

Βάσεις Δεδομένων

Κάκια Παναγίδα
kakiar@di.uoa.gr



Σκοπός του μαθήματος

Τρεις βασικοί στόχοι:

1. Να μπορείτε να σχεδιάζετε και υλοποιείτε ένα σύστημα βάσεων δεδομένων χρησιμοποιώντας ένα (σχεσιακό) ΣΔΒΔ
 - Τι σημαίνει αυτό:
 - Μοντελοποίηση
 - Προγραμματισμός (σε SQL)
 - Καθώς και τη σχετική θεωρία στην οποία βασίζονται τα παραπάνω (κανονικές μορφές, σχεσιακή άλγεβρα)

Σκοπός του μαθήματος

2. Θέματα υλοποίησης ενός ΣΔΒΔ (το εσωτερικό του)
 - Βασικές δομές δεδομένων για προσπέλαση δεδομένων από το δίσκο (ευρετήρια)
 - Βασικές αρχές βελτιστοποίησης ερωτήσεων
3. Γενικές γνώσεις και δεξιότητες για τη διαχείριση δεδομένων

Διαδικαστικά

Θεωρία

- Εργαστήρια και Ασκήσεις
- Τελικός Βαθμός 60% τελική εξέταση και 40% βαθμός ασκήσεων- εργαστηρίων

Εργαστηρια

- Αρτιοι Περιττοι Δευτέρα 12:00-14:00

Υλικό

- Eclass – διαλέξεις, ανακοινώσεις και εργασίες

Βιβλιογραφία

- **Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων - 7η Έκδοση Αναθεωρημένη**, R. Elmasri, S.B. Navathe, Εκδόσεις Δίαυλος

Διαδικαστικά

Κάνετε ερωτήσεις -και μια «άσχετη» ερώτηση είναι καλύτερη από καμία ερώτηση

Θεωρία

- Εργαστήρια και Ασκήσεις
- Τελικός Βαθμός 10% τελική εξέταση και 90% βαθμός εισηγήσεων- εργαστηρίων

Εργαστήρια

- Αρτιοι Περιττοι Δευτέρα 12:00-14:00

Υλικό

- Eclass - υλικό, ξένοι στίχοι, κείμενα κ.λπ.

Βιβλιογραφία

- Θεμελιώδεις Αρχές Συστημάτων Βάσεων Δεδομένων, Έκδοση Ανάπτυξη, R. Elmasri, S.B. Navathe, Δόσις Δίπλα



Εργαστήρια 12:00-14:00

Εργαστήρια	Άρτιοι	Περιττοί
1 ^ο εργαστήριο	30/10	6/11
2 ^ο εργαστήριο	13/11	20/11
3 ^ο εργαστήριο	27/11	4/12
4 ^ο εργαστήριο	11/12	18/12

Τι θα δούμε σήμερα

- I. Τι είναι οι ΒΔ, γιατί μας ενδιαφέρουν
- II. Γενική εικόνα μαθήματος
- III. Ιστορία των ΒΔ

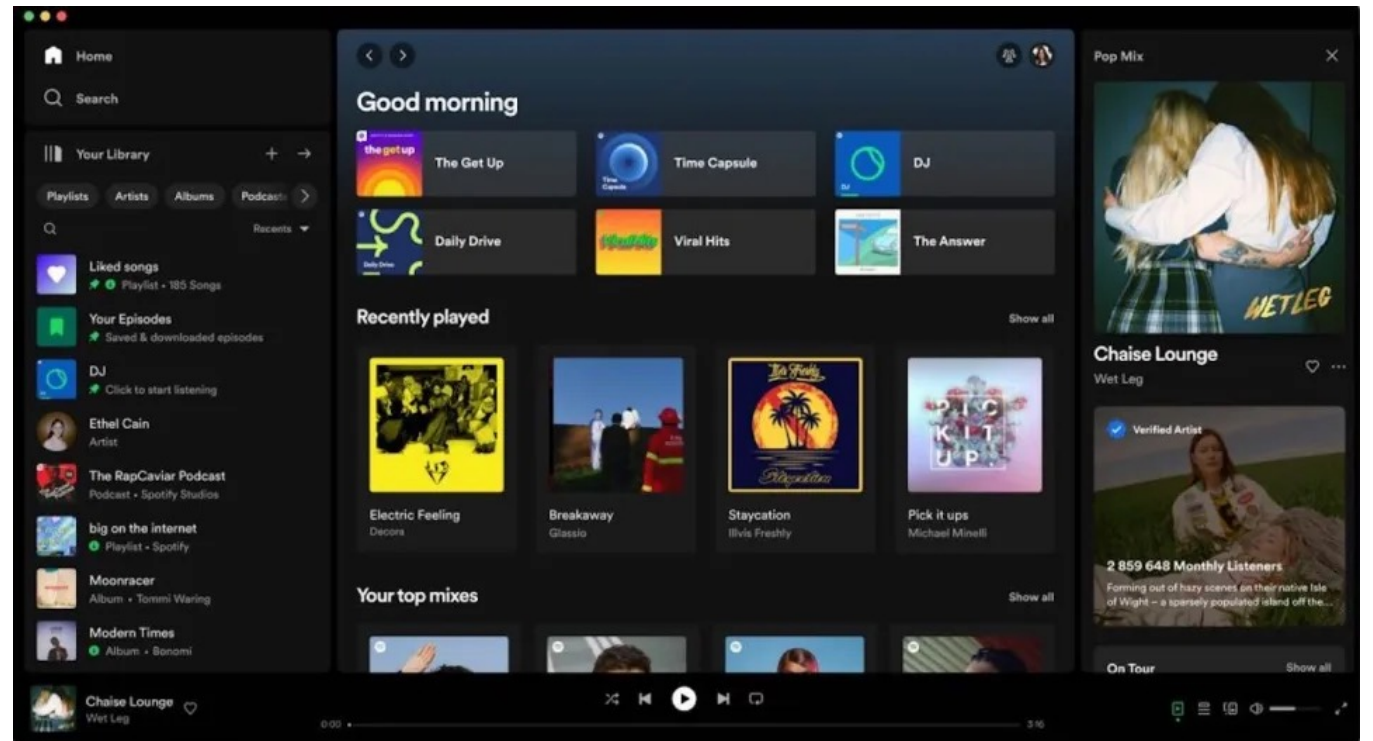
DATABASES

Γιατι μας ενδιαφέρουν οι
ΒΔ;

Οι Βάσεις δεδομένων και τα Συστήματα Διαχείρισης Βάσεων
Δεδομένων είναι παντού.

DATABASES EVERYWHERE

Καταλογοι Οργάνωσης συλλογων μουσικής



Διαδικτυακές αγορές

Amazon.com: databases: Books - Mozilla Firefox

File Edit View History Bookmarks Tools Help

https://www.amazon.com/s/ref=sr_il_ti_stripbooks?l 90% Search

< Any Category

Books

- Computers & Technology (28,294)
- Databases & Big Data (19,455)
- Network Administration (5,173)
- Network Storage & Retrieval Administration (4,775)
- Data Modeling & Design (1,008)
- Software (9,823)
- Software Development (2,220)
- Information Theory (1,058)
- Schools & Teaching (762)
- Computer Simulation (440)
- + See more

Refine by

Delivery Day

Get It by Tomorrow

Amazon Prime

Prime

Prime | FREE One-Day

New Releases

- Last 30 days (165)
- Last 90 days (457)
- Coming Soon (175)

Book Format: Hardcover | Paperback | Kindle Edition

Database Design for Mere Mortals: A Hands-On Guide to Relational Database Design (3rd Edition) Feb 14, 2013
\$17.12 - \$28.86 Prime | FREE One-Day
Paperback, Kindle Edition
★★★★★ 98

Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement May 21, 2012
\$26.93 - \$28.35 Prime | FREE One-Day
Paperback, Kindle Edition
★★★★★ 43

Database Systems: Introduction to Databases and Data Warehouses Jun 2, 2016
\$63.88 Prime
Paperback
★★★★★ 16

FUNDAMENTALS OF DATABASE SYSTEMS

Database System Concepts

Databases DeMYSTiFieD

Κρατήσεις
ταξιδιών,
ξενοδοχείων,
κ.λπ.

