Clustering Consumers Through their Consumption Behavior: Analysis on the Fashion Industry

Andreea-Ionela PUIU[1]

ABSTRACT

The actual paper aimed to confirm the validity of the consumer styles inventory scale, developed by Sproles and Kendall (1986) on people aged 18-45 years based on their consumption behaviour when it comes to the clothing industry. In this respect, I applied the consumer styles inventory scale on individuals from the Romanian space aged 18-45 years, the responses being collected via online platforms. To the gathered data were applied, in the first part, an exploratory analysis, followed by confirmatory factor analysis. In the second part, the K-means cluster analysis was used to segment individuals in distinct groups based on their demographic characteristics and their financial involvement in purchasing fashion items. The output of the k-means cluster analysis distinguished eight segments of consumers defined as: „casual consumer", „quality-seeker consumer", „irrational consumer", „indifferent consumer", „habitual consumer", „occasional consumer", „fashion consciousness consumer" and „passive consumer".

KEYWORDS: *cluster analysis, consumer behaviour, consumer styles inventory, fashion industry.*

JEL CLASSIFICATION: *D12, D90, C38, L67.*

1. INTRODUCTION

The identification of factors that alter the decision-making process of the consumer was of interest for the researchers in consumer behavior. Those decision-making features are self-characteristics of consumers that are of interest if those remain constant over time, becoming significant determinants of consumer behavior.

There were developed several theories to explain and to predict the behavior of the consumer and to classify them based on their decision-making techniques. One of the most universally applied frameworks is the consumer styles inventory that grouped consumers based on their decision-making styles in the eight-factor model. The consumer style inventory was tested and validated in multiple societies to get a better knowledge of cross-cultural consumer decision-making processes.

The present research proposes to test the validity of the consumer styles inventory scale on the individuals aged 18-45 years and to cluster respondents in distinct groups based on their demographic characteristics and their financial involvement in purchasing fashion items.

The article is designed as follows: the next section is dedicated to presenting the theoretical background in consumer decision-making. The third section outlines the data and the methodology, while section four presents the results. The last sections conclude and outlines some limits of the research.

---

[1] Bucharest University of Economic Studies, Romania, e-mail: andreeaipuiu@gmail.com.

2. LITERATURE REVIEW

The consumers' decision-making process was described as a mental orientation toward realizing selections considering the occurrence of multiple available alternatives in the market.

In the acquisition process of any type of goods and services, several essential factors influence personal buying decision-making, the importance of each factor varying by market segments. In this respect, the acquisition of fashion items is recognized as a complex phenomenon.

The specialized literature realized a distinction among multiple sorts of approaches in classifying several psychological traits related to consumer behavior. Some of the most well-known are the consumer typology approach (Moschis, 1976), that classifies consumers in several classes, lifestyle approach (Lastovicka, 1982), that identifies diverse psychological peculiarities related to consumer behavior. Then is the diffusion of innovation model (Rogers, 2003), that classifies consumers based on their receptiveness to innovations relative to other consumers and their decision approaches, that concentrate on appropriate cognitive features of consumer decision-making (Sproles and Kendall, 1986).

I constructed the actual study based on the consumer styles inventory instrument developed by Sproles and Kendal (1986), mainly because of two reasons. Firstly, the validity of the consumer styles inventory scale was examined in various dissimilar cultural and economic circumstances, recognizing it as one of the most widely used instruments for the diagnosis of the decision-making process of consumers. This argument is supported by the fact that the instrument was included in studies conducted in diverse societies, like US (Kang et al., 2013; Leng & Botelho, 2010), Germany (Bauer et al., 2006), Swedish (Andersson et al., 2016) or China (Siu et al., 2001).

Secondly, the consumer-style inventory theory was the first framework that elaborated the consumer features approach within consumer behavior. Moreover, it was recognized as a widely utilized instrument to cluster consumers, being, in this case, a conventional tool for investigating decision-making behavior in diverse contexts.

To conceptualize the consumer decision-making styles, Sproles and Kendall (1986) developed the consumer styles inventory scale. Their exploratory factor analysis revealed the existence of eight-factor models identified through a selection algorithm that encompasses four criteria presented in the following. The first criteria of the selection algorithm is to include the mental characteristic of a consumer decision-making process that is among the essential consumer features. The second criterion specified that the algorithm should contain autonomous and comprehensive decision-making features. The third criteria specified that the algorithm should evaluate and measure each consumer on all components to profile the consumer into several styles, while the last criteria said that the method should be of importance for consumer-interest professionals (Sproles & Kendall, 1986).

