

ΕΝΟΤΗΤΑ 2.5 ΠΙΝΑΚΕΣ ΣΥΝΑΦΕΙΑΣ ΚΑΙ ΡΑΒΔΟΓΡΑΜΜΑΤΑ

Διερεύνηση

*Κατανομή
μαθητών/μαθητριών
σε ομάδες
προσανατολισμού
σπουδών*

Σε ένα Γενικό Λύκειο καταγράφηκαν το φύλο και η ομάδα προσανατολισμού σπουδών που επέλεξαν οι 50 μαθητές/μαθήτριες της Β' τάξης. Συγκεκριμένα, 11 αγόρια επέλεξαν την ομάδα προσανατολισμού Ανθρωπιστικών Σπουδών και 18 αγόρια επέλεξαν την ομάδα προσανατολισμού Θετικών Σπουδών. Αντίστοιχα, 9 κορίτσια επέλεξαν την ομάδα προσανατολισμού Ανθρωπιστικών Σπουδών και 12 κορίτσια την ομάδα προσανατολισμού Θετικών Σπουδών.

α) Να οργανώσετε τα παραπάνω δεδομένα σε έναν πίνακα συνάφειας (όπως δίνεται παρακάτω) που να περιέχει τις απόλυτες συχνότητες και τις σχετικές συχνότητες:

- ως προς το σύνολο των μαθητών/τριών της Β' Λυκείου
- ως προς το φύλο (γραμμές)
- ως προς την ομάδα προσανατολισμού Σπουδών (στήλες)

		Ομάδα Προσανατολισμού Σπουδών		
		Ανθρωπιστικών	Θετικών	Σύνολο
Φύλο	Αγόρια			
	Κορίτσια			
	Σύνολο			

β) Να κατασκευάσετε κατάλληλα γραφήματα συχνοτήτων και σχετικών συχνοτήτων

- ως προς το σύνολο των μαθητών/τριών της Β Λυκείου
- ως προς το φύλο (γραμμές)
- ως προς την ομάδα προσανατολισμού Σπουδών (στήλες)

γ) Θα μπορούσατε να πείτε εάν το φύλο παίζει ρόλο στην επιλογή της ομάδας προσανατολισμού σπουδών;

Βασικές μαθηματικές έννοιες – Ιδέες – Διεργασίες

Στα διάφορα προβλήματα που εξετάσαμε έως τώρα στη Στατιστική ασχοληθήκαμε κάθε φορά με μία μόνο μεταβλητή ξεχωριστά, ποιοτική ή ποσοτική, π.χ. με το φύλο των μαθητών, με τον χρόνο των αθλητών σε έναν αγώνα δρόμου, τον μηνιαίο μισθό των υπαλλήλων μιας εταιρείας. Μια πραγματική στατιστική ανάλυση

δεν περιορίζεται συνήθως στη μελέτη μιας μεταβλητής, αλλά απαιτείται η μελέτη της σχέσης μεταξύ δύο ή περισσότερων μεταβλητών. Στην παράγραφο αυτή θα ασχοληθούμε με την αναζήτηση σχέσεων ανάμεσα σε δύο ποιοτικές μεταβλητές.

Ας δούμε τις βασικές ιδέες της διερεύνησης ενός προβλήματος με δύο ποιοτικές μεταβλητές μέσα από ένα παράδειγμα. Σε μια μελέτη που αφορά τη χρήση ζώνης ασφαλείας από οδηγούς ηλικίας 18 – 24 ετών μετρήθηκαν 198 οδηγοί σύμφωνα με το φύλο και το ενδεχόμενο να είχαν ή όχι κάποιο τροχαίο ατύχημα τα τελευταία πέντε χρόνια. Η εύρεση της πιθανής σχέσης μεταξύ του φύλου (χαρακτηριστικό A) και τροχαίου ατυχήματος (χαρακτηριστικό B) επιτυγχάνεται μέσω της κατασκευής του πίνακα συνάφειας (contingency table), ο οποίος είναι ένας πίνακας συχνοτήτων της μορφής:

**πίνακας συνάφειας
συχνοτήτων**

		Τροχαίο Ατύχημα		
		Ναι	Όχι	Σύνολο
Φύλο	Ανδρες	69	58	127
	Γυναίκες	27	44	71
	Σύνολο	96	102	198

Πίνακας 1: Πίνακας Συνάφειας απόλυτων συχνοτήτων φύλου και συμμετοχής σε ατύχημα.

Οι γραμμές του πίνακα συνάφειας (Πίνακας 1) αντιστοιχούν στις κατηγορίες της μίας ποιοτικής μεταβλητής (π.χ. του φύλου) και οι στήλες στις κατηγορίες της άλλης ποιοτικής μεταβλητής (π.χ. του ατυχήματος), ενώ στο εσωτερικό του παρατίθενται οι συχνότητες που αντιστοιχούν σε όλους τους δυνατούς συνδυασμούς των κατηγοριών των δύο μεταβλητών.

Το πλήθος των ατόμων στις κατηγορίες του Πίνακα 1 είναι διαφορετικό, οπότε δεν μπορούμε να εξηγήσουμε ποια είναι η σχέση μεταξύ φύλου και συμμετοχής σε ατύχημα. Προκειμένου να κατανοήσουμε καλύτερα αυτή τη σχέση, θα πρέπει να εξετάσουμε τα ποσοστά αντί για τις συχνότητες. Ο πίνακας 1 μπορεί να μετατραπεί σε πίνακα σχετικών συχνοτήτων (%) ως προς το σύνολο των παρατηρήσεων του δείγματος n (Πίνακας 2), ως προς το πλήθος των παρατηρήσεων του χαρακτηριστικού A (Πίνακας 3) και ως προς το πλήθος των παρατηρήσεων του χαρακτηριστικού B (Πίνακας 4). Από τον Πίνακα 2, προκύπτει ότι :

**πίνακες συνάφειας
σχετικών συχνοτήτων**

- Στο δείγμα της έρευνάς μας συμμετείχαν περισσότεροι άνδρες οδηγοί από γυναίκες οδηγούς.
- Οι περισσότεροι συμμετέχοντες στην έρευνα δε συμμετείχαν σε ατύχημα, ωστόσο δεν είναι και λίγοι εκείνοι που συμμετείχαν.



		Τροχαίο Ατύχημα		
		Ναι	Όχι	Σύνολο
Φύλο	Άνδρες	34,9%	29,3%	64,2%
	Γυναίκες	13,6%	22,2%	35,8%
	Σύνολο	48,5%	51,5%	100%

Πίνακας 2: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και ατυχήματος ως προς το σύνολο των παρατηρήσεων του δείγματος

Η εμφάνιση των ποσοστών στο εσωτερικό ενός πίνακα συνάφειας μας δίνει τη δυνατότητα να διακρίνουμε τη μορφή της σχέσης που ενδεχομένως υπάρχει μεταξύ των μεταβλητών ενός πίνακα. Στην περίπτωση δύο μεταβλητών, η μια μεταβλητή μπορεί να παίζει τον ρόλο της ανεξάρτητης μεταβλητής (το φύλο) και η άλλη να παίζει τον ρόλο της εξαρτημένης (από το φύλο) μεταβλητής (το τροχαίο ατύχημα).

