

Waarom is niet iedereen crimineel?

Over de rol van eenzaamheid binnen het verband tussen sociale cohesie en crimineel gedrag

Marieke Schelhaas

m.e.schelhaas@student.rug.nl

S4132734

Jaap Nieuwenhuis

8 Juni 2022

Inhoudsopgave

Inhoudsopgave	3
Samenvatting	7
Inleiding	8
Theorie	12
Sociale cohesie & Criminaliteit op wijkniveau	12
Delinquentie in de eigen wijk	14
Sociale cohesie in de wijk en eenzaamheid	15
Eenzaamheid en crimineel gedrag	16
Probleemgedrag	18
Controlevariabelen	19
Methoden	21
Beschrijving van de dataset	21
Operationalisaties	22
Analyseplan	27
Resultaten	29
Beschrijvende statistieken	29
Onderlinge correlaties	30
Modevaluatie	33
Hypothesetoetsing	36
Conclusie & Discussie	46
Conclusie	46
Discussie	47
Literatuur	50
Bijlage 1	55
Selectie niet-Missing variabelen	55
Sociale Cohesie	56
Eenzaamheid	62
Probleemgedrag	69
Crimineel gedrag	75
Leeftijd	82
Geslacht	84
Werkend	85
Inkomen	86

Opleiding	88
Partner	95
Werkend	97
Bijlage 2	99
Regressie-analyse 1: Eenzaamheid	99
Regressie-analyse 2: Probleemgedrag	102
Regressie-analyse 3: Crimineel Gedrag	105
Bijlage 3	109
Assumptietoetsing lineaire regressie	109
Missing-data analyse	113
Outliers	116
Multi-collineariteit	130

Samenvatting

In de sociologische literatuur is de samenhang tussen een hoge sociale cohesie en lage criminaliteitscijfers al langere tijd bekend. Lange tijd werd dit verband gezocht in een hoger sociaal vertrouwen en grotere sociale controle. Dit onderzoek stelt echter een mediërend verband voor, gebaseerd op eenzaamheid. Wanneer een individu crimineel gedrag vertoont, kan er een stigma op deze persoon komen te rusten en kunnen er zware sociale sancties dreigen. Wanneer er een persoon eenzaam is, zijn de sociale sancties voor deze persoon minder groot en kan deze dus relatief 'goedkoop' de voordelen van crimineel gedrag behalen. Om dit verband extra kracht bij te zetten, wordt er tevens ook gekeken naar probleemgedrag: gedrag dat vaak samengaat met of de voorloper kan zijn van crimineel gedrag.

Het verband tussen een hoge sociale cohesie en lage criminaliteitscijfers is in dit verslag opnieuw aangetoond. Ook kon het verband tussen probleemgedrag en sociale cohesie vast worden gesteld. Bovendien werd er een statistisch verband gevonden tussen een hoge sociale cohesie en minder eenzaamheid. Er was daarentegen geen verband te vinden tussen crimineel gedrag en eenzaamheid. Het verwachte mediatie-effect kon dus niet statistisch worden ondersteund. Ook kon er niet gevonden worden dat een individu vaker probleemgedrag vertoont, wanneer die zich eenzaam voelt. De theorie, methoden en resultaten die horen bij deze resultaten zijn in meer detail te vinden in het hierna volgende verslag.

Inleiding

Waarom is eigenlijk niet iedereen crimineel? Binnen de sociologie wordt deze vraag een tikkeltje anders gesteld dan normaalgesproken in de samenleving . Sociologen vragen zich immers af hoe het toch kan dat een paar individuen juist wel crimineel gedrag vertonen. Het antwoord zit hem in het feit dat crimineel gedrag veel voordelen kan hebben voor een individu. Fraude en oplichting kan veel rijkdom voortbrengen, het plegen van wraak (door bijvoorbeeld geweld toe te passen) kan voor macht en een gevoel van gerechtigheid zorgen. Als criminaliteit dus zoveel voordelen voor het individu kan hebben, hoe kan het dan dat eigenlijk niet iedereen crimineel is?

Sociologen zoeken het antwoord in de sociale verbindingen die mensen aangaan. Namelijk, volgens de keuzetheorie van Herrnstein wegen mensen de voor- en nadelen tegen elkaar af voordat ze een beslissing maken (Akçomak en Ter Weel, 2012). Zo kan criminaliteit veel materiële voordelen hebben, maar sociaal gezien zijn hier juist veel nadelen. Er kan een stigma op het individu gaan liggen, en belangrijke connecties zouden deze persoon af kunnen wijzen naar aanleiding van de straf die deze krijgt. Mensen willen zich houden aan de normen en waarden die gelden binnen een samenleving. Crimineel gedrag is een schending van deze geldende normen, waarna vaak sancties volgen. Crimineel gedrag kan dus zorgen voor sancties, die dan weer schaamte of schuld ten gevolg zouden kunnen hebben. Deze negatieve sociale effecten die crimineel gedrag met zich meebrengen, zouden nog belangrijker zijn voor het individu, dan de sancties zelf (Nagin, 1998). Criminaliteit en crimineel gedrag zullen worden gedefinieerd als “verzamelbegrip voor alle gedragingen, acties, handelingen en-of gebeurtenissen die volgens de wet strafbaar zijn” (Encyclo - Thesaurus Politiekunde, 2012). Het is tevens goed om als kanttekening te benoemen dat criminaliteit contextafhankelijk is van de

regio of cultuur die ergens geldt. Het roken van wiet wordt bijvoorbeeld in Nederland niet als crimineel gedrag gezien, maar in veel Afrikaanse en Aziatische landen wel.

Niet alleen op individueel niveau zijn sociale mechanismen van invloed op crimineel gedrag. Binnen wijken zijn er ook vele verschillende sociale mechanismen die invloed hebben op criminaliteit. Zo doen sociaal vertrouwen en maatschappelijke betrokkenheid criminaliteit afnemen, omdat het de effectiviteit van sociale normen verhoogt (Rosenfeld, Messner en Baumer, 2001). Stuart & Taylor (2021) beargumenteren juist dat hoe meer je ingebed bent in een samenleving, hoe groter de kans dat je misdaad en identiteit ontdekt zullen worden. Hoe groter deze kans, hoe groter weer de kans op een veroordeling. Bovendien, zoals eerder vermeld, kan het stigma dat om deze veroordeling hangt grote sociale kosten met zich meebrengen voor het individu. Zo houden deze sociale mechanismen de normen van de gemeenschap in bedwang.

Al deze sociale mechanismen lijken dus fundamenteel voor het behoud van deze normen. Wat gebeurt er dan wanneer deze sociale mechanismen geen invloed meer hebben op het individu? Wat als men bijvoorbeeld buiten deze sociale mechanismen staat omdat men eenzaam is? In dit onderzoek zal eenzaamheid worden gedefinieerd als “het subjectief ervaren van een onplezierig of ontoelaatbaar gemis aan (kwaliteit van) bepaalde sociale relaties.” Hierbij wordt als kanttekening toegevoegd: “het kan zijn dat het aantal contacten dat men heeft met andere mensen geringer is dan men wenst. Het kan ook zijn dat de kwaliteit van de gerealiseerde relaties achterblijft bij de wensen.” (Een tegen Eenzaamheid, 2021).

Eenzaamheid is een groot probleem in de Nederlandse samenleving. Maar liefst 9 procent van de Nederlanders ouder dan 15 voelde in 2019 sterke gevoelens van eenzaamheid. 26 procent van de bevolking voelde zich enigszins eenzaam, en de overige 66 procent voelde zich niet eenzaam (Centraal Bureau voor de Statistiek, 2020). In het actieprogramma “Eén tegen

eenzaamheid” wordt zelfs genoemd dat maar liefst 43 procent van de volwassen Nederlanders eenzaamheid ervaart (Ministerie van Volksgezondheid, Welzijn en Sport, 2018). In dit rapport wordt bevestigd dat deze individuen weinig sociale connecties of alleen maar zwakke vriendschappen hebben, en niemand hebben om naar toe te gaan als deze problemen hebben.

Het Rijksinstituut voor Volksgezondheid en Milieu (2012) voorspelt dat het probleem van eenzaamheid alleen maar toe gaat nemen. Het instituut voorspelt dat het aantal eenzamen van 5,2 miljoen in 2015, naar bijna 5,9 miljoen toe zal nemen in 2040. Deze toename zou vooral te wijden zijn aan de vergrijzing, maar ook aan de toename van het aantal eenpersoonshuishoudens.

Alleen al wanneer men kijkt naar de gezondheid van alle betrokkenen, wordt het belang van het tegengaan van eenzaamheid al duidelijk. Eenzaamheid vergroot de kans op hartziekten, dementie, verhoogt de bloeddruk en het stressniveau (Verrycken, 2020). Er ligt echter mogelijk ook een sociaal-maatschappelijk probleem op de loer. Zoals eerder al vermeld, zijn sociale mechanismen en het risico op sociale sancties fundamenteel voor het behoud van normen. Wanneer men deze eenzame individuen onder de loep neemt, zijn deze sociale mechanismen minder op hen van invloed en zijn de sociale sancties minder zwaar. Voor de eenzamen zijn hierdoor de sociale kosten van misdaad een stuk minder hoog. Het zijn juist de eenzamen die in staat zijn om relatief goedkoop veel financiële voordelen te behalen door middel van crimineel gedrag. Wanneer deze redenering zou kloppen zou in een steeds eenzamer wordende samenleving, het risico bestaan dat crimineel gedrag ook steeds meer zou toenemen.

Vanuit deze vragen rijst de volgende onderzoeksvraag:

“Wat is de rol van die eenzaamheid speelt in het verband tussen sociale cohesie en crimineel gedrag?”

Wanneer eenzaamheid en isolatie naast de mentale en fysieke schaduwzijde, nog een ongewenst effect veroorzaakt (namelijk extra criminele activiteit) dan zijn er waarschijnlijk meer belangstellenden die interesse zouden hebben in het oplossen van dit probleem. Eenzaamheid zal niet langer als een individuele tragedie worden beschouwd, maar als een maatschappelijk probleem dat voor het algemene belang moet worden opgelost. Dit kan extra nadruk leggen op het oplossen van eenzaamheid en een rol spelen bij het vrijmaken van budgetten voor bijvoorbeeld preventie van eenzaamheid.

Theorie

In deze scriptie zal worden onderzocht of een lage sociale cohesie de oorzaak is van criminaliteit of dat het verband misschien anders ligt dan gedacht. De verwachting is namelijk dat niet de afwezigheid van sociale cohesie, maar de eenzaamheid die hieruit voortkomt de oorzaak is van crimineel gedrag. Hieronder zal uitgelegd worden waarom ik vermoed dat crimineel gedrag direct in verband staat met eenzaamheid, met de afwezigheid van sociale cohesie als oorzaak van eenzaamheid. Dit in tegenstelling tot het hiervoor geldende beeld, dat de afwezigheid van sociale cohesie de oorzaak is van crimineel gedrag.

Sociale cohesie & Criminaliteit op wijkniveau

Zoals in de inleiding al is genoemd, zijn er veel sociale mechanismen die invloed hebben op criminaliteit. Een van de mechanismen die het meest effectief wordt gezien in de preventie van criminaliteit, is sociale cohesie. De definitie die ik voor sociale cohesie zal gebruiken komt uit het onderzoek van Cramm et al. (2012): "Sociale cohesie verwijst naar de hoeveelheid betrokkenheid en solidariteit die er is tussen individuen en groepen in een bepaalde gemeenschap heerst". Er wordt steeds opnieuw bewezen dat de criminaliteitscijfers in de wijk lager zijn, wanneer er veel sociale cohesie aanwezig is binnen een wijk. Rosenfeld, Messner en Baumer (2001) stellen dat maatschappelijke betrokkenheid en sociaal vertrouwen criminaliteit doen afnemen, omdat dit zowel informeel als formeel vertrouwen verhoogt. Een hoog sociaal vertrouwen houdt in dat men een positief mensbeeld heeft en op deze manier vertrouwen heeft in de mensen en de instanties om zich heen.

Volgens Steenbeek en Hipp (2011) is informele sociale controle en het geven van sancties onmisbaar voor het verklaren van een lage desorganisatie in een wijk. Door dit informele vertrouwen in elkaar, maar ook door onderlinge sociale controle, worden de kosten van

conflicthantering minder hoog en is er ook minder sprake van geweldsconflicten (Akçomak en Ter Weel, 2012). Het idee achter deze informele controle wordt verder uitgelegd door Stuart & Taylor (2021). Zij stelden dat hoe meer men is ingebed in een samenleving, hoe groter de kans dat je misdaad en identiteit ontdekt wordt. Dit wordt ondersteund door Akçomak en Ter Weel (2012). Zij beargumenteren dat wanneer iemands sociaal kapitaal hoger is, men meer te verliezen hebt. Sociaal kapitaal is een subset van sociale cohesie en wordt veel teruggevonden in een betrokken maatschappij. Sociaal kapitaal is de sociale relaties met de mensen die men kent, die men kan gebruiken als bron voor verschillende voordelen (Savage & Kanazawa, 2002). Wanneer deze dus hoog is, heeft men meer sociale relaties en heeft men daardoor dus meer te verliezen. Een veroordeling kan zorgen voor een stigma, wat ervoor kan zorgen dat belangrijke sociale relaties niets meer met je te maken willen hebben (Nagin, 1998). Volgens de keuzetheorie van Herrnstein worden deze materiële voordelen en sociale nadelen zorgvuldig overwogen, voordat er een beslissing wordt gemaakt. Hoe hoger je sociale kapitaal, hoe groter de kans dat criminele activiteiten je sociaal gezien te veel kosten en je deze activiteiten dus niet pleegt.

Maatschappelijke betrokkenheid en sociaal vertrouwen hebben nog een extra voordeel. Naast de toename van sociale controle, verhogen ze de effectiviteit van sociale normen (Rosenfeld, Messner, Baumer, 2001). Criminaliteit is, zoals eerder genoemd, een schending van een geldende norm. Zowel door een hogere effectiviteit van sociale normen, als door een betere conflicthantering, neemt criminaliteit dus af in wijken met een hoge sociale cohesie.

In het onderzoek van Sampson en Groves (1989) wordt bewezen dat in wijken met weinig vriendschapsnetwerken en weinig participatie van de bewoners, disproportioneel meer criminaliteit en delinquentie was.

Hierbij is het belangrijk om wel één kanttekening te zetten over de aanname van oorzaak en gevolg. Zoals gezegd, wordt er steeds opnieuw bewijs gevonden dat de criminaliteitscijfers in

deze wijk lager zijn, wanneer er veel sociale cohesie aanwezig is binnen een wijk. Dit kan daarentegen natuurlijk ook betekenen dat het heel lastig is om sociale cohesie te vormen, wanneer er veel criminaliteit in een wijk plaatsvindt. Wanneer dit het geval zou zijn, ligt de oorzaak-gevolg relatie dus feitelijk andersom. Zolang er geen experimenteel onderzoek is gedaan naar de causatie van het verband, kan hier nog weinig uitsluitel over worden gegeven. Het wetenschappelijk debat heeft echter veelal aangenomen dat de relatie hoogstwaarschijnlijk in de richting 'hoge sociale cohesie veroorzaakt weinig criminaliteit' ligt.

Delinquentie in de eigen wijk

Een grote betrokkenheid en solidariteit in een maatschappij zijn dus belangrijke oorzaken in het voorkomen van criminaliteit binnen de wijk. Het is daarentegen alsnog goed mogelijk dat de inwoners die zelf wel crimineel gedrag vertonen, naar een wijk gaan met een minder hoge sociale cohesie om hier hun criminele activiteiten te voltrekken. Op deze manier zouden ze wel de voordelen van crimineel gedrag kunnen verkrijgen, maar een veroordeling en het bijbehorende sociaal stigma kunnen ontlopen. Vanuit hieruit volgt de vraag: wanneer er sociale cohesie aanwezig is in een wijk, zijn individuen die deze wijk bewonen dan ook minder crimineel? Volgens het onderzoek van Bernasco et al. (2012) reizen de meeste criminelen niet ver om een delict te plegen. Een van de meest belangrijke factoren in het kiezen van een locatie is, of deze locatie niet te ver verwijderd is van het huis van de dader. Daders plegen hun criminele activiteiten vaak in de eigen leefomgeving en dus waarschijnlijk ook in de eigen wijk. Uit het onderzoek van Lammers et al. (2015) wordt bovendien gevonden dat een delinquent een locatie bij een volgend delict sneller uitgekozen wordt, wanneer deze delinquent hier eenmaal eerder een delict gepleegd heeft. Ook kiest een dader een locatie eerder uit wanneer deze persoon hier in het verleden een delict gepleegd heeft. Dit betekent dus dat daders vaak in hetzelfde gebied actief zijn. Wanneer iemand ervoor kiest om criminele activiteiten te begaan in de eigen wijk, dan zal deze de eigen woonwijk dus vaker als locatie kiezen.

Vanuit deze twee voorgaande paragraaf, vloeit de eerste hypothese van deze scriptie:

Hypothese 1: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, vertoont dit individu minder crimineel gedrag.

Sociale cohesie in de wijk en eenzaamheid

Wanneer er een hogere hoeveelheid betrokkenheid en solidariteit is tussen individuen in de wijk, verwacht je dat individuen binnen deze wijk ook minder eenzaamheid ervaren. Het onderzoek van Yu et al. (2020) stelt dat door het verbeteren van gevoelens van verbondenheid met buren, eenzaamheid bij ouderen kan verminderen. Ook heeft sociale cohesie invloed op welke manier je je eenzaam voelt. Wanneer je status in de wijk hoog is, dan voel je je sneller emotioneel eenzaam. Wanneer je status in de wijk lager is, voel je je sneller sociaal eenzaam. Het verschil is hier dat je bij emotionele eenzaamheid je je niet verbonden voelt met je omgeving, ondanks een mogelijk hoog aantal sociale contacten. Bij sociale eenzaamheid voel je je eenzaam, omdat je een klein sociaal netwerk hebt. Er is dus in de literatuur al een aanwijzing dat sociale cohesie in de wijk individuele eenzaamheid kan doen verminderen. Het omgekeerde kan natuurlijk ook het geval zijn: wanneer er weinig sociale cohesie in de wijk aanwezig is, is de individuele eenzaamheid groter.

Om dit meer te generaliseren naar de gehele populatie in plaats van op alleen ouderen, is de tweede hypothese van deze scriptie opgesteld:

Hypothese 2: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, zal dit individu zich minder eenzaam voelen.

Eenzaamheid en crimineel gedrag

Vanuit de eerste paragraaf wordt beargumenteerd dat sociale cohesie belangrijk is voor het tegengaan van crimineel gedrag. Er zijn veel sociale mechanismen die ervoor zorgen dat bij zorgvuldige overweging van de voor- en nadelen, criminaliteit te nadelig wordt. Wanneer je gezicht bekend is binnen de samenleving, hoe groter je identificatie en hoe groter je pakkans wanneer je normen overtreedt (Stuart & Taylor, 2021). Het schenden van normen zou voor sociale sancties kunnen zorgen, waardoor je belangrijke sociale connecties/je sociale kapitaal kwijt kan raken (Nagin, 1998). Informele sociale controle en het geven van sancties voor het niet volgen van normen worden ook gezien als mechanismen van groot belang (Steenbeek en Hipp, 2011). Deze mechanismen hebben 1 ding gemeen: ze gaan ervan uit dat je sociale connecties hebt waarvan je bang bent om deze kwijt te raken. Wanneer je deze echter niet hebt, kun je deze ook niet kwijtraken. Wanneer je eenzaam bent, heb je weinig sociale connecties of zijn de sociale connecties die je wel hebt weinig waardevol. Op deze manier ben je dus vrij je doelen op criminele wijze te bereiken met relatief weinig sociale consequenties. Het is dus niet de sociale cohesie in de wijk die je aanzet tot crimineel gedrag, het is de eenzaamheid waarin men zich op dat moment begeeft.

Er is niet veel bekend over wat eenzaamheid doet met de normen en waarden van een individu. Hieronder zullen enige bronnen besproken worden die verandering van normen en waarden als gevolg van sociaal isolement aanhalen. Het onderzoek van Trammell et al., (2020) keek naar de invloed van administratieve segregatie (sociaal isolement als straf) op gedetineerden. Hier kwam naar voren dat de sociale cohesie sterk daalde en de wederkerigheid tussen gedetineerden afnam waar het verboden was om te netwerken en administratieve segregatie werd ingezet. Door een constante dreiging van isolatie was er weinig onderling vertrouwen en

ontstond er eenzaamheid en wanhoop. Uiteindelijk hadden de gedetineerden niks meer te verliezen, wat zich uitte in geweld tegenover elkaar en de medewerkers van de gevangenis. Wanneer je dit doortrekt naar geïsoleerde individuen binnen de samenleving, zou het dus kunnen dat eenzame mensen gefrustreerd raken en dit gaan uiten in geweld. Dit zou, in combinatie met een gevoel niets te verliezen te hebben, zich kunnen gaan uiten in crimineel gedrag.

Ook bij intergenerationale mobiliteit kan er sprake zijn van sociale isolatie. De individuen die de sociaal-economische status van hun achtergrond ontvluchten, passen niet meer bij hun oorspronkelijke sociaal-economische herkomst, maar ook niet volledig bij hun nieuwe status. Als resultaat voelt men zich geïsoleerd, gestresst en gefrustreerd. Deze isolatie zorgt voor een gevoel van onzekerheid, en deze individuen voelen zich niet gesteund vanuit de omgeving. Bovendien kan het leiden tot radicale politieke ideeën (bijvoorbeeld een minachting voor minderheden) en nalatigheid in het zijn van een “goede burger”. Dit gemis in burgerschap kan zich uiten in een gemis in politieke activiteit en solidariteit tot anderen. Ook kan het zich uiten in het niet navolgen van wetten en regels en een gebrek aan tolerantie en coöperatie. Het onderzoek wijst uit dat deze effecten van korte duur zijn, doordat mensen zich snel aanpassen aan de nieuwe situatie (Sieben, 2017). Sociaal isolement is daarentegen in de meeste gevallen niet van korte duur, zoals de isolatie die intergenerationale mobiliteit met zich meebrengt. Het zou goed kunnen dat sociaal isolement dezelfde effecten tot gevolg heeft, als in dit onderzoek worden beschreven.

Het is wel goed om aan te geven dat sociaal isolement en eenzaamheid aan elkaar verwant kunnen zijn, maar dat dit zeker niet altijd zo hoeft te zijn. Dit komt opnieuw neer op de eerder aangehaalde sociale en emotionele eenzaamheid. Het komt ook terug in het onderzoek van Vangelisti en Perlman (2006); een individu met een goede sociale positie kan zich overal op het

spectrum van eenzaamheid bevinden. Waar deze persoon zich uiteindelijk bevindt, hangt af van de normen die die persoon heeft over sociale relaties. Zo zijn er mensen die het hebben van weinig sociale relaties als een sociaal isolement zouden kunnen ervaren, maar ook mensen met hetzelfde aantal sociale relaties, die diep ongelukkig zijn.

Daaruit volgt dan ook de derde hypothese die in deze scriptie geanalyseerd zal worden:

Hypothese 3: Naarmate een individu zich eenzaam voelt, zal dit individu meer crimineel gedrag vertonen.

Probleemgedrag

Om de zeggingskracht van het model extra kracht bij te zetten, zal er ook gekeken worden naar probleemgedrag. Eén van de meest invloedrijke theorieën binnen de ontwikkelingscriminologie stelt dat persistente delinquenten hun gehele leven al patronen van probleemgedrag laten zien. Het gaat hier om probleemgedrag dat leeftijdsgebonden is en door het gehele leven heen escaleert (Dijkstra & Veenstra, 2005). Voorbeelden van probleemgedrag in kinderen zijn bijvoorbeeld spijbelen en weglopen van school. In volwassenen uit dit zich vaker in (overmatig) drank- en drugsgebruik (Burssens, 2019). Het verband tussen probleemgedrag, zoals bijvoorbeeld (problematisch) middelengebruik, en crimineel gedrag wordt ook buiten de ontwikkelingscriminologie als evident gezien. Het gebruik van drank en/of drugs zou onder andere bijdragen aan gewelddadig gedrag (Lammers et al., 2014). Ook rondhangen op straat wordt in verband gebracht met crimineel gedrag (Boutellier & van Marissing, 2011). Dit gedrag op zichzelf is vaak niet per definitie schadelijk of crimineel, maar komt over het algemeen vaker voor bij delinquenten. Bovendien stelt dit onderzoek dat probleemgedrag een voorloper van crimineel gedrag kan zijn. Juist omdat probleemgedrag zo consistent in verband wordt gebracht met crimineel gedrag, zou het kunnen zijn dat het veronderstelde verband tussen eenzaamheid

en criminaliteit, ook geldt voor probleemgedrag. Het verband tussen crimineel gedrag en eenzaamheid zou dus sterker onderbouwd kunnen worden als probleemgedrag eenzelfde positief verband vertoont.

Vanuit dit idee zijn de volgende hypothesen opgesteld:

Hypothese 4: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, vertoont dit individu minder probleemgedrag.

&

Hypothese 5: Naarmate een individu zich eenzaam voelt, zal dit individu meer probleemgedrag vertonen.

Controlevariabelen

Vanuit de theorie is er besloten om te controleren voor verschillende variabelen.

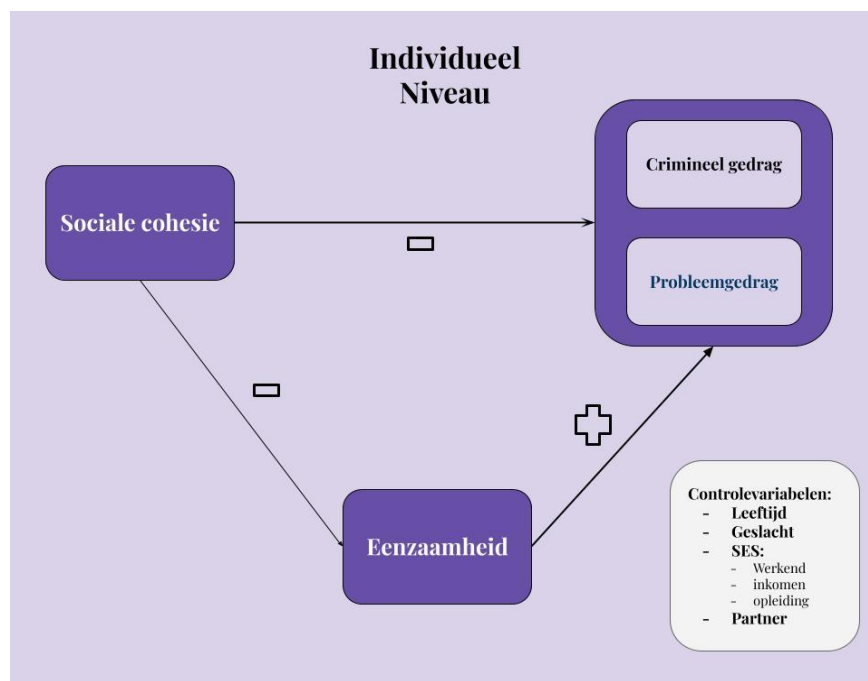
Wegens de developmental theory is er gekozen om te controleren voor leeftijd, omdat antisociaal gedrag zoals criminaliteit, vooral voorkomt bij jongeren en adolescenten (Moffitt, 1993). Bovendien is het bekend dat vrouwen minder vaak criminele activiteiten plegen dan mannen (Block et al., 2010). Vanuit deze kennis wil ik graag controleren voor leeftijd en geslacht.

