



---

# HURTOWNIE DANYCH

---

Lista 6 – Hierarchie, partycje, prosta analiza danych



ALEKSANDER STEPANIUK

NR. INDEKSU: 272644

Politechnika Wrocławska, Informatyka Stosowana

## Rozwiązania:

Zadanie 1.

a)

The screenshot shows the 'Create Named Calculation' dialog box. It has three main input fields: 'Column name:' with the value 'Names', 'Description:' which is empty, and 'Expression:' with the value 'FirstName + ' ' + LastName'. At the bottom, there are three buttons: 'OK', 'Cancel', and 'Help'.

The screenshot shows the SQL Server Data Tools interface with three panes:

- Attributes:** A list of attributes for the 'DIM SALESPERSON' table. The attributes are: Country Region Code, Gender, Group, Names (highlighted), Sales Person ID, and Title.
- Hierarchies:** A list of hierarchies for the 'DIM SALESPERSON' table. The hierarchies are: Hierarchy (checked), Group, Country Region Code, Names, and <new level>. A note on the right says: 'To create a new hierarchy, drag an attribute here.'
- Data Source View:** A list of data source views for the 'DIM SALESPERSON' table. The data source views are: SalesPersonID, FirstName, LastName, Title, Gender, CountryRegionCode, Group, and Names (highlighted).

Attributes

⌵ DIM\_CUSTOMER

City

Country Region Code

Customer ID

Group

Names

Territory Name

Title

Hierarchies

⚠ Hierarchy

▫ Group

▫ Country Region Code

▫ Names

<new level>

To create a new hierarchy, drag an attribute here.

Data Source View

⌵ DIM\_CUSTOMER

CustomerID

FirstName

LastName

Title

City

TerritoryName

CountryRegionCode

Group

Names

Attributes

⌵ DIM\_PRODUCT

Category Name

Color

List Price

Name

Product ID

Sub Category Name

Hierarchies

⚠ Hierarchy

▫ Category Name

▫ Sub Category Name

▫ Name

<new level>

To create a new hierarchy, drag an attribute here.

Data Source View

⌵ DIM\_PRODUCT

ProductID

Name

ListPrice

Color

SubCategoryName

CategoryName

Weight

Size

IsPurchased

Attributes	Hierarchies	Data Source View
DIM TIME <ul style="list-style-type: none"> <li> Day Of Month</li> <li> Month</li> <li> Month Name</li> <li> PK TIME</li> <li> Quarter</li> <li> Weekday Name</li> <li> Year</li> </ul>	Hierarchy <ul style="list-style-type: none"> <li>Year</li> <li>Quarter</li> <li>Month</li> <li>Day Of Month</li> <li>&lt;new level&gt;</li> </ul> <p>To create a new hierarchy, drag an attribute here.</p>	DIM_TIME <ul style="list-style-type: none"> <li> PK_TIME</li> <li>Year</li> <li>Quarter</li> <li>Month</li> <li>MonthName</li> <li>WeekdayName</li> <li>DayOfMonth</li> </ul>

b)

<b>Source</b>	
CustomRollupColumn	(none)
CustomRollupPropertiesColumn	(none)
KeyColumns	DIM_SALESPERSON.SalesPersonID (Integer)
NameColumn	<b>DIM_SALESPERSON.Names (WChar)</b>
ValueColumn	(none)

<b>Source</b>	
CustomRollupColumn	(none)
CustomRollupPropertiesColumn	(none)
KeyColumns	DIM_CUSTOMER.CustomerID (Integer)
NameColumn	<b>DIM_CUSTOMER.Names (WChar)</b>
ValueColumn	(none)

<b>Source</b>	
CustomRollupColumn	(none)
CustomRollupPropertiesColumn	(none)
KeyColumns	DIM_PRODUCT.ProductID (Integer)
NameColumn	<b>DIM_PRODUCT.Name (WChar)</b>
ValueColumn	(none)

c)

**New Measure**

Usage: Distinct count

Source table: FACT\_SALES

Source column:

