



POLITECHNIKA RZESZOWSKA

im. Ignacego Łukasiewicza

WYDZIAŁ MATEMATYKI I FIZYKI STOSOWANEJ

Projekt

Wielowymiarowa analiza danych

Wykonali:

Kamil Baran 166631

Adam Gardyas 166725

Inżynieria i analiza danych

Spis treści

1.	Opis danych	3
1.1.	Opis danych z tabeli sprzedaż	4
1.2.	Opis danych z tabeli ludzie	4
2.	Stworzenie bazy danych	5
3.	Import naszych danych	6
3.1.	Obsługa błędów	11
4.	Oczyszczanie załadowanych danych	13
5.	Tworzenie kluczy głównych i kluczy obcych	19
6.	Połączenie naszej bazy ze środowiskiem Tableau.....	20
6.1.	Stworzenie źródła danych ODBC	20
6.2.	Połączenie Tableau z bazą danych	23
7.	Integracja danych ETL	24
7.1.	Ilość sprzedanych jednostek odzieży w krajach europejskich.	24
7.2.	Ilość wystąpień zamówień online i offline dla każdego państwa Azji.....	26
7.3.	Ilość wystąpień danego sklepu wraz ze sprzedawanym towarem oraz zarobkiem całkowity, który z niego płynie	27
7.4.	Ilość osób w określonym przedziale wiekowym z podziałem na płeć	28
8.	Wizualizacje	29
8.1.	Ilość sprzedanych jednostek odzieży w krajach europejskich.	29
8.2.	Ilość wystąpień zamówień online i offline dla każdego państwa Azji.....	30
8.3.	Ilość wystąpień danego sklepu wraz ze sprzedawanym towarem oraz zarobkiem całkowity, który z niego płynie	31
8.4.	Ilość osób w określonym przedziale wiekowym z podziałem na płeć	32
9.	Wnioski	33

1. Opis danych

Dane które użyliśmy w naszym projekcie są to dane sprzedażowe, dotyczą sprzedaży określonego produktu w określonym regionie. Tabela sprzedaż_gotowe wygląda następująco.

	Order ID	Region	Country	Item Type	Sales Channel	Order Priority	Order Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit	Place of delivery	Customer ID
1	443368995	Sub-Saharan Africa	South Africa	Fruits	Offline	M	2012-07-27	1593	9.33	6.92	14862.69	11023.56	3839.13	grocery store	206055
2	667593514	Middle East and North Africa	Morocco	Clothes	Online	M	2013-09-14	4611	109.28	35.84	503890.08	165258.24	338631.84	clothes shop	66005
3	940995585	Australia and Oceania	Papua New Guinea	Meat	Offline	M	2015-05-15	360	421.89	364.69	151880.4	131288.4	20592	grocery store	54808
4	880811536	Sub-Saharan Africa	Djibouti	Clothes	Offline	H	2017-05-17	562	109.28	35.84	61415.36	20142.08	41273.28	clothes shop	172628
5	174590194	Europe	Slovakia	Beverages	Offline	L	2016-10-26	3973	47.45	31.79	188518.85	126301.67	62217.18	grocery store	279254
6	830192887	Asia	Sri Lanka	Fruits	Online	L	2011-07-11	1379	9.33	6.92	12866.07	9542.68	3323.39	grocery store	116555
7	425793445	Sub-Saharan Africa	Seychelles	Beverages	Online	M	2013-01-18	597	47.45	31.79	28327.65	18978.63	9349.02	grocery store	197684
8	659878194	Sub-Saharan Africa	Tanzania	Beverages	Online	L	2016-11-30	1476	47.45	31.79	70036.2	46922.04	23114.16	grocery store	27199
9	601245963	Sub-Saharan Africa	Ghana	Office Supplies	Online	L	2017-03-23	896	651.21	524.96	583484.16	470364.16	113120	stationery store	288416
10	739008080	Sub-Saharan Africa	Tanzania	Cosmetics	Offline	L	2016-05-23	7768	437.2	263.33	3396169.6	2045547.44	1350622.16	cosmetics store	211512
11	732588374	Asia	Taiwan	Fruits	Offline	M	2014-09-02	8034	9.33	6.92	74957.22	55595.28	19361.94	grocery store	78315
12	761723172	Middle East and North Africa	Algeria	Cosmetics	Online	M	2011-02-18	9669	437.2	263.33	4227286.8	2546137.77	1681149.03	cosmetics store	175646
13	176461303	Asia	Singapore	Snacks	Online	C	2013-01-28	7676	152.58	97.44	1171204.08	747949.44	423254.64	clothes shop	212760
14	647164094	Australia and Oceania	Papua New Guinea	Clothes	Offline	L	2011-06-20	9092	109.28	35.84	993573.76	325857.28	667716.48	clothes shop	182465
15	314505374	Asia	Vietnam	Personal Care	Online	M	2010-04-04	7984	81.73	56.67	652532.32	452453.28	200079.04	cosmetics store	239254
16	539471471	Sub-Saharan Africa	Uganda	Personal Care	Online	M	2014-06-19	451	81.73	56.67	36860.23	25558.17	11302.06	cosmetics store	166334
17	953361213	Sub-Saharan Africa	Zimbabwe	Office Supplies	Offline	C	2011-03-28	9623	651.21	524.96	6266593.83	5051690.08	1214903.75	stationery store	106372
18	807785928	Sub-Saharan Africa	Ethiopia	Cosmetics	Online	M	2011-07-07	662	437.2	263.33	289426.4	174324.46	115101.94	cosmetics store	274135

Posiada ona 1 milion wiersz i jest ona połączona z tabelą ludzie, która posiada 300 000 wierszy. Customer ID z tabeli sprzedaż_gotowe odpowiada kolumnie ID z tabeli ludzie.

Tabela ludzie wygląda następująco:

	ID	First Name	Last Name	Gender	E Mail	Date of Birth	Phone No	City
1	1	Hemila	Suhr	F	hemila.suhr@gmail.com	1992-04-09	479-539-4593	Peach Orchard
2	2	Antonio	Joy	M	antonio.joy@yahoo.com	1989-12-24	229-234-6154	Rocky Ford
3	3	Sebastian	Moores	M	sebastian.moores@ibm.com	1980-09-23	212-231-9912	Antwerp
4	4	Alec	Rittenhouse	M	alec.rittenhouse@gmail.com	1974-01-06	229-873-6796	Milledgeville
5	5	Reggie	Doughty	M	reggie.doughty@gmail.com	1966-11-27	314-677-4501	Springfield
6	6	Elisha	Bottom	M	elisha.bottom@yahoo.com	1972-09-23	206-233-8897	Dryden
7	7	Danilo	Irwin	M	danilo.irwin@gmail.com	1964-12-22	319-326-7935	Urbandale
8	8	Madalene	Dierks	F	madalene.dierks@yahoo.com	1957-10-29	503-778-8049	Deer Island
9	9	Micheal	Campanella	F	micheal.campanella@gmail.com	1969-11-15	216-821-3070	Leavittsburg
10	10	Wilma	Trail	F	wilma.trail@yahoo.com	1966-11-28	505-988-5623	Organ
11	11	Jorge	Mohamed	M	jorge.mohamed@exxonmobil.com	1970-06-23	262-939-1044	Houlton
12	12	Isaac	Cumberland	M	isaac.cumberland@rediffmail.com	1982-09-15	209-304-3773	El Monte
13	13	Emerson	Rayner	M	emerson.rayner@gmail.com	1972-02-22	316-451-6275	Shawnee Mission
14	14	Booker	Olmo	M	booker.olmo@yahoo.com	1980-05-17	236-384-0434	Edwardsville
15	15	Margarito	Cushman	M	margarito.cushman@microsoft.com	1990-07-01	503-264-1465	Rickreall
16	16	Debbie	Diebold	F	debbie.diebold@bp.com	1984-02-20	219-549-4320	Schererville
17	17	Michal	Sturtevant	M	michal.sturtevant@btintemet.com	1974-11-02	803-374-2594	Greenville
18	18	Jonnie	Wingate	F	jonnie.wingate@aol.com	1977-02-10	270-434-9928	Saint Catharine

Dane pobraliśmy z następujących stron i przerobiliśmy, tak że zostawiliśmy tylko według nas najistotniejsze kolumny. Jest to tabela danych klientów, którzy dokonali określone zamówienie.

<https://excelbianalytics.com/wp/downloads-16-sample-csv-files-data-sets-for-testing/>

<https://excelbianalytics.com/wp/downloads-18-sample-csv-files-data-sets-for-testing-sales/>

1.1. Opis danych z tabeli sprzedaż

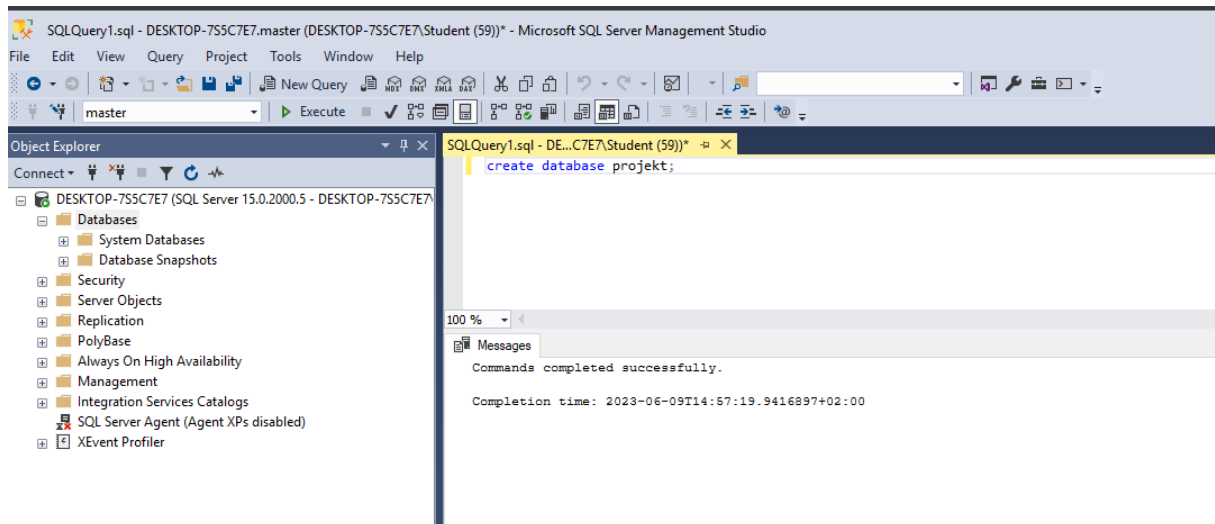
- Order ID – ID zamówienia
- Region – Region, gdzie dokonano zamówienia
- Country- kraj gdzie dokonano zamówienia
- Item Type – zamówiony towar
- Sales Channel- jak zostało wykonane zamówienie Offline czy Online
- Order Priority – Priorytet zamówienia
- Order Date – Data zamówienia
- Units Sold – ilość sprzedanych jednostek produktu
- Unit Price – cena jednostkowa
- Unit Cost- koszt jednostkowy
- Total Revenue- dochód całkowity
- Total Cost – całkowity koszt
- Total Profit – całkowity profit
- Place of delivery – miejsce dostarczenia
- Customer ID – ID klienta

1.2. Opis danych z tabeli ludzie

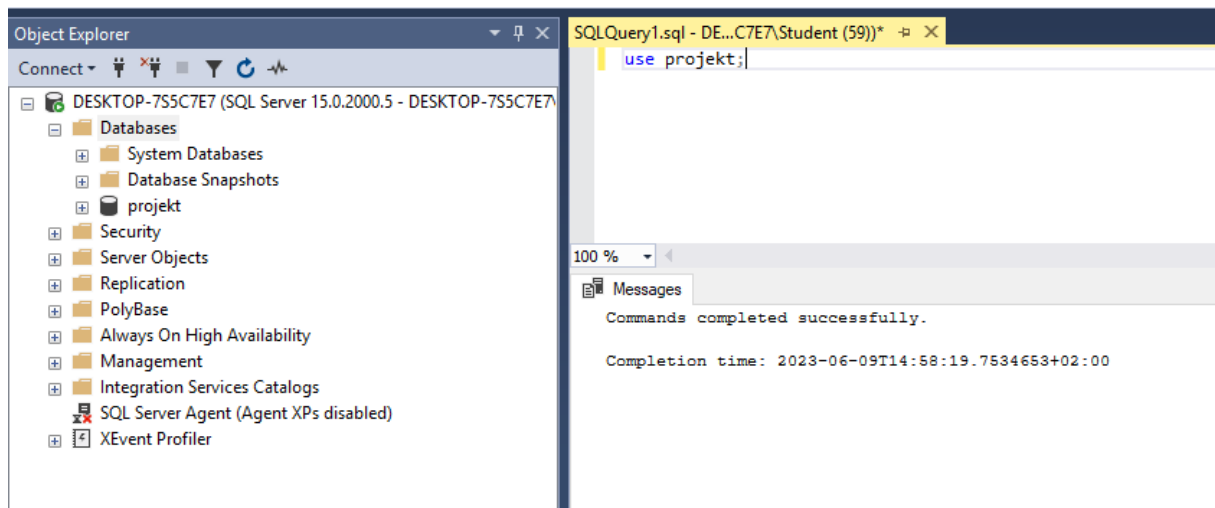
- ID – ID danej osoby
- First Name - Imię
- Last Name - Nazwisko
- Gende - płeć
- E Mail – e-mail
- Date of Birth – data urodzenia
- Phone No – numer telefonu
- City – miasto

2. Stworzenie bazy danych

Na początku w programie SQL Server Management Studio stworzyliśmy naszą bazę. Nazwaliśmy ją projekt.



