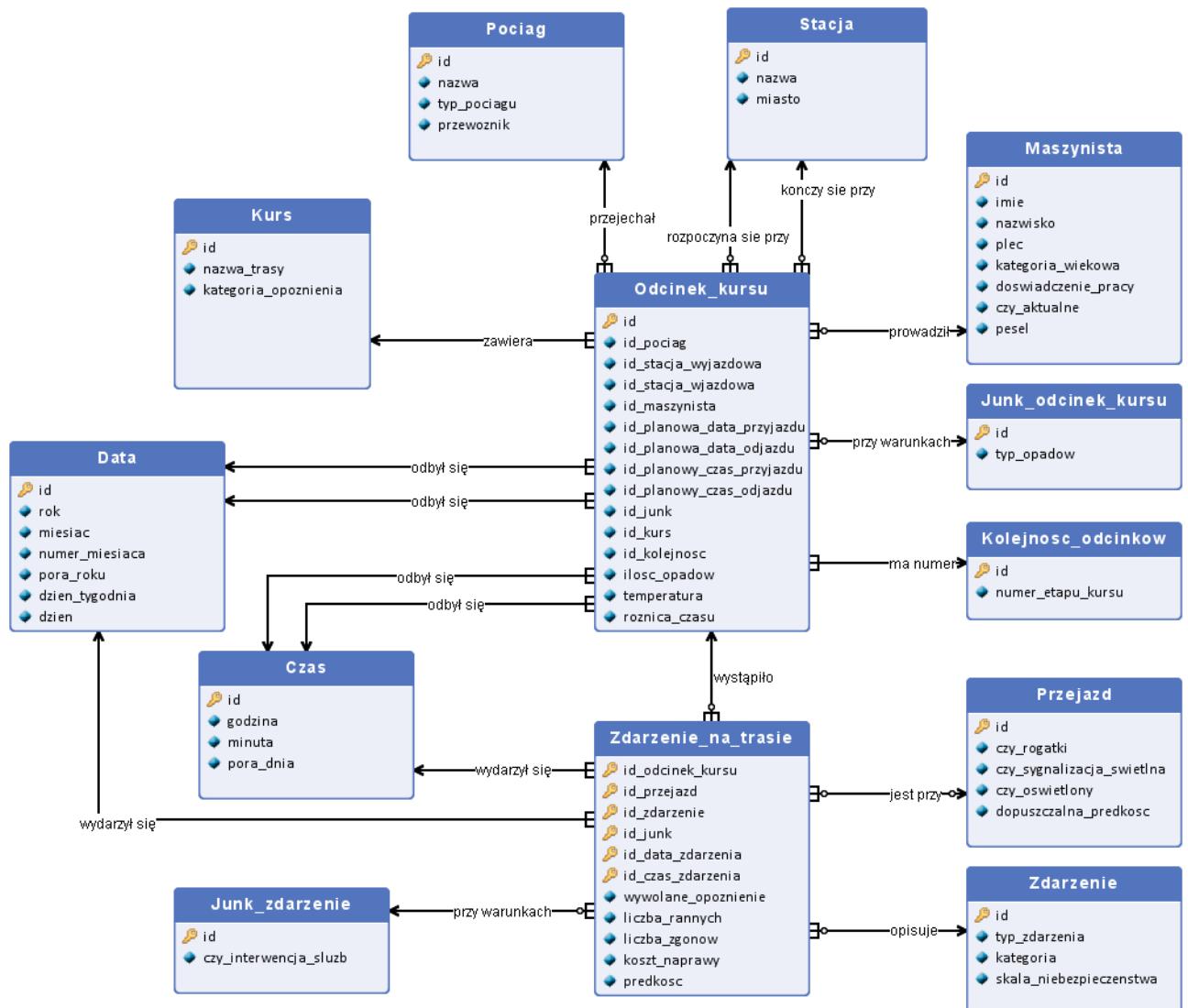


Przejazdy pociągów – Hurtownia danych

Proces biznesowy

Hurtownia danych jest przygotowana dla procesu przejazdu pociągów opisanego w pliku ProcessSpecification.pdf.

