

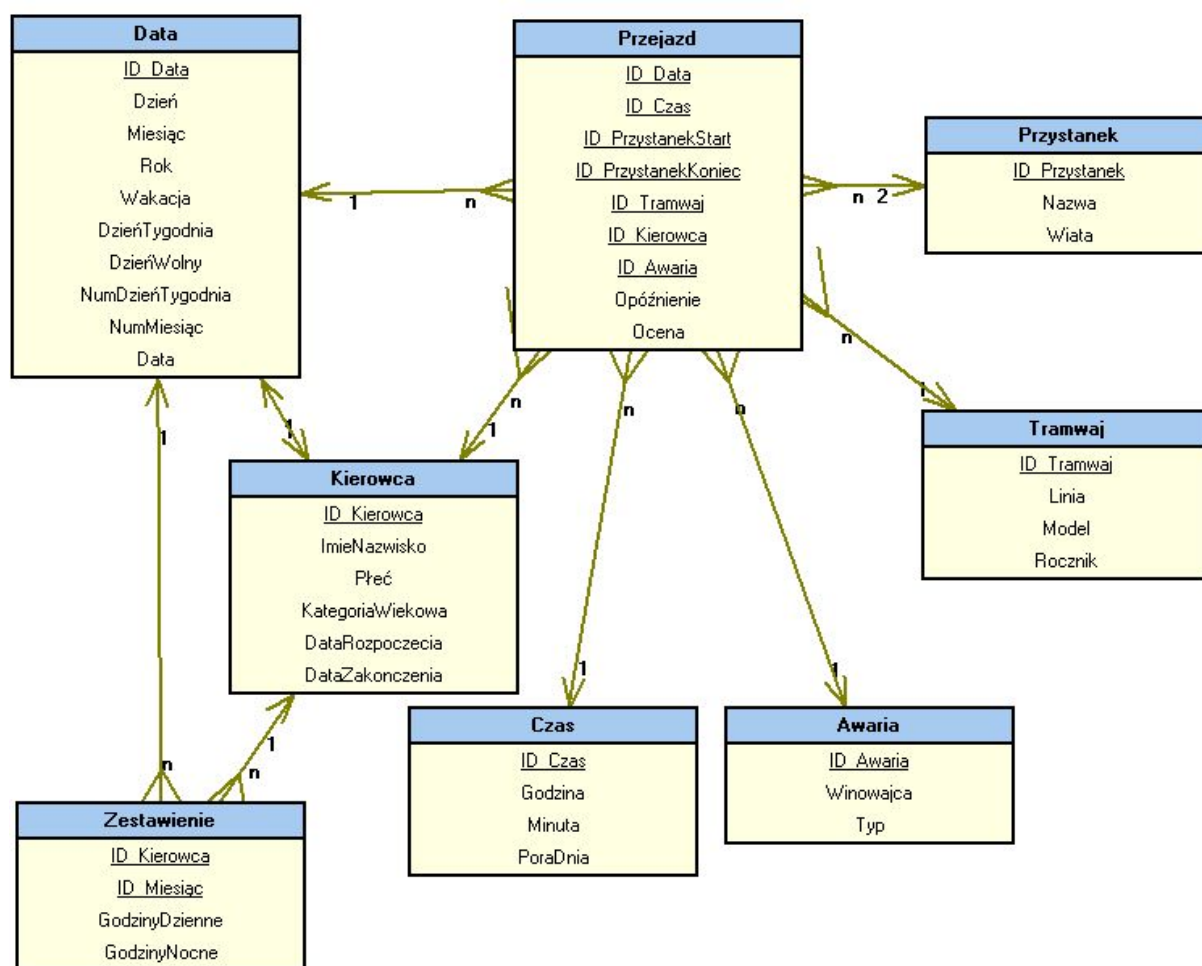
Projekt hurtowni danych

Przyjazdy tramwajów

Business process

Hurtownia danych zaprojektowana jest dla procesu przejazdu tramwajów. Został on opisany w odrębnym dokumencie "Specyfikacja procesów".

