JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

The case for using Factor Score based Two-step Clustering in E-commerce Consumer Segmentation

Dr. Rachna Mohan

Associate Professor, Fabric and Apparel Science,
Department of Home Science, Bhagini Nivedita College,
University of Delhi, India

Abstract:
Background: Over the last 50 years there has been a widespread application of market segmentation to differentiate products, target customers, organize distribution and use the 4 Ps marketing mix. In the new millennium, a new concept of customer segmentation has also emerged, thanks to the advancements in Artificial Intelligence (AI) and Machine Learning (ML), aimed at helping businesses to retain existing customers. However, the macro issue of consumer segmentation, which is very vital for understanding the scope for future growth, new products and innovations and increasing the market share remains relatively less-addressed. E-commerce is a new growth space. It is founded upon the principle of aggregation and not segmentation; its core value proposition is more about service than product; its operations are also place and time neutral. And thus they present a new paradigm - more consumer-centric than product-centric and more aligned to Lauterborn's 4 Cs than McCarthy's 4 Ps of Marketing Mix. Therefore, we need to examine whether ecommerce is better served by the 4 Cs based consumer segmentation as opposed to the traditional market segmentation. We would also need to examine whether the current segmentation strategies focused on selling products and retaining customers will serve the long-term growth prospects of businesses.
Materials and Methods: This research is in two parts with the first focused on providing an overview of the approaches and methods used in extant research concerning market and customer segmentation. The second part presents a case study of research done by the author using factor score based two-step clustering for segmentation and sub-segmentations.
Results: The research papers surveyed brought out a plethora of approaches and methods used for market/consumer segmentation. It also showed the increasing use of data science based behavioral segmentation in ecommerce settings. Also notable was the adoption of 4 Cs based approach in digital marketing scenarios. The examination of the factor score based two-step clustering method used by the author in two of her studies showed that this method helped address the deficiencies identified in other approaches and methods.
Conclusion: While geographic segmentation may not be necessary for ecommerce operations, they need more than the online purchase-data based behavioral segmentation. Psychographic segmentation with demographic sub-segmentation, using intuitive methods such as factor score based two-step clustering will help ecommerce entities proactively engage with the present and prospective customers and achieve long-term growth and strategic advantages.

Key Words: Market segmentation, Ecommerce Consumer Segmentation, Factor-score based clustering, Intuitive consumer sub-segmentation

I. Introduction

What must be clarified at the outset is the distinction between the three types of segmentation approaches found in extant research namely: 1) Market Segmentation, 2) Customer Segmentation and 3) Consumer Segmentation. Although these segmentation strategies use the same or similar methodologies and face common data challenges, each serve a different purpose. This clarification becomes very necessary since most authors use these interchangeably leaving room for confusion to creep in.

1.1 Market Segmentation

It is part of the marketing strategy of a business to find groups of consumers whose genuine needs and desires match the products or service that a company can supply [1]. Market segmentation is done on the basis of variables that can account for the variations in the consumption of the concerned product or service [2]. Once the segments are identified, businesses can decide on the segments to serve based on the differences among them and how far they can customize their offerings to suit the target segments [3]. It is a tool in the hands of a marketing manager to select a target market and design the marketing mix [4]. Market segmentation also helps companies recognize their competitive strengths relative to the identified segments and the competitors operating in these segments [5]. It also helps companies to sustain their brand leadership [6]. Market segmentation shows how a company views the market through the lens of its products, services, strengths and weaknesses to find profitable opportunities. It yields short term strategies within the long term goals of a business. In strategic marketing, segmentation is also considered as a crucial instrument for developing a marketing strategy [7].

1.2 Customer Segmentation

It involves grouping the existing customers based on common traits so that the company can tailor its communications, messages, promotions and loyalty programs accordingly [8]. Salminen et al. in their systematic literature review on customer segmentation clearly point out that market segmentation and customer segmentation are different despite their similarities. While the former focuses on the overall market the latter confines itself specifically to the current customer-base of the concerned business entity [9]. Customer segmentation has become integral part of both ecommerce and retail platforms. This is also an area intense research for developing algorithms for segmentation. However, most of the current research uses the Recency, Frequency and Monetary value (RFM) of customer's purchases as the basis for customer segmentation.

1.3 Consumer Segmentation

Unlike, market and customer segmentation which are focused on specific products and their buyers, consumer segmentation is concerned more about the factors, which fall outside the market but play a role in the buying decisions of people. There are many variables that influence what people but and consume. These work behind the scenes simultaneously in the minds of the consumers. One of the methods used for profiling consumers based on their activities interests and opinions (AIO) is a lifestyle based segmentation [10]. Lifestyle topologies, also known as psychographics have also been widely used for targeting marketing communications [11]. Lifestyle segmentation has evolved into a framework for segmenting consumers based on a combined set of demographic and psychographic variables. Several researchers have also used the VALS framework in different countries to test its applicability [12, 13]. While lifestyle segmentation relies on quantitative data, there is also a consumer culture theory based approach which advocates that qualitative research must also be added to gain correct understanding of consumer perspectives [14]. Another variety of consumer segmentation is the proprietary PRIZM Premier, which uses geo-demographic parameters to classify US households into 68 consumer segments on the basis of their purchasing preferences [15].