Schemat relacyjnej bazy danych



Opis zbioru encji

Odcinek_kursu			
Tablica faktów. Opisuje odcinek kursu, od jednego do drugiego przystanku.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int	PK - klucz główny surogatowy, żeby pozwolić na połączenie z zdarzeniami na trasie
id_pociag	Nie	int	id jadącego pociągu
id_stacja_wyjazdowa	Nie	int	id stacji z której pociąg odjeżdża
id_stacja_wjazdowa	Nie	int	id stacji do której pociąg dojeżdża
id_maszynista	Nie	int	id prowadzącego maszynisty
id_planowa_data_przyjazdu	Nie	int	data o której pociąg miał planowo przyjechać na przystanek
id_planowa_data_odjazdu	Nie	int	data o której pociąg miał planowo ruszyć z przystanku
id_planowy_czas_przyjazdu	Nie	int	czas o której pociąg miał planowo przyjechać na przystanek
id_planowy_czas_odjazdu	Nie	int	czas o której pociąg miał planowo ruszyć z przystanku
id_junk	Nie	int	id junk
id_kurs	Nie	int	id kursu, którego fragmentem jest odcinek
id_kolejnosc	Nie	int	numer, który jest to etap kursu
ilosc_opadow	Nie	int	liczba opadów w [mm/h] w trakcie tego przejazdu pociągu
temperatura	Nie	double (jedno miejsce po przecinku)	liczba stopni Celsjusza podczas pomiaru, wykonanego w trakcie przejazdu
roznica_czasu	Nie	int	opóźnienie w minutach w porównaniu z planowym przyjazdem, lub przyspieszenie jeśli wartość jest ujemna

Zdarzenie_na_trasie			
Tablica faktów. Zawiera incydenty, które miały miejsce w trakcie jazdy pociągu.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id_odcinek_kursu	Tak	int	id odcinka na którym wystąpiło zdarzenie
id_przejazd	Tak	int	id przejazdu kolejowego, przy którym nastąpiło zdarzenie (jeśli nie wystąpiło przy przejeździe to brak)
id_zdarzenie	Tak	int	id zdarzenia, które zaszło
id_junk	Tak	int	id junk
id_data_zdarzenia	Tak	int	data o której wystąpiło zdarzenie
id_czas_zdarzenia	Tak	int	data, kiedy wystąpiło zdarzenie
wywolane_opoznienie	Nie	int	minuty opóźnienia wywołanego zdarzeniem
liczba_rannych	Nie	int	liczba rannych osób
liczba_zgonow	Nie	int	liczba zmarłych osób
koszt_naprawy	Nie	double (do dwóch miejsc po przecinku)	szacowana liczba złotych na naprawę skutków zdarzenia
predkosc	Nie	int	średnia prędkość pociągu na chwilę przed zdarzeniem [km/h]

Pociag			
Tablica wymiaru. Jeżdżące po terenie Polski pociągi różnych operatorów.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
nazwa	Nie	varchar (20)	nazwa pociągu np. ICC4302, EIP12345
typ_pociagu	Nie	varchar(30)	typ pociągu np: passenger, cargo
przewoznik	Nie	varchar(40)	nazwa przewoźnika np: PKP Intercity, POLREGIO, PKP Cargo lub DB Cargo Polska

Maszynista			
Tablica wymiaru. Zbiór maszynistów prowadzących różne pociągi w trakcie kursów. (SCD 2)			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
imie	Nie	varchar(30)	imię maszynisty
nazwisko	Nie	varchar(30)	nazwisko maszynisty
plec	Nie	varchar(10)	płeć: man lub woman
kategoria_wiekowa	Nie	varchar(30)	kategoria wiekowa do której wpada maszynista, możliwe wartości: od 20 do 30, od 30 do 40,.... od 70 do 80
doswiadczenie_pracy	Nie	varchar(30)	przedziały lat doświadczenia, możliwe wartości: od 0 do 3, od 3 do 5, od 5 do 10, od 10 do 20, wiecej niż 20
czy_aktualne	Nie	boolean	1 jeśli informacja jest obecna, 0 jeśli przestarzała
pesel	Nie	varchar(11)	klucz biznesowy, pozwalający na przeszukiwanie poprzednich danych o maszyniste

Stacja			
Tablica wymiaru. Poszczególne stacje na terenie Polski przez które przejeżdżają pociągi.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
nazwa	Nie	varchar (40)	nazwa stacji np: Gdańsk Wrzeszcz, Gdańsk Główny
miasto	Nie	varchar(40)	miasto w którym znajduje się stacja np: Gdańsk

Kurs			
Tablica wymiaru. Pozwala na grupowanie fragmentów przejazdu w cały kurs pociągu.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
nazwa_trasy	Nie	varchar(40)	nazwa trasy kursu np: Gedania, Sudety
kategoria_opoznienia	Nie	varchar(40)	kategorie punktualnosci pociągu po ukończeniu całego kursu, możliwe wartości: przyspieszenie, opoznienie 1-5, opoznienie 5- 20,opoznienie 20-60, opoznienie >60, nieukonczenie