Relational Database schema



Nazwa tabeli	Atrybut	Typ atrybutu	Opis
Przejazd (fakt)	wiersz przedstawia jeden przejazd od przystanku do przystanku		
	ID_Data	Numeryczny	FK Data, Data przejazdu
	ID_Czas	Numeryczny	FK Czas, Czas przejazdu
	ID_PrzystanekStart	Numeryczny	FK Przystanek, Przystanek startowy
	ID_PrzystanekKoniec	Numeryczny	FK Przystanek, Przystanek końcowy
	ID_Awaria	Numeryczny	FK Awaria, Awaria podczas przejazdu
	ID_Tramwaj	Numeryczny	FK Tramwaj, tramwaj wykonujący przejazd
	ID_Kierowca	Numeryczny	FK Kierowca, Motorniczy prowadzący tramwaj
	Opóźnienie	Numeryczny	Opóźnienie wyrażane w minutach
	Ocena	Numeryczny	Ocena wyrażana jako liczba całkowita 0-10
Data (wymiar)	jeden wiersz przedstawia jeden dzień		
	ID_Data	Numeryczny	PK
	Dzień	2 cyfry	Dzień miesiąca 1-31
	Miesiąc	Varchar(11)	Miesiąc. Dozwolone wartości: <ul style="list-style-type: none"> - Styczeń, - Luty, - Marzec, - Kwiecień, - Maj, - Czerwiec, - Lipiec, - Sierpień, - Wrzesień, - Październik, - Listopad, - Grudzień
	Rok	4 cyfry	Rok
	Wakacje	Varchar(3)	Czy podana data to wakacje. Dozwolone

			wartości: - tak, - nie
	DzieńTygodnia	Varchar(12)	Dzień tygodnia. Dozwolone wartości: - Poniedziałek, - Wtorek, - Środa, - Czwartek, - Piątek, - Sobota, - Niedziela
	DzieńWolny	Varchar(3)	Czy podana data to dzień wolny od pracy. Dozwolone wartości: - tak, - nie
	NumDzienTygodnia	Numeryczny	wartość liczbowa opisująca dzień tygodnia (1-7)
	NumMiesiąc	Numeryczny	wartość liczbowa opisująca miesiąc (1-12)
Czas (wymiar)	jeden wiersz przedstawia jedną godzinę		
	ID_Czas	Numeryczny	PK
	Godzina	Numeryczny	Godzina 0-23
	Minuta	Numeryczny	Minuta 0-59
	Pora_dnia	Varchar(10)	Pora dnia. Dozwolone wartości: - noc, - poranek, - południe, - popołudnie, - wieczór
Przystanek (wymiar)	jeden wiersz przedstawia jeden przystanek		
	ID_Przystanek	Numeryczny	PK
	Nazwa	Varchar(40)	Nazwa przystanku
	Wiata	Varchar(3)	Czy przystanek posiada wiatę. Dozwolone wartości: - tak, - nie

Awaria (wymiar)	jeden wiersz przedstawia pojedynczą awarię		
	ID_Awaria	Numeryczny	PK
	Typ	Varchar(25)	Typ awarii. Dozwolone wartości: <ul style="list-style-type: none"> - wykolejenie, - kolizja, - problem z silnikiem, - drobna usterka
	Winowajca	Varchar(16)	Osoba odpowiedzialna za awarię. Dozwolone wartości: <ul style="list-style-type: none"> - motorniczy, - kierowca osobowy, - kierowca inny, - pieszy, - rowerzysta, - nieznane
Tramwaj (wymiar)	jeden wiersz przedstawia jeden tramwaj		
	ID_Tramwaj	Numeryczny	PK
	Model	Varchar(80)	Model tramwaju
	Rocznik	Numeryczny	Rok produkcji tramwaju
	Linia	Varchar(8)	Numer linii tramwaju zapisana w formacie "Linia XX" , gdzie XX to jej numer
Kierowca (wymiar)	jeden wiersz przedstawia jednego kierowcę		
	ID_Kierowca	Numeryczny	PK
	ImieNazwisko	Varchar(80)	Imię i nazwisko kierowcy
	Płeć	Varchar(9)	Płeć kierowcy. Dozwolone wartości: mężczyzna, kobieta
	KategoriaWiekowa	Varchar(20)	Kategoria wiekowa. Dozwolone wartości: pomiędzy 18 a 21, pomiędzy 22 a 29, pomiędzy 30 a 39, pomiędzy 40 a 50, ponad 50
	RokRozpoczecia	4 cyfry	Data rozpoczęcia