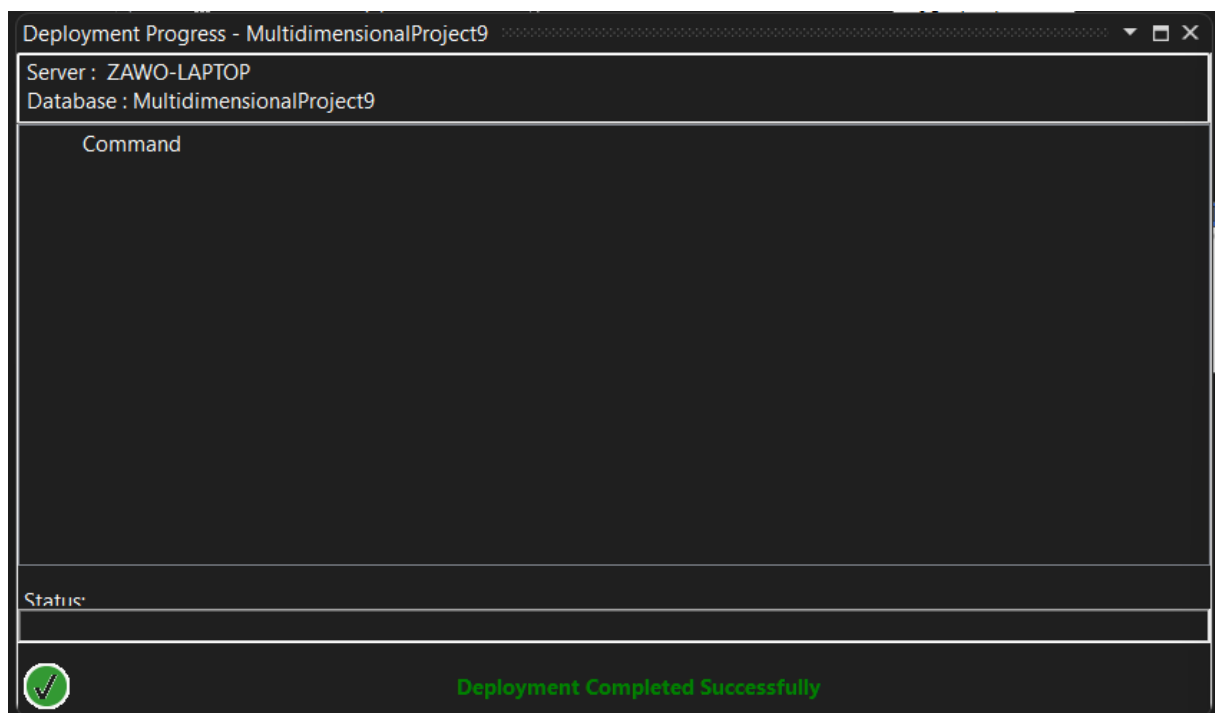
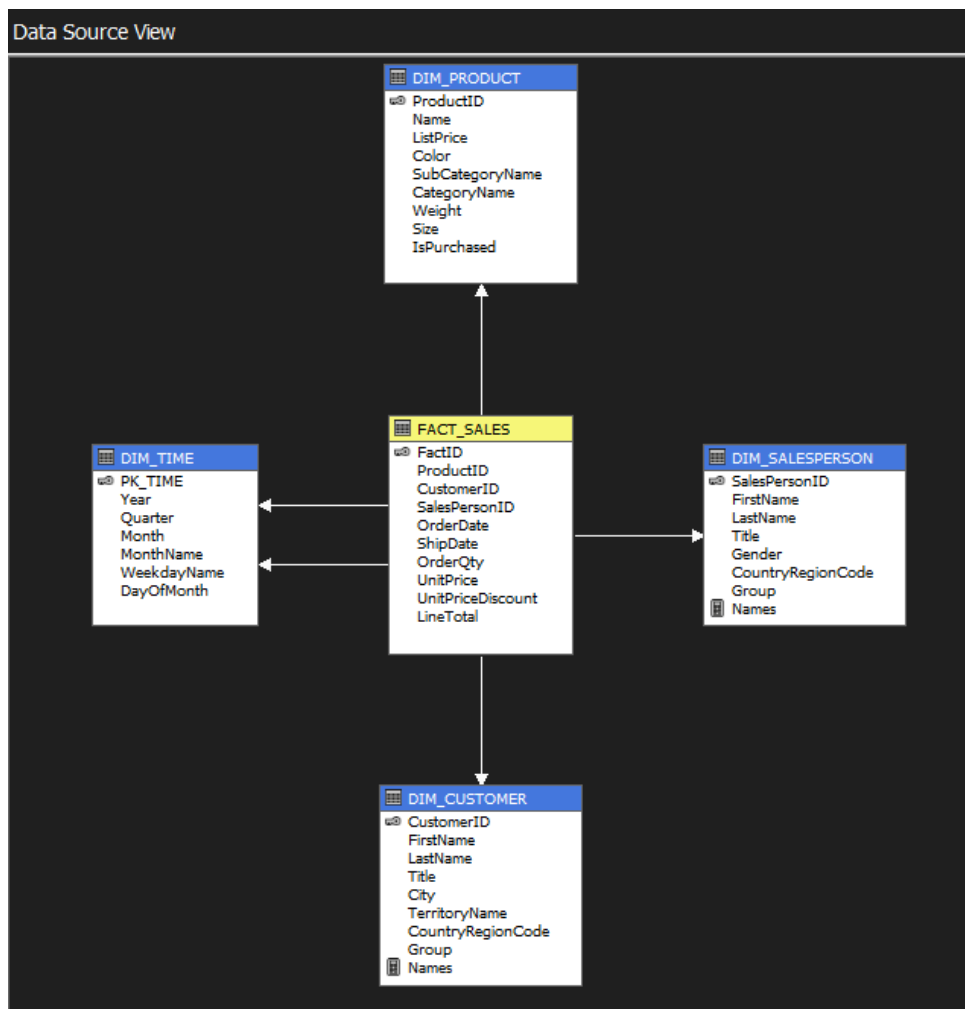
- FactID
- ProductID
- CustomerID**
- SalesPersonID
- OrderDate
- ShipDate
- OrderQty
- UnitPrice
- UnitPriceDiscount
- LineTotal