Omdat de mate van sociaal-economische nadelen die samengaat met een lage sociaal-economische status, duidelijke associatie heeft met het al dan niet plegen van misdaad (Fergusson et al., 2004) wil ik hier ook graag voor controleren. Volgens Bureau of Justice Statistics U.S. Department of Justice et al. (2014) is sociaal-economische status een combinatie van inkomen, opleiding en het soort baan wat iemand heeft. Deze hebben allen hun imperfecties, waardoor ervoor is gekozen om voor alledrie te controleren. De dataset geeft geen gegevens over het soort baan van de respondent, waarom er als substituut gekeken wordt naar of de respondent een betaalde baan heeft. Dit is ook theoretisch te verantwoorden: wanneer je

een baan hebt, wil je deze niet verliezen. Bovendien wijst het Amerikaanse onderzoek van Wadsworth (2006) uit dat een zinvolle baan leidt tot een afname in gewelds- en eigendomsdelicten. Zinvol wordt hier gedefinieerd als een baan met vakantiedagen en pensioen, iets wat niet vanzelfsprekend is in Amerika. Deze data heeft wel als groot nadeel dat het uitblijven van een baan ook simpelweg kan betekenen dat iemand nog studeert of werkt als vrijwilliger.

Vanuit het meta-onderzoek van Skardhamar et al. (2015) wordt aangedragen dat delinquentie afneemt wanneer criminelen getrouwd zijn. Tegenwoordig zijn er echter veel mensen die wel een partner hebben, maar niet zijn getrouwd. Vangelisti en Perlman (2006) stellen dat iemand met een partner over het algemeen beter beschermd is tegen eenzaamheid dan iemand zonder partner. Wanneer je een partner hebt, heb je meer te verliezen dan dat je zonder partner hebt. Van deze redenering wordt ook voor partner gecontroleerd.

Alle 3 de hypothesen en de controle-variabelen zijn terug te vinden in het conceptueel model van het onderzoek, figuur 1.



Figuur 1: Conceptueel model / Visuele weergave van de hypothesen.

Methoden

In dit hoofdstuk zal ik uitleggen hoe de gegevens in elkaar steken waarmee in dit onderzoek is gewerkt. Eerst wordt er een beschrijving geven van de gebruikte dataset. Ten tweede zal verteld worden hoe de verschillende variabelen zijn bewerkt ter voorbereiding op de analyses, de operationalisatie van de variabelen. Ten derde zal beschreven worden hoe deze data is verdeeld. Als laatste zal het volledige analyseplan worden gepresenteerd.

Beschrijving van de dataset

De data waarmee wordt gewerkt is de data uit de Nederlandse Levensloopstudie (hierna NELLS) welke is opgezet om de sociologische onderzoeksgemeenschap te dienen. Alle vragen zijn opgesteld om de effecten van sociale cohesie, ongelijkheid en normen en waarden te meten. Deze dataset bestaat uit twee onderzoeksgolven; wave 1 (2009) en een wave 2 (2013) met oorspronkelijk 5312 Nederlandse inwoners tussen de leeftijd van 15-45. Er is expres gekozen voor een overrepresentatie van etnische minderheden (vooral Turken en Marokkanen), om onderzoek te kunnen doen naar de verschillen tussen de autochtone bevolking en minderheden. Ook is er gekozen voor gevarieerde methoden van informatieverzameling. Zo is in wave één het eerste deel van de vragenlijst verzameld door face-to-face interviews en het andere tweede deel door middel van een enquête die thuis ingevuld kon worden. Van de mensen die deelnamen aan de tweede wave hebben enige respondenten geparticipeerd door middel van een online vragenlijst en het andere deel door face-to-face interview.

De steekproef

De gemiddelde leeftijd van de steekproef is 31,24 jaar oud met een standaarddeviatie van 9,07. Vrouwen (53,2%) zijn licht oververtegenwoordigd in deze steekproef tegenover de mannen

(46,8%). De precieze getallen van alle gebruikte variabelen zijn terug te vinden in tabel 1 in het resultatenhoofdstuk. Van de 5312 respondenten hebben 2829 (75% van het gevraagde aantal) respondenten meegedaan aan wave 2. Omdat er in deze wave 2 veel informatie verloren gaat, besluit ik ervoor om in eerste instantie alleen naar de wave 1 te kijken. De gestelde vragen binnen wave 1 & 2 zijn veelal hetzelfde. “Onveranderbare vragen” zijn niet een tweede keer gesteld (zie: wanneer ben je geboren). Er zijn enige “update”-vragen opgenomen, zoals bijvoorbeeld vragen of de persoon nog met dezelfde partner samen is en of de ouders van de respondent nog bij elkaar zijn. Sommige vragen zijn op een betere manier gesteld in het tweede survey. Een goed voorbeeld hiervan is de variabele opleiding, die in wave 1 was gesteld als 14 losstaande vragen of de respondent een specifiek soort opleiding wel gevolgd heeft. Deze werden daarna opgevolgd met 14 vragen of deze opleiding ook met een diploma was afgerond. In wave 2 is simpelweg gevraagd wat de hoogste opleiding is die de respondent afgerond heeft.

Non-Respons

Het codeboek beschrijft enige patronen die te zien zijn in de non-respons van de niet-meegewerkte benaderden. Zo is te zien dat mannen, etnische minderheden en respondenten tussen de 15-24 jaar minder vaak meededen aan wave 2. Dit gold ook voor mensen die tussen wave 1 en wave 2 verhuisd waren. Bovendien was de algemene non-respons hoger voor inwoners van steden en Nederlandse inwoners uit het westen van het land. In totaal zijn 12310 mensen benaderd om mee te doen. Hiervan heeft 52% van alle respondenten meegedaan aan wave 1. Van deze 5312 respondenten, hebben 2829 (75% van het gevraagde aantal) respondenten meegedaan aan de Wave 2.

Operationalisaties

Om de verschillende analyses in goede baan te leiden, is de data eerst nog bewerkt. Het proces hiervan zal hieronder worden beschreven. Binnen de onderzoeksvraag zijn er drie

hoofdvariabelen het meest belangrijk: Sociale Cohesie, Eenzaamheid en Crimineel Gedrag. Als additie worden probleemgedrag en de controle-variabelen beschreven.

Sociale Cohesie

Voor het beantwoorden van de hypothesen aangaande sociale cohesie is een schaal opgesteld op basis van 6 verschillende stellingen uit de dataset. De stellingen begonnen met de vraag:

“Kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt?”, waarna 6 verschillende stellingen werden genoemd, namelijk: “mensen in deze buurt groeten elkaar”, “de mensen in deze buurt kan men vertrouwen”, “de mensen in deze buurt kunnen over het algemeen goed met elkaar opschieten”, “veel mensen kennen elkaar in deze buurt”, “de mensen hier willen elkaar graag helpen” & “de mensen in de buurt zouden er iets van zeggen als jongeren overlast bezorgen in de buurt”. Hierop konden de antwoorden van 1 tot en met 4 gegeven worden. Deze waarden zijn omgedraaid om te zorgen voor een makkelijkere interpretatie. Van deze antwoorden is het gemiddelde genomen. Binnen deze schaal betekent binnen 4 ‘dit klopt helemaal’ en 1 ‘dit klopt helemaal niet’. Hoe hoger de score op deze schaal, hoe hoger dit individu de sociale cohesie in de eigen wijk ervaart.

De schaal heeft een Cronbach’s Alpha van 0,849. In dit verslag zal aangenomen worden dat een schaal met een Cronbach’s Alpha boven de 0,6 acceptabel is. Boven de 0,7 is de schaal betrouwbaar en boven de 0,8 is deze goed. De schaal van sociale cohesie is aan de hand van deze maatstaf dus ‘goed’.

Eenzaamheid

Voor eenzaamheid is de ingekorte schaal van De Jong-Gierveld & van Tilburg (2008) gebruikt.

Deze is terug te vinden in de vragenlijst als zes stellingen, waarop geantwoord kon worden op een schaal van 1 tot 4. De stellingen waren als volgt: “er zijn voldoende mensen met wie ik me nauw verbonden voel”, “ik ervaar een leegte om me heen”, “ik heb veel mensen op wie ik

volledig kan vertrouwen”, “er zijn genoeg mensen op wie ik in geval van narigheid kan terugvallen”, “ik mis mensen om me heen” & “vaak voel ik me in de steek gelaten”.

Zoals hier te lezen is, staan niet alle stellingen in dezelfde “richting”, wat betekent dat een hoge score op alle stellingen in sommige gevallen wel eenzaamheid betekent en in andere gevallen juist weer niet. Daarom zijn de ingevulde scores eerst gedeeltelijk omgedraaid, en is daarna het gemiddelde hiervan genomen. De gegeven antwoorden variëren nu van 1 tot 4, waarop 4 ‘zeer van toepassing’ en 1 ‘helemaal niet van toepassing’ betekent. Hoe hoger de score op deze schaal, hoe eenzamer het individu zich voelt.

De Cronbach’s alpha van deze schaal is 0,788. Dit is met onze aangenomen beoordelingscriteria een betrouwbare schaal. Het verwijderen van items had niet voor een hogere betrouwbaarheid van de schaal kunnen zorgen.

Crimineel gedrag

Ook voor crimineel gedrag is een schaal opgezet, die bestaat uit de antwoorden van een vijftal vragen. Hierop wordt geantwoord of de respondent in de afgelopen 12 maanden ‘iets van een winkel of persoon gestolen heeft’, ‘een eigendom van iemand anders expres kapot gemaakt of beschadigd heeft’, ‘een wapen bij zich gedragen heeft’, ‘iemand bedreigd heeft’ en ten slotte of de respondent ‘iemand geslagen/geschopt of heeft deelgenomen heeft in een gevecht’. Hier lopen opnieuw de antwoordcategorieën waarop geantwoord kan worden van 1 (‘nooit’) tot 4 (‘4 keer of vaker’), maar er is ook een optie 5 voor ‘wil ik niet zeggen’. Om deze schaal te maken is ten eerste de ‘5’ als antwoordoptie als system ‘missing’ gezet, om geen verkeerd beeld te geven van de gemiddelde hoeveelheid crimineel gedrag. Van de overige antwoorden op de stellingen is het gemiddelde gepakt, waarbij geldt: hoe hoger de score van de respondent, hoe criminelier het gedrag van de respondent in de afgelopen 12 maanden.

De Cronbach’s alpha van deze schaal is 0,650. De schaal kan dus worden gezien als een betrouwbaar meetinstrument in crimineel gedrag gedrag.

Probleemgedrag

Om de resultaten van crimineel gedrag extra kracht bij te zetten zullen dezelfde analyses die uitgevoerd worden voor crimineel gedrag, ook worden uitgevoerd voor probleemgedrag. Deze variabele is opgebouwd uit 4 vragen aangaande probleemgedrag, waar wordt gevraagd hoe vaak de respondent in de afgelopen maand school of werk gemist heeft zonder reden, teveel gedronken heeft (meer dan 10 glazen), drugs gebruikt heeft (marijuana, wiet, hash, cocaïne, XTC etc.) of in de laatste maand 's avonds of 's nachts op straat rondgehangen heeft. Er kon worden geantwoord op een schaal van 1 tot 4, waarbij 1 staat voor 'nooit' en 4 voor '4 keer of vaker'. Hiervoor is een schaal gemaakt door het gemiddelde van al deze vragen te pakken. Hoe hoger de score op deze schaal, des te hoger het probleemgedrag.

De Cronbach's alpha van deze schaal blijft hangen op een 0,523. Deze schaal kan dus niet worden gezien als een betrouwbaar meetinstrument voor mogelijk crimineel gedrag versterkend gedrag. Er zijn geen items die men weg zou kunnen laten om op deze manier de betrouwbaarheid van de schaal te kunnen verhogen; het weglaten van één van de items zou de betrouwbaarheid alleen maar verlagen. Binnen de dataset is er geen ander meetinstrument dat we zouden kunnen gebruiken om probleemgedrag te meten. Toch is ervoor gekozen om de analyse uit te voeren om nog een mogelijk verband te kunnen duiden. Bovendien is de schending in betrouwbaarheid licht. Met de generaliseerbaarheid van de uitspraken moet echter voorzichtig worden omgegaan. De analyse kan mogelijk inspiratie geven tot een nieuwe uitvoering van de analyse, waarbij de tweede keer een betrouwbaarder meetinstrument gebruikt zou kunnen worden.

Controlevariabelen

De bewerkingen die zijn gedaan bij de controle-variabelen als voorbereiding op de analyse, zullen hieronder kort worden beschreven .

Geslacht is gehercodeerd naar een 0-1 notatie in plaats van de oorspronkelijke 1-2 notatie, met als doel om de interpretatie gemakkelijker te maken. Hierbij is 0 gecodeerd als vrouw en 1 als man. Hetzelfde is gedaan bij de variabele Werkend en Partner. Bij Werkend is 0 gecodeerd als “geen betaalde arbeid” en 1 als “wel betaalde arbeid”. Bij Partner is 0 gecodeerd als “geen partner” en 1 als “wel partner”.

Op leeftijd zijn geen bewerkingen uitgevoerd: deze variabele correspondeert gelijk met de leeftijd van de respondent.

Om iets over SES te kunnen zeggen wordt er naar drie verschillende variabelen gekeken, waarvan de eerste opleiding is. Aangezien deze variabele bestond uit 14 afzonderlijke vragen die los van elkaar vroegen of er een diploma is behaald bij een van de verschillende opleidingen die de respondent gevolgd heeft, is deze iets inzichtelijker gecodeerd. Ten eerste is een variabele gemaakt met de daarin de hoogste opleiding die de respondent heeft gevolgd. Daarna werden deze verschillende opleidingen ingedeeld in “hoog”, “midden”, “laag”, en “lager dan middelbaar” in een tweede variabele.

Inkomen is opgedeeld in 16 verschillende categorieën, die beginnen met “minder dan 150 euro per maand”, “150 - 299 euro per maand” & “300 - 499 euro per maand”. Hierna lopen de categorieën steeds verder omhoog in stappen van 500 euro, tot de laatste categorie die “7000 euro of meer per maand” per maand betekent. De enige bewerking die hier is gedaan is dat het antwoord ‘99’ (wil ik niet zeggen) op system ‘missing’ is gezet.

De codering van ‘werkend’ was al eerder te lezen.

Er is gekozen om alleen de respondenten mee te nemen die op minder dan 3 van alle variabelen een system ‘missing’ hebben. Het idee hierachter is dat respondenten die te veel vragen niet ingevuld hebben of na enige vragen zijn gestopt met het onderzoek, niet mee worden genomen in de analyse.

Er is gekozen voor het getal 3 omdat ik graag een antwoord zou willen zien op alle 4 de schalen en ten minste 3 controlevariabelen. In totaal zijn er 10 variabelen. Om deze reden is er gekozen voor 3 variabelen waarop een 'missing' ingevuld mag worden. In theorie zouden alsnog alle controlevariabelen ingevuld kunnen worden en enige schalen op onbekend kunnen staan. Dit zorgt er wel voor dat de n niet voor iedere regressie gelijk is. De N van het tweede model, waarbij eenzaamheid als afhankelijke variabele wordt gebruikt is voor de regressie, is 1699. De N van het derde model, waarbij probleemgedrag als afhankelijke variabele wordt gebruikt voor de regressie, is 1692. De N van het derde model, waarbij crimineel gedrag als afhankelijke variabele wordt gebruikt voor de logistische regressie, is 1680.

Analyseplan

Om de onderzoeksvraag te beantwoorden ("Wat is de rol die eenzaamheid speelt in het verband tussen sociale cohesie en crimineel gedrag?") zijn meerdere regressie-analyses nodig, die zorgvuldig zullen worden opgebouwd.

Lineaire regressie 1 = Eenzaamheid & sociale cohesie

1. Mijn eerste model zal de afhankelijke variabele (eenzaamheid) zijn met mijn controlevariabelen als onafhankelijke variabelen (geslacht, leeftijd, partner, werkend, inkomen & opleiding).
2. In het tweede model voeg ik mijn 'officiële' onafhankelijke variabelen toe. In dit geval is dit 1 variabele: sociale cohesie. Dit model zou het antwoord moeten geven op hypothese 2.

Lineaire regressie 2 = Probleemgedrag en vaststellen mediërend effect

1. Mijn eerste model zal bestaan uit de afhankelijke variabele 'probleemgedrag', in combinatie met mijn controlevariabelen als onafhankelijke variabelen (geslacht, leeftijd, partner, werkend, inkomen & opleiding).

2. In het tweede model wordt de eerste onafhankelijke variabele toegevoegd, in dit plan is dat sociale cohesie.
3. In het derde en tevens laatste model wordt de tweede onafhankelijke variabele toegevoegd: eenzaamheid. Eenzaamheid heeft mogelijk een mediërend effect. Door het verband tussen sociale cohesie & probleemgedrag tussen model 2 en 3 van deze regressie te vergelijken, zou je vast kunnen stellen of het hier om een mediërend effect gaat. Hierna kunnen hypothese 4 & hypothese 5 beantwoord worden.

Logistische regressie = Crimineel gedrag en vaststellen mediërend effect

1. Mijn eerste model zal bestaan uit de afhankelijke variabele 'crimineel gedrag' in combinatie met mijn controlevariabelen als onafhankelijke variabelen (geslacht, leeftijd, partner, werkend, inkomen & opleiding).
2. In het tweede model wordt de eerste onafhankelijke variabele toegevoegd, opnieuw sociale cohesie.
3. In het derde en tevens laatste model wordt de tweede onafhankelijke variabele 'eenzaamheid toegevoegd. Ook in deze regressie zal gekeken worden of sociale cohesie en crimineel gedrag een mediërend effect met elkaar lijken te vertonen. Hierna kunnen hypothese 1 & hypothese 3 beantwoord worden.

Resultaten

Beschrijvende statistieken

In deze paragraaf zal worden gekeken naar de verdeling van waarden binnen de variabelen om goed inzicht in de resultaten te verschaffen. Deze resultaten zijn terug te vinden in tabel 1 en worden uitgebreider besproken in bijlage 1.

De gemiddelde leeftijd van de steekproef is 31,30 jaar oud met een standaarddeviatie van 9,02. Vrouwen (53,2%) zijn licht oververtegenwoordigd in deze steekproef tegenover mannen (46,8%). De meeste respondenten uit de dataset hebben een partner (68,4%) en een baan (55,2%). Bovendien scoort de gemiddelde respondent een score van 6,09 op inkomen, wat neerkomt op een inkomen tussen de 1999 en de 2499 euro per maand. De gemiddelde waarde voor opleiding is een 1,64, wat door de categorische aard van deze vraag wijst op een opleiding tussen middelbare school-diploma en het mbo-diploma in.

Voor de schalen geldt dat er vrij hoog op sociale cohesie (3,31) gescoord wordt, echter is deze variabele vrij rechtsscheef verdeeld. De schaal eenzaamheid (1,89) ligt daarentegen iets minder hoog en de variabele eenzaamheid is redelijk normaal verdeeld. Op probleemgedrag (1,20) & crimineel gedrag (1,03) wordt daarentegen laag gescoord. Bovendien valt op dat deze waarden uiterst scheef verdeeld zijn. De implicaties en oplossingen hiervan zullen besproken worden in de modevaluatie.

Tabel 1: Beschrijvende statistieken afhankelijke, onafhankelijke en controlerende variabele^a

	BHI (95%)							
	Gemiddelde (SD)	N	Minimum	Maximum	Onderste grens	Bovenste grens	Skewness (SD)	Kurtosis (SD)
Eenzaamheid	1,89 (0,50)	4832	1,00	4,00	1,87	1,90	0,18(0,04)	0,066(0,07)
Sociale Cohesie	3,31 (0,57)	5019	1,00	4,00	3,30	3,33	-0,79(0,04)	0,40(0,07)
Probleemgedrag	1,20 (0,37)	4839	1,00	4,00	1,18	11,20	2,83(0,04)	10,60(0,07)
Crimineel gedrag	1,03 (0,15)	4820	1,00	3,20	1,03	1,04	7,09(0,04)	64,90(0,07)
Geslacht	53,2% man, 46,8% vrouw	5020	0,00	1,00	-	-	0,13 (0,04)	-1,98(0,07)
Leeftijd	31,24 (9,07)	5020	14	49	27,33	27,33	-0,19(0,04)	-1,20(0,07)
Partner	68,4% partner; 31,6% geen partner	5020	0,00	1,00	-	-	-0,79(0,04)	-1,38(0,07)
Werkend	55,2% betaald werk, 44,8% geen betaald werk	2201	0,00	1,00	-	-	-0,21(0,05)	-1,96(0,10)
Opleiding	1,64 (0,99)	4691	0,00	3,00	3,40	3,54	-0,078(0,04)	-1,07(0,07)
Inkomen	6,09(3,09)	4494	1,00	16,00	4,34	4,60	0,36(0,04)	0,12(0,07)

^a Bij binominale variabelen (geslacht, gehuwd & aanwezigheid baan) is de frequentieverdeling weergegeven in percentages. Bij inkomen zijn de getallen weergegeven in duizendtallen.

Onderlinge correlaties

De correlaties tussen de variabelen geven een indicatie over de onderlinge relatie tussen de variabelen. De correlaties duiden mogelijke multi-collineariteit aan. De meest belangrijke correlaties om te beschouwen zijn de afhankelijke variabelen binnen de verschillende regressies. Voor de interpretatie van de sterkte van de correlaties zal het onderzoek van Senthilnathan (2019) worden gebruikt. De veronderstelde relatie tussen probleemgedrag en crimineel gedrag vertonen een redelijke correlatie binnen tabel 2 ($r=0,303; p<0,01$). Volgens Senthilnathan's opgestelde spectrum is deze positieve relatie redelijk zwak, maar wel iets waar rekening mee gehouden moet worden. Dit betekent dat probleemgedrag en crimineel gedrag wel redelijk met elkaar geassocieerd mogen worden, maar het betekent zeker niet dat iedereen die probleemgedrag vertoont ook criminele activiteit zal vertonen. Een tweede redelijk zwakke positieve relatie is te vinden tussen probleemgedrag en geslacht ($r=0,213; p<0,01$). Door de manier waarop geslacht gecodeerd is, geeft dit aan dat mannen eerder geneigd zijn om

probleemgedrag te vertonen. Zowel leeftijd ($r=-0,285;p<0,01$), inkomen ($r=-0,240;p<0,01$) als het hebben van een partner ($r=-0,248;p<0,01$) hangen negatief samen met probleemgedrag. Dit houdt in dat wanneer men ouder wordt, meer verdient of een partner heeft, men minder snel probleemgedrag vertoont. Alle drie de relaties zijn zwak, maar hebben zijn wel groot genoeg dat het een verband aantoont waar enige rekening mee gehouden moet worden. Voor de correlatie met leeftijd is het goed om te beseffen dat de jongste respondent 14 jaar oud was. Het gaat hier dus om oudere tieners, adolescenten en volwassenen, niet om kinderen.

De correlaties tussen verschillende variabelen met eenzaamheid zijn ook goed om te bekijken. De positieve relatie tussen sociale cohesie is vrij zwak, maar wel significant ($r=0,160;p<0,01$). Wanneer men aangeeft een hogere ervaring van sociale cohesie in de eigen wijk te ervaren, voelt men zich dus minder eenzaam. Probleemgedrag ($r=0,033;p<0,05$) en crimineel gedrag ($r=0,078;p<0,01$) hangen beide positief samen met eenzaamheid, wat zou betekenen dat respondenten die dit soort gedrag vertonen, zich eenzamer voelen dan respondenten die dit niet doen. Dit is in overeenstemming met de verwachtingen uit het theorie-hoofdstuk, alhoewel de zojuist genoemde verbanden heel erg zwak zijn. Sociale cohesie hangt ook zwak negatief samen met zowel probleemgedrag ($r=-0,106;p<0,01$) als crimineel gedrag ($r=-0,092;p<0,01$). Dit zou betekenen dat iemand die een hogere sociale cohesie in de eigen wijk ervaart, minder snel probleemgedrag of crimineel gedrag zou vertonen.

Tabel 2: Correlaties tussen de continue & categorische variabelen

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1 Eenzaamheid	-	**-.160 ^a	*0,033 ^a	**0,078 ^a	0,016 ^a	**0,045 ^a	**0,106 ^a	**-.0,129 ^a	**-.0,147 ^a	**-.0,155 ^a
2 Sociale Cohesie		-	**-.0,106 ^a	**-.0,092 ^a	**-.0,052 ^a	**0,103 ^a	**0,093 ^a	0,027 ^a	**0,141 ^a	**0,068 ^a
3 Probleemgedrag			-	**0,303 ^a	**0,213 ^a	**-.0,285 ^a	**-.0,248 ^a	0,010 ^a	**-.0,240 ^a	**-.0,136 ^a
4 Crimineel gedrag				-	**0,127 ^a	**-.0,064 ^a	**-.0,123 ^a	*-.0,020 ^a	**-.0,102 ^a	**-.0,088 ^a
5 Geslacht					-	0,014 ^a	**-.0,049 ^c	**0,058 ^c	0,001 ^a	*-.0,036 ^a
6 Leeftijd						-	**0,434 ^a	*0,046 ^a	**0,571 ^a	**0,259 ^a
7 Partner							-	**0,109 ^c	**0,546 ^a	**0,267 ^a
8 Werkend								-	**0,185 ^a	**0,174 ^a
9 Inkomen									-	**0,486 ^a
10 Opleiding										-

* Tweezijdig significant op 0,05 niveau, ** Tweezijdig significant op 0,01 niveau. ^a Pearson correlatie, ^b Cramer's V, ^c Correlatie op basis van ANOVA.

Binnen tabel 2 zijn enige relaties te vinden die sterker zijn dan de relaties die genoemd zijn in voorgaand paragraaf; zoals die tussen inkomen en leeftijd ($r=0,571; p<0,01$). Nog steeds redenerend van de door Senthilnathan's vastgestelde criteria, is een aanzienlijke en redelijk hoge correlatie. Deze duidt waarschijnlijk op het feit dat wanneer iemands werkervaring stijgt, diens loon ook stijgt. Wanneer men snel kijkt, is de hoge positieve correlatie tussen partner en inkomen opvallend ($r=0,546; p<0,01$). Dit kan echter verklaard worden door het feit dat de vraag uit de dataset vraagt naar het inkomen van de respondent en de partner samen. Ook de gematigd positieve relatie tussen inkomen en opleiding is snel te verklaren ($r=0,486; p<0,01$), met een hogere opleiding kan men betere banen bemachtigen met een hoger loon.

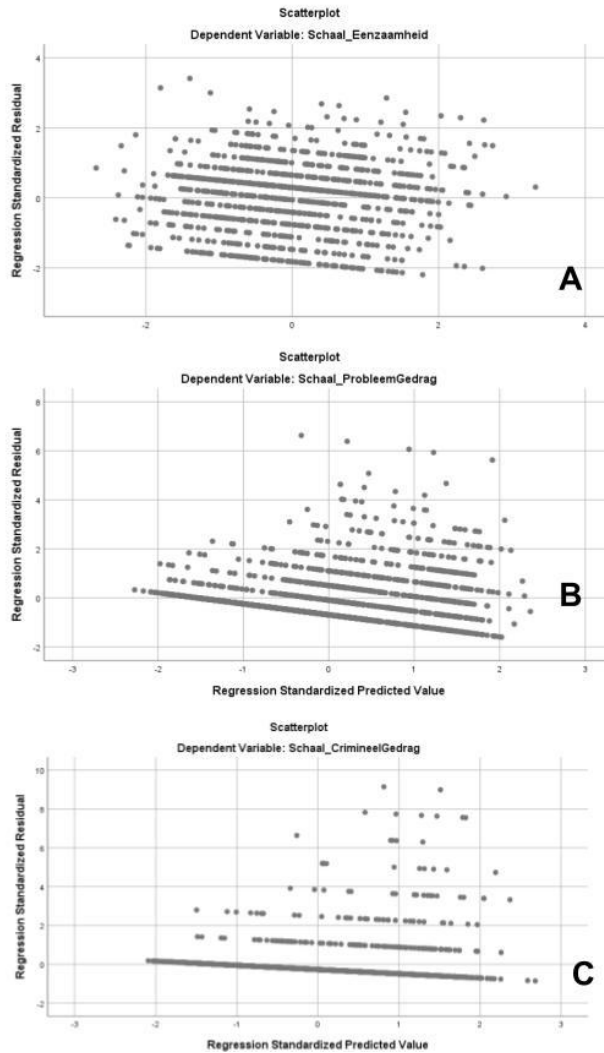
Zoals te zien is in tabel 2 zijn de correlaties niet erg hoog. Volgens Duncan et al. (1986) zou men zich vanaf 0,8 zorgen moeten maken over multicollineariteit. Dit geeft dus geen indicatie dat dit binnen de analyse problemen op gaat leveren.

