

Bazy danych i SQL

OMÓWIENIE

Celem tego zadania jest przygotowanie zapytań SQL, pobierających wybrane informacje z bazy danych. W ramach zadania dostarczony zostanie jej schemat, przykładowe dane oraz skrypty pomocne przy pracy. Dostarczony schemat bazy danych jest uproszczony i przygotowany specjalnie na potrzeby konkursu.

Podstawowe założenia:

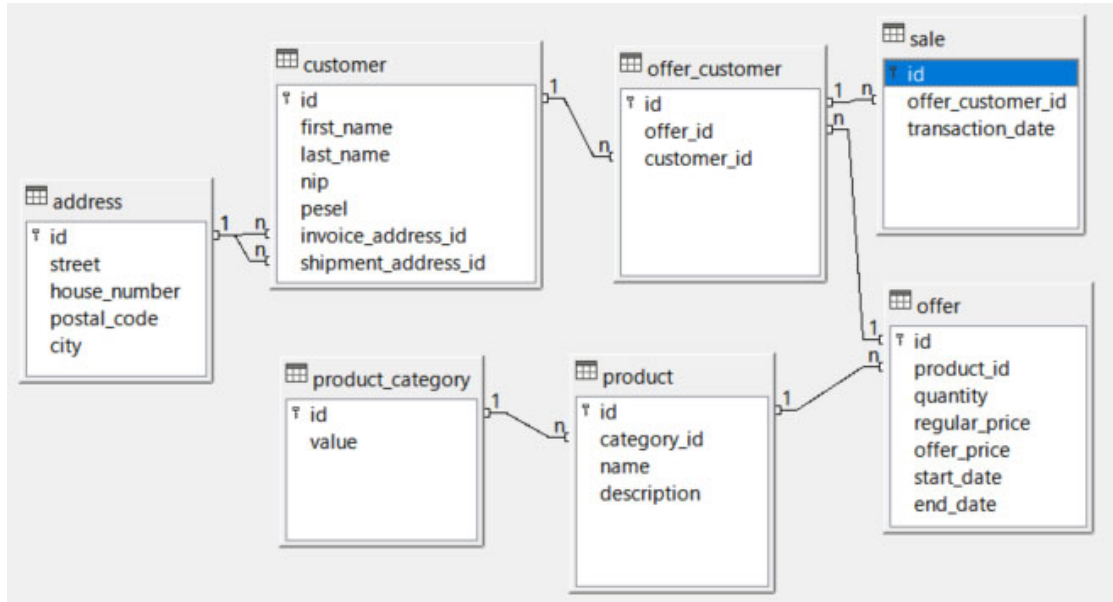
- rozwiązania zadań będą testowane w oparciu o bazę danych PostgreSQL 12.5 (lub nowszą z linii 12.x)
- struktura bazy danych, użytkownicy oraz definicja uprawnień zostaną dostarczone w ramach zadania i nie mogą być zmieniane (oznacza to że dodawanie nowych tabel lub zmienianie uprawnień jest zabronione)
- dozwolone jest nadawanie indeksów poprzez dodanie odpowiednich komend do pliku *index.sql* - może to znacząco przyspieszyć przeszukiwanie bazy danych
- opracowane zapytania nie powinny zwracać więcej danych niż jest to wymagane w konkretnym zadaniu, a ich czas wykonania powinien być jak najkrótszy
- zapytania zapisujemy w osobnych plikach o nazwach odpowiadających numerowi zadania. Pliki zapisujemy z rozszerzeniem *.sql*.

Zestaw danych, który załączamy stanowi jedynie próbkę, prezentującą przykładowe dane. Ma to Wam pomóc w zrozumieniu struktury bazy oraz ułatwić rozpoczęcie pracy. Ocena rozwiązań zostanie przeprowadzona na innym, znacznie większym zbiorze danych. Na potrzeby własnych testów możecie jednak przygotować swoje zestawy danych, pamiętając jedynie, że nie stanowią one części rozwiązania tego zadania.

Zwróćcie szczególną uwagę na dwa pliki które otrzymujecie w ramach zadania :

- *readme.md* zawiera opis poszczególnych plików oraz kilka dodatkowych wskazówek
- *run.bat* przygotowuje schemat bazy danych wraz danymi oraz odpala rozwiązanie konkretnego zadania w sposób analogiczny jak to się będzie odbywać w naszych testach

Schemat bazy danych



CO OCENIAMY

1. Poprawność zwracanych wyników
2. Szybkość wykonania każdego zapytania

Jeżeli rezultat zwracany przez zapytanie jest prawidłowy, przyznawane jest 75% punktów, przeznaczonych na to zadanie. Tylko w przypadku poprawnych wyników przechodzimy do oceny szybkości.

Każde zapytanie dostarczone przez drużynę zostanie uruchomione wielokrotnie, a następnie obliczymy średni czas jego wykonania. Ma to na celu zminimalizowanie wpływu systemu operacyjnego oraz ewentualnego obciążenia maszyny na uzyskane rezultaty.