		Τροχαίο Ατύχημα		
		Ναι	Όχι	Σύνολο
Φύλο	Άνδρες	54,3%	45,7%	100%
	Γυναίκες	38%	62%	100%
	Σύνολο			

Πίνακας 3: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και συμμετοχής σε ατύχημα ως προς το φύλο

Από τον Πίνακα 3 φαίνεται ότι οι άνδρες οδηγοί του δείγματος είχαν σε μεγαλύτερο ποσοστό ατυχήματα τα τελευταία πέντε χρόνια, σε σχέση με τις γυναίκες οδηγούς. Φαίνεται πως στο δείγμα μας υπάρχει μια σχέση ανάμεσα στο φύλο και στη συμμετοχή σε ατύχημα στο ιστορικό του οδηγού: οι άνδρες οδηγοί του δείγματος είναι πιο πιθανό να είχαν ατύχημα τα τελευταία 5 χρόνια. Το αν αυτή η σχέση που ανακαλύψαμε στο δείγμα μας θα μπορούσε να γενικευθεί στον πληθυσμό όλων των οδηγών δεν είναι προφανές. Εξαρτάται πρώτα απ' όλα από το αν το δείγμα που επιλέξαμε είναι αντιπροσωπευτικό του πληθυσμού. Ακόμα κι αν αυτό έχει εξασφαλιστεί με τη δειγματοληψία, το εύρημά μας θα μπορούσε να είναι τυχαίο (έτυχε ανάμεσα στους άνδρες οδηγούς που επιλέξαμε στο δείγμα να υπάρχουν πολλοί με ατύχημα τα τελευταία χρόνια σε σχέση με τον πληθυσμό των ανδρών ή κάτι αντίστοιχο για τις γυναίκες). Όσο μεγαλύτερο είναι το μέγεθος του δείγματος και όσο πιο μεγάλη είναι η διαφορά που βρήκαμε στο δείγμα μας, τόσο λιγότερο

πιθανό είναι το εύρημα να είναι τυχαίο και με μεγαλύτερη εμπιστοσύνη μπορούμε να ισχυριστούμε ότι το εύρημα στο δείγμα μας αντανακλά μια διαφορά στον πληθυσμό.

		Τροχαίο Ατύχημα		
		Ναι	Όχι	Σύνολο
Φύλο	Άνδρες	71,9%	56,9%	
	Γυναίκες	28,1%	43,1%	
	Σύνολο	100%	100%	

Πίνακας 4: Πίνακας Συνάφειας σχετικών συχνοτήτων φύλου και της συμμετοχής σε ατύχημα ως προς τη συμμετοχή σε ατύχημα

Από τον πίνακα 4 φαίνεται ότι τόσο στο σύνολο των οδηγών που είχαν ατύχημα όσο και στο σύνολο των οδηγών που δεν είχαν ατύχημα, οι περισσότεροι οδηγοί ήταν άνδρες.

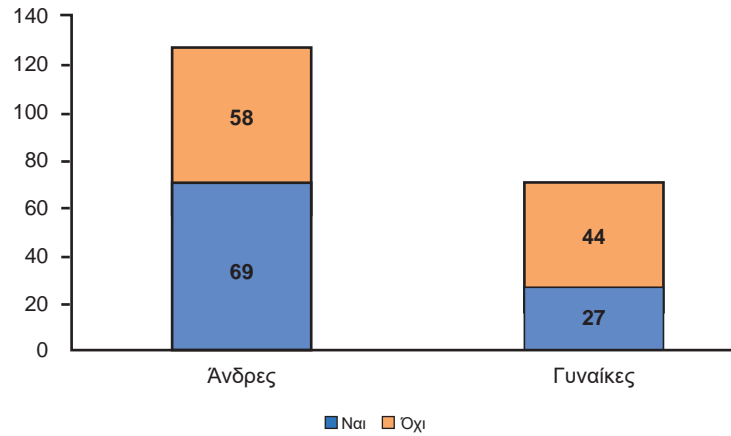
Όπως είδαμε στην περίπτωση της μίας ποιοτικής μεταβλητής οι πληροφορίες που αντλούνται από τους πίνακες συχνοτήτων ή σχετικών συχνοτήτων αναπαριστώνται με τα ραβδογράμματα και τα κυκλικά διαγράμματα. Στην περίπτωση των δύο ποιοτικών μεταβλητών, οπτικοποιούμε τα αποτελέσματα ενός πίνακα συνάφειας συχνοτήτων με το στοιβαγμένο ραβδόγραμμα (stacked bar plot) και τα αποτελέσματα ενός πίνακα συνάφειας σχετικών συχνοτήτων με το ομαδοποιημένο ραβδόγραμμα (clustered bar plot).

Το στοιβαγμένο ραβδόγραμμα συχνοτήτων (σχήμα 1) του πίνακα 2 αποτελείται από ορθογώνιες στήλες, οι βάσεις των οποίων βρίσκονται πάνω στον οριζόντιο άξονα, στον οποίο αντιστοιχούν οι κατηγορίες του ενός χαρακτηριστικού π.χ. του φύλου. Το πλήθος των ορθογώνιων στηλών είναι ίσο με το πλήθος των κατηγοριών του φύλου. Κάθε ορθογώνια στήλη διαιρείται σε τόσα τμήματα διαφορετικού χρώματος, όσες και οι κατηγορίες του άλλου χαρακτηριστικού (ατυχήματος). Συνήθως, στον οριζόντιο άξονα τοποθετούμε τις κατηγορίες της ανεξάρτητης μεταβλητής (φύλου, στο παράδειγμά μας) και στον κατακόρυφο άξονα τις κατηγορίες της εξαρτημένης μεταβλητής (ατυχήματος, στο παράδειγμά μας) ώστε να ελέγξουμε εάν το φύλο σχετίζεται με τα τροχαία ατυχήματα.

Ραβδογράμματα

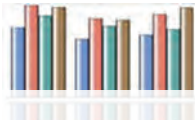
Στοιβαγμένο Ραβδόγραμμα



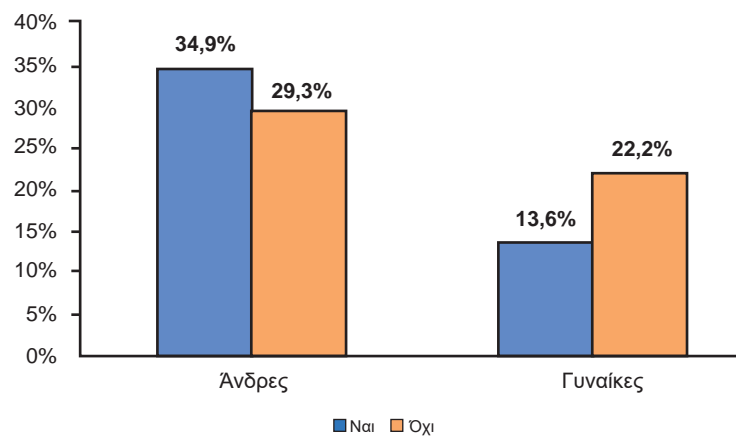


Σχήμα 1: Στοιβαγμένο ραβδόγραμμα συχνοτήτων φύλου και συμμετοχής σε ατύχημα του Πίνακα 1

Ομαδοποιημένο Ραβδόγραμμα



Το ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων αποτελείται από ορθογώνιες στήλες, οι βάσεις των οποίων βρίσκονται πάνω στον οριζόντιο άξονα στον οποίο αντιστοιχούν οι κατηγορίες του ενός χαρακτηριστικού π.χ. του φύλου. Σε κάθε κατηγορία του φύλου αντιστοιχούν τόσες ορθογώνιες στήλες διαφορετικού χρώματος όσες το πλήθος των κατηγοριών του άλλου χαρακτηριστικού π.χ. της συμμετοχής σε ατύχημα. Το ύψος κάθε στήλης είναι ίσο με την αντίστοιχη σχετική συχνότητα.

