

2. sprawozdanie z laboratorium Hurtownie Danych

Mikołaj Kubś, 272662

12 marca 2025


1 Zadanie 1. Ekstrakcja danych

1.1

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

1.1.1 Wersja bez pivot

```
1 SELECT
2 YEAR(OrderDate) AS "Rok",
3 MONTH(OrderDate) AS "Miesiąc",
4 COUNT(DISTINCT CustomerID) AS "Liczba_różnych_klientów"
5 FROM Sales.SalesOrderHeader
6 GROUP BY YEAR(OrderDate), MONTH(OrderDate)
7 ORDER BY YEAR(OrderDate), MONTH(OrderDate)
```



| | Rok | Miesiąc | Liczba_różnych_klientów |
|----|------|---------|-------------------------|
| 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 |
| 9 | 2012 | 1 | 336 |
| 10 | 2012 | 2 | 219 |
| 11 | 2012 | 3 | 304 |
| 12 | 2012 | 4 | 269 |
| 13 | 2012 | 5 | 293 |
| 14 | 2012 | 6 | 390 |
| 15 | 2012 | 7 | 385 |
| 16 | 2012 | 8 | 285 |
| 17 | 2012 | 9 | 352 |
| 18 | 2012 | 10 | 321 |
| 19 | 2012 | 11 | 383 |

Rysunek 1: Wynik wykonania kwerendy 1

1.1.2 Wersja z użyciem pivot

```
1 WITH UniqueCustomers AS (  
2     SELECT  
3         YEAR(OrderDate) AS OrderYear,  
4         MONTH(OrderDate) AS OrderMonth,  
5         CustomerID  
6     FROM Sales.SalesOrderHeader  
7     GROUP BY YEAR(OrderDate), MONTH(OrderDate), CustomerID  
8 )  
9 SELECT * FROM UniqueCustomers  
10 PIVOT (  
11     COUNT(CustomerID)  
12     FOR OrderMonth IN ([1], [2], [3], [4],  
13                        [5], [6], [7], [8],  
14                        [9], [10], [11], [12])  
15 ) AS PivotTable  
16 ORDER BY OrderYear;
```

| Rok | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|------|------|------|------|------|------|-----|------|------|------|------|------|------|
| 2011 | 0 | 0 | 0 | 0 | 43 | 141 | 231 | 250 | 157 | 327 | 230 | 228 |
| 2012 | 336 | 219 | 304 | 269 | 293 | 390 | 385 | 285 | 352 | 321 | 383 | 378 |
| 2013 | 400 | 325 | 441 | 428 | 426 | 713 | 1675 | 1727 | 1741 | 1893 | 2041 | 1970 |
| 2014 | 2073 | 1713 | 2342 | 2050 | 2350 | 898 | 0 | 0 | 0 | 0 | 0 | 0 |

Rysunek 2: Wynik wykonania kwerendy 1 z pivot

1.2

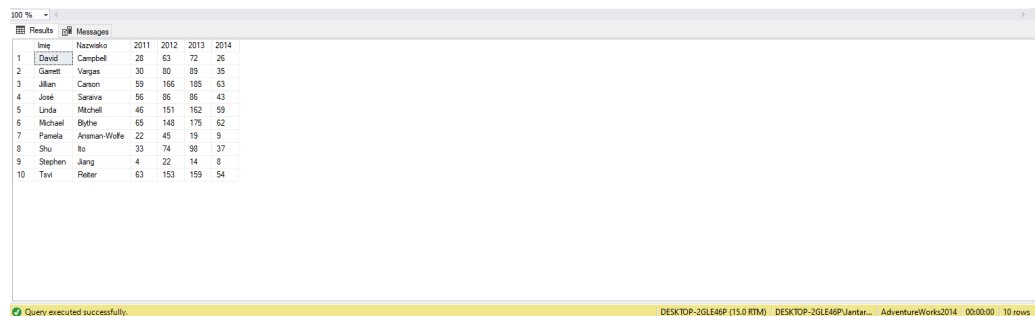
Utworzyć zestawienie zawierające w wierszach imiona i nazwiska sprzedawców, a w kolumnach kolejne lata. Wartością będzie liczba obsłużonych transakcji. Wyświetlić tylko tych sprzedawców, którzy pracowali przez wszystkie 4 lata.

```
1 SELECT * FROM  
2 (  
3     SELECT  
4         FirstName, LastName, SalesOrderID,  
5         YEAR(OrderDate) AS OrderYear FROM Sales.SalesPerson  
6     JOIN HumanResources.Employee ON  
7         Employee.BusinessEntityID = SalesPerson.BusinessEntityID  
8     JOIN Person.Person ON
```

```

9         Person.BusinessEntityID = Employee.BusinessEntityID
10    JOIN Sales.SalesOrderHeader ON
11         SalesOrderHeader.SalesPersonID = SalesPerson.BusinessEntityID
12    WHERE YEAR(HireDate) = 2011
13 ) AS SourceTable
14 PIVOT (
15     COUNT(SalesOrderID)
16     FOR OrderYear IN ([2011], [2012], [2013], [2014])
17 ) AS PivotedTable
18 ORDER BY FirstName

```



| Imię | Nazwisko | 2011 | 2012 | 2013 | 2014 |
|---------|--------------|------|------|------|------|
| David | Campbell | 28 | 63 | 72 | 26 |
| Garret | Vargas | 30 | 80 | 89 | 35 |
| Allan | Carson | 59 | 166 | 125 | 63 |
| Jose | Saravia | 56 | 86 | 86 | 43 |
| Linda | Mitchell | 46 | 151 | 162 | 59 |
| Michael | Blythe | 65 | 148 | 175 | 62 |
| Pamela | Ansman-Wolfe | 22 | 45 | 19 | 9 |
| Shu | Itō | 33 | 74 | 98 | 37 |
| Stephen | Jiang | 4 | 22 | 14 | 8 |
| Tavi | Reiter | 63 | 153 | 159 | 54 |

