



ANALIZA PRZESTĘPCZOŚCI W CHICAGO W LATACH 2018-2022

Z UWZGLĘDNIENIEM CZYNNIKÓW EKONOMICZNO-DEMOGRAFICZNYCH

Hurtownie Danych 2025

Dominik Czech (272725)

Contents

Etap 1 – Zakres Opracowania Projektu	3
1. Tytuł Projektu	3
2. Charakterystyka dziedziny problemowej.....	3
3. Cel przedsięwzięcia oraz zakres analizy.....	3
Cel przedsięwzięcia	3
Zakres analizy	4
4. Źródła Danych	5
5. Zakres Danych.....	6
Crimes 2001 to Present	6
ACS_5_Year_Data_By_Ward	7
IUCR Codes.....	9
6. Definicja Typów Encji/Klas.....	10
7. Przykładowe analizy wielowymiarowe	10
8. Utworzenie Bazy Danych	12
Etap 2 – Proces ETL	14
1. Przygotowanie Danych Źródłowych	14
2. Architektura Hurtowni Danych.....	14
3. Struktura procesu ETL (SSIS)	15
1. Delete Tables	15
2. Create Tables.....	16
3. Load Dim_TimeOfDay.....	18
4. Load Dim_Date	18
5. Load Dim_WardDemographics.....	19
6. Load Dim_IUCR.....	20
7. Load Dim_Location	21
8. Load Fact_Crime.....	22
9. Constraints (References)	24
10. Success Message	26

11. Error Handling	26
Etap 3 – Kostka OLAP w SSAS	27
1. Opis Celu Etapu	27
2. Struktura Kostki	27
3. Wymiary (dimensions)	28
Dim_Date:	28
Dim_TimeOfDay.....	29
Dim_Location	29
Dim_WardDemographics.....	29
Dim_IUCR.....	30
Dim_CrimeFlags (degenerate dimension):	30
4. Analizy – Etap 3 (realizacja punktu 7 Etapu 1)	31
Analiza 1: Rozkład kategorii przestępstw wg dominującej grupy etnicznej	31
Analiza 2: Typy przestępstw vs status ekonomiczny wardu	36
Analiza 3: Przestępstwa w wardach z młodą populacją	41
Analiza 4: Skuteczność aresztowań względem statusu ekonomicznego	43
Analiza 5: Wskaźnik przestępstw domowych a profil wiekowy i etniczny	48
Analiza 6: Trend przestępczości według Statusu Ekonomicznego Wardu	49
Analiza 7: Liczba przestępstw wg dnia tygodnia i lokalizacji.	51
Analiza 8: Liczba przestępstw wg typu i rezultatu (areszt)	53
Analiza 9: Skuteczność aresztowań na przestrzeni lat.	58
Analiza 10: Kategoria przestępstwa w zależności od pory dnia.	63
5. KPI + Zestawienia.....	66
6. Wnioski z Analizy Danych:	67

Etap 1 – Zakres Opracowania Projektu

1. Tytuł Projektu

Analiza przestępczości w Chicago z uwzględnieniem czynników ekonomiczno-demograficznych i klasyfikacji UCR

2. Charakterystyka dziedziny problemowej

Dziedziną problemową projektu jest **analiza przestępczości miejskiej** w kontekście **uwarunkowań socjoekonomicznych** na terenie miasta **Chicago** w latach **2018–2022**. Współczesne miasta, zwłaszcza metropolie takie jak Chicago, borykają się z dużą zmiennością przestępczości w zależności od lokalizacji, czasu, a także warunków społecznych i ekonomicznych zamieszkującej populacji.

Miasto Chicago publikuje bogate źródła danych publicznych dotyczących **przestępstw** (np. rozbojów, kradzieży, napaści, morderstw), które zawierają informacje **o typie przestępstwa, miejscu, czasie, czy aresztowaniu sprawcy**. Samo zestawienie tych informacji może dostarczyć wartościowych statystyk, jednak dopiero połączenie ich z dodatkowymi wymiarami — np. **dochodami mieszkańców, strukturą demograficzną i etniczną** — pozwala na uzyskanie pełniejszego obrazu uwarunkowań sprzyjających przestępczości.

Celem jest zbadanie, jak czynniki takie jak wiek, płeć, dochód czy pochodzenie etniczne populacji w różnych rejonach Chicago mogą korelować z poziomem przestępczości oraz skutecznością egzekwowania prawa. Taka analiza może pomóc w identyfikacji obszarów wymagających interwencji społecznej, zmiany polityki bezpieczeństwa, czy lepszego alokowania zasobów policji i służb miejskich.

3. Cel przedsięwzięcia oraz zakres analizy

Cel przedsięwzięcia

Głównym celem projektu jest **opracowanie hurtowni danych**, która umożliwi analizę przestępczości w Chicago w latach 2018–2022 z uwzględnieniem danych socjoekonomicznych i demograficznych pochodzących z American Community Survey (ACS) oraz innych źródeł publicznych. Projekt zakłada stworzenie **kostki OLAP**, która umożliwi prowadzenie zaawansowanych analiz wielowymiarowych w celu:

- identyfikacji **trendów przestępczości** w zależności od lokalizacji (dzielnica/Ward), czasu (rok, miesiąc) oraz typu przestępstwa,
- znalezienia **zależności między przestępczością a strukturą demograficzną, dochodami i pochodzeniem etnicznym mieszkańców**,

- wsparcia **procesu decyzyjnego** dla służb miejskich, planistów polityki społecznej oraz władz miasta,
- stworzenia podstaw do **budowy modeli predykcyjnych** w przyszłości, które mogłyby przewidywać wzrost zagrożenia przestępczością w określonych obszarach.

Dzięki hurtowni danych i analizom OLAP użytkownicy końcowi będą mogli szybko uzyskać odpowiedzi na pytania typu:

- Czy wyższe dochody w danym rejonie korelują z niższą przestępczością?
- Jakie typy przestępstw dominują w dzielnicach o przeważającej młodej populacji?
- Czy określone grupy etniczne są częściej ofiarami lub sprawcami określonych przestępstw?
- Jak zmieniała się liczba aresztowań w różnych latach i czy ma to związek z danymi demograficznymi?

Zakres analizy

Zakres analiz obejmie dane z lat **2018–2022**, aby zapewnić zgodność z rocznym zakresem danych ACS 2023 (które są agregatem 5-letnim). Projekt zakłada analizę następujących aspektów:

1. Czas:

- Rok, miesiąc, dzień tygodnia, pora dnia – pozwala analizować sezonowość przestępstw.

2. Lokalizacja:

- Ward (okręg polityczny),
- Community Area (obszar geograficzny),
- Dzielnica, rejon policyjny (District, Beat),
- Koordynaty geograficzne (do ewentualnej wizualizacji na mapach).

3. Typy przestępstw:

- Klasyfikacja według rodzaju przestępstwa (np. rozbój, napaść, włamanie),
- Opis szczegółowy przestępstwa (Description, IUCR),
- Informacja, czy doszło do aresztowania.

4. Czynniki społeczno-ekonomiczne:

- Przedziały dochodów mieszkańców w danym Wardzie,
- Rozkład populacji wg wieku i płci,
- Struktura rasowa i etniczna populacji,
- Wielkość populacji ogólnej.

5. Miary analityczne (fakty):

- Liczba przestępstw,
- Liczba przestępstw zakończonych aresztem,
- Udział danego typu przestępstw względem wszystkich przestępstw w obszarze,
- Liczba przestępstw na 1000 mieszkańców (miara nieaddytywna).

4. Źródła Danych

Lp.	Plik	Typ	Liczba rekordów	Rozmiar [MB]	Opis
1.	Crimes 2001 to Present (https://catalog.data.gov/dataset/crimes-2001-to-present)	csv	8315520	~1850	Dane o przestępstwach w Chicago z lat 2001-2025. Na potrzeby analizy czynników socjoekonomicznych w analizie zostanie użyty zakres lat 2018-2022.
2.	ACS_5_Year_Data_by_Ward.csv (https://catalog.data.gov/dataset/acs-5-year-data-by-ward)	csv	51	0.008	Wyniki 5-letnich ankiet podsumowane w 2023 roku. Dane ekonomiczno-demograficzne dla Wardów w Chicago (okres 2018-2022).
3.	Chicago Police Department - Illinois Uniform Crime Reporting (IUCR) Codes	Csv	410	0.02	Opis kodów przestępstw w Illinois. Zawiera

					główny i dodatkowy opis przestępstwa.
--	--	--	--	--	---------------------------------------

5. Zakres Danych

Crimes 2001 to Present

Lp.	Atrybut	Typ danych	Zakres wartości	Uwagi – ocena jakości
1.	ID	INT	634-13837364	Dane unikalne, 0% braków, klucz główny tabeli faktów.
2.	Case Number	STRING	6-9 znaków	ID śledztwa (przydatne w przypadku paru przestępstw w jednej sprawie). 0% braków.
3.	Date	DATE	Lata 2001-2025	Dane kompletne (0% braków). Zapisane w formacie MM/DD/YYYY HH:MM:SS XM.
4.	Block	STRING	4-38 znaków	Adres blokowy, 0% braków, możliwe powtórzenia.
5.	IUCR	STRING	4 znaki	Kod klasyfikacji przestępstwa (Uniform Crime Reporting), 0% braków.
6.	Primary Type	STRING	5–33 znaków	Główna kategoria przestępstwa, np. THEFT, ASSAULT. 0% braków.
7.	Description	STRING		Opis szczegółowy zdarzenia. 0% braków.
8.	Location Description	STRING		Miejsce zdarzenia, 0.17% braków – do rozważenia imputacja lub kategoria „Unknown”.
9.	Arrest	BOOLEAN	True/False	Czy zatrzymano przestępcę (0% braków).
10.	Domestic	BOOLEAN	True/False	Czy przestępstwo miało charakter domowy (0% braków).
11.	Beat	INT	111 – 2535	Kod jednostki patrolowej (0% braków).
12.	District	FLOAT	1.0-31.0	Rejon policyjny, 0% braków, liczba rzeczywista – może wymagać konwersji na INT.

13.	Ward	FLOAT	1.0-50.0	Okręg administracyjny, 7.39% braków – należy rozważyć uzupełnienie.
14.	Community Area	FLOAT	0.0-77.0	Jednostka geograficzna, 7.38% braków – podobnie jak Ward, potrzebna imputacja.
15.	FBI CODE	STRING	2-3 znaki	Klasyfikacja FBI – 0% braków
16.	X Coordinate	Float	0.0-125119.00	Współrzędne X – 4.06% braków.
17.	Y Coordinate	Float	0.0-1951622.00	Współrzędne Y – 4.06% braków.
18.	Year	INT	2001-2025	Zbędne, można wyciągnąć z daty, 0% braków.
19.	Updated On	Date	Lata 2001-2025	Data aktualizacji rekordu (0% braków).
20.	Lattitude	Float	36.619 – 42.022	Szerokość geograficzna 4.06% braków.
21.	Longitude	Float	-91.687 – -87.525	Długość geograficzna 4.06% braków.
22.	Location	STRING	23-29 znaków	Sklejone współrzędne (lat, long). 4.06% braków

ACS_5_Year_Data_By_Ward

Lp.	Atrybut	Typ Danych	Zakres wartości	Uwagi
1.	ACS YEAR	INT	2023	Dane zagregowane w 2023 roku, zbierane w latach 2018-2022. Kolumna zbędna
2.	WARD	INT	1-50	Numer Wardu.
3.	Under \$25.000	INT	0-11000	Liczba gospodarstw domowych o łącznym rocznym przychodzie poniżej \$25.000.
4.	\$25,000 to \$49,999	INT	0-11000	Liczba gospodarstw domowych o łącznym

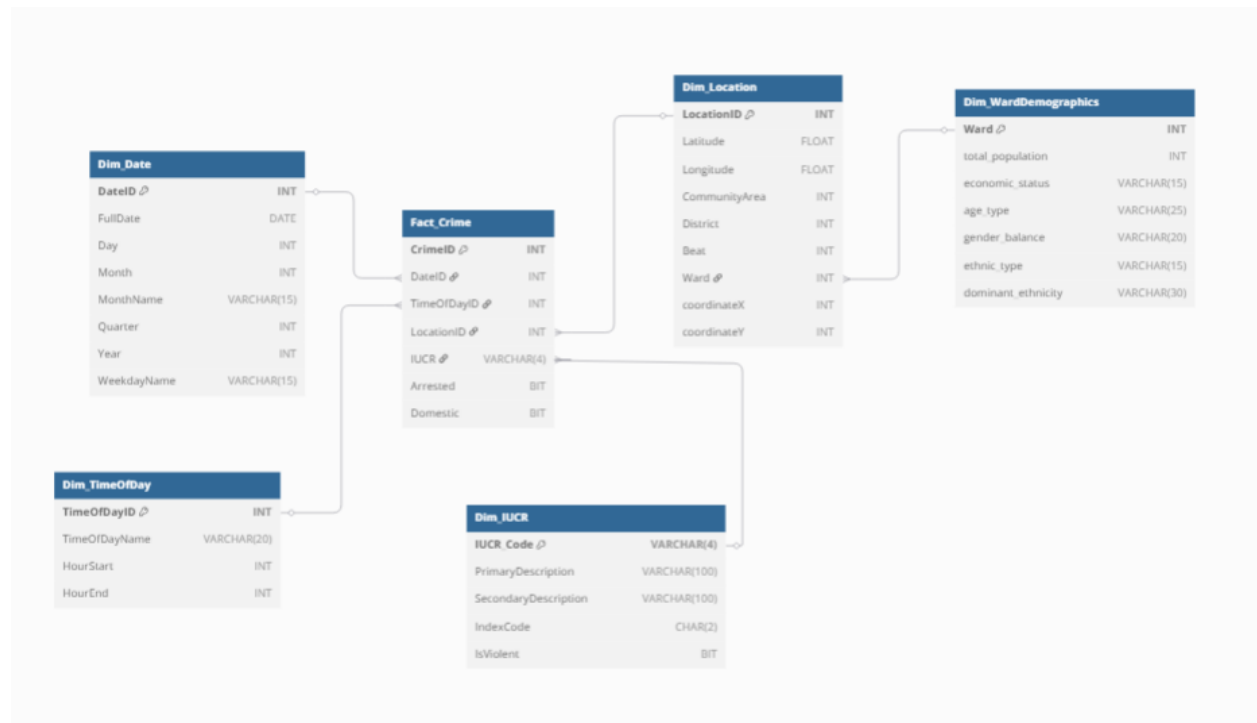
				rocznym przychodzie \$25.000- \$49.999.
5.	\$50.000 to \$74.999	INT	0-11000	Liczba gospodarstw domowych o łącznym rocznym przychodzie \$50.000- \$74.999.
6.	\$75,000 to \$125,000	INT	0-11000	Liczba gospodarstw domowych o łącznym rocznym przychodzie \$75.000- \$125.000.
7.	\$125,000+	INT	0-11000	Liczba gospodarstw domowych o łącznym rocznym przychodzie większym niż \$125,000.
8-13	Male 0 to 17, 18 to 24, 25 to 34, 35 to 49, 50 to 65, 65+	INT	0-15000	Liczba mężczyzn w danym przedziale wiekowym
14-19	Female 0 to 17, 18 to 24, 25 to 34, 35 to 49, 50 to 65, 65+	INT	0-15000	Liczba kobiet w danym przedziale wiekowym.
20.	Total Population	INT	0-75000	Liczba ludności w Wardzie.
21.-27.	White, Black or African, American Indian or Alaska Native, Asian, Native Hawaiian or	INT	0-20000	Liczba ludności z danej grupy etnicznej w Wardzie.

	Pacific Islander, Other Race, Multiracial			
28.	White not Hispanic or Latino	INT	0-50000	Liczba ludności o białym kolorze skóry, niebędąca pochodzenia latynoskiego.
29.	Hispanic or Latino	INT	0-50000	Liczba ludności pochodzenia latynoskiego.
30.	RecordID	STRING	2023_XX	Łączona wartość roku z numerem Wardu, kolumna zbędna.

IUCR Codes

Lp.	Atrybut	Typ danych	Zakres wartości	Uwagi – ocena jakości
1.	IUCR	STRING	4 znaki	Kod typu przestępstwa.
2.	Primary Description	STRING	Do 30 znaków	Główny typ przestępstwa (np. HOMICIDE, THEFT, ASSAULT)
3.	Secondary Description	STRING	Do 100 znaków.	Dokładniejszy opis przestępstwa.
4.	Index Code	STRING	I/N	Czy przestępstwo należy do „Index Crimes” (czy ścigane przez FBI).
5.	Active	BOOLEAN	True/False	Czy typ przestępstwa jest dalej aktywny.

6. Definicja Typów Encji/Klas



Miary addytywne (kalkulowane):

- ArrestedCount
- DomesticCount
- CrimeCount

Miary nieaddytywne (kalkulowane):

- CrimesPer1000Inhabitants
- PercentArrested

7. Przykładowe analizy wielowymiarowe

1. Liczba przestępstw wg dominującej grupy etnicznej

- **Wymiary:** dominant_ethnicity, PrimaryDescription (z Dim_IUCR)
- **Miara:** CrimeCount
- **Cel:** Zidentyfikowanie różnic kulturowych/systemowych w rozkładzie przestępstw w zależności od dominującej grupy etnicznej w Wardzie.

2. Liczba i typ przestępstw w zależności od statusu ekonomicznego

- **Wymiary:** economic_status, PrimaryDescription
- **Miary:** CrimeCount, ArrestCount
- **Cel:** Zbadanie powiązań między statusem majątkowym dzielnic a typami przestępstw i skutecznością aresztowań.

3. Dominujące typy przestępstw w Wardach o młodej populacji

- **Wymiary:** age_type, PrimaryDescription
- **Miara:** CrimeCount
- **Cel:** Ocena zagrożeń i dostosowanie programów prewencyjnych dla młodzieży.

4. Skuteczność aresztowań względem statusu ekonomicznego

- **Wymiary:** economic_status, PrimaryDescription
- **Miara:** ArrestRate (obliczany: aresztowane / wszystkie)
- **Cel:** Analiza nierówności w egzekwowaniu prawa między bogatszymi a biedniejszymi Wardami.

5. Wskaźnik przestępstw domowych a profil wiekowy i etniczny

- **Wymiary:** age_type, dominant_ethnicity
- **Miara:** DomesticCrimeRate (udział przestępstw domowych)
- **Cel:** Ujawnienie potencjalnych korelacji między wiekiem i pochodzeniem a przemocą domową.

6. Trend przestępczości w najbiedniejszych obszarach

- **Wymiary:** Year, Ward
- **Filtr:** economic_status = 'poor'

- **Miara:** CrimeCount
 - **Cel:** Monitorowanie zmian w najbardziej zagrożonych ekonomicznie dzielnicach.
-

7. Liczba przestępstw wg dnia tygodnia i lokalizacji

- **Wymiary:** WeekdayName, Latitude, Longitude (z Dim_Location)
 - **Miara:** CrimeCount
 - **Cel:** Analiza czasowo-przestrzenna (np. do heatmapy w Power BI).
-

8. Liczba przestępstw wg typu i rezultatu (areszt)

- **Wymiary:** PrimaryDescription, Arrested
 - **Miara:** CrimeCount
 - **Cel:** Porównanie skuteczności zatrzymań między typami przestępstw.
-

9. Skuteczność aresztowań wg grup etnicznych

- **Wymiary:** dominant_ethnicity, Year
 - **Miara:** ArrestRate
 - **Cel:** Badanie możliwych różnic w podejściu służb do różnych społeczności.
-

10. Przestępstwa w zależności od pory dnia i wieku mieszkańców

- **Wymiary:** age_type, TimeOfDayName
- **Miara:** CrimeCount
- **Cel:** Określenie, kiedy i gdzie młode społeczności są najbardziej narażone.

8. Utworzenie Bazy Danych

```
-- Dim_Date
CREATE TABLE Projekt.Dim_Date (
    DateID INT,
    FullDate DATE,
    Day INT,
    Month INT,
```

```

    MonthName VARCHAR(15),
    Quarter INT,
    Year INT,
    WeekdayName VARCHAR(15)
);

-- Dim_TimeOfDay
CREATE TABLE Projekt.Dim_TimeOfDay (
    TimeOfDayID INT,
    TimeOfDayName VARCHAR(20), -- Night, Morning, Afternoon, Evening
    HourStart INT,
    HourEnd INT
);

-- Dim_WardDemographics
CREATE TABLE Projekt.Dim_WardDemographics (
    Ward INT,
    total_population INT,
    economic_status VARCHAR(15), -- poor, average, rich, extreme
    age_type VARCHAR(25), -- young-majority, balanced, elderly-majority
    gender_balance VARCHAR(20), -- male-dominant, mixed, female-dominant
    ethnic_type VARCHAR(15), -- mixed, dominant
    dominant_ethnicity VARCHAR(30) -- np. Black, White, Hispanic, etc.
);

-- Dim_Location
CREATE TABLE Projekt.Dim_Location (
    LocationID INT IDENTITY(1,1) PRIMARY KEY,
    Latitude FLOAT,
    Longitude FLOAT,
    CommunityArea INT,
    District INT,
    Beat INT,
    Ward INT,
    coordinateX INT,
    coordinateY INT
);

-- Dim_IUCR
CREATE TABLE Projekt.Dim_IUCR (
    IUCR_Code VARCHAR(4),
    PrimaryDescription VARCHAR(100),
    SecondaryDescription VARCHAR(100),
    IndexCode CHAR(2), --I, N
    IsViolent BIT
);

-- Fact_Crime
CREATE TABLE Projekt.Fact_Crime (
    CrimeID INT,
    DateID INT,
    TimeOfDayID INT,
    LocationID INT,
    IUCR VARCHAR(4),
    Arrested BIT,
    Domestic BIT
);

```

Etap 2 – Proces ETL

1. Przygotowanie Danych Źródłowych

Jako, że analiza dotyczy między innymi czynników socjoekonomicznych zbieranych w latach 2018-2022 pierwszym etapem było przefiltrowanie danych źródłowych.

