

Zad 1. Przygotowanie schematu



The screenshot shows a SQL IDE window with three tabs: 'lista4.sql - JEDREK....(JEDREK\Jędrzej (60))', 'SQLQuery1.sql - JE...EDREK\Jędrzej (63))*', and '2112.sql - JEDREK.m...EDREK\Jędrzej (62))*'. The active tab is '2112.sql'. The editor contains the SQL statement 'CREATE SCHEMA KOCIECKI;' with a red squiggly line under the word 'SCHEMA'.

```
CREATE SCHEMA KOCIECKI;
```

Zad 2. Tworzenie tabel wymiarów i tabeli faktów

```
CREATE TABLE KOCIECKI.DIM_CUSTOMER (  
    CustomerID INT PRIMARY KEY,  
    FirstName NVARCHAR(50),  
    LastName NVARCHAR(50),  
    Title NVARCHAR(10),  
    City NVARCHAR(50),  
    TerritoryName NVARCHAR(50),  
    CountryRegionCode NVARCHAR(10),  
    [Group] NVARCHAR(50),  
);  
  
CREATE TABLE KOCIECKI.DIM_PRODUCT (  
    ProductID INT PRIMARY KEY,  
    Name NVARCHAR(100) NOT NULL,  
    ListPrice DECIMAL(18,2) NOT NULL,  
    Color NVARCHAR(50),  
    SubCategoryName NVARCHAR(100),  
    CategoryName NVARCHAR(100),  
    Weight DECIMAL(18,2),  
    Size NVARCHAR(50),  
    IsPurchased BIT NOT NULL,  
);  
  
CREATE TABLE KOCIECKI.DIM SALESPERSON (  
    SalesPersonID INT PRIMARY KEY,  
    FirstName NVARCHAR(50),  
    LastName NVARCHAR(50),  
    Title NVARCHAR(10),  
    Gender NVARCHAR(1),  
    CountryRegionCode NVARCHAR(3),  
    [Group] NVARCHAR(50),  
);  
  
CREATE TABLE KOCIECKI.FACT_SALES (  
    FactID INT IDENTITY(1,1) PRIMARY KEY,  
    ProductID INT,  
    CustomerID INT,  
    SalesPersonID INT,  
    OrderDate INT NOT NULL,  
    ShipDate INT,  
    OrderQty INT NOT NULL,  
    UnitPrice DECIMAL(18, 2) NOT NULL,  
    UnitPriceDiscount DECIMAL(18, 2) NOT NULL,  
    LineTotal DECIMAL(18, 2) NOT NULL,  
);
```

Kolumny OrderDate i ShipDate są typu całkowitego. W tabeli faktów sprzedawca nie jest wymagany. Wartości NULL wynikają z braku połączenia z inną tabelą produkt bez podkategorii nie może mieć kategorii, lub z opcjonalności.

Zad. 3. Wypełnianie danych – denormalizacja źródłowej bazy

```
INSERT INTO KOCIECKI.DIM_CUSTOMER (
    CustomerID,
    FirstName,
    LastName,
    Title,
    City,
    TerritoryName,
    CountryRegionCode,
    [Group]
)
SELECT DISTINCT
    C.CustomerID,
    P.FirstName,
    P.LastName,
    P.Title,
    A.City,
    T.[Name],
    T.CountryRegionCode,
    T.[Group]
FROM Sales.Customer C
JOIN Person.Person P ON C.PersonID = P.BusinessEntityID
JOIN Sales.SalesTerritory T ON C.TerritoryID = T.TerritoryID
JOIN Sales.SalesOrderHeader SOH ON C.CustomerID = SOH.CustomerID
JOIN Person.Address A ON A.AddressID = SOH.ShipToAddressID;

INSERT INTO KOCIECKI.DIM_PRODUCT
SELECT
    P.ProductID,
    P.Name,
    P.ListPrice,
    P.Color,
    PS.Name,
    PC.Name,
    P.Weight,
    P.Size,
    ~P.MakeFlag
FROM
    Production.Product P
LEFT JOIN Production.ProductSubcategory PS ON P.ProductSubcategoryID = PS.ProductSubcategoryID
LEFT JOIN Production.ProductCategory PC ON PS.ProductCategoryID = PC.ProductCategoryID;

INSERT INTO KOCIECKI.DIM_SALESPERSON
SELECT
    sp.BusinessEntityID AS SalesPersonID,
    p.FirstName,
    p.LastName,
    p.Title,
    e.Gender,
    st.CountryRegionCode,
    st.[Group]
FROM Sales.SalesPerson sp
LEFT JOIN Person.Person p ON sp.BusinessEntityID = p.BusinessEntityID
LEFT JOIN HumanResources.Employee e ON p.BusinessEntityID = e.BusinessEntityID
LEFT JOIN Sales.SalesTerritory st ON sp.TerritoryID = st.TerritoryID;
```

```

INSERT INTO KOCIECKI.FACT_SALES
(ProductID, CustomerID, SalesPersonID, OrderDate, ShipDate, OrderQty, UnitPrice, UnitPriceDiscount, LineTotal)
SELECT
    sod.ProductId,
    soh.CustomerID,
    soh.SalesPersonID,
    DATEPART(YYYY, soh.OrderDate) * 10000 + DATEPART(M, soh.OrderDate) * 100 + DATEPART(D, soh.OrderDate),
    DATEPART(YYYY, soh.ShipDate) * 10000 + DATEPART(M, soh.ShipDate) * 100 + DATEPART(D, soh.ShipDate),
    sod.OrderQty,
    sod.UnitPrice,
    sod.UnitPriceDiscount,
    sod.LineTotal
FROM Sales.SalesOrderHeader soh
JOIN Sales.SalesOrderDetail sod ON soh.SalesOrderID = sod.SalesOrderID;

