

**SVEUČILIŠTE U ZAGREBU  
FAKULTET ORGANIZACIJE I INFORMATIKE  
VARAŽDIN**

**Sara Sušac**

# **ANALIZA POLICIJSKIH PUCNJAVA U SAD-u**

**PROJEKT IZ KOLEGIJA SKLADIŠTA PODATAKA I POSLOVNA  
INTELIGENCIJA**

**Varaždin, 2023.**

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET ORGANIZACIJE I INFORMATIKE**  
**V A R A Ž D I N**

**Sara Sušac**

**JMBAG: 0016143183**

**Studij: Baze podataka i baze znanja**

**ANALIZA POLICIJSKIH PUCNJAVA U SAD-u**

**PROJEKT IZ KOLEGIJA SKLADIŠTA PODATAKA I POSLOVNA INTELIGENCIJA**

**Mentor/Mentorica:**

Prof.dr.sc. Kornelije Rabuzin

**Varaždin, svibanj 2023.**

## **Sažetak**

Tema ovog seminarskog rada je skladišta podataka i njihova analiza koja se temelji na podacima iz samog skladišta podataka. Kako bi proveli analizu prvo je potrebno i izraditi samo skladište podataka. Za izradu skladišta biti će prikazan i ETL postupak kao i objašnjenje svakog koraka postupka. Nakon toga, skladište ćemo učitati u MySQL Workbench-u. Dodatnu analizu provesti ćemo kroz izvješća u Microsoft Power BI alat, gdje ćemo prikazati cijeli proces izgradnje i pripreme samih izvješća. Za analizu sam uzela podatke o policijskim pucnjavama u Sjedinjenim Američkim državama.

**Ključne riječi:** Skladište podataka, Excel, ETL, MySQL Workbench, Power BI

## Sadržaj

1. Uvod.....	1
2. Metode i tehnike rada.....	2
3. Skladište podataka .....	3
3.1. Izrada skladišta podataka.....	3
3.1.1. Izrada skladišta podataka.....	3
3.1.2. ETL postupak .....	5
3.1.2.1. Transformacija podataka u Excelu .....	5
3.1.2.2. Transformacija podataka u MySQL Workbench-u .....	6
4. Dimenzijsko modeliranje.....	10
4.1. Dimenzijska tablica <i>person_info</i> .....	11
4.2. Dimenzijska tablica <i>person_category</i> .....	13
4.3. Dimenzijska tablica <i>place_of_shooting</i> .....	15
4.4. Dimenzijska tablica <i>arms</i> .....	17
4.5. Dimenzijska tablica <i>police_report</i> .....	19
4.6. Dimenzijska tablica <i>date_and_time</i> .....	21
4.7. Činjenična tablica <i>evidention_of_shootings</i> .....	23
4.8. Model zvijezde.....	25
5. Power BI analiza .....	26
5.1. Power BI Desktop .....	26
6. Izrada izvještaja.....	29
6.1. Izvještaj o prosjeku godina osoba ubijenih u policijskoj pucnjava.....	29
6.2. Izvještaj o postotku muškaraca i žena ubijenih u policijskoj pucnjava .....	30
6.3. Izvještaj o gradovima odvijanja pucnjava.....	31
6.4. Izvještaj o postotku rasa koje su sudjelovale u pucnjavama .....	32
6.5. Izvještaj o vrstama oružja korištenim od strane upucane osobe .....	33
7. Zaključak .....	34
Popis literature .....	35
Popis slika.....	36

# 1. Uvod

U poslovnom svijetu, u današnje vrijeme sve više se koriste skladišta podataka (eng. data warehouse). Podaci se mijenjaju svakim danom, nadopunjuju se i njihova količina raste. Vrlo je teško izdvojiti samo bitne podatke među tolikom količinom, a analize podataka su vrlo važne za donošenje odluka. Zbog toga se izdvajaju samo bitni podaci i radi se analiza na temelju njih. Postoje razni načini analiza, a jedan od njih je izrada izvještaja. U poslovnom svijetu, odluke se donose na temelju analiza, dizajniraju se nove usluge i procjenjuje zadovoljstvo klijenata.

Podaci za ovaj rad pronađeni su na stranici Keggla. U radu će biti prikazan primjer analize policijskih pucnjava u Americi kroz izradu skladišta podataka i izradu izvještaja. Na početku rada će biti objašnjeno koje metode i tehnike, tj. procesi i tehnologije su korišteni za izradu samog projekta. Radi se o tehnologijama MySQL Workbench i Microsoft Power BI. Poglavlje nakon toga se bavi izradom skladišta podataka. Na početku poglavlja bit će opisani sami podaci, njihova svrha i struktura. Osim toga, u jednom od poglavlja opisan je i ETL (eng. Extract, Transform, Load) postupak kroz tehnologije. Na samom kraju su podaci učitani u skladište. U idućem poglavlju je opisano dimenzijsko modeliranje, tj. popunjavanje dimenzijskih i činjeničnih tablica. Posljednje poglavlje se odnosi na učitavanje podataka, pregled modela zvijezde i izradu izvještaja u alatu Power BI. Svi izvještaji će biti prikazani i opisani na samom kraju.

## **2. Metode i tehnike rada**

U ovom seminarskom radu će se koristiti različite metode i tehnike izrade. ETL postupak, koji će biti objašnjen, će se koristiti za čišćenje podataka. Za izradu projekta i pomoć pri dovršenju završnog dijela izvješća će se koristiti model zvijezde. Prvo je potrebno pronaći podatke za implementaciju svih metoda i tehnika, što se čini pretraživanjem web stranica.

Za početnu konverziju će se koristiti poznati program MS Excel. Nakon toga, većinu transformacija će obavljati MySQL Workbench alat. Korištena je besplatna verzija tog alata. MySQL Workbench je grafički alat za modeliranje podataka koji uključuje sve osnovne funkcije za takav proces. Nakon provedenih transformacija, podaci će biti spremni za analizu, pri čemu će se koristiti desktop verzija Microsoft Power BI alata. Microsoft Power BI je grafički alat koji je vrlo jednostavan i nudi mnoštvo opcija za prikaz i analizu podataka, stvaranjem grafikona i izvješća.

### 3. Skladište podataka

Današnje poslovno okruženje karakterizira snažna konkurencija u gotovo svim sektorima, što predstavlja izazov za suvremene organizacije. U svijetu u kojem se tehnologija neprestano razvija, modernim organizacijama je sve teže zadovoljiti potrebe korisnika kako bi stekle njihovo povjerenje. Cilj svake organizacije je privući što više zadovoljnih kupaca kako bi ostvarila konkurentске prednosti i profit. U ostvarenju tih ciljeva iznimno važnu ulogu imaju skladišta podataka i poslovna inteligencija.

Postoje mnoge definicije skladišta podataka u literaturi, no u ovom seminarskom radu ćemo se fokusirati na definiciju W.H. Inmona, koji se smatra osnivačem skladišta podataka. Prema Inmonu, skladište podataka je subjektivno orijentiran, integriran, postojan i vremenski različit skup podataka koji služi kao podrška upravljačkim odlukama. To znači da skladište podataka ima četiri glavna svojstva: subjektivnu orijentaciju, integraciju, postojanost i vremensku varijabilnost. Skladišta podataka imaju ključnu ulogu u poslovanju jer olakšavaju analizu različitih procesa kao što su uspješnost prodaje, nabava, projekti i drugi. Tri ključne razlike između baza podataka i skladišta podataka su sljedeće[1]:

1. Skladišta podataka se koriste za analizu podataka i omogućuju pristup većem broju zapisa nego što je moguće kod baza podataka.
2. Vrijeme odziva skladišta podataka je značajno veće nego kod baza podataka.
3. Skladišta podataka koriste manji broj korisnika.

#### 3.1. Izrada skladišta podataka

U ovome ćemo poglavlju prikazati odabrani skup podatak i način kako smo pripremili ove podatke za kreiranje skladišta. Nakon što smo pronašli samo skladište, pripremanje podataka i skladišta je najvažniji proces kako bi izvršili analizu te kreirali izvješća.

##### 3.1.1. Izrada skladišta podataka

Odabrani skup podataka koji ćemo koristiti su podaci o policijskim pucnjavama u Americi, točnije u SAD-u. Ovaj skup podataka sam pronašla na web stranici Keggles.com [2], a podaci su preuzeti u .csv dokumentu. Podaci sadrže 4896 zapisa, a svaki zapis sastoji se od 15 atributa:

- **id** – id upucane osobe
- **name** – ime i prezime upucane osobe
- **date** – datum kada je osoba upucana (D.M.Y.)

- **manner\_of\_death** – način na koji je upucana (shot - upucan, tasered - imobiliziran električnim šokom)
- **armed** - s čime je osoba bila naoružana ili ako je bila nenaoružana (gun – pištolj, unarmed- nenaoružana, toy weapon – dječja igračka, nail gun – pištolj za čavle, knife – nož, shovel – lopata itd. )
- **age** – koliko je osoba bila stara
- **gender** – spol (M – muško, F - žensko)
- **race** – rasa osobe (Asian – Azijac, White – Bijelac, Hispanic – Hispanoamerikanac, Black - crnac)
- **city** – grad u kojem je osoba upucana
- **state** – pokrajina u kojoj je osoba upucana
- **signs\_of\_mental\_illness** – ima li naznaka mentalnih bolesti kod upucane osobe (TRUE -istina, FALSE - neistina)
- **threat\_level** – jesu li napali (attack – napad, undetermined – neodlučeno, other – nešto drugo)
- **flee** – je li osoba pokušala pobjeći (Not fleeing – nije bježala, car – autom, foot - pješke)
- **body\_camera** – je li policajac nosio kameru kada se incident odvio (TRUE – istina, FALSE - laž)
- **arms\_category** – kategorija oružja koju je osoba imala (Guns – pištolji, Unarmed – nenaoružan, Other unusual objects – drugi neobični objekti, sharp objects – oštri objekti itd.)

Ovako izgleda tek preuzeta .csv datoteka:

id,name,date,manner_of_death,armed,age,gender,race,city,state,signs_of_mental_illness,threat_level,flee,body_camera,arms_category														
3, Tim Elliot, 2015-01-02, shot, gun, 53.0, M, Asian, Shelton, WA, True, attack, Not fleeing, False, Guns														
4, Lewis Lee Lembke, 2015-01-02, shot, gun, 47.0, M, White, Aloha, OR, False, attack, Not fleeing, False, Guns														
5, John Paul Quintero, 2015-01-03, shot and Tasered, unarmed, 23.0, M, Hispanic, Wichita, KS, False, other, Not fleeing, False, Unarmed														
8, Matthew Hoffman, 2015-01-04, shot, toy weapon, 32.0, M, White, San Francisco, CA, True, attack, Not fleeing, False, Other unusual objects														
9, Michael Rodriguez, 2015-01-04, shot, nail gun, 39.0, M, Hispanic, Evans, CO, False, attack, Not fleeing, False, Piercing objects														
11, Kenneth Joe Brown, 2015-01-04, shot, gun, 18.0, M, White, Guthrie, OK, False, attack, Not fleeing, False, Guns														
13, Kenneth Arnold Buck, 2015-01-05, shot, gun, 22.0, M, Hispanic, Chandler, AZ, False, attack, Car, False, Guns														
15, Brock Nichols, 2015-01-06, shot, gun, 35.0, M, White, Assaria, KS, False, attack, Not fleeing, False, Guns														
16, Autumn Steele, 2015-01-06, shot, unarmed, 34.0, F, White, Burlington, IA, False, other, Not fleeing, True, Unarmed														
17, Leslie Sapp III, 2015-01-06, shot, toy weapon, 47.0, M, Black, Knoxville, PA, False, attack, Not fleeing, False, Other unusual objects														
19, Patrick Wetter, 2015-01-06, shot and Tasered, knife, 25.0, M, White, Stockton, CA, False, attack, Not fleeing, False, Sharp objects														
21, Ron Sneed, 2015-01-07, shot, gun, 31.0, M, Black, Freeport, TX, False, attack, Not fleeing, False, Guns														
22, Hashim Hanif Ibn Abdul-Rasheed, 2015-01-07, shot, knife, 41.0, M, Black, Columbus, OH, True, other, Not fleeing, False, Sharp objects														
25, Nicholas Ryan Brickman, 2015-01-07, shot, gun, 30.0, M, White, Des Moines, IA, False, attack, Car, False, Guns														
27, Omarr Julian Maximilian Jackson, 2015-01-07, shot, gun, 37.0, M, Black, New Orleans, LA, False, attack, Foot, True, Guns														
29, Loren Simpson, 2015-01-08, shot, unknown, 28.0, M, White, Huntley, MT, False, undetermined, Not fleeing, False, Unknown														
32, James Dudley Barker, 2015-01-08, shot, shovel, 42.0, M, White, Salt Lake City, UT, False, attack, Not fleeing, True, Blunt instruments														
36, Artago Damon Howard, 2015-01-08, shot, unarmed, 36.0, M, Black, Strong, AR, False, attack, Not fleeing, False, Unarmed														
37, Thomas Hamby, 2015-01-08, shot, gun, 49.0, M, White, Syracuse, UT, False, attack, Not fleeing, True, Guns														
38, Jimmy Foreman, 2015-01-09, shot, gun, 71.0, M, White, England, AR, False, attack, Not fleeing, False, Guns														
325, Andy Martinez, 2015-01-09, shot, gun, 33.0, M, Hispanic, El Paso, TX, False, attack, Not fleeing, False, Guns														
42, Tommy Smith, 2015-01-11, shot, gun, 39.0, M, White, Arcola, IL, True, attack, Not fleeing, False, Guns														
43, Brian Barbosa, 2015-01-11, shot, gun, 23.0, M, Hispanic, South Gate, CA, False, attack, Not fleeing, False, Guns														
45, Salvador Figueroa, 2015-01-11, shot and Tasered, gun, 29.0, M, Hispanic, North Las Vegas, NV, False, attack, Foot, False, Guns														
46, John Edward O'Keefe, 2015-01-13, shot, gun, 34.0, M, White, Albuquerque, NM, False, attack, Foot, True, Guns														
48, Richard McClendon, 2015-01-13, shot, knife, 43.0, M, White, Jourdan, TX, True, other, Not fleeing, False, Sharp objects														
49, Marcus Golden, 2015-01-14, shot, unknown, 24.0, M, Black, St. Paul, MN, False, attack, Not fleeing, False, Unknown														
50, Michael Goebel, 2015-01-14, shot, unknown, 29.0, M, White, Franklin County, MO, False, attack, Not fleeing, False, Unknown														
51, Mario Jordan, 2015-01-14, shot, gun, 34.0, M, Black, Chesapeake, VA, True, attack, Not fleeing, False, Guns														
52, Talbot Schroeder, 2015-01-14, shot, knife, 75.0, M, White, Old Bridge, NJ, False, attack, Not fleeing, False, Sharp objects														
54, Robert Edwards, 2015-01-14, shot, gun, 68.0, M, White, Lake Jackson, TX, False, attack, Not fleeing, False, Guns														
55, Jeffrey R. Nielson, 2015-01-14, shot, knife, 34.0, M, White, Draper, UT, False, attack, Not fleeing, True, Sharp objects														
56, Donte Sowell, 2015-01-15, shot, gun, 27.0, M, Black, Indianapolis, IN, False, attack, Foot, False, Guns														
57, Kavonda Earl Payton, 2015-01-15, shot, gun, 39.0, M, Black, Aurora, CO, False, attack, Not fleeing, False, Guns														
60, Quincy Reed Reindl, 2015-01-15, shot, gun, 24.0, M, White, Bloomington, MN, True, attack, Not fleeing, False, Guns														
63, Jose Ceja, 2015-01-15, shot, knife, 36.0, M, Hispanic, Fairfield, CA, True, other, Not fleeing, False, Sharp objects														
64, Nathan Massey, 2015-01-15, shot, gun, 33.0, M, White, Evangeline Parish, LA, False, attack, Not fleeing, False, Guns														
67, Scott Hall, 2015-01-16, shot, gun, 41.0, M, White, Mabank, TX, False, attack, Car, False, Guns														

Slika 1. Preuzeta .csv datoteka



Na temelju ovakvog skupa podataka sa Slike 1. može se ustanoviti da je broj policijskih pucnjava u SAD-u iznimno izražen, stoga ćemo se unutar ovoga rada posvetiti informacijama vezanim uz same uzroke pucnjava.

Unutar ove .csv datoteke podaci su jako loše raspoređeni i teško se snaći u njima. Stupac name sadrži i ime i prezime, dok je prihvatljivije da se to ipak razdvoji na dva stupca, kako bi na kraju dobili što preciznije i kvalitetnije rješenje.

### 3.1.2.ETL postupak

ETL, poznat i kao Ekstrakcija, Transformacija, Učitavanje(Extract, Transfrom, Load), predstavlja ključnu komponentu skladištenja podataka odgovornu za njihovo ekstrahiranje, transformiranje i učitavanje podataka u samo skladište podataka. Proces ekstrakcije obuhvaća čitanje i prikupljanje podataka iz različitih izvora kao što su tekstualne datoteke, baze podataka i skladišta podataka. Nakon ekstrakcije, prikupljeni podaci prolaze kroz transformaciju kako bi se osigurala njihova usklađenost s željenim formatom skladišta podataka. Nakon transformacije, podaci se učitavaju u skladište podataka. U ovom slučaju, konverzija podataka će biti olakšana upotrebom alata kao što su Excel i MySQL Workbench [3].

#### 3.1.2.1. Transformacija podataka u Excelu

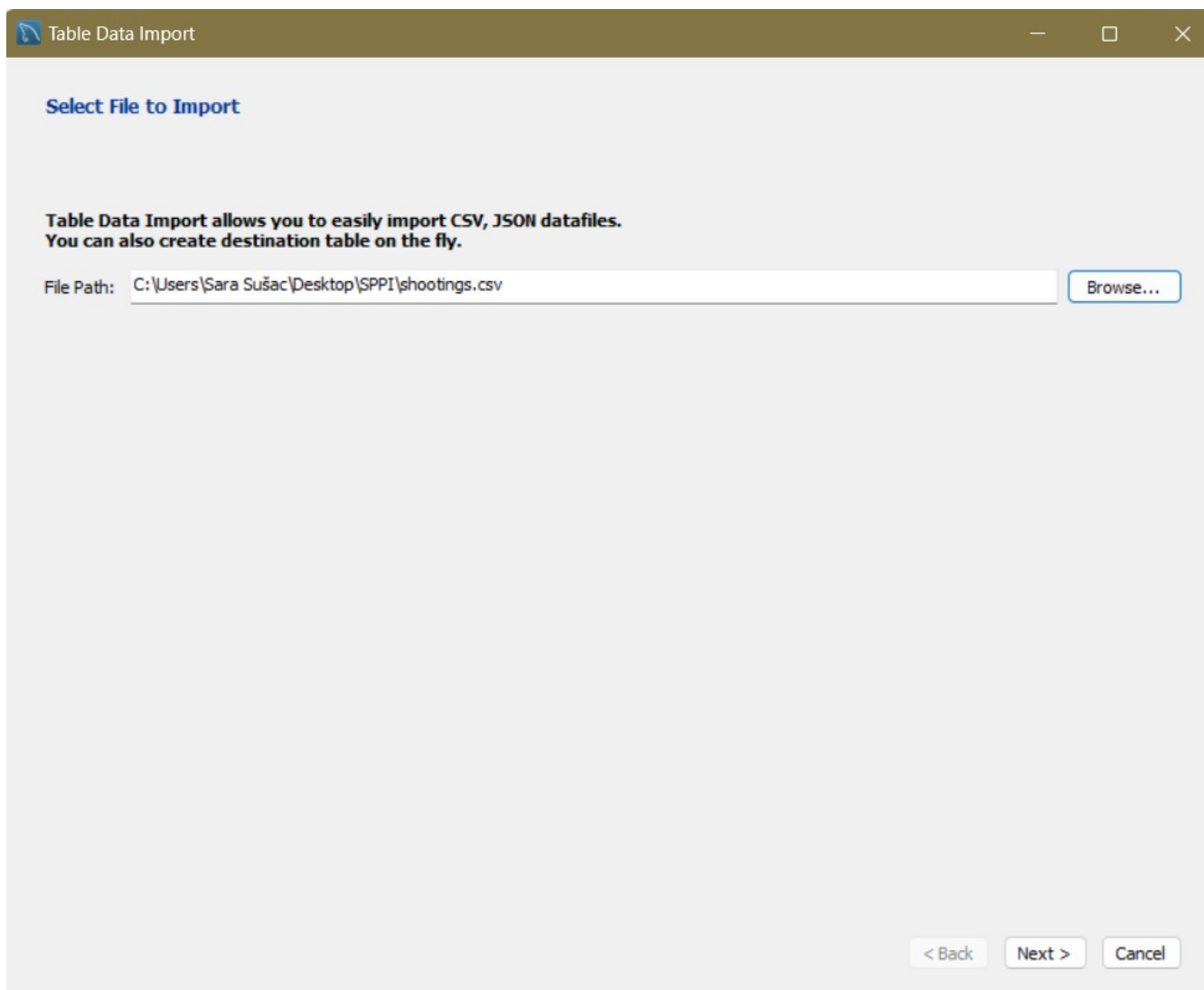
Prije učitavanja podataka u MySQL bazu podataka moramo napraviti par transformacija. U originalnoj .csv datoteci podaci su dosta nepregledni i odvojeni zarezom kao što vidimo na Slici 1. te se nalaze unutar jednog stupca u redovima. Za početak smo pretvorili tekst u stupce i maknuli smo stupac id, jer će se kasnije unutar MySQL-a stupac id ponovo dodati. Sada nam tablica u Excelu izgleda ovako:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	name	date	manner_of_death	armed	age	gender	race	city	state	signs_of_mental_illness	threat_level	flee	body_camera	arms_category
2	Tim Elliot	2.1.2015	shot	gun	53.0	M	Asian	Shelton	WA	TRUE	attack	Not fleeing	FALSE	Guns
3	Lewis Lee Lembke	2.1.2015	shot	gun	47.0	M	White	Aloha	OR	FALSE	attack	Not fleeing	FALSE	Guns
4	John Paul Quintero	3.1.2015	shot and Tasered	unarmed	23.0	M	Hispanic	Wichita	KS	FALSE	other	Not fleeing	FALSE	Unarmed
5	Matthew Hoffman	4.1.2015	shot	toy weapon	32.0	M	White	San Francisco	CA	TRUE	attack	Not fleeing	FALSE	Other unusual objects
6	Michael Rodriguez	4.1.2015	shot	nail gun	39.0	M	Hispanic	Evans	CO	FALSE	attack	Not fleeing	FALSE	Piercing objects
7	Kenneth Joe Brown	4.1.2015	shot	gun	18.0	M	White	Guthrie	OK	FALSE	attack	Not fleeing	FALSE	Guns
8	Kenneth Arnold Buck	5.1.2015	shot	gun	22.0	M	Hispanic	Chandler	AZ	FALSE	attack	Car	FALSE	Guns
9	Brock Nichols	6.1.2015	shot	gun	35.0	M	White	Assaria	KS	FALSE	attack	Not fleeing	FALSE	Guns
10	Autumn Steele	6.1.2015	shot	unarmed	34.0	F	White	Burlington	IA	FALSE	other	Not fleeing	TRUE	Unarmed
11	Leslie Sapp III	6.1.2015	shot	toy weapon	47.0	M	Black	Knoxville	PA	FALSE	attack	Not fleeing	FALSE	Other unusual objects
12	Patrick Wetter	6.1.2015	shot and Tasered	knife	25.0	M	White	Stockton	CA	FALSE	attack	Not fleeing	FALSE	Sharp objects
13	Ron Sneed	7.1.2015	shot	gun	31.0	M	Black	Freeport	TX	FALSE	attack	Not fleeing	FALSE	Guns
14	Hashim Hanif Ibn Abdul-Rasheed	7.1.2015	shot	knife	41.0	M	Black	Columbus	OH	TRUE	other	Not fleeing	FALSE	Sharp objects
15	Nicholas Ryan Brickman	7.1.2015	shot	gun	30.0	M	White	Des Moines	IA	FALSE	attack	Car	FALSE	Guns
16	Omarr Julian Maximillian Jackson	7.1.2015	shot	gun	37.0	M	Black	New Orleans	LA	FALSE	attack	Foot	TRUE	Guns
17	Loren Simpson	8.1.2015	shot	unknown	28.0	M	White	Huntley	MT	FALSE	undetermined	Not fleeing	FALSE	Unknown
18	James Dudley Barker	8.1.2015	shot	shovel	42.0	M	White	Salt Lake City	UT	FALSE	attack	Not fleeing	TRUE	Blunt instruments
19	Artago Damon Howard	8.1.2015	shot	unarmed	36.0	M	Black	Strong	AR	FALSE	attack	Not fleeing	FALSE	Unarmed
20	Thomas Hamby	8.1.2015	shot	gun	49.0	M	White	Syracuse	UT	FALSE	attack	Not fleeing	TRUE	Guns
21	Jimmy Foreman	9.1.2015	shot	gun	71.0	M	White	England	AR	FALSE	attack	Not fleeing	FALSE	Guns
22	Andy Martinez	9.1.2015	shot	gun	33.0	M	Hispanic	El Paso	TX	FALSE	attack	Not fleeing	FALSE	Guns
23	Tommy Smith	11.1.2015	shot	gun	39.0	M	White	Arcola	IL	TRUE	attack	Not fleeing	FALSE	Guns
24	Brian Barbosa	11.1.2015	shot	gun	23.0	M	Hispanic	South Gate	CA	FALSE	attack	Not fleeing	FALSE	Guns
25	Salvador Figueroa	11.1.2015	shot and Tasered	gun	29.0	M	Hispanic	North Las Vegas	NV	FALSE	attack	Foot	FALSE	Guns
26	John Edward O'Keefe	13.1.2015	shot	gun	34.0	M	White	Albuquerque	NM	FALSE	attack	Foot	TRUE	Guns
27	Richard McClendon	13.1.2015	shot	knife	43.0	M	White	Jourdanton	TX	TRUE	other	Not fleeing	FALSE	Sharp objects
28	Marcus Golden	14.1.2015	shot	unknown	24.0	M	Black	St. Paul	MN	FALSE	attack	Not fleeing	FALSE	Unknown
29	Michael Goebel	14.1.2015	shot	unknown	29.0	M	White	Franklin County	MO	FALSE	attack	Not fleeing	FALSE	Unknown
30	Mario Jordan	14.1.2015	shot	gun	34.0	M	Black	Chesapeake	VA	TRUE	attack	Not fleeing	FALSE	Guns
31	Talbot Schroeder	14.1.2015	shot	knife	75.0	M	White	Old Bridge	NJ	FALSE	attack	Not fleeing	FALSE	Sharp objects
32	Robert Edwards	14.1.2015	shot	gun	68.0	M	White	Lake Jackson	TX	FALSE	attack	Not fleeing	FALSE	Guns
33	Jeffrey R. Nielson	14.1.2015	shot	knife	34.0	M	White	Draper	UT	FALSE	attack	Not fleeing	TRUE	Sharp objects
34	Donte Sowell	15.1.2015	shot	gun	27.0	M	Black	Indianapolis	IN	FALSE	attack	Foot	FALSE	Guns
35	Kavonda Earl Payton	15.1.2015	shot	gun	39.0	M	Black	Aurora	CO	FALSE	attack	Not fleeing	FALSE	Guns
36	Quincy Daed Raindi	15.1.2015	shot	gun	74.0	M	White	Blainvinton	MM	TRUE	attack	Not fleeing	FALSE	Guns

