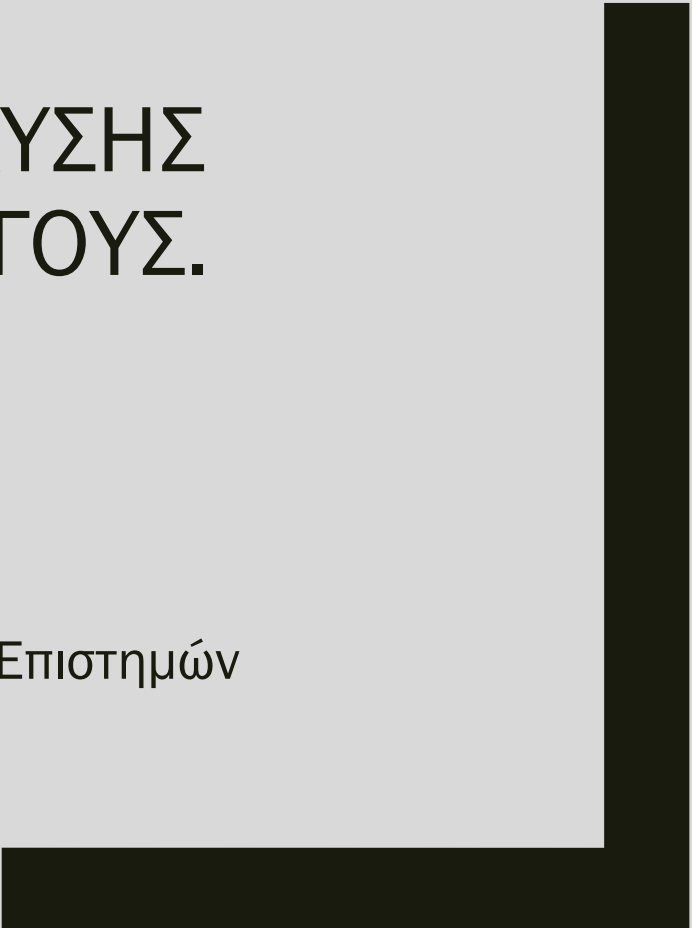




ΑΝΟΙΚΤΑ ΛΟΓΙΣΜΙΚΑ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΨΥΧΟΛΟΓΟΥΣ. Η R & ΤΟ JAMON

Φώτης Φωτιάδης
Τμήμα Ψυχολογίας
Πάντειο Πανεπιστήμιο Κοινωνικών & Πολιτικών Επιστημών



Διάλεξη 11

■ ΠΕΡΙΕΧΟΜΕΝΑ

1. Περιγραφική στατιστική.
 - 1.1. Περιγραφικά Μέτρα
 - 1.2. Έλεγχος Κανονικότητας
 - 1.3. Αποθήκευση, Διαγραφή, Εξαγωγή Αποτελεσμάτων
 - 1.4. Περιγραφικά Μέτρα Ανά Ομάδα
2. Γραφήματα από το μενού Exploration
 - 2.1. Επεξεργασία Γραφημάτων
 - 2.2. Αντιγραφή, Εξαγωγή Γραφημάτων
3. Έλεγχος χ^2

1. Περιγραφική Στατιστική

Εισαγωγή

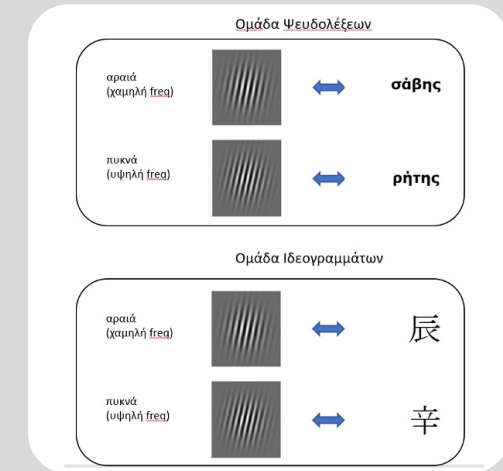
Όταν αναφέρουμε τα αποτελέσματα μια έρευνας, πρέπει:

Να αναφέρουμε δημογραφικά στοιχεία του δείγματος που έλαβε μέρος στην έρευνά μας, δηλαδή ΜΟ και ΤΑ ηλικίας, καθώς και στοιχεία για το φύλλο των συμμετεχόντων.

Να αναφέρουμε περιγραφικά στατιστικά για την εξαρτημένη μεταβλητή, όπως είναι ο ΜΟ και η ΤΑ.

Το Πείραμα και τα Δεδομένα

Θα χρησιμοποιήσουμε τα δεδομένα από το γνωστό πείραμα, όπου συμμετέχοντες μαθαίνουν δύο κατηγορίες. Στη μία ομάδα, οι ετικέτες των κατηγοριών είναι ψευδολέξεις, ενώ στην άλλη ομάδα οι ετικέτες είναι ιδεογράμματα.



- Κατεβάζουμε το αρχείο **independent_samples.zip** από το eclass (Εγγραφα> Διάλεξη 10)
- Αποσυμπιέζουμε το αρχείο, και μεταφέρουμε το απόσυμπιεσμένο αρχείο (**independent_samples.omv**) στο φάκελο **jamovi_mathima**.
- Ανοίγουμε το αρχείο **independent_samples.omv** με το jamovi.

Δεδομένα

Τα δεδομένα είναι κάπως έτσι:

Μεταβλητές:

- **sex**: ονομαστική με τρία επίπεδα: F, M, other
- **age**: συνεχής
- **sbj**: ID
- **group**: ονομαστική, με δύο επίπεδα, ideo και pseudo
- **acc**: συνεχής

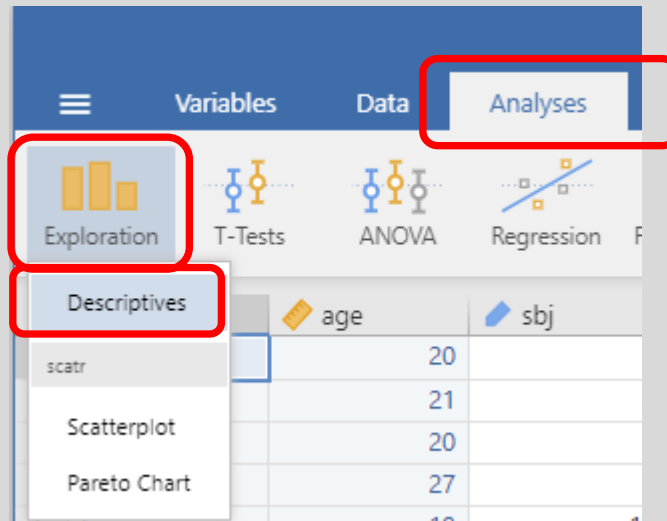
	sex	age	sbj	group	acc
1	F	20	1	ideo	0.575
2	M	21	4	ideo	0.625
3	other	20	6	ideo	0.863
4	F	27	8	ideo	0.850
5	F	19	10	ideo	0.588
6	F	20	12	ideo	0.887
7	M	19	14	ideo	0.887
8	F	20	15	ideo	0.738
9	M	20	16	ideo	0.637
10	F	22	18	ideo	0.863
11	F	22	19	ideo	0.713
12	F	31	24	ideo	0.400
13	other	26	2	pseudo	0.500
14	F	20	3	pseudo	0.738
15	M	19	5	pseudo	0.738
16	M	20	7	pseudo	0.863
17	M	20	9	pseudo	0.900
18	F	22	11	pseudo	0.750
19	M	21	13	pseudo	0.800
20	F	19	17	pseudo	0.738
21	F	20	20	pseudo	0.613
22	F	29	21	pseudo	0.887
23	F	20	22	pseudo	0.787
24	F	20	23	pseudo	0.487
25					
26					
27					
28					

1.1. Περιγραφικά Μέτρα

Descriptives I

Για να υπολογίσουμε το ΜΟ και την ΤΑ της ηλικίας του δείγματός μας, πρέπει να κάνουμε **κλικ** στο **μενού Analyses**, και στη συνέχεια στο **Exploration**.

Στο πτυσσόμενο μενού που εμφανίζεται κάνουμε κλικ στο **Descriptives**.



Descriptives II

Εμφανίζεται το διπλανό παράθυρο:

- Στα αριστερά εμφανίζονται όλες οι μεταβλητές
- Στο πεδίο **Variables** μεταφέρουμε τις μεταβλητές εκείνες για τις οποίες θέλουμε να υπολογίσουμε περιγραφικά στατιστικά μεγέθη.
- Στο πεδίο **Split by** μεταφέρουμε τις (ονομαστικές) μεταβλητές εκείνες, οι οποίες διαχωρίζουν το δείγμα μας σε ομάδες.

Descriptives

Variables

Split by

Descriptives Variables across columns

Frequency tables

> | Statistics

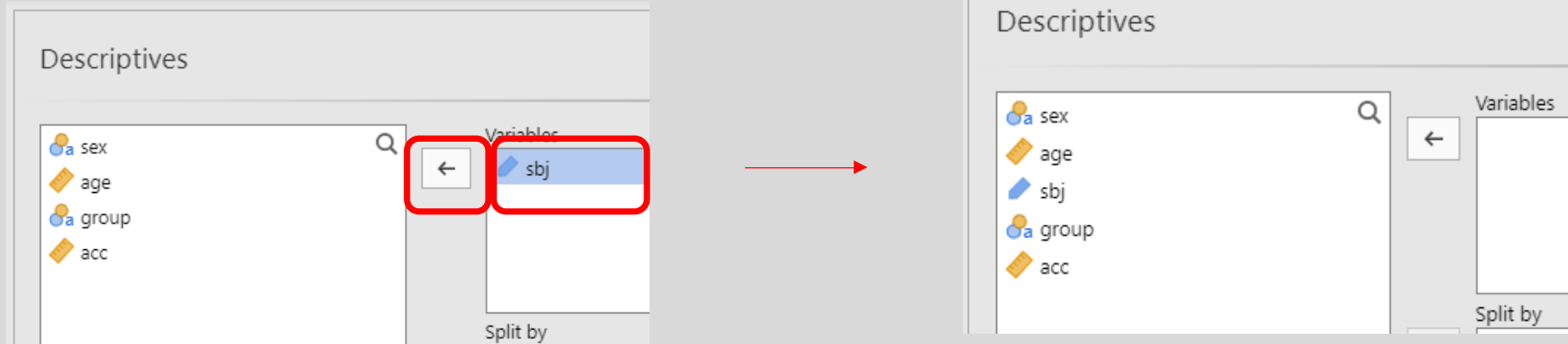
> | Plots

Ηλικία

Αρχικά μεταφέρουμε στο πλαίσιο **Variables** τη μεταβλητή **age**



Σε περίπτωση που κάνουμε λάθος, και μεταφέρουμε άλλη μεταβλητή (π.χ, την **sbj**), μπορούμε να την επαναφέρουμε στη λίστα



Ηλικία II

Αυτόματα βλέπουμε πως δημιουργείται ένα πίνακάκι με κάποια στατιστικά μεγέθη.

Αν χρειαζόμαστε άλλα μεγέθη, πρέπει να κάνουμε κλικ στο **Statistics**.

The screenshot shows the JAMOVI software interface with the 'Analyses' tab selected. In the 'Descriptives' panel, the variable 'age' is selected. The 'Statistics' option is highlighted with a red box. The 'References' panel on the right displays the following table:

Descriptives	
	age
N	24
Missing	0
Mean	21.5
Median	20.0
Standard deviation	3.28
Minimum	19
Maximum	31

Below the table, the 'References' section lists two sources:

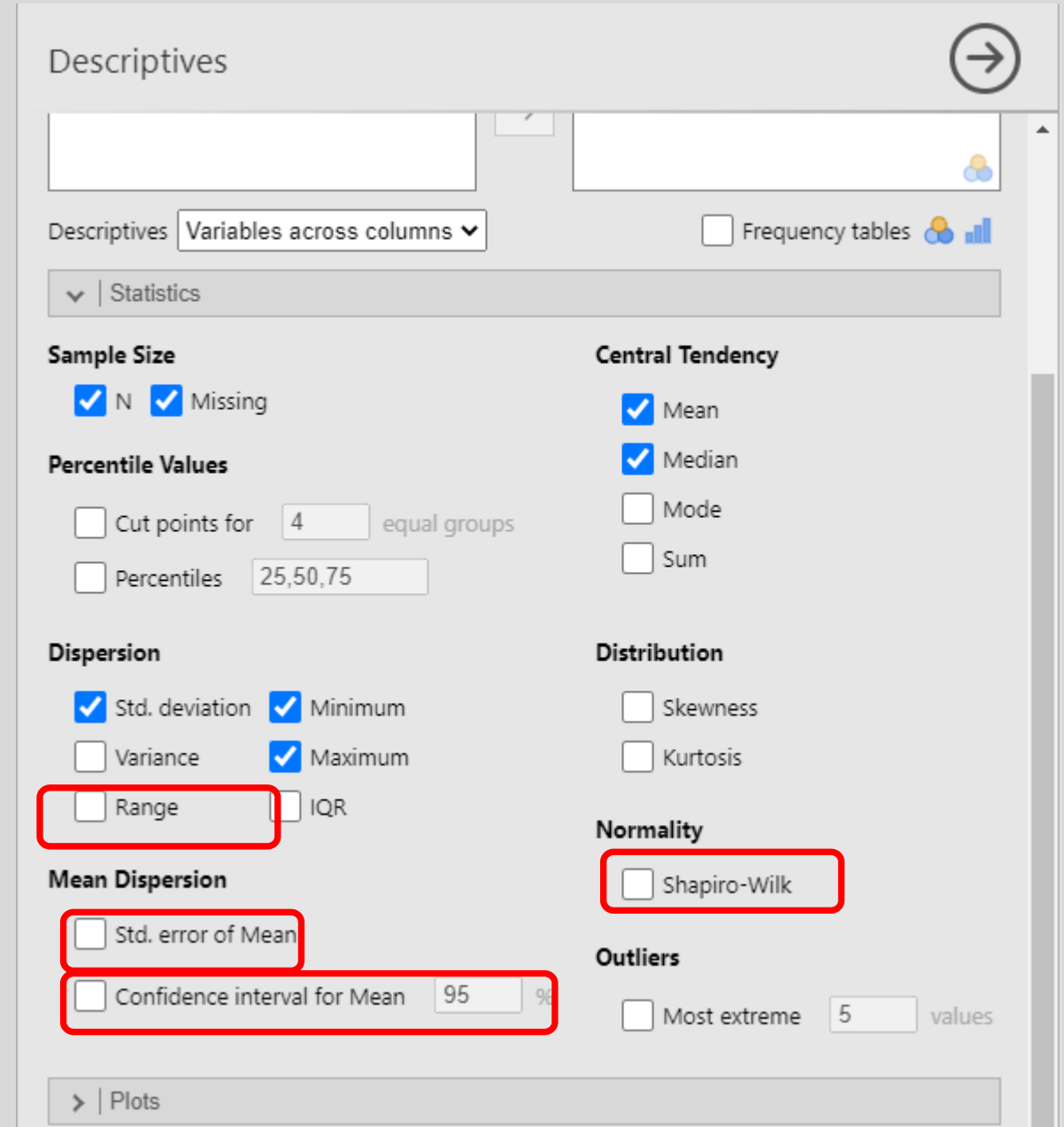
- [1] The jamovi project (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- [2] R Core Team (2021). *R: A Language and environment for statistical computing*. (Version 4.1) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from MRAN snapshot 2022-01-01).

