

Εξόρυξη Δεδομένων

ΑΠΟΘΗΚΕΣ ΔΕΔΟΜΕΝΩΝ Data Warehouses



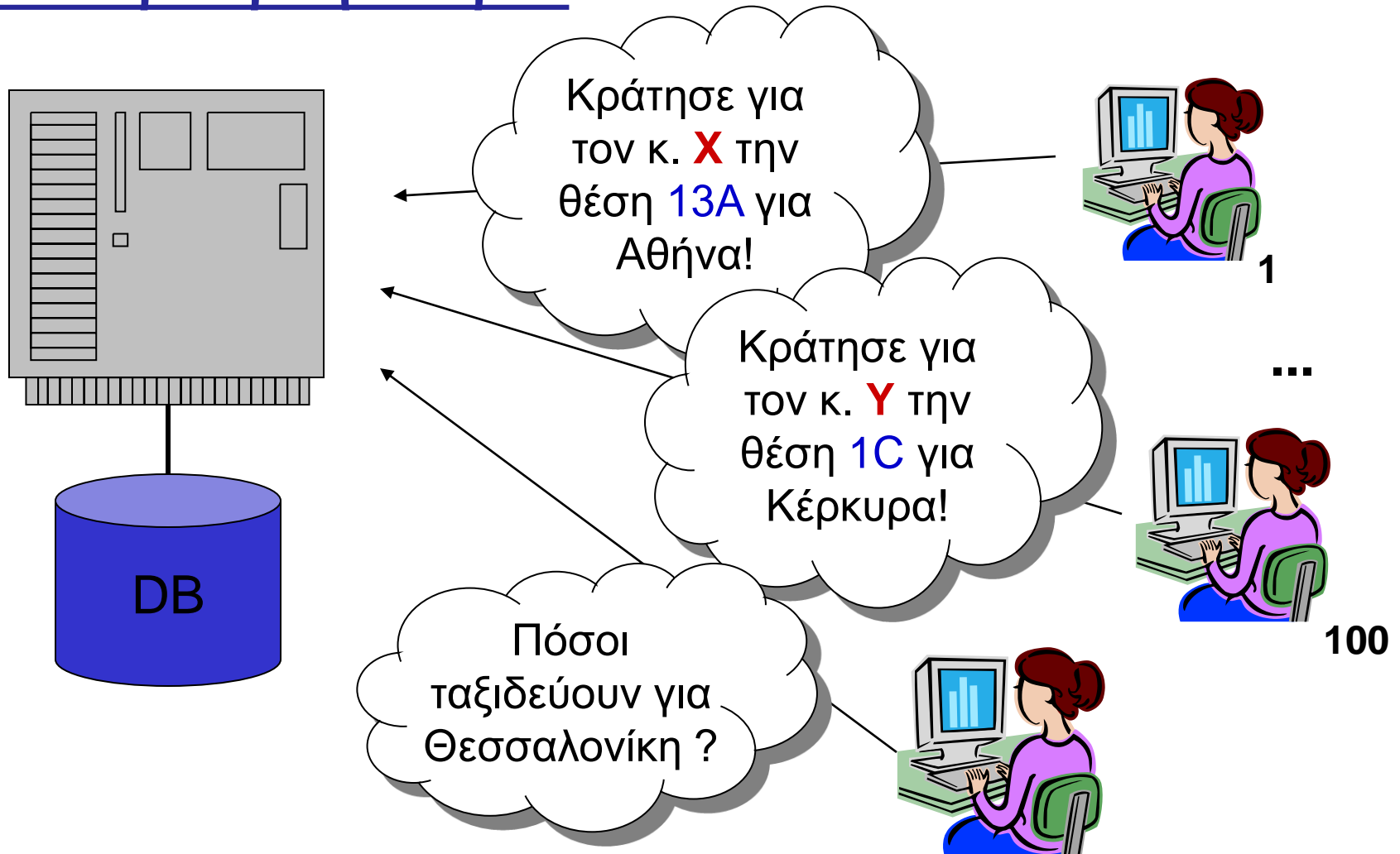
On-Line Transaction Processing (OLTP)

On-Line Transaction Processing (OLTP)

- Το **Σύστημα Επεξεργασίας Δοσοληψιών** είναι ένα πλήρες σύστημα που περιέχει εργαλεία για τον προγραμματισμό εφαρμογών, την εκτέλεση και τη διαχείριση δοσοληψιών.
- Μια τέτοια εφαρμογή πρέπει να δουλεύει συνεχώς, να αντεπεξέρχεται αποτυχιών, να εξελίσσεται συνεχώς. Είναι συνήθως κατανεμημένη και περιλαμβάνει:
 - **Σχεσιακή Βάση Δεδομένων**
 - **Δίκτυο**
 - **Προγράμματα εφαρμογής**
- Είναι εξαιρετικά κρίσιμη για τη λειτουργία κάθε οργανισμού.

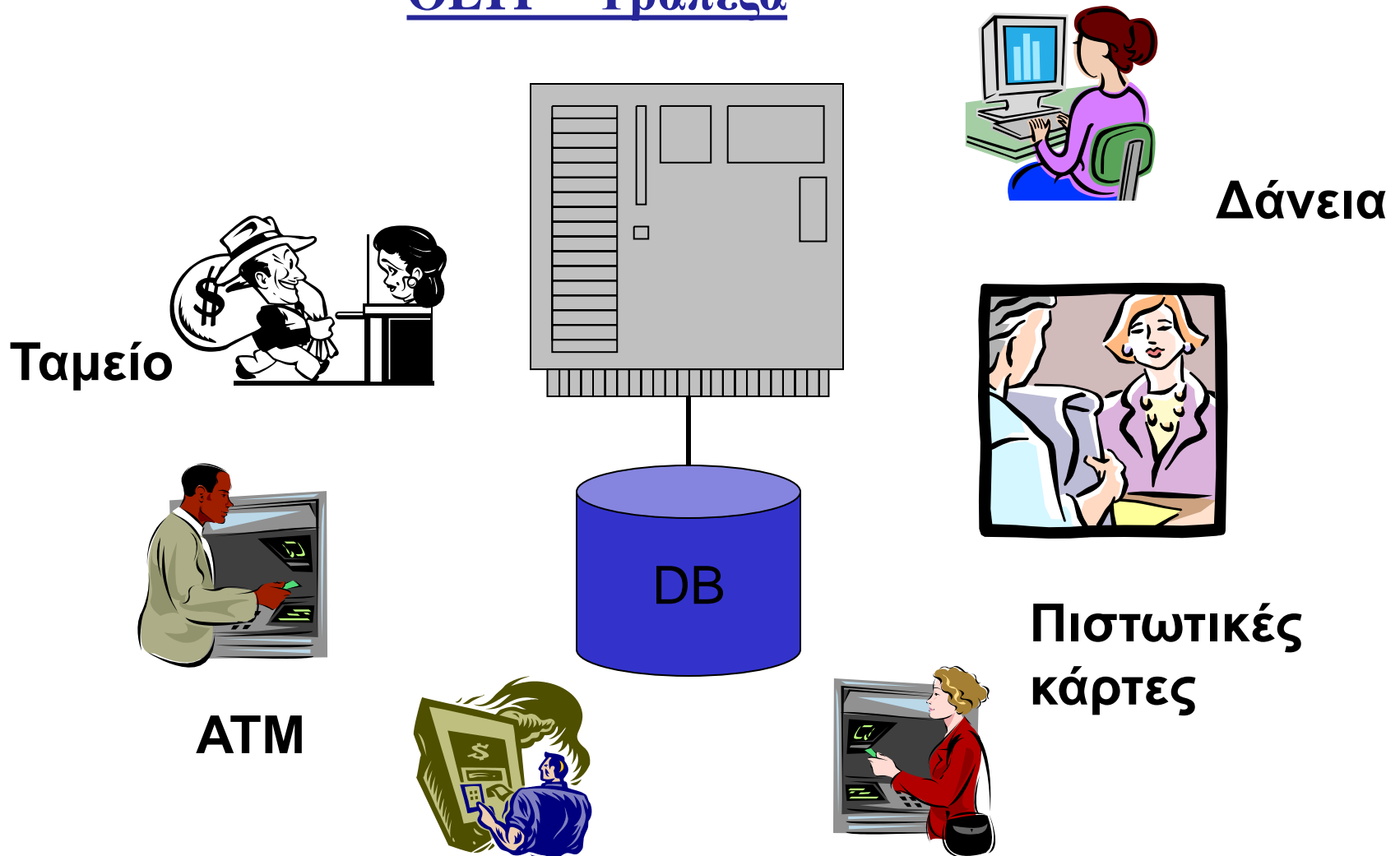
Παραδείγματα On-Line Transaction Processing

OLTP – Αεροπορική Εταιρεία



Παραδείγματα On-Line Transaction Processing

OLTP – Τράπεζα



Τα βασικά χαρακτηριστικά ενός OLTP

- Ελάχιστος χρόνος διαθέσιμος για την εκτέλεση μιας δοσοληψίας.
- Περιορισμένος αριθμός προσβάσεων στον δίσκο (ή δίσκους).
- Περιορισμένος αριθμός υπολογισμών.
- Κάτω όριο λειτουργικών απαιτήσεων:
 - τουλάχιστον 100 TPS σε μια ΒΔ της τάξης του 1 GB
- Άνω όριο λειτουργικών απαιτήσεων:
 - μέχρι 50000 TPS σε μία ΒΔ μεγαλύτερη του 1 TB.

