Use of Factor Scores in Multiple Regression Model for Predicting Customer Satisfaction in Online Shopping

Ali J. Ramadhan[1*], *S. R.* Krishna Priya[2], *N* Naranammal[2], *Rajani* Gautam[3], *Pradeep* Mishra[4], *Soumik* Ray[5], *Mostafa* Abotaleb[6], *Hussein* Alkattan[6] and *Zainalabideen* Albadran[1]

[1]University of Alkafeel, Najaf, Iraq
[2]PSG College of Arts and Science, Tamil Nadu, India
[3]Pandit S. N. Shukla University, Madhya Pradesh, India
[4]JNKVV College of Agriculture, Rewa, India
[5]Centurion University of Technology and Management, Odisha, India
[6]South Ural State University, Chelyabinsk, Russia

Abstract. Online shopping can be done from our convenient places like home, office, etc., and the product will be delivered to the respective places. There are many factors influencing online shopping. The purpose of this study is to develop a statistical model that is used to determine the factors that influence online shopping. In this study, using factor analysis five main factors have been obtained from 15 variables that influence online shopping. These five factors have significant effects on satisfaction of customers and accounted up to 56% of total variation. Using the factor scores as independent variables, multiple regression model has been developed for predicting customers satisfaction in online shopping. Customer satisfaction has been used as dependent variable in the regression model. The five main factors that contribute online shopping are: preference of consumers towards online shopping, the risk involved in purchasing products through online, time effectiveness in online shopping, difficulties faced during online shopping and getting products from trustworthy websites.

1 Introduction

Using a computer, smartphone, or other internet-connected device, a shopper can transact business with an online merchant. Customers can find their product of interest in online websites/retailers and can compare with other website/retailers to get better quality of products at lower cost. They can place a product on their shopping cart and can buy it any time. Customers shopping online require internet access and a valid payment method to make a transaction. Payment can be done through many ways. The most common way of payment is using debit or credit card transaction, net banking and cash on delivery. Once the payment is accepted, the product will be delivered to the address mentioned.

There are many advantages in online shopping. Online shopping is available at any time. Shopping can be done from any convenient place like home, office etc. It is very easy to shop

* Corresponding author: ali.j.r@alkafeel.edu.iq

and customers can compare the prices of products. There are more choices than buying in a retail store. Another advantage of online shopping is detailed information and detailed review. Online retailers allow customers to review the product they are going buy, so that they can check the reviews of other customers and can buy it at their interest. The biggest advantage of online shopping is placing the order for the product of interest can be done from home or office, and it will be delivered to the respective places, so that there is no need of going outside. It makes our work easy. One of the disadvantages of online shopping is sometimes delivery can be delayed. Inadequate long-term stock management causes shipping delays which lead to delay in delivery. Other disadvantages of online shopping are not getting the product ordered, damage of products, etc.

2 Overview

Comparisons have been made between offline and online purchasing, and the factors that influence online shoppers to purchase were discovered [1,2]. Internet personalities, perspectives, and practical considerations all played roles in shaping the online purchasing behaviours of today's youth [3,4]. The most critical aspects of online shopping were the customer's experience and the level of security they feel while making a purchase [5,6]. Customers' satisfaction and after-sales service were important metrics for both men and women when making online purchases [7,8]. Multiple Regression analysis using Factor scores were used to estimate body weight in Brown Trout, Broiler Chickens, Akkeci Kids, Morkaraman Sheep and Romano Lambs [9,10]. Norduz goats' daily milk output was also predicted using multiple linear regression analysis. [11,12]. In order to forecast faculty performance in higher education institutions factor scores were used in multiple regression model [13].

Till now, not much studies have been carried out using factor scores in multiple regression analysis for estimating customer satisfaction in online shopping. The present study has been conducted to overcome the above draw backs.

