Dokumentacja **Zadania nr 1** z przedmiotu: „Hurtownie danych” (IZZ001103)

2020

pt. **„Hurtownie danych – Model bazy i hurtowni danych dla holdingu sklepów**

Sklep nr (smb): **19**

Sklep nazwa: **HUDA\_S1IZ18-19; Bellmont - Retail**

Pozostałe sklepy holdingu (Z1): **18,22,23**

Kategoria struktury DW (Z1): **płatek śniegu**

Wskaźnik efektywności KPI (Z1): **wskaźnik rentowności sprzedaży**

# Polecenia i punktacja do Zadania nr 1 (12,0).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Polecenie** | **Kryterium oceny** | **Max. pkt** | **Pkt**  **oc.** | ***Suma pkt*** |
| Obrona rozwiązania zadania Z1. | 01. Ocena za odp. wyb. osoby na jedno pytanie | 1,0 |  |  |
| 1. Sporządzić charakterystykę sklepu | 02. Charakterystyka sklepu na podst. zaciągniętych danych | 1,0 |  |  |
| 1.1. Utworzyć diagram ERD modelu danych sklepu handlującego artykułami pierwszej potrzeby („Szwarc mydło i powidło”) typu Biedronka (ERD, Visio) | 03. Ocena merytoryczna ERD (encje, atrybuty i związki | 0,5 |  |  |
| 04. Ocena formalna ERD sklepu | 0,5 |  |
| 1.2. Zaproponować szczegółowe definicje tabel, np. *Sprzedaz, Klienci, Produkty, KlasyProduktów, Marki, Promocje*, etc. (MS ACCESS) | 05. ocena zgodności ERD sklepu z mod. fizycz. (zupełność rozw.) | 0,5 |  |  |
| 06. Ocena definicji pól tabel | 0,5 |  |
| 07. Ocena definicji kluczy i związków | 0,5 |  |
| 08. Ocena mod. fizycznego – relacje (ACCESS) | 0,5 |  |
| 1.3. Przeanalizować możliwe KPI dla holdingu, sformułować pytania zarządcze i zarys OLAP; wybrać i uzasadnić kategorię modelu DW (gwiazda, płatek śniegu albo mieszany) i utworzyć ten model (ERD, MS Visio) | 09. Ocena merytoryczna pyt. zarz. i OLAP | 0,5 |  |  |
| 10. Ocena merytoryczna ERD (KPI) | 1,0 |  |
| 11. Ocena wyboru mod. DW (gwiazda, …) | 0,5 |  |
| 12.. Ocena formalna diagramu ERD holdingu | 0,5 |  |
| 1.4. Utworzyć fizyczną postać DW. Uzupełnić definicje tabel z punktu 1.3 o niezbędne klucze obce. Zdefiniować nowe tabele, np. *RejonySprzedazy, KlasySklepów*, etc. (MS ACCESS) **[2,0]** | 13. m.fiz. DW ocena zupełności rozw. (ERD, KPI) | 0,5 |  |  |
| 14. m.fiz. DW oc. def. pól, kluczy i związków | 0,5 |  |
| 15. m.fiz. oc. utworzenia tabeli CZAS | 0,5 |  |
| 16. m.fiz. DW relacje w MS ACCESS | 0,5 |  |
| 1.5. Opisać założenia i wykonanie procedury ładowania danych do hurtowni i ich integracji | 17. Proc. ETL: założenia i kompletność procedury | 0,5 |  |  |
| 18. Proc. ETL: wynik i ‘staging area’ | 0,5 |  |
| 19. Model fizyczny: nazewnictwo obiektów BD | 1,0 |  |
| 1.6. Sporządzić dokumentację przeprowadzonych działań | 20. Jakość dokumentacji (formatowanie, spisy i podpisy, teksty, etc ) | 0,5 |  |  |
|  |  | Razem: | |  |

|  |  |
| --- | --- |
| Opieka dydaktyczna: | **dr inż. L. Szczurowski** |

**Wrocław, maj 2021 r.**

Spis treści

[Polecenia i punktacja do Zadania nr 1 (12,0). 1](#_Toc163750140)

[Ważne uwagi do zadania Z1 (12,0). 2](#_Toc163750141)

[Opis rozwiązania ZAD nr 1. 2](#_Toc163750142)

[1. Archiwum sprzedaży i opis sklepu (wprowadzenie do problemu) 2](#_Toc163750143)

[2. Założenia modelu danych sklepu (diagram ERD, encje, atrybuty, związki) 5](#_Toc163750144)

[3. Baza danych sprzedaży sklepu typu Biedronka (tabele i relacje w MS ACCESS) 7](#_Toc163750145)

[4. Wskaźnik efektywności / skuteczności (KPI) działania holdingu sklepów 9](#_Toc163750146)

[5. Model danych hurtowni danych (gwiazda lub płatek śniegu) dla holdingu sklepów. 9](#_Toc163750147)

[6. Hurtownia danych sprzedaży holdingu sklepów (tabele i relacje w MS ACCESS) 14](#_Toc163750148)

[7. Założenia procedury ładowania danych do hurtowni danych i ich integracji 17](#_Toc163750149)

# Ważne uwagi do zadania Z1 (12,0).

# Opis rozwiązania ZAD nr 1.

## 1. Archiwum sprzedaży i opis sklepu (wprowadzenie do problemu)

**Charakteryzacja sklepu:**  
Nazwa sklepu: HUDA\_S1IZ18-19 Bellmont – Retail  
Dane adresowe: 6644 Sudance Drive, Vancouver, BC 75900 Canada

Liczba departamentów produktów: 22 (Kwerenda: HUDA19\_Departamenty produktów)

|  |
| --- |
| Departamenty produktów |
| **Alcoholic Beverages** |
| **Baked Goods** |
| **Baking Goods** |
| **Beverages** |
| **Breakfast Foods** |
| **Canned Foods** |
| **Canned Products** |
| **Carousel** |
| **Checkout** |
| **Dairy** |
| **Deli** |
| **Eggs** |
| **Frozen Foods** |
| **Health and Hygiene** |
| **Household** |
| **Meat** |
| **Periodicals** |
| **Produce** |
| **Seafood** |
| **Snack Foods** |
| **Snacks** |
| **Starchy Foods** |

Liczba kategorii produktów: 45 (Kwerenda: HUDA19\_Kategorie produktów)

|  |  |
| --- | --- |
| Liczba kategorii produktów | 45 |

Liczba promocji: 14 (Kwerenda: HUDA19\_ Liczba promocji z zerową)

|  |  |
| --- | --- |
| Liczba promocji | 14 |

Liczba klientów sklepu: 891 (Kwerenda: HUDA19\_ Liczba klientów)

|  |  |
| --- | --- |
| Liczba klientów | 891 |

Liczba marek: 111 (Kwerenda: HUDA19\_Liczba marek)

|  |  |
| --- | --- |
| Liczba marek | 111 |

**Liczba faktów sprzedaży:**

1. Poszczególne kwartały (Kwerenda: HUDA19\_Liczba faktów sprzedaży w kwartałach)

|  |  |
| --- | --- |
| Kwartał | Liczba faktów sprzedaży |
| 1 | 2673 |
| 2 | 3297 |
| 3 | 3349 |
| 4 | 3449 |

1. Poszczególne miesiące (Kwerenda: HUDA19\_Liczba faktów sprzedaży w miesiącach)

|  |  |
| --- | --- |
| Miesiąc Roku | Liczba faktów sprzedaży |
| 1 | 803 |
| 2 | 935 |
| 3 | 935 |
| 4 | 1032 |
| 5 | 1085 |
| 6 | 1180 |
| 7 | 1229 |
| 8 | 946 |
| 9 | 1174 |
| 10 | 1082 |
| 11 | 1244 |
| 12 | 1123 |