Προβλήματα ενός OLTP

- **Αρκετά δεδομένα μπορεί να είναι ελλιπή (missing data)**

ΠΡΟΤΕΙΝΟΜΕΝΕΣ ΛΥΣΕΙΣ

- Απόρριψη εγγραφών των οποίων κάποιες τιμές λείπουν
- Αντικατάσταση των τιμών που λείπουν με την μέση τιμή της κατηγορίας
- Αντικατάσταση των τιμών που λείπουν από κάποια χαρακτηριστικά με τις αντίστοιχες τιμές παρόμοιων χαρακτηριστικών

Προβλήματα ενός OLTP

- **Αρκετά δεδομένα μπορεί να έχουν θόρυβο και ασυνέπειες (noisy and inconsistent data)**

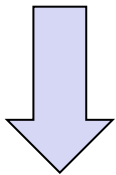
ΣΤΡΑΤΗΓΙΚΕΣ

- **Εντοπισμός διπλότυπων εγγραφών**
 - π.χ. Ιωάννης & Γιάννης
- **Εντοπισμός λανθασμένων τιμών**
 - π.χ. Μηδενική τιμή σε ένα χαρακτηριστικό που αφορά το βάρος ενός προϊόντος
- **Εξομάλυνση δεδομένων (κβαντισμός)**
 - π.χ. περιορισμός του αριθμού των τιμών ενός αριθμητικού χαρακτηριστικού σε συγκεκριμένες τιμές (ή στάθμες)

Προβλήματα ενός OLTP

- Τα δεδομένα μπορεί να είναι ετερογενή και να παρουσιάζουν συνωνυμίες ή αμφισημίες
- π.χ. ένας πελάτης είναι καταχωρημένος με διαφορετικά στοιχεία στη βάση του τμήματος πωλήσεων από ότι στη βάση του τμήματος μάρκετινγκ.

ΣΤΡΑΤΗΓΙΚΕΣ



Εννοιολογική εναρμόνιση των δεδομένων

Source 1:
Personnel
(Cobol)

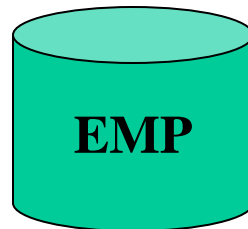
<u>EMP ID</u>	Name	Birth	Salary	Total Income	DeptID
110	Kostas	1/1/72	1500	1200	132
...

Source 2: Accounting
(DB2)

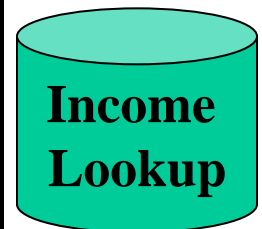
<u>EMP ID</u>	<u>IL ID</u>	Amount
110	10	1500
110	30	300



<u>EMP ID</u>	Name	Age
110	Kostas	30
120	Vasilis	48
130	Roula	29



<u>IL ID</u>	Descr
10	Μισθός
20	Επίδομα Τέκνων
30	Φόρος
...	...

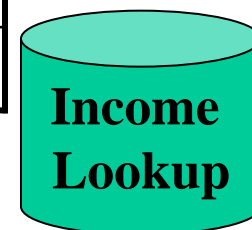


Αποκανονικοποίηση - Denormalization

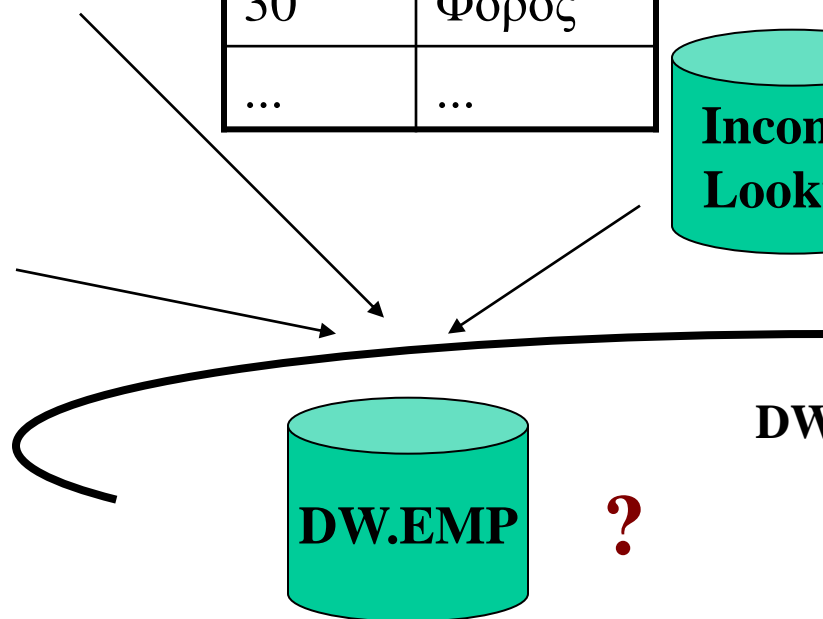
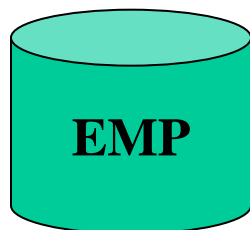
<u>EMP ID</u>	<u>IL ID</u>	Amount
110	10	1500
110	30	300



<u>IL ID</u>	Descr
10	Μισθός
20	Επίδομα Τέκνων
30	Φόρος
...	...



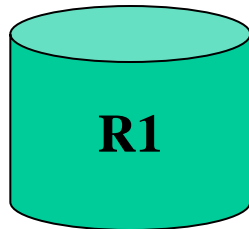
<u>EMP ID</u>	Name	Age
110	Kostas	30
120	Vasilis	48
130	Roula	29



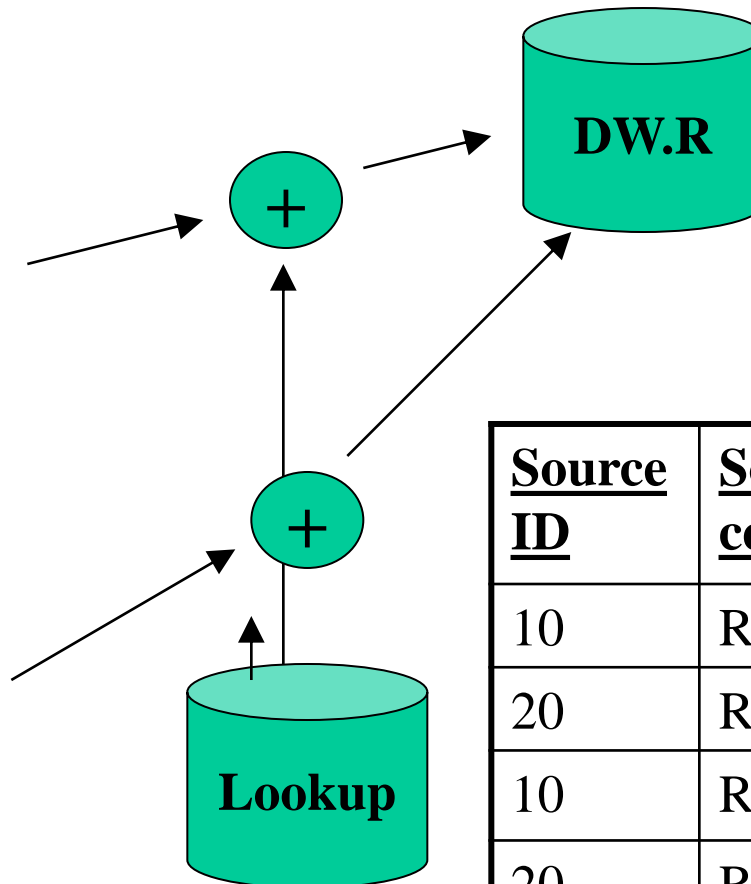
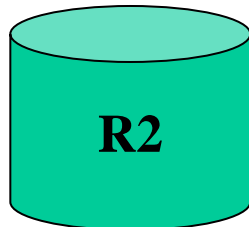
Καθαρισμός των δεδομένων

Ομογενοποίηση κλειδιών

<u>ID</u>	Descr
10	Coca
20	Pepsi



<u>ID</u>	Descr
10	Pepsi
20	Epsa



<u>ID</u>	Descr
100	Coca
110	Pepsi
120	Epsa

<u>Source ID</u>	<u>Sour ce</u>	Surrogate Key
10	R1	100
20	R1	110
10	R2	110
20	R2	120

Μετασχηματισμός Δεδομένων

■ Κανονικοποίηση Δεδομένων

■ Δεκαδική κλιμάκωση

■ Κανονικοποίηση Ελαχίστου-Μεγίστου

$$newValue = \frac{originalValue - oldMin}{oldMax - oldMin}$$

■ Κανονικοποίηση με χρήση τιμών Z-scores

$$newValue = \frac{originalValue - \mu}{\sigma}$$

■ Μετατροπή τύπων δεδομένων

■ Επιλογή χαρακτηριστικών και στιγμιοτύπων

Προβλήματα ενός OLTP

- **Λόγω διαδικασιών ενημέρωσης, εισαγωγής, διαγραφής διατηρούν δεδομένα μόνο για την τρέχουσα κατάσταση.**
 - π.χ., στη βάση του τμήματος προμηθειών διατηρούνται μόνο όσοι προμηθευτές συνεργάζονται αυτή τη στιγμή ενώ μπορεί να χρειασθούν δεδομένα και για προμηθευτές που συνεργαζόταν στο παρελθόν προκειμένου να συγκριθούν οι τιμές τους
- **Η ανάλυση των δεδομένων δεν είναι εύκολο να επιτευχθεί με εργαλεία όπως η γλώσσα SQL.**
 - Προκύπτουν περίπλοκα ερωτήματα SQL τα οποία δεν είναι εύκολο να συνταχθούν.

Προβλήματα ενός OLTP

- Σε φυσικό επίπεδο οργάνωσης (π.χ. δενδρικές μέθοδοι προσπέλασης) δεν είναι σχεδιασμένα για να ανταποκρίνονται στις **απαιτήσεις περίπλοκων ερωτημάτων**.
- Τα δεδομένα οργανώνονται με **διάγραμμα Οντοτήτων-Συσχετίσεων (ER)** και αρχών όπως η **κανονικοποίηση**, που παράγουν περίπλοκες βάσεις στο νοητικό επίπεδο.