Wykorzystałem do tego bibliotekę skrypt w języku Python (biblioteka Pandas).

```
import pandas as pd

df = pd.read_csv("Crimes_-_2001_to_Present.csv", parse_dates=['Date'],
low_memory=False)

df_filtered = df[df['Date'].dt.year.between(2018, 2022)]

df_filtered.to_csv("filtered_crimes.csv", index=False)
```

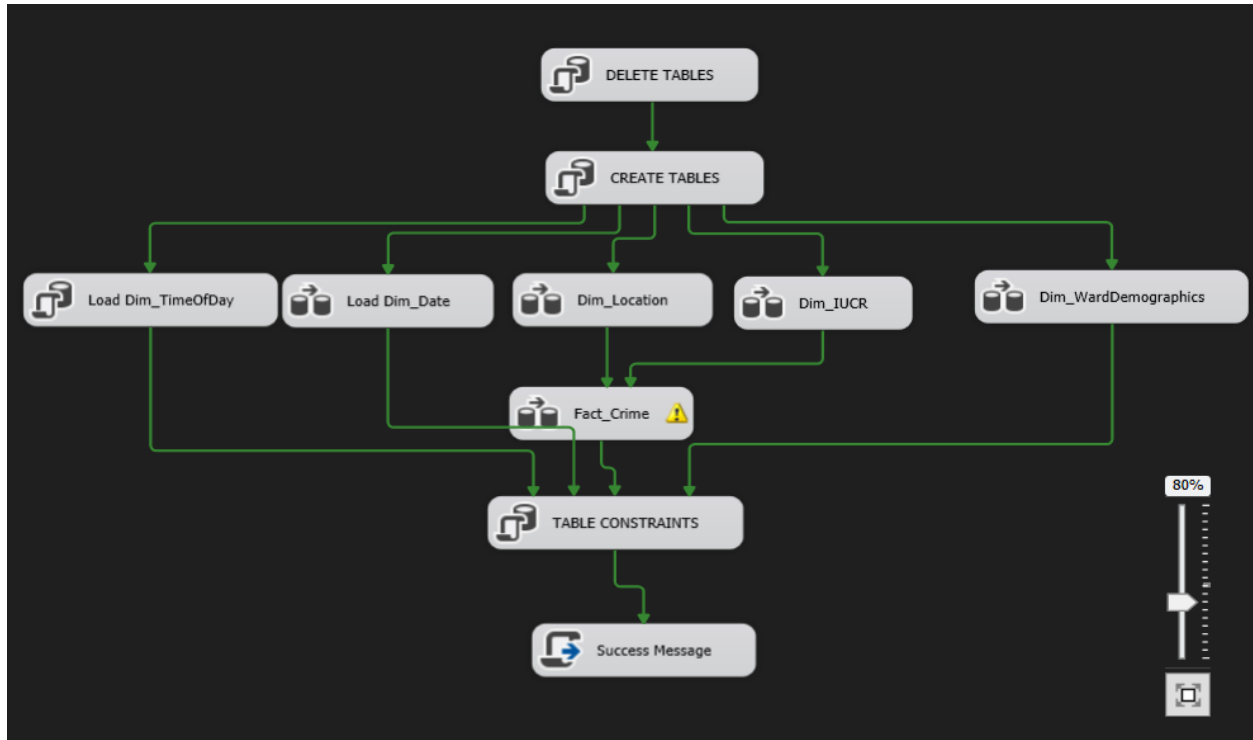
2. Architektura Hurtowni Danych

W hurtowni danych zaimplementowano następujące tabele:

- **Fakt:** Fact_Crime
- **Wymiary:** Dim_Date, Dim_TimeOfDay, Dim_Location, Dim_IUCR, Dim_WardDemographics

3. Struktura procesu ETL (SSIS)

Proces ETL podzielony został na **modularne pakiety SSIS**. Każdy pakiet realizuje określoną funkcję. Wszystkie pakiety uruchamiane są sekwencyjnie i zapisane zostały w rozwiązaniu SSIS „Projekt_ETL”.



1. Delete Tables

Opis:

To zadanie odpowiada za czyszczenie danych z tabel docelowych w hurtowni danych, aby proces ETL mógł zostać przeprowadzony od zera (pełne załadowanie danych).

Typ zadania: Execute SQL Task

Kod SQL:

```
IF EXISTS (  
    SELECT *  
    FROM INFORMATION_SCHEMA.TABLES
```



```

        WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Fact_Crime'
    )
DROP TABLE Projekt.Fact_Crime;
GO

IF EXISTS (
    SELECT *
    FROM INFORMATION_SCHEMA.TABLES
    WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Dim_Date'
)
DROP TABLE Projekt.Dim_Date;
GO

IF EXISTS (
    SELECT *
    FROM INFORMATION_SCHEMA.TABLES
    WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Dim_TimeOfDay'
)
DROP TABLE Projekt.Dim_TimeOfDay;
GO

IF EXISTS (
    SELECT *
    FROM INFORMATION_SCHEMA.TABLES
    WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Dim_Location'
)
DROP TABLE Projekt.Dim_Location;
GO

IF EXISTS (
    SELECT *
    FROM INFORMATION_SCHEMA.TABLES
    WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Dim_WardDemographics'
)
DROP TABLE Projekt.Dim_WardDemographics;
GO

IF EXISTS (
    SELECT *
    FROM INFORMATION_SCHEMA.TABLES
    WHERE TABLE_SCHEMA = 'Projekt' AND TABLE_NAME = 'Dim_IUCR'
)
DROP TABLE Projekt.Dim_IUCR;
GO

```

2. Create Tables

Opis:

Zadanie tworzy strukturę tabel faktów i wymiarów w bazie danych Projekt, jeśli nie została jeszcze utworzona. W praktyce może być wykonywane tylko raz — podczas inicjalizacji projektu.

Typ zadania: Execute SQL Task

SQL – struktury tabel:

```
-- Dim_Date
CREATE TABLE Projekt.Dim_Date (
    DateID INT,
    FullDate DATE,
    Day INT,
    Month INT,
    MonthName VARCHAR(15),
    Quarter INT,
    Year INT,
    WeekdayName VARCHAR(15)
);

-- Dim_TimeOfDay
CREATE TABLE Projekt.Dim_TimeOfDay (
    TimeOfDayID INT,
    TimeOfDayName VARCHAR(20), -- Noc, Rano, Popołudnie, Wieczór
    HourStart INT,
    HourEnd INT
);

-- Dim_WardDemographics
CREATE TABLE Projekt.Dim_WardDemographics (
    Ward INT,
    total_population INT,
    economic_status VARCHAR(15), -- poor, average, rich, extreme
    age_type VARCHAR(25), -- young-majority, balanced, elderly-majority
    gender_balance VARCHAR(20), -- male-dominant, mixed, female-dominant
    ethnic_type VARCHAR(15), -- mixed, dominant
    dominant_ethnicity VARCHAR(30) -- np. Black, White, Hispanic, etc.
);

-- Dim_Location
CREATE TABLE Projekt.Dim_Location (
    LocationID INT IDENTITY(1,1) PRIMARY KEY,
    Latitude FLOAT,
    Longitude FLOAT,
    CommunityArea INT,
    District INT,
    Beat INT,
    Ward INT,
    coordinateX INT,
    coordinateY INT
);

-- Dim_IUCR
CREATE TABLE Projekt.Dim_IUCR (
    IUCR_Code VARCHAR(4),
    PrimaryDescription VARCHAR(100),
    SecondaryDescription VARCHAR(100),
    IndexCode CHAR(2), -- np. I, II, N
    IsViolent BIT
);

-- Fact_Crime
CREATE TABLE Projekt.Fact_Crime (
    CrimeID INT,
```

```

DateID INT,
TimeOfDayID INT,
LocationID INT,
IUCR VARCHAR(4),
Arrested BIT,
Domestic BIT
);

```

3. Load Dim_TimeOfDay

Cel: Ręczne wypełnienie tabeli Dim_TimeOfDay.

Opis: Dane statyczne – 4 pory dnia: Noc (0–6), Rano (6–12), Popołudnie (12–18), Wieczór (18–24). Wprowadzono jako Insert T-SQL.

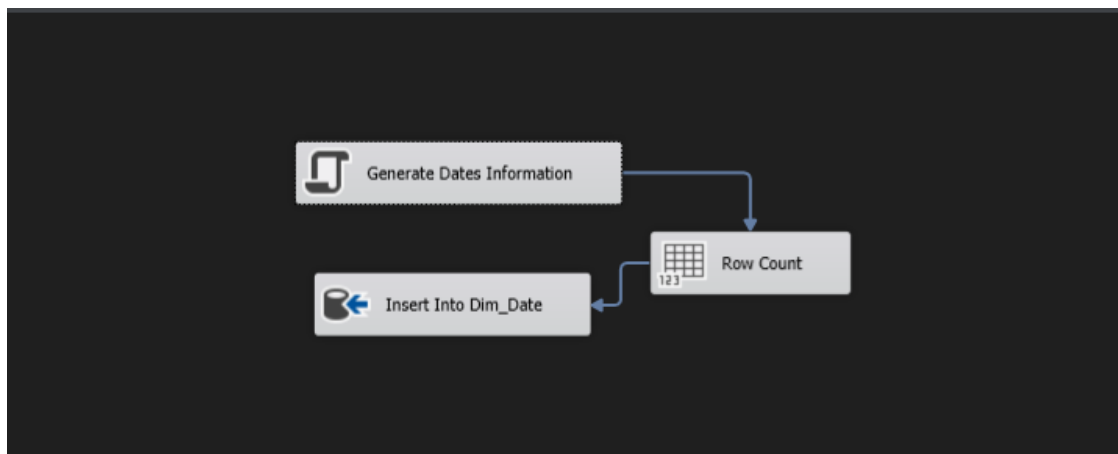
SQL – wypełnianie danych:

```

INSERT INTO Projekt.Dim_TimeOfDay (TimeOfDayID, TimeOfDayName, HourStart, HourEnd)
VALUES
(1, 'Morning', 6, 11),
(2, 'Afternoon', 12, 17),
(3, 'Evening', 18, 23),
(4, 'Night', 0, 5);

```

4. Load Dim_Date



Cel: Wypełnienie tabeli Dim_Date zakresem dat występujących w zbiorze.

Opis: Zastosowano skrypt C# generujący daty w zakresie 2018–2022 wraz z atrybutami daty (dzień, miesiąc, nazwa dnia, kwartał, itd.).

Skrypt C#:

```

DateTime startDate = new DateTime(2018, 1, 1);
DateTime endDate = new DateTime(2022, 12, 31);

```

```

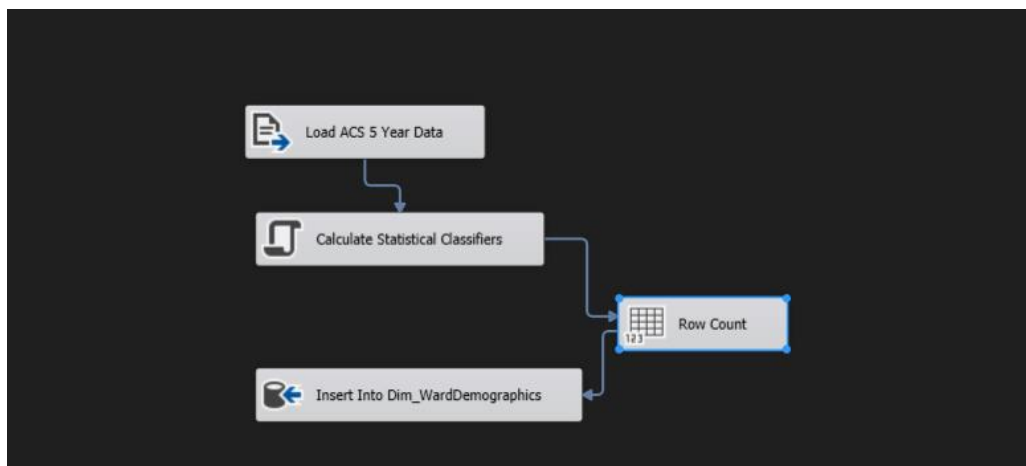
for (DateTime date = startDate; date <= endDate; date = date.AddDays(1))
{
    Output0Buffer.AddRow();

    Output0Buffer.FullDate = date;
    Output0Buffer.Day = date.Day;
    Output0Buffer.Month = date.Month;
    Output0Buffer.MonthName = date.ToString("MMMM",
CultureInfo.InvariantCulture);
    Output0Buffer.Quarter = (date.Month - 1) / 3 + 1;
    Output0Buffer.Year = date.Year;
    Output0Buffer.WeekdayName = date.ToString("dddd",
CultureInfo.InvariantCulture);

    // DateID in format YYYYMMDD
    Output0Buffer.DateID = int.Parse(date.ToString("yyyymmdd"));
}
}

```

5. Load Dim_WardDemographics



Źródło: ACS 5 Year Data by Ward

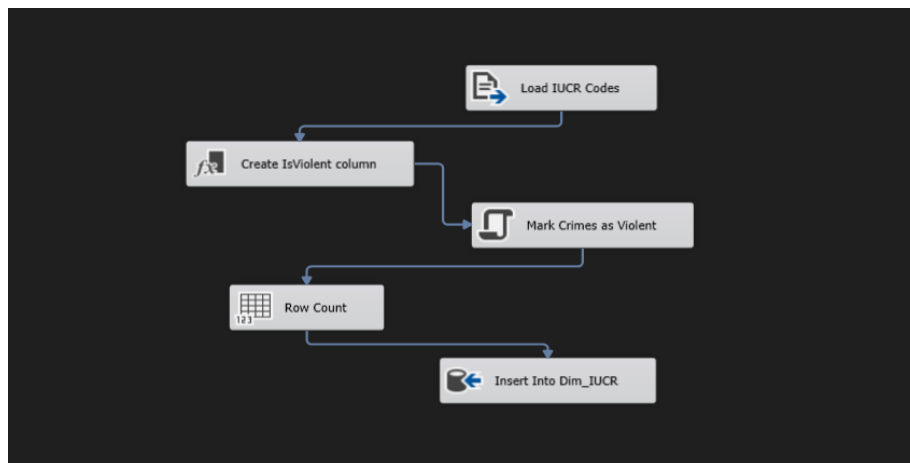
Cel: Wczytanie i transformacja danych socjodemograficznych do Dim_WardDemographics

Opis: Dane sklasyfikowane ręcznie na potrzeby projektu, zawiera klasyfikacje:

- economic_status:
 - „extreme” – jeśli przychód ponad 40% gospodarstw domowych wynosi poniżej \$50,000 oraz przychód ponad 40% gospodarstw domowych wynosi powyżej \$125,000
 - “poor” – jeśli mediana przychodu na gospodarstwo domowe wynosi poniżej \$50,000
 - „rich” – jeśli mediana przychodu na gospodarstwo domowe wynosi powyżej \$125,000

- „average” – w reszcie przypadków
- age_type
 - young-dominant – jeśli ponad 55% mieszkańców wardu ma poniżej 35 roku życia
 - elderly-dominant – jeśli ponad 35% mieszkańców wardu ma powyżej 50 roku życia
 - balanced – w reszcie przypadków
- gender_balance
 - male-dominant – jeśli ponad 55% mieszkańców wardu to mężczyźni
 - female-dominant – jeśli ponad 55% mieszkańców wardu to kobiety
 - mixed – w reszcie przypadków
- ethnic_type
 - dominant – jeśli jedna grupa etniczna stanowi ponad 50% mieszkańców wardu
 - mixed – w przeciwnym przypadku
- dominant_ethnicity – najlicniejsza grupa etniczna w wardzie (np. White, Hispanic, Black)

6. Load Dim_IUCR



Źródło: Chicago Police Department IUCR Codes

Cel: Wczytanie kodów policyjnych IUCR oraz klasyfikacja przestępstw jako przestępstwa z wykorzystaniem przemocy.

Opis: Przestępstwa z wykorzystaniem przemocy klasyfikujemy jako:

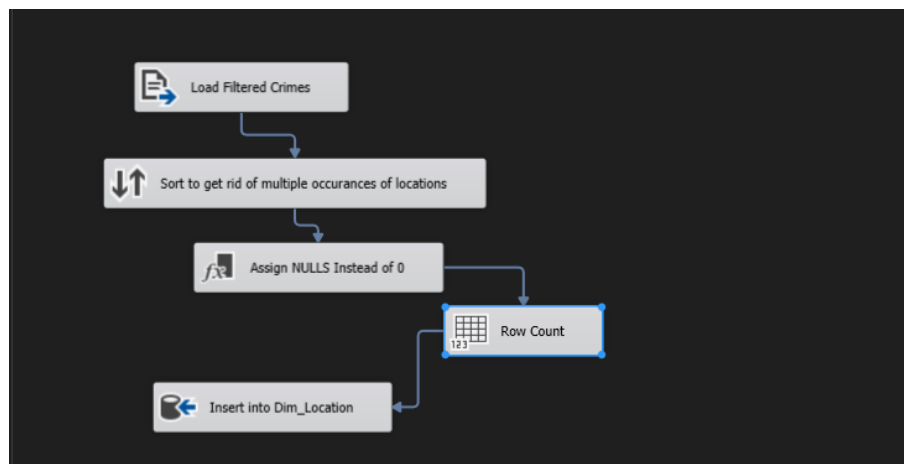
- Zabójstwo

- Kryminalne napady seksualne
- Rozbój
- Pobicie
- Przestępstwa na tle rytualnym
- Napaść
- Handel Ludźmi
- Przestępstwo seksualne (seksualne wykorzystanie dziecka/kryminalnego nadużycie seksualne/kryminalne przenoszenie HIV/próba zastrzonego kryminalnego nadużycia seksualnego)
- Przestępstwo z udziałem dzieci (pornografia dziecięca/kryminalne nadużycie seksualne przez członka rodziny/napaść na tle seksualnym przez członka rodziny)

7. Load Dim_Location

Źródło: Filtered Crimes

Cel: Załadowanie do tabeli wymiarowej Dim_Location unikalnych kombinacji współrzędnych geograficznych i danych terytorialnych (np. Community Area, Beat, Ward), co pozwala analizować przestępczość z perspektywy przestrzennej.



Opis:

Z powodu braku możliwości Lookup'u dla wartości typu float (Longitude i Latitude), zachowujemy poniższe dane o lokalizacji, które pozwalają nam unikalnie zidentyfikować lokalizację w momencie Lookupu w procesie ładowania danych Fact_Crimes:

- Community Area
- District
- Beat
- Ward

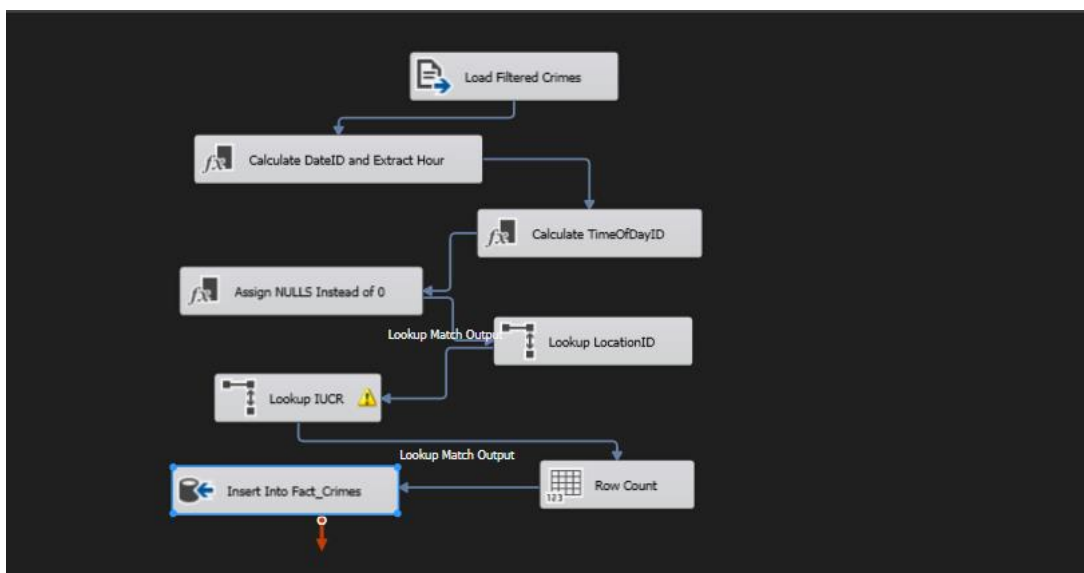
- X Coordinate
- Y Coordinate

Oraz dane które pozwolą nam na analizę przestrzenną przestępstw:

- Longitude
- Latitude

Jeśli w którejś z kolumn występuje 0 (niemożliwe w danych dla Chicago) – zamieniamy tę wartość na NULL.

8. Load Fact_Crime



Cel: Załadowanie tabeli faktów Fact_Crime na podstawie danych z pliku źródłowego, połączonych z wymiarami Dim_Date, Dim_TimeOfDay, Dim_Location, Dim_IUCR.

Etapy przetwarzania (w SSIS - Data Flow):

1. Wczytanie danych z pliku CSV

- **Źródło:** Plik CSV z przestępstwami (np. Crimes_2018_2022.csv)
- **Zadanie:** Flat File Source
- **Opis:** Wczytuje wszystkie rekordy z surowego zbioru danych, zawierające m.in. datę, godzinę, lokalizację, aresztowanie, itp.

2. Derived Column – wyodrębnienie DateID i Hour

- **Zadanie:** Derived Column

- **Opis:** Ekstrahuje datę i godzinę z pola Date w formacie tekstowym i tworzy:
 - DateID w formacie yyyymmdd
 - Hour jako liczba od 0 do 23

3. Derived Column – wyznaczenie TimeOfDayID na podstawie Hour

- **Zadanie:** Derived Column
- **Opis:** Dla każdej godziny przypisywany jest odpowiedni TimeOfDayID:
 - 1: Morning (6–11)
 - 2: Afternoon (12–17)
 - 3: Evening (18–23)
 - 4: Night (0–5)

4. Derived Column – czyszczenie błędnych lokalizacji (0 -> NULL)

- **Zadanie:** Kolejny komponent **Derived Column**
- **Opis:** Zamiana wartości 0 w polach lokalizacyjnych na NULL, które nie mają znaczenia analitycznego.

5. Lookup – pobranie LocationID z Dim_Location

- **Zadanie:** Lookup
- **Opis:** Dopasowanie lokalizacji do wymiaru Dim_Location na podstawie:
 - Latitude, Longitude, CommunityArea, District, Ward, Beat, coordinateX, coordinateY

6. Lookup – sprawdzenie czy kod występuje w IUCR

- **Zadanie:** Lookup
- **Opis:** Dopasowanie kodu IUCR do istniejących w tabeli Dim_IUCR. Jeśli dopasowanie nie istnieje, wiersz jest odrzucany.

