

PISA 4U

Οδηγός Χρήσης

IEA IDB Analyzer

Version 5.0

Στατιστική Ανάλυση Δεδομένων
Διεθνών Εκπαιδευτικών Αξιολογήσεων
με χρήση SPSS και R

Παραδείγματα από τη Διεθνή Μελέτη PISA 2022
Ελλάδα – Πορτογαλία



Πανεπιστήμιο
Ιωαννίνων



Εργαστήριο
Διδακτικής
& Σχολικής
Παιδαγωγικής
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ



Εθνικό
Πρόγραμμα
Ανάπτυξης
2021-2025



Τομεακό
Πρόγραμμα
Ανάπτυξης
2021 - 2025

Η εκπόνηση του παραδοτέου εντάσσεται στο Έργο «Δημιουργία Ερευνητικών Υποδομών, Επεξεργασία Ερευνητικού Υλικού και Επικοινωνία Αποτελεσμάτων Διεθνών Εκπαιδευτικών Ερευνών» με Επιστημονικό Υπεύθυνο τον Καθηγητή του Πανεπιστημίου Ιωαννίνων Αναστάσιο Εμβαλωτή. Χρηματοδοτήθηκε βάσει της Προγραμματικής Σύμβασης του Υπουργείου Παιδείας, Θρησκευμάτων και Αθλητισμού με το Πανεπιστήμιο Ιωαννίνων μέσω του Τομεακού Προγράμματος Ανάπτυξης 2021-2025.

Επιστημονικός Υπεύθυνος:

Αναστάσιος Εμβαλωτής, Καθηγητής Πανεπιστημίου Ιωαννίνων
Διευθυντής Εργαστηρίου Διδακτικής & Σχολικής Παιδαγωγικής

Επιστημονικοί Συνεργάτες:

Αφροδίτη Τσάτση, Υποψήφιος Διδάκτορας στο Πανεπιστήμιο Ιωαννίνων

Βιβλιογραφική αναφορά:

Τσάτση, Α. & Εμβαλωτής, Α. (2025)., *Οδηγός Χρήσης IEA IDB Analyser (v.6)*. Πανεπιστήμιο Ιωαννίνων.
<https://doi.org/10.5281/zenodo.19654782>.

Το υλικό του PISA4U διατίθεται με άδεια 'Creative Commons Attribution-NonCommercial-ShareAlike 4.0'.
[ελεύθερη χρήση, αναπαραγωγή, αναδιανομή, παρουσίαση και αξιοποίηση, με την προϋπόθεση να μην υπάρχει
πρόθεση εμπορικής εκμετάλλευσης. Απαιτείται αναφορά του δημιουργού ή του δικαιούχου της άδειας.
Οποιοδήποτε παράγωγο έργο μπορεί να διανεμηθεί μόνο με την ίδια ή παρόμοια άδεια]

<https://creativecommons.org/licenses/by-nc-sa/4.0/>



Πίνακας Περιεχομένων

1. Εισαγωγή στο IEA IDB Analyzer.....	4
2. Μεταφόρτωση και Εγκατάσταση.....	5
2.1 Συνοπτική Παρουσίαση της Εφαρμογής.....	5
3. Χρήση της Ενότητας Ανάλυσης (Analysis Module)	6
Υπολογισμός του συντελεστή συσχέτισης Pearson.....	7
Υπολογισμός του συντελεστή συσχέτισης Spearman	15
Υπολογισμός γραμμικής παλινδρόμησης.....	22
Υπολογισμός ποσοστών και μέσης τιμής.....	32
Υπολογισμός ποσοστών	41
Υπολογισμός Ποσοστημορίων (Percentiles).....	48
Μελέτη επιδόσεων με βάση τα ορόσημα (Benchmarks)	54
4. Ανάλυση με SPSS.....	61
Υπολογισμός του συντελεστή συσχέτισης Pearson.....	62
Υπολογισμός του συντελεστή συσχέτισης Spearman	70
Υπολογισμός γραμμικής παλινδρόμησης.....	77
Υπολογισμός λογιστικής παλινδρόμησης.....	87
Υπολογισμός ποσοστών και μέσης τιμής.....	96
Υπολογισμός ποσοστών	105
Υπολογισμός Ποσοστημορίων (Percentiles).....	112
Μελέτη επιδόσεων με βάση τα ορόσημα (Benchmarks)	118

1. Εισαγωγή στο IEA IDB Analyzer

Ο παρών οδηγός απευθύνεται σε ερευνητές και εκπαιδευτικούς που επιθυμούν να αξιοποιήσουν τα δεδομένα διεθνών εκπαιδευτικών αξιολογήσεων. Παρουσιάζεται η πλήρης διαδικασία ανάλυσης με χρήση της εφαρμογής IEA IDB Analyzer v5.0, τόσο μέσω της γλώσσας R όσο και μέσω του SPSS.

2. Μεταφόρτωση και Εγκατάσταση

2.1 Συνοπτική Παρουσίαση της Εφαρμογής

Το IEA IDB Analyzer είναι μία εφαρμογή της IEA (International Association for the Evaluation of Educational Achievement), δωρεάν και σχεδιασμένη για να αναλύει δεδομένα από μεγάλες διεθνείς εκπαιδευτικές αξιολογήσεις (όπως TIMSS, PIRLS, PISA κ.ά.).

Το πρόγραμμα αποτελείται από δύο βασικές ενότητες. Η πρώτη είναι η Merge Module, για συνένωση αρχείων από διαφορετικές χώρες ή επίπεδα (μαθητές, σχολεία, δασκάλους) και η δεύτερη είναι η Analysis Module για στατιστικές αναλύσεις. Στην Version 5.0 της εφαρμογής, στην οποία βασίζεται ο οδηγός, έχει προστεθεί και η Convert Module που μετατρέπει SPSS (.sav) σε RData, όταν ο χρήστης επιθυμεί να δουλέψει με την γλώσσα προγραμματισμού R. Επιπλέον το πρόγραμμα δημιουργεί κώδικα σε SPSS, SAS ή R με ενσωματωμένη τη μεθοδολογία δειγματοληπτικών βαρών και Plausible Values. Φυσικά, απαραίτητη προϋπόθεση για χρήση της εφαρμογής είναι να είναι εγκατεστημένο το SPSS, το SAS, ή η R στον υπολογιστή. Το εγχειρίδιο χρησιμοποιεί παραδείγματα από τη διεθνή μελέτη PISA (Programme for International Student Assessment), όπου υποστηρίζεται η Ανάλυση (Analysis) και δίνεται η δυνατότητα υλοποίησής της σε SPSS/SAS/R. Να σημειωθεί ότι στον συγκεκριμένο οδηγό η παραγωγή των αρχείων θα γίνει με το SPSS και την R. Τέλος απαραίτητη προϋπόθεση για να γίνει η παραγωγή αρχείων syntax στο SPSS και scripts στην R είναι να είναι ήδη εγκατεστημένο το SPSS και η R (προαιρετικά και το RStudio όπως χρησιμοποιείται και στον οδηγό παρακάτω).

3. Χρήση της Ενότητας Ανάλυσης (Analysis Module)

Το Analysis Module, πιο αναλυτικά, προσφέρει ένα ολοκληρωμένο σύνολο στατιστικών διαδικασιών, επιτρέποντας τον υπολογισμό μέσων όρων, ποσοστών, τυπικών αποκλίσεων, συσχετίσεων και συντελεστών παλινδρόμησης. Παρέχει τη δυνατότητα εκτέλεσης αναλύσεων τόσο σε εθνικό επίπεδο όσο και για συγκεκριμένες υποομάδες του πληθυσμού, όπως το φύλο ή οι κοινωνικοοικονομικές κατηγορίες. Επιπλέον, υπολογίζει με ακρίβεια τα ποσοστά μαθητών που βρίσκονται εντός, πάνω ή κάτω από καθορισμένα επίπεδα επίδοσης (benchmarks), καθώς και ποσοστημόρια (percentiles) για την επίδοση ή οποιαδήποτε άλλη συνεχής μεταβλητή. Ένα σημαντικό πλεονέκτημα του Module είναι ότι δημιουργεί αυτόματα το απαραίτητο αρχείο εντολών (syntax) σε SPSS, SAS ή R, ενσωματώνοντας ορθά τα sampling weights και τα replicate weights, εξασφαλίζοντας έτσι αναλύσεις συμβατές με τη μεθοδολογία των διεθνών αξιολογήσεων.

Ακολουθεί η υλοποίηση των στατιστικών διαδικασιών αρχικά με χρήση της R και ύστερα με τη χρήση του SPSS. Στο πλαίσιο της ανάλυσης, έχουν επιλεγεί μεταβλητές από τη βάση του PISA με τις χώρες ενδιαφέροντος να είναι η Ελλάδα και η Πορτογαλία. Τέλος να σημειωθεί ότι η στατιστική διαδικασία της λογιστικής παλινδρόμησης δεν μπορεί να εκτελεστεί μέσω της R ως επιλογή ανάλυσης.

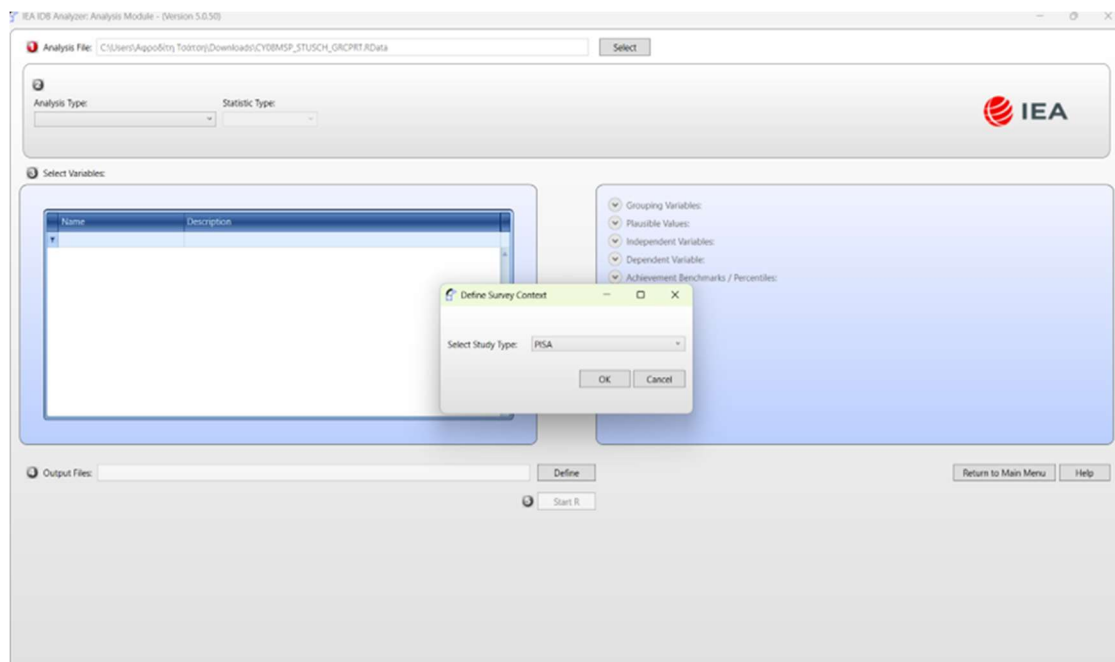
Υπολογισμός του συντελεστή συσχέτισης Pearson

Ερευνητική Υπόθεση: Σκοπός της ανάλυσης συσχέτισης Pearson είναι να διερευνηθεί αν και σε ποιο βαθμό υπάρχει στατιστικά σημαντική γραμμική σχέση μεταξύ της αυτο-αποτελεσματικότητας στα Μαθηματικά και του άγχους για τα Μαθηματικά.

Επιλογή μεταβλητών: MATHEFF, ANXMAT

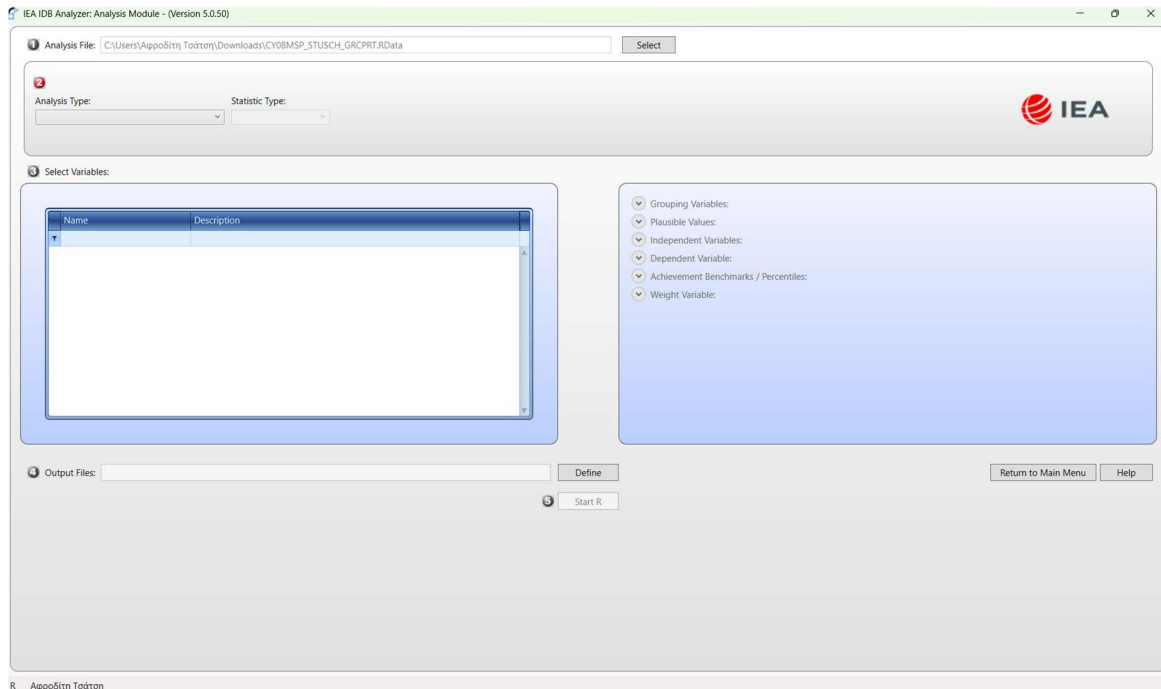
Χαρακτηριστικά μεταβλητών: MATHEFF συνεχής μεταβλητή, ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση συσχέτισης Pearson μεταξύ της μεταβλητής MATHEFF (δείκτης που μετρά την αυτο-αποτελεσματικότητα των μαθητών στα Μαθηματικά) και της μεταβλητής ANXMAT, η οποία αποτυπώνει το άγχος των μαθητών απέναντι στα Μαθηματικά. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPR1.RData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 1).



Εικόνα 1. Επιλογή τύπου και αρχείου δεδομένων PISA στο πεδίο «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 2).

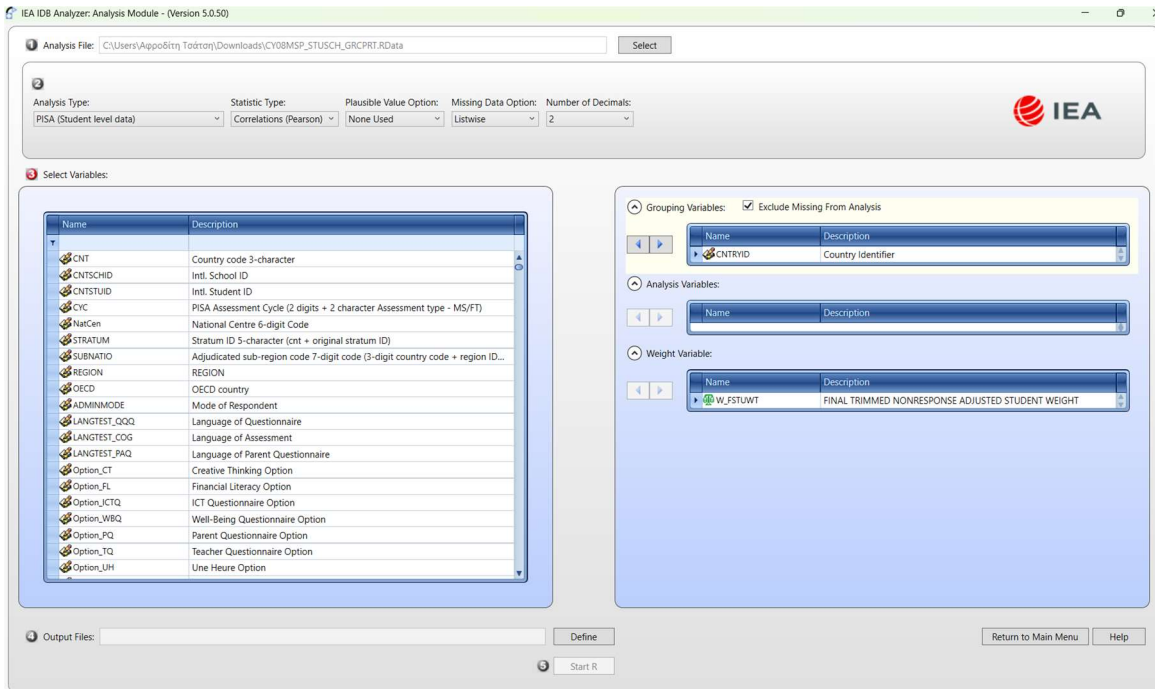


Εικόνα 2. Επιλογή επιπέδου ανάλυσης «PISA (Student level)» στο πεδίο «Analysis Type» (R)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

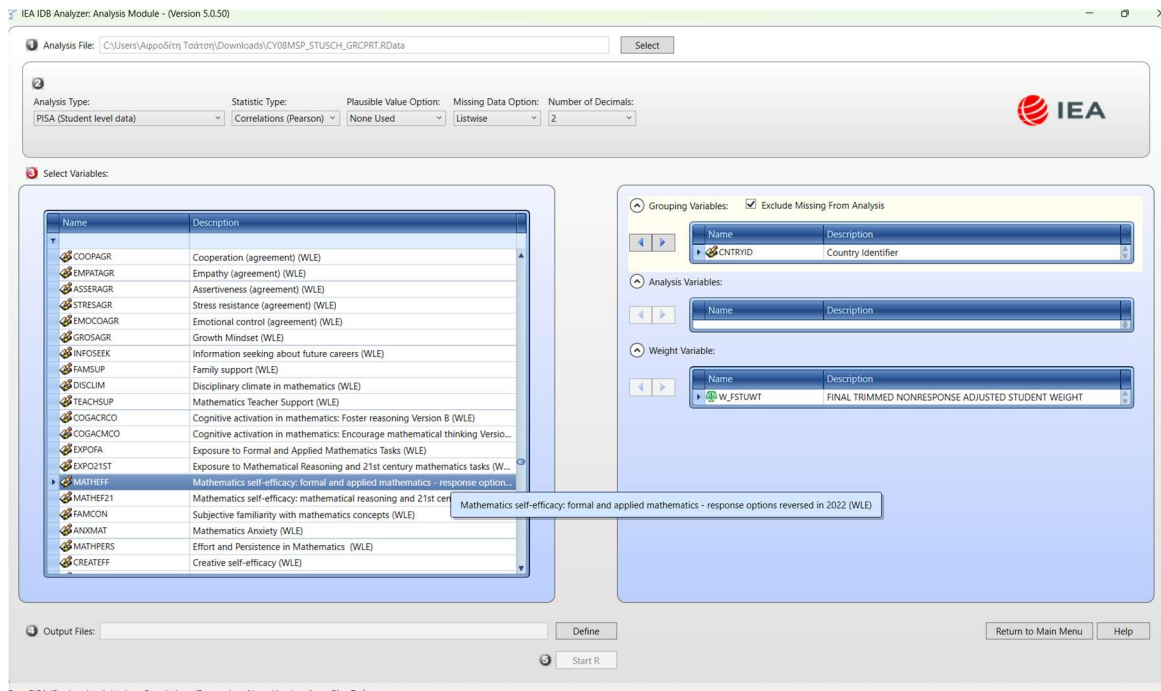
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Correlation (Pearson)».



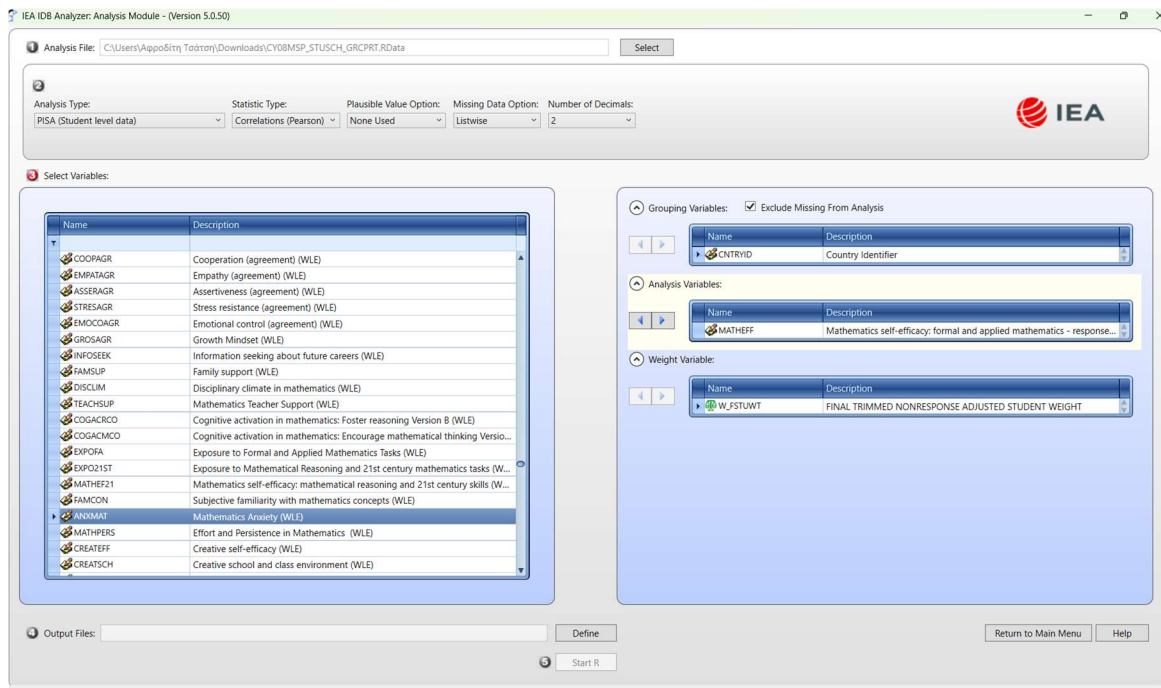
Εικόνα 3. Επιλογή τύπου στατιστικής ανάλυσης «Correlation (Pearson)» στο πεδίο «Statistic Type» (R)

Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Missing Data Option» υπάρχουν διαθέσιμες οι επιλογές «Pairwise», που αν την επιλέξετε χρησιμοποιούνται όλα τα διαθέσιμα δεδομένα για ανάλυση και «Listwise» που αν την επιλέξετε χρησιμοποιούνται μόνο οι περιπτώσεις που έχουν πλήρη δεδομένα. Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων (Εικόνα 3).



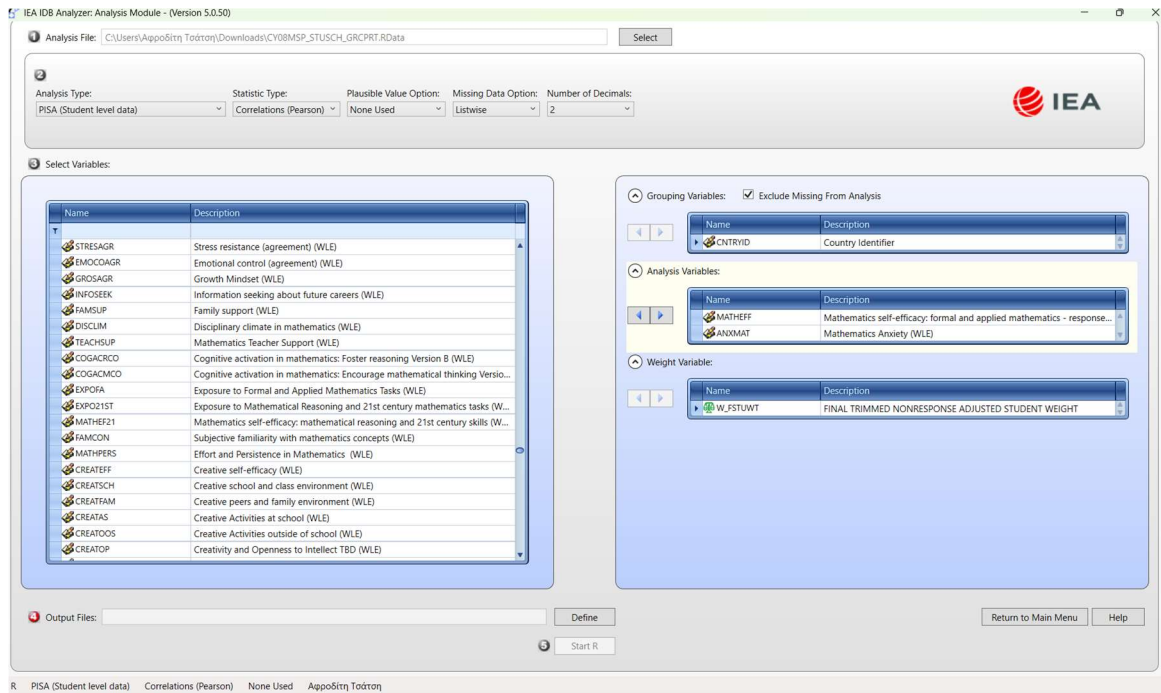
Εικόνα 4. Επισκόπηση των πεδίων ρύθμισης της ανάλυσης Pearson – «Grouping Variables» και «Missing Data Option» (R)

Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 4 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλειπείς τιμές στις μεταβλητές ομαδοποίησης. Στην συνέχεια μας ενδιαφέρει να υπολογίσουμε τον συντελεστή συσχέτισης Pearson μεταξύ της μεταβλητής «MATHEFF», και της «ANXMAT». Επομένως βρίσκετε αρχικά την μεταβλητή «MATHEFF» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» την τοποθετείτε. Επαναλαμβάνετε την διαδικασία για την μεταβλητή «ANXMAT». Τέλος στο πλαίσιο «Weight Variable» υπάρχει ως προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 5).



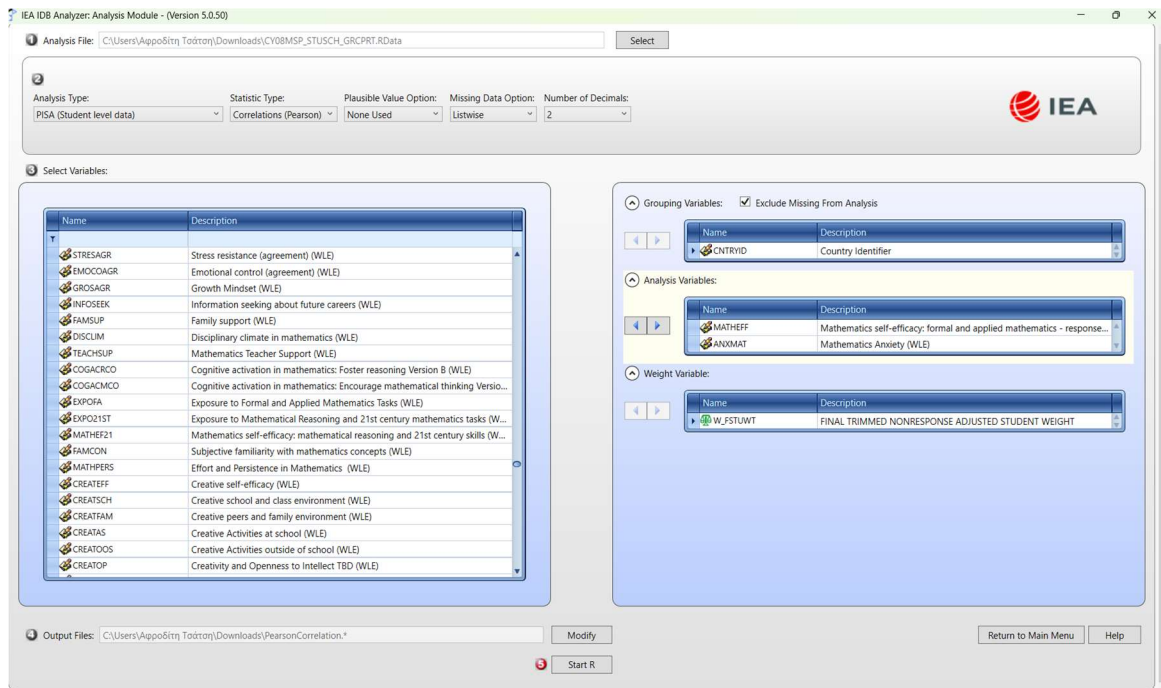
Εικόνα 5. Τοποθέτηση μεταβλητών MATHEFF και ANXMAT στο πεδίο «Analysis Variables» (R)

Επόμενο βήμα, Εικόνα 6, είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PearsonCorrelation».



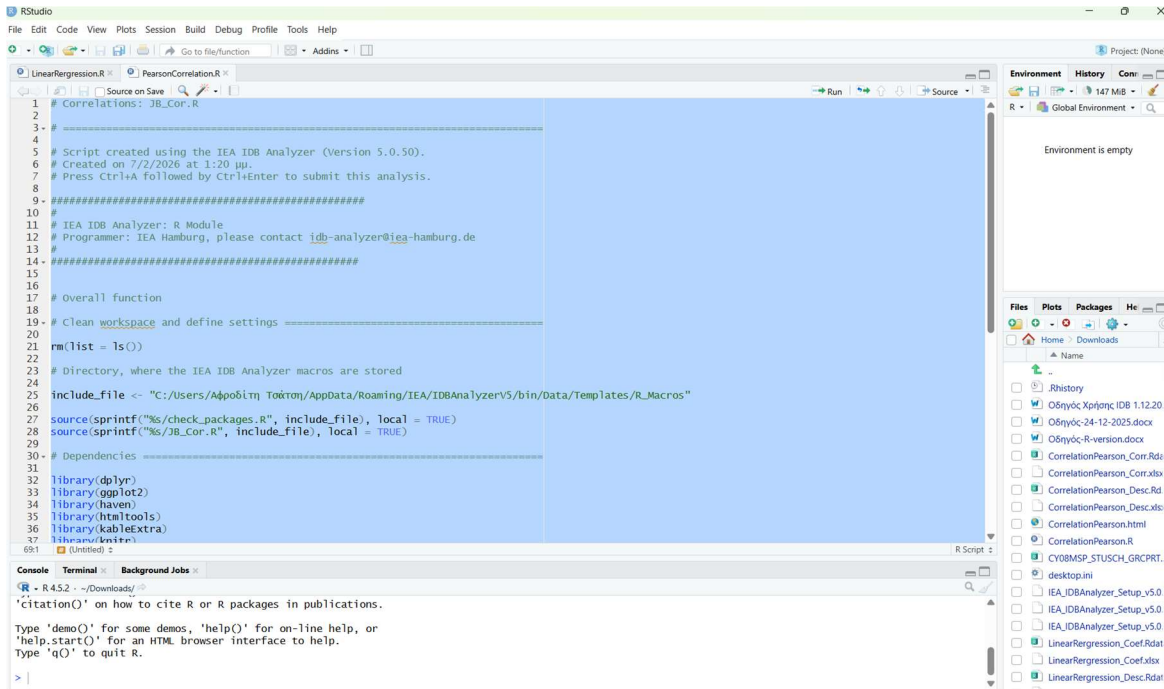
Εικόνα 6. Ορισμός ονόματος αρχείων εξόδου (Output Files) για την ανάλυση Pearson (R)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio (Εικόνα 7).



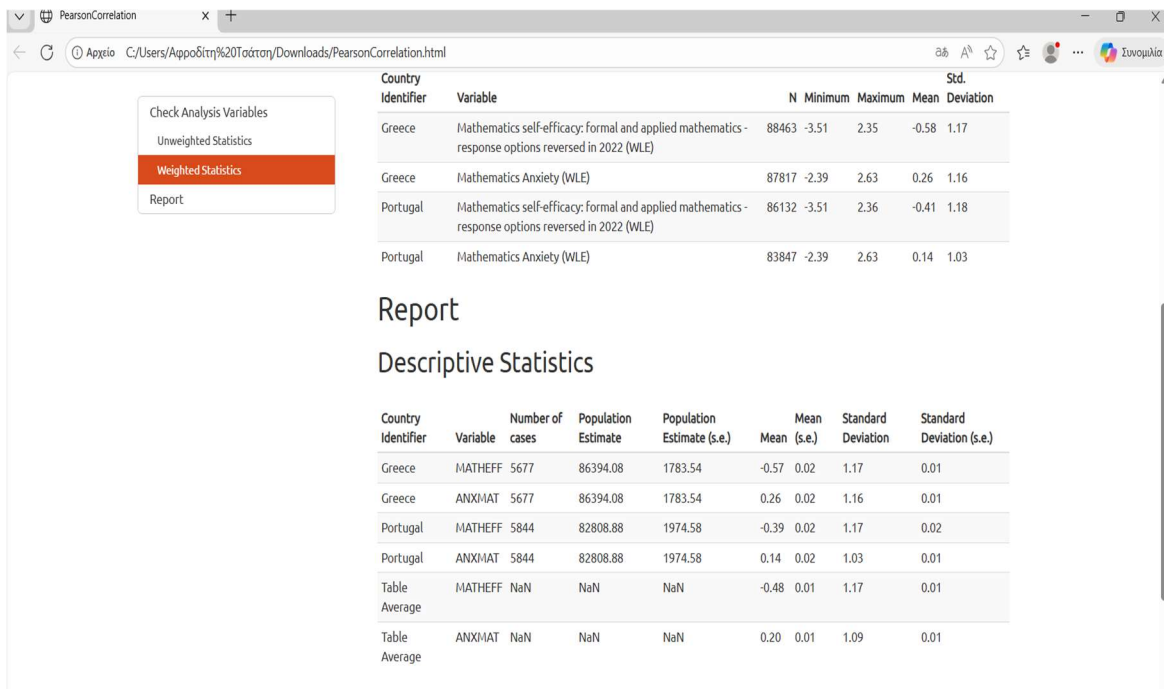
Εικόνα 7. Εκκίνηση παραγωγής script αρχείου μέσω του κουμπιού «Start R»

Αφού μεταβείτε στο περιβάλλον του Rstudio, στο αρχείο PearsonCorrelation.R που εμφανίζεται, με Ctrl+A επιλέξετε όλο το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 8).



Εικόνα 8. Εκτέλεση κώδικα *PearsonCorrelation.R* στο περιβάλλον *RStudio* (*Ctrl+A* → *Run*)

Στο αρχείο *PearsonCorrelation*, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 9 και 10, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο *RStudio* και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



Εικόνα 9. Αποτελέσματα περιγραφικής στατιστικής (*N*, *Mean*, *SD*) ανά χώρα – Ανάλυση *Pearson* (*R*)

Οι στήλες που φαίνονται στην Εικόνα 9, περιγράφουν τα εξής:

- **Country Identifier:** Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- **Variable:** Το όνομα της μεταβλητής (π.χ. ANXMAT).
- **Number of cases:** Ο αριθμός των μαθητών στο δείγμα της χώρας που απάντησαν στη συγκεκριμένη ερώτηση-μεταβλητή.

- Population Estimate: Εκτίμηση του συνολικού πληθυσμού που αντιπροσωπεύει το δείγμα σε εθνικό επίπεδο.
- Population Estimate (s.e.): Το τυπικό σφάλμα της εκτίμησης του πληθυσμού το οποίο δείχνει την ακρίβεια της εκτίμησης.
- Mean: Ο μέσος όρος των τιμών της μεταβλητής στο δείγμα της χώρας.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Standard Deviation: Η τυπική απόκλιση η οποία δείχνει πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο.
- Standard Deviation (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης το οποίο δείχνει πόσο αξιόπιστη είναι η εκτίμηση της διασποράς.

The screenshot shows the PearsonCorrelation software interface. On the left, there is a navigation menu with options: 'Check Analysis Variables', 'Report', 'Descriptive Statistics' (highlighted in orange), and 'Correlation Coefficients'. The main area displays two tables.

Table 1: Descriptive Statistics

Country Identifier	Variable	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)
Greece	MATHEFF	5677	86394.08	1783.54	-0.57	0.02	1.17	0.01
Greece	ANXMAT	5677	86394.08	1783.54	0.26	0.02	1.16	0.01
Portugal	MATHEFF	5844	82808.88	1974.58	-0.39	0.02	1.17	0.02
Portugal	ANXMAT	5844	82808.88	1974.58	0.14	0.02	1.03	0.01
Table Average	MATHEFF	NaN	NaN	NaN	-0.48	0.01	1.17	0.01
Table Average	ANXMAT	NaN	NaN	NaN	0.20	0.01	1.09	0.01

Table 2: Correlation Coefficients

Country Identifier	Variable	MATHEFF	ANXMAT	MATHEFF.se	ANXMAT.se
Greece	MATHEFF	1.00	-0.40	0.00	0.02
Greece	ANXMAT	-0.40	1.00	0.02	0.00
Portugal	MATHEFF	1.00	-0.33	0.00	0.02
Portugal	ANXMAT	-0.33	1.00	0.02	0.00
Table Average	ANXMAT	-0.36	1.00	0.01	0.00
Table Average	MATHEFF	1.00	-0.36	0.00	0.01

Εικόνα 10. Πίνακας συντελεστών συσχέτισης Pearson μεταξύ MATHEFF και ANXMAT ανά χώρα (R)

Οι στήλες που φαίνονται στην Εικόνα 10, περιγράφουν τα εξής:

- Country Identifier: Αναφέρεται στη χώρα από την οποία προέρχονται τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία). Όλες οι πληροφορίες που περιέχονται στις αντίστοιχες γραμμές του πίνακα αφορούν αποκλειστικά την εκάστοτε χώρα.
- Variable: Δείχνει τη μεταβλητή στην οποία αναφέρεται ο συντελεστής συσχέτισης.
- MATHEFF: Ο συντελεστής συσχέτισης της μεταβλητής που αναφέρεται στη στήλη "Variable" με τη μεταβλητή MATHEFF.
- ANXMAT: Ο συντελεστής συσχέτισης της μεταβλητής που αναφέρεται στη στήλη "Variable" με τη μεταβλητή ANXMAT. Η συσχέτιση είναι συμμετρική, επομένως η τιμή μεταξύ των MATHEFF και ANXMAT είναι ίδια ανεξάρτητα από τη σειρά τους.
- MATHEFF (s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής MATHEFF. Δείχνει την ακρίβεια της εκτίμησης.
- ANXMAT(s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής ANXMAT. Δείχνει την ακρίβεια της εκτίμησης.

Ερμηνεία αποτελεσμάτων

Πραγματοποιήθηκε ανάλυση συσχέτισης Pearson προκειμένου να διερευνηθεί η σχέση μεταξύ της αυτο-αποτελεσματικότητας στα Μαθηματικά (MATHEFF) και του άγχους για τα Μαθηματικά (ANXMAT) στην Ελλάδα και την Πορτογαλία. Στην Ελλάδα, οι μαθητές παρουσίασαν χαμηλότερα επίπεδα αυτο-

αποτελεσματικότητας στα Μαθηματικά ($M = -0.57$, $SD = 1.17$) και υψηλότερα επίπεδα άγχους για τα Μαθηματικά ($M = 0.26$, $SD = 1.16$). Αντίστοιχα, στην Πορτογαλία, ο μέσος όρος της αυτο-αποτελεσματικότητας στα Μαθηματικά ήταν υψηλότερος ($M = -0.39$, $SD = 1.17$), ενώ τα επίπεδα άγχους για τα Μαθηματικά ήταν χαμηλότερα ($M = 0.14$, $SD = 1.03$). Η ανάλυση συσχέτισης έδειξε ότι υπάρχει στατιστικά σημαντική αρνητική γραμμική συσχέτιση μεταξύ της αυτο-αποτελεσματικότητας και του άγχους για τα Μαθηματικά και στις δύο χώρες. Συγκεκριμένα, στην Ελλάδα παρατηρήθηκε μέτρια αρνητική συσχέτιση ($r = -0.40$), ενώ στην Πορτογαλία η συσχέτιση ήταν ελαφρώς ασθενέστερη ($r = -0.33$, $p < .001$). Ο μέσος όρος της συσχέτισης για τις δύο χώρες ήταν $r = -0.36$.

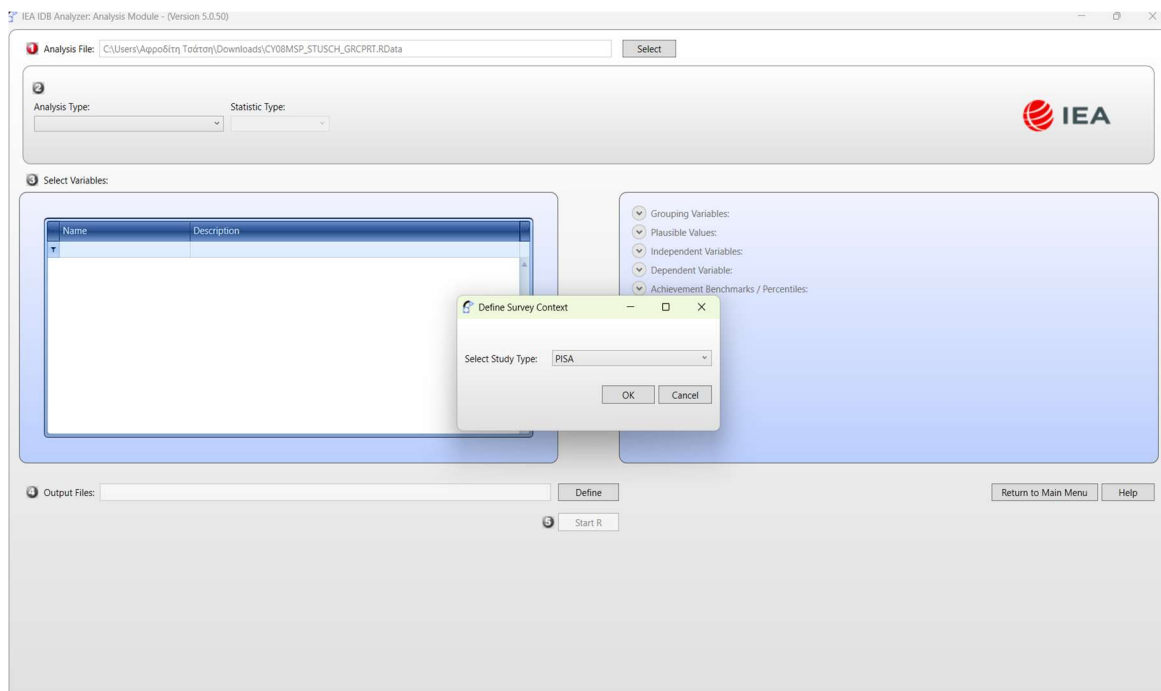
Υπολογισμός του συντελεστή συσχέτισης Spearman

Ερευνητική Υπόθεση: Διερευνάται η ύπαρξη και η κατεύθυνση της σχέσης μεταξύ του άγχους για τα Μαθηματικά και του δείκτη μεταναστευτικού υποβάθρου σε μαθητές της Ελλάδας και της Πορτογαλίας, μέσω της συσχέτισης Spearman.

Επιλογή μεταβλητών: IMMIG, ANXMAT

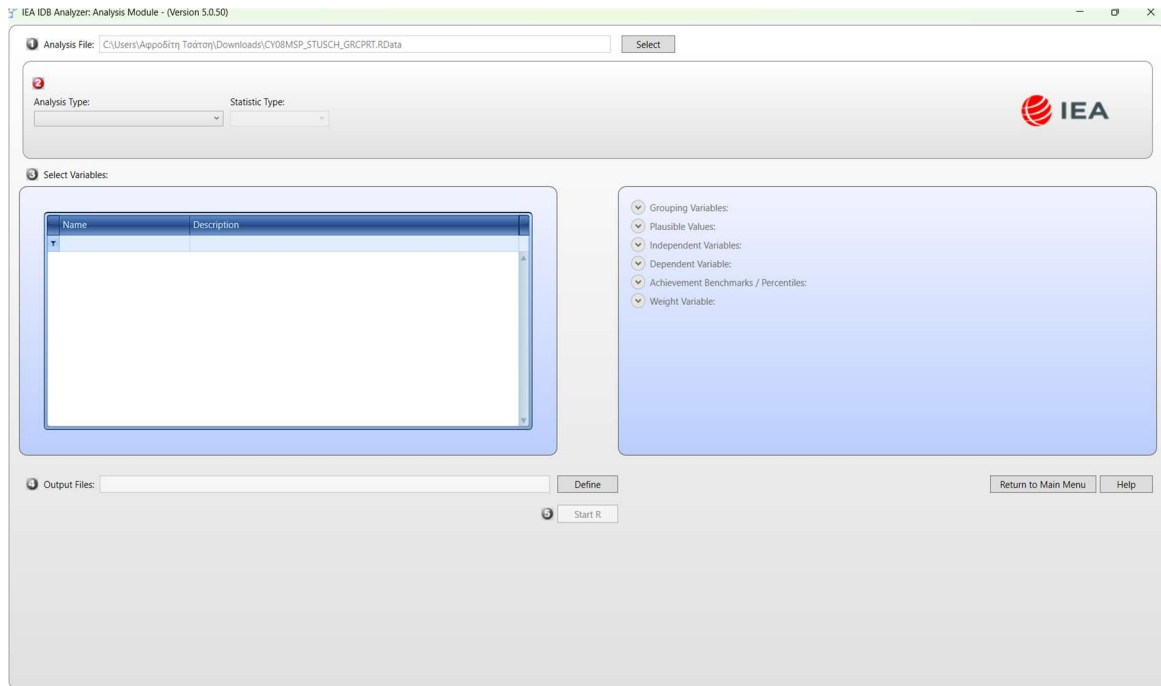
Χαρακτηριστικά μεταβλητών: IMMIG κατηγορική μεταβλητή, ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση συσχέτισης Spearman προκειμένου να διερευνηθεί η σχέση μεταξύ του άγχους για τα Μαθηματικά (ANXMAT) και του δείκτη μεταναστευτικού υποβάθρου (IMMIG) σε μαθητές της Ελλάδας και της Πορτογαλίας. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTRData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 11).



Εικόνα 11. Επιλογή αρχείου δεδομένων για ανάλυση Spearman – «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 12).

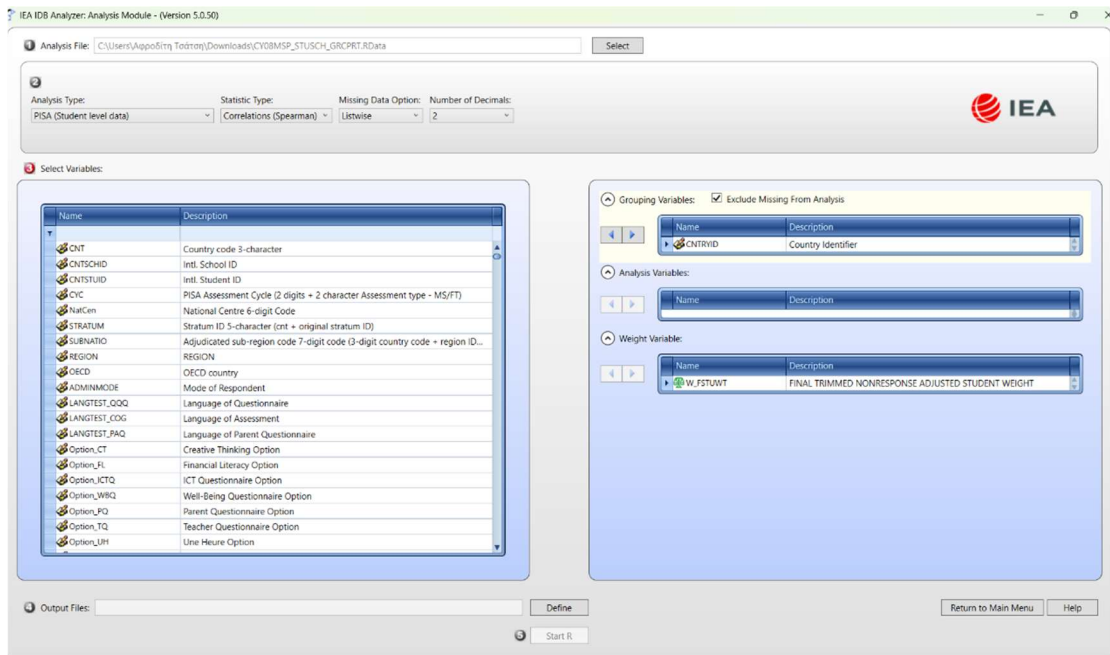


Εικόνα 12. Επιλογή «PISA (Student level)» στο πεδίο «Analysis Type» – Ανάλυση Spearman (R)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

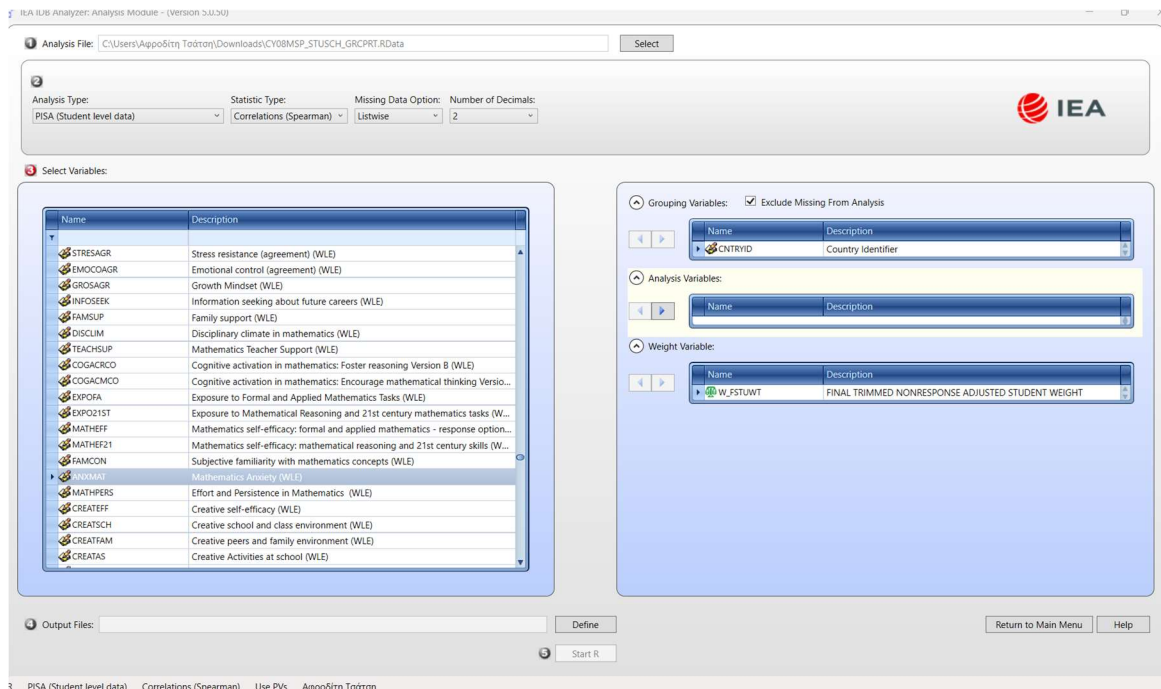
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Correlation (Spearman)».



Εικόνα 13. Επιλογή «Correlation (Spearman)» στο πεδίο «Statistic Type» (R)

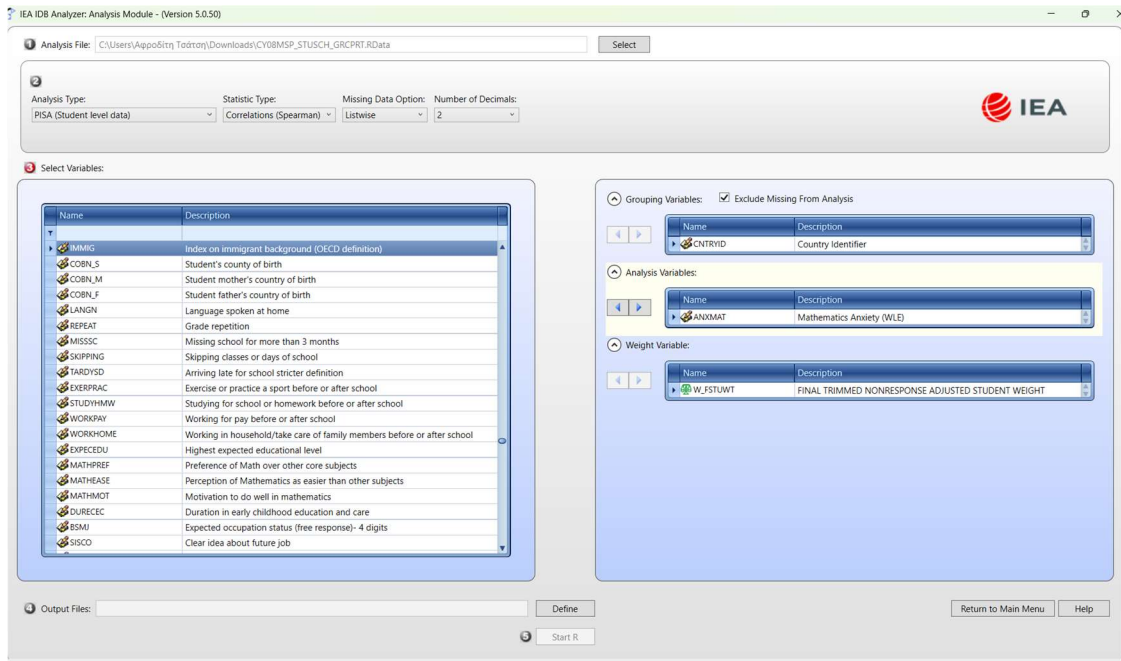
Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Missing Data Option» υπάρχουν διαθέσιμες οι επιλογές «Pairwise», που αν την επιλέξετε χρησιμοποιούνται όλα τα διαθέσιμα δεδομένα για ανάλυση και «Listwise» που αν την επιλέξετε χρησιμοποιούνται μόνο οι περιπτώσεις που έχουν πλήρη δεδομένα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 13).



Εικόνα 14. Ρύθμιση παραμέτρων ανάλυσης Spearman – «Grouping Variables» και «Missing Data Option» (R)

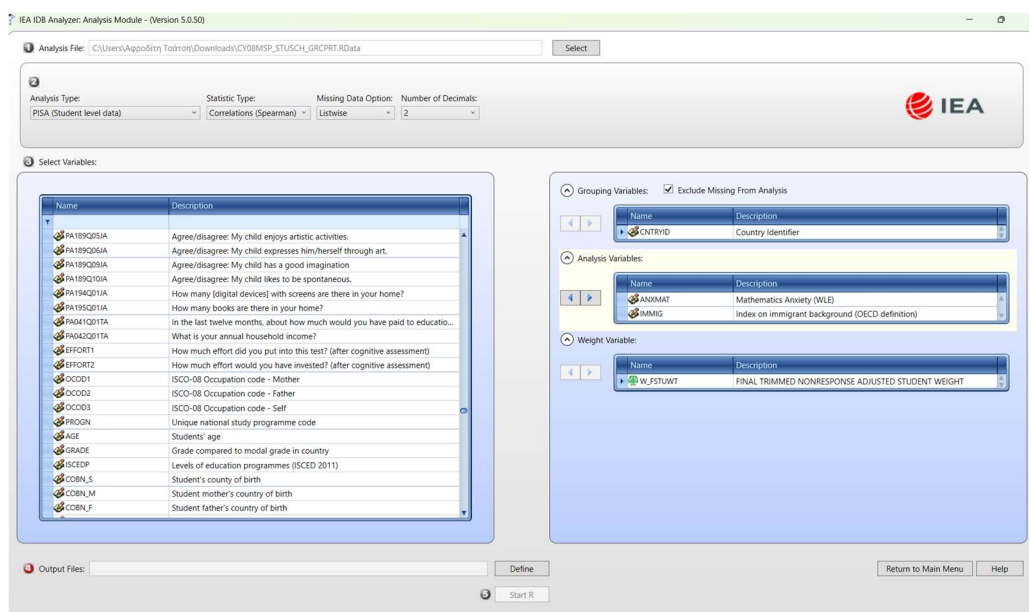
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 14 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλείψεις

τιμές στις μεταβλητές ομαδοποίησης. Στην συνέχεια μας ενδιαφέρει να υπολογίσουμε τον συντελεστή συσχέτισης Spearman μεταξύ της μεταβλητής «ANXMAT», και της «IMMIG». Επομένως βρίσκετε αρχικά την μεταβλητή «ANXMAT» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» την τοποθετείτε (Εικόνα 15). Επαναλαμβάνετε την διαδικασία για την μεταβλητή «IMMIG». Τέλος στο πλαίσιο «Weight Variable» υπάρχει ως προεπιλογή η μεταβλητή «W_FSTUWT». Να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 15).



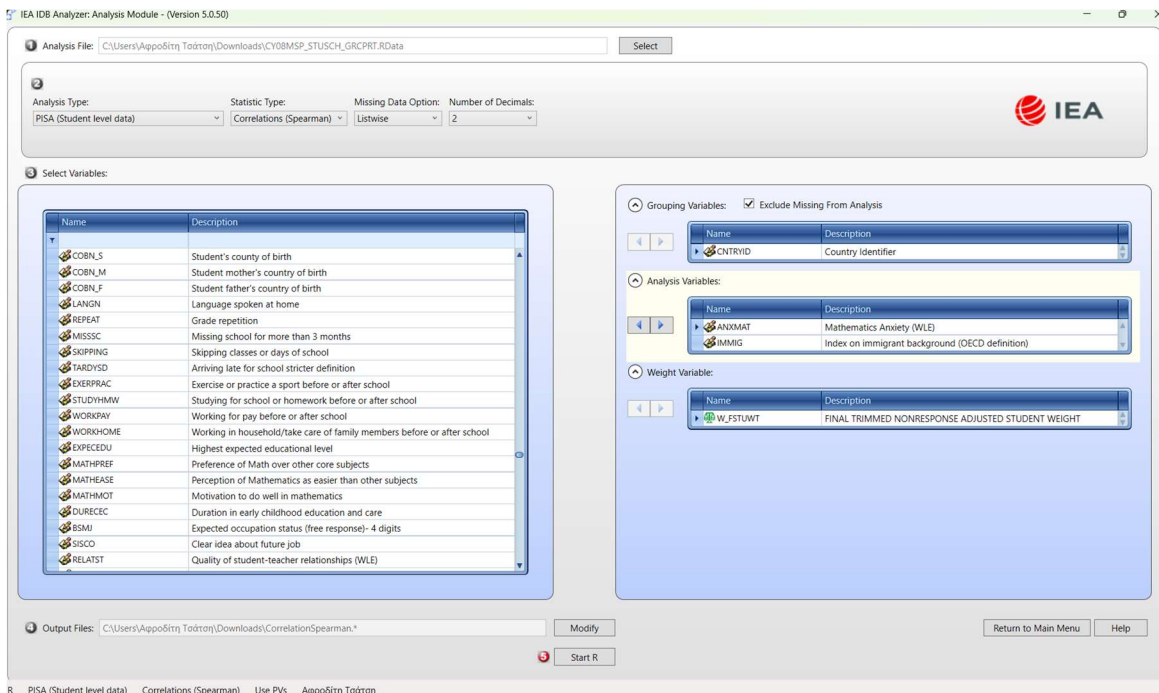
Εικόνα 15. Τοποθέτηση μεταβλητών ANXMAT και IMMIG στο πεδίο «Analysis Variables» – Spearman (R)

Το επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «CorrelationSpearman» (Εικόνα 16).



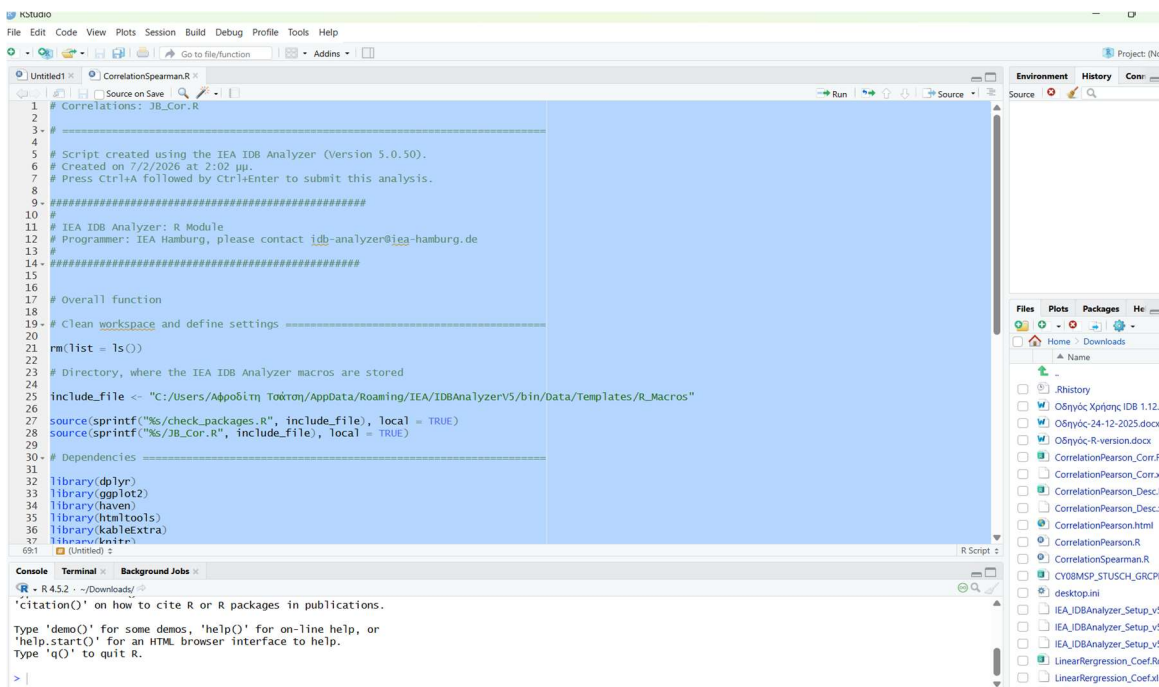
Εικόνα 16. Ορισμός ονόματος αρχείων εξόδου για ανάλυση Spearman (R)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio. (Εικόνα 17).



