

# Gun Violence in de VS

## een analyse

Jonas Kouwenhoven

Martijn Maiwald

Folkert Stijnman

Casper Wortmann

28 juni 2018

## I. Inleiding

Gedurende de jaren 2003 tot 2012, de meeste recente jaren waarover officiële data beschikbaar is met betrekking tot *gun violence*, stierven in de Verenigde Staten 313.045 personen aan vuurwapen gerelateerde verwondingen. Dit zijn meer doden dan het aantal Amerikaanse soldaten dat stierf tijdens de Tweede Wereldoorlog, zelfs dan het aantal soldaten van de Verenigde Staten dat stierf in alle oorlogen uit haar historie [1]. *Gun violence* is daarom niet geheel verrassend een veelbesproken onderwerp in de Verenigde Staten en daarbuiten.

Meer recentelijk worden archieven over *gun violence* online en openbaar bijgehouden en daarmee kunnen databases worden gemaakt die deze incidenten en bijbehorende data verzamelen. Incidenten met vuurwapens zijn door de politie geregistreerde incidenten waarbij wapens waren betrokken. Dit zijn incidenten waarbij geschoten is, of enkel wapens zijn gevonden. Voorbeelden zijn zelfmoord, overvallen, huiszoekingen etc. De dataset die in dit onderzoek gebruikt wordt bevat alle vuurwapen gerelateerde incidenten in de Verenigde Staten van 1/1/2013 tot 3/31/2018. Deze incidenten worden geanalyseerd op basis van vijf deelvragen. De antwoorden op deze deelvragen kunnen wellicht duidelijkheid scheppen over de maatschappelijk problemen rondom gun violence.

Als eerste is onderzocht hoe de incidenten zich verdelen over de verschillende staten en de verschillende jaren. De eerste deelvraag was dus of er aanzienlijke verschillen zijn tussen de jaren en staten. De tweede vraag was welke onverwachte patronen in de dataset te vinden waren. Als derde werd er een antwoord gezocht op de vraag wat de verhouding is tussen verdachten en slachtoffers bij alle incidenten. Hierbij werd ook gekeken naar hoe de verdachten en slachtoffers in verhouding staan tot elkaar. Zijn zij bijvoorbeeld vaker familie van elkaar, of vinden incidenten vooral tussen criminele bendes plaats?

Uit onderzoek is gebleken dat incidenten waarbij vuurwapens zijn betrokken een disproportioneel effect kunnen hebben op jonge mannen, en dat deze incidenten grote gevolgen hebben voor zowel deze jonge slachtoffers als voor de gehele maatschappij [2]. Omdat betrokkenheid in incidenten met vuurwapens

zeer schadelijk kan zijn voor jongeren wordt in dit onderzoek daarom als vierde geprobeerd een antwoord te geven op de vraag: in welke maat en welke vorm zijn jong volwassenen (18-25) betrokken bij vuurwapen gerelateerde incidenten? Als laatste is gekeken naar de mate van betrokkenheid van de politie, aangezien zij de controle uitvoeren over de maatschappij. De laatste deelvraag is daarom of er een verschil is tussen incidenten waarbij politie was betrokken als slachtoffer of verdachte, en incidenten de incidenten waarbij politie niet op een dergelijke manier was betrokken.

## II. Methode

De dataset die werd gebruikt is beschikbaar via github <sup>1</sup>. Deze dataset is verkregen met behulp van *web scraping techniques*, waarmee gegevens zijn opgehaald uit de database van de *Gun Violence Archive*<sup>2</sup>. In de daaruit verkregen dataset is elk incident vastgelegd, en beschreven op basis van verscheidene categorieën. Zo heeft elk incident een adres, datum en het aantal betrokkenen. In Table 1 wordt een volledig overzicht gegeven van alle categorieën. Alle categorieën die doorverwijzen naar extra beschikbare informatie over een incident op het internet, zoals 'incident\_url' zijn bij dit onderzoek buiten beschouwing gelaten. Dit is omdat het onderzoek zich alleen richt op de feiten die uit de dataset zelf te halen zijn. Ook zijn categorieën met een te diverse inhoud niet gebruikt, zoals 'location\_description' en 'Sources'. De inhoud van die categorieën is immers zo divers, dat zij voor het doel en de reikwijdte van dit onderzoek geen toegevoegde waarde hebben. Tot slot zijn de categorieën 'Adress', 'State\_house\_discript', 'State\_senate\_discript', en 'Congressional\_district' ook buiten beschouwing gelaten, omdat ze voor de grootte van dit onderzoek triviaal werden bevonden. Vergelijkbare informatie kan immers ook worden opgehaald uit de categorieën 'state' en 'city\_or\_county'.

