

Σεμινάριο Τελειοφοίτων



9- Δευτερογενή δεδομένα & σύνθετοι δείκτες

Δευτερογενή δεδομένα και ανάλυσή τους



Δευτερογενή δεδομένα

- ∞ Δεδομένα που έχουν ήδη συλλεχθεί στο παρελθόν για άλλους σκοπούς και ξαναχρησιμοποιούνται για τη δική σας έρευνα.
- ∞ Επειδή υπάρχουν εξαιρετικά πολλές πηγές δευτερογενών δεδομένων θα είναι ευκολότερο να χρησιμοποιήσετε τέτοιου είδους δεδομένα σε αντίθεση με τα πρωτογενή που πρέπει να συλλεχθούν ειδικά για το σκοπό της έρευνας.

Αξιολόγηση των πηγών

- ✂ Αν και τα δευτερογενή δεδομένα απαιτούν πολύ λιγότερο χρόνο από τα πρωτογενή πρέπει να αξιολογείται από πριν αν μπορούν να χρησιμοποιηθούν.
- **Διαθεσιμότητα:** είναι διαθέσιμα;
 - **Συνάφεια:** τα δεδομένα που υπάρχουν είναι συναφή με τις ανάγκες σας;
 - **Ακρίβεια:** χωρίς πρόσβαση στα πρωτογενή δεδομένα από τα οποία προήλθαν δεν είναι πάντοτε δυνατόν να προσδιοριστεί η ακρίβειά τους
 - **Επάρκεια:** αν όλα τα παραπάνω είναι εντάξει και πάλι τα δεδομένα μπορεί να μην επαρκούν για τις ανάγκες σας

Αξιολογήστε τα δεδομένα

- ☞ Οι απαντήσεις στις παρακάτω ερωτήσεις θα σας βοηθήσουν να αξιολογήστε καλύτερα την καταλληλότητα των δεδομένων:
- πόσο παλιά είναι τα δεδομένα, μήπως είναι ήδη παρωχημένα;
 - πόσο γρήγορα αλλάζουν;
 - για ποιο σκοπό συλλέχθηκαν; μήπως για να υποστηρίξουν πολιτικές ή άλλες σκοπιμότητες;
 - ποιος τα συνέλεξε;
 - ποιες πληροφορίες είναι διαθέσιμες;
 - πώς συλλέχθηκαν, είναι γνωστό;
 - είναι τα δεδομένα συμβατά με άλλες πληροφορίες που έχετε;

Εσωτερικές επιχειρησιακές πηγές

- ✎ Το λογιστικό σύστημα και το σύστημα ελέγχου μιας επιχείρησης ή ενός οργανισμού αποτελούν βασικές πηγές δεδομένων που μπορούν να ταξινομηθούν σε πέντε κύριες κατηγορίες:
 - λογιστικά βιβλία
 - εκθέσεις πωλήσεων
 - εσωτερικές εκθέσεις ειδικών
 - επιχειρησιακά δεδομένα
 - διάφορα αρχεία
- ✎ Τέτοιου είδους δεδομένα αφορούν τη συγκεκριμένη εταιρεία και μπορεί να μην είναι σε κατάλληλη μορφή ή να μην είναι επίκαιρα ή ακόμα να μην είναι κατάλληλα για χρήση εκτός της εταιρείας.

- ✎ Υπάρχει μεγάλη ποικιλία εξωτερικών πηγών πληροφόρησης:
 - Eurostat
(<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>)
 - Στοιχεία διεθνών, εθνικών και περιφερειακών κυβερνητικών οργανισμών
 - Εφημερίδες - περιοδικά
 - Ετήσιες εταιρικές εκθέσεις και ισολογισμοί
 - Εκθέσεις επαγγελματικών ενώσεων
 - Εκθέσεις εμπειρογνομόνων
 - Άλλες δημοσιευμένες πηγές

☞ Εμπορικές έρευνες, αξιολογήσεις και πάνελ.

