

ATTENTIE! BUURTPREVENTIE

De invloed van sociale cohesie
en overlast op de ervaren
veiligheid in de buurt



door:
Nanouk Zielman
S442644
n.zielman@student.rug.nl

Begeleider:
Jaap Nieuwenhuis

Bachelorwerkstuk 2024

Abstract

In tijden van toenemend individualisme en criminaliteit is het belangrijk om onderzoek te blijven doen naar de sociale cohesie in buurten. Onderlinge verbondenheid tussen burens kan invloed hebben op de veiligheid die iemand ervaart in zijn of haar buurt. Sociale cohesie kan ervoor zorgen dat overlast in de buurt afneemt door bijvoorbeeld sociale controle, en daardoor neemt de ervaren veiligheid in de buurt weer toe. De invloed van sociale cohesie en het mediërende effect van overlast zijn onderzocht door het uitvoeren van lineaire en binaire logistische regressie met de dataset 'Buurtbeleving' van het LISS-panel. In het onderzoeksmodel werden, naast sociale cohesie en overlast, ook de factoren geslacht, leeftijd en percentage burens met buitenlandse komaf meegenomen. De data is verzameld door een online enquête onder leden van het panel van 18 jaar en ouder. Uit de resultaten is gebleken dat er inderdaad invloed is van sociale cohesie op ervaren veiligheid en ook dat er een mediërend effect van overlast is. Vrouwen gaven vaker aan dan mannen zich onveilig te voelen en een hoger percentage burens met buitenlandse komaf in de buurt kan zorgen voor minder ervaren veiligheid. Dit effect wordt echter kleiner wanneer sociale cohesie ook wordt meegenomen. Een implicatie van het onderzoek was de oververtegenwoordiging van 60-75 jarigen in de dataset. Hierdoor kon niet goed gekeken worden naar de ervaren veiligheid van hele jonge respondenten of juist senioren. Voor meer duidelijkheid over het onderwerp van ervaren veiligheid in de buurt zou het interessant zijn specifieke groepen te onderzoeken zoals jonge of juist oudere respondenten, buurten die etnisch zeer divers zijn, of de invloed van politiek op polarisatie en sociale cohesie.

Inleiding

Het sociale netwerk van mensen in de tijden van nu is erg veranderd. Zo zijn steeds minder mensen lid van een kerk of vakbond, groeit het aantal contacten via internet door digitalisering en hebben mensen vaker contact met vrienden en kennissen dan burenen. Deze veranderingen kunnen ertoe leiden dat bepaalde mensen zich niet meer thuis voelen in groepen en geen aansluiting vinden; zij kunnen sociaal geïsoleerd raken (Centraal Bureau voor de Statistiek, 2017). Toenemend individualisme zorgt ervoor dat mensen hun burenen steeds minder goed kennen (Forrest & Kearns, 2001). Dit terwijl de sociale cohesie in een buurt wel belangrijk is; burenen delen dan dezelfde normen en waarden, en er is een hoog vertrouwen in de ander (Sampson et al., 1997). Wanneer de sociale cohesie in een buurt hoog is, willen mensen minder snel verhuizen omdat ze zich veiliger voelen. Sociale cohesie vergroot de solidariteit tussen burenen en omdat ze zich meer betrokken voelen met elkaar, draagt sociale cohesie bij aan de leefbaarheid van de buurt door het voorkomen van problemen (Bolt & Torrance, 2005).

Ook is er een toename te zien in criminaliteit; sinds 2021 zijn er steeds meer personen die slachtoffer zijn geweest van bijvoorbeeld geweld of inbraak. Steeds meer mensen voelen zich onveilig op straat (Ten Doeschot & Huijgen, 2024). Uit onderzoek van het CBS (2024) is gebleken dat in de provincie Groningen de grootste stijging was van het aantal slachtoffers. Met een toename van criminaliteit voelen mensen zich steeds vaker onveilig; zo kunnen vrouwen zich bijvoorbeeld steeds vaker onveilig voelen door de toename van femicide, waarvan dit jaar nog een voorbeeld was dichtbij huis: een 17-jarige vrouw werd vermoedelijk doodgestoken door de man met wie ze een relationele band had (NOS, 2024). Dit maakt het belangrijk om onderzoek te doen naar factoren die invloed hebben op hoe veilig personen zich in hun buurt voelen. Het specifieke model van dit werkstuk is in deze tijden relevant omdat er weer een stijging is in het aantal mensen met onveiligheidsgevoelens na een heel lange daling tussen 2004 en 2021 (Centraal Bureau voor de Statistiek, 2022). Ook ervaren mensen, vooral in sterk verstedelijkte gebieden, overlast door bijvoorbeeld geluid of jongeren (Nederlands Jeugdinstituut, 2023). Het is daarom goed om onderzoek naar onveiligheidsgevoelens opnieuw uit te voeren en zo te kijken hoe het er nu voor staat in buurten in Nederland en ook vooral of individualisering en de toename in criminaliteit en overlast ook gedeeltelijk kunnen verklaren waarom er een stijging is in onveiligheidsgevoelens.

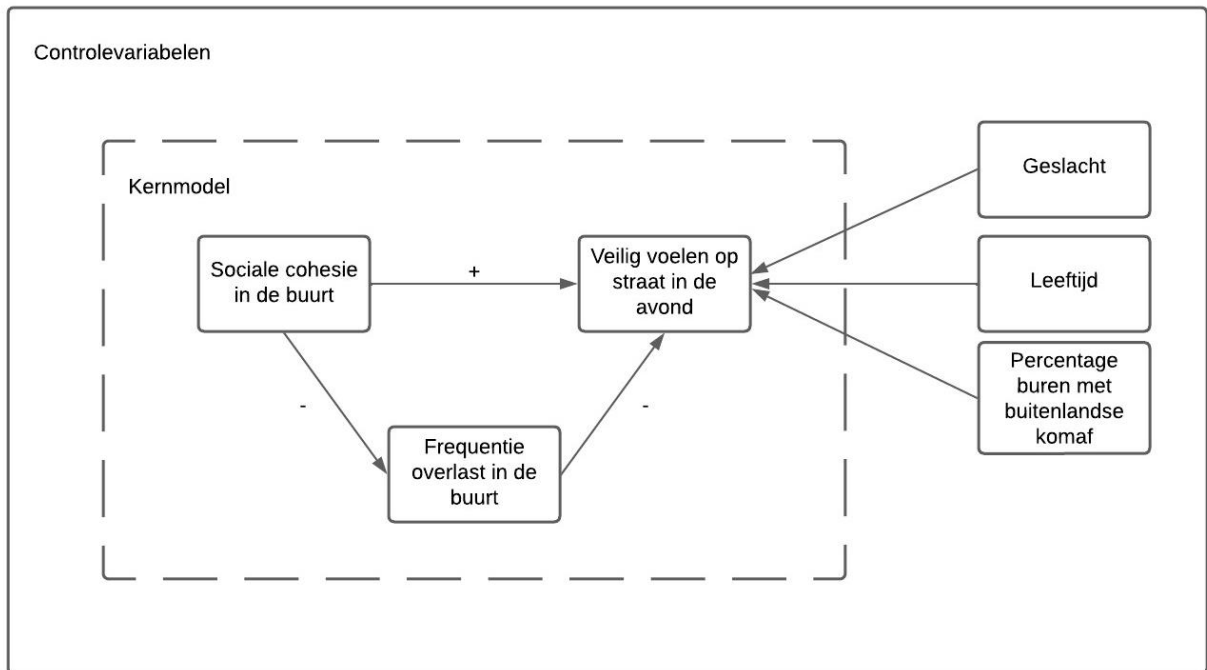
In dit onderzoek wordt dus gekeken naar de invloed van sociale cohesie op de ervaren veiligheid in de buurt en ook wat de invloed van overlast is; hogere sociale cohesie zou kunnen zorgen voor minder overlast, waardoor mensen zich veiliger voelen. Er is dan sprake van een mediërend effect. Verder controleren we voor geslacht, leeftijd en het percentage burenen met buitenlandse komaf. Bij het model hoort de onderzoeksvraag: ‘Wat is de invloed van de ervaren sociale cohesie in de buurt op iemands gevoel van veiligheid in de buurt? Wat is de rol van de ervaren overlast in de buurt?’

Wonen is een primaire levensbehoefte; het is daarom relevant om dit onderzoek uit te voeren en er zo voor te zorgen dat we meer kennis vergaren over hoe mensen kunnen wonen in een veilige buurt. Om de invloeden van de factoren te begrijpen, is gebruik gemaakt van het *fear-decline model*, de *broken window theory* en ook de *busy streets theory* is meegenomen. Verder is een dataset van het LISS panel gebruikt om ondersteuning te vinden voor de theorie. Dit alles geeft ons dan nieuwe inzichten in het vraagstuk van veiligheid in de buurt, wat wetenschappelijk relevant is omdat veel literatuur gaat over onderzoeken in de Verenigde Staten. Over de leefbaarheid van buurten in Nederland wordt ook onderzoek gedaan, maar vaak niet met de diepte waarmee dit onderzoek kijkt naar het verband tussen sociale cohesie en veiligheid in de buurt en wat de invloed op dit verband is met overlast. Dit wordt in specifieke casussen bekeken, maar dit onderzoek is ingegaan op het bredere plaatje van de Nederlandse maatschappij.

Theoretisch kader

Dit onderzoek kijkt naar de invloed van sociale cohesie in de buurt op het gevoel van veiligheid dat iemand ervaart en wat de rol van overlast speelt bij deze ervaring. Wanneer sociale cohesie in een buurt hoger is, kan er minder overlast zijn waardoor het gevoel van veiligheid weer beter wordt.

In model 1 is het onderzoeksmodel schematisch weergegeven.



Model 1: onderzoeksmodel met alle invloedsfactoren

In het onderzoeksmodel is te zien dat het onderzoek kijkt naar de invloed van sociale cohesie op de ervaren veiligheid in de buurt. Het concept van sociale cohesie is wat abstract en daarom lastiger om direct te meten. Sociale cohesie gaat over hoe mensen met elkaar verbonden zijn; in welke mate zij zich met elkaar verbonden voelen en hoe zij invulling geven aan deze verbindingen. Zo is er een hogere mate van ervaren sociale cohesie als buren dezelfde waarden delen en een gevolg van een hoge mate van verbondenheid is dat er veel sociale controle in de buurt is (Sampson et al., 1997). Buren delen dezelfde waarden als er sterke sociale cohesie is; bijvoorbeeld het wonen in een nette buurt waarbij men let op geluidsvolume en zwerfafval om andere buren niet tot last te zijn. Door waarden te delen, vertrouwen buren elkaar ook meer. Dan wordt het ook makkelijker om buren te vertrouwen met jouw veiligheid (Hew et al., 2020). Zo kan sociale controle ontstaan. Buren kunnen invulling geven aan de verbondenheid doordat ze bijvoorbeeld op elkaars kinderen passen, een praatje maken tijdens het uitlaten van de hond of buurtfeestjes organiseren. Deze activiteiten kunnen er weer voor zorgen dat de ervaren verbondenheid bij buren onderling sterker wordt. Wanneer de mensen in de buurt constant verhuizen waardoor de samenstelling van de buurtbewoners keer op keer verandert, kan de onderlinge

verbondenheid juist afnemen (Sampson et al., 1997). Bij een hoog gehalte aan sociale cohesie in de buurt, voelen mensen zich ook veiliger in hun buurt. Buurten kunnen opgebouwd worden tot positieve contexten waarin burens veel interactie hebben en een veilige buurt creëren met elkaar. Dit is te zien in de busy streets theory (Rupp et al., 2019). Deze theorie focust op het belang van het creëren van veilige buurten door pro-sociale interactie, waardoor sociale controle plaatsvindt en burens in elkaar investeren. Zo kunnen burens bijvoorbeeld samenwerken aan het doel van een veilige buurt waardoor sociale controle wordt versterkt, maar ook de onderlinge verbondenheid. Deze transformatie ontstaat echter niet spontaan; de ondersteuning van een organisatie of overheidsorgaan helpt om de transformatie op gang te brengen. Samenvattend kunnen we dan het volgende stellen: wanneer de buurt door sterke sociale cohesie een positieve context is, voelen mensen zich veiliger in hun buurt. Hier hoort de volgende hypothese bij: ‘Een hogere sociale cohesie zorgt voor meer ervaren veiligheid in de buurt.’

Wanneer we kijken naar het effect van sociale cohesie op de ervaren veiligheid in de buurt, loopt deze invloed deels via de overlast in de buurt. De overlast die mensen ervaren in de buurt kan door verschillende dingen komen; zo kan er sprake zijn van vandalisme of geluidsoverlast, maar het kan ook gaan over kleinere dingen zoals hondenpoep en zwerfvuil op straat. Sociale cohesie kan invloed hebben op de overlast in een buurt. Wanneer er sterke sociale cohesie is, kan er ook meer sociale controle zijn waardoor de overlast in een buurt afneemt. Omgekeerd geldt het effect ook; uit onderzoek van Sampson et al. (1997) is gebleken dat een lage mate van sociale cohesie kan zorgen voor meer geweld in de buurt. Dit kan bijvoorbeeld komen doordat in wijken met een lage mate van sociale cohesie ook minder sociale controle is en daarmee meer overlast. Een reden dat er meer overlast is door lage sociale controle is bijvoorbeeld dat er minder wederkerigheid is tussen burens; wanneer de ene persoon zich goed gedraagt, zal de andere persoon in de buurt dat ook doen. Het is als een soort verplichting om de wederdienst te bewijzen. Dit is dus een bepaalde soort macht, gebaseerd op sociale controle (Raven, 1999). Het bijzondere aan wederkerigheid is dat het van beide partijen komt. Deze vorm van sociale controle werkt in bepaalde gevallen beter om overlast terug te dringen dan andere oplossingen tegen overlast, zoals het bouwen van hekken en het ophangen van camera's. Dat zijn ook goed werkende maatregelen, maar het maakt de inwoners van een buurt ook meer gericht zijn op zichzelf en minder sociale interacties aangaan. Hierdoor vertrouwen zij hun burens alsnog niet en is er dus ook geen sociale cohesie (Hew et al., 2020).

Vervolgens is er weer invloed van overlast op het gevoel van veiligheid in de buurt. In onderzoek van Hew et al. (2020) wordt het fear-decline model besproken. In dit model wordt gesteld dat de angst voor criminaliteit in een buurt kan zorgen voor slappe verbindingen van burens onderling omdat ze elkaar niet vertrouwen. Wanneer er overlast is in de buurt, kan dit ook toenemen. In buurten waarin bijvoorbeeld veel zwerfvuil ligt, graffiti gespoten is en waar het er over het algemeen niet verzorgd uitziet, wordt vaker asociaal gedrag vertoond. Dit is het principe van de broken window theory (Keizer et al., 2008). Zo kunnen buurtgenoten, maar ook buitenstaanders, overlast veroorzaken

in een buurt. De overlast kan er weer voor zorgen dat buurtbewoners zich niet veilig voelen en zelfs uit de buurt verhuizen (Hew et al., 2020). Wel is het belangrijk om mee te nemen dat overlast door bijvoorbeeld geweld niet altijd direct effect heeft op hoe onveilig personen hun buurt ervaren. Wanneer zij hun dagelijkse dingen, zoals het huishouden doen, slapen en leven met familie, kunnen doen zonder gestoord te worden door de overlast, blijft voor deze personen hun situatie redelijk stabiel (Forrest & Kearns, 2001).

Samenvattend wordt de ervaren veiligheid van personen in hun buurt dus beïnvloed door sociale cohesie. Sociale cohesie heeft op zijn plaats weer invloed op de frequentie van overlast, wat deels zorgt voor een bepaald niveau van ervaren veiligheid in de buurt. Hieruit komt de volgende hypothese tot stand: ‘Een hogere sociale cohesie kan ervoor zorgen dat overlast daalt waardoor iemand zich veiliger voelt in zijn of haar buurt.’

Bij dit onderzoek is het interessant om te controleren voor verschillende variabelen. Zo zal gekeken worden naar de invloed van geslacht, leeftijd en percentage buren met buitenlandse komaf.

Het meenemen van geslacht in dit onderzoek is belangrijk omdat mannen en vrouwen veiligheid wellicht anders ervaren. Zo zouden vrouwen zich vaker onveilig kunnen voelen op straat omdat zij zich meer kwetsbaar voelen als ze alleen over straat lopen dan mannen; ze kunnen zichzelf minder goed verdedigen omdat ze over het algemeen minder sterk zijn dan mannen. Vrouwen melden veel vaker een bedreiging voor hun persoonlijke veiligheid dan mannen. De kennis dat er dreiging kan zijn van geweld, is voor een vrouw vaak al genoeg om zich onveilig te voelen, zonder dat zij daadwerkelijk bloedgesteld is aan het geweld (Hoffman et al., 2017). Ook zijn vrouwen vaker bang voor geweld omdat een overval kan leiden tot seksueel misbruik, dit fenomeen staat bekend als de *shadow of sexual assault theory* (Hoffman et al., 2017). Het is daarom goed om te kijken of er een verschil is in ervaren veiligheid in de buurt tussen mannen en vrouwen.

Ook het meenemen van leeftijd is belangrijk in dit onderzoek, zo kunnen bejaarde personen bijvoorbeeld vaak kwetsbaar zijn waardoor ze het slachtoffer worden van overlast op straat. Oude personen zijn minder goed ter been en zijn daardoor minder mobiel. Dit zorgt ervoor dat zij zichzelf moeilijk kunnen verdedigen bij overlast. De *vulnerability hypothesis* (Haynes & Rader, 2015) laat zien dat personen die zichzelf niet goed kunnen beschermen omdat ze niet capabel zijn voor zelfverdediging, zich vaker angstig voelen op straat. Maar ook hele jonge personen zouden angstiger kunnen zijn op straat; zij zouden fysiek in de minderheid kunnen raken bij overlast waardoor zij kwetsbaarder zijn. Dit maakt dat het relevant is om te controleren voor leeftijd in dit onderzoek.