Slika 2. Transformacija podataka u Excelu

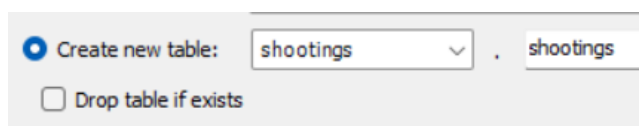
### 3.1.2.2. Transformacija podataka u MySQL Workbench-u

MySQL sustav upravljanja bazom koristi MySQL Workbench za kreiranje skladišta podataka. Za početak se treba spojiti na bazu lokano. Baza se zove sara, a shema unutar baze shootings. Nakon što smo kreirali shemu pozicioniramo se u nju i odaberemo opciju *Table Data Import*. U prozoru koji nam se otvori biramo shooting.csv datoteku kao što je prikazano na idućoj slici:



Slika 3. Table data import 1.dio

Nakon što smo odabrali opciju Next, odabiremo opciju Create new table kao što je prikazano na idućoj slici:



Slika 4. Table data import 2. dio

Slika 5. prikazuje vrste podataka koje smo odabrali za učitavanje u tablicu shootings:

name	date	manner_of...	armed	age	gender	race	city	state	signs_o
Tim Elliot	2.1.2015	shot	gun	53.0	M	Asian	Shelton	WA	TRUE
Lewis Lee...	2.1.2015	shot	gun	47.0	M	White	Aloha	OR	FALSE
John Paul...	3.1.2015	shot and T...	unarmed	23.0	M	Hispanic	Wichita	KS	FALSE
Matthew...	4.1.2015	shot	toy weapon	32.0	M	White	San Francis...	CA	TRUE

Slika 5. Table data import 3. dio

Nakon što smo to sve odradili podaci iz .csv datoteke nam se učitavaju u MySQL.:

	name	date	manner_of_death	armed	age	gender	race	city	state	signs_of_mental_illness	threat_level	flee	body_camera	arms_category
▶	Tim Elliot	2002-01-20 15:00:00	shot	gun	53	M	Asian	Shelton	WA	TRUE	attack	Not fleeing	FALSE	Guns
	Lewis Lee Lembke	2002-01-20 15:00:00	shot	gun	47	M	White	Aloha	OR	FALSE	attack	Not fleeing	FALSE	Guns
	John Paul Quintero	2003-01-20 15:00:00	shot and Tasered	unarmed	23	M	Hispanic	Wichita	KS	FALSE	other	Not fleeing	FALSE	Unarmed
	Matthew Hoffman	2004-01-20 15:00:00	shot	toy weapon	32	M	White	San Francisco	CA	TRUE	attack	Not fleeing	FALSE	Other unusual objec
	Michael Rodriguez	2004-01-20 15:00:00	shot	nail gun	39	M	Hispanic	Evans	CO	FALSE	attack	Not fleeing	FALSE	Piercing objects
	Kenneth Joe Brown	2004-01-20 15:00:00	shot	gun	18	M	White	Guthrie	OK	FALSE	attack	Not fleeing	FALSE	Guns
	Kenneth Arnold Buck	2005-01-20 15:00:00	shot	gun	22	M	Hispanic	Chandler	AZ	FALSE	attack	Car	FALSE	Guns
	Brock Nichols	2006-01-20 15:00:00	shot	gun	35	M	White	Assaria	KS	FALSE	attack	Not fleeing	FALSE	Guns
	Autumn Steele	2006-01-20 15:00:00	shot	unarmed	34	F	White	Burlington	IA	FALSE	other	Not fleeing	TRUE	Unarmed
	Leslie Sapp III	2006-01-20 15:00:00	shot	toy weapon	47	M	Black	Knoxville	PA	FALSE	attack	Not fleeing	FALSE	Other unusual objec
	Patrick Wetter	2006-01-20 15:00:00	shot and Tasered	knife	25	M	White	Stockton	CA	FALSE	attack	Not fleeing	FALSE	Sharp objects
	Ron Sneed	2007-01-20 15:00:00	shot	gun	31	M	White	Freeport	TX	FALSE	attack	Not fleeing	FALSE	Guns
	Hashim Hanif Ibn A...	2007-01-20 15:00:00	shot	knife	41	M	Black	Columbus	OH	TRUE	other	Not fleeing	FALSE	Sharp objects
	Nicholas Ryan Brick...	2007-01-20 15:00:00	shot	gun	30	M	White	Des Moines	IA	FALSE	attack	Car	FALSE	Guns
	Omarr Julian Maximi...	2007-01-20 15:00:00	shot	gun	37	M	Black	New Orleans	LA	FALSE	attack	Foot	TRUE	Guns
	Loren Simpson	2008-01-20 15:00:00	shot	unknown	28	M	White	Huntley	MT	FALSE	undetermined	Not fleeing	FALSE	Unknown
	James Dudley Barker	2008-01-20 15:00:00	shot	shovel	42	M	White	Salt Lake City	UT	FALSE	attack	Not fleeing	TRUE	Blunt instruments
	Artago Damon How...	2008-01-20 15:00:00	shot	unarmed	36	M	Black	Strong	AR	FALSE	attack	Not fleeing	FALSE	Unarmed
	Thomas Hamby	2008-01-20 15:00:00	shot	gun	49	M	White	Syracuse	UT	FALSE	attack	Not fleeing	TRUE	Guns
	Jimmy Foreman	2009-01-20 15:00:00	shot	gun	71	M	White	England	AR	FALSE	attack	Not fleeing	FALSE	Guns
	Andy Martinez	2009-01-20 15:00:00	shot	gun	33	M	Hispanic	El Paso	TX	FALSE	attack	Not fleeing	FALSE	Guns
	Tommy Smith	2011-01-20 15:00:00	shot	gun	39	M	White	Arcola	IL	TRUE	attack	Not fleeing	FALSE	Guns
	Brian Barbosa	2011-01-20 15:00:00	shot	gun	23	M	Hispanic	South Gate	CA	FALSE	attack	Not fleeing	FALSE	Guns

Slika 6. Učitani podaci u MySQL Workbench

Kao što smo već ranije spomenuli, stupac name bi bilo poželjno podijeliti na ime i prezime. Ono što još vidimo je da neke osobe imaju više prezimena tako da ćemo kreirati i drugi stupac tj. tablicu person\_help koja će u prvom stupcu sadržavati name(ime), a u drugom last\_name\_and\_family\_name (prezime\_i\_\_obiteljsko\_prezime). Upit nam stoga izgleda ovako:

```
1 • CREATE TABLE person_help
2   select substring_index(substring_index(name, ' ', 1), ' ', -1) as fist_name,
3   trim(substr(name, locate(' ', name )) ) as last_name_and_faily_name
4   from shootings.shootings;
5
```

*Slika 7. Kreiranje tablice person\_help*

Nakon što smo razdvojili stupac name u novu tablicu person\_help ona nam izgleda ovako

	fist_name	last_name_and_faily_name
►	Tim	Elliot
	Lewis	Lee Lembke
	John	Paul Quintero
	Matthew	Hoffman
	Michael	Rodriguez
	Kenneth	Joe Brown
	Kenneth	Arnold Buck
	Brock	Nichols
	Autumn	Steele
	Leslie	Sapp III
	Patrick	Wetter
	Ron	Sneed
	Hashim	Hanif Ibn Abdul-Rasheed
	Nicholas	Ryan Brickman
	Omarr	Julian Maximillian Jackson
	Loren	Simpson
	James	Dudley Barker
	Artago	Damon Howard
	Thomas	Hamby
	Jimmv	Foreman

*Slika 8. Tablica person\_help*

Istu transformaciju odradili smo i za stupac date. Kreirali smo tablicu date\_help, koja u sebi sadrži stupce date\_of\_shooting i time\_of\_shooting.

```
CREATE TABLE date_help
select substring_index(substring_index(date, ' ', 1), ' ', -1) as date_of_shooting,
trim(substr(date, locate(' ', date) )) as time_of_shooting
from shootings.shootings;
```

