

Zad. 1. Ekstrakcja danych

1. Utworzyć zestawienie, które dla poszczególnych miesięcy i lat przedstawi informację o liczbie różnych klientów. Przygotuj zapytanie z i bez użycia polecenia pivot.

--zadanie 1a

```
SELECT YEAR(OrderDate) 'Rok', MONTH(OrderDate) 'Miesiac',  
COUNT(DISTINCT CustomerID) 'Rozni klienci'  
FROM Sales.SalesOrderHeader  
GROUP BY YEAR(OrderDate), MONTH(OrderDate)  
ORDER BY 1, 2;
```

--zadanie 1b

```
SELECT * FROM  
(  
    SELECT DISTINCT YEAR(OrderDate) 'Rok', MONTH(OrderDate)  
    'Miesiac', CustomerID 'Rozni_klienci'  
    FROM Sales.SalesOrderHeader  
) src  
PIVOT  
(  
    COUNT(Rozni_klienci)  
    FOR Miesiac IN ( [1], [2], [3], [4], [5], [6], [7], [8], [9],  
    [10], [11], [12])  
) AS pivot_table  
ORDER BY Rok;
```

Results		Messages	
	Rok	Miesiac	Rozni klienci
1	2011	5	43
2	2011	6	141
3	2011	7	231
4	2011	8	250
5	2011	9	157
6	2011	10	327
7	2011	11	230
8	2011	12	228

	Rok	1	2	3	4	5	6	7	8	9	10	11	12
1	2011	0	0	0	0	43	141	231	250	157	327	230	228
2	2012	336	219	304	269	293	390	385	285	352	321	383	378
3	2013	400	325	441	428	426	713	1675	1727	1741	1893	2041	1970
4	2014	2...	1...	2...	2...	2...	898	0	0	0	0	0	0

2. Utworzyć zestawienie zawierające w wierszach imiona i nazwiska sprzedawców, a w kolumnach kolejne lata. Wartością będzie liczba obsłużonych transakcji.

```

SELECT * FROM
(
    SELECT  CONCAT(FirstName, ' ', LastName) 'Imie i nazwisko',
    YEAR(OrderDate) 'rok'
    FROM
        Sales.SalesOrderHeader SOH
        JOIN Person.Person P ON P.BusinessEntityID = SOH.SalesPersonID
) src
PIVOT
(
    COUNT(rok)
    FOR rok IN ([2011], [2012], [2013], [2014])
) pivot_table
WHERE
    [2011] != 0
    AND [2012] != 0
    AND [2013] != 0
    AND [2014] != 0;

```

	Imie i nazwisko	2011	2012	2013	2014
1	David Campbell	28	63	72	26
2	Garrett Vargas	30	80	89	35
3	Jillian Carson	59	166	185	63
4	José Saraiva	56	86	86	43
5	Linda Mitchell	46	151	162	59
6	Michael Blythe	65	148	175	62
7	Pamela Ansman-Wolfe	22	45	19	9
8	Shu Ito	33	74	98	37
9	Stephen Jiang	4	22	14	8
10	Tsvi Reiter	63	153	159	54

3. Zdefiniować zapytanie wyznaczające sumę kwot sprzedaży towarów oraz liczbę różnych produktów w zamówieniach w poszczególnych latach, miesiącach, dniach.ch sprzedawców, którzy pracowali przez wszystkie 4 lata.

--zadanie 3

```
WITH DailyProductCounts AS (  
    SELECT  
        YEAR(h.OrderDate) AS Rok,  
        MONTH(h.OrderDate) AS Miesiac,  
        DAY(h.OrderDate) AS Dzień,  
        COUNT(DISTINCT d.ProductID) AS LiczbaRoznychProduktow  
    FROM Sales.SalesOrderHeader h  
    JOIN Sales.SalesOrderDetail d ON h.SalesOrderID = d.SalesOrderID  
    GROUP BY YEAR(h.OrderDate), MONTH(h.OrderDate), DAY(h.OrderDate)  
)  
  
SELECT  
    YEAR(soh.OrderDate) AS Rok,  
    MONTH(soh.OrderDate) AS Miesiac,  
    DAY(soh.OrderDate) AS Dzień,  
    CAST(SUM(soh.TotalDue) AS DECIMAL(15,2)) AS 'Suma',  
    dpc.LiczbaRoznychProduktow AS 'Liczba roznych produktow'  
FROM Sales.SalesOrderHeader soh  
JOIN DailyProductCounts dpc ON  
    dpc.Rok = YEAR(soh.OrderDate) AND  
    dpc.Miesiac = MONTH(soh.OrderDate) AND  
    dpc.Dzień = DAY(soh.OrderDate)  
GROUP BY YEAR(soh.OrderDate), MONTH(soh.OrderDate), DAY(soh.OrderDate),  
dpc.LiczbaRoznychProduktow  
ORDER BY 1, 2, 3;
```