Statistics

Στο πεδίο statistics έχουμε διαθέσιμους όλους του υπολογισμούς που μπορούν να εκτελεστούν.

Για παράδειγμα μπορούμε να υπολογίσουμε

- το εύρος τιμών (**range**)
- το τυπικό σφάλμα του μέσου όρου (**Std. error of Mean**)
- τον στατιστικό έλεγχο κανονικότητας **Shapiro-Wilk**
- το διάστημα εμπιστοσύνης 95%



The screenshot shows the 'Descriptives' dialog box in SPSS. The 'Statistics' section is expanded, showing various options. The following options are highlighted with red boxes:

- Range** (under Dispersion)
- Std. error of Mean** (under Mean Dispersion)
- Confidence interval for Mean** (under Mean Dispersion, with 95% selected)
- Shapiro-Wilk** (under Normality)

Other visible options include:

- Sample Size:** N (checked), Missing (checked)
- Central Tendency:** Mean (checked), Median (checked), Mode, Sum
- Percentile Values:** Cut points for 4 equal groups, Percentiles 25,50,75
- Dispersion:** Std. deviation (checked), Variance, Range, Minimum (checked), Maximum (checked), IQR
- Distribution:** Skewness, Kurtosis
- Outliers:** Most extreme 5 values

Ηλικία III

Επιλέγουμε τα προηγούμενα μέτρα.

Όσο τα επιλέγουμε, παρατηρούμε πως προστίθενται γραμμές στον πίνακά μας.

Descriptives

Descriptives Variables across columns

Statistics

Sample Size

- ☒ N
- ☒ Missing

Percentile Values

- ☐ Cut points for 4 equal groups
- ☐ Percentiles 25,50,75

Dispersion

- ☒ Std. deviation
- ☐ Variance
- ☒ Range
- ☐ IQR
- ☒ Minimum
- ☒ Maximum

Mean Dispersion

- ☒ Std. error of Mean
- ☒ Confidence interval for Mean 95 %

Central Tendency

- ☒ Mean
- ☒ Median
- ☐ Mode
- ☐ Sum

Distribution

- ☐ Skewness
- ☐ Kurtosis

Normality

- ☒ Shapiro-Wilk

Outliers

- ☐ Most extreme 5 values

Results

Descriptives

age	
N	24
Missing	0
Mean	21.5
Std. error mean	0.670
95% CI mean lower bound	20.2
95% CI mean upper bound	22.9
Median	20.0
Standard deviation	3.28
Range	12
Minimum	19
Maximum	31
Shapiro-Wilk W	0.695
Shapiro-Wilk p	< .001

Note. The CI of the mean assumes sample means follow a t-distribution with N - 1 degrees of freedom

1.2. Έλεγχος Κανονικότητας

Shapiro-Wilk

Έκτός από τα μέτρα περιγραφικής στατιστικής, εκτελέσαμε και τον έλεγχο Shapiro-Wilk για την κανονικότητα των δεδομένων μας (ηλικία).

Στα αποτελέσματα εμφανίζονται το στατιστικό μέτρο W καθώς το αντίστοιχο p .

Επομένως, τα αποτελέσματα του ελέγχου είναι $W = 0.695, p < .001$

Shapiro-Wilk W	0.695
Shapiro-Wilk p	< .001

Σημείωση: Στις περιπτώσεις των προϋποθέσεων (κανονικότητα, ομοσκεδαστικότητα κ.λ.π.), θέλουμε το p να είναι μεγαλύτερο του 0.05. Στην περίπτωση του ελέγχου Shapiro-Wilk η μηδενική υπόθεση είναι πως δεν υπάρχει διαφορά μεταξύ της κατανομής των δεδομένων του δείγματος και της κανονικής κατανομής (που θεωρούμε πως χαρακτηρίζει τα δεδομένα όλου του πληθυσμού). Αν λοιπόν $p > 0.05$, τότε τα δεδομένα μας χαρακτηρίζονται από κανονική κατανομή.

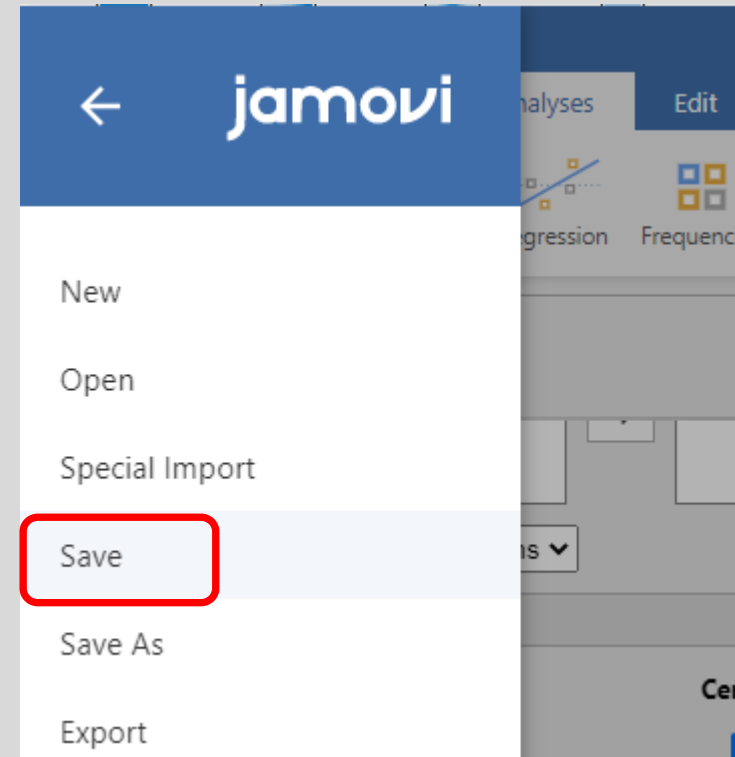
1.3. Αποθήκευση, Διαγραφή, Εξαγωγή Αποτελεσμάτων

Αποθήκευση Αποτελεσμάτων

Στο jamoní, όταν αποθηκεύουμε ένα αρχείο, αποθηκεύονται μαζί και τα αποτελέσματα των υπολογισμών ή των αναλύσεων, στο ίδιο αρχείο οmn.

Δεν έχουμε δηλαδή ξεχωριστά αρχεία για τα δεδομένα και για τα αποτελέσματα των υπολογισμών/αναλύσεων (αυτό ισχύει μόνο στο SPSS).

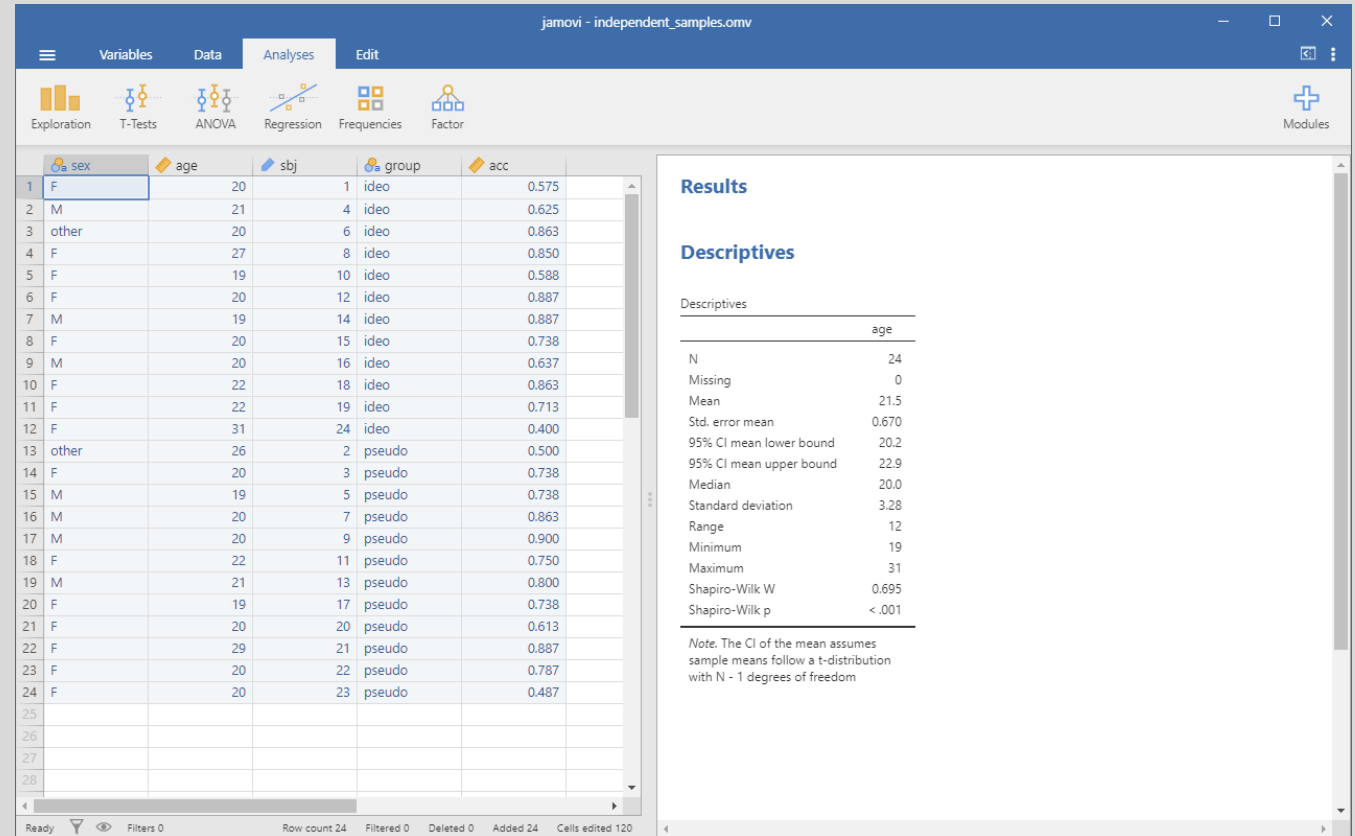
- Για να αποθηκεύσουμε τα αποτελέσματα (μαζί με τα δεδομένα), κάνουμε κλικ στο **μενού File** και στη συνέχεια στην εντολή **Save**.



Άνοιγμα Αρχείου .omv

- Κλείνουμε το jamovi
- Το ξανα-ανοίγουμε
- Ανοίγουμε το αρχείο independent_samples.omv

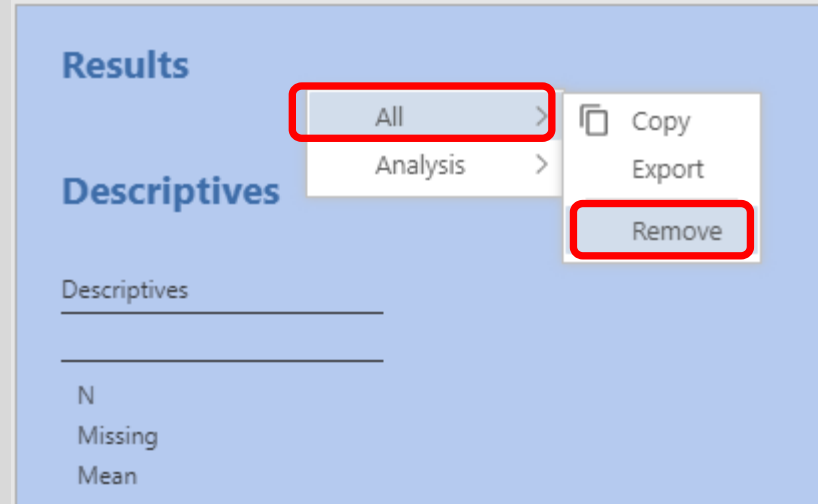
Παρατηρούμε πως εμφανίζονται και τα αποτελέσματα των υπολογισμών/αναλύσεων μας.



