DILEMMA IN ACQUIRING A MOBILE PHONE: AN EXPLORATION AND OVERVIEW WITH EXPLORATORY FACTOR ANALYSIS

A PROJECT REPORT

Submitted by

AYANTIKA MAITY, KORAK ROYGHATAK, SHRESTHA DE

Supervised by

Prof. Prasanta Narayan Dutta, Ms. Anwesha Sengupta, Dr. Banashree Sen

in partial fulfillment for the award of the degree

of

Master of Science

IN

Applied Statistics and Analytics, Year: 2020-21

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

(Formerly known as West Bengal University of Technology)

NH-12 (Old NH-34) Simhat, Haringhata, Nadia 741249, West Bengal

BONAFIDE CERTIFICATE

Certified that this project report **"DILEMMA IN ACQUIRING A MOBILE PHONE: AN EXPLORATION AND OVERVIEW WITH EXPLORATORY FACTOR ANALYSIS"** is the bonafide work of "**AYANTIKA MAITY, KORAK ROYGHATAK, SHRESTHA DE"** who carried out the project work under my supervision.

----------------------- ---------------------- ----------------------

SUPERVISORS

HOD, Department of Applied Science

SONASS

External Examiner

Acknowledgement

We would like to express our profound gratitude to our project supervisors Prof. P.N. Dutta, Ms. Anwesha Sengupta and Dr. Banashree Sen for their help, guidance and suggestions without which this project would not have been a reality.

We express our sincere thanks to Prof. R. Bhar, Head of the Department of Applied Science under SONASS, MAKAUT for his cooperation extended during the period of our research work.

We are gratefully indebted to Prof. S. Samajdar, Director, SONASS, MAKAUT for his great inspiration, encouragement and valuable guidance at every stage of our project work.

We would also like to thank all those who actively responded to our questionnaire through Google form for generation of required data.

Lastly, we thank MAKAUT for providing us the opportunity to carry out the project as well as all those concerned persons who have been directly or indirectly responsible for the completion of our project work.

---------------------------- ---------------------------------- ----------------------------------

Ayantika Maity Korak RoyGhatak Shrestha De

TABLE OF CONTENT

ABSTRACT

In the market research analysis, high-dimensional data is available more often, making it difficult for researchers to visualize and analyze the data. In our present study, an overview of factor analysis has been explored and applied the same to extract the important information by investigating the main factors affecting a consumer's choice of mobile phone. A set of related parameters were identified and corresponding data has been collectedthrough *Online Questionnaire*. Using Exploratory Factor Analysis (EFA) it has been possible to determine meaningful underlying construct or factors which capture a reasonableproportion of the total variance. The attempts have been made to find out a lesser number of factors responsible for customers' choice regarding selection of mobile phones without the loss of extra variance.

KEY WORDS

Exploratory factor analysis, Factor extraction, Factor loading, Factor rotation, Number of factors

1. INTRODUCTION

We begin with the question: "Are there a number of unseen factors acting beneath some measures which determine their observed values?" Answer to the question leads to exploratory factor analysis. The goal of factor analysis is to reduce "the dimensionality of the original space and to give an interpretation to the new space, spanned by a reduced number of new dimensions which are supposed to underlie the old ones" (Rietveld and Van Hout, 1993). Working with high dimensional data becomes much difficult due to inconsistency in data set and increase in computation time. Factor analysis helps us to reduce dimension by removing redundant data without loss of any important information. Factor analysis is performed by examining the pattern of correlations (or covariances) between the observed measures. Measures that are highly correlated (either positively or negatively) are likely influenced by the same factors, while those that are relatively uncorrelated are likely influenced by different factors. The number of factors in Exploratory Factor Analysis (EFA) may not always be stated in advanced by the researcher. Some of the programming languages like R, SPSS have been used to calculate relationships between all the measurement items, placing those highly correlated into factors which are then tried to match with the researcher's theoretically posited constructs.

When we don't have any prior knowledge about the data, then EFA attempts to explore the nature of the constructs influencing a set of observed variables. Here the data determines the factors. But when you have the prior information, and want to do modeling, testing and confirm hypotheses, then we should use confirmatory factor analysis (CFA). CFA tests whether a specified set of constructs is influencing observed variables in a predicted way (Jamie DeCoster, 1998).

1.1 Three Stages of Factor Analysis

First, a correlation matrix is generated for all the variables. Second, factors are extracted from the correlation matrix based on the correlation coefficients of the variables. (i.e. EXTRACTION) Here extraction method is based on Principal Component Analysis. Third, the factors are rotated in order to maximize the relationship between the variables and some of the factors and having a better understanding. (i.e. ROTATION) Here in our project, we have used the varimax rotational method.

1.2 Unpredictable stability with sample size

Sample size is a critical factor which affects estimation accuracy for any statistical entity, including population distribution parameters such as the mean, median, and standard deviation and other statistical parameters such as correlation coefficients, beta weights in regression, test scale reliability, and item and scale factor loading patterns. Sample size always plays a part in the calculation of variability of a parameter estimate. In general, the fewer observations made of a variable, the larger the variability of an estimate of any parameters estimated from them. Costello and Osborne (2005) examined the effect of sample size on an instrument with a very clear, strong factor structure (e.g., the Rosenberg Self-View Inventory (Rosenberg, 1965)) with respect to aspects of factor analysis. It was investigated whether items were assigned to the correct factor, and when they were, how much the factor loadings varied. After hundreds of simulations with real educational data varying sample size, EFAs were found to be relatively unstable. There are no strict rules regarding sample size in EFA. In the studies made by Fabrigar et al., 1999 and MacCallumet al., 1999, it has been pointed out that adequate sample size is partly determined by the nature of the data. In general, the sample size may be lesser with the use of stronger data. A data used in EFA is “Strong ” if there are uniformly high communalities without cross loadings and high loadings of several factors are present on each factor.