Εικόνα 17. Εκκίνηση παραγωγής script αρχείου Spearman μέσω «Start R»

Αφού μεταβείτε στο περιβάλλον του Rstudio, στο αρχείο CorrelationSpearman.R που εμφανίζεται, με Ctrl+A επιλέξτε όλο το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 18).



Εικόνα 18. Εκτέλεση κώδικα CorrelationSpearman.R στο περιβάλλον RStudio (Ctrl+A → Run)

Στο αρχείο CorrelationSpearman, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 19 και 20, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

The screenshot shows a web browser window with the URL 'C:/Users/Approdiτη%20τόστη/Downloads/CorrelationSpearman.html'. The page displays 'Weighted Statistics' and 'Report' sections. The 'Weighted Statistics' section includes a table with columns: Country Identifier, Variable, N, Minimum, Maximum, Mean, and Std. Deviation. The 'Report' section includes 'Descriptive Statistics' and 'Correlation Coefficients' tables.

Country Identifier	Variable	N	Minimum	Maximum	Mean	Std. Deviation
Greece	Mathematics Anxiety (WLE)	87817	-2.39	2.63	0.26	1.16
Greece	Index on immigrant background (OECD definition)	94798	1.00	3.00	1.16	0.43
Portugal	Mathematics Anxiety (WLE)	83847	-2.39	2.63	0.14	1.03
Portugal	Index on immigrant background (OECD definition)	93917	1.00	3.00	1.18	0.53

Country Identifier	Variable	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)
Greece	RANXMAT	5679	86404.11	1764.58	2840.86	24.18	1641.26	10.10
Greece	RIMMIG	5679	86404.11	1764.58	2869.59	18.63	949.72	20.66
Portugal	RANXMAT	5846	82812.19	2019.09	2939.93	26.23	1691.17	9.86
Portugal	RIMMIG	5846	82812.19	2019.09	2934.17	20.58	902.77	26.64
Table Average	RANXMAT	NaN	NaN	NaN	2890.40	17.84	1666.22	7.06
Table Average	RIMMIG	NaN	NaN	NaN	2901.88	13.88	926.24	16.85

Country Identifier	Variable	RANXMAT	RIMMIG	RANXMAT.se	RIMMIG.se
Greece	RANXMAT	1.00	0.09	0.00	0.02

Εικόνα 19. Αποτελέσματα ανάλυσης Spearman – περιγραφικά στατιστικά ανά χώρα (R)

Οι στήλες που φαίνονται στην Εικόνα 19, περιγράφουν τα εξής:

- Country Identifier: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Variable: Το όνομα της μεταβλητής.
- Number of cases: Ο αριθμός των μαθητών στο δείγμα της χώρας που απάντησαν στη συγκεκριμένη ερώτηση-μεταβλητή.
- Population Estimate: Εκτίμηση του συνολικού πληθυσμού που αντιπροσωπεύει το δείγμα σε εθνικό επίπεδο.
- Population Estimate (s.e.): Το τυπικό σφάλμα της εκτίμησης του πληθυσμού το οποίο δείχνει την ακρίβεια της εκτίμησης.
- Mean: Ο μέσος όρος των τιμών της μεταβλητής στο δείγμα της χώρας.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Standard Deviation: Η τυπική απόκλιση η οποία δείχνει πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο.
- Standard Deviation (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης το οποίο δείχνει πόσο αξιόπιστη είναι η εκτίμηση της διασποράς.

Country	Variable	Count	Mean	SE	Median	Q1	Q3	Min	Max
Greece	RANXMAT	5846	82812.19	17.64.58	2019.09	2939.93	26.23	1691.17	9.86
Portugal	RIMMIG	5846	82812.19	2019.09	2934.17	20.58	902.77	26.64	
Table Average	RANXMAT	NaN	NaN	NaN	2890.40	17.84	1666.22	7.06	
Table Average	RIMMIG	NaN	NaN	NaN	2901.88	13.88	926.24	16.85	

Country Identifier	Variable	RANXMAT	RIMMIG	RANXMAT.se	RIMMIG.se
Greece	RANXMAT	1.00	0.09	0.00	0.02
Greece	RIMMIG	0.09	1.00	0.02	0.00
Portugal	RANXMAT	1.00	0.06	0.00	0.01
Portugal	RIMMIG	0.06	1.00	0.01	0.00
Table Average	RANXMAT	1.00	0.08	0.00	0.01
Table Average	RIMMIG	0.08	1.00	0.01	0.00

Εικόνα 20. Πίνακας συντελεστών συσχέτισης Spearman μεταξύ ANXMAT και IMMIG (R)

Οι στήλες που φαίνονται στην Εικόνα 20, περιγράφουν τα εξής:

- Country Identifier: Αναφέρεται στη χώρα από την οποία προέρχονται τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία). Όλες οι πληροφορίες που περιέχονται στις αντίστοιχες γραμμές του πίνακα αφορούν αποκλειστικά την εκάστοτε χώρα.
- Variable: Δείχνει τη μεταβλητή στην οποία αναφέρεται ο συντελεστής συσχέτισης.
- RANXMAT: Ο συντελεστής συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής RANXMAT.
- RIMMIG: Ο συντελεστής συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής RIMMIG. Η συσχέτιση είναι συμμετρική, επομένως η τιμή μεταξύ των RANXMAT και RIMMIG είναι ίδια ανεξάρτητα από τη σειρά τους.
- RANXMAT (s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής της στήλης "Variable" και της RANXMAT.
- RIMMIG (s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής της στήλης "Variable" και της RIMMIG

Ερμηνεία αποτελεσμάτων

Η ανάλυση συσχέτισης έδειξε ότι υπάρχει ασθενής αλλά θετική συσχέτιση μεταξύ του άγχους για τα Μαθηματικά και του δείκτη μεταναστευτικού υπόβαθρου και στις δύο χώρες. Στην Ελλάδα, παρατηρήθηκε ασθενής θετική συσχέτιση ($\rho = 0.09$, $SE = 0.02$), υποδηλώνοντας ότι οι μαθητές με μεταναστευτικό υπόβαθρο τείνουν να εμφανίζουν ελαφρώς αυξημένα επίπεδα άγχους για τα Μαθηματικά. Αντίστοιχα, στην Πορτογαλία η συσχέτιση ήταν επίσης ασθενής και θετική ($\rho = 0.06$, $SE = 0.01$), αλλά μικρότερης έντασης σε σύγκριση με την Ελλάδα. Ο μέσος όρος της συσχέτισης για τις δύο χώρες ήταν $\rho = 0.08$, γεγονός που ενισχύει ότι η σχέση μεταξύ των δύο μεταβλητών είναι γενικά χαμηλής ισχύος.

Υπολογισμός γραμμικής παλινδρόμησης

Ερευνητική Υπόθεση: Διερευνάται αν το άγχος για τα Μαθηματικά προβλέπεται από το φύλο, το κοινωνικοοικονομικό επίπεδο και την αντίληψη των μαθητών ότι μπορούν να εφαρμόζουν μαθηματικό συλλογισμό σε δραστηριότητες του 21ου αιώνα, καθώς και η κατεύθυνση αυτών των σχέσεων.

Επιλογή μεταβλητών: ANXMAT, ST004D01T, ESCS, MATHEF21

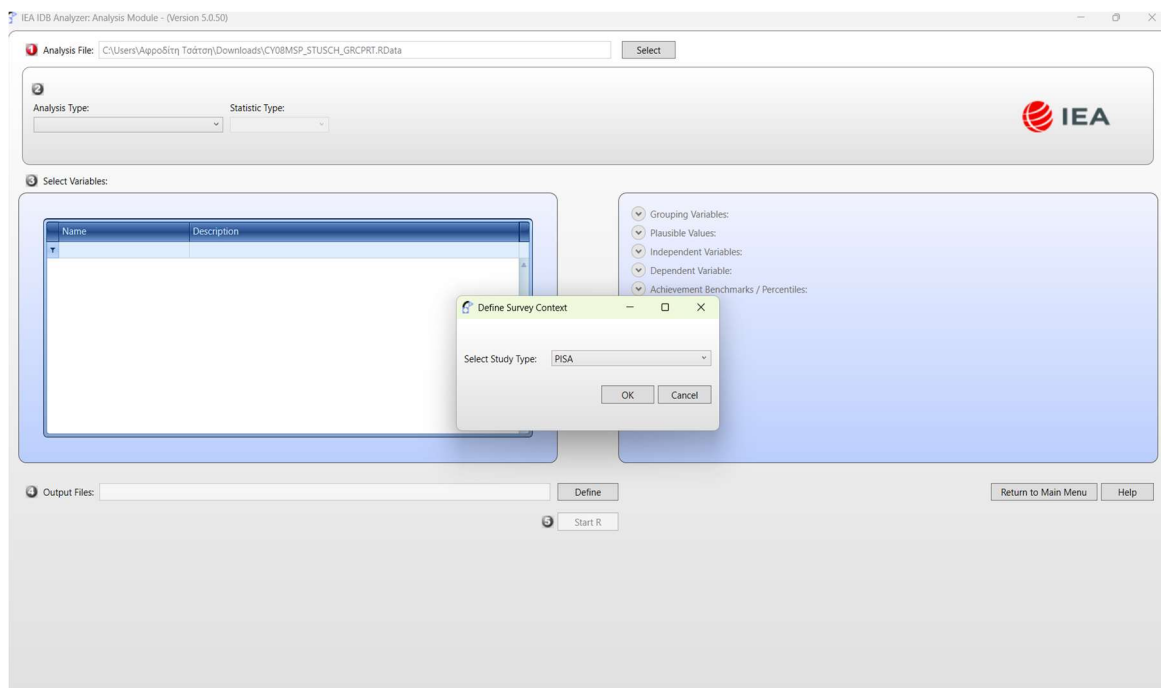
Χαρακτηριστικά μεταβλητών: ANXMAT συνεχής μεταβλητή, ST004D01T κατηγορική μεταβλητή, ESCS συνεχής μεταβλητή, MATHEF21 συνεχής μεταβλητή

Μεταβλητές: Μοντέλο (σε μορφή εξίσωσης)

Όπου:

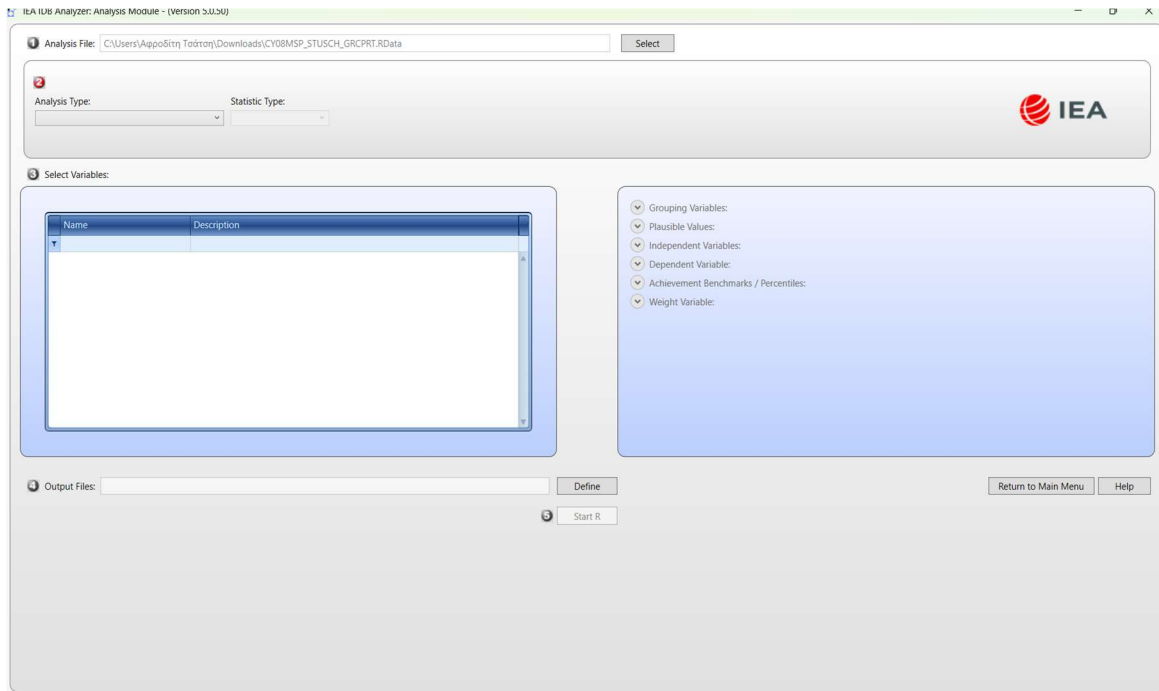
Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση γραμμικής παλινδρόμησης, όπου ως εξαρτημένη μεταβλητή χρησιμοποιείται το άγχος για τα Μαθηματικά (ANXMAT) και ως ανεξάρτητες μεταβλητές το φύλο του μαθητή (ST004D01T), το κοινωνικοοικονομικό επίπεδο (ESCS) και η αυτο-αντίληψη των μαθητών ότι μπορούν να εφαρμόσουν μαθηματικό συλλογισμό σε δραστηριότητες του 21ου αιώνα (MATHEF21).

Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTRData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 21).



Εικόνα 21. Επιλογή αρχείου δεδομένων για γραμμική παλινδρόμηση – «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 22).

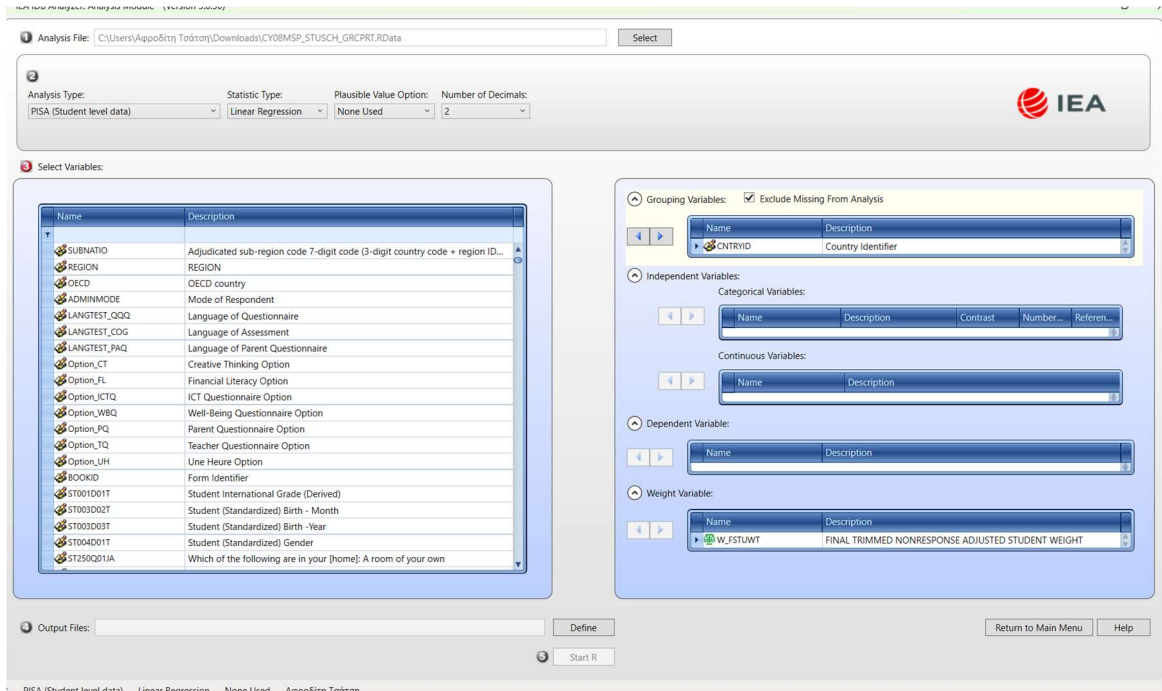


Εικόνα 22. Επιλογή «PISA (Student level)» στο πεδίο «Analysis Type» – Γραμμική Παλινδρόμηση (R)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

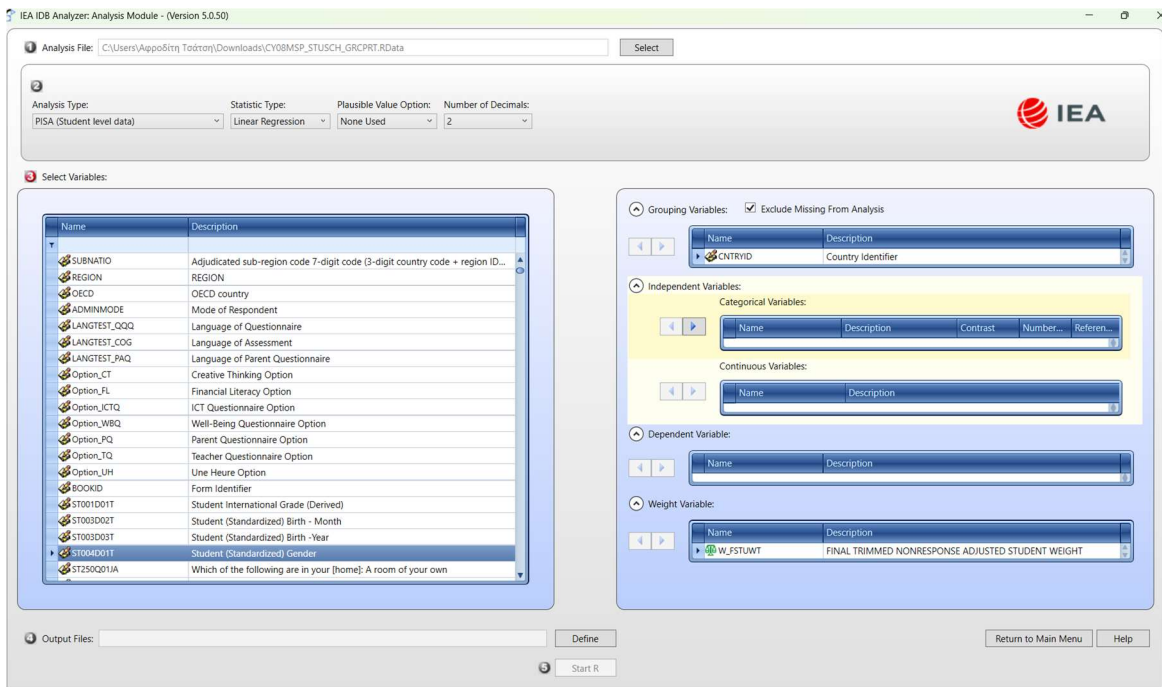
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλিপών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Linear Regression». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 23).



Εικόνα 23. Επιλογή «Linear Regression» στο πεδίο «Statistic Type» και ρύθμιση επιλογών (R)

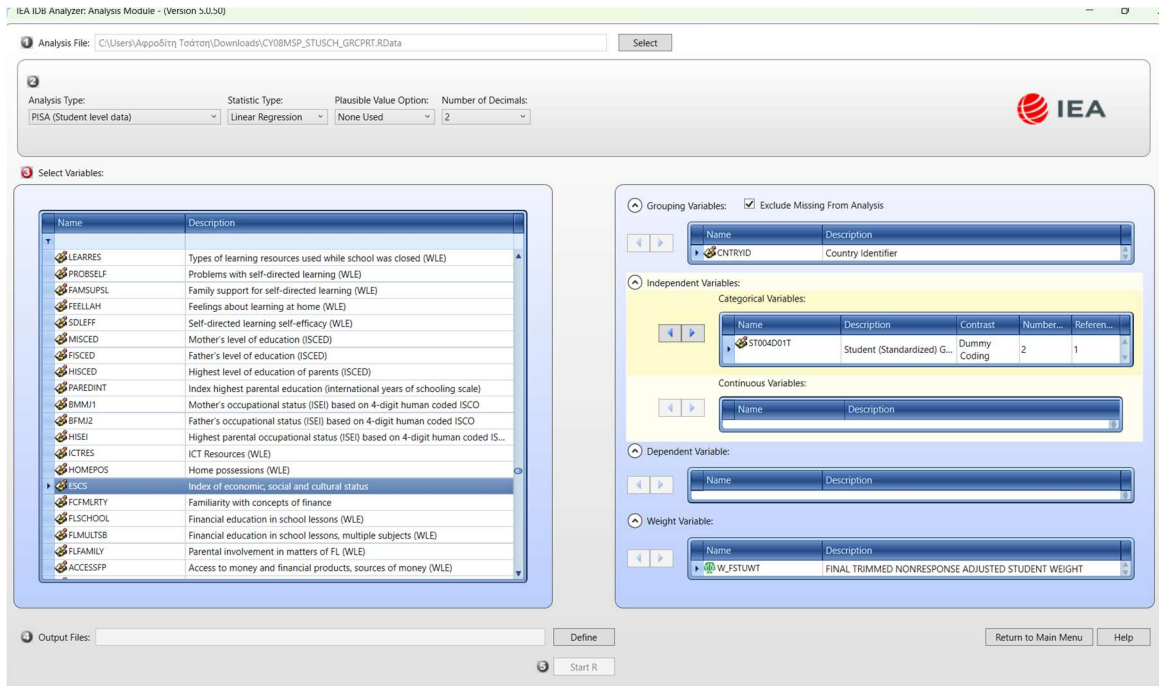
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 23, στο πλαίσιο «Grouping Variables», το οποίο περιλαμβάνει τη λίστα των μεταβλητών εμφανίζεται από προεπιλογή η μεταβλητή «CNTRYID». Επιπλέον, εφόσον η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, στην ανάλυση θα συμπεριληφθούν μόνο οι περιπτώσεις με μη ελλιπείς τιμές στις μεταβλητές.



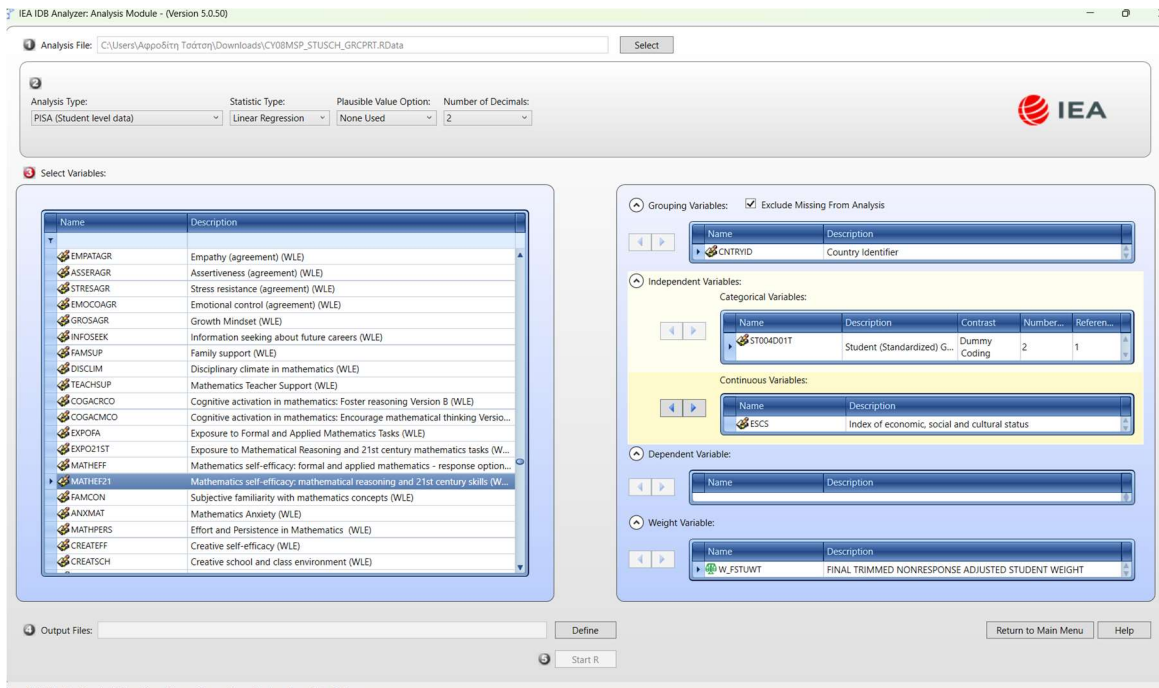
Εικόνα 24. Τοποθέτηση κατηγορικής μεταβλητής ST004D01T στο πεδίο «Categorical Variables» (R)

Για την υλοποίηση της γραμμικής παλινδρόμησης, είναι απαραίτητος ο εντοπισμός και η σωστή τοποθέτηση των απαιτούμενων μεταβλητών στο πλαίσιο «Select Variables». Αρχικά, εντοπίζεται η μεταβλητή «ST004D01T» (φύλο μαθητή), η οποία επιλέγεται και, με το μπλε βελάκι μεταφέρεται στο πλαίσιο «Independent Variables» και συγκεκριμένα στο πλαίσιο «Categorical Variables» (Εικόνα 24). Στη

συνέχεια, η ίδια διαδικασία επαναλαμβάνεται για τις μεταβλητές «ESCS» και «MATHEF21», οι οποίες μεταφέρονται στο πλαίσιο «Continuous Variables», καθώς πρόκειται για συνεχείς μεταβλητές (Εικόνα 25 και Εικόνα 26).

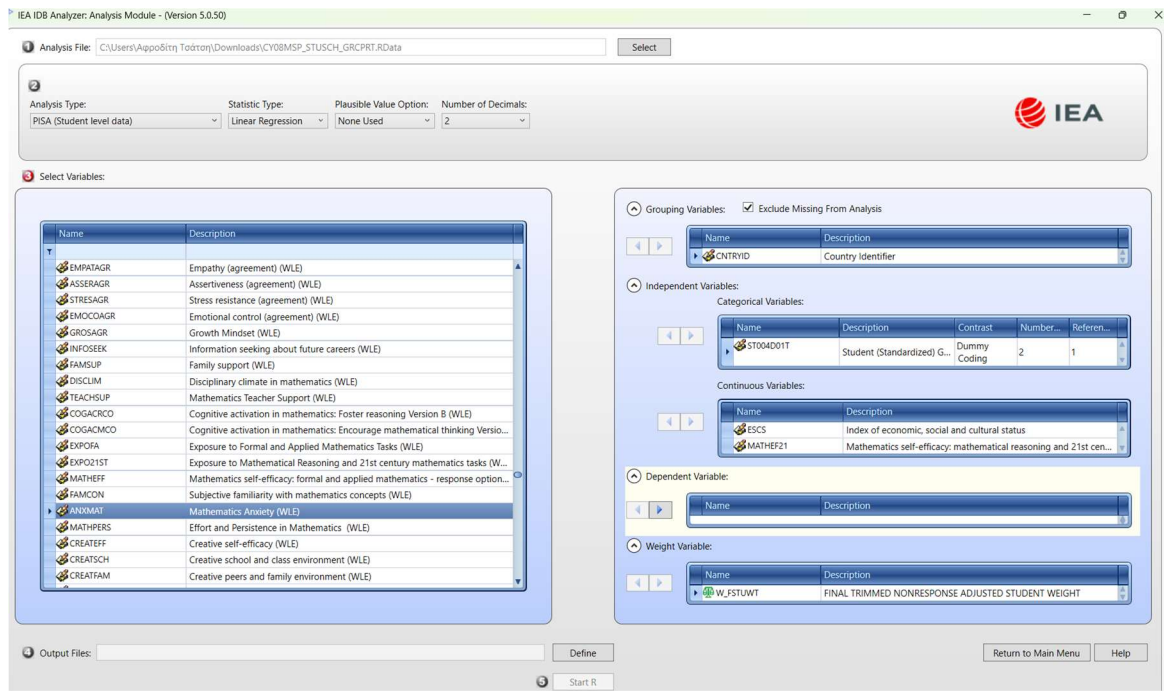


Εικόνα 25. Τοποθέτηση συνεχούς μεταβλητής ESCS στο πεδίο «Continuous Variables» (R)



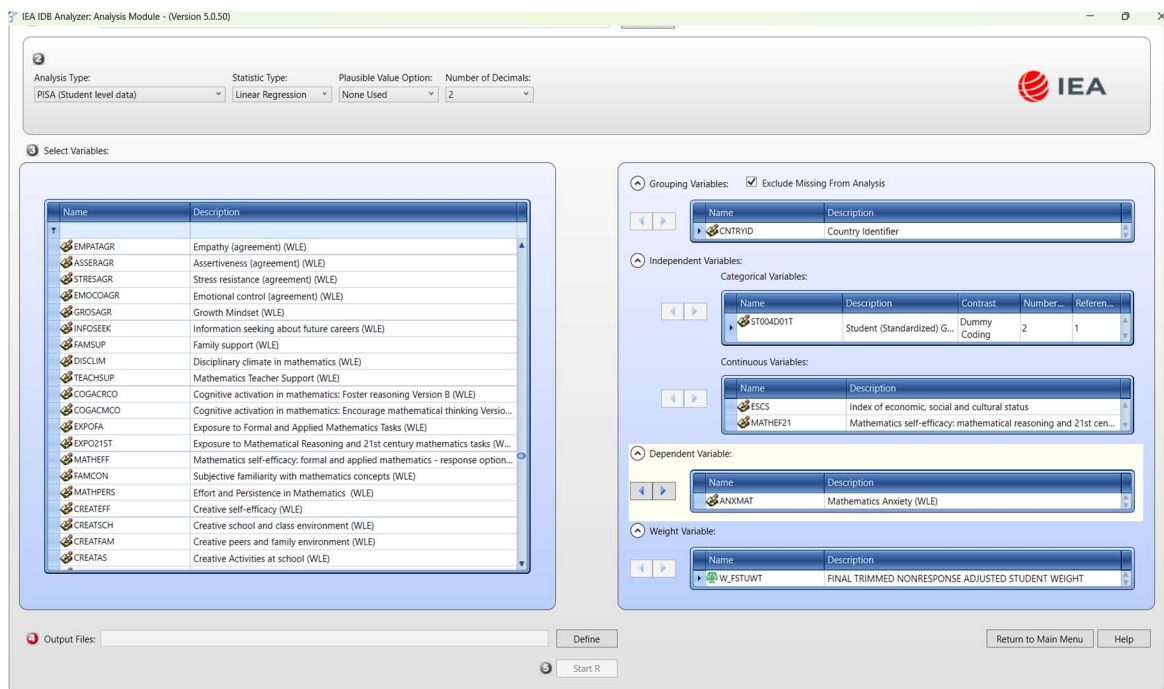
Εικόνα 26. Τοποθέτηση συνεχούς μεταβλητής MATHEF21 στο πεδίο «Continuous Variables» (R)

Ακολούθως, στο πλαίσιο «Dependent Variable», επιλέγεται η μεταβλητή «ANXMAT», η οποία χρησιμοποιείται ως εξαρτημένη μεταβλητή στη γραμμική παλινδρόμηση. Τέλος, στο πλαίσιο «Weight Variable» εμφανίζεται από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα την κατάλληλη μεταβλητή βαρών για την εκτέλεση της ανάλυσης (Εικόνα 27).



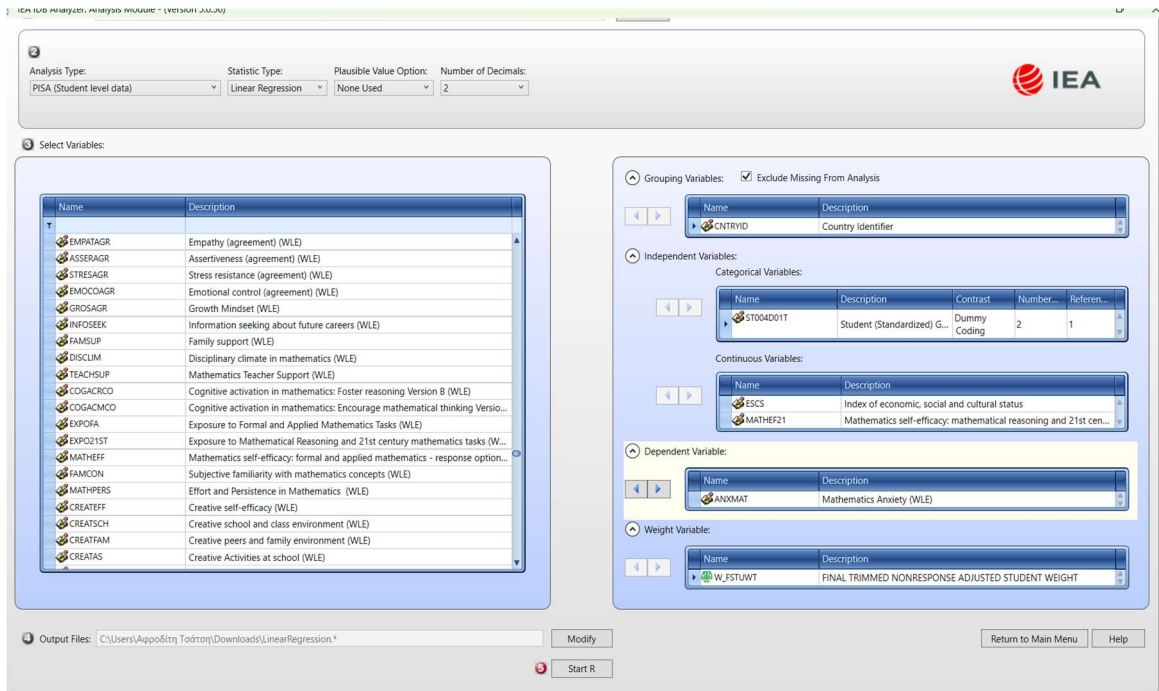
Εικόνα 27. Επιλογή εξαρτημένης μεταβλητής ANXMAT στο πεδίο «Dependent Variable» (R)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «LinearRegression» (Εικόνα 28).



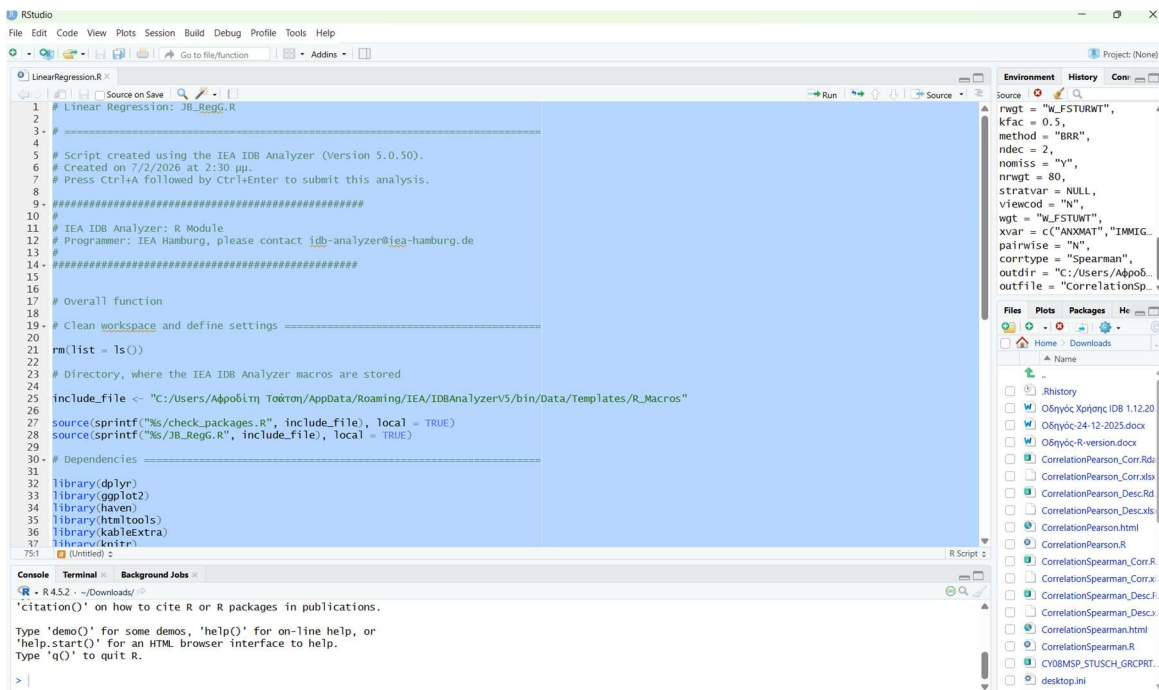
Εικόνα 28. Ορισμός ονόματος αρχείων εξόδου για γραμμική παλινδρόμηση (R)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio (Εικόνα 29).



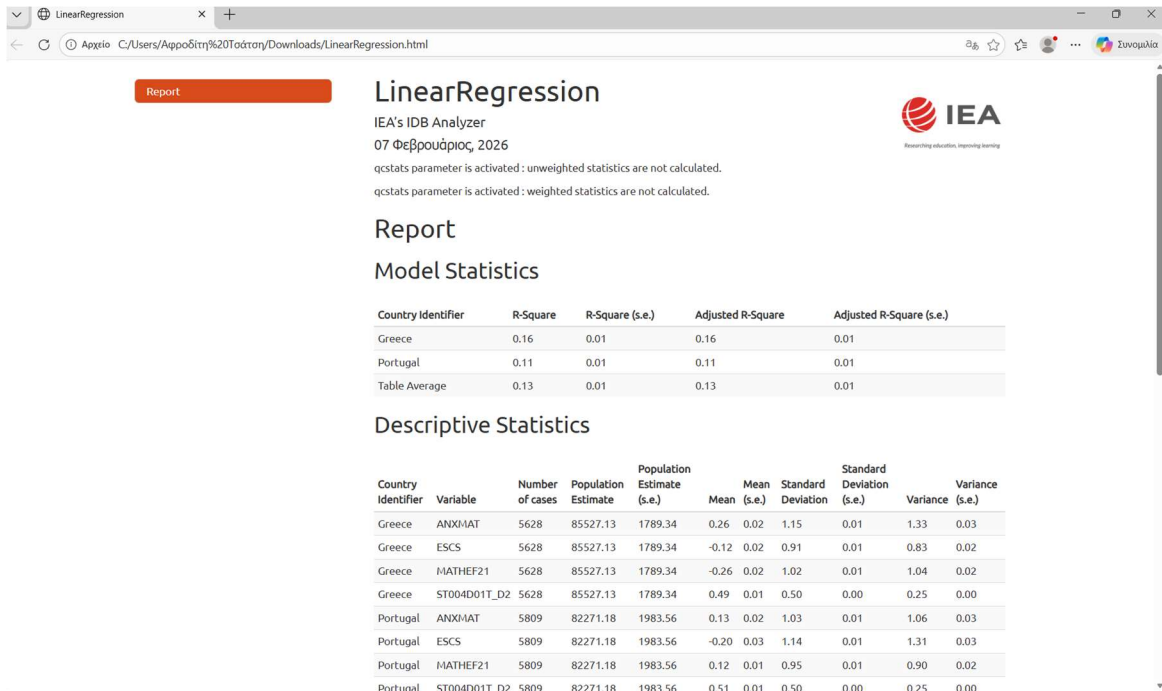
Εικόνα 29. Εκκίνηση παραγωγής script αρχείου γραμμικής παλινδρόμησης μέσω «Start R»

Ύστερα από τη μετάβαση στο Rstudio, στο αρχείο LinearRegression.R που εμφανίζεται, με Ctrl+A επιλέξετε το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 30).



Εικόνα 30. Εκτέλεση κώδικα LinearRegression.R στο RStudio

Στο αρχείο LinearRegression, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 31, 32 και 33, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



Εικόνα 31. Αποτελέσματα R-Square και Adjusted R-Square ανά χώρα – Γραμμική Παλινδρόμηση (R)

Οι στήλες στην Εικόνα 31 παρουσιάζουν τα στατιστικά του μοντέλου για διάφορες χώρες που επιλέγονται από τον χρήστη. Συγκεκριμένα:

- **Country Identifier:** Το όνομα της χώρας, εδώ συγκεκριμένα Ελλάδα, Πορτογαλία και στο τέλος δίνεται ο μέσος όρος.
- **R-Square:** Ο συντελεστής προσδιορισμού, όπου είναι ουσιαστικά το ποσοστό διακύμανσης του άγχους για τα Μαθηματικά που εξηγείται από τις ανεξάρτητες μεταβλητές (εδώ «ST004D01T», «ESCS», «MATHEF21»)
- **R-Square (s.e.):** Το τυπικό σφάλμα του R-Square.
- **Adjusted R-Square:** Ο προσαρμοσμένος συντελεστής προσδιορισμού, που λαμβάνει υπόψη τον αριθμό μεταβλητών.
- **Adjusted R-Square (s.e.):** Το τυπικό σφάλμα του προσαρμοσμένου R-Square.

The screenshot shows the 'LinearRegression' window with a sidebar menu containing 'Report', 'Model Statistics', 'Descriptive Statistics' (highlighted), and 'Regression Coefficients'. The main area displays two tables: 'Descriptive Statistics' and 'Regression Coefficients'.

Table Average

Table Average	0.13	0.01	0.13	0.01
---------------	------	------	------	------

Descriptive Statistics

Country Identifier	Variable	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)	Variance	Variance (s.e.)
Greece	ANXIMAT	5628	85527.13	1789.34	0.26	0.02	1.15	0.01	1.33	0.03
Greece	ESCS	5628	85527.13	1789.34	-0.12	0.02	0.91	0.01	0.83	0.02
Greece	MATHEF21	5628	85527.13	1789.34	-0.26	0.02	1.02	0.01	1.04	0.02
Greece	ST004D01T_D2	5628	85527.13	1789.34	0.49	0.01	0.50	0.00	0.25	0.00
Portugal	ANXIMAT	5809	82271.18	1983.56	0.13	0.02	1.03	0.01	1.06	0.03
Portugal	ESCS	5809	82271.18	1983.56	-0.20	0.03	1.14	0.01	1.31	0.03
Portugal	MATHEF21	5809	82271.18	1983.56	0.12	0.01	0.95	0.01	0.90	0.02
Portugal	ST004D01T_D2	5809	82271.18	1983.56	0.51	0.01	0.50	0.00	0.25	0.00
Table Average	ANXIMAT	NaN	NaN	NaN	0.20	0.01	1.09	0.01	1.20	0.02
Table Average	ESCS	NaN	NaN	NaN	-0.16	0.02	1.03	0.01	1.07	0.02
Table Average	MATHEF21	NaN	NaN	NaN	-0.07	0.01	0.99	0.01	0.97	0.02
Table Average	ST004D01T_D2	NaN	NaN	NaN	0.50	0.00	0.50	0.00	0.25	0.00

Regression Coefficients

EqVar	Country Identifier	Variable	Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stndrdzd. Coefficient	Stndrdzd. Coefficient (s.e.)	Stndrdzd. Coefficient (t-value)
(CONSTANT)	Greece	(CONSTANT)	0.30	0.02	13.04	NA	NA	NA

Εικόνα 32. Περιγραφικά στατιστικά μεταβλητών γραμμικής παλινδρόμησης ανά χώρα (R)

Ο πίνακας της Εικόνας 32 παρουσιάζει περιγραφικά στατιστικά για τις μεταβλητές που συμμετέχουν στη γραμμική παλινδρόμηση, για κάθε χώρα χωριστά (εδώ Ελλάδα και Πορτογαλία), όπως επιλέγονται από τον χρήστη. Συγκεκριμένα:

- Country Identifier: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Variable: Το όνομα της μεταβλητής.
- Number of cases: Ο αριθμός των μαθητών στο δείγμα της χώρας που απάντησαν στη συγκεκριμένη ερώτηση-μεταβλητή.
- Population Estimate: Εκτίμηση του συνολικού πληθυσμού που αντιπροσωπεύει το δείγμα σε εθνικό επίπεδο.
- Population Estimate (s.e.): Το τυπικό σφάλμα της εκτίμησης του πληθυσμού το οποίο δείχνει την ακρίβεια της εκτίμησης.
- Mean: Ο μέσος όρος των τιμών της μεταβλητής στο δείγμα της χώρας.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Standard Deviation: Η τυπική απόκλιση η οποία δείχνει πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο.
- Standard Deviation (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης.
- Variance: Η διακύμανση.
- Variance (s.e.): Το τυπικό σφάλμα της διακύμανσης.

EqVar	Country Identifier	Variable	Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stndrdzd. Coefficient	Stndrdzd. Coefficient (s.e.)	Stndrdzd. Coefficient (t-value)
(CONSTANT)	Greece	(CONSTANT)	0.30	0.02	13.04	NA	NA	NA
ESCS	Greece	ESCS	-0.07	0.02	-4.30	-0.06	0.01	-4.28
MATHEF21	Greece	MATHEF21	-0.38	0.02	-20.87	-0.34	0.02	-21.09
ST004D01T_D2	Greece	ST004D01T_D2	-0.31	0.03	-9.40	-0.13	0.01	-9.59
(CONSTANT)	Portugal	(CONSTANT)	0.29	0.02	13.85	NA	NA	NA
ESCS	Portugal	ESCS	-0.07	0.01	-5.35	-0.07	0.01	-5.30
MATHEF21	Portugal	MATHEF21	-0.30	0.02	-13.04	-0.27	0.02	-12.98
ST004D01T_D2	Portugal	ST004D01T_D2	-0.25	0.03	-9.05	-0.12	0.01	-9.26
(CONSTANT)	Table Average	(CONSTANT)	0.29	0.02	18.96	NaN	NA	NaN
ESCS	Table Average	ESCS	-0.07	0.01	-6.60	-0.07	0.01	-6.78
MATHEF21	Table Average	MATHEF21	-0.34	0.01	-23.28	-0.31	0.01	-23.14
ST004D01T_D2	Table Average	ST004D01T_D2	-0.28	0.02	-13.03	-0.13	0.01	-13.33

Εικόνα 33. Συντελεστές παλινδρόμησης (B , SE , t , θ) ανά μεταβλητή και χώρα (R)

Ο πίνακας της Εικόνας 33 παρουσιάζει τους συντελεστές παλινδρόμησης από το γραμμικό μοντέλο που εφαρμόστηκε, για κάθε χώρα χωριστά (π.χ. Ελλάδα, Πορτογαλία), όπως επιλέγονται από τον χρήστη. Συγκεκριμένα:

- Country Identifier: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Variable: Το όνομα της ανεξάρτητης μεταβλητής που εισέρχεται στο μοντέλο (π.χ. ESCS, MATHEF21, ST004D01T_D2) καθώς και η σταθερά (CONSTANT)).
- Regression Coefficient: Ο εκτιμώμενος συντελεστής παλινδρόμησης. Δείχνει κατά πόσο μεταβάλλεται η εξαρτημένη μεταβλητή (π.χ. επίδοση) όταν η ανεξάρτητη αυξάνεται κατά μία μονάδα, κρατώντας τις υπόλοιπες σταθερές.
- Regression Coefficient (s.e.): Το τυπικό σφάλμα της εκτίμησης του συντελεστή. Μετράει την αβεβαιότητα της εκτίμησης.
- Regression Coefficient (t-value): Η τιμή του t-στατιστικού για τον συντελεστή.
- Stndrdzd. Coefficient: Ο τυποποιημένος συντελεστής παλινδρόμησης.
- Stndrdzd. Coefficient (s.e.): Το τυπικό σφάλμα της εκτίμησης του τυποποιημένου συντελεστή.
- Stndrdzd. Coefficient (t-value): Η τιμή του t-στατιστικού για τον τυποποιημένο συντελεστή. Χρησιμοποιείται για τον έλεγχο της στατιστικής σημαντικότητας στην τυποποιημένη κλίμακα.

Ερμηνεία αποτελεσμάτων

Διεξήχθη πολλαπλή γραμμική παλινδρόμηση με εξαρτημένη μεταβλητή το άγχος για τα Μαθηματικά (ANXMAT) και ανεξάρτητες μεταβλητές το κοινωνικοοικονομικό επίπεδο (ESCS), τον δείκτη (MATHEF21) και το φύλο (ST004D01T_D2) για την Ελλάδα και την Πορτογαλία.

Ελλάδα

Στην Ελλάδα, το ANXMAT είχε μέση τιμή $M = 0.26$ ($SE = 0.02$) και τυπική απόκλιση $SD = 1.15$. Το ESCS είχε $M = -0.12$ ($SE = 0.02$), ενώ το MATHEF21 είχε $M = -0.26$ ($SE = 0.02$). Η μεταβλητή φύλου (ST004D01T_D2) παρουσίασε μέση τιμή 0.49, υποδηλώνοντας περίπου ισόρροπη κατανομή των κατηγοριών.

Το μοντέλο εξήγησε το 16% της διακύμανσης του άγχους για τα Μαθηματικά ($R^2 = .16$, $SE = .01$; adjusted $R^2 = .16$). Οι συντελεστές παλινδρόμησης έδειξαν ότι το ESCS σχετίστηκε αρνητικά και στατιστικά σημαντικά με το άγχος ($\beta = -0.07$, $t = -4.30$), υποδηλώνοντας ότι αύξηση κατά 1 μονάδα στο ESCS συνδέεται με μείωση περίπου 0.07 μονάδων στο ANXMAT διατηρώντας τις λοιπές μεταβλητές σταθερές. Ο δείκτης MATHEF21 παρουσίασε επίσης αρνητική και ισχυρή σχέση με το άγχος ($\beta = -0.38$, $t = -20.87$), δηλαδή υψηλότερες τιμές MATHEF21 συνδέθηκαν με χαμηλότερο ANXMAT. Τέλος, το φύλο (ST004D01T_D2) ήταν αρνητικός και στατιστικά σημαντικός προγνωστικός παράγοντας ($\beta = -0.31$, $t = -9.40$), υποδηλώνοντας χαμηλότερο άγχος κατά περίπου 0.31 μονάδες για τα αγόρια σε σχέση με τα κορίτσια

Με βάση τους τυποποιημένους συντελεστές, ο ισχυρότερος προγνωστικός παράγοντας ήταν το MATHEF21 ($\beta = -.38$), ακολούθησε το φύλο ($\beta = -.31$) και στη συνέχεια το ESCS ($\beta = -.07$).

Πορτογαλία

Στην Πορτογαλία, το ANXMAT είχε $M = 0.13$ ($SE = 0.02$) και $SD = 1.03$, υποδηλώνοντας χαμηλότερο μέσο άγχος σε σύγκριση με την Ελλάδα. Το ESCS είχε $M = -0.20$ ($SE = 0.03$) και το MATHEF21 $M = 0.12$ ($SE = 0.01$). Η μεταβλητή φύλου (ST004D01T_D2) είχε μέση τιμή 0.51, επίσης ενδεικτική ισόρροπης κατανομής. Το μοντέλο εξήγησε το 11% της διακύμανσης του ANXMAT ($R^2 = .11$, $SE = .01$; adjusted $R^2 = .11$). Το ESCS συσχετίστηκε αρνητικά και στατιστικά σημαντικά με το άγχος ($b = -0.07$, $t = -5.35$). Αντίστοιχα, το MATHEF21 ήταν αρνητικός και στατιστικά σημαντικός προγνωστικός παράγοντας ($b = -0.30$, $t = -13.04$). Τέλος, και το φύλο (ST004D01T_D2) παρουσίασε αρνητική και στατιστικά σημαντική σχέση με το άγχος ($b = -0.25$, $t = -9.05$), με τα αγόρια να εμφανίζουν χαμηλότερο άγχος κατά περίπου 0.25 μονάδες σε σχέση με τα κορίτσια

Οι τυποποιημένοι συντελεστές έδειξαν ότι ο ισχυρότερος παράγοντας ήταν και εδώ το MATHEF21 ($\beta = -.30$), ακολούθησε το φύλο ($\beta = -.25$) και στη συνέχεια το ESCS ($\beta = -.07$).

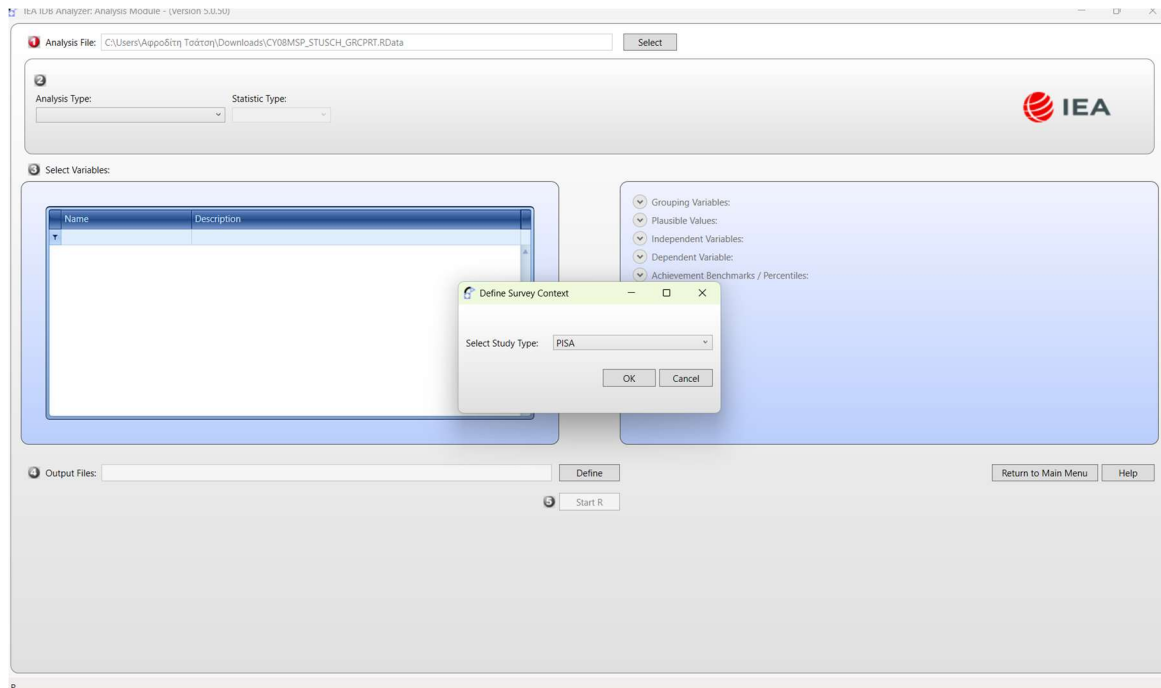
Υπολογισμός ποσοστών και μέσης τιμής

Ερευνητική υπόθεση: Διερευνώνται διαφορές στα ποσοστά και στις μέσες τιμές του άγχους για τα Μαθηματικά μεταξύ των δύο χωρών, λαμβάνοντας υπόψη το φύλο των μαθητών.

Επιλογή μεταβλητών: ANXMAT, ST004D01T

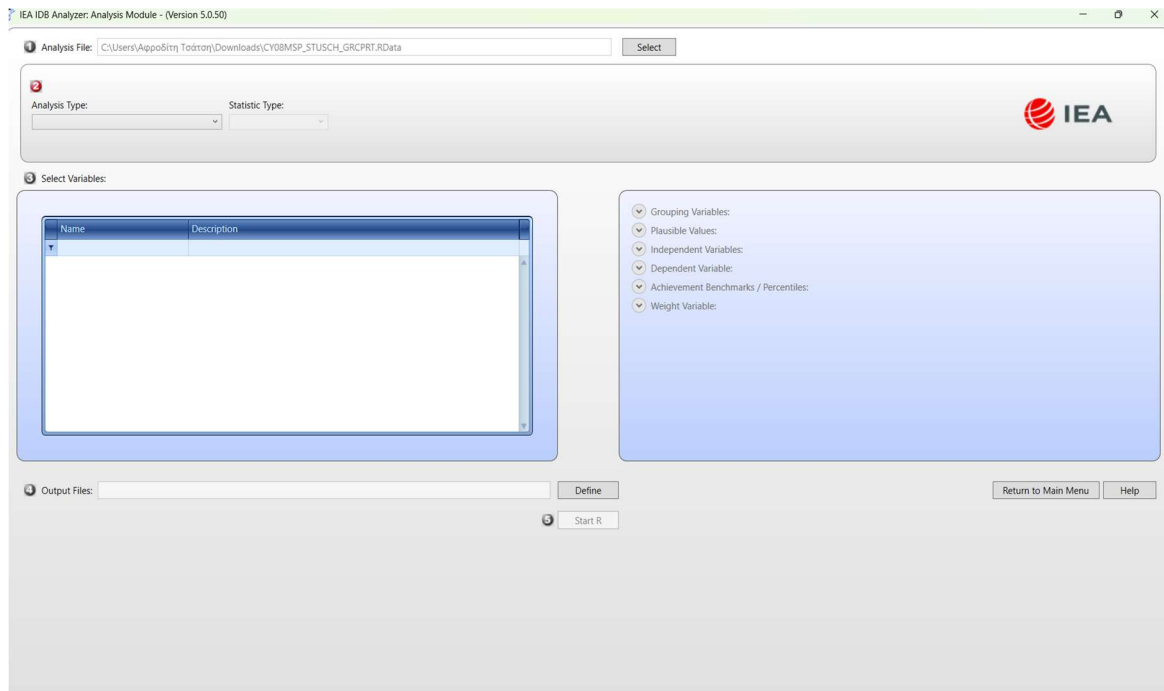
Χαρακτηριστικά μεταβλητών: ANXMAT συνεχής μεταβλητή, ST004D01T κατηγορική μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί σύγκριση ποσοστών ανάμεσα στην Ελλάδα και την Πορτογαλία, καθώς και ο υπολογισμός της μέσης τιμής, με μεταβλητή ενδιαφέροντος το άγχος για τα Μαθηματικά (ANXMAT). Η δεύτερη μεταβλητή που λαμβάνεται υπόψη είναι η «ST004D01T», η οποία καταγράφει το φύλο των μαθητών. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTRData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 34).



Εικόνα 34. Επιλογή αρχείου δεδομένων για ανάλυση Percentages & Means – «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 35).

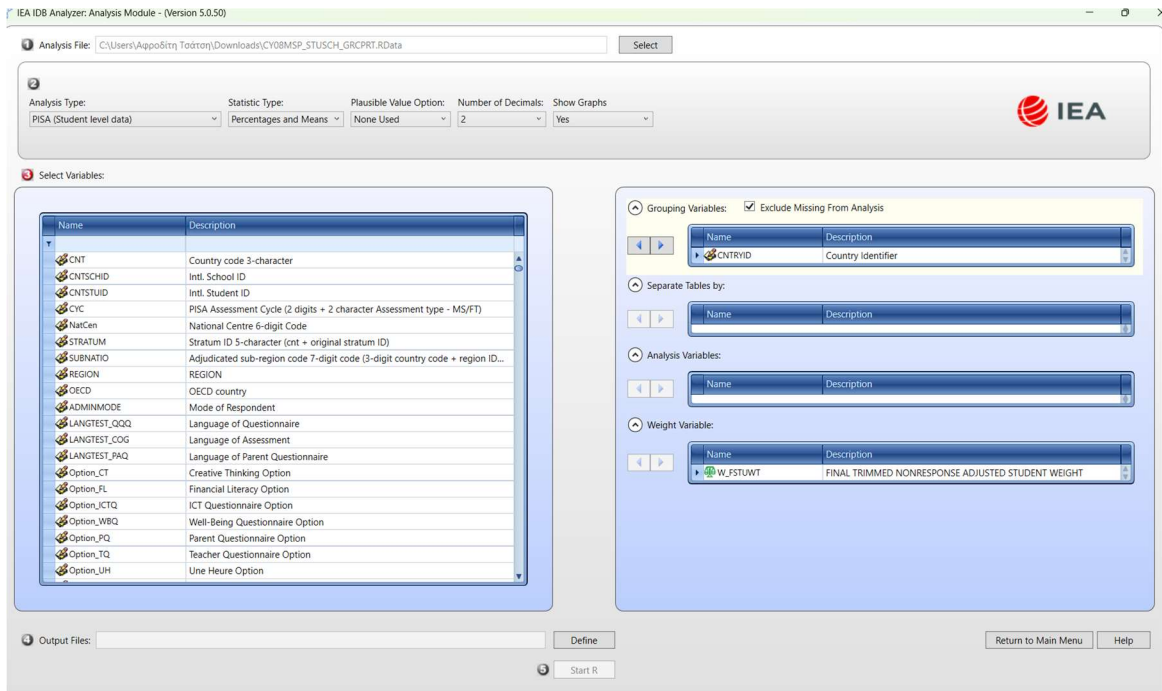


Εικόνα 35. Επιλογή «PISA (Student level)» – Ανάλυση Percentages & Means (R)

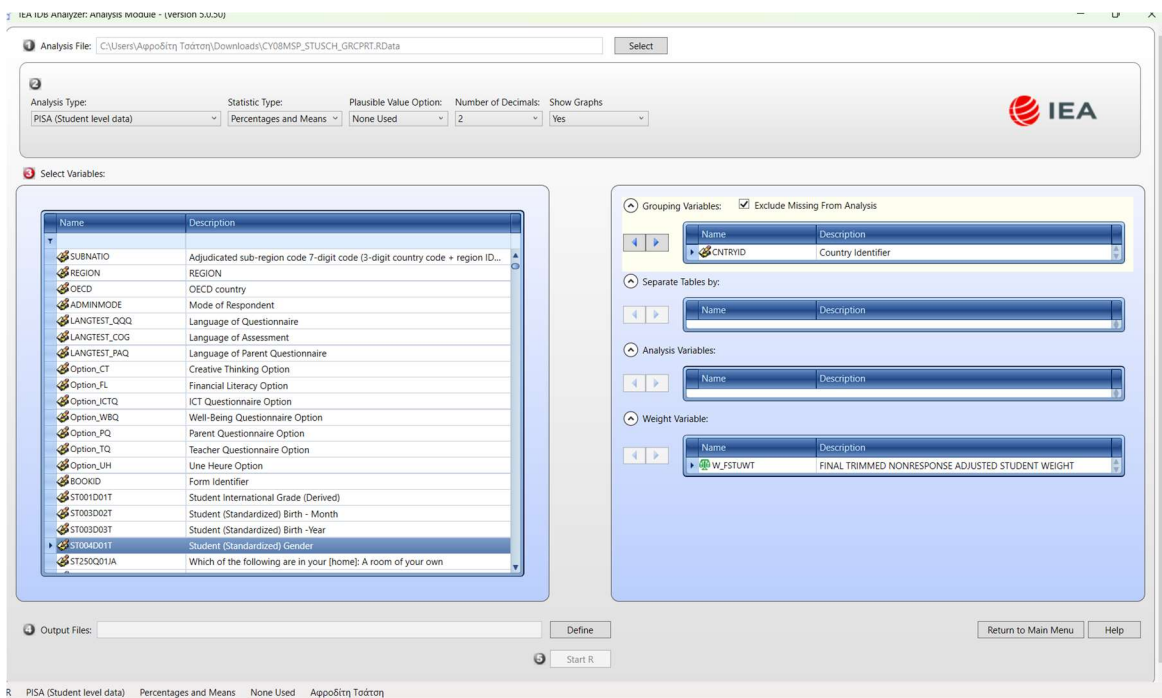
Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλিপών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγεται «Percentages and Means». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων (Εικόνα 36).