Αθήνα (ATH) → Μάντσεστερ (MAN)
Σαβ 04/03 - Περ 09/03 | Επιβάτες: 1 ενήλικας

Αλλαγή

Λίστα Φίλτρα

Αεροπορία	Ημερομηνία	Από	Προς	Χρόνος	Στάση	Λειτουργίες	Τιμή
Lufthansa	Σαβ 04/03	Αθήνα	Μάντσεστερ	6h 35m	1 στάση (FRA)	Λεπτομέρειες	214,85 € 208,85 € Τελική τιμή
swiss	Περ 09/03	Μάντσεστερ	Αθήνα	16h 30m	1 στάση (ZRH)	Λεπτομέρειες	Συνέχεια
				19h	1 στάση (ZRH)	Λεπτομέρειες	
TURKISH AIRLINES	Σαβ 04/03	Αθήνα	Μάντσεστερ	7h 15m	1 στάση (IST)	Λεπτομέρειες	215,89 € 209,81 € Τελική τιμή
TURKISH AIRLINES	Περ 09/03	Μάντσεστερ	Αθήνα				Συνέχεια

Οργάνωση Βιβλιογραφικών Αναφορών

Recipe Shopping items

Title:

Language:

Picture:

Category:

Source:

Nb servings: Portion size or yield:

Preparation time: Total time:

Profit margin:

Quantity	Ingredient	Comment
<input type="text" value="200 ml"/>	<input type="text" value="olive oil"/>	<input type="text"/>
<input type="text" value="7 cloves"/>	<input type="text" value="garlic"/>	<input type="text" value="chopped"/>
<input type="text" value="20 leaves"/>	<input type="text" value="basil"/>	<input type="text" value="chopped"/>
<input type="text" value="1 small"/>	<input type="text" value="chile"/>	<input type="text" value="such as cayenne"/>
<input type="text" value="to taste"/>	<input type="text" value="salt"/>	<input type="text"/>
<input type="text" value="to taste"/>	<input type="text" value="pepper"/>	<input type="text"/>
<input type="text" value="12 (150 g each)"/>	<input type="text" value="tuna steaks"/>	<input type="text"/>
<input type="text" value="1.5 kg"/>	<input type="text" value="cooked spinach"/>	<input type="text" value="buttered"/>
<input type="text" value="500 ml"/>	<input type="text" value="sauce vierge"/>	<input type="text"/>

- Combine the olive oil, garlic, basil, chile, salt, and pepper to make a marinade.
- Marinate the tuna for 1 hour.
- Grill the tuna steaks, keeping them medium rare.
- For each portion, place 4 oz (125 g) hot spinach on a plate. Place the tuna steak on top. Drizzle the sauce around.

Mendeley Desktop

File Edit View Tools Help

Search: Maria

Filter by Authors

Authors	Title	Year	Published In	Added
Aarseth, Espen	Playing Research: Methodological approaches to game analysis	2003	Proceedings of the Digital ...	07-Jun-15
Aarseth, Espen	A Narrative Theory of Games	2012	Proceedings of the Internat...	14-Feb-15
Abt, Clark C.	Serious Games	1970		12-Feb-12
Ahdell, Rolf; Andr...	Games and simulations in workplace eLearning: How to align eLearning conte...	2001	Science	30-May-11
Amazon	Amazon Lumberyard	2016		11-Feb-16
Amory, Alan; Naic...	The use of computer games as an educational tool: Identification of appropr...	1999	British Journal of Educatio...	30-Jun-11
Amory, Alan; Naic...	The use of computer games as an educational tool: Identification of appropr...	1999	British Journal of Educatio...	07-Jun-14
Anagnostou, Kost...	Video Game Genre Affordances for Physics Education	2011	International Journal of ...	07-Jun-14
Anderson, Eike Fa...	Serious Games in Cultural Heritage	2009	The 10th Internation...	07-Jun-14
Ang, Chee Siang; ...	Computer Games and Sociocultural Play: An Activity Theoretical Perspective	2010	Games and Culture	05-Dec-11
Annetta, Leonard A.	The "I"s" have it: A framework for serious educational game design.	2010	Review of General Psy...	13-Nov-12
Apostolellis, Panag...	Small group learning with games in museums: Effects of Interactivity as Medi...	2015	Proceedings of the 14th Int...	06-Jun-16
Aronowsky, Audr...	Building the next science generation through game-based learning in museums	2013	GAMES + LEARNING...	11-May-15
BBC	Syrian Journey: Choose your own escape route			18-May-16

Details Notes Contents

Type: Conference Proceedings

A Narrative Theory of Games

Authors: E. Aarseth

Proc. Title: *Proceedings of the International Conference on the Foundations of Digital Games - F...*

Year: 2012

Pages: 129-133

Abstract:
This paper presents a narrative theory of games, building on standard narratology, as a solution to the conundrum that has haunted computer game studies from the start: How to approach software that combines games and stories?

Tags:

Author Keywords:
ludonarrative model; narratology; storygames

City:

Editors:

Publisher:

URL:
<http://dl.acm.org/citation.cfm?doi=2282338.2282365>



Βάσεις Δεδομενων ΠΑΝΤΟΥ

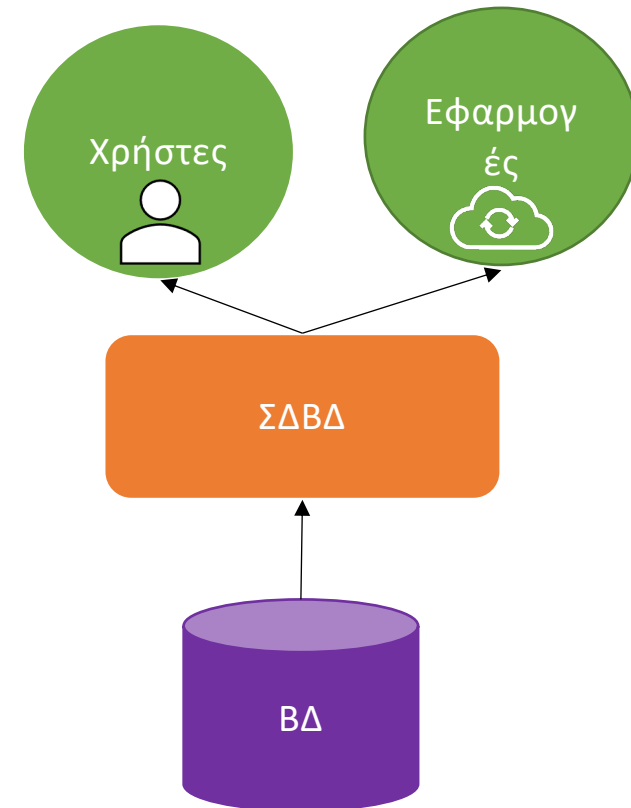
- Τηλεφωνικός κατάλογος
- Τραπεζικοί λογαριασμοί
- κ.ά.

...όπου τα δεδομένα είναι κυρίως κείμενο κι αριθμοί

- Αλλά και μη παραδοσιακές μορφές δεδομένων, π.χ.:
- posts (Facebook)
- tweets (Twitter)
- εικόνες (Flickr, Instagram)
- video clips (YouTube)
- κ.ά.