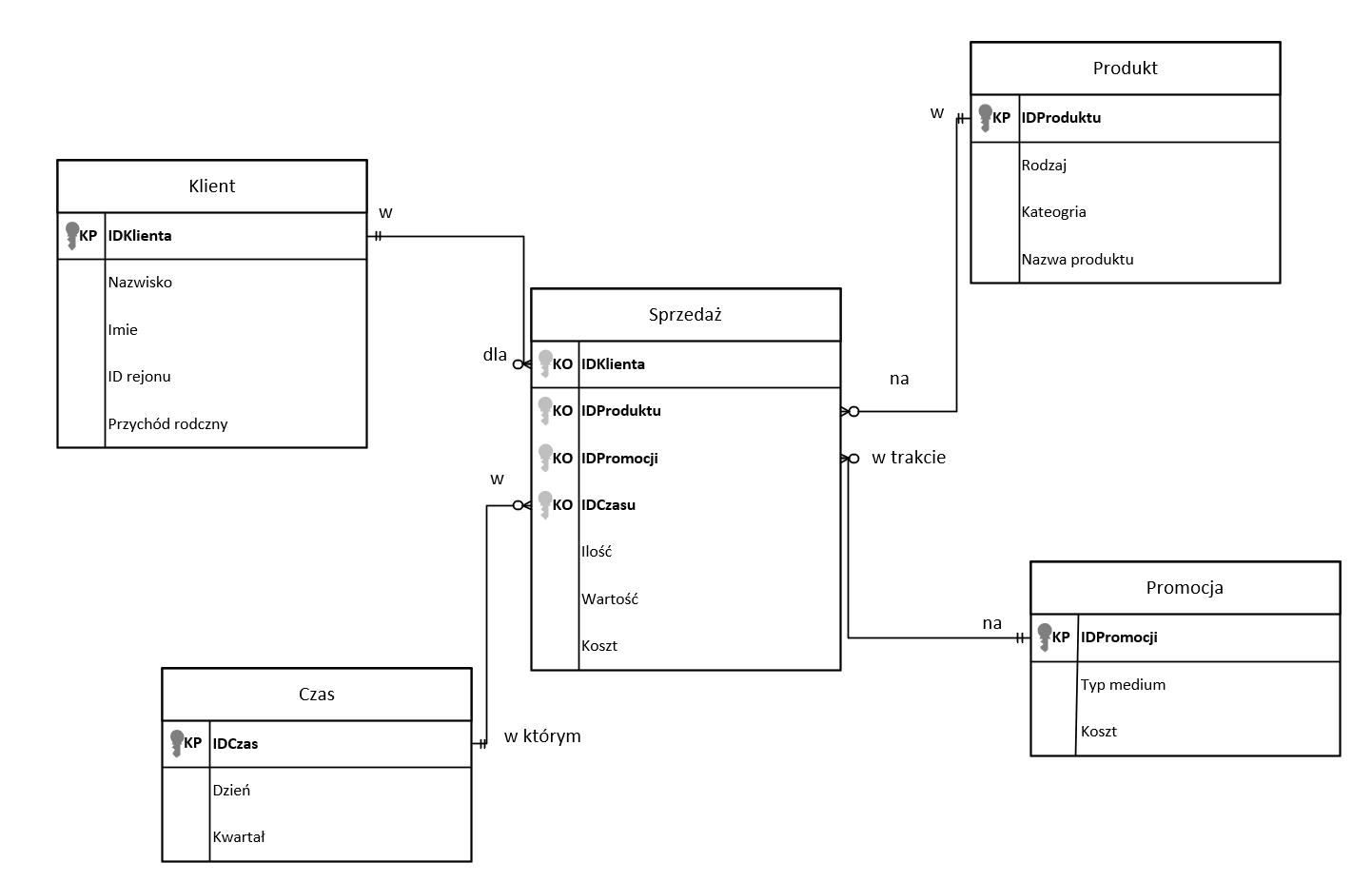
1. Poszczególne tygodnie (Kwerenda: HUDA19\_Liczba faktów sprzedaży w tygodniach)

|  |  |
| --- | --- |
| Tydzień Roku | Liczba faktów sprzedaży |
| 1 | 186 |
| 2 | 456 |
| 4 | 161 |
| 7 | 210 |
| 8 | 525 |
| 9 | 200 |
| 10 | 352 |
| 12 | 583 |
| 16 | 400 |
| 17 | 425 |
| 18 | 207 |
| 19 | 412 |
| 21 | 465 |
| 22 | 208 |
| 23 | 277 |
| 24 | 210 |
| 25 | 502 |
| 26 | 191 |
| 28 | 209 |
| 29 | 1020 |
| 31 | 392 |
| 32 | 160 |
| 33 | 215 |
| 34 | 179 |
| 36 | 453 |
| 37 | 241 |
| 38 | 480 |
| 41 | 227 |
| 42 | 181 |
| 43 | 482 |
| 44 | 192 |
| 45 | 278 |
| 46 | 270 |
| 47 | 216 |
| 48 | 480 |
| 50 | 460 |
| 51 | 222 |
| 52 | 441 |

## 2. Założenia modelu danych sklepu (diagram ERD, encje, atrybuty, związki)

Pytania zarządcze jakie stawia sobie sklep:

* W którym kwartale sprzedaż była największa?
* Czy ilość sprzedanych produktów jest zależna od przychodów rocznych klientów?
* Na jaki typ medium promocja okazała się najzyskowniejsza a która najmniej?
* Czy niektóre produkty są bardziej popularne w różnych obszarach sprzedaży?
* Czy jakiś dzień tygodnia wyróżniał się wyraźnie na tle pozostałych pod względem sprzedaży?



*Rys. 1. Diagram ERD sklepu[[1]](#footnote-1)*

Nasz sklep zajmuje się sprzedażą produktów 3 rodzajów. Nasz sklep oferuje jedzenie np. mięso, jajka, owoce i warzywa. Również w naszym sklepie można zakupić napoje oraz produkty nie nadające się do konsumpcji np. środki higieny, produkty do łazienki, sprzęty elektryczne. Istnieje również zakupienia tych towarów podczas trwających promocji.

*Tabela 1. Encje w modelu danych sklepu*

|  |  |  |  |
| --- | --- | --- | --- |
| **NES** | **Nazwa encji[[2]](#footnote-2)** | **Klucz główny** | **Opis encji** |
| 1 | Sprzedaż | - | Znajdują się tutaj informacje na temat dokonanej sprzedaży. Kiedy ona się odbyła, kto kupił, co kupił, za ile. |
| 2 | Klient | IDKlienta | Znajdują się tutaj informacje na temat klienta. Między innymi informacje o jego rachunku, miejscu zamieszkania czy wieku. |
| 3 | Produkt | IDProduktu | Znajdują się tutaj informacje na temat produktu, jego rodzaju, kategorii, marce, ile sztuk znajduje się w opakowaniu itd. |
| 4 | Promocja | IDPromocji | Znajdują się tutaj informacje na temat promocji. Jej nazwy, od kiedy trwała i do kiedy. |
| 5 | Czas | IDCzas | Znajdują się tutaj informacje na temat daty transakcji oraz jej wyszczególnienia na dni, kwartały itd. |

Najważniejszą encją jest sprzedaż. Jest ona związana z innymi encjami. Znajdują się w niej identyfikatory, które pozwalają się odnieść do innych informacji na temat klienta, produktów, czasu i promocji. Również znajdują się w niej dane na temat ilości, wartość i kosztów sprzedaży, które są bardzo istotne.

*Tabela 2. Atrybuty encji w modelu danych sklepu*

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa encji** | **Nazwa atrybutu** | **Kategoria atrybutu[[3]](#footnote-3)** | **Typ (rozmiar) atrybutu** |
| Sprzedaż | IDKlienta | Obw | Liczba |
| IDProduktu | Obw | Liczba |
| IDPromocji | Obw | Liczba |
| IDCzasu | Obw | Liczba |
| Ilość | Obw | Liczba |
| Wartość | Obw | Liczba |
| Koszt | Obw | Liczba |
| Klient | IDKlienta | Id | Liczba |
| Nazwisko | Obw. | Krótki tekst |
| Imie | Obw. | Krótki tekst |
| ID rejonu | Obw. | Liczba |
| Przychód roczny | Obw. | Krótki tekst |
| Produkt | IDProduktu | Id | Liczba |
| Rodzaj | Obw | Krótki tekst |
| Kategoria | Obw | Krótki tekst |
| Nazwa produktu | Obw | Krótki tekst |
| Promocja | IDPromocji | Id | Liczba |
| Typ medium | Obw | Krótki tekst |
| Koszt | Obw | Liczba |
| Czas | IDCzas | Id | Liczba |
| Dzień | Obw | Data/Godzina |
| Kwartał | Obw | Liczba |

Najistotniejszymi atrybutami jest ilość, wartość, koszt, które informują nas o sprzedaży i pozwalają na obliczenie KPI. Również klucze główne są istotne, znajdujące się w encjach Klient, Produkt, Promocja, Czas dzięki którym można związać encje z Sprzedaż.