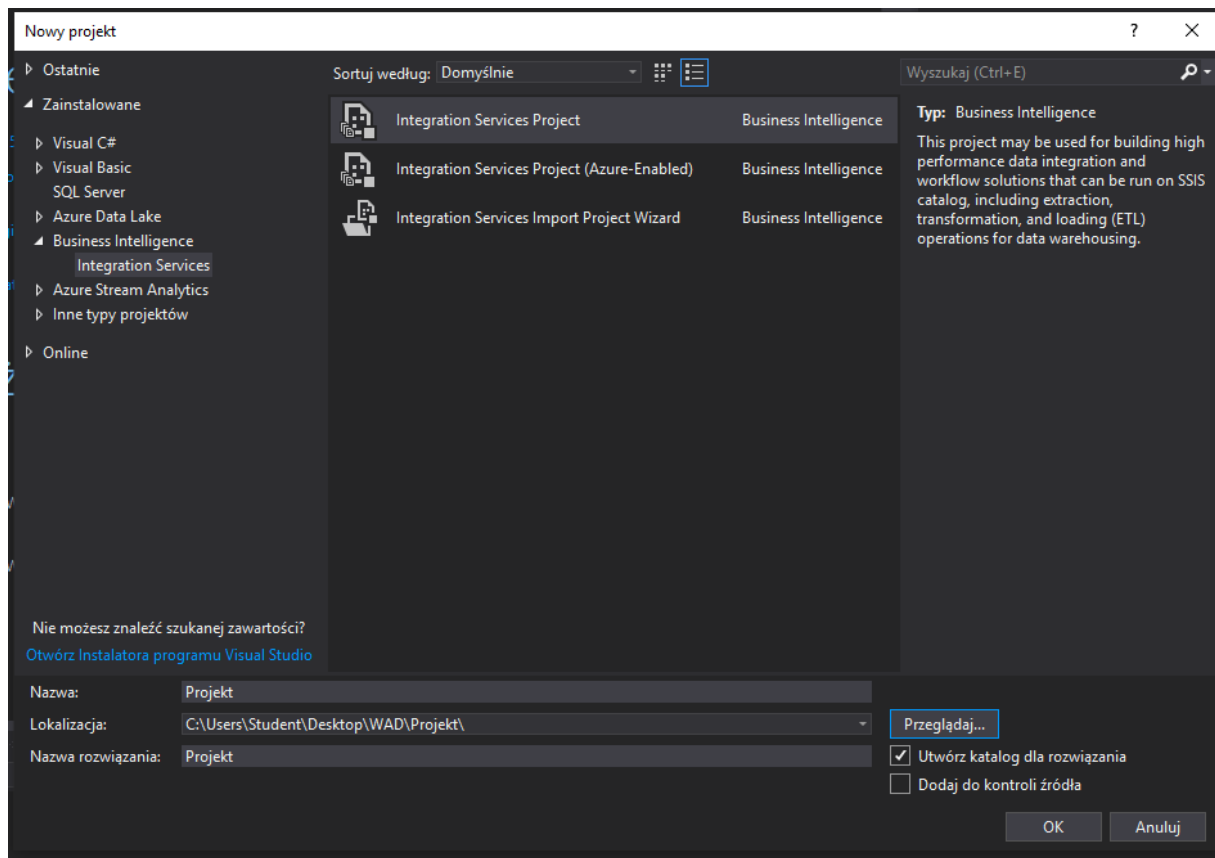
Jak widzimy nasza baza została poprawnie stworzona



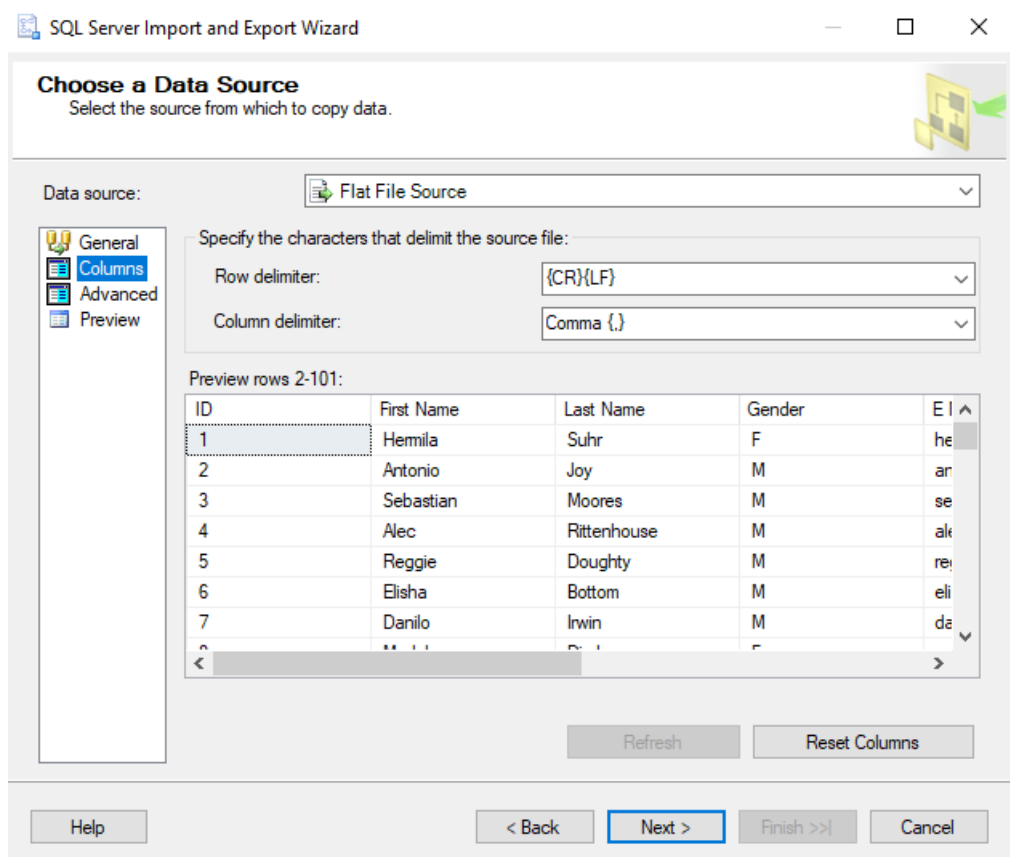
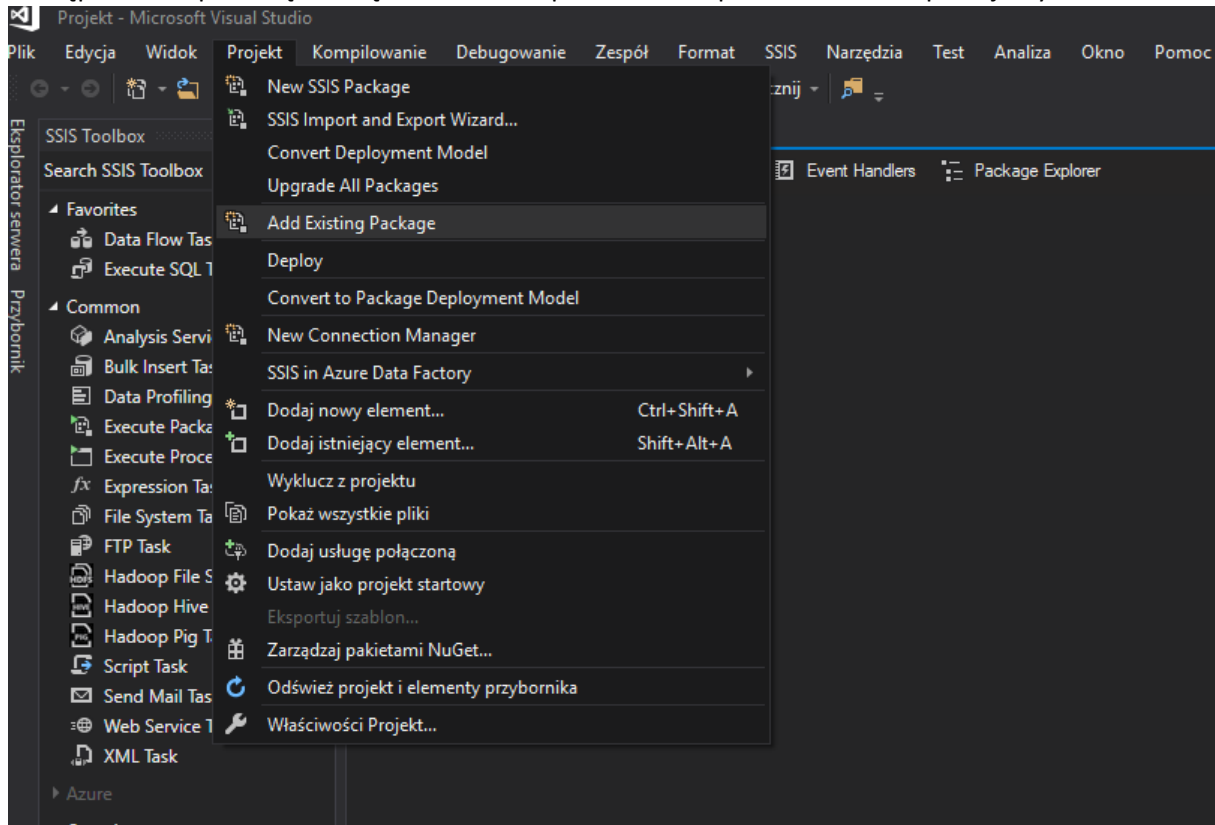
3. Import naszych danych

Nasze dane zostały zaimportowane za pomocą wbudowanego kreatora w Visual Studio jakim jest SSIS Import and Export Wizard. Proces wgrywania danych wyglądał następująco:

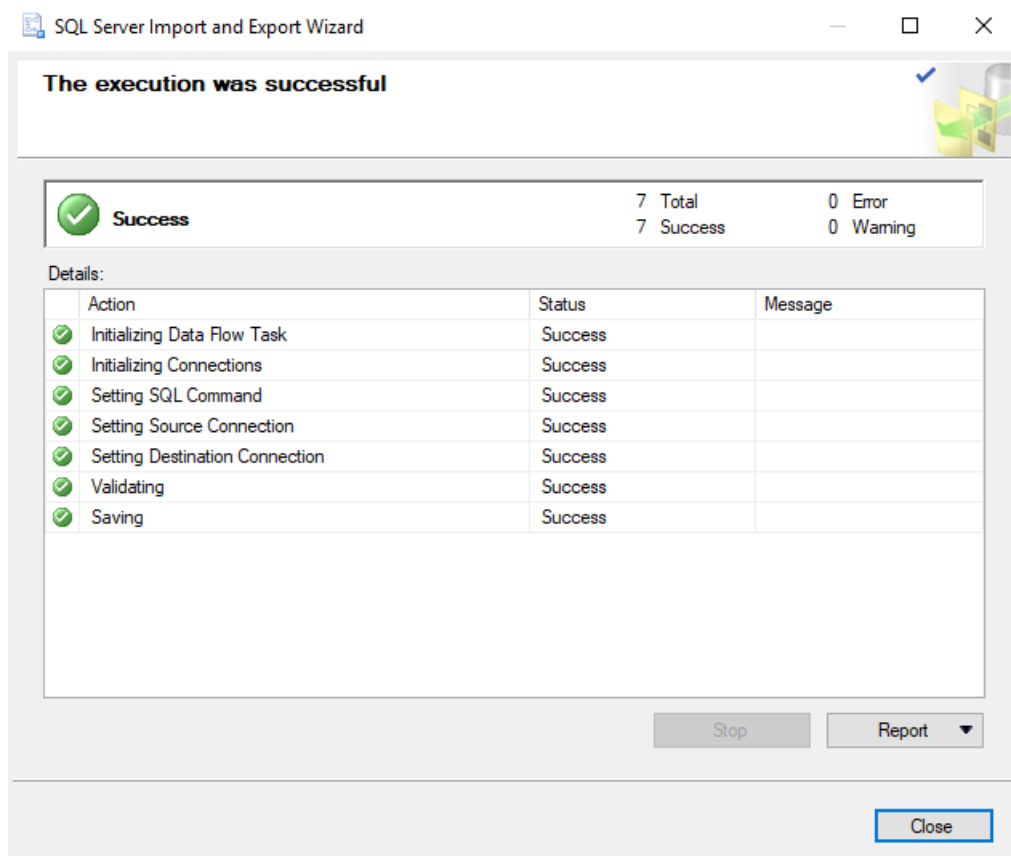
Na początku stworzyliśmy nowy projekt w Visual Studio



Następnie za pomocą narzędzia SSIS Import and Export Wizard importujemy nasze dane



Jak widzimy dane zostały poprawnie zaimportowane



- Dodaliśmy również opcję usuwania tabeli docelowej, jeśli ta znajduje się już w bazie

Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

General
Parameter Mapping
Result Set
Expressions

▼ General	
Name	Usowanie_tabeli_ludzie
Description	Execute SQL Task
▼ Options	
TimeOut	0
CodePage	1250
TypeConversionMode	Allowed
▼ Result Set	
ResultSet	None
▼ SQL Statement	
ConnectionType	OLE DB
Connection	DestinationConnectionOLEDB
SQLSourceType	Direct input
SQLStatement	drop table [dbo].[ludzie]
IsQueryStoredProcedure	False
BypassPrepare	True

BypassPrepare
Indicates whether the task should prepare the query before executing it.

Browse... Build Query... Parse Query

OK Cancel Help

- Oraz opcję czyszczenia tabeli, gdzie znajdują się potencjalne błędy

Execute SQL Task Editor

Configure the properties required to run SQL statements and stored procedures using the selected connection.

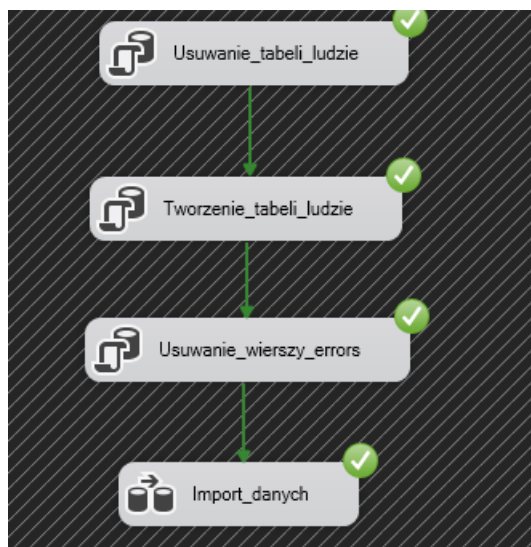
General
 Parameter Mapping
 Result Set
 Expressions

General	
Name	Usuwanie_wierszy_errors
Description	Execute SQL Task
Options	
TimeOut	0
CodePage	1250
TypeConversionMode	Allowed
Result Set	
ResultSet	None
SQL Statement	
ConnectionType	OLE DB
Connection	DestinationConnectionOLEDB
SQLSourceType	Direct input
SQLStatement	delete from errors
IsQueryStoredProcedure	False
BypassPrepare	True

Name
 Specifies the name of the task.

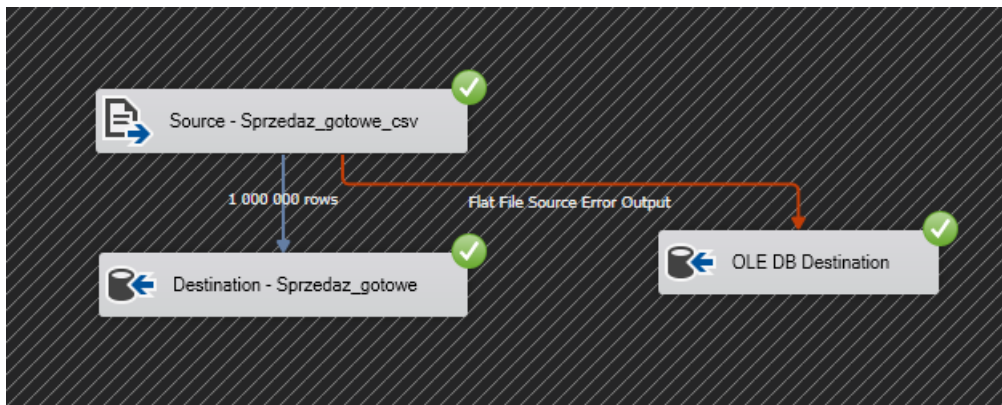
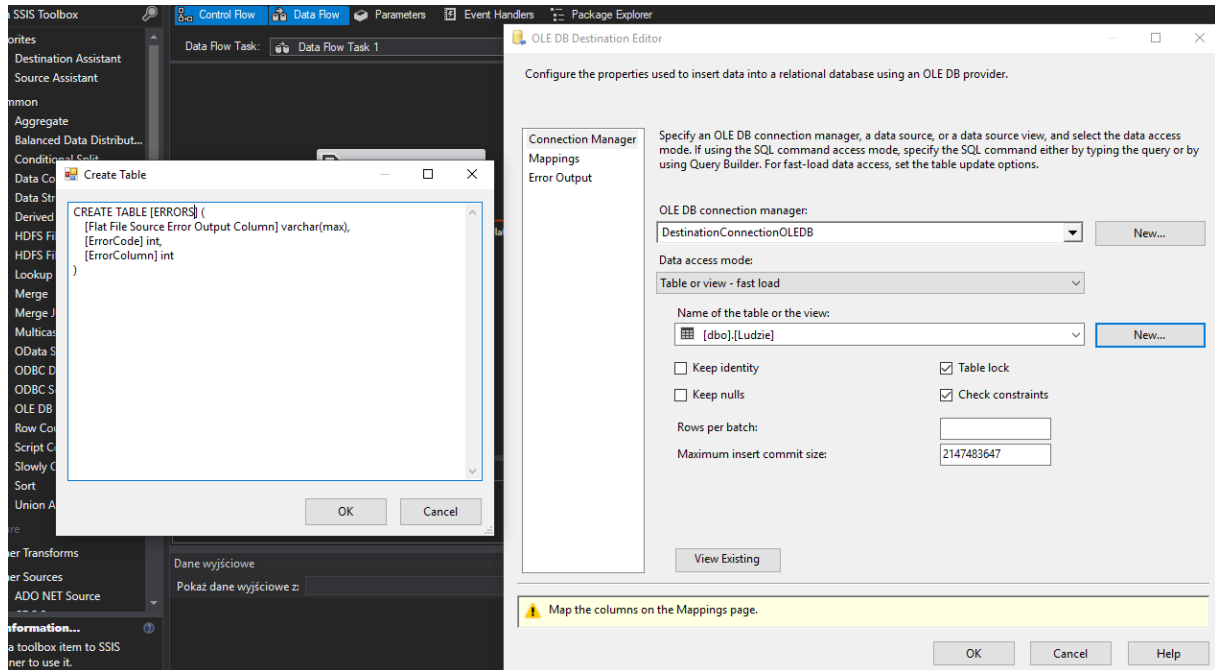
Browse... Build Query... Parse Query

OK Cancel Help



3.1. Obsługa błędów

Stworzyliśmy tabelę ERRORS gdzie zapisywane będą błędy, które wystąpią podczas importu danych do tabeli ludzie, oraz sprzedaż_gotowe



Dla tabel, które zawierają wiele wierszy ustawiliśmy parametr Rows per batch w OLE DB Destination na 10 000. Spowodowane jest to tym, że przy mniejszych wartościach Visual Studio może się zresetować w czasie wykonywania danego procesu.

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:
DestinationConnectionOLEDB New...

Data access mode:
Table or view - fast load

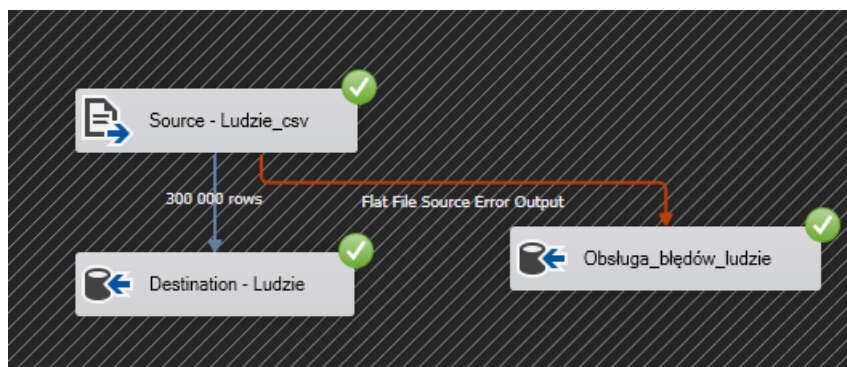
Name of the table or the view:
[ERRORS] New...

