

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΜΥΕ030/
ΠΛΕ045

ΠΡΟΧ. ΘΕΜΑΤΑ ΤΕΧΝΟΛΟΓΙΑΣ & ΕΦΑΡΜΟΓΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ
Π. Βασιλειάδης

ΑΝΟΙΞΗ 2026

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΑΣΚΗΣΗ

Ημερομηνία Παράδοσης: Τρίτη 26-05-2026, 16.59

Η προγραμματιστική άσκηση για το μάθημα είναι υποχρεωτική και αφορά τη σχεδίαση, υλοποίηση και ρύθμιση ενός ολοκληρωμένου πληροφοριακού συστήματος. Η εργαστηριακή άσκηση προσφέρει **3 μονάδες** στον τελικό βαθμό του μαθήματος. Φυσικά, πρέπει να πιάσετε τουλάχιστον τη βάση στην εργασία, όπως και στο διαγώνισμα. Σε περιπτώσεις εξαιρετικών εργασιών, η επίδοση επιβραβεύεται με bonus που μπορεί να φτάσει ως και μία μονάδα στον τελικό βαθμό.

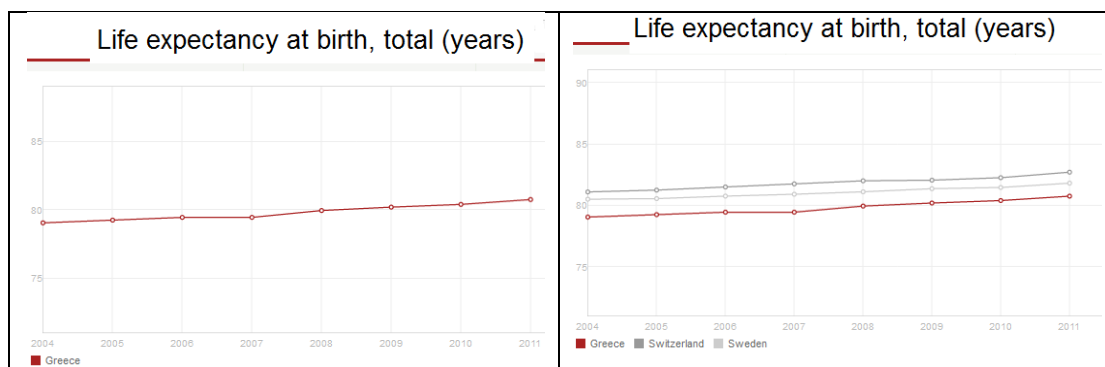
Οι προθεσμίες είναι ιερές.

Είναι υποχρεωτικό να υλοποιήσετε τουλάχιστον ένα σύστημα με σχεσιακό back-end business logic, και γραφική διαπροσωπεία καθώς και μια τελική αναφορά.

Η προγραμματιστική άσκηση του μαθήματος ασχολείται με το πρόβλημα (α) της **ενοποίησης** και (β) της **οπτικοποίησης δεδομένων**. Ο στόχος των τεχνικών ενοποίησης είναι να εντάξουν πληροφορίες με διαφορετική δομή και κωδικοποίηση σε ένα ενιαίο σχήμα, ώστε να είναι ενιαία επερωτήσιμες. Ο στόχος των τεχνικών οπτικοποίησης είναι να δώσουν στον χρήστη την πληροφορία με τρόπο που αναδεικνύει οπτικά ιδιότητες, τάσεις και πρότυπα που βρίσκονται κρυμμένα στα δεδομένα. *Why bother? Κυρίως, γιατί ζούμε σε μια εποχή που έχουμε όλο και πιο πολλά δεδομένα γύρω μας, και γίνεται όλο και πιο δύσκολο να τα αξιοποιήσουμε, ρωτώντας τα. Οι απαντήσεις στις ερωτήσεις πλέον δεν αρκούν: στους χρήστες πρέπει να παρουσιάζονται και ενδιαφέρουσες ιδιότητες εντός των δεδομένων.*

