

# ΣΥΣΧΕΤΙΣΗ ΚΑΙ ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ ΠΟΙΟΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ

Γιαννούλα Φλώρου  
Καθηγήτρια  
Τμήμα Λογιστικής και Χρηματοοικονομικής  
Δημοκρίτειο Πανεπιστήμιο Θράκης

Μεταπτυχιακό Πρόγραμμα  
Λογιστική και Ελεγκτική

# ΠΕΡΙΕΧΟΜΕΝΑ

- ΣΥΣΧΕΤΙΣΗ
- ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ
- ΕΦΑΡΜΟΓΗ ΠΑΡΑΓΟΝΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

# Έννοια συσχέτισης

- Όταν η μεταβολή της μιας μεταβλητής, προκαλεί κάποιο αποτέλεσμα στην άλλη.
- Π.χ η λήψη αντιπυρετικού φαρμάκου και η θερμοκρασία σώματος
- Η τιμή ενός προϊόντος και το πλήθος πωλούμενης ποσότητάς του
- Το φύλο ενός ανθρώπου και τα χρόνια ζωής του
- Η γνώμη του για την οικονομική κατάσταση και το κόμμα που επιλέγει να ψηφίσει

# Μέτρηση συσχέτισης;

- Στην περίπτωση ποσοτικών μεταβλητών, επειδή πρόκειται για αριθμητικές τιμές είναι εύκολο να κάνουμε υπολογισμούς και να μετρήσουμε το πόσο επηρεάζεται η μια από την άλλη.
- Στη περίπτωση ποιοτικών μεταβλητών δεν μπορούμε να κάνουμε αντίστοιχους υπολογισμούς, αλλά μπορούμε να δούμε μόνο «ποιοτικά» την μεταξύ τους «συσχέτιση».
- Στην περίπτωση που μια μεταβλητή είναι ποσοτική και η άλλη ποιοτική μπορεί να μετρηθεί αν υπάρχει διαφορά των μέσων τιμών της ποσοτικής μεταβλητής μεταξύ των κατηγοριών της ποιοτικής μεταβλητής.

# Συσχέτιση ποσοτικών μεταβλητών

Γραφική απεικόνιση

Μέτρηση με συντελεστή συσχέτισης

Τρόπος υπολογισμού

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Ερμηνεία

# Συσχέτιση ποιοτικών μεταβλητών

- Αν μπορούν να διαταχθούν από το μικρότερο στο μεγαλύτερο, τότε μπορούμε να υπολογίσουμε την «τάξη τους» δηλαδή τη «σειρά» τους, πρώτο, δεύτερο κλπ
- Στην περίπτωση αυτή συσχετίζουμε τις «τάξεις» (σειρές) των δύο μεταβλητών.

# Συντελεστής συσχέτισης Spearman

- Ονομάζεται και συντελεστής συσχέτισης τάξεων, γιατί κατατάσσει τις τιμές των μεταβλητών σε τάξεις.
- Χρησιμοποιείται για μεταβλητές των οποίων οι τιμές μπορούν να **διαταχθούν**.
- Ελέγχει αν υπάρχει σχέση μεταξύ των δύο μεταβλητών.

# Υπολογισμός τιμής Spearman

- $S_a$  η τυπική απόκλιση των τάξεων της  $a$  μεταβλητής
- $S_b$  η τυπική απόκλιση των τάξεων της  $b$  μεταβλητής
- $S_{ab}$  η συμμεταβλητότητα των τάξεων των  $a, b$  μεταβλητών

$$r_s = S_{ab} / S_a S_b$$

- Αν η τιμή μπορεί να θεωρηθεί  $=0$  δεν υπάρχει σχέση, διαφορετικά υπάρχει θετική ή αρνητική συσχέτιση.



# παράδειγμα

Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις

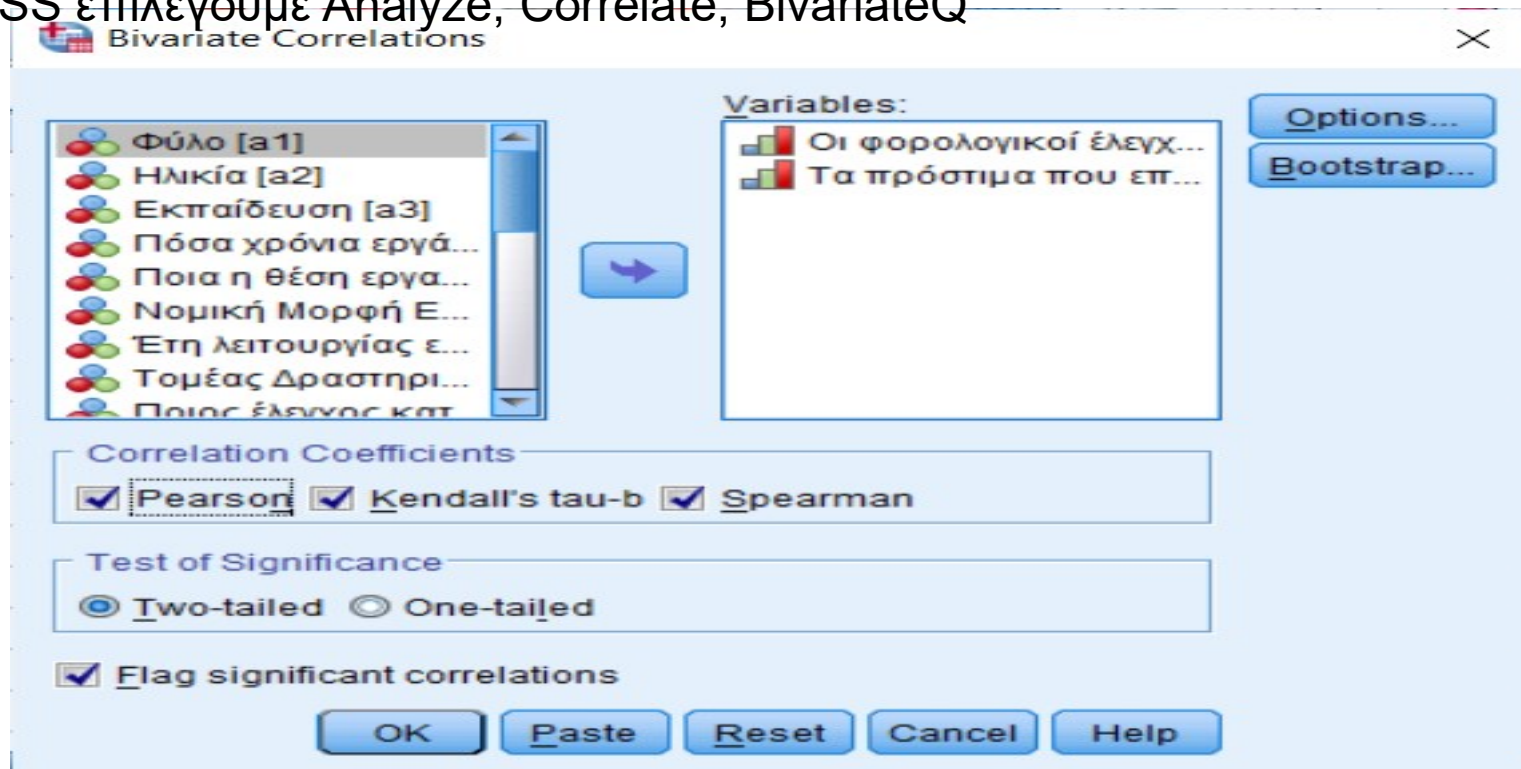
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	9	11,3	11,3	11,3
2	8	10,0	10,0	21,3
3	13	16,3	16,3	37,5
4	25	31,3	31,3	68,8
5	25	31,3	31,3	100,0
Total	80	100,0	100,0	

Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	36	45,0	45,0	45,0
2	21	26,3	26,3	71,3
3	15	18,8	18,8	90,0
4	5	6,3	6,3	96,3
5	3	3,8	3,8	100,0
Total	80	100,0	100,0	

Θέλουμε τη συσχέτιση μεταξύ των 2 μεταβλητών.

Στο SPSS επιλέγουμε Analyze, Correlate, BivariateQ



# Παράδειγμα αποτελέσματα συσχέτισης

Correlations		
	Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις	Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων
Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις	Pearson Correlation Sig. (2-tailed)	1 ,413**
	N	80

Correlations		
Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων	Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις	Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων
Kendall's tau_b	Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις	Correlation Coefficient Sig. (2-tailed) N
		1,000 ,359** 80
	Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων	Correlation Coefficient Sig. (2-tailed) N
		,359** ,000 80
Spearman's rho	Οι φορολογικοί έλεγχοι αποτελούν απαραίτητη διαδικασία στην οποία πρέπει να υποβάλλονται οι επιχειρήσεις	Correlation Coefficient Sig. (2-tailed) N
		1,000 ,426** 80
	Τα πρόστιμα που επιβάλλονται είναι ανάλογα των φορολογικών παραπτώματων	Correlation Coefficient Sig. (2-tailed) N
		,426** ,000 80

\*\* . Correlation is significant at the 0.01 level

## Τρόποι μέτρησης σχέσης δύο μεταβλητών, ανάλογα με το είδος τους

1 <sup>η</sup> μεταβλητή	2 <sup>η</sup> μεταβλητή	Μέτρηση συσχέτισης
κατηγορική	κατηγορική ή διάταξης	$\chi^2$ ή Cramer's V (1)
διχοτομική	διχοτομική	$\chi^2$ ή phi (2)
διάταξης	διάταξης	Spearman's rho ή Kendall's tau
διάταξης	διαστήματος	Αν η διάταξης έχει πολλές κατηγορίες Spearman's rho ή Kendall's tau Αν η διάταξης έχει λίγες κατηγορίες πίνακας συμπτώσεων, $\chi^2$ , ομαδοποιώντας την διαστήματος. Αν η διαστήματος μπορεί να θεωρηθεί εξαρτημένη και η διάταξης έχει λίγες κατηγορίες, ανάλυση διακύμανσης
κατηγορική ή διχοτομική	διαστήματος	Πίνακας συμπτώσεων, $\chi^2$ αν η διαστήματος χωρισθεί σε κατηγορίες. Αν η διαστήματος θεωρείται εξαρτημένη, μπορεί να γίνει και έλεγχος μέσω τιμών

- (1) **Cramer's V** = η τιμή του συντελεστή Cramer's V αναφέρεται (επίσης) στην ισχύ της σχέσης ανάμεσα στις δύο (ονομαστικές) μεταβλητές. Εμπεριέχει τον συντελεστή  $\phi$  και η διαφορά ως προς τον  $\phi$  έγκειται στο γεγονός ότι η  $\phi$  μπορεί να πάρει τιμές  $>1$ , ενώ το εύρος τιμών του Cramer's V είναι ανάμεσα στο 0 και το 1.
- (2) **Phi ( $\phi$ )** = πρόκειται για τιμή η οποία αναφέρεται στην εκτίμηση της έντασης στη σχέση δύο ονομαστικών μεταβλητών. Ο συντελεστής Phi είναι συνάρτηση του  $\chi^2$ , του μεγέθους του δείγματος και ανεξάρτητος από την ταξινόμηση των τιμών των μεταβλητών και το πλήθος των τιμών. Αν οι μεταβλητές είναι ανεξάρτητες, τότε το  $\phi$  τείνει στο μηδέν.

# FACTOR ANALYSIS

# Παραγοντική ανάλυση αναγκαιότητα

- Πολλές φορές όταν μελετούμε ένα φαινόμενο, έχουμε στη διάθεσή μας μεγάλο πλήθος στατιστικών δεδομένων, με την καταγραφή πολλών μεταβλητών, οι οποίες αλληλοεπηρεάζονται μεταξύ τους.
- Ο ερευνητής προσπαθεί να εντοπίσει τις τάσεις του φαινομένου και να εξαγάγει την καλύτερη πληροφορία για τις διαστάσεις του φαινομένου.
- Για να το πετύχει αυτό, εκτός από την στατιστική ανάλυση μιας-μιας μεταβλητής ή της σχέσης ανά δύο των μεταβλητών, χρειάζεται και η ανάλυση πολλών αλληλοσχετιζόμενων μεταβλητών μαζί.
- Οι μέθοδοι που εξετάζουν ταυτόχρονα πολλές μεταβλητές ανήκουν στο πεδίο της πολυμεταβλητής ανάλυσης δεδομένων.
- Μια από τις μεθόδους αυτές, για ποιοτικές μεταβλητές είναι η Παραγοντική Ανάλυση (Factor Analysis)

# Παραγοντική ανάλυση Ορισμός

Η παραγοντική ανάλυση είναι μια πολυμεταβλητή τεχνική, η οποία προσπαθεί να ερμηνεύσει συσχετίσεις μεταξύ παρατηρούμενων μεταβλητών, υποθέτοντας ότι αυτές οφείλονται στην ύπαρξη κάποιων κοινών παραγόντων, οι οποίοι δεν φαίνονται άμεσα ούτε μπορούν να μετρηθούν άμεσα.

Ως αποτέλεσμα η εφαρμογή της μεθόδου, πετυχαίνει τη μείωση των αρχικών μεταβλητών και την αντικατάστασή τους με νέες, που ονομάζονται **παράγοντες**.

