**HURTOWNIE DANYCH**

Laboratorium 6

Maciej Kopiński 254578

**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

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).
2. 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
3. wdrożyć i przeprocesować kostkę.

**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.**

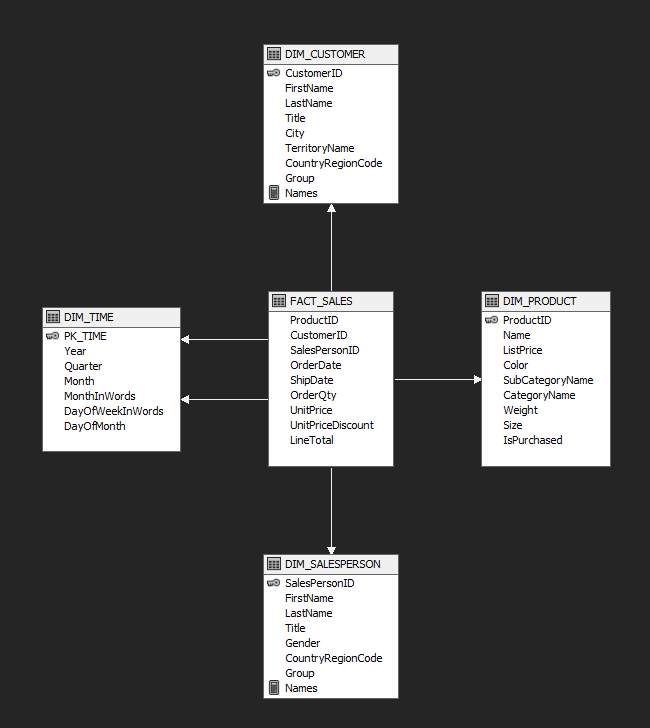
Korzystając z bazy danych AdventureWorksDW:

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:**

**Zad.1**

1. Kostka z atrybutami Names zamiast FirstName i LastName



Hierarchie:



Obraz zawierający tekst

Opis wygenerowany automatycznie



Obraz zawierający tekst

Opis wygenerowany automatycznie









Obraz zawierający tekst

Opis wygenerowany automatycznie

Obraz zawierający tekst

Opis wygenerowany automatycznie

**Zad.2**

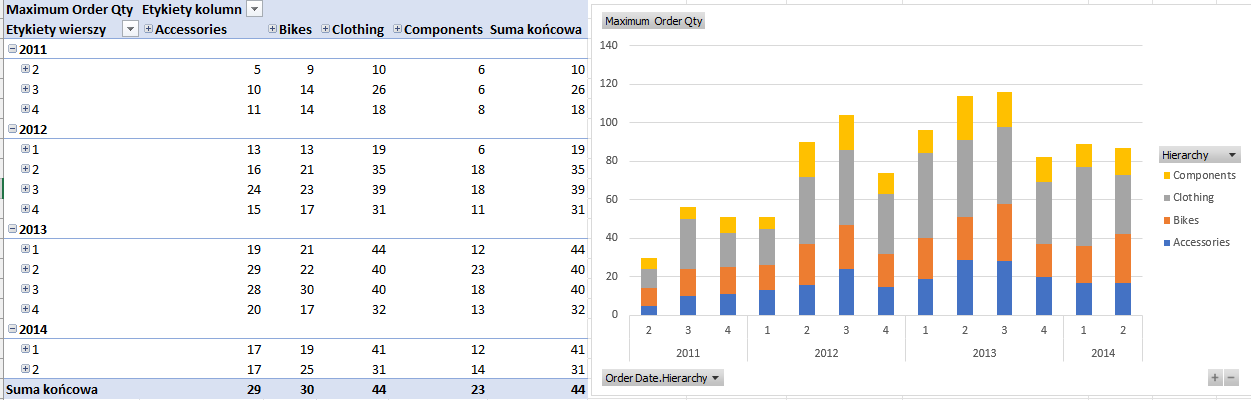
Obraz zawierający stół

Opis wygenerowany automatycznie

Obraz zawierający tekst

Opis wygenerowany automatycznie

*1.Maksymalny rozmiar zamówienia dla kategorii w zależności od roku i kwartału*

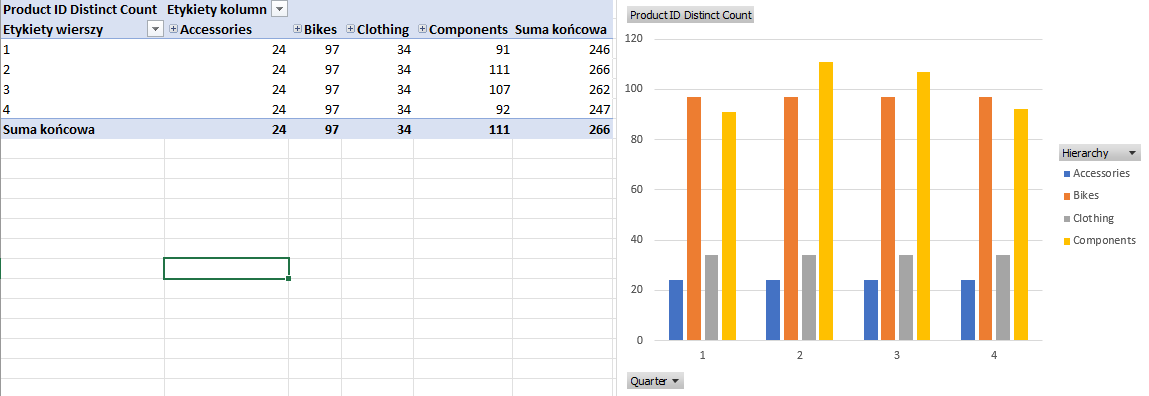


*2.Liczba różnych produktów w zależności od roku i kwartału dla poszczególnych kontynentów*

Obraz zawierający tekst, wewnątrz, zrzut ekranu

Opis wygenerowany automatycznie

*3.Liczba różnych produktów w zależności od kwartału i kategorii*



**Zad.3**

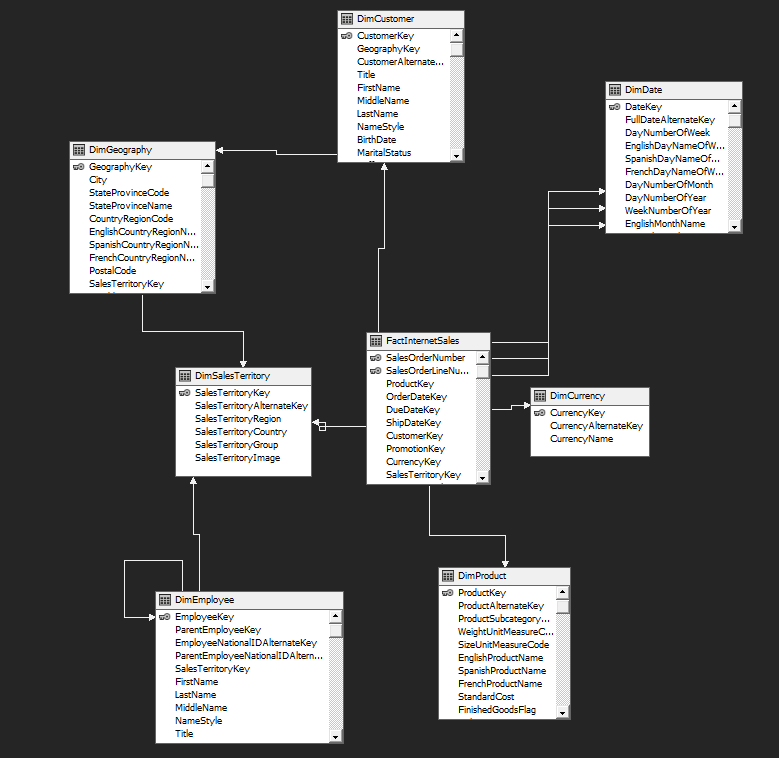
Partycje:

Obraz zawierający tekst, wewnątrz, ekran, zrzut ekranu

Opis wygenerowany automatycznie

**Zad.4**

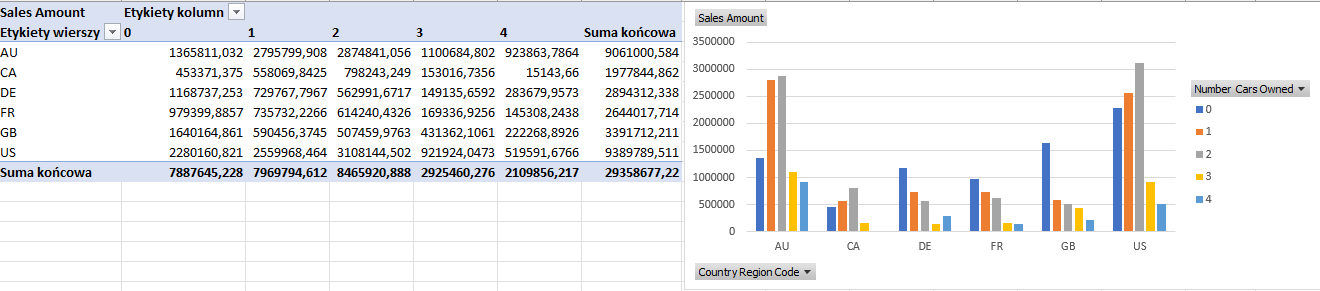
1. oraz b)



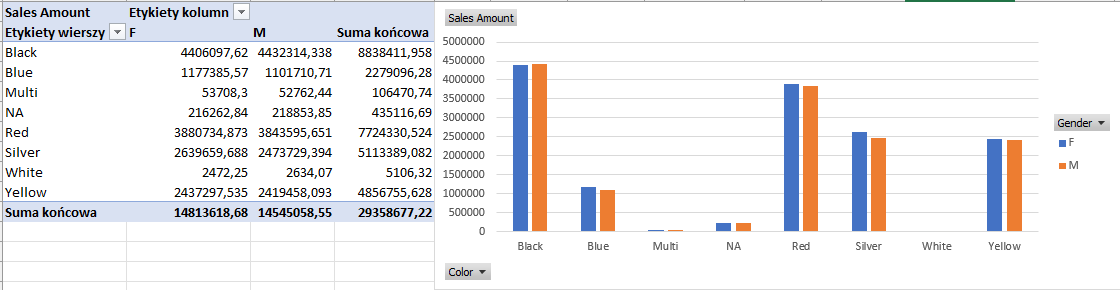
Obraz zawierający tekst

Opis wygenerowany automatycznie

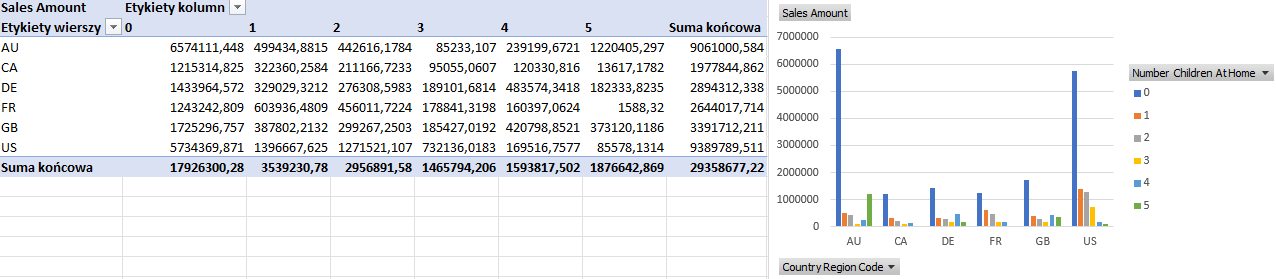
*1.Suma sprzedaży w zależności od kodu regionu i ilości posiadanych samochodów*