1.3 Factor Rotation and its Aim

There are many different types of rotation, but they all try make your factors each highly responsive to a small subset of your items (as opposed to being moderately responsive to a broad set). Most of the logic behind factor rotation can be found in Thurstone (1947) and Cattell (1977) which are still relevant in current literatures. According to the authors, factor rotation simplifies the factor structure with easier interpretation and more reliability.

Five criteria with respect to the matrix of loadings were suggested by Thurstone for the identification of the simple structure. There are two major categories of rotations, orthogonal rotations, which produce uncorrelated factors, and oblique rotations, which produce correlated factors. Orthogonal rotation can be specified through a rotation matrix R having the property $RR^T = I$ and whose elements are cosines of the angles between original and rotated axes. The best orthogonal rotation is widely believed to be Varimax.

The VARIMAX rotation which is an orthogonal rotation has been found as mostly used by the researchers among all available orthogonal rotations. In VARIMAX, a rotation of the original factors is done in such a way that variance of the loadings is maximized. There are other orthogonal rotations like QUARTIMAX which minimizes the number of factors needed to explain each variable and the rotation EQUIMAX has some properties of VARIMAX as well as of QUARTIMAX.

Oblique rotations are less distinguishable, with the three most commonly used being Direct Quartimin, Promax, and Harris-Kaiser Orthoblique.

In oblique rotation, the new axes may be inclined to each other at any angle except orthogonal directions. Using oblique rotation, simplicity in the interpretation may be gained. Rotations are always performed in the subspace generated by factors and the new axes will always explain less variance than the original factors but the part of the total variance explained by the original factors remains same after rotation. The only change is that part of the total variance is redistributed.

Fabrigar et al. (1999) emphasized that many factors are correlated, and therefore oblique rotations more accurately and realistically depict the “true factor structure.”

Harman (1976) further stated that if factors are orthogonal, a successful oblique rotation would accurately estimate inter factor correlations near zero and arrive at solutions close to those produced by an orthogonal rotation. Sass, Daniel A. and Schmitt (2010) suggested thatoblique rotation criteria will provide valid solutions for factors’ structures that have either correlated or uncorrelated factors and therefore provide a more flexible analytic approach.

Field (2000: 439) states “the choice of rotation depends on whether there is a good theoretical reason to suppose that the factors should be related or independent, and also how sthe variables cluster on the factors before rotation. If the oblique rotation demonstrates a negligible correlation between the extracted factors then it is reasonable to use the orthogonally rotated solution”.

The decision of choosing the type of factor rotation is a real problem as rotation of axes are not determined according to any statistical criteria but depends on the requirement of interpretations by the researcher.

1.4 Selection of the number of factors

One of the important questions that arise is about the number of reliable and interpretable factors that can be obtained from a given data set which would also summarize the pattern of correlations in the correlation matrix. Extraction of more factors implies improvement of fit between observed and reproduced correlation matrices and explanation of greater percentage of variance in the data. At the same time lesser choice of factors make the solution more parsimonious. Consequently, a trade-off is required to explain more variance with parsimonious solution.

The number of factors is equal to the number of variables present in the data set. But it is not necessary that all factors are providing significant amount of information about the common variance among the variables. Following criteria (thumb rules) are generally followed in the selection of number of factors. A single or more than one criteria can be jointly used for this purpose.

i) The Kaiser criterion states that you should use a number of factors equal to the number of the Eigen values of the correlation matrix that are greater than one.

ii) Retain those factors which can account for 70 – 80 % of the total variance.

iii) Estimate the numbers from scree-plot. In scree-plot Eigen values are plotted against factors. In scree-plot, the Eigen values decrease with the increase of factor number. The "Scree test" states that you should plot the Eigen values of the correlation matrix in descending order, and then use a number of factors equal to the number of Eigen values that occur prior to the last major drop in Eigen value magnitude.

There is controversy about the number of factors to be retained. Sometimes the researcher have a specific hypothesis that will determine the number factors you will include, while other times you simply want your final model to account for as much of the covariance in your data with as few factors as possible. The final choice rests with the interpretation and choice of the researcher.

After determination of the number of factors, the rotated loading matrix is to be judged to find the number of variables which load on each factor. If it is seen that the variables which are loaded on a factor are highly correlated and poorly correlated with other variables, the corresponding factor may be reliable.

1.5 Interpret your factor structure

Each of your measures will be linearly related to each of your factors. The strength of this relationship is contained in the respective factor loading, produced by the rotation. This loading can be interpreted as a standardized regression coefficient, regressing the factor on the measures. A factor is defined by considering the possible theoretical constructs that could be responsible for the observed pattern of positive and negative loadings. To ease interpretation you have the option of multiplying all of the loadings for a given factor by -1. This essentially reverses the scale of the factor.