Οι παράγοντες αναδεικνύονται μέσα από τα δεδομένα, χωρίς αρχικές υποθέσεις, γίνονται «μετρήσιμοι» και είναι αυτοί που περιγράφουν πλήρως το μελετώμενο φαινόμενο.

Η παραγοντική ανάλυση, εξαιτίας της ικανότητας της να εντοπίζει τους «αφανείς» παράγοντες, βρήκε ευρεία εφαρμογή κυρίως στα κοινωνικές επιστήμες, την οικονομία, τις έρευνες αγοράς.

# Παραγοντική ανάλυση Σκοποί

Αποκάλυψη μοτίβων σχέσεων μεταξύ των μεταβλητών.

Αποκάλυψη συστάδων μεταβλητών, που κάθε μία περιέχει αλληλοσχετιζόμενες μεταβλητές, οπότε αρκεί η μία για να περιγράψει τη συστάδα.

Μείωση μεγάλου αριθμού μεταβλητών σε μικρότερο αριθμό μη συσχετιζόμενων μεταβλητών που ονομάζονται παράγοντες (factors).

# Παραγοντική ανάλυση Χρησιμότητα

Χρησιμοποιούμε την παραγοντική ανάλυση  
σε ποιοτικές μεταβλητές, όταν:

Έχουμε μεγάλο πλήθος δεδομένων (και μεταβλητών)

Θέλουμε μια συνολική ανάλυση ερωτήσεων και όχι  
ανάλυση μία-μία ή ανά δύο μεταβλητών

Αναζητούμε την εύρεση των κυρίαρχων τάσεων των  
δεδομένων μας



# Κύρια σημεία

- Ταυτόχρονη επεξεργασία πολλών ποιοτικών μεταβλητών
- Μείωση διαστάσεων, διατήρηση πληροφορίας, ανάδειξη παραγόντων
- Μεθοδολογία χρήσης στατιστικού λογισμικού για την εφαρμογή της μεθόδου

# Πολλές μεταβλητές - διευρέυση

- Ερευνητής προσπαθεί να προβλέψει μια μεταβλητή, χρησιμοποιώντας μεγάλο αριθμό επεξηγηματικών μεταβλητών (15-40 μεταβλητές)
- Ερωτηματολόγιο με πολλές ερωτήσεις (πχ. 20-50) χρησιμοποιείται για να μετρήσει κάποια χαρακτηριστικά ή τάσεις ατόμων.

# Μείωση διαστάσεων

- Στις παραπάνω περιπτώσεις ο αριθμός των απλών συσχετίσεων μεταξύ μεταβλητών είναι πολύ μεγάλος και δύσκολο να εξαγάγουμε συμπεράσματα. Αν οι μεταβλητές μας είναι 30, υπάρχουν 435 τιμές των ανά δύο συσχετίσεων!!!
- Προσπαθούμε να βρούμε τρόπους μείωσης του αριθμού των μεταβλητών ομαδοποιώντας κάποιες από αυτές, μετά την συλλογή των δεδομένων. Φροντίζουμε να αναδεικνύονται οι κύριες τάσεις των δεδομένων.

# παράδειγμα

Ερώτηση	καθόλου	κάποιες φορές	μέτρια	συχνά	πολύ συχνά
κωδικός	1	2	3	4	5
1. Προσέχω να είμαι υγιής.	2,7	10,8	35,1	37,8	13,5
2. Γυμνάζομαι	24,3	35,1	29,7	8,1	2,7
3. Καπνίζω	83,8	5,4	8,1	2,7	,0
4. Καταναλώνω την απαραίτητη ποσότητα νερού καθημερινά	10,8	2,7	32,4	29,7	24,3
5. Εγώ αγοράζω τα τρόφιμα μου	5,4	16,2	18,9	27,0	32,4
6. Παραλείπω το πρωινό γεύμα	51,4	18,9	10,8	2,7	16,2
7. Παραλείπω το μεσημεριανό γεύμα.	48,6	35,1	13,5	2,7	,0
8. Παραλείπω το βραδινό γεύμα.	32,4	35,1	10,8	16,2	5,4
9. Λαμβάνω φάρμακα.	54,1	21,6	2,7	5,4	16,2
10. Λαμβάνω βιταμίνες.	56,8	24,3	2,7	8,1	8,1
11. Λαμβάνω συμπληρώματα διατροφής.	81,1	2,7	13,5	,0	2,7
12. Φροντίζω να καταναλώνω τρόφιμα πλούσια σε φυτικές ίνες.	8,1	10,8	32,4	24,3	24,3
13. Φροντίζω να καταναλώνω τρόφιμα με χαμηλά λιπαρά.	13,5	32,4	27,0	21,6	5,4
14. Καταναλώνω τρόφιμα εμπλουτισμένα (ω3, ω6, προβιοτικά, φυτοστερόλες)	10,8	18,9	27,0	32,4	10,8
15. Καταναλώνω τρόφιμα χωρίς γλουτένη η χωρίς λακτόζη.	43,2	29,7	13,5	10,8	2,7
16. Καταναλώνω εκχυλίσματα βοτάνων	40,5	16,2	27,0	10,8	5,4
17. Καταναλώνω δυναμωτικά ποτά.	83,8	13,5	2,7	,0	,0

"ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΕΡΕΥΝΑΣ"

Γ. ΦΛΩΡΟΥ

# Συνολική ανάλυση

- Κυρίως ενδιαφέρει η συνολική ανάλυση των δεδομένων, ώστε να προκύψουν οι ομάδες ατόμων που απαντούν με παρόμοιο τρόπο στις ερωτήσεις καθώς και οι μεταβλητές που χαρακτηρίζουν τις ομάδες αυτές.
- Δεν απαιτεί να ορίσουμε κάποια μεταβλητή ως εξαρτημένη από άλλες μεταβλητές, αλλά γίνεται συνολική ανάλυση όλων των μεταβλητών. Έχει στόχο να εξηγήσει τις συσχετίσεις που υπάρχουν στο σύνολο των μεταβλητών. Αρκετές φορές οι συσχετίσεις αυτές δεν εξηγούν αιτιώδη σχέση μεταξύ μεταβλητών αλλά οφείλονται στο γεγονός ότι οι συσχετισμένες μεταβλητές έχουν κοινή εξάρτηση από κάποια ή κάποιες άλλες μεταβλητές

# παράγοντες

- Οι μεταβλητές όμως που επηρεάζουν τις υπόλοιπες δεν είναι γνωστές, ή δεν μπορούν να καταγραφούν. Για το λόγο αυτό υπολογίζονται οι «παράγοντες» που αντιστοιχούν στις άγνωστες μεταβλητές.
- Οι παράγοντες εντοπίζονται με τη σειρά σπουδαιότητας τους (πρώτος, δεύτερος κλπ) και είναι ανεξάρτητοι μεταξύ τους.
- Για κάθε παράγοντα υπολογίζονται οι συσχετίσεις του με τις αρχικές μεταβλητές. Οι μεταβλητές με μεγάλες συσχετίσεις «ερμηνεύουν» το συγκεκριμένο παράγοντα.

# Μέθοδοι μείωσης διαστάσεων

## Factor analysis

- Εντοπίζουμε γραμμικούς συνδυασμούς των αρχικών μεταβλητών τους οποίους ονομάζουμε παράγοντες (factors)

Διακρίνουμε την

Διερευνητική παραγοντική ανάλυση

Επιβεβαιωτική παραγοντική ανάλυση

# μεθοδολογία εύρεσης παραγόντων

- Η ολική διασπορά των δεδομένων (άθροισμα από όλες τις διασπορές των μεταβλητών), διαχωρίζεται, βρίσκοντας έναν παράγοντα (δηλαδή γραμμικό συνδυασμό μεταβλητών) που έχει τη μέγιστη δυνατή διασπορά.
- Κατόπιν βρίσκουμε άλλον παράγοντα (νέο γραμμικό συνδυασμό μεταβλητών) με μέγιστη δυνατή διασπορά, αλλά ασυσχέτιστο με τον προηγούμενο παράγοντα και συνεχίζουμε με στόχο να εντοπίσουμε όλους τους «αφανείς» παράγοντες ....

Π.χ.

$$Y_1 = \alpha X_1 + \beta X_2 + \gamma X_3 \quad Y_2 = \alpha X_4 + \beta X_5 + \gamma X_9 \quad \dots$$

Οι νέες μεταβλητές που προκύπτουν από τους γραμμικούς συνδυασμούς ονομάζονται παράγοντες και η μέθοδος παραγοντική ανάλυση



# Προϋπόθεση μη συσχέτισης

Για την εύρεση των κυρίων συνιστωσών (παραγόντων) προσπαθούμε να βρούμε γραμμικούς συνδυασμούς συσχετισμένων μεταβλητών και κατόπιν ασυσχέτιστων κυρίων συνιστωσών.

Αν όλες οι αρχικές μεταβλητές είναι ασυσχέτιστες δεν υπάρχει λόγος να προχωρήσουμε τη διαδικασία.

Ο έλεγχος για την Υπόθεση

$H_0$  ότι όλες οι μεταβλητές είναι μεταξύ τους ασυσχέτιστες

είναι ο έλεγχος με το κριτήριο Bartlett's sphericity test.

Αν η  $H_0$  απορρίπτεται ( $\text{sig} < 0.05$ ) συνεχίζουμε την ανάλυση.

# Μετατροπή σε ασυσχέτιστους παράγοντες

- Έτσι με την χρήση της παραγοντικής ανάλυσης, ένα σύνολο συσχετισμένων μεταβλητών, μετατρέπεται σε ένα σύνολο **ασυσχέτιστων παραγόντων**.
- Με την μετατροπή αυτή, ευελπιστούμε
  - να έχουμε λιγότερες από τις αρχικές μεταβλητές,
  - να ερμηνεύουν ένα μεγάλο μέρος της διασποράς των δεδομένων,
  - να μπορούμε να εξηγήσουμε – περιγράψουμε.

# Μείωση διαστάσεων

- Συνήθως οι κύριες συνιστώσες που προκύπτουν είναι 2 έως 6 και ερμηνεύουν μεγάλο μέρος της διασποράς των δεδομένων (περίπου 75%).
- Οι παράγοντες ερμηνεύονται – περιγράφονται χρησιμοποιώντας τους συντελεστές των αρχικών μεταβλητών στο γραμμικό συνδυασμό των συνιστωσών. Οι συντελεστές αυτοί ονομάζονται βαρύτητες στους παράγοντες (factor loadings).

# Ονομασία παραγόντων

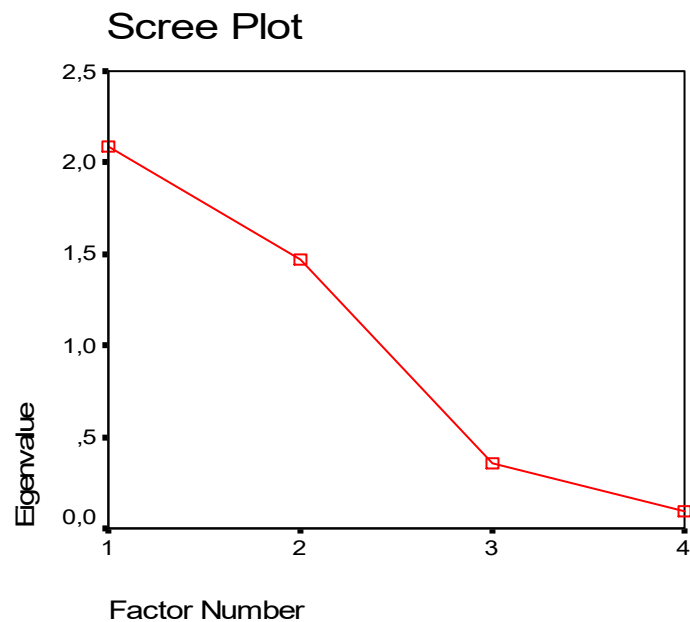
- Ανάλογα με τις μεταβλητές που έχουν τη μεγαλύτερη βαρύτητα σε κάποιο παράγοντα, ονομάζουμε τον παράγοντα αυτόν με γενικό όνομα που περιλαμβάνει όσο μπορούμε καλύτερα τις μεταβλητές με τη μεγάλη βαρύτητα.
  - Για παράδειγμα ένας παράγοντας που αποτελείται από βαθμολογία μαθητών στα καλλιτεχνικά, την έκθεση και τη γλώσσα, θα μπορούσε να ονομασθεί παράγοντας έκφρασης.

# Πόσους παράγοντες;

- Η εύρεση κυρίων παραγόντων είναι μια επαναλαμβανόμενη διαδικασία και τελειώνει όταν το πλήθος των παραγόντων είναι ίσο με το πλήθος των αρχικών μεταβλητών.
- Ο σκοπός μας όμως είναι η μείωση του αριθμού των μεταβλητών που περιγράφουν τα δεδομένα μας.
  - Διάφορα κριτήρια προτάθηκαν. Το πιο διαδεδομένο είναι αυτό που χρησιμοποιεί την ερμηνευόμενη διασπορά με τη γραφική παράσταση **scree test**. Παρουσιάζεται η ερμηνευμένη διασπορά ως προς τον αύξοντα αριθμό του παραγοντα.  
Σταματάμε εκεί που υπάρχει «απότομη» μεταβολή, προς το κάτω.