Based on this criterion were obtained the eight factors that measure perfectly independent decision-making characteristics of consumers. The research identified the perfectionist consumer, characterized as the individual that search for the best quality of the products that he buys, the brand-conscious consumer, characterized as the one that prefers the most expensive and famous brands, the novelty conscious consumer is the type of consumer that is prone to accept the latest innovations introduced in the industry, also being in style is very important for him. Recreational consumers find shopping as a relaxing and entertaining activity, while the price-conscious consumer looks for items at discounts and they prefer to buy clothes during the discounts period. Impulsive consumers do not organize their shopping sessions and they are inclined to regret their irrational spending. The consumer that feels confused by the multiple available options manifests a lack of self-confidence and encounters difficulties in realizing choices. The last category, the brand-loyal consumer is the one that has favorite brands and store.

As previously specified the scale was tested in various distinct cultural and economic contexts, different factor solutions being obtained, and some-specific structures were more prominent depending on the country characteristics.

The actual study proposes to test the validity of the consumer styles inventory scale on the individuals from Romania, aged 18-45 years. In this respect, were used only six dimensions from the Sproles and Kendall instrument (1986) namely brand consciousness dimension, confused by overchoice, perfectionist, fashion-conscious, recreational, and impulsive consumers dimensions. Also, in the second part, the K-means cluster analysis was used to segment individuals in distinct groups based on their demographic characteristics and their financial involvement in purchasing fashion items.

3. METHODOLOGY

3.1 The Instrument and the Data

The actual paper proposes to estimate the consumption styles of individuals aged 18-45 years, using six dimensions of the consumer styles inventory scale, developed by Sproles and Kendall (1986). The scale was translated in Romanian and back translated into English to provide coherence and equivalence of meanings.

To the collected data I applied, in the first step, exploratory factor analysis and confirmatory factor analysis to verify if the items load correctly to the corresponding identified factors since this scale was used in a different cultural context and were reached different factor distributions. Also, I used the confirmatory factor analysis to verify the number of dimensions obtained from the exploratory analysis and to demonstrate the unidimensionality of the scale.

In the second part, I performed a k-means cluster analysis to distinguish consumers profiles. To cluster consumers, based on their consumption behaviours, I used demographics characteristics, involvement in purchasing fashion items and their financial situation.

3.2 Sample and Data Collection

I collect the data used in the present research through a survey, designed in Google Forms Application, questionnaire applied to people aged 18-45 years that follow an education system at the moment when they participated to this study. In this respect, I gathered a total of 300 responses, mainly people aged 18-25, from the urban area. The demographic characteristics of respondents are presented in the following table (Table 1).

Table 1. Demographic Statistics

Variable	Categories	N = 300	Percentage (%)
Age	18-25	229	76.5%
	26-30	42	13.9%
	31-35	9	3.0%
	36-40	9	3.0%
	41-45	11	3.6%
Gender	Male	259	86.33%
	Female	41	13.67%
Residence	Urban	232	77.33%
	Rural	68	22.67%

Source: Author's calculus.

Firstly, I studied the construct validity of the measurement using the exploratory and confirmatory factor analysis. Furthermore, K-means cluster analysis was employed to group respondents to distinguish their consumer profile. To perform all the proposed analysis, I used the R Statistical and Programming Software.

3.3 Data Analysis and Results

In the first part was conducted exploratory factor analysis using the *lavaan (Rossel, 2012)* package from the R library, where six-factor were obtained. The primary factor explained 10.3% of the variance, while the remaining five factors accounted for 9.5%, 8.6%, 8.2%, 8.1% and 5.8%. The obtained dimensions registered Cronbach Alpha scores among 0.78 and 0.88. Therefore, I performed confirmatory factor analysis of the obtained dimensions using the maximum likelihood estimator. The six-factor model previously tested fit the data well, registering good scores of the fit indices (X^2 = 614.567/p-value < 0.05, GFI = 0.974, AGFI = 0.977, TLI = 0.879, CFI = 0.895, RMSEA = 0.067, SRMR = 0.063). Table 2 presents the results of the confirmatory factor analysis model.