4. Wykorzystując polecenie CASE przygotować podsumowania do zestawienia z poprzedniego zadania tak, aby sumowane były kwoty zamówień oraz obliczana liczba różnych produktów dla poszczególnych miesięcy i dni tygodnia.

```
SELECT
    MONTH(OrderDate) AS 'Miesiac',
    CASE DATEPART(weekday, OrderDate)
        WHEN 1 THEN 'Niedziela'
        WHEN 2 THEN 'Poniedziałek'
        WHEN 3 THEN 'Wtorek'
        WHEN 4 THEN 'Środa'
        WHEN 5 THEN 'Czwartek'
        WHEN 6 THEN 'Piątek'
        WHEN 7 THEN 'Sobota'
    END AS 'Dzien tygodnia',
    CAST(ROUND(SUM(LineTotal), 2) AS DECIMAL(10,2)) AS 'Suma',
    COUNT(DISTINCT ProductID) AS 'Liczba roznych produktow'
FROM Sales.SalesOrderHeader SOH
JOIN Sales.SalesOrderDetail SOD ON SOH.SalesOrderID = SOD.SalesOrderID
GROUP BY MONTH(OrderDate), DATEPART(weekday, OrderDate)
ORDER BY 1, 2;
```

	Rok	Miesiac	Dzien	Suma	Liczba roznych produktow
1	2011	5	31	567020.95	47
2	2011	6	1	15394.33	4
3	2011	6	2	16588.46	4
4	2011	6	3	7907.98	2
5	2011	6	4	16588.46	4
6	2011	6	5	15815.95	3
7	2011	6	6	8680.48	3
8	2011	6	7	8680.48	2
9	2011	6	8	23105.31	5
10	2011	6	9	11664.97	3
11	2011	6	10	15815.95	3
12	2011	6	11	15618.95	4
13	2011	6	12	7907.98	2
14	2011	6	13	27677.92	4

5. Przygotować zestawienie, w którym dla wybranych klientów przygotujemy kartę lojalnościową: a. srebrną, jeśli klient wykonał co najmniej 2 transakcje w sklepie; b. złotą, jeśli wykonał co najmniej 4 transakcje w sklepie, w tym co najmniej 2 transakcje, których łączna kwota przekraczała 250% średniej wartości zamówień w bazie; c. platynową, jeśli klient spełniał warunki otrzymania karty złotej oraz w co najmniej jednej transakcji kupił jednocześnie produkty ze wszystkich kategorii.

```
DROP TABLE Karty;
```

```
CREATE TABLE Karty(  
    Imie VARCHAR(255),  
    Nazwisko VARCHAR(255),  
    "Liczba transakcji" int,  
    "Laczna kwota transakcji" int,  
    "Kolor karty" VARCHAR(255)
```

```
);
```

```
DECLARE @Srednia DECIMAL(18,2);  
SET @Srednia = (SELECT AVG(TotalDue) FROM Sales.SalesOrderHeader);
```

```
WITH Transakcje AS (  
    SELECT  
        c.CustomerID,  
        p.FirstName,  
        p.LastName,  
        COUNT(soh.SalesOrderID) AS LiczbaTransakcji,  
        SUM(soh.TotalDue) AS LacznaKwotaTransakcji  
    FROM Sales.SalesOrderHeader soh  
    JOIN Sales.Customer c ON soh.CustomerID = c.CustomerID  
    JOIN Person.Person p ON c.PersonID = p.BusinessEntityID  
    GROUP BY c.CustomerID, p.FirstName, p.LastName  
),
```

```
ZloteKryterium AS (  
    SELECT  
        soh.CustomerID  
    FROM Sales.SalesOrderHeader soh  
    WHERE soh.TotalDue > 2.5 * @Srednia  
    GROUP BY soh.CustomerID  
    HAVING COUNT(soh.SalesOrderID) >= 2  
),
```

```
PlatynoweKryterium AS (  
    SELECT DISTINCT soh.CustomerID  
    FROM Sales.SalesOrderHeader soh  
    JOIN Sales.SalesOrderDetail sod ON soh.SalesOrderID = sod.SalesOrderID  
    JOIN Production.Product p ON sod.ProductID = p.ProductID  
    JOIN Production.ProductSubcategory ps ON p.ProductSubcategoryID =  
    ps.ProductSubcategoryID
```

```

        GROUP BY soh.CustomerID, soh.SalesOrderID
        HAVING COUNT(DISTINCT ps.ProductCategoryID) = (SELECT COUNT(*) FROM
Production.ProductCategory)
    )
INSERT INTO Karty (Imie, Nazwisko, "Liczba transakcji", "Laczna kwota
transakcji", "Kolor karty")
SELECT
    t.FirstName,
    t.LastName,
    t.LiczbaTransakcji,
    t.LacznaKwotaTransakcji,
    CASE
        WHEN t.LiczbaTransakcji >= 4 AND z.CustomerID IS NOT NULL AND
p.CustomerID IS NOT NULL THEN 'Platynowa'
        WHEN t.LiczbaTransakcji >= 4 AND z.CustomerID IS NOT NULL THEN
'Złota'
        WHEN t.LiczbaTransakcji >= 2 THEN 'Srebrna'
        ELSE 'Brak'
    END AS KolorKarty
FROM Transakcje t
LEFT JOIN ZloteKryterium z ON t.CustomerID = z.CustomerID
LEFT JOIN PlatynoweKryterium p ON t.CustomerID = p.CustomerID
ORDER BY t.LacznaKwotaTransakcji DESC;

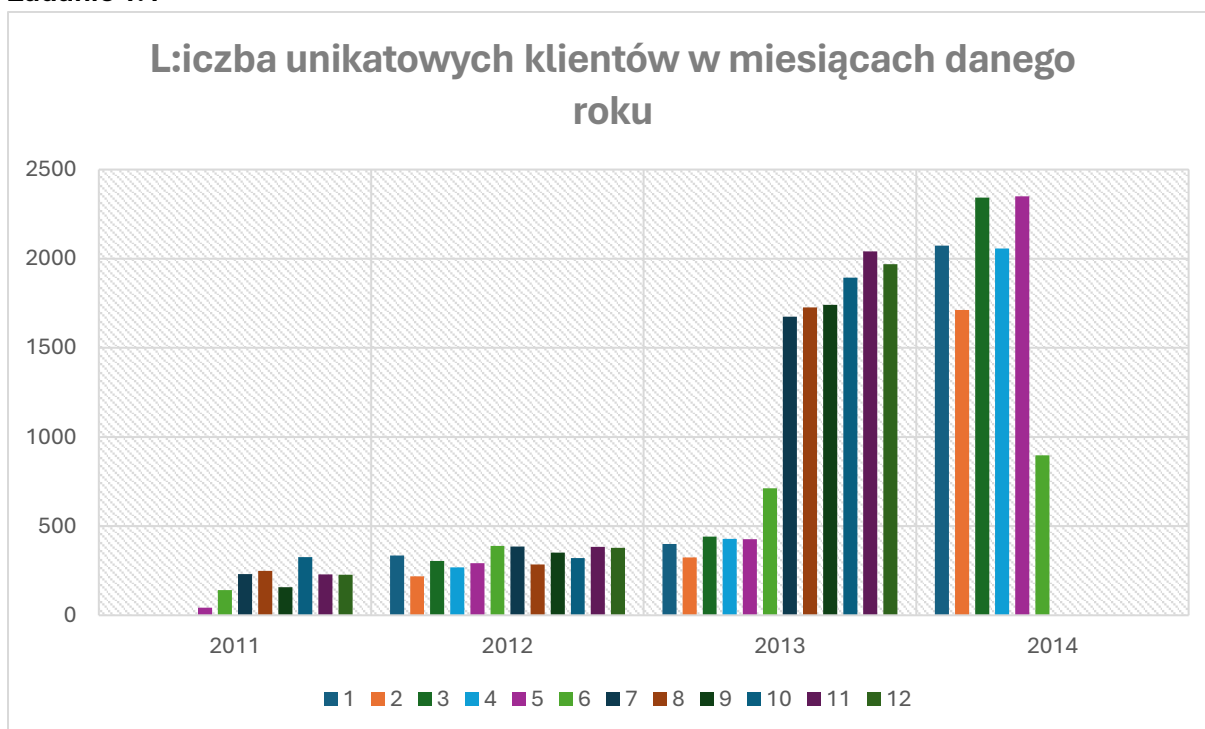
SELECT * FROM Karty;

