



Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Ποσοτικές Μέθοδοι Ανάλυσης στις Κοινωνικές Επιστήμες

Ασκήσεις

Θεόδωρος Χατζηπαντελής

Τμήμα Πολιτικών Επιστημών Α.Π.Θ.

Άδειες Χρήσης

Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons. Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα. Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.



Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Περιεχόμενα

Άδειες Χρήσης.....	2
Χρηματοδότηση.....	2
Ενότητα 1η: Ανάλυση κατά συστάδες	4
Ενότητα 2η: Πολυδιάστατη κλιμάκωση	5
Ενότητα 3η: Ανάλυση αντιστοιχιών	7
Ενότητα 4η: Κυρίων συνιστωσών και Παραγοντική ανάλυση.....	9

Ενότητα 1η: Ανάλυση κατά Συστάδες.

- Εκφώνηση άσκησης: Εφαρμογή της Μεθόδου **Cluster Analysis** (Ανάλυση κατά συστάδες). Τα δεδομένα του παραδείγματος δίνονται στην σελίδα του Τμήματος Στατιστικής του Πανεπιστημίου του Λονδίνου [Department of Statistics at the London School of Economics & Political Science (<http://www.lse.ac.uk/statistics/research/Social-Statistics/Multivariate-Data-Analysis/Data-and-Syntax.aspx>)].

Για περισσότερα παραδείγματα επισκεφτείτε την προτεινόμενη ηλεκτρονική διεύθυνση. This is an example where variables, rather than cases are clustered. The data are correlations between variables relating to home and school circumstances of children. The variables are coded as follows:

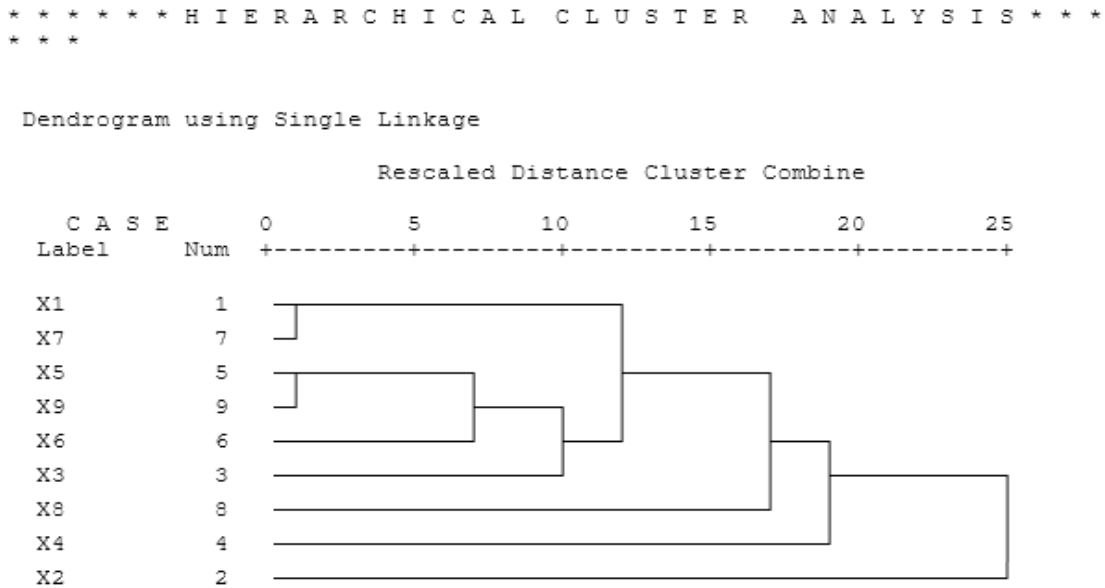
- X1=Parental circumstances in 1964, X2=Details of class teacher in 1964, X3=School-parent interaction in 1964, X4=Girl's attitude in 1964, X5=Test score in 1964, X6=Type of school in 1968, X7=Parental circumstances in 1968, X8=School-parent interaction in 1968, X9=Test score in 1968.
- The syntax for reading these correlations/similarities, and for single (nearest neighbour) and for complete (farthest neighbour) linkage cluster analysis are given in educCLA.sps. The file educCLA.spo (and educCLA.pdf) gives the output for nearest neighbour and for farthest neighbour cluster analysis, including both icicle plots and dendrograms.
 - Οδηγίες επίλυσης (Ακολουθώντας τις εντολές που παρουσιάστηκαν στην ενότητα καταλήγουμε στους παρακάτω πίνακες και δενδρογράμματα).