II. RESEARCH OBJECTIVES

The overall aim of this paper is to state the case for using factor score based two-step clustering to segment consumers in general and ecommerce consumers in particular. The following are the specific objectives of the paper:

- To provide an overview of the analytical methods and tools used by researchers across the market, customer and consumer based segmentation approaches and the observed limitations.
- To explain, with the help of two research case studies, the usefulness of factor score based two-step clustering to overcome some of the limitations in the currently popular segmentation techniques.

III. MATERIALS AND METHODS

This research paper uses a mixed methodology of literature review and case studies. The literature review was done to identify the analytical methods used by researchers so far and the observed limitations. The papers were looked up through the research repositories and open access journals. Web search engines were also used to collect additional data. The research case studies used in this paper were done by the author. Both the researches used factor score based two-step clustering as the main analytical tool.

This research was survey based and done in 2007 to segment ecommerce consumers based on primary data collected through a structured and representative sample survey. The survey instrument gathered demographic, psychographic and behavioral data. Likert Scale was used for obtaining psychographic data.

IV. RESULTS AND DISCUSSION

4.1 Results of the Literature Survey

4.1.1 Tools and techniques used

Segmentation may be done on the basis of geographic, demographic, psychographic, behavioral, and situational variables as well as preferences [17]. Although the stated objective of segmentation is to group consumers or customers based on their similarities, the strategic purpose is to understand the distinct differences among the segments to which so that marketing managers can choose the ones, which they can target viably and profitably. In recent times, segmentation has been integrated with the competitive strategies of targeting and positioning [16]. Researchers have used several types of statistical techniques such as factor, cluster, conjoint, latent class analyses and logistic regression depending upon the specific objectives of the segmentation exercise as well as the nature of the variables that went into the analysis. However, until the availability statistical software packages, only simple and univariate methods found favor because of the complexities associated with advanced multivariate procedures. Recent advances in infotech have not only made it easy to use advanced statistics but also ushered in a wide array of AI and ML based solutions.

J. Salminen et al. present a systematic review of 172 papers in the area of algorithmic customer segmentation. Their review found that a majority (77.9%) of the reviewed studies segmented customers using 46 different algorithms - with K-means clustering used in 20.1% of these studies. Other studies used a wide variety such as K-means clustering variants (7.5%), fuzzy algorithms, and latent class (6.0%), Recency, Frequency and Monetary gain (RFM) and its variants 4.5%), Self-Organizing Maps (SOM) and Genetic Algorithms (GA) 2.25% each. There were also others like the Louvain, Ward's, and hierarchical clustering algorithms used two or three times. 80% of these studies only a single algorithm while others have adopted multiple like K-means clustering, SOM and RFM. Most researchers had extracted 4 to 5 clusters in their studies using data-driven to decide the number of segments, such as the elbow method. 81.7% of the studies used the number of segments as the hyper parameter in their models. The review also identified 14 types of evaluation metrics - with separation-focused metrics getting used slightly more (57.1%) than the statistics-focused. Only 7 studies (4.1%) marketing practitioners were approached for evaluating the Algorithmic Consumer Segmentation (ACS) outputs.

4.1.2 Critique of current trends in segmentation research

The flurry of recent research activities in the area of market/customer segmentation appears to be more technology driven than business or academia driven. Reviewers of ongoing research identify not only a knowledge gap but also technical deficiencies and a disconnect between marketing practitioners and the technology developers. Johannes cites several reviewers who have expressed their concerns about the growing divide. One of them is of the view that the current segmentation research is not amenable to confident generalizability since it is about how it should be done and not about its actual use. Another reviewer points to the lack of focus on developing models that help in solving critical and strategic issues. Yet another reviewer laments over the absence practical advice on doing the segmentation right by choosing the correct variables and segments and controlling the interventions. There are also others who doubt if the marketing practitioners shared the optimism of the researchers about the potential benefits of segmentation. It is also pointed out that the current research outputs do add up to a uniform and consistent framework for the practitioners to follow [19]. Bruce Cooil et al. point out the mismatch between the statistically identifiable segments and what exists in reality [17]. Saminen et al. also echo the previously mentioned concerns about the lack of synthesis about processes and parameters concerning the segmentation process. They also draw attention to the need for greater involvement and participation of experts in evaluating the research outputs [9]. According to Daniel Yankelovich, market segmentation should not be viewed as a tool for only targeting advertisement campaigns. The approach needs to be broadened to make it inform product innovation, pricing and distribution channel selection. It should also be the basis for building relationship between the consumer and the product or the product category. He also refers to the 2004 survey of 200 senior executives of large companies for example by Marakon Associates and the Economist Intelligence Unit which revealed that only 14% derived benefits while 59% of them had conducted a significant segmentation exercise in the course of the previous two years [20].