7. Wybór i mapowanie danych do Fact_Crime

1. **Output Columns:**
 - CrimeID (z pliku)
 - DateID (z Derived Column)

- TimeOfDayID (z Derived Column)
- LocationID (z Lookup)
- IUCR (z pliku)
- Arrested (z pliku)
- Domestic (z pliku)

9. Constraints (References)

Cel: Nadanie poprawnych ograniczeń (PK, FK) tabelom wypełnionym danymi.

Zadanie: Execute SQL Task

SQL:

```
-- PRIMARY KEYS

-- Dim_Date
IF NOT EXISTS (
    SELECT * FROM sys.key_constraints
    WHERE name = 'PK_Dim_Date' AND parent_object_id = OBJECT_ID('Projekt.Dim_Date')
)
ALTER TABLE Projekt.Dim_Date
ADD CONSTRAINT PK_Dim_Date PRIMARY KEY (DateID);

-- Dim_TimeOfDay
IF NOT EXISTS (
    SELECT * FROM sys.key_constraints
    WHERE name = 'PK_Dim_TimeOfDay' AND parent_object_id =
OBJECT_ID('Projekt.Dim_TimeOfDay')
)
ALTER TABLE Projekt.Dim_TimeOfDay
ADD CONSTRAINT PK_Dim_TimeOfDay PRIMARY KEY (TimeOfDayID);

-- Dim_WardDemographics
IF NOT EXISTS (
    SELECT * FROM sys.key_constraints
    WHERE name = 'PK_Dim_WardDemographics' AND parent_object_id =
OBJECT_ID('Projekt.Dim_WardDemographics')
)
ALTER TABLE Projekt.Dim_WardDemographics
ADD CONSTRAINT PK_Dim_WardDemographics PRIMARY KEY (Ward);

-- Dim_IUCR
IF NOT EXISTS (
    SELECT * FROM sys.key_constraints
    WHERE name = 'PK_Dim_IUCR' AND parent_object_id = OBJECT_ID('Projekt.Dim_IUCR')
)
ALTER TABLE Projekt.Dim_IUCR
ADD CONSTRAINT PK_Dim_IUCR PRIMARY KEY (IUCR_Code);

-- Fact_Crime
IF NOT EXISTS (
```

```

        SELECT * FROM sys.key_constraints
        WHERE name = 'PK_Fact_Crime' AND parent_object_id = OBJECT_ID('Projekt.Fact_Crime')
    )
    ALTER TABLE Projekt.Fact_Crime
    ADD CONSTRAINT PK_Fact_Crime PRIMARY KEY (CrimeID);

-- FOREIGN KEYS

-- Fact_Crime.DateID → Dim_Date
IF NOT EXISTS (
    SELECT * FROM sys.foreign_keys
    WHERE name = 'FK_Fact_Crime_Date'
)
    ALTER TABLE Projekt.Fact_Crime
    ADD CONSTRAINT FK_Fact_Crime_Date
    FOREIGN KEY (DateID) REFERENCES Projekt.Dim_Date(DateID);

-- Fact_Crime.TimeOfDayID → Dim_TimeOfDay
IF NOT EXISTS (
    SELECT * FROM sys.foreign_keys
    WHERE name = 'FK_Fact_Crime_TimeOfDay'
)
    ALTER TABLE Projekt.Fact_Crime
    ADD CONSTRAINT FK_Fact_Crime_TimeOfDay
    FOREIGN KEY (TimeOfDayID) REFERENCES Projekt.Dim_TimeOfDay(TimeOfDayID);

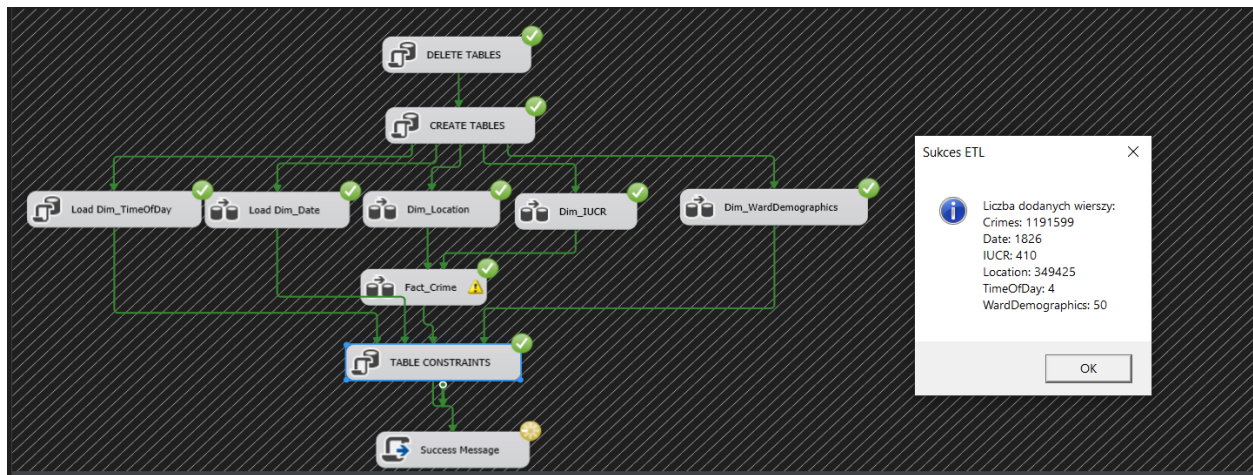
-- Fact_Crime.LocationID → Dim_Location
IF NOT EXISTS (
    SELECT * FROM sys.foreign_keys
    WHERE name = 'FK_Fact_Crime_Location'
)
    ALTER TABLE Projekt.Fact_Crime
    ADD CONSTRAINT FK_Fact_Crime_Location
    FOREIGN KEY (LocationID) REFERENCES Projekt.Dim_Location(LocationID);

-- Fact_Crime.IUCR → Dim_IUCR
IF NOT EXISTS (
    SELECT * FROM sys.foreign_keys
    WHERE name = 'FK_Fact_Crime_IUCR'
)
    ALTER TABLE Projekt.Fact_Crime
    ADD CONSTRAINT FK_Fact_Crime_IUCR
    FOREIGN KEY (IUCR) REFERENCES Projekt.Dim_IUCR(IUCR_Code);

-- Dim_Location.Ward → Dim_WardDemographics.Ward
IF NOT EXISTS (
    SELECT * FROM sys.foreign_keys
    WHERE name = 'FK_Location_WardDemographics'
)
    ALTER TABLE Projekt.Dim_Location
    ADD CONSTRAINT FK_Location_WardDemographics
    FOREIGN KEY (Ward) REFERENCES Projekt.Dim_WardDemographics(Ward);

```

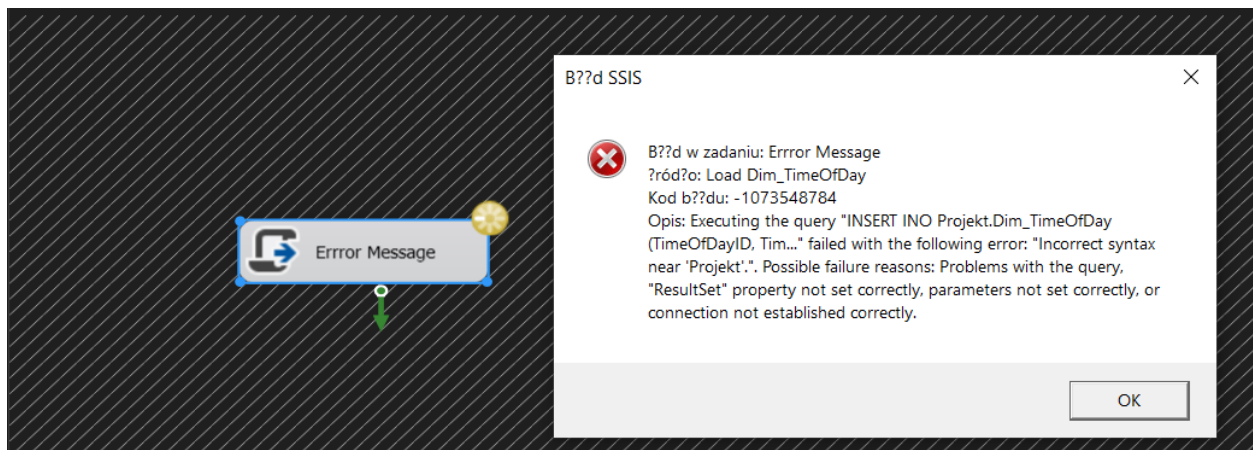
10. Success Message



Cel: Powiadomienie użytkownika o poprawnym zakończeniu procesu ETL oraz wypisanie liczby wierszy w każdej z tabel.

Opis: Skrypt w C#, który wykorzystuje zmienne SSIS odpowiadające liczbie wierszy (wypełnione za pomocą RowCount).

11. Error Handling



Cel: Poinformowanie użytkownika o błędzie w procesie ETL oraz wypisanie kodu błędu i opisu błędu.

Opis: Skrypt w C# który wykorzystuje systemowe zmienne dotyczące Error Handlingu.

Etap 3 – Kostka OLAP w SSAS

1. Opis Celu Etapu

Celem etapu trzeciego było zbudowanie kostki OLAP w SQL Server Analysis Services (SSAS), która umożliwi wielowymiarową analizę przestępczości w Chicago w latach 2018–2022 z uwzględnieniem danych społeczno-ekonomicznych. Kostka została zbudowana w oparciu o dane z wcześniej przygotowanej hurtowni, zasilanej przez proces ETL zrealizowany w SSIS.

Zadaniem kostki jest umożliwienie analiz m.in.:

- typów przestępstw,
- efektywności aresztowań,
- przestępczości w kontekście warunków demograficznych,
- zależności przestrzennych i czasowych.

2. Struktura Kostki

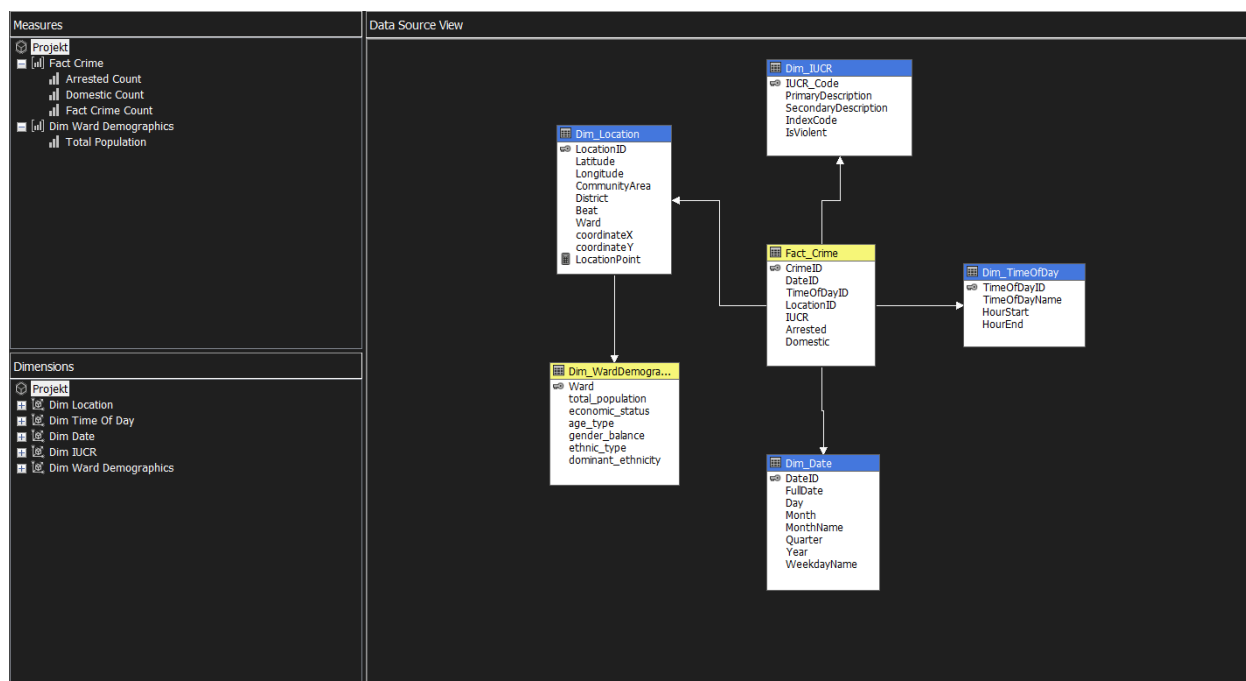


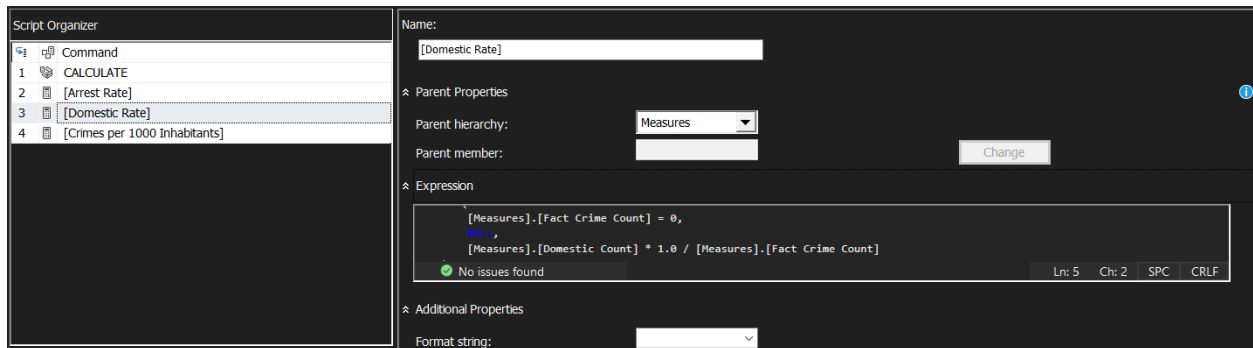
Tabela faktów:

- Fact_Crime – zawiera informacje o pojedynczych przypadkach przestępstw.

Miary (fakty) - addytywne:

- Crime Count – liczba wszystkich przestępstw (ilość rekordów).
- Arrested Count – liczba przestępstw zakończonych aresztowaniem.
- Domestic Count – liczba przestępstw o charakterze domowym.

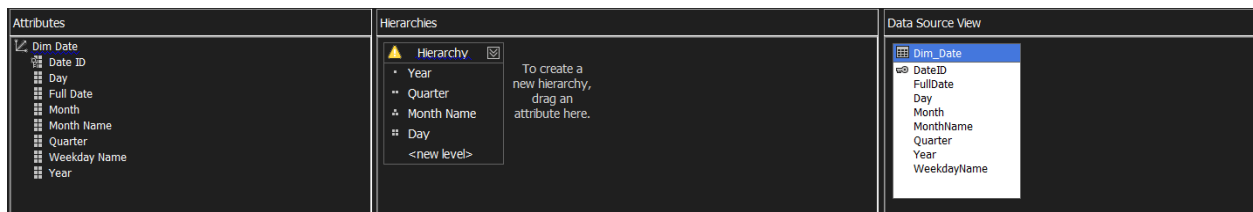
Miary Kalkulowalne:



- Arrest Rate – stosunek Arrested Count do Crime Count
- Domestic Rate – stosunek Domestic Count do Crime Count
- Crimes per 1000 Inhabitants – miara obliczana jako (Crime Count / populacja wardu) * 1000 (nieaddytywna).

3. Wymiary (dimensions)

Dim_Date:



- Atrybuty:
 - DateID - unikalny identyfikator daty,
 - FullDate - pełna data (format RRRR-MM-DD),
 - Year, Quarter, Month, Day - liczby określające jednostki czasu,
 - MonthName - nazwa miesiąca,
 - WeekdayName - dzień tygodnia (np. Monday),
- Hierarchia: Year → Quarter → Month → Day

Dim_TimeOfDay

Attributes	Hierarchies	Data Source View
<ul style="list-style-type: none">Dim Time Of Day<ul style="list-style-type: none">Time Of Day IDTime Of Day Name	<p>To create a new hierarchy, drag an attribute here.</p>	<ul style="list-style-type: none">Dim_TimeOfDay<ul style="list-style-type: none">TimeOfDayIDTimeOfDayNameHourStartHourEnd

- Atrybuty:
 - TimeOfDayID - identyfikator pory dnia
 - TimeOfDayName – nazwa pory dnia (np. Night, Morning)
- Brak hierarchii (wymiar płaski)

Dim_Location

Attributes	Hierarchies	Data Source View
<ul style="list-style-type: none">Dim Location<ul style="list-style-type: none">BeatCommunity AreaDistrictLocation IDLocation PointWard	<p>Police Geography</p> <ul style="list-style-type: none">District<ul style="list-style-type: none">Beat<new level> <p>To create a new hierarchy, drag an attribute here.</p>	<ul style="list-style-type: none">Dim_Location<ul style="list-style-type: none">LocationIDLatitudeLongitudeCommunityAreaDistrictBeatWardcoordinateXcoordinateYLocationPoint

- Atrybuty:
 - Ward, District, Beat, CommunityArea – jednostki podziału administracyjnego miasta
 - RoundedLatitude, RoundedLongitude – współrzędne zaokrąglone do 4 miejsc po przecinku – używane do agregacji na mapach
- Hierarchie:
 - Police Geography:** District → Beat

Dim_WardDemographics

Attributes	Hierarchies	Data Source View
<ul style="list-style-type: none">Dim Ward Demographics<ul style="list-style-type: none">Age TypeDominant EthnicityEconomic StatusEthnic TypeGender BalanceTotal PopulationWard	<p>To create a new hierarchy, drag an attribute here.</p>	<ul style="list-style-type: none">Dim_WardDemograp...<ul style="list-style-type: none">Ward<ul style="list-style-type: none">total_populationeconomic_statusage_typegender_balanceethnic_typedominant_ethnicity

- Atrybuty:
 - economic_status – klasyfikacja wardu na biedny / średni / bogaty
 - age_type – dominująca grupa wiekowa (np. young-majority)
 - gender_balance – równowaga płci

- ethnic_type – typ etniczny (mieszany / dominujący)
 - dominant_ethnicity – dominująca grupa etniczna
 - total_population – liczba mieszkańców wardu
- Brak hierarchii – dane opisowe do analiz korelacyjnych

Dim_IUCR

Attributes	Hierarchies	Data Source View
<ul style="list-style-type: none"> Dim_IUCR <ul style="list-style-type: none"> Index Code Is Violent IUCR Code Primary Description Secondary Description 	<p>Hierarchy</p> <ul style="list-style-type: none"> Primary Description Secondary Description <new level> <p>To create a new hierarchy, drag an attribute here.</p>	<ul style="list-style-type: none"> Dim_IUCR <ul style="list-style-type: none"> IUCR_Code PrimaryDescription SecondaryDescription IndexCode IsViolent

- Atrybuty:
 - PrimaryDescription, SecondaryDescription – podstawowa i szczegółowa klasyfikacja przestępstwa
 - IndexCode – kod raportowy
 - IsViolent – flaga przestępstw z użyciem przemocy
 - CrimeCategory – kategoryzacja nadrzędna (np. Property Crime, Violent Crime)
 - SeverityWeight – waga powagi przestępstwa, przygotowana na podstawie PrimaryDescription, w celu obliczenia wskaźników ważonych (np. KPI)
- Hierarchia: CrimeCategory -> PrimaryDescription → SecondaryDescription

Dim_CrimeFlags (degenerate dimension):

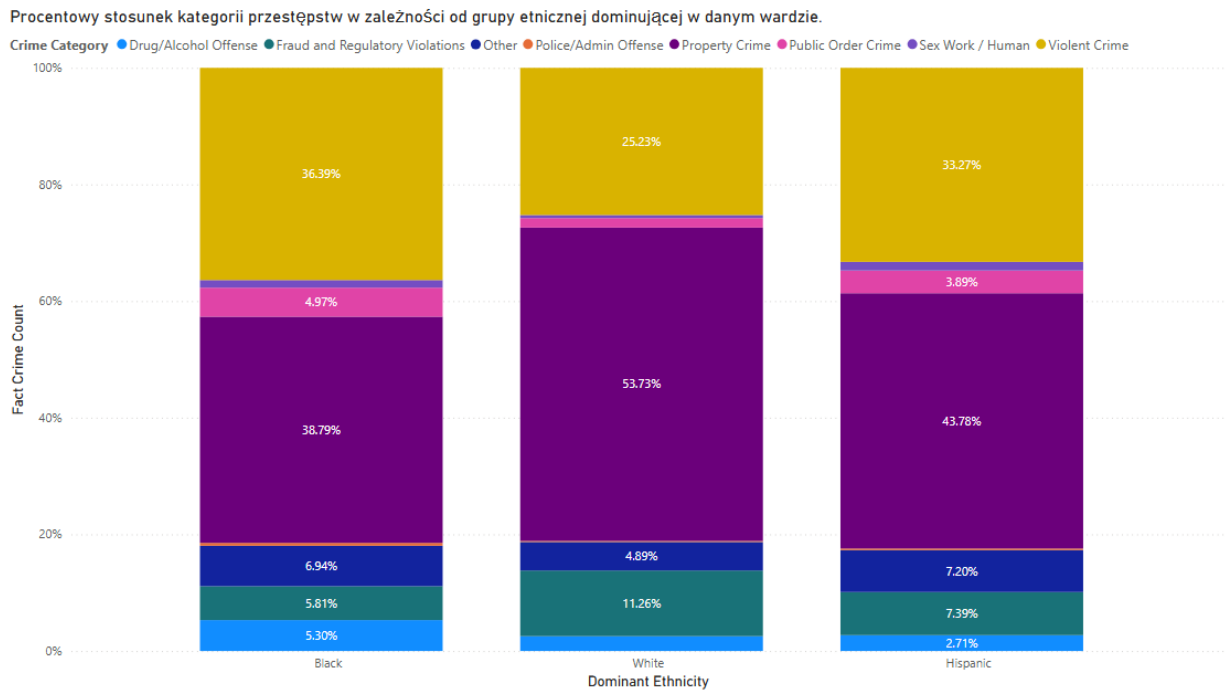
Hierarchies						Data Source View	
Attributes							
	Name	Usage	Type	Key Column	Name Column		
	Arrested	Regular	Regular	Integer	Same as key		
	Crime ID	Key	Regular	Integer	Same as key		
	Domestic	Regular	Regular	Integer	Same as key		
	To create a new attribut...						

