An Implementation of Stock Allocation Plan of Retail Store Based on Cluster Analysis

Mr. Punnakrit Natekijcharoen

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Engineering Management
(CU-Warwick)
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2019

การประยุกต์วิธีการวิเคราะห์กลุ่มเพื่อนำมาปรับปรุงการวางแผนการกระจายสินค้าสำหรับร้านค้าปลีก

นายปัณกฤด เนตรกิจเจริญ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาการจัดการทางวิศวกรรม ศูนย์ระดับภูมิภาคทางวิศวกรรมระบบการผลิต
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562

Thesis Title: An Implementation of Stock Allocation Plan of Retail Store Based on Cluster Analysis

By: Mr. Punnakrit Natekijcharoen

Field of Study: Engineering Management

Thesis Advisor: Assistant Professor Pisit Jarumaneeroj, Ph.D.

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

.... Dean of the FACULTY OF ENGINEERING

(Professor SUPOT TEACHAVORASINSKUN, Ph.D.)

THESIS COMMITTEE

.... Chairman

(Professor PARAMES CHUTIMA, Ph.D.)

.... Thesis Advisor

(Assistant Professor Pisit Jarumaneeroj, Ph.D.)

.... Examiner

(Assistant Professor ARISARA JIAMSANGUANWONG, Ph.D.)

.... External Examiner

(Associate Professor Chuvej Chansa-ngavej, Ph.D.)

ปัณกฤด เนตรกิจเจริญ : การประยุกต์วิธีการวิเคราะห์กลุ่มเพื่อนำมาปรับปรุงการวางแผนการกระจายสินค้าสำหรับร้านค้าปลีก. (An Implementation of Stock Allocation Plan of Retail Store Based on Cluster Analysis) อ.ที่ปรึกษาหลัก : ผศ. ดร.พิศิษฎ์ จารุมณีโรจน์

งานวิจัยนี้มุ่งเน้นการประยุกต์ใช้วิธีการวิเคราะห์กลุ่มในการวางแผนการกระจายสินค้าสำหรับกลุ่มร้านค้าปลีกของบริษัทกระจายสินค้าแห่งหนึ่ง ซึ่งกิจกรรมดังกล่าวถือเป็นหนึ่งในกิจกรรมหลักที่สำคัญของบริษัท โดยเฉพาะอย่างยิ่งในช่วงวันหยุดเทศกาลที่มียอดการขายสูง ทั้งนี้ การจัดกลุ่มร้านค้าเพื่อการกระจายสินค้าที่ไม่เหมาะสมอาจส่งผลให้เกิดการสูญเสียโอกาสในการขาย และอาจส่งผลต่อการจัดเก็บสินค้าที่มีมากเกินความต้องการในบางกลุ่มร้านค้าปลีก จากการศึกษาเบื้องต้น ผู้วิจัยพบว่า วิธีการวิเคราะห์กลุ่มที่บริษัทกรณีศึกษาใช้อยู่ในปัจจุบันยังขาดประสิทธิภาพอยู่มาก โดยวิธีการดังกล่าวเป็นเพียงการแบ่งจำนวนร้านค้าปลีกออกเป็น 4 กลุ่มย่อยๆ ที่มีขนาดเท่าๆ กัน จากค่าประสิทธิภาพทางการขายเพียงเท่านั้น นอกจากนี้ ค่าสัมประสิทธิ์การแปรผันของสินค้าที่กระจายและจัดเก็บในแต่ละกลุ่มยังมีค่าที่ค่อนข้างสูง กล่าวคือ มีค่าระหว่าง 30.1% ถึง 51.1%

ผู้วิจัยได้นำเสนอแนวคิดในการวิเคราะห์กลุ่มใหม่สองวิธีแก่บริษัทกรณีศึกษา อันประกอบไปด้วยวิธีการวิเคราะห์กลุ่มแบบ K-Means และ วิธีการวิเคราะห์กลุ่มแบบ Agglomerative เพื่อปรับปรุงประสิทธิภาพของการวิเคราะห์กลุ่มและการกระจายสินค้าของบริษัทให้ดียิ่งขึ้น ผู้วิจัยพบว่า การจัดกลุ่มร้านค้าปลีกด้วยวิธีการวิเคราะห์กลุ่มแบบ K-Means และวิธีการวิเคราะห์กลุ่มแบบ Agglomerative นี้ ให้ผลลัพธ์ที่ดีกว่าวิธีการวิเคราะห์กลุ่มที่บริษัทกรณีศึกษาใช้อยู่ในปัจจุบัน โดยค่าสัมประสิทธิ์การแปรผันของสินค้าที่กระจายและจัดเก็บในแต่ละกลุ่มร้านค้ามีลดลงเหลือ 9.5% ถึง 9.3% ในขณะที่ค่าผลต่างรวมระหว่างค่าการกระจายสินค้าระดับกลุ่มร้านค้าและค่าการกระจายสินค้าระดับร้านค้าก็มีค่าลดลง โดยมีค่าลดลงจาก 17,818,056 ชิ้น เหลือเพียง 15,672,717 ชิ้นสำหรับวิธีการวิเคราะห์กลุ่มแบบ K-Means และลดลงจาก 17,818,056 ชิ้น เหลือเพียง 15,830,644 ชิ้นสำหรับวิธีการวิเคราะห์กลุ่มแบบ Agglomerative ตามลำดับ นอกจากนี้หากทำการเปรียบเทียบผลลัพธ์ระหว่างสองวิธีการวิเคราะห์กลุ่มใหม่ที่นำเสนอ ผู้วิจัยพบว่า วิธีการวิเคราะห์กลุ่มแบบ K-Means ให้ผลลัพธ์ที่ดีกว่า ทั้งในมุมของค่าสัมประสิทธิ์การแปรผัน และค่าผลต่างรวม ระหว่างค่าการกระจายสินค้าระดับกลุ่มร้านค้าและค่าการกระจายสินค้าระดับร้านค้า

สาขาวิชา	การจัดการทางวิศวกรรม	ลายมือชื่อนิสิต	
ปีการศึกษา	2562	ลายมือชื่อ อ.ที่ปรึกษาหลัก	

6071216621 : MAJOR ENGINEERING MANAGEMENT
KEYWORD: Abstract_Eng

Punnakrit Natekijcharoen : An Implementation of Stock Allocation Plan of Retail Store Based on Cluster Analysis. Advisor: Asst. Prof. Pisit Jarumaneeroj, Ph.D.

This thesis presents an implementation of cluster analysis on retail store clustering so that better cluster-based stock allocation plans could be effectively devised and efficiently executed. For the case study company, stock allocation is one of its key strategic decisions as improper stock allocation, especially during special occurrences with high sale volumes, may lead to loss of sales – and so overstocking at some stores/clusters. Based on our initial investigations, we find that the current clustering technique is somewhat inefficient as it simply divides the stores into four groups with equal members based on store's sales performance. Besides, the coefficient of variation of allocated stocks in each cluster is comparatively high, around 30.1% – 51.1%

To better improve the efficiency of current clustering operation, two more systematic clustering techniques have been therefore introduced and compared with the current technique, namely K-Means and Agglomerative clustering techniques. We find that both K-Means and Agglomerative clustering techniques provide clusters with much less coefficients of variations, about 9.5% and 9.3% respectively. Besides, the total differences between allocated stock target by store cluster and actual stock target by store are also improved from 17,818,056 units to 15,672,717 units and from 17,818,056 units to 15,830,644 units by these two techniques, respectively. When compared among these two new approaches, it can be seen that K-Means clustering technique outperforms Agglomerative clustering technique in terms of both coefficient of variation and total difference between allocated stock target by store cluster and actual stock target by store.

Field of Study:	Engineering Management	Student's Signature ....
Academic Year:	2019	Advisor's Signature ....

ACKNOWLEDGEMENTS

First of all, I would like to express my deep and sincere gratitude to my thesis advisor, Asst. Prof. Dr. Pisit Jarumaneeroj of the Chula System Engineering at Chulalongkorn University, for supporting and providing guidance and recommendation throughout this research. Moreover, I feel very grateful to the organisation and my colleagues that understand and support me during my study at the graduate school. Last but not least, I would like to thank you my parent and family who always supporting and taking care of me both financially and mentally throughout my years of study in master's degree.

Punnakrit Natekijcharoen

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The implementation of data science has been becoming a significant tool for business to help achieve higher competitive advantage over their competitors. Lots of businesses are recognised the importance of value that is derived from the data and have been implementing the use of data science to identify their business's opportunities and elevate the sustainability of their business performance. Moreover, data mining which is one of the data science field that specifically study on the historical data where it is the process of analysing the data and identify the pattern of the data in order to identify and highlight the relationship and characteristic which it will turn into a meaningful information through predictive analysis for business to further develop their business plan and business strategy.

This research will exemplify how the data science and data mining can be applied and implemented to the retailer service provider. The study of cluster analysis will be outlined and implemented in this research with the objective to improve the store classification based on the store performance variables which will result in greater project management and greater efficiency of the stock allocation plan for the organisation during holiday event or seasonal event such as New Year festival and etc.

1.1 Background

1.1.1 Company background

The selected company of this research is one of the leading retailing service providers in Thailand, namely Tesco Lotus, operated by Ek-Chai distribution system Co., LTD, which originally founded in 1994. Tesco Lotus has three different store formats including hypermarket format, supermarket format and convenient store format as shown in Figure 1-1 to Figure 1-3 before introducing a Tesco online shopping platform in 2013. Tesco Lotus offers wide range of the product category such as grocery goods, household products, fresh food, apparel, electrical appliances, and etc. Currently, there are about 2,100 stores of Tesco Lotus's store in total across the country including about 200 stores of hypermarket format, about 200 stores of supermarket format and about 1,700 stores of convenient store format. On the other hand, Tesco Lotus also has their own distribution centres to supply the stock from suppliers before delivering the stock to the store. The distribution centres of Tesco lotus consist of ten distribution centres from four areas across Thailand including central region, north region, northeast region and south region where three distribution centres are for fresh food and frozen product which are located in central, northeast and south region of Thailand and another seven distribution centres are for ambient product which are located across four regions of Thailand.

Figure 1-1: Tesco Lotus Hypermarket Store

Figure 1-2: Tesco Lotus Supermarket Store

Figure 1-3: Tesco Lotus Convenient Store

Image source: http://www.tescolotusthailand.com

1.2 Problem statement

As a project manager of ordering system team under supply chain department of Ek-Chai distribution system company or Tesco Lotus who take the responsibility for order and stock planning during the holiday event and seasonal event such as New Year festival, Songkran festival (Thai New Year festival) and etc. The owner of this project will be leading all related party in supply chain team in order to develop and manage the stock and ordering plan together. Moreover, project manager also coordinates with the external team such as store operation team and distribution centre (DC) operation team in order to feedback the developed stock allocation plan so that other team can manage and prepare their resource and capacity according to the developed plan. Furthermore, the overall objective and scope of working of this project is to ensure the stock availability at stores whereas holding an optimal level of stock so that supply chain team can achieve the target of the stock budget during the holiday event or seasonal period. What is more, the planning will be consisting of both inbound side (stock delivery from suppliers to distribution centres) and outbound side (stock delivery from distribution centres to store).

On the other hand, there are a number of limitations and constraints during the seasonal or holiday event. The first limitation is the issue of limited resource and capacity of the Tesco Lotus's distribution centres where the ability of distribution centres to receive, pick and deliver the stock cannot cope with the peak of demand during holiday or seasonal event. Therefore, the stock allocation plan is helping distribution centres in terms of allocating and phasing the inbound and outbound volume before the event started. On the other hand, vendor closure during the holiday event or seasonal period is also another constraint or obstacle for the organisation. Because the majority of the vendors cannot support to deliver their product to the Tesco Lotus's distribution centres during the holiday or seasonal event. Therefore, these vendors will be delivering their product or stock prior to the distribution centres of Tesco Lotus which it will make the traffic at the distribution centres jam during the holiday or seasonal period. Thus, to overcome these limitations and constraints. There are a number of preparation and tasks need to be planned and prepared by project manager. To illustrate, the researcher will be using the New Year festival period as an example or case study in this case. As can be seen from Figure 1-4, where two lines graph describe the peak of sales volume during week 44 and week 45 which is a New Year period between the year of 2018 and 2019, according to the Tesco Lotus's calendar.

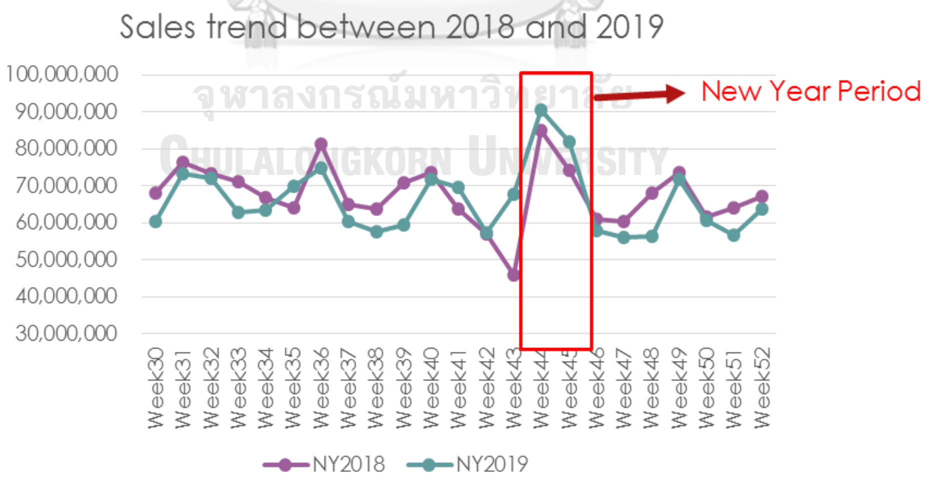

Figure 1-4: Trend of sales volume during 2018 and 2019

Therefore, Tesco Lotus has to deal with higher volume during the event. Project manager has to manage and develop the stock allocation plan to align with the acceptable capability and capacity of the company. What is more, the process to develop the stock allocation plan can be divided into three stage including key item selection, store cluster analysis and estimate the stock target by store cluster level. Thus, the first step to develop the stock allocation plan is to identify a group of key items that contribute the most sales performance during the New Year seasonal period before performing the store cluster analysis in order to group the store that has a similar sales performance together. Then, the stock target of selected key items by store cluster level will be estimated. At the end, the stock allocation plan for an outbound side which is a plan between distribution centres and stores is developed and completed.

Moreover, the reason to develop a stock allocation plan by store cluster level is because there are lots of combination between store and items level. To illustrate, given that there are 100 items and 200 stores needed to develop the stock allocation by store level. In this case, the combination between item and store is equal to 200,000 relations (100 items multiply with 200 stores). On the other hand, if consider that there is four store clusters and 100 items need to developed stock allocation plan by store cluster level. The total relation will be equal to 400 relations (4 store clusters multiply with 100 items) which is a significantly less than the plan by store and item level. However, the approach and technique of store cluster analysis in the past was done by using the store cluster feedbacked by store operation team to develop the stock allocation plan which it did not reflect the actual sales performance of the selected New Year key items and the stores. While, the implementation of store cluster analysis was introduced in later year but still lack off solid clustering technique which lead to an inefficiency of the stock allocation plan. To illustrate, the current approach for store cluster analysis is done by ranking the sales performance during the New Year festival event by store level. Then, split the store equally according to the specified number of store cluster which normally will be divided around three to four store clusters. As can be seen from the Table 1-1 where there are 196 stores of hypermarket store. Then, from total 196 stores are split into 49 stores for each store cluster from cluster A to cluster D. Moreover, Table 1-1 also shows the coefficient of variation by store cluster. In this case, it can be seen that there are high level of coefficient of variation on cluster A and D compared to cluster B and C which it means that cluster A and cluster D have higher level of sales performance disperse around the mean or average of sales performance within the store cluster.

Store format	Cluster	Count of Store	Average of New Year 2019 Sales Volume	SD of New Year 2019 Sales Volume	Coefficient of Variation
Hypermarket	Cluster A	49	369,505	65,247	18%
	Cluster B	49	244,187	20,446	8%
	Cluster C	49	189,111	12,565	7%
	Cluster D	49	135,280	23,377	17%
		196	**234,521**	**94,346**	**40%**

Table 1-1: An example of store cluster from current organisation's clustering technique

As a result of the current store clustering technique, it can lead to an inefficiency of stock allocation plan when using this store clustering technique to allocate the stock target stock by store cluster level. Therefore, this research is intended to study the cluster analysis and identify the suitable clustering technique to improve the performance of store cluster analysis and stock allocation plan.

1.3 Research objective

The project is expected to achieve the main objectives which can be summarised as below points.

1. Identify and recommend an appropriate and efficient store clustering technique to use for store cluster analysis in the future based on the performance evaluation of each clustering technique.
2. Improve the efficiency of the store cluster analysis and improve the efficiency when develop the stock allocation plan for any holiday event or seasonal period such as New Year festival.

1.4 Research question

According to the problem statement and objective of the research mentioned in previous part lead to research question as below:

1. Can generated data and variables be applicable for the store cluster analysis?
2. Which clustering technique give the most efficient result when performing the store cluster analysis and develop the stock allocation plan?
3. Can the implementation of the recommended clustering technique improve the efficiency of the store cluster analysis and the stock allocation plan?
4. Is recommended clustering technique give better results than the current clustering technique that company have been using?

1.5 Hypothesis development

From the research question and research objective, the hypothesis statement can be developed by using the existing knowledge and related research as follows.

1. The improvement of the store cluster analysis can be achieved through the implementation of a recommended clustering technique.
2. The improvement of the stock allocation plan for New Year Seasonal event can be achieved through the implementation of a recommended clustering technique.

1.6 Scope of the research

The research is focused on the study of cluster analysis in order to identify the most efficient clustering technique to implement and apply during the process of store cluster analysis when developing the stock allocation plan. Furthermore, the introduced clustering techniques will be reviewed and selected to perform the store cluster analysis in statistical program, namely SPSS program, to classify and group the store that has similar sales performance together during New Year Seasonal event before estimating the stock target by store cluster level. At the end, each clustering techniques together with the current clustering technique that company have been using will be evaluated and compared the performance of store cluster analysis in order to highlight the result and identify the most efficient clustering technique to use for the store cluster analysis task in the future. In this case, two performance measurement will be used to evaluate the performance of store cluster analysis. The first measurement is the coefficient of variation which is a measure of data dispersion around the mean. Therefore, it is expected that the implementation of two introduced clustering techniques will give a significant improvement on the coefficient of variation when compared with the current clustering technique that company have been using. The lower coefficient of variation, the better efficiency of the store cluster analysis. It means that within the store cluster, all the stores in that store cluster have similar store performance grouped together. Thus, the efficiency of the stock allocation plan will be improved when estimate the stock target by store cluster level. Furthermore, the second measurement is to develop the stock allocation plan by using the result of store cluster analysis from each clustering technique. Then, compare and determine which clustering technique give the least total difference between allocated stock target by store cluster and actual stock target by store. In this case, it means that the clustering technique that give the least total difference will be the most efficient and appropriate clustering technique to use when develop the stock allocation plan.

1.7 Outcome

The outcome of this research can be summarised as shown below.

1. The most efficient clustering technique for store cluster analysis is identified and recommended through the performance evaluation of store cluster analysis.
2. The efficiency of the store cluster analysis is improved through the recommended clustering technique.
3. The efficiency of the stock allocation plan is improved through the implementation of the recommended clustering technique when perform the store cluster analysis.

2. LITERATURE REVIEW

In this chapter will be outlined the concepts and theories of the cluster analysis together with the different types of clustering technique which will be used to study for the implementation of cluster analysis of the retail stores throughout the research. Moreover, the related case study and implementations that associated with cluster analysis will be discussed in this chapter.

2.1 Cluster analysis

Cluster analysis is one of the data mining method that aim to identify the hidden structure of the data set by grouping the similar data that has the same characteristic together from a set of data points into a number of clusters. According to Scoltock (1982), the benefit of grouping or classifying the data set allows business to reveal and distinguish the pattern which might give a significant and meaningful interpretation of the data and help business to suggest a better decision making together with develop a better business strategy and business plan through this analysis technique.

2.2 Type of clustering technique

There are several clustering techniques. According to Charu and Chandan (2014), had outlined a number of clustering techniques such as partitioning clustering, agglomerative clustering, probabilistic clustering, density-based clustering and etc. Furthermore, Charu and Chandan (2014), also mentioned that there are different types of data such as numerical data, categorical data, time series data and etc. Since the organisation's data that this research will be used to perform the store cluster analysis is solely numerical data. The most commonly use and recommended clustering technique that suit for numerical data is distance-based algorithms. Moreover, the distance-based algorithms are focusing on minimising the Euclidean distance of the data set. Therefore, it is suggested that the clustering technique of this research will be applied the distance-based algorithms which it suits with the data type that will be generated from the organisation's database.

Overall, there are two types of clustering technique of distance-based algorithms including flat or partitioning method and hierarchical method which can be summarise as below.

2.2.1 Partitioning method

The first clustering technique namely partitioning method or can be called as flat method. It is the clustering technique that classify each data point based on the similarity. This technique will be clustered the data points into K group according to the number of cluster that analyst is specified. According to Soni (2012), there are several techniques of partitioning clustering algorithms such as K-Means, K-Medians, K-Medoids and etc where the most commonly use technique is the K-Means clustering technique. In this case, the K-means clustering technique is selected to review and perform the store cluster analysis.

K-means clustering technique

K-means clustering technique is one of the partitioning clustering methods. According to Celebi Emre and Hassan (2015) had define the concept of K-Means clustering technique where the objective is to combine the data points based on the minimum distance between the data points and centroids before grouping into the desired K, or the desired number of cluster. Referred from Fahim and et al. (2006), K-Means clustering technique is one of the most efficient clustering techniques that widely used and proposed in literature. On the other hand, there are some factors affecting the performance of K-Means clustering method. Raval and Chaita (2016), had also identified two factors that affecting the K-Means clustering performance. Firstly, it needs to identify the number of cluster or K. The second factor, it requires the analyst to identify the initial centroids. However, there are several methods proposed to overcome each factor. The Elbow method and Silhouette Coefficient is introduced in order to identify the optimal number of cluster or K. On the other hand, the technique of K-Means++ is proposed in order to identify the initial centroids or using the statistical program to perform the cluster analysis such as SPSS program which it will automatically select the initial centroid for the analyst.

Furthermore, according to Andreopoulou and et al. (2017), had described the step to perform the K-Means clustering algorithm. Overall, there are five steps of K-Means clustering algorithm which can be summarised as below.

Step 1: Identify the target K or the number of cluster.

Step 2: Select the randomly initial centroids as an initial cluster centre which it has to align with the number of cluster or K that analyst is selected. For example, if the analyst would like to classify the data points into 2 clusters. Then, two initial centroids are needed to select in this case.

Step 3: Calculate the distance between each data point and each centroid. Then, assign that data point to the nearest cluster centre or the nearest centroid. The formula of Euclidean distance is presented as below equation 2-1

$$Euclidean\ distance = \sqrt{\sum_{k=1}^{n}(p_k - q_k)^2} \qquad (2\text{-}1)$$

Step 4: Assign the new centroid as the average of the data points of each centroid.

Step 5: Repeat the step 2 to 4 until the cluster centroid is no longer change.

Moreover, the sample of data set is given to illustrate how the algorithm of K-Means clustering technique is working. Given that the example of data set has five data points and it would like to classify these five data points into two clusters. The illustration of how K-Means clustering technique is classifying the given five data points is outlined as below.

1. Given that there are five data points from this data set which are {5, 8, 3, 2, and 10}
2. K or the number of cluster is identified which is 2 clusters
3. The two initial centroids are randomly selected which are (4, and 8)
4. Calculate the Euclidean distance between these five data points with each centroid as formula presented in step 3 above. The first iteration of K-Means clustering is presented in Table 2-1 where the calculated Euclidean distance of each point is updated in the table. Then, the recommended cluster of each data point is selected based on the minimum distance between two centroids as presented on the last column from Table 2-1.

1st iteration		Centroids		Cluster
		4	8	
Data points	5	1	3	1
	8	4	0	2
	3	1	5	1
	2	2	6	1
	10	6	2	2

Table 2-1: The first iteration of the example of K-Means clustering calculation

5. On the second iteration, the new centroids are updated as presented in Table 2-2 including 3.3 and 9 which it calculated by the average of each data point from the assigned cluster of the first iteration. Furthermore, the Euclidean distance is re-calculated according to the new calculated centroids. At the end, the recommended cluster of each data point is suggested according to the minimum distance between these two centroids as shown in Table 2-2 below.

2nd iteration		Centroids		Cluster
		3.3	9	
Data points	5	1.7	4	1
	8	4.7	1	2
	3	0.3	6	1
	2	1.3	7	1
	10	6.7	1	2

Table 2-2: The second iteration of the example of K-Means clustering calculation

6. For the third iteration as presented in Table 2-3. Since the assigned cluster of each data point is the same when compared with the first and second iteration, it can be seen that the cluster centroids are no longer change which it means that the calculation and classification can be ended at this process. At the end, the final cluster of each data point can be referred from the previous iteration.

2nd iteration		Centroids		Cluster
		3.3	9	
Data points	5	1.7	4	1
	8	4.7	1	2
	3	0.3	6	1
	2	1.3	7	1
	10	6.7	1	2

3nd iteration	Centroids		Cluster
	3.3	9	

Table 2-3: The second and third iteration of the example of K-Means clustering calculation

Therefore, the result of cluster analysis computed by K-Means clustering algorithm can be summarised as follow. Cluster number one which it contains three data points including 5, 3, and 2. On the other hand, cluster number two contains two data points including 8 and 10.

2.2.2 Hierarchical method

There are two types of hierarchical method including agglomerative clustering technique and divisive clustering technique. According to Murtagh and Contreras (2011), had described the difference between these two clustering techniques where the algorithm of agglomerative clustering technique starts with every case or each data point being an individual cluster before merging each data point into a larger group until getting one final cluster. This clustering technique can be considered as a bottom-up approach. On the other hand, the algorithm of divisive clustering technique starts with all the data points are merged into one single cluster before breaking into the smaller group until all the data points are clustered as an individual cluster, a top-down approach. In this case, the agglomerative clustering technique is used to perform the store cluster analysis on this research.

Agglomerative clustering technique

This clustering technique is considered as one of the hierarchical methods. According to Nielson (2016), agglomerative clustering technique is a bottom-up approach where it starts to cluster the individual data point separately before aggregating all the data points into one final cluster based on the distance between each data point. It means that at the beginning, the number of cluster will be equal to the number of data point until it grouped into one final cluster at the end. The result of this clustering technique normally presented in a form of cluster hierarchy, tree-based structure, or can be called as "Dendrogram" as an example shown in Figure 2-1.