Rysunek 3: Wynik wykonania kwerendy 2

1.3

Zdefiniować zapytanie wyznaczające sumę kwot sprzedaży towarów oraz liczbę różnych produktów w zamówieniach w poszczególnych latach, miesiącach, dniach.

```

1 SELECT
2     YEAR(OrderDate) AS "Rok",
3     MONTH(OrderDate) AS "Miesiąc",
4     DAY(OrderDate) AS "Dzień",
5     SUM(LineTotal) AS "Suma",
6     COUNT(DISTINCT ProductID) AS "Liczba_różnych_produktyw"
7 FROM Sales.SalesOrderHeader
8 JOIN Sales.SalesOrderDetail ON
9     SalesOrderDetail.SalesOrderID = SalesOrderHeader.SalesOrderID
10 GROUP BY YEAR(OrderDate), MONTH(OrderDate), DAY(OrderDate)
11 ORDER BY YEAR(OrderDate), MONTH(OrderDate), DAY(OrderDate)

```

| | Rok | Miesiąc | Dzień | Suma | Liczba różnych produktów |
|----|------|---------|-------|---------------|--------------------------|
| 1 | 2011 | 5 | 31 | 503805.916900 | 47 |
| 2 | 2011 | 6 | 1 | 13931.520000 | 4 |
| 3 | 2011 | 6 | 2 | 15012.178200 | 4 |
| 4 | 2011 | 6 | 3 | 7156.540000 | 2 |
| 5 | 2011 | 6 | 4 | 15012.178200 | 4 |
| 6 | 2011 | 6 | 5 | 14313.080000 | 3 |
| 7 | 2011 | 6 | 6 | 7855.638200 | 3 |
| 8 | 2011 | 6 | 7 | 7855.638200 | 2 |
| 9 | 2011 | 6 | 8 | 20909.780000 | 5 |
| 10 | 2011 | 6 | 9 | 10556.530000 | 3 |
| 11 | 2011 | 6 | 10 | 14313.080000 | 3 |
| 12 | 2011 | 6 | 11 | 14134.800000 | 4 |
| 13 | 2011 | 6 | 12 | 7156.540000 | 2 |
| 14 | 2011 | 6 | 13 | 25047.890000 | 4 |
| 15 | 2011 | 6 | 14 | 11230.628200 | 4 |
| 16 | 2011 | 6 | 15 | 14313.080000 | 2 |
| 17 | 2011 | 6 | 16 | 14134.800000 | 4 |
| 18 | 2011 | 6 | 17 | 6953.260000 | 2 |
| 19 | 2011 | 6 | 18 | 25568.708200 | 5 |

Rysunek 4: Wynik wykonania kwerendy 3

1.4

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. Uwaga: Pamiętaj o wybraniu właściwego atrybutu funkcji datepart tak, aby zgadzała się nazwa dnia tygodnia

```

1 SET DATEFIRST 1;
2 SET LANGUAGE Polish;
3
4 SELECT
5     YEAR(OrderDate) AS "Rok",
6     DATENAME(month, OrderDate) AS "Miesiąc",
7     CASE DATEPART(dw, OrderDate)
8         WHEN 1 THEN 'Poniedziałek'
9         WHEN 2 THEN 'Wtorek'
10        WHEN 3 THEN 'Środa'
11        WHEN 4 THEN 'Czwartek'
12        WHEN 5 THEN 'Piątek'
13        WHEN 6 THEN 'Sobota'
14        WHEN 7 THEN 'Niedziela'
15    END AS "Dzień_tygodnia",
16     SUM(LineTotal) AS "Suma",
17     COUNT(DISTINCT ProductID) AS "Liczba_różnych_produktyw"
18 FROM Sales.SalesOrderHeader
19 JOIN Sales.SalesOrderDetail ON
20     SalesOrderDetail.SalesOrderID = SalesOrderHeader.SalesOrderID
21 GROUP BY

```

```

22     YEAR(OrderDate),
23     DATENAME(month, OrderDate),
24     MONTH(OrderDate),
25     DATEPART(dw, OrderDate)
26 ORDER BY
27     YEAR(OrderDate),
28     MONTH(OrderDate),
29     DATEPART(dw, OrderDate)

```

| Rok | Miesiac | Dzień tygodnia | Suma | Liczba różnych produktów |
|------|----------|----------------|----------------|--------------------------|
| 2011 | maj | Worek | 503805.916900 | 47 |
| 2011 | czerwiec | Poniedziałek | 67001.972800 | 9 |
| 2011 | czerwiec | Worek | 75015.626400 | 10 |
| 2011 | czerwiec | Sroda | 78568.896400 | 12 |
| 2011 | czerwiec | Czwartek | 65450.496400 | 8 |
| 2011 | czerwiec | Piatek | 39856.782200 | 7 |
| 2011 | czerwiec | Sobota | 72250.476400 | 10 |
| 2011 | czerwiec | Niedziela | 60786.562200 | 8 |
| 2011 | lipiec | Poniedziałek | 64344.832200 | 10 |
| 2011 | lipiec | Worek | 78981.692200 | 10 |
| 2011 | lipiec | Sroda | 65743.034600 | 11 |
| 2011 | lipiec | Czwartek | 63799.020000 | 7 |
| 2011 | lipiec | Piatek | 1626141.704338 | 49 |
| 2011 | lipiec | Sobota | 59864.190000 | 8 |
| 2011 | lipiec | Niedziela | 85725.436400 | 11 |
| 2011 | sierpień | Poniedziałek | 1227426.765976 | 49 |
| 2011 | sierpień | Worek | 133209.346400 | 9 |
| 2011 | sierpień | Sroda | 902836.672670 | 58 |
| 2011 | sierpień | Czwartek | 49739.220000 | 6 |

Rysunek 5: Wynik wykonania kwerendy 4

1.5

Przygotować zestawienie, w którym dla wybranych klientów przygotujemy kartę lojalnościową:

- srebrną, jeśli klient wykonał co najmniej 2 transakcje w sklepie;
- 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;
- platynową, jeśli klient spełniał warunki otrzymania karty złotej oraz w co najmniej jednej transakcji kupił jednocześnie produkty ze wszystkich kategorii

```

1 WITH AvgOrderValue AS (
2     SELECT AVG(TotalOrderValue) AS AvgValue
3     FROM (
4         SELECT SalesOrderID, SUM(LineTotal) AS TotalOrderValue
5         FROM Sales.SalesOrderDetail
6         GROUP BY SalesOrderID
7     ) AS OrderValues
8 ),