1.6 Use of Factor Analysis in Research Methodology: A Survey

In this section, we describe a brief survey of wide variety of application of factor analysis and principal component analysis together with other multivariate techniques in some areas of Astrophysics, Information security, Landscape Pattern and behavioral research.

D. Fraix-Burnet et al. (2012) in their study of galaxy diversification using a six parameter space, utilized the techniques of principal component analysis and cluster analysis together with minimum contradiction analysis and cladistics to derive an explanatory classification of galaxies based on the physical causes of the diverse properties of galaxies.

A.K. Chattopadhyay et al. (2012) used the methods of principal component analysis, K-means clustering analysis and Independent component analysis to carry out an objective classification of the globular clusters.

A large scale exploratory factor analysis was carried out with 2504 observations to determine the underlying constructs that comprise barriers to distance education and 10 factors were identified Muilenburg and Berge (2001).

In Riitters et al.(1995), a multivariate factor analysis was performed to identify the common axes or dimensions of pattern and structure out of 55 metrics which were calculated for 85 maps of land use and land cover. A reduced set of 26 metrics were found out and it was observed that 87% of the total variance was explained by first six factors.

Muilenburg and Bergeb (2005) used EFA study to determine the underlying constructs that comprise student barrier to online learning. Fight factors (a) administrative issues, (b) social interaction, (c) academic skills, (d) technical skills, (e) learner motivation, (f) time and support for studies, (g) cost and access to the Internet, and (h) technical problems, were found.

Hu et al. (1999) examined TAM (Technology Acceptance Model) in a professional setting, investigating the factors affecting physicians' acceptance of telemedicine technology.

A principal component factor analysis was also performed in which four components were extracted, precisely matching the number of constructs included in TAM. Furthermore, items intended to measure the same construct exhibited prominently and distinctly higher factor loadings on a single component than on other components, suggesting adequate convergent and discriminate validity of the measurements adopted by the authors.

Chang and Ernett (1999) investigated to assess factors associated with Web site success in the context of electronic commerce. They first hypothesized some possible factors: (1) information and service quality, (2) system use, (3) playfulness, and (4) system design quality and justified those using EFA. KMO measure of the sampling adequacy was found to be 0.86. The rule of Eigen value greater than 1 was applied for retaining the number of factors and after Promax rotation four factors were extracted which matches with the hypothesis.

In the following section attempts have been made to extract the latent factors which captures a significant variance of all parameters.

2. RESEARCH METHODOLOGY

2.1 Collection of data

We have collected the data by quantitative research technique .The data here is collected by online survey. The data has been collected by forming a questionnaire on "Survey onthe parameters for a smartphone selection". We made a Google form consisting of 21 parameters regarding different aspects of a mobile phone and ask the respondents to rate them according to their perspective that which is "Doesn't matter at all" and "Matters most"for them on a 7-point likert scale. Structured data has been collected through the given questionnaire. The data were collected from 191 respondents which include students, officegoers, teachers, housewives and professionals over a period of one month by using Google form survey. In this collection of data some limitations are noticed (Muhammad Turki Alshurideh,2016).

The chosen parameters are as follows:

Table 1

1. Price	2. Battery life and fast charging support	3. Sound quality
4. Availability of service centre	5. Screen size	6. Fingerprint scanner or face unlock features
7. Brand name	8. Processor speed	9. Gorilla glass protection,
10. Offers	11. RAM	12. Accessories supplied with the phone
13. Colour	14. Latest Model	15. Dual sim
16. Camera quality	17. Display or screen resolution	18. Water and dust proof
19. Internal storage	20. Weight of the device	21. Heating issues

2.2 Analysis of collected data with Different Number of Extracted Factors

In order to identify major factors considered to be important by the consumer for mobile phone an exploratory factor analysis study had been conducted.

First of all the rotated component matrix has been analyzed and to obtain the factors without cross loading proper cut off level of cross loading have been applied along with theomission of the less relevant variables.

3. OUTCOME

3.1 Descriptive Statistics

Various descriptive statistical measures are used to compare variables. The mean, median, mode are calculated in each observed variable and they are compared to find out the most important variable which impacts on mobile phone selection preference. Standard deviation is calculated to see in which variable the most variation in customers' response.

Table 2

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Price	5.885	1.2554	191
Availability of Service centre	5.686	1.5647	191
Brand Name	5.623	1.3511	191
Offers	5.042	1.6506	191
Colour	4.440	1.7816	191
Camera quality	6.131	1.0705	191
Internal Storage	6.623	.6839	191
Battery life and fast charging support	6.665	.7418	191
Screen size	5.534	1.1777	191
Processor speed	6.581	.7626	191
RAM	6.592	.7683	191
Latest model	5.539	1.4057	191
Display or screen resolution	6.021	1.0660	191
Weight of the device	4.796	1.6937	191
Sound quality	6.267	1.0192	191
Fingerprint scanner or face unlock features	5.450	1.7249	191
Gorilla glass protection	5.817	1.4556	191
Accessories supplied with the phone	5.429	1.6137	191
Dual sim	5.743	1.6582	191
Water and dust proof	5.942	1.3186	191
Heating issues	6.257	1.2235	191