<ul style="list-style-type: none"> Fact_Crime <ul style="list-style-type: none"> CrimeID DateID TimeOfDayID LocationID IUCR Arrested Domestic
--

- Atrybuty:
 - Arrested – flaga aresztowania (True/False)
 - Domestic – flaga przestępstwa domowego (True/False)
- Uwagi: Wymiar ten jest tzw. *degenerate dimension* – jego atrybuty pochodzą bezpośrednio z tabeli faktów, ale są odseparowane dla potrzeb analitycznych i czytelności modelu.

4. Analizy – Etap 3 (realizacja punktu 7 Etapu 1)

Analiza 1: Rozkład kategorii przestępstw wg dominującej grupy etnicznej



Wnioski:

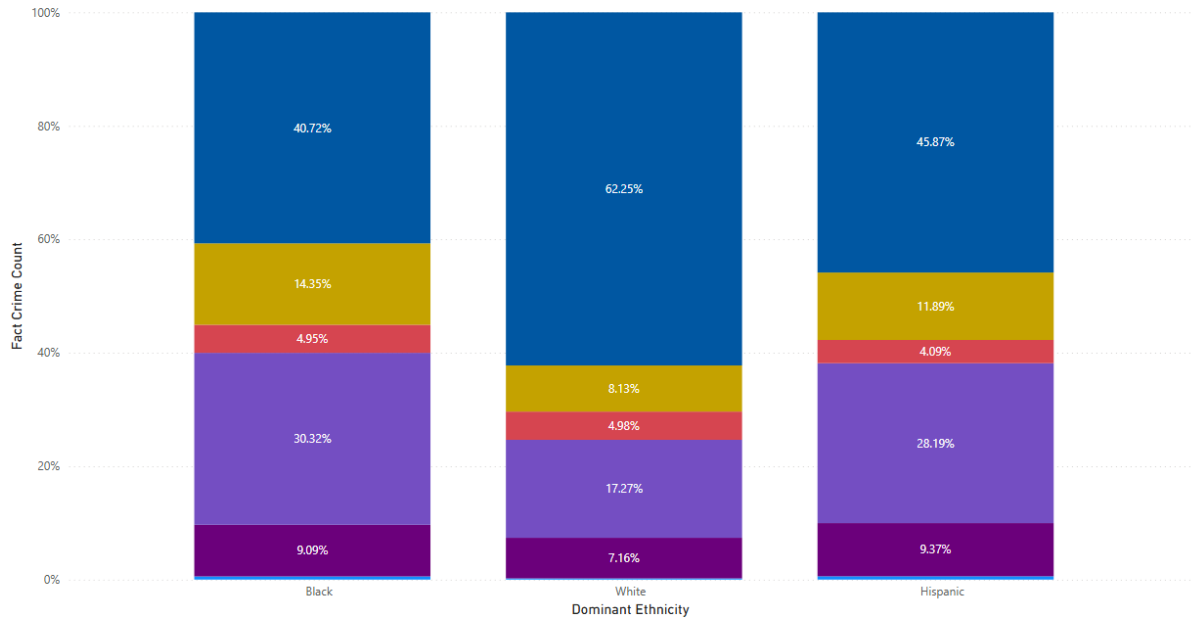
- Struktura przestępstw w wardach z różnymi dominującymi grupami etnicznymi jest ogólnie podobna — we wszystkich przypadkach dominują **przestępstwa przeciwko mieniu** (włamania, kradzieże, uszkodzenia mienia) oraz **przestępstwa z użyciem przemocy** (napady, pobicia, zabójstwa).
- W wardach, w których przeważa **populacja White (non-Hispanic)**, szczególnie widoczny jest **wysoki udział przestępstw własnościowych**, które stanowią aż **53,73%** wszystkich przestępstw. To znacząco więcej niż w wardach z dominującą populacją **Black or African American** oraz **Hispanic or Latino**.
- Wardy z dominującą populacją **White (non-Hispanic)** wykazują także wyraźnie większy udział **oszustw i naruszeń formalno-prawnych** („Fraud and Regulatory Violations”), a jednocześnie niższy udział przestępstw z użyciem przemocy niż wardy z dominującą ludnością czarnoskórą lub latynoską.

Przeanalizujemy zatem rozkład typów przestępstw w dwóch najpopularniejszych kategoriach.

Szczegółowy rozkład w kategorii "Property Crime":

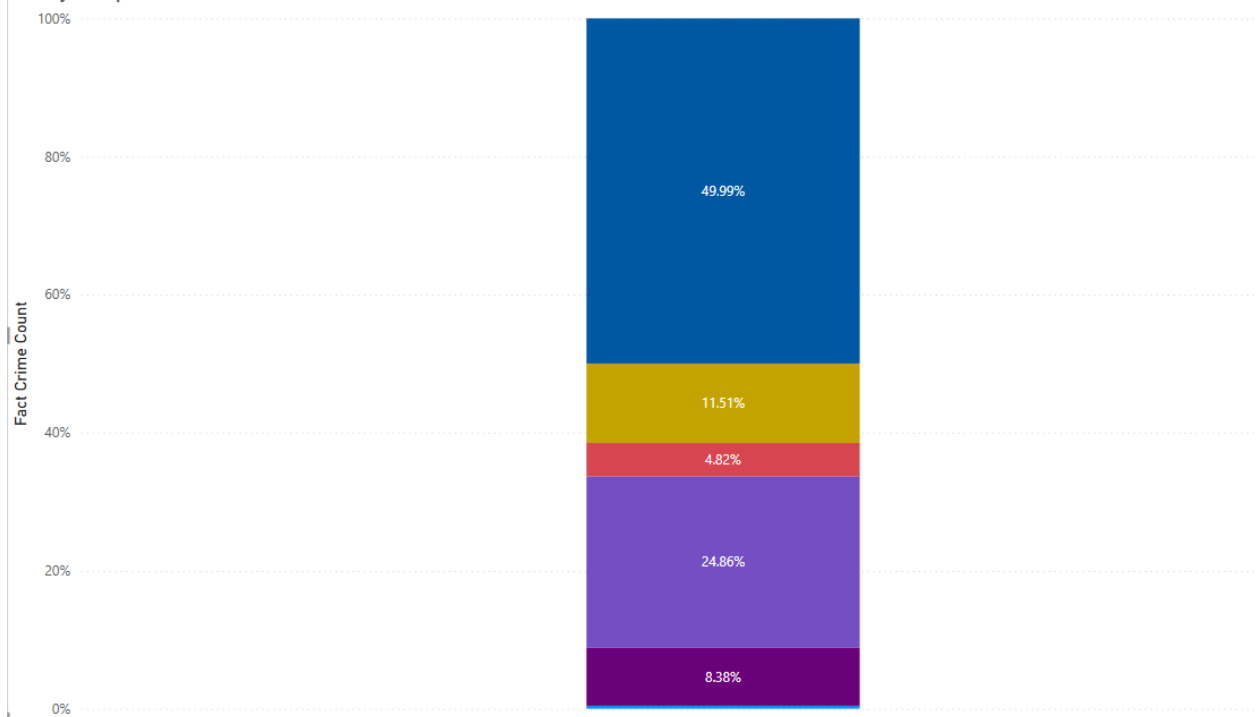
Procentowy stosunek typu przestępstwa własnościowego w zależności od grupy etnicznej dominującej w danym wardzie.

Primary Description ARSON BURGLARY CRIMINAL DAMAGE CRIMINAL TRESPASS MOTOR VEHICLE THEFT THEFT



Procentowy stosunek typu przestępstwa własnościowego.

Primary Description ARSON BURGLARY CRIMINAL DAMAGE CRIMINAL TRESPASS MOTOR VEHICLE THEFT THEFT

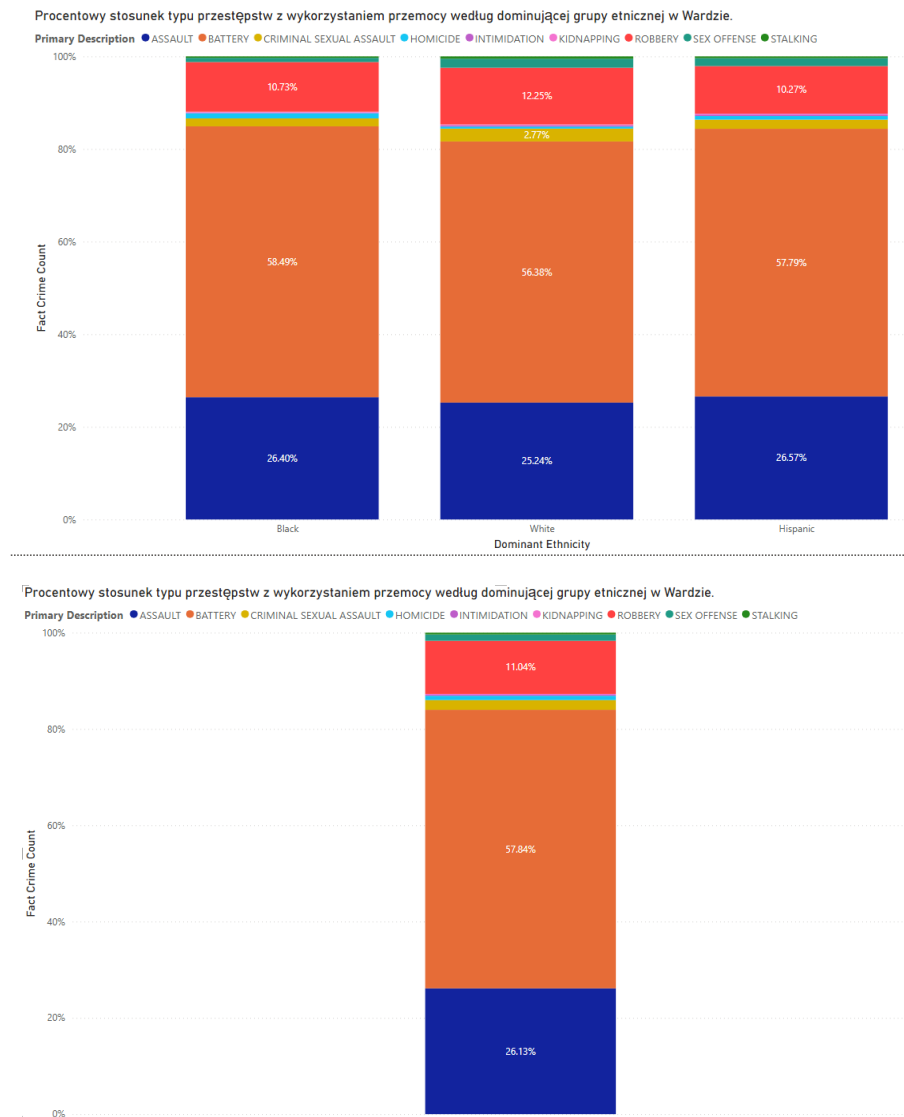


- Rozkład typów przestępstw przeciwko mieniu w wardach z dominującą populacją **Black or African American** i **Hispanic or Latino** jest bardzo zbliżony — struktura tych przestępstw nie różni się istotnie między tymi grupami.

- W wardach z dominującą populacją **White (non-Hispanic)** widoczna jest natomiast **znacznie większa dominacja kradzieży (Theft)**, przy jednocześnie mniejszym udziale **uszkodzeń mienia (Criminal Damage)** i **kradzieży pojazdów (Motor Vehicle Theft)** w porównaniu do pozostałych wardów.
- Szczegółowy rozkład wewnątrz kategorii „Property Crime”:
 - **Kradzieże (Theft): 50,00%**
 - **Uszkodzenia mienia (Criminal Damage): 24,86%**
 - **Kradzieże pojazdów (Motor Vehicle Theft): 11,51%**

Może to świadczyć o różnicach w dostępności i zabezpieczeniach majątku w zależności od dominującej grupy społecznej lub o specyfice lokalnych przestępstw przeciwko mieniu.

Szczegółowa analiza: Violent Crime (przestępstwa z użyciem przemocy)

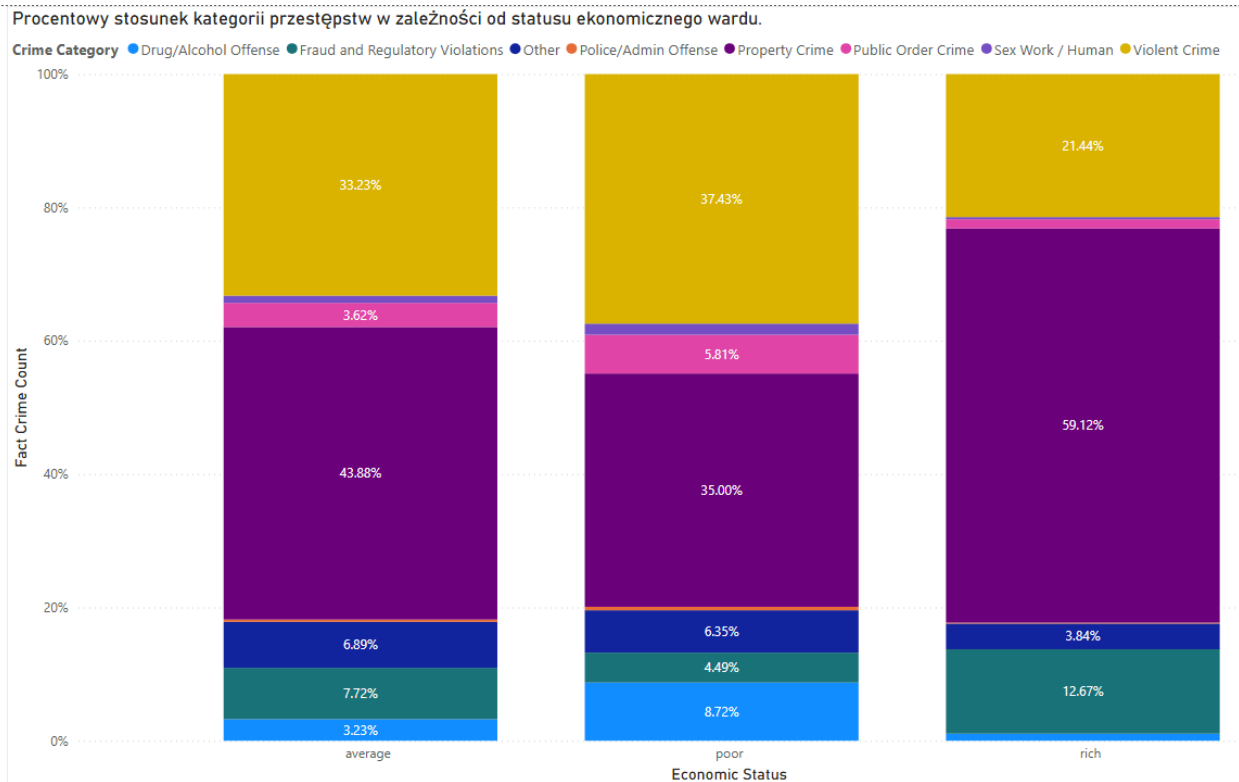


- Rozkład typów przestępstw z użyciem przemocy w wardach z dominującymi grupami **Black or African American, Hispanic or Latino i White (non-Hispanic)** jest bardzo podobny — najczęściej występują:
 - Pobicia (Battery): 57,84%**
 - Napady (Assault): 26,13%**
 - Rozboje (Robbery): 11,04%**
- Warto jednak zauważyć, że w wardach z dominującą populacją **White (non-Hispanic)** udział przestępstw na tle seksualnym (Criminal Sexual Assault) jest wyraźnie wyższy niż w pozostałych:

- **White (non-Hispanic): 2,77%**
- **Black or African American: 1,71%**
- **Hispanic or Latino: 2,08%**

Różnice te mogą wynikać zarówno z lokalnych uwarunkowań, jak i ze sposobu zgłaszania i klasyfikowania przestępstw na tle seksualnym w różnych społecznościach.

Analiza 2: Typy przestępstw vs status ekonomiczny wardu



- **Struktura przestępstw istotnie różni się w zależności od poziomu zamożności dzielnicy.** W wardach **biednych** dominuje **większy udział przestępstw z użyciem przemocy** (Violent Crime – 37,43%) w porównaniu do wardów bogatych (21,44%) i średniozamożnych (33,23%).
- W wardach **bogatych** wyraźnie dominuje kategoria **przestępstw przeciwko mieniu (Property Crime)**, która stanowi **59,12%** wszystkich przestępstw — znacznie więcej niż w wardach biednych (35,00%) oraz wardach średniozamożnych (43,88%). Może to sugerować, że w zamożniejszych dzielnicach dominuje przestępczość o charakterze oportunistycznym (kradzieże, włamania), a nie agresywnym.
- **Przestępstwa związane z porządkiem publicznym (Public Order Crime)** oraz **przestępstwa narkotykowe i alkoholowe (Drug/Alcohol Offense)** mają **największy udział w wardach biednych**, co może wskazywać na częstsze działania prewencyjno-policyjne lub większą ekspozycję mieszkańców tych obszarów na tego typu interwencje.
- **Fraud and Regulatory Violations** występują zdecydowanie częściej w Wardach średniozamożnych (7.72%) i bogatych (12.67%) niż w Wardach biednych (4.49%).

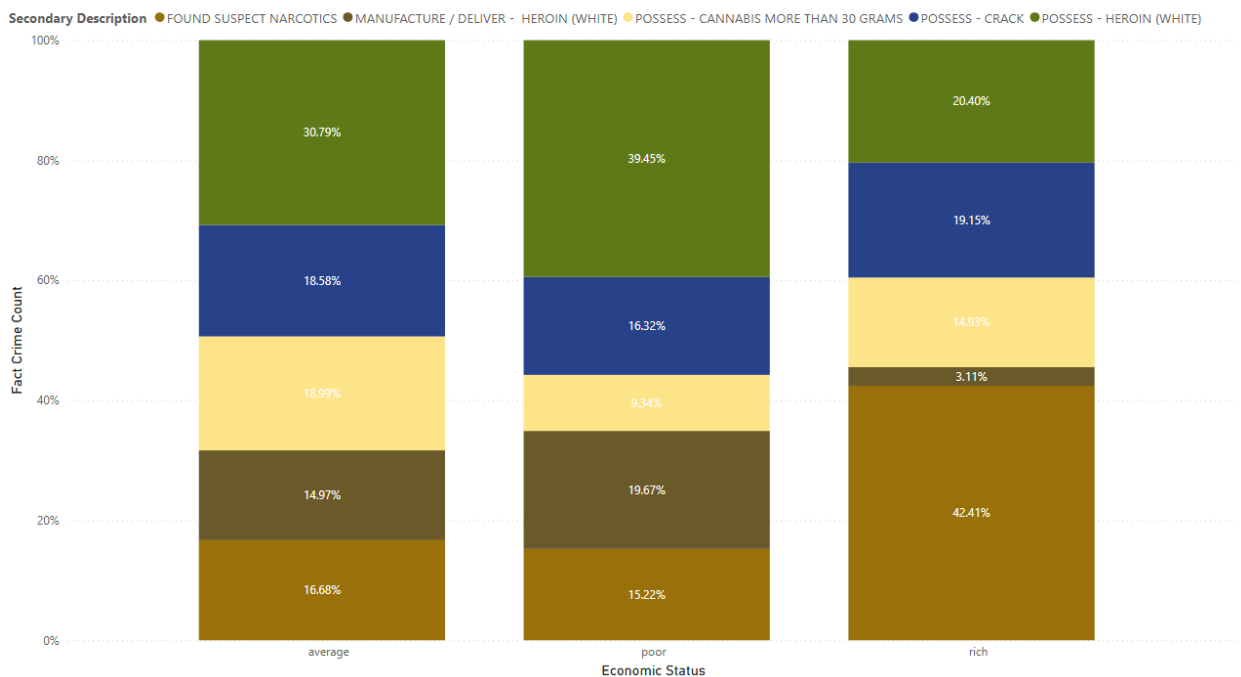
Przeanalizujemy Fraud and Regulatory Violations oraz Drug/Alcohol Offense z powodu widocznej różnicy między wardami biednymi/średniozamożnymi oraz bogatymi.

Procentowy stosunek typu przestępstw formalnych oraz narkotykowych w zależności od statusu ekonomicznego wardu.



Widzimy, że przestępstwa narkotykowe dominują w wardach biednych dominują przestępstwa narkotykowe, a w średnich oraz bogatych oszustwa (finansowe oraz zeznaniowe).

Procentowy rozkład top 5 przestępstw narkotykowych według typu ekonomicznego Wardu.



Wśród pięciu najczęściej występujących przestępstw związanych z narkotykami, w zależności od statusu ekonomicznego wardu, znalazły się następujące kategorie:

- **Posiadanie heroiny** – stanowi najczęstszy typ przestępstwa narkotykowego w **wardach o niskim statusie ekonomicznym**.
- **Posiadanie cracku** – regularnie występuje w każdej grupie wardów.
- **Posiadanie ponad 30 gramów marihuany** – istotna kategoria szczególnie w wardach średniozamożnych i bogatych.
- **Wytwarzanie i rozprowadzanie heroiny** – częściej występuje w wardach **biednych oraz średniozamożnych**, co może świadczyć o obecności zorganizowanej przestępczości narkotykowej w tych obszarach.
- **Posiadanie podejrzanych substancji** – najczęściej rejestrowane w **wardach bogatych**, gdzie odpowiada aż za **42% przestępstw narkotykowych**.

Wysoki udział przestępstw typu "posiadanie podejrzanych substancji" w wardach bogatszych może sugerować **odmienny sposób klasyfikowania wykroczeń lub mniej surowe podejście organów ścigania**. Może to również wskazywać na **różnice w praktykach policyjnych** lub na **potencjalne uprzywilejowanie społeczności bardziej zamożnych** – co warto byłoby dalej zbadać w kontekście **sprawiedliwości społecznej i systemowej równości**.

Procentowy rozkład top 5 oszustw i przestępstw na tle legalnym według typu ekonomicznego Wardu.