Modelevaluatie

Assumpties lineaire regressie

De steekproef is niet geheel willekeurig getrokken, maar door middel van twee-fase stratified sampling. In 35 geselecteerde gemeenten is er een willekeurige selectie gemaakt op basis van leeftijd, geboorteland en ouders van de respondent. In een ideale situatie zou er gekozen moeten worden voor een multi-level analyse, maar dit is boven het niveau van dit onderzoek. Omdat er is gecorrigeerd voor stedelijkheid en regio is besloten dat de assumptie van een willekeurige steekproef niet is geschonden.

Wegens de categorische aard van de schalen, zijn er meerdere streeppatronen te vinden in de residual plots (figuur 2). Eenzaamheid is hierin nog redelijk willekeurig verdeeld, maar door de inhoudelijke betekenis van probleemgedrag en crimineel gedrag is de onderkant van deze verdeling vrij zwaar. Hierbij schendt crimineel gedrag de assumptie van lineariteit sterker dan probleemgedrag. De homoscedasticiteit kan getoetst worden door te kijken naar een “band” binnen de residual plots. Wanneer deze door de gehele plot een gelijke afstand tot de nullijn houdt, is deze assumptie niet geschonden. Bij eenzaamheid kan deze assumptie nog redelijk goed worden aangehouden, maar bij probleemgedrag is toch enige scheefheid van de grafiek te zien. Bij crimineel gedrag is er bijna geen enkele punt boven de nullijn te vinden. Crimineel gedrag schendt deze assumptie dus sterk, probleemgedrag schendt deze redelijk.



Figuur 2: Residual plots van A. Eenzaamheid, B. Probleemgedrag & C. Crimineel gedrag

Eenzelfde soort conclusies kunnen worden getrokken voor de normaalverdeling van de residuen. Waarvan het bijpassende figuur (figuur 8) in bijlage 3 staat. Bij eenzaamheid is een piek te zien, die redelijk normaal verdeeld is. Probleemgedrag is opnieuw enigszins problematisch en de residuen van crimineel gedrag zijn totaal niet normaal verdeeld. Een volledige en uitgebreide analyse van de assumpties, inclusief alle figuren zijn terug te vinden in bijlage 3.

De conclusie die uit deze analyse kan worden getrokken is dat de variabele eenzaamheid prima past bij een lineaire regressie. Probleemgedrag schendt alle assumpties enigzins. Deze zal nog wel geschat worden met een lineaire regressie, maar om hiervoor te corrigeren zal de alpha hier worden gesteld op 0,01. Ook zal er voorzichtiger worden omgegaan met de generalisatie van de conclusies. Als laatste voldoet crimineel gedrag aan geen van de assumpties, waardoor deze geschat zal worden door middel van een logistische regressie.

Missing data

Om mogelijke biases in de dataset op te sporen, zal er worden gekeken naar of er patronen kunnen worden gevonden in de missende waarden binnen de dataset. Gelukkig zijn er binnen dit onderzoek niet erg veel patronen gevonden in de missende waarden. Het is opvallend dat het meest voorkomende patroon wat gevonden kan worden, bestaat uit missende antwoorden op de probleemgedrag-schaal en de vraag of iemand betaalt werk verricht. Maar liefst 46,65% van de respondenten hebben deze twee vragen niet ingevuld. De missende antwoorden zijn snel te verklaren om de volgende reden: Ten eerste is deze vraag niet gesteld aan respondenten die voltijd onderwijs volgen, ten tweede stond de vraag op de ene pagina en de antwoordmogelijkheden op de volgende pagina. Hierdoor is de kans groot dat veel respondenten dit over het hoofd hebben gezien. Door een combinatie van deze factoren hebben in totaal maar liefst 2819 respondenten deze vraag niet beantwoord (56,2%). De reden dat respondenten de vragen aangaande probleemgedrag niet ingevuld hebben heeft waarschijnlijk met sociale wenselijkheid te maken. Vier van de tien variabelen hadden geen enkele missende waarde: dit waren geslacht, partner, de sociale cohesie-schaal en leeftijd. Maar 33,39% van de respondenten had geen enkel missend antwoord op de gebruikte variabelen. Een meer uitgebreide analyse van de missing-data en de daarbij horende grafieken is te vinden in bijlage 3.

Outliers

Binnen de verschillende analyses was een groot aantal uitbijters te vinden. Wanneer er gekeken wordt naar alleen de gestandaardiseerde residuen lag het aantal uitbijters tussen de 23 (eenzaamheid) en de 42 (crimineel gedrag). Dit is geen ontzettend hoog aantal voor een steekproef van 4.382 respondenten. Voor probleemgedrag lag het grootste gestandaardiseerde residu op 6,63 en voor probleemgedrag lag deze op 9,14. Deze hoge residuen kunnen verklaard worden door het lage gemiddelde van beide variabelen (Probleemgedrag: 1,20 & Crimineel gedrag: 1,03) en de grote n van de steekproef. Hierdoor worden de standaarddeviaties erg klein en krijgt men men bij een hogere score al snel een groter residu. Het meest opvallende was het grote aantal uitbijters gebaseerd op de leverage-waarden. Bij eenzaamheid waren dit er 507 en bij crimineel gedrag waren dit er 437. Bij probleemgedrag waren dit er in verhouding heel weinig, aangezien maar 80 respondenten hier een hoge leverage-waarde hadden. Na het herschatten van de modellen steeg bij alle regressie de verklaarde variantie, alhoewel dit bij probleemgedrag minimaal was. Er kon echter geen systematische fout worden gevonden in de dataset, waardoor de uitbijters in de dataset zijn behouden. Voor crimineel gedrag en probleemgedrag was hier bovendien de inhoudelijke overweging dat beide juist de uitzondering is. Wanneer de hoge scores hierop zouden worden verwijderd, verdwijnt de zeggingskracht van het model.

Hypothesetoetsing

Er is gepoogd om de hypothesen (die al eerder zijn genoemd in het theorie-hoofdstuk) te verklaren door middel van 2 lineaire en 1 logistische regressie-analyse. In dit hoofdstuk zullen de hypothesen stuk voor stuk worden besproken, met het daarbij behorende regressie-model. Aan de hand van deze regressies kan er geanalyseerd worden welke variabelen een invloed hebben op eenzaamheid, probleemgedrag en crimineel gedrag. Hierbij wordt gebruik gemaakt van analyseplan die in het methode-hoofdstuk is besproken.

Soms zal de chronologische volgorde van de hypothesen enigszins genegeerd worden, omdat ik ervan overtuigd ben dat dit de leesbaarheid ten goede komt.

Hypothese 1: Sociale Cohesie & Crimineel gedrag

De eerste hypothese gaat over het verband aangaande sociale cohesie en crimineel gedrag, namelijk:

Hypothese 1: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, vertoont dit individu minder crimineel gedrag.

In de eerste alinea zal er een onderbouwing gegeven worden, waarom er gekozen is om te kijken naar het laatste model. Hierna zal het resultaat genoemd worden wat hoort bij hypothese 1. De overige variabelen zullen worden besproken in bijlage 2.

Omdat het hier gaat om een logistische regressie wordt hier gebruik gemaakt van een Chi-kwadraattoets om de kwaliteit van het model te beoordelen. Deze test vergelijkt model 1 met een leeg model, dit model geeft aan dat de kans op crimineel gedrag voor iedere respondent 0,5 is. In tabel 3 is te zien dat model 1 een betere voorspelling kan doen dan het lege model ($X^2(6) = 78,01; P < 0,0005$). Model 2 geeft weer een significant betere voorspelling dan model 1 ($X^2(1) = 12,47; P < 0,0005$), wat aangeeft dat sociale cohesie een waardevolle toevoeging is voor het voorspellen van crimineel gedrag. Voor model 3 kan dit niet over eenzaamheid gezegd worden ($X^2(1) = 0,27; P = 0,603$). Het laatste model is niet significant beter dan model 2.

Ook de deviance geeft een waardevolle inschatting over de kwaliteit van het model. De deviance vertelt ons “hoe slecht” een model bij de data past. Hiervoor geldt: hoe lager de waarde, hoe beter het model past bij de data. In tabel 3 is te zien dat de waarden voor ieder model steeds meer afnemen. Op basis van de deviance is het laatste model dus wel een

verbetering van model 2, maar via de Chi-kwadraattoets kan niet bewezen worden dat het toevoegen van eenzaamheid een significante verbetering is voor het voorspellen van crimineel gedrag. Op basis van van de chi-kwadraat toets wordt ervoor gekozen om de hypothese aangaande sociale cohesie te verklaren vanuit model 2.

Tabel 3

Logistische regressie (Crimineel gedrag als afhankelijke variabele), met Deviance en X²-Toets van model 3

	Model 1			Model 2			Model 3		
	β (SE)	Odd-Ratio	P	β (SE)	Odds-Ratio	P	β (SE)	Odds-Ratio	P
Constante (Crimineel gedrag)	-1,81(0,31)	0,16	<0,0005	-0,11(0,57)	0,89	0,839	-0,30(0,66)	0,75	0,656
Geslacht (0=vrouw, 1=man)	1,02 (0,116)	2,76	<0,0005	0,98 (0,18)	2,67	<0,0005	0,99(0,18)	2,68	<0,0005
Partner (0=geen partner, 1=wel partner)	-0,24 (0,20)	0,79	0,255	-0,24 (0,20)	0,78	0,228	-0,23(0,20)	0,80	0,260
Leeftijd	-0,02(0,02)	0,98	0,119	-0,03(0,02)	0,98	0,088	-0,03(0,015)	0,97	0,076
Werkend (0=niet werkend, 1=wel werkend)	0,21 (0,18)	1,23	0,25	0,21 (0,18)	1,23	0,243	0,22(0,18)	1,25	0,226
Inkomen	-0,10(0,06)	0,90	0,08	-0,098(0,059)	0,91	0,099	-0,097(0,059)	0,91	0,104
Opleiding	-0,03(0,11)	1,03	0,799	0,015(0,11)	1,02	0,885	0,018(0,11)	1,02	0,861
Sociale Cohesie	-	-	-	-0,51 (0,14)	0,60	<0,0005	-0,50(0,14)	0,60	<0,0005
Eenzaamheid	-	-	-	-	-	-	0,09(0,17)	1,09	0,603
Deviance	1005,46			992,99			992,72		
X ² -Toets / P	78,01			<0,0005			12,47		
							<0,0005		
							0,27		
							0,603		

Aangezien de hellingen van een logistische regressie de stijging van het aantal log-odds binnen het model aangeven, is er gekozen om gebruik te maken van een “gemiddelde respondent”. Dit is een fictieve respondent die gemiddeld scoort op alle verschillende variabelen. Het verschil in kans met deze “gemiddelde respondent” zal als indicatie voor een hoge of lage stijging van de kans in crimineel gedrag gebruikt worden. Wanneer men alle gemiddelde waarden (te vinden in tabel 1) invult in de regressie, komt men uit op een kans van 0,084 op crimineel gedrag. Een respondent die gemiddeld scoort op alle waarden van de regressie zal dus een 0,084 kans hebben om crimineel gedrag te vertonen. Het is goed om te benoemen dat door de manier

waarop deze regressie is opgebouwd, dit kan betekenen dat deze persoon afgelopen maand een mes bij heeft gedragen, maar ook dat deze persoon wekelijks doet aan vernielingen, vechtpartijen en bedreigingen.

Om de variatie in crimineel gedrag te kunnen laten zien in combinatie met sociale cohesie, wordt de kans van de “gemiddelde respondent” vergeleken met een respondent die veel sociale cohesie ervaart en een respondent die weinig sociale cohesie ervaart. Wanneer een respondent een hele lage sociale cohesie ervaart, is de kans dat deze dat deze crimineel gedrag pleegt 0,232 ($b=-0,51; p<0,0005$). Niet alleen is deze helling significant, de gevonden kans is bovendien de hoogste kans die binnen deze regressie is gevonden. Ter vergelijking had de eerder genoemde fictieve ‘gemiddelde respondent’ een kans van 0,084 om crimineel gedrag te vertonen. Wanneer iemand een hoge sociale cohesie ervaart is deze kans 0,061. Men kan dus zeggen dat de afwezigheid van sociale cohesie de kans dat een respondent crimineel gedrag vertoont, vergroot. Het verschil tussen een hoge sociale cohesie en een gemiddelde sociale cohesie is daarentegen niet heel groot. Ter conclusie, wordt hypothese 1 hiermee dus ondersteund.

Binnen deze regressie is er geen sprake geweest over multi-collineariteit. Een uitgebreider verslag hiervan staat in bijlage 3.

Hypothese 4: Sociale Cohesie & Probleemgedrag

Om meer gewicht te kunnen geven aan de resultaten wordt er ook gekeken naar het effect van sociale cohesie op probleemgedrag, dat terug te vinden is in tabel 4. Deze hypothese luidt:

Hypothese 4: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, vertoont dit individu minder probleemgedrag.

Ook hier is er gekozen om in de eerste alinea te vertellen waarom er gekozen is om te resultaten te rapporteren uit het laatste model en daarna om de resultaten te noemen die betrekking hebben op de hypothese. De overige variabelen zullen besproken worden in bijlage 2.

Tabel 4

Lineaire regressie (Probleemgedrag als afhankelijke variabele), met F_{change} , R^2_{adjusted} en VIF-waarden van model 3

	Model 1		Model 2		Model 3		VIF
	β (SE)	P	β (SE)	P	β (SE)	P	
Constante (Probleemgedrag)	1,52(0,036)	<0,0005	1,77(0,068)	<0,0005	1,79(0,090)	<0,0005	-
Geslacht (0=vrouw, 1=man)	0,22 (0,21)	<0,0005	0,22 (0,21)	<0,0005	0,22(0,021)	<0,0005	1,033
Partner (0=geen partner, 1=wel partner)	-0,066 (0,025)	0,008	-0,068 (0,025)	0,006	-0,067 (0,25)	0,007	1,501
Leeftijd	-0,011 (0,002)	<0,0005	-0,011 (0,002)	<0,0005	-0,011 (0,002)	<0,0005	2,229
Werkend (0=niet werkend, 1=wel werkend)	0,020 (0,022)	0,363	0,020 (0,021)	0,352	0,020 (0,022)	0,343	1,094
Inkomen	-0,004(0,006)	0,537	-0,002(0,006)	0,7740	-0,002(0,006)	0,748	2,786
Opleiding	0,003(0,12)	0,802	0,002(0,012)	0,872	0,002(0,012)	0,862	1,268
Sociale Cohesie	-	-	-0,076 (0,018)	<0,0005	-0,076 (0,018)	<0,0005	1,041
Eenzaamheid	-	-	-	-	0,005 (0,021)	0,829	1,083
R^2_{adjusted}		0,151		0,160		0,159	
F_{change} / P	51,165	<0,0005	18,319	<0,0005	0,047	0,829	

Het eerste model heeft (op basis van de R^2_{adjusted}) een verklaarde variantie van 15,1%, wat in het derde model zelfs oploopt tot 15,9%. Dit betekent dat sociale cohesie, in combinatie met eenzaamheid en de controlevariabelen, 15,9% beter voorspellingen doet over de waarde van probleemgedrag dan dat het gemiddelde dat doet. Hieruit kunnen we opmaken dat een van de modellen beter gebruikt kan worden om onze hypothese te beantwoorden dan het gemiddelde. De F-change laat (ons) zien tot in hoeverre ieder model beter is dan het voorgaande model. Hierbij is het opvallend om te zien dat niet ieder model een significant betere voorspelling geeft

van de waarde van probleemgedrag dan het vorige model. Zo is het laatste model niet beter dan het tweede model is ($F(6, 1685) = 51,165; p < 0,0005$ / $F(1, 1684) = 18,319; p < 0,0005$ / $F(1, 1683) = 0,047; p = 0,829$). Dit geeft aan dat sociale cohesie een waardevolle bijdrage levert in het voorspellen van probleemgedrag, eenzaamheid waarschijnlijk niet. Dit maakt dat we het best het tweede model kunnen gebruiken voor het beantwoorden van hypothese 4.

Binnen dit tweede model is het antwoord op de hypothese te vinden. Deze hypothese gaat over wat voor invloed het voelen van sociale cohesie in de eigen wijk heeft op het probleemgedrag van dit individu. Het voelen van sociale cohesie in de buurt heeft een redelijk sterk verminderend effect op probleemgedrag ($b = -0,076; P < 0,0005$). Ook hier de schaal loopt van 1 tot 4, wat betekent dat deze afname kan toenemen tot een daling van $-0,304$. Dit is een gematigd sterke daling van het gemiddelde ten opzichte van respondenten die geen hoge sociale cohesie in de wijk voelen. Zo is de gemiddelde waarde op probleemgedrag voor een respondent die geen sociale cohesie voelt $1,47$ en voor een respondent die deze wel heel sterk voelt $1,17$. Dit effect is bovendien ook nog niet significant onder de eerder vastgestelde alpha van $0,01$. Hiermee wordt onze nulhypothese dus niet verworpen en de alternatieve hypothese wel.

De overige variabelen worden besproken in bijlage 2. De VIF-waarden in tabel 4 geven aan dat multicollineariteit geen probleem is binnen deze analyse, alle waarden liggen ver onder de gebruikelijke waarde van vier. Een uitgebreidere uitleg over multi-collineariteit staat in bijlage 3.

Hypothese 2: Sociale Cohesie & Eenzaamheid

De tweede hypothese gaat over de relatie tussen sociale cohesie en eenzaamheid, namelijk:

Hypothese 2: Naarmate een individu een hogere sociale cohesie in diens eigen wijk ervaart, zal dit individu zich minder eenzaam voelen.

De indeling zal verder op dezelfde manier verlopen als in de vorige secties.

Tabel 5

Lineaire regressie (Eenzaamheid als afhankelijke variabele), met F_{change} , R^2_{adjusted} en VIF-waarden van model 2

	Model 1		Model 2		VIF
	β (SE)	P	β (SE)	P	
Constante	1,79(0,04)	<0,0005	2,22(0,079)	<0,0005	-
Geslacht (0=vrouw, 1=man)	0,014 (0,025)	0,560	0,002 (0,024)	0,941	1,033
Partner (0=geen partner, 1=wel partner)	-0,119(0,029)	<0,0005	-0,122 (0,029)	<0,0005	1,485
Leeftijd	0,012 (0,002)	<0,0005	0,012 (0,002)	<0,0005	2,162
Werkend (0=niet werkend, 1=wel werkend)	-0,11(0,025)	<0,0005	-0,107 (0,025)	<0,0005	1,083
Inkomen	-0,015(0,007)	0,028	-0,012(0,007)	0,075	2,785
Opleiding	-0,033(0,014)	0,020	-0,035(0,014)	0,013	1,264
Sociale Cohesie	-	-	-0,134(0,021)	<0,0005	1,016
R^2_{adjusted}		0,051		0,073	
F_{change} / P	16,225	<0,0005	41,571	<0,0005	

Het eerste model van de eerste regressie (tabel 5) geeft op op basis van de R^2_{adjusted} een vrij lage verklaarde variantie. Ten opzichte van het gemiddelde kan het eerste model 5,1% betere voorspellingen doen, wat gelijk blijkt voor het tweede model ($R^2_{\text{adjusted}} = 0,051$). Op basis van de F-toets is model 1 significant beter in het voorspellen van eenzaamheid dan het gemiddelde ($F(6, 1692) = 16,255$; $p < 0,0005$). Op basis van de R^2_{adjusted} doet het tweede

model nog betere voorspellingen dan model 1. De F-Toets geeft ook nogmaals aan dat het toevoegen van sociale cohesie aan het model significant betere voorspellingen geeft ($F(1, 1691) = 41,571; p < 0,0005$). Dit is onder andere de reden waarom er naar het laatste model gekeken zal worden. Deze verbetering toont de waarde aan van het toevoegen van sociale cohesie aan het model.

Zoals al uit de analyse van de twee verschillende modellen viel af te leiden is het toevoegen van sociale cohesie een verbetering voor het doen van voorspellingen over eenzaamheid. De richting van het effect gaat precies de kant op die in het theorie-hoofdstuk wordt verwacht. Het ervaren van sociale cohesie in de buurt laat de respondent minder eenzaam voelen ($b = -0,134; P < 0,0005$). Het effect is aanzienlijk, en daarnaast significant. Hiermee kan onze nulhypothese worden verworpen en is dus statistische grond gevonden voor de alternatieve hypothese.

Ook bij deze regressie is multicollineariteit geen probleem, omdat ook hier alle waarden onder de gebruikelijke VIF-grenswaarde van vier blijven. Dit is ook in lijn met de verwachting die was gesteld op basis van de onderlinge correlatie. Meer hierover is te vinden in bijlage 3.

Hypothese 3: Eenzaamheid & Crimineel gedrag

De derde hypothese gaat over het verband tussen eenzaamheid en crimineel gedrag, namelijk:

Hypothese 3: Naarmate een individu zich eenzaam voelt, zal dit individu meer crimineel gedrag vertonen.

In de sectie 'Hypothese 1: Eenzaamheid & Crimineel gedrag' en in tabel 3 werd al kort besproken dat het derde model, het model waarin eenzaamheid als enige variabele wordt toegevoegd, niet significant beter is dan voorgaande modellen. De deviance geeft aan dat het model wel minder fouten maakt, maar op basis van de Chi-kwadraattoets kan er niet gezegd worden dat het hier gaat om een significant beter model. Voor deze hypothese zal er gekeken

worden naar het laatste model, met als reden dat deze als enige de gegevens heeft die nodig zijn om uitspraken over deze hypothese te doen.

Aangezien we het hier opnieuw hebben over een logistische regressie zorgt een stijging in de waarde van een variabele niet één op één voor eenzelfde verandering in de kans op crimineel gedrag, maar is eerst een rekensom nodig. De 'gemiddelde' respondent heeft een kans van 0,072 om crimineel gedrag te vertonen. Het toevoegen van eenzaamheid aan het model geeft een lagere kans op crimineel gedrag dan het model zonder eenzaamheid als variabele erin ($P=0,044$). Een gemiddelde respondent met een hoge eenzaamheid ($b=0,09; p=0,603$) geeft een kans van 0,086 om crimineel gedrag te vertonen. Deze in vergelijking tot respondenten die weinig eenzaamheid ervaren: deze hebben een gemiddelde kans van 0,067 op het vertonen van crimineel gedrag. Het voelen van een hoge eenzaamheid verhoogt dus de kans enigszins dat de respondent crimineel gedrag vertoont. Wanneer iemand weinig eenzaamheid ervaart ligt deze kans juist iets lager. De verschillen zijn echter niet groot, en de significantie van de helling komt niet in de buurt van de eerder opgestelde alpha van 0,05. Op basis hiervan is er niet genoeg basis om de nulhypothese te verwerpen.

Hypothese 5: Eenzaamheid & Probleemgedrag

De laatste hypothese waar nog antwoord op gegeven moet worden luidt als volgt:

Hypothese 5: Naarmate een individu zich eenzaam voelt, zal dit individu meer probleemgedrag vertonen.

Zoals te zien was in sectie 'Hypothese 1: Sociale Cohesie & Probleemgedrag' en in tabel 4, geeft het toevoegen van eenzaamheid een verandering in de F-Toets van 0,047 ($F(1, 1683)=0,047; p=0,829$). Dit geeft aan dat het toevoegen van eenzaamheid aan het model niet zorgt voor significant betere voorspellingen als het gaat om probleemgedrag. Toch zal dit laatste

model gebruikt worden om de hypothese te testen, simpelweg omdat model 2 geen gegevens over eenzaamheid en diens effect op probleemgedrag. De effecten van de sociale cohesie worden in bovenstaande sectie benoemd. De effecten van de controle-variabelen zijn terug te vinden in bijlage 2.

Uit model 3 is af te lezen dat des te eenzamer het individu, des te groter het probleemgedrag ($b=0,005; p=0,829$). Dit is geen sterke helling, en deze is al helemaal niet significant. Op basis hiervan kan de nulhypothese niet verworpen worden. Er kan dus wanneer men kijkt naar deze gegevens geen verband gevonden worden tussen probleemgedrag en eenzaamheid.

Mediatie-effect

Zoals eerder genoemd in het theorie-hoofdstuk is de verwachting om een mediatie-effect te zien, waarbij de relatie tussen sociale cohesie en crimineel gedrag weg zou vallen wanneer eenzaamheid in het model zou worden geïntroduceerd. Voor de volledigheid is het goed om te benoemen dat in tabel 3 te zien is dat de verwachte mediatie-relatie uitblijft. In model 2 is te zien dat sociale cohesie crimineel gedrag verlaagt met een helling van $-0,51$ ($p < 0,0005$). Wanneer in model 3 eenzaamheid wordt toegevoegd aan de regressie wordt deze helling weliswaar kleiner ($b=-0,50; p < 0,0005$), maar van het verdwijnen van de helling is geen sprake. In de hieraan voorafgaande paragrafen is al geconstateerd dat eenzaamheid geen verband heeft met crimineel gedrag ($b=0,09; p=0,17$). Dit bevestigt alleen maar meer het uitblijven van dit mediatie-effect.

Ook bij probleemgedrag is de mediatie niet te zien. Zowel voor als na het toevoegen van eenzaamheid blijft de helling van sociale cohesie $-0,076$ ($p=0,018$). De niet-significante helling van eenzaamheid laat ook zien dat eenzaamheid niet de eigenlijke oorzaak binnen dit verband is.

Conclusie & Discussie

Conclusie

Dit onderzoek begon met de vraag “waarom is eigenlijk niet iedereen crimineel?”. De literatuur in het theorie-hoofdstuk gaf aan dat sociale cohesie in de eigen wijk hierin een grote rol speelt, maar mijn hypothese was dat eenzaamheid hier ook een rol in speelde. Hierdoor ontstond de onderzoeksvraag “Wat is de rol die eenzaamheid speelt in het verband tussen sociale cohesie en crimineel gedrag?”. Door middel van vijf hypothesen is gepoogd om hier een antwoord op te geven.

In het resultatenhoofdstuk kon gelezen worden dat de literatuur aangaande sociale cohesie en criminaliteit door de resultaten in dit onderzoek opnieuw wordt bevestigd. Ook in de NELLS dataset kon gevonden worden dat een individu die een hogere sociale cohesie in diens eigen wijk ervaart, minder crimineel gedrag vertoont (hypothese 1). Voor probleemgedrag kon dit effect niet worden vastgesteld (hypothese 4). In het onderzoek van Kalff (2001) is al eerder voorgesteld dat lage sociaal-economische status een mediatie-effect zou vertonen op de relatie tot het wonen in een wijk met lage SES en het vertonen van probleemgedrag in kinderen. Zover ik heb kunnen vinden is er geen literatuur die het verband tussen probleemgedrag en sociale cohesie binnen de wijk direct vaststelt.

Er is statische basis gevonden worden voor het idee dat een individu een hogere sociale cohesie voelt in de eigen wijk, wanneer deze zich minder eenzaam voelt (hypothese 2).