☐ Keep identity ☒ Table lock
☐ Keep nulls ☒ Check constraints

Rows per batch: 100000
Maximum insert commit size: 2147483647

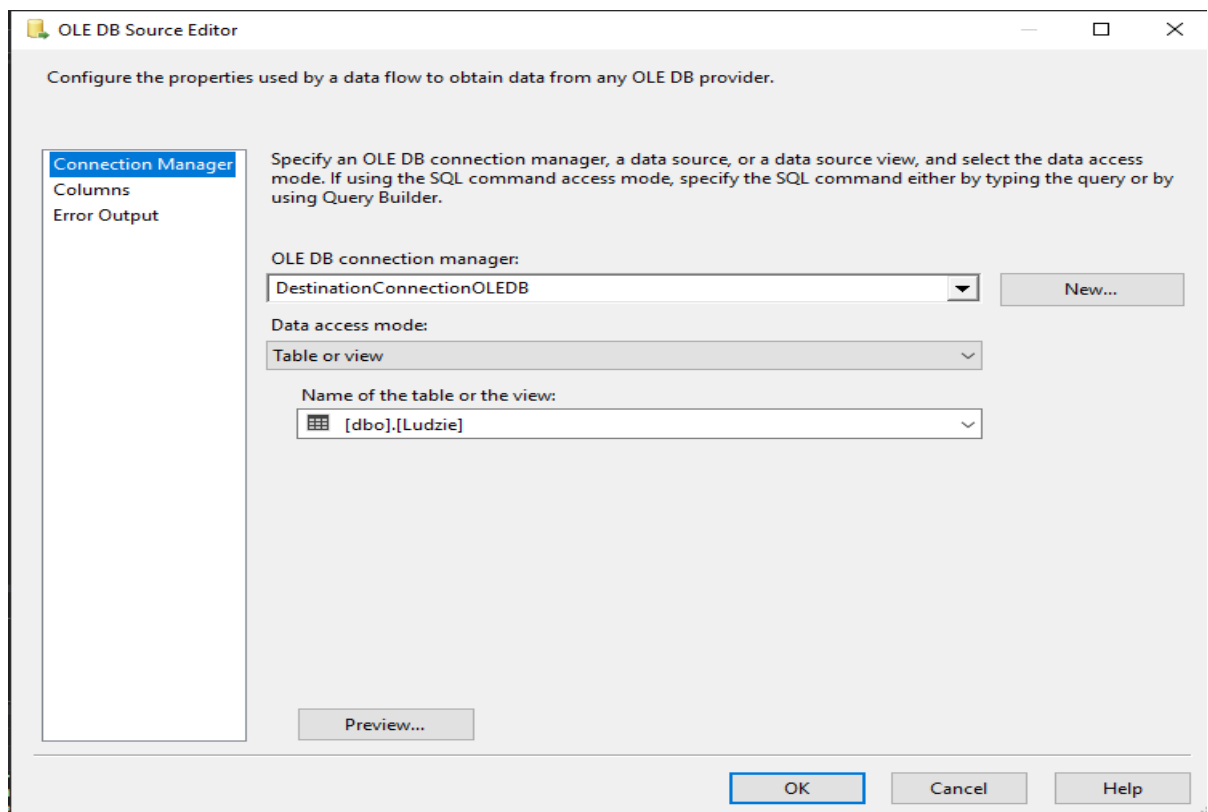
View Existing

OK Cancel Help



4. Oczyszczanie załadowanych danych

Po wczytaniu danych robimy proces oczyszczania z duplikatów, czyli mówiąc prościej usuwamy zduplikowane wiersze znajdujące się w naszej tabeli. Robimy to z użyciem narzędzia sort. Na Wejściu wrzucamy wszystkie dane z istniejącej tabeli, następnie w Sort ustawiamy Remove rows with duplicate sort value i zapisujemy. Robimy tak również dla drugiej tabeli.



Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass Throu...	▲
<input checked="" type="checkbox"/>	ID	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	First Name	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Last Name	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Gender	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	E Mail	<input type="checkbox"/>	▼

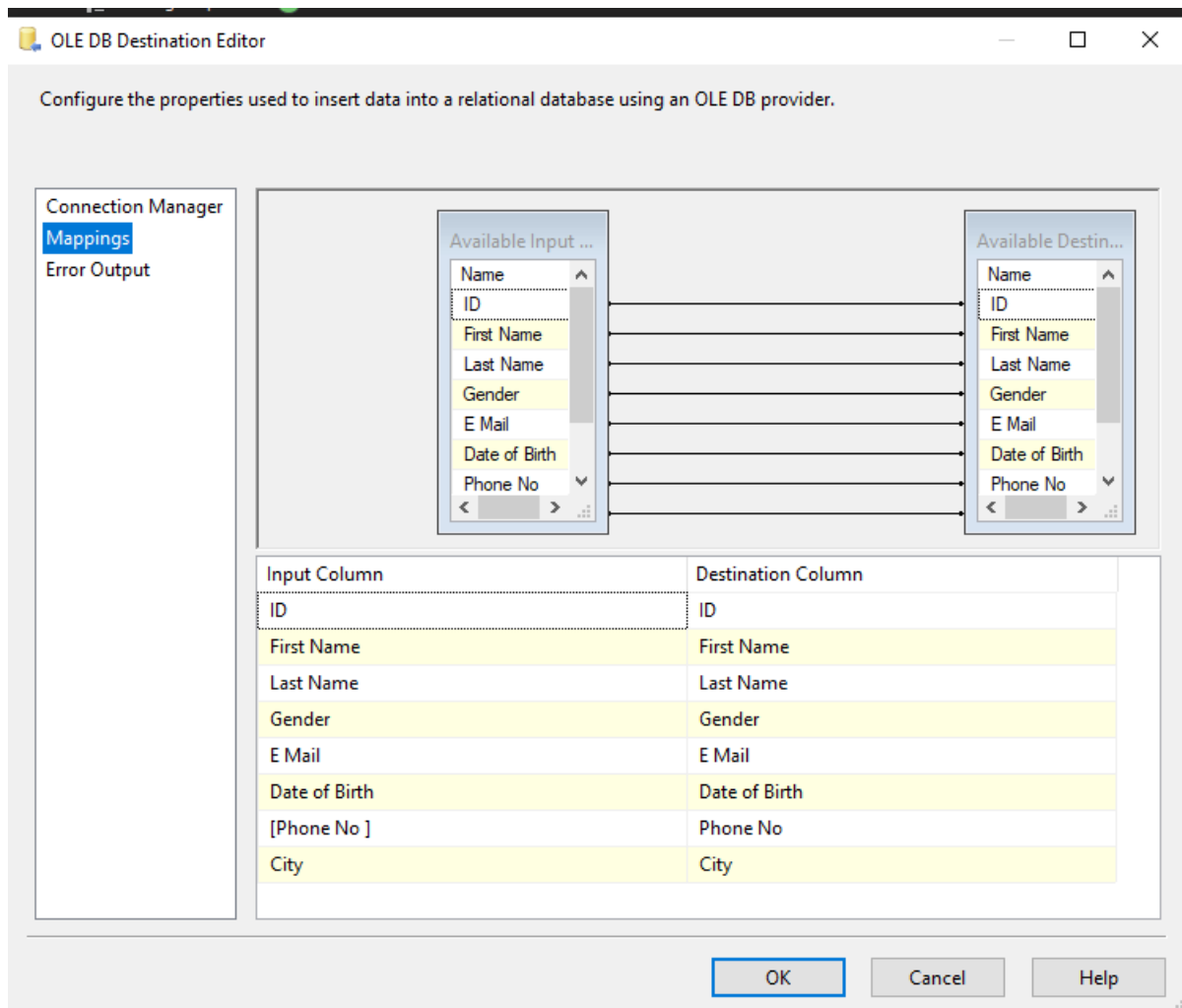
Input Column	Output Alias	Sort Type	Sort Order	C ▲
ID	ID	ascending	1	
First Name	First Name	ascending	2	
Last Name	Last Name	ascending	3	
Gender	Gender	ascending	4	
E Mail	E Mail	ascending	5	
Date of Birth	Date of Birth	ascending	6	
Phone No	Phone No	ascending	7	
City	City	ascending	8	

☒ Remove rows with duplicate sort values

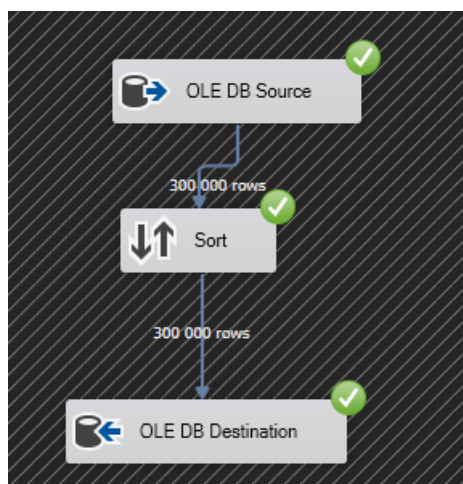
OK

Cancel

Help



Uruchamiamy i sprawdzamy poprawność działania



Jak widzimy wszystko wykonało się poprawnie to samo robimy dla drugiej tabeli

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:

DESKTOP-7S5C7E7.projekt ▼

New...

Data access mode:

Table or view ▼

Name of the table or the view:

[dbo].[Sprzedaz_gotowe] ▼

Preview...

OK

Cancel

Help

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns

<input checked="" type="checkbox"/>	Name	Pass Throu...	^
<input checked="" type="checkbox"/>	Order ID	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Region	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Country	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Item Type	<input type="checkbox"/>	
<input checked="" type="checkbox"/>	Sales Channel	<input type="checkbox"/>	v

Input Column	Output Alias	Sort Type	Sort Order	
Order ID	Order ID	ascending	1	
Region	Region	ascending	2	
Country	Country	ascending	3	
Item Type	Item Type	ascending	4	
Sales Channel	Sales Channel	ascending	5	
Order Priority	Order Priority	ascending	6	
Order Date	Order Date	ascending	7	
Units Sold	Units Sold	ascending	8	
Unit Price	Unit Price	ascending	9	

☒ Remove rows with duplicate sort values

OK

Cancel

Help

OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Connection Manager
Mappings
 Error Output

Available Input Columns

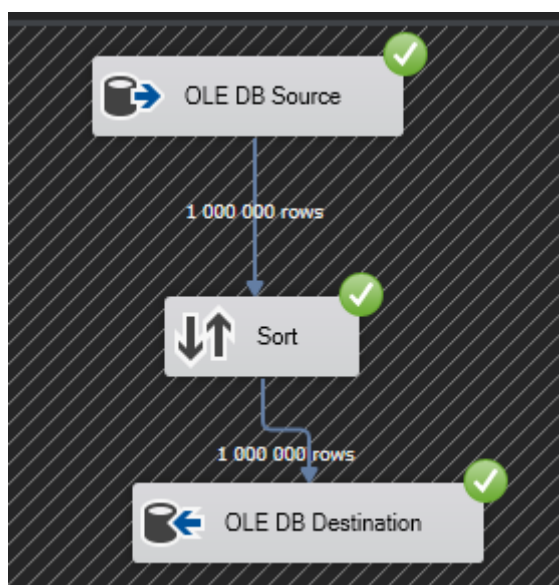
Name
Order ID
Region
Country
Item Type
Sales Channel
Order Priority
Order Date

