Clustering Internet Shoppers: An Empirical Finding from Indonesia

Fadhilah S. Minako
Department of Industrial Engineering,
Diponegoro University
Jl. Prof. Soedarto, S.H., Tembalang
Semarang, Indonesia 50275
+62 24 7460052
fadhilahminako@gmail.com

M. Mujiya Ulkhaq
Department of Industrial Engineering,
Diponegoro University
Jl. Prof. Soedarto, S.H., Tembalang
Semarang, Indonesia 50275
+62 24 7460052
ulkhaq@live.undip.ac.id

Dwi 'Sa Nu
Department of Industrial Engineering,
Diponegoro University
Jl. Prof. Soedarto, S.H., Tembalang
Semarang, Indonesia 50275
+62 24 7460052

Angela R. A. Pratiwi
Department of Industrial Engineering,
Diponegoro University
Jl. Prof. Soedarto, S.H., Tembalang
Semarang, Indonesia 50275
+62 24 7460052
angelaa.ratih@gmail.com

Pradita Y. Akshinta
Department of Industrial Engineering,
Diponegoro University
Jl. Prof. Soedarto, S.H., Tembalang
Semarang, Indonesia 50275
+62 24 7460052
pradita.yusi@gmail.com

ABSTRACT
The objective of this research is to do market segmentation of internet shoppers based on internet psychographics. There are thirty-eight indicators from six criteria, i.e., "internet shopping is easy and fun", "internet shopping is a hassle", "I don't know how", "fear of financial theft", "like the energy of brick-and-mortar stores", and "internet has good prices and quality". To do clustering, *k*-means cluster analysis was employed. Result shows that there are eight segments of internet shoppers, namely, shopping lovers, adventuresome explorers, business users, coward shoppers, suspicious learners, fun seekers, technology muddlers, and shopping avoiders. The first five segments are considered to be more likely to purchase products online, i.e., the online shoppers; while the rest three are the non-online shoppers. The ANOVA test confirmed that the eight segments were appropriate since it created more differentiated and consistent clusters. This research is expected to give a contribution both to the theoretical and empirical literature on customer segmentation where different marketing strategies could be generated for each segment.

CCS Concepts
• **Information systems application→ Data mining→ Clustering**

Keywords
Clustering; internet shoppers; *k*-means; segmentation.

1. INTRODUCTION
The internet today has developed rapidly to serve millions of users in the world with numerous purposes. The internet users are accounted with more than 40% of the world population and it is still increasing [1]. It is believed that the internet nowadays becomes the universal source of information for tons of people. Over the past decades, the internet has consolidated itself as a powerful platform that has changed the way people do business. With the aid of the internet, the companies could enhance themselves with to pursue competitive advantages. They also could interact with their customers by combining the internet and information technology to perform electronic commerce (e-commerce).

The e-commerce is then fundamentally changing the way people buy the products. Gradually, people have to learn how to act in an ever-changing electronic market environment. For some customers, shopping and purchasing products online have become part of their daily lives since online shopping gains several benefits, such as it could save time, offer a greater product selection, and allow for cost saving in terms of transportation, tax, and the price of the product as well [2]. However, for some people, they are still reluctant to buy products online due to various reasons, such as lack of knowledge on using the internet and fear of financial theft.

The firms that engaged in e-commerce business have to understand that the market is heterogeneous and there exist distinct groups or segment of customers which have different needs, characteristics, and behaviors. Therefore, they might require separate services or marketing strategies. Analyzing this phenomenon by segmenting the internet shoppers would allow the firms to formulate the design of more efficient marketing strategies. Once the market is segmented, the firms are able to target which segment can be served best and set the competitive positioning for the particular segment rather than trying to compete in an entire market [3].

Market segmentation is broadly used since it could help marketers to offer better services for a targeted segment of customers [4]. The application of market segmentation can be seen in various studies, for example: identifying the particularities of shopping centers [5], [6]; segmenting market in tourism sector [7], [8]; in ecological products [9]; and in commercial fields [10]. One of the most essential decisions in market segmentation is determining the most useful variable that will be used to identify different cluster of customers.