# Παράδειγμα scree plot

Απότομη μεταβολή παρατηρείται από το 2 στο 3.  
Επομένως ερμηνεύουμε τους δύο πρώτους παράγοντες.



# Δυσκολίες

- Πολλές φορές η σύνθεση των παραγόντων, είναι τέτοια που δεν προκύπτει εύκολα η ερμηνεία τους.
- Πιο είναι το καταλληλότερο μέγεθος δείγματος για αξιόπιστη παραγοντική ανάλυση;  
(προτείνεται από πολλούς 10-20 παρατηρήσεις ανά μεταβλητή)
- Είναι αξιόπιστοι οι παράγοντες; Αν αλλάξουμε το δείγμα θα βρούμε τους ίδιους ή διαφορετικούς;

# Χρησιμότητα αποτελεσμάτων

- Αν είναι εύκολη η ερμηνεία των παραγόντων, περιγράφουμε τα δεδομένα μας μόνο μ' αυτές, χωρίς να χρειάζεται η αναφορά όλων των αρχικών μεταβλητών.
- Μπορούμε να χρησιμοποιήσουμε τις συντεταγμένες των παραγόντων σε περαιτέρω ανάλυση των δεδομένων μας με άλλη μέθοδο, όπως παλινδρόμηση.
- Μπορούμε να απεικονίσουμε γραφικά τα δεδομένα μας χρησιμοποιώντας το επίπεδο των δύο πρώτων παραγόντων όπου βρίσκεται το μεγαλύτερο ποσοστό της πληροφορίας (ερμηνευμένης διασποράς).

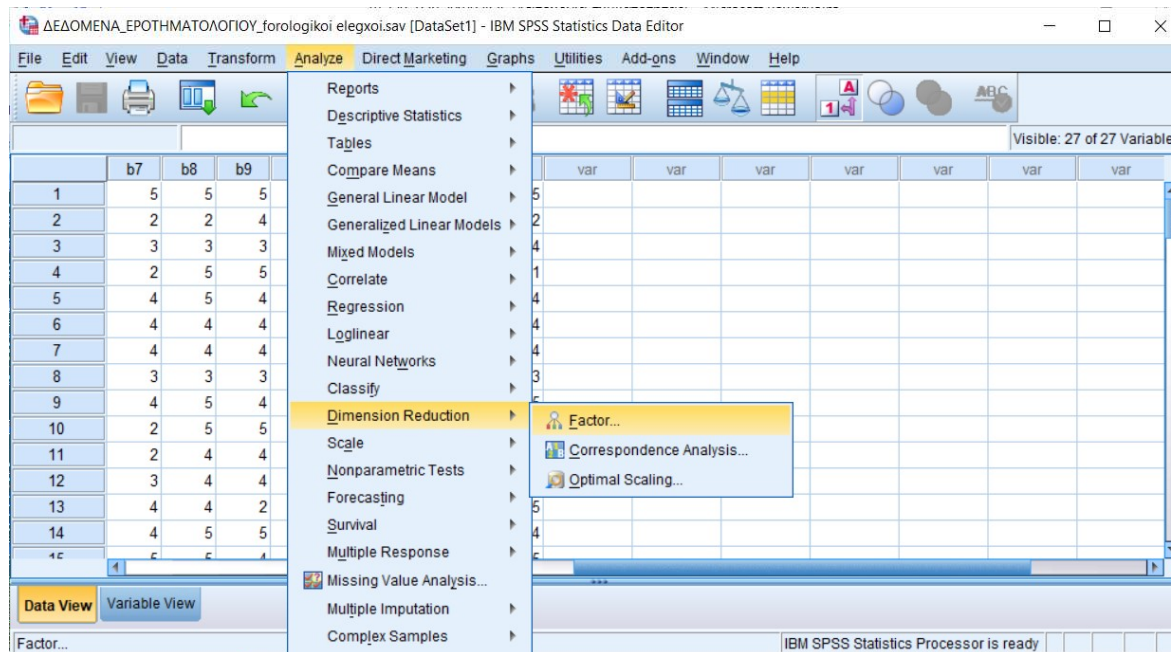


# Εφαρμογή με το SPSS

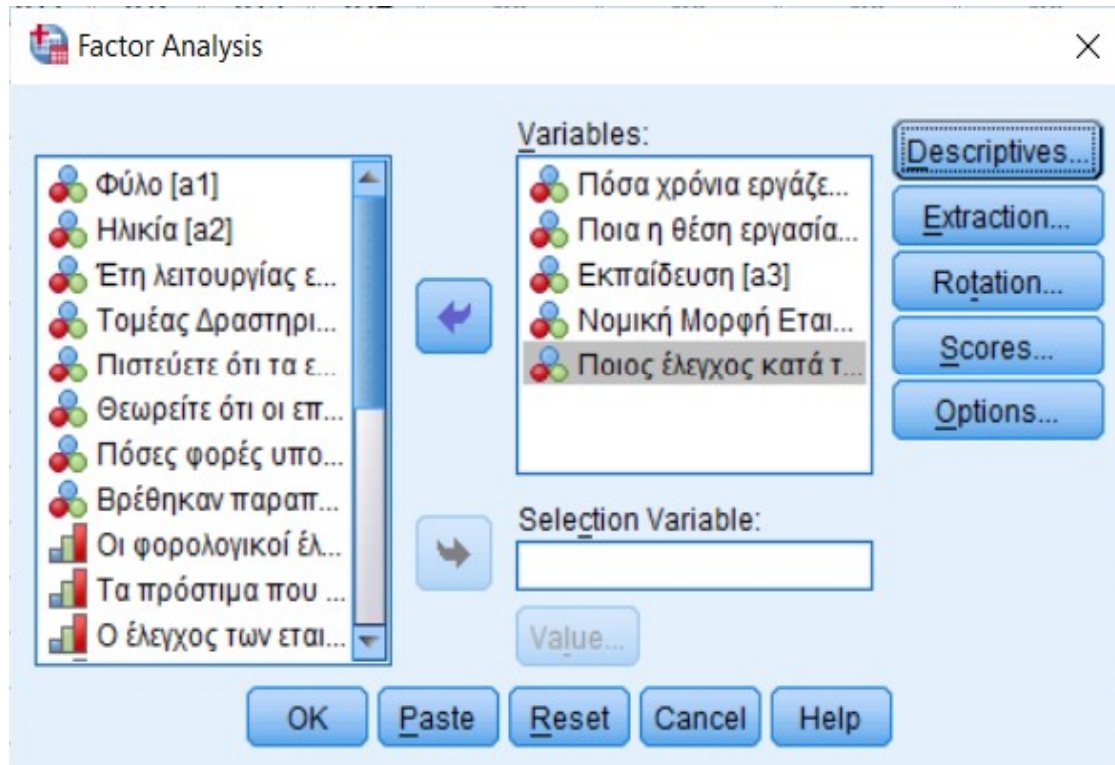
Αν θέλουμε να εφαρμόσουμε την παραγοντική ανάλυση σε αρχείο δεδομένων στο SPSS διαλέγουμε από το μενού

● Analyze, Dimension reduction, Factor...

Στη συνέχεια επιλέγουμε τις μεταβλητές που θα χρησιμοποιήσουμε στην ανάλυση.



# Ανάλυση με το SPSS



Επιπρόσθετα

Πλήκτρο Descriptives  
επιλέγουμε **KMO**

Πλήκτρο Extraction  
επιλέγουμε **Scree plot**

Πλήκτρο Rotation  
επιλέγουμε **Loading plots**

Πλήκτρο Scores επιλέγουμε  
**Save as variable** για αποθήκευση  
των νέων συντεταγμένων

Και πατάμε το OK.

# Επιλογές πλήκτρου Extraction

- 1. Principal components
- 2. Unweighted least squares
- 3. Generalized least squares
- 4. Maximum likelihood
- 5. Principal axis factoring
- 6. Alpha factoring
- 7. Image factoring

# Περιστροφή των παραγόντων

- Με την περιστροφή των παραγόντων προσπαθούμε να κάνουμε τους παράγοντες πιο ερμηνεύσιμους.
- Με αυτή την μέθοδο δεν αλλάζουν κάποια από τα χαρακτηριστικά του μοντέλου, όπως η καλή προσαρμοστικότητα και το ποσό διακύμανσης συνδιακύμανσης που ερμηνεύεται από το μοντέλο, παρά μόνο οι τιμές των επιβαρύνσεων.
- Κάνοντας περιστροφή, ελπίζουμε ότι οι επιβαρύνσεις κάποιων παραγόντων θα είναι μεγάλες σε απόλυτη κλίμακα μόνο για κάποιες από τις μεταβλητές και έτσι, βλέποντας ποιες μεταβλητές εξαρτώνται με ποιους παράγοντες, να μπορέσουμε να δώσουμε ερμηνεία σε αυτούς. (Πραμαγγιούλης, 2008)
- Μέθοδοι περιστροφής
  - Varimax: Προσπαθεί να ελαχιστοποιήσει τον αριθμό των μεταβλητών που έχουν μεγάλες επιβαρύνσεις για κάθε παράγοντα.
  - Quartimax: Προσπαθεί να ελαχιστοποιήσει τον αριθμό των παραγόντων που εξηγούν μια μεταβλητή
  - Equimax: Συνδυασμός των δύο παραπάνω
  - Oblique: Μη ορθογώνια περιστροφή, οι άξονες που προκύπτουν δεν είναι πια ορθογώνιοι (και άρα οι παράγοντες δεν είναι ανεξάρτητοι). Η ερμηνεία των αποτελεσμάτων είναι πιο δύσκολη

## Επιλογές πλήκτρου Rotation

- 1. Varimax
- 2. Direct Oblimin
- 3. Quatrimax
- 4. Equamax
- 5. Promax

## Επιλογές πλήκτρου Scores

- 1. Regression
- 2. Barthlett
- 3. Anderson- Rubin

# Παράδειγμα ανάλυσης ερωτηματολογίου με παραγοντική ανάλυση

- 1.6. Ποιό είναι το επάγγελμά σας ;  
 1. Ιδιωτικός υπάλληλος 2. Δημόσιος υπάλληλος 3. Ελεύθερος επαγγελματίας 4. Ιδιοκτήτης εμπορικής επιχείρησης 5. Αγρότης
- 2.5. Για πόσο χρονικό διάστημα χρησιμοποιείτε τη Εταιρεία Λογιστών με τη οποία αποπερατώνετε στη παρούσα συγκυρία τις Λογιστικές σας εργασίες;  
 \_1. 0-5 χρόνια \_2. 6-10 χρόνια \_3. 11-15 χρόνια \_4. 16-20 χρόνια \_5. άνω των 22 χρόνων
- 2.6. Ποια είναι η σημερινή θέση σας στη Εταιρεία στη οποία εργάζεσθε;  
 \_1 Πρόεδρος \_2 Διευθύνων Σύμβουλος \_3 Διευθυντής \_4 Ανώτερο στέλεχος \_5 Στέλεχος της Γραμματείας της Εταιρείας. \_6. Άλλη θέση
- 2.7. Έχετε βαρύνουσα και αποφασιστική γνώμη στη εκλογή μιας Εταιρείας Λογιστών για τη Επιχείρησή σας; \_1 ΝΑΙ \_0 ΟΧΙ

## ΕΤΟΙΜΟΤΗΤΑ ΕΞΥΠΗΡΕΤΗΣΗΣ ΤΟΥ ΠΡΟΣΩΠΙΚΟΥ (responsiveness)

- 3.1. Είστε ικανοποιημένος από το βαθμό επιθυμίας να σας εξυπηρετήσουν στη Εταιρεία Λογιστών; \_1 \_2 \_3 \_4 \_5
- 3.2. Είστε ικανοποιημένος από τη πρόθυμη εξυπηρέτηση των πελατών; \_1 \_2 \_3 \_4 \_5
- 3.3. Είστε ικανοποιημένος από τη ενημέρωση των πελατών για το χρόνο εξυπηρέτησης; \_1 \_2 \_3 \_4 \_5
- 3.4. Είστε ικανοποιημένος από τη ετοιμότητα να απαντήσει η Εταιρεία Λογιστών σε ένα αίτημα ενός πελάτη \_1 \_2 \_3 \_4 \_5

## ΑΥΤΟΠΕΠΟΙΘΗΣΗ ΓΙΑ ΤΗ ΓΝΩΣΗ ΚΑΙ ΣΩΣΤΗ ΕΞΥΠΗΡΕΤΗΣΗ ΤΩΝ ΥΠΑΛΛΗΛΩΝ (assurance)

- 3.5. Είστε ικανοποιημένος από την εμπιστοσύνη που εμπνέουν οι υπάλληλοι στους πελάτες; \_1 \_2 \_3 \_4 \_5
- 3.6. Είστε ικανοποιημένος από τη ευγένεια πού έχουν οι υπάλληλοι με συνέπεια προς τους πελάτες; \_1 \_2 \_3 \_4 \_5
- 3.7. Είστε ικανοποιημένος από τις γνώσεις πού έχουν οι υπάλληλοι για να απαντήσουν στις ερωτήσεις των πελατών; \_1 \_2 \_3 \_4 \_5
- 3.8. Είστε ικανοποιημένος από τη ικανότητα των υπαλλήλων να εμπνέουν ασφάλεια στους πελάτες όσο αφορά τις συναλλαγές τους; \_1 \_2 \_3 \_4 \_5

