



**Wojskowa Akademia Techniczna
Wydział Cybernetyki
Instytut Systemów Informatycznych**

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych

Wersja 3.0

Autorzy: Alicja Szmigiel, Adrianna Bocial, Adam Mikołajewski

Nazwa grupy: WCY22IJ1S1

Numer podgrupy: 5

Prowadzący laboratorium: dr inż. Stefan Rozmus

Warszawa, 2024/2025 r.

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Metryka dokumentu			
Nazwa dokumentu	Hurtownie danych – sprawozdanie z zajęć laboratoryjnych.		
Opis dokumentu	Opis przebiegu wykonanych prac i uzyskanych wyników w ramach realizowanego projektu z przedmiotu Hurtownie danych.		
Autorzy	Alicja Szmigiel, Adrianna Bocial, Adam Mikołajewski AS, AB, AM		
Nazwa pliku	HD-LAB-WCY22IJ1S1_5_v3.0	Liczba stron	29

Historia zmian dokumentu						
Wersja	Data wydania	Opis	Akcja (*)	Rozdziały (**)	Autorzy (***)	Data kontroli
0.1	2024-11-14	Utworzenie dokumentu	N	W	AM	nd
0.2	2024-11-14	Korekta dokumentu Wypełnienie rozdziału 3	W	2, 3	AM, AS, AB	nd
0.3	2024-11-19	Korekta dokumentu, Wypełnienie rozdziału 4	Z, W	3,4	AM, AS, AB	nd
1.0	2024-11-21	Korekta dokumentu	W	1, 2, 3, 4	AM	nd
1.0	2024-11-21	Kontrola jakości dokumentu	We	W	Stefan Rozmus	2024-11-22
1.1	2024-11-27	Korekta rozdziału 4.	Z	4	AM, AS, AB	nd
1.2	2024-12-02	Korekta dokumentu.	Z, W	3, 4	AM, AS, AB	nd
1.3	2024-12-06	Korekta dokumentu. Wypełnienie rozdziału 5 i 6.	W	5, 6	AM, AS, AB	nd
2.0	2024-12-09	Korekta dokumentu	Z	W	AM	Nd
2.0	2024-12-09	Kontrola jakości dokumentu	We	W	Stefan Rozmus	2024-12-10
2.1	2025-01-10	Korekta dokumentu. Wypełnienie rozdziału 7, 8 i 9.	Z, W	W	AM, AS, AB	nd
3.0	2025-01-12	Korekta dokumentu.	Z	W	AM, AS, AB	nd

(*) Akcje: W = Wstaw, Z = Zamień, We = Weryfikuj, N = Nowy

(**) Rozdziały: W = Wszystkie

(***) Autorzy: patrz metryka dokumentu

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Spis treści

1.	Opis dokumentu	5
1.1.	Cel dokumentu.....	5
1.2.	Wymagana wiedza	5
1.3.	Dokumenty powiązane.....	5
1.4.	Wykaz skrótów i oznaczeń	6
2.	Opis źródeł danych.....	6
3.	Wstępna lista raportów	10
4.	Model punktowy.....	11
5.	Model logiczny bazy danych dla hurtowni danych	14
6.	Projekt procesu ETL.....	16
7.	Opis kostki wielowymiarowej.....	19
8.	Prezentacja raportów	22
9.	Wnioski końcowe	29

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Spis rysunków

Rysunek 1. Model bazy danych Northwind	6
Rysunek 2. Model punktowy HD – część graficzna	11
Rysunek 3. Model relacyjny HD	14
Rysunek 4. Zasilanie tabeli Klient	16
Rysunek 5. Zasilanie tabeli Pracownik.....	16
Rysunek 6. Zasilanie tabeli Produkt.....	17
Rysunek 7. Zasilanie tabeli Sprzedaż.....	18
Rysunek 8. Model kostki OLAP.....	19
Rysunek 9. Wymiar czasu.....	20
Rysunek 10. Wymiar zamówienia.....	20
Rysunek 11. Wymiar klienta	20
Rysunek 12. Wymiar produktu.....	21
Rysunek 13. Wymiar pracownika.....	21
Rysunek 14. Raport 1. w formie graficznej.....	22
Rysunek 15. Raport 1. w formie tabelarycznej	23
Rysunek 16. Raport 2. w formie graficznej.....	24
Rysunek 17. Raport 2. w formie tabelarycznej	24
Rysunek 18. Raport 3. w formie tabelarycznej	25
Rysunek 19. Raport 4. w formie graficznej i tabelarycznej.....	26
Rysunek 20. Raport 5. w formie graficznej.....	27
Rysunek 21. Raport 6. w formie graficznej.....	28

Spis tabel

Tabela 1. Dokumenty powiązane	5
Tabela 2. Wykaz skrótów i oznaczeń	6
Tabela 3. Model punktowy – część opisowa	11

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

1. Opis dokumentu

1.1. Cel dokumentu

Celem tego dokumentu jest przedstawienie sprawozdania z realizacji projektu z zakresu budowy i wdrożenia HD, realizowanego w ramach przedmiotu Hurtownie danych. Ten dokument stanowi dokumentację wszystkich etapów projektu. Te etapy to:

- Opis źródeł danych,
- Wstępna lista raportów,
- Model punktowy,
- Model logiczny bazy danych dla hurtowni danych ,
- Projekt procesu ETL,
- Opis kostki wielowymiarowej,
- Prezentacja raportów,
- Wnioski końcowe.

1.2. Wymagana wiedza

Kompetencje merytoryczne wymagane od osób, do których adresowany jest ten dokument to:

- Podstawowa znajomość relacyjnych baz danych,
- Zrozumienie zależności między tabelami i kluczami (głównymi i obcymi),
- Podstawy analizy danych w kontekście sprzedaży i zarządzania zamówieniami,
- Umiejętność interpretacji prostych modeli punktowych i raportów.

1.3. Dokumenty powiązane

Wykaz dokumentów powiązanych został zawarty w Tabela 1.

Tabela 1. Dokumenty powiązane

Nr	Nazwa	Wykaz wykorzystanych części dokumentu	Wersja	Data wydania
1	Opis bazy Northwind.docx	Diagram bazy Northwind, Opisy tabel	1.0	2024-10-17

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

1.4. Wykaz skrótów i oznaczeń

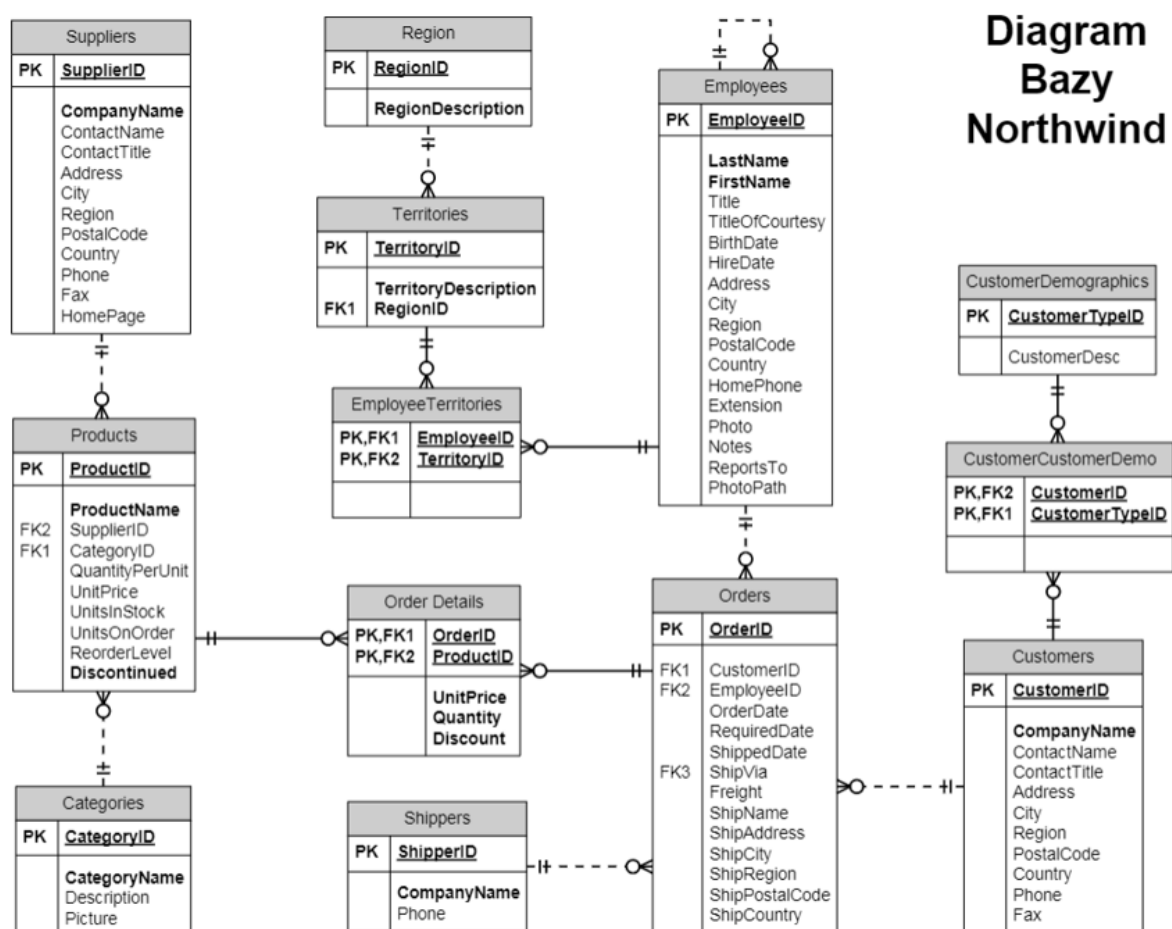
Wykaz skrótów i oznaczeń został zawarty w Tabeli 2.