3 Materials and methods

3.1 Data description

Primary data has been used in the study which has been collected through the well-structured questionnaire. Data has been collected from area of Tiruchirappalli and Kovilpattiusing simple random sampling method. Data from 148 respondents have beencollected to conduct the study. 15 variables which influence the customer satisfactionhave been taken for the study. The variables are:Quality of information, Safety of the ordered products, Preference towards cash on delivery, Minimum purchase time, Easy comparison, Not getting the ordered product(s), Product Damage, Unworthy product(s), Duration for delivery, Choice between Online and Retail, Time saving, Cost Effectiveness, Difficulty in returning products, Ontime Delivery, Trustworthy websites.

3.2 Checking the adequacy of data

3.2.1 KMO test

To evaluate the usefulness of the data for component analysis, the Kaiser-Meyer-Olkin (KMO) test was conducted. Each model variable is tested for sufficient sampling. The value of KMO testranges from 0 to 1.

3.2.2 Bartlett's sphericity test:

The claim that the correlation matrix is an identity matrix can be tested with Bartlett's sphericity test.. For a factor analysis to be useful, it must have a significance level of less than 0.05.

3.3 Factor analysis

Understanding the correlation structure of the data and determining the most influential variables contributing to that structure are the goals of factor analysis. (Padro et.al., 1993). In the current study, factor analysis was used to determine what aspects of online shopping are most important to consumers.Using factor analysis,15 variables which are mentioned in the data description,have been reduced to 5 key factors which account formore than 56% of the total variance applying Varimax rotation.

4 Results and discussion

4.1 Checking the adequacy of data

KMO test and Bartlett's sphericity test have been conducted for adequacy checking of the data and the results are presented in table 1.

Table 1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.738
Bartlett's Test of Sphericity	Approx. Chi-square	348.169
	Df	105
	Sig.	0.01

From table 1, KMO measure is 0.73 which indicates that the value is acceptable and satisfactory for factor analysis.

From Bartlett's test of Sphericity, the associated probability is0.01(< 0.05)which makes correlation matrix as identity matrix. The results from both the tests indicate that the data is fit for factor analysis.

4.2 Factor analysis

Factor analysis is performed using the 15variables mentioned in table 2 and are reduced to 5 factors.

Table 2. Factor analysis results

Variables	Rotated factor loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor5
1.Quality of information	0.755				
2.Safety of ordered products	0.682				
3.Preference towards cash on delivery	0.658				
4. Minimum purchase time	0.442				
5.Easycomparison	0.453				
6.Not getting the ordered product(s)		0.762			
7. Product Damage		0.747			
8.Unworthy product(s)		0.647			
9.Duration of delivery			0.718		
10.Choicebetween Online and Retail			0.686		
11.Time saving			0.470		
12.Cost effectiveness				0.642	
13.Difficulty in returning products				0.547	
14. Ontime Delivery					0.833
15.Trustworthy websites					0.470

The results of table 2 indicatesfive factors are obtainedfrom 15 variables. The five key factors are preference of consumers towards online shopping, the risk involved in purchasing products through online, time effectiveness in online shopping, difficulties faced during online shopping and getting products from trustworthy websites.

The scree plot helps to determine the optimal number of factors. It is the graph of the Eigne values against all the factors.

