
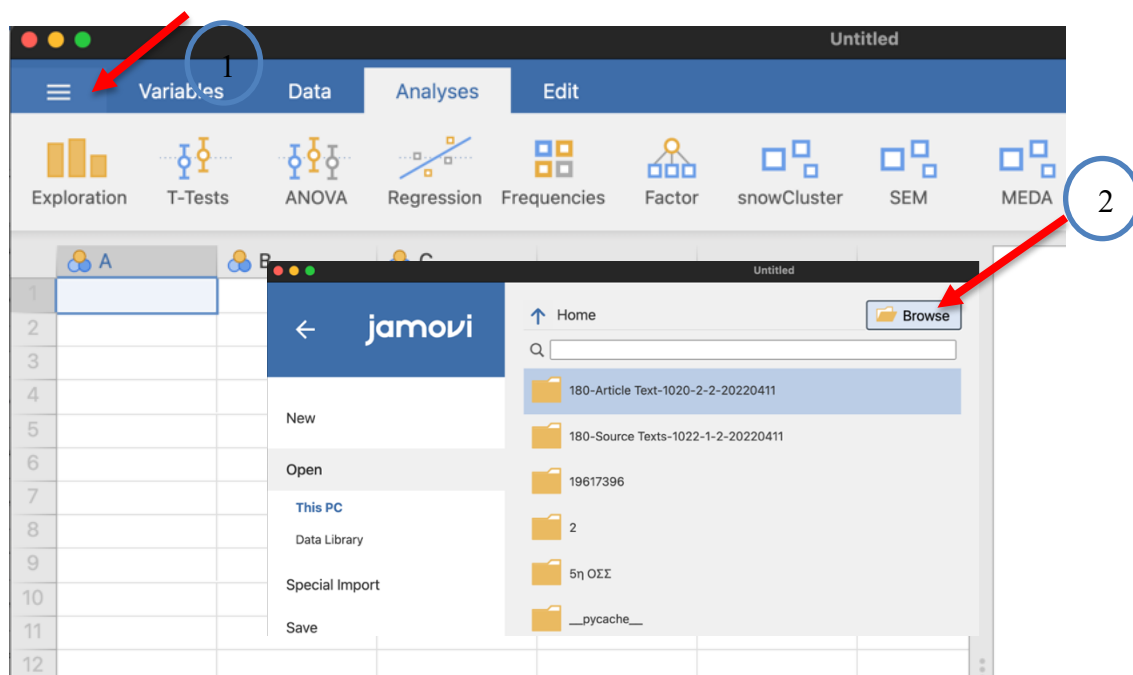



# Παραγοντική Ανάλυση Αντιστοιχιών και Ανιούσα Ιεραρχική Ταξινόμηση με το JAMOVI

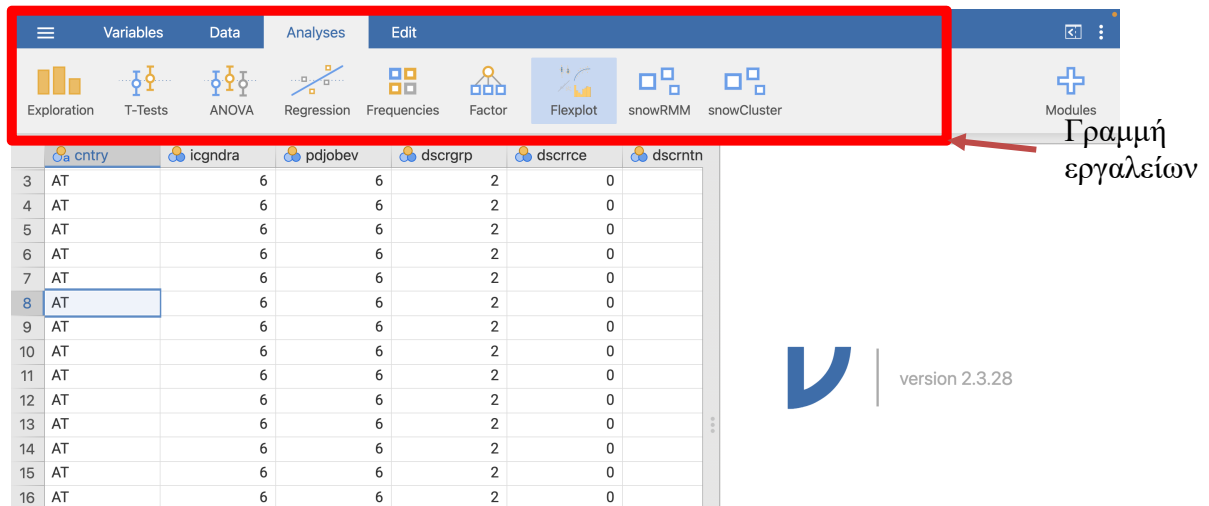
## 1. Εισαγωγή Πίνακα Δεδομένων στο JAMOVI

Στο πρώτο παράθυρο που εμφανίζεται με την εκτέλεση του προγράμματος επιλέγουμε  και **Open (This PC)** → **Browse** (Εικόνα 1). Εντοπίζουμε το αρχείο δεδομένων μας (π.χ. αρχείο .csv) και επιλέγουμε **Open**. *Προσοχή:* η πρώτη γραμμή του αρχείου πρέπει να περιέχει τα ονόματα των μεταβλητών και η πρώτη στήλη τα ονόματα των αντικειμένων/υποκειμένων.