Tabela 2. Wykaz skrótów i oznaczeń

Skrót lub oznaczenie	Rozwinięcie
HD	Hurtownie danych

2. Opis źródeł danych

Model bazy danych Northwind przedstawiono na Rysunek 1.



Rysunek 1. Model bazy danych Northwind

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Tabela Categories - zawiera informacje o kategoriach produktów.

Opis kolumn:

CategoryID int(4) - identyfikator kategorii
 CategoryName nvarchar(15) - nazwa kategorii
 Description ntext(16) - opis
 Picture image(16) - zdjęcie

Tabela Products – zawiera informacje o produktach i towarach.

Opis kolumn:

ProductID int(4) - identyfikator produktu
 ProductName nvarchar(40) - nazwa produktu
 SupplierID int(4) - identyfikator dostawcy
 CategoryID int(4) - identyfikator kategorii produktu
 QuantityPerUnit nvarchar(20) - informacja o ilości w jednostce (np 20 sztuk w opakowaniu)
 UnitPrice money(8) cena jednostkowa
 UnitsInStock smallint(2) - stan magazynu, ilość jednostek towaru w magazynie
 UnitsOnOrder smallint(2) - ilość zamówiona, ilość zamówionych jednostek towaru
 ReorderLevel smallint(2) - minimalna ilość w magazynie
 Discontinued bit(1) - wycofany (tak/nie)

Tabela Suppliers – zawiera informacje o dostawcach.

Opis kolumn:

SupplierID int(4) - identyfikator dostawcy
 CompanyName nvarchar(40) - nazwa dostawcy (firmy będącej dostawcą)
 ContactName nvarchar - nazwisko i imię reprezentanta
 ContactTitle nvarchar(30) - stanowisko/tytuł reprezentanta
 Address nvarchar(60) - adres dostawcy (firmy będącej dostawcą)
 City nvarchar(15) - miasto
 Region nvarchar(15) - region
 PostalCode nvarchar(10) - kod pocztowy
 Country nvarchar(15) - kraj
 Phone nvarchar(24) - telefon
 Fax nvarchar(24) - fax
 HomePage ntext(16) - strona www firmy

Tabela Shippers - zawiera informacje o spedytoraх, tzn. firmach zajmujących się dostawą towarów.

Opis kolumn:

ShipperID int(4) - identyfikator spedytora
 CompanyName nvarchar(40) - nazwa firmy
 Phone nvarchar(24) - telefon

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Tabela Customers – zawiera informacje o klientach.

Opis kolumn:

CustomerID nchar(5) - identyfikator klienta
 CompanyName nvarchar(40) - nazwa klienta (firmy będącej klientem)
 ContactName nvarchar(30) - nazwisko i imię reprezentanta
 ContactTitle nvarchar(30) - stanowisko/tytuł reprezentanta
 Address nvarchar(60) - nazwa klienta (firmy będącej klientem)
 City nvarchar(15) - miasto
 Region nvarchar(15) - region
 PostalCode nvarchar(10) - kod pocztowy
 Country nvarchar(10) - kraj
 Phone nvarchar(24) - nr telefonu
 Fax nvarchar(24) - nr faxu

Tabela Orders – zawiera informacje o zamówieniach.

Opis kolumn:

OrderID int(4) - identyfikator zamówienia
 CustomerID nchar(5) - identyfikator klienta
 EmployeeID int(4) - identyfikator pracownika (obsługującego, wystawiającego zamówienie)
 OrderDate datetime(8) - data zamówienia
 RequiredDate datetime(8) - wymagana data realizacji
 ShippedDate datetime(8) - data wysyłki
 ShipVia int(4) - identyfikator spedytora
 Freight money(8) - opłata za przesyłkę (fracht)
 ShipName nvarchar(40) - nazwa/nazwisko odbiorcy
 ShipAddres nvarchar(60) - adres odbiorcy
 ShipCity nvarchar(15) - miasto odbiorcy
 ShipRegion nvarchar(15) - region odbiorcy
 ShipPostalCode nvarchar(10) - kod pocztowy
 ShipCountry nvarchar(15) - kraj odbiorcy

Tabela OrderDetails – zawiera informacje o szczegółach zamówienia (pozycjach zamówienia).

Opis kolumn:

OrderID int(4) - identyfikator zamówienia
 ProductID int(4) - identyfikator zamawianego produktu
 UnitPrice money(8) - cena jednostkowa
 Quantity smallint(2) - ilość
 Discount real(4) - rabat

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Tabela Employees – zawiera informacje o pracownikach.

Opis kolumn:

EmployeeID int(4) - identyfikator pracownika
 LastName nvarchar(20) - nazwisko
 FirstName nvarchar(10) - imię
 Title nvarchar(30) - tytuł/stanowisko
 TitleOfCourtesy nvarchar(25) - zwrot grzecznościowy (np. używany w korespondencji)
 BirthDate datetime(8) - data urodzenia
 HireDate datetime(8) - data zatrudnienia
 Notes ntext(16) - uwagi dotyczące pracownika
 ReportTo int(4) - identyfikator przełożonego
 PhotoPath nvarchar(255) - zdjęcie (adres strony www)
 Address nvarchar(60) - adres pracownika
 City nvarchar(15) - miasto
 Region nvarchar(15) - region
 PostalCode nvarchar(10) - kod pocztowy
 Country nvarchar(15) - kraj
 HomePhone nvarchar(24) - telefon domowy
 Extension nvarchar(4) - telefon wewnętrzny
 Photo image(16) - zdjęcie

Tabela Region – zawiera informacje o regionach.

Opis kolumn:

RegionID int(4) - identyfikator regionu
 RegionDescription nchar(50) - opis regionu

Tabela Territories - zawiera informacje o terytoriach/obszarach/miastach.

Opis kolumn:

TerritoryID nvarchar(20) - identyfikator obszaru
 TerritoryDescription nchar(50) - opis, nazwa
 RegionID int(4) - identyfikator regionu

Tabela EmployeeTerritories - zawiera informacje o terytoriach/obszarach/miastach (obsługiwanych przez poszczególnych pracowników).

Opis kolumn:

EmployeeID int(4) - identyfikator pracownika
 TerritoryID nvarchar(20) - identyfikator obszaru

Tabela CustomerDemographics - zawiera informacje o grupach klientów.

Opis kolumn:

CustomerTypeID int(4) identyfikator grupy
 CustomerDesc nvarchar(15) - opis, charakterystyka grupy

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

3. Wstępna lista raportów

Wstępna lista raportów została opisana w następujący sposób

< problem> - <zapytanie(-a) decyzyjne> – <raport(-y)>

Lista raportów została przedstawiona poniżej

- < Zapas określonych produktów osiągnął minimalny próg > - < O ile zwiększyć zapasy (tzn. zakupić) danego produktu? >-< Liczba sprzedanych produktów w poszczególnych miesiącach na przestrzeni ostatnich 2 lat >
- < Określenie wielkości zniżek dla poszczególnych klientów >-< Jakie wielkości zniżek przydzielić klientom? >-< Łączna wartość sprzedaży dla poszczególnych klientów w ostatnim roku. >
- < Przyznanie nagród rocznych pracownikom >-< Których pracowników należy wyróżnić za realizację największych zamówień w ostatnim roku? >-< Zestawienie 10 największych zamówień pod względem wartości sprzedaży zrealizowanych w ostatnim roku. >
- < Przyznanie nagród dla pracowników realizujących największą liczbę zamówień w miesiącu>-<Których pracowników należy nagrodzić za największą wydajność?>-< Liczba obsługiwanych zamówień przez pracowników w ostatnim miesiącu.>
- < Zwiększony popyt na kategorię produktów w danym okresie >-< W jakich okresach zwiększyć stan magazynowy danego produktu? >-< Ilość sprzedawanych produktów danej kategorii w poszczególnych miesiącach na przestrzeni ostatniego roku. >
- < Współpraca z firmami z poszczególnych krajów przynosi niezadowalający zysk. > - < Z klientami, z których krajów należy rozwijać współpracę? W jakich krajach reklama może przyczynić się do wzrostu zysków? > - < Wartość sprzedaży produktów klientom z różnych krajów>

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

4. Model punktowy

Model punktowy HD przedstawiono na Rysunek 2.



Rysunek 2. Model punktowy HD – część graficzna

Model punktowy HD szczegółowo opisano w Tabeli 3.

