



ΠΕΡΙΕΧΟΜΕΝΑ:

1. Συναρτήσεις Σώρευσης
2. Ομαδοποίηση
 1. GROUP BY
 1. WITH ROLLUP
 2. HAVING

Ιωάννης Α.

Σμαραγδένιος Χορηγός Μαθήματος

Ανδριάννα Α.

Χρυσός Χορηγός Μαθήματος

ΜΑΘΗΜΑ 1.3: Ομαδοποίηση με τη GROUP BY

1. Συναρτήσεις Σώρευσης

Οι συναρτήσεις σώρευσης (aggregate functions) είναι:

Συνάρτηση	Επεξήγηση
COUNT(column)	Πλήθος στοιχείων στήλης
SUM(column)	Άθροισμα στήλης
AVG(column)	Μέσος όρος στήλης
MIN(column)	Ελάχιστο στοιχείο στήλης
MAX(column)	Μέγιστο στοιχείο στήλης

- Δεν καταμετρώνται/συνεκτιμώνται στον υπολογισμό τα NULL
- Θέτοντας **DISTINCT** μπροστά από το όνομα της στήλης, ο υπολογισμός αφορά μόνο τις διακριτές τιμές της στήλης.
- Ειδικά για την COUNT: με **COUNT(*)** υπολογίζεται το πλήθος των γραμμών.
- Λέμε ότι με χρήση αυτών, κάνουμε μία ερώτηση περίληψης (summary query)

1.1. Διαφορετικές χρήσης της COUNT:

```
-- example 1.1 (DB: world)
SELECT COUNT(*) FROM country;

SELECT COUNT(IndepYear) FROM country;

SELECT COUNT(DISTINCT IndepYear) FROM country;
```

Παρατηρήσεις:

- Η SUM και η AVG δουλεύουν μόνο για αριθμητικά δεδομένα.
- Η στήλη μπορεί να είναι και υπολογιζόμενη.

1.2. Ελάχιστη και μέγιστη συμβολοσειρά

-- example 1.2

```
SELECT MIN(Name), MAX(Name)
FROM country;
```

1.3. Άθροισμα και Μέσος Όρος (μόνο αριθμητικά δεδομένα)

-- example 1.3

```
SELECT SUM(Population), AVG(Population)
FROM country;
```

Άσκηση 1:

- 1) Μετρήστε πόσες χώρες εμφανίζονται στον πίνακα country
- 2) Μετρήστε πόσες χώρες εμφανίζονται στον πίνακα city

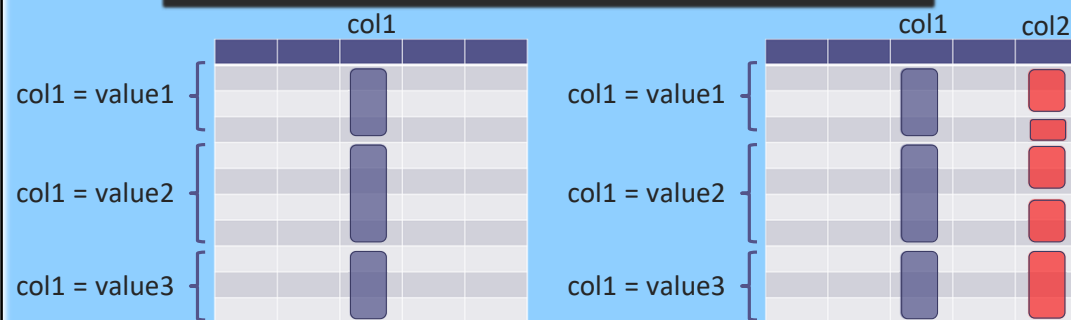
ΜΑΘΗΜΑ 1.3: Ομαδοποίηση με τη GROUP BY

2.1. Ομαδοποίηση (GROUP BY)

Ομαδοποίηση (clause GROUP BY):

- Γράφεται μετά από το WHERE και πριν το ORDER BY
- Χωρίζει τα δεδομένα σε ομάδες με βάση την τιμή της στήλης (ή των στήλων) που ακολουθούν το λεκτικό **GROUP BY**:

GROUP BY column1 [, column2 [, ...]]