Slika 9. Kreiranje tablice date\_help

Tablica date\_help nam zatim izgleda ovako

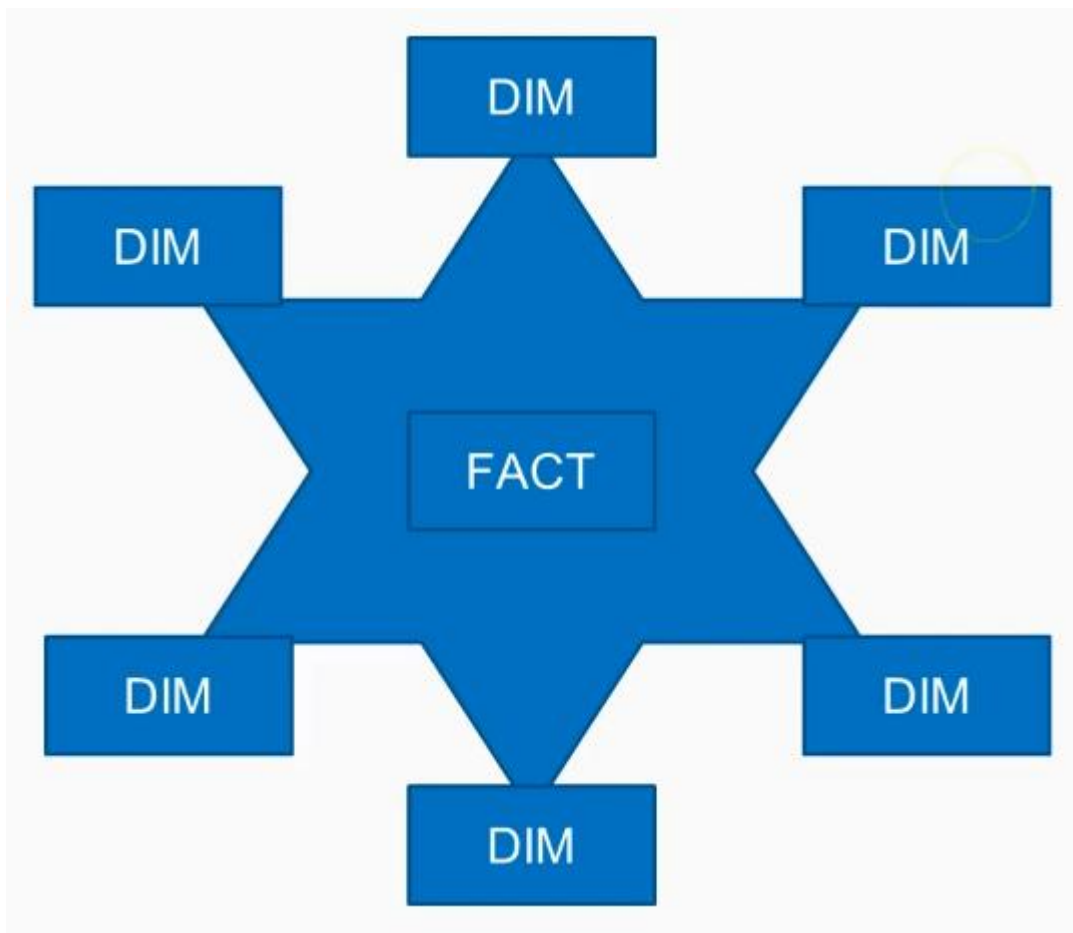
	date_of_shooting	time_of_shooting
►	2002-01-20	15:00:00
	2002-01-20	15:00:00
	2003-01-20	15:00:00
	2004-01-20	15:00:00
	2004-01-20	15:00:00
	2004-01-20	15:00:00
	2005-01-20	15:00:00
	2006-01-20	15:00:00
	2006-01-20	15:00:00
	2006-01-20	15:00:00
	2006-01-20	15:00:00
	2007-01-20	15:00:00
	2007-01-20	15:00:00
	2007-01-20	15:00:00
	2007-01-20	15:00:00
	2008-01-20	15:00:00
	2008-01-20	15:00:00
	2008-01-20	15:00:00

Slika 10. Tablica date\_help

Nakon što smo obavili potrebne transformacije, spremni smo za dimenzijsko modeliranje.

## 4. Dimenzijsko modeliranje

Nakon što smo završili transformacije pomoću MySQL Workbench-a možemo krenuti dalje na dimenzijsko modeliranje. U ovome dijelu koristiti ćemo se modelom zvijezde. Model zvijezde je popularan model dizajna baze podataka koji se koristi za organizaciju podataka u obliku zvijezde. Središnja tablica u modelu zvijezde je često velika činjenična tablica koja sadrži ključne podatke koje želimo analizirati. Povezane dimenzijske tablice, s druge strane, sadrže detaljne podatke o dimenzijama koje opisuju činjeničnu tablicu. Dimenzijske tablice su manje i jednostavnije u usporedbi s činjeničnom tablicom te se često koriste za filtriranje, grupiranje i sortiranje podataka [4].



*Slika 11. Model zvijezde*

Analizom podataka, odredili smo da nam je potrebno šest dimenzijskih tablica i jedna činjenična.



## 4.1. Dimenzijska tablica *person\_info*

Prva tablica koje je predodređena za dimenzijsku je tablica *person\_info*, čija je struktura prikazana na Slici 12.

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_person_info	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
first_name	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
last_name_and_family_name	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Slika 12. Kreiranje dimenzijske tablice *person\_info*

Tablica *person\_info* sastoji se od 3 atributa: *id\_person\_info*, *first\_name* i *last\_name\_and\_family\_name*. Primarni ključ je atribut *id\_person\_info* pa je zato i auto increment. Ostali atributi su varchar i ne smiju biti null.

Ranije smo kreirali tablicu *person\_help* koja nam je skužila da odvojimo ime od prezime tako da ćemo nju koristiti za dohvaćanje podataka kao što je prikazano na Slici 13.

```
1 • INSERT INTO person_info (first_name, last_name_and_family_name)
2   SELECT first_name, last_name_and_family_name FROM person_help;
```

Slika 13. Upit za učitavanje podataka u tablicu *person\_info*

Nakon provedenog upita može vidjeti da se ona ispunila točnim podacima:

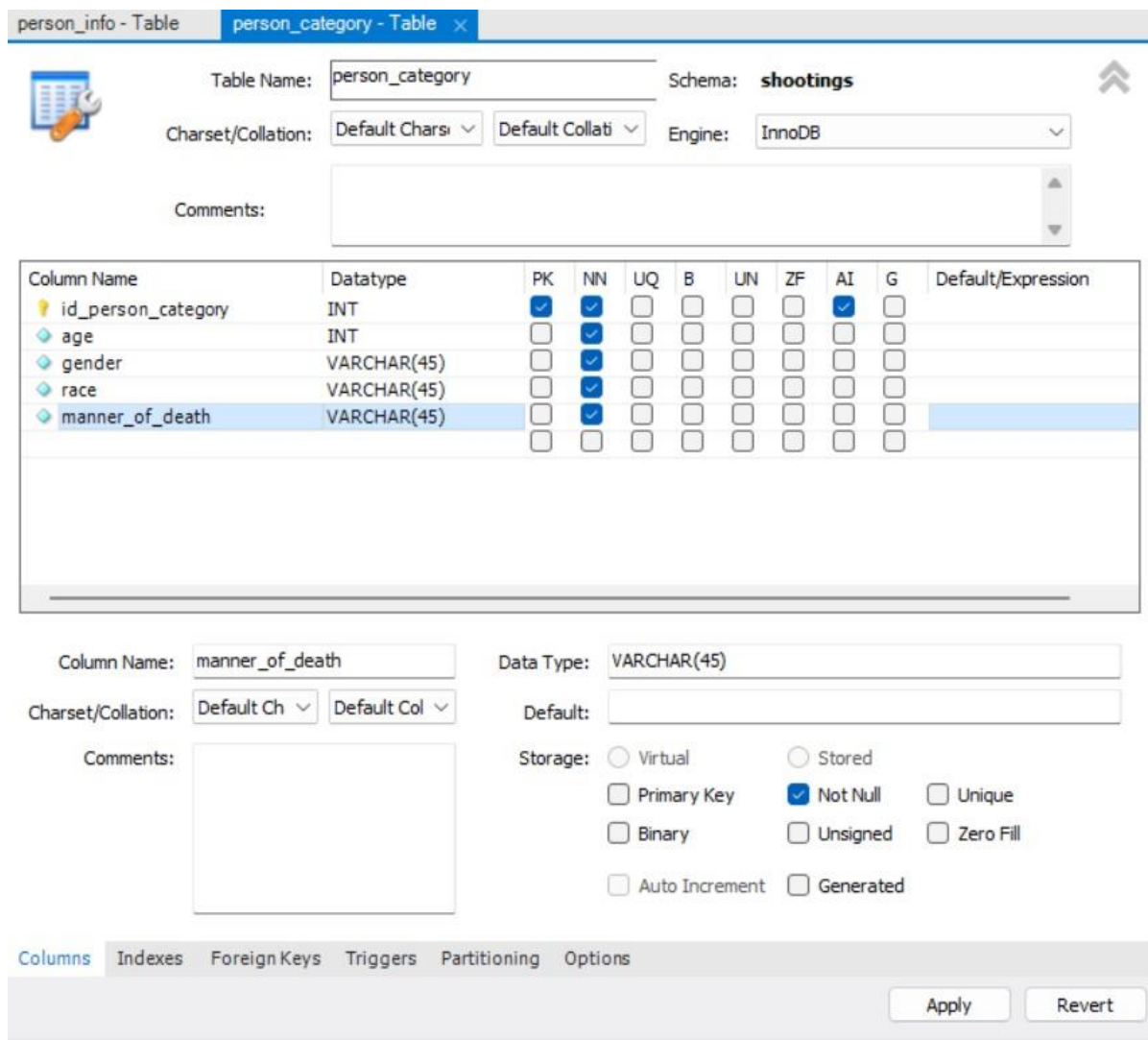
	id_person_info	first_name	last_name_and_family_name
▶	1	Tim	Elliot
	2	Lewis	Lee Lembke
	3	John	Paul Quintero
	4	Matthew	Hoffman
	5	Michael	Rodriguez
	6	Kenneth	Joe Brown
	7	Kenneth	Arnold Buck
	8	Brock	Nichols
	9	Autumn	Steele
	10	Leslie	Sapp III
	11	Patrick	Wetter
	12	Ron	Sneed
	13	Hashim	Hanif Ibn Abdul-Rasheed
	14	Nicholas	Ryan Brickman
	15	Omarr	Julian Maximillian Jackson
	16	Loren	Simpson
	17	James	Dudley Barker
	18	Artago	Damon Howard
	19	Thomas	Hamby
	20	Jimmy	Foreman

*Slika 14. Tablica person\_info.*



## 4.2. Dimenzijska tablica person\_category

Iduća dimenzijska tablica zove se person\_category. Ona u sebi sadržava atribute id\_person\_category, age, gender, race i manner\_of death. U ovom slučaju primarni ključ koji je ujedno i auto inkrement je atribut id\_person\_category. Atribut age je integer i not null dok su ostali atributi varchar i not null. Kreiranje tablice person\_category prikazano je na Slici 15.



The screenshot shows the MySQL Workbench interface for creating a new table named 'person\_category' in the 'shootings' schema. The table is configured with the following columns:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_person_category	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
age	INT	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
gender	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
race	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
manner_of_death	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

The 'manner\_of\_death' column is selected for configuration. The configuration panel shows the following settings:

- Column Name: manner\_of\_death
- Data Type: VARCHAR(45)
- Charset/Collation: Default Ch, Default Col
- Default: (empty)
- Storage: ☐ Virtual, ☐ Stored
- ☐ Primary Key, ☒ Not Null, ☐ Unique
- ☐ Binary, ☐ Unsigned, ☐ Zero Fill
- ☐ Auto Increment, ☐ Generated

The bottom of the interface shows tabs for Columns, Indexes, Foreign Keys, Triggers, Partitioning, and Options. The 'Columns' tab is active. The 'Apply' and 'Revert' buttons are at the bottom right.

Slika 15. Kreiranje dimenzijske tablice person\_category

Nakon što smo kreirali tablicu sa idućim upitom dohvaćamo podatke iz primarne tablice shootings.

```
INSERT INTO person_category (age, gender, race, manner_of_death)
SELECT age, gender, race, manner_of_death FROM shootings;
```

Slika 16. Upit za učitavanje podataka u tablicu person\_category

Upit je uspješno proveden te tablica izgleda ovako:

	id_person_category	age	gender	race	manner_of_death
▶	1	53	M	Asian	shot
	2	47	M	White	shot
	3	23	M	Hispanic	shot and Tasered
	4	32	M	White	shot
	5	39	M	Hispanic	shot
	6	18	M	White	shot
	7	22	M	Hispanic	shot
	8	35	M	White	shot
	9	34	F	White	shot
	10	47	M	Black	shot
	11	25	M	White	shot and Tasered
	12	31	M	Black	shot
	13	41	M	Black	shot
	14	30	M	White	shot
	15	37	M	Black	shot

*Slika 17. Tablica person\_category*

### 4.3. Dimenzijska tablica place\_of\_shooting

Slijedeća dimenzijska tablica je tablica place\_of\_shooting koja nam služi kako bi povezali grad i pokrajinu gdje se odvila pucnjava. Ona u sebi sadrži attribute id\_place\_of\_shooting, state i city. Primarni ključ ove tablice je atribut id\_place\_of\_shooting koji je ujedno not null i auto ikrement. Ostala dva atributa, state i city, su tipa varchar i oba su not null.

