

## Specyfikacja wymagań dla procesu realizacji patroli drogowych

### 1. Ogólny opis procesu biznesowego

1. Ogólny opis procesu biznesowego wraz z opisem metryk wydajności generowanych przez ten proces, ewentualne aktualne problemy analityczne.

Realizacja patroli drogowych przebiega w następujący sposób. Funkcjonariusze rozpoczynają patrol od zalogowania się do tabletu w radiowozie, co rejestruje czas rozpoczęcia służby, skład patrolu i przypisany pojazd. Dane z tabletów są przesyłane do chmury, z której generowany jest cotygodniowy plik CSV opisujący odbyte patrole.

Podczas służby funkcjonariusze patrolują wyznaczony rejon. W przypadku wykrycia zdarzenia drogowego lub otrzymania zgłoszenia, podejmują interwencję – może być to poza ich przypisaną dzielnicą. Interwencja kończy się nałożeniem kary w postaci mandatu, pouczenia lub skierowania wniosku do sądu. Dane dotyczące kary, w tym informacje o osobie notowanej i podstawie prawnej, są wprowadzane do policyjnej bazy danych.

Mimo zbierania danych, największym problemem analitycznym jest aktualnie brak szczegółowej wiedzy na temat tego, które rodzaje kar najskuteczniej wpływają na poprawę bezpieczeństwa.

#### **Coroczny spadek liczby wypadków drogowych o 5%.**

**Comiesięczne zwiększenie stosunku liczby wystawionych kar do ilości patroli o 0,5% miesięcznie względem miesiąca poprzedniego.**

2. Typowe zapytania

Ile patroli przeprowadzono w danym miesiącu i ilu funkcjonariuszy brało w nich udział?

Jaki jest stosunek liczby patroli do liczby wystawionych mandatów?

W jakich okresach należy zwiększyć liczbę patroli?

Którzy funkcjonariusze są najbardziej aktywni?

Ile interwencji przypada średnio na patrol?

Jaka jest zależność między liczbą patroli a liczbą wypadków drogowych?

Ile mandatów i pouczeń wystawiono w danym miesiącu?

Jak rozkładają się kary na poszczególne rodzaje przewinień?

Które wykroczenia są najczęściej karane pouczeniem zamiast mandatem?

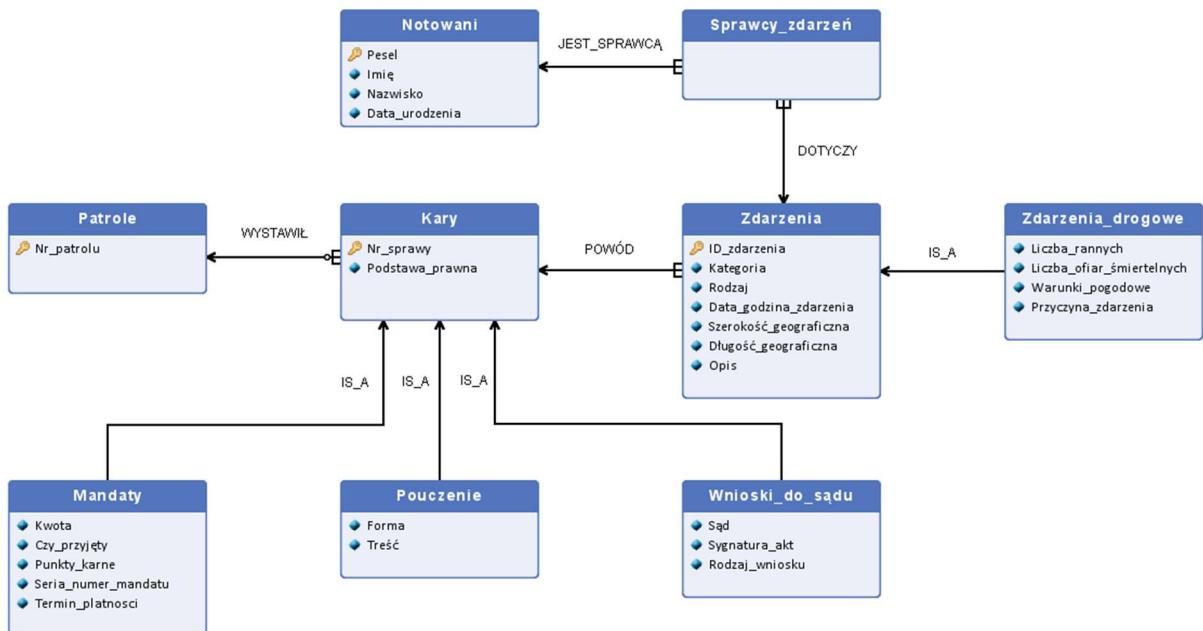
3. Dane

Wszystkie dane dotyczące patroli są pobierane z pliku CSV generowanego przez chmurę, z którą są połączone tablety patroli oraz z bazy danych, która opisuje wystawione kary.