Wanneer er een sterke sociale cohesie in de wijk aanwezig is, verlaagt dit dus de eenzaamheid.

Op basis van dit onderzoek kan dus worden gezegd dat sociale cohesie en eenzaamheid een verband met elkaar vertonen. Dit ligt in lijn met de theorie uit het theorie-hoofdstuk.

Voor hypothese 3 zijn geen ondersteunende gegevens te vinden: eenzaamheid heeft geen sterke invloed op crimineel gedrag. Het gevonden verband was zwak en niet significant. Voor

hypothese 5 lijkt het verband wel het verwachte positieve verband te vertonen: wanneer een individu zich eenzamer voelt, vertoont deze vaker probleemgedrag. De helling is echter niet heel groot en ook niet significant.

Concluderend kon het verwachte mediatie-effect tussen crimineel gedrag, sociale cohesie en eenzaamheid niet op basis van deze resultaten worden aangetoond. De effecten van sociale cohesie op crimineel gedrag & eenzaamheid daarentegen wel.

Het uitblijven van dit mediatie-effect zou erop kunnen wijzen dat eenzaamheid maar in weinig gevallen aanzet tot crimineel gedrag. Mijn speculatie is dat eenzaamheid niet wordt ervaren als een vrijheid om maar te doen en te laten wat deze persoon wil, maar eerder als een gevangenis waarin men bang is om de sociale contacten die deze persoon *nog wel* heeft te verliezen. Dit zou kunnen leiden tot het *niet* af willen wijken van de norm, niet tot het expres breken van deze norm. Wanneer dit zo zou zijn, zou een eenzaam persoon dus inderdaad minder probleem- en crimineel gedrag vertonen. Verder onderzoek zou deze relatie echter uit moeten wijzen.

Discussie

Dan rest de vraag, tot in hoeverre zijn deze uitkomsten betrouwbaar en tot in hoeverre kunnen we dit tot de Nederlandse populatie betrekken?

Ten eerste zijn er limitaties die ertoe leiden dat de resultaten aangaande probleemgedrag voorzichtiger geïnterpreteerd moeten worden. Zo zijn de assumpties van de lineaire regressie op meerdere onderdelen geschonden. Hiervoor zijn maatregelen getroffen door het verlagen van het betrouwbaarheidsniveau. Hiervoor is dus enigszins gecorrigeerd. Omdat de schaal zelf ook onbetrouwbaar is, moet er echter nog steeds voorzichtig met de generaliseerbaarheid van de resultaten worden omgegaan. Vanaf een Cronbach's Alpha van 0,600 wordt een schaal gezien als betrouwbaar; de schaal probleemgedrag zelf heeft een Cronbach's Alpha van 0,523. Dit verschil is klein genoeg om nog enige uitspraken over probleemgedrag te kunnen doen,

maar echt onomstotelijke conclusies kunnen hieruit niet getrokken worden. Zo zou het effect tussen sociale cohesie en probleemgedrag een stukje lager of hoger uit kunnen vallen of zou de relatie tussen eenzaamheid en probleemgedrag sterker kunnen zijn. Verder onderzoek naar de relaties tussen sociale cohesie, eenzaamheid en probleemgedrag zal hierdoor gedaan moeten worden. Voor het onderzoeken van probleemgedrag schijnt de schaal van 'Primary Prevention Awareness, Attitude, and Usage Scales'-schaal van Swisher (1983) een Cronbach's alpha van tussen de 7,5 en 8 te hebben. Deze zou ik aanraden voor volgend onderzoek.

Ten tweede moet er rekening mee gehouden worden t dat er geen algemene waarden te vinden waren over de sociale cohesie in de wijk zelf. De schaal is gebaseerd op het gevoel van sociale cohesie binnen de wijk van de respondent zelf. Dit is iets wat vanuit een eigen perspectief fout kan worden ingeschat. Zo zou een respondent wanneer deze eenzaam is het gevoel kunnen krijgen dat de hele wijk niet met elkaar overweg kan. De respondent zelf ziet namelijk niks terug van deze cohesie binnen de wijk. Andersom kan natuurlijk ook, wanneer een respondent veel contacten heeft binnen de wijk, voelt deze zich waarschijnlijk minder eenzaam. Deze respondent zou de sociale cohesie dan juist heel hoog in kunnen schatten. In praktijk zou de sociale cohesie anders kunnen liggen dan de respondenten zelf inschatten. Voor een volgend onderzoek zou men multi-level analyse kunnen gebruiken om dit te vermijden. Dit is echter boven het niveau van dit onderzoek.

Ten derde is er een overrepresentatie aan Turks- en Marokkaans-Nederlandse inwoners. Nederlanders met een migratieachtergrond hebben meestal andere kenmerken dan Nederlanders zonder migratieachtergrond. Deze doelgroep voelt de sociale cohesie minder hoog in de eigen wijk of voelen zich eenzamer omdat deze zich minder goed kunnen mengen met Nederlanders zonder migratie-achtergrond. Zo is het bekend dat kinderen van Nederlanders met migratieachtergrond hun 'sense of belonging', het gevoel dat je ergens hoort,

een stuk lager ervaren (Kirova, 2001). Ook senioren met migratieachtergrond voelen zich vaker eenzaam en sociaal geïsoleerd (Johnson et al., 2019). Wanneer hetzelfde onderzoek gedaan zou worden met een doelgroep die uit een samenstelling bestaat die representatief zou zijn aan de Nederlandse samenleving, zouden de onderzoeksresultaten anders kunnen liggen.

Het laatste punt van discussie is dat de literatuur over sociale cohesie niet eenduidig is. Veel onderzoekers gebruiken verschillende woorden voor sociale cohesie of gebruiken begrippen die hier sterk aan verwant zijn. Begrippen als sociaal maatschappelijke betrokkenheid (binnen de wijk), sociaal kapitaal, sociaal vertrouwen, sociale cohesie en meer worden door elkaar heen gebruikt, zonder dat het duidelijk wordt gedefinieerd waar de grens tussen verschillende termen ligt. Deze onduidelijkheid maakt het moeilijk om literatuur te vinden die bij elkaar hoort en ook om deze literatuur aan elkaar te verbinden. Dit is onhandig en tevens verwarrend voor onderzoekers die op deze kennis willen voortbouwen. Vandaar dat ik wil pleiten voor een meer coherente opstelling van de verschillende termen vanuit de sociologische gemeenschap. Mijn voorstel is een soort woordenboek, specifiek voor de sociale wetenschappen met daarin definities en betrouwbare schalen die duidelijk bij elkaar horen. Dit vergt de tijd en inzet van vele verschillende sociale wetenschappers, maar mijn vermoeden is dat dit het veld van de sociale wetenschap uiterst vooruit zou helpen.

Literatuur

- Akçomak, S., & Ter Weel, B. (2012). The impact of social capital on crime: Evidence from the Netherlands. *Regional Science and Urban Economics*, 42(1–2), 323–340.
<https://doi.org/10.1016/j.regsciurbeco.2011.09.008>
- Bernasco, W., Block, R., & Ruiters, S. (2012). Go where the money is: modeling street robbers' location choices. *Journal of Economic Geography*, 13(1), 119–143.
<https://doi.org/10.1093/jeg/lbs005>
- Bureau of Justice Statistics U.S. Department of Justice, Berzofsky, M., Smiley-McDonald, H., Moore, A., & Krebs, C. (2014, januari). *Measuring Socioeconomic Status (SES) in the NCVS: Background, Options, and Recommendations*. RTI International.
https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/measuring_ses-paper_authorship_corrected.pdf
- Burssens, E. G. (2019). *Turbulente leerlingen* (1ste editie). Gompel&Svacina Bv.
https://books.google.nl/books?hl=en&lr=&id=_7-qDwAAQBAJ&oi=fnd&pg=PA13&dq=probleem+gedrag+en+criminaliteit&ots=pcfzQOsjaD&sig=oHMq0aLvyITQvDQiooMiBFUBbhY&redir_esc=y#v=onepage&q=probleem%20gedrag%20en%20criminaliteit&f=false
- Block, C., Blokland, A. A., van der Werff, C., van Os, R., & Nieuwbeerta, P. (2010). Long-Term Patterns of Offending in Women. *Feminist Criminology*, 5(1), 73-107.
- Centraal Bureau voor de Statistiek. (2020, 27 maart). *Bijna 1 op de 10 Nederlanders voelde zich sterk eenzaam in 2019*. Geraadpleegd op 10 februari 2022, van

<https://www.cbs.nl/nl-nl/nieuws/2020/13/bijna-1-op-de-10-nederlanders-voelde-zich-sterk-eenzaam-in-2019>

- Cramm, J. M., Van Dijk, H. M., & Nieboer, A. P. (2012). The Importance of Neighborhood Social Cohesion and Social Capital for the Well Being of Older Adults in the Community. *The Gerontologist*, 53(1), 142–152. <https://doi.org/10.1093/geront/gns052>
- Duncan, G. M., Berry, W. D., & Feldman, S. (1986). Multiple Regression in Practice. *Journal of Marketing Research*, 23(3), 309. <https://doi.org/10.2307/3151494>
- Dijkstra, J. C., & Veenstra, R. (2005). Jong probleemgedrag als voorspeller voor latere criminaliteit. *Tijdschrift voor Criminologie*, 03. https://www-bjutijdschriften-nl.proxy-ub.rug.nl/tijdschrift/tijdschriftcriminologie/2005/03/TvC_2005_048_003_006
- Een tegen Eenzaamheid. (2021, 26 mei). Over eenzaamheid. Geraadpleegd op 15 februari 2022, van <https://www.eentegeneenzaamheid.nl/over-eenzaamheid/over-eenzaamheid/>
- Encyclo, - Thesaurus Politiekunde. (2012, 4 januari). [encyclo.nl](https://www.encyclo.nl). Geraadpleegd op 15 februari 2022, van <https://www.encyclo.nl/lokaal/10663>
- Fergusson, D., Swain-Campbell, N., & Horwood, J. (2004). How does childhood economic disadvantage lead to crime? *Journal of Child Psychology and Psychiatry*, 45(5), 956–966. <https://doi.org/10.1111/j.1469-7610.2004.t01-1-00288.x>
- Hans Boutellier en Erik van Marissing. (2011). Veiligheidsarrangementen in IJburg. *Tijdschrift voor Veiligheid* (10) 1, 59-68.
- Johnson, S., Bacsu, J., McIntosh, T., Jeffery, B., & Novik, N. (2019). Social isolation and loneliness among immigrant and refugee seniors in Canada: a scoping review. *International Journal of Migration, Health and Social Care*, 15(3), 177–190. <https://doi.org/10.1108/ijmhsc-10-2018-0067>

- Kalff, A. C. (2001). Neighbourhood level and individual level SES effects on child problem behaviour: a multilevel analysis. *Journal of Epidemiology & Community Health*, 55(4), 246–250. <https://doi.org/10.1136/jech.55.4.246>
- Kirova, A. (2001). Loneliness in Immigrant Children. *Childhood Education*, 77(5), 260. <https://link.gale.com/apps/doc/A76510420/ITOF?u=groning&sid=bookmark-ITOF&xid=279c17ac>
- Ministerie van Volksgezondheid, Welzijn en Sport. (2018, maart). Actieprogramma - Eén tegen eenzaamheid. <https://open.overheid.nl/repository/ronl-fbda2b9c-f4f5-474e-a6d8-72fa311fb08b/1/pdf/Actieprogramma%20Eenzaamheid.pdf>
- Lammers, M., Menting, B., Ruiter, S., & BERNASCO, W. (2015). Biting once, twice: The influence of prior on subsequent crime location choice. *Criminology*, 53(3), 309–329. <https://doi.org/10.1111/1745-9125.12071>
- Lammers, S. M. M., Soe-Agnie, S. E., De Haan, H. A., Bakkum, G. A. M., Pomp, E. R., & Nijman, H. L. I. (2014). Middelengebruik en criminaliteit: Een overzicht. *Tijdschrift voor Psychiatrie*, 56(1), 32–39. <https://repository.ubn.ru.nl/bitstream/handle/2066/134393/134393.pdf>
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674–701. <https://doi.org/10.1037/0033-295x.100.4.674>
- Nagin, D. S. (1998). Criminal Deterrence Research at the Outset of the Twenty-First Century. *Crime and Justice*, 23, 1–42. <https://doi.org/10.1086/449268>
- Rosenfeld, R., Baumer, E. P., & Messner, S. F. (2001). Social Capital and Homicide. *Social Forces*, 80(1), 283–310. <https://doi.org/10.1353/sof.2001.0086>
- Rijksinstituut voor Volksgezondheid en Milieu. (2012). *Gezondheid | Volksgezondheid Toekomst Verkenning*. Volksgezondheid Toekomst Verkenning. Geraadpleegd op 4

maart 2022, van

<https://www.vtv2018.nl/gezondheid#:~:text=In%20totaal%20neemt%20het%20aantal,in%20de%20toekomst%20vaker%20voor.>

- Sampson, R. J., & Groves, W. B. (1989). Community Structure and Crime: Testing Social-Disorganization Theory. *American Journal of Sociology*, 94(4), 774–802.
<https://doi.org/10.1086/229068>
- Senthilnathan, S. (2019). Usefulness of Correlation Analysis. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.3416918>
- Sieben, I. (2017). Child-rearing values: The impact of intergenerational class mobility. *International Sociology*, 32(3), 369–390. <https://doi.org/10.1177/0268580917693954>
- Skardhamar, T., Savolainen, J., Aase, K. N., & Lyngstad, T. H. (2015). Does Marriage Reduce Crime? *Crime and Justice*, 44(1), 385–446. <https://doi.org/10.1086/681557>
- Steenbeek, W., & Hipp, J. R. (2011). A Longitudinal Test Of Social Disorganization Theory:a Longitudinal Test Of Social Disorganization Theory: Feedback Effects Among Cohesion, Social Control, And Disorder* Feedback Effects Among Cohesion, Social Control, And Disorder*. *Criminology*, 49(3), 833–871.
<https://doi.org/10.1111/j.1745-9125.2011.00241.x>
- Stuart, B. A., & Taylor, E. J. (2021). The Effect of Social Connectedness on Crime: Evidence from the Great Migration. *The Review of Economics and Statistics*, 103(1), 18–33. https://doi.org/10.1162/rest_a_00860
- Trammell, R., Rundle, M., & Borrego, A. R. (2020). Anger, Frustration, and Snitching: Inmates Describe Structured Isolation in a High-Tech Prison. *Deviant Behavior*, 42(9), 1067–1085. <https://doi.org/10.1080/01639625.2020.1720936>
- Vangelisti, A. L., & Perlman, D. (2006). *The Cambridge Handbook of Personal Relationships - Chapter 26*. Cambridge University Press.

<https://research.vu.nl/ws/portalfiles/portal/2173239/2006+Handbook+Loneliness+dJG+vT+D.pdf>

- Verrycken, R. (2020, 3 oktober). 'We leven in een eeuw van eenzaamheid'. De Tijd. Geraadpleegd op 4 maart 2022, van <https://www.tijd.be/dossiers/coronavirus/we-leven-in-een-eeuw-van-eenzaamheid/10255796.html>
- Wadsworth, T. (2006). The Meaning of Work: Conceptualizing the Deterrent Effect of Employment on Crime among Young Adults. *Sociological Perspectives*, 49(3), 343–368. <https://doi.org/10.1525/sop.2006.49.3.343>
- Yu, R., Leung, G., Chan, J., Yip, B. H. K., Wong, S., Kwok, T., & Woo, J. (2020). Neighborhood Social Cohesion Associates with Loneliness Differently among Older People According to Subjective Social Status. *The journal of nutrition, health & aging*, 25(1), 41–47. <https://doi.org/10.1007/s12603-020-1496-z>

Bijlage 1

Selectie niet-Missing variabelen

* MISSING DATA

INDICATOREN.

FREQUENCIES VARIABLES=Inkomen Cat_opleiding w1csex w1cage w1fb1 w1fa35
/ORDER=ANALYSIS.

COMPUTE m1=0.

IF MISSING(Schaal_SocCohesie) m1=1.

COMPUTE m2=0.

IF MISSING(Schaal_Eenzaamheid) m2=1.

COMPUTE m3=0.

IF MISSING(Schaal_ProbleemGedrag) m3=1.

COMPUTE m4=0.

IF MISSING(Binair_SchaalCrimineelGedrag) m4=1.

COMPUTE m5=0.

IF MISSING(Geslacht_Binair) m5=1.

COMPUTE m6=0.

IF MISSING(Werkend_Binair) m6=1.

COMPUTE m7=0.

IF MISSING(Partner_Binair) m7=1.

COMPUTE m8=0.

IF MISSING(w1cage) m8=1.

COMPUTE m9=0.

IF MISSING(Inkomen) m9=1.

COMPUTE m10=0.

IF MISSING(Cat_Opleiding) m10=1.

EXECUTE.

COMPUTE miss = m1+m2+m3+m4+m5+m6+m7+m8+m9+m10.

EXECUTE.

FREQUENCIES VARIABLES=miss

/ORDER=ANALYSIS.

DATASET COPY CompleteData.

DATASET ACTIVATE CompleteData.

FILTER OFF.

USE ALL.

SELECT IF (miss<4).
EXECUTE.

Sociale Cohesie

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES=w1fk1a w1fk1b w1fk1c w1fk1d w1fk1e w1fk1f
/ORDER=ANALYSIS.

		Statistics					
		kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? mensen in	kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? veel mense	kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen
N	Valid	5312	5312	5312	5312	5312	5312
	Missing	0	0	0	0	0	0

mensen in deze buurt groeten elkaar

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	3046	57,3	57,3	57,3
	klopt een beetje	1858	35,0	35,0	92,3
	klopt niet echt	308	5,8	5,8	98,1
	klopt helemaal niet	100	1,9	1,9	100,0
Total		5312	100,0	100,0	

de mensen in deze buurt kan men vertrouwen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	1940	36,5	36,5	36,5
	klopt een beetje	2378	44,8	44,8	81,3
	klopt niet echt	833	15,7	15,7	97,0
	klopt helemaal niet	161	3,0	3,0	100,0
Total		5312	100,0	100,0	

de mensen in deze buurt kunnen over het algemeen goed met elkaar opschieten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	2581	48,6	48,6	48,6
	klopt een beetje	2166	40,8	40,8	89,4
	klopt niet echt	477	9,0	9,0	98,3
	klopt helemaal niet	88	1,7	1,7	100,0
	Total	5312	100,0	100,0	

veel mensen kennen elkaar in deze buurt

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	2792	52,6	52,6	52,6
	klopt een beetje	1720	32,4	32,4	84,9
	klopt niet echt	681	12,8	12,8	97,8
	klopt helemaal niet	119	2,2	2,2	100,0
	Total	5312	100,0	100,0	

de mensen hier willen elkaar graag helpen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	2054	38,7	38,7	38,7
	klopt een beetje	2283	43,0	43,0	81,6
	klopt niet echt	839	15,8	15,8	97,4
	klopt helemaal niet	136	2,6	2,6	100,0
	Total	5312	100,0	100,0	

de mensen in de buurt zouden er iets van zeggen als jongeren overlast bezorgen in de buurt

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	klopt helemaal	2832	53,3	53,3	53,3
	klopt een beetje	1654	31,1	31,1	84,5
	klopt niet echt	625	11,8	11,8	96,2
	klopt helemaal niet	201	3,8	3,8	100,0
	Total	5312	100,0	100,0	

DESCRIPTIVES VARIABLES=w1fk1a w1fk1b w1fk1c w1fk1d w1fk1e w1fk1f
 /STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics									
	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? mensen in	5019	1	4	1,52	,687	1,305	,035	1,607	,069
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	5019	1	4	1,85	,784	,638	,035	-,107	,069
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	5019	1	4	1,63	,710	,923	,035	,498	,069
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? veel mense	5019	1	4	1,64	,783	,991	,035	,176	,069
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	5019	1	4	1,82	,780	,638	,035	-,205	,069
kunt u aangeven of deze uitspraken wel of niet kloppen voor uw buurt? de mensen	5019	1	4	1,66	,827	1,106	,035	,450	,069
Valid N (listwise)	5019								

Bewerkingen / Operationalisatie

RECODE w1fk1a w1fk1b w1fk1c w1fk1d w1fk1e
 (1=4) (2=3) (3=2) (4=1) INTO W1_SocCoh_A W1_SocCoh_B W1_SocCoh_C
 W1_SocCoh_D W1_SocCoh_E.
 EXECUTE.

COMPUTE Schaal_SocCohesie=(W1_SocCoh_A + W1_SocCoh_B + W1_SocCoh_C +
 W1_SocCoh_D + W1_SocCoh_E +
 W1_SocCoh_F) / 6.

Uiteindelijke variabele

FREQUENCIES VARIABLES=Schaal_SocCohesie
 /ORDER=ANALYSIS.

Statistics		
Schaal_SocCohesie		
N	Valid	5019
	Missing	0

Schaal_SocCohesie

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	11	,2	,2	,2
	1,17	6	,1	,1	,3
	1,33	10	,2	,2	,5
	1,50	13	,3	,3	,8
	1,67	28	,6	,6	1,4
	1,83	20	,4	,4	1,8
	2,00	52	1,0	1,0	2,8
	2,17	83	1,7	1,7	4,4
	2,33	147	2,9	2,9	7,4
	2,50	177	3,5	3,5	10,9
	2,67	275	5,5	5,5	16,4
	2,83	363	7,2	7,2	23,6
	3,00	537	10,7	10,7	34,3
	3,17	427	8,5	8,5	42,8
	3,33	410	8,2	8,2	51,0
	3,50	489	9,7	9,7	60,7
	3,67	496	9,9	9,9	70,6
	3,83	571	11,4	11,4	82,0
	4,00	904	18,0	18,0	100,0
	Total		5019	100,0	100,0

DESCRIPTIVES VARIABLES=Schaal_SocCohesie
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Schaal_SocCohesie	5019	1,00	4,00	3,3134	,57613	-,784	,035	,392	,069
Valid N (listwise)	5019								

RELIABILITY

/VARIABLES=W1_SocCoh_A W1_SocCoh_B W1_SocCoh_C W1_SocCoh_D W1_SocCoh_E
W1_SocCoh_F
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR ANOVA
/SUMMARY=TOTAL.

Case Processing Summary

		N	%
Cases	Valid	5019	100,0
	Excluded ^a	0	,0
	Total	5019	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,849	,852	6

Inter-Item Correlation Matrix

	W1_SocCoh_A	W1_SocCoh_B	W1_SocCoh_C	W1_SocCoh_D	W1_SocCoh_E	W1_SocCoh_F
W1_SocCoh_A	1,000	,499	,551	,534	,539	,359
W1_SocCoh_B	,499	1,000	,599	,398	,569	,356
W1_SocCoh_C	,551	,599	1,000	,538	,614	,396
W1_SocCoh_D	,534	,398	,538	1,000	,577	,389
W1_SocCoh_E	,539	,569	,614	,577	1,000	,431
W1_SocCoh_F	,359	,356	,396	,389	,431	1,000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
W1_SocCoh_A	16,3969	8,843	,645	,431	,823
W1_SocCoh_B	16,7300	8,485	,624	,445	,826
W1_SocCoh_C	16,5154	8,500	,712	,527	,810
W1_SocCoh_D	16,5226	8,466	,630	,437	,824
W1_SocCoh_E	16,6995	8,128	,723	,532	,806
W1_SocCoh_F	16,5388	8,871	,486	,239	,854

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
19,8807	11,949	3,45677	6

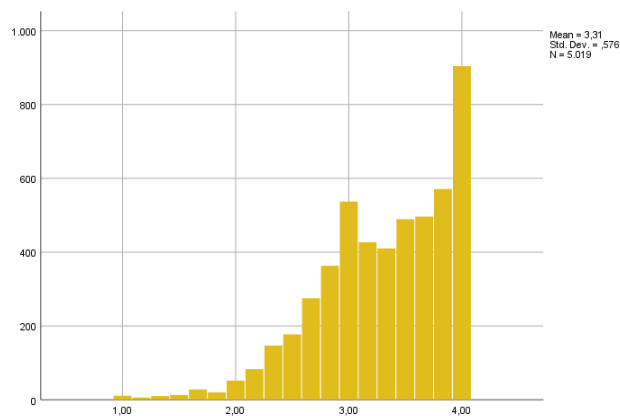
ANOVA

		Sum of Squares	df	Mean Square	F	Sig
Between People		9993,585	5018	1,992		
Within People	Between Items	394,030	5	78,806	261,928	,000
	Residual	7548,804	25090	,301		
	Total	7942,833	25095	,317		
Total		17936,419	30113	,596		

Grand Mean = 3,3134

Beschrijving

Sociale Cohesie heeft een verdeling van 1 tot 4, waarbij 4 staat voor een hoge ervaring van sociale cohesie in de buurt door de respondent. Aangezien het gemiddelde 3,31 is (tabel 1), kan men concluderen dat de gemiddelde respondent een vrij hoge sociale cohesie in de eigen buurt ervaart. Bij nadere inspectie van de data (figuur 2) is het opvallend dat deze data erg rechts verdeeld is met enige uitschieters naar links, wat deze hoge sociale cohesie binnen de dataset ondersteunt. Deze linkse scheefheid van de data is in enige mate ook terug te zien in de skewness van deze schaal (-0,784). Een hoge positieve skewness doelt op een sterk rechtsscheve verdeling, een sterk negatieve skewness duidt op een sterke linksscheve verdeling. Hierbij wijst een hoge kurtosis wijst op een sterk piekende grafiek, een negatieve kurtosis op een platte verdeling.



Figuur 2: Histogram van de verdeling van de schaal 'Sociale cohesie'

Eenzaamheid

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES=w1scf7a w1scf7b w1scf7c w1scf7d w1scf7e w1scf7f

/ORDER=ANALYSIS.

		Statistics					
		er zijn voldoende mensen met wie ik me nauw verbonden voel	ik ervaar een leegte om me heen	ik heb veel mensen op wie ik volledig kan vertrouwen	er zijn genoeg mensen op wie ik in geval van narigheid kan terugvallen	ik mis mensen om me heen	vaak voel ik me in de steek gelaten
N	Valid	4850	4843	4853	4846	4849	4851
	Missing	462	469	459	466	463	461

er zijn voldoende mensen met wie ik me nauw verbonden voel

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	1493	28,1	30,8	30,8
	van toepassing	2766	52,1	57,0	87,8
	niet van toepassing	505	9,5	10,4	98,2
	helemaal niet van toepassing	86	1,6	1,8	100,0
	Total	4850	91,3	100,0	
Missing	System	462	8,7		
Total		5312	100,0		

ik ervaar een leegte om me heen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	128	2,4	2,6	2,6
	van toepassing	678	12,8	14,0	16,6
	niet van toepassing	2537	47,8	52,4	69,0
	helemaal niet van toepassing	1500	28,2	31,0	100,0
	Total	4843	91,2	100,0	
Missing	System	469	8,8		
Total		5312	100,0		

ik heb veel mensen op wie ik volledig kan vertrouwen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	1211	22,8	25,0	25,0
	van toepassing	2841	53,5	58,5	83,5
	niet van toepassing	690	13,0	14,2	97,7
	helemaal niet van toepassing	111	2,1	2,3	100,0
	Total	4853	91,4	100,0	
Missing	System	459	8,6		
Total		5312	100,0		

er zijn genoeg mensen op wie ik in geval van narigheid kan terugvallen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	1319	24,8	27,2	27,2
	van toepassing	2897	54,5	59,8	87,0
	niet van toepassing	535	10,1	11,0	98,0
	helemaal niet van toepassing	95	1,8	2,0	100,0
	Total	4846	91,2	100,0	
Missing	System	466	8,8		
Total		5312	100,0		

ik mis mensen om me heen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	232	4,4	4,8	4,8
	van toepassing	855	16,1	17,6	22,4
	niet van toepassing	2444	46,0	50,4	72,8
	helemaal niet van toepassing	1318	24,8	27,2	100,0
	Total	4849	91,3	100,0	
Missing	System	463	8,7		
Total		5312	100,0		

vaak voel ik me in de steek gelaten

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	zeer van toepassing	123	2,3	2,5	2,5
	van toepassing	491	9,2	10,1	12,7
	niet van toepassing	2487	46,8	51,3	63,9
	helemaal niet van toepassing	1750	32,9	36,1	100,0
	Total	4851	91,3	100,0	
Missing	System	461	8,7		
Total		5312	100,0		

DESCRIPTIVES VARIABLES=w1scf7a w1scf7b w1scf7c w1scf7d w1scf7e w1scf7f

/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics									
	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
er zijn voldoende mensen met wie ik me nauw verbonden voel	4846	1	4	1,83	,674	,564	,035	,556	,070
ik ervaar een leegte om me heen	4839	1	4	3,12	,736	-,584	,035	,192	,070
ik heb veel mensen op wie ik volledig kan vertrouwen	4847	1	4	1,94	,692	,496	,035	,431	,070
er zijn genoeg mensen op wie ik in geval van narigheid kan terugvallen	4843	1	4	1,88	,667	,538	,035	,696	,070
ik mis mensen om me heen	4845	1	4	3,00	,799	-,561	,035	-,031	,070
vaak voel ik me in de steek gelaten	4847	1	4	3,21	,720	-,742	,035	,568	,070
Valid N (listwise)	4832								

Bewerkingen / Operationalisaties

RECODE w1scf7b w1scf7e w1scf7f

(1=4) (2=3) (3=2) (4=1) (ELSE=Copy) INTO W1_Eenzaamheid_B W1_Eenzaamheid_E
W1_Eenzaamheid_F.