Available Destination Columns

Name
Order ID
Region
Country
Item Type
Sales Channel
Order Priority
Order Date

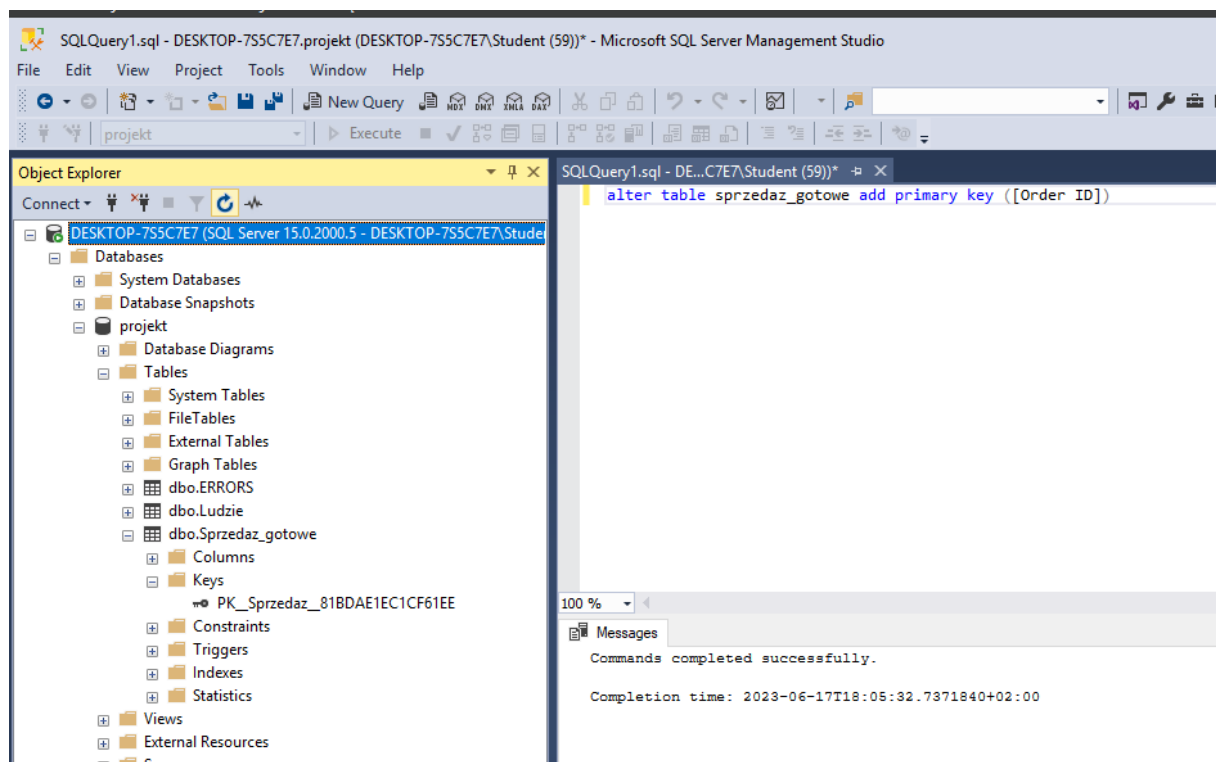
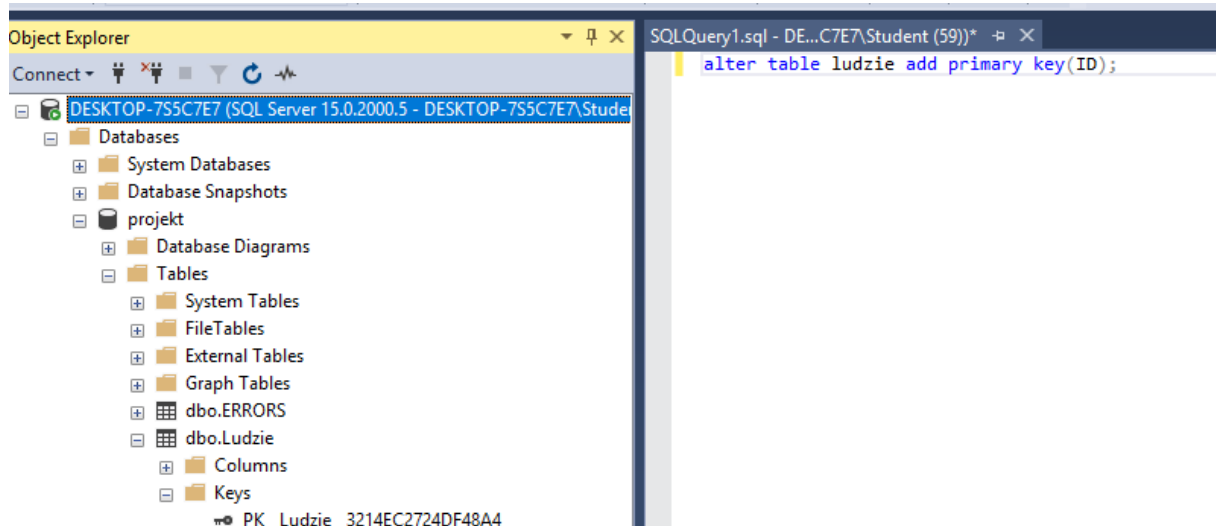
Input Column	Destination Column
Order ID	Order ID
Region	Region
Country	Country
Item Type	Item Type
Sales Channel	Sales Channel
Order Priority	Order Priority
Order Date	Order Date
Units Sold	Units Sold
Unit Price	Unit Price

OK Cancel Help

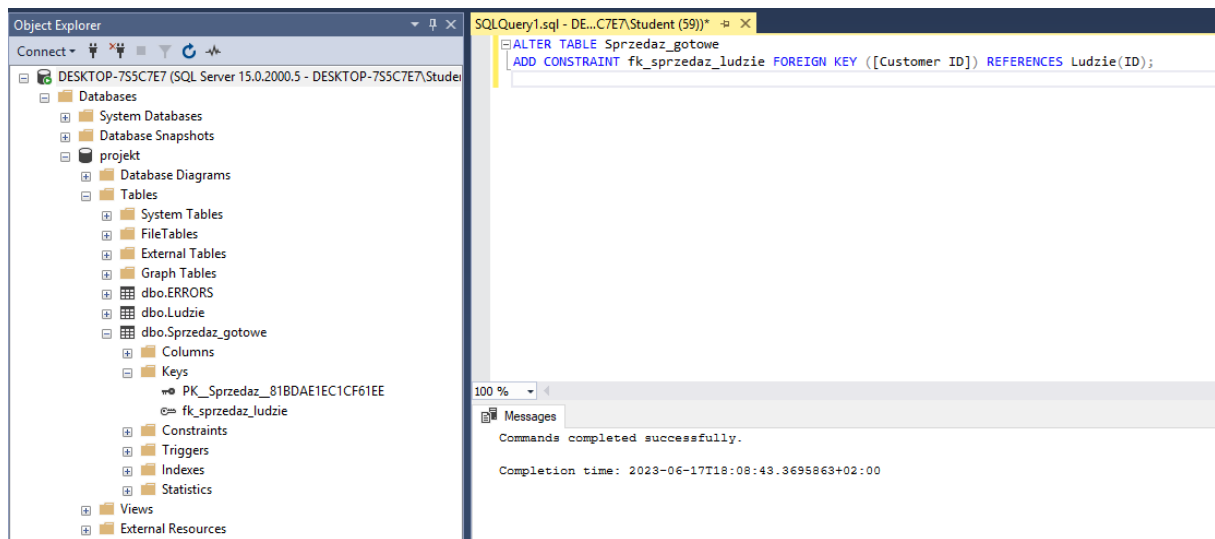


5. Tworzenie kluczy głównych i kluczy obcych

Do tabel, które stworzyliśmy dodajemy klucze główne i klucze obce w celu połączenia obu tabel. Wykonujemy to za pomocą następujących poleceń:



Po dodaniu kluczy głównych przystępuje do dodania klucza obcego

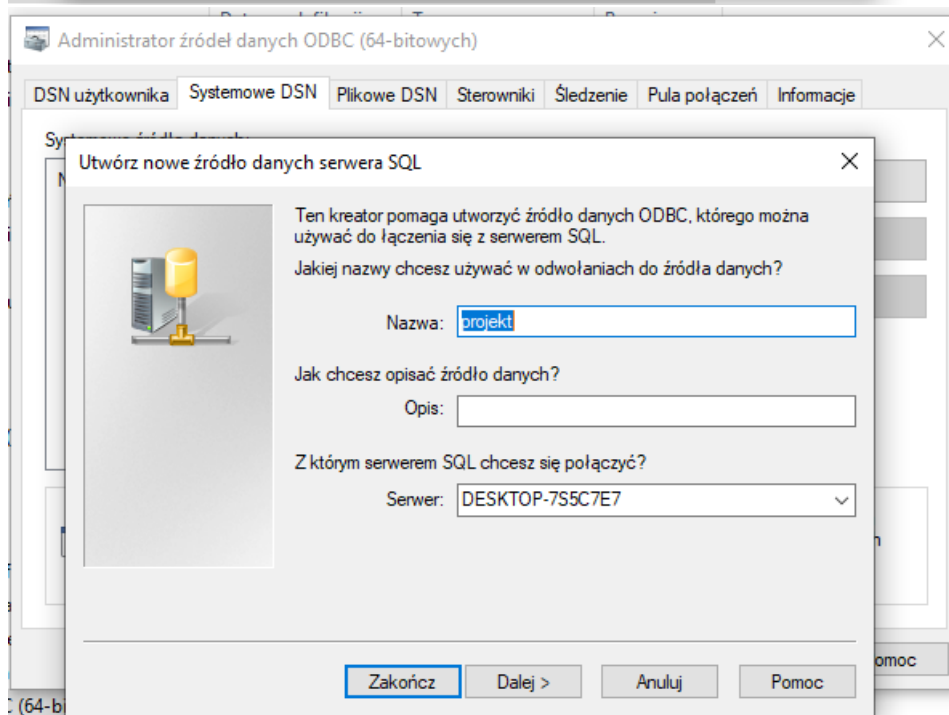
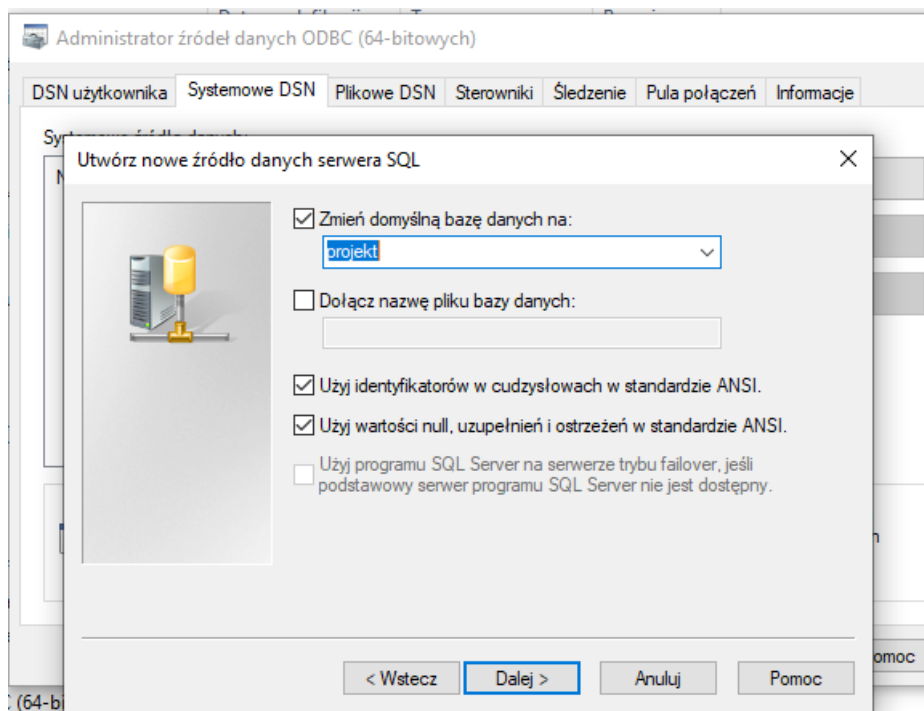


Jak widzimy wszystkie klucze zostały poprawnie dodane

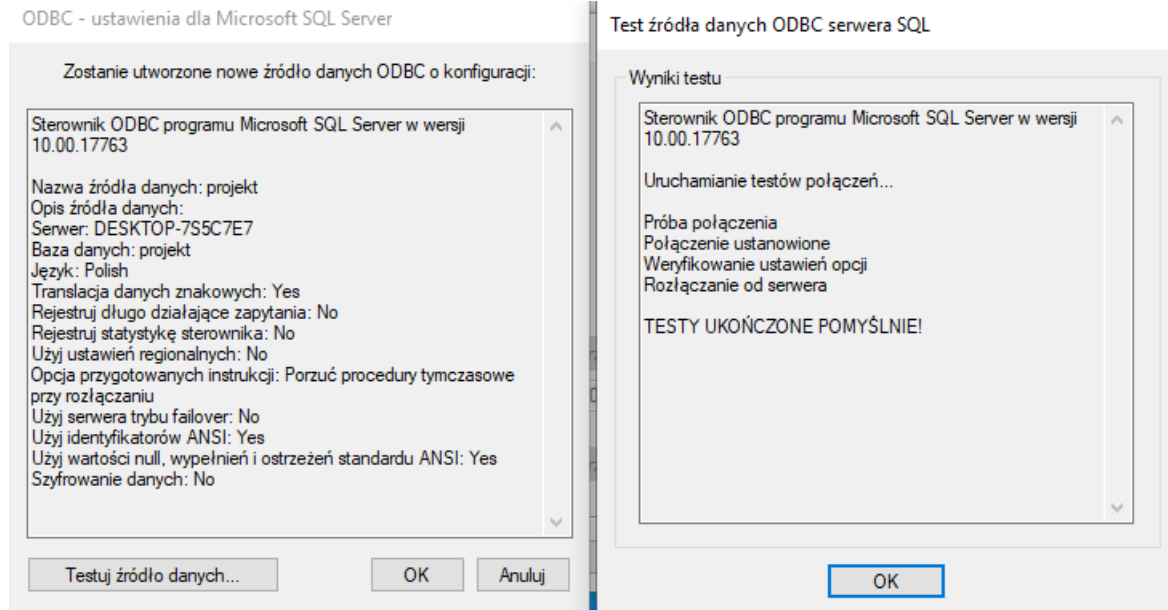
6. Połączenie naszej bazy ze środowiskiem Tableau

6.1. Stworzenie źródła danych ODBC

W celu połączenia Tableau z naszą bazą danych na początku naszym zadaniem jest stworzenie źródła danych, robimy to za pomocą Administratora źródeł danych ODBC. Narzędzie to znajduje się w panelu sterowania -> w narzędziach administracyjnych. Po otwarciu tego narzędzia klikamy Systemowe DNS i dodajemy nowe źródło danych.