Dodatkowo tak uzyskane przez każdą drużynę rezultaty, zostaną zebrane w celu obliczenia średniego czasu wykonania zadania wśród wszystkich uczestników. Drużyny, których czas wykonania zapytania jest lepszy lub równy medianie czasów uzyskanych przez wszystkie drużyny, dostaną kolejne 25% punktów jakie jest w puli danego zadania. W przypadku ekstremalnie małych różnic pomiędzy najlepszym i najgorszym czasem wykonania zapytań, wszystkie drużyny otrzymają 25% punktów z puli.

Suma punktów za wszystkie zadania, będzie stanowić ostateczny wynik.

ZADANIA

1. Ilu jest wszystkich klientów, których nazwisko zaczyna się od litery a (zarówno mała i duża litera)?
2. Który klient dokonał najmniej transakcji zakupu (a dokonał przynajmniej jednej)? W odpowiedzi na zapytanie zwróć ID klienta. W przypadku większej ilości klientów spełniających kryteria, posortuj nazwiska w kolejności alfabetycznej i zwróć ID pierwszego z nich.
3. Który/-zy klient/-ci dokonał/-li najwięcej transakcji zakupu? W odpowiedzi na zapytanie zwróć ID klienta/-ów posortowanych w odwrotnej kolejności alfabetycznej względem nazwiska.
4. Który/-zy klient/-ci kupił/-li produkty z kategorii *hardware* oraz z kategorii: *licencja_rok* lub *licencja_3*? W odpowiedzi na zapytanie zwróć ID klienta/-ów oraz jego imię i nazwisko.
5. Który/-zy klient/-ci kupił/-li produkt typu *hardware* i pochodził z Krakowa? W odpowiedzi na zapytanie zwróć ID klienta. Uwzględnij adres klienta znajdujący się na fakturze.
6. Ilu klientów z Warszawy lub Wrocławia kupiło produkt typu *hardware* lub *licencja_rok* przed 2018 rokiem? W odpowiedzi na zapytanie zwróć ID klientów
7. Ile transakcji sprzedaży zostało zawartych w dniu zakończenia oferty?
8. Jaka jest średnia wartość wszystkich transakcji sprzedaży? Wynik podaj jako liczbę zaokrągloną do trzeciego miejsca po przecinku (przykład: 54.123)
9. Jaka jest średnia wartość transakcji sprzedaży z udziałem klientów z Krakowa? Uwzględnij adres klienta znajdujący się na fakturze.
10. Jaka jest średnia wartość transakcji sprzedaży z udziałem klientów z Krakowa, którzy dokonali zakupu w cenie co najmniej 10% poniżej ceny regularnej? Wynik podaj jako liczbę zaokrągloną do drugiego miejsca po przecinku (przykład: 54.12)
11. Ilu klientów z różnym adresem faktury i wysyłki dokonali co najmniej 3 zakupów w 2019 roku?
12. Który produkt sprzedawał się najczęściej (największa ilość transakcji sprzedaży z jego udziałem)? W odpowiedzi na zapytanie zwróć ID produktu. W przypadku, gdy kilka produktów spełnia kryteria zapytania zwróć produkt z najmniejszym ID.
13. Który produkt sprzedał się w największej ilości (łączna ilość sprzedanych sztuk produktu)? W odpowiedzi na zapytanie zwróć ID produktu oraz ilość sprzedanych sztuk towaru. W przypadku, gdy kilka produktów spełnia kryteria zapytania zwróć informacje o produkcie z największym ID

14. Które produkty sprzedawały się zawsze pojedynczo (nigdy 2 sztuki lub więcej naraz)?
W odpowiedzi na zapytanie zwróć ID produktów spełniających kryteria i których ID mieści się między 5 a 50.
15. Które produkty sprzedały się przynajmniej jeden raz w cenie regularnej w dniu rozpoczęcia oferty? W odpowiedzi na zapytanie zwróć ID produktów, cenę oraz datę sprzedaży. W przypadku, gdy kilka produktów spełnia kryteria zapytania zwróć ID wszystkich wylistowanych produktów uporządkowanych po cenie od najwyższej.
16. Ilu klientów ma adres wysyłki taki sam jak przynajmniej jeden inny klient? O takim samym adresie wysyłki mówimy, gdy adres zawiera tę samą ulicę, numer lokalu oraz nazwę miejscowości.
17. Jaki procent klientów ma podany jedynie numer pesel (bez podanego numeru NIP)?
Wynik podaj jako liczbę zaokrągloną do pierwszego miejsca po przecinku (przykład: 54.1)
18. Podaj listę klientów, którzy otrzymali ofertę z ceną co najmniej 20% poniżej ceny regularnej. W odpowiedzi na zapytanie zwróć nazwisko oraz numer NIP a w przypadku jego braku numer pesel. Wyniki zapytania uporządkuj według numerów NIP/PESEL rosnąco.
19. Podaj listę klientów, którzy otrzymali co najmniej o 20 więcej ofert niż klient który otrzymał najmniej ofert. Nie uwzględniamy tutaj klientów którzy nie otrzymali żadnej oferty. W odpowiedzi na zapytanie zwróć numer NIP a w przypadku jego braku numer pesel. Wyniki zapytania uporządkuj w kolejności malejącej.
20. W ramach jednego zapytania sprawdź ilu jest klientów którzy wydali mniej oraz ilu jest klientów którzy wydali więcej na zakupy niż klient o numerze NIP 1000000000.
W odpowiedzi na zapytanie zwróć dwie liczby (w dwóch kolumnach).