			bytowania w danej kategorii wiekowej
Zestawienie (fakt)	jeden wiersz przedstawia jedno zestawienie godzin dla danego pracownika w danym miesiącu		
	ID_Kierowca	Numeryczny	FK Kierowca, Kierowca
	ID_Miesiąc	Numeryczny	FK Data, Miesiąc
	GodzinyDzienne	Numeryczny	Liczba godzin przepracowanych w ciągu dnia
	GodzinyNocne	Numeryczny	Liczba godzin przepracowanych w ciągu nocy

Model wymiarowy

Definicja faktów

Fakt pierwszy: Przejazdy

Przejazd tramwaju pomiędzy dwoma przystankami konkretnego dnia o konkretnej godzinie. Tramwaj jest prowadzony przez określonego motorniczego, a przejazd ma określoną ocenę przez pasażerów, opóźnienie (wyliczane na podstawie czasu na rozkładzie i rzeczywistego momentu dotarcia na przystanek rozpoczynający przejazd) oraz informację o tym, czy wystąpiła awaria

Ziarnistość:

- określony tramwaj
- określony kierowca
- określona data
- określona godzina rozpoczęcia
- określony przystanek początkowy
- określony przystanek końcowy

Miary i funkcje agregujące:

- Suma opóźnień → SUM Opóźnienie
- Suma przejazdów z awariami → COUNT NON EMPTY (ID_Awaria)
- Średnie opóźnienie → AVG Opóźnienie
- Średnia ocena → AVG Ocena
- Liczba przejazdów → COUNT (*)
- Koszt opóźnienia → Opóźnienie * 10

Fakt drugi: Zestawienia

Zestawienie danego kierowcy w danym miesiącu, które posiada liczbę godzin przepracowanych w ciągu dnia oraz liczbę godzin przepracowanych w ciągu nocy.

Ziarnistość:

- określony kierowca
- określony miesiąc

Miary i funkcje agregujące:

- Średnia ilość godzin nocnych → AVG GodzinyNocne
- średnia ilość godzin dziennych → AVG GodzinyDzienne

Definicja wymiarów

Wymiar (atrybut)	Tabela (kolumna)	Typ
Data		Wymiar
ID_Data	Data.ID_Data	atrybut wymiaru
Dzień	Data.Dzień	atrybut wymiaru
Miesiąc	Data.Miesiąc	atrybut wymiaru
Rok	Data.Rok	atrybut wymiaru
Wakacje	Data.Wakacje	atrybut wymiaru
DzieńTygodnia	Data.DzieńTygodnia	atrybut wymiaru
DzieńWolny	Data.DzieńWolny	atrybut wymiaru
NumDzieńTygodnia	Data.NumDzieńTygodnia	atrybut wymiaru
NumMiesiąc	Data.NumMiesiąc	atrybut wymiaru
Czas		Wymiar
ID_Czas	Czas.ID_Czas	atrybut wymiaru
Godzina	Czas.Godzina	atrybut wymiaru
Minuta	Czas.Minuta	atrybut wymiaru
Pora_dnia	Czas.Pora_dnia	atrybut wymiaru
Przystanek		Wymiar
ID_Przystanek	Przystanek.ID_Przystanek	atrybut wymiaru
Nazwa	Przystanek.Nazwa	atrybut wymiaru
Wiata	Przystanek.Wiata	atrybut wymiaru
Awaria		Wymiar
ID_Awaria	Awaria.ID_Awaria	atrybut wymiaru
Typ	Awaria.Typ	atrybut wymiaru
Winowajca	Awaria.Winowajca	atrybut wymiaru
Tramwaj		Wymiar
ID_Tramwaj	Tramwaj.ID_Tramwaj	atrybut wymiaru
Model	+Model.Model ++Model.Rocznik	wymiar hierarchiczny