## ΠΡΟΣΩΠΙΚΗ ΦΡΟΝΤΙΔΑ ΓΙΑ ΤΟΥΣ ΠΕΛΑΤΕΣ (empathy)

- 3.9. Είστε ικανοποιημένος από τον βαθμό καταλληλότητας των ωρών εξυπηρέτησης; \_1 \_2 \_3 \_4 \_5
- 3.10. Είστε ικανοποιημένος από το βαθμό της προσωπικής φροντίδας με τη οποία αντιμετωπίζεται ο κάθε πελάτης; \_1 \_2 \_3 \_4 \_5
- 3.11. Είστε ικανοποιημένος από τον βαθμό της αντίληψης από το προσωπικό των αναγκών των πελατών; \_1 \_2 \_3 \_4 \_5
- 3.12. Είστε ικανοποιημένος από το βαθμό κατά τον οποίο οι υπάλληλοι ενεργούν έχοντας ως πρώτη προτεραιότητα το συμφέρον των πελατών; \_1 \_2 \_3 \_4 \_5

## ΕΜΦΑΝΙΣΗ ΤΩΝ ΥΠΑΛΛΗΛΩΝ ΚΑΙ ΥΠΑΡΞΗ ΜΗΧΑΝΗΜΑΤΩΝ (tangibles)

- 3.13. Είστε ικανοποιημένος από την κομψή και επαγγελματική εμφάνιση των υπαλλήλων; \_1 \_2 \_3 \_4 \_5
- 3.14. Είστε ικανοποιημένος από τη εμφάνιση των γραφείων της Εταιρείας και τις ευκολίες που παρέχονται στους πελάτες κατά τη παραμονή τους (έπιπλα , υποδομή, parking , κ.τ.λ.); \_1 \_2 \_3 \_4 \_5
- 3.15. Είστε ικανοποιημένος από τα μοντέρνα μηχανήματα πού υπάρχουν στα γραφεία της Εταιρείας ( Η/Υ, εκτυπωτές,τηλέφωνα, fax, internet, κ.τ.λ.); \_1 \_2 \_3 \_4 \_5

## ΑΞΙΟΠΙΣΤΙΑ ΣΤΗ ΕΞΥΠΗΡΕΤΗΣΗ (reliability)

- 3.16. Είστε ικανοποιημένος από τη εξυπηρέτηση που παρέχεται στο χρόνο που οι υπάλληλοι της Εταιρείας έχουν υποσχεθεί στους πελάτες; \_1 \_2 \_3 \_4 \_5
- 3.17. Είστε ικανοποιημένος από τη αξιοπιστία στη αντιμετώπιση του κάθε ενός προβλήματος που έχει σχέση με τις συναλλαγές των πελατών; \_1 \_2 \_3 \_4 \_5
- 3.18. Είστε ικανοποιημένος από τον βαθμό παροχής της εξυπηρέτησης , ακριβώς όπως έχει προσυμφωνηθεί μεταξύ Εταιρείας και πελατών; \_1 \_2 \_3 \_4 \_5
- 3.19. Είστε ικανοποιημένος από τη καταχώρηση και διατήρηση Λογιστικών Εγγραφών απαλλαγμένων από σφάλματα; \_1 \_2 \_3 \_4 \_5

# Παράδειγμα αποτελεσμάτων μιας παραγοντικής ανάλυσης

**KMO and Bartlett's Test**

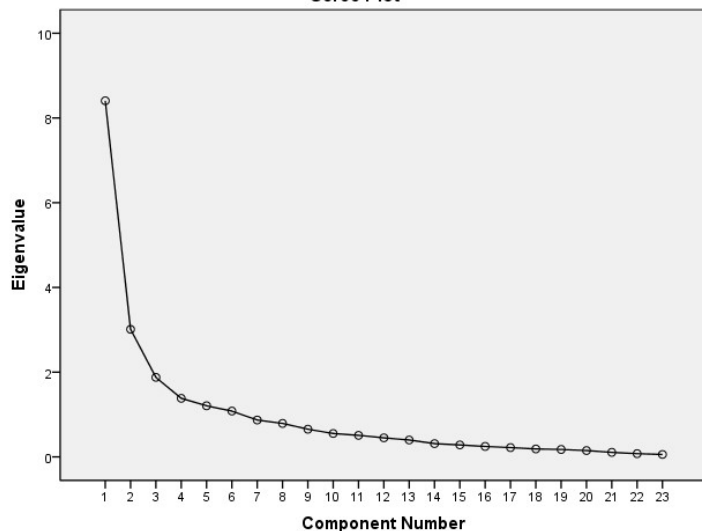
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,755
Bartlett's Test of Sphericity	Approx. Chi-Square	768,372
	df	253
	Sig.	,000

**Total Variance Explained**

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8,406	36,548	36,548	8,406	36,548	36,548
2	3,011	13,093	49,641	3,011	13,093	49,641
3	1,876	8,157	57,798	1,876	8,157	57,798
4	1,382	6,008	63,806	1,382	6,008	63,806
5	1,207	5,246	69,052	1,207	5,246	69,052
6	1,081	4,700	73,752	1,081	4,700	73,752
7	,871	3,786	77,538			
8	,790	3,435	80,973			
9	,653	2,838	83,811			
10	,552	2,400	86,211			
11	,509	2,212	88,423			
12	,450	1,958	90,381			
13	,398	1,732	92,113			
14	,313	1,359	93,472			
15	,282	1,226	94,698			
16	,247	1,074	95,772			
17	,220	,956	96,727			
18	,188	,817	97,544			
19	,175	,761	98,305			
20	,149	,647	98,952			
21	,107	,466	99,419			
22	,076	,331	99,750			
23	,057	,250	100,000			

Extraction Method: Principal Component Analysis.

**Scree Plot**



# Συντεταγμένες των μεταβλητών

Component Matrix<sup>a</sup>

	Component					
	1	2	3	4	5	6
1.6	,066	,765	,097	,159	,417	-,086
2.5	,121	,557	,269	-,017	-,190	,587
2.6	-,079	-,839	-,086	,178	-,015	,047
2.7	,234	,672	,047	,070	,195	-,466
3.1	,569	,358	-,276	-,088	-,355	,187
3.2	,667	,126	,282	-,380	-,202	,180
3.3	,710	,063	,300	-,242	,067	-,262
3.4	,689	-,304	,234	,140	-,255	-,108
3.5	,748	-,127	-,424	-,021	,249	-,030
3.6	,591	-,162	-,311	-,170	,381	,174
3.7	,632	-,411	,062	-,009	,319	,017
3.8	,690	-,369	,195	-,111	,320	,253
3.9	,634	,006	,536	,200	,153	,046
3.10	,744	-,088	,218	-,036	-,087	,042
3.11	,609	-,398	,289	-,014	-,042	-,065
3.12	,645	-,152	-,200	-,140	-,261	-,436
3.13	,611	-,024	-,555	,251	-,213	,077
3.14	,589	,160	-,177	,656	,040	,130
3.15	,495	,004	,235	,662	-,170	-,075
3.16	,663	,242	-,347	-,101	-,067	,059
3.17	,816	,209	-,080	-,042	,251	,113
3.18	,729	,153	-,337	-,166	-,086	-,064
3.19	,670	,177	,272	-,121	-,243	-,142

Extraction Method: Principal Component Analysis.

a. 6 components extracted.

Ο πρώτος παράγοντας περιγράφεται από τις ερωτήσεις 3.2, 3.3, 3.4, 3.5, 3.8, 3.9, 3.10, 3.16, 3.17, 3.18, 3.19

«ικανοποίηση»

Ο δεύτερος παράγοντας από τις 1.6, 2.5, 2.6, 2.7 «επάγγελμα»



# Ανάλυση σε κύριες συνιστώσες

- Όταν οι μεταβλητές είναι ποσοτικές, παρόμοια διαδικασία γίνεται για την εύρεση των **κυρίων συνιστωσών**.
- Δηλαδή χρησιμοποιούμε τεχνική για την μείωση των ποσοτικών μεταβλητών, μέσω της εύρεσης κυρίων συνιστωσών, που είναι συσχετισμένες με τις αρχικές μεταβλητές αλλά ασυσχέτιστες μεταξύ τους.
- Δεν υποθέτουμε ότι υπάρχουν «αφανείς» μεταβλητές που προσπαθούμε να «ανακαλύψουμε» και να μετρήσουμε όπως συμβαίνει με την παραγοντική ανάλυση.

# παράδειγμα

- Τα παρακάτω δεδομένα αφορούν στο σκορ 6 παιχνιδιών για 9 φοιτητές.
- Αναζητούμε συσχετίσεις μεταξύ των επιδόσεων

Case Summaries<sup>a</sup>

	PING_PONG	STAUROLEXO	BELAKIA	SCRABLE	JOGLER	ORTHOGRAFIA
1	10	15	8	26	15	8
2	6	16	5	25	12	9
3	2	11	1	22	7	6
4	5	16	3	28	11	9
5	7	15	4	24	12	7
6	8	13	4	23	14	6
7	6	17	3	29	10	9
8	2	18	1	28	8	8
9	5	14	2	25	10	6

a. Limited to first 100 cases.

# Ανάλυση με το SPSS

- Analyze, Dimension reduction, Factor...

Επιλέγουμε τις μεταβλητές που θα χρησιμοποιήσουμε στην ανάλυση.

Επιπρόσθετα

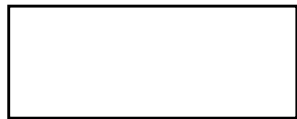
Πλήκτρο Descriptives επιλέγουμε **KMO**

Πλήκτρο Rotation επιλέγουμε **Loading plots**

Πλήκτρο Extraction επιλέγουμε **Scree plot**

Πλήκτρο Scores επιλέγουμε **Save as variable** για αποθήκευση των νέων συντεταγμένων

Και OK.



# Αποτελέσματα ανάλυσης

**Total Variance Explained**

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,951	49,186	49,186	2,951	49,186	49,186
2	2,579	42,981	92,167	2,579	42,981	92,167
3	,264	4,401	96,567			
4	,124	2,062	98,630			
5	,058	,974	99,604			
6	,024	,396	100,000			

Extraction Method: Principal Component Analysis.

Η διασπορά που ερμηνεύεται από την πρώτη κύρια συνιστώσα είναι 49,186.

Η διασπορά που ερμηνεύεται από τη δεύτερη κύρια συνιστώσα είναι 42,981.

Η διασπορά που ερμηνεύεται από τις δύο κύριες συνιστώσες μαζί είναι 92,187.

# Αποτελέσματα ανάλυσης

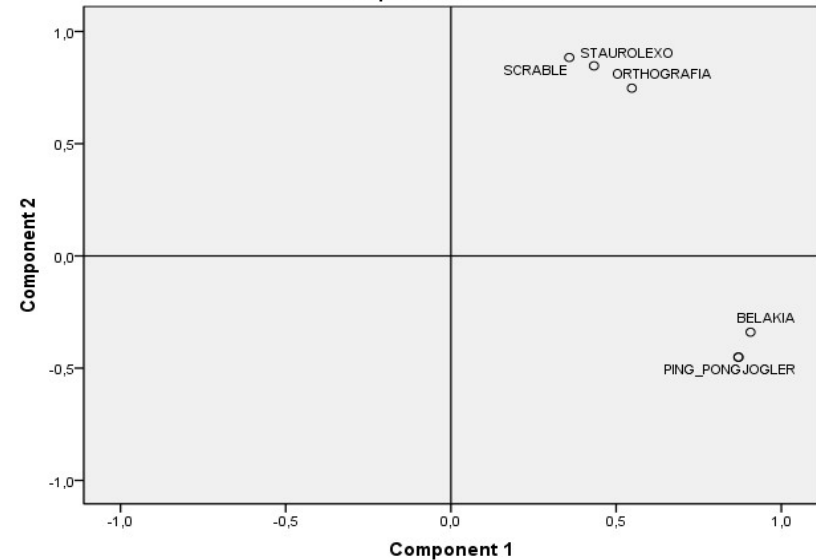
Component Matrix<sup>a</sup>

	Component	
	1	2
PING_PONG	,870	-,451
STAUROLEXO	,433	,847
BELAKIA	,907	-,340
SCRABLE	,358	,884
JOGLER	,870	-,451
ORTHOGRAFIA	,548	,747

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Component Plot



Η συσχέτιση της μεταβλητής PING PONG με την πρώτη κύρια συνιστώσα είναι 0,870 ενώ με τη δεύτερη κύρια συνιστώσα είναι -0,451.....