Διαγραφή Αποτελεσμάτων

Αν θέλουμε να διαγράψουμε κάποια αποτελέσματα:

- Κάνουμε δεξί κλικ στα συγκεκριμένα αποτελέσματα.
- Παρατηρούμε πως επιλέγεται ο πίνακας των αποτελεσμάτων, και πως εμφανίζεται ένα πτυσσόμενο μενού.
- Κάνουμε κλικ στο All
- Κάνουμε κλικ στο Remove



Σημείωση: Επιλογή Αποτελεσμάτων

- Επαναλαμβάνουμε τους υπολογισμούς (Range, Shapiro-Wilk, Std error, Confidence Intervals).

Τα αποτελέσματα δομούνται υπό τη μορφή στοιχείων πινάκων.

Άλλο στοιχείο είναι ο τίτλος Results

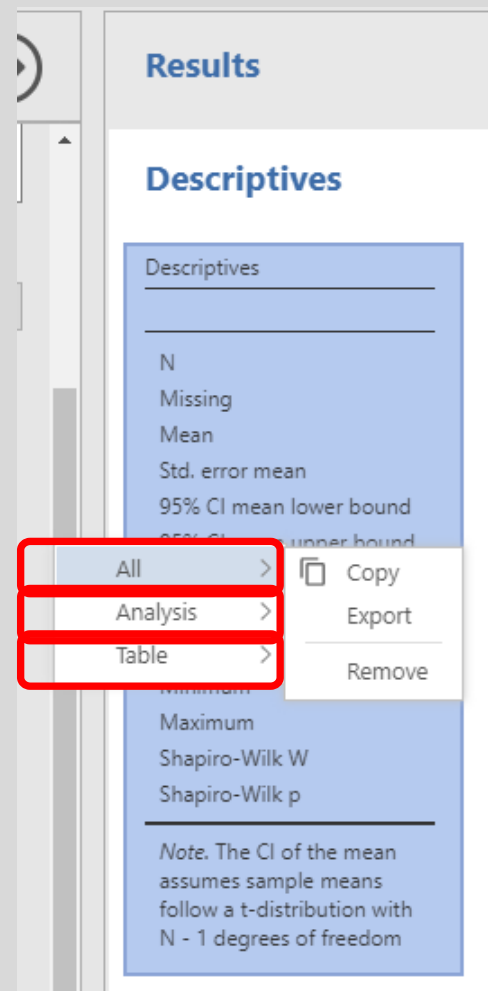
Άλλο στοιχείο είναι ο τίτλος Descriptives

Άλλο στοιχείο είναι ο πίνακας Descriptives

Παρατηρούμε πως αν κάνουμε δεξί κλικ σε διαφορετικά σημεία του δεξιού πλαισίου, επιλέγονται διαφορετικά στοιχεία.

Μπορούμε, επίσης, να κάνουμε δεξί κλικ στον πίνακα, και να επιλέξουμε ένα από:

- All
- Analyses
- Table



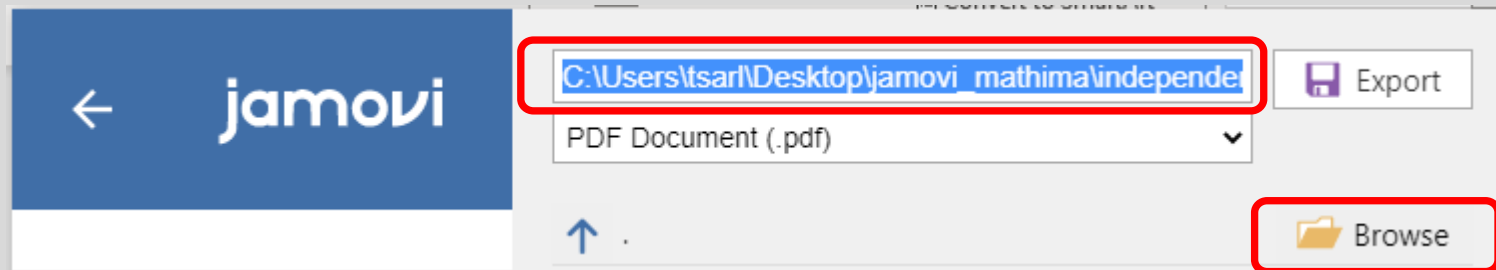
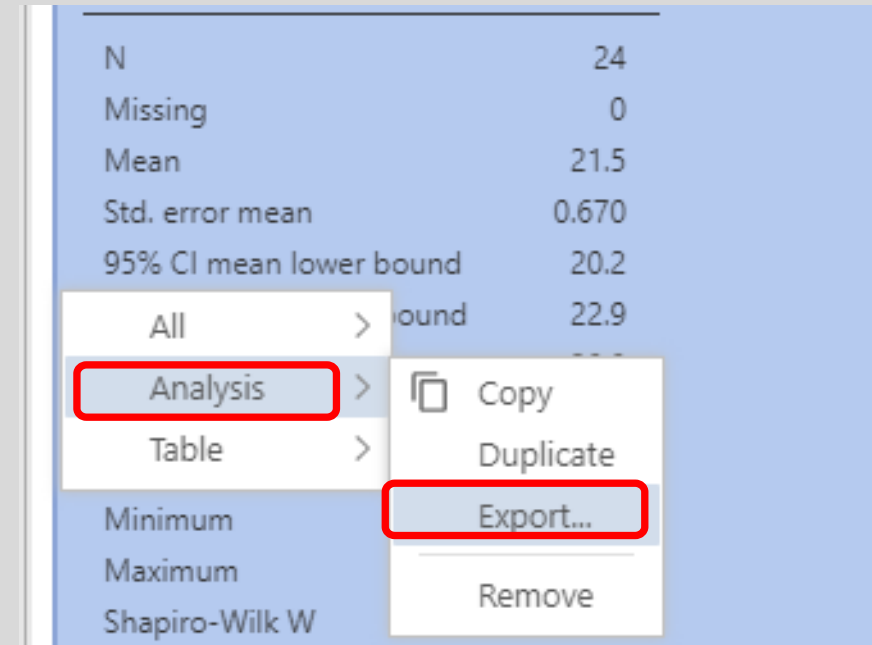
Εξαγωγή Αποτελεσμάτων σε pdf

Στο jamoni, έχουμε τη δυνατότητα να κάνουμε εξαγωγή (export) των αποτελεσμάτων σε μορφή pdf.

- Κάνουμε **δεξί κλικ** στα αποτελέσματα
- Κάνουμε κλικ στο **Analysis**
- Στη συνέχεια, επιλέγουμε την εντολή **Export...**

Στο πλαίσιο που εμφανίζεται υπάρχει η δυσκολία πως εμφανίζεται η πλήρης διαδρομή του αρχείου.

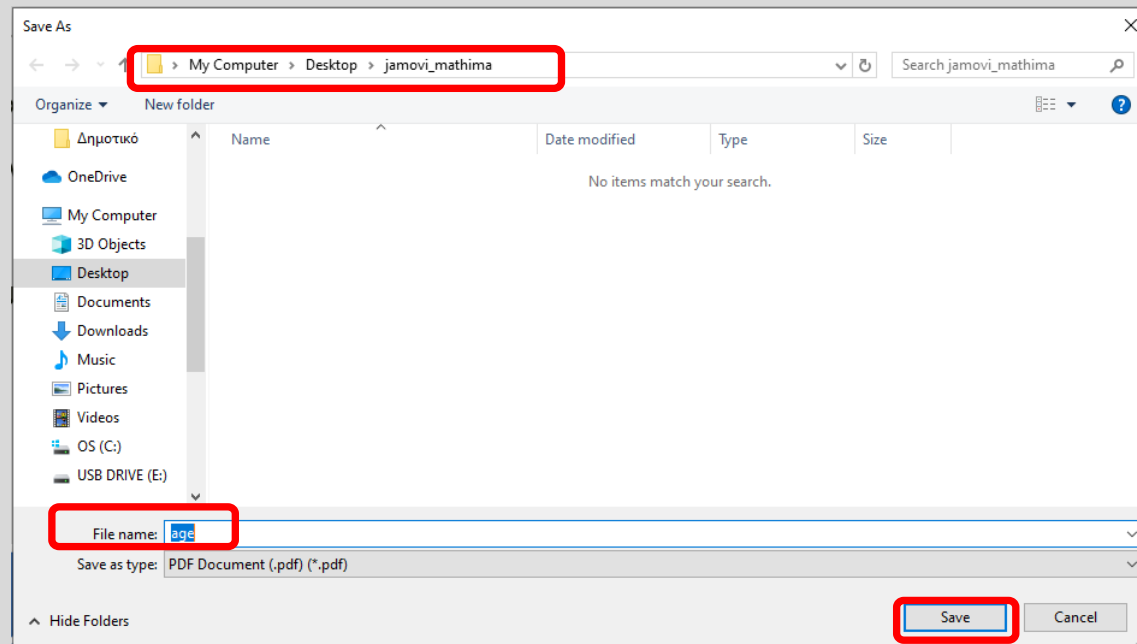
Για να ξεπεράσουμε το ζήτημα αυτό, κάνουμε κλικ στο **Browse**.



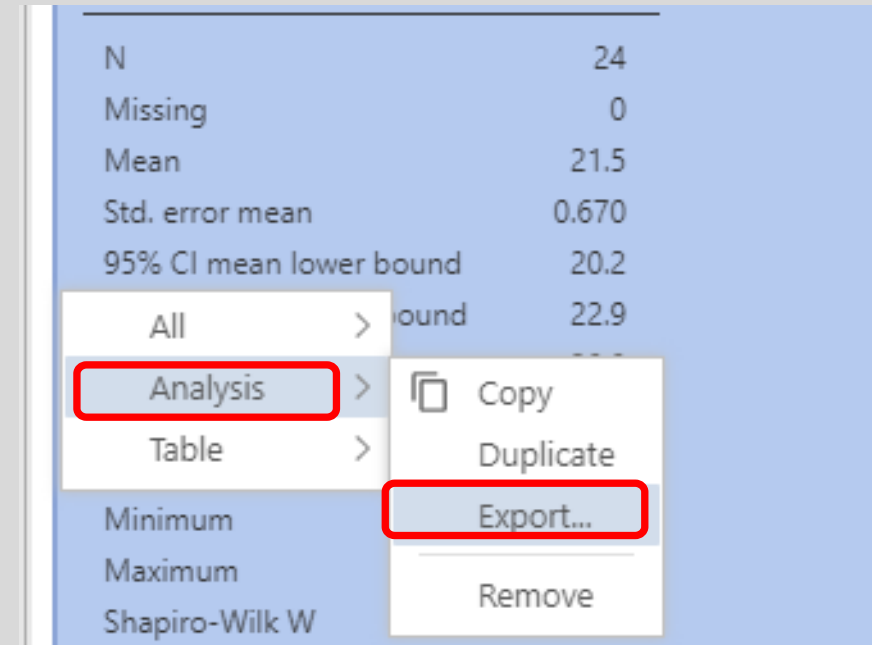
Εξαγωγή Αποτελεσμάτων σε pdf II

Στο πλαίσιο που εμφανίζεται επιλέγουμε

- όνομα (age)
- φάκελο (jamovi_mathima)



- κάνουμε κλικ στο Save



Εξαγωγή Αποτελεσμάτων σε pdf III

Εντοπίζουμε το αρχείο **age.pdf** στον φάκελο `jamoni_mathima`, και το ανοίγουμε (κάνοντας διπλό κλικ).

Descriptives

Descriptives

	age
N	24
Missing	0
Mean	21.5
Std. error mean	0.670
95% CI mean lower bound	20.2
95% CI mean upper bound	22.9
Median	20.0
Standard deviation	3.28
Range	12
Minimum	19
Maximum	31
Shapiro-Wilk W	0.695
Shapiro-Wilk p	< .001

Note. The CI of the mean assumes sample means follow a t-distribution with N - 1 degrees of freedom

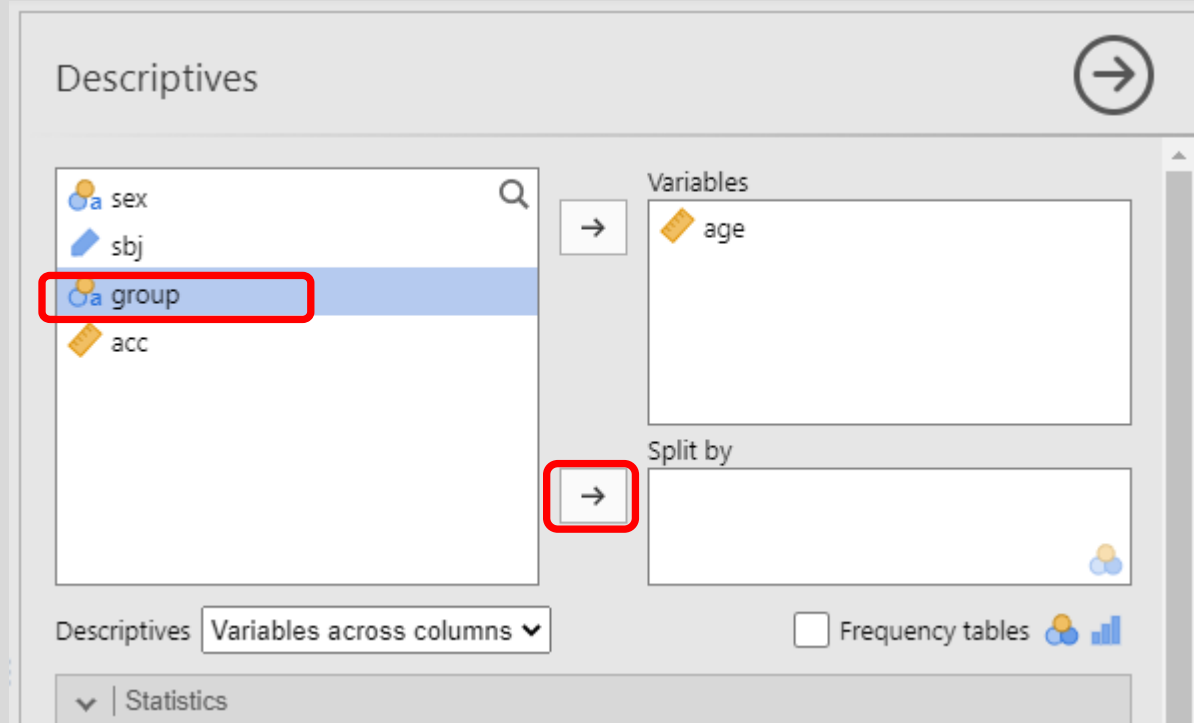
1.4. Περιγραφικά Μέτρα Ανά Ομάδα

Ηλικία ανά Ομάδα

Σε κάποιες περιπτώσεις, όταν, π.χ., έχουμε πειραματική ομάδα και ομάδα ελέγχου, πρέπει να αναφέρουμε τα δημογραφικά στοιχεία ανά ομάδα.

