PRINCIPAL COMPONENT ANALYSIS IN TOURISM MARKETING

Erika KULCSÁR
Babes-Bolyai University, Cluj-Napoca, Romania
Horea str., no. 7, Cluj-Napoca
e-mail: erika.kulcsar@econ.ubbcluj.ro

Abstract. *The analysis methods of the interdependences are meant to give a meaning to a set of variables or to group variables in a certain way. This work includes analysis in principal components. In ACP (Analysis of Principal Components) I included nine variables. Starting from the nine variables I sought to identify three principal components (factors) that summarize most of the information held by these variables and simplify the process of interpretation of results. To achieve the ACP which is a multivariate analysis method of marketing data I worked with primary data collected through quantitative marketing research.*

Keywords: correlation, factor, items (variable), variance.

Management & Marketing (2010) Vol. 5, No. 2, pp. 151-158

1. Introduction

The Analysis of principal components is classified among the descriptive methods analyzing interdependencies between variables. Therefore there are no dependent variables and independent variables, the simultaneous combination of analyzed variables (interdependences) is important (Constantin, 2006, p. 246).

The Analysis of principal components consists in identifying, based on a set of variables, a few factors that can synthesize most of the total information contained in the original variables. These factors are those common elements, latent, which is the basis of the variables intercollinearity (Lefter, 2004, p. 398).

A sufficiently large sample of subjects (300-400 subjects) can compensate for both the lower factorial saturation and for the reduced number of variables per factor (Labar, 2008, p. 311).

With the Analysis of principal components, the variables are measured by Likert scales, scales of interval or proportions (Labar, 2008, p. 310).

The quantitative research was conducted between 15.05.2009-17.10.2009. I distributed over 2,000 questionnaires. The questionnaires were distributed to hotels in Brasov, Predeal, Poiana-Brasov, Sfantu Gheorghe, Covasna, Miercurea-Ciuc, Gheorgheni, Tusnad, Targu-Mures, Sighisoara, Sibiu, Alba – Iulia and other localities that have linked this route.

2. Principal components analysis

For the principal components analysis, we analyzed the interrelationships between several variables, based on which tourists staying in hotels in Centre Development Region appreciated the furnishing of the hotel rooms, the offer of culinary products in the location where they most often served meals and the serving personnel.

Assessed items/variables are:

- V1: The room offers a nice familiar ambiance
- V2: The room offers special comfort.
- V3: The room is furnished in good taste.
- V4: The menu variety.
- V5: The quality of food.
- V6: Novelty of food.
- V7: Specific offer, traditional meals.
- V8: The way of serving up the ordered food.
- V9: Hospitality of the staff can make this hotel become one of the favorite places for tourists.

Principal components analysis results are presented below.

The Correlation Matrix Table (Table 1) presents the matrix of correlations between variables, it can be seen that there are several sets of correlations above 0.30, therefore the application of the factorial analysis on these variables is appropriate. You can also see that there are no correlation coefficients above 0.80, therefore there are no variables to correlate very strongly with each other (avoiding multicollinearity). Since the determinant is greater than 0.00001 (in this case 0.008) it results that there is no multicollinearity or singularity between variables.

Table 1

Matrix of correlation coefficients between the variables analyzed

Correlation Matrix[a]

		V1	V2	V3	V4	V5	V6	V7	V8	V9
Correlation	V1	1.000	.678	.710	.349	.398	.433	.387	.457	.507
	V2	.678	1.000	.720	.356	.472	.530	.413	.458	.474
	V3	.710	.720	1.000	.291	.409	.435	.371	.365	.443
	V4	.349	.356	.291	1.000	.632	.538	.620	.422	.545
	V5	.398	.472	.409	.632	1.000	.574	.580	.520	.458
	V6	.433	.530	.435	.538	.574	1.000	.648	.456	.449
	V7	.387	.413	.371	.620	.580	.648	1.000	.477	.462
	V8	.457	.458	.365	.422	.520	.456	.477	1.000	.522
	V9	.507	.474	.443	.545	.458	.449	.462	.522	1.000

a. Determinant = .008

KMO Table and Bartlett's test (Table no. 2) presents the results to the Bartlett and KMO tests. It may be noted that for the Bartlett sphericity test $\chi^2(36)$ = 1785.162, p <0.001 and therefore the correlations matrix is significantly different from the identity matrix in which the variables would not correlate with each other, the variables being appropriate for factorization. KMO Index = 0.884 characterizes the set of variables as being very good for factorial analysis.