```

```

9 OrderCount AS (
10     SELECT CustomerID, COUNT(DISTINCT SalesOrderID) AS TransactionCount,
11           SUM(TotalDue) AS TotalTransactionValue
12     FROM Sales.SalesOrderHeader
13     GROUP BY CustomerID
14 ),
15 HighValueOrders AS (
16     SELECT CustomerID, COUNT(*) AS HighValueOrderCount
17     FROM (
18         SELECT SalesOrderID, SUM(LineTotal) AS TotalOrderValue
19         FROM Sales.SalesOrderDetail
20         GROUP BY SalesOrderID
21     ) AS OrderValues
22     JOIN Sales.SalesOrderHeader
23         ON SalesOrderHeader.SalesOrderID = OrderValues.SalesOrderID
24     CROSS JOIN AvgOrderValue A
25     WHERE TotalOrderValue > 2.5 * A.AvgValue
26     GROUP BY CustomerID
27 ),
28 UniqueCategories AS (
29     SELECT
30         C.CustomerID,
31         COUNT(DISTINCT PC.ProductCategoryID) AS UniqueCategories
32     FROM Sales.Customer C
33     JOIN Sales.SalesOrderHeader SOH
34         ON SOH.CustomerID = C.CustomerID
35     JOIN Sales.SalesOrderDetail SOD
36         ON SOD.SalesOrderID = SOH.SalesOrderID
37     JOIN Production.Product PR ON PR.ProductID = SOD.ProductID
38     JOIN Production.ProductSubcategory PSC
39         ON PSC.ProductSubcategoryID = PR.ProductSubcategoryID
40     JOIN Production.ProductCategory PC
41         ON PC.ProductCategoryID = PSC.ProductCategoryID
42     GROUP BY C.CustomerID
43 )
44 SELECT
45     P.FirstName AS Imie,
46     P.LastName AS Nazwisko,
47     COALESCE(OrderCount.TransactionCount, 0) AS "Liczba_transakcji",
48     COALESCE(OrderCount.TotalTransactionValue, 0) AS "Łączna_kwota_transakcji",
49     CASE

```

```

50      WHEN COALESCE(OrderCount.TransactionCount, 0) >= 4
51          AND COALESCE(HighValueOrders.HighValueOrderCount, 0) >= 2
52          AND COALESCE(UniqueCategories.UniqueCategories, 0) =
53              (SELECT COUNT(*) FROM Production.ProductCategory)
54          THEN 'Platynowa'
55      WHEN COALESCE(OrderCount.TransactionCount, 0) >= 4
56          AND COALESCE(HighValueOrders.HighValueOrderCount, 0) >= 2
57          THEN 'Złota'
58      WHEN COALESCE(OrderCount.TransactionCount, 0) >= 2
59          THEN 'Srebrna'
60      ELSE 'Brak_karty'
61  END AS "Kolor_karty"
62  FROM Sales.Customer C
63  JOIN Person.Person P ON P.BusinessEntityID = C.PersonID
64  LEFT JOIN HighValueOrders ON HighValueOrders.CustomerID = C.CustomerID
65  LEFT JOIN OrderCount ON OrderCount.CustomerID = C.CustomerID
66  LEFT JOIN UniqueCategories ON UniqueCategories.CustomerID = C.CustomerID
67  ORDER BY OrderCount.TotalTransactionValue DESC;

```

| | Imię | Nazwisko | Liczba transakcji | Łączna kwota transakcji | Kolor karty |
|-----|----------|-----------|-------------------|-------------------------|-------------|
| 307 | Pei | Chow | 4 | 61333.7327 | Platynowa |
| 308 | Adam | Barr | 8 | 60010.9459 | Platynowa |
| 309 | Nicky | Cheesnut | 4 | 59709.5091 | Złota |
| 310 | Jane | Greer | 8 | 59808.1531 | Platynowa |
| 311 | Calla | Edridge | 12 | 59418.5364 | Srebrna |
| 312 | Robert | Stolka | 8 | 55823.3779 | Srebrna |
| 313 | Frances | Giglio | 8 | 55608.3548 | Platynowa |
| 314 | Kathleen | Harding | 4 | 54905.1516 | Złota |
| 315 | Kristine | Olney | 4 | 53526.5664 | Platynowa |
| 316 | Pat | Coleman | 4 | 52742.2969 | Złota |
| 317 | John | Berger | 9 | 51533.2559 | Złota |
| 318 | Jim | Stewart | 8 | 51183.5661 | Srebrna |
| 319 | Walter | Mays | 8 | 51015.5565 | Platynowa |
| 320 | Jovita | Carmody | 4 | 50997.8805 | Złota |
| 321 | Leo | Gaskoun | 8 | 50673.1856 | Srebrna |
| 322 | Yao-Qi | Cheng | 8 | 50149.2315 | Srebrna |
| 323 | David | Birk | 8 | 50117.8904 | Srebrna |
| 324 | Douglas | Groncki | 4 | 49614.9179 | Złota |
| 325 | Ruth | Choi | 7 | 49438.7051 | Platynowa |
| 326 | Kyle | Abdelaziz | 8 | 48061.4725 | Srebrna |
| 327 | Gina | Clark | 4 | 48054.0631 | Złota |
| 328 | Alice | Clark | 4 | 47094.0902 | Złota |
| 329 | Aaron | Con | 12 | 46795.976 | Srebrna |
| 330 | Michael | Brundage | 8 | 46391.1048 | Srebrna |
| 331 | Richard | Iwini | 6 | 45650.2544 | Złota |
| 332 | Gary | Dunry | 12 | 45348.5877 | Srebrna |
| 333 | Eva | Coneta | 8 | 45316.8917 | Platynowa |
| 334 | Mary | Alexander | 8 | 45166.9405 | Srebrna |
| 335 | Comett | Gibbens | 4 | 44581.5065 | Platynowa |
| 336 | Susan | Cheesnut | 4 | 44104.1959 | Złota |
| 337 | Dev | Desilvo | 8 | 44055.4622 | Srebrna |
| 338 | Kay | Krane | 12 | 43717.8005 | Srebrna |
| 339 | Jeanette | Cole | 11 | 41979.0807 | Srebrna |
| 340 | Eric | Bumfield | 4 | 41264.0391 | Złota |
| 341 | Mindy | Martin | 8 | 40188.4765 | Srebrna |

