

Notatka z zajęć plus praca domowa

1. Wyszukaj wszystkie książki, które mają w tytule słowo „the”.
2. Wyświetl dwie kolumny jako jedną, tytuł i rating.
3. Wyświetl książki, w których nazwisko autora kończy się na ‘en’

FUNKCJE AGREGUJĄCE

ZLICZANIE (count) i distinct

funkcja COUNT pozwala na zliczanie przypadków:

```
SELECT COUNT(*) FROM książki;
```

policzy liczbę wszystkich przypadków w tabeli (wierszy, liczba kolumn jest nieistotna), analogicznie:

```
SELECT COUNT(author) FROM książki;
```

wskazanie kolumny nie zmieni wyniku (liczymy przypadki rekordów - każdy przypadek odnosi się do rekordu)

Aby policzyć wartość unikalną, należy dodać słówko DISTINCT:

```
SELECT COUNT(DISTINCT author) FROM książki;
```

DISTINCT author jest wewnątrz, jest ciałem, na którym zliczamy przypadki. Gdyby konstrukcja była na odwrót, czyli najpierw distinct a potem count to wynik byłby dokładnie 100, czyli taki jak liczba rekordów.

4. Policz ile książek zostało opublikowanych w 2012 roku.
5. Policz ile książek w naszej bazie zostało napisanych przez autora o imieniu John.

GRUPOWANIE

Jeżeli chcemy policzyć coś przypadkami, np. liczbę książek w bazie każdego autora, należy wykorzystać mechanizm grupowania, który pozwala kompresować wynik zapytania do zmiennej porządkującej. Następnie wykonuje zapytanie select wewnątrz grup stworzonych wcześniej:

```
1. SELECT
2.     author,
3.     COUNT(*) AS 'ile książek'
4. FROM książki
5. GROUP BY author;
```

Funkcji GROUP BY zbiera wszystko do jednego wiersza i wywołuje na nim instrukcje zawarte w SELECT. Powyższe wywołanie funkcji COUNT działa w obszarze tego właśnie wiersza, a nie wszystkich wierszy w tabeli. Jeżeli nie zastosujemy żadnej funkcji, tylko wypiszemy bezpośrednio rekordy z tabeli z zastosowaniem GROUP BY, dla każdego przypadku zostanie wyświetlony tylko pierwszy rekord jaki dla niego był znaleziony w tabeli.

6. Wyświetl ile dany autor ma książek w bazie danych

7. Policz liczbę książek wydanych w różnych latach, wynik podaj w kolejności malejącej według liczby książek.

STATYSTYKI OPISOWE – ZESTAWIENIA

min, max, suma, avg, round

Aby zwrócić najniższą ocenę książki w bazie używamy funkcji MIN:

```
SELECT MIN(rating) FROM ksiazki;
```

analogicznie można zapisać:

```
SELECT rating FROM ksiazki ORDER BY rating LIMIT 1;
```

Najwyższą wartość zwraca funkcja MAX:

```
SELECT MAX(rating) FROM ksiazki;
```

Poniższe zapytanie nie zadziała poprawnie:

```
SELECT author, title, MIN(rating) FROM ksiazki;
```

dlatego, że zgodnie z kolejnością podaną w zapytaniu, wykonana została następująca czynność: wypisz autorów i tytuły książek po kolei z bazy, wypisz najniższą ocenę jaka jest w całej bazie (3 kolumna) -> najniższa ocena jest tylko jedna, nie została więc wypisana cała reszta przypadków, a jedynie pierwsza pozycja w całej tabeli i najniższa ocena w całej tabeli, które nie są sobie tożsame.

Aby otrzymać prawidłowy wynik tego zapytania, należy go zmodyfikować:

```
SELECT author, title, rating FROM ksiazki WHERE rating = (SELECT MIN(rating) FROM ksiazki);
```

W pierwszym członie zapytania poprosiliśmy bazę o podanie autora, tytułu i oceny książki, następnie podaliśmy warunek, że ma to być książka o najniższej ocenie. Nie możemy tak po prostu użyć „MIN(rating)”, bo nasza baza nie będzie wiedzieć skąd go wziąć, dlatego musimy stworzyć drugie zapytanie typu SELECT, które zwraca nam wartość minimalną, która dalej używana jest do pierwszego zapytania.