```

```

SELECT COUNT(*) FROM kociecki.DIM_CUSTOMER;
SELECT COUNT(*) FROM kociecki.DIM_PRODUCT;
SELECT COUNT(*) FROM kociecki.DIM SALESPERSON;
SELECT COUNT(*) FROM kociecki.FACT_SALES;

```

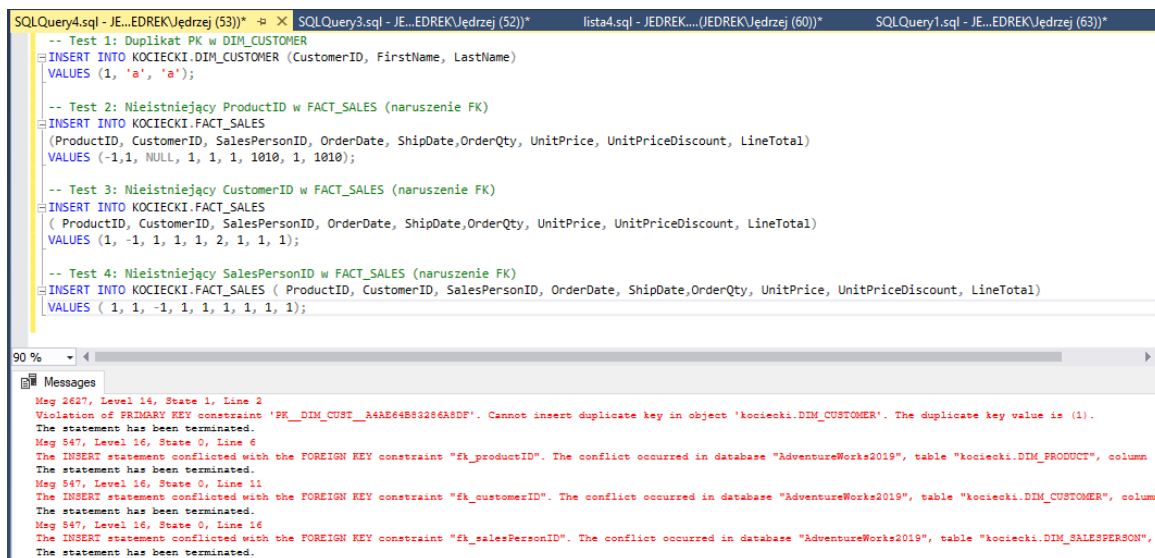
90 %

Results Messages

(No column name)	
1	19119
(No column name)	
1	504
(No column name)	
1	17
(No column name)	
1	121317

Zad. 4. Więzy integralności

```
--3
ALTER TABLE KOCIECKI.FACT_SALES
ADD
    CONSTRAINT
    fk_productID FOREIGN KEY (ProductID) REFERENCES KOCIECKI.DIM_PRODUCT(ProductID),
    CONSTRAINT
    fk_customerID FOREIGN KEY (CustomerID) REFERENCES KOCIECKI.DIM_CUSTOMER(CustomerID),
    CONSTRAINT
    fk_salesPersonID FOREIGN KEY (SalesPersonID) REFERENCES KOCIECKI.DIM SALESPERSON(SalesPersonID);
```