Als laatste is het belangrijk om het percentage van buren met buitenlandse komaf mee te nemen. Mensen vinden het fijn om zich te omringen met personen die op hen lijken; niet alleen qua uiterlijk, maar ook qua normen, waarden en cultuur. In een multiculturele wijk kunnen verschillen in meningen ontstaan door cultuurverschillen, wat ervoor kan zorgen dat er minder sociale cohesie is (Sampson et al., 1997). Personen delen minder waarden en vertrouwen elkaar wellicht minder. Dit

wordt ondersteund door de *constrict hypothesis* en de *sub-cultural hypothesis* (Glas et al., 2021). Wanneer er een grote out-group is in de buurt, dus de groep waartoe de respondent niet behoort, zijn personen van de in-group of meerderheid vaak ook angstiger voor overlast en criminaliteit. Dit maakt het relevant om te controleren voor het percentage burens met buitenlandse komaf in dit onderzoek.

Methoden

Dataverzameling

Dit onderzoek maakt gebruik van een kwantitatief onderzoeksmodel en gebruikt daarbij een dataset afkomstig van het LISS panel. De factoren die worden gebruikt in het kernmodel (sociale cohesie, overlast en ervaren veiligheid in de avond) zijn opgenomen in een dataset over buurtbeleving. De studie over buurtbeleving heeft als doel om de individuele buurtbeleving te onderzoeken en te kijken hoe buurtbeleving gevormd wordt (LISS Panel, 2021). Om het onderzoek te doen, is een random selectie van panelleden gemaakt. Deze panelleden zijn allemaal 18 jaar en ouder. In totaal deden 2663 respondenten mee aan het onderzoek over buurtbeleving, waarvan 2646 een complete vragenlijst ingevuld hebben. De afname van deze vragenlijst was in juli 2020.

Voor de factoren die dit onderzoek gebruikt, waren uiteindelijk 2644 respondenten die alle vragen voor de desbetreffende factoren volledig hadden ingevuld. Dit is ook het aantal respondenten in de steekproef die voor het eindmodel van het onderzoek is gebruikt. Er is dus een lage non-respons, wat betekent dat de data generaliseerbaar is naar de populatie.

De factoren geslacht en leeftijd komen uit het databestand van achtergrond variabelen van het LISS panel. Participanten moeten bij het starten van een vragenlijst de informatie over hun achtergrond invullen en kunnen elke maand aanpassingen doen als dat van toepassing is (LISS Panel, 2021). De factor percentage burens met buitenlandse komaf komt uit de dataset van buurtbeleving.

De data over buurtbeleving is onderdeel van een *single wave study*, wat inhoudt dat het onderzoek niet meermaals is uitgevoerd nadat een periode van bepaalde tijd is verstreken. Ondanks dat het onderzoek niet meermaals is uitgevoerd, is de data nog steeds waardevol omdat de perceptie van participanten over hun buurt veel zegt en een goede bron is voor onderzoek en beleidsvorming. Op basis van de resultaten kan gekeken worden hoe het ervoor staat in het algemeen in Nederland met buurtbeleving en waar aandachtspunten liggen om te zorgen dat buurtbeleving verbetert. De vragenlijsten zijn bij elke respondent op dezelfde manier afgenomen en de vragenlijsten zijn opgebouwd uit heldere formuleringen van vragen over buurtbeleving.

Meetinstrumenten

Ervaren veiligheid is de afhankelijke factor in het kernmodel en het is een ordinale factor waarbij respondenten konden aangeven of ze zich veilig voelen in hun buurt als ze 's avonds alleen op straat lopen. Respondenten gaven dit aan op een schaal van vier waarbij 1 betekent dat ze zich heel onveilig voelen en 4 dat ze zich heel veilig voelen. Bij deze stelling gaven de meeste respondenten aan dat ze zich erg veilig voelen op straat. Respondenten hebben hoogstwaarschijnlijk in eerlijkheid kunnen antwoorden in plaats van sociaal wenselijke antwoorden te geven, aangezien de data is verzameld in

een online vragenlijst waarin de respondenten anoniem konden zijn (LISS panel, 2021). Dit bevordert de eerlijkheid bij sensitieve onderwerpen (Krumpal, 2023) zoals ervaren veiligheid.

Sociale cohesie is in het kernmodel de onafhankelijke factor en het is een schaalverdeling (Cronbach's $\alpha = 0.831$) van drie verschillende factoren. De schaalverdeling is gebaseerd op ideeën van de respondenten over of mensen in de buurt elkaar helpen, of mensen uit de buurt te vertrouwen zijn en of mensen in de buurt dezelfde waarden delen. Dit is gemeten doordat respondenten konden aangeven op schaal van één tot vijf hoe erg ze het eens waren met de stelling. Ze konden aangeven met 1 dat ze het helemaal niet mee eens waren en met vijf 5 dat ze het helemaal mee eens waren. De schaal is gevormd door de items bij elkaar op te tellen en te delen door drie. Waar op gelet moet worden bij de conclusies over sociale cohesie is dat mensen wellicht een vertekend beeld hebben van de sociale cohesie in hun buurt. Het is niet vanzelfsprekend dat personen zomaar nieuwe contacten maken in hun buurt, de bewoners die het langst wonen in de buurt zijn vaak sterk verbonden met elkaar, maar weer minder met nieuwkomers. Ook kan er sterke cohesie zijn onder personen die in nieuwbouwhuizen wonen en ze hebben dan weer weinig contact met personen in sociale huurwoningen. Dit concept waarbij verschillende bevolkingsgroepen wel in dicht bij elkaar in de buurt zijn, maar elkaar nauwelijks raken heet het proces van *social tectonics* (Weltevrede et al., 2018) en is dus een concept wat goed is om in ons achterhoofd te houden tijdens het onderzoek.

Overlast is de derde factor die bij het kernmodel hoort en deze dient als de mediator. Ook deze factor is een schaalverdeling (Cronbach's $\alpha = 0.822$), gebaseerd op 11 stellingen over overlast in de buurt. Respondenten gaven aan in deze stellingen of ze overlast ervoeren in hun buurt van afval naast de prullenbak, kapotte fietsen, hinder veroorzaakt door jongeren, hinder veroorzaakt door ouderen, geluidsoverlast, ruzies of vechtpartijen, de geur van marihuana, mensen die alcohol drinken op straat, mensen die drugs gebruiken, vandalisme en slecht onderhouden voortuinen. Hierbij werd een schaal van vier gebruikt waarbij 1 betekent dat het nooit voorkomt en 4 dat het heel vaak voorkomt. De Schaalverdeling is gemaakt door de variabelen bij elkaar op te tellen en te delen door 11. In de dataset geven mensen vaak aan dat overlast in hun buurt niet vaak voorkomt.

Geslacht is een nominale factor, zowel mannen als vrouwen worden goed gerepresenteerd in het onderzoek. In het onderzoek gaven mensen aan of ze een man zijn (in de data aangegeven met het getal 1) of een vrouw (in de data aangegeven met het getal 0). Doordat beide groepen goed gerepresenteerd zijn, maakt dit dat de data goed generaliseerbaar is naar de populatie. Geslacht dient in het onderzoek als controlevariabele.

Leeftijd is een ratio factor, respondenten konden hun leeftijd invullen in de vragenlijst. Voor alle leeftijden zijn respondenten. Dit is goed om mee te nemen wanneer we kijken naar de resultaten. Leeftijd dient in het onderzoek als controlevariabele.

Voor het percentage burens met buitenlandse komaf werden respondenten gevraagd om een inschatting te maken van het percentage in hun buurt. Ze konden dit aangeven in percentages met stappen van 10%. Hierbij gaf 0% aan dat in de buurt geen burens zijn met buitenlandse komaf en 100%

dat de buurt bestaat uit alleen bewoners met buitenlandse komaf. Uit de data kunnen we zien dat het lijkt dat de meeste respondenten in relatief homogene buurten te wonen. Dit is goed om te bedenken wanneer het onderzoek de conclusies bespreekt. Percentage burens met buitenlandse komaf dient in het onderzoek als controlevariabele.

In bijlage 1 zijn alle analyses van de factoren en de bewerkingen uitgebreid toegelicht.

Analyse

Om de onderzoeksvraag te kunnen beantwoorden is gekozen voor een kwantitatieve mediatie analyse. Het onderzoek kijkt naar de invloed van sociale cohesie op ervaren veiligheid in de buurt en wat het mediërende effect van overlast is op dit verband. Model 1 is het lege model; hierin zijn alleen de controlevariabelen geslacht, leeftijd en percentage burens met buitenlandse komaf terug te vinden en de afhankelijke variabele ervaren veiligheid in de buurt. In model 2 wordt de variabele sociale cohesie hieraan toegevoegd. Model 3 laat zien of er wel echt een verband is tussen sociale cohesie en overlast, hierbij is overlast de afhankelijke variabele. In het laatste model is ervaren veiligheid de afhankelijke variabele en zijn sociale cohesie en overlast toegevoegd -tezamen met de controlevariabelen- om het mediërende effect weer te geven.

Om te controleren of de resultaten robuust zijn, is er ook een ordinale logistische regressie en binaire logistische regressie gedaan. Dit is nodig omdat de assumptie van normaliteit bij de lineaire regressie is geschonden. De daarop volgende analyse is ordinale logistische regressie. Daarbij hoort de assumptie van de *test of parallel lines*. Omdat deze ook geschonden is, volgt een binaire logistische regressie. Als uit de resultaten van deze logistische regressies blijkt dat we hetzelfde concluderen dan bij lineaire regressie, zijn de resultaten dus robuust omdat de geschonden assumpties de conclusies niet beïnvloeden. De regressieanalyses zijn uitgevoerd met het programma SPSS. Uit de resultaten van de data zijn conclusies getrokken over de hypothesen die eerder genoemd werden in het onderzoek.

Resultaten

Beschrijvende statistieken

Voor het onderzoek is het goed om te kijken naar de verbanden tussen de variabelen. Hiervoor kijken we naar de verdeling van de variabelen en de samenhang tussen de variabelen in termen van correlatie en Cramers V. In tabel 1 zijn de univariatie statistieken te zien voor alle variabelen in het onderzoek.

Tabel 1: beschrijving van de in de analyse opgenomen variabelen: gemiddelde, standaarddeviatie, minimum- en maximum waarden en totaal aantal respondenten (N = 2644)

	Gemiddelde	Standaarddeviatie	Minimum	Mediaan	Maximum
Sociale cohesie	3.567	0.815	1.000	3.667	5.000
Overlast	1.563	0.401	1.000	1.546	4.000
Ervaren veiligheid	3.550	0.678	1.000	4.000	4.000
Geslacht	1.530	0.499	1.000	1.000	2.000
Leeftijd	55.430	17.242	18.000	58.000	95.000
% buren buitenlandse komaf	2.190	1.879	0.000	2.000	10.000

Uit de tabel zijn conclusies te trekken over de verdelingen van de variabelen. Een aantal variabelen vallen op, zo is het gemiddelde van overlast ligt relatief laag; het tweede kwartiel ligt zelfs nog onder de tweede antwoordmogelijkheid. Dit betekent dat respondenten uit het onderzoek relatief weinig overlast ervaren. Ook is te zien dat het gemiddelde van leeftijd relatief hoog is; ook aan de kwartielen is te zien dat de leeftijd van de meeste respondenten wat hoger ligt. De data is dus wat rechtsscheef. Wel zijn in alle leeftijdsgroepen respondenten. Als laatste valt op dat het gemiddelde van percentage buren met buitenlandse komaf juist weer erg laag ligt. Ook is aan de kwartielen te zien dat de data linksscheef is. Dit betekent dat de meeste respondenten hebben aangegeven in buurten te wonen waar weinig mensen wonen met buitenlandse komaf.

Nu we weten hoe de variabelen eruit zien, kunnen we kijken naar de samenhang. In tabel 2 zijn alle maten van samenhang in een tabel samengevoegd.

Tabel 2: Correlaties, Cramer's V waarden en ANOVA r als maten voor samenhang voor variabelen onderling

	Sociale cohesie (continu)	Overlast (continu)	Ervaren veiligheid (categorisch)	Geslacht (categorisch)	Leeftijd (continu)	Percentage buren met buitenlandse komaf (continu)
Sociale cohesie	1	-0.410**	0.345b**	0.000 ^b	0.096**	-0.337**
Overlast		1	0.390b**	0.000 ^b	-0.172**	0.504**

Ervaren veiligheid	1	0.228 ^{***}	0.063 ^{b*}	0.349 ^{**}
Geslacht		1	0,105 ^{b**}	0.090 ^{b*}
Leeftijd			1	-0.223 ^{**}
Percentage buren met buitenlandse komaf				1

^aCramer's *V* waarde in plaats van correlatie

^b*r* uit ANOVA in plaats van correlatie

* $p < 0.05$

** $p < 0.001$

De tabel laat alle correlaties en associatiematen (zonder de missende data) zien. Er is te zien dat er een redelijk sterk verband is tussen sociale cohesie en ervaren veiligheid ($r = 0.345$; $p < 0.001$). Wanneer mensen meer sociale cohesie ervaren, ervaren ze ook meer veiligheid op straat.

Tussen sociale cohesie en overlast is ook een redelijk sterk verband, deze is negatief ($r = -0.410$; $p < 0.001$). Dit houdt in dat respondenten die meer sociale cohesie ervaren, vaak minder overlast ervaren.

Ook zijn er nog een aantal opvallende correlaties en associatiematen te zien. Tussen sociale cohesie en percentage buren met buitenlandse komaf is een redelijk sterk negatief verband ($r = -0.337$; $p < 0.001$). Dit houdt in dat respondenten die meer sociale cohesie ervaren, vaak aangaven te wonen in buurten waar het percentage buren met buitenlandse komaf lager is. Wel moet meegenomen worden dat respondenten over het algemeen in buurten woonden die relatief homogeen zijn. Ook valt op dat er een matig, positief verband is tussen ervaren veiligheid en geslacht ($r = 0.0228$; $p < 0.001$). Dit betekent dat vrouwen vaker aangeven zich veilig te voelen dan mannen, maar de samenhang is dus niet heel sterk.

Modevaluatie

Dit onderzoek maakt gebruik van lineaire regressie, in model 1 is het lege model te zien met alleen de controlevariabelen. Model 2 laat het hoofdeffect zien; het effect van de sociale cohesie op de ervaren veiligheid, gecontroleerd voor geslacht, leeftijd en percentage buren met buitenlandse komaf. Model 3 laat het verband zien tussen sociale cohesie en overlast, hierin is de factor overlast de afhankelijke variabele. Zo is te zien of sociale cohesie ook daadwerkelijk invloed heeft op overlast. Het laatste model laat het mediërende effect zien, hierbij is ervaren veiligheid de afhankelijke variabele en zijn sociale cohesie en overlast opgenomen voor het hoofdmodel. Ook zijn de controlevariabelen weer opgenomen. Passend bij de modellen zijn verschillende hypothesen opgesteld. De eerste hypothese gaat over het hoofdeffect en is als volgt: 'Een hogere sociale cohesie zorgt voor meer ervaren veiligheid in de buurt.' De tweede hypothese gaat over het mediërende effect van overlast en is als volgt geformuleerd: 'Een hogere sociale cohesie kan ervoor zorgen dat overlast daalt waardoor iemand zich veiliger voelt in zijn of haar buurt'.

Tabel 3: parameterschattingen (n=2644)

	Model 1 ^a		Model 2 ^a		Model 3 ^b		Model 4 ^a	
	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>
<i>Constante</i>	4.312 (0.062)	<0.001	3.486 (0.083)	<0.001	1.936 (0.046)	0.046	4.219 (0.105)	<0.001
<i>Geslacht</i>	-0.269 (0.024)	<0.001	-0.278 (0.024)	<0.001	-0.009 (0.001)	0.013	-0.281 (0.023)	<0.001
<i>Leeftijd</i>	-0.001 (0.001)	0.051	-0.02 (0.001)	<0.017	-0.001 (0.001)	<0.001	-0.002 (0.001)	0.001
<i>Percentage buren met buitenlandse komaf</i>	-0.123 (0.007)	<0.001	-0.091 (0.007)	<0.001	0.086 (0.007)	0.004	-0.059 (0.007)	<0.001
<i>Sociale cohesie</i>			0.220 (0.015)	<0.001	-0.133 (0.015)	0.008	0.169 (0.016)	<0.001
<i>Overlast</i>							-0.379 (0.035)	<0.001
R²	0.158		0.219		0.0323		0.253	
<i>(Partiële) F</i>	164.743	<0.001	185.431	<0.001	314.349	<0.001	179.103	<0.001

^a Afhankelijke variabele: ervaren veiligheid; ^b Afhankelijke variabele: overlast

De eerste hypothese ('Een hogere sociale cohesie zorgt voor meer ervaren veiligheid in de buurt') toetsen we vooral met model 2. We zien dat er een positieve helling is voor sociale cohesie ($\beta = 0.220$; $p < 0.001$); dus de ervaren veiligheid is hoger als mensen meer sociale cohesie ervaren. In model 2 zien we ook dat het effect van het percentage buren met buitenlandse komaf ($\beta = -0.091$; $p < 0.001$) wat kleiner wordt vergeleken met model 1; het houdt in dat stijging in percentage buren met buitenlandse komaf zorgt voor een daling in ervaren veiligheid, maar met het toevoegen van sociale cohesie wordt dit effect wel minder sterk. Het effect van leeftijd is verwaarloosbaar omdat er maar een hele kleine helling is. Voor geslacht is een negatieve helling, dit houdt in dat vrouwen zich vaker onveilig voelen op straat dan mannen.

