

Grupa G	266469	266839	
------------	--------	--------	--

MINI-PROJEKT

REALIZACJA W GRUPACH

MINI-PROJEKT 1 – ANALIZA DZIEDZINY I PROJEKT MODELU WIELOWYMIAROWEGO

Opracowanie dobrego systemu BI wymaga właściwego zrozumienia potrzeb biznesowych i rzeczywistości (obszaru problemowego – domeny problemu), a także dostępnych zasobów danych. Kluczowe jest zatem uświadomienie sobie sposobu postrzegania rzeczywistości danych przez użytkowników, tzw. model korporacyjny danych, na określonym (wybranym) odcinku rzeczywistości. Umożliwieni to identyfikację (określenie) potrzeb, celu i możliwości analizy biznesowej wspierającej procesy decyzyjne (podejmowanie właściwych decyzji biznesowych). Należy pamiętać, iż skupiamy się tutaj na wielowymiarowej analizie danych, u podstaw której leży możliwość analizy podsumowań i zestawień pomiarów (realizowanych na danych operacyjnych) w kontekście różnych kontekstów (wymiarów analizy).

ZADANIE – ANALIZA DZIEDZINY I CEL PROJEKTU – ROZWIĄZANIA:

1.1 TYTUŁ PROJEKTU

Platforma analizy danych dotyczących strzelanin w USA

1.2 OGÓLNY OPIS DOMENY

Proszę krótko przedstawić wybraną domenę. Opis powinien być biznesowy, nie techniczny. Powinien wprowadzać do obszaru rzeczywistości i umożliwić zrozumienie podstawowych pojęć i terminów domeny.

W Stanach Zjednoczonych często dochodzi do aktów działalności przestępczej, Akty te rozumiane są jako działanie przestępcze mające na celu pozbawienie kogoś np. życia, zdrowia, majątku lub praw, przy użyciu broni palnej (rodzaj broni miotającej pociski energią gazów powstałych ze spalania ładunku miotającego). W związku z naturą tych zdarzeń, są one problematyczne i wymagają podjęcia odpowiednich decyzji, zarówno tych w odpowiedzi na zdarzenie jak i tych zapobiegawczych. Ilość przestępstw dokonywanych z bronią jest powiązana z dostępnością broni palnej, która jest regulowana prawnie przez każdy ze stanów niezależnie.

1.3 OPIS OBSZARU ANALIZY WRAZ Z UZASADNIENIEM

Proszę podać i skomentować wybrany fragment domeny przeznaczonej dla systemu BI. Skoncentruj się na identyfikacji i wprowadzeniu podstawowych tematów biznesowych (procesów, zdarzeń).

Procesy biznesowe:

1. **Zarządzanie danymi incydentów:** Gromadzenie, przechowywanie i przetwarzanie danych o incydentach z bronią palną.
2. **Analiza ofiar i podejrzanych:** Analiza liczby ofiar i podejrzanych, w tym rannych i zabitych.
3. **Analiza geograficzna:** Analiza lokalizacji incydentów, co pozwala na identyfikację obszarów o wysokiej przestępczości.
4. **Przestrzeganie prawa:** Analiza skuteczności prawa związanego z posiadaniem broni.

Zdarzenia:

5. **Wystąpienie incydentu:** Rejestracja nowego incydentu z użyciem broni palnej.

6. **Raportowanie liczby ofiar i podejrzanych:** Aktualizacja danych dotyczących liczby ofiar i podejrzanych.
7. **Analiza lokalizacji incydentu:** Analiza miejsca wystąpienia incydentu, co pomaga w planowaniu działań prewencyjnych.
8. **Ocena legalności posiadania broni:** Ocena czy posiadanie broni bez zezwolenia w danym stanie jest legalne

1.4 DOCELOWI UŻYTKOWNICY

Zidentyfikuj podstawowych użytkowników (typy) systemu BI - podaj krótki opis każdego typu i określ oczekiwania każdego typu. Nie wybieraj zbyt wielu typów – ponieważ korzystne może być skupienie się na większej ilości szczegółów dla 1-3 typów użytkowników.