Najczęściej występujące przestępstwa z kategorii **oszustw i naruszeń formalno-prawnych** różnią się w zależności od statusu ekonomicznego wardów:

- **Kradzież tożsamości finansowej powyżej 300 dolarów** (*Financial Identity Theft over \$300*) stanowi największy odsetek tej kategorii w wardach o niskim i średnim statusie ekonomicznym:
 - **Poor wards: aż 40,99%**
 - **Average wards: 33,33%**
 - W wardach zamożnych występuje rzadziej
- **Kradzież tożsamości finansowej do 300 dolarów** (*\$300 and under*) również jest obecna, ale nie dominuje wyraźnie w żadnej klasie ekonomicznej
- **Oszustwa kartą kredytową** (*Credit Card Fraud*) są **znacznie częstsze w wardach zamożnych**:
 - **Rich wards: 37,12%**
 - Dla porównania: tylko **15,08%** w wardach biednych i **17,06%** w średniozamożnych
- **Oszustwa związane z grami losowymi i wykorzystaniem kart płatniczych** (*Confidence Game, Illegal Use of Cash Card*) występują w **podobnych proporcjach** niezależnie od klasy ekonomicznej wardu

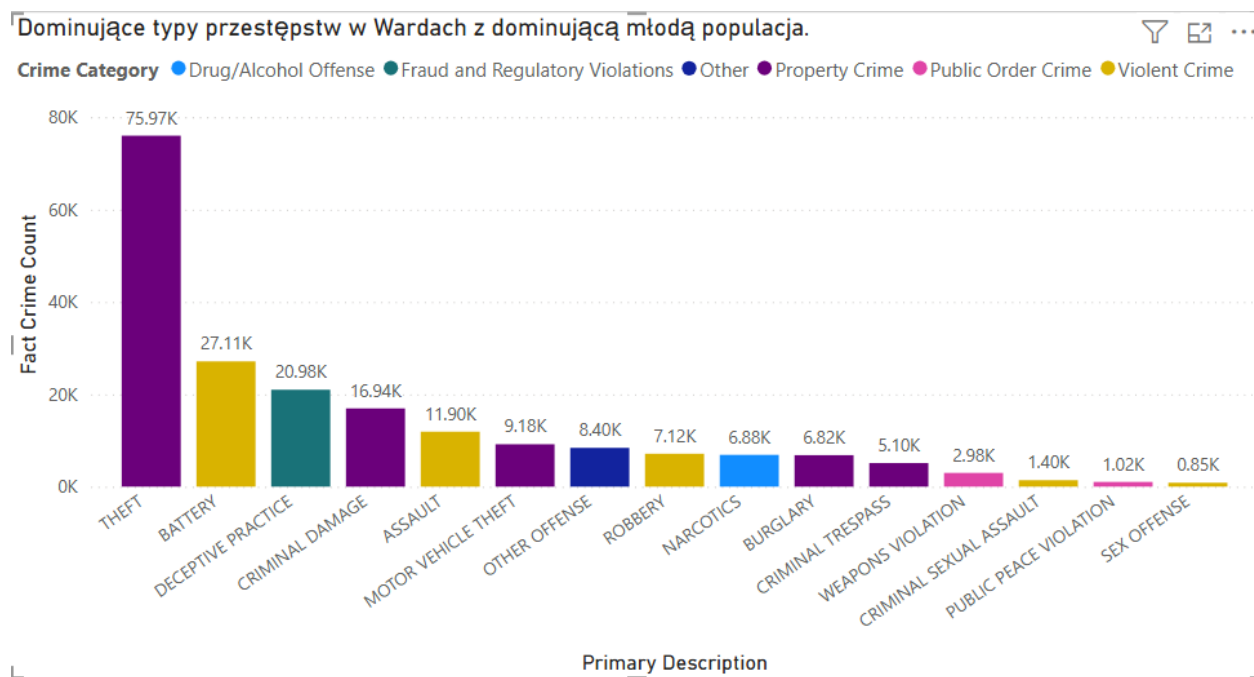
Wnioski: Typ przestępstwa a status ekonomiczny wardu

Analiza danych wskazuje na **istotne zależności między klasą ekonomiczną wardów a strukturą popełnianych przestępstw**:

- **W wardach biedniejszych** (poor) dominują klasyczne, bardziej bezpośrednie przestępstwa, takie jak:
 - **Posiadanie heroiny**
 - **Kradzież tożsamości finansowej powyżej \$300**
 - **Produkcja i dystrybucja narkotyków**
- **W wardach zamożnych** (rich) częściej występują:
 - **Oszustwa elektroniczne i z użyciem kart kredytowych**
 - **Posiadanie podejrzanych substancji**, co może być efektem mniej jednoznacznego egzekwowania prawa lub innej klasyfikacji wykroczeń
- **Wardów średniozamożnych** dotyczą w dużym stopniu przestępstwa o mieszanym charakterze, m.in. posiadanie marihuany w większych ilościach czy kradzieże danych osobowych na mniejszą skalę

Wnioski te sugerują, że **profil przestępczości odzwierciedla lokalne warunki społeczno-ekonomiczne**, a być może również **różnice w sposobie klasyfikowania i ścigania przestępstw** w zależności od dzielnicy. Dane te stanowią punkt wyjścia do dalszych badań nad równością systemu karnego oraz ukrytymi mechanizmami uprzywilejowania lub marginalizacji poszczególnych grup społecznych.

Analiza 3: Przestępstwa w wardach z młodą populacją



Analiza danych dla wardów, w których **dominującą grupą demograficzną są osoby młode (young-majority)**, ujawnia charakterystyczny profil przestępczości:

- **Zdecydowanie najczęstszym typem przestępstwa jest kradzież („Theft”),** która wystąpiła aż **75,97 tys. razy**. Oznacza to, że **przestępstwa przeciwko mieniu dominują** w tej grupie wardów.
- Na kolejnych miejscach znajdują się przestępstwa z użyciem przemocy:
 - **Pobicia („Battery”) – 27,11 tys. przypadków**
 - **Napaści („Assault”) – 11,90 tys. przypadków**

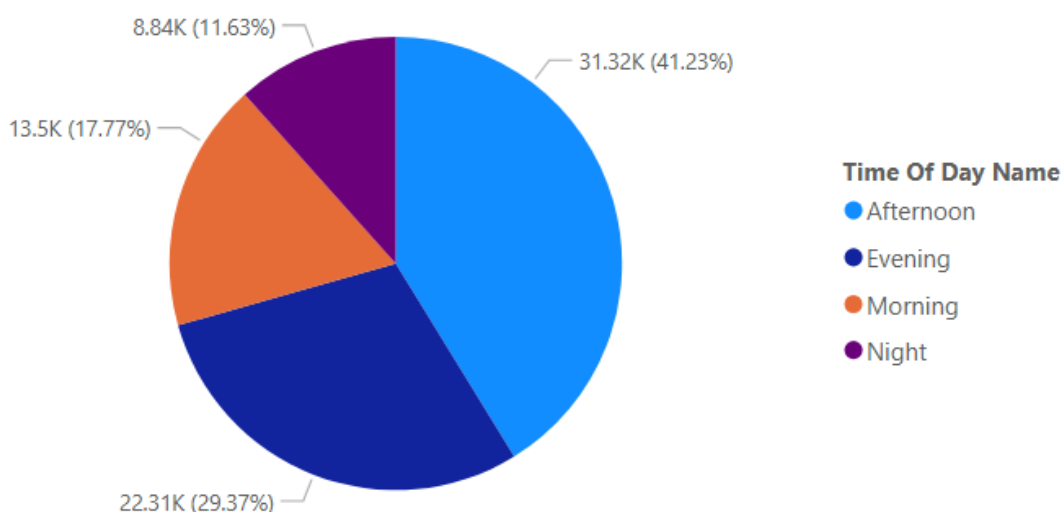
To sugeruje, że młoda populacja jest **bardziej narażona zarówno na utratę mienia, jak i na fizyczną przemoc**.

- **Przestępstwa związane z oszustwami i naruszeniami regulacyjnymi** (np. „Deceptive Practice”) wystąpiły **20,98 tys. razy**, co wskazuje na **istotną obecność przestępczości gospodarczej** wśród młodych mieszkańców — być może powiązaną z większym udziałem transakcji cyfrowych i konsumpcyjnym stylem życia.
- **Przestępstwa narkotykowe („Narcotics”),** choć mniej liczne (**6,88 tys.**), nadal stanowią **zauważalny element lokalnej przestępczości**.

- **Wysoka liczba kradzieży niebędących rozbojami** może mieć związek z działalnością **kieszonkowców i drobnych złodziei**, szczególnie w kontekście **wydarzeń masowych, życia nocnego i imprez**, które są popularne wśród młodszych mieszkańców miasta.

W celu zweryfikowania tej hipotezy przeprowadzona zostanie **dodatkowa analiza zależności kradzieży od pory dnia**, która pozwoli ocenić, czy wzmożona aktywność przestępcza przypada na godziny nocne.

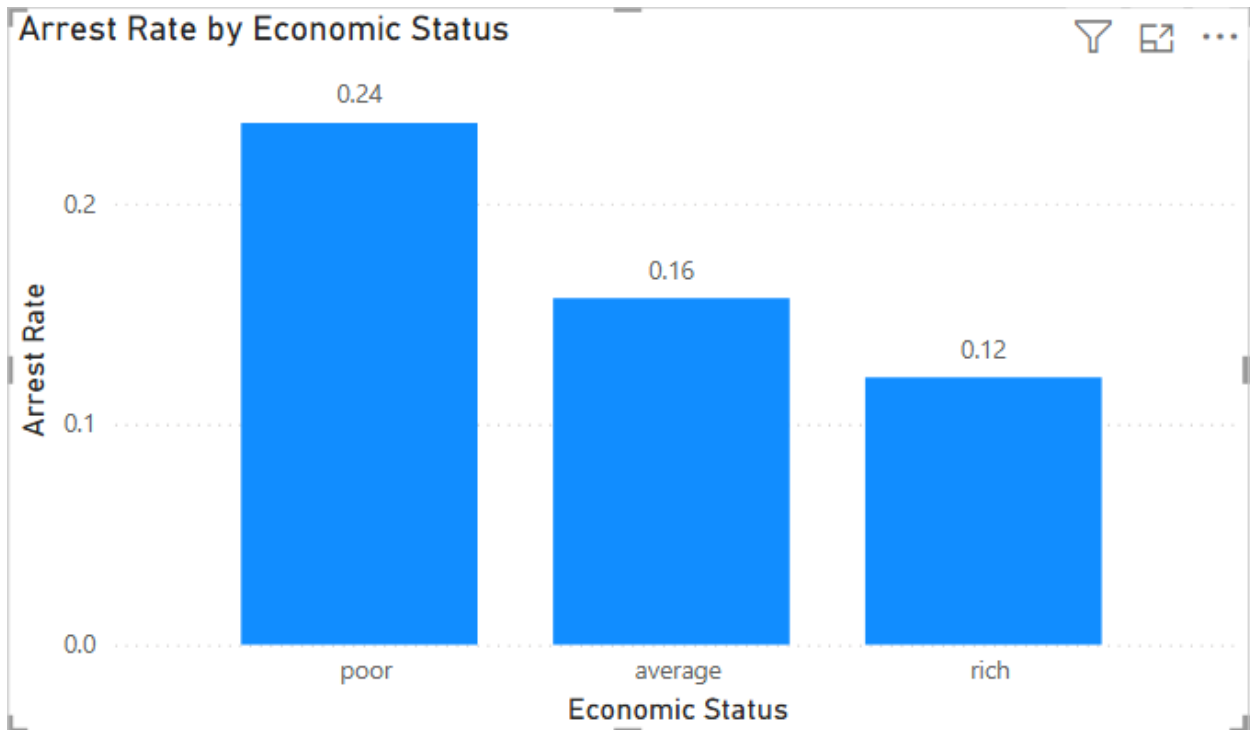
Rozkład czasu dnia kradzieży w Wardach z dominującym młodym społeczeństwem.



Analiza rozkładu przestępstw typu „Theft” względem pory dnia (TimeOfDayName) **nie potwierdza jednoznacznie hipotezy**, że kradzieże są najczęstsze wieczorem lub w nocy — czyli w godzinach typowo związanych z imprezami i aktywnością społeczną osób młodych.

- **Najwięcej kradzieży** odnotowano **w godzinach popołudniowych (Afternoon)** – **31 323 przypadki**, co stanowi największy udział w tej analizie.
- Na drugim miejscu znajdują się **godziny wieczorne (Evening)** – **22 312 przypadków**.
- **Rano (Morning): 13 497 kradzieży**
- **W nocy (Night)** odnotowano wyraźnie **najmniej kradzieży** – **8 836 przypadków**

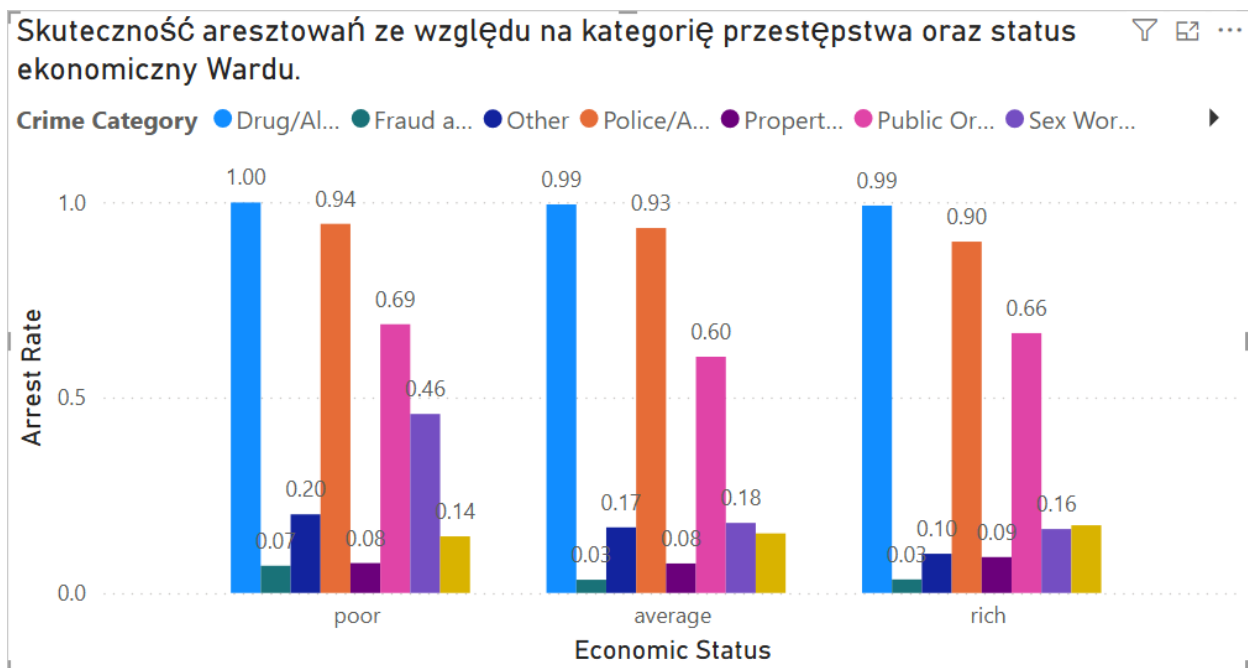
Analiza 4: Skuteczność aresztowań względem statusu ekonomicznego



Analiza skuteczności aresztowań (Arrest Rate) względem **statusu ekonomicznego wardów** wskazuje na wyraźną zależność:

- W wardach biednych (poor): średni wskaźnik aresztowań wynosi **24%**
- W wardach średniozamożnych (average): **16%**
- W wardach bogatych (rich): tylko **12%**

Dane te sugerują, że **im wyższy status ekonomiczny wardu, tym niższa ogólna skuteczność aresztowań**. Może to być jednak częściowo wynikiem **różnic w strukturze typów popełnianych przestępstw**.



Po rozszerzeniu analizy o dodatkowy wymiar – **kategorię przestępstwa (Crime Category)** – zauważamy, że:

- **W większości kategorii skuteczność aresztowań nie różni się znacząco między klasami wardów.** Sugeruje to, że różnice ogólne są raczej pochodną rozkładu typów przestępstw niż różnic w działaniach policji.
- **Wyjątkiem są przestępstwa z kategorii „Sex Work / Human Crimes”,** gdzie widać silną rozbieżność (różnica ta będzie analizowana w dalszej części raportu):
 - **Ward poor: 46%**
 - **Ward average: 18%**
 - **Ward rich: 16%**
- **Kategorie przestępstw o najwyższej skuteczności aresztowań:**
 - **Drug/Alcohol Offense: 99–100%**
 - Zrozumiałe, ponieważ są to przestępstwa ujawniane bezpośrednio przez funkcjonariuszy podczas zatrzymania
 - **Police / Admin Offense: 90–94%**
 - Również wynikające z bezpośredniego kontaktu z policją
 - **Public Order Crime: 60–69%**

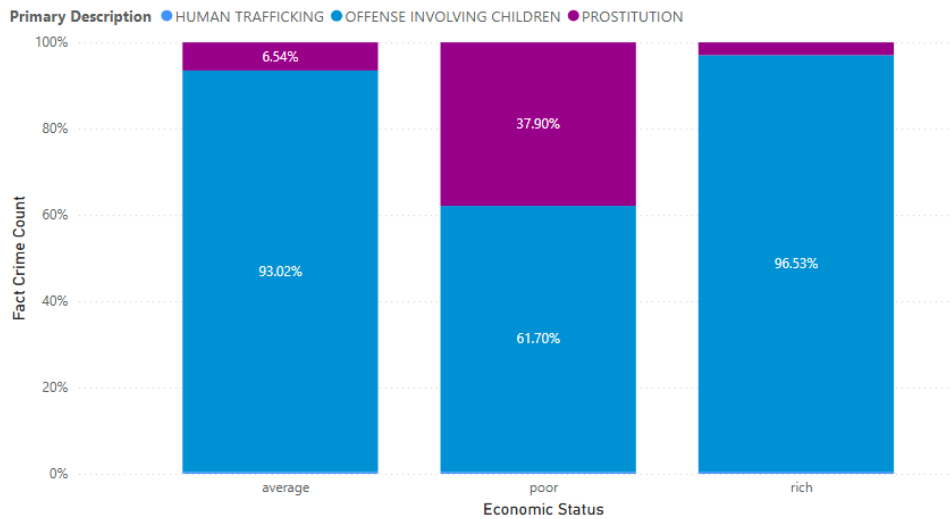
- Prawdopodobnie także ujawniane w sposób bezpośredni (np. zakłócanie porządku publicznego)
- **Kategorie przestępstw o najniższej skuteczności aresztowań:**
 - **Violent Crime: 8–9%**
 - Możliwe trudności z identyfikacją sprawców, brak świadków lub zbierania dowodów
 - **Fraud and Regulatory Violations: 3–7%**
 - Część tych przestępstw nie zawsze wymaga aresztowania (np. nadużycia administracyjne), część może pozostać niewykryta
 - **Property Crime: 14–18%**
 - Podobnie jak w przypadku przemocy – trudności z wykryciem sprawców, zwłaszcza przy kradzieżach bez świadków

Różnice w skuteczności aresztowań między wardami o różnym statusie ekonomicznym wynikają głównie ze zróżnicowanej struktury typów przestępstw, a nie z samego podejścia organów ścigania. Najwyższa skuteczność dotyczy przestępstw, w których sprawcy są ujawniani bezpośrednio przez policję, natomiast w przypadku przestępstw przeciwko mieniu, przemocy czy oszustw – identyfikacja sprawców jest znacząco trudniejsza. Istotna różnica w kategorii „Sex Work / Human Crimes” może wskazywać na różnice w egzekwowaniu prawa lub celowości działań operacyjnych w zależności od dzielnicy, co wymaga dalszej analizy.

Skuteczność aresztowań ze względu na typ przestępstwa Sex / Human Crime i status ekonomiczny Wardu.



Stosunek liczby przestępstw ze względu na typ przestępstwa Sex / Human Crime i status ekonomiczny Wardu.



Szczegółowa analiza podkategorii przestępstw w ramach grupy „**Sex Work / Human Crimes**” pozwala lepiej zrozumieć różnice w skuteczności aresztowań między wardami o różnym statusie ekonomicznym.

Skuteczność aresztowań ze względu na typ przestępstwa:

- **Prostitution** – skuteczność aresztowań jest bardzo wysoka i praktycznie identyczna we wszystkich wardach: **99–100%**
 - Wynika to najprawdopodobniej z bezpośredniego zatrzymania sprawców na miejscu.
- **Offense Involving Children** – skuteczność na poziomie **12–14%**, niezależnie od klasy wardu
 - Trudności w wykrywaniu, delikatny charakter spraw, możliwy brak świadków
- **Human Trafficking** – wyraźne różnice:
 - 0 przypadków w wardach biednych
 - 8% aresztowań w wardach średniozamożnych
 - 33% aresztowań w wardach bogatych
 - Może to sugerować różnice w podejściu służb, nierówną skuteczność dochodzeń, a także możliwość występowania czynników instytucjonalnych, takich jak **niedoinwestowanie służb w biedniejszych rejonach** czy potencjalna **presja społeczno-polityczna** w dzielnicach zamożniejszych. Lub najzwyczajniej **zbyt małą liczbę przypadków żeby wyciągać sensowne statystyki**.