Βάση Δεδομένων

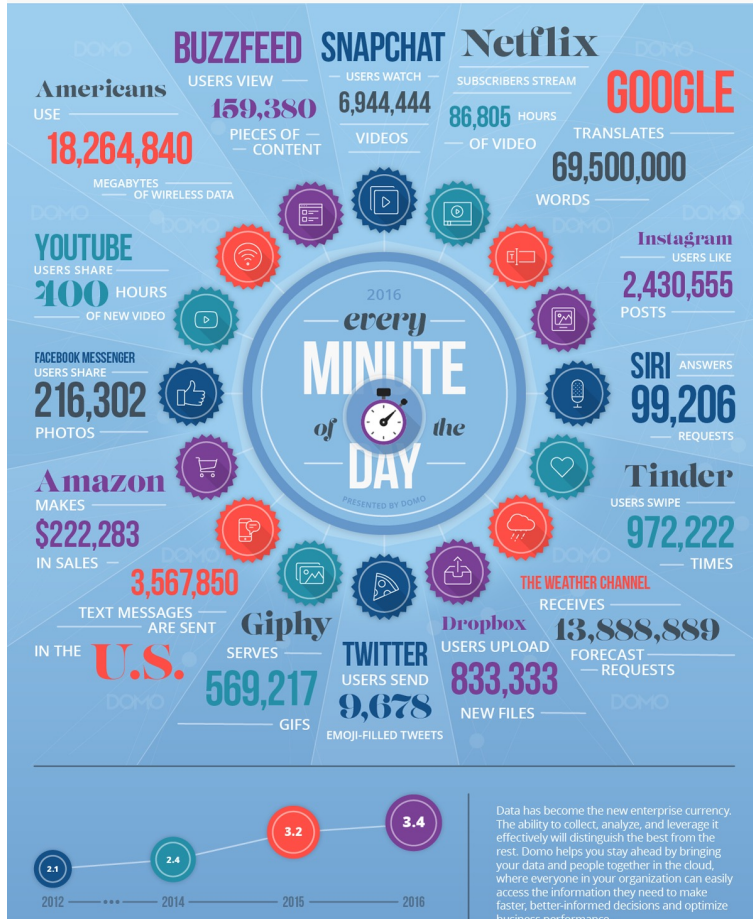
- Μια **οργανωμένη, ολοκληρωμένη** συλλογή (μεγάλου όγκου) δεδομένων που συνήθως αποθηκεύεται στο δίσκο.
- Μοντελοποιεί τις δραστηριότητες μιας πραγματικής επιχείρησης/οργανισμού:
 - Αντικείμενα (π.χ., φοιτητές, μαθήματα)
 - Συσχετίσεις (π.χ., φοιτητής γράφεται σε μάθημα)
- Βρίσκεται “πίσω” από οποιοδήποτε πληροφοριακό σύστημα





DATA NEVER SLEEPS 4.0

How much data is generated every minute? In the fourth annual edition of Data Never Sleeps, newcomers like Giphy and Facebook Messenger illustrate the rise of our multimedia messaging obsession, while veterans like YouTube and Snapchat highlight our insatiable appetite for video. Just how many GIFs, videos, and emoji-filled Tweets flood the Internet every minute? See for yourself below.



Big Data

Ο όρος «big data» αναφέρεται σε δεδομένα που είναι τόσο μεγάλα, γρήγορα ή πολύπλοκα που είναι δύσκολο ή αδύνατο να επεξεργαστούν με παραδοσιακές μεθόδους (3Vs).

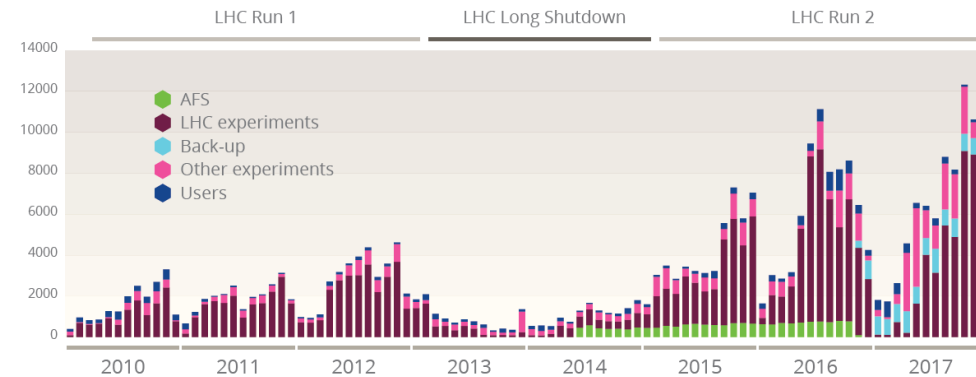
- Όγκος: Οι οργανισμοί συλλέγουν δεδομένα από διάφορες πηγές, συμπεριλαμβανομένων επιχειρηματικών συναλλαγών, έξυπνων συσκευών (IoT), βιομηχανικού εξοπλισμού, βίντεο, κοινωνικών μέσων και πολλά άλλα. Στο παρελθόν, η αποθήκευση θα ήταν πρόβλημα –
- Ταχύτητα: Με την ανάπτυξη του Διαδικτύου των Πραγμάτων, η ροή δεδομένων προς τις επιχειρήσεις με πρωτοφανή ταχύτητα και πρέπει να αντιμετωπιστεί εγκαίρως. Οι ετικέτες RFID, οι αισθητήρες και οι έξυπνοι μετρητές οδηγούν στην ανάγκη αντιμετώπισης αυτών των χειμάρρων δεδομένων σε σχεδόν πραγματικό χρόνο.
- Ποικιλία: Τα δεδομένα έρχονται σε όλους τους τύπους μορφών - από δομημένα, αριθμητικά δεδομένα σε παραδοσιακές βάσεις δεδομένων έως μη δομημένα έγγραφα κειμένου, μηνύματα ηλεκτρονικού ταχυδρομείου κ.α.

Επιστημονικά Δεδομένα

Large Hadron Collider (LHC)



Data recorded on tapes at CERN on a monthly basis in TB



Τα σωματίδια συγκρούονται στους ανιχνευτές Large Hadron Collider (LHC) περίπου 1 δισεκατομμύριο φορές ανά δευτερόλεπτο, δημιουργώντας περίπου ένα petabyte δεδομένων σύγκρουσης ανά δευτερόλεπτο

Τα πειράματα LHC παράγουν περίπου **90 petabytes** δεδομένων ετησίως και επιπλέον **25 petabytes δεδομένων** παράγονται ετησίως για δεδομένα από άλλα (μη LHC) πειράματα στο CERN. Η αρχειοθέτηση τεράστιων ποσοτήτων δεδομένων είναι μια ουσιαστική λειτουργία στο CERN.

<https://home.cern/science/computing/storage>

Χρήση Βάσεων δεδομένων

Μοντελοποίηση, έλεγχος πλεονασμού, περιορισμοί ακεραιότητας

Αποδοτική επεξεργασία ερωτήσεων (queries) (ευρετήρια, βελτιστοποίηση)

Μόνιμη αποθήκευση (persistent storage)

Ενημέρωση δεδομένων

Ορθότητα λειτουργίας: Πως θα διασφαλίσουμε την ορθότητά τους κατά τη διάρκεια αποτυχιών και ταυτόχρονης προσπέλασης από πολλούς χρήστες, ανάκαμψη από σφάλματα

Επίσης: θέματα ασφάλειας, δικαιωμάτων/εξουσιοδότηση προσπέλασης, ...

Γιατί
χρειαζόμαστε
ειδικό
λογισμικό;

Κοινή λειτουργικότητα ήδη υλοποιημένη

Σωστή υλοποίηση

Αποδοτική προσπάθεια

Γρήγορη ανάπτυξη εφαρμογών

Ανεξαρτησία δεδομένων (θα δούμε περισσότερα)

Μόνιμη αποθήκευση

Πότε δεν χρησιμοποιούμε ΒΔ

- Επένδυση σε λογισμικό και υλικό, καθώς και για εκπαίδευση
- Η γενικότητα που παρέχει προκαλεί χρονική επιβάρυνση (overhead)
- Δε χρειάζονται όσα προσφέρει

