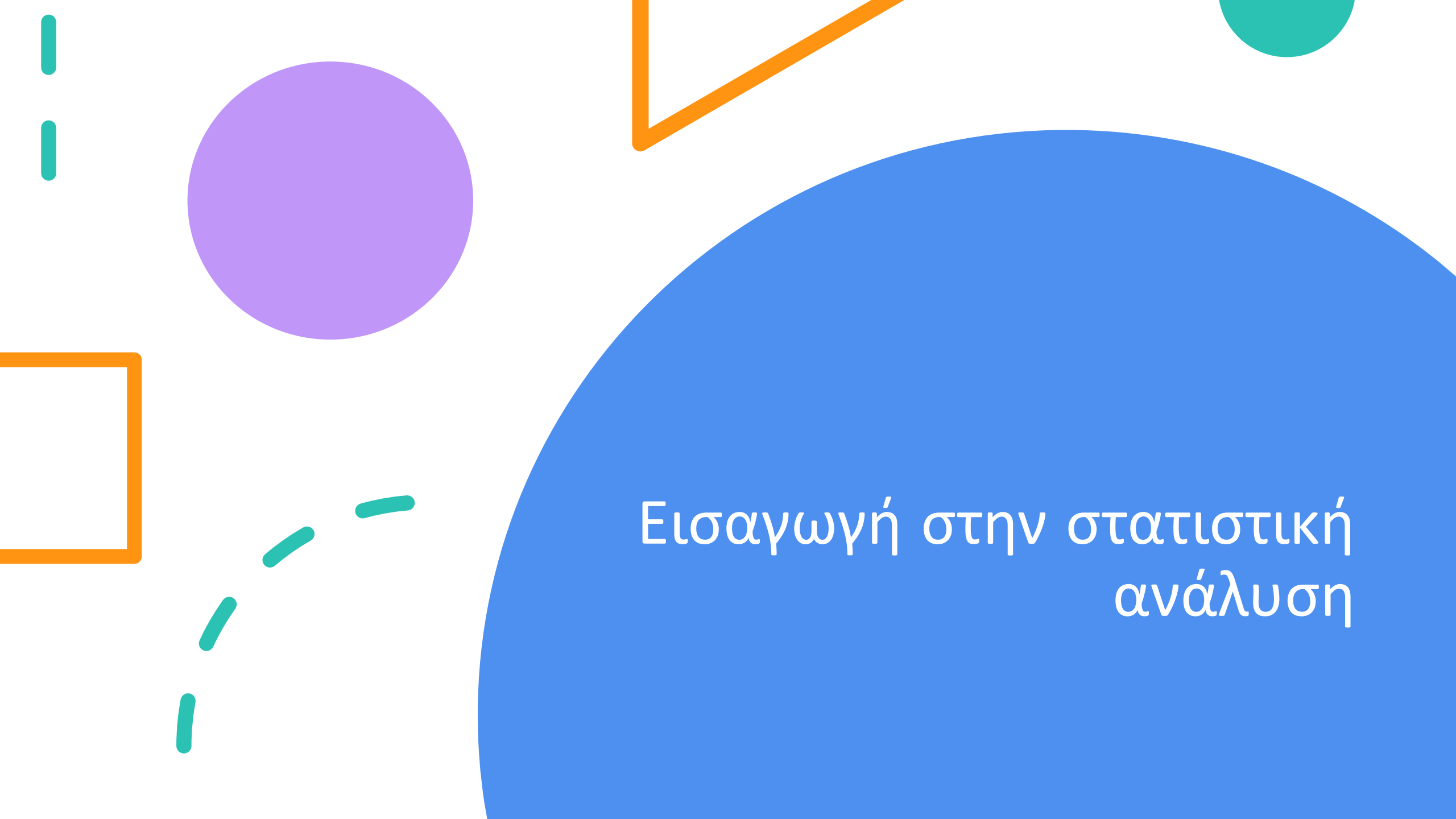
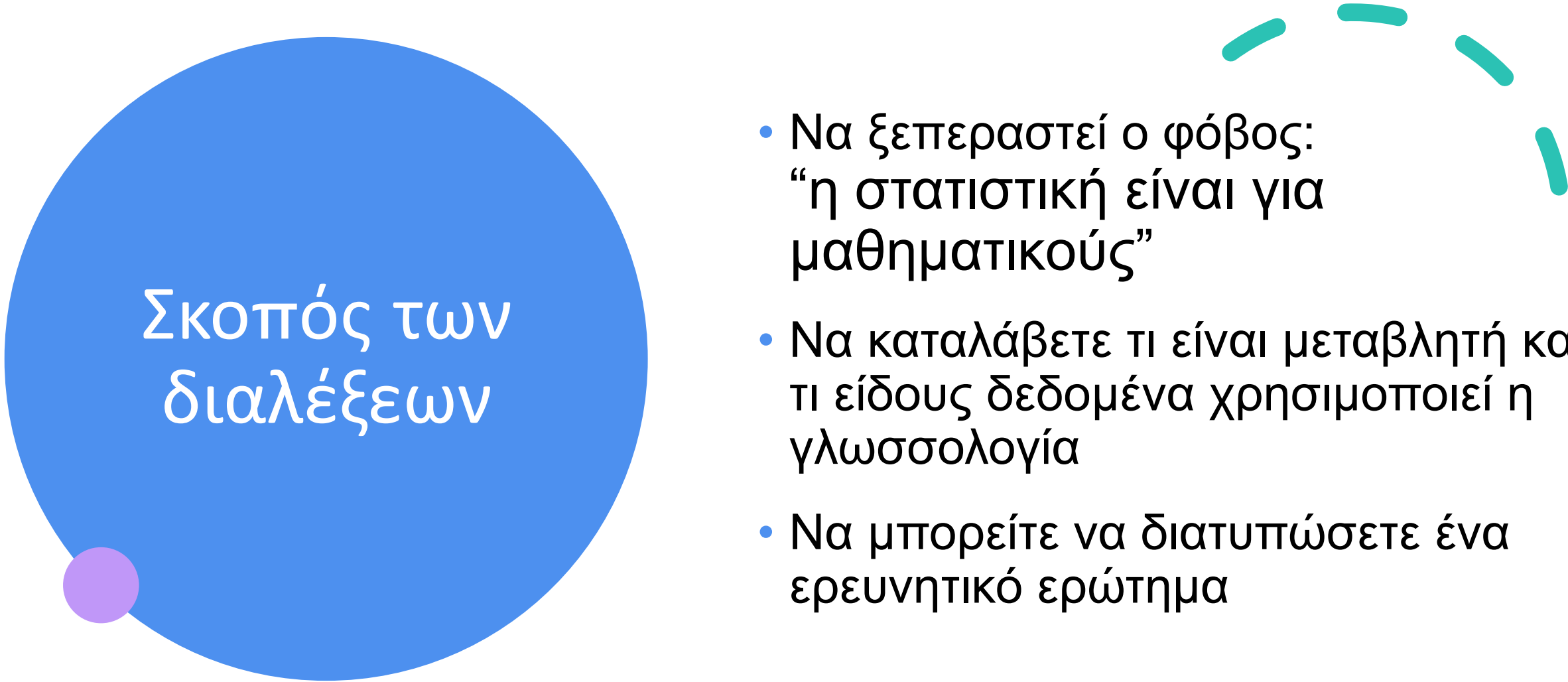