Table 2. The factor loadings and constructs reliability

Construct	Item	Loadings	Reliability
Brand Consciousness	The well-known national brands are best for me.	1.419	
	The more expensive brands are usually my choices.	1.364	0.85
	The higher the price of a product, the better its quality	1.044	
	Nice departments and specialty stores offer me the best products.	0.937	
	I prefer buying best-selling products.	1.016	
	The most advertised brands are usually exceptionally good choices.	0.956	
Confused by overchoice	There are so many brands to choose from that, often I feel confused	1.523	0.87
	Sometimes it is hard to choose which stores to shop.	1.571	
	The more I learn about products, the harder it seems to choose the best.	1.607	
	All the information I get on different products confuses me.	1.401	
Perfectionist	Getting good quality is important to me.	1.009	
	In general, I usually try to buy the best overall quality.	1.220	
	I make a special effort to choose the best quality products.	1.265	0.78
	I really do not give my purchases much thought or care.	-0.654	
Fashion-conscious	I usually have one or more outfits of the very newest styles	1.498	0.88
	I keep my wardrobe up to date with the changing fashions.	1.284	

Construct	Item	Loadings	Reliability
Recreational	Fashionable, attractive styling is important for me.	1.454	
	Shopping is not a pleasant activity for me.	1.473	
	Going shopping is one of the enjoyable activities in my life.	-1.529	0.80
	Shopping the stores wastes my time.	1.130	
	I enjoy shopping just for the fun of it.	0.934	
Impulsive	I should plan my shopping more carefully than I do.	1.256	0.80
	I am impulsive when purchasing.	1.349	
	Often, I make careless purchases I later wish I had not.	1.318	
	I carefully watch how much I spend.	-0.536	

Source: Author's calculus.

In the second part, to group respondents based on their consumption styles, I performed K-means cluster analysis.

To execute this analysis were used the following packages from the R library; *tidyverse* (Wickham, 2017) for data manipulation, *cluster* (Maechler et al., 2019) for clustering algorithm and *factoextra* (Kassambara & Mundt, 2019) for clustering algorithm and data visualization.

Before running the analysis, data were standardized. To cluster observations, there is required to measure the distance between each pair of observations. The selected distance settles the association of two elements and impacts the shape of clusters. For the present research was used the Euclidian distance.

K-means clustering procedure is the most employed machine learning algorithms of segmentation a given data set in a set of *k* clusters, where *k* designates the number of clusters designates by the analyst. This method lists objects in multiple associations such that objects from the corresponding clusters are as similar as possible (high intra-group similarity), while objects from different clusters are as different as possible (low inter-group similarity).

Therefore, the basic idea behind the k-means clustering technique is that clusters should be defined such that total-within cluster variation is minimized.

Before starting the algorithm, the analyst should specify the number of clusters, being recommended to test several distinct values of *k* and to compare the results that were obtained. In this respect was performed k-means cluster analysis with 2,3,4 and 5 clusters, graphical representations being presented in the following (Figure 1).
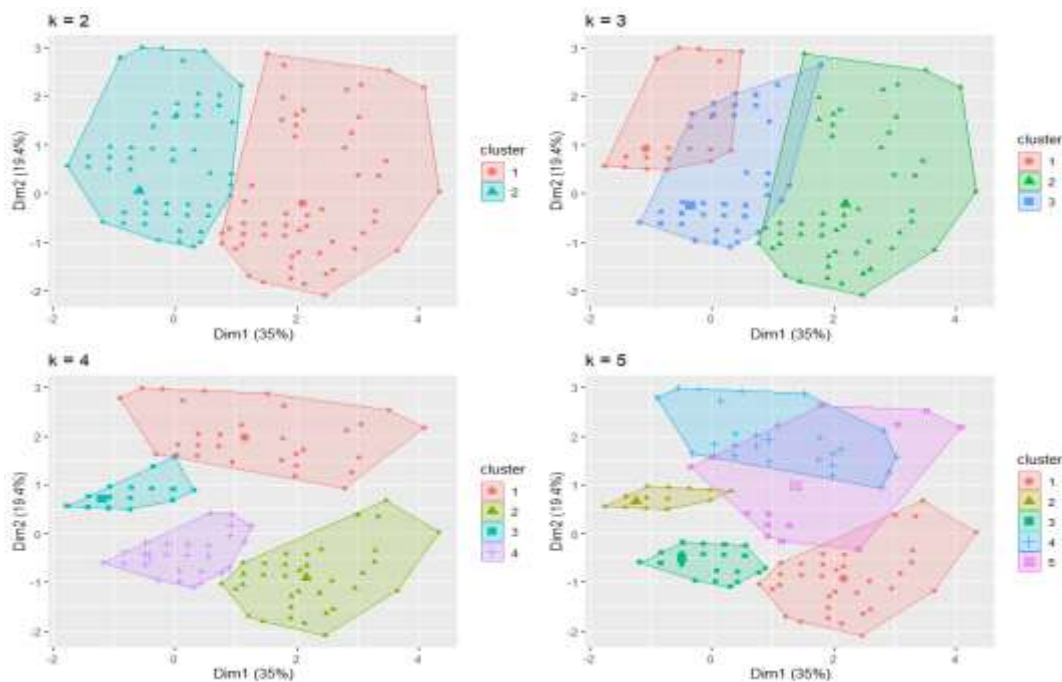
Figure 1. Perform k-means cluster with different k values