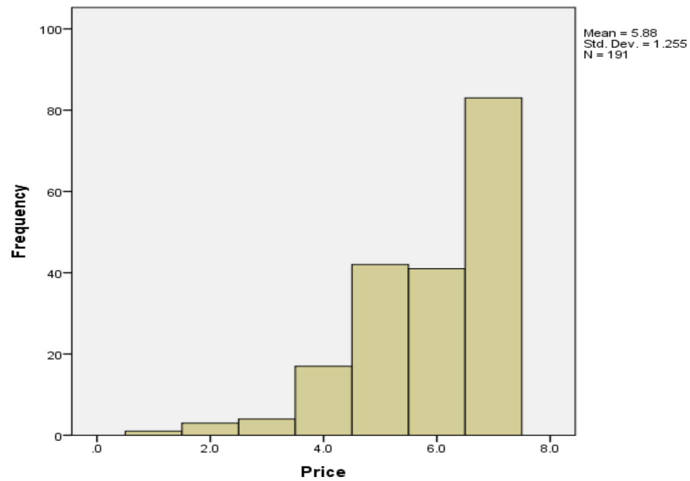

Fig.1

This is the histogram of the observed variable “Price” with highest communality. It may be observed that the variable has a negatively skewed distribution.

The observed variable having the highest mean is seen in Battery life and fast charging support. It's standard deviation is 0.7418. The lowest dispersion is seen in Internal storage response. Most of the variables follows a distribution which are negatively skewed.

3.2 KMO and Bartlett’s sphericity test

Kaiser-Meyer-Olkin(KMO) and Bartlett's Tests measures strength of the relationship among variables. Before performing factor analysis, two important conditions must be made:

i)The sample size is adequate and

ii) The correlation matrix is not an identity matrix

If the p value of Bartlett's sphericity test is less than 0.05 then we can reject null hypothesis, which assumes that no correlation is present among the variables. The Bartlett'stest has a strong drawback. It tends to be always statistically significant when the numberof instances 'n' increases. Some references advise to use this test only if the ratio 'n:p'(number of instances divided by the number of variables) is lower than 5.

The sampling adequacy is also tested through KMO test. It checks if we can factorize efficiently the original variables. KMO test measures the proportion of variance that might be a common variance among the variables. According to KMO test if the KMO index is high (≈1), the PCA can act efficiently; if KMO is low (0), the PCA is not relevant, values more than 0.6 is satisfactory. Larger proportions are expected as it represents more correlation is present among the variables, thereby giving way for the factor analysis.

Table 3

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.833
Bartlett's Test of Sphericity	Approx. Chi-Square	1168.861
	Df	210
	Sig.	.000

Here KMO measure is 0.833. This signifies that the dataset with 191 observations is adequate.

The p-value for Bartlett's test is 9.940785e-133<0.05. We reject the null hypothesis that assumes that the correlation matrix is identity matrix. If number of respondents would be 5 times of number of statement, then we can say the sample size is adequate.

This clearly suggests that factor analysis can be used to extract research factors/components.

3.3 Communalities

Table 4

Communalities

	Initial	Extraction
Price	1.000	.826
Availability of Service centre	1.000	.810
Brand Name	1.000	.818
Offers	1.000	.729
Colour	1.000	.705
Camera quality	1.000	.669
Internal Storage	1.000	.709
Battery life and fast charging support	1.000	.682
Screen size	1.000	.656
Processor speed	1.000	.784
RAM	1.000	.758
Latest model	1.000	.710
Display or screen resolution	1.000	.672
Weight of the device	1.000	.682
Sound quality	1.000	.766
Fingerprint scanner or face unlock features	1.000	.743
Gorilla glass protection	1.000	.667
Accessories supplied with the phone	1.000	.686
Dual sim	1.000	.836
Water and dust proof	1.000	.698
Heating issues	1.000	.545

Extraction Method: Principal Component Analysis.

This table gives details of the proportion of variance explained by the common factors associated with the various variables. The third column of Table 4 headed as Extraction reveals that, the least percentage contribution of variation in a variable by the observed common factors is 54.5%.