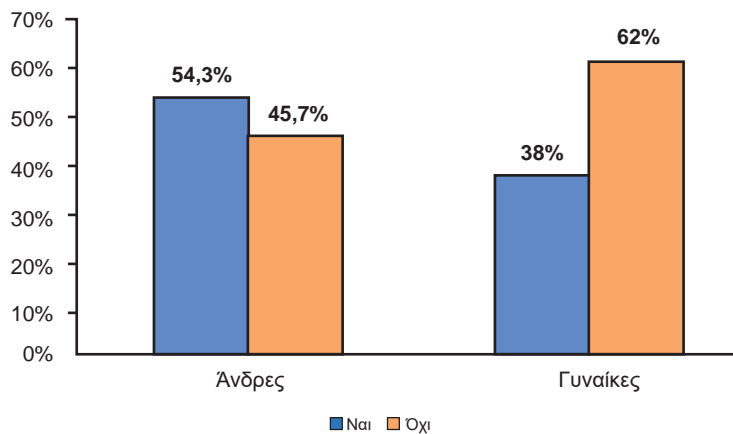


Σχήμα 2: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων φύλου και ατυχήματος του πίνακα 2

Στο σχήμα 2 φαίνονται όλα τα ποσοστά του πίνακα 2. Για παράδειγμα, εάν αθροίσετε όλα τα ποσοστά που φαίνονται στο σχήμα προκύπτει το 100%, ενώ εάν αθροίσετε τα ποσοστά των δύο ράβδων στους άνδρες οδηγού προκύπτει το 64,2% που είναι το ποσοστό τους στο δείγμα. Επίσης, παρατηρούμε ότι οι ράβδοι σε κάθε κατηγορία του φύλου δεν έχουν το ίδιο μοτίβο. Δηλαδή, στους άνδρες

οδηγούς οι περισσότεροι είχαν ατύχημα, ενώ στις γυναίκες οδηγούς οι περισσότερες δεν είχαν ατύχημα.

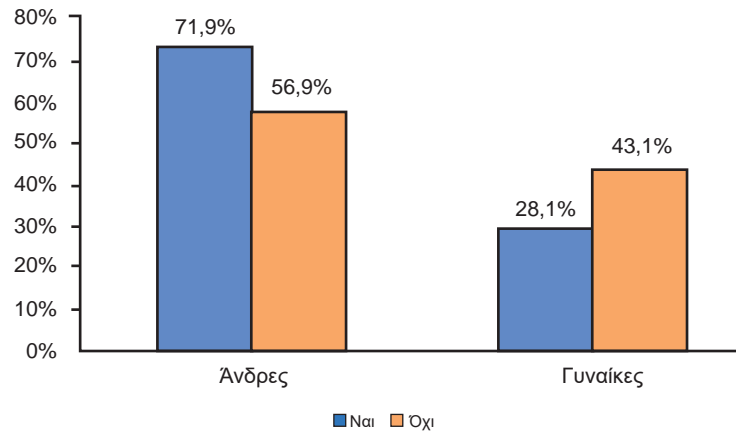
Στα σχήματα 3 και 4 φαίνονται τα ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων ως προς το φύλο και τη συμμετοχή σε ατύχημα, αντίστοιχα. Στο σχήμα 3 τα ποσοστά των ράβδων έχουν άθροισμα 100% τόσο ανάμεσα στους άνδρες όσο και ανάμεσα στις γυναίκες οδηγούς. Βλέπουμε, δηλαδή, στους άνδρες (γυναίκες) τα ποσοστά των οδηγών που είχαν/δεν είχαν ατύχημα στο σύνολο των ανδρών (γυναικών) οδηγών.



Σχήμα 3: Ομαδοποιημένο ραβδογράμμο σχετικών συχνοτήτων φύλου και της συμμετοχής σε ατύχημα ως προς το φύλο (αντ. Πίνακα 3)

Από το σχήμα 3, το μοτίβο των ράβδων σε κάθε κατηγορία του φύλου είναι διαφορετικό, οπότε φαίνεται να υπάρχει μια σχέση ανάμεσα στο φύλο και στα ατυχήματα. Εάν το μοτίβο των ράβδων σε κάθε κατηγορία του φύλου ήταν περίπου ίδιο, αυτό θα ήταν μία ένδειξη ότι οι μεταβλητές φύλο και ατύχημα τα τελευταία πέντε χρόνια δε σχετίζονται.

Στο σχήμα 4 τα ποσοστά των ομοιόμορφα χρωματισμένων ράβδων έχουν άθροισμα 100%. Βλέπουμε, δηλαδή, τα ποσοστά των μπλε ράβδων να αντιστοιχούν στους άνδρες και γυναίκες οδηγούς που είχαν ατύχημα στο σύνολο των οδηγών που είχαν ατύχημα, ενώ τα ποσοστά των ροζ ράβδων αντιστοιχούν στους άνδρες και γυναίκες οδηγούς που δεν είχαν ατύχημα στο σύνολο των οδηγών που δεν είχαν ατύχημα. Από αυτό το σχήμα φαίνεται πιο εύκολα ότι το φύλο και η συμμετοχή σε ατύχημα σχετίζονται, αφού το ύψος των ράβδων σε κάθε κατηγορία του φύλου δεν είναι περίπου το ίδιο.



Σχήμα 4: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων φύλου και συμμετοχής σε ατύχημα ως προς τη συμμετοχή σε ατύχημα (αντ. Πίνακα 4)

Σχέση Αιτιότητας

Με ανάλογο τρόπο προκύπτουν στοιβαγμένα και ομαδοποιημένα ραβδογράμματα, εάν στον οριζόντιο άξονα τοποθετήσουμε τις κατηγορίες του ατυχήματος. Στο σημείο αυτό, καλό είναι να γίνει η παρακάτω διευκρίνιση. Το γεγονός ότι δύο μεταβλητές φαίνεται να σχετίζονται, όπως στο παράδειγμα μας, είναι λάθος να ερμηνεύεται με όρους αιτίου και αποτελέσματος, δηλαδή η ανεξάρτητη μεταβλητή να είναι η αιτία που προκλήθηκε η εξαρτημένη μεταβλητή (το αποτέλεσμα). Στο παράδειγμά μας δε θα ήταν σωστό για μια ασφαλιστική εταιρεία να ερμηνεύσει το αποτέλεσμα ως ότι το φύλο ενός οδηγού είναι υπεύθυνο για την πρόκληση ατυχημάτων. Με δεδομένο, λοιπόν, ότι δύο μεταβλητές A και B σχετίζονται, αυτό μπορεί να σημαίνει ότι:

- η μεταβλητή A είναι η αιτία για τη μεταβλητή B
- η μεταβλητή B είναι η αιτία για τη μεταβλητή A
- υπάρχει ένας τρίτος (συγχυτικός) παράγοντας ο οποίος να είναι η αιτία τόσο για το A, όσο και για το B.
- είναι απλά μια σύμπτωση, διότι απλά ένα τυχαίο γεγονός συνέβη στο δείγμα μας, ενώ στην πραγματικότητα, δηλαδή στον πληθυσμό, οι δύο μεταβλητές είναι ασυσχέτιστες.

