SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

**Antonio Petak**

Analiza ubojstava u SAD-u od 2015. do 2020.

SEMINARSKI RAD

Varaždin, 2023.

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

Antonio Petak

Studij: Informacijske tehnologije i digitalizacija poslovanja

(*ili Poslovni sustavi, Ekonomika poduzetništva, Primjena informacijske tehnologije u poslovanju, Informacijsko i programsko inženjerstvo, Baze podataka i baze znanja, Organizacija poslovnih sustava, Informatika u obrazovanju*)

ANALIZA UBOJSTVA U SAD-u OD 2015. do 2020.

SEminarski rad

Mentor/Mentorica:

Mag. educ. inf. Maja Cerjan

Varaždin, prosinac 2023.

*Antonio Petak*

Izjava o izvornosti

Izjavljujem da je moj završni/diplomski rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

*Autor/Autorica potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Sažetak

Moj seminarski rad fokusira se na analizu ubojstava u SAD-u od 2015. do 2020. godine, koristeći pristup koji obuhvaća analizu i vizualizaciju podataka. Središnje pitanje istraživanja je postoji li rasna diskriminacija prema žrtvama.

Teorijsko-metodološki okvir rada temelji se na integraciji poslovne inteligencije s analizom podataka o ubojstvima, kao što su je li žrtva bila naoružana, pokušala pobjeći, mentalno bolesna itd. Vizualizacije su korištene za jasno prikazivanje rezultata, olakšavajući čitateljima razumijevanje kompleksnih informacija.

Glavne teze rada usmjerene su na identifikaciju uzročnika i faktora rizika povezanih s ubojstvima, istražujući geografske varijacije, dob žrtava te učestalost oružanih zločina. Analize su obuhvatile i promjene tijekom vremena, naglašavajući mjesece s najvećim i najmanjim stopama ubojstava.

Ključne riječi: ubojstvo; dob; rasa; spol; grad; oružje

Sadržaj

[Sadržaj iii](#_Toc154924555)

[1. Uvod 1](#_Toc154924556)

[2. Metode i tehnike rada 2](#_Toc154924557)

[3. Opis odabranog dataset-a 3](#_Toc154924558)

[4. Izrada skladišta podataka 4](#_Toc154924559)

[4.1. Skladište podataka 4](#_Toc154924560)

[4.1.1. Arhitektura skladišta podataka 4](#_Toc154924561)

[4.2. Model zvijezde 5](#_Toc154924562)

[4.3. Kreiranje dimenzijskih tablica 5](#_Toc154924563)

[4.3.1. Tablica armed (naoružan) 6](#_Toc154924564)

[4.3.2. Tablica arms\_categories (kategorija oružja) 6](#_Toc154924565)

[4.3.3. Tablica dates (datumi ubojstva) 7](#_Toc154924566)

[4.3.4. Tablica flee (pokušaj bijega) 7](#_Toc154924567)

[4.3.5. Tablica locations (lokacije ubojstva) 8](#_Toc154924568)

[4.3.6. Tablica manner\_of\_death (način smrti) 8](#_Toc154924569)

[4.3.7. Tablica persons (žrtve) 9](#_Toc154924570)

[4.3.8. Tablica threat\_level (razina prijetnje) 9](#_Toc154924571)

[4.4. Kreiranje činjenične tablice murders (ubojstva) 10](#_Toc154924572)

[5. ETL proces 11](#_Toc154924573)

[5.1. Extract – ekstrakcija podataka 11](#_Toc154924574)

[5.2. Transform – transformacija podataka 11](#_Toc154924575)

[5.3. Load – učitavanje podataka 11](#_Toc154924576)

[5.3.1. Učitavanje podataka u tablicu armed 11](#_Toc154924577)

[5.3.2. Učitavanje podataka u tablicu arms\_categories (kategorija oružja) 12](#_Toc154924578)

[5.3.3. Učitavanje podataka u tablicu dates (datumi ubojstva) 12](#_Toc154924579)

[5.3.4. Učitavanje podataka u tablicu flee (pokušaj bijega) 12](#_Toc154924580)