Single Linkage

Agglomeration Schedule

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	1	7	.770	0	0	5
2	5	9	.758	0	0	3
3	5	6	.613	2	0	4
4	3	5	.556	0	3	5
5	1	3	.501	1	4	6
6	1	8	.382	5	0	7
7	1	4	.339	6	0	8
8	1	2	.184	7	0	0

Πίνακας 1.1: Agglomeration Schedule.



Πίνακας 1.2: Το δενδρόγραμμα.

Ενότητα 2η: Πολυδιάστατη Κλιμάκωση

- Εκφώνηση άσκησης: Εφαρμογή της Μεθόδου **Multidimensional Scaling** (Πολυδιάστατη κλιμάκωση). Τα δεδομένα του παραδείγματος δίνονται στην σελίδα του Τμήματος Στατιστικής του Πανεπιστημίου του Λονδίνου [Department of Statistics at the London School of Economics & Political Science(<http://www.lse.ac.uk/statistics/research/Social-Statistics/Multivariate-Data-Analysis/Data-and-Syntax.aspx>)].

Για περισσότερα παραδείγματα επισκεφτείτε την προτεινόμενη ηλεκτρονική διεύθυνση.

The data are from an experiment in which subjects were asked to look at a screen which had two circular opaque glass windows. These windows were lit from two projectors behind the screen. Different colour filters could be inserted in the projectors. Fourteen colour filters were used. Each stimulus was combined with each other stimulus in a random order. The subjects were then asked to rate the degree of 'qualitative similarity' between each pair of colour filters on a 5-point scale. The similarity matrix is given in Table 3.4. The file colour.txt contains the full similarity matrix. Syntax for reading the similarity matrix into SPSS and carrying out 1- and 2-dimensional ordinal scaling is given in file colour.sps.

2-Dimensional Solution**Stress and Fit Measures**

Normalized Raw Stress	.00106
Stress-I	.03263 ^a
Stress-II	.07972 ^a
S-Stress	.00282 ^b
Dispersion Accounted For (D.A.F.)	.99894
Tucker's Coefficient of Congruence	.99947

PROXSCAL minimizes Normalized Raw Stress.

a. Optimal scaling factor = 1.001.

b. Optimal scaling factor = .999.

- – Οδηγίες επίλυσης: (Ακολουθώντας τις εντολές που παρουσιάστηκαν στην ενότητα καταλήγουμε στους παρακάτω πίνακες και διαγράμματα αξόνων).

Πίνακας 2.1: 2 Dimensional Solution.