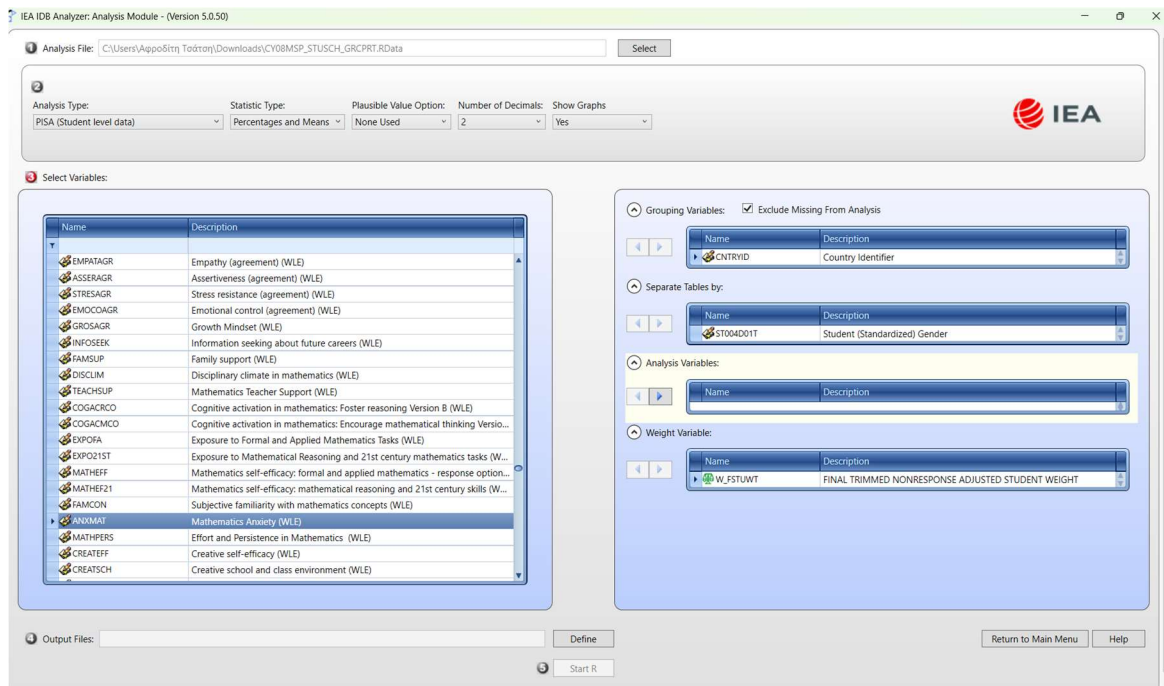
Εικόνα 36. Επιλογή «Percentages and Means» στο πεδίο «Statistic Type» (R)



Εικόνα 37. Ρύθμιση μεταβλητών ανάλυσης: ST004D01T ως «Separate Tables by» (R)

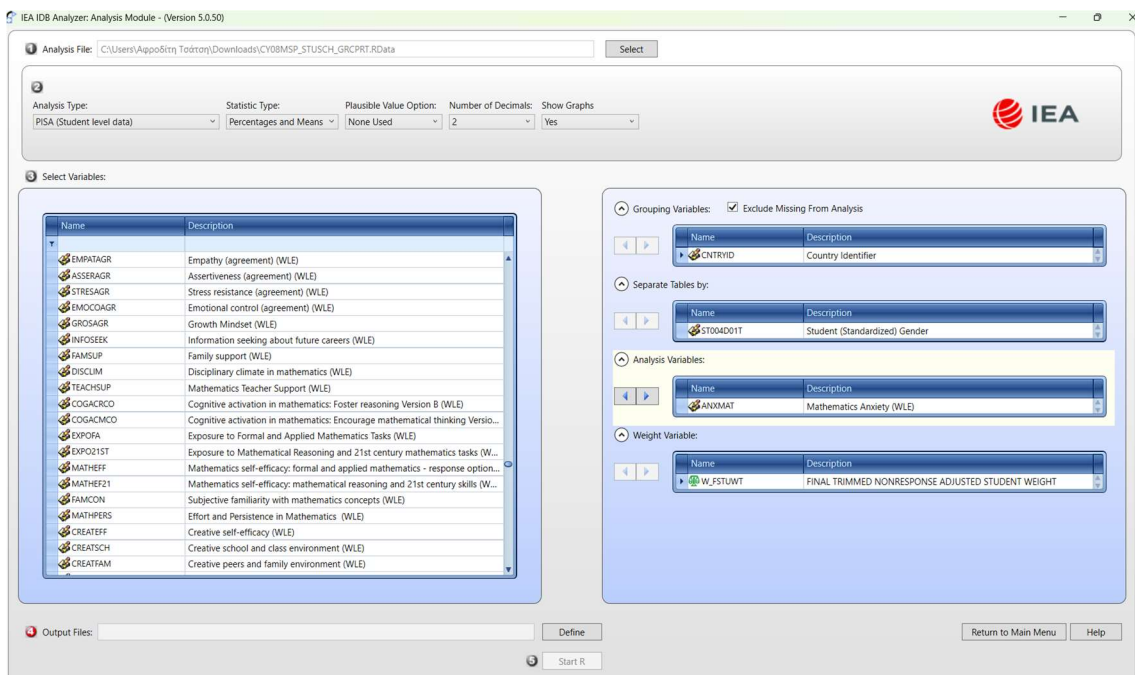
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 37 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλείψεις τιμές στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί η σύγκριση, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή «ST004D01T» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Separate Tables by» την τοποθετείτε (Εικόνα 38). Επαναλαμβάνετε την διαδικασία

για την μεταβλητή «ANXMAT» την οποία τοποθετείτε στο πλαίσιο «Analysis Variables». Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 38).

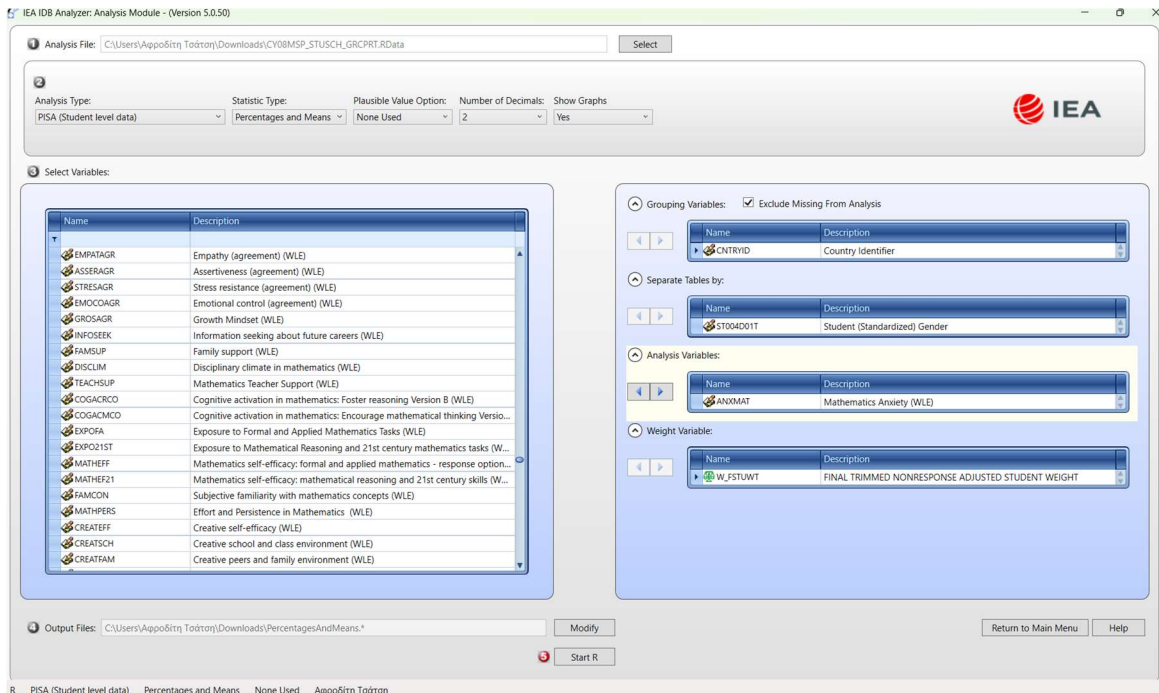


Εικόνα 38. Ρύθμιση μεταβλητών ανάλυσης: ANXMAT ως «Analysis Variable» (R)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PercentagesAndMeans» (Εικόνα 39). Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio (Εικόνα 40).

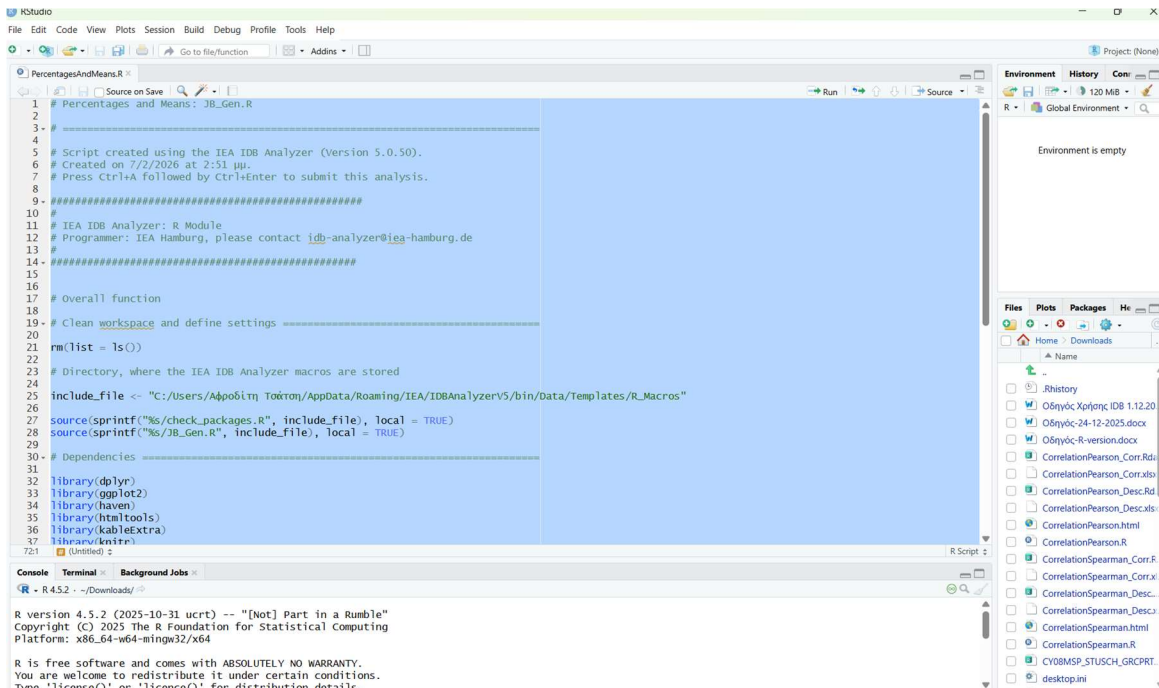


Εικόνα 39. Ορισμός ονόματος εξόδου για ανάλυση Percentages & Means (R)



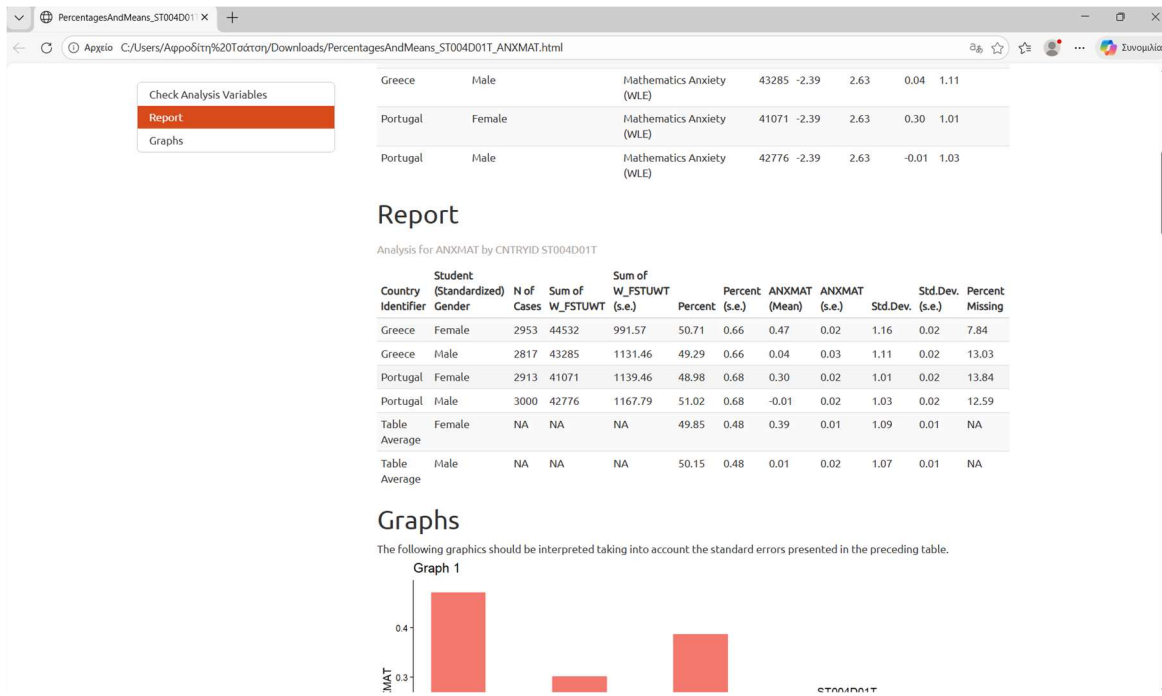
Εικόνα 40. Εκκίνηση παραγωγής script μέσω «Start R» – Percentages & Means

Ύστερα από τη μετάβαση στο Rstudio, στο αρχείο PercentagesAndMeans.R που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 41).



Εικόνα 41. Εκτέλεση κώδικα PercentagesAndMeans.R στο RStudio

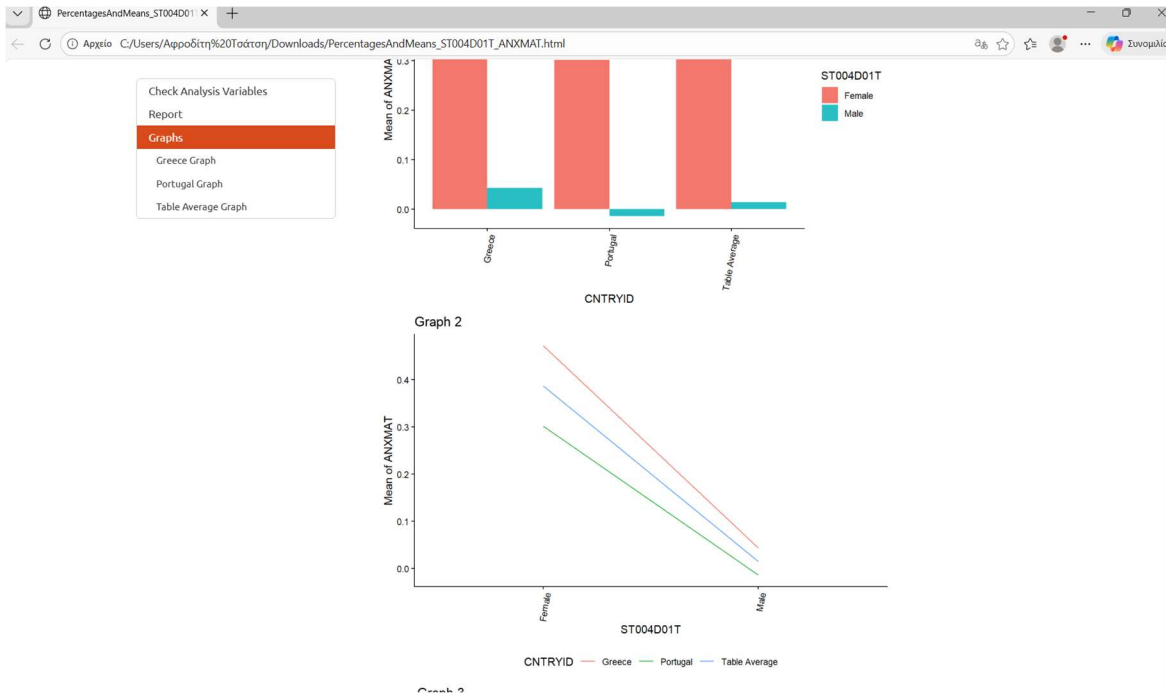
Στο αρχείο PercentagesAndMeans, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 42, 43, 44 και 45, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



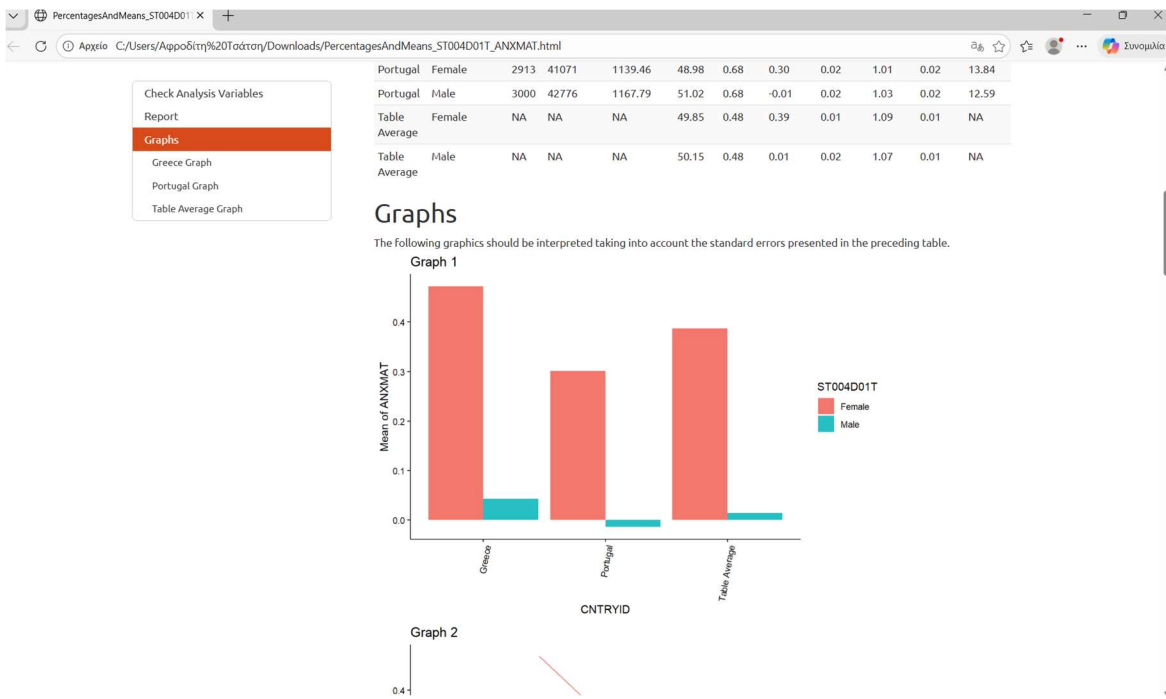
Εικόνα 42. Πίνακας μέσων όρων, ποσοστών και τυπικών αποκλίσεων ANXMAT ανά φύλο και χώρα (R)

Οι στήλες της Εικόνας 42 παρουσιάζουν την κατανομή και τα μέσα επίπεδα άγχους για τα Μαθηματικά (ANXMAT) κατά φύλο (ST004D01T) για κάθε χώρα. Συγκεκριμένα:

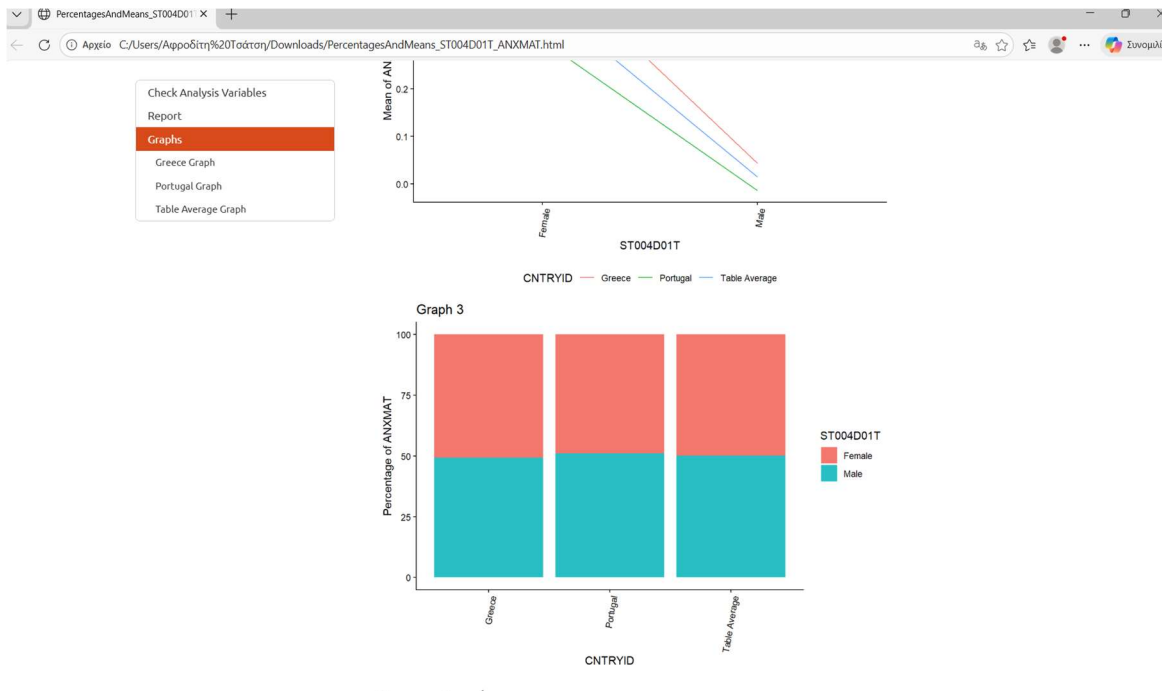
- Country Identifier: Η χώρα στην οποία αναφέρονται τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Student Gender (Standardized): Το φύλο των μαθητών, κωδικοποιημένο ως Female (κορίτσια) και Male (αγόρια).
- N of Cases: Ο αριθμός μαθητών στο δείγμα κάθε κατηγορίας φύλου.
- Sum of W_FSTUWT: Το συνολικό σταθμισμένο άθροισμα βαρών του δείγματος, το οποίο αντιστοιχεί στον πληθυσμό που εκπροσωπείται.
- Sum of W_FSTUWT (s.e.): Το τυπικό σφάλμα του συνολικού σταθμισμένου αθροίσματος βαρών.
- Percent: Το ποσοστό του πληθυσμού που αντιπροσωπεύει η κάθε κατηγορία φύλου μέσα στη χώρα.
- Percent (s.e.): Το τυπικό σφάλμα της ποσοστιαίας εκτίμησης.
- ANXMAT (Mean): Ο μέσος όρος της μαθηματικής επίδοσης (ANXMAT) για κάθε φύλο.
- ANXMAT (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Std. Dev: Η τυπική απόκλιση της επίδοσης, δηλαδή η διασπορά των επιδόσεων γύρω από τον μέσο όρο.
- Std. Dev (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης.
- Percent Missing: Ποσοστό των ελλείψεων.



Εικόνα 43. Γράφημα μέσων τιμών ANXMAT ανά χώρα και φύλο (R)



Εικόνα 44. Γράφημα μέσων τιμών ANXMAT ανά χώρα και φύλο (R)



Εικόνα 45. Γράφημα ποσοστών ANXMAT ανά χώρα και φύλο (R)

Επιπλέον, οι πληροφορίες που έχουμε από τις Εικόνες 43, 44 και 45 όπου ουσιαστικά απεικονίζουν διαγράμματα, είναι οι εξής: Στην Εικόνα 43 απεικονίζεται ένα Line chart που μας δείχνει τον μέσο όρο του άγχους για τα Μαθηματικά (ANXMAT) ανά φύλο (Female/Male) για Ελλάδα και Πορτογαλία. Στην Εικόνα 44 εμφανίζεται ένα ομαδοποιημένο ραβδόγραμμα (Bar chart) όπου δείχνει το μέσο όρο του ANXMAT ανά χώρα ξεχωριστά για τα αγόρια και τα κορίτσια και τέλος το διάγραμμα της Εικόνας 45 δείχνει ένα bar chart κατανομής φύλου (ποσοστά), στους μαθητές κάθε χώρας.

Ερμηνεία αποτελεσμάτων

Για την Ελλάδα, ο μέσος όρος του άγχους για τα Μαθηματικά (ANXMAT) ήταν υψηλότερος στα κορίτσια ($M = 0.47$, $SE = 0.02$) σε σύγκριση με τα αγόρια ($M = 0.04$, $SE = 0.03$). Οι τυπικές αποκλίσεις ($SD = 1.16$ για τα κορίτσια και $SD = 1.11$ για τα αγόρια) υποδηλώνουν επαρκή διασπορά γύρω από τους μέσους όρους. Η κατανομή φύλου ήταν σχεδόν ισόποση (50.71% κορίτσια, 49.29% αγόρια), γεγονός που μειώνει την πιθανότητα οι παρατηρούμενες διαφορές να οφείλονται σε ανισοκατανομή του δείγματος. Οπτικά, στο διάγραμμα (Εικόνα 43), η γραμμή της Ελλάδας παρουσίασε σαφή πτώση από τα κορίτσια προς τα αγόρια, αποτυπώνοντας την υψηλότερη μέση τιμή άγχους των κοριτσιών σε σχέση με τα αγόρια.

Στην Πορτογαλία, ο μέσος όρος του ANXMAT ήταν επίσης υψηλότερος στα κορίτσια ($M = 0.30$, $SE = 0.02$) σε σχέση με τα αγόρια ($M = -0.01$, $SE = 0.02$). Οι τυπικές αποκλίσεις ήταν ελαφρώς μικρότερες σε σύγκριση με την Ελλάδα ($SD = 1.01$ για τα κορίτσια και $SD = 1.03$ για τα αγόρια), υποδηλώνοντας μικρότερη διασπορά. Και εδώ, η κατανομή φύλου ήταν σχεδόν ισόποση (48.98% κορίτσια, 51.02% αγόρια). Στην Εικόνα 43, η γραμμή της Πορτογαλίας εμφάνισε επίσης πτωτική πορεία από τα κορίτσια προς τα αγόρια, ωστόσο η διαφορά μεταξύ φύλων ήταν ηπιότερη σε σχέση με την Ελλάδα.

Η συγκριτική απεικόνιση στο ομαδοποιημένο ραβδόγραμμα (Εικόνα 44) έδειξε ότι, και για τα δύο φύλα, τα επίπεδα άγχους για τα Μαθηματικά ήταν υψηλότερα στην Ελλάδα από ό,τι στην Πορτογαλία. Επιπλέον, το διάγραμμα ποσοστών (Εικόνα 45) κατέδειξε ότι η συμμετοχή αγοριών και κοριτσιών ήταν σχεδόν ισότιμη και στις δύο χώρες, στοιχείο που ενισχύει την ερμηνεία ότι οι διαφοροποιήσεις στο ANXMAT αντανακλούν ουσιαστικές διαφορές μεταξύ φύλων και χωρών. Συνολικά, τα αποτελέσματα

υποδεικνύουν ότι οι μαθήτριες παρουσιάζουν υψηλότερα επίπεδα άγχους για τα Μαθηματικά από τους μαθητές τόσο στην Ελλάδα όσο και στην Πορτογαλία, με τη διαφορά να είναι εντονότερη στην Ελλάδα. Παράλληλα, το μέσο άγχος ήταν υψηλότερο στην Ελλάδα σε σχέση με την Πορτογαλία, ανεξαρτήτως φύλου.

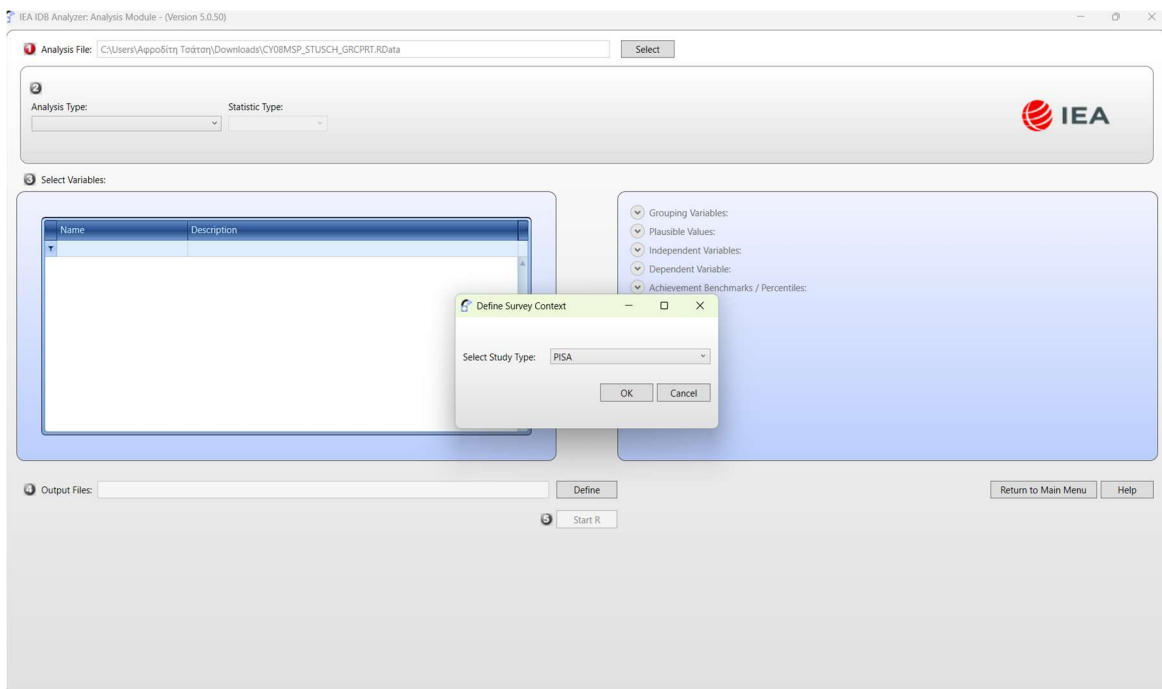
Υπολογισμός ποσοστών

Ερευνητική υπόθεση: Διερευνώνται διαφορές στα ποσοστά προτίμησης για τα Μαθηματικά μεταξύ των δύο χωρών που επιλέχθηκαν.

Επιλογή μεταβλητής: MATHPREF

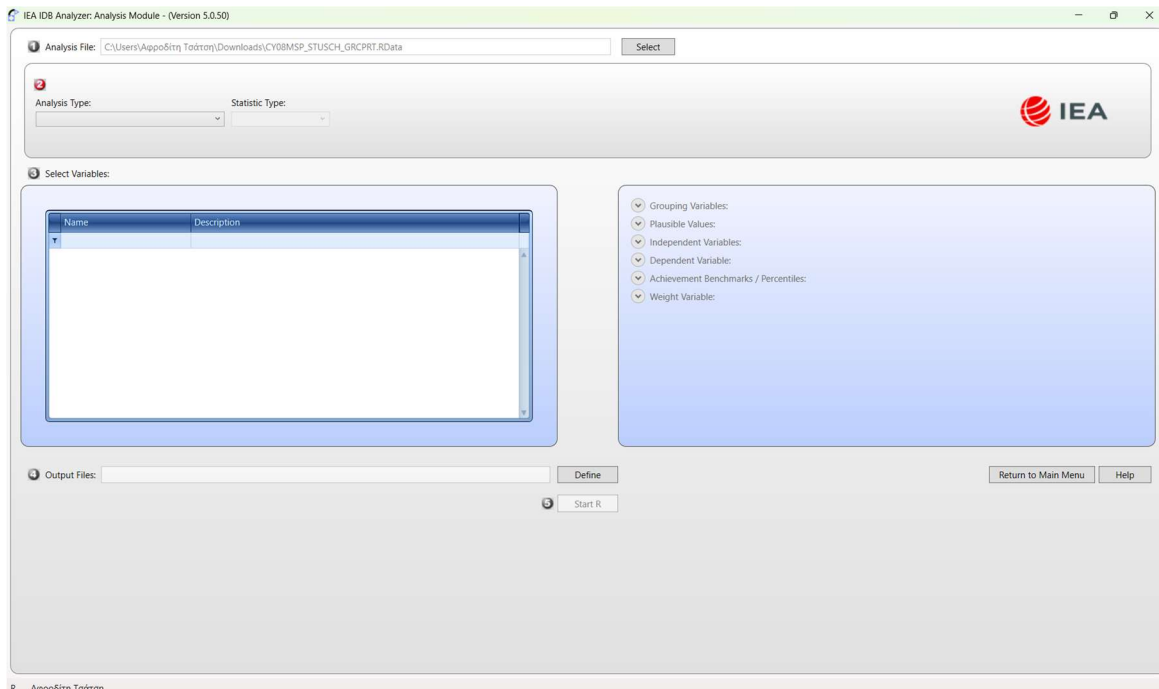
Χαρακτηριστικά μεταβλητής: MATHPREF συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί σύγκριση ποσοστών ανάμεσα στην Ελλάδα και την Πορτογαλία, με μεταβλητή ενδιαφέροντος τη προτίμηση των μαθητών για τα Μαθηματικά, σε σχέση με τα άλλα μαθήματα. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTR.RData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 46).



Εικόνα 46. Επιλογή αρχείου δεδομένων για ανάλυση Percentages Only – «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 47).

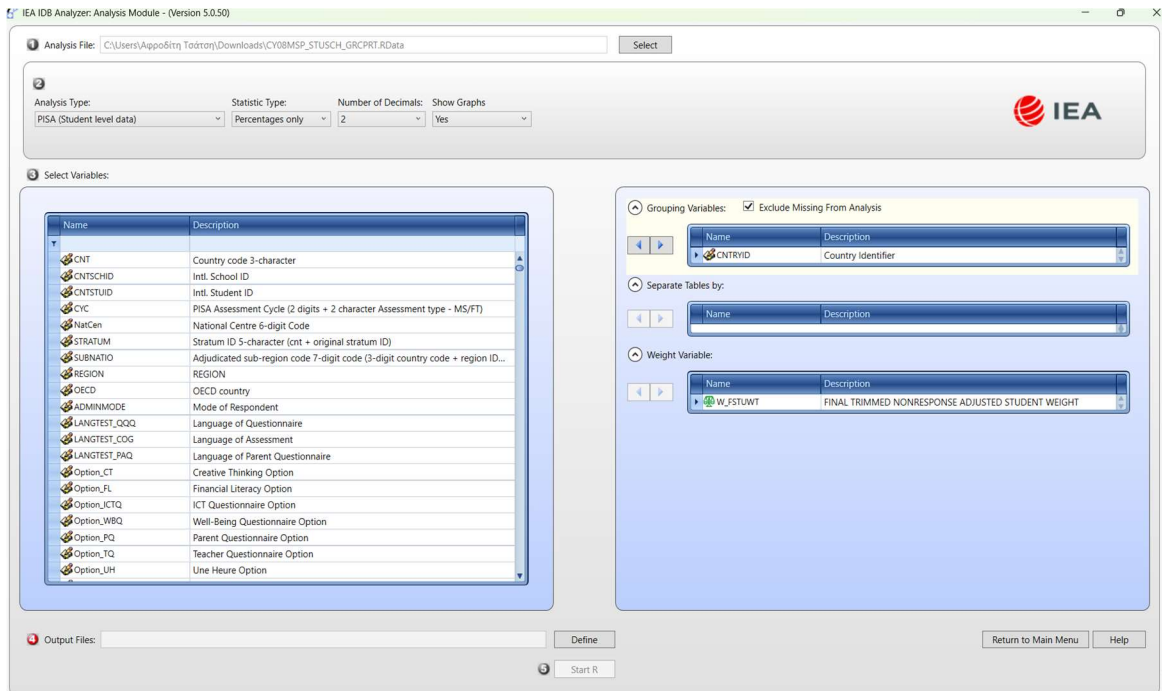


Εικόνα 47. Επιλογή «PISA (Student level)» – Ανάλυση Percentages Only (R)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

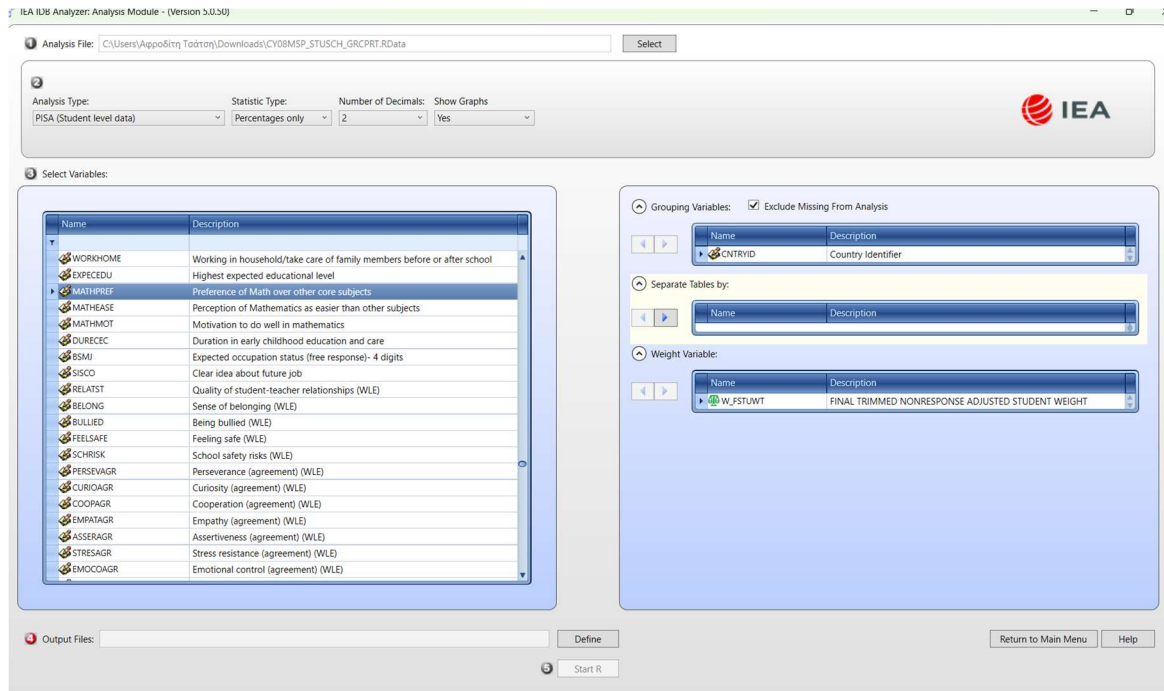
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Percentages only». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 48).



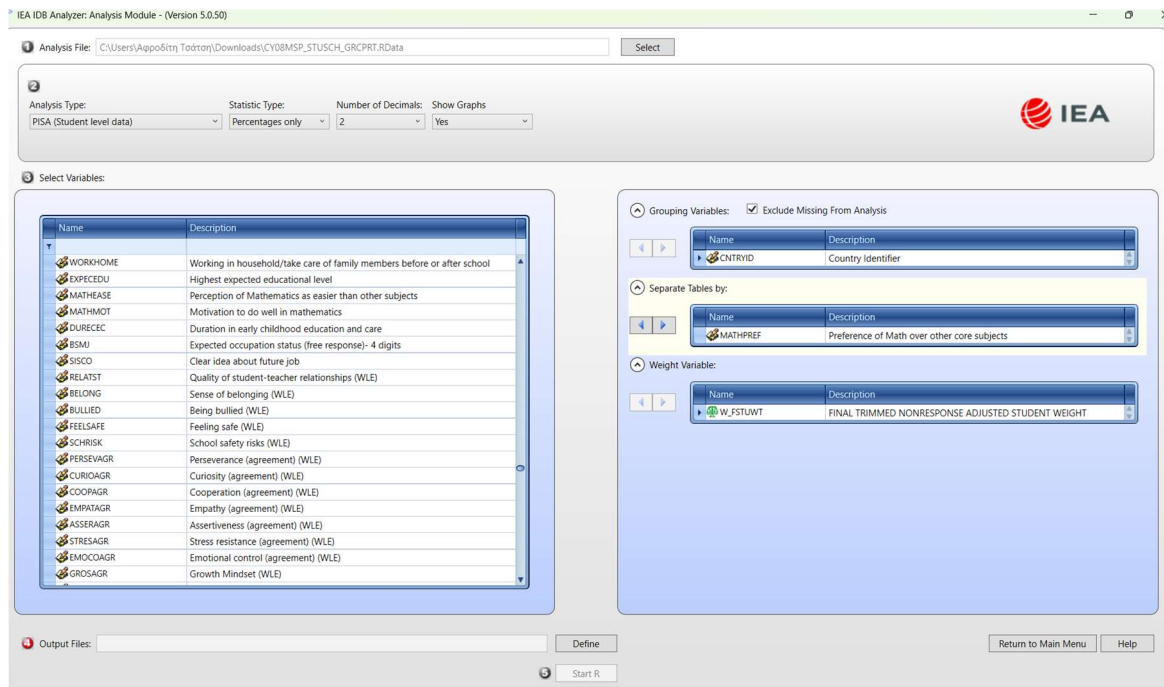
Εικόνα 48. Επιλογή «Percentages only» στο πεδίο «Statistic Type» (R)

Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 48 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλείψεις τιμές στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί η σύγκριση, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή «MATHPREF» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Separate Tables by» την τοποθετείτε. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 49).

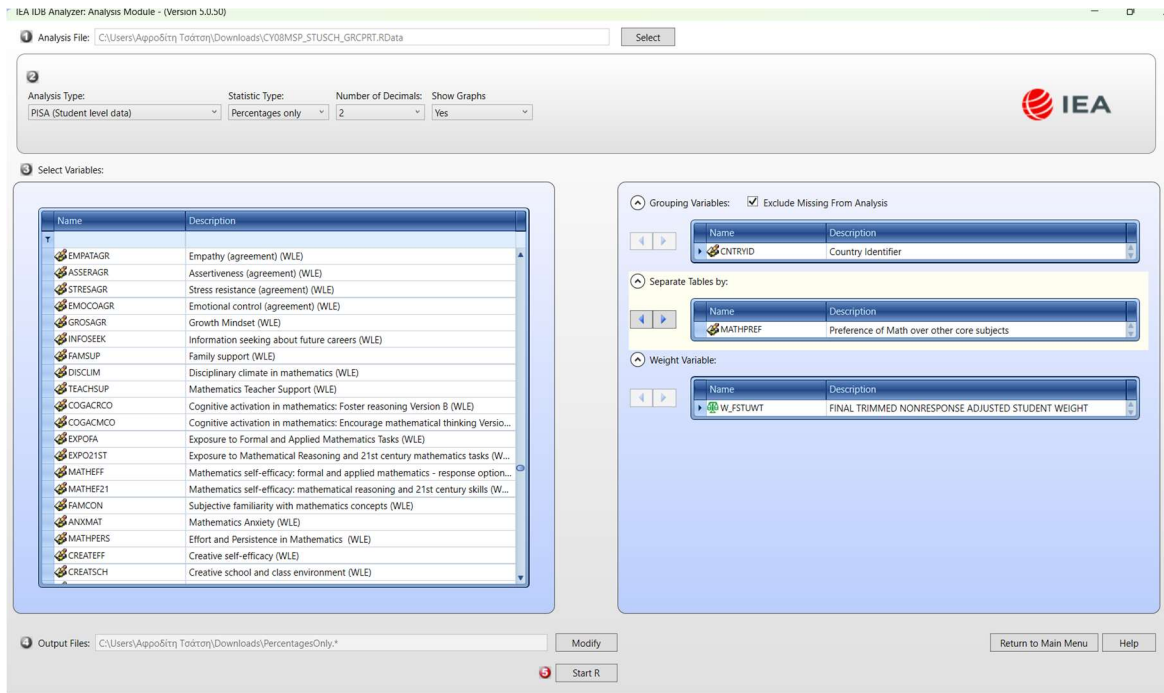


Εικόνα 49. Τοποθέτηση μεταβλητής MATHPREF στο πεδίο «Separate Tables by» (R)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια (Εικόνα 50). Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PercentagesOnly». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio (Εικόνα 51).

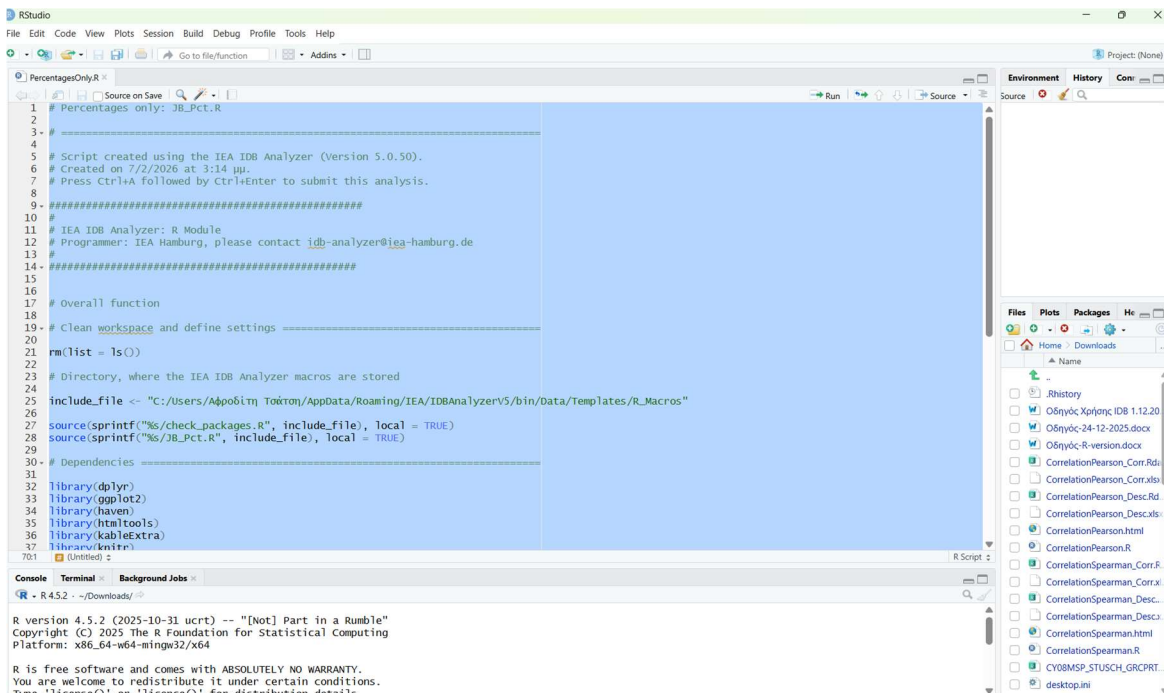


Εικόνα 50. Ορισμός ονόματος εξόδου για ανάλυση Percentages Only (R)



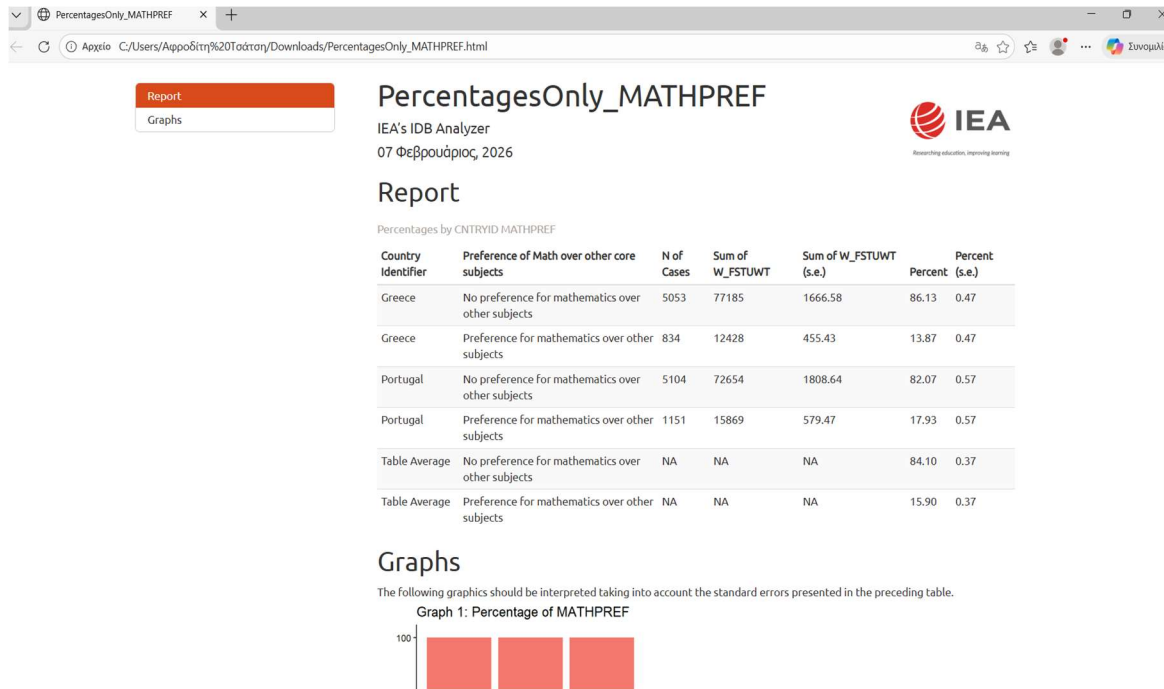
Εικόνα 51. Εκκίνηση παραγωγής script μέσω «Start R» – Percentages Only

Ύστερα από τη μετάβαση στο Rstudio, στο αρχείο PercentagesOnly.R που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 52).



Εικόνα 52. Εκτέλεση κώδικα PercentagesOnly.R στο RStudio

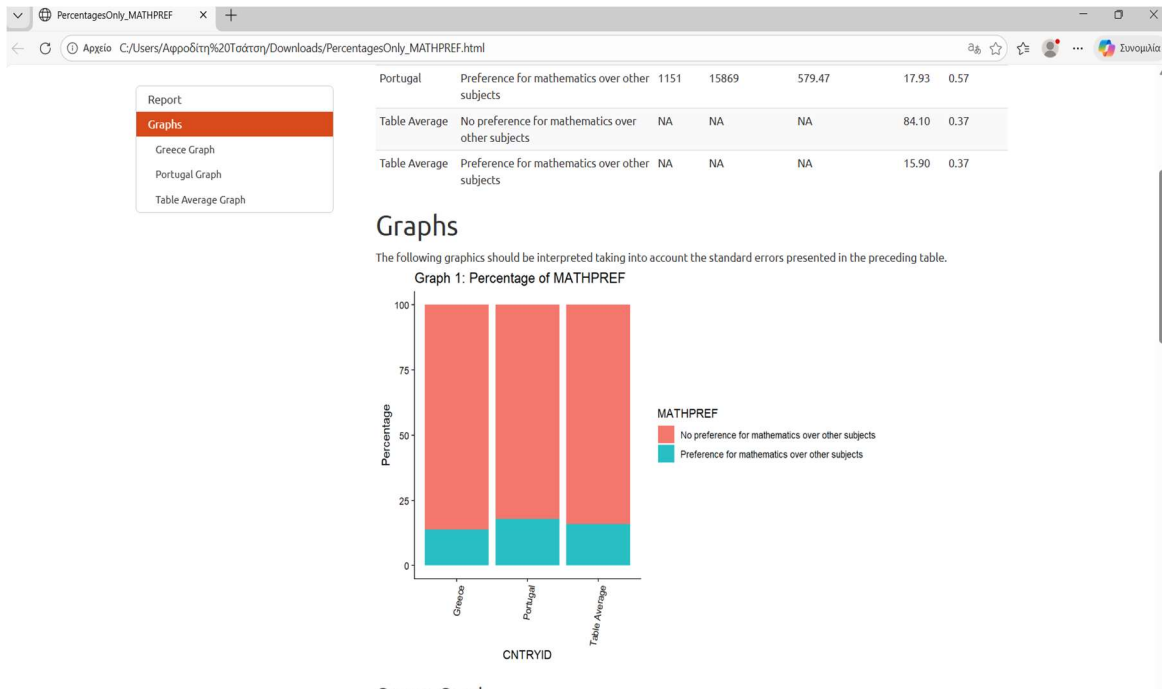
Στο αρχείο PercentagesOnly, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 53 και 54, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



Εικόνα 53. Πίνακας ποσοστών προτίμησης για τα Μαθηματικά ανά χώρα (R)

Οι στήλες της Εικόνας 53 παρουσιάζουν την κατανομή των μαθητών ως προς την προτίμησή τους για τα Μαθηματικά σε σύγκριση με άλλα βασικά μαθήματα, ανά χώρα, αποτυπώνοντας τα αντίστοιχα ποσοστά στον μαθητικό πληθυσμό. Συγκεκριμένα:

- Country Identifier: Η χώρα (π.χ. Ελλάδα, Πορτογαλία).
- Preference of Math over other core subjects: Αν ο μαθητής προτιμά τα Μαθηματικά ή όχι σε σχέση με άλλα βασικά μαθήματα.
- N of Cases: Πόσοι μαθητές του δείγματος ανήκουν σε αυτή την κατηγορία.
- Sum of W_FSTUWT: Το σταθμισμένο πλήθος μαθητών που αντιπροσωπεύουν τον συνολικό πληθυσμό.
- Sum of W_FSTUWT (s.e.): Το τυπικό σφάλμα.
- Percent: Το ποσοστό (%) των μαθητών που ανήκουν στην αντίστοιχη κατηγορία.
- Percent (s.e.): Το τυπικό σφάλμα του ποσοστού.



Εικόνα 54. Γράφημα κατανομής MATHPREF ανά χώρα – Percentages Only (R)

Το γράφημα που απεικονίζεται στην Εικόνα 54 παριστάνει την κατανομή των μαθητών ως προς την προτίμησή τους για τα Μαθηματικά για τις δύο χώρες.

Ερμηνεία αποτελεσμάτων

Παρουσιάζονται τα αποτελέσματα της ανάλυσης ποσοστών (percentages) για τη μεταβλητή MATHPREF, η οποία αποτυπώνει εάν οι μαθητές δηλώνουν προτίμηση για τα Μαθηματικά σε σχέση με άλλα μαθήματα. Τα ποσοστά υπολογίστηκαν με χρήση του βάρους W_FSTUWT , ώστε οι εκτιμήσεις να είναι αντιπροσωπευτικές του μαθητικού πληθυσμού. Για την Ελλάδα, η πλειονότητα των μαθητών δεν δήλωσε προτίμηση για τα Μαθηματικά έναντι άλλων βασικών μαθημάτων. Συγκεκριμένα, το 86.13% (SE = 0.47) ανήκε στην κατηγορία No preference for mathematics over other subjects, ενώ το 13.87% (SE = 0.47) δήλωσε Preference for mathematics over other subjects. Αντίστοιχα, στην Πορτογαλία παρατηρήθηκε παρόμοιο μοτίβο, με τη μη προτίμηση να υπερισχύει. Ωστόσο, το ποσοστό των μαθητών που δήλωσαν προτίμηση για τα Μαθηματικά ήταν υψηλότερο σε σχέση με την Ελλάδα: 82.07% (SE = 0.57) ανέφερε No preference for mathematics over other subjects και 17.93% (SE = 0.57) ανέφερε Preference for mathematics over other subjects. Συνολικά, και στις δύο χώρες υπερισχύει η μη προτίμηση για τα Μαθηματικά, αν και η Πορτογαλία εμφάνισε μεγαλύτερο ποσοστό μαθητών που προτιμούν τα Μαθηματικά σε σύγκριση με την Ελλάδα (διαφορά = 4.06 ποσοστιαίες μονάδες).

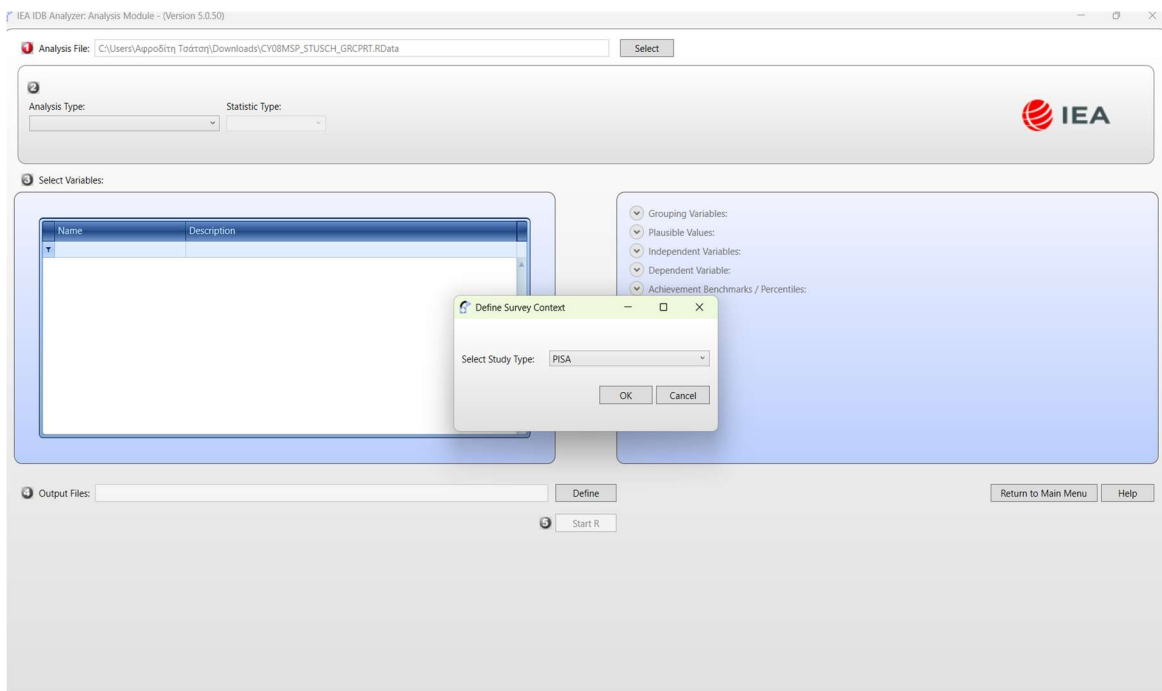
Υπολογισμός Ποσοστημορίων (Percentiles)

Ερευνητική Υπόθεση: Διερευνώνται διαφορές στα ποσοστημόρια του άγχους για τα Μαθηματικά μεταξύ των δύο χωρών, ώστε να εξεταστεί αν η κατανομή του άγχους διαφοροποιείται σε διαφορετικά επίπεδα της κατανομής.

Επιλογή μεταβλητής: ANXMAT

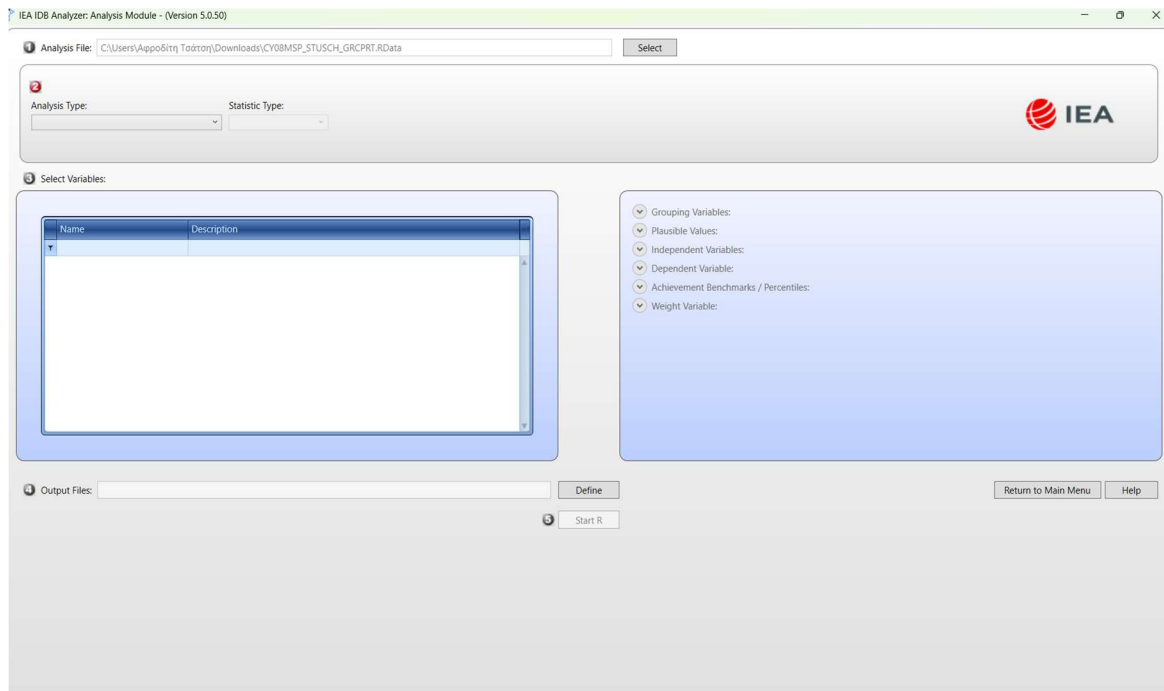
Χαρακτηριστικά μεταβλητής: ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί υπολογισμός των ποσοστημορίων για τη μεταβλητή ANXMAT, η οποία αποτυπώνει το άγχος των μαθητών απέναντι στα Μαθηματικά, προκειμένου να ελεγχθούν οι διαφορές στην κατανομή του άγχους για τα Μαθηματικά για τις δύο χώρες που επιλέχθηκαν. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTRData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 55).