Linia	Tramwaj.Linia	atrybut wymiaru
Kierowca		Wymiar SCD
ID_Kierowca	Kierowca.ID_Kierowca	atrybut wymiaru
ImięNazwisko	Kierowca.ImięNazwisko	atrybut wymiaru
Płeć	Kierowca.Płeć	atrybut wymiaru
KategoriaWiekowa	Kierowca.KategoriaWiekowa	atrybut wymiaru
RokRozpoczecia	Kierowca.RokRozpoczecia	atrybut wymiaru

Sprawdzenie wykonalności zapytań

Porównaj całkowite opóźnienia tramwajów z tego miesiąca i poprzedniego.	
Miara	Suma opóźnień
Wymiar	Data (atrybut miesiąc)
Jaka linia miała największe sumaryczne opóźnienie w tym miesiącu w porównaniu do miesiąca poprzedniego?	
Miara	Suma opóźnień
Wymiar	Data (atrybut miesiąc)
Wymiar	Tramwaj (atrybut linia)
Na jakich przystankach są największe sumaryczne opóźnienia w tym i poprzednim miesiącu?	
Miara	Suma opóźnień
Wymiar	Data (atrybut miesiąc)
Wymiar	Przystanek start (atrybut nazwa)
Porównaj liczbę przepracowanych godzin dziennych i nocnych kierowców, którzy prowadzili sumarycznie najbardziej opóźnione tramwaje w tym i poprzednim miesiącu.	
Miara	Suma opóźnień
Wymiar	Data (atrybut miesiąc)
Porównaj procent opóźnień w tym i poprzednim miesiącu przejazdów, które były kierowane przez kierowców męskich i żeńskich.	
Miara	Suma opóźnień
Wymiar	Data (atrybut miesiąc)
Wymiar	Kierowca (atrybut płeć)
Porównaj ilość awarii w każdym dniu tygodnia z obecnego i poprzedniego miesiąca.	
Miara	Suma przejazdów z awarią
Wymiar	Data (atrybut miesiąc, dzień tygodnia)
O której godzinie wystąpiło najwięcej awarii w miesiącu analizowanym i poprzednim?	

<i>Miara</i>	Suma przejazdów z awarią
<i>Wymiar</i>	Data (atrybut miesiąc)
<i>Wymiar</i>	Czas (atrybut godzina)
Jakie były średnie oceny najbardziej awaryjnych tramwajów z tego i poprzedniego miesiąca (nie wiem czy można brać dwie miary, ale imo można xd)	
<i>Miara</i>	Suma przejazdów z awarią
<i>Miara</i>	Średnia ocena
<i>Wymiar</i>	Data (atrybut miesiąc)
Czy dane modele tramwajów miały więcej awarii od innych biorąc pod uwagę aktualny i poprzedni miesiąc?	
<i>Miara</i>	Suma przejazdów z awarią
<i>Wymiar</i>	Data (atrybut miesiąc)
<i>Wymiar</i>	Tramwaj (atrybut model)
Ile awarii w tym i poprzednim miesiącu było spowodowanych przez kierowców aut osobowych ?	
<i>Miara</i>	Suma przejazdów z awarią
<i>Wymiar</i>	Awaria (atrybut winowajca)
<i>Wymiar</i>	Data (atrybut miesiąc)
Przy jakim przystanku dokładnie było najwięcej awarii w tym i poprzednim miesiącu (chodzi o przystanek początkowy) ?	
<i>Miara</i>	Suma przejazdów z awarią
<i>Wymiar</i>	Data (atrybut miesiąc)
<i>Wymiar</i>	Przystanek (atrybut nazwa)
Ile było awarii tramwajów w dni wolne od pracy porównując ten miesiąc i poprzedni?	
<i>Miara</i>	Suma przejazdów z awarią
<i>Wymiar</i>	Data (atrybut miesiąc, dzień wolny od pracy)