Tabela 3. Model punktowy – część opisowa

Nazwa obiektu	Typ	Zmiennosc w czasie	Uwagi	Typ danych	Wymagalnośc	Źródło
Sprzedaż	Encja	Stała	Sprzedaż dotyczy jednego produktu, nie ma możliwości edycji	n/d	n/d	n/d
Zamówienie	Encja	Stała	Nie ma możliwości anulowania zamówienia.	n/d	n/d	n/d
Produkt	Encja	Tak	Istnieje możliwość wycofania produktu ze sprzedaży.	n/d	n/d	n/d
Kategoria	Encja	Stała	Nie ma możliwości usunięcia kategorii	n/d	n/d	n/d
Pracownik	Encja	Tak	Istnieje możliwość zakończenia pracy.	n/d	n/d	n/d
Klient	Encja	Nie	Klient, który raz korzystał z usług nie zostaje zapomniany.	n/d	n/d	n/d
Dzień	Encja	Stała	Każdy miesiąc ma stałą ilość dni	n/d	n/d	n/d
Miesiąc	Encja	Stała	Każdy rok ma 12 miesięcy	n/d	n/d	n/d
Rok	Encja	Stała	Stała wartość dla każdego roku od roku najstarszej danej w hurtowni	n/d	n/d	n/d
Miesiąc→Rok	Relacja	Stała	Relacja między miesiącem a rokiem jest stała	n/d	Tak	n/d

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Dzień→Miesiąc	Relacja	Stała	Relacja między dniem a miesiącem jest stała	n/d	Tak	n/d
Dzień→Sprzedaż	Relacja	Stała	Relacja między dniem a sprzedażą jest stała	n/d	Tak	n/d
Klient→Sprzedaż	Relacja	Stała	Nie można zmienić przypisanego klienta podczas sprzedaży	n/d	Tak	n/d
Produkt→Sprzedaż	Relacja	Stała	Nie można zmieniać zamówionych produktów podczas sprzedaży	n/d	Tak	n/d
Produkt→Kategoria	Relacja	Stała	Nie można zmieniać kategorii produktom	n/d	Tak	n/d
Pracownik → Sprzedaż	Relacja	Stała	Nie można unieważnić sprzedaży realizowanej przez danego pracownika.	n/d	Tak	n/d
Zamówienie → Sprzedaż	Relacja	Stała	Sprzedaż jest przypisana do jednego zamówienia.	n/d	Tak	n/d
Dzien.ID	Atrybut	Stała	Identyfikator dnia, unikalny	int(4)	Tak	n/d
Miesiąc.ID	Atrybut	Stała	Identyfikator miesiąca, unikalny	int(4)	Tak	n/d
Rok.ID	Atrybut	Stała	Identyfikator roku, unikalny	int(4)	Tak	n/d
Kategoria.ID	Atrybut	Stała	Identyfikator kategorii, unikalny w bazie danych	int(4)	Tak	n/d
Kategoria.Nazwa	Atrybut	Stała	Nazwa kategorii, niezmienna w czasie	nvarchar(40)	Tak	n/d
Produkt.ID	Atrybut	Stała	Identyfikator produktu, unikalny w bazie danych.	int(4)	Tak	Northwind.Products.ProductID
Produkt.Nazwa	Atrybut	Stała	Nazwa produktu, niezmienna w czasie.	nvarchar(40)	Tak	Northwind.Products.Product Name
Zamówienie.ID	Atrybut	Stała	Identyfikator zamówienia, unikalny w bazie danych.	int(4)	Tak	Northwind.Orders.OrderID

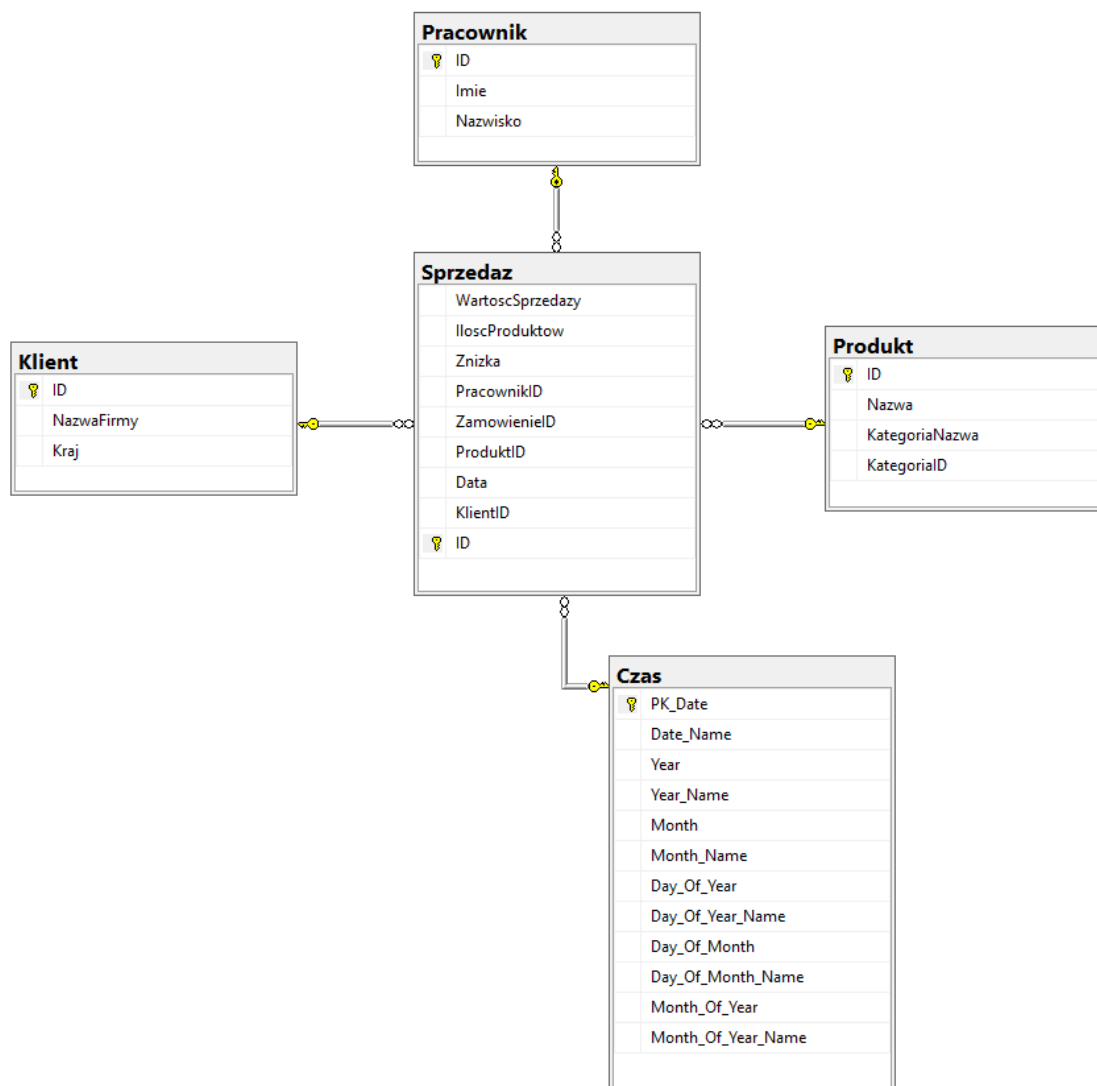
Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Sprzedaz.ID	Atrybut	Stała	Identyfikator sprzedaży, przypisany przy zasilaniu bazy danych.	int(4)	Tak	n/d
Sprzedaz.Wartosc Sprzedazy	Atrybut	Stała	Wartość pozycji zamówienia przy uwzględnieniu rabatu	numeric(10,2)	Tak	Northwind.OrderDetails.UnitPrice * Northwind.OrderDetails.Quantity * (1 - Northwind.OrderDetails.Discount)
Sprzedaz.IloscProduktow	Atrybut	Stała	Liczba produktów zamawianych w ramach pozycji zamówienia. Wartość stała dla zamówienia.	int(4)	Tak	Northwind.OrderDetails.Quantity
Sprzedaz.Zniżka	Atrybut	Stała	Wartość zniżki przydzielonej do pozycji zamówienia.	real	Nie	(Northwind.OrderDetails.UnitPrice * Northwind.OrderDetails.Quantity * (1 - Northwind.OrderDetails.Discount)) * Northwind.OrderDetails.Discount
Klient.ID	Atrybut	Stała	Identyfikator klienta, unikalny w bazie danych.	nchar(5)	Tak	Northwind.Customers.CustomerID
Klient.NazwaFirmy	Atrybut	Tak	Nazwa firmy klienta może być zmieniona	nvarchar(40)	Tak	Northwind.Customers.CompanyName
Klient.Kraj	Atrybut	Tak	Klient może zmienić kraj zamieszkania	nvarchar(15)	Tak	Northwind.Customers.Country
Pracownik.ID	Atrybut	Stała	Identyfikator pracownika, przypisany na stałe w systemie.	int(4)	Tak	Northwind.Employees.EmployeeID
Pracownik.Imie	Atrybut	Tak	Pracownik może zmienić imię	nvarchar(30)	Tak	Northwind.Employees.FirstName
Pracownik.Nazwisko	Atrybut	Tak	Pracownik może zmienić nazwisko	nvarchar(30)	Tak	Northwind.Employees.LastName

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

5. Model logiczny bazy danych dla hurtowni danych

Na podstawie opisu modelu punktowego z rozdziału 4, stworzono model relacyjny HD. Model przedstawiono na Rysunek 3.



Rysunek 3. Model relacyjny HD

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Opis tabeli modelu relacyjnego HD umieszczono poniżej.

Tabela Pracownik - zawiera informacje o pracownikach.

Opis kolumn:

ID int(4) - identyfikator pracownika
Imie nvarchar(30) – imię pracownika
Nazwisko nvarchar(30) – nazwisko pracownika

Tabela Produkt – zawiera informacje o produktach.

Opis kolumn:

ID int(4) - identyfikator kategorii
Nazwa nvarchar(40) – nazwa produktu
KategoriaNazwa nvarchar(40) – nazwa kategorii produktu
KategoriaID int(4) – identyfikator kategorii

Tabela Czas – zawiera informacje o czasie z zakresu od 04.07.1996 do 05.06.1998.