- **Εμπορικές έρευνες:** βασίζονται σε πληροφορίες που συλλέγονται από καταναλωτές και θεωρούνται σχετικά αμερόληπτες πηγές
- **Σύνθετες (Omnibus) έρευνες:** οι πληροφορίες στις οποίες βασίζονται καλύπτουν ένα ευρύτερο φάσμα θεμάτων και συνεπώς η χρησιμότητά τους μπορεί να είναι περιορισμένη ειδικά για τη δική σας έρευνα
- **Έρευνες μέσων ενημέρωσης (Media):** γενικά αφορούν τον προγραμματισμό, παρακολούθηση και αξιολόγηση των μέσων ενημέρωσης
- **Αξιολογήσεις λιανικών πωλήσεων:** πληροφορίες πωλήσεων προς καταναλωτές μέσα από ένα πάνελ καταστημάτων λιανικής πώλησης
- **Πάνελ καταναλωτών:** δείγμα καταναλωτών που απαντούν σε ερωτήσεις σε τακτική βάση

Και αν οι πηγές δεν συμφωνούν;

- ✎ Δεν είναι ασυνήθιστο να συναντήσετε πηγές δεδομένων με αντικρουόμενες πληροφορίες. Τι πρέπει να κάνετε
 - Αξιολογήστε εάν η πηγή σας είναι έγκυρη ή όχι. Αν υπάρχει τεκμηρίωση των δεδομένων τότε ΟΚ. Αν απλά αναφέρονται αριθμοί τότε μπορεί να μην είναι ακριβείς.
 - Αν δεν μπορείτε να εντοπίσετε την αρχική πηγή των δεδομένων, αναζητήστε άλλες πηγές που καλύπτουν το θέμα και επίσης εξετάστε τα ευρέως διαδεδομένα συμπεράσματα.
 - Συμβουλευτείτε κάποιον ειδικό για να σας βοηθήσει να επιλύσετε ένα θέμα, να απαντήσει στις ερωτήσεις σας, ή να σας κατευθύνει προς τις απαντήσεις

Πλεονεκτήματα - μειονεκτήματα δευτερογενών πηγών

☞ Πλεονεκτήματα

- Ταχύτερη συλλογή και ανάλυση δεδομένων
- Αποφυγή αλληλεπικάλυψης προσπαθειών συλλογής δεδομένων
- Προσφέρονται για την ανάλυση τάσεων, δεδομένου ότι προσφέρουν δυνατότητες παρακολούθησης των αλλαγών στην πάροδο του χρόνου
- Συμπληρώνουν την πρωτογενή συλλογή δεδομένων

☞ Μειονεκτήματα

- Κάποιες φορές δεν βοηθούν στη κατανόηση των φαινομένων
- Δύσκολο να προσδιοριστεί η ποιότητα των δεδομένων
- Οι πηγές μπορεί να έρχονται σε σύγκρουση
- Εύκολο να χαθείτε στον τεράστιο όγκο δεδομένων
- Τα διαθέσιμα δευτερογενή δεδομένα συνήθως έχουν συγκεντρωθεί για σκοπούς άλλους από αυτούς της δικής σας έρευνας

Σύγκριση πρωτογενών και δευτερογενών δεδομένων

Συλλογή	Πρωτογενή	Δευτερογενή
<i>Σκοπός</i>	Για το συγκεκριμένο πρόβλημα	Για άλλα προβλήματα
<i>Διαδικασία</i>	Ειδικά σχεδιασμένη	Γρήγορη και σχετικά απλή
<i>Κόστος</i>	Σχετικά υψηλό	Σχετικά χαμηλό
<i>Χρόνος</i>	Μακρύς	Σύντομος

Πίνακας: A Comparison of Primary and Secondary Data

Πηγή: Malhotra and Birks

Συμπερασματικά

- ✎ Τα δευτερογενή δεδομένα μπορεί να αποτελέσουν πολύτιμες πηγές πληροφόρησης.
- ✎ Η αξιολόγηση και ανάλυση των δευτερογενών δεδομένων βοηθούν στην αντιμετώπιση των θεμάτων, στη διενέργεια συγκρίσεων καθώς και στο καθορισμό της κατεύθυνσης και του μεγέθους των αλλαγών
- ✎ Συμπληρώνουν, αλλά δεν αντικαθιστούν τα πρωτογενή δεδομένα

Σύνθετοι δείκτες

∞ 13 ∞

Τι είναι ένας σύνθετος δείκτης (ΣΔ);

- ✎ Ένας δείκτης καταγράφει μια τάση ή φαινόμενο που διαφορετικά δεν είναι ανιχνεύσιμα.
- ✎ Ένας δείκτης ποσοτικοποιεί και απλοποιεί πληροφορίες.
- ✎ **Η σύνθεση επί μέρους δεικτών δημιουργεί ένα σύνθετο δείκτη.**