*Tabela 3. Związki między encjami w modelu danych sklepu*

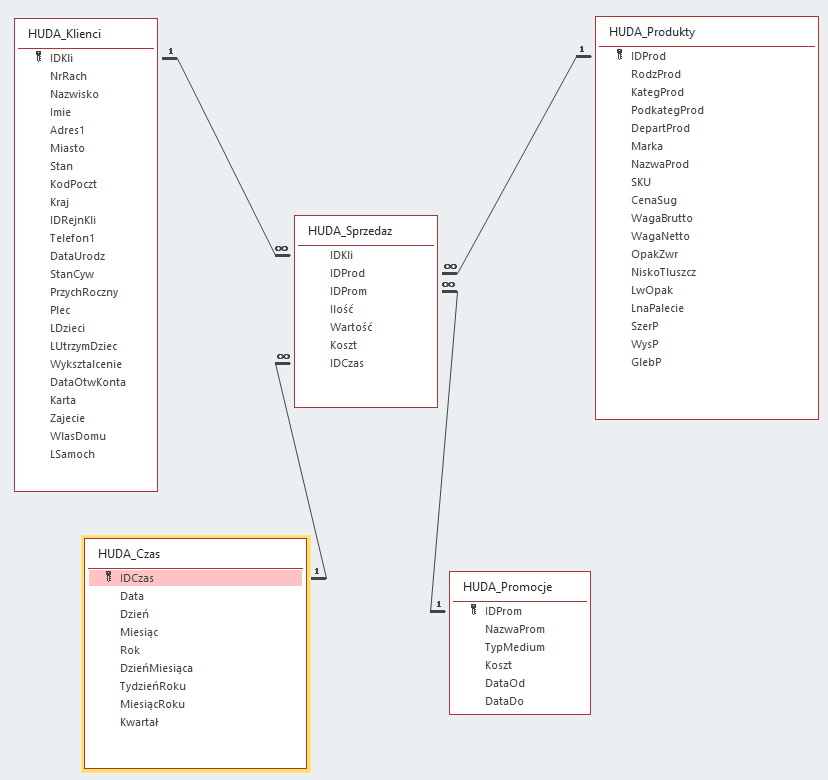
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NZS** | **Nazwa encji pierwotnej** | **Typ i nazwa związku[[4]](#footnote-4)** | **Kardynalność związku[[5]](#footnote-5)** | **Nazwa encji związanej** |
| 1 | Klient | ID „w” | 1/\* | Sprzedaż |
| 2 | Produkt | ID „w” | 1/\* | Sprzedaż |
| 3 | Promocja | ID „na” | 1/\* | Sprzedaż |
| 4 | Czas | ID „w którym” | 1/\* | Sprzedaż |

Najistotniejsze związki to Klient ze Sprzedażą, związek pozwala nam określić sprzedaż na podstawie regionu zamieszkanego przez klienta oraz sprzedaży na podstawie jego dochodów rocznych.

Istotny związek to również Czas ze Sprzedażą, który pozwala określić sprzedaż dla danych przedziałów czasu.

Również związek Produkt ze Sprzedażą, na podstawie którego jesteśmy w stanie dowiedzieć się, jakie wyniki dane produkty osiągają w sprzedaży.

## 3. Baza danych sprzedaży sklepu typu Biedronka (tabele i relacje w MS ACCESS)

**

*Rys. 2. Relacje między tabelami bazy danych sprzedaży dla sklepu typu Biedronka*

Na rysunku znajdują się relacje między tabelami. Tabela Sprzedaż posiada 4 klucze obce, które odwołują się do 4 kluczy głównych w innych tabelach. Relacje między tabelami są związane więzami integralnością i są relacjami jeden do wielu.

*Tabela 4. Tabele w bazie danych sklepu*

|  |  |  |  |
| --- | --- | --- | --- |
| **NTS** | **NES** | **Nazwa tabeli[[6]](#footnote-6)** | **Klucz główny** |
| 1 | 1 | Sprzedaż | - |
| 2 | 2 | Klienci | IDKli |
| 3 | 3 | Produkty | IDProd |
| 4 | 4 | Promocje | IDProm |
| 5 | 5 | Czas | IDCzas |

*Tabela 5. Pola tabel w bazie danych sklepu*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NTS** | **Nazwa tabeli** | **Nazwa pola** | **Kategoria pola[[7]](#footnote-7)** | **Typ (rozmiar) pola** |
| 1 | Sprzedaż | IDKli | IdO | Liczba |
| IDProd | IdO | Liczba |
| IDProm | IdO | Liczba |
| IDCzas | IdO | Liczba |
| Ilość | Opc | Liczba |
| Wartość | Opc | Liczba |
| Koszt | Opc | Liczba |
| 2 | Klienci | IDKli | Id | Liczba |
| NrRach | Obw | Liczba |
| Nazwisko | Obw | Krótki tekst |
| Imie | Obw | Krótki tekst |
| Adres1 | Opc | Krótki tekst |
| Miasto | Obw | Krótki tekst |
| Stan | Obw | Krótki tekst |
| KodPoczt | Opc | Krótki tekst |
| Kraj | Obw | Krótki tekst |
| IDRejnKli | Obw | Liczba |
| Telefon1 | Opc | Krótki tekst |
| DataUrodz | Opc | Data/Godzina |
| StanCyw | Opc | Krótki tekst |
| PrzychRoczny | Obw | Krótki tekst |
| Plec | Opc | Krótki tekst |
| LDzieci | Opc | Liczba |
| LUtrzymDziec | Opc | Liczba |
| Wykształcenie | Opc | Krótki tekst |
| DataOtwKonta | Opc | Data/Godzina |
| Karta | Opc | Krótki tekst |
| Zajecie | Opc | Krótki tekst |
| WlasDomu | Opc | Krótki tekst |
| LSamoch | Opc | Liczba |
| 3 | Produkty | IDprom | Id | Liczba |
| RodzProd | Obw | Krótki tekst |
| KategProd | Obw | Krótki tekst |
| PodkategProd | Opc | Krótki tekst |
| DepartProd | Opc | Krótki tekst |
| Marka | Opc | Krótki tekst |
| NazwaProd | Obw | Krótki tekst |
| SKU | Opc | Liczba |
| CenaSug | Obw | Liczba |
| WagaBrutto | Opc | Liczba |
| WagaNetto | Opc | Liczba |
| OpakZwr | Opc | Tak/Nie |
| NiskoTluszcz | Opc | Tak/Nie |
| LwOpak | Opc | Liczba |
| LnaPalecie | Opc | Liczba |
| SzerP | Opc | Liczba |
| WysP | Opc | Liczba |
| GlebP | Opc | Liczba |
| 4 | Promocje | IDProm | Id | Liczba |
| NazwaProm | Obw | Krótki tekst |
| TypMedium | Obw | Krótki tekst |
| Koszt | Obw | Liczba |
| DataOd | Obc | Data/Godzina |
| DataDo | Obc | Data/Godzina |
| 5 | Czas | IDCzas | Id | Liczba |
| Data | Obw | Data/Godzina |
| Dzień | Obw | Data/Godzina |
| Miesiąc | Obc | Krótki tekst |
| Rok | Obc | Liczba |
| DzieńMiesiąca | Opc | Liczba |
| TydzieńRoku | Opc | Liczba |
| MiesiącRoku | Opc | Liczba |
| Kwartał | Obw | Liczba |

Najistotniejszymi polami będą klucze obce w tabeli Sprzedaż, które odwołują się również do istotnych kluczy głównych w pozostałych tabelach.

## 4. Wskaźnik efektywności / skuteczności (KPI) działania holdingu sklepów

Wskaźnik na który się zdecydowaliśmy to wskaźnik rentowności sprzedaży (return on sales, ROS). Wskaźnik informuje nas o tym ile zysku netto uzyskujemy ze sprzedaży. Wskaźnik jest stosunkiem zysku netto do przychodów.