1. Biura podróży - biura podróży (zarówno z USA jak i poza) oferujące wycieczki do rejonów na terenie USA - ogólne dane statystyczne i rankingowe co do popełnianych przestępstw z podziałem na rejony

2. Mieszkańcy - mieszkańcy danego rejonu - ogólne dane statystyczne co do popełnianych przestępstw w ich rejonie

3. Służby medyczne - służby medyczne - dane statystyczne co do ofiar śmiertelnych i osób rannych, które ucierpiały w strzelaninie

1.5 ZIDENTYFIKOWANE PROCESY BIZNESOWE W DOMENIE

Proszę podać i skomentować zidentyfikowane procesy w założonej domenie. Spróbuj uzasadnić znaczenie możliwości analizy danych z tych procesów dla docelowych typów użytkowników (użyj prostej skali low, medium i high impact; uzasadnij).

9. **Zarządzanie danymi incydentów:** Gromadzenie, przechowywanie i przetwarzanie danych o incydentach z bronią palną.
10. **Analiza ofiar i podejrzanych:** Analiza liczby ofiar i podejrzanych, w tym rannych i zabitych.
11. **Analiza geograficzna:** Analiza lokalizacji incydentów, co pozwala na identyfikację obszarów o wysokiej przestępczości.
12. **Przestrzeganie prawa:** Analiza skuteczności prawa związanego z posiadaniem broni.

Zidentyfikowane procesy

13. Zarządzanie danymi incydentów
14. Analiza ofiar i podejrzanych
15. Analiza geograficzna
16. Przestrzeganie prawa

Komentarz do zidentyfikowanych procesów

1. Zarządzanie danymi incydentów

Opis: Proces obejmujący gromadzenie, przechowywanie, aktualizowanie i weryfikację danych dotyczących incydentów z użyciem broni palnej.

Znaczenie dla użytkowników:

- **Organy ścigania:** High impact – precyzyjne dane są kluczowe do prowadzenia dochodzeń i strategii prewencyjnych.
- **Analizy statystyczne i polityki publiczne:** High impact – rzetelne dane są niezbędne do tworzenia skutecznych polityk publicznych oraz oceny ich efektów.
- **Media i społeczeństwo:** Medium impact – dostęp do dokładnych danych zwiększa transparentność i świadomość społeczną.

2. Analiza ofiar i podejrzanych

Opis: Proces polegający na analizie danych dotyczących liczby i charakterystyki ofiar oraz podejrzanych w incydentach z użyciem broni palnej.

Znaczenie dla użytkowników:

- **Organy ścigania:** High impact – pozwala na identyfikację wzorców przestępczości oraz planowanie działań operacyjnych.
- **Instytucje zdrowia publicznego:** High impact – analiza wpływu przemocy z użyciem broni na zdrowie publiczne, planowanie interwencji medycznych.
- **Naukowcy i badacze:** High impact – dostarcza danych do badań nad przyczynami i skutkami przemocy z użyciem broni.

3. Analiza geograficzna

Opis: Proces obejmujący analizę lokalizacji incydentów z użyciem broni palnej w celu identyfikacji obszarów o wysokiej przestępczości.

Znaczenie dla użytkowników:

- **Organy ścigania:** High impact – identyfikacja "hotspotów" przestępczości pozwala na efektywniejsze rozmieszczenie zasobów i działania prewencyjne.
- **Władze lokalne:** High impact – informacje o koncentracji przestępczości wspierają planowanie urbanistyczne i interwencje społeczne.
- **Społeczność i organizacje pozarządowe:** Medium impact – zwiększa świadomość lokalnej społeczności i wspiera działania prewencyjne.

4. Przestrzeganie prawa

Opis: Proces obejmujący ocenę czy broń użyta w incydencie była legalnie posiadana zgodnie z lokalnymi przepisami.

Znaczenie dla użytkowników:

- **Organy ścigania:** High impact – informacje o legalności posiadania broni są kluczowe w dochodzeniach i prewencji.
- **Prawodawcy:** High impact – dane te są istotne przy tworzeniu lub modyfikacji przepisów dotyczących posiadania broni.
- **Społeczność:** Medium impact – zwiększa świadomość na temat skuteczności i egzekwowania prawa dotyczącego broni.

