

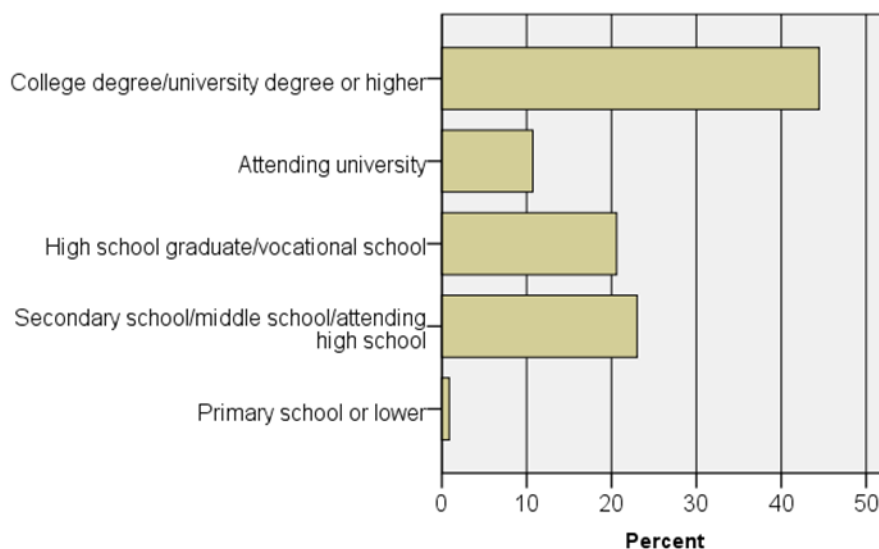
Θέματα Περιγραφικής Στατιστικής

Για τα παρακάτω θέματα υπάρχουν παραπομπές σε σύντομα βίντεο που βρίσκονται στον φάκελο «Έγγραφα/ Ποσοτική Μεθοδολογία – Στατιστική/ Video στατιστικής με το SPSS» της e-class. Σε μια περίπτωση σας παραπέμπω σε video του youtube. Τα θέματα 1,2, 3, θα τα απαντήσετε μελετώντας τις έννοιες:

- Κατανομή συχνοτήτων (ποσοτικών και κατηγορικών μεταβλητών
 - Κεντρική θέση (μέση τιμή, διάμεσος, επικρατέστερη τιμή) της κατανομής μιας μεταβλητής
 - Μεταβλητότητα (διακύμανση, τυπική απόκλιση) της κατανομής μιας ποσοτικής μεταβλητής
 - Συμμετρία (συμμετρική, θετική ασυμμετρία, αρνητική ασυμμετρία) της κατανομής μιας ποσοτικής μεταβλητής
- Βιβλιογραφία
(στο φάκελο «Έγγραφα/ Ποσοτική Μεθοδολογία – Στατιστική/ Βιβλιογραφία» της e-class
 1. [Στατιστικές μέθοδοι και τεχνικές στις κοιν. επιστήμες με χρήση SPSS.pdf](#)
 2. [A Handbook of Statistical Analyses using SPSS - Excellent !!!..pdf](#)
 3. Τα video που αναφέρθηκαν παραπάνω

=====

1. (1 μονάδα) Σε μια έρευνα σχετική με το προφίλ του Ελληνοκύπριου χρήστη διαδικτύου (Κύπρος 2010) που βασίστηκε σε ένα δείγμα 1000 ατόμων, παρουσιάζεται το γράφημα που περιγράφει τα ποσοστά των 587 χρηστών σε κάθε μια από τις κατηγορίες εκπαίδευσης (0,5 μονάδα).



- a. Ποιος είναι ο πληθυσμός ο οποίος αναφέρεται στο παραπάνω γράφημα;

πληθυσμός των ελληνοκυπρίων χρηστών διαδικτύου

- b. Αν ο μέσος χρήστης είναι Άνδρας, 33 ετών, παντρεμένος, κάτοικος αστικής περιοχής και τα στοιχεία αυτά προκύπτουν από την επεξεργασία τεσσάρων μεταβλητών, να ονομαστούν οι μεταβλητές.