Εικόνα 55. Επιλογή αρχείου δεδομένων για ανάλυση Percentiles – «Analysis File» (R)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 56).

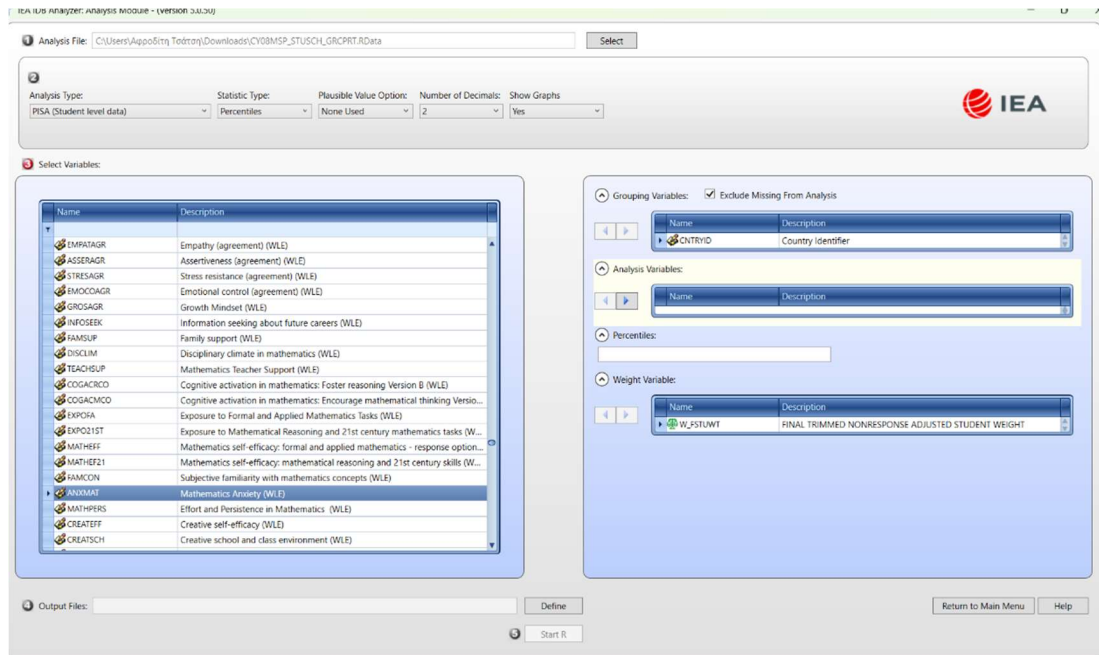


Εικόνα 56. Επιλογή «PISA (Student level)» – Ανάλυση Percentiles (R)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

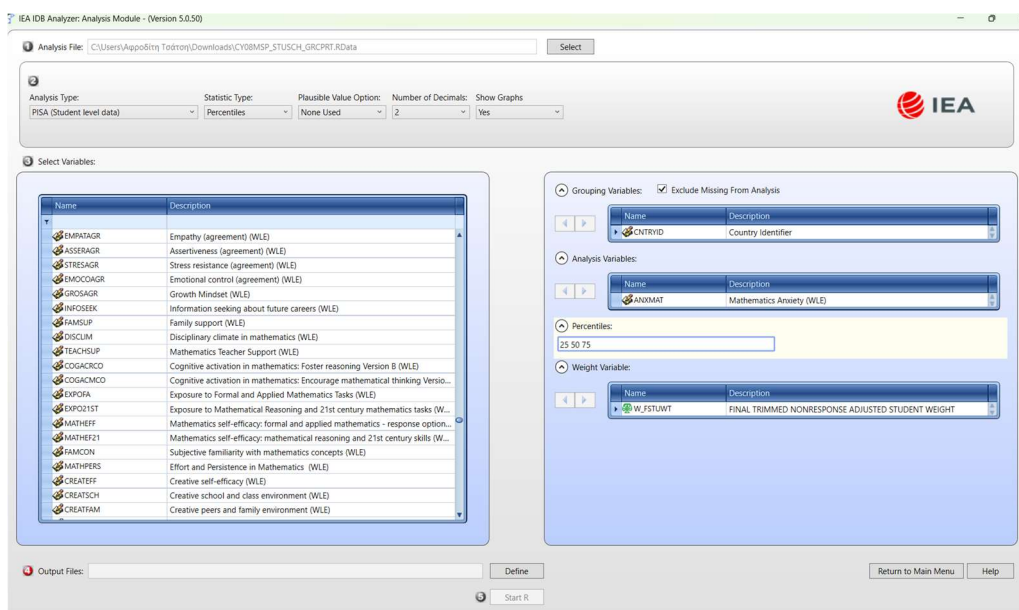
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλিপών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγεται «Percentiles». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων (Εικόνα 57).



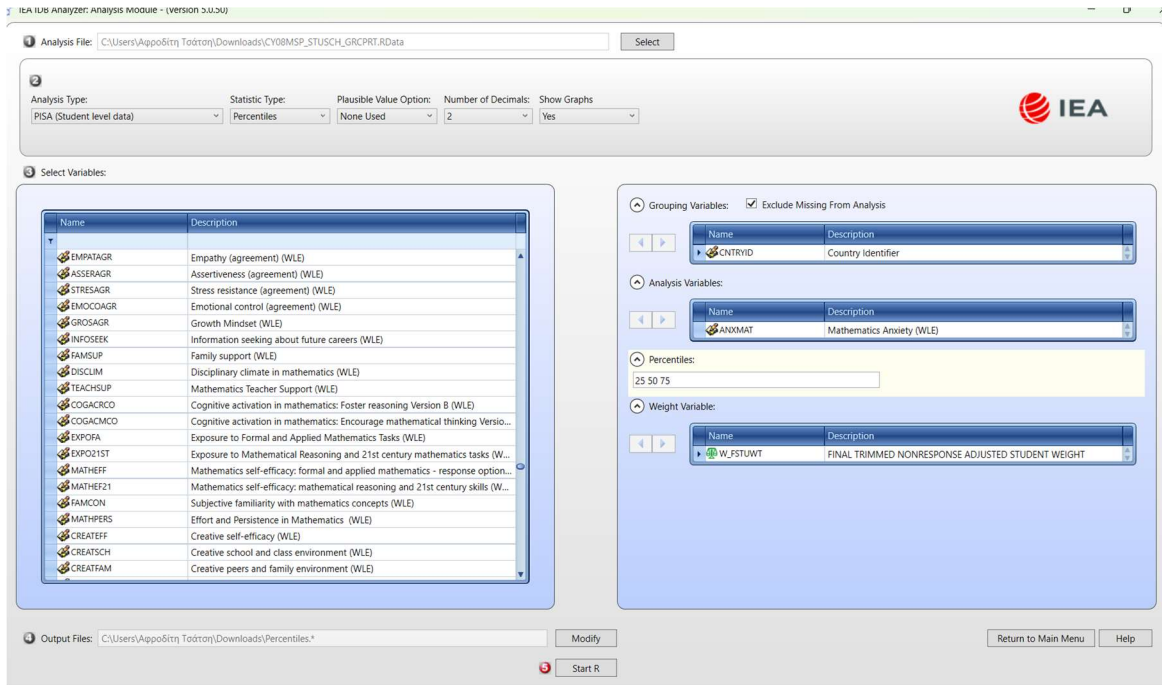
Εικόνα 57. Επιλογή «Percentiles» στο πεδίο «Statistic Type» και ορισμός ποσοστημορίων (R)

Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 57 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλειπείς τιμές στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί ο υπολογισμός των ποσοστημορίων για την μεταβλητή που επιλέγετε, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή την μεταβλητή «ANXMAT» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» και την τοποθετείτε (Εικόνα 58). Στην συνέχεια στο πλαίσιο «Percentiles» εισάγουμε τα ποσοστημόρια που θέλουμε να μελετήσουμε και συγκεκριμένα το 25ο, το 50ο (διάμεσος) και το 75ο ποσοστημόριο. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 58).



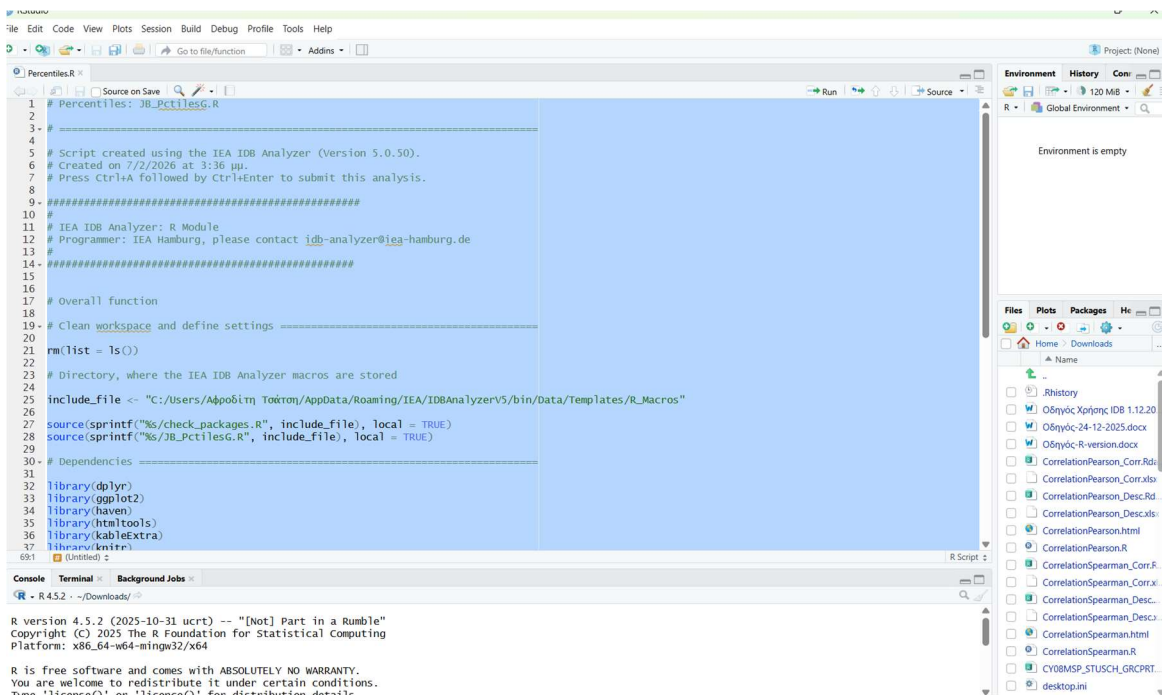
Εικόνα 58. Τοποθέτηση μεταβλητής ANXMAT και εισαγωγή τιμών ποσοστημορίων (p25, p50, p75) (R)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια (Εικόνα 59). Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «Percentiles». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio (Εικόνα 59).



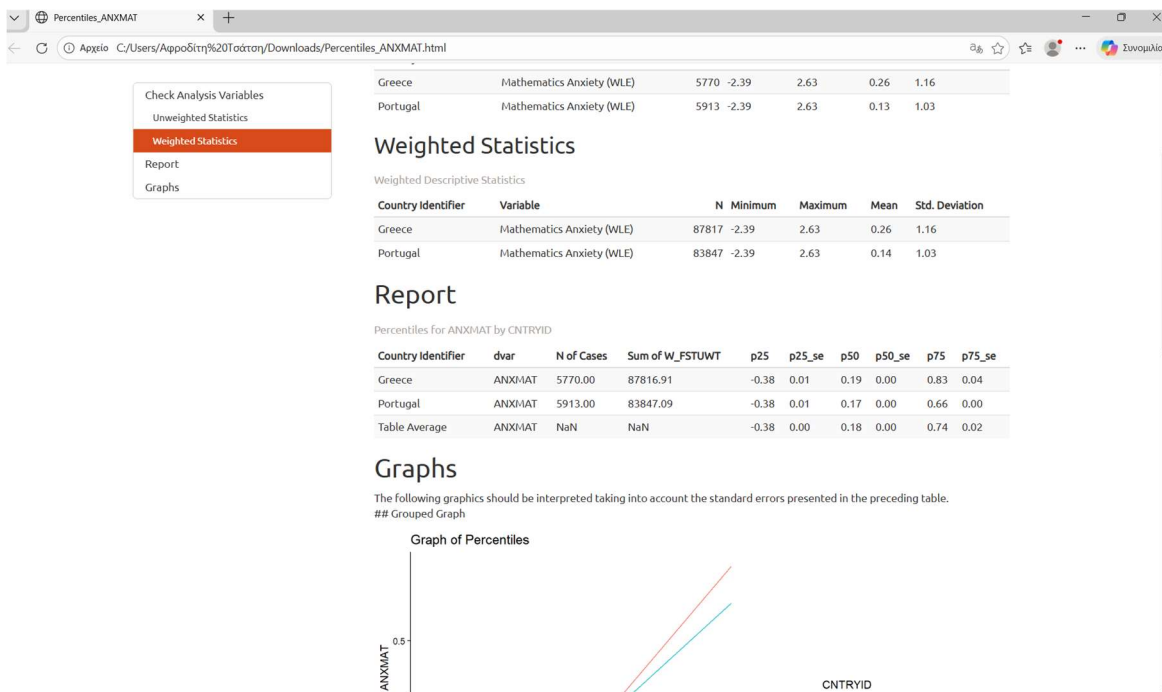
Εικόνα 59. Ορισμός ονόματος εξόδου και Εκκίνηση παραγωγής script αρχείου Percentiles μέσω «Start R»

Αφού μεταβείτε στο περιβάλλον του Rstudio, συγκεκριμένα στο αρχείο Percentiles.R που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 60).



Εικόνα 60. Εκτέλεση κώδικα Percentiles.R στο RStudio

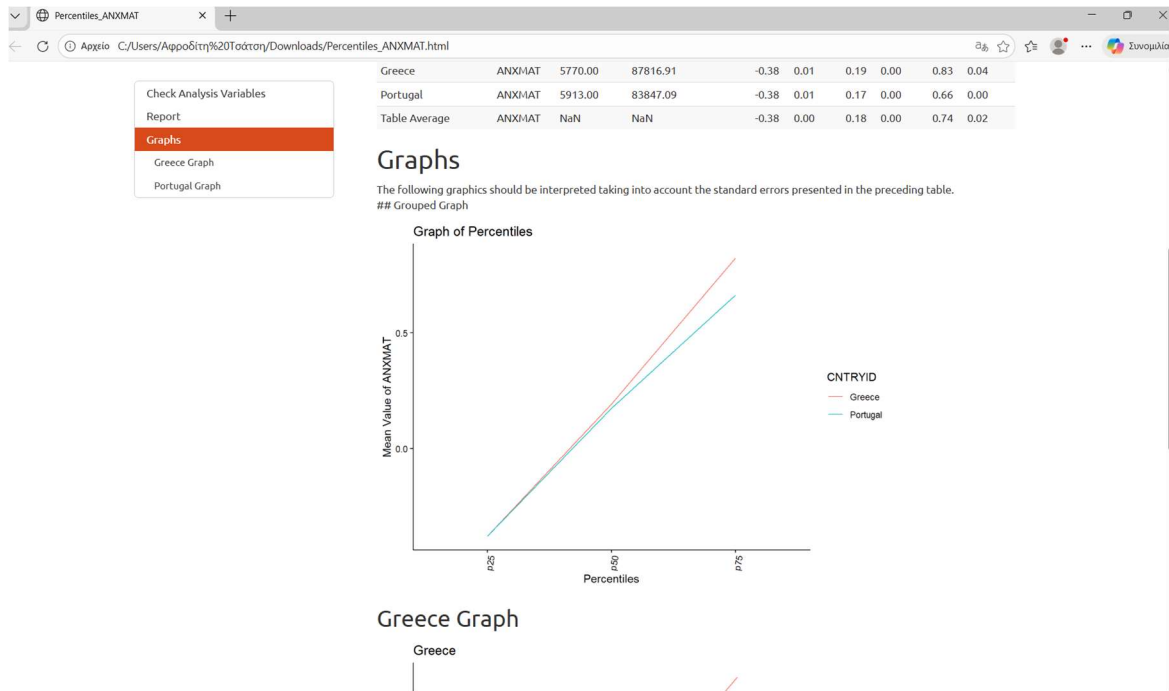
Στο αρχείο Percentiles, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 61 και 62, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



Εικόνα 61. Πίνακας ποσοστημορίων p_{25} , p_{50} , p_{75} της ANXMAT ανά χώρα (R)

Παρακάτω αναλύεται το περιεχόμενο των στηλών της Εικόνας 61, και πιο συγκεκριμένα:

- Country Identifier: Η χώρα στην οποία αναφέρονται τα αποτελέσματα (π.χ. Ελλάδα, Πορτογαλία).
- dvar: Το όνομα της μεταβλητής ενδιαφέροντος.
- N of Cases: Ο αριθμός των μαθητών που συμπεριλήφθηκαν στην ανάλυση για κάθε χώρα.
- Sum of W_FSTUWT: Το άθροισμα των μαθητικών σταθμίσεων (weights), το οποίο αντιπροσωπεύει το εκτιμώμενο πληθυσμιακό μέγεθος για τη συγκεκριμένη χώρα.
- p25: Το 25ο ποσοστιαίο σημείο (1ο τεταρτημόριο) της μεταβλητής ANXMAT.
- p25_se: Το τυπικό σφάλμα της εκτίμησης του 25ου ποσοστημορίου.
- p50: Η διάμεσος (50ο ποσοστιαίο σημείο) της ANXMAT.
- p50_se: Το τυπικό σφάλμα της εκτίμησης της διαμέσου.
- p75: Το 75ο ποσοστιαίο σημείο (3ο τεταρτημόριο) της ANXMAT.
- p75_se: Το τυπικό σφάλμα της εκτίμησης του 75ου ποσοστημορίου.



Εικόνα 62. Γράφημα κατανομής ποσοστημορίων ANXMAT – Ελλάδα vs Πορτογαλία (R)

Επιπλέον, το γράφημα της Εικόνας 62 απεικονίζει την κατανομή του άγχους για τα Μαθηματικά (ANXMAT) στα ποσοστημόρια για την Ελλάδα και την Πορτογαλία,

Ερμηνεία αποτελεσμάτων

Με βάση τα αποτελέσματα των ποσοστημορίων για τη μεταβλητή ANXMAT (άγχος για τα Μαθηματικά), παρατηρούνται τα εξής. Για την Ελλάδα, το 25% των μαθητών έχει τιμή ANXMAT κάτω από -0.38 ($p_{25} = -0.38$, $SE = 0.01$), το 50% των μαθητών έχει τιμή έως 0.19 ($p_{50} = 0.19$, $SE = 0.00$) και το 75% των μαθητών έχει τιμή κάτω από 0.83 ($p_{75} = 0.83$, $SE = 0.04$). Όσον αφορά την Πορτογαλία, το 25% των μαθητών έχει επίσης τιμή ANXMAT κάτω από -0.38 ($p_{25} = -0.38$, $SE = 0.01$), το 50% έχει τιμή έως 0.17 ($p_{50} = 0.17$, $SE = 0.00$) και το 75% έχει τιμή κάτω από 0.66 ($p_{75} = 0.66$, $SE = 0.00$). Συνεπώς, συγκεντρωτικά φαίνεται ότι το κάτω μέρος της κατανομής (p_{25}) είναι ίδιο στις δύο χώρες, ενώ από τη διάμεσο και πάνω η Ελλάδα εμφανίζει υψηλότερες τιμές άγχους. Η διαφορά είναι πιο έντονη στο άνω 25% της κατανομής, όπου το p_{75} στην Ελλάδα (0.83) είναι υψηλότερο από το p_{75} στην Πορτογαλία (0.66) κατά περίπου 0.17 μονάδες. Αυτό επιβεβαιώνεται και από το γράφημα, όπου η γραμμή της Ελλάδας βρίσκεται πάνω από της Πορτογαλίας στα υψηλότερα ποσοστημόρια, δείχνοντας ότι οι μαθητές με μεγαλύτερο άγχος στην Ελλάδα τείνουν να εμφανίζουν ακόμη υψηλότερα επίπεδα άγχους σε σχέση με τους αντίστοιχους μαθητές στην Πορτογαλία.

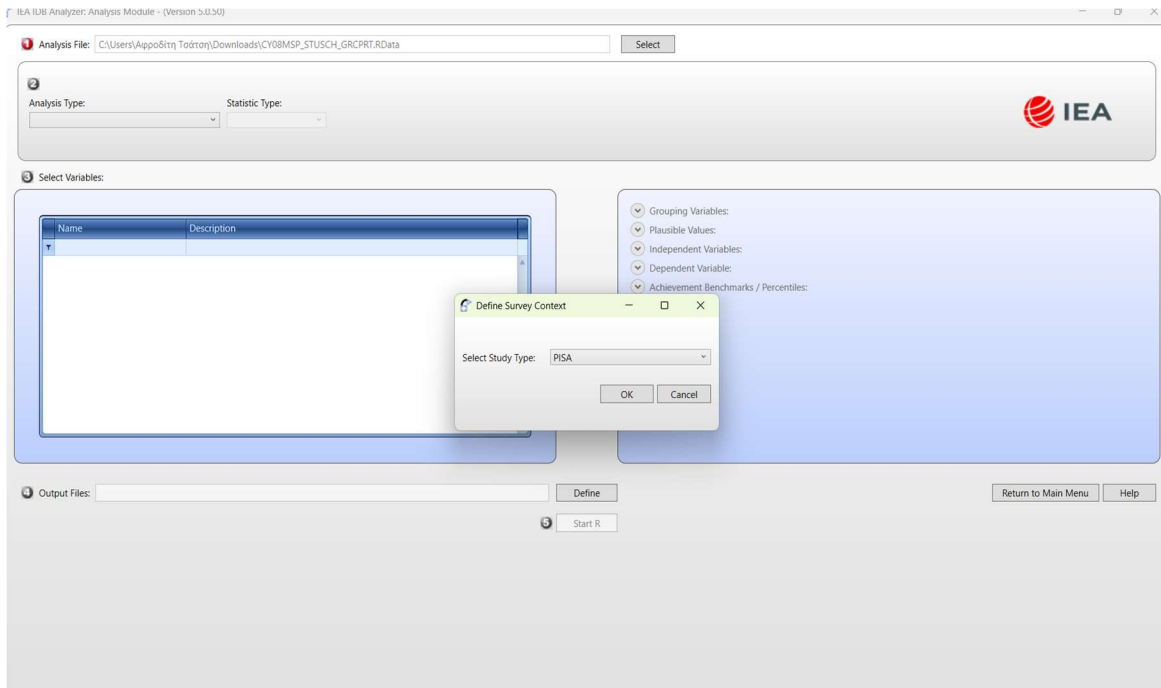
Μελέτη επιδόσεων με βάση τα ορόσημα (Benchmarks)

Ερευνητική Υπόθεση: Διερευνώνται διαφορές στα ποσοστά των μαθητών που επιτυγχάνουν τα καθορισμένα ορόσημα μαθηματικής επίδοσης μεταξύ των χωρών που επιλέχθηκαν

Επιλογή μεταβλητών: PVMATH01-10

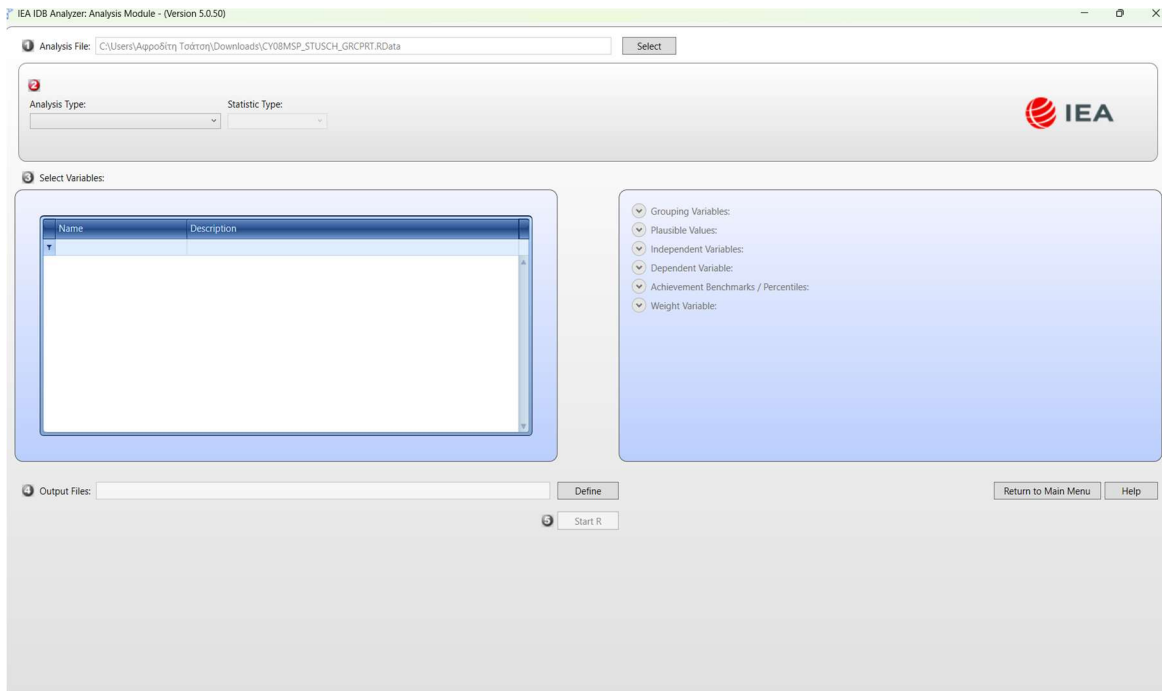
Χαρακτηριστικά μεταβλητών: PVMATH01-10 συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση των επιδόσεων των μαθητών με βάση τα ορόσημα (Benchmarks) που πετυχαίνουν στα μαθηματικά. Η μεταβλητή όπου καταγράφεται η επίδοση των μαθητών, με την χρήση των Plausible Values, είναι η «PVMATH01-10». Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRTR.RData» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 63).



Εικόνα 63. Επιλογή αρχείου δεδομένων για ανάλυση Benchmarks – «Analysis File» (R)

Έπειτα στο πεδίο «Analysis Type» επιλέξτε τον τύπο της ανάλυσης που θέλετε να πραγματοποιήσετε, για παράδειγμα «Pisa (Student level)» (Εικόνα 64).



Εικόνα 64. Επιλογή «PISA (Student level)» – Ανάλυση Benchmarks (R)

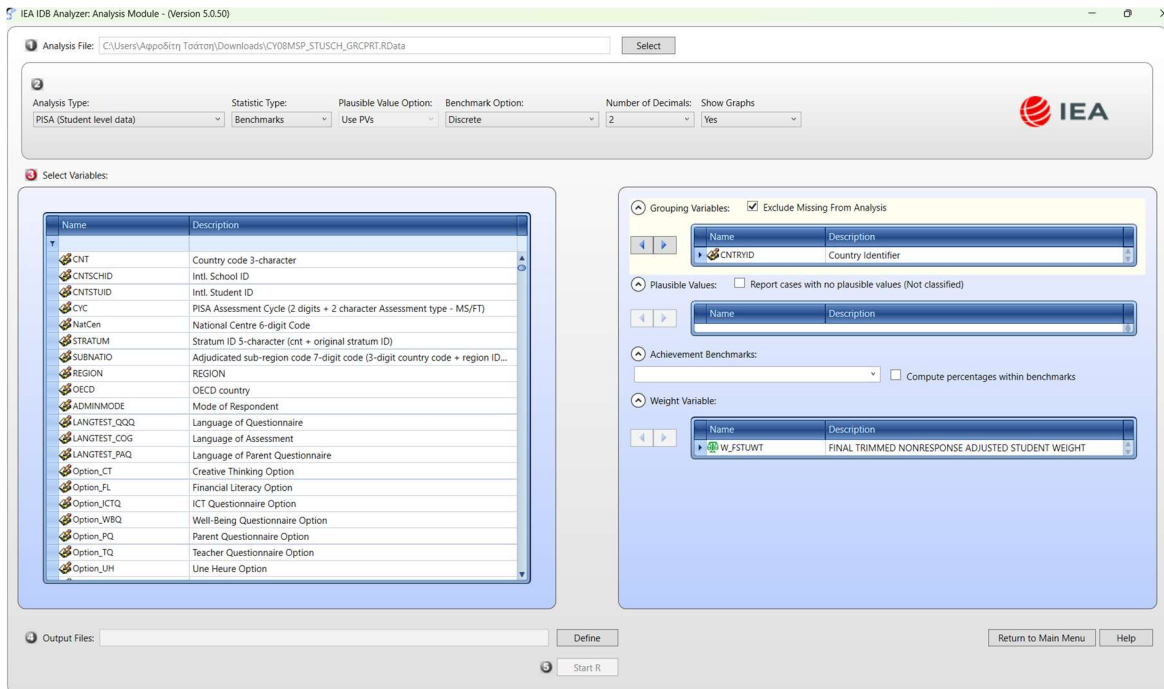
Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

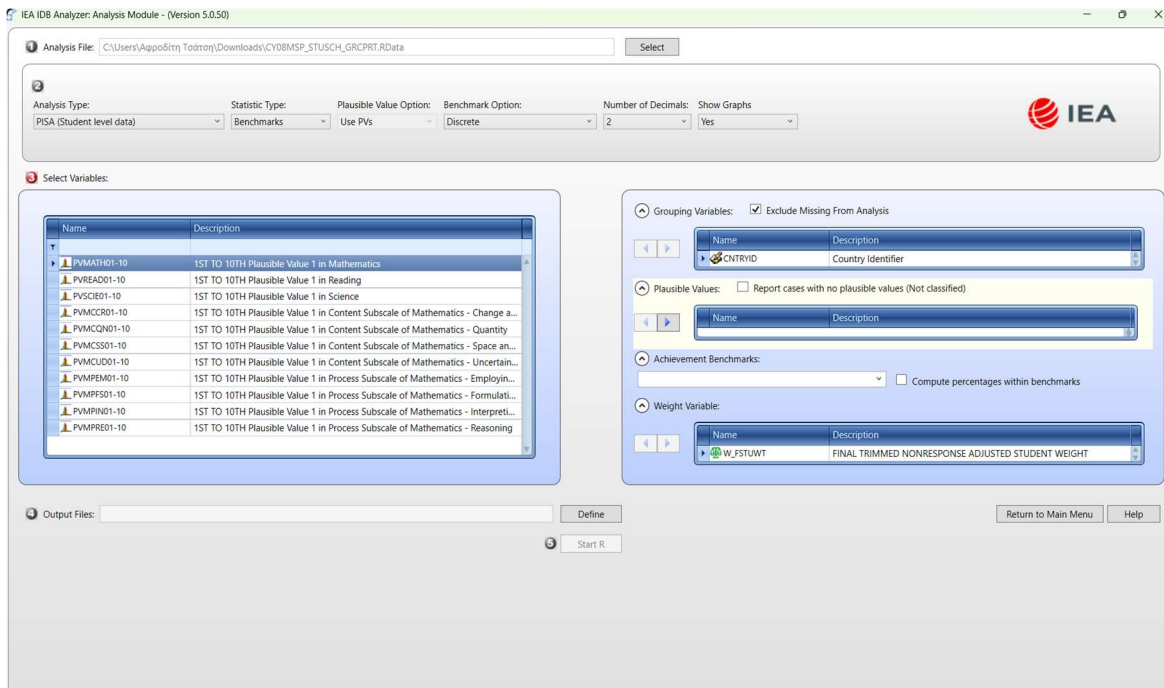
Επιλέγετε «Benchmarks». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «Use PVs». Επιπλέον στο πεδίο «Benchmark Option» υπάρχουν διαθέσιμες οι επιλογές:

- Cumulative: Υπολογίζεται το ποσοστό των ατόμων που βρίσκονται «στο συγκεκριμένο όριο ή πάνω από αυτό».
- Discrete: Υπολογίζεται το ποσοστό των ατόμων μέσα σε κάθε κατηγορία που ορίζεται από τα όρια επίδοσης (τα cut-off points λειτουργούν ως κατώτερο όριο κάθε ομάδας).
- Discrete with Analysis Variable(s): Εκτός από τα ποσοστά ανά κατηγορία επίδοσης, υπολογίζεται και τον μέσο όρο μιας (ή περισσότερων) μεταβλητών ανάλυσης μέσα σε κάθε κατηγορία (π.χ. μέσος όρος «κίνησης για διάβασμα» ανά επίπεδο επίδοσης).

Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων και στο πεδίο «Show Graphs» αν επιθυμείτε να εμφανιστούν γραφήματα στο τελικό αποτέλεσμα (Εικόνα 65).



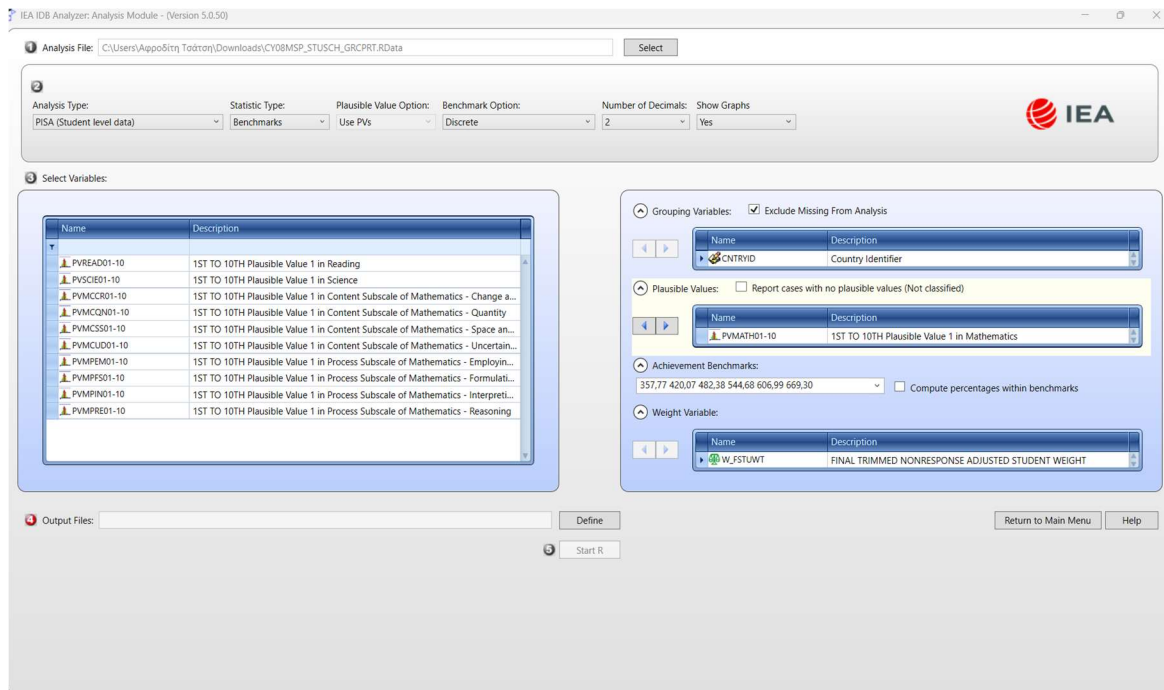
Εικόνα 65. Επιλογή «Benchmarks» στο πεδίο «Statistic Type» (R)



Εικόνα 66. Ρύθμιση «Plausible Value Option: Use PVs» (R)

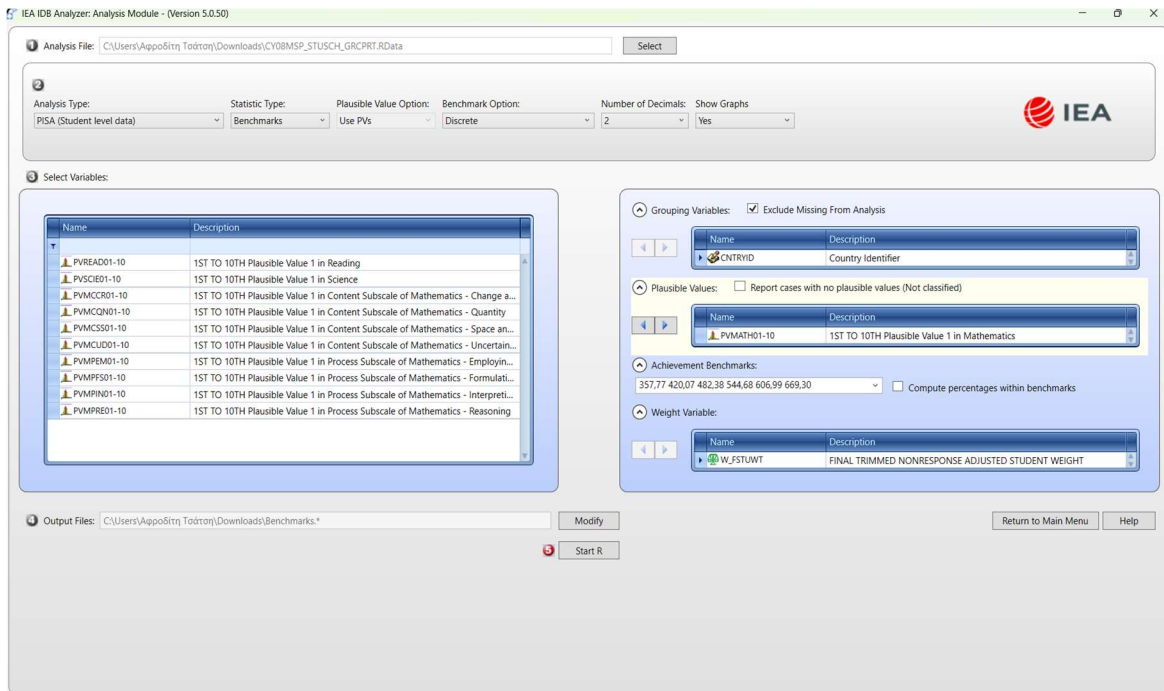
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 65 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλιπείς τιμές στις μεταβλητές ομαδοποίησης. Στη συνέχεια, εφόσον ο σκοπός είναι να αναλυθεί η επίδοση των μαθητών στα μαθηματικά, θα πρέπει να επιλέξετε τη μεταβλητή “PVMATH01-10” και να την

τοποθετήσετε στο πλαίσιο “Plausible Values” (Εικόνα 66). Έπειτα, στο πλαίσιο “Achievement Benchmarks” να εισάγετε τις τιμές που θεωρούμε ως ορόσημα (benchmarks). Εδώ εισάγονται οι τιμές 357,77 / 420,07 / 482,54 / 544,60 / 606,91 / 669,33 που αφορούν τα μαθηματικά. Ωστόσο, επειδή το IEA IDB Analyzer δεν έχει ενημερώσει την εφαρμογή με τα νέα cut-off points, θα πρέπει να προστεθούν χειροκίνητα, όπως ορίζονται στη βιβλιογραφία του εκάστοτε κύκλου ανάλυσης. Τα νέα cut-off points είναι 233,17 / 295,47 / 357,77 / 420,07 / 482,54 / 544,60 / 606,91 / 669,33.. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 67).



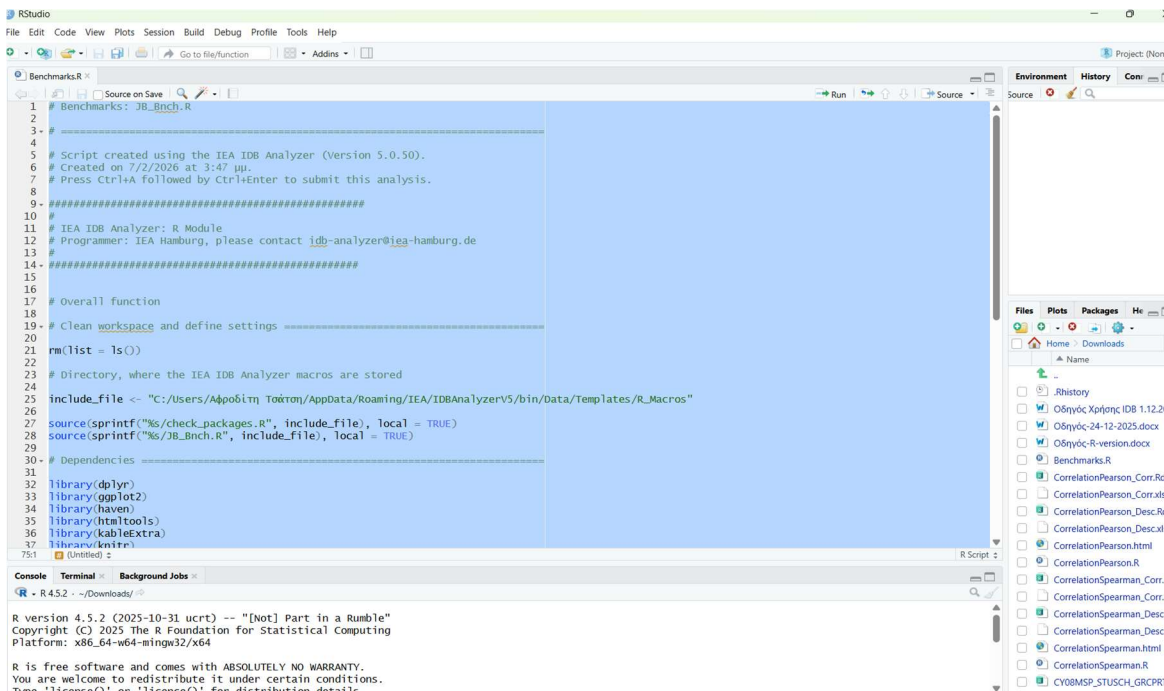
Εικόνα 67. Τοποθέτηση PVMATH01-10 στο πεδίο «Plausible Values» και εισαγωγή ορίων επίδοσης (R)

Το επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «Benchmarks». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start R», για να δημιουργηθεί το script αρχείο στο Rstudio. (Εικόνα 68).



Εικόνα 68. Ορισμός ονόματος εξόδου και Εκκίνηση παραγωγής script αρχείου Benchmarks μέσω «Start R»

Αφού μεταβείτε στο περιβάλλον του Rstudio, συγκεκριμένα στο αρχείο Benchmarks.R που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί “Run” για να εκτελέσετε τις εντολές (Εικόνα 69).



Εικόνα 69. Εκτέλεση κώδικα Benchmarks.R στο RStudio

Στο αρχείο Benchmarks, μεταξύ άλλων πληροφοριών, θα παρατηρήσετε τις Εικόνες 70 και 71, οι οποίες παρουσιάζουν το HTML output που παράγεται μετά την εκτέλεση του κώδικα στο RStudio και απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

Report
Percentage by Performance Group of PVMATH

Country Identifier	cutvar	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent (s.e.)
Greece	1. Below 357.77	1219	20001	1086.99	20.39 1.02
Greece	2. From 357.77 to Below 420.07	1693	26303	929.29	26.82 0.77
Greece	3. From 420.07 to Below 482.38	1702	25511	877.76	26.01 0.77
Greece	4. From 482.38 to Below 544.68	1154	16971	731.03	17.30 0.70
Greece	5. From 544.68 to Below 606.99	502	7341	510.04	7.48 0.52
Greece	6. From 606.99 to Below 669.3	124	1814	271.91	1.85 0.28
Greece	7. At or Above 669.3	10	146	68.74	0.15 0.07
Portugal	1. Below 357.77	627	10039	914.29	10.39 0.87
Portugal	2. From 357.77 to Below 420.07	1265	18670	868.56	19.33 0.73
Portugal	3. From 420.07 to Below 482.38	1714	24167	910.38	25.02 0.76
Portugal	4. From 482.38 to Below 544.68	1626	22189	842.86	22.97 0.78

Εικόνα 70. Πίνακας ποσοστών μαθητών ανά επίπεδο επίδοσης (Benchmarks) ανά χώρα (R)

Report
Percentage by Performance Group of PVMATH

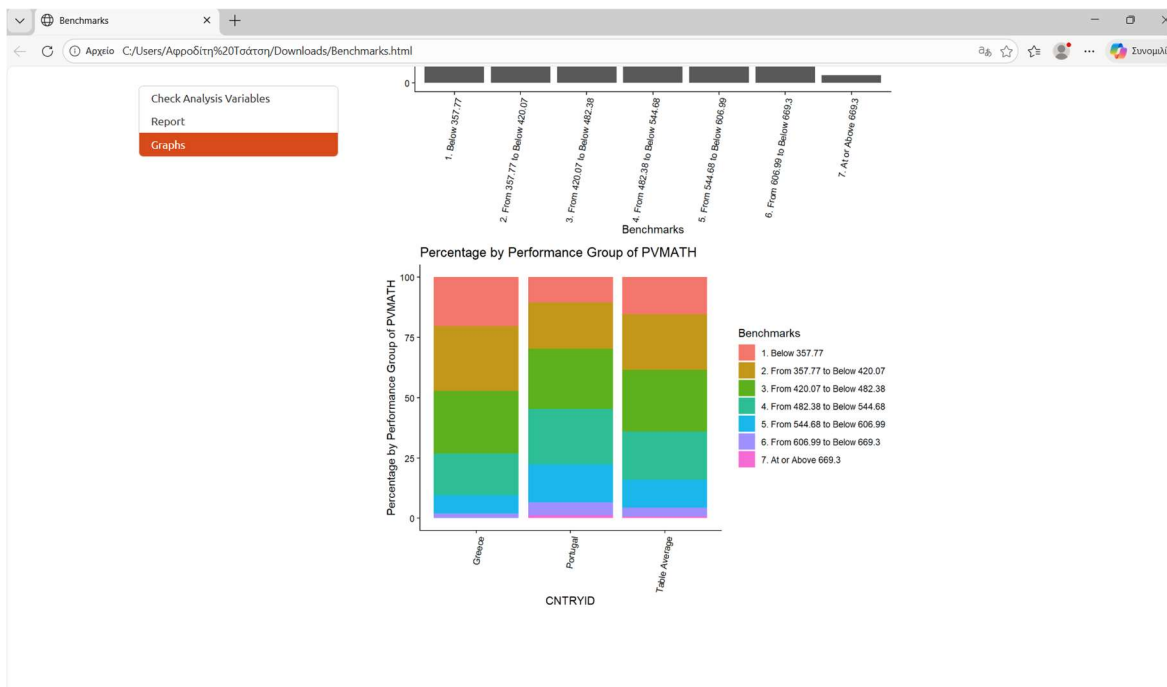
Country Identifier	cutvar	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent (s.e.)
Portugal	4. From 482.38 to Below 544.68	1626	22189	842.86	22.97 0.78
Portugal	5. From 544.68 to Below 606.99	1101	15094	646.33	15.62 0.67
Portugal	6. From 606.99 to Below 669.3	384	5349	351.00	5.54 0.37
Portugal	7. At or Above 669.3	76	1099	188.94	1.14 0.19
Table Average	1. Below 357.77	NA	NA	NA	15.39 0.67
Table Average	2. From 357.77 to Below 420.07	NA	NA	NA	23.07 0.53
Table Average	3. From 420.07 to Below 482.38	NA	NA	NA	25.51 0.54
Table Average	4. From 482.38 to Below 544.68	NA	NA	NA	20.14 0.52
Table Average	5. From 544.68 to Below 606.99	NA	NA	NA	11.55 0.43
Table Average	6. From 606.99 to Below 669.3	NA	NA	NA	3.69 0.23
Table Average	7. At or Above 669.3	NA	NA	NA	0.64 0.10

Εικόνα 71. Πίνακας ποσοστών μαθητών ανά επίπεδο επίδοσης (Benchmarks) ανά χώρα (R)

Αναλυτικά οι στήλες της Εικόνας 70 και 71 παρουσιάζουν:

- Country Identifier: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- cutvar: Οι κατηγορίες επίδοσης που ορίζονται από τα Benchmarks, όπου κάθε γραμμή δείχνει ένα εύρος τιμών.
- N of Cases: Αριθμός ατόμων στην κάθε κατηγορία.
- Sum of W_FSTUWT: Το σταθμισμένο άθροισμα των μαθητών, δηλαδή πόσους μαθητές αντιπροσωπεύουν στο σύνολο του πληθυσμού.
- Sum of W_FSTUWT (s.e.): Το τυπικό σφάλμα του σταθμισμένου αθροίσματος των μαθητών.

- Percent: Το ποσοστό (%) του πληθυσμού της χώρας που βρίσκεται στη συγκεκριμένη κατηγορία επίδοσης.
- Percent (s.e.): Το τυπικό σφάλμα του ποσοστού του πληθυσμού της χώρας που βρίσκεται στη συγκεκριμένη κατηγορία επίδοσης.



Εικόνα 72. Γράφημα κατανομής μαθητών ανά επίπεδο επίδοσης – Ελλάδα vs Πορτογαλία (R)

Ερμηνεία αποτελεσμάτων

Με βάση τα αποτελέσματα που παρουσιάζονται στις Εικόνες 70 και 71, καθώς και στο γράφημα της Εικόνας 72, παρατηρείται ότι στην Ελλάδα στα δύο χαμηλότερα επίπεδα επίδοσης (κάτω από 420.07 μονάδες) συγκεντρώνεται το 47.21% των μαθητών, ενώ στην Πορτογαλία το αντίστοιχο ποσοστό είναι 29.70%. Επομένως, το ποσοστό των μαθητών με χαμηλή επίδοση είναι σαφώς υψηλότερο στην Ελλάδα σε σχέση με την Πορτογαλία. Όσον αφορά το μεσαίο επίπεδο (ομάδες 3 και 4), τα ποσοστά για τις δύο χώρες είναι σχεδόν ίδια. Συγκεκριμένα, αθροιστικά στο μεσαίο επίπεδο εντάσσεται το 43.31% των μαθητών στην Ελλάδα και το 47.99% στην Πορτογαλία. Αντίθετα, στα ανώτερα επίπεδα επίδοσης (πάνω από 544.68 μονάδες), στην Ελλάδα βρίσκεται το 9.48% των μαθητών, ενώ στην Πορτογαλία το 22.30%. Συμπερασματικά, στην Ελλάδα η επίδοση στα Μαθηματικά συγκεντρώνεται κυρίως στις χαμηλότερες βαθμίδες της κλίμακας, ενώ στην Πορτογαλία η κατανομή εμφανίζεται μετατοπισμένη προς τα μεσαία και υψηλά επίπεδα.

4. Ανάλυση με SPSS

Στην ενότητα αυτή παρουσιάζεται η ίδια σειρά στατιστικών αναλύσεων με τη χρήση του SPSS, αξιοποιώντας τις ίδιες μεταβλητές και δεδομένα PISA που χρησιμοποιήθηκαν στην ενότητα με τη χρήση της R.

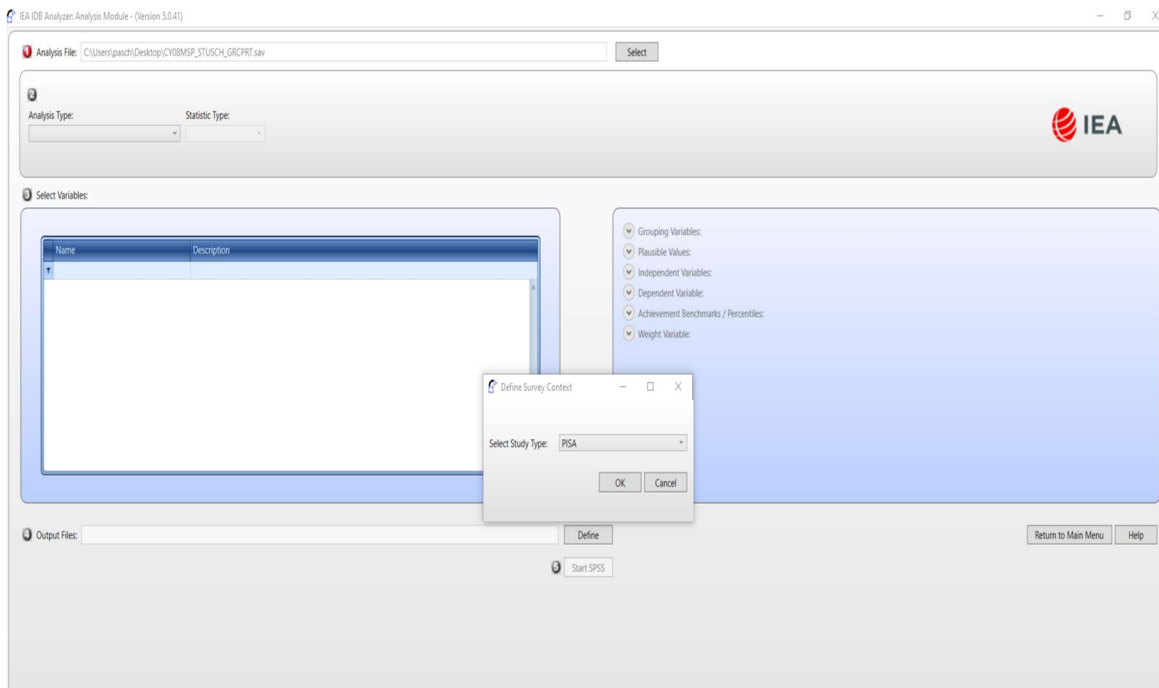
Υπολογισμός του συντελεστή συσχέτισης Pearson

Ερευνητική Υπόθεση: Σκοπός της ανάλυσης συσχέτισης Pearson είναι να διερευνηθεί αν και σε ποιο βαθμό υπάρχει στατιστικά σημαντική γραμμική σχέση μεταξύ της αυτο-αποτελεσματικότητας στα Μαθηματικά και του άγχους για τα Μαθηματικά.

Επιλογή μεταβλητών: MATHEFF, ANXMAT

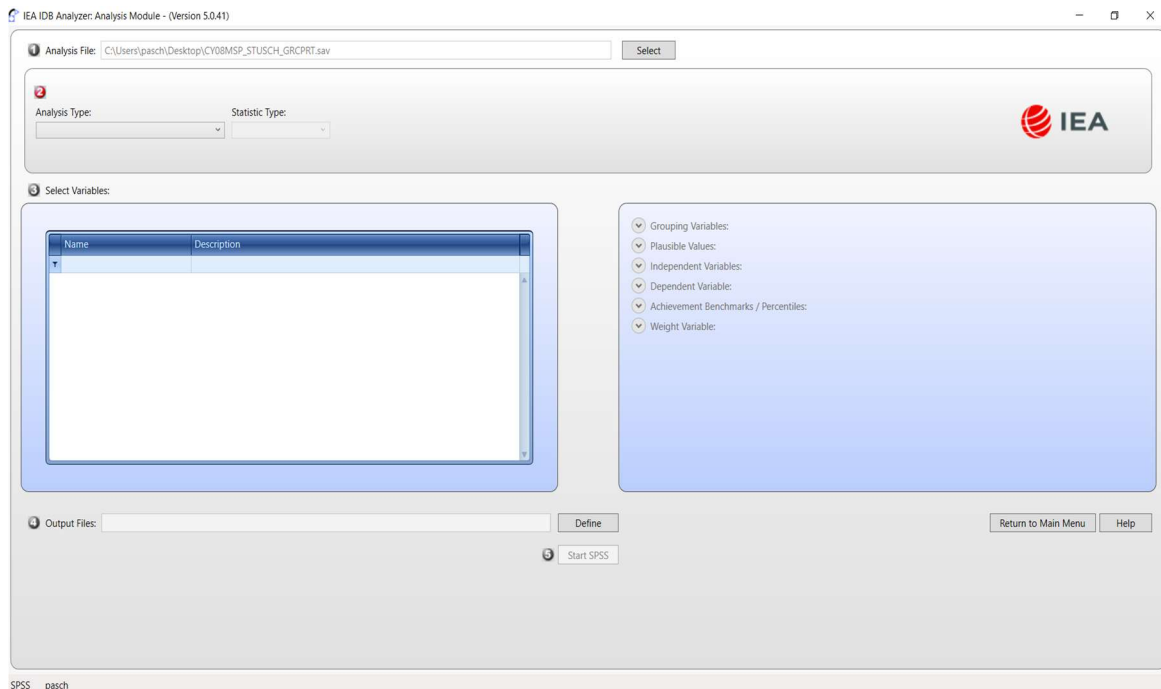
Χαρακτηριστικά μεταβλητών: MATHEFF συνεχής μεταβλητή, ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση συσχέτισης Pearson μεταξύ της μεταβλητής MATHEFF (δείκτης που μετρά την αυτο-αποτελεσματικότητα των μαθητών στα Μαθηματικά) και της μεταβλητής ANXMAT, η οποία αποτυπώνει το άγχος των μαθητών απέναντι στα Μαθηματικά. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPR1.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 73).



Εικόνα 73. Επιλογή αρχείου δεδομένων PISA (.sav) στο πεδίο «Analysis File» (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 74).

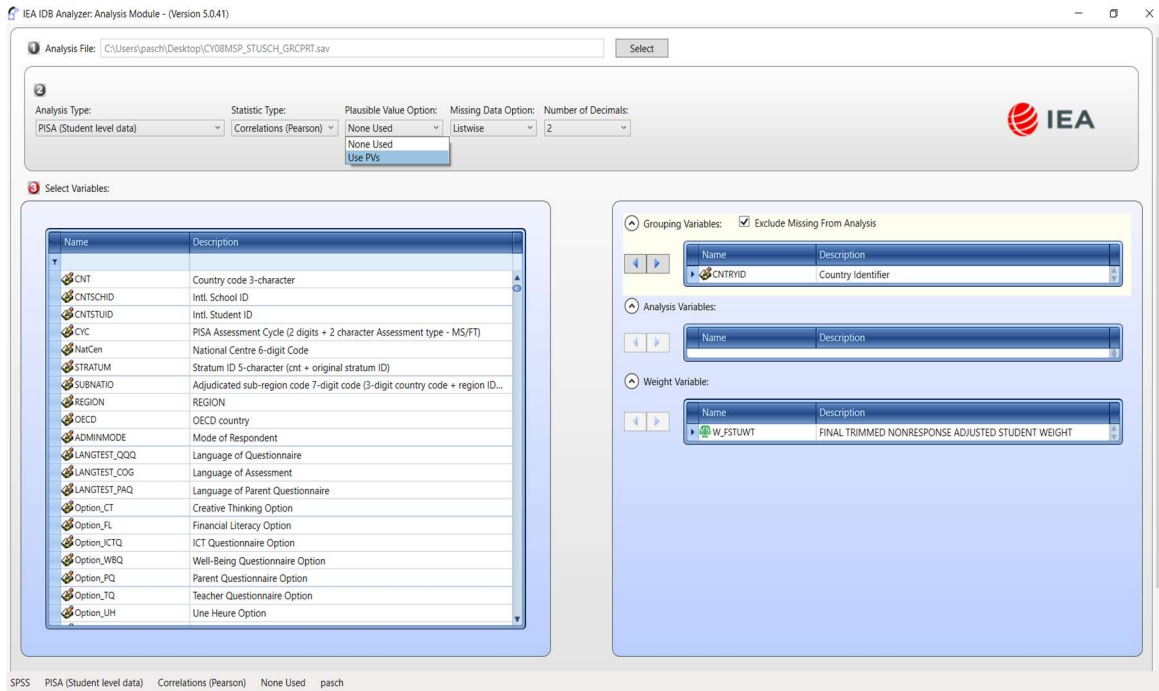


Εικόνα 74. Επιλογή «PISA (Student level)» στο πεδίο «Analysis Type» (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

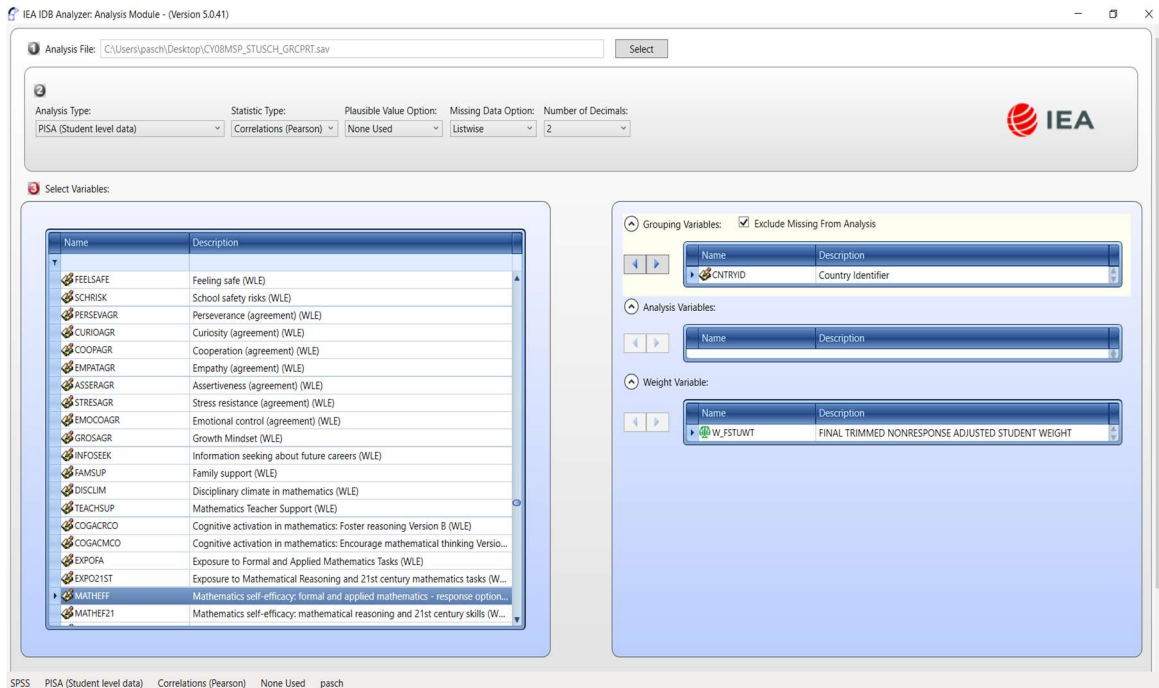
- **Benchmarks:** Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- **Correlation (Pearson):** Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- **Correlation (Spearman):** Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- **Linear regression:** Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- **Logistic regression:** Εξήγηση της σχέσης μεταξύ μιας εξαρτημένης, ποιοτικής, μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- **Percentages and Means:** Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- **Percentages only:** Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- **Percentiles:** Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Correlation (Pearson)».