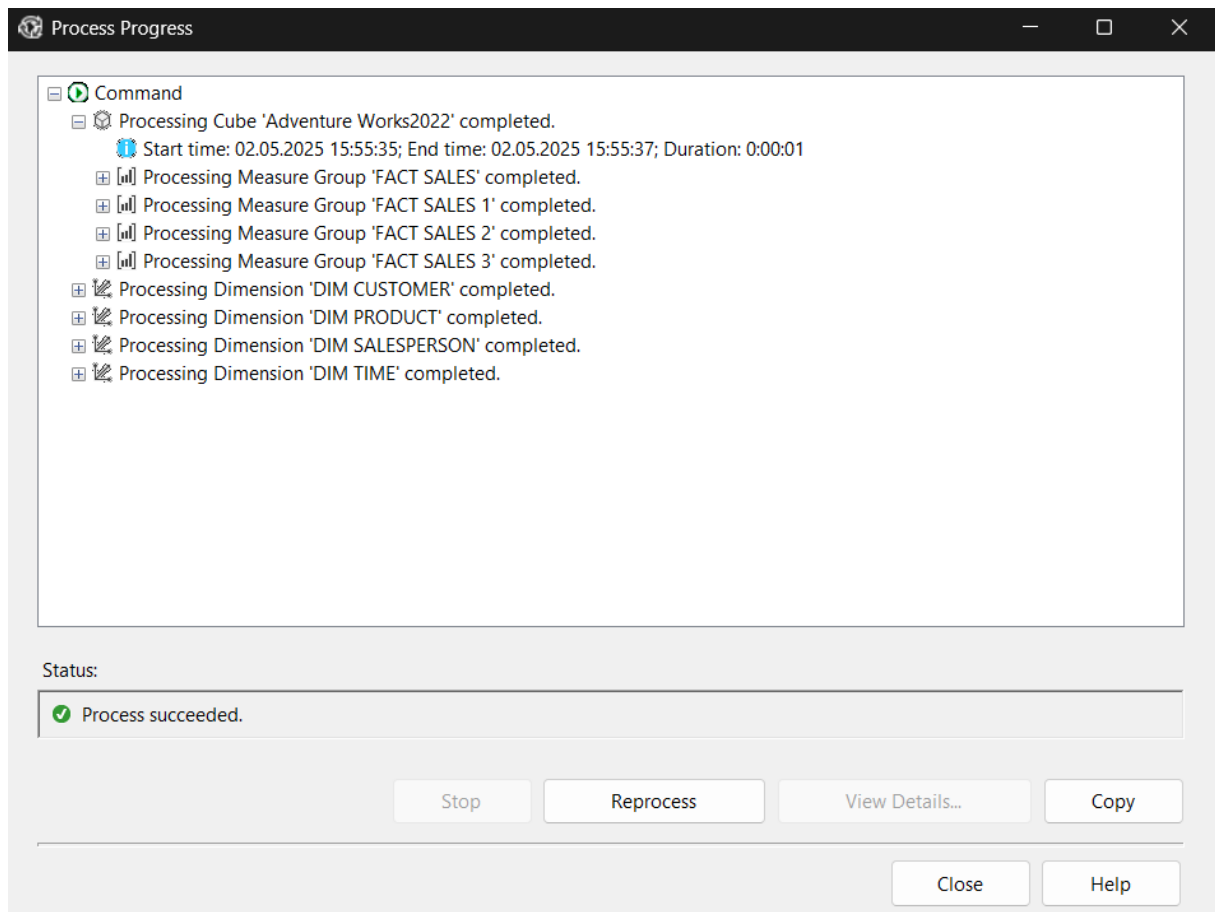
☒ Show all columns

OK Cancel Help

Measures	
Adventure Works2022	
FACT SALES	
Order Qty	
Unit Price Discount	
Line Total	
Maximum Unit Price Discount	
Maximum Order Qty	
FACT SALES 1	
Customer ID Distinct Count	
FACT SALES 2	
Product ID Distinct Count	
FACT SALES 3	
Sales Person ID Distinct Count	

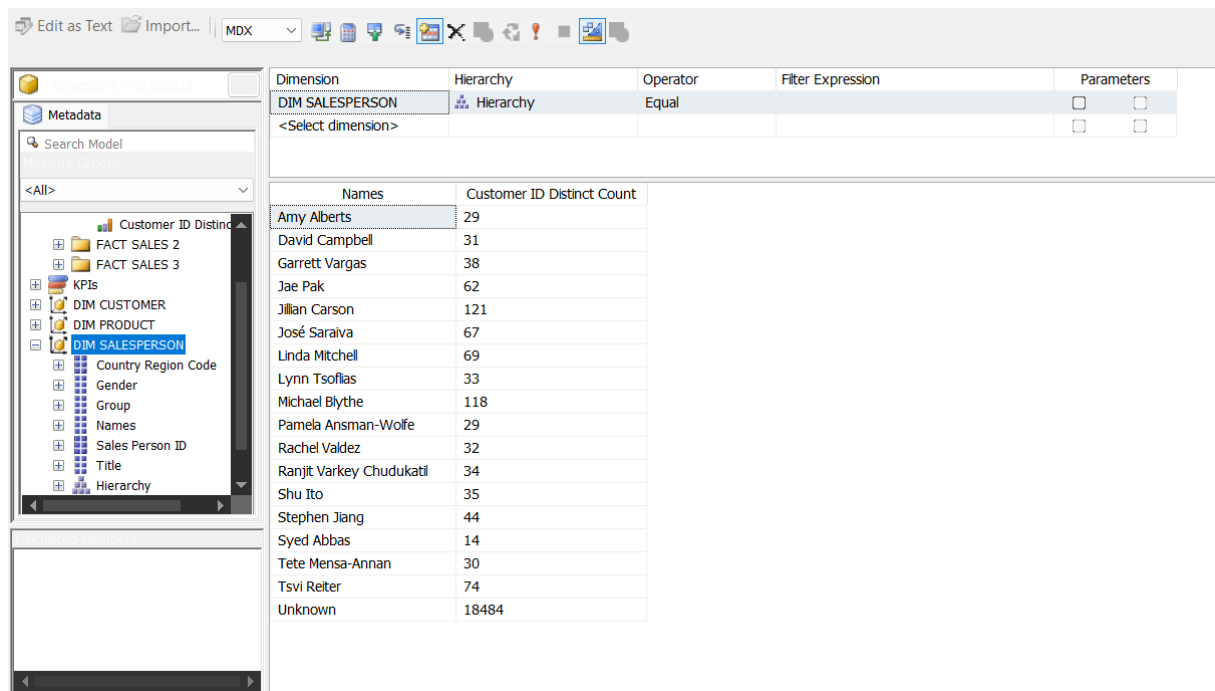
d)