```

	Imie	Nazwisko	Liczba transakcji	Laczna kwota transakcji	Kolor karty
1	Catherine	Abel	4	127380	Platynowa
2	Kim	Abercrombie	12	584949	Platynowa
3	Humberto	Acevedo	11	77585	Złota
4	Gustavo	Achong	7	147805	Platynowa
5	Pilar	Ackerman	4	249805	Platynowa
6	Aaron	Adams	1	130	Brak
7	Adam	Adams	1	157	Brak
8	Alex	Adams	1	1918	Brak
9	Alexandra	Adams	1	3954	Brak
10	Allison	Adams	2	1771	Srebrna
11	Amanda	Adams	2	149	Srebrna
12	Amber	Adams	2	2212	Srebrna

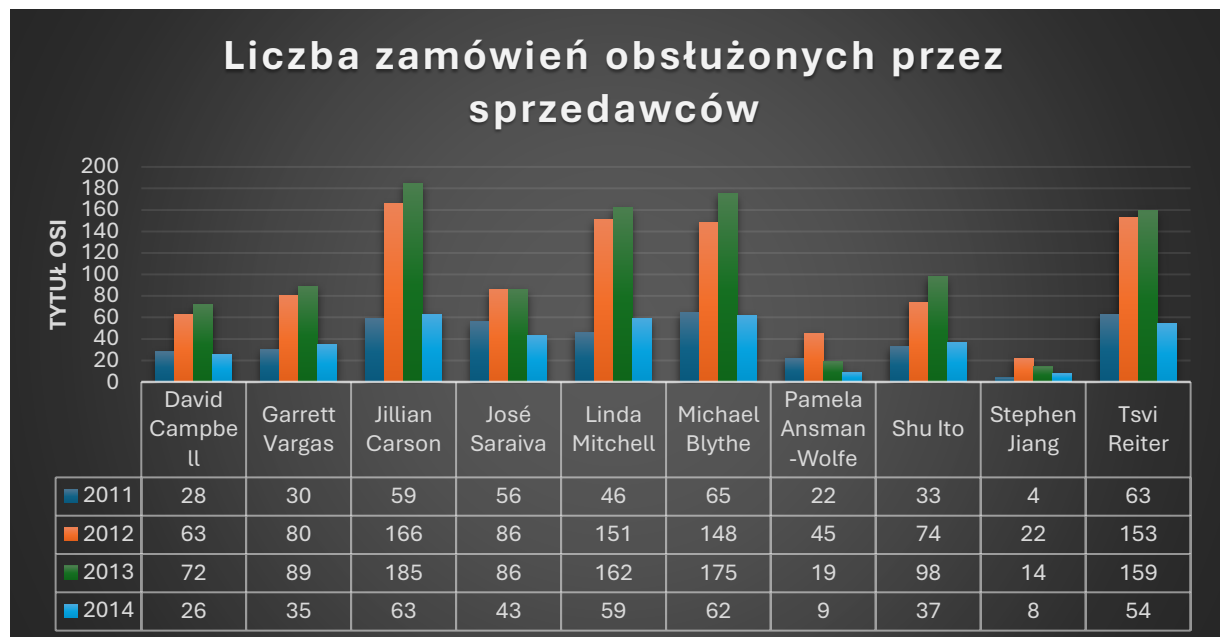
2. Przedstaw wyniki zadania 1 w postaci tabel i wykresów przestawnych w programie MS Excel. Zinterpretuj wyniki.

