



rijksuniversiteit  
groningen

faculteit gedrags- en  
maatschappijwetenschappen

*De relatie tussen seksuele oriëntatie en middelengebruik en de rol van  
uitgaansgedrag in deze relatie onder Nederlandse jongeren*

Auteur: Ryanne Schonewille (S4878663)

[t.j.schonewille@student.rug.nl](mailto:t.j.schonewille@student.rug.nl)

Begeleider: Wouter Kiekens

Tweede lezer: René Veenstra

Datum: 5-6-2024

Cursus: bachelorwerkstuk Sociologie

# Abstract

Uit onderzoek blijkt dat LHB+ jongeren (lesbisch, homoseksueel, biseksueel of andere seksuele oriëntaties) meer middelen gebruiken dan heteroseksuele jongeren, ook in Nederland. Door een stijging in het aantal LHB+ jongeren en de mogelijke risico's van veelvuldig middelengebruik is het van belang dat er inzicht verkregen wordt in de factoren die meespelen bij het middelengebruik van LHB+ jongeren. Dit inzicht kan helpen het verschil in middelengebruik tussen LHB+ en heteroseksuele jongeren te verklaren. De huidige literatuur verklaart het verschil in middelengebruik tussen LHB+ jongeren en heteroseksuele jongeren vooral aan de hand van (sociaal) psychologische factoren, zoals stress. Verschillen in uitgaansgedrag kunnen mogelijk een deel van dit verschil verklaren en dit is nog niet onderzocht voor Nederlandse jongeren. In dit onderzoek is daarom de invloed van seksuele oriëntatie op middelengebruik onderzocht, en tot in hoeverre uitgaansgedrag deze relatie kan verklaren onder Nederlandse jongeren. Ik heb meerdere soorten middelengebruik onderzocht: alcohol drinken, bingedrinken, tabak roken en wiet roken. De onderzoeksvraag heb ik beantwoord aan de hand van data van de eerste wave van The Peers and the Emergence of Adolescent Romance (PEAR) study. Met deze data heb ik vier binaire logistische regressieanalyses en een lineaire regressieanalyse uitgevoerd. Uit de resultaten is gebleken dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren en dat uitgaansgedrag weinig van deze relatie kan verklaren. De bevinding dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren is in lijn met eerdere bevindingen en de theorie. Verder onderstreept deze bevinding het belang van beleid dat het middelengebruik van LHB+ jongeren kan verminderen, zoals antidiscriminatiebeleid. Daarnaast moet vervolgonderzoek opnieuw kijken naar tot in hoeverre uitgaansgedrag de relatie tussen seksuele oriëntatie en middelengebruik kan verklaren, aan de hand van een representatievere steekproef, een betere vragenlijst en een kritische blik op de theorie.

# Inhoud

Abstract .....	2
Inleiding.....	6
Theorie .....	9
Stress en middelengebruik .....	9
Minority stress framework .....	9
Middelengebruik om met stress om te gaan .....	12
Uitgaansgedrag en middelengebruik.....	12
Uitgaansgedrag van LHB+ jongeren .....	12
Normen van middelengebruik .....	12
Toepasbaarheid theorie op (plattelandse) jongeren.....	14
Controlevariabelen .....	15
Methoden.....	17
Data en procedure .....	17
Operationalisaties .....	18
Middelengebruik.....	18
Seksuele oriëntatie.....	19
Uitgaansgedrag .....	20
Controlevariabelen.....	21
Analyse-opzet .....	22
Resultaten.....	25
Beschrijvende statistieken .....	25
Univariate statistieken.....	25
Bivariate statistieken .....	26
Modevaluatie .....	29
Hypothesetoetsing .....	32
Invloed van seksuele oriëntatie op uitgaansgedrag .....	32
Alcohol drinken .....	33
Bingedrinken.....	35
Tabak roken.....	36
Wiet roken .....	37
Conclusie .....	44
Seksuele oriëntatie en middelengebruik.....	44
Uitgaansgedrag .....	44
Discussie.....	46
Vervolgonderzoek.....	48
Tot slot .....	48

Literatuurlijst.....	50
Bijlage 1: variabelen in de analyse en hun totstandkoming.....	56
De afhankelijke variabele: middelengebruik.....	56
Alcohol drinken .....	56
Bingedrinken.....	58
Tabak roken .....	61
Wiet roken .....	63
De eerste onafhankelijke variabele: seksuele oriëntatie .....	65
Seksuele oriëntatie.....	65
De tweede onafhankelijke variabele (mediator): uitgaansgedrag.....	66
Zitten in een caravan doordeweeks .....	66
Zitten in een caravan 's weekends.....	68
Uitgaan in disco's en bars doordeweeks.....	70
Uitgaan in disco's en bars 's weekends.....	72
Betrouwbaarheidsanalyse van de nieuwe items .....	74
Samenvoegen van de items tot de schaalvariabele en de beschrijvende statistieken voor de nieuwe variabele.....	75
Controlevariabelen .....	77
Leeftijd.....	77
Geslacht.....	78
Migratieachtergrond van de respondent .....	79
Migratieachtergrond van de vader van de respondent.....	80
Migratieachtergrond van de moeder van de respondent .....	82
Missende waarden .....	84
Bijlage 2 Statistische analyses.....	86
Univariate analyses.....	86
Output univariate analyses.....	86
Bivariate analyses.....	93
Output bivariate analyses.....	93
Regressieanalyses.....	158
Output regressieanalyses .....	158
Bijlage 3 Assumptiecontrole, uitbijters en multicollineariteit.....	182
Assumptiecontrole .....	182
Output van de assumptiecontrole .....	183
Invloedrijke punten en uitbijters.....	186
Syntax voor het controleren van invloedrijke punten en uitbijters.....	190
Analyses opnieuw uitvoeren zonder uitbijters.....	191
Syntax voor het uitfilteren van de invloedrijke punten en uitbijters.....	192

Multicollineariteit.....	194
Syntax voor het berekenen van de VIF-scores ter controle van multicollineariteit .....	195
Bijlage 4: gebruik van AI-software .....	197

# Inleiding

Uit onderzoek blijkt dat lesbische, homoseksuele, biseksuele en jongeren met een andere seksuele oriëntatie (LHB+) meer middelen gebruiken dan heteroseksuele jongeren (Marshall et al., 2008; Meyer, 2003). In dit onderzoek is een jongere een persoon tussen de 11 en 21 jaar oud. Seksuele oriëntatie heeft betrekking tot iemands gevoelens, dus tot wie iemand zich aangetrokken voelt (Moser, 2015). Veelvuldig middelengebruik, denk aan het gebruik van alcohol, wiet of tabak, kan gevaarlijk zijn door het risico op verslaving en aantasting van organen (Newcomb & Locke, 2006). Voor jongeren brengt veelvuldig middelengebruik extra risico's met zich mee, omdat hun hersenen nog niet vol-groeid zijn en door veel middelen te gebruiken de groei van de hersenen wordt verstoord. Jongeren kunnen hierdoor op latere leeftijd problemen krijgen, zoals gedragsproblemen of concentratiestoornissen (Degenhardt et al., 2016). Verder lijkt het aantal jongeren dat zich identificeert als LHB+ toe te nemen. Zo blijkt uit Amerikaans onderzoek dat de stijging van het aantal LHB+ mensen tussen 2017 en 2021 onder Millennials minder groot was dan in de jongere generatie Generation Z<sup>1</sup>. Zo steeg het aantal LHB+ Millennials van 7,8% naar 10,5%, en voor Gen Z'ers van 10,5% naar 20,8% (Jones, 2022). Gezien de mogelijke risico's van veelvuldig middelengebruik en de toename van het aantal LHB+ jongeren, is het belangrijk om inzicht te verkrijgen in de factoren die invloed hebben op het middelengebruik van LHB+ jongeren. Dit inzicht kan helpen om het verschil in middelengebruik tussen LHB+ en heteroseksuele jongeren te verklaren. Met deze informatie kan de politiek beleid ontwikkelen om de oorzaken van veelvuldig middelengebruik bij LHB+ jongeren te verminderen.

De bestaande literatuur verklaart het verschil in middelengebruik tussen LHB+ en heteroseksuele jongeren vooral aan de hand van (sociaal) psychologische theorieën. Zo blijkt dat LHB+ jongeren meer stress ervaren door bijvoorbeeld discriminatie of afwijzing, en dat ze middelen gebruiken om hiermee om te gaan (Mereish, 2019). Er wordt minder onderzoek gedaan naar de invloed van sociale contexten op het middelengebruik van (LHB+) jongeren, zoals de invloed van de mensen om hen heen. Het onderzoeken van sociale contexten is relevant, omdat bijvoorbeeld blijkt dat LHB+ jongeren meer het gevoel hebben dat middelengebruik gedoogd wordt door de mensen om hen heen dan heteroseksuele jongeren (Litt et al., 2015).

Het uitgaansleven is een voorbeeld van een sociale context die relevant is voor het begrijpen van het middelengebruik van (LHB+) jongeren. Onder het uitgaansleven versta ik uitgaansgelegenheden zoals bars, disco's en clubs. Echter, uitgaansgelegenheden buiten de stad, zoals een keet, worden steeds meer erkend als uitgaansgelegenheden, zeker op en rondom het platteland (Stemmingmakers, 2022; Haartsen & Strijker, 2010). Daarom neem ik keten ook mee als uitgaansgelegenheid in dit onderzoek.

---

<sup>1</sup> Een Millennial is geboren tussen 1981 en 1996. Een Gen Z'er is geboren tussen 1997 en 2012 (Jones, 2022)

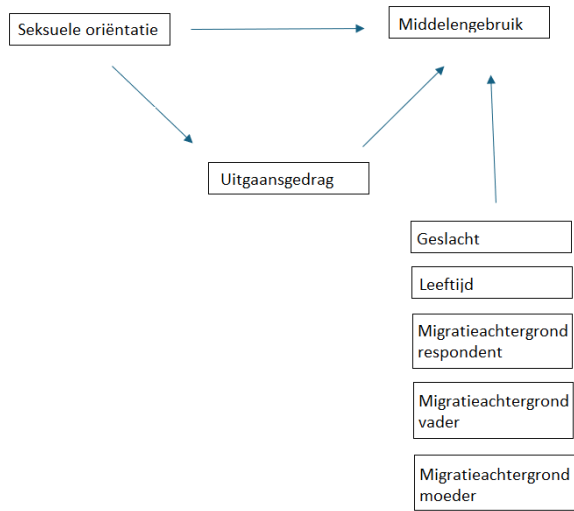
Uit onderzoek in de Verenigde Staten blijkt dat LHB+ mensen meer uitgaan dan heteroseksuelen (Trocki et al., 2005). Daarnaast is bekend dat voor veel LHB+ mensen het LHB+ uitgaansleven een populaire plek is om uit te gaan, naast het algemene uitgaansleven (Green & Feinstein, 2012). Het is nog niet bekend of deze verschillen in uitgaansgedrag ook gelden voor Nederlandse jongeren, maar het is wel interessant om te onderzoeken. Dit omdat mensen die vaak uitgaan meer een norm en acceptatie van middelengebruik ervaren dan mensen die weinig uitgaan, omdat ze meer in contact komen met middelengebruik. Hierdoor gaan ze meer middelen gebruiken (Van de Goor et al., 2011). In LHB+ uitgaansgelegenheden worden doorgaans meer middelen gebruikt dan in het algemene uitgaansleven, waardoor de ervaren norm en acceptatie mogelijk sterker is (Green & Feinstein, 2012). Uitgaansgedrag kan dus mogelijk een deel van de relatie tussen seksuele oriëntatie en middelengebruik verklaren.

Er zijn al eerder onderzoeken gedaan naar het middelengebruik en het uitgaansgedrag van Nederlandse jongeren (Goossens et al., 2013; Kiekens et al., 2020). De onderzoeken naar uitgaansgedrag maken daarentegen geen onderscheid tussen LHB+ en heteroseksuele jongeren. Onderzoeken over middelengebruik theoretiseren dat meer uitgaan kan zorgen voor het ervaren van een norm van middelengebruik, waardoor men meer middelen gaat gebruiken. Echter, dit is nog niet onderzocht voor Nederlandse jongeren. Dit onderzoek is dus een toevoeging aan de huidige literatuur, omdat het nog niet onderzocht is of er verschillen zijn in uitgaansgedrag tussen Nederlandse LHB+ en heteroseksuele jongeren, en of dit verschil de relatie tussen seksuele oriëntatie en middelengebruik gedeeltelijk kan verklaren.

De onderzoeksvraag van dit onderzoek luidt als volgt:

*Wat is de invloed van seksuele oriëntatie op middelengebruik en tot in hoeverre kan uitgaansgedrag deze invloed verklaren onder Nederlandse jongeren?*

Deze vraag beantwoord ik aan de hand van vier binaire logistische regressieanalyses waarin middelengebruik verklaard wordt aan de hand van seksuele oriëntatie en waarin gekeken wordt of uitgaansgedrag (een deel) van de invloed van seksuele oriëntatie op middelengebruik kan verklaren. Verder voer ik een lineaire regressieanalyse uit om het effect van seksuele oriëntatie op uitgaansgedrag vast te stellen. Tot slot controleer ik voor geslacht, leeftijd, de migratieachtergrond van de respondent en de migratieachtergrond van de ouders van de respondent, zie voor het onderzoeksmodel Figuur 1. Verwachtingen voor de weergegeven invloeden in Figuur 1 verklaar ik aan de hand van onder andere het minority stress framework van Meyer (2003) en the social norms approach (Perkins, 2003; Berkowitz, 2004). Voor dit onderzoek gebruik ik data van de eerste wave van The Peers and the Emergence of Adolescent Romance (PEAR) study.



*Figuur 1: Het onderzoeksmodel*



# Theorie

## Stress en middelengebruik

### Minority stress framework

LHB+ jongeren hebben meer stress dan heteroseksuele jongeren en dit kan een verklaring zijn voor het verschil in middelengebruik. LHB+ jongeren hebben meer stress, omdat ze meer stressoren ervaren dan heteroseksuele jongeren. Stressoren zijn situaties die het leven van een persoon veranderen, zoals baanverlies, overbelasting of ziekte (Meyer, 2003). Deze stressoren vergen soms dat een persoon zich aanpast aan de situatie; dit kan stressvol zijn. Er zijn verschillende manieren om stress te definiëren, maar ik ga uit van de definitie die stelt dat stress een staat is van zorgen en spanning als reactie op lastige situaties, oftewel stressoren (World Health Organization, 2023). LHB+ jongeren ervaren meer stressoren dan heteroseksuele jongeren, omdat ze behoren tot een gestigmatiseerde minderheid in de samenleving. Dit houdt in dat er negatieve opvattingen, stereotypen en vooroordelen over hen heersen in de samenleving (Meyer, 2003). Door deze negatieve opvattingen, stereotypen en vooroordelen ervaren LHB+ jongeren stressoren, waardoor ze zogeheten minderheidsstress ervaren (Meyer, 2003). Het minority stress framework van Meyer (2003) beschrijft deze minderheidsstress en de stressoren die LHB+ jongeren ervaren. Ook verklaart het hoe de beschreven stressoren tot stress leiden.

### Assumpties van minderheidsstress

Te beginnen met de beschrijving van minderheidsstress. Meyer (2003) benoemt in zijn onderzoek een drietal assumpties die minderheidsstress karakteriseren. De eerste assumptie is dat minderheidsstress uniek is voor gestigmatiseerde minderheden en dat het bovenop algemene stressoren komt die iedereen kan ervaren. Denk bij algemene stressoren aan prestatiedruk of het verlies van een dierbare. Dit zorgt ervoor dat gestigmatiseerde minderheidsgroepen meer stressoren ervaren en daardoor moeten ze zich meer aanpassen om met deze extra stressoren om te gaan (Meyer, 2003).

De tweede assumptie is dat minderheidsstress chronisch is. Chronisch betekent dat de gestigmatiseerde groep in kwestie er langdurig en voortdurend last van heeft. Dit komt omdat minderheidsstress veroorzaakt wordt door negatieve opvattingen in de samenleving. Deze negatieve opvattingen zijn vastgelegd in sociale en culturele structuren (Meyer, 2003). Sociale en culturele structuren geven weer hoe maatschappijen en culturen zijn vormgegeven, in de vorm van bijvoorbeeld normen, waarden en tradities. De manier waarop mensen denken over bepaalde bevolkingsgroepen, bijvoorbeeld LHB+ jongeren, hoort hier ook bij (Alexander, 2003). Deze structuren zijn stabiel van aard en daardoor veranderen de negatieve opvattingen niet zomaar.

De laatste assumptie is dat minderheidsstress sociaal vastgelegd is. Het komt bijvoorbeeld voort uit sociale processen, zoals interacties met mensen. Daarnaast komt het voort uit instituties, zoals bepaald beleid. Een voorbeeld van beleid die LHB+ mensen discrimineert, is het beleid omtrent bloeddonatie en LHB+ mannen. Homoseksuele mannen mogen bloed doneren als ze minimaal 12 maanden in een monogame relatie zitten (COC, 2021). Deze voorwaarde geldt niet voor heteroseksuele donoren. Dit beleid geldt voor alle LHB+ mannen van 18 jaar en ouder, omdat men pas bloed mag doneren vanaf deze leeftijd. Dit beleid gaat dus niet op voor alle LHB+ jongeren. Toch kan dit beleid door LHB+ jongeren als discriminerend ervaren worden, omdat dit beleid zorgt voor een andere behandeling van LHB+ mensen ten opzichte van heteroseksuelen. Tot slot komt minderheidsstress voort uit sociale structuren, bijvoorbeeld de opvattingen die men heeft over LHB+ jongeren. Minderheidsstress gaat dus verder dan een individueel incident; de maatschappij als geheel heeft invloed op de vorming van minderheidsstress.

### De vier stressoren

Meyer (2003) beschrijft in zijn onderzoek vier stressoren waar LHB+ mensen mee te maken krijgen die tot minderheidsstress leiden: gebeurtenissen van discriminatie, geweld en afwijzing; de verwachting dat deze gebeurtenissen zullen plaatsvinden; het verbergen van je seksuele oriëntatie en het eigen maken van de negatieve opvattingen van buitenaf. Deze stressoren gaan ook op voor LHB+ jongeren. In deze paragraaf leg ik de stressoren van Meyer (2003) uit en zet ik uiteen hoe deze stressoren tot stress leiden. Daarna leg ik uit hoe deze extra stressoren ervoor kunnen zorgen dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren.

De eerste stressor die kan leiden tot minderheidsstress is discriminatie, geweld en afwijzing. Uit onderzoek blijkt dat LHB+ mensen meer gediscrimineerd worden en zij meer geweld tegen hen meemaken dan heteroseksuelen (Mays & Cochran, 2001; De Muijnck et al., 2024). Dit komt omdat er negatieve opvattingen heersen over LHB+ mensen in de samenleving. De verwachting is dat deze negatieve opvattingen en de daaruit voortvloeiende discriminatie ook gelden voor LHB+ jongeren. Discriminatie, angst en afwijzing zorgen voor stress, omdat het onder andere frustratie en angst oplevert. Daarnaast kan meermaals gediscrimineerd worden men hopeloos laten voelen, omdat men uitgesloten wordt en minder kansen heeft in de samenleving (Meyer, 2003). Denk hierbij aan moeite hebben met vrienden maken of het vinden van een baan. Hierdoor ervaren LHB+ jongeren meer stress dan heteroseksuele jongeren.

De tweede stressor die kan leiden tot minderheidsstress is de verwachting van LHB+ jongeren dat ze gediscrimineerd of afgewezen zullen worden op basis van hun seksuele oriëntatie (Meyer, 2003). Door deze verwachting zijn LHB+ jongeren alert in het dagelijks leven, omdat ze zich ervan bewust

zijn dat er een kans bestaat dat de mensen om hen heen een negatieve opvatting hebben over hun seksuele oriëntatie, waardoor zij hen mogelijk zullen discrimineren, naropen of zelfs geweld tegen hen zullen gebruiken. Dit bewustzijn zorgt voor angstgevoelens en spanning, aangezien de kans bestaat dat er nare situaties kunnen plaatsvinden (Steele, 1997). Door alert te zijn hopen LHB+ jongeren nare situaties te kunnen vermijden. Leven in spanning en angst voor discriminatie, geweld of afwijzing zorgt voor stress, omdat er weinig ruimte is om tot rust te komen en even niet na te denken over de negatieve opvattingen die men over je heeft en de mogelijke gevolgen daarvan (Steele, 1997). Deze angstgevoelens zorgen ervoor dat LHB+ jongeren meer stress ervaren dan heteroseksuele jongeren.

De derde stressor die kan leiden tot minderheidsstress is dat LHB+ jongeren hun seksuele oriëntatie verbergen, om afwijzing, geweld of discriminatie te voorkomen of uit schaamte voor hun seksuele oriëntatie. Ze blijven dus in de kast. Ze komen niet uit voor hun seksuele oriëntatie en gedragen zich anders, denk aan hun omgang met anderen, kledingstijl of hun manier van praten (Meyer, 2003). Zij gedragen zich op een manier die tegenstrijdig is met hun identiteit. Dit zorgt voor stress, omdat ze veel bezig zijn met hoe ze zich gedragen en continu alert zijn op hoe zij overkomen op anderen (Palsane, 2005; Steele, 1997). Door het verbergen van hun seksuele oriëntatie ervaren LHB+ jongeren dus meer stress dan heteroseksuele jongeren.

De laatste stressor stelt dat LHB+ jongeren de negatieve opvattingen vanuit de samenleving na verloop van tijd eigen zullen maken, oftewel ze krijgen een slecht zelfbeeld (Meyer, 2003). Dit wordt ook wel geïnternaliseerde homofobie genoemd (Meyer, 2003). Doordat er negatieve opvattingen over LHB+ jongeren heersen in de samenleving, zullen LHB+ jongeren geconfronteerd worden met deze opvattingen in het dagelijks leven, bijvoorbeeld door het zien van nieuwsberichten over geweld tegen LHB+ jongeren (Meyer, 2003). Ook spelen socialisatieprocessen vanaf jonge leeftijd een rol. Socialisatie is het bewust en onbewust internaliseren van de geldende normen, waarden en opvattingen in de samenleving door bijvoorbeeld opvoeding of onderwijs (Fromme, 2006). LHB+ jongeren kunnen via interacties met andere mensen ook geconfronteerd worden met de negatieve opvattingen, doordat ze bijvoorbeeld gediscrimineerd, afgewezen of nageroepen worden. Door blootstelling aan de negatieve opvattingen die over hen heersen en door socialisatie, ervaren LHB+ jongeren meer stress door het internaliseren van deze negatieve opvattingen en het daaruit voortvloeiende negatieve zelfbeeld (Meyer, 2003).

## Middelengebruik om met stress om te gaan

LHB+ jongeren ervaren meer stressoren dan heteroseksuele jongeren en dit zorgt ervoor dat ze meer stress ervaren. Een manier om om te gaan met stress is middelengebruik, zoals alcohol drinken of tabak roken (Mereish, 2019). Middelen zorgen ervoor dat LHB+ jongeren hun problemen even kunnen vergeten en kunnen ontvluchten van de realiteit (Mereish, 2019). Ook zorgen middelen voor minder stress, omdat ze hormonen laten vrijkomen die tijdelijk een fijn en euforisch gevoel geven (Duffing et al., 2014). Aangezien ik verwacht dat LHB+ jongeren meer stress ervaren dan heteroseksuelen, verwacht ik ook dat zij meer middelen gebruiken om met deze stress om te gaan.

De hierboven besproken theorie brengt mij tot de eerste hypothese:

*LHB+ jongeren gebruiken meer middelen dan heteroseksuele jongeren.*

## Uitgaansgedrag en middelengebruik

### Uitgaansgedrag van LHB+ jongeren

Het uitgaansgedrag van LHB+ jongeren verklaart mogelijk voor een deel waarom ze meer middelen gebruiken dan heteroseksuele jongeren. Uitgaansgedrag betekent in dit onderzoek het aantal uren dat een jongere per week besteedt aan uitgaan in disco's, bars, clubs en keten. De verwachting is dat LHB+ jongeren meer uitgaan dan heteroseksuele jongeren. Door afwijzing in de samenleving kunnen LHB+ jongeren het beste steun vinden bij elkaar. LHB+ jongeren ervaren namelijk vergelijkbare situaties, zoals discriminatie, waardoor ze elkaar beter begrijpen (Mereish, 2019). Er zijn uitgaansgelegenheden speciaal voor LHB+ mensen, waar LHB+ jongeren naartoe kunnen om met gelijkgezinden te kunnen praten. Ook zijn deze uitgaansgelegenheden veiliger voor LHB+ jongeren, aangezien de negatieve opvattingen over LHB+ jongeren niet heersen binnen deze gelegenheden en er dus minder risico is op discriminatie, geweld of afwijzing (Maliapaard, 2013). Doordat LHB+ jongeren minderheidsstressoren zoals discriminatie meemaken, kan het fijn zijn dat er veiligere uitgaansgelegenheden zijn voor hen, zodat LHB+ jongeren minder alert hoeven te zijn op mogelijk nare situaties. Ook gaan LHB+ jongeren meer uit dan heteroseksuele jongeren, omdat uitgaan een goede manier is om plezier te hebben en je stress te vergeten (Van de Goor et al., 2011). Aangezien ik verwacht dat LHB+ jongeren meer stress ervaren dan heteroseksuele jongeren, verwacht ik ook dat zij meer uitgaan om deze stress te kunnen vergeten. Concluderend brengt uitgaan voordelen met zich mee voor LHB+ jongeren, waardoor ik verwacht dat ze meer uitgaan dan heteroseksuele jongeren.

### Normen van middelengebruik

Het ontstaan van een norm omtrent middelengebruik in het uitgaansleven kan verklaren waarom meer uitgaan tot een hoger middelengebruik leidt. Uit onderzoek blijkt dat LHB+ mensen het idee hebben dat andere LHB+ mensen meer middelen gebruiken dan heteroseksuelen (Litt et al., 2015).

Daarnaast blijkt dat LHB+ mensen meer het idee hebben dat andere LHB+ mensen middelengebruik accepteren (Litt et al., 2015; Boyle et al., 2020). Echter, de meeste LHB+ mensen geven aan middelengebruik niet te accepteren (Cochran et al., 2012). Er is daarmee een tegenstelling tussen de geschatte acceptatie van middelengebruik en de echte acceptatie van middelengebruik onder LHB+ mensen. Deze overschatting van acceptatie en gebruik van middelen is te verklaren met de social norms approach (Perkins, 2003; Berkowitz, 2004). Deze theorie beschrijft hoe mensen het gedrag en de manier van denken van gelijke anderen verkeerd inschatten, bijvoorbeeld het middelengebruik en de acceptatie ervan. De mate van ongezonde gedragingen van anderen, zoals middelengebruik, worden overschat. Deze misvatting ontstaat doordat men momenten van middelengebruik van gelijke anderen beter herinnert dan momenten waarin er geen middelen worden gebruikt (Perkins, 2003; Berkowitz, 2004). Ondanks dat dit laatste voor de meeste groepen vaker voorkomt, zijn momenten van middelengebruik zichtbaarder en ongebruikelijker. Hierdoor herinnert men dit zich beter en ervaren ze een norm van middelengebruik (Perkins, 2003; Berkowitz, 2004). Ook hierbij geldt dat de verwachting is dat deze bevindingen ook opgaan voor LHB+ jongeren.