The screenshot displays a SQL Server environment. The top pane shows a query window with the following SQL script:

```
-- Test 1: Duplikat PK w DIM_CUSTOMER
INSERT INTO KOCIECKI.DIM_CUSTOMER (CustomerID, FirstName, LastName)
VALUES (1, 'a', 'a');

-- Test 2: Nieistniejący ProductID w FACT_SALES (naruszenie FK)
INSERT INTO KOCIECKI.FACT_SALES
(ProductID, CustomerID, SalesPersonID, OrderDate, ShipDate, OrderQty, UnitPrice, UnitPriceDiscount, LineTotal)
VALUES (-1, 1, NULL, 1, 1, 1, 1010, 1, 1010);

-- Test 3: Nieistniejący CustomerID w FACT_SALES (naruszenie FK)
INSERT INTO KOCIECKI.FACT_SALES
(ProductID, CustomerID, SalesPersonID, OrderDate, ShipDate, OrderQty, UnitPrice, UnitPriceDiscount, LineTotal)
VALUES (1, -1, 1, 1, 1, 2, 1, 1, 1);

-- Test 4: Nieistniejący SalesPersonID w FACT_SALES (naruszenie FK)
INSERT INTO KOCIECKI.FACT_SALES (ProductID, CustomerID, SalesPersonID, OrderDate, ShipDate, OrderQty, UnitPrice, UnitPriceDiscount, LineTotal)
VALUES (1, 1, -1, 1, 1, 1, 1, 1, 1);
```

The bottom pane shows the Messages window with the following error messages:

```
Msg 2627, Level 16, State 1, Line 2
Violation of PRIMARY KEY constraint 'PK_DIM_CUST_A4AE64B83286A5DF'. Cannot insert duplicate key in object 'kociecki.DIM_CUSTOMER'. The duplicate key value is (1).
The statement has been terminated.
Msg 547, Level 16, State 0, Line 6
The INSERT statement conflicted with the FOREIGN KEY constraint "fk_productID". The conflict occurred in database "AdventureWorks2019", table "kociecki.DIM_PRODUCT", column "ProductID".
The statement has been terminated.
Msg 547, Level 16, State 0, Line 11
The INSERT statement conflicted with the FOREIGN KEY constraint "fk_customerID". The conflict occurred in database "AdventureWorks2019", table "kociecki.DIM_CUSTOMER", column "CustomerID".
The statement has been terminated.
Msg 547, Level 16, State 0, Line 16
The INSERT statement conflicted with the FOREIGN KEY constraint "fk_salesPersonID". The conflict occurred in database "AdventureWorks2019", table "kociecki.DIM SALESPERSON", column "SalesPersonID".
The statement has been terminated.
```

Zad. 5. Tworzenie kostki

Data Source View Wizard

Select Tables and Views

Select objects from the relational database to be included in the data source view.

Available objects:

Name	Type
AWBuildVersion (dbo)	Table
DatabaseLog (dbo)	Table
ErrorLog (dbo)	Table
Karty (dbo)	Table
sysdiagrams (dbo)	Table
Department (HumanResources)	Table
Employee (HumanResources)	Table
EmployeeDepartmentHistory (HumanRes...	Table
EmployeePayHistory (HumanResources)	Table
JobCandidate (HumanResources)	Table
Shift (HumanResources)	Table
Address (Person)	Table
AddressType (Person)	Table
BusinessEntity (Person)	Table
BusinessEntityAddress (Person)	Table
BusinessEntityContact (Person)	Table
ContactType (Person)	Table
CountryRegion (Person)	Table
EmailAddress (Person)	Table
Password (Person)	Table
Person (Person)	Table
PersonPhone (Person)	Table
PhoneNumberType (Person)	Table
StateProvince (Person)	Table
BillOfMaterials (Production)	Table
Culture (Production)	Table
Document (Production)	Table
Illustration (Production)	Table
Location (Production)	Table
Product (Production)	Table
ProductCategory (Production)	Table
ProductCostHistory (Production)	Table

Filter:

☐ Show system objects

Included objects:

Name	Type
DIM_CUSTOMER (KOCIECKI)	Table
DIM_PRODUCT (KOCIECKI)	Table
DIM SALESPERSON (KOCIECKI)	Table
FACT_SALES (KOCIECKI)	Table

>

<

>>

<<

Add Related Tables

< Back

Next >

Finish >>

Cancel

Cube Wizard

Select Measure Group Tables

Select a data source view or diagram and then select the tables that will be used for measure groups.