Οι μεταβλητές που συσχετίζονται με την πρώτη κύρια συνιστώσα είναι οι PING PONG, BELAKIA, JOGLER. Μπορούμε να την χαρακτηρίσουμε **ΕΠΙΔΕΞΙΟΤΗΤΑ**

Οι μεταβλητές που συσχετίζονται με τη δεύτερη κύρια συνιστώσα είναι οι STAUROLEXO, SCRABLE, ORTHOGRAFIA. Μπορούμε να τον χαρακτηρίσουμε **ΣΚΕΨΗ**

Παράδειγμα επεξεργασία δεδομένων που περιγράφονται από μεγάλο πλήθος μεταβλητών

- Χρησιμοποιείτε το αρχείο car-sales του SPSS, όπου περιγράφονται δεδομένα πωλήσεων αυτοκινήτων και εφαρμόστε τη μέθοδο της ανάλυσης σε κύριες συνιστώσες.
- Χρησιμοποιείτε τις ποσοτικές μεταβλητές **Vehicle type** έως **Fuel efficiency**.
- Σχολιάστε τα αποτελέσματα.
- (αντίστοιχο παράδειγμα - Πραμαγγιούλης, 2008, σελ 116)

# FACTOR ANALYSIS

Παραγοντική ανάλυση

ΕΦΑΡΜΟΓΗ ΣΕ ΠΡΑΓΜΑΤΙΚΟ  
ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

# ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ

- Έρευνα για τις αγοραστικές συνήθειες καταναλωτών πριν και μετά την κρίση.
- Διαθέτουμε απαντήσεις στις παρακάτω ερωτήσεις:
  - Ηλικία (σε έτη)
  - Φύλο (άνδρας - γυναίκα)
  - Τόπος κατοικίας (μεγάλη πόλη – άλλες περιοχές)
  - Επάγγελμα (φοιτητής – άνεργος- δημόσιος τομέας – ιδιωτικός τομέας)
  - Κατηγορία μηνιαίου εισοδήματος (πριν το 2010, μετά το 2010)
    - 1:<300 2:300-600 3:600-900 4: 900-1200 5:>1200
- 26 ερωτήσεις με 5 πολλαπλές απαντήσεις στη μορφή
  - 1-πάρα πολύ 2-πολύ 3-μέτρια 4-λίγο 5-καθόλου ( γνώμη-σκέψη)
  - 1:πολύ συχνά 2: 1φορά/εβδομάδα 3: 1φορά/15μερο 4:1φορά/μήνα 5-σπάνια (συχνότητα)
  - 1:<150 2:150-300 3:300-450 4: 450-600 5:>600 (δαπάνη)



# Ερωτήσεις με πέντε πολλαπλές απαντήσεις

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
εισόδημα πριν	660	1	5	2,47	1,230
εισόδημα μετά	660	1	5	2,27	1,204
γνώμη πριν	660	1	5	2,46	,853
γνώμη μετά	660	1	5	2,96	1,010
συχνότητα αγοράς ενδυσης πριν	660	1	5	3,90	1,123
σκέψη αγοράς ενδυσης πριν	660	1	5	3,23	1,107
δαπάνη αγοράς ενδυσης πριν	660	1	5	1,58	,880
συχνότητα αγοράς ενδυσης μετά	660	1	5	4,22	1,034
σκέψη αγοράς ενδυσης μετά	660	1	5	2,43	1,110
δαπάνη αγοράς ενδυσης μετά	660	1	5	1,33	,693
συχνότητα αγοράς καλλυντικά πριν	660	1	5	3,97	1,136
σκέψη αγοράς καλλυντικά πριν	660	1	5	3,26	1,184
δαπάνη αγοράς καλλυντικά πριν	660	1	5	1,29	,715
συχνότητα αγοράς καλλυντικά μετά	659	1	5	4,32	1,008
σκέψη αγοράς καλλυντικά μετά	659	1	5	2,57	1,231
δαπάνη αγοράς καλλυντικά μετά	660	1	5	1,19	,645
συχνότητα αγοράς τεχνολογίας πριν	660	1	5	4,49	,919
σκέψη αγοράς τεχνολογίας πριν	660	1	5	2,65	1,195
δαπάνη αγοράς τεχνολογίας πριν	660	1	5	1,56	,939
συχνότητα αγοράς τεχνολογίας μετά	660	1	5	4,74	,673
σκέψη αγοράς τεχνολογίας μετά	660	1	5	2,08	1,195
δαπάνη αγοράς τεχνολογίας μετά	660	1	5	1,24	,590
συχνότητα αγοράς διασκέδαση πριν	658	1	5	2,34	1,057
σκέψη αγοράς διασκέδαση πριν	659	1	5	3,68	1,101
δαπάνη αγοράς διασκέδαση πριν	659	1	5	2,07	1,102
συχνότητα αγοράς διασκέδαση μετά	659	1	5	2,90	1,145
σκέψη αγοράς διασκέδαση μετά	658	1	5	2,73	1,213
δαπάνη αγοράς διασκέδαση μετά	659	1	5	1,59	,830
Valid N (listwise)	655				

# Συσχετίσεις ανά δύο

Correlation Matrix

	εισόδημα πριν	εισόδημα μετά	γνώμη πριν	γνώμη μετά	συχνότητα αγοράς ενδυσης πριν	σκέψη αγοράς ενδυσης πριν	δαπάνη αγοράς ενδυσης πριν	συχνότητα αγοράς ενδυσης μετά	σκέψη αγοράς ενδυσης μετά	δαπάνη αγοράς ενδυσης μετά	συχνότητα αγοράς καλλυντικά πριν	σκέψη αγοράς καλλυντικά πριν	δαπάνη αγοράς καλλυντικά πριν	συχνότητα αγοράς καλλυντικά μετά	σκέψη αγοράς καλλυντικά μετά	δαπάνη αγοράς καλλυντικά μετά	συχνότητα αγοράς τεχνολογίας πριν	σκέψη αγοράς τεχνολογίας πριν	δαπάνη αγοράς τεχνολογίας πριν	συχνότητα αγοράς τεχνολογίας μετά	σκέψη αγοράς τεχνολογίας μετά	δαπάνη αγοράς τεχνολογίας μετά	
Correlation	1,000	,877	-,171	-,132	-,069	,216	,474	-,044	,148	,366	-,119	,221	,269	-,131	,137	,166	-,163	,243					
εισόδημα μετά	,877	1,000	-,168	-,225	-,037	,199	,389	-,040	,178	,373	-,095	,185	,230	-,095	,175	,170	-,100	,207					
γνώμη πριν	-,171	-,168	1,000	,725	,267	-,314	-,234	,212	-,245	-,199	,192	-,300	-,162	,193	-,301	-,100	,155	-,345					
γνώμη μετά	-,132	-,225	,725	1,000	,210	-,140	-,221	,296	-,330	-,284	,111	-,191	-,151	,243	-,378	-,158	,109	-,222					
συχνότητα αγοράς ενδυσης πριν	-,069	-,037	,267	,210	1,000	-,303	-,325	,654	-,230	-,235	,579	-,328	-,254	,457	-,219	-,125	,160	-,174					
σκέψη αγοράς ενδυσης πριν	,216	,199	-,314	-,140	-,303	1,000	,297	-,140	,539	,218	-,282	,624	,173	-,197	,412	,037	-,112	,438					
δαπάνη αγοράς ενδυσης πριν	,474	,389	-,234	-,221	-,325	,297	1,000	-,259	,235	,658	-,253	,271	,483	-,278	,223	,350	-,205	,187					
συχνότητα αγοράς ενδυσης μετά	-,044	-,040	,212	,296	,654	-,140	-,259	1,000	-,329	-,369	,500	-,181	-,287	,577	-,254	-,244	,281	-,120					
σκέψη αγοράς ενδυσης μετά	,148	,178	-,245	-,330	-,230	,539	,235	-,329	1,000	,251	-,237	,417	,147	-,371	,598	,135	-,181	,338					
δαπάνη αγοράς ενδυσης μετά	,366	,373	-,199	-,284	-,235	,218	,658	-,369	,251	1,000	-,192	,160	,480	-,367	,235	,432	-,180	,130					
συχνότητα αγοράς καλλυντικά πριν	-,119	-,095	,192	,111	,579	-,282	-,253	,500	-,237	-,192	1,000	-,277	-,353	,675	-,150	-,171	,205	-,133					
σκέψη αγοράς καλλυντικά πριν	,221	,185	-,300	-,191	-,328	,624	,271	-,181	,417	,160	-,277	1,000	,108	-,207	,626	,076	-,155	,480					
δαπάνη αγοράς καλλυντικά πριν	,269	,230	-,162	-,151	-,254	,173	,483	-,287	,147	,480	-,353	,108	1,000	-,473	,093	,547	-,142	,060					
συχνότητα αγοράς καλλυντικά μετά	-,131	-,095	,193	,243	,457	-,197	-,278	,577	-,371	-,367	,675	-,207	-,473	1,000	-,297	-,310	,193	-,121					
σκέψη αγοράς καλλυντικά μετά	,137	,175	-,301	-,378	-,219	,412	,223	-,254	,598	,235	-,150	,626	,093	-,297	1,000	,113	-,126	,442					
δαπάνη αγοράς καλλυντικά μετά	,166	,170	-,100	-,158	-,125	,037	,350	-,244	,135	,432	-,171	,076	,547	-,310	,113	1,000	-,306	,023					
συχνότητα αγοράς τεχνολογίας πριν	-,163	-,100	,155	,109	,160	-,112	-,205	,281	-,181	-,180	,205	-,155	-,142	,193	-,126	-,306	1,000	-,220					
σκέψη αγοράς τεχνολογίας πριν	,243	,207	-,345	-,222	-,174	,438	,187	-,120	,338	,130	-,133	,480	,060	-,121	,442	,023	-,220	1,000					
δαπάνη αγοράς τεχνολογίας πριν	,270	,207	-,203	-,118	-,144	,185	,389	-,138	,133	,343	-,092	,194	,279	-,138	,134	,292	-,299	,193					
συχνότητα αγοράς τεχνολογίας μετά	-,128	-,083	,133	,191	,141	-,050	-,162	,196	-,157	-,251	,103	-,101	-,216	,215	-,140	-,286	,585	-,135					
σκέψη αγοράς τεχνολογίας μετά	,121	,135	-,335	-,369	-,167	,247	,146	-,223	,476	,191	-,119	,324	,100	-,262	,525	,136	-,208	,686					
δαπάνη αγοράς τεχνολογίας μετά	,223	,209	-,205	-,226	-,182	,147	,412	-,224	,235	,418	-,092	,177	,233	-,211	,256	,235	-,252	,154					
συχνότητα αγοράς διασκέδαση πριν	-,035	,017	,258	,199	,261	-,324	-,177	,118	-,204	-,068	,203	-,325	-,015	,095	-,198	,046	,067	-,204					
σκέψη αγοράς διασκέδαση πριν	,206	,138	-,292	-,143	-,236	,533	,262	-,090	,265	,175	-,203	,494	,099	-,101	,308	,021	-,072	,352					
δαπάνη αγοράς διασκέδαση πριν	,473	,363	-,214	-,115	-,237	,308	,495	-,154	,189	,385	-,269	,314	,293	-,182	,177	,260	-,219	,278					
συχνότητα αγοράς διασκέδαση μετά	,061	,034	,293	,350	,264	-,254	-,116	,256	-,296	-,170	,183	-,247	-,037	,213	-,316	-,075	,070	-,146					
σκέψη αγοράς διασκέδαση μετά	,081	,104	-,300	-,340	-,190	,395	,153	-,231	,481	,233	-,144	,366	,108	-,217	,449	,162	-,117	,282					
δαπάνη αγοράς διασκέδαση μετά	,398	,362	-,266	-,271	-,223	,267	,447	-,254	,351	,472	-,245	,273	,316	-,301	,300	,296	-,190	,259					

