

## Hurtownie Danych - laboratorium

### Lista 4

#### *Podstawy Integration Services oraz Analysis Services*

##### Wstęp teoretyczny

W ramach niniejszej listy wykorzystane zostaną podstawowe skrypty DDL oraz DML.

Tworzenie schematu:

```
CREATE SCHEMA schema_name [ <schema_element> [ ...n ] ]
```

Tworzenie tabeli:

```
CREATE TABLE  
  { database_name.schema_name.table_name. |  
    schema_name.table_name | table_name }  
  ( { <column_definition> } [ ,...n ] )
```

Wypełnianie tabeli danymi:

```
[ WITH <common_table_expression> [ ,...n ] ]  
INSERT  
{  
  [ TOP ( expression ) [ PERCENT ] ]  
  [ INTO ]  
  { <object> | rowset_function_limited  
    [ WITH ( <Table_Hint_Limited> [ ...n ] ) ]  
  }  
  {  
    [ ( column_list ) ]  
    [ <OUTPUT Clause> ]  
    { VALUES ( { DEFAULT | NULL | exp } [ ,...n ] )  
      | derived_table  
      | execute_statement  
      | <dml_table_source>  
      | DEFAULT VALUES  
    }  
  }  
}
```

Dodawanie więzów integralności:

```
ALTER TABLE table_name  
ADD CONSTRAINT constraint_name  
{  
  [ NULL | NOT NULL ]  
  { PRIMARY KEY | UNIQUE }  
}
```

**„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”**

```
[ CLUSTERED | NONCLUSTERED ]
[ WITH FILLFACTOR = fillfactor ]
[ WITH ( index_option [, ...n ] ) ]
[ ON { partition_scheme_name (partition_column_name)
    | filegroup | "default" } ]
FOREIGN KEY attribute_name
REFERENCES [ schema_name . ] referenced_table_name
    [ ( ref_column ) ]
[ ON DELETE { NO ACTION | CASCADE | SET NULL | SET DEFAULT } ]
[ NOT FOR REPLICATION ]
| CHECK [ NOT FOR REPLICATION ] ( logical_expression )
}
```

Źródła:

<https://docs.microsoft.com/en-us/sql/t-sql/statements/create-schema-transact-sql?view=sql-server-ver15>  
<https://docs.microsoft.com/en-us/sql/t-sql/statements/create-table-transact-sql?view=sql-server-ver15>  
<https://docs.microsoft.com/en-us/sql/t-sql/statements/insert-transact-sql?view=sql-server-ver15>  
<https://docs.microsoft.com/en-us/sql/t-sql/statements/alter-table-column-constraint-transact-sql?view=sql-server-ver15>

---

## Zadania do wykonania

Baza danych: **AdventureWorks**

### Zad 1. Przygotowanie schematu

W bazie danych należy utworzyć schemat, którego nazwa będzie odpowiadać nazwisku wykonującego ćwiczenie (zapisać zapytanie tworzące ten schemat).

### Zad 2. Tworzenie tabel wymiarów i tabeli faktów

W nowo utworzonym schemacie utworzyć tabele wymiarów: klienta, produktu i sprzedawcy (zapisać skrypt `CREATE TABLE`), opisane w następujących schematach:

- DIM\_CUSTOMER (CustomerID, FirstName, LastName, Title, City, TerritoryName, CountryRegionCode, Group)
- DIM\_PRODUCT (ProductID, Name, ListPrice, Color, SubCategoryName, CategoryName, Weight, Size, IsPurchased)
- DIM SALESPERSON (SalesPersonID, FirstName, LastName, Title, Gender, CountryRegionCode, Group)

oraz tabelę faktów:

- FACT\_SALES (ProductID, CustomerID, SalesPersonID, OrderDate, ShipDate, OrderQty, UnitPrice, UnitPriceDiscount, LineTotal)

*Uwaga 1: Proszę sprawdzić, dlaczego niektóre atrybuty mają wartość NULL. Czy zawsze wynika to z braku danych?*

*Uwaga 2: Kolumny OrderDate oraz ShipDate powinny przechowywać dane typu całkowitego, gdzie cztery pierwsze cyfry oznaczają rok, dwie następne miesiąc, a dwie ostatnie dzień. Do pobrania poszczególnych części daty użyć funkcji datepart.*

*Uwaga 3: Atrybuty OrderQty oraz LineTotal będą miarami w naszej hurtowni.*

**„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”**

**Zad. 3. Wypełnianie danych – denormalizacja źródłowej bazy**

Wypełnić nowoutworzone tabele danymi znajdującymi się w tabelach źródłowych. Do wypełnienia użyć instrukcji INSERT INTO. Proszę sprawdzić liczbę skopiowanych rekordów.

