

White Paper



Store clustering

More than one right answer

Subhashis Nath

Abstract

Growing competition has made customer retention a critical issue for retailers. While they realize that sophisticated clustering tools do provide market insights, retailers are often unable to choose the clustering approach best suited for their business.

This paper explains how a clustering approach is effective only if it addresses a retailer's business objectives. It demonstrates how the right clustering approach – by performance benchmarking, price optimization, and assortment and space planning – can help meet these objectives.

Introduction

Even as retail chains span multiple regions and markets, customers expect their shopping experience to match up to their individual expectations. For stores to ensure customer satisfaction and increased sales in a rapidly changing environment, relevance has become the watchword in retail today. However, tectonic shifts in consumer demographics and increased consumer choice have made it increasingly difficult to provide consumers a relevant experience in all aspects of retail. This challenge has driven mass merchants to use segmentation and clustering tools to understand the market better and increase their ability to be relevant.

Today, most retailers adopt store clustering to ensure their strategies are designed as per store attributes. For instance, store clustering strategies could be based on the location or market type of stores – strategies will vary for stores on a beach, in a student township, an urban neighborhood, a suburban locale, etc. However, most retailers leverage clustering practices at the ‘basic’ level.

Through discussions with retailers, we have observed that once they are convinced about adopting a more sophisticated clustering process, we often encounter several questions. They range from queries about the best approach to clustering stores or going by demographic data vis-à-vis competition intensity. Retailers often seek to know whether they should analyze POS (point of sale) data and group stores with similar patterns, or should go by syndicated data (like Nielsen or IRI) and group stores as per market behavior obtained from such syndicated data. We always give them with the same answer – For what do you want to use your clusters?

Clustering Best Practice: Align the Clustering Process to Meet your Business Objectives

The concept of clustering can be applied across many retail functions – assortment planning, space planning, pricing, promotion planning, store performance benchmarking, etc. Since no single clustering scheme (clustering methodology and process) can fulfill all the requirements of a company, the methodology and design of the clustering process must be defined by the company’s business objectives. To explain the need for this approach, we have put together three examples, as follows:

Example 1 Use Clustering to Achieve Store Performance Benchmarking

It is important to cluster together stores with similar selling potential so that the comparison of performance is fair. If we were to use demand patterns exhibited by POS data for such an objective, the purpose would be defeated. This is because POS data depicts actual performance and by using such data one would end up with clusters of stores “having similar performance” rather than with clusters of stores that “should have similar performance”. Therefore, in this case, stores would need to be grouped by similar opportunity of performance – probably an attribute grouping – starting with size, retail trade area (RTA) population, income levels, freeway access, competition, etc. It is important to note that in practice, this exercise would be more rigorous, and the following provides an overview of the process:

- The process is initiated by using a superset of attributes that can affect performance (it could contain as many as 300 attributes)
- This is followed by using historical data across stores and the historical values of these attributes. These help us perform a regression analysis to identify the most significant attributes
- In most cases, about seven attributes will explain 70% of the performance differences. These seven attributes are then used to perform a k-means-based clustering of all the stores. The ‘k’ in k-means stands for a user-defined number of clusters, i.e., the user stipulates to the statistical tool how many clusters are to be formed. Using the attribute data, the package groups the stores into the stipulated number of clusters. Please note that there are statistical methods that allow one to infer the correct number of clusters as well.

Example 2 Use Clustering to Achieve Price Optimization

Elasticity analysis is ideal in the case of pricing. If it is coupled with a price optimization package, a store can easily carry out analysis and price derivation. However, the optimization package will operate within rules and constraints set by the user. To overcome this issue, pricing clusters (or zones) need to be used to prevent the following:

- Application of similar pricing rules for the entire chain
- Creation of rules for every store

Thus, these clusters are groups of stores for which the same pricing rules can be applied. To create such clusters, one should not take POS data into account. POS performance depends on right pricing, so we could have a cyclic self-fulfilling effect if we set pricing rules based on POS performance.

Among the attributes to be used here include:

- Income levels in the RTA
- Competition presence and type (a Wal-mart has a different implication on pricing than a Target)
- Regulatory zones
- Education levels
- Primary life-stage (e.g., if a large percentage of the population are students, or senior citizens)
- Typical lifestyle (modern – more singles, or conservative – larger families)

These clusters also have to be grouped by category, since, for example, if you have a Walgreens next door, you might have to be aggressive in pricing the pharmacy and health and beauty care (HBC) category.