4.1.3 The Rationale for Using Factor Score based two-step Clustering

Factor analysis is not just a multivariate tool. It is also a tool with multiple procedural options, which helps in intuitively applying it suit the specific needs of data analysis. A very valuable feature of factor analysis is the factor score data generated by it for use in further analysis. This feature enables analysts to squeeze in more variables into procedures like cluster analysis. This facility is particularly useful for analyzing human behaviour which invariably result from the simultaneous interplay of a multitude of thought processes, working at the conscious and subconscious levels. In addition to the convenience of helping to compress a large number of variables into fewer manageable factors, it also provides the additional advantage of normalizing the data fed into it, which eliminates the impact of varying dimensionalities embodied in the data coming from different measurement systems. Equally important are the checks like sample adequacy and sphericity it performs, which helps ascertain the suitability of the data set as well as the communality estimates that help decide on the inclusion of each variable in the data set. There are also well established conventions that have set clear cut off points for accepting the results of the procedural options used in the analysis. Although factor analysis has been used as a segmentation tool by itself, this process limits the scope for performing further analysis on the characteristics of the different segments. This limitation can be overcome with the help of the factor scores generated along with the factors extracted.

Factors scores provide a true representation of the same old data collected from each participant or subject - compressed into a new latent score for each extracted factor. In other words, factor score stands for the contribution made the by the concerned individual's original data in the newly formed factors. There are also several methods to choose from for generating the factor scores depending on their end use. Some of the popular factor score generation methods include: (1) Regression; (2) Bartlett; (3) Anderson-Rubin; and (4) Thompson [21]. Factor scores also come with a very convenient feature. They are inherently standardized and normalized and can be directly used for further analysis [22].

The method of rotation used during the factor extraction process will also affect the factors and the factor score. The choice is between orthogonal and oblique methods; the former creates factors that are not correlated among themselves and the latter allows for some degree of inter-correlation among the extracted factor. Here again, the end use will decide. If the objective is to create a simple structure or a parsimonious theory, orthogonal rotation is the answer. If the aim is to create realistic cluster, then oblique rotation is preferable. Since orthogonal rotation makes it easier to interpret the clusters, the general advice is that one should first do oblique rotation to find the extent of inter-correlation and then decide to stick with it or opt for orthogonal rotation depending on the degree of inter-correlation [23].

The two-step clustering is preferred because of the opportunity it offers to use a categorical variable along with the factor scores. This creates options for experimenting with different demographic variables as is needed. Also, this procedure provides the option to use automatic clustering or predefine the number of clusters to be created. This procedure also offers two key advantages for consumer segmentation. First, the variable wise and factor wise plots generated by the procedure helps understand the cluster characteristics very clearly. Second, the cluster membership data can be imported into the original data sheet and used for segmenting it and repeat the factor extraction, factor score generation and clustering for each to create sub-clusters within the original clusters to gain deeper insights.

4.2 Research Case Study

This case study has been extracted from a larger research, done in 2007, aimed at making a comparative analysis of the preferences of apparel consumers towards buying from retail and etail stores. It gathered data through a structured sample in Delhi. The survey instrument included three sets of variables covering preferences towards selecting i) stores, ii) apparel, and iii) buying specific types of apparel, separately for retail and etail segments, besides common demographic variables. There were 784 complete responses. Analysis of the response data was done with the help of factor score based two-step clustering. The illustrative excerpts presented here relate only to their online apparel store selection related preferences. The aim of this presentation is to demonstrate how this kind of detailed consumer segmentation can be helpful to the marketing teams.

4.2.1 Principal Component Analysis (PCA)

17 out of the 22 consumer preference variables relating to online apparel shopping met the criteria for inclusion in the PCA. The rest were reserved for reintroduction later in the cluster based PCA. PCA extracted six distinct components using the correlation matrix and Eigen values of over 1 condition. The extracted components were rotated using the Varimax method. These six factors together accounted for 66.98% of the variance. The variance explained by the factors fell in to two categories. While factors 1 to 3 explained approximately 14% variance each, the next 3 factors around 8% each. The Rotated Component Matrix generated by the PCA process is given in Table 1.

Table 1: Rotated Component Matrix

	Consumer Preference Variables	Components					
		1	**2**	**3**	**4**	**5**	**6**
1	I look for good quality pictures to assess the visual appeal	.802					
2	Quality marks/labels help assess the quality of clothes sold online.	.751					
3	Detailed quality information would give me confidence to buy online	.722					
4	I would prefer zoom facilities to take a closer look at the finer details	.639			.365		
5	I would prefer to buy clothes of my favorite brands from online stores		.835				
6	I would feel more confident buying branded clothes online		.794				
7	I would prefer to buy branded clothes from their company web stores		.771				
8	I would prefer online stores that offer alteration services for perfect fitting			.772			
9	Detailed size specifications of clothes would give me confidence to buy online.			.648			
10	Not being able to try out the garment discourages online shopping			.600			
11	I look for simulated trials on 3D figures to check the fitting			.590			
12	I would miss the experience of price bargains in online stores				.800		
13	I would miss the touch and feel experience in online stores			.507	.553		
14	I would prefer to buy online, ready-mades than custom stitched clothes					.746	
15	I find online stores a good source for comparative price information		.328			.658	
16	I would prefer online shopping if it offered prices lower than retail						.788
17	I find online stores a good source for buying clothes at discounts					.340	.762

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 8 iterations.

Based on the Component Matrix the extracted factors were identified as given in Table 2.