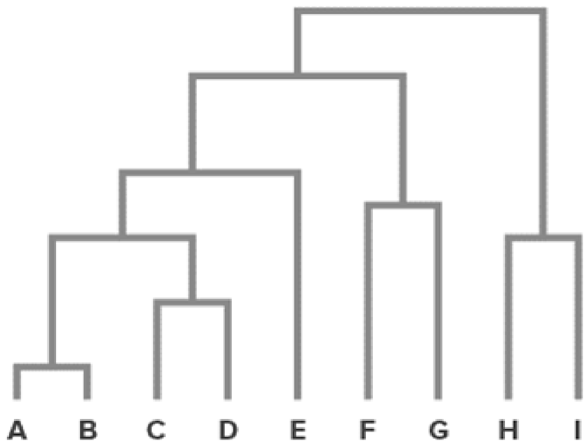

Figure 2-1: An example of Dendrogram, tree-based structure

According to Sławomir and Mieczysław (2018), had described and concluded the step of performing the cluster analysis by using agglomerative clustering technique. In this case, the algorithm of agglomerative cluster analysis can be divided into four steps which can be summarised as below.

Step 1: Start by preparing the data and consider each data point as its own individual cluster

Step 2: Compute the distance between each cluster

Step 3: The most similar or closest distance between two clusters are grouped together into a new bigger cluster in a form of hierarchical cluster tree.

Step 4: Repeat the step 2-3 until the cluster are merged into one single cluster.

Furthermore, the sample of data set is given in order to illustrate how the algorithm of agglomerative clustering technique is working. In this case, the single linkage method of agglomerative clustering method is used which select the minimum distance between each pair or it can be said that this technique is select the closest distance between each the data point. Given that the example of data set is as same as the example of K-Means clustering method which including five data points from the data set.

1. Five data points of this data set which includes {5, 8, 3, 2, and 10}
2. From these five data points will be named as {a, b, c, d, and e}
3. At the initial stage, the number of cluster is equal to the number of data points which is 5 clusters.
4. The distance between each data points is calculated and updated in the matrix as Table 2-4 shown below. In this step, the minimum distance between two data points will be selected to combine into another cluster. Therefore, data points between c and d will be joined into one bigger cluster.

1st iteration			Data points				
			a	b	c	d	e
			5	8	3	2	10
Data points	a	5	0	3	2	3	5
	b	8	3	0	5	6	2
	c	3	2	5	0	1	7
	d	2	3	6	1	0	8
	e	10	5	2	7	8	0

Table 2-4: The distance between each data points from the first iteration matrix

5. In the second iteration, after update the data points as shown in Table 2-5. Moreover, the distance between each point are re-calculated as presented in below table. Then, select the minimum distance between each data point again. This round, the data points between [a, (c,d)] together with [b,e] will be merged into one bigger cluster.

2nd iteration			Data points			
			a	b	c,d	e
			5	8	-	10
Data points	a	5	0	3	2	3
	b	8	3	0	5	2
	c,d	-	2	5	0	7
	e	10	5	2	7	0

Table 2-5: The distance between each data points from the second iteration matrix

6. The last iteration shown no data point left to merge as presented in Table 2-6. Therefore, the process of clustering by agglomerative method is ended in this stage. At the end, the agglomerative clustering technique grouped all the data points into one single cluster.

3rd iteration			Data points	
			a,(c,d)	b,e
			-	-
Data pc	a,(c,d)	-	0	2
	b,e	-	2	0

Table 2-6: The distance between each data points from the last iteration matrix

7. In this case, the result of this cluster analysis by using agglomerative clustering technique can be constructed as the Dendrogram, step by step as below Figure 2-2,

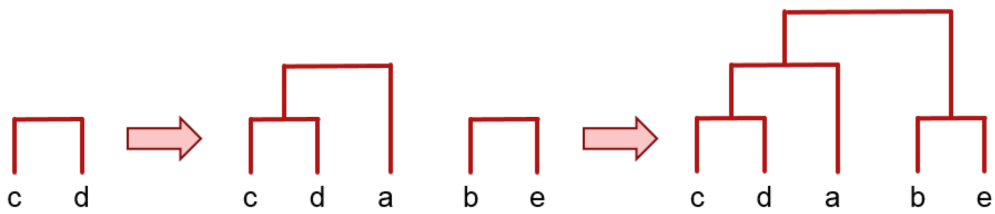

Figure 2-2: The Dendrogram from example dataset

At the end, if consider these five data points to group into two clusters as same as the previous example of K-Means clustering technique. Therefore, the result of the cluster analysis by using Agglomerative clustering technique can be summarised as follow. Cluster number one will be contained three data points including c, d and a or 3, 2, and 5. On the other hand, cluster number two will be contained two data points including b and e or 8 and 10.

2.3 Existing of case study and implementation that related to cluster analysis

According to Marr (2015), big data analytics is a game changer in the retail industry where it is being applied and implemented at every stage of the business plan such as identifying and predicting trends, forecasting the demand which can be expected to achieve higher competitive advantage over their competitors.

What is more, the technique of cluster analysis has been widely used and implemented to many businesses when analysing the data and identifying or predicting the pattern. There are a number of implementation and application of the cluster analysis. According to Frei (2013), had studied and implemented the application of cluster analysis in medication industry with the objective to identify a group of obese patients that lose weight after the bariatric surgery in order to determine the root cause of patient's weight loss based on medical statistic variables. On the other hand, Dardac and Boitan (2009), had applied the concept of cluster analysis to group the credit institutions that have the similar in terms of their risk profile and profitability to classify the performance of credit institutions. Moreover, there was an application of K-Means cluster algorithm applied by Oyelade, Oladipupo, and Obagbuwa (2010), to classify the students' academic performance of a private school in Nigeria with the objective to analyse and classify the students' performance in order to monitor the development of student's performance and develop an efficient academic planning for these students.

2.4 Summary of literature review

In conclusion, cluster analysis is the technique to identify the hidden structure of the data set by grouping the data based on the similarity. Moreover, the objective of clustering technique of distance-based algorithms is to minimise the distance between each data point and group the closest or nearest distance between data points together. Overall, the clustering technique of distance-based algorithms can be divided into two types including the partitioning method and hierarchical method namely K-Means clustering technique and Agglomerative clustering techniques respectively. Furthermore, there were several implementation and application of cluster analysis such as the use and implementation in the medications industry, credit constitutions and education institution as mentioned in previous session. Therefore, it can be seen that the use of cluster analysis can be a useful tool and technique to help business interpret those data to develop their business strategy and business plan. Lastly, after the literature review part, the following chapters will be using these two clustering techniques including Agglomerative clustering technique and K-Means clustering technique to implement and apply when performing the store cluster analysis which will be performed in the statistical program namely SPSS program. At the end, the result of each clustering technique will be outlined, evaluated, compared and suggested for the future work of store cluster analysis task in the future.

3. RESEARCH METHODOLOGY

In this chapter will be outlined and summarised the approach of how the research will be conducted in order to achieve the research objectives. After literature review stage where the concept of cluster analysis together with the algorithm of different clustering techniques are reviewed and outlined, the researcher will be using these reviewed clustering techniques including K-Means clustering technique and Agglomerative clustering technique to perform the store cluster analysis in the analysis chapter. Moreover, the performance measurement will be outlined and discussed together with evaluated the result of each clustering technique in the fifth chapter. At the end, the conclusion of the research will be outlined in order to highlight and summarise the research objective, finding and result of the research.

Overall, the research activities are consisting of five main processes which can be summarised as below flow chart.

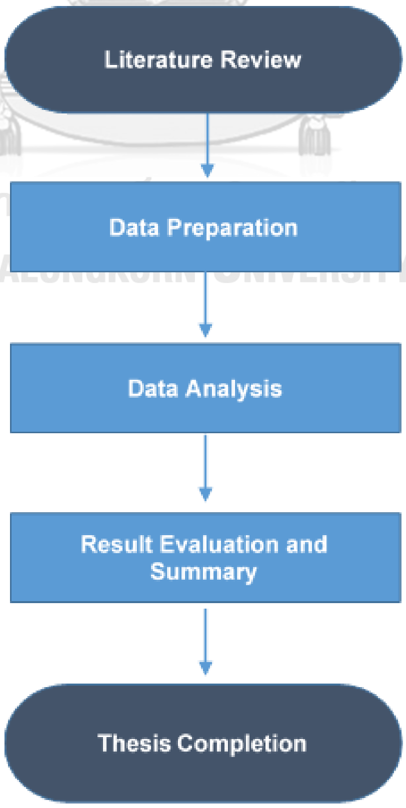

Figure 3-1: Research Activities Flow Chart

To illustrate, the first step begins with literature review of cluster analysis. The next procedure after literature review stage followed by the data preparation process. After generating and collecting all the data that needed for the store cluster analysis. Then, the analysis of store cluster will be performed in the following process. Next stage after conducting the store cluster analysis is to evaluate the result of each clustering technique and compare the result. At the end, the recommendation and summarisation of the result from the store cluster analysis will be outlined and summarised before completing the thesis book at the final step.

3.1 Data preparation

In this stage, all the data are prepared before performing the store cluster analysis. Overall, there are two steps for data preparation including data generation and data normalisation. Furthermore, the first step of the data preparation is to generate the data from the organisation's database. What is more, the generated data will be collected by store level which includes a set of data to use to perform the store cluster analysis as follows.

1. Sales Volume by store during New Year 2018 (Overall total sales)
2. Sales Volume by store during New Year 2019 (Overall total sales)
3. Sales Volume by store during New Year 2018 (of Selected Key Items)
4. Sales Volume by store during New Year 2019 (of Selected Key Items)
5. %Sales Uplift of New Year 2018 by store (Overall total sales)
6. %Sales Uplift of New Year 2019 by store (Overall total sales)
7. %Sales Uplift of New Year 2018 by store (of Selected Key Items)
8. %Sales Uplift of New Year 2019 by store (of Selected Key Items)

The generated data are consisting of eight variables in total which are the variables that company have been using to perform the store cluster analysis in the past. Therefore, these eights variables will be used when performing the store cluster analysis in statistical program namely SPSS program. What is more, the first four variables are the total sales volume during New Year period between 2018 and 2019 by store level which split by the overall total sales volume and the total sales volume of selected key items. In this case the New Year period is between two weeks including week 44 and week 45 according to week number of company's calendar. What is more, the last four variables are the percentage of sales uplift between New Year period and normal period by store level which split by the overall total sales uplift and the total sales uplift of selected key items. Moreover, the percentage of sales uplift came from the sales uplift between New Year period 2018 and normal period of 2017 together with the percentage of sales uplift between New Year period of 2019 and normal period of 2018. Where the percentage of sales uplift can be calculated as below equation 3-1.

$$\%\text{Sales uplift of year X of store number Y} = \frac{((\mathrm{A} + B) \div 2)}{\mathrm{C}} - 1 \qquad (3\text{-}1)$$

Where

A = Weekly sales of New Year period of week 44

B = Weekly sales of New Year period of week 45

C = An average of weekly sales during normal period

At the end, these eight variables will be generated and performed the store cluster analysis as an example shown in Table 3-1. Even though, it can be seen that these eight variables are presented in numerical data. However, there is a difference in the scale of data between these eight variables. According to Eessaar (2016), to perform the data analysis, it is recommended that all the data must be normalise in order to eliminate data inconsistency and reduce the redundancy of the data set. Moreover, there are a number of techniques for normalising the data such as Z-score technique, the data conversion into the range between 0 and 1.

In this research will be selected the Z-score technique as a method to normalise all the eight variables that generated from the organisation's database. The data normalisation by using the Z-score technique can be performed by using the below equation 3-2.

$$Z = (x - \mu) \div \sigma \qquad (3\text{-}2)$$

Where

Z = Z - score

x = individual value in the data set

μ = mean of all values in the data set

σ = the standard deviation of a population

Thus, the calculation of Z-score will be performed in order to normalise the data and convert all eight variables into the same scale. Finally, the process of data preparation is done and ready for the next stage which is store cluster analysis.

Format	Store No.	Sales Volume (Overall) 2018	Sales Volume (Overall) 2019	Sales Volume (Key item) 2018	Sales Volume (Key item) 2019	%Sales Uplift (Overall) 2018	%Sales Uplift (Overall) 2019	%Sales Uplift (Key item) 2018	%Sales Uplift (Key item) 2019
Hyper	**HPX0001**	486,295	534,666	349,337	427,682	41.6%	55.7%	39.4%	70.6%
Hyper	**HPX0002**	170,648	164,053	132,087	137,632	61.1%	54.9%	57.4%	64.0%
Hyper	**HPX0003**	266,758	243,518	202,004	197,489	102.5%	84.8%	102.1%	97.5%
Hyper	**HPX0004**	377,484	368,003	244,942	255,891	48.6%	44.9%	46.3%	52.8%
Hyper	**HPX0005**	509,864	582,251	383,630	478,711	61.6%	84.5%	58.0%	97.2%
Hyper	**HPX0006**	274,639	374,260	211,387	329,790	94.3%	164.8%	91.5%	198.7%
Hyper	**HPX0007**	414,662	400,861	323,556	336,181	118.6%	111.3%	128.2%	137.1%
Hyper	**HPX0008**	294,054	321,329	224,522	268,446	75.5%	91.7%	76.3%	110.8%
Hyper	**HPX0009**	412,715	445,504	315,680	369,698	25.5%	35.5%	21.7%	42.6%
Hyper	**HPX0010**	264,711	256,637	202,969	208,862	89.2%	83.4%	87.6%	93.0%
Hyper	**HPX0011**	300,732	321,062	221,585	258,404	41.1%	50.7%	36.1%	58.7%
Hyper	**HPX0012**	394,229	431,123	290,290	342,713	19.4%	30.5%	16.6%	37.7%
Hyper	**HPX0013**	236,536	238,647	174,672	184,209	70.5%	72.0%	71.4%	80.7%
Hyper	**HPX0014**	449,501	447,161	335,021	353,103	72.9%	72.0%	71.3%	80.6%
Hyper	**HPX0015**	237,172	222,544	152,288	161,629	32.7%	24.5%	25.7%	33.4%
Hyper	**HPX0016**	506,633	520,197	376,042	416,758	28.6%	32.0%	24.0%	37.4%
Hyper	**HPX0017**	369,488	375,487	248,596	285,497	34.1%	36.3%	23.7%	42.0%
Hyper	**HPX0018**	401,858	394,315	294,177	308,126	22.8%	20.5%	18.4%	24.0%
Hyper	**HPX0019**	408,796	376,054	314,584	308,504	54.6%	42.2%	55.3%	52.3%
Hyper	**HPX0020**	396,453	406,761	291,857	329,613	19.6%	22.7%	12.0%	26.5%

Table 3-1: An example of generated data from organisation's database to use for store cluster analysis

3.2 Data analysis

After preparing, consolidating and normalising all the data in previous stage, these data which consisting of eight variables are taken to perform the store cluster analysis in this part. At the initial stage, the algorithm of current clustering technique that company have been using will be outlined and used to perform the store cluster analysis. After that, the clustering techniques introduced and outlined in the literature review session including K-Means clustering technique and Agglomerative clustering technique will be used to perform the store cluster analysis in statistical program namely SPSS program. At the end, the results of store cluster analysis from different clustering techniques will be used to evaluate the performance of store cluster analysis in the next stage.

3.3 Result evaluation and summary

After performing the store cluster analysis, the results of store cluster analysis from the current clustering technique that company have been using together with two introduced clustering techniques that performed in SPSS program including K-Means clustering technique and Agglomerative clustering technique will be compared and evaluated in this stage in order to identify the most efficient clustering technique for store cluster analysis task in the future. The performance measurement of store cluster analysis will be outlined and evaluated which consists of two performance measurement. The first performance measurement is the coefficient of variation which is a measure of data dispersion around the mean. In this case, the comparison of coefficient of variation between each clustering technique will be made and summarised. Moreover, another performance measurement of store cluster analysis is to use the result of store cluster analysis from each clustering technique to develop the stock allocation plan in order to compare the total difference between allocated stock target by store cluster and actual stock target per store.

3.4 Thesis completion

The final stage of the research is to consolidate and summarise all the result and complete the thesis book. The thesis book will be followed the guideline, structure and requirement from both Chulalongkorn University and the University of Warwick before finalising all the contents and perform the proofreading to submit to graduate school.

4. ANALYSIS

Overall, the process of developing the stock allocation plan of each event can be summarised into three steps including key item selection, store cluster analysis and stock allocation. At the initial stage, project manager has to identify a group of key items based on the study of historical sales data during last year or past event to prepare for the stock build up and stock allocation plan. After that, the second stage, the project owner will be classifying the store characteristic in order to group the store that has the similarity in terms of sales performance together in order to estimate the stock target by store cluster level. The last step is to develop the stock allocation plan, the estimated stock volume of key item will be allocated according to the study of store cluster analysis. Therefore, it can be seen that the analysis of store cluster will be the key factor that determine how well of the stock allocation plan is. In this chapter, the implementation of different clustering techniques will be applied and analysed in order to identify an appropriate and suitable technique to do the store classification which it will help project manager to maximise the efficiency of stock allocation plan of retail store during special event or seasonal period that usually has high demand or high sales.

4.1 Data preparation

As there are three store formats of Tesco lotus including hypermarket, supermarket, convenient store. Figure 4-1 presents the proportion of sales volume during New Year 2019 festival which breakdown by store formats. It can be seen that between these three formats, hypermarket is the most contributed sales volume to the company. Therefore, in this research, hypermarket will be the selected store format to do the analysis.

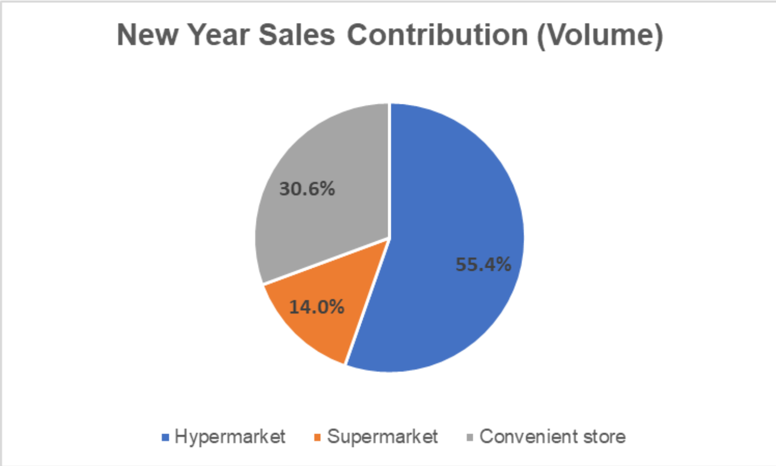

Figure 4-1: New Year 2019 Sales Contribution (Volume) by store format

Moreover, since the objective of the stock allocation plan is to prepare the right amount of stock deliver prior to the store before the event is started to ensure stock availability at the store. As there will be a peak of demand during New Year festival, the plan will also help company to manage and utilise the capacity such as DC space utilisation, manpower and workload during this peak demand. Furthermore, as mentioned at the beginning of this chapter that there are three stages for preparing and developing the stock allocation plan. The first stage is key item selection, in this part project manager will be identifying a group of key items that contribute the most sales volume contribution to the company during the event before suggesting the stock target by item to the store. In this case, New Year 2020's key items of hypermarket store will be identified not only to use for the stock allocation plan of key item but also to use with the store cluster analysis later in the analysis part. The process of identifying a group of key items is to generate the overall sales volume during last New Year 2019 by format and item level as the example of Table 4-1 below presents the first top ten item that contribute the most sales volume during the event. Then, the concept of pareto analysis will be applied in order to identify a group of key items for New Year festival event.

Store Format	Item no.	Sales Volume
Hypermarket	Item 1	1,863,666
Hypermarket	Item 2	1,195,180
Hypermarket	Item 3	1,084,145
Hypermarket	Item 4	1,075,034
Hypermarket	Item 5	894,617
Hypermarket	Item 6	823,547
Hypermarket	Item 7	735,156
Hypermarket	Item 8	546,270
Hypermarket	Item 9	531,219
Hypermarket	Item 10	458,605

Table 4-1: An example of sales volume during New Year 2019 by item and store format

According to Leavengood and Reeb (2002), the pareto analysis is one of the statistical techniques to use when the analyst would like to identify the significant impact on an outcome while considering only a limited number of input factors. According to Koch (1998), the 80/20 principle of pareto analysis will be applied where it considers 20 percent of the total input factors which it contributes 80 percent of the impact on an outcome. In this case, the pareto analysis is applied to select the key item during New Year festival event. Overall, it can be seen from Table 4-2 that key items were selected at 2,308 items out of 26,424 items in total which it considers just only 9 percent of total items in hypermarket store. While, these selected key items contribute about 80 percent of total sales volume of hypermarket store during last New Year 2019.

Store Format	Key item?	Count of item	%to total item	Sales Volume	%Sales Contribution
Hypermarket	Selected Key Item	2,308	8.7%	43,170,416	80.0%
Hypermarket	Other Non-key Item	24,116	91.3%	10,792,201	20.0%
Grand Total		26,424	100.0%	53,962,617	100.0%

Table 4-2: The summary of selected key items during New Year 2019 by applying the pareto analysis

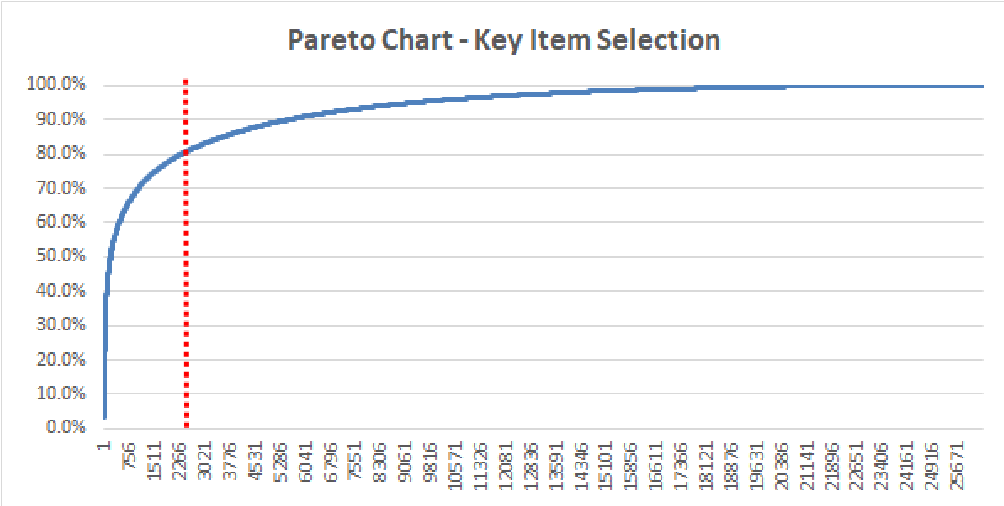

Figure 4-2: Pareto chart of key item selection

Moving to the second stage, the store cluster analysis which is the main part that will be studied and analysed in this research. There are several steps to prepare for the store cluster analysis which can be divided into two parts including generating the data and normalising the data. The first step before performing the store cluster analysis is to generate all the related data from company's database which it consists of eight variables as listed below. These eight variables are the variables that company have been using to perform the store cluster analysis when develop the stock allocation plan. Overall, the variables can be separated into two group including sales volume and sales uplift where sales volume is the total sales volume during New Year period. On the other hand, sales uplift is the percentage difference of sales volume between New Year period and Normal period. In addition, the data from these two group will be further split into the overall total sales volume and total sales volume of selected key item. In this case, all the data will be using the data during the period of 2018 and 2019 as reference to perform the store cluster analysis. Therefore, the total variables are consisting of eight variables as listed below where an example of generated data by store level is presented as Table 4-3 below which will be used to perform the store cluster analysis in next step.

1. Sales Volume by store during New Year 2018 (Overall total sales)
2. Sales Volume by store during New Year 2019 (Overall total sales)

3. Sales Volume by store during New Year 2018 (of Selected Key Items)
4. Sales Volume by store during New Year 2019 (of Selected Key Items)
5. %Sales Uplift of New Year 2018 by store (Overall total sales)
6. %Sales Uplift of New Year 2019 by store (Overall total sales)
7. %Sales Uplift of New Year 2018 by store (of Selected Key Items)
8. %Sales Uplift of New Year 2019 by store (of Selected Key Items)

Before Normalising The Data					
		Variable no.1	Variable no.2	Variable no.3	Variable no.4
Format	Store No.	Sales Volume (Overall) 2018	Sales Volume (Overall) 2019	Sales Volume (Key item) 2018	Sales Volume (Key item) 2019
Hyper	HPX0001	486,295	534,666	349,337	427,682
Hyper	HPX0002	170,648	164,053	132,087	137,632
Hyper	HPX0003	266,758	243,518	202,004	197,489
Hyper	HPX0004	377,484	368,003	244,942	255,891
Hyper	HPX0005	509,864	582,251	383,630	478,711
Hyper	HPX0006	274,639	374,260	211,387	329,790
Hyper	HPX0007	414,662	400,861	323,556	336,181
Hyper	HPX0008	294,054	321,329	224,522	268,446
Hyper	HPX0009	412,715	445,504	315,680	369,698
Hyper	HPX0010	264,711	256,637	202,969	208,862
Hyper	HPX0011	300,732	321,062	221,585	258,404
Hyper	HPX0012	394,229	431,123	290,290	342,713
Hyper	HPX0013	236,536	238,647	174,672	184,209
Hyper	HPX0014	449,501	447,161	335,021	353,103
Hyper	HPX0015	237,172	222,544	152,288	161,629
Hyper	HPX0016	506,633	520,197	376,042	416,758
Hyper	HPX0017	369,488	375,487	248,596	285,497
Hyper	HPX0018	401,858	394,315	294,177	308,126
Hyper	HPX0019	408,796	376,054	314,584	308,504
Hyper	HPX0020	396,453	406,761	291,857	329,613

Table 4-3: An example of generated data of variable number 1-4 (Sales volume) by store level

Before Normalising The Data					
		Variable no.5	Variable no.6	Variable no.7	Variable no.8
Format	Store No.	%Sales Uplift (Overall) 2018	%Sales Uplift (Overall) 2019	%Sales Uplift (Key item) 2018	%Sales Uplift (Key item) 2019
Hyper	HPX0001	41.6%	55.7%	39.4%	70.6%
Hyper	HPX0002	61.1%	54.9%	57.4%	64.0%
Hyper	HPX0003	102.5%	84.8%	102.1%	97.5%
Hyper	HPX0004	48.6%	44.9%	46.3%	52.8%
Hyper	HPX0005	61.6%	84.5%	58.0%	97.2%
Hyper	HPX0006	94.3%	164.8%	91.5%	198.7%
Hyper	HPX0007	118.6%	111.3%	128.2%	137.1%
Hyper	HPX0008	75.5%	91.7%	76.3%	110.8%
Hyper	HPX0009	25.5%	35.5%	21.7%	42.6%
Hyper	HPX0010	89.2%	83.4%	87.6%	93.0%
Hyper	HPX0011	41.1%	50.7%	36.1%	58.7%
Hyper	HPX0012	19.4%	30.5%	16.6%	37.7%
Hyper	HPX0013	70.5%	72.0%	71.4%	80.7%
Hyper	HPX0014	72.9%	72.0%	71.3%	80.6%
Hyper	HPX0015	32.7%	24.5%	25.7%	33.4%
Hyper	HPX0016	28.6%	32.0%	24.0%	37.4%
Hyper	HPX0017	34.1%	36.3%	23.7%	42.0%
Hyper	HPX0018	22.8%	20.5%	18.4%	24.0%
Hyper	HPX0019	54.6%	42.2%	55.3%	52.3%
Hyper	HPX0020	19.6%	22.7%	12.0%	26.5%

Table 4-4: An example of generated data of variable number 5-8 (Sales uplift) by store level

The second step of data preparation is to normalise the data, since there is a difference in value or scale between these eight variables. As can be seen between Table 4-3 and Table 4-4 above, where the data from the first table is the integer while the second table is the percentage. Therefore, to get rid of data inconsistency and reduce the redundancy of the data, all the variables will be normalised before performing the store cluster analysis by using Z-score technique to normalise all the generated data which consisting of eight variables. In this case, the illustration of data normalisation is examined as below Table 4-5.

