

Χρήση δευτερογενών δεδομένων

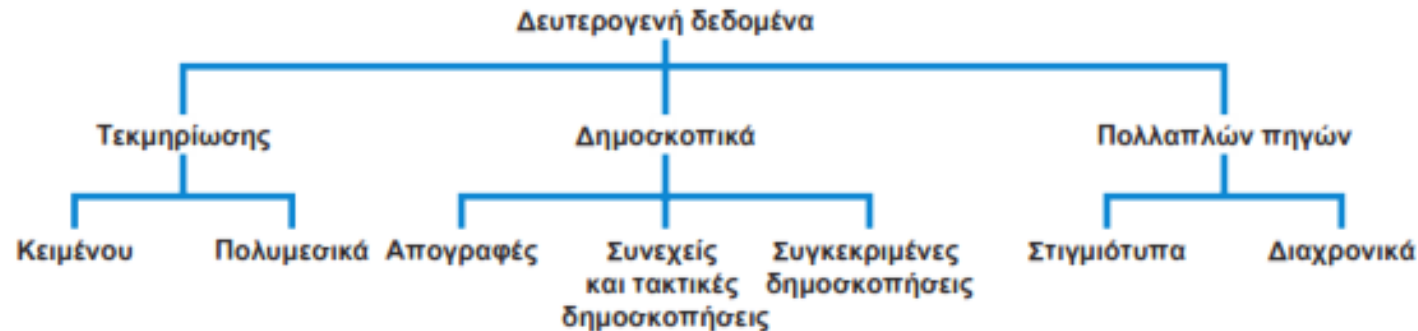
# Μαθησιακοί στόχοι κεφαλαίου

- Να αναγνωρίζετε όλα τα είδη δευτερογενών δεδομένων,
- Να εκτιμάτε τους τρόπους με τους οποίους τα δευτερογενή δεδομένα μπορούν να χρησιμοποιηθούν για να σας βοηθήσουν να απαντάτε στο ερευνητικό σας ερώτημα (ή ερωτήματα) και να πετυχαίνετε τους στόχους σας,
- Να κατανοείτε τα πλεονεκτήματα και τα μειονεκτήματα της χρήσης δευτερογενών δεδομένων,
- Να χρησιμοποιείτε διάφορες τεχνικές για τον εντοπισμό δευτερογενών δεδομένων,
- Να αξιολογείτε την καταλληλότητα των δευτερογενών δεδομένων για το ερευνητικό σας έργο όσον αφορά την κάλυψη, την εγκυρότητα, την αξιοπιστία και τη μεροληψία μέτρησης,

# Δευτερογενή δεδομένα

- ήδη υπάρχοντα δεδομένα που έχουν συλλεχθεί για άλλους λόγους.
- περιλαμβάνουν τόσο ακατέργαστα όσο και συνοπτικά, δημοσιευμένα, δεδομένα.
- μπορούν να αναλυθούν περαιτέρω, ώστε να παρέχουν συμπληρωματική ή διαφορετική γνώση, ερμηνείες ή συμπεράσματα

# Τύποι δευτερογενών δεδομένων



**Παραδείγματα:**

Βάσεις δεδομένων οργανισμών (π.χ., προσωπικού ή παραγωγής)  
 Επικοινωνία οργανισμών, όπως e-mail, επιστολές, σημειώματα  
 Σχόλια στο web  
 Εκθέσεις και πρακτικά επιτροπών  
 Περιοδικά  
 Εφημερίδες  
 Ημερολόγια  
 Μετεγγραφές συνεντεύξεων

**Παραδείγματα:**

Από μέσα ενημέρωσης όπως τηλεόραση και ραδιόφωνο  
 Ηχογραφήσεις  
 Εγγραφές βίντεο  
 Εικόνες και φωτογραφίες  
 Εικόνες στο web

**Παραδείγματα:**

Κρατικές απογραφές: απογραφή πληθυσμού, απογραφή απασχόλησης

**Παραδείγματα:**

Κυβέρνηση: Οικονομικές δαπάνες, τάσεις αγοράς εργασίας  
 Οργανισμοί: Δημοσκοπήσεις συμπεριφοράς εργαζομένων, Δημοσκοπήσεις συμπεριφοράς καταναλωτών

**Παραδείγματα:**

Κυβερνητικές δημοσκοπήσεις  
 Δημοσκοπήσεις οργανισμών  
 Ακαδημαϊκές δημοσκοπήσεις

**Παραδείγματα:**

Επεξεργασμένα δεδομένα: Σε Κυβερνητικές εκδόσεις  
 Βιβλία  
 Περιοδικά

**Παραδείγματα:**

Επεξεργασμένα δεδομένα: Σε στατιστικές και σε εκθέσεις για τη βιομηχανία  
 Κυβερνητικές εκδόσεις  
 Εκδόσεις της Ευρωπαϊκής Ένωσης  
 Ρεπορτάζ εφημερίδων  
 Βιβλία  
 Περιοδικά

# Δευτερογενή δεδομένα τεκμηρίωσης

- Συχνά χρησιμοποιούνται σε συνδυασμό με πρωτογενή δεδομένα.
- Χρησιμοποιούνται επίσης από μόνα τους ή σε συνδυασμό με άλλες πηγές δευτερογενών δεδομένων.