[5.3.5. Učitavanje podataka u tablicu locations (lokacije ubojstva) 13](#_Toc154924581)

[5.3.6. Učitavanje podataka u tablicu manner\_of\_death (način smrti) 13](#_Toc154924582)

[5.3.7. Učitavanje podataka u tablicu persons (žrtve) 13](#_Toc154924583)

[5.3.8. Učitavanje podataka u tablicu threat\_level (razina prijetnje) 13](#_Toc154924584)

[5.3.9. Učitavanje podataka u tablicu murders (ubojstva) 14](#_Toc154924585)

[6. Vizualizacija podataka 14](#_Toc154924586)

[6.1. MS Power BI 14](#_Toc154924587)

[6.2. Izrada izvještaja 15](#_Toc154924588)

[6.2.1. Interaktivni dashboard 15](#_Toc154924589)

[6.2.2. Lokacije ubojstava 16](#_Toc154924590)

[6.2.3. Hijerarhija datuma 17](#_Toc154924591)

[7. Zaključak 18](#_Toc154924592)

[Popis literature 19](#_Toc154924593)

[Popis slika 20](#_Toc154924594)

1. Uvod

Analiza ubojstava otvara vrata istraživanju dinamike kriminaliteta u Sjedinjenim Američkim Državama tijekom proteklog petogodišnjeg razdoblja. U godinama koje su obilježene društvenim promjenama, političkim nemirima i ekonomskim izazovima, pitanje ubojstava postaje ključno za dublje razumijevanje složenosti američkog društva. Ujedno, stope rasizma rastu iz dana u dan stoga je svrha ovog istraživanja otkriti koliko su zapravo bijelci „superiorniji“ naspram ostalim rasama.

Kroz analizu podataka o ubojstvima, istraživanje će prikazivati različite geografske lokacije, razlike u spolu žrtava, starosti, mentalnim bolestima, otporu prema policijskim službenicima, posjedovanju oružja, razini opasnosti žrtava prema društvu te promjene tijekom vremena kako bi prikazao kakva je međuovisnost. U ovom kontekstu, fokus će biti usmjeren na identifikaciju ključnih faktora koji utječu na razloge ubojstava, pružajući sagledavanje šireg društvenog konteksta.

Motivacija za odabir teme, koja se prožima s problematikom rasizma, proizlazi iz sažaljenja koje je izniklo iz tragičnog slučaja ubojstva Georgea Floyda. George Floyd postao je simbol nepravde i nehumanog postupanja prema osobama crne boje kože, pokazujući na izazove s kojima se suočavaju pripadnici određenih zajednica u suvremenom društvu. Brutalnost koja se dogodila dok je policajac držao koljeno na vratu Georgea Floyda i ignorirao njegove vapaje za pomoć izazvala je potrebu za dubljim razumijevanjem uzroka i posljedica takvih događaja. Upravo iz tog potresnog iskustva proizlazi volja za istraživanjem korijena rasizma, razumijevanjem mehanizama diskriminacije te traženjem načina kako se suprotstaviti toj poteškoći. Osim toga, pitanja koja se tiču preferencija žrtava, poput češćeg izlaganja riziku crne populacije ili muškaraca dovode do daljnjeg interesa u istraživanju ove teme.

1. Metode i tehnike rada

Za istraživanje teme koristit ću alat MS Access putem kojeg ću upravljati skladištem podataka i programski jezik SQL. U to spada organizacija, spajanje, čišćenje podataka s ciljem uređivanja podataka kako bi se mogli koristiti u daljnjoj analizi te vizualizaciji.

Vizualizaciju podataka provest ću kroz alat MS Power BI. Ovaj alat omogućuje stvaranje interaktivnih grafova, dashboard-a čime ću lakše prikazati uzroke, trendove i ključne zaključke ovog istraživanja.

Što se tiče samih podataka, koristit ću dataset sa platforme „Kaggle“. Dataset ću prilagoditi svojim potrebama i željama.