Zad 2.

a)



b)

**MDX**

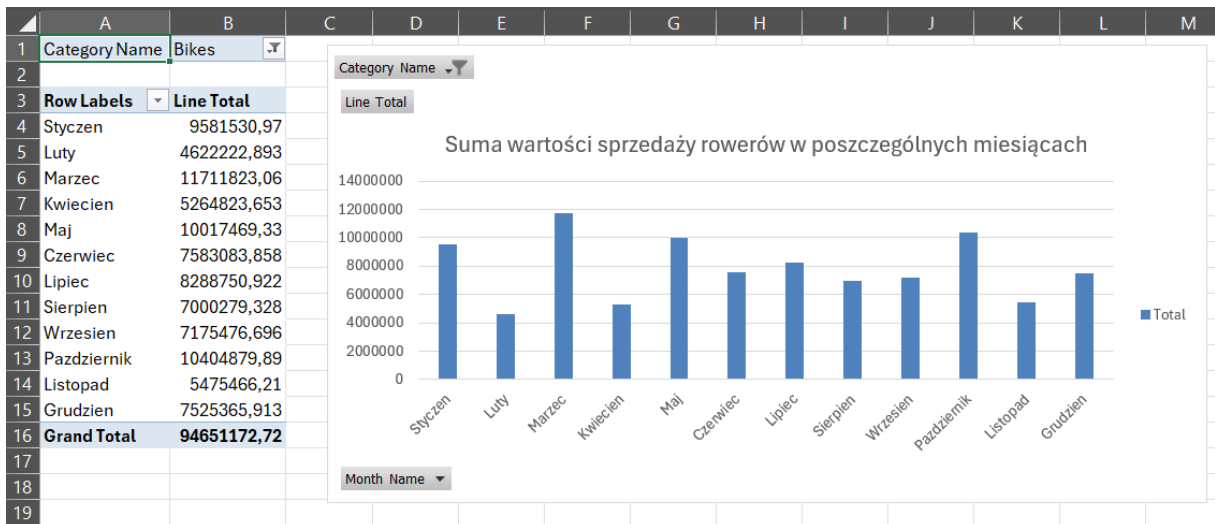
Dimension	Hierarchy	Operator	Filter Expression	Parameters
DIM PRODUCT	Category Name	Equal	{ Bikes, Clothing }	<input type="checkbox"/> <input type="checkbox"/>
<Select dimension>				

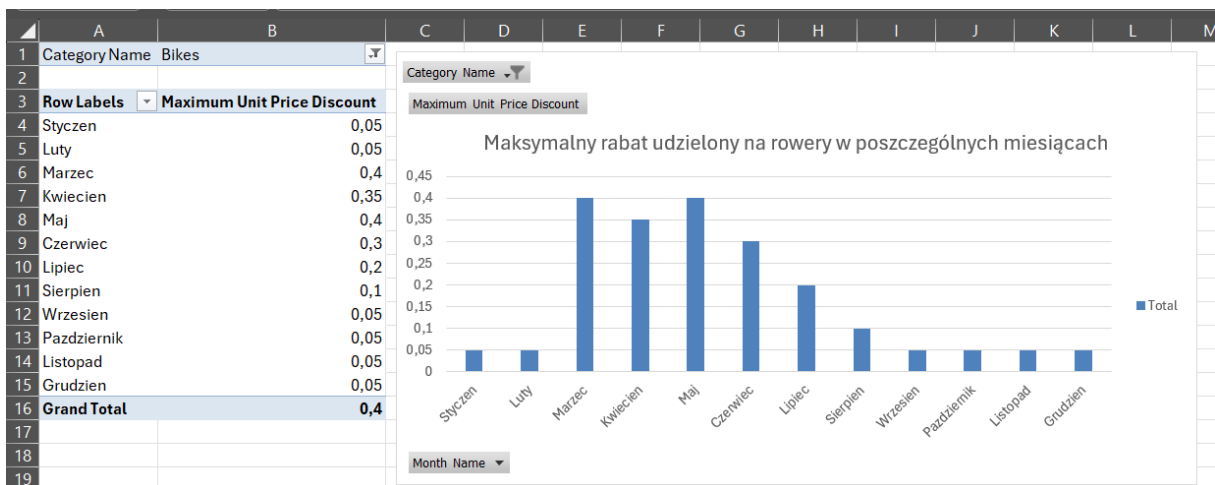
Names	Customer ID Distinct Count
Amy Alberts	27
David Campbell	30
Garrett Vargas	37
Jae Pak	58
Jillian Carson	115
José Saraiva	65
Linda Mitchell	64
Lynn Tsolias	29
Michael Blythe	112
Pamela Ansman-Wolfe	26
Rachel Valdez	30
Ranjit Varkey Chudukatti	32
Shu Ito	35
Stephen Jiang	41
Syed Abbas	13
Tete Mensa-Annan	28
Tsvi Retter	70
Unknown	12812

c)