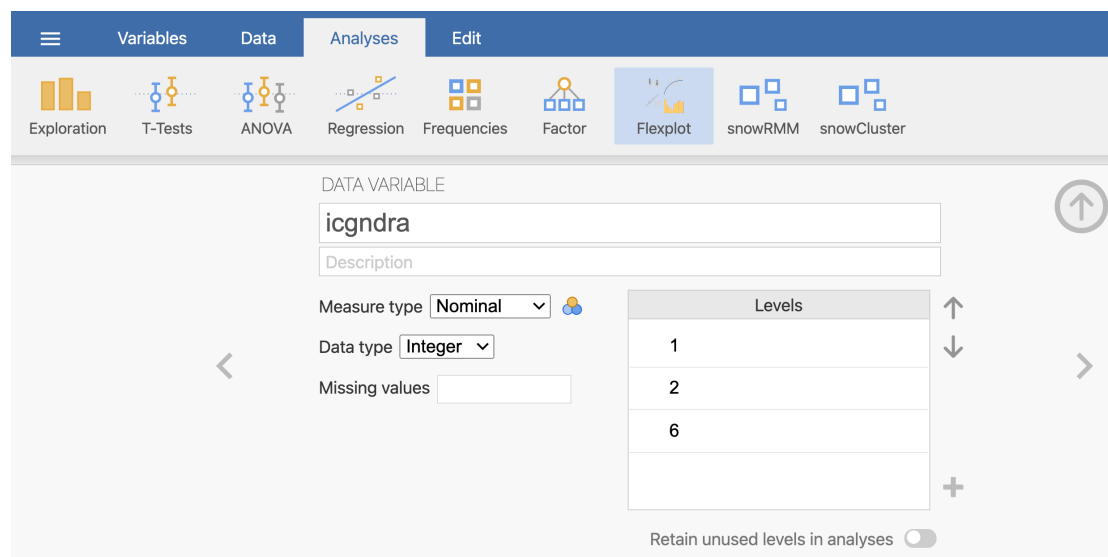
Εικόνα 1. Εισαγωγή αρχείου δεδομένων στο JAMOVI

Τα δεδομένα έχουν εισαχθεί στον πίνακα που βρίσκεται κάτω από την γραμμή εργαλείων (βλ. Εικόνα 2). Στο παράδειγμά μας, πρόκειται για 21,187 υποκείμενα (άτομα) και 16 μεταβλητές (ερωτήσεις σχετικά με τις διακρίσεις που πιθανόν υφίστανται τα άτομα). Μπορούμε να αποθηκεύσουμε τα δεδομένα σε αρχείο του JAMOVI (με επέκταση .omv) με επιλογή του  και **Save/Save as...**



Εικόνα 2. Τα δεδομένα μας στο JAMOVI και η γραμμή εργαλείων

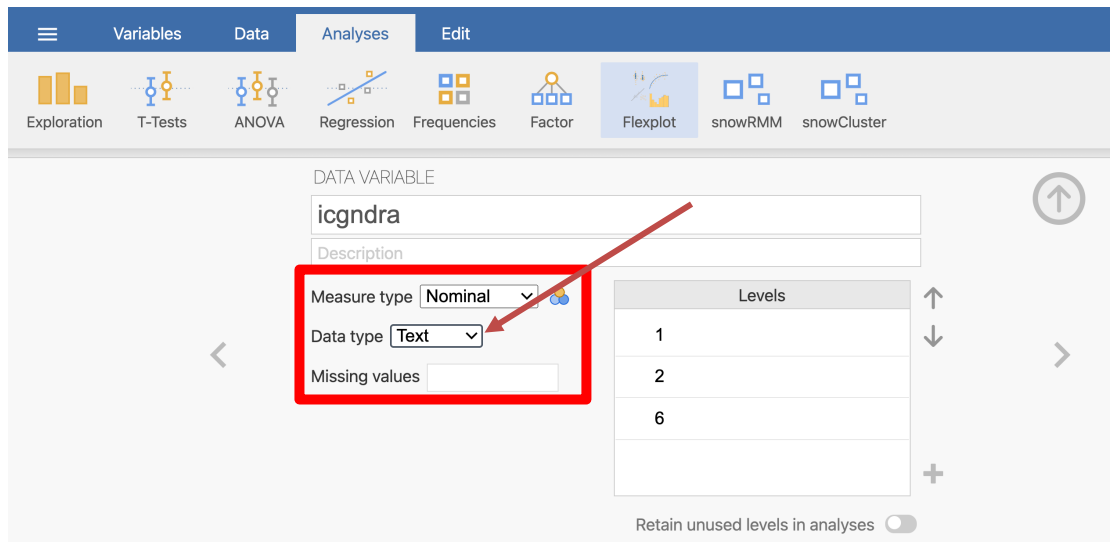
Ελέγχουμε ότι τα δεδομένα έχουν εισαχθεί σωστά στην καρτέλα **Data**. Παρατηρήστε ότι το JAMOVI έχει αναγνωρίσει αυτόματα την κλίμακα μέτρησης των μεταβλητών ως Nominal (κατηγορική μεταβλητή σε ονομαστική κλίμακα). Με διπλό κλικ στο όνομα μιας μεταβλητής μπορούμε να δούμε περισσότερες πληροφορίες γι' αυτήν, όπως την κλίμακα μέτρησής της (Measure type) (Εικόνα 3).