Kombinacija ovih metoda i alata omogućit će mi provođenje temeljite analize podataka o ubojstvima u Americi. SQL će poslužiti kao osnovna tehnika za ekstrakciju podataka, MS Access za njihovo organiziranje, a Power BI za vizualizaciju podataka. Dataset će mi pružiti autentične i relevantne informacije, osiguravajući da moje istraživanje bude temeljeno na relevantnim podacima.

1. Opis odabranog dataset-a

Dataset sam uzeo sa platforme Kaggle pod nazivom US Police Shootings autora Ahsen Nazir. Ovaj sadržaj sadrži osnovne podatke o ljudima poput njihovog imena, dobi, spola i rase. Uz to, sadrži informacije o ubojstvima, poput datuma događaja, mjesta gdje se dogodilo, načina kako su pogođeni, jesu li napali, jesu li držali oružje, jesu li pokazivali znakove mentalnih bolesti, je li policajac nosio kameru/je li incident snimljen, je li osumnjičeni pobjegao? Osim toga, stupac kategorija sadrži vrstu oružja koje je koristio osumnjičeni. Podatci su strukturirani u jednu tablicu pod nazivom shootings. Ta tablica se sastoji od id-a ubojstva, imena i prezimena žrtve, datuma događaja, načina smrti, jesu li bili naoružani, starost, spol, rasa, grad, država, znakovi mentalnih bolesti, razinu prijetnje, je li osumnjičeni pokušao trčati, je li incident snimljen i kategoriju oružja koju je osumnjičeni posjedovao.

Slika na kojoj se prikazuje tekst, snimka zaslona, broj, Font

Opis je automatski generiran

Slika 1. Struktura dataset-a

1. Izrada skladišta podataka
   1. Skladište podataka

„Data warehouse ili skladište podataka tvrtke (EDW), sustav je koji agregira podatke iz različitih izvora u jedno, centralno i konzistentno skladište podataka kako bi podržao analizu podataka, rudarenje podataka, umjetnu inteligenciju (AI) i strojno učenje. Sustav data warehousea omogućuje organizaciji izvođenje moćnih analiza na ogromnim količinama povijesnih podataka na načine koje standardna baza podataka ne može.

Sustavi za pohranu podataka bili su dio rješenja poslovne inteligencije (BI) već više od tri desetljeća, ali su se nedavno razvijali s pojavom novih vrsta podataka i metoda pohrane podataka. Tradicionalno, data warehouse bio je smješten na lokaciji - često na glavnom računalu i njegova funkcionalnost bila je usmjerena na izvlačenje podataka iz drugih izvora, čišćenje i pripremu podataka te učitavanje i održavanje podataka u relacijsku bazu podataka. Nedavno se data warehouse može smjestiti na posebni uređaj ili u oblak, a većina data warehousea dodala je mogućnosti analize te alate za vizualizaciju i prezentaciju podataka.“

(<https://www.ibm.com/topics/data-warehouse>)

* + 1. Arhitektura skladišta podataka

„Općenito govoreći, skladišta podataka imaju tro stupanjsku arhitekturu koja se sastoji od:

Donji stupanj: Donji stupanj sastoji se od poslužitelja za skladište podataka, obično relacijskog sustava baza podataka, koji prikuplja, čisti i transformira podatke iz različitih izvora podataka putem procesa poznatog kao Ekstrakcija, Transformacija i Učitavanje (ETL) ili procesa poznatog kao Ekstrakcija, Učitavanje i Transformacija (ELT).

Srednji stupanj: Srednji stupanj sastoji se od OLAP (online analitička obrada) poslužitelja koji omogućuje brze brzine upita. U ovom se stupnju mogu koristiti tri vrste OLAP modela, poznate kao ROLAP, MOLAP i HOLAP. Vrsta korištenog OLAP modela ovisi o vrsti sustava baza podataka koji postoji.