Εισαγωγή στην στατιστική ανάλυση



Σκοπός των διαλέξεων

- Να ξεπεραστεί ο φόβος:
“η στατιστική είναι για
μαθηματικούς”
- Να καταλάβετε τι είναι μεταβλητή και
τι είδους δεδομένα χρησιμοποιεί η
γλωσσολογία
- Να μπορείτε να διατυπώσετε ένα
ερευνητικό ερώτημα

Γιατί η στατιστική στη Φιλολογία;

- Πόσο συχνά εμφανίζεται ένα φωνολογικό φαινόμενο σε ένα ιδίωμα;
- Γράφουν οι μαθητές πιο σύντομες προτάσεις στα αγγλικά όταν γράφουν χειρόγραφα ή στον υπολογιστή;
- Χρησιμοποιούν οι άνδρες και οι γυναίκες διαφορετικά συχνά τον ενεστώτα στον προφορικό λόγο;
- Βελτιώνεται η ανάγνωση παιδιών με δυσλεξία μετά από συγκεκριμένη παρέμβαση;

→ Σε όλες αυτές τις ερωτήσεις χρειαζόμαστε αριθμούς και στατιστική για να απαντήσουμε με αξιοπιστία.



Σκοπός των διαλέξεων

- Να καταλάβετε τι είναι στατιστική ως εργαλείο
- Να καταφέρετε να υλοποιείτε στατιστικούς ελέγχους
- Να καταφέρετε να διατυπώσετε τα συμπεράσματα της έρευνάς σας



Τι είναι η Στατιστική

Είναι ένα σύνολο αρχών και μεθοδολογιών για:

- το σχεδιασμό της διαδικασίας συλλογής **δεδομένων**
- τη συλλογή, οργάνωση, περιγραφή και παρουσίαση των **δεδομένων** αυτών
- τη διερεύνηση και ανάλυση των **δεδομένων** αυτών για την εξαγωγή έγκυρων συμπερασμάτων

ΔΕΔΟΜΕΝΑ



Ποσοτική έρευνα

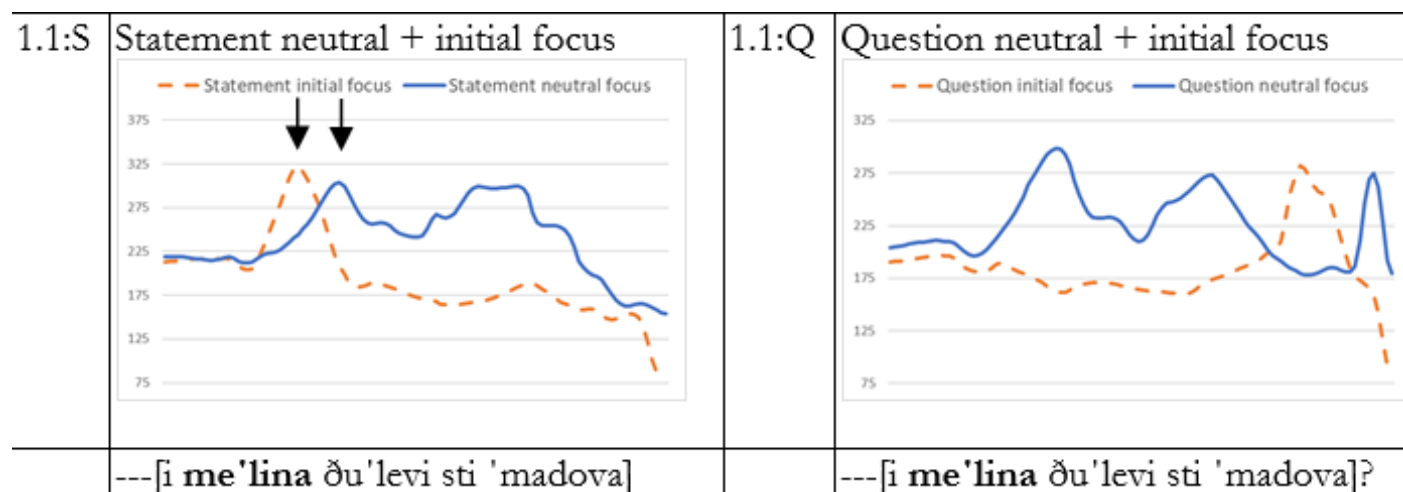
- Ασχολείται με τη συστηματική συλλογή αριθμητικών δεδομένων τον έλεγχο και την ανάλυσή τους.
- Σκοπός είναι η εύρεση σχέσεων/επιδράσεων/συσχετίσεων μεταξύ μεταβλητών με έμφαση στην διατύπωση και τον έλεγχο υποθέσεων.
- Έχω μεγάλο δείγμα συμμετεχόντων (>30)

Ποσοτική έρευνα

- Η ερευνητική διαδικασία είναι δομημένη και γραμμική και γίνεται μέσω:
 - ελεγχόμενων πειραμάτων
 - ερωτηματολογίων κλειστού τύπου
 - δομημένων συνεντεύξεων
 - μη συμμετοχικής παρατήρησης

Ποσοτική έρευνα

- **Ελεγχόμενο πείραμα**
- Στο ελεγχόμενο πείραμα, οι επιστήμονες αλλάζουν μόνο μία μεταβλητή και κρατούν όλες τις άλλες σταθερές.
- Σκοπός του είναι να διερευνήσουν τη σχέση μεταξύ της αλλαγής αυτής της μεταβλητής και των αποτελεσμάτων που προκύπτουν.



Ποσοτική έρευνα

- Ερωτηματολόγια κλειστού τύπου
- Περιέχει συγκεκριμένες απαντήσεις που έχουν προκαθοριστεί από τον ερευνητή.
- Οι απαντήσεις συνήθως είναι σε μορφή κλίμακας Likert ή σε μορφή ναι/όχι ή πολλαπλής επιλογής.

ΕΡΩΤΗΜΑΤΟΛΟΓΙΟ ΓΙΑ ΤΗ ΣΥΝΟΛΙΚΗ ΕΚΤΙΜΗΣΗ ΤΗΣ ΙΚΑΝΟΠΟΙΗΣΗΣ ΑΠΟ ΤΗ ΝΟΣΗΛΕΙΑ							
Ι. Σχετικά με τους/τις γιατρούς αυτού του νοσοκομείου, πώς θα αξιολογούσατε:		Κακό 1	Μέτριο 2	Καλό 3	Πολύ καλό 4	Άριστο 5	Μη εφαρμόσιμο/ μη σχετικό
A1	Τον τρόπο που σας εξέτασαν (ψηλάφηση, ακρόαση κλπ)	☐	☐	☐	☐	☐	☐
A2	Τη σημασία που έδωσαν στο ιστορικό της υγείας σας	☐	☐	☐	☐	☐	☐
A3	Το πόσο καλά φαίνεται να γνωρίζουν την ασθένειά σας	☐	☐	☐	☐	☐	☐
A4	Τη θεραπεία και την ιατρική παρακολούθηση που οργάνωσαν	☐	☐	☐	☐	☐	☐
A5	Τις ερωτήσεις που σας έκαναν για τα σωματικά σας προβλήματα	☐	☐	☐	☐	☐	☐