Γενικά περί Οπτικοποίησης

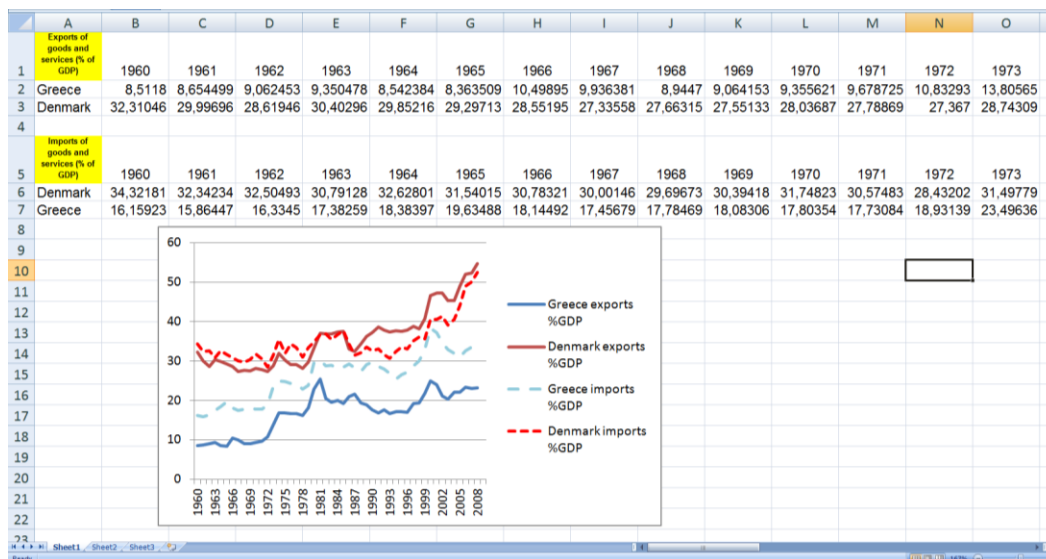
Linecharts. Αν θέλουμε να δείξουμε την εξέλιξη ενός ή περισσότερων δεικτών στο χρόνο, το πιο συχνά χρησιμοποιούμενο μέσο είναι οι line charts. Ο χρόνος απεικονίζεται στον άξονα των x και το μετρούμενο μέγεθος στον άξονα των y . Αν αντί για χρόνο έχουμε άλλο ποσό στον άξονα των x (π.χ., ο πληθυσμός μιας χώρας, η έκτασή της κλπ.) τότε εμπίπτουμε στη γενικότερη κατηγορία των trendlines. Εδώ έχουμε προσδόκιμο ζωής για (α) μία και (β) 3 χώρες στο χρόνο.



Οι υποκατηγορίες που μπορεί να έχουμε είναι:

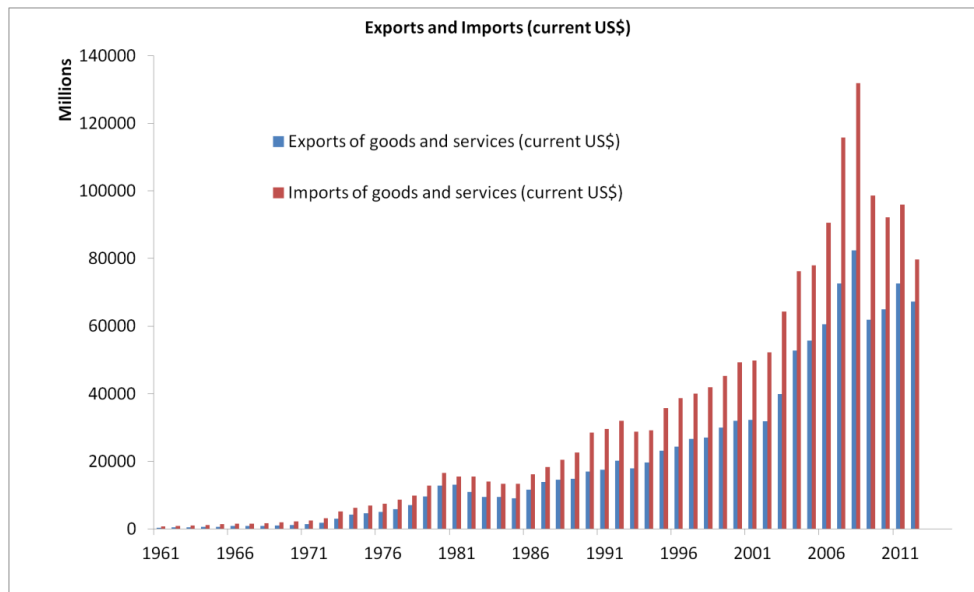
- Για k -το πλήθος χώρες, για ένα δείκτη, πώς εξελίσσεται στο χρόνο (απλή περίπτωση: 1 χώρα)
- Για k_c -το πλήθος χώρες και για k_m -το πλήθος δείκτες, πώς εξελίσσεται στο χρόνο ο καθένας

Είναι υποχρεωτικό σχεδιαστικά στον άξονα X να έχουμε ένα μέγεθος ισομορφικό στους ακέραιους (π.χ., χρόνος) και στον άξονα Y ένα μετρήσιμο μέγεθος ισομορφικό στους πραγματικούς αριθμούς.