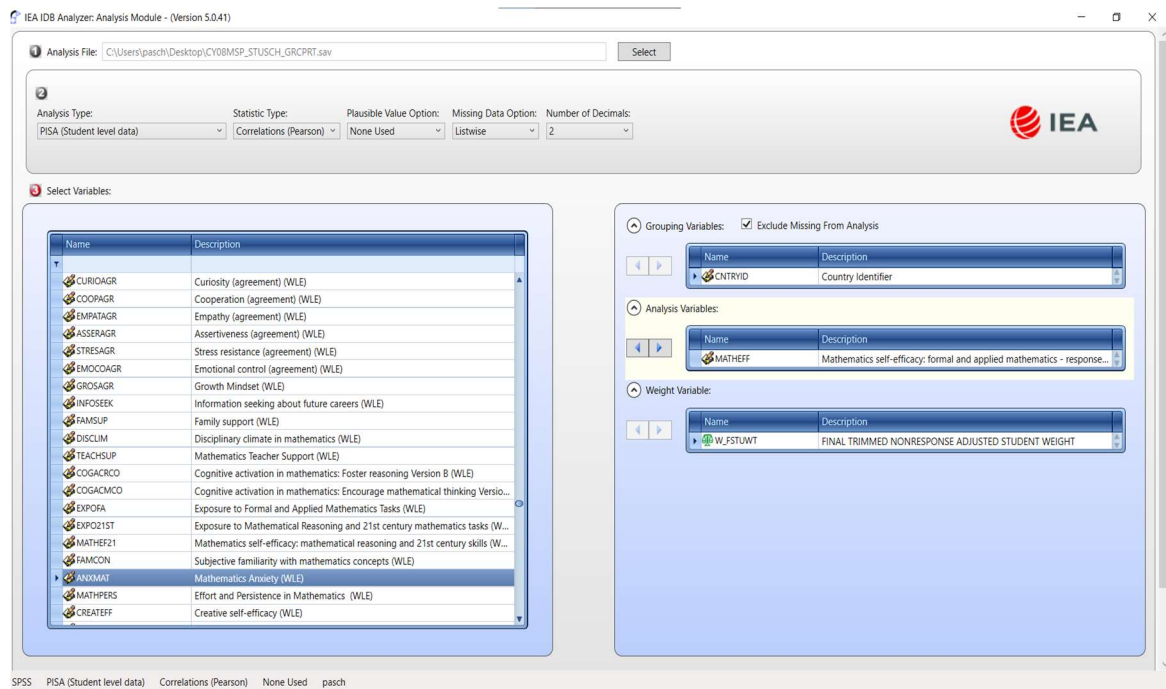
Εικόνα 75. Επιλογή «Correlation (Pearson)» στο πεδίο «Statistic Type» (SPSS)

Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Missing Data Option» υπάρχουν διαθέσιμες οι επιλογές «Pairwise», που αν την επιλέξετε χρησιμοποιούνται όλα τα διαθέσιμα δεδομένα για ανάλυση και «Listwise» που αν την επιλέξετε χρησιμοποιούνται μόνο οι περιπτώσεις που έχουν πλήρη δεδομένα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα75).



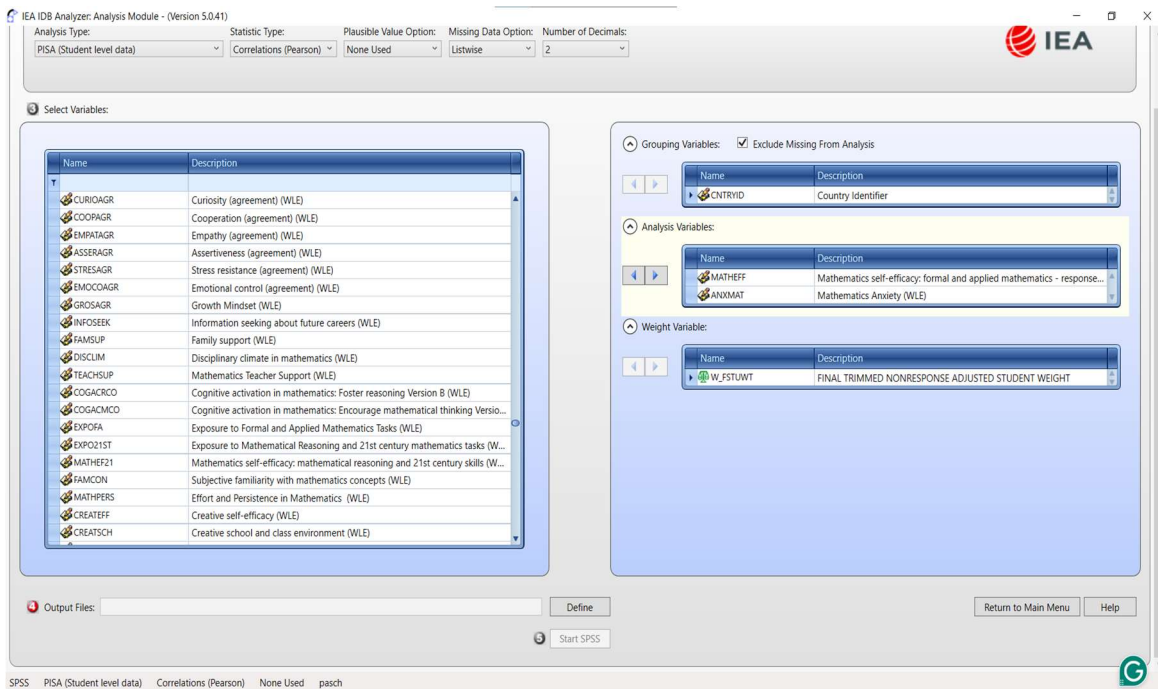
Εικόνα 76. Ρύθμιση παραμέτρων ανάλυσης Pearson – «Grouping Variables» και «Missing Data Option» (SPSS)

Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 76 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει αποπροεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλειπείς τιμές στις μεταβλητές ομαδοποίησης. Στην συνέχεια μας ενδιαφέρει να υπολογίσουμε τον συντελεστή συσχέτισης Pearson μεταξύ της μεταβλητής «MATHEFF», και της «ANXMAT». Επομένως βρίσκουμε αρχικά την μεταβλητή «MATHEFF» από το πλαίσιο «Select Variables», την επιλέγουμε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» την τοποθετείτε. Επαναλαμβάνετε την διαδικασία για την μεταβλητή «ANXMAT». Τέλος στο πλαίσιο «Weight Variable» υπάρχει ως προεπιλογή «W_FSTUWT». Να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 77).



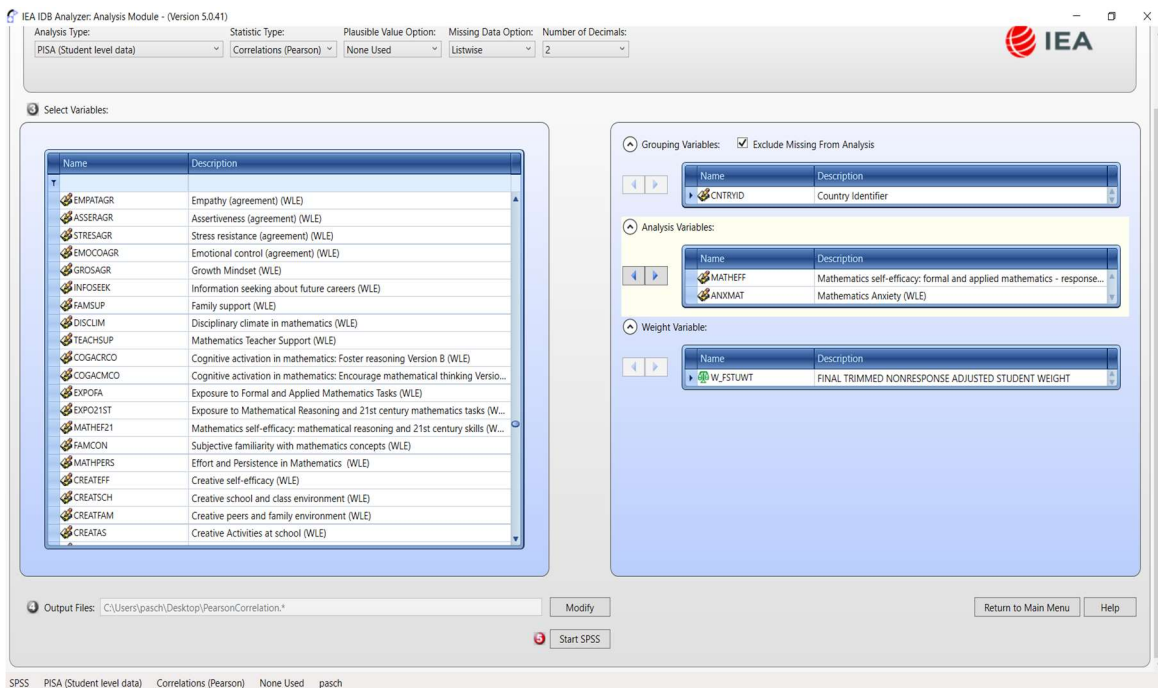
Εικόνα 77. Τοποθέτηση μεταβλητών MATHEFF και ANXMAT στο πεδίο «Analysis Variables» (SPSS)

Επόμενο βήμα, Εικόνα 78, είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PearsonCorrelation».



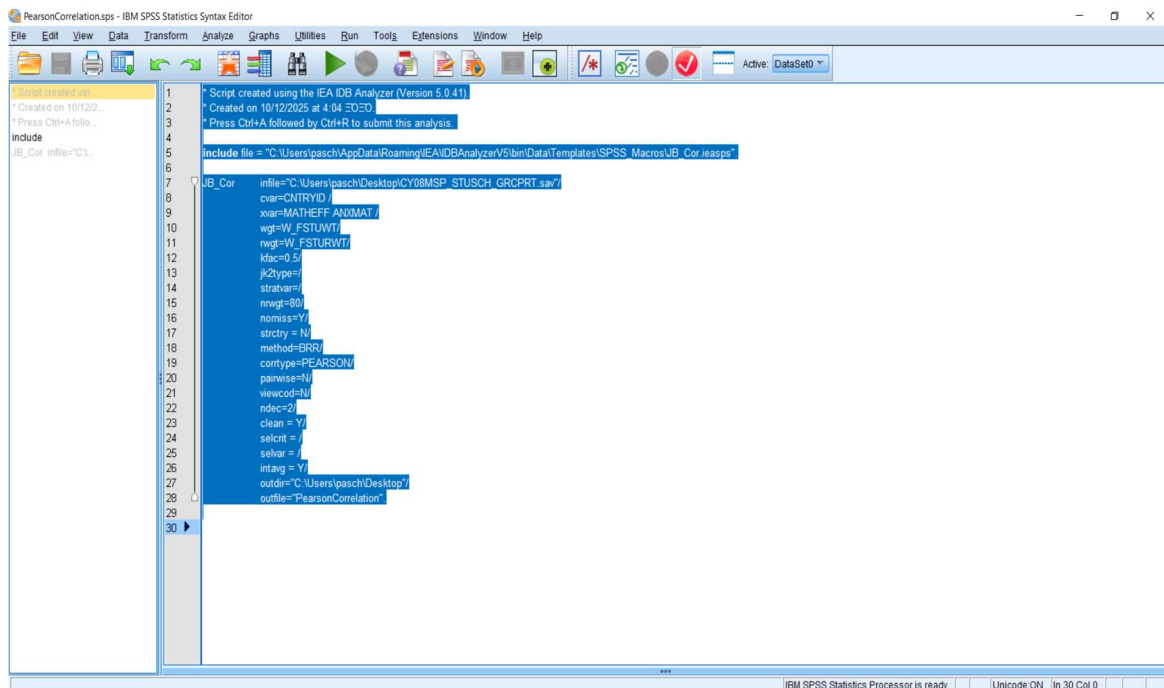
Εικόνα 78. Ορισμός ονόματος αρχείων εξόδου για ανάλυση Pearson (SPSS)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 79).



Εικόνα 79. Εκκίνηση παραγωγής syntax αρχείου μέσω «Start SPSS»

Αφού μεταβείτε στο περιβάλλον του SPSS, στο αρχείο PearsonCorrelation.sps που εμφανίζεται, με Ctrl+A επιλέξτε όλο το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 80).

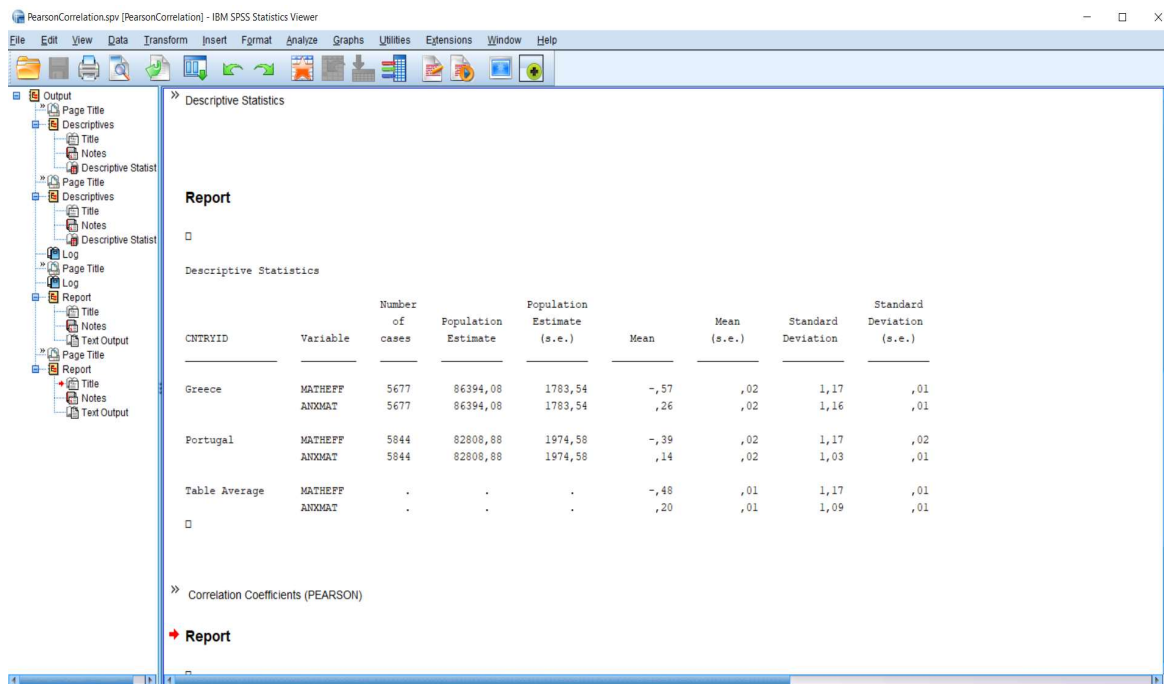


```

1 * Script created using the IEA IDB Analyzer (Version 5.0.41)
2 * Created on 10/12/2025 at 4:04:50:50
3 * Press Ctrl+A followed by Ctrl+R to submit this analysis.
4
5 include file = "C:\Users\pasch\AppData\Roaming\IEA\IDBAnalyzer\5\bin\Data\Templates\SPSS_Macros\UB_Cor.leaps"
6
7 JB_Cor
8   infile="C:\Users\pasch\Desktop\CY08MSP_STUSCH_GRCPR1.sav"
9   onar=CNTRYID
10  xvar=MATHEFF ANXMAT
11  wgt=W_FSTURWT
12  nwt=W_FSTURWT
13  kfac=0.5
14  jk2type=
15  stravar=
16  nwtgt=80
17  nomiss=Y
18  strctry = N
19  method=BRR
20  comtpe=PEARSON
21  pairwise=
22  viewcod=
23  ndec=
24  clean = Y
25  selcrit =
26  selvar =
27  intavg = Y
28  outdir="C:\Users\pasch\Desktop\
29  outfile="PearsonCorrelation"
30
  
```

Εικόνα 80. Εκτέλεση κώδικα PearsonCorrelation.sps στο SPSS (Ctrl+A → Run)

Στο αρχείο PearsonCorrelation.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 81 και 82 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.



Report

Descriptive Statistics

Variable	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)
Greece MATHEFF	5677	86394,08	1783,54	-,57	,02	1,17	,01
Greece ANXMAT	5677	86394,08	1783,54	,26	,02	1,16	,01
Portugal MATHEFF	5844	82808,88	1974,58	-,39	,02	1,17	,02
Portugal ANXMAT	5844	82808,88	1974,58	,14	,02	1,03	,01
Table Average MATHEFF	.	.	.	-,48	,01	1,17	,01
Table Average ANXMAT	.	.	.	,20	,01	1,09	,01

Correlation Coefficients (PEARSON)

Report

Εικόνα 81. Αποτελέσματα περιγραφικής στατιστικής (N, Mean, SD) – Ανάλυση Pearson (SPSS)

Οι στήλες που φαίνονται στην Εικόνα 81, περιγράφουν τα εξής:

- CNTRYID: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).

- **Correlation with ANXMAT (s.e.):** Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής ANXMAT. Δείχνει την ακρίβεια της εκτίμησης.

Ερμηνεία αποτελεσμάτων

Πραγματοποιήθηκε ανάλυση συσχέτισης Pearson προκειμένου να διερευνηθεί η σχέση μεταξύ της αυτο-αποτελεσματικότητας στα Μαθηματικά (MATHEFF) και του άγχους για τα Μαθηματικά (ANXMAT) στην Ελλάδα και την Πορτογαλία. Στην Ελλάδα, οι μαθητές παρουσίασαν χαμηλότερα επίπεδα αυτο-αποτελεσματικότητας στα Μαθηματικά ($M = -0.57$, $SD = 1.17$) και υψηλότερα επίπεδα άγχους για τα Μαθηματικά ($M = 0.26$, $SD = 1.16$). Αντίστοιχα, στην Πορτογαλία, ο μέσος όρος της αυτο-αποτελεσματικότητας στα Μαθηματικά ήταν υψηλότερος ($M = -0.39$, $SD = 1.17$), ενώ τα επίπεδα άγχους για τα Μαθηματικά ήταν χαμηλότερα ($M = 0.14$, $SD = 1.03$). Η ανάλυση συσχέτισης έδειξε ότι υπάρχει στατιστικά σημαντική αρνητική γραμμική συσχέτιση μεταξύ της αυτο-αποτελεσματικότητας και του άγχους για τα Μαθηματικά και στις δύο χώρες. Συγκεκριμένα, στην Ελλάδα παρατηρήθηκε μέτρια αρνητική συσχέτιση ($r = -0.40$), ενώ στην Πορτογαλία η συσχέτιση ήταν ελαφρώς ασθενέστερη ($r = -0.33$, $p < .001$). Ο μέσος όρος της συσχέτισης για τις δύο χώρες ήταν $r = -0.36$.

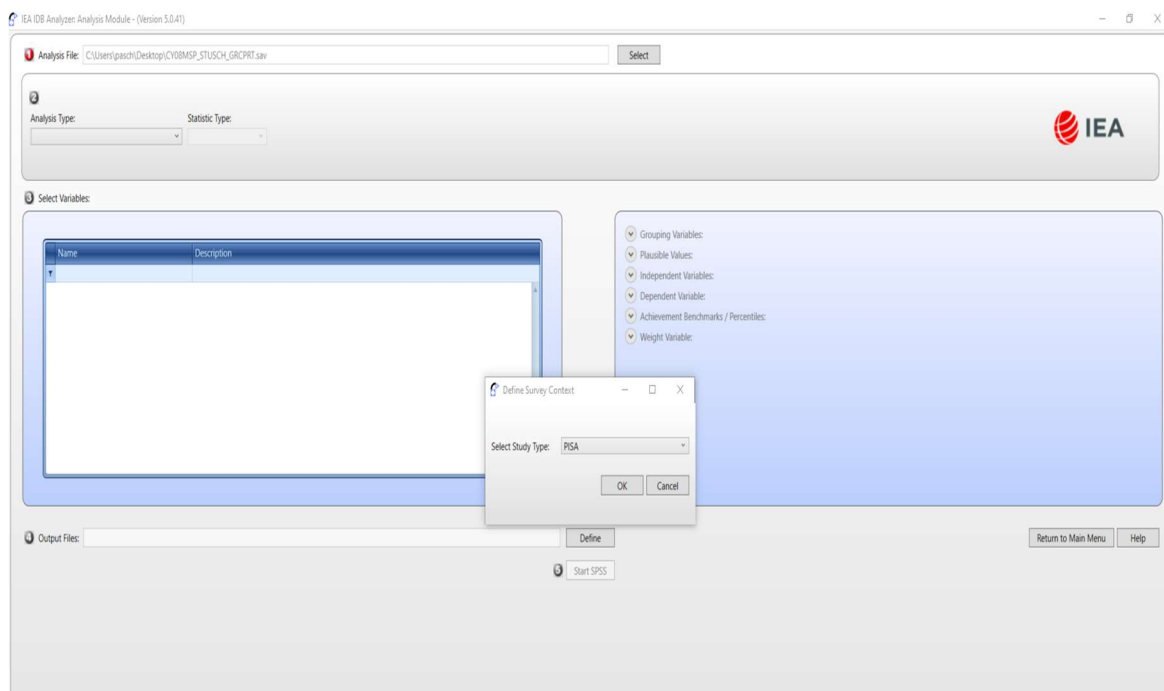
Υπολογισμός του συντελεστή συσχέτισης Spearman

Ερευνητική Υπόθεση: Διερευνάται η ύπαρξη και η κατεύθυνση της σχέσης μεταξύ του άγχους για τα Μαθηματικά και του δείκτη μεταναστευτικού υποβάθρου σε μαθητές της Ελλάδας και της Πορτογαλίας, μέσω της συσχέτισης Spearman.

Επιλογή μεταβλητών: IMMIG, ANXMAT

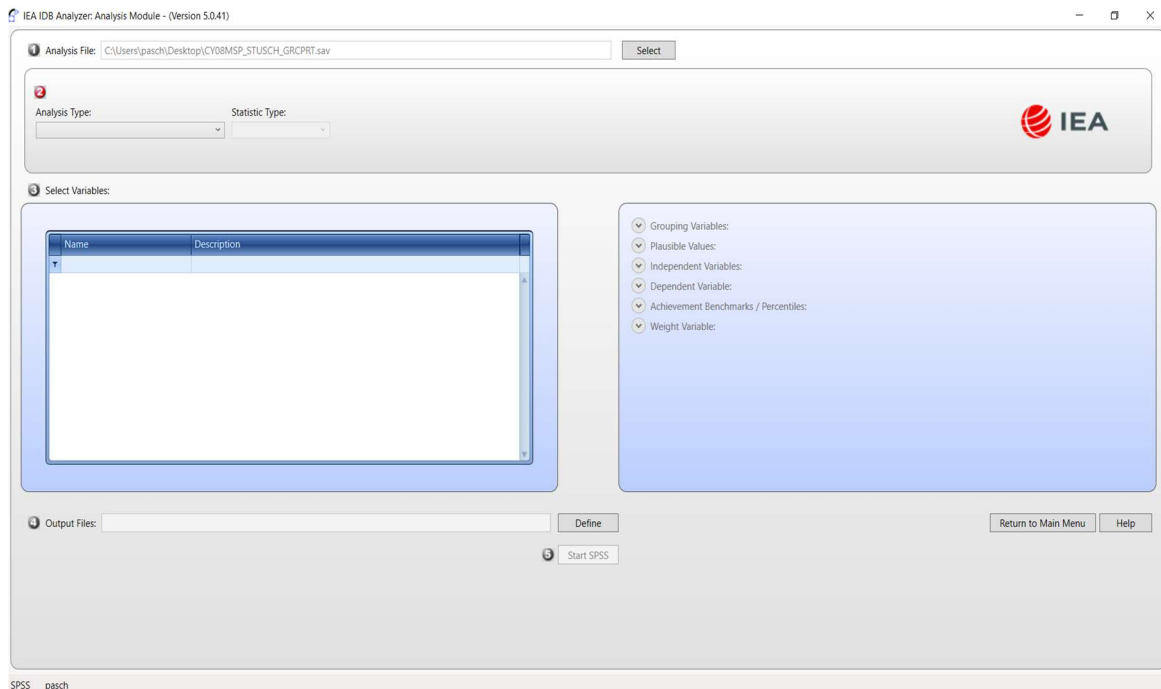
Χαρακτηριστικά μεταβλητών: IMMIG κατηγορική μεταβλητή, ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση συσχέτισης Spearman προκειμένου να διερευνηθεί η σχέση μεταξύ του άγχους για τα Μαθηματικά (ANXMAT) και του δείκτη μεταναστευτικού υποβάθρου (IMMIG) σε μαθητές της Ελλάδας και της Πορτογαλίας. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCprt.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 83).



Εικόνα 83. Επιλογή αρχείου δεδομένων για ανάλυση Spearman – «Analysis File» (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 84).

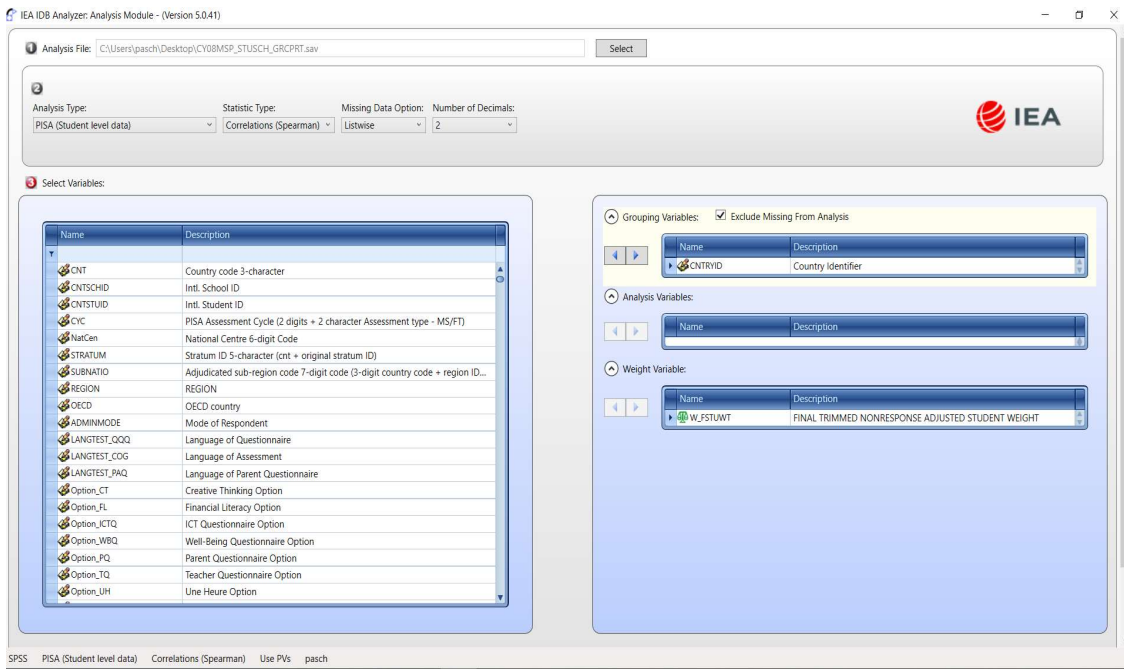


Εικόνα 84. Επιλογή «PISA (Student level)» – Ανάλυση Spearman (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

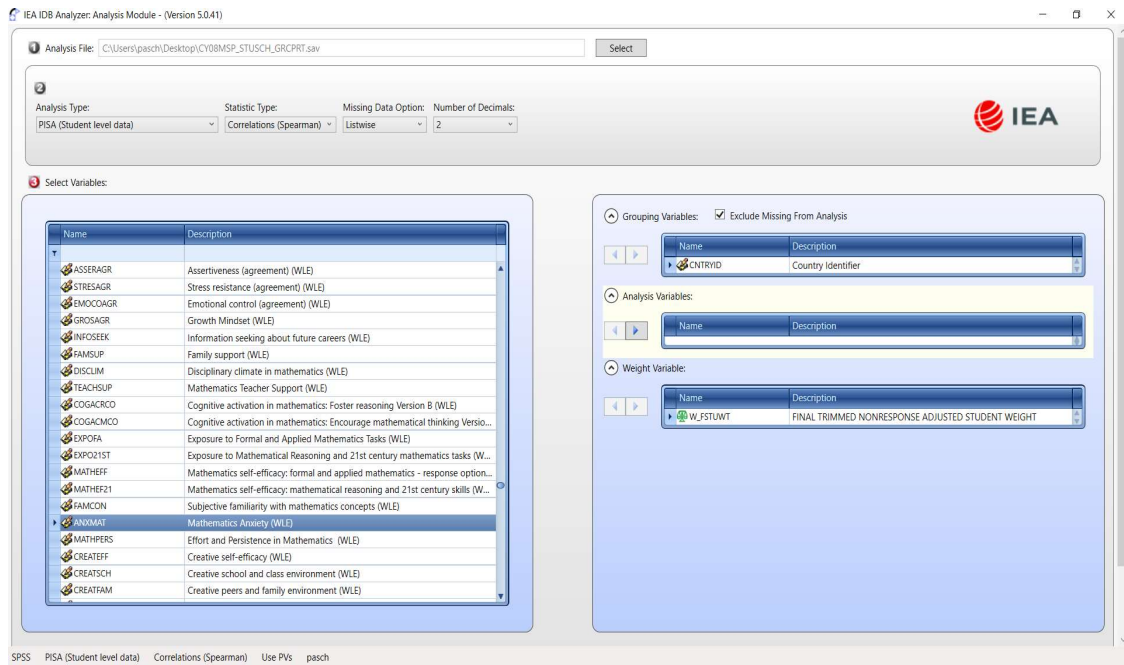
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Logistic regression: Εξήγηση της σχέσης μεταξύ μιας εξαρτημένης, ποιοτικής, μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Correlation (Spearman)».



Εικόνα 85. Επιλογή «Correlation (Spearman)» στο πεδίο «Statistic Type» (SPSS)

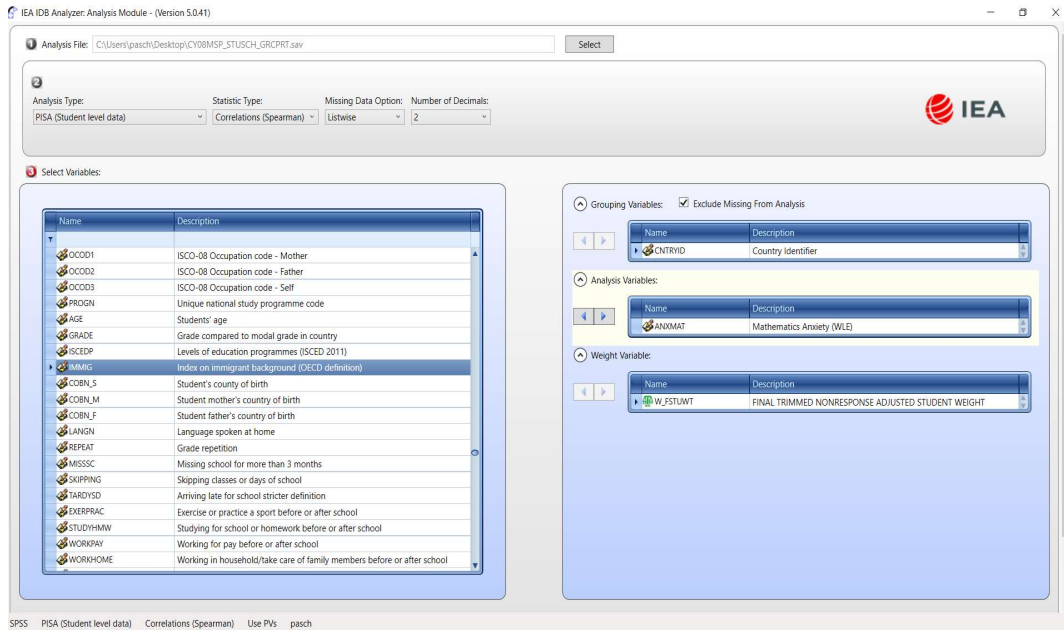
Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Missing Data Option» υπάρχουν διαθέσιμες οι επιλογές «Pairwise», που αν την επιλέξετε χρησιμοποιούνται όλα τα διαθέσιμα δεδομένα για ανάλυση και «Listwise» που αν την επιλέξετε χρησιμοποιούνται μόνο οι περιπτώσεις που έχουν πλήρη δεδομένα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 85).



Εικόνα 86. Ρύθμιση παραμέτρων ανάλυσης Spearman (SPSS)

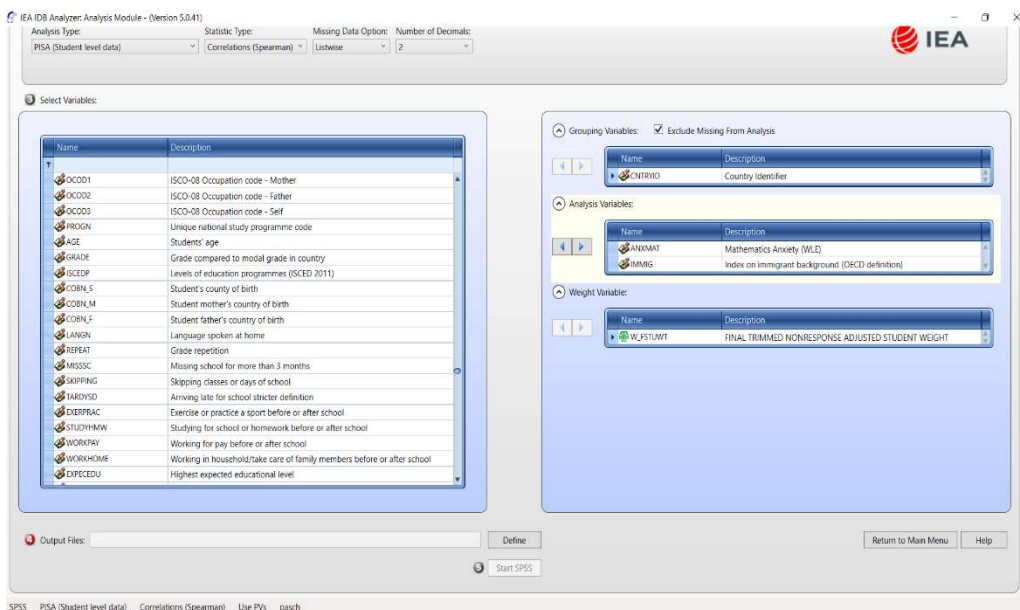
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 86 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλιπείς

τιμές στις μεταβλητές ομαδοποίησης. Στην συνέχεια μας ενδιαφέρει να υπολογίσουμε τον συντελεστή συσχέτισης Spearman μεταξύ της μεταβλητής «ANXMAT», και της «IMMIG». Επομένως βρίσκετε αρχικά την μεταβλητή «ANXMAT» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» την τοποθετείτε (Εικόνα 87). Επαναλαμβάνετε την διαδικασία για την μεταβλητή «IMMIG». Τέλος στο πλαίσιο «Weight Variable» υπάρχει ως προεπιλογή η μεταβλητή «W_FSTUWT». Να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 87).



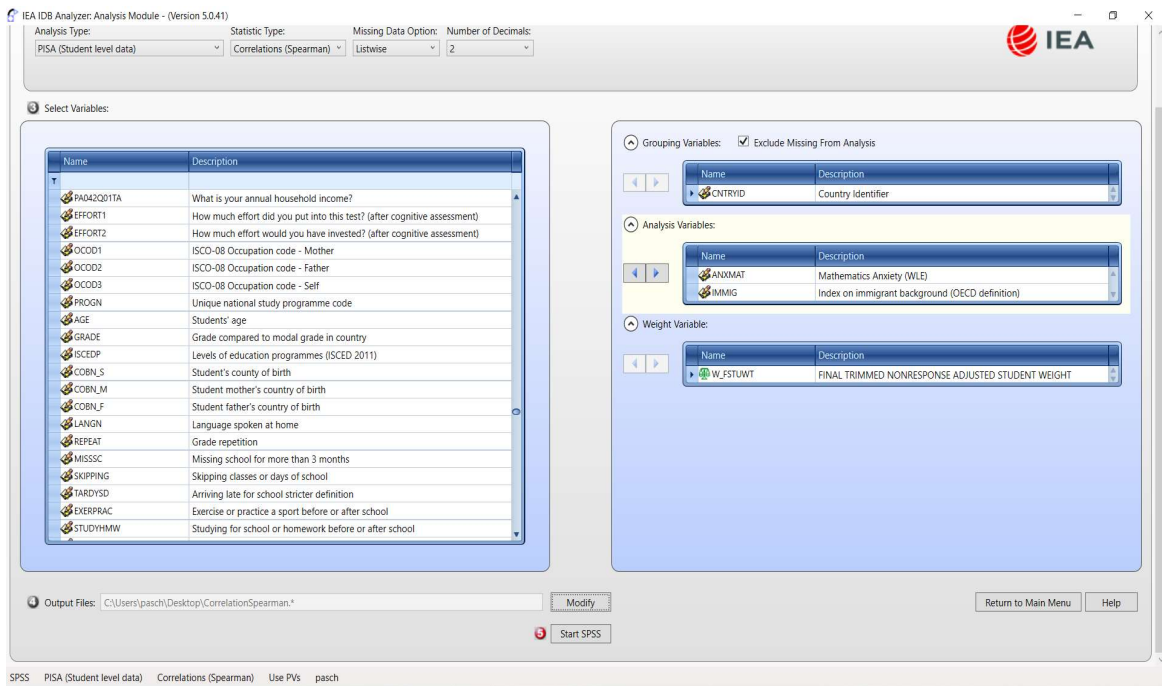
Εικόνα 87. Τοποθέτηση μεταβλητών ANXMAT και IMMIG στο πεδίο «Analysis Variables» (SPSS)

Το επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «CorrelationSpearman» (Εικόνα 88).



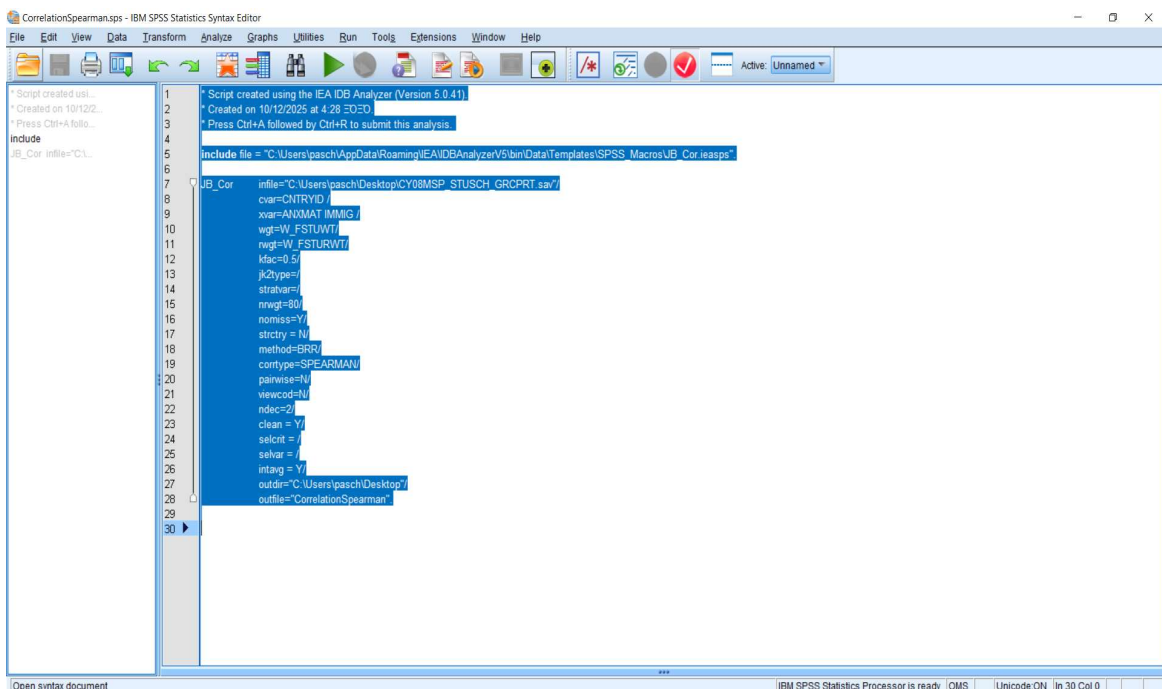
Εικόνα 88. Ορισμός ονόματος εξόδου για ανάλυση Spearman (SPSS)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 89).



Εικόνα 89. Εκκίνηση παραγωγής syntax αρχείου Spearman μέσω «Start SPSS»

Αφού μεταβείτε στο περιβάλλον του SPSS, στο αρχείο CorrelationSpearman.sps που εμφανίζεται, με Ctrl+A επιλέξετε όλο το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 90).



Εικόνα 90. Εκτέλεση κώδικα CorrelationSpearman.sps στο SPSS

Στο αρχείο CorrelationSpearman.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 91 και 92 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

The screenshot shows the SPSS Statistics Viewer interface. The main window displays a report titled 'Descriptive Statistics' for Spearman correlation analysis. The report includes a table with the following data:

CNTRYID	Variable	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)
Greece	RANXMAT	5679	86404,11	1764,58	2840,86	24,18	1641,26	10,10
	RIMMIG	5679	86404,11	1764,58	2869,59	18,63	949,72	20,66
Portugal	RANXMAT	5846	82812,19	2019,09	2939,93	26,23	1691,17	9,86
	RIMMIG	5846	82812,19	2019,09	2934,17	20,58	902,77	26,64
Table Average	RANXMAT	.	.	.	2890,40	17,84	1666,22	7,06
	RIMMIG	.	.	.	2901,88	13,88	926,24	16,85

Εικόνα 91. Αποτελέσματα ανάλυσης Spearman – περιγραφικά στατιστικά ανά χώρα (SPSS)

Οι στήλες που φαίνονται στην Εικόνα 91, περιγράφουν τα εξής:

- CNTRYID: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Variable: Το όνομα της μεταβλητής.
- Number of cases: Ο αριθμός των μαθητών στο δείγμα της χώρας που απάντησαν στη συγκεκριμένη μεταβλητή.
- Population Estimate: Εκτίμηση του συνολικού πληθυσμού που αντιπροσωπεύει το δείγμα σε εθνικό επίπεδο.
- Population Estimate (s.e.): Το τυπικό σφάλμα της εκτίμησης του πληθυσμού το οποίο δείχνει την ακρίβεια της εκτίμησης.
- Mean: Ο μέσος όρος των τιμών της μεταβλητής στο δείγμα της χώρας.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Standard Deviation: Η τυπική απόκλιση η οποία δείχνει πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο.
- Standard Deviation (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης το οποίο δείχνει πόσο αξιόπιστη είναι η εκτίμηση της διασποράς.

The screenshot shows the IBM SPSS Statistics Viewer interface. The main window displays a report titled 'Correlation Coefficients (SPEARMAN)'. The report includes a table with the following data:

CNTRYID	Variable	Correlation with RANXMAT	Correlation with RANXMAT (s.e.)	Correlation with RIMMIG	Correlation with RIMMIG (s.e.)
Greece	RANXMAT	1,00	,00	,09	,02
	RIMMIG	,09	,02	1,00	,00
Portugal	RANXMAT	1,00	,00	,06	,01
	RIMMIG	,06	,01	1,00	,00
Table Average	RANXMAT	1,00	,00	,08	,01
	RIMMIG	,08	,01	1,00	,00

Εικόνα 92. Πίνακας συντελεστών συσχέτισης Spearman (SPSS)

Οι στήλες που φαίνονται στην Εικόνα 92, περιγράφουν τα εξής:

- CNTRYID: Αναφέρεται στη χώρα από την οποία προέρχονται τα δεδομένα (π.χ Ελλάδα, Πορτογαλία). Όλες οι πληροφορίες που περιέχονται στις αντίστοιχες γραμμές του πίνακα αφορούν αποκλειστικά αυτήν τη χώρα.
- Variable: Δείχνει τη μεταβλητή στην οποία αναφέρεται ο συντελεστής συσχέτισης.
- Correlation with RANXMAT: Ο συντελεστής συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής RANXMAT.
- Correlation with RANXMAT (s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής της στήλης "Variable" και της RANXMAT.
- Correlation with RIMMIG: Ο συντελεστής συσχέτισης μεταξύ της μεταβλητής που αναφέρεται στη στήλη "Variable" και της μεταβλητής RIMMIG. Η συσχέτιση είναι συμμετρική, επομένως η τιμή μεταξύ των RANXMAT και RIMMIG είναι ίδια ανεξάρτητα από τη σειρά τους.
- Correlation with RIMMIG (s.e.): Το τυπικό σφάλμα του συντελεστή συσχέτισης μεταξύ της μεταβλητής της στήλης "Variable" και της RIMMIG.

Ερμηνεία αποτελεσμάτων

Η ανάλυση συσχέτισης έδειξε ότι υπάρχει ασθενής αλλά θετική συσχέτιση μεταξύ του άγχους για τα Μαθηματικά και του δείκτη μεταναστευτικού υπόβαθρου και στις δύο χώρες. Στην Ελλάδα, παρατηρήθηκε ασθενής θετική συσχέτιση ($\rho = 0.09$, $SE = 0.02$), υποδηλώνοντας ότι οι μαθητές με μεταναστευτικό υπόβαθρο τείνουν να εμφανίζουν ελαφρώς αυξημένα επίπεδα άγχους για τα Μαθηματικά. Αντίστοιχα, στην Πορτογαλία η συσχέτιση ήταν επίσης ασθενής και θετική ($\rho = 0.06$, $SE = 0.01$), αλλά μικρότερης έντασης σε σύγκριση με την Ελλάδα. Ο μέσος όρος της συσχέτισης για τις δύο χώρες ήταν $\rho = 0.08$, γεγονός που ενισχύει ότι η σχέση μεταξύ των δύο μεταβλητών είναι γενικά χαμηλής ισχύος.

Υπολογισμός γραμμικής παλινδρόμησης

Ερευνητική Υπόθεση: Διερευνάται αν το άγχος για τα Μαθηματικά προβλέπεται από το φύλο, το κοινωνικοοικονομικό επίπεδο και την αντίληψη των μαθητών ότι μπορούν να εφαρμόζουν μαθηματικό συλλογισμό σε δραστηριότητες του 21ου αιώνα, καθώς και η κατεύθυνση αυτών των σχέσεων.

Επιλογή μεταβλητών: ANXMAT, ST004D01T, ESCS, MATHEF21

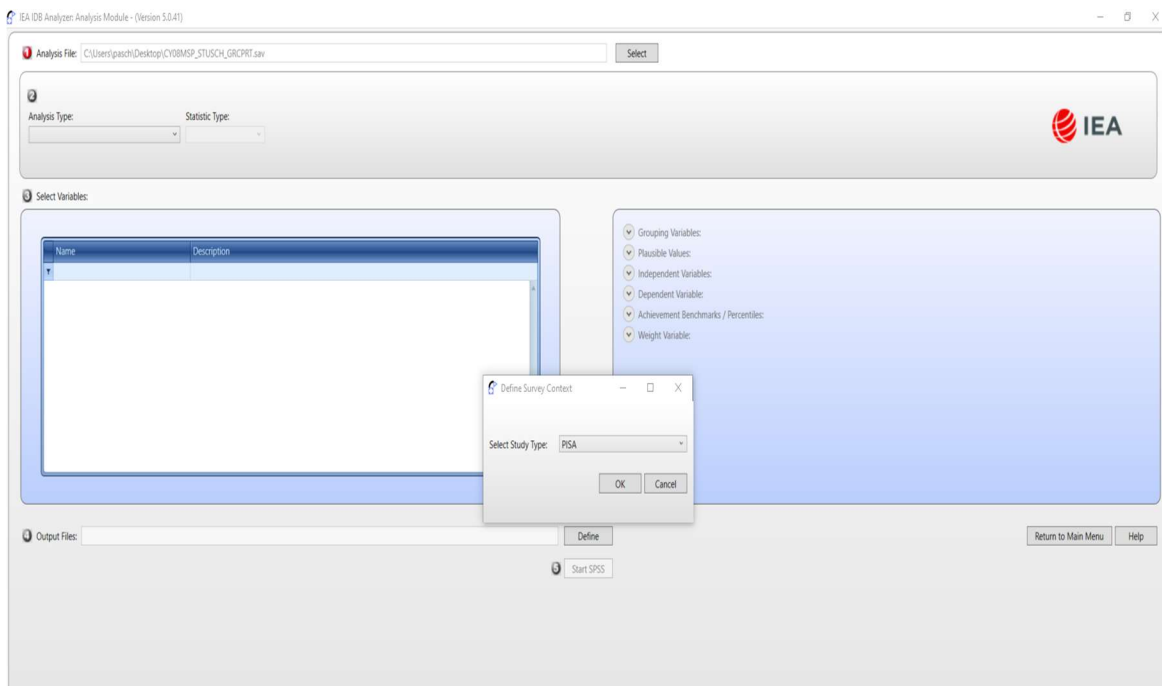
Χαρακτηριστικά μεταβλητών: ANXMAT συνεχής μεταβλητή, ST004D01T κατηγορική μεταβλητή, ESCS συνεχής μεταβλητή, MATHEF21 συνεχής μεταβλητή

Μεταβλητές: Μοντέλο (σε μορφή εξίσωσης)

Όπου:

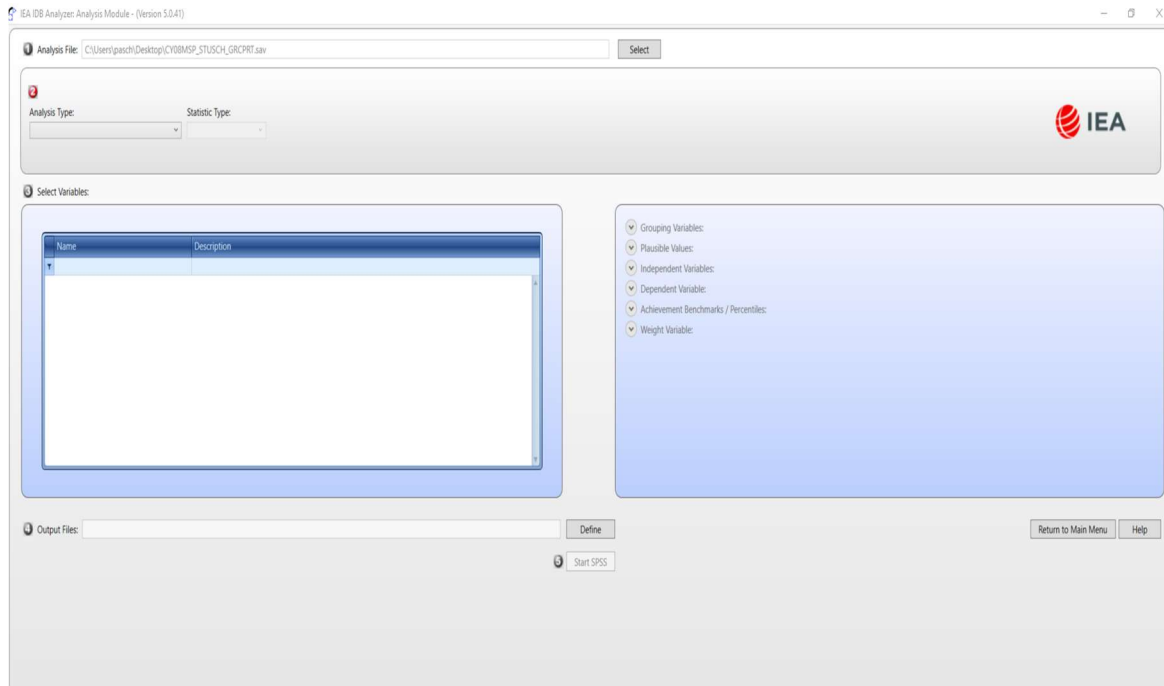
Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση γραμμικής παλινδρόμησης, όπου ως εξαρτημένη μεταβλητή χρησιμοποιείται το άγχος για τα Μαθηματικά (ANXMAT) και ως ανεξάρτητες μεταβλητές το φύλο του μαθητή (ST004D01T), το κοινωνικοοικονομικό επίπεδο (ESCS) και η αυτο-αντίληψη των μαθητών ότι μπορούν να εφαρμόσουν μαθηματικό συλλογισμό σε δραστηριότητες του 21ου αιώνα (MATHEF21).

Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPR1.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 93).



Εικόνα 93. Επιλογή αρχείου δεδομένων για γραμμική παλινδρόμηση (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 94).

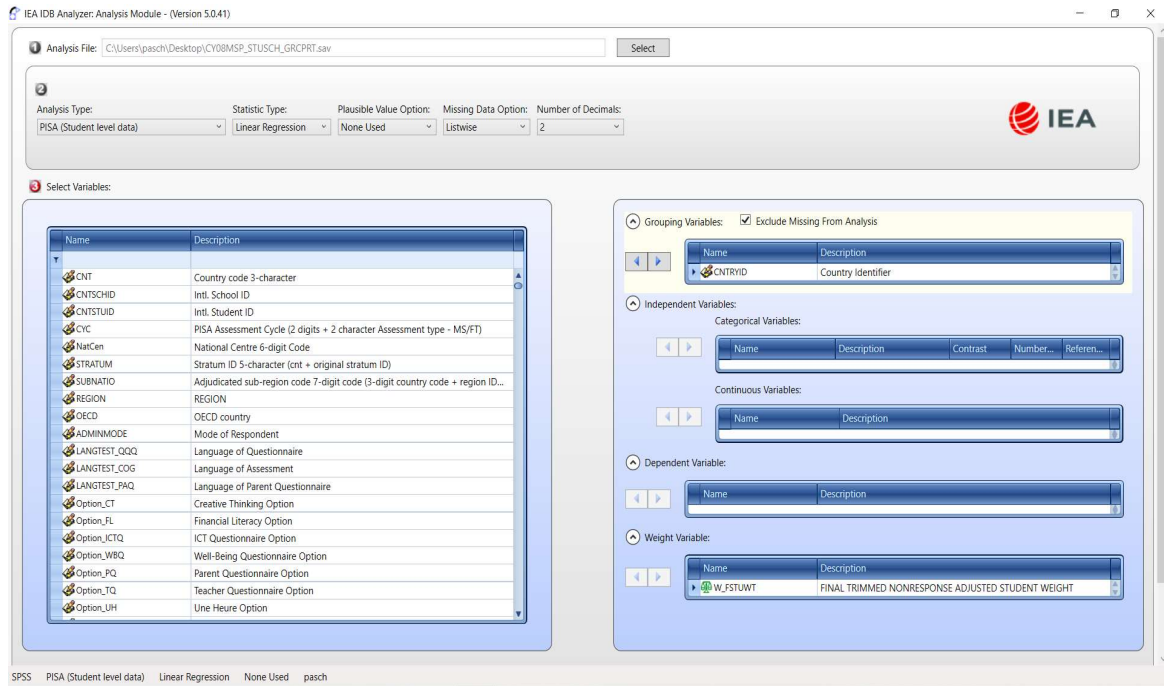


Εικόνα 94. Επιλογή «PISA (Student level)» – Γραμμική Παλινδρόμηση (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

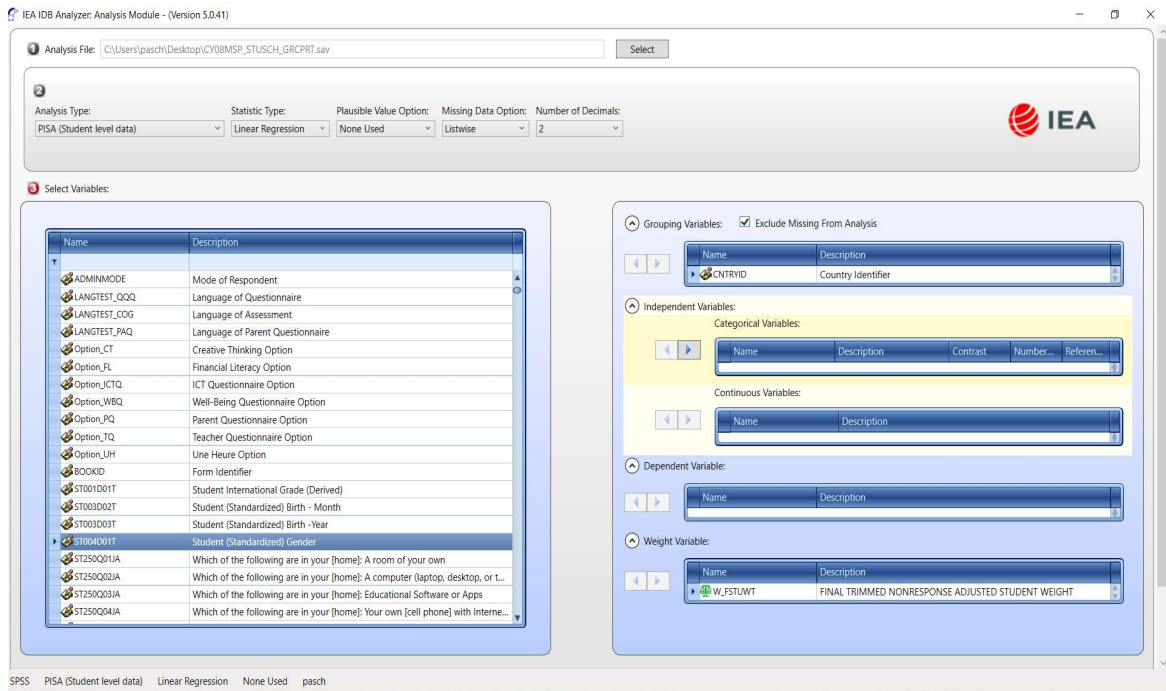
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Linear Regression». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Missing Data Option» υπάρχουν διαθέσιμες οι επιλογές «Pairwise», οπότε αν την επιλέξετε χρησιμοποιούνται όλα τα διαθέσιμα δεδομένα για ανάλυση και «Listwise» οπότε αν την επιλέξετε χρησιμοποιούνται μόνο οι περιπτώσεις που έχουν πλήρη δεδομένα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 95).

