

Μάθημα 2

§ 1.5 Πηγές Συλλογής Στατιστικών Στοιχείων

Η στατιστική ανάλυση ενός φαινομένου βασίζεται σε δεδομένα τα οποία προέρχονται από πρωτογενείς ή δευτερογενείς πηγές στατιστικής πληροφόρησης. Τα πρωτογενή δεδομένα (primary data) προέρχονται από πρωτογενή πηγή και συλλέγονται υπό την ευθύνη του ερευνητή που καλείται να δώσει απάντηση σε ερωτήματα όπως "Κατά πόσο είναι πιθανό ένας καπνιστής να αναπτύξει καρκίνο του πνεύματος;".

Η συλλογή των πρωτογενών δεδομένων επιτυγχάνεται με τη διανομή και τη διακίνηση ερωτηματολογίων (questionnaires) ή συμπλήρωση των οποίων γίνεται από ειδικά εκπαιδευμένους ερευνητές.

Τα δευτερογενή δεδομένα (secondary data) αφορούν στοιχεία που συγκεντρώνονται από υπηρεσίες στατιστικής πληροφόρησης (π.χ. Τράπεζα της Ελλάδος) **μέσα** από συνεχείς προγραμματισμένες και τυποποιημένες διαδικασίες συλλογής και δημοσίευσης δεδομένων. Η τεχνική συλλογής πρωτογενών στοιχείων είναι γνωστή ως έρευνα πεδίου (field research). Ο όρος έρευνα γραφείου (desk research) αφορά την επεξεργασία και την ανάλυση είτε πρωτογενών είτε δευτερογενών στατιστικών στοιχείων.

Η Εθνική Στατιστική Υπηρεσία κάθε χώρας διενεργεί σε τακτά χρονικά διαστήματα δειγματοληπτικές έρευνες για να πάρει πληροφορίες για την απασχόληση και την ανεργία στην κάθε χώρα. Πολλές φορές ανάλογα με τα αποτελέσματα διαμορφώνεται και η κυβερνητική πολιτική

στα θέματα αυτά.

Η Εθνική Στατιστική Υπηρεσία της Ελλάδος είναι η κύρια επίσημη υπηρεσία υπεύθυνη για τη συλλογή και τη δημοσίευση δεδομένων που αφορούν τον πληθυσμό της χώρας και τους περισσότερους τομείς της Εθνικής Οικονομίας. Επιπλέον, η ΕΣΥΕ δημοσιεύει πλήθος στοιχείων που αφορούν μακροοικονομικά δάνεια και γεωργικές επεξεργασίες και χρηματοδοτήσεις, νομισματικά και τραπεζο-οικονομικά μεγέθη κλπ. Άλλοι οργανισμοί όπως το Κέντρο Προγραμματισμού και Οικονομικών Ερευνών (ΚΕΠΕ) ή ο Οργανισμός Απασχόλησης Εργατικού Δυναμικού (ΟΑΕΔ) αλλά και ιδιωτικές εταιρείες συλλέγουν και παρέχουν πληθώρα πληροφοριών για διάφορους τομείς του δημοσίου και του ιδιωτικού τομέα. Συχνικότερά δεσφνέται στοιχεία δημοσιεύονται σε ειδικούς τόμους από τις στατιστικές υπηρεσίες της Ευρωπαϊκής Κοινότητας και του Οργανισμού Ηνωμένων Εθνών.

ΚΕΦΑΛΑΙΟ 2 ΜΕΤΑΒΛΗΤΕΣ

§ 2.1 Τύποι Δεδομένων

Η φύση των στοιχείων που έχουμε στη διάθεσή μας προσδιορίζει το είδος των γραφικών και υπολογιστικών τεχνικών που θα πρέπει να χρησιμοποιηθούν κατά τη στατιστική ανάλυση.

Ας δούμε ένα παράδειγμα. Για τις ανάγκες μίας έρευνας με σκοπό τη μελέτη των καπνιστικών συνηθειών του Ελληνικού πληθυσμού κατασκευάστηκε ειδικός ερωτηματολόγιος. Ο υπό μελέτη πληθυσμός ήταν άτομα ηλικίας 18-60 ετών κάτοικοι αστικών περιοχών της χώρας.