Remark	Variable no.1	Variable no.2	Variable no.3	Variable no.4
	Sales Volume (Overall) 2018	Sales Volume (Overall) 2019	Sales Volume (Key item) 2018	Sales Volume (Key item) 2019
Average or Mean	234,521	244,789	176,447	199,126
Standard deviation or SD	94,587	102,593	70,104	83,675
Data of store number HPX0001	486,295	534,666	349,337	427,682
Z-score = (Individual data - Mean) ÷ SD	Z-score of variable no.1 = (486,295 - 234,521) ÷ 94,587 = 2.66			
Normalised data of store number HPX0001	**2.66**	**2.83**	**2.47**	**2.73**

Table 4-5: An example of Z-score calculation in order to normalise the data

Furthermore, Table 4-6 and Table 4-7 show the descriptive statistics of the data of both before normalising the data and after normalising the data. If comparing between these two tables, it can be seen that after normalising the data, the mean or average of all variables will be equal to zero, while the standard deviation is equal to one as presented below which it means that all of the eight variables are completely normalised by the Z-score technique.

Descriptive Statistics of Hyper Format "Before Normalising The Data"						
No.	**Variable Name**	**Store**	**Minimum**	**Maximum**	**Mean**	**Deviation**
1	Sales Volume by store during NY2018 (Overall)	196	71,125	517,800	234,521	94,587
2	Sales Volume by store during NY2019 (Overall)	196	73,500	591,557	244,789	102,593
3	Sales Volume by store during NY2018 (Selected Key Item)	196	58,880	384,875	176,447	70,104
4	Sales Volume by store during NY2019 (Selected Key Item)	196	60,551	507,623	199,126	83,675
5	%Sales Uplift of NY2018 by store (Overall)	196	-10.6%	234.0%	73.4%	39.6%
6	%Sales Uplift of NY2019 by store (Overall)	196	10.0%	197.4%	80.0%	39.1%
7	%Sales Uplift of NY2018 by store (Selected Key Item)	196	-7.2%	270.8%	73.0%	43.1%
8	%Sales Uplift of NY2019 by store (Selected Key Item)	196	12.7%	236.4%	93.7%	45.8%

Table 4-6: The descriptive statistics of hypermarket format "Before normalising the data"

Descriptive Statistics of Hyper Format "After Normalising The Data"						
No.	Variable Name	Store	Minimum	Maximum	Mean	Deviation
1	Sales Volume by store during NY2018 (Overall)	196	-1.73	2.99	0.00	1.00
2	Sales Volume by store during NY2019 (Overall)	196	-1.67	3.38	-0.00	1.00
3	Sales Volume by store during NY2018 (Selected Key Item)	196	-1.68	2.97	0.00	1.00
4	Sales Volume by store during NY2019 (Selected Key Item)	196	-1.66	3.69	-0.00	1.00
5	%Sales Uplift of NY2018 by store (Overall)	196	-2.12	4.05	-0.00	1.00
6	%Sales Uplift of NY2019 by store (Overall)	196	-1.79	3.00	-0.00	1.00
7	%Sales Uplift of NY2018 by store (Selected Key Item)	196	-1.86	4.59	-0.00	1.00
8	%Sales Uplift of NY2019 by store (Selected Key Item)	196	-1.77	3.11	-0.00	1.00

Table 4-7: The descriptive statistics of hypermarket format "After normalising the data"

At the end, all the generated data which consisting of eight variables are normalised as shown in Table 4-6 to Table 4-7. Therefore, the data are ready to use to perform the store cluster analysis in next stage.

After Normalising The Data					
Format	Store No.	Variable no.1 Sales Volume (Overall) 2018	Variable no.2 Sales Volume (Overall) 2019	Variable no.3 Sales Volume (Key item) 2018	Variable no.4 Sales Volume (Key item) 2019
Hyper	**HPX0001**	2.66	2.83	2.47	2.73
Hyper	**HPX0002**	-0.68	-0.79	-0.63	-0.73
Hyper	**HPX0003**	0.34	-0.01	0.36	-0.02
Hyper	**HPX0004**	1.51	1.20	0.98	0.68
Hyper	**HPX0005**	2.91	3.29	2.96	3.34
Hyper	**HPX0006**	0.42	1.26	0.50	1.56
Hyper	**HPX0007**	1.90	1.52	2.10	1.64
Hyper	**HPX0008**	0.63	0.75	0.69	0.83
Hyper	**HPX0009**	1.88	1.96	1.99	2.04
Hyper	**HPX0010**	0.32	0.12	0.38	0.12
Hyper	**HPX0011**	0.70	0.74	0.64	0.71
Hyper	**HPX0012**	1.69	1.82	1.62	1.72
Hyper	**HPX0013**	0.02	-0.06	-0.03	-0.18
Hyper	**HPX0014**	2.27	1.97	2.26	1.84
Hyper	**HPX0015**	0.03	-0.22	-0.34	-0.45
Hyper	**HPX0016**	2.88	2.68	2.85	2.60
Hyper	**HPX0017**	1.43	1.27	1.03	1.03
Hyper	**HPX0018**	1.77	1.46	1.68	1.30
Hyper	**HPX0019**	1.84	1.28	1.97	1.31
Hyper	**HPX0020**	1.71	1.58	1.65	1.56

Table 4-8: An example of the data after normalising of variable number 1-4 (Sales uplift) by store level

After Normalising The Data					
		Variable no.5	Variable no.6	Variable no.7	Variable no.8
Format	Store No.	%Sales Uplift (Overall) 2018	%Sales Uplift (Overall) 2019	%Sales Uplift (Key item) 2018	%Sales Uplift (Key item) 2019
Hyper	HPX0001	- 0.80	- 0.62	- 0.78	- 0.50
Hyper	HPX0002	- 0.31	- 0.64	- 0.36	- 0.65
Hyper	HPX0003	0.73	0.12	0.67	0.08
Hyper	HPX0004	- 0.62	- 0.90	- 0.62	- 0.89
Hyper	HPX0005	- 0.30	0.11	- 0.35	0.08
Hyper	HPX0006	0.53	2.17	0.43	2.29
Hyper	HPX0007	1.14	0.80	1.28	0.95
Hyper	HPX0008	0.05	0.30	0.08	0.37
Hyper	HPX0009	- 1.21	- 1.14	- 1.19	- 1.12
Hyper	HPX0010	0.40	0.09	0.34	- 0.02
Hyper	HPX0011	- 0.81	- 0.75	- 0.86	- 0.76
Hyper	HPX0012	- 1.36	- 1.27	- 1.31	- 1.22
Hyper	HPX0013	- 0.07	- 0.20	- 0.04	- 0.28
Hyper	HPX0014	- 0.01	- 0.21	- 0.04	- 0.29
Hyper	HPX0015	- 1.03	- 1.42	- 1.10	- 1.32
Hyper	HPX0016	- 1.13	- 1.23	- 1.14	- 1.23
Hyper	HPX0017	- 0.99	- 1.12	- 1.15	- 1.13
Hyper	HPX0018	- 1.28	- 1.52	- 1.27	- 1.52
Hyper	HPX0019	- 0.47	- 0.97	- 0.41	- 0.90
Hyper	HPX0020	- 1.36	- 1.46	- 1.42	- 1.47

Table 4-9: An example of the data after normalising of variable number 5-8 (Sales uplift) by store level

4.2 Data analysis

Once all the data are generated and normalised, the store cluster analysis will be performed in this stage. Before performing the store cluster analysis with the introduced clustering techniques, the current store clustering technique that company have been using will be outlined and clustered first in order to compare the performance of the current clustering technique with the new clustering techniques in the fifth chapter. To begin, the technique of clustering the store that company have been using nowadays is to consider and rank the variables according to the sales performance of each store before splitting the number of store equally from the total 196 stores into the designed number of cluster, usually the past projects have been splitting the store cluster into four clusters. In this case, the normalised values of eight variables will be averaged and ranked as an example shown in Table 4-6. Then, the total 196 stores are split equally into four clusters which consisting of 49 stores for each cluster. The result of this clustering technique and average of the variables by each cluster are presented as below Table 4-7. Later, this result of store clustering technique will be compared and analysed with the other two new clustering techniques in the fifth chapter.

Format	Store No.	Average	Variable No. 1	Variable No. 2	Variable No. 3	Variable No. 4	Variable No. 5	Variable No. 6	Variable No. 7	Variable No. 8
Hyper	**HPX0156**	2.04	0.54	0.04	1.03	0.27	4.05	2.83	4.59	3.00
Hyper	**HPX0073**	1.94	0.57	0.20	0.73	0.38	3.80	3.00	3.68	3.11
Hyper	**HPX0052**	1.53	1.96	2.43	2.12	2.63	0.33	1.00	0.52	1.27
Hyper	**HPX0005**	1.51	2.91	3.29	2.96	3.34	- 0.30	0.11	- 0.35	0.08
Hyper	**HPX0007**	1.42	1.90	1.52	2.10	1.64	1.14	0.80	1.28	0.95
Hyper	**HPX0031**	1.35	1.61	3.38	1.57	3.69	- 0.79	0.95	- 0.78	1.16
Hyper	**HPX0177**	1.25	0.73	1.15	1.04	1.45	0.96	1.86	0.98	1.81
Hyper	**HPX0184**	1.21	0.24	0.15	0.43	0.26	2.22	2.15	2.21	2.03
Hyper	**HPX0045**	1.16	2.99	2.71	2.97	2.67	- 0.46	- 0.60	- 0.46	- 0.55
Hyper	**HPX0006**	1.14	0.42	1.26	0.50	1.56	0.53	2.17	0.43	2.29

Table 4-10: The data by store ranked by an average of normalised data of eight variables

Clustering Technique	Count of store	Variable no. 1	Variable no. 2	Variable no. 3	Variable no. 4	Variable no. 5	Variable no. 6	Variable no. 7	Variable no. 8
Current Clustering Technique	**196**	**234,521**	**244,789**	**176,447**	**199,126**	**73.4%**	**80.0%**	**73.0%**	**93.7%**
Cluster no. 1	49	311,094	330,266	239,179	274,635	94.8%	105.4%	97.1%	124.0%
Cluster no. 2	49	237,313	249,523	177,743	203,054	79.1%	87.8%	78.5%	102.5%
Cluster no. 3	49	215,336	215,301	162,794	174,025	73.0%	70.9%	72.4%	81.9%
Cluster no. 4	49	174,342	184,067	126,072	144,792	46.6%	56.0%	44.2%	66.6%

Table 4-11: The results table of store cluster analysis of current clustering technique

Moving to the new techniques that will be used to perform the store cluster analysis, the analysis will be analysed and performed the store cluster analysis by using the SPSS program, SPSS refer to Statistical Package for the Social Sciences Program, in order to analyse and classify the store cluster according to the normalised variables listed in 4.1. What is more, there are two techniques of cluster analysis that will be using and performing in the SPSS program including K-Means clustering technique and Agglomerative clustering technique. In this case, the analysis of store cluster will be grouped into three groups which separated by the use of eight variables. At the end, from these three groups, there are six sets in total to perform store cluster analysis in the SPSS program which can be summarised as below.

Group 1: Clustering the store with eight variables

1. K-Means clustering technique with eight variables
2. Agglomerative clustering technique with eight variables

Group 2: Clustering the store with the first four variables

1. K-Means clustering technique with the first four variables
2. Agglomerative clustering technique with the first four variables

Group 3: Clustering the store with the last four variables

1. K-Means clustering technique with the last four variables
2. Agglomerative clustering technique with the last four variables

As can be seen from above, the reason to split the variables into three group to perform store cluster analysis (including eight variables, the first four variables, and the last four variables) is to identify which variable and clustering technique will be the most efficient combination to do the store cluster analysis before developing the stock allocation plan at the end which will be illustrated, tested and recommended later in the fifth chapter. On the other hand, in terms of the number of store cluster, as the stock allocation plan in the past normally analysed and clustered the store cluster into four clusters before estimating the stock target and allocating the stock into each store cluster and item. In this case, four clusters are considered to be the acceptable level to manage and handle for the analyst when developing the stock allocation plan. Therefore, the number of store cluster will be specified at four clusters when performing the store clustering analysis for every group and each clustering technique. At the end, the result from each group and each clustering technique will be outlined and compared in the following part.

4.2.1 Result of the store cluster analysis

Group 1: Clustering the store with eight variables

As performing the store cluster analysis with eight variables in the SPSS program. The result of store cluster analysis with eight variables between K-Means clustering technique and Agglomerative clustering technique can be summarised as in Table 4-8 which shown the number of store split by each cluster together with its average of the variables by each cluster. Moreover, Figure 4-3 shows the comparison of the number of store by each cluster between K-Means clustering technique and Agglomerative clustering technique. In this case, it can be seen that each clustering technique has clustered the total of 196 stores into four clusters differently depending on its algorithm of each clustering technique as described in chapter number two. Moreover, if looking at the average of normalised value by cluster on Figure 4-4 and Figure 4-6, it can also be seen that both clustering techniques have classified the first four variables in descending order from cluster number one to cluster number four. However, both clustering techniques have not clustered the last four variables in descending order as the first four variables. As Figure 4-5 and Figure 4-7 show that there is a fluctuation in average of normalised value between the fifth to the eighth variable where cluster number 2 from both clustering techniques recorded highest value while cluster number 1 is the lowest value. In this case, according to the fluctuation between variables number five to variable number eight. The analysis will be further split the variables into another two group to perform the store cluster analysis including the variable number one to four which are the sales volume variables as Group 2 and the variables number five to eight which are the sales uplift variables as Group 3 to ensure that if clustering the store separately between each type of variable will give a better result when allocating the stock allocation plan. At the end, the result will be outlined, compared and recommended whether which group, between Group 1 to Group 3, and which clustering technique will be the most appropriate approach to do the store cluster analysis in the fifth chapter.

Clustering Technique	Count of store	Variable no. 1	Variable no. 2	Variable no. 3	Variable no. 4
K-Means 8 Variables	**196**	**234,521**	**244,789**	**176,447**	**199,126**
Cluster no. 1	44	377,242	399,594	279,580	321,986
Cluster no. 2	10	240,375	268,571	193,432	233,597
Cluster no. 3	75	206,353	214,054	152,729	170,823
Cluster no. 4	67	171,452	173,982	132,733	144,980
Agglomerative 8 Variables	**196**	**234,521**	**244,789**	**176,447**	**199,126**
Cluster no. 1	54	356,754	377,126	263,892	303,297
Cluster no. 2	4	251,873	236,132	205,521	203,812
Cluster no. 3	87	187,083	197,802	145,096	165,632
Cluster no. 4	51	184,661	185,503	135,059	145,599

Table 4-12: The results of store cluster analysis of Group 1 : Eight variables (Variable no: 1-4)

Clustering Technique	Count of store	Variable no. 5	Variable no. 6	Variable no. 7	Variable no. 8
K-Means 8 Variables	**196**	**73.4%**	**80.0%**	**73.0%**	**93.7%**
Cluster no. 1	44	46.3%	54.8%	44.5%	66.0%
Cluster no. 2	10	148.7%	172.4%	155.8%	202.0%
Cluster no. 3	75	52.8%	59.3%	51.1%	69.7%
Cluster no. 4	67	103.0%	106.0%	104.0%	122.7%
Agglomerative 8 Variables	**196**	**73.4%**	**80.0%**	**73.0%**	**93.7%**
Cluster no. 1	54	44.6%	52.7%	42.6%	63.4%
Cluster no. 2	4	202.7%	183.7%	217.4%	214.5%
Cluster no. 3	87	96.6%	107.3%	97.6%	124.4%
Cluster no. 4	51	53.9%	54.4%	52.2%	64.0%

Table 4-13: The results of store cluster analysis of Group 1 : Eight variables (Variable no: 5-8)

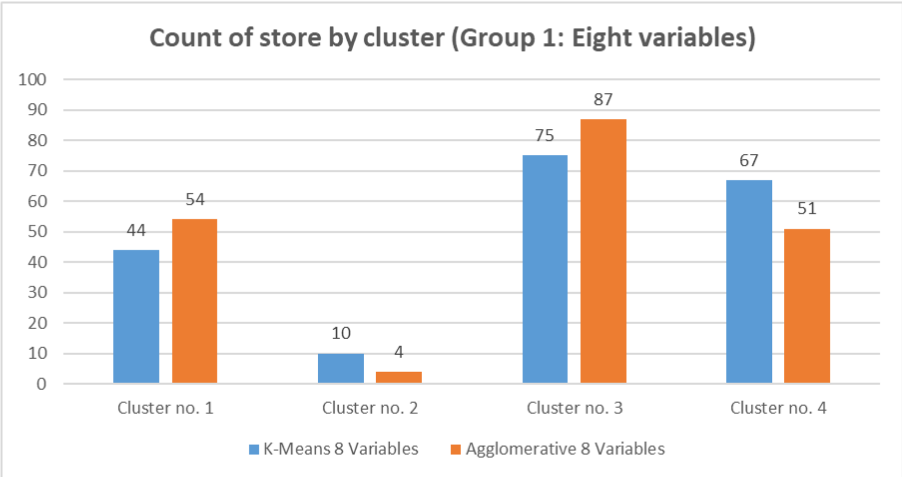

Figure 4-3: Count of store by cluster (Group1: Eight variables)

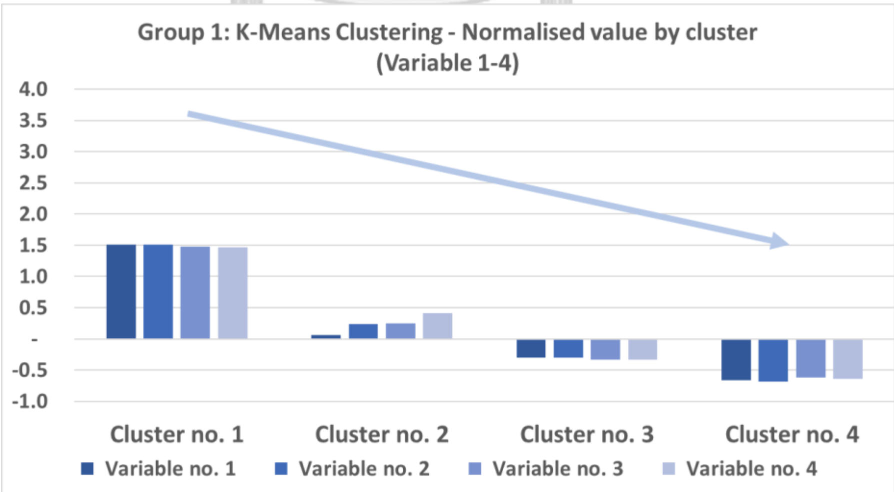

Figure 4-4: The average of normalised value by cluster from K-Means clustering of Group 1 (Variable no. 1-4)

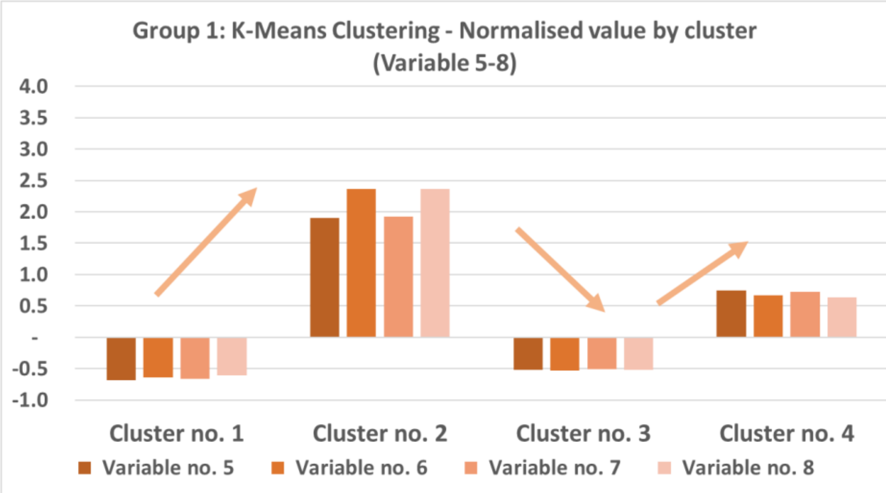

Figure 4-5: The average of normalised value by cluster from K-Means clustering of Group 1 (Variable no. 5-8)

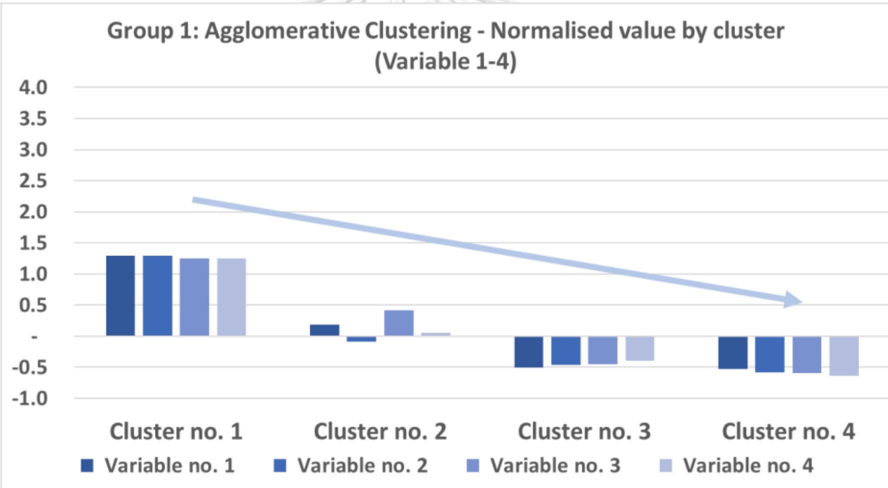

Figure 4-6: The average of normalised value by cluster from Agglomerative clustering of Group 1 (Variable no. 1-4)

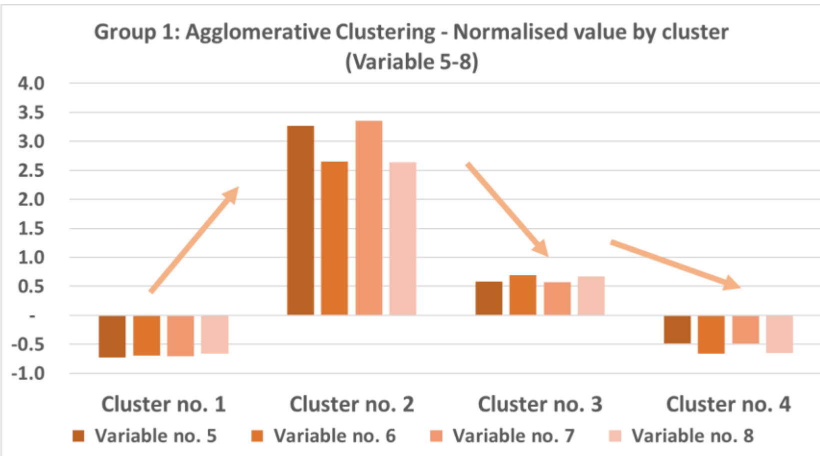

Figure 4-7: The average of normalised value by cluster from Agglomerative clustering of Group 1 (Variable no. 5-8)

Group 2: Clustering the store with the first four variables

Moving to the second group of store cluster analysis which is clustering the store with the first four variables by using the K-Means clustering technique and Agglomerative clustering technique. Overall, it can be seen that after excluding the last four variables to perform store cluster analysis, the result as presented in Table 4-10 shows a significant difference in terms of the number of store by each cluster when compared with the previous result of store cluster analysis from group 1. In this case, if comparing the number of store by each cluster between Figure 4-3 and Figure 4-8 of Group 1 and Group 2 of store cluster analysis, it can be seen that the number of store in cluster number 1 of group 2 is significant less than cluster number 1 of group 1 while the number of store in cluster number 2 of group 2 is remarkable higher than cluster number 1 of group 1. Therefore, after removing the last four variables and perform store cluster analysis in SPSS program, it indicates that the last four variables are a factor of this noticeable change in number of store when clustering the store into four clusters. Moreover, both two clustering techniques have clustered all the first four variables in descending order from cluster 1 to cluster 4 as shown in Figure 4-9 and Figure 4-10 which illustrate the average of normalised value by each cluster.