Για να υπολογίσουμε περιγραφικά στατιστικά μεγέθη για κάθε ομάδα, πρέπει να εισάγουμε στο πεδίο **Split by** την (ονομαστική) μεταβλητή που διακρίνει τους συμμετέχοντες.

- Εν προκειμένω, στο πλαίσιο Descriptives πρέπει να εισάγουμε στο πεδίο Split by την μεταβλητή group.



Ηλικία ανά Ομάδα II

Βλέπουμε πως αυτόματα εμφανίζονται οι υπολογισμοί της μεταβλητής μας (ηλικία) ανά ομάδα.

The screenshot shows the SPSS Descriptives dialog box on the left and the Results window on the right. In the Descriptives dialog, the variable 'age' is selected in the 'Variables' list, and 'group' is selected in the 'Split by' list. The 'Statistics' section at the bottom is expanded, showing various statistical options. The Results window displays a table of Descriptives for the variable 'age' split by 'group'.

Descriptives

Variables: age

Split by: group

Statistics: ☒ Mean ☒ Std. deviation ☒ Sum of squares ☒ Minimum ☒ Maximum ☒ Range ☒ Variance ☒ Skewness ☒ Kurtosis ☒ Std. error mean ☒ Std. error sum of squares ☒ Std. error variance ☒ Std. error skewness ☒ Std. error kurtosis ☒ Std. error range ☒ Std. error minimum ☒ Std. error maximum ☒ Std. error variance ☒ Std. error skewness ☒ Std. error kurtosis ☒ Std. error range ☒ Std. error minimum ☒ Std. error maximum

Results

Descriptives

	group	age
N	ideo	12
	pseudo	12
Missing	ideo	0
	pseudo	0
Mean	ideo	21.8
	pseudo	21.3
Std. error mean	ideo	1.05
	pseudo	0.882

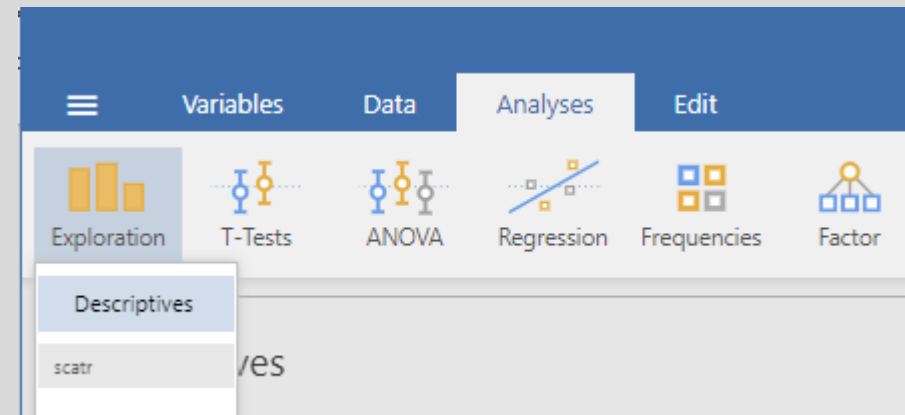
Εξαρτημένη Μεταβλητή ανά Ομάδα

Στην ενότητα Αποτελέσματα (σε μια πτυχιακή ή σε ερευνητική αναφορά), ξεκινάμε με τα περιγραφικά μέτρα της εξαρτημένης μεταβλητής μας.

Επειδή στο πείραμά μας έχουμε δύο ομάδες (pseudo ideο), καλό είναι να αναφέρουμε ΜΟ και ΤΑ ακρίβειας για κάθε μία ομάδα.

Για να μην διαγραφεί η ανάλυση που έχουμε κάνει (ηλικία ανά ομάδα) πρέπει να κάνουμε κλικ:

- στο μενού **Analyses** (αν δεν είναι ήδη επιλεγμένο)
- στο **Exploration**
- στην επιλογή **Descriptives**



Εξαρτημένη Μεταβλητή ανά Ομάδα II

- Μετάφέρουμε την μεταβλητή acc στο πλαίσιο Variables
- Μεταφέρουμε την μεταβλητή group στο πλαίσιο Split by

The screenshot shows the SPSS Descriptives dialog box on the left and the resulting output window on the right.

Descriptives Dialog Box:

- Variables:** acc
- Split by:** group
- Statistics:** Sample Size (N, Missing), Central Tendency (Mean, Median), Percentile Values (Cut points for 4 equal groups).

Descriptives Output:

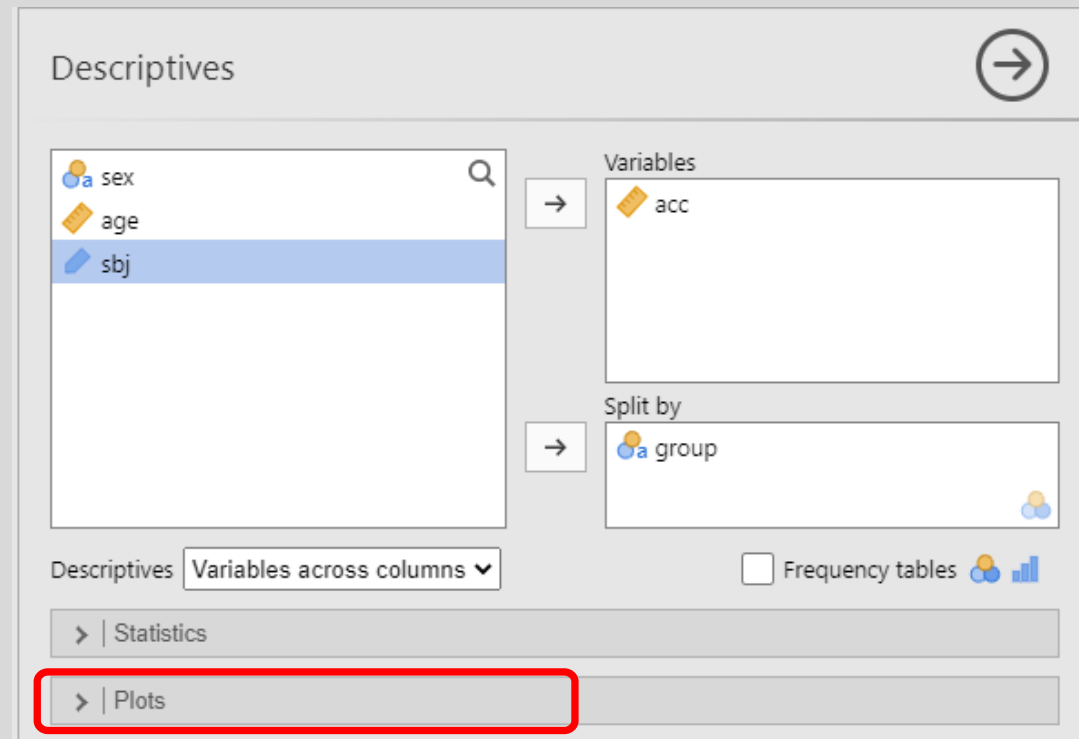
	group	acc
N	ideo	12
	pseudo	12
Missing	ideo	0
	pseudo	0
Mean	ideo	0.719
	pseudo	0.733
Median	ideo	0.725
	pseudo	0.744
Standard deviation	ideo	0.157
	pseudo	0.137

Παρατηρούμε πως αυτόματα εκτελούνται οι υπολογισμοί.

2. Γραφήματα από το μενού Exploration

Γραφήματα

- (Ανοίγουμε το αρχείο independent_samples.omv.)
- (Κάνουμε κλικ στο μενού Exploration, και μετά στο Descriptives.)
- (Εισάγουμε την μεταβλητή acc στο Πλαίσιο Variables και την μεταβλητή group στο Split by.)
- Κάνουμε κλικ στο **Plots**



Γραφήματα II

Εμφανίζονται στο πεδίο Plots κάποιες επιλογές:

- Histograms
- Box Plots
- Bar Plots
- Q-Q Plots

The screenshot shows a software interface with a 'Plots' menu. The menu is open, displaying several options organized into columns. The options are as follows:

Histograms	Box Plots	Bar Plots
☐ Histogram	☐ Box plot	☐ Bar plot
☐ Density	☐ Label outliers	
Q-Q Plots	☐ Violin	
☐ Q-Q	☐ Data	
	Jittered v	
	☒ Mean	

Ιστογράμματα

Μπορούμε να εμφανίσουμε είτε

- μόνο το ιστόγραμμα

Histograms

☒ Histogram

☐ Density

- μόνο την συνάρτηση πυκνότητας πιθανότητας (κατανομή)

Histograms

☐ Histogram

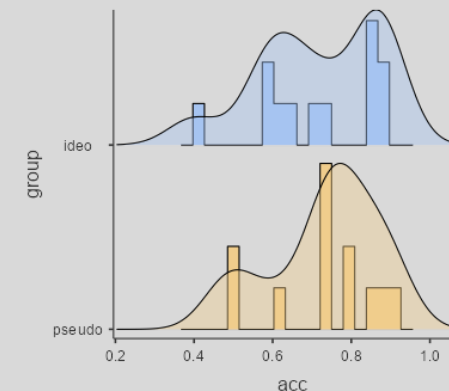
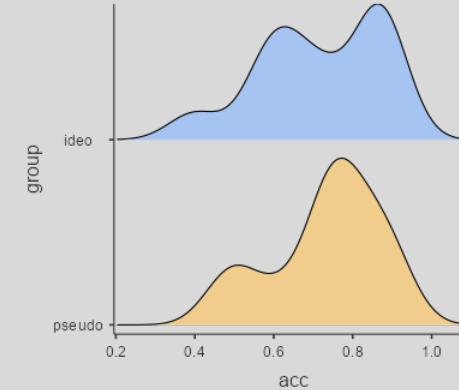
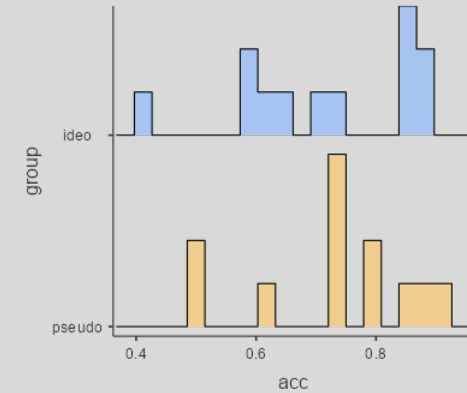
☒ Density

- και τα δύο

Histograms

☒ Histogram

☒ Density



Θηκογράμματα

Μπορούμε να εμφανίσουμε είτε

- μόνο το θηκόγραμμα
(το label outliers εμφανίζει τη γραμμή στα δεδομένα που υπάρχουν οι ακραίες τιμές)

Box Plots

☒ Box plot

☐ Label outliers

- μόνο το violin plot

Box Plots

☐ Box plot

☐ Label outliers

☒ Violin

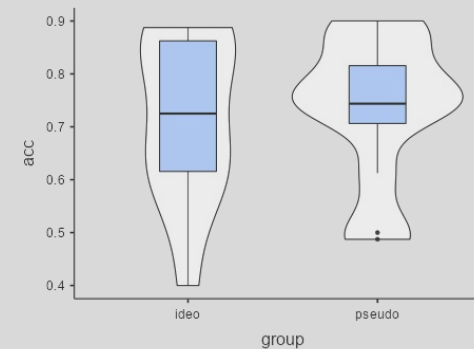
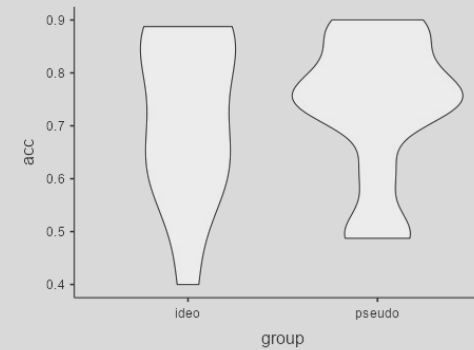
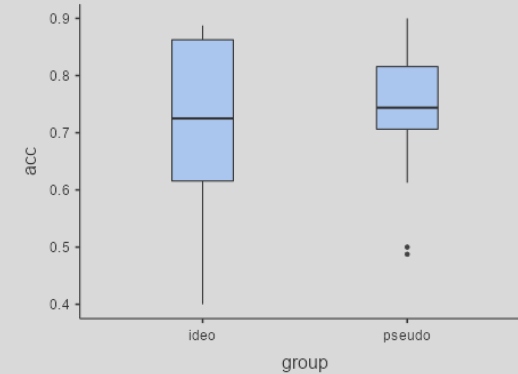
- και τα δύο