On-Line Analytical Processing (OLAP & OLAM)

- Ανάλυση μεγάλου όγκου σύνθετων δεδομένων από διαφορετικές εφαρμογές
- Ειδικού τύπου ερωτήματα συνάθροισης
- Οπτικοποίηση / στατιστική ανάλυση / πολυδιάστατη ανάλυση
- Εξόρυξη Γνώσης (**Knowledge Discovery / Data Mining**)
- Εξεύρεση **προτύπων** σε τεράστιες βάσεις δεδομένων

OLAP + Data Mining = On-Line Analytical Mining

Παραδείγματα ερωτημάτων στο OLAP

- Ποιος ήταν ο όγκος πωλήσεων ανά περιοχή και κατηγορία προϊόντος την περασμένη χρονιά;
- Πόσο σχετίζονται οι αυξήσεις τιμών των υπολογιστών με τα κέρδη των πωλήσεων τα 10 τελευταία χρόνια;
- Ποια ήταν τα δέκα πρώτα καταστήματα σε πωλήσεις CD;
- Τι ποσοστό από τους πελάτες που αγοράζουν αναψυκτικά αγοράζουν και πατατάκια;

Λειτουργικά Χαρακτηριστικά Απαιτήσεων OLAP

- Πρόσβαση σε **μεγάλο όγκο** δεδομένων
- Συμμετοχή **αθροιστικών** και **ιστορικών** δεδομένων σε πολύπλοκες ερωτήσεις
- Μεταβολή της **οπτικής γωνίας** της παρουσίασης των δεδομένων (π.χ. από τις πωλήσεις ανά περιοχή στις πωλήσεις ανά τμήμα κλπ.)
- Συμμετοχή πολύπλοκων υπολογισμών (π.χ. **στατιστικές συναρτήσεις**)
- Γρήγορη απάντηση σε οποιαδήποτε χρονική στιγμή τεθεί ένα ερώτημα (**On-Line-AP**).

OLTP vs. OLAP

	OLTP	OLAP
Λειτουργίες	Αυτοματισμός καθημερινών εργασιών	Χάραξη στρατηγικής
Τύπος Δεδομένων	Λεπτομερή	Αθροιστικά
Όγκος Δεδομένων	~ 100 GB	~ 1 TB και πάνω
Φύση Δεδομένων	Δυναμικά	Στατικά, Ιστορικά

OLTP vs. OLAP

	OLTP	OLAP
I/O Τύποι	Περιορισμένο I/O	Εκτεταμένο I/O
Τροποποιήσεις	Συνεχείς Ενημερ.	Περιοδικές Ενημερ.
Φόρτος	Δοσοληψίες με πρόσβαση λίγων εγγραφών	Ερωτήσεις που σαρώνουν εκατομμύρια εγγραφές
Σχεδίαση ΒΔ	Κατευθυνόμενη από την Εφαρμογή	Κατευθυνόμενη από το Περιεχόμενο

OLTP vs. OLAP

OLTP

Χαμηλόβαθμοι Υπ.

Μέσω
προκατασκευασμένων
φορμών

Χιλιάδες

Εισαγωγή
Δεδομένων

OLAP

Υψηλόβαθμοι Υπ.

Ad-hoc

Δεκάδες

Εξαγωγή
Πληροφοριών

Τυπικοί Χρήστες

Χρήση

Αριθμός Χρηστών

Εστίαση

Αποθήκες δεδομένων

- Η τεχνολογία των αποθηκών δεδομένων προσφέρει
 - ολοκλήρωση **ετερογενών πηγών δεδομένων** και
 - πλατφόρμα για αποδοτική ανάλυση **ιστορικών δεδομένων**
- Μία αποθήκη δεδομένων αποτελεί μία συλλογή δεδομένων που
 - επιλέγονται από τις **Επιχειρησιακές Βάσεις Δεδομένων**,
 - Ολοκληρώνονται (**integrated**),
 - τα δεδομένα αναλύονται με διαδικασίες όπως η **On-line Analytical Processing (OLAP)** ή η εξόρυξη δεδομένων.

Αποθήκες δεδομένων

- Είναι μια κεντριοποιημένη ΒΔ με σκοπό:
 - την ολοκλήρωση (**integration**) ετερογενών πηγών πληροφοριών (**data sources**) που σημαίνει **την συνάθροιση όλης της ενδιαφέρουσας πληροφορίας σε μία τοποθεσία**
 - την αποφυγή της σύγκρουσης μεταξύ OLTP και OLAP (DSS) συστημάτων με αποδοτικές εφαρμογές και διαχείριση της διαθεσιμότητα του συστήματος
- Μπορεί να συμπληρώνεται και από εξειδικευμένα θεματικά υποσύνολα (**Data Marts**) για περαιτέρω απόδοση των OLAP εφαρμογών

Αποθήκες δεδομένων

Λειτουργικά Χαρακτηριστικά

Ιστορικά Δεδομένα

- Ο χρονικός ορίζοντας μιας αποθήκης δεδομένων είναι πολύ μεγαλύτερος από ότι ενός συστήματος σε συνεχή λειτουργία
- Η ΒΔ έχει τα τωρινά δεδομένα ενώ οι αποθήκες διατηρούν και παλιά δεδομένα (π.χ. των προηγούμενων 5-10 χρόνων)

Τροποποιήσεις

- Οι τροποποιήσεις στις πηγές δεδομένων δεν φαίνονται άμεσα στις αποθήκες δεδομένων καθώς εκτελούνται συνήθως περιοδικά
- Έχουν δύο βασικές λειτουργίες: το αρχικό φόρτωμα των δεδομένων (loading) και την προσπέλαση δεδομένων (access)

Ιστορικά Στοιχεία για Αποθήκες Δεδομένων

- Πρώτες προσπάθειες στα μέσα της δεκαετίας του 1990 που οι αποθήκες δεδομένων εξελίχθηκαν σε αγορά της τάξης των **2 δις \$**.
- Μέχρι τα τέλη της δεκαετίας του 1990, **το 95% των 1000 επιχειρήσεων του Fortune** ανέπτυσαν αποθήκες δεδομένων, οπότε η αγορά των αποθηκών δεδομένων ανήρθε οικονομικά στην τάξη των **7 δις \$**.
- Εκτιμήθηκε ότι σε **3** χρόνια από την ανάπτυξη μίας αποθήκης δεδομένων, η απόσβεση γίνεται σε ποσοστό **400%**.

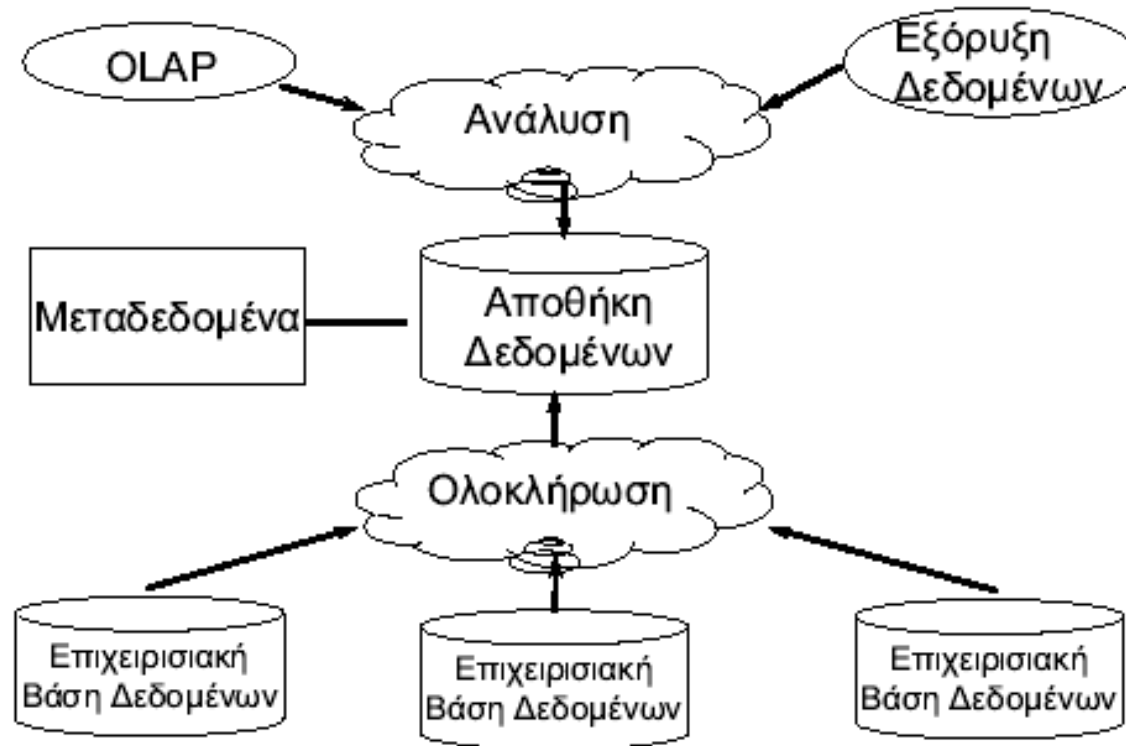
Ιστορικά Στοιχεία για Αποθήκες Δεδομένων