# Μέρος πίνακα συσχετίσεων

Correlation Matrix

	εισόδημα πριν	εισόδημα μετά	γνώμη πριν	γνώμη μετά	συχνότητα αγοράς πριν	σκέψη αγοράς πριν	δαπάνη αγοράς πριν	συχνότητα αγοράς μετά	σκέψη αγοράς μετά	δαπάνη αγοράς μετά	συχνότητα αγοράς καλλυντικά πριν	σκέψη αγοράς καλλυντικά πριν	δαπάνη αγοράς καλλυντικά πριν	συχνότητα αγοράς καλλυντικά μετά	σκέψη αγοράς καλλυντικά μετά	δαπάνη αγοράς καλλυντικά μετά
εισόδημα πριν	1,000	,878	-,174	-,138	-,071	,220	,477	-,051	,155	,367	-,117	,219	,267	-,129	,138	,164
εισόδημα μετά	,878	1,000	-,171	-,232	-,038	,203	,393	-,048	,186	,375	-,094	,183	,228	-,093	,175	,169
γνώμη πριν	-,174	-,171	1,000	,724	,267	-,317	-,237	,216	-,249	-,203	,196	-,291	-,160	,189	-,296	-,099
γνώμη μετά	-,138	-,232	,724	1,000	,211	-,145	-,226	,302	-,337	-,289	,111	-,187	-,149	,240	-,375	-,157
συχνότητα αγοράς πριν	-,071	-,038	,267	,211	1,000	-,303	-,325	,651	-,230	-,232	,574	-,327	-,254	,456	-,220	-,125
σκέψη αγοράς πριν	,220	,203	-,317	-,145	-,303	1,000	,300	-,146	,542	,223	-,283	,616	,172	-,194	,408	,036
δαπάνη αγοράς πριν	,477	,393	-,237	-,226	-,325	,300	1,000	-,264	,240	,660	-,253	,266	,481	-,276	,220	,348
συχνότητα αγοράς μετά	-,051	-,048	,216	,302	,651	-,146	-,264	1,000	-,336	-,373	,497	-,176	-,284	,571	-,252	-,242
σκέψη αγοράς μετά	,155	,186	-,249	-,337	-,230	,542	,240	-,336	1,000	,257	-,238	,408	,144	-,365	,591	,133
δαπάνη αγοράς μετά	,367	,375	-,203	-,289	-,232	,223	,660	-,373	,257	1,000	-,196	,152	,477	-,363	,228	,429
συχνότητα αγοράς καλλυντικά πριν	-,117	-,094	,196	,111	,574	-,283	-,253	,497	-,238	-,196	1,000	-,263	-,350	,669	-,141	-,169
σκέψη αγοράς καλλυντικά πριν	,219	,183	-,291	-,187	-,327	,616	,266	-,176	,408	,152	-,263	1,000	,110	-,209	,629	,077
δαπάνη αγοράς καλλυντικά πριν	,267	,228	-,160	-,149	-,254	,172	,481	-,284	,144	,477	-,350	,110	1,000	-,473	,093	,547
συχνότητα αγοράς καλλυντικά μετά	-,129	-,093	,189	,240	,456	-,194	-,276	,571	-,365	-,363	,669	-,209	-,473	1,000	-,298	-,311
σκέψη αγοράς καλλυντικά μετά	,138	,175	-,296	-,375	-,220	,408	,220	-,252	,591	,228	-,141	,629	,093	-,298	1,000	,113
δαπάνη αγοράς καλλυντικά μετά	,164	,169	-,099	-,157	-,125	,036	,348	-,242	,133	,429	-,169	,077	,547	-,311	,113	1,000

# Απλές συσχετίσεις ανά δύο των μεταβλητών

- Στις παραπάνω περιπτώσεις ο αριθμός των απλών συσχετίσεων μεταξύ μεταβλητών είναι πολύ μεγάλος και δύσκολο να εξαγάγουμε συμπεράσματα. Αν οι μεταβλητές μας είναι 30, υπάρχουν 435 τιμές των ανά δύο συσχετίσεων!!!
- Προσπαθούμε να βρούμε τρόπους μείωσης του αριθμού των μεταβλητών ομαδοποιώντας κάποιες από αυτές, μετά την συλλογή των δεδομένων. Η μέθοδος που χρησιμοποιούμε εξαρτάται από τον τρόπο μείωσης διαστάσεων.

# Factor analysis

Εντοπίζουμε γραμμικούς συνδυασμούς των αρχικών μεταβλητών τους οποίους ονομάζουμε παράγοντες (factors)

- Οι νέες μεταβλητές που προκύπτουν από τους γραμμικούς συνδυασμούς ονομάζονται κύριες συνιστώσες principal components
- Συνήθως οι κύριες συνιστώσες που προκύπτουν είναι 2 έως 6 και ερμηνεύουν μεγάλο μέρος της διασποράς των δεδομένων (περίπου 75%).
- Οι κύριες συνιστώσες ερμηνεύονται – περιγράφονται χρησιμοποιώντας τους συντελεστές των αρχικών μεταβλητών στο γραμμικό συνδυασμό των συνιστωσών. Οι συντελεστές αυτοί ονομάζονται βαρύτητες στους παράγοντες (factor loadings).

# Ανάλυση με το SPSS

- Analyze, Dimension reduction, Factor...

Επιλέγουμε τις μεταβλητές που θα χρησιμοποιήσουμε στην ανάλυση.

Επιπρόσθετα

Πλήκτρο **Descriptives** επιλέγουμε **KMO**

Πλήκτρο **Rotation** επιλέγουμε **Loading plots**

Πλήκτρο **Extraction** επιλέγουμε **Scree plot**

Πλήκτρο **Scores** επιλέγουμε **Save as variable** για αποθήκευση των νέων συντεταγμένων

Και OK.

# Προϋπόθεση μη συσχέτισης

Για την εύρεση των κυρίων συνιστωσών (παραγόντων) προσπαθούμε να βρούμε γραμμικούς συνδυασμούς συσχετισμένων μεταβλητών και κατόπιν ασυσχέτιστων κυρίων συνιστωσών.

Αν όλες οι αρχικές μεταβλητές είναι ασυσχέτιστες δεν υπάρχει λόγος να προχωρήσουμε τη διαδικασία.

**KMO and Bartlett's Test**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,795
Bartlett's Test of Sphericity	Approx. Chi-Square	10340,798
	df	378
	Sig.	,000

st.

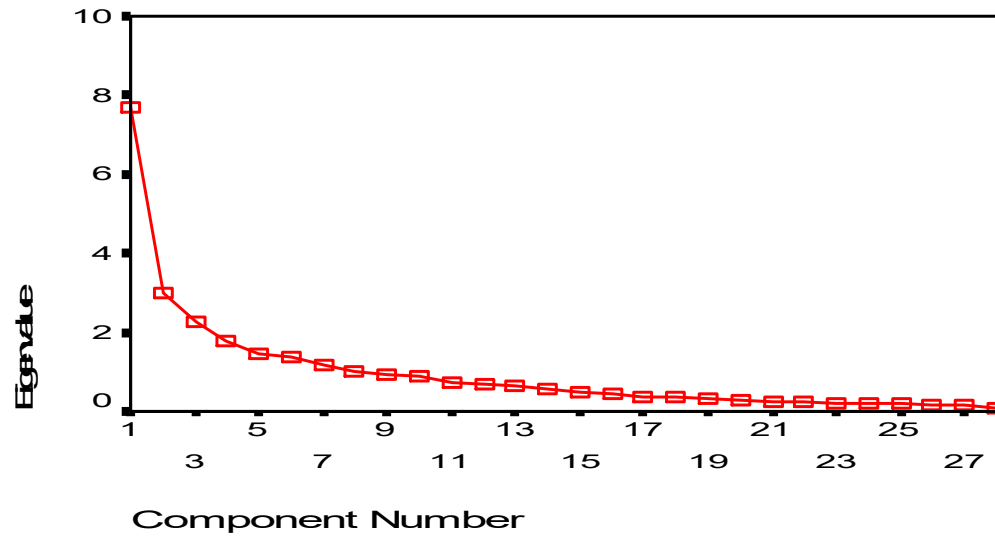
$H_0$  ότι όλες οι μεταβλητές είναι μεταξύ τους ασυσχέτιστες  
Αν η  $H_0$  απορρίπτεται (sig < 0.05) συνεχίζουμε την ανάλυση.

# Πόσους παράγοντες;

- Η εύρεση κυρίων παραγόντων (συνιστωσών) είναι μια επαναλαμβανόμενη διαδικασία και τελειώνει όταν το πλήθος των παραγόντων είναι ίσο με το πλήθος των αρχικών μεταβλητών.
- Ο σκοπός μας όμως είναι η μείωση του αριθμού των μεταβλητών.
  - Διάφορα κριτήρια προτάθηκαν. Το πιο διαδεδομένο είναι αυτό που χρησιμοποιεί τη γραφική παράσταση scree test. Παρουσιάζεται η ερμηνευμένη διασπορά ως προς τον αύξοντα αριθμό του παράγοντα. Σταματάμε εκεί που υπάρχει «απότομη» μεταβολή.



## Scree Plot



Επιλέγουμε να ερμηνεύσουμε τους 4 κύριους παράγοντες που ερμηνεύουν το 52,5% της συνολικής μεταβλητότητας των δεδομένων.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,680	27,430	27,430	7,680	27,430	27,430
2	2,991	10,683	38,114	2,991	10,683	38,114
3	2,277	8,130	46,244	2,277	8,130	46,244
4	1,766	6,306	52,550	1,766	6,306	52,550
5	1,466	5,235	57,786	1,466	5,235	57,786
6	1,390	4,964	62,750	1,390	4,964	62,750
7	1,173	4,190	66,940	1,173	4,190	66,940
8	,997	3,561	70,501			
9	,922	3,294	73,795			
10	,882	3,149	76,944			
11	,736	2,627	79,572			
12	,696	2,485	82,057			
13	,662	2,366	84,422			
14	,572	2,043	86,465			
15	,504	1,800	88,265			
16	,450	1,605	89,870			
17	,372	1,327	91,197			
18	,361	1,291	92,488			
19	,328	1,171	93,660			
20	,289	1,033	94,693			
21	,229	,817	95,510			
22	,227	,810	96,320			
23	,206	,735	97,055			
24	,203	,725	97,780			
--						

Θα μπορούσαμε να ερμηνεύσουμε και τους 5, 6, ή 7 πρώτους αν θέλουμε.

Γενικά παράγοντες με eigenvalue <1 δεν ερμηνεύονται.

# Γενική παρουσίαση

Component Matrix <sup>a</sup>

	Component			
	1	2	3	4
εισόδημα πριν	,463	,406	,531	,196
εισόδημα μετά	,428	,363	,518	,128
γνώμη πριν	-,535	,217	-3,26E-02	,180
γνώμη μετά	-,511	,165	5,487E-02	,333
συχνότητα αγοράς ενδυσης πριν	-,530	5,850E-02	,496	-,280
σκέψη αγοράς ενδυσης πριν	,592	-,344	,192	,238
δαπάνη αγοράς ενδυσης πριν	,633	,393	,125	,218
συχνότητα αγοράς ενδυσης μετά	-,527	-5,96E-02	,609	-1,72E-02
σκέψη αγοράς ενδυσης μετά	,616	-,297	-1,60E-02	-,172
δαπάνη αγοράς ενδυσης μετά	,613	,421	4,940E-06	3,251E-02
συχνότητα αγοράς καλλυντικά πριν	-,493	-3,97E-02	,517	-,361
σκέψη αγοράς καλλυντικά πριν	,609	-,369	,181	,116
δαπάνη αγοράς καλλυντικά πριν	,482	,467	-,218	,135
συχνότητα αγοράς καλλυντικά μετά	-,560	-,127	,561	-7,92E-02
σκέψη αγοράς καλλυντικά μετά	,610	-,376	8,779E-02	-,242
δαπάνη αγοράς καλλυντικά μετά	,408	,456	-,181	-,150
συχνότητα αγοράς τεχνολογίας πριν	-,388	-,219	,131	,391
σκέψη αγοράς τεχνολογίας πριν	,528	-,311	,294	-,233
δαπάνη αγοράς τεχνολογίας πριν	,461	,350	,156	-,138
συχνότητα αγοράς τεχνολογίας μετά	-,356	-,285	,138	,499
σκέψη αγοράς τεχνολογίας μετά	,536	-,292	6,901E-02	-,481
δαπάνη αγοράς τεχνολογίας μετά	,498	,281	7,488E-02	-,233
συχνότητα αγοράς διασκέδαση πριν	-,391	,440	5,338E-03	-,359
σκέψη αγοράς διασκέδαση πριν	,513	-,389	,232	,341
δαπάνη αγοράς διασκέδαση πριν	,584	,271	,276	,286
συχνότητα αγοράς διασκέδαση μετά	-,438	,450	,181	-9,38E-02
σκέψη αγοράς διασκέδαση μετά	,545	-,407	3,329E-02	-7,96E-02
δαπάνη αγοράς διασκέδαση μετά	,650	,236	,169	4,845E-02

Extraction Method: Principal Component Analysis.

a. 7 components extracted.

# Ευμηνεία παραγόντων

Component Matrix <sup>a</sup>

	Component			
	1	2	3	4
εισόδημα πριν	,463	,406	,531	,196
εισόδημα μετά	,428	,363	,518	,128
γνώμη πριν	-,535	,217	-3,26E-02	,180
γνώμη μετά	-,511	,165	5,487E-02	,333
συχνότητα αγοράς ενδυσης πριν	-,530	5,850E-02	,496	-,280
σκέψη αγοράς ενδυσης πριν	,592	-,344	,102	,238
δαπάνη αγοράς ενδυσης πριν	,633	,393	,125	,218
συχνότητα αγοράς ενδυσης μετά	-,527	-5,96E-02	,609	-1,72E-02
σκέψη αγοράς ενδυσης μετά	,616	-,297	-1,60E-02	-,172
δαπάνη αγοράς ενδυσης μετά	,613	,421	4,940E-06	3,251E-02
συχνότητα αγοράς καλλυντικά πριν	-,493	-3,97E-02	,517	-,361
σκέψη αγοράς καλλυντικά πριν	,609	-,369	,181	,116
δαπάνη αγοράς καλλυντικά πριν	,402	,467	-,218	,135
συχνότητα αγοράς καλλυντικά μετά	-,560	-,127	,561	-7,92E-02
σκέψη αγοράς καλλυντικά μετά	,610	-,376	6,775E-02	-,242
δαπάνη αγοράς καλλυντικά μετά	,408	,456	-,181	-,150
συχνότητα αγοράς τεχνολογίας πριν	-,388	-,219	,131	,391
σκέψη αγοράς τεχνολογίας πριν	,528	-,311	,294	-,233
δαπάνη αγοράς τεχνολογίας πριν	,461	,350	,156	-,138
συχνότητα αγοράς τεχνολογίας μετά	-,356	-,285	,138	,499
σκέψη αγοράς τεχνολογίας μετά	,536	-,292	6,901E-02	-,481
δαπάνη αγοράς τεχνολογίας μετά	,498	,281	7,488E-02	-,233
συχνότητα αγοράς διασκέδαση πριν	-,391	,440	5,338E-03	-,359
σκέψη αγοράς διασκέδαση πριν	,513	-,309	,232	,341
δαπάνη αγοράς διασκέδαση πριν	,584	,271	,276	,286
συχνότητα αγοράς διασκέδαση μετά	-,438	,450	,181	-9,38E-02
σκέψη αγοράς διασκέδαση μετά	,545	-,407	3,329E-02	-7,96E-02
δαπάνη αγοράς διασκέδαση μετά	,650	,236	,169	4,845E-02

Extraction Method: Principal Component Analysis.

a. 7 components extracted.

