

# 4. sprawozdanie z laboratorium Hurtownie Danych

Mikołaj Kubś, 272662

10 kwietnia 2025

1. DDL (Data Definition Language) - tworzenie struktur danych i schema z użyciem poleceń CREATE, ALTER, DROP
2. DML (Data Manipulation Language) - manipulacja danymi w tabelach z użyciem poleceń INSERT, UPDATE, DELETE
3. DCL (Data Control Language) - zarządzanie uprawnieniami, za pomocą poleceń GRANT, DENY, REVOKE
4. DQL (Data Query Language) - pozyskiwanie danych z bazy, za pomocą polecenia SELECT

## 1 Zadanie 1 - przygotowanie schematu

```
1 CREATE SCHEMA Kubs
```

Utworzenie dedykowanego schematu pozwala na logiczne odseparowanie obiektów stworzonych na potrzeby laboratorium od pozostałych struktur bazy danych.

## 2 Zadanie 2 - Tworzenie tabel wymiarów i tabeli faktów

```
1  
2 CREATE TABLE Kubs.DIM_CUSTOMER (  
3     CustomerID INT NOT NULL,  
4     FirstName NVARCHAR(50) NULL,  
5     LastName NVARCHAR(50) NULL,
```

```

6      Title NVARCHAR(8) NULL,
7      City NVARCHAR(30) NULL,
8      TerritoryName NVARCHAR(50) NULL,
9      CountryRegionCode NVARCHAR(3) NULL,
10     [Group] NVARCHAR(50) NULL
11 );
12
13 CREATE TABLE Kubs.DIM_PRODUCT (
14     ProductID INT NOT NULL,
15     Name NVARCHAR(50) NOT NULL,
16     ListPrice MONEY NULL,
17     Color NVARCHAR(15) NULL,
18     SubCategoryName NVARCHAR(50) NULL,
19     CategoryName NVARCHAR(50) NULL,
20     Weight DECIMAL(8, 2) NULL,
21     Size NVARCHAR(5) NULL,
22     IsPurchased BIT NULL DEFAULT 0
23 );
24
25 CREATE TABLE Kubs.DIM SALESPERSON (
26     SalesPersonID INT NOT NULL,
27     FirstName NVARCHAR(50) NULL,
28     LastName NVARCHAR(50) NULL,
29     Title NVARCHAR(8) NULL,
30     Gender NCHAR(1) NULL,
31     CountryRegionCode NVARCHAR(3) NULL,
32     [Group] NVARCHAR(50) NULL
33 );
34
35 CREATE TABLE Kubs.FACT_SALES (
36     ProductID INT NOT NULL,
37     CustomerID INT NOT NULL,
38     SalesPersonID INT NULL,
39     OrderDate INT NOT NULL,
40     ShipDate INT NULL,
41     OrderQty SMALLINT NOT NULL,
42     UnitPrice MONEY NOT NULL,
43     UnitPriceDiscount DECIMAL(8, 4) NOT NULL DEFAULT 0,
44     LineTotal DECIMAL(19, 4) NOT NULL
45 );

```

Zgodnie z poleceniem, kolumny OrderDate oraz ShipDate będą przechowywać dane typu całkowitego. W tabeli faktów sprzedawca nie jest wymagany. W tabeli wymiaru dla produktu kategoria i podkategoria również mogą

być NULL. Tabele DIM\_ pełnią oczywiście role wymiarów, a FACT\_SALES jest tabelą faktów zawierającą miary dotyczące transakcji sprzedaży.

Wartości NULL w niektórych kolumnach wynikają z opcjonalności tych danych w oryginalnym źródle danych lub mogą być nieobecne (również z powodu braku połączenia z inną tabelą, np. produkt bez podkategorii nie może mieć kategorii).

### 3 Zadanie 3 - wypełnianie danych

```
1  INSERT INTO Kubs.DIM_CUSTOMER (
2      CustomerID,
3      FirstName,
4      LastName,
5      Title,
6      City,
7      TerritoryName,
8      CountryRegionCode,
9      [Group]
10 )
11 SELECT DISTINCT
12     c.CustomerID,
13     p.FirstName,
14     p.LastName,
15     p.Title,
16     a.City,
17     st.Name AS TerritoryName,
18     st.CountryRegionCode,
19     st.[Group]
20 FROM Sales.Customer AS c
21 LEFT JOIN Person.Person AS p ON c.PersonID = p.BusinessEntityID
22 LEFT JOIN Sales.SalesTerritory AS st
23     ON c.TerritoryID = st.TerritoryID
24 LEFT JOIN Person.BusinessEntityAddress bea
25     ON p.BusinessEntityID = bea.BusinessEntityID
26 LEFT JOIN Person.Address AS a ON bea.AddressID = a.AddressID
27 WHERE c.PersonID IS NOT NULL;
28
29 INSERT INTO Kubs.DIM_PRODUCT (
30     ProductID,
31     Name,
32     ListPrice,
33     Color,
```

```

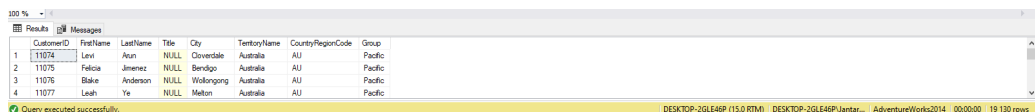
34     SubCategoryName,
35     CategoryName,
36     Weight,
37     Size,
38     IsPurchased
39 )
40 SELECT DISTINCT
41     p.ProductID,
42     p.Name,
43     p.ListPrice,
44     p.Color,
45     psc.Name AS SubCategoryName,
46     pc.Name AS CategoryName,
47     p.Weight,
48     p.Size,
49     1 AS IsPurchased
50 FROM Production.Product AS p
51 INNER JOIN Sales.SalesOrderDetail AS sod
52     ON p.ProductID = sod.ProductID
53 LEFT JOIN Production.ProductSubcategory AS psc
54     ON p.ProductSubcategoryID = psc.ProductSubcategoryID
55 LEFT JOIN Production.ProductCategory AS pc
56     ON psc.ProductCategoryID = pc.ProductCategoryID;
57
58 INSERT INTO Kubs.DIM_SALESPERSON (
59     SalesPersonID,
60     FirstName,
61     LastName,
62     Title,
63     Gender,
64     CountryRegionCode,
65     [Group]
66 )
67 SELECT
68     sp.BusinessEntityID AS SalesPersonID,
69     p.FirstName,
70     p.LastName,
71     p.Title,
72     e.Gender,
73     st.CountryRegionCode,
74     st.[Group]
75 FROM Sales.SalesPerson AS sp
76 INNER JOIN Person.Person AS p