Average	Count of Store No.	Average of Variable 1	Average of Variable 2	Average of Variable 3	Average of Variable 4
K-Means First 4 Variables	**196**	**234,521**	**244,789**	**176,447**	**199,126**
Cluster no. 1	7	458,803	537,504	340,669	444,148
Cluster no. 2	30	376,420	388,748	280,020	312,322
Cluster no. 3	65	251,249	259,313	189,466	211,173
Cluster no. 4	94	160,966	167,004	122,160	136,424
Agglomerative First 4 Variables	**196**	**234,521**	**244,789**	**176,447**	**199,126**
Cluster no. 1	23	427,134	448,374	318,995	362,268
Cluster no. 2	61	278,794	294,860	209,281	240,558
Cluster no. 3	78	190,019	193,152	143,923	156,937
Cluster no. 4	34	126,886	135,701	95,724	111,221

Table 4-14: The results table of store cluster analysis of Group 2 : The first four variables

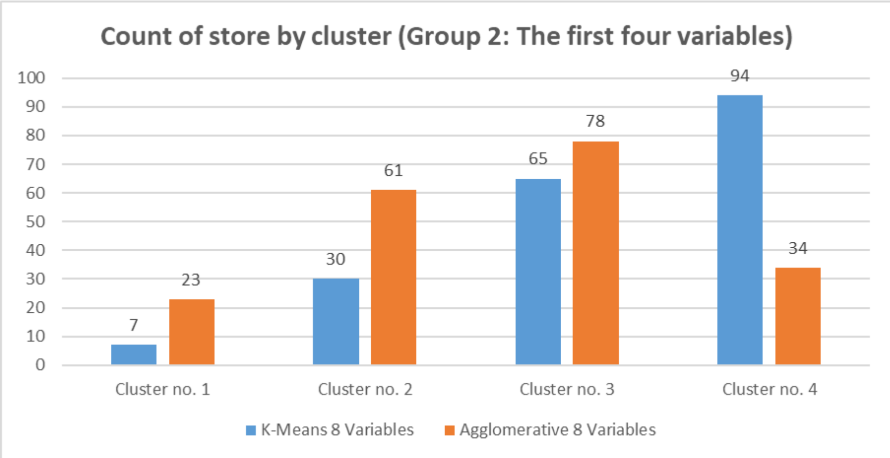

Figure 4-8: Count of store by cluster (Group 2: The first four variables)

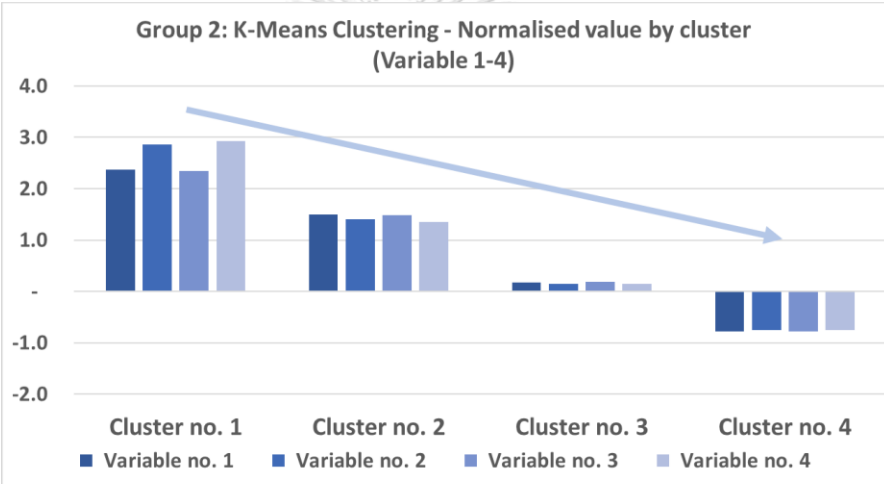

Figure 4-9: The average of normalised value by cluster from K-Means clustering of Group 2 (Variable no. 1-4)

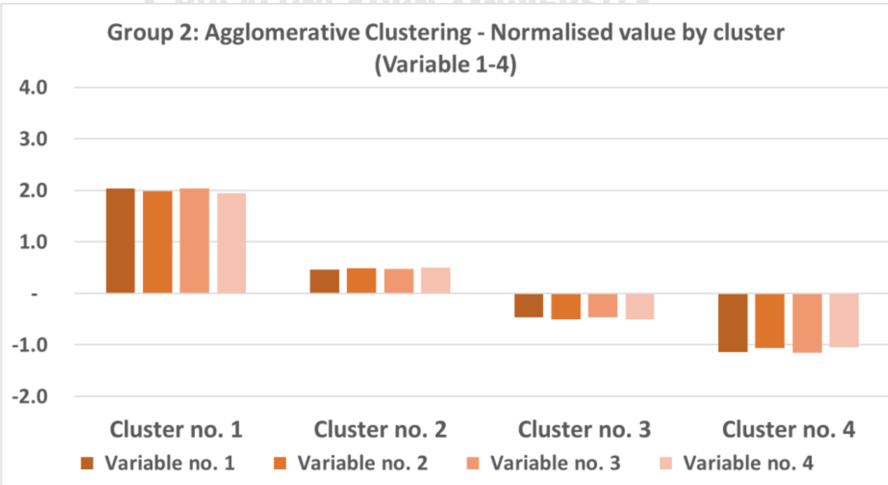

Figure 4-10: The average of normalised value by cluster from Agglomerative clustering of Group 2 (Variable no. 1-4)

Group 3: Clustering the store with the last four variables

The last group of store cluster analysis which clustering the store with the last four variables. In this case, the result of store clustering with the last four variables presents in Table 4-11. Overall, the number of store by cluster of this group almost has the same trend between K-Means clustering technique and Agglomerative clustering technique. It can be seen from the Figure 4-11 that the first and last cluster almost has the same number of store while the difference is in between the middle cluster which are cluster 2 and cluster 3. On the other hand, if considering the average of normalised value between the last four variables as presented in Figure 4-16 and Figure 4-17. It shows that both clustering techniques have clustered the store in descending order from cluster 1 to cluster 4 as Figure 4-12 and Figure 4-13 present below which is unlike the result of store cluster analysis from Group 1. At the end, the result of this store cluster analysis will be evaluated and compared with the other two groups later in the fifth chapter.

Average	Count of Store No.	Average of Variable 5	Average of Variable 6	Average of Variable 7	Average of Variable 8
K-Means Last 4 Variables	**196**	**73.4%**	**80.0%**	**73.0%**	**93.7%**
Cluster no. 1	4	202.7%	183.7%	217.4%	214.5%
Cluster no. 2	34	113.9%	133.8%	117.5%	156.5%
Cluster no. 3	75	85.6%	86.0%	85.6%	99.7%
Cluster no. 4	83	39.5%	47.7%	36.6%	56.8%
Agglomerative Last 4 Variables	**196**	**73.4%**	**80.0%**	**73.0%**	**93.7%**
Cluster no. 1	4	202.7%	183.7%	217.4%	214.5%
Cluster no. 2	66	106.1%	115.7%	107.6%	134.5%
Cluster no. 3	48	74.1%	77.0%	74.1%	89.7%
Cluster no. 4	78	38.6%	46.4%	35.7%	55.5%

Table 4-15: The results table of store cluster analysis of Group 3: The last four variables

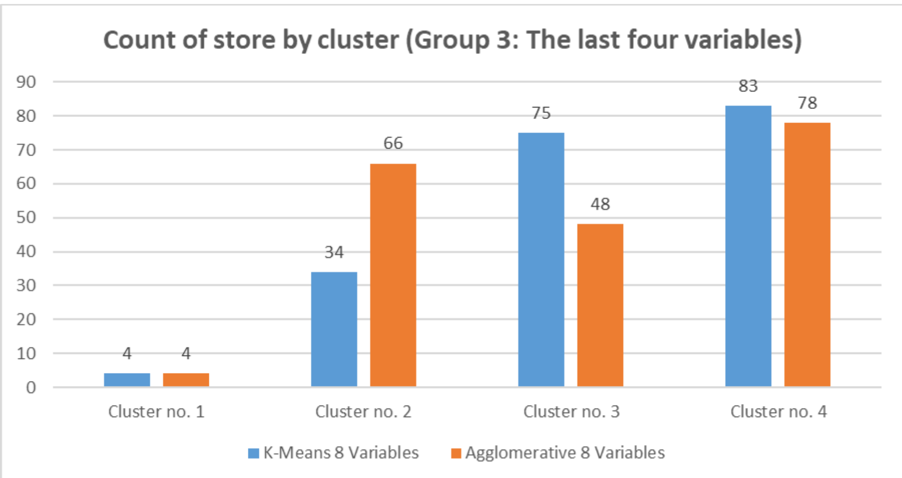

Figure 4-11: Count of store by cluster (Group 3: The last four variables)

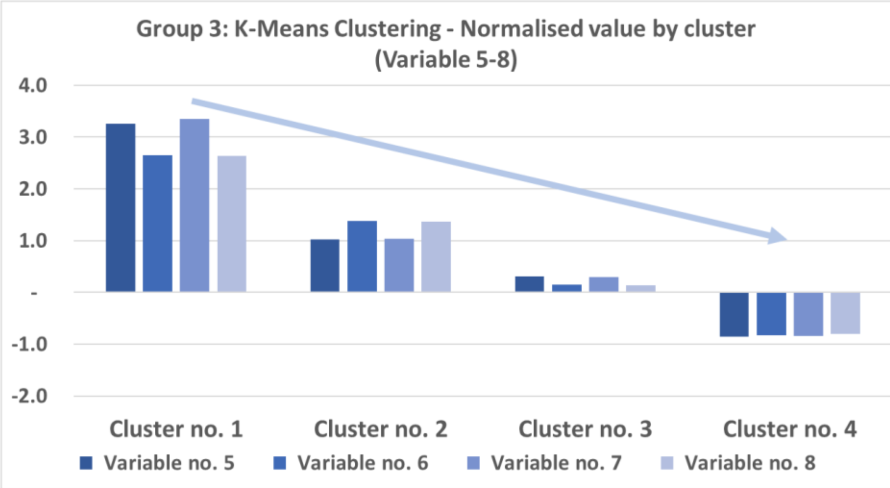

Figure 4-12: The average of normalised value by cluster from K-Means clustering of Group 3 (Variable no. 5-8)

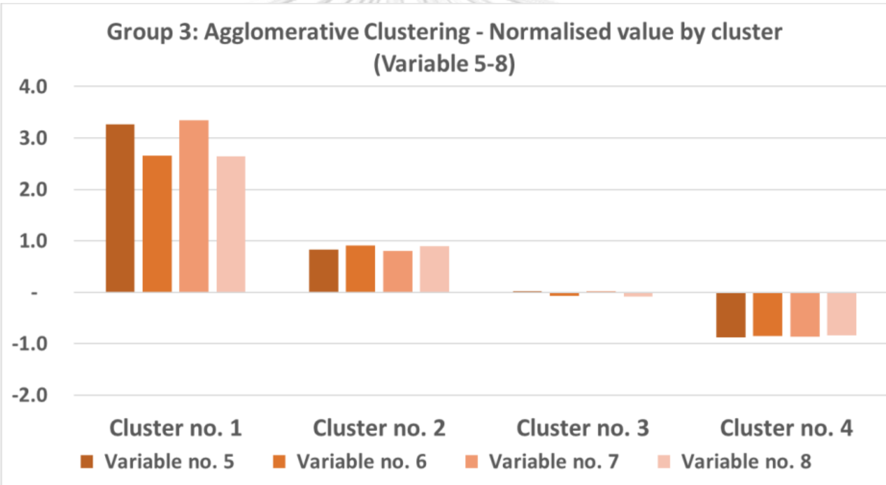

Figure 4-13: The average of normalised value by cluster from Agglomerative clustering of Group 3 (Variable no. 5-8)

4.3 Analysis conclusion

In summary, the analysis chapter is consisting of two parts in overall including data preparation and data analysis. For the first part which is data preparation, there are two processes of data preparation including data generation and data normalisation. In this case, the first step of data preparation is to generate all the data from organisation’s database which consisting of eight variables. Furthermore, the second step of data preparation. Since there is a difference in scale of the data among each variable, all the data will be normalised before performing the store cluster analysis by using the technique of Z-score to normalise the data.

What is more, the second part of this chapter is data analysis. In this part, two introduced clustering techniques including K-Means clustering technique and Agglomerative clustering techniques are performed the store cluster analysis in the SPSS program. Moreover, the store cluster analysis is split into three group including Group 1 that clustered the store with all of eight variables, Group 2 that clustered the store with the first four variables and Group 3 that clustered the store with the last four variables. At the end, the result of store cluster analysis from these three groups together with the result of current clustering technique that company have been using will be used to evaluate and compare the performance of the store cluster analysis between each clustering technique. Then, the suggestion will be made whether which clustering technique will give the most efficient result when develop a stock allocation plan in the following chapter.

5. RESULT AND DISCUSSION

After performing the store cluster analysis in SPSS program. The result of clustering the store from two new clustering techniques including K-Means clustering technique and Agglomerative clustering technique will be measured and compared with the current clustering technique that company have been using nowadays. In this case, the performance measurement of store cluster analysis is consisting of two measurements. Firstly, the coefficient of variation or CV will be calculated and compared between each clustering technique. Furthermore, the second measurement is to develop the stock allocation plan by using the result of store cluster analysis of each clustering technique. Then, the result will be measured and compared which clustering technique will give the minimum difference between allocated stock target by store cluster and actual sales performance by individual store. At the end, the most efficient clustering technique will be recommended according to these performance measurement reviews.

5.1 Performance measurement of store cluster analysis

As mentioned at the introduction part of this chapter that there are two measurement of store cluster analysis including the coefficient of variation and the difference between allocated stock target by store cluster and actual sales performance by individual store. Both measurements will be outlined and measured the performance as following.

5.1.1 The coefficient of variation

The first measurement which is the coefficient of variation, according to Herv́e (2010), the coefficient of variation is the measure of the dispersion of data points around the mean or average. The coefficient of variation can be calculated by taking the standard deviation divided by mean or average. Then, multiplying by 100 percent to convert the value into percentage format as presented in below equation 5-1.

$$\textbf{The coefficient of variation} = (\sigma \div \mu) \times 100\% \qquad (5\text{-}1)$$

Where:

σ = the standard deviation

μ = the mean or average

Therefore, the concept of the coefficient of variation is applied to measure the performance of the store cluster analysis. According to the result of store cluster analysis in chapter four, all of eight variables as listed in section 4.1, will be calculated and compared the coefficient of variation by each group and each clustering technique. According to Marcisz (2013), had described the meaning and interpretation of the coefficient of variation ratio. In this case, the lower of the coefficient of variation indicates the better performance of the store cluster analysis which it means that within the store cluster, all of the stores in that store cluster have no significant difference in terms of the store performance which are the eight variables from sales volume and sales uplift that selected to perform the store cluster analysis. In this case, the coefficient of variation from the result of store cluster analysis that using the current clustering technique together with three group of store cluster analysis that had performed by using K-Means clustering technique and Agglomerative clustering technique will be calculated as below.

Current clustering technique

In this group, the calculated average or mean together with the standard deviation and the coefficient of variation by store cluster from the result of store cluster analysis by using the current clustering technique that company have been using are shown in Table 5-1 to Table 5-3 respectively which will be used to compare the coefficient of variation with the other three group in the next stage. Overall, the range of coefficient of variation in each cluster is between 30.1% – 51.1%

Average	Count of Store No.	Average of Variable 1	Average of Variable 2	Average of Variable 3	Average of Variable 4	Average of Variable 5	Average of Variable 6	Average of Variable 7	Average of Variable 8
Current Clustering technique									
Cluster no. 1	49	311,094	330,266	239,179	274,635	94.8%	105.4%	97.1%	124.0%
Cluster no. 2	49	237,313	249,523	177,743	203,054	79.1%	87.8%	78.5%	102.5%
Cluster no. 3	49	215,336	215,301	162,794	174,025	73.0%	70.9%	72.4%	81.9%
Cluster no. 4	49	174,342	184,067	126,072	144,792	46.6%	56.0%	44.2%	66.6%

Table 5-1: The average of store cluster analysis of current clustering technique

Standard Deviation	Count of Store No.	S.D. of Variable 1	S.D. of Variable 2	S.D. of Variable 3	S.D. of Variable 4	S.D. of Variable 5	S.D. of Variable 6	S.D. of Variable 7	S.D. of Variable 8
Current Clustering technique									
Cluster no. 1	49	110,754	124,092	80,546	100,642	47.9%	48.0%	53.0%	56.3%
Cluster no. 2	49	82,878	86,711	55,427	64,539	35.2%	34.7%	37.8%	40.2%
Cluster no. 3	49	64,078	68,299	46,558	53,864	31.2%	25.3%	33.1%	29.2%
Cluster no. 4	49	50,025	49,176	34,194	36,993	23.1%	24.0%	24.8%	29.1%

Table 5-2: The standard deviation of store cluster analysis of current clustering technique

Coefficient of Variation	Count of Store No.	C.V. of Variable 1	C.V. of Variable 2	C.V. of Variable 3	C.V. of Variable 4	C.V. of Variable 5	C.V. of Variable 6	C.V. of Variable 7	C.V. of Variable 8
Current Clustering Technique									
Cluster no. 1	49	35.6%	37.6%	33.7%	36.6%	50.6%	45.5%	54.6%	45.4%
Cluster no. 2	49	34.9%	34.8%	31.2%	31.8%	44.5%	39.5%	48.1%	39.2%
Cluster no. 3	49	29.8%	31.7%	28.6%	31.0%	42.7%	35.7%	45.8%	35.7%
Cluster no. 4	49	28.7%	26.7%	27.1%	25.5%	49.5%	42.8%	56.1%	43.7%
Average of CV - Current Clustering Technique		**32.2%**	**32.7%**	**30.1%**	**31.2%**	**46.8%**	**40.9%**	**51.1%**	**41.0%**

Table 5-3: The coefficient of variation of store cluster analysis of current clustering technique

Group 1: Store clustering with eight variables

For Group 1 that clustering the store with eight variables by using the clustering technique of K-means clustering and Agglomerative clustering technique. The average or mean are calculated by each store cluster from cluster number one to cluster number four as shown in Table 5-4 while the standard deviation is calculated and presented in Table 5-5. At the end, the coefficient of variation is calculated and presented in Table 5-6.

Average	Count of Store No.	Average of Variable 1	Average of Variable 2	Average of Variable 3	Average of Variable 4	Average of Variable 5	Average of Variable 6	Average of Variable 7	Average of Variable 8
K-Means 8 Variables									
Cluster no. 1	44	377,242	399,594	279,580	321,986	46.3%	54.8%	44.5%	66.0%
Cluster no. 2	10	240,375	268,571	193,432	233,597	148.7%	172.4%	155.8%	202.0%
Cluster no. 3	75	206,353	214,054	152,729	170,823	52.8%	59.3%	51.1%	69.7%
Cluster no. 4	67	171,452	173,982	132,733	144,980	103.0%	106.0%	104.0%	122.7%
Agglomerative 8 Variables									
Cluster no. 1	54	356,754	377,126	263,892	303,297	44.6%	52.7%	42.6%	63.4%
Cluster no. 2	4	251,873	236,132	205,521	203,812	202.7%	183.7%	217.4%	214.5%
Cluster no. 3	87	187,083	197,802	145,096	165,632	96.6%	107.3%	97.6%	124.4%
Cluster no. 4	51	184,661	185,503	135,059	145,599	53.9%	54.4%	52.2%	64.0%

Table 5-4: The average of store cluster analysis of Group 1: Eight variables

Standard Deviation	Count of Store No.	S.D. of Variable 1	S.D. of Variable 2	S.D. of Variable 3	S.D. of Variable 4	S.D. of Variable 5	S.D. of Variable 6	S.D. of Variable 7	S.D. of Variable 8
K-Means 8 Variables									
Cluster no. 1	44	64,986	75,617	50,355	65,791	23.5%	26.0%	27.4%	32.4%
Cluster no. 2	10	50,382	63,738	43,916	58,173	48.6%	17.5%	56.7%	24.9%
Cluster no. 3	75	51,167	49,459	39,229	40,918	22.3%	21.6%	24.5%	25.3%
Cluster no. 4	67	41,333	40,496	32,542	35,169	24.9%	24.4%	26.6%	29.5%
Agglomerative 8 Variables									
Cluster no. 1	54	72,920	83,278	56,230	71,571	22.3%	24.8%	25.8%	30.9%
Cluster no. 2	4	45,231	38,053	41,170	36,298	28.6%	12.4%	38.2%	20.2%
Cluster no. 3	87	54,915	59,841	42,753	52,193	32.7%	30.7%	34.2%	37.3%
Cluster no. 4	51	49,167	48,052	37,799	37,014	22.6%	19.1%	25.8%	22.4%

Table 5-5: The standard deviation of store cluster analysis of Group 1: Eight variables

Coefficient of Variation	Count of Store No.	C.V. of Variable 1	C.V. of Variable 2	C.V. of Variable 3	C.V. of Variable 4	C.V. of Variable 5	C.V. of Variable 6	C.V. of Variable 7	C.V. of Variable 8
K-Means 8 Variables									
Cluster no. 1	44	17.2%	18.9%	18.0%	20.4%	50.8%	47.4%	61.5%	49.1%
Cluster no. 2	10	21.0%	23.7%	22.7%	24.9%	32.7%	10.1%	36.4%	12.3%
Cluster no. 3	75	24.8%	23.1%	25.7%	24.0%	42.3%	36.5%	47.9%	36.3%
Cluster no. 4	67	24.1%	23.3%	24.5%	24.3%	24.1%	23.0%	25.6%	24.1%
Average of CV K-Means 8 Variables		**21.8%**	**22.3%**	**22.7%**	**23.4%**	**37.5%**	**29.3%**	**42.8%**	**30.4%**
Agglomerative 8 Variables									
Cluster no. 1	54	20.4%	22.1%	21.3%	23.6%	49.9%	47.1%	60.7%	48.7%
Cluster no. 2	4	18.0%	16.1%	20.0%	17.8%	14.1%	6.7%	17.5%	9.4%
Cluster no. 3	87	29.4%	30.3%	29.5%	31.5%	33.8%	28.6%	35.1%	30.0%
Cluster no. 4	51	26.6%	25.9%	28.0%	25.4%	42.0%	35.1%	49.4%	35.0%
Average of CV Agglomerative 8 Variables		**23.6%**	**23.6%**	**24.7%**	**24.6%**	**35.0%**	**29.4%**	**40.7%**	**30.8%**

Table 5-6: The coefficient of variation of store cluster analysis of Group 1: Eight variables

Group 2: Store clustering with the first four variables

For Group 2 that clustering the store with the first four variables by using the clustering technique of K-Means and Agglomerative clustering technique. The average or mean together with the standard deviation and the coefficient of variation are calculated and presented in Table 5-7 to Table 5-9 respectively.

Average	Count of Store No.	Average of Variable 1	Average of Variable 2	Average of Variable 3	Average of Variable 4
K-Means First 4 Variables					
Cluster no. 1	7	458,803	537,504	340,669	444,148
Cluster no. 2	30	376,420	388,748	280,020	312,322
Cluster no. 3	65	251,249	259,313	189,466	211,173
Cluster no. 4	94	160,966	167,004	122,160	136,424
Agglomerative First 4 Variables					
Cluster no. 1	23	427,134	448,374	318,995	362,268
Cluster no. 2	61	278,794	294,860	209,281	240,558
Cluster no. 3	78	190,019	193,152	143,923	156,937
Cluster no. 4	34	126,886	135,701	95,724	111,221

Table 5-7: The average of store cluster analysis of Group 2: The first four variables

Standard Deviation	Count of Store No.	S.D. of Variable 1	S.D. of Variable 2	S.D. of Variable 3	S.D. of Variable 4
K-Means First 4 Variables					
Cluster no. 1	7	55,263	33,324	41,577	32,495
Cluster no. 2	30	48,472	40,925	37,818	34,450
Cluster no. 3	65	32,163	37,502	26,368	32,546
Cluster no. 4	94	33,034	32,459	25,156	27,249
Agglomerative First 4 Variables					
Cluster no. 1	23	43,233	70,732	34,062	63,137
Cluster no. 2	61	42,593	49,192	30,709	41,613
Cluster no. 3	78	25,974	25,505	18,892	20,692
Cluster no. 4	34	22,431	24,376	16,357	21,803

Table 5-8: The standard deviation of store cluster analysis of Group 2: The first four variables

Coefficient of Variation	Count of Store No.	C.V. of Variable 1	C.V. of Variable 2	C.V. of Variable 3	C.V. of Variable 4
K-Means First 4 Variables					
Cluster no. 1	7	12.0%	6.2%	12.2%	7.3%
Cluster no. 2	30	12.9%	10.5%	13.5%	11.0%
Cluster no. 3	65	12.8%	14.5%	13.9%	15.4%
Cluster no. 4	94	20.5%	19.4%	20.6%	20.0%
Average of CV K-Means First 4 Variables		**14.6%**	**12.7%**	**15.1%**	**13.4%**
Agglomerative First 4 Variables					
Cluster no. 1	23	10.1%	15.8%	10.7%	17.4%
Cluster no. 2	61	15.3%	16.7%	14.7%	17.3%
Cluster no. 3	78	13.7%	13.2%	13.1%	13.2%
Cluster no. 4	34	17.7%	18.0%	17.1%	19.6%
Average of CV Agglomerative First 4 Variables		**14.2%**	**15.9%**	**13.9%**	**16.9%**

Table 5-9: The coefficient of variation of store cluster analysis of Group 2: The first four variables

Group 3: Store clustering with the last four variables

For Group 3 that clustering the store with the last four variables by using the clustering technique of K-means and Agglomerative clustering technique. The average or mean together with the standard deviation and the coefficient of variation are calculated and presented in Table 5-10 to Table 5-12 respectively.

Average	Count of Store No.	Average of Variable 5	Average of Variable 6	Average of Variable 7	Average of Variable 8
K-Means Last 4 Variables					
Cluster no. 1	4	202.7%	183.7%	217.4%	214.5%
Cluster no. 2	34	113.9%	133.8%	117.5%	156.5%
Cluster no. 3	75	85.6%	86.0%	85.6%	99.7%
Cluster no. 4	83	39.5%	47.7%	36.6%	56.8%
Agglomerative Last 4 Variables					
Cluster no. 1	4	202.7%	183.7%	217.4%	214.5%
Cluster no. 2	66	106.1%	115.7%	107.6%	134.5%
Cluster no. 3	48	74.1%	77.0%	74.1%	89.7%
Cluster no. 4	78	38.6%	46.4%	35.7%	55.5%

Table 5-10: The average of store cluster analysis of Group 3: The last four variables

Standard Deviation	Count of Store No.	S.D. of Variable 5	S.D. of Variable 6	S.D. of Variable 7	S.D. of Variable 8
K-Means Last 4 Variables					
Cluster no. 1	4	28.6%	12.4%	38.2%	20.2%
Cluster no. 2	34	24.7%	24.6%	25.8%	30.6%
Cluster no. 3	75	19.7%	15.9%	20.6%	19.5%
Cluster no. 4	83	17.1%	18.2%	18.7%	21.9%
Agglomerative Last 4 Variables					
Cluster no. 1	4	28.6%	12.4%	38.2%	20.2%
Cluster no. 2	66	25.0%	27.1%	26.7%	33.6%
Cluster no. 3	48	10.3%	12.5%	11.5%	15.8%
Cluster no. 4	78	17.8%	18.4%	20.2%	22.2%

Table 5-11: The standard deviation of store cluster analysis of Group 3: The last four variables

Coefficient of Variation	Count of Store No.	C.V. of Variable 5	C.V. of Variable 6	C.V. of Variable 7	C.V. of Variable 8
K-Means Last 4 Variables					
Cluster no. 1	4	14.1%	6.7%	17.5%	9.4%
Cluster no. 2	34	21.7%	18.4%	22.0%	19.6%
Cluster no. 3	75	23.1%	18.5%	24.1%	19.5%
Cluster no. 4	83	43.2%	38.2%	51.2%	38.5%
Average of CV K-Means Last 4 Variables		**25.5%**	**20.5%**	**28.7%**	**21.7%**
Agglomerative Last 4 Variables					
Cluster no. 1	4	14.1%	6.7%	17.5%	9.4%
Cluster no. 2	66	23.6%	23.4%	24.8%	25.0%
Cluster no. 3	48	13.9%	16.2%	15.5%	17.6%
Cluster no. 4	78	46.2%	39.7%	56.6%	40.0%
Average of CV Agglomerative Last 4 Variables		**24.4%**	**21.5%**	**28.6%**	**23.0%**

Table 5-12: The coefficient of variation of store cluster analysis of Group 3: The last four variables

After calculated the coefficient of variation to all group and all clustering techniques above. In this part, the average of coefficient of variation from cluster number one to cluster number four by each clustering technique will be compared between each group and also compare with the current clustering technique. To begin with Group 1 which is clustering the store with eight variables, Table 5-13 and Figure 5-1 shows the result and the comparison of the average of coefficient of variation from cluster number one to cluster number four by each clustering technique including K-Means clustering technique and Agglomerative clustering technique and current clustering technique that company have been using. Overall, it can be seen that both K-Means clustering technique and Agglomerative clustering technique have a significant better result in terms of the overall average of coefficient of variation, around 9.5 percent and 9.3 percent less respectively when compared with the current clustering technique that company have been using while K-Means clustering technique has slightly lower coefficient of variation than Agglomerative clustering technique of around 0.2 percent.