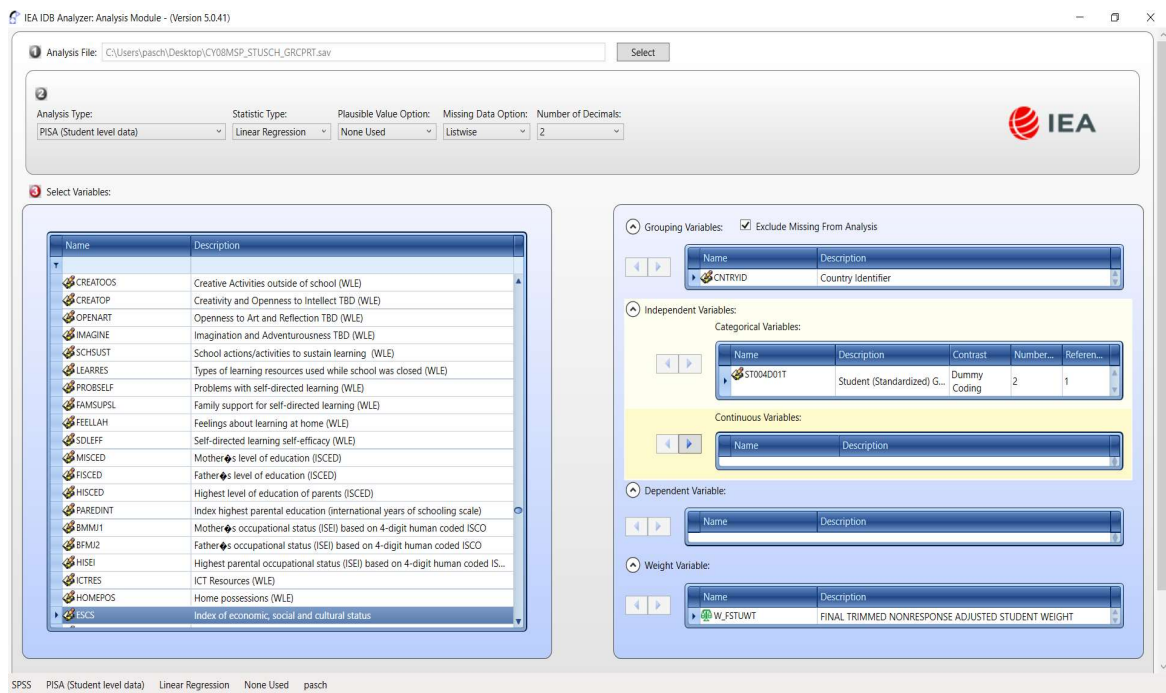


Εικόνα 95. Επιλογή «Linear Regression» στο πεδίο «Statistic Type» (SPSS)

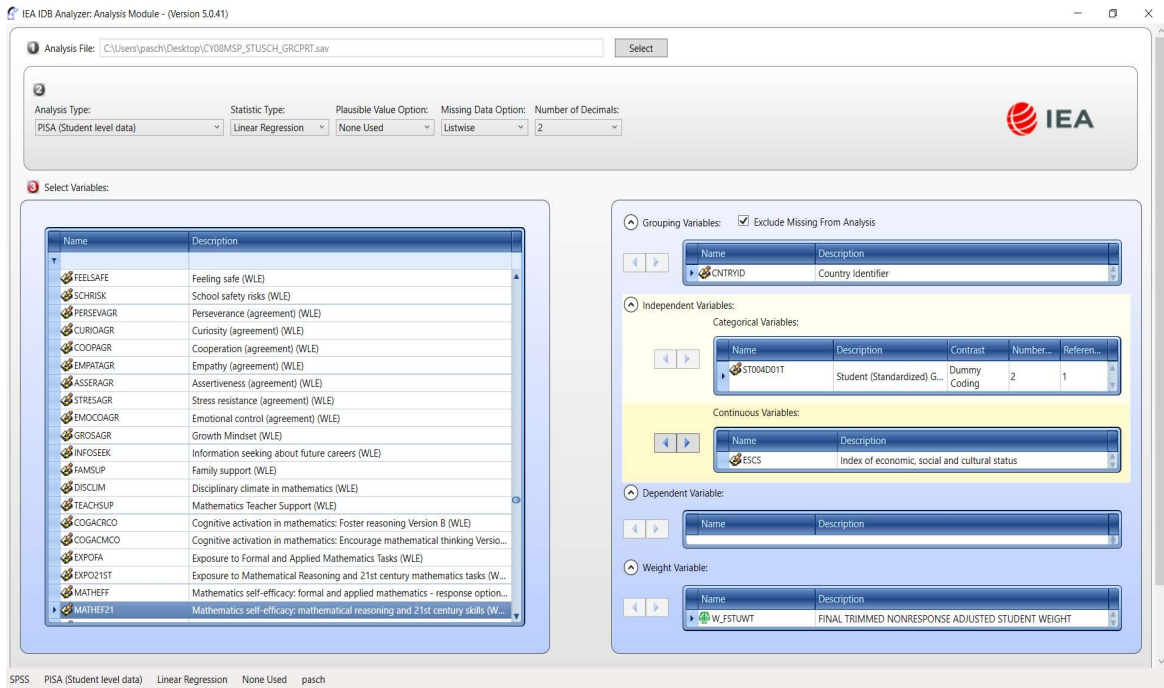
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 95, στο πλαίσιο «Grouping Variables», το οποίο περιλαμβάνει τη λίστα των μεταβλητών εμφανίζεται από προεπιλογή η μεταβλητή «CNTRYID». Επιπλέον, εφόσον η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, στην ανάλυση θα συμπεριληφθούν μόνο οι περιπτώσεις με μη ελλιπείς τιμές στις μεταβλητές. Για την υλοποίηση της γραμμικής παλινδρόμησης, είναι απαραίτητος ο εντοπισμός και η σωστή τοποθέτηση των απαιτούμενων μεταβλητών στο πλαίσιο «Select Variables». Αρχικά, εντοπίζεται η μεταβλητή «ST004D01T» (φύλο μαθητή), η οποία επιλέγεται και, με το μπλε βελάκι μεταφέρεται στο πλαίσιο «Independent Variables» και συγκεκριμένα στο πλαίσιο «Categorical Variables» (Εικόνα 96). Στη συνέχεια, η ίδια διαδικασία επαναλαμβάνεται για τις μεταβλητές «ESCS» και «MATHEF21», οι οποίες μεταφέρονται στο πλαίσιο «Continuous Variables», καθώς πρόκειται για συνεχείς μεταβλητές (Εικόνα 97 και Εικόνα 98).



Εικόνα 96. Τοποθέτηση κατηγορικής μεταβλητής ST004D01T (SPSS)

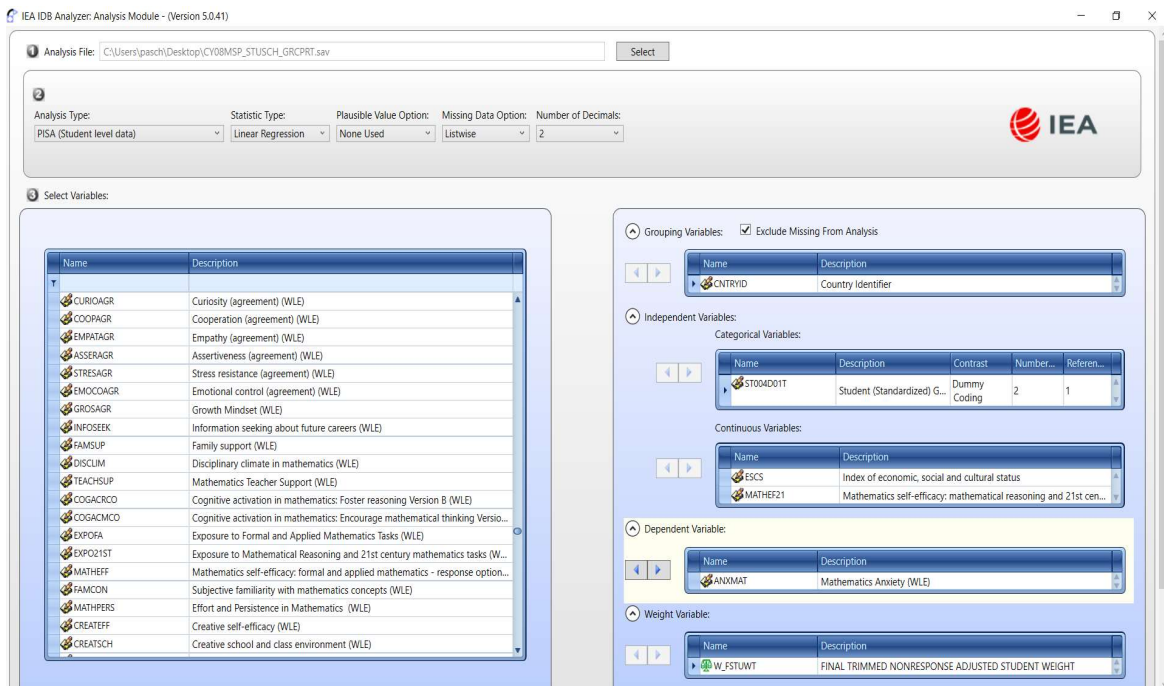


Εικόνα 97. Τοποθέτηση συνεχούς μεταβλητής ESCS (SPSS)



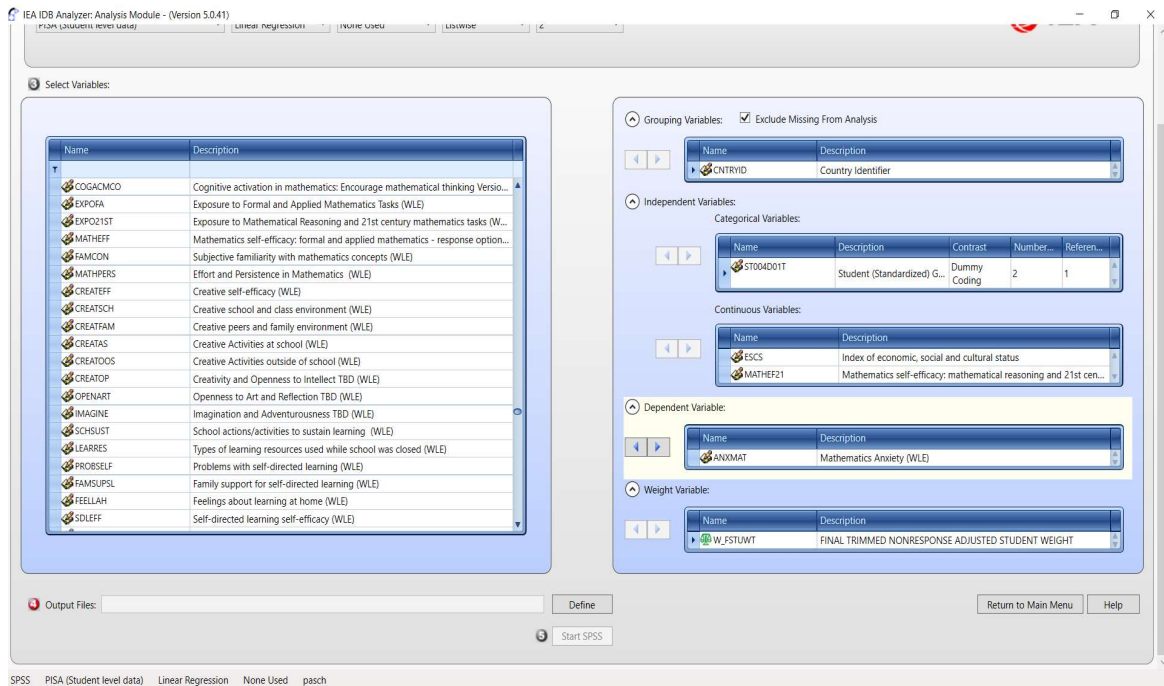
Εικόνα 98. Τοποθέτηση συνεχούς μεταβλητής MATHEF21 (SPSS)

Ακολούθως, στο πλαίσιο «Dependent Variable», επιλέγεται η μεταβλητή «ANXMAT», η οποία χρησιμοποιείται ως εξαρτημένη μεταβλητή στη γραμμική παλινδρόμηση. Τέλος, στο πλαίσιο «Weight Variable» εμφανίζεται από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα την κατάλληλη μεταβλητή βάρων για την εκτέλεση της ανάλυσης (Εικόνα 98).



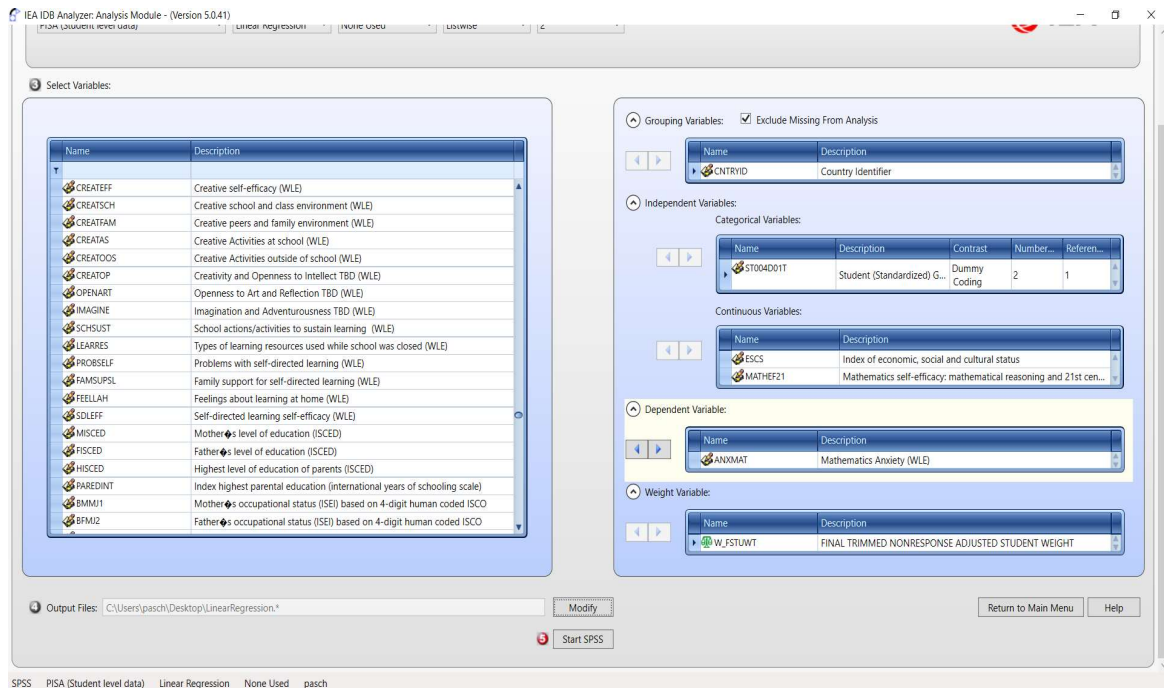
Εικόνα 99. Επιλογή εξαρτημένης μεταβλητής ANXMAT – Γραμμική Παλινδρόμηση (SPSS)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «LinearRegression» (Εικόνα 100).



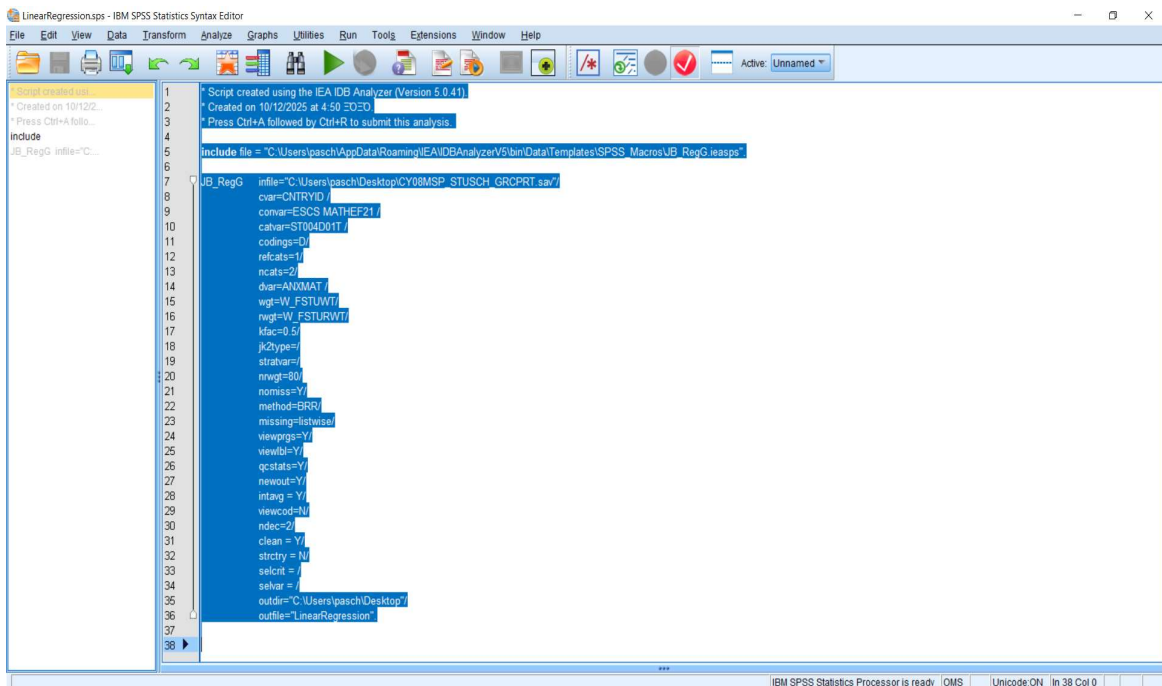
Εικόνα 100. Ορισμός ονόματος εξόδου για γραμμική παλινδρόμηση (SPSS)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 101).



Εικόνα 101. Εκκίνηση παραγωγής syntax γραμμικής παλινδρόμησης μέσω «Start SPSS»

Ύστερα από τη μετάβαση στο SPSS, στο αρχείο LinearRegression.sps που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 102).



Εικόνα 102. Εκτέλεση κώδικα LinearRegression.sps στο SPSS

Στο αρχείο LinearRegression.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 103 και 104 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

	R-Square	R-Square (s.e.)	Adjusted R-Square	Adjusted R-Square (s.e.)
Greece	,16	,01	,16	,01
Portugal	,11	,01	,11	,01
Table Average	,13	,01	,13	,01

Εικόνα 103. Αποτελέσματα R-Square – Γραμμική Παλινδρόμηση (SPSS)

Οι στήλες στην Εικόνα 103 παρουσιάζουν τα στατιστικά μοντέλου για διάφορες χώρες που επιλέγονται από τον χρήστη. Συγκεκριμένα:

- CNTRYID: Το όνομα της χώρας, εδώ συγκεκριμένα Ελλάδα, Πορτογαλία και στο τέλος δίνεται ο μέσος όρος.
- R-Square: Ο συντελεστής προσδιορισμού, είναι ουσιαστικά το ποσοστό διακύμανσης του άγχους για τα Μαθηματικά που εξηγείται από τις ανεξάρτητες μεταβλητές (εδώ «ST004D01T», «ESCS», «MATHEF21»)
- R-Square (s.e.): Το τυπικό σφάλμα του R-Square.
- Adjusted R-Square: Ο προσαρμοσμένος συντελεστής προσδιορισμού, που λαμβάνει υπόψη τον αριθμό μεταβλητών.
- Adjusted R-Square (s.e.): Το τυπικό σφάλμα του προσαρμοσμένου R-Square.

The screenshot shows the SPSS Statistics Viewer interface with a report titled 'Report'. The report displays descriptive statistics for four variables: ANKMAT, ESCS, MATHEF21, and ST004D01T_D2, grouped by country (Greece and Portugal). The variables are also grouped under a 'Table Average' section. The statistics include the number of cases, population estimate, mean, standard deviation, and variance for each variable and group.

CNTRYID	EqVar	Number of cases	Population Estimate	Population Estimate (s.e.)	Mean	Mean (s.e.)	Standard Deviation	Standard Deviation (s.e.)	Variance	Variance (s.e.)
Greece	ANKMAT	5628	85527,13	1789,34	,26	,02	1,15	,01	1,33	,03
	ESCS	5628	85527,13	1789,34	-,12	,02	,91	,01	,83	,02
	MATHEF21	5628	85527,13	1789,34	-,26	,02	1,02	,01	1,04	,02
	ST004D01T_D2	5628	85527,13	1789,34	,49	,01	,50	,00	,25	,00
Portugal	ANKMAT	5809	82271,18	1983,56	,13	,02	1,03	,01	1,06	,03
	ESCS	5809	82271,18	1983,56	-,20	,03	1,14	,01	1,31	,03
	MATHEF21	5809	82271,18	1983,56	,12	,01	,95	,01	,90	,02
	ST004D01T_D2	5809	82271,18	1983,56	,51	,01	,50	,00	,25	,00
Table Average	ANKMAT	.	.	.	,20	,01	1,09	,01	1,20	,02
	ESCS	.	.	.	-,16	,02	1,03	,01	1,07	,02
	MATHEF21	.	.	.	-,07	,01	,99	,01	,97	,02
	ST004D01T_D2	.	.	.	,50	,00	,50	,00	,25	,00

Εικόνα 104. Περιγραφικά στατιστικά μεταβλητών γραμμικής παλινδρόμησης (SPSS)

Ο πίνακας της Εικόνα 104 παρουσιάζει περιγραφικά στατιστικά για τις μεταβλητές που συμμετέχουν στη γραμμική παλινδρόμηση, για κάθε χώρα χωριστά (εδώ Ελλάδα και Πορτογαλία), όπως επιλέγονται από τον χρήστη. Συγκεκριμένα:

- CNTRYID: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- EqVar: Το όνομα της μεταβλητής.
- Number of cases: Ο αριθμός των μαθητών στο δείγμα της χώρας που απάντησαν στη συγκεκριμένη μεταβλητή.
- Population Estimate: Εκτίμηση του συνολικού πληθυσμού που αντιπροσωπεύει το δείγμα σε εθνικό επίπεδο.
- Population Estimate (s.e.): Το τυπικό σφάλμα της εκτίμησης του πληθυσμού το οποίο δείχνει την ακρίβεια της εκτίμησης.
- Mean: Ο μέσος όρος των τιμών της μεταβλητής στο δείγμα της χώρας.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Standard Deviation: Η τυπική απόκλιση η οποία δείχνει πόσο διασκορπισμένες είναι οι τιμές γύρω από τον μέσο όρο.
- Standard Deviation (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης.

- Variance: Η διακύμανση.
- Variance (s.e.): Το τυπικό σφάλμα της διακύμανσης.

		Regression Coefficient	Regression Coefficient (s.e.)	Regression Coefficient (t-value)	Stdrdzsd. Coefficient	Stdrdzsd. Coefficient (s.e.)	Stdrdzsd. Coefficient (t-value)
Greece	(CONSTANT)	,30	,02	13,04	.	.	.
	ESCS	-,07	,02	-4,30	-,06	,01	-4,28
	MATHEF21	-,38	,02	-20,87	-,34	,02	-21,09
	ST004D01T_D2	-,31	,03	-9,40	-,13	,01	-9,59
Portugal	(CONSTANT)	,29	,02	13,85	.	.	.
	ESCS	-,07	,01	-5,35	-,07	,01	-5,30
	MATHEF21	-,30	,02	-13,04	-,27	,02	-12,98
	ST004D01T_D2	-,25	,03	-9,05	-,12	,01	-9,26
Table Average	(CONSTANT)	,29	,02	18,96	.	.	.
	ESCS	-,07	,01	-6,60	-,07	,01	-6,78
	MATHEF21	-,34	,01	-23,28	-,31	,01	-23,14
	ST004D01T_D2	-,28	,02	-13,03	-,13	,01	-13,33

Εικόνα 105. Συντελεστές παλινδρόμησης ανά μεταβλητή και χώρα (SPSS)

Ο πίνακας της Εικόνας 105 παρουσιάζει τους συντελεστές παλινδρόμησης από το γραμμικό μοντέλο που εφαρμόστηκε, για κάθε χώρα χωριστά (π.χ. Ελλάδα, Πορτογαλία), όπως επιλέγονται από τον χρήστη. Συγκεκριμένα:

- CNTRYID: Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- EqVar: Το όνομα της ανεξάρτητης μεταβλητής που εισέρχεται στο μοντέλο (π.χ. ESCS, MATHEF21, ST004D01T_D2) καθώς και η σταθερά (CONSTANT)).
- Regression Coefficient: Ο εκτιμώμενος συντελεστής παλινδρόμησης. Δείχνει κατά πόσο μεταβάλλεται η εξαρτημένη μεταβλητή (π.χ. επίδοση) όταν η ανεξάρτητη αυξάνεται κατά μία μονάδα, κρατώντας τις υπόλοιπες σταθερές.
- Regression Coefficient (s.e.): Το τυπικό σφάλμα της εκτίμησης του συντελεστή. Μετράει την αβεβαιότητα της εκτίμησης.
- Regression Coefficient (t-value): Η τιμή του t-στατιστικού ελέγχου για τον συντελεστή.
- Stdrdzsd. Coefficient: Ο τυποποιημένος συντελεστής παλινδρόμησης.
- Stdrdzsd. Coefficient (s.e.): Το τυπικό σφάλμα της εκτίμησης του τυποποιημένου συντελεστή.
- Stdrdzsd. Coefficient (t-value): Η τιμή του t-στατιστικού ελέγχου για τον τυποποιημένο συντελεστή. Χρησιμοποιείται για τον έλεγχο της στατιστικής σημαντικότητας στην τυποποιημένη κλίμακα.

Ερμηνεία αποτελεσμάτων

Διεξήχθη πολλαπλή γραμμική παλινδρόμηση με εξαρτημένη μεταβλητή το άγχος για τα Μαθηματικά (ANXMAT) και ανεξάρτητες μεταβλητές το κοινωνικοοικονομικό επίπεδο (ESCS), τον δείκτη (MATHEF21) και το φύλο (ST004D01T_D2) για την Ελλάδα και την Πορτογαλία.

Ελλάδα

Στην Ελλάδα, το ANXMAT είχε μέση τιμή $M = 0.26$ ($SE = 0.02$) και τυπική απόκλιση $SD = 1.15$. Το ESCS είχε $M = -0.12$ ($SE = 0.02$), ενώ το MATHEF21 είχε $M = -0.26$ ($SE = 0.02$). Η μεταβλητή φύλου (ST004D01T_D2) παρουσίασε μέση τιμή 0.49, υποδηλώνοντας περίπου ισόρροπη κατανομή των κατηγοριών.

Το μοντέλο εξήγησε το 16% της διακύμανσης του άγχους για τα Μαθηματικά ($R^2 = .16$, $SE = .01$; adjusted $R^2 = .16$). Οι συντελεστές παλινδρόμησης έδειξαν ότι το ESCS σχετίστηκε αρνητικά και στατιστικά σημαντικά με το άγχος ($\beta = -0.07$, $t = -4.30$), υποδηλώνοντας ότι αύξηση κατά 1 μονάδα στο ESCS συνδέεται με μείωση περίπου 0.07 μονάδων στο ANXMAT διατηρώντας τις λοιπές μεταβλητές σταθερές. Ο δείκτης MATHEF21 παρουσίασε επίσης αρνητική και ισχυρή σχέση με το άγχος ($\beta = -0.38$, $t = -20.87$), δηλαδή υψηλότερες τιμές MATHEF21 συνδέθηκαν με χαμηλότερο ANXMAT. Τέλος, το φύλο (ST004D01T_D2) ήταν αρνητικός και στατιστικά σημαντικός προγνωστικός παράγοντας ($\beta = -0.31$, $t = -9.40$), υποδηλώνοντας χαμηλότερο άγχος κατά περίπου 0.31 μονάδες για τα αγόρια σε σχέση με τα κορίτσια

Με βάση τους τυποποιημένους συντελεστές, ο ισχυρότερος προγνωστικός παράγοντας ήταν το MATHEF21 ($\beta = -.38$), ακολούθησε το φύλο ($\beta = -.31$) και στη συνέχεια το ESCS ($\beta = -.07$).

Πορτογαλία

Στην Πορτογαλία, το ANXMAT είχε $M = 0.13$ ($SE = 0.02$) και $SD = 1.03$, υποδηλώνοντας χαμηλότερο μέσο άγχος σε σύγκριση με την Ελλάδα. Το ESCS είχε $M = -0.20$ ($SE = 0.03$) και το MATHEF21 $M = 0.12$ ($SE = 0.01$). Η μεταβλητή φύλου (ST004D01T_D2) είχε μέση τιμή 0.51, επίσης ενδεικτική ισόρροπης κατανομής.

Το μοντέλο εξήγησε το 11% της διακύμανσης του ANXMAT ($R^2 = .11$, $SE = .01$; adjusted $R^2 = .11$). Το ESCS συσχετίστηκε αρνητικά και στατιστικά σημαντικά με το άγχος ($b = -0.07$, $t = -5.35$). Αντίστοιχα, το MATHEF21 ήταν αρνητικός και στατιστικά σημαντικός προγνωστικός παράγοντας ($b = -0.30$, $t = -13.04$). Τέλος, και το φύλο (ST004D01T_D2) παρουσίασε αρνητική και στατιστικά σημαντική σχέση με το άγχος ($b = -0.25$, $t = -9.05$), με τα αγόρια να εμφανίζουν χαμηλότερο άγχος κατά περίπου 0.25 μονάδες σε σχέση με τα κορίτσια

Οι τυποποιημένοι συντελεστές έδειξαν ότι ο ισχυρότερος παράγοντας ήταν και εδώ το MATHEF21 ($\beta = -.30$), ακολούθησε το φύλο ($\beta = -.25$) και στη συνέχεια το ESCS ($\beta = -.07$).

Υπολογισμός λογιστικής παλινδρόμησης

Ερευνητική Υπόθεση: Διερευνάται κατά πόσο η πιθανότητα εμφάνισης σχολικής απουσίας προβλέπεται από το φύλο, το κοινωνικοοικονομικό επίπεδο και την εμπειρία εκφοβισμού, μέσω λογιστικής παλινδρόμησης.

Επιλογή μεταβλητών: SKIPPING, ST004D01T, ESCS, BULLIED

Χαρακτηριστικά μεταβλητών: SKIPPING κατηγορική μεταβλητή, ST004D01T κατηγορική μεταβλητή, ESCS συνεχής μεταβλητή, BULLIED συνεχής μεταβλητή

Μεταβλητές: Μοντέλο (σε μορφή εξίσωσης)

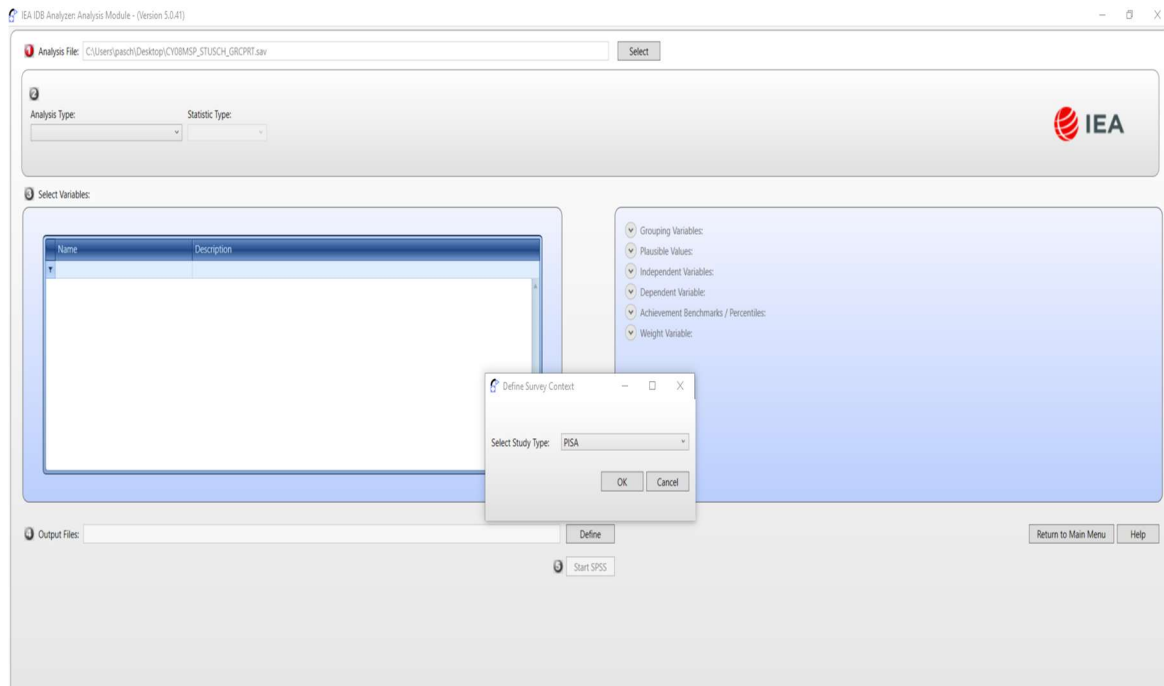
Όπου:

, δηλαδή η πιθανότητα να ισχύει για το i άτομο

, (SKIPPING)

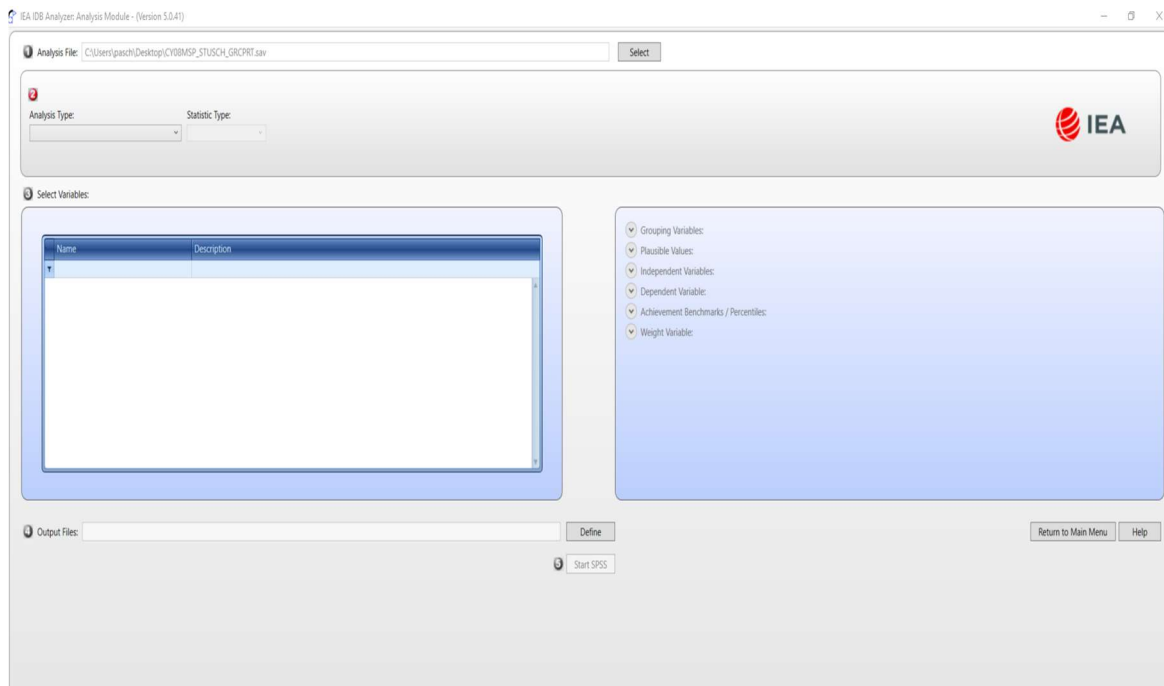
(ST004D01T, ESCS, BULLIED)

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση λογιστικής παλινδρόμησης, όπου ως εξαρτημένη μεταβλητή χρησιμοποιείται η μεταβλητή SKIPPING, η οποία αποτυπώνει την εμφάνιση ή μη στο σχολείο και λαμβάνει την τιμή 1 όταν ο μαθητής δήλωσε ότι έχει παραλείψει μαθήματα ή ημέρες σχολείου και 0 σε διαφορετική περίπτωση. Ως ανεξάρτητες μεταβλητές χρησιμοποιούνται το φύλο του μαθητή (ST004D01T), το κοινωνικοοικονομικό επίπεδο (ESCS) και η μεταβλητή BULLIED, η οποία αποτελεί δείκτη εμπειρίας εκφοβισμού. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCprt.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 106).



Εικόνα 106. Επιλογή αρχείου δεδομένων για λογιστική παλινδρόμηση (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 107).

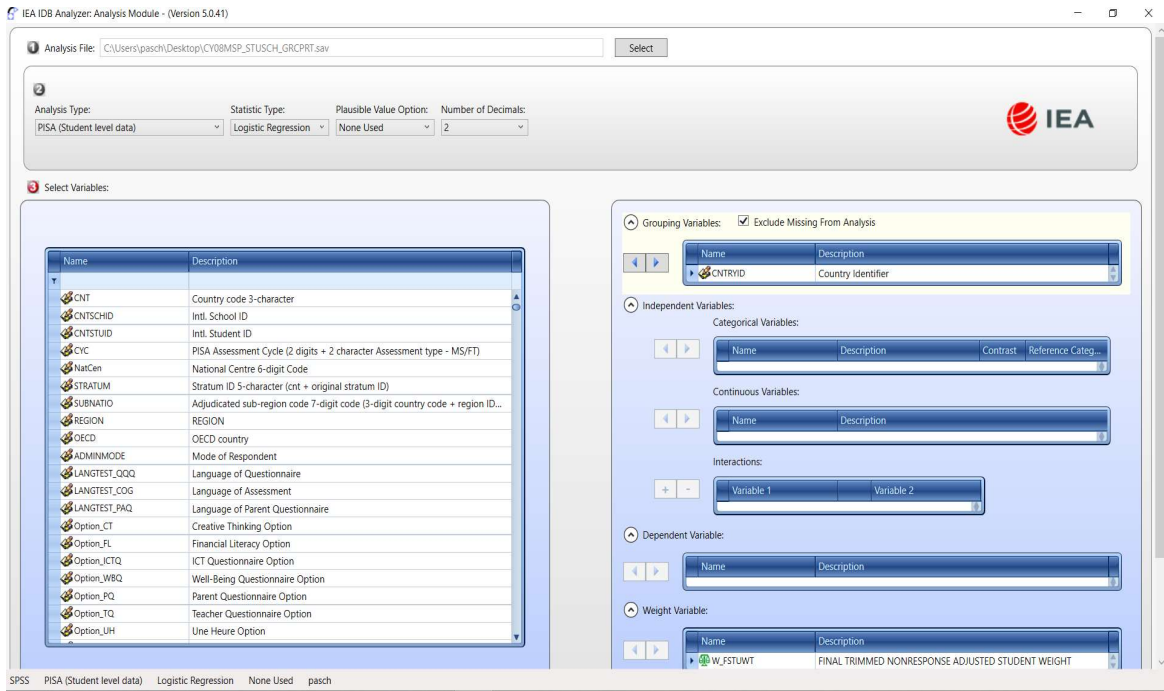


Εικόνα 107. Επιλογή «PISA (Student level)» – Λογιστική Παλινδρόμηση (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

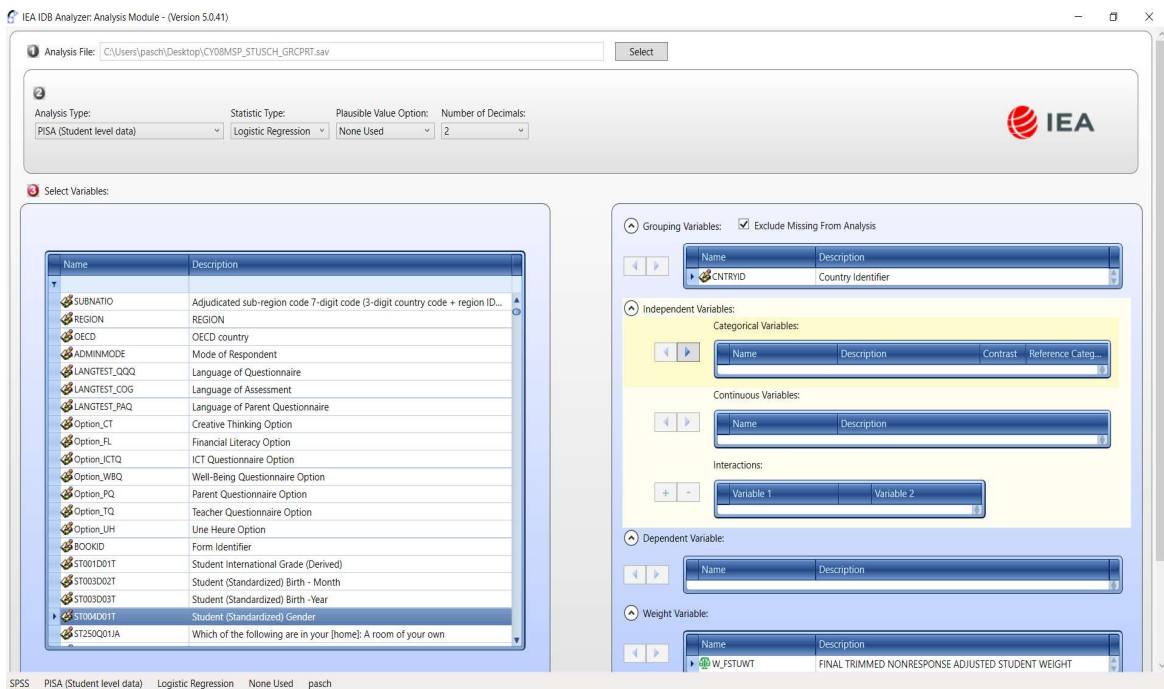
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανεμημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλিপών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγεται «Logistic Regression». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων (Εικόνα 108).



Εικόνα 108. Επιλογή «Logistic Regression» στο πεδίο «Statistic Type» (SPSS)

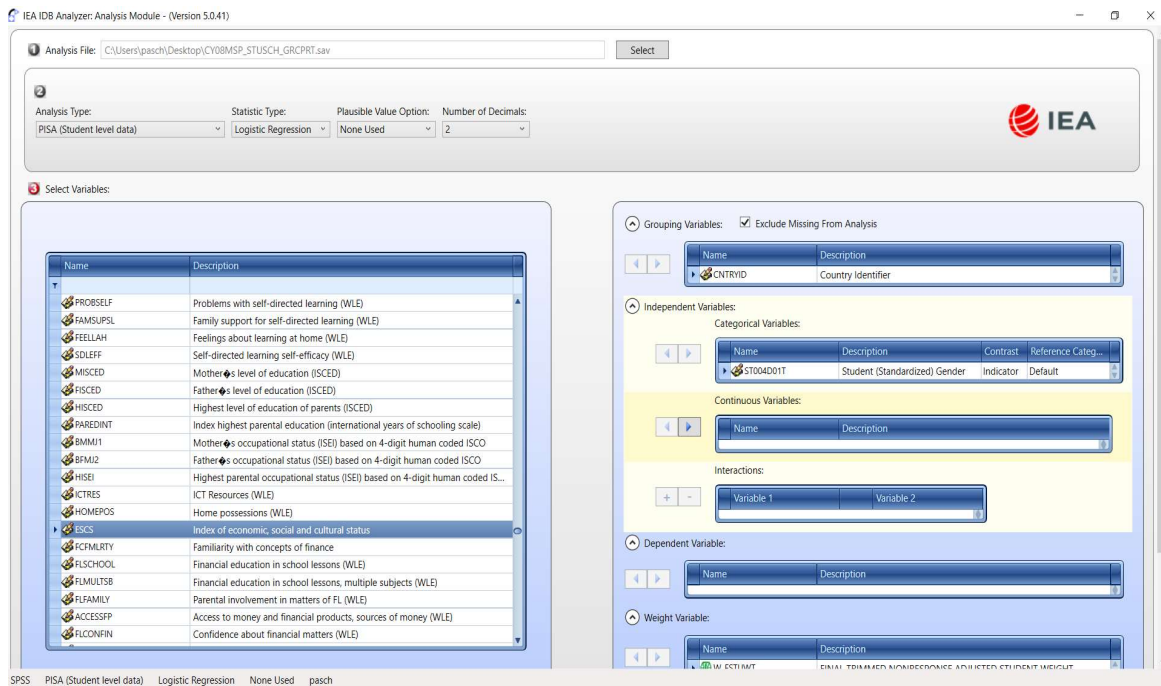
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 108, στο πλαίσιο «Grouping Variables», το οποίο περιλαμβάνει τη λίστα των μεταβλητών εμφανίζεται από προεπιλογή η μεταβλητή «CNTRYID». Επιπλέον, εφόσον η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, στην ανάλυση θα συμπεριληφθούν μόνο οι περιπτώσεις με μη ελλείψεις στις μεταβλητές.



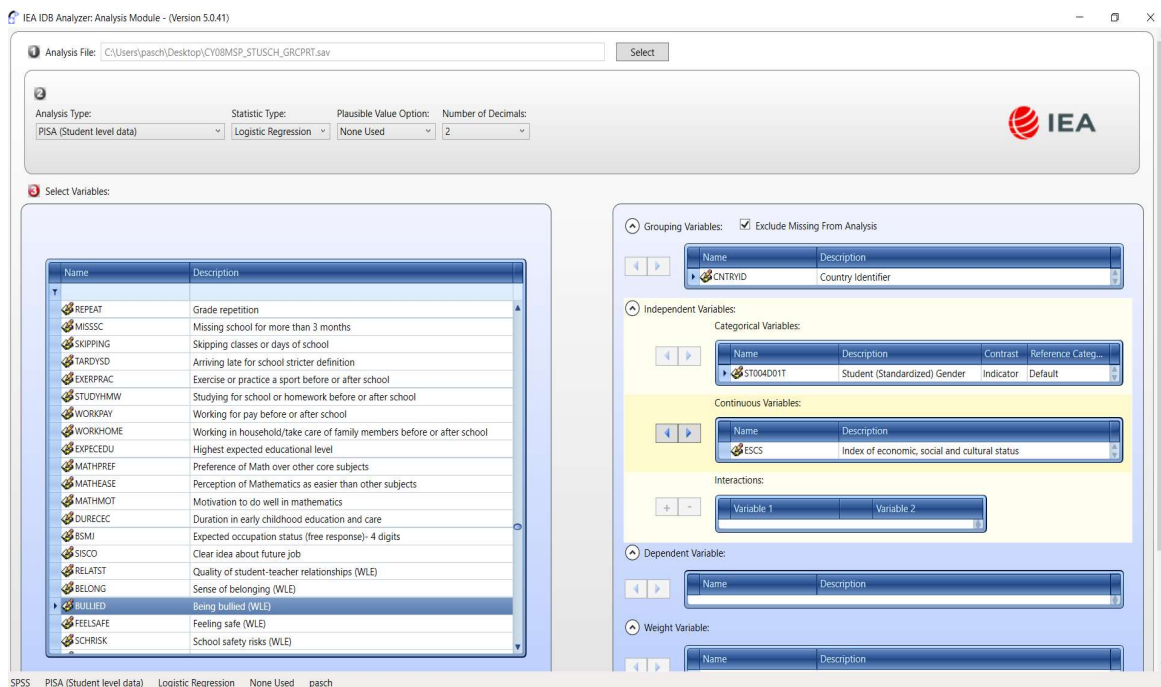
Εικόνα 109. Τοποθέτηση μεταβλητών λογιστικής παλινδρόμησης (SPSS)

Για την υλοποίηση της λογιστικής παλινδρόμησης, είναι απαραίτητος ο εντοπισμός και η σωστή τοποθέτηση των απαιτούμενων μεταβλητών στο πλαίσιο «Select Variables». Αρχικά, εντοπίζεται η μεταβλητή «ST004D01T» (φύλο μαθητή), η οποία επιλέγεται και, με το μπλε βελάκι μεταφέρεται στο πλαίσιο «Independent Variables» και συγκεκριμένα στο πλαίσιο «Categorical Variables» (Εικόνα 109).

Στη συνέχεια, η ίδια διαδικασία επαναλαμβάνεται για τις μεταβλητές «ESCS» και «BULLIED», οι οποίες μεταφέρονται στο πλαίσιο «Continuous Variables», καθώς πρόκειται για συνεχείς μεταβλητές (Εικόνα 110 και Εικόνα 111).

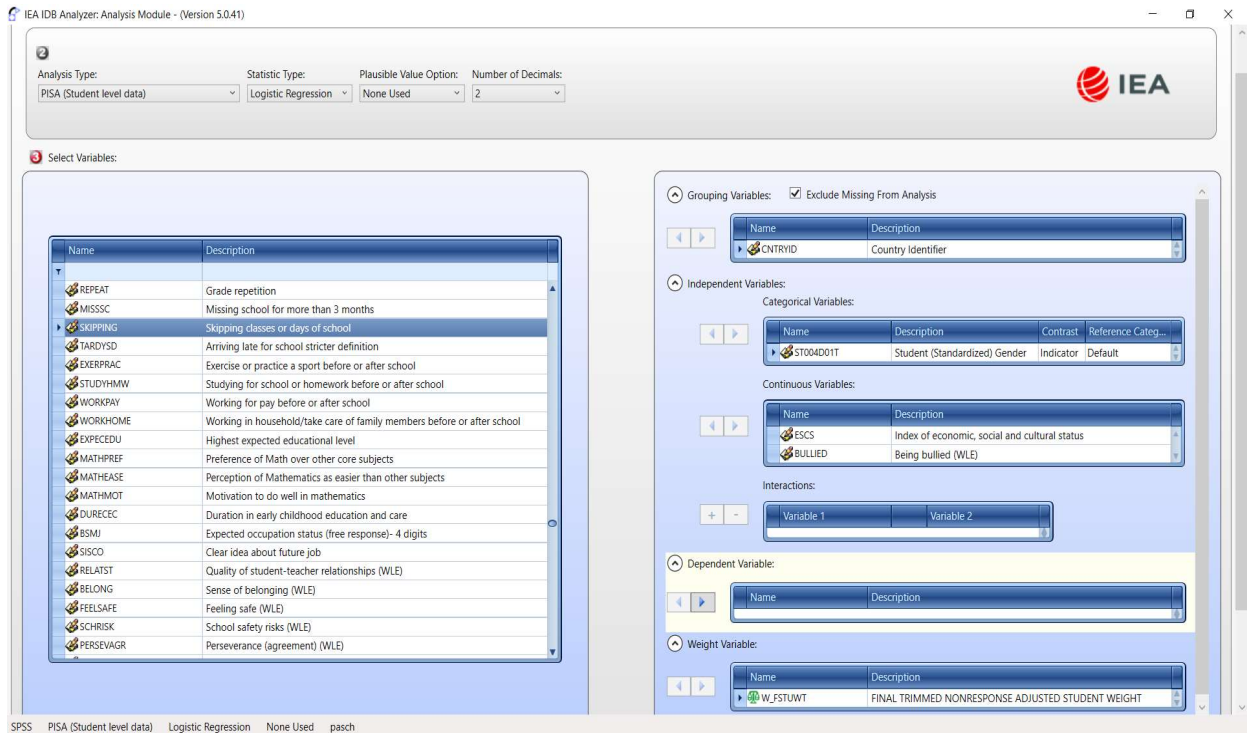


Εικόνα 110. Επιλογή εξαρτημένης μεταβλητής ESCS (SPSS)



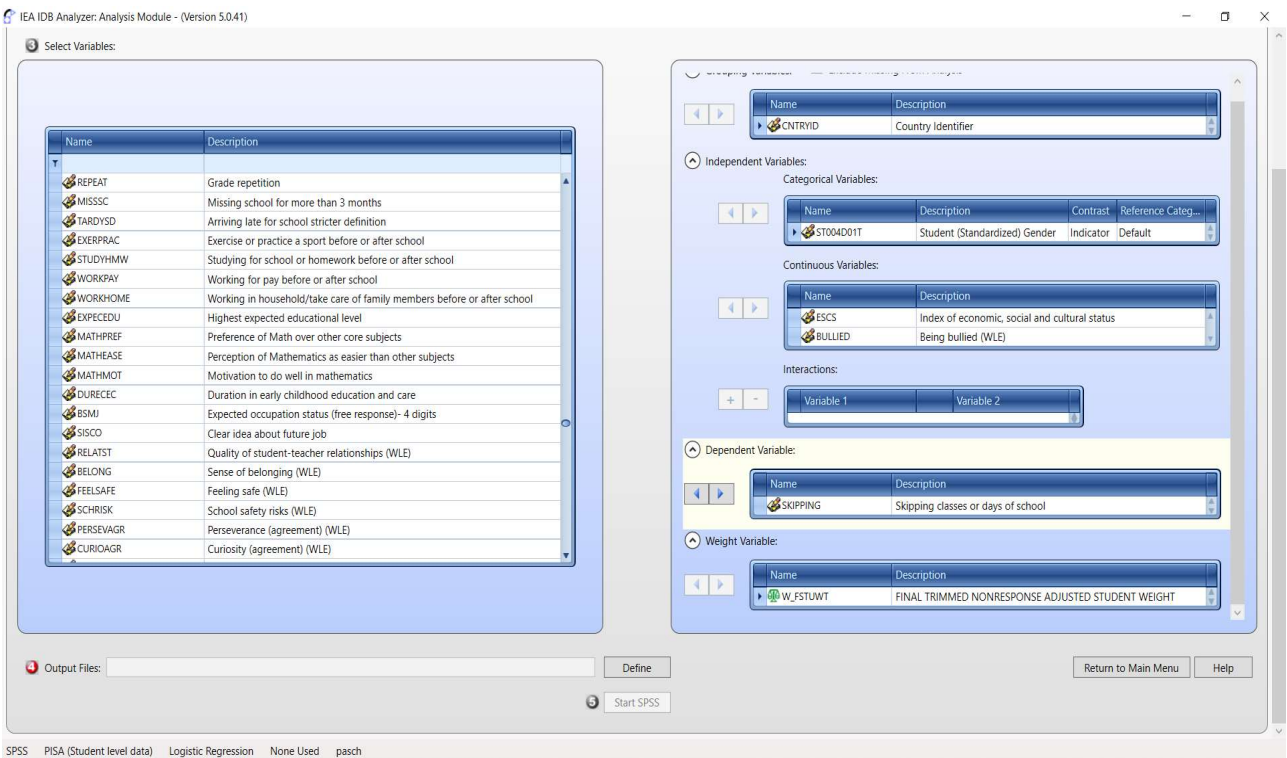
Εικόνα 111. Επιλογή εξαρτημένης μεταβλητής BULLIED (SPSS)

Ακολούθως, στο πλαίσιο «Dependent Variable», επιλέγεται η μεταβλητή «SKIPPING», η οποία χρησιμοποιείται ως εξαρτημένη μεταβλητή στη γραμμική παλινδρόμηση. Τέλος, στο πλαίσιο «Weight Variable» εμφανίζεται από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα την κατάλληλη μεταβλητή βαρών για την εκτέλεση της ανάλυσης (Εικόνα 112).

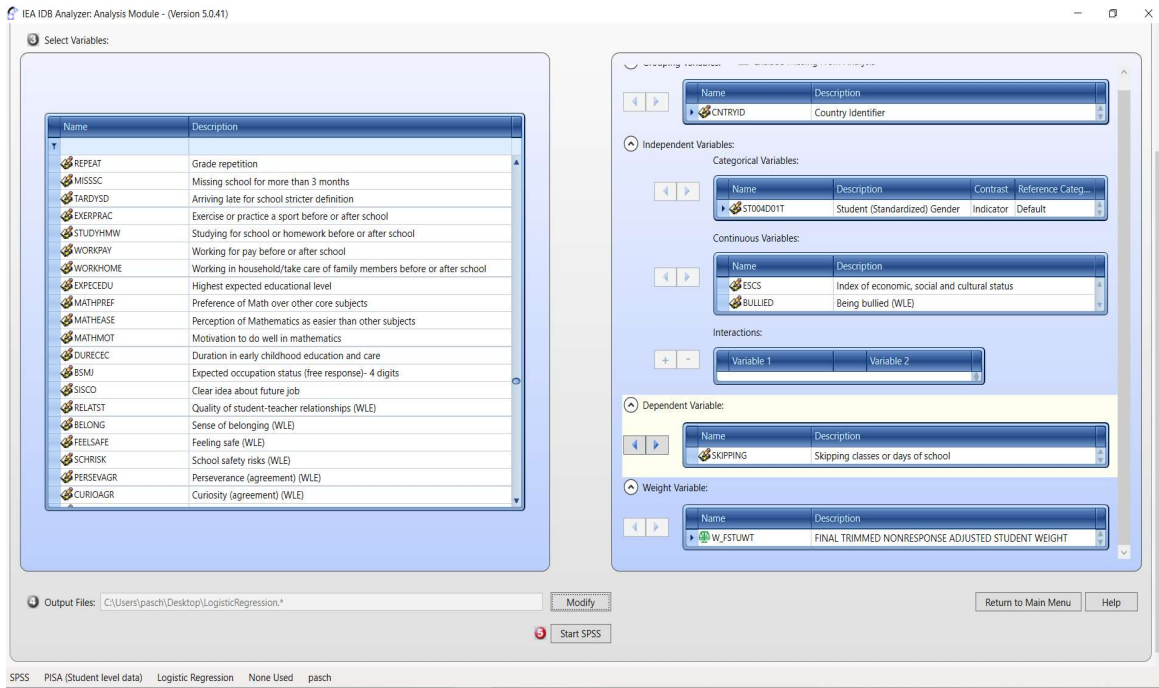


Εικόνα 112. Επιλογή εξαρτημένης μεταβλητής SKIPPING στο πεδίο «Dependent Variable» (SPSS)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «LogisticRegression» (Εικόνα 113). Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 114).

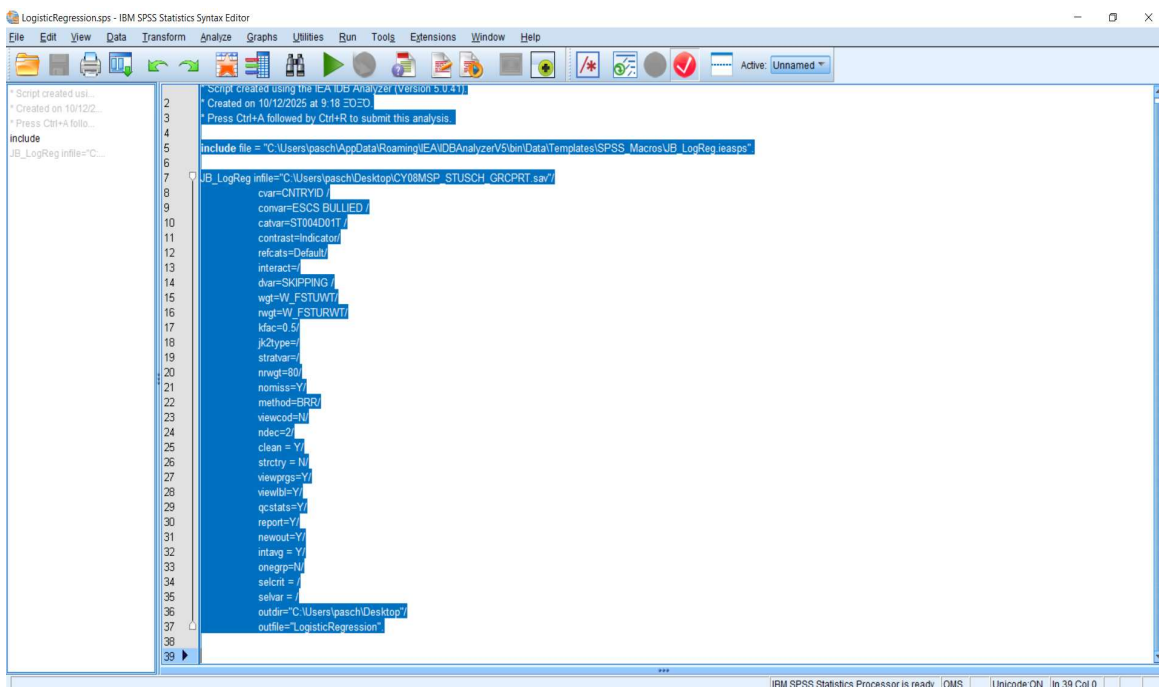


Εικόνα 113. Ορισμός ονόματος εξόδου για λογιστική παλινδρόμηση (SPSS)



Εικόνα 114. Εκκίνηση παραγωγής syntax λογιστικής παλινδρόμησης μέσω «Start SPSS»

Ύστερα από τη μετάβαση στο SPSS, στο αρχείο LogisticRegression.sps που εμφανίζεται, με Ctrl+A επιλέξετε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 115).



Εικόνα 115. Εκτέλεση κώδικα LogisticRegression.sps στο SPSS

Στο αρχείο LogisticRegression.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 116 και 117 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

Computing variance terms...

Model Statistics

Report

Model Statistics

CNTRYID	-2 Log Likelihood	-2 LL (s.e.)	Cox & Snell R Square	Cox & Snell (s.e.)	Nagelkerke R Square	Nagelkerke (s.e.)
Greece	126194,80	2390,45	,01	,00	,02	,00
Portugal	67926,29	2359,49	,03	,01	,06	,01
Table Av	.	.	,02	,00	,04	,01

Regression Coefficients

Report

Εικόνα 116. Αποτελέσματα λογιστικής παλινδρόμησης – στατιστικά (SPSS)

Ο πίνακας της Εικόνας 116 παρουσιάζει, στατιστικά μοντέλου για κάθε χώρα χωριστά (εδώ Ελλάδα, Πορτογαλία), όπως επιλέγονται από τον χρήστη. Συγκεκριμένα οι στήλες δείχνουν:

- CNTRYID: Η χώρα (π.χ. Ελλάδα, Πορτογαλία).
- -2 Log Likelihood (-2LL): Δείκτης προσαρμογής του μοντέλου. Όσο μικρότερος, τόσο καλύτερα «ταιριάζει» το μοντέλο στα δεδομένα.
- -2 LL (s.e.): Το τυπικό σφάλμα (standard error) του -2LL.
- Cox & Snell R-square: Συντελεστής προσδιορισμού. Όσο μεγαλύτερο είναι τόσο καλύτερα εξηγεί το μοντέλο τα δεδομένα.
- Cox & Snell (s.e.): Τυπικό σφάλμα του Cox & Snell R-square.
- Nagelkerke R-square: Εναλλακτικό R-square (παραλλαγή του Cox & Snell).
- Nagelkerke (s.e.): Τυπικό σφάλμα του Nagelkerke R-square.

CNTRYID	Variable	Regression Coefficient	Regression Coefficient (s.e.)	Wald Statistic	Degrees of Freedom	Significance	Exp (b)	Exp (b) (s.e.)
Greece	CONSTANT	,42	,06	55,644	1	,000	1,52	,08
	ESCS	-,09	,04	4,552	1	,033	,92	,04
	BULLIED	,21	,04	36,987	1	,000	1,24	,04
	ST004D01T (1)	-,01	,05	,050	1	,823	,99	,05
Portugal	CONSTANT	-1,74	,07	681,181	1	,000	,18	,01
	ESCS	-,25	,05	29,929	1	,000	,78	,04
	BULLIED	,45	,03	164,958	1	,000	1,56	,05
	ST004D01T (1)	-,20	,10	3,804	1	,051	,82	,09
Table Average	CONSTANT	-,66	,04	231,780	1	,000	,85	,04
	ESCS	-,17	,03	30,280	1	,000	,85	,03
	BULLIED	,33	,02	178,378	1	,000	1,40	,04
	ST004D01T (1)	-,11	,06	3,436	1	,064	,90	,05

Εικόνα 117. Συντελεστές λογιστικής παλινδρόμησης – Odds Ratios (SPSS)

Ο πίνακας της Εικόνας 117 παρουσιάζει, πληροφορίες σχετικά με τα μοντέλα για κάθε χώρα χωριστά (εδώ Ελλάδα, Πορτογαλία). Συγκεκριμένα οι στήλες δείχνουν:

- CNTRYID: Η χώρα επιλογής (π.χ. Ελλάδα, Πορτογαλία).
- Variable: Το όνομα της μεταβλητής που περιλαμβάνεται στο μοντέλο (ESCS, BULLIED, ST004D01T, CONSTANT).
- Regression Coefficient: Ο συντελεστής παλινδρόμησης που δείχνει αν και πόσο η μεταβλητή αυξάνει ή μειώνει την πιθανότητα εμφάνισης του φαινομένου (π.χ. SKIPPING).
- Regression Coefficient (s.e.): Το τυπικό σφάλμα του συντελεστή παλινδρόμησης.
- Wald Statistic: Η τιμή του στατιστικού ελέγχου.
- Degrees of Freedom: Οι βαθμοί ελευθερίας του ελέγχου.
- Significance: Το επίπεδο στατιστικής σημαντικότητας (p-value), το οποίο δείχνει αν η επίδραση της μεταβλητής είναι στατιστικά σημαντική.
- Exp(b): Ο λόγος πιθανοτήτων (odds ratio), που δείχνει πόσες φορές αυξάνονται ή μειώνονται οι πιθανότητες εμφάνισης του φαινομένου για αύξηση της μεταβλητής κατά μία μονάδα.
- Exp(b) (s.e.): Το τυπικό σφάλμα του λόγου πιθανοτήτων.