ICEMC 2019, May 22–24, 2019, Taichung, Taiwan

ACM ISBN 978-1-4503-7182-7/19/05…$15.00.
DOI: https://doi.org/10.1145/3332324.3332325

This research tried to utilize internet psychographics to cluster internet shoppers. It has been shown to be useful in identifying internet shoppers [11]. This research then tried to extend the scope of the previous study by conducting the research in Indonesia. Different characteristics of internet shoppers in different country could aid the generalization study for segmenting internet shoppers. Therefore, this research attempted to identify and classify types of online shoppers as well as to discuss the characteristics of each type. In Indonesia, the internet is no longer a new thing. According to the Ministry of Communication and Information Technology, in 2017, there are 143.26 million internet users in Indonesia or around 54.68% of the total population of Indonesia, which 49.52% of them on the age of 19 to 34 years old [12]. The retail e-commerce sales expenses 7.056 billion USD in revenues in 2017 and projected to reach the value of 16.475 in 2022 [13], such a promising market.

2. RESEARCH DESIGN

The objective of this research is to perform a segmentation to identify different clusters or groups of internet shoppers according to internet psychographic. Thirty-eight indicators of internet psychographic adapted from [11] were used in this research. These indicators are grouped into six criteria, i.e., "internet shopping is easy and fun", "internet shopping is a hassle", "I don't know how", "fear of financial theft", "like the energy of brick-and-mortar stores", and "internet has good prices and quality". The criteria and indicators are presented in Table 1. A survey was conducted in Semarang, a capital city of Central Java Province, to complete the above-mentioned objective. Thus, those six criteria had been modified since the research was conducted in Indonesia, which has different characteristics than the original study.

The first criterion is "internet shopping is easy and fun". It refers to how the internet is making it easier and fun to do shopping; for instance, there is no need for internet shoppers to go to the physical stores and sometimes buyers could get a merchandise if they buy a product online. The second criterion is "internet shopping is a hassle". It refers to how internet actually makes people feel more complex and harder to buy something from internet. It seems that sometimes it is difficult to estimate the quality of the product by only seeing it from the internet. Another unpleasant example is the lead time of shipping since it could take a long time for the product the shoppers purchased to come in their hands. The third criterion is "I don't know how". This is about why people do not like to do online shopping because they do not know how to do it. They sometimes feel that online shopping stores do not always give them the products they want, or the internet itself is hard to understand somehow for them.

The fourth criterion is "fear of financial theft". It refers to how people still feel worried and fear to buy products online. It could be that they worry about their credit card numbers or other private information to be stolen. Therefore, they think that all of their private information must not be given. The fifth criterion is "like the energy of brick-and-mortar stores". Some people are fond of shopping in physical stores since they like to "connect" with products they want to purchase, a feeling like smelling, touching, and hearing. In addition, people also like to go to the physical store with their friends and meet with the local stores due to the friendliness. The last criterion is "internet has good prices and quality". This criterion refers to the ability of online stores that offer shoppers with good price or promotion and better quality than the offline stores.

To do clustering, *k*-means cluster analysis was employed to distinguish and specify each cluster's characteristic by determining the relationship of those six criteria. This technique has been widely used in clustering analysis, see for example [14]–[16]. The method partitions the data into the cluster so that the data that has the similar characteristics are gathered in the same cluster and the data that has different characteristics are grouped in other clusters. The farther the distance between one to another cluster, the greater variant characteristics among groups, whereas the internal group consisted similar characteristics that were alike.

3. FINDINGS

The research was conducted in Semarang, a capital city of Central Java Province, Indonesia, to cluster internet shoppers into different market segments; and thus, it employed a questionnaire to be fulfilled by respondents. There are two parts of the questionnaires. The first part aims to collect demographic data of the respondents, such as age, gender, and occupation. The second part utilizes the thirty-eight indicators of internet psychographic mentioned in the previous section. All the indicators were measured using the 7-point Likert scale, ranging from 1 which represents strongly disagree to 7 which represents strongly agree. After administering the survey, we gained 316 valid answers from the respondents.

As mentioned in the previous section, *k*-means clustering procedure was applied to explore the cluster solution. The optimal number of clusters was found to be eight as they have the greatest differences among the groups. Validation procedure using ANOVA test shows that the eight groups was the best solution which it created more differentiated and consistent clusters. The largest group member was Group 4 (21.97%), followed by Group 7 (18.79%), Group 6 (17.83%), Group 5 (14.01%), Group 8 (11.46%), Group 4 (9.24%), Group 3 (4.14%) and the last is Group 1 (2.55%). Note that the differences were statistically significant since the *p*-value is 0.000.