Table 2: Interpretation of Factors

		Identified Factors	Variance Explained
1	• I look for good quality pictures to assess the visual appeal • Quality marks and labels help assess the quality of clothes sold online. • Detailed quality information would give me confidence to buy online • I would prefer zoom facilities to take a closer look at the finer details	Quality assessment information and tools	14.250 %
2	• I would prefer to buy clothes of my favorite brands from online stores • I would feel more confident buying branded clothes online • I would prefer to buy branded clothes from their respective company web stores	Preference for brands	14.222 %
3	• I would prefer online stores that offer alteration services for perfect fitting • Detailed size specifications of clothes would give me confidence to buy online. • Not being able to try out the garment discourages online shopping • I look for simulated trials on 3D figures to check the fitting	Size and fit assessment information and tools	13.712 %
4	• I would miss the experience of price bargains in online stores • I would miss the touch and feel experience in online stores	Absence of touch-n-feel and bargains	8.560 %
5	• I would prefer to buy online, ready-mades than custom stitched clothes • I find online stores a good source for comparative price information	Preference for ready-mades and price comparisons	8.241 %
6	• I would prefer online shopping if it offered prices lower than retail • I find online stores a good source for buying clothes at discounts	Low price and discounts	7.999 %

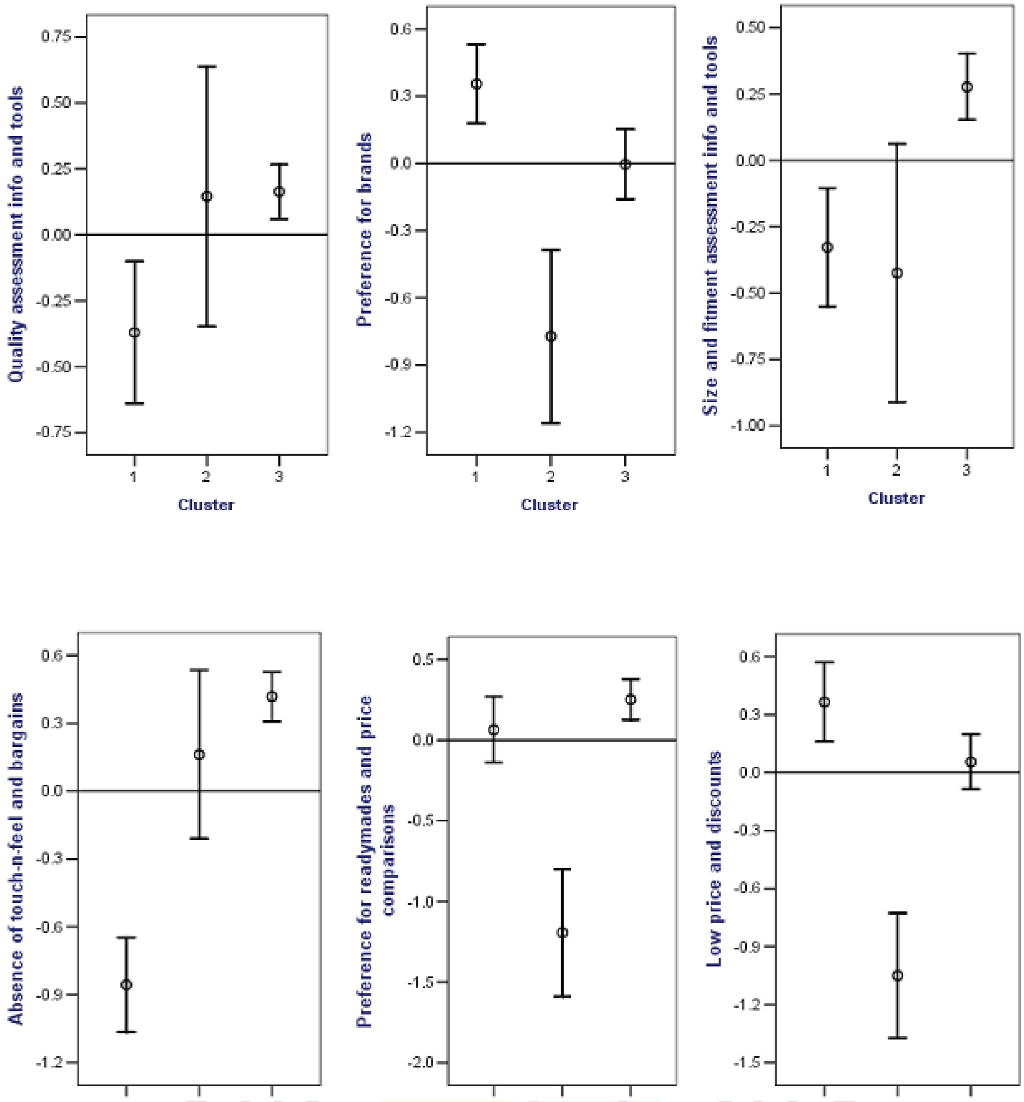
4.2.2 Cluster analysis

Two Step Cluster Analysis (TSCA) was performed based on Schwarz's Bayesian Criterion (BIC) on the factor scores produced by PCA. TSCA produced 3 clusters as shown in Table 3.

