

## Σεμινάριο SPSS

Δραστηριότητα εξάσκησης

Τρίτη, 19 Μαρτίου 2024

Κωνσταντίνος Λαβίδας

## Ενότητα Περιγραφική Στατιστική

### Προετοιμασία για το εργαστήριο

- Δημιουργία προσωπικού φακέλου στον υπολογιστή.  
Στα «έγγραφα μου» φτιάξτε φάκελο με όνομα τον Αριθμό Μητρώου σας. Εκεί θα αποθηκεύετε τα αρχεία με τα οποία θα ασχοληθούμε σε όλες τις εργαστηριακές συναντήσεις. Μπορείτε αν θέλετε, να αντιγράψετε τα αρχεία αυτά και σε κάποιο flash memory.
- Τρέξτε (φορτώστε) το spss ....

### Δραστηριότητα 1<sup>η</sup>.

#### A) Εισαγωγή δεδομένων τύπου san του SPSS

Ας υποθέσουμε ότι έχουμε στη διάθεσή μας δεδομένα από έρευνα αξιολόγησης ενός προϊόντος.

- 1.A.1) Φορτώστε το αρχείο δεδομένων: (Εντολή: **file – open - data**): αρχείο «yogurt».
- 1.A.2) Παρατηρείστε τις δύο επιλογές Data View και Variable View, αριστερά κάτω.  
Με την επιλογή Data View βλέπω .....  
Με την επιλογή Variable View βλέπω .....
- 1.A.3) Ποια είναι τα στοιχεία παρατήρησης (μεταβλητές) του θέματος που μελετάω;
- 1.A.4) Υποδείξτε ποιες μεταβλητές είναι Ποιοτικές και ποιες Ποσοτικές.

#### B) Ρυθμίσεις εμφάνισης των μεταβλητών

Στο περιβάλλον δημιουργίας και αναδόμησης χαρακτηριστικών – ιδιοτήτων των μεταβλητών (**variable view**, κάτω μέρος και αριστερά) βλέπω τις ιδιότητες που μπορώ να δώσω σε μια μεταβλητή.

- 1.B.1) Ποιες είναι οι ιδιότητες αυτές; .....
- 1.B.2) Να σημειώσετε στην στήλη **Label**, αναλυτικές περιγραφές για τις υπόλοιπες μεταβλητές: advertise, rating, cost, calories.

## «Ποσοτικοποίηση» των δεδομένων

### Κλειστού τύπου παρατήρησης ή ερωτήσεις

Σε όλες τις ποιοτικές μεταβλητές μπορούμε να εισάγουμε και αριθμούς δηλώνοντας ταυτόχρονα με ποιες λέξεις ή αλφαριθμητικά (string) αντιστοιχούν οι αριθμοί.

1.B.3) Να δημιουργήσετε, σύμφωνα με τον πίνακα στα δεξιά, λεκτικές αντιστοιχίες (Values) των τιμών 1, 2, 3, 4 και 5 της μεταβλητής rating (αξιολογική κλίμακα)

*Κάνω κλικ στο κελί της στήλης values, για τη μεταβλητή που επιθυμώ. Κάθε φορά αφού συμπληρώσω το αντίστοιχο value και label πατώ add. Επιπλέον αν επιθυμώ έχω στη διάθεσή μου και επιλογές για αλλαγή (change) αλλά και για διαγραφή (remove) της αντίστοιχης ετικέτας*

Value	Label
1	φτωχό
2	μέτριο
3	καλό
4	πολύ καλό
5	εξαιρετικό

### Ανοικτού τύπου ερωτήσεις (πιο ελεύθερη παρατήρηση)

1.B.4) Για την μεταβλητή comments να δημιουργήσετε μια καινούργια μεταβλητή (με όνομα: location) στην οποία θα αντιστοιχήσετε τις φράσεις της αρχικής μεταβλητής σε μοναδικές τιμές (σαν να έχετε μια ερώτηση κλειστού τύπου). Μια ενδεικτική πρόταση (μελέτησα όλες τις απαντήσεις και κατέληξα σε μια συνοπτική λύση) είναι να δημιουργήσετε μια μεταβλητή που να παίρνει τις τιμές: (0. όχι, 1. περίπου, 2. ναι, 99. Δεν απαντώ, 999. άσχετο).

1.B.5) Για ποιο είδος μεταβλητών πρέπει να φροντίζω να έχω λεκτικές; αντιστοιχίες (ετικέτες) των τιμών τους; .....