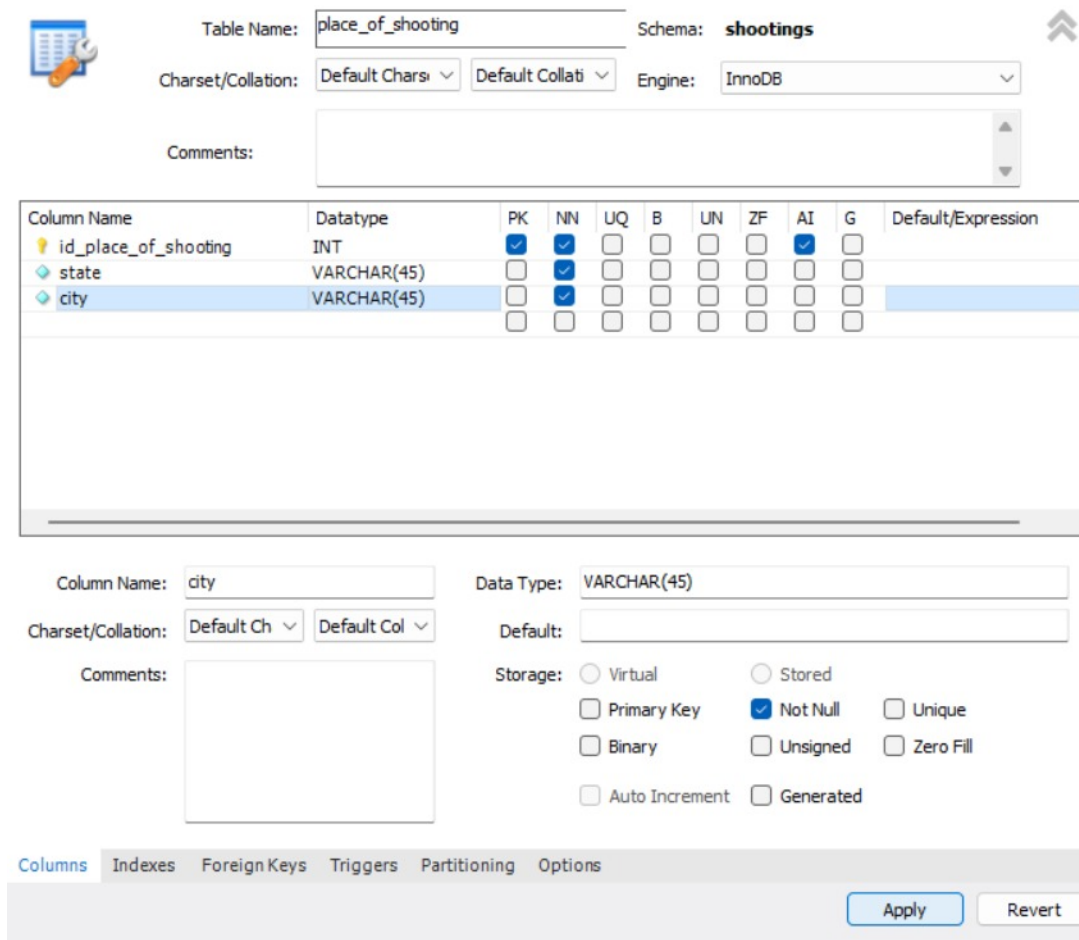


Table Name:  Schema: **shootings**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_place_of_shooting	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
state	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
city	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column Name:  Data Type:

Charset/Collation:

Default:

Comments:

Storage: ☐ Virtual ☐ Stored

☐ Primary Key ☒ Not Null ☐ Unique

☐ Binary ☐ Unsigned ☐ Zero Fill

☐ Auto Increment ☐ Generated

Columns Indexes Foreign Keys Triggers Partitioning Options

Slika 18. Kreiranje dimenzijske tablice place\_of\_shooting

Upit koji smo koristili za kreiranje ove tablice povlači podatke iz tablice shootings.

```
INSERT INTO place_of_shooting (state, city)
SELECT state, city FROM shootings;
```

Slika 19. Upit za učitavanje podataka u tablicu place\_of\_shooting

Nakon učitavanja podataka tablica nam izgleda kao na slici 20.

	id_place_of_shooting	state	city
▶	1	WA	Shelton
	2	OR	Aloha
	3	KS	Wichita
	4	CA	San Francisco
	5	CO	Evans
	6	OK	Guthrie
	7	AZ	Chandler
	8	KS	Assaria
	9	IA	Burlington
	10	PA	Knoxville
	11	CA	Stockton
	12	TX	Freeport
	13	OH	Columbus
	14	IA	Des Moines
	15	LA	New Orleans

*Slika 20. Tablica place\_of\_shootings*

## 4.4. Dimenzijska tablica arms

Tablica arms je iduća dimenzijska tablica koju ćemo kreirati kako bi povezali tip oružja koji je napadač koristio. Na slici 21. prikazano je kreiranje same tablice:

Table Name:  Schema: **shootings**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_arms	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
armed	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
arms_category	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column Name:  Data Type:

Charset/Collation:

Default:

Comments:

Storage: ☐ Virtual ☐ Stored

☐ Primary Key ☒ Not Null ☐ Unique

☐ Binary ☐ Unsigned ☐ Zero Fill

☐ Auto Increment ☐ Generated

Columns Indexes Foreign Keys Triggers Partitioning Options

Slika 21. Kreiranje dimenzijske tablice arms

Kao što se vidi iz priloženog. Tablica se sastoji od primarnog ključa id\_arms i dva atributa armed i arms\_category.

Podaci za ova atribute se nalaze u tablici shootings pa ih treba dohvatiti upitom.

```
INSERT INTO arms (armed, arms_category)
SELECT armed, arms_category FROM shootings;
```

Slika 22. Upit za učitavanje podataka u tablicu arms

Nakon što smo učitali podatke dobivamo idući izgled tablice gdje armed prikazuje vrstu oružja, a arms\_category opis kojoj vrsti oružja navedeno oružje točno spada.

	id_arms	armed	arms_category
►	1	gun	Guns
	2	gun	Guns
	3	unarmed	Unarmed
	4	toy weapon	Other unusual objects
	5	nail gun	Piercing objects
	6	gun	Guns
	7	gun	Guns
	8	gun	Guns
	9	unarmed	Unarmed
	10	toy weapon	Other unusual objects
	11	knife	Sharp objects
	12	gun	Guns
	13	knife	Sharp objects
	14	gun	Guns
	15	gun	Guns
	16	unknown	Unknown

*Slika 23. Tablica arms*

## 4.5. Dimenzijska tablica police\_report

Ova dimenzijska tablica prikazuje nam podatke koje je policijski službenik zabilježio nakon incidenta. Tablica police\_report u sebi sadrži atribute: id\_police\_report, signs\_of\_mental\_illness, threat\_level, flee i body\_camera kao što je prikazano na slici 24.

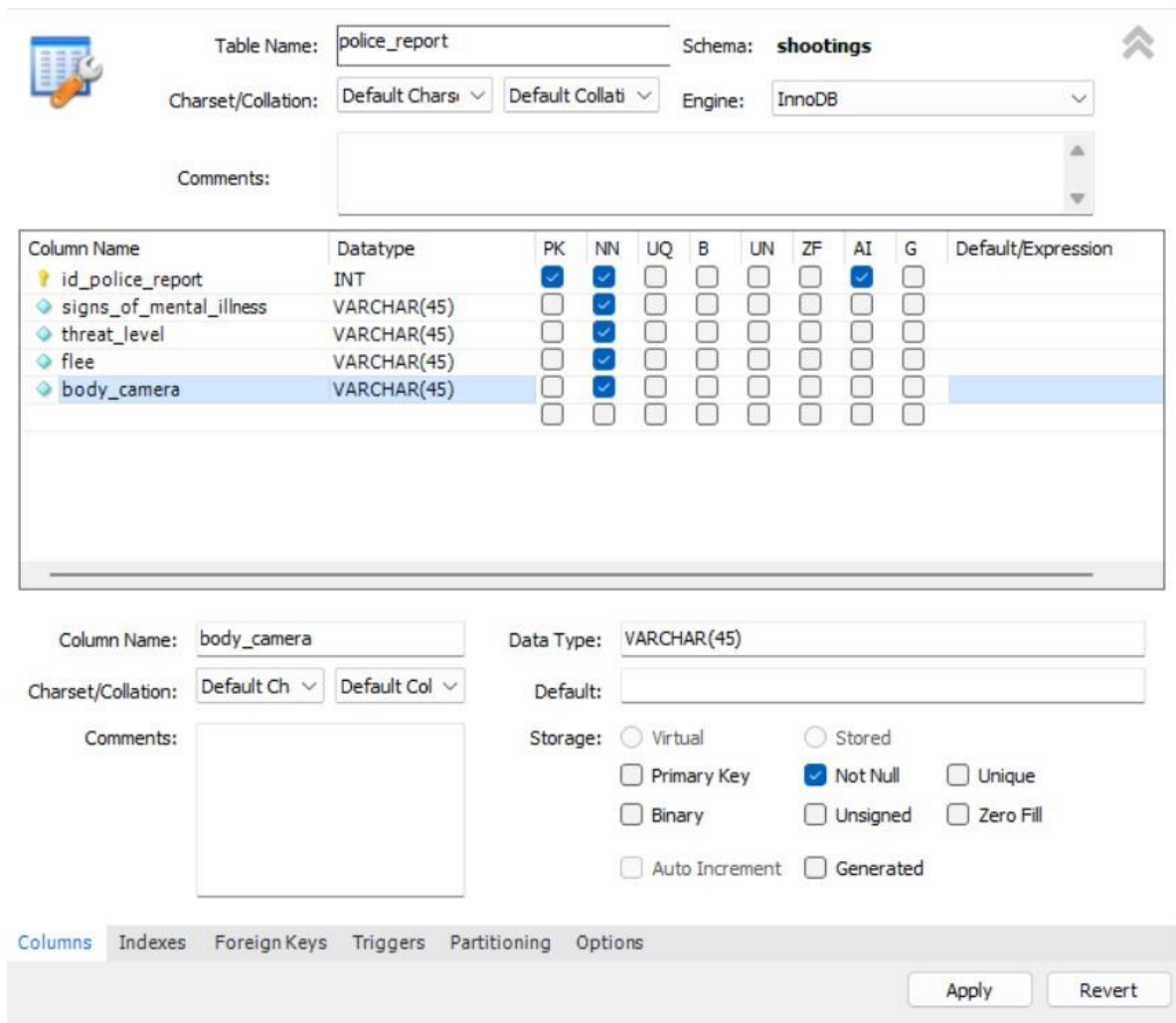







Table Name:  Schema: **shootings**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
 id_police_report	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
 signs_of_mental_illness	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
 threat_level	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
 flee	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
 body_camera	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column Name:  Data Type:

Charset/Collation:

Comments:

Storage: ☐ Virtual ☐ Stored

☐ Primary Key ☒ Not Null ☐ Unique

☐ Binary ☐ Unsigned ☐ Zero Fill

☐ Auto Increment ☐ Generated

Columns | Indexes | Foreign Keys | Triggers | Partitioning | Options

Slika 24. Kreiranje dimenzijske tablice police\_report

Nakon što je tablica kreirana uz pomoć upita koji povlači podatke iz tablice shootings ispunjavamo samu tablicu.

```
INSERT INTO police_report (signs_of_mental_illness, threat_level, flee, body_camera)
SELECT signs_of_mental_illness, threat_level, flee, body_camera FROM shootings;
```

Slika 25. Upit za učitavanje podataka u tablicu police\_report



Nakon određenog upita dobivamo tablicu kao na Slici 26.

	id_police_report	signs_of_mental_illness	threat_level	flee	body_camera
►	1	TRUE	attack	Not fleeing	FALSE
	2	FALSE	attack	Not fleeing	FALSE
	3	FALSE	other	Not fleeing	FALSE
	4	TRUE	attack	Not fleeing	FALSE
	5	FALSE	attack	Not fleeing	FALSE
	6	FALSE	attack	Not fleeing	FALSE
	7	FALSE	attack	Car	FALSE
	8	FALSE	attack	Not fleeing	FALSE
	9	FALSE	other	Not fleeing	TRUE
	10	FALSE	attack	Not fleeing	FALSE
	11	FALSE	attack	Not fleeing	FALSE
	12	FALSE	attack	Not fleeing	FALSE
	13	TRUE	other	Not fleeing	FALSE
	14	FALSE	attack	Car	FALSE
	15	FALSE	attack	Foot	TRUE

*Slika 26. Tablica police\_report*



## 4.6. Dimenzijska tablica date\_and\_time

Zadnja dimenzijska tablica koja je kreirana je tablica date\_and\_time koja ima attribute id\_date\_and\_time(primarni ključ), te date\_of\_shooting i time\_of\_shooting.

