

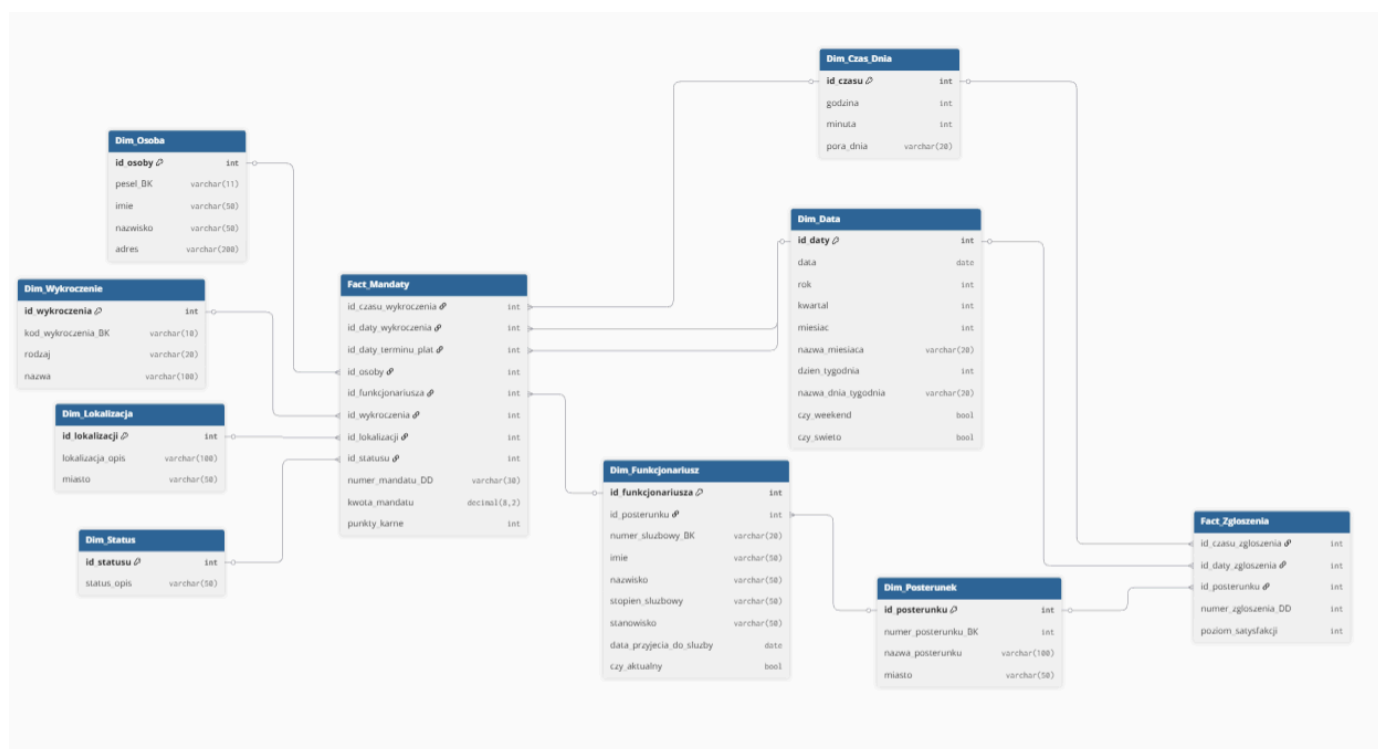
# Policja - data warehouse design

Kacper Kozak 198223, Marcin Mazuro 198019

Proces biznesowy

Hurtownia danych jest stworzona do procesu biznesowego wystawiania mandatów.

Schemat



## Opis

### MANDAT (TABELA FAKTÓW)

Pojedynczy wiersz opisuje jeden fakt wystawienia mandatu.