EXECUTE.

RECODE w1scf7a w1scf7c w1scf7d (ELSE=Copy) INTO W1_Eenzaamheid_A

W1_Eenzaamheid_C W1_Eenzaamheid_D.

EXECUTE.

COMPUTE Schaal_Eenzaamheid=(W1_Eenzaamheid_A + W1_Eenzaamheid_B +

W1_Eenzaamheid_C +

W1_Eenzaamheid_D + W1_Eenzaamheid_E + W1_Eenzaamheid_F) / 6.

Uiteindelijke variabele

FREQUENCIES VARIABLES= Schaal_Eenzaamheid

/ORDER=ANALYSIS.

Statistics		
Schaal_Eenzaamheid		
N	Valid	4832
	Missing	188

Schaal_Eenzaamheid

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	386	7,7	8,0	8,0
	1,17	212	4,2	4,4	12,4
	1,33	243	4,8	5,0	17,4
	1,50	582	11,6	12,0	29,4
	1,67	363	7,2	7,5	37,0
	1,83	463	9,2	9,6	46,5
	2,00	1080	21,5	22,4	68,9
	2,17	478	9,5	9,9	78,8
	2,33	347	6,9	7,2	86,0
	2,50	333	6,6	6,9	92,9
	2,67	139	2,8	2,9	95,7
	2,83	83	1,7	1,7	97,5
	3,00	64	1,3	1,3	98,8
	3,17	33	,7	,7	99,5
	3,33	14	,3	,3	99,8
	3,50	4	,1	,1	99,8
	3,67	2	,0	,0	99,9
	3,83	2	,0	,0	99,9
	4,00	4	,1	,1	100,0
		Total	4832	96,3	100,0
Missing	System	188	3,7		
	Total	5020	100,0		

DESCRIPTIVES VARIABLES= Schaal_Eenzaamheid
 /STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Schaal_Eenzaamheid	4832	1,00	4,00	1,8866	,49909	,183	,035	,066	,070
Valid N (listwise)	4832								

RELIABILITY

/VARIABLES=W1_Eenzaamheid_A W1_Eenzaamheid_B W1_Eenzaamheid_C
 W1_Eenzaamheid_D W1_Eenzaamheid_E
 W1_Eenzaamheid_F
 /SCALE('ALL VARIABLES') ALL
 /MODEL=ALPHA
 /STATISTICS=DESCRIPTIVE SCALE CORR ANOVA
 /SUMMARY=TOTAL.

Case Processing Summary

		N	%
Cases	Valid	4832	96,3
	Excluded ^a	188	3,7
	Total	5020	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,788	,789	6

Item Statistics

	Mean	Std. Deviation	N
W1_Eenzaamheid_A	1,8324	,67455	4832
W1_Eenzaamheid_B	1,8822	,73587	4832
W1_Eenzaamheid_C	1,9394	,69228	4832
W1_Eenzaamheid_D	1,8771	,66789	4832
W1_Eenzaamheid_E	1,9986	,79911	4832
W1_Eenzaamheid_F	1,7899	,71991	4832

Inter-Item Correlation Matrix

	W1_Eenzaamheid_A	W1_Eenzaamheid_B	W1_Eenzaamheid_C	W1_Eenzaamheid_D	W1_Eenzaamheid_E	W1_Eenzaamheid_F
W1_Eenzaamheid_A	1,000	,211	,531	,540	,217	,255
W1_Eenzaamheid_B	,211	1,000	,290	,249	,558	,579
W1_Eenzaamheid_C	,531	,290	1,000	,686	,240	,316
W1_Eenzaamheid_D	,540	,249	,686	1,000	,224	,278
W1_Eenzaamheid_E	,217	,558	,240	,224	1,000	,596
W1_Eenzaamheid_F	,255	,579	,316	,278	,596	1,000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
W1_Eenzaamheid_A	9,4872	6,839	,474	,347	,770
W1_Eenzaamheid_B	9,4373	6,408	,542	,414	,755
W1_Eenzaamheid_C	9,3802	6,482	,569	,522	,749
W1_Eenzaamheid_D	9,4425	6,649	,543	,516	,755
W1_Eenzaamheid_E	9,3210	6,249	,521	,426	,762
W1_Eenzaamheid_F	9,5296	6,320	,588	,461	,744

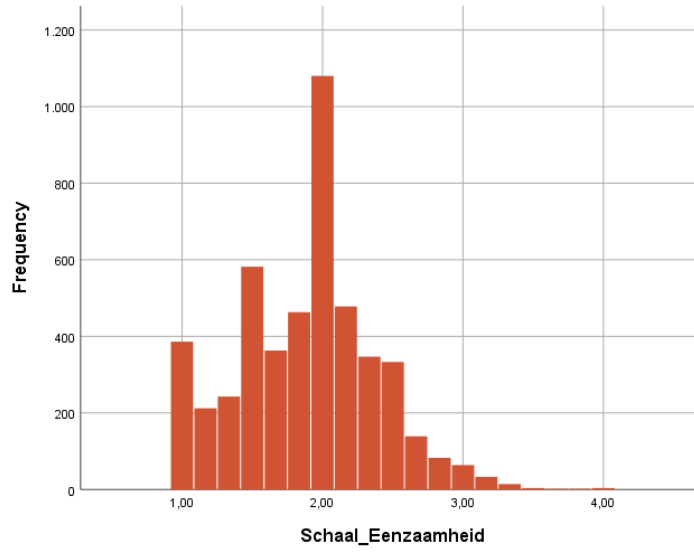
<0,0005

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
11,3195	8,967	2,99453	6

Beschrijving

In tabel 1 is te zien dat de gemiddelde respondent een 11,32 op eenzaamheid scoort, op een schaal van 1 tot 4. Dit betekent dat de gemiddelde respondent zich best wel eenzaam voelt. Dit in tegenstelling tot het beeld wat ik persoonlijk veel waarschijnlijker vond: de gemiddelde mens voelt zich niet eenzaam en een kleinere groep af en toe wel. Eenzaamheid schijnt redelijk normaal verdeeld te zijn (figuur 3), alhoewel het daarbij wel opvallend dat er erg veel waarnemingen worden gedaan bij een score van 2,0. Dit zorgt niet voor een erg hoge kurtosis (0,066).



Figuur 3: Histogram van de verdeling van de schaal 'Eenzaamheid'

Probleemgedrag

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES=w1scf10a w1scf10b w1scf10c w1scf10d

/ORDER=ANALYSIS.

		Statistics			
		afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer	afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	afgelopen maand gedaan: 's avonds of 's nachts op straat hangen
N	Valid	4846	4857	4855	4858
	Missing	466	455	457	454

afgelopen maand gedaan: verzuimen van werk of school zonder goede reden

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4195	79,0	86,6	86,6
	1 keer	432	8,1	8,9	95,5
	2-3 keer	149	2,8	3,1	98,6
	4 keer of vaker	70	1,3	1,4	100,0
	Total	4846	91,2	100,0	
Missing	System	466	8,8		
Total		5312	100,0		

afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer dan 10 glazen)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	3787	71,3	78,0	78,0
	1 keer	630	11,9	13,0	90,9
	2-3 keer	290	5,5	6,0	96,9
	4 keer of vaker	150	2,8	3,1	100,0
	Total	4857	91,4	100,0	
Missing	System	455	8,6		
Total		5312	100,0		

afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4521	85,1	93,1	93,1
	1 keer	151	2,8	3,1	96,2
	2-3 keer	71	1,3	1,5	97,7
	4 keer of vaker	112	2,1	2,3	100,0
	Total	4855	91,4	100,0	
Missing	System	457	8,6		
Total		5312	100,0		

afgelopen maand gedaan: 's avonds of 's nachts op straat hangen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4497	84,7	92,6	92,6
	1 keer	169	3,2	3,5	96,0
	2-3 keer	122	2,3	2,5	98,6
	4 keer of vaker	70	1,3	1,4	100,0
	Total	4858	91,5	100,0	
Missing	System	454	8,5		
Total		5312	100,0		

DESCRIPTIVES VARIABLES=w1scf10a w1scf10b w1scf10c w1scf10d
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	4843	1	4	1,19	,551	3,223	,035	10,653	,070
afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer	4853	1	4	1,34	,728	2,228	,035	4,237	,070
afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	4851	1	4	1,13	,528	4,453	,035	19,536	,070
afgelopen maand gedaan: 's avonds of 's nachts op straat hangen	4854	1	4	1,13	,499	4,254	,035	18,196	,070
Valid N (listwise)	4839								

Bewerkingen / Operationalisaties

COMPUTE Schaal_Probleemgedrag=(w1scf10a + w1scf10b + w1scf10c + w1scf10d) / 4.
EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES= Schaal_AanzetCrimi
/ORDER=ANALYSIS.

Statistics

Schaal_ProbleemGedrag

N	Valid	4838
	Missing	182

Schaal_ProbleemGedrag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	3161	63,0	65,3	65,3
	1,25	726	14,5	15,0	80,3
	1,50	435	8,7	9,0	89,3
	1,75	230	4,6	4,8	94,1
	2,00	121	2,4	2,5	96,6
	2,25	59	1,2	1,2	97,8
	2,50	51	1,0	1,1	98,9
	2,75	20	,4	,4	99,3
	3,00	13	,3	,3	99,5
	3,25	11	,2	,2	99,8
	3,50	4	,1	,1	99,9
	4,00	7	,1	,1	100,0
		Total	4838	96,4	100,0
Missing	System	182	3,6		
Total		5020	100,0		

DESCRIPTIVES VARIABLES=Schaal_ProbleemGedrag
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Schaal_ProbleemGedrag	4838	1,00	4,00	1,1983	,37441	2,829	,035	10,583	,070
Valid N (listwise)	4838								

RELIABILITY

```

/VARIABLES=w1scf10a w1scf10b w1scf10c w1scf10d
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR ANOVA
/SUMMARY=TOTAL.
    
```

Case Processing Summary

		N	%
Cases	Valid	4838	96,4
	Excluded ^a	182	3,6
	Total	5020	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,523	,535	4

Item Statistics

	Mean	Std. Deviation	N
afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	1,19	,552	4839
afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer	1,34	,729	4839
afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	1,13	,529	4839
afgelopen maand gedaan: 's avonds of 's nachts op straat hangen	1,13	,497	4839

Inter-Item Correlation Matrix

	afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer)	afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	afgelopen maand gedaan: 's avonds of 's nachts op straat hangen
afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	1,000	,159	,155	,243
afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer)	,159	1,000	,286	,234
afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	,155	,286	1,000	,261
afgelopen maand gedaan: 's avonds of 's nachts op straat hangen	,243	,234	,261	1,000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
afgelopen maand gedaan: verzuimen van werk of school zonder goede reden	3,60	1,586	,253	,075	,499
afgelopen maand gedaan: te veel alcohol drinken (bijvoorbeeld op een avond meer)	3,45	1,192	,326	,116	,454
afgelopen maand gedaan: drugs gebruiken (marihuana, wiet, hasj, coke, xtc, etc.)	3,66	1,513	,345	,127	,427
afgelopen maand gedaan: 's avonds of 's nachts op straat hangen	3,67	1,554	,355	,131	,424

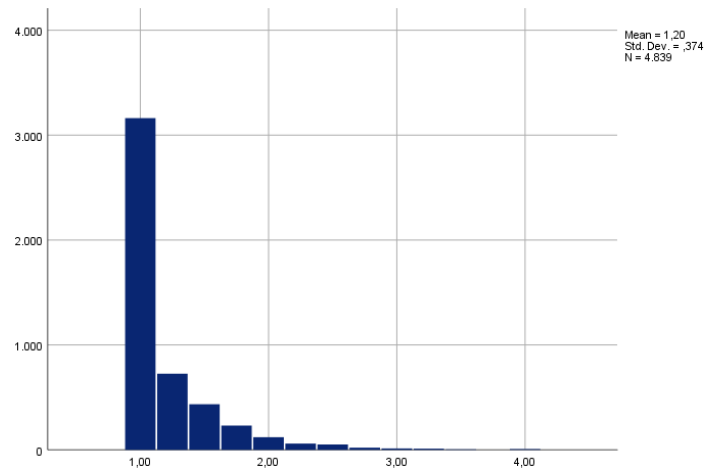
Scale Statistics

Mean	Variance	Std. Deviation	N of Items
4,79	2,242	1,497	4

Beschrijving

Probleemgedrag heeft een gemiddelde van 1,20. Op een schaal die van 1 tot 4 loopt betekent dat de gemiddelde respondent ergens tussen de 'geen' en '1 keer' probleemgedrag heeft vertoond in de afgelopen maand. Het gaat hier om een variabele met een zwaar rechtsscheve

grafiek Dit is ook te zien aan figuur 4 en de scheve skewness (2,827) en kurtosis waarden(10,57).



Figuur 4: Histogram van de verdeling van de schaal 'Probleemgedrag'

Crimineel gedrag

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES= w1scf11a w1scf11b w1scf11c w1scf11d w1scf11e
/ORDER=ANALYSIS.

Statistics

		afgelopen 12 maanden gedaan: iets gestolen van iemand of uit een winkel	afgelopen 12 maanden gedaan: met opzet dingen van anderen beschadigd of vernield	afgelopen 12 maanden gedaan: een wapen gedragen (mes, pistool)	afgelopen 12 maanden gedaan: iemand bedreigd	afgelopen 12 maanden gedaan: iemand geslagen of geschopt of meegedaan aan een ve
N	Valid	4860	4857	4858	4858	4859
	Missing	452	455	454	454	453

afgelopen 12 maanden gedaan: iets gestolen van iemand of uit een winkel

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4744	89,3	97,6	97,6
	1 keer	75	1,4	1,5	99,2
	2-3 keer	26	,5	,5	99,7
	4 keer of vaker	6	,1	,1	99,8
	wil niet zeggen	9	,2	,2	100,0
	Total	4860	91,5	100,0	
Missing	System	452	8,5		
Total		5312	100,0		

afgelopen 12 maanden gedaan: met opzet dingen van anderen beschadigd of vernield

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4755	89,5	97,9	97,9
	1 keer	65	1,2	1,3	99,2
	2-3 keer	27	,5	,6	99,8
	4 keer of vaker	4	,1	,1	99,9
	wil niet zeggen	6	,1	,1	100,0
	Total	4857	91,4	100,0	
Missing	System	455	8,6		
Total		5312	100,0		

afgelopen 12 maanden gedaan: een wapen gedragen (mes, pistool)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4753	89,5	97,8	97,8
	1 keer	36	,7	,7	98,6
	2-3 keer	24	,5	,5	99,1
	4 keer of vaker	28	,5	,6	99,7
	wil niet zeggen	17	,3	,3	100,0
	Total	4858	91,5	100,0	
Missing	System	454	8,5		
Total		5312	100,0		

afgelopen 12 maanden gedaan: iemand bedreigd

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4736	89,2	97,5	97,5
	1 keer	80	1,5	1,6	99,1
	2-3 keer	22	,4	,5	99,6
	4 keer of vaker	5	,1	,1	99,7
	wil niet zeggen	15	,3	,3	100,0
	Total	4858	91,5	100,0	
Missing	System	454	8,5		
Total		5312	100,0		

afgelopen 12 maanden gedaan: iemand geslagen of geschopt of meegedaan aan een vechtpartij

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	nooit	4664	87,8	96,0	96,0
	1 keer	129	2,4	2,7	98,6
	2-3 keer	39	,7	,8	99,4
	4 keer of vaker	10	,2	,2	99,7
	wil niet zeggen	17	,3	,3	100,0
	Total	4859	91,5	100,0	
Missing	System	453	8,5		
Total		5312	100,0		

DESCRIPTIVES VARIABLES=w1scf11a w1scf11b w1scf11c w1scf11d w1scf11e
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
afgelopen 12 maanden gedaan: iets gestolen van iemand of uit een winkel	4856	1	5	1,04	,276	9,582	,035	108,052	,070
afgelopen 12 maanden gedaan: met opzet dingen van anderen beschadigd of vernield	4854	1	5	1,03	,249	9,978	,035	118,978	,070
afgelopen 12 maanden gedaan: een wapen gedragen (mes, pistool)	4855	1	5	1,05	,360	8,473	,035	75,746	,070
afgelopen 12 maanden gedaan: iemand bedreigd	4855	1	5	1,04	,297	9,682	,035	108,308	,070
afgelopen 12 maanden gedaan: iemand geslagen of geschopt of meegedaan aan een ve	4856	1	5	1,06	,359	7,455	,035	64,749	,070
Valid N (listwise)	4852								

Bewerkingen / Operationalisaties

```
RECODE w1scf11a w1scf11b w1scf11c w1scf11d w1scf11e (5=SYSMIS) (ELSE=Copy) INTO
W1_CrimiGedr_A
```

```
W1_CrimiGedr_B W1_CrimiGedr_C W1_CrimiGedr_D W1_CrimiGedr_E.
EXECUTE.
```

```
COMPUTE Schaal_CrimineelGedrag=(W1_CrimiGedr_A + W1_CrimiGedr_B +
W1_CrimiGedr_C + W1_CrimiGedr_D +
```

```
W1_CrimiGedr_E) / 5.
EXECUTE.
```

Uiteindelijke variabele

```
FREQUENCIES VARIABLES= Schaal_CrimineelGedrag
/ORDER=ANALYSIS.
```

Statistics

Schaal_CrimineelGedrag

N	Valid	4820
	Missing	200

Schaal_CrimineelGedrag

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	4453	88,7	92,4	92,4
	1,20	191	3,8	4,0	96,3
	1,40	74	1,5	1,5	97,9
	1,60	48	1,0	1,0	98,9
	1,80	21	,4	,4	99,3
	2,00	8	,2	,2	99,5
	2,20	11	,2	,2	99,7
	2,40	7	,1	,1	99,9
	2,60	1	,0	,0	99,9
	2,80	1	,0	,0	99,9
	3,00	3	,1	,1	100,0
	3,20	2	,0	,0	100,0
	Total		4820	96,0	100,0
Missing	System	200	4,0		
Total		5020	100,0		

DESCRIPTIVES VARIABLES=Schaal_CrimineelGedrag
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Schaal_CrimineelGedrag	4820	1,00	3,20	1,0328	,15112	7,088	,035	64,904	,071
Valid N (listwise)	4820								

RELIABILITY

/VARIABLES=W1_CrimiGedr_A W1_CrimiGedr_B W1_CrimiGedr_C W1_CrimiGedr_D
W1_CrimiGedr_E
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR ANOVA

/SUMMARY=TOTAL.

Case Processing Summary

		N	%
Cases	Valid	4820	96,0
	Excluded ^a	200	4,0
	Total	5020	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,650	,667	5

Item Statistics

	Mean	Std. Deviation	N
W1_CrimiGedr_A	1,0295	,21551	4820
W1_CrimiGedr_B	1,0263	,20354	4820
W1_CrimiGedr_C	1,0340	,27800	4820
W1_CrimiGedr_D	1,0266	,19359	4820
W1_CrimiGedr_E	1,0477	,26758	4820

Inter-Item Correlation Matrix

	W1_CrimiGedr_A	W1_CrimiGedr_B	W1_CrimiGedr_C	W1_CrimiGedr_D	W1_CrimiGedr_E
W1_CrimiGedr_A	1,000	,351	,278	,290	,228
W1_CrimiGedr_B	,351	1,000	,256	,356	,316
W1_CrimiGedr_C	,278	,256	1,000	,253	,187
W1_CrimiGedr_D	,290	,356	,253	1,000	,344
W1_CrimiGedr_E	,228	,316	,187	,344	1,000

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
W1_CrimiGedr_A	4,1346	,410	,415	,186	,593
W1_CrimiGedr_B	4,1378	,407	,470	,234	,571
W1_CrimiGedr_C	4,1301	,376	,344	,128	,636
W1_CrimiGedr_D	4,1376	,418	,460	,220	,579
W1_CrimiGedr_E	4,1164	,375	,379	,170	,613

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
5,1641	,571	,75558	5

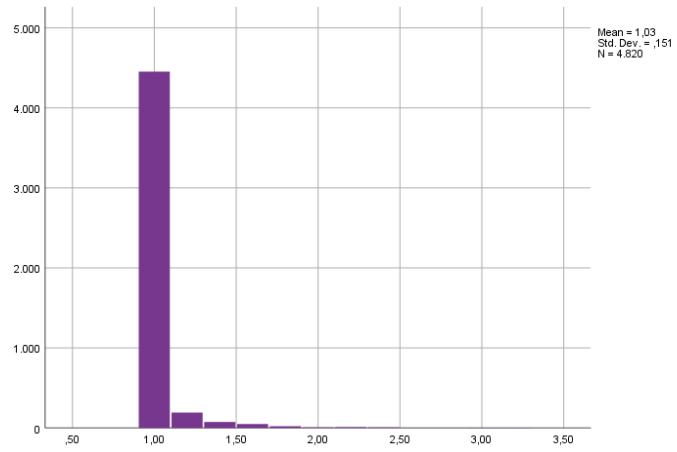
ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Between People		550,238	4819	,114		
Within People	Between Items	1,522	4	,381	9,511	,000
	Residual	771,278	19276	,040		
	Total	772,800	19280	,040		
Total		1323,038	24099	,055		

Grand Mean = 1,0328

Beschrijving

Crimineel gedrag heeft een gemiddelde score van 1,03. Dat duidt erop dat de gemiddelde respondent ergens tussen de 'nooit' en '1 keer' crimineel gedrag heeft vertoond in de afgelopen 12 maanden. Het gemiddelde ligt echter wel erg dicht bij de 1, wat erop duidt dat de meeste respondenten 'nooit' in hebben gevuld. Dit wordt ondersteund door het histogram in figuur 5, die laat zien dat de schaal heel erg rechtsscheef verdeeld is. Deze gepiektheid is ook terug te zien in de skewness (7,088) die aangeeft dat de grafiek een zware verdeling aan de rechterkant heeft. De kurtosis (106,949) van deze variabele geeft aan dat het hier gaat om een vrij zwaar gepiekte verdeling van de waarden.



Figuur 5: Histogram van de verdeling van de schaal 'Crimineel gedrag'.

Controle-variabelen¹

Beschrijving

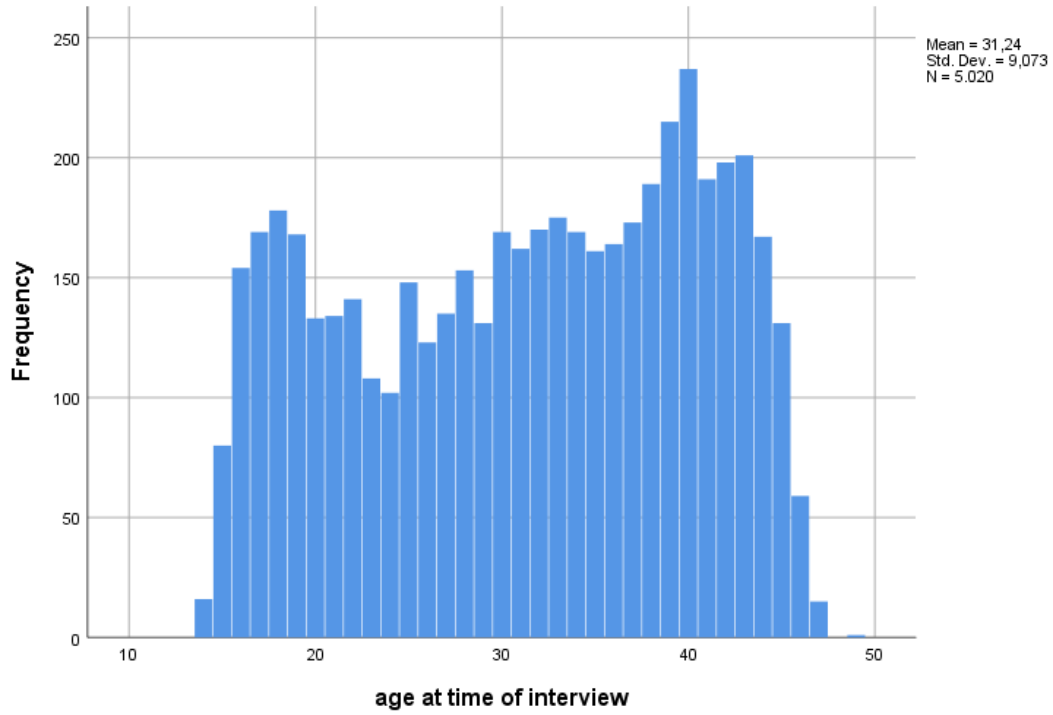
De gemiddelde respondent heeft een leeftijd van 31,23 jaar oud. De jongste respondent uit de steekproef was 14 en de oudste respondent was 49 jaar oud. Binnen de gebruikte steekproef heeft 68,4% van de respondenten een partner en was 31,6% van de respondenten single. Iets meer dan de helft had een betaalde baan (53,2%) en iets minder dan de helft had geen betaald werk (48,2%). Mannen zijn iets meer vertegenwoordigd dan vrouwen in de steekproef, maar liefst 53,2% van de steekproef was man. Bovendien scoort deze gemiddelde respondent met een score van 6,09 op inkomen, wat neerkomt op een inkomen boven de 1999 en onder de 2499 euro per maand. De gemiddelde waarde voor opleiding is een 1,64 op opleiding, wat door de categorische aard van deze vraag wijst op een opleiding tussen middelbare school-diploma en het mbo-diploma in in.

Leeftijd

GRAPH

/HISTOGRAM=w1cage.

¹ Aangezien er geen bewerkingen zijn uitgevoerd bij de variabele leeftijd is hier geen onderscheid gemaakt tussen oorspronkelijke & huidige variabele.



DESCRIPTIVES VARIABLES=w1cage
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness Statistic	Std. Error	Kurtosis Statistic	Std. Error
age at time of interview	5020	14	49	31,24	9,073	-,192	,035	-1,196	,069
Valid N (listwise)	5020								

Geslacht

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES= w1csex
/ORDER=ANALYSIS.

