**Hurtownie Danych - laboratorium**

**Lista 6**

***Hierarchie, partycje, prosta analiza danych***

*Uwaga: do realizacji zadań niezbędne jest przygotowanie listy 4.*

**Zad. 1. Modyfikacja wymiarów i tabeli faktów**

Bazując na kostce utworzonej przy realizacji listy 4, należy:

1. zmodyfikować definicję wymiarów tak, aby:

* w wymiarach CUSTOMER i SALESPERSON nie można było korzystać z atrybutów FirstName oraz LastName. W zamian dodać atrybut Names
* w wymiarze SALESPERSON pojawiła się hierarchia Group – CountryRegionCode – Names
* w wymiarze CUSTOMER pojawiła się hierarchia Group – CountryRegionCode – Names
* w wymiarze PRODUCT pojawiła się hierarchia CategoryName – SubCategoryName – Name
* w wymiarze TIME pojawiła się hierarchia Rok – Kwartał – Miesiąc – Dzień miesiąca



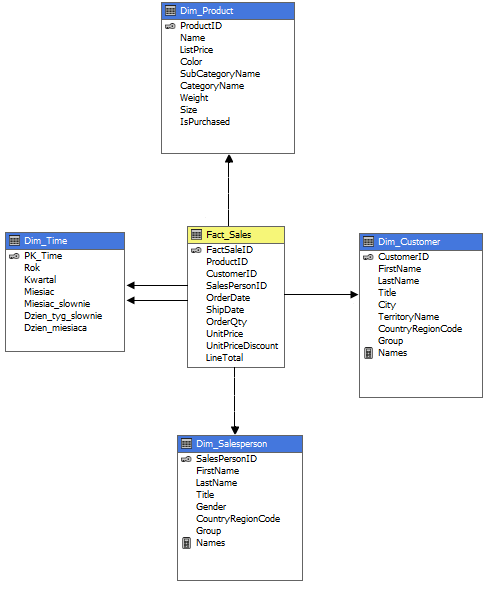
*Przykładowe rozwiązanie.*

1. 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).



*Przykładowe rozwiązanie.*

1. utworzyć nowe miary, które będą odzwierciedlać:
   * Liczbę różnych klientów (aggregatedFunction: distinct count)
   * Liczbę różnych produktów
   * Maksymalną wartość rabatu (aggregatedFunction: max)
   * Maksymalną liczbę zamówionych produktów
   * Liczbę różnych sprzedawców realizujących zamówienia
2. wdrożyć i przeprocesować kostkę.



*Przykładowe rozwiązanie – schemat wynikowej kostki.*

**Zad. 2. Przegląd danych i tworzenie zestawień**

Przy użyciu zakładki Browser:

1. Sprawdzić, czy dane zapisane w kostce zgadzają się z danymi zapisanymi w tabelach, przeciągając za pomocą myszy:
   * atrybuty wymiarów w region wierszy
   * 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!)

**Zad. 3. Partycje**

Podzielić zawartość kostki 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.

**Zad. 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. Zbadać, 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.

**Rozwiązania:**

**Wnioski:**

***Uwaga!***

* Sprawozdanie, bez wniosków podsumowujących aspekt zagadnień analizowanych na zajęciach laboratoryjnych i zawartych w sprawozdaniu, jest automatycznie oceniane negatywnie!