Atrybut	Typ	Opis
id_daty_wykroczenia	int	FK do Dim_Data. Data wystawienia mandatu.
id_czasu_wykroczenia	int	FK do Dim_Czas_Dnia. Godzina wystawienia mandatu.
id_daty_terminu_plat	int	FK do Dim_Data. Kiedy mandat powinien zostać zapłacony.
id_osoby	int	FK do Dim_Osoba. Osoba ukarana mandatem.
id_funkcjonariusza	int	FK do Dim_Funkcjonariusz. Funkcjonariusz wystawiający.
id_wykroczenia	int	FK do Dim_Wykroczenie. Powód wystawienia mandatu.
id_lokalizacji	int	FK do Dim_Lokalizacja. Miejsce wystawienia mandatu.

id_statusu	int	FK do Dim_Status. Status opłacenia mandatu.
numer_mandatu_DD	varchar(30)	Wymiar zdegenerowany. Oryginalny numer mandatu.
kwota_mandatu	decimal(8,2)	<b>Miara.</b> Wartość mandatu.
punkty_karne	int	<b>Miara.</b> Liczba przyznanych punktów karnych

## DATA (TABELA WYMIARU)

Pojedynczy wiersz opisuje jeden dzień w kalendarzu.

Atrybut	Typ	Opis
id_daty	int	PK (Klucz zastępczy, np. RRRRMMDD)
data	date	Pełna data, np. '2025-11-10'
rok	int	Rok
kwartal	int	Kwartał (1-4)
miesiac	int	Miesiąc (1-12)

nazwa_miesiaca	varchar(30)	Np. 'Listopad'
dzien_tygodnia	int	Dzień tygodnia (1-7)
nazwa_dzien_tygodnia	varchar(20)	Np. 'Wtorek'
czy_weekend	bool	Tak lub nie
czy_swieto	bool	Tak lub nie

## CZAS\_DNIA (TABELA WYMIARU)

Pojedynczy wiersz opisuje konkretną godzinę/minutę w ciągu doby.

Atrybut	Typ	Opis
id_czasu	int	PK
godzina	int	Godzina (0-23)
minuta	int	Minuta (0-59)
pora_dnia	varchar(20)	Np. 'Rano', 'Popołudnie', 'Noc'

## OSOBA (TABELA WYMIARU)

Pojedynczy wiersz opisuje jedną osobę ukaraną mandatem.

Atrybut	Typ	Opis
id_osoby	int	PK (Klucz zastępczy)
pesel_BK	varchar(11)	BK (Klucz biznesowy)
imie	varchar(50)	Imię
nazwisko	varchar(50)	Nazwisko
adres	varchar(200)	Adres zamieszkania

## FUNKCJONARIUSZ (TABELA WYMIARU)

Pojedynczy wiersz opisuje jednego funkcjonariusza.

Atrybut	Typ Danych	Opis
id_funkcjonariusza	int	Klucz Główny (PK)
id_posterunku	int	Powiązanie z wymiarem Dim_Posterunek. Klucz Obcy (FK)

numer_sluzbowy_BK	varchar(20)	Klucz Biznesowy (Business Key). Łączy wszystkie rekordy historyczne tego samego funkcjonariusza.
imie	varchar(50)	Imię funkcjonariusza.
nazwisko	varchar(50)	Nazwisko funkcjonariusza.
stopien_sluzbowy	varchar(50)	Stopień (Rank) służbowy.
stanowisko	varchar(50)	Zajmowane stanowisko (Position).
data_przyjecia_do_sluzby	date	Data rozpoczęcia służby (StartDate).
czy_aktualny	bool	Flaga (1 / 0) oznaczająca, czy dany rekord jest bieżącym stanem funkcjonariusza (SCD2)

## WYKROCZENIE (TABELA WYMIARU)

Pojedynczy wiersz opisuje jeden typ wykroczenia.

Atrybut	Typ	Opis
id_wykroczenia	int	PK (Klucz zastępczy)
kod_wykroczenia_BK	varchar(10)	BK (Klucz biznesowy)

rodzaj	varchar(20)	Rodzaj wykroczenia
nazwa	varchar(100)	Nazwa wykroczenia

## LOKALIZACJA (TABELA WYMIARU)

Pojedynczy wiersz opisuje lokalizację.

Atrybut	Typ	Opis
id_lokalizacji	int	PK (Klucz zastępczy)
lokalizacja_opis	varchar(100)	Opis tekstowy lokalizacji
miasto	varchar(50)	Nazwa miasta

## STATUS (TABELA WYMIARU)

Pojedynczy wiersz opisuje status płatności.