Opis kolumn:

PK_Date datetime – data z dokładnością do dnia
Date_name nvarchar(50) – pełna nazwa daty
Year datetime – data z dokładnością do roku
Year_Name nvarchar(50) – nazwa daty z dokładnością do roku
Month datetime – data z dokładnością do miesiąca
Month_Name nvarchar(50) – nazwa daty z dokładnością do miesiąca
Day_Of_Year int – numer dnia roku
Day_Of_Year_Name nvarchar(50) – nazwa dnia roku
Day_Of_Month int – numer dnia miesiąca
Day_Of_Month_Name nvarchar(50) – nazwa dnia miesiąca
Month_Of_Year int – numer miesiąca roku
Month_Of_Year_Name nvarchar(50) – nazwa miesiąca roku

Tabela Klient – zawiera informacje o klientach.

Opis kolumn:

ID nchar(5) - identyfikator klienta
NazwaFirmy nvarchar(40) – nazwa firmy
Kraj nvarchar(15) – nazwa kraju pochodzenia klienta

Tabela Sprzedaz – tabela faktów, zawiera informacje o sprzedaży.

Opis kolumn:

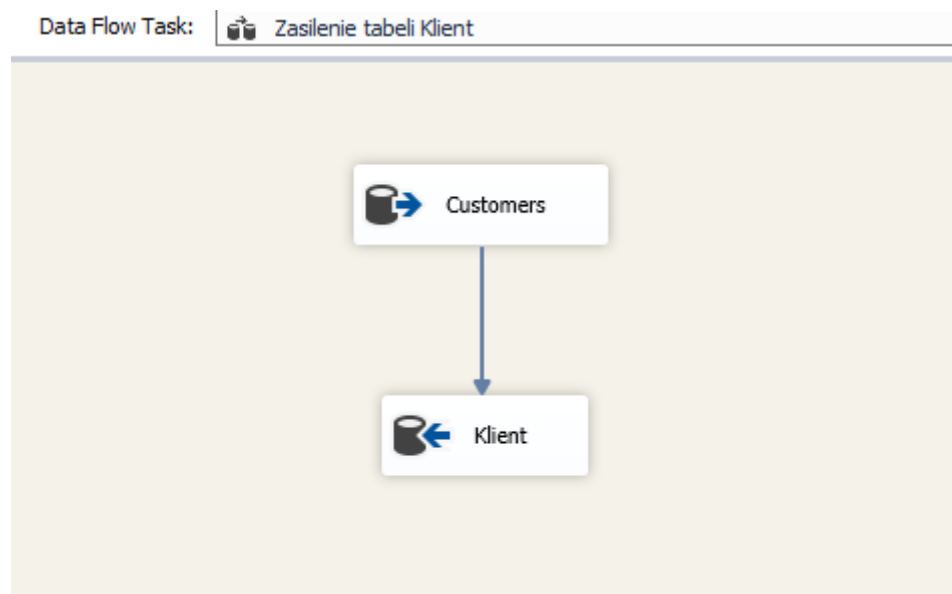
ID int(4) - identyfikator sprzedaży
WartoscSprzedazy numeric(10,2) – wartość sprzedaży, z uwzględnieniem zniżki
IloscProduktow int(4) – liczba sprzedanych produktów
Znizka real – wielkość określająca zniżkę
PracownikID int(4) – id pracownika obsługującego sprzedaż
ZamowienieID int(4) – id zamówienia, którego elementem jest opisywana sprzedaż
ProduktID int(4) – id sprzedawanego produktu
Data datetime – data sprzedaży
KlientID int(4) – id klienta

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

6. Projekt procesu ETL

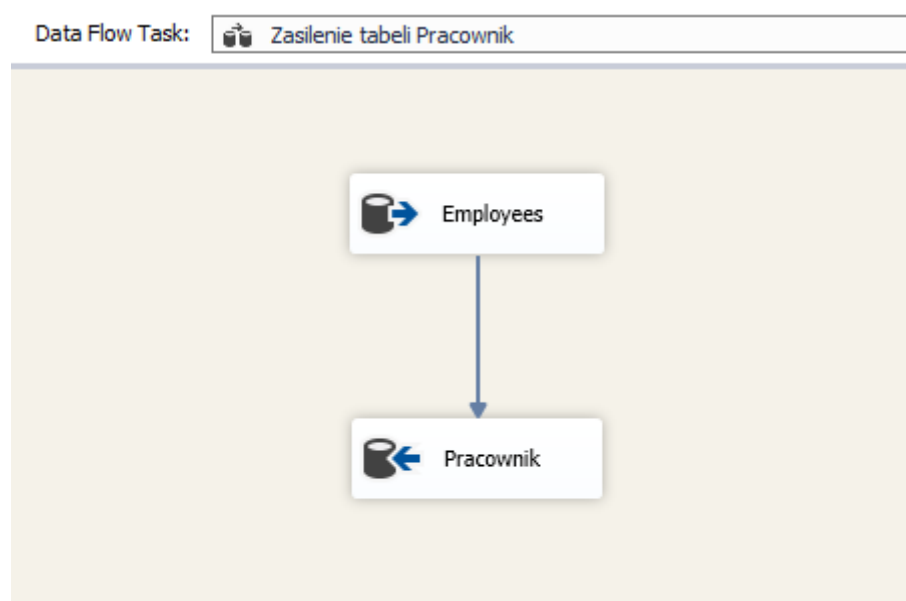
Zaimplementowano proces ETL zasilania bazy danych wykorzystując narzędzie SSIS. Poniżej przedstawiono procesy zasilania kolejnych tabel HD.

- Zasilanie tabeli Klient – tabelę Klient w HD zasilono pobierając dane z tabeli źródłowej Customers bazy danych Northwind.



Rysunek 4. Zasilanie tabeli Klient

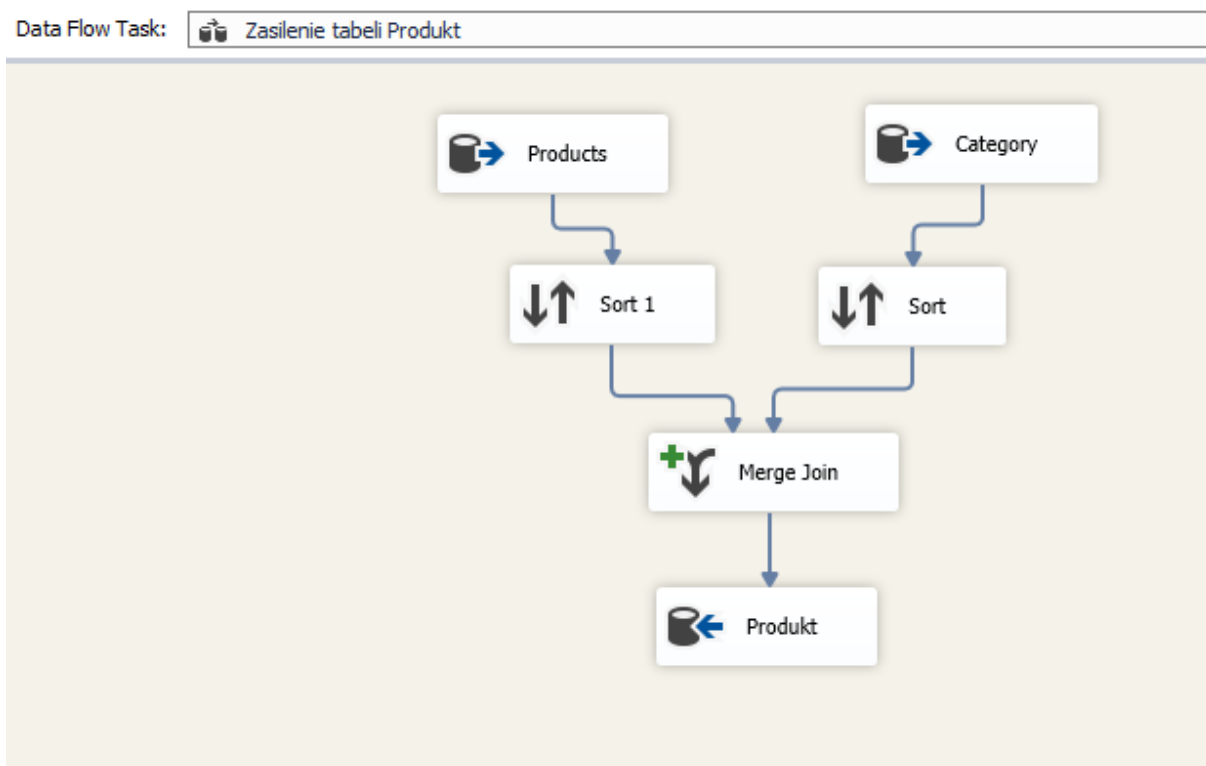
- Zasilanie tabeli Pracownik – tabelę Pracownik zasilono pobierając dane, z tabeli źródłowej Employees.