Φύλο, Ηλικία, Οικογενειακή κατάσταση, Αστικότητα τόπου διαμονής

- c. Ποια είναι η μεταβλητή στην οποία αναφέρεται το παραπάνω γράφημα;

Μορφωτικό επίπεδο

2. (1 μονάδα) Ποιο από τα παρακάτω είναι κατάλληλο στο να εκφράσει την κεντρική θέση στην ανάλυση των απαντήσεων της ερώτησης «ποια είναι η συσκευή με τη οποία συνδέεστε συχνότερα στο διαδίκτυο;»:

1. Η τυπική απόκλιση
2. Θετική ασυμμετρία
3. **Η επικρατέστερη τιμή**
4. Η διάμεσος
5. Το ενδοτεταρτημοσιακό εύρος
6. Η αθροιστική συχνότητα της τιμής με τη μεγαλύτερη συχνότητα
7. Ο αριθμός κατηγοριών μιας κατηγορικής μεταβλητής
8. Η Σχετική συχνότητα τη τιμής με τη μεγαλύτερη συχνότητα

3. (3 μονάδες) Προκειμένου να επεξεργαστείτε τα δεδομένα του αρχείου `statist_background.xlsx` θα εκτελέσετε τα παρακάτω :

- a. Να μετατρέψετε το αρχείο σε αρχείο δεδομένων SPSS με το ίδιο όνομα **statist_background.sav**. Με το τελευταίο θα εργαστείτε στα επόμενα ερωτήματα του θέματος και στα υπόλοιπα θέματα (βίντεο: [εισαγωγή δεδομένων Excel-συχνότητες.mp4](#))
- b. Ονομασία των μεταβλητών των στηλών B έως G σε q1, q2,...q6 (αυτό θα γίνει από τη λίστα των μεταβλητών **variable view** στην στήλη **name**)
- c. Αριθμητική κωδικοποίηση των μεταβλητών των μεταβλητών q1, q2,...q5 σε q1_n, q2_n,...,q5_n αντίστοιχα, με χρήση της εντολής **automatic recode** από την επιλογή **Transform**. Αυτό είναι απαραίτητο αφού οι τιμές των μεταβλητών είναι ονόματα και πρέπει να μετατραπούν σε αριθμούς. Στο SPSS όλες οι μεταβλητές είναι αριθμητικού τύπου είτε είναι κατηγορικές είτε ποσοτικές. Θα εκτελέσετε την εντολή για τις q1 έως q4 μαζί, επειδή έχουν τον ίδιο τρόπο απάντησης, επιλέγοντας πριν την εκτέλεση τις «**use the same recoding scheme for all variables**» και «**Treat blank values as user-missing**». Η πρώτη εξασφαλίζει ότι το 1 και το 2 θα έχουν ακριβώς το ίδιο νόημα σ' όλες τις μεταβλητές ενώ η δεύτερη θα δώσει την τιμή 3 στο κενό δηλαδή την απουσία τιμής και θα δηλώσει στην στήλη **missing** στο **variable view** ότι το 3 σημαίνει ελλείπουσα τιμή.
Θα εκτελέσετε ξανά μια δεύτερη φορά την **automatic recode** για την q5_n μόνη της (Βίντεο: <https://www.youtube.com/watch?v=NZlgkRHva5Y>)
- d. Με την εντολή **recode into same variable** επανακωδικοποιήσετε q5_n αφού μελετήστε το πως έχει δώσει τις τιμές 1 έως 5 το SPSS. Αν για παράδειγμα έχει

δώσει την τιμή 1 στην δήλωση «δεν είμαι βέβαιος/η» αυτό δεν είναι επιθυμητό για την μελλοντική επεξεργασία τέτοιων ποιοτικών μεταβλητών. Μετατρέψετε τις τιμές έτσι ώστε να ισχύει:

1= καθόλου

2= λίγο

3 = δεν είμαι βέβαιος/η

4= πολύ

5= πάρα πολύ

(Βίντεο: [δημιουργία κλιμάκων.mp4](#))