Statistics

sample geslacht rp

N	Valid	5312
	Missing	0

sample geslacht rp

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	man	2508	47,2	47,2	47,2
	vrouw	2804	52,8	52,8	100,0
Total		5312	100,0	100,0	

DESCRIPTIVES VARIABLES=w1csex

/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
sample geslacht rp	5312	1	2	1,53	,499	-,112	,034	-1,988	,067
Valid N (listwise)	5312								

Bewerkingen / Operationalisaties

RECODE w1csex (2=0) (ELSE=Copy) INTO Geslacht_Binair.
EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES=Geslacht_Binair
/ORDER=ANALYSIS.

Statistics

sample geslacht rp

N	Valid	5020
	Missing	0

sample geslacht rp

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	man	2347	46,8	46,8	46,8
	vrouw	2673	53,2	53,2	100,0
Total		5020	100,0	100,0	

DESCRIPTIVES VARIABLES=Geslacht_Binair
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Geslacht_Binair	5020	,00	1,00	,4675	,49899	,130	,035	-1,984	,069
Valid N (listwise)	5020								

Werkend

Oorspronkelijke variabelen

FREQUENCIES VARIABLES= w1fa35
/ORDER=ANALYSIS.

Statistics

verricht u op dit moment betaald

N	Valid	2201
	Missing	2819

verricht u op dit moment betaalde arbeid?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	1214	24,2	55,2	55,2
	nee	987	19,7	44,8	100,0
	Total	2201	43,8	100,0	
Missing	System	2819	56,2		
Total		5020	100,0		

Bewerkingen / Operationalisaties

RECODE w1fa35 (2=0) (ELSE=Copy) INTO Werkend_Binair.
EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES= Werkend_Binair
/ORDER=ANALYSIS.

Statistics

Werkend_Binair

N	Valid	2201
	Missing	2819

Werkend_Binair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	987	19,7	44,8	44,8
	1,00	1214	24,2	55,2	100,0
	Total	2201	43,8	100,0	
Missing	System	2819	56,2		
Total		5020	100,0		

Inkomen

Oorspronkelijke variabele

Statistics

wat is het netto inkomen per ma

N	Valid	5019
	Missing	1

wat is het netto inkomen per maand %1109./van u en uw partner samen?
/van u?/ %11

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	minder dan 150 euro per maand	407	8,1	8,1	8,1
	150 - 299 euro per maand	262	5,2	5,2	13,3
	300 - 499 euro per maand	223	4,4	4,4	17,8
	500 - 999 euro per maand	429	8,5	8,5	26,3
	1000 - 1499 euro per maand	650	12,9	13,0	39,3
	1500 - 1999 euro per maand	599	11,9	11,9	51,2
	2000 - 2499 euro per maand	499	9,9	9,9	61,1
	2500 - 2999 euro per maand	460	9,2	9,2	70,3
	3000 - 3499 euro per maand	384	7,6	7,7	78,0
	3500 - 3999 euro per maand	260	5,2	5,2	83,1
	4000 - 4499 euro per maand	138	2,7	2,7	85,9
	4500 - 4999 euro per maand	67	1,3	1,3	87,2
	5000 - 5499 euro per maand	41	,8	,8	88,0
	5500 - 5999 euro per maand	19	,4	,4	88,4
	6000 - 6999 euro per maand	17	,3	,3	88,8
	7000 euro of meer per maand	39	,8	,8	89,5
	weet niet, wil niet zeggen	525	10,5	10,5	100,0
	Total	5019	100,0	100,0	
Missing	System	1	,0		
Total		5020	100,0		

Bewerkingen / Operationalisaties

RECODE w1fa61 (99=SYSMIS) (ELSE=Copy) INTO Inkomen.
EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES= Inkomen
/ORDER=ANALYSIS.

Statistics

Inkomen

N	Valid	4494
	Missing	526

Inkomen

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	407	8,1	9,1	9,1
	2,00	262	5,2	5,8	14,9
	3,00	223	4,4	5,0	19,8
	4,00	429	8,5	9,5	29,4
	5,00	650	12,9	14,5	43,9
	6,00	599	11,9	13,3	57,2
	7,00	499	9,9	11,1	68,3
	8,00	460	9,2	10,2	78,5
	9,00	384	7,6	8,5	87,1
	10,00	260	5,2	5,8	92,9
	11,00	138	2,7	3,1	95,9
	12,00	67	1,3	1,5	97,4
	13,00	41	,8	,9	98,3
	14,00	19	,4	,4	98,8
	15,00	17	,3	,4	99,1
	16,00	39	,8	,9	100,0
Total		4494	89,5	100,0	
Missing	System	526	10,5		
Total		5020	100,0		

Opleiding

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES= w1fa35 w1fa61 w1fa23b02 w1fa23b03 w1fa23b04 w1fa23b05
w1fa23b06 w1fa23b07 w1fa23b08 w1fa23b09 w1fa23b10 w1fa23b11 w1fa23b12 w1fa23b13
w1fa23b14 w1fa23b15
/ORDER=ANALYSIS.

heeft u deze opleiding met een diploma afgerond? lagere school

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	3833	76,4	92,2	92,2
	nee	325	6,5	7,8	100,0
	Total	4158	82,8	100,0	
Missing	System	861	17,2		
Total		5019	100,0		

heeft u deze opleiding met een diploma afgerond? lbo, vmbo-kb/bbl

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	1007	20,1	81,5	81,5
	nee	228	4,5	18,5	100,0
	Total	1235	24,6	100,0	
Missing	System	3784	75,4		
Total		5019	100,0		

heeft u deze opleiding met een diploma afgerond? mavo, vmbo-tl

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	1362	27,1	88,1	88,1
	nee	184	3,7	11,9	100,0
	Total	1546	30,8	100,0	
Missing	System	3473	69,2		
Total		5019	100,0		

heeft u deze opleiding met een diploma afgerond? havo

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	800	15,9	78,2	78,2
	nee	223	4,4	21,8	100,0
	Total	1023	20,4	100,0	
Missing	System	3996	79,6		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
vwo/gymnasium**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	586	11,7	80,2	80,2
	nee	145	2,9	19,8	100,0
	Total	731	14,6	100,0	
Missing	System	4288	85,4		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond? mbo-kort
(kmbo)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	490	9,8	72,6	72,6
	nee	185	3,7	27,4	100,0
	Total	675	13,4	100,0	
Missing	System	4344	86,6		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond? mbo-
tussen/lang (mbo)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	1190	23,7	80,5	80,5
	nee	289	5,8	19,5	100,0
	Total	1479	29,5	100,0	
Missing	System	3540	70,5		
Total		5019	100,0		

heeft u deze opleiding met een diploma afgerond? hbo

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	826	16,5	70,9	70,9
	nee	339	6,8	29,1	100,0
	Total	1165	23,2	100,0	
Missing	System	3854	76,8		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
universiteit (bachelor)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	171	3,4	52,9	52,9
	nee	152	3,0	47,1	100,0
	Total	323	6,4	100,0	
Missing	System	4696	93,6		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
universiteit (master, doctoraal)**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	252	5,0	76,8	76,8
	nee	76	1,5	23,2	100,0
	Total	328	6,5	100,0	
Missing	System	4691	93,5		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
promotietraject**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	18	,4	66,7	66,7
	nee	9	,2	33,3	100,0
	Total	27	,5	100,0	
Missing	System	4992	99,5		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, laag**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	39	,8	76,5	76,5
	nee	12	,2	23,5	100,0
	Total	51	1,0	100,0	
Missing	System	4968	99,0		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, midden**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	69	1,4	76,7	76,7
	nee	21	,4	23,3	100,0
	Total	90	1,8	100,0	
Missing	System	4929	98,2		
Total		5019	100,0		

**heeft u deze opleiding met een diploma afgerond?
buitenlandse opleiding, hoog**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	40	,8	83,3	83,3
	nee	8	,2	16,7	100,0
	Total	48	1,0	100,0	
Missing	System	4971	99,0		
Total		5019	100,0		

DESCRIPTIVES VARIABLES=w1fa35 w1fa61 w1fa23b02 w1fa23b03 w1fa23b04 w1fa23b05
w1fa23b06 w1fa23b07 w1fa23b08 w1fa23b09 w1fa23b10 w1fa23b11 w1fa23b12 w1fa23b13
w1fa23b14 w1fa23b15
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
verricht u op dit moment betaalde arbeid?	2201	1	2	1,45	,497	,208	,052	-1,959	,104
wat is het netto inkomen per maand %1109,/van u en uw partner samen? /van u?/ %11	5018	1	99	15,85	28,636	2,527	,035	4,493	,069
heeft u deze opleiding met een diploma afgerond? lagere school	4158	1	2	1,08	,268	3,144	,038	7,890	,076
heeft u deze opleiding met een diploma afgerond? lbo, vmbo-kb/bbl	1235	1	2	1,18	,388	1,628	,070	,651	,139
heeft u deze opleiding met een diploma afgerond? mavo, vmbo-tl	1546	1	2	1,12	,324	2,355	,062	3,553	,124
heeft u deze opleiding met een diploma afgerond? havo	1023	1	2	1,22	,413	1,368	,076	-,129	,153
heeft u deze opleiding met een diploma afgerond? vwo/gymnasium	731	1	2	1,20	,399	1,516	,090	,299	,181
heeft u deze opleiding met een diploma afgerond? mbo-kort (kmbo), primair leerli	675	1	2	1,27	,446	1,015	,094	-,972	,188
heeft u deze opleiding met een diploma afgerond? mbo-tussen/lang (mbo), secundair	1479	1	2	1,20	,397	1,538	,064	,366	,127
heeft u deze opleiding met een diploma afgerond? hbo	1165	1	2	1,29	,454	,922	,072	-1,153	,143
heeft u deze opleiding met een diploma afgerond? universiteit (bachelor)	323	1	2	1,47	,500	,118	,136	-1,998	,271
heeft u deze opleiding met een diploma afgerond? universiteit (master, doctoraal)	328	1	2	1,23	,423	1,278	,135	-,370	,268
heeft u deze opleiding met een diploma afgerond? promotietraject	27	1	2	1,33	,480	,749	,448	-1,560	,872
heeft u deze opleiding met een diploma afgerond? buitenlandse opleiding, niet go	51	1	2	1,24	,428	1,286	,333	-,361	,656
heeft u deze opleiding met een diploma afgerond? buitenlandse opleiding, niet go	90	1	2	1,23	,425	1,282	,254	-,364	,503
heeft u deze opleiding met een diploma afgerond? buitenlandse opleiding, niet go	48	1	2	1,17	,377	1,847	,343	1,471	,674
Valid N (listwise)	0								

Bewerkingen / Operationalisaties

IF (w1fa23b02 = 1) Opleiding=2.

EXECUTE.

IF (w1fa23b03 = 1) Opleiding=3.

EXECUTE.

IF (w1fa23b04 = 1) Opleiding=4.

EXECUTE.

IF (w1fa23b05 = 1) Opleiding=5.

EXECUTE.

IF (w1fa23b06 = 1) Opleiding=6.

EXECUTE.

IF (w1fa23b13 = 1) Opleiding=13.

EXECUTE.

IF (w1fa23b08 = 1) Opleiding=7.

EXECUTE.

IF (w1fa23b07 = 1) Opleiding=8.

EXECUTE.

IF (w1fa23b14 = 1) Opleiding=14.

EXECUTE.

IF (w1fa23b09 = 1) Opleiding=9.

EXECUTE.

IF (w1fa23b10 = 1) Opleiding=10.

EXECUTE.

IF (w1fa23b11 = 1) Opleiding=11.

EXECUTE.

IF (w1fa23b12 = 1) Opleiding=12.

EXECUTE.

IF (w1fa23b15 = 1) Opleiding=15.

EXECUTE.

RECODE Opleiding (2=0) (13=1) (14=2) (15=3) (3 thru 6=1) (7 thru 8=2) (9 thru 12=3)

(ELSE=Copy)

INTO Cat_Opleiding.

Uiteindelijke variabele

FREQUENCIES VARIABLES= Cat_Opleiding

/ORDER=ANALYSIS.

Statistics

Cat_Opleiding

N	Valid	4691
	Missing	329

Cat_Opleiding

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	637	12,7	13,6	13,6
	1,00	1527	30,4	32,6	46,1
	2,00	1396	27,8	29,8	75,9
	3,00	1131	22,5	24,1	100,0
	Total	4691	93,4	100,0	
Missing	System	329	6,6		
Total		5020	100,0		

DESCRIPTIVES VARIABLES=Cat_Opleiding
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Cat_Opleiding	4691	,00	3,00	1,6440	,99159	-,078	,036	-1,067	,071
Valid N (listwise)	4691								

Partner

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES= w1fb1
/ORDER=ANALYSIS.

Statistics

heeft u op dit moment een partn

N	Valid	5312
	Missing	0

heeft u op dit moment een partner? hieronder verstaan we iemand met wie u ten mi

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	3650	68,7	68,7	68,7
	nee	1662	31,3	31,3	100,0
Total		5312	100,0	100,0	

DESCRIPTIVES VARIABLES=w1fb1
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics									
	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
heeft u op dit moment een partner? hieronder verstaan we iemand met wie u ten mi	5312	1	2	1,31	,464	,807	,034	-1,349	,067
Valid N (listwise)	5312								

Bewerkingen / Operationalisaties

RECODE w1fb1 (2=0) (ELSE=Copy) INTO Partner_Binair.
EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES=Partner_Binair
/ORDER=ANALYSIS.

Statistics

Partner_Binair

N	Valid	Missing
	5019	0

Partner_Binair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	1588	31,6	31,6	31,6
	1,00	3431	68,4	68,4	100,0
Total		5019	100,0	100,0	

DESCRIPTIVES VARIABLES=Partner_Binair
/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Partner_Binair	5019	,00	1,00	,6836	,46512	-,790	,035	-1,377	,069
Valid N (listwise)	5019								

Werkend

Oorspronkelijke Variabelen

FREQUENCIES VARIABLES= w1fa35

/ORDER=ANALYSIS.

Statistics

verricht u op dit moment betaald

N	Valid	Missing
	2248	3064

verricht u op dit moment betaalde arbeid?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	ja	1229	23,1	54,7	54,7
	nee	1019	19,2	45,3	100,0
	Total	2248	42,3	100,0	
Missing	System	3064	57,7		
Total		5312	100,0		

DESCRIPTIVES VARIABLES=w1fa35

/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
verricht u op dit moment betaalde arbeid?	2248	1	2	1,45	,498	,188	,052	-1,966	,103
Valid N (listwise)	2248								

Bewerkingen / Operationalisaties

RECODE w1fa35 (2=0) (ELSE=Copy) INTO Werkend_Binair.

EXECUTE.

Uiteindelijke variabele

FREQUENCIES VARIABLES=Werkend_Binair

/ORDER=ANALYSIS.

Statistics

Werkend_Binair

N	Valid	2201
	Missing	2818

Werkend_Binair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	,00	987	19,7	44,8	44,8
	1,00	1214	24,2	55,2	100,0
	Total	2201	43,9	100,0	
Missing	System	2818	56,1		
Total		5019	100,0		

DESCRIPTIVES VARIABLES=Werkend_Binair

/STATISTICS=MEAN STDDEV MIN MAX KURTOSIS SKEWNESS.

Descriptive Statistics

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness		Kurtosis	
						Statistic	Std. Error	Statistic	Std. Error
Werkend_Binair	2201	,00	1,00	,5516	,49745	-,208	,052	-1,959	,104
Valid N (listwise)	2201								

Bijlage 2

- Bijlage 2 bevat syntax, output en toelichting van de analyses die je hebt uitgevoerd. Structureer per analyse: dus geef per analyse de syntax, output en toelichting. Zorg dat SPSS de syntax als onderdeel van de output weergeeft (zie Deelopdracht 3 voor instructies hierover).

Regressie-analyse 1: Eenzaamheid

```
REGRESSION  
/MISSING LISTWISE  
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL CHANGE  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT Schaal_Eenzaamheid  
/METHOD=ENTER Geslacht_Binair Partner_Binair w1cage Werkend_Binair Inkomen  
Cat_Opleiding  
/METHOD=ENTER Schaal_SocCohesie  
/SCATTERPLOT=(*ZRESID ,*ZPRED)  
/SAVE COOK LEVER ZRESID.
```

Model Summary^c

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,233 ^a	,054	,051	,49525	,054	16,225	6	1692	,000
2	,278 ^b	,077	,073	,48942	,023	41,571	1	1691	,000

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Dependent Variable: Schaal_Eenzaamheid

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t		Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,789	,043		41,864	,000	1,705	1,873		
	Geslacht_Binair	,014	,025	,014	,583	,560	-,034	,063	,974	1,027
	Partner_Binair	-,119	,029	-,117	-4,056	,000	-,176	-,061	,674	1,485
	age at time of interview	,012	,002	,241	6,935	,000	,009	,016	,463	2,162
	Werkend_Binair	-,107	,025	-,104	-4,231	,000	-,157	-,058	,923	1,083
	Inkomen	-,015	,007	-,087	-2,203	,028	-,029	-,002	,361	2,771
	Cat_Opleiding	-,033	,014	-,062	-2,326	,020	-,061	-,005	,791	1,264
2	(Constant)	2,222	,079		27,993	,000	2,066	2,378		
	Geslacht_Binair	,002	,024	,002	,074	,941	-,046	,050	,968	1,033
	Partner_Binair	-,122	,029	-,120	-4,202	,000	-,179	-,065	,673	1,485
	age at time of interview	,012	,002	,240	6,986	,000	,009	,016	,463	2,162
	Werkend_Binair	-,107	,025	-,103	-4,256	,000	-,156	-,057	,923	1,083
	Inkomen	-,012	,007	-,069	-1,779	,075	-,026	,001	,359	2,785
	Cat_Opleiding	-,035	,014	-,065	-2,481	,013	-,063	-,007	,791	1,264
	Schaal_SocCohesie	-,134	,021	-,152	-6,448	,000	-,175	-,093	,984	1,016

a. Dependent Variable: Schaal_Eenzaamheid

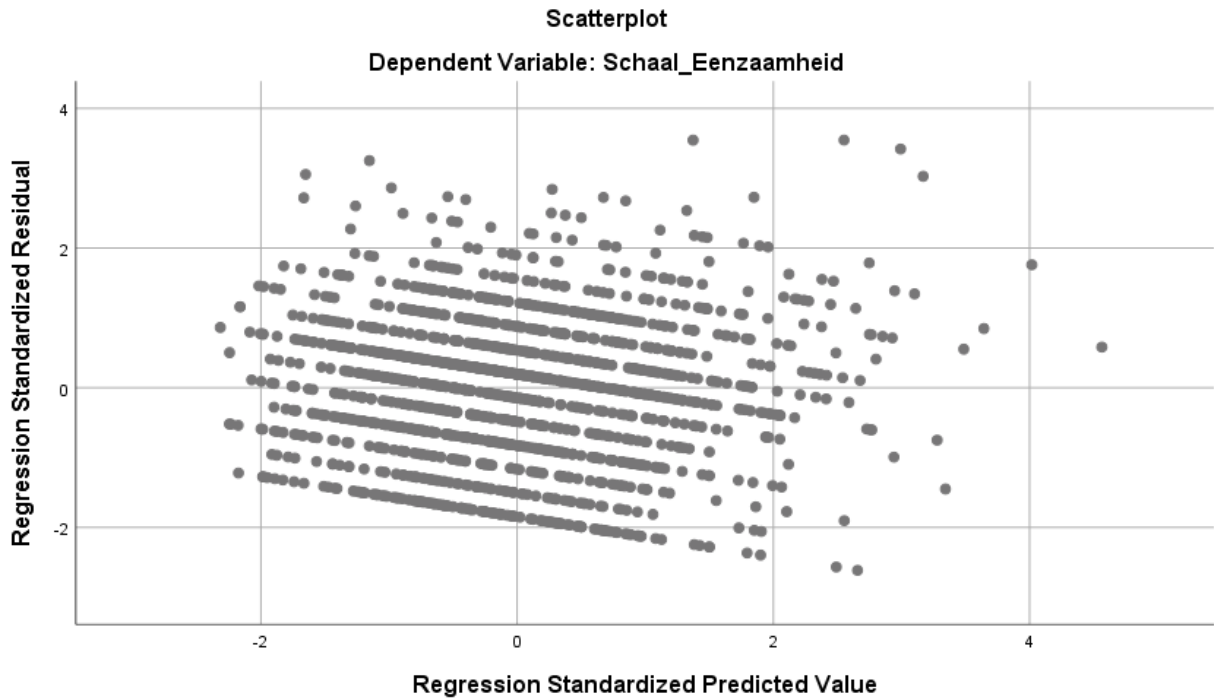
ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	23,877	6	3,980	16,225	,000 ^b
	Residual	415,005	1692	,245		
	Total	438,883	1698			
2	Regression	33,835	7	4,834	20,179	,000 ^c
	Residual	405,048	1691	,240		
	Total	438,883	1698			

a. Dependent Variable: Schaal_Eenzaamheid

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

c. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie



Beschrijving

Hieronder zullen de variabelen besproken worden, die het resultatenhoofdstuk niet heeft benoemd. Binnen het tweede model vertelt de constante ons in een 'beginsituatie' (waarin alle waarden worden ingevuld als nul) een neutrale situatie tegenover eenzaamheid heeft ($b=2,22$; $p<0,0005$). 2,22 bevindt zich namelijk tussen "van toepassing" en tussen "niet van toepassing" in. Wanneer deze respondent ouder wordt, stijgt deze eenzaamheid significant iets ($b=0,012$; $P=0,001$). Mannen scoren gemiddeld hoger op eenzaamheid dan vrouwen ($b=0,002$; $P=0,951$), maar dit effect is niet erg sterk en helemaal niet significant. Ook respondenten met een hoger inkomen scoren gemiddeld iets lager op eenzaamheid ($b=-0,012$; $P=0,075$), maar dit effect is klein en niet significant. Het hebben van een partner verlaagd eenzaamheid ($b=-0,122$; $P<0,0005$), net als het hebben van een hogere opleiding ($b=-0,035$; $p=0,013$). Het hebben van werk heeft ook een verlagend effect op eenzaamheid ($b=-0,107$; $P<0,0005$). Deze drie effecten zijn alledrie significant. De overige beschrijvingen aangaande het 3e model zijn te vinden in het resultatenhoofdstuk.

Regressie-analyse 2: Probleemgedrag

REGRESSION

/MISSING LISTWISE

/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL CHANGE

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

/NOORIGIN

/DEPENDENT Schaal_ProbleemGedrag

/METHOD=ENTER Geslacht_Binair Partner_Binair w1cage Werkend_Binair Inkomen

Cat_Opleiding

/METHOD=ENTER Schaal_SocCohesie

/METHOD=ENTER Schaal_Eenzaamheid

/SCATTERPLOT=(*ZRESID ,*ZPRED)

/SAVE COOK LEVER ZRESID.

Model Summary^d

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,393 ^a	,154	,151	,42127	,154	51,165	6	1685	,000
2	,404 ^b	,163	,160	,41912	,009	18,319	1	1684	,000
3	,404 ^c	,163	,159	,41924	,000	,047	1	1683	,829

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie, Schaal_Eenzaamheid

d. Dependent Variable: Schaal_ProbleemGedrag

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	54,480	6	9,080	51,165	,000 ^b
	Residual	299,028	1685	,177		
	Total	353,507	1691			
2	Regression	57,698	7	8,243	46,923	,000 ^c
	Residual	295,810	1684	,176		
	Total	353,507	1691			
3	Regression	57,706	8	7,213	41,041	,000 ^d
	Residual	295,802	1683	,176		
	Total	353,507	1691			

a. Dependent Variable: Schaal_ProbleemGedrag

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

c. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

d. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie, Schaal_Eenzaamheid

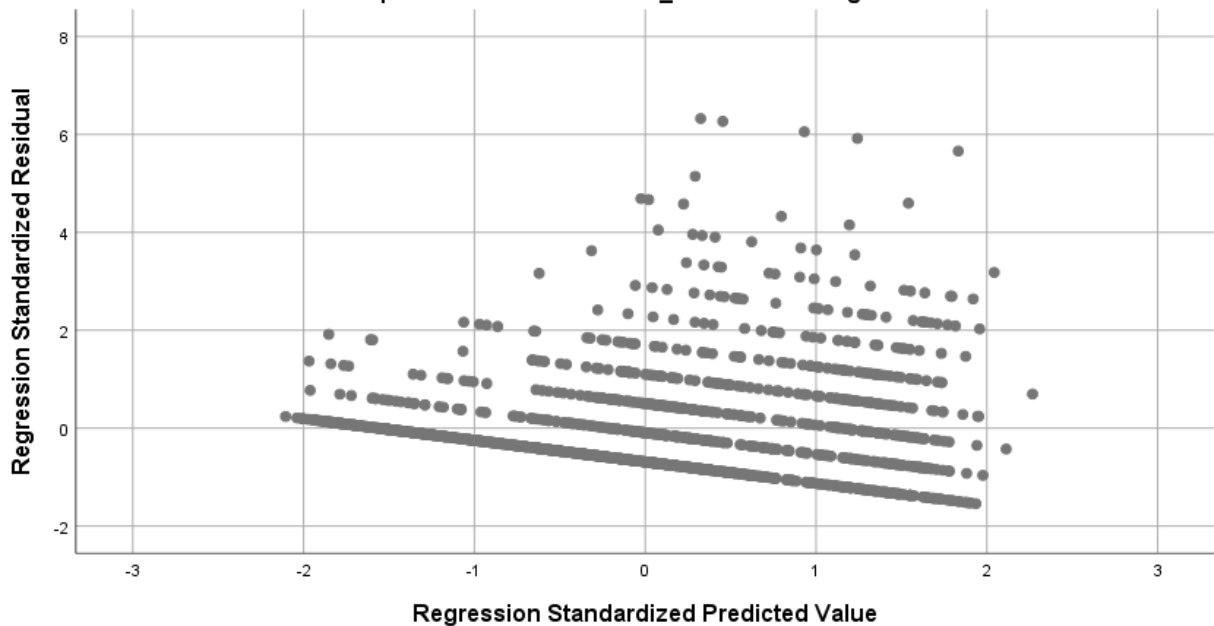
Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t		Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,525	,036		41,928	,000	1,454	1,597		
	Geslacht_Binair	,224	,021	,243	10,703	,000	,183	,266	,975	1,026
	Partner_Binair	-,066	,025	-,073	-2,656	,008	-,115	-,017	,673	1,485
	age at time of interview	-,011	,002	-,228	-6,914	,000	-,014	-,008	,462	2,166
	Werkend_Binair	,020	,022	,021	,910	,363	-,023	,062	,924	1,082
	Inkomen	-,004	,006	-,023	-,617	,537	-,015	,008	,361	2,768
	Cat_Opleiding	,003	,012	,006	,251	,802	-,021	,027	,792	1,263
2	(Constant)	1,772	,068		26,031	,000	1,639	1,906		
	Geslacht_Binair	,217	,021	,235	10,380	,000	,176	,258	,969	1,033
	Partner_Binair	-,068	,025	-,074	-2,730	,006	-,117	-,019	,673	1,486
	age at time of interview	-,011	,002	-,229	-6,966	,000	-,014	-,008	,462	2,166
	Werkend_Binair	,020	,021	,022	,932	,352	-,022	,062	,924	1,082
	Inkomen	-,002	,006	-,012	-,331	,740	-,014	,010	,360	2,781
	Cat_Opleiding	,002	,012	,004	,161	,872	-,022	,026	,792	1,263
3	(Constant)	1,762	,082		21,383	,000	1,600	1,924		
	Geslacht_Binair	,217	,021	,235	10,377	,000	,176	,258	,969	1,033
	Partner_Binair	-,067	,025	-,074	-2,693	,007	-,116	-,018	,666	1,501
	age at time of interview	-,011	,002	-,230	-6,903	,000	-,014	-,008	,449	2,229
	Werkend_Binair	,020	,022	,022	,949	,343	-,022	,063	,914	1,094
	Inkomen	-,002	,006	-,012	-,322	,748	-,014	,010	,359	2,786
	Cat_Opleiding	,002	,012	,004	,174	,862	-,022	,026	,789	1,268
Schaal_SocCohesie	-,076	,018	-,095	-4,193	,000	-,111	-,040	,960	1,041	
Schaal_Eenzaamheid	,005	,021	,005	,216	,829	-,036	,045	,923	1,083	

a. Dependent Variable: Schaal_ProbleemGedrag

Scatterplot

Dependent Variable: Schaal_ProbleemGedrag



Beschrijving

Om de invloed van sociale cohesie op probleemgedrag te begrijpen, is het nodig om het model te begrijpen waarin deze gegevens zich bevinden. Om te beginnen, de constante van model 3 geeft aan wat de gemiddelde score op probleemgedrag is wanneer geen enkele van de waarden binnen de regressie is ingevuld ($b=1,79$; $p<0,0005$). Het is echter realistischer de stijging van het gemiddelde te vergelijken met de “gemiddelde” respondent, welke een waarde van 1,30 toont op probleemgedrag. Dit ligt ergens tussen de “nooit” en “1 keer” probleemgedrag vertoont in de afgelopen maand.