Mensen ervaren sterker een norm en acceptatie rondom middelengebruik in het uitgaansleven, omdat er meer middelen worden gebruikt in het uitgaansleven dan in het alledaagse leven en men dus meer momenten van middelengebruik ziet (Van de Goor et al., 2011). Uit onderzoek blijkt dat middelen in het uitgaansleven makkelijk verkrijgbaar zijn en er dus lagere barrières zijn om aan middelen te komen, waardoor men sneller middelen gaat gebruiken (Dimova et al., 2023). Zo kan iedereen boven de 18 gemakkelijk alcohol bestellen aan de bar, maar uit sommige onderzoeken is gebleken dat jongeren onder de 18 soms ook aan alcohol kunnen komen aan de bar, door gebrekkige leeftijdscontroles in uitgaansgelegenheden (Dimova et al., 2023). Daarnaast zijn er dealers op straat van wie drugs gekocht kan worden en zijn bars tot laat open (Dimova et al., 2023). Voornamelijk alcohol is populair in het uitgaansleven en door het gebruik van alcohol zijn mensen sneller geneigd andere middelen te gebruiken, omdat ze door het alcoholgebruik minder goed nadenken over mogelijke risico's van andere middelen (Egginton et al., 2002). Verder blijkt dat in LHB+ uitgaansgelegenheden meer middelen worden gebruikt dan in algemene uitgaansgelegenheden (Green & Feinstein, 2012). Mensen die daar uitgaan komen dus meer in aanraking met middelengebruik dan in het algemene uitgaansleven, waardoor de ervaren norm van middelengebruik sterker is. Aangezien ik verwacht dat LHB+ jongeren meer uitgaan in LHB+ uitgaansgelegenheden en meer uitgaan in het algemeen dan heteroseksuele jongeren, verwacht ik dat ze een sterkere norm van middelengebruik ervaren en daardoor meer middelen gebruiken dan heteroseksuele jongeren.

## Toepasbaarheid theorie op (plattelandse) jongeren

Bovengenoemde theorieën en bevindingen hebben niet allemaal expliciet betrekking tot (plattelandse) jongeren en houden mogelijk geen rekening met de omstandigheden van (plattelandse) jongeren, dus het is onzeker tot in hoeverre deze theorieën ook opgaan voor hen. Het is wel van belang dat deze bevindingen en theorieën opgaan voor (plattelandse) jongeren, aangezien de steekproef van dit onderzoek bestaat uit jongeren die naar middelbare scholen gaan in semi-landelijke gebieden. Dit betekent dat de regio zowel plattelandse als stedelijke kenmerken heeft (Pini, 2019). Hierdoor is het aannemelijk dat er jongeren in de steekproef zitten die op of rondom het platteland wonen. Er zijn twee redenen waarom het aannemelijk is dat bovengenoemde theorieën en bevindingen wel op kunnen gaan voor (plattelandse) jongeren. Ten eerste is de periode van de middelbare school voor veel jongeren een periode waarin het gedrag van andere jongeren veel invloed heeft op hun eigen gedrag (Brown et al., 2008). Het is dus aannemelijk dat als andere jongeren in je omgeving middelen gebruiken, de jongere in kwestie sterk het gevoel kan krijgen dat middelengebruik de norm is en geaccepteerd wordt. Hierdoor gaat de jongere in kwestie het mogelijk zelf ook (meer) doen om erbij te horen.

De tweede reden heeft betrekking tot de mogelijkheid voor jongeren om uit te gaan, in het specifiek jongeren onder de 18. Voor hen kan gelden dat ze niet naar bepaalde uitgaansgelegenheden mogen, omdat sommige uitgaansgelegenheden een leeftijdsrestrictie hebben. Dit is vaak het geval voor clubs of disco's (Peters, 2023). Dan is een uitgaansgelegenheid bij iemand thuis een alternatief, denk hierbij aan een caravan, keet of tuinhuisje. Hier kunnen jongeren bijeen komen om dezelfde activiteiten te ondernemen als in het uitgaansleven in de stad: middelen gebruiken, praten en dansen (Haartsen & Strijker, 2010).

De laatste reden heeft voornamelijk betrekking op plattelandse jongeren. Voor het platteland geldt dat uitgaansgelegenheden zoals disco's en clubs wellicht ver weg zijn en men ver moet reizen om in de stad te komen. Voor deze situatie geldt ook dat de keet een goed alternatief kan zijn als uitgaansgelegenheid (Haartsen & Strijker, 2010). Hierdoor zijn er ook voor jongeren die op of rondom het platteland wonen mogelijkheden om uit te gaan en een norm van middelengebruik te ervaren. Dus ondanks dat niet alle bovengenoemde bevindingen en theorieën expliciet betrekking hebben tot (plattelandse) jongeren, is het wel aannemelijk dat ze ook voor hen kunnen opgaan. Het is daarom interessant om te kijken of meer uitgaan inderdaad tot meer middelengebruik leidt voor jongeren.

De hierboven besproken theorie brengt mij tot de tweede hypothese:

*Het positieve effect van seksuele oriëntatie op middelengebruik wordt gedeeltelijk verklaard door uitgaansgedrag.*

## Controlevariabelen

In dit onderzoek controleer ik voor een aantal variabelen, omdat sommige persoonskenmerken ervoor zorgen dat een persoon meer middelen gebruikt en/of uitgaat dan andere personen, los van iemands seksuele oriëntatie. Aangezien de hypothesen van dit onderzoek zich enkel richten op het effect van seksuele oriëntatie op middelengebruik en uitgaansgedrag, is het goed om voor bepaalde variabelen te controleren (Bhandari, 2023).

Ik controleer ten eerste voor het effect van geslacht, omdat het middelengebruik en uitgaansgedrag van mannen en vrouwen verschillen (Center for Behavioral Health Statistics and Quality, 2017). Over het algemeen nemen jonge mannen meer risico's dan jonge vrouwen, en dit komt omdat mannen eerder de positieve gevolgen zien van risicovol gedrag, zoals alcohol drinken (Halpern-Felsher & Biehl, 2004). Dit geldt voor zowel jongeren als (jong)volwassenen. Dit risicogedrag zorgt ervoor dat mannen meer middelen gebruiken dan vrouwen. Jonge mannen gaan ook meer uit dan jonge vrouwen (Brajša-Žganec et al., 2010).

Ten tweede controleer ik voor leeftijd. Over het algemeen tonen veel jongeren risicovol gedrag, waardoor ze sneller geneigd zijn middelen te gebruiken en uit te gaan voor de spanning en sensatie dan oudere mensen (Van Leijenhorst & Crone, 2009). Echter, middelengebruik en uitgaan neemt toe naarmate jongeren ouder worden (Van Laar et al., 2010; Goossens et al., 2013).

Als laatste controleer ik voor migratieachtergrond en maak ik een onderscheid tussen geen migratieachtergrond, oftewel de persoon is in Nederland geboren, of een niet-westerse migratieachtergrond. Met migratieachtergrond bedoel ik of iemand een eerste generatie migratieachtergrond heeft. Dit verwijst naar het land waar iemand geboren is en vanuit dit land is diegene naar Nederland gemigreerd (Centraal Bureau voor de Statistiek, z.d.-a). De rest van dit onderzoek zal ik enkel migratieachtergrond gebruiken, maar daarmee bedoel ik eerste generatie migratieachtergrond. Iemands eigen migratieachtergrond kan een rol spelen bij hun middelengebruik en uitgaansgedrag, omdat je door interacties met anderen de cultuur van je geboorteland leert (Gabbert, 2006). De migratieachtergrond van iemands ouders speelt ook mee, aangezien zij hun kind opvoeden met normen en waarden die hen aangeleerd zijn in hun geboorteland (Barni et al., 2017). Ik controleer voor migratieachtergrond, omdat in westerse landen, zoals Nederland, meer middelen worden gebruikt dan in niet-westerse landen, omdat westerse mensen toleranter zijn naar middelengebruik (Haidt, 2013). Wanneer mensen uit niet-westerse culturen verhuizen naar het westen, houden ze zich doorgaans vast aan hun niet-westerse cultuur of combineren ze de culturen (Bhugra, 2004). Hierdoor blijven ze minder middelen gebruiken dan westerse mensen. Wat betreft uitgaansgedrag is het gebleken dat niet-westerse

mensen minder uitgaan dan westerse mensen, dit komt bijvoorbeeld doordat niet-westerse mensen doorgaans een lager inkomen hebben (Van den Broek & Keuzenkamp, 2008).



# Methoden

In dit hoofdstuk beschrijf ik hoe dit onderzoek is uitgevoerd. Eerst beschrijf ik de gebruikte dataset, inclusief informatie over steekproeftrekking, werving en non-respons. Ten tweede vertel ik hoe de variabelen in dit onderzoek zijn gemeten. Als er aanpassingen zijn gedaan op de variabelen, bespreek ik deze ook. Tot slot beschrijf ik de analyse-opzet.

## Data en procedure

In dit onderzoek heb ik gebruik gemaakt van data van The Peers and the Emergence of Adolescent Romance (PEAR) study. In deze studie is onderzoek gedaan naar romantische en seksuele relaties van jongeren. Deze informatie is verzameld aan de hand van enquêtes. Deze enquêtes zijn twee keer afgenomen bij vier middelbare scholen in Noord-Nederland. De scholen staan geografisch afgezonderd in een semi-plattelands gebied. Dit betekent dat het gebied zowel plattelands als stedelijke kenmerken heeft (Pini, 2019). Het gemiddelde besteedbare inkomen in deze regio ligt rond het nationale gemiddelde. Verder hebben het overgrote deel van de inwoners van dit gebied geen migratieachtergrond (96,3%). Het eerste afnamemoment was in de herfst van 2014 en de tweede keer in de lente van 2015. Scholen hebben informatie over het onderzoek gestuurd naar de ouders, waarbij ouders de mogelijkheid hadden om deelname van hun kind aan het onderzoek te weigeren. Als ouders deelname wilden weigeren, moesten ze een ingevuld formulier terugsturen. Ook de leerlingen hadden het recht om te weigeren en ze moesten schriftelijk consent geven om mee te doen. Er zijn 2159 leerlingen uitgenodigd om mee te doen, uiteindelijk heeft 94% (2029 leerlingen) minstens één van de enquêtes ingevuld. De data die ik voor dit onderzoek heb gebruikt, is afkomstig van het eerste afnamemoment ( $N_{\text{met missende waarden}}=2159$ ;  $\text{gemiddelde}_{\text{leeftijd}}= 14,87$ ; 50,1% mannen). De enquêtes zijn op papier afgenomen binnen een lesuur en de uiteindelijke data is geanonimiseerd. Dit houdt in dat de gegeven antwoorden niet terug te herleiden zijn naar een specifiek persoon (Csányi et al., 2021).

In de data zelf zijn missende waarden, oftewel oningevulde vragen, en veelal was dit het geval bij de vragen over seks, middelengebruik of crimineel gedrag. Aan het einde van de vragenlijst hadden de respondenten de mogelijkheid een opmerking te plaatsen en veel gaven aan dat ze sommige vragen te persoonlijk vonden en bepaalde informatie niet wilden delen. Dit kan de non-respons op deze vragen verklaren. Verder had een van de deelnemende scholen een respons van lager dan 75%. Dit komt omdat dit een kleine beroepsschool is waarbij leerlingen voor lange tijd stages lopen op een locatie buiten school (N= 242 leerlingen). Deze school is wel meegenomen in de analyse, omdat de leerlingen die de vragenlijst wel hebben ingevuld bruikbare informatie leveren voor dit onderzoek.

Ik heb alle respondenten die een bruikbaar antwoord hebben gegeven op de benodigde variabelen voor dit onderzoek opgenomen in de analyse. Ik heb antwoorden als 'weet ik niet' of 'geen antwoord' als onbruikbare antwoorden beschouwd. Deze antwoorden heb ik daarom uit de dataset gehaald door ze als missende waarde te coderen, zie Bijlage 1. 'Weet ik niet' en 'geen antwoord' waren antwoordmogelijkheden voor de vraag over seksuele oriëntatie. 'Weet ik niet' is door 120 respondenten ingevuld en 105 respondenten hebben 'geen antwoord' ingevuld. Ook oningevulde vragen zijn onbruikbaar; deze worden in de data al weergegeven als missende waarden. De antwoordmogelijkheid 'anders' heb ik ook als missende waarde gecodeerd, zie Bijlage 1. 'Anders' was een antwoordmogelijkheid voor de vraag over het geboorteland van de respondent en het geboorteland van diens ouders, waarbij de respondent de mogelijkheid had zelf een ander land in te vullen. In totaal hebben 251 respondenten 'anders' ingevuld, waarvan 41 voor de migratieachtergrond van de respondent, 110 voor de migratieachtergrond van de vader van de respondent en 100 voor de migratieachtergrond van de moeder van de respondent. Aan de hand van een filter heb ik alle missende waarden uit de dataset gehaald. Na deze procedure bleven 1231 respondenten over ( $N_{\text{zonder missende waarden}} = 1231$ ;  $\text{gemiddelde}_{\text{leeftijd}} = 14,98$ ; 52,0% vrouwen).

## Operationalisaties

### Middelengebruik

De afhankelijke variabele is middelengebruik. Middelengebruik is gemeten aan de hand van vier verschillende variabelen: alcohol drinken, bingedrinken, tabak roken en wiet roken. Deze variabelen heb ik apart geanalyseerd.

Het eerste middel is alcohol, dit is gemeten door te vragen welke uitspraak het beste bij de respondent past. Hierbij kon de respondent kiezen uit (1) Ik heb nog nooit alcohol gedronken, zelfs geen slokje; (2) Ik heb ooit 1 of 2 keer alcohol gedronken; (3) Ik drink 1 of 2 keer in de maand alcohol; (4) Ik drink 1 of 2 keer per week alcohol of (5) Ik drink elke dag alcohol. Een hoger antwoord betekent dat de respondent vaker alcohol drinkt.

Ten tweede is er, naast de frequentie van alcohol drinken, gevraagd naar de hoeveelheid alcohol die gedronken wordt. In het specifiek wordt er gekeken naar bingedrinken en is respondenten gevraagd "Hoe vaak heb je VIJF OF MEER drankjes met alcohol gedronken tijdens één gelegenheid (bijv. tijdens een feestje of op een avond)?". De respondent kon antwoorden met (1) Nooit; (2) 1 keer; (3) 2 keer; (4) 3 of 4 keer; (5) 5 of 6 keer; (6) 7 of 8 keer of (7) 9 keer of meer. Een hoger antwoord betekent dat de respondent vaker aan bingedrinken heeft gedaan. Respondenten konden deze vraag alleen invullen wanneer ze bij de vraag over alcohol aangegeven hadden dat ze ooit alcohol gedronken hebben. Wanneer een respondent nooit alcohol gedronken heeft, kreeg deze respondent een

missende waarde bij bingedrinken. Als een respondent nooit alcohol gedronken heeft, heeft diegene ook nooit aan bingedrinken gedaan. Daarom heb ik de variabele gehercodeerd zodat deze respondenten geen missende waarde meer hebben, maar onder de categorie 'nooit' vallen bij bingedrinken.

Ten derde tabak roken, dit is gemeten door aan de respondent te vragen welke uitspraak het beste bij hen past. Hoe hoger het antwoord, hoe meer de respondent tabak rookt. De respondent kon antwoorden met (1) Ik heb nooit gerookt, zelfs niet één trekje; (2) Ik heb ooit 1 of 2 keer gerookt; (3) Ik rook 1 of 2 keer per maand; (4) Ik rook 1 of 2 keer per week; (5) Ik rook op zijn minst elke dag of (6) Ik heb gerookt maar ben gestopt.

Tot slot de vraag over wiet roken. De respondent is gevraagd welke uitspraak het beste bij hen past. Respondenten konden hier antwoorden met (1) Ik heb nooit gerookt, zelfs niet één trekje; (2) Ik heb ooit 1 of 2 keer gerookt; (3) Ik rook 1 of 2 keer per maand; (4) Ik rook 1 of 2 keer per week; (5) Ik rook op zijn minst elke dag. Hoe hoger de waarde, hoe meer de respondent wiet rookt.

Ik heb elke variabele van middelengebruik gehercodeerd naar een dummyvariabele met als antwoordcategorieën (1) Nee en (2) Ja. De variabele richt zich nu op of een respondent ooit het middel heeft gebruikt of niet. Antwoordcategorie 1 van de oorspronkelijke variabele is 1 gebleven, dus 'nooit' is nu 'nee'. De overige antwoordcategorieën vallen onder 'ja', aangezien deze antwoordmogelijkheden inhouden dat de respondent ooit het middel heeft gebruikt of nog steeds gebruikt en dus automatisch ook ooit heeft gebruikt, zie Bijlage 1. Ik heb deze variabelen zo gecodeerd, omdat ze scheef verdeeld zijn. De meeste respondenten hebben nog nooit het middel gebruikt of maar één of twee keer (61,2% alcohol drinken; 64,3% bingedrinken; 80,65 tabak roken; 93,6% wiet roken).

De vragen en antwoordmogelijkheden die gebruikt worden om middelengebruik te meten zijn niet gevalideerd, wat betekent dat het onzeker is of deze vragen middelengebruik meten. Echter, er wordt specifiek gevraagd naar hoe vaak respondenten het middel gebruikt hebben. Verder staat er bij de vraag van bingedrinken wat bingedrinken is, zodat het voor de respondenten duidelijk is. Dus ondanks dat de schaal niet gevalideerd is, is het aannemelijk dat de vragen over middelengebruik ook middelengebruik meten.

## Seksuele oriëntatie

De eerste onafhankelijke variabele is seksuele oriëntatie. Seksuele oriëntatie is gemeten door de vraag "Wat denk jij dat je bent?". Respondenten konden hierop antwoorden met (1) Heteroseksueel; (2) Homoseksueel; (3) Biseksueel; (4) Weet ik niet en (5) Geen antwoord. Deze variabele heb ik gehercodeerd door homoseksueel en biseksueel samen te voegen tot de antwoordcategorie 'LHB+',

zie Bijlage 1. De reden hiervoor is dat er weinig homoseksuelen en biseksuelen in de steekproef zitten en kleine groepen zorgen voor grotere standaardfouten en dus minder accurate schattingen voor het model (Agresti & Finlay, 2018). Ook vallen homoseksuelen en biseksuelen onder de categorie LHB+ en dit onderzoek gaat over LHB+ jongeren. Verder heb ik de antwoordcategorieën 'weet ik niet' en 'geen antwoord' als missende waarde gecodeerd, zie Bijlage 1. 'Geen antwoord' heb ik gecodeerd als een missende waarde, omdat dit onderzoek gericht is op heteroseksuele en LHB+ jongeren. Voor iemand die 'geen antwoord' heeft ingevuld, is het onbekend of diegene heteroseksueel of LHB+ is, en daardoor kan diegene niet met zekerheid onder één van deze groepen geplaatst worden. 'Weet ik niet' heb ik gecodeerd als missende waarde, omdat het onzeker is waarom een respondent 'weet ik niet' heeft ingevuld. Als een respondent diens seksuele oriëntatie niet weet, vult diegene 'weet ik niet' in. Wellicht vullen sommige respondenten 'weet ik niet' in, omdat ze bijvoorbeeld panseksueel zijn. Dan hebben ze wel een antwoord, maar is het geen antwoordmogelijkheid. Door de onzekerheid over de exacte betekenis die respondenten aan deze optie geven, laat ik het uit de analyse.

Wat betreft de validiteit van deze vraag, is het onzeker wat deze vraag precies meet. Er zijn namelijk verschillende manieren om seksualiteit te definiëren. Ten eerste seksuele identiteit; dit is hoe iemand zichzelf ziet en identificeert in termen van seksualiteit. Ten tweede de seksuele interesses; dit refereert naar wat iemand wil doen. Dit staat los van of iemand dit daadwerkelijk doet. Ten derde seksueel gedrag, dit is wat iemand daadwerkelijk doet, los van of het consistent is met diens seksuele identiteit of interesses. Tot slot seksuele oriëntatie, dit is tot wie of wat iemand zich aangetrokken voelt (Moser, 2015). De vraag over iemands seksuele oriëntatie specificceert niet over welke vorm van seksualiteit het gaat, waardoor het onzeker is of alle respondenten hun seksuele oriëntatie opgegeven hebben of een andere vorm van seksualiteit.

## Uitgaansgedrag

De tweede onafhankelijke variabele en de mediator is uitgaansgedrag. Uitgaansgedrag is een schaalvariabele en bestaat uit vier verschillende items. Twee items over uitgaan in disco's, bars en club en twee items over met vrienden rondhangen in een caravan, keet, hok of schuur. Voor beide sets van twee items geldt dat het ene item het aantal uur meet dat iemand aan deze activiteit besteedt doordeweeks (maandag tot en met donderdag), de ander meet het voor het weekend (vrijdag tot en met zondag). Met vrienden rondhangen in een caravan, keet, hok of schuur heb ik opgenomen in de analyse, omdat rondhangen in de keet erkend wordt als uitgaansgelegenheid in plattelandse gebieden en de steekproef van dit onderzoek bestaat uit jongeren op middelbare scholen in semi-landelijke gebieden (Haartsen & Strijker, 2010). De items zijn gemeten door respondenten te vragen hoeveel tijd ze besteden aan bovengenoemde activiteiten, waarbij een hoger antwoord betekent dat de respondent er meer tijd aan besteedt. Respondenten konden antwoorden

met (0) 0-1 uur; (1) 1-2 uur; (2) 2-3 uur; (3) 3-4 uur; (4) 4-5 uur of (5) 5-6 uur. Aangezien alle andere variabelen een schaal hebben die begint bij 1, heb ik deze items gehercodeerd zodat deze schaal ook bij 1 begint. Voordat ik deze items ging samenvoegen, heb ik de Chronbach's alpha uitgerekend. Dit is een maat die weergeeft tot in hoeverre verschillende items consistent met elkaar zijn en of ze dus hetzelfde concept lijken te meten. De vier items samen levert een Chronbach's alpha op van 0,80. Dit is een hoge waarde en dus zijn deze items goed samen te gebruiken als schaal. Daarna heb ik de schaal gemaakt, door het gemiddelde te nemen van deze items, zie Bijlage 1. Voor de nieuwe schaalvariabele geldt ook dat een hogere score betekent dat de respondent meer uitgaat.

Wat betreft de validiteit van deze schaal, lijkt het te meten wat het moet meten. Aangezien zowel bars en disco's als de keet als uitgaansgelegenheid worden gezien op het platteland. Daardoor is het aannemelijk dat als men hier tijd aan besteedt, dit ook als uitgaan telt volgens hen.

## Controlevariabelen

De eerste controlevariabele is geslacht. De respondenten werden gevraagd wat hun geslacht is. De antwoorden waren (1) Man of (2) Vrouw. Het is daarentegen twijfelachtig of deze vraag altijd iemands geslacht meet. Geslacht gaat over biologische categorisaties op basis van bepaalde kenmerken, zoals chromosomen: mannen zijn geboren met XY-chromosomen, vrouwen met XX-chromosomen. Gender gaat over hoe iemand zich voelt en dit hoeft niet altijd overeen te komen met iemands geslacht (Newman, 2012). Het is onbekend of alle respondenten op de hoogte zijn van deze definities en of dus voor iedereen het geslacht is gemeten in plaats van gender.

De tweede controlevariabele is leeftijd. Leeftijd is gemeten door de respondenten te vragen naar hun geboortedatum. De geboortedatum is in de dataset omgerekend naar de leeftijd op het moment van invullen, in dit geval tijdens het afnemen van de eerste wave van enquêtes. Iemands geboortedatum is een goede manier om iemands leeftijd te berekenen op het moment van de afname van de enquête, aangezien je tot op de dag precies iemands leeftijd kan berekenen. Zolang de respondenten hun geboortedatum naar waarheid hebben ingevuld, meet deze vraag wat het moet meten.

De derde en laatste controlevariabele is migratieachtergrond. De controlevariabele van migratieachtergrond verwijst naar drie aparte variabelen: de migratieachtergrond van de respondent, de migratieachtergrond van de vader van de respondent en die van de moeder van de respondent. De respondent werd gevraagd naar diens eigen geboorteland en die van diens ouders en kon antwoorden met (1) Nederland; (2) Turkije; (3) Marokko; (4) Suriname; (5) Antillen; Aruba, Curacao, Sint Maarten; (6) Indonesië, Molukken of (7) Anders. Westerse landen, oftewel Nederland, heb ik gecodeerd als 1, niet-westerse landen als 2, zie Bijlage 1. Verder heb ik de categorie 'anders' als missende waarde gecodeerd, omdat het in de gegeven tijd niet haalbaar was om alle antwoorden

handmatig te categoriseren. De uiteindelijke antwoordmogelijkheden zijn (1) geen migratieachtergrond en (2) niet-westerse migratieachtergrond. Deze vraag meet iemands migratieachtergrond, omdat er is gevraagd naar het geboorteland van de respondent en diens ouders. Als een persoon van een niet-westers land naar Nederland migreert, heeft diegene een niet-westerse eerste generatie migratieachtergrond. Personen die in Nederland geboren zijn, hebben geen migratieachtergrond. Deze vraag is dus valide.

## Analyse-opzet

Voor de eerste stap van de analyse heb ik gekeken naar de beschrijvende statistieken. Dit zijn de univariate en de bivariate statistieken. Deze statistieken heb ik uitgerekend voor de complete dataset, oftewel de dataset met enkel de respondenten die alle benodigde vragen ingevuld hebben met een bruikbaar antwoord. Eerst heb ik gekeken naar de univariate statistieken. Deze statistieken geven de verdeling van de variabelen weer. Ik heb de gemiddelden, standaarddeviaties, modus, het minimum en maximum, de mediaan en het eerste en derde kwartiel bekeken. De verdeling van de variabelen is van belang, omdat een variabele die erg scheef verdeeld is mogelijk invloed kan hebben op de analyses door afwijkende observaties die de resultaten beïnvloeden. Daarna heb ik gekeken naar de bivariate statistieken. Deze statistieken geven de onderlinge relaties tussen alle variabelen weer en hiermee geven ze een indicatie van de te verwachten verbanden in de analyses. Ik heb de onderlinge relaties ook bekeken om multicollineariteit vast te stellen, daarnaast heb ik de VIF-scores bekeken. Multicollineariteit betekent dat er een (te) grote samenhang is tussen de onafhankelijke variabelen. Hierdoor zijn de schattingen mogelijk minder zuiver, omdat multicollineariteit zorgt voor grotere standaardfouten (Agresti & Finlay, 2018). Daarnaast zorgt multicollineariteit ervoor dat de onafhankelijke variabelen veel overlappen en dezelfde variantie verklaren in de afhankelijke variabele. Hierdoor voegen ze weinig toe aan het model in termen van verklaarde variantie. De samenhang tussen continue variabelen heb ik vastgesteld aan de hand van Pearson's correlaties. Voor continue en categorische variabelen heb ik eenweg-ANOVA analyses uitgevoerd en aan de hand van de  $R^2$  heb ik de correlatie berekend, door de wortel te nemen van de  $R^2$ . Ook heb ik t-toetsen uitgevoerd, om de verschillen tussen de scores van de groepen te toetsen. Tot slot heb ik kruistabellen gemaakt om de richting van de samenhang tussen de categorische variabelen te berekenen. Cramer's V heb ik hierbij als samenhangsmaat gebruikt.

Om de onderzoeksvraag te beantwoorden, heb ik de twee opgestelde hypothesen getoetst. Dit heb ik gedaan aan de hand van binaire logistische regressieanalyses en een lineaire regressieanalyse. Voor de logistische regressieanalyses heb ik vier verschillende analyses uitgevoerd, waarbij iedere analyse een ander middel had als afhankelijke variabele. Iedere logistische regressieanalyse heb ik op een

hiërarchische wijze opgebouwd, wat betekent dat er elke keer een nieuwe (set van) variabele(n) wordt toegevoegd aan het model.