1.6 CEL PROJEKTU

Przedstaw szczegółowe potrzeby użytkowników poprzez określenie 10-15 typów zapytań OLAP. Zapytania sformułuj biznesowo, nie są to zapytania SQL, MDX, DAX, czy inne. Uwaga – postaraj się, aby pytania były odpowiednio różnego typu (nie zmieniaj jedynie wartości atrybutu, np. roku z 2019 na 2020). Dla każdego zapytania/potrzeby dopasuj zidentyfikowany typ użytkowników oraz wskaż, w jaki sposób użytkownik może wykorzystać wynikowe informacje. Dodatkowo dla każdego zapytania określ elementy składowe zapytania – fakt, miary, wymiary. Wyniki przedstaw w punkcie 1.6.1.

Wykonaj podstawową analizę zapytań (1.6.1), spróbuj wywnioskować pewne ogólne wymagania użytkownika. Skoncentruj się na identyfikacji zdarzeń i perspektyw, którymi interesuje się użytkownik. Ponadto wykonaj podstawową analizę zapytań użytkowników (1.6.1), spróbuj wywnioskować które elementy składowe są kluczowe dla wymagań użytkownika. W szczególności, podsumuj wszystkie możliwe fakty, miary i wszystkie możliwe indywidualne wymiary - wyniki przedstaw w sekcji 1.6.2.

1.6.1 OCZEKIWANIA I SZCZEGÓŁOWE POTRZEBY W ZAKRESIE WSPARCIA DECYZYJNEGO

1. Liczba incydentów z bronią palną w podziale na miesiące w danym roku

- a. **Kontekst:** Analiza sezonowych trendów przestępczości.
- b. **Znaczenie:** Organy ścigania mogą identyfikować okresy o zwiększonej aktywności przestępczej i planować odpowiednie działania prewencyjne.

2. Średnia liczba ofiar na incydent w podziale na stany

- c. **Kontekst:** Porównanie skutków incydentów między różnymi stanami.
- d. **Znaczenie:** Władze stanowe mogą ocenić efektywność lokalnych strategii przeciwdziałania przemocy.

3. Liczba incydentów w miastach powyżej 500,000 mieszkańców vs. poniżej 500,000 mieszkańców

- e. **Kontekst:** Analiza wpływu wielkości populacji na przestępczość.
- f. **Znaczenie:** Organy ścigania i władze lokalne mogą dostosować działania do specyfiki różnych obszarów miejskich.

4. Procent incydentów, przed i po zmianie prawa

- g. **Kontekst:** Ocena przestrzegania prawa dotyczącego posiadania broni.
- h. **Znaczenie:** Prawodawcy i organy ścigania mogą lepiej rozumieć skuteczność przepisów prawnych.

5. Trendy liczby aresztowań podejrzanych na przestrzeni ostatnich 5 lat

- i. **Kontekst:** Analiza zmian w egzekwowaniu prawa.
- j. **Znaczenie:** Organy ścigania mogą ocenić skuteczność swoich działań i wprowadzać odpowiednie korekty.

6. Porównanie liczby incydentów i liczby ofiar w dużych miastach vs. małych miastach w danym stanie

- k. **Kontekst:** Analiza wpływu urbanizacji na przestępczość.
- l. **Znaczenie:** Władze lokalne mogą lepiej rozumieć specyfikę przestępczości w różnych środowiskach.

7. Liczba incydentów z bronią palną w podziale na dni tygodnia

- m. **Kontekst:** Identyfikacja dni o najwyższej przestępczości.
- n. **Znaczenie:** Organy ścigania mogą zwiększyć obecność policji w dniach o wyższej aktywności przestępczej.

8. Średnia liczba rannych ofiar na incydent w podziale na lata

- o. **Kontekst:** Monitorowanie zmian w nasileniu przemocy.
- p. **Znaczenie:** Instytucje zdrowia publicznego mogą lepiej planować zasoby medyczne i programy wsparcia dla ofiar.