Aan de hand van model 3 is te zien of er überhaupt een verband is tussen sociale cohesie en overlast. Model 3 laat zien dat er een significant verband is ($\beta = -0.133$; $p = 0.008$); wanneer mensen meer sociale cohesie ervaren, ervaren ze minder overlast.

Ook kijken we naar de hypothese over het mediërende effect ('Een hogere sociale cohesie kan ervoor zorgen dat overlast daalt waardoor iemand zich veiliger voelt in zijn of haar buurt'). Als we kijken naar model 4, zien we dat het effect van sociale cohesie nog steeds positief is, maar wel kleiner ten opzichte van model 2. Er is een negatieve helling voor overlast; dit houdt in dat een stijging in de ervaren overlast zorgt voor een daling in de ervaren veiligheid. Deze helling is redelijk groot en heeft dus een redelijk groot effect op de ervaren veiligheid. Ook zien we aan de verklaarde variantie dat

model 4 de meeste variantie verklaard ($R^2 = 0.253$). Dit houdt in dat de toevoeging van overlast waardevol is voor het begrijpen van het effect van sociale cohesie op ervaren veiligheid.

Assumpties

Bij lineaire regressie gaat het model uit van verschillende assumpties. Bij het controleren van deze assumpties was te zien dat er een aselecte steekproef heeft plaatsgevonden en dat er geen sprake was van homoscedasticiteit. Wel was er een schending in de assumptie van normaliteit en ook de residuen waren niet helemaal normaal verdeeld. Dit is dan ook de reden dat er ook ordinale logistische regressie uitgevoerd is en binaire logistische regressie.

Gevoeligheidsanalyses

Omdat de assumptie van normaliteit is geschonden, is er voor dit onderzoek ook ordinale logistische regressie en binaire logistische regressie gedaan. De resultaten van de ordinale regressie bevestigden de resultaten, maar uit de test of parallel lines bleek dat de assumptie van gelijke hellingen bij elke antwoordcategorie geschonden was. Daarom is er ook binaire logistische regressie gedaan. Hieruit is gebleken dat de resultaten van de lineaire regressie alsnog bevestigd zijn. Uit de resultaten van de modellen van de binaire logistische regressie konden namelijk dezelfde conclusies getrokken als bij de lineaire regressie. Deze resultaten en uitwerkingen zijn terug te vinden in bijlage 2. Ondanks dat de assumpties voor lineaire regressie geschonden zijn, zijn de resultaten dus nog steeds betrouwbaar en gecontroleerd, de assumptieschendingen hebben geen invloed op de resultaten en zijn dus “robuust”.

Multicollineariteit en uitbijters

Multicollineariteit treedt op als er een sterk verband is tussen verklarende variabelen. Het is dus niet altijd goed om steeds meer variabelen toe te voegen voor meer verklaarde variantie. Daarom moet er getest worden op multicollineariteit. Uit de *Variation Inflation Factor* (VIF) is gebleken dat de vuistregels niet geschonden zijn en dat er dus geen sprake is van multicollineariteit.

Ook is er gekeken of er sprake was van invloedrijke punten; uitbijters die te veel aan de waarden trekken. Dit is getest door te kijken naar de DFFIT, DFBETA, Cook's distance en de residuen. Een aantal cases kwamen naar voren maar uiteindelijk bleek dat deze punten te weinig invloed hadden op de hellingen om ze te verwijderen uit de dataset. De hellingen bleven namelijk nagenoeg gelijk toen de uitbijters uit de dataset waren gehaald.

Conclusie en discussie

Conclusie

In dit onderzoek is geprobeerd te verklaren wat de invloed is van sociale cohesie op de ervaren veiligheid die iemand ervaart in zijn of haar buurt. Ook is er gekeken wat sociale cohesie doet met overlast en wat deze in zijn plaats voor invloed heeft op de ervaren veiligheid. Uit de resultaten van het onderzoek is gebleken dat een hogere sociale cohesie ook daadwerkelijk leidt tot meer ervaren veiligheid. Wanneer mensen zich dus meer verbonden voelen met hun buren en het gevoel hebben hun buren te kunnen vertrouwen, voelen zij zich ook veiliger in hun buurt. Het onderzoek bevestigt dus de bevindingen van Hew et al. (2020) dat meer sociale cohesie ertoe leidt dat buren elkaar sneller vertrouwen met hun veiligheid.

Ook bleek dat er een mediërend effect was van overlast; als sociale cohesie omhoog ging, was er minder overlast in een buurt waardoor deze ook veiliger voelde voor haar bewoners. Dit kan bijvoorbeeld door sociale controle, personen voelen de verplichting zich goed te gedragen als hun buren dat ook doen (Raven, 1999). Hiermee zijn de hypothesen van dit onderzoek bevestigd. Een belangrijk punt waar wel op gelet moet worden bij overlast is dat respondenten die aangaven overlast te ervaren in hun buurt, wel maar weinig overlast ervoeren. Dit zou kunnen komen doordat personen in Nederland over het algemeen weinig criminaliteit ervaren en veel veiligheid ervaren in hun buurt (Akkermans et al., 2022). Het zou ook kunnen komen doordat personen al lange tijd in hun buurt wonen waardoor hun idee van veiligheid in hun buurt ‘verzacht’, personen raken gewend aan de overlast die in hun buurt plaats vindt (Fowler et al., 2009).

Uit de controlevariabelen kunnen ook conclusies getrokken worden. Zo was de invloed van het percentage buren met buitenlandse komaf klein op de ervaren veiligheid in de buurt. Dit zou kunnen komen doordat in het grootste gedeelte van Nederland de buurten geen hoge percentages bewoners van buitenlandse komaf kennen en de buurten dus erg homogeen zijn (Centraal Bureau van de Statistiek, z.d.). In andere landen zoals de Verenigde Staten zou er dus meer invloed kunnen zijn van deze factor omdat men daar meer variatie kent in etniciteit (US Census Bureau, 2023) in vergelijking met Nederland. Verder is uit het onderzoek gebleken dat vrouwen aangaven zich vaker onveilig te voelen dan mannen. Deze bevinding bevestigt ook het beeld wat de ‘veiligheidsmonitor’ van het CBS schetst; vrouwen voelen zich vaker onveilig dan mannen, zowel in hun buurt als in het algemeen. Vooral jonge vrouwen in de leeftijd van 15-25 jaar en 65-plussers (Akkermans et al., 2024). De invloed van leeftijd was te verwaarlozen in het onderzoek. Dit is opvallend want uit eerdere onderzoeken is gebleken dat leeftijd wel degelijk invloed had op de ervaren veiligheid, zo zouden ouderen bij bepaalde plaatsen en tijden bang zijn voor het risico om slachtoffer te worden van criminaliteit (Mesch, 2000; Rosen et al., 2019).

Discussie

Ondanks dat de resultaten uit het onderzoek bruikbaar zijn, zijn er nieuwe vragen naar boven gekomen tijdens het doen van dit onderzoek. Om deze reden zijn er een aantal aanbevelingen voor nieuw onderzoek. Zo zou het interessant zijn om niet alleen naar de invloed van sociale cohesie en overlast op de ervaren veiligheid in het algemeen in Nederland te kijken, maar ook naar specifieke buurten waar de ervaren veiligheid opvallend laag ligt. Nu zijn de conclusies getrokken op basis van data uit alle buurten, terwijl er verschil kan zijn in buurten onderling. Er bestaan aanzienlijke verschillen tussen buurten en steden in hoe criminaliteit is gespreid (Wittebrood, 2000). Het is daarom goed om bij buurten waar vaak problemen opspelen en waar veiligheid niet altijd gewaarborgd kan worden, specifiek onderzoek te doen naar de buurtkenmerken omdat deze sterke invloed kunnen hebben op bijvoorbeeld de criminaliteit in de buurt (Wittebrood, 2000).

Verder was er in dit onderzoek een tekortkoming bij de factor leeftijd; de leeftijdsgroep van 60- tot en met 75-jarigen was oververtegenwoordigd. Uit de resultaten is gebleken dat er geen invloed was van leeftijd op de ervaren veiligheid, dit spreekt eerder onderzoek tegen. Zo bleek uit onderzoek van Hoffman et al. (2017) dat juist jonge personen en dan vooral vrouwen zich vaker onveilig kunnen voelen. Dit kan bijvoorbeeld komen omdat vrouwen meer angst hebben voor geweld en diefstal omdat deze gebeurtenissen zouden kunnen leiden tot seksueel misbruik bijvoorbeeld. Ook zouden oudere mensen zich onveiliger kunnen voelen omdat ze kwetsbaarder zijn (Rosen et al., 2019). Het is daarom interessant om verder onderzoek te doen naar de ervaringen van specifiek jonge respondenten of oude respondenten om zo meer inzicht te krijgen in wat de rol van sociale cohesie en overlast is bij hun ervaring van veiligheid in hun buurt. Daarnaast is uit eerder onderzoek gebleken dat het goed is om niet alleen onderscheid te maken tussen mannen en vrouwen bij het onderzoeken van ervaren veiligheid, maar ook te kijken naar andere groepen zoals personen die bij de LHBTIQ+ gemeenschap horen omdat ze andere ervaringen hebben op dit gebied (Akkerman et al., 2024). Zo zouden zij hun veiligheid anders kunnen ervaren omdat zij vaker het slachtoffer zijn van geweld (Espino et al., 2024). De kennis dat er geweld in een buurt aanwezig kan zijn, is al genoeg voor iemand om zich onveilig te voelen (Hoffman et al., 2017).

Als laatste zou het in het huidige politieke landschap goed zijn om specifiek te kijken naar de ervaren veiligheid van personen met buitenlandse komaf. Uit de data van dit onderzoek is gebleken dat ze in buurten vaak in de minderheid zijn. In Europa is er een toename van polarisatie en politisering. Door de opkomst van rechts populisme in de politiek wordt er steeds meer nadruk gelegd op de etnische verschillen in een land (Van der Tol & Becker, 2024). Hierdoor kan polarisatie in een buurt toenemen. Een persoon van buitenlandse komaf zou zich sneller onveilig kunnen voelen als gevolg van de combinatie van politieke uitingen en het feit dat de persoon behoort tot de minderheid. Bijvoorbeeld omdat er afname is in sociale cohesie (Ministerie van Volksgezondheid, Welzijn en Sport, 2022).

Daarom zou het goed zijn om deze doelgroep verder onder de loep te nemen en te kijken naar hoe zij de veiligheid ervaren in hun buurt; wellicht verschilt dit van de autochtone inwoner in dezelfde buurt die behoort tot de meerderheid.

Concluderend heeft dit onderzoek dus aangetoond dat sociale cohesie invloed heeft op hoe iemand de veiligheid ervaart in zijn of haar buurt; meer sociale cohesie zorgt ervoor dat iemand zich veiliger voelt. Ook is er invloed van overlast; wanneer er meer sociale cohesie is in een buurt, neemt de overlast af waardoor de ervaren veiligheid weer hoger wordt. Er is gebleken dat vrouwen zich vaker onveilig voelen dan mannen en dat het percentage burens met buitenlandse komaf in de buurt geen grote rol speelde bij de ervaren veiligheid. Omdat de maatschappij constant verandert, is onderzoek doen over de ervaren veiligheid in de buurt nog lang niet overbodig; het is goed om de maatschappelijke veranderingen op de voet te blijven volgen in de sociale wetenschappen.

Literatuurlijst

- Akkermans, M., Derksen, E., Kennis, M., Kloosterman, R., & Moons, E. (2024). 3. *Veiligheidsbeleving*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/longread/rapportages/2024/veiligheidsmonitor-2023/3-veiligheidsbeleving#:~:text=Vrouwen%20voelen%20zich%20vaker%20onveilig,jarigen%20en%20vooral%2065%2Dplussers>
- Akkermans, M., Kloosterman, R., Moons, E., Reep, C., Tummers-van Der Aa, M., & Centraal Bureau voor de Statistiek. (2022, 28 februari). 2. *Leefbaarheid en overlast in woonbuurt*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/longread/rapportages/2022/veiligheidsmonitor-2021/2-leefbaarheid-en-overlast-in-woonbuurt>
- Bolt, G. and Torrance, M.I., (2005). *Stedelijke herstructurering en sociale cohesie*. Utrecht: DGW/NETHUR-partnership-30
- Centraal Bureau voor de Statistiek. (2017). *Worden we individualistischer?* *Centraal Bureau Voor De Statistiek*. <https://www.cbs.nl/nl-nl/nieuws/2017/52/worden-we-individualistischer->
- Centraal Bureau voor de Statistiek. (2022). 3. *Veiligheidsbeleving*. Centraal Bureau Voor De Statistiek. <https://www.cbs.nl/nl-nl/longread/rapportages/2022/veiligheidsmonitor-2021/3-veiligheidsbeleving#:~:text=In%20de%20periode%202005%E2%80%932021,onveiligheidsgevoelens%20met%2034%20procent%20afgenomen>
- Centraal Bureau voor de Statistiek. (z.d.). *Cijfers op de kaart*. Centraal Bureau Voor de Statistiek. https://www.cbs.nl/nl-nl/visualisaties/cijfers-op-de-kaart?location=gem-0828862b6db3ff61c3c5806d498f00a2&subject=H007933_3&year=2023&level=Gemeente
- Espino, E., Jiménez-Díaz, O., Del Rey, R., & Elipe, P. (2024). Outlining Individual and Contextual Factors Related to LGBTQ+ Bullying: A Systematic Review of Two Decades of Research. *Trauma, Violence, & Abuse*, 25(2), 898-917. <https://doi-org.proxy-ub.rug.nl/10.1177/15248380231165724>
- Forrest, R. & Kearns, A. (2001) *Social Cohesion, Social Capital and the Neighbourhood*. *Urban*

Studies, 38(12), 2125-2143. <https://doi.org/10.1080/00420980120087081>

- Fowler, P. J., Tompsett, C. J., Braciszewski, J. M., Jacques-Tiura, A. J., & Baltes, B. B. (2009). Community violence: A meta-analysis on the effect of exposure and mental health outcomes of children and adolescents. *Development and Psychopathology*, 21, 227–259. <https://doi.org/10.1017/S0954579409000145>
- Glas, I., Engbersen, G., & Snel, E. (2019). The street level and beyond: The impact of ethnic diversity on neighborhood cohesion and fear of crime among Dutch natives and nonnatives. *Journal of Urban Affairs*, 41(6), 737–755. <https://doi-org.proxy-ub.rug.nl/10.1080/07352166.2018.1533379>
- Glas, I., Jennissen, R., & Engbersen, G. (2021). Estimating Diversity Effects in the Neighborhood: On the Role of Ethnic Diversity and Out-group Size and their Associations with Neighborhood Cohesion and Fear of Crime. *Social Indicators Research*, 158(1), 1–22. <https://doi-org.proxy-ub.rug.nl/10.1007/s11205-021-02704-9>
- Haynes, S. H., & Rader, N. E. (2015). Concerns About Crime for Self and Others: An Analysis of Individual and Contextual Effects. *Criminal Justice Review*, 40(3), 303–321. <https://doi.org/10.1177/0734016814568011>
- Hew, W., Guan-Gan Goh, G., Boon-Yee Low, A., & Siok-Hwa Lau. (2020). Flight Intentions among Residents in Crime-ridden Neighborhoods: A PLS-SEM Assessment. *Asia-Pacific Social Science Review*, 20(3), 89–101. <https://www.dlsu.edu.ph/wp-content/uploads/pdf/research/journals/apssr/2020-September-vol20-3/8-flight-intentions-among-residents-in-crime-ridden-neighborhoods-a-pls-sem-assessment.pdf>
- Hoffman, E. E., Mair, T. T. M., Hunter, B. A., Prince, D. M., & Tebes, J. K. (2017). Neighborhood sexual violence moderates women’s perceived safety in urban neighborhoods. *Journal of Community Psychology*, 46(1), 79–94. <https://doi-org.proxy-ub.rug.nl/10.1002/jcop.21917>
- Krumpal, I. (2023). Social Desirability Bias and Context in Sensitive Surveys. In *Journal of Applied Psychology* (pp. 6527–6532). Springer Ebooks. https://doi.org/10.1007/978-3-031-17299-1_4086
- LISS Panel. (2021). *Neighbourhood Perceptions*. dataarchive.lissdata.nl. Retrieved March 27, 2024, from <https://www.dataarchive.lissdata.nl/study-units/view/1095>

- LISS Panel. (2022). *background Variables*. dataarchive.lissdata.nl. Retrieved March 27, 2024, from <https://www.dataarchive.lissdata.nl/study-units/view/322>
- Mesch, G. S. (2000). Perceptions of risk, lifestyle activities, and fear of crime. *Deviant behavior*, 21(1), 47-62. <https://doi-org.proxy-ub.rug.nl/10.1080/016396200266379>
- Ministerie van Volksgezondheid, Welzijn en Sport. (2022, 20 december). *Driekwart van de Nederlanders denkt dat polarisatie toeneemt*. Nieuwsbericht | Sociaal en Cultureel Planbureau. <https://www.scp.nl/actueel/nieuws/2022/12/29/ergernis-over-harde-toon-en-extreme-uitingen-in-politieke-en-publieke-debat>
- Nederlands Jeugdinstituut. (2023) *Cijfers over overlast | Nederlands Jeugdinstituut*. <https://www.nji.nl/cijfers/overlast#:~:text=In%20Nederland%20zijn%20in%202022,respectie%20velijk%20137.947%20en%20124.486%20meldingen>.
- NOS. (2024). Winsum neemt afscheid van doodgestoken tiener. *NOS*. <https://nos.nl/artikel/2511140-winsum-neemt-afschied-van-doodgestoken-tiener>
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, 277(5328), 918–924. <https://doi.org/10.1126/science.277.5328.918>
- Raven, B. H. (1999). Kurt Lewin Address: Influence, Power, Religion, and the Mechanisms of Social Control. *Journal of Social Issues*, 55(1), 161–186. <https://doi.org/10.1111/0022-4537.00111>
- Rosen, T., Makaroun, L. K., Conwell, Y., & Betz, M. (2019). Violence in Older Adults: scope, impact, challenges, and Strategies for Prevention. *Health Affairs*, 38(10), 1630–1637. <https://doi.org/10.1377/hlthaff.2019.00577>
- Rupp, L. A., Zimmerman, M. A., Sly, K. W., Reischl, T. M., Thulin, E. J., Wyatt, T. A., & Stock, J. P. (2020). Community-Engaged Neighborhood Revitalization and Empowerment: Busy Streets Theory in Action. *American Journal of Community Psychology*, 65(1/2), 90–106. <https://doi-org.proxy-ub.rug.nl/10.1002/ajcp.12358>
- Ten Doeschot, F., & Huijgen, G. (2024). *Veiligheidsmonitor 2023: toename traditionele*

criminaliteit. Ipsos I&O. <https://www.ioresearch.nl/actueel/veiligheidsmonitor-2023-toename-traditionele-criminaliteit/>

US Census Bureau. (2023). *Some of the Most Diverse Census Tracts in 2020 Were in Anchorage, Alaska and Queens County, New York*. Census.gov.