The characteristics of the eight groups are as following, see and Figure 1. Table 2 shows the percentage of each segment based on their demographic variables while Figure 1 shows the radar charts for each segment. Note that the values in Figure 1 are the average values of the standardized values from the answers of the respondents. Note that for simplicity and readability, all values are multiplied by 10^{17} except the first group that is multiplied by 10^{3}.

The first group is called the shopping lovers. They are people who really fond of shopping, not only on the internet (purchasing products online), but also in physical stores. They tend to be competent in using computers and familiar with online shopping methods. Therefore, they think that internet shopping is easy and fun. However, even though they love to do online shopping, they disagree that shopping online could gain a product with good quality and price. This rationale was represented as people in this segment still like to shop in brick-and-mortar stores. Based on the survey, only one of eight people in this segment is a college graduate. Also in this segment, it is about 62.50% of the members have incomes between 1,000,000 IDR to 2,000,000 IDR.

The next segment is fun seekers. They represent individuals who prefer to purchase products through physical stores than purchase products on the internet. People in this segment strongly agree with statements that "I want my purchase to be absolutely private" and "I want to see things in person before I buy". These people do not think that internet shopping is a fun and convenient thing to do. This segment consists of 11.59% of married couple and about 36.23% of them are college graduates. Around 28.89% and 23.19% individuals have incomes from 1,000,000 to 2,000,000 IDR and between 2,000,000 to 3,000,000 IDR respectively.

The third group is the adventuresome explorers. It is one of the significant online buying segments. The adventurous explorers strongly agree with statements that "I go to the internet for reviews or recommendations", "I like having merchandise delivered to me at home", and "I search for the lowest price in everything I buy". This segment represents those people who find that internet shopping is such an adventure for them to explore. Besides, the member of this segment think that internet could offer products with good price and quality. Interestingly, in this segment, there is no married couple. About 23.08% of the members are college graduates. This segment only consists of shoppers with income below 3,000,000 IDR.

The business users, as the fourth segment, are more likely to make online purchasing for business-related work compared to other segments. Business users do not really mind with any issue that other segments struggle with, such as fear of financial information to be theft, or lack of trust to internet retailers. Although business users can be called as online shoppers, they do not find internet shopping is a fun activity to do since they do internet shopping only for work. This segment consists of 6.90% of married couple, one of the segments with only a few amounts of married couple. About 24.14% of them are college graduates and 48.28% of this segment are those people with income between 1,000,000 to 2,000,000 IDR.

The fifth segment is coward shoppers. They represent people who like to buy online but having the big fear of financial theft. However, people in this segment do not think that internet shopping is a hassle or find any difficulties of it. This segment consists of 34.09% male, 65.91% female, and 95.45% of married couple. About 4.55% of them are college graduates. Around 47.74% of people in this segment have incomes between 1,000,000 to 2,000,000 IDR.

Technology muddlers as the sixth group represent people who find the difficulties of online shopping. This segment is not considered as an attractive target market for online selling. People in this segment are also afraid of financial theft. They strongly agree with the statement either "I worry about my credit card number being stolen on the internet" or "I want my purchases absolutely private". This segment consists of 3.57% of married couple. About 17.86% of the members are college graduates. Around 51.79% people in this segment have incomes between 1,000,000 to 2,000,000 IDR.

Next, there is the suspicious learners as the seventh segment. This segment represents individuals who find internet shopping is something enjoyable, yet they are still lack on understanding on how to do transaction online. Therefore, the members of this segment might need a guidance on how to do online purchasing easily. These people also do not think that online buying is a hassle and they are not scared of purchasing products online although there might be possibility of financial theft. This segment consists of around 18.64% of married couple. About 28.81% of them are college graduates. Around 32.20% and 27.12% of the individuals in this segment have incomes below 1,000,000 IDR and between 1,000,000 to 2,000,000 IDR respectively.