Average of Coefficient of Variation	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Average
K-Means - 8 Variables	21.8%	22.3%	22.7%	23.4%	37.5%	29.3%	42.8%	30.4%	28.8%
Agglomerative - 8 Variables	23.6%	23.6%	24.7%	24.6%	35.0%	29.4%	40.7%	30.8%	29.0%
Current Clustering Technique	32.2%	32.7%	30.1%	31.2%	46.8%	40.9%	51.1%	41.0%	38.3%

Table 5-13: The overall average of coefficient of variation of Group 1: Clustering with eight variables

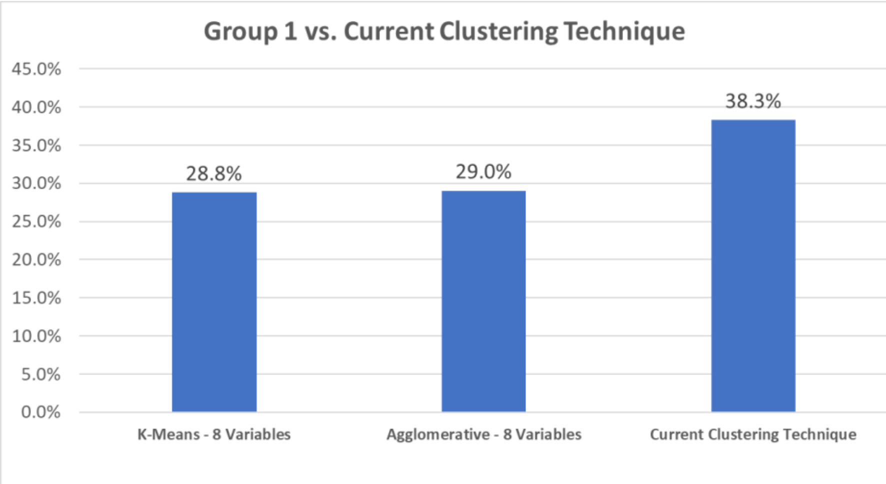

Figure 5-1: The comparison of overall average of coefficient of variation between Group 1 and Current clustering technique

For Group 2 which is clustering the store with the first four variables, the result of this group also shows the same trend as Group 1 where K-Means clustering technique and Agglomerative clustering technique have a remarkable better result in terms of the overall average of the coefficient of variation when compared with the current clustering technique. Around 15 percent lower in terms of the overall average of the coefficient of variation as shown in Table 5-14 and Figure 5-2. Furthermore, in this group, the K-Means clustering technique has better performance than the Agglomerative clustering technique.

Average of Coefficient of Variation	Variable 1	Variable 2	Variable 3	Variable 4	Average
K-Means - First 4 Variables	14.6%	12.7%	15.1%	13.4%	13.9%
Agglomerative - First 4 Variables	14.2%	15.9%	13.9%	16.9%	15.2%
Current Clustering Technique	32.2%	32.7%	30.1%	31.2%	31.6%

Table 5-14: The overall average of coefficient of variation of Group 2: Clustering with the first four variables

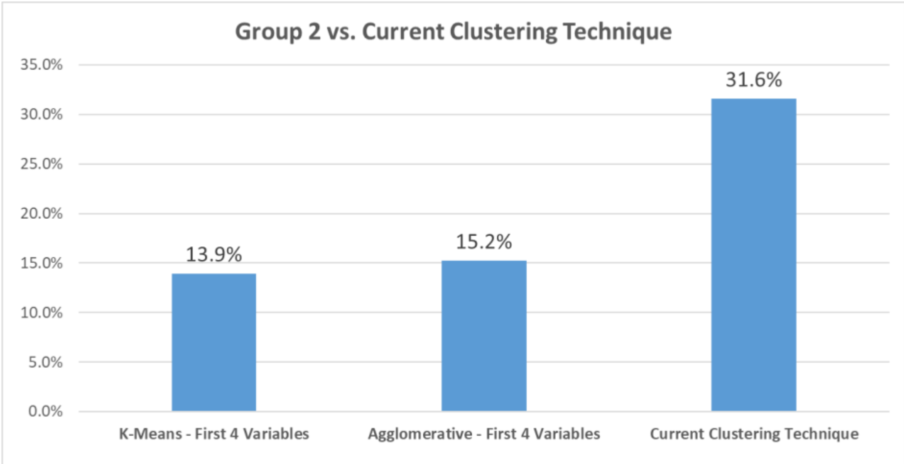

Figure 5-2: The comparison of overall average of coefficient of variation between Group 2 and Current clustering technique

Lastly, Group 3 which is clustering the store with the last four variables. In this group, the result also shows the same trend as Group 1 and Group 2 where the overall average of coefficient of variation from both K-Means clustering technique and Agglomerative clustering technique are notable less than the current clustering technique that company have been using which is around 20 percent less as shown in Table 5-15 and Figure 5-3. In this group, K-Means clustering technique also give a better result than the Agglomerative clustering as same as Group 2.

Average of Coefficient of Variation	Variable 5	Variable 6	Variable 7	Variable 8	Average
K-Means - Last 4 Variables	25.5%	20.5%	28.7%	21.7%	24.1%
Agglomerative - Last 4 Variables	24.4%	21.5%	28.6%	23.0%	24.4%
Current Clustering Technique	46.8%	40.9%	51.1%	41.0%	45.0%

Table 5-15: The overall average of coefficient of variation of Group 3: Clustering with the last four variables

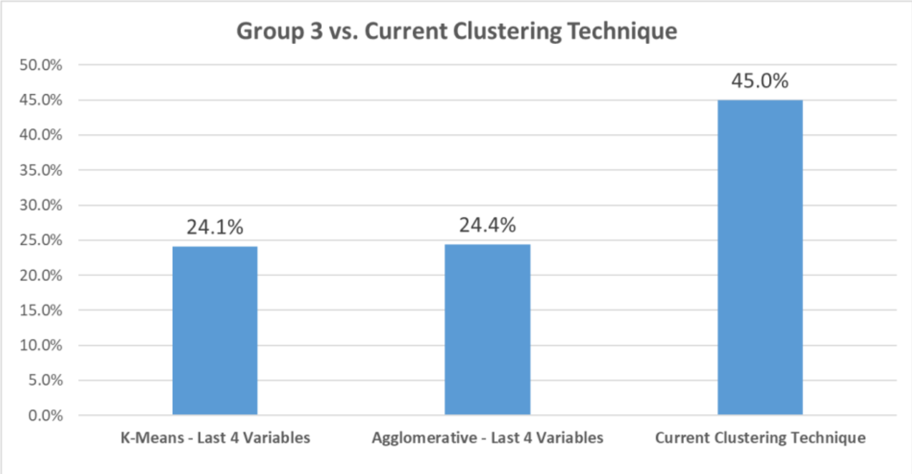

Figure 5-3: The comparison of overall average of coefficient of variation between Group 3 and Current clustering technique

At the end, to summarise the comparison of the first performance measurement above which is the coefficient of variation. It can be summarised that all of the clustering techniques that performed in SPSS program including K-Means clustering technique and Agglomerative clustering technique have a significant better result than the current clustering technique that company have been using. While, K-Means clustering technique give the least coefficient of variation among these three group.

5.1.2 The result of stock allocation plan from each clustering technique

Moving to the second performance measurement which is the difference between allocated stock target by store cluster and actual sales performance by individual store. Normally, stock allocation plan will be developed at the third or final stage of the project after key item selection and store cluster analysis that researcher have already analysed in the fourth chapter of this research. In this case, the result of both selected key items together with store cluster analysis will be used to develop a stock allocation plan which project manager will be using this plan for the stock buildup activity during the event preparation. Therefore, another measurement to determine the performance of store cluster analysis can be measured by developing a stock allocation plan of selected key items based on the result of each store clustering technique. Then, the comparison will be made between each clustering technique in order to identify and suggest which clustering technique will give a better result of stock allocation plan. At the end, the result of each clustering technique will be measured by the total difference between allocated stock target of key items by store cluster for each individual store and compared with the actual stock target of key items of each individual store.

What is more, there are four steps needed to develop a stock allocation plan as below bullet points before determining the performance of store cluster analysis from each clustering technique at the last step.

- Identify the actual stock target of selected key items by each individual store
- Identify the allocated stock target by store cluster level of each clustering technique for each individual store and each selected key item
- Calculate the total difference between allocated stock target per store and actual stock target by individual store
- Compare the result between each clustering technique and identify the best clustering technique

Step 1: Identify the stock target of selected key items by each individual store

To begin, the actual sales volume of each selected key item by store level during past event will be considered as the stock target of that store. In this case, sales volume of selected key items by store level during past New Year 2019 event are generated as an example shown in Table 5-16. It can be seen from the table below that the first row of the table which is item number 1 of store number HPX0001 where it shows sales volume recorded at 44,108 units during past New Year 2019. Thus, the stock target of item number 1 and store HPX0001 will be equal to 44,108 units in this case.

Store no.	Item no.	Sales Volume = Stock target by store
HPX0001	Item 1	44,108
HPX0002	Item 1	9,174
HPX0003	Item 1	7,976
HPX0004	Item 1	2,115
HPX0005	Item 1	16,635
HPX0006	Item 1	48,956
HPX0007	Item 1	19,044
HPX0008	Item 1	9,930
HPX0009	Item 1	11,754
HPX0010	Item 1	2,591
HPX0001	Item 2	6,348
HPX0002	Item 2	1,992
HPX0003	Item 2	4,836
HPX0004	Item 2	5,136
HPX0005	Item 2	15,733
HPX0006	Item 2	26,623
HPX0007	Item 2	4,344
HPX0008	Item 2	13,680
HPX0009	Item 2	4,944
HPX0010	Item 2	4,369

Table 5-16: An example of stock target by store level of selected key item

Step 2: Identify the stock target by store cluster level

In this step, the result of store cluster analysis from the fourth chapter will be using in order to determine the stock target by store cluster level. In this case, the stock target by store cluster will be calculated according to all of the result of store cluster analysis which it includes three group of store cluster analysis with two introduced clustering technique including K-Means clustering technique, Agglomerative clustering technique and another stock target by store cluster level from using the current clustering technique that company have been using. Therefore, the result of stock target by store cluster can be summarised as following before comparing the result in the next step.

- Stock target by store cluster of Group 1: Clustering the store with eight variables
- Stock target by store cluster of Group 2: Clustering the store with the first four variables
- Stock target by store cluster of Group 3: Clustering the store with the last four variables

- Stock target by store cluster of current clustering technique

After that the stock target by store cluster will be calculated based on the average of sales volume during the past event by store cluster. To begin, sales volume of selected key item during New Year 2019 event will be generated by store level. As can be seen from Table 5-17 where the table shows the example of sales volume of key item number one by store level

Store no.	Item no.	Sales Volume
HPX0001	Item 1	44,108
HPX0002	Item 1	9,174
HPX0003	Item 1	7,976
HPX0004	Item 1	2,115
HPX0005	Item 1	16,635
HPX0006	Item 1	48,956
HPX0007	Item 1	19,044
HPX0008	Item 1	9,930
HPX0009	Item 1	11,754
HPX0010	Item 1	2,591

Table 5-17: An example of generated sales volume of key item number one by store level

Then, the store will be mapped with the result of store cluster analysis of each clustering technique. In this case, the result of store cluster analysis of Group 1 including K-Means clustering and Agglomerative clustering are examined and mapped as shown in below Table 5-18. It can be seen from the fourth and fifth column of the table that item number one of store number HPX0002, the K-Means clustering technique that had clustered this store with eight variables show that store number HPX0002 is in the store cluster number three while the Agglomerative clustering technique that had clustered this store with eight variables show that store number HPX0002 is in the cluster number four.

Store no.	Item no.	Sales Volume	Clustering Technique	Cluster no
HPX0001	Item 1	44,108	K-Means - 8 variables	1
HPX0002	Item 1	9,174	K-Means - 8 variables	3
HPX0003	Item 1	7,976	K-Means - 8 variables	4
HPX0004	Item 1	2,115	K-Means - 8 variables	1
HPX0005	Item 1	16,635	K-Means - 8 variables	1
HPX0006	Item 1	48,956	K-Means - 8 variables	2
HPX0007	Item 1	19,044	K-Means - 8 variables	1
HPX0008	Item 1	9,930	K-Means - 8 variables	1
HPX0009	Item 1	11,754	K-Means - 8 variables	1
HPX0010	Item 1	2,591	K-Means - 8 variables	3
HPX0001	Item 1	44,108	Agglomerative - 8 variables	1
HPX0002	Item 1	9,174	Agglomerative - 8 variables	4
HPX0003	Item 1	7,976	Agglomerative - 8 variables	3
HPX0004	Item 1	2,115	Agglomerative - 8 variables	1
HPX0005	Item 1	16,635	Agglomerative - 8 variables	1
HPX0006	Item 1	48,956	Agglomerative - 8 variables	3
HPX0007	Item 1	19,044	Agglomerative - 8 variables	1
HPX0008	Item 1	9,930	Agglomerative - 8 variables	1
HPX0009	Item 1	11,754	Agglomerative - 8 variables	1
HPX0010	Item 1	2,591	Agglomerative - 8 variables	3

Table 5-18: An example of generated sales volume by item and store level mapped with the result of store cluster analysis

At the end, sales volume will be summed up into store cluster and item level by each clustering technique. Then, calculate the average of sales volume per store by each store cluster level and each selected key item. Finally, this average value will be considered as the allocated stock target per store for each item and each store cluster as presented in Table 5-19. From the table, it can be seen that cluster number one of K-means clustering technique that clustered the store with eight variables has the allocated stock target per store of item number 1 equal to 14,926 units. Therefore, it means that all of the store which consists of 44 stores, in cluster number one from this clustering technique will have the stock target of item number one equal to 14,926 units.

Clustering Technique	Cluster no.	Item no.	Sum of Sales Volume	Count of store	Average per store cluster
K-Means - 8 variables	**1**	**Item 1**	656,744	44	14,926
K-Means - 8 variables	**2**	**Item 1**	263,264	10	26,326
K-Means - 8 variables	**3**	**Item 1**	484,354	75	6,458
K-Means - 8 variables	**4**	**Item 1**	459,304	67	6,855
Agglomerative - 8 variables	**1**	**Item 1**	738,023	54	13,667
Agglomerative - 8 variables	**2**	**Item 1**	87,349	4	21,837
Agglomerative - 8 variables	**3**	**Item 1**	836,696	87	9,617
Agglomerative - 8 variables	**4**	**Item 1**	201,598	51	3,953

Table 5-19: A calculated example of stock target by store cluster for each selected key item

Step 3: Calculate the total difference between allocated stock target and actual stock target

In this step, the result from step 1 and step 2 will be used to calculate the difference between allocated stock target by store cluster and actual stock target per store. As presented in Table 5-20 below where the first row which is item number one of store number HPX0001 has the absolute difference of 29,182 units which came from the absolute difference between allocated stock target per store of 14,926 units and actual stock target of this store at 44,108 units.

Store no.	Item no.	Clustering Technique	Cluster no	Stock target by store cluster (Step 2)	Stock target by individual store (Step 1)	Absolute difference
HPX0001	Item 1	K-Means - 8 variables	1	14,926	44,108	**|14,026 - 44,108| = 29,182**
HPX0002	Item 1	K-Means - 8 variables	3	6,458	9,174	**2,716**
HPX0003	Item 1	K-Means - 8 variables	4	6,855	7,976	**1,121**
HPX0004	Item 1	K-Means - 8 variables	1	14,926	2,115	**12,811**
HPX0005	Item 1	K-Means - 8 variables	1	14,926	16,635	**1,709**
HPX0006	Item 1	K-Means - 8 variables	2	26,326	48,956	**22,630**
HPX0007	Item 1	K-Means - 8 variables	1	14,926	19,044	**4,118**
HPX0008	Item 1	K-Means - 8 variables	1	14,926	9,930	**4,996**
HPX0009	Item 1	K-Means - 8 variables	1	14,926	11,754	**3,172**
HPX0010	Item 1	K-Means - 8 variables	3	6,458	2,591	**3,867**
HPX0001	Item 1	Agglomerative - 8 variables	1	13,667	44,108	**30,441**
HPX0002	Item 1	Agglomerative - 8 variables	4	3,953	9,174	**5,221**
HPX0003	Item 1	Agglomerative - 8 variables	3	9,617	7,976	**1,641**
HPX0004	Item 1	Agglomerative - 8 variables	1	13,667	2,115	**11,552**
HPX0005	Item 1	Agglomerative - 8 variables	1	13,667	16,635	**2,968**
HPX0006	Item 1	Agglomerative - 8 variables	3	9,617	48,956	**39,339**
HPX0007	Item 1	Agglomerative - 8 variables	1	13,667	19,044	**5,377**
HPX0008	Item 1	Agglomerative - 8 variables	1	13,667	9,930	**3,737**
HPX0009	Item 1	Agglomerative - 8 variables	1	13,667	11,754	**1,913**
HPX0010	Item 1	Agglomerative - 8 variables	3	9,617	2,591	**7,026**

Table 5-20: An example of the difference between allocated stock target by store cluster and actual stock target per store

In addition, all of the clustering techniques will be calculated the difference between allocated stock target by store cluster and actual stock target per store before summing up the total difference by each of the clustering technique in order to compare and analyse the performance in the next step. In this case, the total difference will be calculated and split into four groups for the comparison including the total difference from top 10 of key items, top 50 of key items, top 100 of key items and top 500 of key items respectively. Furthermore, as can be seen from Table 5-21 where there are 26,424 items available in hypermarket store format, the top 500 items in terms of sales volume already contributed to 62 percent of the total sales performance of hypermarket store. Thus, it means that these four group can be represented the overall sales performance of the company in hypermarket store format.

Item	Count of item	Sales Volume	%Sales contribution
Top 10 items	10	9,207,438	17%
Top 50 items	50	19,124,312	35%
Top 100 items	100	23,391,214	43%
Top 500 items	500	33,196,160	62%
All item	**26,424**	**53,962,617**	**100%**

Table 5-21: Key item split into four layer for making the comparison

To begin, the result of the total difference between allocated stock target by store cluster of the current clustering technique that company have been using and actual stock target per store is shown as below Table 5-22.

Current Clustering Technique	Total difference
Top 10 items	5,036,849
Top 50 items	10,355,567
Top 100 items	12,568,777
Top 500 items	17,818,056

Table 5-22: The result of the total difference from current clustering technique

Furthermore, the result of the total difference between allocated stock target by store cluster of other clustering technique and actual stock target per store will be calculated and summarised into three group as below.

Group 1: Store clustering with eight variables

For the first group, the total difference between allocated stock target by store cluster and actual stock target per store of Group 1: Store clustering with eight variables is presented as shown in Table 5-23.

Top 10 items	Total difference
K-Means - 8 Variables	4,521,020
Agglomerative - 8 Variables	4,562,247

Top 50 items	Total difference
K-Means - 8 Variables	9,321,888
Agglomerative - 8 Variables	9,393,369

Top 100 items	Total difference
K-Means - 8 Variables	11,226,577
Agglomerative - 8 Variables	11,327,278

Top 500 items	Total difference
K-Means - 8 Variables	15,672,717
Agglomerative - 8 Variables	15,830,644

Table 5-23: The result of the total difference from Group 1: Custering the store with eight variables

Group 2: Store clustering with the first four variables

In this group, the total difference between allocated stock target by store cluster and actual stock target per store of Group 2: Store clustering with the first four variables is presented as Table 5-24.

Top 10 items	Total difference
K-Means - The first four variables	4,583,096
Agglomerative - The first four variables	4,648,002

Top 50 items	Total difference
K-Means - The first four variables	9,424,472
Agglomerative - The first four variables	9,449,249

Top 100 items	Total difference
K-Means - The first four variables	11,296,321
Agglomerative - The first four variables	11,316,449

Top 500 items	Total difference
K-Means - The first four variables	15,690,766
Agglomerative - The first four variables	15,677,432

Table 5-24: The result of the total difference from Group 2: Custering the store with the first four variables

Group 3: Store clustering with the last four variables

In the last group, the total difference between allocated stock target by store cluster and actual stock target per store of Group 3: Store clustering with the last four variables is presented as Table 5-25.

Top 10 items	Total difference
K-Means - The last four variables	5,055,806
Agglomerative - The last four variables	5,059,770

Top 50 items	Total difference
K-Means - The last four variables	10,341,151
Agglomerative - The last four variables	10,316,142

Top 100 items	Total difference
K-Means - The last four variables	12,534,052
Agglomerative - The last four variables	12,491,009

Top 500 items	Total difference
K-Means - The last four variables	17,609,820
Agglomerative - The last four variables	17,522,141

Table 5-25: The result of the total difference from Group 2: Custering the store with the last four variables

Step 4: The comparison of the result between each clustering technique

In this step, the result of the total difference between allocated stock target by store cluster and actual stock target per store in step 3 will be compared and analysed between each clustering technique together with the percentage of improvement from the total difference when compared with the current clustering technique that company have been using. The comparison below will be split into three group as same as when performing the store cluster analysis on the SPSS program in the fourth chapter. At the end, the clustering technique that give the minimum result of the total difference between allocated stock target by store cluster and actual stock target per store will be identified and recommended as a preferred clustering technique for the company when performing the store cluster analysis.

Group 1: Store clustering with eight variables

For Group 1 that had clustered the store with eight variables. Overall, it can be seen from the Table 5-26 that all of the clustering techniques that had performed in SPSS program including K-Means clustering technique and Agglomerative clustering technique have better result in terms of the total difference between allocated stock target by store cluster and actual stock target per store than the current clustering technique around 12 percent in overall. On the other hand, K- Means clustering technique has the least total difference in this case which it indicates that K-Means clustering technique provides the most efficient performance when classifying the store.

Top 10 items	Total difference	%Improvement from current clustering technique
K-Means - 8 Variables	4,521,020	11.4%
Agglomerative - 8 Variables	4,562,247	10.4%
Current Clustering Technique	5,036,849	0.0%
Top 50 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - 8 Variables	9,321,888	11.1%
Agglomerative - 8 Variables	9,393,369	10.2%
Current Clustering Technique	10,355,567	0.0%
Top 100 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - 8 Variables	11,226,577	12.0%
Agglomerative - 8 Variables	11,327,278	11.0%
Current Clustering Technique	12,568,777	0.0%
Top 500 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - 8 Variables	15,672,717	13.7%
Agglomerative - 8 Variables	15,830,644	12.6%
Current Clustering Technique	17,818,056	0.0%

Table 5-26: The total difference and the percentage improvement from current clustering technqiue of Group 1

Group 2: Store clustering with the first four variables

As same as the first group, the result from clustering the store with the first four variables from both K-means clustering technique and Agglomerative clustering technique also show a significant better result, around 10 percent better, of the total difference between allocated stock target by store cluster and actual stock target per store when compared with the current clustering technique that company have been using. It also can be seen from the Table 5-27 below that K-Means clustering technique has the best result among these three clustering techniques in group Top 10 items, Top 50 items and Top 100 items. While in group Top 500 items, K-Means and Agglomerative clustering technique recorded about the same percentage of improvement from the current clustering technique of around 13.7 percent where Agglomerative clustering technique show a slightly better result of around 0.1 percent.

Top 10 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The first four variables	4,583,096	9.9%
Agglomerative - The first four variables	4,648,002	8.4%
Current Clustering Technique	5,036,849	0.0%
Top 50 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The first four variables	9,424,472	9.9%
Agglomerative - The first four variables	9,449,249	9.6%
Current Clustering Technique	10,355,567	0.0%
Top 100 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The first four variables	11,296,321	11.3%
Agglomerative - The first four variables	11,316,449	11.1%
Current Clustering Technique	12,568,777	0.0%
Top 500 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The first four variables	15,690,766	13.6%
Agglomerative - The first four variables	15,677,432	13.7%
Current Clustering Technique	17,818,056	0.0%

Table 5-27: The total difference and the percentage improvement from current clustering technqiue of Group 2

Group 3: Store clustering with the last four variables

In this group, the result of total difference between allocated stock target by store cluster and actual stock target per store from two clustering techniques including K-Means clustering technique and Agglomerative clustering technique have no significant difference when compared with the current clustering technique that company have been. The percentage improvement from the current clustering technique is differed about 1 percent improvement in overall. Especially, if considering a group of Top 10 items, it can be seen from Table 5-28 below that when allocating the stock target by store cluster into each individual store, both two clustering techniques give poorer result when compared with the current clustering technique that company have been using. In this case, it can be concluded that clustering the store solely with the last four variables either K-Means clustering technique or Agglomerative clustering technique are not show a significant improvement or benefit when develop a stock allocation plan.

Top 10 items	Total difference	%Improvement from current clustering technique
K-Means - The last four variables	5,055,806	-0.4%
Agglomerative - The last four variables	5,059,770	-0.5%
Current Clustering Technique	5,036,849	0.0%
Top 50 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The last four variables	10,341,151	0.1%
Agglomerative - The last four variables	10,316,142	0.4%
Current Clustering Technique	10,355,567	0.0%
Top 100 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The last four variables	12,534,052	0.3%
Agglomerative - The last four variables	12,491,009	0.6%
Current Clustering Technique	12,568,777	0.0%
Top 500 items	**Total difference**	**%Improvement from current clustering technique**
K-Means - The last four variables	17,609,820	1.2%
Agglomerative - The last four variables	17,522,141	1.7%
Current Clustering Technique	17,818,056	0.0%

Table 5-28: The total difference and the percentage improvement from current clustering technqiue of Group 3

On the other hand, if consolidate all the result together between each clustering technique by group of item including Top 10, Top 50, Top 100 and Top 500 items of the total difference between allocated stock target by store cluster and actual stock target per store, it can be clearly noticed from the Figure 5-4 to Figure 5-7 that there are three clustering techniques that show a noticeable poorer result than the other four clustering techniques. The three clustering techniques that gave poorer result including K-Means clustering technique and Agglomerative clustering technique that had clustered the store with the last four variables. The other one is the current clustering technique that company have been using.