Table 3: Cluster Distribution

Cluster	No. of Respondents	% of Total
1	118	30.1%
2	53	13.5%
3	221	56.4%
Total	**392**	**100.0%**

Variable Importance (VI) plots were produced using TSCA for identifying the defining characteristics of each of these clusters applying 95% confidence level as the condition for each factor and cluster. Figure 1 contains VI plots for all the online apparel selection factors across all the 3 clusters extracted by TSCA procedure.



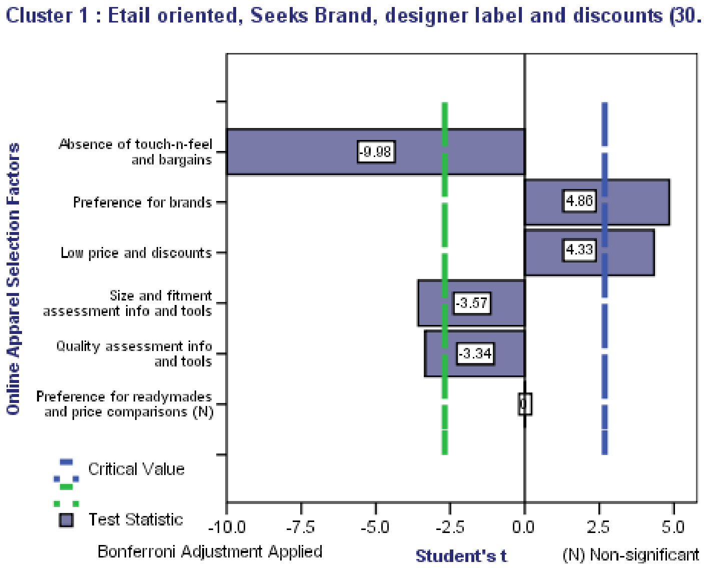
One of the points revealed by these plots is the fact that, cluster 2 exhibits very high variability with respect to components 1 and 3 and relatively high variability with respect to all other components. It could be inferred that this relatively small (13.5%) cluster consists of respondents with widely differing perceptions about buying apparel from online stores. The following interpretations are made based on this analysis.

- Quality assessment information and tools: This factor had a positive influence on cluster 3 with 56.4% of the respondents and least important for 1, with 30.1%, in significant variability. Cluster 2 showed the maximum variability in terms of the 95% confidence interval of means.
- Preference for brands: This factor had a distinct positive influence on cluster 1. Cluster 2 is least influenced by this factor while it exerts a neutral influence on cluster 3. In all, this factor is important for 30.1% of the respondents. Clusters 2 exhibited the highest variability in the extent to which they were influenced by this factor.
- Size and fitment assessment information and tools: This factor had the maximum influence on cluster 3, thereby being important to 56.4% of the respondents. Clusters 1 and 2 were least influenced by this factor. Clusters 2 exhibited the highest variability in the extent to which they were influenced by this factor.
- Absence of touch-n-feel and bargains: This factor had a distinct positive and negative influence on clusters 3 and 1 respectively. Cluster 2 was positively influenced by this factor to some extent. Cluster 2 also exhibited the highest variability in the extent to which they were influenced by this factor.
- Preference for ready-mades and price comparisons: This factor had a positive influence on cluster 3, thus being important to 56.4% of the respondents. Cluster 1 was relatively neutral to this factor. Clusters 2 exhibited a negative influence with highest variability towards this factor.
- Low price and discounts: This factor had a positive influence on clusters 1, thus being important to 30.1% of the respondents. Cluster 3 was relatively neutral to this factor. Clusters 2 exhibited a negative influence with highest variability towards this factor.

4.2.2.1 Analysis of the online apparel selection preference pattern of cluster 1

As is revealed by the variable wise importance plot, the consumers who belonged to this cluster were positively influenced by brands and low prices. This cluster is distinctly brand oriented since the disinterest in quality and fit assessment tools is attributable to the positive influence of brands. Brands in this case become the surrogate guarantor of quality and fit. So much so that this cluster is considerably disinterested in touch and feel. The significantly positive influence of low price and discounts could also interpreted as an additional preference for buying branded clothes at lower prices or discounts.

Fig. 2: Online Apparel Selection related Variable Importance Plot of cluster 1



In order to fully understand the additional factors that could influence the online apparel choice of this category of consumers, Cluster Based (CB)-PCA was performed on the variables previously excluded from the PCA done for the entire sample in the first step on account of low communalities. The same conditionality of communality above 0.05 and explanation of at least approximately two thirds of variance was applied. 4 out of the 5 residual variables met the criteria set for CB-PCA. The rotated component matrix of the PCA on these residual variables is given in Table 4

Table 4: Rotated Component Matrix for Members of Cluster 1

Variables		Component	
		1	2
1	Internet is better than retail for finding latest styles & fashion more easily	.885	
2	I find online shopping better than retail for buying designer labels	.786	
3	I look for online stores that offer wider choice in terms of sizes	.328	.773
4	I would prefer online stores, which offer wider variety of brands	.383	-.713
Variance explained %		**41.37**	**27.61**

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 3 iterations

This Cluster Based PCA (CB-PCA) revealed that this group of consumers were interested in two additional factors:

1. Belief about Internet being better than retail for buying latest styles and designer labels. This factor brings out that online stores selling latest designs and designer labels could attract online consumers, who are in favor of buying branded garments. This factor is important for this group of consumers as it accounts for a relatively larger variance of 41.37% among the 5 residual variables.
2. Preference for buying from online stores providing wider variety of sizes. The negative loading of the variable No. 4 also underscores the fact consumers who are interested in buying branded clothes online are interested in wider collection of sizes pertaining to a brand and not a wider collection of brands per se. This factor also confirms the preference for buying branded apparel from original company web sites. The cross loading of these two variables on component 1 is relatively small and therefore ignored.