9. Liczba incydentów z bronią palną w podziale na przedziały czasowe w ciągu doby

- q. **Kontekst:** Identyfikacja godzin o najwyższej przestępczości.
- r. **Znaczenie:** Organy ścigania mogą zwiększyć patrole w godzinach szczytu przestępczości.

10. Procent incydentów z aresztowaniami podejrzanych w podziale na miasta

- s. **Kontekst:** Ocena skuteczności działań policji w różnych miastach.

- t. **Znaczenie:** Organy ścigania mogą analizować i optymalizować swoje procedury operacyjne.

11. Analiza trendów liczby incydentów z bronią palną przed i po wprowadzeniu nowych przepisów dotyczących broni

- u. **Kontekst:** Ocena wpływu zmian legislacyjnych na przestępczość.
v. **Znaczenie:** Prawodawcy mogą ocenić skuteczność nowych przepisów i wprowadzać odpowiednie modyfikacje.

1.6.2 ZAKRES ANALIZY – BADANE ASPEKTY

- Wymiar czasowy incydentu (dzień, miesiąc, rok)
- Zakres czasowy wprowadzenia prawa (rok wprowadzenia, rok zmiany)
- Wymiar dotyczący lokalizacji (miasto, stan)
- Dane dotyczące aresztowań (czy aresztowano, czy zabito)

ZADANIE – ANALIZA ŹRÓDEŁ DANYCH – ROZWIĄZANIA:

Przygotuj krótki opis wybranych źródeł danych - skup się na ogólnej strukturze, opcjach dostępu, charakterystyce aktualizacji (zmienność danych), objętości (rozmiar, liczba rekordów), liczbie atrybutów, przedziale czasowym dla faktów itp. Następnie przeprowadź wstępną ocenę jakości dostępnych danych – skup się na liczbie rekordów, liczbie unikalnych rekordów, liczbie brakujących rekordów oraz skróconej ocenie zastosowania/przydatności.

2.1 LOKALIZACJA, FORMAT, DOSTĘPNOŚĆ

Pliki xlsx:

State-Level Estimates of Household Firearm Ownership.xlsx

(<https://www.rand.org/pubs/tools/TL354.html>)

Gunviolences.xlsx

(<https://www.gunviolencearchive.org/>)

OpenWeather OneCallAPI

(<https://openweathermap.org/api/one-call-3>)

2.2 WSTĘPNA OCENA (LICZBA REKORDÓW, PRZEDZIAŁ CZASOWY DLA FAKTÓW, ITP.)

State-Level Estimates of Household Firearm Ownership.xlsx

1. Liczba rekordów: 1850
2. Liczba kolumn: 20
3. Zakres czasowy prezentowanych danych: 1980-2016 (dane zbierane rok do roku)

Gunviolences.xlsx

1. Liczba rekordów: 4681
2. Liczba kolumn: 11
3. Zakres czasowy prezentowanych danych: 01.01.2014- 31.12.2023

OpenWeather OneCallAPI

1. Dane: historyczne od 1979 roku
2. Forma: Dzielne podsumowania pogodowe dla daty i lokalizacji
3. Liczba kolumn/trybutów: 11
4. REST API – komunikacja poprzez protokół HTTP/S
5. 1000/requests per dzień

2.3 OCENA PRZYDATNOŚCI DANYCH

2.3.1 OCENA JAKOŚCI ŹRÓDŁA DANYCH

	Plik	Liczba wierszy	Liczba trybutów	Rozmiar	Zmienność	Ziarno	Ocena
1	State-Level Estimates of Household Firearm Ownership.xlsx	1851	8	260KB	Brak	1 zestaw metryk dostępu do broni w stanie w danym czasie	Dobre źródło danych, dużo wskaźników
2	Gunviolences.xlsx	4681	11	387KB	Jeżeli weźmiemy pod uwagę datę, to na jedną datę przypada około 3 incydentów	1 zarejestrowana w mediach strzelanina	Dobre źródło danych, aby używać nie trzeba obrabiać danych
3	OpenWeather OneCallAPI	NA	11	NA	Dziennie	1 podsumowanie dzienne pogody w danym dniu	Ograniczenia w dostępie do danych (1000 zapytań na dzień), dane częściowo wybrakowane

Komentarz ...