Πλεονεκτήματα – Μειονεκτήματα ΣΔ

∞ Πλεονεκτήματα

- Μειώνουν τον όγκο ενός συνόλου δεικτών χωρίς να απορρίπτουν την ενσωματωμένη πληροφόρηση
- Αξιολογούν αλλαγές στην πάροδο του χρόνου
- Είναι ευκολότερο να ερμηνευθούν από πολλούς επί μέρους δείκτες χωριστά
- Συνοψίζουν περίπλοκα ζητήματα

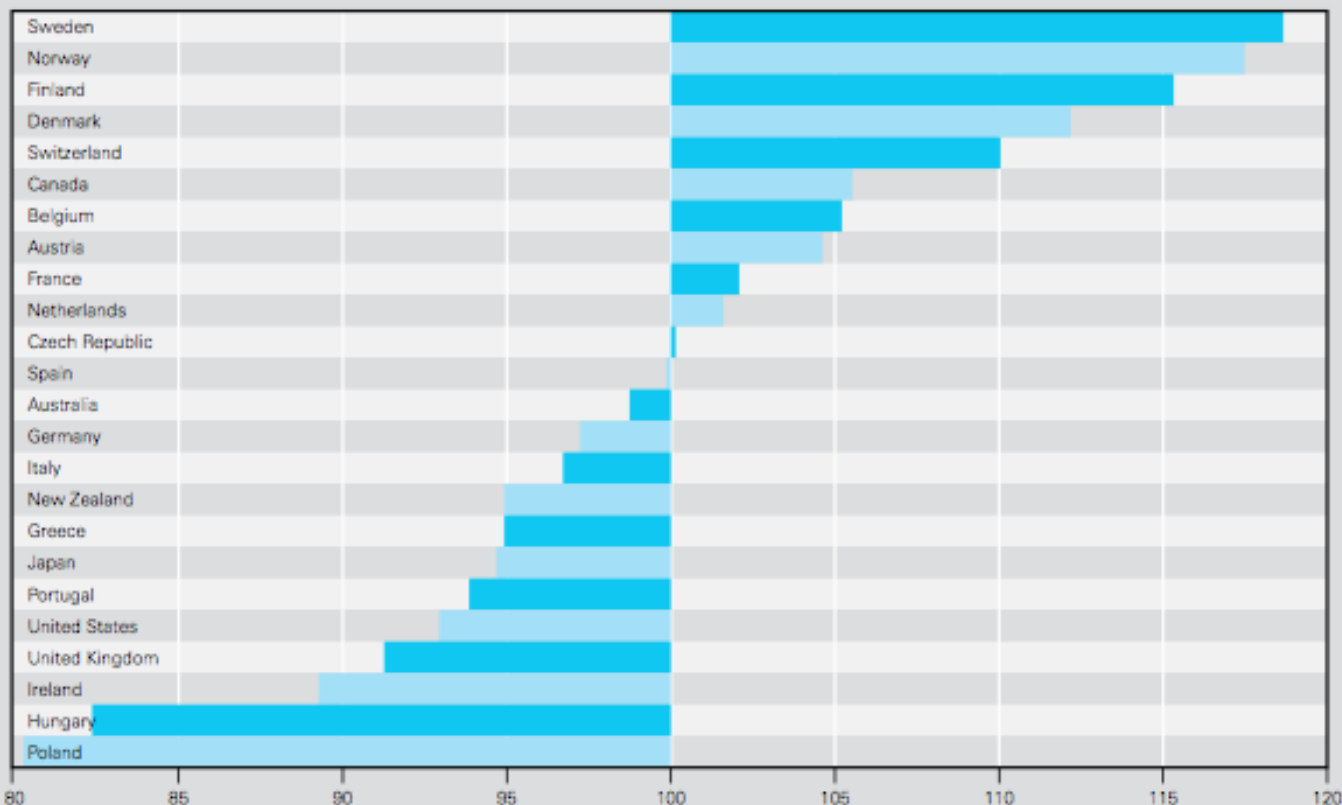
∞ Μειονεκτήματα

- Πιθανόν οδηγούν σε απλοϊκά συμπεράσματα
- Μπορεί να χρησιμοποιούνται λάθος αν ο τρόπος κατασκευής τους δεν είναι διαφανής και/ή δεν βασίζονται σε αξιόπιστα δεδομένα ή εννοιολογικές αρχές
- Η επιλογή των επί μέρους δεικτών και η βαρύτητά τους πρέπει να εξεταστεί ως προς τη θεωρητική τους ορθότητα.

