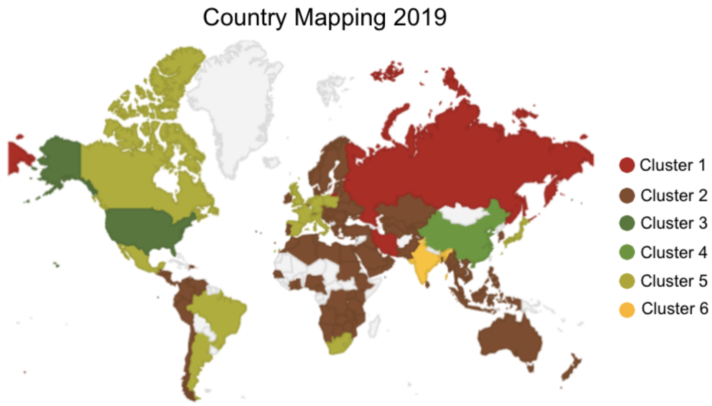


Figure 5. Countries Mapping by Cluster from 2015 to 2019

5.5 Variable Importance for Clustering - PCA Loadings

Using the 9 principle components, presented in Figure 6, the most significant loadings have been selected for visualization from the years 2015 to 2019. The bar graphs provide valuable insights into the contribution of features to the captured variation by each principal component, indicating their importance in understanding data patterns and clustering.

In the clustering analysis, it is observed that *Energy Drink with Portion Pack* has the highest PCA loading in PCA1, which significantly influences the clustering for the years 2015-2017. However, for the years 2018 and 2019, *Flavoured Milk with Portion Pack* exhibits the highest loading. Table 1 shows that PCA1 captures nearly 30% of the variation, indicating that these variables are crucial in differentiating the clusters and have a strong influence on their formation.

The study also finds that clusters sharing similar levels of the variable with high PCA loadings are more likely to be grouped together in the analysis. Additionally, PCA2, which captures approximately 16% of the variance according to Table 1, reveals that *Soy Milk with Loose* has the highest loading score. Both PCA1 and PCA2 exhibit loadings ranging from 0.17 to 0.19.

On the other hand, *Sweetened Condensed Milk with Large Size Containers* had the highest PCA loadings, ranging from 0.363 to 0.399 specifically in the years 2016-2019. Close behind are *Butter Milk with Loose* and *Traditional Cultured Drinks with Family Pack*. While these variables play an important role in specific components, they do not have a significant impact on clustering the market when compared to the first and second components.

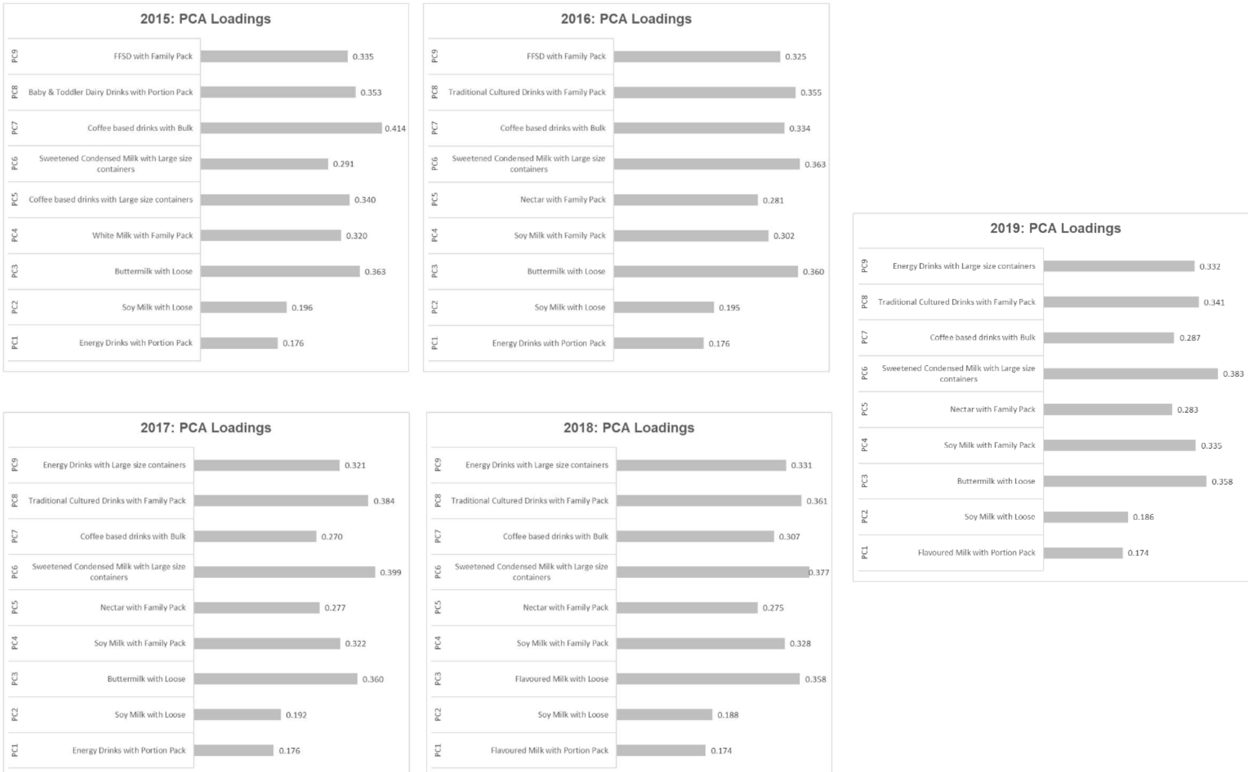


Figure 6. The Highest PCA Loading Score (PCA1 - PCA9) from 2015 to 2019

For a more detailed analysis of the PCA Loadings, particularly PCA1 and PCA2 (with PC3 - PC9 provided in appendix C), refer to Figure 7 and Figure 8. These figures highlight the consistent trends observed from 2015 to 2019, with a focus on the most recent year to gain deeper insights.

Figure 7 showcases the remarkable performance of *Tea Based Drink*, while Figure 8 highlights the significance of *Soy Milk*. Interestingly, *Flavoured Milk* also holds a prominent position, ranking among the top three variables with the highest loading in PCA1 in total, particularly when considering the *Portion Pack* size. Conversely, *Soy Milk* stands out among other packaging groups, with the *Loose* package size registering the highest loading. These findings align with the results presented in Figure 6, which demonstrates the high scores for both *Flavoured Milk* and *Soy Milk* in the package category.

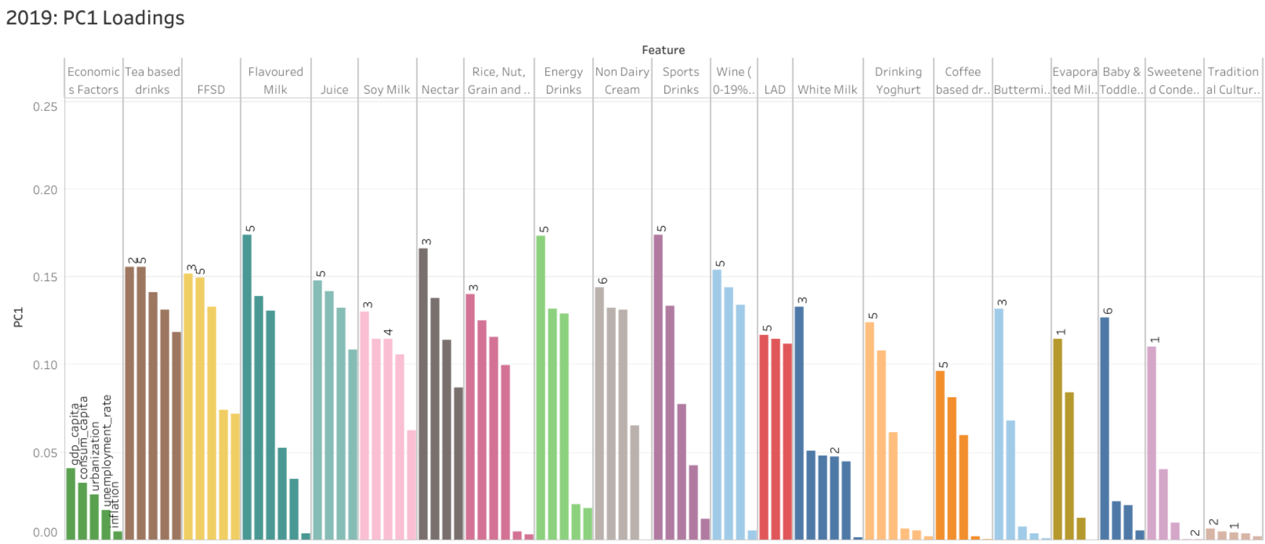


Figure 7. Illustrating PCA1 splitted by Packed Product Group. Overall, every variable in PCA1 has quite similar PCA Loading Score. In total, *Tea Based Drinks*, *FFSD* are the first and second highest PCA Loading Score.

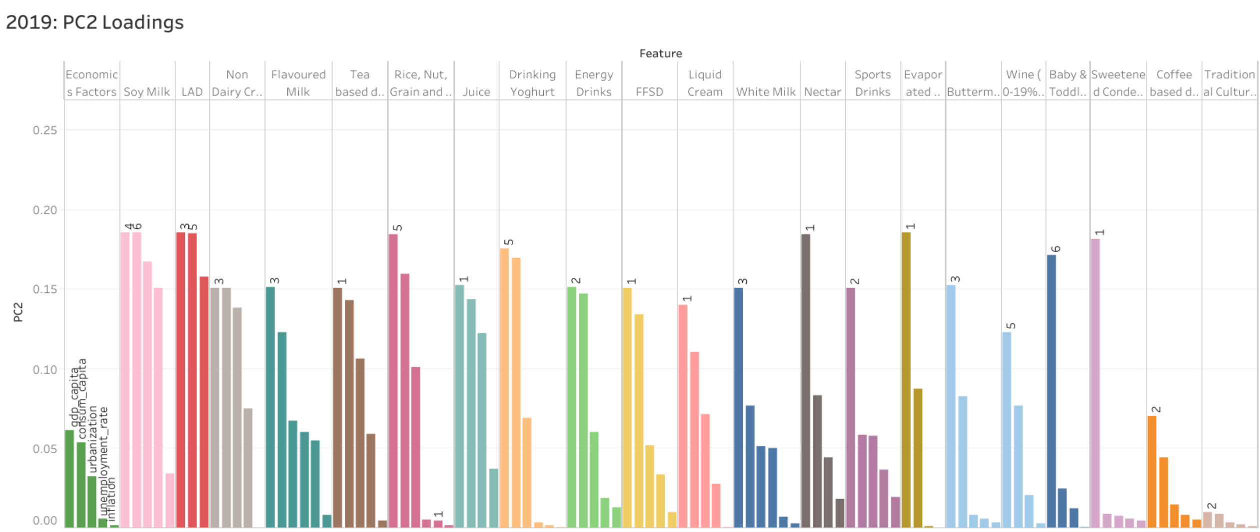


Figure 8. Illustrating PCA2. Overall, every variable in PCA2 has quite similar PCA Loading Score. *Soy Milk with Loose* size is the highest PCA2. In total, *Soy Milk and LAD* are the first and second highest PCA Loading Score.

Remark: Referring to Figure 7 and Figure 8, the index from 1 to 6 stands for *Bulk*, *Family Pack, Large Size Container, Loose, Portion Pack* and *Powder* respectively.

5.6 Descriptive Analysis of Clustering

In this section, the clustering results, characteristics, and dynamics will be explored and analyzed. By examining the distribution of market and movement within and between the clusters, the authors aim to gain a comprehensive understanding of the underlying patterns and trends. This analysis will provide valuable insights and behavior of the data.

Table 3. Total Number of Countries in each Cluster from 2015 - 2019

Cluster	2015	2016	2017	2018	2019
1	1	11	4	3	2
2	108	98	97	95	96
3	1	1	1	1	1
4	1	1	1	1	1
5	1	1	10	13	13
6	1	1	1	1	1
Total Country	**113**	**113**	**114**	**114**	**114**

Remark: Argentina did not have available data in 2015. The data has been included starting from 2016 onwards.

Cluster 2 encompasses over 80% of the markets (total country), indicating a significant concentration within this cluster. Certain movements can be observed in clusters 1, 2, and 5. Since 2017, these clusters have established a high degree of stability, suggesting consistent market dynamics within them. Only the United States of America, China, and India have been assigned their own distinct clusters, namely Cluster 3, Cluster 5, and Cluster 6, respectively. This highlights the unique market characteristics and dynamics of these influential economies.

Even though cluster 2 comprises more than 80% of the total country, it contributes approximately one-third of the market share, shown by Figure 16. Notably, the proportions of market share display a consistent trend across the years 2017 to 2019, while exhibiting variation in 2015 and 2016. Since 2017, Cluster 2 consistently contributes around 30% of the market share, closely followed by Cluster 5. Moreover, China and India, represented by their respective clusters (Cluster 3 and Cluster 4), make substantial market share contributions at 16% and 13% respectively. Conversely, Cluster 1 exhibits a relatively low average market share contribution of 4% over the last three years, which marks a notable decrease from the previous year's contribution of 22%. These findings highlight the varying market dynamics and the importance of understanding the market share distribution among the different clusters.

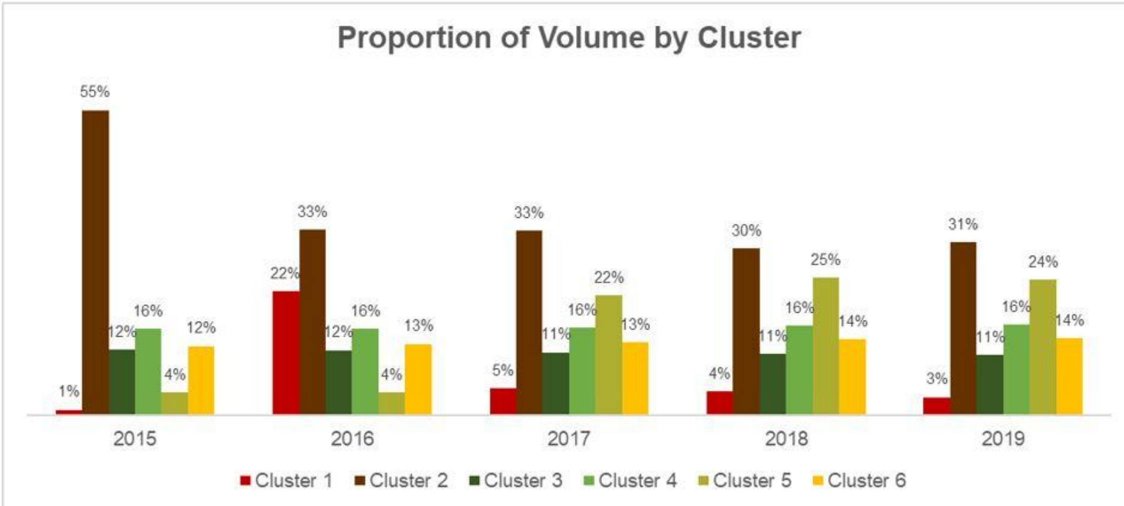


Figure 9. The Proportion of Volume by Cluster from 2015 to 2019

The analysis of volume differences by product package group is facilitated by the examination of the proportion of volume within each cluster, as outlined in Table 4. Notably, in both 2018 and 2019, the trend of proportion remains consistent across all clusters. It is observed that the highest proportion of volume is attributed to *White Milk* in all clusters, with the exception of Cluster 4. Within Cluster 4, *Soy Milk* holds the highest volume share at 15%, followed closely by FFSD at 14%. These findings highlight the dominance of *White Milk* as the primary product package group across clusters, while also emphasizing the unique preference for *Soy Milk* and *FFSD* within Cluster 4.