## Δευτερογενή δεδομένα τεκμηρίωσης

υλικό κειμένου

Ανακοινώσεις, αλληλογραφία, πρακτικά συνεδριάσεων, εκθέσεις προς μετόχους, ημερολόγια, πρακτικά συνεδριάσεων, γραπτές καταγραφές ομιλιών και συνομιλιών, διοικητικά και δημόσια αρχεία και κείμενα από ιστοσελίδες

μη γραπτό υλικό

εγγραφές ήχου και βίντεο, εικόνες, σχέδια, ταινίες και τηλεοπτικά προγράμματα, DVD και CD-ROM, καθώς και ιστοσελίδες

διαδικτυακό υλικό, το οποίο παράγεται από κοινότητες του Internet

# Δημοσκοπικά δευτερογενή δεδομένα

συλλέγονται μέσω: απογραφής, συνεχούς/τακτικής δημοσκόπησης ή ειδικής δημοσκόπησης

Απογραφές	συνήθως διεξάγονται από κυβερνήσεις και είναι μοναδικές διότι, αντίθετα από τις δημοσκοπήσεις, η συμμετοχή είναι υποχρεωτική, με συνέπεια να παρέχουν πολύ μεγάλη κάλυψη του πληθυσμού
συνεχείς και τακτικές δημοσκοπήσεις	εκείνες που επαναλαμβάνονται. Περιλαμβάνουν δημοσκοπήσεις όπου τα δεδομένα συλλέγονται καθ' όλη τη διάρκεια του χρόνου ή άλλες που επαναλαμβάνονται σε τακτά χρονικά διαστήματα
ειδικές δημοσκοπήσεις	συνήθως εφάπαξ δημοσκοπήσεις και πολύ συγκεκριμένες ως προς το αντικείμενό τους. Περιλαμβάνουν δεδομένα από ερωτηματολόγια που έχουν διακινηθεί από ανεξάρτητους ερευνητές, καθώς και συνεντεύξεις που έχουν διεξαχθεί από οργανισμούς και κυβερνήσεις.

## Δευτερογενή δεδομένα πολλαπλών πηγών

μπορεί να προκύψουν εξ ολοκλήρου από δεδομένα τεκμηρίωσης ή δημοσκόπησης, ή μπορεί να είναι συνδυασμός των δύο.

διαφορετικά σύνολα δεδομένων έχουν συνδυαστεί για να διαμορφώσουν ένα άλλο σύνολο δεδομένων πριν την πρόσβασή σας σε αυτό.

# Πλεονεκτήματα δευτερογενών δεδομένων

<ul style="list-style-type: none"><li>• Λιγότερες απαιτήσεις πόρων</li></ul>	<ul style="list-style-type: none"><li>• κύριο πλεονέκτημα η τεράστια εξοικονόμηση πόρων, χρόνου και χρημάτων</li><li>• Θα έχετε περισσότερο χρόνο να ασχοληθείτε με τους θεωρητικούς σας στόχους και άλλα ουσιαστικά ζητήματα</li></ul>
<ul style="list-style-type: none"><li>• Διευκολύνονται οι διαχρονικές μελέτες</li></ul>	<ul style="list-style-type: none"><li>• Για πολλά ερευνητικά έργα, οι χρονικοί περιορισμοί σημαίνουν πως τα δευτερογενή δεδομένα παρέχουν τη μόνη δυνατότητα υλοποίησης διαχρονικών μελετών.</li><li>• είτε με τη δημιουργία ενός δικού σας είτε με τη χρήση ενός υπάρχοντος συνόλου δεδομένων πολλαπλών πηγών .</li></ul>
<ul style="list-style-type: none"><li>• Μπορούν να προσφέρουν συγκριτικά και εξειδικευμένα δεδομένα</li></ul>	<ul style="list-style-type: none"><li>• Συχνά είναι χρήσιμο να συγκρίνετε πρωτογενή δεδομένα, που έχετε συλλέξει, με δευτερογενή δεδομένα.</li><li>• Μπορείτε, έτσι, να τοποθετήσετε τα δικά σας ευρήματα εντός ενός πιο γενικού πλαισίου ή, εναλλακτικά, να τριγωνίσετε τα δικά σας ευρήματα</li></ul>
<ul style="list-style-type: none"><li>• Μπορούν να οδηγήσουν σε απρόβλεπτες ανακαλύψεις</li></ul>	<ul style="list-style-type: none"><li>• Η εκ νέου ανάλυση δευτερογενών δεδομένων μπορεί να οδηγήσει σε απρόβλεπτες ή απροσδόκητες νέες ανακαλύψεις.</li><li>• Π.χ. η σχέση μεταξύ καπνίσματος και καρκίνου του πνεύμονα</li></ul>
<ul style="list-style-type: none"><li>• Μονιμότητα</li></ul>	<ul style="list-style-type: none"><li>• τα δευτερογενή δεδομένα είναι συχνά μόνιμα διαθέσιμα σε μορφή που μπορεί να ελεγχθεί σχετικά εύκολα από άλλους</li></ul>