Zadanie 1.1



- W 2013 roku w okresie wakacyjnym nastąpił znaczący wzrost liczby klientów, co może świadczyć o ekspansji działalności firmy, skuteczniejszym marketingu lub zwiększonym popycie na oferowane produkty.
- Wzrost ten utrzymał się, nie był chwilowy. Może to świadczyć o zmianach modelu biznesowego, ekspansji rynku, czy działaniach promocyjnych/marketingowych stosowanych od okresu wakacyjnego 2013 roku.
- W 2014 widoczne są spore wahania wartości, co może sugerować okresowe działania promocyjne lub inne czynniki wpływające na zainteresowanie klientów.
- W latach 2011 – 2012 widoczny niewielki wzrost wartości – sugeruje, że firma była dopiero na etapie zdobywania rynku i klientów

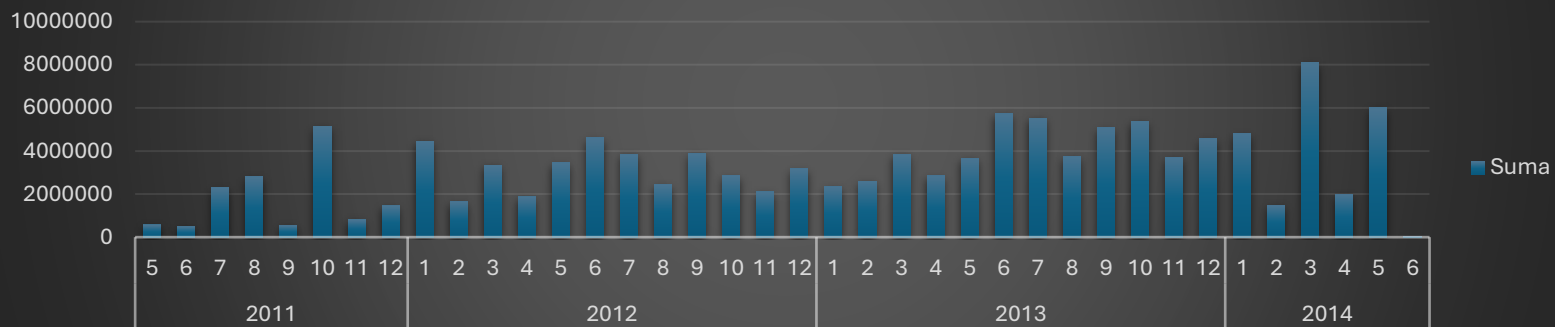
Zadanie 1.2



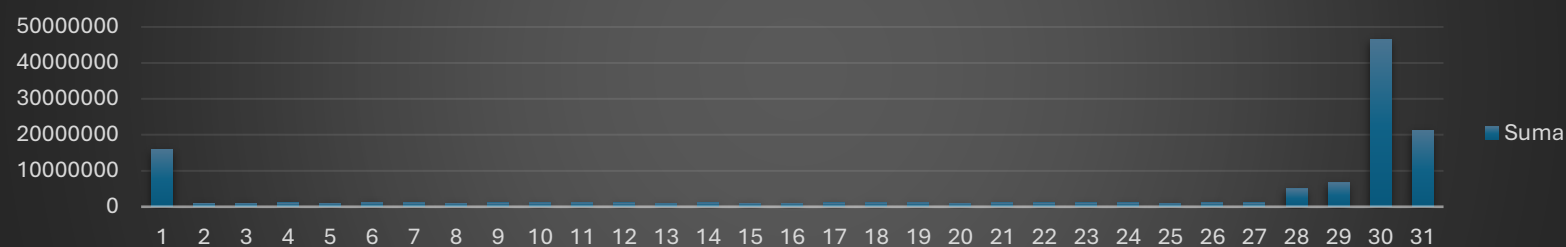
- Widać duże różnice w liczbie zamówień obsługiwanych przez poszczególnych sprzedawców
- Niektórzy sprzedawcy, np. Michael Blythe, mają znacznie więcej obsługiwanych zamówień niż inni.
- Pamela Ansman-Wolfe i Stephen Jiang jako jedyni sprzedawcy nie utrzymali tendencji wzrostowej w liczbie zamówień z 2012 na 2013 rok.
- Pomimo znacznego wzrostu unikalnych klientów w 2013 roku, liczba zamówień obsługiwanych przez sprzedawców w 2013 niewiele różni się od tej z roku 2012.
- Należy zwrócić uwagę, że sprzedawcy realizujący małą liczbę zamówień nie muszą być uznani za nieefektywnych – ich zadaniem może być realizacja np. czasochłonnych zamówień hurtowych na duże kwoty. Przykładowo Pamela Ansman-Wolfe w 2012 wygenerowała podobny przychód co David Campbell i Garrett Vargas.
- Na podium pod względem liczby zamówień konsekwentnie możemy zauważyć tych samych sprzedawców