Atrybut	Typ	Opis
<u>id_statusu</u>	int	PK (Klucz zastępczy)
<u>status_opis</u>	varchar(50)	Np. 'Zapłacony', 'Niezapłacony', 'Anulowany'

## POSTERUNEK (TABELA WYMIARU)

Pojedynczy wiersz opisuje posterunek policji.

Atrybut	Typ	Opis
id_posterunku	int	<b>PK (Klucz Zastępczy)</b> . Unikalny identyfikator w hurtowni danych.
numer_posterunku_BK	varchar(20)	<b>BK (Klucz Biznesowy)</b> . Służy do łączenia z systemem źródłowym.
nazwa_posterunku	varchar(100)	Pełna nazwa jednostki policji.
miasto	varchar(50)	Miasto, w którym znajduje się posterunek.

## ZGŁOSZENIE (TABELA FAKTÓW)

Pojedynczy wiersz opisuje jeden fakt przyjęcia anonimowego zgłoszenia.

Atrybut	Typ	Opis
id_czasu_zgloszenia	int	<b>FK do Dim_Czas_Dnia</b> . Czas przyjęcia zgłoszenia.
id_daty_zgloszenia	int	<b>FK do Dim_Data</b> . Data przyjęcia zgłoszenia.
id_posterunku_dw	int	<b>FK do Dim_Posterunek</b> . Posterunek, który przyjął zgłoszenie.

numer_zgloszenia_DD	varchar(30)	<b>Wymiar zdegenerowany.</b> Oryginalny numer zgłoszenia (DD).
poziom_satysfakcji	int	<b>Miara.</b> Ocena w skali 1-10.

## Definicje faktów

**Fakt 1 Mandat:** Wystawienie jednego mandatu przez określonego funkcjonariusza, konkretnej osobie (z uwzględnieniem jej adresu w danym momencie), za określone wykroczenie, w określonym czasie, miejscu, o określonym terminie i statusie płatności.

**Tabela faktów:** Mandat

**Poziom szczegółowości (Granularity):**

- określony numer mandatu (transakcja),
- określona data i godzina wykroczenia,
- określona data terminu płatności,
- określona osoba ukarana (w konkretnym momencie historii jej adresu),
- określony funkcjonariusz,
- określone wykroczenie (rodzaj i nazwa),
- określona lokalizacja wykroczenia,
- określony status płatności.

**Miary i funkcje agregujące:**

- Liczba mandatów - DISTINCT COUNT(numer\_mandatu\_DD)
- Liczba faktów mandatów - COUNT(1)
- Suma kwot mandatów - SUM(kwota\_mandatu)
- Suma punktów karnych - SUM(punkty\_karne)
- Średnia kwota mandatu - AVG(kwota\_mandatu)
- Średnia liczba punktów karnych - AVG(punkty\_karne)
- Wskaźnik Ściągalności -  $[SUM(mandaty\_zaplacone) / SUM(licznik\_mandatow)] * 100$

## Definicja Faktu: Zgłoszenie

**Fakt 2 Zgłoszenie:** Przyjęcie jednego zgłoszenia przez określony posterunek, w określonym czasie i miejscu, z nadaniem mu unikalnego numeru.

**Tabela faktów:** Fact\_Zgloszenia

**Poziom szczegółowości (Granularity):**

Każdy rekord w tabeli faktów reprezentuje unikalne, **przyjęte zgłoszenie**.

- Określony numer zgłoszenia (transakcja).
- Określona data i godzina przyjęcia zgłoszenia.
- Określony posterunek przyjmujący zgłoszenie.