Gornji stupanj: Gornji stupanj predstavlja neku vrstu korisničkog sučelja ili izvještajnog alata koji omogućuje krajnjim korisnicima provođenje ad-hoc analize podataka o njihovim poslovnim podacima.“ (<https://www.ibm.com/topics/data-warehouse>)

* 1. Model zvijezde

Model zvijezde organizira podatke u dvije osnovne vrste tablica: činjenične tablice ( fact table) koje sadrže numeričke vrijednosti i dimenzijske tablice ( dimension table) koje sadrže tekstualne vrijednosti. Činjenične tablice daju odgovor na pitanje što, a dimenzijske tablice prema čemu mjerimo. Struktura ovog modela podsjeća na oblik zvijezde u kojem činjenična tablica predstavlja središnju točku, a dimenzijske tablice okružuju činjeničnu tablicu. Model zvijezde idealan je za procesiranje zbog redundantnosti. Ovakav pristup olakšava analizu i izvještavanje te se vrlo često u poslovnoj inteligenciji. (prezentacije s predavanja)

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, Trokut

Opis je automatski generiran

Slika 2. Model zvijezde

Model zvijezde sastoji se od činjenične tablice murders te dimenzijskih tablica: manner\_of\_death, flee, persons, locations, dates, armed, threat\_level i arms\_categories.

* 1. Kreiranje dimenzijskih tablica
     1. Tablica armed (naoružan)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font, broj

Opis je automatski generiran

Slika 3. Kreiranje tablice armed

Opis tablice:

* Naziv tablice armed
* armed\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* armed – tip podataka TEXT, može primiti do 255 znakova, sadrži vrstu oružja koju je posjedovala žrtva
  + 1. Tablica arms\_categories (kategorija oružja)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 4. Kreiranje tablice arms\_categories

Opis tablice:

* Naziv tablice arms\_categories
* arms\_category\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* category\_name – tip podataka TEXT, može primiti do 255 znakova, sadrži vrstu kategorije oružja koju je posjedovala žrtva
  + 1. Tablica dates (datumi ubojstva)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 5. Kreiranje tablice dates

Opis tablice:

* Naziv tablice dates
* date\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* incident\_date – tip podataka DATETIME, sadrži datum kada je žrtva ubijena
  + 1. Tablica flee (pokušaj bijega)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font, broj

Opis je automatski generiran

Slika 6. Kreiranje tablice flee

Opis tablice:

* Naziv tablice flee
* flee\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* incident\_date – tip podataka TEXT, može primiti do 255 znakova, sadrži oblik pokušaja bijega
  + 1. Tablica locations (lokacije ubojstva)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 7. Kreiranje tablice locations

Opis tablice:

* Naziv tablice locations
* location\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* city – tip podataka TEXT, može primiti do 255 znakova, sadrži naziv grada ubojstva
* state – tip podataka CHAR, prima isključivo 2 znaka, sadrži kraticu imena države u kojoj je žrtva ubijena
  + 1. Tablica manner\_of\_death (način smrti)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 8. Kreiranje tablice manner\_of\_death

Opis tablice:

* Naziv tablice manner\_of\_death
* manner\_of\_death\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* manner\_of\_death – tip podataka TEXT, može primiti do 255 znakova, sadrži način ubojstva
  + 1. Tablica persons (žrtve)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font, softver

Opis je automatski generiran

Slika 9. Kreiranje tablice persons

Opis tablice:

* Naziv tablice persons
* person\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* name – tip podataka TEXT, može primiti do 255 znakova, sadrži ime i prezime žrtve
* age – tip podataka INT, sadrži starost žrtve
* gender – tip podataka CHAR, sadrži jedno slovo, M za muškarac ili F za žena
* race – tip podataka TEXT, može primiti do 20 znakova, sadrži rasu žrtve
* signs\_of\_mental\_ilness – tip podataka YESNO (BOOLEAN), prikazuje je li žrtva pokazivala znakove duševne bolesti
  + 1. Tablica threat\_level (razina prijetnje)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font, broj

Opis je automatski generiran

Slika 10. Kreiranje tablice threat\_level

Opis tablice:

* Naziv tablice threat\_level
* threat\_level\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* threat\_level – tip podataka TEXT, može primiti do 255 znakova, sadrži razinu prijetnje
  1. Kreiranje činjenične tablice murders (ubojstva)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 11. Kreiranje tablice murders

Opis tablice:

* Naziv tablice murders
* murders\_ID – autoincrement, primarni ključ koji u sebi sadrži NOT NULL i UNIQUE
* person\_ID – tip podataka LONG, vanjski ključ na tablicu persons, atribut person\_ID
* date\_ID - tip podataka LONG, vanjski ključ na tablicu dates, atribut date\_ID
* location\_ID - tip podataka LONG, vanjski ključ na tablicu locations, atribut location\_ID
* armed\_ID - tip podataka LONG, vanjski ključ na tablicu armed, atribut armed\_ID
* manner\_of\_death\_ID - tip podataka LONG, vanjski ključ na tablicu manner\_of\_death, atribut manner\_of\_death\_ID
* threat\_level\_ID - tip podataka LONG, vanjski ključ na tablicu threat\_level, atribut threat\_level\_ID
* flee\_ID - tip podataka LONG, vanjski ključ na tablicu flee, atribut flee\_ID
* body\_camera – tip podataka YESNO (BOOLEAN), pokazuje je li policajac nosio kameru za tijelo tokom ubojstva
* arms\_category\_ID - tip podataka LONG, vanjski ključ na tablicu arms\_categories, atribut arms\_category\_ID

1. ETL proces
   1. Extract – ekstrakcija podataka

Sirove podatke sam preuzeo sa stranice Kaggle. Naziv dataseta je US Police Shootings, autora Ahsen Nazir. Nakon preuzimanja podataka ubacio sam ih u MS Access kako bi ih mogao prilagoditi i transformirati. Podaci su pohranjeni u obliku jedne tablice.

* 1. Transform – transformacija podataka

Kod ubacivanja podataka u MS Access, podaci o starosti žrtava su dobili nulu na kraju brojeva, npr. umjesto 53 se prikazivalo 530. Problem je riješen izvođenjem sljedećeg koda: Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 12. Popravljanje godina

Sljedeći problem je bio taj da su neki redovi poprimili null vrijednosti koje sam zatim izbrisao iz dataseta. Ostalih potreba za transformacijom nije bilo.

* 1. Load – učitavanje podataka
     1. Učitavanje podataka u tablicu armed

Slika na kojoj se prikazuje tekst, Font, snimka zaslona, broj

Opis je automatski generiran

Slika 13. Učitavanje podataka u tablicu armed

Podaci iz dataseta (tablica shootings), atributa armed, učitavaju se u atribut armed dimenzijske tablice armed na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu arms\_categories (kategorija oružja)

Slika na kojoj se prikazuje tekst, Font, snimka zaslona

Opis je automatski generiran

Slika 14. Učitavanje podataka u tablicu arms\_categories

Podaci iz dataseta (tablica shootings), atributa arms\_category, učitavaju se u atribut category\_name dimenzijske tablice arms\_categories na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu dates (datumi ubojstva)

Slika na kojoj se prikazuje tekst, Font, snimka zaslona

Opis je automatski generiran

Slika 15. Učitavanje podataka u tablicu dates

Podaci iz dataseta (tablica shootings), atributa date, učitavaju se u atribut incident\_date dimenzijske tablice dates na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu flee (pokušaj bijega)

Slika na kojoj se prikazuje tekst, Font, snimka zaslona, broj

Opis je automatski generiran

Slika 16. Učitavanje podataka u tablicu flee

Podaci iz dataseta (tablica shootings), atributa flee, učitavaju se u atribut flee dimenzijske tablice flee na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu locations (lokacije ubojstva)

Slika na kojoj se prikazuje tekst, Font, snimka zaslona

Opis je automatski generiran

Slika 17. Učitavanje podataka u tablicu locations

Podaci iz dataseta (tablica shootings), atributi city i state, učitavaju se u atribute city i state dimenzijske tablice locations na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu manner\_of\_death (način smrti)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 18. Učitavanje podataka u tablicu manner\_of\_death

Podaci iz dataseta (tablica shootings), atributa manner\_of\_death, učitavaju se u atribut manner\_of\_death dimenzijske tablice manner\_of\_death na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu persons (žrtve)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 19. Učitavanje podataka u tablicu persons

Podaci iz dataseta (tablica shootings), atributi name, age, gender, race, signs\_of\_mental\_illness, učitavaju se u atribute name, age, gender, race, signs\_of\_mental\_ilness dimenzijske tablice persons na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu threat\_level (razina prijetnje)

Slika na kojoj se prikazuje tekst, Font, snimka zaslona, broj

Opis je automatski generiran

Slika 20. Učitavanje podataka u tablicu threat\_level

Podaci iz dataseta (tablica shootings), atributa threat\_level, učitavaju se u atribut threat\_level dimenzijske tablice threat\_level na način da se biraju samo različite vrijednosti i gdje su vrijednosti različite od nule.