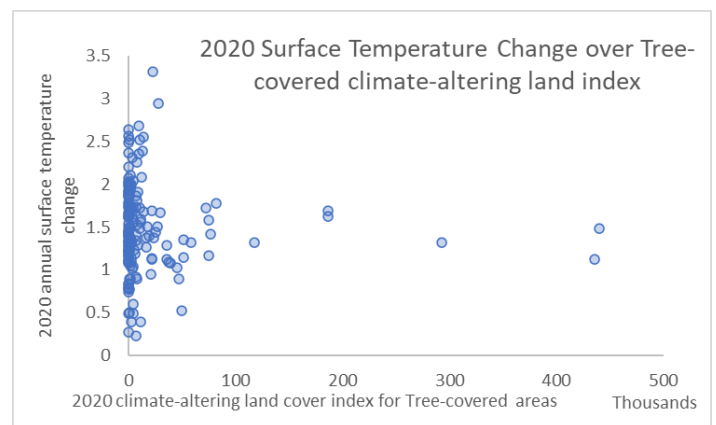


Bar charts. Η εν λόγω τεχνική χρησιμοποιείται για να συγκρίνει δύο ή περισσότερες μετρικές (y -axis) πάνω στις ίδιες τιμές του άξονα των x . Γενικά, μπορούμε να γενικεύσουμε το παραπάνω σε περισσότερες από 2 μετρικές, k -το πλήθος στη γενική περίπτωση, αλλά με μικρό k (σκεφθείτε πόσο άσχημο θα ήταν το διάγραμμα για την περίπτωση 2 δεικτών και 2 χωρών).

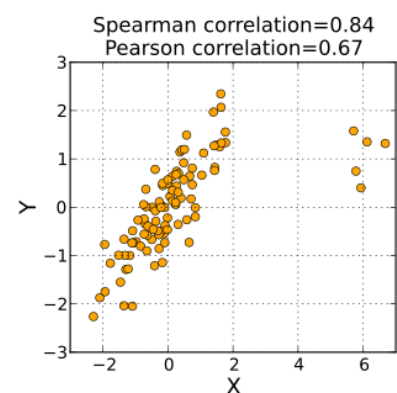
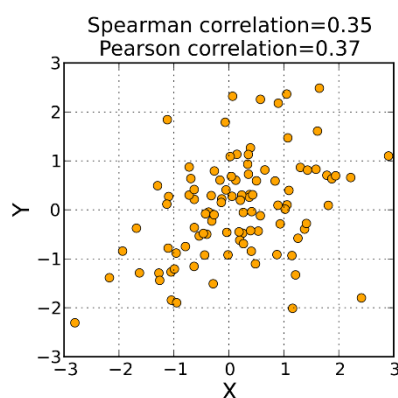
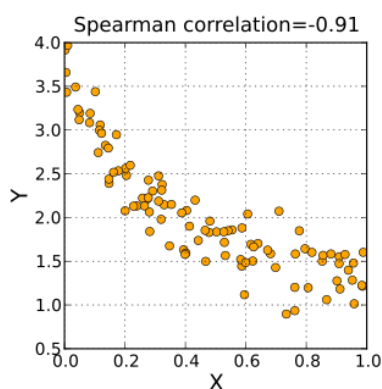
Τα bar charts είναι παντός καιρού διαγράμματα σε ότι αφορά τον άξονα X : με εξαίρεση το ότι υποφέρουν όταν υπάρχει «υπερπληθυσμός» γενικώς μπορούν να απεικονίσουν τα πάντα. Ο άξονας Y πρέπει να είναι μετρήσιμο μέγεθος. Ο μόνος πολύ βασικός περιορισμός είναι ότι ο άξονας Y πρέπει υποχρεωτικά να ξεκινάει από το 0.



Scatter Plots. Η εν λόγω τεχνική οπτικοποίησης συσχετίζει περισσότερες της μίας μετρικές και προσπαθεί να δείξει το βαθμό συσχέτισής τους. Για παράδειγμα, αν θέλουμε να δούμε πώς σχετίζεται η αύξηση της θερμοκρασίας με ένα δείκτη εδαφικής κάλυψης από δάση για όλες τις χώρες του κόσμου, κάνουμε ένα scatter plot. Είναι σημαντικό σχεδιαστικά ότι κάθε *reference "object"* είναι ένα σημείο με δύο *real-valued* μεγέθη: εδώ κάθε σημείο είναι μια χώρα, και έχει δύο διαστάσεις, την εδαφική κάλυψη (x-coordinate) και την αύξηση της θερμοκρασίας (y-coordinate).



Δείτε πώς ένα scatter plot αναδεικνύει την συσχέτιση δύο μεγεθών (εδώ: Spearman / Pearson correlation). Όσο πιο κοντά στο 0 η στατιστική συσχέτιση, τόσο πιο «τυχαία» είναι απλωμένα τα σημεία στο διάγραμμα.



Δεδομένα που θα χρησιμοποιήσουμε φέτος

Θα χρησιμοποιήσουμε κάποια σύνολα βιβλιογραφικών δεδομένων τα οποία θα βρείτε στο σχετικό φάκελο που εμφανίζεται στη σελίδα της εκφώνησης.