*Uwaga 1. Do tabeli DIM\_PRODUCT należy także skopiować produkty, które nie mają przypisanej podkategorii.*

*Uwaga 2. Do tabeli FACT\_SALES należy skopiować również transakcje, które nie mają sprzedawcy.*

Przykładowe rozwiązania:

	CustomerID	FirstName	LastName	Title	City	Name	CountryRegionCode	Group
1...	21258	Alfredo	Serrano	N...	Mat...	Aus...	AU	Paci...
1...	23844	Alexandra	Gonzales	N...	Port...	Nort...	US	Nort...
1...	12598	Alexandra	Bryant	N...	Sto...	Unit...	GB	Eur...
1...	24775	Alexandra	Alexander	N...	Colma	Sou...	US	Nort...
1...	20456	Alfredo	Ramos	N...	Sun...	Aus...	AU	Paci...
1...	26288	Alexandra	Russell	N...	Bur...	Sou...	US	Nort...
1...	16178	Alfredo	Carlson	N...	Syd...	Aus...	AU	Paci...
1...	16245	Alexandra	Griffin	N...	San...	Sou...	US	Nort...

*Przykładowe rozwiązanie – tabela DIM\_CUSTOMER.*

	ProductID	Name	ListPrice	Color	SubCategoryName	CategoryName	Weight	Size	IsPurchased
3.	853	Women's Tights, M	74,99	Black	Tights	Clothing	NULL	M	1
3.	854	Women's Tights, L	74,99	Black	Tights	Clothing	NULL	L	1
3.	855	Men's Bib-Shorts, S	89,99	Multi	Bib-Shorts	Clothing	NULL	S	1
3.	856	Men's Bib-Shorts, M	89,99	Multi	Bib-Shorts	Clothing	NULL	M	1
3.	857	Men's Bib-Shorts, L	89,99	Multi	Bib-Shorts	Clothing	NULL	L	1
3.	858	Half-Finger Gloves, S	24,49	Black	Gloves	Clothing	NULL	S	1
3.	859	Half-Finger Gloves, M	24,49	Black	Gloves	Clothing	NULL	M	1
3.	860	Half-Finger Gloves, L	24,49	Black	Gloves	Clothing	NULL	L	1
3.	861	Full-Finger Gloves, S	37,99	Black	Gloves	Clothing	NULL	S	1
3.	862	Full-Finger Gloves, M	37,99	Black	Gloves	Clothing	NULL	M	1
3.	863	Full-Finger Gloves, L	37,99	Black	Gloves	Clothing	NULL	L	1
3.	864	Classic Vest, S	63,50	Blue	Vests	Clothing	NULL	S	1
3.	865	Classic Vest, M	63,50	Blue	Vests	Clothing	NULL	M	1
3.	866	Classic Vest, L	63,50	Blue	Vests	Clothing	NULL	L	1
3.	867	Women's Mountain ...	69,99	Black	Shorts	Clothing	NULL	S	1
3.	868	Women's Mountain ...	69,99	Black	Shorts	Clothing	NULL	M	1

*Przykładowe rozwiązanie – tabela DIM\_PRODUCT.*

	SalesPersonID	FirstName	LastName	Title	Gender	CountryRegionCode	Group
1	274	Stephen	Jiang	NULL	NULL	US	North America
2	275	Michael	Blythe	NULL	NULL	US	North America
3	276	Linda	Mitchell	NULL	NULL	US	North America
4	277	Jillian	Carson	NULL	NULL	US	North America
5	278	Garrett	Vargas	NULL	NULL	CA	North America
6	279	Tsvi	Reiter	NULL	NULL	US	North America
7	280	Pamela	Ansman-Wolfe	NULL	NULL	US	North America
8	281	Shu	Ito	NULL	NULL	US	North America
9	282	José	Saraiva	NULL	NULL	CA	North America
10	283	David	Campbell	NULL	NULL	US	North America
11	284	Tete	Mensa-Annan	Mr.	M	US	North America
12	285	Syed	Abbas	Mr.	M	US	North America
13	286	Lynn	Tsofrlias	NULL	NULL	AU	Pacific
14	287	Amy	Alberts	NULL	NULL	US	North America
15	288	Rachel	Valdez	NULL	NULL	DE	Europe
16	289	Jae	Pak	NULL	NULL	GB	Europe
17	290	Ranjit	Varkey Chudukatil	NULL	NULL	FR	Europe