Επιτυχημένη περίπτωση εφαρμογής

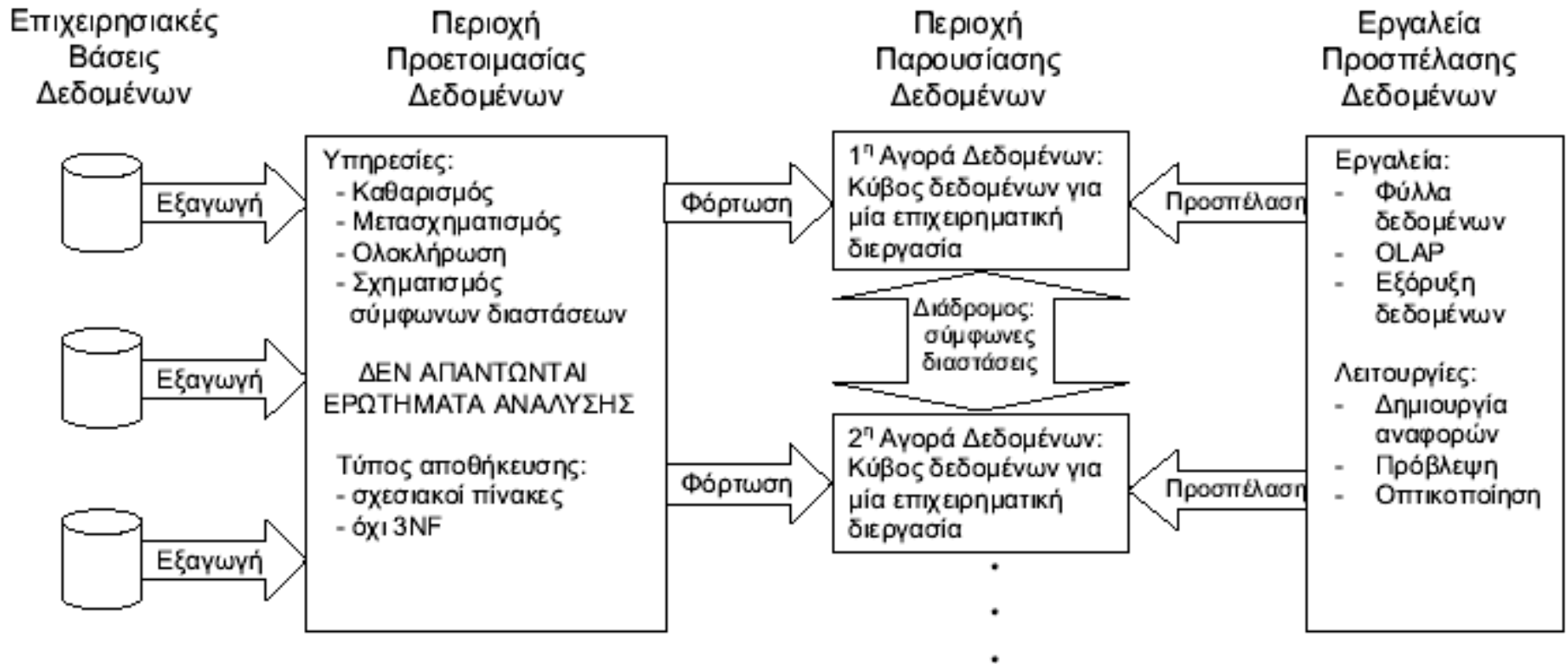
- Στην αλυσίδα υπεραγορών **Walmart**, η οποία δραστηριοποιήθηκε με 2000 υποκαταστήματα,
- βοήθησε την Walmart να βελτιστοποιήσει τις διαδικασίες προμήθειας προϊόντων και να μειώσει το κόστος αγοράς τους κατά 20%.
- Ο όγκος των δεδομένων στην αποθήκη δεδομένων της Walmart ανέρχεται στα 24 TB, σε ένα σύστημα 96 κόμβων με 900 επεξεργαστές και 2700 δίσκους.

Αποθήκες δεδομένων

Βασικές Λειτουργίες



Αρχιτεκτονική Αποθηκών Δεδομένων



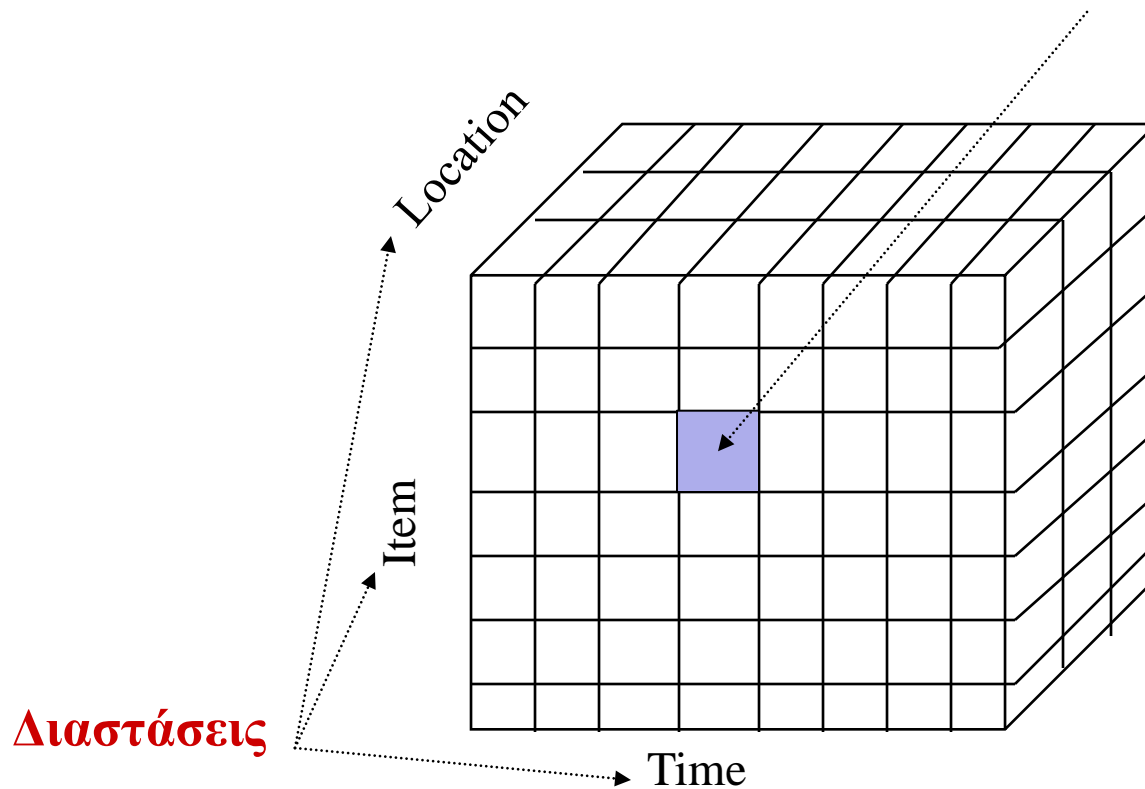
Κύβος Δεδομένων

Ορισμός

- Μια αποθήκη δεδομένων βασίζεται σε ένα **πολυδιάστατο μοντέλο δεδομένων** (multidimensional data model) που αναπαριστά τα δεδομένα με τη μορφή ενός **κύβου** δεδομένων (data cube)
- Ένας κύβος δεδομένων (data cube) επιτρέπει την μοντελοποίηση και την θεώρηση των δεδομένων από πολλές οπτικές γωνίες: τις **διαστάσεις** (dimensions) του
- Για συγκεκριμένες τιμές στις διαστάσεις του κύβου μπορούμε να κάνουμε μία **μέτρηση** (measure) του στοιχείου που μας ενδιαφέρει

Παράδειγμα Κύβου Δεδομένων

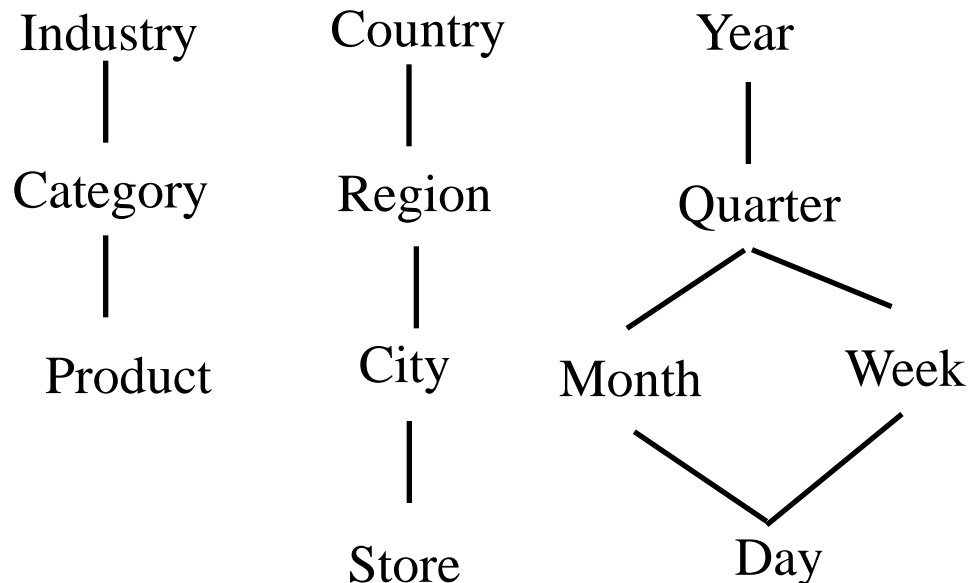
- Μέτρηση: **Πωλήσεις** για κάθε συγκεκριμένο συνδυασμό των τιμών των διαστάσεων (Location, Item, Time)