Zadanie 1.3

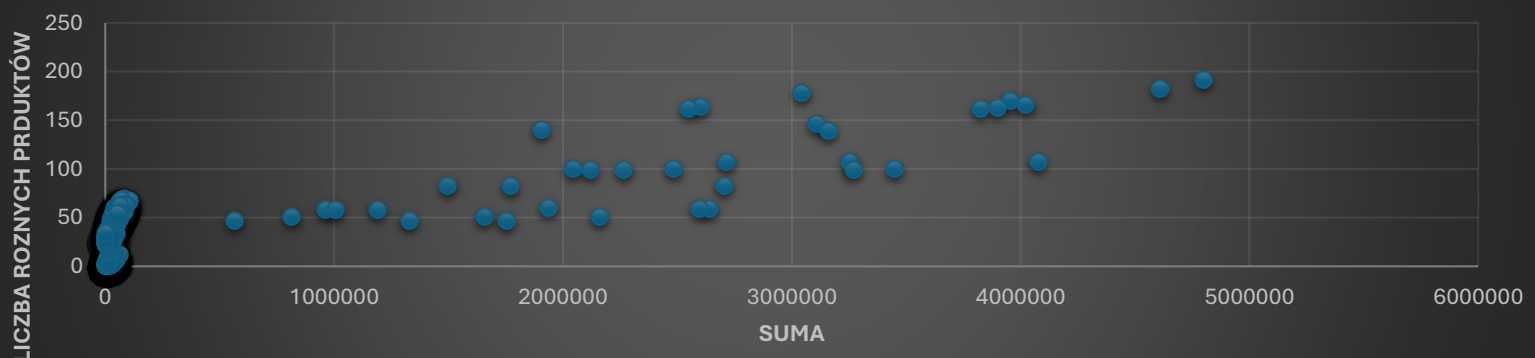
Całkowita wartość sprzedaży w podziale na miesiące i lata



Całkowita wartość sprzedaży w zależności od dnia miesiąca



Liczba różnych produktów od sumy zamówień liczone w obrębie jednego dnia



- Sprzedaż nie jest równomierna w ciągu roku – są miesiące o znacznie wyższych wartościach.
- Widoczne są okresowe wzrosty sprzedaży, szczególnie w końcówkach lat (np. listopad, grudzień).
- Największe skoki sprzedaży wystąpiły w 2014 roku, co może świadczyć o większej liczbie zamówień lub droższych transakcjach.

- Widać wyraźne szczyty sprzedaży w kilku miesiącach, szczególnie pod koniec każdego roku. Możliwe, że firma korzysta z sezonowych promocji lub wpływa na to moc magii świąt.
- Trend zależności liczby produktów od sumy zamówienia jest wzrostowy – intuicyjnie im większa liczba różnych produktów tym większa jest sumaryczna wartość sprzedaży.
- Największa wartość sprzedaży przypada na ostatnie dni miesiąca (głównie 30. i 31. dzień). Może to wynikać z końca okresów rozliczeniowych, rabatów lub zamykania budżetów przez klientów.