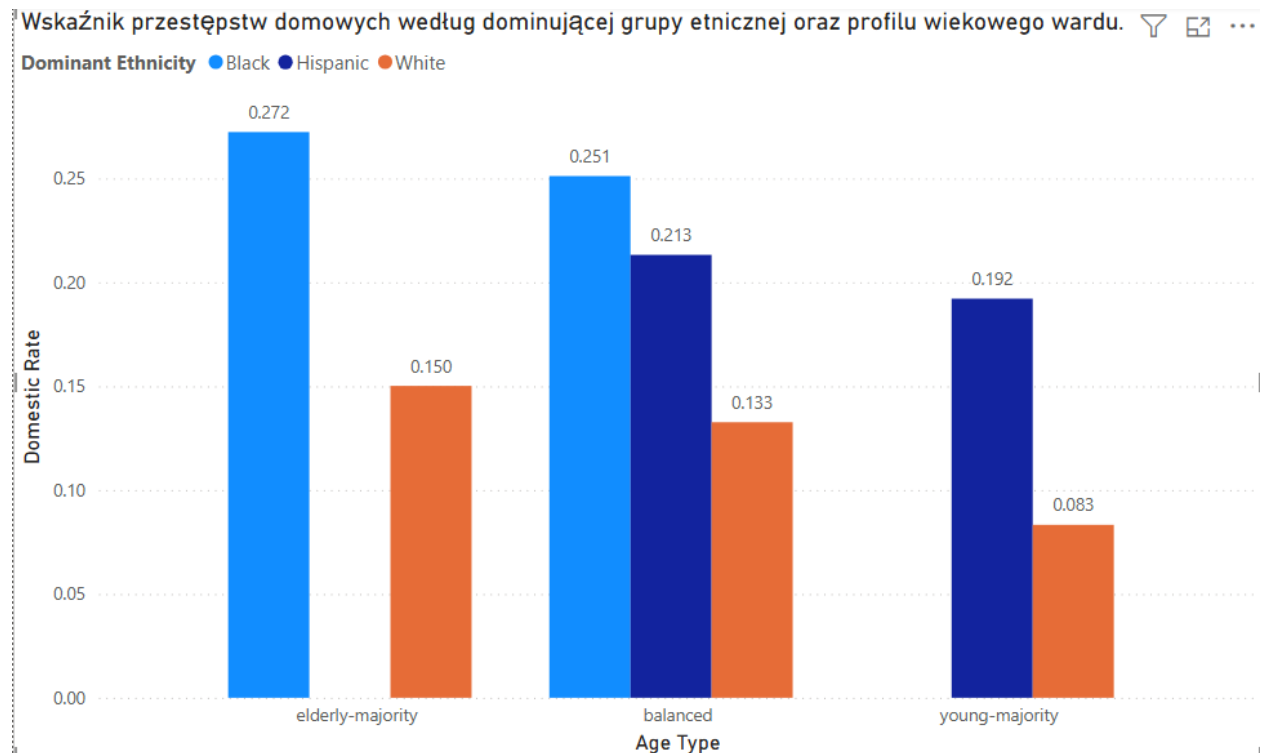
Struktura przestępstw:

- **Ward poor:**
 - 0 przypadków Human Trafficking
 - 37,9% stanowią przestępstwa związane z prostytutką
 - 61,7% – przestępstwa z udziałem dzieci
- **Ward average:**
 - 37 przypadków Human Trafficking
 - 93,02% – Offense Involving Children
 - 6,54% – Prostitution
- **Ward rich:**
 - 3 przypadki Human Trafficking
 - 96,53% – Offense Involving Children
 - 2,97% – Prostitution

Różnice w skuteczności aresztowań w tej kategorii **nie wynikają wyłącznie z podejścia służb**, lecz również z **odmiennych struktur przestępczości w poszczególnych wardach**. Przestępstwa związane z prostytutką są bardziej powszechne w wardach biednych, natomiast przypadki handlu ludźmi pojawiają się wyłącznie w wardach średniozamożnych i bogatych – choć bardzo rzadko.

Niska skuteczność w sprawach najpoważniejszych (np. Human Trafficking) może wskazywać na potrzebę lepszego finansowania śledztw, rozwoju procedur międzyinstytucjonalnych lub ujawniać luki systemowe. Wysoka skuteczność aresztowań za prostytutkę kontrastuje z niską skutecznością w sprawach bardziej złożonych i społecznie istotnych, co może wymagać dalszej refleksji.

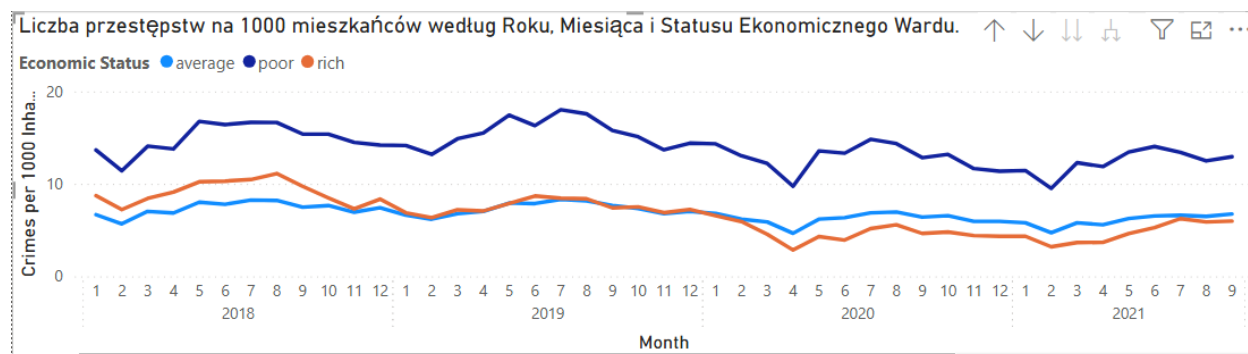
Analiza 5: Wskaźnik przestępstw domowych a profil wiekowy i etniczny



Analiza wskaźnika przestępstw domowych (DomesticCrimeRate) względem dominującego profilu **wiekowego** i **etnicznego** wardów ujawnia wyraźne różnice w skali tego zjawiska:

- **Najwyższy wskaźnik przestępstw domowych** (27,2%) odnotowano w wardach o starszej populacji (elderly-majority) i dominującej czarnoskórej społeczności (Black).
- **Wysoki poziom** (25,1%) występuje również w wardach zrównoważonych wiekowo (balanced) i czarnoskórych.
- Wardy **latynoskie (Hispanic)** – zarówno zrównoważone, jak i z młodą populacją – mają umiarkowane wskaźniki (21,3% i 19,2%), ale nadal wyższe niż wardy białe.
- **Najniższe wskaźniki** występują w wardach z **dominacją białej populacji (White)** – zarówno w grupie starszej (15%), jak i młodszej (8,3%).

Analiza 6: Trend przestępczości według Statusu Ekonomicznego Wardu



1. Wardy o niskim statusie ekonomicznym (poor) mają zdecydowanie najwyższy wskaźnik przestępczości

- Przez cały okres 2018–2022, wardy określone jako *poor* utrzymują **dwukrotnie lub nawet trzykrotnie wyższy poziom przestępczości** niż wardy bogate.
- Przykład: w lipcu 2019 r. było to **18,0** przestępstw na 1000 mieszkańców w wardach biednych, przy **8,4** w wardach bogatych.

2. Wardy bogate (rich) cechują się najniższym wskaźnikiem – i to konsekwentnie

- W żadnym miesiącu wskaźnik nie przekroczył **11 przestępstw na 1000 mieszkańców**, a po 2020 roku rzadko przekraczał 7.
- W szczycie pandemii (kwiecień 2020), spadł nawet do **2,84**, co jest jednym z najniższych poziomów w całym zestawie danych.

3. Wardy średniozamożne (average) wykazują umiarkowany poziom przestępczości

- Wskaźnik przestępstw oscyluje najczęściej między **6 a 8**.
- Zmiany są sezonowe, ale mniej skrajne niż w wardach biednych.

4. Sezonowość widoczna we wszystkich grupach, ale na różnym poziomie

- W miesiącach letnich (maj–sierpień) wskaźniki rosną, szczególnie w wardach biednych (np. **14–18**), co może być powiązane z aktywnością społeczną, pogodą i większą liczbą interakcji międzyludzkich.
- W miesiącach zimowych wskaźniki spadają — szczególnie w wardach rich, gdzie różnica między zimą a latem może wynosić nawet **ponad 3 pkt proc.**

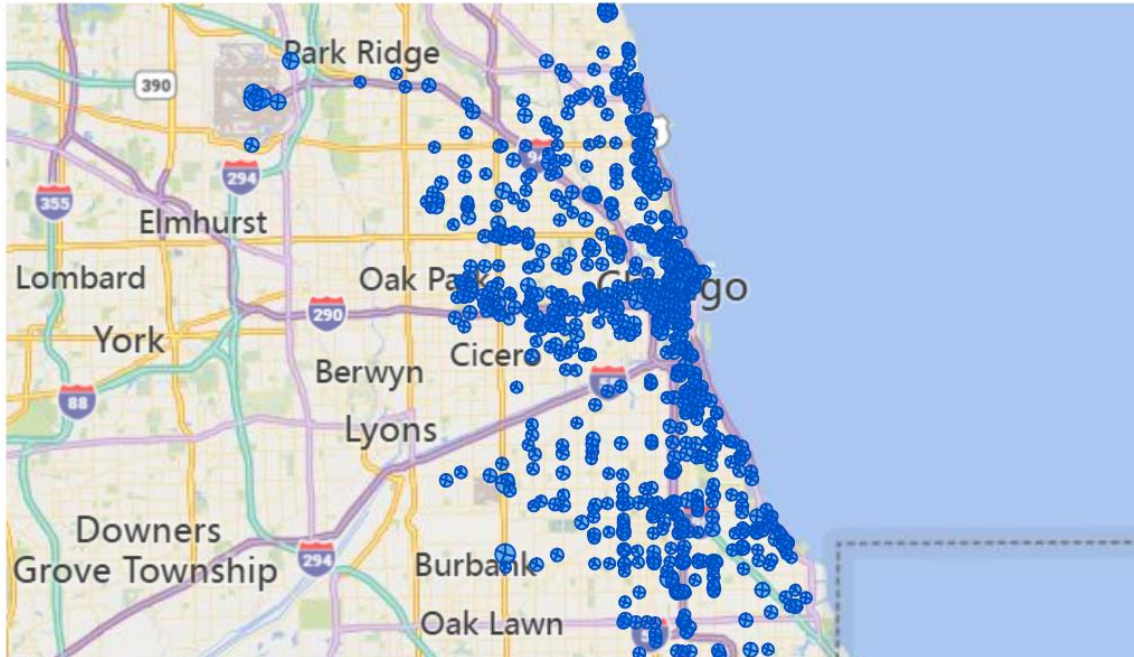
5. Pandemia COVID-19 (2020) wpłynęła na wszystkie grupy, ale najbardziej na wardy bogate

- Największy spadek przestępczości na 1000 mieszkańców dotyczy wardów rich – w kwietniu 2020 spadek do **2,83**, w porównaniu z ~5–7 w innych latach.
- Wardy biedne również zanotowały spadki, ale ich poziom nadal był wyższy niż jakiegokolwiek wynik w wardach rich.

Analiza 7: Liczba przestępstw wg dnia tygodnia i lokalizacji.

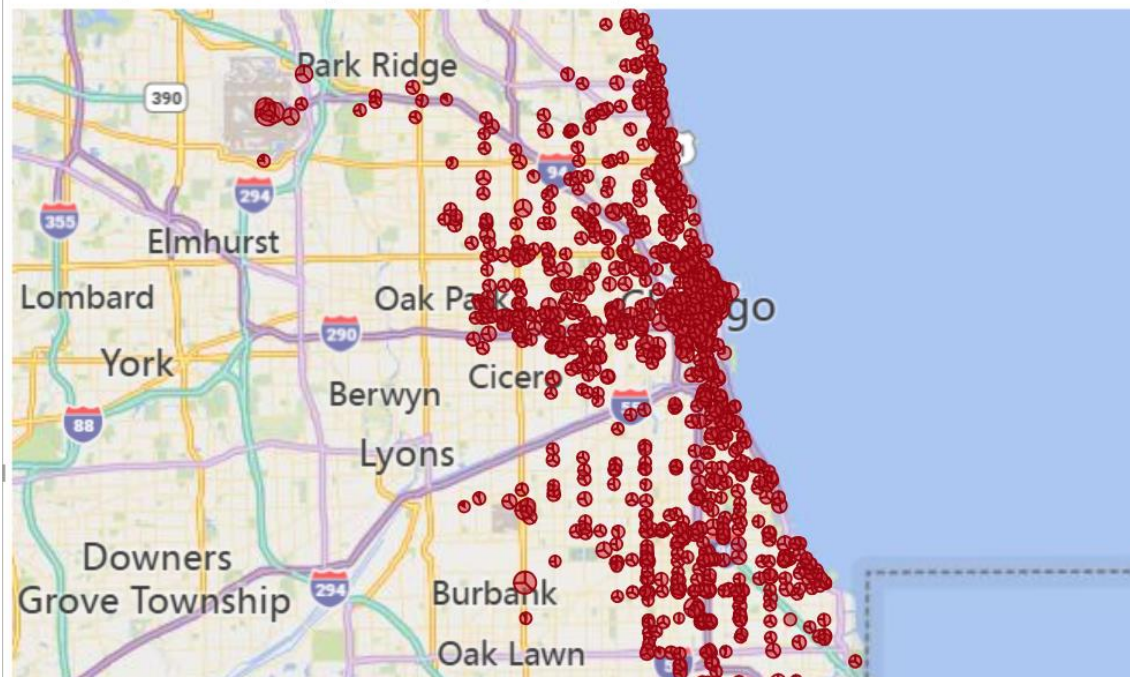
Mapa przestępstw w Chicago w czasie tygodnia (poniedziałek-czwartek)

Weekday Name ● Monday ● Thursday ● Tuesday ● Wednesday



Mapa przestępstw w Chicago w czasie weekendu (piątek - niedziela)

Weekday Name ● Friday ● Saturday ● Sunday

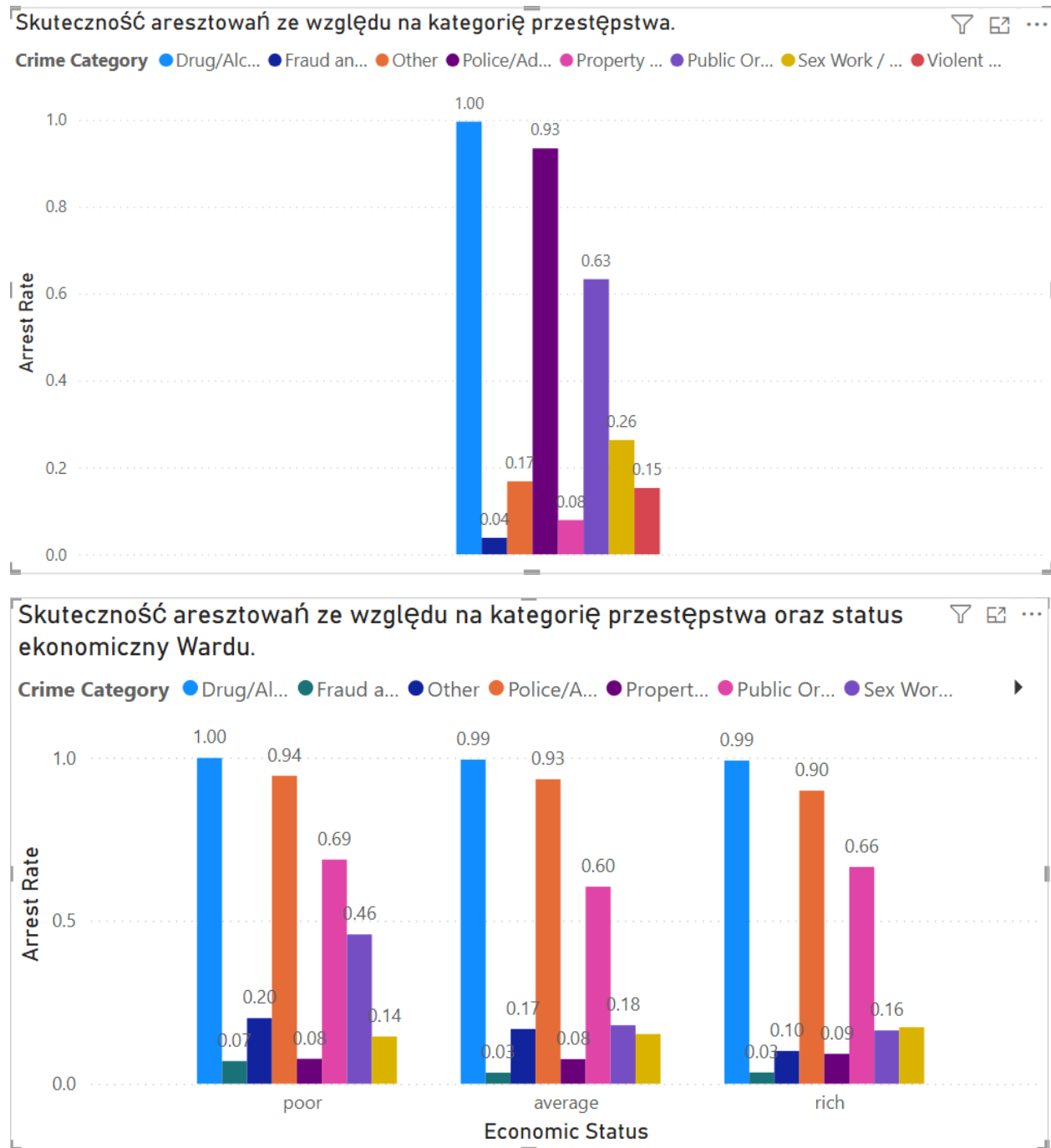


Analiza przestrzenna rozmieszczenia przestępstw w zależności od dnia tygodnia ujawnia, że:

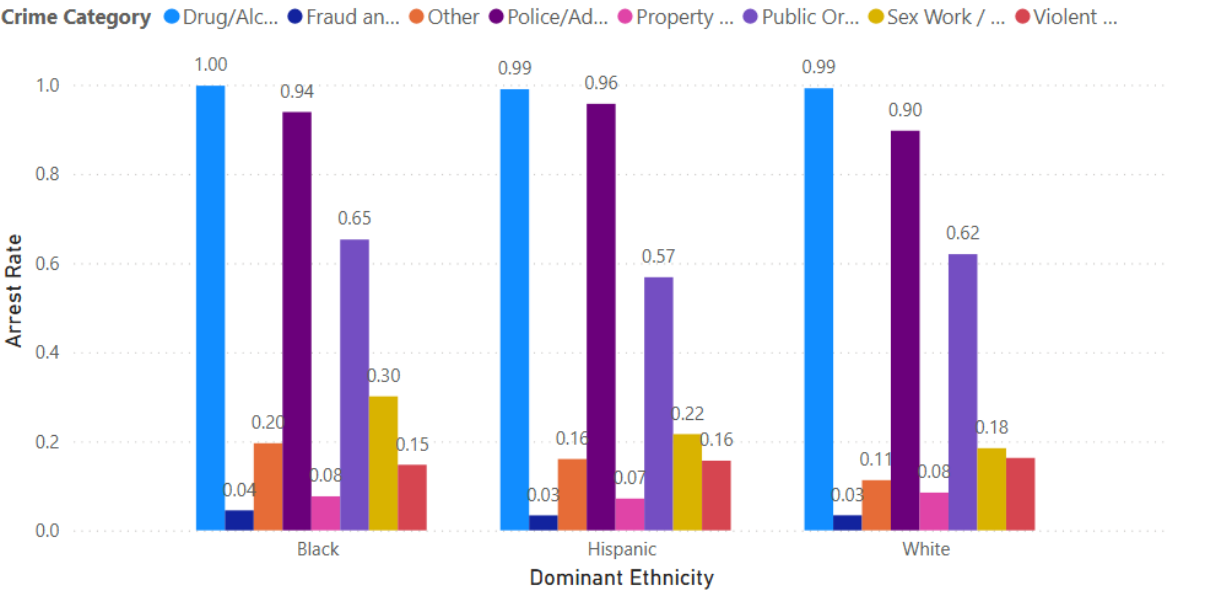
- **Ogólny rozkład lokalizacji przestępstw w weekendy i dni robocze jest bardzo podobny** – przestępstwa koncentrują się w tych samych, najgęściej zaludnionych rejonach miasta.
- **W weekendy zauważalny jest nieco większy udział incydentów w rejonach śródmieścia oraz wzdłuż nabrzeża**, co może być powiązane z aktywnością nocną, życiem towarzyskim i turystyką.
- Poza tym, **większość lokalizacji o wysokim natężeniu przestępczości występuje zarówno w tygodniu, jak i w weekendy**, co sugeruje stały charakter zagrożeń w tych obszarach.

Analiza 8: Liczba przestępstw wg typu i rezultatu (areszt)

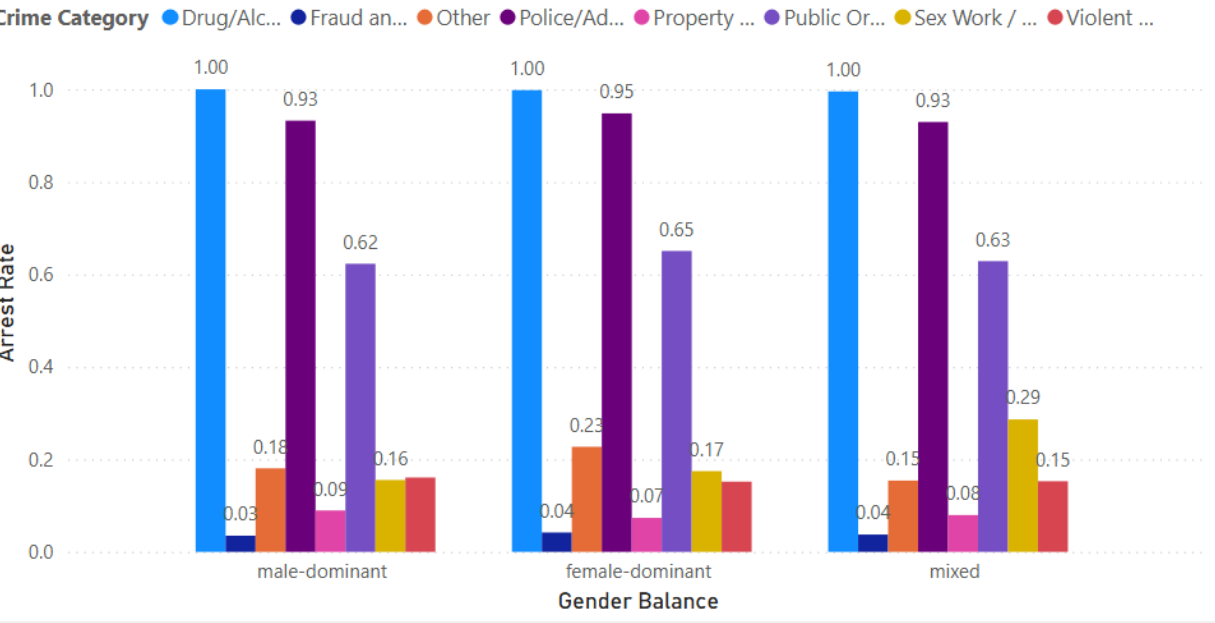
Część analizy została już dokonana w trakcie analizy rezultatu przestępstwa według statusu ekonomicznego wardu w Analizie 4, uzupełnimy więc ją o pozostałe czynniki demograficzne. Oraz zagłębimy się w Violent Crimes.