Przejazd			
Tablica wymiaru. Lokalizacja przejazdu kolejowo-samochodowego, przy którym wystąpił wypadek, bądź inne zdarzenie. Zdarzenie nie musi być powiązane z przejazdem, gdy np. pociąg z jakiś powodów nie mógł ruszyć.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
czy_rogatki	Nie	boolean	czy przejazd ma zamontowane rogateki
czy_sygnalizacja_swietlna	Nie	boolean	czy przejazd ma zamontowaną sygnalizację świetlną
czy_oswietlony	Nie	boolean	czy przejazd jest oswietlony
dopuszczalna_predkosc	Nie	varchar (30)	prędkość dopuszczalna dla pojazdów przy tym przejeździe kolejowym, możliwe wartości: ponizej 20, ponizej 30, ponizej 40, ... ponizej 80

Zdarzenie			
Tablica wymiaru. Zbiór możliwych zdarzeń podczas jazdy pociągiem.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
typ_zdarzenia	Nie	varchar (30)	Ogólna klasyfikacja zdarzenia: wypadek, incydent, awaria, zdarzenie techniczne
kategoria	Nie	varchar (40)	Dokładny typ zdarzenia: potrącenie pieszego, wykolejenie, zderzenie z innym pociągiem, wyłamani rogatek, przerwa w zasilaniu...
skala_niebezpieczenstwa	Nie	varchar(40)	skala niebezpieczenia, możliwe wartości: niskie,srednie,wysokie, bardzo_wysokie

Data			
Tablica wymiaru. Pojedynczy wiersz przedstawia jedną jednostkę czasu z dokładnością do dnia.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
rok	Nie	4 cyfry	liczba reprezentująca rok - 4 cyfry
miesiac	Nie	varchar(20)	możliwe nazwy miesiąca: styczeń, luty, marzec...
numer_miesiąca	Nie	liczby 1- 12	numer pozwalający na sortowanie po miesiącu
pora_roku	Nie	varchar(20)	możliwe pory roku: zima, wiosna, lato, jesień
dzien_tygodnia	Nie	varchar (20)	możliwe dni tygodnia: poniedziałek, wtorek, środa...
dzien	Nie	liczby 1 - 31	liczba reprezentująca dzień miesiąca 1-31

Czas			
Tablica wymiaru. Pojedynczy wiersz przedstawia jedną jednostkę czasu z dokładnością do minuty.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
godzina	Nie	liczby 0 - 23	Godzina 0-23
minuta	Nie	liczby 0 - 59	Minuta 0-59
pora_dnia	Nie	varchar(10)	możliwe wartości: rano, południe, wieczór, noc

Junk_odecinek_kursu			
Tablica wymiaru. Wszystkie możliwe wartości przyjmowane przez typ_opadow.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
typ_opadow	Nie	varchar(20)	deszcz, śnieg lub grad lub brak

Junk_zdarzenie			
Tablica wymiaru. Wszystkie możliwe wartości przyjmowane przez czy_interwencja_sluzb.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
czy_interwencja_sluzb	Nie	bool	czy musiały interweniować służby ratownicze w trakcie zdarzenia

Kolejnosc_odcinkow			
Tablica wymiaru. Zawiera numery odcinków, które mówią o kolejności mijanych stacji w trakcie kursu.			
Nazwa	Klucz główny	Typ/Dziedzina	Opis
id	Tak	int autoinkrement	PK
numer_etapu_kursu	Nie	kolejne liczby 1,2,3...	kolejne liczby reprezentujące, który to przystanek od początku

Model wymiarowy

Definicja faktów

Fakt 1: Przejazd odcinku kursu

Fakt przejazdu odcinka kursu reprezentuje przejazd pociągu pomiędzy dwiema stacjami w ramach konkretnego kursu, realizowanego przez określonego maszynistę, w danym dniu i czasie oraz określonych warunkach pogodowych.

Tablica faktów: Odcinek_kursu

Ziarnistość:

- określony pociąg
- określony kurs
- określony etap kursu
- określony maszynista
- określony dzień i czas przyjazdu
- określony dzień i czas odjazdu
- określona stacja początkowa
- określona stacja końcowa

Miary i funkcje agregujące:

- Liczba przejechanych odcinków – COUNT(*)
- Średnie, maksymalne, minimalne opóźnienie/przyspieszenie – AVG,MAX,MIN(roznica_czasu)
- Maksymalna, minimalna temperatura i opady – MAX,MIN(ilosc_opadow), MAX, MIN(temperatura)
- Liczba przejazdów z opadami COUNT(CASE WHEN ilosc_opadow > 0 THEN 1 ELSE 0 END)
- Średnia liczba odcinków na kurs – COUNT(*) / COUNT (DISTINCT id_kursu)

Fakt 2: Zdarzenia na trasie

Fakt opisuje zdarzenia, które wystąpiły w trakcie przejazdu pociągu między dwiema stacjami (np. wypadki, awarie). Każde zdarzenie jest powiązane z konkretnym odcinkiem kursu, przejazdem kolejowym, jeśli doszło na nim do tego zdarzenia i typem zdarzenia opisującym jego skale niebezpieczeństwa. Zawarte są dane o wywołanym przez zdarzenie opóźnieniu, szacowanej prędkości pociągu w trakcie zdarzenia, stratach w ludziach, rannych i szacowanych kosztach naprawy po zdarzeniu.