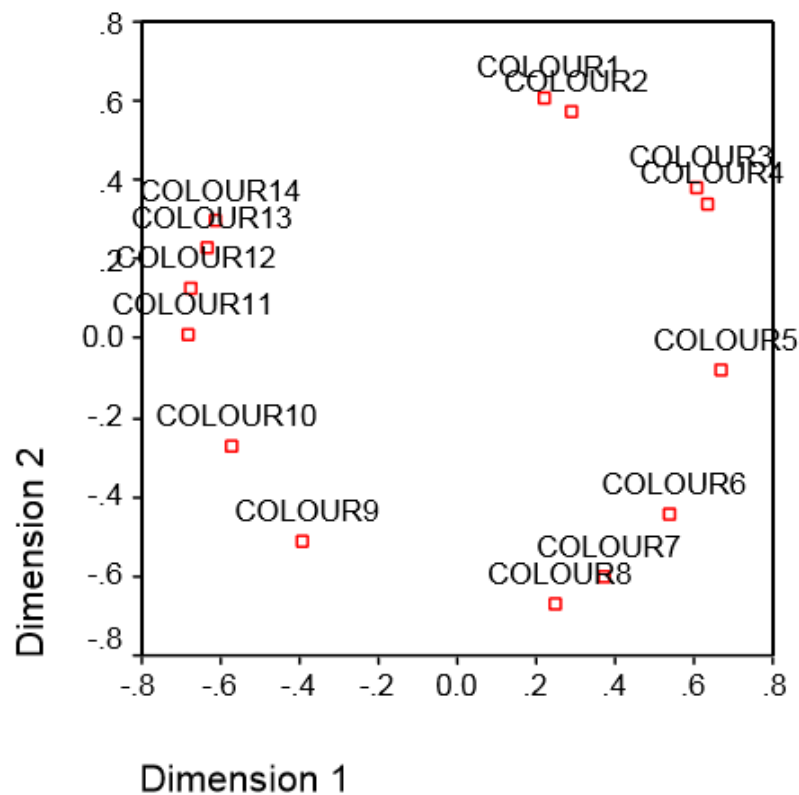
Final Coordinates

	Dimension	
	1	2
COLOUR1	.221	.611
COLOUR2	.289	.576
COLOUR3	.605	.380
COLOUR4	.633	.340
COLOUR5	.669	-.080
COLOUR6	.535	-.443
COLOUR7	.376	-.602
COLOUR8	.247	-.667
COLOUR9	-.396	-.513
COLOUR10	-.573	-.271
COLOUR11	-.681	.010
COLOUR12	-.678	.130
COLOUR13	-.633	.228
COLOUR14	-.614	.302

Πίνακας 2.2: Coordinates.

Object Points

Common Space



Διάγραμμα 2.3.

Ενότητα 3η: Ανάλυση Αντιστοιχιών

- Εκφώνηση άσκησης: Εφαρμογή της Μεθόδου **Correspondence Analysis** (Ανάλυση αντιστοιχιών). Τα δεδομένα του παραδείγματος δίνονται στην σελίδα του Τμήματος Στατιστικής του Πανεπιστημίου του Λονδίνου [Department of Statistics at the London School of Economics & Political Science (<http://www.lse.ac.uk/statistics/research/Social-Statistics/Multivariate-Data-Analysis/Data-and-Syntax.aspx>)]. Για περισσότερα παραδείγματα επισκεφτείτε την προτεινόμενη ηλεκτρονική διεύθυνση. The data are from the British General Election Study 1992. Respondents were asked what party they voted for in the 1992 general election, and the main reason for their choice. The data file contains 3 variables:
- REASON Reason for voting for chosen party, 1=Always vote that way, 2=Best party, 3=My party had no chance, PARTY Party voted for, 1=Conservative, 2=Labour, 3=Liberal Democrat, 4=Other, 5=Refused to give a response, COUNT Frequency of respondents.

- Οδηγίες επίλυσης: (Ακολουθώντας τις εντολές που παρουσιάστηκαν στην ενότητα καταλήγουμε στους παρακάτω πίνακες και διαγράμματα αξόνων).

Correspondence Table

REASON	PARTY						Active Margin
	Conservative	Labour	LibDem	Other	Refused		
Always vote that way	244	405	48	39	18		754
Best party	933	542	305	127	46		1953
My party had no chan	59	74	87	31	5		256
Active Margin	1236	1021	440	197	69		2963

Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlatio
								n
1	.253	.064			.683	.683	.018	.207
2	.172	.030			.317	1.000	.018	
Total		.094	277.916	.000 ^a	1.000	1.000		

a. 8 degrees of freedom

Πίνακας 3.1 - 3.2: Correspondence Tables.

Overview Row Points^a

REASON	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
Always vote that way	.254	.829	-.193	.048	.691	.055	.964	.036	1.000
Best party	.659	-.226	.233	.015	.133	.208	.579	.421	1.000
My party had no chan	.086	-.720	-1.212	.033	.177	.737	.341	.659	1.000
Active Total	1.000			.094	1.000	1.000			