Table 4. The Proportion of Volume by Product Package Group

2018

Product Group	Cluster						
	1	2	3	4	5	6	Total
Baby & Toddler Dairy Drinks	3%	4%	1%	8%	1%	0%	3%
Buttermilk	0%	0%	0%	0%	0%	4%	1%
Coffee based drinks	0%	1%	2%	1%	3%	0%	1%
Drinking Yoghurt	1%	2%	0%	5%	2%	0%	2%
Energy Drinks	2%	3%	5%	4%	2%	0%	3%
Evaporated Milk	0%	1%	0%	0%	0%	0%	0%
FFSD	6%	14%	15%	14%	20%	4%	14%
Flavoured Milk	2%	5%	5%	4%	2%	2%	4%
Juice	3%	2%	9%	1%	6%	0%	3%
LAD	0%	1%	0%	9%	0%	0%	2%
Liquid Cream	1%	1%	1%	0%	2%	0%	1%
Nectar	9%	3%	2%	2%	4%	0%	3%
Non Dairy Cream	0%	1%	2%	2%	0%	0%	1%
Rice, Nut, Grain and Seed Based Drinks	0%	1%	2%	5%	1%	0%	2%
Soy Milk	0%	2%	0%	15%	1%	0%	3%
Sports Drinks	0%	1%	8%	2%	2%	0%	2%
Sweetened Condensed Milk	1%	1%	0%	0%	0%	0%	0%
Tea based drinks	3%	4%	12%	13%	8%	0%	7%
Traditional Cultured Drinks	17%	4%	0%	0%	0%	2%	2%
White Milk	46%	46%	29%	12%	35%	87%	42%
Wine (0-19% alcohol)	5%	4%	6%	3%	9%	0%	5%
Grand Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

2019

Product Group	Cluster						
	1	2	3	4	5	6	Total
Baby & Toddler Dairy Drinks	3%	3%	1%	9%	1%	0%	3%
Buttermilk	0%	0%	0%	0%	0%	4%	1%
Coffee based drinks	0%	1%	2%	1%	3%	0%	1%
Drinking Yoghurt	2%	2%	0%	5%	2%	0%	2%
Energy Drinks	3%	3%	5%	4%	3%	0%	3%
Evaporated Milk	0%	1%	0%	0%	0%	0%	0%
FFSD	5%	14%	15%	14%	20%	4%	14%
Flavoured Milk	2%	5%	5%	4%	2%	2%	4%
Juice	4%	2%	9%	1%	6%	0%	3%
LAD	0%	1%	0%	9%	0%	0%	2%
Liquid Cream	1%	1%	1%	0%	2%	0%	1%
Nectar	8%	3%	2%	2%	4%	0%	3%
Non Dairy Cream	0%	1%	2%	2%	0%	0%	1%
Rice, Nut, Grain and Seed Based Drinks	0%	1%	2%	4%	1%	0%	2%
Soy Milk	0%	2%	0%	15%	1%	0%	3%
Sports Drinks	0%	1%	8%	2%	2%	1%	2%
Sweetened Condensed Milk	1%	1%	0%	0%	0%	0%	0%
Tea based drinks	2%	4%	11%	14%	8%	0%	7%
Traditional Cultured Drinks	14%	4%	0%	0%	0%	2%	2%
White Milk	48%	46%	28%	12%	35%	86%	41%
Wine (0-19% alcohol)	6%	3%	6%	3%	9%	0%	5%
Grand Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

5.6.1 Characteristics of the 6 Clusters

To illustrate the distinctive characteristics of each cluster or country group, the histograms presented in Figure 17 provide an overview of the distribution of the 6 clusters based on Package Group Size, while Figure 18 further elucidates the characteristics of the clusters, encompassing *Coffee Based Drinks*, *FFSD*, *Soy Milk*, *Sweetened Condensed Milk*, *Tea Based Drinks*, *Traditional Cultured Drinks*, and *White Milk*.

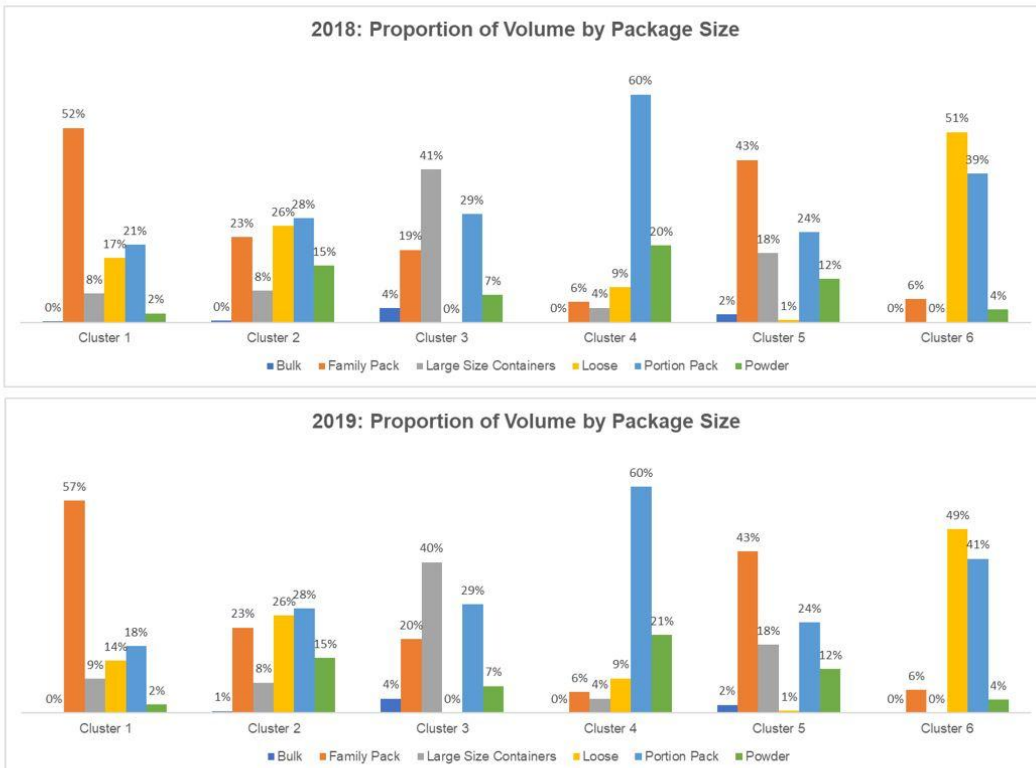


Figure 10. Volume Distribution by Packed Size

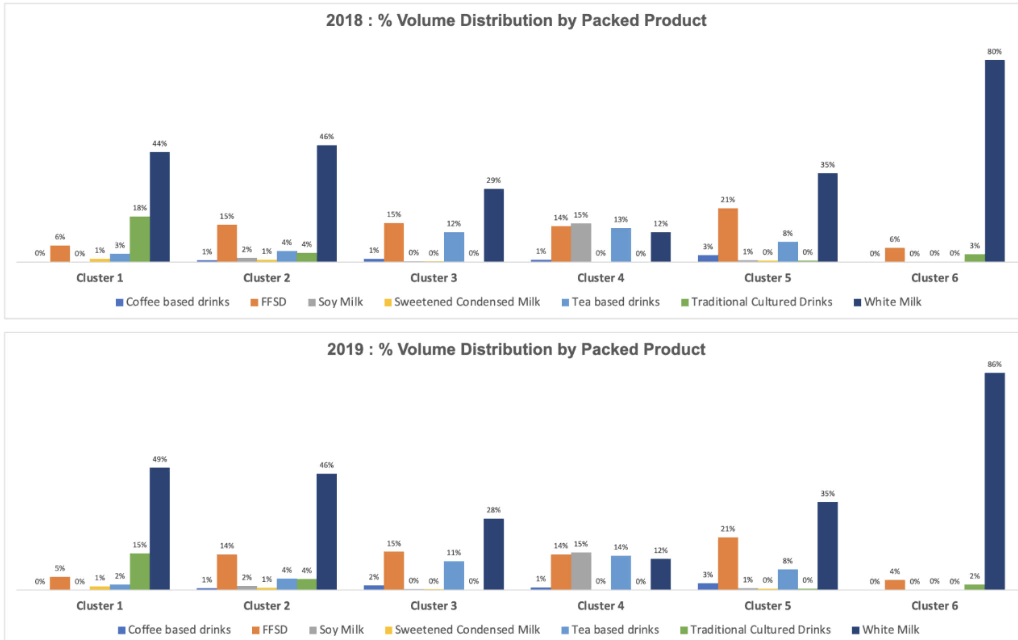


Figure 11. Volume Distribution by Packed Product and The Highest PCA Loading Score

Examining these figures reveals consistent patterns in the characteristics of the clusters across the years 2018 and 2019. The histograms offer valuable insights into the customer profiles within each cluster:

1. **Cluster 1** primarily consists of countries with a strong preference for *Family Pack* size and a notable inclination towards *Traditional Cultured Drinks*.
2. **Cluster 2** demonstrates a relatively balanced distribution of *Family Pack, Loose, and Portion Pack* sizes, with considerable consumption of *FFSD*. Moreover, it ranks second highest in terms of *Traditional Cultured Drinks.*
3. **Cluster 3** (United States) stands out with its significant consumption of package *Large Size Containers* and a complete absence of *Loose* packaging. This cluster also exhibits high consumption levels of *FFSD and Tea Based Drinks*.
4. **Cluster 4** (China) is distinguished by its preference for *Portion Pack* size and relatively lower consumption of *White Milk*. The consumption of *FFSD, Soy Milk, and Tea Based Drinks* contribute similarly to the overall volume distribution within this cluster.
5. **Cluster 5** showcases a substantial inclination towards *Family Pack* size, with minimal contribution from *Loose* packaging. This cluster also exhibits the highest distribution of *FFSD* among all clusters.
6. **Cluster 6** (India) displays a prominent consumption of *White Milk in Loose* package size *and Portion Pack* size.

By analyzing the histograms and discerning the specific characteristics associated with each cluster, valuable insights into consumption patterns can be obtained.

5.7 Clustering Movement

As discussed in the previous section, the period from 2017 to 2019 exhibits a notable degree of similarity in the clustering patterns of countries. However, it is observed that certain countries undergo transitions from one cluster to another. A summary of these cluster movements is presented in Table 5. Over the course of the last three years, a total of seven countries are observed to have changed their clusters. Notably, the majority of these movements occur between Cluster 2 and Cluster 5, with the exception of Iraq and Turkey.

Table 5. The Country Movement between Clusters from 2015 - 2019

Country/Market	2015	2016	2017	2018	2019	move last 4 years	move last 3 years
Argentina			2	5	5	yes	yes
Brazil	2	1	5	5	5	yes	no
Canada	2	1	2	5	5	yes	yes
France	2	1	5	5	5	yes	no
Germany	2	1	5	5	5	yes	no
Iran (Islamic Republic of)	2	2	1	1	1	yes	no
Iraq	2	2	1	2	2	yes	yes
Italy	2	1	5	5	5	yes	no
Mexico	2	1	5	5	5	yes	no
Netherlands	2	2	5	5	5	yes	no
Poland	2	1	2	5	5	yes	yes
Russian Federation	2	1	1	1	1	no	no
South Africa	2	1	2	5	5	yes	yes
Spain	2	1	5	5	5	yes	no
Thailand	1	2	5	2	2	yes	yes
Turkey	2	2	1	1	2	yes	yes
United Kingdom of Great Britain and Northern Ireland	2	1	5	5	5	yes	no

During the transition from 2017 to 2018, significant movement between clusters was observed for certain countries. Canada and Poland experienced a slight improvement in total volume for *Tea Based Drinks*, particularly in the *Family Pack* and *Portion Pack* categories, leading to a shift from Cluster 2 to Cluster 5. Similarly, South Africa demonstrated a gradual rise in *FFSD* and *Juice*, specifically in the *Portion Pack* and *Family*

Pack segments, causing a shift from Cluster 2 to Cluster 5. Conversely, Thailand transitioned from Cluster 5 to Cluster 2 due to a decline in *Tea Based Drink* and *Family Pack* volumes. Iraq moved from Cluster 1 to Cluster 2 as a result of increased *FFSD* and total *Portion Pack* size. While, from 2018 to 2019, only Turkey shifts from Cluster 1 to Cluster 2 due to the reduction in total volume of *Family Pack* size.

6. Conclusion

6.1 Conclusion and Implications

The primary objective of this study is to employ unsupervised machine learning algorithms, namely K-Means and Hierarchical Clustering, to cluster countries or markets based on packaging product and packaging size for the companies in the food packaging market. The rationale behind this clustering approach is to extract meaningful insights into the underlying factors driving market demand. By understanding the distinctive characteristics of different clusters, the company can tailor its product portfolio and marketing strategies to enhance its competitiveness and profitability in each specific market.

Upon conducting the clustering analysis, it is observed that K-Means with 6 clusters has a higher Silhouette Score compared to Hierarchical Clustering. This implies that the clustering results obtained from K-Means exhibit a better degree of separation and compactness within each cluster, thus indicating a more robust clustering solution for identifying distinct market segments.

Additionally, an examination of the clustering trends from 2015 to 2019 is conducted in this study. The analysis reveals a consistent pattern in the country clusters during the period from 2017 to 2019, indicating a degree of stability and similarity in terms of country characteristics and the overall volume of packaging. However, notable differences are observed in the clustering patterns of 2015 and 2016, suggesting variations in country characteristics and package volume during those particular years. These findings highlight the significance of considering temporal trends and dynamics when analyzing the clustering results and drawing conclusions regarding country characteristics and packaging volumes.

PCA Loading Scores are employed in this study to ascertain the relative importance of variables. Specifically, PCA with 9 components is selected as it captures approximately 70% of the cumulative percent variance, thus providing a comprehensive representation of the dataset. This approach enables us to identify the key variables that contribute significantly to the overall variability in the data, allowing for a more focused and meaningful analysis of the underlying factors and patterns.

Based on the analysis on cluster characteristics, Cluster 1 has strong preference for *Family Pack size* and the highest volume of *Traditional Cultured Drinks*, Cluster 2 has equal distribution of *Family Pack, Loose, and Portion Pack Size* with high concentration of *FFSD,* Cluster 3, which is United States, has high proportion of *Large Container Size* but no *Loose* packaging contributing to the market. This cluster addresses a large volume of *FFSD and Tea Based Drinks* compared with other clusters. Cluster 4 (China) has a high preference for *Portion Pack Size* and relatively low consumption of *White Milk*. Cluster 5 has a high distribution of *Family Pack Size* and the highest distribution of *FFSD* among other clusters. Lastly, Cluster 6, which is only India, has a dominant consumption of *White Milk in Loose* Package Size and *Portion* Packed Size.

Over time 2017 - 2019, there are movements of some countries from one cluster to another cluster. The reason is because of the changes in package group as well as package size, mostly occurring in the variables having high PCA loadings or influence on the forming of clustering.

6.2 Limitations and Future Work

The clustering technique used in this study has several important limitations. Firstly, it struggles with handling mixed and panel data, requiring the conversion of categorical variables into numerical ones or the use of dimensionality reduction techniques. However, these approaches can introduce challenges and potentially lead to information loss. Secondly, the inclusion of a large number of variables in the clustering process can result in the formation of clusters that lack clear alignment with distinct market segments. This can be attributed to the introduction of noise or irrelevant information. Additionally, this limitation can give rise to duplicate markets within clusters, where similar market characteristics are duplicated across different clusters.

In this thesis, clustering techniques have been applied to identify the market segment for the company in the food packaging industry. However, to enhance the practical applicability of our research, further work can be done to incorporate forecasting methods. The inclusion of forecasting will allow for the prediction of future trends with each market segment, providing valuable insights for businesses and decision-makers. Future work will focus on exploring and implementing various forecasting models, such as time series analysis or classification models, to enable accurate predictions of country segments. This additional analysis will contribute to a comprehensive understanding of the market dynamics and support the company on decision-making processes.

References

Abdulhafedh, A. (2021). “Customer Segmentation.” *Journal of City and Development*, vol. 3, no. 1, pp. 12–30. Available online: https://doi.org/10.12691/jcd-3-1-3.

Akay, Ö. and Yüksel, G. (2017). Clustering the mixed panel dataset using Gower’s distance and k-prototypes algorithms. Communications in Statistics - Simulation and Computation. Available online: https://doi.org/10.1080/03610918.2017.1367806.

Ao, X. and Wei, E. (2022). Review of Panel Data Clustering Methods and Applications. IEEE Xplore. Available online: https://doi.org/10.1109/ICAIBD55127.2022.9820005.

Bayer, J. (2010). Customer segmentation in the telecommunications industry. Journal of Database Marketing & Customer Strategy Management. Available online: https://doi.org/10.1057/dbm.2010.21.

Brennan, L., Langley, S., Verghese, K., Lockrey, S., Ryder, M., Francis, C., Phan-Le, N.T. and Hill, A. (2021). The role of packaging in fighting food waste: A systematised review of consumer perceptions of packaging. Journal of Cleaner Production. Available online: https://doi.org/10.1016/j.jclepro.2020.125276.

de Amorim, R.C. (2015). Feature Relevance in Ward’s Hierarchical Clustering Using the L p Norm. Journal of Classification, 32(1), pp.46–62. Available online: https://doi.org/10.1007/s00357-015-9167-1.

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise.

Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). Hierarchical clustering. In Cluster Analysis, pp. 57-86.

Everitt, Brian. (2004). *An R and S-plus Companion to Multivariate Analysis*. London, Springer, p. 41.

European Economic and Social Committee. (2022). EU plastic and packaging reforms: EESC concerns about the impact on the environment, consumers and industrial competitiveness. Available online: https://www.eesc.europa.eu/en/news-media/news/eu-plastic-and-packaging-reforms-eesc-concerns-about-impact-environment-consumers-and-industrial-competitiveness.

Fortune Business Insights. (2019). Food Packaging Market Size, Share, Analysis, Growth, Research Report 2026. Available online: https://www.fortunebusinessinsights.com/industry-reports/food-packaging-market-101941.

Grand View Research. North America Food Packaging Market Report, 2028. Available online: https://www.grandviewresearch.com/industry-analysis/north-america-food-packaging-market-report.

Grein, Andreas F, et al. (2010). "A Dynamic Analysis of Country Clusters, the Role of Corruption, and Implications for Global Firms 1." *Journal of ECONOMICS and BUSINESS*, vol. Vol. XIII, No 2 (33-60), no. 2, 2011, pp. 33–60.

Kansal, T., Bahuguna, S., Singh, V. and Choudhury, T. (2018). Customer Segmentation using K-means Clustering. 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS). Available online: https://doi.org/10.1109/ctems.2018.8769171.

Lindholm, Andreas. (2022). *Machine Learning a First Course for Engineers and Scientists*. Cambridge, Uk New York, Ny Cambridge University Press, p. 25.

---. (2022). *Machine Learning a First Course for Engineers and Scientists*. Cambridge, Uk New York, Ny Cambridge University Press, pp. 279–284.

Mahmoud, M.A., Tsetse, E.K.K., Tulasi, E.E. and Muddey, D.K. (2022). Green Packaging, Environmental Awareness, Willingness to Pay and Consumers' Purchase Decisions. Sustainability. Available online: https://doi.org/10.3390/su142316091.

Markets and Markets. (2020). Military Simulation and Training Market Size, Global Growth Drivers, Opportunities 2031. Available online: https://www.marketsandmarkets.com/Market-Reports/military-simulation-virtual-training-market-661.html.