De eerste analyse toetste het verschil in alcohol gebruik tussen LHB+ jongeren en heteroseksuele jongeren, de tweede bingedrinken, de derde tabak roken en de vierde wiet roken. Ik heb de modellen opgebouwd door eerst de afhankelijke variabele te schatten aan de hand van de controle variabelen: geslacht, leeftijd en de migratieachtergrond van de respondent en diens ouders. Vervolgens heb ik dit model uitgebreid door seksuele oriëntatie toe te voegen. Dit tweede model diende om te toetsen of LHB+ jongeren meer van het middel gebruiken dan heteroseksuele jongeren, oftewel de eerste hypothese. Het derde model bestond uit uitgaansgedrag als afhankelijke variabele, en seksuele oriëntatie, geslacht, leeftijd en de migratieachtergrond van de respondent en diens ouders als onafhankelijke variabelen. Dit model heb ik geschat aan de hand van lineaire regressie. Dit model diende om te kijken wat de invloed van seksuele oriëntatie op uitgaansgedrag is. Deze invloed is van belang om te voldoen aan de voorwaarde van mediatie, want als er geen significant effect is van seksuele oriëntatie op uitgaansgedrag, kan uitgaansgedrag waarschijnlijk weinig verklaren in de relatie tussen seksuele oriëntatie en middelengebruik. Tot slot heb ik de mediator uitgaansgedrag toegevoegd aan het tweede model om de tweede hypothese te toetsen, die stelt dat het positieve effect van seksuele oriëntatie op middelengebruik (gedeeltelijk) verklaard wordt door uitgaansgedrag. De mate van mediatie heb ik vastgesteld door het effect van seksuele oriëntatie op middelengebruik in het tweede en vierde model met elkaar te vergelijken. Een grote verandering in de helling van seksuele oriëntatie in model 4 ten opzichte van de helling in model 2, duidt op een mediatie.

Naast het schatten van de modellen, heb ik ook de modelfit bekeken en getoetst. De modelfit geeft weer hoe goed het model in staat is de afhankelijke variabele te voorspellen en of het model dus goed bij de data past. Ook heb ik de modellen met het vorige model of het lege model vergeleken, om te kijken of de modellen meer variantie konden verklaren. Voor de logistische regressies heb ik eerst naar de Hosmer-Lemeshowtoets gekeken voor een indicatie van de modelfit. Deze toets toetst het verschil tussen de geschatte en de geobserveerde waarden. Een insignificant resultaat betekent dat er weinig verschil zit tussen de geobserveerde en geschatte waarden, en de fit van het model dus niet slecht is. Verder heb ik gekeken naar de deviance, die een indicatie van de modelverbetering geeft op basis van de kans op het vinden van de geobserveerde data, gegeven het model. Hoe lager de deviance, hoe beter het model. Met de likelihood-ratio heb ik getoetst of het model beter is dan het lege model of het model ervoor in termen van deviance. Met deze toets heb ik ook getoetst of de toegevoegde variabele(n) een significante verbetering in fit voor het model opleverden en of ze goed in staat waren de afhankelijke variabele te voorspellen. Voor het lineaire regressiemodel heb ik de  $R^2$

–(*adjusted*) bekeken en voerde ik een F-toets uit om te toetsen of het model significant beter is dan het lege model en om een indicatie te verkrijgen van de modelfit. De  $R^2$  geeft het percentage verklaarde variantie weer; een hoge  $R^2$  betekent dat de variabelen in het model goed in staat zijn de afhankelijke variabele te voorspellen. De F-toets toetst het verschil in  $R^2$  tussen twee modellen.

De uitgevoerde regressieanalyses moeten voldoen aan een aantal assumpties; deze assumpties heb ik gecontroleerd. Voor zowel de logistische regressies als de lineaire regressie geldt dat er voldaan moet worden aan de assumptie van onafhankelijkheid. Daarnaast moet lineaire regressie voldoen aan lineariteit, homoscedasticiteit en een normale verdeling van de residuen. Lineariteit en homoscedasticiteit heb ik gecontroleerd door naar de residual plot en partial plots te kijken. De normale verdeling van de residuen heb ik gecontroleerd door naar het histogram en PP-plot van de residuen te kijken.

Tot slot heb ik gekeken of er uitbijters en invloedrijke punten in de dataset zijn. Dit zijn cases die veel invloed kunnen hebben op de onderzoeksresultaten, doordat ze ver afwijken van de overige observaties. Ik heb de leverage, Cooks's distance, gestandaardiseerde residuen en DFFIT gebruikt om uitbijters vast te stellen voor de lineaire regressieanalyse. Voor de logistische regressieanalyses heb ik gekeken naar de leverage en DFBETA. De meest extreme observaties heb ik uit de data gefilterd en toen heb ik alle analyses opnieuw gedaan om te kijken of er grote verschillen waren in de modelfit en effecten.



# Resultaten

In dit hoofdstuk bespreek ik de resultaten van dit onderzoek. Ten eerste een bespreking van de beschrijvende statistieken, te beginnen met de univariate statistieken en daarna de bivariate statistieken. Ten tweede zal ik de resultaten van de regressieanalyses bespreken. Eerst de modelfit en de aannames, multicollineariteit en uitbijters. Dan zal ik de twee opgestelde hypothesen toetsen aan de hand van de gevonden resultaten.

## Beschrijvende statistieken

### Univariate statistieken

In deze paragraaf bespreek ik de univariate statistieken, zie Tabel 1 voor de resultaten. De opvallendste en interessantste bevindingen uit Tabel 1 bespreek ik hieronder.

Ten eerste valt het op dat het merendeel van de respondenten heteroseksueel is (96,5% heteroseksuelen; 3,5% LHB+). Ten tweede blijkt het dat de meeste respondenten weinig uitgaan (*gemiddelde*= 1,81; *sd*= 1,10). De spreiding is relatief hoog, omdat de score op uitgaansgedrag met 1,10 kan variëren op een schaal van 1 tot en met 7. Het blijkt dus dat er veel verschillen zitten tussen de scores.

Ten tweede zijn de statistieken van middelengebruik interessant. Het valt op dat het merendeel van de respondenten wel ooit alcohol gedronken heeft (76,5%). Voor de andere vormen van middelengebruik geldt het omgekeerde; het merendeel van de respondenten heeft het nooit gebruikt (54,2% bingedrinken; 60,4% tabak roken; 82,9% wiet roken).

Tot slot zijn een aantal statistieken van de controlevariabelen interessant. Deze zijn interessant, omdat ze een indicatie geven van de representativiteit van deze steekproef. Dit wordt uitgebreider besproken in de Discussieparagraaf. De eerste interessante bevinding is dat de meeste respondenten geen migratieachtergrond hebben en dus in Nederland geboren zijn (98,8%). Ook de ouders van de respondenten zijn meestal in Nederland geboren (vaders migratieachtergrond: 93,8%; moeders migratieachtergrond: 95,0%). Tot slot ligt het gemiddelde van leeftijd ongeveer in het midden van de verdeling (*gemiddelde*=14,98; *sd*= 1,56; *minimum*=11,75; *maximum*= 20,75).

Tabel 1: Beschrijving van de in de analyse opgenomen variabelen: gemiddelde (standaarddeviatie) en de vijf-getallensamenvatting (N=1231)

Variabele	Gemiddelde (standaarddeviatie)	Minimum	Maximum	Eerste kwartiel	Mediaan	Derde kwartiel	N
Geslacht*	Man= 48,0% Vrouw= 52,0%	1	2				1231
Leeftijd	14,98 (1,56)	11,75	20,75	13,75	15,00	16,17	1231
Migratieachtergrond respondent*	Geen migratieachtergrond =98,8%  Niet-westerse migratieachtergrond= 1,2%	1	2				1231
Migratieachtergrond vader*	Geen migratieachtergrond = 93,8%  Niet-westerse migratieachtergrond = 6,2%	1	2				1231
Migratieachtergrond moeder*	Geen migratieachtergrond= 95,0%  Niet-westerse migratieachtergrond = 5,0%	1	2				1231
Seksuele oriëntatie*	Heteroseksueel= 96,5%  LHB+= 3,5%	1	2				1231
Uitgaansgedrag	1,81 (1,10)	1	7	1,00	1,50	2,25	1231
Alcohol drinken*	Nee= 23,5% Ja= 76,5%	1	2				1231
Bingedrinken*	Nee= 54,2% Ja= 45,8%	1	2				1231
Tabak roken*	Nee= 60,4% Ja= 39,6%	1	2				1231
Wiet roken*	Nee= 82,9% Ja= 17,1%	1	2				1231

\* bij nominale variabelen is de frequentieverdeling vermeld in percentages

## Bivariate statistieken

In deze paragraaf bespreek ik de bivariate statistieken, zie Tabel 2 voor de berekende samenhangsmaten. De interessantste en opvallendste bevindingen uit Tabel 2 bespreek ik hieronder.

Ten eerste is er in dit onderzoek interesse voor het verband tussen seksuele oriëntatie en middelengebruik. In Tabel 2 valt het op dat er zwakke verbanden zijn tussen seksuele oriëntatie en middelengebruik ( $r_{alcohol\ drinken} = 0,09$ ;  $p = <0,01$ ;  $r_{bingedrinken} = 0,10$ ;  $p = <0,01$ ;  $r_{tabak\ roken} = 0,11$ ;  $p = 0,01$ ;  $r_{wiet\ roken} = 0,08$ ;  $p = <0,01$ ). Deze verbanden zijn significant, ondanks dat ze niet erg groot zijn. Om de richting van deze verbanden te bepalen kijk ik naar de kruistabellen. Uit de kruistabellen blijkt dat LHB+ jongeren meer middelen gebruiken dan heteroseksuelen, dus de samenhang is positief, zie Bijlage 2.

Ten tweede is er interesse voor het verband tussen seksuele oriëntatie en uitgaansgedrag. Het blijkt dat seksuele oriëntatie een zwak en insignificant verband heeft met uitgaansgedrag ( $r=0,04$ ;  $p=0,17$ ). Uit een eenweg-ANOVA blijkt dat het een positief verband is, zie Bijlage 2.

Ten derde is de samenhang tussen uitgaansgedrag en middelengebruik interessant voor dit onderzoek. Uit Tabel 2 blijkt dat uitgaansgedrag een redelijk sterke samenhang heeft met middelengebruik ( $r_{alcohol\ drinken}= 0,25$ ;  $p= <0,01$ ;  $r_{bingedrinken}= 0,45$ ;  $p= <0,01$ ;  $r_{tabak\ roken}= 0,36$ ;  $p= <0,01$ ;  $r_{wiet\ roken}= 0,37$ ;  $p= <0,01$ ). Alle verbanden zijn significant. Bij het uitvoeren van een eenweg-ANOVA blijkt dat mensen die meer uitgaan, meer middelen gebruiken en andersom, zie Bijlage 2. Ondanks dat het een redelijk sterke verbanden zijn, zijn ze niet sterk genoeg om multicollineariteit te veroorzaken. Over het algemeen worden correlaties boven de 0,80 of 0,90 als indicator voor multicollineariteit gezien (Agresti & Finlay, 2018).

Tot slot bestaan er significante verbanden tussen de verschillende vormen van middelengebruik, bijvoorbeeld de verbanden tussen alcohol drinken en de overige vormen van middelengebruik ( $r_{bingedrinken}= 0,50$ ;  $p= <0,01$ ;  $r_{tabak\ roken}= 0,35$ ;  $p= <0,01$ ;  $r_{wiet\ roken}= 0,24$ ;  $p= <0,01$ ). Uit de kruistabellen blijkt dat de verbanden tussen de middelen positief zijn, dus mensen die meer van het ene middel gebruiken, in dit geval alcohol, gebruiken ook meer van het andere middel.

Tabel 2: Samenhangsmaten van alle variabelen die zijn opgenomen in de analyse (N=1231)

	1. Gsl.	2. Lft.	3. Mig.acht. resp	4. Mig.acht. vad.	5. Mig.acht. moed.	6. Seks. or.	7. Uit. Gedr.	8. Alc. drink.	9. Bing. drink.	10. Tab. roken	11. Wiet roken
1. Geslacht (1= man, 2= vrouw)	-										
2. Leeftijd	<0,01 <sup>c</sup>	-									
3. Migratieachtergrond respondent (1= geen, 2= niet-westerse)	0,05 <sup>b</sup>	0,05 <sup>c</sup>	-								
4. Migratieachtergrond vader (1= geen, 2= niet-westerse)	0,02 <sup>b</sup>	<0,01 <sup>c</sup>	**0,28 <sup>b</sup>	-							
5. Migratieachtergrond moeder (1= geen, 2= niet-westerse)	0,01 <sup>b</sup>	0,03 <sup>c</sup>	**0,35 <sup>b</sup>	**0,68 <sup>b</sup>	-						
6. Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)	**0,08 <sup>b</sup>	**0,09 <sup>c</sup>	*0,06 <sup>b</sup>	0,01 <sup>b</sup>	0,04 <sup>b</sup>	-					
7. Uitgaansgedrag	0,05 <sup>c</sup>	**0,26 <sup>a</sup>	0,05 <sup>c</sup>	*0,06 <sup>c</sup>	<0,01 <sup>c</sup>	0,04 <sup>c</sup>	-				
8. Alcohol drinken (1= nee, 2= ja)	0,02 <sup>b</sup>	**0,30 <sup>c</sup>	0,01 <sup>b</sup>	**0,13 <sup>b</sup>	**0,43 <sup>b</sup>	**0,09 <sup>b</sup>	**0,25 <sup>c</sup>	-			
9. Bingedrinken (1= nee, 2= ja)	*0,07 <sup>b</sup>	**0,40 <sup>c</sup>	0,05 <sup>b</sup>	0,05 <sup>b</sup>	0,06 <sup>b</sup>	**0,10 <sup>b</sup>	**0,45 <sup>c</sup>	**0,50 <sup>b</sup>	-		
10. Tabak roken (1= nee, 2= ja)	0,05 <sup>b</sup>	**0,25 <sup>c</sup>	*0,06 <sup>b</sup>	*0,06 <sup>b</sup>	0,04 <sup>b</sup>	**0,11 <sup>b</sup>	**0,36 <sup>c</sup>	**0,35 <sup>b</sup>	**0,52 <sup>b</sup>	-	
11. Wiet roken (1= nee, 2= ja)	0,04 <sup>b</sup>	**0,28 <sup>c</sup>	0,01 <sup>b</sup>	0,05 <sup>b</sup>	0,05 <sup>b</sup>	**0,08 <sup>b</sup>	**0,37 <sup>c</sup>	**0,24 <sup>b</sup>	**0,42 <sup>b</sup>	**0,53 <sup>b</sup>	-

<sup>a</sup> Pearson correlatie; <sup>b</sup> Cramer's V; <sup>c</sup> correlatie op basis van ANOVA R<sup>2</sup>

\* significant bij p<0,05, \*\* significant bij p<0,01; tweezijdige toets; N=1231

## Modevaluatie

In deze paragraaf bespreek ik de modelfit, aannames, multicollineariteit en uitbijters. Ten eerste de fit van model 3. In Tabel 3 staan de resultaten van een lineaire regressieanalyse op uitgaansgedrag. Dit model heeft een lage verklaarde variantie ( $R^2 = 0,08$ ;  $R^2\text{-adjusted} = 0,07$ ). Dit betekent dat dit model 8% van de variantie in uitgaansgedrag kan verklaren en de fit van dit model niet erg goed is. De lage verklaarde variantie kan betekenen dat er belangrijke voorspellers voor uitgaansgedrag ontbreken in dit onderzoek. De F-toets laat daarentegen wel een significant resultaat zien, wat betekent dat het model significant meer variantie kan verklaren dan het model zonder variabelen ( $F(6, 1224) = 16,94$ ;  $p = <0,01$ ). De fit van model 3 is dus beter dan de fit van het lege model, ondanks dat het weinig variantie in uitgaansgedrag kan verklaren. Dit komt omdat er variabelen zijn toegevoegd die een significant effect op uitgaansgedrag hebben, zoals leeftijd ( $b = 0,18$ ;  $p = <0,01$ ). Seksuele oriëntatie heeft geen significant effect op uitgaansgedrag, gegeven de andere variabelen ( $b = 0,14$ ;  $p = 0,39$ ). Dus iemands seksuele oriëntatie kan niet goed iemands uitgaansgedrag voorspellen.

Ten tweede de fit van de modellen van alcohol drinken. In Tabel 4 is te zien dat de deviance significant afneemt per model. De toevoeging van seksuele oriëntatie heeft gezorgd voor een afname in deviance van 7,57. Dit is niet heel veel, maar het is wel een significante verbetering (LR-toets<sub>model 2</sub> = 7,57;  $p = <0,01$ ). Dit betekent dat seksuele oriëntatie een goede voorspeller is van alcohol drinken. De toevoeging van uitgaansgedrag zorgt voor een afname in deviance van 64,88. Dit is een redelijk grote afname en ook een significante verbetering van de fit (LR-toets<sub>model 4</sub> = 64,88;  $p = <0,01$ ). Uitgaansgedrag is dus een goede voorspeller van alcohol drinken. Aangezien de afname in deviance groter is voor uitgaansgedrag, betekent dit dat uitgaansgedrag een betere voorspeller is voor alcohol drinken dan seksuele oriëntatie. De Hosmer-Lemeshowtoets bevestigt dat de modellen een goede fit hebben en verbeteren ten opzichte van het vorige model, want zowel model 2 als model 4 hebben een insignificant resultaat voor deze toets en de p-waarde van model 4 is groter dan de p-waarde van model 2, zie Tabel 4 ( $Hosmer\text{-Lemeshow}_{model\ 2} = 5,62$ ;  $p = 0,69$ ;  $Hosmer\text{-Lemeshow}_{model\ 4} = 3,37$ ;  $p = 0,91$ ). De insignificant resultaten betekenen dat er geen grote verschillen zitten tussen de geobserveerde en geschatte data.

Ten derde de fit van de modellen van bingedrinken. Uit Tabel 5 blijkt dat de deviance significant afneemt per model, en dus is ieder model beter dan het vorige model. De toevoeging van zowel seksuele oriëntatie als uitgaansgedrag levert een significante verbetering op (LR-toets<sub>model 2</sub> = 7,47  $p = <0,01$ ; LR-toets<sub>model 4</sub> = 212,44;  $p = <0,01$ ). De afname in deviance door toevoeging van seksuele oriëntatie is niet erg groot; het is namelijk 7,47. De toevoeging van uitgaansgedrag leidt tot een grote afname in deviance van 212,44. De significante resultaten geven een indicatie dat seksuele oriëntatie en uitgaansgedrag goede voorspellers zijn van bingedrinken, maar uitgaansgedrag is een betere

voorspeller dan seksuele oriëntatie, aangezien de daling in deviance door toevoeging van uitgaansgedrag het grootst is. De Hosmer-Lemeshowtoets bevestigt dat de modellen een goede fit hebben, zie Tabel 5 (*Hosmer-Lemeshow*<sub>model 2</sub> = 5,47; *p* = 0,71; *Hosmer-Lemeshow*<sub>model 4</sub> = 8,13; *p* = 0,42). Echter, de Hosmer-Lemeshowtoets laat zien dat model 4 niet beter is dan model 2, aangezien de *p*-waarde voor model 4 kleiner is en er dus grotere verschillen zitten in de geobserveerde en geschatte data. Het is dus onduidelijk of de fit van model 4 beter is dan model 2, omdat de resultaten verschillende bevindingen opleveren.

Ten vierde de fit van de modellen van tabak roken. De deviance neemt per model significant af, zie Tabel 6. Seksuele oriëntatie en uitgaansgedrag zijn goede voorspellers van tabak roken, omdat de resultaten van de likelihood-ratio test significant zijn (LR-toets<sub>model 2</sub> = 10,53; *p* < 0,01; LR-toets<sub>model 4</sub> = 116,87; *p* < 0,01). De afname in deviance door seksuele oriëntatie is 10,53, dus niet erg groot. De afname in deviance door toevoeging van uitgaansgedrag is 116,87; dit is een grote afname en dus is uitgaansgedrag een betere voorspeller van tabak roken dan seksuele oriëntatie. De Hosmer-Lemeshow toets geeft een insignificant resultaat voor model 4, maar een significant resultaat voor model 2, zie Tabel 6 (*Hosmer-Lemeshow*<sub>model 2</sub> = 16,25; *p* = 0,04; *Hosmer-Lemeshow*<sub>model 4</sub> = 87,69; *p* = 0,47). Volgens de Hosmer-Lemeshow toets zijn er significante verschillen tussen de geobserveerde en geschatte data voor model 2, ondanks dat seksuele oriëntatie een goede voorspeller is voor tabak roken volgens de likelihood-ratio toets. Het is dus onduidelijk of de fit van model 2 goed is en of seksuele oriëntatie een goede indicator is voor tabak roken, aangezien de resultaten verschillende bevindingen opleveren.

Tot slot de modevaluatie voor de modellen van wiet roken, zie Tabel 7. De deviance neemt voor ieder model significant af, dus ieder model is beter dan het vorige model. Seksuele oriëntatie en uitgaansgedrag zorgen voor een significante afname in deviance en zijn dus goede voorspellers van wiet roken (LR-toets<sub>model 2</sub> = 4,32; *p* = 0,04; LR-toets<sub>model 4</sub> = 97,43; *p* < 0,01). De deviance neemt met 4,32 af door seksuele oriëntatie. Dit is niet veel, maar wel een significante afname. De afname in deviance van 97,43 door uitgaansgedrag is wel groot, dus uitgaansgedrag is een betere voorspeller van wiet roken dan seksuele oriëntatie. De Hosmer-Lemeshowtoets geeft insignificante resultaten voor model 2 en 4, waardoor deze toets bevestigt dat de fit van deze modellen goed is en seksuele oriëntatie en uitgaansgedrag goede voorspellers zijn voor wiet roken, zie Tabel 7 (*Hosmer-Lemeshow*<sub>model 2</sub> = 14,98; *p* = 0,06; *Hosmer-Lemeshow*<sub>model 4</sub> = 9,32; *p* = 0,32). Ook is te zien dat de *p*-waarde van model 4 groter is dan van model 2, dus model 4 heeft een betere fit dan model 2 volgens deze toets.

Naast een evaluatie van de fit van de modellen worden de assumpties, multicollineariteit en uitbijters gecontroleerd. Een schending van de assumpties, te hoge multicollineariteit of veel uitbijters kan

grote invloed hebben op de geschatte modellen en dus op de resultaten. Het is daarom van belang om inzicht te verkrijgen in deze factoren, zodat er eventueel rekening mee gehouden kan worden in de interpretatie van de resultaten. Hieronder een korte beschrijving van de controle van assumpties, multicollineariteit en uitbijters. Zie Bijlage 3 voor een uitgebreide beschrijving hiervan.

Te beginnen met de assumpties. Logistische en lineaire regressieanalyses moeten voldoen aan de assumptie van onafhankelijke waarnemingen. Dit houdt in dat er geen samenhang bestaat tussen de geobserveerde scores van de respondenten. In het geval van mijn steekproef zijn er vier middelbare scholen benaderd, binnen deze scholen zijn alle studenten benaderd om de vragenlijst in te vullen. Studenten binnen een bepaalde school of klas zijn niet volledig onafhankelijk van elkaar. Zo worden ze bijvoorbeeld beïnvloed door dezelfde docent. Deze assumptie wordt dus geschonden.

Verder moet lineaire regressie voldoen aan drie andere assumpties. Deze assumpties gelden dus voor model 3. De eerste assumptie is lineariteit. Deze assumptie stelt dat er sprake moet zijn van een lineair verband tussen de afhankelijke en de onafhankelijke variabelen. Voor het controleren van deze assumptie kijk ik naar de residual plot. Hieruit blijkt dat de assumptie geschonden is.

De tweede assumptie is homoscedasticiteit. Dit houdt in dat de spreiding van de residuen voor elke waarde van de onafhankelijke variabele gelijk moet zijn. Ik kijk wederom naar de residual plot en concludeer dat deze assumptie ook geschonden wordt.

Tot slot de assumptie van normaliteit. Deze assumptie stelt dat de residuen normaal verdeeld moeten zijn. Deze assumptie controleer ik door naar het histogram en PP-plot te kijken van de residuen. Hieruit concludeer ik dat de assumptie van normaliteit geschonden wordt.

Concluderend worden alle assumpties geschonden. Het is daarom van belang om kritisch naar de verkregen resultaten te kijken, omdat ze mogelijk vertekend zijn. Om enigszins te corrigeren voor de schendingen, zullen de hellingen in de lineaire regressie getoetst worden met een strenger significantieniveau van 0,01 in plaats van 0,05. Op deze manier is er meer zekerheid dat er niet ten onrechte een hypothese verworpen wordt. De hellingen van de logistische regressie zullen getoetst worden met een significantieniveau van 0,05. Dit omdat logistische regressie niet uitgaat van een lineair verband, normaal verdeelde residuen en homoscedasticiteit, en dus robuuster is als het gaat om schendingen van assumpties.

Daarnaast kijk ik naar de multicollineariteit. Om multicollineariteit te controleren kijk ik, naast de bivariate statistieken, naar de VIF-scores. Deze scores geven weer hoeveel de standaardfouten van de onafhankelijke variabelen toenemen door sterke samenhang tussen de onafhankelijke variabelen. Er zijn meerdere grenswaarden die gehanteerd kunnen worden om vast te stellen of een VIF-score hoog

is, zoals 2, 4 of 10 (Agresti & Finlay, 2018). In dit onderzoek hanteer ik een grenswaarde van 4, omdat het niet de strengste of de coulantste grenswaarde is. Het blijkt dat geen enkele VIF-score hoger is dan 4, zie Tabel 3 tot en met 7. Hieruit concludeer ik dat er geen sprake is van veel multicollineariteit.

Ten slotte kijk ik naar of er uitbijters in de dataset zitten, zie Bijlage 3 voor een uitgebreide bespreking. Uitbijters zijn observaties die afwijken van de rest van de observaties, waardoor ze mogelijk veel invloed uitoefenen op de geschatte modellen en resultaten. Er zijn verschillende maten die bekeken kunnen worden om vast te stellen of een observatie een uitbijter is. Voor de lineaire regressie kijk ik naar de leverage, Cook's Distance en DFFIT en voor de logistische regressie naar de leverage en DFBETA. Deze waarden zijn berekend voor de complete modellen, dus model 3 voor de lineaire regressie en model 4 voor de logistische regressies. Uit de analyse blijkt dat er een aantal observaties zijn die afwijken van de rest van de observaties, maar ik kijk naar de meest extreme observaties. Er zijn zes uitbijters voor model 3, 14 voor model 4 van alcohol drinken en 15 voor model 4 van bingedrinken, tabak roken en wiet roken. Als ik de analyses opnieuw doe zonder de uitbijters, zie ik dat er bij de meeste analyses weinig veranderd is. Wat opvalt is dat de invloed van seksuele oriëntatie op alcohol drinken steker is door verwijdering van de uitbijters. Ook valt het op dat de fit van sommige modellen beter is door verwijdering van de uitbijters, maar andere modellen krijgen een slechtere fit. De uitbijters zorgen dus niet altijd voor een slechtere fit en dragen soms bij aan het accuraat voorspellen van de afhankelijke variabele. Aangezien de verwijderde uitbijters geldige observaties zijn, worden ze niet definitief uit de data verwijderd.

## Hypothesetoetsing

Dit onderzoek richt zich op twee hypothesen. In deze paragraaf toets ik deze hypothesen zoals beschreven in de analyse-opzet. De eerste hypothese is: *LHB+ jongeren gebruiken meer middelen dan heteroseksuele jongeren*. De tweede hypothese is: *Het positieve effect van seksuele oriëntatie op middelengebruik wordt gedeeltelijk verklaard door uitgaansgedrag*. De hypothesen toets ik voor iedere vorm van middelengebruik. Echter, om deze paragraaf niet te uitgebreid te maken, kies ik ervoor om de analyse van alcohol drinken uitgebreid te bespreken. De rest van de analyses bespreek ik korter, tenzij er opvallende bevindingen zijn.