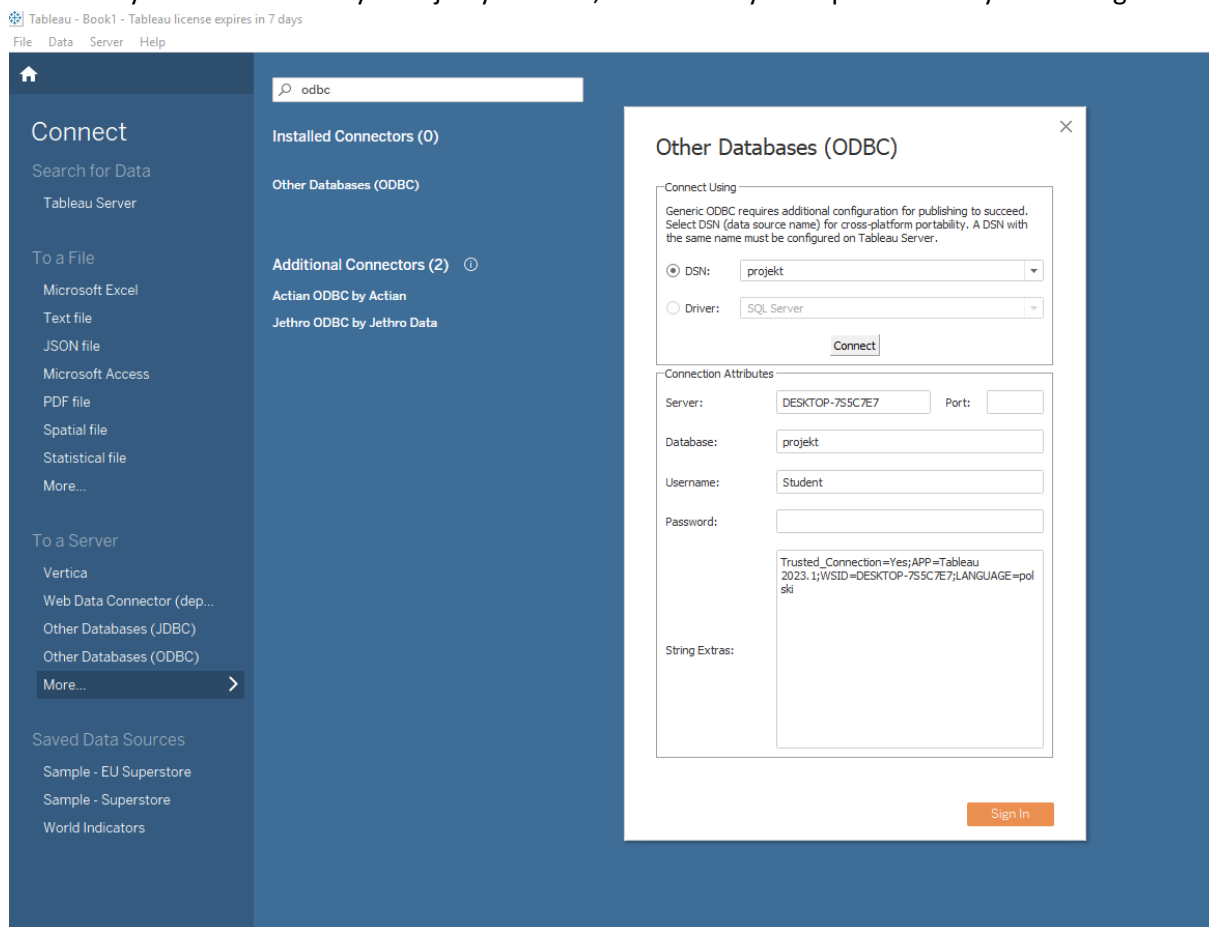
Po utworzeniu nowego źródła danych wykonujemy jego test.



Jak widzimy powyżej test został wykonany pomyślnie oznacza to, że poprawnie stworzyliśmy źródło naszych danych.

6.2. Połączenie Tableau z bazą danych

Dzięki źródle danych, które utworzyliśmy połączymy teraz naszą hurtownię danych z tableau. Otwieramy tableau i wyszukujemy ODBC, otwieramy i przechodzimy do logowania.



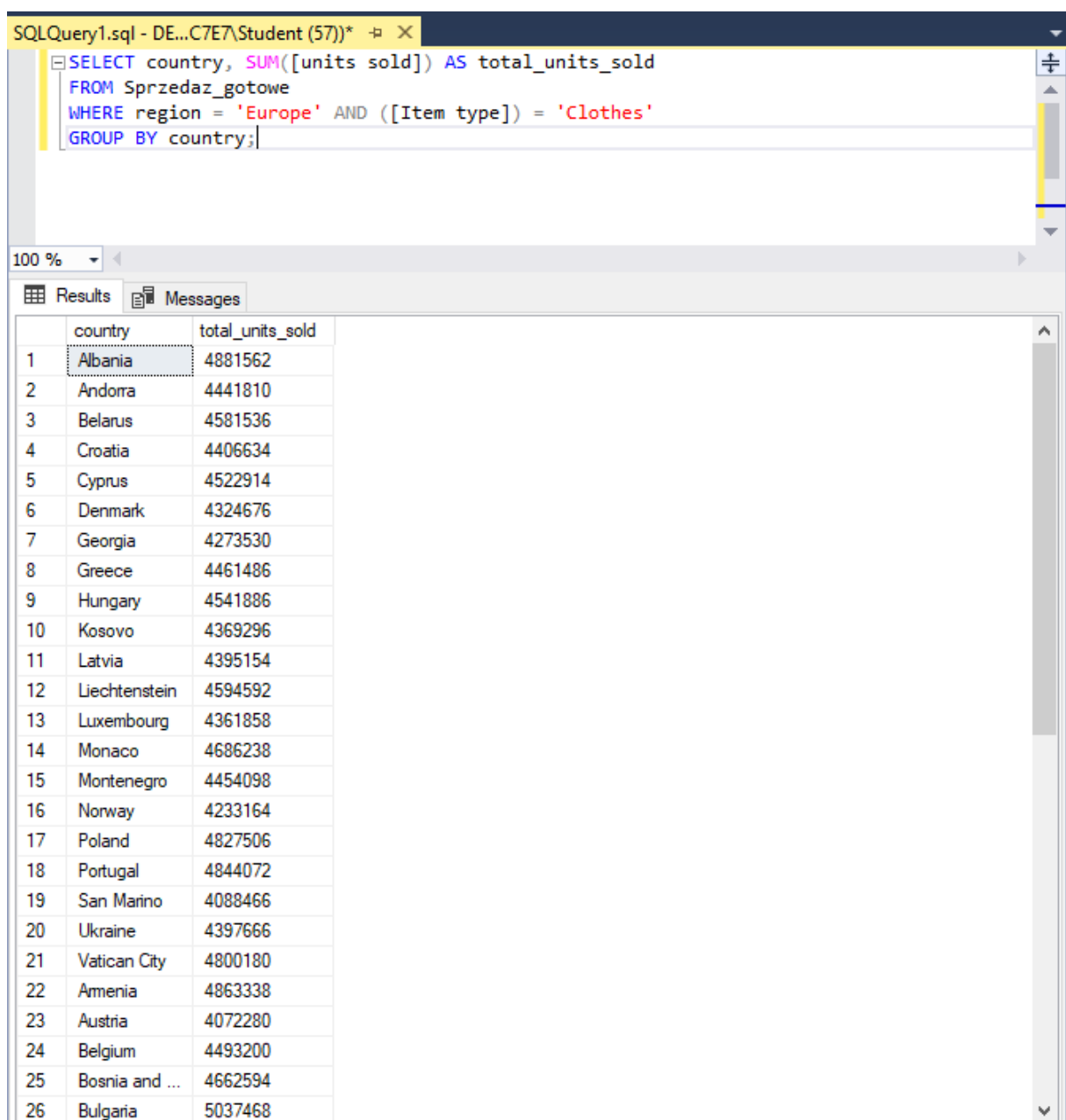
Wybieramy utworzone źródło danych, w polu DNS klikamy connect a następnie Sing In. Po tym następuje połączenie źródła danych ze środowiskiem tableau.

7. Integracja danych ETL

Do tego zadania wykorzystamy Visual Studio, aby stworzyć procesy ETL. Mają one na celu wyciągnięcie z bazy interesujących nas kolumn. Polecenie, które wpisaliliśmy do Ole DB Source sprawdziliśmy najpierw w SQL Management Studio.

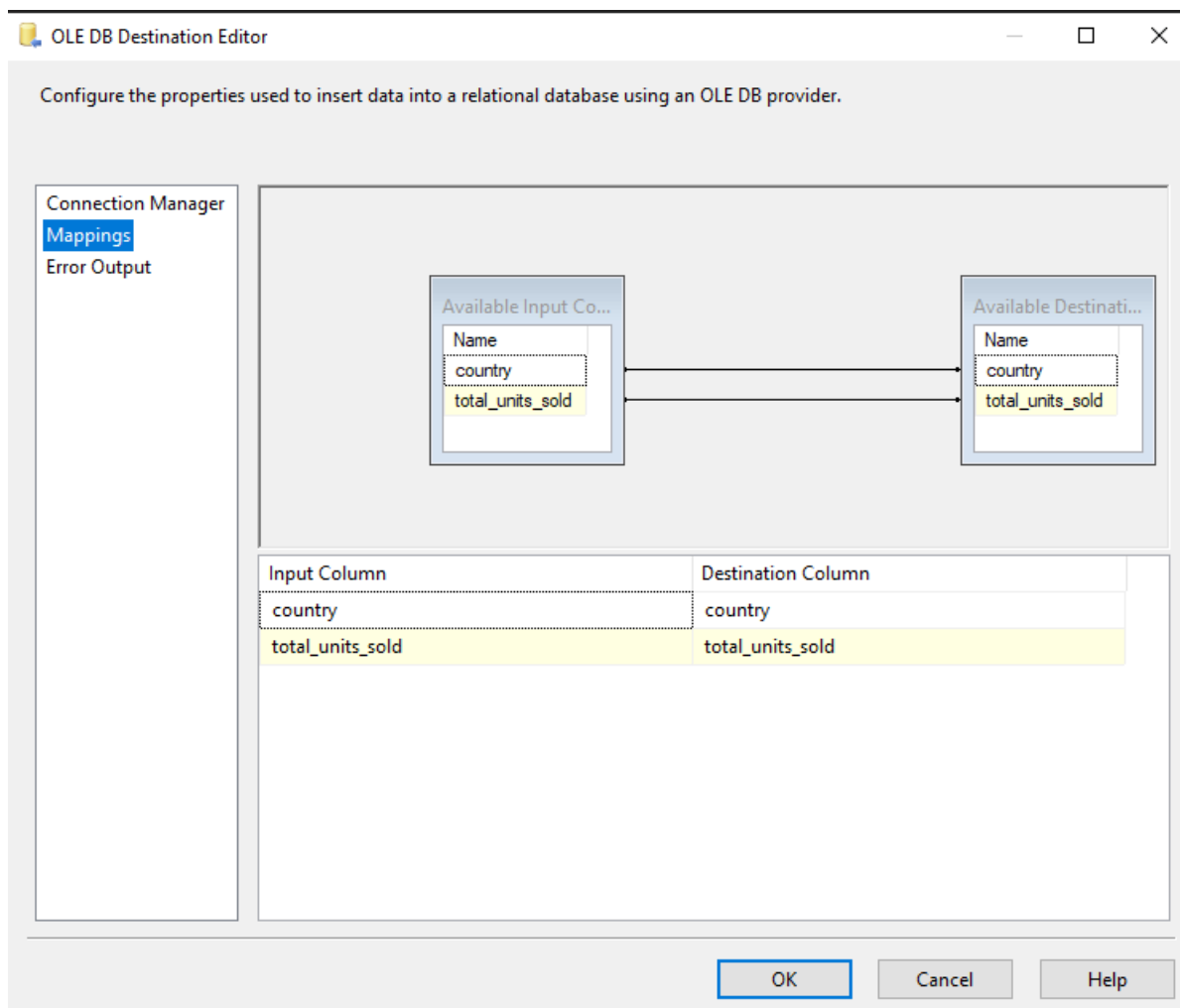
7.1. Ilość sprzedanych jednostek odzieży w krajach europejskich.