Source: Author's Representation Using the R Statistical and Programming Software

The visual representation tells us where the overlapping occurs in clusters (e.g. k = 3, k = 5) or when it does not occur (e.g. k = 2, k = 4), but it doesn't tell anything about the ideal number of clusters for the actual data.

To settle the ideal number of clusters could be applied three distinct approaches: elbow method, silhouette method and gap statistic. In the case of the elbow procedure, I computed the clustering algorithm for different values of *k*, for each *k* calculating the total within-cluster sum of squares. Therefore, I plotted the total within-cluster sum of squares *k*, the inflexion point, in the graph, being usually considered an indicator of the optimal representation of clusters. In the present case, the elbow method suggests that the optimal number of clusters is eight.
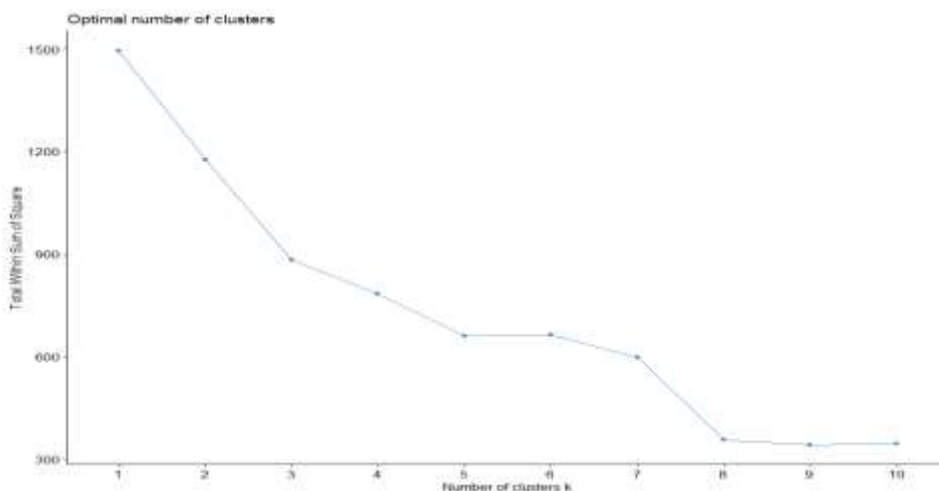

Figure 2. The optimal number of clusters – Elbow Method

Source: Author's Representation Using the R Statistical and Programming Software

The silhouette method estimates the quality of clusters, deciding how properly fits an object within its cluster. A high average silhouette width designates a good clustering. This method showed that four clusters maximize the average silhouette values (Figure 3), while eight clusters come in as the second optimal number of clusters (Figure 4).
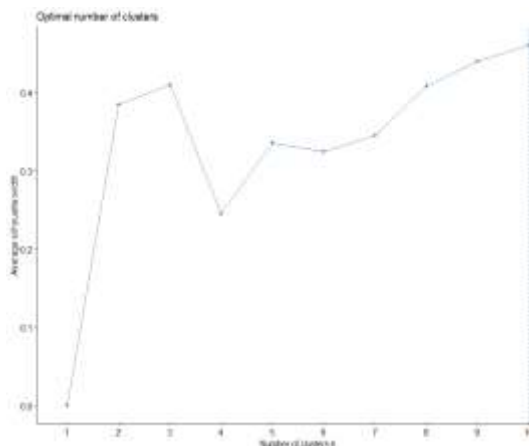