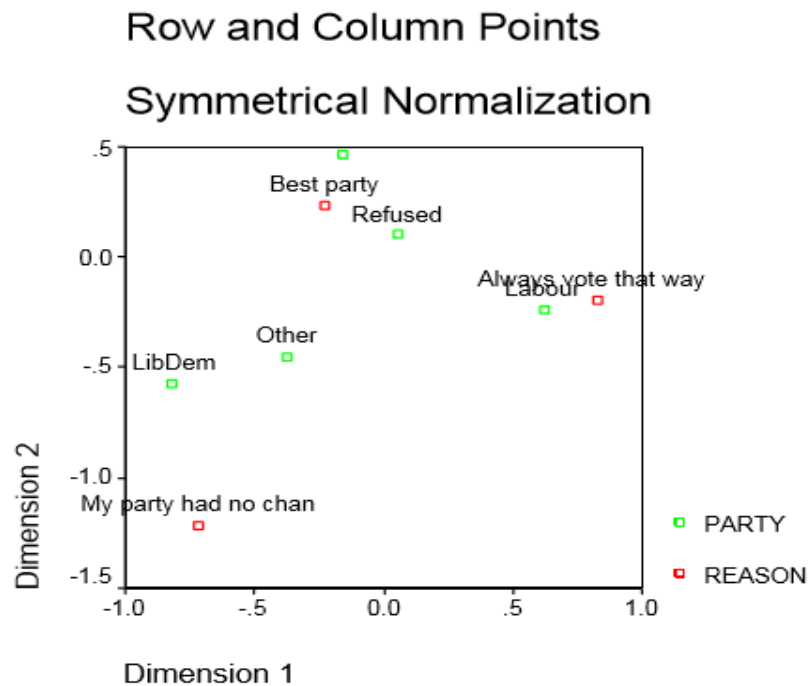
a. Symmetrical normalization

Overview Column Points^a

PARTY	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
Conservative	.417	-.162	.465	.018	.043	.524	.151	.849	1.000
Labour	.345	.620	-.235	.037	.522	.110	.911	.089	1.000
LibDem	.148	-.823	-.574	.034	.397	.284	.751	.249	1.000
Other	.066	-.374	-.456	.005	.037	.080	.497	.503	1.000
Refused	.023	.054	.101	.000	.000	.001	.294	.706	1.000
Active Total	1.000			.094	1.000	1.000			

a. Symmetrical normalization

Πίνακας 3.2 - 3.4: Γραμμές και Στήλες.



Διάγραμμα 3.3.

Ενότητα 4η: Κυρίων συνιστωσών και Παραγοντική Ανάλυση

- Εκφώνηση άσκησης: Εφαρμογή της Μεθόδου **Principal Components Analysis** (Κυρίων συνιστωσών). Τα δεδομένα του παραδείγματος δίνονται στην σελίδα του Τμήματος Στατιστικής του Πανεπιστημίου του Λονδίνου [Department of Statistics at the London School of Economics & Political Science (<http://www.lse.ac.uk/statistics/research/Social-Statistics/Multivariate-Data-Analysis/Data-and-Syntax.aspx>)]. Για περισσότερα παραδείγματα επισκεφτείτε την προτεινόμενη ηλεκτρονική διεύθυνση. The data are the percentage of the workforce employed in each of 9 industries for 26 European countries in 1979. The data file contains 10 variables: COUNTRY Country label (a string variable), AGRIC % employed in agriculture, MINING % employed in mining, MANU % employed in manufacturing, POWER % employed in power supply industries, CONSTR % employed in construction, SERVICE % employed in service industries, FINANCE % employed in finance, SOCIAL % employed in social and personal services, TRANS % employed in transport.
 - Οδηγίες επίλυσης: (Ακολουθώντας τις εντολές που παρουσιάστηκαν στην ενότητα καταλήγουμε στους παρακάτω πίνακες και διαγράμματα αξόνων).

Correlation Matrix

Correlation	AGRIC	MINING	MANU	POWER	CONSTR	SERVICE	FINANCE	SOCIAL	TRANS
	1.000	.038	-.671	-.400	-.538	-.737	-.220	-.747	-.585
		1.000	.445	.405	-.028	-.397	-.443	-.281	.157
			1.000	.385	.494	.204	-.158	.154	.351
				1.000	.060	.202	.110	.132	.375
					1.000	.366	.016	.158	.388
						1.000	.366	.572	.188
							1.000	.108	-.246
								1.000	.568
									1.000

3 Components

Communalities

	Initial	Extraction
AGRIC	1.000	.965
MINING	1.000	.858
MANU	1.000	.714
POWER	1.000	.719
CONSTR	1.000	.400
SERVICE	1.000	.776
FINANCE	1.000	.837
SOCIAL	1.000	.735
TRANS	1.000	.711

Extraction Method: Principal Component Analysis.