Εικόνα 3. Κλίμακα μέτρησης μιας μεταβλητής (Measure type)

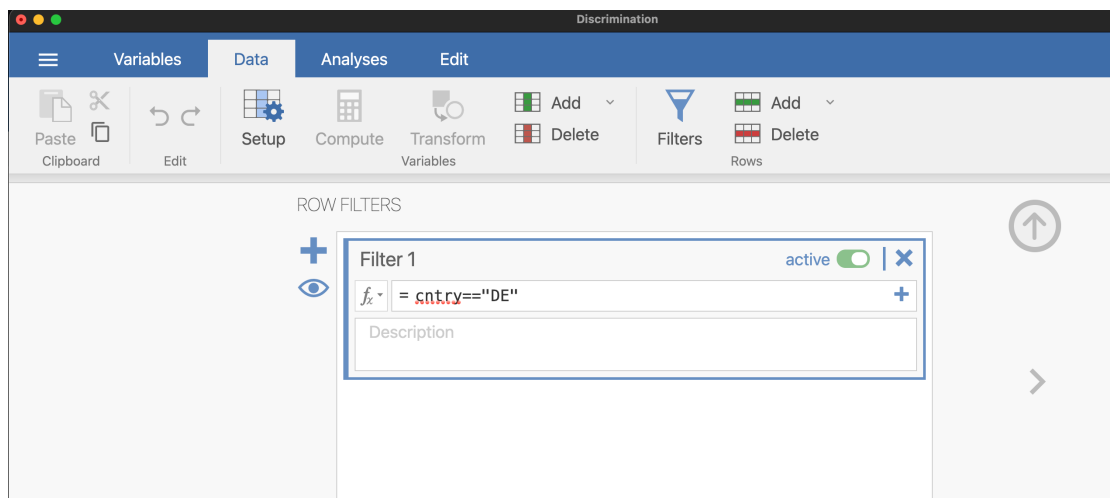
Αν η κλίμακα μέτρησης δεν έχει αναγνωριστεί σωστά από το λογισμικό τότε χρειάζεται να την ορίσουμε εμείς.

**Προκειμένου να εφαρμόσουμε την Παραγοντική Ανάλυση των Αντιστοιχιών (Multiple Correspondence Analysis), πρέπει να αλλάξουμε, για κάθε μεταβλητή, το πεδίο Data type από Integer σε Text, όπως στην Εικόνα 4. Επαναλαμβάνουμε το ίδιο για όλες τις μεταβλητές που πρόκειται να αναλύσουμε με τη συγκεκριμένη μέθοδο.**



Εικόνα 4. Αλλαγή τύπου δεδομένων σε 'Text'


Επειδή το αρχείο δεδομένων περιλαμβάνει τα άτομα από κάθε χώρα που συμμετείχαν στην έρευνα EVS, πρέπει να επιλέξουμε μία συγκεκριμένη χώρα. Για παράδειγμα, θα επιλέξουμε την Γερμανία, η οποία αντιστοιχεί στην τιμή "DE" της μεταβλητής *cntry*. Στην καρτέλα **Data** επιλέγουμε **Filters**. Στο πεδίο *fx* πληκτρολογούμε `cntry=="DE"`, όπως στην παρακάτω Εικόνα.