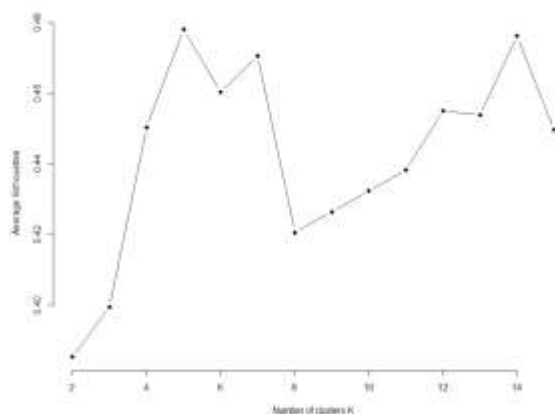**Figure 3. Maximize the average silhouette**   **Figure 4. The optimal number**
                                                        of clusters - Silhouette method

Source: Author's Representation Using the R Statistical and Programming Software

The gap statistic method matches the total-within cluster variation for different value of *k* with their expected values under the null reference distribution of the data. Also, this approach suggests the optimal number of clusters as being eight (Figure 5).
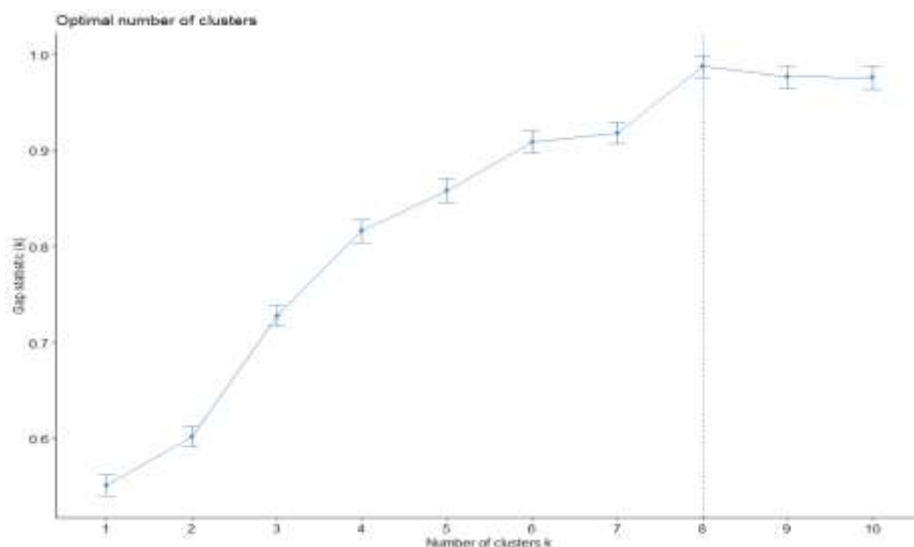

Figure 5. Optimal number of clusters – Gap Statistic Method
Source: Author's Representation Using the R Statistical and Programming Software