Data source view:

Adventure Works2019

Measure group tables:

☐ DIM_CUSTOMER

☐ DIM_PRODUCT

☐ DIM SALESPERSON

☒ FACT_SALES

Suggest

< Back

Next >

Finish >>

Cancel

Cube Wizard

Completing the Wizard

Name the cube, review its structure, and then click Finish to save the cube.

Cube icon

✓

Cube name:

Adventure Works2019

Preview:

Measure groups

FACT SALES

Order Qty

Unit Price

Unit Price Discount

Line Total

FACT SALES Count

Dimensions

DIM PRODUCT

DIM SALESPERSON

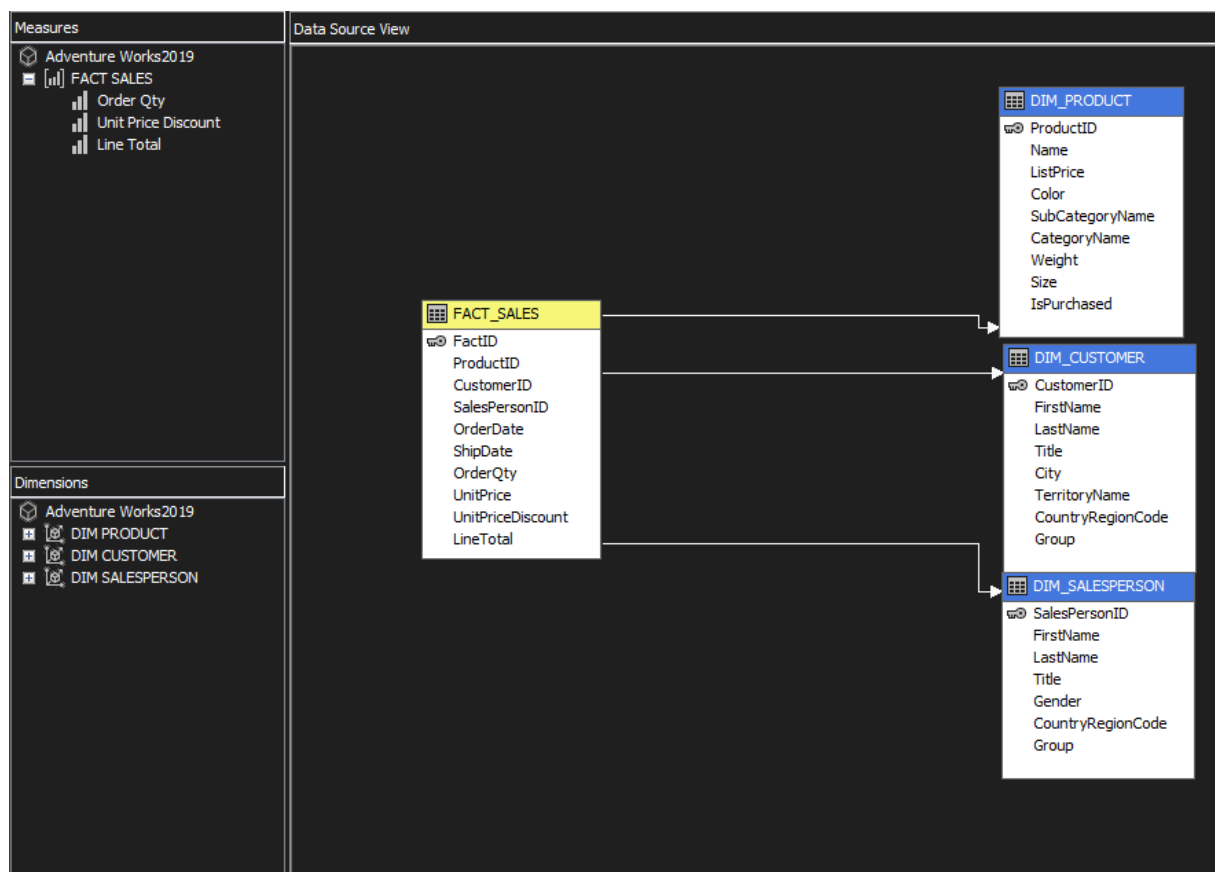
DIM CUSTOMER

< Back

Next >

Finish

Cancel



Data Source View

FACT_SALES

FactID

ProductID

CustomerID

SalesPersonID

OrderDate

ShipDate

OrderQty

UnitPrice

UnitPriceDiscount

LineTotal

DIM_PRODUCT

ProductID

Name

ListPrice

Color

SubCategoryName

CategoryName

Weight

Size

IsPurchased

DIM_CUSTOMER

CustomerID

FirstName

LastName

Title

City