Marsh, K. and Bugusu, B. (2007). Food Packaging - Roles, Materials, and Environmental Issues. Journal of Food Science. Available online: :https://doi.org/10.1111/j.1750-3841.2007.00301.x.

Mordor Intelligence. (2021). Food Packaging Market | 2022 - 27 | Industry Share, Size, Growth - Mordor Intelligence. Available online: https://www.mordorintelligence.com/industry-reports/food-packaging-market.

Mordor Intelligence. (2021). Latin America Active & Intelligent Packaging Market Size & Share Analysis - Industry Research Report - Growth Trends. Available online: https://www.mordorintelligence.com/industry-reports/latin-american-active-and-intelligent-packaging-market-industry.

Popat, S, et al. (2014). Review and Comparative Study of Clustering Techniques. International Journal of Computer Science and Information Technologies.

Rajagukguk, S.A. and Fudholi, D.H. (2022). Rich Kid, Poor Kid: Students Clustering using K-Prototype Algorithm. 2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA). Available online: https://doi.org/10.1109/icitda55840.2022.9971445.

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, pp.53–65. Available online: https://doi.org/10.1016/0377-0427(87)90125-7.

Sasirekha, K. and Baby, P. (2013). Agglomerative Hierarchical Clustering Algorithm-A Review. International Journal of Scientific and Research Publications.

Scutariu, A.-L, and Danny Ho. (2022). "Of the COVID-19 Pandemic." J. Theor. Appl. Electron. Commer. Res, p. 17. Available online: https://doi.org/10.3390/jtaer17010003

Shi, Congming, et al. (2021). "A Quantitative Discriminant Method of Elbow Point for the Optimal Number of Clusters in a Clustering Algorithm." *EURASIP Journal on Wireless Communications and Networking*, no. 1, 15 Feb. 2021. Available online: https://doi.org/10.1186/s13638-021-01910-w.

Suhr, Diana. (2005). *Paper 203-30 Principal Component Analysis vs. Exploratory Factor Analysis*.

Šulc, Z. and Řezanková, H. (2019). Comparison of Similarity Measures for Categorical Data in Hierarchical Clustering. Journal of Classification. Available online: https://doi.org/10.1007/s00357-019-09317-5.

Teichert, T., Shehu, E. and von Wartburg, I. (2008). Customer segmentation revisited: The case of the airline industry. Transportation Research Part A: Policy and Practice. Available online: https://doi.org/10.1016/j.tra.2007.08.003.

Valle, Sergio, et al. (1999). "Selection of the Number of Principal Components: The Variance of the Reconstruction Error Criterion with a Comparison to Other Methods." *Industrial & Engineering Chemistry Research*, vol. 38, no. 11, pp. 4389–4401. Available online: https://doi.org/10.1021/ie990110i.

Verain, M.C.D., Sijtsema, S.J. and Antonides, G. (2016). Consumer segmentation based on food-category attribute importance: The relation with healthiness and sustainability perceptions. Food Quality and Preference, 48, pp.99–106. Available online: https://doi.org/10.1016/j.foodqual.2015.08.012.

Versino, F., Ortega, F., Monroy, Y., Rivero, S., López, O.V. and García, M.A. (2023). Sustainable and Bio-Based Food Packaging: A Review on Past and Current Design Innovations. Foods. Available online: https://doi.org/10.3390/foods12051057.

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association, p. 236-244.

Wedel, M. and Kamakura, W.A. (2000). The Historical Development of the Market Segmentation Concept. International Series in Quantitative Marketing. Available online: https://doi.org/10.1007/978-1-4615-4651-1_1.

Wold, Svante, et al. (1987). “Principal Component Analysis.” *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1-3, pp. 37–52. Available online: https://doi.org/10.1016/0169-7439(87)80084-9.

Yuan, Chunhui, and Haitao Yang. (2019). “Research on K-Value Selection Method of K-Means Clustering Algorithm.” *J*, vol. 2, no. 2, pp. 226–235. Available online: https://doi.org/10.3390/j2020016.

Appendix

A. 107 Variables Used for Clustering

No	Feature	Categories
1	Country Characteristic	consum_capita
2		gdp_capita
3		inflation
4		unemployment_rate
5		urbanization
6	Baby & Toddler Dairy Drinks	Family Pack
7		Large size containers
8		Portion Pack
9		Powder
10	Buttermilk	Bulk
11		Family Pack
12		Large size containers
13		Loose
14		Portion Pack
15	Coffee based drinks	Bulk
16		Family Pack
17		Large size containers
18		Portion Pack
19		Powder
20	Drinking Yogurt	Bulk
21		Family Pack
22		Large size containers
23		Loose
24		Portion Pack
25		Powder
26	Energy Drinks	Bulk
27		Family Pack
28		Large size containers
29		Portion Pack
30		Powder

No	Feature	Categories
53	Liquid Cream	Bulk
54		Family Pack
55		Large size containers
56		Loose
57		Portion Pack
58	Nectar	Bulk
59		Family Pack
60		Large size containers
61		Portion Pack
62	Non Dairy Cream	Bulk
63		Family Pack
64		Large size containers
65		Portion Pack
66		Powder
67	Rice, Nut, Grain and Seed Based Drinks	Bulk
68		Family Pack
69		Large size containers
70		Loose
71		Portion Pack
72		Powder
73	Soy Milk	Family Pack
74		Large size containers
75		Loose
76		Portion Pack
77		Powder
78	Sports Drinks	Bulk
79		Family Pack
80		Large size containers
81		Portion Pack
82		Powder

31	Evaporated Milk	Bulk
32		Family Pack
33		Large size containers
34		Portion Pack
35	FFSD	Bulk
36		Family Pack
37		Large size containers
38		Portion Pack
39		Powder
40	Flavoured Milk	Bulk
41		Family Pack
42		Large size containers
43		Loose
44		Portion Pack
45		Powder
46	Juice	Bulk
47		Family Pack
48		Large size containers
49		Portion Pack
50	LAD	Family Pack
51		Large size containers
52		Portion Pack

83	Sweetened Condensed Milk	Bulk
84		Family Pack
85		Large size containers
86		Loose
87		Portion Pack
88	Tea based drinks	Bulk
89		Family Pack
90		Large size containers
91		Portion Pack
92	Traditional Cultured Drinks	Powder
93		Bulk
94		Family Pack
95		Large size containers
96		Loose
97		Portion Pack
98	White Milk	Bulk
99		Family Pack
100		Large size containers
101		Loose
102		Portion Pack
103		Powder
104	Wine (0-19% alcohol)	Bulk
105		Family Pack
106		Large size containers
107		Portion Pack

B. Agglomerative Clustering's process to perform clustering

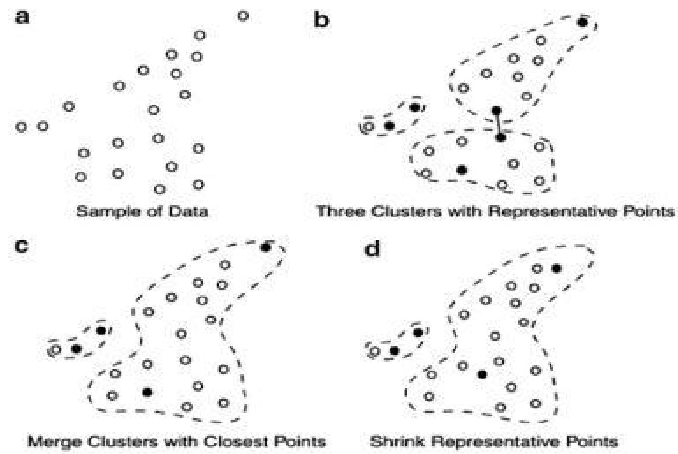


Figure 12. Showing how Hierarchical Clustering performs clustering (Abdullah and Hamdan, 2015)

C. The Highest PCA Loading Score in 2019 (PCA3 - PCA9)

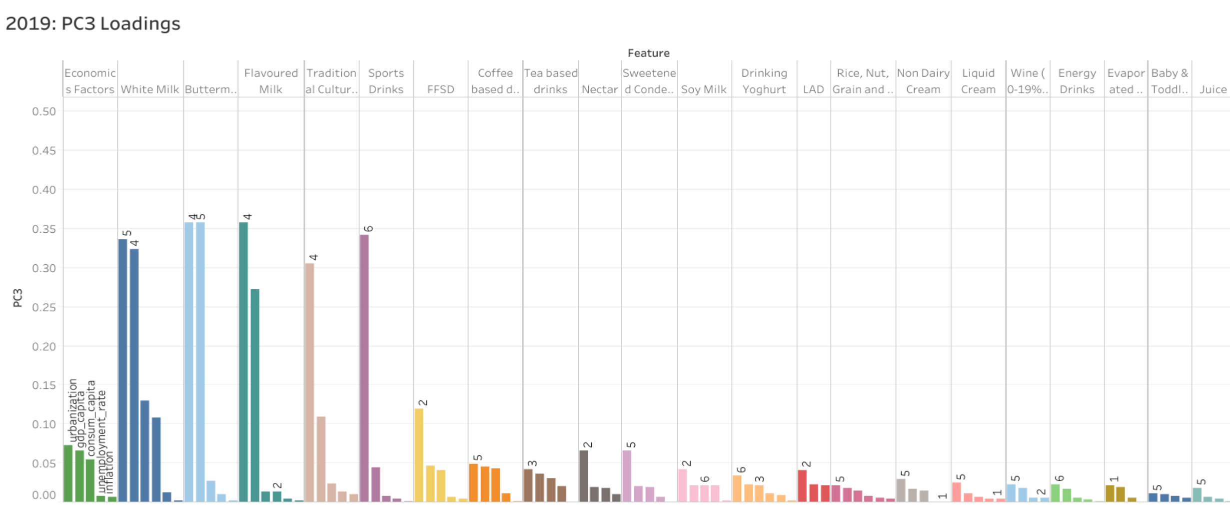


Figure C.1 Illustrating PCA3. Only 7 variables that have a high PCA3 Loading Score. *Buttermilk with Loose* and *Buttermilk with Portion Pack* have the highest first and second PCA3 Loading, followed by *Sport Drinks with Powder*, *White Milk with Portion Pack* and *Loose*.

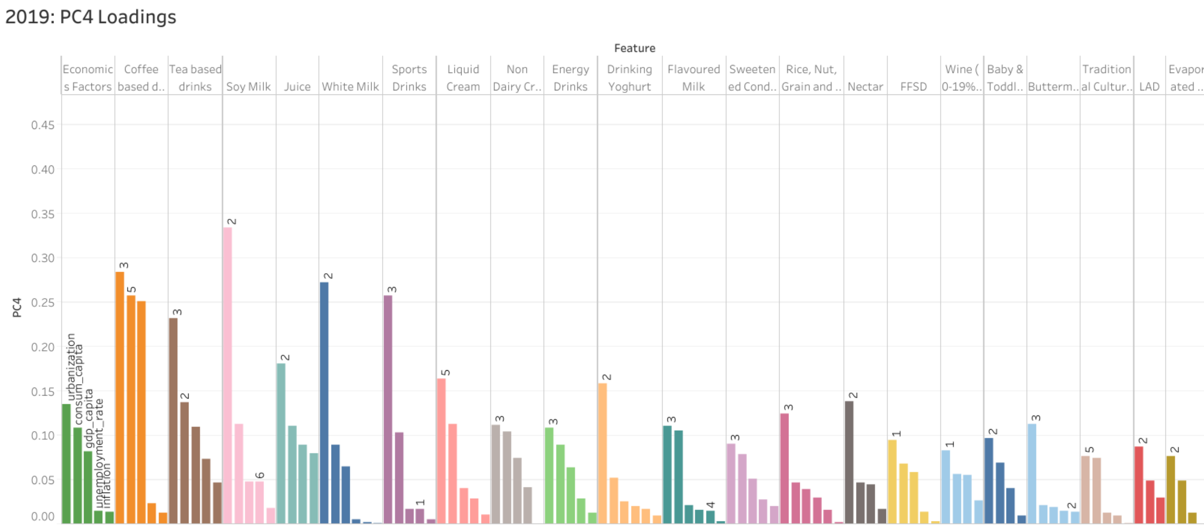


Figure C.2 Illustrating PCA4. *Soy Milk with Family Pac*k size is the highest PCA4, followed by *Coffee Based Drinks with Large Size Containers* and *White Milk with Family Pack*. In total *Coffee based Drinks* has the highest PCA4.

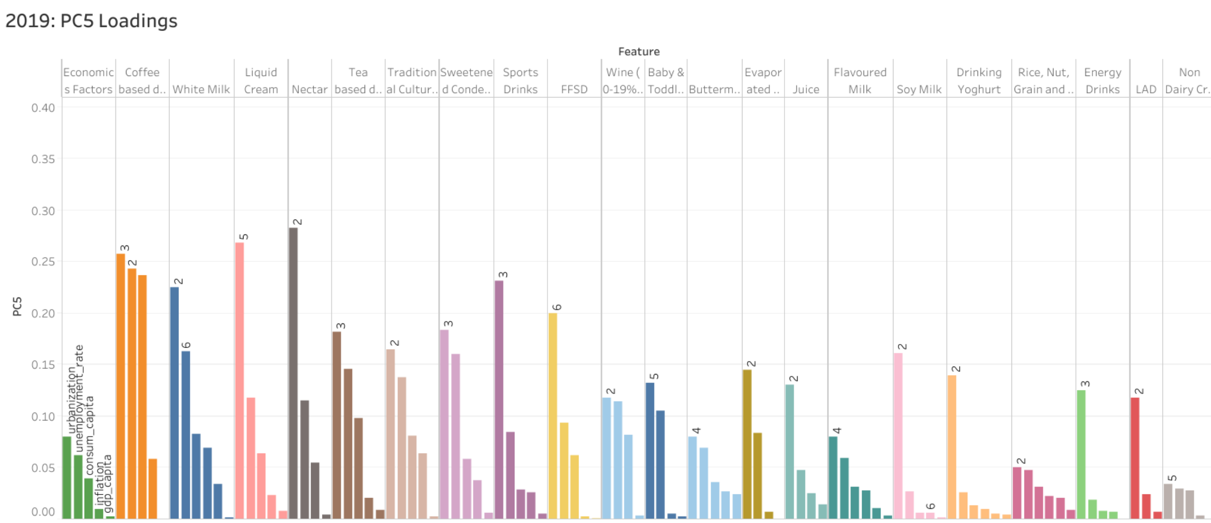


Figure C.3 Illustrating PCA5. There are three variables that have a PCA score above 0.25, namely *Nectar with Family Pack, Liquid Cream with Family Pack,* and *Coffee Based Drink with Large Size Containers*. In total, Coffee Based Drinks has the highest PCA5

2019: PC6 Loadings

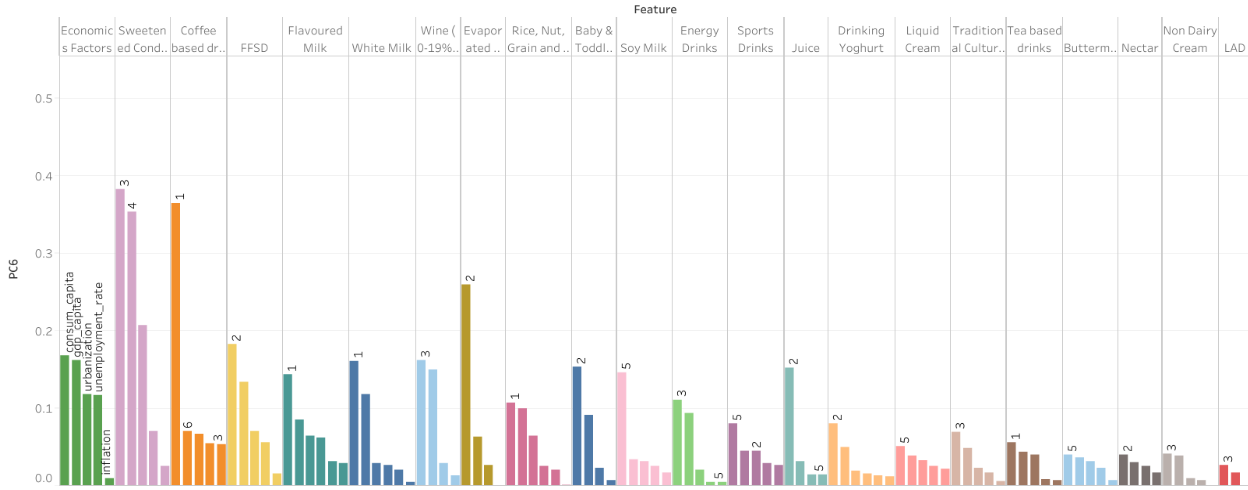


Figure C.4 Illustrating PCA6. Four variables have more than 0.25 PCA Scores, namely *Sweetened Condensed Milk with Large Size Containers and Loose*, *Coffee Based Drinks with Bulk* and *Evaporated Milk with Family Pack*

2019: PC7 Loadings

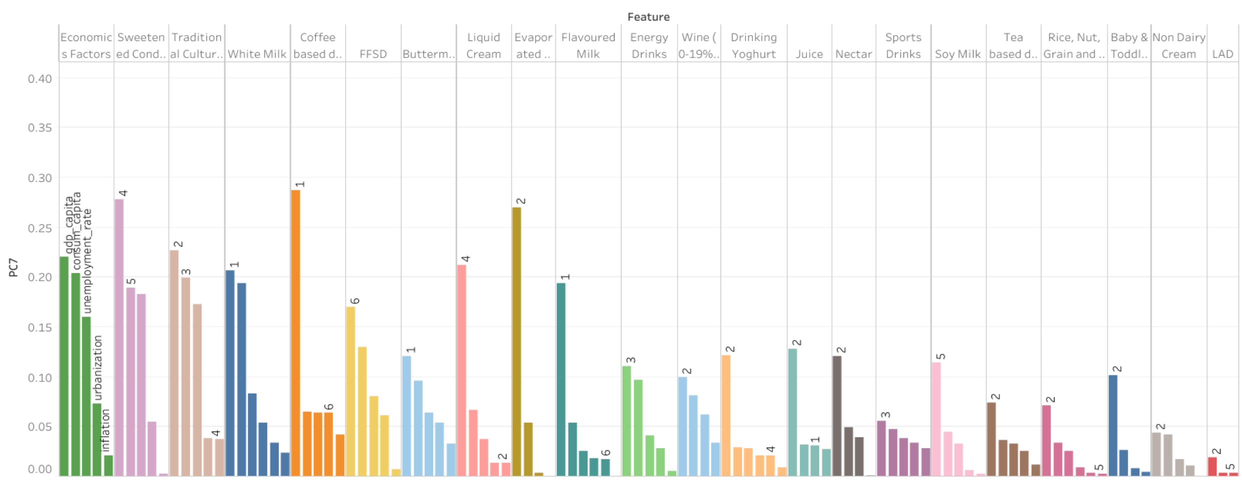


Figure C.5 Illustrating PCA7. *GDP per Capita* and *Consumption per Capita* have higher PCA7 compared with other PCA Scores. *Coffee Based Drinks with Bulk, Sweetened Condensed Milk with Loose* and *Evaporated Milk with Family Pack* are the top 3 PCA7.