Considering that all the previous methods suggests eight as the number of optimal clusters, I performed the final analysis and extract the results using eight clusters (Figure 6).
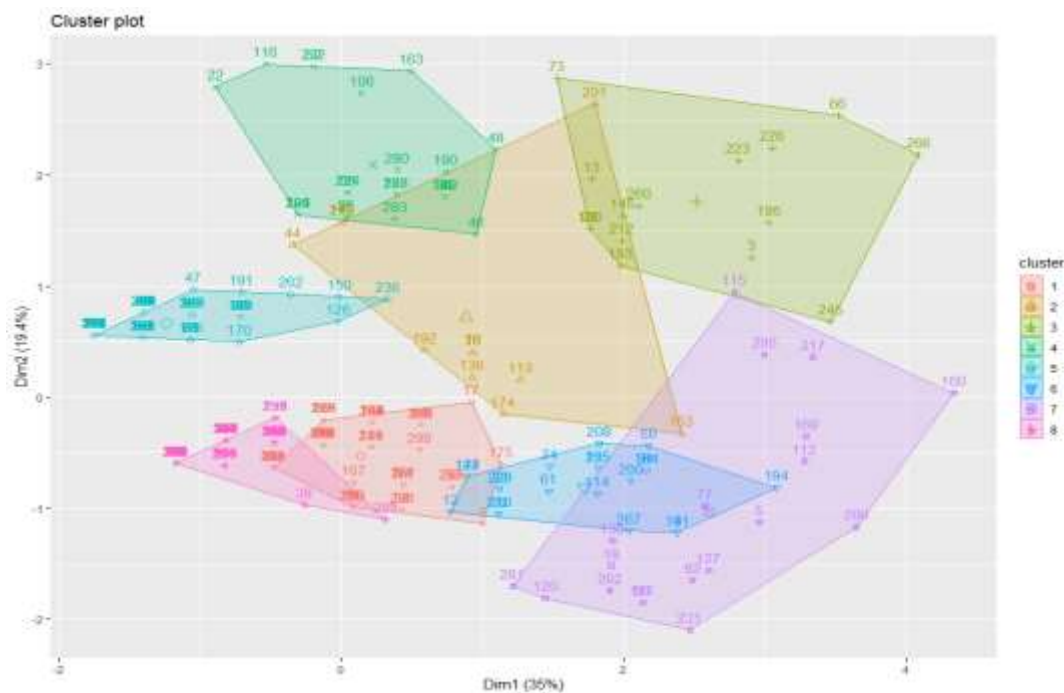
Figure 6. Cluster Plot
Source: Author's Representation Using the R Statistical and Programming Software

Information about each cluster regarding the distribution of demographics characteristics, involvement in purchasing fashion items and financial situation within each cluster were presented in the following.

Cluster 1 – Casual Consumer: people from this segment place a moderately low value on shopping of fashion items and they are somehow indifferent to brands. They go for shopping when they feel a need, but otherwise, they do not spend lots of money on shopping for clothes. This cluster is constituted, mostly, by females, with an average monthly income of $785 and expenditures on clothes on average of $67 per month.

Cluster 2 – Quality Seeker Consumer: this segment is composed entirely of men with an average age of 35 years. All of them reported monthly incomes higher than $1345 and expenditures on clothes higher than $200 per month. Quality seekers consumers are willing to invest more money on their image, they search for best products, with the high overall quality.

Cluster 3 – Irrational Consumer: the segment of people from this cluster is constituted preponderantly by females with incomes lower than $450 per month and spending's on clothes higher than $200 per month. People from this cluster spend half or more of their revenues on clothing acquisition, manifesting, in this respect, an irrational and compulsive behavior towards shopping fashion items.

Cluster 4 – Indifferent Consumer: people from this cluster are, in general, women with a mean age of 30 years, with relatively low incomes, that do not exceed $450 per month and that spend, in average $45 per clothing acquisition/per month. We could infer that people from this category manifest the most rational and balanced behavior because they spend on clothes proportionally with their monthly revenues.

Cluster 5 – Habitual Consumer: people from this cluster are mostly men, from the rural area, aged 18-25 years, with monthly incomes that oscillate among $450-$670, with spending's on clothes that do not exceed $67/per month. This type of consumer is the most cautious one, investing money on clothes only there is a need.

Cluster 6 – Occasional Consumer: in this cluster were mainly female, aged 35-45 years, from the urban area, with an average income of $785 and expenditures that vary among $35 to $90 per month.

Cluster 7 – Fashion Consciousness Consumer: for people from this cluster it is important to be in style, shopping being an important part from those people purpose of being fashionable. The segment of people from this category is mainly composed of females, aged 40-45 years, from the rural area, with a mean income of $1120 and average monthly spending in clothes of $157.

Cluster 8 – Passive Consumer: People from this cluster spend on clothes among $67-$90 per month, registering revenues that vary from $675 to $112. In general people from this cluster are females of 36-45 years.

4. CONCLUSIONS

The purpose of this research paper was to contribute to the existing literature by investigating the validity of consumer styles inventory scale on people aged 18-45 years and because of convenience, the sample was constituted of Romanians. However, the unbalanced gathered sample do not permit the generalizability of the results to the entire population.