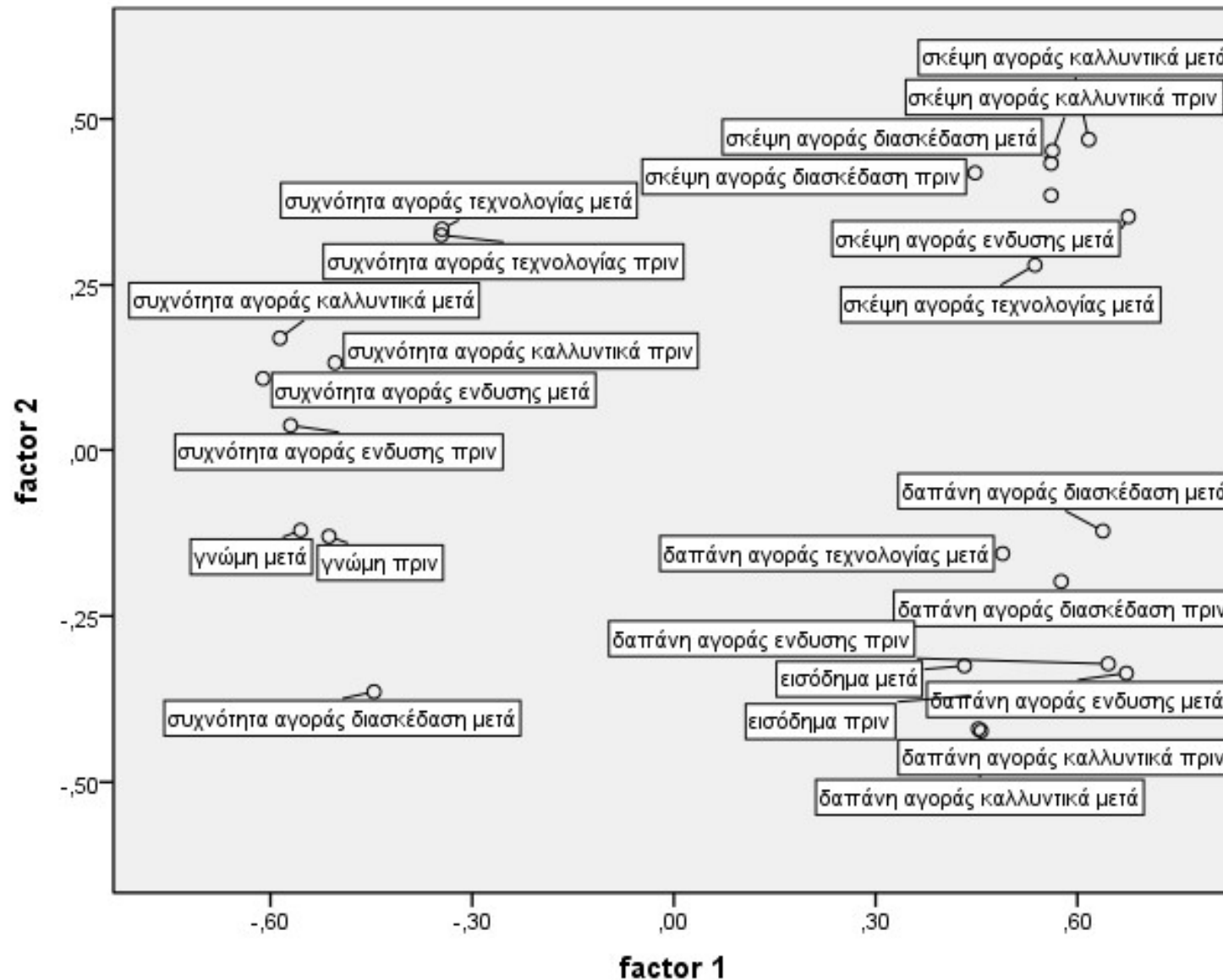
# ερμηνεία παραγόντων

- 1ος παράγοντας σχετίζεται θετικά με τις μεταβλητές
  - Δαπάνη αγοράς ένδυσης πριν και μετά
  - Δαπάνη διασκέδασης μετά
  - Σκέψη αγοράς ένδυσης μετά
  - Σκέψη αγοράς καλλυντικών πριν και μετά
- 2ος παράγοντας σχετίζεται θετικά με τις μεταβλητές
  - Εισόδημα πριν
  - Δαπάνη αγοράς καλλυντικών πριν και μετά
  - συχνότητα διασκέδασης πριν και μετά
  - Αρνητικά με τη Σκέψη διασκέδασης μετά
- 3ος παράγοντας σχετίζεται θετικά με τις μεταβλητές
  - Εισόδημα πριν και μετά
  - συχνότητα αγοράς ένδυσης πριν και μετά
  - Σκέψη αγοράς καλλυντικών μετά
- 4ος παράγοντας σχετίζεται θετικά με τις μεταβλητές
  - συχνότητα αγοράς τεχνολογίας μετά
  - Αρνητικά με τη Σκέψη αγοράς τεχνολογίας μετά

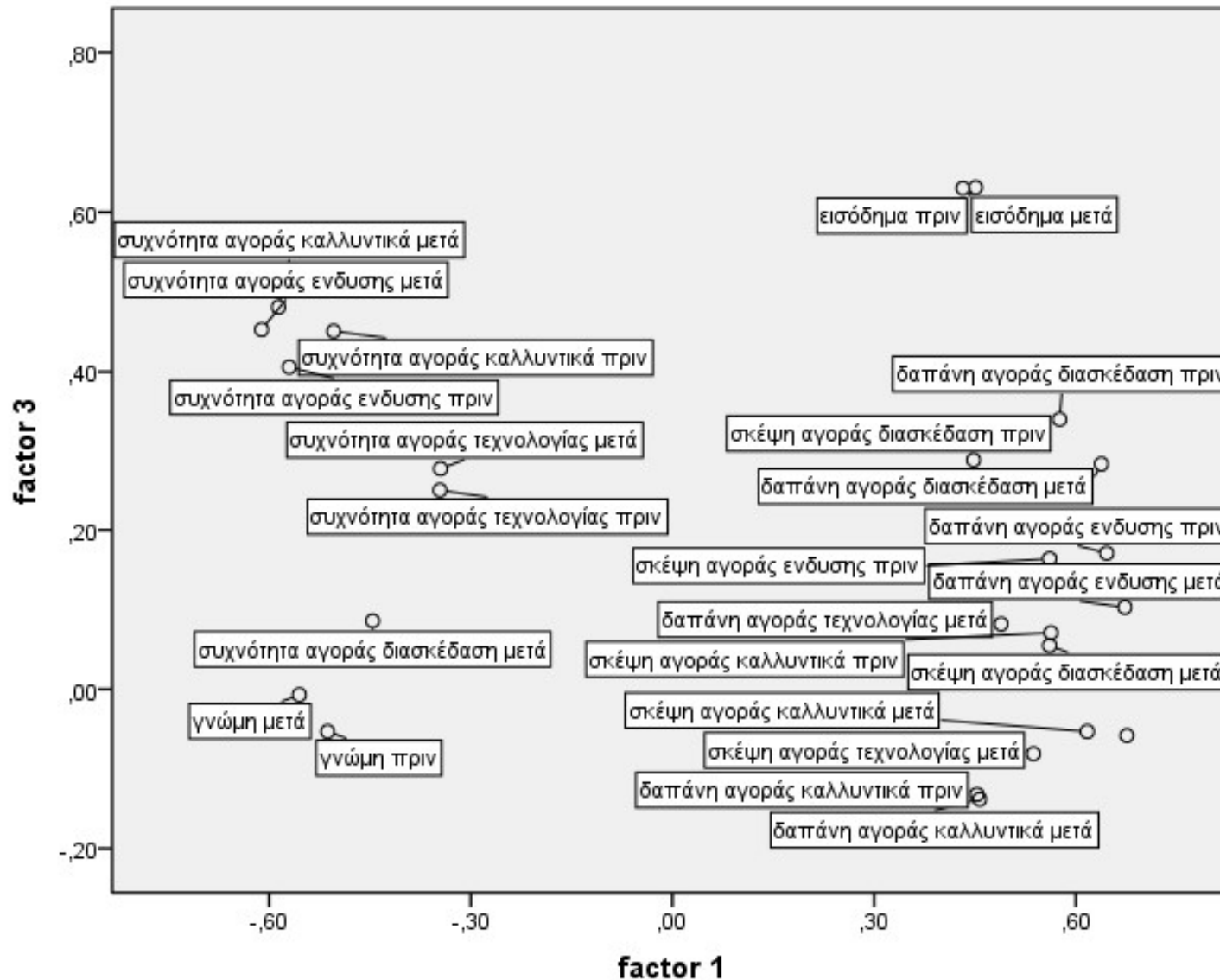
# ονομασία παραγόντων

- Ανάλογα με τις μεταβλητές που έχουν τη μεγαλύτερη βαρύτητα σε κάποιο παράγοντα, ονομάζουμε τον παράγοντα αυτόν με γενικό όνομα που περιλαμβάνει όσο μπορούμε καλύτερα τις μεταβλητές με τη μεγάλη βαρύτητα.
- **1ος παράγοντας**
  - Δαπάνη αγοράς ένδυσης πριν και μετά
  - Δαπάνη διασκέδασης μετά
  - Σκέψη αγοράς ένδυσης μετά
  - Σκέψη αγοράς καλλυντικών πριν και μετά
  - Μεγάλες δαπάνες πριν και μετά χωρίς πολύ σκέψη αγοράς
- **2ος παράγοντας**
  - Εισόδημα πριν
  - Δαπάνη αγοράς καλλυντικών πριν και μετά
  - συχνότητα διασκέδασης πριν και μετά
  - Αρνητικά με τη Σκέψη διασκέδασης μετά
  - Μεγάλες δαπάνες πριν και μετά χωρίς αλλαγή συχνότητας διασκέδασης αλλά και χωρίς σκέψη διασκέδασης μετά
- **3ος παράγοντας**
  - Εισόδημα πριν και μετά
  - συχνότητα αγοράς ένδυσης πριν και μετά
  - Σκέψη αγοράς καλλυντικών μετά
  - Μέτρια προς υψηλά εισοδήματα με περισσότερη σκέψη αγοράς καλλυντικών μετά
- **4ος παράγοντας**
  - συχνότητα αγοράς τεχνολογίας μετά
  - Αρνητικά με τη Σκέψη αγοράς τεχνολογίας μετά
  - Μέτρια προς χαμηλά εισοδήματα με περισσότερη σκέψη αγοράς τεχνολογίας μετά