Παραδείγματα ΣΒΔΒ

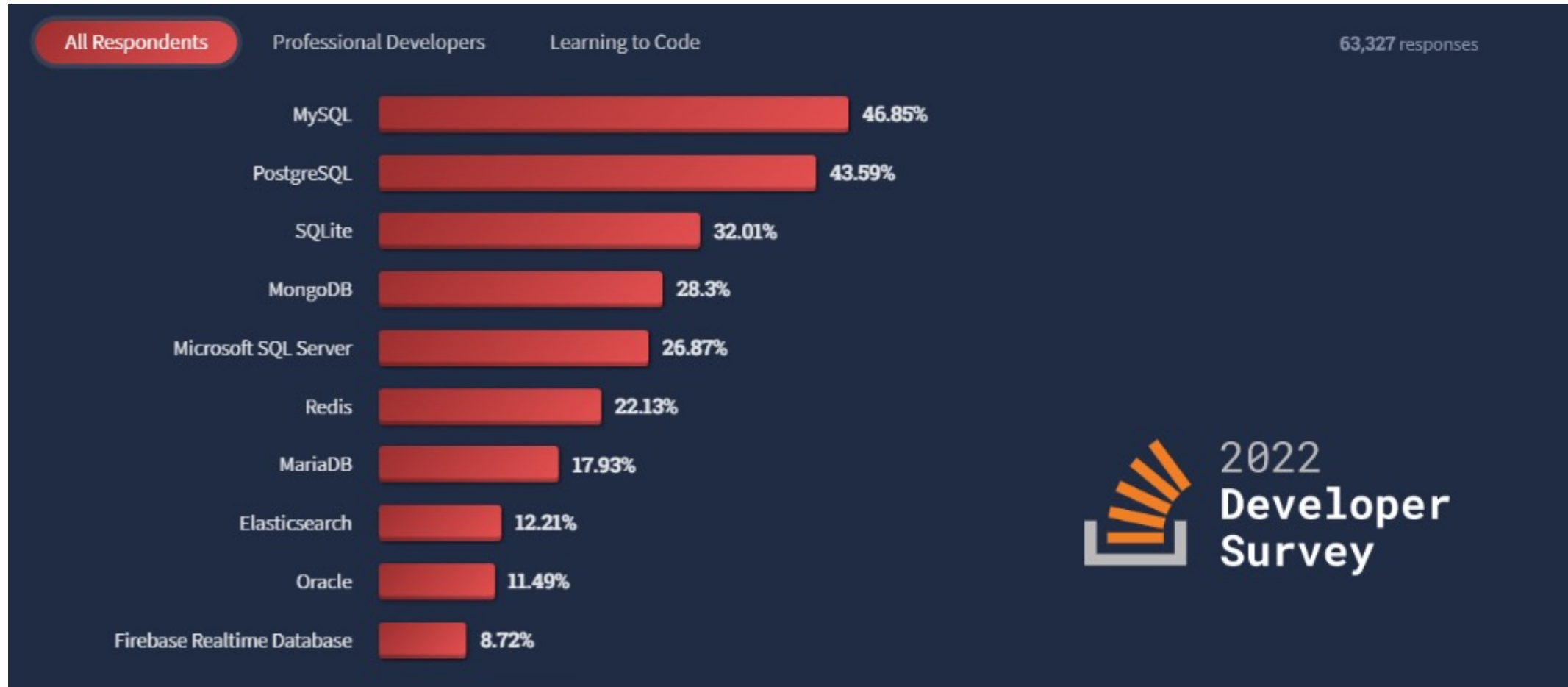
Εμπορικά

- Oracle
- IBM
- MS SQL Server
- Sybase
- MS Access ...

Ελεύθερα

- MySQL
- PostgreSQL
- MongoDB
- MariaDb
-

Σειρα κατάταξης 2022



Κατηγορίες Βάσεων Δεδομένων

Σχεσιακή Βάση Δεδομένων

- Το SQL είναι συντομογραφία για Structured Query Language, που σημαίνει βασικά έναν πολύ σταθερό τρόπο ταξινόμησης δεδομένων με τη μορφή πινάκων, στηλών και γραμμών.

Όνομα	Επώνυμο	Αρ. Μητρώου	Εξαμηνιο
Μάριος	Παπαδοπουλος	11115202100333	1
Μαρία	Παπαδοπούλου	11115202101001	1

Μη- Σχεδιακή Βάση Δεδομένων

Είναι λιγότερο δομημένη/περιορισμένη σε μορφή, και έτσι, επιτρέπει μεγαλύτερη ευελιξία και προσαρμοστικότητα.

Ένα σύνολο δεδομένων που δεν είναι σαφώς καθορισμένο, που σημαίνει ότι δεν είναι οργανωμένο ή δομημένο

Π.χ. το Facebook Messenger

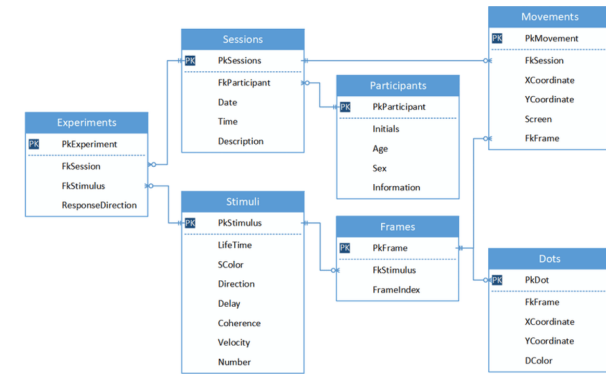
Κατηγορίες Βάσεων Δεδομένων

Σχεσιακή

Τα δεδομένα διαρθρώνονται εύκολα σε κατηγορίες.

Τα δεδομένα σας είναι συνεπή ως προς την εισαγωγή, το νόημα και είναι εύκολο στην πλοήγηση.

Οι σχέσεις μπορούν εύκολα να οριστούν μεταξύ των σημείων δεδομένων.



Μη σχεσιακή

Τα δεδομένα δεν περιορίζονται σε μια δομημένη ομάδα.

Μπορείτε να εκτελέσετε λειτουργίες που επιτρέπουν μεγαλύτερη ευελιξία.

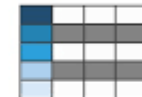
Τα δεδομένα και η ανάλυσή σας μπορεί να είναι πιο δυναμικά και να επιτρέπουν περισσότερες παραλλαγές.

NoSQL

Key-Value



Column-Family



Graph



Document



Μια γενική εικόνα: Δημιουργία ΣΒΔ

- Ένα μοντέλο δεδομένων είναι μια γλώσσα περιγραφής της δομής και οργάνωσης των δεδομένων μέσω νοητικών όρων/ εννοιών /εργαλείων
 - Πχ Το σχεσιακό μοντέλο εκφράζει τα δεδομένα και την οργάνωση τους χρησιμοποιώντας μόνο πίνακες με γραμμές και στήλες
- Ένα σχήμα δεδομένων είναι η περιγραφή μιας συγκεκριμένης βάσης δεδομένων όπως προκύπτει από τη χρήση ενός μοντέλου δεδομένων
 - Πχ
 - Φοιτητής (κωδ, ονομα, τηλ)
 - Μάθημα (κωδ, τίτλος, εξαμ)
 - Δήλωση (κωδφ, κωδμ, βαθμός)

Τι είναι η «ερώτηση» (query)

- Ερώτηση (query): Προκαλεί την ανάκτηση κάποιων δεδομένων από μια ΒΔ



- Δοσοληψία (transaction): αφορά την ανάγνωση κι εγγραφή δεδομένων στη ΒΔ

Λειτουργίες ΣΔΒΔ

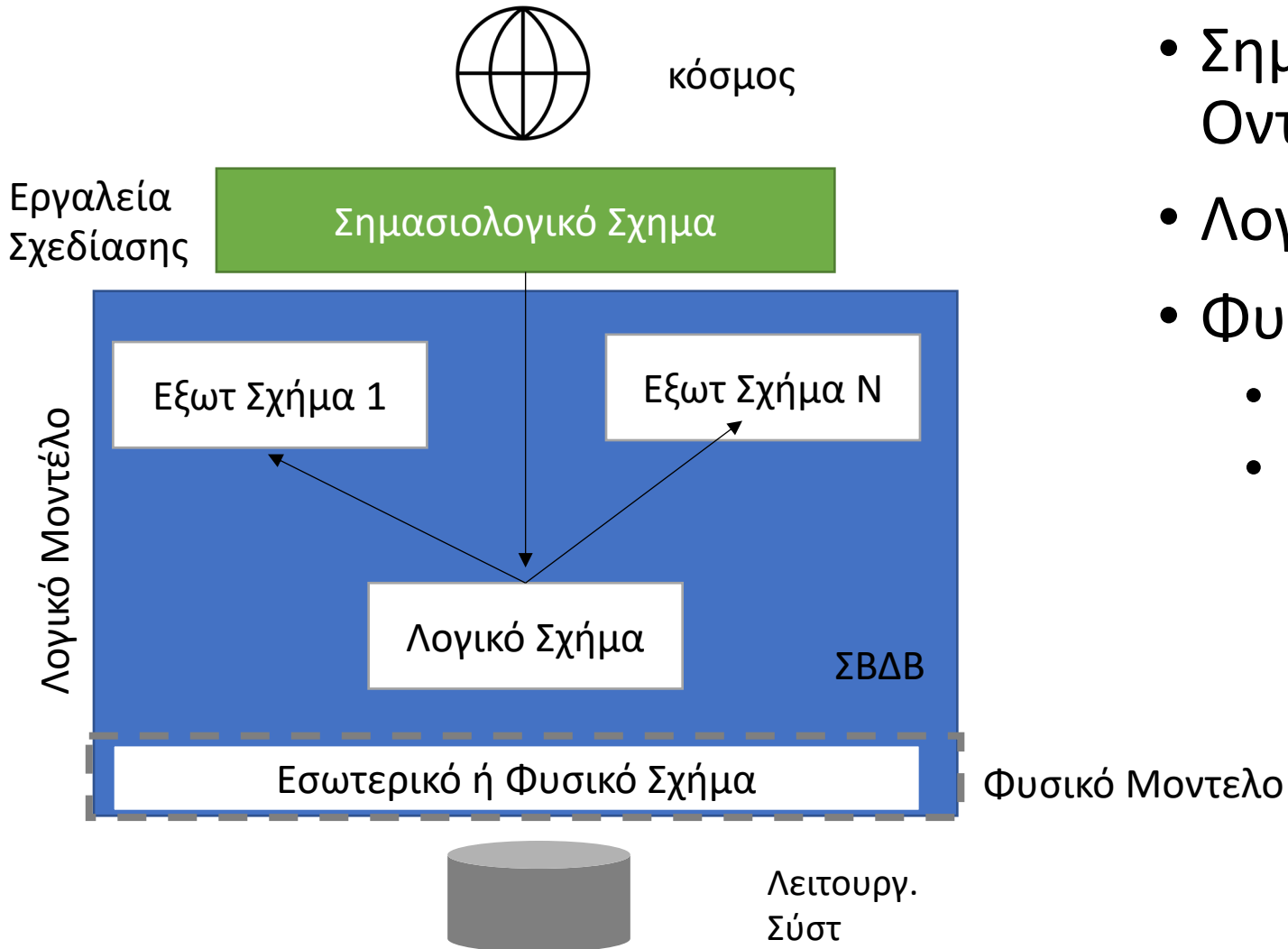
Ορισμός/διαχείριση δομής βάσεων δεδομένων