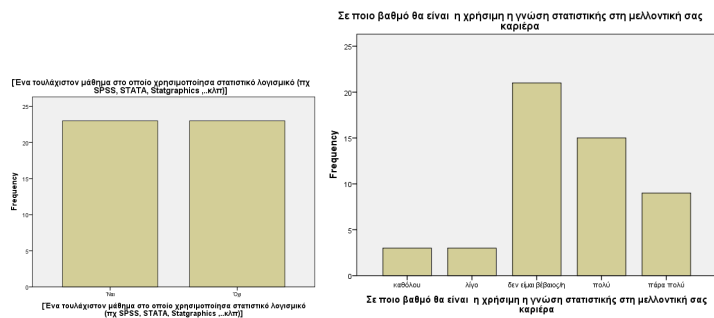
Επίσης, θα δηλώστε στη στήλη “Values” την σωστή σημασία των τιμών 1 έως 5 για την q5_n όπως προαναφέρθηκε.

Τα αποτελέσματα της εργασίας αυτής πρέπει να παρουσιάζονται στο αρχείο δεδομένων

4. (1 μονάδα) Να σχολιαστούν τόσο η εμπειρία των φοιτητών/τριών όσο και στάση τους απέναντι στην μελλοντική χρησιμότητα της ποσοτικής μεθοδολογία και στατιστικής με τη βοήθεια ραβδόγραμματος για κάθε μια από τις μεταβλητές q1_n έως q5_n. (υπόδειξη: κατανομή συχνοτήτων από την επιλογή **analyze->descriptive statistics->frequencies**, Βίντεο: [κατανομές συχνοτήτων και στατιστικά.mp4](#) , [ραβδόγραμμα & κυκλικό διάγραμμα.mp4](#))

Στα παρακάτω ραβδόγραμματα απεικονίζονται οι απαντήσεις των φοιτητών/τριών αναφορικά με την εμπειρία τους και την άποψή τους για μελλοντική χρησιμότητα της στατιστικής. Από τα γραφήματα προκύπτει ότι η συντριπτική πλειοψηφία των δαπανώντων έχει διδαχθεί μεθοδολογία της έρευνας (περίπου 75%). Ένας/μία στους δύο έχουν διδαχθεί ένα μάθημα περιγραφικής στατιστικής και η ίδια αναλογία ισχύει για μάθημα που χρησιμοποιήθηκε στατιστικό λογισμικό. Η μικρότερη εμπειρία που δηλώθηκε αφορούσε την επαγωγική στατιστική (περίπου 45%). Τέλος από το ραβδόγραμμα της «μελλοντικής χρησιμότητας της γνώσης στη στατιστική» διαπιστώνεται ότι η συχνότερη απάντηση ήταν η αβεβαιότητα για την μελλοντική χρησιμότητα. Όμως, από του φοιτητές/τριες που είχαν βεβαιότητα, η μεγάλη πλειοψηφία αποτελούμενη από 24 άτομα, δήλωσε πολύ ή πάρα πολύ έναντι 6 που δήλωσαν καθόλου ή λίγο χρήσιμη.





5. (1 μονάδα) Να χρησιμοποιηθεί η επιλογή **compare means-> means** προκειμένου να διερευνηθεί περιγραφικά το ερώτημα της συνάφειας της «μελλοντικής χρησιμότητας της στατιστικής (q5_n) με την εμπειρία των φοιτητών σε κάποιο μάθημα περιγραφικής στατιστικής (q2_n) (θα θεωρήσουμε ότι μπορεί να χρησιμοποιηθεί στην διερεύνηση αυτή η μέση τιμή της q5_n). Το ίδιο ερώτημα να απαντηθεί για το μάθημα της επαγωγικής στατιστικής και για την χρήση στατιστικού λογισμικού.

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα * q2_n [Εναι τουλάχιστον μάθημα που περιλάμβανε μεθόδους περιγραφικής στατιστικής]

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα

q2_n [Εναι τουλάχιστον μάθημα που περιλάμβανε μεθόδους περιγραφικής στατιστικής]	Mean	N	Std. Deviation
1 'Ναι	3,78	23	1,043
2 'Όχι	3,32	22	1,041
Total	3,56	45	1,056

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα * q3_n [Εναι τουλάχιστον μάθημα που περιλάμβανε μεθόδους επαγωγικής στατιστικής]

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα

q3_n [Ένα τουλάχιστον μάθημα που περιλάμβανε μεθόδους επαγωγικής στατιστικής]	Mean	N	Std. Deviation
1 'Ναι	3,80	20	1,056
2 'Όχι	3,38	24	1,056
Total	3,57	44	1,065