The above procedure was repeated for the remaining two clusters (Cluster 2 and Cluster 3)

4.2.2.2 Demographic Influences on Online Apparel Store Selection Factors (OASSF)

TSCA was performed on the online store selection factors extracted by PCA for the entire sample using the demographic data as categorical variables to assess if these had any significant influence on the way in which consumers chose etail apparel stores.

4.2.2.2.1 Influence of Gender on OASSF

The sample included equal number of males and females. The TSCA revealed that in the case of for buying readymade items online and making of online stores for price comparisons males were positively interested in it to a significant degree, females were significantly disinterested in it almost to the to the same extent. Fig. 3 and Fig 4 present VI plots for male and female clusters.

Fig. 3 VI plot for Males with reference to OASSF

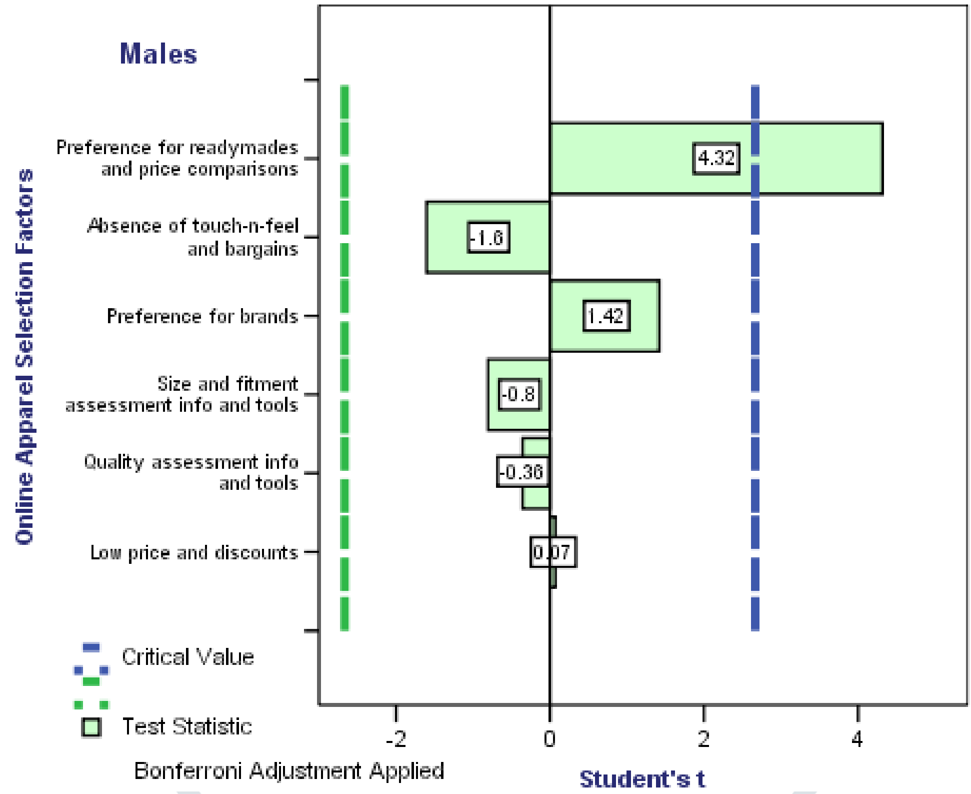
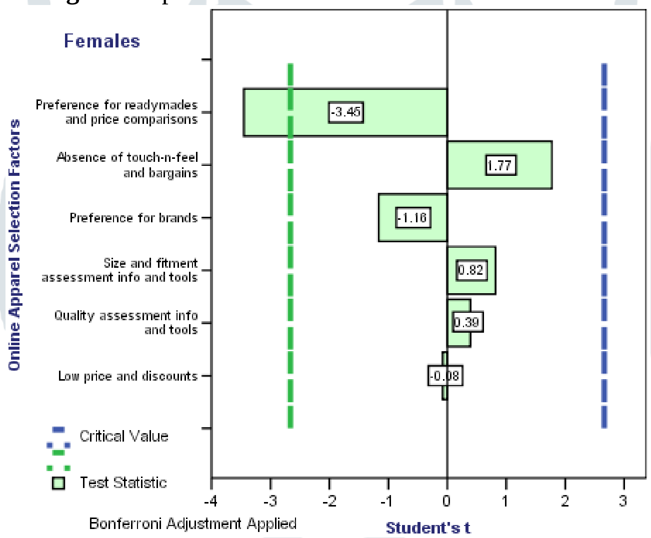


Fig. 4: VI plot for Females with reference to OASSF



The analysis also revealed that males and females differed with respect to all the factors relating to apparel selection from online stores. Though these differences were not statistically insignificant at the 95% confidence level, they were still worth taking note. While males were in favour of brands, females showed interest in size and quality assessment information and tools. In the case of the touch and feel too males and females differed; while males were disinterested in it, females were somewhat interested in it. Similar analysis was also performed and interpreted on all other demographic categories like i) age-group, ii) occupation, iii) annual spending, iv) Internet experience, etc.