Jak widzimy poniżej zapytanie wykonuje się poprawnie, więc możemy przejść do tworzenia ETL

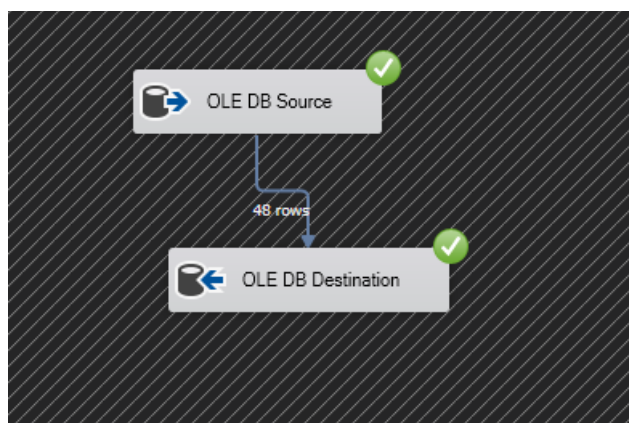


The screenshot shows the SQL Server Enterprise Manager interface. The top pane displays a SQL query in the 'SQLQuery1.sql' file. The query is: `SELECT country, SUM([units sold]) AS total_units_sold FROM Sprzedaz_gotowe WHERE region = 'Europe' AND ([Item type]) = 'Clothes' GROUP BY country;`. The bottom pane shows the 'Results' tab with a table of 26 rows. The first row is highlighted. The table has two columns: 'country' and 'total_units_sold'.

	country	total_units_sold
1	Albania	4881562
2	Andorra	4441810
3	Belarus	4581536
4	Croatia	4406634
5	Cyprus	4522914
6	Denmark	4324676
7	Georgia	4273530
8	Greece	4461486
9	Hungary	4541886
10	Kosovo	4369296
11	Latvia	4395154
12	Liechtenstein	4594592
13	Luxembourg	4361858
14	Monaco	4686238
15	Montenegro	4454098
16	Norway	4233164
17	Poland	4827506
18	Portugal	4844072
19	San Marino	4088466
20	Ukraine	4397666
21	Vatican City	4800180
22	Armenia	4863338
23	Austria	4072280
24	Belgium	4493200
25	Bosnia and ...	4662594
26	Bulgaria	5037468



Jak widzimy wszystko wykonało się poprawnie



7.2. Ilość wystąpień zamówień online i offline dla każdego państwa Azji

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
DESKTOP-755C7E7.projekt

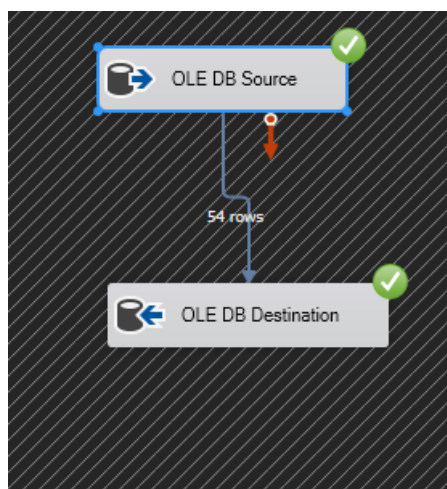
Data access mode:
SQL command

SQL command text:
SELECT Region, [Sales Channel], Country, count(*) as Count
FROM Sprzedaz_gotowe
WHERE Region = 'Asia' AND ([Sales Channel] = 'Offline' OR [Sales
Channel] = 'Online')
GROUP BY Region, [Sales Channel], Country;

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help



7.3. Ilość wystąpień danego sklepu wraz ze sprzedawanym towarem oraz zarobkiem całkowity, który z niego płynie

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
DESKTOP-755C7E7.projekt

Data access mode:
SQL command

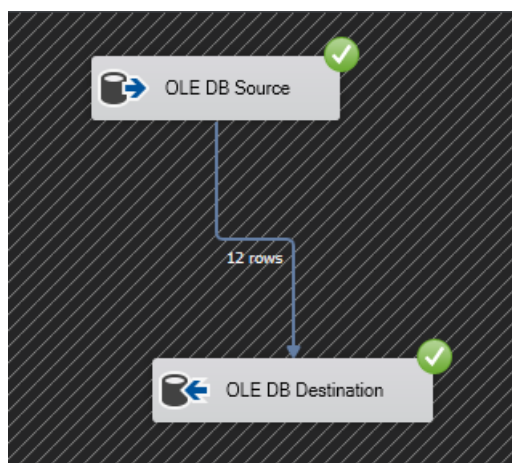
SQL command text:

```
SELECT  
[place of delivery] AS sklep,  
[item type],  
COUNT(*) AS liczba_wystapien,  
ROUND(SUM([total revenue]), 2) AS zarobek  
FROM  
sprzedaz_gotowe  
WHERE  
country = 'Poland'  
GROUP BY  
[place of delivery],  
[item type]  
ORDER BY  
[place of delivery],  
[item type];
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help



7.4. Ilość osób w określonym przedziale wiekowym z podziałem na płeć

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
Columns
Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
DESKTOP-755C7E7.projekt

Data access mode:
SQL command

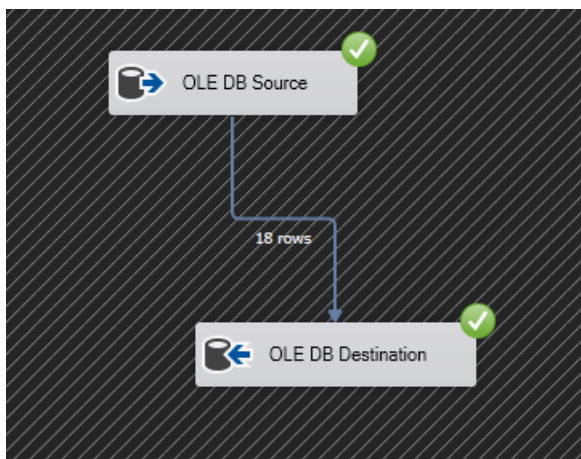
SQL command text:

```
SELECT
gender,
age_group,
COUNT(DISTINCT city) AS city_count
FROM
(
SELECT
CASE
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 15 AND 20 THEN '15-20'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 21 AND 25 THEN '21-25'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 26 AND 30 THEN '26-30'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 31 AND 35 THEN '31-35'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 36 AND 40 THEN '36-40'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 41 AND 45 THEN '41-45'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 46 AND 50 THEN '46-50'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 51 AND 55 THEN '51-55'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 56 AND 60 THEN '56-60'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 61 AND 65 THEN '61-65'
WHEN YEAR(CURRENT_TIMESTAMP) - YEAR(CAST([date of birth] AS date)) BETWEEN 66 AND 70 THEN '66-70'
ELSE 'N/A'
END AS age_group,
gender,
city
FROM Iudzie
) AS subquery
GROUP BY gender, age_group
ORDER BY gender, MIN(CAST(SUBSTRING(age_group, 1, 2) AS INT));
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help



8. Wizualizacje

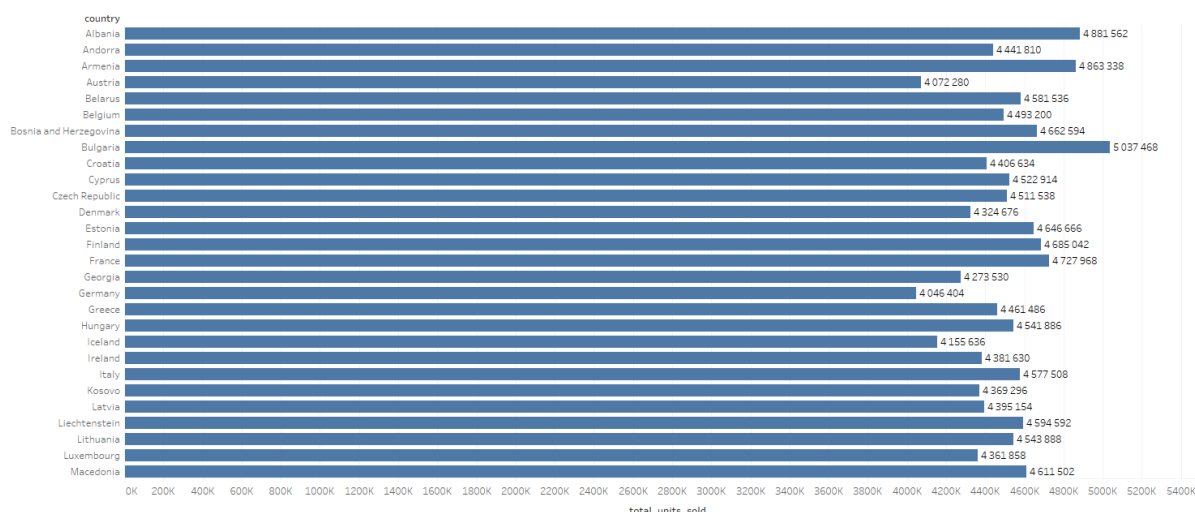
8.1. Ilość sprzedanych jednostek odzieży w krajach europejskich.

Analizując dane z wykresu, można dostrzec ciekawą dynamikę sprzedaży odzieży w różnych krajach Europy. Obserwujemy tu interesujący kontrast pomiędzy państwami, gdzie zróżnicowanie wolumenu sprzedaży jest dość znaczne.

Bułgaria przoduje w tym zestawieniu, demonstrując imponującą liczbę sprzedanych ubrań. Jest to prawdopodobnie efekt połączenia różnych czynników - od silnej lokalnej produkcji odzieży, poprzez skuteczne strategie marketingowe, aż po możliwe preferencje konsumentów. Bez względu na powody, Bułgaria zdecydowanie wyróżnia się na tle innych krajów w zakresie sprzedaży odzieży.

Z drugiej strony, mamy Niemcy, które, mimo swojej gospodarczej siły i dużej populacji, znajdują się na końcu tego zestawienia. Może to wynikać z różnych czynników, takich jak różnice w stylach i preferencjach zakupowych, różnorodność dostępnych na rynku marek, a nawet wpływ klimatu na rodzaj i ilość zakupionej odzieży.