Rysunek 5. Zasilanie tabeli Pracownik

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

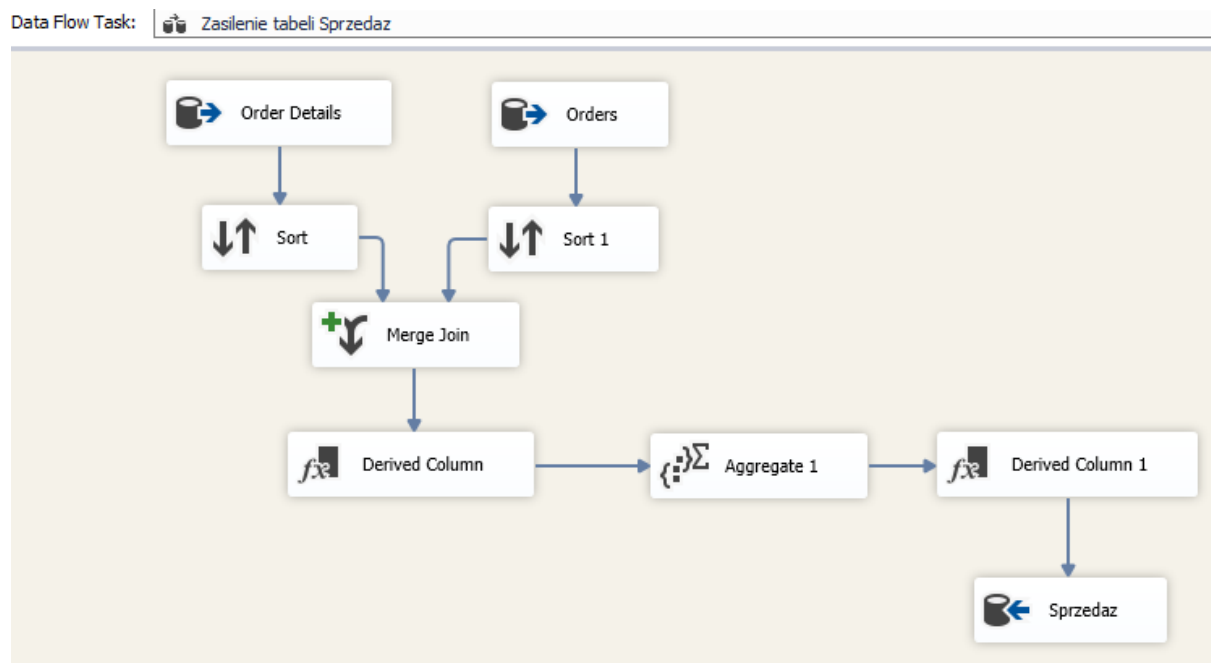
- Zasilanie tabeli Produkt – tabelę Produkt zasilono pobierając dane z dwóch tabel źródłowych opisujących produkty, tzn. z tabeli Products i tabeli Category. Dane przed zasileniem tabeli Produkt zostały posortowane i połączone.



Rysunek 6. Zasilanie tabeli Produkt

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

- Zasilanie tabeli Sprzedaz – tabelę sprzedaż zasilono na podstawie danych z tabel źródłowych Order Details i Orders. Dane poddano sortowaniu i łączeniu. Następnie na podstawie tych danych wyznaczono wartości sprzedaży, i dokonano agregacji danych. Tak zmodyfikowanymi danymi zasilono tabelę Sprzedaż.



Rysunek 7. Zasilanie tabeli Sprzedaż

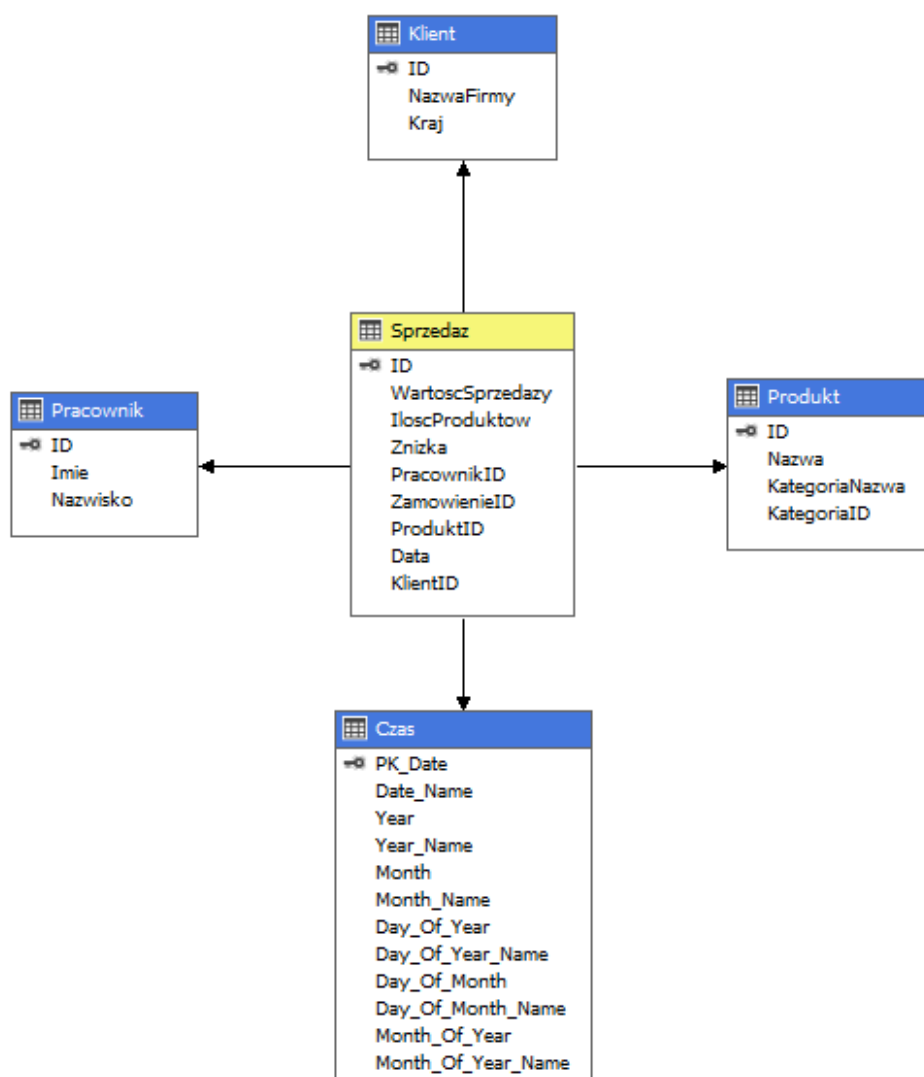
Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

7. Opis kostki wielowymiarowej

Kostka wielowymiarowa została zbudowana na podstawie danych z opracowanej hurtowni danych, wykorzystując tabelę faktów oraz powiązane z nią tabele wymiarów. Poniżej przedstawiono opis uwzględnionych miar, oraz wymiarów.

- **Model kostki OLAP**

Na Rysunek 8 przedstawiono model kostki OLAP.



Rysunek 8. Model kostki OLAP

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

- **Opis wymiarów kostki**

Poniżej przedstawiono opisy wymiarów zbudowanej kostki.

- **Wymiar czasu**

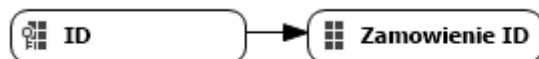
Hierarchia wymiaru czasu została przedstawiona na Rysunek 9. Atrybuty tego wymiaru to dzień, miesiąc, rok, oraz dzień miesiąca, dzień roku i numer miesiąca.



Rysunek 9. Wymiar czasu

- **Wymiar zamówienia**

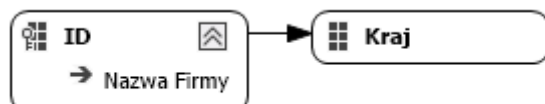
Hierarchia wymiaru zamówienia została przedstawiona na Rysunek 10. Atrybuty tego wymiaru to ID sprzedaży, i ID zamówienia. Jest to wymiar zdegenerowany, jego atrybuty, zamiast być przechowywane w osobnej tabeli wymiarów, zostały bezpośrednio pobrane z tabeli faktów.



Rysunek 10. Wymiar zamówienia

- **Wymiar klienta**

Hierarchia wymiaru klienta została przedstawiona na Rysunek 11. Atrybuty tego wymiaru to ID klienta, nazwa firmy i kraj klienta.

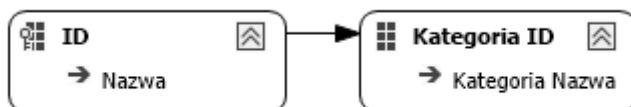


Rysunek 11. Wymiar klienta

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

- **Wymiar produktu**

Hierarchia wymiaru produktu została przedstawiona na Rysunek 12. Atrybuty tego wymiaru to ID produktu, nazwa produktu, ID kategorii, nazwa kategorii.



Rysunek 12. Wymiar produktu

- **Wymiar pracownika**

Hierarchia wymiaru pracownika została przedstawiona na Rysunek 13. Atrybuty tego wymiaru to ID pracownika, imię i nazwisko.



Rysunek 13. Wymiar pracownika

- **Opis miar kostki**

Poniżej przedstawiono opisy miar zbudowanej kostki.

- **Miara wartości sprzedaży**

Miara ta wyznacza sumę wartości sprzedaży, względem analizowanych atrybutów wymiarów.

- **Miara ilości produktów**

Wyznacza sumę ilości produktów, względem analizowanych atrybutów wymiarów.

- **Miara zniżki**

Wyznacza sumę wartości zniżek, względem analizowanych atrybutów wymiarów.

- **Miara liczby zamówień**

Zlicza wszystkie unikalne id zamówienia, względem analizowanych atrybutów wymiarów.

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

8. Prezentacja raportów

W oparciu o utworzoną kostkę wielowymiarową przygotowano zestaw raportów odpowiadający wstępnej liście raportów z Rozdziału 3. Raporty zostały opracowane w programie Power BI. Poniżej przedstawiono listę raportów, wraz z analizą ich wpływu na podejmowane decyzje biznesowe.