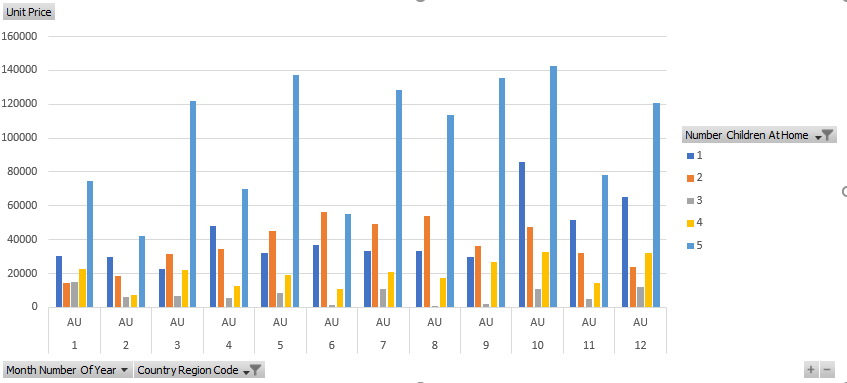
*2.Suma sprzedaży w zależności od koloru produktu i płci*



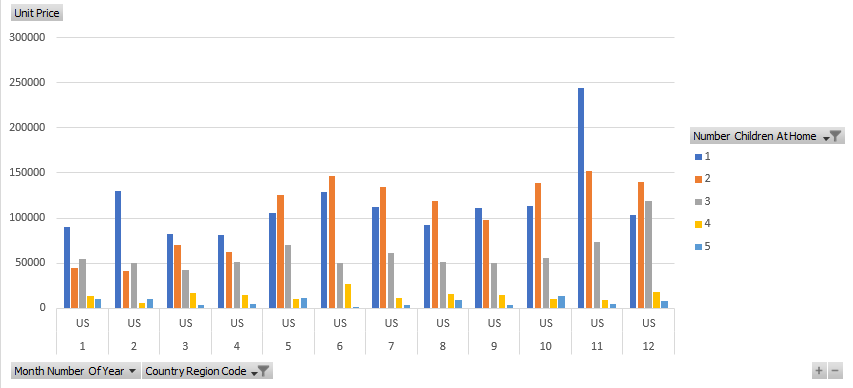
*3.Suma sprzedaży w zależności od kodu regionu i liczby dzieci*

****

*4.Suma sprzedaży dla osób z dziećmi według miesięcy w Australii*

****

*Dla Stanów Zjednoczonych*

****

**Wnioski:**

Wykres 2.c.1:

Klienci kupują w jednym zamówieniu najwięcej produktów z kategorii ubrania.

Wykres 2.c.2:

Sklep z roku na rok oferuje coraz więcej różnych produktów.

Wykres 2.c.3:

Najwięcej różnych produktów jest kupowane z kategorii rowery i komponenty.

Wykres 4.e.1:

Im więcej klient posiada samochodów, tym mniej kupuje.

Wykres 4.e.2:

Płeć klienta nie wpływa na kolory produktów, które kupuje.

Wykres 4.e.3:

Im większa liczba dzieci, tym niższe ilości produktów w zamówieniu. Wyjątkiem jest Australia, gdzie wciąż najwięcej kupują osoby bezdzietne, natomiast w przypadku posiadania 5 dzieci kupują więcej niż w przypadku posiadania mniejszej ilości dzieci.

Projekt typu Analysis Services pozwala na wygodne wybieranie, przetwarzanie i prezentowanie danych. Pozwala on na modyfikowanie przedstawionych danych bez konieczności modyfikowania ich w bazie poprzez funkcję Named Calculations czy tworzenie hierarchii pozwalającej na późniejsze wygodniejsze doprecyzowanie danych przykładowo na wykresie. Nie wymaga pisania zapytań do samej bazy w celu utworzenia tabel przestawnych prezentujących potrzebne nam dane.