Sprawdzenie czy w źródłach są wszystkie dane do hurtowni

Nazwa tabeli	Atrybut	Źródło
Przejazd	ID_Data	Data przejazdu. Klucz obcy z wymiaru Data. Pozyskano z bazy danych o przejazdach, tabela przejazdy, kolumna planowany czas startu..
	ID_Czas	Moment przejazdu. Klucz obcy z wymiaru Czas. Pozyskano z bazy danych o przejazdach, tabela przejazdy, kolumna planowany czas startu..
	ID_Przystanek_start	Przystanek początkowy. Klucz obcy z wymiaru Przystanek. Pozyskano z bazy danych o przejazdach, tabela przejazdy, kolumna przystanek początkowy.
	ID_Przystanek_koniec	Przystanek końcowy. Klucz obcy z wymiaru Przystanek. Pozyskano z bazy danych o przejazdach, tabela przejazdy, kolumna przystanek końcowy
	ID_Awaria	Awaria. Klucz obcy z wymiaru Awaria. Pozyskano z bazy danych o przejazdach. tabela przejazdy, kolumna awaria
	ID_Tramwaj	Tramwaj. Klucz obcy z wymiaru Tramwaj. Pozyskano z bazy danych o przejazdach. tabela przejazdy, kolumna tramwaj
	ID_Kierowca	Kierowca. Klucz obcy z wymiaru Kierowca. Pozyskano z bazy danych o przejazdach. tabela przejazdy, kolumna kierowca
	Opóźnienie	Opóźnienie przejazdu. Pozyskano z bazy danych o przejazdach. tabela przejazdy, kolumna planowany czas startu i końca oraz rzeczywisty czas startu i końca
	Ocena	Ocena przejazdu. Nowo pozyskiwane dane.
Data	Wszystkie informacje są generowane na podstawie zwykłego kalendarza.	
	Dzień_wolny_od_pracy	Informacja czy dzień jest świętem lub jest wolny od pracy. Nowo pozyskana informacja.
Czas	Wszystkie informacje są generowane na podstawie zwykłego kalendarza	
Przystanek	ID_Przystanek	ID Przystanek. Klucz surogatowy generowany przez bazę danych.

	Nazwa	Nazwa przystanku. Pozyskano z bazy danych o przejazdach. tabela przystanki, kolumna nazwa
	Wiata	Informacja o wiacie na przystanku. Pozyskano z bazy danych o przejazdach. tabela przystanki, kolumna wiata.
Awaria	ID_Awaria	ID Awaria. Klucz surogatowy generowany przez bazę danych.
	Typ	Typ awarii pozyskano z bazy danych o przejazdach. tabela awaria, kolumna typ awarii
	Winowajca	Winowajca. Nowo pozyskane dane.
Tramwaj	ID_Tramwaj	ID Tramwaj. Klucz surogatowy generowany przez bazę danych.
	Model	Model. Pozyskano z bazy danych o przejazdach. tabela tramwaj, kolumna model
	Rocznik	Rocznik. Pozyskano z bazy danych o przejazdach. tabela tramwaj, kolumna rocznik
	Linia	Linia. Pozyskano z bazy danych o przejazdach. tabela tramwaj, kolumna linia
Kierowca	ID_Kierowca	ID_Kierowca Klucz surogatowy generowany przez bazę danych.
	ImieNazwisko	Imię i nazwisko kierowcy pozyskiwany z arkusza o kierowcach (zakładka kierowcy)
	Płeć	Płeć kierowcy wzięta z arkusza kalkulacyjnego o kierowcach (zakładka kierowcy)
	RokRozpoczecia	Rok rozpoczęcia danej kategorii wiekowej, pozyskano na podstawie daty urodzenia z arkusza kalkulacyjnego o kierowcach (zakładka kierowcy)
	Kategoria_wiekowa	Kategoria wiekowa. Dozwolone wartości: pomiędzy 18 a 21, pomiędzy 22 a 29, pomiędzy 30 a 39, pomiędzy 40 a 50, ponad 50
Zestawienie	ID_Kierowca	Kierowca. Klucz obcy z wymiaru Kierowca. Pozyskano z bazy danych o przejazdach. tabela przejazdy, kolumna kierowca.
	ID_Miesiąc	Miesiąc. Klucz obcy w wymiaru Data.
	Godziny_dzienne	Godziny przepracowane w ciągu dnia generowane na podstawie arkusza kalkulacyjnego o kierowcach (zakładka godziny

		dzienne)
	Godziny_nocne	Godziny przepracowane w ciągu nocy generowane na podstawie arkusza kalkulacyjnego o kierowcach (zakładka godziny nocne)