Ερμηνεία αποτελεσμάτων

Παρουσιάζονται τα αποτελέσματα της λογιστικής παλινδρόμησης (logistic regression) για κάθε χώρα, με εξαρτημένη μεταβλητή το SKIPPING (0 = δεν δήλωσε παράλειψη μαθημάτων/ημερών σχολείου, 1 = δήλωσε παράλειψη). Ως ανεξάρτητες μεταβλητές συμπεριλήφθηκαν το φύλο (ST004D01T), το κοινωνικοοικονομικό επίπεδο (ESCS) και ο δείκτης εκφοβισμού (BULLIED). Η ερμηνεία των συντελεστών βασίζεται κυρίως στο Exp(b) (odds ratio), το οποίο δείχνει τη μεταβολή στις πιθανότητες (odds) εμφάνισης του σχολικού skipping για αύξηση της ανεξάρτητης μεταβλητής κατά μία μονάδα, με τις υπόλοιπες μεταβλητές σταθερές.

Ελλάδα

Για την Ελλάδα, οι δείκτες προσαρμογής του μοντέλου ήταν χαμηλοί (Cox & Snell $R^2 = .01$, Nagelkerke $R^2 = .02$), υποδηλώνοντας περιορισμένη εξηγηματική ικανότητα. Το ESCS αποτέλεσε στατιστικά σημαντικό αρνητικό προγνωστικό παράγοντα για την εμφάνιση σχολικού skipping ($p = .033$), με $\text{Exp}(b) = 0.92$, δηλαδή κάθε αύξηση του ESCS κατά 1 μονάδα συνδέθηκε με μείωση περίπου 8% στις πιθανότητες εμφάνισης skipping. Αντίθετα, ο δείκτης εκφοβισμού (BULLIED) ήταν στατιστικά σημαντικός θετικός παράγοντας ($p < .001$), με $\text{Exp}(b) = 1.24$, γεγονός που υποδηλώνει αύξηση περίπου 24% στις πιθανότητες εμφάνισης σχολικού skipping για κάθε αύξηση κατά 1 μονάδα στον δείκτη εκφοβισμού. Η επίδραση του φύλου (ST004D01T) δεν ήταν στατιστικά σημαντική ($p = .823$, $\text{Exp}(b) = 0.99$). Συνολικά, στην Ελλάδα η εμφάνιση σχολικού skipping σχετίστηκε κυρίως με υψηλότερες εμπειρίες εκφοβισμού, ενώ το υψηλότερο κοινωνικοοικονομικό επίπεδο συνδέθηκε με μειωμένη πιθανότητα skipping.

Πορτογαλία

Για την Πορτογαλία, οι δείκτες προσαρμογής ήταν υψηλότεροι σε σχέση με την Ελλάδα, αν και συνολικά παρέμειναν σχετικά χαμηλοί (Cox & Snell $R^2 = .03$, Nagelkerke $R^2 = .06$). Το ESCS ήταν στατιστικά σημαντικός αρνητικός παράγοντας ($p < .001$), με $\text{Exp}(b) = 0.78$, δηλαδή κάθε αύξηση κατά 1 μονάδα στο κοινωνικοοικονομικό επίπεδο συνδέθηκε με μείωση περίπου 22% στις πιθανότητες εμφάνισης skipping. Ο δείκτης εκφοβισμού (BULLIED) αποτέλεσε επίσης στατιστικά σημαντικό θετικό παράγοντα ($p < .001$), με $\text{Exp}(b) = 1.56$, υποδηλώνοντας αύξηση περίπου 56% στις πιθανότητες εμφάνισης σχολικού skipping για κάθε αύξηση κατά 1 μονάδα στον εκφοβισμό. Η επίδραση του φύλου (ST004D01T) δεν ήταν στατιστικά σημαντική ($p = .051$, $\text{Exp}(b) = 0.82$). Συνολικά, στην Πορτογαλία ο εκφοβισμός αναδείχθηκε ως ο ισχυρότερος παράγοντας που σχετίζεται με το σχολικό skipping, με επίδραση εντονότερη από την αντίστοιχη στην Ελλάδα, ενώ το υψηλότερο κοινωνικοοικονομικό επίπεδο συνδέθηκε με σημαντικά μειωμένη πιθανότητα skipping.

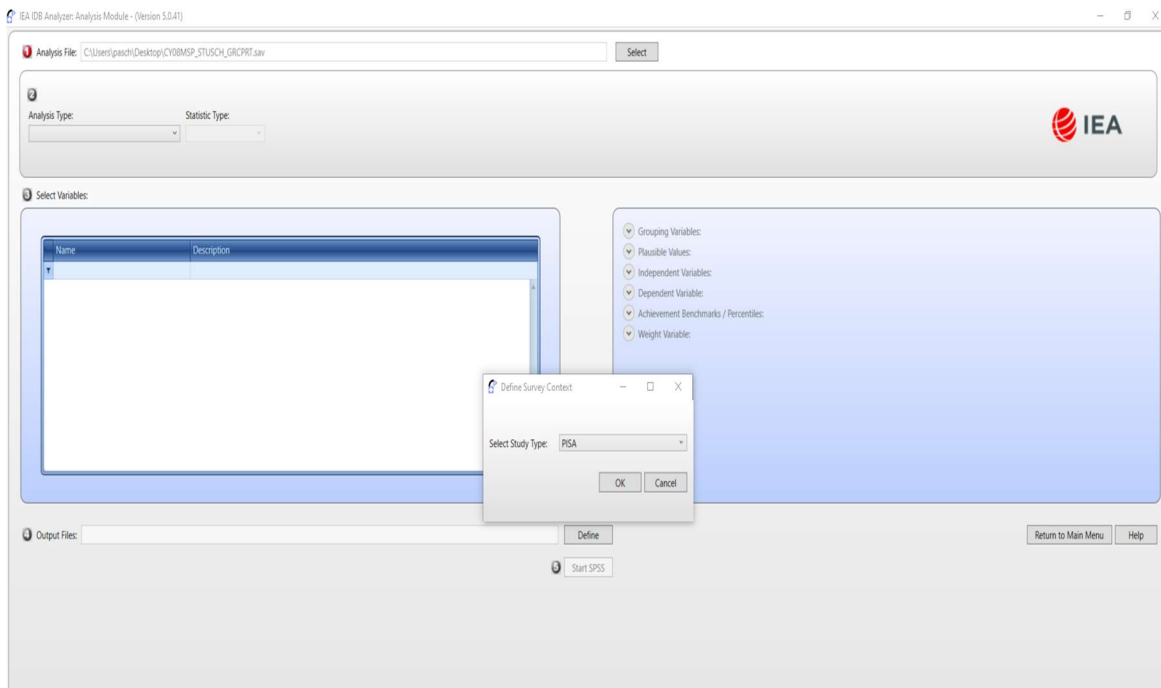
Υπολογισμός ποσοστών και μέσης τιμής

Ερευνητική υπόθεση: Διερευνώνται διαφορές στα ποσοστά και στις μέσες τιμές του άγχους για τα Μαθηματικά μεταξύ των δύο χωρών, λαμβάνοντας υπόψη το φύλο των μαθητών.

Επιλογή μεταβλητών: ANXMAT, ST004D01T

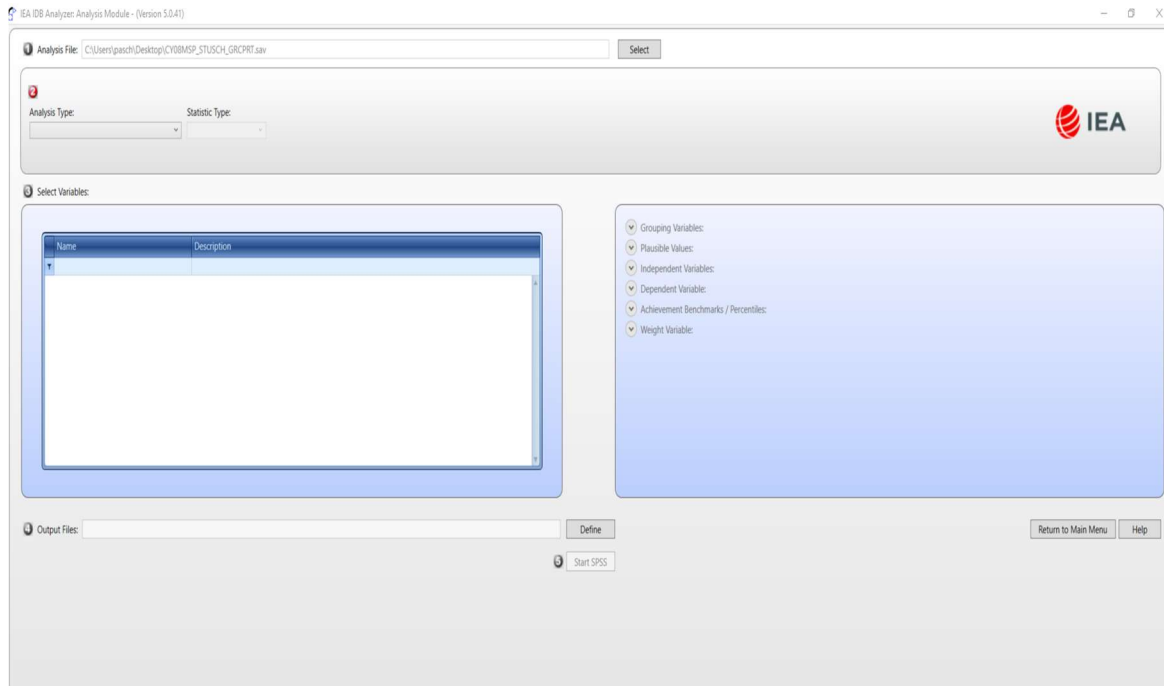
Χαρακτηριστικά μεταβλητών: ANXMAT συνεχής μεταβλητή, ST004D01T κατηγορική μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί σύγκριση ποσοστών ανάμεσα στην Ελλάδα και την Πορτογαλία, καθώς και ο υπολογισμός της μέσης τιμής, με μεταβλητή ενδιαφέροντος το άγχος για τα Μαθηματικά (ANXMAT). Η δεύτερη μεταβλητή που λαμβάνεται υπόψη είναι η «ST004D01T», η οποία καταγράφει το φύλο των μαθητών. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPR1.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 118).



Εικόνα 118. Επιλογή αρχείου δεδομένων για ανάλυση Percentages & Means (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 119).

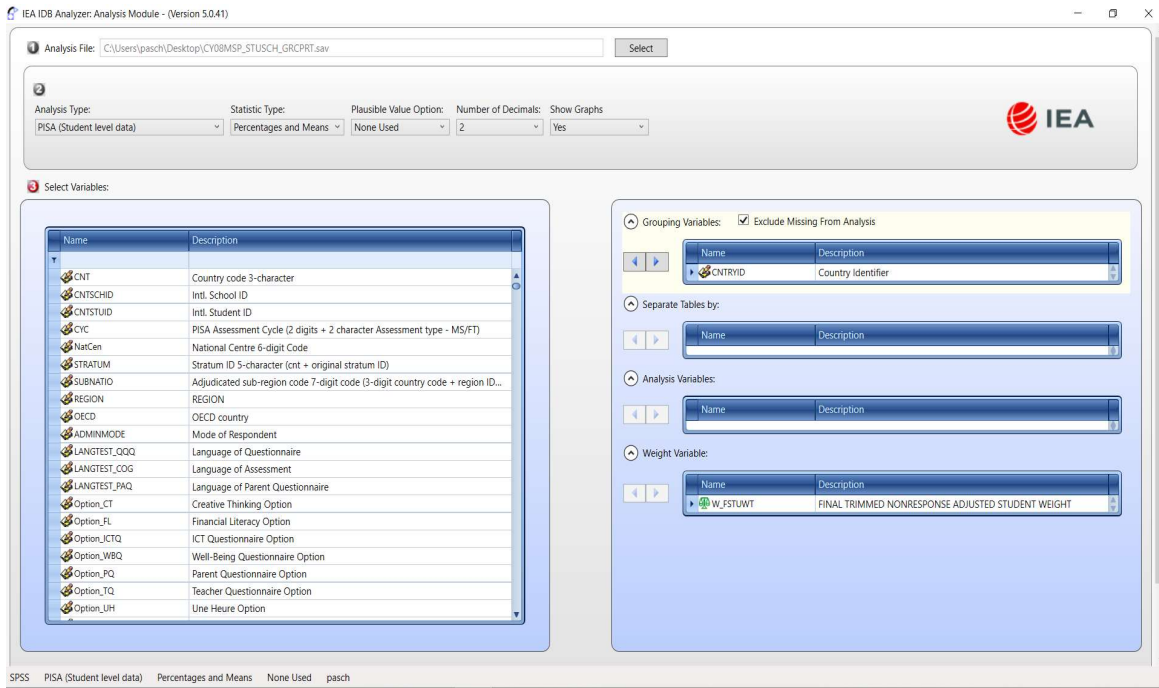


Εικόνα 119. Επιλογή «PISA (Student level)» – Percentages & Means (SPSS)

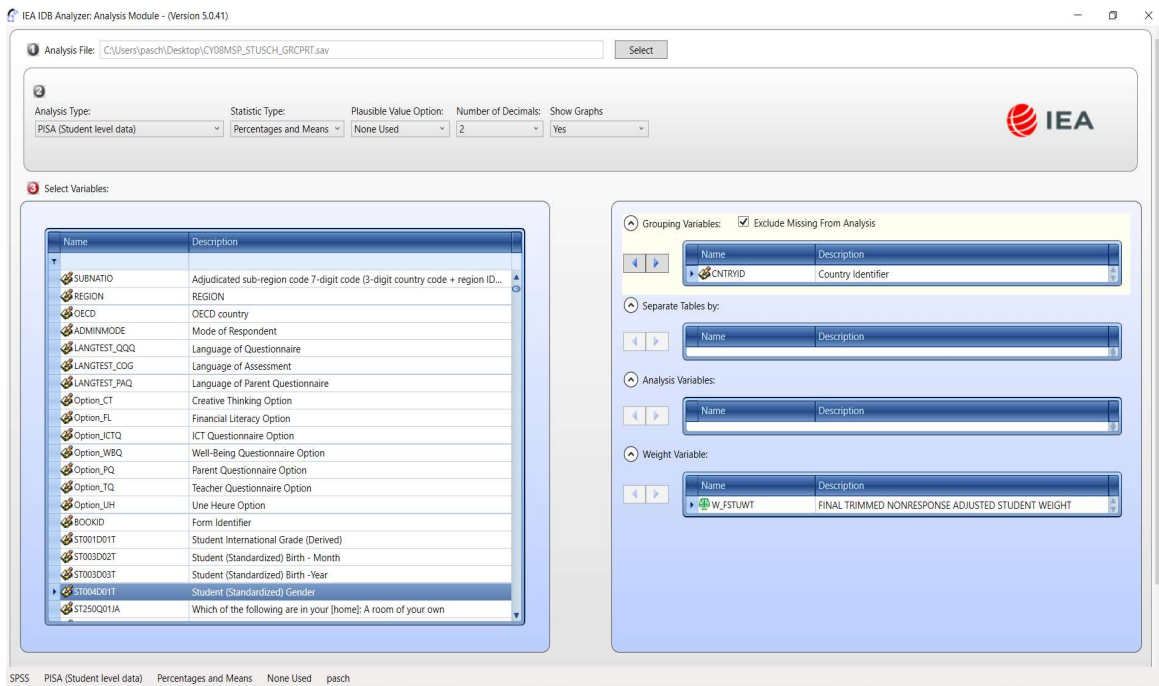
Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγεται «Percentages and Means». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων (Εικόνα 120).



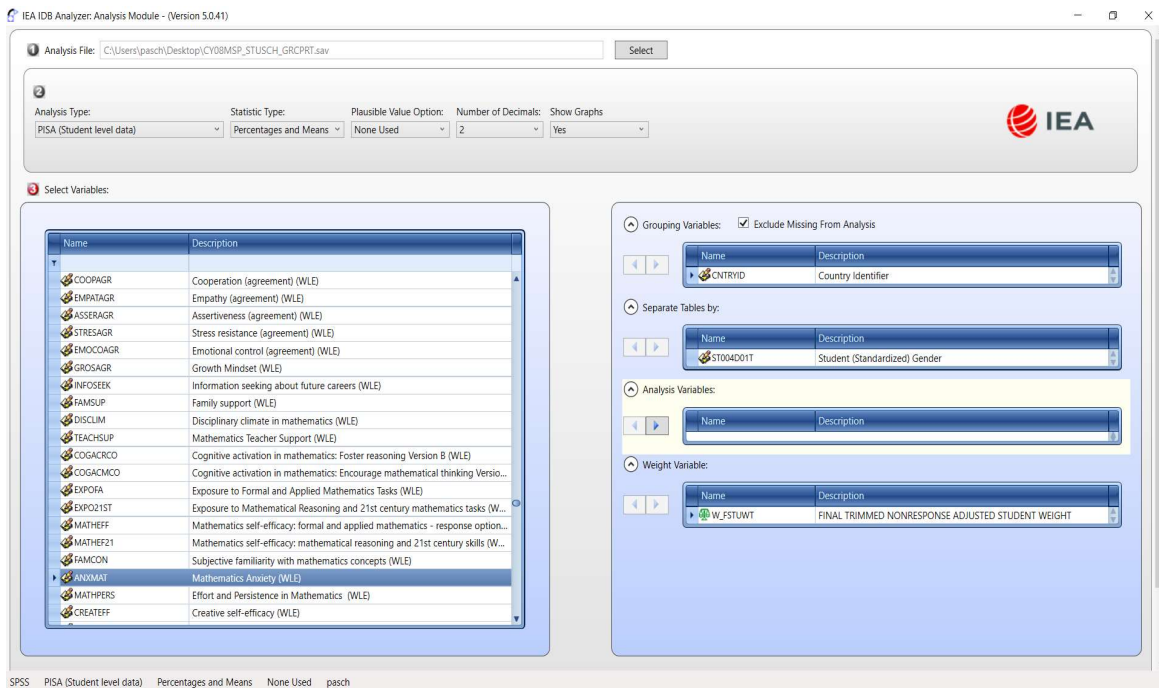
Εικόνα 120. Επιλογή «Percentages and Means» (SPSS)



Εικόνα 121. Τοποθέτηση μεταβλητής ST004D01T (SPSS)

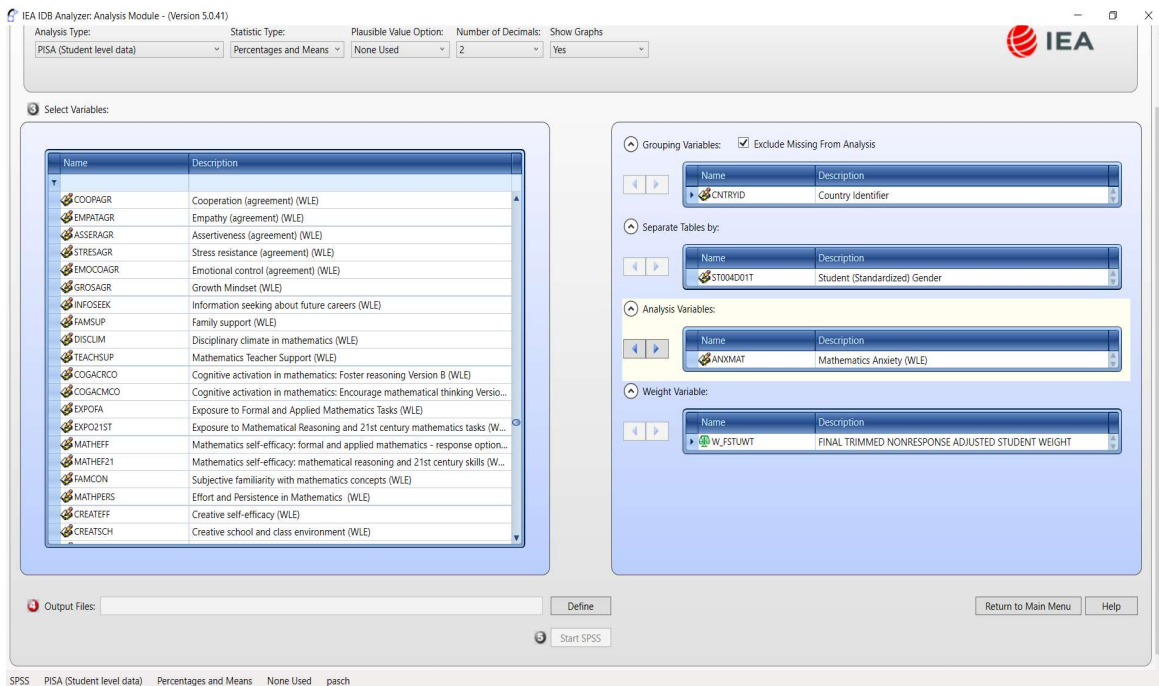
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 121 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλείψεις στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί η σύγκριση, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή «ST004D01T» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Separate Tables by» την τοποθετείτε (Εικόνα 121). Επαναλαμβάνετε την διαδικασία

για την μεταβλητή «ANXMAT» την οποία τοποθετείτε στο πλαίσιο «Analysis Variables». Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 122).



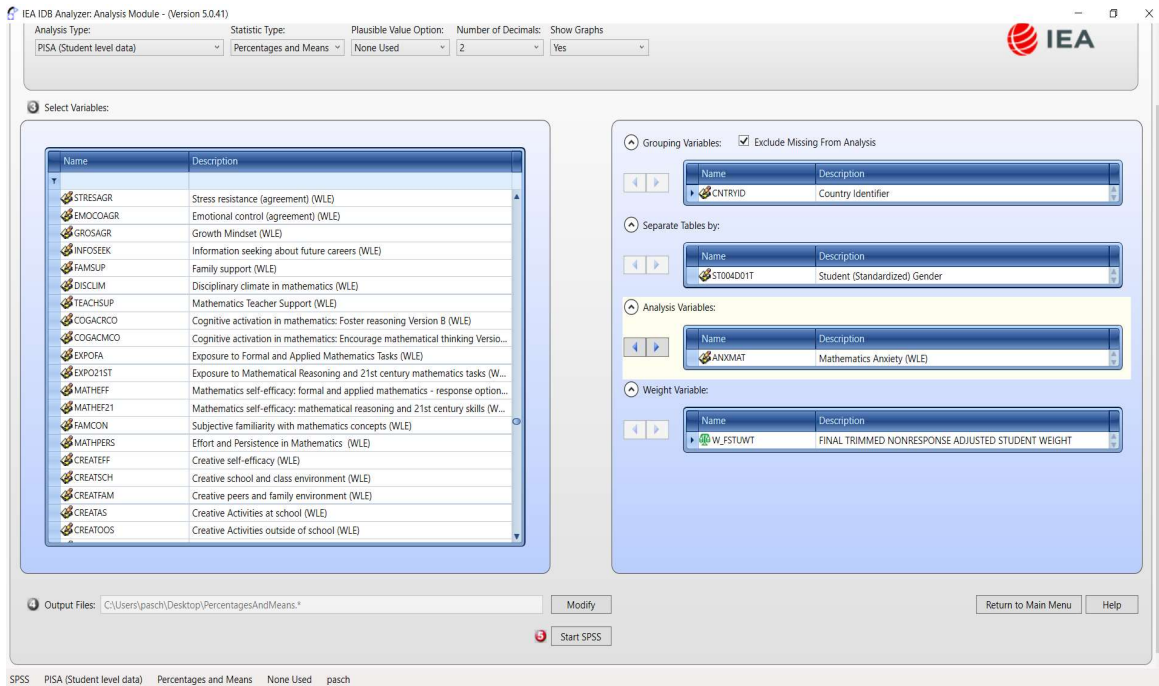
Εικόνα 122. Τοποθέτηση μεταβλητής ANXMAT (SPSS)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PercentagesAndMeans» (Εικόνα 123).



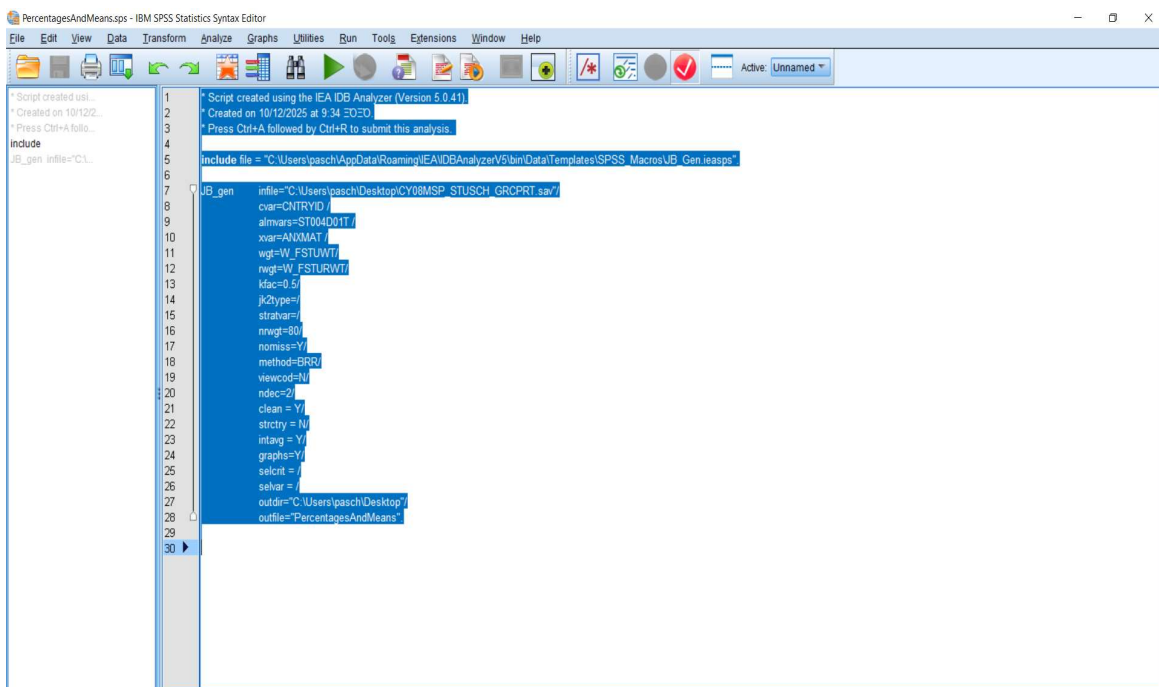
Εικόνα 123. Ορισμός ονόματος εξόδου – Percentages & Means (SPSS)

Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 124).



Εικόνα 124. Εκκίνηση παραγωγής syntax μέσω «Start SPSS» – Percentages & Means

Ύστερα από τη μετάβαση στο SPSS, στο αρχείο PercentagesAndMeans.sps που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 125).



Εικόνα 125. Εκτέλεση κώδικα PercentagesAndMeans.sps στο SPSS

Στο αρχείο PercentagesAndMeans.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 126 έως και 129 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

Report

C:\Users\pasch\Desktop\tmpANXMAT

Average for ANXMAT by (COUNTRYID ST004D01T)

Country Identifier	Student (Standardized) Gender	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent	Percent (s.e.)	ANXMAT (Mean)	ANXMAT (s.e.)	Std.Dev.	Std.Dev. (s.e.)	Percent Missing
Greece	Female	2953	44531,83	991,57	50,71	,66	,47	,02	1,16	,02	7,84
	Male	2817	43285,08	1131,46	49,29	,66	,04	,03	1,11	,02	13,03
Portugal	Female	2913	41071,37	1139,46	48,98	,68	,30	,02	1,01	,02	13,84
	Male	3000	42775,73	1167,79	51,02	,68	-,01	,02	1,03	,02	12,59
Table Average	Female	.	.	.	49,85	,48	,39	,01	1,09	,01	.
	Male	.	.	.	50,15	,48	,01	,02	1,07	,01	.

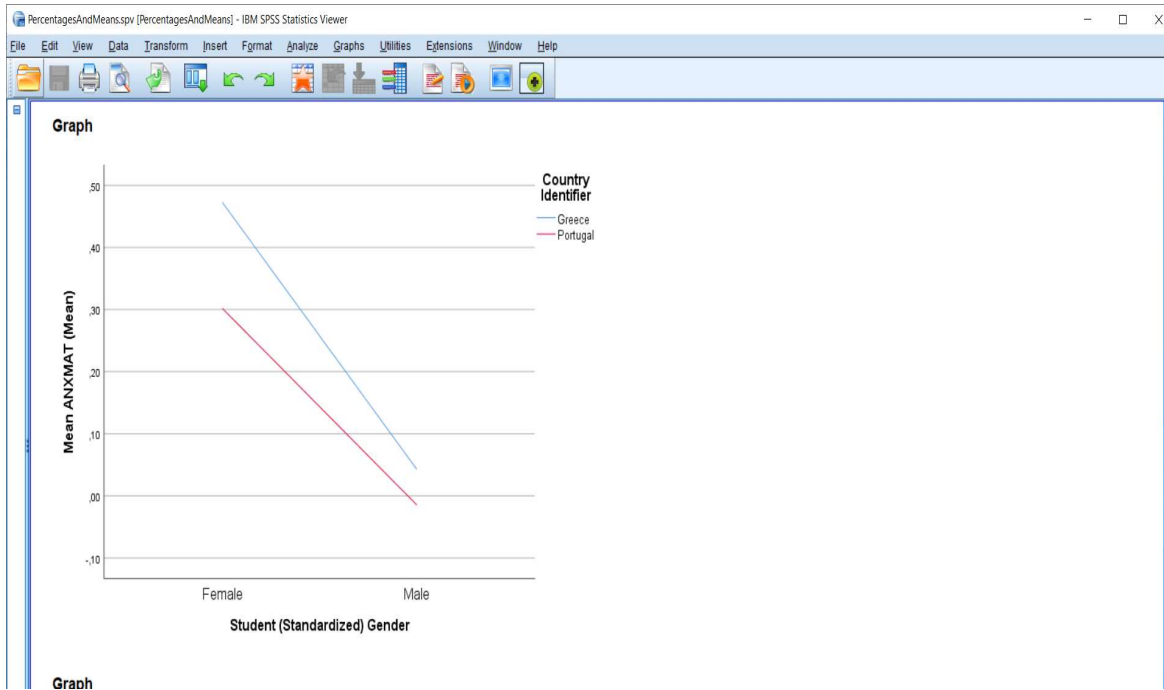
PLEASE NOTE:

THE FOLLOWING GRAPHICS SHOULD BE INTERPRETED TAKING INTO ACCOUNT THE STANDARD ERRORS PRESENTED IN THE PRECEDING TABLE.

Εικόνα 126. Πίνακας ποσοστών και μέσων όρων ανά χώρα και φύλο (SPSS)

Οι στήλες της Εικόνας 126 παρουσιάζουν την κατανομή και τα μέσα επίπεδα άγχους για τα Μαθηματικά (ANXMAT) κατά φύλο (ST004D01T) για κάθε χώρα. Συγκεκριμένα:

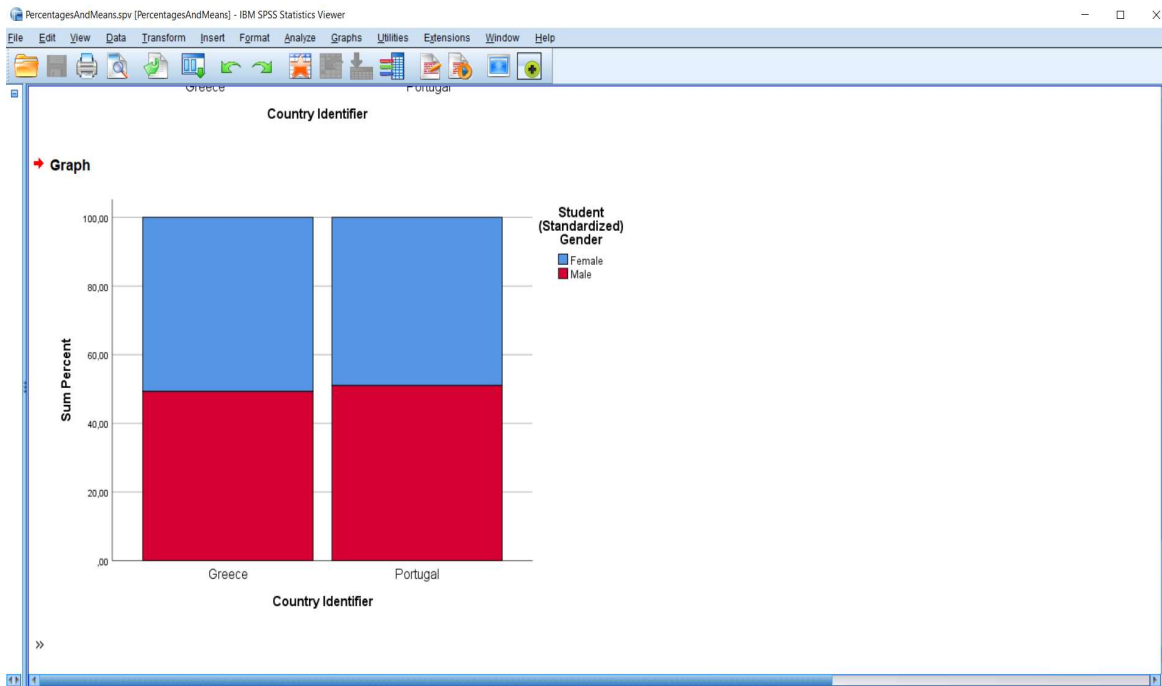
- Country Identifier: Η χώρα στην οποία αναφέρονται τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- Student Gender (Standardized): Το φύλο των μαθητών, κωδικοποιημένο ως Female (κορίτσια) και Male (αγόρια).
- N of Cases: Ο αριθμός μαθητών στο δείγμα κάθε κατηγορίας φύλου.
- Sum of W_FSTUWT: Το συνολικό σταθμισμένο άθροισμα βαρών του δείγματος, το οποίο αντιστοιχεί στον πληθυσμό που εκπροσωπείται.
- Sum of W_FSTUWT (s.e.): Το τυπικό σφάλμα του συνολικού σταθμισμένου αθροίσματος βαρών.
- Percent: Το ποσοστό του πληθυσμού που αντιπροσωπεύει η κάθε κατηγορία φύλου μέσα στη χώρα.
- Percent (s.e.): Το τυπικό σφάλμα της ποσοστιαίας εκτίμησης.
- Mean: Ο μέσος όρος της μαθηματικής επίδοσης (ANXMAT) για το φύλο.
- Mean (s.e.): Το τυπικό σφάλμα του μέσου όρου.
- Std. Dev: Η τυπική απόκλιση της επίδοσης, δηλαδή η διασπορά των επιδόσεων γύρω από τον μέσο όρο.
- Std. Dev (s.e.): Το τυπικό σφάλμα της τυπικής απόκλισης.
- Percent Missing: Ποσοστό ελλιπών δεδομένων.



Εικόνα 127. Γράφημα σύγκρισης μέσων τιμών ανά χώρα και φύλο (SPSS)



Εικόνα 128. Γράφημα σύγκρισης μέσων τιμών ανά χώρα και φύλο (SPSS)



Εικόνα 129. Γράφημα σύγκρισης ποσοστών ανά χώρα και φύλο (SPSS)

Επιπλέον, οι πληροφορίες που έχουμε από τις Εικόνα 127, Εικόνα 128 και Εικόνα 129 όπου ουσιαστικά απεικονίζουν διαγράμματα, είναι οι εξής: στην Εικόνα 127 απεικονίζεται ένα Line chart που μας δείχνει τον μέσο όρο του άγχους για τα Μαθηματικά (ANXMAT) ανά φύλο (Female/Male) για Ελλάδα και Πορτογαλία. Στην Εικόνα 128 εμφανίζεται ένα ομαδοποιημένο ραβδόγραμμα (Bar chart) όπου δείχνει το μέσο όρο του ANXMAT ανά χώρα ξεχωριστά για τα αγόρια και τα κορίτσια. Τέλος το διάγραμμα της Εικόνα 129 δείχνει ένα bar chart κατανομής φύλου (ποσοστά), στους μαθητές κάθε χώρας.

Ερμηνεία αποτελεσμάτων

Για την Ελλάδα, ο μέσος όρος του άγχους για τα Μαθηματικά (ANXMAT) ήταν υψηλότερος στα κορίτσια ($M = 0.47$, $SE = 0.02$) σε σύγκριση με τα αγόρια ($M = 0.04$, $SE = 0.03$). Οι τυπικές αποκλίσεις ($SD = 1.16$ για τα κορίτσια και $SD = 1.11$ για τα αγόρια) υποδηλώνουν επαρκή διασπορά γύρω από τους μέσους όρους. Η κατανομή φύλου ήταν σχεδόν ισόποση (50.71% κορίτσια, 49.29% αγόρια), γεγονός που μειώνει την πιθανότητα οι παρατηρούμενες διαφορές να οφείλονται σε ανισοκατανομή του δείγματος. Οπτικά, στο διάγραμμα (Εικόνα 127), η γραμμή της Ελλάδας παρουσίασε σαφή πτώση από τα κορίτσια προς τα αγόρια, αποτυπώνοντας την υψηλότερη μέση τιμή άγχους των κοριτσιών σε σχέση με τα αγόρια. Στην Πορτογαλία, ο μέσος όρος του ANXMAT ήταν επίσης υψηλότερος στα κορίτσια ($M = 0.30$, $SE = 0.02$) σε σχέση με τα αγόρια ($M = -0.01$, $SE = 0.02$). Οι τυπικές αποκλίσεις ήταν ελαφρώς μικρότερες σε σύγκριση με την Ελλάδα ($SD = 1.01$ για τα κορίτσια και $SD = 1.03$ για τα αγόρια), υποδηλώνοντας μικρότερη διασπορά. Και εδώ, η κατανομή φύλου ήταν σχεδόν ισόποση (48.98% κορίτσια, 51.02% αγόρια). Στην Εικόνα 127, η γραμμή της Πορτογαλίας εμφάνισε επίσης πτωτική πορεία από τα κορίτσια προς τα αγόρια, ωστόσο η διαφορά μεταξύ φύλων ήταν ηπιότερη σε σχέση με την Ελλάδα. Η συγκριτική απεικόνιση στο ομαδοποιημένο ραβδόγραμμα (Εικόνα 128) έδειξε ότι, και για τα δύο φύλα, τα επίπεδα άγχους για τα Μαθηματικά ήταν υψηλότερα στην Ελλάδα από ό,τι στην Πορτογαλία. Επιπλέον, το διάγραμμα ποσοστών (Εικόνα 129) κατέδειξε ότι η συμμετοχή αγοριών και κοριτσιών ήταν σχεδόν ισότιμη και στις δύο χώρες, στοιχείο που ενισχύει την ερμηνεία ότι οι διαφοροποιήσεις στο ANXMAT αντανακλούν ουσιαστικές διαφορές μεταξύ φύλων και χωρών. Συνολικά, τα αποτελέσματα υποδεικνύουν ότι οι μαθήτριες παρουσιάζουν υψηλότερα επίπεδα άγχους για τα Μαθηματικά από τους μαθητές τόσο στην

Ελλάδα όσο και στην Πορτογαλία, με τη διαφορά να είναι εντονότερη στην Ελλάδα. Παράλληλα, το μέσο άγχος ήταν υψηλότερο στην Ελλάδα σε σχέση με την Πορτογαλία, ανεξαρτήτως φύλου.

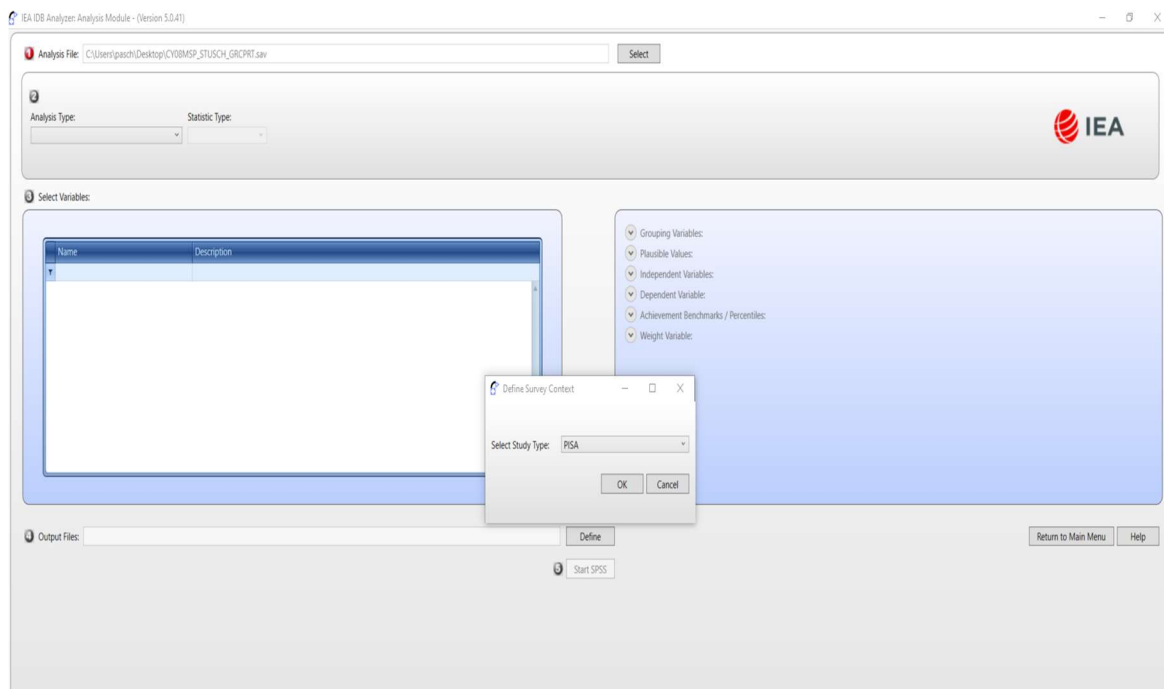
Υπολογισμός ποσοστών

Ερευνητική υπόθεση: Διερευνώνται διαφορές στα ποσοστά προτίμησης για τα Μαθηματικά μεταξύ των δύο χωρών που επιλέχθηκαν.

Επιλογή μεταβλητών: MATHPREF

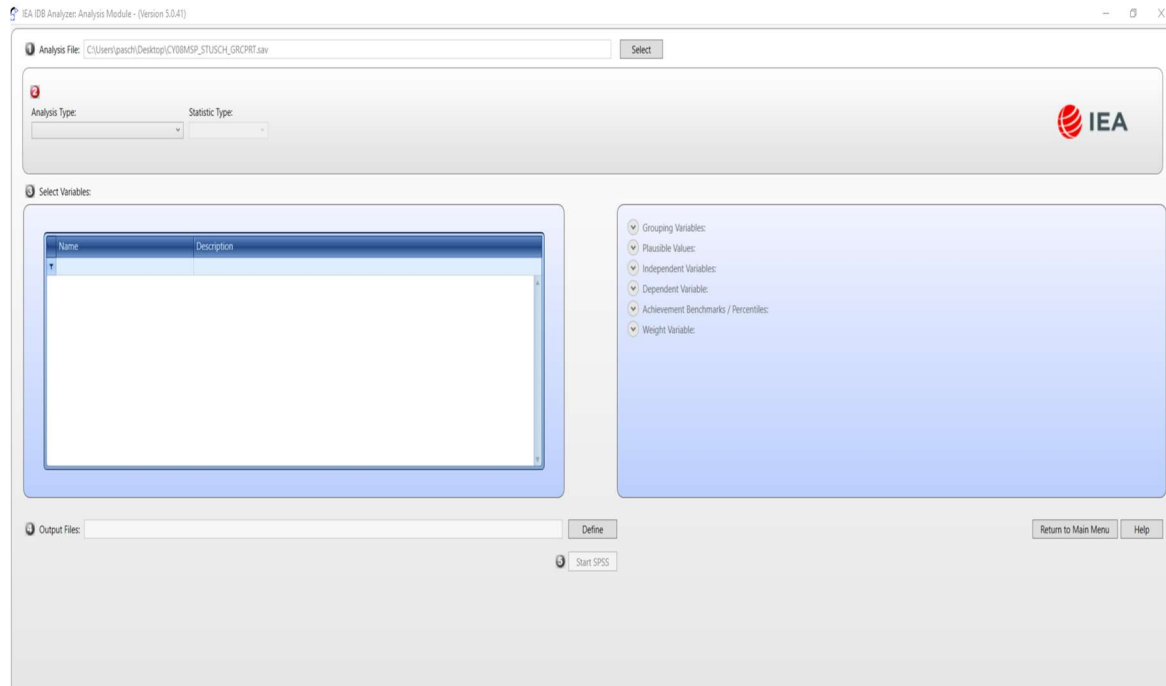
Χαρακτηριστικά μεταβλητών: MATHPREF συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί σύγκριση ποσοστών ανάμεσα στην Ελλάδα και την Πορτογαλία, με μεταβλητή ενδιαφέροντος τη προτίμηση των μαθητών για τα Μαθηματικά, σε σχέση με τα άλλα μαθήματα. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPRT.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 130).



Εικόνα 130. Επιλογή αρχείου δεδομένων για ανάλυση Percentages Only – «Analysis File» (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέγεται ο τύπος της ανάλυσης που πρόκειται να πραγματοποιηθεί, για παράδειγμα «PISA (Student level)» (Εικόνα 131).

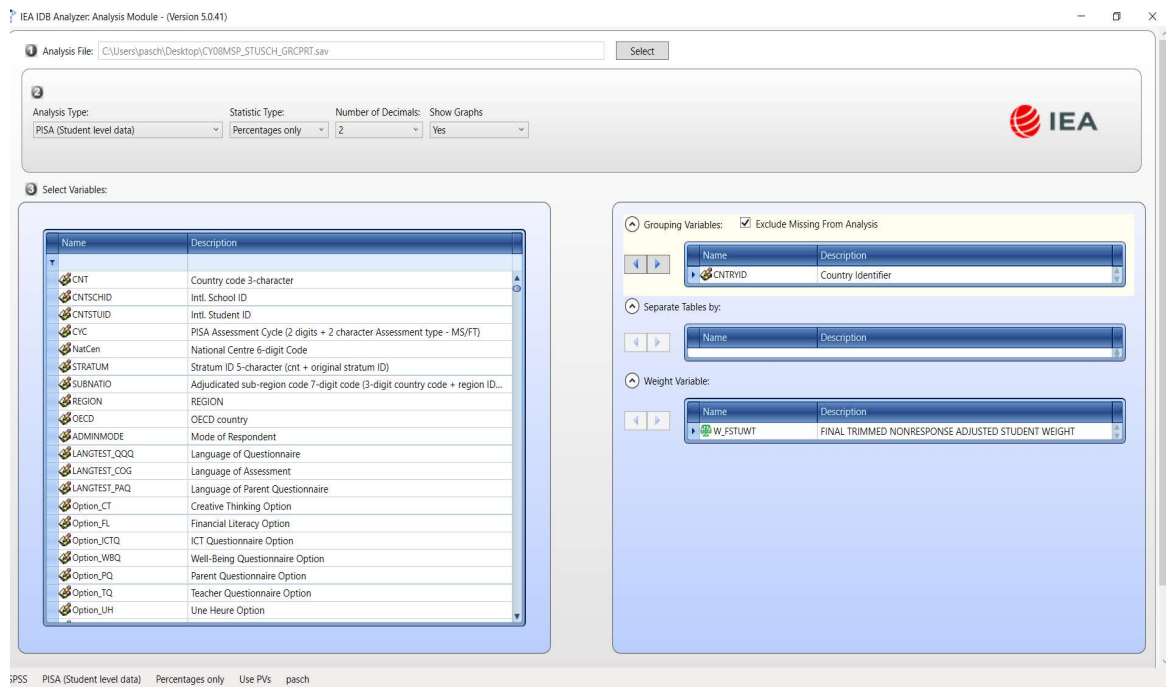


Εικόνα 131. Τοποθέτηση μεταβλητής MATHPREF (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

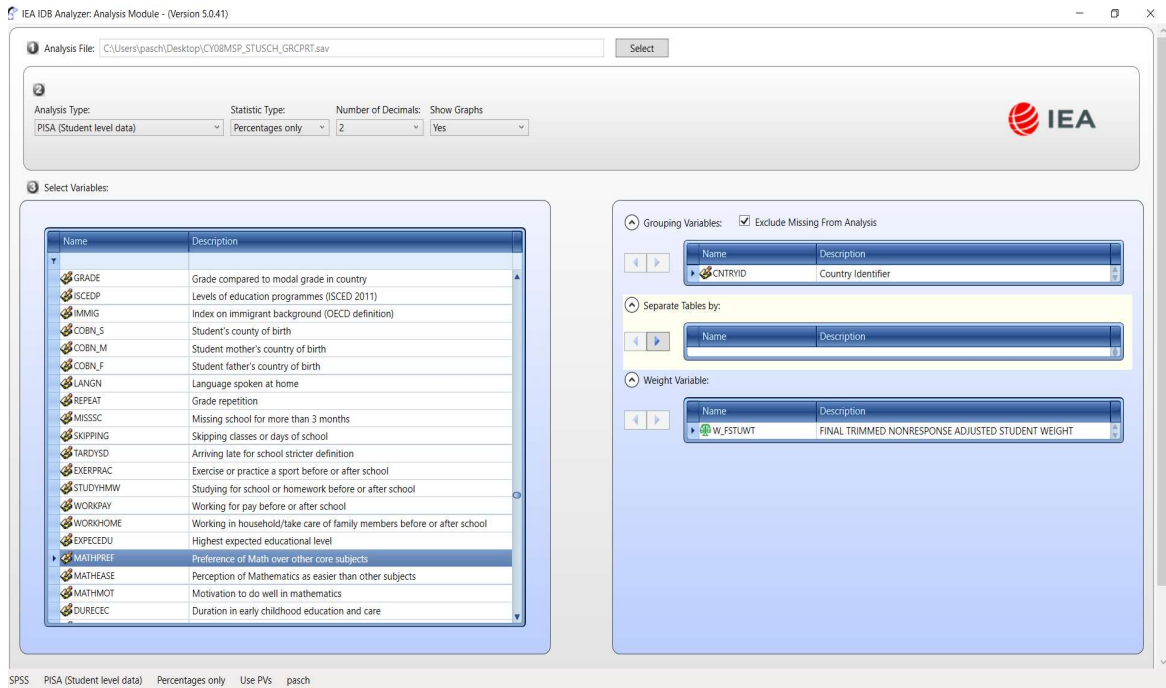
- **Benchmarks:** Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- **Correlation (Pearson):** Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- **Correlation (Spearman):** Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- **Linear regression:** Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- **Percentages and Means:** Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- **Percentages only:** Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- **Percentiles:** Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Percentages only». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 132).



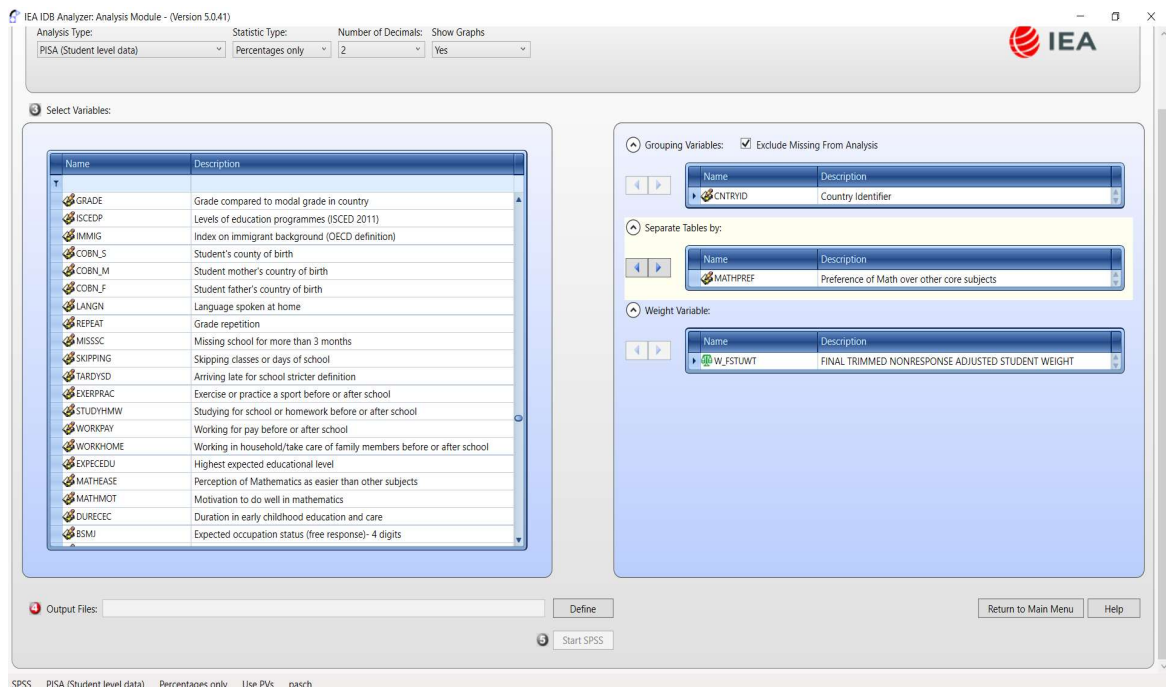
Εικόνα 132. Επιλογή «Percentages only» στο πεδίο «Statistic Type» (SPSS)

Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 132 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλείψεις τιμές στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί η σύγκριση, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή «MATHPREF» από το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Separate Tables by» την τοποθετείτε. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 133).

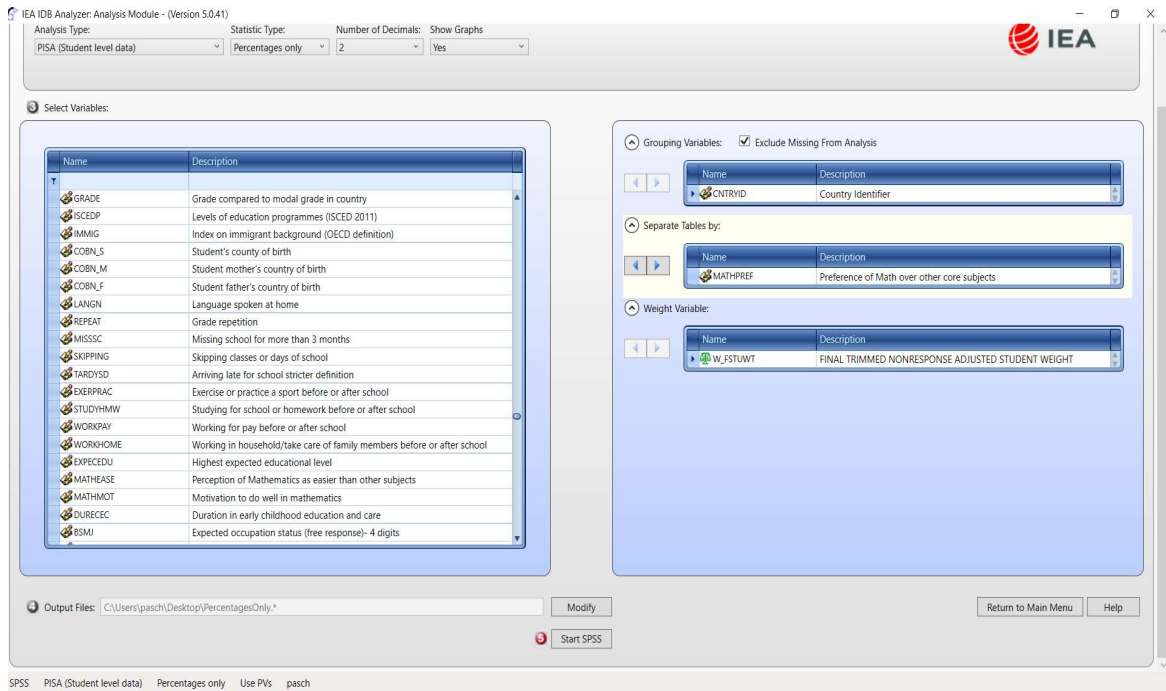


Εικόνα 133. Τοποθέτηση μεταβλητής MATHPREF στο πεδίο «Separate Tables by» (SPSS)

Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια (Εικόνα 134). Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «PercentagesOnly». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 135).

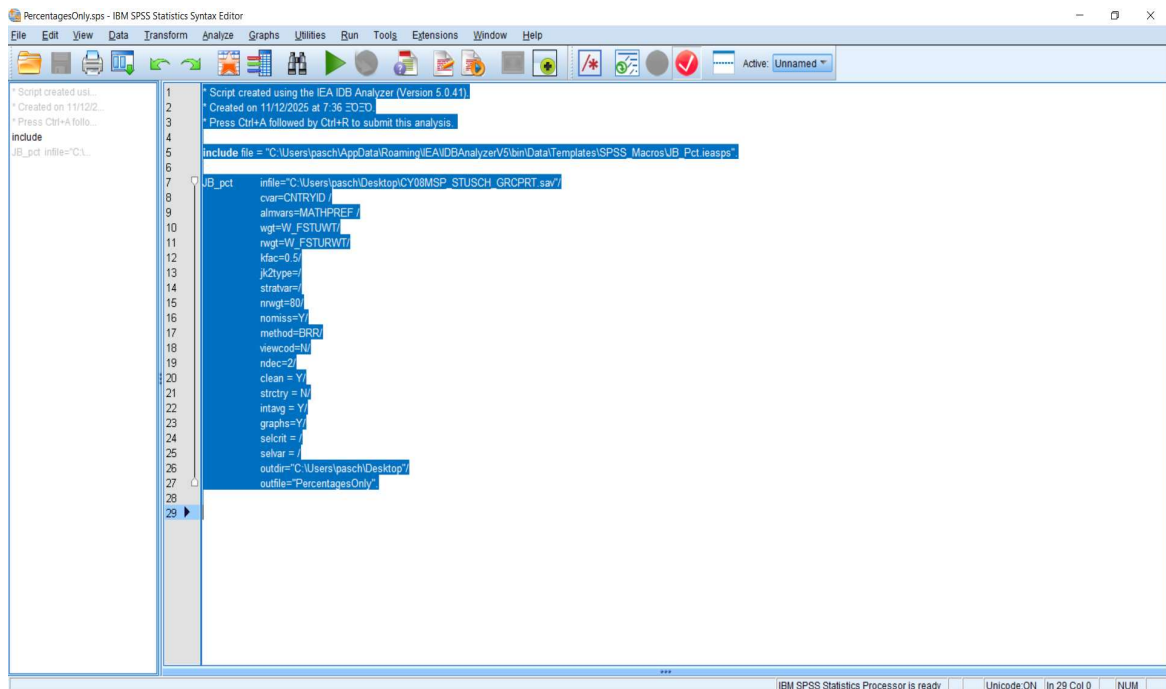


Εικόνα 134. Ορισμός ονόματος εξόδου για ανάλυση Percentages Only (SPSS)



Εικόνα 135. Εκκίνηση παραγωγής script μέσω «Start R» – Percentages Only

Αφού μεταβαίνετε στο SPSS, συγκεκριμένα στο αρχείο PercentagesOnly.sps που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών. (Εικόνα 136).



Εικόνα 136. Εκτέλεση κώδικα PercentagesOnly.sps στο SPSS

Στο αρχείο PercentagesOnly.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 137 και 138 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

*** Processing 80 replicates...
Computing variance terms...

>> Percentages by (COUNTRY MATHPREF)

Report

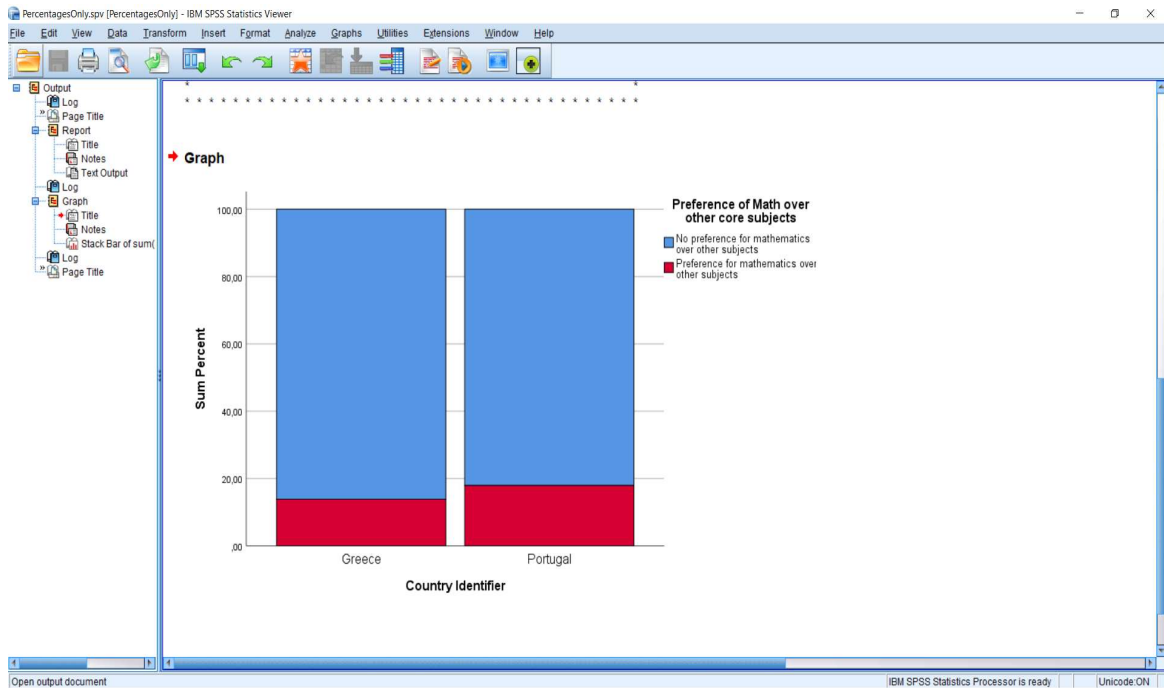
Percentages by (COUNTRY MATHPREF)

Country Identifier	Preference of Math over other core subjects	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent	Percent (s.e.)
Greece	No preference for mathematics over other subjects	5053	77185	1666,58	86,13	,47
	Preference for mathematics over other subjects	834	12428	455,43	13,87	,47
Portugal	No preference for mathematics over other subjects	5104	72654	1808,64	82,07	,57
	Preference for mathematics over other subjects	1151	15869	579,47	17,93	,57
Table Average	No preference for mathematics over other subjects	.	.	.	84,10	,37
	Preference for mathematics over other subjects	.	.	.	15,90	,37

Εικόνα 137. Πίνακας ποσοστών προτίμησης για τα Μαθηματικά ανά χώρα (SPSS)

Οι στήλες της Εικόνα 137 παρουσιάζουν την κατανομή των μαθητών ως προς την προτίμησή τους για τα Μαθηματικά σε σύγκριση με άλλα βασικά μαθήματα, ανά χώρα, αποτυπώνοντας τα αντίστοιχα ποσοστά στον μαθητικό πληθυσμό. Συγκεκριμένα:

- Country Identifier: Η χώρα (π.χ. Ελλάδα, Πορτογαλία).
- Preference of Math over other core subjects: Αν ο μαθητής προτιμά τα Μαθηματικά ή όχι σε σχέση με άλλα βασικά μαθήματα.
- N of Cases: Πόσοι μαθητές του δείγματος ανήκουν σε αυτή την κατηγορία.
- Sum of W_FSTUWT: Το σταθμισμένο πλήθος μαθητών που αντιπροσωπεύουν τον συνολικό πληθυσμό.
- Sum of W_FSTUWT (s.e.): Το τυπικό σφάλμα.
- Percent: Το ποσοστό (%) των μαθητών που ανήκουν στην αντίστοιχη κατηγορία.
- Percent (s.e.): Το τυπικό σφάλμα του ποσοστού.