Ποσοτική έρευνα

- **Δομημένη συνέντευξη**
- Είναι μία τεχνική συλλογής δεδομένων κατά την οποία ο ερευνητής υποβάλει συγκεκριμένο αριθμό προκαθορισμένων ερωτήσεων που γίνονται με ελεγχόμενο τρόπο. Όλες οι απαντήσεις είναι κωδικοποιημένες.
 - Π.χ. δημοσκοπήσεις, έρευνες αγοράς

Δεδομένα έρευνας

- Πρωτογενή δεδομένα
 - Είναι τα δεδομένα που συλλέγονται για τον σκοπό της έρευνας με χρήση συγκεκριμένων εργαλείων (πείραμα, ερωτηματολόγια, παρατήρηση, συνέντευξη)
 - Μεθοδολογικά: μελέτη περίπτωσης, ποσοτική ή ποιοτική ή μεικτή έρευνα

Δεδομένα έρευνας

- Δευτερογενή δεδομένα
 - Συλλέγονται αφού έχει καθοριστεί και διευκρινιστεί το ερευνητικό ερώτημα
 - Είναι τα δεδομένα που υπάρχουν ήδη σε μία μελέτη που έχει διαφορετικό σκοπό από τη δική μας. Τροφοδοτούν τη δική μας μελέτη μέσω συμπληρωματικής ή διαφορετικής γνώσης, διαφορετικών ερμηνειών και συμπερασμάτων

Δευτερογενή δεδομένα – πλεονεκτήματα

- Οικονομία (κόστους, χρόνου)
- Πηγή ιδεών:
 - Για τη διευκρίνιση του υπό έρευνα προβλήματος
 - Για πιθανή λύση ή αστοχίες στο υπό έρευνα πρόβλημα
 - Για επιλογή της ερευνητικής μεθόδου και της ανάλυσης δεδομένων (όταν έχω πρωτογενή δεδομένα)
- Σημείο αναφοράς και σύγκρισης
 - Για την εξέλιξη ενός φαινομένου
 - Για την επαλήθευση των πρωτογενών δεδομένων

Δευτερογενή δεδομένα – περιορισμοί

- Καταλληλότητα των δεδομένων
 - Ταιριάζουν με την έρευνά σας;
 - Είναι μήπως παρωχημένα;
 - Είναι διαθέσιμα; (περιορισμός πρόσβασης λόγω πνευματικών δικαιωμάτων)
 - Είναι έγκυρα κι αξιόπιστα ή ανακριβή κι ανεπαρκή;

Ένα παράδειγμα πρωτογενών – δευτερογενών δεδομένων

«Η θέση των Φαράσων βρισκόταν στην περιοχή της Νοτιοανατολικής Καππαδοκίας, στο ενδιάμεσο του Ταύρου και Αντιταύρου, μέσα σε ένα βαθύ φαράγγι. Διάφορες κοινότητες απαρτίζουν την περιοχή αυτή όπως το Τασσί, το Πεσκαρντάς, το Σατή, το Τσουχούρι, ο Βαρασός, η Κίσκα, το Αφσιάρι. **Σκοπός της έρευνας αυτής είναι να** παρουσιάσει τα λαογραφικά στοιχεία, το γλωσσικό ιδίωμα, την εκπαίδευση και την οικονομία που αναπτύχθηκαν στα Φάρασα και πώς στοιχεία αυτών μεταφέρθηκαν στην περιοχή των Πετρανών και διαμόρφωσαν την συνείδηση και τη συμπεριφορά των κατοίκων της. Το θέμα της παρούσας εργασίας έχει ελάχιστα διερευνηθεί και περιορίζεται σε μεμονωμένες εργασίες, περισσότερο αφηγηματικού χαρακτήρα. Επομένως η εργασία αυτή ίσως αποτελέσει μία συγκεντρωτική πηγή για όσους θέλουν στο μέλλον να προσπαθήσουν να πραγματοποιήσουν παρόμοιες έρευνες. Η έρευνα θεμελιώθηκε με βάση τις αρχές της ιστορικής εθνογραφίας. **Χρησιμοποιήθηκαν πρωτογενή και δευτερογενή δεδομένα. Στα πρωτογενή συμπεριλαμβάνεται η επιτόπια έρευνα, η οποία βασίστηκε στην άμεση και συμμετοχική παρατήρηση, καθώς και στη συνέντευξη, ενώ οι δευτερεύουσες πηγές αντλήθηκαν από άρθρα επιστημονικών περιοδικών, πρακτικά συνεδρίων, μεταπτυχιακές διπλωματικές εργασίες, καθώς και από συνεντεύξεις κατοίκων των Πετρανών.** Από τη συστηματική εξέταση των δεδομένων προκύπτει ότι μεταφέρθηκαν λαογραφικά, γλωσσικά και πολιτισμικά στοιχεία του πολιτισμού της περιοχής των Φαράσων στην περιοχή των Πετρανών Κοζάνης. Οι κάτοικοι των Φαράσων παρόλο που ζούσαν σε μια ευρύτερη αλλόθρησκη και τουρκόφωνη περιοχή, κατόρθωσαν να διατηρήσουν την εθνική και θρησκευτική τους ταυτότητα και τις παραδόσεις.»

Πηγή: Αποθετήριο Π.Δ.Μ. (<http://dspace.uowm.gr/xmlui/handle/123456789/1810>)

Συγγραφέας: Καραγκίοζη Ιωάννα, «Από τα Φάρασα της Καππαδοκίας στα Πετρανά Κοζάνης: ιστορική διαδρομή, γλώσσα και λαογραφικά στοιχεία, εκπαίδευση και οικονομία», Πτυχιακή εργασία

Άλλο ένα παράδειγμα πρωτογενών – δευτερογενών δεδομένων

2.1 Μεθοδολογία

Παρουσιάζουμε στη συνέχεια τα αλλόφωνα του /l/ και τα ακουστικά χαρακτηριστικά τους, με βάση υλικό από δομημένες ηχογραφήσεις ομιλίας δύο γυναικών ηλικίας 80 ετών περίπου. Οι δύο γυναίκες χρησιμοποιούν τη διάλεκτο στην καθημερινή τους επικοινωνία. Επιλέξαμε λέξεις και φράσεις όπου το /l/ ακολουθείται από όλα τα φωνήεντα, τονισμένα και άτονα (τα /o/ και /e/ πραγματώνονται μόνο τονισμένα), με προηγούμενα όλα τα φωνήεντα, τονισμένα και άτονα. Επίσης εξετάσαμε και εκφωνήματα με τα [l] και [ɫ] στο τέλος της λέξης. Κάθε λέξη ή φράση επαναλαμβανόταν τρεις φορές. Η ανάλυση του υλικού έγινε με το πρόγραμμα Praat, 6.0.36.