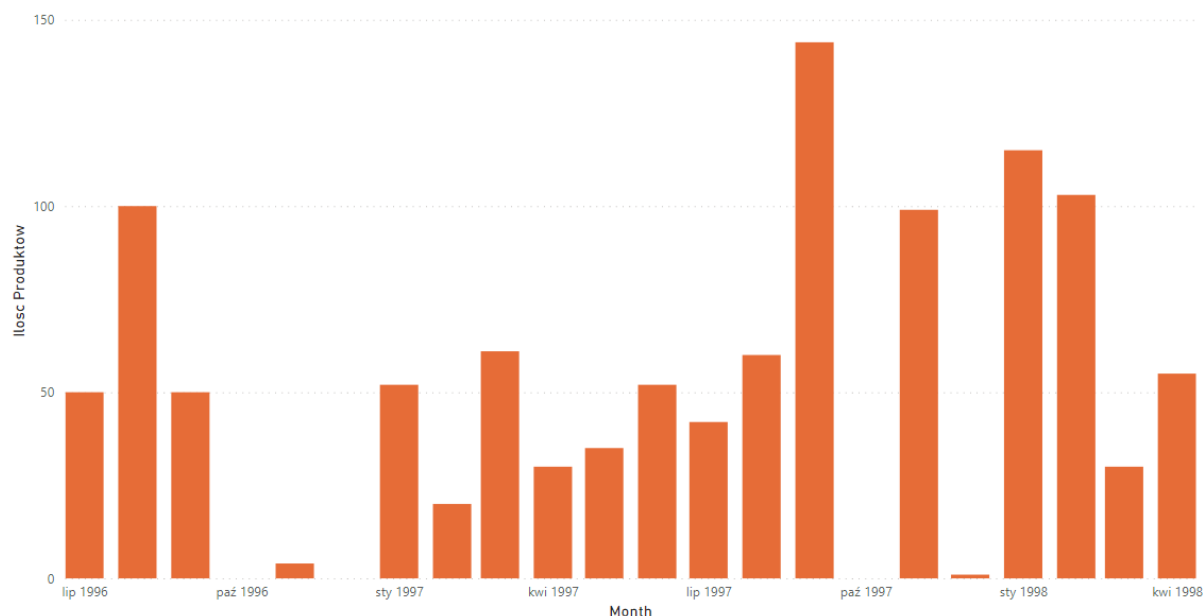
- **Problem 1:** Zapas określonych produktów osiągnął minimalny próg

Zapytanie decyzyjne: O ile zwiększyć zapasy (tzn. ile zakupić) danego produktu?

Raport 1. Założono, że produktem, którego zapas osiągnął minimalny próg jest “Boston Crab Meat”. Stworzono raport liczby sprzedanych produktów “Boston Crab Meat” w poszczególnych miesiącach na przestrzeni ostatnich 2 lat. Raport w formie graficznej i tabelarycznej przedstawiono na Rysunek 14. i Rysunek 15.

Miesięczna liczba sprzedanych produktów na przestrzeni ostatnich 2 lat

Nazwa ● Boston Crab Meat



Rysunek 14. Raport 1. w formie graficznej

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Month	Boston Crab Meat
1996-07-01	50
1996-08-01	100
1996-09-01	50
1996-11-01	4
1997-01-01	52
1997-02-01	20
1997-03-01	61
1997-04-01	30
1997-05-01	35
1997-06-01	52
1997-07-01	42
1997-08-01	60
1997-09-01	144
1997-11-01	99
1997-12-01	1
1998-01-01	115
1998-02-01	103
1998-03-01	30
1998-04-01	55

Rysunek 15. Raport 1. w formie tabelarycznej

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport przedstawia miesięczną sprzedaż danego produktu na przestrzeni ostatnich dwóch lat, co pozwala na identyfikację sezonowych trendów oraz miesięcy o najwyższym i najniższym popycie. Dane te pomagają w prognozowaniu zapotrzebowania oraz planowaniu odpowiedniego poziomu zapasów, zapewniając ciągłość dostępności produktu.

Decyzja: Na podstawie raportu należy zwiększyć zapasy produktu do poziomu minimum 50 sztuk w miesiącach w których w obydwu latach powtórzył się taki popyt (np. wrzesień i sierpień). W miesiącach, w których powtórzyła się mała sprzedaż (np. marzec) zapasy można utrzymywać na niższym poziomie, aby uniknąć nadmiernego magazynowania. Dodanie zapasu bezpieczeństwa w okresach wzmożonej sprzedaży (np. początek roku) zapewni ciągłość dostępności produktu.

- **Problem 2:** Określenie wielkości zniżek dla poszczególnych klientów

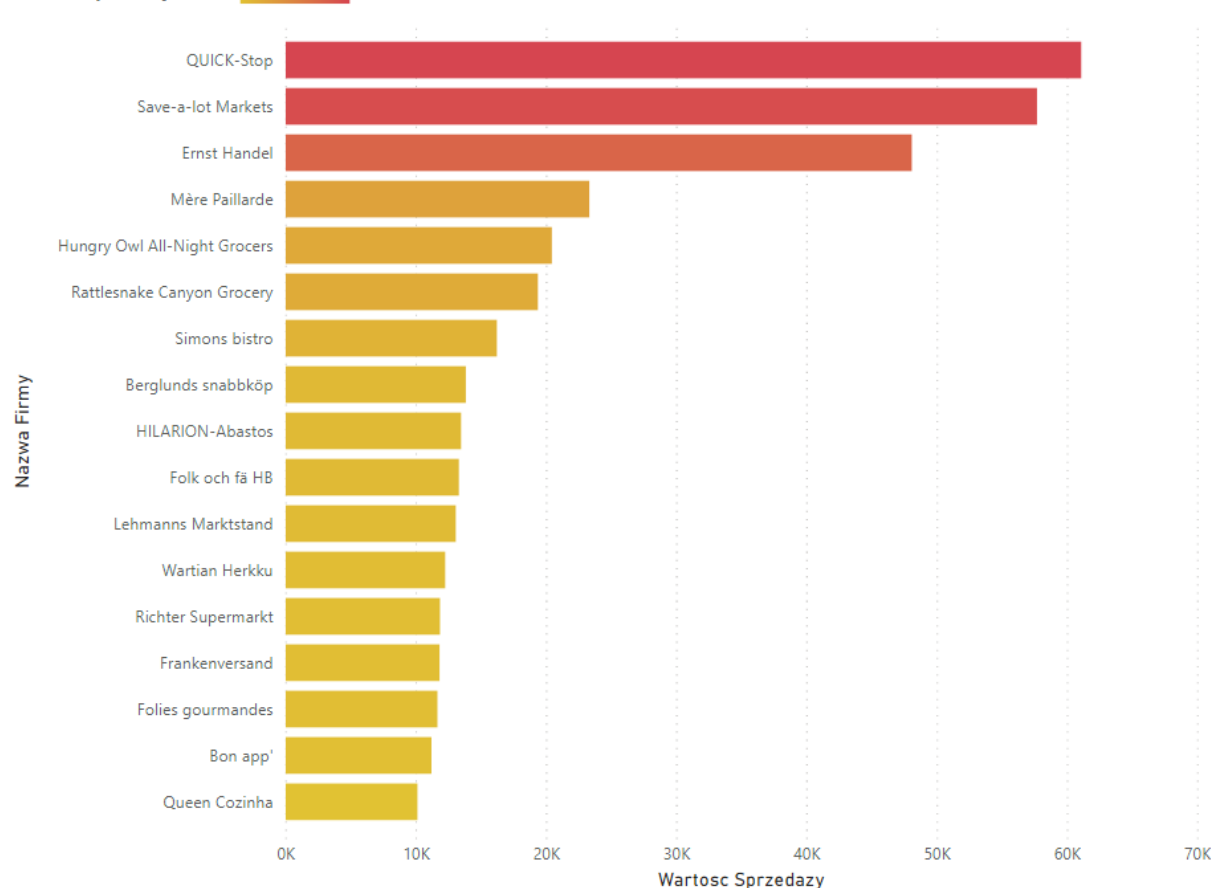
Zapytanie decyzyjne: Jakie wielkości zniżek przydzielić klientom?

Raport 2. Założono, że minimalna wartość sprzedaży w ostatnim roku, kwalifikująca daną firmę do zniżek to 10000. Stworzono raport łącznej wartości sprzedaży dla poszczególnych klientów w ostatnim roku. Raport w formie graficznej i tabelarycznej przedstawiono na Rysunek 16. i Rysunek 17.

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Łączna wartości sprzedaży dla poszczególnych klientów w 1997 roku

Wartosc Sprzedazy 10,13K 61,11K



Rysunek 16. Raport 2. w formie graficznej

Nazwa Firmy	Wartosc Sprzedazy
QUICK-Stop	61 109,92
Save-a-lot Markets	57 713,57
Ernst Handel	48 096,26
Mère Paillarde	23 332,32
Hungry Owl All-Night Grocers	20 454,40
Rattlesnake Canyon Grocery	19 383,75
Simons bistro	16 232,42
Berglunds snabbköp	13 849,01
HILARION-Abastos	13 482,74
Folk och fä HB	13 314,67
Lehmanns Marktstand	13 076,12
Wartian Herkku	12 262,94
Richter Supermarkt	11 864,42
Frankenversand	11 829,78
Folies gourmandes	11 666,90
Bon app'	11 208,35
Queen Cozinha	10 132,77

Rysunek 17. Raport 2. w formie tabelarycznej

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport przedstawia łączną wartość sprzedaży dla poszczególnych klientów, którzy w 1997 roku przekroczyli minimalny próg wartości sprzedaży, kwalifikujący klientów do stałej zniżki (10 000 złotych). Pozwala to, na określenie klientów kwalifikujących się do zniżek oraz wyznaczenie ich wielkości. Dane te wspierają proces decyzyjny, wskazując, którzy klienci wnieśli największy wkład do sprzedaży oraz umożliwiając zaplanowanie zniżek motywujących do dalszej współpracy.