1. Suma wartości sprzedaży rowerów w poszczególnych miesiącach

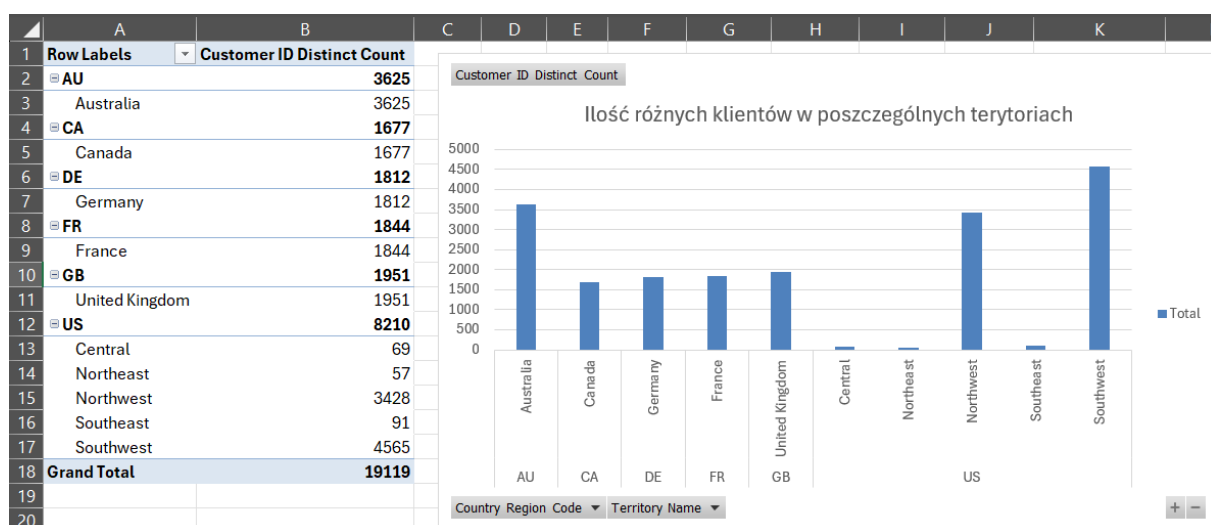


## 2. Maksymalny rabat udzielony na rowery w poszczególnych miesiącach

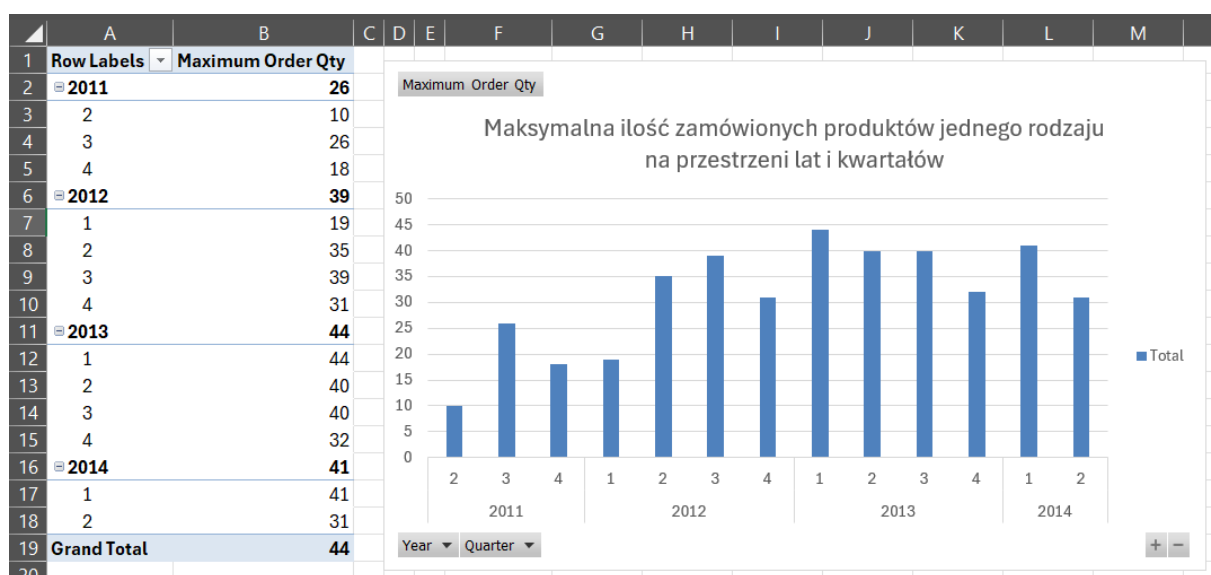




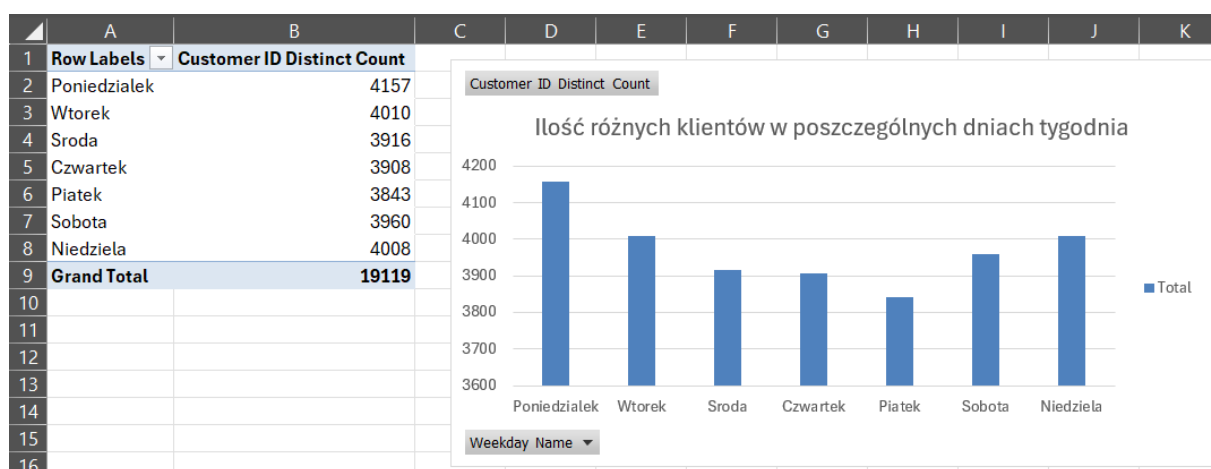
### 3. Ilość różnych klientów w poszczególnych terytoriach




### 4. Maksymalna ilość zamówionych produktów jednego rodzaju na przestrzeni lat i kwartałów



### 5. Ilość różnych klientów w poszczególnych dniach tygodnia



Zad 3.



Create Named Calculation

Column name:

WeightedQty

Description:

Expression:

[OrderQty] \* [UnitPrice]

OK

Cancel

Help

Name:

[Srednia wazona liczba towarów na zamówienie]

Parent Properties

Parent hierarchy:

Measures

Parent member:

Change

Expression

[Measures].[Weighted Qty] / [Measures].[Unit Price]

✓ No issues found

Ln: 1Ch: 52SPCCRLF

Additional Properties

Format string:

Visible:

True

Non-empty behavior:

Associated measure group:

(Undefined)

Display folder:

Color Expressions

Font Expressions

Name: [Srednia liczba zamowionych towarow na zamowienie]

Parent Properties

Parent hierarchy: Measures

Parent member:

Change

Expression

[Measures].[Order Qty] / [Measures].[Fact ID Distinct Count]

No issues found Ln: 1 Ch: 61 SPC CRLF

Additional Properties

Format string:

Visible: True

Non-empty behavior:

Associated measure group: (Undefined)

Display folder:

Color Expressions

Font Expressions

Zad 4.

a)

```
SQLQuery1.sql - ZA...-LAPTOP\aliks (68)) *
CREATE TABLE Stepaniuk.FACT_SALES_2011 (
    FactID INT IDENTITY(1,1) PRIMARY KEY,
    ProductID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_PRODUCT(ProductID),
    CustomerID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_CUSTOMER(CustomerID),
    SalesPersonID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_SALESPERSON(SalesPersonID),
    OrderDate INT NOT NULL,
    ShipDate INT,
    OrderQty INT NOT NULL,
    UnitPrice MONEY NOT NULL,