Πηγή: Σ. Αρμοστής, Μ. Μαργαρίτη – Ρόγκα & Μ. Τσολάκη, «Το /l/ στο ιδίωμα της Σαμοθράκης»

Άλλο ένα παράδειγμα πρωτογενών – δευτερογενών δεδομένων

Το /l/ στο ιδίωμα της Σαμοθράκης, ήδη στις παλαιότερες μαρτυρίες, εμφανίζει αποβολή, όπως και το /r/, ή τροπή (ενδεχομένως, σε ορισμένες περιπτώσεις, αντικατάσταση) σε άλλο φθόγγο. Συγκεκριμένα, ο Conze (1860, 52–54), που επισκέφθηκε το νησί το 1858, παραθέτει μερικές λέξεις και φράσεις του ιδιώματος, όπου παρατηρούμε:⁵

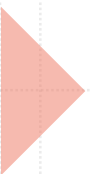
(α) αποβολή, π.χ.

[1] καὸ ‘καλό’, μαξιιάι ‘μαξιλάρι’, όγια ‘λόγια’, άογο ‘άλογο’.

(β) τροπή του αλλοφώνου [ɬ] σε [j] (με το [ɬ] δηλώνουμε ουρανικοποιημένο φατνιακό ήχο [lʰ]),⁶ π.χ.

[2] γήγιος [ˈjijos] ‘ήλιος’, να γιτιϊγήσουμ [naʝitiːˈjisum] ‘να λειτουργήσουμε’.

Πηγή: Σ. Αρμოსτής, Μ. Μαργαρίτη – Ρόγκα & Μ. Τσολάκη, «Το /l/ στο ιδίωμα της Σαμοθράκης»



Στατιστική Ανάλυση

1. Σχεδιασμός της διαδικασίας συλλογής δεδομένων

2. Συλλογή, οργάνωση, περιγραφή και παρουσίαση των δεδομένων αυτών

3. Διερεύνηση και ανάλυση των δεδομένων αυτών για την εξαγωγή έγκυρων συμπερασμάτων

- Καθορίζουμε τον σκοπό της έρευνας
- Ορίζουμε τους όρους που θα χρησιμοποιηθούν στην έρευνα
- Προσδιορίζουμε τις πηγές από όπου θα συλλεγούν τα στατιστικά δεδομένα:
 - συμμετέχοντες
 - μεθοδολογία (υλικό πειρ. διαδικασίας)
 - διαδικασία (περιγραφή πειρ. διαδικασίας)
- Συλλέγουμε τα δεδομένα

Στατιστική Ανάλυση

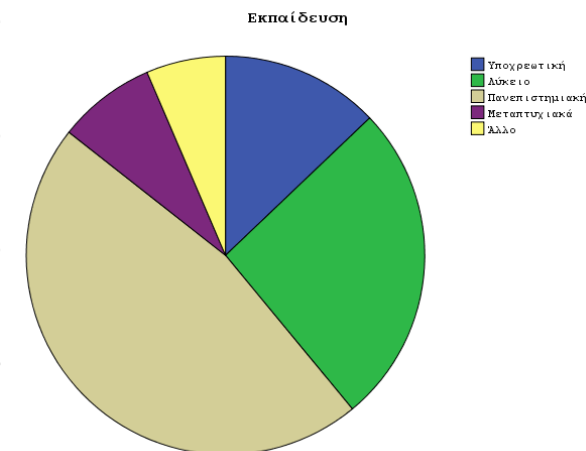
1. Σχεδιασμός της διαδικασίας συλλογής δεδομένων
- 2. Συλλογή, οργάνωση, περιγραφή και παρουσίαση των δεδομένων αυτών**
3. Διερεύνηση και ανάλυση των δεδομένων αυτών για την εξαγωγή έγκυρων συμπερασμάτων

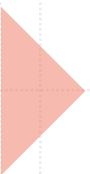
Περιγραφική Στατιστική

- Παρουσίαση των δεδομένων σε κατάλληλη μορφή:

Εκπαίδευση

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Υποχρεωτική	35	13,1	13,1	13,1
	Λύκειο	72	27,0	27,0	40,1
	Πανεπιστήμιο	122	45,7	45,7	85,8
	Μεταπτυχιακό	20	7,5	7,5	93,3
	Άλλο	18	6,7	6,7	100,0
	Total	267	100,0	100,0	



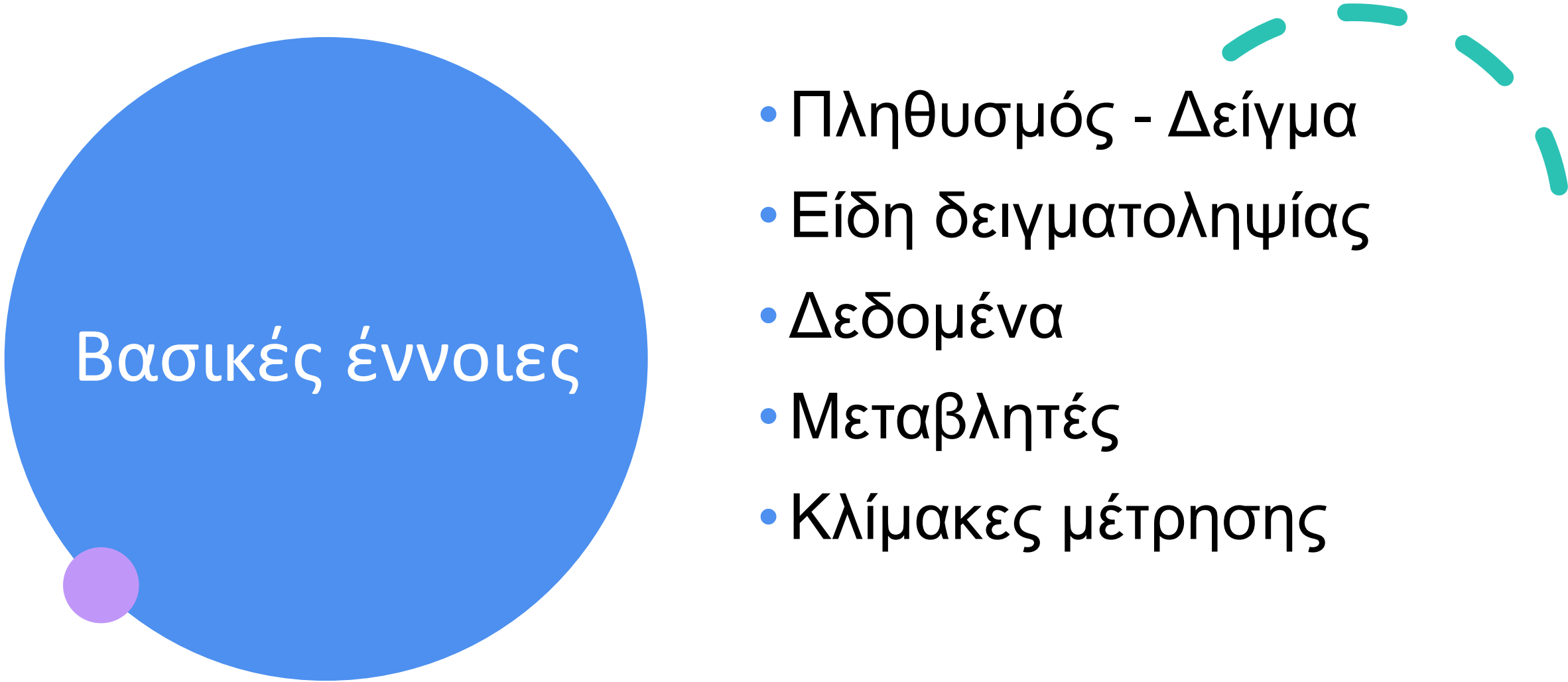


Στατιστική Ανάλυση

1. Σχεδιασμός της διαδικασίας συλλογής δεδομένων
2. Συλλογή, οργάνωση, περιγραφή και παρουσίαση των δεδομένων αυτών
3. **Διερεύνηση και ανάλυση των δεδομένων αυτών για την εξαγωγή έγκυρων συμπερασμάτων**