### Invloed van seksuele oriëntatie op uitgaansgedrag

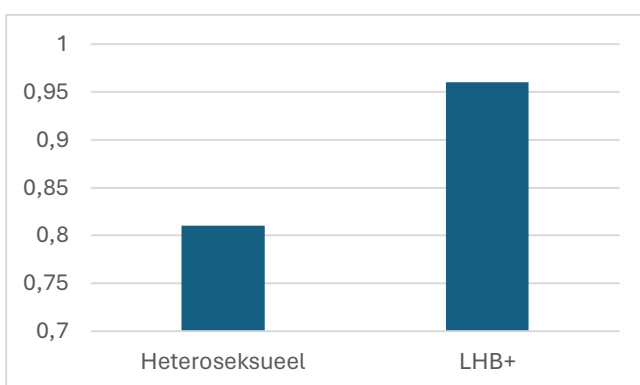
Eerst kijk ik naar de invloed van seksuele oriëntatie op uitgaansgedrag, gegeven de andere variabelen in het model. In Tabel 3 is te zien dat seksuele oriëntatie een zwakke positieve invloed heeft op uitgaansgedrag, gegeven de andere variabelen in het model ( $b = 0,14$ ;  $p = 0,39$ ). Het gevonden effect van seksuele oriëntatie op uitgaansgedrag is klein, dus het kan verwacht worden dat uitgaansgedrag niet een groot deel van het effect van seksuele oriëntatie op middelengebruik kan verklaren. Het



effect van seksuele oriëntatie op uitgaansgedrag is in termen van grootte niet relevant, maar het is wel relevant omdat het effect een indicatie geeft of er een mediatie-effect zal optreden. Het effect is overigens niet significant ( $t(1224) = 0,87$ ;  $p = 0,39$ ). Er wordt dus niet voldaan aan de voorwaarde van een mediatie-effect van uitgaansgedrag, die stelt dat er een significant effect van seksuele oriëntatie op uitgaansgedrag moet zijn.

## Alcohol drinken

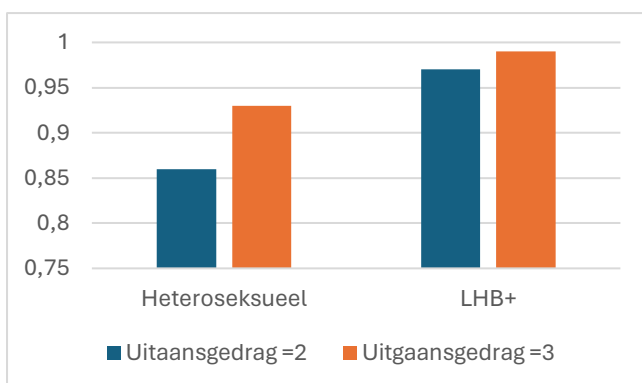
Nu ik de voorwaarde van het mediatie-effect getoetst heb, kan ik de hypothesen toetsen. Het eerste middel waarvoor ik de hypothesen toets is alcohol drinken. In Tabel 4 bij model 2 is te zien dat seksuele oriëntatie een positieve invloed heeft op alcohol drinken, gegeven de andere variabelen ( $b = 1,64$ ;  $p = 0,03$ ;  $odds-ratio = 5,15$ ). Het gevonden effect verschilt ook significant van 0 ( $Wald(1) = 4,84$ ;  $p = 0,03$ ). Dit betekent dat LHB+ jongeren meer alcohol drinken dan heteroseksuele jongeren. Een heteroseksuele jongen van 15 jaar zonder een migratieachtergrond en wiens ouders geen migratieachtergrond hebben, oftewel een jongen die gemiddeld scoort op de variabelen, heeft een kans van 0,82 op alcohol drinken, zie Figuur 2. Eenzelfde LHB+ jongen heeft een kans van 0,96. De bovengenoemde kansen heb ik berekend door eerst de vergelijking in model 2 in Tabel 4 in te vullen en de log-odds uit te rekenen, dus voor de heteroseksuele jongen:  $-7,28 + 0,52 \times 15 - 0,19 \times 1 + 1,29 \times 1 - 0,37 \times 1 - 1,44 \times 1 + 1,64 \times 1 = 1,45$ . Dan neem ik de e-macht van de berekende log-odds en dit deel ik door 1 plus de e-macht van de log-odds:  $e^{1,45} / 1 + e^{1,45} = 0,81$ . Voor de LHB+ jongere geldt:  $-7,28 + 0,52 \times 15 - 0,19 \times 1 + 1,29 \times 1 - 0,37 \times 1 - 1,44 \times 1 + 1,64 \times 2 = 3,09$ . Dus  $e^{3,09} / 1 + e^{3,09} = 0,96$ . Het verschil tussen de kansen is best groot, zie Figuur 2. Seksuele oriëntatie heeft daarmee een redelijk effect op alcohol drinken. Deze bevindingen ondersteunen de eerste hypothese.



Figuur 2: Kans op alcohol drinken voor gemiddeld scorende heteroseksuele en LHB+ jongens

De tweede hypothese stelt dat het positieve effect van seksuele oriëntatie op middelengebruik voor een deel verklaard wordt door uitgaansgedrag. Uit Tabel 3 heb ik geconcludeerd dat seksuele oriëntatie een positieve invloed heeft op uitgaansgedrag, maar dat dit effect insignificant en klein is.

In Tabel 4 bij model 4 is te zien dat uitgaansgedrag een positieve invloed heeft op alcoholgebruik, gegeven de andere variabelen. ( $b = 0,81$ ;  $p = <0,01$ ;  $odds-ratio = 2,24$ ). Dit effect is ook significant ( $Wald(1) = 43,49$ ;  $p = <0,01$ ). Jongeren die meer uitgaan, drinken dus meer alcohol. Een gemiddeld scorende heteroseksuele jongen heeft een kans op bingedrinken van 0,86. Eenzelfde jongen met een punt hoger op uitgaansgedrag heeft een kans van 0,93, zie Figuur 3. Verder heeft een gemiddeld scorende LHB+ jongen een kans op alcohol drinken van 0,97. Dezelfde jongen met een punt hoger op uitgaansgedrag heeft een kans van 0,99, zie Figuur 3. Deze kansen heb ik op dezelfde manier berekend zoals hierboven besproken, maar in dit geval is de helling van uitgaansgedrag toegevoegd aan de vergelijking. Het verschil tussen de kansen tussen de heteroseksuele jongens is matig en tussen de LHB+ jongens klein, dus het effect van uitgaansgedrag alcohol drinken is matig tot klein. Verder is in Tabel 4 bij model 4 een positief effect van seksuele oriëntatie op alcoholgebruik te zien, gegeven de andere variabelen ( $b = 1,59$ ;  $p = 0,04$ ;  $odds-ratio = 4,91$ ). Als ik deze helling vergelijk met de helling van model 2 in Tabel 4, constateer ik dat het effect van seksuele oriëntatie op alcoholgebruik licht is afgenomen ( $b_{model\ 2} = 1,64$ ;  $p = 0,03$ ;  $b_{model\ 4} = 1,59$ ;  $p = 0,04$ ). Als ik het verschil in kans op alcohol drinken tussen een gemiddeld scorende heteroseksuele jongen en LHB+ jongen in model 2 en 4 vergelijk, is te zien dat het eerste verschil ongeveer 15% is en het tweede verschil 11%, zie Figuur 2 en 3. Het verschil in kans is dus ook afgenomen tussen model 2 en 4, maar niet veel. Het blijkt dat uitgaansgedrag weinig kan verklaren in de relatie tussen seksuele oriëntatie en alcohol drinken. Aangezien er niet wordt voldaan aan de voorwaarde van mediatie door het insignificante effect van seksuele oriëntatie op uitgaansgedrag en uitgaansgedrag weinig kan verklaren in de relatie tussen seksuele oriëntatie en alcoholgebruik, kan er niet gesproken worden van een mediatie-effect. Deze bevindingen ondersteunen de tweede hypothese niet.

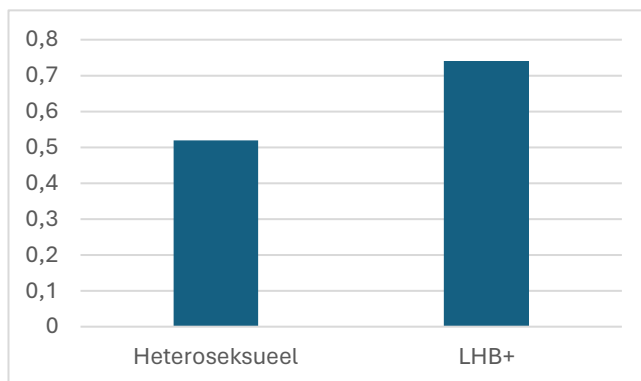


Figuur 3: Kans op alcohol drinken voor gemiddeld scorende heteroseksuele en LHB+ jongens met verschillende scores op uitgaansgedrag

Wat betreft de controlevariabelen zijn er weinig opvallende bevindingen; de effecten zijn zoals verwacht, zie Tabel 4. Het enige wat opvalt is de invloed van de migratieachtergrond van de respondent. De verwachting was dat jongeren met een niet-westerse migratieachtergrond minder middelen gebruiken dan jongeren zonder migratieachtergrond, maar de invloed van de migratieachtergrond van de respondent is positief ( $b_{\text{model 1}} = 1,28$ ;  $p = 0,09$ ;  $\text{odds-ratio}_{\text{model 1}} = 3,58$ ;  $b_{\text{model 2}} = 1,29$ ;  $p = 0,09$ ;  $\text{odds-ratio}_{\text{model 2}} = 3,63$ ;  $b_{\text{model 4}} = 1,03$ ;  $p = 0,19$ ;  $\text{odds-ratio}_{\text{model 4}} = 2,80$ ).

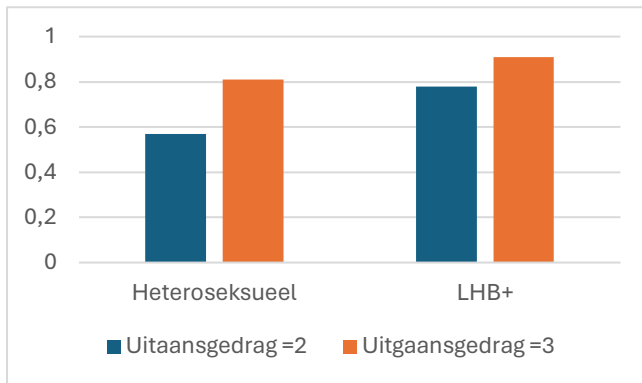
## Bingedrinken

Ten tweede toets ik de hypothesen voor bingedrinken. Uit model 2 in Tabel 5 blijkt dat seksuele oriëntatie een positief effect heeft op bingedrinken, gegeven de andere variabelen ( $b = 0,96$ ;  $p < 0,01$ ;  $\text{odds-ratio} = 2,62$ ). Het effect is ook significant, dus LHB+ jongeren doen meer aan bingedrinken dan heteroseksuele jongeren ( $\text{Wald}(1) = 6,85$ ;  $p < 0,01$ ). Het verschil tussen de kansen is groot, dus het effect is ook groot. Deze bevindingen ondersteunen de eerste hypothese, zie Figuur 4 voor de kansen.



Figuur 4: Kans op bingedrinken voor gemiddeld scorende heteroseksuele en LHB+ jongens

De tweede hypothese veronderstelt een mediatie-effect van uitgaansgedrag, oftewel dat uitgaansgedrag een deel van het effect van seksuele oriëntatie op middelengebruik verklaard. In model 4 in Tabel 5 is te zien dat uitgaansgedrag een positieve invloed heeft op bingedrinken, gegeven de andere variabelen ( $b = 1,13$ ;  $p < 0,01$ ). Tevens is dit effect significant ( $\text{Wald}(1) = 139,09$ ;  $p < 0,01$ ). Jongeren die meer uitgaan, doen dus meer aan bingedrinken, zie Figuur 5 voor de kansen. Het effect is redelijk groot, gezien het verschil tussen de kansen. De helling van seksuele oriëntatie in model 4 is weinig afgenomen in vergelijking met de helling in model 2 ( $b_{\text{model 2}} = 0,96$ ;  $p < 0,01$ ;  $b_{\text{model 4}} = 0,94$ ;  $p = 0,02$ ). Ook het verschil in de kans op bingedrinken tussen een gemiddeld scorende heteroseksuele jongen en een LHB+ jongen in model 2 en 4 is klein, zie Figuur 4 en 5. Het blijkt dat uitgaansgedrag weinig kan verklaren in de relatie tussen seksuele oriëntatie en bingedrinken. Deze bevindingen ondersteunen de tweede hypothese niet.

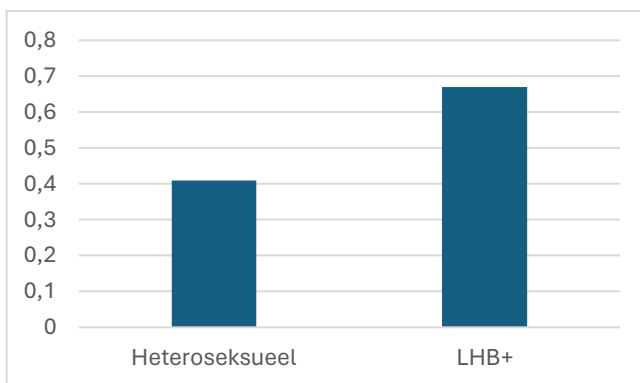


Figuur 5: Kans op bingedrinken voor gemiddeld scorende heteroseksuele en LHB+ jongens met verschillende scores op uitgaansgedrag

Tot slot de controlevariabelen. Hier zijn geen interessante bevindingen, naast dat het effect van de migratieachtergrond van de respondent wederom positief is, zie Tabel 5 ( $b_{\text{model 1}} = 1,54$ ;  $p = 0,02$ ;  $\text{odds-ratio}_{\text{model 1}} = 4,68$ ;  $b_{\text{model 2}} = 1,53$ ;  $p = 0,03$ ;  $\text{odds-ratio}_{\text{model 2}} = 4,59$ ;  $b_{\text{model 4}} = 1,28$ ;  $p = 0,09$ ;  $\text{odds-ratio}_{\text{model 4}} = 3,62$ ).

## Tabak roken

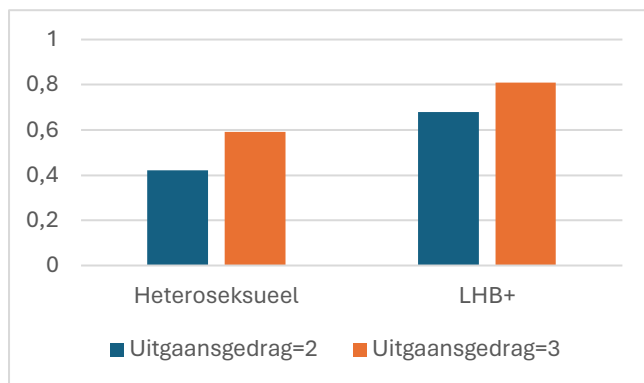
Ten derde tabak roken. De bevindingen ondersteunen de eerste hypothese, want in Tabel 6 bij model 2 is een positief effect van seksuele oriëntatie op tabak roken te zien, gegeven de andere variabelen ( $b = 1,07$ ;  $p < 0,01$ ;  $\text{odds-ratio} = 2,91$ ). Het effect is ook significant ( $Wald(1) = 9,82$ ;  $p < 0,01$ ). Dit betekent dat LHB+ jongeren meer tabak roken dan heteroseksuele jongeren, zie Figuur 6 voor de kansen. Het is een groot effect, want het verschil tussen de kansen is groot.



Figuur 6: Kans op tabak roken voor gemiddeld scorende heteroseksuele en LHB+ jongens

Vervolgens de tweede hypothese, oftewel de verwachting dat uitgaansgedrag een deel van het effect van seksuele oriëntatie op middelengebruik verklaard. Bij model 4 in Tabel 6 is te zien dat uitgaansgedrag een positieve invloed heeft op tabak roken, gegeven de andere variabelen ( $b = 0,68$ ;

$p = <0,01$ ;  $odds-ratio = 1,97$ ). Het effect is ook significant ( $Wald(1) = 93,32$ ;  $p = <0,01$ ). Jongeren die meer uitgaan, roken dus meer tabak, zie Figuur 7 voor de kansen. Het effect is redelijk groot, gezien de verschillen tussen de kansen. De helling van seksuele oriëntatie in model 4 is gedaald ten opzichte van de helling in model 2, maar de daling is niet erg groot ( $b_{model\ 2} = 1,07$ ;  $p = <0,01$ ;  $b_{model\ 4} = 1,05$ ;  $p = <0,01$ ). Ook het verschil in de kans op tabak roken tussen een gemiddeld scorende heteroseksuele jongen en een LHB+ jongen in model 2 en 4 is klein, zie Figuur 6 en 7. Het blijkt dat uitgaansgedrag weinig kan verklaren in de relatie tussen seksuele oriëntatie en tabak roken. De tweede hypothese wordt niet ondersteund door deze bevindingen.

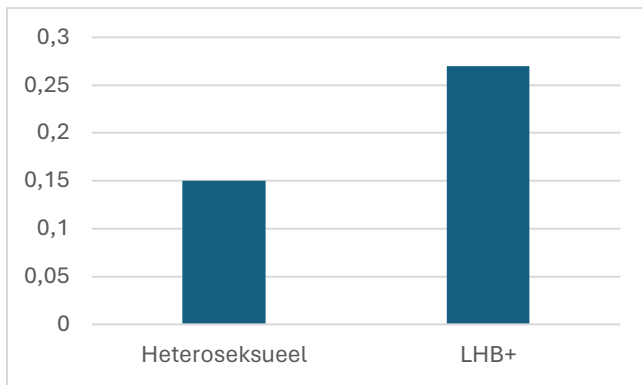


*Figuur 7: Kans op tabak roken voor gemiddeld scorende heteroseksuele en LHB+ jongens met verschillende scores op uitgaansgedrag*

Tot slot de controle variabelen. Ook hier valt het op dat de migratieachtergrond van de respondent een positieve invloed heeft, zie Tabel 6. Verder valt het ook op dat de migratieachtergrond van de vader van de respondent een positieve invloed heeft op tabak. Dit gaat tegen de verwachtingen in, omdat de verwachting was dat jongeren met ouders met een niet-westerse migratieachtergrond minder middelen gebruiken.

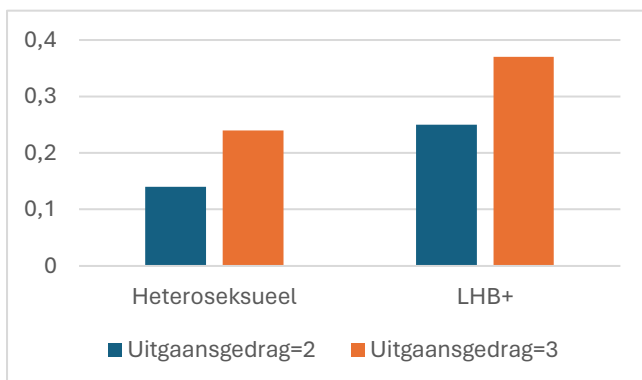
## Wiet roken

Als laatste de analyse van wiet roken, te beginnen met de eerste hypothese. Bij model 2 in Tabel 7 is te zien dat seksuele oriëntatie een positieve invloed heeft op wiet roken, zie Figuur 8 voor de kansen ( $b = 0,75$ ;  $p = 0,03$ ;  $odds-ratio = 2,13$ ). Het effect is ook significant ( $Wald(1) = 4,68$ ;  $p = 0,03$ ). LHB+ jongeren roken dus meer wiet dan heteroseksuele jongeren, gegeven de andere variabelen. Het effect is redelijk groot, gezien het verschil tussen de kansen. Deze bevindingen ondersteunen de eerste hypothese.



Figuur 8: Kans op wiet roken voor gemiddeld scorende heteroseksuele en LHB+ jongens

Voor de tweede hypothese die uitgaat van een mediatie-effect van uitgaansgedrag, kijk ik naar model 4 in Tabel 7. Het is te zien dat uitgaansgedrag een positieve invloed heeft op wiet roken, gegeven de andere variabelen, zie Figuur 9 voor de kansen ( $b = 0,65$ ;  $p = <0,01$ ;  $odds-ratio = 1,91$ ). Het effect is significant ( $Wald(1) = 88,79$ ;  $p = <0,01$ ). Jongeren die meer uitgaan, roken dus meer wiet, gegeven de andere variabelen. Het effect is redelijk groot, gezien het verschil tussen de kansen. De helling van seksuele oriëntatie in model 4 is gedaald ten opzichte van de helling in model 2, maar de daling is niet erg groot ( $b_{model\ 2} = 0,75$ ;  $p = 0,03$ ;  $b_{model\ 4} = 0,74$ ;  $p = 0,04$ ). Ook het verschil in de kans op wiet roken tussen een gemiddeld scorende heteroseksuele jongen en een LHB+ jongen in model 2 en 4 is klein, zie Figuur 8 en 9. Het blijkt dat uitgaansgedrag weinig kan verklaren in de relatie tussen seksuele oriëntatie en wiet roken. Deze bevindingen ondersteunen de tweede hypothese niet.



Figuur 9: Kans op wiet roken voor gemiddeld scorende heteroseksuele en LHB+ jongens met verschillende scores op uitgaansgedrag

Tot slot de controlevariabelen, zie Tabel 7. Het valt op dat de migratieachtergrond van de vader van de respondent een positieve invloed heeft ( $b_{model\ 1} = 0,38$ ;  $p = 0,40$ ;  $odds-ratio_{model\ 1} = 1,46$ ;  $b_{model\ 2} = 0,44$ ;  $p = 0,33$ ;  $odds-ratio_{model\ 2} = 1,55$ ;  $b_{model\ 4} = 0,17$ ;  $p = 0,70$ ;  $odds-ratio_{model\ 4} = 1,19$ ). Wat ook opvallend is, is dat de invloed van de migratieachtergrond van de moeder van de respondent positief

is; dit gaat ook tegen de verwachtingen in ( $b_{\text{model 1}} = 0,34$ ;  $p = 0,49$ ;  $\text{odds-ratio}_{\text{model 1}} = 1,40$ ;  $b_{\text{model 2}} = 0,28$ ;  $p = 0,56$ ;  $\text{odds-ratio}_{\text{model 2}} = 1,33$ ;  $b_{\text{model 4}} = 0,56$ ;  $p = 0,26$ ;  $\text{odds-ratio}_{\text{model 4}} = 1,74$ ).

Tabel 3: Resultaten van een lineaire regressieanalyse met uitgaansgedrag als afhankelijke en seksuele oriëntatie als onafhankelijke variabele (N=1231)

	Model 3 <i>b (se)</i>	<i>p</i>	VIF
Constante	-1,35 (0,43)	<0,01	
Leeftijd	0,18 (0,02)	<0,01	1,01
Geslacht (1= man, 2= vrouw)	-0,14 (0,06)	0,02	1,01
Migratieachtergrond respondent (1= geen, 2= niet-westerse)	0,38 (0,30)	0,20	1,15
Migratieachtergrond vader (1= geen, 2= niet-westerse)	0,49 (0,17)	<0,01	1,94
Migratieachtergrond moeder (1= geen, 2= niet-westerse)	-0,38 (0,19)	0,05	2,05
Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)	0,14 (0,17)	0,39	1,01
<i>R</i> <sup>2</sup>	0,08		
<i>R</i> <sup>2</sup> <i>adjusted</i>	0,07		
<i>F</i>	16,94	<0,01	

Tabel 4: Resultaten van een stapsgewijze regressieanalyse met alcohol drinken als afhankelijke, seksuele oriëntatie als onafhankelijke en uitgaansgedrag als mediërende variabele (N=1231)

	Model 1 <i>b (se)</i>	OR	<i>p</i>	Model 2 <i>b (se)</i>	OR	<i>p</i>	Model 4 <i>b (se)</i>	OR	<i>p</i>	VIF
Constante	-5,82 (1,04)		<0,01	-7,28 (1,26)		<0,01	-6,75 (1,28)		<0,01	
Leeftijd	0,53 (0,05)	1,70	<0,01	0,52 (0,05)	1,68	<0,01	0,43 (0,05)	1,54	<0,01	1,08
Geslacht (1= man, 2= vrouw)	-0,14 (0,14)	0,87	0,34	-0,16 (0,15)	0,85	0,27	-0,12 (0,15)	0,88	0,41	1,01
Migratieachtergrond respondent (1= geen, 2= niet-westerse)	1,28 (0,75)	3,58	0,09	1,29 (0,76)	3,63	0,09	1,03 (0,79)	2,80	0,19	1,15
Migratieachtergrond vader (1= geen, 2= niet-westerse)	-0,42 (0,36)	0,66	0,24	-0,37 (0,37)	0,69	0,32	-0,76 (0,40)	0,47	0,06	1,96
Migratieachtergrond moeder (1= geen, 2= niet-westerse)	-1,34 (0,40)	0,26	<0,01	-1,44 (0,42)	0,24	<0,01	-1,27 (0,45)	0,28	<0,01	2,06
Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)				1,64 (0,75)	5,15	0,03	1,59 (0,76)	4,91	0,04	1,02
Uitgaansgedrag							0,81 (0,12)	2,24	<0,01	1,08
-2LL	1192,80			1185,24			1120,36			
$\chi^2$ -toets	148,92		<0,01	7,57		<0,01	64,88		<0,01	
Hosmer-Lemeshow	9,32		0,32	5,62		0,69	3,37		0,91	



Tabel 5: Resultaten van een stapsgewijze regressieanalyse met bingedrinken als afhankelijke, seksuele oriëntatie als onafhankelijke en uitgaansgedrag als mediërende variabele (N=1231)

	Model 1			Model 2			Model 4			VIF
	<i>b (se)</i>	OR	<i>p</i>	<i>b (se)</i>	OR	<i>p</i>	<i>b (se)</i>	OR	<i>p</i>	
Constante	-8,98 (0,95)	-	<0,01	-9,78 (1,01)	-	<0,01	-9,66 (1,06)		<0,01	
Leeftijd	0,59 (0,05)	1,81	<0,01	0,59 (0,05)	1,79	<0,01	0,49 (0,05)	1,64	<0,01	1,08
Geslacht (1= man, 2= vrouw)	-0,38 (0,13)	0,69	<0,01	-0,40 (0,13)	0,67	<0,01	-0,41 (0,14)	0,67	<0,01	1,01
Migratieachtergrond respondent (1= geen, 2= niet-westerse)	1,54 (0,68)	4,68	0,02	1,53 (0,69)	4,59	0,03	1,28 (0,75)	3,62	0,09	1,15
Migratieachtergrond vader (1= geen, 2= niet-westerse)	-0,10 (0,37)	0,90	0,79	-0,06 (0,37)	0,95	0,88	-0,72 (0,42)	0,49	0,09	1,96
Migratieachtergrond moeder (1= geen, 2= niet-westerse)	-0,93 (0,42)	0,39	0,03	-1,00 (0,43)	0,37	0,02	-0,74 (0,48)	0,48	0,12	2,06
Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)				0,96 (0,37)	2,62	<0,01	0,94 (0,39)	2,56	0,02	1,02
Uitgaansgedrag							1,13 (0,10)	3,08	<0,01	1,08
-2LL	1473,87			1466,40			1253,97			
$\chi^2$ -toets	224,03		<0,01	7,47		<0,01	212,44		<0,01	
Hosmer-Lemeshow	5,47		0,71	5,47		0,71	8,13		0,42	

Tabel 6: Resultaten van een stapsgewijze regressieanalyse met tabak roken als afhankelijke, seksuele oriëntatie als onafhankelijke en uitgaansgedrag als mediërende variabele (n=1231)

	Model 1 b (se)	OR	p	Model 2 b (se)	OR	p	Model 4 b (se)	OR	p	VIF
Constante	-6,56 (0,88)		<0,01	-7,40 (0,93)		<0,01	-6,94 (0,95)		<0,01	
Leeftijd	0,35 (0,04)	1,42	<0,01	0,34 (0,04)	1,41	<0,01	0,24 (0,04)	1,28	<0,01	1,08
Geslacht (1= man, 2= vrouw)	-0,26 (0,12)	0,77	0,04	-0,29 (0,12)	0,75	0,02	-0,23 (0,13)	0,80	0,07	1,01
Migratieachtergrond respondent (1= geen, 2= niet-westerse)	0,86 (0,62)	2,36	0,17	0,78 (0,63)	2,19	0,22	0,52 (0,66)	1,67	0,43	1,15
Migratieachtergrond vader (1= geen, 2= niet-westerse)	0,68 (0,34)	1,98	0,05	0,74 (0,34)	2,10	0,03	0,47 (0,37)	1,60	0,20	1,96
Migratieachtergrond moeder (1= geen, 2= niet-westerse)	-0,31 (0,39)	0,73	0,42	-0,38 (0,39)	0,68	0,33	-0,14 (0,41)	0,87	0,74	2,06
Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)				1,07 (0,34)	2,91	<0,01	1,05 (0,35)	2,84	<0,01	1,02
Uitgaansgedrag							0,68 (0,07)	1,97	<0,01	1,08
-2LL	1563,26			1552,73			1435,86			
X <sup>2</sup> -toets	90,06		<0,01	10,53		<0,01	116,87		<0,01	
Hosmer-Lemeshow	16,01		0,04	16,25		0,04	7,69		0,47	

Tabel 7: Resultaten van een stapsgewijze regressieanalyse met wiet roken als afhankelijke, seksuele oriëntatie als onafhankelijke en uitgaansgedrag als mediërende variabele (N=1231)

	Model 1			Model 2			Model 4			
	<i>b (se)</i>	<i>OR</i>	<i>p</i>	<i>b (se)</i>	<i>OR</i>	<i>p</i>	<i>b (se)</i>	<i>OR</i>	<i>p</i>	<i>VIF</i>
Constante	-9,23 (1,12)		<0,01	-9,85 (1,12)		<0,01	-9,96 (1,21)		<0,01	
Leeftijd	0,52 (0,06)	1,68	<0,01	0,51 (0,06)	1,67	<0,01	0,43 (0,06)	1,54	<0,01	1,08
Geslacht (1= man, 2= vrouw)	-0,24 (0,16)	0,79	0,13	-0,28 (0,16)	0,76	0,9	-0,16 (0,17)	0,85	0,34	1,01
Migratieachtergrond respondent (1= geen, 2= niet-westerse)	-0,69 (0,75)	0,50	0,36	-0,76 (0,74)	0,47	0,31	-0,93 (0,73)	0,40	0,21	1,15
Migratieachtergrond vader (1= geen, 2= niet-westerse)	0,38 (0,45)	1,46	0,40	0,44 (0,45)	1,55	0,33	0,17 (0,46)	1,19	0,70	1,96
Migratieachtergrond moeder (1= geen, 2= niet-westerse)	0,34 (0,49)	1,40	0,49	0,28 (0,48)	1,33	0,56	0,56 (0,49)	1,74	0,26	2,06
Seksuele oriëntatie (1= heteroseksueel, 2= LHB+)				0,75 (0,35)	2,13	0,03	0,74 (0,37)	2,09	0,04	1,02
Uitgaansgedrag							0,65 (0,07)	1,91	<0,01	1,08
-2LL	1020,70			1016,38			918,95			
$\chi^2$ -toets	104,00			4,32			97,43			<0,01
Hosmer-Lemeshow	16,70			14,98			9,23			0,32

## Conclusie

In dit onderzoek stond de invloed van seksuele oriëntatie op middelengebruik centraal, en tot in hoeverre uitgaansgedrag deze invloed kan verklaren. Aan de hand van binaire logistische en lineaire regressie heb ik de volgende onderzoeksvraag beantwoord: *Wat is de invloed van seksuele oriëntatie op middelengebruik en tot in hoeverre kan uitgaansgedrag deze invloed verklaren onder Nederlandse jongeren?* Met behulp van literatuuronderzoek heb ik aan de hand van deze onderzoeksvraag twee hypothesen opgesteld en getoetst. De conclusies die over de hypothesen getrokken zijn op basis van de statistische analyses worden hieronder besproken.