```

```

77     ON sp.BusinessEntityID = p.BusinessEntityID
78 INNER JOIN HumanResources.Employee AS e
79     ON sp.BusinessEntityID = e.BusinessEntityID
80 LEFT JOIN Sales.SalesTerritory AS st
81     ON sp.TerritoryID = st.TerritoryID;
82
83 INSERT INTO Kubs.FACT_SALES (
84     ProductID,
85     CustomerID,
86     SalesPersonID,
87     OrderDate,
88     ShipDate,
89     OrderQty,
90     UnitPrice,
91     UnitPriceDiscount,
92     LineTotal
93 )
94 SELECT
95     sod.ProductID,
96     soh.CustomerID,
97     soh.SalesPersonID,
98     DATEPART(YEAR, soh.OrderDate) * 10000 +
99     DATEPART(MONTH, soh.OrderDate) * 100 +
100    DATEPART(DAY, soh.OrderDate) AS OrderDate,
101    DATEPART(YEAR, soh.ShipDate) * 10000 +
102    DATEPART(MONTH, soh.ShipDate) * 100 +
103    DATEPART(DAY, soh.ShipDate) AS ShipDate,
104    sod.OrderQty,
105    sod.UnitPrice,
106    sod.UnitPriceDiscount,
107    sod.LineTotal
108 FROM Sales.SalesOrderDetail AS sod
109 INNER JOIN Sales.SalesOrderHeader AS soh
110     ON sod.SalesOrderID = soh.SalesOrderID;

```



	CustomerID	FirstName	LastName	Title	City	TerritoryName	CountryRegionCode	Group
1	11074	Levi	Arin	NULL	Cloverdale	Australia	AU	Pacific
2	11075	Felicia	Jimenez	NULL	Berdingo	Australia	AU	Pacific
3	11076	Blake	Anderson	NULL	Wollongong	Australia	AU	Pacific
4	11077	Leah	Ye	NULL	Melton	Australia	AU	Pacific

Query executed successfully. DESKTOP-2GLE4BP (15.0 RTM) DESKTOP-2GLE4BP\jantar... AdventureWorks2014 00:00:00 19130 rows

Rysunek 1: Pierwsze 3 wiersze DIM\_CUSTOMER i liczba wierszy - 19130

ProductID	Name	ListPrice	Color	SubCategoryName	CategoryName	Weight	Size	IsPurchased
984	Mountain-500 Silver, 40	564.99	Silver	Mountain Bikes	Bikes	27.35	40	1
722	LL Road Frame - Black, 58	337.22	Black	Road Frames	Components	2.46	58	1
964	Classic Vest, S	63.90	Blue	Vests	Clothing	NULL	S	1

Rysunek 2: Pierwsze 3 wiersze DIM\_PRODUCT i liczba wierszy - 266

SalesPersonID	FirstName	LastName	Title	Gender	CountryRegionCode	Group
274	Stephen	Jiang	NULL	M	NULL	NULL
275	Michael	Blythe	NULL	M	US	North America
276	Linda	Mitchell	NULL	F	US	North America

Rysunek 3: Pierwsze 3 wiersze DIM\_SALESPERSON i liczba wierszy - 17

ProductID	CustomerID	SalesPersonID	OrderDate	ShipDate	OrderQty	UnitPrice	UnitPriceDiscount	LineTotal
870	29485	276	20140129	20140205	5	2.994	0.0000	14.9700
964	29485	276	20140129	20140205	4	38.10	0.0000	152.4000
970	29485	276	20140129	20140205	3	729.91	0.0000	2189.7300

Rysunek 4: Pierwsze 3 wiersze FACT\_SALES i liczba wierszy - 121317

Zgodnie z poleceniem wypełniono tabele danymi z bazy danych AdventureWorks2014. Ponieważ analizy będą dotyczyć sprzedaży produktów, pominięto wszystkie produkty, które nigdy nie zostały kupione. Okazało się, że wszystkie produkty, które nie miały kategorii lub podkategorii, nigdy nie były kupione.

Liczby wierszy:

DIM\_CUSTOMER - 19130

DIM\_PRODUCT - 266

DIM\_SALESPERSON - 17

FACT\_SALES - 121317