The aim was to test the applicability of six dimensions of consumer styles inventory scale, namely, brand consciousness, confused by over choice, perfectionist, fashion-conscious, recreational, impulsive.

The gathered results revealed the good performance of the investigated dimensions through the results of the exploratory and confirmatory factor analysis, even if the segment of the population of interest was different of the one tested in the initial article.

In the second part, the actual paper proposed to segment individuals in dissimilar groups based on their demographic characteristics and their financial involvement in purchasing fashion items. The output of the k-means cluster analysis recommended the ideal number of clusters as being eight. Therefore, were identified the following typologies of consumers: „casual consumer ", „quality-seeker consumer", „irrational consumer", „indifferent consumer", „habitual consumer", „occasional consumer", „fashion consciousness consumer", „passive consumer". Again, the obtained clusters are specific to the characteristics of the gathered sample.

However, there are some characteristics of the gathered clusters that could be connected to some Romanian social contexts.

First of all, the quality seeker consumer represents that typology of consumer that worked to obtain a higher level in the social status hierarchy and to confirm his position and to own the respect of the others he is willing to invest more money in clothing acquisition because this investment is somehow refunded from the fact that he imposes respect through the clothes that he is wearing. In other words, people from this group perceive clothing acquisition as being an investment.

On the other side, the habitual consumer represents a group that is formed mainly from men, that gain high revenues, but that are not likely to invest so much money on their styles. Therefore, this segment is somehow caution and more rational, considering that money was acquired working hard and they do not believe that the spending is justified.

The last cluster that I would like to discuss is the group of the irrational consumer, composed of that type of people who spend half or more of their earnings on clothing acquisition. In this cluster could be incorporated also people that take a loan to buy clothes, even if this does not represent a necessity.

LIMITATIONS AND FUTURE RESEARCH

When generalizing the results obtained in this research, there are some aspect that should be considered. Firstly, the dimension of the sample is relatively small reported to the Romanian population of resident peoples aged 18-45 years, therefore the generalizability of the results is decreasing.

Secondly, the characteristics of individuals that are part of specifically identified clusters are context-dependent and are specific to the individuals that were part of this study. To permit generalizability of the group diagnosis should be studied diverse groups of individuals.

REFERENCES

Andersson, A., Hallberg, E. & Caroline, I. (2016). Examining the Applicability of the Consumer Style Inventory in a Swedish Context a Qualitative Exploration of Male Generation Y Students on Fashion Wear.

Bauer, H. H., Sauer, N. E. & Becker, C. (2006). Investigating the relationship between product involvement and consumer decision-making styles. *Journal of Consumer Behaviour*, 5(4), 342-354.

Kassambara, A. & Mundt, F. (2019). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R package version 1.0.6.

Lastovicka, J. L. (1982). On the validation of lifestyle traits: a review and illustration. *Journal of Marketing Research*, 19(1), 126-138.

Leng, C. Y. & Botelho, D. (2010). How does national culture impact on consumers decision making styles? A cross cultural study in Brazil, the United States and Japan. *BAR-Brazilian Administration Review*, 7(3), 260-275.

Kang, J. Y. M., Johnson, K. K. P. & Wu, J. (2014). Consumer style inventory and intent to social shop online for apparel using social networking sites. *Journal of Fashion Marketing and Management*, 18(3), 301-320.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2019). C*luster: Cluster Analysis Basics and Extensions,* R package version 2.1.0.

Moschis, G. P. (1976). Shopping orientations and consumers uses of information. *Journal of Retailing*, 52(1), 61-70.

Rogers, E. M. (2003). *Diffusion of innovations*. New York: The Free Press.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36.

Siu, N.Y.M., Wang, C.C.L., Chang, L.W.K. & Hui, A. S.Y. (2001). Adapting Consumer Style Inventory to Chinese Consumers. *Journal of International Consumer Marketing*, 13(2), 29-47.

Sproles G. B. & Kendall E. L. (1986). A methodology for profiling consumers' decision-making styles. *The Journal of Consumer Affairs*, 20(2), 267-279.

Wickham, H. (2017). tidyverse: Easily Install and Load the 'Tidyverse.