ΑΡΘΡΑ ΣΕ ΣΥΝΕΔΡΙΑ (Source: dblp)

id	authors (attn: separated by ' ')	booktitle	...	crossref	...	key	mdate	...	pages	...
98643	Anthony K. H. Tung Zhan Su Zhenjie Zhang	EDBT		conf/edbt/2012	...	conf/edbt/SuTZ12	09-06-2012		97-108	
172736	Dik Lun Lee Qijun Zhu Wang-Chien Lee	ICDE		conf/icde/2011	...	conf/icde/ZhuLL11	23-01-2012		279-290	

...	Title	title-bibtex	tt	url	year
...	Supporting top-K item exchange recommendations in large online communities.			db/conf/edbt/edbt2012.html#SuTZ12	2012
...	Collaborative caching for spatial queries in Mobile P2P Networks.			db/conf/icde/icde2011.html#ZhuLL11	2011

ΑΡΘΡΑ ΣΕ ΠΕΡΙΟΔΙΚΑ (Source: dblp)

id	authors	booktitle	...	ee	journal	key	mdate	...
2931365	Yiu-ming Cheung			...	IEEE Trans. Knowl. Data Eng.	journals/tkde/Cheung05a	10-05-2006	
3378487	Ismail Hakki Toroslu Pinar Senkul			...	Inf. Syst.	journals/is/SenkulT05	06-01-2006	

...	number	pages	...	title	title-bibtex	tt	url
	11	1583-1588		On Rival Penalization Controlled Competitive Learning for Clustering with Automatic Cluster Number Selection.	db/journals/tkde/tkde17.html#Cheung05a	17	2005
	5	399-422		An architecture for workflow scheduling under resource allocation constraints.	db/journals/is/is30.html#SenkulT05	30	2005

ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΣΥΝΕΔΡΙΩΝ (Source: iCore26)

ID	Title	Acronym	Source	Rank	DBLP	PrimaryFoR
528	Extending Database Technology	EDBT	ICORE2026	B	Yes	4605
986	International Conference on Data Engineering	ICDE	ICORE2026	A*	Yes	4605

Δείτε το σχετικό αρχιείακι για τα PrimaryFoR

ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ ΠΕΡΙΟΔΙΚΩΝ (Source: Kaggle)

Rank	Title	OA	Country	SJR index	Cite Score	H index	Best Quartile	...	Total Refs	Total Cites 3y	Citable Docs 3y	Cites /Doc 2y	Refs / Doc	...
910	IEEE Transactions on Knowledge and Data Engineering	FALSE	United States	2.358	12.4	190	Q1		22496	5269	646	7	27.2	
3935	Information Systems	FALSE	United Kingdom	0.976	7.9	92	Q1		6371	1072	238	5	45.51	

Στόχος

Θα χρειαστεί να υποστηρίξουμε ένα αναλυτή με έτοιμες αναφορές για διάφορα επί μέρους ερωτήματα που μπορεί να τον ενδιαφέρουν. **Ο τελικός σκοπός σας είναι να μπορείτε να υλοποιήσετε μια εφαρμογή ενοποίησης και οπτικής εξαγωγής συμπερασμάτων η οποία θα αξιοποιεί δεδομένα που θα έχουν ενσωματωθεί σε μια βάση δεδομένων. Στο τέλος θέλουμε: (1) τα δεδομένα να έχουν φορτωθεί σε μια (α) καθαρή και (β) επερωτήσιμη βάση δεδομένων, (2) να έχει αναπτυχθεί πληροφοριακό σύστημα με front- και back-end, καθώς και ενδιάμεση λογική, και (3) να υπάρχει διαδραστική εφαρμογή που να μπορεί να επιτρέψει την πλοήγηση στα δεδομένα.**

ΣΥΝΕΔΡΙΟ / ΠΕΡΙΟΔΙΚΟ

Για οποιοδήποτε συνέδριο ή περιοδικό ο αναλυτής να μπορεί να δει το προφίλ του, τουτέστιν

- Το σχετικό ranking καθώς και κάποια linecharts, που δείχνουν, για κάθε χρονιά, πόσα άρθρα δημοσιεύτηκαν, συνολικό αριθμό συγγραφέων, και, αρ. διακριτών συγγραφέων
- Στατιστικά, π.χ.: (*) πρώτη/τελευταία γνωστή χρονιά, (*) αρ. συγγραφέων που έχουν δημοσιεύσει, μέσος # συγγραφέων ανά άρθρο καθώς και ανά χρονιά, (*) σύνολο άρθρων, και αντίστοιχα με συγγραφείς, Μ.Ο. άρθρων ανά χρονιά

Επιπλέον, να υπάρχει η δυνατότητα να βγουν:

- Τα παραπάνω στατιστικά στοιχεία, αλλά με φίλτρο για το εύρος χρονιών
- Ένα report με τις πληροφορίες για όλα τα άρθρα του σχετικού συνεδρίου/περιοδικού σαν πινακάκι (αν έχει υπάρξει φίλτρο, μόνο για τις χρονιές του φίλτρου)

ΧΡΟΝΙΕΣ

Θέλουμε να μπορεί να δει ο αναλυτής:

- Προφίλ μιας χρονιάς, ήτοι, πόσα άρθρα δημοσιεύθηκαν, σε πόσα διακριτά περιοδικά και πόσα διακριτά συνέδρια, πόσοι συγγραφείς (συνολικά και διακριτά) δημοσίευσαν κλπ.
- Επίσης, για τη συγκεκριμένη χρονιά, τις σχετικές δημοσιεύσεις σαν πινακάκι, (και με φίλτρο για συνέδριο, περιοδικό ή συγγραφέα).