2019: PC8 Loadings

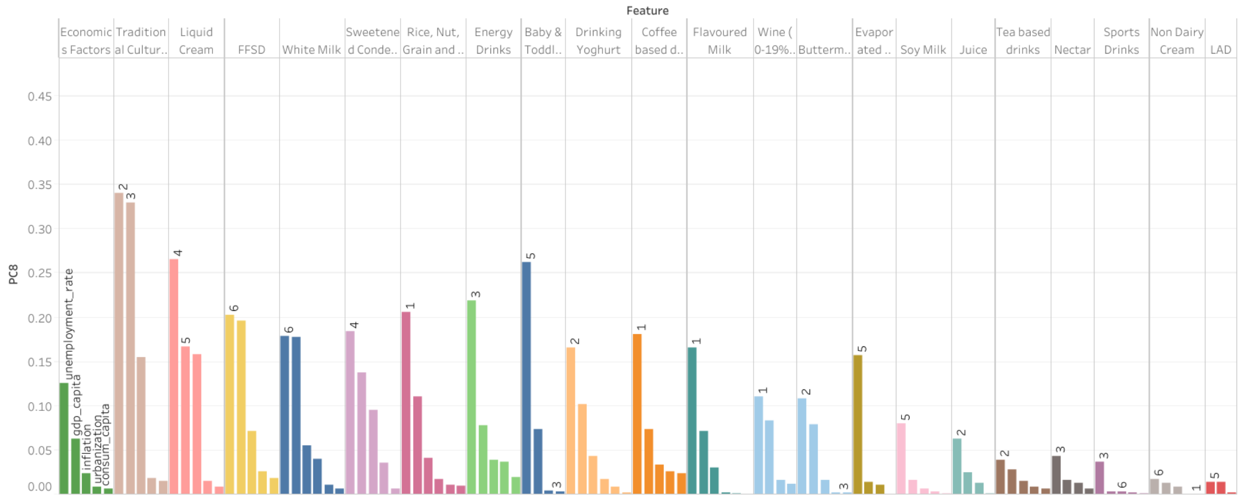


Figure C.6 Illustrating PCA8. *Traditional Cultured Drinks with Family Pack and Large Size Containers* followed by *Liquid Cream with Loose* and *Baby Toddler & Dairy Drinks with Portion Pack*. In total, *Traditional Cultured Drinks* contribute the highest PCA8

2019: PC9 Loadings

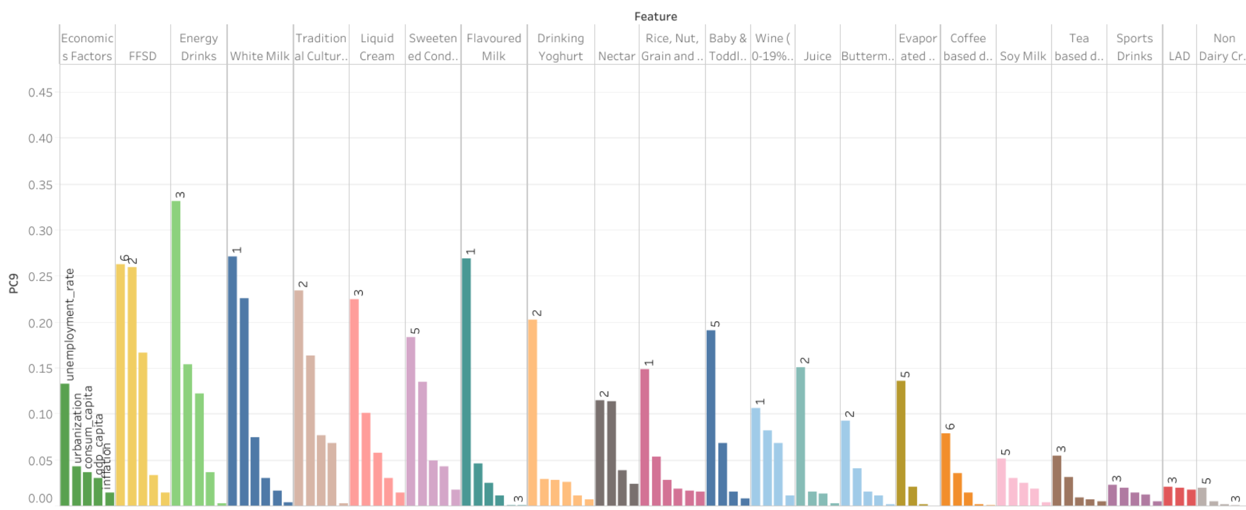


Figure C.7 Illustrating PCA9. *Energy Drinks with Large Size Containers* is the highest PCA9 Loading Score. *FFSD with Powder and with Family Pack*, *White Milk with Bulk and Flavored Milk with Bulk* have quite similar PCA9 Score. In total *FFSD* has the highest contribution of PCA9.

D. PCA Loadings by Package Group

PCA Loadings in 2015

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks	0.20	0.25	0.04	0.29	0.26	0.36	0.07	**0.47**	0.20
Buttermilk	0.23	0.24	**0.75**	0.17	0.23	0.28	0.37	0.27	0.25
Coffee based drinks	0.21	0.10	0.19	**0.60**	**1.02**	**0.52**	**0.58**	0.31	0.26
Drinking Yoghurt	0.28	0.42	0.10	0.34	0.13	0.24	0.16	0.43	0.25
Energy Drinks	0.49	0.33	0.06	0.32	0.17	0.31	0.24	0.41	**0.67**
Evaporated Milk	0.22	0.29	0.06	0.17	0.16	0.15	**0.45**	0.30	0.15
FFSD	**0.61**	0.36	0.18	0.27	0.26	**0.54**	0.25	**0.49**	**0.71**
Flavoured Milk	**0.55**	**0.49**	**0.66**	0.32	0.25	**0.45**	0.19	0.29	**0.48**
Juice	**0.55**	0.44	0.03	**0.44**	0.17	0.29	0.16	0.07	0.11
LAD	0.32	**0.56**	0.09	0.12	0.15	0.06	0.05	0.01	0.12
Liquid Cream	0.47	0.39	0.05	0.41	0.36	0.26	0.28	**0.49**	0.35
Nectar	0.47	0.34	0.11	**0.44**	0.41	0.10	0.28	0.19	0.27
Non Dairy Cream	0.48	0.46	0.06	0.35	0.06	0.13	0.11	0.04	0.08
Rice, Nut, Grain and Seed Based Drinks	0.44	**0.47**	0.08	0.29	0.13	0.34	0.14	0.32	**0.45**
Soy Milk	**0.53**	**0.76**	0.11	**0.46**	0.26	0.27	0.25	0.11	0.17
Sports Drinks	0.46	0.34	**0.41**	0.34	**0.45**	0.25	0.16	0.11	0.13
Sweetened Condensed Milk	0.16	0.22	0.12	0.43	0.35	**0.91**	**0.94**	0.45	0.37
Tea based drinks	**0.72**	**0.46**	0.14	0.40	**0.51**	0.21	0.12	0.11	0.16
Traditional Cultured Drinks	0.03	0.03	**0.44**	0.36	**0.45**	0.28	**0.57**	**0.77**	0.40
White Milk	0.35	0.35	**0.90**	**0.55**	**0.46**	0.44	0.34	**0.59**	**0.63**
Wine (0-19% alcohol)	0.43	0.23	0.04	0.32	0.29	0.42	0.13	0.22	0.22
Economics Factors	0.13	0.16	0.24	0.35	0.15	**0.81**	**0.52**	0.12	0.16

PCA Loadings in 2016

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks	0.19	0.23	0.04	0.23	0.30	0.31	0.12	**0.44**	0.23
Buttermilk	0.23	0.24	**0.75**	0.20	0.24	0.17	0.37	0.31	0.20
Coffee based drinks	0.22	0.11	0.19	**0.82**	**0.84**	**0.57**	**0.56**	0.28	0.25
Drinking Yoghurt	0.29	0.43	0.10	0.29	0.17	0.21	0.24	0.37	0.27
Energy Drinks	0.49	0.35	0.06	0.30	0.16	0.25	0.29	0.44	**0.61**
Evaporated Milk	0.21	0.28	0.05	0.15	0.19	0.34	0.37	0.32	0.12
FFSD	**0.61**	0.36	0.19	0.23	0.29	**0.46**	0.46	**0.46**	**0.72**
Flavoured Milk	**0.55**	**0.48**	**0.65**	0.31	0.25	**0.41**	0.28	0.25	**0.50**
Juice	**0.55**	0.44	0.03	**0.45**	0.22	0.24	0.21	0.14	0.11
LAD	0.32	**0.56**	0.09	0.16	0.14	0.06	0.03	0.03	0.11
Liquid Cream	0.48	0.38	0.05	0.34	0.44	0.18	0.25	**0.59**	0.36
Nectar	0.48	0.35	0.12	0.28	**0.48**	0.07	0.24	0.12	0.28
Non Dairy Cream	0.47	**0.47**	0.06	0.33	0.09	0.11	0.14	0.05	0.09
Rice, Nut, Grain and Seed Based Drinks	0.46	**0.46**	0.08	0.30	0.14	0.34	0.11	0.43	0.42
Soy Milk	**0.53**	**0.76**	0.11	**0.52**	0.18	0.30	0.24	0.09	0.19
Sports Drinks	0.46	0.32	**0.41**	0.40	0.41	0.23	0.18	0.10	0.15
Sweetened Condensed Milk	0.16	0.22	0.12	0.32	0.43	**1.02**	**0.88**	0.37	0.35
Tea based drinks	**0.72**	0.46	0.15	**0.54**	**0.44**	0.17	0.17	0.12	0.16
Traditional Cultured Drinks	0.02	0.03	**0.45**	0.22	**0.51**	0.15	**0.52**	**0.83**	**0.54**
White Milk	0.34	0.34	**0.91**	**0.46**	**0.55**	0.36	**0.55**	**0.47**	**0.65**
Wine (0-19% alcohol)	0.43	0.22	0.05	0.24	0.34	0.38	0.27	0.18	0.21
Economics Factors	0.13	0.16	0.22	0.34	0.17	**0.65**	**0.60**	0.24	0.18

PCA Loadings in 2017

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks	0.18	0.22	0.04	0.22	0.29	0.29	0.14	0.41	0.22
Buttermilk	0.23	0.24	**0.76**	0.20	0.23	0.12	0.38	0.27	0.19
Coffee based drinks	0.22	0.12	0.18	**0.84**	**0.83**	**0.62**	**0.53**	0.29	0.19
Drinking Yoghurt	0.29	0.42	0.10	0.29	0.19	0.16	0.26	0.34	0.28
Energy Drinks	0.48	0.35	0.05	0.30	0.16	0.23	0.31	**0.42**	**0.68**
Evaporated Milk	0.22	0.28	0.05	0.14	0.22	**0.41**	0.32	0.25	0.17
FFSD	**0.59**	0.38	0.19	0.24	0.34	0.37	0.51	**0.46**	**0.73**
Flavoured Milk	**0.55**	**0.48**	**0.66**	0.30	0.23	**0.37**	0.32	0.25	**0.44**
Juice	**0.54**	0.44	0.04	**0.46**	0.23	0.20	0.22	0.12	0.14
LAD	0.33	**0.55**	0.09	0.15	0.14	0.06	0.02	0.03	0.08
Liquid Cream	0.48	0.38	0.05	0.33	**0.46**	0.17	0.28	**0.60**	0.44
Nectar	0.49	0.35	0.12	0.24	**0.47**	0.11	0.22	0.12	0.28
Non Dairy Cream	0.47	**0.48**	0.06	0.33	0.10	0.10	0.14	0.06	0.08
Rice, Nut, Grain and Seed Based Drinks	0.48	**0.47**	0.07	0.28	0.14	0.27	0.13	0.33	0.28
Soy Milk	**0.53**	**0.74**	0.11	**0.54**	0.19	0.28	0.22	0.08	0.14
Sports Drinks	0.44	0.32	**0.40**	0.40	0.40	0.22	0.20	0.09	0.08
Sweetened Condensed Milk	0.16	0.21	0.11	0.28	0.44	**1.06**	**0.74**	0.36	0.38
Tea based drinks	**0.71**	0.46	0.14	**0.55**	0.45	0.15	0.18	0.11	0.14
Traditional Cultured Drinks	0.02	0.03	**0.45**	0.18	**0.47**	0.12	**0.58**	**0.95**	**0.52**
White Milk	0.33	0.34	**0.91**	**0.45**	**0.57**	0.30	**0.65**	**0.46**	**0.64**
Wine (0-19% alcohol)	0.43	0.22	0.05	0.23	0.34	0.35	0.31	0.19	0.23
Economics Factors	0.13	0.15	0.21	0.36	0.20	**0.56**	**0.67**	0.21	0.20

PCA Loadings in 2018

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks	0.18	0.21	0.04	0.22	0.26	0.28	0.14	0.38	0.27
Buttermilk	0.22	0.25	**0.75**	0.17	0.25	0.15	0.36	0.23	0.17
Coffee based drinks	0.23	0.13	0.15	**0.83**	**0.82**	**0.61**	**0.53**	0.32	0.16
Drinking Yoghurt	0.31	0.42	0.10	0.29	0.19	0.20	0.22	0.33	0.30
Energy Drinks	0.48	0.39	0.05	0.30	0.16	0.24	0.27	0.40	**0.65**
Evaporated Milk	0.21	0.28	0.05	0.14	0.23	0.34	0.34	0.19	0.16
FFSD	**0.58**	0.38	0.21	0.23	0.35	**0.46**	0.45	**0.49**	**0.75**
Flavoured Milk	**0.54**	**0.47**	**0.66**	0.28	0.23	**0.41**	0.32	0.27	0.37
Juice	**0.53**	0.45	0.03	**0.46**	0.22	0.23	0.20	0.11	0.17
LAD	0.34	**0.54**	0.09	0.15	0.14	0.05	0.02	0.03	0.06
Liquid Cream	0.48	0.36	0.06	0.34	**0.46**	0.17	0.32	**0.61**	0.42
Nectar	0.50	0.33	0.14	0.25	**0.46**	0.11	0.24	0.08	0.30
Non Dairy Cream	0.47	**0.50**	0.07	0.33	0.10	0.11	0.14	0.06	0.09
Rice, Nut, Grain and Seed Based Drinks	0.49	0.45	0.08	0.27	0.16	0.30	0.12	0.35	0.28
Soy Milk	**0.53**	**0.73**	0.11	**0.55**	0.20	0.26	0.21	0.10	0.14
Sports Drinks	0.45	0.33	**0.40**	0.41	0.39	0.23	0.20	0.06	0.08
Sweetened Condensed Milk	0.16	0.21	0.12	0.28	0.45	**1.03**	**0.74**	**0.43**	**0.43**
Tea based drinks	**0.71**	**0.46**	0.13	**0.58**	**0.46**	0.16	0.18	0.10	0.11
Traditional Cultured Drinks	0.02	0.03	**0.46**	0.18	0.45	0.13	**0.67**	**0.89**	**0.52**
White Milk	0.33	0.34	**0.93**	**0.44**	**0.57**	0.37	**0.60**	**0.47**	**0.64**
Wine (0-19% alcohol)	0.43	0.22	0.05	0.23	0.33	0.37	0.27	0.21	0.25
Economics Factors	0.12	0.15	0.21	0.35	0.19	**0.59**	**0.64**	0.21	0.26