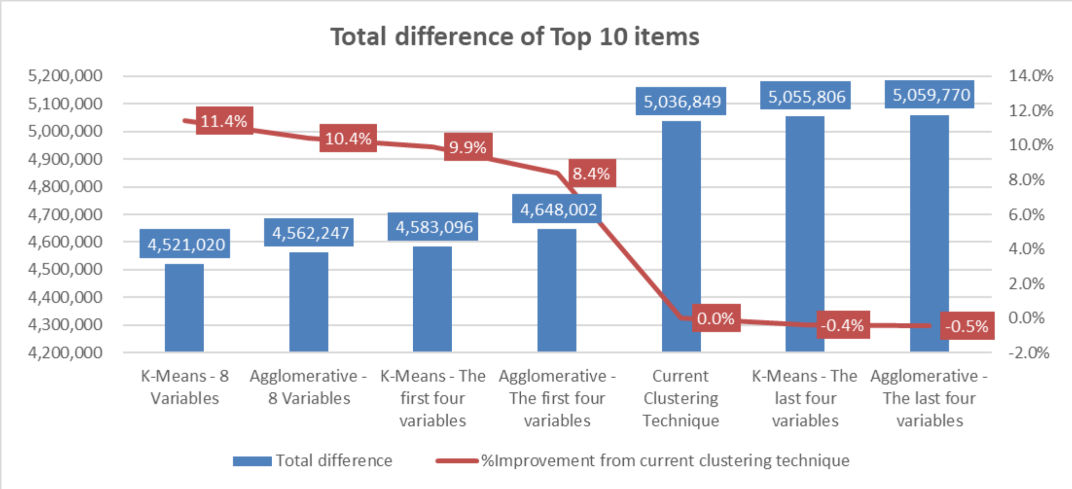

Figure 5-4: The total difference and the percentage improvement of Top 10 items

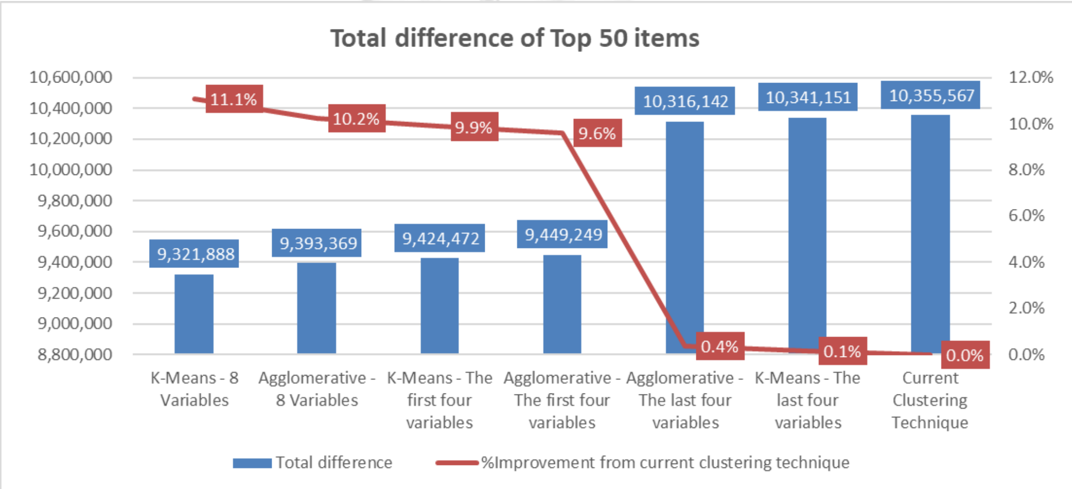

Figure 5-5: The total difference and the percentage improvement of Top 50 items

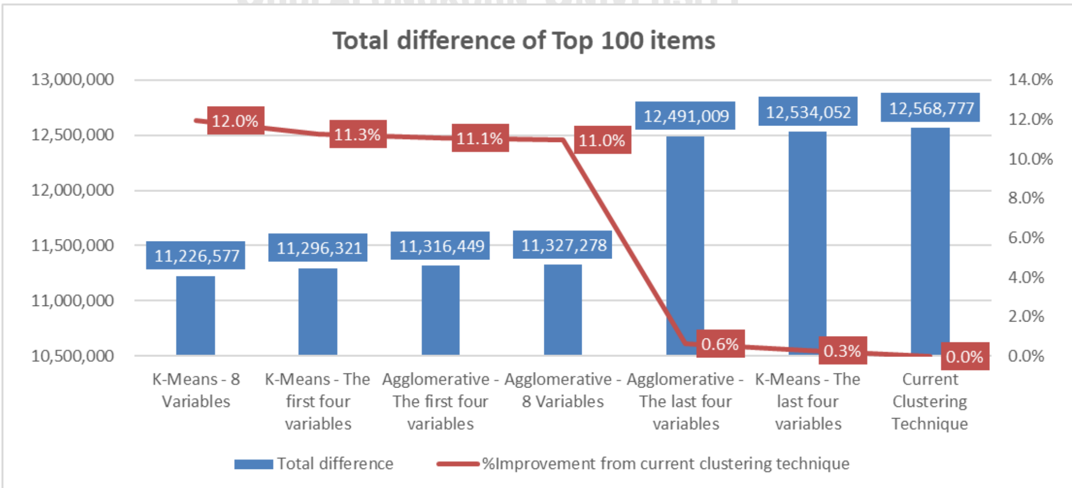

Figure 5-6: The total difference and the percentage improvement of Top 100 items

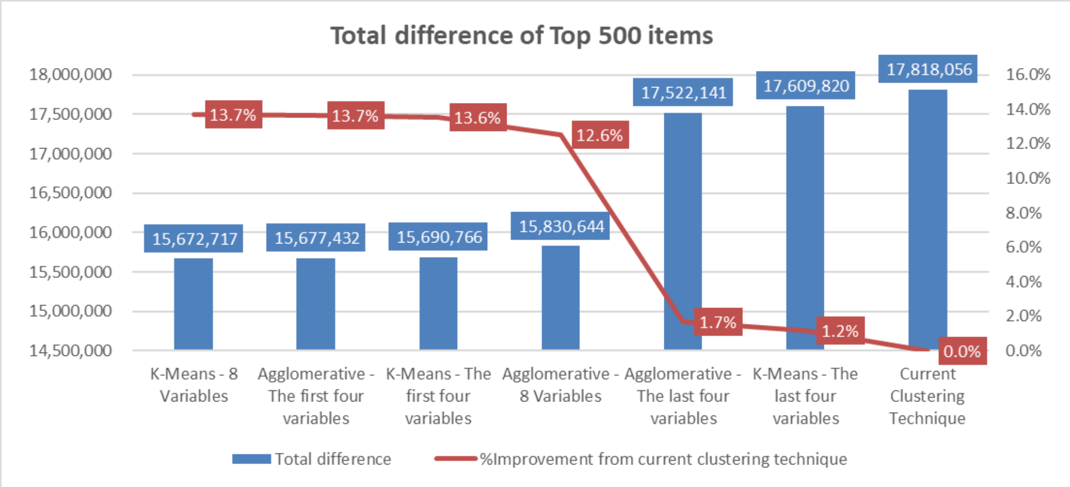

Figure 5-7: The total difference and the percentage improvement of Top 500 items

At the end, all the results between each clustering technique are ranked by the least total difference between allocated stock target by store cluster and actual stock target per store. From Table 5-29 to Table 5-30, it can be seen that the result of clustering the store by K-Means clustering technique with eight variables shows the least total difference between allocated stock target by store cluster and actual stock target per store on every group of items together with the highest percentage improvement from current clustering technique when develop a stock allocation plan. On the other hand, the overall worst performance of store clustering is the current clustering technique that company have been using.

	Total difference			
Clustering Technique	**Top 10 items**	**Top 50 items**	**Top 100 items**	**Top 500 items**
K-Means - 8 Variables	4,521,020	9,321,888	11,226,577	15,672,717
Agglomerative - The first four variables	4,648,002	9,449,249	11,316,449	15,677,432
K-Means - The first four variables	4,583,096	9,424,472	11,296,321	15,690,766
Agglomerative - 8 Variables	4,562,247	9,393,369	11,327,278	15,830,644
Agglomerative - The last four variables	5,059,770	10,316,142	12,491,009	17,522,141
K-Means - The last four variables	5,055,806	10,341,151	12,534,052	17,609,820
Current Clustering Technique	5,036,849	10,355,567	12,568,777	17,818,056

Table 5-29: The result of total difference between allocated stock target by store cluster and actual stock per store from all clustering techniques

	The percentage improvement from current clustering technique			
Clustering Technique	**Top 10 items**	**Top 50 items**	**Top 100 items**	**Top 500 items**
K-Means - 8 Variables	11.41%	11.09%	11.96%	13.69%
Agglomerative - The first four variables	8.37%	9.59%	11.07%	13.65%
K-Means - The first four variables	9.90%	9.88%	11.26%	13.56%
Agglomerative - 8 Variables	10.40%	10.24%	10.96%	12.55%
Agglomerative - The last four variables	-0.45%	0.38%	0.62%	1.69%
K-Means - The last four variables	-0.37%	0.14%	0.28%	1.18%
Current Clustering Technique	0.0%	0.0%	0.0%	0.0%

Table 5-30: The summary of percentage improvement of stock allocation plan from the current clustering technique

	Ranking			
Clustering Technique	**Top 10 items**	**Top 50 items**	**Top 100 items**	**Top 500 items**
K-Means - 8 Variables	1	1	1	1
Agglomerative - The first four variables	4	4	3	2
K-Means - The first four variables	3	3	2	3
Agglomerative - 8 Variables	2	2	4	4
Agglomerative - The last four variables	7	5	5	5
K-Means - The last four variables	6	6	6	6
Current Clustering Technique	5	7	7	7

Table 5-31: Ranking of each clustering technique from the least total difference between allocated stock target by store cluster and actual stock per store

5.1.3 Summary of the result from both performance measurement

To summarise the result from two performance measurements of the store cluster analysis including the coefficient of variation and the total difference between allocated stock target by store cluster and actual stock per store. For the first measurement, the overall coefficient of variation shows that both clustering technique that have performed in the SPSS program including K-Means clustering technique and Agglomerative clustering technique have a significant better value in terms of the coefficient of variation when compared with the current clustering technique that company have been using.

On the other hand, the second measurement indicates that clustering the store with eight variables by using K-Means clustering technique will give the minimum total difference between allocated stock target by store cluster and actual stock per store. Therefore, according to the result of both performance measurements. The recommendation of the store clustering technique is to use K-Means clustering technique with eight variables to perform the store cluster analysis when developing a stock allocation plan for the future event.

5.2 Other interpretation from the result of store cluster analysis

Since the performance measurement indicates that clustering the store by using K-means clustering technique with eight variables recorded the most efficient performance in terms of the total difference between allocated stock target by store cluster and actual stock per store as evaluated in 5.1.2. Therefore, the result of store cluster analysis of this technique will be used to further analyse and interpret the result of store cluster analysis in this case. To begin, if considering the result of store cluster analysis in terms of the geographical aspects. In this case, the regions of Thailand can be divided into six group including Bangkok (The capital city of Thailand) and vicinities, Central region, Northeast region, East region, North region and South region. Furthermore, the result of store cluster analysis by using K-Means clustering technique with eight variables is summarised and counted the store by each region of Thailand as presented in Table 5-32 together with the average of each variables by store cluster as in Table 5-33. What is more, the result of store cluster analysis is also highlighted by province of Thailand by each store cluster as shown in Figure 5-8 to Figure 5-11.

Overall, it can be seen that the majority of the store in store cluster number one came from Bangkok and vicinities area where these stores have the overall highest performance in terms of sales volume among the other three store clusters. On the other hand, store cluster number 2, in this cluster there is 70 percent of total store in this cluster came from northeast region of Thailand which this cluster recorded the number two ranked in terms of sales volume while it is the cluster that has the overall highest sales uplift from all of four store clusters. Moreover, for store cluster number three which can be considered as the store that has a medium performance in terms of both sales volume and sales uplift, the majority of the store in this cluster came from central region of Thailand together with Bangkok and vicinities. Lastly, store cluster number four which has the least in terms of sales volume performance when compared with the other three store cluster. What is more, in this store cluster the majority of the store came from northeast region which it contributes of more than 50 percent from the total store in this store cluster.

The study of the result of store cluster analysis by geographical aspect helps project manager be able to identify the area to focus and highlight when developing the stock allocation plan for any event. In this case as highlighted and summarised above, the store that contributes high sales volume during New Year festival generally located in Bangkok and vicinities area which came from store cluster number one. Therefore, it can be expected that these stores from store cluster number one will get higher stock volume compared to the other store when developing the stock allocation plan. On the other hand, store cluster number four where the majority of the stores located in northeast region of Thailand. These stores can be expected to have lower stock target than the other three store clusters since there have the least performance in terms of sales volume. What is more, it can be seen that stores located in northeast region have higher sales uplift since these stores mainly are in cluster number two and four which are the first two highest sales uplift store cluster. At the end, from this interpretation of the result of store cluster analysis, the project manager can also be able to coordinate and feedback to store operation and DC operation to prepare and plan the capacity and resource beforehand according to the result of store cluster analysis which will help each team to plan and utilise their resource at the right place and right decision and be able to prevent and eliminate the unexpected obstacle during the event.

	Count of store by region				
Region	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Total**
BKK & Vicinities	25		17		42
Central	4	1	20	12	37
Northeast	4	7	12	35	58
East	5		7	2	14
North	3	2	8	6	19
South	3		11	12	26
Total	**44**	**10**	**75**	**67**	**196**

Table 5-32: Count of store by region from the result of store cluster analysis

Cluster no	Count of Store	Average of Variable 1	Average of Variable 2	Average of Variable 3	Average of Variable 4	Average of Variable 5	Average of Variable 6	Average of Variable 7	Average of Variable 8
Cluster no. 1	44	377,242	399,594	279,580	321,986	46.3%	54.8%	44.5%	66.0%
Cluster no. 2	10	240,375	268,571	193,432	233,597	148.7%	172.4%	155.8%	202.0%
Cluster no. 3	75	206,353	214,054	152,729	170,823	52.8%	59.3%	51.1%	69.7%
Cluster no. 4	67	171,452	173,982	132,733	144,980	103.0%	106.0%	104.0%	122.7%

Table 5-33: An average of each variables by store cluster of K-Means Clustering technique with eight variables

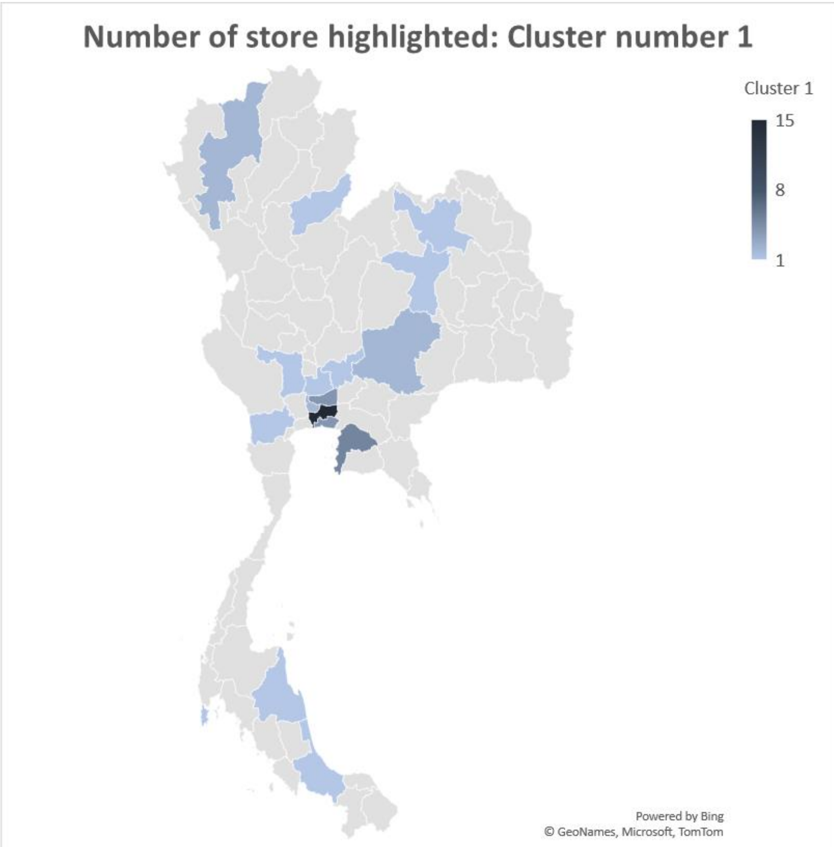

Figure 5-8: Number of store highlighted by province of Cluster number 1

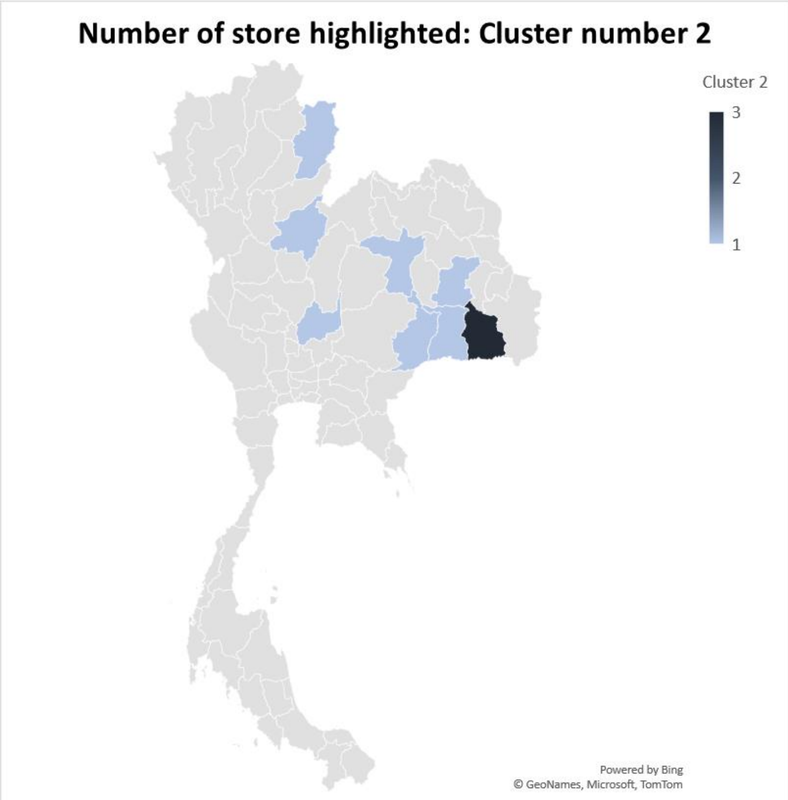

Figure 5-9: Number of store highlighted by province of Cluster number 2

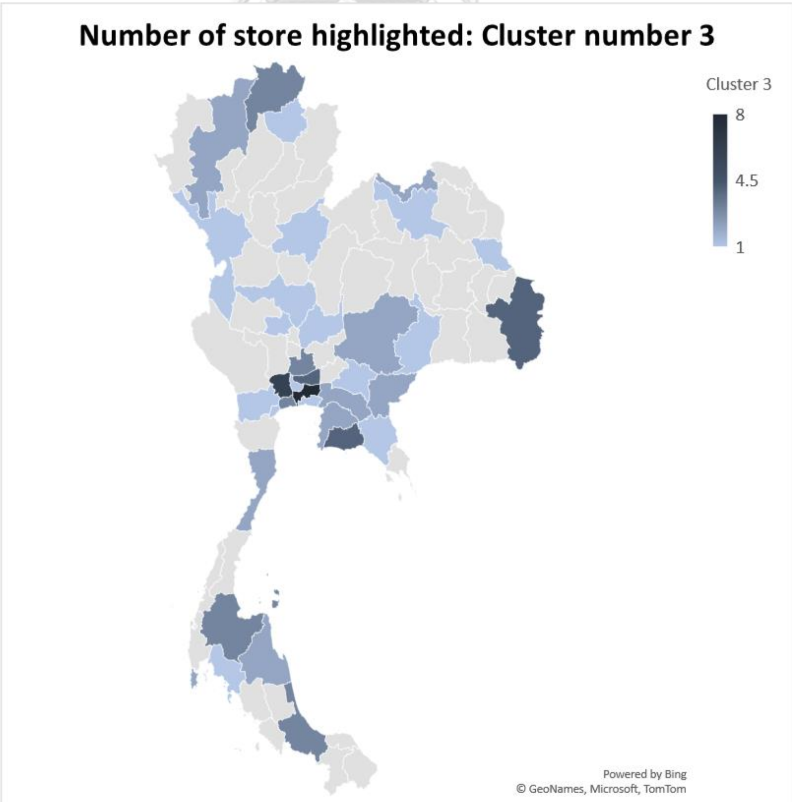

Figure 5-10: Number of store highlighted by province of Cluster number 3

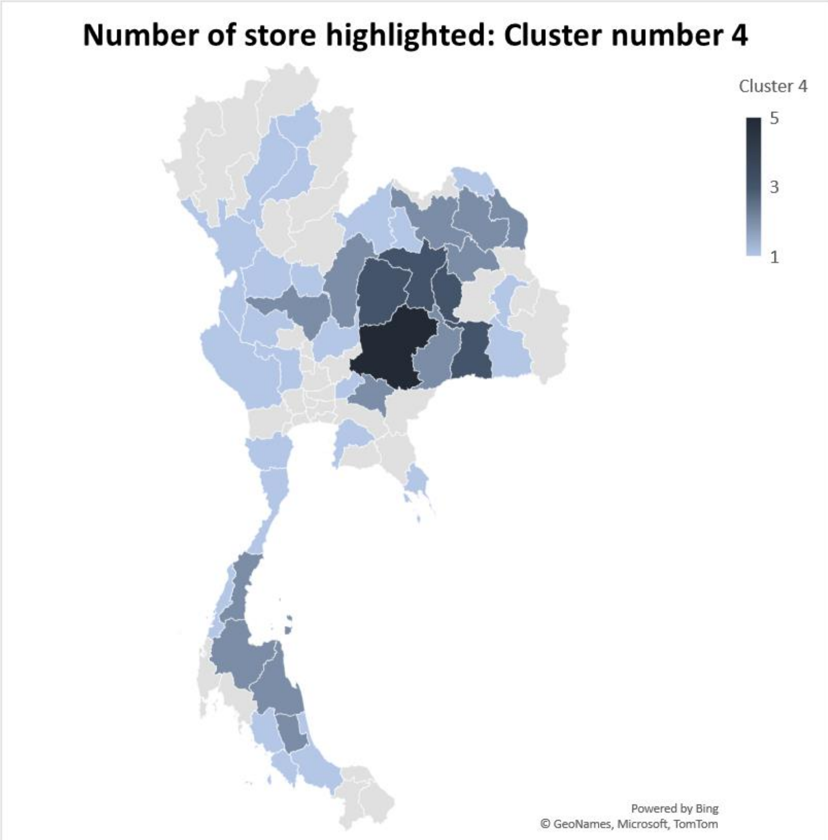

Figure 5-11: Number of store highlighted by province of Cluster number 4

5.3 The possible risks associated with store cluster analysis

There are several points to consider about the risk associated with the store cluster analysis. Firstly, as sales trend can be changed over time, the analysis of store cluster needed to be analysed and updated every time when developing the stock allocation plan. The use of past store cluster analysis may not reflect the current sales performance of each store which it may results in poor and inefficient of stock allocation plan. Thus, the refresh of store cluster analysis will help to capture the current sales trend or sales performance of the store and it also help to maximise the efficiency of stock allocation plan.

What is more, another point to consider is about the new store. Since the analysis is using the historical data such as sales volume and sales uplift in the past to perform the store cluster analysis. In this case, new store that has just opened will not have the data in the past available to perform the store cluster analysis. To illustrate this issue, if the analyst generates the data during the period of New Year 2020 to perform the store cluster analysis while there was a new store opened in February of 2020. Therefore, the data during the period of New Year 2020 of this new store is not available as this store still not open yet during that time. To overcome this issue, the analyst can be estimated the sales performance of this store based on the available sales performance data during the normal period and compensate with the overall percentage of sales uplift during the New Year 2020 event of hypermarket store format so that this new store can be included for the store cluster analysis

6. CONCLUSION

6.1 Scope and objective of the research

As there is a peak of demand during holiday or seasonal event such as New Year festival or Songkran festival (Thai New Year Festival). One major obstacle of the organisation during the seasonal period or holiday event is to handle a huge number of stock volume flowing between both inbound, from suppliers deliver to the distribution centres, and outbound, from distribution centres deliver to the store. Project manager from Tesco Lotus's supply chain team has the responsibility to plan, manage and coordinate all the related tasks and preparation of stock ordering with both internal supply chain team and external team including store operation team and distribution centres team to ensure the stock availability at store together with achieve the optimised level of stock before the event is started. The project manager also needed to identify the solution to overcome any limitation of the organisation during the event or seasonal period such as manpower, limited distribution centre's capacity and store space utilisation. Therefore, the stock allocation plan is developed in order to smoothen the incoming stock to the distribution centres and stores. Overall, the stock allocation plan consists of three stages including key item selection, store cluster analysis and stock allocation plan.

On the other hand, one factor that determine the efficiency of stock allocation plan is the store cluster analysis. Since there are lots of combination between stores and items to consider when developing the plan. The store cluster analysis is help classifying and grouping the store that has a similar in terms of sales performance together which it helps reduce the level of difficulty when developing the stock allocation plan. Therefore, developing the stock allocation plan by item and store cluster level is much less combination than the store and item level. However, different clustering technique has different approach and algorithm to clustering the store. The current clustering technique that company have been using is to rank the overall average of sales performances during the event by store level including sales volume variables and sales uplift variables which it consists of eight variables as listed in the fourth chapter. Then, these stores will be split equally into the designed store cluster which usually cluster into four store clusters as it is an acceptable level to develop and manage the stock allocation plan for the project manager.

What is more, two clustering techniques are introduced in this research including K-Means clustering technique and Agglomerative clustering technique. The concept and algorithm of K-Means clustering technique is to combine the data point together based on the minimum distance between the data point and centroids until all the data point are grouped into the desired number of cluster. On the other hand, the algorithm of the Agglomerative clustering technique is the bottom-up approach where each individual data point is considered as one cluster before grouping the data into one final cluster based on the distance between each data point.

6.2 Analysis and result

To summarise the analysis part, after generated all the data. Since there is a difference in scale of the data between each variable, these data will be normalised by using z-score technique After that, the store cluster analysis is performed in SPSS program by using two introduced clustering techniques including K-Means clustering technique and Agglomerative clustering technique. The variables are split into three group when performing store cluster analysis including Group 1 that has clustered the store with eight variables, Group 2 that have clustered the store with the first four variables, Group 3 that have clustered the store with the last four variables. Moreover, the current clustering technique that company have been using will also be performed the store cluster analysis in order to evaluate and compare the result of store cluster analysis with those two clustering techniques. To summarise, there is seven set of store cluster analysis to perform which can be summarised as below Table 6-1.