ΣΥΓΓΡΑΦΕΙΣ (θεωρούμε ότι δεν έχουμε συνωνυμίες, ελλείψει άλλων δεδομένων)

Για κάθε συγγραφέα, θέλουμε να μπορεί να δει ο αναλυτής το προφίλ του:

- πρώτη/τελευταία γνωστή χρονιά που δημοσίευσε, σύνολο άρθρων, Μ.Ο. άρθρων ανά χρονιά
- με το σχετικό linechart, ανά χρονιά, τον αρ. δημοσιευθέντων άρθρων (και από συνέδρια και από περιοδικά)

ΣΤΟΧΕΥΜΕΝΕΣ ΓΡΑΦΙΚΕΣ ΠΑΡΑΣΤΑΣΕΙΣ

Linecharts

- Έχοντας διαλέξει κάποια συνέδρια/περιοδικά: αρ. άρθρων (ή αρ. συγγραφέων ή ...) ανά χρονιά (ενδεχομένως και με φίλτρο για τις χρονιές)
- για κάθε κατηγορία συνεδρίων (PrimaryFor) ή περιοδικών (BestSubjectArea) (ενδεχομένως και με φίλτρο στην κατηγορία): αρ. συνεδρίων/περιοδικών που εντάσσονται σε αυτή ανά χρονιά

BarCharts

- Για κάθε συνέδριο/περιοδικό (ενδεχομένως και με φίλτρο για κάποια συνέδρια/περιοδικά): σύνολο δημοσιευθέντων άρθρων / Μ.Ο. άρθρων ανά χρονιά / Μ.Ο. συγγραφέων ανά χρονιά
- Για κάθε publisher περιοδικού (ή με φίλτρο ίσως για τους publishers):
 - αριθμό περιοδικών που εκδίδει συνολικά
 - για κάθε quartile: αρ. περιοδικών (x-axis: publisher, y-axis: count, 4 series Q1-Q4)

ScatterPlots

Συνδυάστε 2 μετρήσιμα μεγέθη, π.χ.,

- Οποιαδήποτε δύο από TotalDocs., TotalDocs3y, TotalRefs, TotalCites3y, CitableDocs3y, Cites/Doc2y, Refs/Doc, ... από τις κατατάξεις περιοδικών (κάθε περιοδικό είναι ένα σημείο)
- Μ.Ο. συγγραφέων σε σχέση με #άρθρων ανά χρονιά, για συνέδρια ή περιοδικά

WORKPLAN

Το project έχει τρεις φάσεις: (I) setup & προεπεξεργασία DBMS και δεδομένων, (II) υλοποίηση του κορμού μιας εφαρμογής και (III) τελειοποίηση της εφαρμογής και μελέτη.

ΦΑΣΗ I: ενοποίηση σε μία (1) ενιαία βάση δεδομένων

Η ενοποίηση έχει στόχο (α) να φτιαχτούν σωστά οι όποιοι *reference, single-version-of-the-truth, lookup tables* με lookup values (άρα με το σωστό πρωτεύον κλειδί), και (β) να μετασχηματισθούν οι υπόλοιπες factual πληροφορίες και να φορτωθούν σε *factual tables με την ουσία της πληροφορίας*, με τα σωστά πρωτεύοντα κλειδιά αφενός, και εξωτερικά κλειδιά στους lookup πίνακες αφετέρου, ώστε να έχουμε πίνακες με καθαρά δεδομένα και εγγυήσεις ορθότητας.

A. Σχεδίαση υπό το πρίσμα των προβλημάτων και ιδιαιτεροτήτων των δεδομένων. Μερικά ζητήματα σχεδίασης και φόρτωσης για τα δεδομένα που έχετε ανά χείρας:

- Συγγραφείς: οι συγγραφείς έρχονται σε λίστες, ενωμένοι με |. Πρέπει να φτιαχτεί πίνακας αναφοράς για τους συγγραφείς (θεωρήστε ότι δεν υπάρχουν συνωνυμίες). Προσοχή, η σχέση με τα άρθρα είναι N:M.
- Άρθρα σε περιοδικά και συνέδρια: έχουμε παρόμοια αλλά όχι ίδια δομή. Επίσης, τα πεδία χρησιμοποιούνται διαφορετικά (π.χ. booktitle & journal). Τι πίνακες και τι views θα φτιάξετε, και ποια τα tradeoff's?
- ΠΡΟΣΟΧΗ: άλλο ΆρθροΣεΠεριοδικό και άλλο Περιοδικό (ομοίως με συνέδρια)
- Invalid rows / cells: η παρουσία NULL values μπορεί να δυσκολέψει την φόρτωση των δεδομένων. Ο κώδικας προεπεξεργασίας των δεδομένων θα πρέπει να μετασχηματίζει / απομονώνει προβληματικές εγγραφές PRIN τη φόρτωση.
- Τι θα κάνετε με τους πίνακες των αξιολογήσεων? Χωριστοί πίνακες από τους πίνακες άρθρων και περιοδικών, ή ένας?
- Υποψήφιοι lookup tables για τους πίνακες των άρθρων: συγγραφείς, συνέδρια / περιοδικά, PrimaryFoR, BestSubjectArea, ...
- Non-matching values: δείτε ότι ειδικά για τα περιοδικά, πρέπει να χρησιμοποιήσετε regular expressions για να κάνετε match τα στοιχεία των άρθρων και τα στοιχεία των αξιολογήσεων
- Κλειδιά. Τι θα είναι primary key σε κάθε πίνακα? Τα πρωτεύοντα κλειδιά που είναι συμβολοσειρές είναι απολύτως ακατάλληλα για τη δουλειά.