TerritoryName

CountryRegionCode

Group

DIM SALESPERSON

SalesPersonID

FirstName

LastName

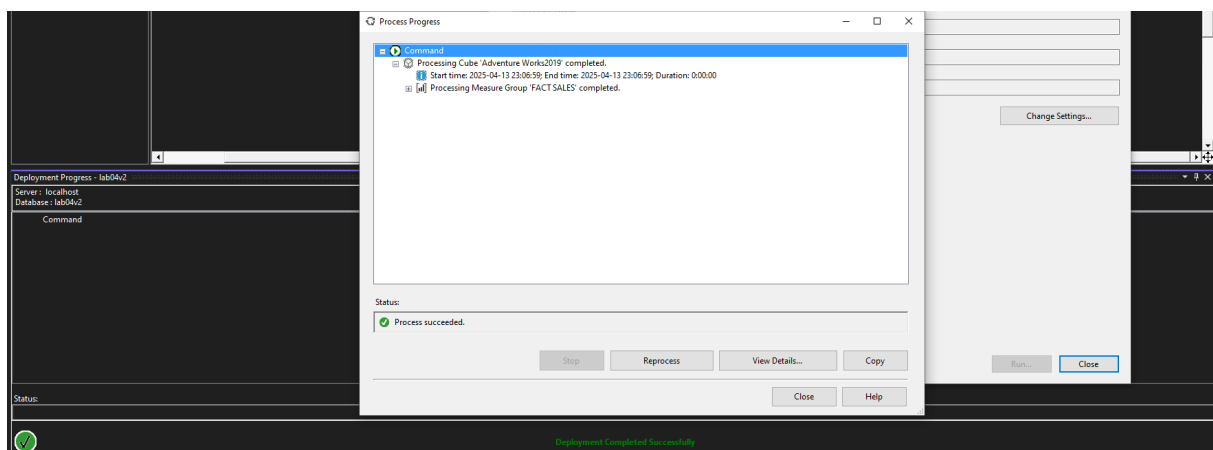
Title

Gender

CountryRegionCode

Group

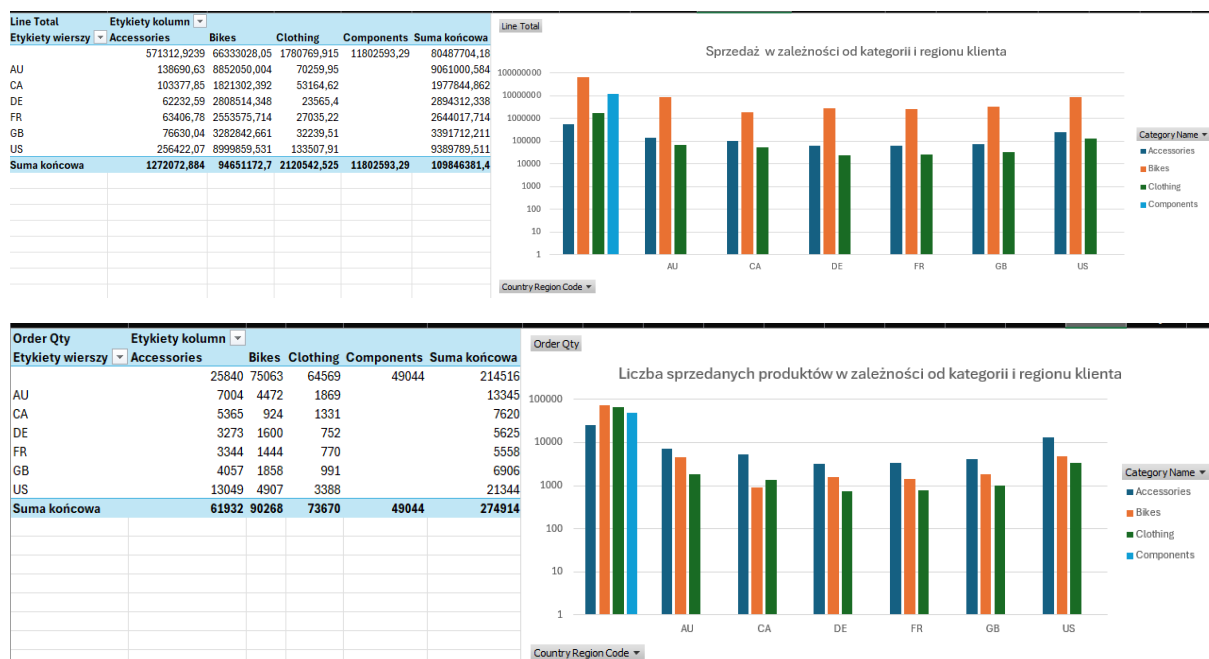
Zad. 6. Uruchomienie kostki



Zad. 7

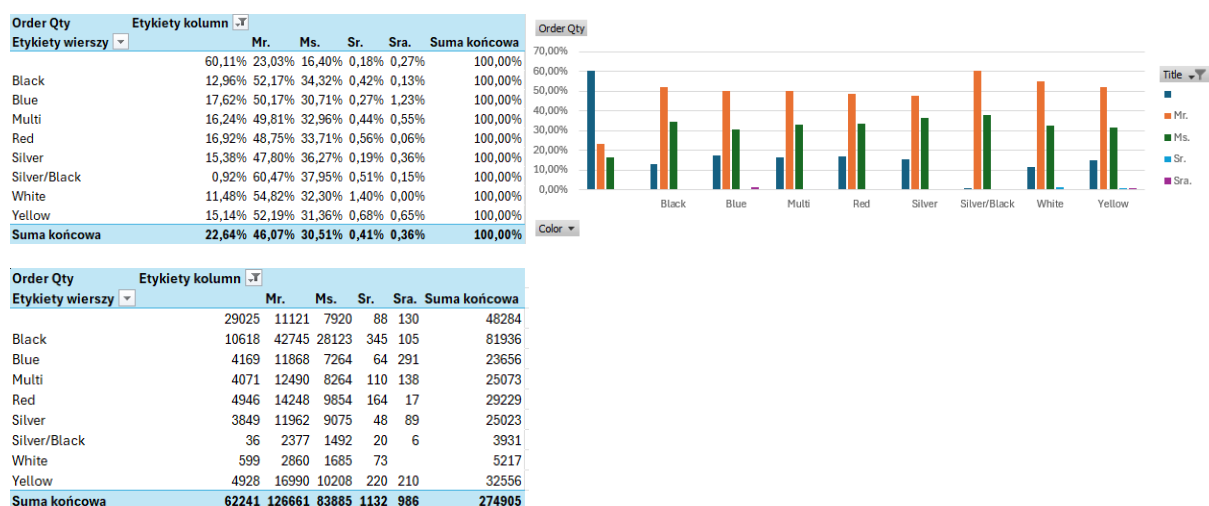
Proste raporty Połączyć się z programem MS Excel i przygotować raporty (tabele i wykresy przestawne), w których zostaną ujęte ciekawe zależności pomiędzy danymi.

1. Sprzedaż w danych kategoriach w danych regionach



- Rowery są zdecydowanym liderem przychodów, stanowią ponad 86% całkowitej sprzedaży (94,65mln z 109,85mln).
- Akcesoria miały spory udział ilościowy w sprzedaży – we wszystkich regionach zajmują pierwsze miejsce, mają za to zdecydowanie mniejszy udział wartościowy spowodowany niższą ceną jednostkową.
- Najwięcej produktów sprzedano w US – 9,39 mln i w Australii 9,06mln.
- US i AU to dwa najważniejsze rynki dla rowerów, dodatkowo akcesoria w US generują ponad 256tys., co stanowi 20% całkowitej sprzedaży (256tys. / 1.27mln)
- Produkty odzieżowe generują najmniejsze przychody i są sprzedawane w najmniejszej liczbie sztuk niezależnie od regionu.
- Kategoria Components wygenerowała 11.8mln przychodu i sprzedaż 49 044 sztuk, jednak nie jest przypisana do poszczególnych regionów.
- Australia zaskakująco osiąga drugie miejsce, wyprzedzając kraje europejskie, Wielka Brytania, Francja, Niemcy mają zdecydowanie większą liczbę ludności, dodatkowo w krajach europejskich kolarstwo jest popularnym sportem. Zapewne spora konkurencja na rynku europejskim utrudnia wybicie się.
- Analiza może również wskazywać na potencjał optymalizacji oferty produktów odzieżowych, która generuje najmniejsze przychody i jest sprzedawana w najmniejszej liczbie sztuk niezależnie od regionu.