Ziarnistość:

- określony odcinek kursu
- określony przejazd (jeśli zdarzenie wystąpiło przy przejezdzie)
- określona kategoria i typ zdarzenia
- określona data i czas zdarzenia

Miary i funkcje agregujące:

- Suma kosztów napraw SUM(koszt_naprawy)
- Liczba zdarzeń COUNT(*)
- Średnie opóźnienie wywołane zdarzeniami AVG(wywolane_opoznienie)
- Suma ludzi rannych - SUM(liczba_rannych)

- Suma ludzi zmarlych - $\text{SUM}(\text{liczba_zgonow})$
- Suma poszkodowanych - $\text{SUM}(\text{liczba rannych}) + \text{SUM}(\text{liczba_zgonow})$
- Średnia prędkość pociągu w trakcie zdarzeń $\text{AVG}(\text{predkosc})$

Definicja wymiarów

Wymiary dla faktu 1: Przejazd odcinku kursu

Wymiar/ atrybut wymiaru	Tablica / kolumna	Typ
Pociag	Pociag	Wymiar
Nazwa pociągu	Pociag.nazwa	Atrybut wymiaru
Typ pociągu	Pociag.typ_pociagu	Atrybut wymiaru
Przewoźnik pociągu	Pociag.przewoznik	Atrybut wymiaru
Model pociągu	-Pociag.typ_pociagu --Pociag.operator ---Pociag.nazwa	Wymiar hierarchiczny
Kurs	Kurs	Wymiar
Nazwa trasy kursu	Kurs.nazwa_trasy	Atrybut wymiaru
Kategoria opoznienia kursu	Kurs.kategoria_opoznienia	Atrybut wymiaru
Stacja wyjazdowa	Stacja	Wymiar
Nazwa stacji wyjazdowej	Stacja.nazwa	Atrybut wymiaru
Miasto, w którym jest stacja wyjazdowa	Stacja.miasto	Atrybut wymiaru
Stacja wjazdowa	Stacja	Wymiar
Nazwa stacji wjazdowej	Stacja.nazwa	Atrybut wymiaru
Miasto, w którym jest stacja wjazdowa	Stacja.miasto	Atrybut wymiaru
Maszynista	Maszynista	Wymiar
Imię maszynisty	Maszynista.imie	Atrybut wymiaru
Nazwisko maszynisty	Maszynista.nazwisko	Atrybut wymiaru
Pesel maszynisty	Maszynista.pesel	Atrybut wymiaru
Płeć maszynisty	Maszynista.plec	Atrybut wymiaru
Kategoria wiekowa maszynisty	Maszynista.kategoria_wiekowa	Atrybut wymiaru
Doświadczenie maszynisty	Maszynista.doswiadczenie_pracy	Atrybut wymiaru
Junk_odcinek_kursu	Junk_odcinek_kursu	Wymiar
Typ opadów	Junk_odcinek_kursu.typ_opadow	Atrybut wymiaru
Data przyjazdu	Data	Wymiar
Rok przyjazdu	Data.rok	Atrybut wymiaru
Miesiąc przyjazdu	Data.miesiac	Atrybut wymiaru
Dzień przyjazdu	Data.dzien	Atrybut wymiaru

Pora roku przyjazdu	Data.pora_roku	Atrybut wymiaru
Dzień tygodnia odjazdu	Data.dzien_tygodnia	Atrybut wymiaru
Hierarchia daty przyjazdu	-Data.rok --Data.miesiac ---Data.dzien	Wymiar hierarchiczny
Data odjazdu	Data	Wymiar
Rok odjazdu	Data.rok	Atrybut wymiaru
Miesiąc odjazdu	Data.miesiac	Atrybut wymiaru
Dzień odjazdu	Data.dzien	Atrybut wymiaru
Pora roku odjazdu	Data.pora_roku	Atrybut wymiaru
Dzień tygodnia odjazdu	Data.dzien_tygodnia	Atrybut wymiaru
Hierarchia daty odjazdu	-Data.rok --Data.miesiac ---Data.dzien	Wymiar hierarchiczny
Czas przyjazdu	Czas	Wymiar
Godzina przyjazdu	Czas.godzina	Atrybut wymiaru
Minuty przyjazdu	Czas.minuta	Atrybut wymiaru
Pora roku przyjazdu	Czas.pora_roku	Atrybut wymiaru
Hierarchia czasu przyjazdu	-Czas.godzina --Czas.minuta	Wymiar hierarchiczny
Czas odjazdu	Czas	Wymiar
Godzina odjazdu	Czas.godzina	Atrybut wymiaru
Minuty odjazdu	Czas.minuta	Atrybut wymiaru
Pora roku odjazdu	Czas.pora_roku	Atrybut wymiaru
Hierarchia czasu odjazdu	-Czas.godzina --Czas.minuta	Wymiar hierarchiczny
Kolejność odcinków	Kolejnosc_odcinkow	Wymiar
Numer etapu odcinka	Kolejnosc_odcinkow.numer_etapu_odcinka	Atrybut wymiaru

