Hurtownie danych  
Laboratorium Czw 11:15  
  
Lista 5  
  
Kajetan Pynka 254495

**Zad 1**

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'FACT\_SALES'))

BEGIN

    DROP TABLE Pynka.FACT\_SALES

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'DIM\_TIME'))

BEGIN

    DROP TABLE Pynka.DIM\_TIME

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'ENUM\_MONTHS'))

BEGIN

    DROP TABLE Pynka.ENUM\_MONTHS

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'ENUM\_WEEKDAYS'))

BEGIN

    DROP TABLE Pynka.ENUM\_WEEKDAYS

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'DIM\_CUSTOMER'))

BEGIN

    DROP TABLE Pynka.DIM\_CUSTOMER

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'DIM\_PRODUCT'))

BEGIN

    DROP TABLE Pynka.DIM\_PRODUCT

END;

IF (EXISTS (SELECT \*

                 FROM INFORMATION\_SCHEMA.TABLES

                 WHERE TABLE\_SCHEMA = 'Pynka'

                 AND  TABLE\_NAME = 'DIM\_SALESPERSON'))

BEGIN

    DROP TABLE Pynka.DIM\_SALESPERSON

END;

**Zad 2.**

CREATE TABLE Pynka.ENUM\_MONTHS (

    Miesiac INT PRIMARY KEY,

    Miesiac\_slownie VARCHAR(11) NOT NULL

);

INSERT INTO Pynka.ENUM\_MONTHS VALUES(1, 'styczeń');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(2, 'luty');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(3, 'marzec');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(4, 'kwiecień');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(5, 'maj');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(6, 'czerwiec');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(7, 'lipiec');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(8, 'sierpień');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(9, 'wrzesień');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(10, 'październik');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(11, 'listopad');

INSERT INTO Pynka.ENUM\_MONTHS VALUES(12, 'grudzień');

CREATE TABLE Pynka.ENUM\_WEEKDAYS (

    Dzien INT PRIMARY KEY,

    Dzien\_tygodnia VARCHAR(12) NOT NULL

);

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(7, 'poniedziałek');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(2, 'wtorek');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(3, 'środa');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(4, 'czwartek');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(5, 'piątek');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(6, 'sobota');

INSERT INTO Pynka.ENUM\_WEEKDAYS VALUES(1, 'niedziela');

CREATE TABLE Pynka.DIM\_TIME (

    PK\_TIME INT PRIMARY KEY,

    Rok INT NOT NULL,

    Kwartal INT NOT NULL,

    Miesiac INT NOT NULL,

    Miesiac\_slownie VARCHAR(11) NOT NULL,

    Dzien\_tygodnia VARCHAR(12) NOT NULL,

    Dzien\_miesiaca INT NOT NULL

);

INSERT INTO Pynka.DIM\_TIME

SELECT CAST(FORMAT(DateRec, 'yyyy') + FORMAT(DateRec, 'MM') + FORMAT(DateRec, 'dd') AS INT),

    YEAR(DateRec), DATEPART(quarter, DateRec), MONTH(DateRec), EM.Miesiac\_slownie,

    EW.Dzien\_tygodnia, DAY(DateRec)

    FROM (

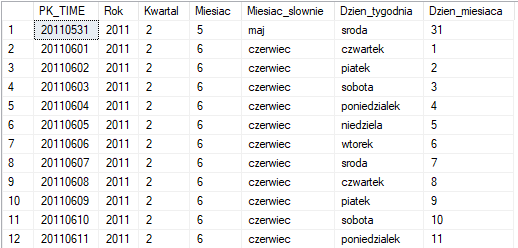
        SELECT DISTINCT OrderDate "DateRec" FROM Sales.SalesOrderHeader

        UNION SELECT ShipDate FROM Sales.SalesOrderHeader) I

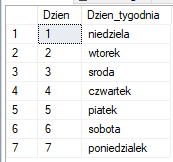
    JOIN Pynka.ENUM\_MONTHS EM ON EM.Miesiac=MONTH(DateRec)

    JOIN Pynka.ENUM\_WEEKDAYS EW ON EW.Dzien=DATEPART(WEEKDAY, DateRec);