3.4 Extraction

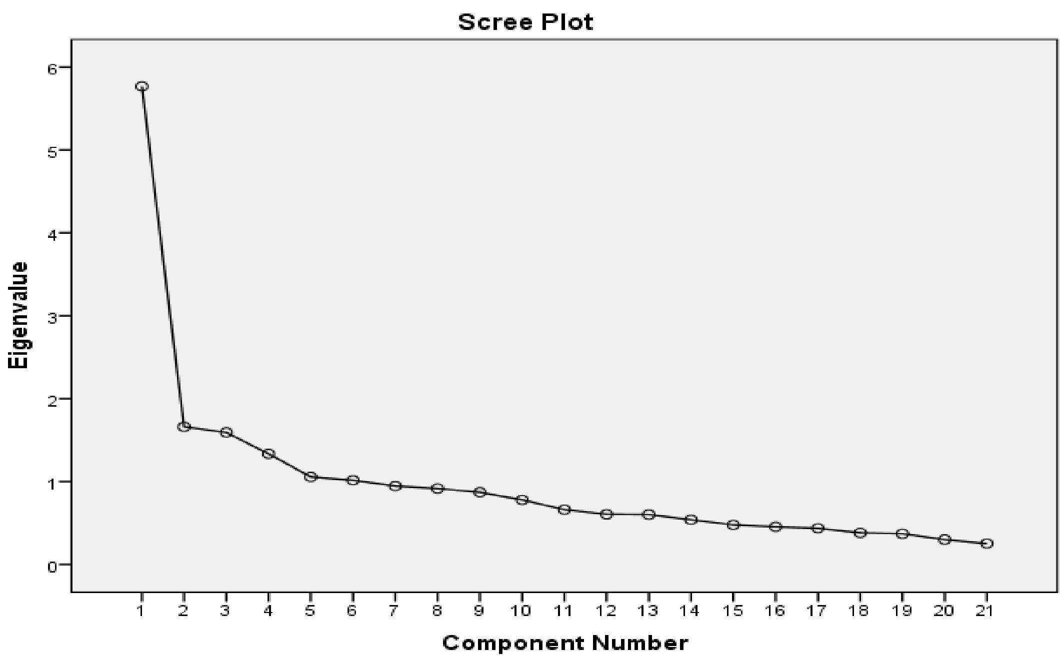

Fig. 2

From Scree Plot (see Fig.2) we have got the 1st elbow point at 2nd component, then there is an elbow point at the 5th component. After 5^{th} component there is no significant drop of the Eigen value level.

Table 5

Communalities

	Initial	Extraction
Price	1.000	.554
Availability of Service centre	1.000	.596
Offers	1.000	.587
Colour	1.000	.631
Camera quality	1.000	.638
Internal Storage	1.000	.562
Battery life and fast charging support	1.000	.682
Screen size	1.000	.582
Processor speed	1.000	.793
RAM	1.000	.738
Latest model	1.000	.506
Display or screen resolution	1.000	.528
Weight of the device	1.000	.505
Sound quality	1.000	.683
Fingerprint scanner or face unlock features	1.000	.572
Gorilla glass protection	1.000	.629
Accessories supplied with the phone	1.000	.597
Dual sim	1.000	.805
Water and dust proof	1.000	.682

Extraction Method: Principal Component Analysis.

Table 6

Rotated Component Matrix[a]

	Component					
	1	2	3	4	5	6
Price					.720	
Availability of Service centre					.654	
Brand Name				.640		
Offers					.451	
Colour				.752		
Camera quality		.713				
Internal Storage		.456	.457			
Battery life and fast charging support		.700	.420			
Screen size				.601		
Processor speed			.787			
RAM			.781			
Latest model	.496					
Display or screen resolution	.463					
Weight of the device	.466					
Sound quality		.734				
Fingerprint scanner or face unlock features	.570					
Gorilla glass protection	.715					
Accessories supplied with the phone	.620					
Dual sim						.878
Water and dust proof	.661					
Heating issues	.564					

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.[a]
a. Rotation converged in 9 iterations.

Table 7

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.768	27.466	27.466	5.768	27.466	27.466	3.247	15.460	15.460
2	1.660	7.904	35.370	1.660	7.904	35.370	2.344	11.163	26.622
3	1.592	7.581	42.951	1.592	7.581	42.951	2.138	10.182	36.804
4	1.334	6.351	49.302	1.334	6.351	49.302	2.014	9.592	46.397
5	1.055	5.023	54.325	1.055	5.023	54.325	1.461	6.956	53.353
6	1.015	4.832	59.157	1.015	4.832	59.157	1.219	5.804	59.157
7	.944	4.496	63.653						
8	.915	4.355	68.008						
9	.870	4.145	72.153						
10	.778	3.706	75.860						
11	.660	3.144	79.004						
12	.605	2.879	81.883						
13	.601	2.861	84.744						
14	.538	2.562	87.306						
15	.476	2.268	89.574						
16	.453	2.158	91.732						
17	.436	2.076	93.808						
18	.380	1.811	95.618						
19	.369	1.757	97.375						
20	.300	1.427	98.802						
21	.252	1.198	100.000						

Extraction Method: Principal Component Analysis.

The Kaiser criterion says that the components having Eigen value greater than 1 must be chosen. When minimum Eigen value 1 is taken, 6 components are extracted. We have observed that 6 factors can account only 59.157% variance, which falls behind the expected capture of the variance whose minimum value is usually considered as 60%. So, it does not meet the minimum expectation.

Table 8

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.768	27.466	27.466	5.768	27.466	27.466	2.331	11.102	11.102
2	1.660	7.904	35.370	1.660	7.904	35.370	2.208	10.514	21.615
3	1.592	7.581	42.951	1.592	7.581	42.951	2.129	10.137	31.753
4	1.334	6.351	49.302	1.334	6.351	49.302	1.857	8.841	40.594
5	1.055	5.023	54.325	1.055	5.023	54.325	1.780	8.477	49.071
6	1.015	4.832	59.157	1.015	4.832	59.157	1.316	6.266	55.337
7	.944	4.496	63.653	.944	4.496	63.653	1.228	5.849	61.186
8	.915	4.355	68.008	.915	4.355	68.008	1.177	5.604	66.790
9	.870	4.145	72.153	.870	4.145	72.153	1.126	5.363	72.153
10	.778	3.706	75.860						
11	.660	3.144	79.004						
12	.605	2.879	81.883						
13	.601	2.861	84.744						
14	.538	2.562	87.306						
15	.476	2.268	89.574						
16	.453	2.158	91.732						
17	.436	2.076	93.808						
18	.380	1.811	95.618						
19	.369	1.757	97.375						
20	.300	1.427	98.802						
21	.252	1.198	100.000						

Extraction Method: Principal Component Analysis.