* + 1. Učitavanje podataka u tablicu murders (ubojstva)

Slika na kojoj se prikazuje tekst, snimka zaslona, Font

Opis je automatski generiran

Slika 21. Učitavanje podataka u tablicu murders

Dimenzijske tablice armed, arms\_categories, dates, flee, locations, manner\_of\_death, persons i threat\_level spajaju se kako bi se podaci iz njih mogli učitati u činjeničnu tablicu murders.

1. Vizualizacija podataka
   1. MS Power BI

„Microsoft Power BI je platforma za vizualizaciju podataka koja se koristi prvenstveno u svrhu poslovne inteligencije. Namijenjena je poslovnim profesionalcima s različitim razinama poznavanja podataka, a nadzorna ploča Power BI-ja sposobna je izvještavati i vizualizirati podatke u različitim stilovima, uključujući grafove, karte, dijagrame, raspršene dijagrame i više. Funkcionalnost "AI Insights" u Power BI-ju koristi umjetnu inteligenciju kako bi pronašla uvide unutar skupova podataka za korisnike.“ (<https://www.coursera.org/articles/what-is-power-bi>)

Razlog zašto sam odlučio koristiti MS Power BI umjesto Tableau je taj da mi je korisničko sučelje intuitivnije i lakše za shvatiti. Time je kreiranje vizualizacija brže i jednostavnije.

* 1. Izrada izvještaja
     1. Interaktivni dashboard

Slika na kojoj se prikazuje tekst, dijagram, snimka zaslona, radnja

Opis je automatski generiran

Slika 22. Interaktivni dashboard

Dashboard se sastoji od sljedećih elemenata:

* Linijski grafikon koji prikazuje broj ubojstava po godini
* Složeni stupčasti grafikon koji prikazuje distribuciju oružanih kategorija koje je žrtva posjedovala
* Tortni grafikon koji prikazuje brojeve i postotke ubojstava prema rasama
* Broj ubojstava

Filteri:

* Starost – raspon od 12 do 91 godina
* Spol – F za ženski i M za muški
* Pokušaj bijega – autom, pješice, nije bježao/la, ostalo
* Znakovi duševne bolesti – da ili ne
* Način smrti – upucan ili šokiran elektrošokom i upucan
* Kamera za tijelo – prisutna ili ne

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, radnja

Opis je automatski generiran

Slika 23. Primjer filtera

Uključeni su filteri raspon godina od 12 do 91, samo muškarci, žrtva nije pokušala pobjeći, žrtva je pokazivala znakove duševnih bolesti, policajac nije imao kameru za tijelo i žrtva je bila crne rase. Prema ovoj statistici možemo vidjeti da je ubijeno 188 osoba koje odgovaraju našem filteru, najviše žrtava je imalo kod sebe neku vrstu vatrenog oružja te da se najviše takvih ubojstava desilo u 2018. godini i to njih 47.

* + 1. Lokacije ubojstava

Slika na kojoj se prikazuje tekst, snimka zaslona, dijagram, karta

Opis je automatski generiran

Slika 24. Lokacije ubojstava

Na slici 24. je prikazana karta SAD-a i države u kojima su se desila ubojstva. Složenim trakastim grafikonom prikazani su gradovi sa brojem ubojstava po pojedinom gradu te broj ubojstava koji se mijenja ovisno o odabranom gradu ili državi. Veličina kružića na mapi pokazuje količinu ubojstava po državi. Može se vidjeti da se najviše ubojstava desilo u gradu Houston i to njih 230, dok je država s najviše ubojstava Kalifornija sa 1054 ubojstva.