Skuteczność aresztowań ze względu na kategorię przestępstwa i dominującą grupę etniczną wardu.

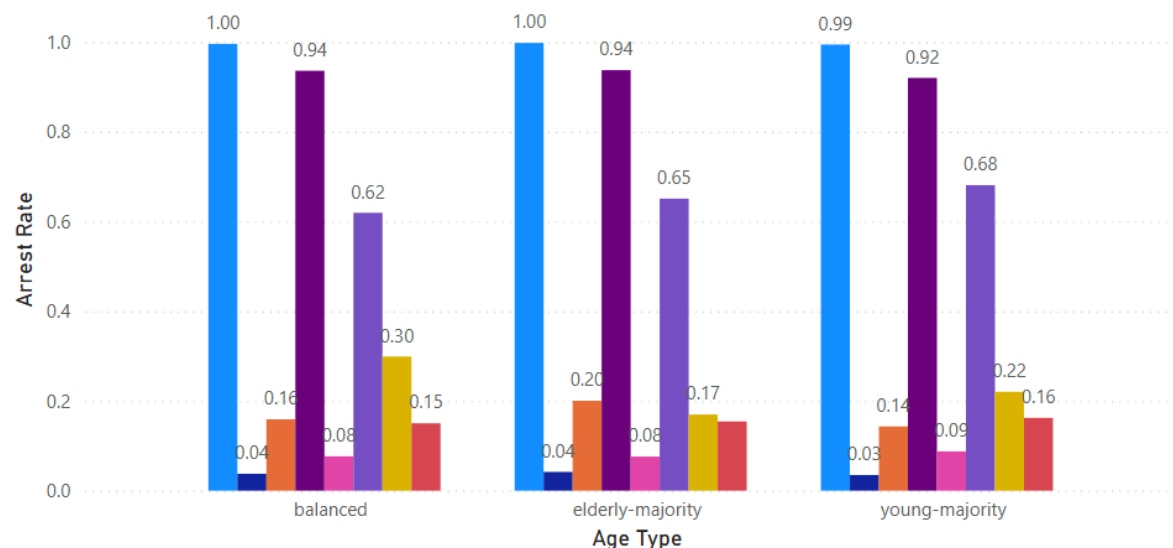


Skuteczność aresztowań ze względu na kategorię przestępstwa i balans płci wardu.



Skuteczność aresztowań ze względu na kategorię przestępstwa i strukturę wiekową.

Crime Category ● Drug/Alc... ● Fraud an... ● Other ● Police/Ad... ● Property ... ● Public Or... ● Sex Work / ... ● Violent ...

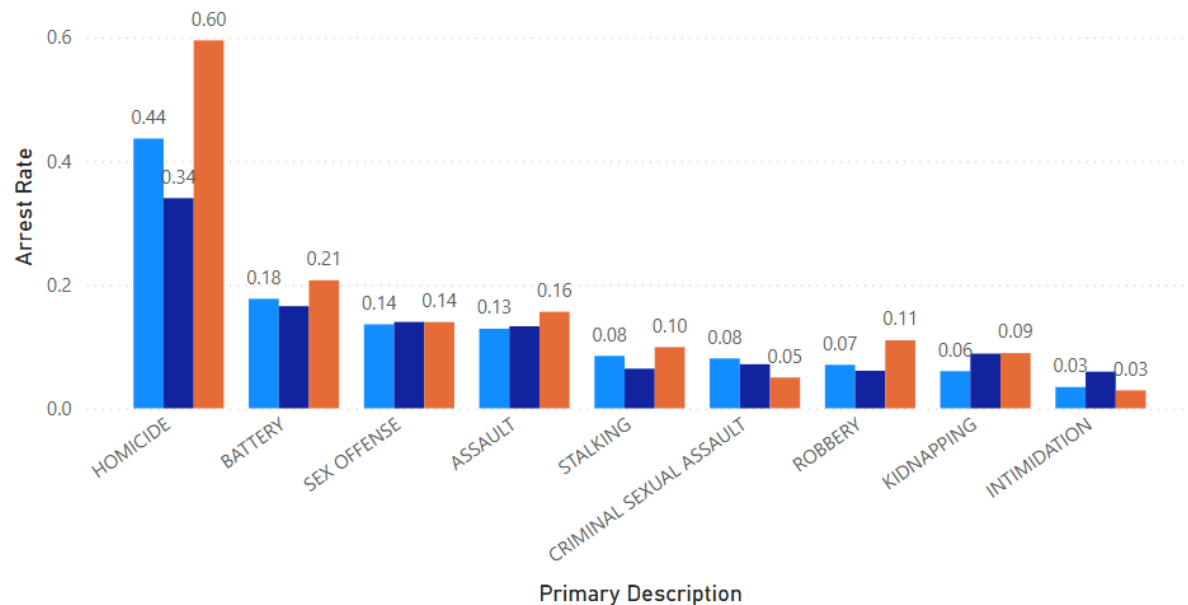


- **Analiza skuteczności aresztowań w podziale na różne czynniki demograficzne i ekonomiczne** (status majątkowy, dominująca grupa etniczna, struktura wiekowa) **nie wykazała istotnych różnic w sposobie działania organów ścigania między wardami.**
- **Poziom zatrzymań dla poszczególnych kategorii przestępstw pozostaje zbliżony niezależnie od klasyfikatora,** co sugeruje względnie spójne egzekwowanie prawa w skali całego miasta.
- **Możliwe są jednak wyjątki dla konkretnych przestępstw lub obszarów,** jednak ze względu na zbyt wysoką agregację danych na poziomie kategorii ogólnych, **nie jesteśmy w stanie ich jednoznacznie wykryć** w tej analizie.
- W związku z tym, **przeprowadzimy pogłębioną analizę jednej z wybranych kategorii przestępstw,** aby sprawdzić, czy w przypadku szczegółowego typu czynu ujawniają się różnice w skuteczności zatrzymań między wardami o różnym profilu społeczno-ekonomicznym.

Skuteczność aresztowań według typu przestępstwa z wykorzystaniem przemocy i statusu ekonomicznego wardu.

Skuteczność Aresztowań według typu przestępstwa z wykorzystaniem przemocy i statusu ekonomicznego wardu.

Economic Status ● average ● poor ● rich



- **Najwyższa skuteczność aresztowań dotyczy zabójstw (*HOMICIDE*), co jest zrozumiałe ze względu na priorytet ich ścigania:**
 - W wardach bogatych wynosi aż **59.5%**,
 - w średniozamożnych **43.6%**,
 - a w biednych jedynie **33.9%**.
 - Może to wskazywać na większe zasoby śledcze lub większą skuteczność organów ścigania w zamożniejszych dzielnicach.
- Dla innych przestępstw z użyciem przemocy, takich jak *BATTERY* i *ASSAULT*, skuteczność aresztowań jest znacznie niższa i stopniowo spada wraz z obniżeniem statusu ekonomicznego:
 - *BATTERY*: od **20.7%** w bogatych wardach do **16.5%** w biednych.
 - *ASSAULT*: od **15.6%** (rich) do **13.2%** (poor).
- W przypadku przestępstw seksualnych:

- Skuteczność aresztowań dla *SEX OFFENSE* jest stosunkowo niska i **niezależna od zamożności wardu** (13–14% we wszystkich grupach).
- Dla *CRIMINAL SEXUAL ASSAULT* natomiast istnieje istotna różnica – **8% w wardach średnich i 7.1% w biednych**, ale zaledwie **5% w bogatych**, co może wskazywać na różnice w sposobie klasyfikacji lub podejściu śledczych.
- Najniższa skuteczność aresztowań występuje przy przestępstwach takich jak:
 - *INTIMIDATION* (2.8% w wardach bogatych, 5.9% w biednych),
 - *KIDNAPPING* i *STALKING* również utrzymują się na poziomach poniżej 10%, niezależnie od statusu wardu.

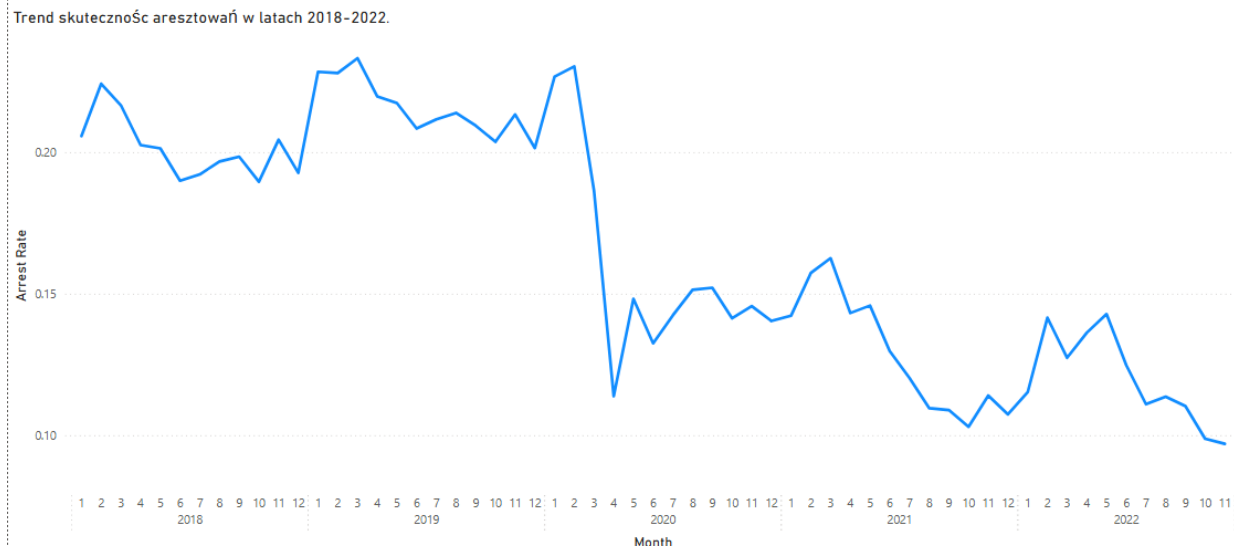
Podsumowanie:

W przypadku przestępstw z wykorzystaniem przemocy:

- **Status ekonomiczny dzielnicy wpływa na skuteczność aresztowań**, szczególnie przy poważniejszych przestępstwach takich jak *HOMICIDE*.
- W wielu przypadkach **im biedniejszy ward, tym niższy odsetek zatrzymań**, co może wynikać z mniejszych zasobów policji, słabszej współpracy społecznej lub większego obciążenia organów ścigania.
- Różnice w aresztowaniach mogą również sugerować **nierówności systemowe** i warto byłoby je pogłębić, analizując np. czas dochodzenia, wskaźnik nierozwiązanych spraw lub obecność monitoringu miejskiego.

Analiza 9: Skuteczność aresztowań na przestrzeni lat.

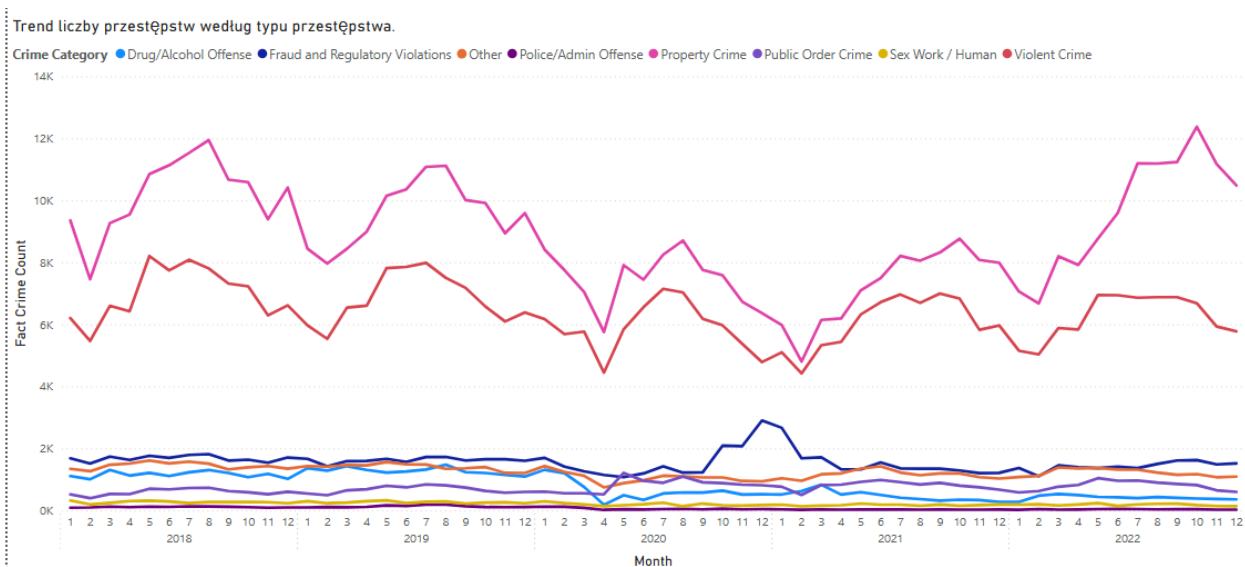
W ramach analizy 9 śledzono zmiany wskaźnika skuteczności aresztowań (Arrest Rate) w kolejnych latach, ze szczególnym uwzględnieniem wpływu pandemii COVID-19 oraz zmian w strukturze przestępstw. Celem było określenie, czy zaobserwowane spadki skuteczności wynikają ze zmian systemowych, operacyjnych lub strukturalnych.



Widzimy gwałtowny spadek skuteczności aresztowań w kwietniu roku 2020, wraz z początkiem kwarantanny spowodowanej przez COVID-19.

Hipotezy do pogłębienia:

- Gwałtowny spadek w 2020 roku może być związany z:
 - **ograniczeniami pandemicznymi**, wpływającymi na operacyjność służb porządkowych,
 - **zmianami w typach przestępstw** – np. wzrostem drobnych wykroczeń, gdzie nie dochodzi do aresztu,
 - **zmniejszoną liczbą interwencji bezpośrednich** lub obostrzeniami w zatrzymywaniu.

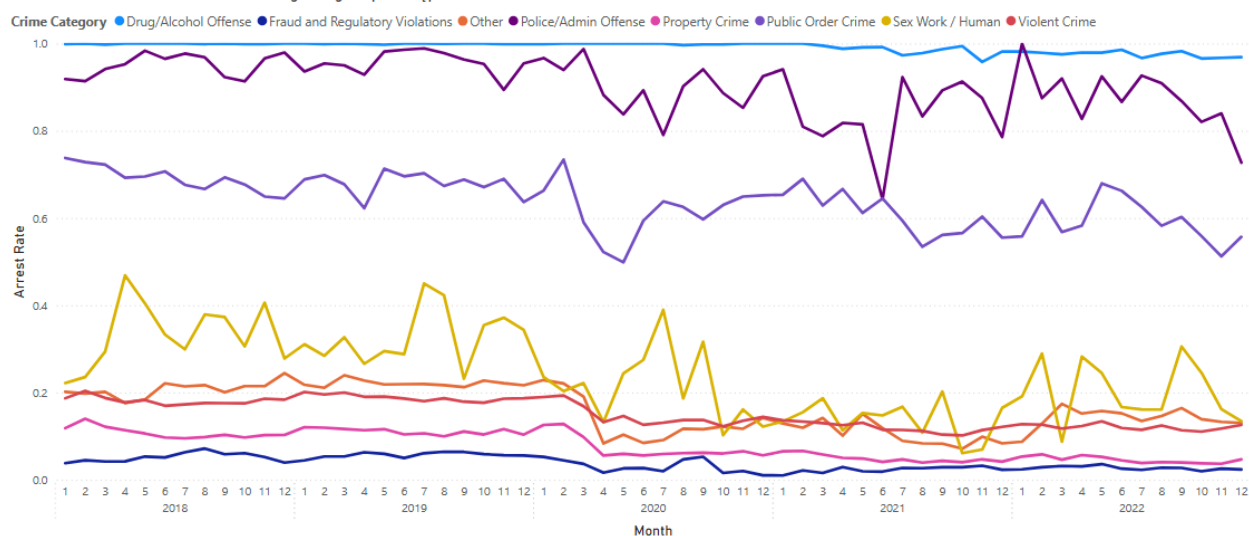


Wnioski:

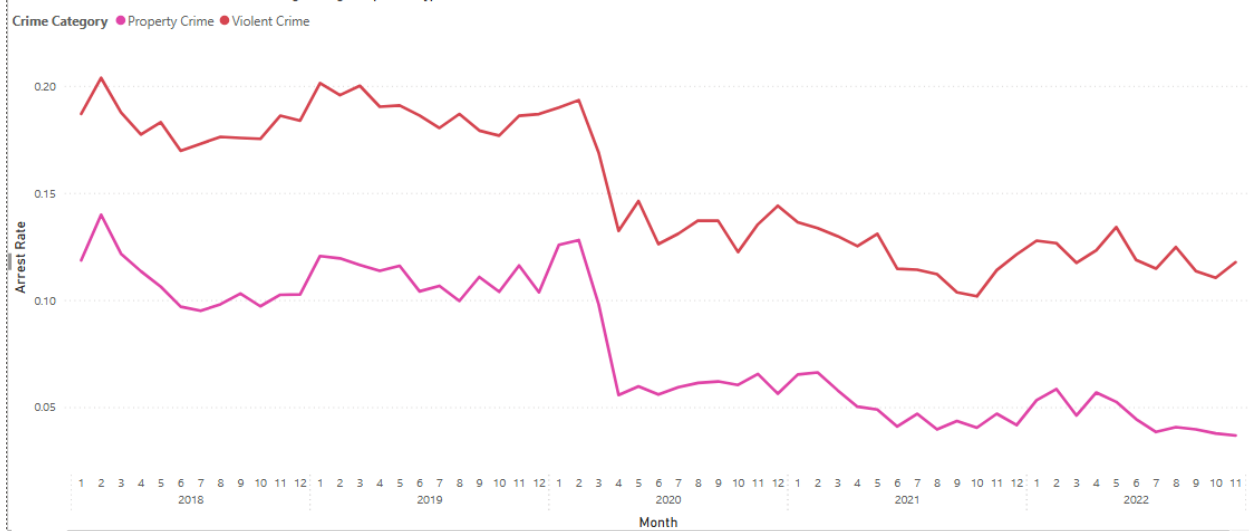
- W kwietniu i maju 2020 roku rzeczywiście zauważalny jest **znaczny spadek liczby przestępstw** w większości kategorii (np. *Violent Crime*, *Drug/Alcohol Offense*, *Sex Work*), co pokrywa się z początkową fazą pandemii COVID-19 i wprowadzonymi restrykcjami.
- Po tych miesiącach większość kategorii wraca stopniowo do poziomów sprzed pandemii, a liczba przestępstw w **drugiej połowie 2020 r. oraz 2021 r.** pozostaje względnie stabilna.
- **Nie widać jednak znaczącej zmiany struktury typów przestępstw** – proporcje między kategoriami (np. *Violent*, *Property*, *Fraud*) pozostają dość stałe na przestrzeni lat.

Przeanalizujemy zatem skuteczność na przestrzeni lat według typu przestępstwa.

Trend skuteczności aresztowań według kategorii przestępstwa.



Trend skuteczności aresztowań według kategorii przestępstwa.



1. Przestępstwa z użyciem przemocy (Violent Crime)

W latach 2018–2019 skuteczność aresztowań w tej kategorii utrzymywała się na stabilnym poziomie, zazwyczaj w przedziale 17–20%. Począwszy od marca 2020 roku, wraz z początkiem pandemii COVID-19, zaobserwowano gwałtowny spadek skuteczności:

- Kwiecień 2020: 13,2%
- Czerwiec 2020: 12,6%
- Lipiec 2021: 11,4%
- Październik 2021: najniższy odnotowany poziom – 10,1%

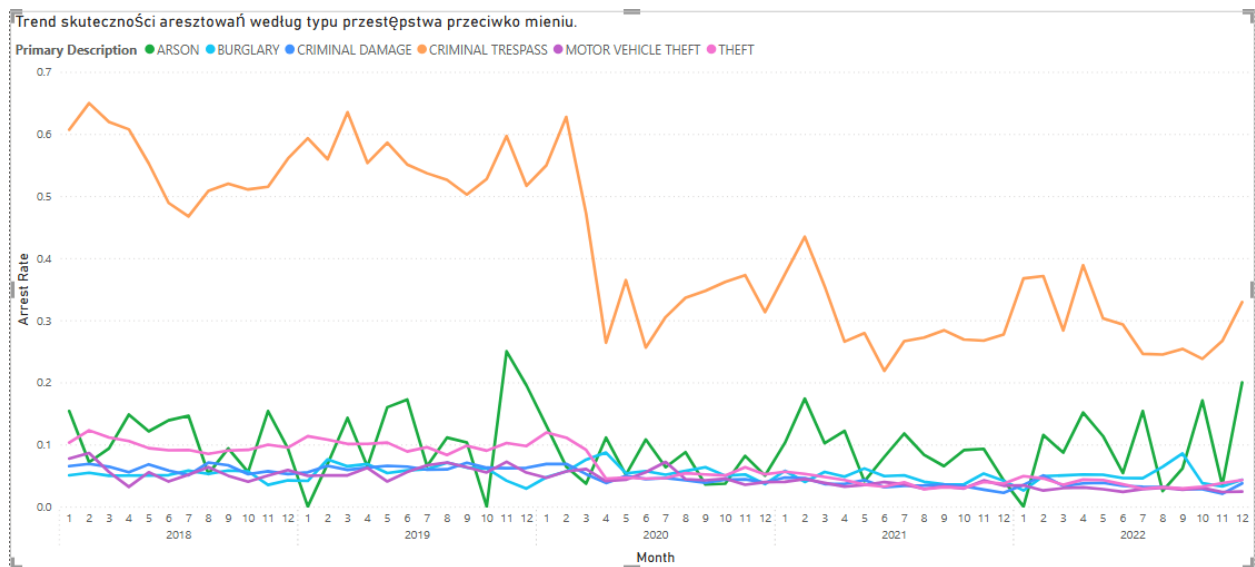
W kolejnych miesiącach wskaźnik ten stopniowo wzrastał, jednak do końca 2022 roku nie powrócił do poziomów sprzed pandemii.