Παράδειγμα ΣΔ: the child well-being index

- ✎ Ο δείκτης **Ευημερίας παιδιών και νέων της Unicef** ([Child and Youth Well-Being Index](#)) μετρά τεκμηριωμένα την τάση στο χρόνο της ποιότητας ζωής ή ευημερίας των παιδιών και των νέων.
- ✎ Αποτελεί συνδυασμό ετήσιων χρονοσειρών πολλών επί μέρους δεικτών σχετικά με την ευημερία των παιδιών και των νέων.
- ✎ Καταδεικνύει τη συνολική κατεύθυνση της αλλαγής (βελτίωση ή επιδείνωση) ανά χώρα.

Παράδειγμα ΣΔ: the child well-being index



Note: Each country has been placed on a scale determined by the average score for the group as a whole. The unit used is the standard deviation (the average deviation from the average). To ease interpretation, the results are presented on a scale with a mean of 100 and a standard deviation of 10.

UNICEF (2007)

ΤΕΙ Κρήτης/Τμήμα Λ&Χ/Δρ Ευάγγελος
Τσουκάτος

Παράγοντες δημιουργίας ενός ΣΔ

- ✧ Θεωρητικό πλαίσιο
- ✧ Επιλογή – Συλλογή δεδομένων
- ✧ Ομογενοποίηση
- ✧ Στάθμιση
- ✧ Σύνθεση

- ✎ Απαραίτητο για να υπάρξει ξεκάθαρη κατανόηση και ορισμός του πολυδιάστατου φαινομένου στο οποίο αναφέρεται ο δείκτης
- ✎ Παρέχει τη βάση για την επιλογή και συνδυασμό επί μέρους δεικτών σε ένα κατάλληλο ΣΔ για ένα συγκεκριμένο σκοπό

Επιλογή – Συλλογή δεδομένων

- ✎ Θα πρέπει να βασίζεται σε αναλυτική αξιοπιστία, μετρησιμότητα, κάλυψη της χώρας/περιοχής, καταλληλότητα επί μέρους δεικτών και τη μεταξύ τους σχέση.
- ✎ Ποιότητα διαθέσιμων δεικτών.
 - Χρησιμοποιήστε παραπλήσιους δείκτες όταν δεν υπάρχουν άμεσα διαθέσιμα στοιχεία
 - Εκτίμηση – συμπλήρωση ελλειπουσών τιμών

☞ Αφορά:

- επεξεργασία των επί μέρους δεικτών ώστε να γίνουν συγκρίσιμες οι τιμές τους αλλά με σεβασμό του θεωρητικού πλαισίου και των ιδιοτήτων των δεδομένων
- εξάλειψη διαφορετικών μονάδων μέτρησης
- αντιμετώπιση της παρουσίας ακραίων τιμών
- πιθανές προσαρμογές κλίμακας
- επεξεργασία – τροποποίηση ασύμμετρων δεικτών

Μέθοδοι ομογενοποίησης

Κατάταξη (ranking)	Η απλούστερη μέθοδος
Τυποποίηση (Standardisation)	Μετατρέπει τις μεταβλητές σε κοινή κλίμακα με μέση τιμή μηδέν και τυπική απόκλιση ένα.
Min-Max	Εξομαλύνει τις μεταβλητές ώστε να εκφράζονται στο κλειστό διάστημα $[0, 1]$ αφαιρώντας από κάθε τιμή της μεταβλητής την ελάχιστη τιμή και διαιρώντας με το εύρος των τιμών.
Απόσταση από σημείο αναφοράς (Distance to a reference)	Μετράται η σχετική απόσταση κάθε τιμής από ένα σημείο αναφοράς.
Κατηγορική κλίμακα (Categorical scale)	Αντιστοιχίζει τα αποτελέσματα σε κατηγορίες που μπορεί να εκφράζονται αριθμητικά (1, 2,...), ποιοτικά ("πλήρως", "εν μέρει", "καθόλου"), ή με βάση τα εκατοστημόρια της κατανομής της μεταβλητής.