<https://www.census.gov/library/stories/2024/01/racial-ethnic-diversity-neighborhoods.html>

Van der Tol, M., & Becker, E. (2024). What's ethnicity got to do with it? Religious and racial politics in Europe. *Ethnic And Racial Studies*, 1–20. <https://doi.org/10.1080/01419870.2024.2328318>

Weltevrede, A., Van Den Heerik, A., Helmer, N., & De Boom, J. (2018). *Nieuwe Buren: Een onderzoek naar de veranderende sociale compositie in drie Rotterdamse wijken*. Erasmus Universiteit Rotterdam/ Risbo. https://offgridstudio.nl/images/onderzoeksrapport_KL04.pdf

Wittebrood, K. (2000). *Buurten en geweldscriminaliteit: een multilevel-analyse*.

<https://ugp.rug.nl/MenM/article/view/18599/16075>

Bijlage 1

In deze bijlage wordt gekeken naar de verschillende variabelen die in het onderzoek gebruikt worden. Zo krijgt de lezer inzicht in de verdeling van de variabelen en zijn de bewerkingen van de variabelen ook controleerbaar.

Sociale cohesie

De variabele *sociale cohesie* bestaat uit drie verschillende variabelen. Eerst zijn de frequentieverdelingen en descriptieven van de losse variabelen te zien, vervolgens volgt de data van de schaalverdeling.

*Frequentieverdelingen 'sociale cohesie' losse variabelen.

DATASET ACTIVATE DataSet1.

FREQUENCIES VARIABLES=sr20a008 sr20a009 sr20a010

/NTILES=4

/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN

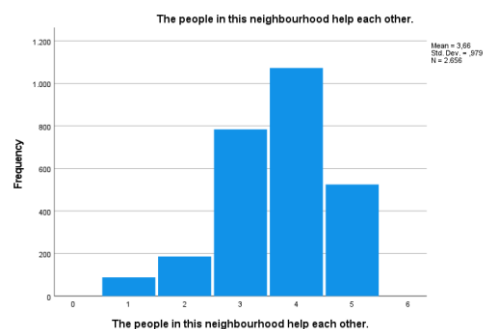
/HISTOGRAM

/ORDER=ANALYSIS.

sr20a008 - De mensen in de buurt helpen elkaar

The people in this neighbourhood help each other.

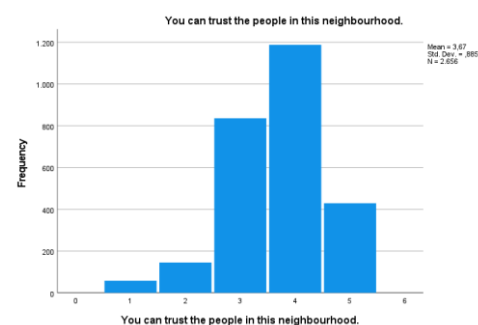
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 completely disagree	88	3,3	3,3	3,3
	2	186	7,0	7,0	10,3
	3	784	29,5	29,5	39,8
	4	1073	40,3	40,4	80,2
	5 completely agree	525	19,7	19,8	100,0
	Total	2656	99,8	100,0	
Missing	System	5	,2		
Total		2661	100,0		



sr20a009 - Je kunt de mensen in deze buurt vertrouwen

You can trust the people in this neighbourhood.

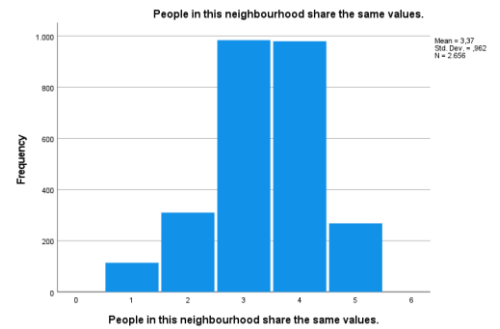
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 completely disagree	58	2,2	2,2	2,2
	2	145	5,4	5,5	7,6
	3	836	31,4	31,5	39,1
	4	1188	44,6	44,7	83,8
	5 completely agree	429	16,1	16,2	100,0
	Total	2656	99,8	100,0	
Missing	System	5	,2		
Total		2661	100,0		



sr20a010 - De mensen in deze buurt delen dezelfde waarden

People in this neighbourhood share the same values.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 completely disagree	114	4,3	4,3	4,3
	2	310	11,6	11,7	16,0
	3	984	37,0	37,0	53,0
	4	980	36,8	36,9	89,9
	5 completely agree	268	10,1	10,1	100,0
	Total	2656	99,8	100,0	
Missing	System	5	,2		
	Total	2661	100,0		



Deze variabelen samen kunnen de schaalverdeling van sociale cohesie vormen. Ze geven samen betekenis aan wat sociale cohesie is. Het samenvoegen was mogelijk, omdat deze variabelen samen een Cronbach's alpha had van 0.831.

*Controleren betrouwbaarheid sociale cohesie.

RELIABILITY

```

/VARIABLES=sr20a008 sr20a009 sr20a010
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE
/SUMMARY=TOTAL.

```

Reliability Statistics

Cronbach's Alpha	N of Items
,831	3

*Schaalverdeling sociale cohesie.

```
COMPUTE SocialeCohesie=(sr20a008 + sr20a009 + sr20a010) / 3.
```

```
EXECUTE.
```

Schaalverdeling sociale cohesie

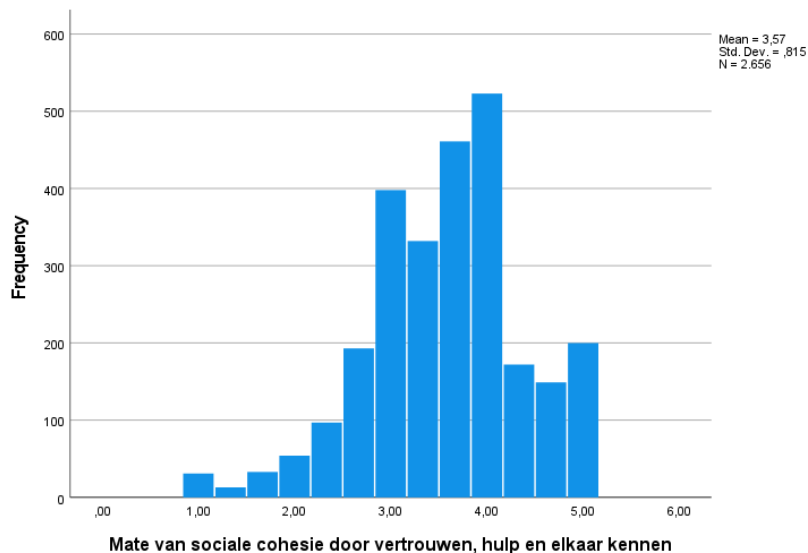
Naam in dataset: 'SocialeCohesie'.

Omschrijving: Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen

Type: ordinaal

GRAPH

```
/HISTOGRAM=SocialeCohesie.
```

Sociale cohesie is gemeten door verschillende stellingen;

- De mensen in deze buurt helpen elkaar
- Je kunt de mensen in deze buurt vertrouwen
- De mensen in deze buurt delen dezelfde waarden

Respondenten konden de vraag beantwoorden door aan te geven op schaal van één tot en met vijf hoe erg ze het eens waren met de stelling; waarbij 1 helemaal niet mee eens is en 5 helemaal mee eens.

SocialeCohesie				
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	31	1,2	1,2
	1,33	13	,5	1,7
	1,67	33	1,2	2,9
	2,00	54	2,0	4,9
	2,33	97	3,6	8,6
	2,67	193	7,3	15,9
	3,00	398	15,0	30,8
	3,33	332	12,5	43,3
	3,67	461	17,3	60,7
	4,00	523	19,7	80,4
	4,33	172	6,5	86,9
	4,67	149	5,6	92,5
	5,00	200	7,5	100,0
	Total	2656	99,8	100,0
Missing	System	5	,2	
Total		2661	100,0	

De verdeling voor sociale cohesie is niet helemaal normaal. Wel liggen de meeste antwoorden in het midden; het antwoord 4 – mee eens is het vaakst gegeven. De verdeling is niet echt scheef en daarom wel goed te gebruiken in dit onderzoek. In elke antwoordcategorie zitten respondenten; daardoor is er redelijke verdeling ontstaan.

Ondanks dat de variabele bestaat uit verschillende categorieën; zal ik de variabele meenemen in het model als een continue variabele.

Overlast

De variabele *overlast* bestaat uit 11 verschillende variabelen. Eerst zijn de frequentieverdelingen en descriptieven van de losse variabelen te zien, vervolgens volgt de data van de schaalverdeling.

De oorspronkelijke variabelen die samen de schaalverdeling vormen van overlast hadden eerst de interpretatie van de waarden als volgt:

- 1 – heel vaak
- 2 – vaak
- 3 – soms
- 4 – nooit

Dit maakte de interpretatie van de variabelen wat lastig. Daarom zijn de variabelen eerst gespiegeld en vervolgens zijn ze samengevoegd tot de schaalverdeling *overlast*.

Hiervoor zijn de volgende syntax commando's voltooid:

*Omgekeerde waarden van alle overlast-variabelen.

```
RECODE sr20a047 sr20a048 sr20a049 sr20a050 sr20a051 sr20a052 sr20a053 sr20a054 sr20a055
sr20a056
```

```
sr20a057 sr20a058 (1=4) (2=3) (3=2) (4=1).
```

```
EXECUTE.
```

*Frequentieverdelingen 'overlast' losse variabelen.

```
FREQUENCIES VARIABLES=sr20a047 sr20a049 sr20a050 sr20a051 sr20a052 sr20a053 sr20a054
sr20a055 sr20a056 sr20a057 sr20a058
/NTILES=4
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
/HISTOGRAM
/ORDER=ANALYSIS.
```

sr20a047 - Afval naast de prullenbak

How often does it occur in your neighbourhood? - Litter or waste next to waste c

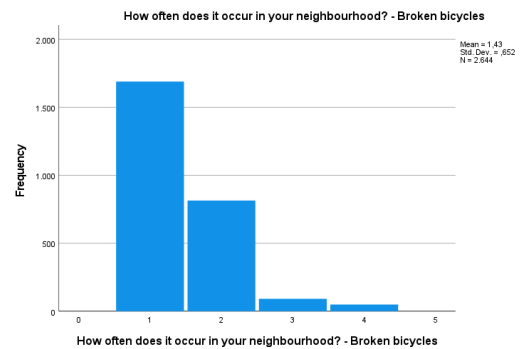
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	829	31,2	31,3	31,3
	soms	1277	48,0	48,3	79,6
	vaak	330	12,4	12,5	92,1
	heel vaak	209	7,9	7,9	100,0
	Total	2645	99,4	100,0	
Missing	System	16	,6		
Total		2661	100,0		



Sr20a049 - Kapotte fietsen

How often does it occur in your neighbourhood? - Broken bicycles

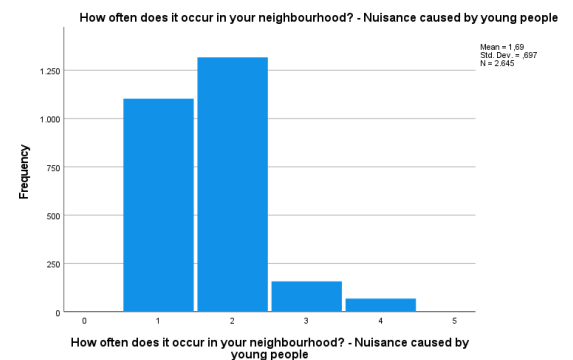
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1690	63,5	63,9	63,9
	soms	814	30,6	30,8	94,7
	vaak	91	3,4	3,4	98,1
	heel vaak	49	1,8	1,9	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a050 - Hinder veroorzaakt door jongeren

How often does it occur in your neighbourhood? - Nuisance caused by young people

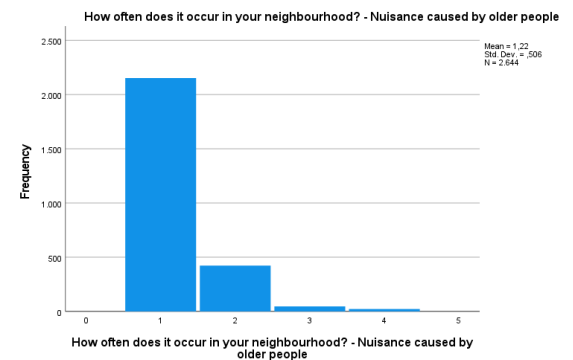
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1103	41,5	41,7	41,7
	soms	1317	49,5	49,8	91,5
	vaak	157	5,9	5,9	97,4
	heel vaak	68	2,6	2,6	100,0
	Total	2645	99,4	100,0	
Missing	System	16	,6		
Total		2661	100,0		



Sr20a051 - Hinder veroorzaakt door ouderen

How often does it occur in your neighbourhood? - Nuisance caused by older people

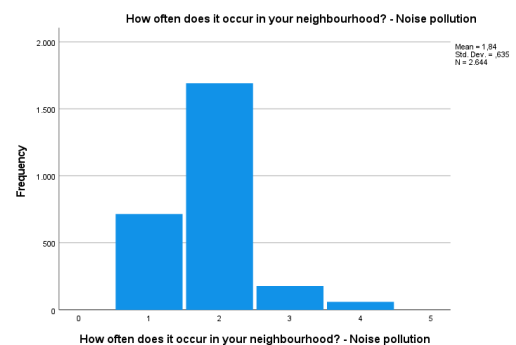
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	2153	80,9	81,4	81,4
	soms	423	15,9	16,0	97,4
	vaak	46	1,7	1,7	99,2
	heel vaak	22	,8	,8	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a052 - Geluidsoverlast

How often does it occur in your neighbourhood? - Noise pollution

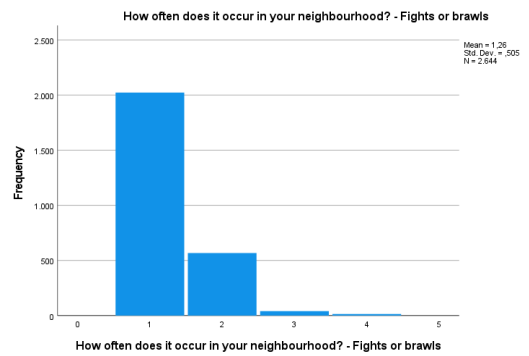
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	715	26,9	27,0	27,0
	soms	1691	63,5	64,0	91,0
	vaak	178	6,7	6,7	97,7
	heel vaak	60	2,3	2,3	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a053 - Ruzies en vechtpartijen

How often does it occur in your neighbourhood? - Fights or brawls

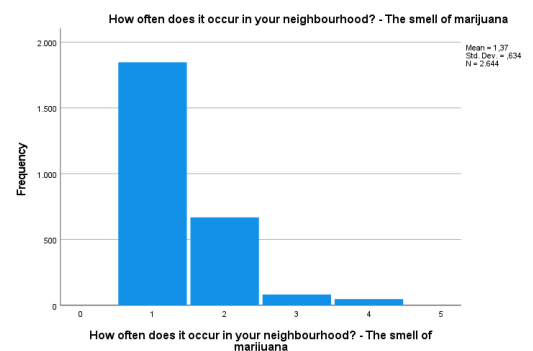
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	2023	76,0	76,5	76,5
	soms	567	21,3	21,4	98,0
	vaak	40	1,5	1,5	99,5
	heel vaak	14	,5	,5	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a054 - De geur van marihuana

How often does it occur in your neighbourhood? - The smell of marijuana

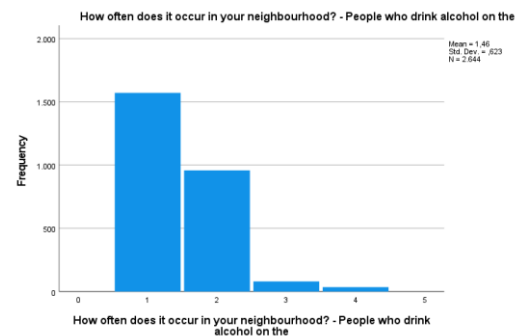
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1847	69,4	69,9	69,9
	soms	668	25,1	25,3	95,1
	vaak	82	3,1	3,1	98,2
	heel vaak	47	1,8	1,8	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a055 - Mensen die alcohol drinken op straat