2.3.2 OCENA JAKOŚCI ATRYBUTÓW:

	Plik	Nazwa atr.	Skala pomiarowa (typ trybutu)	Typ danych	Liczba wart. o zduplik. danych	Unikalnych wartości	Pustych wartości	Ocena (jakości/zastosowania)
1	State-Level Estimates of Household Firearm Ownership.xlsx	FIP	Porządkowy	Integer	1800	50	0	Dane nie wymagają dod. Przetwarzania – pozwala unikalnie zidentyfikować wpis FIP
2		STATE	Tekstowy	String	1800	50	0	Dane nie wymagają dod. Przetwarzania
3		HFR	Zmiennoprzecinkowy	Float	1288	562	0	Dane nie wymagają dod. Przetwarzania
4		HFR_se	Zmiennoprzecinkowy	Float	1835	15	0	Dane nie wymagają dod. Przetwarzania

5		universl	Prawda/Fałsz	Boolean	1848	2	0	Dane nie wymagają dod. Przetwarzania
6		permit	Prawda/Fałsz	Boolean	1848	2	0	Dane nie wymagają dod. Przetwarzania
7		FEM_FS_S	Zmiennoprzecinkowy	Float	553	1297	0	Dane nie wymagają dod. Przetwarzania
8		MALE_FS_S	Zmiennoprzecinkowy	Float	147	1703	0	Dane nie wymagają dod. Przetwarzania
9	Gunviolences.xls	IncidentID	Porządkowy	Integer	0	4681	0	Dane nie wymagają dod. Przetwarzania – pozwala unikalnie zidentyfikować incydent
10		Incident Date	Data	Data	2405	2276	0	Dane nie wymagają dod. Przetwarzania
11		State	Tekstowy	String	4630	51	0	Dane nie wymagają dod. Przetwarzania
12		City or Country	Tekstowy	String	3542	1139	0	Dane nie wymagają dod. Przetwarzania
13		Address	Tekstowy	String	60	4621	14	Rozdział na pojedyncze adresy
14		Victims Killed	Numeryczny	Integer	4660	21	0	Dane nie wymagają dod. Przetwarzania
15		Victims Injured	Numeryczny	Integer	4650	31	0	Dane nie wymagają dod. Przetwarzania
16		Suspects Killed	Numeryczny	Integer	4677	4	0	Dane nie wymagają dod. Przetwarzania
17		Suspects Injured	Numeryczny	Integer	4675	6	0	Dane nie wymagają dod. Przetwarzania
18		Suspects Arrested	Numeryczny	Integer	4668	13	0	Dane nie wymagają dod. Przetwarzania
19		Operations	NaN	NaN	0	0	4681	Brak danych
20	OpenWeather OneCallAPI	lat	Numeryczny	Zmiennoprzecinkowa	0	NA	0	Pozwala jasno zidentyfikować lokalację
21		lon	Numeryczny	Zmiennoprzecinkowa	0	NA	0	Pozwala jasno zidentyfikować lokalację,
22		tz	Czasowy	Unix timestamp	0	Wszystkie są unikalne	0	Pozwala jasno zidentyfikować moment w czasie

2 3		date	Data	Data	NA	NA	0	Dane nie wymagają dod. Przetwarzania
2 4		cloud_cover	Numeryczne	Procentowe pokrycie	NA	NA	NA	Dane nie wymagają dod. Przetwarzania
2 5		humidity	Numeryczne	Procentowe	NA	NA	NA	Dane nie wymagają dod. Przetwarzania
2 6		precipitation	Numeryczne	Liczba rzeczywista	NA	NA	NA	Dane nie wymagają dod. Przetwarzania
2 7		pressure	Numeryczne	Liczba rzeczywista	NA	NA	NA	Dane nie wymagają dod. Przetwarzania
2 8		temperature	Numeryczne	Zmiennoprzecinkowa	NA	NA	NA	Potrzebne uzgodnienie jednostek, potrzebne uzgodnienie czasu
2 9		Wind.speed	Numeryczne	Zmiennoprzecinkowa	NA	NA	NA	Dane nie wymagają dod. Przetwarzania
3 0		Wind.direction	Numeryczne	Liczba naturalna (kąty)	NA	NA	NA	Dane nie wymagają dod. Przetwarzania

Komentarz ...