By the extraction method of PCA we can see that up to 9 components eigen value is more than 0.8 . The 9 components account for 72.153% of the total variance. This variance captured is sufficiently good.

Table 9

Rotated Component Matrix[a]

	Component								
	1	2	3	4	5	6	7	8	9
Price							.881		
Availability of Service centre						.802			
Brand Name									.804
Offers							.539		
Colour					.795				
Camera quality			.762						
Internal Storage		.453	.561						
Battery life and fast charging support		.410	.671						
Screen size					.632				
Processor speed		.798							
RAM		.804							
Latest model				.778					
Display or screen resolution				.588					
Weight of the device					.485	.488			
Sound quality			.632			.482			
Fingerprint scanner or face unlock features				.682					
Gorilla glass protection	.667								
Accessories supplied with the phone	.725								
Dual sim								.884	
Water and dust proof	.688								
Heating issues	.586								

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 12 iterations.

We are suppressing the loadings below 0.4. From this table we are getting information about how these items are related to a particular component. The component matrix has many cross loadings of the variables in different components. A few variables like 'Brand name', 'Internal storage', 'Processor speed' have a moderate correlation with various extracted components. So it is not clear about their belongingness a particular factor. Here we also see few variables like “Internal storage”, “Weight of the device”, “Battery life and fast charging support” have significant loading in more than one component. We have triedby taking different cut offs for Eigen values and factor loadings to minimize cross loadings.

Table 10

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.421	28.532	28.532	5.421	28.532	28.532	2.541	13.372	13.372
2	1.656	8.717	37.249	1.656	8.717	37.249	2.285	12.025	25.397
3	1.495	7.869	45.118	1.495	7.869	45.118	2.273	11.962	37.359
4	1.289	6.783	51.901	1.289	6.783	51.901	2.150	11.315	48.675
5	1.019	5.361	57.262	1.019	5.361	57.262	1.429	7.518	56.193
6	.988	5.201	62.463	.988	5.201	62.463	1.191	6.270	62.463
7	.918	4.831	67.294						
8	.861	4.533	71.828						
9	.747	3.931	75.759						
10	.623	3.280	79.039						
11	.604	3.179	82.218						
12	.600	3.159	85.376						
13	.521	2.742	88.119						
14	.457	2.407	90.526						
15	.439	2.310	92.836						
16	.397	2.088	94.924						
17	.380	1.999	96.923						
18	.315	1.659	98.582						
19	.269	1.418	100.000						

Extraction Method: Principal Component Analysis.

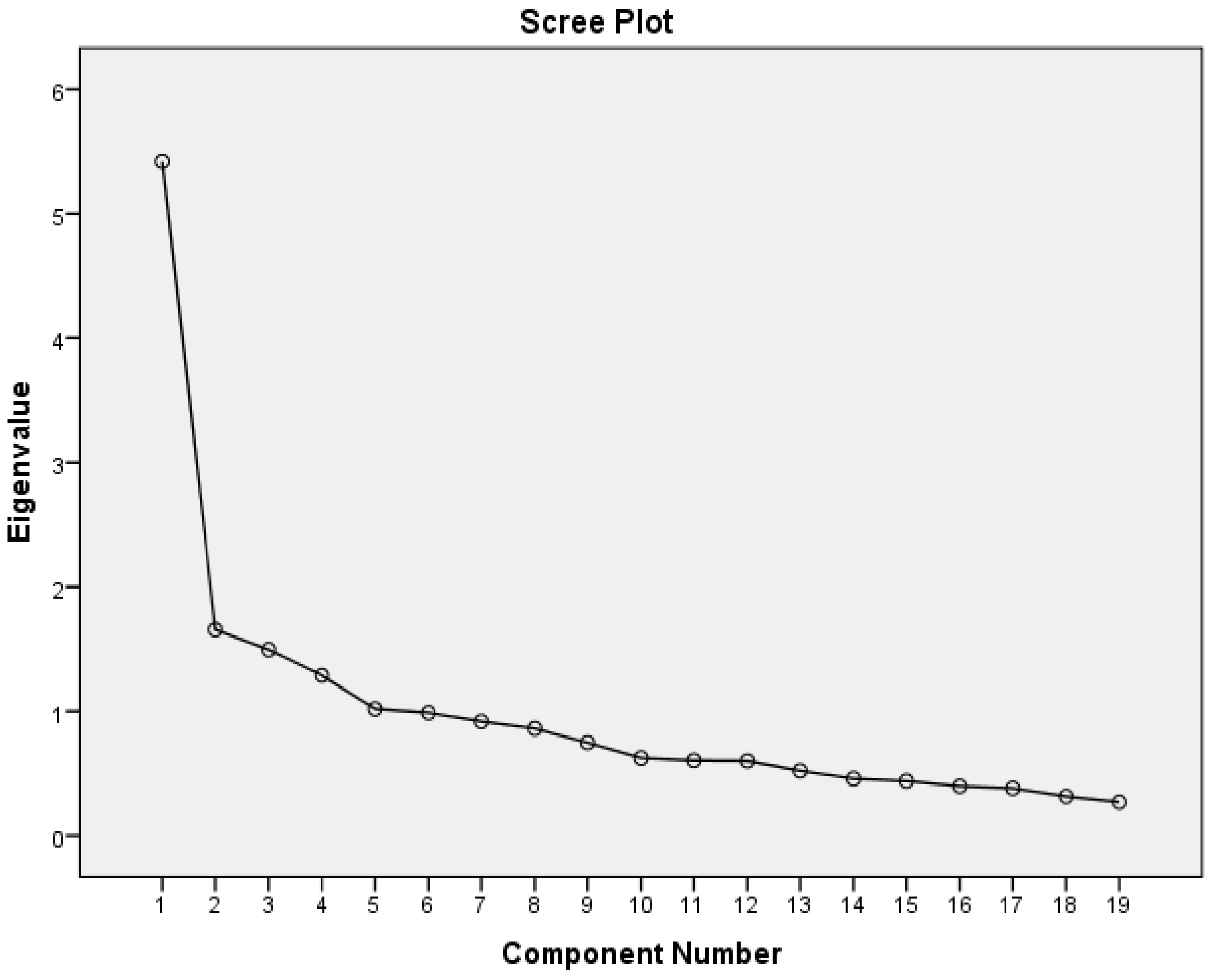
Scree Plot
Eigenvalue
6
5
4
3
2
1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
Component Number