### Miary i funkcje agregujące:

- Liczba zgłoszeń - DISTINCT COUNT(numer\_zgloszenia\_DD)
- Liczba faktów zgłoszeń - COUNT(1)
- Średnia liczba zgłoszeń - AVG(licznik\_zgloszen)
- Średni poziom satysfakcji - AVG(poziom\_satysfakcji)

## Definicje wymiarów

### Wymiary faktu: Mandaty

WYMIAR / ATRYBUT WYMIARU	TABELA / KOLUMNA	TYP
NUMER MANDATU	Fact_Mandaty.numer_mandatu_DD	Wymiar zdegenerowany
DATA WYKROCZENIA	Dim_Data	Wymiar
ROK WYKROCZENIA	Dim_Data.rok	Atrybut wymiaru
MIESIĄC WYKROCZENIA	Dim_Data.nazwa_miesiaca	Atrybut wymiaru
DZIEŃ WYKROCZENIA	Dim_Data.data	Atrybut wymiaru
CZY WEEKEND	Dim_Data.czy_weekend	Atrybut wymiaru
HIERARCHIA DATY	Dim_Data.rok -> Dim_Data.kwartal -> Dim_Data.miesiac -> Dim_Data.data	Wymiar hierarchiczny
DATA TERMINU PŁATNOŚCI	Dim_Data	Wymiar
ROK TERMINU	Dim_Data.rok	Atrybut wymiaru
MIESIĄC TERMINU	Dim_Data.nazwa_miesiaca	Atrybut wymiaru
CZAS WYKROCZENIA	Dim_Czas_Dnia	Wymiar
PORA DNIA	Dim_Czas_Dnia.pora_dnia	Atrybut wymiaru

GODZINA	Dim_Czas_Dnia.godzina	Atrybut wymiaru
HIERARCHIA CZASU	Dim_Czas_Dnia.pora_dnia -> Dim_Czas_Dnia.godzina -> Dim_Czas_Dnia.minuta	Wymiar hierarchiczny
OSOBA UKARANA	Dim_Osoba	Wymiar
PESEL	Dim_Osoba.pesel	Atrybut wymiaru
IMIĘ I NAZWISKO	Dim_Osoba.imie, Dim_Osoba.nazwisko	Atrybut wymiaru
ADRES	Dim_Osoba.adres	Atrybut wymiaru
WYKROCZENIE	Dim_Wykroczenie	Wymiar
KOD WYKROCZENIA	Dim_Wykroczenie.kod_wykroczenia	Atrybut wymiaru
RODZAJ WYKROCZENIA	Dim_Wykroczenie.rodzaj	Atrybut wymiaru
NAZWA WYKROCZENIA	Dim_Wykroczenie.nazwa	Atrybut wymiaru
FUNKCJONARIUSZ (SCD2)	Dim_Funkcjonariusz	Wymiar
NUMER SŁUŻBOWY	Dim_Funkcjonariusz.numer_sluzbowy	Atrybut wymiaru
NAZWISKO I IMIĘ	Dim_Funkcjonariusz.nazwisko, imie	Atrybut wymiaru
STOPIEŃ SŁUŻBOWY	Dim_Funkcjonariusz.stopien_sluzbowy	Atrybut wymiaru
STANOWISKO	Dim_Funkcjonariusz.stanowisko	Atrybut wymiaru
POSTERUNEK	Dim_Posterunek	Wymiar
NAZWA POSTERUNKU	Dim_Posterunek.nazwa_posterunku	Atrybut wymiaru
MIASTO POSTERUNKU	Dim_Posterunek.miasto	Atrybut wymiaru
HIERARCHIA JEDNOSTKI	Dim_Posterunek.miasto -> Dim_Posterunek.nazwa_posterunku -> Dim_Funkcjonariusz.numer_sluzbowy	Wymiar hierarchiczny
LOKALIZACJA	Dim_Lokalizacja	Wymiar
MIASTO ZDARZENIA	Dim_Lokalizacja.miasto	Atrybut wymiaru
OPIS LOKALIZACJI	Dim_Lokalizacja.lokalizacja_opis	Atrybut wymiaru

HIERARCHIA LOKALIZACJI	Dim_Lokalizacja.miasto -> Dim_Lokalizacja.lokalizacja_opis	Wymiar hierarchiczny
STATUS PŁATNOŚCI	Dim_Status	Wymiar
OPIS STATUSU	Dim_Status.status_opis	Atrybut wymiaru