Επαγωγική Στατιστική

- Ανάλυση των δεδομένων
- Ερμηνεία των αποτελεσμάτων
- Παρουσίαση των σχετικών συμπερασμάτων που προέκυψαν από τη στατιστική έρευνα:
 - απάντηση στο ερευνητικό ερώτημα
 - γενίκευση δεδομένων
 - συζήτηση



Βασικές έννοιες

- Πληθυσμός - Δείγμα
- Είδη δειγματοληψίας
- Δεδομένα
- Μεταβλητές
- Κλίμακες μέτρησης

Πληθυσμός - Δείγμα

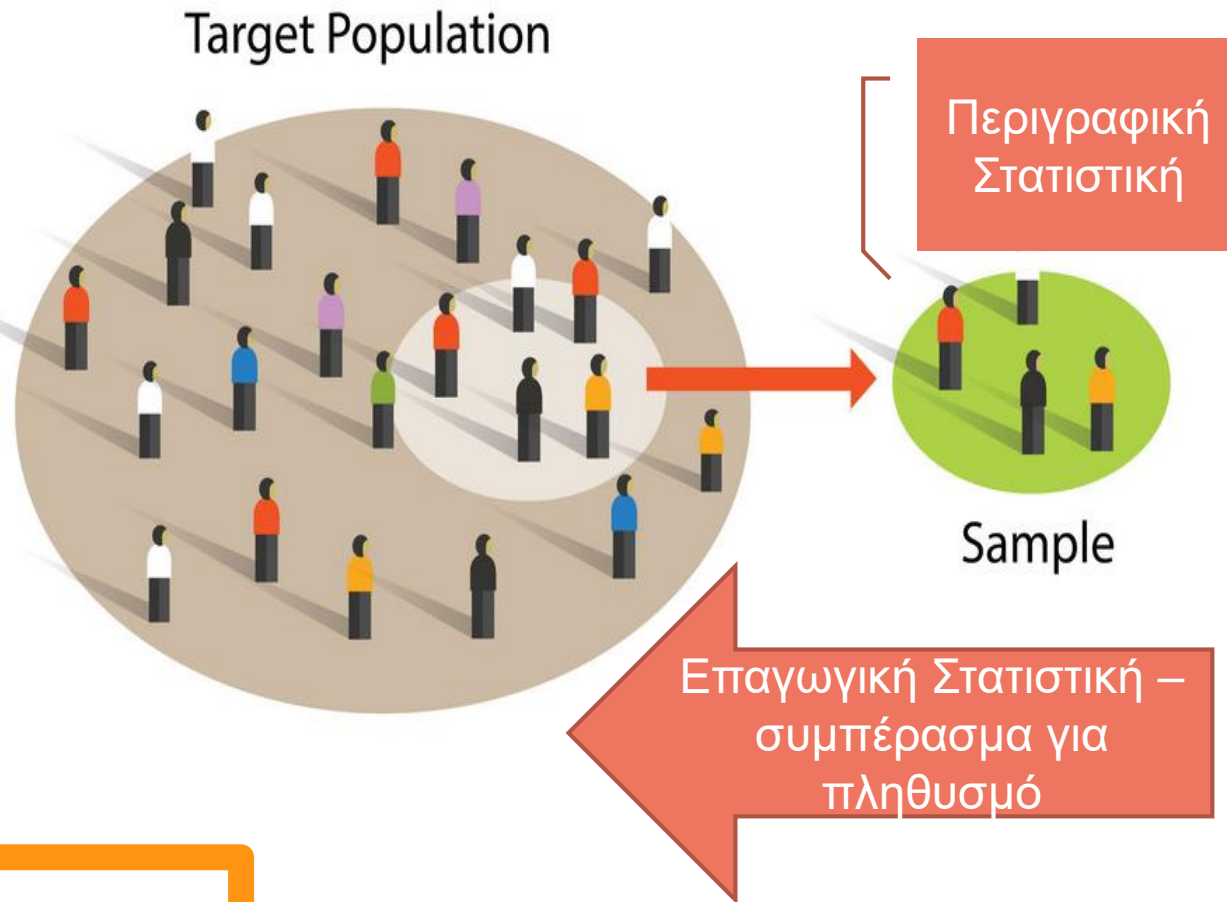
Πληθυσμός:

Το σύνολο των παρατηρήσεων που μελετάμε



Δείγμα:

Το υποσύνολο ενός πληθυσμού



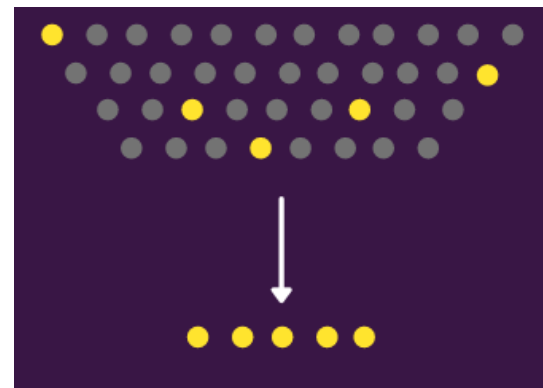
-Τεχνικές που επιτρέπουν την ανάλυση των δεδομένων που έχω συλλέξει από το δείγμα μου.

-Στη συνέχεια μπορώ να εξάγω χρήσιμα και γενικεύσιμα συμπεράσματα για τον πληθυσμό.

Είδη δειγματοληψίας

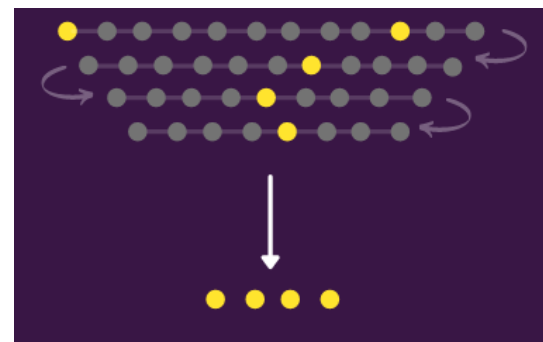
ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