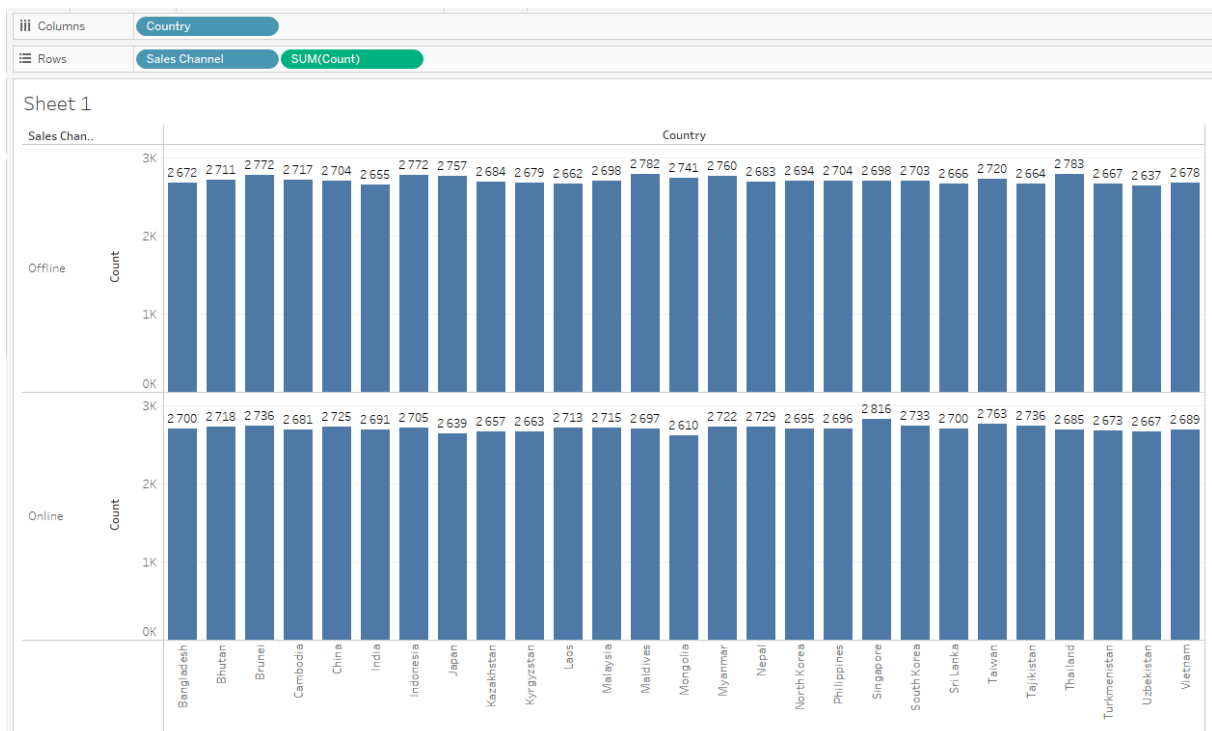
Sheet 1



8.2. Ilość wystąpień zamówień online i offline dla każdego państwa Azji

Analizując dane dotyczące zamówień z różnych krajów Azji, zauważamy, że kanały sprzedaży są równomiernie rozłożone między opcjami online i offline. Wygląda na to, że klienci z różnych krajów Azji składają zamówienia z równą częstotliwością za pośrednictwem obu tych kanałów.

Dla każdego analizowanego kraju, rozkład zamówień między kanałami online i offline jest zaskakująco zrównoważony, z wartościami oscylującymi wokół połowy wszystkich zamówień. W związku z tym, niezależnie od kraju, oba kanały sprzedaży są równie popularne.



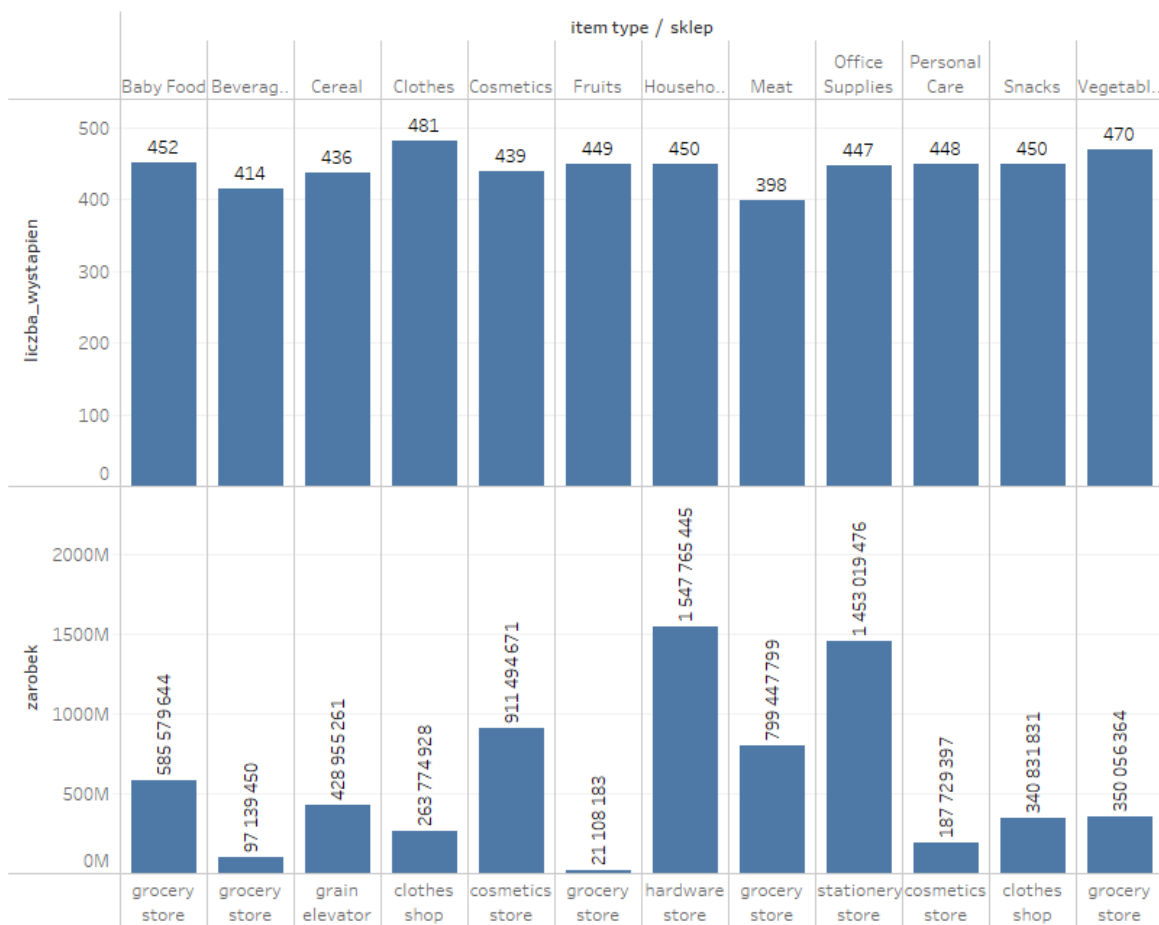
8.3. Ilość wystąpień danego sklepu wraz ze sprzedawanym towarem oraz zarobkiem całkowity, który z niego płynie

Zajrzeliśmy głębiej w strukturę sprzedaży różnych rodzajów artykułów w sklepach, próbując zrozumieć, które z nich przynoszą największe zyski. Wyniki tej analizy zostały zaprezentowane na poniższym wykresie, który ilustruje dystrybucję asortymentu, jak i generowane zyski.

Wykres wyraźnie pokazuje, że sprzęt elektroniczny stanowi najbardziej dochodową kategorię produktów. Jego sprzedaż przynosi największe zyski, co może wynikać z wysokich marż na tego typu produkty oraz stale rosnącego popytu na najnowsze technologie. To świadczy o niezaprzeczalnej roli, jaką sprzęt elektroniczny odgrywa w współczesnym świecie konsumpcji, będąc nieodłącznym elementem naszego codziennego życia.

Z drugiej strony, mamy owoce, które, mimo że są podstawowym elementem naszej diety, generują najmniejsze zyski. Może to wynikać z niskich cen tych produktów oraz małych marż. Chociaż owoce mogą nie przynosić tak wysokich zysków jak inne kategorie produktów, ich stała obecność na półkach sklepowych jest kluczowa dla zrównoważonego asortymentu i zaspokajania potrzeb wszystkich klientów.

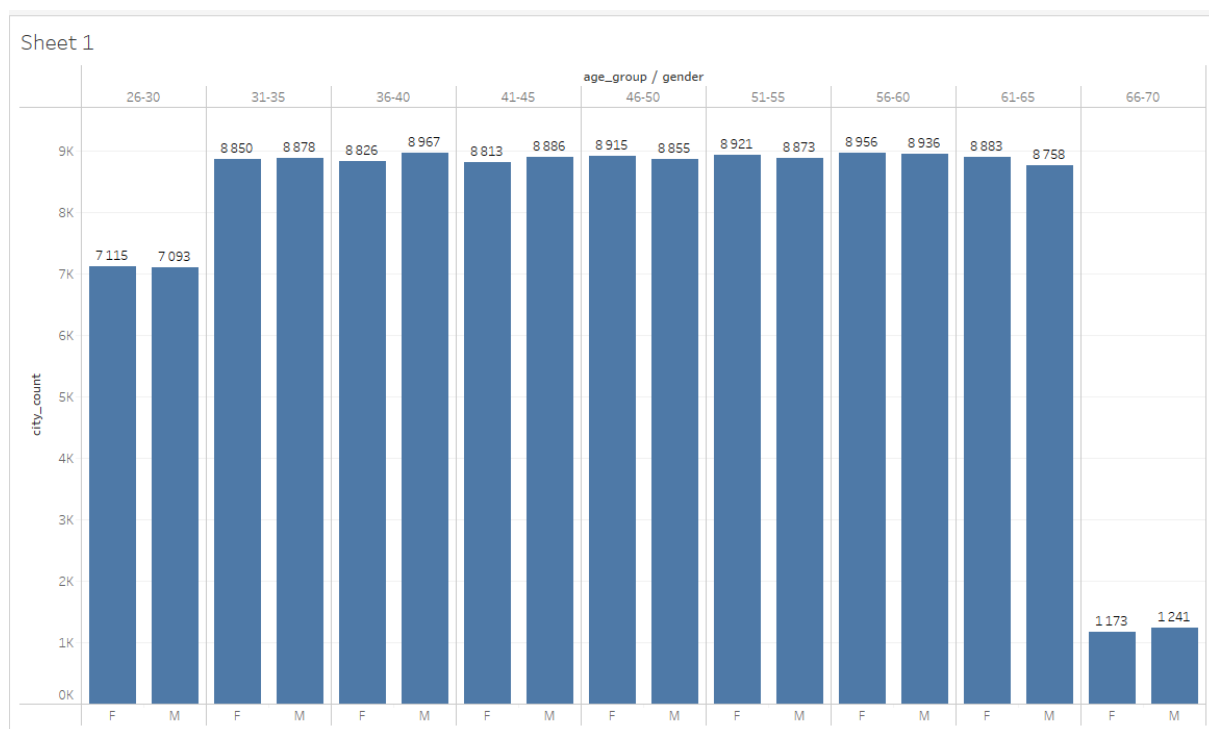
Sheet 1



8.4. Ilość osób w określonym przedziale wiekowym z podziałem na płeć

Wykres, który analizujemy, prezentuje zbalansowany rozkład popytu na produkty w bazie danych w różnych grupach wiekowych, biorąc pod uwagę płeć klientów. Widać, że zarówno kobiety, jak i mężczyźni z poszczególnych grup wiekowych wykazują podobne zainteresowanie produktami.

Ta analiza podkreśla uniwersalność i szeroki zasięg produktów dostępnych w bazie danych. Niezależnie od płci czy wieku, oferowane produkty znajdują swoich odbiorców, co przekłada się na równomierne rozłożenie zamówień.



9. Wnioski

Przeanalizowaliśmy dane dotyczące sprzedaży, korzystając z różnorodnych narzędzi, takich jak MS SQL Server, MS Visual Studio i Tableau. Dzięki temu, zdobyliśmy głębsze zrozumienie trendów, wzorców i relacji występujących w danych sprzedażowych. Procesy ETL umożliwiły nam skuteczną transformację i ładowanie danych do naszej bazy danych, zapewniając spójny i niezwykle precyzyjny zbiór danych do dalszej analizy.

Wykorzystując funkcje wizualizacyjne w Tableau, mogliśmy kreować interaktywne i dynamiczne wizualizacje, które pozwoliły nam identyfikować kluczowe trendy i zależności pomiędzy różnymi czynnikami. Analiza danych sprzedażowych pozwoliła nam bacznie przyjrzeć się takim elementom jak region czy kategoria produktu.

Byliśmy w stanie zidentyfikować największą sprzedaż poszczególnych produktów dla różnych państw czy chociażby to, czy proporcja kupców płci żeńskiej znacząco różni się od kupców płci męskiej.

Podsumowując, analiza danych sprzedażowych, przeprowadzona przy użyciu różnorodnych narzędzi, pozwoliła nam na lepsze zrozumienie danych. Te informacje są niezwykle cenne dla przedsiębiorstw, pomagając w podejmowaniu trafnych decyzji biznesowych, opracowaniu strategii sprzedażowych i osiągnięciu sukcesu na konkurencyjnym rynku.