Dla struktury DW obliczyliśmy wskaźnik ROS z pomocą kwerendy „20\_Pn1315\_19\_Obliczanie KPI w DW” która pozwoliła posumować wartości sprzedanych produktów oraz koszta jakie poniosła firma przy każdej dokonanej transakcji. Następnie podstawiając do wzoru na ROS podane dane (Wartość: 195116,48, Koszt: 78109,7) obliczyliśmy jego wartość na 59,96766%. Interpretować to można jako, że sieć holdingowa przy każdej transakcji zyskiwała około 60% ceny towaru zakupionego przez klienta.

## 5. Model danych hurtowni danych (gwiazda lub płatek śniegu) dla holdingu sklepów.

* Popularność jakich produktów wzrosła a jakich zmalała?
* Jakie produkty są sezonowe?
* Którzy klienci składają regularnie takie same zamówienia?
* Czy niektóre produkty są bardziej popularne w różnych obszarach?

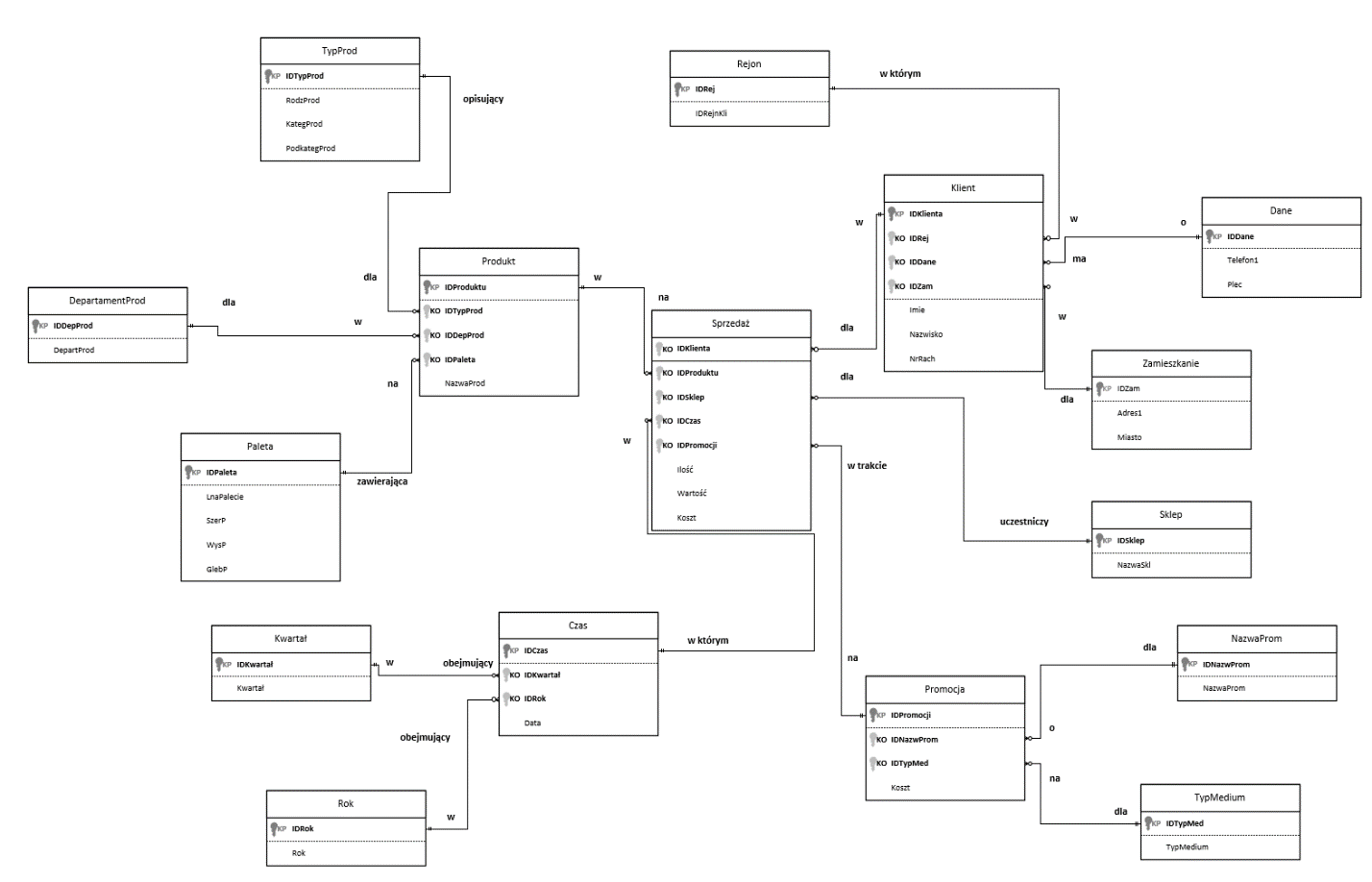
*Tabela 6. Pytania zarządcze i analizy w zarzadząniu modelu danych sklepu*

|  |  |  |
| --- | --- | --- |
| **NPZ** | **Nazwa pytania zarządczego** | **Opis analiz związanych z pytaniem** |
| 1 | Popularność jakich produktów wzrosła a jakich zmalała? | Podstawą analizy odpowiedzi na to pytanie stałaby się analiza zmienności ilości nabywania konkretnych produktów w czasie, oraz ustalenie czy istniała tendencja wzrostowa lub spadkowa dla konkretnego produktu np. na wykresie. |
| 2 | Jakie produkty są sezonowe | W tym miejscu podstawą analizy byłoby kwartalne sprawdzenie tendencji wzrostowych dla konkretnego rodzaju produktów. Na podstawie wiedzy o ilości sprzedaży w danym kwartale możliwe stałoby się czy popularność danego produktu jest związana z sezonem. |
| 3 | Którzy klienci składają regularnie takie same zamówienia | Konieczne byłoby ustalenie regularności np. powyżej jednego takiego samego zamówienia przez klienta miesięcznie, następnie wynikami byłby dane zwracające nazwisko klienta wraz z produktem oraz jego ilością, która została przez klienta zakupiona. |
| 4 | Czy niektóre produkty są bardziej popularne w różnych obszarach. | Popularność będziemy definiować jako wiele, niekoniecznie dużych zamówień danego produktu przez różnych klientów, następnie skupiając się na jednym produkcie z pomocą wykresu łatwo odczytalibyśmy dane odnośnie większej popularności produktu w pewnym rejonie. |

Z punktu widzenia holdingu sklepów, najważniejszymi pytaniami zarządczymi są pytania 1 oraz 4.

Pierwsze pytanie w klarowny sposób pokaże trendy nabywcze klientów oraz będzie służył jako podpowiedź odnośnie tego jakie towary są najopłacalniejsze w kwestii sprzedaży, zarówno w chwili obecnej jak i w przyszłości.

Pytanie czwarte natomiast przedstawi jak zróżnicowany jest rynek w sposób geograficzny i pozwoli poznać, który region składa się z dużej liczby klientów jednostkowych a który region opiewa na klientów nabywających w naszym sklepie duże ilości produktów.