Box Plots

☒ Box plot

☐ Label outliers

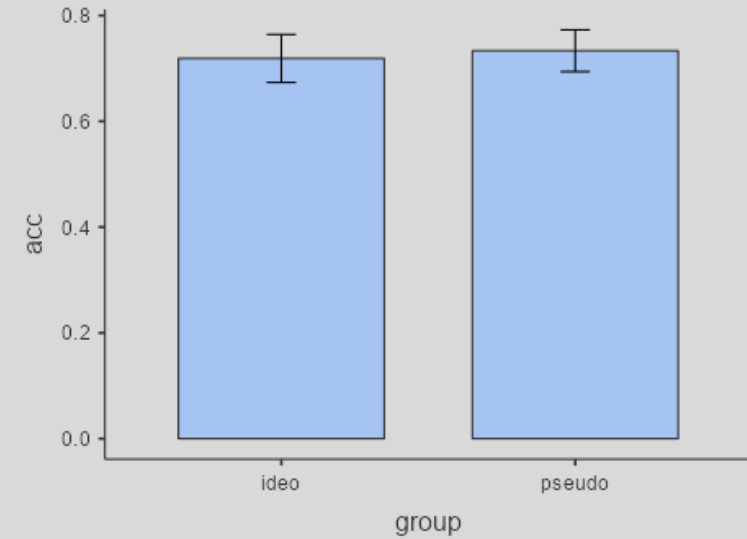
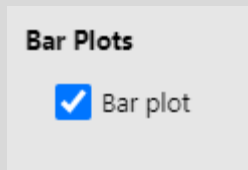
☒ Violin



Ραβδογράμματα

Μπορούμε να εμφανίσουμε είτε

- το θηκόγραμμα

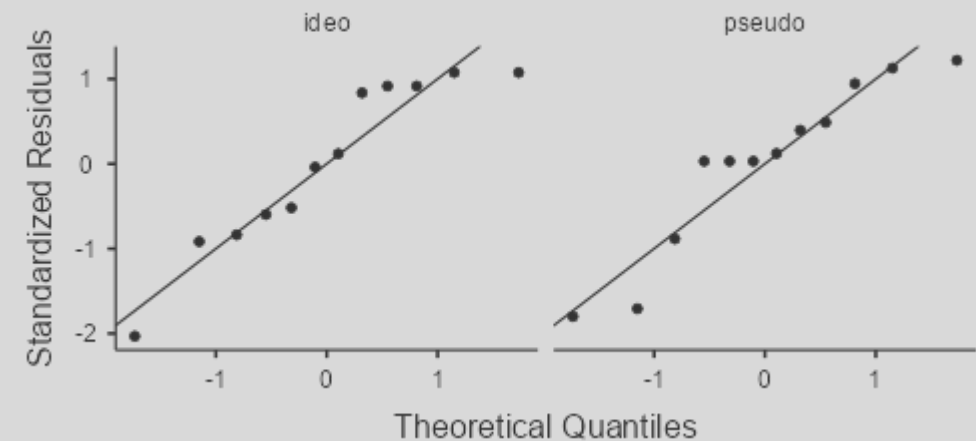


Q-Q Plots

Σημείωση: Τα Q-Q plots είναι γραφήματα που μας επιτρέπουν να εκτιμήσουμε την κανονικότητα των δεδομένων μας.

Μπορούμε να εμφανίσουμε:

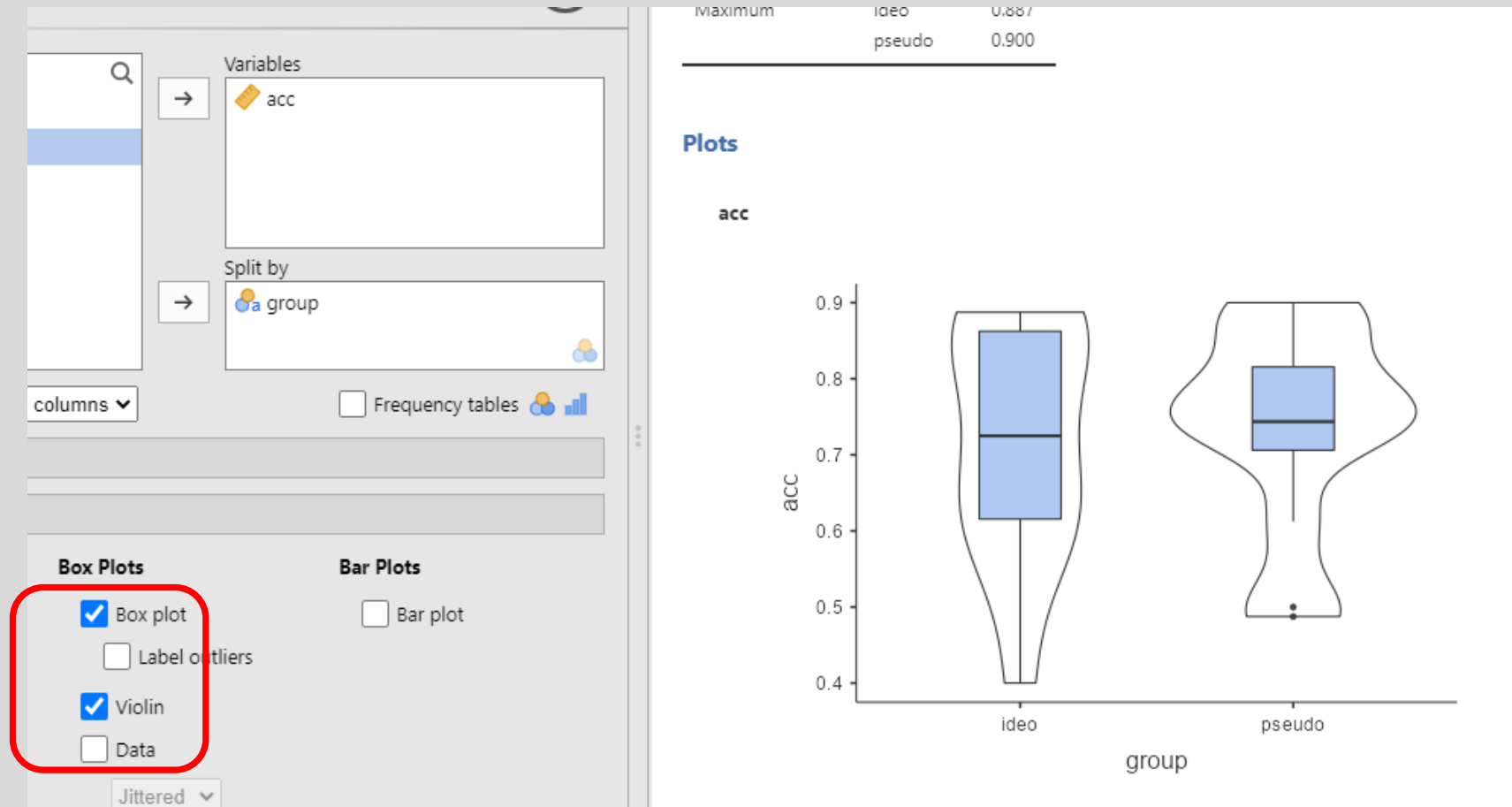
- τα Q-Q plot



2.1. Επεξεργασία Γραφημάτων

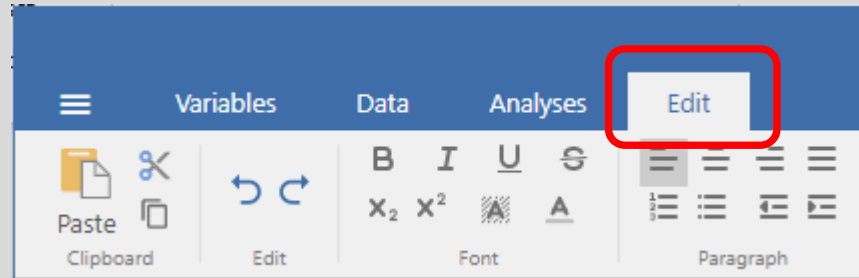
Boxplot – Violin Plot

Δημιουργούμε ξανά το θηκόγραμμα μαζί με το violin plot.

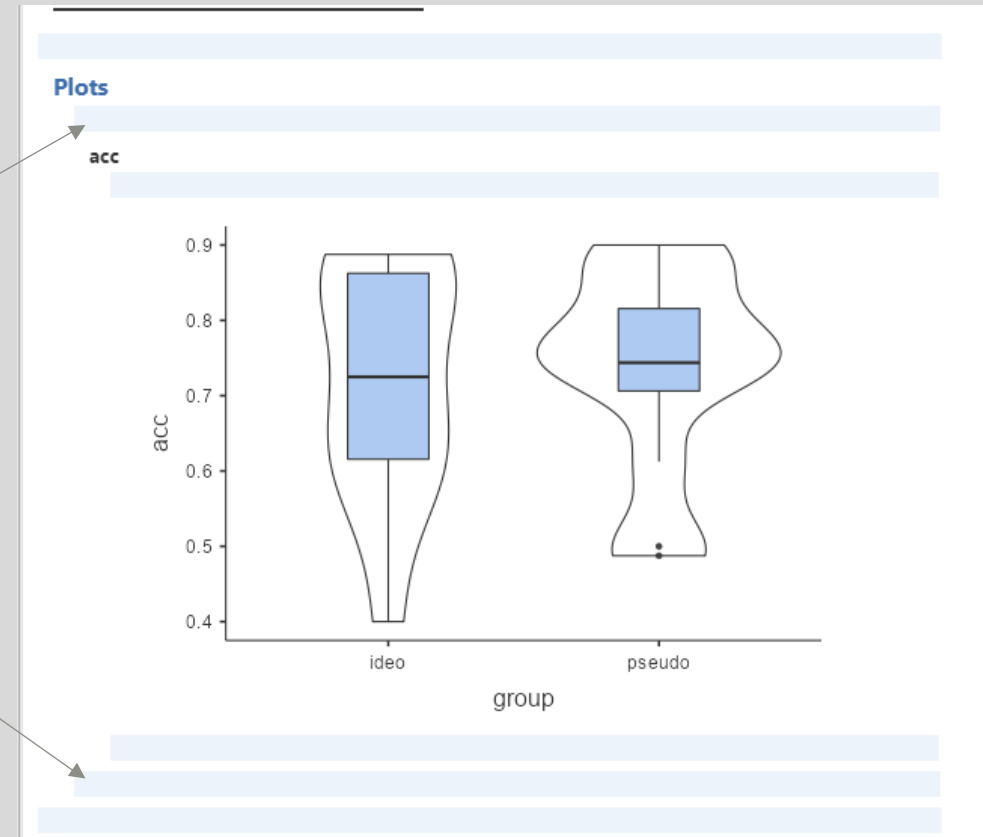


Μενού Edit

Για να επεξεργαστούμε το γράφημα, πρέπει να κάνουμε κλικ στο **μενού Edit**.



Αυτόματα εμφανίζονται τα πεδία στα οποία μπορούμε να προσθέσουμε και να μορφοποιήσουμε κείμενο.



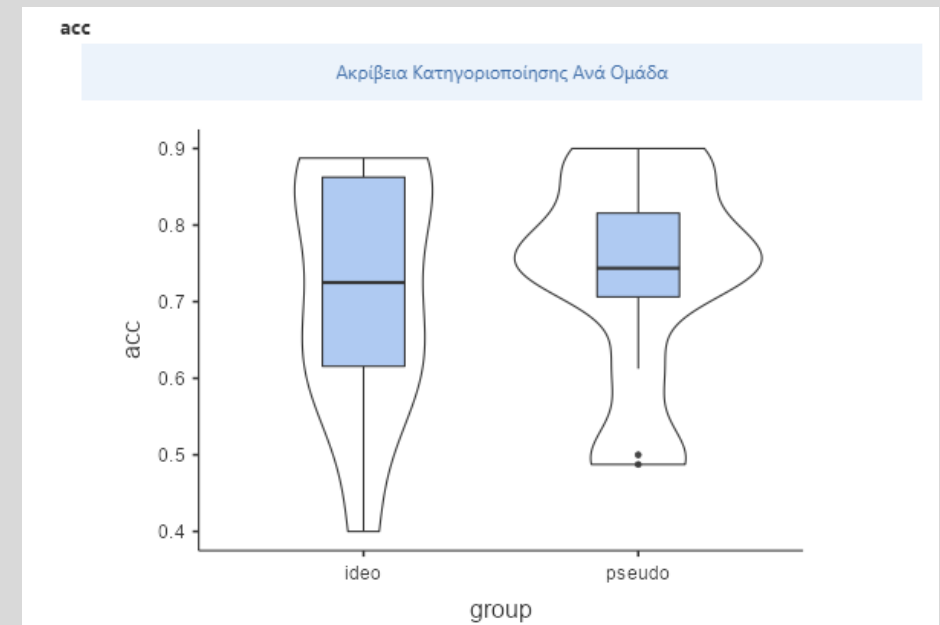
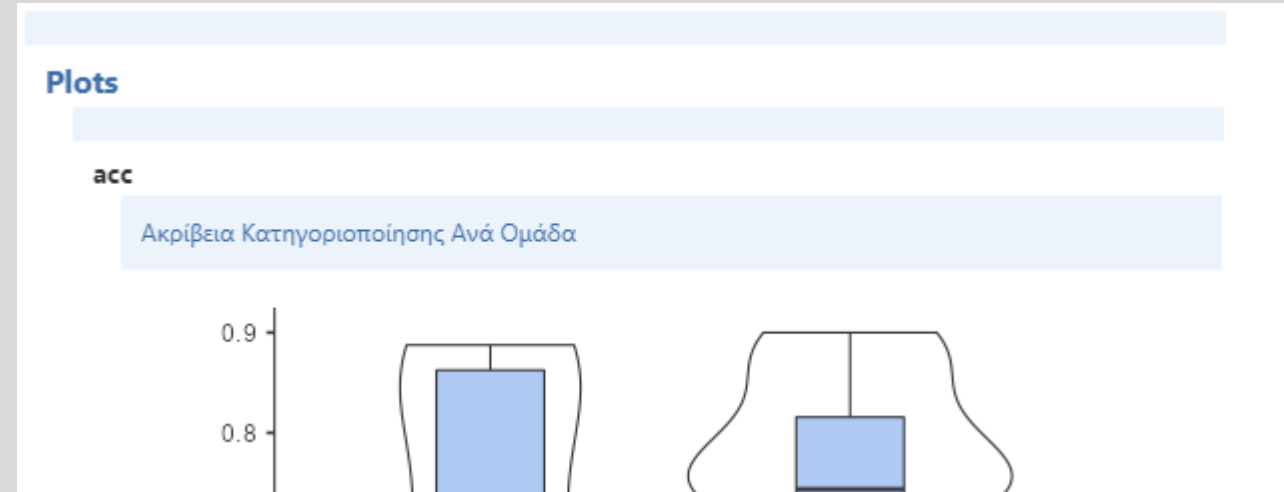
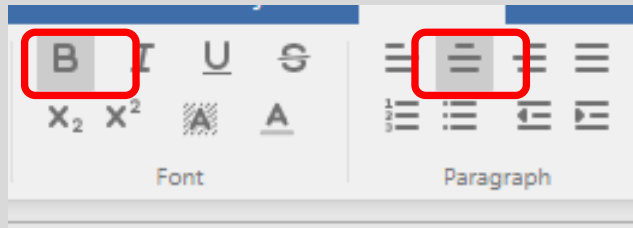
Προσθήκη Τίτλου