Wanneer men kijkt naar de uitkomsten van de regressie, neemt probleemgedrag af met leeftijd. Het effect van leeftijd lijkt op het eerste gezicht vrij klein ($b=-0,011$; $p<0,0005$), maar wanneer men bedenkt dat de oudste respondent in de dataset 49 jaar is, dan levert dit al snel een daling van -0,539 op. Samengenomen met ons gemiddelde is dit een verschil van 1,30 tot 0,761. Mannen vertonen een iets eerder probleemgedrag ($b=0,22$; $p<0,0005$) dan vrouwen, een gematigd verschil. Het hebben van een hoger inkomen verlaagd dit soort gedrag heel licht ($b=-0,002$; $p=0,748$). Dit effect is niet significant. Het hebben van een partner ontmoedigd dit gedrag daarentegen iets meer, met een significant effect ($b=-0,061$; $p=0,007$). Het hebben van werk schijnt dit soort gedrag juist aan te sporen ($b=0,020$; $p=0,343$). Hetzelfde geldt voor het hebben van een opleiding ($b=0,002$; $p=0,862$). Deze beide effecten zijn niet significant. Het positieve verband zou echter te herleiden kunnen zijn naar het feit dat een van de vragen waarop de schaal opgebouwd is, vraagt naar of de respondent de afgelopen maand school of werk heeft gemist zonder reden. Om dit te kunnen doen, is het natuurlijk wel van belang dat men een opleiding of baan heeft. De overige variabelen worden besproken in het resultatenhoofdstuk.

Regressie-analyse 3: Crimineel Gedrag

LOGISTIC REGRESSION VARIABLES Binair_SchaalCrimineelGedrag
 /METHOD=ENTER Geslacht_Binair Partner_Binair w1cage Werkend_Binair Inkomen
 Cat_Opleiding
 /METHOD=ENTER Schaal_SocCohesie
 /METHOD=ENTER Schaal_Eenzaamheid
 /CRITERIA=PIN(.05) POUT(.10) ITERATE(20) CUT(.5).

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	78,005	6	,000
	Block	78,005	6	,000
	Model	78,005	6	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1005,458 ^a	,045	,095

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Geslacht_Binair	1,016	,175	33,666	1	,000	2,761
	Partner_Binair	-,235	,203	1,344	1	,246	,791
	age at time of interview	-,023	,015	2,429	1	,119	,978
	Werkend_Binair	,208	,179	1,340	1	,247	1,231
	Inkomen	-,103	,059	3,050	1	,081	,902
	Cat_Opleiding	,027	,105	,065	1	,799	1,027
	Constant	-1,805	,310	33,892	1	,000	,164

a. Variable(s) entered on step 1: Geslacht_Binair, Partner_Binair, age at time of interview, Werkend_Binair, Inkomen, Cat_Opleiding.

Block 2: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	12,470	1	,000
	Block	12,470	1	,000
	Model	90,475	7	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	992,988 ^a	,052	,110

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	992,988 ^a	,052	,110

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Block 3: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	,270	1	,603
	Block	,270	1	,603
	Model	90,745	8	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	992,718 ^a	,053	,111

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than ,001.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Geslacht_Binair	,985	,176	31,344	1	,000	2,678
	Partner_Binair	-,229	,203	1,270	1	,260	,795
	age at time of interview	-,026	,015	3,142	1	,076	,974
	Werkend_Binair	,219	,181	1,467	1	,226	1,245
	Inkomen	-,097	,059	2,648	1	,104	,908
	Cat_Opleiding	,018	,105	,031	1	,861	1,019
	Schaal_SocCohesie	-,501	,144	12,179	1	,000	,606
	Schaal_Eenzaamheid	,090	,172	,270	1	,603	1,094
	Constant	-,295	,662	,198	1	,656	,745

a. Variable(s) entered on step 1: Schaal_Eenzaamheid.

Beschrijving

Om de variatie in crimineel gedrag te kunnen laten zien afhankelijk van de persoonlijkheidskenmerken van de respondent, zal er steeds 1 waarde van een variabele veranderen, terwijl de andere waarden gelijk blijven. In onderstaande beschrijving zullen de variabele die nog niet besproken zijn in het resultatenhoofdstuk verder worden beschreven.

Met deze techniek, kan men zien dat wanneer men een vrouw is, de kans op het plegen van crimineel gedrag 0,034 is ($b=0,99; p<0,0005$), in tegenstelling tot de 0,085 die een 'gemiddelde' respondent ons geeft. Het zijn van een vrouw doet de kans op crimineel gedrag dus significant dalen. Wanneer men geen partner heeft, ligt de kans daarentegen op 0,106 ($b=-0,23; p=0,260$). Het hebben van een partner geeft dus een verhoogd dus de kans op crimineel gedrag enigszins. Wanneer men invult dat de respondent 14 is, geeft dit een vrij sterke stijging in kans ($P= 0,135$) om crimineel gedrag te plegen ($b=-0,03; p=0,076$). Op 49-jarige leeftijd ligt deze kans op 0,052. Een niet-werkende respondent heeft een kans van 0,070 om crimineel gedrag te plegen ($b=0,21; p=0,226$). Dit ligt iets hoger dan de "gemiddelde" kan van 0,072. Het hebben van de laagste inkomenscategorie (minder dan 150 euro in de maand netto) geeft een kans van 0,133 op crimineel gedrag, waar het vallen in de hoogste inkomenscategorie (7000 euro of

meer per maand netto) een kans van 0,034 ($b=-0,098;p=0,104$). Het hebben van een laag inkomen versterkt de kans op crimineel gedrag enigszins, waar deze daalt wanneer men een hoog inkomen heeft. Wanneer de respondent in de laagste categorie van opleiding valt (geen middelbare school diploma) is de kans op crimineel gedrag 0,083. Wanneer men in de hoogste opleidingscategorie valt (HBO of hoger), dan is de kans op crimineel gedrag 0,089 ($b=0,015;p=0,880$). Allebei deze kansen liggen ongeveer gelijk met de “gemiddelde” respondent. Zowel opleiding als inkomen zijn niet significant voor het bepalen van crimineel gedrag.

Bijlage 3

Assumptietoetsing lineaire regressie

Om een lineaire regressie uit te mogen voeren zijn er 4 assumpties waaraan moet worden voldaan om te kunnen concluderen dat de regressie ook geldige resultaten geeft die naar een gehele populatie gegeneraliseerd mogen worden. In deze paragraaf wordt duidelijk waarom lineaire regressie voor de variabele crimineel gedrag niet mogelijk zal zijn. Verder zal duidelijk worden waarom er voor de andere twee regressies wel voor voor een lineaire regressie gekozen is.

Onafhankelijke observaties

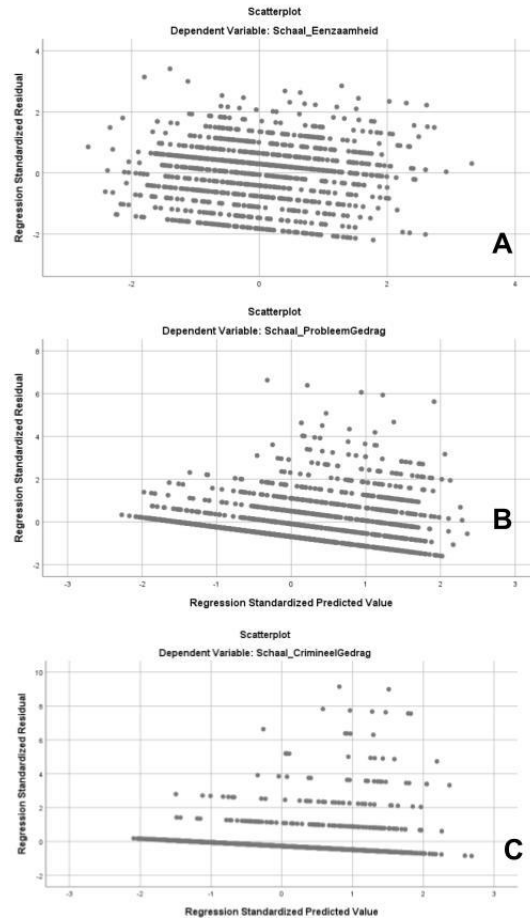
De eerste assumptie waaraan de regressie moet voldoen is dat de verschillende observaties onafhankelijk van elkaar zijn. Hiervoor is het nemen van een willekeurige steekproef van belang. De NELLS-dataset is niet volledig willekeurig getrokken, maar is gebeurd door middel van twee-fase stratified sampling. Eerst zijn er 35 gemeenten geselecteerd, en daarna werd er een willekeurige selectie gemaakt uit het bevolkingsregister op basis van leeftijd, geboorteland en de ouders van de respondent. Binnen deze quasi-willekeurige steekproef is er rekening gehouden met stedelijkheid en regio.

Binnen deze gemeenten zou een overlappende ervaring kunnen gelden die invloed heeft op de antwoorden van de respondent. In het ideale geval zou hier dus gekozen moeten worden voor multi-level analyse. Dit is echter boven het niveau van dit onderzoek.

Omdat het hier gaat om groot aantal gemeenten, waarbij daarna weer gecorrigeerd is voor de stedelijkheid en de regio, concludeer ik dat de steekproef willekeurig genoeg getrokken is voor het uitvoeren van een regressie-analyse. Deze assumptie is dus niet geschonden.

Lineair verband

Door het maken van een residual plot kan de tweede assumptie van regressie-analyse, lineariteit, gecontroleerd worden. Wanneer deze residual plots geen patroon bevatten, maar lijken op een willekeurige puntenwolk, wordt er aan deze assumptie voldaan. In figuur 6 valt te zien dat er aan deze assumptie bij alle regressies niet volledig voldaan wordt, omdat er meerdere streeppatronen te zien zijn binnen de residual plot. Dit is het gevolg van het feit dat de afhankelijke variabelen worden samengesteld uit het gemiddelde van meerdere categorische variabelen. Hierdoor ontstaat een redelijk continue schaal door de vele antwoordmogelijkheden, maar wel eentje waarbij het aantal antwoordmogelijkheden beperkt is en dus in een soort strepen uit elkaar valt. Bij figuur 6A kan men zeggen dat er een willekeurige puntenwolk te vinden is, ondanks dat deze wel redelijk in strepen uiteenvalt. Voor figuur 6B en 6C is dit minder het geval. Het beperkte aantal antwoorden in combinatie met de inhoudelijke betekenis van deze antwoorden, maakt hier dat de verdeling van de antwoorden aan de onderkant vrij zwaar is. Veel respondenten vallen in de laagste categorie, namelijk weinig problematisch of crimineel gedrag. Dit maakt dat 6B nog enige willekeurigheid toont, maar bij 6C wordt de verdeling van de variabelen problematisch. Voor 6C is deze assumptie duidelijk geschonden. Figuur 6A is nog redelijk acceptabel, figuur 6B ligt op de grens van acceptabel en geschonden. Dit is opnieuw een aanwijzing dat er extra voorzichtig gedaan zal moeten worden met de resultaten aangaande probleemgedrag.

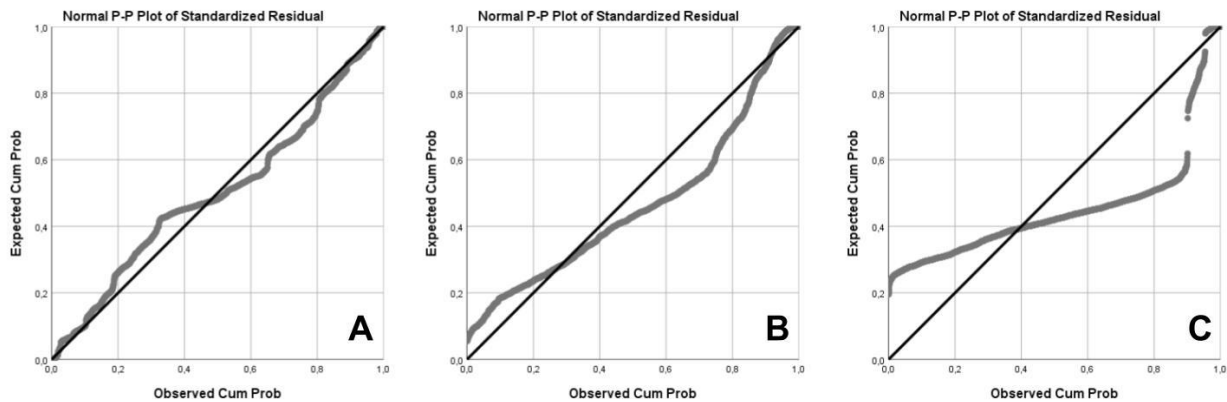


Figuur 6: Residual plots van A. Eenzaamheid, B. Probleemgedrag & C. Crimineel gedrag

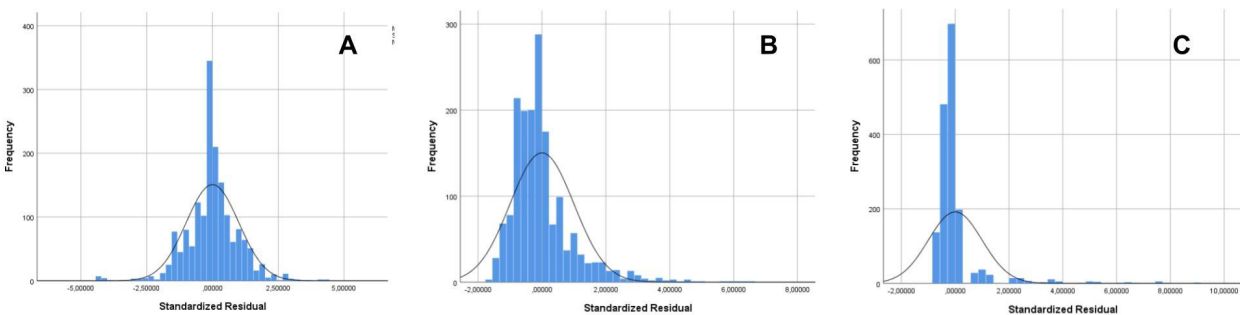
Homoscedasticiteit

Homoscedasticiteit betekent dat de variantie van de residuen voor iedere variabele ongeveer gelijk is. Dit houdt in dat wanneer je het gemiddelde van iedere waarde van X af zou trekken, dat wat overblijft gelijk verspreid is voor iedere waarde van X. In bovenstaande residual plots zou zich dat uiten in een soort “band” die even breed is over de gehele nullijn. In figuur 6A is een goed voorbeeld te zien van een assumptie die niet geschonden wordt. In deze redelijk willekeurige puntenwolk is om en nabij overal even breed. Opnieuw vertonen figuur 6B en 6C een patroon wat niet geheel binnen deze assumptie past. Zoals eerder vermeldt ligt bij beide figuren ligt het gewicht van de regressie vooral onderin, en zijn bovenin maar heel weinig waarden te vinden. Bovendien is bij beide ook te zien dat de residuen niet recht maar naar

rechtsonder gekanteld zijn. Bij figuur 6B is nog sprake van een enigzins overal even brede band, ookal is deze wel redelijk scheef. Opnieuw is hier dus oppassen geboden. Bij figuur 6C is van een overal even brede band is geen sprake. Onderin is een enkele streep te zien met daarboven enige los rondzwevende punten. De assumptie van homoscedasticiteit wordt hierbij dus door crimineel gedrag geschonden.



Figuur 7: PP Plots van de residuen van A. Eenzaamheid, B. Probleemgedrag & Crimineel gedrag



Figuur 8: Histogrammen van de residuen van A. Eenzaamheid, B. Probleemgedrag & Crimineel gedrag

Normaal-verdeelde residuen

De assumptie die hier het laatst gecontroleerd gaat worden is of de residuen van de onafhankelijke variabele normaal zijn verdeeld. Dit wordt getest aan de hand van histogrammen en PP plots. Voor de PP plots geldt, dat wanneer alle waarden langs de zwarte lijn zouden liggen, aan deze assumptie wordt voldaan. Voor de histogrammen geldt dat wanneer de staven dezelfde vorm aannemen als de zwarte lijn er sprake is van een normaalverdeling. Voor deze assumptie geldt dat eigenlijk geen enkele variabele hier perfect aan voldoet. Bij figuur 8A is, op

de pieken in het midden na, te zien dat de verdeling nog redelijk normaal is. Wanneer deze naast figuur 7A legt, is te zien eenzaamheid deze toch iets schever is dan dat graag wordt gezien. Dit zou bij de analyse echter weinig problemen op moeten leveren. Bij figuur 8B is nog een redelijke normaalverdeling te herkennen, ookal wordt dit wel een stuk lastiger. Figuur 8B laat ook zien dat de residuen zeker niet netjes langs de zwarte lijn in het plot liggen.

Probleemgedrag zit dus echt dicht op het schenden van de assumptie van normaal-verdeelde residuen. In figuur 8C is van een normaalverdeling geen sprake meer, wat hier nog gevonden kan worden is een steile piek gevolgd door een paar kleine piekjes die steeds kleiner worden in omvang. In de PP plot is ook te zien dat dat de residuen zwaar afwijken van het te verwachten patroon.

Op basis van deze grafieken en plots is gebaseerd dat de hypothesen aangaande crimineel gedrag niet te testen zijn door middel van een lineaire regressie. Als alternatief wordt er daarom voor een logistische regressie gekozen. Ook bij probleemgedrag zullen enige aanpassingen plaatsvinden. Aangezien er veel schendingen Er zal hier gekozen worden voor een alpha van 0,01 ipv 0,05. Opnieuw zal er hier voorzichtig om worden gegaan met de generaliseerbaarheid van de resultaten.

De hypothesen aangaande eenzaamheid zullen daarentegen wel aan de hand van een lineaire regressie geschat worden. Aangezien er wel enige lichte schendingen waren, zal er daarom gekozen worden om voorzichtig te zijn over het generaliseren van de onderzoeksresultaten.

Missing-data analyse

Binnen bijna iedere dataset zijn er wel missende waarden te vinden. Dit is overal het algemeen geen probleem, zolang het zich hier niet om patronen handelt. Om dit te controleren, zal er een missing-data analyse uitgevoerd worden.

Wanneer men kijkt naar de missing data, kan er worden gezien de meeste missende waarden te vinden zijn bij de vraag of de respondent werk heeft of niet. Er zijn maar liefst 56,2% van de respondenten die deze vraag niet hebben beantwoord. Wanneer er gekeken wordt in de vragenlijst, kan worden gezien dat de vraag niet gesteld is aan alle respondenten die aangeven dat ze voltijd onderwijs genieten. Dit laat de antwoorden van scholieren, studenten en mensen met voltijd bijscholing uit de analyse. Ten tweede zullen hier ook respondenten zijn die deze vraag per ongeluk hebben gemist, omdat de antwoordmogelijkheden en de vraag zelf op twee verschillende pagina's stonden. Door combinatie van deze twee factoren is deze vraag niet ingevuld door maar liefst 2819 respondenten. De variabele die hierna de meeste missende data heeft is inkomen. Door de gevoeligheid van de vraag en mede doordat ik de "weet ik niet" als missing data heb gecodeerd, heeft dit een missing waarde van 10,5%. Hierna komen opleiding (6,6%), crimineel gedrag (4,0%), eenzaamheid (3,7%) en probleemgedrag (3,6%).

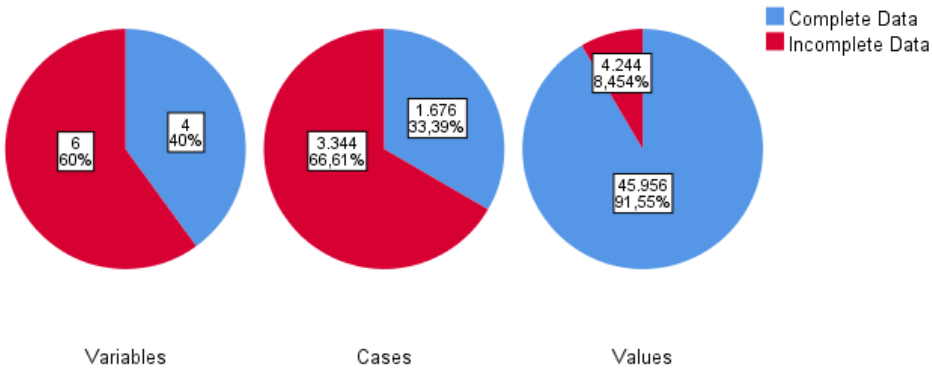
Univariate Statistics

	N	Mean	Std. Deviation	Missing		No. of Extremes ^a	
				Count	Percent	Low	High
Inkomen	4494	6,0946	3,09429	526	10,5	0	75
Schaal_CrimineelGedrag	4820	1,0328	,15112	200	4,0	.	.
Schaal_Eenzaamheid	4832	3,1134	,49909	188	3,7	59	0
Schaal_ProbleemGedrag	4838	1,1983	,37441	182	3,6	0	516
Schaal_SocCohesie	5020	3,3132	,57696	0	,0	69	0
w1 cage	5020	31,24	9,073	0	,0	0	0
Geslacht_Binair	5020			0	,0		
Opleiding	4691			329	6,6		
Partner_Binair	5020			0	,0		
Werkend_Binair	2201			2819	56,2		

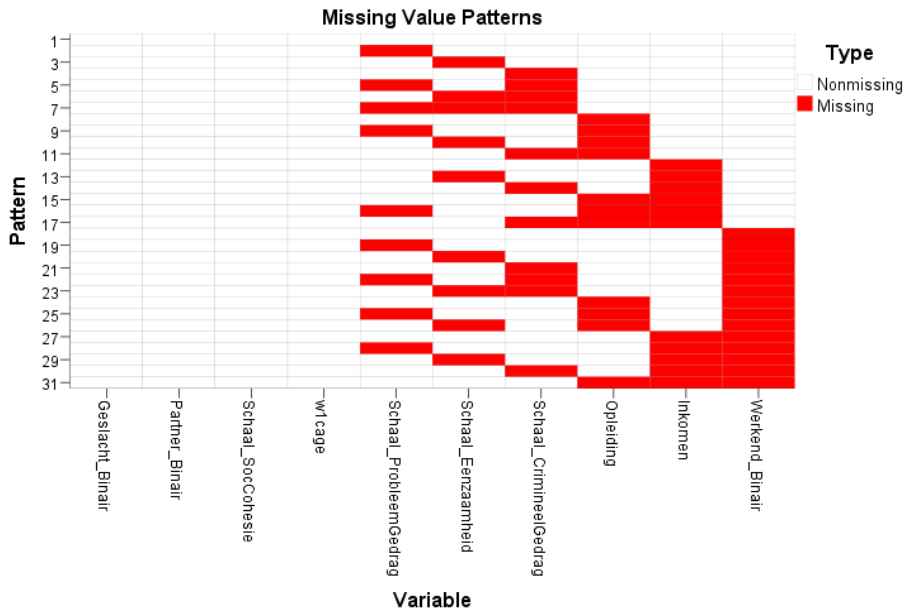
a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

Van alle variabelen zijn maar liefst 6 van de 4 variabelen niet compleet. Van het aantal cases is 66,61% niet compleet. En van alle ingevulde vragen mist maar liefst 8,4% van de waarden. Dit zijn er in totaal 4.244 van de 50200 vragen.

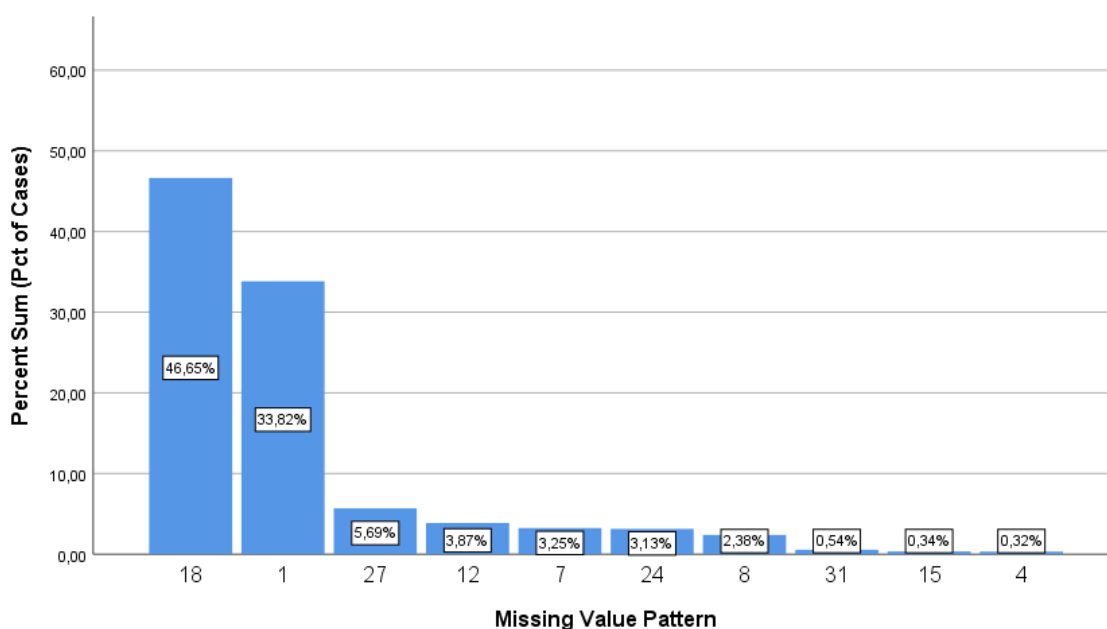
Overall Summary of Missing Values



Onderstaande figuur geeft de verschillende patronen weer die er bestaan waarbij de data niet is ingevuld. De bovenste rij geeft het meest voorkomende patroon aan. De meeste respondenten hebben alle vragen netjes ingevuld. Opvallend is dat de er bij geen enkele respondent missende waarden zijn voor geslacht, het hebben van een partner, de schaal voor sociale cohesie en leeftijd. Ook heeft geen enkele respondent meer dan 3 variabelen niet ingevuld. De meeste patronen waren te ontdekken waarbij werkend niet ingevuld is. Dit is waarschijnlijk wegens bovenstaande redenen het geval. Omdat er niet al te veel overlap binnen de missing values te vinden is, is er geen reden om ons hier zorgen over te maken.