Table 2

KMO and Barlet's Test Table

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.884
Bartlett's Test of Sphericity	Approx. Chi-Square	1785.162
	df	36
	Sig.	.000

Anti-image Matrices Table (Table no. 3) consists of a series of coefficients for assessing the suitability of variables to the factorial model.

I examined the lower half of the table, namely the principal diagonal of the Anti-image Correlation field; since on this principal diagonal there are no values under 0.50 these coefficients are very good, indicating that they are suitable for the factorial analysis, it can be seen that with item V1, the coefficient is 0.878, for V2 0.880, for V3 0.836 etc.

Communalities table (Table no. 4) presents in the Extraction column the communalities for each variable after extraction of the three factors. Thus, for the item V1, the communality is 0.787, which means that the three factors extracted explain 78.8% of the V1 item variance.

Table 3

Coefficients for assessing the suitability of variables to the factorial model

Anti-image Matrices

		V1	V2	V3	V4	V5	V6	V7	V8	V9
Anti-image Covariance	V1	.403	-.098	-.163	-.016	.023	.002	-.005	-.069	-.069
	V2	-.098	.374	-.155	.016	-.039	-.087	.018	-.046	-.023
	V3	-.163	-.155	.382	.041	-.038	-.008	-.021	.046	-.034
	V4	-.016	.016	.041	.447	-.163	-.045	-.129	.029	-.149
	V5	.023	-.039	-.038	-.163	.458	-.070	-.053	-.115	.017
	V6	.002	-.087	-.008	-.045	-.070	.464	-.166	-.021	-.009
	V7	-.005	.018	-.021	-.129	-.053	-.166	.450	-.067	-.011
	V8	-.069	-.046	.046	.029	-.115	-.021	-.067	.578	-.139
	V9	-.069	-.023	-.034	-.149	.017	-.009	-.011	-.139	.534
Anti-image Correlation	V1	.878[a]	-.253	-.415	-.039	.053	.004	-.013	-.144	-.150
	V2	-.253	.880[a]	-.411	.038	-.094	-.209	.043	-.099	-.050
	V3	-.415	-.411	.836[a]	.099	-.091	-.020	-.052	.097	-.076
	V4	-.039	.038	.099	.851[a]	-.360	-.099	-.288	.056	-.304
	V5	.053	-.094	-.091	-.360	.898[a]	-.152	-.116	-.224	.034
	V6	.004	-.209	-.020	-.099	-.152	.909[a]	-.364	-.040	-.019
	V7	-.013	.043	-.052	-.288	-.116	-.364	.890[a]	-.131	-.023
	V8	-.144	-.099	.097	.056	-.224	-.040	-.131	.907[a]	-.250
	V9	-.150	-.050	-.076	-.304	.034	-.019	-.023	-.250	.909[a]

a. Measures of Sampling Adequacy(MSA)

Table 4

Communalities

Communalities

	Initial	Extraction
V1	1.000	.787
V2	1.000	.799
V3	1.000	.835
V4	1.000	.727
V5	1.000	.679
V6	1.000	.747
V7	1.000	.746
V8	1.000	.719
V9	1.000	.756

Extraction Method: Principal Component Analysis.

Total Variance Explained table (Table no. 5) is one of the most important, because it contains eigenvalue values for each factor, the percentage of variance explained by each extracted factor as well as the percentages of cumulative variance explained by all factors extracted before and after rotation. (Labar, 2008, p. 325)

Thus, factor 1 explains 54.51% of items variance, factor 2 explains 13.61% of the items variance, factor 3 explains 7.39% of total items variance and the three factors altogether explain 75.5% of total variance of items. Eigenvalue is the variance explained by each factor of the total variance of items.

Table 5

Values of the components and the variance explained

Total Variance Explained

Compo-nent	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	4.906	54.512	54.512	4.906	54.512	54.512	2.675	29.722	29.722
2	1.224	13.605	68.117	1.224	13.605	68.117	2.477	27.524	57.246
3	.665	7.387	75.503	.665	7.387	75.503	1.643	18.257	75.503
4	.545	6.061	81.564						
5	.460	5.110	86.674						
6	.370	4.109	90.783						
7	.304	3.376	94.159						
8	.282	3.134	97.293						
9	.244	2.707	100.000						

Extraction Method: Principal Component Analysis.

The check of the suitability is made through the percentage of non-redundant residues which are greater than 0.05; in this case, the percentage is 38% (note a from the Table no. 6 Reproduced Correlations). For a better suitability the percentage should be as small as possible. The rule is that the percentage of non-redundant residues above 0.05 to be under 50% (Labar, 2008, p. 328).