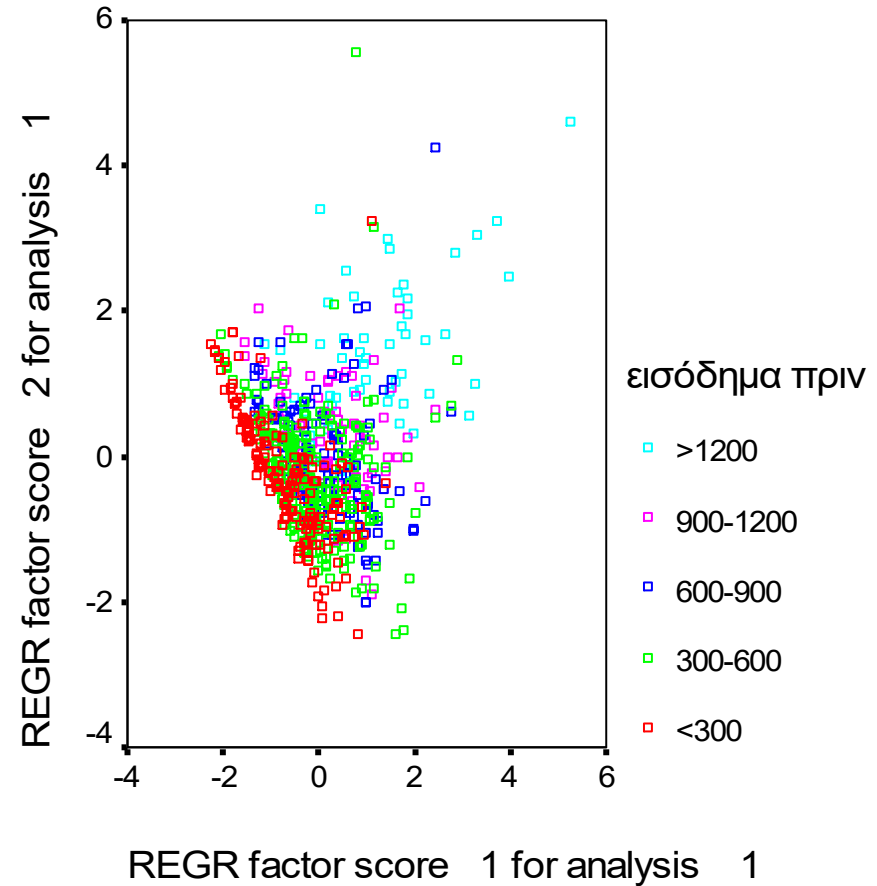
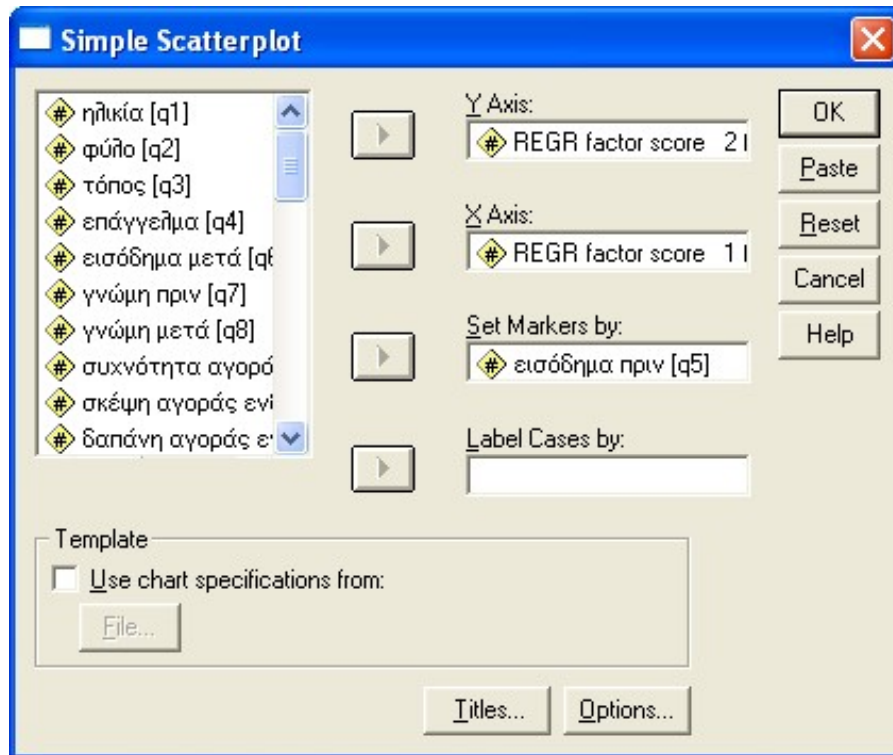
# Οπτικοποίηση συντεταγμένων (νέο αρχείο με scatterplot)



# Οπτικοποίηση συντεταγμένων (νέο αρχείο με scatterplot

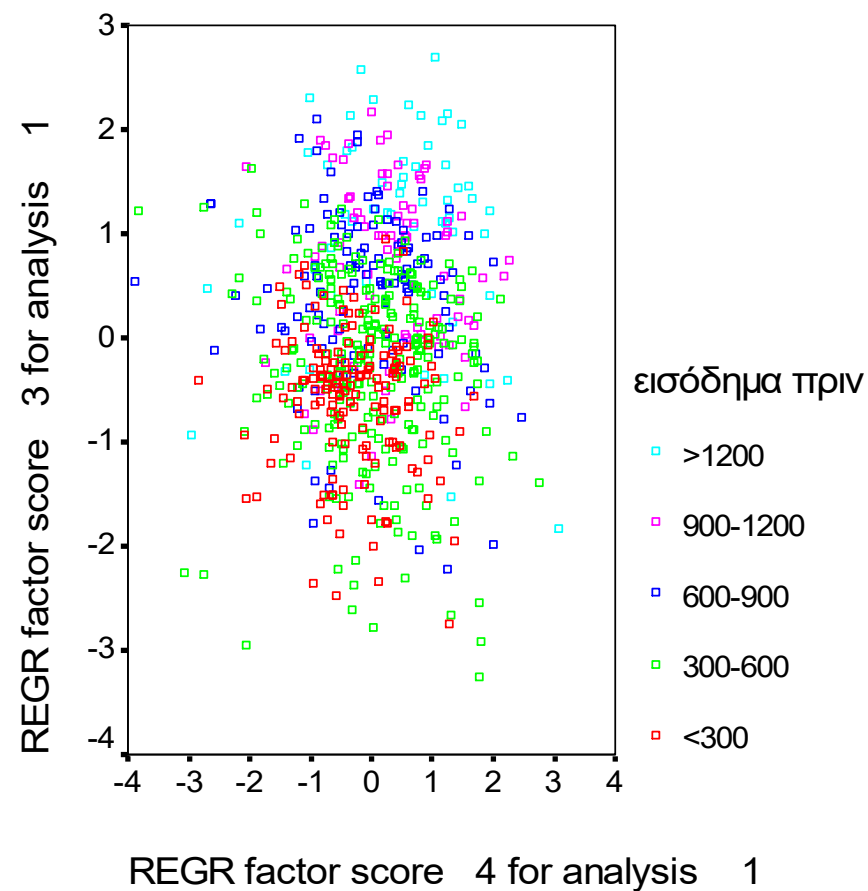
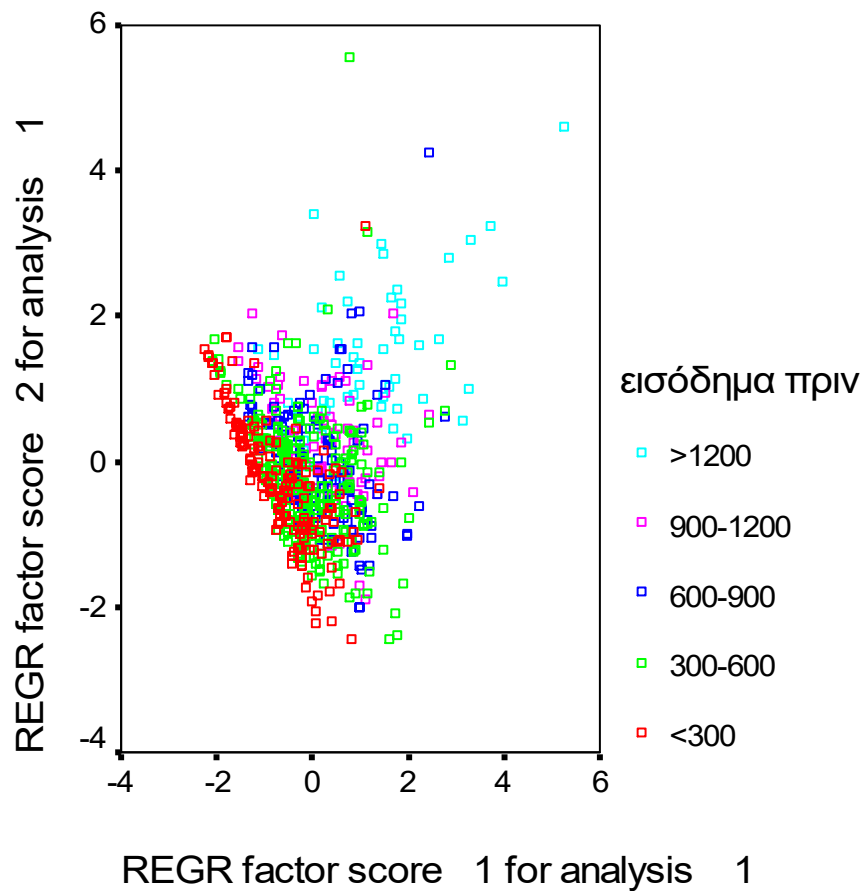


# «Οπτικοποίηση» αποτελεσμάτων



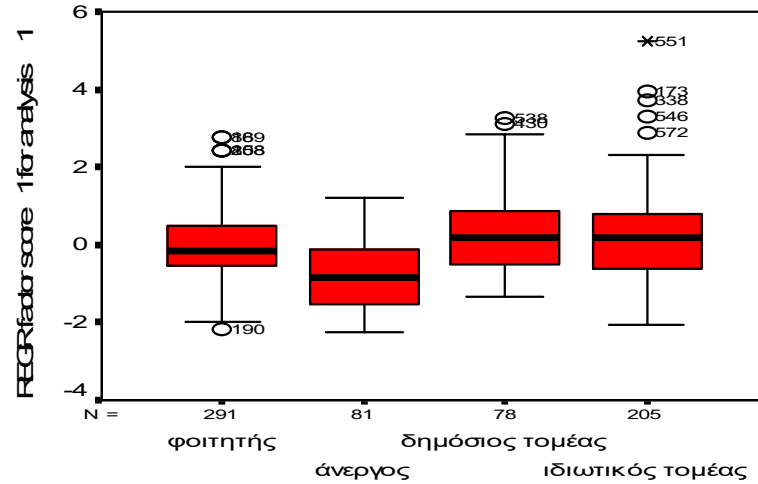


# 3ος και 4ος παράγοντας με εισόδημα

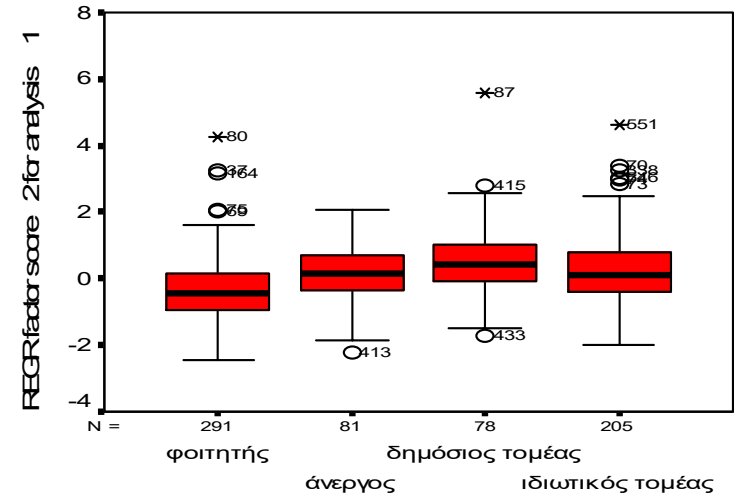




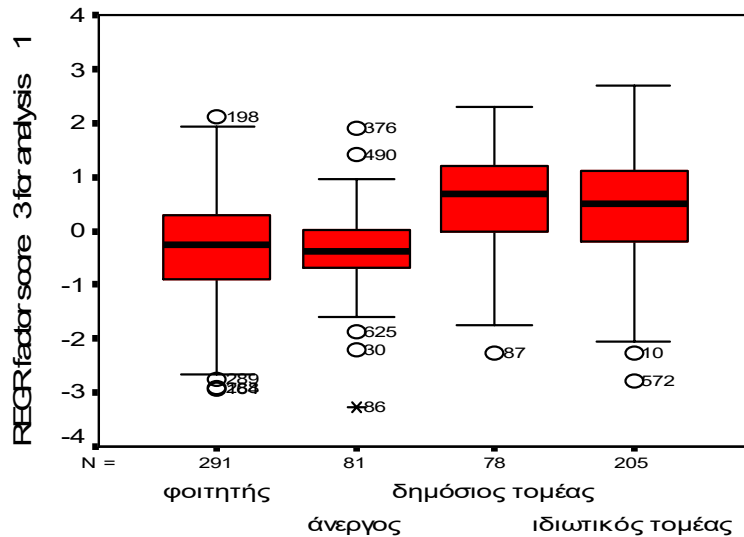
# παράγοντες με επάγγελμα (box-plot)



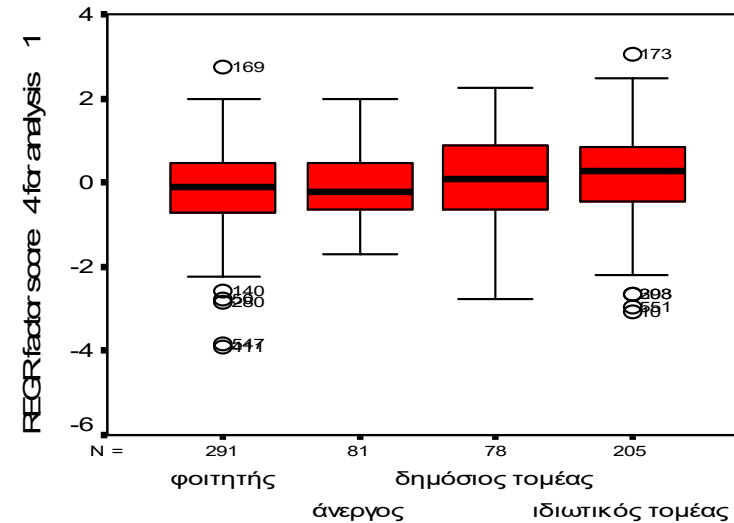
επάγγελμα



επάγγελμα



επάγγελμα



επάγγελμα