# Μειονεκτήματα δευτερογενών δεδομένων

- |   |  |
|---|--|
| <ul style="list-style-type: none"><li>Μπορεί να έχουν συλλεχθεί για κάποιο σκοπό που δεν ταιριάζει με τις ανάγκες σας</li></ul> | <ul style="list-style-type: none"><li>τα δευτερογενή δεδομένα συνήθως έχουν συλλεχθεί για κάποιον άλλο συγκεκριμένο σκοπό που διαφέρει από το δικό σας</li><li>τα υπάρχοντα δεδομένα μπορεί να μην είναι πρόσφατα, ή να διαφέρουν οι μέθοδοι συλλογής των αρχικών δεδομένων, τα οποία στη συνέχεια συγχωνεύτηκαν, ώστε να διαμορφώσουν το σύνολο των δευτερογενών δεδομένων που σκοπεύετε να χρησιμοποιήσετε</li></ul> |
| <ul style="list-style-type: none"><li>Η πρόσβαση μπορεί να είναι δύσκολη ή δαπανηρή</li></ul>                                   | <ul style="list-style-type: none"><li>αν τα δεδομένα έχουν συλλεχθεί για εμπορικούς λόγους, η απόκτηση πρόσβασης μπορεί να είναι δύσκολη ή δαπανηρή</li><li>είτε με τη δημιουργία ενός δικού σας είτε με τη χρήση ενός υπάρχοντος συνόλου δεδομένων πολλαπλών πηγών .</li></ul>  |
| <ul style="list-style-type: none"><li>Μη ουσιαστικός έλεγχος της ποιότητας</li></ul>  | <ul style="list-style-type: none"><li>απαιτείται προσοχή και οι πηγές των δεδομένων πρέπει να αξιολογούνται προσεκτικά</li></ul>   |
| <ul style="list-style-type: none"><li>Ο αρχικός σκοπός μπορεί να επηρεάσει την παρουσίαση</li></ul>                             | <ul style="list-style-type: none"><li>όταν χρησιμοποιείτε δεδομένα που έχουν ήδη παρουσιαστεί πρέπει να γνωρίζετε το σκοπό της συλλογής και επεξεργασίας και τον αντίκτυπο που αυτός έχει στον τρόπο παρουσίασης των δεδομένων</li></ul>   |



# Αξιολόγηση πηγών δευτερογενών δεδομένων

**1. Αποτίμηση της συνολικής καταλληλότητας των δεδομένων για το ερευνητικό ερώτημα (ερωτήματα) και τους στόχους σας**

*Δώστε ιδιαίτερη προσοχή:*

- στην εγκυρότητα της μέτρησης
- στην κάλυψη, συμπεριλαμβανομένων των μεταβλητών που δεν έχουν μετρηθεί

**2. Αξιολογήστε ακριβώς την καταλληλότητα των δεδομένων για τις αναλύσεις που απαιτούνται ώστε να απαντήσετε στο ερευνητικό σας ερώτημα (ερωτήματα) και να επιτύχετε τους στόχους σας**