### Seksuele oriëntatie en middelengebruik

Ten eerste de invloed van seksuele oriëntatie op middelengebruik. Op basis van het minority stress framework van Meyer (2003) heb ik de volgende hypothese opgesteld: *LHB+ jongeren gebruiken meer middelen dan heteroseksuele jongeren*. Volgens Meyer (2003) ervaren LHB+ jongeren meer stress dan heteroseksuele jongeren. Dit komt doordat LHB+ jongeren, bovenop alledaagse stress, ook minderheidsstress ervaren (Meyer, 2003). Een manier om om te gaan met stress is middelengebruik en daarom verwachtte ik dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren (Mereish, 2019). Uit de resultaten van de analyses is gebleken dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren, dus de eerste hypothese wordt ondersteund. Deze bevindingen zijn in lijn met het minority stress framework (Meyer, 2003). Deze bevinding geeft daarentegen geen informatie over de mate van stress die LHB+ jongeren ervaren, dus dat deel van de theorie van Meyer (2003) kan niet bevestigd worden aan de hand van deze bevinding.

Verder was het opvallend dat de migratieachtergrond van de respondent een positieve invloed heeft op alcohol drinken, bingedrinken en tabak roken. Daarnaast heeft de migratieachtergrond van de vader een positieve invloed op tabak en wiet roken. De migratieachtergrond van de moeder heeft een positieve invloed op wiet roken. Deze bevindingen gaan tegen de verwachtingen in.

### Uitgaansgedrag

Ten tweede de rol van uitgaansgedrag in de relatie tussen seksuele oriëntatie en middelengebruik. De verwachting is dat LHB+ jongeren meer uitgaan in LHB+ uitgaansgelegenheden en meer uitgaan in het algemeen dan heteroseksuele jongeren (Mereish, 2019; Maliepaard, 2013; Van de Goor et al., 2011). Er zijn theorieën die stellen dat door blootstelling aan middelengebruik mensen het idee krijgen dat middelengebruik de norm is en geaccepteerd wordt door anderen (Perkins, 2003; Berkowitz, 2004). Aangezien er meer middelen gebruikt worden in het uitgaansleven dan in het dagelijks leven, zullen mensen (meer) middelen gebruiken om aan de waargenomen norm te voldoen (Perkins, 2003). Uit onderzoek blijkt ook dat er meer middelen worden gebruikt in LHB+ uitgaansgelegenheden dan in

het algemene uitgaansleven (Green & Feinstein, 2012). Uitgaan in LHB+ uitgaansgelegenheden en meer uitgaan in het algemeen zorgt daardoor voor meer middelengebruik en daarom was de verwachting dat uitgaansgedrag een deel van de relatie tussen seksuele oriëntatie en middelengebruik verklaren kan. Op basis hiervan heb ik de tweede hypothese opgesteld: *Het positieve effect van seksuele oriëntatie op middelengebruik wordt gedeeltelijk verklaard door uitgaansgedrag.*

Uit de resultaten is gebleken dat LHB+ jongeren niet significant meer uitgaan dan heteroseksuele jongeren. Dit gaat tegen de verwachtingen in. Verder is gebleken dat jongeren die meer uitgaan, meer middelen gebruiken. Dit is wel in lijn met de verwachtingen. Ondanks dat de data niks weergeeft over de normen die jongeren ervaren omtrent middelengebruik in het uitgaansleven, levert deze bevinding wel een indicatie dat er mogelijk normen van middelengebruik in het uitgaansleven zijn. De theorie van Perkins (2003) en Berkowitz (2004) wordt dus niet direct bevestigd door deze bevinding, maar de verwachting op basis van deze theorie wel. Uit iedere analyse is gebleken dat uitgaansgedrag een klein deel van de invloed van seksuele oriëntatie op middelengebruik kan verklaren, maar het is te klein om te spreken van een mediatie-effect. Daarnaast is er niet voldaan aan de voorwaarde voor een mediatie, aangezien er geen effect van seksuele oriëntatie op uitgaansgedrag gevonden is. De tweede hypothese wordt dus niet ondersteund. Dit kan komen doordat de relatie tussen seksuele oriëntatie en uitgaansgedrag mogelijk niet accuraat is geschat, omdat alle assumpties van de lineaire regressie geschonden zijn. Daarnaast is het mogelijk dat de theorieën over uitgaan niet of niet volledig opgaan voor (plattelandse) jongeren. Wellicht dat jongeren niet genoeg in aanraking komen met het uitgaansleven, omdat er niet veel keten zijn in de regio en dat deze keten klein zijn en niet iedereen daarheen kan. Ook is het mogelijk dat leeftijd toch een belemmering is en jongeren onder de 18 daarvoor geen toegang krijgen tot uitgaansgelegenheden of dat hun ouders niet toestaan dat ze uitgaan. De steekproef van dit onderzoek bestaat voornamelijk uit jongeren onder de 18 (98,3%). Als deze restrictie opgaat voor de jongeren in deze steekproef, is dat een verklaring voor het lage gemiddelde van uitgaansgedrag, waardoor er mogelijk geen groot verschil in uitgaansgedrag gevonden is tussen LHB+ en heteroseksuele jongeren. Vervolgonderzoek moet daarom een representatievere steekproef gebruiken in termen van leeftijd, zodat de jongeren van 18 jaar of ouder ook gerepresenteerd worden in de resultaten en er dan mogelijk wel een significant verschil in uitgaansgedrag gevonden wordt tussen LHB+ en heteroseksuele jongeren. Dan wordt er wel aan de voorwaarde van mediatie voldaan.

Concluderend is het antwoord op de onderzoeksvraag dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren, maar dat uitgaansgedrag weinig van deze invloed kan verklaren onder Nederlandse jongeren. Nederlandse LHB+ jongeren bevestigen daarmee eerdere bevindingen over het verschil in middelengebruik tussen LHB+ jongeren en heteroseksuele jongeren. Ze bevestigen

daarentegen geen theorieën die stellen dat frequent uitgaan een deel van de relatie tussen middelengebruik en seksuele oriëntatie kan verklaren.

## Discussie

Dit onderzoek kent een aantal beperkingen die in acht genomen moeten worden. Een eerste beperking is de representativiteit van de steekproef met betrekking tot de leeftijd, de migratieachtergrond van zowel de respondent als de ouders en de regio waar de respondenten vandaan komen. In deze steekproef zitten namelijk enkel jongeren uit Noord-Nederland en deze regio wijkt af van andere regio's in Nederland door bijvoorbeeld andere heersende normen en waarden. Ook is het aantal respondenten in de steekproef met 43% gedaald door het uitfilteren van respondenten met missende waarden, mogelijk zijn dit respondenten met specifieke kenmerken en zijn de uitkomsten niet meer representatief voor deze groep. Verder is de groep LHB+ jongeren klein in deze steekproef (3,5%). Dit zorgt ervoor dat de gevonden resultaten mogelijk niet goed te generaliseren zijn naar de gehele populatie Nederlandse jongeren, omdat een kleine steekproef van LHB+ jongeren uit een bepaalde regio wellicht geen accuraat beeld geeft van de gehele populatie Nederlandse LHB+ jongeren. De kleine groep LHB+ jongeren in deze steekproef kan te verklaren zijn door de plaats van dataverzameling, namelijk Noord-Nederland. In Noord-Nederland gelden er veelal Protestants conservatieve waarden (Brons, 2006). Protestants conservatieve mensen zijn minder tolerant naar homoseksuelen dan de algemene Nederlandse bevolking (Burdette et al., 2005; NOS, 2018). Het is mogelijk dat LHB+ jongeren schaamte voelen voor hun seksuele oriëntatie en daardoor in de kast blijven of sociaal wenselijk antwoorden, omdat ze weten dat hun seksuele oriëntatie niet geaccepteerd wordt (Meyer, 2003). Hierdoor wordt de groep van LHB+ jongeren mogelijk onderschat. Sociaal wenselijk antwoorden kan ook opgetreden zijn bij middelengebruik en uitgaansgedrag. Mogelijk hebben respondenten die veel uitgaan en/of middelen gebruiken een lagere score opgegeven uit schaamte. Het kan ook dat respondenten een hogere score hebben opgegeven om stoer over te komen tegenover hun klasgenoten (Tourangeau & Yan, 2007). Het is niet bekend welk van deze mechanismen het grootste effect heeft gehad voor dit onderzoek, maar het is goed om rekening te houden met mogelijke vertekening van de resultaten door sociaal wenselijk antwoorden. Echter, aangezien ik de variabelen van middelengebruik heb gehercodeerd naar 'ja' of 'nee', is het voor deze variabelen een minder groot probleem als mensen sociaal wenselijk antwoorden. Eenmalig gebruik van een middel is namelijk voldoende om in dit onderzoek in de categorie 'ja' te worden geplaatst. Dus zolang respondenten niet aangeven dat ze nooit een middel hebben gebruikt terwijl dat wel het geval is, is het geen probleem. Vervolgonderzoek zou sociaal wenselijk antwoorden (gedeeltelijk) tegen kunnen gaan door de vragenlijst digitaal te verspreiden, waardoor de

respondenten het in hun eigen tijd en alleen kunnen invullen. Het is namelijk mogelijk dat de aanwezigheid van klasgenoten of de onderzoekers ervoor zorgt dat respondenten sociaal wenselijker antwoorden (Tourangeau & Yan, 2007).

Een andere beperking van dit onderzoek heeft betrekking op de validiteit en de antwoordopties van de variabelen seksuele oriëntatie en geslacht, zoals eerder uitgebreid besproken in de Methodenparagraaf bij de operationalisaties. Te beginnen met seksuele oriëntatie. Er werd gevraagd wat iemand denkt te zijn en daarmee specificeert de vraag niet of het naar bijvoorbeeld iemands seksuele oriëntatie of identiteit vraagt. Het is onzeker of alle respondenten hun seksuele oriëntatie hebben ingevuld, hierdoor is het mogelijk dat de hoeveelheid LHB+ jongeren in de steekproef in het echt groter of kleiner is. Verder houdt deze vraag geen rekening met andere seksuele oriëntaties, zoals panseksualiteit. Dit is een probleem, omdat respondenten met een andere seksuele oriëntatie mogelijk 'weet ik niet' of 'geen antwoord' hebben ingevuld of de vraag overgeslagen hebben, en daardoor niet meegenomen zijn in het onderzoek. Hierdoor kan de groep LHB+ jongeren in de steekproef onderschat zijn wat problematisch is voor onder andere de representativiteit. Bij toekomstig onderzoek is het van belang dat gespecificeerd wordt naar welke vorm van seksualiteit er gevraagd wordt, en moeten er meer seksuele oriëntaties opgenomen worden in de antwoordopties.

De tweede variabele waarbij de validiteit en de antwoordopties mogelijk problematisch is, is geslacht. Het is namelijk onzeker of geslacht of gender gemeten is. Daarnaast wordt er geen rekening gehouden met transgenders of non-binaire mensen. Het is onbekend of deze mensen tot de steekproef behoren, maar als ze in de steekproef zitten is het onduidelijk hoe deze personen deze vraag beantwoord hebben. Mogelijk vertonen transgenders of non-binaire mensen ander gedrag dan cisgender personen, en kan dit interessante bevindingen opleveren voor vervolgonderzoek. In toekomstig onderzoek moet dus duidelijk gemaakt worden of er naar geslacht of gender gevraagd wordt, samen met een definitie. Ook moet er rekening gehouden worden met transgenders of non-binaire mensen, door bijvoorbeeld een optie 'anders' te geven of een open antwoordmogelijkheid te geven zodat respondenten die zich niet herkennen in de antwoordopties, daar het antwoord kunnen geven dat bij hen past.

Een andere beperking is dat de assumpties van zowel de logistische als de lineaire regressieanalyse worden geschonden. Hierdoor kunnen de resultaten niet volledig betrouwbaar en/of accuraat zijn. Ondanks dat er voor de lineaire regressieanalyse strenger getoetst is met een lager significantieniveau van 0,01, is het goed om rekening te houden met mogelijke onbetrouwbaarheid van de resultaten.

Dit onderzoek kent ook sterke punten. Het eerste sterke punt is dat mijn onderzoek een toevoeging is aan de huidige wetenschappelijke literatuur. Het was namelijk nog niet onderzocht of er tussen Nederlandse LHB+ jongeren en heteroseksuele jongeren verschillen zaten in uitgaansgedrag, en of uitgaansgedrag een deel van het effect van seksuele oriëntatie op middelengebruik kon verklaren. Ik heb geen mediatie-effect van uitgaansgedrag gevonden. Echter, dit onderzoek geeft hierdoor wel aanleiding tot vervolgonderzoek, dat kan leren van de beperkingen van dit onderzoek.

Ten tweede was onderzoek naar de relatie tussen seksuele oriëntatie en middelengebruik nog niet gedaan met een steekproef van middelbare scholieren uit Noord-Nederland. Dit onderzoek heeft interessante bevindingen opgeleverd met betrekking tot de mogelijke restricties die (plattelands) jongeren ervaren met betrekking tot uitgaan.

## Vervolgonderzoek

De resultaten en beperkingen van dit onderzoek geven aanleiding voor vervolgonderzoek. Zo is er geen mediatie-effect gevonden voor uitgaansgedrag, maar de theorie en eerdere bevindingen maken het aannemelijk dat er mogelijk wel een mediatie-effect van uitgaansgedrag is. Vervolgonderzoek moet letten op een aantal zaken. Zo was de steekproef van dit onderzoek niet representatief. Het is daarom van belang dat er vervolgonderzoek uitgevoerd wordt dat gebruikmaakt van een representatieve steekproef van Nederlandse jongeren. Daarnaast moet vervolgonderzoek kritisch kijken naar de mogelijkheden voor jongeren om uit te gaan, bijvoorbeeld door leeftijdsrestricties of door een te grote afstand van uitgaansgelegenheden. Vervolgonderzoek kan mogelijk eerst kijken naar de restricties die jongeren ervaren betreffende uitgaan aan de hand van enquêtes of interviews. Verder moet vervolgonderzoek meer aandacht besteden aan het meten van seksuele oriëntatie en geslacht, door duidelijker aan te geven welk concept gemeten wordt en meer antwoordopties te bieden. Wellicht kunnen onderzoekers eerst in gesprek gaan met (LHB+) jongeren, om na te gaan of de opgestelde vragen begrijpelijk zijn en de antwoordopties toereikend.

Een ander aandachtspunt voor vervolgonderzoek zijn andere sociale contexten waarin jongeren middelen kunnen gebruiken, zoals bij iemand thuis of in een parkje met vrienden. Mogelijk zijn deze locaties toegankelijker voor jongeren om middelen te gebruiken, en kunnen deze sociale contexten wel de relatie tussen seksuele oriëntatie en middelengebruik (gedeeltelijk) verklaren.

## Tot slot

Dit onderzoek bevestigt eerdere bevindingen dat LHB+ jongeren meer middelen gebruiken dan heteroseksuele jongeren. Er is geen mediatie-effect gevonden van uitgaansgedrag. Ondanks de beperkingen van dit onderzoek zijn deze bevindingen relevant. Er is namelijk een toename in het



aantal LHB+ mensen, waaronder LHB+ jongeren (Jones, 2022). De mogelijke risico's van veelvuldig middelengebruik onderstrepen het belang van beleid dat zich richt op het verminderen van het middelengebruik van LHB+ jongeren. Zo kunnen middelbare scholen bijvoorbeeld inzetten op campagnes, die leerlingen bewust maken van de negatieve gevolgen van discriminatie op LHB+ jongeren. Mogelijk kan minderheidsstress verminderd worden door dit beleid, omdat er daadwerkelijk minder gediscrimineerd wordt of doordat het voeren van dergelijk beleid een positieve boodschap afgeeft aan LHB+ jongeren waardoor ze zich meer gewaardeerd voelen. Daarnaast moet het mediatie-effect van uitgaansgedrag nogmaals onderzocht worden met een betere steekproef, een betere vragenlijst en een kritische blik op de theorie. Als uit dit onderzoek blijkt dat er wel een mediatie-effect optreedt, kan er beleid gevoerd worden dat invloed uitoefent op de norm van middelengebruik in het uitgaansleven, bijvoorbeeld door campagnes die nuchter uitgaan bevorderen.

# Literatuurlijst

Agresti, A., & Finlay, B. (2018). *Statistical Methods for the Social Sciences, Global Edition*.

Alexander, J. C. (2003). *The Meanings of Social Life: A Cultural Sociology*.

[https://is.muni.cz/el/1423/jaro2012/SOC403/um/ALEXANDER\\_The\\_meanings\\_of\\_social\\_life.pdf](https://is.muni.cz/el/1423/jaro2012/SOC403/um/ALEXANDER_The_meanings_of_social_life.pdf)

Barni, D., Ranieri, S., Donato, S., Tagliabue, S., & Scabini, E. (2017). Personal and Family Sources of Parents' Socialization Values: A Multilevel Study. *Avances en Psicología Latinoamericana*, 35(1), 9. <https://doi.org/10.12804/revistas.urosario.edu.co/apl/a.3468>

Bhandari, P. (2023, 16 juni). *Een introductie tot controlevariabelen*. Scribbr.

<https://www.scribbr.nl/onderzoeksmethoden/controlevariabelen/>

Berkowitz, A. D. (2004). *The Social Norms Approach: Theory, Research, and Annotated Bibliography*. [social\\_norms.pdf](#)

Bhugra, D. (2004). Migration, distress and cultural identity. *British Medical Bulletin*, 69(1), 129–141. <https://doi.org/10.1093/bmb/ldh007>

Boyle, S., LaBrie, J. W., & Omoto, A. M. (2020). Normative substance use antecedents among sexual minorities: A scoping review and synthesis. *Psychology Of Sexual Orientation And Gender Diversity*, 7(2), 117–131. <https://doi.org/10.1037/sgd0000373>

Brajša-Žganec, A., Merkaš, M., & Šverko, I. (2010). Quality of Life and Leisure Activities: How do Leisure Activities Contribute to Subjective Well-Being? *Social Indicators Research*, 102(1), 81–91. <https://doi.org/10.1007/s11205-010-9724-2>

Brons, L. L. (2006). INDIRECT MEASUREMENT OF REGIONAL CULTURE IN THE NETHERLANDS. *Tijdschrift Voor Economische en Sociale Geografie*, 97(5), 547–566. <https://doi.org/10.1111/j.1467-9663.2006.00363.x>

Brown, B. B., Bakken, J. P., Ameringer, S. W., & Mahon, S. D. (2008). A comprehensive conceptualization of the peer influence process in adolescence. *Understanding peer influence in children and adolescents*, 13, 17-44. [BrownBakkenAmeringerandMahon\\_2008.pdf](#)

- Burdette, A. M., Ellison, C. G., & Hill, T. D. (2005). Conservative Protestantism and Tolerance toward Homosexuals: An Examination of Potential Mechanisms\*. *Sociological Inquiry*, 75(2), 177–196. <https://doi.org/10.1111/j.1475-682x.2005.00118.x>
- Center for Behavioral Health Statistics and Quality (2017). *Results from the 2016 National Survey on Drug Use and Health: Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2017. <https://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2016/NSDUH-DetTabs-2016.pdf>.
- Centraal Bureau voor de Statistiek. (z.d.-a). *Migratieachtergrond*. Centraal Bureau Voor de Statistiek. <https://www.cbs.nl/nl-nl/onze-diensten/methoden/begrippen/migratieachtergrond>
- COC. (2021, 31 augustus). *Homo- en biseksuele mannen mogen vanaf vandaag bloed doneren*. Ge raadpleegd op 19 maart 2024, van <https://coc.nl/nieuws-en-publicaties/homo-en-biseksuele-mannen-mogen-vanaf-vandaag-bloed-doneren/>
- Cochran, S. D., Grella, C. E., & Mays, V. M. (2012). Do substance use norms and perceived drug availability mediate sexual orientation differences in patterns of substance use? Results from the California Quality of Life Survey II. *Journal of studies on alcohol and drugs*, 73(4), 675–685. <https://doi.org/10.15288/jsad.2012.73.675>
- Csányi, G. M., Nagy, D., Vági, R., Vadász, J. P., & Orosz, T. (2021). Challenges and Open Problems of Legal Document Anonymization. *Symmetry*, 13(8), 1490. <https://doi.org/10.3390/sym13081490>
- Degenhardt, L., Stockings, E., Patton, G. C., Hall, W., & Lynskey, M. T. (2016). The increasing global health priority of substance use in young people. *The Lancet Psychiatry*, 3(3), 251–264. [https://doi.org/10.1016/s2215-0366\(15\)00508-8](https://doi.org/10.1016/s2215-0366(15)00508-8)
- De Muijnck, J. A., Schoonbeek, I., Snippe, J., Pieper, R., & Breuer&Intraval. (2024). De regenboog kleuren: Evaluatie van het Actieplan Veiligheid Lhbt 2019-2022. In *Breuer&Intraval – Evaluatie van het Actieplan Veiligheid Lhbt 2019-2022* [Report]. [pdf \(overheid.nl\)](#)

- Dimova, E., Lekkas, P., Maxwell, K. L., Clemens, T., Pearce, J., Mitchell, R., Emslie, C., & Shortt, N. (2023). Exploring the influence of local alcohol availability on drinking norms and practices: A qualitative scoping review. *Drug And Alcohol Review*, 42(3), 691–703. <https://doi.org/10.1111/dar.13596>
- Duffing, T. M., Greiner, S. G., Mathias, C. W., & Dougherty, D. M. (2014). Stress, Substance Abuse, and Addiction. In *Current topics in behavioral neurosciences* (pp. 237–263). [https://doi.org/10.1007/7854\\_2014\\_276](https://doi.org/10.1007/7854_2014_276)
- Eggington, R., Williams, L., & Parker, H. (2002). Going out drinking: the centrality of heavy alcohol use in English adolescents' leisure time and poly-substance-taking repertoires. *Journal Of Substance Use*, 7(3), 125–135. <https://doi.org/10.1080/14659890209169342>
- Fromme, J. (2006). Socialisation in the Age of New Media. *MedienPädagogik: Zeitschrift Für Theorie Und Praxis Der Medienbildung*, 11(Media Education), 1–29. <https://doi.org/10.21240/mpaed/11/2006.01.17.x>
- Gabbert, W. (2006). Concepts of Ethnicity. *Latin American And Caribbean Ethnic Studies*, 1(1), 85–103. <https://doi.org/10.1080/17486830500510034>
- Goossens, F.X., Frijns, T., van Hasselt, N.E. & van Laar, M.W. (2013). Het grote uitgaansonderzoek 2013. Uitgaanspatronen, middelengebruik en risicogedrag onder uitgaande jongeren en jongvolwassensen. Utrecht: Trimbos [Monitoring middelengebruik in het uitgaansleven \(trimbos.nl\)](https://www.trimbos.nl)
- Green, K. E., & Feinstejn, B. A. (2012). Substance use in lesbian, gay, and bisexual populations: An update on empirical research and implications for treatment. *Psychology Of Addictive Behaviors*, 26(2), 265–278. <https://doi.org/10.1037/a0025424>
- Haartsen, T., & Strijker, D. (2010). Rural youth culture: Keten in the Netherlands. *Journal Of Rural Studies*, 26(2), 163–172. <https://doi.org/10.1016/j.jrurstud.2009.11.001>
- Haidt, J. (2013). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Penguin UK.

