

SYSTEMY BAZ DANYCH SQL



TEMAT 22-04: Grupowanie danych.

Autor dokumentu: Wojciech Galiński poniedziałek, 12 grudnia 2016 r.

351203 Technik informatyk

ŹRÓDŁA WIEDZY: <http://dev.mysql.com/doc/>.

Zagadnienia obowiązkowe

1. **Grupowanie wyników** – to operacji tworzenia grup z danych w tabeli. Oto przykładowe dane (tabela „samochod”):

Marka	Model	Rok
Skoda	Octavia	1996
Skoda	Fabia	1999
Opel	Astra	1991
Opel	Zafira	1997
Seat	Alhambra	1996

Grupowanie wg
kolumny „Marka”
(3 rekordy):

Opel
Astra,
Zafira

Seat
Alhambra

Skoda
Fabia,
Octavia

Grupowanie wg
kolumny „Rok”
(4 rekordy):

1991
Astra

1996
Alhambra,
Octavia

1997
Zafira

1999
Fabia

ZAPAMIĘTAJ: pogrupowane dane są automatycznie sortowane według grupowanych pól.

2. **Grupowanie wyników w języku SQL** – do grupowania wyników służy opcja „**GROUP BY**”. Grupowanie wyników ma sens, gdy użyjemy funkcji agregującej. Oto wzorzec:

SELECT pole1, ... FROM tabela GROUP BY poleM;

PRZYKŁAD:

-- wyświetla liczbę miast we wszystkich krajach bez porządkowania wyników
SELECT kraj, COUNT(*) AS 'Liczba miast' FROM miasto GROUP BY kraj;

3. **Grupowanie według kombinacji pól** - grupować możemy na podstawie kilku pól (kolumn).

Oto wzorzec:

SELECT pole1, ... FROM tabela GROUP BY poleM, ...;

4. **Grupowanie zagnieżdżone** - czyli wielokrotne grupowanie tabeli (grupowanie tabeli z pogrupowanymi danymi).

Oto wzorzec:

**SELECT x2.pole1, ... FROM
(SELECT x1.pole1, ... FROM tabela x1 GROUP BY x1.poleM, ...) x2
GROUP BY x2.poleP, ...);**

ZAPAMIĘTAJ! Tabela wewnętrzna musi mieć zdefiniowany alias.

5. **Porządkowanie grup wyników** – umożliwia to opcja „**ORDER BY**” użyta po wystąpieniu opcji „**GROUP BY**”. Należy pamiętać, że dane są już posortowane, więc opcji tej używamy tylko do zmiany sposobu sortowania.

Oto wzorzec:

SELECT pole1, ... FROM tabela GROUP BY poleM ORDER BY poleP;

PRZYKŁADY:

-- wyświetla liczbę miast w krajach na literę P malejąco według liczby miast
SELECT kraj, COUNT(*) AS 'Liczba miast' FROM miasto
GROUP BY kraj HAVING kraj LIKE "P%" ORDER BY COUNT(*) DESC;

6. **Stosowanie warunków dla grup wyników** – zamiast opcji „WHERE” stosujemy opcję „HAVING”. Tylko wtedy warunki będą dotyczyły grup danych. Oto wybrany wzorzec:

```
SELECT pole1, ... FROM tabela GROUP BY poleK HAVING warunek;
```

PRZYKŁAD:

-- wyświetla liczbę miast tylko w krajach na literę P bez porządkowania wyników

```
SELECT kraj, COUNT(*) AS 'Liczba miast' FROM miasto  
GROUP BY kraj HAVING kraj LIKE 'P%';
```

7. **Słowo kluczowe „WHERE”, a słowo kluczowe „HAVING”** – obydwa słowa kluczowe umożliwiają filtrowanie danych. Różnica pomiędzy nimi polega na tym, że:

- słowo kluczowe „WHERE” używamy przed pogrupowaniem danych;
- słowo kluczowe „HAVING” stosujemy po pogrupowaniu danych.

Oto wzorzec:

```
SELECT pole1, ... FROM tabela WHERE warunek_dla_danych_przed_grupowaniem  
GROUP BY poleK HAVING warunek_dla_danych_po_grupowaniu;
```

PRZYKŁADY:

-- Wyświetla te państwa na literę „P”, w których znajdują się co najmniej 2 miasta na literę „W”

-- (wyświetlana jest liczba miast w tych państwach)

```
SELECT kraj, COUNT(*) AS 'Liczba miast' FROM miasto  
WHERE kraj LIKE 'P%' AND miasto LIKE 'W%'  
GROUP BY kraj HAVING COUNT(*) > 1 ORDER BY COUNT(*) DESC;
```

-- Wyświetla liczbę miast na literę S w krajach na literę C według liczby miast w porządku malejącym

```
SELECT kraj, COUNT(*) AS 'Liczba miast' FROM miasto  
WHERE miasto LIKE 'S%' GROUP BY kraj HAVING kraj LIKE "C%" ORDER BY COUNT(*) DESC;
```

8. **Modyfikator „WITH ROLLUP”** – po zastosowaniu tego słowa kluczowego dodawane są dane sumaryczne dotyczące wszystkich rekordów w tabeli (gdy grupujemy według jednej kolumny) albo dane sumarycznych grupowanych kolumn (gdy grupujemy według wielu kolumn).

Oto wzorzec:

```
SELECT pole1, ... FROM tabela GROUP BY poleK WITH ROLLUP;
```

PRZYKŁADY:

-- Na końcu tego zestawienia pojawi się dodatkowy wiersz z sumą wszystkich dochodów

```
SELECT IFNULL(rok, "Razem") as rok, SUM(dochod) FROM dochody GROUP BY rok WITH ROLLUP;
```

-- Wyświetla liczbę miast w poszczególnych krajach pogrupowanych według kontynentów

```
SELECT kontynent, kraj, COUNT(miasto) FROM miasta_swiate  
GROUP BY kontynent, kraj WITH ROLLUP;
```

Pytania kontrolne

1. Ile byłoby grup i co by zawierały, gdyby tabelę z punktu 1 pogrupować według kolumny „Model”?
2. Na czym polega grupowanie wyników? Gdzie w praktyce można to wykorzystać?
3. W jaki sposób można grupować dane według więcej niż 1 kolumny?
4. Na czym polega zagnieżdżanie grupowania danych? Czy można zagnieżdżać grupowanie danych w języku SQL?
5. Jak stosować warunki na grupach danych?
6. Jakie są różnice pomiędzy klauzulami: „WHERE” i „HAVING”?
7. Do czego służy modyfikator „WITH ROLLUP”?

Zadania

1. Wyświetl średnią wieku uczniów w poszczególnych klasach.
2. Wyświetl 3 klasy z najwyższą średnią ocen.
3. Wyświetl imiona uczniów bez powtórzeń w poszczególnych klasach.
4. Wyświetl imiona i nazwiska uczniów, którzy mają średnią ocen wyższą niż 4.00.
5. Wyświetl średnie ocen poszczególnych uczniów z wybranej klasy.
6. Wyświetl imiona i nazwiska uczniów, którzy mają co najmniej 3 oceny pozytywne (wyświetl także ilość ocen pozytywnych).