How often does it occur in your neighbourhood? - People who drink alcohol on the

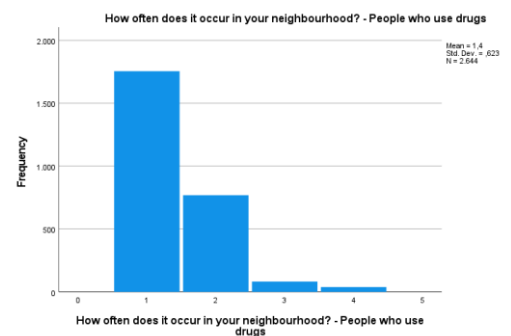
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1571	59,0	59,4	59,4
	soms	958	36,0	36,2	95,7
	vaak	80	3,0	3,0	98,7
	heel vaak	35	1,3	1,3	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a056 - Mensen die drugs gebruiken

How often does it occur in your neighbourhood? - People who use drugs

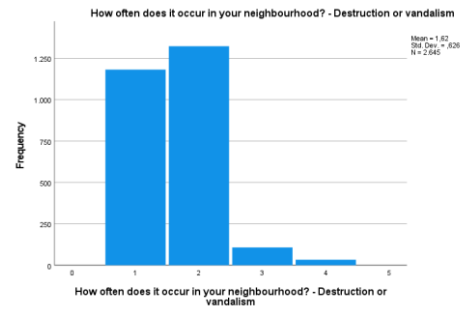
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1755	66,0	66,4	66,4
	soms	769	28,9	29,1	95,5
	vaak	82	3,1	3,1	98,6
	heel vaak	38	1,4	1,4	100,0
	Total	2644	99,4	100,0	
Missing	System	17	,6		
Total		2661	100,0		



Sr20a057 - Vandalisme en vernieling

How often does it occur in your neighbourhood? - Destruction or vandalism

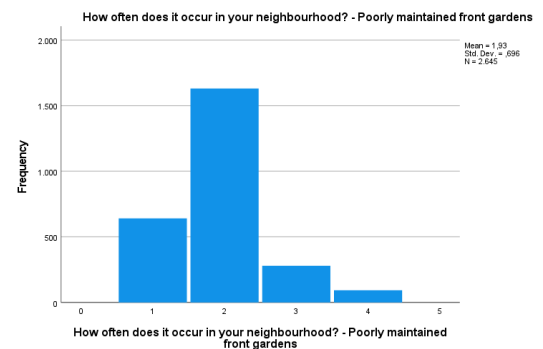
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	1182	44,4	44,7	44,7
	soms	1323	49,7	50,0	94,7
	vaak	107	4,0	4,0	98,8
	heel vaak	33	1,2	1,2	100,0
	Total	2645	99,4	100,0	
Missing	System	16	,6		
Total		2661	100,0		



Sr20a058 - Slecht onderhouden voortuinen

How often does it occur in your neighbourhood? - Poorly maintained front gardens

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	641	24,1	24,2	24,2
	soms	1631	61,3	61,7	85,9
	vaak	280	10,5	10,6	96,5
	heel vaak	93	3,5	3,5	100,0
	Total	2645	99,4	100,0	
Missing	System	16	,6		
Total		2661	100,0		



Deze variabelen samen kunnen de schaalverdeling van overlast vormen. Overlast is namelijk een breed begrip waarbij verschillende factoren horen, daarom zijn ze samengevoegd. Het samenvoegen was mogelijk, omdat deze variabelen samen een Cronbach's alpha had van 0.822.

*Controleren betrouwbaarheid overlast.

RELIABILITY

```
/VARIABLES=sr20a048 sr20a049 sr20a050 sr20a051 sr20a052 sr20a053 sr20a054 sr20a055
sr20a056
```

```
sr20a057 sr20a058
```

```
/SCALE('ALL VARIABLES') ALL
```

```
/MODEL=ALPHA
```

```
/STATISTICS=DESCRIPTIVE SCALE
```

```
/SUMMARY=TOTAL.
```

Reliability Statistics

Cronbach's Alpha	N of Items
,822	11

Om de schaalverdeling te maken, is het volgende commando gebruikt in de syntax:

*Schaal maken voor overlast.

```
COMPUTE Overlast_schaal=(sr20a047 + sr20a049 + sr20a050 + sr20a051 + sr20a052 + sr20a053 +  
sr20a054 +  
sr20a055 + sr20a056 + sr20a057 + sr20a058) / 11.
```

EXECUTE.

Schaalverdeling overlast

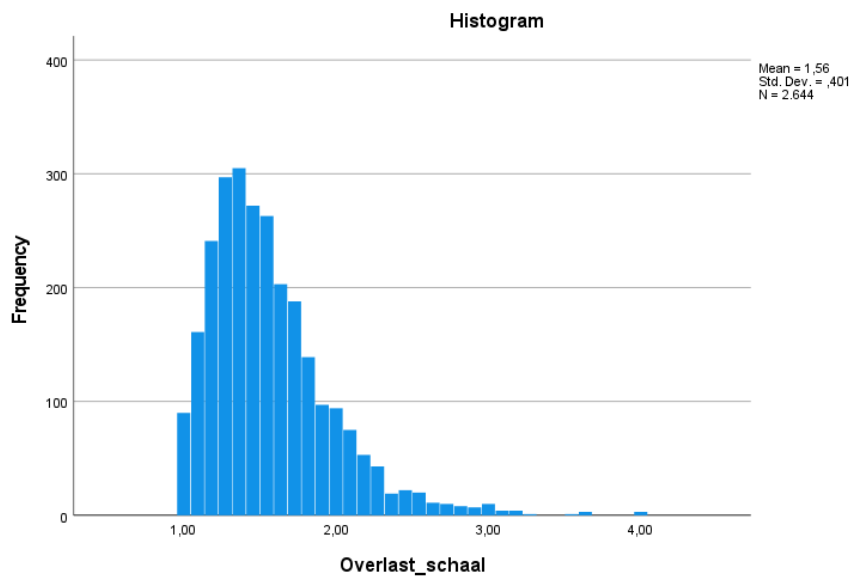
Naam in dataset: 'Overlast'

Omschrijving: Overlast op straat

Type: ordinaal

*Frequentieverdeling voor overlast.

```
FREQUENCIES VARIABLES=Overlast_schaal  
/NTILES=4  
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN  
/HISTOGRAM  
/ORDER=ANALYSIS.
```



Overlast is gemeten door respondenten te vragen hoe vaak de volgende zaken voorkwamen in hun buurt:

- Afval naast de prullenbak
- Kapotte fietsen
- Hinder veroorzaakt door jongeren
- Hinder veroorzaakt door ouderen

- Geluidsoverlast
- Ruzies en vechtpartijen
- De geur van marihuana
- Mensen die alcohol drinken op straat
- Mensen die drugs gebruiken
- Vandalisme en vernieling
- Slecht verzorgde voortuinen

Respondenten konden de stellingen beantwoorden door aan te geven op schaal van een tot en met vier hoe vaak de verschillende zaken voorkwamen; waarbij de score 1 aangeeft dat het nooit voorkomt en vier aangeeft dat het heel vaak voorkomt.

Overlast_schaal					
	Valid	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	90	3,4	3,4	3,4
	1,09	181	6,1	6,1	9,5
	1,18	241	9,1	9,1	18,6
	1,27	297	11,2	11,2	29,8
	1,36	305	11,5	11,5	41,4
	1,45	272	10,2	10,3	51,7
	1,55	283	9,9	9,9	61,6
	1,64	203	7,6	7,7	69,3
	1,73	188	7,1	7,1	76,4
	1,82	139	5,2	5,3	81,7
	1,91	97	3,6	3,7	85,3
	2,00	94	3,5	3,6	88,9
	2,09	75	2,8	2,8	91,7
	2,18	53	2,0	2,0	93,7
	2,27	43	1,6	1,6	95,3
	2,36	19	,7	,7	96,1
	2,45	22	,8	,8	96,9
	2,55	20	,8	,8	97,7
	2,64	11	,4	,4	98,1
	2,73	10	,4	,4	98,4
	2,82	8	,3	,3	98,8
	2,91	7	,3	,3	99,0
	3,00	10	,4	,4	99,4
	3,09	4	,2	,2	99,5
	3,18	4	,2	,2	99,7
	3,27	1	,0	,0	99,7
	3,55	1	,0	,0	99,8
	3,64	3	,1	,1	99,9
	4,00	3	,1	,1	100,0
	Total	2644	99,4	100,0	
Missin	System	17	,8		
Total		2661	100,0		

De verdeling voor overlast op straat is wat linksscheef; dit houdt in dat respondenten niet heel vaak overlast ervaren. Ook is te zien aan de spreiding dat deze niet heel groot is; er zijn bijna geen respondenten die de vraag beantwoorden met de antwoordcategorie 4 – heel vaak. De meeste respondenten geven aan dat ze bijna nooit tot nooit overlast ervaren in hun buurt. Respondenten die aangeven vaak last te hebben van overlast, zijn uitbijters omdat zij buiten de normale verdeling vallen. Ondanks dat de variabele bestaat uit verschillende categorieën, zal ik de variabele meenemen in de dataset als continu. Dit omdat er door het samenvoegen om tot een schaalverdeling te komen, veel verschillende categorieën zijn ontstaan.

Ervaren veiligheid op straat in de avond

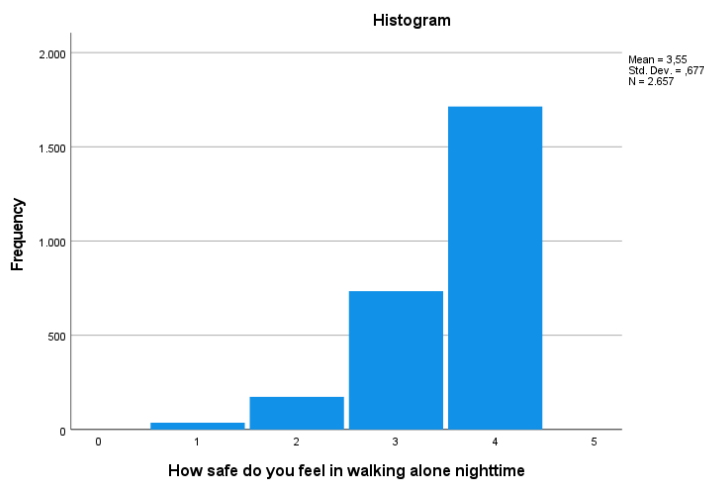
Naam in dataset: 'sr20a007'

Omschrijving: How safe do you feel in walking alone nighttime

Type: Ordinaal

GRAPH

```
/BAR(SIMPLE)=COUNT BY sr20a007.  
FREQUENCIES VARIABLES=sr20a007  
/NTILES=4  
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN  
/HISTOGRAM  
/ORDER=ANALYSIS.
```



De ervaren veiligheid is geen schaalverdeling van ervaren veiligheid overdag en in de avond omdat deze een matige correlatie hadden. Ze meten dus niet helemaal hetzelfde en daarom heb ik gekozen voor de ervaren veiligheid die mensen ervaren in de avond. Als mensen zich overdag onveilig voelen, hebben ze dat vaak in de avond ook denk ik en daarom heb ik gekozen voor ervaren veiligheid in de avond.

De ervaren veiligheid is gemeten doordat respondenten konden aangeven op schaal van een tot en met vier hoe veilig ze zich voelen op straat in de avond als ze alleen lopen; hierbij is één heel onveilig en vier heel veilig.

*Frequentieverdeling voor ervaren veiligheid avond.

```
DATASET ACTIVATE DataSet1.
```

```
FREQUENCIES VARIABLES=sr20a007
```

```
/NTILES=4
```

```
/STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
```

```
/HISTOGRAM
```

```
/ORDER=ANALYSIS.
```


How safe do you feel in walking alone nighttime					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	very unsafe	36	1,4	1,4	1,4
	a bit unsafe	173	6,5	6,5	7,9
	a bit safe	734	27,6	27,6	35,5
	very safe	1714	64,4	64,5	100,0
	Total	2657	99,8	100,0	
Missing	System	4	,2		
Total		2661	100,0		

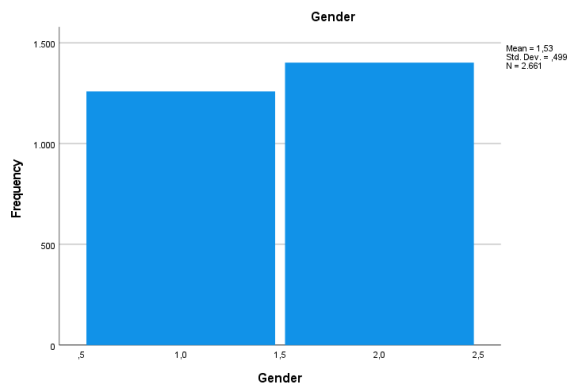
De verdeling van ervaren veiligheid is rechtsscheef; dit geeft aan dat de meeste respondenten aangeven zich heel veilig of veilig te voelen in hun buurt. Er zijn minder respondenten die hebben aangegeven zich erg onveilig te voelen in hun buurt als ze 's avonds alleen over straat lopen. Personen die zich erg onveilig voelen, zijn dan ook een uitbijter. Zij vallen buiten de normale verdeling. Het is goed om tijdens het onderzoek in ons achterhoofd te houden dat de verdeling van ervaren veiligheid linksscheef is. Dit kan namelijk invloed hebben op de resultaten.

Geslacht

Naam in dataset: 'Geslacht'

Omschrijving: Gender

Type: Nominaal



Respondenten konden aangeven of ze man of vrouw zijn; hierbij is één man en twee vrouw.

Gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	1259	47,3	47,3	47,3
	Female	1402	52,7	52,7	100,0
Total		2661	100,0	100,0	

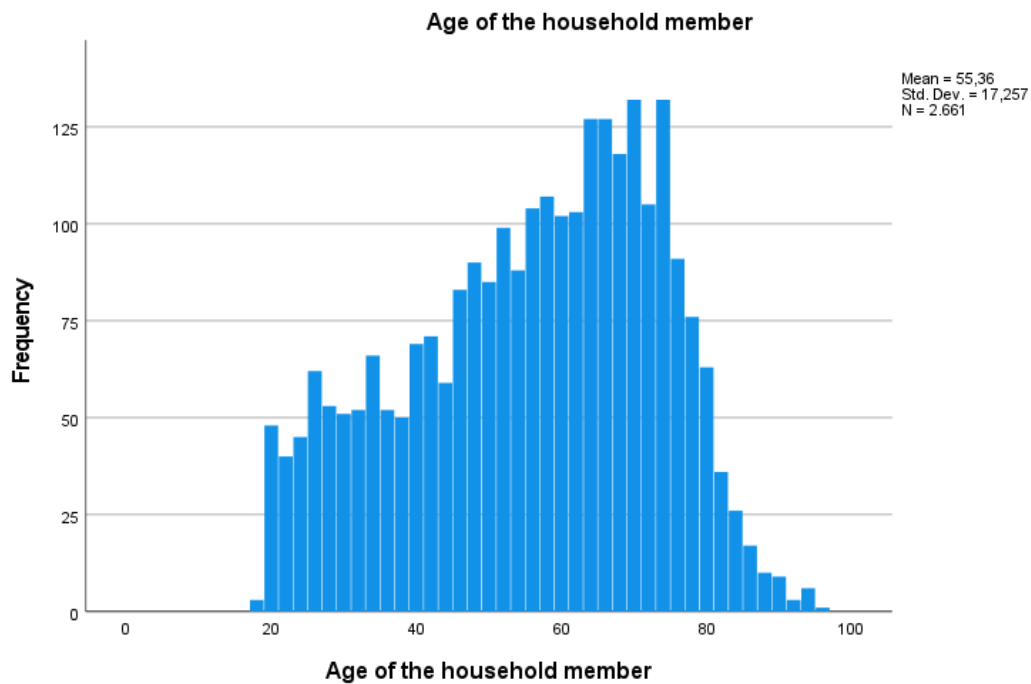
Geslacht is verdeeld in 2 categorieën; mannen en vrouwen. De categorieën zijn allebei ongeveer even groot; van de respondenten was een iets groter deel vrouw. Maar het verschil is verwaarloosbaar. Dit is goed omdat zo de mannen en vrouwen ongeveer even goed gerepresenteerd worden in het onderzoek.

Leeftijd

Naam in dataset: 'Leeftijd'

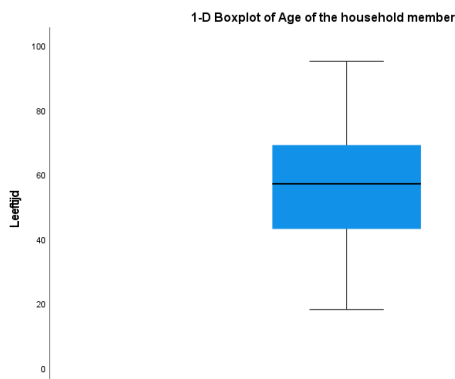
Omschrijving: Age of the household member

Type: Ratio



Respondenten konden aangeven welke leeftijd ze zijn door hun huidige leeftijd in te vullen. De jongste persoon in de dataset was 18 en de oudste was 95.

Leeftijd is een continue variabele; het gemiddelde is 55.36 en de standaarddeviatie is 17.257.