Κάθε μέλος του πληθυσμού έχει τις ίδιες πιθανότητες να επιλεγεί



ΣΥΣΤΗΜΑΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

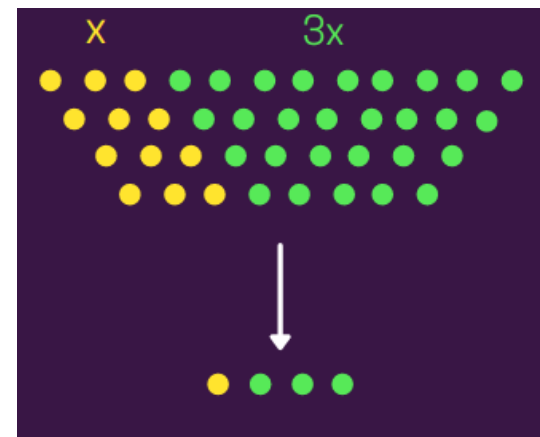
Επιλογή του δείγματος από κατάλογο βάσει συστηματικού σχεδίου



Είδη δειγματοληψίας

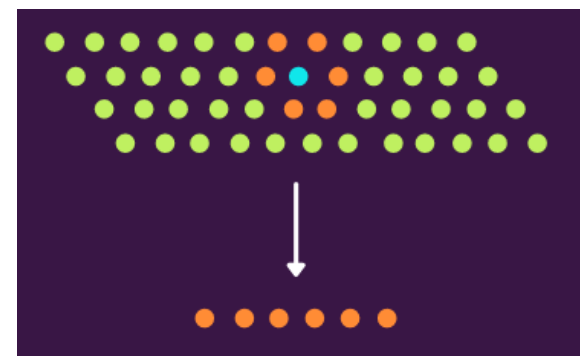
ΔΕΙΓΜΑΤΟΛΗΨΙΑ ΚΑΤΑ ΣΤΡΩΜΑΤΑ

Επιβολή περιορισμών στο δείγμα μέσω διαστρωμάτωσης



ΣΥΜΠΤΩΜΑΤΙΚΗ ΔΕΙΓΜΑΤΟΛΗΨΙΑ

Αλλιώς ευκαιριακό δείγμα. Χρησιμοποίηση δειγμάτων απλά και μόνο γιατί είναι διαθέσιμα, π.χ. φοιτητές



Δεδομένα

Είναι οι τιμές (παρατηρήσεις) που καταγράφουμε κατά την υλοποίηση της έρευνάς μας

Τα δεδομένα καταγράφονται συνήθως σε πίνακα

Φύλο	Βάρος
Άνδρας	120
Γυναίκα	80
Γυναίκα	55
Άνδρας	90
Άνδρας	75
Γυναίκα	60

Μεταβλητές

Οι τιμές του προηγούμενου πίνακα παριστάνουν τις τιμές που προέκυψαν μετά από κάθε επανάληψη του πειράματος.

Οι ποσότητες αυτές επειδή επαναλαμβάνονται κατά τυχαίο τρόπο ονομάζονται **μεταβλητές**

Μεταβλητή 1 ↓	Μεταβλητή 2 ↓
Φύλο	Βάρος
Άνδρας	120
Γυναίκα	80
Γυναίκα	55
Άνδρας	90
Άνδρας	75
Γυναίκα	60

Είδη μεταβλητών i

Κατηγορικές ή Ποιοτικές μεταβλητές (Categorical or Qualitative Variables)

- Είναι αυτές που οι τιμές τους δεν είναι αριθμητικές
- Μεταβάλλονται σε είδος (π.χ. χρώμα ματιών, μέτρια / καλή / άριστη φυσική κατάσταση κλπ)
- Ονομαστικές ή Διατακτικές

Αριθμητικές ή Ποσοτικές Μεταβλητές (Numerical or Quantitative Variables)

- Είναι αυτές που οι τιμές τους είναι αριθμητικές
- Μεταβάλλονται σε ποσό (π.χ. βάρος, μισθός κλπ)

Είδη μεταβλητών i

Παραδείγματα από τη γλωσσολογία:

- Κατηγορικές: γλώσσα, φύλο ομιλητή, είδος διαλέκτου
- Αριθμητικές: αριθμός λέξεων ανά λεπτό, βαθμός κατανόησης κειμένου, χρόνος αντίδρασης σε πείραμα



Είδη μεταβλητών i

Μπορείτε να σκεφτείτε δύο παραδείγματα
κατηγορικών και δύο παραδείγματα
αριθμητικών μεταβλητών;



Σημείωση

- Για να μπορέσει ένα πρόγραμμα (όπως το jamon ή το SPSS) να κάνει στατιστική ανάλυση, όλες οι κατηγορικές μεταβλητές πρέπει να μετατραπούν σε αριθμούς. Αυτό σημαίνει ότι κάθε κατηγορία παίρνει έναν μοναδικό αριθμό. Οι αριθμοί αυτοί χρησιμοποιούνται μόνο για την κωδικοποίηση — **δεν έχουν ποσοτική σημασία.**

Μεταβλητή: *Βιολογικό φύλο*

- Αρσενικό = **1**
- Θηλυκό = **2**

Μεταβλητή: *Μητρική Γλώσσα*

- Ελληνικά = **1**
- Αγγλικά = **2**
- Αλβανικά = **3**

Κλίμακες μέτρησης

Κατηγορική Κλίμακα (categorical scale)

Οι αριθμοί δεν έχουν ποσοτικές ιδιότητες (δεν μπορούμε να κάνουμε πράξεις), απλώς αντιπροσωπεύουν τις διάφορες κατηγορίες

What is your gender?

- M – Male
- F – Female

What is your hair color?

- 1 – Brown
- 2 – Black
- 3 – Blonde
- 4 – Gray
- 5 – Other

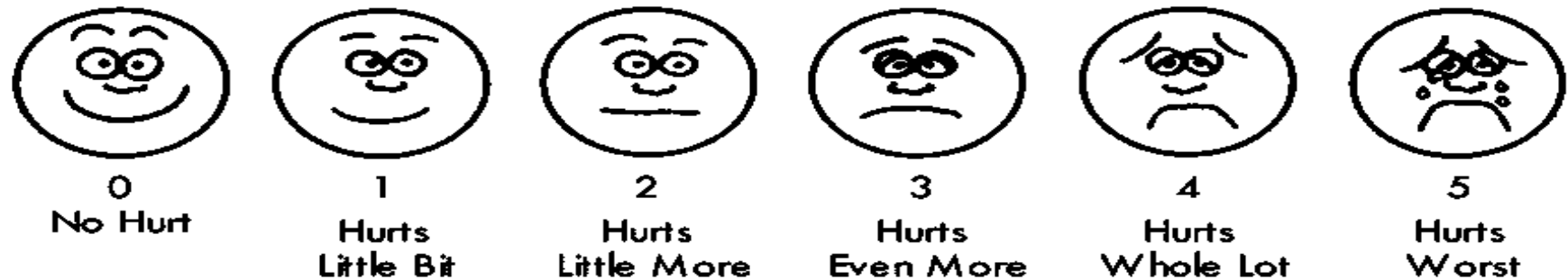
Where do you live?