Er zijn een aantal voorverwerkingen gedaan voordat de dataset kon worden geanalyseerd. Ten eerste is de CSV dataset opgeschoond. Op verschei-

---

<sup>1</sup><https://github.com/jamesqo/gun-violence-data>

<sup>2</sup><http://www.gunviolencearchive.org/>

Table 1: Categorieën

<b>Gebruikte categorieën</b>	<b>Beschrijving</b>
incident_ID	nummer incident in database
date	de datum van het incident
state	de staat waarin het incident plaatsvond
city_or_county	de staat waarin het incident plaatsvond
n_killed	het aantal doden ten tijde van registratie
n_injured	het aantal gewonden ten tijde van registratie
incident_characteristics	kenmerken zoals, geschoten, wapens gevonden, zelfmoord, incident waarbij politie was betrokken, incident op een school
latitude	de breedtegraad van het incident
location_description	Een locatie omschrijving zoals een terras, een club, restaurant, theater etc.
longitude	de hoogtegraad van het incident
participant_characteristics	leeftijd, leeftijdsgroep, geslacht, naam, mogelijke relatie tot andere betrokken, status m.b.t. incident (gearresteerd, omgekomen, gewond), type (verdachte of slachtoffer)
<b>Niet gebruikte categorieën</b>	<b>Beschrijving</b>
Address	Het adres van het incident
Incident_url	Link naar gunviolencearchive.org met details van het incident
Source_url	Link naar online nieuws verslag van het incident
Incident_url_fields_missing	geeft aan of er een Incident_url mist middels False
location_description	Een locatie omschrijving zoals een terras, een club, restaurant, theater etc.
Congressional district	congress district
Notes	mogelijke notities bij het incident
gun_stolen	aantal gevonden wapens dat gestolen was
gun_type	het soort wapen
n_guns_involved	aantal wapens dat geregistreerd werd die betrokken
Sources	links naar online nieuwsverslag over het incident
State_house_district	state house district
State_senate_district	state senate district

*Overzicht van naam en beschrijving van categorieën uit de dataset*

dene plekken zaten witregels en dubbele spaties, deze zijn verwijderd. Alle cellen die geen waarde hadden, zijn opgevuld door 'NaN' (Not a Number). Van alle waarden in de dataset, was slechts 82,5% ingevuld. De rest is vervangen door 'NaN'. Daarna is de dataset omgezet van CSV naar JSON. Om twee redenen is daarvoor gekozen. Ten eerste heeft JSON bij grote datasets een voordeel voor de *leesbaarheid* van de dataset. In een CSV bestand worden alle waarden per incident gescheiden door comma's, waarbij elke nieuwe comma een nieuwe kolom representeert. In een JSON worden alle waardes voor ieder incident gekoppeld aan de bijbehorende categorie. Zo is de informatie per incident geclassificeerd en kan er makkelijker door de dataset worden gezocht. Ten tweede werd in deze CSV in sommige kolommen gebruik gemaakt van een *dictionary structure*. Deze structuur heeft sterke overeenkomsten met JSON.