- Halpern-Felsher, B. L., & Biehl, M. (2004). Developmental and environmental influences on underage drinking: A general overview. *Reducing underage drinking: A collective responsibility*, 402-416. [Developmental and Environmental Influences on Underage Drinking: A General Overview - Reducing Underage Drinking - NCBI Bookshelf \(nih.gov\)](#)
- Jones, B. J. M. (2022). LGBT Identification in U.S. Ticks Up to 7.1%. *Gallup.com*. <https://news.gallup.com/poll/389792/lgbt-identification-ticks-up.aspx>
- Kieken, W., La Roi, C., Bos, H. M. W., Kretschmer, T., Van Bergen, D. D., & Veenstra, R. (2020). Explaining Health Disparities between Heterosexual and LGB Adolescents by Integrating the Minority Stress and Psychological Mediation Frameworks: Findings from the TRAILS Study. *Journal Of Youth And Adolescence*, 49(9), 1767–1782. <https://doi.org/10.1007/s10964-020-01206-0>
- Litt, D. M., Lewis, M. A., Rhew, I. C., Hodge, K. A., & Kaysen, D. L. (2015). Reciprocal relationships over time between descriptive norms and alcohol use in young adult sexual minority women. *Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors*, 29(4), 885–893. <https://doi.org/10.1037/adb0000122>
- Maliepaard, E. (2013). Seksuele interacties in het nachtleven: case study uit Brighton & Hove (UK). *Vrijtijdsstudies*, 31(2), 23–33. [https://www.nritmedia.nl/file/VTS\\_2013\\_2\\_Maliepaard.pdf](https://www.nritmedia.nl/file/VTS_2013_2_Maliepaard.pdf)
- Marshal, M. P., Friedman, M., Stall, R., King, K. M., Miles, J. C., Gold, M. A., Bukstein, O. G., & Morse, J. Q. (2008). Sexual orientation and adolescent substance use: a meta-analysis and methodological review. *Addiction*, 103(4), 546–556. <https://doi.org/10.1111/j.1360-0443.2008.02149.x>
- Mays, V. M., & Cochran, S. D. (2001). Mental Health Correlates of Perceived Discrimination Among Lesbian, Gay, and Bisexual Adults in the United States. *American Journal Of Public Health*, 91(11), 1869–1876. <https://doi.org/10.2105/ajph.91.11.1869>
- Mereish, E. H. (2019). Substance use and misuse among sexual and gender minority youth. *Current Opinion in Psychology*, 30, 123–127. <https://doi.org/10.1016/j.copsyc.2019.05.002>

- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, 129(5), 674–697.  
<https://doi.org/10.1037/0033-2909.129.5.674>
- Moser, C. (2015). Defining sexual orientation. *Archives Of Sexual Behavior*, 45(3), 505–508.  
<https://doi.org/10.1007/s10508-015-0625-y>
- Newcomb, M. D., & Locke, T. F. (2006). Health, Social, and Psychological Consequences of Drug Use and Abuse. In *Springer eBooks* (pp. 45–59). [https://doi.org/10.1007/0-387-24416-6\\_4](https://doi.org/10.1007/0-387-24416-6_4)
- Perkins, H. W. (2003). The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians. <http://psycnet.apa.org/psycinfo/2003-02237-000>
- Peters, S. (2023, 1 april). *Nachthoreca verhogen toegangsleeftijd naar 21 jaar in steeds meer steden - Hormax*. Hormax. Geraadpleegd op 19 maart 2024, van <https://hormax.nl/nachthoreca-verhogen-toegangsleeftijd-naar-21-jaar/>
- Pini, N. (2019). Semi-urban or semi-rural settlements: A new definition of urban centres required?. *Archaeology and Economy in the Ancient World. Judaea/Palaestina and Arabia: Cities and Hinterlands in Roman and Byzantine Times*, 44, 5-17. <https://doi.org/10.11588/propylaeum.552>
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629.  
<https://doi.org/10.1037/0003-066x.52.6.613>
- Stemmingmakers. (2022). Alternatief uitgaan in coronatijd: Een onderzoek naar de ervaringen van jongeren (16 t/m 25) op gebied van veiligheid in het alternatieve uitgaansleven tijdens de coronacrisis. In *hetccv.nl*. Centrum voor Criminaliteitspreventie en Veiligheid. Geraadpleegd op 25 februari 2024, van [https://hetccv.nl/app/uploads/2023/11/Rapportage\\_CCV\\_Uitgaan\\_in\\_Coronatijd.pdf](https://hetccv.nl/app/uploads/2023/11/Rapportage_CCV_Uitgaan_in_Coronatijd.pdf)

- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>
- Trocki, K. F., Drabble, L., & Midanik, L. (2005). Use of heavier drinking contexts among heterosexuals, homosexuals and bisexuals: results from a National Household Probability Survey. *Journal Of Studies On Alcohol*, 66(1), 105–110. <https://doi.org/10.15288/jsa.2005.66.105>
- Van de Goor, I. A. M., Spijkerman, R., Van Den Eijnden, R., & Knibbe, R. A. (2011). Drinking Patterns and Going-Out Behavior as Predictors of Illicit Substance Use: An Analysis among Dutch Adolescents. *Journal Of Child & Adolescent Substance Abuse*, 20(2), 99–113. <https://doi.org/10.1080/1067828x.2011.555267>
- Van den Broek, A. & Keuzenkamp, S. (red). (2008). Het dagelijkse leven van allochtone stedelingen. Den Haag: Sociaal en Cultureel Planbureau. [Het dagelijks leven van allochtone stedelingen \(scp.nl\)](https://www.scp.nl)
- Van Laar, M., Monshouwer, K., & Van Den Brink, W. (2010). Roken, drinken en blowen door de Nederlandse jeugd. *Kind en Adolescent*, 31(4), 204–220. <https://doi.org/10.1007/s12453-010-0722-6>
- Van Leijenhorst, L., & Crone, E. A. (2009). Het adolescentenbrein: inzichten in risicovol gedrag in de adolescentie uit de cognitieve neurowetenschappen. *Neuropraxis*, 13(1), 3–7. <https://doi.org/10.1007/bf03080140>
- Van Zwieten, M., & Willems, D. (2004). Waardering van kwalitatief onderzoek. *Huisarts en Wetenschap*, 47(13), 38–43. <https://doi.org/10.1007/bf03083653>
- World Health Organization. (2023, februari). *Stress*. World Health Organization. Geraadpleegd op 25 februari 2024, van <https://www.who.int/news-room/questions-and-answers/item/stress>

# Bijlage 1: variabelen in de analyse en hun totstandkoming

Een aantal van de variabelen in dit onderzoek zijn gehercodeerd en/of zijn tot een schaal geconstrueerd. In deze bijlage staat omschreven welke hercoderingen er zijn gedaan en waarom, ook wordt de bijbehorende syntax van de hercoderingen gegeven. Daarnaast worden van zowel de oorspronkelijke variabele als de nieuwe variabele de beschrijvende statistieken gegeven.

## De afhankelijke variabele: middelengebruik

### Alcohol drinken

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van alcohol drinken\*

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

## Frequencies

### Statistics

drink How often did you drink alcohol?

N	Valid	1779
	Missing	380
Mean		2,28
Median		2,00
Mode		2
Std. Deviation		1,062
Minimum		1
Maximum		5
Percentiles	25	1,00
	50	2,00
	75	3,00



### drink How often did you drink alcohol?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 I have never drunk alcohol, not even a sip	500	23,2	28,1	28,1
	2 I drank alcohol 1 or 2 times	589	27,3	33,1	61,2
	3 I drink alcohol one to two times a month	397	18,4	22,3	83,5
	4 I drink alcohol one to two times a week	278	12,9	15,6	99,2
	5 I drink alcohol every day	15	,7	,8	100,0
	Total	1779	82,4	100,0	
Missing	System	380	17,6		
Total		2159	100,0		

\*Hercodering van de variabele van alcohol drinken\*

\*Hercoderen naar (1)Nee en (2)Ja, omdat de variabele te scheef verdeeld is voor een betrouwbare analyse. Zo zorgt een scheve verdeling voor grotere standaardfouten en dus minder zuivere schattingen. Ook is de kans op significante effecten kleiner.\*

```
RECODE drink (1=1) (2 thru 5=2) INTO ooit_alcohol.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van alcohol drinken\*

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

## Frequencies

### Statistics

ooit\_alcohol

N	Valid	1779
	Missing	380
Mean		1,7189
Median		2,0000
Mode		2,00
Std. Deviation		,44964
Minimum		1,00
Maximum		2,00
Percentiles	25	1,0000
	50	2,0000
	75	2,0000

		ooit_alcohol			Cumulative Per-
		Frequency	Percent	Valid Percent	cent
Valid	1,00 No	500	23,2	28,1	28,1
	2,00 Yes	1279	59,2	71,9	100,0
	Total	1779	82,4	100,0	
Missing	System	380	17,6		
Total		2159	100,0		

## Bingedrinken

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van bingedrinken\*

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

## Frequencies

### Statistics

bingedrink How often did you drink FIVE OR MORE drinks (alcoholic) in a row? For example, at a party or at one evening.

N	Valid	1333
	Missing	826
Mean		2,96
Median		2,00
Mode		1
Std. Deviation		2,242
Minimum		1
Maximum		7
Percentiles	25	1,00
	50	2,00
	75	4,00

**bingedrink How often did you drink FIVE OR MORE drinks (alcoholic) in a row? For example, at a party or at one evening.**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Never	580	26,9	43,5	43,5
	2 1 time	160	7,4	12,0	55,5
	3 2 times	117	5,4	8,8	64,3
	4 3 or 4 times	145	6,7	10,9	75,2
	5 5 or 6 times	80	3,7	6,0	81,2
	6 7 or 8 times	38	1,8	2,9	84,0
	7 9 times or more	213	9,9	16,0	100,0
	Total	1333	61,7	100,0	
Missing	System	826	38,3		
Total		2159	100,0		

\*Hercodering van de variabele van bingedrinken\*

\*Als een respondent 'nooit' heeft ingevuld bij alcohol, dan is het aannemelijk dat de respondent ook nooit aan bingedrinken heeft gedaan. Dus als een respondent 'nooit' heeft ingevuld bij alcohol en een missende waarde heeft bij bingedrinken (omdat je de vraag van bingedrinken enkel mocht invullen als je ooit alcohol gedronken hebt) dan wordt de respondent ook bij nooit bingedrinken gezet.\*

```
COMPUTE bingedrinknew = bingedrink.
IF (drink = 1 AND MISSING(bingedrink)) bingedrinknew = 1.
EXECUTE.
```

\*Hercoderen naar (1)Nee en (2)Ja, omdat de variabele te scheef verdeeld is voor een betrouwbare analyse. Zo zorgt een scheve verdeling voor grotere standaardfouten en dus minder zuivere schattingen. Ook is de kans op significante effecten kleiner.\*

```
RECODE bingedrinknew (1=1) (2 thru 7=2) INTO ooit_bingedrink.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van bingedrinken.\*

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

## Frequencies

### Statistics

ooit\_bingedrink

N	Valid	1765
	Missing	394
Mean		1,4266
Median		1,0000
Mode		1,00
Std. Deviation		,49473
Minimum		1,00
Maximum		2,00
Percentiles	25	1,0000
	50	1,0000
	75	2,0000

### ooit\_bingedrink

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	1012	46,9	57,3	57,3
	2,00 Yes	753	34,9	42,7	100,0
	Total	1765	81,8	100,0	
Missing	System	394	18,2		
Total		2159	100,0		

## Tabak roken

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van tabak roken\*

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

## Frequencies

### Statistics

smoke How often did you smoke cigarettes?

N	Valid	1781
	Missing	378
Mean		1,91
Median		1,00
Mode		1
Std. Deviation		1,485
Minimum		1
Maximum		6
Percentiles	25	1,00
	50	1,00
	75	2,00

### smoke How often did you smoke cigarettes?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 I have never smoked, not even a puff	1087	50,3	61,0	61,0
	2 I have smoked 1 or 2 times	348	16,1	19,5	80,6
	3 I smoke 1 or 2 times per month	70	3,2	3,9	84,5
	4 I smoke 1 or 2 times per week	56	2,6	3,1	87,6
	5 I smoke at least once a day	137	6,3	7,7	95,3
	6 I have smoked, but I stopped	83	3,8	4,7	100,0
	Total	1781	82,5	100,0	
Missing	System	378	17,5		
Total		2159	100,0		

\*Hercodering van de variabele van tabak roken\*

\*Hercoderen naar (1)Nee en (2)Ja, omdat de variabele te scheef verdeeld is voor een betrouwbare analyse. Zo zorgt een scheve verdeling voor grotere standaardfouten en dus minder zuivere schattingen. Ook is de kans op significante effecten kleiner.\*

```
RECODE smoke (1=1) (2 thru 6=2)INTO ooit_tabak.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van tabak roken.\*

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

## Frequencities

### Statistics

ooit\_tabak

N	Valid	1781
	Missing	378
Mean		1,3897
Median		1,0000
Mode		1,00
Std. Deviation		,48781
Minimum		1,00
Maximum		2,00
Percentiles	25	1,0000
	50	1,0000
	75	2,0000

### ooit\_tabak

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	1087	50,3	61,0	61,0
	2,00 Yes	694	32,1	39,0	100,0
	Total	1781	82,5	100,0	
Missing	System	378	17,5		
Total		2159	100,0		

## Wiet roken

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van wiet roken\*

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

## Frequencies

### Statistics

weed How often did you smoke cannabis?

N	Valid	1778
	Missing	381
Mean		1,28
Median		1,00
Mode		1
Std. Deviation		,741
Minimum		1
Maximum		5
Percentiles	25	1,00
	50	1,00
	75	1,00

### weed How often did you smoke cannabis?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 I have never smoked cannabis, not even a puff	1487	68,9	83,6	83,6
	2 I have smoked cannabis 1 or 2 times	177	8,2	10,0	93,6
	3 I smoke cannabis 1 or 2 times per month	56	2,6	3,1	96,7
	4 I smoke cannabis 1 or 2 times per week	31	1,4	1,7	98,5
	5 I smoke cannabis at least once a day	27	1,3	1,5	100,0
	Total	1778	82,4	100,0	
Missing	System	381	17,6		
Total		2159	100,0		

\*Hercodering van de variabele van wiet roken\*

\*Hercoderen naar (1)Nee en (2)Ja, omdat de variabele te scheef verdeeld is voor een betrouwbare analyse. Zo zorgt een scheve verdeling voor grotere standaardfouten en dus minder zuivere schattingen. Ook is de kans op significante effecten kleiner.\*

```
RECODE weed (1=1) (2 thru 7=2) INTO ooit_wiet.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van wiet roken.\*

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

## Frequencies

### Statistics

ooit\_wiet

N	Valid	1778
	Missing	381
Mean		1,1637
Median		1,0000
Mode		1,00
Std. Deviation		,37008
Minimum		1,00
Maximum		2,00
Percentiles	25	1,0000
	50	1,0000
	75	1,0000

### ooit\_wiet

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	1487	68,9	83,6	83,6
	2,00 Yes	291	13,5	16,4	100,0
	Total	1778	82,4	100,0	
Missing	System	381	17,6		
Total		2159	100,0		



# De eerste onafhankelijke variabele: seksuele oriëntatie

## Seksuele oriëntatie

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van seksuele oriëntatie\*

```
FREQUENCIES VARIABLES=sex_id  
  /STATISTICS=MODE  
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

sex\_id What do you think you are  
?

N	Valid	1783
	Missing	376
Mode		1

### sex\_id What do you think you are ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 heterosexual	1509	69,9	84,6	84,6
	2 homosexual	14	,6	,8	85,4
	3 bisexual	35	1,6	2,0	87,4
	4 don't know	120	5,6	6,7	94,1
	5 no answer	105	4,9	5,9	100,0
	Total		1783	82,6	100,0
Missing	System	376	17,4		
Total		2159	100,0		

\*Hercodering van de variabele van seksuele oriëntatie\*

\*'Geen antwoord' en 'weet ik niet' coderen als missende waarden. 'Weet ik niet' wordt missing, omdat het onzeker is waarom iemand 'weet ik niet' invult. Als een respondent diens seksuele oriëntatie niet weet, vult diegene 'weet ik niet' in. Het kan ook zo zijn dat iemand 'weet ik niet' invult, omdat ze bijvoorbeeld panseksueel zijn. Dan hebben ze wel een antwoord, maar is het geen antwoordmogelijkheid. Door de onzekerheid over de betekenis die respondenten aan deze optie geven, laat ik hem uit de analyse. 'Geen antwoord' wordt een missende waarde, omdat ik me richt op heteroseksuele en LHB+ jongeren. Als een respondent 'geen antwoord' invult weet ik niet tot welke groep diegene behoort. Ook heb ik homoseksueel en biseksueel samengevoegd tot LHB+, omdat de twee groepen apart te klein zijn om er betrouwbare analyses mee te doen en de kans op een significant resultaat erg klein zal zijn.\*

```
RECODE sex_id (1=1) (2 thru 3= 2) (ELSE=SYSMIS) INTO NEWsex.  
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van seksuele oriëntatie\*

```
FREQUENCIES VARIABLES=NEWsex
  /STATISTICS=MODE
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

NEWsex What do you think you are ?

N	Valid	1558
	Missing	601
Mode		1,00

### NEWsex What do you think you are ?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 Heterosexual	1509	69,9	96,9	96,9
	2,00 LHB+	49	2,3	3,1	100,0
	Total	1558	72,2	100,0	
Missing	System	601	27,8		
Total		2159	100,0		

## De tweede onafhankelijke variabele (mediator): uitgaansgedrag

### Zitten in een caravan doorde weeks

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van zitten in een caravan doorde weeks\*

```
FREQUENCIES VARIABLES=ts_caravan1
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

ts\_caravan1 Hanging out in caravans -

Mon-Thu

N	Valid	1716
	Missing	443
Median		,00
Minimum		0
Maximum		7
Percentiles	25	,00
	50	,00
	75	,00

### ts\_caravan1 Hanging out in caravans - Mon-Thu

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 0-1 hours	1349	62,5	78,6	78,6
	1 1-2 hours	186	8,6	10,8	89,5
	2 2-3 hours	70	3,2	4,1	93,5
	3 3-4 hours	44	2,0	2,6	96,1
	4 4-5 hours	24	1,1	1,4	97,5
	5 5-6 hours	16	,7	,9	98,4
	6 more than 6 hours	24	1,1	1,4	99,8
	7	3	,1	,2	100,0
	Total	1716	79,5	100,0	
Missing	System	443	20,5		
Total		2159	100,0		

\*Herocodering van de variabele van zitten in een caravan doordeeweeks\*

\*Schaal veranderen zodat het net als de rest van de variabelen bij 1 begint. Dit om het interpreteren van de hellingen gemakkelijker te maken\*

```
RECODE ts_caravan1 (0=1) (1=2) (2=3) (3=4) (4=5) (5=6) (6=7) INTO NEW_caravan1.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van zitten in een caravan doordeeweeks\*

```
FREQUENCIES VARIABLES=new_caravan1
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

NEW\_caravan1

N	Valid	1713
	Missing	446
Median		1,0000
Minimum		1,00
Maximum		7,00
Percentiles	25	1,0000
	50	1,0000
	75	1,0000

### NEW\_caravan1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	1349	62,5	78,8	78,8
	2,00	186	8,6	10,9	89,6
	3,00	70	3,2	4,1	93,7
	4,00	44	2,0	2,6	96,3
	5,00	24	1,1	1,4	97,7
	6,00	16	,7	,9	98,6
	7,00	24	1,1	1,4	100,0
	Total	1713	79,3	100,0	
Missing	System	446	20,7		
Total		2159	100,0		

## Zitten in een caravan 's weekends

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van zitten in een caravan 's weekends\*

```
FREQUENCIES VARIABLES=ts_caravan2
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

ts\_caravan2 Hanging out in caravans -

Fri-Sun

N	Valid	1703
	Missing	456
Median		,00
Minimum		0
Maximum		7
Percentiles	25	,00
	50	,00
	75	1,00

### ts\_caravan2 Hanging out in caravans - Fri-Sun

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 0-1 hours	1089	50,4	63,9	63,9
	1 1-2 hours	235	10,9	13,8	77,7
	2 2-3 hours	129	6,0	7,6	85,3
	3 3-4 hours	82	3,8	4,8	90,1
	4 4-5 hours	59	2,7	3,5	93,6
	5 5-6 hours	43	2,0	2,5	96,1
	6 more than 6 hours	64	3,0	3,8	99,9
	7	2	,1	,1	100,0
Total		1703	78,9	100,0	
Missing	System	456	21,1		
Total		2159	100,0		

\*Hercodering van de variabele van zitten in een caravan 's weekends\*

\*Schaal veranderen zodat het net als de rest van de variabelen bij 1 begint. Dit om het interpreteren van de hellingen gemakkelijker te maken \*

```
RECODE ts_caravan2 (0=1) (1=2) (2=3) (3=4) (4=5) (5=6) (6=7) INTO NEW_caravan2.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van zitten in een caravan 's weekends\*

```
FREQUENCIES VARIABLES=new_caravan2
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

NEW_caravan2		
N	Valid	1701
	Missing	458
Median		1,0000
Minimum		1,00
Maximum		7,00
Percentiles	25	1,0000
	50	1,0000
	75	2,0000

NEW_caravan2					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	1089	50,4	64,0	64,0
	2,00	235	10,9	13,8	77,8
	3,00	129	6,0	7,6	85,4
	4,00	82	3,8	4,8	90,2
	5,00	59	2,7	3,5	93,7
	6,00	43	2,0	2,5	96,2
	7,00	64	3,0	3,8	100,0
	Total		1701	78,8	100,0
Missing	System	458	21,2		
Total		2159	100,0		

## Uitgaan in disco's en bars doordeeweeks

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van uitgaan in disco's en bars doordeeweeks\*

```
FREQUENCIES VARIABLES=ts_goingout1
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

ts\_goingout1 Going out (e.g. to the bar, clubbing) - Mon-Thu

N	Valid	1690
	Missing	469
Median		,00
Minimum		0
Maximum		7
Percentiles	25	,00
	50	,00
	75	,00

### ts\_goingout1 Going out (e.g. to the bar, clubbing) - Mon-Thu

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 0-1 hours	1309	60,6	77,5	77,5
	1 1-2 hours	159	7,4	9,4	86,9
	2 2-3 hours	72	3,3	4,3	91,1
	3 3-4 hours	59	2,7	3,5	94,6
	4 4-5 hours	38	1,8	2,2	96,9
	5 5-6 hours	19	,9	1,1	98,0
	6 more than 6 hours	32	1,5	1,9	99,9
	7	2	,1	,1	100,0
	Total	1690	78,3	100,0	
Missing	System	469	21,7		
Total		2159	100,0		

\*Herocodering van de variabele van uitgaan in disco's en bars doordeeweeks\*

\*Schaal veranderen zodat het net als de rest van de variabelen bij 1 begint\*

```
RECODE ts_goingout1 (0=1) (1=2) (2=3) (3=4) (4=5) (5=6) (6=7) INTO NEW_go-
ingout1.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van uitgaan in disco's en bars doordeeweeks\*

```
FREQUENCIES VARIABLES=NEW_goingout1
/NTILES=4
/STATISTICS=MINIMUM MAXIMUM MEDIAN
/ORDER=ANALYSIS.
```

## Frequencies

### Statistics

NEW\_goingout1

N	Valid	1688
	Missing	471
Median		1,0000
Minimum		1,00
Maximum		7,00
Percentiles	25	1,0000
	50	1,0000
	75	1,0000

### NEW\_goingout1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00	1309	60,6	77,5	77,5
	2,00	159	7,4	9,4	87,0
	3,00	72	3,3	4,3	91,2
	4,00	59	2,7	3,5	94,7
	5,00	38	1,8	2,3	97,0
	6,00	19	,9	1,1	98,1
	7,00	32	1,5	1,9	100,0
	Total		1688	78,2	100,0
Missing	System	471	21,8		
Total		2159	100,0		

## Uitgaan in disco's en bars 's weekends

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van uitgaan in disco's en bars 's weekends\*

```
FREQUENCIES VARIABLES=ts_goingout2
/NTILES=4
/STATISTICS=MINIMUM MAXIMUM MEDIAN
/ORDER=ANALYSIS.
```



## Frequencies

### Statistics

ts\_goingout2 Going out (e.g. to the bar, clubbing) - Fri-Sun

N	Valid	1692
	Missing	467
Median		1,00
Minimum		0
Maximum		7
Percentiles	25	,00
	50	1,00
	75	3,00

### ts\_goingout2 Going out (e.g. to the bar, clubbing) - Fri-Sun

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 0-1 hours	844	39,1	49,9	49,9
	1 1-2 hours	220	10,2	13,0	62,9
	2 2-3 hours	143	6,6	8,5	71,3
	3 3-4 hours	165	7,6	9,8	81,1
	4 4-5 hours	140	6,5	8,3	89,4
	5 5-6 hours	65	3,0	3,8	93,2
	6 more than 6 hours	113	5,2	6,7	99,9
	7	2	,1	,1	100,0
	Total	1692	78,4	100,0	
Missing	System	467	21,6		
	Total	2159	100,0		

\*Hercodering van de variabele van uitgaan in disco's en bars 's weekends

\*Schaal veranderen zodat het net als de rest van de variabelen bij 1 begint. Dit om het interpreteren van de hellingen gemakkelijker te maken.\*

```
RECODE ts_goingout2 (0=1) (1=2) (2=3) (3=4) (4=5) (5=6) (6=7) INTO NEW_goingout2.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel voor de nieuwe variabele van uitgaan in disco's en bars 's weekends\*

```
FREQUENCIES VARIABLES=NEW_goingout2
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

NEW_goingout2		
N	Valid	1690
	Missing	469
Median		2,0000
Minimum		1,00
Maximum		7,00
Percentiles	25	1,0000
	50	2,0000
	75	4,0000

		NEW_goingout2			Cumulative Per-
		Frequency	Percent	Valid Percent	cent
Valid	1,00	844	39,1	49,9	49,9
	2,00	220	10,2	13,0	63,0
	3,00	143	6,6	8,5	71,4
	4,00	165	7,6	9,8	81,2
	5,00	140	6,5	8,3	89,5
	6,00	65	3,0	3,8	93,3
	7,00	113	5,2	6,7	100,0
	Total		1690	78,3	100,0
Missing	System	469	21,7		
Total		2159	100,0		

## Betrouwbaarheidsanalyse van de nieuwe items

\*Cronbachs alpha berekenen. Deze waarde geeft een indicatie van hoe goed de verschillende variabelen bij elkaar passen en hetzelfde lijken te meten. De Chronbachs alpha is in dit geval 0,80, dus het lijkt dat de vier variabelen van uitgaansgedrag goed bij elkaar passen\*

```
RELIABILITY
/VARIABLES=NEW_caravan1 NEW_goingout1 NEW_caravan2 NEW_goingout2
/SCALE('ALL VARIABLES') ALL
/MODEL=ALPHA.
```

### Case Processing Summary

		N	%
Cases	Valid	1642	76,1
	Excluded <sup>a</sup>	517	23,9
	Total	2159	100,0

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

### Reliability Statistics

Cronbach's	
Alpha	N of Items
,802	4

### Samenvoegen van de items tot de schaalvariabele en de beschrijvende statistieken voor de nieuwe variabele

\*Creeëren van de schaal van uitgaansgedrag uitgaan door de items bij elkaar op te tellen en te delen door het totaal aantal variabelen, oftewel het gemiddelde van de itemscores wordt berekend\*

```
COMPUTE Schaal_uitgaan=(NEW_caravan1 + NEW_goingout1 + NEW_caravan2 +  
NEW_goingout2) / 4.  
EXECUTE.
```

\*Beschrijvende statistieken en histogram van de schaalvariabele van uitgaansgedrag\*

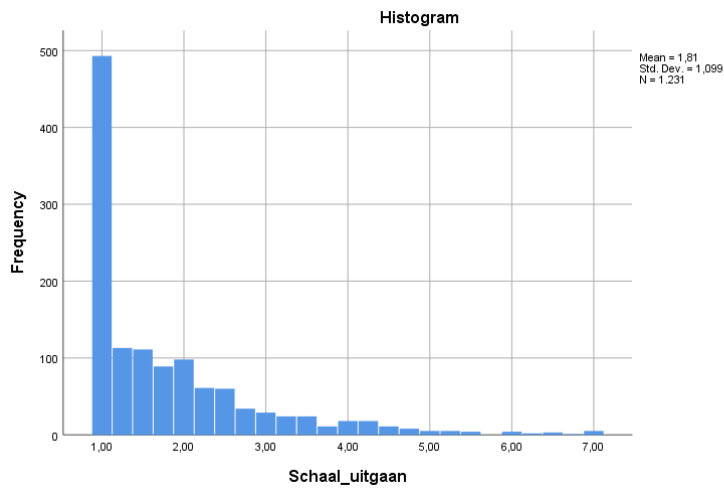
```
EXAMINE VARIABLES=Schaal_uitgaan  
/PLOT HISTOGRAM  
/STATISTICS DESCRIPTIVES  
/CINTERVAL 95  
/MISSING LISTWISE  
/NOTOTAL.
```

## Explore

### Descriptives

		Statistic	Std. Error
Schaal_uitgaan	Mean	1,7922	,03197
	95% Confidence Interval for Mean	Lower Bound	1,7294
		Upper Bound	1,8549
	5% Trimmed Mean	1,6495	
	Median	1,2500	
	Variance	1,200	
	Std. Deviation	1,09528	
	Minimum	1,00	
	Maximum	7,00	
	Range	6,00	
	Interquartile Range	1,25	
	Skewness	1,980	,071
	Kurtosis	4,415	,143

## Schaal\_uitgaan



## Controlevariabelen

### Leeftijd

\*Beschrijvende statistieken en histogram van de variabele van leeftijd  
leeftijd\*

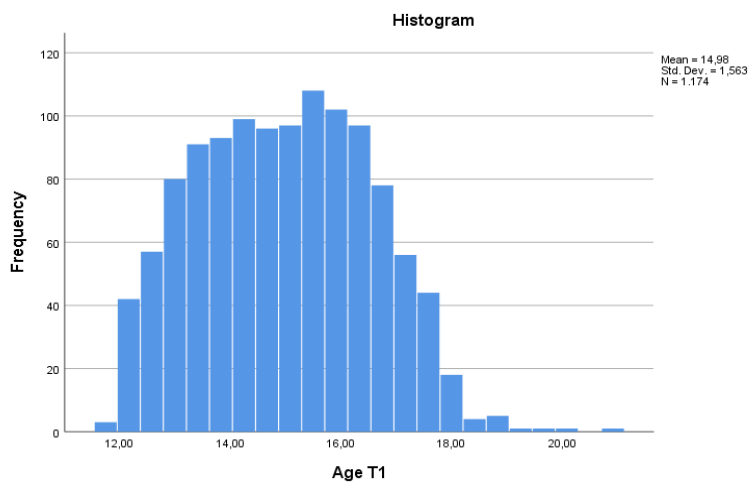
```
EXAMINE VARIABLES=age_w1  
  /PLOT HISTOGRAM  
  /STATISTICS DESCRIPTIVES  
  /CINTERVAL 95  
  /MISSING LISTWISE  
  /NOTOTAL.
```

### Explore

#### Descriptives

		Statistic	Std. Error	
age_w1 What is your age?	Mean	14,9831	,04561	
	95% Confidence Interval for Mean	Lower Bound	14,8936	
		Upper Bound	15,0726	
	5% Trimmed Mean	14,9654		
	Median	15,0000		
	Variance	2,442		
	Std. Deviation	1,56272		
	Minimum	11,75		
	Maximum	20,75		
	Range	9,00		
	Interquartile Range	2,42		
	Skewness	,129	,071	
	Kurtosis	-,617	,143	

## age\_w1 What is your age?



## Geslacht

\*Beschrijvende statistieken en frequentietabel van de variabele van geslacht\*

```
FREQUENCIES VARIABLES=gender_final
  /NTILES=4
  /STATISTICS=MINIMUM MAXIMUM MEDIAN
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

gender\_final What is your gender?