### 2. Struktura źródeł danych

1. Policyjna baza danych

## Hurtownie danych: Wymagania dla policji



| Nazwa tabeli    | Atrybut   | Typ atrybutu        | Opis  |
|-----------------|---|---------------------|---|
| Notowani        | Osiobe notowane w systemie – sprawcy przynajmniej 1 zdarzenia, identyfikowani numerem pesel   |                     |   |
|                 | Pesel   | Znakowy – 11 cyfr   | PK  |
|                 | Imię  | Znakowy – 20 znaków | Imię osoby notowanej  |
|                 | Nazwisko  | Znakowy – 30 znaków | Nazwisko osoby notowanej  |
|                 | Data_urodzenia  | DateTime            | Data urodzenia osoby notowanej  |
| Sprawcy_zdarzeń | Implementacja związku wiele do wiele pomiędzy Notowani i Zdarzenia oznaczająca bycie sprawcą danego zdarzenia. Identyfikowany kluczami obcymi tabel Notowani i Zdarzenia. |                     |   |
|                 | FK_Nutowani   | Znakowy – 11 cyfr   | Element PK, FK wskazujący na notowaną osobę. Implementacja związku 1:N pomiędzy Notowani a Sprawcy_zdarzeń (JEST SPRAWCA) |
|                 | FK_Zdarzenia  | Numeryczny          | Element PK, FK wskazujący na zdarzenie. Implementacja związku 1:N pomiędzy Sprawcy_zdarzeń a Zdarzenia (DOTYCZY)          |
| Zdarzenia       | Unikalne zdarzenia (przewinienia), spowodowane przez pewnego sprawcę, skutkujące wystawieniem pewnej kary.  |                     |   |
|                 | ID_zdarzenia  | Numeryczny          | PK, generowany automatycznie  |

Hurtownie danych: Wymagania dla policji

|                          |   |                      |   |
|--------------------------|---|----------------------|---|
|                          | Kategoria   | Znakowy – 30 znaków  | Kategoria zdarzenia   |
|                          | Rodzaj  | Znakowy – 50 znaków  | Dokładny rodzaj zdarzenia   |
|                          | Data_godzina_zdarzenia  | DateTime             | Data i czas wystąpienia zdarzenia   |
|                          | Szerokość_geograficzna  | Numeryczny           | Współrzędna geograficzna miejsca zdarzenia  |
|                          | Długość_geograficzna  | Numeryczny           | Współrzędna geograficzna miejsca zdarzenia  |
|                          | Opis  | Znakowy – 500 znaków | Opis zdarzenia  |
| <i>Zdarzenia_drogowe</i> | Zdarzenia z kategorii „Drogowe” – zdarzenia uzupełnione przez dodatkowe dane.     |                      |   |
|                          | FK_Zdarzenia  | Numeryczny           | PK, FK wskazujący na Zdarzenia. Implementacja związku IS_A pomiędzy Zdarzenia_drogowe a Zdarzenia |
|                          | Liczba_rannych  | Numeryczny           | Całkowita liczba osób rannych w zdarzeniu   |
|                          | Liczba_ofiar_śmiertelnych   | Numeryczny           | Całkowita liczba ofiar śmiertelnych w zdarzeniu   |
|                          | Warunki_pogodowe  | Znakowy - 50 znaków  | Opis warunków pogodowych w momencie zdarzenia   |
|                          | Przyczyna_zdarzenia   | Znakowy – 200 znaków | Główna przyczyna zdarzenia  |
| <i>Kary</i>              | Kary wystawiane przez patrole za spowodowanie zdarzeń (wykroczeń, wypadków itp.). |                      |   |
|                          | Nr_sprawy   | Znakowy – 20 znaków  | PK, unikalny numer akt sprawy   |
|                          | FK_Zdarzenia  | Numeryczny           | FK wskazujący na Zdarzenia. Implementacja związku 1:N pomiędzy Kary a Zdarzenia (POWÓD)           |
|                          | FK_Patrole  | Znakowy – 10 znaków  | FK wskazujący na Patrole. Implementacja związku 1:0...N pomiędzy Patrole a Kary (WYSTAWIŁ)        |
|                          | Podstawa_prawna   | Znakowy – 20 znaków  | Numer artykułu KK będącego podstawą prawną nałożonej kary   |
| <i>Mandaty</i>           | Jeden z rodzajów kary, z dodatkowymi atrybutami charakterystycznymi dla mandatów. |                      |   |