- A – North of the equator
- B – South of the equator
- C – Neither: In the international space station

Κλίμακες μέτρησης

Ιεραρχική Κλίμακα (ordinal scale)

Οι αριθμοί δεν έχουν ποσοτικές ιδιότητες (δεν μπορούμε να κάνουμε πράξεις), αντιπροσωπεύουν τις διάφορες κατηγορίες και υπάρχει ταξινόμηση σε σειρά ή θέση (διάταξη)



Κλίμακες μέτρησης

Κλίμακα Ίσων Διαστημάτων (interval scale)

Οι αριθμοί αντιπροσωπεύουν μία κατηγορία και μας δίνουν πληροφορίες και για τις διαφορές που υπάρχουν ανάμεσα στις θέσεις κατάταξης

Τα διαστήματα ανάμεσα στις τιμές της κλίμακας είναι ίσα

Μπορούμε να κάνουμε πράξεις

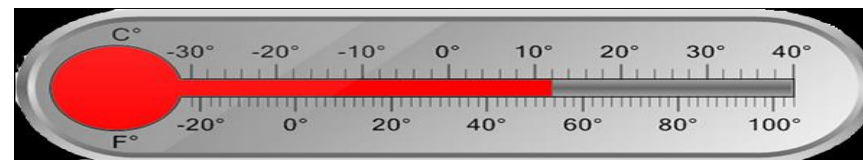
Δεν υπάρχει το απόλυτο 0

Κλίμακες μέτρησης

Κλίμακα Ίσων Διαστημάτων (interval scale)

Παράδειγμα 1. Θερμοκρασία αντικειμένων: 27° & 35°

- Το δεύτερο αντικείμενο είναι θερμότερο από το πρώτο (ιεραρχική κλίμακα)
- Η διαφορά μεταξύ τους είναι 8° C



Παράδειγμα 2. Ομαδοποιημένα αριθμητικά δεδομένα

- Συχνότητα άσκησης στο γυμναστήριο: 1= 1-5 φορές/ μήνα, 2= 6-10 φορές/μήνα, 3= 11-15 φορές/ μήνα, 4= 16-20 φορές/ μήνα

Και στις δύο περιπτώσεις τα διαστήματα ανάμεσα στις τιμές είναι ίσα

Κλίμακες μέτρησης

Αναλογική Κλίμακα (ratio scale)

- Περιλαμβάνει τις ιδιότητες όλων των προηγούμενων κλιμάκων
- Εξασφαλίζει ταξινόμηση, ίσα διαστήματα και αναγνώριση
- Ο αριθμός 0 έχει νόημα (π.χ. 0 έτη προϋπηρεσίας)
- Οι τιμές ερμηνεύονται από όλους το ίδιο
- Οι τιμές είναι συγκρίσιμες μεταξύ τους (20 έτη προϋπηρεσίας είναι διπλάσια από τα 10 έτη)

Κλίμακες μέτρησης

Δώστε ένα παράδειγμα μιας μέτρησης που μπορεί να θεωρηθεί σε μερικές περιπτώσεις ιεραρχική και σε άλλες αναλογική

Κλίμακες μέτρησης

Βαθμολογία σε διαγωνισμό ή τεστ

Αντίληψη ιεραρχίας: Πρώτος, δεύτερος, τρίτος — δείχνει μόνο τη σειρά κατάταξης (ordinal).

Αναλογική χρήση: Αριθμητική βαθμολογία 0–100 — μπορεί να μετρηθεί αναλογικά (ratio), γιατί υπάρχει απόλυτο μηδέν και διαστήματα με νόημα.

Ύψος ή βάρος ατόμων

Ιεραρχική χρήση: Κατηγοριοποίηση σε «χαμηλό», «μέτριο», «ψηλό».

Αναλογική χρήση: Ακριβές ύψος σε μέτρα ή βάρος σε κιλά.

Χρόνος ολοκλήρωσης ενός αγώνα

Ιεραρχική χρήση: Πρώτος, δεύτερος, τρίτος (κατάταξη).

Αναλογική χρήση: Ακριβής χρόνος σε δευτερόλεπτα (με αναλογία, διότι υπάρχει απόλυτο μηδέν).

Εισόδημα

Ιεραρχική χρήση: Κατηγορίες όπως «χαμηλό», «μεσαίο», «υψηλό».

Αναλογική χρήση: Ακριβές ποσό σε ευρώ ή δολάρια.

Κλίμακες μέτρησης

		Κατηγορίες	Διάταξη	Ίσα διαστήματα σε όλο το μήκος της κλίμακας	Υπάρχει το απόλυτο μηδέν
Κατηγορικές μεταβλητές	Ονομαστική κλίμακα	x			
	Ιεραρχική κλίμακα	x	x		
Ποσοτικές μεταβλητές	Κλίμακα ίσων διαστημάτων		x	x	
	Αναλογική κλίμακα		x	x	x

Είδη Μεταβλητών ii

Εξαρτημένη Μεταβλητή (Dependent Variable)

- Είναι η μεταβλητή που μετράμε (δηλαδή τα δεδομένα της έρευνάς μας)

Ανεξάρτητη Μεταβλητή (Independent Variable)

- Είναι η μεταβλητή για την οποία ο ερευνητής χρειάζεται να διαπιστώσει αν ασκεί κάποια επίδραση πάνω στην εξαρτημένη μεταβλητή. Αυτή μπορεί ο ερευνητής να αποφασίσει πως θα την χρησιμοποιήσει σε μία έρευνα

Σενάρια – ερευνητικά ερωτήματα

- Σε μια κοινότητα, κάποια παιδιά εκπαιδεύονται ταυτόχρονα σε δύο γλώσσες, τη μητρική και μια ξένη γλώσσα, ενώ τα υπόλοιπα μόνο στη μητρική. Οι εκπαιδευτικοί θέλουν να μελετήσουν αν υπάρχουν διαφορές στην ανάπτυξη λεξιλογίου.
- Ποιες είναι οι μεταβλητές (είδος-εξαρτημένη-ανεξάρτητη);
- Ποιο είναι ένα πιθανό ερευνητικό ερώτημα;

Σενάρια – ερευνητικά ερωτήματα

- Σε μια κοινότητα, κάποια παιδιά εκπαιδεύονται ταυτόχρονα σε δύο γλώσσες, τη μητρική και μια ξένη γλώσσα, ενώ τα υπόλοιπα μόνο στη μητρική. Οι εκπαιδευτικοί θέλουν να μελετήσουν αν υπάρχουν διαφορές στην ανάπτυξη λεξιλογίου.
- **Μεταβλητές:**
- **Ανεξάρτητη:** Τύπος εκπαίδευσης (μονόγλωσση, δίγλωσση) – **Κατηγορική.**
- **Εξαρτημένη:** Ανάπτυξη λεξιλογίου στη μητρική γλώσσα – **Αριθμητική.**
- **Ερευνητική Υπόθεση (H1):**
Η δίγλωσση εκπαίδευση επηρεάζει την ανάπτυξη του λεξιλογίου στη μητρική γλώσσα.