Tijdens de eerste analyses kwam naar voren dat in de dataset een aantal vermoedelijk incorrect uitschieters zaten. Zo werden er leeftijden van 209 en 311 gevonden. De incidenten met dergelijke waarden zijn uit de dataset verwijderd, om betrouwbaarheid van het onderzoek te kunnen waarborgen. Tot slot is er, om geografische vergelijkingen te kunnen doen, een dataset toegevoegd met aantal inwoners per staat. De betreffende dataset met aantal inwoners is afkomstig van the United States Census Bureau<sup>3</sup>, het officiële overheidsorgaan van de VS dat volkstellingen doet. Er vinden immers meer incidenten plaats in dichtgevolkte gebieden. Bij alle analyses waarbij verschillende staten werden vergeleken, is daarom genormaliseerd door het aantal incidenten te delen door het aantal inwoners van die staat in dat jaar. Zo werd rekening gehouden met een mogelijke groei of daling van de bevolking.

---

<sup>3</sup><https://www.census.gov>

## Resultaten

Allereerst werd gekeken naar het aantal incidenten per jaar. Zoals te zien is in Figure 1, waren er in 2017 meer incidenten per hoofd van de bevolking dan in 2014. In Figure 2 is te zien hoe deze groei of daling per staat is. Om een duidelijker overzicht te krijgen van de geografische verdeling, zijn de incidenten per hoofd van de bevolking per staat ook weergegeven op een kaart. Zoals in Figure 6 te zien is, is er een cluster van staten te zien waar de meeste incidenten plaatsvonden, aan de oostkant van centraal Amerika.