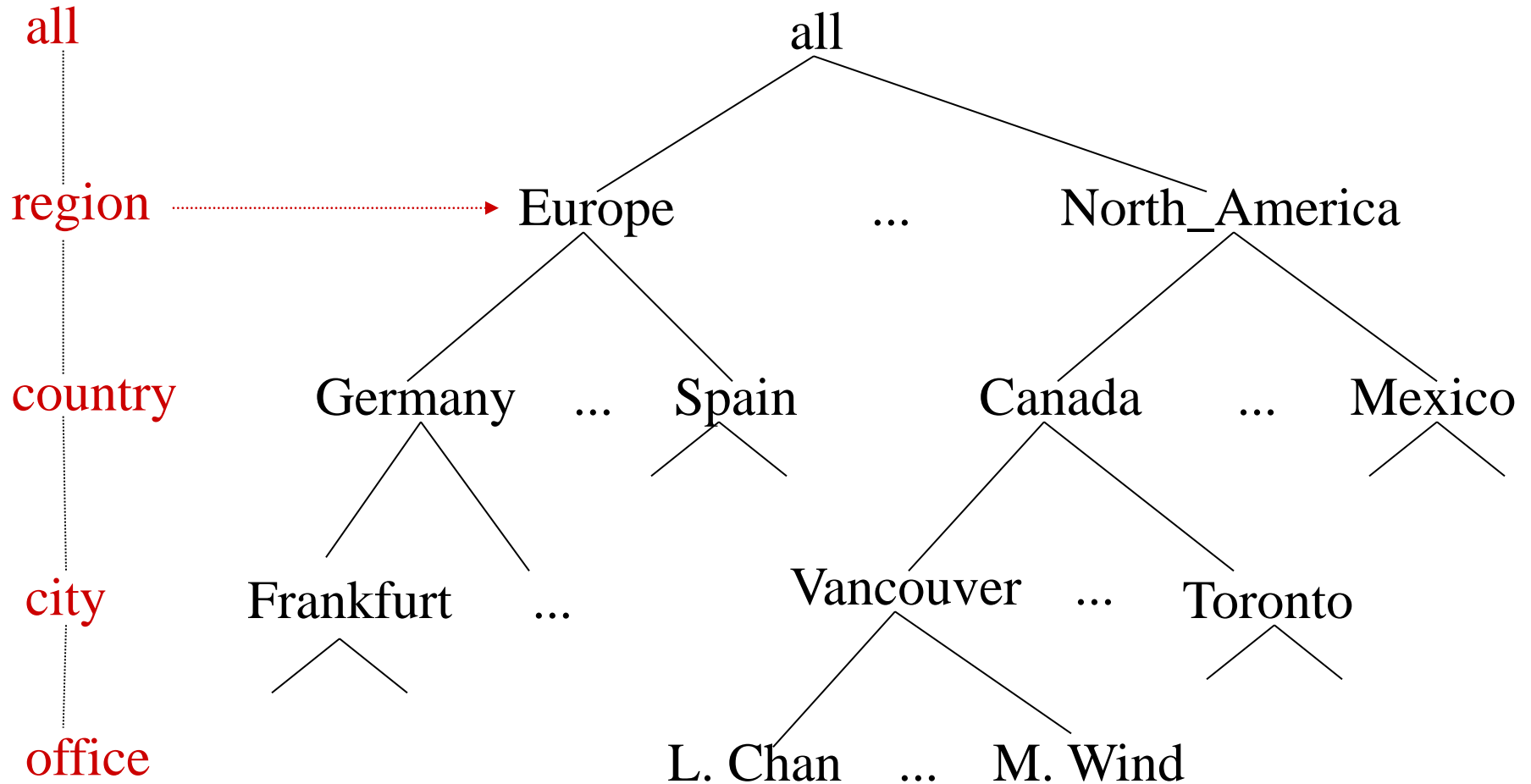
Κύβος ΠΩΛΗΣΕΩΝ

Ιεραρχίες Διαστάσεων

- Κάθε διάσταση παίρνει τιμές από διαφορετικά επίπεδα και μπορεί να εκφραστεί σε διαφορετικά **επίπεδα λεπτομέρειας** (ιεραρχίες)
- Για παράδειγμα για τις διαστάσεις: Product, Region, Date, οι ιεραρχίες θα μπορούσαν να είναι:



Εννοιολογική ιεραρχία (Concept Hierarchy) Παράδειγμα για Location



Αναπαράσταση του Κύβου Μοντέλο Δεδομένων Σχήματος Αστέρα

■ Σχήμα Αστέρα (Star schema)

- Ένας **Πίνακας Γεγονότων** βρίσκεται στο κέντρο και συνδέεται με ένα σύνολο από πίνακες διαστάσεων

■ Ο Πίνακας Γεγονότων (Fact Table) έχει:

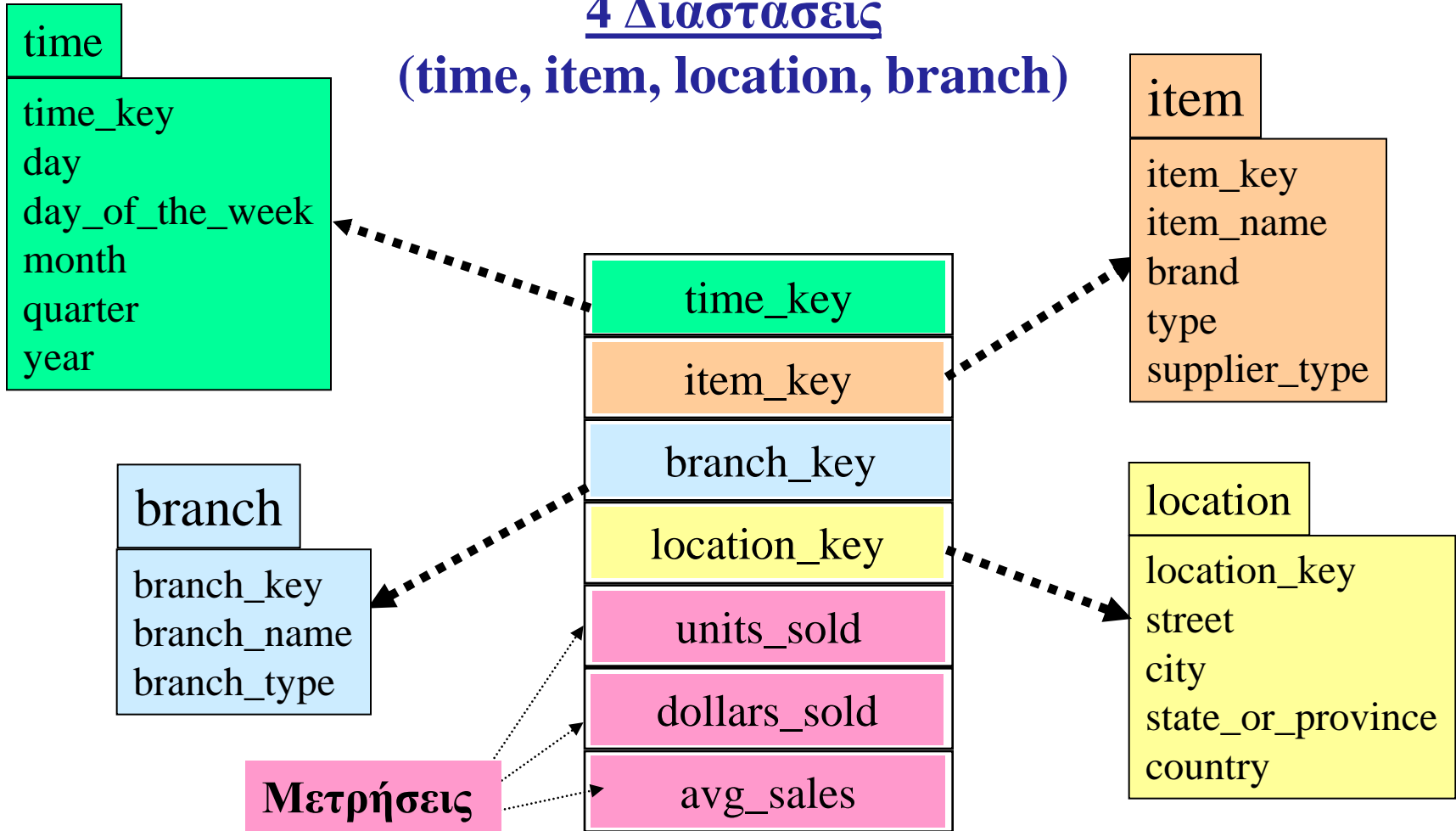
- τις **μετρήσεις** που μας ενδιαφέρουν στο συνδυασμό των διαστάσεων (π.χ. αριθμός πωλήσεων, τιμή σε δολάρια, κλπ.)
- το **πρωτεύον κλειδί** κάθε σχετικού πίνακα διαστάσεων

■ Πίνακες Διαστάσεων

- Πίνακες με πληροφορίες σχετικά με τις τιμές και τις ιεραρχίες της κάθε διάστασης
- π.χ. Item (item_name, brand, type),
- π.χ. Time (day, week, month, quarter, year)

Παράδειγμα Σχήματος Αστέρα

4 Διαστάσεις
(time, item, location, branch)



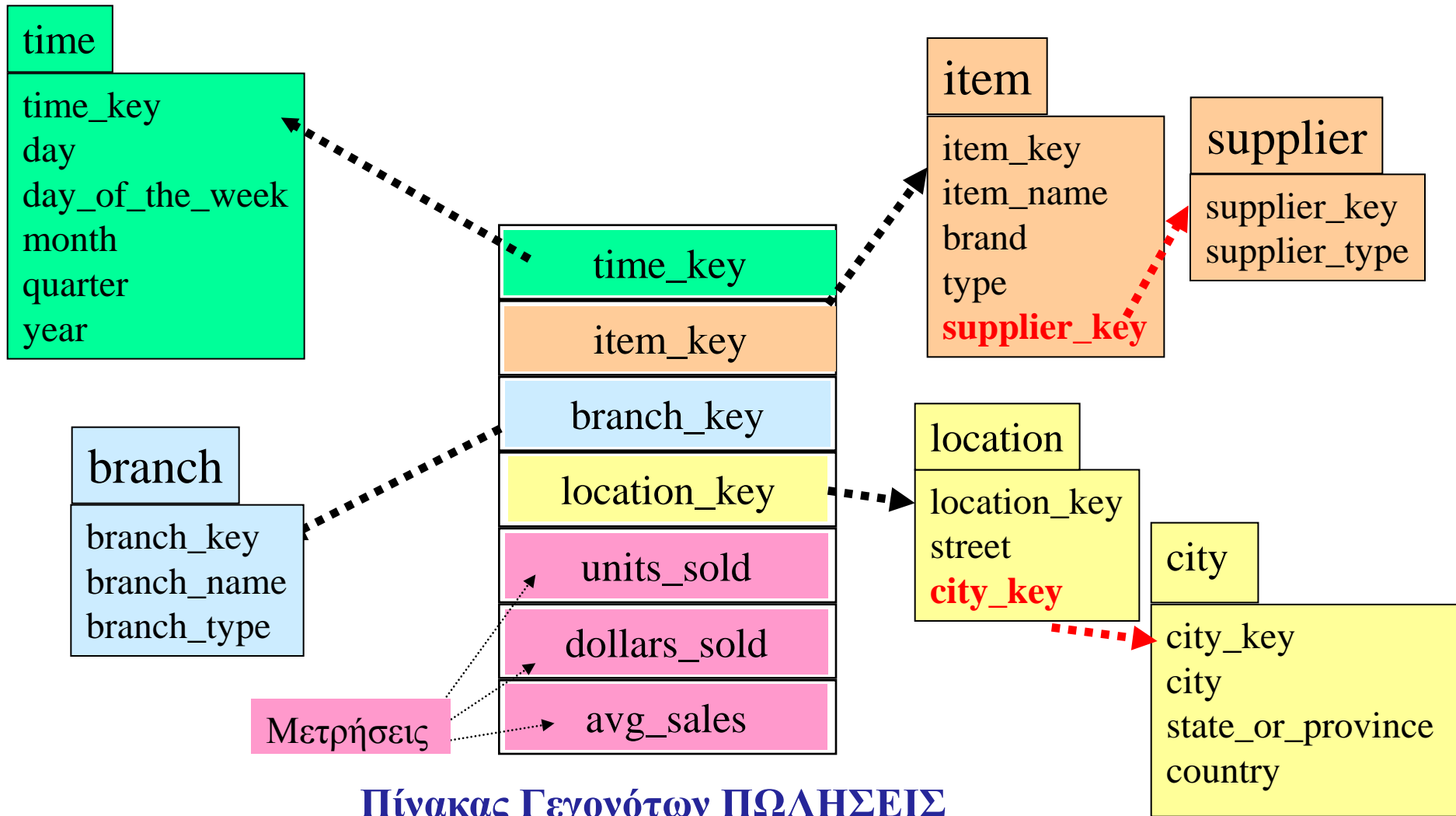
Πίνακας Γεγονότων (ΠΩΛΗΣΕΙΣ)