No.	Group	Clustering Technique	Variables
1	1	**K-Means Clustering**	All variables (8 variables)
2		**Agglomerative Clustering**	All variables (8 variables)
3	2	**K-Means Clustering**	The first four variables
4		**Agglomerative Clustering**	The first four variables
5	3	**K-Means Clustering**	The last four variables
6		**Agglomerative Clustering**	The last four variables
7	**Current Clustering Technique**		All variables (8 variables)

Table 6-1: The summary of all type of store cluster analysis that performed in this research

After performing the store cluster analysis, the result of store cluster analysis between each clustering technique above is measured and compared. In this case, the result of store clustering between each clustering technique can be measured by two performance measurements including the coefficient of variation as the first performance measurement. Another performance measurement is to develop the stock allocation plan by using the result of store cluster analysis that had performed with K-Means clustering technique, Agglomerative clustering technique and the current clustering technique that company have been using. Then, compare the total difference between allocated stock target by store cluster and actual stock per store by each clustering technique. The result of these two performance measurements shows that two introduced clustering technique including K-Means clustering technique and Agglomerative clustering technique have a significant less coefficient of variation when compared with the current clustering technique that company have been using. While, K-Means clustering technique give the least total coefficient of variation. On the other hand, clustering the store by using K-Means clustering technique with eight variables recorded the minimum result of the total difference between allocated stock target by store cluster and actual stock per store. Therefore, it is recommended to use K-Means clustering technique with eight variables to perform the store cluster analysis when developing a stock allocation plan in the future.

6.3 The benefits bring to the organisation

From the study of the concept of cluster analysis together with the implementation of different clustering techniques in this research, there are a number of benefits bring to the organisation which can be summarised the benefits as follow. To begin, the first benefit, the performance of store cluster analysis is improved together with enhance the efficiency of stock allocation plan. Since the implementation of introduced clustering technique give a significant better result than the current clustering technique that company have been using as highlighted in the previous chapter. Moreover, the second benefit brings to organisation is to be able to achieve better store availability and optimise level of stock during seasonal event. As the performance measurement of the store cluster analysis indicates that the recommended clustering technique give a significant less total difference between allocated stock target by store cluster and actual stock target per store. Moving to the third benefit brings to the organisation. Since the efficiency of the store cluster analysis is improved. It leads to enhance the quality of planning and coordinating the related work with other team such as Store operation team and DC operation team as those teams will use the feedback of stock allocation plan developed by Supply Chain team to further manage and plan their capacity and resource. Lastly, the opportunity to further develop and implement the concept of cluster analysis with other business aspects in the future which will be outlined in the following section 6.4.

6.4 The future implementation of the analysis and other constraints

6.4.1 The implementation of store cluster analysis with the other store format

Even though, the selected store format to perform the store cluster analysis in this research is performed and analysed solely on hypermarket store format as it is the store format that contributes the most sales participation to the overall company's sales, as shown in Figure 4-1 in the fourth chapter. However, the recommended clustering technique is also applicable and can be implemented with the other store formats as well including supermarket format or convenient store format. Since the data needed for the store cluster analysis will be using the variables as same as when performing the store cluster analysis of hypermarket store format which it includes eight variables such as sales volume and sales uplift as listed in section 4.1.

6.4.2 The implementation of the store cluster analysis and the recommended clustering technique with other work

The concept of store cluster analysis and recommended clustering technique can be used and implemented with other work as well. In this case, the store cluster analysis can be implemented not only for the stock allocation plan during holiday or seasonal event as analysed and outlined in this research, but it can also be classified or grouped the store in terms of the overall sales performance during the specific period and specific product category. For example, the position namely category manager in supply chain team, who responsible and manage the order and stock level specifically in the dry grocery non-food product group such as Toiletries product and Household chemical product, can use the store cluster analysis to identify and cluster the hypermarket store format based on sales performance of the dry grocery non-food product group during last three months so that the category manager can be able to capture the current performance of the store and can be identified which group of store is the potential store group needed to focus which has good sales performance in dry grocery non-food product and etc.

On the other hand, the recommended clustering technique can also be implemented with other aspect as well apart from store cluster analysis. For example, the analysis of item clustering in order to identify and classify the similar characteristic of item based on sales volume performance and sales uplift during holiday event and seasonal period and etc. At the end, the analyst will get a group of items to focus during the specific event in order to prepare and plan the order and manage the stock level in the later stage. Therefore, it can be seen that the concept of store cluster analysis together with the recommended clustering technique can be adapted and implemented with other works as well which it helps analyst to classify or group the data that has the similar pattern or characteristic together so that the analyst can identify and be able to develop the business plan and business strategy according to the analysis.

6.4.3 Software requirement and execution time constraint

In terms of the implementation of store cluster analysis with the recommended clustering technique which is K-Means clustering technique that has performed the store cluster analysis in the SPSS program. In this case, the organisation has a license of the SPSS program which can be requested and installed on the user's laptop that owned by the organisation. Therefore, there is no concern on the license of the SPSS program to use as a tool when perform the store cluster analysis. On the other hand, in terms of the time constraint. The SPSS program is executed and performed the store cluster analysis promptly and efficiently. The overall execution time is less than five minutes based on the store cluster analysis that researcher had performed in this research. Thus, time constraint is also not an issue to work in this case when performing the store cluster analysis in the SPSS program.

REFERENCES

[1] Andreopoulou Z. et al. (2017), Multicriteria and Clustering: Classification Techniques in Agrifood and Environment, Springer

[2] Celebi Emre M. and Hassan A. (2015), Linear, Deterministic and Order-Invariant Initialization Methods for the K-Means Clustering Algorithms, Springer, pp.79-86.

[3] Charu C. and Chandan K. (2014), Data Clustering Algorithms and Applications: Distance-Based Algorithms, Taylor & Francis Group.

[4] Dardac N. and Boitan I. (2009), A Cluster Analysis Approach For Bank's Risk Profile: The Romanian Evidence, European Research Studies, 12(1), pp. 109-116.

[5] Eessaar E. (2016), The Database Normalization Theory and The Theory of Normalized Systems: Finding A Common Ground, Baltic J. Modern Computing, Tallinn University of Technology, Estonia, pp. 5-33.

[6] Fahim A. M. et al. (2006), An Efficient Enhanced K-Means Clustering Algorithm, Journal of Zhejiang University SCIENCE A, 7(10), pp. 1626-1633.

[7] Frei F. (2013), Cluster Analysis for Classification of Patients Submitted to Fobi-Capella Bariatric Surgery, ABCD Arg Bras Cir Dig, pp.33-38.

[8] Herv´e A. (2010), Coefficient of Variation, The University of Texas at Dallas, Richardson, pp. 1-5

[9] Koch R. (1998), The 80/20 Principle: The Secret of Achieving More with Less, Nicholas Brealey

[10] Leavengood S. and Reeb J. (2002), Performance Excellence in The Wood Industry: Statistical Process Control, Part3: Pareto Analysis and Check Sheets, Oregon State University, The United States, pp. 1-8.

[11] Marcisz M. (2013), Practical Application of Coefficient of Variation, Faculty of Mining and Geology, Institute of Applied Geology, Silesian University of Technology, Poland

[12] Marr B. (2015), Big Data: A Game Changer in The Retail Sector Available at: https://www.forbes.com/sites/bernardmarr/2015/11/10/big-data-a-game-changer-in-the-retail-sector/#6bc6adfd9f37 [Accessed on 15 February 2020]

[13] Morissette L. and Chartier S. (2013), The K-Means Clustering Technique: General Considerations and Implementation in Mathematica, University of Ottawa, 9(1), pp. 15-24.

[14] Murtagh F. and Contreras P. (2011), Methods of Hierarchical Clustering, University of Hudderfield and University of Warwick, pp.1-16

[15] Neeti A. and Motwani M. (2014), A Distance Based Clustering Algorithm, IJCET, 5(5), pp.109-119.

[16] Nielson F. (2016), Introduction of HPC with MPI for Data Science: Hierarchical Clustering, Springer.

[17] Oyelade, O. J, Oladipupo, O. O and Obagbuwa, I. C. (2010), Application of K-Means Clustering Algorithm for Prediction of Students' Academic Performance, IJCSIS, 7(1), pp. 292-295

[18] Raval R. and Chaita J. (2016), Implementing & Improvisation of K-means Clustering Algorithm, IJCSMC, 5(5), pp.191-203.

[19] Santini M. (2016), Advantages & Disadvantages of K-means and Hierarchical Clustering, Uppsala University, Sweden.

[20] Scoltock J. (1982), A Survey of the Literature of Cluster Analysis, The Computer Journal, Heyden & Son, 25(1), pp.130-134

[21] Sławomir T. and Mieczysław A. (2018), Modern Algorithms of Cluster Analysis: Hierarchical Method of Cluster Analysis, Springer

[22] Soni M. T. (2012), An Overview on Clustering Methods: Partitional Clustering, IOSR Journal of Engineering, 2(4), pp.719-725.

[23] Stojčetović B. and Šarkoćević Z. (2015), Application of The Pareto Analysis in Project Management, High technical school of professional education, Zvečan, Kosovo, pp. 655-657

[24] SPSS Software IBM (2019), SPSS Statistics
Available at: https://www.ibm.com/analytics/spss-statistics-software
[Accessed on 17 November 2019]

[25] Tapas K. and IEE. (2002), An Efficient K-Means Clustering Algorithm: Analysis and Implementation, IEEE, pp.881-892.

[26] Tesco Lotus (2019), Tesco Lotus: Our Business
Available at: https://www.tescolotus.com/en/about/our-business
[Accessed on 13 October 2019]

[27] Tesco Lotus Online Shopping (2019), Tesco Lotus Shop Online: About Tesco Lotus Shop Online
Available at: https://www.tescolotus.com/shoponline/en/help-about-shoponline.php [Accessed on 13 October 2019]

[28] Yuan C. and Yang H. (2019), Research on K-Value Selection Method of K-Means Clustering Algorithm, J Multidisciplinary Scientific Journal, pp.226-235.

APPENDICES

Appendix A: The generated data by store level

Table 7-1 to Table 7-4 are the generated data by store level to use for store cluster analysis in the fourth chapter which consisting of eight variables as presented below.

Variable 1: Sales Volume by store during New Year 2018 (Overall total sales)

Variable 2: Sales Volume by store during New Year 2019 (Overall total sales)

Variable 3: Sales Volume by store during New Year 2018 (of Selected Key Items)

Variable 4: Sales Volume by store during New Year 2019 (of Selected Key Items)

Variable 5: %Sales Uplift of New Year 2018 by store (Overall total sales)

Variable 6: %Sales Uplift of New Year 2019 by store (Overall total sales)

Variable 7: %Sales Uplift of New Year 2018 by store (of Selected Key Items)

Variable 8: %Sales Uplift of New Year 2019 by store (of Selected Key Items)

Store No.	Raw data before normalising the data								Normalised Data							
	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
No. 001	486,295	534,666	349,337	427,682	41.6%	55.7%	39.4%	70.6%	2.66	2.83	2.47	2.73	-0.80	-0.62	-0.78	-0.50
No. 002	170,648	164,053	132,087	137,632	61.1%	54.9%	57.4%	64.0%	-0.68	-0.79	-0.63	-0.73	-0.31	-0.64	-0.36	-0.65
No. 003	266,758	243,518	202,004	197,489	102.5%	84.8%	102.1%	97.5%	0.34	-0.01	0.36	-0.02	0.73	0.12	0.67	0.08
No. 004	377,484	368,003	244,942	255,891	48.6%	44.9%	46.3%	52.8%	1.51	1.20	0.98	0.68	-0.62	-0.90	-0.62	-0.89
No. 005	509,864	582,251	383,630	478,711	61.6%	84.5%	58.0%	97.2%	2.91	3.29	2.96	3.34	-0.30	0.11	-0.35	0.08
No. 006	274,639	374,260	211,387	329,790	94.3%	164.8%	91.5%	198.7%	0.42	1.26	0.50	1.56	0.53	2.17	0.43	2.29
No. 007	414,662	400,861	323,556	336,181	118.6%	111.3%	128.2%	137.1%	1.90	1.52	2.10	1.64	1.14	0.80	1.28	0.95
No. 008	294,054	321,329	224,522	268,446	75.5%	91.7%	76.3%	110.8%	0.63	0.75	0.69	0.83	0.05	0.30	0.08	0.37
No. 009	412,715	445,504	315,680	369,698	25.5%	35.5%	21.7%	42.6%	1.88	1.96	1.99	2.04	-1.21	-1.14	-1.19	-1.12
No. 010	264,711	256,637	202,969	208,862	89.2%	83.4%	87.6%	93.0%	0.32	0.12	0.38	0.12	0.40	0.09	0.34	-0.02
No. 011	300,732	321,062	221,585	258,404	41.1%	50.7%	36.1%	58.7%	0.70	0.74	0.64	0.71	-0.81	-0.75	-0.86	-0.76
No. 012	394,229	431,123	290,290	342,713	19.4%	30.5%	16.6%	37.7%	1.69	1.82	1.62	1.72	-1.36	-1.27	-1.31	-1.22
No. 013	236,536	238,647	174,672	184,209	70.5%	72.0%	71.4%	80.7%	0.02	-0.06	-0.03	-0.18	-0.07	-0.20	-0.04	-0.28
No. 014	449,501	447,161	335,021	353,103	72.9%	72.0%	71.3%	80.6%	2.27	1.97	2.26	1.84	-0.01	-0.21	-0.04	-0.29
No. 015	237,172	222,544	152,288	161,629	32.7%	24.5%	25.7%	33.4%	0.03	-0.22	-0.34	-0.45	-1.03	-1.42	-1.10	-1.32
No. 016	506,633	520,197	376,042	416,758	28.6%	32.0%	24.0%	37.4%	2.88	2.68	2.85	2.60	-1.13	-1.23	-1.14	-1.23
No. 017	369,488	375,487	248,596	285,497	34.1%	36.3%	23.7%	42.0%	1.43	1.27	1.03	1.03	-0.99	-1.12	-1.15	-1.13
No. 018	401,858	394,315	294,177	308,126	22.8%	20.5%	18.4%	24.0%	1.77	1.46	1.68	1.30	-1.28	-1.52	-1.27	-1.52
No. 019	408,796	376,054	314,584	308,504	54.6%	42.2%	55.3%	52.3%	1.84	1.28	1.97	1.31	-0.47	-0.97	-0.41	-0.90
No. 020	396,453	406,761	291,857	329,613	19.6%	22.7%	12.0%	26.5%	1.71	1.58	1.65	1.56	-1.36	-1.46	-1.42	-1.47
No. 021	321,597	363,925	233,527	298,955	20.1%	35.9%	11.7%	43.1%	0.92	1.16	0.81	1.19	-1.35	-1.13	-1.42	-1.11
No. 022	417,847	469,494	269,318	340,419	18.5%	33.1%	17.0%	47.9%	1.94	2.19	1.32	1.69	-1.38	-1.20	-1.30	-1.00
No. 023	305,587	326,305	220,159	261,026	26.7%	35.3%	19.4%	41.6%	0.75	0.79	0.62	0.74	-1.18	-1.14	-1.25	-1.14
No. 024	330,181	334,148	254,169	277,733	64.8%	66.8%	60.6%	75.5%	1.01	0.87	1.11	0.94	-0.22	-0.34	-0.29	-0.40
No. 025	401,992	424,402	296,095	342,915	53.7%	62.3%	49.5%	73.2%	1.77	1.75	1.71	1.72	-0.50	-0.45	-0.55	-0.45
No. 026	269,607	279,667	202,067	225,080	50.9%	56.5%	51.4%	68.6%	0.37	0.34	0.37	0.31	-0.57	-0.60	-0.50	-0.55
No. 027	470,624	421,535	352,054	332,683	52.4%	36.5%	51.0%	42.7%	2.50	1.72	2.50	1.60	-0.53	-1.11	-0.51	-1.11
No. 028	446,183	446,435	329,824	354,015	32.9%	33.0%	32.7%	42.4%	2.24	1.97	2.19	1.85	-1.02	-1.20	-0.94	-1.12
No. 029	321,670	296,111	242,317	235,031	54.0%	41.7%	51.7%	47.1%	0.92	0.50	0.94	0.43	-0.49	-0.98	-0.50	-1.02
No. 030	362,642	369,047	262,610	286,155	30.6%	32.9%	27.4%	38.8%	1.35	1.21	1.23	1.04	-1.08	-1.21	-1.06	-1.20
No. 031	386,447	591,557	286,556	507,623	41.9%	117.2%	39.4%	146.9%	1.61	3.38	1.57	3.69	-0.79	0.95	-0.78	1.16
No. 032	361,982	339,354	268,709	266,273	27.2%	19.3%	22.7%	21.6%	1.35	0.92	1.32	0.80	-1.16	-1.55	-1.17	-1.57
No. 033	271,736	249,323	181,097	183,146	47.1%	34.9%	35.3%	36.9%	0.39	0.04	0.07	-0.19	-0.66	-1.15	-0.88	-1.24
No. 034	345,668	361,249	251,086	287,450	39.4%	45.7%	36.3%	56.0%	1.18	1.14	1.06	1.06	-0.86	-0.88	-0.85	-0.82
No. 035	245,624	271,783	187,284	232,732	92.2%	112.6%	89.1%	135.0%	0.12	0.26	0.15	0.40	0.47	0.83	0.37	0.90
No. 036	448,177	293,710	358,581	230,865	100.6%	31.4%	117.6%	40.1%	2.26	0.48	2.60	0.38	0.69	-1.24	1.03	-1.17
No. 037	337,478	437,635	250,691	367,146	36.2%	76.6%	32.1%	93.5%	1.09	1.88	1.06	2.01	-0.94	-0.09	-0.95	-0.01
No. 038	384,627	517,170	278,946	436,519	15.1%	54.7%	6.1%	66.1%	1.59	2.65	1.46	2.84	-1.47	-0.65	-1.55	-0.60
No. 039	199,657	168,114	140,039	123,194	68.9%	42.3%	71.8%	51.2%	-0.37	-0.75	-0.52	-0.91	-0.11	-0.97	-0.03	-0.93
No. 040	204,132	180,105	142,642	137,369	24.6%	10.0%	17.0%	12.7%	-0.32	-0.63	-0.48	-0.74	-1.23	-1.79	-1.30	-1.77
No. 041	194,175	235,345	121,451	167,955	35.4%	64.2%	26.9%	75.4%	-0.43	-0.09	-0.78	-0.37	-0.96	-0.41	-1.07	-0.40
No. 042	206,589	203,442	155,592	169,315	98.4%	95.4%	97.6%	115.1%	-0.30	-0.40	-0.30	-0.36	0.63	0.39	0.57	0.47
No. 043	277,912	219,661	213,713	172,808	78.6%	41.2%	85.9%	50.3%	0.46	-0.24	0.53	-0.31	0.13	-0.99	0.30	-0.95
No. 044	282,634	289,400	207,076	233,148	57.6%	61.4%	56.9%	76.7%	0.51	0.43	0.44	0.41	-0.40	-0.48	-0.37	-0.37
No. 045	517,800	522,972	384,875	422,393	55.1%	56.7%	53.4%	68.4%	2.99	2.71	2.97	2.67	-0.46	-0.60	-0.46	-0.55
No. 046	246,761	249,617	171,223	187,847	51.4%	53.2%	49.4%	64.0%	0.13	0.05	-0.07	-0.13	-0.55	-0.69	-0.55	-0.65
No. 047	298,938	361,907	236,969	310,227	63.8%	98.4%	73.2%	126.7%	0.68	1.14	0.86	1.33	-0.24	0.47	0.00	0.72
No. 048	300,928	235,372	232,646	181,837	62.7%	27.2%	71.2%	33.8%	0.70	-0.09	0.80	-0.21	-0.27	-1.35	-0.04	-1.31
No. 049	383,668	367,645	289,281	282,423	32.2%	26.7%	36.3%	33.0%	1.58	1.20	1.61	1.00	-1.04	-1.36	-0.85	-1.32

Store No.	Raw data before normalising the data								Normalised Data							
	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
No. 050	236,968	370,840	169,043	316,299	16.2%	81.9%	8.3%	102.6%	0.03	1.23	-0.11	1.40	-1.44	0.05	-1.50	0.19
No. 051	194,147	216,311	134,106	164,322	6.9%	19.1%	0.8%	23.5%	-0.43	-0.28	-0.60	-0.42	-1.68	-1.56	-1.68	-1.53
No. 052	419,955	493,716	325,294	419,350	86.5%	119.2%	95.3%	151.7%	1.96	2.43	2.12	2.63	0.33	1.00	0.52	1.27
No. 053	225,223	194,284	171,281	158,334	97.2%	70.1%	97.2%	82.3%	-0.10	-0.49	-0.07	-0.49	0.60	-0.25	0.56	-0.25
No. 054	365,371	398,618	275,698	320,248	47.7%	61.1%	51.1%	75.5%	1.38	1.50	1.42	1.45	-0.65	-0.48	-0.51	-0.40
No. 055	221,095	274,442	140,081	199,756	32.4%	64.3%	33.8%	90.8%	-0.14	0.29	-0.52	0.01	-1.03	-0.40	-0.91	-0.06
No. 056	249,468	249,775	187,193	200,962	32.8%	33.0%	30.1%	39.7%	0.16	0.05	0.15	0.02	-1.02	-1.20	-1.00	-1.18
No. 057	137,025	111,873	103,671	96,132	130.3%	88.0%	116.4%	100.7%	-1.03	-1.30	-1.04	-1.23	1.44	0.20	1.01	0.15
No. 058	332,403	378,306	256,691	317,808	62.0%	84.4%	62.2%	100.8%	1.03	1.30	1.14	1.42	-0.29	0.11	-0.25	0.16
No. 059	206,499	197,515	160,349	161,032	57.5%	50.6%	60.0%	60.7%	-0.30	-0.46	-0.23	-0.46	-0.40	-0.75	-0.30	-0.72
No. 060	262,147	298,369	187,477	234,555	43.4%	63.2%	37.5%	72.0%	0.29	0.52	0.16	0.42	-0.76	-0.43	-0.83	-0.47
No. 061	277,263	224,082	223,392	179,586	56.9%	26.8%	63.3%	31.3%	0.45	-0.20	0.67	-0.23	-0.42	-1.36	-0.23	-1.36
No. 062	217,719	216,793	160,531	173,311	77.4%	76.6%	75.3%	89.3%	-0.18	-0.27	-0.23	-0.31	0.10	-0.09	0.05	-0.10
No. 063	244,653	208,204	193,102	169,909	88.0%	60.0%	91.6%	68.6%	0.11	-0.36	0.24	-0.35	0.37	-0.51	0.43	-0.55
No. 064	179,533	178,000	126,890	131,921	61.7%	60.4%	62.5%	68.9%	-0.58	-0.65	-0.71	-0.80	-0.29	-0.50	-0.25	-0.54
No. 065	288,492	324,398	215,636	264,143	23.0%	38.3%	17.4%	43.9%	0.57	0.78	0.56	0.78	-1.27	-1.07	-1.29	-1.09
No. 066	409,038	445,707	303,261	354,105	50.0%	63.5%	48.8%	73.7%	1.85	1.96	1.81	1.85	-0.59	-0.42	-0.56	-0.44
No. 067	184,127	195,935	139,833	157,637	33.7%	42.2%	27.3%	43.5%	-0.53	-0.48	-0.52	-0.50	-1.00	-0.97	-1.06	-1.10
No. 068	206,812	204,551	142,936	151,968	36.6%	35.1%	37.7%	46.4%	-0.29	-0.39	-0.48	-0.56	-0.93	-1.15	-0.82	-1.03
No. 069	275,520	274,124	199,581	214,091	31.6%	30.9%	29.1%	38.5%	0.43	0.29	0.33	0.18	-1.05	-1.26	-1.02	-1.21
No. 070	176,324	167,788	122,460	126,929	48.0%	40.8%	44.4%	49.6%	-0.62	-0.75	-0.77	-0.86	-0.64	-1.00	-0.67	-0.96
No. 071	222,933	243,203	175,117	206,376	51.9%	65.7%	45.7%	71.7%	-0.12	-0.02	-0.02	0.09	-0.54	-0.37	-0.63	-0.48
No. 072	199,993	186,249	149,381	146,066	96.8%	83.3%	98.5%	94.1%	-0.37	-0.57	-0.39	-0.63	0.59	0.08	0.59	0.01
No. 073	288,670	265,147	227,884	231,056	223.8%	197.4%	231.8%	236.4%	0.57	0.20	0.73	0.38	3.80	3.00	3.68	3.11
No. 074	139,265	135,362	97,461	103,892	84.8%	79.6%	86.0%	98.2%	-1.01	-1.07	-1.13	-1.14	0.29	-0.01	0.30	0.10
No. 075	160,050	144,358	123,043	120,408	152.0%	127.3%	154.0%	148.5%	-0.79	-0.98	-0.76	-0.94	1.98	1.21	1.88	1.20
No. 076	123,036	102,266	87,949	76,965	143.2%	102.1%	148.9%	117.8%	-1.18	-1.39	-1.26	-1.46	1.76	0.56	1.76	0.53
No. 077	226,871	254,644	169,620	214,067	69.0%	89.7%	66.0%	109.5%	-0.08	0.10	-0.10	0.18	-0.11	0.25	-0.16	0.34
No. 078	212,709	249,810	158,014	207,820	64.9%	93.7%	62.5%	113.8%	-0.23	0.05	-0.26	0.10	-0.21	0.35	-0.24	0.44
No. 079	185,407	200,935	137,883	167,175	80.3%	95.4%	71.4%	107.8%	-0.52	-0.43	-0.55	-0.38	0.18	0.39	-0.04	0.31
No. 080	233,947	230,558	161,860	170,227	52.5%	50.2%	51.9%	59.8%	-0.01	-0.14	-0.21	-0.35	-0.53	-0.76	-0.49	-0.74
No. 081	263,544	284,294	193,858	229,410	29.4%	39.5%	24.0%	46.7%	0.31	0.39	0.25	0.36	-1.11	-1.04	-1.14	-1.03
No. 082	142,443	176,129	87,367	120,776	26.2%	56.1%	19.6%	65.3%	-0.97	-0.67	-1.27	-0.94	-1.19	-0.61	-1.24	-0.62
No. 083	360,559	340,772	258,107	263,757	73.3%	63.7%	66.0%	69.7%	1.33	0.94	1.16	0.77	-0.00	-0.42	-0.16	-0.52
No. 084	135,147	155,867	102,923	130,979	82.1%	110.0%	80.6%	129.8%	-1.05	-0.87	-1.05	-0.81	0.22	0.77	0.17	0.79
No. 085	90,913	133,036	71,160	117,342	58.3%	131.6%	58.6%	161.6%	-1.52	-1.09	-1.50	-0.98	-0.38	1.32	-0.33	1.48
No. 086	225,822	234,872	169,102	188,234	70.2%	77.0%	67.0%	85.8%	-0.09	-0.10	-0.10	-0.13	-0.08	-0.08	-0.14	-0.17
No. 087	290,540	333,592	218,148	271,569	17.5%	34.9%	16.4%	44.9%	0.59	0.87	0.59	0.87	-1.41	-1.15	-1.31	-1.06
No. 088	213,513	225,744	152,464	179,791	86.7%	97.4%	82.8%	115.6%	-0.22	-0.19	-0.34	-0.23	0.34	0.44	0.23	0.48
No. 089	130,596	110,881	95,821	85,187	147.4%	110.0%	145.2%	118.0%	-1.10	-1.31	-1.15	-1.36	1.87	0.77	1.67	0.53
No. 090	311,544	303,383	222,211	229,784	47.2%	43.3%	42.1%	47.0%	0.81	0.57	0.65	0.37	-0.66	-0.94	-0.72	-1.02
No. 091	229,555	325,673	180,988	282,950	100.9%	185.0%	108.3%	225.6%	-0.05	0.79	0.06	1.00	0.69	2.68	0.82	2.88
No. 092	187,119	151,704	148,204	122,188	147.5%	100.7%	157.5%	112.3%	-0.50	-0.91	-0.40	-0.92	1.87	0.53	1.96	0.41
No. 093	132,195	125,730	92,780	92,270	76.9%	68.3%	86.1%	85.0%	-1.08	-1.16	-1.19	-1.28	0.09	-0.30	0.30	-0.19
No. 094	200,312	259,970	149,745	223,211	55.1%	101.3%	48.8%	121.8%	-0.36	0.15	-0.38	0.29	-0.46	0.54	-0.56	0.61
No. 095	186,376	212,735	120,486	150,930	27.3%	45.3%	21.1%	51.7%	-0.51	-0.31	-0.80	-0.58	-1.16	-0.89	-1.21	-0.92
No. 096	172,480	155,287	132,174	124,924	116.8%	95.2%	119.7%	107.6%	-0.66	-0.87	-0.63	-0.89	1.10	0.39	1.08	0.30
No. 097	239,614	246,067	184,140	205,389	79.7%	84.5%	77.6%	98.0%	0.05	0.01	0.11	0.07	0.16	0.11	0.10	0.09
No. 098	303,354	345,153	221,384	279,110	46.2%	66.3%	46.5%	84.7%	0.73	0.98	0.64	0.96	-0.69	-0.35	-0.62	-0.20