Μπορούμε να βάλουμε έναν τίτλο στο γράφημά μας:

«Ακρίβεια Κατηγοριοποίησης ανά Ομάδα»

Στη συνέχεια, μπορούμε με τα εργαλεία του μενού, να μορφοποιήσουμε τον τίτλο.

- Επιλέγουμε όλο τον τίτλο
- Κάνουμε τα γράμματα έντονα (B)
- Στοιχίζουμε στο κέντρο
- Όταν τελειώσουμε, κάνουμε κλικ στο μενού Analyses



Σημείωση

Δυστυχώς, δεν παρέχεται η δυνατότητα να επεξεργαστούμε άλλα στοιχεία του γραφήματος, όπως:

- χρώματα
- τίτλους και εύρος αξόνων

Επομένως, οι δυνατότητες του jamoní είναι περιορισμένες ως προς αυτό το θέμα.

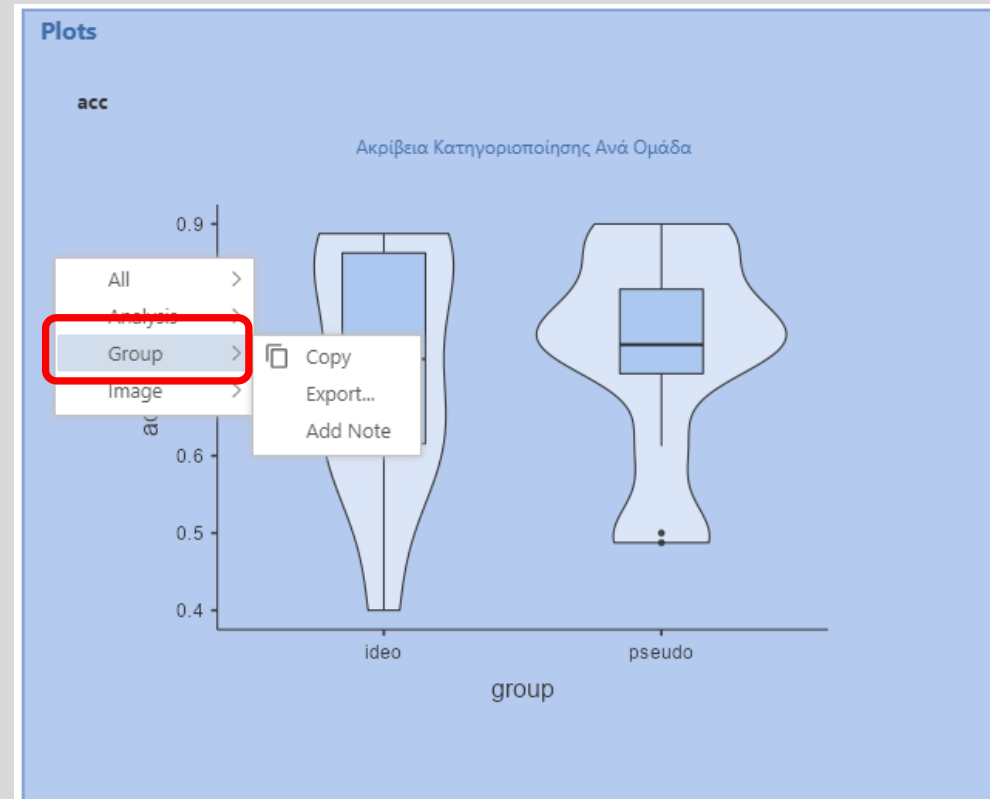
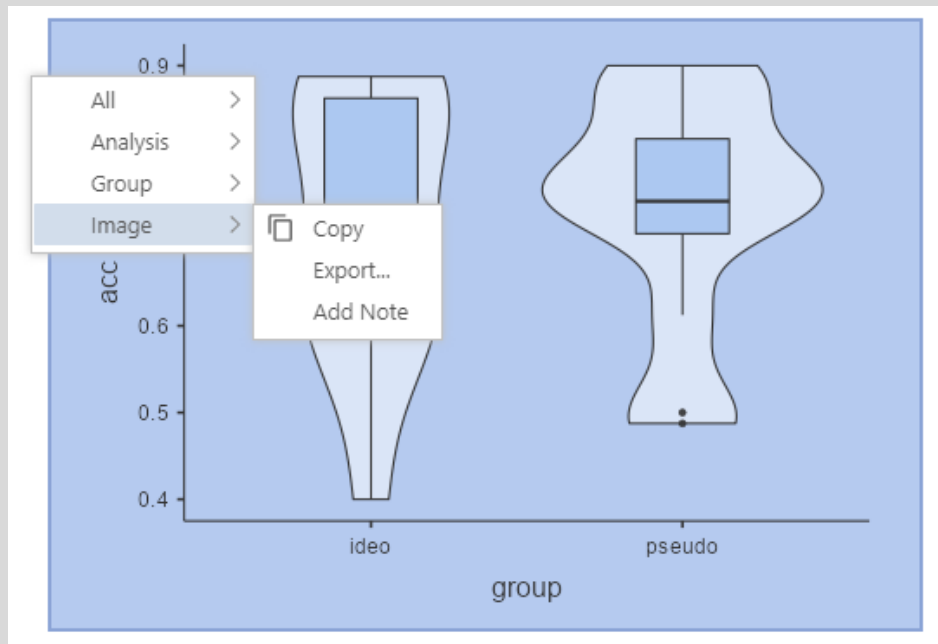
2.2. Αντιγραφή και Εξαγωγή Γραφημάτων

Επιλογή Γραφήματος

- Για να επιλέξουμε ένα γράφημα, πρέπει να κάνουμε **δεξί κλικ** πάνω του.

Παρατηρούμε πως επιλέγεται μόνο το γράφημα (όχι οι τίτλοι).

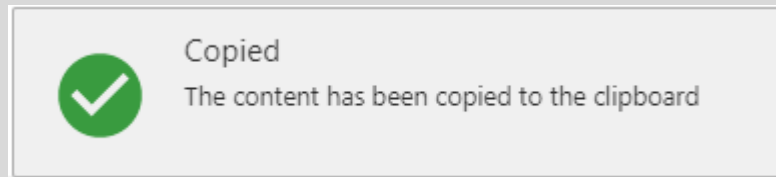
- Αν κανουμε κλικ στο Group, επιλέγεται όλο το γράφημα



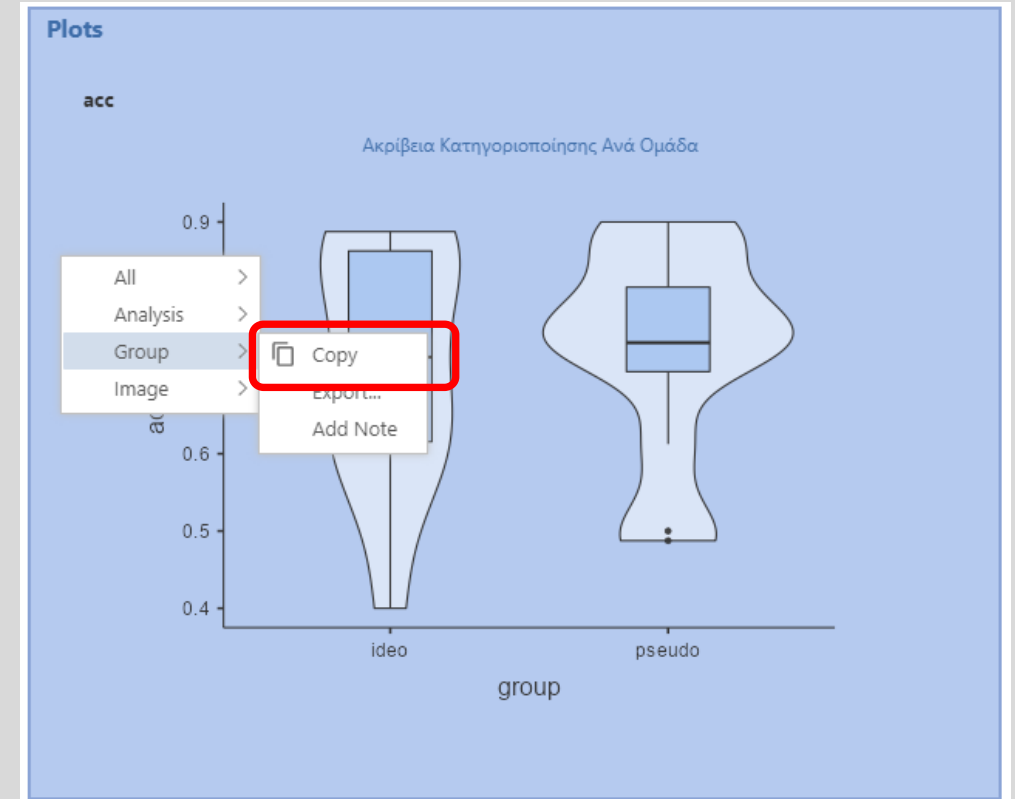
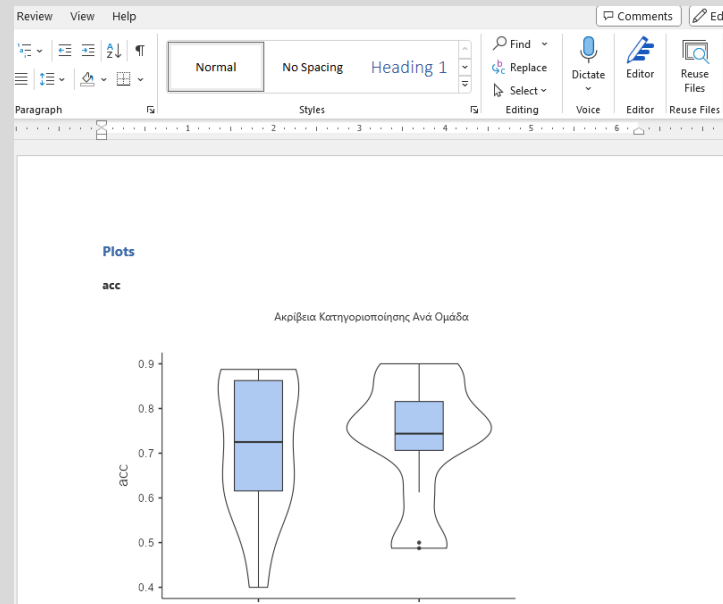
Αντιγραφή Γραφήματος

- Για να αντιγράψουμε το γράφημα, κάνουμε κλικ στην εντολή Copy

Εμφανίζεται ένα πλαίσιο:



Στη συνέχεια, μπορούμε να το επικολλήσουμε (Paste) σε ένα αρχείο Κειμένου

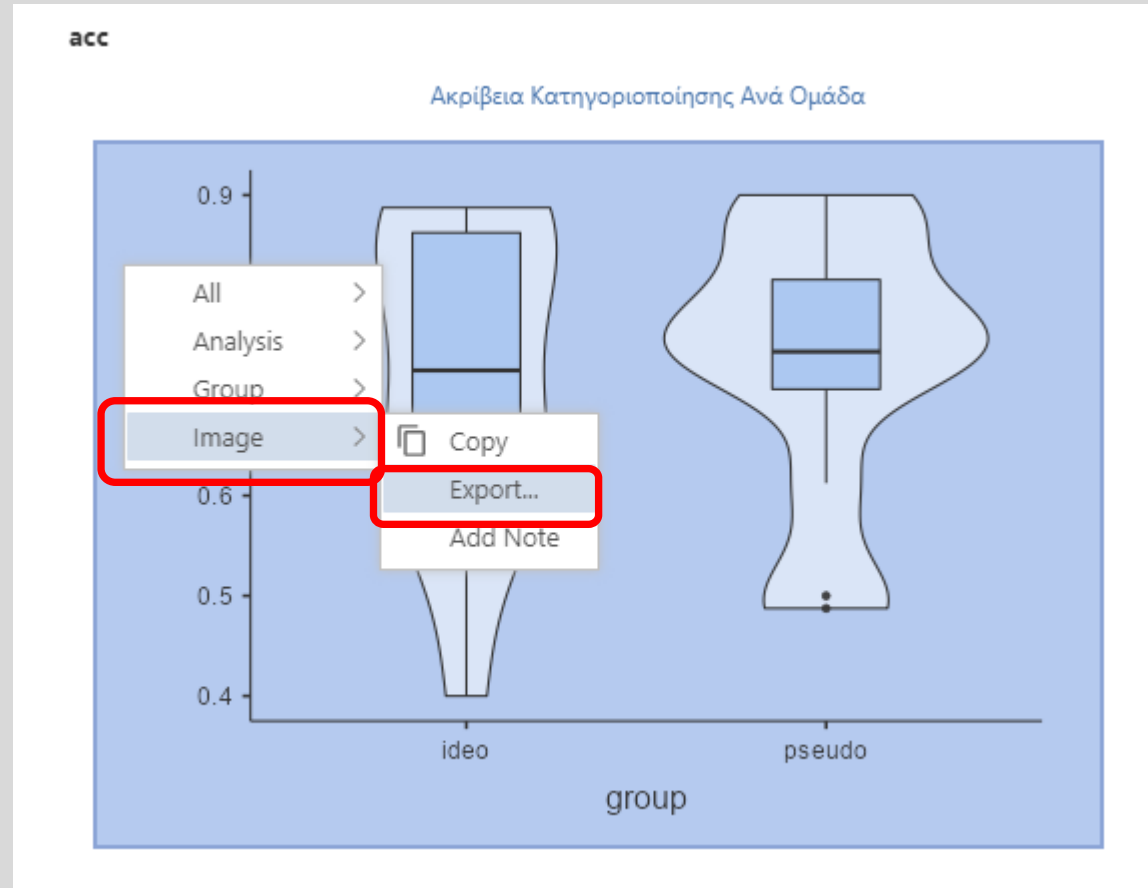


Εξαγωγή Γραφήματος

Στο jamonι έχουμε τη δυνατότητα να εξαγάγουμε ένα γράφημα σαν αρχείο εικόνας.

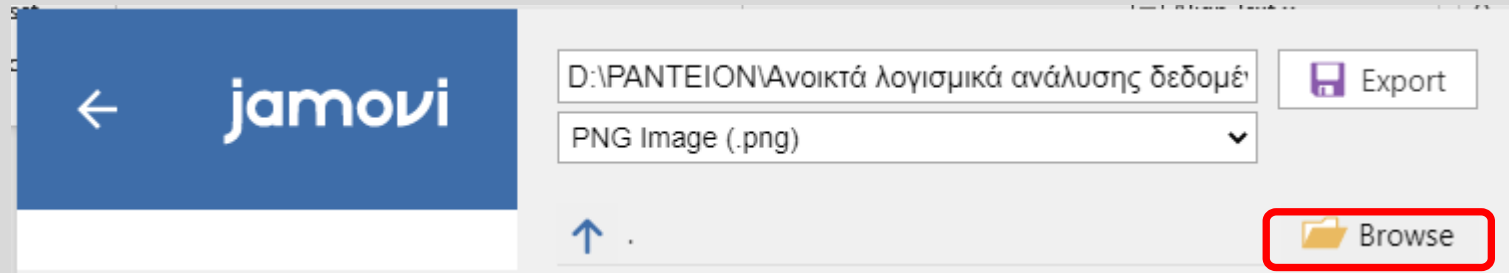
Για να το κάνουμε αυτό, πρέπει να είναι επιλεγμένο το στοιχείο Image

- Στη συνέχεια, κάνουμε κλικ στην εντολή **Export**



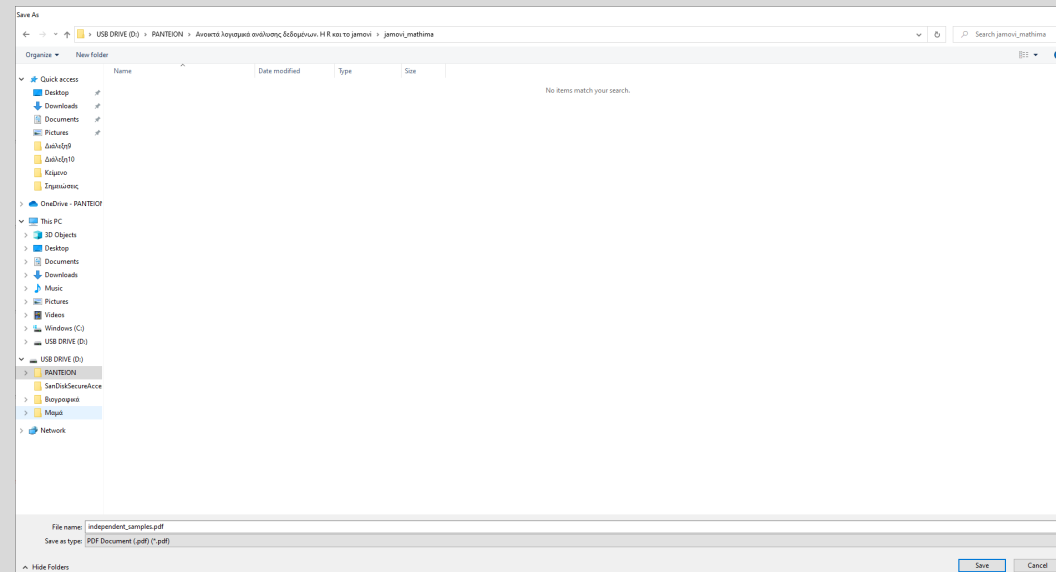
Εξαγωγή Γραφήματος II

Στο πλαίσιο που εμφανίζεται, υπάρχει η δυσκολία πως αναγράφεται η πλήρης διαδρομή του αρχείου.



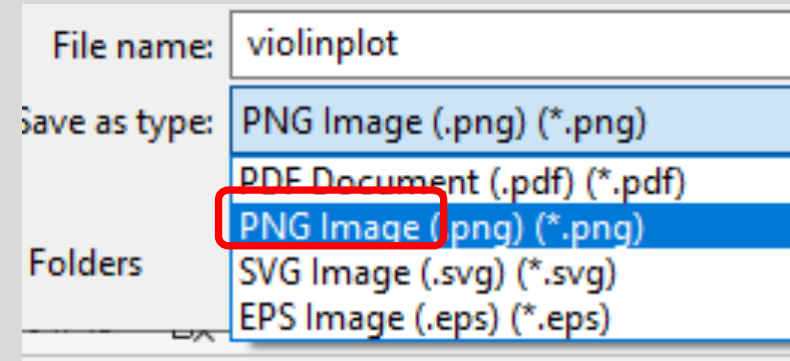
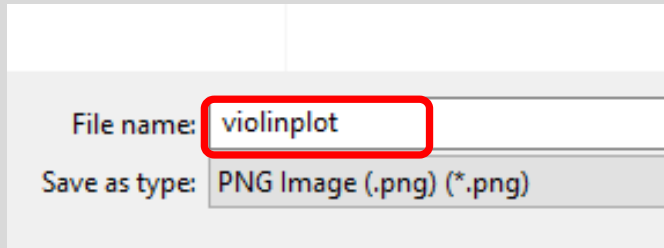
Για να προσπεράσουμε τη δυσκολία, κάνουμε κλικ κλικ στο κουμπί **Browse**

Στο παράθυρο που εμφανίζεται, είναι πιο εύκολο να επιλέξουμε φάκελο, όνομα, και τύπο αρχείου:

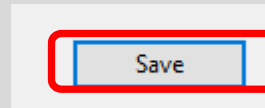


Εξαγωγή Γραφήματος III

- Επιλέγουμε τον φάκελο `jamoni_mathima`
- Δίνουμε το όνομα **violinplot**
- Επιλέγουμε είδος αρχείου: **PNG**



- Κάνουμε κλικ στο **Save**



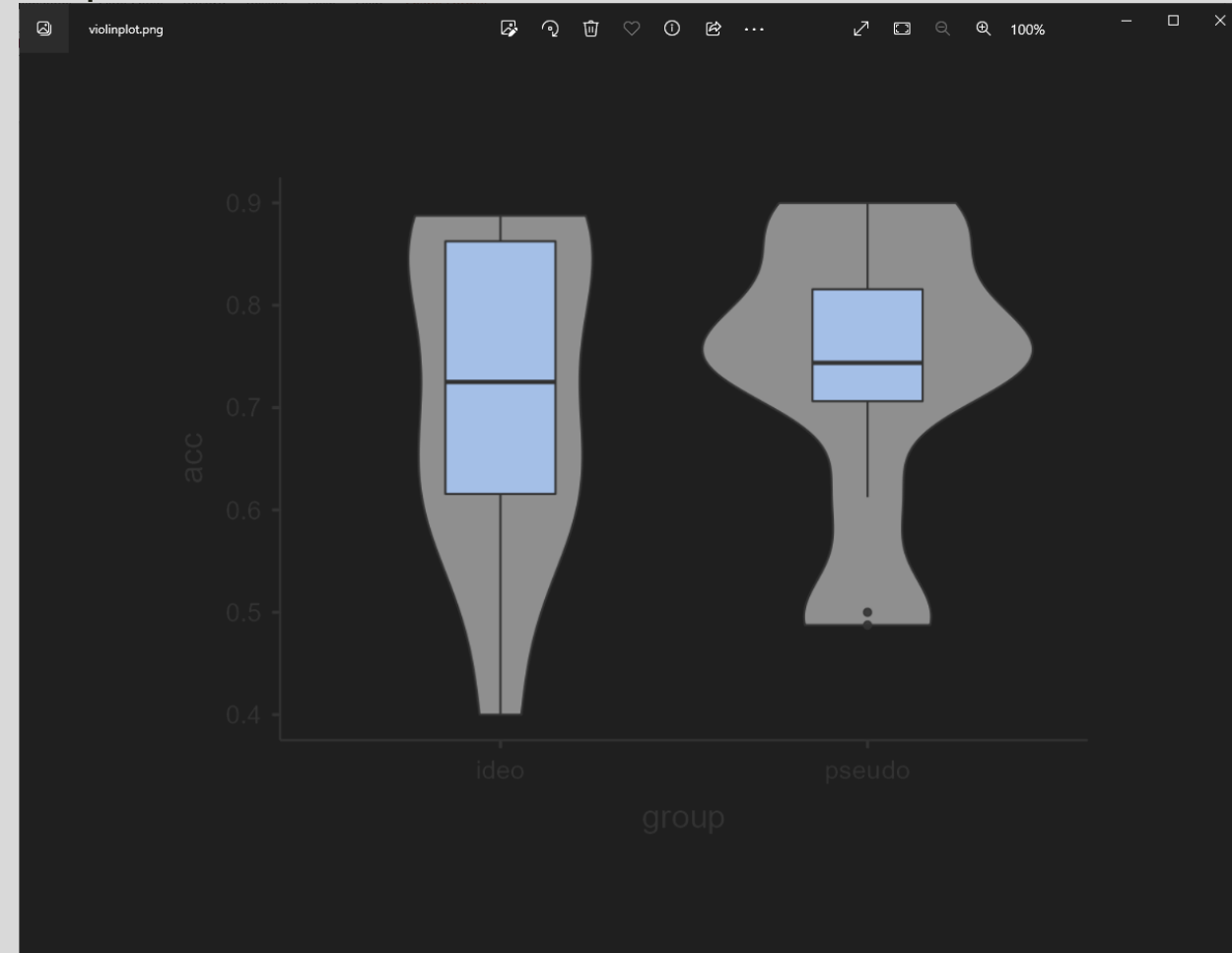
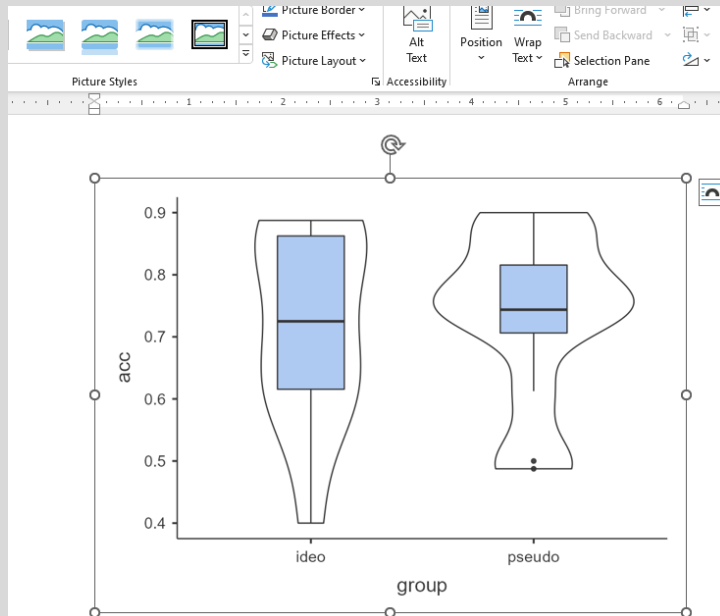
Αποθηκευμένο Γράφημα

- Εντοπίζουμε το αρχείο violinplot.png και κάνουμε διπλό κλικ επάνω του.

Στο πρόγραμμα των Windows
το γράφημα εμφανίζεται έτσι:

Αυτό οφείλεται στο ότι τα γραφήματα
του jamoni είναι **διαφανή**.

Αν το εισάγουμε σε ένα αρχείο κειμένου
δεν θα υπάρξει θέμα.