Μοντέλο Δεδομένων Σχήματος Χιονονιφάδας (Snowflake schema)

- Είναι μια βελτίωση του σχήματος αστέρα σύμφωνα με την οποία **η ιεραρχία** διαστάσεων των πινάκων διάστασης **κανονικοποιείται** σε ένα σύνολο από **μικρότερους πίνακες** διαστάσεων



Παράδειγμα Σχήματος Χιονονιφάδας



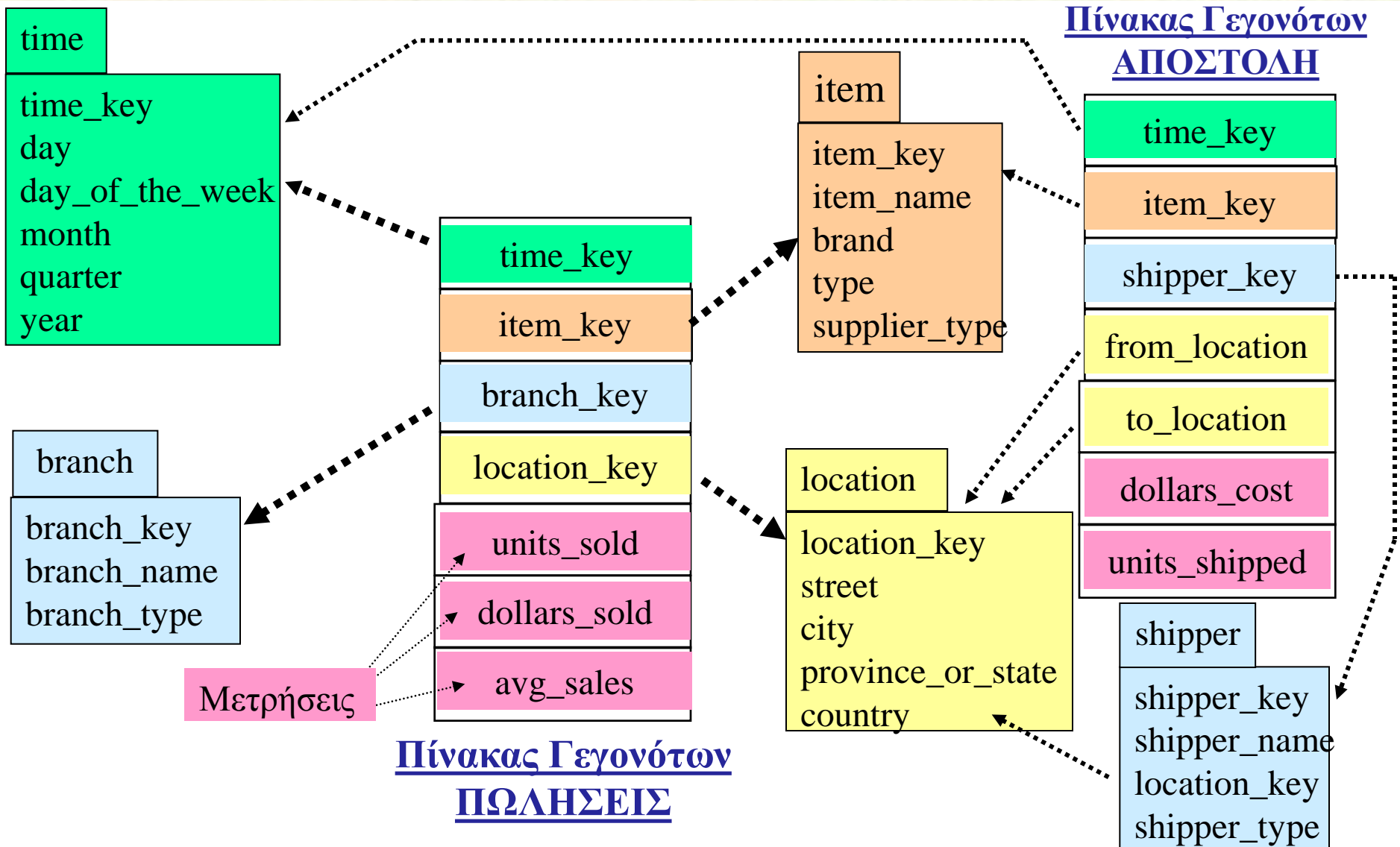
Πίνακας Γεγονότων ΠΩΛΗΣΕΙΣ

Μοντέλο Δεδομένων - Αστερισμοί Γεγονότων (Fact constellations)

- Βασίζεται σε **Πολλαπλούς Πίνακες Γεγονότων** που μοιράζονται τους Πίνακες Διαστάσεων.
- Μπορούμε να τους παρομοιάσουμε ως μία συλλογή από αστέρια ή γαλαξίες



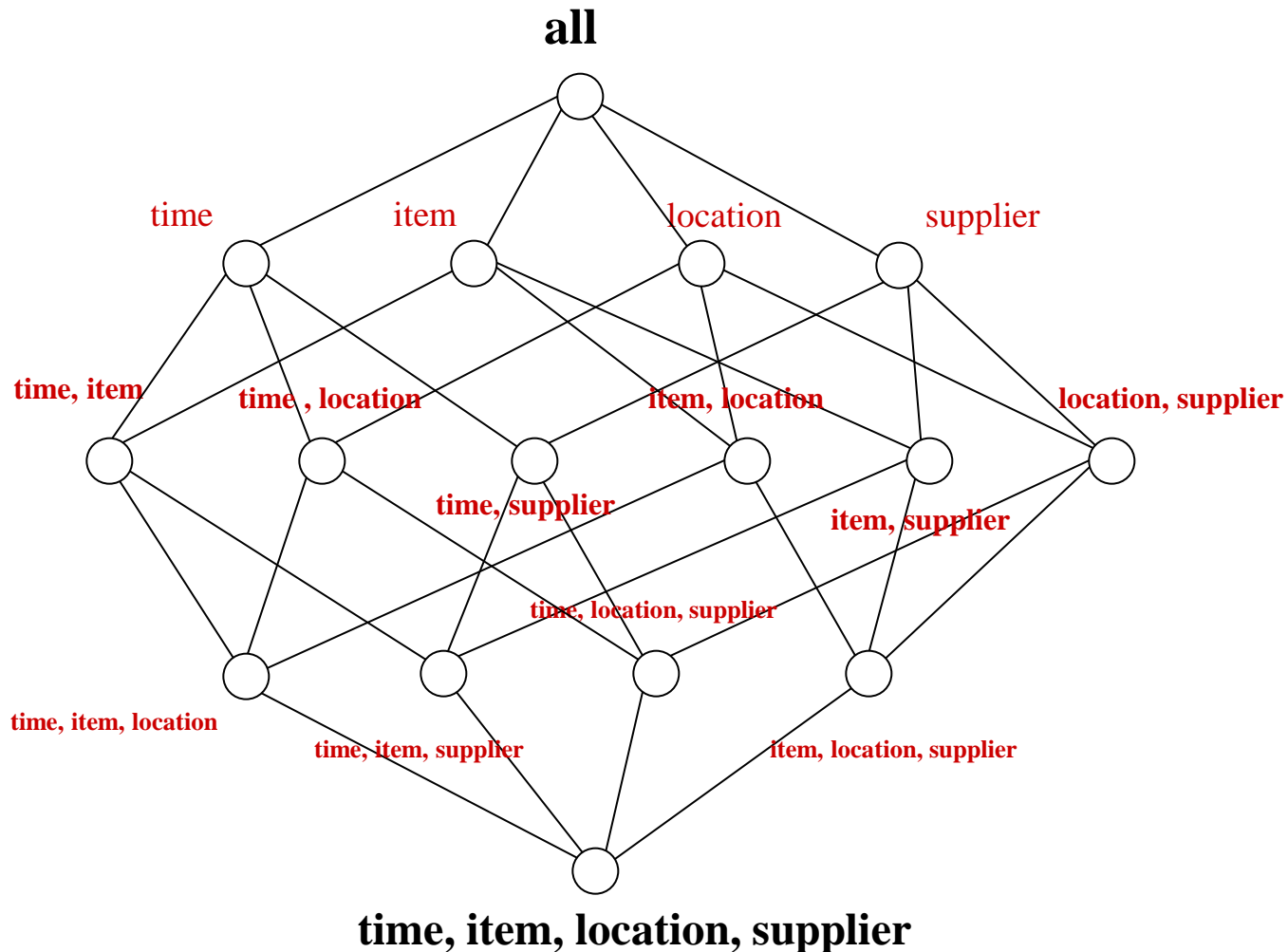
Παράδειγμα Σχήματος Αστερισμών Γεγονότων



Κύβος Δεδομένων

- Συχνά ο n -D κύβος λέγεται **βασικός κυβοειδής** (base cuboid). Περιέχει τον μεγαλύτερο βαθμό λεπτομέρειας.
 - π.χ. Για κύβο με τέσσερις διαστάσεις (Item, Time, Branch, Location) έχουμε 4-D βασικό κυβοειδή
- Ο **κυβοειδής μηδενικής διάστασης** (0-D cuboid) περιέχει το μεγαλύτερο δυνατό επίπεδο περίληψης (apex cuboid).
- Το **πλέγμα όλων των κυβοειδών** κάθε διάστασης από 0 έως και n ονομάζεται κύβος δεδομένων.