## Wymiary faktu: Zgłoszenie

WYMIAR / ATRYBUT WYMIARU	TABELA / KOLUMNA	TYP
NUMER ZGŁOSZENIA	Fact_Zgloszenia.numer_zgloszenia_DD	Wymiar zdegenerowany
DATA ZGŁOSZENIA	Dim_Data	Wymiar
ROK ZGŁOSZENIA	Dim_Data.rok	Atrybut wymiaru
MIESIĄC ZGŁOSZENIA	Dim_Data.nazwa_miesiaca	Atrybut wymiaru
HIERARCHIA DATY	Dim_Data.rok -> Dim_Data.kwartal -> Dim_Data.miesiac -> Dim_Data.data	Wymiar hierarchiczny
CZAS ZGŁOSZENIA	Dim_Czas_Dnia	Wymiar
PORA DNIA	Dim_Czas_Dnia.pora_dnia	Atrybut wymiaru
GODZINA	Dim_Czas_Dnia.godzina	Atrybut wymiaru
HIERARCHIA CZASU	Dim_Czas_Dnia.pora_dnia -> Dim_Czas_Dnia.godzina -> Dim_Czas_Dnia.minuta	Wymiar hierarchiczny
POSTERUNEK	Dim_Posterunek	Wymiar
NAZWA POSTERUNKU	Dim_Posterunek.nazwa_posterunku	Atrybut wymiaru
MIASTO POSTERUNKU	Dim_Posterunek.miasto	Atrybut wymiaru

HIERARCHIA POSTERUNKU	Dim_Posterunek.miasto -> Dim_Posterunek.nazwa_posterunku	Wymiar hierarchiczny
--------------------------	---	----------------------

## Sprawdzenie wykonalności zapytań

Problem analityczny 1: Dlaczego nastąpił wzrost/spadek liczby mandatów w tym miesiącu?

1. Porównaj liczbę wystawionych mandatów według kategorii wykroczeń w analizowanym miesiącu względem poprzednich miesięcy.  
Miara: Liczba mandatów (agregacja COUNT)  
Wymiar: Dim\_Wykroczenie (atrybut: rodzaj)  
Wymiar: Dim\_Data (atrybuty: rok, nazwa\_miesiaca)
2. Porównaj liczbę mandatów wystawionych w godzinach szczytu (7:00-9:00 i 15:00-17:00) w bieżącym i poprzednim miesiącu.  
Miara: Liczba mandatów (agregacja COUNT)  
Wymiar: Dim\_Czas\_Dnia (atrybut: godzina, filtrowany do wartości 7, 8, 15, 16)  
Wymiar: Dim\_Data (atrybuty: rok, nazwa\_miesiaca)
3. W jakim stopniu zmiana (wzrost/spadek) liczby przyjmowanych zgłoszeń w tym miesiącu jest skorelowana ze zmianą (wzrostem/spadkiem) liczby wystawionych mandatów?  
Miara 1: Liczba mandatów (COUNT)  
Miara 2: Liczba zgłoszeń (COUNT)  
Wymiar (wspólny): Dim\_Data (atrybuty: rok, nazwa\_miesiaca)
4. Porównaj efektywność poszczególnych funkcjonariuszy (liczba mandatów) w bieżącym i poprzednim miesiącu.  
Miara: Liczba mandatów (agregacja COUNT)  
Wymiar: Dim\_Funkcjonariusz (atrybuty: imie, nazwisko, numer\_sluzbowy)  
Wymiar: Dim\_Data (atrybuty: rok, nazwa\_miesiaca)
5. Porównaj liczbę mandatów według posterunków w bieżącym i poprzednim miesiącu.  
Miara: Liczba mandatów (agregacja COUNT)  
Wymiar: Dim\_Posterunek (atrybut: nazwa\_posterunku)  
Wymiar: Dim\_Data (atrybuty: rok, nazwa\_miesiaca)

Problem analityczny 2: Jakie czynniki determinują powtarzalność wykroczeń (recydywę)?