Decyzja: Zniżki powinny być przyznane wszystkim klientom pojawiającym się w raporcie. Ich wielkość może być uzależniona od wartości sprzedaży – dla największych klientów, takich jak "QUICK-Stop" (61 109,92) czy "Save-a-lot Markets" (57 713,57), zniżki mogą być wyższe, aby wzmocnić relacje biznesowe. Klientom o niższych wartościach sprzedaży (np. "Queen Cozinha" – 10 132,77), można przyznać zniżki w mniejszym wymiarze, aby zachęcić do zwiększenia zakupów.

- **Problem 3:** Przyznanie nagród rocznych pracownikom.

Zapytanie decyzyjne: Których pracowników należy wyróżnić za realizację największych zamówień w ostatnim roku?

Raport 3. Stworzono raport obrazujący 10 największych zamówień pod względem wartości sprzedaży. Raport w formie tabelarycznej przedstawiono na Rysunek 18.

Year	Zamowienie ID	Imie	Nazwisko	Wartosc Sprzedazy
1997-01-01	10417	Margaret	Peacock	11 188,40
1997-01-01	10479	Janet	Leverling	10 495,60
1997-01-01	10540	Janet	Leverling	10 191,70
1997-01-01	10691	Andrew	Fuller	10 164,80
1997-01-01	10515	Andrew	Fuller	9 921,30
1997-01-01	10424	Robert	King	9 194,56
1997-01-01	10514	Janet	Leverling	8 623,45
1997-01-01	10776	Nancy	Davolio	6 635,27
1997-01-01	10607	Steven	Buchanan	6 475,40
1997-01-01	10612	Nancy	Davolio	6 375,00
Total				89 265,48

Rysunek 18. Raport 3. w formie tabelarycznej

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport przedstawia 10 największych zamówień zrealizowanych w 1997 roku, uwzględniając ich wartość oraz pracowników odpowiedzialnych za ich obsługę. Analiza tych danych pozwala wskazać osoby, które przyczyniły się do osiągnięcia największych wyników sprzedażowych w danym roku. Raport wspiera proces decyzyjny w zakresie wyróżnienia najbardziej efektywnych pracowników.

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

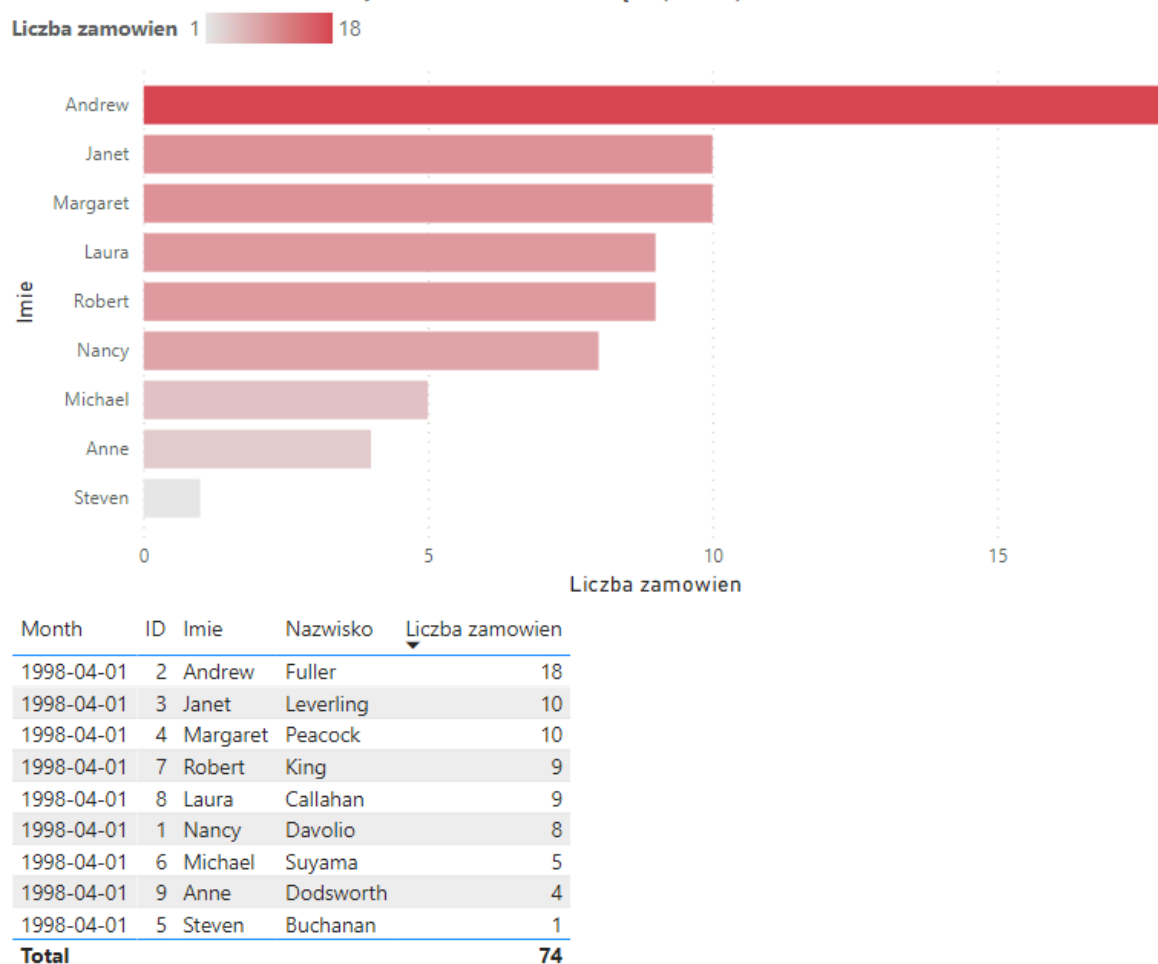
Decyzja: Na podstawie raportu należy wyróżnić pracowników, którzy zrealizowali największe zamówienia pod względem wartości. W szczególności należy nagrodzić Margaret Peacock za realizację zamówienia o wartości 11 188,40 i Janet Leverling, która odpowiada za trzy zamówienia z pierwszej dziesiątki o łącznej wartości prawie 30 000. Przyznanie nagród tym pracownikom doceni ich zaangażowanie i zmotywuje zespół do osiągania wysokich wyników w kolejnych latach.

- **Problem 4:** Przyznanie nagrody pracownika miesiąca.

Zapytanie decyzyjne: Których pracowników należy nagrodzić za największą wydajność?

Raport 4. Stworzono raport ilości obsługiwanych zamówień przez pracowników w ostatnim miesiącu. Raport w formie graficznej i tabelarycznej przedstawiono na Rysunek 19.

Liczba zamówień zrealizowanych w ostatnim miesiącu przez pracowników



Rysunek 19. Raport 4. w formie graficznej i tabelarycznej

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport przedstawia ilość zamówień obsługiwanych przez poszczególnych pracowników w ostatnim miesiącu, co pozwala ocenić efektywność każdego z nich. Daje także wgląd w równomierność rozkładu pracy

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

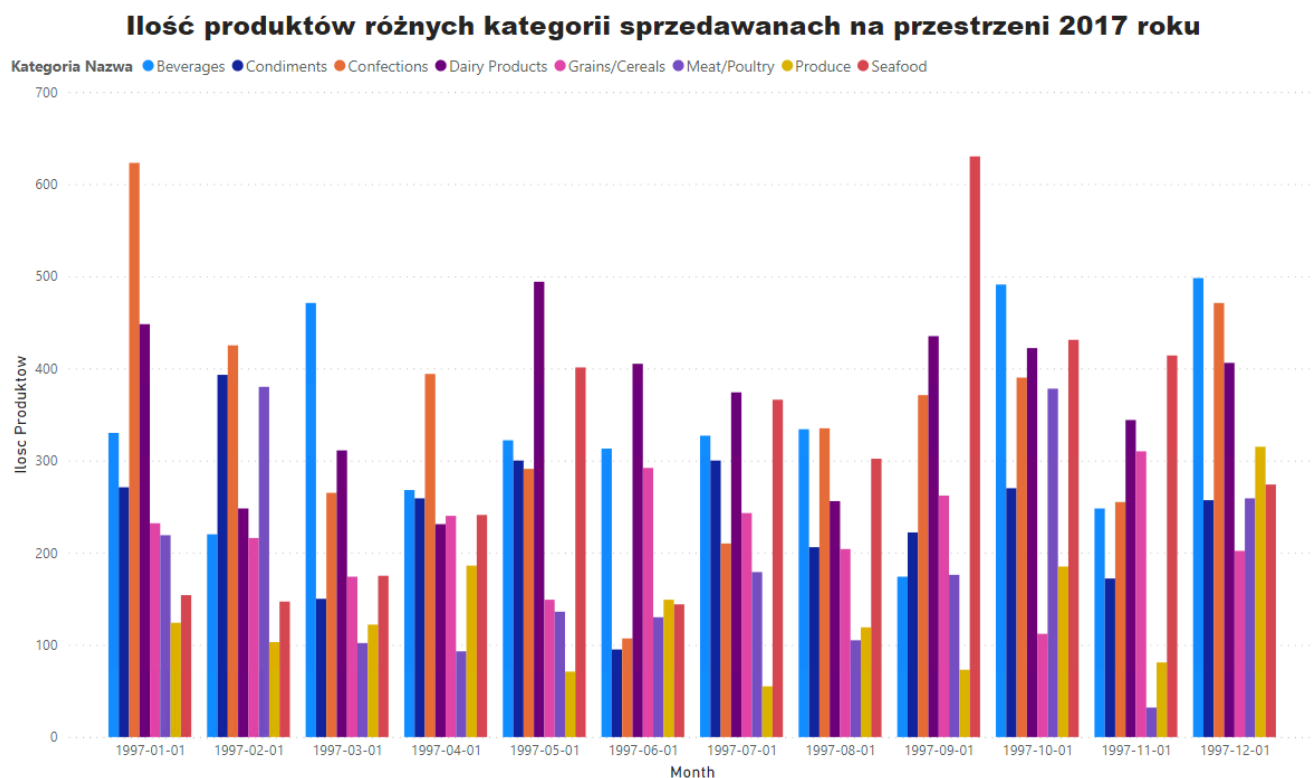
w zespole. Dane te pomagają w zidentyfikowaniu najbardziej wydajnych pracowników oraz potencjalnych kandydatów do wyróżnienia.