    UnitPriceDiscount MONEY NOT NULL,
    LineTotal MONEY NOT NULL
)
CREATE TABLE Stepaniuk.FACT_SALES_2012 (
    FactID INT IDENTITY(1,1) PRIMARY KEY,
    ProductID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_PRODUCT(ProductID),
    CustomerID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_CUSTOMER(CustomerID),
    SalesPersonID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_SALESPERSON(SalesPersonID),
    OrderDate INT NOT NULL,
    ShipDate INT,
    OrderQty INT NOT NULL,
    UnitPrice MONEY NOT NULL,
    UnitPriceDiscount MONEY NOT NULL,
    LineTotal MONEY NOT NULL
)
CREATE TABLE Stepaniuk.FACT_SALES_2013 (
    FactID INT IDENTITY(1,1) PRIMARY KEY,
    ProductID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_PRODUCT(ProductID),
    CustomerID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_CUSTOMER(CustomerID),
    SalesPersonID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_SALESPERSON(SalesPersonID),
    OrderDate INT NOT NULL,
    ShipDate INT,
    OrderQty INT NOT NULL,
    UnitPrice MONEY NOT NULL,
    UnitPriceDiscount MONEY NOT NULL,
    LineTotal MONEY NOT NULL
)
CREATE TABLE Stepaniuk.FACT_SALES_2014 (
    FactID INT IDENTITY(1,1) PRIMARY KEY,
    ProductID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_PRODUCT(ProductID),
    CustomerID INT FOREIGN KEY REFERENCES Stepaniuk.DIM_CUSTOMER(CustomerID),
```

```

WITH data AS (
    SELECT
        Stepaniuk.FACT_SALES.ProductID,
        Stepaniuk.FACT_SALES.CustomerID,
        Stepaniuk.FACT_SALES.SalesPersonID,
        Stepaniuk.FACT_SALES.OrderDate,
        Stepaniuk.FACT_SALES.ShipDate,
        Stepaniuk.FACT_SALES.OrderQty,
        Stepaniuk.FACT_SALES.UnitPrice,
        Stepaniuk.FACT_SALES.UnitPriceDiscount,
        Stepaniuk.FACT_SALES.LineTotal
    FROM
        Stepaniuk.FACT_SALES
    WHERE
        OrderDate >= 20130101 AND OrderDate < 20140101
)
INSERT INTO Stepaniuk.FACT_SALES_2013
SELECT * FROM data;

```

FACT SALES (4 Partitions)

