

# Ανάλυση σε ομάδες με τη χρήση του στατιστικού πακέτου SPSS (Cluster Analysis using SPSS)

---

## Περιεχόμενα:

1. Μη-ιεραρχικές μέθοδοι ομαδοποίησης
2. Στάδια ομαδοποίησης
3. Παράδειγμα
  - 3.1. Εισαγωγή δεδομένων
  - 3.2. Ομαδοποίηση
4. Μειονεκτήματα – Πλεονεκτήματα μη-ιεραρχικών διαδικασιών ομαδοποίησης

1. Μη-ιεραρχικές μέθοδοι ομαδοποίησης
- 

**Άσκηση 1: Κάτω από ποιες προϋποθέσεις χρησιμοποιούμε τις μη-ιεραρχικές μεθόδους ομαδοποίησης.**

2. Στάδια ομαδοποίησης
- 

Τα βασικά στάδια ομαδοποίησης που εφαρμόζουμε στις μη-ιεραρχικές μεθόδους είναι:

I. Επιλέγουμε ένα αρχικό αριθμό ομάδων. Πώς επιλέγουμε τον αρχικό αριθμό ομάδων;

.....

.....

II. Κατατάσσουμε κάθε παρατήρηση στην κοντινότερη ομάδα. Με ποια μέθοδο;

.....

III. Επανατοποθετούμε κάθε παρατήρηση σε μια από τις ομάδες βάση κανόνα. Πότε σταματά η διαδικασία επανατοποθέτησης;

.....

### 3. Παράδειγμα

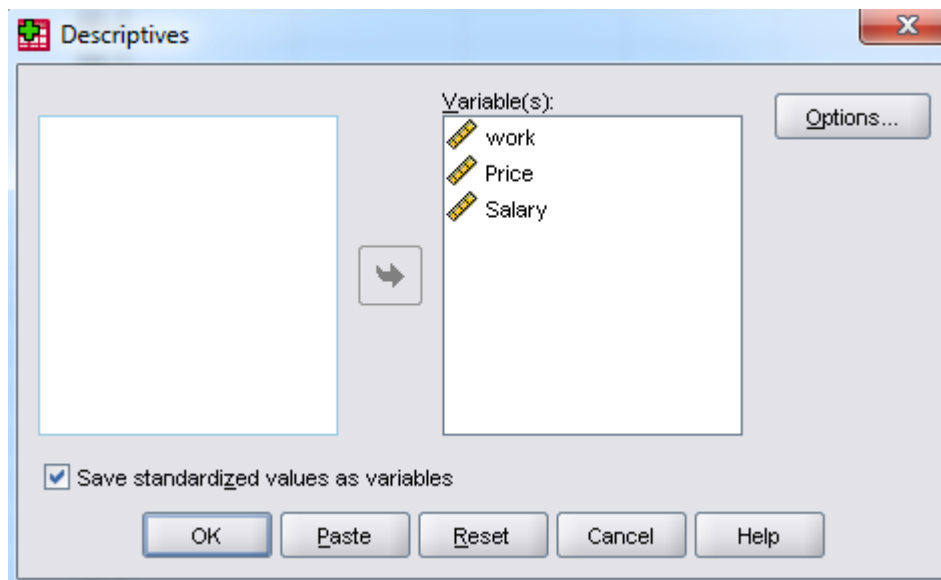
---

#### 3.1. Εισαγωγή δεδομένων

**Χρησιμοποιείτε τα πιο κάτω δεδομένα στο SPSS.**

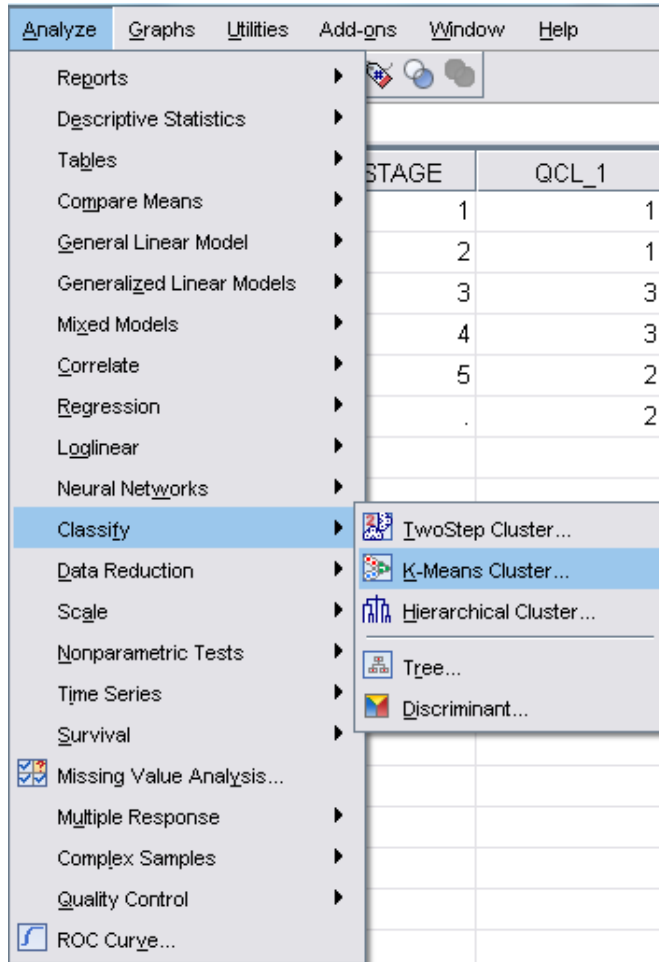
Income	5	6	15	16	25	30
Education	5	6	14	15	20	19
NAME	S1	S2	S3	S4	S5	S6

Σημείωση: Σε περίπτωση τυποποίησης των δεδομένων χρησιμοποιήστε την επιλογή *Analyze* → *Descriptive Statistics* → *Descriptives* → *Save standardized values as variables* → *OK*.