Wymiary dla faktu 2: Zdarzenia na trasie

Wymiar/ atrybut wymiaru	Tablica / kolumna	Typ
Pociąg	Pociąg	Wymiar
Nazwa pociągu	Pociąg.nazwa	Atrybut wymiaru
Typ pociągu	Pociąg.typ_pociagu	Atrybut wymiaru
Przewoźnik pociągu	Pociąg.przewoznik	Atrybut wymiaru
Model pociągu	-Pociąg.typ_pociagu --Pociąg.operator ---Pociąg.nazwa	Wymiar hierarchiczny
Kurs	Kurs	Wymiar
Nazwa trasy kursu	Kurs.nazwa_trasy	Atrybut wymiaru

Kategoria opoznienia kursu	Kurs.kategoria_opoznienia	Atrybut wymiaru
Stacja wyjazdowa	Stacja	Wymiar
Nazwa stacji wyjazdowej	Stacja.nazwa	Atrybut wymiaru
Miasto, w którym jest stacja wyjazdowa	Stacja.miasto	Atrybut wymiaru
Stacja wjazdowa	Stacja	Wymiar
Nazwa stacji wjazdowej	Stacja.nazwa	Atrybut wymiaru
Miasto, w którym jest stacja wjazdowa	Stacja.miasto	Atrybut wymiaru
Maszynista	Maszynista	Wymiar
Imię maszynisty	Maszynista.imie	Atrybut wymiaru
Nazwisko maszynisty	Maszynista.nazwisko	Atrybut wymiaru
Pesel maszynisty	Maszynista.pesel	Atrybut wymiaru
Płeć maszynisty	Maszynista.plec	Atrybut wymiaru
Kategoria wiekowa maszynisty	Maszynista.kategoria_wiekowa	Atrybut wymiaru
Doświadczenie maszynisty	Maszynista.doswiadczenie_pracy	Atrybut wymiaru
Junk_odcinek_kursu	Junk_odcinek_kursu	Wymiar
Typ opadów	Junk_odcinek_kursu.typ_opadow	Atrybut wymiaru
Data przyjazdu	Data	Wymiar
Rok przyjazdu	Data.rok	Atrybut wymiaru
Miesiąc przyjazdu	Data.miesiac	Atrybut wymiaru
Dzień przyjazdu	Data.dzien	Atrybut wymiaru
Pora roku przyjazdu	Data.pora_roku	Atrybut wymiaru
Dzień tygodnia odjazdu	Data.dzien_tygodnia	Atrybut wymiaru
Hierarchia daty przyjazdu	-Data.rok --Data.miesiac ---Data.dzien	Wymiar hierarchiczny
Data odjazdu	Data	Wymiar
Rok odjazdu	Data.rok	Atrybut wymiaru
Miesiąc odjazdu	Data.miesiac	Atrybut wymiaru
Dzień odjazdu	Data.dzien	Atrybut wymiaru
Pora roku odjazdu	Data.pora_roku	Atrybut wymiaru
Dzień tygodnia odjazdu	Data.dzien_tygodnia	Atrybut wymiaru
Hierarchia daty odjazdu	-Data.rok --Data.miesiac ---Data.dzien	Wymiar hierarchiczny
Czas przyjazdu	Czas	Wymiar
Godzina przyjazdu	Czas.godzina	Atrybut wymiaru
Minuty przyjazdu	Czas.minuta	Atrybut wymiaru
Pora roku przyjazdu	Czas.pora_roku	Atrybut wymiaru
Hierarchia czasu przyjazdu	-Czas.godzina --Czas.minuta	Wymiar hierarchiczny