Hurtownie danych: Wymagania dla policji

|                        |  |                      |  |
|------------------------|--|----------------------|--|
|                        | FK_Kary  | Znakowy – 20 znaków  | PK, FK wskazujący na Kary.<br>Implementacja związku IS_A pomiędzy Mandaty a Kary         |
|                        | Kwota  | Numeryczny           | Kwota wystawionego mandatu w złotówkach z dokładnością do groszy                         |
|                        | Czy_przyjęty   | Boolean              | Informacja czy mandat został przyjęty  |
|                        | Punkty_karne   | Numeryczny           | Całkowita liczba przyznanych punktów karnych   |
|                        | Seria_numer_mandatu  | Znakowy – 20 znaków  | Unikalne seria i numer mandatu   |
|                        | Termin_płatności   | DateTime             | Data i czas, do którego mandat musi zostać opłacony                                      |
| <i>Pouczenia</i>       | Jeden z rodzajów kary, z dodatkowymi atrybutami charakterystycznymi dla pouczeń.           |                      |  |
|                        | FK_Kary  | Znakowy – 20 znaków  | PK, FK wskazujący na Kary.<br>Implementacja związku IS_A pomiędzy Pouczenia a Kary       |
|                        | Forma  | Znakowy – 30 znaków  | Forma udzielonego pouczenia (ustne, pisemne)   |
|                        | Treść  | Znakowy – 500 znaków | Przybliżona treść udzielonego pouczenia  |
| <i>Wnioski_do_sądu</i> | Jeden z rodzajów kary, z dodatkowymi atrybutami charakterystycznymi dla wniosków sądowych. |                      |  |
|                        | FK_Kary  | Znakowy – 20 znaków  | PK, FK wskazujący na Kary.<br>Implementacja związku IS_A pomiędzy Wnioski_do_sądu a Kary |
|                        | Sąd  | Znakowy – 30 znaków  | Sąd, do którego został złożony wniosek   |
|                        | Sygnatura_akt  | Znakowy – 10 znaków  | Sygnatura sądowych akt sprawy  |
|                        | Rodzaj_wniosku   | Znakowy – 30 znaków  | Rodzaj złożonego wniosku (o ukaranie, o zastosowanie środka wychowawczego)               |

## Hurtownie danych: Wymagania dla policji

|         |   |                     |                                    |
|---------|---|---------------------|------------------------------------|
| Patrole | Tabela na potrzeby połączenia z plikiem CSV generowanym z chmury zsynchronizowanej z tabletami policjantów patrolujących. |                     |                                    |
|         | Nr_patrolu  | Znakowy – 10 znaków | PK, unikalny identyfikator patrolu |

### 2. CSV generowany z tabletów

Każdy wiersz reprezentuje jeden patrol (czyli jedną służbę jednego/dwóch policjantów, z konkretną datą rozpoczęcia i zakończenia, z unikalnym numerem patrolu). Dane są dopisywane:

- Kolumna 1 – nr\_patrolu (unikalny ciąg 10 znaków)
- Kolumna 2 – data\_godzina\_rozpoczecia (TimeStamp)
- Kolumna 3 – data\_godzina\_zakończenia (TimeStamp)
- Kolumna 4 – dzielnica\_patrolu (ciąg 20 znaków)
- Kolumna 5 – radiowóz (model auta, ciąg 30 znaków)
- Kolumna 6 – nr\_rejestracyjny (ciąg 7-8 znaków)
- Kolumna 7 – nr\_odznaki\_kierowcy (ciąg 6 znaków)
- Kolumna 8 – nr\_odznaki\_partnera (ciąg 6 znaków)

Uwaga!

W przypadku zmiany dzielnicy patrolu lub radiowozu, dane w arkuszu są aktualizowane. W arkuszu jest zawsze tylko jeden wiersz dla jednego patrolu.