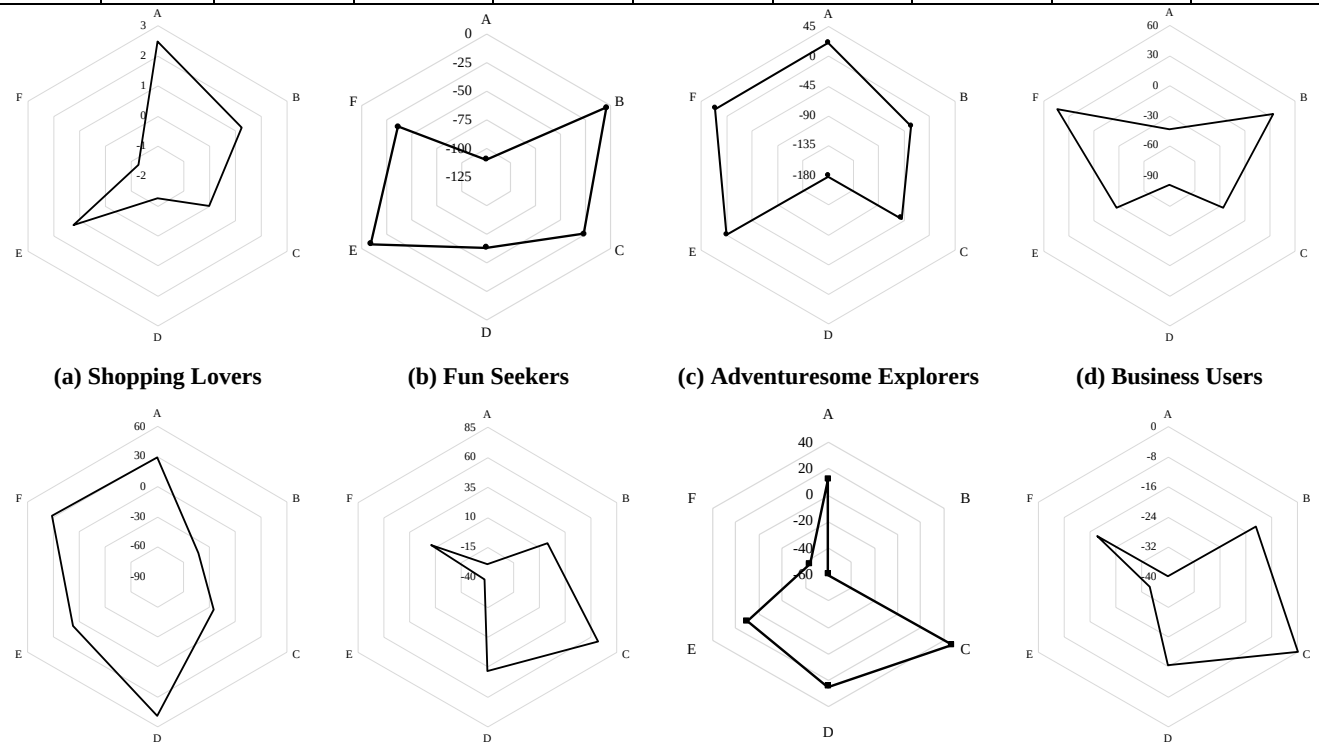
The shopping avoiders as the last segment represent people who really do not know how to do online shopping. People in this group dislike shipping charges on the internet and find the hassle of returning products that have been bought online. They want to view

Table 1. Internet Psychographics

No	Characteristics
A	**Internet shopping is easy and fun**
A1	I like that no car is necessary on the Internet
A2	I like not having to leave home when shopping
A3	Internet shopping is easier than local
A4	I like having merchandise delivered to me at home
A5	Online buying is fun
A6	I enjoy buying things on the Internet
A7	I'd shop more on the Internet if prices were lower
A8	Shopping in stores is a hassle
B	**Internet shopping is a hassle**
B1	I don't like waiting for products to arrive
B2	It's a hassle to return merchandise bought online
B3	It's hard to judge the quality on the Internet
B4	Internet buying has delivery problems
B5	I dislike shipping charges on the Internet
B6	Stores have better services policies
B7	I want to see things in person before I buy
B8	None of my friends shop on the Internet
C	**Don't know how to shop on the Internet**
C1	I don't know much about using the Internet
C2	I'm not good at finding what I want on Internet
C3	Internet ordering is hard to understand and use
C4	Internet stores don't carry things I want
C5	I go to the Internet for reviews or recommendations
C6	I like browsing on the Internet
C7	I go to the Internet to review products
D	**Fear of financial theft on the Internet**
D1	I worry about my credit card number being stolen
D2	I want my purchases absolutely private
D3	I don't want to give my credit-card-number
D4	Buying things on the Internet scares me
D5	I just don't trust Internet retailers
D6	I search for the lowest price in everything
E	**Like the energy of brick-and-mortar stores**
E1	I like to go shopping with my friends
E2	I like the energy at local retail stores
E3	I like friendliness of local stores
E4	I buy using layaway programs
E5	I often return items that I have purchased
F	**The internet has good prices and quality**
F1	Internet offers lowers prices than local stores
F2	Internet shopping offers better selection
F3	Internet has better quality than stores
F4	Local stores have better prices, promos