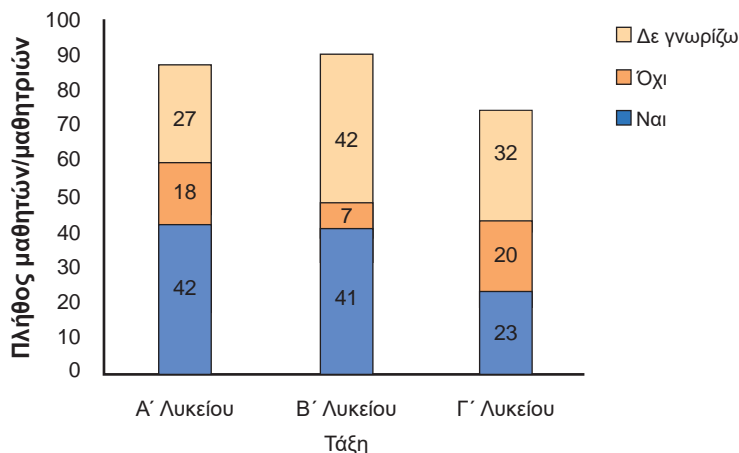
Για παράδειγμα, μια έρευνα έδειξε ότι τα μικρά παιδιά που κοιμούνται με αναμμένο το φως είναι πολύ πιθανότερο να εμφανίσουν μυωπία αργότερα. Μπορούμε άραγε να συμπεράνουμε ότι ο ύπνος με αναμμένο το φως προκαλεί μυωπία; Μεταγενέστερη έρευνα έδειξε ότι υπάρχει μια ισχυρή σχέση μεταξύ της γονικής μυωπίας και της εμφάνισης μυωπίας στα παιδιά τους, σημειώνοντας επίσης ότι οι μυωπικοί γονείς είναι πιο πιθανό να αφήσουν ένα φως στην κρεβατοκάμαρα των παιδιών τους. Ο συγχυτικός παράγοντας εδώ είναι η γονική μυωπία που επηρεάζει τόσο τη μεταβλητή «αναμμένο φως» όσο και τη μεταβλητή «μυωπία παιδιών».

Τέλος, αξίζει να αναφέρουμε το επόμενο παράδειγμα. Το αποτέλεσμα του τελευταίου τοπικού παιχνιδιού ποδοσφαίρου από την ομάδα των Washington Redskins πριν από τις προεδρικές εκλογές των Η.Π.Α. προέβλεπε το αποτέλεσμα της επόμενης προεδρικής εκλογής σε κάθε εκλογή από το 1940 έως το 2000. Όταν κέρδιζαν οι Redskins, το κόμμα του τότε προέδρου διατηρούσε την προεδρία, ενώ όταν οι Redskins έχαναν, κέρδιζε το κόμμα της αντιπολίτευσης. Φυσικά αυτό το εύρημα είναι μάλλον τυχαίο, αφού δεν είναι λογικό να υπάρχει σχέση ανάμεσα στα αποτελέσματα των τοπικών ποδοσφαιρικών αγώνων και στο αποτέλεσμα των προεδρικών εκλογών.

Εφαρμογές

Εφαρμογή 1

Στα πλαίσια μιας έρευνας που διενεργήθηκε σε 252 μαθητές/μαθήτριες ενός Λυκείου σχετικά με τους Παραολυμπιακούς αγώνες τέθηκε η ερώτηση: «Υπάρχει διαφορά μεταξύ των Special Olympics (SO) και των Παραολυμπιακών Αγώνων (ΠΑ);» Οι απαντήσεις των μαθητών/τριών του σχολείου αναπαριστώνται στο επόμενο στοιβαγμένο ραβδόγραμμα:



Σχήμα 5: Στοιβαγμένο ραβδόγραμμα συχνοτήτων για το ερώτημα «Υπάρχει διαφορά μεταξύ των Special Olympics και των Παραολυμπιακών Αγώνων (ΠΑ);»



- Ποιες ήταν οι πιθανές απαντήσεις των μαθητών/τριών του Λυκείου;
- Αθροίστε τις επιμέρους τιμές των μπλε ράβδων. Ποιο είναι το άθροισμά τους; Να ερμηνεύσετε το αποτέλεσμα. Να βρείτε τα ποσοστά των μαθητών/τριών του Λυκείου που έδωσαν κάθε διαφορετική απάντηση.
- Αθροίστε τις επιμέρους τιμές των ράβδων στη Β' τάξη. Ποιο είναι το άθροισμά τους; Να ερμηνεύσετε το αποτέλεσμα. Να εκφράσετε τις επιμέρους τιμές των

ράβδων της Β΄ τάξης ως ποσοστά επί του συνόλου των μαθητών της Β΄ τάξης, ερμηνεύοντάς τα.

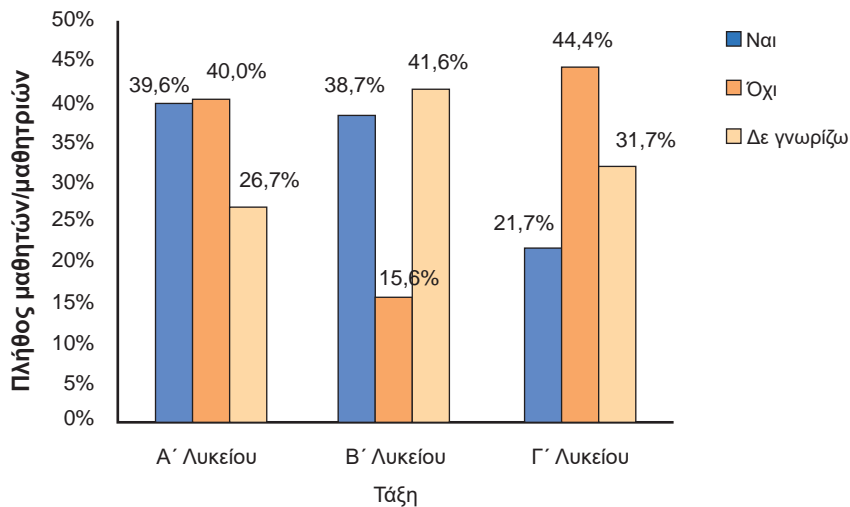
- δ) Στο σύνολο των μαθητών/τριών του Λυκείου, πόσοι μαθητές/μαθήτριες από κάθε τάξη ήταν σίγουροι ότι δεν υπάρχει διαφορά μεταξύ (SO) και (ΠΑ); Τι ποσοστό ήταν από κάθε τάξη;
- ε) Στο σύνολο των μαθητών/τριών του Λυκείου, ποιας τάξης οι μαθητές/μαθήτριες φαίνεται να είναι καλύτερα ενημερωμένοι ως προς το γεγονός ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) και σε τι ποσοστό, επί του συνόλου των μαθητών/τριών του Λυκείου;