Δύο από τις ερωτήσεις που απευθύνθηκαν στους ερωτώμενους ήταν:

- (α) Τι ηλικία έχετε;
- (β) Είστε καπνιστής;

Μια πιθανή απάντηση στην Ερώτηση (α) είναι π.χ. 27 ετών. Οι πληροφορίες που αντλούνται από τη συγκεκριμένη ερώτηση είναι αριθμητικές δηλαδή ποσοτικές. Τέτοιου είδους στατιστικά στοιχεία καλούνται αριθμητικά ή ποσοτικά δεδομένα (numeric, quantitative data).

Μία πιθανή απάντηση στην Ερώτηση (β) είναι ΝΑΙ ή ΟΧΙ. Η πληροφορία που παρέχεται από το ερωτώμενο άτομο είναι ποιοτικής μορφής και περιγράφεται από μία λέξη που εκφράζει κατάσταση ή ιδιότητα.

Τέτοιου είδους στοιχεία καλούνται ονομαστικά ή ποιοτικά δεδομένα (nominal, qualitative data).

Το ύψος, το βάρος, η ηλικία, το εισόδημα, η θερμοκρασία οι τιμές προϊόντων, το εμβαδόν ενός οικοπέδου ή διαμερίσματος αποτελούν παραδείγματα αριθμητικών δεδομένων.

Το φύλο, η οικογενειακή κατάσταση, το επίπεδο εκπαίδευσης, το θρήσκευμα, οι πολιτικές προτιμήσεις αποτελούν παραδείγματα ονομαστικών (ποιοτικών) δεδομένων.

Για τους σκοπούς της ίδιας έρευνας υποβλήθηκαν επίσης οι εξής ερωτήσεις: (α) Τι ύψος έχετε; (β) Πόσα παιδιά έχετε;

Οι απαντήσεις 5 ατόμων σχετικά με το ύψος τους είναι: 1,70, 1,82, 1,83, 1,68, 1,71

Οι απαντήσεις των πέντε ατόμων σχετικά με τον αριθμό των παιδιών είναι: 2 1 2 0 3

Και στις δύο περιπτώσεις τα δεδομένα είναι αριθμητικά. Το ύψος όμως δηλώνεται και εκφράζεται με δεκαδικό αριθμό σε αντίθεση με το πλήθος των παιδιών που εκφράζεται με ακέραιο αριθμό. Δεν είναι δυνατόν κάποιος να έχει 1,5 παιδιά ή μπορεί όμως να έχει ύψος 1,71 ή 1,712 ή 1,7119 κτλ ανάλογα με το πόσο ακριβείς επιθυμούμε να είναι οι μετρήσεις μας.

Αριθμητικά δεδομένα όπως το ύψος τα οποία μπορούν να εκφραστούν καλύπτοντας ένα ολόκληρο διάστημα τιμών καλούνται συνεχή δεδομένα (continuous data).

Τα αριθμητικά στατιστικά στοιχεία που εκφράζονται με
αέριες τρές μόνον λέγονται ασυνεχή ή διακριτά
δεδομένα (discrete, non-continuous data). Τα συνεχή
δεδομένα είναι μετρήσιμα (με κάποια μονάδα μέτρησης
π.χ. μέτρο, κιλό κλπ.) ενώ τα διακριτά δεδομένα είναι
απαριθμήσιμα.

Το ύψος, το βάρος, η ηλικία, το εισόδημα, τα ποσά
κατανάλωσης, η θερμοκρασία, το μήκος χαλιών κλπ.
αποτελούν παραδείγματα συνεχών στατιστικών στοιχεί-
ων. Ο αριθμός παιδιών, ο αριθμός μελών ενός νοικοκυριού,
ο αριθμός ατυχημάτων κλπ., αποτελούν παραδείγματα
διακριτών στατιστικών στοιχείων.