Czas odjazdu	Czas	Wymiar
Godzina odjazdu	Czas.godzina	Atrybut wymiaru
Minuty odjazdu	Czas.minuta	Atrybut wymiaru
Pora roku odjazdu	Czas.pora_roku	Atrybut wymiaru
Hierarchia czasu odjazdu	-Czas.godzina --Czas.minuta	Wymiar hierarchiczny
Przejazd	Przejazd	Wymiar
Czy są rogatki	Przejazd.godzina	Atrybut wymiaru
Czy jest sygnalizacja świetlna	Przejazd.czy_sygnalizacja_swietlna	Atrybut wymiaru
Czy jest oświetlony	Przejazd.czy_oswietlony	Atrybut wymiaru
Predkość dopuszczalna dla aut na tym przejeździe	Przejazd.dopuszczalna_predkosc	Atrybut wymiaru
Zdarzenie	Zdarzenie	Wymiar
Typ zdarzenia	Zdarzenie.typ_zdarzenia	Atrybut wymiaru
Kategoria zdarzenia	Zdarzenie.kategoria	Atrybut wymiaru
Skala niebezpieczenstwa zdarzenia	Zdarzenie.skala_niebezpieczenstwa	Atrybut wymiaru
Hierarchia ustalenia kategorii zdarzenia	-Zdarzenie.typ_zdarzenia --Zdarzenie.kategoria	Wymiar hierarchiczny
Junk_zdarzenie	Junk_zdarzenie	Wymiar
Czy interweniowały służby przy zdarzeniu	Junk_zdarzenie.czy_interwencja_sluzb	Atrybut wymiaru
Kolejnosc odcinkow	Kolejnosc_odcinkow	Wymiar
Numer etapu odcinka	Kolejnosc_odcinkow.numer_etapu_odcinka	Atrybut wymiaru
Data zdarzenia	Data	Wymiar
Rok zdarzenia	Data.rok	Atrybut wymiaru
Miesiąc zdarzenia	Data.miesiac	Atrybut wymiaru
Dzień zdarzenia	Data.dzien	Atrybut wymiaru
Pora roku zdarzenia	Data.pora_roku	Atrybut wymiaru
Dzień tygodnia zdarzenia	Data.dzien_tygodnia	Atrybut wymiaru
Hierarchia daty zdarzenia	-Data.rok --Data.miesiac ---Data.dzien	Wymiar hierarchiczny
Czas zdarzenia	Czas	Wymiar
Godzina zdarzenia	Czas.godzina	Atrybut wymiaru
Minuty zdarzenia	Czas.minuta	Atrybut wymiaru
Pora roku zdarzenia	Czas.pora_roku	Atrybut wymiaru
Hierarchia czasu zdarzenia	-Czas.godzina --Czas.minuta	Wymiar hierarchiczny
Odcinek_kursu	Odcinek_kursu	Wymiar

Zapytania

1. Porównaj średnie opóźnienie z kursów dla każdego przewoźnika pociągów w tym roku.
 - Miara: Średnie opóźnienie z odcinka kursu
 - Wymiar: Pociąg (atrybut wymiaru: Przewoźnik pociągu)
 - Wymiar: Data (atrybut wymiary: Rok)
2. Które pociągi średnio spóźniają się dłużej – te prowadzone przez doświadczone (> 5 lat w zawodzie) w zawodzie kobiety, czy prowadzone przez niedoświadczonych (< 3 lat w zawodzie) mężczyzn?
 - Miara: Średnie opóźnienie z odcinka kursu
 - Wymiar: Maszynista (atrybuty wymiaru: Doświadczenie maszynisty, Płeć maszynisty)
3. Podaj listę TOP 10 stacji, przy których średnie opóźnienie pociągu jest największe.
 - Miara: Średnie opóźnienia z odcinka kursu
 - Wymiar: Stacja wjazdowa (atrybut wymiaru: Nazwa stacji)
4. Ile było takich fragmentów kursów, które mimo zajścia incydentu na trasie zakończyły się o czasie?
 - Miara: Liczba odcinków kursu
 - Wymiar: Zdarzenie (atrybut wymiaru: typ_zdarzenia)
5. Dla ilu odcinków tras podczas których padał śnieg wystąpiło opóźnienie w tym miesiącu?
(dane z obu źródeł)
 - Miara: Liczba odcinków kursu
 - Wymiar: Junk odcinek kursu (atrybut wymiaru: typ_opadow)
 - Wymiar: Wymiar: Data (atrybut wymiary: Miesiąc)
6. Porównaj, liczbę wypadków na przejazdach kolejowych ze światłami i rogatkami, jak i bez nich.
 - Miara: Liczba zdarzeń
 - Wymiar: Zdarzenie (atrybut wymiaru: typ_zdarzenia)
 - Wymiar: Przejazd (atrybuty wymiary: czy_sygnalizacja_swietlna, czy_rogatki)
7. Którego przewoźnika pociągi wywołują incydenty o największych kosztach napraw?
 - Miara: Suma kosztów napraw
 - Wymiar: Pociąg (atrybut wymiaru: Przewoźnik pociągu)
8. Pomiędzy którymi stacjami doszło do zranienia i śmierci największej liczby osób?
 - Miara: Liczba poszkodowanych (ranni + zgony)
 - Wymiar: Stacja_wjazdowa (nazwa_stacji)
 - Wymiar: Stacja_wyjazdowa (nazwa_stacji)
9. Jakiej skali niebezpieczeństwa zdarzenia generują średnio najmniejsze dodatkowe opóźnienia?
 - Miara: Średnie wywołane opóźnienie
 - Wymiar: Zdarzenie (atrybut wymiaru: skala_niebezpieczeństwa)
10. Podczas, jak wielu wypadków z pojazdem padał deszcz i interweniowały wtedy służby?
(dane z obu źródeł)
 - Miara: Liczba zdarzeń
 - Wymiar: Junk odcinek kursu (atrybut wymiaru: typ_opadow)
 - Wymiar: Zdarzenie (atrybuty wymiaru: typ_zdarzenia, kategoria_zdarzenia)
 - Wymiar: Junk zdarzenie (atrybut wymiaru: czy_interwencja_sluzb)