Λύση

- α) Οι μαθητές/μαθήτριες απάντησαν «ναι», εάν πίστευαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), «όχι», εάν πίστευαν ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) και «δε γνωρίζω», εάν δε γνώριζαν την απάντηση.
- β) Αθροίζοντας τις τιμές των μπλε ράβδων έχουμε 106 μαθητές/μαθήτριες (42 από την Α΄, 41 από τη Β΄ και 23 από τη Γ΄) που απάντησαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ). «Ναι» απάντησε το $\frac{106}{252} = 42,1\%$, «όχι» το $\frac{45}{252} = 17,8\%$ και «δε γνωρίζω» το $\frac{101}{252} = 40,1\%$ των μαθητών/τριών.
- γ) Αθροίζοντας τις επιμέρους τιμές στη στήλη της Β΄ τάξης έχουμε 90 μαθητές/μαθήτριες της Β΄ Λυκείου εκ των οποίων οι 41 απάντησαν ότι υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), οι 7 απάντησαν ότι δεν υπάρχει διαφορά και οι 42 δε γνώριζαν την απάντηση. Τα ποσοστά των επιμέρους τιμών είναι: «Ναι» απάντησε περίπου το $\frac{41}{90} = 45,5\%$, «όχι» απάντησε περίπου το $\frac{7}{90} = 7,8\%$ και «δε γνωρίζω» απάντησε περίπου το $\frac{42}{90} = 46,7\%$ των μαθητών/τριών της Β΄ Λυκείου.
- δ) Από τους 252 μαθητές/μαθήτριες του Λυκείου, οι 18 από την Α΄ τάξη, οι 7 από τη Β΄ τάξη και οι 20 από τη Γ΄ τάξη απάντησαν «όχι». Τα ποσοστά είναι 7,1%, 2,8% και 7,9%, αντίστοιχα.
- ε) Οι μαθητές/μαθήτριες που είναι καλύτερα ενημερωμένοι σύμφωνα με το γράφημα είναι οι 42 της Α΄ τάξης του Λυκείου. Το αντίστοιχο ποσοστό ως προς το σύνολο των μαθητών/μαθητριών του σχολείου είναι 16,7%. Ωστόσο και οι μαθητές/μαθήτριες της Β΄ τάξης θα μπορούσαμε να πούμε ότι είναι εξίσου καλά ενημερωμένοι.

Εφαρμογή 2

Στα πλαίσια της ίδιας έρευνας για την ίδια ερώτηση κατασκευάζουμε το παρακάτω ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς τις απαντήσεις του ερωτήματος.



Σχήμα 5: Ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς τις απαντήσεις του ερωτήματος «Υπάρχει διαφορά μεταξύ των Special Olympics και των Παραολυμπιακών Αγώνων (ΠΑ);»

- α) Από τους μαθητές/μαθήτριες του Λυκείου που απάντησαν Ναι στην ερώτηση, ποιο είναι το ποσοστό των μαθητών/τριών που αντιστοιχεί σε κάθε τάξη;
- β) Από τους μαθητές/μαθήτριες του Λυκείου που απάντησαν Όχι στην ερώτηση, ποιο είναι το ποσοστό των μαθητών/τριών που αντιστοιχεί στην τάξη της Β Λυκείου και ποιο το ποσοστό που αντιστοιχεί στην τάξη της Γ΄ Λυκείου;
- γ) Συγκρίνοντας τα ύψη των ράβδων του γραφήματος, μπορείτε να πείτε εάν τελικά οι απαντήσεις των μαθητών/τριών στο ερώτημα σχετίζονται με την τάξη στην οποία βρίσκονται;

Λύση

- α) Το ποσοστό των μαθητών/μαθητριών της Α΄ τάξης του Λυκείου που απάντησαν Ναι στην ερώτηση είναι 39,6%, ενώ για τις άλλες δύο τάξεις είναι 38,7% και 21,7%, αντίστοιχα.
- β) Το ποσοστό των μαθητών/μαθητριών της Β΄ και Γ΄ τάξης του Λυκείου που απάντησαν Όχι στην ερώτηση είναι 15,6% και 44,4%, αντίστοιχα. Παρατη-

ρούμε ότι το ποσοστό των μαθητών/μαθητριών της Γ΄ είναι πολύ μεγαλύτερο συγκριτικά με αυτό της Β΄. Αξιοσημείωτο είναι ότι οι μαθητές/μαθήτριες ήταν σίγουροι ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ), κάτι το οποίο δεν ισχύει.

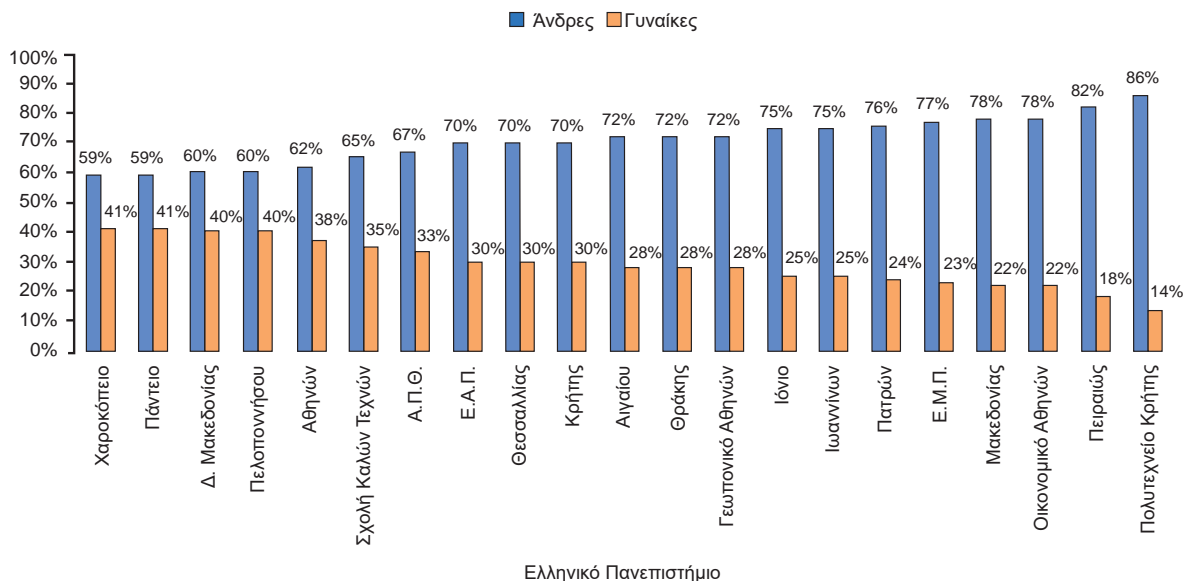
- γ) Από το σχήμα 5 βλέπουμε ότι οι ράβδοι σε κάθε τάξη δεν έχουν περίπου το ίδιο ύψος, γεγονός που μας υποδηλώνει ότι η τάξη και οι απαντήσεις των μαθητών σχετίζονται. Εάν όλες οι ράβδοι είχαν περίπου το ίδιο ύψος τότε ο παράγοντας της τάξης δε θα έπαιζε κανένα ρόλο στη διαμόρφωση των απαντήσεων των μαθητών/μαθητριών. Η καλύτερα ενημερωμένη τάξη είναι η Α΄ Λυκείου και η λιγότερο καλά ενημερωμένη είναι η Γ΄ τάξη. Αξίζει να σημειωθεί ότι οι μαθητές/μαθήτριες της Α΄ και της Γ΄ τάξης πίστευαν λανθασμένα ότι δεν υπάρχει διαφορά μεταξύ των (SO) και των (ΠΑ) σε πολύ μεγαλύτερο βαθμό από τη Β΄. Τέλος, να σημειώσουμε ότι είναι αρκετά μεγάλα τα ποσοστά της απάντησης «δε γνωρίζω», γεγονός που υποδεικνύει την ελλιπή ενημέρωση των μαθητών/μαθητριών σχετικά με αυτές τις σημαντικές διοργανώσεις των Special Olympics και των Παραολυμπιακών Αγώνων.