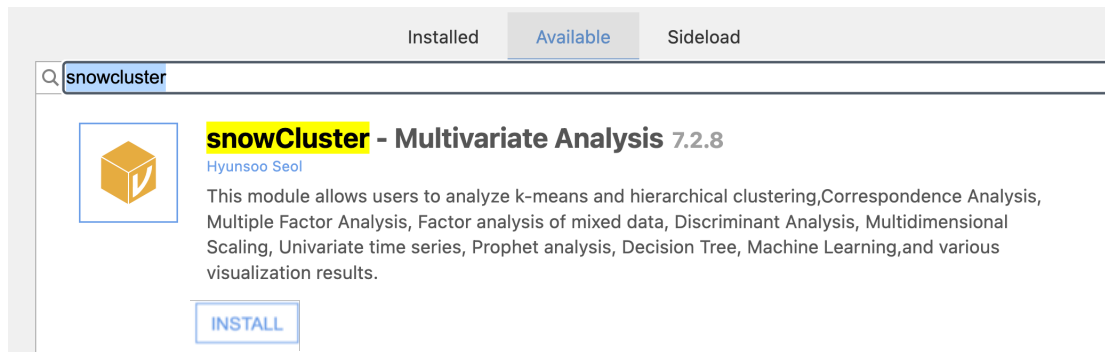
Εικόνα 5. Φιλτράρισμα των υποκειμένων (ατόμων) που αντιστοιχούν σε μια συγκεκριμένη χώρα

## 2. Εφαρμογή της Παραγοντικής Ανάλυσης των Αντιστοιχιών (Multiple Correspondence Analysis)

Για να εφαρμόσουμε την Πολλαπλή Παραγοντική Ανάλυση Αντιστοιχιών (Multiple Correspondence Analysis - MCA) χρειάζεται να έχουμε εγκαταστήσει το πρόσθετο (module) snowCluster.

Κλικ στο  → jamoni library, εντοπίζουμε το πρόσθετο snowCluster και κλικ

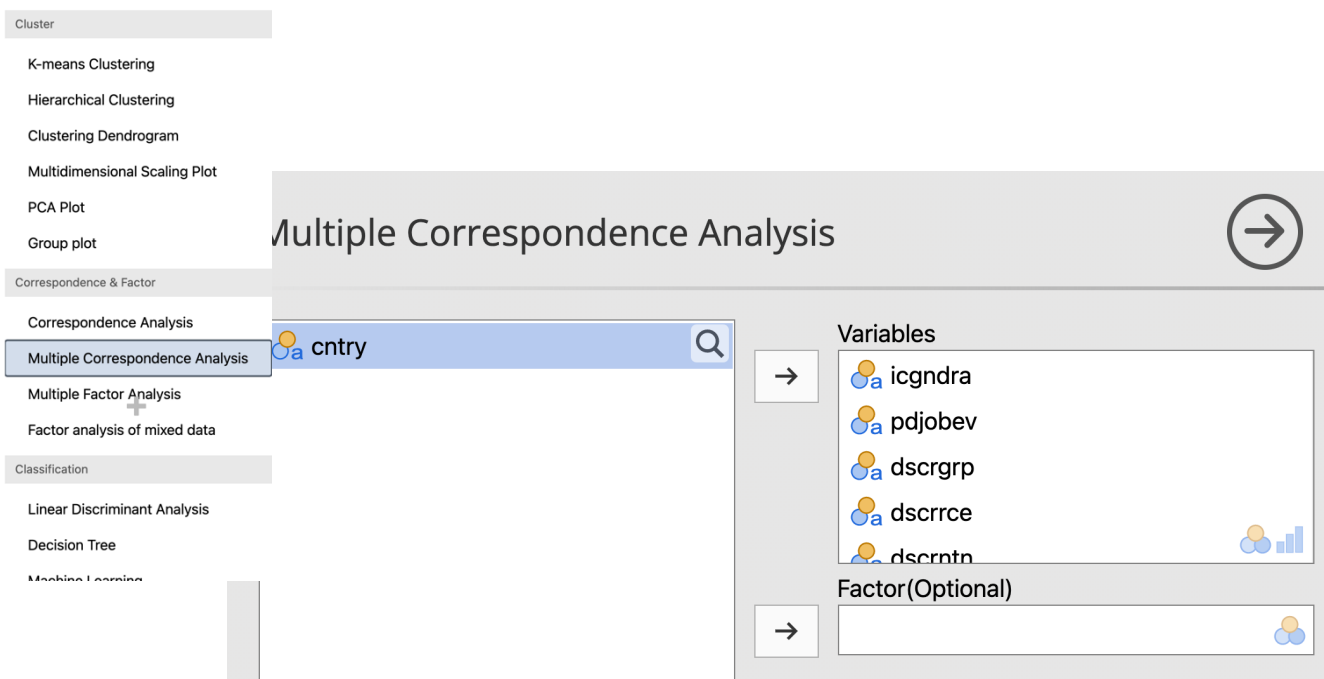
στο **INSTALL** (Εικόνα 6).



Εικόνα 6. Εγκατάσταση module snowCluster από την βιβλιοθήκη του JAMOVI

Βεβαιωνόμαστε ότι το πρόσθετο έχει εγκατασταθεί και εμφανίζεται στη γραμμή εργαλείων (καρτέλα Analyses).