Rysunek 6: Wynik wykonania kwerendy 5

2 Zadanie 2. Analiza danych

2.1

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

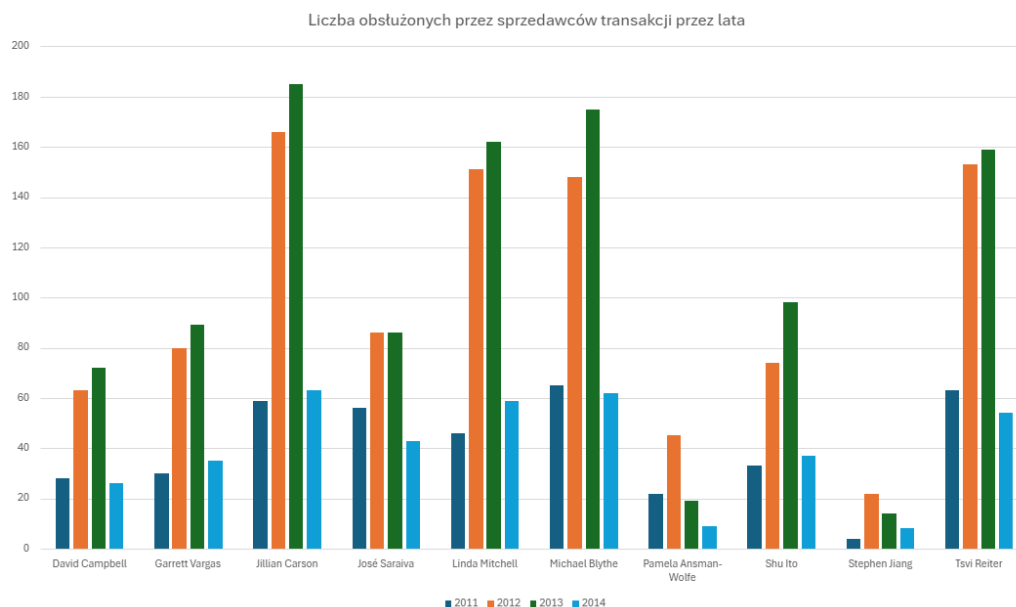
2.1.1

| RokMiesiąc | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | Suma końcowa |
|--------------|------|------|------|------|------|------|------|------|------|------|------|------|--------------|
| 2011 | | | | | 43 | 141 | 231 | 250 | 157 | 327 | 230 | 228 | 1607 |
| 2012 | 336 | 219 | 304 | 269 | 293 | 390 | 385 | 285 | 352 | 321 | 383 | 378 | 3915 |
| 2013 | 400 | 325 | 441 | 428 | 426 | 713 | 1675 | 1727 | 1741 | 1893 | 2041 | 1970 | 13780 |
| 2014 | 2073 | 1713 | 2342 | 2058 | 2350 | 898 | | | | | | | 11434 |
| Suma końcowa | 2809 | 2257 | 3087 | 2755 | 3112 | 2142 | 2291 | 2262 | 2250 | 2541 | 2654 | 2576 | 30736 |

Rysunek 7: Tabela przestawna na podstawie wyników kwerendy 1

Można zauważyć stabilny trend wzrostowy w liczbie klientów wraz z czasem. W lipcu 2013 roku do firmy przybyło aż prawie 900 klientów, co było zdecydowanie największym wzrostem. Kolejną anomalią jest lipiec 2014 roku. Mimo posiadania danych z całego miesiąca, liczba różnych klientów zmniejszyła się ponad dwukrotnie i osiągnęła najniższy poziom od czerwca 2013 roku. Dane za cały miesiąc były niskie. Nie było też żadnego dnia, w którym sklep miał ponad dwukrotność średniej klientów na miesiąc, jak to bywało w innych miesiącach.

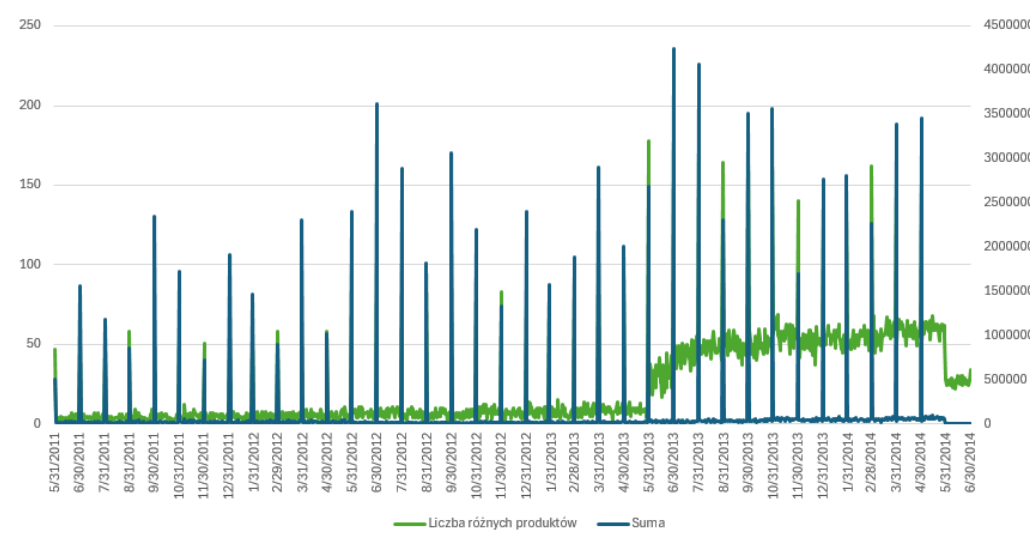
2.1.2



Rysunek 8: Wykres na podstawie wyników kwerendy 2

Różnice między liczbą obsługiwanych klientów między sprzedawcami są bardzo duże. Jest 4 sprzedawców, którzy pozytywnie wyróżniają się na tle reszty, i dwóch sprzedawców w odwrotnej sytuacji. Na tym wykresie również można zauważyć potwierdzony w innych kwerendach trend wzrostowy sklepu w liczbie klientów. Jak widać na wykresie, dla prawie każdego sprzedawcy liczba obsługiwanych klientów w 2012 lub 2013 jest znacznie wyższa niż w 2011. 2014 jeszcze się nie skończył i jest nadzieja na dobry sezon wakacyjny, dzięki któremu poziom poprzednich 2 lat może być utrzymany.