Ασκήσεις - Προβλήματα - Δραστηριότητες

- 1) Για ποια από τα επόμενα ζεύγη μεταβλητών θα κατασκευάζατε πίνακα συνάφειας; Για εκείνα που θα επιλέξετε, να προσδιορίσετε την ανεξάρτητη και την εξαρτημένη μεταβλητή για την έρευνά σας.
 - α) φύλο και επίπεδο εκπαίδευσης
 - β) φύλο και χρήση διαδικτύου (σε ώρες/ημέρα)
 - γ) βαθμός στην Άλγεβρα και ώρες μελέτης (ανά εβδομάδα)
 - δ) ικανοποίηση από το σχολείο και φύλο
 - ε) ευτυχία στον γάμο και πίστη στα θαύματα
- 2) Σε μια έρευνα του τρόπου ζωής και των καθημερινών συνθηκών στη χώρα μας συλλέχθηκε ένα δείγμα 833 ατόμων απ' όλη την επικράτεια, προκειμένου να εξετάσουμε εάν υπάρχει σχέση μεταξύ του επιπέδου εκπαίδευσης (δευτεροβάθμια, αν έχουν ολοκληρώσει το Λύκειο και τριτοβάθμια, αν έχουν ολοκληρώσει τις σπουδές τους σε ΑΕΙ) και της συστηματικής σωματικής άσκησης τους (ναί: τουλάχιστον 2 φορές την εβδομάδα και όχι: το πολύ μια φορά την εβδομάδα) σε όλη τη διάρκεια ενός έτους. Τα δεδομένα οργανώθηκαν στον παρακάτω πίνακα συνάφειας:

		Άσκηση		
		Ναι	Όχι	Σύνολο
Εκπαίδευση	Δευτεροβάθμια	366	75	
	Τριτοβάθμια	303	89	
	Σύνολο			

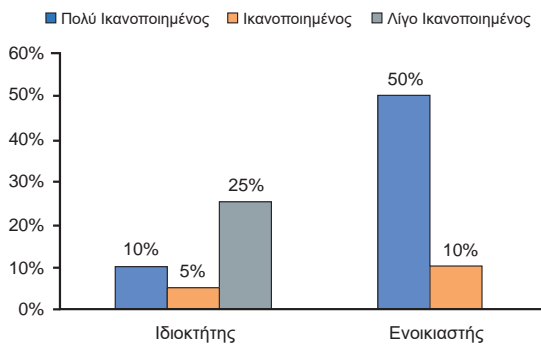
- α) Να συμπληρώσετε στον παραπάνω πίνακα τα κενά κελιά και να κατασκευάσετε το στοιβαγμένο ραβδόγραμμα των συχνοτήτων του πίνακα συνάφειας.
 - β) Να κατασκευάσετε τον πίνακα συνάφειας σχετικών συχνοτήτων ως προς το επίπεδο εκπαίδευσης και το αντίστοιχο ομαδοποιημένο ραβδόγραμμα του.
 - γ) Μπορείτε να υποστηρίξετε την άποψη ότι το επίπεδο εκπαίδευσης επηρεάζει το αν κάποιο άτομο ασκείται συστηματικά ή όχι;
- 3) Το επόμενο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων δίνει τα ευρήματα μιας μελέτης που έγινε το 2016 στο Οικονομικό Πανεπιστήμιο Αθηνών για τη θέση των γυναικών στα Ελληνικά Πανεπιστήμια, η οποία δημοσιεύθηκε σε κυριακάτικη εφημερίδα. Πιο συγκεκριμένα, δίνονται τα ποσοστά ανδρών και γυναικών μόνιμων διδασκόντων στα Πανεπιστήμια.



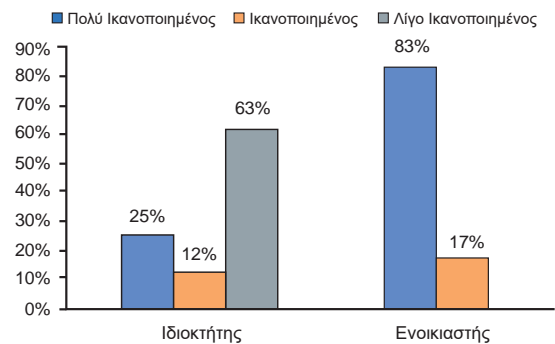
- α) Σε ποια Πανεπιστήμια παρατηρείται το μεγαλύτερο ποσοστό γυναικών και σε ποια το χαμηλότερο;

- β)** Τη συγκεκριμένη χρονιά δίδαξαν συνολικά 8.483 μόνιμοι διδάσκοντες εκ των οποίων οι 5.894 ήταν άνδρες και οι 2.589 ήταν γυναίκες. Να βρείτε σε ποια πανεπιστήμια το ποσοστό των γυναικών ξεπέρασε το γενικό μέσο ποσοστό τους;
- γ)** Ποιος πιστεύετε ότι μπορεί να ήταν ο τίτλος του άρθρου και για ποιον λόγο μια εφημερίδα το ανέδειξε ως ένα από τα κύρια κοινωνικά άρθρα της;
- 4)** Στη Μυτιλήνη, ο δήμος εγκατέστησε ένα καινοτόμο πρόγραμμα ηλεκτρονικών υπηρεσιών του δήμου από το οποίο εξυπηρετούνται οι κάτοικοι του νησιού, ιδιοκτήτες και ενοικιαστές σπιτιών. Η ικανοποίηση των κατοίκων του νησιού από τις ηλεκτρονικές υπηρεσίες του δήμου φαίνονται από τα παρακάτω ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων (α) και (β).
- α)** Ποιο από τα δύο ομαδοποιημένα ραβδογράμματα σχετικών συχνοτήτων αντιστοιχεί στο σύνολο των παρατηρήσεων του δείγματος; Τι δείχνει το άλλο ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων;
- β)** Ποιο ποσοστό του δείγματος είναι ιδιοκτήτες και ποιο ενοικιαστές σπιτιών;
- γ)** Ποιο ποσοστό του δείγματος είναι πολύ ικανοποιημένοι από τις ηλεκτρονικές υπηρεσίες του δήμου;
- δ)** Τι εκφράζουν τα ποσοστά 63% και 83% στο σχήμα (β);
- ε)** Αν το μέγεθος του δείγματος είναι 200 κάτοικοι, να βρείτε πόσοι είναι οι πολύ ικανοποιημένοι ιδιοκτήτες και πόσοι οι ικανοποιημένοι ενοικιαστές σπιτιών στο νησί.