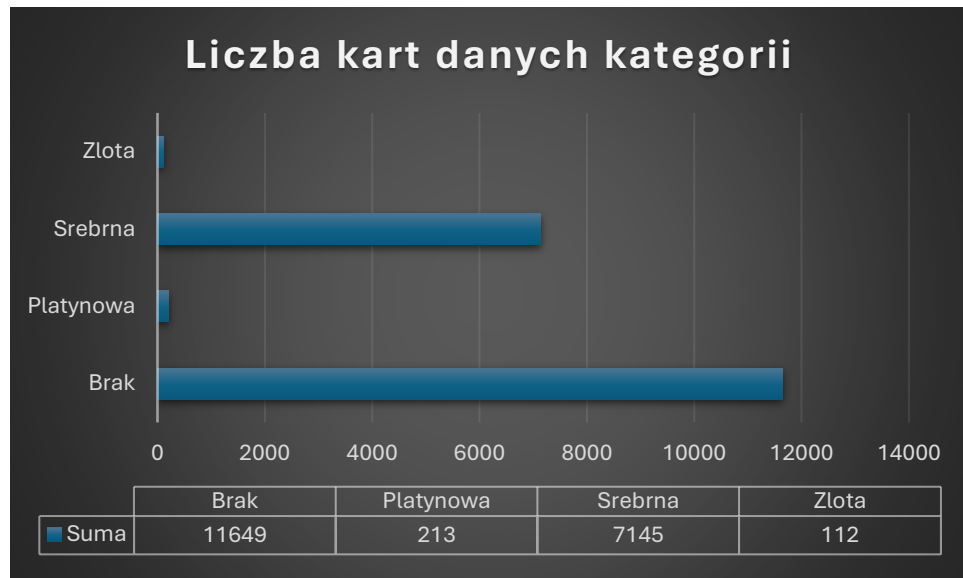
Zadanie 1.4



- Środa i poniedziałek to dni największej różnorodności zamówień

- Weekendowe spadki (sobota, niedziela, piątek) sugerują, że klienci składają mniej zróżnicowane zamówienia, co może wskazywać na dominację popularnych produktów. Sobota ma wysoką sprzedaż, ale niską różnorodność produktów.
- Piątek ma zarówno najniższą wartość sprzedaży, jak i najmniejszą liczbę różnych produktów. Możliwe, że firma zamyka transakcje wcześniej w tygodniu i nie robią dużych zamówień w piątki. Być może klienci odwlekają składania zamówień przed weekendem z wiedząc, że zostaną one wysłane dopiero po weekendzie.

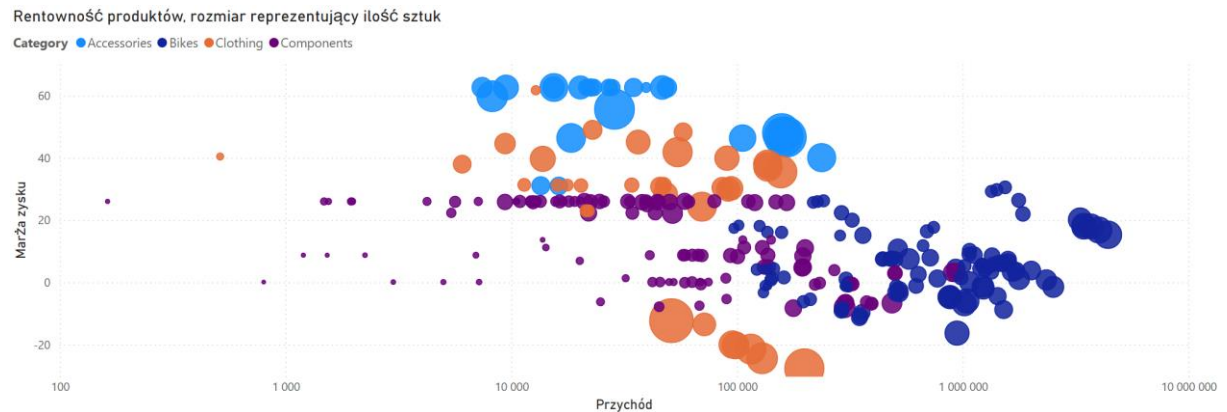
Zadanie 1.5



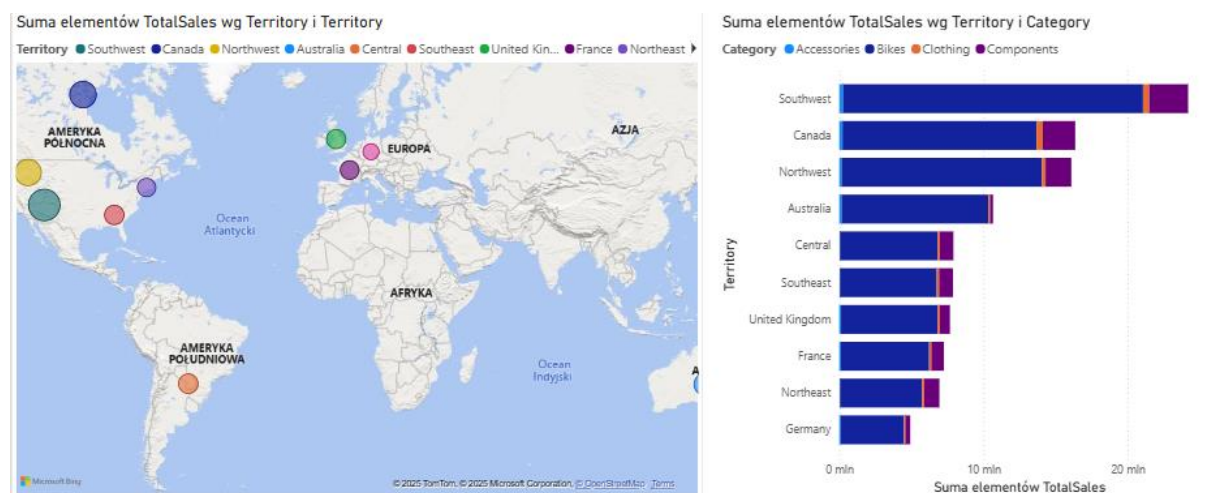
- Bardzo niewielu klientów spełnia warunki platyny, ale ci, którzy ją mają, generują spore przychody, ale
- Większość klientów dokonuje tylko jednej transakcji, co oznacza, że nie wracają do sklepu, ale generują 6% całkowitej wartości sprzedaży, co oznacza, że nawet jednorazowi kupujący zostawiają istotne kwoty.

- Srebrnych klientów jest dużo, ale ich udział w wartości sprzedaży jest stosunkowo niski, co oznacza, że kupują często, ale na mniejsze kwoty.

2. Przygotuj 5 dodatkowych tabel/wykresów, które pokażą ciekawe zależności w bazie AdventureWorks przy użyciu Power BI lub Tableau. Przedstaw wnioski biznesowe wynikające z tych zestawień.