Aby zsumować liczby używamy funkcji SUM:

```
SELECT SUM(ratings_no) FROM ksiazki;
```

Analogicznie liczymy średnią (z angielskiego average, skrót: AVG):

```
SELECT AVG(ratings_no) FROM ksiazki;
```

Wynik można zaokrąglić, np do 2 miejsc po przecinku:

```
SELECT ROUND(4.21351357, 2);
```

zaokrąglanie działa tak jak uczyli nas tego na matematyce, czyli od 5 zaokrąglamy do góry:

```
SELECT ROUND(4.25, 1);
```

8. Podaj rok pierwszej publikacji każdego autora w naszej bazie.

9. Podaj listę 10 najczęściej komentowanych autorów.

W operacjach logicznych możemy tworzyć zapytania, które zwracają jako wartość wynik 0 (fałsz) lub 1 (prawda). W tym celu możemy używać operatorów porównania:

1. SELECT 5 = 5; -- równości
2. SELECT 5 != 5; -- różności
3. SELECT 8 > 5; -- większości
4. SELECT 5 < 6; -- mniejszości
5. SELECT 5 >= 5; -- większości lub równości
6. SELECT 5 <= 6; -- mniejszości lub równości

i stosować je w zapytaniach

```
SELECT title, publication_year FROM ksiazki WHERE publication_year != 2015;
```

Warunki można łączyć w sekwencje, poprzez łączniki logiczne:

1. AND && (koniunkcja)
2. OR || (alternatywa)

AND zwraca prawdę, gdy wszystkie wyrażenia są prawdziwe

OR natomiast gdy choć jeden warunek jest prawdziwy

Przykładowo, warunek który zwróci książki wydane po 2010 roku i jednocześnie których ocena jest większa niż 4.0:

1. SELECT
2. title,
3. publication_year,
4. rating
5. FROM ksiazki

```
6. WHERE
7.     publication_year > 2010 AND
8.     rating > 4.0;
```

lub warunek, który zwróci książki wydane po 2010 roku lub z oceną większą niż 4.0:

```
1. SELECT
2.     title,
3.     publication_year,
4.     rating
5. FROM ksiazki
6. WHERE
7.     publication_year > 2010 OR
8.     rating > 4.0;
```

Jeżeli zawieramy warunek pomiędzy dwiema wartościami, można użyć operatora BETWEEN:

```
1. SELECT
2.     title,
3.     publication_year
4. FROM ksiazki
5. WHERE
6.     publication_year BETWEEN 2000 AND 2010;
```

Zarówno BETWEEN jak i LIKE można zaprzeczać, aby uzyskać wszystko poza zasięgiem zwracanym przez funkcje, poprzez operator NOT:

```
1. SELECT
2.     author,
3.     title
4. FROM ksiazki
5. WHERE
6.     author NOT LIKE '%John%';
```

LOGIKA w sql

or, and, between

6.1. Ćwiczenia #29 - #35 **+

10. Znajdź tytuły książek które zostały opublikowane nie wcześniej niż w 2005 roku oraz nie później niż w 2010.

11. Oblicz średni rating książek wydanych pomiędzy 1990 a 2005, dla każdego roku wydawniczego, wyniki podaj chronologicznie.

12. Policz ile każdy z autorów dostał po jednej gwiazdce dla wszystkich swoich książek, ale uwzględnij tylko te pozycje dla których oddano co najmniej milion głosów w sumie, lub umieszczono co najmniej 50 000 komentarzy.

13. Jaki procent wśród książek o ocenie co najmniej 4, stanowią oceny pięciogwiazdkowe?

14. Jaka jest średnia ocena dla autorów których nazwiska są krótsze niż 6 albo dłuższe niż 10 znaków?

*15. Dla każdego autora, dla każdego roku w którym opublikował książkę (lub książki) policz sumę komentarzy do tych pozycji, uwzględnij tylko pozycje, które miały co najmniej 600 tys. głosów, listę podaj posortowaną od Z do A.

[Podpowiedź: należy zastosować podwójne grupowanie - nie omawialiśmy tego tematu ale rozwiązanie jest prostsze niż Ci się wydaje – tylko pamiętaj o kolejności!]

*16. Podaj listę autorów i tytułów książek wydanych w latach parzystych.