Table 2. Segmentation of Internet Customers

Demographics Variable		Shopping Lovers	Fun Seekers	Adventure-some Explorers	Business Users	Coward Shoppers	Techno-logy Muddlers	Suspicious Learners	Shopping Avoiders
Gender	Male	50.00%	24.64%	23.08%	34.48%	34.09%	32.14%	40.68%	22.22%
	Female	50.00%	75.36%	76.92%	65.52%	65.91%	66.07%	59.32%	77.78%
Married		12.50%	11.59%	0.00%	6.90%	95.45%	3.57%	18.64%	2.78%
Collage Graduate		12.50%	36.23%	23.08%	24.14%	4.55%	17.86%	28.81%	33.33%
Income (in million IDR)	< 1	12.50%	17.39%	15.38%	20.69%	11.36%	26.79%	32.20%	33.33%
	1 – 2	62.50%	28.99%	7.69%	48.28%	47.73%	51.79%	27.12%	41.67%
	2 – 3	12.50%	23.19%	38.46%	13.79%	15.91%	8.93%	8.47%	5.56%
	3 – 4	0.00%	2.90%	0.00%	10.34%	6.82%	1.79%	5.08%	5.56%
	4 – 5	0.00%	8.70%	0.00%	3.45%	11.36%	7.14%	10.17%	11.11%
	> 5	12.50%	18.84%	0.00%	3.45%	6.82%	3.57%	16.95%	2.78%



(a) Shopping Lovers **(b) Fun Seekers** **(c) Adventuresome Explorers** **(d) Business Users**

(e) Coward Shoppers **(f) Technology Muddlers** **(g) Suspicious Learners** **(h) Shopping Avoiders**

Figure 1. The Radar Charts for each Segment of Internet Shoppers

the products in person to know their quality before deciding to buy it. This segment consists of 2.78% of married couple. About 33.33% of the members are college graduates. Around 33.33% and 41.67% people in this segment have incomes below 1,000,000 IDR and between 1,000,000 IDR to 2,000,000 IDR.

4. CONCLUSION

The objective of this research was to determine the cluster or segment of internet shoppers based on internet psychographics. From the theoretical perspective, this study aims to give a contribution to the literature of segmentation study for segmenting internet shoppers. The result of the study determined that there are eight segments for internet shoppers, i.e., shopping lovers, fun seekers, adventuresome explorers, business users, coward shoppers, technology muddlers, suspicious learners, and shopping avoiders. The characteristics of each segment was presented in the previous section.

Several recommendations could be addressed to managers as well as businessmen that have e-commerce businesses. Establishing the strategical plan based on customers segmentation is considered as essential [17], [18]. According to the result of this study, Group 1 (shopping lovers), Group 3 (adventuresome explorers), Group 4 (business users), Group 5 (coward shoppers), and Group 6 (suspicious learners) are more likely to buy products online; hence, businesses could prioritize their advertising or marketing strategies to gain customer loyalty better than other segments.

Finally, the limitations of this study are twofold. First is related to the selected variables, i.e., internet psychographics. For the upcoming research, more variables have to be considered, such as motivation, values, lifestyle, etc. The second limitation is related to the area (location) of this research, i.e., the study was conducted in Semarang, Indonesia. Therefore, to make a generalization, the upcoming research should broaden and expand the research area.

5. REFERENCES

[1] Internet Live Stats. 2018. Available at: http://www.internetlivestats.com/internet-users/

[2] Ulkhaq, M. M., Wijayanti, W. R., Kusumawati, A., Aulia, F. S., Wijayanti, R. S., and Wiganingrum, R. 2017. Combining the eTransQual scale and importance-performance analysis to assess service quality of online shopping. In *Proceedings of the International Conference on Industrial Engineering and Applications* (Nagoya, Japan, April 21-23, 2017). ICIEA '17. IEEE, 146-150. DOI= http://10.1109/IEA.2017.7939196.