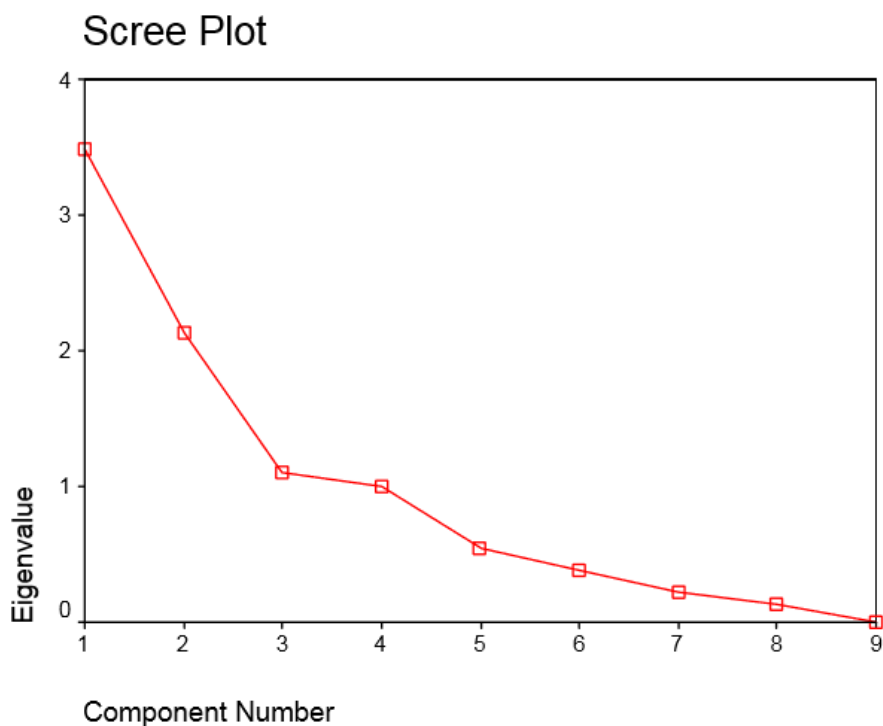
Πίνακας 4.1 - 4.2.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.487	38.746	38.746	3.487	38.746	38.746
2	2.130	23.669	62.415	2.130	23.669	62.415
3	1.099	12.211	74.625	1.099	12.211	74.625
4	.994	11.050	85.675			
5	.543	6.036	91.711			
6	.383	4.260	95.971			
7	.226	2.508	98.480			
8	.137	1.520	99.999			
9	4.563E-05	5.069E-04	100.000			

Extraction Method: Principal Component Analysis.

Πίνακας 4.3.



Διάγραμμα 4.4.

Component Matrix^a

	Component		
	1	2	3
AGRIC	-.978	7.822E-02	-5.10E-02
MINING	-2.47E-03	.902	.211
MANU	.649	.518	.158
POWER	.478	.381	.588
CONSTR	.607	7.486E-02	-.161
SERVICE	.708	-.511	.121
FINANCE	.139	-.662	.616
SOCIAL	.723	-.323	-.327
TRANS	.685	.296	-.393

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Πίνακας 4.5.

- Εκφώνηση άσκησης: Εφαρμογή της Μεθόδου **Factor Analysis** (Παραγοντική Ανάλυση). Τα δεδομένα του παραδείγματος δίνονται στην σελίδα του Τμήματος Στατιστικής του Πανεπιστημίου του Λονδίνου [Department of Statistics at the London School of Economics & Political Science (<http://www.lse.ac.uk/statistics/research/Social-Statistics/Multivariate-Data>)]

[Analysis/Data-and-Syntax.aspx](#)]. Για περισσότερα παραδείγματα επισκεφτείτε την προτεινόμενη ηλεκτρονική διεύθυνση. The data are from the text anxiety inventory which is used to assess overall anxiety associated with taking tests. The data are correlations between the responses to 20 items given by a sample of 335 male students in British Columbia. Students were asked to report how frequently they experienced various symptoms of anxiety in taking the tests.

- Brief descriptions of the tests (in the order they appear in the data file) are as follows.
 - X1=Lack of confidence during tests, X2=Uneasy, upset feeling, X3=Thinking about grades, X4=Freeze up, X5=Thinking about getting through school, X6=The harder I work, the more confused I get, X7=Thoughts interfere with concentration, X8=Jittery when taking tests, X9=Even when prepared, get nervous, X10=Uneasy before getting the test back, X11=Tense during test, X12=Exams bother me, X13=Tense/stomach upset, X14=Defeat myself during tests, X15=Panicky during tests, X16=Worry before important tests, X17=Think about failing, X18=Heart beating fast during tests, X19=Can't stop worrying, X20=Nervous during tests, forget facts.
- Οδηγίες επίλυσης: (Ακολουθώντας τις εντολές που παρουσιάστηκαν στην ενότητα καταλήγουμε στους παρακάτω πίνακες και διαγράμματα αξόνων).

Communalities

	Initial	Extraction
X1	.430	.393
X2	.443	.406
X3	.378	.356
X4	.477	.435
X5	.427	.517
X6	.302	.279
X7	.512	.539
X8	.511	.512
X9	.451	.480
X10	.270	.254
X11	.564	.569
X12	.449	.436
X13	.397	.350
X14	.454	.441
X15	.638	.643
X16	.588	.612
X17	.508	.529
X18	.470	.461
X19	.485	.431
X20	.493	.438

Extraction Method: Maximum Likelihood.