Παράδειγμα: ομογενοποίηση με min-max

∞ Για θετικούς δείκτες
όπου περισσότερο = καλύτερο

$$I = \frac{(X_i - X_{\min})}{(X_{\max} - X_{\min})}$$

∞ Για αρνητικούς δείκτες
όπου λιγότερο = καλύτερο

$$I = \frac{(X_{\max} - X_i)}{(X_{\max} - X_{\min})}$$

- ✎ Αφορά την επιλογή κατάλληλων συντελεστών στάθμισης με βάση τόσο το θεωρητικό πλαίσιο όσο και τις ιδιότητες των δεδομένων.
- ✎ Οι συντελεστές στάθμισης συνεισφέρουν σημαντικά στην ολική συμπεριφορά του ΣΔ.
- ✎ Οι συντελεστές στάθμισης μπορεί να προέρχονται από στατιστικές επεξεργασίες, γνώμες ειδικών, προτεραιότητες πολιτικής ή τη θεωρητική βάση.
- ✎ Όποια μέθοδο και αν επιλέξετε η στάθμιση δεν είναι άλλο παρά αξιολογική κρίση.

- ✎ Οι ομογενοποιημένες τιμές των επί μέρους δεικτών συνθέτονται σταθμικά στο σύνθετο δείκτη. Π.χ.
 - $\Sigma\Delta = \Sigma(\text{συντελεστής στάθμισης} * \text{επί μέρους δείκτης})$
- ✎ Άλλες προσεγγίσεις περιλαμβάνουν γεωμετρική σύνθεση και ανάλυση πολλαπλών κριτηρίων



the sustainable cities index

ranking the largest 20 British cities

Σκοπός των δεικτών ποιότητας ζωής

- ✂ **Ανταγωνιστικότητα:** προσέλκυση κεφαλαίων και εργαζομένων σε μια πόλη ή περιφέρεια.
- ✂ *‘Οι άνθρωποι αγαπούν τις λίστες, ιδιαίτερα αυτές που κατατάσσουν ένα μέρος ή ένα άτομο πάνω από άλλα’ (Rogerson, 1999).*

Monday, March 09, 2009

BRIGHTON BUSINESS the Business Forum and the Economic Partnership in Brighton & Hove

Homepage ■ BCRP Intranet ■ Funding ■ Events ■ Partners ■ Knowledgebase ■ Links ■ About us ■ Contact us

Key projects

- Crime reduction
- Workforce development
- Transport
- Business & climate change
- Property & Development sites

News - 16 November 2008

Bristol displaces Brighton as the most sustainable city

Forum for the Future's index has placed Bristol at the head of the league tables for the top 20 most sustainable cities in Britain, pushing Brighton into second place.

[Email this to a friend](#) Cities are ranked according to three categories - environment, quality of life and future-proofing or readiness to face the challenges ahead.

Bristol came top because of a significant rise in its recycling and composting rates. Brighton was rated highest in terms of both quality of life and future-proofing but it was not enough to maintain its 2007 first place ranking..

(photo source: Ian Britton)

RECESSION

Δείκτες ποιότητας ζωής: δεδομένα για το ΗΒ

Neighbourhood Statistics - Home Page - Windows Internet Explorer

http://www.neighbourhood.statistics.gov.uk/dissemination/LeadHome.do;jessionid=ac1f930d30d7117649a309f444ada56dda

Neighbourhood Statistics - Home Page

Office for National Statistics UK Snapshot | Neighbourhood | Economic

Neighbourhood Statistics

FIND STATISTICS FOR AN AREA >More about areas

This search allows you to find **detailed statistics** within specific geographic areas, for example in **neighbourhood regeneration**.

- 1 First enter the name of an area OR full postcode:**

e.g. Clapham Park e.g. N51JP
- 2 Then select the type of area you need statistics for:**
☒ Local Authority
☐ Ward
☐ New Deal For Communities
☐ Super Output Area
[More areas](#) ▼
- 3 Search**

NEIGHBOURHOOD SUMMARY

This search allows you to find **local neighbourhood**. If you live in a neighbourhood you live or work in.