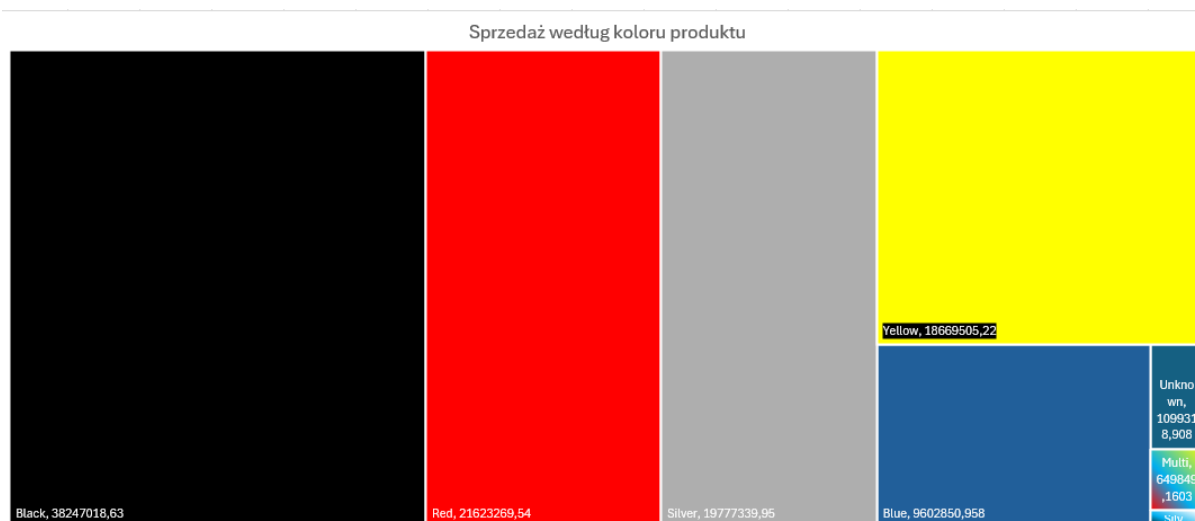
2. Kolorystyka produktu a sprzedaż



- Czarny kolor produktów stanowi największy udział w sprzedaży, co nie jest zaskakujące zważywszy na jego uniwersalność.
- Mężczyźni stanowią największą grupę klientów. Najpopularniejsze kolory produktów w zależności od płci są takie same: czarny, żółty, czerwony, multi, srebrny. Ciężko zauważyć jakieś różnice w preferencjach w zależności od płci.
- Istnieje niewielki odsetek hiszpańskich klientów z tytułami Sr. i Sra.
- Zaskakująca może być niska liczba sprzedanych białych produktów. Biały kolor tak jak i czarny wydają się być tymi najbardziej pospolitymi i najczęściej wybieranymi, okazuje się jednak, że ludzie częściej wybierają nawet produktu o rażących kolorach np. żółty czy czerwony.



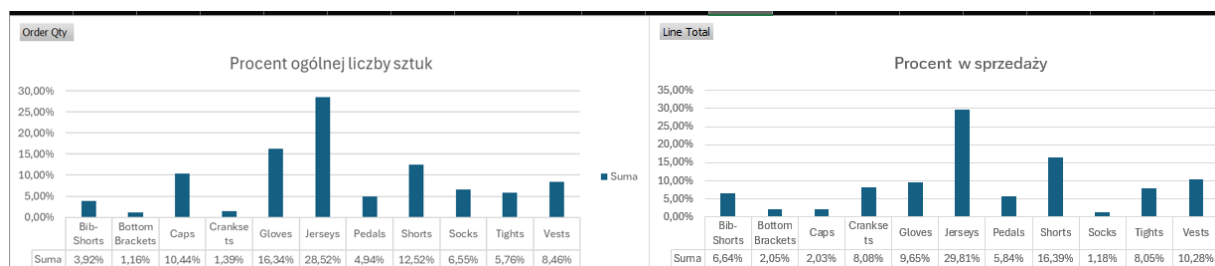
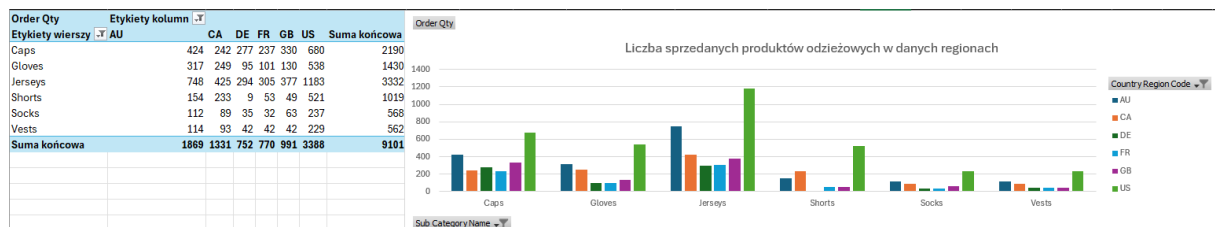
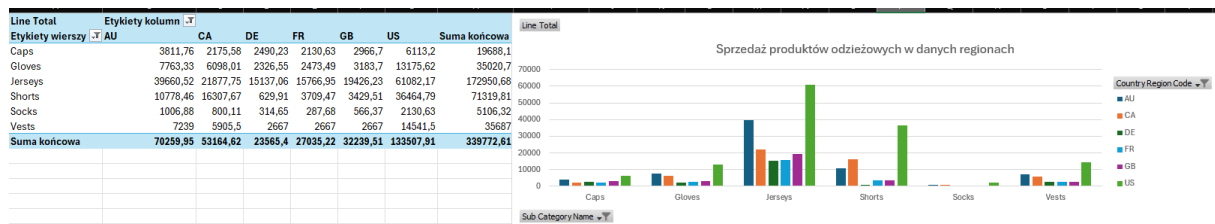
- Spora liczba produktów nie ma zdefiniowanego koloru.
- W regionach bez większych zaskoczeń, czarny jest zawsze najczęściej wybieranym kolorem produktu.
- Najpopularniejsze kolory plasują się w tej samej kolejności w różnych regionach, ze zmianą w Australii gdzie na 2 miejsce zajmuje żółty zamiast, czerwonego tak jak to jest w innych regionach. Jednak mimo to różnica jest minimalna.