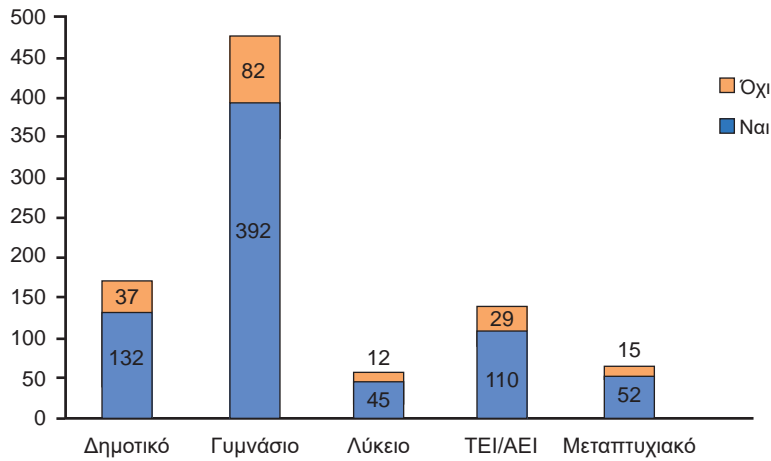
Σχήμα (α)



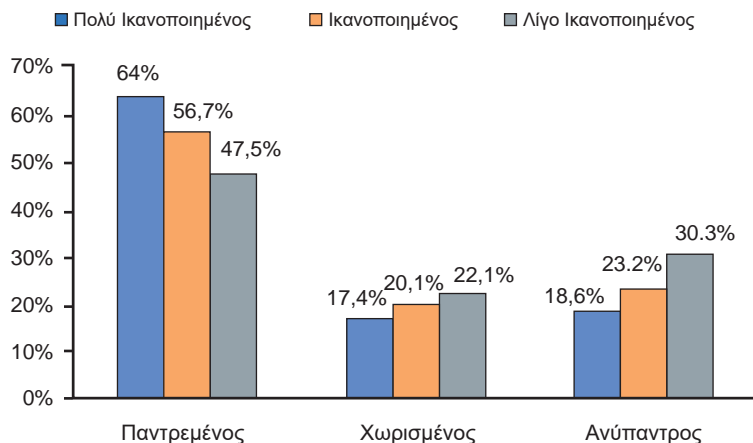
Σχήμα (β)



- 5)** Πρόσφατα πραγματοποιήθηκε μια έρευνα προκειμένου να ερευνηθεί εάν υπάρχει σχέση μεταξύ της πίστης του ανθρώπου στα θαύματα και του επιπέδου εκπαίδευσης που έχει αποκτήσει. Τα αποτελέσματα της έρευνας συνοψίζονται στο επόμενο στοιβαγμένο ραβδόγραμμα συχνοτήτων.



- α) Τι ποσοστό αποφοίτων Γυμνασίου, Λυκείου και ΑΕΙ/ΤΕΙ έχουμε στο δείγμα μας;
 - β) Τι ποσοστό αποφοίτων Λυκείου φαίνεται να πιστεύουν στα θαύματα και τι ποσοστό αποφοίτων ΤΕΙ/ΑΕΙ φαίνεται ότι δεν πιστεύουν στα θαύματα;
 - γ) Να μετατρέψετε το παραπάνω στοιβαγμένο ραβδόγραμμα συχνοτήτων σε ομαδοποιημένο ραβδόγραμμα σχετικών συχνοτήτων ως προς την πίστη στα θαύματα.
 - δ) Με βάση τα αποτελέσματα της συγκεκριμένης έρευνας φαίνεται να υπάρχει κάποια σχέση μεταξύ του επιπέδου εκπαίδευσης και της πίστης στα θαύματα;
- 6) Μια κυριακάτικη εφημερίδα περιέχει ένα άρθρο σχετικά με την ικανοποίηση των ανθρώπων από την εργασία τους και την οικογενειακή τους κατάσταση μαζί με το παρακάτω γράφημα. Οι συντάκτες του άρθρου καταλήγουν στο συμπέρασμα ότι ο γάμος κάνει τους ανθρώπους περισσότερο ικανοποιημένους με την εργασία τους, δεδομένου ότι το 64% των πολύ ικανοποιημένων ανθρώπων είναι παντρεμένοι, ενώ μόνο το 18,6% δεν έχουν παντρευτεί ποτέ.



- α) Συμφωνείτε με το συμπέρασμα του συντάκτη του άρθρου;
- β) Συμφωνείτε με την αιτιολόγηση του συντάκτη του άρθρου;

Πρόσθετο Υλικό

Θέματα για διερεύνηση



- 1) Το φθινόπωρο του 1973 παρατηρήθηκε το ακόλουθο παράδοξο κατά την εισαγωγή μεταπτυχιακών φοιτητών στο Πανεπιστήμιο του Berkeley στην Καλιφόρνια. Τα στοιχεία της έρευνας έδειξαν ότι οι άνδρες που υπέβαλλαν αίτηση ήταν πιο πιθανό από ό,τι οι γυναίκες να γίνουν δεκτοί. Στον πίνακα (α) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές τη χρονιά του 1973. Στον πίνακα (β) δίνονται τα στοιχεία αιτούντων και εισακτέων ανδρών και γυναικών σε έξι διακεκριμένα τμήματα του Πανεπιστημίου του Berkeley για μεταπτυχιακές σπουδές τη χρονιά του 1973.

	Αιτούντες	Εισακτέοι	Ποσοστό εισακτέων
Άνδρες	8442	3714
Γυναίκες	4321	1512

Πίνακας (α): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές στο Πανεπιστήμιο του Berkeley

- α) Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (α). Τι διαπιστώνετε, συγκρίνοντας τα ποσοστά των ανδρών και των γυναικών εισακτέων;

	Άνδρες			Γυναίκες		
	Αιτούντες	Εισακτέοι	Ποσοστό εισακτέων	Αιτούντες	Εισακτέοι	Ποσοστό εισακτέων
Τμήμα Α	825	512	108	89
Τμήμα Β	560	353	25	17
Τμήμα Γ	325	120	593	202
Τμήμα Δ	417	138	375	131
Τμήμα Ε	191	54	393	94
Τμήμα Ζ	373	23	341	24

Πίνακας (β): Πίνακας αιτούντων και εισακτέων ανδρών και γυναικών για μεταπτυχιακές σπουδές σε έξι διακεκριμένα τμήματα Πανεπιστημίου του Berkeley

- β) Συμπληρώστε τα ποσοστά των εισακτέων του πίνακα (β) και συγκρίνετε:
- (i) τα ποσοστά των ανδρών εισακτέων στα τμήματα Γ και Ε με τα αντίστοιχα ποσοστά των γυναικών,

(ii) τα ποσοστά των εισακτέων γυναικών στα τμήματα Α, Β, Δ και Ζ με τα αντίστοιχα ποσοστά των ανδρών.