Εικόνα 138. Γράφημα κατανομής MATHPREF – Percentages Only (SPSS)

Το γράφημα που απεικονίζεται στην Εικόνα 138 παριστάνει την κατανομή των μαθητών ως προς την προτίμησή τους για τα Μαθηματικά για τις δύο χώρες.

Ερμηνεία αποτελεσμάτων

Παρουσιάζονται τα αποτελέσματα της ανάλυσης ποσοστών (percentages) για τη μεταβλητή MATHPREF, η οποία αποτυπώνει εάν οι μαθητές δηλώνουν προτίμηση για τα Μαθηματικά σε σχέση με άλλα μαθήματα. Τα ποσοστά υπολογίστηκαν με χρήση του βάρους W_FSTUWT , ώστε οι εκτιμήσεις να είναι αντιπροσωπευτικές του μαθητικού πληθυσμού. Για την Ελλάδα, η πλειονότητα των μαθητών δεν δήλωσε προτίμηση για τα Μαθηματικά έναντι άλλων βασικών μαθημάτων. Συγκεκριμένα, το 86.13% (SE = 0.47) ανήκε στην κατηγορία No preference for mathematics over other subjects, ενώ το 13.87% (SE = 0.47) δήλωσε Preference for mathematics over other subjects. Αντίστοιχα, στην Πορτογαλία παρατηρήθηκε παρόμοιο μοτίβο, με τη μη προτίμηση να υπερισχύει. Ωστόσο, το ποσοστό των μαθητών που δήλωσαν προτίμηση για τα Μαθηματικά ήταν υψηλότερο σε σχέση με την Ελλάδα: 82.07% (SE = 0.57) ανέφερε No preference for mathematics over other subjects και 17.93% (SE = 0.57) ανέφερε Preference for mathematics over other subjects. Συνολικά, και στις δύο χώρες υπερισχύει η μη προτίμηση για τα Μαθηματικά, αν και η Πορτογαλία εμφάνισε μεγαλύτερο ποσοστό μαθητών που προτιμούν τα Μαθηματικά σε σύγκριση με την Ελλάδα (διαφορά = 4.06 ποσοστιαίες μονάδες).

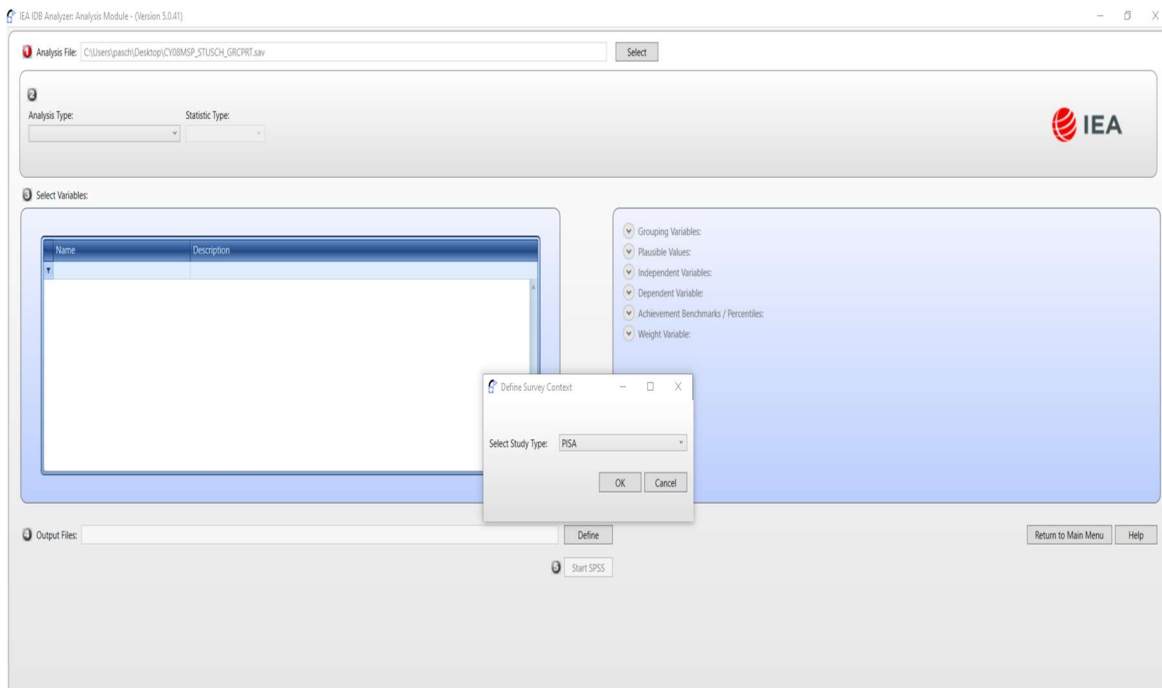
Υπολογισμός Ποσοστημορίων (Percentiles)

Ερευνητική Υπόθεση: Διερευνώνται διαφορές στα ποσοστημόρια του άγχους για τα Μαθηματικά μεταξύ των δύο χωρών, ώστε να εξεταστεί αν η κατανομή του άγχους διαφοροποιείται σε διαφορετικά επίπεδα της κατανομής.

Επιλογή μεταβλητής: ANXMAT

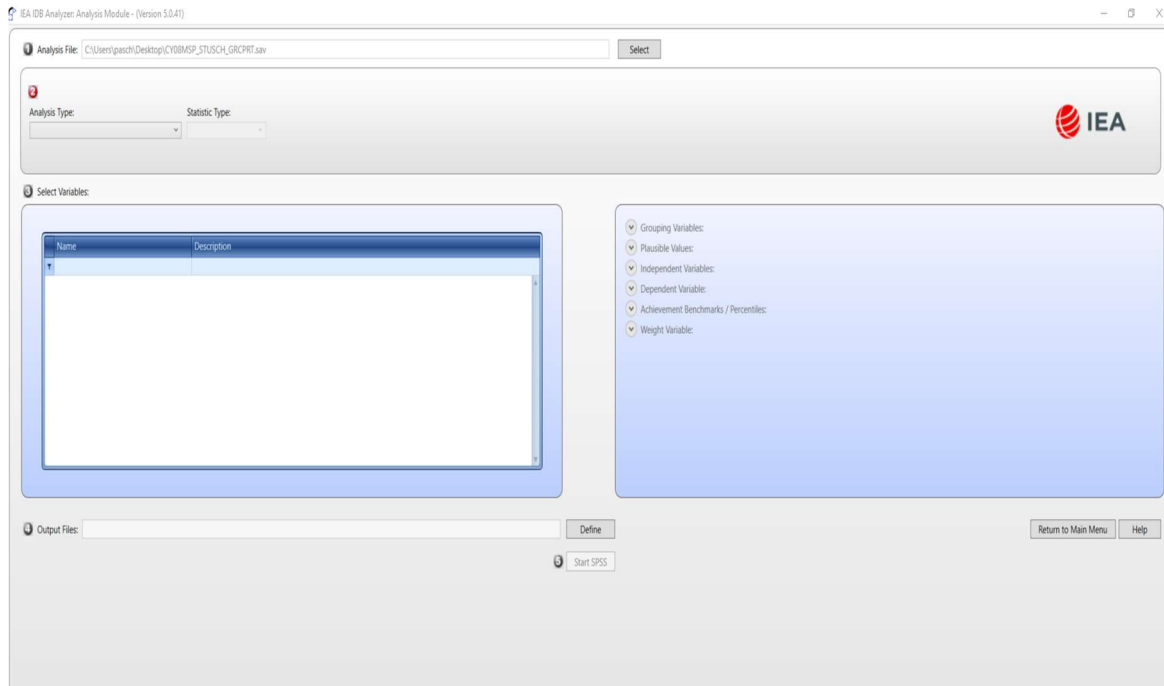
Χαρακτηριστικά μεταβλητών: ANXMAT συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί υπολογισμός των ποσοστημορίων για τη μεταβλητή ANXMAT, η οποία αποτυπώνει το άγχος των μαθητών απέναντι στα Μαθηματικά, προκειμένου να ελεγχθούν οι διαφορές στην κατανομή του άγχους για τα Μαθηματικά για τις δύο χώρες που επιλέχθηκαν. Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCPR1.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 139).



Εικόνα 139. Επιλογή αρχείου δεδομένων για ανάλυση Percentiles – «Analysis File» (SPSS)

Έπειτα, στο πεδίο «Analysis Type», επιλέξτε τον τύπο της ανάλυσης που θέλετε να πραγματοποιήσετε, για παράδειγμα «PISA (Student level)» (Εικόνα 140).



Εικόνα 140. Επιλογή «PISA (Student level)» – Ανάλυση Percentiles (SPSS)

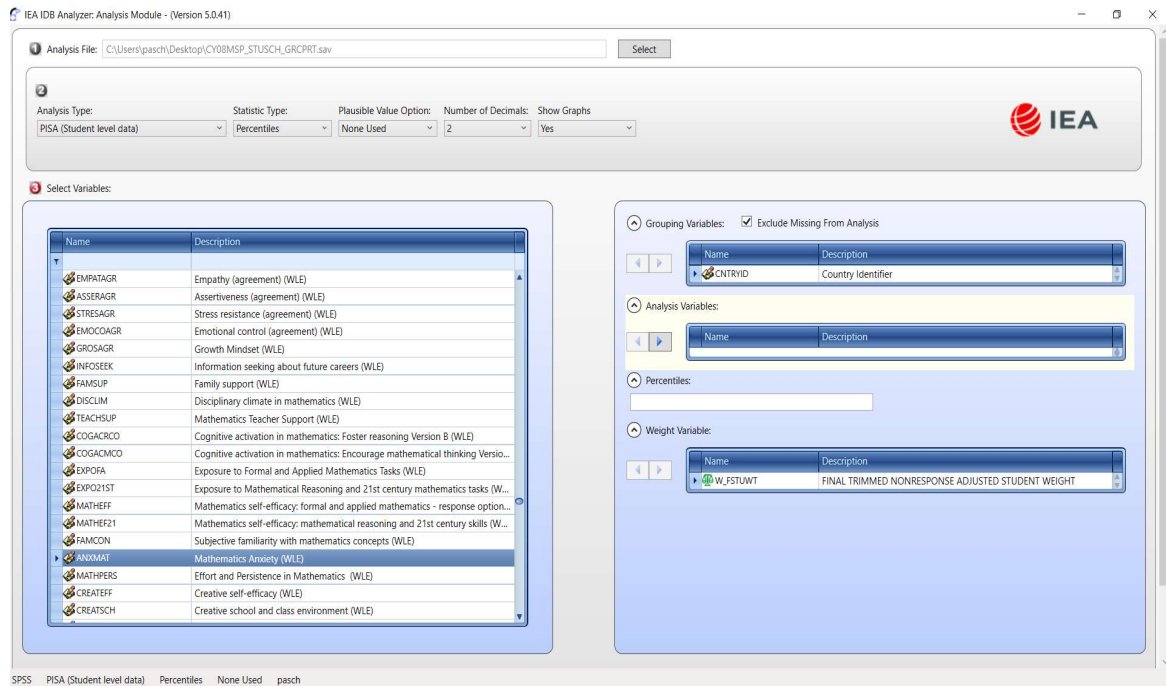
Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε. Παρέχονται οι παρακάτω επιλογές:

- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Percentiles». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «None Used». Επιπλέον στο πεδίο «Show Graphs» δίνεται από προεπιλογή η δυνατότητα να εμφανιστούν διαγράμματα. Τέλος στο πεδίο «Number of Decimals» επιλέγετε τον αριθμό των δεκαδικών ψηφίων (Εικόνα 141).

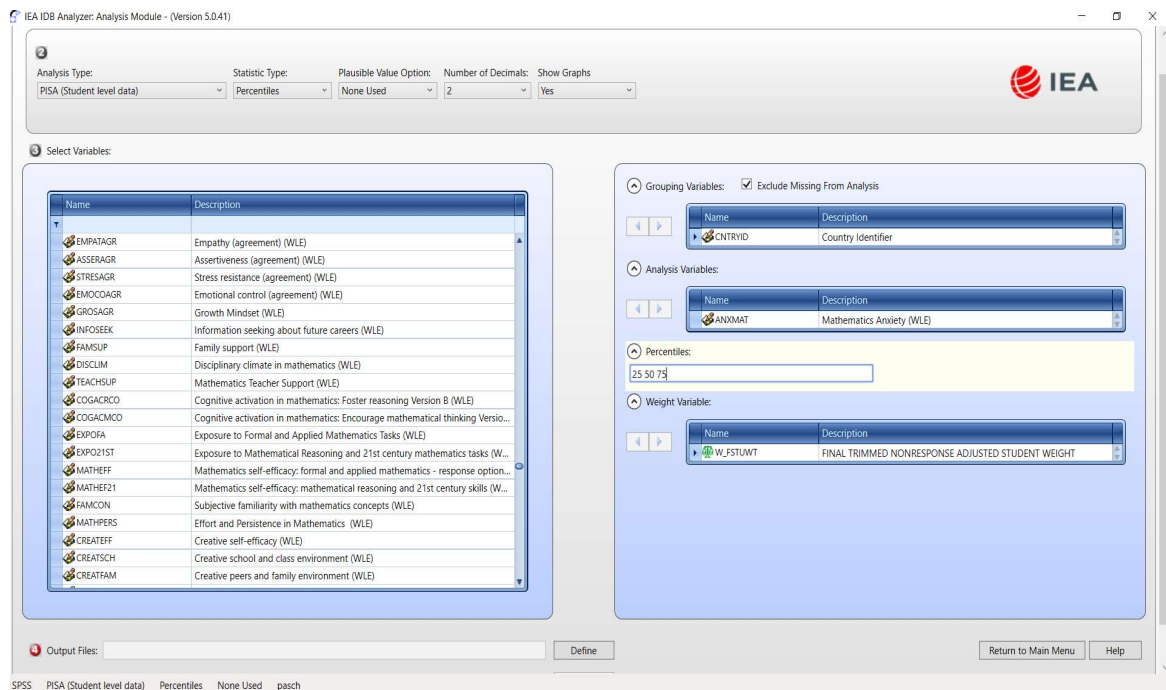
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 141 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλιπείς τιμές στις μεταβλητές ομαδοποίησης. Προκειμένου να υλοποιηθεί ο υπολογισμός των ποσοστιμορίων για την μεταβλητή που επιλέγετε, θα πρέπει να εντοπίσετε τις μεταβλητές που είναι απαραίτητες στο πλαίσιο «Select Variables». Επομένως βρίσκετε αρχικά την μεταβλητή την μεταβλητή «ANXMAT» από

το πλαίσιο «Select Variables», την επιλέγετε και μετά επιλέγοντας το μπλε βελάκι στο πλαίσιο «Analysis Variables» και την τοποθετείτε (Εικόνα 141).



Εικόνα 141. Τοποθέτηση μεταβλητής ANXMAT (SPSS)

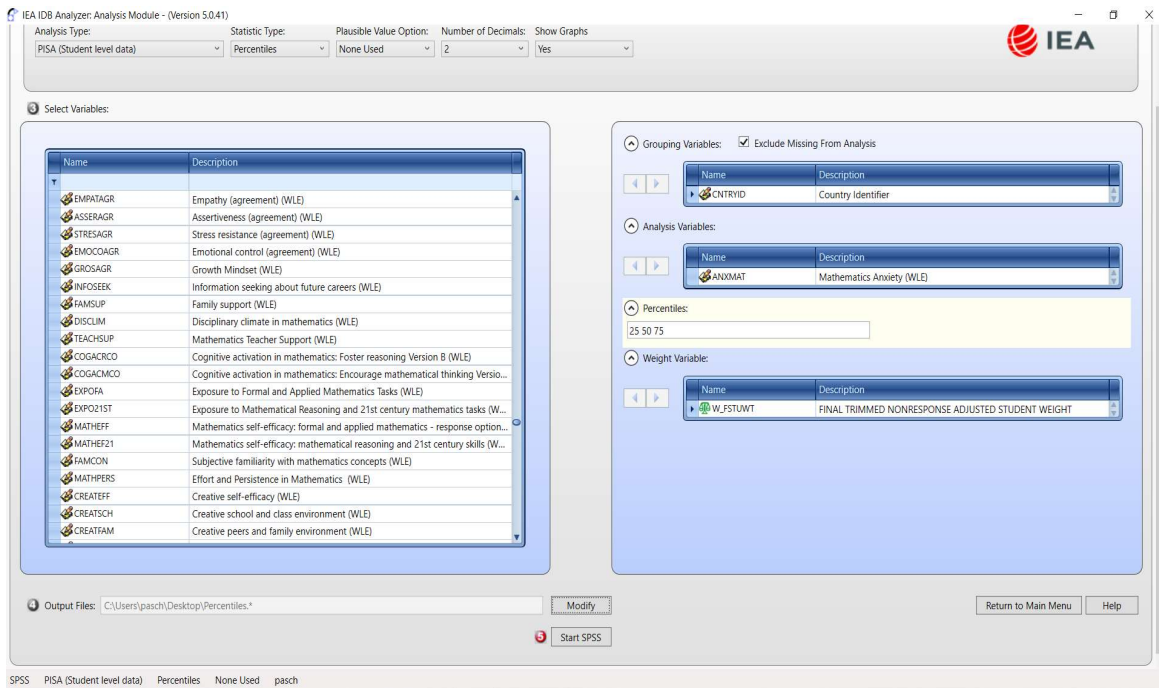
Στην συνέχεια στο πλαίσιο «Percentiles» εισάγουμε τα ποσοστημόρια που θέλουμε να μελετήσουμε και συγκεκριμένα το 25ο, το 50ο (διάμεσος) και το 75ο ποσοστημόριο. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 142).



Εικόνα 142. Εισαγωγή τιμών p25, p50, p75 (SPSS)

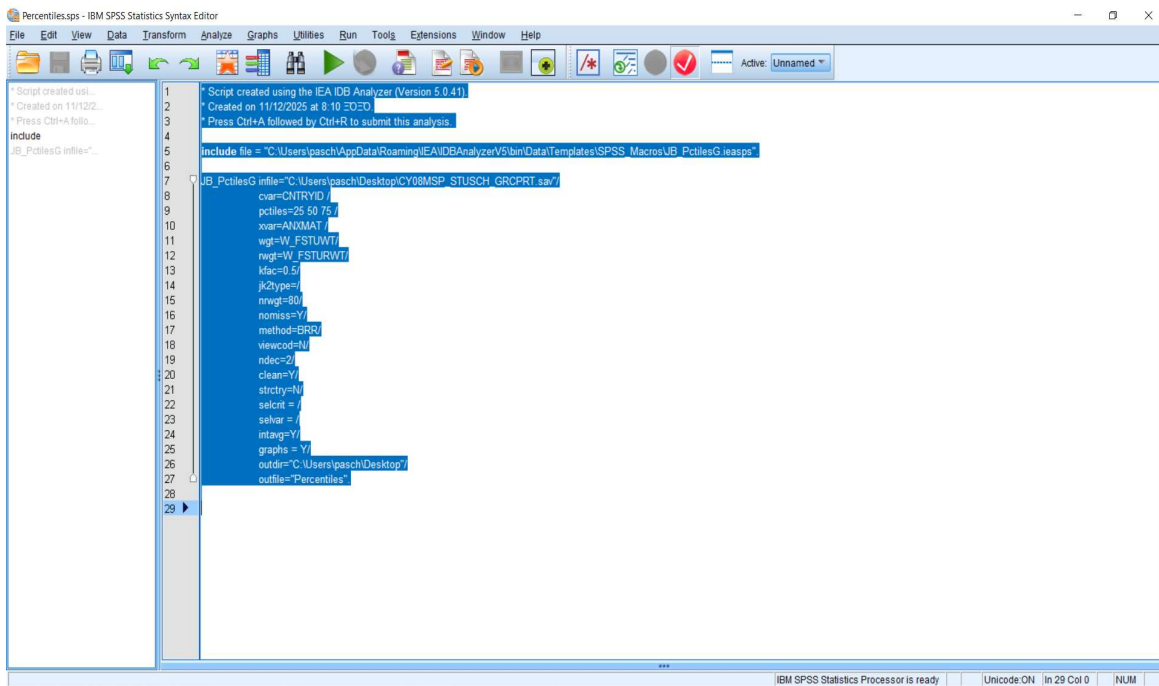
Επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι

σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «Percentiles». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο SPSS (Εικόνα 143).



Εικόνα 143. Εκκίνηση παραγωγής syntax μέσω «Start SPSS» – Percentiles

Αφού μεταβείτε στο περιβάλλον του SPSS, συγκεκριμένα στο αρχείο Percentiles.sps που εμφανίζεται, με Ctrl+A επιλέξετε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 144).



Εικόνα 144. Εκτέλεση κώδικα Percentiles.sps στο SPSS

Στο αρχείο Percentiles.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 145 και 146 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

*** Processing 80 replicates...
Computing variance terms...
Percentiles for ANXMAT by COUNTRYID

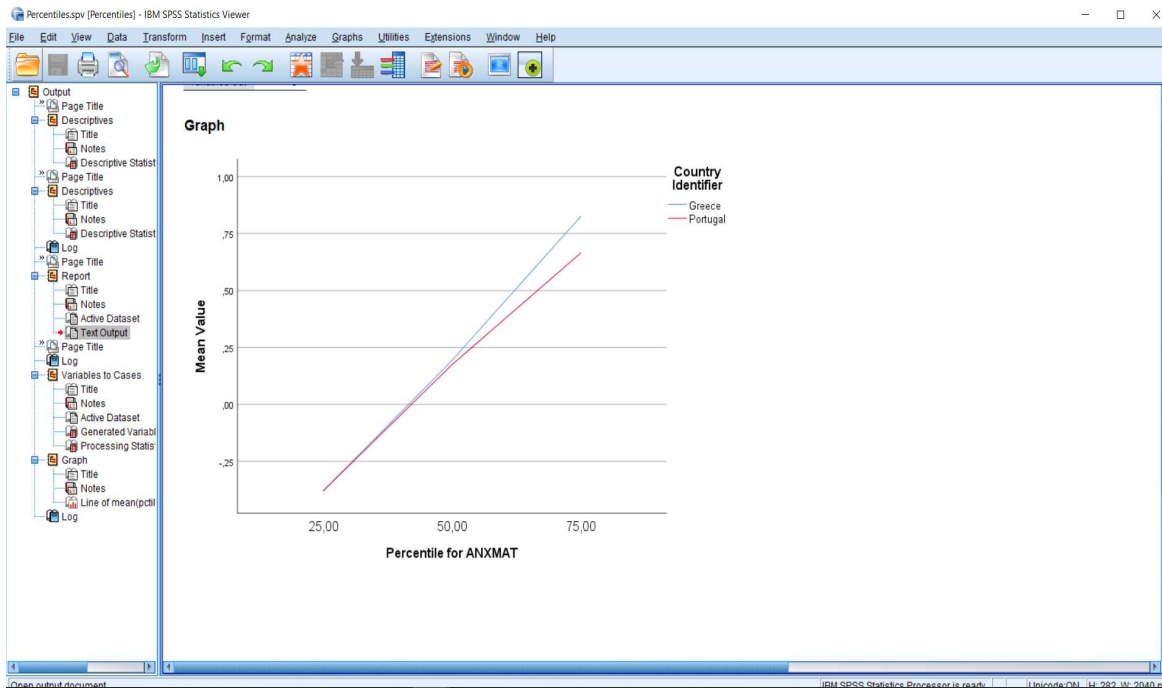
Report
C:\Users\pasach\Desktop\tmp3

Country Identifier	N of Cases	Sum of W_FSTUWT	p25	p25_se	p50	p50_se	p75	p75_se
Greece	5770	87817	-,38	,01	,19	,00	,83	,04
Portugal	5913	83847	-,38	,01	,17	,00	,66	,00
Table Average	.	.	-,38	,00	,18	,00	,74	,02

Εικόνα 145. Πίνακας ποσοστημορίων p_{25} , p_{50} , p_{75} της ANXMAT ανά χώρα (SPSS)

Παρακάτω αναλύεται το περιεχόμενο των στηλών της Εικόνα 145, και πιο συγκεκριμένα:

- Country Identifier: Η χώρα στην οποία αναφέρονται τα αποτελέσματα (π.χ. Ελλάδα, Πορτογαλία).
- N of Cases: Ο αριθμός των μαθητών που συμπεριλήφθηκαν στην ανάλυση για κάθε χώρα.
- Sum of W_FSTUWT: Το άθροισμα των μαθητικών σταθμίσεων (weights), το οποίο αντιπροσωπεύει το εκτιμώμενο πληθυσμιακό μέγεθος για τη συγκεκριμένη χώρα.
- p25: Το 25ο ποσοστιαίο σημείο (1ο τεταρτημόριο) της μεταβλητής ANXMAT.
- p25_se: Το τυπικό σφάλμα της εκτίμησης του 25ου ποσοστημορίου.
- p50: Η διάμεσος (50ο ποσοστιαίο σημείο) της ANXMAT.
- p50_se: Το τυπικό σφάλμα της εκτίμησης της διαμέσου.
- p75: Το 75ο ποσοστιαίο σημείο (3ο τεταρτημόριο) της ANXMAT.
- p75_se: Το τυπικό σφάλμα της εκτίμησης του 75ου ποσοστημορίου.



Εικόνα 146. Γράφημα κατανομής ποσοστημοριών ANXMAT – Ελλάδα vs Πορτογαλία (SPSS)

Επιπλέον, το γράφημα της Εικόνας 146 απεικονίζει την κατανομή του άγχους για τα Μαθηματικά (ANXMAT) στα ποσοστημόρια για την Ελλάδα και την Πορτογαλία,

Ερμηνεία αποτελεσμάτων

Με βάση τα αποτελέσματα των ποσοστημοριών για τη μεταβλητή ANXMAT (άγχος για τα Μαθηματικά), παρατηρούνται τα εξής. Για την Ελλάδα, το 25% των μαθητών έχει τιμή ANXMAT κάτω από -0.38 ($p_{25} = -0.38$, $SE = 0.01$), το 50% των μαθητών έχει τιμή έως 0.19 ($p_{50} = 0.19$, $SE = 0.00$) και το 75% των μαθητών έχει τιμή κάτω από 0.83 ($p_{75} = 0.83$, $SE = 0.04$). Όσον αφορά την Πορτογαλία, το 25% των μαθητών έχει επίσης τιμή ANXMAT κάτω από -0.38 ($p_{25} = -0.38$, $SE = 0.01$), το 50% έχει τιμή έως 0.17 ($p_{50} = 0.17$, $SE = 0.00$) και το 75% έχει τιμή κάτω από 0.66 ($p_{75} = 0.66$, $SE = 0.00$). Συνεπώς, συγκεντρωτικά φαίνεται ότι το κάτω μέρος της κατανομής (p_{25}) είναι ίδιο στις δύο χώρες, ενώ από τη διάμεσο και πάνω η Ελλάδα εμφανίζει υψηλότερες τιμές άγχους. Η διαφορά είναι πιο έντονη στο άνω 25% της κατανομής, όπου το p_{75} στην Ελλάδα (0.83) είναι υψηλότερο από το p_{75} στην Πορτογαλία (0.66) κατά περίπου 0.17 μονάδες. Αυτό επιβεβαιώνεται και από το γράφημα, όπου η γραμμή της Ελλάδας βρίσκεται πάνω από της Πορτογαλίας στα υψηλότερα ποσοστημόρια, δείχνοντας ότι οι μαθητές με μεγαλύτερο άγχος στην Ελλάδα τείνουν να εμφανίζουν ακόμη υψηλότερα επίπεδα άγχους σε σχέση με τους αντίστοιχους μαθητές στην Πορτογαλία.

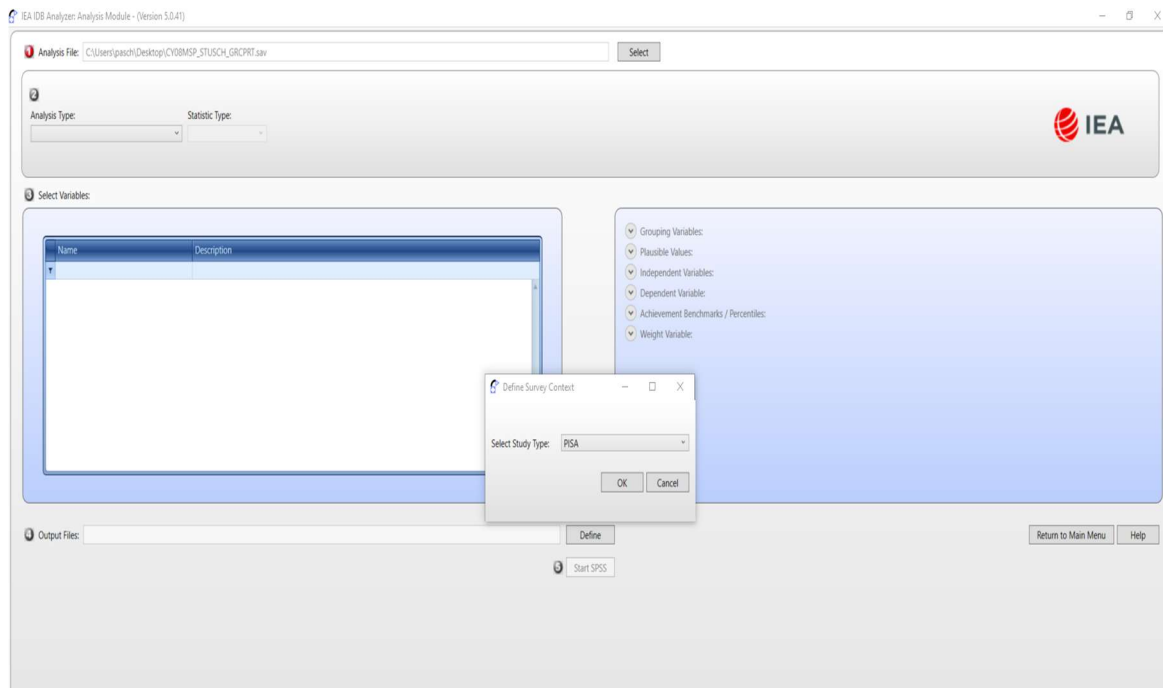
Μελέτη επιδόσεων με βάση τα ορόσημα (Benchmarks)

Ερευνητική Υπόθεση: Διερευνώνται διαφορές στα ποσοστά των μαθητών που επιτυγχάνουν τα καθορισμένα ορόσημα μαθηματικής επίδοσης μεταξύ των χωρών που επιλέχθηκαν

Επιλογή μεταβλητής: PVMATH01-10

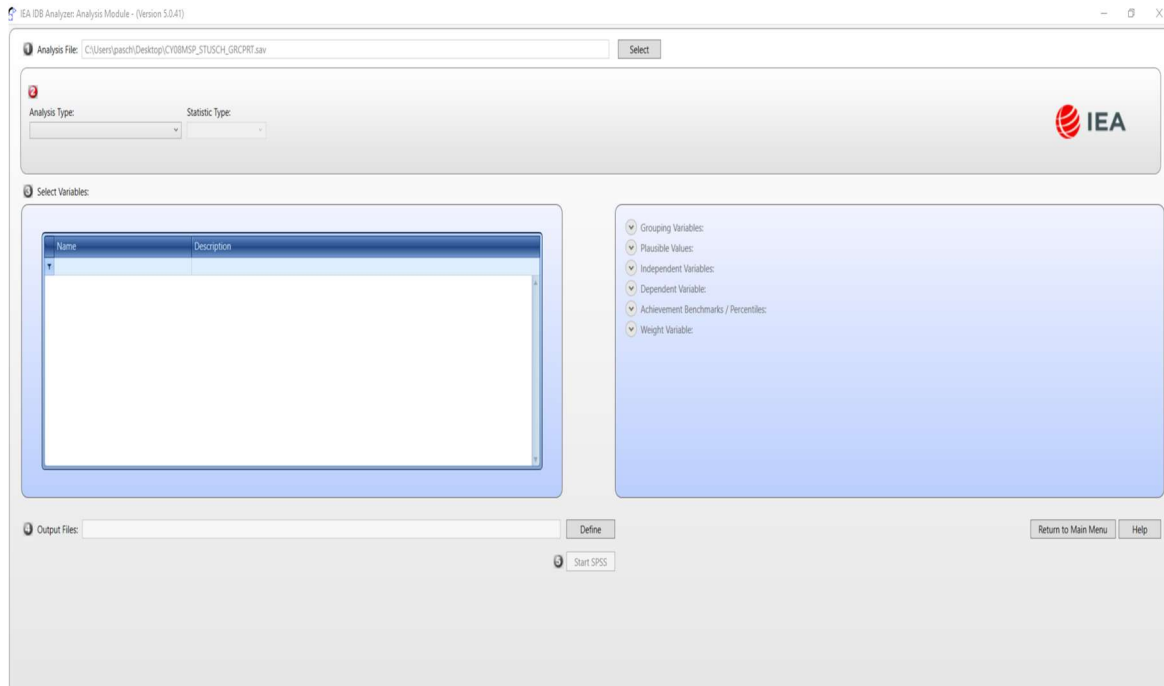
Χαρακτηριστικά μεταβλητών: PVMATH01-10 συνεχής μεταβλητή

Μας ενδιαφέρει να πραγματοποιηθεί ανάλυση των επιδόσεων των μαθητών με βάση τα ορόσημα (Benchmarks) που πετυχαίνουν στα μαθηματικά. Η μεταβλητή όπου καταγράφεται η επίδοση των μαθητών, με την χρήση των Plausible Values, είναι η «PVMATH01-10». Στο πεδίο «Analysis File», με την επιλογή του κουμπιού «Select», επιλέγεται το αρχείο δεδομένων από το σημείο του υπολογιστή όπου το έχετε αποθηκεύσει, στη συγκεκριμένη περίπτωση το αρχείο με δεδομένα της έρευνας «PISA» με όνομα «CY08MSP_STUSCH_GRCprt.sav» και στη συνέχεια το είδος του αρχείου, συγκεκριμένα «PISA» (Εικόνα 147).



Εικόνα 147. Επιλογή αρχείου δεδομένων για ανάλυση Benchmarks (SPSS)

Έπειτα στο πεδίο «Analysis Type» επιλέξτε τον τύπο της ανάλυσης που θέλετε να πραγματοποιήσετε, για παράδειγμα «Pisa (Student level)» (Εικόνα 148).



Εικόνα 148. Επιλογή «PISA (Student level)» – Ανάλυση Benchmarks (SPSS)

Στην συνέχεια επιλέξτε στο πεδίο «Statistic Type» το είδος στατιστικής ανάλυσης που θέλετε να πραγματοποιήσετε, παρέχονται οι παρακάτω επιλογές:

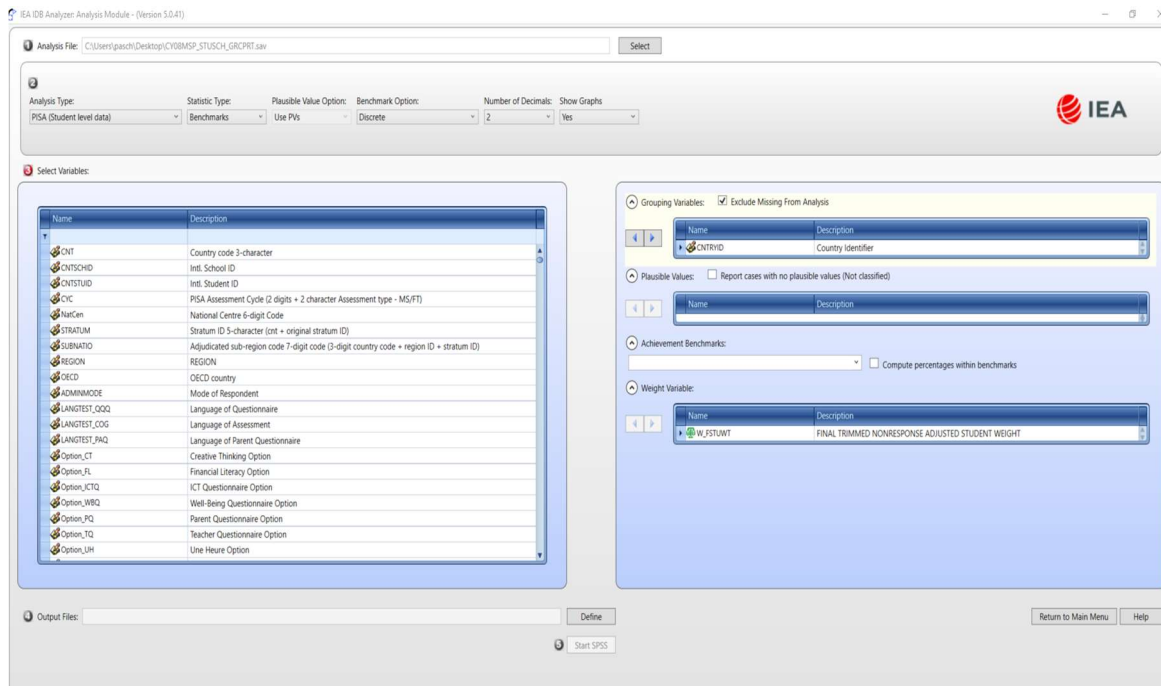
- Benchmarks: Υπολογίζεται το ποσοστό του πληθυσμού που βρίσκεται πάνω ή μέσα σε προκαθορισμένα όρια επίδοσης.
- Correlation (Pearson): Μέτρηση γραμμικής συσχέτισης μεταξύ δύο συνεχών μεταβλητών.
- Correlation (Spearman): Μέτρηση συσχέτισης μεταξύ δύο μεταβλητών, όταν τα δεδομένα δεν είναι κανονικά κατανομημένα.
- Linear regression: Ανάλυση γραμμικής παλινδρόμησης που εξηγεί την σχέση μεταξύ μιας εξαρτημένης, συνεχούς μεταβλητής και μίας ή περισσότερων ανεξάρτητων.
- Percentages and Means: Υπολογισμός ποσοστών, μέσων όρων, design effects και τυπικές αποκλίσεις για επιλεγμένες μεταβλητές, ανά υποομάδες που ορίζονται από τον χρήστη. Περιλαμβάνει το ποσοστό των ελλιπών απαντήσεων. Επιπλέον, εκτελεί t-tests για τη διαφορά μέσων όρων και ποσοστών μεταξύ ομάδων.
- Percentages only: Υπολογισμός ποσοστών που ορίζονται από τον χρήστη.
- Percentiles: Υπολογισμός των ποσοστιαίων σημείων που χωρίζουν την κατανομή των σκορ, ανά υποομάδες που ορίζονται από μεταβλητές ομαδοποίησης («Grouping Variables»).

Επιλέγετε «Benchmarks». Τα υπόλοιπα πεδία συμπληρώνονται αυτόματα και μπορείτε να τα αφήσετε ως έχουν. Σημειώστε ότι στο πεδίο «Plausible Value Option» επιλέγεται «Use PVs».

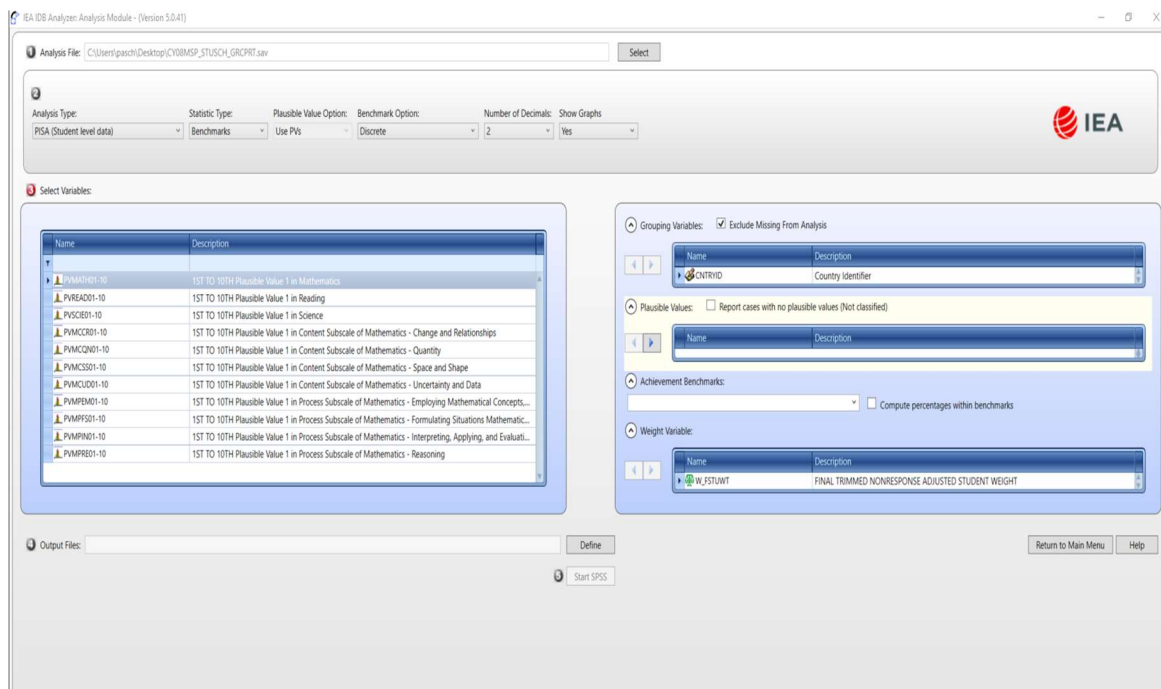
Επιπλέον στο πεδίο «Benchmark Option» υπάρχουν διαθέσιμες οι επιλογές:

- Cumulative: Υπολογίζεται το ποσοστό των ατόμων που βρίσκονται «στο συγκεκριμένο όριο ή πάνω από αυτό».
- Discrete: Υπολογίζεται το ποσοστό των ατόμων μέσα σε κάθε κατηγορία που ορίζεται από τα όρια επίδοσης (τα cut-off points λειτουργούν ως κατώτερο όριο κάθε ομάδας).
- Discrete with Analysis Variable(s): Εκτός από τα ποσοστά ανά κατηγορία επίδοσης, υπολογίζεται και τον μέσο όρο μιας (ή περισσότερων) μεταβλητών ανάλυσης μέσα σε κάθε κατηγορία (π.χ. μέσος όρος «κίνησης για διάβασμα» ανά επίπεδο επίδοσης).

Τέλος στο πεδίο «Number of Decimals» επιλέγεται τον αριθμό των δεκαδικών ψηφίων και στο πεδίο «Show Graphs» αν επιθυμείτε να εμφανιστούν γραφήματα στο τελικό αποτέλεσμα (Εικόνα 149).



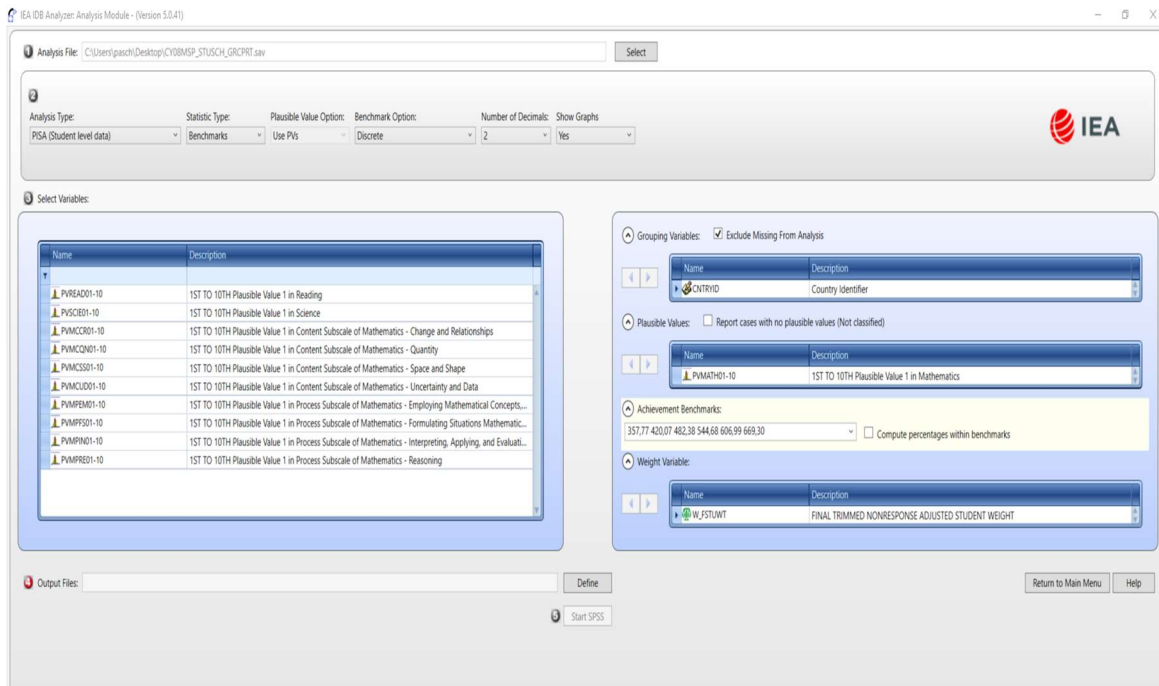
Εικόνα 149. Επιλογή «Benchmarks» και ορισμός «Plausible Value Option: Use PVs» (SPSS)



Εικόνα 150. Τοποθέτηση PVMATH01-10 και εισαγωγή ορίων επίδοσης (Benchmarks) (SPSS)

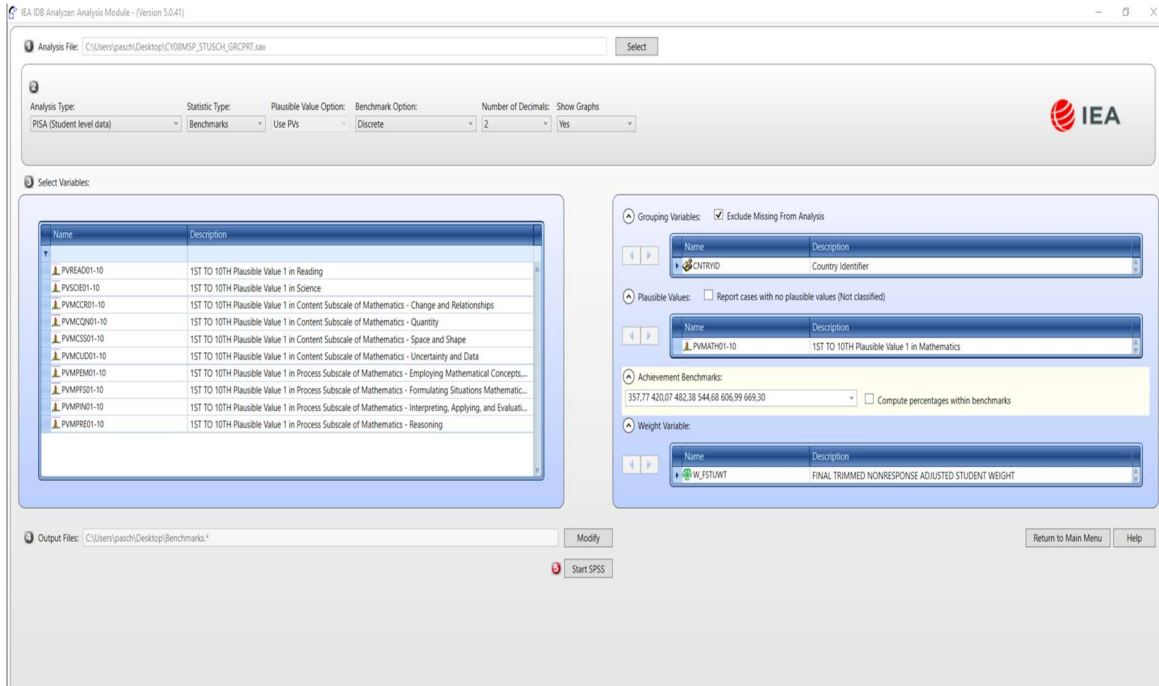
Όπως φαίνεται στο δεξιό τμήμα της Εικόνας 150 στο πλαίσιο «Grouping Variables», όπου είναι η λίστα των μεταβλητών που θα χρησιμοποιηθούν για να οριστούν οι υποομάδες στην ανάλυση, υπάρχει από προεπιλογή η μεταβλητή «CNTTRYID». Επίσης αν η επιλογή «Exclude Missing from Analysis» είναι ενεργοποιημένη, τότε στην ανάλυση θα χρησιμοποιηθούν μόνο οι περιπτώσεις που έχουν μη ελλειπείς τιμές στις μεταβλητές ομαδοποίησης. Στη συνέχεια, εφόσον ο σκοπός είναι να αναλυθεί η επίδοση των μαθητών στα μαθηματικά, θα πρέπει να επιλέξετε τη μεταβλητή “PVMATH01-10” και να την

τοποθετήσετε στο πλαίσιο “Plausible Values” (Εικόνα 150). Έπειτα, στο πλαίσιο “Achievement Benchmarks” να εισάγετε τις τιμές που θεωρούμε ως ορόσημα (benchmarks). Εδώ εισάγονται οι τιμές 357,77 / 420,07 / 482,54 / 544,60 / 606,91 / 669,33 που αφορούν τα μαθηματικά. Ωστόσο, επειδή το IEA IDB Analyzer δεν έχει ενημερώσει την εφαρμογή με τα νέα cut-off points, θα πρέπει να προστεθούν χειροκίνητα, όπως ορίζονται στη βιβλιογραφία του εκάστοτε κύκλου ανάλυσης. Τα νέα cut-off points είναι 233,17 / 295,47 / 357,77 / 420,07 / 482,54 / 544,60 / 606,91 / 669,33. Τέλος στο πλαίσιο «Weight Variable» υπάρχει από προεπιλογή η μεταβλητή «W_FSTUWT». Αξίζει να αναφερθεί ότι το IEA IDB Analyzer επιλέγει αυτόματα τις κατάλληλες μεταβλητές βάρους για την ανάλυση (Εικόνα 151).



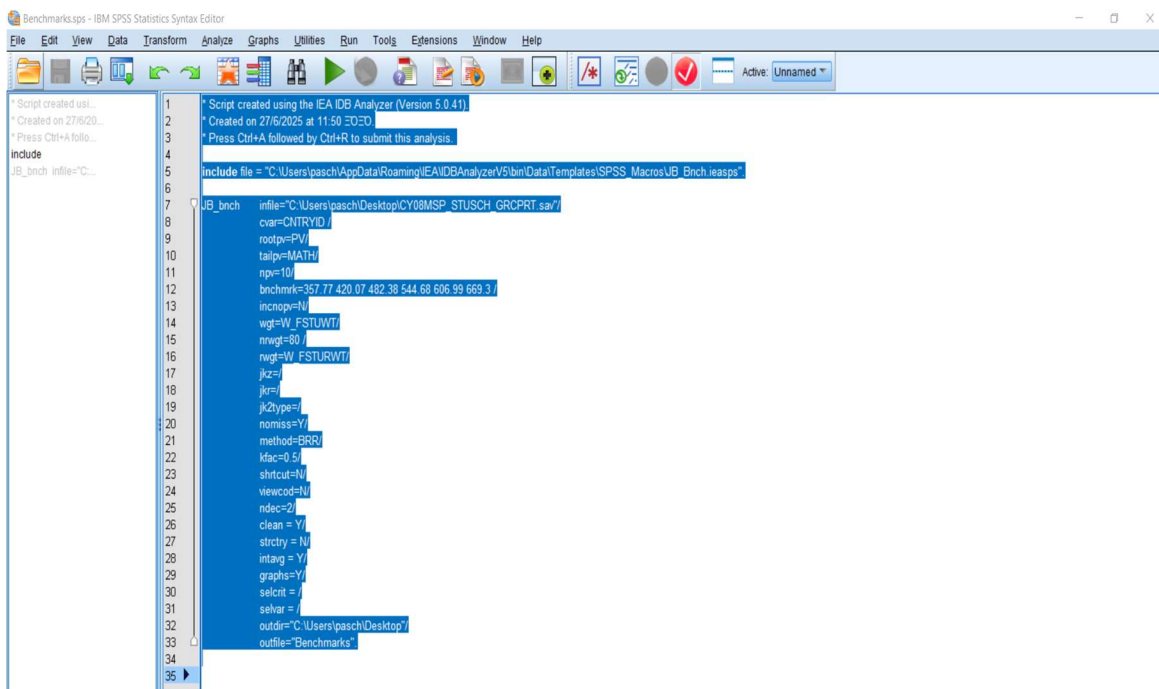
Εικόνα 151. Ορισμός ονόματος εξόδου για ανάλυση Benchmarks (SPSS)

Το επόμενο βήμα είναι να επιλέξετε το κουμπί «Define» δίπλα στο πλαίσιο «Output Files» και να ορίσετε το όνομα των output files που θα εμφανιστούν στη συνέχεια. Η επιλογή του ονόματος καλό θα ήταν να είναι σχετικό με την ανάλυση που πραγματοποιείται για λόγους ευκολίας, εδώ ενδεικτικά «Benchmarks». Τελευταίο βήμα είναι να επιλέξετε το κουμπί «Start SPSS», για να δημιουργηθεί το syntax αρχείο στο spss (Εικόνα 152).



Εικόνα 152. Εκκίνηση παραγωγής syntax μέσω «Start SPSS» – Benchmarks

Αφού μεταβείτε στο SPSS, συγκεκριμένα στο αρχείο Benchmarks.sps που εμφανίζεται, με Ctrl+A επιλέξτε το περιεχόμενο και πατήστε το πράσινο κουμπί, που εμφανίζεται στη δεύτερη γραμμή εντολών (Εικόνα 153).



Εικόνα 153. Εκτέλεση κώδικα Benchmarks.sps στο SPSS

Στο αρχείο Benchmarks.sps, μεταξύ άλλων πληροφοριών θα παρατηρήσετε τις Εικόνες 154 και 155 οι οποίες απεικονίζουν τα αποτελέσματα της παραπάνω ανάλυσης.

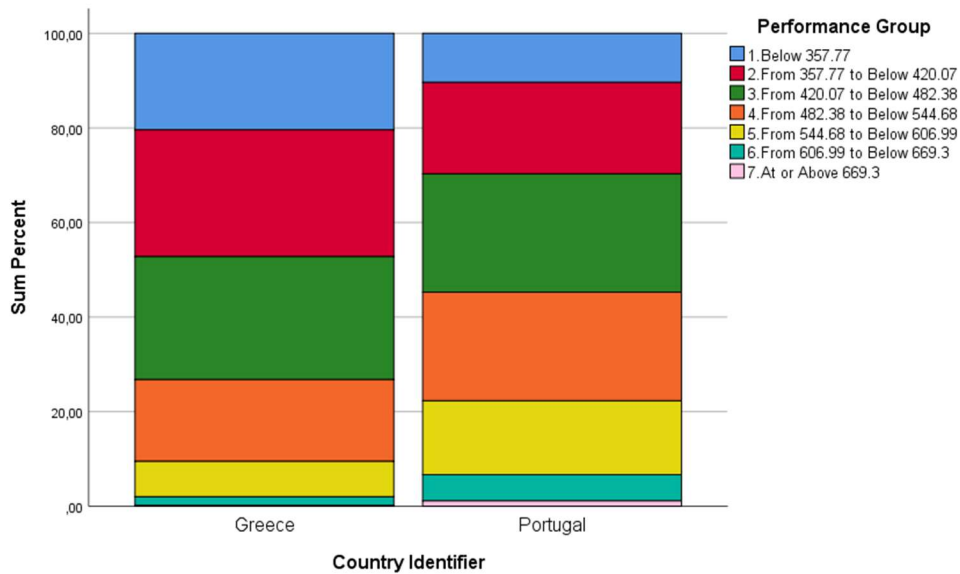
Percents by Performance Groups of PVMATH						
Country Identifier	Performance Group	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent	Percent (s.e.)
Greece	1.Below 357.77	1219	20001	1086,99	20,39	1,02
	2.From 357.77 to Below 420.07	1693	26303	929,29	26,82	,77
	3.From 420.07 to Below 482.38	1702	25511	877,76	26,01	,77
	4.From 482.38 to Below 544.68	1154	16971	731,03	17,30	,70
	5.From 544.68 to Below 606.99	502	7341	510,04	7,48	,52
	6.From 606.99 to Below 669.3	124	1814	271,91	1,85	,28
	7.At or Above 669.3	10	146	68,74	,15	,07
Portugal	1.Below 357.77	627	10039	914,29	10,39	,87
	2.From 357.77 to Below 420.07	1245	18670	868,56	19,33	,73
	3.From 420.07 to Below 482.38	1714	24167	910,38	25,02	,76
	4.From 482.38 to Below 544.68	1626	22189	842,86	22,97	,78
	5.From 544.68 to Below 606.99	1101	15094	646,33	15,62	,67
	6.From 606.99 to Below 669.3	384	5349	351,00	5,54	,37
	7.At or Above 669.3	76	1099	188,94	1,14	,19
Table Average	1.Below 357.77	.	.	.	15,39	,67
	2.From 357.77 to Below 420.07	.	.	.	23,07	,53
	3.From 420.07 to Below 482.38	.	.	.	25,51	,54
	4.From 482.38 to Below 544.68	.	.	.	20,14	,52
	5.From 544.68 to Below 606.99	.	.	.	11,55	,43
	6.From 606.99 to Below 669.3	.	.	.	3,69	,23
	7.At or Above 669.3	.	.	.	,64	,10

Εικόνα 154. Πίνακας ποσοστών μαθητών ανά επίπεδο επίδοσης (SPSS)

Αναλυτικά οι στήλες της Εικόνα 154 παρουσιάζουν:

- **Country Identifier:** Η χώρα στην οποία αντιστοιχούν τα δεδομένα (π.χ. Ελλάδα, Πορτογαλία).
- **Performance Group:** Οι κατηγορίες επίδοσης που ορίζονται από τα Benchmarks, όπου κάθε γραμμή δείχνει ένα εύρος τιμών.
- **N of Cases:** Αριθμός ατόμων στην κάθε κατηγορία.
- **Sum of W_FSTUWT:** Το σταθμισμένο άθροισμα των μαθητών, δηλαδή πόσους μαθητές αντιπροσωπεύουν στο σύνολο του πληθυσμού.
- **Sum of W_FSTUWT (s.e.):** Το τυπικό σφάλμα του σταθμισμένου αθροίσματος των μαθητών (Sum of W_FSTUWT).
- **Percent:** Το ποσοστό (%) του πληθυσμού της χώρας που βρίσκεται στη συγκεκριμένη κατηγορία επίδοσης.
- **Percent (s.e.):** Το τυπικό σφάλμα του ποσοστού του πληθυσμού της χώρας που βρίσκεται στη συγκεκριμένη κατηγορία επίδοσης.

Graph



>>

Εικόνα 155. Γράφημα κατανομής μαθητών ανά επίπεδο επίδοσης – Benchmarks (SPSS)

Ερμηνεία αποτελεσμάτων

Με βάση τα αποτελέσματα που παρουσιάζονται στην Εικόνα 154, καθώς και στο γράφημα της Εικόνας 155, παρατηρείται ότι στην Ελλάδα στα δύο χαμηλότερα επίπεδα επίδοσης (κάτω από 420.07 μονάδες) συγκεντρώνεται το 47.21% των μαθητών, ενώ στην Πορτογαλία το αντίστοιχο ποσοστό είναι 29.70%. Επομένως, το ποσοστό των μαθητών με χαμηλή επίδοση είναι σαφώς υψηλότερο στην Ελλάδα σε σχέση με την Πορτογαλία. Όσον αφορά το μεσαίο επίπεδο (ομάδες 3 και 4), τα ποσοστά για τις δύο χώρες είναι σχεδόν ίδια. Συγκεκριμένα, αθροιστικά στο μεσαίο επίπεδο εντάσσεται το 43.31% των μαθητών στην Ελλάδα και το 47.99% στην Πορτογαλία. Αντίθετα, στα ανώτερα επίπεδα επίδοσης (πάνω από 544.68 μονάδες), στην Ελλάδα βρίσκεται το 9.48% των μαθητών, ενώ στην Πορτογαλία το 22.30%. Συμπερασματικά, στην Ελλάδα η επίδοση στα Μαθηματικά συγκεντρώνεται κυρίως στις χαμηλότερες βαθμίδες της κλίμακας, ενώ στην Πορτογαλία η κατανομή εμφανίζεται μετατοπισμένη προς τα μεσαία και υψηλά επίπεδα.