Κλικ στο **snowCluster** από τη γραμμή εργαλείων και στη συνέχεια επιλέγουμε **Multiple Correspondence Analysis**. Μεταφέρουμε τις μεταβλητές στη λίστα Variables (Εικόνα 7).



Εικόνα 7. Η επιλογή μεταβλητών για την MCA

Από τις επιλογές που βρίσκονται στο κάτω μέρος (κλικ στο Multiple Correspondence Analysis) επιλέγουμε **Variable categories** με Type **cos2** και **Individuals** με Type **coordinates**. Από τα Plots, επιλέγουμε **Biplot** και **Coordinates of variables categories** όπως στην Εικόνα 8.

Εικόνα 8. Οι επιλογές της MCA

Στο δεξί μέρος της οθόνης εμφανίζονται τα αποτελέσματα, οι τιμές του δείκτη  $\cos^2$  (COR) των στηλών (κατηγορίες των μεταβλητών - variable categories) και των γραμμών (άτομα - individuals) για τους πέντε πρώτους παραγοντικούς άξονες, καθώς και το παραγοντικό επίπεδο (biplot) με τα σημεία γραμμών και στηλών (όπως και το παραγοντικό επίπεδο μόνο με τα σημεία στηλών).

**Αντιγράφουμε στο αρχείο κειμένου της εργασίας μας τον πίνακα “Variable categories”. Επιλέγουμε ένα όριο για τον δείκτη ποιότητας προβολής COR ή  $\cos^2$  (π.χ. το 0,2), και σημειώνουμε τα σημεία με δείκτη COR > 0,2 στην πρώτη ή τη δεύτερη στήλη του πίνακα. Αυτά είναι τα σημεία που θα ερμηνεύσουμε στο παραγοντικό επίπεδο (βλ πρότυπο εργασίας).**

Μπορούμε να την επιβεβαιώσουμε ή να την αλλάξουμε/εμπλουτίσουμε αυτά τα αποτελέσματα με βάση τα αποτελέσματα της Ανιούσας Ιεραρχικής Ταξινόμησης στις παραγοντικές συντεταγμένες (βλ. επόμενη ενότητα).

#### **Αντιγραφή του πίνακα συντεταγμένων των ατόμων στα αρχικά δεδομένα**

Πριν εφαρμόσουμε την Ανιούσα Ιεραρχική Ταξινόμηση, πρέπει να μεταφέρουμε στα δεδομένα των πίνακα με τις συντεταγμένες των ατόμων στους πέντε πρώτους παραγοντικούς άξονες. Αυτές οι τιμές βρίσκονται στον πίνακα Individuals across dimensions – coordinates.

Από το παράθυρο με τα αποτελέσματα, αντιγράφουμε με δεξί κλικ και Copy τον πίνακα αυτόν. Κάνουμε επικόλληση (paste) σε ένα κενό αρχείο Excel και σβήνουμε τα περιττά ώστε να μην υπάρχουν ενδιάμεσες κενές στήλες. Μεταβαίνουμε στην καρτέλα Data του Jamonί, και επιλέγουμε το πρώτο κελί που βρίσκεται δίπλα στην τελευταία στήλη-μεταβλητή με τα δεδομένα μας (προσοχή: το κελί αυτό βρίσκεται στη γραμμή που βρίσκεται και το πρώτο άτομο της χώρας που επιλέξαμε να αναλύσουμε). Κάνουμε

εκεί copy-paste από το Excel μόνο τις τιμές (όχι δηλαδή τα ονόματα των μεταβλητών Dim1, Dim2 ...), όπως φαίνεται στην Εικόνα 9. Αρχικά τα ονόματα των μεταβλητών θα είναι γράμματα του αγγλικού αλφαβήτου (π.χ. K, L, M, N, O), αλλά μπορούμε να τα μετονομάσουμε με διπλό κλικ στο όνομα της κάθε στήλης σε Dim1, Dim2, Dim3, Dim4 και Dim 5.