Το να αφήσετε τη δομή των πινάκων ολόδια με τη δομή των πηγαίων αρχείων είναι λόγος αποτυχίας. Σε ό,τι αφορά τη σχεδίαση της βάσης:

- *Obligatorily, you need Lookup Tables with the correct lookup values and primary keys!*
- *For the database that will ultimately hold your single-version-of-the-truth Factual Tables (i.e., not the ones that you will use for loading and transformation purposes, but for the ones hosting the clean v. of the data that will ultimately be queried by the users for info): Primary and Foreign keys to the correct Lookup Tables are obligatory!*
- *Obligatorily, you need to clean, validate and transform the data before loading into the db!*

Κάθε ομάδα πρέπει να προβεί στις παρακάτω ενέργειες για το φόρτωμα:

1. Στήσιμο της βάσης και ενός γραφικού εργαλείου διαχείρισης (π.χ., MySQL & MySQL Workbench) στο μηχάνημά σας.
2. Download το κομμάτι των δεδομένων που θα χρησιμοποιήσετε – τα αντίστοιχα αρχεία δηλαδή.
3. Δημιουργήστε το σχήμα της βάσης για τα δεδομένα σας – όπως θα συζητήσουμε στο μάθημα (βλ. υποδείξεις στο παρακάτω παράδειγμα). Χρησιμοποιήστε InnoDB τύπο αποθήκευσης στη MySQL.
4. Δημιουργία scripts που μετατρέπουν τα εισερχόμενα αρχεία σε αρχεία φόρτωσης δεδομένων – αρχεία δηλαδή, στα οποία τα δεδομένα είναι έτοιμα προς φόρτωση
5. Δημιουργία scripts φόρτωσης των αρχείων φόρτωσης (π.χ., δείτε την εντολή LOAD DATA INFILE στη MySQL)
6. Φόρτωση των αρχείων και εξαγωγή backup της βάσης

B. Extract – Transform – Load (ETL) the data. Τα πηγαία δεδομένα ΔΕΝ είναι στη δομή της τελικής τους μορφής μέσα στη βάση δεδομένων => πιθανώς θα πρέπει να μετασχηματισθούν.

- Μια πιθανή μέθοδος είναι χρησιμοποιήσετε κάποιους βοηθητικούς πίνακες υποδοχής, ώστε να φέρετε τα δεδομένα στη βάση, και μετά από κάποιες μετατροπές (με scripts / views / ... ό,τι χρειαστεί τέλος πάντων) να τροφοδοτήσετε τους πίνακες που θα χρησιμοποιηθούν για τις ερωτήσεις των χρηστών.
- Ο πιο συνήθης τρόπος είναι να κατασκευασθούν κάποιες ροές εργασίας ETL που κάνουν αυτή τη δουλειά. Στο πλαίσιο της εργασίας αυτής, μπορείτε να το κάνετε εύκολα με κάποια script. Μπορείτε όμως και να χρησιμοποιήσετε κάποιο σχετικό εργαλείο.

Θα χρειαστεί να σχεδιάσετε και να τεκμηριώσετε καλά τη διαδικασία μετασχηματισμού και φόρτωσης των δεδομένων. Δείτε το [συνοδευτικό κείμενο](#) και τη [σελίδα project](#) στο [site του μαθήματος](#) για ETL που καλύπτει (α) τη διαδικασία σχεδίασης και (β) εργαλεία που υποστηρίζουν την εκτέλεση ETL ροών.

Υποχρεωτικά στο τέλος της διαδικασίας πρέπει να έχετε (i) DDL scripts for the creation of the database schema if needed, (ii) ETL scripts for the automated loading of the data in the database, (iii) backup of the loaded database. / if there is ever a failure, you (in fact: anyone) need to be able to recreate the database automatically */*

Γ. Query design. Με βάση το σχήμα της βάσης και το τι φορτώθηκε, πρέπει να μπορείτε να απαντήσετε τα βασικά ερωτήματα που έχει η εφαρμογή σας. *[Άρα δοκιμάστε σε ένα περιβάλλον επερώτησης και κρατήστε σαν script αρχεία τις ερωτήσεις που θα χρειαστείτε στη συνέχεια.](#)* Αυτό θα αποδειχθεί πιο δύσκολο απ' όσο φαίνεται, κρατήστε χρόνο και γι' αυτό.

