

Hurtownie danych – sprawozdanie 5

Baza danych: **AdventureWorks2016**

Jeżeli nie potrafisz rozwiązać zadania, to próbuj podać chociaż częściowe rozwiązanie lub uzasadnienie przyczyny braku rozwiązania.

Microsoft SQL Server Management Studio oraz Integration Services

ZADANIE 1

Używając Visual Studio utworzyć projekt typu Analysis Services Project (Menu File -> New Project)

- a) Dodać źródło danych (Solution Explorer → Data Sources → New Data Source), które będzie wskazywało na bazę danych, która przechowuje tabele utworzone podczas rozwiązywania listy 4. Przeanalizować opcje związane z bezpieczeństwem dostępu do danych.
- b) Utworzyć nowy widok źródła danych (Solution Explorer → Data Source Views → New Data Source View). Dodać tabele utworzone podczas rozwiązywania listy 4.
- c) Utworzyć nową kostkę za pomocą asystenta (Solution Explorer → Cubes → New Cube):
 - Wybrać utworzenie kostki na podstawie istniejących tabel (Use existing tables)
 - Wybrać, utworzony w poprzednim punkcie, widok źródła danych
 - Jako tabelę faktów (Measure group tables) wybrać FACT_SALES
 - Na stronie dotyczącej miar wybrać OrderQty, UnitPriceDiscount, Line Total. Zastanowić się nad użytecznością wybranych miar. Dlaczego nie wszystkie atrybuty tabeli FACT_SALES mogą być użyte jako miary?
 - Na stronie dotyczącej wymiarów wybrać wszystkie tabele z przedrostkiem Dim.

ZADANIE 2

Po utworzeniu kostki dokonać edycji wymiarów (Solution Explorer → Dimensions lub zakładka Cube Structure → Dimensions) i wdrożyć projekt (Deploy)

1. Dla każdego z wymiarów zdefiniować potrzebne atrybuty. Przykładowo wymiar produkt powinien zawierać: Nazwę, Cenę, Kolor, Podkategorię i Kategorię. W przypadku tabeli Produkt należy zmienić jej definicję i sposób ładowania, tak aby zawierała ona zarówno kategorię jak i podkategorię (jeśli nie zrobiło się tego już podczas pracy z kreatorem dodawania nowej kostki).
2. Zmodyfikować definicję wymiarów tak, aby:
 - a. w wymiarach CUSTOMER i SALESPERSON nie można było korzystać z atrybutów FirstName oraz LastName. W zamian dodać atrybut Names
 - b. w wymiarze SALESPERSON pojawiła się hierarchia Group – CountryRegionCode – Names
 - c. w wymiarze CUSTOMER pojawiła się hierarchia Group – CountryRegionCode – Names

Sprawozdanie bez wniosków końcowych nie będzie sprawdzane i tym samym ocena jest negatywna!

- d. w wymiarze PRODUCT pojawiła się hierarchia CategoryName – SubCategoryName – Name
3. Dla każdego atrybutu kluczowego wymiaru, którego wartościami są liczby całkowite, zmodyfikować właściwości (Properties). Zmodyfikować parametr NameColumn, tak aby nazwy kolejnych elementów wymiaru nie były liczbami. (Przykładowo dla wymiaru dotyczącego Produktu można wykorzystać atrybut Name).
4. Sprawdzić ustawienia projektu (Solution Explorer → Nazwa projektu → Properties).
5. Sprawdzić ustawienia konfiguracji wdrożenia (Configuration properties → Deployment → Target) projektu – Server: localhost, Database: PierwszyProjektKostek
6. Wdrożyć projekt (Menu Build → Deploy). Przetworzyć kostkę (Przycisk Process).

ZADANIE 3

Przy użyciu zakładki Browser:

1. Sprawdzić, czy dane zapisane w kostce zgadzają się z danymi zapisanymi w tabelach, przeciągając za pomocą myszy:
 - a. atrybuty wymiarów w region wierszy
 - b. miary w część centralną widoku
2. Przetestować możliwości przeglądarki (Browser) – operator wyboru danych (Operator), wyrażenia filtrujące dane (Filter Expression) itp.
3. Przygotować przykładowe tabele i wykresy przestawne (Excel) oraz zinterpretować uzyskane wyniki (wnioski!)

ZADANIE 4

Korzystając z bazy danych AdventureWorksDW:

1. Utworzyć nowe źródło danych
2. Utworzyć nowy widok źródła danych
3. Utworzyć nową kostkę wybierając jako tabelę faktów FactInternetSales, wymiary według własnego uznania – uzasadnić wybór
4. Z badać, jak zmienia się czas przetwarzania kostki w zależności od liczby zdefiniowanych wymiarów.
5. Poćwiczyć edycję wymiarów i korzystanie z przeglądarki kostki danych (Browser) oraz przygotować ciekawe zestawienia – tabele i wykresy przestawne w MS Excel.

ZADANIE 5

Podzielić zawartość kostki z zadania 4 na partycje (zakładka Partitions). Każda partycja powinna odzwierciedlać jeden rok. Istnieją dwa podstawowe sposoby podziału partycjonowania kostek:

- dane do zasilania poszczególnych partycji znajdują się w osobnych tabelach
- dane do zasilania poszczególnych partycji znajdują się w tej samej tabeli, zaś każda z partycji ma przypisanie zapytanie SQL, którego wynik służy do jej zasilenia.

Proszę przygotować partycje na dwa sposoby i znaleźć uzasadnienie dla każdej opcji.