Sprawdzenie czy w źródłach są wszystkie dane do hurtowni

Odcinek_kursu	
Tablica faktów. Opisuje odcinek kursu, od jednego do drugiego przystanku.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
id_pociag	Klucz obcy do tabeli wymiaru Pociag
id_stacja_wyjazdowa	Klucz obcy do tabeli wymiaru Stacja
id_stacja_wjazdowa	Klucz obcy do tabeli wymiaru Stacja
id_maszynista	Klucz obcy do tabeli wymiaru Maszynista
id_planowa_data_przyjazdu	Klucz obcy do tabeli wymiaru Data
id_planowa_data_odjazdu	Klucz obcy do tabeli wymiaru Data
id_planowy_czas_przyjazdu	Klucz obcy do tabeli wymiaru Czas
id_planowy_czas_odjazdu	Klucz obcy do tabeli wymiaru Czas
id_junk	Klucz obcy do tabeli wymiaru Junk_odecinek.
id_kurs	Klucz obcy do tabeli wymiaru Kurs
id_kolejnosc	Klucz obcy do tabeli wymiaru Kolejnosc odcinkow.
ilosc_opadow	liczba opadow w [mm/h] w trakcie tego przejazdu pociągu. Źródło danych: Z pliku csv, kolumna ilosc_opadow, ta sama reprezentacja.
temperatura	liczba stopni Celsjusza podczas pomiaru, wykonanego w trakcie przejazdu. Z pliku csv, kolumna ilosc_opadow, ta sama reprezentacja.
roznica_czasu	opóźnienie w minutach w porównaniu z planowym przyjazdem, lub przyspieszenie jeśli wartość jest ujemna. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.

Zdarzenie_na_trasie	
Tablica faktów. Zawiera incydenty, które miały miejsce w trakcie jazdy pociągu.	
Nazwa	Źródło
id_odecinek_kursu	Klucz obcy do tabeli wymiaru Odcinek_kursu
id_przejazd	Klucz obcy do tabeli wymiaru Przejazd
id_zdarzenie	Klucz obcy do tabeli wymiaru Czas.
id_junk	Klucz obcy do tabeli wymiaru Junk_zdarzenie.
id_data_zdarzenia	Klucz obcy do tabeli wymiaru Data
id_czas_zdarzenia	Klucz obcy do tabeli wymiaru Czas
wywolane_opoznienie	minuty opóźnienia wywołanego zdarzeniem. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
liczba_rannych	liczba rannych osób. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
liczba_zgonow	liczba zmarłych osób. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie

	i tym atrybutem.
koszt_naprawy	szacowana liczba złotych na naprawę skutków zdarzenia. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
predkosc	średnia prędkość pociągu na chwilę przed zdarzeniem [km/h]. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.

Pociag	
Tablica wymiaru. Jeżdżące po terenie Polski pociągi różnych operatorów.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
nazwa	nazwa pociągu np. ICC4302, EIP12345. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
typ_pociagu	typ pociągu np: passenger, cargo. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
przewoznik	nazwa przewoźnika np: PKP Intercity, POLREGIO, PKP Cargo lub DB Cargo Polska. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i atrybutem - operator.