N	Valid	2147
	Missing	12
Mode		1

### gender\_final What is your gender?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	1076	49,8	50,1	50,1
	2 Female	1071	49,6	49,9	100,0
	Total	2147	99,4	100,0	
Missing	System	12	,6		
Total		2159	100,0		

## Migratieachtergrond van de respondent

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van de migratieachtergrond van de respondent\*

```
FREQUENCIES VARIABLES=eth
  /STATISTICS=MODE
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

eth Where were you born?

N	Valid	1826
	Missing	333
Mode		1

### eth Where were you born?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Netherlands	1757	81,4	96,2	96,2
	2 Turkey	6	,3	,3	96,5
	3 Marocco	1	,0	,1	96,6
	4 Surinam	5	,2	,3	96,9
	5 Dutch Antilles, Aruba, Curacao, Saint-Martin	15	,7	,8	97,7
	6 Indonesia, Maluku Islands	1	,0	,1	97,8
	7 Other, namely	41	1,9	2,2	100,0
	Total	1826	84,6	100,0	
Missing	System	333	15,4		
Total		2159	100,0		

\*Hercodering van de variabele van de migratieachtergrond van de respondent\*

\*Een dummy maken van eth. 1= geen migratieachtergrond, 2= niet-westerse migratieachtergrond, 'anders' = missing, omdat 'anders' zowel westerse als niet-westerse landen bevatte en het in de gegeven tijd niet mogelijk was om dit handmatig te categoriseren en er soms onserieuze antwoorden bijstonden, zoals ziekenhuis\*

```
RECODE eth (1=1) (7=SYSMIS) (2 thru 6=2) INTO migratie_respondent.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van de oorspronkelijke variabele van de migratieachtergrond van de respondent\*

```
FREQUENCIES VARIABLES=migratie_respondent
  /STATISTICS=MODE
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

migratie\_respondent

N	Valid	1785
	Missing	374
Mode		1,00

### migratie\_respondent

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No migration background	1757	81,4	98,4	98,4
	2,00 Non-western migration background	28	1,3	1,6	100,0
	Total	1785	82,7	100,0	
Missing	System	374	17,3		
Total		2159	100,0		

## Migratieachtergrond van de vader van de respondent

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van de migratieachtergrond van de vader van de respondent\*

```
FREQUENCIES VARIABLES=eth_f
  /STATISTICS=MODE
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

eth\_f Where was your father born?

N	Valid	1816
	Missing	343
Mode		1



### eth\_f Where was your father born?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Netherlands	1578	73,1	86,9	86,9
	2 Turkey	51	2,4	2,8	89,7
	3 Marocco	6	,3	,3	90,0
	4 Surinam	27	1,3	1,5	91,5
	5 Dutch Antilles, Aruba, Curacao, Saint-Martin	30	1,4	1,7	93,2
	6 Indonesia, Maluku Islands	14	,6	,8	93,9
	7 Other, namely	110	5,1	6,1	100,0
	Total	1816	84,1	100,0	
Missing	System	343	15,9		
Total		2159	100,0		

\*Hercodering van de variabele van de migratieachtergrond van de vader\*

\*Een dummy maken van eth\_f. 1= geen migratieachtergrond, 2= niet-westerse migratieachtergrond, anders = missing, omdat 'anders' zowel westerse als niet-westerse landen bevatte en het in de gegeven tijd niet mogelijk was om dit handmatig te categoriseren en er soms onserieuze antwoorden bijstonden, zoals ziekenhuis \*

```
RECODE eth_f (1=1) (7=SYSMIS) (2 thru 6=2) INTO migratie_vader.
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele de migratieachtergrond van de vader van de respondent\*

```
FREQUENCIES VARIABLES=migratie_vader
/STATISTICS=MODE
/ORDER=ANALYSIS.
```

## Frequencies

### Statistics

migratie\_vader

N	Valid	1706
	Missing	453
Mode		1,00

		migratie_vader			Cumulative Per-
		Frequency	Percent	Valid Percent	cent
Valid	1,00 No migration background	1578	73,1	92,5	92,5
	2,00 Non-western background	128	5,9	7,5	100,0
	Total	1706	79,0	100,0	
Missing	System	453	21,0		
Total		2159	100,0		

## Migratieachtergrond van de moeder van de respondent

\*Beschrijvende statistieken en frequentietabel van de oorspronkelijke variabele van de migratieachtergrond van de moeder van de respondent\*

```
FREQUENCIES VARIABLES=eth_m
  /STATISTICS=MODE
  /ORDER=ANALYSIS.
```

## Frequencies

### Statistics

eth\_m Where was your mother born?

N	Valid	1823
	Missing	336
Mode		1

		eth_m Where was your mother born?			Cumulative Per-
		Frequency	Percent	Valid Percent	cent
Valid	1 Netherlands	1615	74,8	88,6	88,6
	2 Turkey	48	2,2	2,6	91,2
	3 Marocco	7	,3	,4	91,6
	4 Surinam	17	,8	,9	92,5
	5 Dutch Antilles, Aruba, Curacao, Saint-Martin	25	1,2	1,4	93,9
	6 Indonesia, Maluku Islands	11	,5	,6	94,5
	7 Other, namely	100	4,6	5,5	100,0
	Total	1823	84,4	100,0	
Missing	System	336	15,6		
Total		2159	100,0		

\*Herocodering van de variabele van de migratieachtergrond van de moeder\*

\*Een dummy maken van eth\_m. 1= geen migratieachtergrond, 2= niet-westerse migratieachtergrond, anders = missing, omdat 'anders' zowel westerse als niet-westerse landen bevatte en het in de gegeven tijd niet mogelijk was om dit handmatig te categoriseren en er soms onserieuze antwoorden bijstonden, zoals ziekenhuis \*

```
RECODE eth_m (1=1) (7=SYSMIS) (2 thru 6=2) INTO migratie_moeder
EXECUTE.
```

\*Beschrijvende statistieken en frequentietabel van de nieuwe variabele van de migratieachtergrond van de moeder van de respondent \*

```
FREQUENCIES VARIABLES=migratie_moeder
/STATISTICS=MODE
/ORDER=ANALYSIS.
```

## Frequencies

### Statistics

migratie\_moeder

N	Valid	1723
	Missing	436
Mode		1,00

### migratie\_moeder

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No migration back-ground	1615	74,8	93,7	93,7
	2,00 Non-western migration background	108	5,0	6,3	100,0
	Total	1723	79,8	100,0	
Missing	System	436	20,2		
Total		2159	100,0		

## Missende waarden

Als laatste heb ik alle observaties die niet op alle benodigde variabelen een antwoord hebben gegeven, uit de dataset gehaald. Van de 2159 respondenten bleven er nog 1174 over. Hieronder staat de syntax.

Eerst zijn er lineaire regressie analyses uitgevoerd, waarbij de residuen worden opgeslagen. Daarna worden er variabelen aangemaakt waarbij 0 betekent dat er geen residu uitgerekend kon worden en 1 dat er wel een residu uitgerekend kon worden. Als er geen residu uitgerekend kon worden, komt dat doordat een observatie missende waarde(n) heeft, oftewel een respondent heeft een of meerdere vragen niet ingevuld. Tot slot worden alle variabelen die een 0 hebben op de gemaakte variabele uit de dataset gefilterd, zodat er alleen observaties overblijven die op alle benodigde variabelen een antwoord hebben gegeven.

\*Regressies voor het opslaan van de residuen

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT ooit_tabak
  /METHOD=ENTER Schaal_uitgaan NEWsex migratie_respondent migratie_vader
migratie_moeder gender_final age_w1
  /SAVE RESID
```

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT ooit_alcohol
  /METHOD=ENTER Schaal_uitgaan NEWsex migratie_respondent migratie_vader
migratie_moeder gender_final age_w1
  /SAVE RESID
```

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT ooit_bingedrink
  /METHOD=ENTER Schaal_uitgaan NEWsex migratie_respondent migratie_vader
migratie_moeder gender_final age_w1
  /SAVE RESID
```

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT ooit_wiet
  /METHOD= Schaal_uitgaan NEWsex migratie_respondent migratie_vader migra-
tie_moeder gender_final age_w1
  /SAVE RESID
```

\*Dummy maken van de residuen. Als er een residu berekend kon worden, betekent dit dat er geen missende waarden zijn, oftewel de respondent heeft alle benodigde vragen ingevuld. Zodra er een missende is voor het residu, heeft de respondent een of meerdere vragen niet ingevuld. Als er een of meerdere missende waarde zijn, krijgt een observatie de waarde 0. Als er geen missende waarden zijn, krijgt de observatie de waarde 1.\*

```
RECODE RES_1 (SYSMIS=0) (ELSE=1) INTO obs1.  
EXECUTE.
```

```
RECODE RES_2 (SYSMIS=0) (ELSE=1) INTO obs2.  
EXECUTE.
```

```
RECODE RES_3 (SYSMIS=0) (ELSE=1) INTO obs3.  
EXECUTE.
```

```
RECODE RES_4 (SYSMIS=0) (ELSE=1) INTO obs4.  
EXECUTE.
```

\*Observaties met ontbrekende waarden eruit filteren\*

```
USE ALL.  
COMPUTE filter_$=(obs1 = 1 | obs2 = 1 | obs3 = 1 | obs4 = 1).  
VARIABLE LABELS filter_$ 'obs1 = 1 | obs2 = 1 | obs3 = 1 | obs4 = 1 (FILTER)'.  
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.  
FORMATS filter_$ (f1.0).  
FILTER BY filter_$.  
EXECUTE.
```

## Bijlage 2 Statistische analyses

### Univariate analyses

De eerste analyse is van de univariate statistieken. Deze geven een beeld van de verdeling van de variabelen in de analyse aan de hand van gemiddelden, standaarddeviaties, modus, het minimum en maximum, de mediaan en het eerste en derde kwartiel. Voor de continue variabelen zijn histogrammen gegeven en voor de categorische variabelen staafdiagrammen. Ook zijn er voor alle variabelen frequentietabellen gegeven. Hieronder staat de syntax en de bijbehorende output van de analyses. De uitkomsten van deze analyses bespreek ik in de Resultatenparagraaf bij de univariate statistieken.

### Output univariate analyses

\*Beschrijvende statistieken, staafdiagrammen en frequentietabellen van de categorische variabelen\*

```
FREQUENCIES VARIABLES= ooit_alcohol ooit_bingedrink ooit_tabak ooit_wiet
migratie_respondent migratie_vader migratie_moeder NEWsex gender_final
  /NTILES=4
  /BARCHART FREQ
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MODE
  /ORDER=ANALYSIS.
```

### Frequencies

		Statistics								
		ooit_al- cohol	ooit_bin- gedrink	ooit_ta- bak	ooit_wiet	migra- tie_res- pondent	migra- tie_vader	migra- tie_moe- der	NEWsex What do you think you are?	gen- der_final What is your gen- der?
N	Valid	1231	1231	1231	1231	1231	1231	1231	1231	1231
	Mis- sing	0	0	0	0	0	0	0	0	0
Mean		1,7652	1,4582	1,3964	1,1706	1,0122	1,0617	1,0504	1,0349	1,52
Mode		2,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	2
Std. Devia- tion		,42403	,49845	,48935	,37631	,10976	,24078	,21879	,18368	,500
Minimum		1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1
Maximum		2,00	2,00	2,00	2,00	2,00	2,00	2,00	2,00	2
Percen- tiles	25	2,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,00
	50	2,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	2,00
	75	2,0000	2,0000	2,0000	1,0000	1,0000	1,0000	1,0000	1,0000	2,00

## Frequency Table

		oait_alcohol			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	289	23,5	23,5	23,5
	2,00 Yes	942	76,5	76,5	100,0
	Total	1231	100,0	100,0	

		oait_bingedrink			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	667	54,2	54,2	54,2
	2,00 Yes	564	45,8	45,8	100,0
	Total	1231	100,0	100,0	

		oait_tabak			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	743	60,4	60,4	60,4
	2,00 Yes	488	39,6	39,6	100,0
	Total	1231	100,0	100,0	

		oait_wiet			
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No	1021	82,9	82,9	82,9
	2,00 Yes	210	17,1	17,1	100,0
	Total	1231	100,0	100,0	

**migratie\_respondent**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No migration background	1216	98,8	98,8	98,8
	2,00 Non-western migration background	15	1,2	1,2	100,0
	Total	1231	100,0	100,0	

**migratie\_vader**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No migration background	1155	93,8	93,8	93,8
	2,00 Non-western migration background	76	6,2	6,2	100,0
	Total	1231	100,0	100,0	

**migratie\_moeder**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 No migration background	1169	95,0	95,0	95,0
	2,00 Non-western migration background	62	5,0	5,0	100,0
	Total	1231	100,0	100,0	

**NEWsex What do you think you are?**

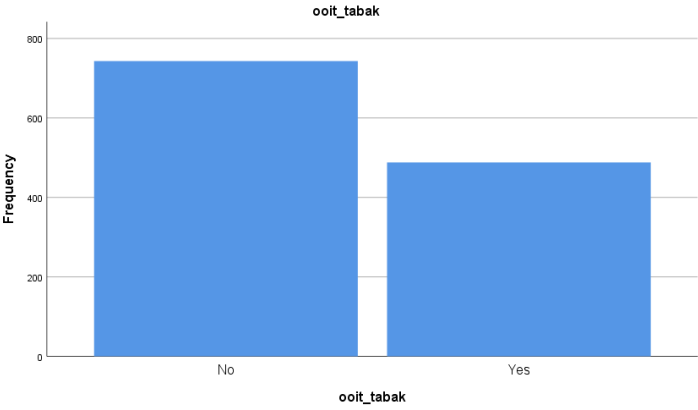
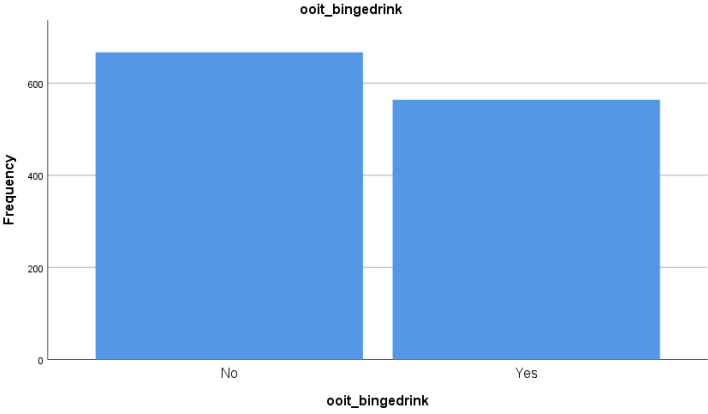
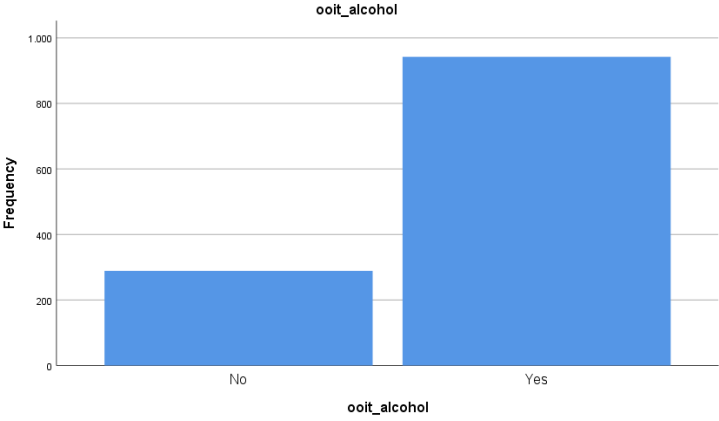
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1,00 Heterosexual	1188	96,5	96,5	96,5
	2,00 LHB+	43	3,5	3,5	100,0
	Total	1231	100,0	100,0	

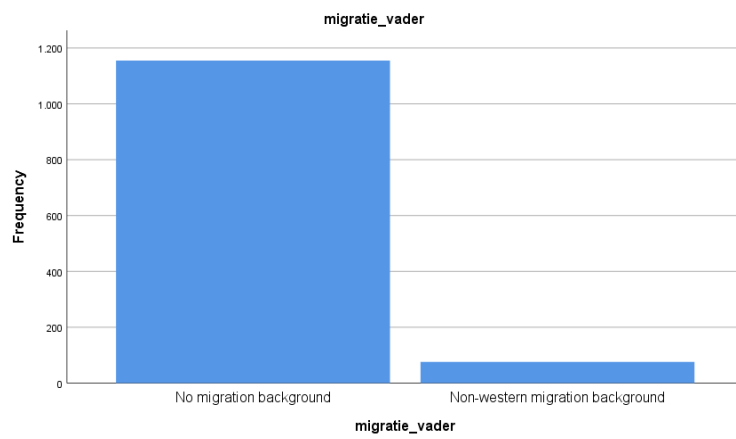
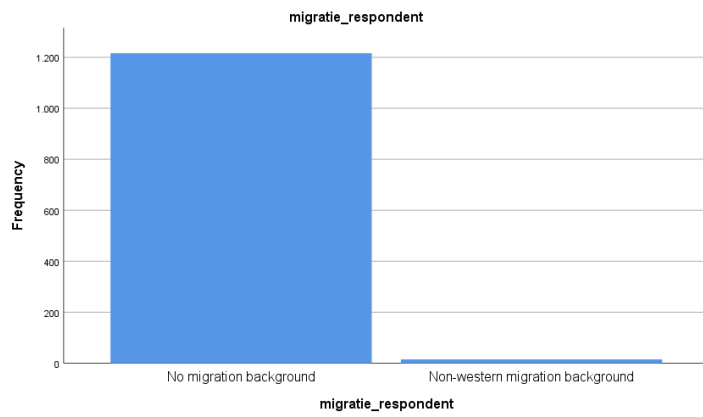
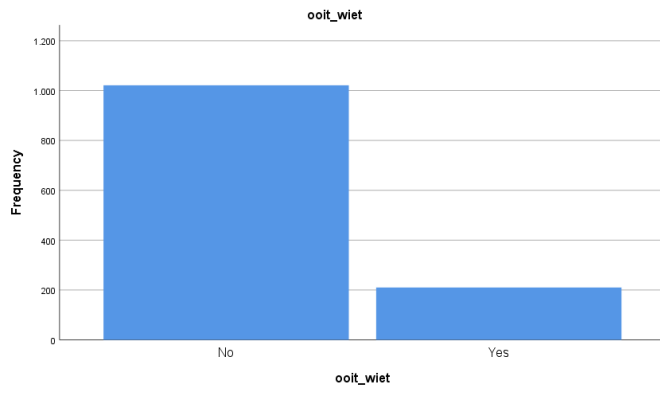
**gender\_final What is your gender?**

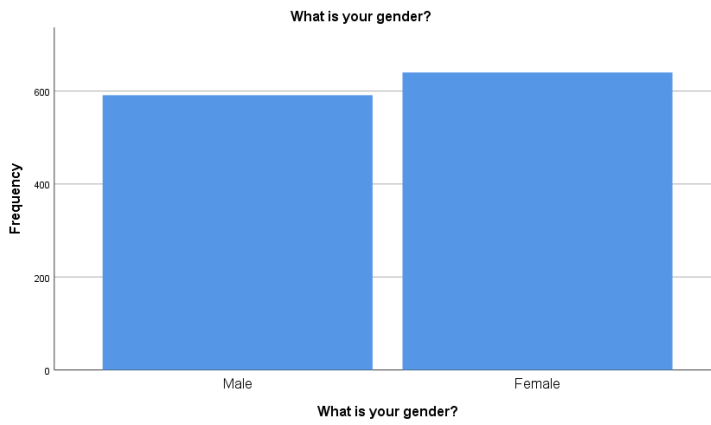
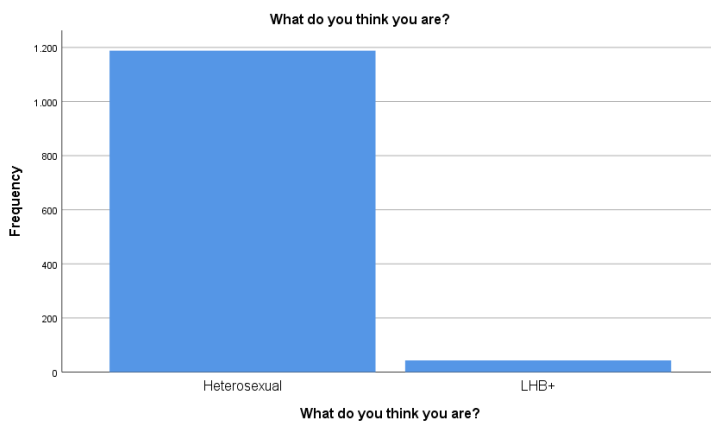
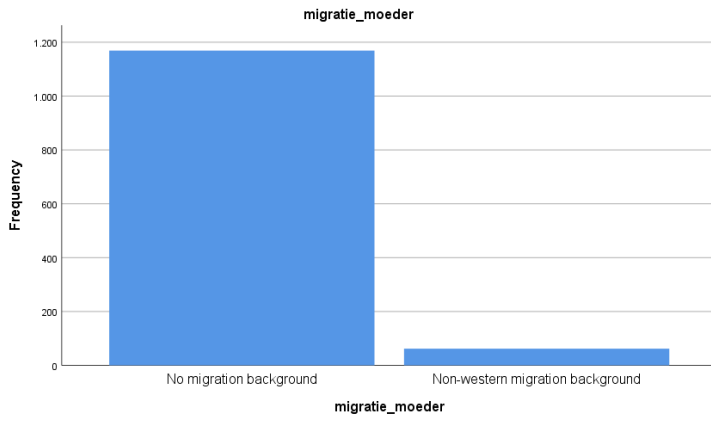
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	591	48,0	48,0	48,0
	2 Female	640	52,0	52,0	100,0
	Total	1231	100,0	100,0	



# Bar Chart







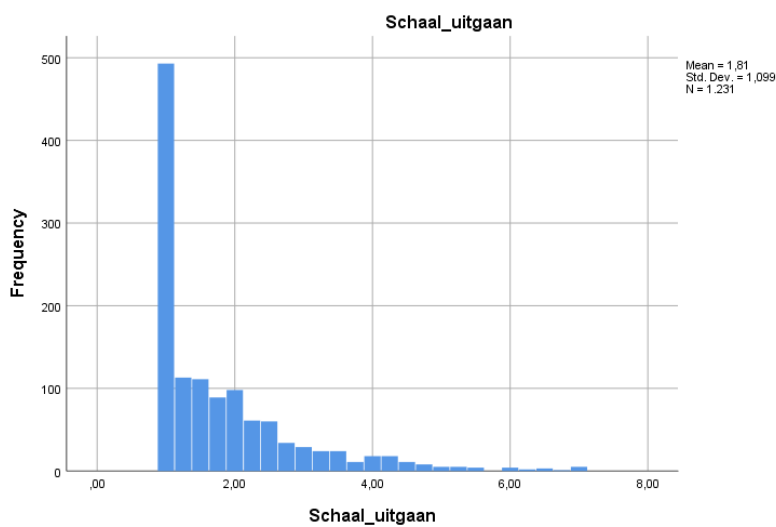
\*Beschrijvende statistieken, histogrammen en frequentietabellen van de continue variabelen\*

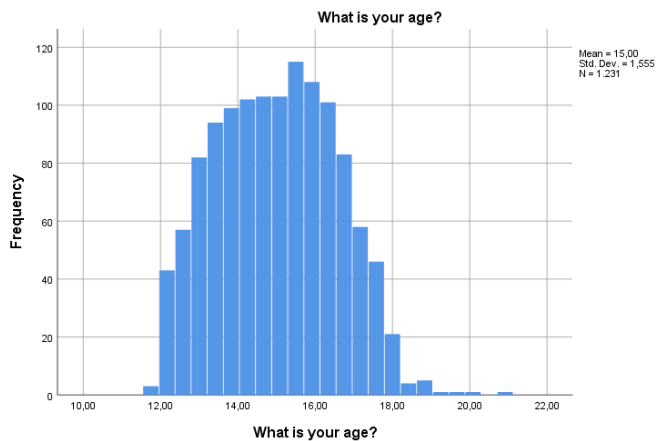
```
FREQUENCIES VARIABLES= Schaal_uitgaan age_w1
  /NTILES=4
  /HISTOGRAM
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MODE
  /ORDER=ANALYSIS.
```

## Frequencies

		Statistics	
		Schaal_uitgaan	age_w1 What is your age?
N	Valid	1231	1231
	Missing	0	0
Mean		1,8119	14,9998
Mode		1,00	15,50
Std. Deviation		1,09918	1,55472
Minimum		1,00	11,75
Maximum		7,00	20,75
Percentiles	25	1,0000	13,7500
	50	1,5000	15,0000
	75	2,2500	16,1667

## Histogram





## Bivariate analyses

De tweede analyse die ik heb uitgevoerd betreft de bivariate statistieken. Deze statistieken geven de samenhang tussen de variabelen weer. Tussen de continue variabelen zijn Pearson's correlaties uitgerekend. Voor continue en categorische variabelen zijn eenweg-ANOVA analyses uitgevoerd en aan de hand van de  $R^2$  is de correlatie berekend, door de wortel te nemen van de  $R^2$ . Ook zijn er t-toetsen uitgevoerd, om de verschillen tussen de groepen te toetsen. Tot slot zijn er kruistabellen gemaakt om de samenhang tussen de categorische variabelen te berekenen. Cramer's V is hierbij als associatiemaat gebruikt. Hieronder de syntax en de output van de uitgevoerde analyses. Een uitgebreide bespreking van de opvallendste en interessantste samenhangen is te vinden in de Resultatenparagraaf.

## Output bivariate analyses

\*Pearsons correlaties berekenen tussen de continue variabelen, oftewel leeftijd en uitgaansgedrag\*

```
CORRELATIONS
/VARIABLES=age_w1 Schaal_uitgaan
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

## Correlations

### Correlations

		age_w1 What is your age?	Schaal_uitgaan
age_w1 What is your age?	Pearson Correlation	1	,255**
	Sig. (2-tailed)		,000
	N	1231	1231
Schaal_uitgaan	Pearson Correlation	,255**	1
	Sig. (2-tailed)	,000	
	N	1231	1231

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

\*Eenweg-anova tussen de continue variabelen en categorische variabelen\*  
 \*Leeftijd en geslacht\*

```
UNIANOVA age_w1 BY gender_final
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(gender_final)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=gender_final
```

## Univariate Analysis of Variance

### Between-Subjects Factors

	Value Label	N
gender_final What is your gender?	1 Male	591
	2 Female	640

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,561 <sup>a</sup>	1	,561	,232	,630
Intercept	276497,304	1	276497,304	114318,538	,000
gender_final	,561	1	,561	,232	,630
Error	2972,529	1229	2,419		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,000 (Adjusted R Squared = -,001)

## Estimated Marginal Means

### What is your gender?

Dependent Variable: age\_w1 What is your age?