Item	Partition Name ↑	Source	Estimated Rows	Storage Mode	Aggregation Design
1	FACT SALES 2011	FACT_SALES_2011	5716	MOLAP	AggregationDesign
2	FACT SALES 2012	FACT_SALES_2012	21689	MOLAP	AggregationDesign 1
3	FACT SALES 2013	FACT_SALES_2013	56573	MOLAP	AggregationDesign 2
4	FACT SALES 2014	FACT_SALES_2014	37339	MOLAP	AggregationDesign 3

[New Partition...](#) [Storage Settings...](#)

Process Progress

Command

Processing Cube 'Adventure Works2022 1' completed.  
 Start time: 02.05.2025 19:41:46; End time: 02.05.2025 19:41:47; Duration: 0:00:00  
 Processing Measure Group 'FACT SALES' completed.  
 Processing Dimension 'DIM CUSTOMER 1' completed.  
 Processing Dimension 'DIM PRODUCT 1' completed.  
 Processing Dimension 'DIM SALESPERSON 1' completed.  
 Processing Dimension 'DIM TIME 1' completed.

Status:

Process succeeded.

Stop

Reprocess

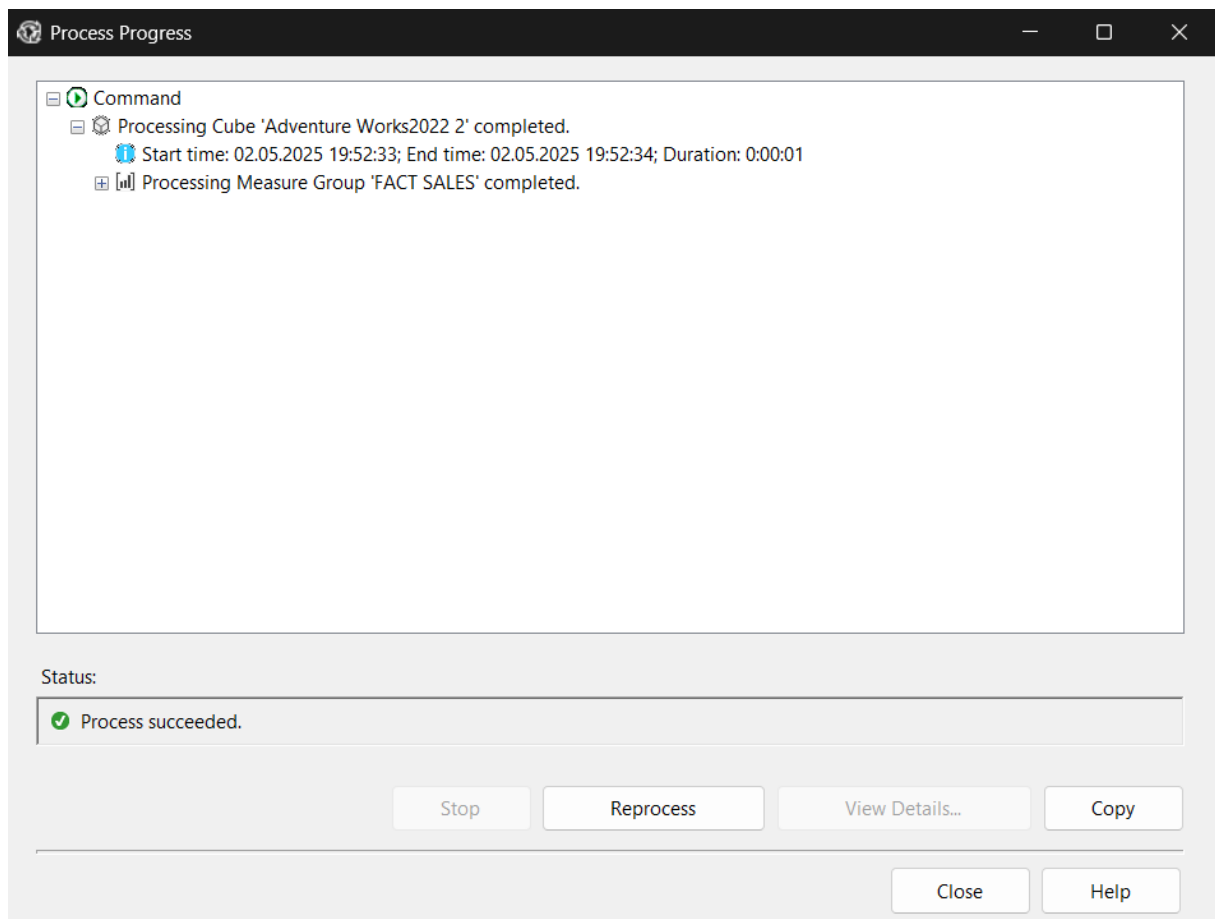
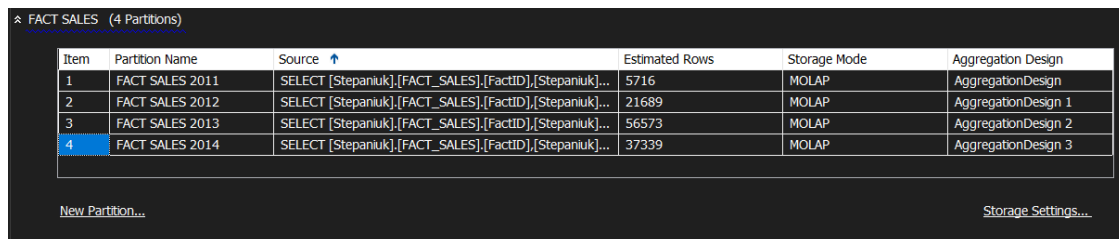
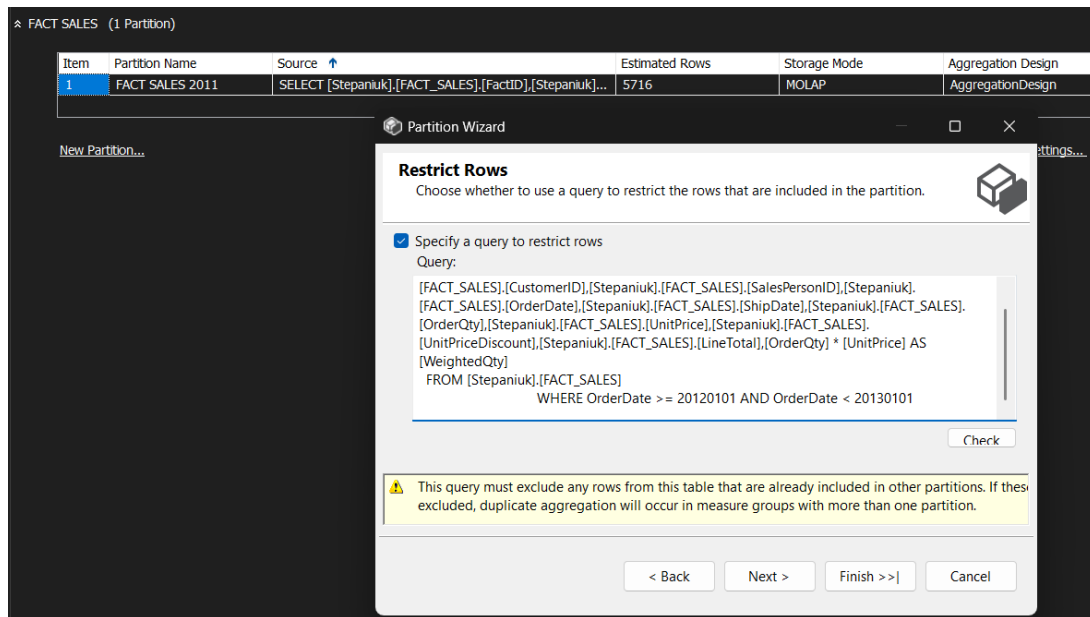
View Details...