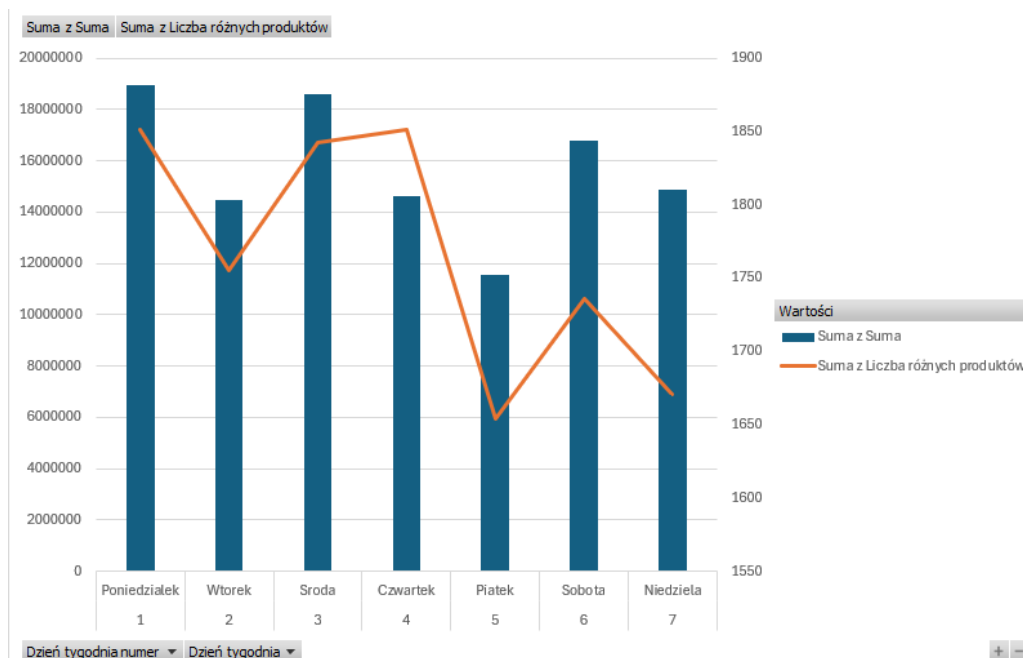
2.1.3



Rysunek 9: Wykres na podstawie wyników kwerendy 3

Od razu zauważalna jest pewna niespójność w danych - mniej więcej co miesiąc (na samym początku lub końcu miesiąca) liczba sprzedanych różnych produktów i suma sprzedaży osiągają wielokrotność wyników w innych dniach. Poza tym widać trend wzrostowy, wraz z nagłym silnym wzrostem sprzedaży, który został zauważony w poprzednich analizach. Tak jak w poprzednich analizach, ostatni miesiąc ma o około połowę niższą sprzedaż niż poprzednie.

2.1.4



Rysunek 10: Wykresy na podstawie wyników kwerendy 4

Wykres pokazuje, że nie ma bezpośredniej korelacji między liczbą różnych produktów a sumą sprzedaży. Na przykład niedziela charakteryzuje się podobną liczbą różnych produktów co piątek, ale wyraźnie wyższą wartością sprzedaży.

Piątek wyróżnia się najniższą liczbą produktów i najniższą sprzedażą, podczas gdy poniedziałek i środa osiągają najwyższe wartości. W weekend sobota generuje wyższe przychody i większą liczbę sprzedanych produktów niż niedziela.

2.1.5



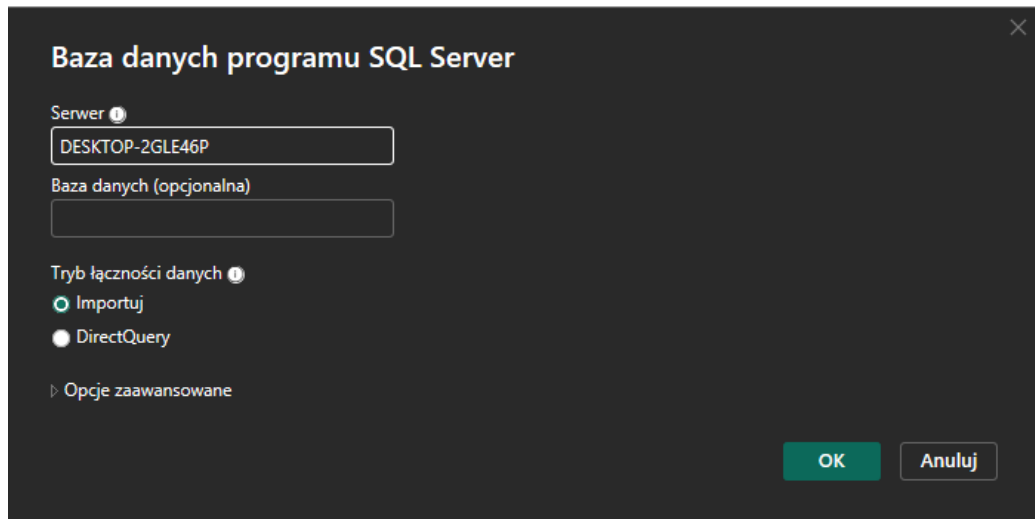
Rysunek 11: Wykres na podstawie wyników kwerendy 5

Z wykresu drzewa można wyczytać, że największa suma wartości przypada na karty platynowe, które zajmują największą powierzchnię na wykresie. Karty srebrne są drugie pod względem sumy, natomiast karty złote i brak karty mają mniejsze wartości, przy czym "brak karty" ma najmniejszy udział. Wykres ten dobrze obrazuje proporcje między kategoriami, pokazując, że klienci z platynowymi kartami mają największy udział w przychodach firmy. Z tego wynika, że za większość obrotów sklepu odpowiadają stali klienci. Przeprowadzili co najmniej 4 razy transakcję, przynajmniej raz bardzo różnorodne zakupy i przynajmniej 2 razy zakupy o wartości znacznie przewyższającej średnią. Złota karta odpowiada za stan przejściowy między srebrną a platynową, co sugeruje możliwość zmiany wielu klientów w tych najbardziej dochodowych. Co ciekawe, dzieje się tak, mimo, iż liczba transakcji klientów z kartą platynową jest ponad 10 razy mniejsza od analogicznej wartości dla klientów z kartą srebrną. Pokazuje to, że liczba transakcji niekoniecznie wyznacza ich łączną kwotę.