Διαχείριση δεδομένων (εισαγωγή/διαγραφή/ενημέρωση).

Ανάκτηση δεδομένων μέσω ερωτημάτων (αιτημάτων)

Διαχείριση συστήματος (χρήστες, ασφάλεια, επαναφορά)

Σχεδίαση ΒΔ



- Σημασιολογικό Μοντέλο: Μοντέλο Οντοτήτων/ Συσχετίσεων
- Λογικό Μοντέλο: Σχεσιακό Μοντέλο
- Φυσικό Μοντέλο:
 - Δομές B+ δέντρα, κατακερματισμός....
 - Οργάνωση: ταξινόμηση....

Παράδειγμα ΒΔ: Πανεπιστήμιο

Φοιτητες

Όνομα	A/M	Έτος	Κατεύθυνση
Λάκης Λαλάκης	2015123456	2	Επιστήμη Η/Υ
Αριστοτέλης Φιλόσοφος	2000123456	8	Τηλεπικοινωνίες

Μαθήματα

Όνομα_μαθήματος	Κωδικός_μαθ.	Διδ_μονάδες	Τμήμα
Σχεδίαση & Χρήση ΒΔ	K29	4	DI
Επικοινωνία Ανθρώπου Η/Υ	ΥΣ08	4	DI
Διακριτά Μαθηματικά	ΜΑ01	4	Math

Διδασκάλια

Κωδικός_διδασκ.	Κωδικός_μαθ.	Εξάμηνο	Ακ_έτος	Διδάσκων
85	K29	Εαρινό	2015	Ιωαννίδης
92	ΥΣ08	Χειμερινό	2016	Ρούσσου

Παράδειγμα ΒΔ: Πανεπιστήμιο

Βαθμολογία

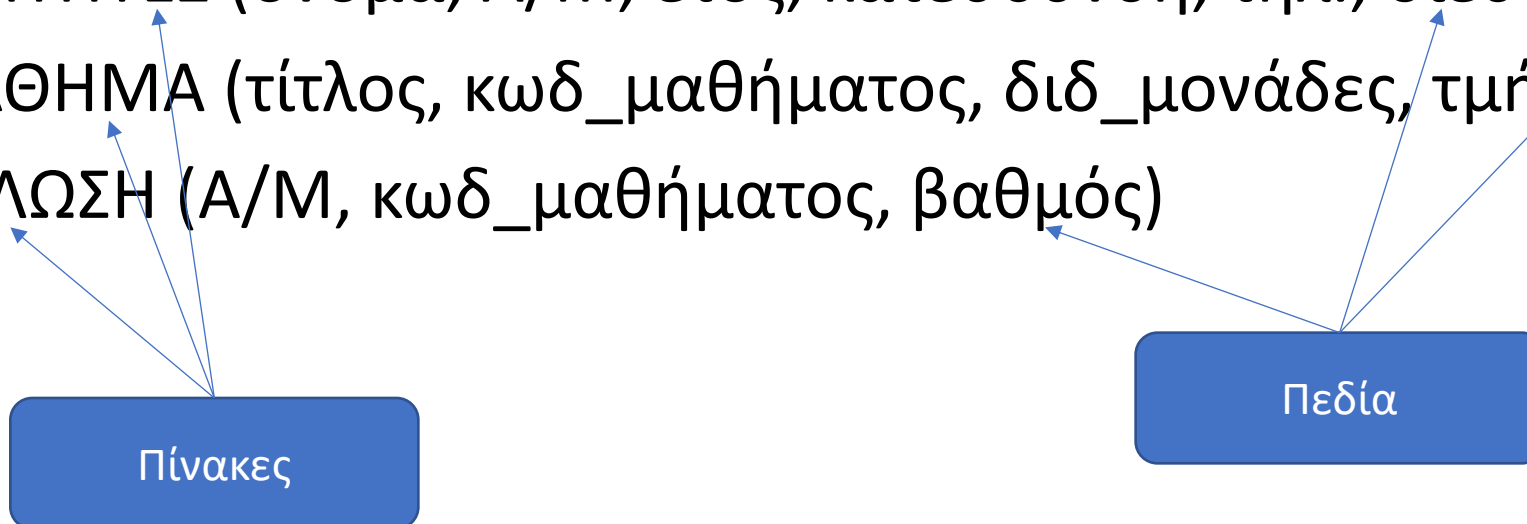
A/M	Κωδικός_διδασκ.	Βαθμός
2015123456	85	10
2000123456	90	4

Μπορούμε να έχουμε κι άλλα «αρχεία», όπως:

- ΔΗΛΩΣΕΙΣ, ΚΑΘΗΓΗΤΕΣ... και ότι άλλο χρειαστούμε για να αναπαραστήσουμε τον «κόσμο» που θέλουμε


Παράδειγμα ΒΔ: Πανεπιστήμιο

- Π.χ. ένα σχεσιακό σχήμα για τη Γραμματεία:
- ΦΟΙΤΗΤΕΣ (όνομα, Α/Μ, έτος, κατεύθυνση, τηλ., διεύθυνση, ...)
- ΜΑΘΗΜΑ (τίτλος, κωδ_μαθήματος, διδ_μονάδες, τμήμα, ...)
- ΔΗΛΩΣΗ (Α/Μ, κωδ_μαθήματος, βαθμός)



Παράδειγμα ΒΔ: Πανεπιστήμιο

- Παραδείγματα ερωτημάτων:
 - Δώσε μου τα ονόματα και τους βαθμούς όλων των φοιτητών που πέρασαν το μάθημα «Σχεδίαση και Χρήση Βάσεων Δεδομένων» το εαρινό εξάμηνο του 2020
 - Δημιούργησε ένα πιστοποιητικό σπουδών (μαθημάτων κι αντίστοιχων βαθμών) για την φοιτήτρια Μαρία Παπαδοπούλου
 - Ποια είναι τα συνιστώμενα προ-απαιτούμενα του μαθήματος ΣΧΒΔ;

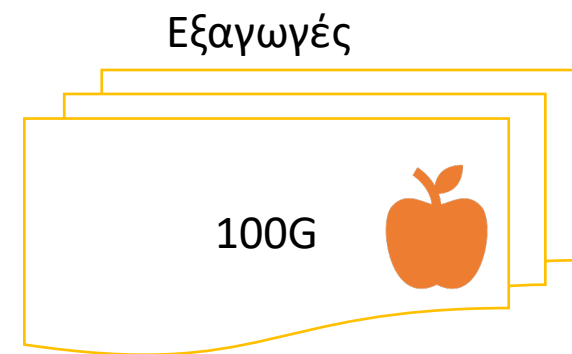
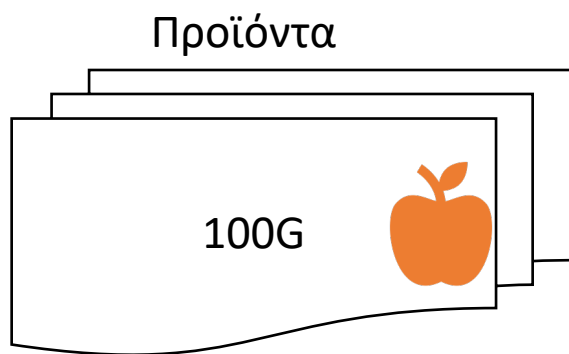
An illustration of a grocery store. In the background, a woman in a red top and teal skirt pushes a shopping cart filled with various items. To her right, a man with a mustache in a yellow sweater and brown pants pushes a cart containing a red umbrella, a white chicken, and other goods. In the foreground, a woman in a yellow top and pink floral skirt is looking at products on a shelf. Next to her, a man in a teal shirt and white t-shirt is talking to a woman in a pink top. The shelves are stocked with various products like bread, jars, and bottles. A large white circle is overlaid on the left side of the image, containing text.