2. Przestępstwa przeciwko mieniu (Property Crime)

Wskaźnik skuteczności aresztowań w tej kategorii przestępstw w latach 2018–2019 oscylował w granicach 10–12%. Od marca 2020 nastąpił gwałtowny spadek:

- Kwiecień 2020: 5,6%
- Czerwiec 2020: 5,5%
- W latach 2021–2022: stabilizacja na niskim poziomie, w przedziale 3,7–6%

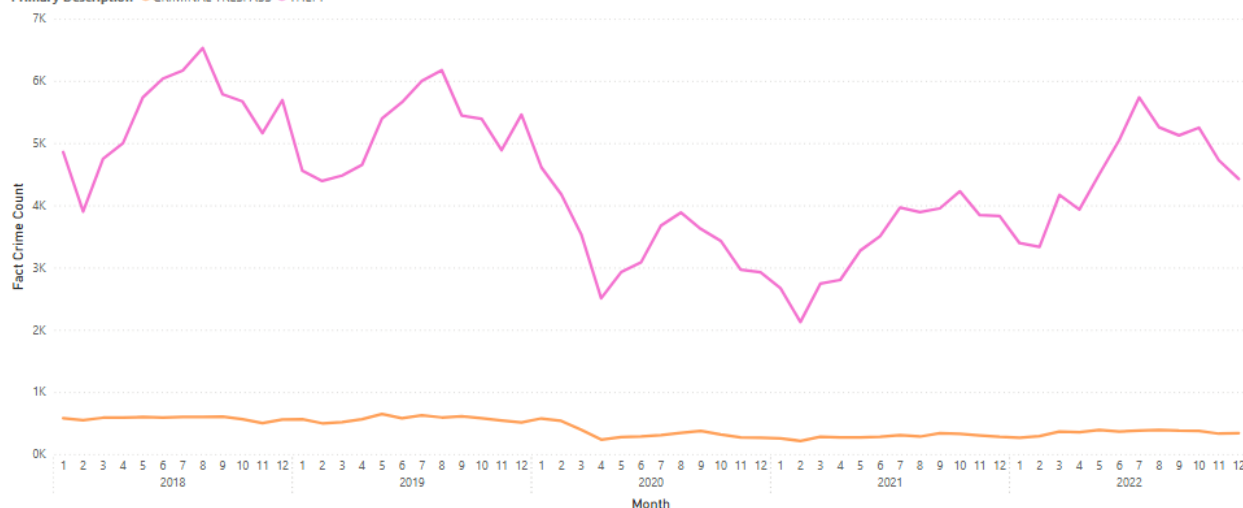
Spadek ten był nie tylko gwałtowny, ale także długotrwały.



W latach 2018–2019 skuteczność aresztowań dla przestępstwa *Criminal Trespass* utrzymywała się na relatywnie wysokim poziomie – najczęściej przekraczając 50%. Od początku 2020 roku obserwujemy jednak wyraźny i trwały spadek tego wskaźnika. W kolejnych miesiącach rzadko przekraczał on 35%, a okresowo spadał nawet poniżej 25%.

Porównanie liczby kradzieży z liczbą wtargnięć na przestrzeni lat 2018-2022.

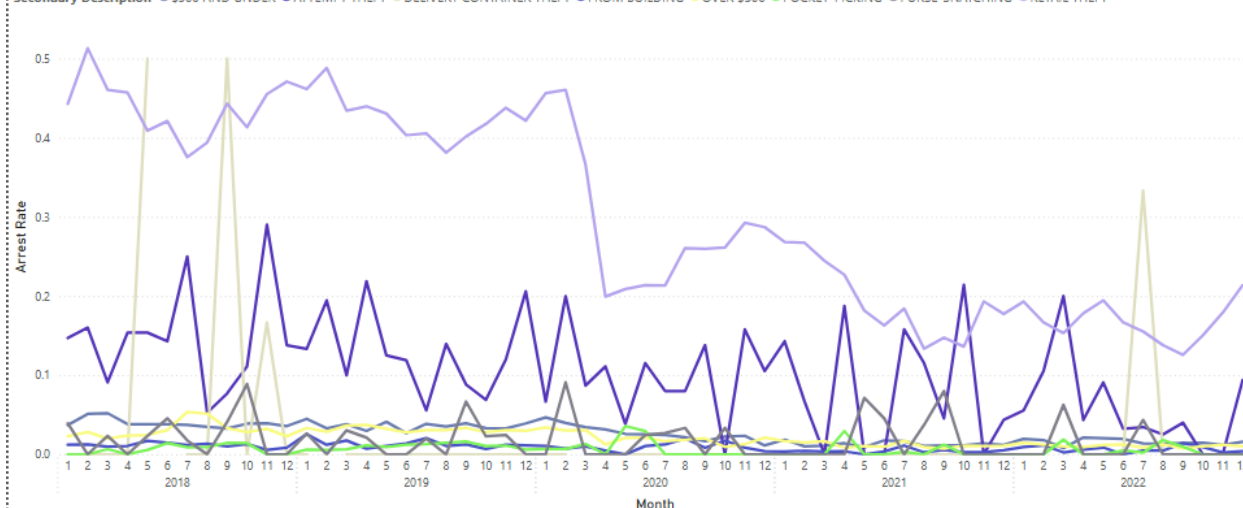
Primary Description ● CRIMINAL TRESPASS ● THEFT



Pomimo tej zmiany, *Criminal Trespass* stanowi jedynie niewielki odsetek ogółu przestępstw przeciwko mieniu. Dominującą kategorią w tej grupie są kradzieże (*Theft*), które co miesiąc występowały w liczbie kilku tysięcy przypadków. To właśnie w tej kategorii, szczególnie w podtypie *Retail Theft*, odnotowano najbardziej zauważalny spadek skuteczności aresztowań. Jeszcze w 2019 roku wskaźnik ten regularnie przekraczał 40%, natomiast od 2020 roku sukcesywnie malał – w niektórych miesiącach spadając nawet poniżej 20%.

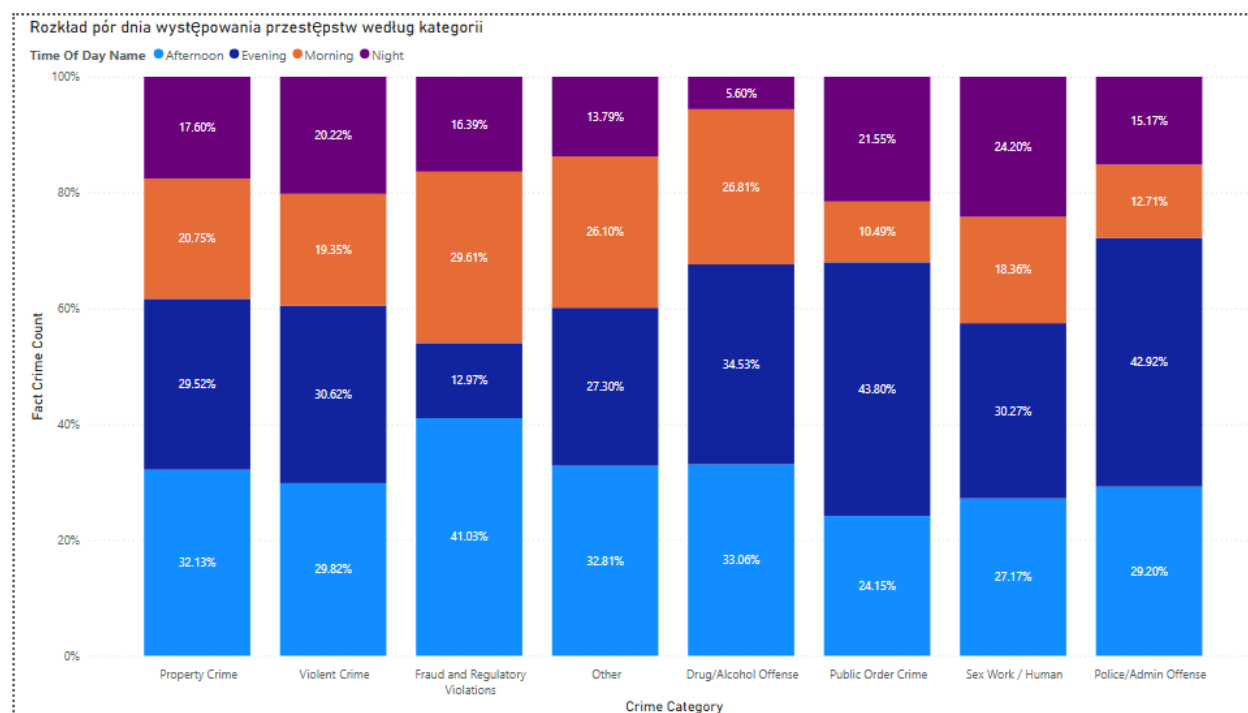
Trend skuteczność aresztowań w przypadku i kradzieży.

Secondary Description ● \$500 AND UNDER ● ATTEMPT THEFT ● DELIVERY CONTAINER THEFT ● FROM BUILDING ● OVER \$500 ● POCKET-PICKING ● PURSE-SNATCHING ● RETAIL THEFT



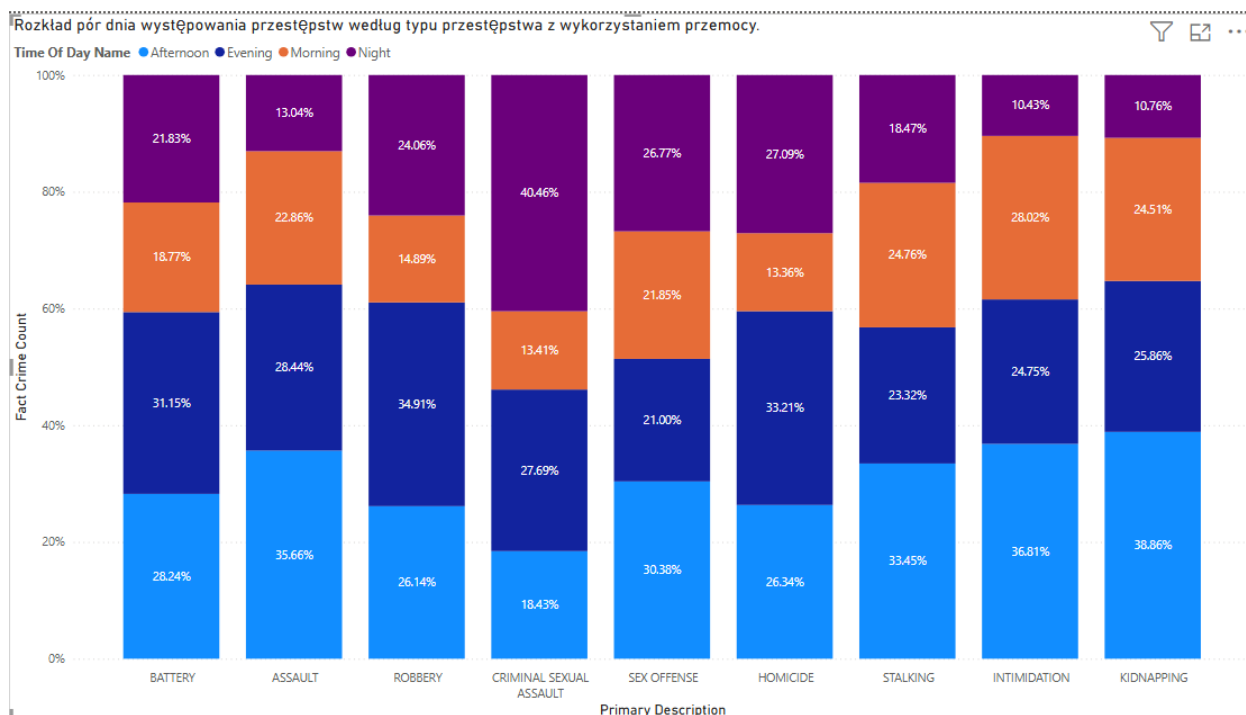
W związku z powyższym można stwierdzić, że pomimo spadku skuteczności dla *Criminal Trespass*, to właśnie *Retail Theft* miało decydujący wpływ na ogólny trend obniżenia skuteczności aresztowań dla przestępstw przeciwko mieniu. Przyczyn tego spadku można upatrywać w ograniczeniach operacyjnych policji podczas pandemii, zmianach priorytetów działań służb, a także możliwym zmniejszeniu nacisku na ściganie drobnych kradzieży sklepów w tym okresie.

Analiza 10: Kategoria przestępstwa w zależności od pory dnia.



Na podstawie danych o liczbie przestępstw z podziałem na porę dnia i kategorię przestępstwa, można zaobserwować wyraźne różnice w rozkładzie czasowym dla poszczególnych typów czynów:

- **Przestępstwa przeciwko mieniu (Property Crime)** najczęściej popełniane są **po południu (33,5%)** oraz **wieczorem (30,8%)**, łącznie te dwie pory dnia stanowią **ponad 64%** wszystkich tego typu przestępstw. Rano występuje 21,6%, a w nocy – 18,4%.
- **Przestępstwa z użyciem przemocy (Violent Crime)** rozkładają się bardziej równomiernie, jednak również dominują **godziny popołudniowe (30,8%)** i **wieczorne (31,7%)**, które razem stanowią **ponad 62%** przypadków. Rano i w nocy – po ok. 19% każda.
- **Przestępstwa narkotykowe i alkoholowe** są najbardziej skoncentrowane **wieczorem (33,7%)** i **po południu (32,3%)**, natomiast **w nocy (5,5%)** są relatywnie rzadkie.
- **Prostytucja i handel ludźmi** mają największy udział **wieczorem (30,6%)**, ale także **w nocy (24,5%)** – co wyróżnia je spośród innych kategorii.
- **Przestępstwa porządkowe i administracyjne** również dominują **wieczorem (37%)**, ale występują stosunkowo rzadko o poranku (11%).



Analiza danych ujawnia ciekawe zróżnicowanie typów przestępstw względem pory dnia. Niektóre przestępstwa wykazują silną koncentrację w konkretnych godzinach doby:

- **Criminal Sexual Assault** (gwałt) najczęściej popełniany jest **w nocy** – aż **3152 przypadki**, co stanowi **ok. 39%** wszystkich zgłoszonych tego typu przestępstw. W porównaniu: wieczorem – 2157 (27%), po południu – 1436 (18%), rano – 1045 (13%). To najwyraźniej **najbardziej nocna kategoria przestępstw z użyciem przemocy**.
- **Battery** (pobicie) dominuje **wieczorem (69241)** i **po południu (62783)**, co łącznie stanowi **ponad 60%** przypadków. Mniej zdarzeń odnotowano w nocy (48536) i rano (41731), co sugeruje większą intensywność tego typu incydentów w godzinach aktywności społecznej.
- **Robbery** (rozbój) również osiąga szczyt **wieczorem (14806)** oraz **w nocy (10206)** – łącznie to **ponad 54%** przypadków. To przestępstwo wyraźnie nasila się w późnych godzinach.
- **Homicide** (zabójstwo) rozkłada się mniej równomiernie – najwięcej przypadków zgłaszanych jest **wieczorem (1146)** i **w nocy (935)**, co łącznie stanowi około **60%** wszystkich zdarzeń.
- **Stalking** (nękanie) oraz **Intimidation** (groźby) częściej mają miejsce **w ciągu dnia**, szczególnie **po południu** – odpowiednio 489 i 293 przypadki.

Wnioski ogólne

- Przestępstwa o charakterze seksualnym oraz te najbardziej brutalne (jak rozboje i zabójstwa) mają tendencję do występowania **wieczorem i w nocy**, kiedy aktywność społeczna odbywa się przy mniejszym nadzorze i ograniczonym dostępie do pomocy publicznej.
- Przestępstwa takie jak stalking, zastraszanie i porwania wykazują **bardziej dzienny charakter**, co może wskazywać na ich powiązanie z relacjami osobistymi lub zawodowymi.
- Dane potwierdzają intuicję, że **noc to pora największego ryzyka dla przestępstw o charakterze seksualnym**, co może być istotne dla działań prewencyjnych oraz planowania patroli policyjnych.

5. KPI + Zestawienia

W ramach projektu zaprojektowano wskaźnik KPI (Key Performance Indicator) służący do monitorowania udziału przestępstw o charakterze domowym w ogólnej liczbie przestępstw – tzw. **Domestic Crime Rate**. KPI ten został wyliczony jako iloraz liczby przestępstw oznaczonych flagą *domestic* oraz całkowitej liczby przestępstw w danej grupie.

Dla potrzeb interpretacji KPI przyjęto następującą skalę wartości progowych:

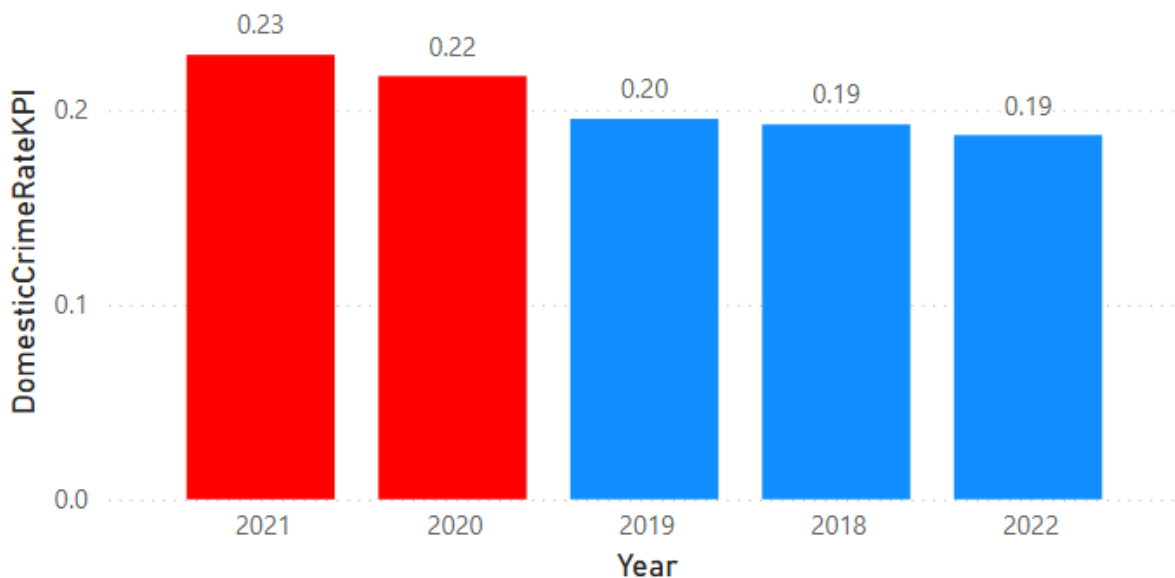
- **Wartość KPI ≤ 0.10** – *Status* = 1 (zielony): Poziom domowej przestępczości uważany za niski
- **$0.10 < \text{KPI} \leq 0.20$** – *Status* = 0 (żółty): Poziom umiarkowany
- **KPI > 0.20** – *Status* = -1 (czerwony): Poziom wysoki, potencjalnie niepokojący

Przykładowe zestawienie: Domestic Crime Rate w podziale na typ dzielnicy (Age Type) i dominującą grupę etniczną (Dominant Ethnicity)

Age Type	Dominant Ethnicity	DomesticCrimeRateKPI	DomesticCrimeRateKPI Goal	DomesticCrimeRateKPI Status
balanced	Black	0.25	0.10	○
balanced	Hispanic	0.21	0.10	○
balanced	White	0.13	0.10	●
elderly-majority	Black	0.27	0.10	○
elderly-majority	White	0.15	0.10	●
young-majority	Hispanic	0.19	0.10	●
young-majority	White	0.08	0.10	●
Total		0.20	0.10	○

Przykładowe zestawienie: Domestic Crime Rate według roku:

DomesticCrimeRateKPI by Year



6. Wnioski z Analizy Danych:

W oparciu o kostkę OLAP zbudowaną w SSAS przeprowadzono szereg wielowymiarowych analiz, które umożliwiły uzyskanie pogłębionego wglądu w strukturę i dynamikę przestępczości w Chicago w latach 2018–2022. Integracja danych faktowych z wymiarami demograficznymi, geograficznymi i czasowymi umożliwiła identyfikację zależności, które byłyby trudne do zauważenia w klasycznych, jednowymiarowych raportach.

Główne obserwacje:

- **Profil przestępczości zależy od statusu ekonomicznego wardu** – w dzielnicach biedniejszych dominują przestępstwa z użyciem przemocy i narkotykowe, podczas gdy w zamożnych przeważają oszustwa finansowe i wykroczenia rejestrowane jako „podejrzane substancje”. To sugeruje odmienne wzorce przestępczości oraz potencjalne różnice w działaniach prewencyjnych lub policyjnych.
- **Dominująca grupa etniczna i wiekowa wpływa na strukturę przestępstw** – wardy o przeważającej populacji młodych mieszkańców charakteryzują się dużym udziałem kradzieży i przestępstw cyfrowych, natomiast w wardach starszych i czarnoskórych częściej występują przestępstwa domowe.

- **Czas i miejsce mają znaczenie** – wiele typów przestępstw wykazuje wyraźną sezonowość (np. wyższe wskaźniki latem) i koncentrację w określonych porach dnia. Np. przestępstwa seksualne dominują nocą, a kradzieże – po południu.
- **Pandemia COVID-19 miała wyraźny wpływ na skuteczność działań policji** – w 2020 roku nastąpił znaczący spadek skuteczności aresztowań, szczególnie w kategoriach przestępstw przeciwko mieniu i z użyciem przemocy. W kolejnych latach trend ten nie wrócił do poziomów sprzed pandemii.
- **Skuteczność aresztowań zależy od rodzaju przestępstwa, ale także od dzielnic** – największą skuteczność wykazują kategorie przestępstw wykrywanych bezpośrednio (np. narkotyki), podczas gdy dla przestępstw złożonych (np. oszustwa, przemoc seksualna) odsetek aresztowań jest zdecydowanie niższy – szczególnie w wardach biedniejszych.
- **Różnice w klasyfikacjach i podejściu śledczym** mogą powodować znaczne rozbieżności w statystykach – np. w kategorii Human Trafficking odnotowano przypadki wyłącznie w wardach zamożniejszych, co może wskazywać na różne podejście lub niedostateczne ujawnianie tego typu przestępstw w innych rejonach.