## 3. Scenariusze problemów analitycznych

1. Dlaczego w tym miesiącu nastąpił wzrost/spadek stosunku liczby wypadków drogowych do ilości patroli?
  1. Porównaj liczbę patroli i wypadków drogowych w poszczególnych dzielnicach w analizowanym miesiącu względem poprzednich.
  2. Porównaj średnią liczbę interwencji przypadającą na jeden patrol w bieżącym i poprzednich miesiącach.
  3. Zidentyfikuj okresy (dni tygodnia, pory dnia) o największej liczbie wypadków w bieżącym i poprzednich miesiącach.
  4. Porównaj liczbę przeprowadzonych patroli oraz liczbę zaangażowanych funkcjonariuszy w bieżącym i poprzednich miesiącach.
  5. Wskaź dzielnice i rodzaje zdarzeń, które dominowały w statystykach w bieżącym i poprzednich miesiącach.
  6. Podaj wpływ natężenia ruchu drogowego na liczbę zdarzeń w poszczególnych dzielnicach w godzinach szczytu.
2. Jaki wpływ na ilość zarejestrowanych przestępstw ma rodzaj wystawianych kar?
  1. Przeanalizuj rozkład rodzajów kar (mandat, pouczenie, wniosek do sądu) dla poszczególnych typów przewinień.
  2. Zidentyfikuj wykroczenia, przy których funkcjonariusze najczęściej decydują się na pouczenie zamiast mandatu.
  3. Wskaź najaktywniejszych funkcjonariuszy pod względem liczby i rodzaju nałożonych kar.
  4. Porównaj liczbę wystawionych mandatów i pouczeń w poszczególnych dzielnicach w bieżącym i poprzednich miesiącach.
  5. Czy w rejonach patrolowanych przez funkcjonariuszy częściej kierujących wnioski do sądu obserwuje się większy spadek liczby wypadków?

6. Czy decyzja o nałożeniu pouczenia była podyktowana niską szkodliwością czynu, czy postawą sprawcy.

## 4. Potrzebne dane dla problemów analitycznych

**Problem analityczny: „Dlaczego w tym miesiącu nastąpił wzrost/spadek stosunku liczby wypadków drogowych do ilości patroli?”**

1. Porównaj liczbę patroli i wypadków drogowych w poszczególnych dzielnicach w analizowanym miesiącu względem poprzednich.
  - Liczba patroli: plik CSV z tabletów, zliczanie unikalnych wartości z kolumny nr\_patrolu (w danym miesiącu)
  - Dzielnica patrolu: plik CSV z tabletów, kolumna dzielnica\_patrolu
  - Liczba wypadków drogowych: Policyjna baza danych, tabela Zdarzenia\_drogowe, zliczanie rekordów
  - Lokalizacja wypadku: Policyjna baza danych, tabela Zdarzenia, kolumny Szerokość\_geograficzna i Długość\_geograficzna. Wymaga dodatkowego przetworzenia w celu przypisania współrzędnych do konkretnej dzielnicy.
  - Miesiąc patrolu/wypadku: plik CSV, kolumna data\_godzina\_rozpoczęcia oraz Baza danych, tabela Zdarzenia, kolumna Data\_godzina\_zdarzenia
2. Porównaj średnią liczbę interwencji przypadającą na jeden patrol w bieżącym i poprzednich miesiącach.
  - Liczba interwencji: Policyjna baza danych, tabela Kary, zliczanie rekordów po Nr\_sprawy
  - Identyfikator patrolu: Policyjna baza danych, tabela Kary, kolumna FK\_Patrole do połączenia z plikiem CSV
  - Liczba patroli: plik CSV z tabletów, zliczanie unikalnych wartości z kolumny nr\_patrolu
  - Miesiąc interwencji: Baza danych, tabela Zdarzenia, kolumna Data\_godzina\_zdarzenia
3. Zidentyfikuj okresy (dni tygodnia, pory dnia) o największej liczbie wypadków w bieżącym i poprzednich miesiącach.
  - Data i godzina wypadku: Policyjna baza danych, tabela Zdarzenia, kolumna Data\_godzina\_zdarzenia
  - Liczba wypadków: Policyjna baza danych, tabela Zdarzenia\_drogowe, zliczanie rekordów
4. Porównaj liczbę przeprowadzonych patroli oraz liczbę zaangażowanych funkcjonariuszy w bieżącym i poprzednich miesiącach.
  - Liczba patroli: plik CSV z tabletów, zliczanie unikalnych wartości z kolumny nr\_patrolu
  - Liczba zaangażowanych funkcjonariuszy: plik CSV z tabletów, zliczanie unikalnych wartości z kolumn nr\_odznaki\_kierowcy i nr\_odznaki\_partnera
  - Miesiąc patrolu: plik CSV z tabletów, kolumna data\_godzina\_rozpoczęcia
5. Wskaź dzielnice i rodzaje zdarzeń, które dominowały w statystykach w bieżącym i poprzednich miesiącach.
  - Dzielnica: plik CSV z tabletów, kolumna dzielnica\_patrolu (połączona z karą przez nr\_patrolu)
  - Rodzaj zdarzenia: Policyjna baza danych, tabela Zdarzenia, kolumna Rodzaj
6. Podaj wpływ natężenia ruchu drogowego na liczbę zdarzeń w poszczególnych dzielnicach w godzinach szczytu.
  - Liczba zdarzeń: Policyjna baza danych, tabela Zdarzenia\_drogowe, zliczanie rekordów
  - Dzielnica patrolu: plik CSV z tabletów, kolumna dzielnica\_patrolu
  - Godziny szczytu – ustalone przedziały czasowe (np. 7:00-9:00, 15:00 – 17:00), wyznaczone na podstawie kolumny Data\_godzina\_zdarzenia w tabeli Zdarzenia
  - Natężenie ruchu drogowego – brak danych w obecnych źródłach (policyjna baza i plik CSV nie zawierają informacji o natężeniu ruchu drogowego)