Store No.	Raw data before normalising the data								Normalised Data							
	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
No. 099	225,612	168,269	182,616	140,604	145.4%	83.1%	152.8%	94.7%	-0.09	-0.75	0.09	-0.70	1.82	0.08	1.85	0.02
No. 100	227,942	212,462	176,349	174,962	92.9%	79.8%	90.4%	88.9%	-0.07	-0.32	-0.00	-0.29	0.49	-0.01	0.40	-0.11
No. 101	165,208	161,565	129,272	134,471	83.3%	79.3%	82.5%	89.8%	-0.73	-0.81	-0.67	-0.77	0.25	-0.02	0.22	-0.09
No. 102	323,180	297,545	243,517	234,542	51.2%	39.2%	49.8%	44.2%	0.94	0.51	0.96	0.42	-0.56	-1.04	-0.54	-1.08
No. 103	174,055	181,274	134,089	148,076	67.7%	74.6%	69.4%	87.0%	-0.64	-0.62	-0.60	-0.61	-0.14	-0.14	-0.09	-0.15
No. 104	278,202	267,165	206,349	212,353	40.7%	35.1%	40.0%	44.0%	0.46	0.22	0.43	0.16	-0.82	-1.15	-0.77	-1.08
No. 105	172,658	159,689	134,595	130,244	139.3%	121.3%	146.4%	138.4%	-0.65	-0.83	-0.60	-0.82	1.66	1.06	1.70	0.97
No. 106	181,224	243,465	138,514	209,660	68.1%	125.9%	62.1%	145.3%	-0.56	-0.01	-0.54	0.13	-0.13	1.17	-0.25	1.13
No. 107	195,006	188,229	151,056	153,643	108.5%	101.3%	110.3%	113.9%	-0.42	-0.55	-0.36	-0.54	0.89	0.54	0.86	0.44
No. 108	235,539	259,290	179,715	216,572	92.4%	111.8%	101.1%	142.4%	0.01	0.14	0.05	0.21	0.48	0.81	0.65	1.06
No. 109	172,949	172,559	125,734	135,479	120.7%	120.3%	116.4%	133.2%	-0.65	-0.70	-0.72	-0.76	1.20	1.03	1.01	0.86
No. 110	242,990	259,492	182,045	212,425	66.6%	77.9%	67.3%	95.2%	0.09	0.14	0.08	0.16	-0.17	-0.05	-0.13	0.03
No. 111	179,046	168,447	128,370	127,990	66.4%	56.6%	65.2%	64.7%	-0.59	-0.74	-0.69	-0.85	-0.17	-0.60	-0.18	-0.63
No. 112	273,320	298,643	194,933	230,413	5.2%	14.9%	0.7%	19.0%	0.41	0.52	0.26	0.37	-1.72	-1.66	-1.68	-1.63
No. 113	117,487	138,523	86,551	116,168	72.8%	103.7%	67.8%	125.3%	-1.24	-1.04	-1.28	-0.99	-0.02	0.60	-0.12	0.69
No. 114	186,533	206,637	137,156	166,551	48.4%	64.4%	41.9%	72.4%	-0.51	-0.37	-0.56	-0.39	-0.63	-0.40	-0.72	-0.47
No. 115	189,069	175,710	142,756	138,883	81.6%	68.7%	78.6%	73.8%	-0.48	-0.67	-0.48	-0.72	0.21	-0.29	0.13	-0.44
No. 116	187,004	167,533	134,855	130,562	55.3%	39.2%	50.6%	45.8%	-0.50	-0.75	-0.59	-0.82	-0.45	-1.04	-0.52	-1.05
No. 117	193,390	179,272	141,549	137,662	81.1%	67.9%	82.5%	77.5%	-0.43	-0.64	-0.50	-0.73	0.20	-0.31	0.22	-0.35
No. 118	195,572	200,167	150,776	166,102	93.6%	98.2%	92.4%	111.9%	-0.41	-0.43	-0.37	-0.39	0.51	0.46	0.45	0.40
No. 119	207,168	204,865	164,330	173,703	135.6%	132.9%	132.9%	146.2%	-0.29	-0.39	-0.17	-0.30	1.57	1.35	1.39	1.14
No. 120	233,653	268,042	165,022	212,875	44.9%	66.2%	35.3%	74.5%	-0.01	0.23	-0.16	0.16	-0.72	-0.35	-0.88	-0.42
No. 121	157,240	163,761	98,724	112,827	34.0%	39.6%	26.8%	45.0%	-0.82	-0.79	-1.11	-1.03	-0.99	-1.03	-1.07	-1.06
No. 122	165,109	177,475	127,981	146,363	77.8%	91.1%	78.3%	103.9%	-0.73	-0.66	-0.69	-0.63	0.11	0.28	0.12	0.22
No. 123	171,937	182,870	125,437	142,779	43.0%	52.1%	35.5%	54.2%	-0.66	-0.60	-0.73	-0.67	-0.77	-0.71	-0.87	-0.86
No. 124	100,528	95,820	72,586	74,192	43.5%	36.8%	43.7%	46.9%	-1.42	-1.45	-1.48	-1.49	-0.75	-1.11	-0.68	-1.02
No. 125	191,173	192,795	139,124	152,053	91.7%	93.3%	87.2%	104.6%	-0.46	-0.51	-0.53	-0.56	0.46	0.34	0.33	0.24
No. 126	105,418	145,498	83,360	129,824	86.8%	157.8%	83.2%	185.3%	-1.36	-0.97	-1.33	-0.83	0.34	1.99	0.24	2.00
No. 127	188,867	212,080	139,354	174,426	43.7%	61.4%	40.9%	76.3%	-0.48	-0.32	-0.53	-0.30	-0.75	-0.48	-0.75	-0.38
No. 128	134,389	133,702	105,739	112,201	62.0%	61.1%	61.8%	71.7%	-1.06	-1.08	-1.01	-1.04	-0.29	-0.48	-0.26	-0.48
No. 129	163,137	149,350	125,797	124,030	82.4%	67.0%	82.0%	79.4%	-0.75	-0.93	-0.72	-0.90	0.23	-0.33	0.21	-0.31
No. 130	217,866	238,365	159,705	190,388	19.3%	30.6%	14.7%	36.7%	-0.18	-0.06	-0.24	-0.10	-1.36	-1.26	-1.35	-1.24
No. 131	165,214	200,059	127,912	171,417	102.2%	144.8%	98.6%	166.1%	-0.73	-0.44	-0.69	-0.33	0.73	1.66	0.59	1.58
No. 132	95,505	99,061	72,524	81,705	78.0%	84.6%	75.6%	97.9%	-1.47	-1.42	-1.48	-1.40	0.12	0.12	0.06	0.09
No. 133	120,224	124,483	95,057	105,716	64.2%	70.0%	63.4%	81.7%	-1.21	-1.17	-1.16	-1.12	-0.23	-0.26	-0.22	-0.26
No. 134	177,269	233,194	137,824	201,550	79.9%	136.7%	80.4%	163.8%	-0.61	-0.11	-0.55	0.03	0.17	1.45	0.17	1.53
No. 135	163,761	205,695	132,961	181,439	120.1%	176.5%	121.6%	202.4%	-0.75	-0.38	-0.62	-0.21	1.18	2.47	1.13	2.37
No. 136	184,459	169,486	150,097	144,719	115.7%	98.2%	119.2%	111.3%	-0.53	-0.73	-0.38	-0.65	1.07	0.46	1.07	0.38
No. 137	148,788	164,974	119,453	143,139	107.0%	129.5%	109.6%	151.1%	-0.91	-0.78	-0.81	-0.67	0.85	1.26	0.85	1.25
No. 138	148,737	134,495	119,137	109,824	132.4%	110.2%	141.5%	122.6%	-0.91	-1.08	-0.82	-1.07	1.49	0.77	1.59	0.63
No. 139	147,450	140,474	116,238	116,897	98.1%	88.7%	104.6%	105.8%	-0.92	-1.02	-0.86	-0.98	0.62	0.22	0.73	0.26
No. 140	161,436	146,718	131,952	124,543	86.1%	69.1%	88.0%	77.4%	-0.77	-0.96	-0.63	-0.89	0.32	-0.28	0.35	-0.36
No. 141	133,056	173,554	109,796	154,975	98.7%	159.2%	106.1%	190.9%	-1.07	-0.69	-0.95	-0.53	0.64	2.02	0.77	2.12
No. 142	203,523	190,337	159,095	155,349	124.5%	110.0%	129.1%	123.7%	-0.33	-0.53	-0.25	-0.52	1.29	0.77	1.30	0.65
No. 143	167,009	173,729	135,458	147,632	102.9%	111.0%	107.8%	126.4%	-0.71	-0.69	-0.58	-0.62	0.74	0.79	0.81	0.71
No. 144	223,440	216,612	178,377	179,742	121.2%	114.5%	123.0%	124.7%	-0.12	-0.27	0.03	-0.23	1.21	0.88	1.16	0.68
No. 145	107,472	117,924	80,155	93,704	11.7%	22.5%	9.2%	27.7%	-1.34	-1.24	-1.37	-1.26	-1.56	-1.47	-1.48	-1.44
No. 146	140,747	139,807	110,774	116,647	59.2%	58.2%	61.6%	70.1%	-0.99	-1.02	-0.94	-0.99	-0.36	-0.56	-0.27	-0.51
No. 147	168,012	211,634	127,363	179,114	116.8%	173.1%	119.4%	208.6%	-0.70	-0.32	-0.70	-0.24	1.10	2.38	1.08	2.51

Store No.	Raw data before normalising the data								Normalised Data							
	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8
No. 148	169,641	177,113	131,461	146,004	88.6%	96.9%	90.9%	112.0%	-0.69	-0.66	-0.64	-0.63	0.38	0.43	0.41	0.40
No. 149	145,955	140,849	113,754	116,749	103.0%	95.9%	101.5%	106.8%	-0.94	-1.01	-0.89	-0.98	0.75	0.41	0.66	0.29
No. 150	149,139	182,948	120,858	158,572	111.0%	158.9%	119.1%	187.5%	-0.90	-0.60	-0.79	-0.48	0.95	2.02	1.07	2.05
No. 151	146,607	161,162	114,293	133,373	52.7%	67.8%	51.3%	76.6%	-0.93	-0.82	-0.89	-0.79	-0.52	-0.31	-0.50	-0.37
No. 152	145,050	159,808	114,824	134,500	124.1%	146.9%	129.7%	169.0%	-0.95	-0.83	-0.88	-0.77	1.28	1.71	1.31	1.64
No. 153	153,998	171,994	123,469	148,960	81.7%	102.9%	83.8%	121.7%	-0.85	-0.71	-0.76	-0.60	0.21	0.58	0.25	0.61
No. 154	151,245	180,708	121,646	153,633	66.0%	98.3%	67.0%	110.9%	-0.88	-0.62	-0.78	-0.54	-0.19	0.47	-0.14	0.38
No. 155	144,865	123,513	111,220	103,176	139.4%	104.1%	136.6%	119.4%	-0.95	-1.18	-0.93	-1.15	1.67	0.62	1.47	0.56
No. 156	285,497	248,482	248,489	221,801	234.0%	190.7%	270.8%	231.0%	0.54	0.04	1.03	0.27	4.05	2.83	4.59	3.00
No. 157	391,352	380,748	312,922	317,506	86.1%	81.1%	87.9%	90.7%	1.66	1.33	1.95	1.41	0.32	0.03	0.35	-0.07
No. 158	220,336	184,572	182,528	157,459	117.7%	82.4%	120.4%	90.2%	-0.15	-0.59	0.09	-0.50	1.12	0.06	1.10	-0.08
No. 159	198,006	224,841	156,943	192,876	79.7%	104.0%	80.6%	121.9%	-0.39	-0.19	-0.28	-0.07	0.16	0.61	0.17	0.61
No. 160	249,018	256,534	186,320	207,275	72.4%	77.6%	69.8%	88.9%	0.15	0.11	0.14	0.10	-0.02	-0.06	-0.08	-0.11
No. 161	133,189	120,087	104,447	98,215	75.3%	58.0%	79.2%	68.5%	-1.07	-1.22	-1.03	-1.21	0.05	-0.56	0.14	-0.55
No. 162	176,543	171,053	138,949	141,313	191.6%	182.5%	198.7%	203.8%	-0.61	-0.72	-0.53	-0.69	2.98	2.62	2.92	2.40
No. 163	116,563	155,710	92,889	137,426	110.7%	181.5%	109.5%	210.0%	-1.25	-0.87	-1.19	-0.74	0.94	2.59	0.85	2.54
No. 164	178,565	164,542	139,920	136,021	124.7%	107.0%	135.8%	129.2%	-0.59	-0.78	-0.52	-0.75	1.30	0.69	1.46	0.77
No. 165	220,620	265,728	171,359	231,044	62.3%	95.5%	57.8%	112.8%	-0.15	0.20	-0.07	0.38	-0.28	0.39	-0.35	0.42
No. 166	121,457	171,180	91,335	146,726	57.6%	122.2%	51.4%	143.2%	-1.20	-0.72	-1.21	-0.63	-0.40	1.08	-0.50	1.08
No. 167	85,573	73,500	66,891	60,551	151.7%	116.2%	167.1%	141.8%	-1.57	-1.67	-1.56	-1.66	1.98	0.92	2.18	1.05
No. 168	191,655	236,784	124,568	185,458	42.1%	75.5%	32.1%	96.7%	-0.45	-0.08	-0.74	-0.16	-0.79	-0.11	-0.95	0.07
No. 169	224,864	196,688	169,846	156,223	111.8%	85.3%	112.2%	95.2%	-0.10	-0.47	-0.09	-0.51	0.97	0.13	0.91	0.03
No. 170	164,717	112,807	118,140	87,670	143.3%	66.7%	128.9%	69.8%	-0.74	-1.29	-0.83	-1.33	1.77	-0.34	1.30	-0.52
No. 171	239,194	261,205	182,051	214,860	84.6%	101.6%	81.1%	113.8%	0.05	0.16	0.08	0.19	0.28	0.55	0.19	0.44
No. 172	203,331	247,817	156,363	208,765	31.4%	60.1%	29.1%	72.3%	-0.33	0.03	-0.29	0.12	-1.06	-0.51	-1.02	-0.47
No. 173	204,502	194,291	161,574	160,512	85.0%	75.7%	86.6%	85.4%	-0.32	-0.49	-0.21	-0.46	0.29	-0.11	0.32	-0.18
No. 174	274,861	275,249	217,687	229,356	87.9%	88.2%	89.6%	99.8%	0.43	0.30	0.59	0.36	0.37	0.21	0.38	0.13
No. 175	152,741	140,594	115,551	113,852	77.3%	63.2%	76.1%	73.6%	-0.86	-1.02	-0.87	-1.02	0.10	-0.43	0.07	-0.44
No. 176	200,945	176,748	158,677	147,788	69.4%	49.0%	62.1%	51.0%	-0.35	-0.66	-0.25	-0.61	-0.10	-0.79	-0.25	-0.93
No. 177	303,713	363,110	249,540	320,757	111.5%	152.9%	115.2%	176.6%	0.73	1.15	1.04	1.45	0.96	1.86	0.98	1.81
No. 178	198,706	205,552	157,476	175,600	99.0%	105.9%	99.5%	122.5%	-0.38	-0.38	-0.27	-0.28	0.65	0.66	0.61	0.63
No. 179	282,658	273,389	219,411	224,377	76.3%	70.5%	76.0%	80.0%	0.51	0.28	0.61	0.30	0.07	-0.24	0.07	-0.30
No. 180	263,691	288,240	205,370	239,709	81.3%	98.2%	81.1%	111.4%	0.31	0.42	0.41	0.49	0.20	0.46	0.19	0.39
No. 181	207,572	191,367	166,571	158,083	83.8%	69.4%	96.5%	86.4%	-0.28	-0.52	-0.14	-0.49	0.26	-0.27	0.54	-0.16
No. 182	184,471	185,306	148,593	157,082	86.5%	87.3%	91.6%	102.6%	-0.53	-0.58	-0.40	-0.50	0.33	0.19	0.43	0.19
No. 183	212,270	190,355	167,136	157,587	93.7%	73.7%	96.6%	85.3%	-0.24	-0.53	-0.13	-0.50	0.51	-0.16	0.55	-0.18
No. 184	256,782	259,847	206,761	221,078	161.2%	164.3%	168.3%	186.9%	0.24	0.15	0.43	0.26	2.22	2.15	2.21	2.03
No. 185	256,574	260,807	209,994	226,668	132.5%	136.3%	131.8%	150.2%	0.23	0.16	0.48	0.33	1.49	1.44	1.36	1.23
No. 186	239,923	291,867	189,018	247,488	36.0%	65.4%	36.1%	78.2%	0.06	0.46	0.18	0.58	-0.94	-0.37	-0.86	-0.34
No. 187	222,634	238,113	164,414	190,259	52.3%	62.8%	50.2%	73.8%	-0.13	-0.07	-0.17	-0.11	-0.53	-0.44	-0.53	-0.43
No. 188	146,990	181,305	109,250	152,231	81.4%	123.8%	83.1%	155.1%	-0.93	-0.62	-0.96	-0.56	0.20	1.12	0.23	1.34
No. 189	130,395	118,480	102,550	100,457	114.5%	94.9%	112.6%	108.2%	-1.10	-1.23	-1.05	-1.18	1.04	0.38	0.92	0.32
No. 190	143,437	156,118	106,881	125,507	86.8%	103.4%	85.8%	118.2%	-0.96	-0.86	-0.99	-0.88	0.34	0.60	0.30	0.53
No. 191	268,102	202,762	213,772	166,830	136.8%	79.1%	143.3%	89.9%	0.36	-0.41	0.53	-0.39	1.60	-0.02	1.63	-0.08
No. 192	128,846	215,106	110,754	194,434	54.8%	158.5%	61.8%	184.0%	-1.12	-0.29	-0.94	-0.06	-0.47	2.00	-0.26	1.97
No. 193	158,640	183,680	135,545	154,285	61.8%	87.4%	75.2%	99.5%	-0.80	-0.60	-0.58	-0.54	-0.29	0.19	0.05	0.12
No. 194	71,125	152,915	58,880	132,415	-10.6%	92.2%	-7.2%	108.8%	-1.73	-0.90	-1.68	-0.80	-2.12	0.31	-1.86	0.33
No. 195	86,348	180,844	68,476	157,860	-0.2%	109.1%	0.4%	131.4%	-1.57	-0.62	-1.54	-0.49	-1.86	0.74	-1.69	0.82
No. 196	120,282	161,583	94,498	139,789	37.2%	84.3%	35.1%	99.8%	-1.21	-0.81	-1.17	-0.71	-0.91	0.11	-0.88	0.13

Appendix B: The result of store cluster analysis

Moreover, Table 7-5 to Table 7-8 presented the result of store cluster analysis by each clustering technique including

- Current clustering technique
- K-Means clustering technique with eight variables
- Agglomerative clustering technique with eight variables
- K-Means clustering technique with the first four variables
- Agglomerative clustering technique with the first four variables
- K-Means clustering technique with the last four variables
- Agglomerative clustering technique with the last four variables

Store No.	Clustering Technique						
	Current Clustering Technique	K-Means - 8 Variables	Agglomerative - 8 Variables	K-Means - First 4 Variables	Agglomerative - First 4 Variables	K-Means - Last 4 Variables	Agglomerative - Last 4 Variables
No. 001	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 4	Cluster 4
No. 002	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 003	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2
No. 004	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 005	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 3	Cluster 3
No. 006	Cluster 1	Cluster 2	Cluster 3	Cluster 2	Cluster 2	Cluster 2	Cluster 2
No. 007	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 2
No. 008	Cluster 1	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 009	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 010	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 011	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 012	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 013	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 014	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 3	Cluster 3
No. 015	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4
No. 016	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 4	Cluster 4
No. 017	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 018	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 019	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 020	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 021	Cluster 3	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 022	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 023	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 024	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 3
No. 025	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 026	Cluster 3	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 027	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 028	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 029	Cluster 2	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 030	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 031	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 3	Cluster 2
No. 032	Cluster 3	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 033	Cluster 4	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 034	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 035	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2
No. 036	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 3	Cluster 4
No. 037	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 038	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 4	Cluster 4
No. 039	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 040	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 041	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 042	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 043	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 044	Cluster 2	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 3
No. 045	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 4	Cluster 4
No. 046	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4

Store No.	Clustering Technique						
	Current Clustering Technique	K-Means - 8 Variables	Agglomerative - 8 Variables	K-Means - First 4 Variables	Agglomerative - First 4 Variables	K-Means - Last 4 Variables	Agglomerative - Last 4 Variables
No. 047	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 3
No. 048	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 049	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 050	Cluster 2	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 051	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 052	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2
No. 053	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 054	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 055	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 056	Cluster 4	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 057	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 058	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 3
No. 059	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 060	Cluster 3	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 061	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 062	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 063	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 064	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 3
No. 065	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 066	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 4	Cluster 4
No. 067	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 068	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 069	Cluster 4	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 070	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 071	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 072	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 073	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 2	Cluster 1	Cluster 1
No. 074	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 075	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 076	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 077	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 078	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 079	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 080	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4
No. 081	Cluster 4	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 082	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 083	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 3
No. 084	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 085	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 086	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 087	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 088	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 2
No. 089	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 090	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 091	Cluster 1	Cluster 2	Cluster 3	Cluster 3	Cluster 2	Cluster 2	Cluster 2
No. 092	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 093	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 094	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 095	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 096	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 097	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 098	Cluster 2	Cluster 1	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 4
No. 099	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 100	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 3
No. 101	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 102	Cluster 3	Cluster 1	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 103	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 104	Cluster 3	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 105	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 106	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 2
No. 107	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 108	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 2	Cluster 2
No. 109	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 110	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3

Store No.	Clustering Technique						
	Current Clustering Technique	K-Means - 8 Variables	Agglomerative - 8 Variables	K-Means - First 4 Variables	Agglomerative - First 4 Variables	K-Means - Last 4 Variables	Agglomerative - Last 4 Variables
No. 111	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 3
No. 112	Cluster 4	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 113	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 114	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 115	Cluster 3	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 116	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 117	Cluster 3	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 118	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 119	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 2
No. 120	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 121	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 122	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 123	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 124	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 125	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 126	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 127	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 128	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3
No. 129	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 130	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4
No. 131	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 132	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 133	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 134	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 135	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 136	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 137	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 138	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 139	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 140	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 141	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 142	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 143	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 144	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 2
No. 145	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 146	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 3
No. 147	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 148	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 149	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 150	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 151	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 152	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 153	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 154	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 155	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 156	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 2	Cluster 1	Cluster 1
No. 157	Cluster 1	Cluster 1	Cluster 1	Cluster 2	Cluster 1	Cluster 3	Cluster 3
No. 158	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 2
No. 159	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 2
No. 160	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 161	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 162	Cluster 1	Cluster 2	Cluster 2	Cluster 4	Cluster 3	Cluster 1	Cluster 1
No. 163	Cluster 1	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 164	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 165	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 166	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 167	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 2	Cluster 2
No. 168	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 169	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 3	Cluster 3	Cluster 2
No. 170	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 171	Cluster 2	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2
No. 172	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 173	Cluster 3	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 174	Cluster 1	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2

Store No.	Clustering Technique						
	Current Clustering Technique	K-Means - 8 Variables	Agglomerative - 8 Variables	K-Means - First 4 Variables	Agglomerative - First 4 Variables	K-Means - Last 4 Variables	Agglomerative - Last 4 Variables
No. 175	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 3	Cluster 3
No. 176	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 4	Cluster 4
No. 177	Cluster 1	Cluster 2	Cluster 3	Cluster 2	Cluster 2	Cluster 2	Cluster 2
No. 178	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 179	Cluster 2	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 3
No. 180	Cluster 1	Cluster 3	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2
No. 181	Cluster 3	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 182	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 2
No. 183	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 184	Cluster 1	Cluster 2	Cluster 2	Cluster 3	Cluster 2	Cluster 1	Cluster 1
No. 185	Cluster 1	Cluster 2	Cluster 3	Cluster 3	Cluster 2	Cluster 2	Cluster 2
No. 186	Cluster 3	Cluster 3	Cluster 1	Cluster 3	Cluster 2	Cluster 4	Cluster 4
No. 187	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4
No. 188	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 189	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 190	Cluster 3	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 3	Cluster 2
No. 191	Cluster 1	Cluster 4	Cluster 3	Cluster 3	Cluster 2	Cluster 3	Cluster 2
No. 192	Cluster 2	Cluster 4	Cluster 3	Cluster 4	Cluster 3	Cluster 2	Cluster 2
No. 193	Cluster 3	Cluster 3	Cluster 3	Cluster 4	Cluster 3	Cluster 3	Cluster 3
No. 194	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 195	Cluster 4	Cluster 3	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4
No. 196	Cluster 4	Cluster 3	Cluster 4	Cluster 4	Cluster 4	Cluster 4	Cluster 4

VITA

NAME	Punnakrit Natekijcharoen
DATE OF BIRTH	23 April 1993
PLACE OF BIRTH	Bangkok
INSTITUTIONS ATTENDED	Chulalongkorn University
HOME ADDRESS	368/80 Maestro 12 residence, Thanonphetburi, Ratchathewi, Bangkok 10400