Enter a full postcode from

e.g. N51JP

Search

I WANT TO:

[view or download data by topic](#) [create a custom table, chart or map](#)
[explore analysis, training and guidance](#)

LATEST NEWS

[26 February 2009 - Concept](#)
[20 February 2009 - Withdrawal](#)
[More news stories...](#)

About Neighbourhood Statistics | News | User Guide | FAQ | Glossary | Accessibility | Developer API

© Crown Copyright | Contact Us | Terms and Conditions | Privacy Statement | Directgov

Κατασκευή ΣΔ: ένα παράδειγμα

(οι πόλεις είναι φανταστικές)

	Positive	Negative	Negative	Positive	Positive
	Health	Worklessness	Crime	Education	Environment
	Life expectancy [Years]	Unemployment [%]	Violent crime [per 1000 popn]	Ave. L3 QCA score [No.]	Green space [m2 per capita]
Ambridge	82	9	46.8	850	105
Borchester	79	5	37.4	570	120
Coketown	77	3	17.5	780	98
Slagness	81	8	16.0	833	89
Trumpton	75	7	49.1	640	69

- ☞ Τρεις δείκτες είναι θετικοί (περισσότερο = καλύτερο):
 - Υγεία (Health), Εκπαίδευση (Education) και Περιβάλλον (Environment)
- ☞ Δύο δείκτες είναι αρνητικοί (περισσότερο = χειρότερο):
 - Ανεργία (Worklessness) και Έγκλημα (Crime)
- ☞ Αλλά οι επί μέρους δείκτες είναι σε διαφορετικές μονάδες, και πρέπει να **ομογενοποιηθούν**

Ομογενοποίηση με min-max: υγεία - θετικός δείκτης

Χρησιμοποιούμε τον τύπο : $I = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$

$$X_{\max} = 82$$

$$X_{\min} = 75$$

$$X_{\text{Borchester}} = 79$$

Το εξομαλυμέν ο σκορ του Borchester είναι :

$$I = \frac{79 - 75}{82 - 75} = 0.571$$

Εξομάλυνση με min-max: έγκλημα –αρνητικός δείκτης

Χρησιμοποιούμε τον τύπο : $I = \frac{X_{\max} - X_i}{X_{\max} - X_{\min}}$

$$X_{\max} = 49.1$$

$$X_{\min} = 16.0$$

$$X_{\text{Borchester}} = 37.4$$

Η εξομαλυμένη τιμή του Borchester είναι :

$$I = \frac{49.1 - 37.4}{49.1 - 16.0} = 0.353$$

Ομογενοποιημένες ενδείξεις για τις 5 πόλεις

- ☞ Όλοι οι δείκτες μετριοούνται τώρα στην ίδια κλίμακα [0,1]
- ☞ Οι αρνητικοί δείκτες (ανεργία και έγκλημα) έχουν αναστραφεί:
 - υψηλότερη ανεργία = 0, χαμηλότερη ανεργία = 1.

	Health	Worklessness	Crime	Education	Environment
Ambridge	1.000	0.000	0.069	1.000	0.706
Borchester	0.571	0.667	0.353	0.000	1.000
Coketown	0.286	1.000	0.955	0.750	0.569
Slagness	0.857	0.167	1.000	0.939	0.392
Trumpton	0.000	0.333	0.000	0.250	0.000