Σενάρια – ερευνητικά ερωτήματα

- Σε μια πολυπολιτισμική πόλη, οι κάτοικοι μιλούν δύο διαλέκτους: τη Διάλεκτο Α (π.χ., τοπική) και τη Διάλεκτο Β (π.χ., πρότυπη). Παρατηρείται ότι η επιλογή διαλέκτου εξαρτάται από την κοινωνική περίσταση (π.χ., επίσημη ή ανεπίσημη).
- Ποιες είναι οι μεταβλητές (είδος-εξαρτημένη-ανεξάρτητη);
- Ποιο είναι ένα πιθανό ερευνητικό ερώτημα;

Σενάρια – ερευνητικά ερωτήματα

- Σε μια πολυπολιτισμική πόλη, οι κάτοικοι μιλούν δύο διαλέκτους: τη Διάλεκτο Α (π.χ., τοπική) και τη Διάλεκτο Β (π.χ., πρότυπη). Παρατηρείται ότι η επιλογή διαλέκτου εξαρτάται από την κοινωνική περίσταση (π.χ., επίσημη ή ανεπίσημη).
- **Μεταβλητές:**
 1. **Διάλεκτος** (Διάλεκτος Α ή Διάλεκτος Β) – **Κατηγορική**
 2. **Κοινωνική περίσταση** (Επίσημη ή Ανεπίσημη) – **Κατηγορική**
- **Ερευνητική Υπόθεση (H1):**
 - Η επιλογή διαλέκτου εξαρτάται από την κοινωνική περίσταση (με τη Διάλεκτο Β να προτιμάται σε επίσημες περιστάσεις και τη Διάλεκτο Α σε ανεπίσημες).

Σενάρια – ερευνητικά ερωτήματα

- Ένας ερευνητής θέλει να εξετάσει αν υπάρχει σχέση μεταξύ της ταχύτητας ανάγνωσης (λέξεις ανά λεπτό) και της κατανόησης κειμένου (βαθμολογία σε ερωτηματολόγιο κατανόησης).
- Ποιες είναι οι μεταβλητές (είδος-εξαρτημένη-ανεξάρτητη);
- Ποιο είναι ένα πιθανό ερευνητικό ερώτημα;

Σενάρια – ερευνητικά ερωτήματα

- Ένας ερευνητής θέλει να εξετάσει αν υπάρχει σχέση μεταξύ της ταχύτητας ανάγνωσης (λέξεις ανά λεπτό) και της κατανόησης κειμένου (βαθμολογία σε ερωτηματολόγιο κατανόησης).
- **Μεταβλητές:**
 1. **Ταχύτητα Ανάγνωσης – Αριθμητική** (Μετριέται σε λέξεις ανά λεπτό)
 2. **Κατανόηση Κειμένου – Αριθμητική** (Μετριέται ως βαθμολογία (π.χ., 0-100) σε ερωτηματολόγιο κατανόησης)
- **Ερευνητική Υπόθεση (H1):**
 - Υπάρχει στατιστικά σημαντική σχέση μεταξύ της ταχύτητας ανάγνωσης και της κατανόησης κειμένου.

Η λογική του ελέγχου υποθέσεων

- Φανταστείτε ότι ρωτάτε: “Μαθαίνουν τα παιδιά που διαβάζουν φωναχτά ορθογραφία πιο γρήγορα;”
- Δεν μπορείτε να μετρήσετε ΟΛΟΥΣ τους μαθητές του κόσμου. Παίρνετε ένα **δείγμα** και κάνετε ένα **έλεγχο υποθέσεων**.
- Η λογική μοιάζει με δικαστήριο:
- Αρχίζουμε με την υπόθεση ότι **δεν υπάρχει καμία διαφορά** (ο κατηγορούμενος είναι “αθώος”)
- Συλλέγουμε **στοιχεία (δεδομένα)**
- Αν τα στοιχεία είναι πολύ ισχυρά, **απορρίπτουμε** την αρχική υπόθεση (“ένοχος”)

Μηδενική υπόθεση (H_0) και Ερευνητική υπόθεση (H_1)

- **Μηδενική Υπόθεση (H_0):** “Δεν υπάρχει διαφορά / σχέση” — αυτή προσπαθούμε να αντικρούσουμε
- **Ερευνητική Υπόθεση (H_1):** “Υπάρχει διαφορά / σχέση” — αυτό θέλουμε να αποδείξουμε
- *Παράδειγμα από τη γλωσσολογία:*
- **H_0 :** Δεν υπάρχει διαφορά στο λεξιλόγιο μεταξύ μονόγλωσσων και δίγλωσσων παιδιών.
- **H_1 :** Τα δίγλωσσα παιδιά έχουν διαφορετική ανάπτυξη λεξιλογίου από τα μονόγλωσσα.
- *→ Ποτέ δεν “αποδεικνύουμε” την H_1 απευθείας. Αντί αυτού, “απορρίπτουμε” ή “δεν απορρίπτουμε” την H_0 .*

Τι είναι το p-value;

- Το **p-value** είναι η πιθανότητα να πάρουμε τα αποτελέσματά μας **αν η H_0 ήταν αληθής**.
- **$p \leq 0.05$** : Απορρίπτουμε την H_0 — το αποτέλεσμα είναι **στατιστικά σημαντικό**
- **$p > 0.05$** : Δεν απορρίπτουμε την H_0 — δεν βρήκαμε αρκετές αποδείξεις
- *Παράδειγμα: συγκρίνω λεξιλόγιο δίγλωσσων / μονόγλωσσων μαθητών:*
- $p = 0.03 \rightarrow$ Η διαφορά είναι πραγματική (όχι τυχαία). Απορρίπτω H_0 .
- $p = 0.42 \rightarrow$ Η διαφορά μπορεί να οφείλεται στην τύχη. Δεν απορρίπτω H_0 .
- *⚠ Το p-value ΔΕΝ μας λέει πόσο “σημαντικό” είναι το εύρημα στην πράξη — μόνο αν είναι τυχαίο ή όχι!*

Ένα πλήρες παράδειγμα ελέγχου υποθέσεων

- **Ερώτημα:** Επηρεάζει η δίγλωσση εκπαίδευση το λεξιλόγιο στη μητρική γλώσσα;
- ① **H₀:** Δεν υπάρχει διαφορά λεξιλογίου μεταξύ δίγλωσσων και μονόγλωσσων παιδιών.
- ② **H₁:** Υπάρχει διαφορά λεξιλογίου μεταξύ των δύο ομάδων.
- ③ Μετράμε λεξιλόγιο σε 40 παιδιά (20 δίγλωσσα, 20 μονόγλωσσα) με τεστ 50 λέξεων.
- ④ Το SPSS βγάζει: **p = 0.02**
- ⑤ **p = 0.02 < 0.05 → Απορρίπτουμε H₀.** Υπάρχει στατιστικά σημαντική διαφορά!
- → *Συμπέρασμα: “Η δίγλωσση εκπαίδευση επηρεάζει σημαντικά την ανάπτυξη λεξιλογίου στη μητρική γλώσσα.”*

Περιγραφική Στατιστική

Είναι μία ομάδα στατιστικών μεθόδων για οργάνωση και περιγραφή των δεδομένων μας

Η παρουσίαση μπορεί να γίνει με μορφή:

- Πινάκων
- Διαγραμμάτων
- Μέτρων θέσης και διασποράς





Βιβλιογραφία

Field A. (2000). *Discovering Statistics using SPSS*. SAGE.

Ρούσσος Π., Τσαούσης Γ. (2010). *Στατιστική εφαρμοσμένη στις κοινωνικές επιστήμες*. Αθηνά: ΕΛΛΗΝΙΚΑ ΓΡΑΜΜΑΤΑ.

<https://www.jamovi.org/user-manual.html>