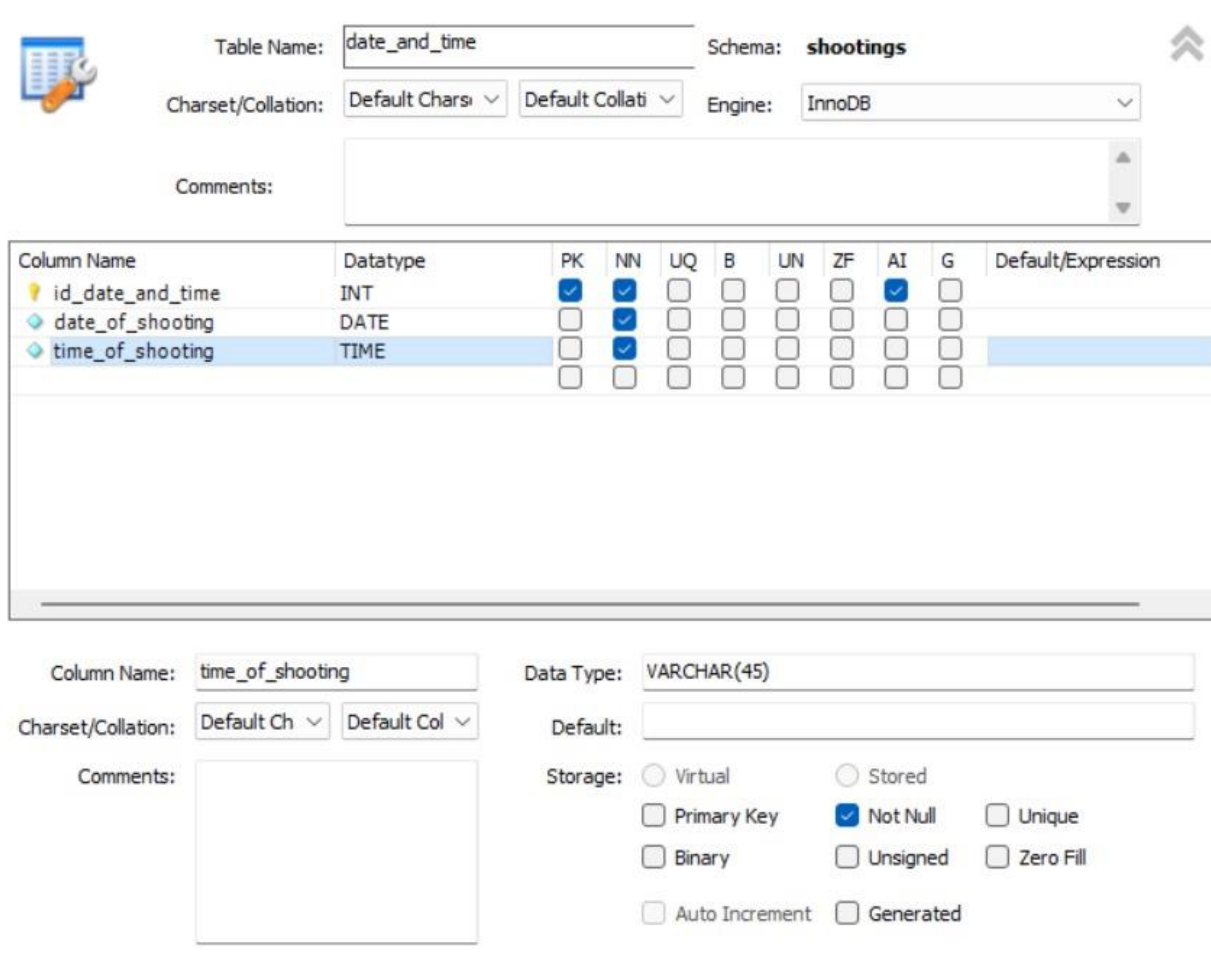


Table Name:  Schema: **shootings**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_date_and_time	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
date_of_shooting	DATE	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
time_of_shooting	TIME	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column Name:  Data Type:

Charset/Collation:

Comments:

Storage: ☐ Virtual ☐ Stored

☐ Primary Key ☒ Not Null ☐ Unique

☐ Binary ☐ Unsigned ☐ Zero Fill

☐ Auto Increment ☐ Generated

Slika 27. Kreiranje dimenzijske tablice date\_and\_time

Kako smo već ranije napravili tablicu date\_help, koja nam je služila za odvajanje datuma i vremena događaja, nju ćemo ona i koristiti prilikom provedbe upita za unos podataka.

```
INSERT INTO date_and_time (date_of_shooting, time_of_shooting)
SELECT date_of_shooting, time_of_shooting FROM date_help;
```

Slika 28. Upit za učitavanje podataka u tablicu date\_and\_time

Nakon uspješno provedenog upita, tablica date\_and\_time popunjena je podacima.

	id_date_and_time	date_of_shooting	time_of_shooting
►	1	2002-01-20	15:00:00
	2	2002-01-20	15:00:00
	3	2003-01-20	15:00:00
	4	2004-01-20	15:00:00
	5	2004-01-20	15:00:00
	6	2004-01-20	15:00:00
	7	2005-01-20	15:00:00
	8	2006-01-20	15:00:00
	9	2006-01-20	15:00:00
	10	2006-01-20	15:00:00
	11	2006-01-20	15:00:00

*Slika 29. Tablica date\_and\_time*

## 4.7. Činjenična tablica evidention\_of\_shootings

Nakon što smo definirali šest dimenzijskih tablica, možemo preći na izradu činjenične tablice. Činjenična tablica, u našem slučaju, zove se evidention\_of\_shootings te sadrži šest vanjskih ključeva s vezama jedan naprema više. Strukturu tablice evidention\_of\_shootings prikazana je na Slici 30.

Table Name: evidention\_of\_shootings Schema: shootings

Charset/Collation: utf8mb4 utf8mb4\_0900 Engine: InnoDB

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
id_person_info	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
id_person_category	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
id_place_of_shooting	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
id_arms	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
id_police_report	INT	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
id_date_and_time	INT	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

Column Name: id\_date\_and\_time Data Type: INT

Charset/Collation: Default Ch Default Col

Default:

Comments:

Storage: ☐ Virtual ☐ Stored

☒ Primary Key ☒ Not Null ☐ Unique

☐ Binary ☐ Unsigned ☐ Zero Fill

☐ Auto Increment ☐ Generated

Columns Indexes Foreign Keys Triggers Partitioning Options

Apply Revert

Slika 30. Kreiranje činjenične tablice evidention\_of\_shootings

Činjenična tablica se sastoji od 6 atributa: id\_person\_info, id\_person\_category, id\_place\_of\_shooting, id\_arms, id\_police\_report i id\_date\_and\_time. Kao što je već ranije spomenuto radi se o vanjskim ključevima na tablice person\_info, person\_category, place\_of\_shooting, arms, police\_report i date\_and\_time.

Na Slici 31. može se vidjeti kako smo povezali vanjske ključeve:

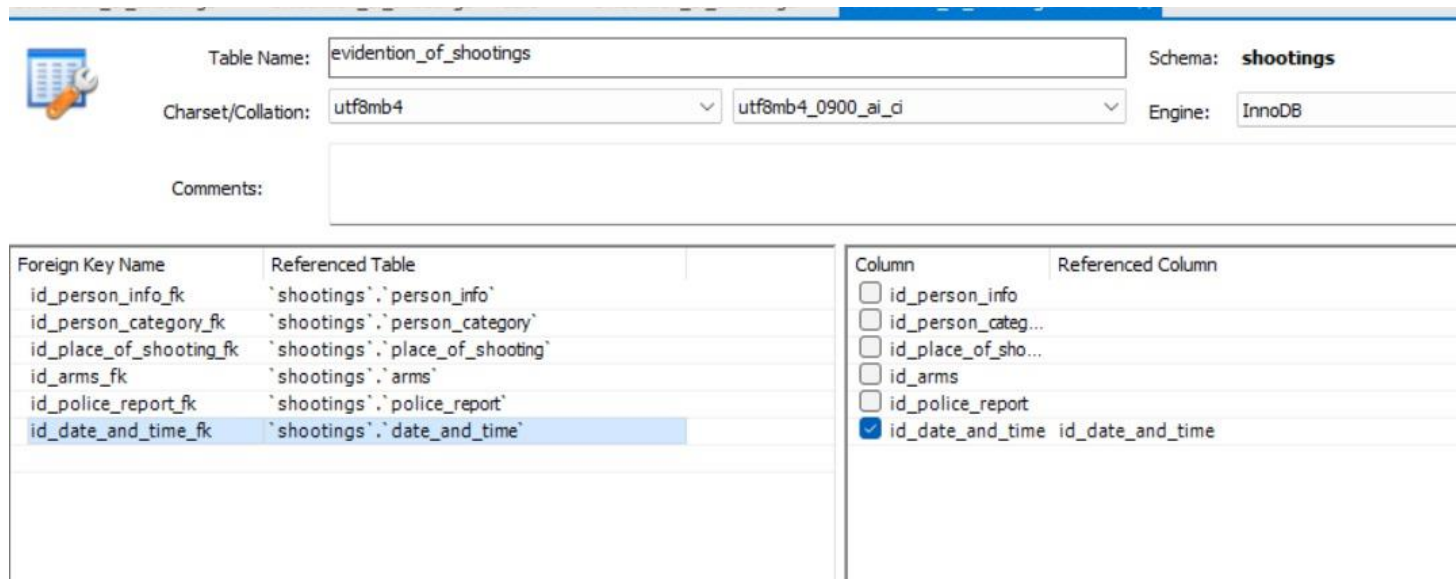


Table Name:  Schema: **shootings**

Charset/Collation:   Engine:

Comments:

Foreign Key Name	Referenced Table	Column	Referenced Column
id_person_info_fk	'shootings`.`person_info`	<input type="checkbox"/> id_person_info	
id_person_category_fk	'shootings`.`person_category`	<input type="checkbox"/> id_person_category	
id_place_of_shooting_fk	'shootings`.`place_of_shooting`	<input type="checkbox"/> id_place_of_shooting	
id_arms_fk	'shootings`.`arms`	<input type="checkbox"/> id_arms	
id_police_report_fk	'shootings`.`police_report`	<input type="checkbox"/> id_police_report	
id_date_and_time_fk	'shootings`.`date_and_time`	<input checked="" type="checkbox"/> id_date_and_time	id_date_and_time

Slika 31. Povezivanje ključeva

Zatim kako bi ispunili činjeničnu tablicu podacima, morali smo izvršiti i slijedeći upit kako bi dohvatili id-eve iz dimenzijskih tablica i uspješno ispunili činjeničnu tablicu podacima.

```
INSERT INTO evidention_of_shootings (id_person_info, id_person_category, id_place_of_shooting, id_arms, id_police_report, id_date_and_time)
SELECT id_person_info, id_person_category, id_place_of_shooting, id_arms, id_police_report, id_date_and_time
FROM person_info, person_category, place_of_shooting, arms, police_report, date_and_time
WHERE id_person_info = id_person_category AND id_person_category = id_place_of_shooting
AND id_place_of_shooting = id_arms AND id_arms = id_police_report AND id_police_report = id_date_and_time;
```

Slika 32. Upit za učitavanje podataka u tablicu evidention\_of\_shootings povezivanjem id-a