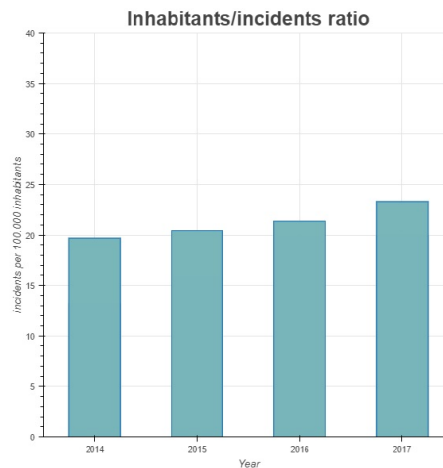


Figure 1: Incidenten per jaar

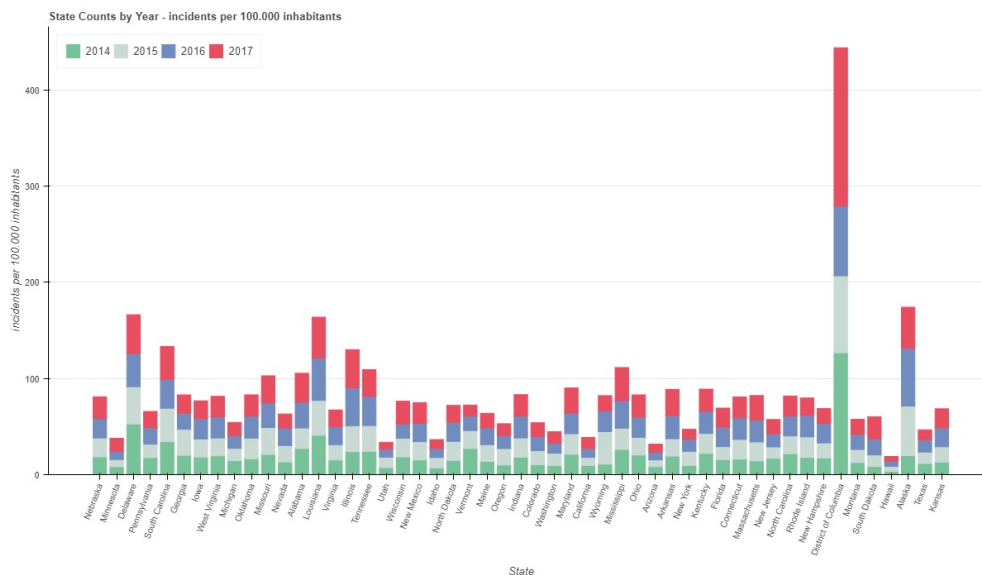


Figure 2: Aantal incidenten per staat

Daarna is onderzocht welke onverwachte patronen in de dataset te vinden waren. Zo werd gevonden dat veel meer mannen zijn betrokken bij incidenten dan vrouwen. Zoals in Figure 3(a) te zien is, zijn er veel meer mannen zowel verdachte als slachtoffer bij incidenten. Ook zijn er bepaalde dagen in het jaar waar bovengemiddeld veel incidenten plaatsvinden. Vaak zijn dit feestdagen. Figure 3(b) laat zien dat de meeste incidenten, in verhouding tot andere feestdagen, op nieuwjaarsdag plaatsvinden. Op *Fourth of July*, een nationale feestdag in de Verenigde Staten, vinden het minste aantal incidenten plaats.

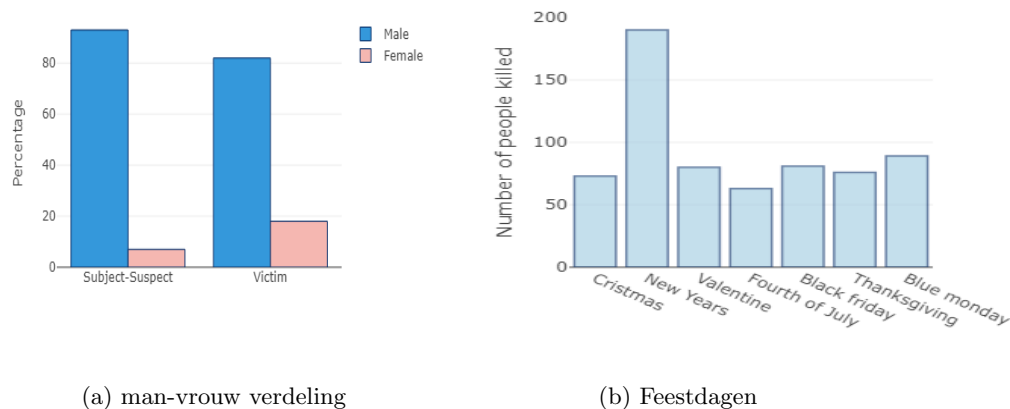


Figure 3: Onverwachte patronen

In Figure 4(a) is te zien dat in de grootste steden de verdeling tussen verdachte en slachtoffer ongeveer gelijk is. In 4(b) wordt gekeken naar de percentages van incidenten met betrekking tot gewapende overvallen, huiszoekingen, vrienden, bendes of familie. Zo is in Figure 4(b) te zien dat de meeste incidenten in 2017 te maken hadden met gewapende overvallen en de minste incidenten te maken hadden met criminele bendes.

Vervolgens werd gekeken naar jong volwassen en de mate van betrokkenheid in incidenten. Hierbij werd elk type incident gerelateerd aan de gemiddelde leeftijd. Zoals in Figure 5(a) te zien is, zijn vooral jonge man-