*Δώστε ιδιαίτερη προσοχή:*

- στην εγκυρότητα
- στην αξιοπιστία
- στη μεροληψία της μέτρησης

**3. Κρίνετε αν θα χρησιμοποιήσετε τα δεδομένα βάσει εκτίμησης κόστους/οφέλους, σε σύγκριση με εναλλακτικές πηγές**

**Αν θεωρείτε τα δεδομένα ως εντελώς ακατάλληλα, μην προχωρήσετε παραπέρα**

# Αποτίμηση πηγών



## Πλαίσιο 8.10

### Λίστα ελέγχου

#### Αποτίμηση των πηγών δευτερογενών δεδομένων

##### Συνολική καταλληλότητα

- ✓ Το σύνολο δεδομένων περιέχει τις πληροφορίες που χρειάζεστε για να απαντήσετε στο ερευνητικό σας ερώτημα (ή ερωτήματα) και να επιτύχετε τους στόχους σας;
- ✓ Τα μέτρα που χρησιμοποιήθηκαν ταιριάζουν μ' αυτά που χρειάζεστε;
- ✓ Το σύνολο δεδομένων προσεγγίζει ικανοποιητικά τα δεδομένα που πράγματι χρειάζεστε;
- ✓ Το σύνολο των δεδομένων καλύπτει τον πληθυσμό που σας ενδιαφέρει;
- ✓ Το σύνολο των δεδομένων καλύπτει τη γεωγραφική περιοχή που σας ενδιαφέρει;
- ✓ Μπορούν τα δεδομένα σχετικά με τον πληθυσμό που σας ενδιαφέρει να «καθαριστούν» από ανεπιθύμητα δεδομένα;
- ✓ Αφορούν τα δεδομένα τη σωστή χρονική περίοδο ή είναι κατάλληλα επικαιροποιημένα;
- ✓ Περιλαμβάνουν τα δεδομένα όλες τις μεταβλητές που χρειάζεστε για να απαντήσετε στο ερευνητικό σας ερώτημα (ή ερωτήματα) και να επιτύχετε τους στόχους σας;
- ✓ Οι μεταβλητές έχουν καθοριστεί σαφώς;

##### Ειδική καταλληλότητα

- ✓ Πόσο αξιόπιστο είναι το σύνολο των δεδομένων που σκοπεύετε να χρησιμοποιήσετε;
- ✓ Πόσο αξιόπιστη είναι η πηγή δεδομένων;
- ✓ Είναι σαφές τι είδους πηγή είναι;
- ✓ Τα διαπιστευτήρια της πηγής δεδομένων (συντάκτης, φορέας ή οργανισμός χρηματοδότησης) υποδεικνύουν την αξιοπιστία της;
- ✓ Έχουν τα δεδομένα σχετική δήλωση πνευματικών δικαιωμάτων;

- ✓ Υπάρχουν σχετικά δημοσιευμένα αρχεία;
- ✓ Η πηγή περιέχει στοιχεία επικοινωνίας για την απόκτηση περαιτέρω πληροφοριών;
- ✓ Έχει περιγραφεί σαφώς η μέθοδος;
- ✓ Αν εφαρμόστηκε η μέθοδος της δειγματοληψίας, ποια ήταν η διαδικασία και ποια ήταν τα σχετικά δειγματοληπτικά σφάλματα και ποσοστά απόκρισης;
- ✓ Ποιος ήταν υπεύθυνος για τη συλλογή ή την καταγραφή των δεδομένων;
- ✓ (Για δημοσκοπήσεις) Περιλαμβάνεται αντίγραφο του ερωτηματολογίου ή λίστα ελέγχου της συνέντευξης;
- ✓ (Για τα επεξεργασμένα δεδομένα) Είστε σίγουροι πως έγινε η ανάλυση και η επεξεργασία των δεδομένων;
- ✓ Υπάρχει πιθανότητα τα δεδομένα να περιέχουν μεροληψία στη μέτρηση;
- ✓ Ποιος ήταν ο αρχικός σκοπός για τον οποίο συλλέχθηκαν τα δεδομένα;
- ✓ Ποιο ήταν το κοινό-στόχος και ποια ήταν η σχέση του με αυτόν που τα συνέλεξε ή αυτόν που τα επεξεργάστηκε;
- ✓ Υπήρξαν καταγεγραμμένες αλλαγές στον τρόπο με τον οποίο τα δεδομένα μετρήθηκαν ή καταγράφηκαν, συμπεριλαμβανομένων και των αλλαγών ορισμού;
- ✓ Πόσο συνεπή είναι τα δεδομένα απ' αυτή την πηγή σε σύγκριση με δεδομένα από άλλες πηγές;
- ✓ Έχουν καταγραφεί τα δεδομένα με ακρίβεια;

##### Κόστος και οφέλη

- ✓ Ποιο είναι το οικονομικό και χρονικό κόστος απόκτησης αυτών των δεδομένων;
- ✓ Μπορούν να ληφθούν τα δεδομένα σε υπολογιστικό φύλλο, σε λογισμικό στατιστικής ανάλυσης ή σε επεξεργαστή κειμένου;
- ✓ Τα συνολικά οφέλη της χρήσης αυτών των πηγών δευτερογενών δεδομένων αντισταθμίζουν το σχετικό κόστος;

Πηγή: Εμπειρία των συγγραφέων, Blumberg και άλλοι (2008), Dale και άλλοι (1988), Dochartaigh (2007), Kervin (1999), Smith (2006), Stewart και Kamins (1993)