*Rys.3. Schemat płatka śniegu hurtowni danych holdingu sklepów*

Wybrany model płatka śniegu jest rozszczepieniem początkowego schematu gwiazdy. Z punktu widzenia osób administrujących bazą danych jest to rozwiązanie logiczne ze względu na uporządkowanie konkretnych, związanych ze sobą danych w tabelach. Dzięki temu możliwa jest prostsza oraz szybsza praca w momencie znajdywania odpowiedzi na konkretne szczegółowe zapytania. Analizy jakich przewiduje się przeprowadzanie w systemie sprzedaży holdingu DW, to przede wszystkim analizy związane z odpowiedziami na pytania zarządcze oraz pozwalające na liczenie różnego rodzaju wskaźników rentowności sprzedaży. Elementami systemu CRM wspieranymi przez analizy danych z DW będą przede wszystkim informacje o kartach posiadanych przez klientów oraz ich numery telefonów. Oprócz tego możliwym będzie wspieranie zarządzania namiarami, automatyzacja sprzedaży oraz zarządzanie klientami kluczowymi. Przed wyborem modelu struktury bazy danych należy zadać sobie pytanie o to w jaki sposób będziemy z niej korzystać tj. czy dane których będziemy często poszukiwać powinny być wyodrębnione w osobnych tabelach dla łatwego ich znajdowania oraz zarządzania czy też powinny zostać zintegrowane w jedność dla większej kumulacji danych przy wyświetlaniu jednej konkretnej. Model ten od przedstawionego na Rys.1 różnił się będzie strukturą oraz ilością tabeli z podtabelami do których zostały rozproszone dane z tabeli głównych z Rys.1. Encja faktów sprzedaży jest najważniejszą encją w całym modelu DW, jako że za jej pomocą możliwa jest nawigacja do innych encji powiązanych. Zawiera ona informacje o tym, w jakim sklepie została dokonana sprzedaż, kto zakupił jaki produkt, w trakcie jakiej promocji oraz informacje tyczące się ilości, wartości oraz kosztów poniesionych przez naszą firmę w danym fakcie sprzedaży. Encja wymiarów składa się przede wszystkim z tabeli Sprzedaż, w której to zawierają się klucze typu ID odwołujące się do dalszych tabel takich jak S19\_Czas, S19\_Produkty, S19\_Promocje itd. Są to tabele wymiaru pierwszego w modelu płatka śniegu (posiadają bezpośrednią relację z tabelą S19\_Sprzedaż), encjami podwymiarów są tabele które swój klucz główny posiadają zapisany w ten sam sposób w jaki w tabelach wymiaru pierwszego zapisane są atrybuty o początku ID, które nie posiadają klucza głównego, do tego typu tabel należą m.in. S19\_Rok, S19\_DepartamentProd, S19\_NazwaProm.

*Tabela 7. Encje w modelu hurtowni danych*

|  |  |  |  |
| --- | --- | --- | --- |
| **NEH** | **Nazwa encji[[8]](#footnote-8)** | **Klucz główny** | **Opis encji / zmian w encji** |
| 1 | Sprzedaż | - | Znajdują się tutaj informacje na temat dokonanej sprzedaży. Kiedy ona się odbyła, kto kupił, co kupił, za ile. Jest ona rozszerzeniem encji modelu danych sklepu. |
| 2 | Klient | IDKlienta | Znajdują się tutaj ogólne informacje na temat klienta. Takie jak imię, nazwisko czy numer rachunku. Dane dodatkowe o kliencie i te związane z lokalizacją zamieszkania zostały przeniesione do podwymiaru encji |
| 3 | Rejon | IDRej | Zawiera dane na temat rejonu, w którym klient mieszka |
| 4 | Zamieszkanie | IDZam | Zawiera informacje na temat lokalizacji zamieszkania klienta. Między innymi jak np. adres w którym mieszka dany klient |
| 5 | Dane | IDDane | Posiada informacje na temat dodatkowych danych o kliencie. Jakiej płci jest klient |
| 6 | Produkt | IDProduktu | Została zwężona w stosunku do encji bazy danych sklepu. Informacje na temat produktu. Między innymi jej nazwy. Dane o palecie i tych związanych z departamentem oraz rodzajem, kategorią i podkategorią zostały przeniesione do podwymiaru encji |
| 7 | TypProd | IDTypProd | Przedstawia rodzaj, kategorie i podkategorie danego produktu |
| 8 | DepartamentProd | IDDepProd | Znajdują się tutaj informacje na temat departamentu produktu |
| 9 | Paleta | IDPaleta | Zawiera informacje na temat palet produktu. Ich długość, szerokość, głębokość, liczba |
| 10 | Promocja | IDPromocji | Znajdują się tutaj informacje na temat promocji, jej kosztów. Dane związane z nazwą promocji i typem medium zostały przeniesione do podwymiaru encji. Jest ona zwężeniem encji z modelu danych sklepu |
| 11 | NazwaProm | IDNazwProm | Zawiera nazwy danej promocji |
| 12 | TypMedium | IDTypMed | Posiada informacje na temat typu medium danej promocji |
| 13 | Czas | IDCzas | Zawiera informacje na temat daty transakcji. Dane związane z kwartałem i rokiem, w którym zostały wykonane transakcje są przeniesione do podwymiaru encji. |
| 14 | Kwartał | IDKwartał | Znajdują się tutaj informacje na temat wyszczególnionych kwartałów poszczególnych dat |
| 15 | Rok | IDRok | Informacje o roku, w którym została wykonana transakcja |
| 16 | Sklep | IDSklep | Znajdują się tutaj informacje na temat danych sklepu |

W modelu hurtowni danych w stosunku do modelu danych sklepu zmieniła się struktura danych. Stworzyliśmy podwymiary dla encji Klienci, Produkty, Promocja, Czas. Encja czas została rozbita na 2 podwymiary, które pomogą nam w analizie jednego z pytań zarządczych, odnośnie sezonowości produktów.

*Tabela 8. Atrybuty encji w modelu hurtowni danych*

|  |  |  |  |
| --- | --- | --- | --- |
| **Nazwa encji** | **Nazwa atrybutu** | **Kategoria atrybutu[[9]](#footnote-9)** | **Typ (rozmiar) atrybutu** |
| Sprzedaż | IDKli | IdO | Liczba |
| IDProd | IdO | Liczba |
| IDSklep | IdO | Liczba |
| IDProm | IdO | Liczba |
| IDCzas | IdO | Liczba |
| Ilość | Opc | Liczba |
| Wartość | Opc | Liczba |
| Koszt | Opc | Liczba |
| Klient | IDKli | Id | Liczba |
| IDRej | IdO | Liczba |
| IDDaneKli | IdO | Liczba |
| IDZam | IdO | Liczba |
| NrRach | Obw | Liczba |
| Nazwisko | Obw | Krótki tekst |
| Imie | Obw | Krótki tekst |
| Rejon | IDRej | Id | Liczba |
| IDRejnKli | Obw | Liczba |
| Zamieszkanie | IDZam | Id | Liczba |
| Miasto | Obw | Krótki tekst |
| Adres1 | Obc | Krótki tekst |
| Dane | IDDane | Id | Liczba |
| Telefon1 | Obc | Krótki tekst |
| Plec | Obc | Krótki tekst |
| Produkt | IDProduktu | Id | Liczba |
| IDTypProd | IdO | Liczba |
| IDDepProd | IdO | Liczba |
| IDPaleta | IdO | Liczba |
| NazwaProd | Obw | Krótki tekst |
| TypProd | IDTypProd | Id | Liczba |
| RodzProd | Obw | Krótki tekst |
| KategProd | Obw | Krótki tekst |
| PodkategProd | Obc | Krótki tekst |
| DepartamentProd | IDDepProd | Id | Liczba |
| DepartProd | Obc | Krótki tekst |
| Paleta | IDPaleta | Id | Liczba |
| LnaPalecie | Obc | Liczba |
| SzerP | Obc | Liczba |
| WysP | Obc | Liczba |
| GlebP | Obc | Liczba |
| Promocja | IDProm | Id | Liczba |
| IDNazwProm | IdO | Liczba |
| IDTypMed | IdO | Liczba |
| Koszt | Obw | Liczba |
| NazwaProm | IDNazwProm | Id | Liczba |
| NazwaProm | Obw | Krótki tekst |
| TypMedium | IDTypMed | Id | Liczba |
| TypMedium | Obw | Krótki tekst |
| Czas | IDCzas | Id | Liczba |
| IDKwartał | IdO | Liczba |
| IDRok | IdO | Liczba |
| Data | Obw | Data/Godzina |
| Kwartał | IDKwartał | Id | Liczba |
| Kwartał | Obw | Liczba |
| Rok | IDRok | Id | Liczba |
| Rok | Obc | Liczba |
| Sklep | IDSklep | Id | Liczba |
| NazwaSkl | Obw | Krótki tekst |