3. Έλεγχος χ^2

Το Πείραμα

Θέλουμε να ελέγξουμε την αποτελεσματικότητα ενός φαρμάκου για μια συγκεκριμένη ιατρική πάθηση. Ας υποθέσουμε ότι έχουμε 105 ασθενείς υπό μελέτη και οι 50 από αυτούς έλαβαν το φάρμακο. Επιπλέον, οι υπόλοιποι 55 ασθενείς διατηρήθηκαν ως ομάδα ελέγχου. Έτσι, η κατάσταση υγείας όλων των ασθενών ελέγχθηκε μετά από μια εβδομάδα.

Καταγράφουμε για κάθε συμμετέχοντα στη μεταβλητή **treatment** το αν έλαβαν αγωγή (**treated**) ή όχι (**not-treated**).

Επίσης, καταγράφουμε στη μεταβλητή **improvement** ο/η συμμετέχων σημείωσε βελτίωση (**improved**) ή όχι (**not-improved**).

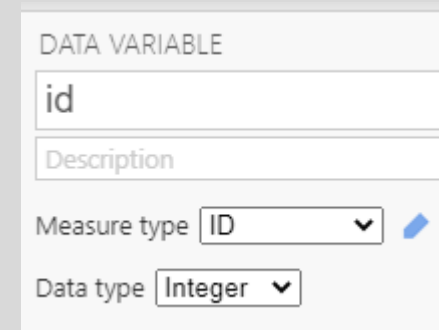
Θέλουμε να ελέγξουμε αν βελτίωση (**improvement**) σχετίζεται ή όχι με την παροχή φαρμάκου (**treatment**).

Τα Δεδομένα

- Κατεβάζουμε το αρχείο **categorical_data.csv** από το eclass (Εγγραφα > Διάλεξη 10), και το μετακινούμε στον φάκελο jamoni_mathima
- Ανοίγουμε με το jamoni (μενού File > Open) το αρχείο **categorical_data.csv**.

Όπως πάντα, ελέγχουμε τις μεταβλητές μας.

- Μετατρέπουμε την μεταβλητή **id** σε **τύπο ID**



DATA VARIABLE

id

Description

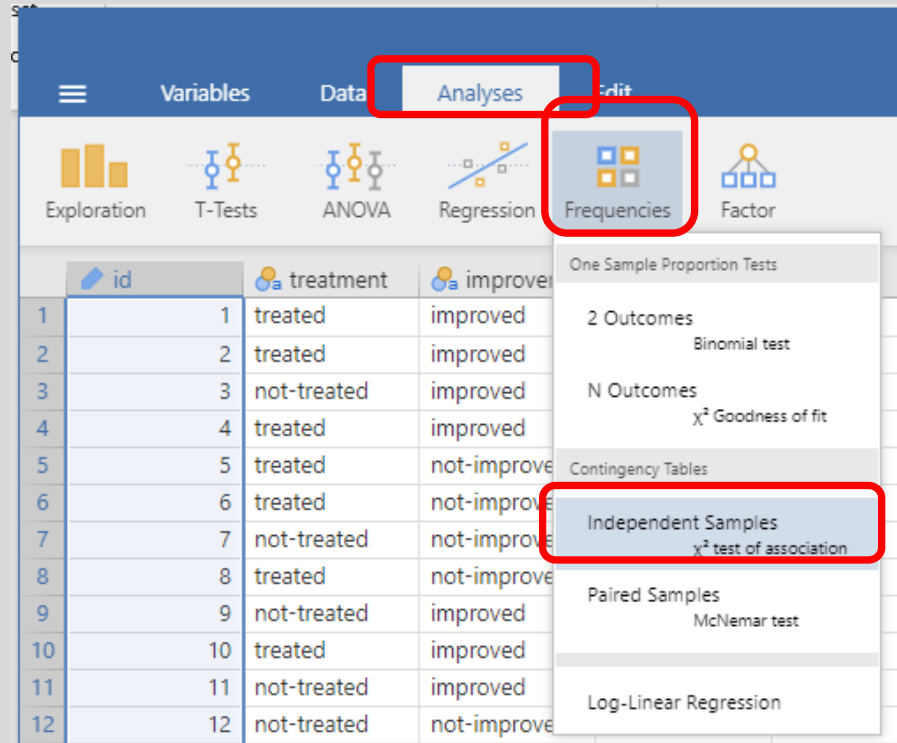
Measure type ID

Data type Integer

Analyses > Frequencies

Για να διεξάγουμε τον έλεγχο χ^2 πρέπει να

- κάνουμε κλικ στο μενού **Analyses** και μετά
- κλικ στο **Frequencies**.
- Στο πτυσσόμενο μενού που εμφανίζεται κάνουμε κλικ στην εντολή **Independent Samples**



Το παράθυρο

Εμφανίζεται το παρακάτω παράθυρο:

jamovi - categorical_data

Variables Data Analyses Edit

Exploration T-Tests ANOVA Regression Frequencies Factor Modules

Contingency Tables

treatment
improvement
id

Rows
Columns
Counts (optional)
Layers

> | Statistics
> | Cells
> | Plots

Results

Contingency Tables

Contingency Tables

			Total
.	.	.	.
.	.	.	.
Total	.	.	.

χ² Tests

	Value	df	p
χ ²	.	.	.
N	.		

References

[1] The jamovi project (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.

Πίνακας Συχνοτήτων

Πρέπει να δημιουργήσουμε τον πίνακα συχνοτήτων.

Τοποθετούμε

- στο πλαίσιο Rows την μεταβλητή treatment
- στο πλαίσιο Columns την μεταβλητή improvement

Παρατηρούμε πως άμεσα εμφανίζεται ο πίνακας συχνοτήτων, και εκτελείται ο έλεγχος χ^2 .

Σημείωση:

Τα αποτελέσματα είναι ίδια με αυτά της R.

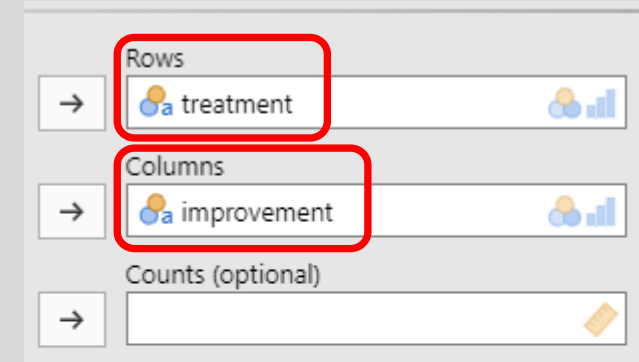
Μόνο που το jamoni κάνει την στρογγυλοποίηση για εμάς.

```
> chisq.test(table, correct = F)
```

Pearson's Chi-squared test

```
data: table
```

```
X-squared = 5.5569, df = 1, p-value = 0.01841
```



Contingency Tables

Contingency Tables

treatment	improvement		Total
	improved	not-improved	
not-treated	26	29	55
treated	35	15	50
Total	61	44	105

χ^2 Tests

	Value	df	p
χ^2	5.56	1	0.018
N	105		

Αναμενόμενες Τιμές

Για να ισχύει ο έλεγχος χ^2 , πρέπει οι αναμενόμενες τιμές (βλέπε Διάλεξη &, Ενότητα 1.1) να είναι μεγαλύτερες του 5.

Δυστυχώς, το jamonι δεν προειδοποιεί αν δεν πληρείται αυτή η προϋπόθεση.

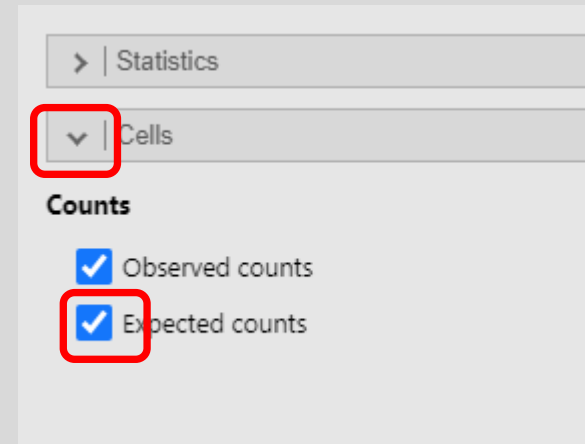
Πρέπει να ελέγξουμε τις αναμενόμενες τιμές μόνοι μας.

Για να εμφανίσουμε τις αναμενόμενες τιμές,

- κάνουμε κλικ στο Cells
- επιλέγουμε το Expected Counts

Παρατηρούμε, πως πλέον εμφανίζονται οι τιμές

Επειδή όλες οι τιμές είναι > 5 , πληρείται η προϋπόθεση.



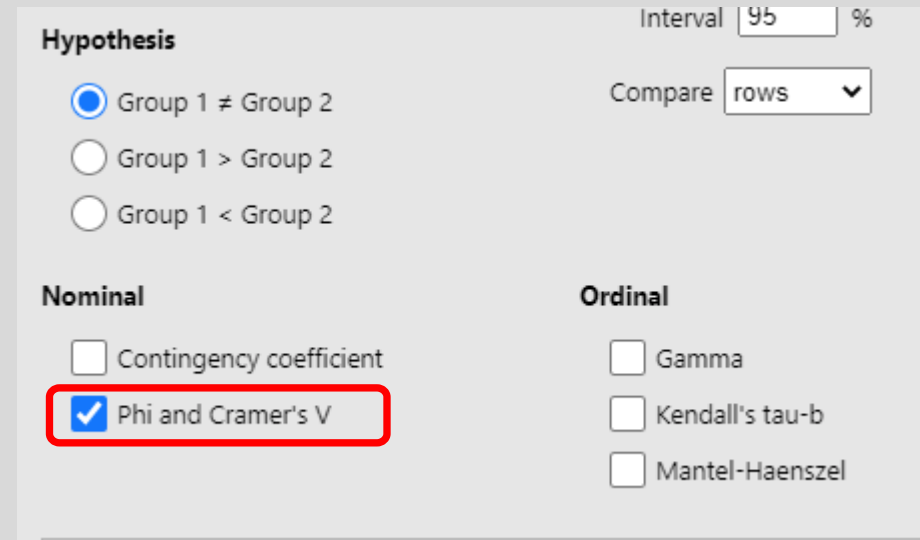
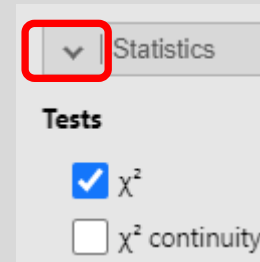
Contingency Tables				
		improvement		Total
		improved	not-improved	
not-treated	Observed	26	29	55
	Expected	32.0	23.0	55.0
treated	Observed	35	15	50
	Expected	29.0	21.0	50.0
Total	Observed	61	44	105
	Expected	61.0	44.0	105.0

Μέγεθος Επίδρασης

Το μέγεθος επίδρασης για τον έλεγχο χ^2 είναι το **Cramer's V**.

Για να υπολογίσουμε το μέγεθος αυτό:

- κάνουμε κλικ στο Statistics
- επιλέγουμε στην κατηγορία **Nominal** το **Phi and Cramer's V**



Παρατηρούμε πως εμφανίζεται ένας νέος πίνακας με το αποτέλεσμα που θέλουμε.

Σημείωση: Η τιμή είναι ίδια με αυτή της R.

```
> cramerv(table)
Cramer V
0.2301
```

Nominal	
	Value
Phi-coefficient	0.230
Cramer's V	0.230

Αναφορά κατά APA

Για την αναφορά των αποτελεσμάτων,
χρησιμοποιούμε τους πίνακες χ^2 Tests και Nominal

χ^2 Tests			
	Value	df	p
χ^2	5.56	1	0.018
N	105		

Nominal	
	Value
Phi-coefficient	0.230
Cramer's V	0.230

Γράφουμε κατά APA:

«Βρέθηκε μια σημαντική συσχέτιση μεταξύ της παροχής φαρμάκου και της βελτίωσης των ασθενών, $\chi^2(1) = 5.56$, $p = .018$, Cramer's $V = 0.23$ »

Ο ακριβής Έλεγχος του Fisher

Σε δεδομένα για τα οποία δεν πληρείται η προϋπόθεση των αναμενόμενων τιμών, ακολουθούμε τα προηγούμενα βήματα, και μπορούμε να:

- κάνουμε κλικ στο **Statistics**
- από-επιλέξουμε το χ^2
- επιλέξουμε το **Fisher's exact test**

Παρατηρούμε πως εμφανίζονται οι υπολογισμοί του ελέγχου.

Σημείωση: Οι τιμή είναι ίδια με της R.

```
> fisher.test(table)

Fisher's Exact Test for Count Data

data: table
p-value = 0.02098
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 1.449481      Inf
sample estimates:
odds ratio
      Inf
```

The screenshot shows the SPSS 'Contingency Tables' dialog box and its output. In the dialog, 'smoker' is assigned to Rows and 'athlete' to Columns. Under the 'Statistics' section, 'Fisher's exact test' is selected with a checked box. The output window displays a contingency table and the results of the Fisher's exact test.

Contingency Tables

smoker		athlete		Total
		athlete	non-athlete	
non-smoker	Observed	7	0	7
	Expected	4.50	2.50	7.00
smoker	Observed	2	5	7
	Expected	4.50	2.50	7.00
Total	Observed	9	5	14
	Expected	9.00	5.00	14.00

χ^2 Tests

	Value	p
Fisher's exact test		0.021
N	14	

Tests

- ☐ χ^2
- ☐ χ^2 continuity correction
- ☐ Likelihood ratio
- ☒ Fisher's exact test
- ☐ z test for difference in 2 proportions

Comparative Measures (2x2 only)

- ☐ Odds ratio
- ☐ Log odds ratio
- ☐ Relative risk
- ☐ Difference in proportions
- ☒ Confidence intervals

Εργασία 11

- Γράφημα Κατηγορικών Δεδομένων
- Τίτλος