Trzeba pamiętać, że niekoniecznie musi istnieć przyczynowość między kolorem karty a wydatkami klientów. Im wyższy "poziom" karty, tym większe wymagania wobec klientów, naturalnie sprawiające, że do np. platynowej karty zgrupowani zostali klienci, którzy najwięcej wydali.

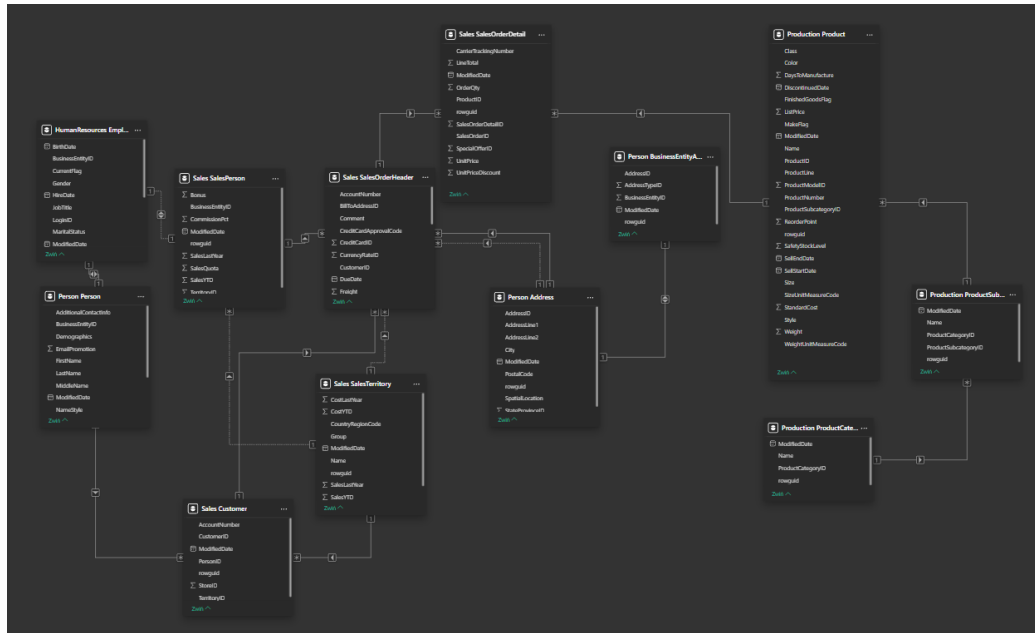
2.2

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



Rysunek 12: Połączenie z Power BI

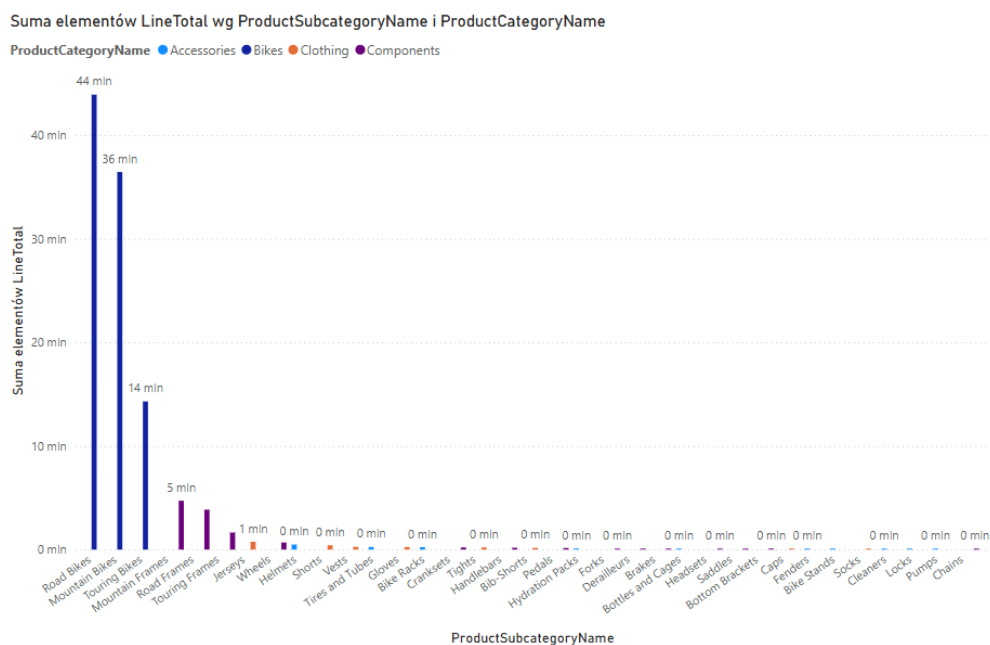
2.2.1



Rysunek 13: Model danych w Power BI

Na wykresie jasno widać, w których państwach sprzedawane są artykuły sklepu. Dodatkowo można porównać wielkości wykresów kołowych, aby zobaczyć wysokość sprzedaży z poprzedniego roku. Wykres kołowy w USA jest podzielony dodatkowo na 5 terytori, typu Northeast, Southeast itd. Jak łatwo można zauważyć, największa suma sprzedaży jest w Stanach Zjednoczonych i Kanadzie, a najmniejsza w Niemczech.