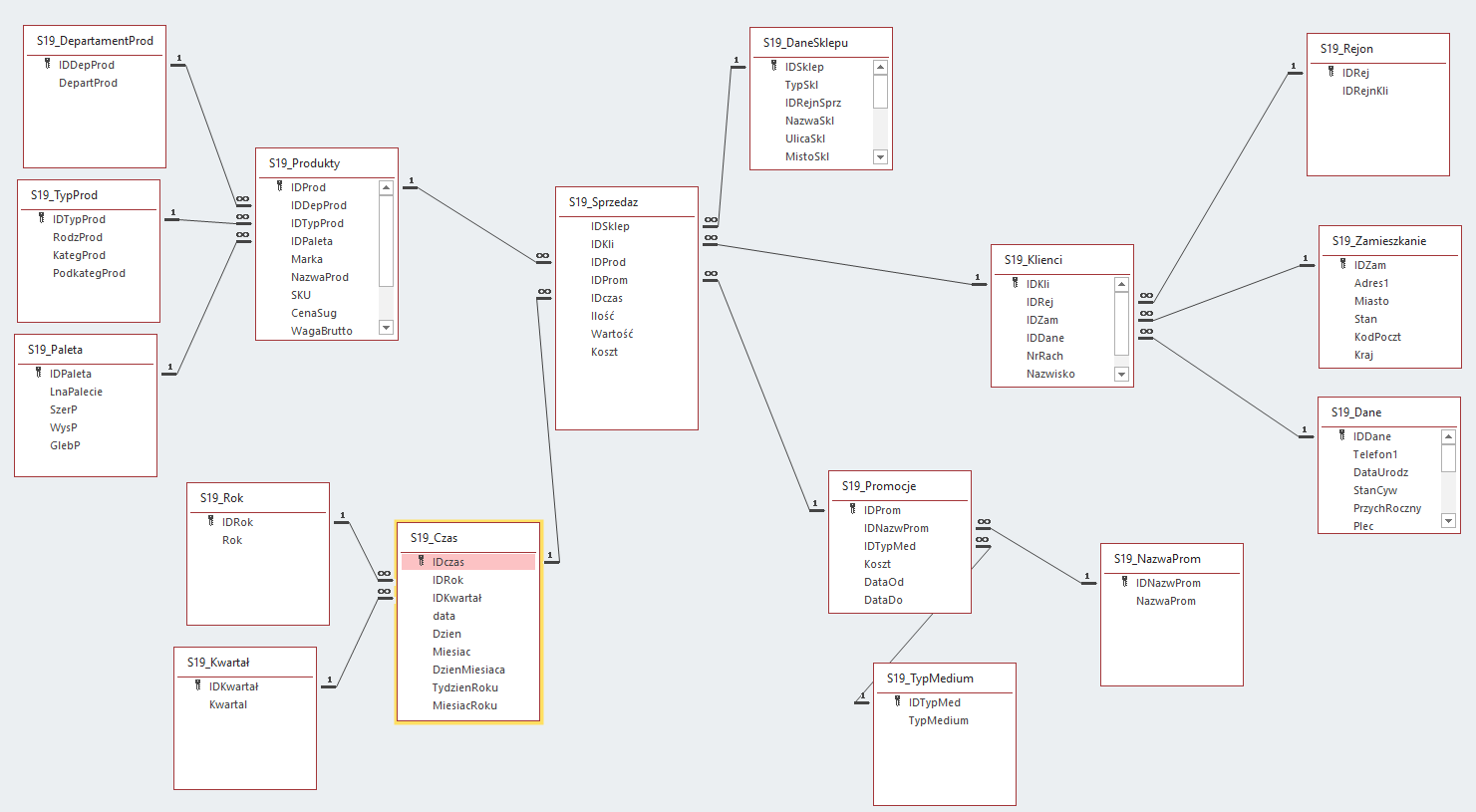
Najistotniejszymi atrybutami są m. in. NrRachunku, Nazwisko klienta, które pozwolą nam na identyfikacje klientów, którzy regularnie składają u nas zamówienia na dane produkty. Również istotne jest IDRejonuKlienta, które pozwoli nam ustalić w jakich obszarach jakie produkty są najpopularniejsze. Także nazwa, typ produktu, dzięki temu dowiemy się jakie produkty ostatnio stały się bardziej popularne, a które mniej.

*Tabela 9. Związki między encjami w modelu hurtowni danych*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NZH** | **Nazwa encji pierwotnej** | **Typ i nazwa związku[[10]](#footnote-10)** | **Kardynalność związku[[11]](#footnote-11)** | **Nazwa encji związanej** |
| 1 | Klient | ID „w” | 1/\* | Sprzedaż |
| 2 | Produkt | ID „w” | 1/\* | Sprzedaż |
| 3 | Promocja | ID „na” | 1/\* | Sprzedaż |
| 4 | Czas | ID „w którym” | 1/\* | Sprzedaż |
| 5 | Sklep | ID „uczestniczy” | 1/\* | Sprzedaż |
| 6 | Rejon | ID „w którym” | 1/\* | Klient |
| 7 | Dane | ID „o” | 1/\* | Klient |
| 8 | Zamieszkanie | ID „dla” | 1/\* | Klient |
| 9 | TypProd | ID „opisujący” | 1/\* | Produkt |
| 10 | DepartamentProd | ID „dla” | 1/\* | Produkt |
| 11 | Paleta | ID „zawierająca” | 1/\* | Produkt |
| 12 | Kwartał | ID „w” | 1/\* | Czas |
| 13 | Rok | ID „w” | 1/\* | Czas |
| 14 | NazwaProm | ID „dla” | 1/\* | Promocja |
| 15 | TypMedium | ID „dla” | 1/\* | Promocja |
| 16 | Sklep | ID „uczestniczy” | 1/\* | Sprzedaż |

Zdecydowaliśmy się na strukturę płatka śniegu więc doszło do normalizacji na skutek, której pojawiły się nowe związki. Najistotniejszymi związkami na pewno będzie związek o numerze 12 i 13, czyli kwartał i rok. Związki pomogą nam w odpowiedzi na pytanie, które produkty są sezonowe. Również ważnym związkiem jest Sklep związany ze Sprzedażą, który informuje nas, w którym sklepie dokonano sprzedaży.

## 6. Hurtownia danych sprzedaży holdingu sklepów (tabele i relacje w MS ACCESS)



*Rys. 4. Relacje między tabelami hurtowni danych sprzedaży dla holdingu sklepów*

Na rysunku znajdują się tabele w hurtowni danych sprzedaży dla holdingu sklepów, które są połączone w relacje. Tabele połączone są w schemacie płatka śniegu przez co dane są znormalizowane. Ponadto względem bazy danych sklepu hurtownia została wzbogacona o tabele DaneSklepu, która pozwala na określenie, w którym sklepie dokonano sprzedaży.

*Tabela 10. Tabele w hurtowni danych holdingu sklepów*

|  |  |  |  |
| --- | --- | --- | --- |
| **NTH** | **NEH** | **Nazwa tabeli[[12]](#footnote-12)** | **Klucz główny** |
| 1 | 1 | Sprzedaż | - |
| 2 | 2 | Klient | IDKlienta |
| 3 | 3 | Rejon | IDRej |
| 4 | 4 | Zamieszkanie | IDZam |
| 5 | 5 | Dane | IDDane |
| 6 | 6 | Produkt | IDProduktu |
| 7 | 7 | TypProd | IDTypProd |
| 8 | 8 | DepartamentProd | IDDepProd |
| 9 | 9 | Paleta | IDPaleta |
| 10 | 10 | Promocja | IDPromocji |
| 11 | 11 | NazwaProm | IDNazwProm |
| 12 | 12 | TypMedium | IDTypMed |
| 13 | 13 | Czas | IDCzas |
| 14 | 14 | Kwartał | IDKwartał |
| 15 | 15 | Rok | IDRok |
| 16 | 16 | DaneSklepu | IDSklep |