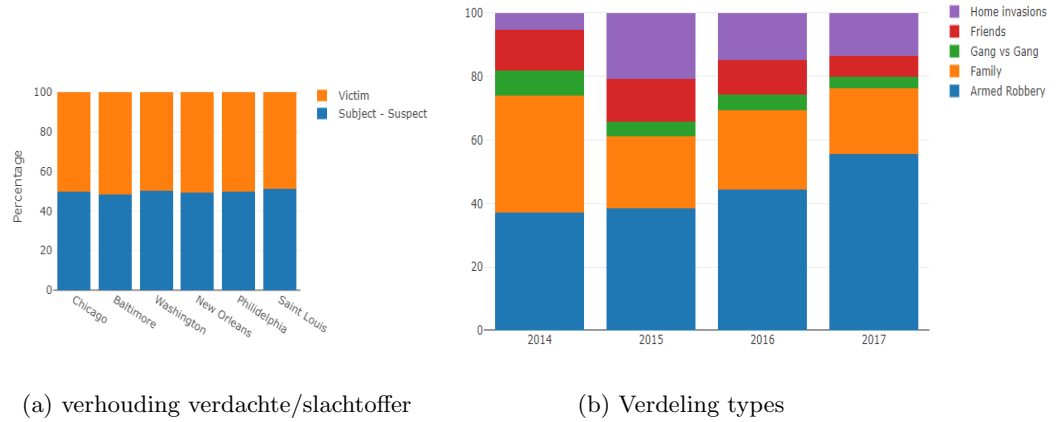


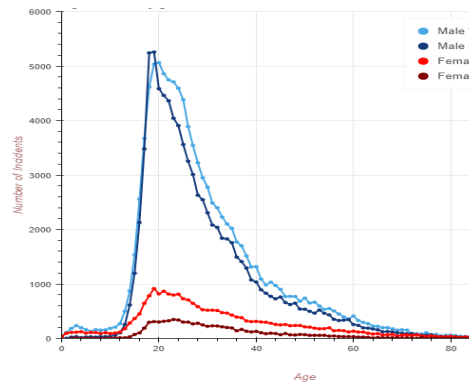
Figure 4: Verdeling mannen en vrouwen

nen betrokken bij incidenten. De verdeling bij de vrouwen is wat meer gelijk verdeeld. Als gekeken wordt naar de gemiddelde leeftijd van een verdachte in verhouding tot het type incident dat hij pleegt, is te zien dat *Gang vs Gang* incidenten vooral voorkomen bij verdachten met een relatief jonge leeftijd. Incidenten waarbij burens betrokken zijn vinden juist op relatief oude leeftijd plaats. Er zit weinig verschil tussen de gemiddelde leeftijd die vrouwen hebben als ze een bepaald type incident plegen, ten opzichte van de leeftijd die mannen hebben als ze een bepaald type incident plegen.

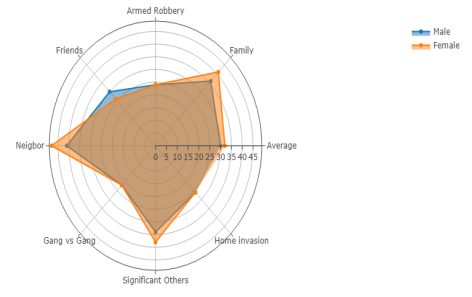
De meeste incidenten waarbij politie is betrokken (zowel als verdachte als als slachtoffer), vinden plaats in de steden. De incidenten vinden zowel in het oosten plaats, als in het westen. Als echter wordt gekeken naar het aantal incidenten per hoofd van de bevolking, is te zien dat deze juist in het centrale gedeelte van de Verenigde staten het meeste plaats vinden (Figure 7(B)). De gemiddelde leeftijd waarmee een burger op de politie schiet, is bovengemiddeld hoog (Figure 8(a)). Een opvallend resultaat wordt verkregen als wordt gekeken naar het aantal onbekende variabelen ('NaN') in de categorieën die alleen op een slachtoffer of verdachte van toepassing zijn. Dat betreft bijvoorbeeld de categorieën 'participant\_age' en 'participant\_gender', en niet 'date' of 'state', die over het gehele incident gaan. Over de gehele dataset ontbreekt 47% van



de waarden die over een slachtoffer of verdachte gaan. Echter bij incidenten waarbij de politie betrokken is ligt dat percentage op 68%.