4.2.2.3 Summary of Cluster and Cluster based PCA, Conclusions and recommendations

The summary of the findings of the cluster analysis and cluster based PCA are given in Table 5.

Table 5: Findings of the cluster analysis and cluster based PCA

1	**Etail oriented, Seeks Brand, designer label and discounts**	**30.1%**
Factors of positive interest	• Preference for brands • Low price and discounts	
Additional factors of interest	• Believes that Internet is better than retail for buying latest styles and designer labels. • Preference for buying from online stores providing wider variety of sizes.	
Neutral factors	• Preference for ready-mades and price comparisons	
Factors of disinterest	• Absence of touch-n-feel and bargains • Size and fitment assessment information and tools • Quality assessment information and tools	

Shopping orientation	• Not concerned about credit card misuse and online stores being impersonal in nature. • Perceives online stores to be better than retail in offering wider variety and quick search.	
Store selection orientation	• No significant pattern observed	
Demographic influences	• Those who attempted to buy clothes online as also those who had bought any product online were significantly influenced by brands. • Those whose online purchases were between Rs. 1000 and 5000 were also positively influenced by brands • Non-working are significantly disinterested in buying branded garments from online stores • Consumers spending more than Rs.20000 annually on clothes were not influenced by low prices and discounts	
Most likely to visit	• Online stores selling wider variety of branded or designer labels, as well as official online stores of established brands and designers.	
2	**No definitive etail orientation**	**13.5%**
Factors of positive interest	• None	
Additional factors of interest	• Preference for shops offering wider choice of brands and sizes and display of all additional charges, • Belief about online stores being better than retail for finding latest styles, designs and buying designer labels.	
Neutral factors	• Size and fitment assessment information and tools • Absence of touch-n-feel and bargains • Preference for brands	
Factors of disinterest	• Low price and discounts • Preference for ready-mades and price comparisons • Preference for brands	
Shopping orientation	• Significantly disinterested in wider variety and quick search • Does not also believe that online shopping offers more comfort and convenience as compared to retail stores	
Store selection orientation	• Significantly disinterested in credibility, loyalty reward and services	
Demographic influences	• None	
Most likely to visit	• No predictable pattern	
3	**Readymade garments oriented, seeks substitutes for retail conveniences**	**56.4%**
Factors of positive interest	• Absence of touch-n-feel and bargains • Size and fitment assessment information and tools • Preference for ready-mades and price comparisons • Quality assessment information and tools	
Additional factors of interest	• Belief about online stores being better than retail for finding latest styles, designs and buying designer labels. • Preference for shops offering wider choice of brands as also display of all additional charges	
Neutral factors	• Low price and discounts • Preference for brands	
Factors of disinterest	• None	
Shopping orientation	• Significantly concerned about the possibility of credit card misuse • Positively influenced to some extent by the comfort and convenience factor	
Store selection orientation	• Has no disinterest in any of the store selection factors	
Demographic influences	• Males are significantly influenced by the preference for buying readymade items online • Non-working are significantly disinterested in buying ready-made garments from online stores. • Internet users spending more than 10 hours a week on the net; those who had brought any item from online stores, as also whose online purchases in a year exceeded Rs. 5000 are significantly influenced by preference for buying readymade garments from online stores • Students as also those who have used Internet for less than three years are influenced significantly by the absence of touch and feel factor. • Those who spent more than 10 hours a day on the Internet are influenced by the quality assessment tools and information factor	
Most likely to visit	• Any online store selling readymade garments and offering facilities for price comparisons, quality, size and fit assessment.	

This table brings together information about 1) the significance of each of the factors extracted by the PCA within each cluster, 2) additional factors extracted by the cluster based PCA, 3) analysis of the apparel selection orientation of the cluster, 4) shopping orientation of the cluster and 5) demographic influences if any, affecting the cluster through significant variables. A closer analysis showed that Cluster 2 had no positive influences except those produced by the CB-PCA. This cluster displayed a pattern of influences, which are difficult to interpret indicating that there are consumers who are still undecided about their etail preferences.

V. CONCLUSION

While market segmentation research has grown phenomenally in recent times, particularly with the evolution of additional analytical tools in the domains of Artificial Intelligence (AI), Machine Learning (ML) and Data Science (DS), the gaps between academic outputs and the needs of the practitioner remain largely unaddressed. The core question that must be answered is: Are businesses market takers or market seekers and makers. The reliance on the RFM paradigm, while necessary to retain and profit from the existing customers, will only make the ecommerce companies introverted. Most market research organizations predict the B2C revenues to reach US$ 7 trillion with its share in total retail rising to 27 to 29% by 2027-28. This implies a massive addition new customers particularly from the younger lot. Future growth and competitiveness of ecommerce companies will thus hinge upon, how well they understand their potential and prospective customers. Whatever may be the extent of commerce is better served by the 4 Cs based consumer segmentation as opposed to the traditional market segmentation. Factor score based two-step clustering would certainly provide deeper insights into the orientations, preferences and expectations of consumers.