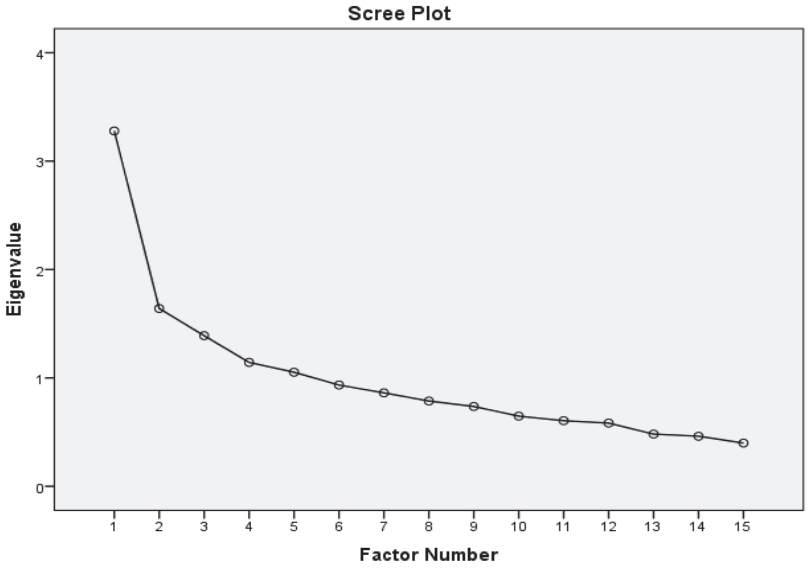

Fig. 1. Scree plot

Figure 1 displays the results, which indicate that the first five factors (out of a total of 15) are retained because their eigen values are larger than 1.

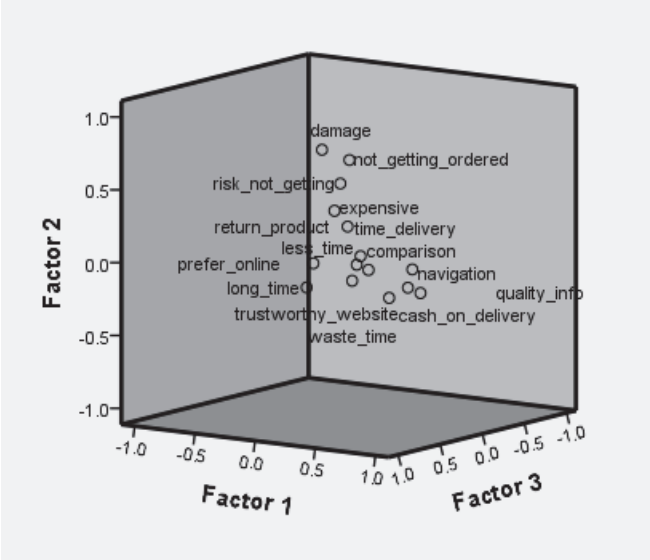

Fig. 2. Factor Plot in Rotated Factor Space

The factor diagram in a rotated factor space is shown in Figure 2. The figure displays the three readily apparent variables, while the other two can be revealed by rotating it. Results from component analysis are then used as independent variables in multiple regression analysis in the form of factor scores.

4.3 Multiple Regression Model

The Multiple regression model is developed to predict customer satisfaction in online shopping using the dependent variable, Y=Customer Satisfaction and independent variables

X_1=Preference of customers towards online shopping; X_2 = Risk involved in purchasing products through online; X_3 = Time effectiveness in online shopping; X_4 = Difficulties faced during online shopping;X_5 = Getting products from trustworthy websites.

Table 3. Results of t test for partial regression coefficients. (**- Significant at 1% level)

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Standard Error			
(Constant)	49.074	0.071		694.765	0.000
X_1	3.507	0.071	0.656	49.484**	0.000
X_2	2.470	0.071	0.462	34.854**	0.000
X_3	2.777	0.071	0.520	39.185**	0.000
X_4	0.388	0.071	0.073	5.478**	0.000
X_5	1.257	0.071	0.235	17.734**	0.000

The significance of the regression coefficients is tested using t-test. The results are presented in table 3. The results show that all the partial regression coefficients are significant at 1% level which indicate that all the five factors have a great influence on the customer satisfaction.

Table 4. Estimated Multiple Regression Equation

Model	Adj R^2	Durbin-Watson
$Y= 49.074+3.507X_1+2.470X_2+2.777X_3+0.388X_4+1.257X_5$	0.974	2.139

The results of table 4 show that the adjusted R square value is 0.974, which indicates that the independent variables explain 97.4 percent of the variation in customer satisfaction. The value of Durbin-Watson statistic is 2.139, indicating that there is no auto correlation.