Παράδειγμα: «φρουτεμπορική»

- 3 τμήματα (προϊόντων, πελατών, εξαγωγών)...

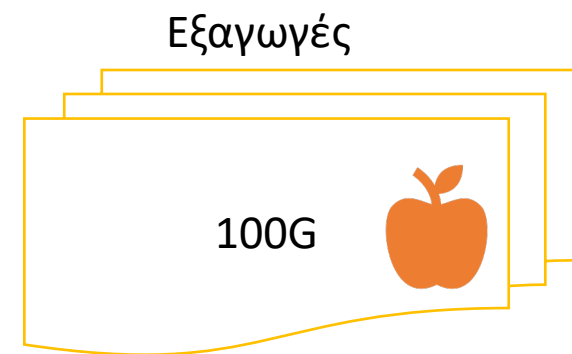
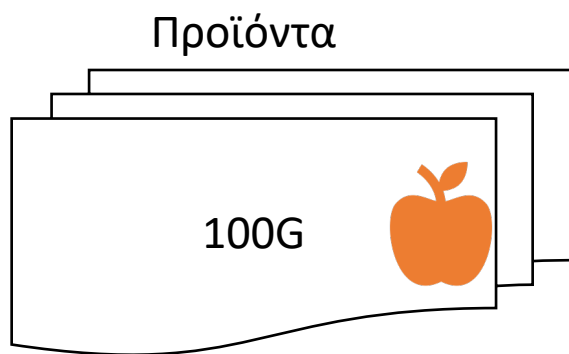
Παράδειγμα: «φρουτεμπορική»

- Πρόβλημα: η τιμή των μήλων (έστω 100 'G') σε 3 διαφορετικά συστήματα αρχείων -τα ίδια δεδομένα επαναλαμβάνονται στα αρχεία κάθε τμήματος...



Παράδειγμα: «φρουτεμπορική»

- Κι έτσι, αν αυξηθεί η τιμή των μήλων από 100 σε 120, θα πρέπει να ενημερωθεί κάθε αρχείο ξεχωριστά, αυξάνοντας την πιθανότητα λάθους...



Παράδειγμα: «φρουτεμπορική»

Πωλήσεις

Όνομα	Τιμή
Αχλάδι	150
Μήλο	120
Πεπόνι	200
Φράουλα	800

Εξαγωγές

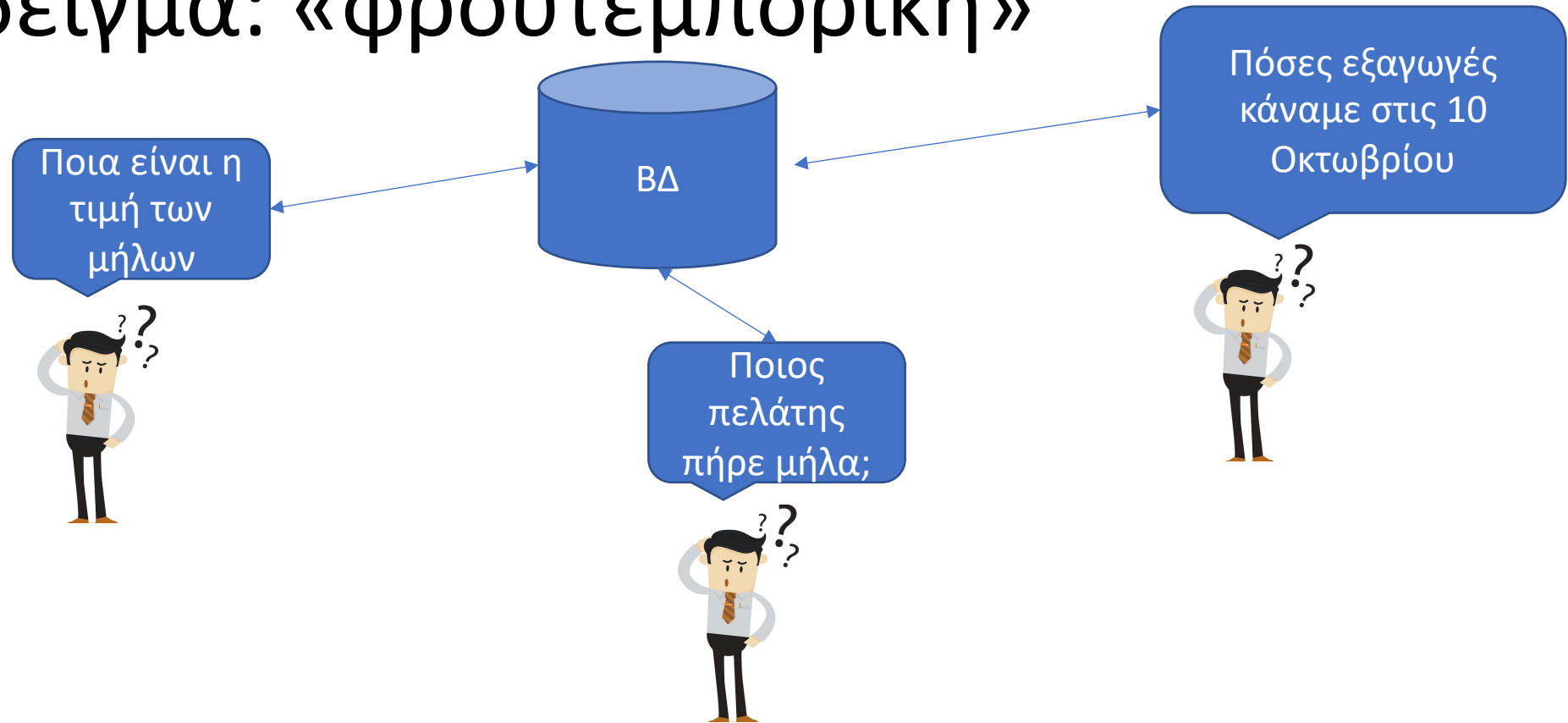
Όνομα	Τιμή
Αχλάδι	150
Μήλο	100
Πεπόνι	200
Φράουλα	800

Εισαγωγές

Όνομα	Τιμή
Αχλάδι	150
Μήλο	300
Πεπόνι	200
Φράουλα	800

...αλλά και την έλλειψη ευελιξίας π.χ. τι θα γίνει αν δημιουργηθεί 4ο τμήμα;

Παράδειγμα: «φρουτεμπορική»



- Όλα τα τμήματα θα έχουν τις ίδιες, κοινές πληροφορίες
- Τα δεδομένα είναι προσπελάσιμα απ'όλους
- Ο καθένας μπορεί να κάνει διαφορετικά ερωτήματα
- Νέα τμήματα μπορούν να προστεθούν με πρόσβαση στην ίδια BΔ

- Αρα....