*Tabela 11. Pola tabel w hurtowni danych holdingu sklepów*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **NTH** | **Nazwa tabeli** | **Nazwa pola** | **Kategoria pola[[13]](#footnote-13)** | **Typ (rozmiar) pola** |
| 1 | Sprzedaż | IDKli | IdO | Liczba |
| IDProd | IdO | Liczba |
| IDSklep | IdO | Liczba |
| IDProm | IdO | Liczba |
| IDCzas | IdO | Liczba |
| Ilość | Opc | Liczba |
| Wartość | Opc | Liczba |
| Koszt | Opc | Liczba |
| 2 | Klient | IDKli | Id | Liczba |
| IDRej | IdO | Liczba |
| IDDaneKli | IdO | Liczba |
| IDZam | IdO | Liczba |
| NrRach | Obw | Liczba |
| Nazwisko | Obw | Krótki tekst |
| Imie | Obw | Krótki tekst |
| 3 | Rejon | IDRej | Id | Liczba |
| IDRejnKli | Obw | Liczba |
| 4 | Zamieszkanie | IDZam | Id | Liczba |
| Adres1 | Opc | Krótki tekst |
| Miasto | Obw | Krótki tekst |
| Stan | Obw | Krótki tekst |
| KodPoczt | Opc | Krótki tekst |
| Kraj | Obw | Krótki tekst |
| 5 | Dane | IDDane | Id | Liczba |
| Telefon1 | Opc | Krótki tekst |
| DataUrodz | Opc | Data/Godzina |
| StanCyw | Opc | Krótki tekst |
| PrzychRoczny | Obw | Krótki tekst |
| Plec | Opc | Krótki tekst |
| LDzieci | Opc | Liczba |
| LUtrzymDziec | Opc | Liczba |
| Wykształcenie | Opc | Krótki tekst |
| DataOtwKonta | Opc | Data/Godzina |
| Karta | Opc | Krótki tekst |
| Zajecie | Opc | Krótki tekst |
| WlasDomu | Opc | Krótki tekst |
| LSamoch | Opc | Liczba |
| 6 | Produkt | IDProduktu | Id | Liczba |
| IDTypProd | IdO | Liczba |
| IDDepProd | IdO | Liczba |
| IDPaleta | IdO | Liczba |
| NazwaProd | Obw | Krótki tekst |
| SKU | Opc | Liczba |
| CenaSug | Obw | Liczba |
| WagaBrutto | Opc | Liczba |
| WagaNetto | Opc | Liczba |
| OpakZwr | Opc | Tak/Nie |
| NiskoTluszcz | Opc | Tak/Nie |
| LwOpak | Opc | Liczba |
| Marka | Opc | Krótki tekst |
| 7 | TypProd | IDTypProd | Id | Liczba |
| RodzProd | Obw | Krótki tekst |
| KategProd | Obw | Krótki tekst |
| PodkategProd | Opc | Krótki tekst |
| 8 | DepartamentProd | IDDepProd | Id | Liczba |
| DepartProd | Opc | Krótki tekst |
| 9 | Paleta | IDPaleta | Id | Liczba |
| LnaPalecie | Opc | Liczba |
| SzerP | Opc | Liczba |
| WysP | Opc | Liczba |
| GlebP | Opc | Liczba |
| 10 | Promocja | IDProm | Id | Liczba |
| IDNazwProm | IdO | Liczba |
| IDTypMed | IdO | Liczba |
| Koszt | Obw | Liczba |
| DataOd | Opc | Data/Godzina |
| DataDo | Opc | Data/Godzina |
| 11 | NazwaProm | IDNazwProm | Id | Liczba |
| NazwaProm | Obw | Krótki tekst |
| 12 | TypMedium | IDTypMed | Id | Liczba |
| TypMedium | Obw | Krótki tekst |
| 13 | Czas | IDCzas | Id | Liczba |
| IDKwartał | IdO | Liczba |
| IDRok | IdO | Liczba |
| Data | Obw | Data/Godzina |
| Dzień | Obw | Data/Godzina |
| Miesiąc | Opc | Krótki tekst |
| DzieńMiesiąca | Opc | Liczba |
| TydzieńRoku | Opc | Liczba |
| MiesiącRoku | Opc | Liczba |
| 14 | Kwartał | IDKwartał | Id | Liczba |
| Kwartał | Obw | Liczba |
| 15 | Rok | IDRok | Id | Liczba |
| Rok | Opc | Liczba |
| 16 | DaneSklepu | IDSklep | Id | Liczba |
| TypSkl | Obw | Krótki tekst |
| IDRejnSprz | Obw | Liczba |
| NazwaSkl | Obw | Krótki tekst |
| UlicaSkl | Obw | Krótki tekst |
| MiastoSkl | Obw | Krótki tekst |
| StanSkl | Obw | Krótki tekst |
| KodPSkl | Obw | Krótki tekst |
| KrajSkl | Obw | Krótki tekst |
| KierSkl | Obw | Krótki tekst |
| TelSkl | Obw | Krótki tekst |
| DataOtw1 | Obw | Data/Godzina |
| DataZmiany1 | Opc | Data/Godzina |

NrRach, Imię, Nazwisko są istotnymi polami, które identyfikują klienta, dzięki tym informacjom jesteśmy w stanie dokonać analizy, którzy klienci kupują u nas najczęściej i jakie produkty. NazwaProduktu i TypProduktu pozwolą nam zweryfikować, popularność jakich ostatnio wzrosła, a która zmalała. Rok, Kwartał również są istotne mianowicie pozwolą nam określić, które produkty są sezonowe.

Pomysłem „zaciągnięcia” danych do DW tabel jest zintegrowanie ich wcześniej czyli połączenie, aby móc już gotowe tabele od razu załadować do Hurtowni Danych.

## 7. Założenia procedury ładowania danych do hurtowni danych i ich integracji

W celu załadowania zintegrowanych niesprzecznych danych sklepów muszą zostać podjęte konkretne działania. Początkowo trzeba zaciągnąć wszystkie dane udostępnione przez udziałowców do nowego pliku bazo danowego. Następnie trzeba połączyć ze sobą tożsame tabele. Jeśli to konieczne wcześniej znormalizować lub zdenomarlizować dane. Przed scaleniem, również konieczne jest sprawdzenie formatów, aby się ze sobą zgadzały. Po zintegrowaniu danych należy przystąpić do budowy naszej struktury hurtowni danych.

Wszystkie powyższe czynności zostały przeprowadzone na pliku bazodanowym HUDA\_S1IZ20\_Pn1315\_19-Z1\_SA (Staging Area)

Każdy z udziałowców udostępnił swoją bazę danych w planowanym terminie 23 kwietnia. Wszystkie bazy danych udziałowców posiadały strukturę gwiazdy zarówno jak nasza baza danych. Zaimportowaliśmy wszystkie tabele do jednego pliku czyli naszej SA, udało się od razu zaimportować wszystkie dane bez problemu.

Jeszcze uprzednio wyeliminowaliśmy błędy na poziomie bazy danych:

* Błąd 1:Promocja trwa od 12 grudnia 2007 roku do 13 miesiąca 19 dnia. Najlogiczniejszym rozwiązaniem jest to, gdy promocja trwa do 19 dnia ale 12 miesiąca (13 miesiąc nie istnieje). 
* Błąd 2:W tabeli Sprzedaż znajduje się ID klienta (99969), którego brakuje w tabeli Klienci. Postanowiliśmy usunąć informacje o sprzedaży skoro klient nie figuruje w Tabeli klienci.   
  Kwerenda: HUDA19\_Zlokalizowanie błędu 2
* Błąd 3:W tabeli sprzedaży bardzo dużo sprzedaży z IDPromocji o numerze 0. Oznacza to po prostu brak promocji i taką promocje o ID 0 i nazwie „Brak promocji” dodaliśmy do tabeli Promocje.
* Błąd 4:  
  W tabeli Sprzedaży produkt, który nie istnieje w tabeli Produkty. Postanowiliśmy usunąć ID tego produktu i jego dane z tabeli Sprzedaż.  
  Kwerenda: HUDA19\_Zlokalizowanie błędu 4
* Błąd 5:  
  W tabeli Sprzedaży produkt o IDProd 616 posiadał zadeklarowaną ilość dnia 05.01.2007 równa -2. Jako, że nie można sprzedać minus dwóch produktów postanowiliśmy zamienić tą ilość na 2. Błąd został znaleziony przy dodawaniu reguły walidacji.
* Błąd6:  
  W tabeli sprzedaży zlokalizowaliśmy datę w której rok sprzedaży zamiast 2007 był rokiem 2008. Jako, że był to jedyny rekord z taką datą uznaliśmy to za błąd i postanowiliśmy zmienić na 2007 tak jak reszta rekordów.