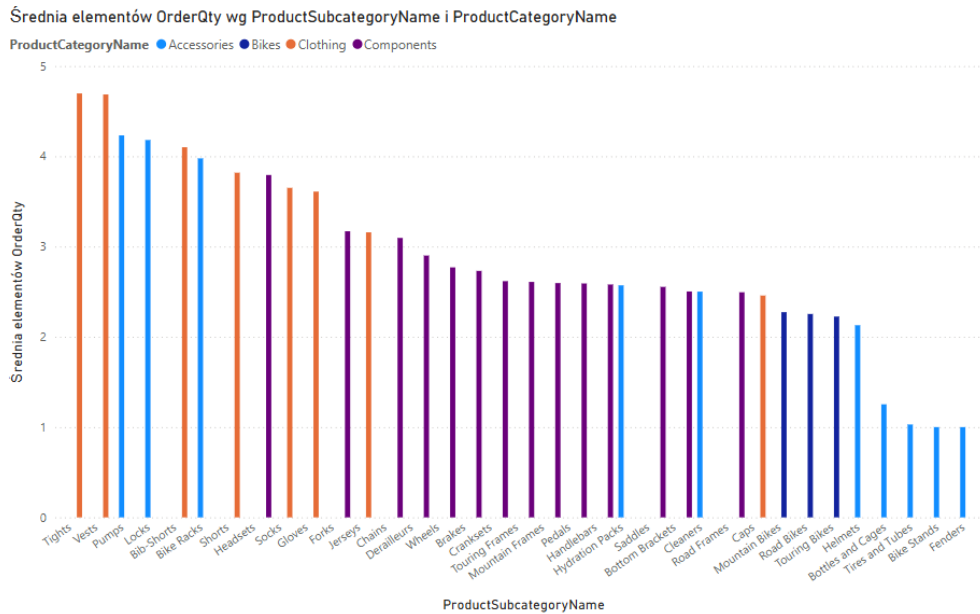
2.2.2



Rysunek 14: 2. wykres w Power BI

Sklep uzyskiwał największą sumę przychodów ze sprzedaży różnego rodzaju rowerów. Następne są różne komponenty, ale przychody z nich były znacznie mniejsze. Dzięki temu wykresowi jasno widać, z czego sklep historycznie czerpał jaką część przychodów.

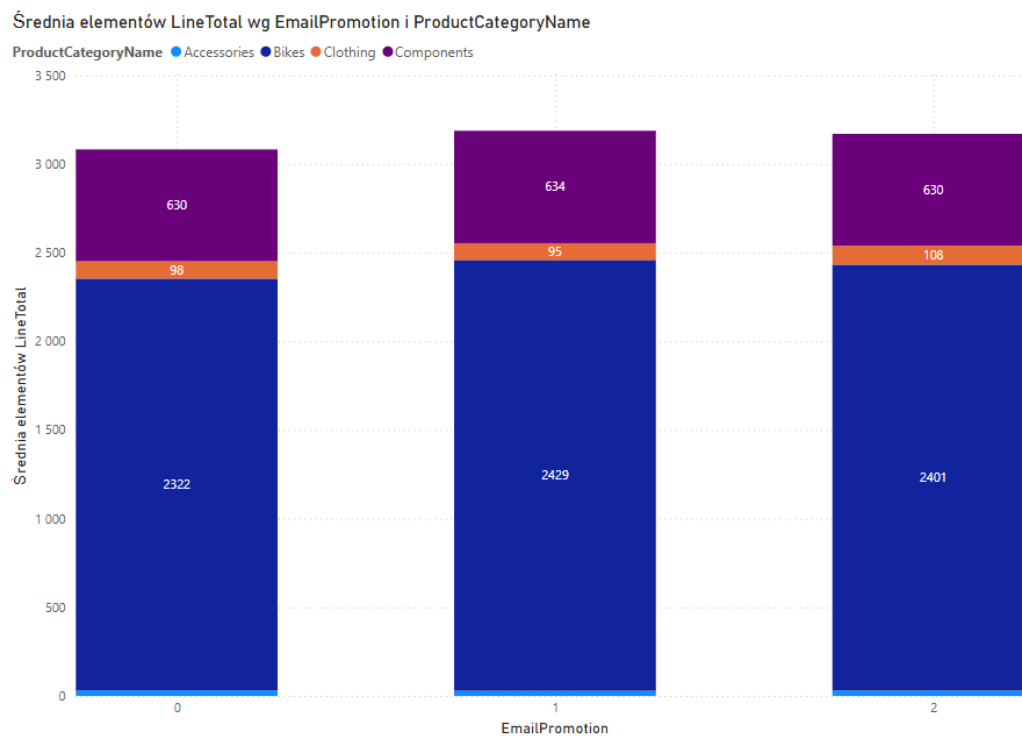
2.2.3



Rysunek 15: 3. wykres w Power BI

Na tym wykresie można zauważyć ciekawe zależności. Mimo, iż sklep większość przychodów uzyskuje ze sprzedaży rowerów, średnia liczba kupionych tych samych rowerów nie jest najwyższa. Jak widać na wykresie, dla danego produktu w danej transakcji klienci kupują średnio najwięcej elementów ubioru lub akcesorii, takich jak rajstopy, kamizelki czy pompki. Być może można by ustalić kandydatów na oferty z obniżoną ceną przy wielosztukowaniu na podstawie tego wykresu.