Ιδιαίτερα χαρακτηριστικά των ΒΔ & ΣΔΒΔ

- Μειωμένη επανάληψη πληροφορίας
- Συνέπεια πληροφορίας
- Επιβολή (σημασιολογικής) ορθότητας
- Ανεξαρτησία δεδομένων
- Έλεγχος συνδρομικότητας
- Ανάκαμψη από βλάβη
- Ασφάλεια -λεπτομερής και κεντρικός έλεγχος

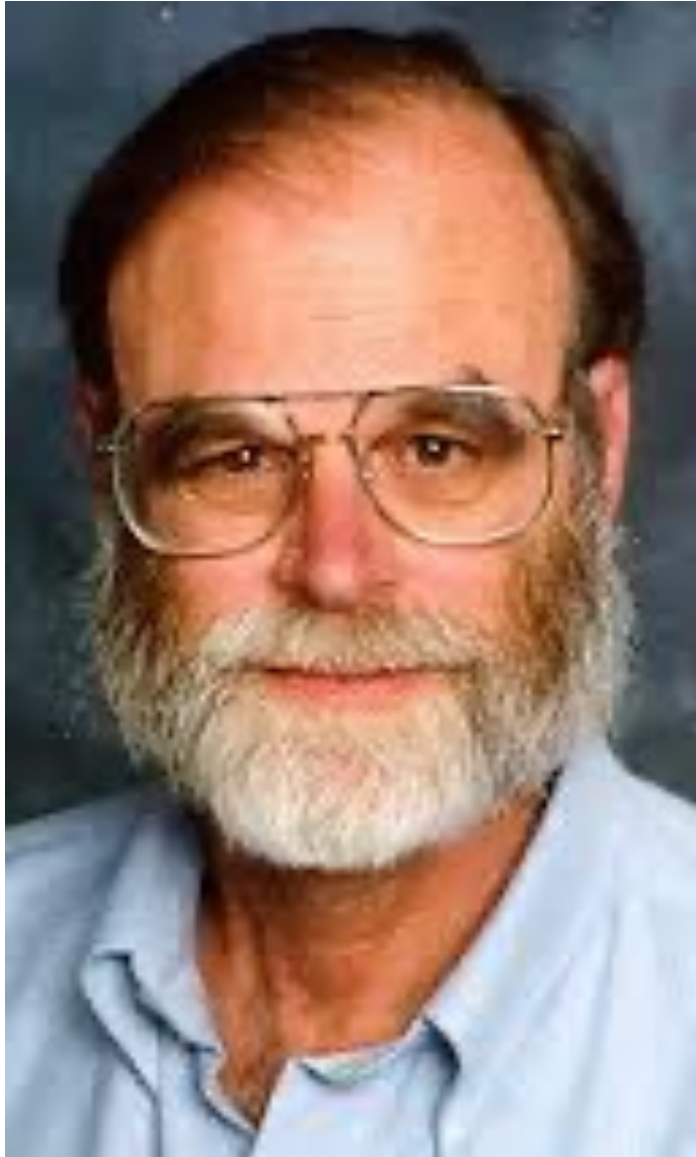
Βασικοί Ρόλοι

- Κατασκευαστής ΣΔΒΔ
 - αυτός που υλοποιεί το λογισμικό ΣΔΒΔ
- Αναλυτής/Σχεδιαστής ΒΔ
 - αυτός που κάνει ανάλυση απαιτήσεων
- Προγραμματιστής ΒΔ
 - αυτός που ασχολείται με κώδικα SQL
- Διαχειριστής Βάσεων Δεδομένων (DBA)
 - υπεύθυνος για τη σωστή λειτουργία και συντήρηση του συστήματος (όχι της ΒΔ). Π.χ. Oracle-certified.



Ιστορία

- **Δεκαετία του 1950**
 - Κάρτες και ταινίες (σειριακή επεξεργασία) – Batch processing
- **Αρχή του 1960**
 - Γενικευμένη χρήση δίσκων
 - πρώτο γενικού-σκοπού ΣΔΒΔ (διαχωρισμός της λειτουργικότητας διαχείρισης δεδομένων από τις εφαρμογές): Integrated Data Store (GE)
 - Charles Bachman (Recipient of the 1st Turing Award, 1973)
- **1970**
 - Edgar Codd (IBM, San Jose) *σχεσιακό μοντέλο δεδομένων* (relational data model) (Recipient of the Turing Award, 1981)
 - Ερευνητικά Προγράμματα: System R, INGRES - Γλώσσες: SEQUEL, QBE, QUEL



Ιστορία

- **Δεκαετία του 1980**
 - SQL (μέρος του **System R**)
 - transaction management (Jim Gray, Turing Award, 1999)
 - [Τάσεις: αντικειμενοστραφή, αρχιτεκτονική πελάτη-εξυπηρέτη, κατακευματισμένες, έμπειρα]
- **Δεκαετία του 1990**
 - εμπορικά αντικειμενοστραφή συστήματα
 - [Τάσεις: πολυβάσεις, χωρικές & χρονικές, πολυμέσα, συμπερασματικές, αποθήκες δεδομένων (αναλυτική επεξεργασία), προγραμματισμό πόρων της επιχείρησης (ERP – Enterprise Resource Planning) και της διαχείρισης τους (MRP – Management Resource Planning), Internet]
- **Δεκαετία του 2000**
 - Σύστημα Διαχείρισης Χρωμοσωμάτων (Human Genome Project)
 - Σύστημα Παρατήρησης της Γης (Earth Observation System)

Ιστορία

- Σήμερα, 2004 - ...
 - Μεγάλος όγκος δεδομένων (BIG DATA)
 - Αλλαγές σε υλικό (επεξεργαστές με πολλούς πυρήνες, κλπ)
 - cloud computing
 - data lakes
 - AI & ML
- *NoSQL*
 - **MapReduce** (2004) και **Bigtable** (2006) by Google, **Dynamo** (2007) by Amazon
 - **Hadoop** (βασισμένο στο MapReduce, 2006), **Cassandra** (επηρεασμένη από Bigtable και Dynamo papers) και **MongoDB** (2009)
- *NotOnlySQL*
 - *Join! SQL interfaces* πάνω στο Hadoop (και αργότερα στη Spark)
- *NewSQL*
 - **Spanner** (2012) by Google **PostgreSQL 10** (native support for JSON, κλπ)

Παράδειγμα: μια πρώτη μοντελοποίηση

- Έχουμε εκατοντάδες CD από εφημερίδες και περιοδικά.
- Τα περισσότερα είναι συλλογές τραγουδιών διάφορων καλλιτεχνών.
- Θέλουμε να κατασκευάσουμε μια εφαρμογή αναζήτησης με βάση λέξεις κλειδιά σχετικές με
 - τον τίτλο του CD ή
 - το όνομα του καλλιτέχνη ή
 - τον τίτλο του τραγουδιού

Παράδειγμα: Διάφορες προσεγγίσεις

- Καρτέλες, μια για κάθε CD, ταξινομημένες αλφαβητικά ως προς τον τίτλο του CD.
- Πιθανόν ok για να βρίσκω τα CD, αλλά στην περίπτωσή μας δεν βοηθούν:
 - Και αν θέλω να ψάξω για όλα τα τραγούδια ενός συγκεκριμένου καλλιτέχνη;
 - Ή όλα τα τραγούδια που περιέχουν μια λέξη κλειδί;

Παράδειγμα: Διάφορες προσεγγίσεις

- Μπορώ να έχω και επιπλέον καρτέλες, μια για κάθε καλλιτέχνη, ταξινομημένες αλφαβητικά ως προς το όνομα του καλλιτέχνη.
- Επίσης, επιπλέον καρτέλες για τα τραγούδια.
- Αντιστοιχία με τις δομές δεδομένων (εγγραφές σε πίνακες, λίστες, ουρές, κλπ.)