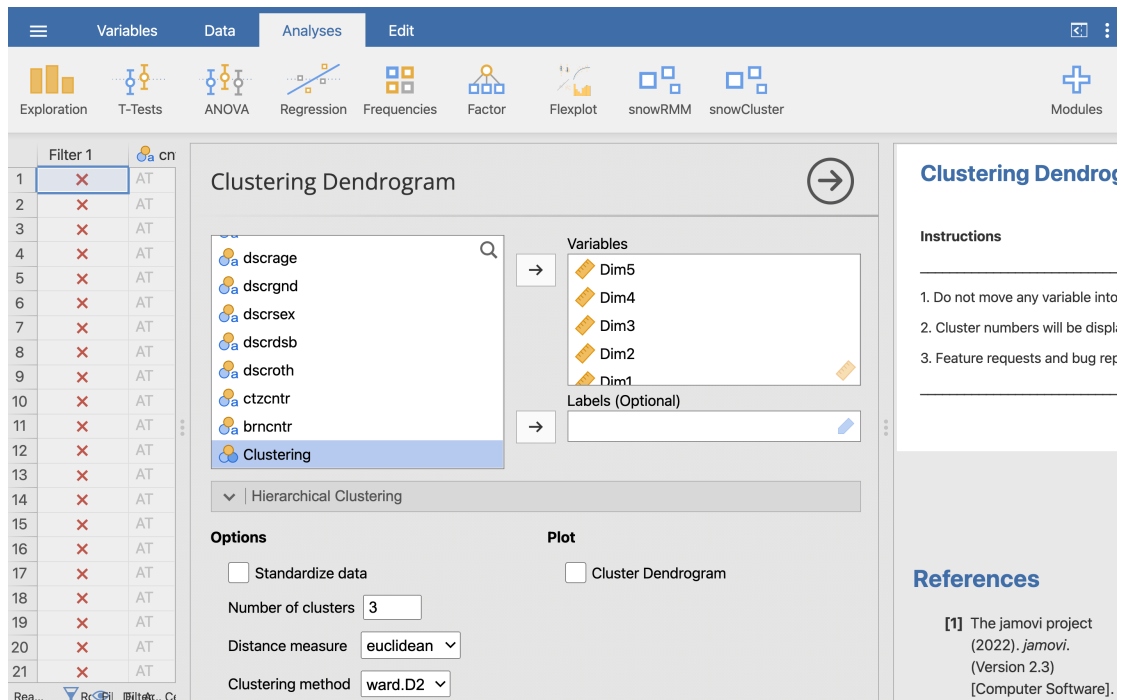
Τελικά, έχουμε δημιουργήσει τις πέντε νέες στήλες που αντιστοιχούν στις συντεταγμένες των ατόμων στους πέντε πρώτους παραγοντικούς άξονες (Εικόνα 9).

jscrdsb	dscroth	ctzcntr	brncntr	Dim1	Dim2	Dim3	Dim4
7481	1	0	1				
7482	1	0	1				
7483	0	1	1				
7484	0	1	1				
7485	0	1	1				
7486	0	1	1				
7487	0	0	1				
7488	0	0	1	-0.198	-0.082	-0.082	0.0
7489	0	0	1	-0.198	-0.082	-0.082	0.0
7490	0	0	1	-0.198	-0.082	-0.082	0.0
7491	0	0	1	-0.198	-0.082	-0.082	0.0
7492	0	0	1	-0.198	-0.082	-0.082	0.0
7493	0	0	1	-0.198	-0.082	-0.082	0.0
7494	0	0	1	-0.198	-0.082	-0.082	0.0
7495	0	0	1	-0.198	-0.082	-0.082	0.0
7496	0	0	1	-0.198	-0.082	-0.082	0.0
7497	0	0	1	-0.198	-0.082	-0.082	0.0
7498	0	0	1	-0.198	-0.082	-0.082	0.0

Εικόνα 9. Αντιγραφή των παραγοντικών συντεταγμένων των ατόμων στους 5 άξονες, ως νέες μεταβλητές

#### 4. Εφαρμογή της Ανιούσας Ιεραρχικής Ταξινόμησης στις παραγοντικές συντεταγμένες

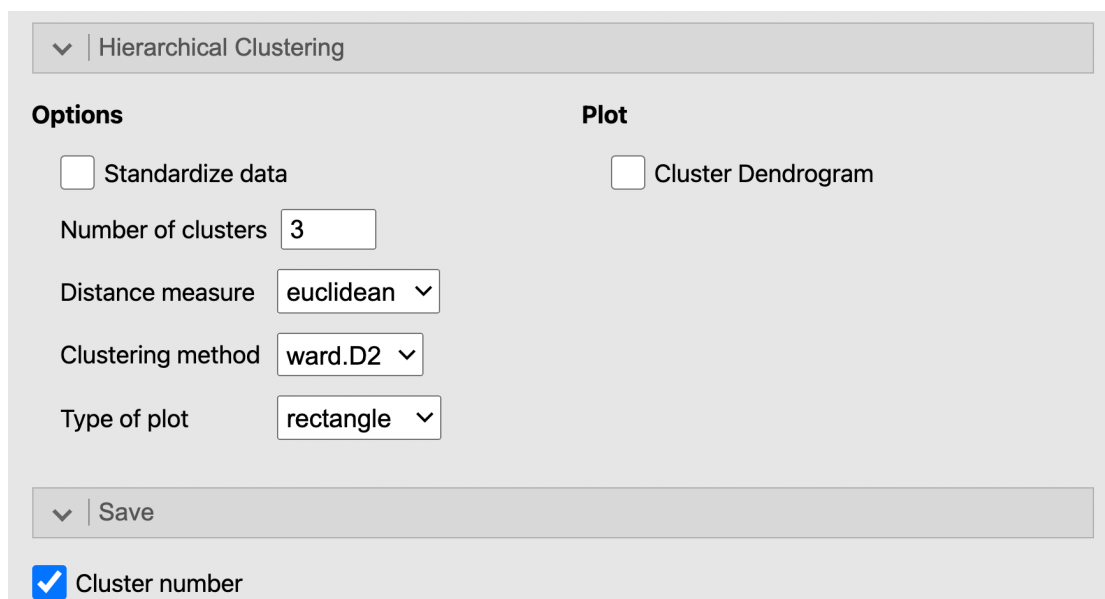
Πηγαίνουμε στο **Analyses** → **snowCluster** από τη γραμμή εργαλείων και στη συνέχεια **Clustering Dendrogram**. Μεταφέρουμε τις πέντε νέες στήλες (Dim1 έως Dim5) στη λίστα Variables (Εικόνα 10).