Θέλουμε ιδανικά όλα τα ερωτήματα, κι αν όχι όλα, όσο δυνατόν περισσότερα, να απαντώνται με SQL views ή έστω direct queries ΚΑΙ ΟΧΙ «φέρνω όλα τα δεδομένα σε μια συλλογή στην κύρια μνήμη και κάνω τη δουλειά εκεί». Άρα **το back-end data processing πρέπει να γίνεται ΜΕΣΑ στο DBMS**. Αυτό δεν μπορεί να γίνει καθολικά, αλλά η ιδέα είναι να μην πετάξουμε το μπαλάκι στον host κώδικα, όσο αυτό είναι εφικτό.

Οι views είναι ένα εξαιρετικό προγραμματιστικό εργαλείο για το σκοπό αυτό. Το ίδιο και οι ερωτήσεις με WITH statement καθώς και οι προσωρινοί πίνακες
<https://dev.mysql.com/doc/refman/8.4/en/with.html>
<https://dev.mysql.com/doc/refman/8.4/en/create-temporary-table.html>

Engineering is all about tradeoffs. You must judge the tradeoffs for whichever part of your design that you decide, and document the decision accordingly.

ΦΑΣΗ II: υλοποίηση κορμού εφαρμογής

Στη φάση II θα φτιάξετε κάποια γρήγορα prototypes από τις ερωτήσεις και τις οπτικοποιήσεις που απαιτούνται (όπως θα δείτε παρακάτω). Θα χρειαστεί:

1. Στήσιμο του προγραμματιστικού περιβάλλοντος στο οποίο θα γίνει η ανάπτυξη.
2. Στήσιμο του περιβάλλοντος στο οποίο θα στηθεί και θα τρέξει η εφαρμογή σας (ενδεχομένως το ίδιο).
3. Πειραματισμός με έτοιμα παραδείγματα από την τεκμηρίωση των τεχνολογιών που θα χρησιμοποιήσετε: φτιάξτε μικρά προγραμματάκια που να τρέχουν.
4. Κατασκευή του κώδικα που προσπελάζει τη βάση δεδομένων και (α) συνδέεται, (β) υποβάλει μια ερώτηση, (γ) διαχειρίζεται το αποτέλεσμα της.
5. Κατασκευή του κώδικα που οπτικοποιεί δεδομένα (όχι απαραίτητα αποτελέσματα ερωτήσεων σε βάση) με τον επιθυμητό τρόπο.
6. Προοδευτική σύνδεση των παραπάνω.

/ Στην φάση III θα εμπλουτίσετε την εφαρμογή σας με την πλήρη γκάμα από ερωτήσεις και οπτικοποιήσεις που ζητούνται */*

Τεχνολογίες που μπορείτε να χρησιμοποιήσετε (ενδεικτικά & όχι περιοριστικά):

- **d3** (<https://d3js.org/>) (αξίζει να αφιερώσετε ώρα να περιηγηθείτε στα παραδείγματα του <https://observablehq.com/@d3/gallery> τα οποία έχουν, όλα, και τον κώδικά τους μαζί).
- **JavaFX σε περιβάλλον Eclipse** (<https://openjfx.io/> && ένα εξαιρετικό tutorial στο σύνδεσμο: <https://code.makery.ch/library/javafx-tutorial/>). Προσοχή: μετά τη Java 8, δεν υπάρχει built-in support της JavaFX στο jdk και πρέπει να συμπεριλάβετε τα σχετικά jar στο project.
- **Swing**. Το Java Swing είναι ένα απλό και πανίσχυρο προγραμματιστικό εργαλείο (<https://docs.oracle.com/javase/tutorial/uiswing/components/index.html> και ιδίως <https://docs.oracle.com/javase/tutorial/uiswing/components/menu.html> <https://docs.oracle.com/javase/tutorial/uiswing/examples/components/>). Υπάρχει και η πολύ απλή library JFreeChart <https://www.baeldung.com/jfreechart-visualize-data> .
- **Οτιδήποτε άλλο περιβάλλον για front-end/application logic σας είναι οικείο.**

ΦΑΣΗ III: τελειοποίηση της εφαρμογής

Στη φάση III, θα πρέπει να προσθέσετε και ένα βαθμό διαδραστικότητας στο παραπάνω. Προσθέστε μενού επιλογής (ή άλλους τρόπους επιλογής) και χρησιμοποιήστε γραφικούς τρόπους αλληλεπίδρασης (π.χ., φόρμες και drop-down list boxes) ώστε να πάρετε από το χρήστη τι ακριβώς επιθυμεί να δει. Συνδέστε το κομμάτι αυτό με ερωτήσεις και οπτικοποιήσεις.