Het meest voorkomende patroon is patroon 19. Dit is het patroon waarbij de schaal van probleemgedrag niet compleet is en waar geen waarde voor ‘werkend’ is ingevuld. Maar liefst 46,63% van de cases volgde dit patroon. Dit is opvallend. De vragenlijst valt hier niet half van een pagina af, wat dus zou betekenen dat de meeste respondenten deze vraag dus gewoon niet graag in willen vullen. Hierna komt de een na laatste groep, die nergens missende waarden heeft (33,82%).



The 10 most frequently occurring patterns are shown in the chart.

Outliers

Outliers (of in het Nederlands, uitbijters) zijn punten die sterk afwijken van de andere punten.

Doordat deze zo ver afliggen van de andere punten, “trekken” deze harder aan de lijn van de regressie: ze geven de gehele regressie een hogere, of lagere, score. Om te kijken tot in hoeverre uitbijters een probleem zijn geweest binnen de analyses van dit onderzoek zal hieronder een uitgebreide beschrijving van de mogelijke outliers volgen. Per analyse zal er eerst worden gekeken naar de gestandaardiseerde residuen, daarna naar de leverage en als laatste

naar de cook's distance. Aan de hand van deze maten en het herschatten van de analyse zal er worden vastgesteld of de uitbijters op systematische fouten zouden kunnen wijzen.

Regressie-analyse 1: Eenzaamheid

De eerste maat die we zullen gebruiken om uitbijters te analyseren is de gestandaardiseerde residuen. Dit is de afstand van dit punt tot het gemiddelde, waarmee rekening gehouden wordt met de spreiding binnen de variabele. Wanneer deze score hoger is dan 3 of lager dan -3, ligt dit punt meer dan 3 standaardfouten van het gemiddelde af. Door de dataset aflopend en oplopend te zetten op basis van de gestandaardiseerde variabele kan worden gezien dat er maar liefst 23 variabelen die deze hierbij scoren als een uitbijter. Deze staan genoemd in tabel 6.

Tabel 6: Gestandaardiseerde residuen eenzaamheid. Geel gemarkeerde respondenten vallen in de top 20 hoogste scores voor Cook's Distance

Respondentnummer	Gestandaardiseerd residue
4030	4,32484
983	4,31580
12028	4,09728
100001154	4,08763
100002143	3,71585
100003159	3,36416
100001343	3,29867
100051322	3,20768
100011448	-4,30240
8534	-4,29283
651	-4,26624
8757	-4,24139
13030	-4,24139
100002557	-4,23779

100002740	-4,22673
2077	-4,17652
100050425	-4,16761
100003603	-4,12449
5520	-4,1086
8275	-3,90593
100002118	-3,44285
12032	-3,01002
100002121	-3,00480

Leverage

Zoals eerder genoemd kunnen uitbijters harder aan de lijn van de regressie “trekken”, en geven zo hogere of lagere resultaten dan de regressie zou geven zonder deze waarde. Een specifieke maat voor hoe hard het punt precies aan de lijn “trekt” is de leverage. Wanneer deze hoger is, hoe hoger deze “hefboomwerking”. Hierbij wordt vaak aangehouden dat wanneer de score hoger is dan $(3p)/n$, de leverage als hoog wordt gezien. Hierbij betekent de p het aantal variabelen en de n het aantal respondenten binnen de regressie. In het geval van deze dataset is de berekening dan:

$$(3 \times 7) / 4.382 = 0,004792$$

Door de data te sorteren op deze variabele, kan men zien dat er maar liefst 507 punten zijn die een leverage hebben dan 0,004792. Dit is niet heel gek wanneer men zich bedenkt dat de eenzaamheid binnen de dataset vrij hoog ligt (3,31 op een schaal van 1 tot 4) in combinatie met de hoge n van de dataset. Door de hoge n wordt de leverage uitermate klein. Door het hoge gemiddelde zijn veel respondenten die lager scoren waarschijnlijk een stuk invloedrijker. Door het grote aantal punten met een hoge leverage score is ervoor gekozen om deze niet allemaal individueel bij naam te noemen. De hoogste leverage-waarde is hier 0,019. De hoogste 10

scores vallen daarentegen niet samen met de hoge residuen of de hoogste 10 scores voor de cook's distance. De respondentnummers zijn: 100050656, 13530, 100010555, 20988, 100006834, 12806, 100003832, 5237, 100002172 & 11064.

Wanneer de regressie opnieuw wordt geschat is te vinden dat het model een hogere R²-adjusted heeft en veel variabelen een sterkere helling krijgen. Deze gegevens staan in figuur 9 & 10. Dit betekent dat het model in staat is om betere voorspellingen te doen, en dat de verschillende variabelen een sterkere invloed hebben. Deze modellen zouden echter alleen mogen worden aangenomen wanneer het zich om een systematische fout handelt. Deze is niet gevonden, waardoor de deze uitbijters niet zijn verwijderd uit de originele regressie.

Model Summary^c

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,155 ^a	,024	,019	,47595	,024	4,381	6	1067	,000
2	,205 ^b	,042	,036	,47180	,018	19,878	1	1066	,000

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Dependent Variable: Schaal_Eenzaamheid

Figuur 9: Tabel Samenvattende statistieken regressie eenzaamheid, zonder punten met hoge leverage

		Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients		Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t		Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,763	,057		30,999	,000	1,651	1,875		
	Geslacht_Binair	,026	,030	,027	,858	,391	-,034	,086	,954	1,049
	Partner_Binair	-,087	,041	-,090	-2,126	,034	-,167	-,007	,507	1,974
	age at time of interview	,010	,003	,194	3,324	,001	,004	,016	,267	3,742
	Werkend_Binair	-,074	,032	-,075	-2,340	,019	-,136	-,012	,884	1,131
	Inkomen	-,014	,012	-,076	-1,140	,254	-,038	,010	,208	4,808
	Cat_Opleiding	-,024	,022	-,040	-1,076	,282	-,068	,020	,653	1,532
2	(Constant)	2,220	,117		18,985	,000	1,990	2,449		
	Geslacht_Binair	,011	,030	,011	,354	,724	-,049	,070	,941	1,062
	Partner_Binair	-,092	,040	-,096	-2,279	,023	-,172	-,013	,506	1,975
	age at time of interview	,011	,003	,203	3,491	,001	,005	,016	,267	3,746
	Werkend_Binair	-,078	,031	-,079	-2,480	,013	-,139	-,016	,884	1,132
	Inkomen	-,011	,012	-,061	-,921	,357	-,035	,013	,207	4,820
	Cat_Opleiding	-,025	,022	-,042	-1,127	,260	-,069	,019	,653	1,533
Schaal_SocCohesie	-,140	,031	-,136	-4,458	,000	-,202	-,078	,965	1,036	

a. Dependent Variable: Schaal_Eenzaamheid

Figuur 10: Tabel Regressie Waarden voor eenzaamheid, zonder punten met hoge leverage

Cook's Distance

Cook's Distance is de laatste maat voor uitbijters waar naar gekeken zal worden. Dit is een product van de studentized residuals en de leverage. Pas wanneer beide componenten hiervan hoog zijn, dan is de Cook's distance ook hoog. Hiervoor geldt dat wanneer de Cook's distance hoger is dan 1, het gaat om een echt probleem. Wanneer de waarde hoger is dan $4/n$ (opnieuw het aantal respondenten) dan gaat het om een mogelijk probleem. Binnen deze dataset is dit een waarde van 0,0009128. Opnieuw is het aantal punten wat boven deze grens ligt erg hoog, maar liefst 298. De hoogste score van op Cook's distance is hier 0,021. Opnieuw zal de oorzaak hier liggen in de hoge n , maar ook in de al reeds hoge leverage van veel van de punten.

Wanneer deze regressie opnieuw worden geschat, ontstaat opnieuw een regressie met een hogere verklaarde variantie en hellingen die enigszins hoger zijn dan de originele regressie. Hier geldt echter nog steeds dat er geen systematische fout binnen deze variabele gevonden kan worden, waardoor het gekozen is om deze punten toch in de originele regressie te laten.

Model Summary^c

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,155 ^a	,024	,019	,47595	,024	4,381	6	1067	,000
2	,205 ^b	,042	,036	,47180	,018	19,878	1	1066	,000

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Dependent Variable: Schaal_Eenzaamheid

Figuur 11: Tabel Samenvattende statistieken regressie eenzaamheid, zonder punten met hoge cook's distance

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,763	,057		30,999	,000	1,651	1,875		
	Geslacht_Binair	,026	,030	,027	,858	,391	-,034	,086	,954	1,049
	Partner_Binair	-,087	,041	-,090	-2,126	,034	-,167	-,007	,507	1,974
	age at time of interview	,010	,003	,194	3,324	,001	,004	,016	,267	3,742
	Werkend_Binair	-,074	,032	-,075	-2,340	,019	-,136	-,012	,884	1,131
	Inkomen	-,014	,012	-,076	-1,140	,254	-,038	,010	,208	4,808
	Cat_Opleiding	-,024	,022	-,040	-1,076	,282	-,068	,020	,653	1,532
2	(Constant)	2,220	,117		18,985	,000	1,990	2,449		
	Geslacht_Binair	,011	,030	,011	,354	,724	-,049	,070	,941	1,062
	Partner_Binair	-,092	,040	-,096	-2,279	,023	-,172	-,013	,506	1,975
	age at time of interview	,011	,003	,203	3,491	,001	,005	,016	,267	3,746
	Werkend_Binair	-,078	,031	-,079	-2,480	,013	-,139	-,016	,884	1,132
	Inkomen	-,011	,012	-,061	-,921	,357	-,035	,013	,207	4,820
	Cat_Opleiding	-,025	,022	-,042	-1,127	,260	-,069	,019	,653	1,533
Schaal_SocCohesie	-,140	,031	-,136	-4,458	,000	-,202	-,078	,965	1,036	

a. Dependent Variable: Schaal_Eenzaamheid

Figuur 12: Tabel Regressie Waarden voor eenzaamheid, zonder punten met hoge Cook's distance

Regressie-analyse 2: Probleemgedrag

Ook voor probleemgedrag zijn er residuen te vinden die een uitbijter zijn op basis van de gestandaardiseerde residuen. Met dezelfde methode zijn er maar liefst 31 punten die hoger scoren dan een 3 en geen enkele respondent die lager scoort dan een -3. Omdat het me niet nuttig lijkt om deze allemaal te benoemen zullen alleen de hoogste 20 waarden worden weergegeven in tabel 7. Dat er respondenten zijn die hoge gestandaardiseerde residuen hebben is niet heel vreemd. Het gemiddelde binnen de dataset lag erg laag, de meeste respondenten hadden een lage score op probleemgedrag. Bovendien is de standaarddeviatie

door de grootte van de steekproef erg klein. Hierdoor liggen hogere scores al snel een groot aantal standaarddeviaties van het gemiddelde af.

Tabel 7: Gestandaardiseerde residuen probleemgedrag. Geel gemarkeerde respondenten vallen in de top 20 hoogste scores voor Cook's Distance.

Respondentnummer	Gestandaardiseerd residue
100011448	6,63184
15059	6,39059
10607	6,06562
2434	5,93673
100005241	5,62758
21104	5,62758
100050183	5,08141
7245	4,67416
9937	4,63222
100004444	4,63216
100010381	4,50637
100004249	4,34292
100001339	4,18902
100000579	4,02685
600060367	4,01597
100002472	3,95414
100001939	3,95414
100000682	3,92161
11412	3,75247
100001238	3,65017

Leverage

Voor de leverage van probleemgedrag is de berekening iets anders, omdat deze 8 variabelen heeft in plaat van 7. Dit komt neer op:

$$(3 \times 8) / 4.382 = 0,005476$$

Binnen Probleemgedrag zijn er maar liefst 80 respondenten die hoger dan deze waarde scoort op leverage. Dit zijn een stuk minder uitschieters dan bij eenzaamheid. De hoogste leverage-waarde is hier 0,019. De tien hoogste waarden hiervan worden gevonden in de respondenten met respondentnummer: 651, 2077, 100003603, 8757, 13030, 4030, 983, 100002557, 5520 & 100002143. De top 20 van de leverage overlappen niet met de gestandaardiseerde residuen en de 20 hoogste cook's distance-waarden.

Zoals kan worden gezien in figuur 13 en 14, doet het model zonder deze hoge leverage scores niet erg veel betere voorspellingen dan het originele model. Hier ligt de uiteindelijke R²-adjusted namelijk op 0,159. Dit verschil is verwaarloosbaar. Mede hierom, en omdat het niet gaat om systematische fouten, is daarom het originele model gebruikt met de outliers er nog in.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,432 ^a	,187	,182	,24473	,187	40,910	6	1070	,000
2	,437 ^b	,191	,186	,24420	,004	5,687	1	1069	,017
3	,437 ^c	,191	,185	,24427	,000	,377	1	1068	,539

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie, Schaal_Eenzaamheid

d. Dependent Variable: Schaal_ProbleemGedrag

Figuur 13: Tabel Samenvattende statistieken regressie probleemgedrag, zonder punten met hoge leverage

		Coefficients ^a								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,325	,029		46,496	,000	1,269	1,381		
	Geslacht_Binair	,129	,016	,228	8,108	,000	,098	,160	,963	1,038
	Partner_Binair	-,032	,021	-,059	-1,527	,127	-,074	,009	,504	1,984
	age at time of interview	-,007	,001	-,228	-4,446	,000	-,009	-,004	,288	3,474
	Werkend_Binair	,054	,016	,098	3,349	,001	,023	,086	,883	1,132
	Inkomen	-,006	,006	-,063	-1,074	,283	-,018	,005	,223	4,487
	Cat_Opleiding	,002	,011	,005	,147	,884	-,020	,023	,664	1,505
2	(Constant)	1,448	,059		24,671	,000	1,332	1,563		
	Geslacht_Binair	,126	,016	,223	7,918	,000	,095	,157	,958	1,044
	Partner_Binair	-,036	,021	-,065	-1,686	,092	-,077	,006	,502	1,993
	age at time of interview	-,006	,001	-,224	-4,368	,000	-,009	-,004	,287	3,479
	Werkend_Binair	,053	,016	,096	3,280	,001	,021	,085	,882	1,133
	Inkomen	-,006	,006	-,055	-,946	,344	-,017	,006	,222	4,500
	Cat_Opleiding	,001	,011	,003	,090	,928	-,020	,022	,664	1,506
3	Schaal_SocCohesie	-,037	,016	-,066	-2,385	,017	-,068	-,007	,979	1,022
	(Constant)	1,470	,069		21,335	,000	1,335	1,605		
	Geslacht_Binair	,126	,016	,222	7,903	,000	,095	,157	,957	1,045
	Partner_Binair	-,037	,021	-,067	-1,728	,084	-,078	,005	,499	2,005
	age at time of interview	-,006	,001	-,221	-4,269	,000	-,009	-,003	,284	3,524
	Werkend_Binair	,052	,016	,095	3,223	,001	,020	,084	,877	1,140
	Inkomen	-,006	,006	-,056	-,965	,335	-,017	,006	,222	4,504
Cat_Opleiding	,001	,011	,002	,069	,945	-,021	,022	,663	1,508	
Schaal_SocCohesie	-,039	,016	-,069	-2,448	,015	-,070	-,008	,957	1,044	
Schaal_Eenzaamheid	-,010	,016	-,017	-,614	,539	-,041	,022	,954	1,048	

a. Dependent Variable: Schaal_ProbleemGedrag

Figuur 14: Tabel Regressie Waarden voor probleemgedrag, zonder punten met hoge leverage

Cook's Distance

De waarde van de Cook's distance verandert niet, aangezien voor iedere regressie de n gelijk blijft. Opnieuw zullen we kijken naar de waarden die hoger zijn dan 0,0009128. Hier zijn maar liefst 205 respondenten met een Cook's distance die hoger ligt dan deze waarde. Geen hiervan ligt hoger dan 1, wat op zware problematiek zou duiden. De hoogste Cook's distance is hier 0,101. Ook wanneer het model opnieuw zou worden opgesteld zonder de punten met een hoge Cook's distance, ligt de R^2 -adjusted niet erg veel hoger dan de R^2 -adjusted van de originele regressie. Bovendien liggen de hellingen niet anders. Ook hier kan dus geen systematische fout gevonden worden in de data, waardoor de originele regressie gebruikt zal worden in het resultatenhoofdstuk. Ook is hiervoor gekozen wegens de inhoudelijke waarde die deze uitbijters hebben voor het model. Wanneer iedere respondent die afwijkt van het gemiddelde niet meer in

de dataset zou zitten, zou als het ware zo goed als niemand meer in de regressie probleemgedrag vertonen. Hiermee zou dus alle zeggingskracht van het model verdwijnen.

Model Summary^d

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	Change Statistics			Sig. F Change
						F Change	df1	df2	
1	,432 ^a	,186	,182	,24521	,186	40,670	6	1065	,000
2	,437 ^b	,191	,185	,24469	,004	5,524	1	1064	,019
3	,437 ^c	,191	,185	,24476	,000	,368	1	1063	,544

a. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen

b. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie

c. Predictors: (Constant), Cat_Opleiding, Geslacht_Binair, Werkend_Binair, Partner_Binair, age at time of interview, Inkomen, Schaal_SocCohesie, Schaal_Eenzaamheid

d. Dependent Variable: Schaal_ProbleemGedrag

Figuur 15: Tabel Samenvattende statistieken voor probleemgedrag, zonder punten met hoge Cook's distance

Coefficients ^a											
Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95,0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta				Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,325	,029			46,303	,000	1,269	1,381		
	Geslacht_Binair	,130	,016	,229		8,132	,000	,099	,161	,963	1,038
	Partner_Binair	-,032	,021	-,059		-1,506	,132	-,073	,010	,506	1,978
	age at time of interview	-,007	,001	-,228		-4,432	,000	-,009	-,004	,289	3,462
	Werkend_Binair	,055	,016	,098		3,352	,001	,023	,087	,885	1,130
	Inkomen	-,006	,006	-,061		-1,045	,296	-,018	,005	,223	4,479
	Cat_Opleiding	,001	,011	,002		,071	,943	-,021	,022	,659	1,517
	(Constant)	1,446	,059			24,565	,000	1,330	1,561		
2	Geslacht_Binair	,127	,016	,224		7,937	,000	,096	,158	,957	1,044
	Partner_Binair	-,035	,021	-,065		-1,666	,096	-,077	,006	,503	1,987
	age at time of interview	-,006	,001	-,223		-4,352	,000	-,009	-,004	,288	3,466
	Werkend_Binair	,053	,016	,096		3,288	,001	,022	,085	,884	1,131
	Inkomen	-,006	,006	-,054		-,927	,354	-,017	,006	,223	4,490
	Cat_Opleiding	,000	,011	,001		,036	,971	-,021	,022	,659	1,517
	Schaal_SocCohesie	-,037	,016	-,066		-2,350	,019	-,068	-,006	,979	1,022
	(Constant)	1,468	,069			21,244	,000	1,332	1,603		
3	Geslacht_Binair	,127	,016	,223		7,922	,000	,095	,158	,957	1,045
	Partner_Binair	-,036	,021	-,067		-1,707	,088	-,078	,005	,500	1,999
	age at time of interview	-,006	,001	-,220		-4,253	,000	-,009	-,003	,285	3,512
	Werkend_Binair	,053	,016	,095		3,233	,001	,021	,085	,879	1,137
	Inkomen	-,006	,006	-,055		-,946	,345	-,017	,006	,222	4,495
	Cat_Opleiding	,000	,011	,001		,016	,987	-,021	,022	,658	1,519
	Schaal_SocCohesie	-,038	,016	-,068		-2,413	,016	-,069	-,007	,957	1,045
	Schaal_Eenzaamheid	-,010	,016	-,017		-,607	,544	-,041	,022	,954	1,048

a. Dependent Variable: Schaal_ProbleemGedrag

Figuur 16: Tabel Regressie Waarden voor probleemgedrag, zonder punten met hoge Cook's distance

Regressie-analyse 3: Crimineel gedrag

Binnen de dataset zijn maar liefst 42 respondent gevonden die een gestandaardiseerd residu hebben boven de 3. Onder de -3 zijn geen gestandaardiseerde residuen te vinden. De 20 hoogste residuen zijn terug te vinden in tabel 8. De uiterst hoge waarden gestandaardiseerde residuen is niet gek wanneer men zich bedenkt dat de gemiddelde score van crimineel gedrag ligt op 1,03. Bovendien is de standaarddeviatie erg klein door de grote steekproef. Er zijn enige respondenten zijn die erg hoog scoren op crimineel gedrag. Deze liggen dan al snel een groot aantal standaarddeviaties van het gemiddelde af en halen hiermee een hoog gestandaardiseerd residue.

Tabel 8: Gestandaardiseerde residuen crimineel gedrag. Geel gemarkeerde respondenten vallen in de top 20 hoogste scores voor Cook's Distance.

Respondentnummer	Gestandaardiseerd residue
100008238	9,13723
14556	8,98549
13062	7,81855
489	7,73541
2434	7,66731
100011131	7,62587
4354	7,55675
6837	7,54872
7407	6,63238
100004042	6,38089
13498	6,37590
100007018	6,36518
7681	6,29463
20277	5,19465
100004444	5,19434
100050515	5,19201
100011338	5,18390
5287	5,00234
13524	4,93422
100001184	4,92142

Leverage

De leverage-waarde waarbij de variabelen als hoog worden gezien, is bij crimineel gedrag hetzelfde als bij probleemgedrag. Dit is het geval omdat de variabelen en het aantal respondenten niet verandert. Dit komt neer op:

$$(3 \times 8) / 4.382 = 0,005476$$

Er zijn maar liefst 437 respondenten die een leverage-waarde hebben die boven deze waarde ligt. De reden dat dit niet heel gek is, is dezelfde reden die zojuist gegeven is voor de hoogte van de gestandaardiseerde residuen. Het gemiddelde van crimineel gedrag ligt ontzettend laag, waardoor veel uitbijters voor de hand ligt. De 10 respondenten met de hoogste leverage zijn: 100050656, 100011448, 100010555, 20988, 100001154, 100003832, 8534, 100002740, 100006834 & 12806. De hoogste leverage-waarde is hier 0,020. Deze komen niet overeen met de 20 hoogste residuen & de 20 respondenten met een hoge Cook's Distance.

Wanneer de punten met een hoge leverage uit de regressie worden gelaten, wordt de -2 log likelihood (de deviance) aanzienlijk kleiner. Deze daalt van 992,72 naar 525,75. Deze lagere deviance kan gedeeltelijk verklaard worden door het feit dat de dataset zonder uitbijters, minder respondenten bevat (hoe minder respondenten, hoe minder fouten er kunnen worden gemaakt, hoe lager de deviance). Dit geeft ons echter geen goede reden om deze uitbijters daadwerkelijk uit de dataset te verwijderen. Omdat de verdeling zo verschrikkelijk rechtsscheef is, hebben we juist deze outliers nodig om iets over crimineel gedrag te kunnen zeggen. Wanneer deze er niet meer in zouden zitten, zou als het ware zo goed als niemand meer in de steekproef crimineel gedrag vertonen en verdwijnt alle zeggingskracht van het model. Dit is ook te zien in de hellingen, waarbij alleen nog maar geslacht (in plaats van bij geslacht en sociale cohesie) significante resultaten vertoont. Vanuit deze inhoudelijke overweging is er gekozen om de waarden met een hoge leverage in het model te laten.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	354,198 ^a	,040	,129

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Figuur 17: Tabel Samenvattende statistieken regressie crimineel gedrag, zonder punten met hoge Leverage

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Geslacht_Binair	1,230	,318	14,917	1	,000	3,420
	Partner_Binair	-,212	,415	,260	1	,610	,809
	age at time of interview	-,075	,040	3,483	1	,062	,928
	Werkend_Binair	,153	,338	,205	1	,650	1,165
	Inkomen	-,043	,137	,099	1	,753	,958
	Cat_Opleiding	,375	,243	2,391	1	,122	1,455
	Schaal_SocCohesie	-,661	,313	4,450	1	,035	,516
	Schaal_Eenzaamheid	-,079	,327	,059	1	,809	,924
	Constant	,134	1,376	,010	1	,922	1,144

a. Variable(s) entered on step 1: Schaal_Eenzaamheid.

Figuur 18: Tabel Regressie Waarden voor crimineel gedrag, zonder punten met hoge leverage

Cook's Distance

De waarde van de Cook's distance is hetzelfde voor de andere regressie-analyses, aangezien bij deze regressie ook de n gelijk blijft. Opnieuw zullen we kijken naar de waarden die hoger zijn dan 0,0009128. Er zijn maar liefst 94 respondenten te vinden met een Cook's Distance die hoger is dan deze waarde. De respondenten die overlappen met de top 20 residuele waarden en de top 20 hoogste waarden Cook's Distance, zijn in tabel 8 gemarkeerd in het geel. De hoogste Cook's distance is hier 0,062. Wanneer dit model opnieuw wordt geschat ligt de deviance opnieuw een stuk hoger dan wanneer deze waarden wel in het model zitten (figuur 19). Opnieuw is er gekozen om deze wel in het model te laten zitten. Dit is opnieuw door de inhoudelijke waarde die deze uitbijters hebben voor het model. Wanneer iedere respondent die afwijkt van het gemiddelde niet meer in de dataset zou zitten, zou als het ware zo goed als niemand meer in de regressie crimineel gedrag vertonen. Hiermee zou dus alle zeggingskracht van het model verdwijnen.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	354,198 ^a	,040	,129

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than ,001.

Figuur 19: Tabel Samenvattende statistieken regressie crimineel gedrag, zonder punten met hoge Cook's Distance

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Geslacht_Binair	1,230	,318	14,917	1	,000	3,420
	Partner_Binair	-,212	,415	,260	1	,610	,809
	age at time of interview	-,075	,040	3,483	1	,062	,928
	Werkend_Binair	,153	,338	,205	1	,650	1,165
	Inkomen	-,043	,137	,099	1	,753	,958
	Cat_Opleiding	,375	,243	2,391	1	,122	1,455
	Schaal_SocCohesie	-,661	,313	4,450	1	,035	,516
	Schaal_Eenzaamheid	-,079	,327	,059	1	,809	,924
	Constant	,134	1,376	,010	1	,922	1,144

a. Variable(s) entered on step 1: Schaal_Eenzaamheid.

Figuur 20: Tabel Regressie Waarden voor crimineel gedrag, zonder punten met hoge Cook's distance

Multi-collineariteit

Binnen dit onderzoek is er weinig sprake geweest van multicollineariteit. Ten eerste kwamen de variabelen inhoudelijk niet al te sterk overeen. De hoogste onderlinge correlatie die je tussen de variabelen kan vinden, staat in tabel 2. Deze was tussen leeftijd en inkomen ($r=0,571$). Dit duidt waarschijnlijk op het fenomeen dat je meer gaat verdienen als je ouder bent. De andere correlatie die boven de 0,500 lag was de relatie met het netto inkomen van de respondent en zijn partner, en het hebben van een partner ($r=0,546$). Ook dit verband is snel te duiden, een werkende respondent met een werkende partner verdient al snel substantieel meer. Binnen de

statistiek wordt vaak aangehouden dat er met een correlatie boven de 0,800 zorgen zouden moeten ontstaan over multi-collineariteit. Dit is binnen dit onderzoek nergens het geval geweest. Een maat die binnen de statistiek gebruikt wordt om aan te geven of er sprake is van multicollineariteit is de VIF-waarden. Dit staat voor Variance inflation factor. Hoe hoger deze score, hoe groter de multicollineariteit. Binnen dit onderzoek wordt aangehouden dat bij een VIF-score van 4 het gaat om problematische waarden. De hoogste waarde binnen de lineaire regressie is 2,78. Dit ligt ver onder de waarde van 4. Vandaar dat er wordt geconcludeerd dat er geen sprake van multicollineariteit binnen dit onderzoek.