[3] Kotler, P., Armstrong, G., and Wong, V. 1996. *Principles of Marketing.* Prentice Hall, London.

[4] González-Hernández, E. M. and Orozco-Gómez, M. 2012. A segmentation of Mexican consumers based on shopping centre attractiveness. *Int. J. Retail Dist. Manag.*, 40, 10, 759–777. DOI= http://10.1108/09590551211263173.

[5] El-Adly, M. I. 2007. Shopping malls attractiveness: a segmentation approach. *Int. J. Retail Dist. M.*, 35, 11, 936-950. DOI= http://10.1108/09590550710828245.

[6] Gilboa, S. 2009. A segmentation study of Israeli mall customers. *J. Retailing and Consumer Services*, 16, 2 (March 2009), 135–144. DOI= http://10.1016/j.jretconser.2008.11.001.

[7] Cho, M., Bonn, M. A., and Brymer, R. A. 2017. A Constraint-Based Approach to Wine Tourism Market Segmentation. *J. Hosp. Tour. Res.*, 41, 4, 415–444. DOI = http://10.1177/1096348014538049.

[8] Iversen, N. M., Hem, L. E., and Mehmetoglu, M. 2016. Lifestyle segmentation of tourists seeking nature-based experiences: the role of cultural values and travel motives. *J. Travel. Tour. Mark.*, 33, sup1, 38–66. DOI= http://10.1080/10548408.2014.998359.

[9] Barr, S., Shaw, G., Coles, T. and Prillwitz, J. 2010. 'A holiday is a holiday': practicing sustainability, home and away. *J. Transp. Geogr.*, 18, 3, 474–481. DOI= http://10.1016/j.jtrangeo.2009.08.007.

[10] Sarki, I., Bhutto, N., and Arshad, I. 2012. Impact of Pakistani university student's cultural values and lifestyles on meaning of brands. *Interdiscip. J. Contemp. Res. Bus.*, 3, 9, 643–654.

[11] Swinyard, W. R. and Smith, S. M. 2003. Why people (don't) shop online: A lifestyle study of the internet consumer. *Psychology & Marketing* 20, 7, 567-597. DOI= http://10.1002/mar.10087.

[12] Kementerian Komunikasi dan Informatika Republik Indonesia. 2018. Jumlah pengguna internet 2017 meningkat, Kominfo terus lakukan percepatan pembangunan broadband [in Indonesia]. Available at: https://www.kominfo.go.id/content/detail/12640/siaran-pers-no-53hmkominfo022018-tentang-jumlah-pengguna-internet-2017-meningkat-kominfo-terus-lakukan-percepatan-pembangunan-broadband/0/siaran_pers.

[13] Statista. 2018. Retail e-commerce sales in Indonesia from 2015 to 2021 (in billion U.S. dollars). Available at: https://www.statista.com/statistics/280925/b2c-e-commerce-sales-in- indonesia/.

[14] Dhanachandra, N., Manglem, K., and Chanu, Y. J. 2015. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54, 764–771. DOI= http://10.1016/j.procs.2015.06.090.

[15] Wu, J., Liu, H., Xiong, H., Cao, J., and Chen, J. 2015. K-means-based consensus clustering: A unified view. *IEEE Transactions on Knowledge and Data Engineering*, 27, 1, 155-169. DOI= http://10.1109/TKDE.2014.2316512.

[16] Cohen, M. B., Elder, S., Musco, C., Musco, C., and Persu, M. 2015. Dimensionality reduction for k-means clustering and low rank approximation. In *Proceedings of the forty-seventh Annual ACM Symposium on Theory of Computing* (June, 2015). STOC '15. ACM, New York, 163–172. DOI = http://10.1145/2746539.2746569.

[17] Griva, A., Bardaki, C., Pramatari, K., and Papakiriakopoulos, D. 2018. Retail business analytics: Customer visit segmentation using market basket data. *Expert Syst. Appl.*, 100, 1–16. DOI= http://10.1016/j.eswa.2018.01.029.

[18] Kabadayi, S. and Paksoy, B. A. 2016. A segmentation of Turkish consumers based on their motives to visit shopping centers. *The International Review of Retail, Distribution and Consumer Research*, 26, 4, 456–476. DOI= http://10.1080/09593969.2016.1157513.