Εικόνα 10. Η επιλογή των μεταβλητών για την Ανιούσα Ιεραρχική Ταξινόμηση

Στη συνέχεια *αφαιρούμε* τις επιλογές **Standardize data** και **Cluster dendrogram**. Στο πεδίο **Number of Clusters** βάζουμε τον αριθμό των ομάδων που εντοπίσαμε προηγουμένως στο παραγοντικό επίπεδο (έστω 3). Επιπλέον, επιλέγουμε **Save Cluster number** (για την αποθήκευση της μεταβλητής ομαδοποίησης – νέα στήλη Clustering – στα δεδομένα), όπως στην Εικόνα 11.

Ελέγχουμε ότι στην καρτέλα **Data** (δλδ. στα δεδομένα μας) έχει δημιουργηθεί στο τέλος μια νέα στήλη, η μεταβλητή ομαδοποίησης – νέα στήλη Clustering (Εικόνα 12).



Εικόνα 11. Οι επιλογές της Ανιούσας Ιεραρχικής Ταξινόμησης

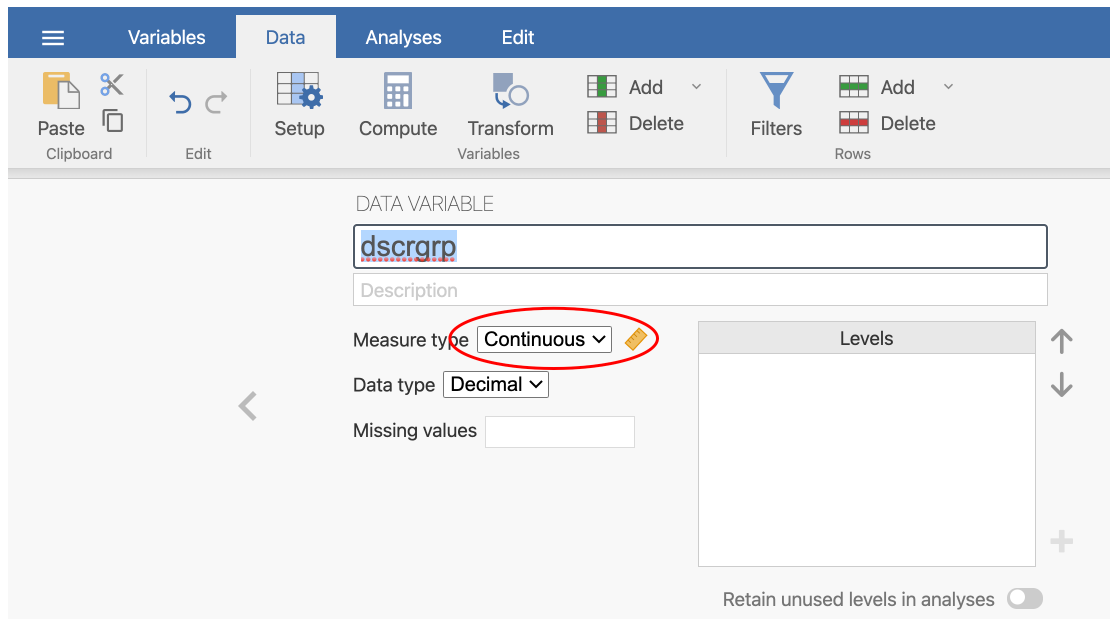
	Dim4	Dim5	Clustering
7473			
7474			
7475			
7476			
7477			
7478			
7479			
7480			
7481			
7482			
7483			
7484			
7485			
7486			
7487			
7488	0.017	0.004	1
7489	0.017	0.004	1
7490	0.017	0.004	1
7491	0.017	0.004	1