- Kolor srebrny, pod względem liczby sprzedanych sztuk, w każdym regionie zajmował niższe miejsce niż kolor żółty. Pomimo to wygenerował większe zyski, może to wskazywać, że produkty srebrnego koloru mają średnio większe ceny jednostkowe.
- Produkty multi, silver/black mają niewielki udział w sprzedaży, może to wskazywać na mniejsze zapotrzebowanie i popularność produktów łączących kolory, być może preferowana jest prostolinijna kolorystyka
- Istnieje znacząca różnica pomiędzy produktami sprzedającymi się najlepiej i najgorzej, może to wskazywać na silne preferencje konsumentów, różnice w cenach produktów w poszczególnych kolorach np. wynikające z większego kosztu produkcji produktów w nie typowych kolorach.

- Kolor czarny zdecydowanie dominuje nad resztą, wyróżnia się 5 dominujących kolorów, wprowadzając nowe produkty na rynek warto upewnić się, że są one dostępne w tych 5 kolorach.

3. Analiza sprzedaży produktów odzieżowych w danych regionach



- Koszulki generują zdecydowanie największy globalny przychód, zdominowały resztę produktów zarówno pod kątem liczby sprzedanych sztuk jak i pod kątem zysku stanowiąc odpowiednio 28.5% i prawie 30% całościowej sprzedaży produktów odzieżowych.
- Koszulki są zdecydowanie najbardziej popularnym towarem sprzedawanym przez sklep w każdym regionie. Nawet regiony cechujące się niskim zainteresowaniem produktami odzieżowymi (Francja, Niemcy) wykazują się stosunkowo sporym zainteresowaniem koszulkami.

- USA zdecydowanie podkreśla swój udział jako zdecydowanie największy rynek. Zdominowali wszystkie produkty odzieżowe. Odpowiadają za 37% (3 388/9 101) sprzedanych sztuk i 39% (133 507 / 399 733) całkowitych przychodów.
- Czapki cechują się stosunkowo sporym zainteresowaniem, ale generują niewielki zysk.
- Istnieją produkty zupełnie niszowe jak skarpetki i kamizelki, niska sprzedaż i udział w przychodach. Możliwe, że nie są promowane lub nie mają dobrej pozycji w ofercie. Być może potrzebna jest zmiana polityki sprzedażowej tych produktów – np. traktowanie ich jako produkty wspierające do zestawów promocyjnych.
- Zagadkowo niska liczba sprzedanych spodenek w Niemczech, inne kraje Europejskie (Francja, GB), które kupują kilkakrotnie razy więcej szortów. Statystyka o tyle dziwna, że rynek niemiecki nie odstaje od Francuskiego czy Brytyjskiego, aż tak w innych kategoriach. Możliwe problemy z dopasowaniem produktu, brak lokalizacji lub być może różnica kulturowa.