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα * q4_n [Ένα τουλάχιστον μάθημα στο οποίο χρησιμοποίησα στατιστικό λογισμικό (πχ SPSS, STATA, Statgraphics ,...κλπ)]

q5_n Σε ποιο βαθμό θα είναι η χρήσιμη η γνώση στατιστικής στη μελλοντική σας καριέρα

q4_n [Ένα τουλάχιστον μάθημα στο οποίο χρησιμοποίησα στατιστικό λογισμικό (πχ SPSS, STATA, Statgraphics ,...κλπ)]	Mean	N	Std. Deviation
1 'Ναι	3,61	23	1,033
2 'Όχι	3,43	23	1,121
Total	3,52	46	1,070

Για την απάντηση στο ερώτημα θα συγκριθούν οι μέσες τιμές μεταξύ αυτών που παρακολούθησαν και αυτών που δεν έχουν παρακολουθήσει κάποιο μάθημα για κάθε ένα από τους παραπάνω πίνακες. Συμπεραίνουμε ότι και στις 3 περιπτώσεις οι φοιτητές/τριες που έχουν παρακολουθήσει κάποιο μάθημα έχουν υψηλότερο μέση τιμή, δηλαδή

θεωρούν χρησιμότερη τη γνώση στατιστικής στην μελλοντική τους καριέρα, σε σχέση με εκείνες/ους που δεν έχουν παρακολουθήσει ένα τέτοιο μάθημα . Η επίδραση από την ισχυρότερη στην περίπτωση της περιγραφικής στατιστικής, και λιγότερο ισχυρή εκείνη του στατιστικού λογισμικού

6. (1,5 μονάδες) Να αναζητηθούν οι κατηγορίες: “χρήση του SPSS ή άλλου λογισμικού επεξεργασία δεδομένων”, “κατανόηση της στατιστικής ανάλυσης των επιστημονικών εργασιών ή συγγραμμάτων» και «χρήση ποσοτικής μεθοδολογίας σε μελλοντική εργασία», έτσι ώστε να δημιουργηθούν 3 μεταβλητές αντίστοιχα (πχ.

P1, P2, P3) με τιμές «0=όχι/1=ναι/» ανάλογα με το αν η κατηγορία βρίσκεται στην απάντηση του φοιτητή/τριας.

Το αποτέλεσμα της επεξεργασίας που δεν είναι άλλη από την ανάγνωση των απαντήσεων και καταγραφή με 0 ή 1 της παρουσίας μια κατηγορίας, παρουσιάζεται στο αρχείο δεδομένων του SPSS.

7. (1,5 μονάδες) Να αποτυπωθεί και να σχολιαστεί το αναμενόμενο μαθησιακό αποτέλεσμα με κάποια μέθοδο περιγραφικής στατιστικής (πίνακες ή γραφήματα, **βίντεο: [ραβδόγραμμα & κυκλικό διάγραμμα.mp4](#)**) χρησιμοποιώντας τα αποτελέσματα του θέματος 7. Πχ. τι επιλέγουν περισσότερο οι φοιτήτριες και οι φοιτητές.

Θα μπορούσε να χρησιμοποιηθεί είτε το ραβδόγραμμα είτε το κυκλικό διάγραμμα στην περίπτωση αυτή για κάθε μια από τις μεταβλητές. Παρακάτω παρουσιάζονται τα ραβδόγραμμα ένα για κάθε μεταβλητή. Από αυτό παρατηρούμε προκύπτει ότι πρώτη επιλογή των φοιτητών/τριων είναι η “χρήση του SPSS ή άλλου λογισμικού επεξεργασία δεδομένων» με 44,9% και ακολουθεί η «χρήση ποσοτικής μεθοδολογίας σε μελλοντική εργασία» με ποσοστό 42,86%. Το μακρότερο ποσοστό 38,78% παρατηρήθηκε στις δηλώσεις που συνδέονται με κατανόηση της στατιστικής ανάλυσης των επιστημονικών εργασιών ή συγγραμμάτων. Ωστόσο οι παρατηρούμενες διαφορές δεν είναι αξιόλογες.