1. Jaki typ wykroczenia najczęściej prowadzi do recydywy?  
 Miara: Liczba mandatow (używana do identyfikacji recydywy przez grupowanie)  
 Wymiar: Dim\_Osoba (używany do grupowania mandatów na osobę)  
 Wymiar: Dim\_Wykroczenie (atrybut: rodzaj lub nazwa, analizowany dla osób z recydywą)
2. Czy surowość kary (wysokość mandatu) wpływa na recydywę?  
 Miara 1: kwota\_mandatu (analizowana miara, np. AVG dla pierwszego mandatu)  
 Miara 2: Liczba mandatow (używana do identyfikacji recydywy)  
 Wymiar: Dim\_Osoba (używany do grupowania)
3. W jakich miastach recydywa jest największa?  
 Miara: Liczba mandatow (używana do identyfikacji recydywy)  
 Wymiar: Dim\_Osoba (używany do grupowania)  
 Wymiar: Dim\_Lokalizacja (atrybut: miasto, analizowany dla mandatów osób z recydywą)
4. Czy osoby, które zignorowały obowiązek zapłaty (lub zapłaciły po terminie) są bardziej skłonne do ponownego popełnienia wykroczenia?  
 Miara: Liczba mandatow (używana do identyfikacji recydywy)  
 Wymiar: Dim\_Osoba (używany do grupowania)  
 Wymiar: Dim\_Status (atrybut: status\_opis, analizowany dla pierwszego mandatu osoby)
5. Wpływ jednostki (Czy funkcjonariusz obsługujący wykroczenie ma wpływ na recydywę zatrzymanej osoby?)  
 Miara: Liczba mandatow (używana do identyfikacji recydywy)  
 Wymiar: Dim\_Osoba (używany do grupowania)  
 Wymiar: Dim\_Funkcjonariusz (atrybut: numer\_sluzbowy, analizowany dla pierwszego mandatu osoby)

## Sprawdzenie źródeł danych

NAZWA TABELI	KOLUMNA	ŹRÓDŁO
Fact_Mandaty	Pojedynczy wiersz opisuje jeden fakt wystawienia mandatu.	
	id_daty_wykroczenia	Klucz obcy z wymiaru Dim_Data. Na podstawie czas_wykroczenia z tabeli Mandaty w bazie Police Record.

	id_czasu_wykroczenia	Klucz obcy z wymiaru Dim_Czas_Dnia. Na podstawie czas_wykroczenia z tabeli Mandaty w bazie Police Record.
	id_daty_terminu_plat	Klucz obcy z wymiaru Dim_Data. Na podstawie termin_platności z tabeli Mandaty w bazie Police Record.
	id_osoby	Klucz obcy z wymiaru Dim_Osoba. Na podstawie pesel_obciazonego z tabeli Mandaty w bazie Police Record.
	id_funkcjonariusza	Klucz obcy z wymiaru Dim_Funkcjonariusz. Na podstawie numer_sluzbowy z tabeli Mandaty w bazie Police Record.
	id_wykroczenia	Klucz obcy z wymiaru Dim_Wykroczenie. Na podstawie kod_wykroczenia z tabeli Mandaty w bazie Police Record.
	id_lokalizacji	Klucz obcy z wymiaru Dim_Lokalizacja. Na podstawie lokalizacja z tabeli Mandaty w bazie Police Record.
	id_statusu	Klucz obcy z wymiaru Dim_Status. Na podstawie status_platności z tabeli Mandaty w bazie Police Record.
	numer_mandatu_BK	Wymiar zdegenerowany. Pobrano z numer_mandatu z tabeli Mandaty w bazie Police Record.
	kwota_mandatu	Miara. Pobrano z kwota_mandatu z tabeli Mandaty w bazie Police Record.
	punkty_karne	Miara. Pobrano z punkty_karne (nowa kolumna) z tabeli Mandaty w bazie Police Record.
Fact_Zgloszenia	Pojedynczy wiersz opisuje jeden fakt przyjęcia anonimowego zgłoszenia.	
	id_czasu_zgloszenia	Klucz obcy z wymiaru Dim_Czas_Dnia. Na podstawie data_zgloszenia z tabeli Zgloszenia w bazie Police Record.
	id_daty_zgloszenia	Klucz obcy z wymiaru Dim_Data. Na podstawie data_zgloszenia z tabeli Zgloszenia w bazie Police Record.
	id_posterunku_dw	Klucz obcy z wymiaru Dim_Posterunek. Na podstawie id_posterunku z tabeli Zgloszenia w bazie Police Record.
	numer_zgloszenia_BK	Wymiar zdegenerowany. Pobrano z id_zgloszenia z tabeli Zgloszenia w bazie Police Record.
	poziom_satysfakcji	Miara. Pobrano z poziom_satysfakcji (nowa kolumna) z tabeli Zgloszenia w bazie Police Record.
Dim_Data	Pojedynczy wiersz opisuje jeden dzień w kalendarzu.	
	Wszystkie kolumny	Dane generowane na podstawie kalendarza, przed procesem ETL.
Dim_Czas_Dnia	Pojedynczy wiersz opisuje konkretną godzinę/minutę w ciągu doby.	
	Wszystkie kolumny	Dane generowane na podstawie zegara (24h), przed procesem ETL.
Dim_Osoba	Pojedynczy wiersz opisuje jedną osobę ukaraną mandatem.	