*Przykładowe rozwiązanie – tabela DIM SALESPERSON.*

**„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”**

	ProductID	CustomerID	SalesPersonID	OrderDate	ShipDate	OrderQty	UnitPrice	UnitPriceDiscount	LineTotal
1	776	29825	279	20110531	20110607	1	2024,994	0,00	2024,994000
2	711	29825	279	20110531	20110607	4	20,1865	0,00	80,746000
3	709	29825	279	20110531	20110607	6	5,70	0,00	34,200000
4	716	29825	279	20110531	20110607	1	28,8404	0,00	28,840400
5	714	29825	279	20110531	20110607	3	28,8404	0,00	86,521200
6	774	29825	279	20110531	20110607	1	2039,994	0,00	2039,994000
7	712	29825	279	20110531	20110607	2	5,1865	0,00	10,373000
8	772	29825	279	20110531	20110607	1	2039,994	0,00	2039,994000
9	771	29825	279	20110531	20110607	1	2039,994	0,00	2039,994000
10	778	29825	279	20110531	20110607	1	2024,994	0,00	2024,994000
11	773	29825	279	20110531	20110607	2	2039,994	0,00	4079,988000
12	777	29825	279	20110531	20110607	3	2024,994	0,00	6074,982000
13	758	29672	279	20110531	20110607	1	874,794	0,00	874,794000
14	762	29672	279	20110531	20110607	1	419,4589	0,00	419,458900
15	741	29734	282	20110531	20110607	2	818,70	0,00	1637,400000
16	708	29734	282	20110531	20110607	5	20,1865	0,00	100,932500
17	777	29734	282	20110531	20110607	2	2024,994	0,00	4049,988000
18	716	29734	282	20110531	20110607	2	20,1865	0,00	40,373000

Przykładowe rozwiązanie – tabela FACT\_SALES.

#### Zad. 4. Więzy integralności

1. Dodać integralność referencyjną i klucze główne do tabel już zdefiniowanych.
2. Przygotować instrukcję INSERT INTO, która sprawdzi poprawność integralności referencyjnej oraz klucze główne.

#### Zad. 5. Tworzenie kostki

Należy utworzyć projekt Analysis Services, w którym zostanie przygotowana kostka zawierająca utworzone wymiary (klienta, produktu i sprzedawcy) oraz tabelę faktów.

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 faktów i wymiarów. Przeanalizować opcje związane z bezpieczeństwem dostępu do danych.
- b) Utworzyć nowy widok źródła danych (Solution Explorer -> Data Source Views -> New Data Source View). Dodać wcześniej utworzone tabele.
- c) Utworzyć nową kostkę za pomocą asystenta (Solution Explorer -> Cubes -> New Cube):
  - o Wybrać utworzenie kostki na podstawie istniejących tabel (Use existing tables)
  - o Wybrać, utworzony w poprzednim punkcie, widok źródła danych
  - o Jako tabelę faktów (Measure group tables) wybrać FACT\_SALES
  - o 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?
  - o Na stronie dotyczącej wymiarów wybrać wszystkie tabele z przedrostkiem DIM.

Po utworzeniu kostki dokonać edycji wymiarów (Solution Explorer -> Dimensions lub zakładka Cube Structure -> Dimensions).

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).

**„ZPR PWr – Zintegrowany Program Rozwoju Politechniki Wrocławskiej”**

**Zad. 6. Uruchomienie kostki**

- a) Sprawdzić ustawienia projektu (Solution Explorer -> Nazwa projektu -> Properties).
- b) Sprawdzić ustawienia konfiguracji wdrożenia (Configuration properties -> Deployment -> Target) projektu – Server: localhost, Database: PierwszyProjektKostek
- c) Wdrożyć projekt (Menu Build -> Deploy). W przypadku błędu:
  - w konfiguracji wdrożenia (Configuration properties -> Deployment -> Target) przekleić nazwę serwera (z ekranu wyboru serwera w MSSQL Management Studio);
  - klikając dwukrotnie na utworzony Data Source, sprawdzić działanie innych opcji „Impersonation information” (VS powinien być uruchamiany z konta o uprawnieniach administratora);
- d) Przetworzyć kostkę (Przycisk Process).

**Zad. 7. Proste raporty**

Połączyć się z programem MS Excel i przygotować raporty (tabele i wykresy przestawne), w których zostaną ujęte ciekawe zależności pomiędzy danymi.

**Rozwiązania:**

**Wnioski:**