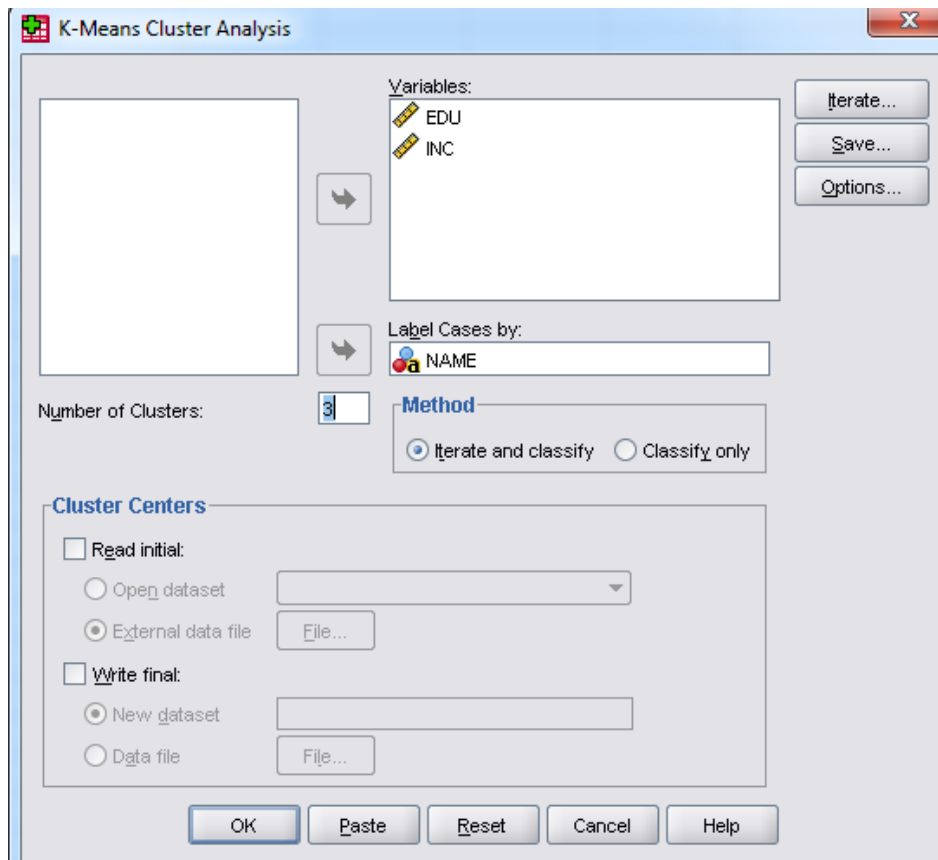


### 3.2. Ομαδοποίηση

Για τη διαδικασία ομαδοποίησης επιλέγουμε *Analyze* → *Classify* → *K-means Cluster Analysis*.



- ✚ Στο παράθυρο που εμφανίζεται επιλέξτε όλες τις μεταβλητές που θέλουμε να διερευνήσουμε. Επιλέξτε το “NAME” σαν **“Label”**.
- ✚ Επιλέγουμε τον αριθμό των ομάδων, **“Number of Clusters”**. Βασικός κανόνας είναι ότι πρέπει να υπάρχουν τουλάχιστον 2 ομάδες και ο αριθμός τους δεν πρέπει να υπερβαίνει τον αριθμό των παρατηρήσεων.



- ✚ Στην επιλογή *Iterate* μπορείτε να επιλέξετε τον μέγιστο αριθμό επαναλήψεων της διαδικασίας ομαδοποίησης καθώς και το κριτήριο επανάληψης. Επιλέξτε *Use running means* μόνο αν θέλετε να επαναυπολογίζετε το κέντρο κάθε φορά που μια παρατήρηση ομαδοποιείται.
- ✚ Στην επιλογή *Save* μπορείτε να επιλέξετε να αποθηκεύσετε με τη μορφή νέας μεταβλητής (α) την τελική ομάδα κάθε μεταβλητής, π.χ. 1-3, (β) την ευκλείδεια. Επιλέξτε και τα δύο.
- ✚ Επιπλέον επιλέξτε *Option* → *Statistics* → *Initial cluster centers/Cluster information for each case*.

**Άσκηση 2: Περιγράψτε τα αποτελέσματα. Δοκιμάστε διαφορετικούς αριθμούς ομάδων. Τι παρατηρείτε;**

4. Μειονεκτήματα - Πλεονεκτήματα μη-ιεραρχικών διαδικασιών ομαδοποίησης

---

**Άσκηση 3: Πώς μπορεί να βελτιωθεί η απόδοση των μη-ιεραρχικών μεθόδων;**

**ΣΥΜΒΟΥΛΗ: ΣΤΗΝ ΤΕΛΙΚΗ ΣΑΣ ΕΡΓΑΣΙΑ ΧΡΗΣΙΜΟΠΟΙΕΙΣΤΕ ΙΕΡΑΡΧΙΚΕΣ ΚΑΙ ΜΗ-ΙΕΡΑΡΧΙΚΕΣ ΜΕΘΟΔΟΥΣ ΟΜΑΔΟΠΟΙΗΣΗΣ.**