Fig: 3

Table 11

Rotated Component Matrix[a]

	Component					
	1	2	3	4	5	6
Price					.733	
Availability of Service centre					.621	
Offers					.534	
Colour	.762					
Camera quality				.736		
Internal Storage				.477		
Battery life and fast charging support				.717		
Screen size	.688					
Processor speed			.825			
RAM			.812			
Latest model	.586					
Display or screen resolution						
Weight of the device	.529					
Sound quality				.693		
Fingerprint scanner or face unlock features	.568					
Gorilla glass protection		.633				
Accessories supplied with the phone		.629				
Dual sim						.874
Water and dust proof		.778				

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

Deciding upon the number of factors that can be retained is difficult. After using different cut offs in Eigen values and factor loadings we have come up with the best possible factor extraction. This table shows the factor loadings, which are the correlations between each extracted factor (Component1, Component2, Component3, etc.) and each variable (Price, Availability of Service centre etc.). A quick scan of these loadings give us an idea of which variables are most highly correlated with each factor. To remove the cross loading, we tried to remove a minimum number of variables. We have observed that the variable "Heating issue" having the least communality and the variable "Brand name" which has no direct correlation with the quality as per biased perception, we omitted these two variables. We have taken the factors which have Eigen values greater than 0.975 and suppressed the loadings below 0.475. In this trial no cross loading is observed and the loadings of the variables in the respective components are high. The variables are evenly distributed between the six factors and there is consistency in the distribution. Our study leads to the conclusion that there are 6 latent constructs which can be interpreted in the following meaningful way.

The following factors (Fig: 4) that have been extracted using the original variables are

Factor 1 : Mobile esthetics and hardware

a. Colour
b. Screen size
c. Latest model
d. Weight of the device
e. Fingerprint scanner or face unlock features

Factor 2 : Accessories

a. Gorilla glass protection
b. Accessories supplied with the phone
c. Water and dust proof

Factor 3: CPU criteria

a. Processor speed
b. RAM

Factor 4: Popular features

a. Camera quality
b. Internal storage
c. Battery life and fast charging support

d. Sound quality

Factor 5: Affordability

a. Price

b. Availability of service centre

c. Offers

Factor 6: Additional facility

a. Dual sim

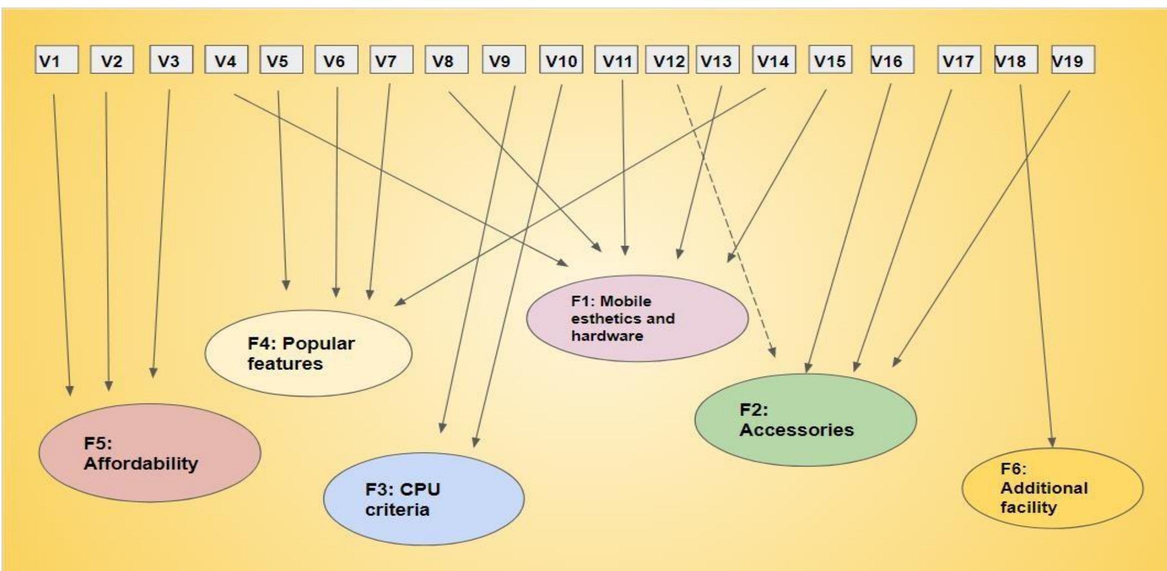

Fig: 4

4. CONCLUSION

This study aims to investigate the factors affecting a customer's choice of mobile phone. It can be concluded that, this is not an easy task for mobile service suppliers, as these factors often appear to overlap. Many psychological factors also affect the selection. Thus, mobile phone providers try to identify those factors that are important from customer perspectives and enhance their marketing activities.