ZADANIE – PROJEKT MODELU – ROZWIĄZANIA:

Przedstaw opracowany model wielowymiarowy. Przedstaw szczegóły modelu – w tym szczegóły dotyczące faktów oraz miar w 3.1 (pamiętaj, aby określić typy oraz metodę agregacji dla wszystkich miar), szczegóły dotyczące wymiarów i ich atrybutów w 3.2 (pamiętaj, aby określić wszystkie wymiary – nie tylko te stanowiące później tabele wymiarów) oraz hierarchie w 3.3 (pamiętaj, aby zidentyfikować unikalny identyfikator dla każdego z poziomów hierarchii). Następnie, w części 3.4, przygotuj schemat logiczny przedstawiający schemat gwiazdy/płatka śniegu/konstelacji faktów (diagram zawierający tabele oraz relacje między nimi) oraz – opcjonalnie – określ typ każdej tabeli wymiarów.

3.1 FAKTY I MIARY

	Fakt	Miara	Typ Miary	Agregacja	Ziarno
1	Strzelanina	Victims Killed	Liczba	Brak	Bardzo szczegółowe
2	Strzelanina	Victims Injured	Liczba	Brak	Bardzo szczegółowe
3	Strzelanina	Suspects Killed	Liczba	Brak	Bardzo szczegółowe
4	Strzelanina	Suspects Injured	Liczba	Brak	Bardzo szczegółowe
5	Strzelanina	Suspect Arrested	Liczba	Brak	Bardzo szczegółowe

3.2 WYMIARY I ATRYBUTY

	Wymiar	Atrybut	Krótki opis	Ziarno
1	Czas	Dekada, Rok, Miesiąc, Dzień	Wymiar, który będzie opisywał czas strzelaniny	Bardzo szczegółowe
2	Lokalizacja	State, City, Adres	Wymiar, który będzie opisywał lokalizację strzelaniny	Bardzo szczegółowe
3	Pogoda	Temperatura, Chmury, Opady, Ciśnienie, Wiatr	Wymiar, który będzie pogodę w trakcie strzelaniny	Bardzo szczegółowe
4	Dostęp do broni (prawo)	Lokalizacja, Rok, Dostęp do broni	Wymiar, który będzie zawierał kontekst o prawie do posiadania broni w danym roku	Ogólne statystyki

3.3 HIERARCHIE – ŚCIEŻKI NAWIGACYJNE

Każdy z wymiarów jest dostępny z tabeli faktów w odległości 1.