γ) Τι διαπιστώνετε, συγκρίνοντας τα παραπάνω ποσοστά σε κάθε τμήμα του Πανεπιστημίου με τα ποσοστά του ερωτήματος α; Υπάρχει κάποια σύγχυση;

δ) Σε τι είδους τμήματα ως προς την ανταγωνιστικότητα και το ποσοστό των εισακτέων φαίνεται να απευθύνονται οι γυναίκες και σε τι είδους τμήματα οι άνδρες;

4) Το σχολικό έτος 2017-2018 η ΕΛ. ΣΤΑΤ. (Ελληνική Στατιστική Αρχή) διοργάνωσε για πρώτη φορά έναν πανελλήνιο διαγωνισμό Στατιστικής. Στη Β΄ φάση οι διαγωνιζόμενοι κλήθηκαν να επεξεργαστούν τα δεδομένα μιας στατιστικής έρευνας χρήσης τεχνολογιών πληροφορικής και επικοινωνίας από 3351 πολίτες το πρώτο τρίμηνο του 2017. Ένα από τα αντικείμενα μελέτης ήταν η χρήση υπηρεσιών διακυβέρνησης από τους πολίτες, ανάλογα με την ηλικία τους, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία τους. Πιο συγκεκριμένα, οι πολίτες ρωτήθηκαν εάν λαμβάνουν πληροφορίες από ιστοσελίδες δημόσιων υπηρεσιών. Οι πιθανές απαντήσεις ήταν δύο: Ναι, εάν έχουν λάβει πληροφορίες από ιστοσελίδες δημόσιων υπηρεσιών και Όχι, εάν δεν έχουν κάνει τη συγκεκριμένη ενέργεια. Τα αποτελέσματα συνοψίζονται στους επόμενους πίνακες συνάφειας.



	Ναι	Όχι	Σύνολο
Ανδρας	1094	573	1667
Γυναίκα	1056	628	1684
Σύνολο	2150	1201	3351

Πίνακας 1: Πίνακας Συνάφειας συχνότητας φύλου και χρήσης ηλ. Διακυβέρνησης

	Ναι	Όχι	Σύνολο
16-34	616	293	909
35-54	1109	529	1638
>54	425	379	804
Σύνολο	2150	1201	3351

Πίνακας 2: Πίνακας Συνάφειας συχνότητας ηλικίας και χρήσης ηλ. Διακυβέρνησης

	Ναι	Όχι	Σύνολο
Έως Γυμνάσιο	175	360	535
Έως Λύκειο	829	622	1451
Έως Ανώτατη Εκπαίδευση	1146	219	1365
Σύνολο	2150	1201	3351

Πίνακας 3: Πίνακας Συνάφειας συχνοτήτων επιπέδου εκπαίδευσης και χρήσης ηλ. Διακυβέρνησης

¹Νήσοι Αιγαίου – Κρήτη:
Βόρειο Αιγαίο, Νότιο Αιγαίο και Κρήτη

²Βόρεια Ελλάδα:
Ανατολική Μακεδονία και Θράκη, Κεντρική Μακεδονία, Δυτική Μακεδονία και Ήπειρος

³Κεντρική Ελλάδα:
Θεσσαλία, Νησιά Ιονίου, Δυτική Ελλάδα, Λοιπή Στερεά Ελλάδα και Πελοπόννησος

	Ναι	Όχι	Σύνολο
Αττική	1002	454	1456
Νήσοι Αιγαίου – Κρήτη¹	185	167	352
Βόρεια Ελλάδα²	556	313	869
Κεντρική Ελλάδα³	407	267	674
Σύνολο	2150	1201	3351

Πίνακας 4: Πίνακας Συνάφειας συχνοτήτων γεωγραφικής περιοχής και χρήσης ηλ. Διακυβέρνησης

	Ναι	Όχι	Σύνολο
Μισθωτός - Ελ.Επ.	1329	499	1828
Άνεργος	309	199	508
Μαθητής/Φοιτητής	165	105	270
Άλλο⁴	347	398	745
Σύνολο	2150	1201	3351

Πίνακας 5: Πίνακας Συνάφειας συχνοτήτων γεωγραφικής περιοχής και χρήσης ηλ. Διακυβέρνησης (⁴Άλλο: Συνταξιούχος, Νοικοκυρά, Στρατιώτης)

Να γράψετε μια ερευνητική έκθεση με τις πληροφορίες που μπορείτε να αντλήσετε από τους παραπάνω πίνακες συνάφειας και την επεξεργασία τους (κατασκευή πινάκων συνάφειας σχετικών συχνοτήτων, κατασκευή κατάλληλων γραφημάτων), απαντώντας στα επόμενα ερωτήματα:

- α) Ποιες είναι οι μεταβλητές της έρευνας και ποιο το είδος τους;
- β) Να κατασκευάσετε κατάλληλο γράφημα για κάθε μεταβλητή χωριστά και να ερμηνεύσετε τα αποτελέσματα.
- γ) Με κατάλληλα γραφήματα, να περιγράψετε τη σχέση της χρήσης υπη-

ρεσιών ηλεκτρονικής διακυβέρνησης, ανάλογα με το φύλο, την ηλικία, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία των ερωτηθέντων.

- δ) Μπορείτε να πείτε εάν η χρήση υπηρεσιών ηλεκτρονικής διακυβέρνησης συσχετίζεται με κάποιους από τους παράγοντες, όπως το φύλο, την ηλικία, το επίπεδο εκπαίδευσης, τη γεωγραφική περιοχή και την κύρια ασχολία.
- 3) Αποφασίζετε με τους φίλους σου να κάνετε μια έρευνα για ένα θέμα που σας απασχολεί ή συζητήσατε σε κάποιο άλλο μάθημα και θέλετε να δείτε τι γνώμη έχουν οι συμμαθητές σας. Συντάξτε ένα ερωτηματολόγιο με ερωτήσεις που αφορούν τα βασικά στοιχεία των συμμαθητών σας (φύλο, τάξη, ομάδα προσανατολισμού) και ερωτήσεις που αφορούν το θέμα που επιλέξατε. Φροντίστε οι ερωτήσεις να είναι σαφώς διατυπωμένες και οι πιθανές απαντήσεις να μην επικαλύπτονται. Χρησιμοποιώντας αυτά που μάθατε στη συγκεκριμένη παράγραφο, γράψτε μια σύντομη ερευνητική έκθεση παρουσιάζοντας τα αποτελέσματά σας. Για παράδειγμα, ο εθισμός στο διαδίκτυο είναι μια νέα μορφή εξάρτησης και αναφέρεται στην καταναγκαστική και υπερβολική χρήση του διαδικτύου και τον εκνευρισμό που προκαλείται από τη στέρηση της.

Αναζητήστε συσχετίσεις ανάμεσα στον χρόνο χρήσης του διαδικτύου (καθόλου, μέχρι μία ώρα την ημέρα, από 1 έως και 3 ώρες την ημέρα, πάνω από 3 ώρες την ημέρα) με διάφορους παράγοντες που πιστεύετε ότι επηρεάζουν την ενασχόληση με το διαδίκτυο.