Copy

Close

Help

b)



# Wnioski:

## Wykres 1:

Widać, że największa sprzedaż rowerów jest w marcu, jest to w końcu pierwszy miesiąc wiosny, a także sezonu rowerowego. Wysoka sprzedaż jest również w miesiącach taki jak styczeń, maj, październik co może być spowodowane na przykład promocjami noworocznymi, promocjami na majówkę, albo wyprzedażą na koniec sezonu. Najmniejsza sprzedaż jest w lutym oraz kwietniu, co może być spowodowaną zimną lub zmienną pogodą.

## Wykres 2:

Widać, że największe rabaty udzielane są w miesiącach ciepłych na początku sezonu rowerowego czyli podczas wiosny (marzec, kwiecień, maj) – później maksymalne rabaty stopniowo się zmniejszają i utrzymują stały poziom od września do lutego. Widać tutaj wyraźnie strategię rabatową sklepu rowerowego, który to w sezonie rowerowym ustawia wysokie promocje na rowery, a poza sezonem te promocje są o wiele mniejsze i wynoszą maksymalnie 5%.

## Wykres 3:

Widać, że jest o wiele więcej unikalnych klientów w Australii, czy poszczególnych regionach zachodnich Ameryki, niż w pozostałych regionach – pozwala to zobaczyć skalę rynków w poszczególnych regionach. Widać też wyraźnie, że Ameryka w znacznej wielkości posiada klientów na zachodzie, gdzie zarówno centrum jak i wschód zawiera jedynie pojedyncze przypadki, znacznie odstające od reszty.

## Wykres 4:

Widać znaczący pik w kwartale drugim oraz trzecim nałożony jednocześnie z ogólnym wzrostem i rozwojem sprzedaży sklepu w latach 2012-2013. Najlepszym kwartałem okazał się pierwszy kwartał 2013 roku, gdzie było historycznie duże zamówienie.

## Wykres 5:

Widać, że ilość różnych klientów ma zdecydowany pik w poniedziałek oraz najniższą swoją wartość osiąga w piątek. Jest to nieintuicyjne i może być spowodowane tym, że poniedziałek jest pierwszym dniem roboczym, więc ludzie po weekendzie decydują się zrobić zakupy w sklepie rowerowym.

## Zad 4:

Podział kostki na partycje pozwala szybciej przetwarzać kostkę oraz poprawia wydajność zapytań, ponieważ jeśli zapytanie nie wymaga danych z innych partycji nie ma potrzeby żeby przeszukiwać pozostałe partycje. Pierwszy sposób na tworzenie partycji jest lepszy, kiedy dane źródłowe już są podzielone na tabele w pewien sposób (na przykład podzielone w czasie na równe odstępy jak kwartały czy lata) – nie istnieje wtedy potrzeba tworzenia nowych tabel w bazie ani ich łączenia. Jeśli jednak dane nie są podzielone (czyli istnieje jedna tabela faktów) to wtedy lepiej się sprawdzi sposób drugi, ponieważ nie trzeba wtedy tworzyć nowych tabel w bazie.