Propozowane sposoby pozyskania tych danych:

- pobieranie danych o natężeniu ruchu z ogólnodostępnych systemów miejskich (np. API GDDKiA)
- uzupełnienie pliku CSV o pole natężenie ruchu generowane automatycznie podczas rozpoczęcia patrolu dzięki danym z Google Maps

## **Problem analityczny: „Jaki wpływ na ilość zarejestrowanych przestępstw ma rodzaj wystawianych kar?”**

1. Przeanalizuj rozkład rodzajów kar (mandat, pouczenie, wniosek do sądu) dla poszczególnych typów przewinień.
  - Rodzaj kary: Informacja wynikająca z istnienia powiązanego rekordu w tabelach Mandaty, Pouczenia lub Wnioski\_do\_sądu w Policyjnej bazie danych.
  - Rodzaj przewinienia: Policyjna baza danych, tabela Zdarzenia, kolumna Rodzaj
2. Zidentyfikuj wykroczenia, przy których funkcjonariusze najczęściej decydują się na pouczenie zamiast mandatu.
  - Rodzaj wykroczenia: Policyjna baza danych, tabela Zdarzenia, kolumna Rodzaj
  - Informacje o karze: Zliczanie rekordów w tabelach Pouczenia i Mandaty dla danego rodzaju wykroczenia
3. Wskaź najaktywniejszych funkcjonariuszy pod względem liczby i rodzaju nałożonych kar.
  - Identyfikator funkcjonariusza: plik CSV z tabletów, kolumny nr\_odznaki\_kierowcy i nr\_odznaki\_partnera
  - Identyfikator patrolu: Policyjna baza danych, tabela Kary, kolumna FK\_Patrole
  - Rodzaj kary: Policyjna baza danych, tabele Mandaty, Pouczenia, Wnioski\_do\_sądu
4. Porównaj liczbę wystawionych mandatów i pouczeń w poszczególnych dzielnicach w bieżącym i poprzednich miesiącach.
  - Liczba mandatów i pouczeń: Policyjna baza danych, zliczanie rekordów w tabelach Mandaty i Pouczenia
  - Dzielnica: plik CSV z tabletów, kolumna dzielnica\_patrolu
  - Miesiąc kary: Policyjna baza danych, tabela Zdarzenia, kolumna Data\_godzina\_zdarzenia
5. Czy w rejonach patrolowanych przez funkcjonariuszy częściej kierujących wnioskami do sądu obserwuje się większy spadek liczby wypadków?
  - Liczba wniosków do sądu: Policyjna baza danych, zliczanie rekordów w tabeli Wnioski\_do\_sądu pogrupowanych po FK\_Patrole
  - Rejon: plik CSV z tabletów, kolumna dzielnica\_patrolu
  - Liczba wypadków w dzielnicy w czasie: Policyjna baza danych, tabela Zdarzenia\_drogowe
6. Czy decyzja o nałożeniu pouczenia była podyktowana niską szkodliwością czynu, czy postawą sprawcy?
  - Rodzaj kary: Policyjna baza danych, tabela Pouczenia
  - Szkodliwość czynu i postawa sprawcy: brak danych w obecnych źródłach (policyjna baza i plik CSV nie zawierają informacji o powodach decyzji funkcjonariusza)Propozowane sposoby pozyskania tych danych:
  - rozszerzenie wpisów w systemie używanym przez patrole o wymagane pole „powód decyzji o karze” (np. niewielka szkodliwość czynu) i automatyczne zapisywanie tej informacji w bazie danych w tabeli Kary
  - okresowe ankietowanie funkcjonariuszy dotyczące powodów stosowania pouczeń w formie raportów wgrywanych do systemu