5 number summary:

- Minimum: 18
- Eerste kwartiel: 43
- Mediaan: 57
- Derde kwartiel: 69
- Maximum: 95

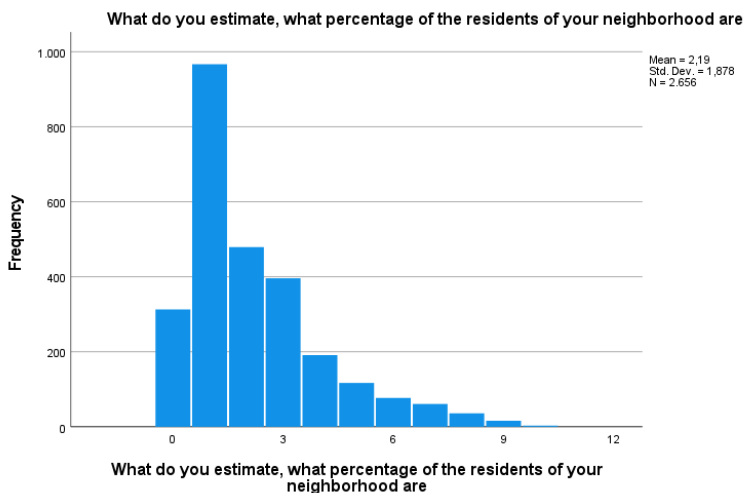
De verdeling van de variabele leeftijd is niet normaal; er hebben meer mensen meegedaan met een hogere leeftijd dan mensen met een jonge leeftijd; er is een piek van respondenten in de leeftijd tussen 60 en 75. Ondanks dat zijn er wel voor alle leeftijden respondenten en daarom is de variabele alsnog goed te gebruiken voor dit onderzoek. In de resultaten moeten we wel stilstaan bij het feit dat de verdeling wat rechtsscheef is. De spreiding in leeftijd is groot; mensen van alle leeftijden hebben deelgenomen aan het onderzoek. Dit kan ervoor zorgen dat mensen verschillende ervaringen hebben door hun leeftijd, dat maakt dat het goed is om te controleren voor deze variabele.

Percentage buren met buitenlandse komaf

Naam in de dataset: 'sr20a019'

Omschrijving: What do you estimate, what percentage of the residents of your neighborhood are

Type: Ratio, gebruikt in het onderzoek als continue variabele



What do you estimate, what percentage of the residents of your neighborhood are

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0%	313	11,8	11,8	11,8
10%	967	36,3	36,4	48,2
20%	479	18,0	18,0	66,2
30%	396	14,9	14,9	81,1
40%	191	7,2	7,2	88,3
50%	117	4,4	4,4	92,7
60%	77	2,9	2,9	95,6
70%	61	2,3	2,3	97,9
80%	36	1,4	1,4	99,3
90%	16	,6	,6	99,9
100%	3	,1	,1	100,0
Total	2656	99,8	100,0	
Missing System	5	,2		
Total	2661	100,0		

Respondenten konden de vraag beantwoorden door aan te geven welk percentage buren naar hun idee van buitenlandse komaf was; 0%; 10%; 20% (...); 100%.

De verdeling van het percentage buren met buitenlandse komaf is linksscheef. Dit houdt in dat het percentage buren van buitenlandse komaf in de buurten van de respondenten niet heel hoog is. De meeste respondenten schatten het percentage op 10%. Een percentage van 90% of 100% zijn uitbijters;

er zijn bijna geen respondenten die in buurten wonen met heel veel burens met buitenlandse komaf. Het is voor het onderzoek goed om te onthouden dat de onderzochte buurten overwegend homogeen zijn.

Bijlage 2

Voor het uitvoeren van dit onderzoek zijn verschillende soorten regressies uitgevoerd. Eerst een lineaire regressie om conclusies te kunnen trekken uit de data die overzichtelijk zijn voor mij als onderzoeker. Door de schending van de assumptie normaliteit, is vervolgens een ordinale logistische regressie gedaan. De assumptie van de test of parallel lines was ook geschonden dus vervolgens ben ik overgegaan op binaire logistische regressie. Deze is uiteindelijk in het onderzoek toegevoegd. In deze bijlagen zijn de uitkomsten van alle lineaire regressies te vinden.

Bivariate analyse

De bivariate verdelingen laten de correlaties en andere associatiematen zien.

In Tabel 1 zijn de gegevens samengevat. Daarna volgen de verschillende resultaten uit SPSS.

Tabel 1: Correlaties, Cramer's V waarden en de r waarden uit ANOVA

	Sociale cohesie (continu)	Overlast (continu)	Ervaren veiligheid (categorisch)	Geslacht (categorisch)	Leeftijd (continu)	Percentage burens met buitenlandse komaf (continu)
Sociale cohesie	1	-0.410	0.345**	0.000**	0.096	-0.337
Overlast	-0,410	1	0.390**	0.000**	-0.172	0.504
Ervaren veiligheid	0,344**	0,390**	1	0.228*	0.063**	0.349
Geslacht	0.000**	0.000**	0,228*	1	0,105**	0.090**
Leeftijd	0,094	-0,172	0,063**	0.105**	1	-0.223
Percentage burens met buitenlandse komaf	0.349	0.508	0,230**	0.090**	0,235	1

*Cramer's V waarde in plaats van correlatie

**r uit ANOVA in plaats van correlatie

Continue variabelen

*Correlaties van alle continue variabelen; sociale cohesie, overlast en leeftijd, % burens.

CORRELATIONS

```
/VARIABLES=leeftijd SocialeCohesie Overlast_schaal sr20a019
/PRINT=TWOTAIL NOSIG FULL
/MISSING=PAIRWISE.
```

Correlations

		Age of the household member	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	Overlast_schaal	What do you estimate, what percentage of the residents of your neighborhood are
Age of the household member	Pearson Correlation	1	,096**	-,172**	-,223**
	Sig. (2-tailed)		<,001	<,001	<,001
	N	2644	2644	2644	2644
Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	Pearson Correlation	,096**	1	-,410**	-,337**
	Sig. (2-tailed)	<,001		<,001	<,001
	N	2644	2644	2644	2644
Overlast_schaal	Pearson Correlation	-,172**	-,410**	1	,504**
	Sig. (2-tailed)	<,001	<,001		<,001
	N	2644	2644	2644	2644
What do you estimate, what percentage of the residents of your neighborhood are	Pearson Correlation	-,223**	-,337**	,504**	1
	Sig. (2-tailed)	<,001	<,001	<,001	
	N	2644	2644	2644	2644

** . Correlation is significant at the 0.01 level (2-tailed).

- De correlatie tussen leeftijd en overlast is -0.172 en deze is significant; $p < 0.001$ en $N = 2644$
- De correlatie tussen leeftijd en sociale cohesie is 0.096 en deze is significant; $p < 0.001$ en $N = 2644$
- De correlatie tussen leeftijd en percentage burens is -0.223 en deze is significant; ; $p < 0.001$ en $N = 2644$
- De correlatie tussen overlast en sociale cohesie is -0.410 en deze is significant; $p < 0.001$ en $N = 2644$
- De correlatie tussen overlast en percentage burens is 0.504 en deze is significant; $p < 0.001$ en $N = 2644$
- De correlatie tussen sociale cohesie en percentage burens is -0.337 en deze is significant; $p < 0.001$ en $N = 2644$

Continue variabele sociale cohesie en categorische variabele ervaren veiligheid

*Gemiddelden in groepen sociale cohesie en ervaren veiligheid.

UNIANOVA SocialeCohesie BY sr20a007

```

/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=sr20a007.

```

Descriptive Statistics

Dependent Variable: Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen

How safe do you feel in walking alone nighttime	Mean	Std. Deviation	N
very unsafe	2,4074	,98274	36
a bit unsafe	2,9499	,85442	173
a bit safe	3,3397	,79677	728
very safe	3,7502	,73713	1707
Total	3,5666	,81535	2644

Tests of Between-Subjects Effects

Dependent Variable: Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	209,203 ^a	3	69,734	118,938	<,001
Intercept	4362,214	1	4362,214	7440,111	,000
sr20a007	209,203	3	69,734	118,938	<,001
Error	1547,859	2640	,586		
Total	35389,778	2644			
Corrected Total	1757,062	2643			

a. R Squared = ,119 (Adjusted R Squared = ,118)

- $F = 118.938$; $p < 0.001$
- $R^2 = 0.119$

Continue variabele leeftijd en categorische variabele ervaren veiligheid

*Gemiddelden in groepen leeftijd en ervaren veiligheid.

UNIANOVA leeftijd BY sr20a007

```

/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=sr20a007.
    
```

Descriptive Statistics			
Dependent Variable: Age of the household member			
How safe do you feel in walking alone nighttime	Mean	Std. Deviation	N
very unsafe	53,19	19,793	36
a bit unsafe	53,59	19,837	173
a bit safe	54,06	18,659	728
very safe	56,24	16,204	1707
Total	55,43	17,242	2644

Tests of Between-Subjects Effects

Dependent Variable: Age of the household member

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3270,338 ^a	3	1090,113	3,678	,012
Intercept	1326852,856	1	1326852,856	4476,901	,000
sr20a007	3270,338	3	1090,113	3,678	,012
Error	782436,720	2640	296,378		
Total	8908590,000	2644			
Corrected Total	785707,057	2643			

a. R Squared = ,004 (Adjusted R Squared = ,003)

- $F = 3.678$; $p = 0.012$
- $R^2 = 0.004$

Continue variabele overlast en categorische variabele ervaren veiligheid

*Gemiddelden in groepen overlast en ervaren veiligheid.

UNIANOVA Overlast BY sr20a007

```

/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/PRINT DESCRIPTIVE
/CRITERIA=ALPHA(.05)
/DESIGN=sr20a007.
    
```

Descriptive Statistics			
Dependent Variable: Overlast op straat			
How safe do you feel in walking alone nighttime	Mean	Std. Deviation	N
very unsafe	2,6641	,70646	36
a bit unsafe	3,0909	,48927	173
a bit safe	3,3360	,39423	728
very safe	3,5310	,33342	1707
Total	3,4367	,40148	2644

Tests of Between-Subjects Effects

Dependent Variable: Overlast op straat

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	64,735 ^a	3	21,578	157,679	<,001
Intercept	4485,602	1	4485,602	32777,386	,000
sr20a007	64,735	3	21,578	157,679	<,001
Error	361,285	2640	,137		
Total	31654,694	2644			
Corrected Total	426,021	2643			

a. R Squared = ,152 (Adjusted R Squared = ,151)

- $F = 157.679$; $p < 0.001$
- $R^2 = 0.152$

Continue variabele percentage buren met buitenlandse komaf en categorische variabele ervaren veiligheid

*Gemiddelden in groepen % buren en ervaren veiligheid.

UNIANOVA sr20a019 BY sr20a007

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/PRINT DESCRIPTIVE

/CRITERIA=ALPHA(.05)

/DESIGN=sr20a007.

Descriptive Statistics

Dependent Variable: What do you estimate, what percentage of the residents of your neighborhood are

How safe do you feel in walking alone nighttime

	Mean	Std. Deviation	N
very unsafe	5,31	2,364	36
a bit unsafe	3,70	2,356	173
a bit safe	2,60	1,898	728
very safe	1,79	1,612	1707
Total	2,19	1,879	2644

Tests of Between-Subjects Effects

Dependent Variable: What do you estimate, what percentage of the residents of your neighborhood are

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1134,923 ^a	3	378,308	121,798	<,001
Intercept	5055,839	1	5055,839	1627,753	<,001
sr20a007	1134,923	3	378,308	121,798	<,001
Error	8199,901	2640	3,106		
Total	22001,000	2644			
Corrected Total	9334,824	2643			

a. R Squared = ,122 (Adjusted R Squared = ,121)

- $F = 121.798$; $p < 0.001$
- $R^2 = 0.122$

Continue variabele sociale cohesie en categorische variabele geslacht

*Gemiddelden in groepen sociale cohesie en geslacht.

UNIANOVA SocialeCohesie BY geslacht

/METHOD=SSTYPE(3)

/INTERCEPT=INCLUDE

/PRINT DESCRIPTIVE

/CRITERIA=ALPHA(.05)

/DESIGN=geslacht.

Descriptive Statistics

Dependent Variable: Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen

Gender	Mean	Std. Deviation	N
Male	3,5641	,78973	1254
Female	3,5688	,83807	1390
Total	3,5666	,81535	2644

Tests of Between-Subjects Effects

Dependent Variable: Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,015 ^a	1	,015	,022	,881
Intercept	33541,426	1	33541,426	50434,870	,000
geslacht	,015	1	,015	,022	,881
Error	1757,047	2642	,665		
Total	35389,778	2644			
Corrected Total	1757,062	2643			

a. R Squared = ,000 (Adjusted R Squared = ,000)

- $F = 0.022$; $p = 0.881$
- $R^2 = 0.000$

Continue variabele overlast en categorische variabele geslacht

*Gemiddelden in groepen overlast en geslacht.

UNIANOVA Overlast_schaal BY geslacht

```
/METHOD=SSTYPE(3)  
/INTERCEPT=INCLUDE  
/PRINT DESCRIPTIVE  
/CRITERIA=ALPHA(.05)  
/DESIGN=geslacht.
```

Descriptive Statistics

Dependent Variable: Overlast_schaal

Gender	Mean	Std. Deviation	N
Male	1,5559	,39011	1254
Female	1,5699	,41150	1390
Total	1,5633	,40148	2644

Tests of Between-Subjects Effects

Dependent Variable: Overlast_schaal

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,130 ^a	1	,130	,804	,370
Intercept	6441,332	1	6441,332	39958,564	,000
geslacht	,130	1	,130	,804	,370
Error	425,891	2642	,161		
Total	6887,421	2644			
Corrected Total	426,021	2643			

a. R Squared = ,000 (Adjusted R Squared = ,000)

- $F = 0.804$; $p = 0.370$
- $R^2 = 0.000$

Continue variabele leeftijd en categorische variabele geslacht

*Gemiddelden in groepen leeftijd en geslacht.

UNIANOVA leeftijd BY geslacht

```
/METHOD=SSTYPE(3)  
/INTERCEPT=INCLUDE
```

/PRINT DESCRIPTIVE
 /CRITERIA=ALPHA(.05)
 /DESIGN=geslacht.

Descriptive Statistics

Dependent Variable: Age of the household member

Gender	Mean	Std. Deviation	N
Male	57,32	16,765	1254
Female	53,72	17,492	1390
Total	55,43	17,242	2644

Tests of Between-Subjects Effects

Dependent Variable: Age of the household member

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	8534,954 ^a	1	8534,954	29,015	<,001
Intercept	8128465,387	1	8128465,387	27632,754	,000
geslacht	8534,954	1	8534,954	29,015	<,001
Error	777172,103	2642	294,161		
Total	8908590,000	2644			
Corrected Total	785707,057	2643			

a. R Squared = ,011 (Adjusted R Squared = ,010)

- F = 29.015; p<0.001
- R² = 0.011

Continue variabele percentage buren met buitenlandse komaf en categorische variabele geslacht

*Gemiddelden in groepen % buren en geslacht.

UNIANOVA sr20a019 BY geslacht

/METHOD=SSTYPE(3)
 /INTERCEPT=INCLUDE
 /PRINT DESCRIPTIVE
 /CRITERIA=ALPHA(.05)
 /DESIGN=geslacht.

Descriptive Statistics

Dependent Variable: What do you estimate, what percentage of the residents of your neighborhood are

Gender	Mean	Std. Deviation	N
Male	2,08	1,847	1254
Female	2,29	1,903	1390
Total	2,19	1,879	2644

Tests of Between-Subjects Effects

Dependent Variable: What do you estimate, what percentage of the residents of your neighborhood are

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	30,443 ^a	1	30,443	8,644	,003
Intercept	12568,948	1	12568,948	3568,981	,000
geslacht	30,443	1	30,443	8,644	,003
Error	9304,381	2642	3,522		
Total	22001,000	2644			
Corrected Total	9334,824	2643			

a. R Squared = ,003 (Adjusted R Squared = ,003)

- F = 8.644; p=0.003
- R² = 0.003

Kruistabel voor geslacht en ervaren veiligheid

*Kruistabel geslacht en ervaren veiligheid.

CROSSTABS

```
/TABLES=geslacht BY sr20a007  
/FORMAT=AVALUE TABLES  
/STATISTICS=CHISQ PHI  
/CELLS=COUNT  
/COUNT ROUND CELL.
```

Gender * How safe do you feel in walking alone nighttime Crosstabulation

Count		How safe do you feel in walking alone nighttime				Total
		very unsafe	a bit unsafe	a bit safe	very safe	
Gender	Male	12	41	251	950	1254
	Female	24	132	477	757	1390
Total		36	173	728	1707	2644

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	137,215 ^a	3	<,001
Likelihood Ratio	140,642	3	<,001
Linear-by-Linear Association	121,075	1	<,001
N of Valid Cases	2644		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 17,07.

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,228	<,001
	Cramer's V	,228	<,001
N of Valid Cases		2644	

- X^2 (df=3; N=2644) = 137.215; $p < 0.001$
- Cramer's V = 0.228

Voor de continue variabelen is de correlatie gebruikt, voor de continue variabelen in relatie tot categorische variabelen zijn de wortels getrokken uit de R^2 van de anova tabellen en voor categorische variabelen is gekeken naar de Cramer's V.

Multivariate analyse

Lineaire regressie

In tabel 2 staan de modellen voor de lineaire regressie.

Eerst wordt de onderzoeksvraag onderzocht door middel van lineaire regressie. Om het eindmodel tot stand te laten komen, zijn verschillende modellen geschat:

- Model 1: Hierin wordt de ervaren veiligheid geschat uit de variabelen geslacht, leeftijd en percentage burens met buitenlandse komaf. Dit is dus het model met alleen de controlevariabelen

- Model 2: In dit model wordt de variabele sociale cohesie toegevoegd. De ervaren veiligheid wordt dus geschat uit geslacht, leeftijd, percentage buren met buitenlandse komaf en sociale cohesie
- Model 3: Als laatste wordt de variabele overlast toegevoegd. De ervaren veiligheid wordt dan dus uit 4 variabelen geschat.

De eerste hypothese gaat over het hoofdeffect en is als volgt: ‘Een hogere sociale cohesie zorgt voor meer ervaren veiligheid in de buurt.’ Om deze hypothese te toetsen, zal gekeken worden naar model 2. De tweede hypothese gaat over het mediërende effect van overlast en is als volgt geformuleerd: ‘Een hogere sociale cohesie kan ervoor zorgen dat overlast daalt waardoor iemand zich veiliger voelt in zijn of haar buurt.’ Om deze hypothese te toetsen, kijken we naar model 3.