Παραδοτέα

Turnin: Θα χρειαστεί να κάνετε turnin ένα text file που θα περιέχει: (α) AM, Ονόματα της ομάδας, (β) το link@Github, στο οποίο θα βρω αυτά που ζητάω στη σχετική λιστούλα, (γ) αν υπάρχει κάτι άλλο που θέλετε να μου πείτε. `turnin prj@mye030 AM1_AM2_AM3_prj.txt`

Στο <project> @ GitHub για το project σας (κατά προτίμηση με ένα διακριτό όνομα) θα αναμένω να δω (Α) το αρχικό README.md όπου θα πρέπει να φαίνονται οι συντελεστές, και οι όποιες τεχνικές λεπτομέρειες, (Β) το φάκελο src και το φάκελο data (τουλάχιστον) και (Γ) ένα φάκελο deliverables που θα περιέχει (α) το pdf για το τελικό παραδοτέο, (β) ένα link to a mp4 video, approx. 15' length, showing your system at work, along with an explanation of the internal structure and design. **Checklist:**

1. Ο κώδικας + δεδομένα + README.md μέσω ενός GitHub project
2. Τελική αναφορά με όνομα AM1_AM2_AM3_projectReport.pdf
3. Ένα link to video with your system at work, similarly named

Θα σας πρότεινα να μπειτε στη διαδικασία να κάνετε το project σας public, οπότε να μεριμνήσετε να έχετε test + documentation όπως πρέπει. Μπορείτε να δείτε παραδείγματα Readme.md στο [OOAD](#), στο [Pythia](#), και για πιο απλά, στο [Delian Cubes](#) και στο [Hecate](#).

Οδηγίες για το video. Προτεινόμενη δομή για το video που θα φτιάξετε.

ΜΕΡΟΣ Α (περίπου 5'): υπόβαθρο του project

- Ποιοι είμαστε στην ομάδα
- Σχήμα της βάσης δεδομένων, σχεδιαστικές επιλογές και trade-off's
- Πώς μετασχηματίσαμε και φορτώσαμε τα δεδομένα

ΜΕΡΟΣ Β (περίπου 5'): οργάνωση του project

- Αρχιτεκτονική του κώδικα (π.χ., δείχνοντας modules/packages/... μέσα σε ένα IDE)
- Εξήγηση της βασικής λειτουργικότητας του project μέσα στον κώδικα (δείχνοντας τα σχετικά κύρια σημεία μέσα σε ένα IDE)

ΜΕΡΟΣ Γ (περίπου 5'): Demo

- Demo

Προφανώς όλα τα παραπάνω είναι ενδεικτικά.

PROJECT CHECKLIST

REPEATABILITY		FRONT END	
☐ Scripts ETL			
☐ Scripts DDL		CONF/JOURNAL	AUTHORS
☐ Backup of the DB		☐ Profile	☐ Profile
		☐ Filtered profile	☐ Annual stats
ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ		☐ Paper details	
☐ Εξηγήστε τι σχεδίαση που ακολουθήσατε και τα trade-offs			
☐ Έχετε PKs? FK's?		YEARS	CHARTS
☐ Έχετε σχέδιο για το πώς θα ρωτάτε?		☐ Profile	☐ LineCharts
		☐ Details	☐ BarCharts
ΠΑΡΑΔΟΤΕΑ		☐ Filtered details	☐ Scatterplots
☐ Git repo with code			
☐ Report			
☐ Video			

Χρονοδιάγραμμα

Στη συνέχεια παρατίθενται ενδεικτικά στάδια της ανάπτυξης με ενδιάμεσες προθεσμίες (milestones) και (το μόνο υποχρεωτικό και παραδόσιμο) καταληκτική ημερομηνία ολοκλήρωσης.

[10/02]	Εκφώνηση
Κάντε την Φάση I και ότι μπορείτε από II	Εκτέλεση των βημάτων της ΦΑΣΗΣ I. Στόχοι: • Exported Σχήμα + workbench screenshot • Φορτωμένη βάση για τα backbone data backup • scripts + 1-page diagram for the transformation process • First cut of view definitions and main queries
[10/03]	Milestone: ΟΛΟΚΛΗΡΩΣΗ ΦΑΣΗΣ I
Η αρχή είναι το ήμισυ του παντός	Στήσιμο προγραμματιστικού περιβάλλοντος/framework και κατανόησή τους 60% του χρόνου να τρέξει η πρώτη αναφορά, 20% του χρόνου να τρέξει η δεύτερη, the rest of the time for the rest. Στόχοι: • Project setup of the application • Code containing the above forms/charts • Πρώτη οπτικοποίηση με dynamically constructed queries • Περιβάλλον αλληλεπίδρασης για τις επιλογές του χρήστη • Πρώτη έκδοση της αναφοράς
[21/04]	Milestone: ΟΛΟΚΛΗΡΩΣΗ ΦΑΣΗΣ II
	Πλήρης υλοποίηση της εφαρμογής Παραδοτέα: P3.1: το σύστημα ως ολοκληρωμένος κώδικας (και ως back-end, και ως front-end, και τα scripts για το μετασχηματισμό και φόρτωση των δεδομένων), καθώς και τα δεδομένα (input, output, backups) P3.2: τελική αναφορά (όπως στο σχετικό πρότυπο που βρίσκεται αναρτημένο στο δικτυακό τόπο του μαθήματος) P3.3: video με τον σύστημα σε δράση, και την τελική αναφορά
[26/05]	Hard Deadline: ΟΛΟΚΛΗΡΩΣΗ ΦΑΣΗΣ III

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!!