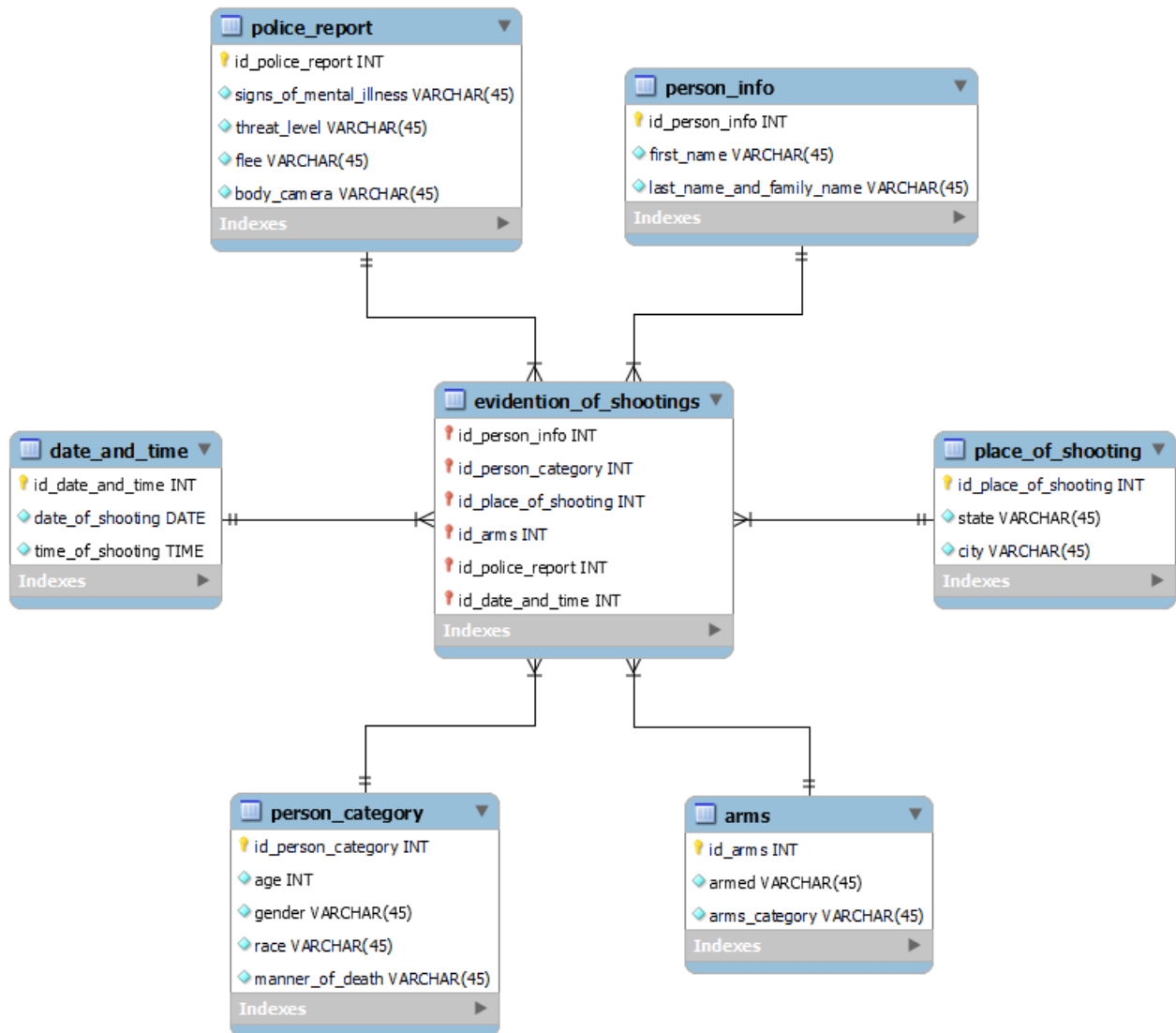
Sada vidimo i izgled popunjene tablice na Slici 33.

	id_person_info	id_person_category	id_place_of_shooting	id_arms	id_police_report	id_date_and_time
▶	1	1	1	1	1	1
	2	2	2	2	2	2
	3	3	3	3	3	3
	4	4	4	4	4	4
	5	5	5	5	5	5
	6	6	6	6	6	6
	7	7	7	7	7	7
	8	8	8	8	8	8
	9	9	9	9	9	9
	10	10	10	10	10	10
	11	11	11	11	11	11
	12	12	12	12	12	12
	13	13	13	13	13	13
	14	14	14	14	14	14
	15	15	15	15	15	15
	16	16	16	16	16	16
	17	17	17	17	17	17
	18	18	18	18	18	18

Slika 33. Tablica evidention\_of\_shootings

## 4.8. Model zvijezde

Ranije smo već objasnili model zvijezde. Trenutno skladište sastoji se od 6 dimenzijskih tablica povezanih sa činjeničnom prema modelu zvijezde. Svih šest tablica povezano je sa činjeničnom vezom jedan na prema više kao što se može vidjeti i iz slike 34.



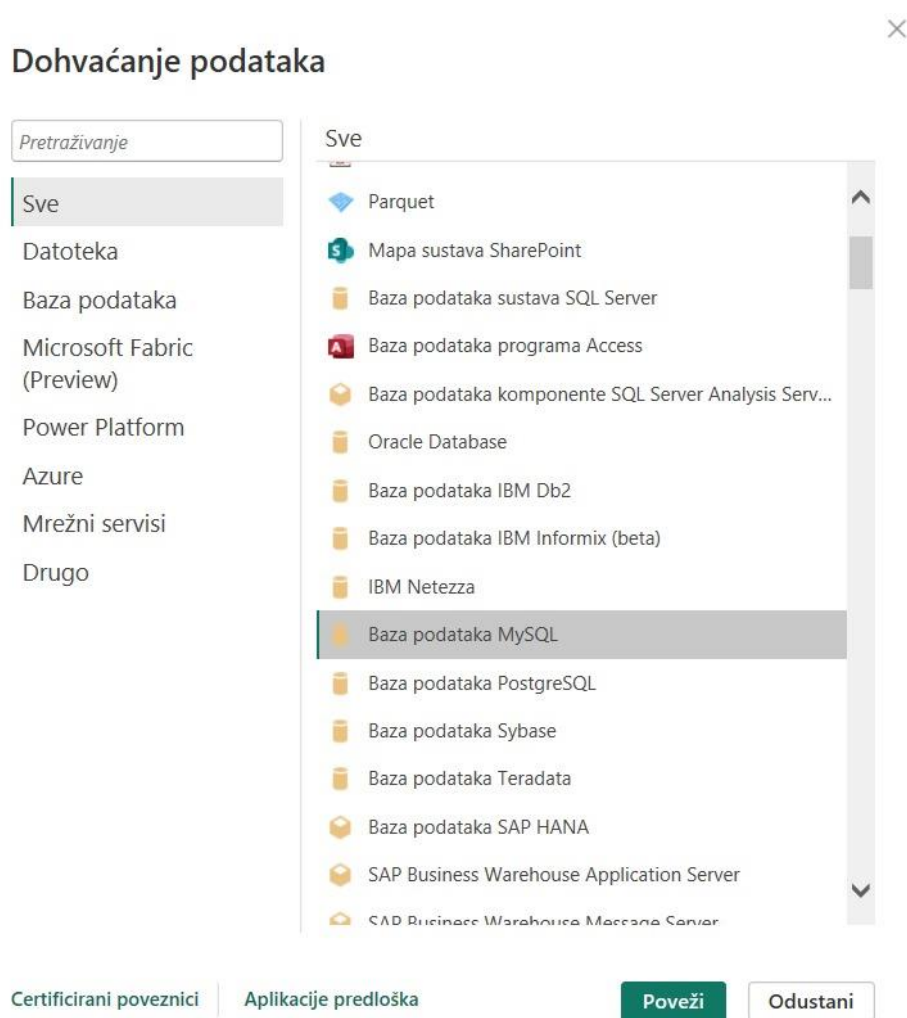
Slika 34. Model zvijezde izrađenog skladišta podataka

## 5. Power BI analiza

Preostala nam je još izrada izvještaja koja se radi na temelju podataka dobivenih ih skladišta podataka u alatu Power BI Desktop. Biti će napravljeno ukupno 5 izvještaja te će biti prikazan proces njihova kreiranja.

### 5.1. Power BI Desktop

Za početak unutar alata Power BI Desktop dohvaćamo podatke našeg skladišta. Odabiremo opciju Dohvati podatke nakon čega odabiremo opciju Baza podataka MySQL kao što je prikazano na Slici 35.



Slika 35. Dohvaćanje podataka u Power BI Desktopu

Nakon što pritisnemo Poveži potrebno je spojiti se na našu bazu:

### Baza podataka MySQL

Poslužitelj

127.0.0.1:3306

Baza podataka

shootings

> Napredne mogućnosti

U redu

Odustani

Slika 36. Spajanje na bazu unutar Power BI Desktopa

Nakon povezivanja na bazu potrebno je odabrati željene tablice. Tablice koje smo u ovom slučaju odabrali mogu se vidjeti na slici 37.

### Navigator

shootings.evidention\_of\_shootings

id_person_info	id_person_category	id_place_of_shooting	id_arms	id_poli
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23

Odabir povezanih tablica

Učitaj

Transformacija podataka

Odustani

Slika 37. Odabir tablica za korištenje u Power BI Desktopu



Pritiskom na tipku Učitaj, dobivamo prikaz modela zvijezde:



Slika 38. Model zvijezde unutar alata Power BI Desktop



## 6. Izrada izvještaja

Sada ćemo prikazati izvještaje. Svi izvještaji su vezani uz prethodno kreiranu bazu i temelje se na podacima vezanim uz tablicu shootings.

### 6.1. Izvještaj o prosjeku godina osoba ubijenih u policijskoj pucnjava

U prvom izvještaju prikazan je prosječan broj godina koje je imala osoba koja je ubijena od strane policajca.

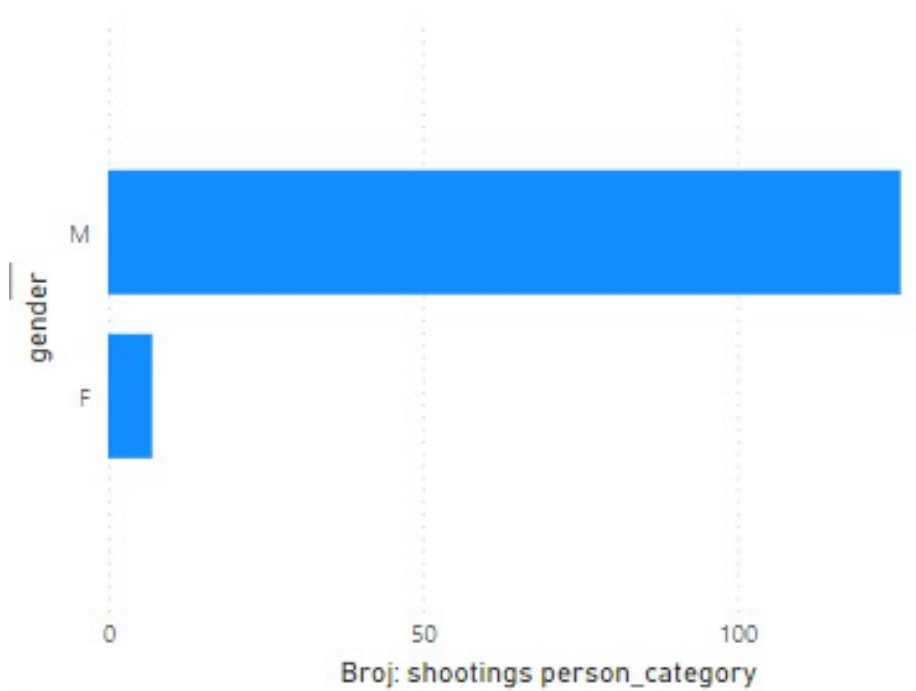


*Slika 39. Prosjek godina ubijene osobe*

Iz dobivenog izvještaja možemo vidjeti da se najviše radi o osobama srednjih godina, točnije osobama koje imaju između 37 i 38 godina.

## 6.2. Izvještaj o postotku muškaraca i žena ubijenih u policijskoj pucnjava

U ovom izvještaju promatramo ukupni udio muškaraca i žena unutar policijskih pucnjava.