(a) man-vrouw verdeling



(b) gem. leeftijden/type

Figure 5: Verdeling mannen en vrouwen

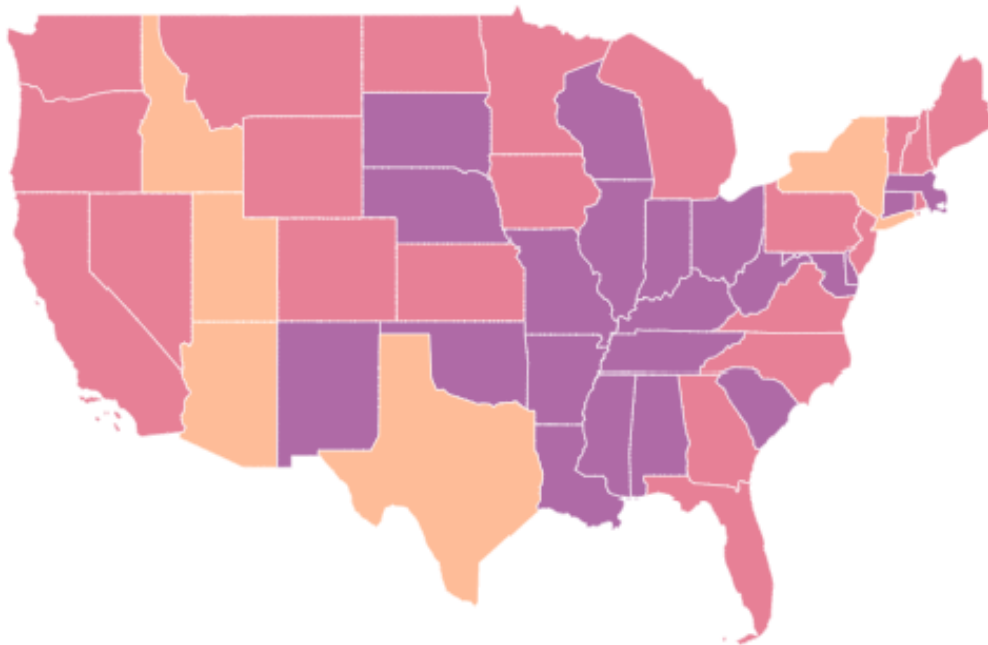
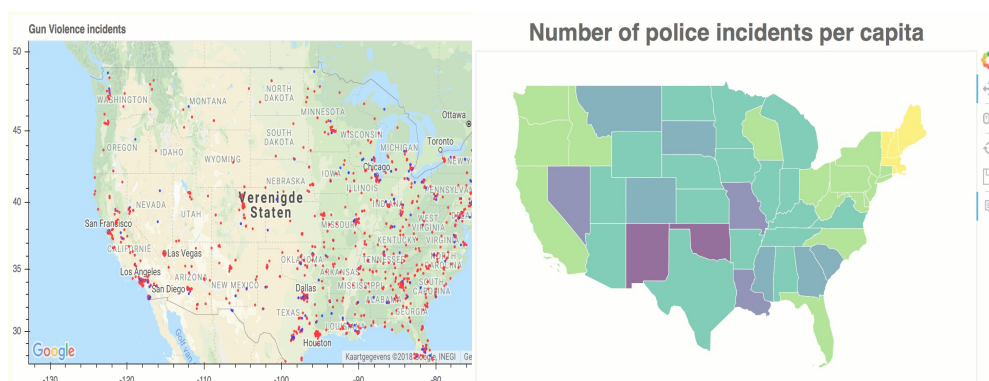


Figure 6: Aantal incidenten per staat per hoofd van de bevolking (2017)