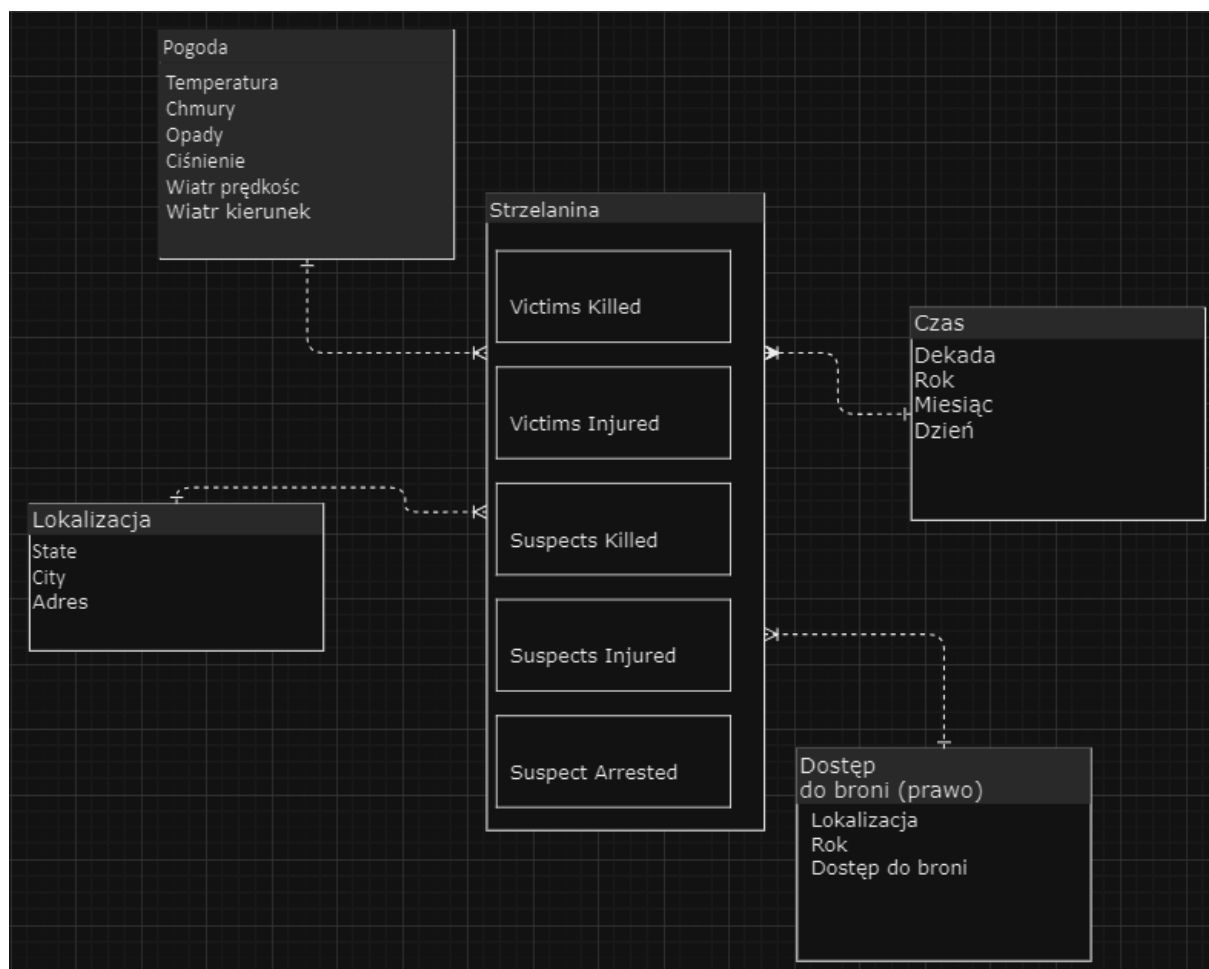
3.4 PROJEKT MODELU WIELOWYMIAROWEGO – SCHEMAT LOGICZNY

Naszym głównym faktem będzie strzelanina.

Od niej będziemy mogli dostać się do:

- czasu, w którym się wydarzyła,
- miejsce strzelaniny,
- pogoda jaka występowała w danym czasie w danym miejscu,
- informacje dotyczące prawa do posiadania broni w danym miejscu w danym czasie.

DIAGRAM



ZADANIE – PLAN ETL

Proces przygotowania danych dla systemu BI powinien być poprzedzony właściwie zdefiniowanym celem – tj. poprawnym modelem wielowymiarowym. Następnie należy odpowiednio zaplanować cały proces ekstrakcji, czyszczenia, uzgadniania i ładowania danych. W tym celu, w raporcie, przedstaw podstawowe mapowanie pomiędzy danymi docelowymi (model wielowymiarowy) a danymi źródłowymi w postaci mapy logicznej. Pamiętaj, aby skrótowo i hasłowo określić wymagane transformacje na danych. Dalej, w oparciu o przygotowaną mapę należy przygotować wysokopoziomowy plan tworzenia tabel wymiarów oraz tabeli faktów – dla każdej tabeli, osobno, przedstaw źródła oraz konieczne do realizacji transformacje, a następnie przedstaw plan weryfikacji kompletności i poprawności danych.

Podstawą do realizacji etapu jest uświadomienie sobie celu procesu, występujących problemów z danymi (tzw. anomalii), i opracowanie planu właściwego przygotowania modelu danych (czyszczenie, uzgadnianie, wzbogacanie danych).