Συντελεστές στάθμισης

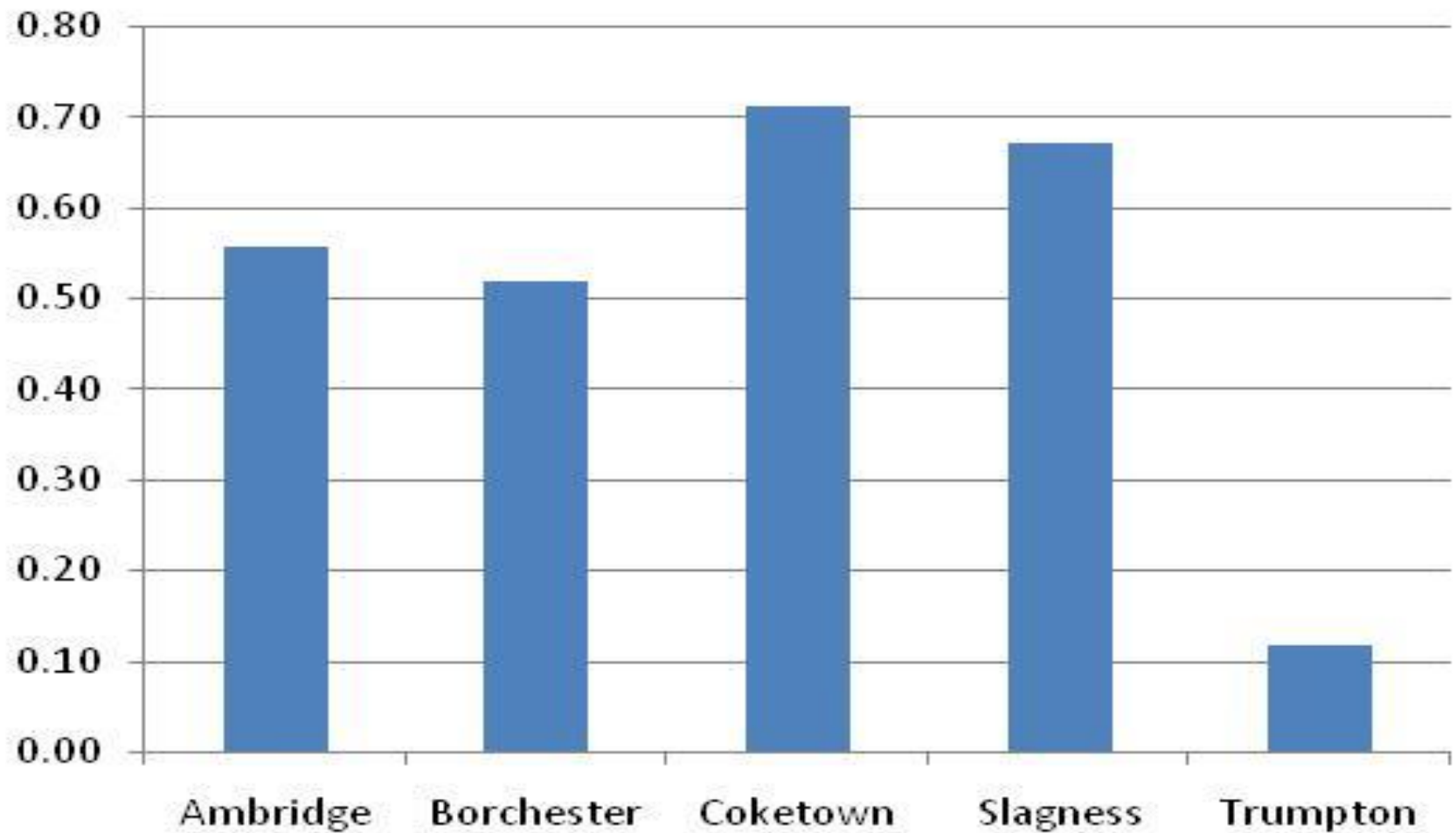
- ✧ Η σχετική σημασία των δεικτών μετριέται από τους συντελεστές στάθμισης.
- ✧ Αυτοί προκύπτουν:
 - άτυπη απόφαση
 - στατιστική επεξεργασία (π.χ. πολυκριτήρια ανάλυση)
- ✧ Θα εξετάσουμε δύο δυνατότητες:
 - Ίσοι συντελεστές
 - Άνισοι συντελεστές (μεγαλύτερη σημασία στην υγεία)

	Υγεία	Ανεργία	Έγκλημα	Εκπαίδευση	Περιβάλλον
Ίσοι	0.2	0.2	0.2	0.2	0.2
Άνισοι	0.6	0.1	0.1	0.1	0.1