(a) Normale kaart

(b) Kleurenkaart

Figure 7: 1 en 2 van politie

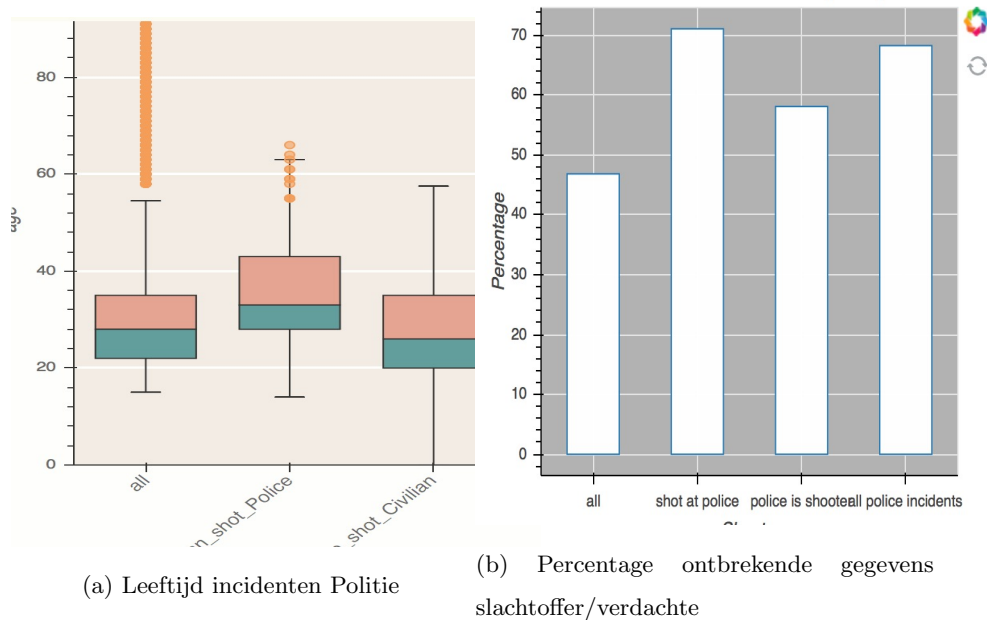


Figure 8: analyse politie incidenten

## IV. Discussie

Op basis van de resultaten kunnen een aantal van de vooraf gestelde vragen beantwoord worden. De eerste vraag was hoe het aantal incidenten zich per staat over de jaren heeft ontwikkeld. Op basis van de resultaten kan worden geconcludeerd dat in de periode van 2014 tot en met 2017 steeds meer incidenten plaatsvonden. Ook lijkt er met name een groei van het aantal incidenten te zijn rondom het oostelijke deel van het midden van de Verenigde Staten. Ook kwamen er een aantal onverwachte patronen uit de dataset naar voren. Er waren aanzienlijk meer mannen dan vrouwen zowel verdachte als slachtoffer. Daarnaast geschieden vuurwapen gerelateerde incidenten op feestdagen, vooral op nieuwjaar, en het minst op de nationale feestdag van de Verenigde Staten, de *Fourth of July*. De derde vraag betrof de verdeling van verdachte en slachtoffers bij incidenten en welke relaties voornamelijk voorkwamen. Opvallend is dat de verdeling van slachtoffer en verdachte voor alle steden ongeveer

gelijk verdeeld is. Een verklaring hiervoor zou kunnen zijn dat het merendeel van de gevallen één dader en één slachtoffer betreft en dat dit de uitschieters uitbalanceert. Een precieze verklaring voor deze verdeling zou mogelijk in verder onderzoek bekeken kunnen worden. De relatie die het meest voorkwam was een gewapende overval. In 2014 lag dit aantal dicht tegen het aantal familiegerelateerde incidenten. Belangrijk om hierbij te vermelden is dat bij zeer weinig incidenten een relatie is geregistreerd: slechts bij 6.5 procent van de incidenten. Verder onderzoek waarbij het type incident preciezer is bijgehouden is daarom van belang.

Een antwoord op de vraag in welke maat jong volwassenen (18-25) betrokken zijn bij vuurwapen gerelateerde incidenten is snel gegeven: bovengemiddeld veel. Uit de resultaten blijkt dat vooral jonge mannen betrokken zijn, terwijl de leeftijden bij vrouwen die betrokken zijn beter is verdeeld. Een verklaring hiervoor zou kunnen zijn dat mannen cultureel gezien op jongere leeftijd in aanraking komen met vuurwapens en dus makkelijker betrokken raken bij dergelijke incidenten. Ook is gekeken naar de verschillende relaties ten opzichte van de gemiddelde leeftijd bij zowel mannen als vrouwen. Daar is terug te zien dat bij alle type incidenten de gemiddelde leeftijd bij vrouwen hoger ligt dan bij mannen. De gemiddelde leeftijd van mannen ligt alleen hoger bij mannen als het een incident tussen burens betreft.