4.1 MAPA LOGICZNA PROCESU ETL

	Cel					Źródło			Transformacje
	Tabela	Typ Tabeli	Atrybut	Typ Danych	SCD Retrospekcja	Źródło	Atrybut	Typ Danych	
1	Gun_permission	Wymiar	Year	Liczbowy	prawdziwa	State-Level Estimates of Household Firearm	YEAR	Liczbowy	Brak
2	Gun_permission	Wymiar	State	Tekstowy	prawdziwa		STATE	Tekstowy	Brak

						Ownership.xlsx			
3	Gun_permission	Wy miar	Permit	Logiczny	prawdziwa		permit	Logiczny	Brak
4	Strzelania	Fakt	Id	Liczbowy	permanenta	Gunviolence.slsx	Incident Id	Liczbowy	Brak
5	Strzelania Czas	Fakt Wy miar	Date Dekada Rok Miesiąc Dzień	Date	permanenta		Incident Date	Tekstowy	Uzgodnienie format, podział na MM DD YYYY
6	Lokalizacja	Wy miar	State	Tekstowy	permanenta		State	Tekstowy	Brak
7	Lokalizacja	Wy miar	City_or_Country	Tekstowy	permanenta		City or Country	Tekstowy	Brak
8	Lokalizacja	Wy miar	Address	Tekstowy	permanenta		Address	Tekstowy	Rozdział na poszczególne lokacje
9	Strzelania	Fakt	Victims_killed	Liczbowy	permanenta	Gunviolences.xlsx	Victims Killed	Liczbowy	Brak
10	Strzelania	Fakt	Victims_injured	Liczbowy	permanenta		Victims Injured	Liczbowy	Brak
11	Strzelania	Fakt	Suspects_killed	Liczbowy	permanenta		Suspects Killed	Liczbowy	Brak
12	Strzelania	Fakt	Suspects_injured	Liczbowy	permanenta		Suspects Injured	Liczbowy	Brak
13	Strzelania	Fakt	Suspects_arrested	Liczbowy	permanenta		Suspects Arrested	Liczbowy	Brak
14	Pogoda	Wy miar	Temperatura	Zmiennoprzecinkowa	Permanenta	OpenWeatherOneCallAPI	temperature	Zmiennoprzecinkowa	Konwersja na odp. jednostki

15	Pogoda	Wy miar	Zachmurze nie	Integer	Pernam etna		cloud_cov er	Integer	brak
16	Pogoda	Wy miar	Opady	Liczba rzeczywista	Pernam etna		precipitati on	Liczba rzeczywista	Brak
17	Pogoda	Wy miar	Ciśnienie	Liczba rzeczywista	Pernam etna		pressure	Liczba rzeczywista	Brak
18	Pogoda	Wy miar	Wiatr prędkość	Liczba rzeczywista	Pernam etna		Wind.spe ed	Liczba rzeczywista	Brak
19	Pogoda	Wy miar	Wiatr kierunek	Liczba naturalna (kąty)	Pernam etna		Wind.dire ction	Liczba naturalna (kąty)	Brak
20	Lokalizacj e	Wy miar	City	Tekstowa	Pernam etna		Lat, lon	Zmiennopr zecinkowa	Mapowa nie koordyn atów na nazwy miast
21	Lokalizacj e	Wy miar	State	Tekstowa	Pernam etna		Lat, lon	Zmiennopr zecinkowa	Mapowa nie koordyn atów na nazwy stanów

4.2 PLAN PROCESU – DIAGRAM ORAZ OPIS

4.2.1 Wymiary



4.2.2 FAKTY

Strzelanina.Operations
Strzelanina.Lokalizacja
Strzelanina.Incident Date

Usunięciu atrybutu Operations,
Przeniesienie danych lokalizacji
do wymiaru Lokalizacja,
Przeniesienie danych o dacie do
wymiaru Czas

Strzelanina

WNIOSKI OGÓLNE:

Skorzystaj z tej sekcji, aby przedstawić ogólne wnioski:

W ramach niniejszego raportu przedstawiliśmy kompleksowe podejście do projektowania rozwiązania Business Intelligence, uwzględniając zarówno analizę danych, jak i specyficzne potrzeby klienta. Nasza analiza objęła identyfikację źródeł danych, ocenę ich jakości oraz określenie metod integracji danych z różnych systemów.