*Lineaire regressie in blokken.

DATASET ACTIVATE DataSet1.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N

/MISSING LISTWISE

/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE

/CRITERIA=PIN(.05) POUT(.10)

/NOORIGIN

/DEPENDENT sr20a007

/METHOD=ENTER geslacht leeftijd sr20a019

/METHOD=ENTER SocialeCohesie

/METHOD=ENTER Overlast_schaal.

Tabel 3: parameterschattingen (n=2644)

	Model 1 ^a		Model 2 ^a		Model 3 ^b		Model 4 ^a	
	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>	<i>b</i> (SE)	<i>p</i>
<i>Constante</i>	4.312 (0.062)	<0.001	3.486 (0.083)	<0.001	1.936 (0.062)	0.046	4.219 (0.105)	<0.001
<i>Geslacht</i>	-0.269 (0.024)	<0.001	-0.278 (0.024)	<0.001	-0.009 (0.024)	0.013	-0.281 (0.023)	<0.001
<i>Leeftijd</i>	-0.001 (0.001)	0.051	-0.02 (0.001)	<0.017	-0.001 (0.001)	<0.001	-0.002 (0.001)	0.001
<i>Percentage buren met buitenlandse komaf</i>	-0.123 (0.007)	<0.001	-0.091 (0.007)	<0.001	0.086 (0.007)	0.004	-0.059 (0.007)	<0.001
<i>Sociale cohesie</i>			0.220 (0.015)	<0.001	-0.133 (0.015)	0.008	0.169 (0.016)	<0.001
<i>Overlast</i>							-0.379 (0.035)	<0.001
R²	0.158		0.219		0.0323		0.253	

(Partiële) F	164.743	<0.001	185.431	<0.001	314.349	<0.001	179.103	<0.001
--------------	---------	--------	---------	--------	---------	--------	---------	--------

^a Afhankelijke variabele: ervaren veiligheid; ^b Afhankelijke variabele: overlast

Model 2:

- We zien dat er een positieve helling is voor sociale cohesie; dus als sociale cohesie met 1 eenheid omhoog gaat, dan gaat ervaren veiligheid met 0.220 eenheden omhoog. Dus de ervaren veiligheid is hoger als mensen meer sociale cohesie ervaren.
- Het effect van de negatieve helling van het percentage burens met buitenlandse komaf wat kleiner wordt.
- Het effect van leeftijd is verwaarloosbaar omdat er maar een hele kleine helling is.
- Voor geslacht is een negatieve helling.

Model 3:

- Dit is het model met de mediator als afhankelijke variabele.
- We zien een significant effect van sociale cohesie, maar wel bij een grens van $\alpha=0.05$. De grens van $\alpha<0.001$ wordt wel overschreden. Maar dit zou geen probleem moeten zijn, gezien de resultaten van de binaire logistische regressie.
-

Model 4:

- Het effect van sociale cohesie is nog steeds positief, maar wel veel kleiner.
- Er is een negatieve helling voor overlast. Deze helling is redelijk groot en heeft dus een redelijk groot effect op de ervaren veiligheid.
- Model 3 heeft de meeste variantie verklaard ($R^2 = 0.253$). Dit houdt in dat de toevoeging van overlast waardevol is voor het begrijpen van het effect van sociale cohesie op ervaren veiligheid.

Ordinale logistische regressie

Omdat de assumptie van normaliteit is geschonden, is het de moeite waard om logistische regressie te doen. In tabel 3 zijn de resultaten voor ordinale logistische regressie te zien. Als de resultaten van de logistische regressie onderstrepen wat de resultaten van de lineaire regressie zeggen, zijn de resultaten robuust.

*Ordinale logistische regressie & test of parallel lines.

```
PLUM sr20a007 WITH SocialeCohesie Overlast_schaal geslacht leeftijd sr20a019
/CRITERIA=CIN(95) DELTA(0) LCONVERGE(0) MXITER(100) MXSTEP(5)
PCONVERGE(1.0E-6) SINGULAR(1.0E-8)
/LINK=LOGIT
/PRINT=FIT KERNEL PARAMETER SUMMARY TPARALLEL.
```

Tabel 3: output ordinale logistische regressie

		Parameter Estimates						
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[sr20a007 = 1]	-7,317	,442	274,232	1	<,001	-8,183	-6,451
	[sr20a007 = 2]	-5,207	,404	166,174	1	<,001	-5,999	-4,415
	[sr20a007 = 3]	-2,924	,391	55,815	1	<,001	-3,691	-2,157
Location	SocialeCohesie	,635	,059	116,331	1	<,001	,519	,750
	Overlast_schaal	-1,236	,124	99,599	1	<,001	-1,479	-,994
	geslacht	-1,144	,091	156,513	1	<,001	-1,323	-,964
	leeftijd	-,007	,003	7,910	1	,005	-,012	-,002
	sr20a019	-,181	,025	50,664	1	<,001	-,231	-,131

Link function: Logit.

Bij ordinale logistische regressie is de aanname dat de hellingen gelijk zijn voor alle constanten, per stap van de dichotomisaties waarbij eerst wordt gekeken van heel onveilig tegenover een beetje onveilig, een beetje veilig en heel veilig. Dan wordt er gekeken naar de dichotomisatie van heel onveilig en een beetje onveilig ten opzichte van een beetje veilig en heel veilig. Als laatste wordt gekeken naar de dichotomisatie van heel onveilig, een beetje onveilig en een beetje veilig ten opzichte van heel veilig. Om te controleren of de hellingen ook gelijk zijn bij elke constante van de dichotomisaties, doen we de test of parallel lines. De uitkomst van deze test is te zien in tabel 4:

Tabel 4: uitkomsten test of parallel lines

Test of Parallel Lines ^a				
Model	-2 Log Likelihood ^b	Chi-Square	df	Sig.
Null Hypothesis	3929,390			
General	3905,225	24,165	10	,007

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The kernel of the log-likelihood function is displayed.

In de test of parallel lines wordt gekeken of de nulhypothese dat de helling voor alle responsiecategorieën gelijk is, in de tabel is te zien dat de test significant is ($\alpha = 0.007$) dus de nulhypothese wordt verworpen. Dit kan komen omdat de steekproef groot is.

De hellingen zijn dus niet voor alle constanten van de dichotomisaties hetzelfde. Daarom is het goed om voor alle categorieën dummy's te maken en dan te kijken bij welke categorie de helling eventueel anders is.

Binaire logistische regressie

Omdat de hellingen verschillen, zijn dummy's gemaakt. De resultaten zijn te zien in tabel 5 ('heel onveilig' tegenover de andere variabelen), tabel 6 ('heel onveilig' en 'beetje onveilig' tegenover 'beetje veilig' en 'heel veilig') en tabel 7 ('heel onveilig', 'beetje onveilig' en 'beetje veilig' tov 'heel veilig').

*Binaire logistische met binaire variabele 1-234.

LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_1v234
 /METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
 /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

*Binaire logistische regressie met binaire variabele 12-34.

LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_12v34
 /METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
 /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

*Binaire logistische regressie met binaire variabele 123-4.

LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_123v4
 /METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
 /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

Tabel 5: binaire logistische regressie 'heel onveilig' tov de andere variabelen

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-,471	,385	1,496	1	,221	,624
	Age of the household member	-,020	,011	3,121	1	,077	,980
	What do you estimate, what percentage of the residents of your neighborhood are	-,326	,087	14,129	1	<,001	,722
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,800	,216	13,677	1	<,001	2,225
	Overlast_schaal	-1,301	,368	12,492	1	<,001	,272
	Constant	7,236	1,521	22,637	1	<,001	1387,983

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Tabel 6: binaire logistische regressie 'heel onveilig en 'beetje onveilig' tov 'beetje veilig' en 'heel veilig'

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-1,213	,181	44,977	1	<,001	,297
	Age of the household member	-,015	,005	9,407	1	,002	,985
	What do you estimate, what percentage of the residents of your neighborhood are	-,236	,041	33,570	1	<,001	,790
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,707	,102	48,323	1	<,001	2,029
	Overlast_schaal	-1,276	,198	41,627	1	<,001	,279
	Constant	5,811	,713	66,469	1	<,001	333,967

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Tabel 7: binaire logistische regressie 'heel onveilig', 'beetje onveilig' en 'beetje veilig' tov 'heel veilig'

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-1,122	,095	139,642	1	<,001	,326
	Age of the household member	-,006	,003	4,299	1	,038	,994
	What do you estimate, what percentage of the residents of your neighborhood are	-,160	,028	33,260	1	<,001	,852
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,606	,063	92,243	1	<,001	1,833
	Overlast_schaal	-1,139	,140	66,176	1	<,001	,320
	Constant	2,711	,420	41,731	1	<,001	15,046

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Na het maken van de dummy's en binaire logistische regressie uit te voeren voor elk model, is te zien dat de hellingen vooral verschillen wanneer de antwoordcategorie heel onveilig wordt genomen ten opzichte van de rest van de antwoordcategorieën. Dit kan komen omdat er maar weinig mensen in de steekproef aangeven zich onveilig te voelen. Vanuit de theorie is het logisch om een dummy te gebruiken waarbij heel onveilig en een beetje onveilig samen worden genomen, en een beetje veilig en heel veilig. Als mensen zich meestal veilig voelen maar niet altijd, kiezen ze waarschijnlijk voor de antwoordcategorie beetje veilig. Het is daarom niet logisch om deze categorie te stoppen bij de categorie waarin mensen aangeven zich onveilig te voelen. Ook is het voor de statistische resultaten logisch om de antwoordcategorieën samen te voegen omdat de groep 'heel onveilig' heel klein is; dan levert het niet echt waardevolle resultaten op. Daarom is het verantwoord om de binaire logistische regressie te doen waarbij de antwoordcategorieën heel onveilig en beetje onveilig te nemen tegenover beetje veilig en heel veilig.

In de binaire logistische regressie wordt gekeken naar de invloed van de verschillende factoren op de binaire variabele ervaren veiligheid. Hierbij is de waarde 0 'onveilig' en de waarde 1 'veilig'.

Tabel 3: resultaten model 1, model 2 en model 3 voor binaire logistische regressie

	Model 1			Model 2			Model 3		
	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p	b(SE)	Odds-ratio	p
Constante	5.911	369.050	<0.001	3.007	20.235	<0.001	5.811	333.967	<0.001
geslacht	-1.079	0.340	<0.001	-1.141	0.320	<0.001	-1.213	0.297	<0.001
Leeftijd	-0.008	0.992	0.061	-0.010	0.990	0.028	-0.015	0.985	0.002
% buren buitenlandse komaf	-0.427	0.653	<0.001	-0.337	0.714	<0.001	-0.236	0.790	<0.001
Sociale cohesie				0.871	2.390	<0.001	0.707	2.029	<0.001

Overlast					-1.276	0.279	<0.001
Deviance (-2LL)		1253.726		1167.916			1125.475
-toets (Chi^2)	208.065	<0.001	85.810	<0.001	42.441		<0.001
n		2644		2644			2644

Het eerste model laat het lege model zien met alleen de controlevariabelen. In dit model is te zien dat de hellingen van alle variabelen negatief zijn. Dit betekent dat een toename in leeftijd ($\beta = -0.008$), een hoger percentage buren met buitenlandse komaf in de buurt ($\beta = -0.427$) en vrouwen ($\beta = -1.079$), een hogere kans hebben dat ze zich minder veilig voelen in hun buurt.

Model 2 laat zien dat percentage buren met buitenlandse komaf minder invloed heeft wanneer de invloedsfactor sociale cohesie wordt toegevoegd ($\beta = -0.337$). De helling van sociale cohesie laat zien dat er een verandering is van 0.871 in de log-odds van ervaren veiligheid als sociale cohesie met 1 eenheid omhoog gaat. Een positieve helling houdt in dat een hogere sociale cohesie zorgt dat de kans dat respondenten zich meer veilig voelen in hun buurt toeneemt. De helling van leeftijd is nog steeds verwaarloosbaar ($\beta = -0.010$) en de invloed van geslacht is iets toegenomen bij het toevoegen van de variabele sociale cohesie ($\beta = -1.141$). Vrouwen hebben dus een grotere kans om zich onveilig te voelen.

Model 3 laat zien dat de toevoeging van overlast zorgt voor een minder groot effect van sociale cohesie op ervaren veiligheid ($\beta = 0.707$), er is dus bevestigd dat er sprake is van een mediatie. De odds-ratio laat wel zien dat er een verdubbeling van de ervaren veiligheid kan zijn als sociale cohesie met 1 eenheid toeneemt ($odds-ratio = 2.029$). De helling van overlast laat zien dat er daling in de log-odds van ervaren veiligheid is van 1.276 als overlast met 1 eenheid stijgt. Als respondenten dus meer overlast ervaren, voelen zij zich minder veilig in hun buurt.

De modellen ondersteunen de resultaten van de lineaire regressie, dit betekent dat de resultaten “robuust” zijn, ze blijven consistent en betrouwbaar ondanks het overschrijden van de assumpties.

Bijlage 3

Controle van assumpties

Lineaire regressie

In het onderzoek kijken we naar verschillende modelassumpties;

- De observaties zijn onafhankelijk, er is een aselecte steekproef geweest
- Er is een lineair verband
- Homoscedasticiteit; er is een constante residuele variantie
- De variantie van de errorterm is constant voor alle waarden van de x-en
- Residuen zijn normaal verdeeld

Nu kunnen we kijken of dit onderzoek voldoet aan de assumpties of dat deze geschonden zijn.

Uit het codeboek blijkt dat er voor het onderzoek een random selectie heeft plaatsgevonden van panelleden (ouder dan 18 jaar). De observaties zijn dus **onafhankelijk** van elkaar.

*Assumpties controleren, residual plot.

```
DATASET ACTIVATE DataSet1.
```

```
REGRESSION
```

```
  /MISSING LISTWISE
```

```
  /STATISTICS COEFF OUTS R ANOVA
```

```
  /CRITERIA=PIN(.05) POUT(.10)
```

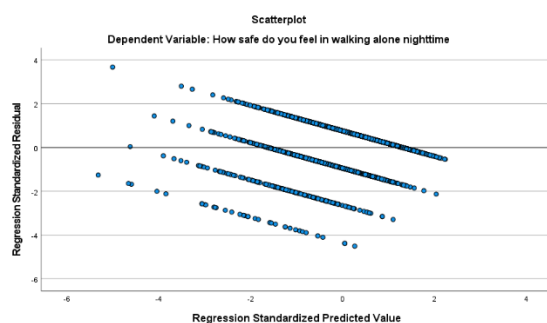
```
  /NOORIGIN
```

```
  /DEPENDENT sr20a007
```

```
  /METHOD=ENTER SocialeCohesie Overlast_schaal geslacht leeftijd sr20a019
```

```
  /SCATTERPLOT=(*ZRESID ,*ZPRED)
```

```
  /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID).
```



Figuur 1: residual plot

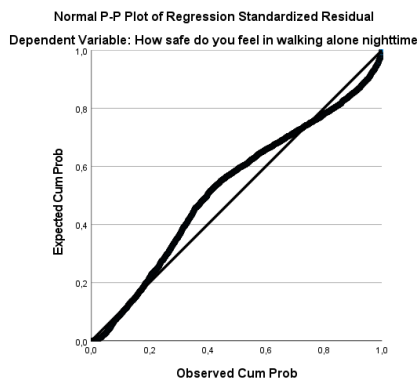
Om te zien of de data **lineair** is, mogen er geen systematische afwijkingen zijn van de nullijn in de residual plot. De residual plot is te zien in figuur 1. Deze assumptie is geschonden; er is een systematische afwijking van de nullijn omdat de variabele 'ervaren veiligheid' 4 antwoordcategorieën heeft. We gaan door met de lineaire regressie, maar om de reden dat er geen lineair verband is, zal ook

een logistische regressie gedaan worden. Daarmee kan gecontroleerd worden of de resultaten van de lineaire regressie kloppen.

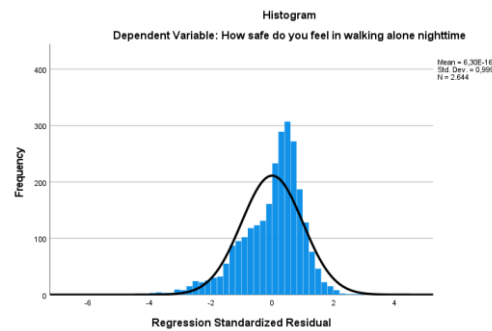
Er is geen sprake van **homoscedasticiteit**; er is in de residual plot in figuur 1 geen toename in de puntenwolk te zien.

Als laatste kijken we naar de normaliteit van de residuen. In figuur 3 zien we de verdeling van de residuen; deze is niet helemaal normaal maar wat rechtsscheef. Aan de linkerkant is te zien dat er wat meer cases zijn, aan de rechterkant zijn er juist minder cases. Het is daarom misschien verstandig om met een kleiner significantieniveau te testen (bijvoorbeeld 0.01 in plaats van 0.05). Ook in de p-p plot van figuur 2 is de scheefheid te zien. De assumptie lijkt dus geschonden maar wellicht is er een oplossing voor.

Doordat er een aantal schendingen zijn van assumpties, zal ook een ordinale logistische regressie gedaan worden om de conclusies te onderstrepen.



Figuur 2: P-P plot



Figuur 3: histogram residuen

Omdat de assumptie van normaliteit is geschonden, is het de moeite waard om logistische regressie te doen.

Ordinale logistische regressie

In tabel 1 zijn de resultaten voor ordinale logistische regressie te zien.