5. LIMITATIONS OF THE STUDY

This study has a few limitations, which can be taken into consideration in any future research. The primary limitation is that, the data collected for the study is limited to mobilephone users only in Kolkata and few parts of West Bengal. The questionnaire was distributed online and a large part of them are filled by college goers. Therefore, the data set is not a true representative of the whole population. Also, the study is limited to one time point, and did not allow for analysis of changing demands of customer over time. Moreover, a Likert scale was used to collect the respondent views regarding the level to which they agreed or disagreed with the study's statements. The assumption that the deviation of each point in the Likert scale is equal may not be true. Future research should address these factors and their influence more deeply.

6. FUTURE SCOPE AND PLAN

In our study, we have reduced the number of variables to a fewer meaningful factors or latent constructs. With this prior knowledge of factors and their correlation with variables, more study is needed to determine the structural relationships in between dependent and independent variables (considering constructs as also variables) to confirm the results which may be done through Structural Equation Modeling (SEM). The existing model of EFA may also be improved by measuring and analyzing the errors associated with constructs.

This study can be extended to a wide area of Psychology, Health informatics, Marketing research and Information technology. In particular, we would like to revisit and improve the study carried out by Muilenberg and Bergeb (2005) to find out the minimum number of factors which may act as students' barrier to online learning and teaching under present restricted environment.

7. REFERENCES

1. Cattell, R.B., “Scientific Use of Factor Analysis in Behavioral and Life Sciences,” 1977 – agris.fao.org.
2. Chang, L., and Arnett, K.P., “Exploring the factors associated with Web site success in the context of electronic commerce,” Information & Management 38 (2000), pp. 23 – 33.
3. Chattopadhyay, A.K., Mondal, Saptarshi and Chattopadhyay, Tanuka, “ Independent Component Analysis for the objective classification of globular clusters of the galaxy NGC 5128, Computational Statistics and Data Analysis 57 (2013), pp. 17–32.
4. Costello, A.B., and Osborne, J.W., “Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis,” Practical Assessment, Research, and Evaluation, 10(7), 2005, pp.1 – 9.
5. Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 4, (1999), 272–299.
6. Field, A., (2000), “Discovering Statistics using SPSS for Windows,” London- Thousand Oaks-New Delhi: Sage Publications.
7. Fraix-Burnet, D., Chattopadhyay, T., Chattopadhyay, A.K., Davoust, E. and Thuillards, M., “A Six Parameter Space to Describe galaxy diversification,” Astronomy & Astrophysics, 545, A80 (2012), DOI: 10.1051/0004-6361/201218769.
8. Harman, H. H. (1976). Modern factor analysis (3rd ed.). Chicago: The University of Chicago Press.
9. Hu, P.J. ; Chau, P.Y.K.; Sheng, O.R.L. and Tam, K.Y.,” Examining the Technology Acceptance Model Using physician Acceptance of Telemedicine Technology,” Journal of Management Information Systems, Vol. 16, No. 2 (Fall, 1999), pp. 91-112.
10. MacCallum, R.C., Widaman, K.F., Zhang, S.B., and Hong, S.H., “Sample Size in Factor Analysis,” Psychological Methods, 4 (1), (1999), pp. 84 – 99.
11. Alshurideh, M.T. (2016), “Exploring the Main Factors Affecting Consumer Choice of Mobile Phone Service
12. Provider Contracts”, Int. J. Communications, Network and System Sciences, 9, 563-581.Muilenburg, L.Y. and Berge, Z.L.,” Barriers to distance education: A factor-analytic

study," The American Journal of Distance Education. 15(2) (2001): 7-22.

13. Muilenburg, L.Y. and Berge, Z.L., "Student Barrier to Online Learning: A Factor Analytic Study," Distance Education, 26(1), 2005, pp. 29 – 48.
14. Rietveld, T., and Van Hout, R., "Statistical Techniques for the Study of Language and Language Behavior, 1993, Berlin-New York: Mouton de Gruyter.
15. Riitters, K.H., O'Neil, Hunsaker, C.T., Wickham, J.D., Yankee, D,H., Timmins, S.P., Jones, K.B., and Jackson, B.L., " A Factor Analysis of Landscape Pattern and Structure Metrics," Landscape Ecology, Vol. 10, No. 1 (1995), pp. 23 – 39. SPB Academic Publishing, Amsterdam.
16. Rosenberg, M., "Soceity and the Adolescent Self Image," (1965), Princeton, NJ: Princeton University Press.
17. Sass, Daniel A. and Schmitt, Thomas A. 'A Comparative Investigation of RotationCriteria within Exploratory Factor Analysis', Multivariate Behavioral Research, 45: 1 (2010), 73 103.
18. Thurstone, L. L. Multiple-factor analysis. 1947, Chicago: University of Chicago Press.