- Επιστρέφει μία γραμμή για κάθε ομάδα (χωρίς ταξινόμηση)
- Όταν έχουμε ομαδοποίηση κατά 2 στήλες, κάθε ομάδα προσδιορίζεται και από τις 2 αυτές στήλες



2.1. Query με GROUP BY

```
-- example 2.1 (DB: world)
SELECT *
FROM country
GROUP BY continent;
```

Ομαδοποίηση, με επιλογή aggregate functions:

- Χρησιμοποιώντας aggregate functions, στο αποτέλεσμα ενός group by ερωτήματος, αυτές **εφαρμόζονται ξεχωριστά σε κάθε ομάδα που έχει κατασκευαστεί!**

2.2. Ομαδοποίηση με Aggregate Function επί ενός πεδίου:

-- example 2.2

```
SELECT Continent, COUNT(*) AS Countries
FROM country
GROUP BY Continent;
```

2.3. Ομαδοποίηση με Aggregate Function σε δύο πεδία:

-- example 2.3

```
SELECT Continent, Region, COUNT(*) AS Countries
FROM country
GROUP BY Continent, Region;
```

2.4. (Σχεδόν) Ολοκληρωμένο συντακτικό της SQL (βλ. βίντεο):

-- example 2.4

```
SELECT Continent, Region, COUNT(*) AS Countries
FROM country
WHERE Continent IN ('Asia', 'Europe', 'Africa')
GROUP BY Continent, Region
ORDER BY Continent, Region DESC
LIMIT 10;
```

Σημείωση:

- Η παραπάνω χρήση δεν έχει ιδιαίτερο ενδιαφέρον

WITH ROLLUP:

- Εφαρμόζεται στο τέλος του GROUP BY
- GROUP BY** column1 [, column2 [, ...]] [**WITH ROLLUP**]
- και εμφανίζει μία γραμμή περίληψης στο τέλος κάθε ομάδας, με τα σωρευτικά αποτελέσματα των aggregate functions
 - Έχοντας προσθέσει NULL στα πεδία ομαδοποίησης.



3.1. Query με GROUP BY WITH ROLLUP

-- example 3.1

```
SELECT Continent, Region, COUNT(*) AS Countries
FROM country
GROUP BY Continent, Region WITH ROLLUP;
```

Άσκηση 2: Πίνακας city

Μετρήστε το πλήθος των πόλεων, το άθροισμα του πληθυσμού και το μέσο όρο του πληθυσμού, για κάθε κωδικό πόλης. Εμφανίστε και τις γραμμές περίληψης για κάθε ομάδα με τη ROLLUP.

Άσκηση 3: Πίνακας countrylanguage

Μετρήστε σε πόσες χώρες ομιλείται κάθε γλώσσα. Εμφανίστε τα αποτελέσματα σε φθίνουσα σειρά πλήθους χωρών.

Άσκηση 4: Πίνακας country

Για όλα τα πολίτευματα, ταξινομημένα αλφαβητικά:

- μετρήστε σε πόσες χώρες αυτό εφαρμόζεται
- το μέγιστο GNP χώρας που έχει αυτό το πολίτευμα
- το μέσο όρο του προσδόκιμου ηλικίας των χωρών που έχουν αυτό το πολίτευμα.
- το άθροισμα της γεωγραφικής έκτασης των χωρών που έχουν αυτό το πολίτευμα.

Άσκηση 5: Πίνακας country

Επαναλάβετε την προηγούμενη άσκηση, αυτή τη φορά μόνο για τις ευρωπαϊκές χώρες.

ΜΑΘΗΜΑ 1.3: Ομαδοποίηση με τη GROUP BY

2.2. Ομαδοποίηση (HAVING)

Ομαδοποίηση (clause HAVING):

- Ακολουθεί (και απαιτεί να υπάρχει) το GROUP BY
- Φιλτράρει τις γραμμές - ομάδες που έχουν κατασκευαστεί, με βάση κάποια συνθήκη.