PCA Loadings in 2019

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks	0.17	0.21	0.04	0.22	0.25	0.28	0.14	0.34	0.28
Buttermilk	0.21	0.25	**0.76**	0.18	0.24	0.14	0.37	0.21	0.16
Coffee based drinks	0.24	0.14	0.15	**0.83**	**0.80**	**0.61**	**0.52**	0.34	0.13
Drinking Yoghurt	0.31	0.42	0.10	0.28	0.20	0.19	0.23	0.34	0.31
Energy Drinks	0.47	0.39	0.05	0.30	0.16	0.23	0.28	0.39	**0.65**
Evaporated Milk	0.21	0.27	0.05	0.14	0.24	0.35	0.33	0.18	0.16
FFSD	**0.58**	0.38	0.22	0.24	0.36	**0.46**	0.45	**0.51**	**0.74**
Flavoured Milk	**0.53**	**0.47**	**0.66**	0.27	0.21	**0.42**	0.31	0.27	0.36
Juice	**0.53**	0.46	0.03	**0.46**	0.22	0.21	0.22	0.10	0.18
LAD	0.34	**0.53**	0.09	0.17	0.15	0.04	0.03	0.03	0.06
Liquid Cream	0.48	0.35	0.05	0.36	**0.48**	0.17	0.34	**0.62**	**0.43**
Nectar	0.51	0.33	0.11	0.25	**0.46**	0.11	0.21	0.08	0.29
Non Dairy Cream	0.47	**0.52**	0.06	0.33	0.10	0.10	0.11	0.04	0.03
Rice, Nut, Grain and Seed Based Drinks	0.49	0.46	0.07	0.26	0.18	0.32	0.15	0.40	0.28
Soy Milk	**0.53**	**0.72**	0.11	**0.56**	0.20	0.26	0.20	0.11	0.13
Sports Drinks	0.44	0.32	**0.40**	0.40	0.38	0.23	0.20	0.05	0.08
Sweetened Condensed Milk	0.16	0.21	0.11	0.27	0.45	**1.04**	**0.71**	**0.46**	0.43
Tea based drinks	**0.70**	**0.47**	0.13	**0.60**	**0.46**	0.16	0.18	0.10	0.11
Traditional Cultured Drinks	0.02	0.03	**0.46**	0.18	0.45	0.17	**0.67**	**0.86**	**0.55**
White Milk	0.33	0.34	**0.91**	**0.44**	**0.58**	0.36	**0.59**	**0.47**	**0.63**
Wine (0-19% alcohol)	0.44	0.22	0.05	0.22	0.32	0.36	0.28	0.22	0.27
Economics Factors	0.12	0.16	0.21	0.36	0.20	**0.57**	**0.68**	0.23	0.26

E. PCA Loadings by Package and Size

PCA Loadings in 2015

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks									
Family Pack	0.03	0.03	0.01	0.12	0.08	0.17	0.05	0.11	0.07
Large size containers	0.01	0.00	0.01	0.01	0.00	0.03	0.01	0.00	0.01
Portion Pack	0.05	0.04	0.02	0.14	0.16	0.13	0.01	0.35	0.11
Powder	0.12	0.18	0.00	0.03	0.02	0.02	0.01	0.00	0.01
Buttermilk									
Bulk	0.00	0.01	0.01	0.03	0.02	0.04	0.14	0.11	0.04
Family Pack	0.07	0.08	0.02	0.01	0.04	0.05	0.12	0.12	0.10
Large size containers	0.14	0.14	0.00	0.12	0.01	0.05	0.02	0.00	0.02
Loose	0.00	0.01	0.36	0.01	0.09	0.05	0.03	0.03	0.02
Portion Pack	0.01	0.00	0.36	0.01	0.07	0.08	0.05	0.01	0.07
Coffee based drinks									
Bulk	0.00	0.01	0.00	0.03	0.02	0.25	0.41	0.15	0.02
Family Pack	0.07	0.05	0.06	0.18	0.33	0.05	0.05	0.02	0.00
Large size containers	0.05	0.03	0.06	0.20	0.34	0.04	0.05	0.02	0.00
Portion Pack	0.08	0.01	0.06	0.18	0.33	0.07	0.02	0.03	0.01
Powder	0.00	0.01	0.02	0.01	0.00	0.11	0.05	0.10	0.23
Drinking Yoghurt									
Bulk	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.02
Family Pack	0.05	0.06	0.00	0.21	0.06	0.10	0.07	0.20	0.12
Large size containers	0.10	0.18	0.02	0.04	0.02	0.02	0.00	0.06	0.04
Loose	0.01	0.00	0.01	0.03	0.02	0.03	0.02	0.00	0.03
Portion Pack	0.12	0.18	0.02	0.02	0.03	0.03	0.05	0.01	0.04
Powder	0.00	0.01	0.04	0.02	0.00	0.06	0.01	0.14	0.00
Energy Drinks									
Bulk	0.14	0.14	0.00	0.07	0.04	0.00	0.01	0.07	0.11
Family Pack	0.14	0.14	0.00	0.09	0.03	0.04	0.03	0.02	0.01
Large size containers	0.02	0.02	0.01	0.12	0.09	0.15	0.10	0.21	0.27
Portion Pack	0.18	0.02	0.01	0.03	0.01	0.00	0.04	0.04	0.00
Powder	0.02	0.01	0.02	0.01	0.00	0.12	0.06	0.06	0.28
Evaporated Milk									
Bulk	0.11	0.20	0.02	0.04	0.01	0.03	0.01	0.00	0.04
Family Pack	0.01	0.00	0.00	0.11	0.08	0.08	0.36	0.11	0.00
Large size containers	0.01	0.00	0.01	0.01	0.00	0.01	0.02	0.00	0.00
Portion Pack	0.09	0.09	0.03	0.02	0.06	0.03	0.07	0.19	0.10
FFSD									
Bulk	0.14	0.14	0.00	0.09	0.03	0.07	0.04	0.03	0.01
Family Pack	0.08	0.01	0.09	0.06	0.06	0.21	0.08	0.14	0.34
Large size containers	0.16	0.01	0.02	0.01	0.03	0.10	0.05	0.03	0.16
Portion Pack	0.15	0.15	0.03	0.01	0.00	0.00	0.00	0.03	0.01
Powder	0.08	0.05	0.04	0.10	0.13	0.15	0.08	0.27	0.20
Flavoured Milk									
Bulk	0.06	0.06	0.01	0.01	0.03	0.17	0.10	0.14	0.31
Family Pack	0.14	0.11	0.02	0.14	0.06	0.08	0.00	0.05	0.04
Large size containers	0.14	0.14	0.00	0.11	0.01	0.04	0.02	0.00	0.02
Loose	0.00	0.01	0.36	0.01	0.09	0.05	0.03	0.03	0.02
Portion Pack	0.17	0.09	0.00	0.03	0.00	0.04	0.00	0.02	0.00
Powder	0.04	0.08	0.26	0.03	0.05	0.07	0.04	0.05	0.09
Juice									
Bulk	0.14	0.14	0.00	0.12	0.01	0.05	0.02	0.00	0.01
Family Pack	0.11	0.04	0.00	0.19	0.07	0.20	0.10	0.03	0.10
Large size containers	0.15	0.14	0.01	0.09	0.01	0.03	0.02	0.02	0.01
Portion Pack	0.15	0.11	0.02	0.05	0.08	0.02	0.02	0.03	0.00
IAD									
Family Pack	0.11	0.17	0.04	0.05	0.13	0.01	0.02	0.01	0.04
Large size containers	0.11	0.20	0.02	0.04	0.01	0.03	0.01	0.00	0.04
Portion Pack	0.11	0.20	0.02	0.03	0.01	0.02	0.01	0.00	0.04
Liquid Cream									
Bulk	0.13	0.13	0.00	0.11	0.01	0.04	0.03	0.01	0.03
Family Pack	0.16	0.09	0.00	0.03	0.03	0.05	0.00	0.00	0.03
Large size containers	0.10	0.10	0.01	0.04	0.09	0.07	0.09	0.19	0.19
Loose	0.00	0.00	0.01	0.02	0.04	0.03	0.13	0.10	0.03
Portion Pack	0.07	0.07	0.02	0.21	0.19	0.08	0.03	0.19	0.07
Nectar									
Bulk	0.10	0.19	0.02	0.04	0.02	0.03	0.02	0.02	0.05
Family Pack	0.08	0.02	0.07	0.23	0.23	0.03	0.13	0.01	0.07
Large size containers	0.16	0.04	0.00	0.07	0.08	0.02	0.08	0.11	0.13
Portion Pack	0.13	0.08	0.01	0.11	0.08	0.02	0.05	0.05	0.03
Non Dairy Cream									
Bulk	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.04
Family Pack	0.14	0.14	0.01	0.11	0.01	0.06	0.03	0.01	0.02
Large size containers	0.14	0.14	0.00	0.12	0.01	0.06	0.03	0.00	0.01
Portion Pack	0.05	0.05	0.03	0.04	0.02	0.02	0.02	0.01	0.01
Powder	0.15	0.13	0.01	0.08	0.02	0.00	0.02	0.02	0.01
Rice, Nut, Grain and Seed Based Drinks									
Bulk	0.00	0.01	0.00	0.00	0.04	0.10	0.04	0.12	0.19
Family Pack	0.09	0.05	0.02	0.04	0.04	0.11	0.02	0.08	0.12
Large size containers	0.15	0.04	0.01	0.13	0.01	0.02	0.03	0.02	0.01
Loose	0.00	0.00	0.01	0.04	0.01	0.00	0.02	0.02	0.03
Portion Pack	0.11	0.20	0.02	0.04	0.01	0.03	0.01	0.00	0.04
Powder	0.09	0.17	0.02	0.03	0.03	0.07	0.02	0.07	0.07
Soy Milk									
Family Pack	0.08	0.05	0.05	0.25	0.22	0.07	0.03	0.02	0.04
Large size containers	0.14	0.14	0.00	0.12	0.01	0.05	0.02	0.00	0.02
Loose	0.11	0.20	0.02	0.04	0.01	0.02	0.00	0.01	0.04
Portion Pack	0.09	0.17	0.02	0.01	0.00	0.11	0.19	0.08	0.03
Powder	0.11	0.20	0.02	0.04	0.01	0.03	0.01	0.00	0.04
Sports Drinks									
Bulk	0.01	0.02	0.01	0.01	0.01	0.04	0.01	0.03	0.04
Family Pack	0.14	0.14	0.00	0.11	0.00	0.06	0.02	0.01	0.01
Large size containers	0.08	0.06	0.05	0.17	0.30	0.04	0.05	0.02	0.01
Portion Pack	0.17	0.07	0.01	0.02	0.05	0.07	0.05	0.03	0.06
Powder	0.06	0.04	0.33	0.03	0.08	0.04	0.03	0.02	0.01
Sweetened Condensed Milk									
Bulk	0.10	0.19	0.02	0.04	0.02	0.02	0.02	0.01	0.04
Family Pack	0.00	0.00	0.02	0.09	0.07	0.11	0.08	0.18	0.06
Large size containers	0.01	0.01	0.02	0.12	0.10	0.29	0.33	0.08	0.04
Loose	0.00	0.01	0.00	0.03	0.02	0.25	0.40	0.15	0.01
Portion Pack	0.05	0.01	0.06	0.15	0.14	0.24	0.12	0.02	0.23
Tea based drinks									
Bulk	0.14	0.14	0.01	0.11	0.01	0.06	0.01	0.01	0.01
Family Pack	0.17	0.02	0.03	0.04	0.11	0.08	0.04	0.05	0.05
Large size containers	0.12	0.05	0.05	0.16	0.25	0.00	0.03	0.02	0.06
Portion Pack	0.16	0.10	0.03	0.02	0.12	0.06	0.02	0.00	0.02
Powder	0.14	0.15	0.02	0.07	0.02	0.00	0.02	0.03	0.01
Traditional Cultured Drinks									
Bulk	0.00	0.00	0.01	0.00	0.01	0.06	0.02	0.03	0.01
Family Pack	0.01	0.01	0.03	0.15	0.16	0.10	0.20	0.33	0.16
Large size containers	0.00	0.00	0.00	0.02	0.06	0.01	0.14	0.15	0.01
Loose	0.00	0.01	0.31	0.02	0.06	0.01	0.03	0.02	0.06
Portion Pack	0.01	0.01	0.08	0.17	0.17	0.09	0.18	0.23	0.15
White Milk									
Bulk	0.05	0.05	0.00	0.00	0.03	0.19	0.11	0.15	0.31
Family Pack	0.05	0.01	0.12	0.32	0.13	0.02	0.06	0.07	0.02
Large size containers	0.14	0.14	0.00	0.10	0.02	0.01	0.02	0.03	0.07
Loose	0.00	0.01	0.34	0.01	0.07	0.03	0.01	0.03	0.02
Portion Pack	0.06	0.06	0.33	0.02	0.09	0.05	0.03	0.04	0.03
Powder	0.04	0.08	0.11	0.10	0.11	0.14	0.12	0.27	0.19
Wine (0-19% alcohol)									
Bulk	0.01	0.00	0.00	0.11	0.07	0.02	0.04	0.09	0.06
Family Pack	0.13	0.08	0.00	0.09	0.11	0.17	0.05	0.04	0.06
Large size containers	0.14	0.02	0.02	0.07	0.09	0.18	0.01	0.09	0.08
Portion Pack	0.15	0.13	0.02	0.05	0.02	0.05	0.03	0.00	0.03
Economics Factors									
consum_capita	0.04	0.06	0.06	0.10	0.01	0.22	0.12	0.02	0.00
gdp_capita	0.04	0.06	0.07	0.07	0.02	0.23	0.14	0.00	0.05
inflation	0.01	0.01	0.02	0.02	0.05	0.10	0.13	0.03	0.03
unemployment_rate	0.02	0.00	0.01	0.01	0.03	0.13	0.10	0.05	0.06
urbanization	0.03	0.04	0.08	0.15	0.03	0.14	0.04	0.01	0.01