* + 1. Hijerarhija datuma

Slika na kojoj se prikazuje tekst, Font, dijagram, snimka zaslona

Opis je automatski generiran

Slika 25. Hijerarhija datuma

Slika na kojoj se prikazuje tekst, snimka zaslona, Font, crta

Opis je automatski generiran

Slika 26. Dashboard hijerarhije datuma

Najviše ubojstava se desilo u mjesecu svibnju i to njih 757. U prvom tromjesečju desilo se najviše ubojstava i taj broj iznosi 2110, dok se 29-tog dana u mjesecu desilo 436 ubojstava što je i najveći broj ubojstava po danu. Pored grafova se nalaze filteri u kojima se može birati određeni mjesec, tromjesečje ili dan.

1. Zaključak

U ovom seminarskom radu koristio sam stranicu Kaggle kako bih pronašao dataset za izradu skladišta podataka i izradu samih izvještaja. Nakon pronalaska dataseta isti sam ubacio u MS Access te krenuo s procesom transformacije podataka. Napravio sam skladište podataka u modelu zvijezde te sam u njega ubacio transformirane podatke. Zatim je slijedila vizualizacija podataka. Odlučio sam koristiti alat MS Power BI zbog svoje jednostavnosti i kvalitetnosti. Kao rezultate vizualizacija možemo vidjeti podatke kako su se kretala ubojstva po danima, mjesecima, kvartalima i godinama. Možemo vidjeti njihove lokacije, spol, starost, znakove duševnih bolesti, jesu li pokušali pobjeći, jesu li imali neko oružje kod sebe, rasu, na koji način su ubijeni i slično.

Popis literature

prezentacije s predavanja

<https://www.ibm.com/topics/data-warehouse> (Pristupano 30.12.2023.)

<https://www.coursera.org/articles/what-is-power-bi> (Pristupano 30.12.2023.)

<https://www.kaggle.com/datasets/ahsen1330/us-police-shootings> (dataset, Pristupano 30.12.2023.)

Popis slika

[Slika 1. Struktura dataset-a 3](#_Toc154924529)

[Slika 2. Model zvijezde 5](#_Toc154924530)

[Slika 3. Kreiranje tablice armed 6](#_Toc154924531)

[Slika 4. Kreiranje tablice arms\_categories 6](#_Toc154924532)

[Slika 5. Kreiranje tablice dates 7](#_Toc154924533)

[Slika 6. Kreiranje tablice flee 7](#_Toc154924534)

[Slika 7. Kreiranje tablice locations 8](#_Toc154924535)

[Slika 8. Kreiranje tablice manner\_of\_death 8](#_Toc154924536)

[Slika 9. Kreiranje tablice persons 9](#_Toc154924537)

[Slika 10. Kreiranje tablice threat\_level 9](#_Toc154924538)

[Slika 11. Kreiranje tablice murders 10](#_Toc154924539)

[Slika 12. Popravljanje godina 11](#_Toc154924540)

[Slika 13. Učitavanje podataka u tablicu armed 11](#_Toc154924541)

[Slika 14. Učitavanje podataka u tablicu arms\_categories 12](#_Toc154924542)

[Slika 15. Učitavanje podataka u tablicu dates 12](#_Toc154924543)

[Slika 16. Učitavanje podataka u tablicu flee 12](#_Toc154924544)

[Slika 17. Učitavanje podataka u tablicu locations 13](#_Toc154924545)

[Slika 18. Učitavanje podataka u tablicu manner\_of\_death 13](#_Toc154924546)

[Slika 19. Učitavanje podataka u tablicu persons 13](#_Toc154924547)

[Slika 20. Učitavanje podataka u tablicu threat\_level 14](#_Toc154924548)

[Slika 21. Učitavanje podataka u tablicu murders 14](#_Toc154924549)

[Slika 22. Interaktivni dashboard 15](#_Toc154924550)

[Slika 23. Primjer filtera 16](#_Toc154924551)

[Slika 24. Lokacije ubojstava 16](#_Toc154924552)

[Slika 25. Hijerarhija datuma 17](#_Toc154924553)

[Slika 26. Dashboard hijerarhije datuma 17](#_Toc154924554)