***Προσοχή.** Δεν ξεχνώ να αποθηκεύσω τις επιλογές μου. Αποθηκεύστε το αρχείο των δεδομένων με το όνομα «yogurt» ( Εντολή: file -save) Για την αποθήκευση είναι χρήσιμο να σκεφτώ που θα το αποθηκεύσω και με ποιο όνομα. Ιδιαίτερα για το όνομα, συνήθως προτιμώ κάτι που να ταιριάζει στο περιεχόμενο του αρχείου.*

### Γ) Σύντομη παρουσίαση των δεδομένων: Πίνακας Συχνοτήτων

1.Γ.1) Περιγράψτε σύντομα τις κατανομές των χαρακτηριστικών advertise και rating, των γιαουρτιών του δείγματος. (Παρόμοια: ποια είναι η κατανομή των τιμών των δύο αυτών μεταβλητών)

«Ζητήστε» από το λογισμικό να εξάγει αντίστοιχους πίνακες συχνοτήτων (**frequency tables**). (Εντολή: **Analyze – Descriptive Statistics – frequencies...**)

*Προσοχή.* τα όποια αποτελέσματα παράγουν οι ενέργειες ανάλυσης που πραγματοποιείτε, εμφανίζονται σε νέο περιβάλλον – παράθυρο «**Output .....**».

1.Γ.2) Αποθηκεύστε το **αρχείο των αποτελεσμάτων** με όνομα «yogurt\_output»

1.Γ.3) Σχολιάστε τις συχνότητες και τις σχετικές συχνότητες – **Ποσοστά (Percent)**.

Ποιο από τα δύο είναι σημαντικότερα να σχολιάζετε σε μια παρουσίαση και γιατί;.....

.....

1.Γ.4) Ποια η διαφορά των Percent και Valid Percent;

1.Γ.5) Γιατί υπάρχει διαφορά στις δύο στήλες, στον πίνακα συχνοτήτων, για τη μεταβλητή Advertise;

1.Γ.6) Οι επόμενες στήλες τι φανερώνουν; Για ποιο είδος μεταβλητών έχουν νόημα;

**Missing Values - Χαμένες τιμές:** είναι οι τιμές για τις οποίες δεν έγινε καταγραφή (system) ή έχουν αποκλειστεί κατά την ανάλυση (user)

1.Γ.7) Θα θέλαμε για επεκτείνουμε την ανάλυση των δεδομένων μας, να είχαμε δηλαδή και έναν πίνακα συχνοτήτων της μεταβλητής «calories». Μπορείτε να εξάγετε έναν πίνακα συχνοτήτων; τι παρατηρείτε από τον συγκεκριμένο πίνακα συχνοτήτων;

1.Γ.8) Για ποιο είδος μεταβλητών έχει νόημα να υπολογίζουμε συχνότητες και σχετικές συχνότητες; .....

1.Γ.9) Για τη συγκεκριμένη μεταβλητή «calories», τι θα μπορούσαμε να κάνουμε έτσι ώστε να την παρουσιάσουμε με συνοπτικό τρόπο, δηλαδή με πίνακα συχνοτήτων;

.....

## **Δραστηριότητα 2<sup>η</sup>**

### **A) Εισαγωγή δεδομένων που έχουν καταχωρηθεί σε άλλο πρόγραμμα**

Δεδομένα που είναι αποθηκευμένα σε κάποιο αρχείο του excel, μπορούν να εισαχθούν στο spss για επεξεργασία.

2.A.1) Άνοιγμα του αρχείου «Marks.xls» μέσω του SPSS (Εντολή: **file-Open-Data**, στην επιλογή **Files of type** επιλέγω **Excel(\*.xls....)**)

2.A.2) Ποια είναι τα στοιχεία παρατήρησης (μεταβλητές) του θέματος που μελετάω;

### **B) Προσθήκη και διόρθωση δεδομένων**

2.B.1) Επιλέξτε το περιβάλλον των δεδομένων **data view**. Εισάγετε στο τέλος τα

**!** Προσοχή με το δεξί κλικ σε κάποιο σημείο μια λίστα από εργαλεία εμφανίζεται, η οποία σας παρέχει την εντολή που ζητάτε.

δεδομένα του φοιτητή «ονοματεπώνυμο δικό σας» και τους βαθμούς που επιθυμείτε.

2.B.2) Να εισάγετε μετά τη δέκατη γραμμή τα δεδομένα του

φοιτητή «Παππά Μάξιμου, 57, 62, 61» (Εντολή: Επιλέγω την επόμενη γραμμή και **Edit - Insert cases**)

*Case ή περίπτωση, λέμε την κάθε εγγραφή δεδομένων για το κάθε υποκείμενο της έρευνας*

2.B.3) Να διαγράψετε την φοιτήτρια «ΚΛΙΑΝΗ ΣΟΦΙΑ»

2.B.4) Να μετακινήσετε την περίπτωση του μαθητή «Πούλιου Παναγιώτα» ακριβώς πριν από τον «Καλόφωνο Νίκο»

2.B.5) Αλλάξτε τη θέση των μεταβλητών Test 3 και Test 2.