PCA Loadings in 2016

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks									
Family Pack	0.03	0.03	0.01	0.10	0.11	0.15	0.10	0.10	0.07
Large size containers	0.01	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.00
Portion Pack	0.03	0.03	0.02	0.09	0.18	0.11	0.01	0.33	0.15
Powder	0.12	0.18	0.00	0.03	0.02	0.02	0.01	0.00	0.01
Buttermilk									
Bulk	0.00	0.01	0.01	0.02	0.02	0.01	0.12	0.12	0.02
Family Pack	0.07	0.08	0.02	0.02	0.04	0.02	0.10	0.14	0.08
Large size containers	0.14	0.14	0.00	0.11	0.02	0.04	0.03	0.00	0.02
Loose	0.00	0.01	0.36	0.02	0.09	0.04	0.05	0.01	0.02
Portion Pack	0.01	0.00	0.36	0.03	0.07	0.06	0.07	0.03	0.06
Coffee based drinks									
Bulk	0.00	0.01	0.00	0.03	0.04	0.35	0.33	0.11	0.01
Family Pack	0.07	0.05	0.06	0.26	0.26	0.04	0.05	0.03	0.01
Large size containers	0.05	0.03	0.06	0.27	0.27	0.04	0.05	0.04	0.00
Portion Pack	0.09	0.01	0.06	0.25	0.26	0.06	0.03	0.03	0.00
Powder	0.00	0.01	0.02	0.01	0.00	0.09	0.10	0.08	0.23
Drinking Yoghurt									
Bulk	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Family Pack	0.05	0.06	0.00	0.19	0.10	0.08	0.13	0.16	0.14
Large size containers	0.10	0.17	0.02	0.05	0.00	0.02	0.01	0.07	0.02
Loose	0.01	0.00	0.01	0.03	0.03	0.02	0.02	0.00	0.03
Portion Pack	0.12	0.18	0.02	0.00	0.02	0.02	0.03	0.00	0.03
Powder	0.00	0.00	0.04	0.02	0.01	0.05	0.04	0.12	0.02
Energy Drinks									
Bulk	0.14	0.14	0.00	0.07	0.02	0.00	0.01	0.09	0.10
Family Pack	0.14	0.14	0.00	0.09	0.01	0.03	0.04	0.04	0.00
Large size containers	0.02	0.02	0.01	0.11	0.12	0.11	0.10	0.26	0.23
Portion Pack	0.18	0.04	0.02	0.03	0.00	0.01	0.04	0.03	0.00
Powder	0.02	0.01	0.02	0.01	0.00	0.10	0.11	0.02	0.27
Evaporated Milk									
Bulk	0.11	0.20	0.02	0.05	0.00	0.03	0.00	0.01	0.04
Family Pack	0.01	0.00	0.00	0.08	0.11	0.25	0.31	0.08	0.02
Large size containers	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.00
Portion Pack	0.09	0.09	0.02	0.02	0.08	0.06	0.04	0.21	0.06
FFSD									
Bulk	0.14	0.14	0.00	0.10	0.00	0.06	0.06	0.02	0.01
Family Pack	0.08	0.01	0.10	0.05	0.09	0.18	0.14	0.16	0.32
Large size containers	0.16	0.02	0.02	0.01	0.03	0.07	0.08	0.04	0.16
Portion Pack	0.15	0.14	0.03	0.01	0.00	0.01	0.01	0.03	0.01
Powder	0.08	0.05	0.04	0.06	0.16	0.14	0.18	0.21	0.22
Flavoured Milk									
Bulk	0.06	0.06	0.00	0.01	0.03	0.14	0.17	0.13	0.30
Family Pack	0.14	0.11	0.02	0.13	0.03	0.07	0.01	0.05	0.05
Large size containers	0.14	0.14	0.00	0.11	0.02	0.03	0.03	0.01	0.02
Loose	0.00	0.01	0.36	0.02	0.09	0.04	0.05	0.01	0.02
Portion Pack	0.17	0.07	0.00	0.03	0.01	0.05	0.02	0.01	0.00
Powder	0.04	0.08	0.26	0.01	0.06	0.08	0.01	0.04	0.10
Juice									
Bulk	0.14	0.14	0.00	0.11	0.02	0.04	0.03	0.01	0.01
Family Pack	0.11	0.04	0.00	0.18	0.12	0.16	0.13	0.08	0.08
Large size containers	0.15	0.14	0.00	0.09	0.01	0.02	0.02	0.02	0.01
Portion Pack	0.15	0.12	0.02	0.08	0.06	0.02	0.02	0.03	0.00
IAD									
Family Pack	0.11	0.17	0.05	0.09	0.12	0.01	0.02	0.01	0.04
Large size containers	0.11	0.20	0.02	0.05	0.00	0.03	0.00	0.01	0.04
Portion Pack	0.11	0.19	0.02	0.03	0.02	0.02	0.01	0.01	0.03
Liquid Cream									
Bulk	0.13	0.13	0.00	0.11	0.02	0.03	0.03	0.02	0.04
Family Pack	0.16	0.08	0.00	0.03	0.02	0.05	0.02	0.01	0.04
Large size containers	0.11	0.11	0.01	0.02	0.11	0.03	0.07	0.21	0.16
Loose	0.00	0.00	0.01	0.01	0.05	0.00	0.13	0.17	0.07
Portion Pack	0.07	0.07	0.03	0.17	0.25	0.07	0.01	0.19	0.05
Nectar									
Bulk	0.10	0.19	0.02	0.04	0.01	0.03	0.01	0.01	0.05
Family Pack	0.08	0.03	0.08	0.16	0.28	0.01	0.13	0.01	0.09
Large size containers	0.16	0.05	0.01	0.01	0.08	0.01	0.05	0.08	0.13
Portion Pack	0.13	0.09	0.02	0.06	0.11	0.01	0.06	0.02	0.01
Non Dairy Cream									
Bulk	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.01	0.04
Family Pack	0.14	0.14	0.02	0.11	0.02	0.04	0.04	0.02	0.02
Large size containers	0.14	0.14	0.00	0.11	0.02	0.05	0.04	0.00	0.01
Portion Pack	0.04	0.04	0.03	0.03	0.03	0.01	0.02	0.01	0.00
Powder	0.14	0.14	0.01	0.08	0.00	0.01	0.02	0.02	0.01
Rice, Nut, Grain and Seed Based Drinks									
Bulk	0.01	0.01	0.00	0.00	0.04	0.11	0.01	0.18	0.18
Family Pack	0.10	0.03	0.02	0.05	0.03	0.11	0.04	0.10	0.11
Large size containers	0.15	0.06	0.01	0.13	0.03	0.01	0.03	0.03	0.01
Loose	0.00	0.00	0.01	0.03	0.02	0.01	0.01	0.02	0.03
Portion Pack	0.11	0.19	0.02	0.04	0.00	0.03	0.00	0.01	0.03
Powder	0.09	0.17	0.02	0.04	0.02	0.07	0.01	0.09	0.06
Soy Milk									
Family Pack	0.08	0.05	0.05	0.30	0.16	0.05	0.04	0.01	0.05
Large size containers	0.14	0.14	0.00	0.11	0.02	0.04	0.04	0.01	0.02
Loose	0.11	0.20	0.02	0.04	0.00	0.02	0.00	0.00	0.04
Portion Pack	0.10	0.17	0.02	0.01	0.00	0.16	0.15	0.06	0.04
Powder	0.11	0.19	0.02	0.04	0.00	0.03	0.00	0.01	0.04
Sports Drinks									
Bulk	0.01	0.02	0.00	0.01	0.01	0.03	0.02	0.02	0.05
Family Pack	0.14	0.14	0.00	0.11	0.02	0.05	0.04	0.00	0.01
Large size containers	0.08	0.05	0.05	0.24	0.25	0.03	0.05	0.05	0.01
Portion Pack	0.17	0.07	0.01	0.03	0.04	0.09	0.02	0.02	0.06
Powder	0.05	0.04	0.34	0.01	0.09	0.03	0.04	0.01	0.02
Sweetened Condensed Milk									
Bulk	0.10	0.19	0.02	0.05	0.00	0.03	0.00	0.00	0.03
Family Pack	0.00	0.00	0.02	0.07	0.09	0.09	0.10	0.16	0.05
Large size containers	0.01	0.01	0.01	0.08	0.13	0.36	0.25	0.07	0.02
Loose	0.00	0.01	0.00	0.02	0.04	0.34	0.32	0.12	0.02
Portion Pack	0.04	0.01	0.06	0.10	0.17	0.21	0.21	0.02	0.23
Tea based drinks									
Bulk	0.14	0.14	0.01	0.11	0.02	0.06	0.02	0.01	0.02
Family Pack	0.17	0.01	0.03	0.09	0.12	0.06	0.06	0.05	0.05
Large size containers	0.12	0.05	0.05	0.22	0.20	0.00	0.03	0.01	0.06
Portion Pack	0.16	0.10	0.03	0.05	0.11	0.05	0.04	0.01	0.02
Powder	0.14	0.15	0.02	0.07	0.00	0.00	0.02	0.03	0.01
Traditional Cultured Drinks									
Bulk	0.00	0.00	0.01	0.00	0.01	0.05	0.04	0.02	0.02
Family Pack	0.01	0.01	0.03	0.10	0.19	0.03	0.17	0.36	0.22
Large size containers	0.00	0.00	0.01	0.01	0.07	0.04	0.11	0.23	0.06
Loose	0.00	0.01	0.31	0.00	0.06	0.00	0.03	0.03	0.07
Portion Pack	0.01	0.00	0.10	0.11	0.18	0.02	0.17	0.21	0.18
White Milk									
Bulk	0.05	0.05	0.01	0.00	0.03	0.16	0.18	0.14	0.30
Family Pack	0.05	0.01	0.12	0.28	0.21	0.03	0.07	0.05	0.02
Large size containers	0.14	0.14	0.00	0.10	0.00	0.00	0.02	0.03	0.07
Loose	0.00	0.01	0.33	0.01	0.08	0.03	0.03	0.02	0.02
Portion Pack	0.05	0.05	0.33	0.00	0.09	0.03	0.05	0.02	0.03
Powder	0.05	0.08	0.11	0.07	0.14	0.11	0.20	0.20	0.21
Wine (0-19% alcohol)									
Bulk	0.01	0.00	0.00	0.08	0.10	0.03	0.05	0.07	0.06
Family Pack	0.13	0.08	0.00	0.06	0.13	0.15	0.10	0.01	0.06
Large size containers	0.14	0.02	0.02	0.04	0.10	0.17	0.08	0.09	0.06
Portion Pack	0.15	0.12	0.02	0.05	0.01	0.04	0.04	0.01	0.02
Economics Factors									
consum_capita	0.04	0.06	0.06	0.10	0.04	0.18	0.18	0.02	0.00
gdp_capita	0.04	0.06	0.07	0.08	0.00	0.18	0.19	0.06	0.05
inflation	0.01	0.01	0.01	0.01	0.02	0.03	0.09	0.04	0.04
unemployment_rate	0.02	0.00	0.00	0.01	0.04	0.13	0.07	0.09	0.07
urbanization	0.03	0.03	0.08	0.13	0.07	0.12	0.07	0.03	0.01

PCA Loadings in 2017

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks									
Family Pack	0.03	0.03	0.01	0.10	0.12	0.13	0.12	0.10	0.09
Large size containers	0.01	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.01
Portion Pack	0.03	0.02	0.01	0.07	0.16	0.12	0.01	0.31	0.12
Powder	0.12	0.17	0.00	0.04	0.01	0.01	0.01	0.00	0.00
Buttermilk									
Bulk	0.00	0.01	0.01	0.02	0.03	0.00	0.11	0.11	0.02
Family Pack	0.07	0.08	0.03	0.02	0.04	0.02	0.09	0.13	0.10
Large size containers	0.14	0.15	0.00	0.11	0.03	0.03	0.03	0.00	0.01
Loose	0.00	0.01	0.36	0.02	0.08	0.03	0.06	0.01	0.01
Portion Pack	0.01	0.00	0.36	0.03	0.06	0.04	0.08	0.03	0.05
Coffee based drinks									
Bulk	0.00	0.01	0.00	0.03	0.06	0.39	0.27	0.11	0.02
Family Pack	0.07	0.06	0.05	0.26	0.26	0.04	0.06	0.03	0.01
Large size containers	0.05	0.03	0.05	0.28	0.27	0.04	0.06	0.04	0.00
Portion Pack	0.09	0.01	0.06	0.26	0.25	0.07	0.04	0.03	0.00
Powder	0.00	0.01	0.02	0.01	0.00	0.08	0.11	0.09	0.16
Drinking Yoghurt									
Bulk	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.01
Family Pack	0.05	0.06	0.00	0.17	0.12	0.06	0.14	0.17	0.17
Large size containers	0.10	0.17	0.02	0.05	0.00	0.02	0.01	0.05	0.02
Loose	0.01	0.00	0.01	0.03	0.03	0.01	0.02	0.01	0.03
Portion Pack	0.12	0.18	0.02	0.01	0.02	0.02	0.03	0.00	0.03
Powder	0.00	0.00	0.03	0.02	0.01	0.04	0.04	0.10	0.02
Energy Drinks									
Bulk	0.13	0.14	0.00	0.06	0.02	0.01	0.00	0.08	0.12
Family Pack	0.14	0.15	0.00	0.09	0.00	0.02	0.04	0.03	0.01
Large size containers	0.02	0.02	0.01	0.11	0.13	0.11	0.11	0.24	0.32
Portion Pack	0.18	0.03	0.01	0.02	0.00	0.01	0.04	0.03	0.01
Powder	0.02	0.01	0.02	0.01	0.00	0.09	0.13	0.04	0.22
Evaporated Milk									
Bulk	0.11	0.19	0.02	0.05	0.00	0.03	0.01	0.01	0.03
Family Pack	0.01	0.00	0.00	0.08	0.13	0.31	0.26	0.05	0.00
Large size containers	0.00	0.00	0.01	0.01	0.00	0.01	0.02	0.01	0.01
Portion Pack	0.09	0.09	0.02	0.01	0.08	0.05	0.03	0.18	0.13
FFSD									
Bulk	0.14	0.15	0.00	0.10	0.00	0.05	0.06	0.02	0.02
Family Pack	0.08	0.01	0.10	0.06	0.10	0.15	0.15	0.15	0.30
Large size containers	0.16	0.03	0.01	0.01	0.05	0.05	0.08	0.04	0.16
Portion Pack	0.15	0.14	0.04	0.01	0.00	0.01	0.01	0.02	0.01
Powder	0.07	0.05	0.04	0.07	0.18	0.11	0.21	0.23	0.23
Flavoured Milk									
Bulk	0.06	0.06	0.00	0.01	0.03	0.11	0.19	0.13	0.31
Family Pack	0.14	0.12	0.02	0.12	0.02	0.07	0.02	0.06	0.06
Large size containers	0.13	0.15	0.00	0.11	0.03	0.03	0.03	0.01	0.01
Loose	0.00	0.01	0.36	0.02	0.08	0.03	0.06	0.01	0.01
Portion Pack	0.17	0.07	0.01	0.03	0.01	0.05	0.02	0.00	0.00
Powder	0.04	0.08	0.27	0.01	0.06	0.08	0.00	0.04	0.06
Juice									
Bulk	0.14	0.15	0.00	0.11	0.02	0.03	0.03	0.00	0.00
Family Pack	0.11	0.04	0.01	0.18	0.12	0.14	0.14	0.07	0.13
Large size containers	0.14	0.14	0.00	0.08	0.02	0.01	0.03	0.02	0.00
Portion Pack	0.15	0.12	0.02	0.09	0.06	0.02	0.02	0.03	0.01
LAD									
Family Pack	0.11	0.17	0.04	0.08	0.11	0.01	0.01	0.01	0.03
Large size containers	0.11	0.19	0.02	0.05	0.00	0.03	0.01	0.01	0.03
Portion Pack	0.11	0.19	0.02	0.03	0.02	0.02	0.00	0.01	0.02
Liquid Cream									
Bulk	0.13	0.13	0.00	0.11	0.02	0.02	0.03	0.02	0.02
Family Pack	0.17	0.06	0.01	0.02	0.02	0.04	0.02	0.00	0.04
Large size containers	0.11	0.11	0.01	0.02	0.10	0.03	0.05	0.17	0.21
Loose	0.00	0.00	0.01	0.01	0.06	0.02	0.16	0.23	0.08
Portion Pack	0.07	0.07	0.02	0.17	0.26	0.07	0.01	0.17	0.08
Nectar									
Bulk	0.11	0.19	0.02	0.04	0.00	0.03	0.00	0.01	0.04
Family Pack	0.09	0.03	0.08	0.15	0.28	0.04	0.12	0.03	0.12
Large size containers	0.16	0.05	0.01	0.00	0.07	0.03	0.04	0.07	0.11
Portion Pack	0.13	0.08	0.02	0.05	0.11	0.01	0.07	0.00	0.02
Non Dairy Cream									
Bulk	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.01	0.05
Family Pack	0.14	0.15	0.02	0.10	0.03	0.04	0.05	0.02	0.01
Large size containers	0.14	0.15	0.00	0.11	0.03	0.04	0.04	0.00	0.00
Portion Pack	0.05	0.05	0.03	0.03	0.03	0.01	0.02	0.01	0.01
Powder	0.15	0.13	0.01	0.08	0.00	0.01	0.02	0.02	0.00
Rice, Nut, Grain and Seed Based Drinks									
Bulk	0.00	0.01	0.00	0.00	0.04	0.09	0.01	0.17	0.14
Family Pack	0.12	0.03	0.02	0.04	0.02	0.08	0.07	0.03	0.04
Large size containers	0.15	0.08	0.01	0.13	0.03	0.01	0.03	0.02	0.01
Loose	0.00	0.00	0.01	0.03	0.02	0.00	0.00	0.01	0.02
Portion Pack	0.11	0.19	0.02	0.05	0.01	0.03	0.01	0.01	0.02
Powder	0.10	0.16	0.02	0.04	0.02	0.06	0.01	0.09	0.05
Soy Milk									
Family Pack	0.07	0.04	0.05	0.32	0.15	0.04	0.05	0.01	0.04
Large size containers	0.13	0.15	0.00	0.11	0.03	0.03	0.03	0.01	0.01
Loose	0.11	0.19	0.02	0.05	0.00	0.02	0.01	0.00	0.03
Portion Pack	0.10	0.17	0.02	0.02	0.01	0.16	0.11	0.06	0.04
Powder	0.11	0.19	0.02	0.05	0.00	0.03	0.01	0.01	0.03
Sports Drinks									
Bulk	0.01	0.02	0.01	0.01	0.01	0.03	0.03	0.02	0.02
Family Pack	0.14	0.15	0.00	0.10	0.03	0.04	0.04	0.00	0.00
Large size containers	0.07	0.05	0.05	0.26	0.24	0.04	0.05	0.04	0.01
Portion Pack	0.17	0.06	0.01	0.02	0.03	0.09	0.01	0.02	0.03
Powder	0.05	0.04	0.34	0.00	0.09	0.03	0.05	0.00	0.01
Sweetened Condensed Milk									
Bulk	0.11	0.19	0.02	0.05	0.00	0.03	0.00	0.00	0.02
Family Pack	0.00	0.00	0.01	0.04	0.05	0.07	0.07	0.18	0.08
Large size containers	0.01	0.01	0.01	0.09	0.16	0.40	0.17	0.05	0.03
Loose	0.00	0.01	0.00	0.02	0.06	0.38	0.26	0.11	0.03
Portion Pack	0.04	0.01	0.06	0.09	0.16	0.18	0.24	0.02	0.21
Tea based drinks									
Bulk	0.14	0.15	0.01	0.11	0.02	0.05	0.03	0.01	0.01
Family Pack	0.16	0.00	0.03	0.10	0.13	0.04	0.07	0.05	0.04
Large size containers	0.12	0.06	0.05	0.22	0.19	0.00	0.03	0.01	0.06
Portion Pack	0.16	0.10	0.03	0.05	0.10	0.05	0.04	0.01	0.02
Powder	0.14	0.15	0.02	0.07	0.01	0.01	0.02	0.03	0.01
Traditional Cultured Drinks									
Bulk	0.00	0.00	0.01	0.00	0.01	0.05	0.05	0.02	0.00
Family Pack	0.01	0.01	0.02	0.08	0.17	0.01	0.19	0.38	0.20
Large size containers	0.00	0.00	0.01	0.01	0.07	0.05	0.14	0.30	0.08
Loose	0.00	0.01	0.31	0.00	0.06	0.01	0.03	0.04	0.07
Portion Pack	0.00	0.00	0.10	0.08	0.15	0.01	0.17	0.21	0.17
White Milk									
Bulk	0.05	0.05	0.01	0.00	0.04	0.13	0.21	0.14	0.31
Family Pack	0.05	0.01	0.12	0.27	0.22	0.04	0.09	0.05	0.02
Large size containers	0.14	0.15	0.00	0.09	0.00	0.00	0.03	0.03	0.08
Loose	0.00	0.01	0.33	0.01	0.07	0.02	0.03	0.01	0.01
Portion Pack	0.05	0.05	0.34	0.01	0.08	0.02	0.06	0.02	0.01
Powder	0.05	0.08	0.11	0.07	0.15	0.09	0.22	0.20	0.20
Wine (0-19% alcohol)									
Bulk	0.01	0.00	0.00	0.08	0.10	0.03	0.06	0.08	0.07
Family Pack	0.13	0.08	0.00	0.07	0.13	0.14	0.11	0.00	0.06
Large size containers	0.14	0.01	0.02	0.03	0.09	0.15	0.10	0.09	0.08
Portion Pack	0.15	0.13	0.02	0.05	0.01	0.03	0.04	0.01	0.01
Economics Factors									
consum_capita	0.03	0.06	0.06	0.11	0.05	0.15	0.22	0.02	0.04
gdp_capita	0.04	0.06	0.07	0.08	0.01	0.16	0.22	0.06	0.02
inflation	0.01	0.00	0.00	0.02	0.01	0.01	0.05	0.03	0.01
unemployment_rate	0.02	0.00	0.00	0.01	0.05	0.12	0.09	0.10	0.09
urbanization	0.03	0.03	0.08	0.14	0.08	0.11	0.08	0.01	0.04