Παράδειγμα πλέγματος κυβοειδών (0-4D)



0-D(apex) cuboid

1-D cuboids

2-D cuboids

3-D cuboids

4-D (base) cuboid

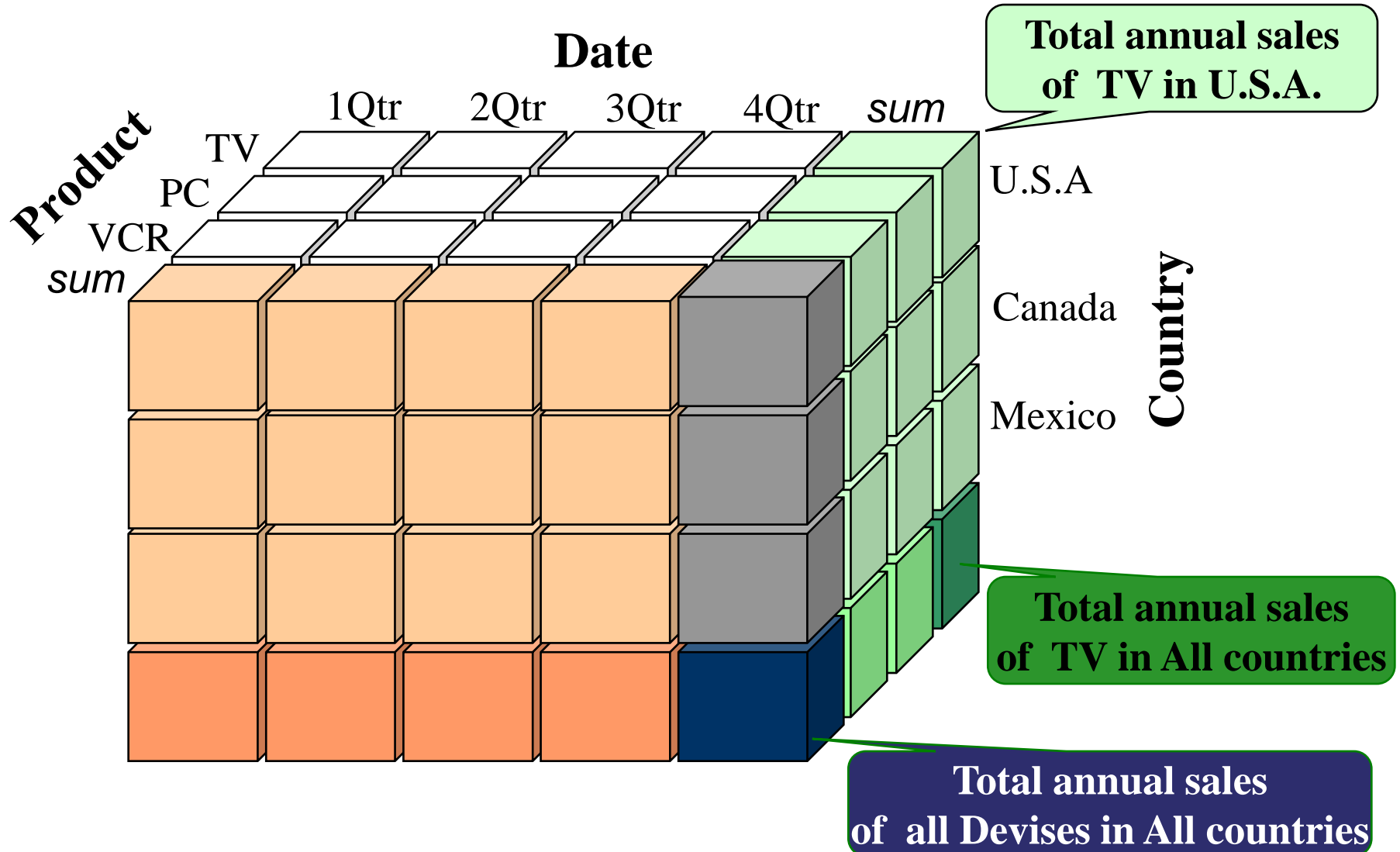
OLAP και Συνάθροιση

- Για την εκτέλεση πράξεων σε ένα σύστημα OLAP απαιτείται ο ορισμός του τρόπου παραγωγής των κύβων-αποτελεσμάτων, μέσω μίας **συναθροιστικής συνάρτησης** (aggregate function)
- Οι βασικές συναθροιστικές συναρτήσεις είναι:
 - αθροίσματος (sum),
 - πλήθους (count),
 - μέσου όρου (avg),
 - μεγίστου (max), και
 - ελαχίστου (min).

Μετρήσεις - Συναθροίσεις

- Εκτός από τις λεπτομερείς πληροφορίες που ορίστηκαν ως μετρήσεις στα fact tables, μπορούν να υπολογιστούν και συναθροίσεις των δεδομένων για καλύτερους χρόνους απόκρισης.
- Για παράδειγμα, αν ο **fact table** είναι:
 - SALES(GeographyCode, ProductCode, TimeCode, BrandCode, **Amount**, **Unit**)
- Τότε μπορούν επιπρόσθετα να υπολογιστούν π.χ.:
 - AVG(Amount) ανά Region, Product, Quarter
 - SUM(Amount) ανά City
 - MAX(Unit) ανά Brand, Month, με Region = Europe

Παράδειγμα Συναθροίσεων



Παράδειγμα Συναθροίσεων σε Πίνακα Γεγονότων

Sales table

RID	City	...	Amount
1	Athens	...	\$100
2	N.Y.	...	\$300
3	Rome	...	\$120
4	Athens	...	\$250
5	Rome	...	\$180
6	Rome	...	\$65
7	N.Y.	...	\$450



City-dimension
sum table

City	Amount
Athens	\$350
N.Y.	\$750
Rome	\$365

δημιουργείται ξεχωριστός
πίνακας/όψη για τα αθροίσματα

Εκτεταμένος πίνακας πωλήσεων:

Ενσωμάτωση των αθροιστικών εγγραφών
στον βασικό fact table με μία επιπλέον στήλη
που εξηγεί το επίπεδο της συνάθροισης

sum



Extended Sales table

RID	City	...	Amount	Level
1	Athens	...	\$100	NULL
2	N.Y.	...	\$300	NULL
3	Rome	...	\$120	NULL
4	Athens	...	\$250	NULL
5	Rome	...	\$180	NULL
6	Rome	...	\$65	NULL
7	N.Y.	...	\$450	NULL
8	Athens	...	\$350	City
9	N.Y.	...	\$750	City
10	Rome	...	\$365	City

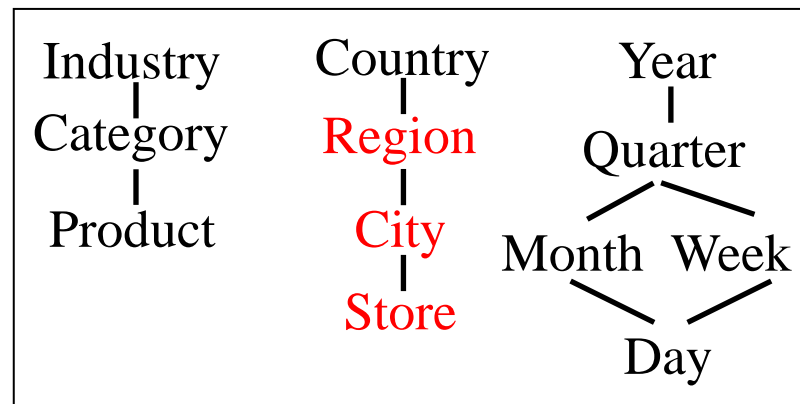


Βασικές Αλγεβρικές Πράξεις

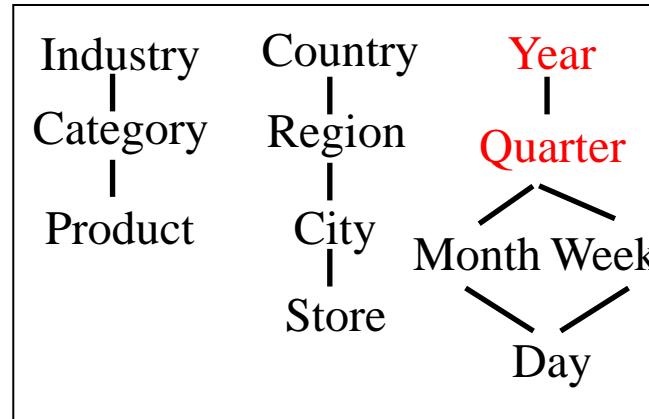
- **Συναθροιστική Άνοδος (Roll-up):** συνάθροιση της πληροφορίας με μετάβαση από χαμηλότερο σε υψηλότερο επίπεδο ιεραρχίας (π.χ. από day σε month)
- **Αναλυτική Κάθοδος (Drill-down):** το αντίστροφο του Roll up (π.χ month σε day)
- **Οριζόντιος Τεμαχισμός (Slice):** επιλογή εγγραφών από μία διάσταση.
- **Κάθετος Τεμαχισμός (Dice):** επιλογή εγγραφών από περισσότερες διαστάσεις.
- **Περιστροφή (Pivot):** αναδιάταξη της προβολής του πολυδιάστατου κύβου (της απεικόνισής του)

Πράξη Συναθροιστικής Ανόδου (Roll-up)

- Η συναθροιστική άνοδος περιλαμβάνει τον υπολογισμό μίας συνολικής τιμής για μία θέση στην ιεραρχία μίας διάστασης δεδομένων.
- Για παράδειγμα, με ένα roll-up, οι πωλήσεις σε επίπεδο τοπικών μαγαζιών (Store) παράγουν τις συνολικές πωλήσεις σε επίπεδο πόλης (City) και αυτές με τη σειρά τους με ένα ακόμα roll-up παράγουν τις πωλήσεις σε επίπεδο περιοχής (Region).