	id_osoby	Klucz zastępczy (PK). Generowany przez bazę danych.
	pesel	Klucz biznesowy (BK). Pobrano z pesel z tabeli Osoby w bazie Police Record.
	imie	Pobrano z imie z tabeli Osoby w bazie Police Record.
	nazwisko	Pobrano z nazwisko z tabeli Osoby w bazie Police Record.
	adres	Pobrano z adres z tabeli Osoby w bazie Police Record.
Dim_Wykroczenie	Pojedynczy wiersz opisuje jeden typ wykroczenia.	
	id_wykroczenia	Klucz zastępczy (PK). Generowany przez bazę danych.
	kod_wykroczenia	Klucz biznesowy (BK). Pobrano z kod z tabeli Wykroczenia w bazie Police Record.
	rodzaj	Pobrano z rodzaj z tabeli Wykroczenia w bazie Police Record.
	nazwa	Pobrano z nazwa z tabeli Wykroczenia w bazie Police Record.
Dim_Status	Pojedynczy wiersz opisuje status płatności.	
	id_statusu	Klucz zastępczy (PK). Generowany przez bazę danych.
	status_opis	Wartości generowane na podstawie unikalnych wartości z status_płatności z tabeli Mandaty w bazie Police Record.
Dim_Lokalizacja	Pojedynczy wiersz opisuje lokalizację.	
	id_lokalizacji	Klucz zastępczy (PK). Generowany przez bazę danych.
	lokalizacja_opis	Wartości generowane na podstawie unikalnych wartości z lokalizacja z tabeli Mandaty w bazie Police Record.
	miasto	Wartość parsowana (wyodrębniana) z pola lokalizacja z tabeli Mandaty w bazie Police Record.
Dim_Posterunek	Pojedynczy wiersz opisuje posterunek policji.	
	id_posterunku_dw	Klucz zastępczy (PK). Generowany przez bazę danych.
	id_posterunku_BK	Klucz biznesowy (BK). Pobrano z id_posterunku z tabeli Posterunki w bazie Police Record.
	nazwa_posterunku	Pobrano z nazwa z tabeli Posterunki w bazie Police Record.
	miasto	Pobrano z miasto z tabeli Posterunki w bazie Police Record.
Dim_Funkcjonariusz	Pojedynczy wiersz opisuje jednego funkcjonariusza.	
	id_funkcjonariusza_dw	Klucz zastępczy (PK). Generowany przez bazę danych. (SCD2)
	id_posterunku_dw	Klucz obcy z wymiaru Dim_Posterunek. Na podstawie id_posterunku (Police Record) lub StationNumber (CSV).
	numer_sluzbowy	Klucz biznesowy (BK). Pobrano z BadgeNumber (CSV) oraz numer_sluzbowy (Police Record).
	imie	Pobrano z FirstName z pliku police_staff.csv.

	nazwisko	Pobrano z LastName z pliku police_staff.csv.
	stopien_sluzbowy	Pobrano z Rank z pliku police_staff.csv.
	stanowisko	Pobrano z Position z pliku police_staff.csv.
	data_przyjecia_do_sluzby	Pobrano z StartDate z pliku police_staff.csv.
	czy_aktualny	Wartość generowana ("1" lub "0") podczas procesu ETL (Implementacja SCD2).