Po zaimportowaniu wszystkich danych do SA, trzeba było przejść do transformacji i czyszczenia danych, by móc je ze sobą połączyć. Poniżej przedstawione są operacje, których celem było dostosowanie danych:

* Usunięcie godziny otwarcia i zamknięcia w HUDA23\_DaneSklepu. Ta tabela jako jedyna ze wszystkich udziałowców posiadały te dane, w celu połączenia ze sobą wszystkich danych postanowiliśmy usunąć te informacje
* Zmienienie rozmiaru pola z liczy całkowitej na podwójną precyzję dla Ilość w tabeli HUDA22\_Sprzedaz.
* Dla tabeli HUDA18\_Sprzedaz została utworzona nowa tabela czas, ponieważ w sprzedaży nie znajdował się identyfikator czasu tylko data, połączona z datą w tabeli czas, w której również nie było identyfikatora.
* Utworzenie HUDA18\_Sprzedaz z IDCzas, żeby móc powiązać nową tabele czas również usunięcie daty z tej tabeli
* W HUDA23\_Sprzedaz usunięcie kolumny data, identyfikator czasu już się tu znajdował. Również zmiana typu danych z Waluty na Liczbę dla Wartości.
* Zmiana typu w tabeli HUDA23\_Czas w polu dzień z krótkiego tekstu na data/czas przy uwzględnieniu formatu „dddd” w celu pokazania dnia miesiąca.
* HUDA22\_Czas zmiana formatu daty z „dd-mmm-rr” na „dd\.mm\.rrrr”
* Utworzenie nowych tabel Sprzedaż z uwzględnieniem IDSklepu. Wszystkie tabele zostały utworzone na podstawie HUDA22\_Czas, ponieważ tabela pokrywała wszystkie daty wszystkich sklepów co umożliwiło zintegrowanie tych tabel na tym czasie. Nadane zostały nowe identyfikatory czasu dla tabeli sprzedaż.
* HUDA23\_Produkty zmiana typu CenaSug z Waluty na Liczbę.
* HUDA22\_Produkty zmiana typu danych dla CenaSug z Waluty na Liczbę.

Następnie należało połączyć ze sobą dane:

* Dla tabeli produkty, identyfikatory pokrywały się z innymi produktami udziałowców, przy imporcie produktów innych sklepów do naszych produktów, pojawił się komunikat ze strony głównego o duplikatach, czyli do tabeli udało się zaimportować tylko nowe produkty, więc udało się scalić produkty
* Połączenie tabel sprzedaż obyło się bez problemów po etapie transformacji udało się od razu połączyć ze sobą tabele w jedną całość
* Klienci podobnie jak produkty posiadali swoje unikalne identyfikatory dla całej SA czyli wystarczyło dodać do klientów z innych sklepów do naszych, oczywiście dodani zostali tylko nowi klienci, nic się zduplikowało.
* Połączenie promocji również obyło się bez problemu wszystkie dane się zgadzały.
* Różnych tabel Czas już nie trzeba było łączyć, ponieważ wcześniej sprzedaże zostały powiązane z jedną tabelą czas, w której znajdywały się wszystkie daty.

Kolejnym etapem była normalizacja danych w celu utworzenia płatka śniegu. Proces polegał na wybraniu pól tabeli np. Produkty, które znajdą się w podwymiarze, tabela podwymiaru była tworzona od razu z pogrupowanymi wartościami żeby nie wystąpiły błędu podczas łączenia tabel w relacji ze strony klucza głównego. Następnie utworzenie tabeli z samymi tymi wartościami i dodanie do niej klucza głównego z autonumeracją. Kolejna trzeba było za pomocą wcześniejszej tabeli czyli Produkty oraz tabeli podwymiarowej utworzyć nową tabele Produkty, w której już nie znajdywały się informacje z tabeli podwymiarowej oraz dodano do identyfikator. Na podstawie tego identyfikatora można połączyć naszą tabele Produkty z drugą. Po wszystkim jesteśmy w stanie połączyć tabele ze sobą w relacje.

Ostatnim etapem było zaciągnięcie danych do hurtowni danych. Obyło się to bez problemu ze względu na wcześniejszą transformację, która przygotowała już nam gotowe tabele do importu. Po zaciągnięciu tabel do hurtowni zostały na nowo połączone w relacje.

Procedury kontrolne przy tworzeniu DW za pomocą procesów ETL obejmowałyby przede wszystkim kontrole agregacji danych, sprawdzania ich poprawności i wykonalności po każdym przeprowadzonym kroku. Do przeprowadzenia tych procedur konieczne byłoby poświęcenie dużej ilości czasu aby sprawdzić przy każdym kolejnym etapie czy wynik uzyskany pokrywa się z wynikiem oczekiwanym.

Przy tworzeniu tabeli CZAS nasz sklep zdecydował przede wszystkim o funkcjonalności tej tabeli. Posiadała ona wszelkie potrzebne do analiz i pytań zarządczych atrybuty takie jak kwartał, tydzień roku, miesiąc roku etc. Tabela ta pomogła również przy organizacji głównej encji Sprzedaż, jako że z pomocą IDCzas, czyli klucza głównego w tabeli CZAS, wszelkie informacje wyżej wymienione znajdowały się w encji Sprzedaż w jednym polu IDCzas, któremu odpowiadają odpowiednie atrybuty w tabeli wymiaru CZAS. Tabela została utworzona przy pomocy programu MS Excel. Za pomocą jego wewnętrznych funkcji utworzyliśmy konkretne pola w tabeli Czas m. in. Kwartał, Miesiąc, Rok. Dla każdej daty w Excelu przypisaliśmy identyfikator, czyli zwykłą numerację od 1. Każda data posiadała jeden unikalny identyfikator.

Po przeprowadzeniu wszystkich działań otrzymaliśmy naszą hurtownie danych ze strukturą płatka śniegu z poprawnie zagregowanymi danymi. Hurtownia znajduję się w pliku o nazwie: HUDA\_S1IZ20\_Pn1315\_19-Z1\_DW

1. Proszę pamiętać, że rysunek lub tabela nie może być umieszczony na początku lub na końcu punktu (przed i po rysunku musi występować tekst!) [↑](#footnote-ref-1)
2. Nazwa encji jest rzeczownikiem w liczbie pojedynczej, np. KLIENT [↑](#footnote-ref-2)
3. Kategoria (Mandatoryjność), tzn. czy atrybut jest kluczem głównym (**Id,**), obowiązkowy (**Obw.**) czy opcjonalny (**Opc.**) [↑](#footnote-ref-3)
4. Typ (**ID** / **musi** / **może**) i *nazwa związku* (**czasownik** łączący encje) [↑](#footnote-ref-4)
5. Kardynalność: **jeden (1)** / **wiele (\*)** (z iloma wystąpieniami encji może/musi być w związku) [↑](#footnote-ref-5)
6. Nazwa tabeli jest rzeczownikiem w liczbie mnogiej, np. KLIENCI [↑](#footnote-ref-6)
7. Kategoria (Mandatoryjność), tzn. czy pole jest kluczem głównym (**Id,**), kluczem obcym (**IdO**), obowiązkowe (**Obw.**) czy opcjonalne (**Opc.**) [↑](#footnote-ref-7)
8. Nazwa encji jest rzeczownikiem w liczbie pojedynczej, dla odróżnienia od encji z tabeli 1 może się zaczynać od litery „H” np. HKLIENT [↑](#footnote-ref-8)
9. Kategoria (Mandatoryjność), tzn. czy atrybut jest kluczem głównym (**Id,**), obowiązkowy (**Obw.**) czy opcjonalny (**Opc.**) [↑](#footnote-ref-9)
10. Typ (**ID** / **musi** / **może**) i *nazwa związku* (**czasownik** łączący encje) [↑](#footnote-ref-10)
11. Kardynalność: **jeden (1)** / **wiele (\*)** (z iloma wystąpieniami encji może/musi być w związku) [↑](#footnote-ref-11)
12. Nazwa tabeli jest rzeczownikiem w liczbie mnogiej, np. HKLIENCI [↑](#footnote-ref-12)
13. Kategoria (Mandatoryjność), tzn. czy pole jest kluczem głównym (**Id,**), kluczem obcym (**IdO**), obowiązkowe (**Obw.**) czy opcjonalne (**Opc.**) [↑](#footnote-ref-13)