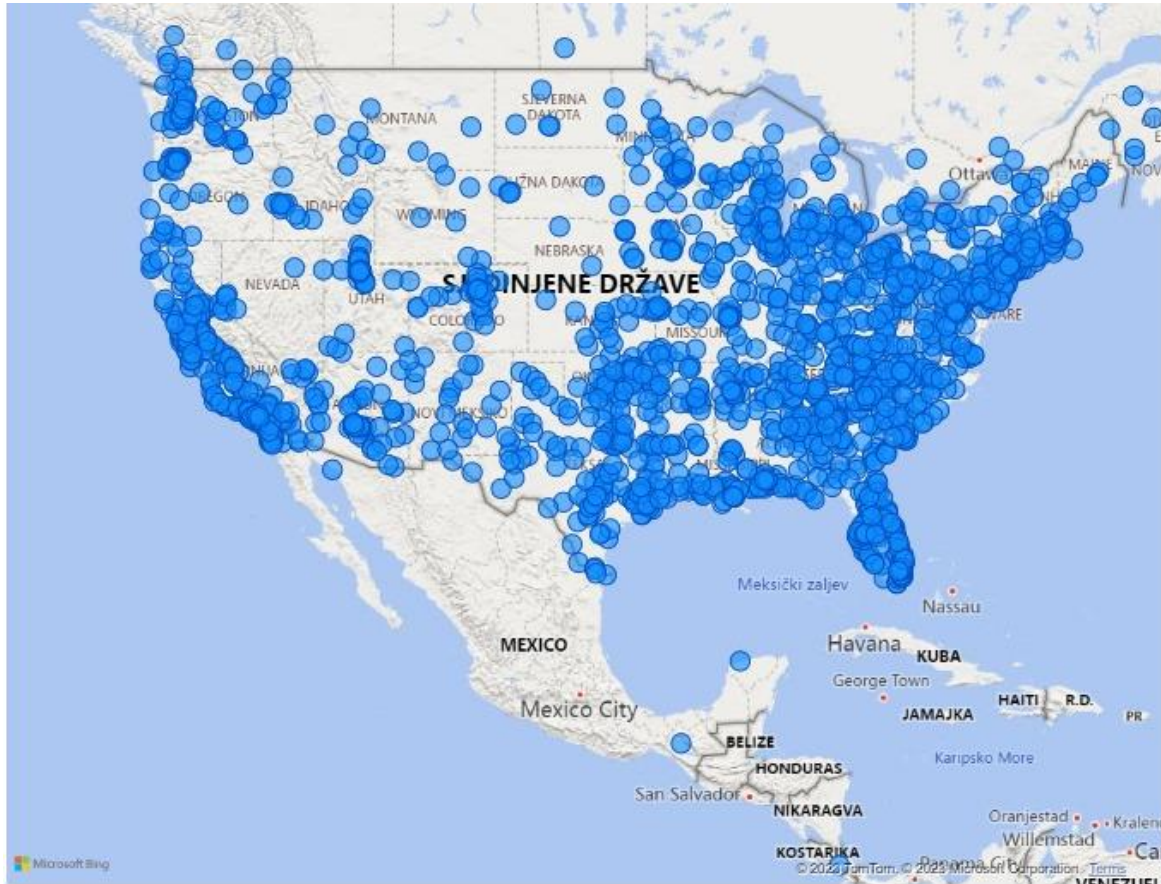


Slika 40. Udio muškaraca i žena koji su ubijeni u policijskoj pucnjava

Iz dobivenog izvještaja možemo vidjeti da se najviše radi o muškim osobama koje su bile ubijene, dok su žene u manjini.

### 6.3. Izvještaj o gradovima odvijanja pucnjava

U idućem izvještaju pokušati ćemo ustanoviti o kojim se gradovima u SAD-u dešavalo najviše pucnjava.

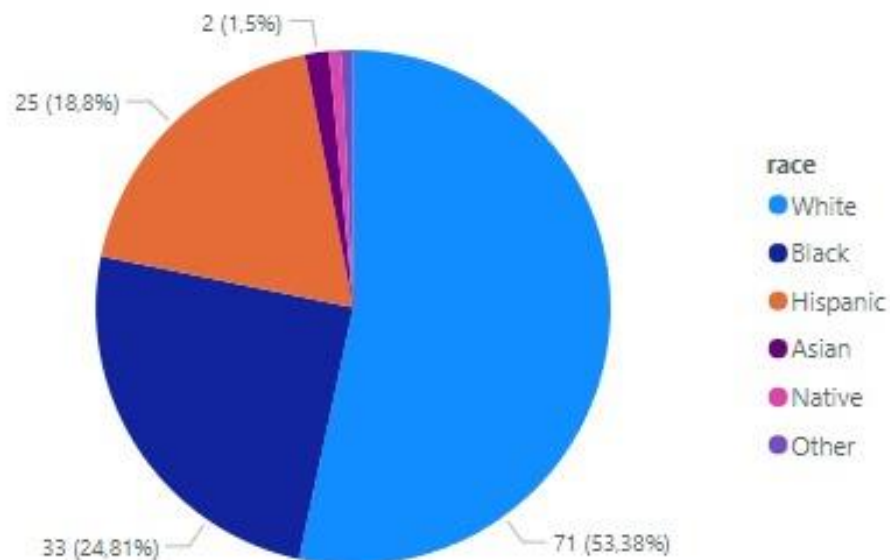


Slika 41. Gradovi u kojima se odvijalo najviše pucnjava

Iz dobivene mape možemo vidjeti da se najviše pucnjava odvijalo u Istočnim Sjedinjenim Američkim Državama, točnije na istočnoj obali. Ovaj podatak pomaže mogućoj reorganizaciji policijskih službenika kako bi znali gdje uvesti veće mjere kazne kako bi se smanjio broj ovakvih slučajeva.

## 6.4. Izvještaj o postotku rasa koje su sudjelovale u pucnjavama

U ovome izvještaju istražiti ćemo zanimljivi postotak kako bi ustanovi koje su rase osobe koje su sudjelovale u pucnjavama i u kojem omjeru. Kao pretpostavku možemo uzeti da se radi o crncima jer takve vijesti se svakodnevno najviše prikazuju na vijestima.

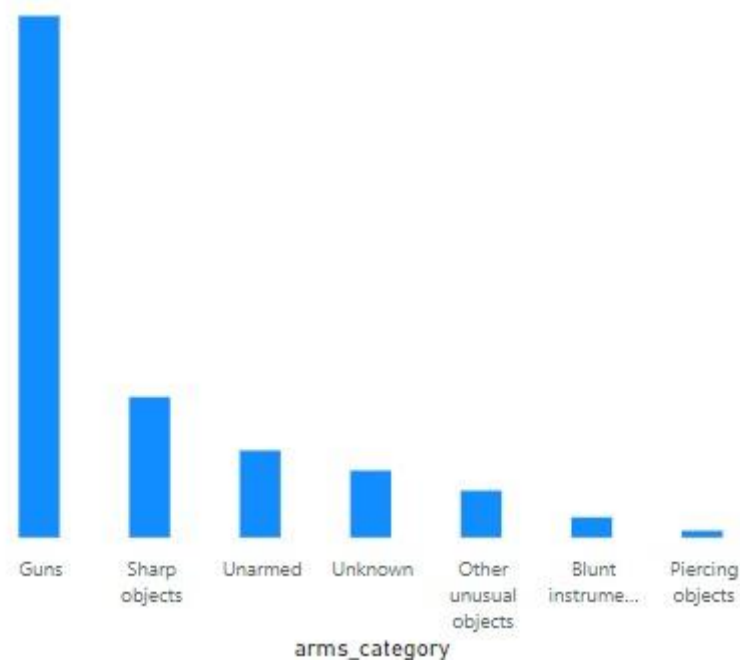


*Slika 42. Udio rasa koje su sudjelovale u policijskim pucnjavama*

Iz dobivenog izvještaja sa Slike 42. vidimo se ipak radi radi o bijelcima kao osobama koje su najviše sudjelovale u pucnjavama sa ukupno 53,38%. Na drugom mjestu su crnci sa 24,81% te zatim hispanci sa 18,8%.

## 6.5. Izvještaj o vrstama oružja korištenim od strane upucane osobe

U zadnjem izvještaju napraviti ćemo dijagram koji nam pokazuje koja su oružja i u kojem udjelu koristile osobe koje su ubijene od strane policijskog službenika. Dobiveni rezultati mogu se vidjeti na Slici 43.



Slika 43. Udio korištenog oružja u pucnjavama

Iz dobivenog grafa vidimo da se najviše radi o pištoljima što i nije čudan odgovor jer je u njihovoj državi dozvoljena legalna kupnja oružja. Većina kućanstava ustvari posjeduje ako ne i više vatrenih oružja.

Na drugom su mjestu oštri predmeti, na trećem da je osoba bila nenaoružana nakon čega slijede nepoznati i drugi čudni objekti.

## 7. Zaključak

Većina ljudi u današnje vrijeme ima pristup i koristi se velikom količinom podataka dostupnih na web preglednicima. Velika i srednja poduzeća se svakoga dana moraju koristiti takve podatke kako bi ispunili dobivene zadatke. Podaci koje koriste moraju biti organizirani i dostupni svima za jednostavno rukovanje. Skladišta podataka tu igraju veliku ulogu jer omogućuju prikupljanje, pohranu, izdvajanje i pretvorbu podataka koji dalje služe za analizu i obradu. Poduzeća sa organiziranim skladištima podataka tako unaprjeđuju svoje poslovanje

Na temelju ovog seminarskog rada došli smo do zanimljivih zaključaka. Kod nekih izvještaja smo moli pretpostaviti ishod kao na primjer da će veći postotak ubijenih biti kod muškaraca jer su oni skloni agresivnijem ponašanju i pokazivanju u društvu. Iz analize ipak vidimo da se ustvari radi o daleko većem broju. Kod analiziranja rasa, pretpostavili bi da se radi su crnci u većem broju radi izražene diskriminacije u njihovoj državi, ali na temelju izvještaja ipak dobivamo podatak da se radi o bijelcima. Također, da su uglavnom u pitanju pištolji i vatreno oružje je očekivan odgovor zbog dopuštenosti korištenja vatrenih dozvola i njihove lake kupnje.

## Popis literature

[1] Rabuzin, K., Predavanje 2: Skladište podataka., nastavni materijali na kolegiju Skladišta podataka i poslovna inteligencija [Moodle]. Sveučilište u Zagrebu, Fakultet organizacije i informatike Varaždin, 2022

[2] <https://www.kaggle.com/datasets/ahsen1330/us-police-shootings>

[3] Beal V., ETL – Extract, Transform, Load, 2019. Preuzeto 29.5.2023. s <https://www.webopedia.com/TERM/E/ETL.html>

[4] Rabuzin, K., Predavanje 6: ETL., nastavni materijali na kolegiju Skladišta podataka i poslovna inteligencija [Moodle]. Sveučilište u Zagrebu, Fakultet organizacije i informatike Varaždin, 2023.

## Popis slika

Slika 1. Preuzeta .csv datoteka .....	4
Slika 2. Transformacija podataka u Excelu .....	5
Slika 3. Table data import 1.dio .....	6
Slika 4. Table data import 2. dio .....	6
Slika 5. Table data import 3. dio .....	7
Slika 6. Učitani podaci u MySQL Workbench .....	7
Slika 7. Kreiranje tablice person_help .....	8
Slika 8. Tablica person_help .....	8
Slika 9. Kreiranje tablice date_help .....	9
Slika 10. Tablica date_help .....	9
Slika 11. Model zvijezde .....	10
Slika 12. Kreiranje dimenzijske tablice person_info.....	11
Slika 13. Upit za učitavanje podataka u tablicu person_info .....	11
Slika 14. Tablica person_info.....	12
Slika 15. Kreiranje dimenzijske tablice person_category .....	13
Slika 16. Upit za učitavanje podataka u tablicu person_category .....	13
Slika 17. Tablica person_category .....	14
Slika 18. Kreiranje dimenzijske tablice place_of_shooting .....	15
Slika 19. Upit za učitavanje podataka u tablicu place_of_shooting.....	15
Slika 20. Tablica place_of_shootings.....	16
Slika 21. Kreiranje dimenzijske tablice arms .....	17
Slika 22. Upit za učitavanje podataka u tablicu arms .....	17
Slika 23. Tablica arms .....	18
Slika 24. Kreiranje dimenzijske tablice police_report .....	19
Slika 25. Upit za učitavanje podataka u tablicu police_report .....	19
Slika 26. Tablica police_report .....	20
Slika 27. Kreiranje dimenzijske tablice date_and_time .....	21
Slika 28. Upit za učitavanje podataka u tablicu date_and_time.....	21
Slika 29. Tablica date_and_time .....	22
Slika 30. Kreiranje činjenične tablice evidention_of_shootings.....	23
Slika 31. Povezivanje ključeva.....	24
Slika 32. Upit za učitavanje podataka u tablicu evidention_of_shootings povezivanjem id-a	24
Slika 33. Tablica evidention_of_shootings .....	24
Slika 34. Model zvijezde izrađenog skladišta podataka.....	25
Slika 35. Dohvaćanje podataka u Power BI Desktopu.....	26
Slika 36. Spajanje na bazu unutar Power BI Desktopa .....	27
Slika 37. Odabir tablica za korištenje u Power BI Desktopu .....	27
Slika 38. Model zvijezde unutar alata Power BI Desktop.....	28
Slika 39. Prosjek godina ubijene osobe.....	29
Slika 40. Udio muškaraca i žena koji su ubijenu u policijskoj pucnjava.....	30
Slika 41. Gradovi u kojima se odvijalo najviše pucnjava .....	31
Slika 42. Udio rasa koje su sudjelovale u policijskim pucnjavama .....	32
Slika 43. Udio korištenog oružja u pucnjavama .....	33