Εικόνα 12. Η νέα στήλη Clustering

## 5. Ερμηνεία των ομάδων με βάση τις αρχικές μεταβλητές

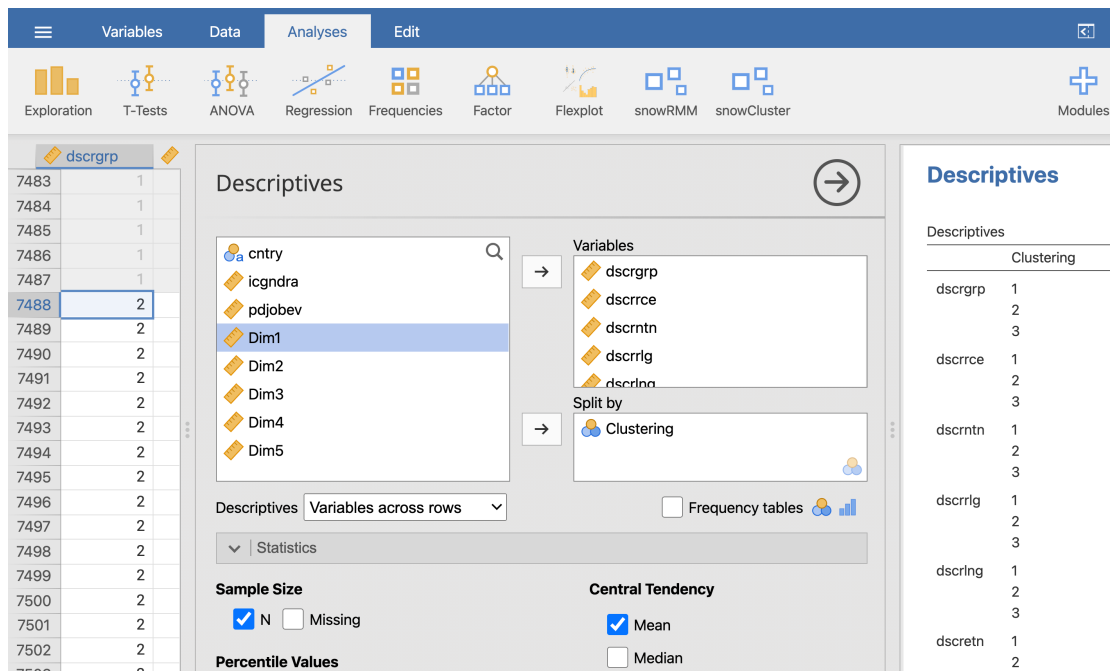
Για να ερμηνεύσουμε τις ομάδες μπορούμε να τις περιγράψουμε ως προς τις αρχικές μεταβλητές. Προηγουμένως, όμως, πρέπει να μετατρέψουμε όλες τις αρχικές μεταβλητές σε ποσοτικές (continuous). Για να το κάνουμε αυτό πηγαίνουμε στα **Data** και κάνουμε διπλό κλικ στο όνομα της κάθε μεταβλητής. Αλλάζουμε το **Measure Type** σε **Continuous**, βλ. Εικόνα 13 για μία μεταβλητή. Αυτό πρέπει να γίνει για όλες τις αρχικές μεταβλητές.





Εικόνα 13. Μετατροπή μεταβλητής σε Continuous

Επιλέγουμε από τη γραμμή εργαλείων **Exploration** → **Descriptives**. Στη λίστα Variables μεταφέρουμε τις αρχικές μεταβλητές και στο πεδίο Split μεταφέρουμε τη μεταβλητή ομαδοποίησης (Clustering), όπως φαίνεται παρακάτω.



Εικόνα 14. Περιγραφή των ομάδων με βάση τις αρχικές μεταβλητές

Στο παράθυρο αυτό επιλέγουμε **Descriptives Variables across rows** και από τα Statistics, μόνο **N** και **Mean**.

Αν κοιτάξουμε τα αποτελέσματα (Εικόνα 15) και εστιάσουμε στους μέσους όρους (Mean) θα διαπιστώσουμε ότι τα άτομα στις ομάδες 1 (1243 άτομα, 79,17%) και 2 (103 άτομα, 6,5%) περιγράφουν τον εαυτό τους ως άτομα που δεν υφίστανται διακρίσεις (μέσοι όροι 1,99 και 1,78 – θυμίζουμε εδώ ότι στη συγκεκριμένη μεταβλητή το 2

σημαίνει Όχι και το 1 = Ναι), ενώ η ομάδα 3 περιλαμβάνει άτομα που υφίστανται διακρίσεις (224 άτομα, 14,26%), με μέσο όρο 1.

### Descriptives

Descriptives			
	Clustering	N	Mean
dscrgrp	1	1243	1.99678
	2	103	1.78641
	3	224	1.00000
dscrce	1	1243	0.00000
	2	103	0.07767
	3	224	0.23661
dscrntn	1	1243	0.00000
	2	103	0.08738
	3	224	0.08482
dscrllg	1	1243	0.00000
	2	103	0.02913
	3	224	0.18750
dscrlnq	1	1243	0.00000
	2	103	0.00000
	3	224	0.03571
dscrcrn	1	1243	0.00000
	2	103	0.00000
	3	224	0.16071

Εικόνα 15. Μέσοι όροι των τριών ομάδων στις αρχικές μεταβλητές