Decyzja: Nagroda „Pracownik miesiąca” zostaje przyznana pracownikowi, który obsłużył największą liczbę zamówień w danym miesiącu - w tym wypadku jest to Andrew Fuller, który zrealizował 18 zamówień. Decyzja opiera się na liczbowych wynikach efektywności, które są obiektywne i motywujące dla pracowników. Przyznanie nagrody promuje produktywność, docenia zaangażowanie i może zwiększać morale zespołu.

- **Problem 5:** Zwiększony popyt na kategorię produktów w danym okresie.

Zapytanie decyzyjne: W jakich okresach zwiększyć stan magazynowy produktów danej kategorii?

Raport 5. Stworzono raport ilości sprzedawanych produktów danej kategorii w poszczególnych miesiącach na przestrzeni ostatniego roku. Raport w formie graficznej przedstawiono na Rysunek 20.



Rysunek 20. Raport 5. w formie graficznej

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport pokazuje liczbę sprzedanych produktów z danej kategorii w poszczególnych miesiącach w ciągu ostatniego roku. Pozwala zidentyfikować sezonowe wzrosty sprzedaży i przewidywać okresy zwiększonego popytu.

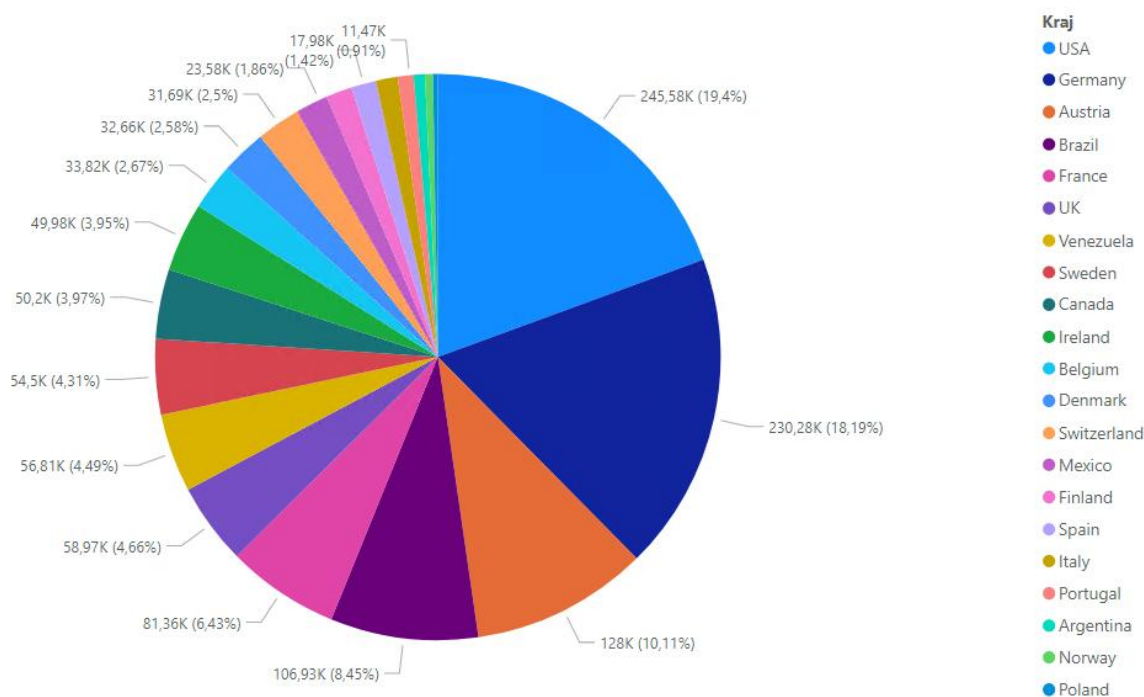
Decyzja: Zwiększono stany magazynowe dla produktów z określonych kategorii na nadchodzące miesiące, w których historycznie odnotowywano wzmożony popyt. W tym wypadku są to: “słodycze” w styczniu oraz “owoce morza” w okresie wrzesień - listopad. W tych przypadkach jest zauważalny znaczący wzrost sprzedaży względem pozostałych miesięcy. Decyzja minimalizuje ryzyko braków towaru w kluczowych okresach sprzedażowych, co zwiększa zadowolenie klientów i zapobiega utracie przychodów. Dane historyczne stanowią solidną podstawę do planowania zapasów i lepszego zarządzania magazynem.

- **Problem 6:** Współpraca z firmami z poszczególnych krajów przynosi niezadowalający zysk.

Zapytanie decyzyjne: Z klientami, z których krajów należy rozwijać współpracę? W jakich krajach reklama może przyczynić się do wzrostu zysków?

Raport 6. Stworzono raport wartości sprzedaży klientom z różnych krajów. Raport w formie graficznej przedstawiono na Rysunek 21.

Wartość sprzedaży klientom z różnych krajów



Rysunek 21. Raport 6. w formie graficznej

Hurtownie danych – sprawozdanie z zajęć laboratoryjnych	Nr ref.	HD-LAB-WCY22IJ1S1_5
	Wersja	3.0

Wsparcie w odpowiedzi na pytanie decyzyjne: Raport obrazuje wartości sprzedaży do klientów z różnych krajów, umożliwiając identyfikację rynków generujących największe i najmniejsze zyski. Dane pozwalają także ocenić, które kraje mają potencjał rozwoju dzięki reklamie lub innym działaniom wspierającym sprzedaż.

Decyzja: Podjęto decyzję o zwiększeniu współpracy z klientami z krajów o niskich zyskach oraz o intensyfikacji działań reklamowych w tych krajach, w celu zwiększenia sprzedaży i poprawy wyników finansowych - w tym wypadku są to: Polska, Norwegia, Argentyna, Portugalia, Włochy, Hiszpania, Finlandia, Meksyk, Szwajcaria, Dania oraz Belgia czyli kraje z wartością sprzedaży poniżej 3%. Decyzja opiera się na analizie rentowności rynków, wskazującej na konieczność wzmocnienia działań reklamowych w krajach o mniejszych zyskach, gdzie istnieje potencjał na poprawę wyników. Zwiększenie inwestycji w reklamę pozwoli na dotarcie do szerszej grupy odbiorców i przyciągnięcie nowych klientów, co może wpłynąć na wzrost sprzedaży i zysków w tych regionach. Strategia ta ma na celu poprawę wyników w mniej rentownych krajach poprzez lepsze wykorzystanie działań marketingowych.

9. Wnioski końcowe

Projekt miał na celu przeprowadzenie kompleksowej analizy danych oraz opracowanie narzędzi wspierających analizę funkcjonowania i pracy przedsiębiorstwa. W pierwszym etapie przeanalizowane zostały podstawowe źródła danych, co pozwoliło na zidentyfikowanie kluczowych informacji - danych o sprzedaży, klientach, pracownikach i zamówieniach. Na tej podstawie opracowana została wstępna lista raportów, które mają wspomóc podejmowanie decyzji dotyczących różnych aspektów funkcjonowania przedsiębiorstwa, takimi jak zarządzanie zapasami, przyznanie zniżek klientom, czy analiza efektywności pracowników. W oparciu o dane źródłowe zbudowany został model punktowy, który odwzorowywał strukturę danych i relacje między nimi oraz model logiczny bazy. Następnie za pomocą narzędzia SSIS zaimplementowano procesy inicjalnego zasilania hurtowni danych ETL. Procesy uwzględniały sortowanie, łączenie i agregację danych, dzięki czemu dane mogły być użyte w dalszej analizie. Na bazie zasilonej hurtowni zaimplementowano wielowymiarową kostkę OLAP, która umożliwia tworzenie analiz z uwzględnieniem hierarchii wymiarów takich jak czas, produkty, klienci, czy pracownicy. Wykorzystując kostkę w narzędziu Power BI stworzono zestaw raportów odpowiadających na zdefiniowane wcześniej pytania. Raporty obejmowały analizę zapasów produktów i sezonowych wzrostów ich sprzedaży, identyfikację najbardziej efektywnych pracowników, określenie zniżek dla najbardziej przynoszących największy dochód klientów oraz identyfikację krajów, w których przedsiębiorstwo powinno postawić na większy rozwój. Opracowane modele, procesy i raporty stanowią podstawę do prowadzenia dalszych analiz w firmie.