2.2.4



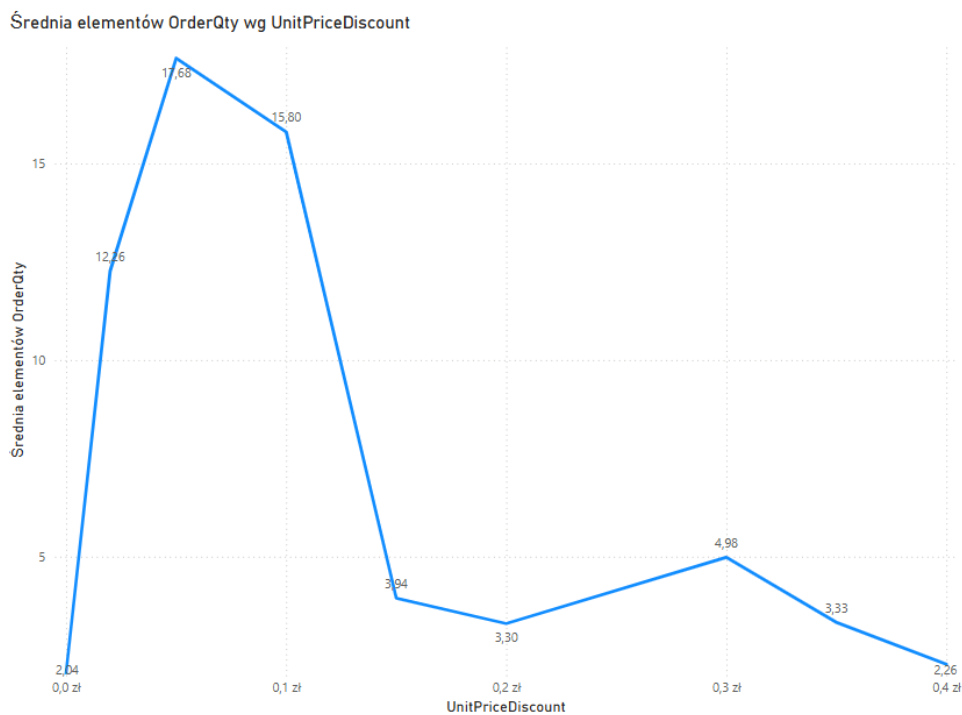
Rysunek 16: 4. wykres w Power BI

Według strony dataedo.com [1] EmailPromotion oznacza:

"0 = Contact does not wish to receive e-mail promotions, 1 = Contact does wish to receive e-mail promotions from AdventureWorks, 2 = Contact does wish to receive e-mail promotions from AdventureWorks and selected partners."

Brakuje korelacji między EmailPromotion a średnią sprzedażą dla danego klienta, co sugeruje, że promocje e-mail nie są specjalnie efektywne.

2.2.5



Rysunek 17: 5. wykres w Power BI

Brakuje zależności między średnią liczbą zamówionego produktu, a rabatem na niego przeznaczonym. Wykres co prawda na początku rośnie, ale potem szybko spada. Ostatecznie średnia liczba zamówionego produktu jest bardzo podobna dla zerowego rabatu, jak i dla maksymalnego rabatu. Być może wysokie rabaty mają klauzulę ograniczającą, na ile maksymalnie produktów jest rabat.

3 Wnioski

Na podstawie przeprowadzonych analiz można sformułować następujące wnioski:

1. **Trend wzrostowy i sezonowość:** Analiza danych wykazała stabilny trend wzrostowy w liczbie klientów, co potwierdzają zarówno wyniki kwerendy 1, jak i wizualizacje w MS Excel. Szczególnie zauważalny jest

gwałtowny wzrost w lipcu 2013 roku oraz wyraźna anomalia w lipcu 2014 - miesiącu, w którym liczba klientów spadła znacząco. Może to sugerować wystąpienie czynników sezonowych lub operacyjnych, które warto zbadać w dalszych analizach.

2. **Różnice w efektywności sprzedaży:** Zestawienie sprzedażowe według sprzedawców (kwerenda 2) ujawnia duże zróżnicowanie wyników. Kilku sprzedawców wyróżnia się bardzo dobrymi rezultatami, podczas gdy inni osiągają znacznie niższe wyniki. Wskazuje to na potrzebę dalszej analizy efektywności poszczególnych pracowników oraz rozważenie wdrożenia szkoleń, aby wyrównać wyniki zespołu, zmiany pracowników lub dalszej analizy przyczyn.
3. **Zależności między ofertą a sprzedażą:** Wyniki kwerend 3 i 4 pokazały, że między liczbą różnych produktów a sumą sprzedaży nie występuje bezpośrednia korelacja. Analiza według dni tygodnia wskazuje, że poszczególne dni charakteryzują się odmienną intensywnością sprzedaży - najwyższe wartości odnotowano w poniedziałki i środy, a najniższe w piątki. Może to stanowić podstawę do optymalizacji harmonogramu promocji i działań marketingowych.
4. **Znaczenie lojalności klientów:** Przygotowana karta lojalnościowa (kwerenda 5) umożliwia segmentację klientów na podstawie liczby transakcji, wartości zakupów oraz różnorodności kategorii produktów. Wyniki wskazują, że stali klienci, czyli klienci z platynowymi kartami generują największy obrót, co podkreśla wagę budowania długoterminowych relacji z klientami. Segmentacja ta może posłużyć do opracowania dalszych programów lojalnościowych. Trzeba mieć na uwadze, że karty grupują klientów o odpowiednio większych wydatkach, co może oznaczać brak przyczynowości między kolorem karty a wydatkami, a inne, wspólne przyczyny.
5. **Wnioski z analiz w Power BI:**
 - **Geograficzna dystrybucja sprzedaży:** Wizualizacja danych geograficznych pokazuje, że największe obroty generowane są w USA i Kanadzie, co może wpłynąć na decyzje dotyczące ekspansji rynkowej lub skoncentrowania działań marketingowych w tych regionach.
 - **Struktura sprzedaży produktów:** Największe przychody pochodzą z kategorii rowerów, jednak analiza ilościowa zakupów wskazuje, że produkty ubieralne i akcesoria kupowane są w większych

ilościach. Otwiera to możliwości wprowadzenia ofert promocyjnych przy zakupach wieloelementowych.

- **Skuteczność promocji e-mailowych:** Brak wyraźnej korelacji między ustawieniem opcji `EmailPromotion` a średnią sprzedażą sugeruje, że obecne kampanie e-mailowe mogą wymagać rewizji lub lepszego targetowania.
- **Strategia rabatowa:** Analiza zależności między średnią ilością zamówionych produktów a przyznawanym rabatem wskazuje, że rabaty nie wpływają znacząco na wielkość zamówień, co może sugerować potrzebę modyfikacji polityki rabatowej lub wprowadzenia dodatkowych ograniczeń.

Podsumowując, przedstawione rozwiązania umożliwiły bardziej dogłębną analizę danych sprzedażowych oraz identyfikację kluczowych czynników wpływających na wyniki firmy. Wnioski te mogą stanowić solidną podstawę do podejmowania decyzji biznesowych, optymalizacji strategii sprzedażowych oraz doskonalenia działań marketingowych. Oczywiście, nadal pozostaje wiele działów, czynników i przyczyn do zbadania.

References

- [1] *AdventureWorks Person Table*. Accessed: 2025-03-12. URL: https://dataedo.com/samples/html/AdventureWorks/doc/AdventureWorks_2/tables/Person_Person_144.html.