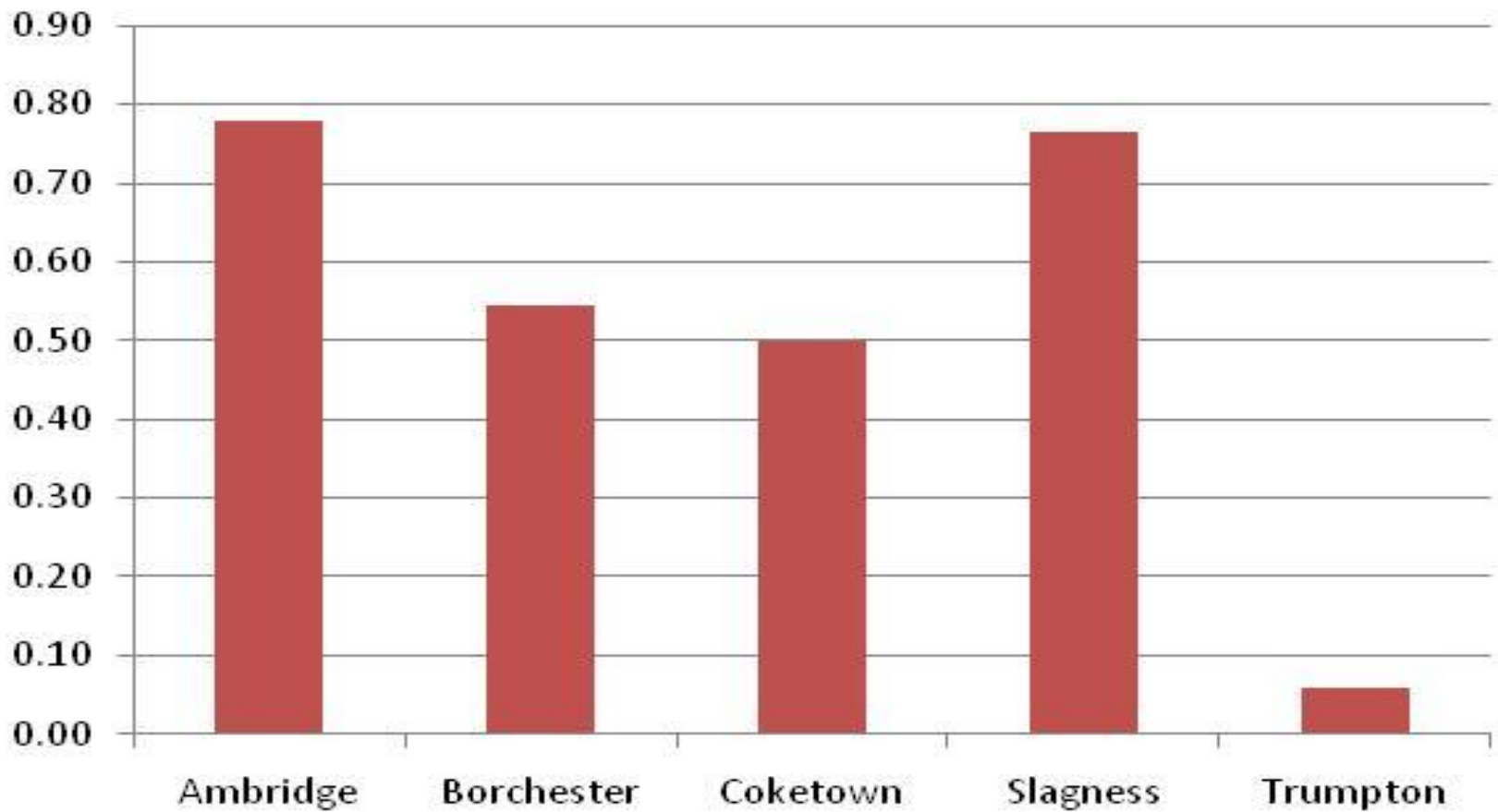
- ✧ Πολλαπλασιάζουμε τις εξομαλυμένες τιμές για κάθε πόλη επί τους αντίστοιχους συντελεστές στάθμισης και προσθέτουμε.
- ✧ Η κατάταξη των πόλεων εξαρτάται από τους συντελεστές βαρύτητας.

	Quality of Life scores	
	Equal weights	Unequal weights
Ambridge	0.56	0.78
Borchester	0.52	0.54
Coketown	0.71	0.50
Slagness	0.67	0.76
Trumpton	0.12	0.06

Quality of Life scores for 5 cities
Assuming equal weights for component indicators



Quality of Life scores for 5 cities
Assuming higher weight given to Health



- ✎ Οι αριθμοδείκτες χρησιμοποιούνται για τη μέτρηση πολύπλοκων και πολυδιάστατων φαινομένων.
 - Ποιότητα ζωής
 - Ανθρώπινη ευημερία
 - Βιωσιμότητα
 - Κοινωνική υστέρηση
 - Κλπ
- ✎ Η μεθοδολογία κατασκευής ΣΔ περιλαμβάνει:
 - Επιλογή
 - Εξομάλυνση
 - Επιλογή συντελεστών στάθμισης
 - Σύνθεση