Πίνακας 4.6: Communalities.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.779	43.895	43.895	8.251	41.255	41.255
2	1.350	6.748	50.643	.831	4.155	45.410
3	.971	4.855	55.498			
4	.888	4.440	59.938			
5	.774	3.872	63.810			
6	.742	3.708	67.518			
7	.706	3.531	71.049			
8	.655	3.274	74.324			
9	.571	2.853	77.177			
10	.543	2.717	79.894			
11	.541	2.705	82.599			
12	.506	2.530	85.129			
13	.472	2.360	87.489			
14	.436	2.180	89.669			
15	.419	2.095	91.764			
16	.377	1.887	93.651			
17	.348	1.741	95.392			
18	.332	1.662	97.054			
19	.314	1.569	98.623			
20	.275	1.377	100.000			

Extraction Method: Maximum Likelihood.

Πίνακας 4.7.

Factor Matrix^a

	Factor	
	1	2
X1	.623	-6.71E-02
X2	.617	-.161
X3	.541	.251
X4	.653	9.360E-02
X5	.513	.504
X6	.489	.200
X7	.676	.287
X8	.693	-.179
X9	.660	-.212
X10	.496	-9.19E-02
X11	.733	-.179
X12	.653	-9.69E-02
X13	.589	-5.74E-02
X14	.639	.181
X15	.799	-7.11E-02
X16	.753	-.212
X17	.642	.342
X18	.665	-.136
X19	.653	-6.42E-02
X20	.662	1.732E-02

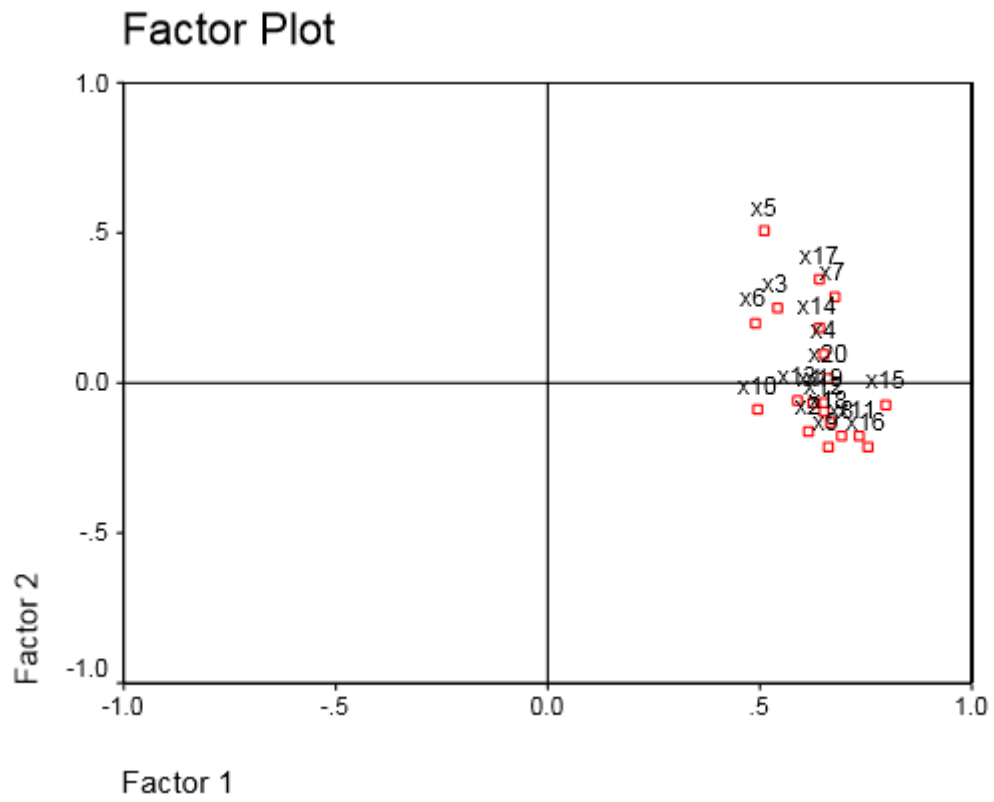
Extraction Method: Maximum Likelihood.

a. 2 factors extracted. 5 iterations required.

Goodness-of-fit Test

Chi-Square	df	Sig.
287.501	151	.000

Πίνακας 4.8.



Διάγραμμα 4.9.