PCA Loadings in 2018

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks									
Family Pack	0.02	0.03	0.00	0.10	0.11	0.15	0.10	0.09	0.07
Large size containers	0.01	0.00	0.01	0.01	0.00	0.02	0.00	0.00	0.02
Portion Pack	0.02	0.02	0.02	0.07	0.14	0.09	0.03	0.29	0.18
Powder	0.13	0.17	0.01	0.04	0.01	0.01	0.01	0.00	0.00
Buttermilk									
Bulk	0.00	0.01	0.01	0.02	0.02	0.01	0.12	0.09	0.02
Family Pack	0.07	0.08	0.02	0.02	0.04	0.03	0.09	0.12	0.09
Large size containers	0.13	0.15	0.00	0.11	0.03	0.04	0.03	0.00	0.00
Loose	0.00	0.01	0.36	0.01	0.09	0.03	0.05	0.00	0.01
Portion Pack	0.01	0.00	0.36	0.02	0.07	0.05	0.07	0.02	0.05
Coffee based drinks									
Bulk	0.00	0.01	0.00	0.02	0.06	0.36	0.31	0.16	0.04
Family Pack	0.08	0.07	0.04	0.25	0.25	0.06	0.06	0.02	0.00
Large size containers	0.06	0.04	0.05	0.29	0.26	0.05	0.06	0.03	0.01
Portion Pack	0.09	0.01	0.05	0.26	0.24	0.07	0.04	0.03	0.00
Powder	0.00	0.01	0.01	0.01	0.00	0.08	0.07	0.08	0.10
Drinking Yoghurt									
Bulk	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.01
Family Pack	0.06	0.07	0.00	0.17	0.13	0.08	0.12	0.16	0.19
Large size containers	0.11	0.17	0.02	0.05	0.00	0.02	0.01	0.04	0.01
Loose	0.01	0.00	0.01	0.03	0.03	0.02	0.02	0.01	0.03
Portion Pack	0.12	0.18	0.02	0.01	0.01	0.02	0.03	0.00	0.03
Powder	0.00	0.00	0.03	0.02	0.01	0.05	0.03	0.10	0.02
Energy Drinks									
Bulk	0.13	0.15	0.00	0.06	0.02	0.01	0.00	0.08	0.12
Family Pack	0.13	0.15	0.00	0.09	0.01	0.02	0.04	0.04	0.00
Large size containers	0.02	0.02	0.01	0.11	0.12	0.12	0.10	0.23	0.33
Portion Pack	0.17	0.06	0.02	0.03	0.01	0.00	0.03	0.03	0.03
Powder	0.02	0.01	0.02	0.01	0.00	0.10	0.10	0.02	0.17
Evaporated Milk									
Bulk	0.11	0.19	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Family Pack	0.01	0.00	0.01	0.08	0.14	0.26	0.28	0.01	0.00
Large size containers	-	-	-	-	-	-	0.00	0.00	-
Portion Pack	0.09	0.09	0.02	0.01	0.08	0.06	0.05	0.16	0.13
FFSD									
Bulk	0.13	0.15	0.00	0.10	0.00	0.06	0.06	0.02	0.03
Family Pack	0.07	0.01	0.12	0.06	0.09	0.18	0.13	0.18	0.28
Large size containers	0.15	0.03	0.01	0.00	0.06	0.07	0.08	0.06	0.17
Portion Pack	0.15	0.14	0.04	0.01	0.00	0.02	0.01	0.02	0.02
Powder	0.07	0.05	0.04	0.07	0.19	0.14	0.17	0.21	0.26
Flavoured Milk									
Bulk	0.06	0.06	0.00	0.01	0.03	0.14	0.19	0.16	0.28
Family Pack	0.14	0.12	0.01	0.11	0.01	0.06	0.01	0.07	0.05
Large size containers	0.13	0.15	0.01	0.11	0.03	0.03	0.03	0.00	0.00
Loose	0.00	0.01	0.36	0.01	0.09	0.03	0.05	0.00	0.01
Portion Pack	0.17	0.06	0.01	0.02	0.00	0.06	0.02	0.00	0.00
Powder	0.04	0.07	0.27	0.02	0.06	0.08	0.02	0.03	0.03
Juice									
Bulk	0.13	0.15	0.00	0.11	0.02	0.03	0.03	0.00	0.00
Family Pack	0.11	0.04	0.00	0.18	0.12	0.16	0.12	0.06	0.15
Large size containers	0.14	0.14	0.01	0.08	0.01	0.02	0.03	0.01	0.01
Portion Pack	0.15	0.12	0.02	0.09	0.05	0.02	0.02	0.03	0.01
LAD									
Family Pack	0.11	0.16	0.04	0.07	0.11	0.00	0.02	0.00	0.02
Large size containers	0.11	0.19	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Portion Pack	0.12	0.19	0.02	0.03	0.02	0.02	0.00	0.01	0.02
Liquid Cream									
Bulk	0.13	0.14	0.01	0.11	0.02	0.03	0.04	0.01	0.02
Family Pack	0.17	0.04	0.01	0.03	0.01	0.04	0.01	0.01	0.04
Large size containers	0.11	0.11	0.01	0.02	0.11	0.03	0.06	0.16	0.21
Loose	0.00	0.00	0.01	0.01	0.06	0.01	0.20	0.25	0.08
Portion Pack	0.07	0.07	0.03	0.17	0.26	0.06	0.02	0.17	0.07
Nectar									
Bulk	0.11	0.19	0.02	0.04	0.00	0.03	0.01	0.00	0.03
Family Pack	0.08	0.02	0.09	0.15	0.28	0.04	0.12	0.01	0.11
Large size containers	0.16	0.05	0.01	0.01	0.07	0.02	0.05	0.05	0.12
Portion Pack	0.13	0.08	0.02	0.06	0.12	0.02	0.06	0.01	0.04
Non Dairy Cream									
Bulk	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.01	0.06
Family Pack	0.13	0.15	0.02	0.10	0.03	0.04	0.04	0.01	0.01
Large size containers	0.13	0.15	0.00	0.11	0.03	0.04	0.04	0.00	0.00
Portion Pack	0.05	0.06	0.03	0.04	0.03	0.01	0.02	0.02	0.02
Powder	0.15	0.14	0.02	0.07	0.00	0.01	0.01	0.02	0.00
Rice, Nut, Grain and Seed Based Drinks									
Bulk	0.00	0.01	0.00	0.00	0.04	0.10	0.02	0.18	0.15
Family Pack	0.13	0.00	0.02	0.03	0.04	0.09	0.06	0.04	0.03
Large size containers	0.15	0.09	0.01	0.13	0.03	0.02	0.03	0.02	0.01
Loose	0.00	0.00	0.01	0.03	0.02	0.00	0.00	0.01	0.02
Portion Pack	0.11	0.19	0.02	0.04	0.01	0.03	0.00	0.01	0.02
Powder	0.10	0.16	0.02	0.04	0.02	0.06	0.01	0.10	0.05
Soy Milk									
Family Pack	0.06	0.04	0.04	0.33	0.16	0.04	0.05	0.00	0.03
Large size containers	0.13	0.15	0.00	0.11	0.02	0.03	0.03	0.00	0.01
Loose	0.11	0.19	0.02	0.05	0.00	0.02	0.01	0.00	0.03
Portion Pack	0.10	0.17	0.02	0.02	0.00	0.14	0.12	0.07	0.05
Powder	0.11	0.19	0.02	0.05	0.00	0.03	0.00	0.01	0.02
Sports Drinks									
Bulk	0.01	0.02	0.01	0.02	0.01	0.03	0.03	0.01	0.02
Family Pack	0.13	0.15	0.00	0.10	0.03	0.05	0.04	0.00	0.01
Large size containers	0.08	0.06	0.04	0.25	0.24	0.05	0.05	0.04	0.02
Portion Pack	0.17	0.06	0.01	0.02	0.03	0.08	0.03	0.01	0.02
Powder	0.05	0.04	0.34	0.01	0.09	0.03	0.05	0.00	0.01
Sweetened Condensed Milk									
Bulk	0.11	0.18	0.02	0.05	0.00	0.02	0.00	0.00	0.02
Family Pack	0.00	0.00	0.01	0.03	0.04	0.07	0.06	0.16	0.13
Large size containers	0.01	0.01	0.02	0.09	0.18	0.38	0.19	0.08	0.05
Loose	0.00	0.01	0.00	0.02	0.06	0.35	0.30	0.17	0.05
Portion Pack	0.04	0.01	0.07	0.08	0.16	0.21	0.19	0.03	0.18
Tea based drinks									
Bulk	0.13	0.15	0.00	0.11	0.02	0.06	0.02	0.01	0.01
Family Pack	0.16	0.00	0.04	0.12	0.14	0.05	0.07	0.04	0.03
Large size containers	0.12	0.06	0.04	0.23	0.19	0.01	0.03	0.01	0.06
Portion Pack	0.16	0.10	0.03	0.05	0.10	0.04	0.04	0.01	0.01
Powder	0.14	0.14	0.02	0.07	0.01	0.01	0.01	0.03	0.01
Traditional Cultured Drinks									
Bulk	0.00	0.00	0.01	0.00	0.01	0.05	0.04	0.02	0.00
Family Pack	0.01	0.01	0.03	0.08	0.17	0.01	0.22	0.36	0.22
Large size containers	0.00	0.00	0.01	0.01	0.07	0.06	0.19	0.31	0.06
Loose	0.00	0.01	0.30	0.01	0.07	0.01	0.04	0.03	0.07
Portion Pack	0.00	0.00	0.10	0.07	0.14	0.01	0.17	0.18	0.16
White Milk									
Bulk	0.05	0.05	0.01	0.00	0.03	0.16	0.21	0.17	0.28
Family Pack	0.05	0.01	0.14	0.27	0.22	0.03	0.09	0.05	0.03
Large size containers	0.13	0.15	0.00	0.09	0.00	0.01	0.03	0.04	0.08
Loose	0.00	0.01	0.33	0.00	0.08	0.02	0.02	0.01	0.01
Portion Pack	0.05	0.05	0.34	0.01	0.09	0.03	0.06	0.01	0.02
Powder	0.05	0.08	0.11	0.07	0.16	0.12	0.19	0.19	0.23
Wine (0-19% alcohol)									
Bulk	0.01	0.00	0.01	0.08	0.11	0.02	0.06	0.10	0.09
Family Pack	0.13	0.08	0.00	0.06	0.12	0.16	0.09	0.01	0.08
Large size containers	0.14	0.02	0.02	0.03	0.09	0.17	0.08	0.09	0.07
Portion Pack	0.15	0.12	0.02	0.05	0.01	0.03	0.03	0.01	0.01
Economics Factors									
consum_capita	0.03	0.05	0.05	0.11	0.04	0.18	0.20	0.01	0.04
gdp_capita	0.04	0.06	0.06	0.08	0.00	0.17	0.21	0.06	0.04
inflation	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.03	0.01
unemployment_rate	0.02	0.00	0.01	0.01	0.06	0.11	0.14	0.11	0.13
urbanization	0.03	0.03	0.07	0.14	0.08	0.12	0.07	0.01	0.04

PCA Loadings in 2019

Features	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9
Baby & Toddler Dairy Drinks									
Family Pack	0.02	0.02	0.01	0.10	0.11	0.15	0.10	0.07	0.07
Large size containers	0.01	0.00	0.01	0.01	0.00	0.02	0.00	0.00	0.02
Portion Pack	0.02	0.01	0.01	0.07	0.13	0.09	0.03	0.26	0.19
Powder	0.13	0.17	0.01	0.04	0.01	0.01	0.01	0.00	0.01
Buttermilk									
Bulk	0.00	0.01	0.01	0.02	0.02	0.01	0.12	0.08	0.02
Family Pack	0.07	0.08	0.03	0.01	0.04	0.02	0.10	0.11	0.09
Large size containers	0.13	0.15	0.00	0.11	0.03	0.04	0.03	0.00	0.00
Loose	0.00	0.01	0.36	0.01	0.08	0.03	0.05	0.00	0.01
Portion Pack	0.01	0.00	0.36	0.02	0.07	0.04	0.06	0.02	0.04
Coffee based drinks									
Bulk	0.00	0.01	0.00	0.02	0.06	0.36	0.29	0.18	0.04
Family Pack	0.08	0.07	0.04	0.25	0.24	0.05	0.06	0.02	0.00
Large size containers	0.06	0.04	0.05	0.28	0.26	0.05	0.07	0.03	0.02
Portion Pack	0.10	0.01	0.05	0.26	0.24	0.07	0.04	0.03	0.00
Powder	0.00	0.01	0.01	0.01	0.00	0.07	0.06	0.07	0.08
Drinking Yoghurt									
Bulk	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.01
Family Pack	0.06	0.07	0.00	0.16	0.14	0.08	0.12	0.17	0.20
Large size containers	0.11	0.17	0.02	0.05	0.00	0.02	0.01	0.04	0.01
Loose	0.01	0.00	0.01	0.03	0.03	0.01	0.02	0.01	0.03
Portion Pack	0.12	0.18	0.02	0.02	0.01	0.02	0.03	0.00	0.03
Powder	0.00	0.00	0.03	0.02	0.01	0.05	0.03	0.10	0.03
Energy Drinks									
Bulk	0.13	0.15	0.00	0.06	0.02	0.00	0.01	0.08	0.12
Family Pack	0.13	0.15	0.00	0.09	0.01	0.02	0.04	0.04	0.00
Large size containers	0.02	0.02	0.01	0.11	0.13	0.11	0.11	0.22	0.33
Portion Pack	0.17	0.06	0.02	0.03	0.01	0.00	0.03	0.04	0.04
Powder	0.02	0.01	0.02	0.01	0.00	0.09	0.10	0.02	0.15
Evaporated Milk									
Bulk	0.11	0.19	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Family Pack	0.01	0.00	0.01	0.08	0.15	0.26	0.27	0.01	0.00
Large size containers	-	-	-	-	-	-	0.00	-	0.00
Portion Pack	0.08	0.09	0.02	0.01	0.08	0.06	0.05	0.16	0.14
FFSD									
Bulk	0.13	0.15	0.00	0.10	0.00	0.06	0.06	0.03	0.03
Family Pack	0.07	0.01	0.12	0.06	0.09	0.18	0.13	0.20	0.26
Large size containers	0.15	0.03	0.01	0.00	0.06	0.07	0.08	0.07	0.17
Portion Pack	0.15	0.13	0.05	0.01	0.00	0.02	0.01	0.02	0.01
Powder	0.07	0.05	0.04	0.07	0.20	0.13	0.17	0.20	0.26
Flavoured Milk									
Bulk	0.05	0.06	0.00	0.00	0.03	0.14	0.19	0.17	0.27
Family Pack	0.14	0.12	0.01	0.11	0.01	0.06	0.00	0.07	0.05
Large size containers	0.13	0.15	0.00	0.11	0.03	0.03	0.03	0.00	0.00
Loose	0.00	0.01	0.36	0.01	0.08	0.03	0.05	0.00	0.01
Portion Pack	0.17	0.06	0.01	0.02	0.00	0.06	0.02	0.00	0.00
Powder	0.03	0.07	0.27	0.02	0.06	0.09	0.02	0.03	0.03
Juice									
Bulk	0.13	0.15	0.00	0.11	0.03	0.03	0.03	0.00	0.00
Family Pack	0.11	0.04	0.01	0.18	0.13	0.15	0.13	0.06	0.15
Large size containers	0.14	0.14	0.00	0.08	0.01	0.01	0.03	0.01	0.01
Portion Pack	0.15	0.12	0.02	0.09	0.05	0.01	0.03	0.03	0.02
IAD									
Family Pack	0.11	0.16	0.04	0.09	0.12	0.00	0.02	0.00	0.02
Large size containers	0.11	0.19	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Portion Pack	0.12	0.18	0.02	0.03	0.02	0.02	0.00	0.01	0.02
Liquid Cream									
Bulk	0.13	0.14	0.00	0.11	0.02	0.03	0.04	0.01	0.01
Family Pack	0.17	0.03	0.00	0.04	0.01	0.04	0.01	0.02	0.03
Large size containers	0.10	0.11	0.01	0.03	0.12	0.03	0.07	0.16	0.23
Loose	0.00	0.00	0.01	0.01	0.06	0.02	0.21	0.27	0.10
Portion Pack	0.07	0.07	0.03	0.16	0.27	0.05	0.01	0.17	0.06
Nectar									
Bulk	0.11	0.18	0.02	0.05	0.00	0.03	0.00	0.01	0.02
Family Pack	0.09	0.02	0.07	0.14	0.28	0.04	0.12	0.02	0.12
Large size containers	0.17	0.04	0.01	0.02	0.05	0.03	0.04	0.04	0.11
Portion Pack	0.14	0.08	0.02	0.04	0.12	0.02	0.05	0.01	0.04
Non Dairy Cream									
Bulk	-	-	-	-	-	-	-	-	-
Family Pack	0.13	0.15	0.02	0.10	0.03	0.04	0.04	0.01	0.01
Large size containers	0.13	0.15	0.00	0.11	0.03	0.04	0.04	0.00	0.00
Portion Pack	0.07	0.08	0.03	0.04	0.03	0.01	0.02	0.01	0.02
Powder	0.14	0.14	0.02	0.07	0.00	0.01	0.01	0.02	0.00
Rice, Nut, Grain and Seed Based Drinks									
Bulk	0.00	0.00	0.00	0.00	0.05	0.11	0.03	0.21	0.15
Family Pack	0.12	0.01	0.02	0.02	0.05	0.10	0.07	0.04	0.03
Large size containers	0.14	0.10	0.01	0.12	0.03	0.02	0.03	0.02	0.02
Loose	0.00	0.00	0.01	0.03	0.02	0.00	0.00	0.01	0.02
Portion Pack	0.12	0.18	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Powder	0.10	0.16	0.01	0.04	0.02	0.06	0.01	0.11	0.05
Soy Milk									
Family Pack	0.06	0.03	0.04	0.33	0.16	0.03	0.04	0.00	0.03
Large size containers	0.13	0.15	0.00	0.11	0.03	0.03	0.03	0.00	0.00
Loose	0.11	0.19	0.02	0.05	0.01	0.02	0.01	0.01	0.03
Portion Pack	0.11	0.17	0.02	0.02	0.00	0.15	0.11	0.08	0.05
Powder	0.11	0.19	0.02	0.05	0.01	0.03	0.00	0.02	0.02
Sports Drinks									
Bulk	0.01	0.02	0.01	0.02	0.01	0.03	0.03	0.00	0.02
Family Pack	0.13	0.15	0.00	0.10	0.03	0.04	0.04	0.00	0.01
Large size containers	0.08	0.06	0.04	0.26	0.23	0.05	0.06	0.04	0.02
Portion Pack	0.17	0.06	0.00	0.02	0.03	0.08	0.03	0.00	0.01
Powder	0.04	0.04	0.34	0.01	0.08	0.03	0.05	0.00	0.02
Sweetened Condensed Milk									
Bulk	0.11	0.18	0.02	0.05	0.01	0.03	0.00	0.01	0.02
Family Pack	0.00	0.00	0.01	0.03	0.04	0.07	0.06	0.14	0.14
Large size containers	0.01	0.01	0.02	0.09	0.18	0.38	0.18	0.10	0.05
Loose	0.00	0.01	0.00	0.02	0.06	0.35	0.28	0.18	0.04
Portion Pack	0.04	0.01	0.07	0.08	0.16	0.21	0.19	0.04	0.18
Tea based drinks									
Bulk	0.13	0.15	0.00	0.11	0.02	0.06	0.03	0.01	0.01
Family Pack	0.16	0.00	0.04	0.14	0.15	0.04	0.07	0.04	0.03
Large size containers	0.12	0.06	0.04	0.23	0.18	0.01	0.03	0.02	0.06
Portion Pack	0.16	0.11	0.03	0.05	0.10	0.04	0.04	0.01	0.01
Powder	0.14	0.14	0.02	0.07	0.01	0.01	0.01	0.03	0.01
Traditional Cultured Drinks									
Bulk	0.00	0.00	0.01	0.00	0.00	0.05	0.04	0.02	0.00
Family Pack	0.01	0.01	0.02	0.08	0.16	0.02	0.23	0.34	0.23
Large size containers	0.00	0.00	0.01	0.01	0.08	0.07	0.20	0.33	0.08
Loose	0.00	0.01	0.31	0.01	0.06	0.01	0.04	0.02	0.07
Portion Pack	0.00	0.00	0.11	0.08	0.14	0.02	0.17	0.16	0.16
White Milk									
Bulk	0.04	0.05	0.01	0.00	0.03	0.16	0.21	0.18	0.27
Family Pack	0.05	0.00	0.13	0.27	0.22	0.03	0.08	0.06	0.03
Large size containers	0.13	0.15	0.00	0.09	0.00	0.01	0.03	0.04	0.07
Loose	0.00	0.01	0.32	0.00	0.07	0.02	0.02	0.01	0.00
Portion Pack	0.05	0.05	0.34	0.00	0.08	0.03	0.05	0.01	0.02
Powder	0.05	0.08	0.11	0.07	0.16	0.12	0.19	0.18	0.23
Wine (0-19% alcohol)									
Bulk	0.01	0.00	0.01	0.08	0.11	0.01	0.06	0.11	0.11
Family Pack	0.13	0.08	0.01	0.06	0.12	0.15	0.10	0.02	0.08
Large size containers	0.14	0.02	0.02	0.03	0.08	0.16	0.08	0.08	0.07
Portion Pack	0.15	0.12	0.02	0.06	0.00	0.03	0.03	0.01	0.01
Economics Factors									
consum_capita	0.03	0.05	0.05	0.11	0.04	0.17	0.20	0.01	0.04
gdp_capita	0.04	0.06	0.07	0.08	0.00	0.16	0.22	0.06	0.03
inflation	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.02
unemployment_rate	0.02	0.01	0.01	0.02	0.06	0.12	0.16	0.13	0.13
urbanization	0.03	0.03	0.07	0.14	0.08	0.12	0.07	0.01	0.04