*Ordinale logistische regressie & test of parallel lines.

```
PLUM sr20a007 WITH SocialeCohesie Overlast_schaal geslacht leeftijd sr20a019  
/CRITERIA=CIN(95) DELTA(0) LCONVERGE(0) MXITER(100) MXSTEP(5)  
PCONVERGE(1.0E-6) SINGULAR(1.0E-8)  
/LINK=LOGIT  
/PRINT=FIT KERNEL PARAMETER SUMMARY TPARALLEL.
```

Tabel 1: output ordinale logistische regressie

		Parameter Estimates						
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[sr20a007 = 1]	-7,317	,442	274,232	1	<,001	-8,183	-6,451
	[sr20a007 = 2]	-5,207	,404	166,174	1	<,001	-5,999	-4,415
	[sr20a007 = 3]	-2,924	,391	55,815	1	<,001	-3,691	-2,157
Location	SocialeCohesie	,635	,059	116,331	1	<,001	,519	,750
	Overlast_schaal	-1,236	,124	99,599	1	<,001	-1,479	-,994
	geslacht	-1,144	,091	156,513	1	<,001	-1,323	-,964
	leeftijd	-,007	,003	7,910	1	,005	-,012	-,002
	sr20a019	-,181	,025	50,664	1	<,001	-,231	-,131

Link function: Logit.

Bij ordinale logistische regressie is de aanname dat de hellingen gelijk zijn voor alle constanten, per stap van de dichotomisaties waarbij eerst wordt gekeken van heel onveilig tegenover een beetje onveilig, een beetje veilig en heel veilig. Dan wordt er gekeken naar de dichotomisatie van heel onveilig en een beetje onveilig ten opzichte van een beetje veilig en heel veilig. Als laatste wordt gekeken naar de dichotomisatie van heel onveilig, een beetje onveilig en een beetje veilig ten opzichte van heel veilig. Om te controleren of de hellingen ook gelijk zijn bij elke constante van de dichotomisaties, doen we de test of parallel lines. De uitkomst van deze test is te zien in tabel 2:

Tabel 2: uitkomsten test of parallel lines

Test of Parallel Lines ^a				
Model	-2 Log Likelihood ^b	Chi-Square	df	Sig.
Null Hypothesis	3929,390			
General	3905,225	24,165	10	,007

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The kernel of the log-likelihood function is displayed.

In de test of parallel lines wordt gekeken of de nulhypothese dat de helling voor alle responsiecategorieën gelijk is, in de tabel is te zien dat de test significant is ($\alpha = 0.007$) dus de nulhypothese wordt verworpen. Dit kan komen omdat de steekproef groot is.

De hellingen zijn dus niet voor alle constanten van de dichotomisaties hetzelfde. Daarom is het goed om voor alle categorieën dummy's te maken en dan te kijken bij welke categorie de helling eventueel anders is.

Binaire logistische regressie

De hellingen zijn dus niet voor alle constanten van de dichotomisaties hetzelfde. Daarom is het goed om voor alle categorieën dummy's te maken en dan te kijken bij welke categorie de helling eventueel anders is.

*Binaire variabelen maken ervaren veiligheid.

```
RECODE sr20a007 (1=0) (ELSE=1) INTO ErvarenVeiligheid_1v234.
VARIABLE LABELS ErvarenVeiligheid_1v234 'Variabele 1: heel onveilig tov beetje onveilig,
beetje '+
'veilig en heel veilig'.
EXECUTE.
```

```
RECODE sr20a007 (1=0) (2=0) (3=1) (4=1) INTO ErvarenVeiligheid_12v34.
VARIABLE LABELS ErvarenVeiligheid_12v34 'Variabele 2: heel onveilig, beetje onveilig tov
beetje '+
'veilig en heel veilig'.
EXECUTE.
```

```
RECODE sr20a007 (1=0) (2=0) (3=0) (4=1) INTO ErvarenVeiligheid_123v4.
VARIABLE LABELS ErvarenVeiligheid_123v4 "'Variabele 3: heel onveilig, beetje onveilig, beetje
"+
'veilig tov heel veilig"'.
EXECUTE.
```

*Binaire logistische met binaire variabele 1-234.

```
LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_1v234
/METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

*Binaire logistische regressie met binaire variabele 12-34.

```
LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_12v34
/METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

*Binaire logistische regressie met binaire variabele 123-4.

```
LOGISTIC REGRESSION VARIABLES ErvarenVeiligheid_123v4
/METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal
/CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).
```

Tabel 3: binaire logistische regressie 'heel onveilig' tov de andere variabelen

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-,471	,385	1,496	1	,221	,624
	Age of the household member	-,020	,011	3,121	1	,077	,980
	What do you estimate, what percentage of the residents of your neighborhood are	-,326	,087	14,129	1	<,001	,722
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,800	,216	13,677	1	<,001	2,225
	Overlast_schaal	-1,301	,368	12,492	1	<,001	,272
	Constant	7,236	1,521	22,637	1	<,001	1387,983

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Tabel 4: binaire logistische regressie 'heel onveilig en 'beetje onveilig' tov 'beetje veilig' en 'heel veilig'

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-1,213	,181	44,977	1	<,001	,297
	Age of the household member	-,015	,005	9,407	1	,002	,985
	What do you estimate, what percentage of the residents of your neighborhood are	-,236	,041	33,570	1	<,001	,790
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,707	,102	48,323	1	<,001	2,029
	Overlast_schaal	-1,276	,198	41,627	1	<,001	,279
	Constant	5,811	,713	66,469	1	<,001	333,967

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Tabel 5: binaire logistische regressie 'heel onveilig', 'beetje onveilig' en 'beetje veilig' tov 'heel veilig'

		Variables in the Equation					
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Gender	-1,122	,095	139,642	1	<,001	,326
	Age of the household member	-,006	,003	4,299	1	,038	,994
	What do you estimate, what percentage of the residents of your neighborhood are	-,160	,028	33,260	1	<,001	,852
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,606	,063	92,243	1	<,001	1,833
	Overlast_schaal	-1,139	,140	66,176	1	<,001	,320
	Constant	2,711	,420	41,731	1	<,001	15,046

a. Variable(s) entered on step 1: Gender, Age of the household member, What do you estimate, what percentage of the residents of your neighborhood are, Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen, Overlast_schaal.

Na het maken van de dummy's en binaire logistische regressie uit te voeren voor elk model, is te zien dat de hellingen vooral verschillen wanneer de antwoordcategorie heel onveilig wordt genomen ten opzichte van de rest van de antwoordcategorieën. Dit kan komen omdat er maar weinig mensen in de steekproef aangeven zich onveilig te voelen. Vanuit de theorie is het logisch om een dummy te gebruiken waarbij heel onveilig en een beetje onveilig samen worden genomen, en een beetje veilig en heel veilig. Als mensen zich meestal veilig voelen maar niet altijd, kiezen ze waarschijnlijk voor de antwoordcategorie beetje veilig. Het is daarom niet logisch om deze categorie te stoppen bij de categorie waarin mensen aangeven zich onveilig te voelen. Ook is het voor de statistische resultaten logisch om de antwoordcategorieën samen te voegen omdat de groep 'heel onveilig' heel klein is; dan levert het niet echt waardevolle resultaten op. Daarom is het verantwoord om de binaire logistische regressie te doen waarbij de antwoordcategorieën heel onveilig en beetje onveilig te nemen tegenover beetje veilig en heel veilig.

Controle van outliers

*Case ID variabele aanmaken.

```
DATASET ACTIVATE DataSet1.  
COMPUTE ID=$CASENUM.  
VARIABLE LABELS ID 'Casenumbers'.  
EXECUTE.
```

*Cook's distance, DFFit en DFbeta.

```
REGRESSION  
  /DESCRIPTIVES MEAN STDDEV CORR SIG N  
  /MISSING LISTWISE  
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE  
  /CRITERIA=PIN(.05) POUT(.10)  
  /NOORIGIN  
  /DEPENDENT sr20a007  
  /METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal  
  /SAVE COOK DFBETA DFFIT.
```

*Uitbijters controleren.

```
SORT CASES BY ZRE_1(A).
```

```
SORT CASES BY COO_1(A).
```

```
DESCRIPTIVES VARIABLES=DFB_1 DFB0_1 DFB1_1 DFB2_1 DFB3_1 DFB4_1 DFB5_1  
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Voor de voorspelling is het belangrijk om te controleren of er uitbijters zijn en of deze invloed zouden hebben op het model. Een uitbijter kan zorgen voor vertekende resultaten. Om te controleren of er uitbijters zijn, kunnen we verschillende methoden inzetten. Als eerste wordt gekeken naar de residuen. Hiermee kunnen we kijken of er in de y-richting uitbijters zijn; normale punten liggen tussen de waarde -3 en 3. Wanneer we de data sorteren en punten tegenkomen die boven deze waarden liggen, zouden dit uitbijters kunnen zijn. De cases 6 (gestandaardiseerde residu = -4.509), 1272 (gestandaardiseerde residu = -4.3840), 7 (gestandaardiseerde residu = -4.385), 1273 (gestandaardiseerde residu = -4.108), 8 (gestandaardiseerde residu = -4.050), 1274 (gestandaardiseerde residu = -3.902) en 2661 (gestandaardiseerde residu = 3.674) hebben de meest extreme waarden en zouden dus uitbijters kunnen zijn.

Ook kunnen we kijken naar de uitbijters met *Cook's distance*. Hierbij gelden de vuistregels dat een waarde van meer dan 1 bij Cook's distance echt een probleem is, en dat bij een waarde van 0.0015 ($4/n = 4/2644$) een mogelijk probleem zou kunnen bestaan. Op basis daarvan zouden meer dan 150 cases problematisch kunnen zijn. De cases met de hoogste Cook's distance zijn 2661, 1281, 1259, 1276, 8, 1265, 6, 22, 1274 en 1305. Maar de waarden van de Cook's distance bij deze cases komt bij

lange na niet bij de 1, de hoogste waarde hoort bij case 2661 en die heeft een Cook's distance van 0.03654.

De cases die zowel naar voren komen bij het controleren met de residuen als bij Cook's distance zijn de nummers 6, 8, 1274, 1276, 1281 en 2661.

De DFBETA kan ons ook meer duidelijkheid geven over uitbijters. We kijken met de DFBETA naar het effect op een geschatte parameter wanneer een bepaalde observatie wordt weggelaten. Waarden die erg afwijken van andere scores kunnen uitbijters zijn. We kunnen dan dus kijken naar de verandering in regressiecoëfficiënten. Hierbij geldt een vuistregel dat een waarde hoger dan $0.058 (3/\sqrt{n} = 3/\sqrt{2644})$ een uitbijter aangeeft.

Tabel 6: descriptieve output voor DFFIT en DFBETA

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
DFFIT	2644	-,01964	,03446	-,0000255	,00215912
DFBETA Intercept	2644	-,01967	,02419	-,0000003	,00222923
DFBETA geslacht	2644	-,00207	,00246	,0000000	,00044343
DFBETA leeftijd	2644	-,00012	,00011	,0000000	,00001401
DFBETA sr20a019	2644	-,00137	,00142	,0000000	,00017370
DFBETA SocialeCohesie	2644	-,00333	,00268	,0000000	,00034230
DFBETA Overlast_schaal	2644	-,00958	,01342	,0000000	,00084958
Valid N (listwise)	2644				

Uit de SPSS output kunnen we halen dat bij de DFBETA geen uitbijters te vinden zijn, alle maximum waarden zijn kleiner dan 0.058.

Als laatste kunnen we kijken naar de DFFIT, Hierbij wordt gekeken naar de verandering in de voorspelde waarde van welzijn uit de regressielijn. Er wordt gekeken naar het effect op de fit als bepaalde waarden worden weggelaten. Hierbij geldt dat een waarde hoger dan $0.087 (2 * \sqrt{p/n} = 2 * \sqrt{5/2644})$ of hoger dan $0.130 (3 * \sqrt{p/n} = 3 * \sqrt{5/2644})$ een uitbijter aangeeft. Nu is te zien in de output van SPSS dat de maximum waarde van de DFFIT 0.345 is en dus beneden de waarde van de vuistregels ligt. Dit houdt in dat er geen uitbijters gevonden zijn als we kijken naar de DFFIT.

Na het weghalen van de cases die zowel bij het controleren met residuen als Cook's distance terugkwamen, kunnen we opnieuw de lineaire regressie uitvoeren. Het resultaat van het uiteindelijke model is te zien in tabel 7:

*Uitbijters verwijderen.

USE ALL.

COMPUTE filter_\$(ID ~= 6 & ID ~= 8 & ID ~= 1274 & ID ~= 1276 & ID ~= 1281 & ID ~= 2661).

VARIABLE LABELS filter_\$ 'ID ~= 6 & ID ~= 8 & ID ~= 1274 & ID ~= 1276 & ID ~= 1281 & ID ~= 2661 '+'

'(FILTER)'.

VALUE LABELS filter_\$ 0 'Not Selected' 1 'Selected'.

FORMATS filter_\$ (f1.0).
 FILTER BY filter_\$.
 EXECUTE.

*Lineaire regressie na verwijderen uitbijters.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE
 /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
 /CRITERIA=PIN(.05) POUT(.10)
 /NOORIGIN
 /DEPENDENT sr20a007
 /METHOD=ENTER geslacht leeftijd sr20a019 SocialeCohesie Overlast_schaal.

Tabel 7: lineaire regressie na het weghalen van de uitbijters

		Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	4,235	,104		40,829	<,001		
	Gender	-,283	,023	-,211	-12,494	<,001	,987	1,013
	Age of the household member	-,002	,001	-,060	-3,436	<,001	,938	1,066
	What do you estimate, what percentage of the residents of your neighborhood are	-,058	,007	-,162	-8,096	<,001	,705	1,419
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,170	,015	,207	11,096	<,001	,809	1,236
	Overlast_schaal	-,384	,034	-,228	-11,192	<,001	,678	1,475

a. Dependent Variable: How safe do you feel in walking alone nighttime

In de tabel is te zien dat er nauwelijks verandering is in de hellingen. De veranderingen zijn zo klein dat ze verwaarloosbaar zijn. Ook zijn er bij het controleren met de DFFIT en DFBETA geen uitbijters gevonden. Het is voor dit onderzoek daarom niet nodig om uitbijters weg te laten in het onderzoek.

Controle van multicollineariteit

Multicollineariteit treedt op als er een sterk lineair verband is tussen verklarende variabelen. Als er een sterk lineair verband is, kan het regressiemodel slechter worden geschat omdat verklarende variabelen elkaar voorspellen. Dan wordt er dus geen extra variantie verklaard in het model, daarom is het goed om te controleren voor multicollineariteit. Het toevoegen van steeds meer variabelen is dus niet per definitie beter voor de verklaarde variantie.

*Multicollineariteit controleren.

DATASET ACTIVATE DataSet1.

REGRESSION

/DESCRIPTIVES MEAN STDDEV CORR SIG N
 /MISSING LISTWISE

```

/STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT sr20a007
/METHOD=ENTER geslacht leeftijd sr20a019
/METHOD=ENTER SocialeCohesie
/METHOD=ENTER Overlast_schaal.
DATASET ACTIVATE DataSet1.

```

Tabel 8: coefficients tabel voor lineaire regressie met VIF

Model		Unstandardized Coefficients		Standardized	t	Sig.	Collinearity Statistics	
		B	Std. Error	Coefficients Beta			Tolerance	VIF
1	(Constant)	4,312	,062		69,494	,000		
	Gender	-,269	,024	-,198	-11,035	<,001	,988	1,012
	Age of the household member	-,001	,001	-,036	-1,953	,051	,942	1,062
	What do you estimate, what percentage of the residents of your neighborhood are	-,123	,007	-,341	-18,601	<,001	,949	1,053
2	(Constant)	3,486	,083		42,172	<,001		
	Gender	-,278	,024	-,205	-11,829	<,001	,987	1,013
	Age of the household member	-,002	,001	-,042	-2,388	,017	,941	1,062
	What do you estimate, what percentage of the residents of your neighborhood are	-,091	,007	-,253	-13,555	<,001	,848	1,179
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,220	,015	,264	14,444	<,001	,885	1,129
3	(Constant)	4,219	,105		40,216	<,001		
	Gender	-,281	,023	-,207	-12,235	<,001	,987	1,013
	Age of the household member	-,002	,001	-,055	-3,180	,001	,937	1,067
	What do you estimate, what percentage of the residents of your neighborhood are	-,059	,007	-,163	-8,145	<,001	,706	1,417
	Mate van sociale cohesie door vertrouwen, hulp en elkaar kennen	,169	,016	,204	10,886	<,001	,809	1,236
	Overlast_schaal	-,379	,035	-,224	-10,967	<,001	,677	1,476

In tabel 8 zijn verschillende coëfficiënten gepresenteerd, waaronder de Variation Inflation Factor (VIF). De VIF laat zien of er sprake is van problematische multicollineariteit. Er zijn verschillende vuistregels voor de VIF, als waarden groter zijn dan 10, is er sprake van zeer problematische multicollineariteit en dan moeten de variabelen bewerkt worden. Een andere vuistregel is dat de waarden niet groter mogen zijn dan 4. Dit zou kunnen duiden op significante multicollineariteit. En soms wordt de vuistregel met de waarde 2 gebruikt, variabelen met een VIF hoger dan 2 zijn dan problematisch. In tabel 8 is te zien dat alle variabelen in alle modellen onder 2 blijven, en dus is er geen sprake van multicollineariteit in de modellen.

Bijlage 4

Deze bijlage gaat in op het gebruik van AI in dit onderzoek. Tijdens het onderzoek is AI alleen gebruikt voor het helpen begrijpen van modellen en tabellen in de resultatensectie. Ik vroeg vragen aan een AI programma om te zorgen dat ik de modellen op de juiste manier zou interpreteren. Daarnaast heb ik AI ook gebruikt voor spellingscontrole. AI gaf dan bijvoorbeeld aan waar een woord met een hoofdletter moest beginnen of dat er een spatie miste. Voor andere doeleinden heb ik AI niet gebruikt, ik heb alle stukken zelf geschreven en opgebouwd in mijn eigen woorden.