### **Γ) Δημιουργία καινούριας μεταβλητής**

2.Γ.1) Να εισάγετε τη νέα μεταβλητή (variable) με όνομα Gender αμέσως τη μεταβλητή name η οποία θα καταγράφει το «φύλο». Για μεταβλητές μπορείτε εναλλακτικά να δουλεύετε και στο περιβάλλον data view. (Εντολή: **Επιλέγω την επόμενη μεταβλητή και Edit - Insert variable**)

2.Γ.2) Δημιουργήστε λεκτικά σχόλια για τις τιμές της καινούριας μεταβλητής: ( 0. «αγόρι», 1. «κορίτσι»). Θα μπορούσα να αντιστοιχήσω άλλους αριθμούς και αν ναι γιατί δεν το επέλεξα;

2.Γ.3) Πληκτρολογήστε τώρα το φύλο για κάθε μαθητή. Αυτό προσδιορίζετε από το όνομα.

### **Δ) Λειτουργική εμφάνιση των δεδομένων**

2.Δ.1) Να εμφανίσετε τις λεκτικές περιγραφές των τιμών των μεταβλητών, σε όσες μεταβλητές υπάρχουν. (Εντολή: **View –value labels**). Βρείτε το αντίστοιχο εικονίδιο στην εργαλειοθήκη που φαίνεται κάτω από το μενού.

2.Δ.2) Να παρουσιάσετε τα δεδομένα ταξινομημένα ως προς το επώνυμο σε αύξουσα σειρά. (δεξί κλικ στην αντίστοιχη στήλη **Sort Ascending**)

2.Δ.3) Να παρουσιάσετε τα δεδομένα ταξινομημένα ως προς το βαθμό του πρώτου τεστ σε φθίνουσα σειρά (δεξί κλικ στην αντίστοιχη στήλη **Sort .....**).

2.Δ.4) Να παρουσιαστούν τα δεδομένα ταξινομημένα σε αύξουσα σειρά ως προς το βαθμό του πρώτου τεστ και μετά ως προς το επώνυμο. (Εντολή: **Data –sort cases.... σε κάθε μεταβλητή ξεχωριστά το είδος της ταξινόμησης**)

### **«Επέκταση των δεδομένων»**

### **Ε) Δημιουργία μεταβλητής βασισμένη σε υπολογισμούς άλλων μεταβλητών**

2.E.1) Πως θα συνοψίζατε την επίδοση των μαθητών σας στα τρία τεστ;

2.Ε.2) Να υπολογίσετε τη μέση βαθμολογία στα τρία test (το όνομα αυτής ΜΟ) για κάθε μαθητή. (Εντολή: **Transform – compute, στην περιοχή Numeric Expression πληκτρολογείτε (TEST1+TEST2+TEST3)/3**)

**ΣΤ) Δημιουργία μεταβλητής βασισμένη στην επανακωδικοποίηση τιμών άλλης μεταβλητής.**

2.ΣΤ.1) Θέλουμε να παρουσιάσουμε συνοπτικά την κατανομή των τιμών της μεταβλητής ΜΟ που συνοψίζει την επίδοση των μαθητών σας. Τι προτείνετε;

2.ΣΤ.2) Να δημιουργήσετε καινούρια μεταβλητή με όνομα «estimate» από επανακωδικοποίηση των τιμών της μεταβλητής τελικός βαθμός (ΜΟ) σε μεταβλητή «με κλάσεις - ομάδες τιμών» ως εξής: ([μικρότερος ή ίσος με 60]=απέτυχε, (κωδικός 1), (60, 70]= μέτριος (κωδικός 2), (70, 80]= καλός (κωδικός 3), (80, 100]= Πολύ καλός ή άριστος (κωδικός 4). (Εντολή: **Transform – Recode Into different variables...**)

2.ΣΤ.3) Τι τύπος κλίμακας μέτρησης αντιστοιχεί στη μεταβλητή «estimate»;

#### **Z) Ανάλυση των δεδομένων**

2.Z.1) Παρουσιάστε συνοπτικά την επίδοση (μεταβλητή estimate) της τάξης.

#### **Εργασία**

Για το τελευταίο ερώτημα φτιάξτε στο word μια μικρή αναφορά επισυνάπτοντας τον πίνακα συχνοτήτων και σχετικών συχνοτήτων που θα υποδεικνύει τα όσα θα γράψετε. Ο πίνακας (δες στα επόμενα την μορφή που πρέπει να έχει) ακολουθεί του κειμένου.

*Για τον πίνακα, σύμφωνα με το σύστημα της American Psychology Assosiation (APA) στη πρώτη γραμμή τοποθετείται ο αριθμός του πίνακα και ο τίτλος του πίνακα: «Πίνακας αριθμός Τίτλος του πίνακα». Ο αριθμός του πίνακα εμφανίζεται μόνος του δίπλα από την λέξη πίνακας με κανένα σημείο στίξης. Μετά τον αριθμό και στην επόμενη γραμμή ακολουθεί με πλάγια γράμματα ο τίτλος του πίνακα. Προσοχή οι στήλες και οι γραμμές του πίνακα πρέπει να έχουν ξεκάθαρους τίτλους και να μην περιέχουν κάθετες γραμμές (χωρίς περίγραμμα κάθετων γραμμών). Τέλος μέσα στο κείμενο που αφορά το σχολιασμό του πίνακα γράφουμε την παραπομπή, συνήθως με τη μορφή (δες πίνακα αριθμός).*