Maszynista	
Tablica wymiaru. Zbiór maszynistów prowadzących różne pociągi w trakcie kursów. (SCD 2)	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
imie	imię maszynisty. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
nazwisko	nazwisko maszynisty. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
plec	płeć: man lub woman. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
kategoria_wiekowa	kategoria wiekowa do której wpada maszynista, możliwe wartości: od 20 do 30, od 30 do 40,..... od 70 do 80. Źródło danych: przekonwertowane wartości liczbowe na opisowe z bazy danych o tabeli o tej nazwie i tym atrybutem.

doswiadczenie_pracy	przedziały lat doświadczenia, możliwe wartości: od 0 do 3, od 3 do 5, od 5 do 10, od 10 do 20, wiecej niż 20. Źródło danych: przekonwertowane wartości liczbowe na opisowe z bazy danych o tabeli o tej nazwie i tym atrybutem.
czy_aktualne	1 jeśli informacja jest obecna, 0 jeśli przestarzała. Do reprezentacji SCD 2
pesel	klucz biznesowy, pozwalający na przeszukiwanie poprzednich danych o maszynie. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem. (Zostanie dodane ze względu na konieczność klucza biznesowego)

Stacja	
Tablica wymiaru. Poszczególne stacje na terenie Polski przez które przejeżdżają pociągi.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
nazwa	nazwa stacji np: Gdańsk Wrzeszcz, Gdańsk Główny. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
miasto	miasto w którym znajduje się stacja np: Gdańsk. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.

Kurs	
Tablica wymiaru. Pozwala na grupowanie fragmentów przejazdu w cały kurs pociągu.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
nazwa_trasy	nazwa trasy kursu np: Gedania, Sudety. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutem.
kategoria_opoznienia	kategorie punktualności pociągu po ukończeniu całego kursu, możliwe wartości: przyspieszenie, opoznienie 1-5, opoznienie 5-20, opoznienie 20-60, opoznienie >60, nieukończenie. Źródło danych: przekonwertowane wartości liczbowe na opisowe z bazy danych o tabeli o tej nazwie i tym atrybutem.

Przejazd	
Tablica wymiaru. Lokalizacja przejazdu kolejowo-samochodowego, przy którym wystąpił wypadek, bądź inne zdarzenie. Zdarzenie nie musi być powiązane z przejazdem, gdy np. pociąg z jakiś powodów nie mógł ruszyć.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
czy_rogatki	czy przejazd ma zamontowane rogatek. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutie.
czy_sygnalizacja_swietlna	czy przejazd ma zamontowaną sygnalizację świetlną. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutie.
czy_oswietlony	czy przejazd jest oświetlony. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutie.
dopuszczalna_predkosc	prędkość dopuszczalna dla pojazdów przy tym przejeździe kolejowym, możliwe wartości: ponizej 20, ponizej 30, ponizej 40, ... ponizej 80. Źródło danych: przekonwertowane wartości liczbowe na opisowe z bazy danych o tabeli o tej nazwie i tym atrybutie.

Zdarzenie	
Tablica wymiaru. Zbiór możliwych zdarzeń podczas jazdy pociągiem.	
Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
typ_zdarzenia	Ogólna klasyfikacja zdarzenia: wypadek, incydent, awaria, zdarzenie techniczne. Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutie.
kategoria	Dokładny typ zdarzenia: potrącenie pieszego, wykolejenie, zderzenie z innym pociągiem, wyłamani rogatek, przerwa w zasilaniu... Źródło danych: taka sama reprezentacja w bazie danych o tabeli o tej nazwie i tym atrybutie.
skala_niebezpieczenstwa	skala niebezpieczenstwa, możliwe wartości: niskie, średnie, wysokie, bardzo_wysokie. Przekonwertowane wartości z tej samej tabeli z bazy danych o tej samej nazwie atrybutu 1 – 10 na te opisowe.

Data

Tablica wymiaru. Pojedynczy wiersz przedstawia jedną jednostkę czasu z dokładnością do dnia.
Wszystkie dane tej tabeli są generowane krotka po krotce przed procesem ETL.

Czas

Tablica wymiaru. Pojedynczy wiersz przedstawia jedną jednostkę czasu z dokładnością do minuty.
Wszystkie dane tej tabeli są generowane krotka po krotce przed procesem ETL.

Junk_odcinek_kursu

Tablica wymiaru. Wszystkie możliwe wartości przyjmowane przez typ_opadow.

Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
typ_opadow	deszcz, snieg lub grad lub brak – są to wszystkie możliwe opcje w tym polu

Junk_zdarzenie

Tablica wymiaru. Wszystkie możliwe wartości przyjmowane przez czy_interwencja_sluzb.

Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
czy_interwencja_sluzb	Tak i nie - są to wszystkie możliwe opcje w tym polu

Kolejnosc_odcinkow

Tablica wymiaru. Zawiera numery odcinków, które mówią o kolejności mijanych stacji w trakcie kursu.

Nazwa	Źródło
id	Klucz surogatowy generowany przez bazę
numer_etapu_kursu	Mogą wygenerować przed procesem ETL kolejne liczby - 1,2,3....