Table 5. Results of One-way ANOVA

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	4093.329	5	818.666	1108.696**	0.000
Residual	104.853	142	0.738		
Total	4198.182	147			

**- Significant at 1% level

Table 5 displays the outcomes of a one-way analysis of variance and the F-test findings to determine whether or not the regression model is adequate. The findings show that the F values are statistically significant at the 1% level, indicating that the model is very good.

5 Conclusion

The world's economy is expanding, and many people are interested in online shopping especially teenagers. The following study is an attempt to determine factors that influence online shopping using factor analysis. The major factors that influence online shopping are preference of customers towards online shopping, risk involved in purchasing products through online, time effectiveness in online shopping, difficulties faced during online shopping and getting products from trustworthy websites. The factor scores obtained from

factor analysis are used as regressors in multiple regression model for predicting the customer satisfaction. So, online retailers should take note of these factors to increase the online sales.

References

1. Daskiran I, Keskin S and Bingol M (2017) Usability of the factor analysis scores in multiple linear regression analyses for the prediction of daily milk yield in Norduz goats. J. Agr. Sci. Tech., **19**:1507-1515.
2. Al-Nuaimi, B. T., Al-Mahdawi, H. K., Albadran, Z., Alkattan, H., Abotaleb, M., & El-kenawy, E. S. M. (2023). Solving of the inverse boundary value problem for the heat conduction equation in two intervals of time. Algorithms, **16**(1), 33.
3. Katawetawarks C and Wang C L (2011) Online Shopper Behavior: Influences of Online Shopping Decision. Asian Journal of Business Research, **1**(2):66-74.
4. Keskin S, Daskiran I and Kor A (2007) Factor analysis scores in multiple linear regression model for the prediction of caracass weight in Akkeci kids. Journal of Applied Animal Research,**31**(2):201-204, DOI: 10.1080/09712119.2007.9706664
5. Keskin S, Kor A and Karaca S (2007) Use of factor analysis scores in multiple linear regression model for determining relationships between milk yield and some udder traits in goats. Journal of Applied Animal Research, **31**(2):185-188, DOI: 10.1080/09712119.2007.9706660
6. Al-Mahdawi, H. K., Albadran, Z., Alkattan, H., Abotaleb, M., Alakkari, K., & Ramadhan, A. J. (2023, December). Using the inverse Cauchy problem of the Laplace equation for wave propagation to implement a numerical regularization homotopy method. AIP Conference Proceedings (Vol. **2977**, No. 1). AIP Publishing.
7. Padro R, Barrado E, Castrillejo Y, Valasco M A and Vaga M (1993) Study of the Contents and Speciation of Heavy Metals in River Sediments by Factor Analysis. Anal. Lett, **26**:1719-1739.
8. Rajesh R (2018) Evaluating the Factors Influencing Online Shopping and its Consumer Satisfaction in Pune Area. International Journal of Social Sciences, **4**(1):54-76.
9. Sharma S and Sitlani M (2013) Online Shopping among Higher Education Students in Indore: A Factor Analysis Approach. Indian Journal of Marketing,**43**(1):44-51.
10. Tahtali Y (2019) Use of factor scores in multiple regression analysis for estimation of body weight by certain body measurements in Romanov lambs. Peer J, **7**:e7434.
11. Akbari, E., Mollajafari, M., Al-Khafaji, H. M. R., Alkattan, H., Abotaleb, M., Eslami, M., & Palani, S. (2022). Improved salp swarm optimization algorithm for damping controller design for multimachine power system. IEEE Access, **10**, 82910-82922.
12. Ehsan khodadadi, S. K. Towfek, Hussein Alkattan. (2023). Brain Tumor Classification Using Convolutional Neural Network and Feature Extraction. Fusion: Practice and Applications, **13**(2), 34-41.
13. Yan Y, Du S (2017) Research on the online shopping customer satisfaction based on confirmatory factor analysis. BoletinTenico, Technical Bulletin, **55**(4):732-737.