REFERENCES

[1] Dolnicar, S., Grün, B., Leisch, F. (2018). Market Segmentation. In: Market Segmentation Analysis. Management for Professionals. Springer, Singapore. https://doi.org/10.1007/978-981-10-8818-6_1

[2] Yoram (Jerry) Wind and David R. Bell. (2006) Market segmentation. In: The Marketing Book, Edition 6. (Eds) Michael Baker, Susan Hart. Routledge. DOI https://doi.org/10.4324/9780080942544

[3] Camilleri, M. A. (2018). Market Segmentation, Targeting and Positioning. In Travel Marketing, Tourism Economics and the Airline Product (Chapter 4, pp. 69-83). Springer, Cham, Switzerland.

[4] Tynan, C. and Drayton, J. (1987) Market segmentation. Journal of Marketing Management, 2 (3) pp. 303–35.

[5] McDonald, M., Christopher, M., Bass, M. (2003). Market segmentation. In: Marketing. Palgrave, London. https://doi.org/10.1007/978-1-4039-3741-4_3

[6] Nadube, P. M., & Didia, J. U. D. (2018). Market targeting and strategic positioning. International Journal of Marketing Research and Management, 8(1), 32-45.

[7] Rozhko, Viktor. (2023). Justification of consumer market segmentation as a mandatory tool of strategic marketing. Technology audit and production reserves. 2. 15-19. 10.15587/2706-5448.2023.277373.

[8] Lizzie Davey. (2023). What Is Customer Segmentation? Retrieved from https://www.shopify.com/blog/what-is-customer-segmentation

[9] Salminen, Joni & Mustak, Mekhail & Sufyan, Muhammad & Jansen, Jim. (2023). How can algorithms help in segmenting users and customers? Journal of Marketing Analytics. 1-16. https://doi.org/10.1057/s41270-023-00235-5

[10] Jain, Reetika. (2019). Analysis of Indian Consumers Behaviour using Lifestyle Segmentation. Journal of Business Thought. 10. 57-65. 10.18311/jbt/2019/23573.

[11] Vyncke, Patrick (2002) Lifestyle Segmentation: From Attitudes, Interests and Opinions, to Values, Aesthetic Styles, Life Visions and Media Preferences. European Journal of Communication. Vol.17, Issue 4. pp 445-463. doi:10.1177/0267323102017004030l

[12] G. G. Rousseau & J. Kruger (1990). Testing the value and life-style model (VALS) of psychographic market segmentation. Journal of Industrial Psychology. Vol.16(1) pp 31-37 DOI:10.4102/sajip.v16i1.502

[13] Herrero-Crespo, Ángel & Pérez, Andrea & Bosque, I.R.. (2014). Values and lifestyles in the adoption of new technologies applying vals scale. Academy of Marketing Studies Journal. Volume 18, Number 2, pp 29-48.

[14] Aaron Ahuvia, Barbara Carroll, Yi Yang. (2006) Consumer Culture Theory and Lifestyle Segmentation. Innovative Marketing, Volume 2, Issue 4, pp 33-43

[15] Claritas Prizm Premier Segment Narratives (2022). https://environicsanalytics.com/docs/default-source/us---data-product-support-documents/claritas-prizm-premier-segment-narratives-ea.pdf

[16] Bronwyn Higgs & Allison C. Ringer. (2007) Trends in Consumer Segmentation. https://api.semanticscholar.org/CorpusID:21123174}

[17] Cooil, Bruce & Aksoy, Lerzan & Keiningham, Timothy. (2007). Approaches to Customer Segmentation. Journal of Relationship Marketing. 6. 9-39. 10.1300/J366v06n03_02.

[18] Wedel, M., Kamakura, W.A. (2000). Tools for Market Segmentation. In: Market Segmentation. International Series in Quantitative Marketing, vol 8. Springer, Boston, MA. https://doi.org/10.1007/978-1-4615-4651-1_4

[19] Johannes, H.V. Bases of market segmentation success: A marketing decion makers' perspective. International Journal of Business and Administrative Studies 2016, 2(3), 75-80

[20] Yankelovich, Daniel & Meer, David. (2006). Rediscovering market segmentation. Harvard business review. 84. 122-31, 166.

[21] Mary, O. (2011) 'Factor scores, structure and communality coefficients: A primer'. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, February 3, 2011.

[22] Christine, D. Min, Z. and Diana, M. (2009) 'Understanding and Using Factor Scores: Considerations for the Applied Researcher'. Practical Assessment, Research & Evaluation [online]. Volume 14, Number 20, October 2009

[23] Kieffer, Kevin M. (1998) Orthogonal versus Oblique Factor Rotation: A Review of the Literature regarding the Pros and Cons. Paper presented at the 27th Annual Meeting of the Mid-South Educational Research Association New Orleans, LA, November 4-6, 1998.