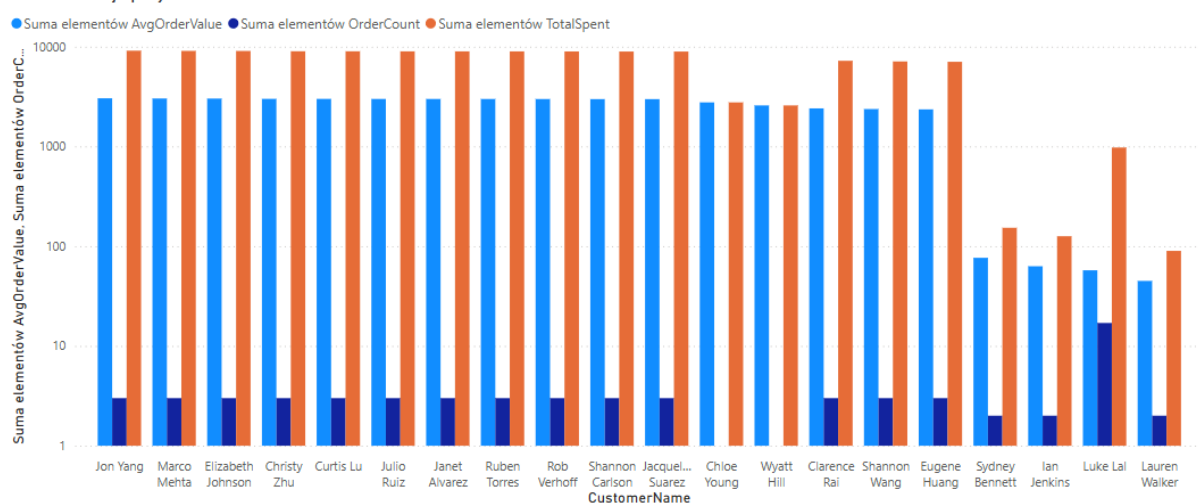


- Rowery generują najwyższe przychody (widoczne po prawej stronie wykresu, w przedziale około 1-10 mln), ale mają zróżnicowaną marżę zysku (od około -20% do +30%).
- Akcesoria mają generalnie wysoką marżę zysku (około 40-60%), ale generują średnie przychody (głównie w przedziale 10-100 tys.).
- Odzież wykazuje dużą zmienność marży zysku (od -30% do +50%) przy średnich przychodach.
- Komponenty mają najniższą marżę zysku (głównie 0-25%) i relatywnie niskie przychody.
- Widoczne są produkty o ujemnej marży co może świadczyć o promocjach lub wyprzedażach magazynowych.
- Największe kółka (oznaczające najwyższą ilość sprzedanych sztuk) można zauważyć głównie wśród akcesoriów i odzieży o średnich przychodach.
- Z perspektywy biznesowej, dane te mogą wskazywać na potrzebę optymalizacji portfolio produktów, ze szczególnym uwzględnieniem tych z ujemną marżą lub niskimi przychodami. Rowery, mimo że generują największe przychody, mogą wymagać analizy kosztów produkcji lub sprzedaży ze względu na zróżnicowaną marżę.



- Największy potencjał sprzedażowy znajduje się w Ameryce Północnej (Southwest, Canada, Northwest).
- We wszystkich regionach rowery stanowią zdecydowanie największe źródło przychodów.
- Rynki o mniejszej sprzedaży (Germany, Northeast, France) mogą stanowić potencjał wzrostu.
- Kategorie Accessories, Clothing i Components mają mniejszy udział w sprzedaży, ale jak wiadomo z powyższych wizualizacji mogą oferować wyższe marże.

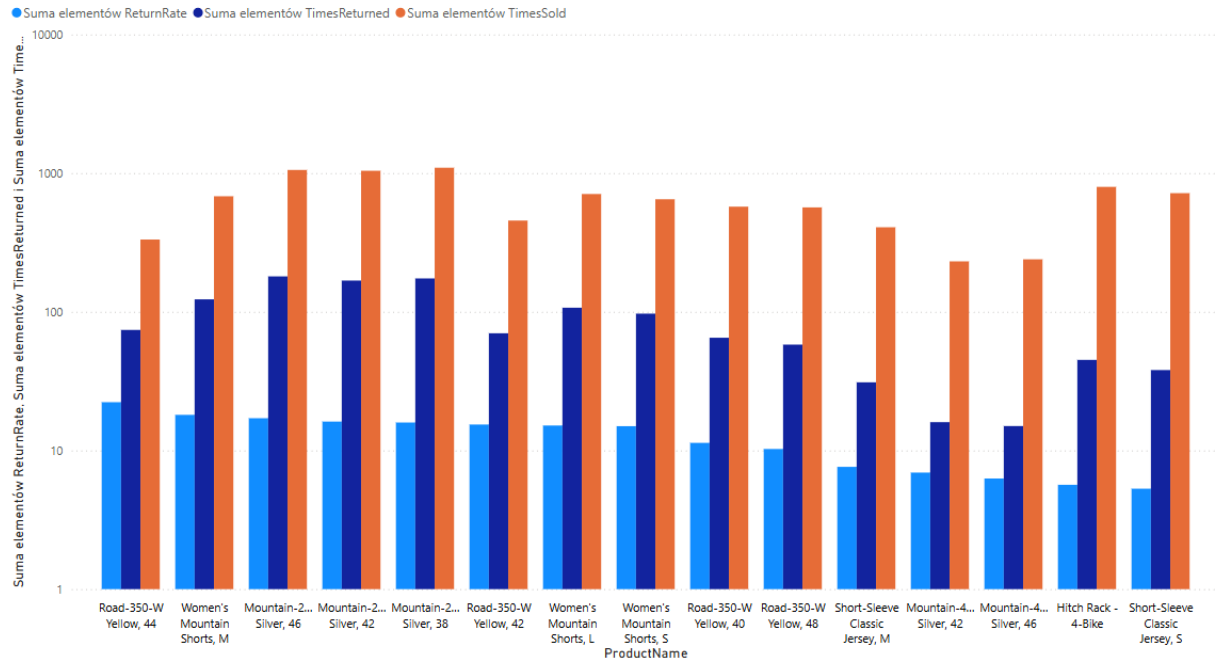
Analiza 20 najlepszych klientów:



- Wśród topowych klientów zauważalne są 3 grupy
 - Grupa premium (od Jon Yang do Eugene Huang) - wysoka wartość zamówień i duża liczba zamówień
 - Grupa średnia (Sydney Bennett, Ian Jenkins, Luke Lai) - średnia wartość zamówień i mniejsza liczba zamówień

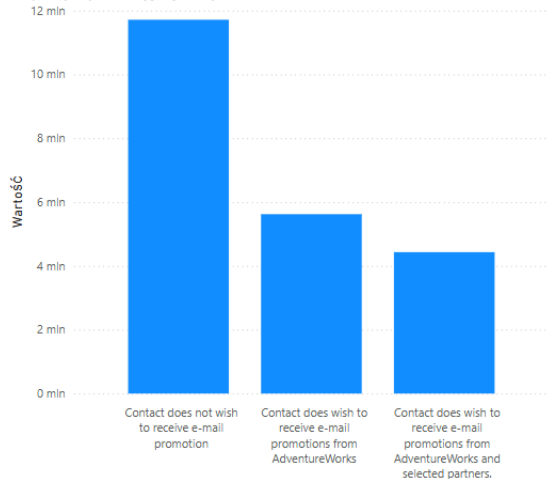
- Grupa niska (Lauren Walker) - najniższe wartości
- Luke Lai ma znacznie wyższą wartość TotalSpent w porównaniu do średniej wartości zamówienia
- Zdecydowanie większą część zasobów marketingowych i obsługi klienta warto przeznaczyć dla grupy premium

Analiza zwrotów pośród 15 najczęściej zwracanych produktów

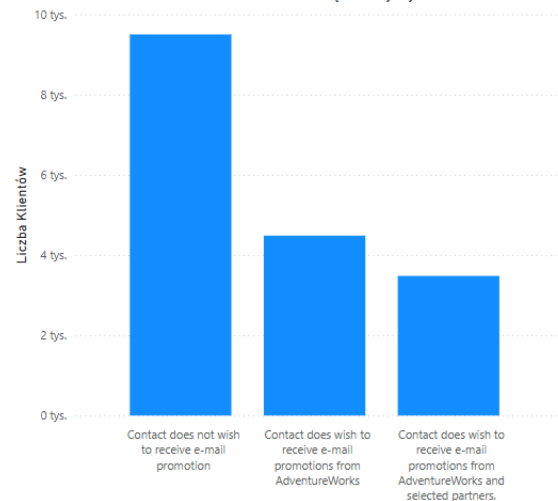


- Rowery Road-350-W Yellow w rozmiarze 44 mają zdecydowanie najwyższy współczynnik zwrotów, co sugeruje potencjalne problemy jakościowe specyficzne dla tego modelu i rozmiaru. Dodatkowo ten sam model w innych rozmiarach pojawia się w zestawieniu czterokrotnie.
- Podobnie, koszulki Short-Sleeve Classic Jersey w rozmiarze S wykazują problemy
- Zauważalny jest wzorec zwrotów dla konkretnych rozmiarów, zwykle najczęściej zwracane są rowery o dużej średnicy koła – powyżej 40 centymetrów.

Sumaryczna wartość sprzedaży w zależności od ustosunkowania klientów do polityki promocyjnej firmy



Liczba klientów w zależności od chęci otrzymywania ofert



- Klienci, którzy nie życzą sobie otrzymywać promocji e-mail, generują zdecydowanie największą wartość sprzedaży - około 11,5 miliona. Wyższa całkowita wartość sprzedaży w tej grupie wynika przede wszystkim z większej liczby klientów, a nie z ich wyższej wartości zakupowej
- Klienci akceptujący promocje od AdventureWorks mają nieznacznie wyższą średnią wartość zakupów od tych akceptujących promocje jedynie od AdventureWorks i partnerów.
- Większość klientów (54%) preferuje brak komunikacji promocyjnej przez e-mail, może stanowić sygnał dla firmy do zmiany sposobu komunikacji lub stosowania zróżnicowanych kanałów komunikacyjnych.
- Wykres ogólnie pokazuje, że różnice w wartościach zakupów, istnieją głównie zważywszy na licznosc danej grupy. Firma może chcieć zatem poprawić strategie marketingowe.