Example 3 Use Clustering to Achieve Assortment and Space Planning

Retailers use clustering most during assortment and space planning. When clustering for assortment planning, one can use POS data and identify stores that show similar demand patterns. However, planning assortments based on such grouping has a cyclic effect. For example, if a store never sold Hispanic products one would never know if selling Hispanic products makes sense for that store since the demand pattern for Hispanic products would not exist for that store.

The trick is to combine RTA syndicated demand data from external sources and also use one's POS data, and then create clusters accordingly, i.e., by exhibited patterns in store sales as well as patterns in market sales.

However, continuing with the Hispanic example, it is possible that no retailer in the RTA has figured out the potential for Hispanic products. In that case, using syndicated data still would not help. It is therefore also important to ensure that clusters are reported to experienced merchandisers who can then modify them based on their experience and the current information.

What retailers do for store clustering in pricing, holds true for assortment and space planning. Retailers would be advised to adopt a category-specific clustering of stores – where stores are grouped based on data specific to the appropriate category, such as POS data, market information, and manufacturer category data.

Using such clusters in assortment and space planning allows the merchandise mix for a particular category to reflect the exhibited sales patterns/ margins and the market potential of that category.

Thus, as seen in the above three examples (and we've tried to keep them simple), clustering methodologies differ according to the various business objectives that need to be achieved. Therefore, every clustering exercise should begin with the question, "What do you want to achieve?"

Conclusion

As numerous methodologies and data inputs are required while creating store clusters to achieve different business objectives, retailers must adopt a clustering tool that has the following characteristics:

- Easy to configure and customize
- Defines item and store attributes effectively
- Defines data collation and cleansing parameters
- Enables parameter setting
- Has the capability of top-down and bottom-up clustering
- Allows reviewing and reporting of clusters in a manner that is understandable to business users

A tool with such characteristics provides the flexibility of designing different clustering processes and techniques, thus enabling a retailer to cluster his store based on business objectives.

Infosys has deep experience in strategic roadmap definition for clustering, requirements gathering and process definition, buy v/s build analysis, package evaluation in store clustering and implementation, and consulting support in the usage of clusters.

The following table outlines high-level differences while clustering for different objectives:

	Objectives	Statistical Methodology	Data for creating the clusters	Data used for describing the clusters
Relevant Promotion	Develop segment-specific promotions	<ul style="list-style-type: none"> • Hierarchical • Classification & Regression Trees • Agglomerative 	<ul style="list-style-type: none"> • Promotion performance • Promotion types 	<ul style="list-style-type: none"> • Store Attributes • Shopper Demographics • Zip code zone
Relevant Merchandise	Create assortment planning rules	<ul style="list-style-type: none"> • Divisive Ward's Method • K-means 	<ul style="list-style-type: none"> • Sales performance • Loyalty card data 	<ul style="list-style-type: none"> • Store Attributes • Shopper Demographics
Relevant store plans	Create store planning template	<ul style="list-style-type: none"> • CHAID 	<ul style="list-style-type: none"> • Sales revenue • Sales profitability • Shopper demographics • Loyalty card data 	<ul style="list-style-type: none"> • Store Attributes • Shopper Demographics
Relevant pricing	Create pricing rules		<ul style="list-style-type: none"> • Competition • Shopper demographics • Shopper life stage 	<ul style="list-style-type: none"> • Store Attributes • Shopper Demographics • Competition
Relevant marketing programs	Create Campaigns with better ROI		<ul style="list-style-type: none"> • Consumer demographics • Lifestyle-related attributes • Store footfall measures 	<ul style="list-style-type: none"> • Shopper demographics/ Lifestyle attributes • Store Attributes

To discuss more, please contact us.

About the Author

[Subhashis Nath](#) is a Principal Consultant with Infosys' Retail and CPG practice. He has extensive experience in marketing, customer relationship and consulting. He has provided strategic and process consulting services to European and US retailers in the areas of CRM, demand planning, merchandising and analytics. He can be reached at subhashis_nath@infosys.com.



For more information, contact askus@infosys.com

About Infosys

Many of the world's most successful organizations rely on Infosys to deliver measurable business value. Infosys provides business consulting, technology, engineering and outsourcing services to help clients in over 30 countries build tomorrow's enterprise.

For more information about Infosys (NASDAQ:INFY), visit www.infosys.com.