De verdeling van het aantal incidenten waarbij de politie betrokken is lijkt op het totaal aan incidenten per hoofd van de bevolking. Als gekeken wordt naar alle incidenten zijn de meeste incidenten aan de kust staten, waar het dichtst bevolkt is. Echter in het midden van de VS zijn per hoofd van de bevolking de meeste incidenten, zowel voor politie als voor niet-politie gerelateerde incidenten. De politie gerelateerde en niet-politie gerelateerde incidenten lijken dus op elkaar, maar er zit wel een groot verschil in het aantal niet-geregistreerde gegevens tussen beiden. Een verklaring hiervoor kan zijn dat bij een schot op een politie de dader makkelijker weg kan komen. Dat zou echter geen goed teken zijn voor de mate waarin politie-incidenten in retrospect worden opgelost. Ook in het geval dat de politie het schot heeft gevuurd zijn er bovengemiddeld

veel ontbrekende gegevens. Met betrekking op de publieke aandacht die *police brutality* de afgelopen jaren heeft gehad, is dat een des te opvallende verschijnsel. Is een verklaring dat zo min mogelijk mensen hun handen willen branden aan gegevens waarbij politie is betrokken? Worden politie-gerelateerde incidenten slechter geregistreerd omdat ze direct van hoger af worden afgehandeld? De resultaten roepen vooral nieuwe vragen op. Meer informatie over politie-gerelateerde incidenten is van belang, ook de komende jaren. Hoewel het nu nog niet te zien is, zou er best een verandering in kunnen zitten over de jaren heen.

Een aantal zaken zijn van belang om in ogenschouw te nemen wanneer gekeken wordt naar de resultaten uit dit onderzoek. Allereerst betreft de dataset enkel geregistreerde gevallen van vuurwapen gerelateerde incidenten. In stedelijke gebieden zullen dergelijke incidenten bijvoorbeeld eerder zijn gemeld dan in de gebieden daarbuiten. Zo bestaat de Verenigde Staten uit een enorme oppervlakte waar soms weinig tot geen mensen wonen en dus ook minder politie actief is. Er zouden zodoende gevallen kunnen zijn waarbij het incident niet geregistreerd is. Daarnaast zijn bepaalde incidenten wellicht onvolledig, omdat niet elke betrokkene geregistreerd heeft kunnen worden. Daders zouden het plaats delict al kunnen hebben ontvlucht voordat de politie is gearriveerd en bij de registratie van dergelijke incidenten kan hier geen rekening mee worden gehouden. Zodoende betreft deze dataset enkel een steekproef onder geregistreerde incidenten en kunnen deze incidenten mogelijk onvolledig zijn. Omwille van deze redenen kan geen sluitend oordeel gegeven worden over de incidenten met betrekking tot jong volwassenen die in dit onderzoek zijn onderzocht. Eveneens geeft de data zelf geen uitleg over de gevonden correlaties. Een hoog zelfmoordcijfer kan correleren met de stijging of daling van het aantal *school shootings*. Redenen daarvoor komen niet uit de data zelf naar voren en zullen door middel van interpretatie moeten worden gezocht. Dit is echter altijd onderhevig aan enige vorm van *bias*.

Zoals eerder vermeld, levert niet elke vraag uit dit onderzoek een eenduidig antwoord op. Tijdens het onderzoek is tussen de regels door ook vaak

opgevallen hoe wijdverspreid de problematiek rondom *gun violence* is. Want of er nu naar leeftijdsverdeling, type slachtoffer, of politie gerelateerde incidenten wordt gekeken, de incidenten komen overal in de Verenigde Staten voor. Er zijn natuurlijk tendensen in de dataset, maar het lijkt alsof er geen gebied, leeftijdsgroep of geslacht is die aan de problematiek ontkomt. Tot die tijd is het belangrijk om alle nieuwe incidenten nauwkeurig te noteren. Want bovendien: elk incident is er één te veel.

## References

- [1] Wintemute, G. J. (2015). *The epidemiology of firearm violence in the twenty-first century United States*. Annual review of public health, 36, 5-19.
- [2] Beardslee, J., Mulvey, E., Schubert, C., Allison, P., Infante, A., Pardini, D. (2018). *Gun-and Non-Gun-Related Violence Exposure and Risk for Subsequent Gun Carrying Among Male Juvenile Offenders..* Journal of the American Academy of Child Adolescent Psychiatry, 57(4), 274-279.