Παράδειγμα Πράξης Roll-up



Sales volume

	Products	Store1	Store2
Q1	Electronics	\$5,2	\$5,6
	Toys	\$1,9	\$1,4
	Clothing	\$2,3	\$2,6
	Cosmetics	\$1,1	\$1,1
Q2	Electronics	\$8,9	\$7,2
	Toys	\$0,75	\$0,4
	Clothing	\$4,6	\$4,6
	Cosmetics	\$1,5	\$0,5

Χρόνος: Επίπεδο Quarter

Sales volume

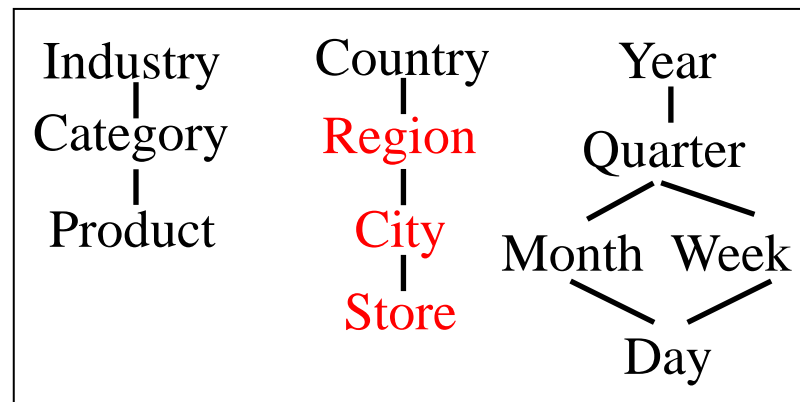
	Products	Store1	Store2
Year 1996	Electronics	\$14,1	\$12,8
	Toys	\$2,65	\$1,8
	Clothing	\$6,9	\$7,2
	Cosmetics	\$2,6	\$1,6

Χρόνος: Επίπεδο Year

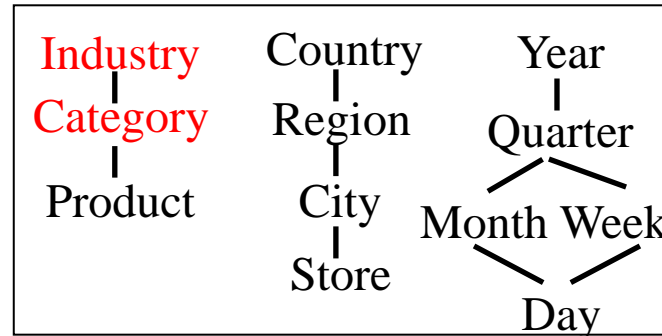
SUM(Sales volumes in dollars)

Πράξη Αναλυτικής Καθόδου (Drill-down)

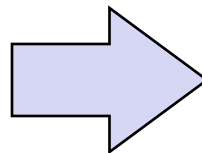
- Ο χρήστης περνά από ένα ανώτερο επίπεδο μίας διάστασης που έχει συγκεντρωτικά δεδομένα σε ένα χαμηλότερο επίπεδο με πιο λεπτομερή δεδομένα. Πρόκειται για την αντίστροφη πράξη του roll-up.
- Για παράδειγμα, κατά το drill-down, ξεκινάμε από τις πωλήσεις ανά περιοχή (Region) και παίρνουμε τις αναλυτικές πωλήσεις ανά πόλη (City) και μετά τις πωλήσεις ανά κατάστημα (Store).



Παράδειγμα Πράξης Drill-down



Sales volume			
Products		Store1	Store2
Q1	Electronics	\$5,2	\$5,6
	Toys	\$1,9	\$1,4
	Clothing	\$2,3	\$2,6
	Cosmetics	\$1,1	\$1,1
Q2	Electronics	\$8,9	\$7,2
	Toys	\$0,75	\$0,4
	Clothing	\$4,6	\$4,6
	Cosmetics	\$1,5	\$0,5



Sales volume			
Electronics		Store1	Store2
Q1	VCR	\$1,4	\$1,4
	Camcorder	\$0,6	\$0,6
	TV	\$2,0	\$2,4
	CD player	\$1,2	\$1,2
Q2	VCR	\$2,4	\$2,4
	Camcorder	\$3,3	\$1,3
	TV	\$2,2	\$2,5
	CD player	\$1,0	\$1,0

Item: Επίπεδο Industry

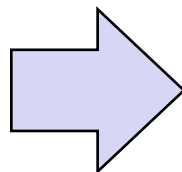
Item: Επίπεδο Category

Πράξη Περιστροφής (Pivot)

- Με την πράξη αυτή γίνεται εναλλαγή των γραμμών και των στηλών του κύβου, όπως αυτός απεικονίζεται.
- Δεν απαιτείται κανένας νέος υπολογισμός.

Παράδειγμα Πράξης Pivot

Sales volume			
	Products	Store1	Store2
Q1	Electronics	\$5,2	\$5,6
	Toys	\$1,9	\$1,4
	Clothing	\$2,3	\$2,6
	Cosmetics	\$1,1	\$1,1
Q2	Electronics	\$8,9	\$7,2
	Toys	\$0,75	\$0,4
	Clothing	\$4,6	\$4,6
	Cosmetics	\$1,5	\$0,5

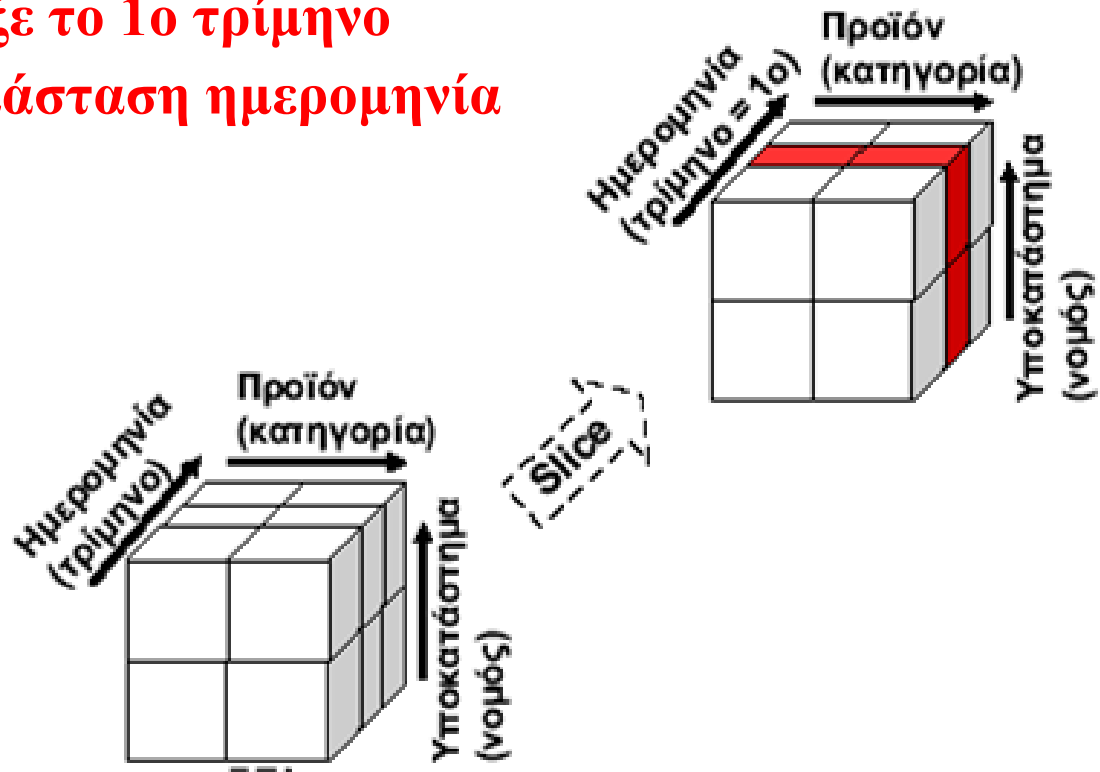


Sales volume			
	Products	Q1	Q2
Store 1	Electronics	\$5,2	\$8,9
	Toys	\$1,9	\$0,75
	Clothing	\$2,3	\$4,6
	Cosmetics	\$1,1	\$1,5
Store 2	Electronics	\$5,6	\$7,2
	Toys	\$1,4	\$0,4
	Clothing	\$2,6	\$4,6
	Cosmetics	\$1,1	\$0,5

Εναλλαγή γραμμών και στηλών

Οριζόντιος Τεμαχισμός (Slice)

- Γίνεται επιλογή συγκεκριμένων τιμών από μία διάσταση (where της SQL),
 - π.χ. επέλεξε το 1ο τρίμηνο από την διάσταση ημερομηνία



Κάθετος Τεμαχισμός (Dice)

- Γίνεται επιλογή συγκεκριμένων τιμών από περισσότερες διαστάσεις (εξάγει έναν υποκύβο)