F. Proportion of Volume by Package Group (2015 - 2017)

2015

Product Group	Cluster						
	1	2	3	4	5	6	Total
Baby & Toddler Dairy Drinks	8%	3%	1%	7%	1%	0%	3%
Buttermilk	0%	0%	0%	0%	0%	3%	0%
Coffee based drinks	4%	0%	1%	1%	16%	0%	1%
Drinking Yoghurt	8%	2%	0%	3%	2%	0%	2%
Energy Drinks	5%	2%	4%	2%	2%	0%	2%
Evaporated Milk	1%	1%	0%	0%	0%	0%	0%
FFSD	12%	17%	16%	16%	9%	4%	14%
Flavoured Milk	10%	3%	5%	5%	4%	2%	4%
Juice	2%	4%	10%	1%	6%	0%	4%
LAD	0%	0%	0%	10%	3%	0%	2%
Liquid Cream	0%	1%	1%	0%	0%	0%	1%
Nectar	1%	5%	2%	2%	2%	0%	3%
Non Dairy Cream	0%	1%	2%	1%	0%	0%	1%
Rice, Nut, Grain and Seed Based Drinks	0%	1%	1%	7%	0%	0%	2%
Soy Milk	16%	1%	1%	16%	2%	0%	3%
Sports Drinks	6%	1%	8%	2%	5%	0%	2%
Sweetened Condensed Milk	2%	1%	0%	0%	0%	0%	0%
Tea based drinks	12%	4%	11%	12%	32%	0%	7%
Traditional Cultured Drinks	0%	4%	0%	0%	0%	2%	2%
White Milk	13%	44%	31%	11%	16%	88%	41%
Wine (0-19% alcohol)	1%	6%	6%	2%	2%	0%	4%
Grand Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

2016

Product Group	Cluster						
	1	2	3	4	5	6	Total
Baby & Toddler Dairy Drinks	2%	3%	1%	8%	1%	0%	3%
Buttermilk	0%	0%	0%	0%	0%	3%	1%
Coffee based drinks	0%	1%	1%	1%	16%	0%	1%
Drinking Yoghurt	2%	2%	0%	4%	2%	0%	2%
Energy Drinks	2%	2%	4%	3%	2%	0%	2%
Evaporated Milk	0%	1%	0%	0%	0%	0%	0%
FFSD	19%	15%	16%	16%	9%	4%	14%
Flavoured Milk	2%	4%	5%	4%	3%	2%	4%
Juice	6%	2%	10%	1%	6%	0%	4%
LAD	0%	1%	0%	10%	3%	0%	2%
Liquid Cream	2%	1%	1%	0%	0%	0%	1%
Nectar	5%	4%	2%	2%	1%	0%	3%
Non Dairy Cream	0%	1%	2%	2%	0%	0%	1%
Rice, Nut, Grain and Seed Based Drinks	1%	1%	2%	6%	0%	0%	2%
Soy Milk	0%	2%	0%	16%	2%	0%	3%
Sports Drinks	1%	1%	8%	2%	5%	0%	2%
Sweetened Condensed Milk	1%	1%	0%	0%	0%	0%	0%
Tea based drinks	3%	5%	12%	13%	33%	0%	7%
Traditional Cultured Drinks	2%	5%	0%	0%	0%	2%	2%
White Milk	41%	45%	30%	11%	15%	87%	41%
Wine (0-19% alcohol)	10%	3%	6%	2%	2%	0%	4%
Grand Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

2017

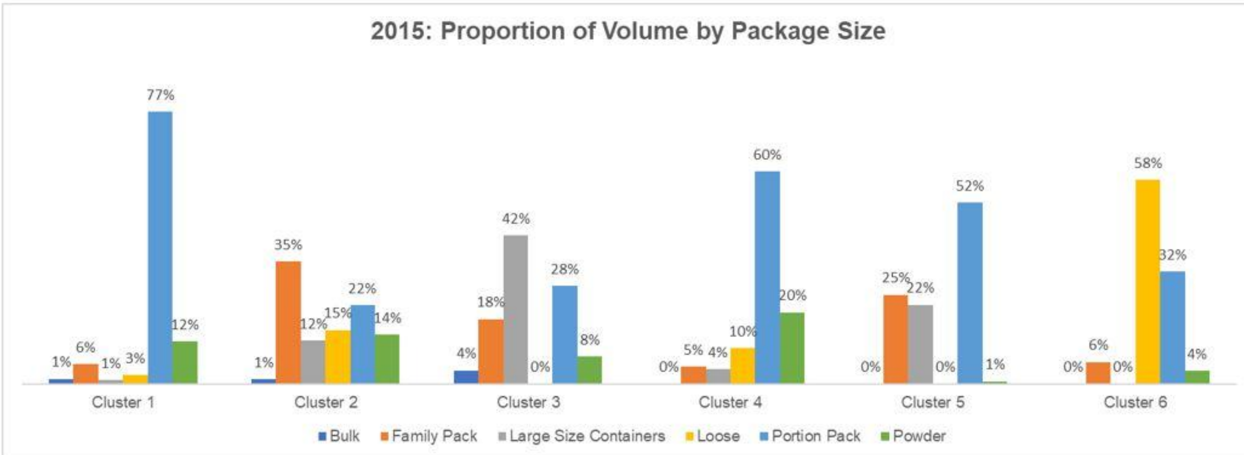
Product Group	Cluster						
	1	2	3	4	5	6	Total
Baby & Toddler Dairy Drinks	3%	3%	1%	8%	2%	0%	3%
Buttermilk	0%	0%	0%	0%	0%	3%	1%
Coffee based drinks	0%	1%	1%	1%	3%	0%	1%
Drinking Yoghurt	1%	2%	0%	5%	2%	0%	2%
Energy Drinks	2%	3%	4%	3%	2%	0%	2%
Evaporated Milk	0%	1%	0%	0%	0%	0%	0%
FFSD	9%	16%	15%	16%	18%	4%	14%
Flavoured Milk	2%	4%	5%	4%	3%	2%	4%
Juice	3%	3%	9%	1%	6%	0%	3%
LAD	0%	1%	0%	10%	0%	0%	2%
Liquid Cream	1%	1%	1%	0%	1%	0%	1%
Nectar	8%	4%	2%	2%	3%	0%	3%
Non Dairy Cream	0%	1%	2%	1%	0%	0%	1%
Rice, Nut, Grain and Seed Based Drinks	0%	2%	2%	5%	1%	0%	2%
Soy Milk	0%	1%	0%	16%	2%	0%	3%
Sports Drinks	0%	1%	8%	2%	2%	0%	2%
Sweetened Condensed Milk	1%	1%	0%	0%	1%	0%	0%
Tea based drinks	3%	4%	12%	13%	9%	0%	7%
Traditional Cultured Drinks	17%	3%	0%	0%	0%	2%	2%
White Milk	45%	45%	30%	11%	35%	87%	41%
Wine (0-19% alcohol)	4%	4%	6%	3%	9%	0%	5%
Grand Total	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

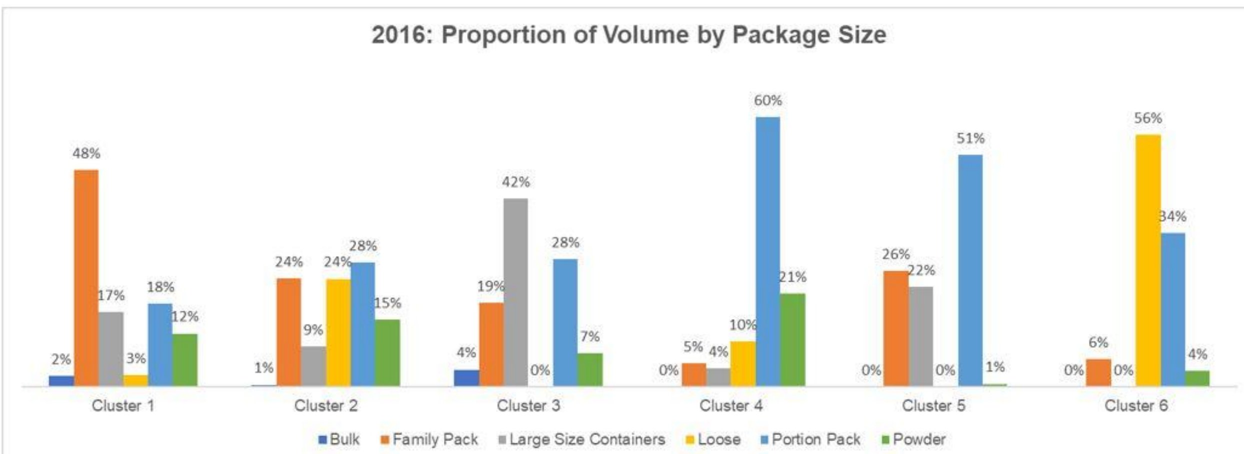
G. Table of Country and Cluster by Year (2015 - 2019)

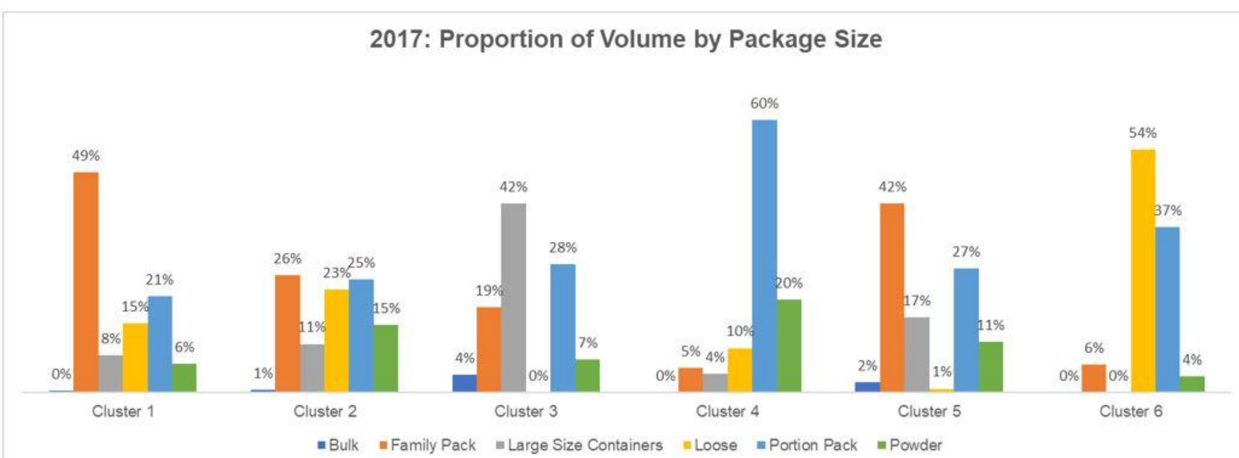
Cluster	2015	2016	2017	2018	2019
1	Thailand	Brazil, Mexico, Russian Federation, United Kingdom of Great Britain and Northern Ireland, Germany, France, Spain,	Russian Federation, Turkey, Iran (Islamic Republic of), Iraq	Russian Federation, Turkey, Iran (Islamic Republic of)	Russian Federation, Iran (Islamic Republic of)

		South Africa, Italy, Canada, Poland			
2	Pakistan, Brazil, Russian Federation, Mexico, United Kingdom of Great Britain and Northern Ireland, Germany, France, Indonesia, Spain, Italy, South Africa, Canada, Poland, Turkey, Colombia, Viet Nam, Australia, Philippines, Iran (Islamic Republic of), Venezuela (Bolivarian Republic of), Nigeria,....	Pakistan, Indonesia, Thailand, Turkey, Colombia, Viet Nam, Philippines, Australia, Saudi Arabia, Iran (Islamic Republic of), Korea, Algeria, Ethiopia, Nigeria, Netherlands, Malaysia, Iraq, Kazakhstan,...	Pakistan, Indonesia, Argentina, South Africa, Canada, Poland, Viet Nam, Colombia, Australia, Philippines, Saudi Arabia, Korea, Algeria, Ethiopia,...	Pakistan, Indonesia, Argentina, South Africa, Canada, Poland, Viet Nam, Colombia, Australia, Philippines, Saudi Arabia, Korea, Algeria, Ethiopia,...	Pakistan, Turkey, Indonesia, Argentina, South Africa, Canada, Poland, Viet Nam, Colombia, Australia, Philippines, Saudi Arabia, Korea, Algeria, Ethiopia,...
3	United States	United States	United States	United States	United States
4	China	China	China	China	China
5	Japan	Japan	Japan, Brazil, Mexico, United Kingdom of Great Britain and Northern Ireland, Germany, France, Spain, Italy, Thailand, Netherlands	Japan, Brazil, Mexico, United Kingdom of Great Britain and Northern Ireland, Germany, France, South Africa, Argentina, Spain, Canada, Italy, Poland, Netherlands	Japan, Brazil, Mexico, United Kingdom of Great Britain and Northern Ireland, Germany, France, South Africa, Argentina, Spain, Canada, Italy, Poland, Netherlands
6	India	India	India	India	India

H. Proportion of Volume by Package Size (2015 - 2017)







I. Proportion of Volume by Packed Product (2015 - 2017)