**Wynikowe tabele:**

****

****

****

**Zad 3.**

UPDATE Pynka.DIM\_PRODUCT SET Color='Unknown' WHERE Color IS NULL;

UPDATE Pynka.DIM\_PRODUCT SET SubCategoryName='Unknown' WHERE SubCategoryName IS NULL;

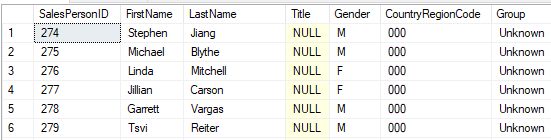
UPDATE Pynka.DIM\_CUSTOMER SET CountryRegionCode='000' WHERE CountryRegionCode IS NULL;

UPDATE Pynka.DIM\_CUSTOMER SET [Group]='Unknown' WHERE [Group] IS NULL;

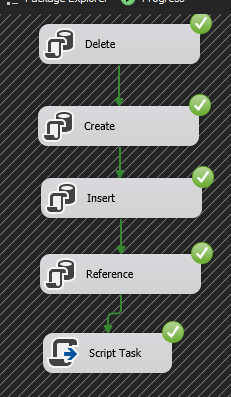
UPDATE Pynka.DIM\_SALESPERSON SET CountryRegionCode='000' WHERE CountryRegionCode IS NULL;

UPDATE Pynka.DIM\_SALESPERSON SET [Group]='Unknown' WHERE [Group] IS NULL;

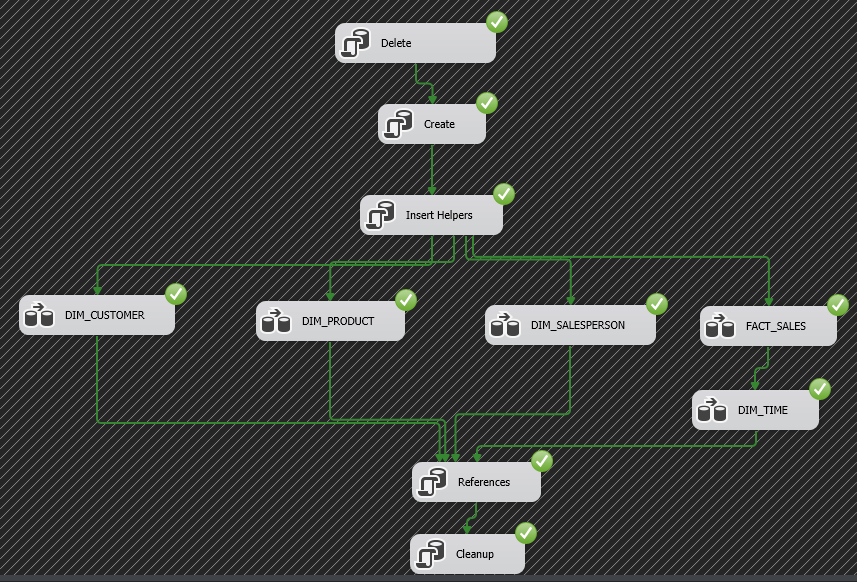
**Po czyszczeniu:**

****

**Zad 4.**

****

**Zad 5.**

****

**Wnioski:**

* W ramach listy przygotowaliśmy proces ETL korzystając z projektu „Integration Services” z Visual Studio, który pozwala nam w bardzo prosty sposób wywoływać kolejne kroki ekstrakcji, transformacji i wczytywania danych. W związku z tym mamy do czynienia w zasadzie z gotowym skryptem i za pomocą jednego kliknięcia możemy wykonać wszystkie wcześniej wymienione kroki (bez ponownego pytania zapytań czy wpisywania ich w dobrej kolejności). W przypadku zmian wystarczy zaktualizować to co potrzebne w ramach projektu „Integration Services”.
* Dodatkowo możemy rozbić poszczególne elementy zapytań SQL’owych na „bloczki” dostarczone przez VS w ramach projektu „Integration Services”. Może to pozwolić szybciej i łatwiej odczytywać poszczególne kroki procesu ETL a co za tym idzie w celu dokonania zmian nie trzeba „grzebać” w kodzie SQL tylko wystarczy „wyklikać” odpowiednie rzeczy w interfejsie graficznym.