

Hurtownie danych – laboratorium

Lista 6

Hierarchie, miary, partycje, MDX

Zad. 1 Modyfikacja wymiarów i tabeli faktów

W oparciu o kostkę przygotowaną podczas realizacji listy 4 należy:

- a) Zmodyfikować wymiary CUSTOMER i SALESPERSON tak, aby nie można było korzystać z atrybutów FirstName oraz LastName. W zamian dodać atrybut Names
- b) Utworzyć hierarchie:
 - W wymiarze SALESPERSON: Group – CountryRegionCode – Names
 - W wymiarze CUSTOMER: Group – CountryRegionCode – Names
 - W wymiarze PRODUCT: CategoryName – SubCategoryName – Name
 - W wymiarze TIME: Rok – Kwartał – Miesiąc – Dzień miesiąca
- c) 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).
- d) Utworzyć 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

Zad. 2 Definiowanie miar kalkulowanych

W zakładce Calculations dodać dwie miary kalkulowane (ang. *calculated members*):

- średnią liczbę zamówionych towarów na zamówienie
- średnią ważoną liczbę towarów na zamówienie. Jako wagę należy wybrać cenę danego produktu.

Wskazówka: w celu utworzenia wyżej wymienionej średniej ważonej można posłużyć się nową kolumną zdefiniowaną w widoku źródła danych (lub w tabeli). Kolumna ta powinna definiować miarę pomocniczą, która pozwoli uzyskać fragment wyrażenia odpowiadającego średniej ważonej.

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 Hurtownia AdventureWorksDW

Korzystając z bazy danych AdventureWorksDW:

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

TEORIA

Aby uruchomić zapytania MDX należy uruchomić Microsoft SQL Server Management Studio połączyć się z Analysis Services i wybrać opcję *New query...*

Uproszczona forma zapytania MDX, która będzie używana podczas rozwiązywania tej listy, ma postać:

```
select {Sc} on columns,{Sr} on rows from NazwaKostki  
where (Tf)
```

gdzie:

- Sc* – to zbiór elementów wymiaru (dimension members), które mają znaleźć się w kolumnach np.
`Sc={ [CUSTOMER].[Hierarchy].[Group].&[Europe],
[CUSTOMER].[Hierarchy].[Group].&[Europe].&[DE].&[Adams Luis]}`
W powyższym przykładzie zbiór ma dwa elementy Europa i Adam Luis
- Sr* - to zbiór elementów wymiaru (dimension members), które mają znaleźć się w wierszach
- NazwaKostki* – nazwa kostki, do której kierowane jest zapytanie
- Tf* - krotka, która będzie filtrowała dane znajdujące się w wierszach i kolumnach (odpowiednik wartości znajdujących się w nagłówku widoku sporządzanego w Management Studio opcją Browse). Krotka w MDX zawiera po jednej wartości każdego atrybutu zdefiniowanego dla każdego wymiaru. Jeśli atrybut jakiegoś wymiaru nie jest wymieniony, brana jest jego wartość domyślna (DefaultMember). Najczęściej jest to wartość sumaryczna dla wszystkich wartości tego atrybutu (np. jeśli nie jest wymieniona płeć handlowca krotka odnosi się do wartości sumarycznej dla mężczyzn, kobiet i płci niezidentyfikowanej).

Przykład poprawnie sformułowanej krotki:

```
Tf=([PRODUCT].[Name].&[Cup-Shaped Race],  
[PRODUCT].[Color].&[Black],[SHIPDATE].[Ship Year].&[2002])
```

Przykład niepoprawnie sformułowanej krotki:

```
([PRODUCT].[Color].&[Blue],  
[PRODUCT].[Color].&[Black],[SHIPDATE].[Ship Year].&[2002])
```

Uwaga! Klauzula where jest opcjonalna.

Więcej informacji na temat MDX można znaleźć w Books Online.

Zad. 5 MDX

Napisać zapytania MDX:

- a) Wyświetl liczbę różnych klientów z każdego regionu.
- b) Przedstaw liczbę różnych klientów z każdego regionu w latach 2012 i 2013 (wskazówka: należy przenieść miarę “liczba różnych klientów” do klauzuli *where*).
- c) Do zapytania z zadania 4 dodać miarę Order Qty. Zarówno miary jak i lata mają znaleźć się w kolumnach. Do umieszczenia dwóch wymiarów w kolumnach użyć funkcji *crossjoin*.
- d) Wyświetl wszystkie kategorie i podkategorie produktów (funkcja *children*), które zostały zakupione przez klientów z poszczególnych regionów oraz liczbę zakupionych produktów
- e) Zmodyfikować zapytanie 6, aby wyświetliło tylko te podkategorie, w których zakupiono co najmniej 10 różnych produktów (do filtrowania użyć funkcji *filter*). Posortować podkategorie wg liczby zakupionych produktów.
- f) Wyświetl i porównaj średnie kroczące wartości transakcji w kolejnych miesiącach w latach 2012 i 2013 (funkcja *LastPeriod* i *Avg*).
- g) Znaleźć dzień w roku 2013, w którym sprzedano największą liczbę produktów (użyć funkcji *descendants* i *head*). Do wyświetlenia numeru miesiąca użyć następującej definicji typu *calculated member*:

with member [Measures].[NazwaMiesiaca]

as '[ORDERDATE].[OrderMonth].CurrentMember.Name'

Powyższa definicja powinna się znaleźć przed słowem kluczowym *select*. Umożliwia ona odwołanie się do miary, która wyświetli nazwę miesiąca, w ten sam sposób, w jaki wyświetla się inne miary.

Zad. 6 Reporting Services

W środowisku Business Intelligence Studio utworzyć nowy projekt Reporting Services, a następnie:

- Utworzyć nowe współdzielone źródło danych (w sekcji *Connection Properties* jako źródło danych wybrać *MS SQL Server Analysis Services*)
- Wygenerować prosty raport z użyciem kreatora raportów (*Solution Explorer->Reports->Add New Report*)
- Zweryfikować poprawność wygenerowanego raportu za pomocą podglądu raportu (*Preview*).