HAVING condition

- όπου η συνθήκη ακολουθεί το ίδιο συντακτικό με τη WHERE αλλά μπορεί να ενσωματώσει και σωρευτικά αποτελέσματα

Σημείωση:

- Συνήθως το φιλτράρισμα γίνεται επί μίας στήλης που έχουμε υπολογίσει κάποιο σωρευτικό αποτέλεσμα.

4.1. Query με GROUP BY και HAVING

-- example 4.1 (DB: world)

```
SELECT Continent, COUNT(*) AS Countries
FROM country
GROUP BY Continent
HAVING Countries > 50;
```

Παρατηρήσεις:

- Η συνθήκη της HAVING μπορεί να αφορά
 - οποιαδήποτε στήλη έχουμε επιλέξει στο SELECT clause.
 - και να ενσωματώνει σωρευτικό αποτέλεσμα επί οποιασδήποτε στήλης του αρχικού πίνακα.

4.2. Query με GROUP BY, HAVING και συνδυασμό σωρευτικών συναρτήσεων.

-- example 2.2

```
SELECT Continent, COUNT(*) AS Countries,
      AVG(Population) as avg_population
FROM country
WHERE IndepYear IS NOT NULL
GROUP BY Continent
HAVING Countries > 10
      AND MIN(SurfaceArea) > 10
ORDER BY Countries DESC
LIMIT 2 ;
```

Σημείωση: Βήματα που ακολουθεί η MySQL για τη διεκπεραίωση ενός ερωτήματος

1. FROM: σε όλες τις εγγραφές του πίνακα
2. WHERE: γίνεται το φιλτράρισμα με βάση τα κριτήρια
3. GROUP BY: Γίνεται η ομαδοποίηση και υπολογίζονται οι σωρευτικές συναρτήσεις του ερωτήματος
4. HAVING: γίνεται το φιλτράρισμα στις ομάδες με βάση τα κριτήρια
5. SELECT: επιλέγονται οι τελικές στήλες
6. ORDER BY: γίνεται η ταξινόμηση
7. LIMIT: περιορίζονται οι εγγραφές που τελικά επιστρέφονται. (προσοχή, η σειρά αυτή δεν είναι “αυστηρή”, βλ. και βίντεο)

ΜΑΘΗΜΑ 1.3: Ομαδοποίηση με τη GROUP BY

(όλες οι ασκήσεις να γίνουν επί της ΒΔ: classicmodels)

Άσκηση 6: Πίνακας customers

Άκολουθεί μία λίστα με τις πόλεις των ΗΠΑ και το πλήθος των πελατών μας σε κάθε πόλη από αυτές:

city	customers
NYC	5
Brickhaven	3
San Francisco	2
New Haven	2
Cambridge	2
Glendale	2
New Bedford	2
Philadelphia	2
Boston	2

Κατασκευάστε ένα ερώτημα που να επιστρέφει το ίδιο αποτέλεσμα.

Ασκήσεις (2/2)

Άσκηση 7: Πίνακας products

Ο πίνακας περιέχει για κάθε προϊόν:

- τον προμηθευτή του (productVendor)
- πόσα αντικείμενα του προϊόντος έχουμε στοκ (quantityInStock)
- τι τιμή έχει ένα τεμάχιο (buyPrice)

Ο manager μας ζητάει ένα report, που να περιέχει εκείνους τους προμηθευτές, των οποίων έχουμε σε στοκ, προϊόντα συνολικής αξίας μεγαλύτερης των 2.500.000. Το report να έχει δύο στήλες ως εξής:

Vendor	total_price
--------	-------------

Η εμφάνιση των ζητούμενων να γίνεται σε φθίνουσα σειρά συνολικής αξίας.

Άσκηση 8: Πίνακας customers

1) Επαναλάβετε την άσκηση 6, αυτήν τη φορά χωρίς τον περιορισμό να αναφερόμαστε σε πελάτες εντός των ΗΠΑ, ομαδοποιώντας τους πελάτες ανάλογα με την πολιτεία (state).

2) Πως μπορείτε να αποφύγετε την ομάδα που έχει στην πολιτεία NULL;