Παράδειγμα: Διάφορες προσεγγίσεις

The screenshot shows a LibreOffice Calc spreadsheet with the following data:

	A	B	C	D	E	F
	artist_name	cd_title	year	pot	song_title	
2	Adele	Adele 21	2011	1	Rolling in the Deep	
3	Adele	Adele 21	2011	2	Rumour Has It	
4	Adele	Adele 21	2011	3	Turning Tables	
5	Adele	Adele 21	2011	4	Don't You Remember	
6	Adele	Adele 21	2011	5	Set Fire to the Rain	
7	Adele	Adele 21	2011	6	He Won't Go	
8	Adele	Adele 21	2011	7	Take It All	
9	Adele	Adele 21	2011	8	I'll Be Waiting	
10	Adele	Adele 21	2011	9	One and Only	
11	Adele	Adele 21	2011	10	Lovesong	
12	Adele	Adele 21	2011	11	Someone Like You	
13	Adele	Adele 21	2011	12	If It Hadn't Been for Love	
14	Adele	Adele 21	2011	13	Hiding My Heart	
15	Buddy Guy	The Definitive Buddy Guy	2009	1	Sit And Cry (The Blues)	
16	Buddy Guy	The Definitive Buddy Guy	2009	2	First Time I Met The Blues	
17	Buddy Guy	The Definitive Buddy Guy	2009	3	Ten Years Ago	
18	Buddy Guy	The Definitive Buddy Guy	2009	4	Stone Crazy	
19	Buddy Guy	The Definitive Buddy Guy	2009	5	When My Left Eye Jumps	
20	Buddy Guy	The Definitive Buddy Guy	2009	6	Hoodoo Man Blues	
21	Buddy Guy	The Definitive Buddy Guy	2009	7	A Man And The Blues	
22	Buddy Guy	The Definitive Buddy Guy	2009	8	I Got My Eyes on You (live)	
23	Buddy Guy	The Definitive Buddy Guy	2009	9	Five Long Years	
24	Buddy Guy	The Definitive Buddy Guy	2009	10	A Man Of Many Words	
25	Buddy Guy	The Definitive Buddy Guy	2009	11	The Things I Used To Do (live)	
26	Buddy Guy	The Definitive Buddy Guy	2009	12	When I Left Home	
27	Buddy Guy	The Definitive Buddy Guy	2009	13	Give Me My Coat And Shoes	

Παράδειγμα: προβλήματα flat table

Όπως και στο παράδειγμα της φρουτεμπορικής πολλαπλή καταχώριση της ίδιας πληροφορίας:

- Δυσκολία εισαγωγής, διαγραφής, ενημέρωσης
- Ασυνέπεια στα δεδομένα
- Μεγάλος όγκος
- Απώλεια πληροφορίας

Βελτίωση;

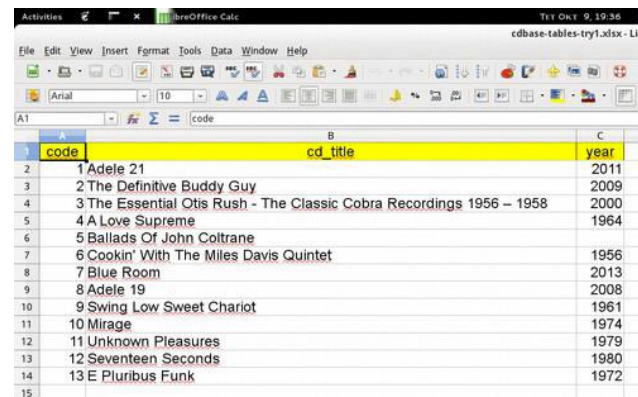
Καλλιτέχνες



A screenshot of the LibreOffice Calc application window. The spreadsheet has two columns: 'code' and 'artist_name'. The data is as follows:

code	artist_name
1	Adele
2	Buddy Guy
3	Otis Rush
4	John Coltrane
5	Miles Davis
6	Madeleine Peyroux
7	The Staples Singers
8	Camel
9	Joy Division
10	Cure
11	Grand Funk Railroad

CD



A screenshot of the LibreOffice Calc application window. The spreadsheet has three columns: 'code', 'cd_title', and 'year'. The data is as follows:

code	cd_title	year
1	Adele 21	2011
2	The Definitive Buddy Guy	2009
3	The Essential Otis Rush - The Classic Cobra Recordings 1956 – 1958	2000
4	A Love Supreme	1964
5	Ballads Of John Coltrane	
6	Cookin' With The Miles Davis Quintet	1956
7	Blue Room	2013
8	Adele 19	2008
9	Swing Low Sweet Chariot	1961
10	Mirage	1974
11	Unknown Pleasures	1979
12	Seventeen Seconds	1980
13	E Pluribus Funk	1972

Τραγούδια

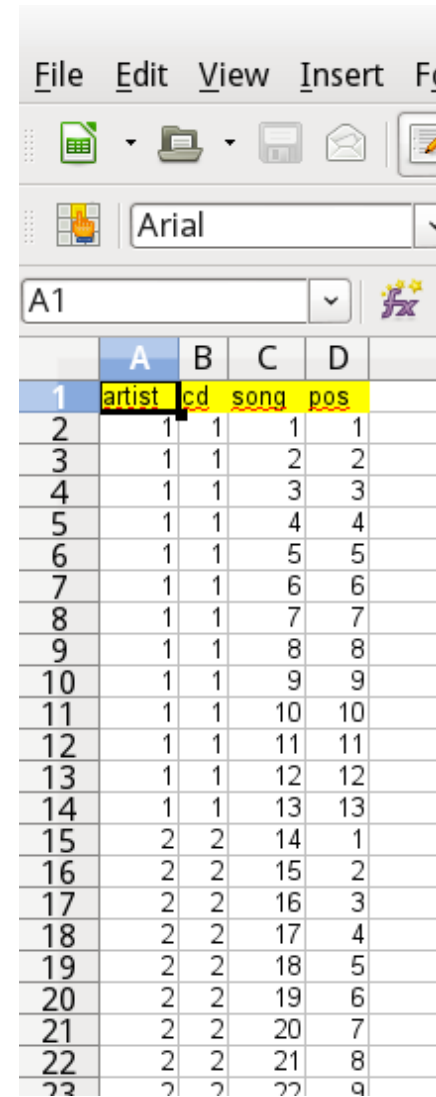


A screenshot of the LibreOffice Calc application window. The spreadsheet has two columns: 'code' and 'song_title'. The data is as follows:

code	song_title
1	Rolling in the Deep
2	Rumour Has It
3	Turning Tables
4	Don't You Remember
5	Set Fire to the Rain
6	He Won't Go
7	Take It All
8	I'll Be Waiting
9	One and Only
10	Lovesong
11	Someone Like You
12	If It Hadn't Been for Love
13	Hiding My Heart
14	Sit And Cry (The Blues)
15	First Time I Met The Blues
16	Ten Years Ago
17	Stone Crazy
18	When My Left Eye Jumps
19	Hoodoo Man Blues
20	A Man And The Blues
21	I Got My Eyes on You (live)
22	Five Long Years
23	A Man Of Many Words
24	The Things I Used To Do (live)
25	When I Left Home

Τι άλλο χρειαζομαστε;

- οι συσχετίσεις ανάμεσα στις οντότητες
 - Είναι η δυναμική πληροφορία που παράγεται κατά τη λειτουργία του πληροφοριακού συστήματος της επιχείρησης /οργανισμού



The image shows a screenshot of a Microsoft Excel spreadsheet. The spreadsheet has a menu bar at the top with 'File', 'Edit', 'View', 'Insert', and 'Fo'. Below the menu bar is a toolbar with icons for various functions. The spreadsheet itself has a grid with columns labeled 'A', 'B', 'C', and 'D' and rows numbered 1 through 22. The first row (row 1) has headers: 'artist' in column A, 'cd' in column B, 'song' in column C, and 'pos' in column D. The data in the spreadsheet is as follows:

	A	B	C	D
1	artist	cd	song	pos
2	1	1	1	1
3	1	1	2	2
4	1	1	3	3
5	1	1	4	4
6	1	1	5	5
7	1	1	6	6
8	1	1	7	7
9	1	1	8	8
10	1	1	9	9
11	1	1	10	10
12	1	1	11	11
13	1	1	12	12
14	1	1	13	13
15	2	2	14	1
16	2	2	15	2
17	2	2	16	3
18	2	2	17	4
19	2	2	18	5
20	2	2	19	6
21	2	2	20	7
22	2	2	21	8
23	2	2	22	9

R. ELMASRI - S.B. NAVATHE

ΘΕΜΕΛΙΩΔΕΙΣ ΑΡΧΕΣ ΣΥΣΤΗΜΑΤΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

7η ΕΚΔΟΣΗ ΑΝΑΘΕΩΡΗΜΕΝΗ

ΜΕΤΑΦΡΑΣΗ - ΕΠΙΜΕΛΕΙΑ
ΜΙΧΑΛΗΣ ΧΑΤΖΟΠΟΥΛΟΣ
ΚΑΘΗΓΗΤΗΣ
ΤΜΗΜΑΤΟΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΑΘΗΝΩΝ

ΔΙΑΥΛΟΣ

Ευχαριστώ!

- <https://eclass.uoa.gr/courses/DIND136/> Έγγραφα > Διαλέξεις
- Βιβλία:
 - Ενότητα 1, Κεφάλαιο 1. *ΒΔ & Χρήστες ΒΔ*