Component Matrix Table (Table no. 7) presents the factorial saturation of the items in factors before rotation.

Table 6

Model Suitability Analysis

Reproduced Correlations

		V1	V2	V3	V4	V5	V6	V7	V8	V9
Reproduced Correlation	V1	.787[b]	.774	.787	.302	.411	.451	.360	.498	.547
	V2	.774	.799[b]	.808	.347	.468	.549	.443	.437	.484
	V3	.787	.808	.835[b]	.252	.392	.490	.363	.376	.428
	V4	.302	.347	.252	.727[b]	.684	.638	.706	.545	.540
	V5	.411	.468	.392	.684	.679[b]	.676	.703	.516	.520
	V6	.451	.549	.490	.638	.676	.747[b]	.723	.392	.402
	V7	.360	.443	.363	.706	.703	.723	.746[b]	.455	.455
	V8	.498	.437	.376	.545	.516	.392	.455	.719[b]	.735
	V9	.547	.484	.428	.540	.520	.402	.455	.735	.756[b]
Residual[a]	V1		-.096	-.078	.048	-.013	-.019	.027	-.040	-.040
	V2	-.096		-.089	.009	.004	-.019	-.030	.021	-.010
	V3	-.078	-.089		.038	.017	-.055	.008	-.011	.015
	V4	.048	.009	.038		-.052	-.100	-.087	-.123	.005
	V5	-.013	.004	.017	-.052		-.102	-.123	.004	-.062
	V6	-.019	-.019	-.055	-.100	-.102		-.075	.063	.047
	V7	.027	-.030	.008	-.087	-.123	-.075		.023	.008
	V8	-.040	.021	-.011	-.123	.004	.063	.023		-.213
	V9	-.040	-.010	.015	.005	-.062	.047	.008	-.213	

Extraction Method: Principal Component Analysis.

a. Residuals are computed between observed and reproduced correlations. There are 14 (38.0%) nonredundant residuals with absolute values greater than 0.05.

b. Reproduced communalities

Table 7

Correlations between the variables and the principal components

Component Matrix[a]

	Component		
	1	2	3
V2	.770	-.433	
V6	.766		-.346
V5	.761	.302	
V7	.748	.373	
V1	.740	-.487	
V9	.729		.472
V4	.714	.466	
V3	.714	-.549	
V8	.700		.469

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

The table Rotated Component Matrix (Table no. 8) is one of the most important output tables of the factorial analysis. In this table, the 0.814 saturation of the item V7 in F1 factor represents the Pearson correlation coefficient between V7 item and F1 factor.

Table 8

Correlations between variables and factors following the rotation of the axis

Rotated Component Matrix[a]

	Component		
	1	2	3
V7	.814		
V6	.776	.370	
V4	.750		.402
V5	.724		.317
V3		.881	
V2		.820	
V1		.804	.340
V9		.300	.764
V8	.304		.754

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Reviewing the composition of the items of the three factors, as a result of principal components analysis, I could define the factors as follows: F1 factor relates to the food offer, F2 factor to the arrangement of the rooms and F3 factor to the hotel staff.

3. Conclusions

Our factorial analysis was a useful tool for reducing the list of variables taken into account, in surveys, when assessing the customer preference for a certain hotel, to a limited group of variables which capture most of the meaning. Thus, the analysis revealed that customers are mostly influenced in their choices by the food offer, by the way rooms are decorated and by the behavior of the hotel staff. A better focus, from the part of the hotels, as service providers, on these issues would ensure an increased customer loyalty and the perspective of widening the pool of potential clients.

Further research on the topic, including qualitative elements, and a diversification of the types of touristic services providers included in the survey may benefit the addressability of the results, and the possibility of generalizing them on a wider, national scale.

References

Borza, A., Bordean, O. (2006), „Performance measuring of tourism small and medium enterprises”, *Management & Marketing*, 1 (4), pp. 106-110

Coita, D.C., Nedelea, A. (2006), „Tourists’ behavior and nationality – criteria for tourists classification and market segmentation”, *Management & Marketing*, 1 (3), pp. 87-98

Constantin, C. (2006*), Marketing computer systems. Analysis and processing of marketing data*, Infomarket Printing House, Brasov

Labar, A.V. (2008) *SPSS for Education*, Polirom, Iasi

Lefter, C. (2004) *Marketing Research.* Infomarket Printing House, Brasov