What is your gender?	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1 Male	14,978	,064	14,852	15,103
2 Female	15,020	,061	14,900	15,141

\*Leeftijd en migratieachtergrond van de respondent\*

```
UNIANOVA age_w1 BY migratie_respondent
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(migratie_respondent)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWETH
```

### Between-Subjects Factors

		Value Label	N
migratie_respondent	1,00	No migration-back- ground	1159
	2,00	Non-western mi- grationback- ground	15

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	7,809 <sup>a</sup>	1	7,809	3,204	,074
Intercept	13932,933	1	13932,933	5716,030	,000
migratie_respondent	7,809	1	7,809	3,204	,074
Error	2856,773	1172	2,438		
Total	266419,917	1174			
Corrected Total	2864,582	1173			

a. R Squared = ,003 (Adjusted R Squared = ,002)

### Estimated Marginal Means

#### migratie\_respondent

Dependent Variable: age\_w1 What is your age?

migratie_respondent	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration back- ground	14,991	,045	14,904	15,079
2,00 Non-western migration background	15,700	,401	14,913	16,487

\*Leeftijd en migratieachtergrond van de vader van de respondent\*

```
UNIANOVA age_w1 BY migratie_vader
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(migratie_vader)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWETH_F
```

## Univariate Analysis of Variance

### Between-Subjects Factors

	Value	Label	N
migratie_vader	1,00	No migration background	1155
	2,00	Non-western migration background	76

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,701 <sup>a</sup>	1	,701	,290	,591
Intercept	63804,187	1	63804,187	26381,247	,000
migratie_vader	,701	1	,701	,290	,591
Error	2972,390	1229	2,419		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,000 (Adjusted R Squared = -,001)



## Estimated Marginal Means

### migratie\_vader

Dependent Variable: age\_w1 What is your age?

migratie_vader	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration back-ground	15,006	,046	14,916	15,096
2,00 Non-western migration background	14,907	,178	14,557	15,257

\*Leeftijd en migratieachtergrond van de moeder van de respondent\*

```
UNIANOVA age_w1 BY migratie_moeder
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(migratie_moeder)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWETH_M
```

## Univariate Analysis of Variance

### Between-Subjects Factors

		Value Label	N
migratie_moeder	1,00	No migration background	1169
	2,00	Non-western migration background	62

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2,839 <sup>a</sup>	1	2,839	1,175	,279
Intercept	53688,066	1	53688,066	22214,498	,000
migratie_moeder	2,839	1	2,839	1,175	,279
Error	2970,251	1229	2,417		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,001 (Adjusted R Squared = ,000)

### Estimated Marginal Means

#### migratie\_moeder

Dependent Variable: age\_w1 What is your age?

migratie_moeder	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration background	14,989	,045	14,900	15,078
2,00 Non-western migration background	15,208	,197	14,821	15,596

\*Leeftijd en seksuele oriëntatie\*

```
UNIANOVA age_w1 BY NEWsex
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(NEWsex)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWsex
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
NEWsex What do you think you are?	1,00 Heterosexual	1188
	2,00 LHB+	43

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	22,307 <sup>a</sup>	1	22,307	9,291	,002
Intercept	39064,427	1	39064,427	16270,319	,000
NEWsex	22,307	1	22,307	9,291	,002
Error	2950,783	1229	2,401		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,008 (Adjusted R Squared = ,007)

### Estimated Marginal Means

#### What do you think you are?

Dependent Variable: age\_w1 What is your age?

What do you think you are?	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 Heterosexual	14,974	,045	14,886	15,062
2,00 LHB+	15,707	,236	15,244	16,171

\*Leeftijd en alcohol drinken\*

```
UNIANOVA age_w1 BY ooit_alcohol
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_alcohol)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=drinknew2
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
ooit_alcohol	1,00 No	289
	2,00 Yes	942

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	269,079 <sup>a</sup>	1	269,079	122,299	,000
Intercept	191343,039	1	191343,039	86967,319	,000
ooit_alcohol	269,079	1	269,079	122,299	,000
Error	2704,011	1229	2,200		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,091 (Adjusted R Squared = ,090)

### Estimated Marginal Means

#### ooit\_alcohol

Dependent Variable: age\_w1 What is your age?

ooit_alcohol	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	14,156	,087	13,985	14,327
2,00 Yes	15,259	,048	15,164	15,354

\*Leeftijd en bingedrinken

```
UNIANOVA age_w1 BY ooit_bingedrink*
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_bingedrink)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=bingedrinknew2
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
ooit_bingedrink	1,00 No	667
	2,00 Yes	564

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	462,291 <sup>a</sup>	1	462,291	226,285	,000
Intercept	276918,625	1	276918,625	135547,695	,000
ooit_bingedrink	462,291	1	462,291	226,285	,000
Error	2510,799	1229	2,043		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,155 (Adjusted R Squared = ,155)

### Estimated Marginal Means

#### ooit\_bingedrink

Dependent Variable: age\_w1 What is your age?

ooit_bingedrink	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	14,436	,055	14,328	14,545
2,00 Yes	15,666	,060	15,548	15,784

\*Leeftijd en tabak roken\*

```
UNIANOVA age_w1 BY ooit_tabak
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_tabak)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=smokenew2
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
ooit_tabak	1,00 No	743
	2,00 Yes	488

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	184,472 <sup>a</sup>	1	184,472	81,301	,000
Intercept	267987,714	1	267987,714	118107,565	,000
ooit_tabak	184,472	1	184,472	81,301	,000
Error	2788,618	1229	2,269		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,062 (Adjusted R Squared = ,061)

### Estimated Marginal Means

#### ooit\_tabak

Dependent Variable: age\_w1 What is your age?

ooit_tabak	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	14,686	,055	14,578	14,794
2,00 Yes	15,477	,068	15,344	15,611

\*Leeftijd en wiet roken\*

```
UNIANOVA age_w1 BY ooit_wiet
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_wiet)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=weednew2
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
ooit_wiet	1,00 No	1021
	2,00 Yes	210

### Tests of Between-Subjects Effects

Dependent Variable: age\_w1 What is your age?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	228,986 <sup>a</sup>	1	228,986	102,556	,000
Intercept	164747,234	1	164747,234	73785,236	,000
ooit_wiet	228,986	1	228,986	102,556	,000
Error	2744,104	1229	2,233		
Total	279940,590	1231			
Corrected Total	2973,090	1230			

a. R Squared = ,077 (Adjusted R Squared = ,076)

### Estimated Marginal Means

#### ooit\_wiet

Dependent Variable: age\_w1 What is your age?

ooit_wiet	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	14,804	,047	14,712	14,896
2,00 Yes	15,951	,103	15,748	16,153

\*Uitgaansgedrag en geslacht\*

```
UNIANOVA Schaal_uitgaan BY gender_final
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(gender_final)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=gender_final
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
gender_final What is your	1 Male	591
gender?	2 Female	640

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan On average, how much time do you spend on going out?

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4,310 <sup>a</sup>	1	4,310	3,575	,059
Intercept	4045,638	1	4045,638	3355,486	,000
gender_final	4,310	1	4,310	3,575	,059
Error	1481,779	1229	1,206		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,003 (Adjusted R Squared = ,002)

### Estimated Marginal Means

#### What is your gender?

Dependent Variable: Schaal\_uitgaan

What is your gender?	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1 Male	1,874	,045	1,785	1,962
2 Female	1,755	,043	1,670	1,840

\*Uitgaansgedrag en migratieachtergrond van de respondent\*

```
UNIANOVA Schaal_uitgaan BY migratie_respondent
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(migratie_respondent)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWETH
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

Value Label	N
-------------	---



migratie_respondent	1,00	No migration background	1216
	2,00	Non-western migration background	15

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4,128 <sup>a</sup>	1	4,128	3,423	,065
Intercept	253,819	1	253,819	210,494	,000
migratie_respondent	4,128	1	4,128	3,423	,065
Error	1481,961	1229	1,206		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,003 (Adjusted R Squared = ,002)

### Estimated Marginal Means

#### migratie\_respondent

Dependent Variable: Schaal\_uitgaan

migratie_respondent	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration background	1,806	,031	1,744	1,867
2,00 Non-western migration background	2,333	,284	1,777	2,890

\*Uitgaansgedrag en migratieachtergrond van de vader van de respondent\*

```
UNIANOVA Schaal_uitgaan BY migratie_vader
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/EMMEANS=TABLES(migratie_vader)
/CRITERIA=ALPHA(0.05)
/DESIGN=NEWETH_F
```

## Univariate Analysis of Variance

### Between-Subjects Factors

		Value Label	N
migratie_vader	1,00	No migration background	1155
	2,00	Non-western migration background	76

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5,494 <sup>a</sup>	1	5,494	4,560	,033
Intercept	1066,412	1	1066,412	885,198	,000
migratie_vader	5,494	1	5,494	4,560	,033
Error	1480,596	1229	1,205		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

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

## Estimated Marginal Means

### migratie\_vader

Dependent Variable: Schaal\_uitgaan

migratie_vader	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration background	1,795	,032	1,731	1,858

2,00 Non-western migration background	2,072	,126	1,825	2,319
---------------------------------------	-------	------	-------	-------

\*Uitgaansgedrag en migratieachtergrond van de moeder van de respondent\*

```
UNIANOVA Schaal_uitgaan BY migratie_moeder
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(migratie_moeder)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWETH_M
```

## Univariate Analysis of Variance

### Between-Subjects Factors

		Value Label	N
migratie_moeder	1,00	No migration background	1169
	2,00	Non-western migration background	62

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	,497 <sup>a</sup>	1	,497	,411	,521
Intercept	808,868	1	808,868	669,160	,000
migratie_moeder	,497	1	,497	,411	,521
Error	1485,592	1229	1,209		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

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

## Estimated Marginal Means

### migratie\_moeder

Dependent Variable: Schaal\_uitgaan

migratie_moeder	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No migration back-ground	1,807	,032	1,744	1,870
2,00 Non-western migration background	1,899	,140	1,625	2,173

\*Uitgaansgedrag en seksuele oriëntatie\*

```
UNIANOVA Schaal_uitgaan BY NEWsex
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(NEWsex)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=NEWsex
```

## Univariate Analysis of Variance

### Between-Subjects Factors

	Value Label	N
NEWsex What do you think you are?	1,00 Heterosexual	1188
	2,00 LHB+	43

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2,332 <sup>a</sup>	1	2,332	1,931	,165
Intercept	613,303	1	613,303	508,000	,000
NEWsex	2,332	1	2,332	1,931	,165
Error	1483,758	1229	1,207		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,002 (Adjusted R Squared = ,001)

## Estimated Marginal Means

### What do you think you are?

Dependent Variable: Schaal\_uitgaan

What do you think you are?	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 Heterosexual	1,804	,032	1,741	1,866
2,00 Homosexual or bisexual(LHB+)	2,041	,168	1,712	2,369

\*Uitgaansgedrag en alcohol drinken\*

UNIANOVA Schaal\_uitgaan BY ooit\_alcohol

```

/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/EMMEANS=TABLES(ooit_alcohol)
/CRITERIA=ALPHA(0.05)
/DESIGN=drinknew2

```

## Univariate Analysis of Variance

### Between-Subjects Factors

	Value Label	N
	2,00 Yes	942

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	91,050 <sup>a</sup>	1	91,050	80,213	,000
Intercept	2384,341	1	2384,341	2100,554	,000
ooit_alcohol	91,050	1	91,050	80,213	,000

Error	1395,039	1229	1,135		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,061 (Adjusted R Squared = ,061)

## Estimated Marginal Means

### ooit\_alcohol

Dependent Variable: Schaal\_uitgaan

ooit_alcohol	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	1,321	,063	1,198	1,444
2,00 Yes	1,963	,035	1,894	2,031

\*Uitgaansgedrag en bingedrinken\*

```
UNIANOVA Schaal_uitgaan BY ooit_bingedrink
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_bingedrink)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=bingedrinknew2
```

## Univariate Analysis of Variance

### Between-Subjects Factors

		Value Label	N
ooit_bingedrink	1,00	No	667
	2,00	Yes	564

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	296,109 <sup>a</sup>	1	296,109	305,819	,000
Intercept	4197,738	1	4197,738	4335,384	,000
ooit_bingedrink	296,109	1	296,109	305,819	,000
Error	1189,980	1229	,968		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,199 (Adjusted R Squared = ,199)

## Estimated Marginal Means

### ooit\_bingedrink

Dependent Variable: Schaal\_uitgaan

ooit_bingedrink	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	1,361	,038	1,286	1,436
2,00 Yes	2,345	,041	2,264	2,427

\*Uitgaansgedrag en tabak roken\*

```
UNIANOVA Schaal_uitgaan BY ooit_tabak
/METHOD=SSTYPE(3)
/INTERCEPT=INCLUDE
/EMMEANS=TABLES(ooit_tabak)
/CRITERIA=ALPHA(0.05)
/DESIGN=smokenew2
```

## Univariate Analysis of Variance

### Between-Subjects Factors

	Value Label	N
ooit_tabak	1,00 No	743
	2,00 Yes	488

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	191,137 <sup>a</sup>	1	191,137	181,402	,000
Intercept	4232,545	1	4232,545	4016,980	,000
ooit_tabak	191,137	1	191,137	181,402	,000
Error	1294,953	1229	1,054		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,129 (Adjusted R Squared = ,128)

### Estimated Marginal Means

#### ooit\_tabak

Dependent Variable: Schaal\_uitgaan

ooit_tabak	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	1,493	,038	1,419	1,566
2,00 Yes	2,298	,046	2,207	2,389

\*Uitgaansgedrag en wiet roken\*

```
UNIANOVA Schaal_uitgaan BY ooit_wiet
  /METHOD=SSTYPE(3)
  /INTERCEPT=INCLUDE
  /EMMEANS=TABLES(ooit_wiet)
  /CRITERIA=ALPHA(0.05)
  /DESIGN=weednew2.
```

### Univariate Analysis of Variance

#### Between-Subjects Factors

	Value Label	N
ooit_wiet	1,00 No	1021



	2,00	Yes	210
--	------	-----	-----

### Tests of Between-Subjects Effects

Dependent Variable: Schaal\_uitgaan

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	202,366 <sup>a</sup>	1	202,366	193,739	,000
Intercept	3271,654	1	3271,654	3132,188	,000
ooit_wiet	202,366	1	202,366	193,739	,000
Error	1283,724	1229	1,045		
Total	5527,625	1231			
Corrected Total	1486,089	1230			

a. R Squared = ,136 (Adjusted R Squared = ,135)

### Estimated Marginal Means

#### ooit\_wiet

Dependent Variable: Schaal\_uitgaan

ooit_wiet	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1,00 No	1,628	,032	1,565	1,691
2,00 Yes	2,706	,071	2,568	2,844

\*T-toetsen van de gemiddelde scores op de continue variabelen per groep van de categorische variabelen\*

\*Migratieachtergrond van de respondent\*

```
T-TEST GROUPS=migratie_respondent(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

### T-Test

### Group Statistics

	migratie_respndent	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 No migration background	1216	14,9912	1,54727	,04437
	2,00 Non-western migration background	15	15,7000	2,01601	,52053
Schaal_uitgaan	1,00 No migration background	1216	1,8055	1,09676	,03145
	2,00 Non-western migration background	15	2,3333	1,20885	,31212

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	4,350	,037	-1,757	1229	,079	-,70884	,40355	-1,50057	,08289
	Equal variances not assumed			-1,357	14,204	,196	-,70884	,52242	-1,82781	,41013
Schaal_uitgaan	Equal variances assumed	1,196	,274	-1,850	1229	,065	-,52782	,28527	-1,08750	,03185
	Equal variances not assumed			-1,683	14,286	,114	-,52782	,31370	-1,19939	,14374

\*Migratieachtergrond van de vader van de respondent\*

```
T-TEST GROUPS=migratie_vader(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI (.95) .
```

### T-Test

### Group Statistics

	migratie_vader	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 No migration background	1155	15,0059	1,54599	,04549
	2,00 Non-western migration background	76	14,9068	1,69006	,19386
Schaal_uitgaan	1,00 No migration background	1155	1,7948	1,08554	,03194
	2,00 Non-western migration background	76	2,0724	1,26874	,14553

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	1,222	,269	,538	1229	,591	,09912	,18417	-,26220	,46043
	Equal variances not assumed			,498	83,470	,620	,09912	,19913	-,29691	,49514
Schaal_uitgaan	Equal variances assumed	1,848	,174	-2,135	1229	,033	-,27756	,12998	-,53257	-,02256
	Equal variances not assumed			-1,863	82,387	,066	-,27756	,14900	-,57395	,01882

\*Migratieachtergrond van de moeder van de respondent\*

```
T-TEST GROUPS=migratie_moeder(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

### T-Test

Group Statistics					
	migratie_moeder	N	Mean	Std. Deviation	Std. Error Mean

age_w1 What is your age?	1,00 No migration background	1169	14,9887	1,55580	,04550
	2,00 Non-western migration background	62	15,2083	1,53153	,19450
Schaal_uitgaan	1,00 No migration background	1169	1,8073	1,09209	,03194
	2,00 Non-western migration background	62	1,8992	1,23180	,15644

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	,341	,559	-1,084	1229	,279	-,21960	,20260	-,61708	,17789
	Equal variances not assumed			-1,099	67,849	,276	-,21960	,19976	-,61822	,17903
Schaal_uitgaan	Equal variances assumed	,201	,654	-,641	1229	,521	-,09188	,14328	-,37299	,18923
	Equal variances not assumed			-,575	66,186	,567	-,09188	,15967	-,41065	,22689

\*Geslacht\*

```
T-TEST GROUPS=gender_final(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI (.95).
```

### T-Test

#### Group Statistics

	gender_final What is your gender?	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1 Male	591	14,9776	1,57934	,06497

	2 Female	640	15,0203	1,53258	,06058
Schaal_uitgaan	1 Male	591	1,8735	1,17467	,04832
	2 Female	640	1,7551	1,02219	,04041

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
age_w1 What is your age?	Equal variances assumed	,373	,542	-,482	1229	,630	-,04273	,08872	-,21680	,13133
	Equal variances not assumed			-,481	1214,381	,631	-,04273	,08883	-,21701	,13154
Schaal_uitgaan	Equal variances assumed	7,982	,005	1,891	1229	,059	,11844	,06264	-,00445	,24134
	Equal variances not assumed			1,880	1173,718	,060	,11844	,06299	-,00514	,24202

\*Seksuele oriëntatie\*

```
T-TEST GROUPS=NEWsex(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

### T-Test

#### Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 Heterosexual	1188	14,9742	1,56314	,04535
	2,00 LHB+	43	15,7074	1,09629	,16718
Schaal_uitgaan	1,00 Heterosexual	1188	1,8037	1,10000	,03191

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	12,682	,000	-3,048	1229	,002	-,73318	,24054	-1,20508	-,26127
	Equal variances not assumed			-4,233	48,399	,000	-,73318	,17322	-1,08139	-,38496
Schaal_uitgaan	Equal variances assumed	,000	,984	-1,390	1229	,165	-,23704	,17057	-,57167	,09760
	Equal variances not assumed			-1,434	45,314	,158	-,23704	,16527	-,56985	,09578

\*Alcohol drinken\*

```
T-TEST GROUPS=ooit_alcohol(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

### T-Test

#### Group Statistics

	ooit_alcohol	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age	1,00 No	289	14,1557	1,35667	,07980
	2,00 Yes	942	15,2588	1,51995	,04952
Schaal_uitgaan	1,00 No	289	1,3209	,70322	,04137
	2,00 Yes	942	1,9626	1,15376	,03759

### Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	6,518	,011	-11,059	1229	,000	-1,10305	,09974	-1,29873	-,90736
	Equal variances not assumed			-11,744	528,530	,000	-1,10305	,09392	-1,28755	-,91854
Schaal_uitgaan	Equal variances assumed	77,881	,000	-8,956	1229	,000	-,64165	,07164	-,78220	-,50109
	Equal variances not assumed			-11,479	794,303	,000	-,64165	,05589	-,75136	-,53193

\*Bingedrinken\*

```
T-TEST GROUPS= ooit_ bingedrink(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI (.95) .
```

## T-Test

### Group Statistics

	ooit_bingedrink	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 No	667	14,4363	1,42224	,05507
	2,00 Yes	564	15,6662	1,43765	,06054
Schaal_uitgaan O	1,00 No	667	1,3609	,67510	,02614
	2,00 Yes	564	2,3453	1,25479	,05284

### Independent Samples Test

Levene's Test for Equality of Variances

t-test for Equality of Means

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	,300	,584	-15,043	1229	,000	-1,22994	,08176	-1,39035	-1,06953
	Equal variances not assumed			-15,029	1190,927	,000	-1,22994	,08184	-1,39050	-1,06938
Schaal_uitgaan	Equal variances assumed	182,140	,000	-17,488	1229	,000	-,98436	,05629	-1,09479	-,87392
	Equal variances not assumed			-16,699	830,288	,000	-,98436	,05895	-1,10006	-,86865

\*Tabak roken\*

```
T-TEST GROUPS=ooit_ tabak(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

## T-Test

### Group Statistics

	ooit_ tabak	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 No	743	14,6861	1,50172	,05509
	2,00 Yes	488	15,4775	1,51332	,06850
Schaal_uitgaan	1,00 No	743	1,4926	,80311	,02946
	2,00 Yes	488	2,2982	1,29473	,05861

### Independent Samples Test

Levene's Test for  
Equality of Vari-  
ances

t-test for Equality of Means



		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	,465	,495	-9,017	1229	,000	-,79139	,08777	-,96358	-,61919
	Equal variances not assumed			-9,002	1036,182	,000	-,79139	,08791	-,96389	-,61889
Schaal_uitgaan	Equal variances assumed	132,449	,000	-13,469	1229	,000	-,80556	,05981	-,92290	-,68822
	Equal variances not assumed			-12,280	733,499	,000	-,80556	,06560	-,93434	-,67677

\*Wiet roken\*

```
T-TEST GROUPS=ooit_wiet(1 2)
/MISSING=ANALYSIS
/VARIABLES=age_w1 Schaal_uitgaan
/CRITERIA=CI(.95).
```

## T-Test

### Group Statistics

	ooit_wiet	N	Mean	Std. Deviation	Std. Error Mean
age_w1 What is your age?	1,00 No	1021	14,8042	1,53666	,04809
	2,00 Yes	210	15,9508	1,26708	,08744
Schaal_uitgaan	1,00 No	1021	1,6281	,92557	,02897
	2,00 Yes	210	2,7060	1,40045	,09664

### Independent Samples Test

Levene's Test for  
Equality of Vari-  
ances

t-test for Equality of Means

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
age_w1 What is your age?	Equal variances assumed	23,209	,000	-10,127	1229	,000	-1,14660	,11322	-1,36873	-,92447
	Equal variances not assumed			-11,490	348,050	,000	-1,14660	,09979	-1,34286	-,95033
Schaal_uitgaan	Equal variances assumed	71,923	,000	-13,919	1229	,000	-1,07789	,07744	-1,22982	-,92596
	Equal variances not assumed			-10,684	247,831	,000	-1,07789	,10089	-1,27660	-,87918

\*Kruistabellen met bijbehorende Cramers V berekenen voor de categorische variabelen\*

\*Geslacht met alle andere variabelen\*

CROSSTABS

```

/TABLES=gender_final BY migratie_respondent migratie_vader migratie_moe-
der NEWsex ooit_alchol ooit_bingedrink ooit_tabak ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Geslacht met migratieachtergrond van de vader van de respondent\*

## gender\_final What is your gender?\* migratie\_vader

### Crosstab

Count

		migratie_vader		Total
		1,00 No migration background	2,00 Non-western migration back-ground	
gender_final What is your gender?	1 Male	557	34	591
	2 Female	598	42	640
Total		1155	76	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,348 <sup>a</sup>	1	,555		
Continuity Correction <sup>b</sup>	,222	1	,638		
Likelihood Ratio	,348	1	,555		
Fisher's Exact Test				,636	,319
Linear-by-Linear Association	,347	1	,556		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 36,49.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,017	,555
	Cramer's V	,017	,555
N of Valid Cases		1231	

\*Geslacht met migratieachtergrond van de moeder van de respondent\*

### gender\_final What is your gender? \* migratie\_moeder

#### Crosstab

Count

		migratie_moeder		Total
		1,00 No migration background	2,00 Non-western migration background	
gender_final What is your gender?	1 Male	563	28	591
	2 Female	606	34	640
Total		1169	62	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,212 <sup>a</sup>	1	,645		
Continuity Correction <sup>b</sup>	,109	1	,741		
Likelihood Ratio	,213	1	,645		
Fisher's Exact Test				,697	,371
Linear-by-Linear Association	,212	1	,645		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 29,77.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,013	,645
	Cramer's V	,013	,645
N of Valid Cases		1231	

\*Geslacht met migratieachtergrond van de respondent\*

### gender\_final What is your gender?\* migratie\_respondent

#### Crosstab

Count

		migratie_respondent		Total
		1,00 No migration background	2,00 Non-western migration background	
gender_final What is your gender?	1 Male	587	4	591
	2 Female	629	11	640
Total		1216	15	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,771 <sup>a</sup>	1	,096		
Continuity Correction <sup>b</sup>	1,973	1	,160		

Likelihood Ratio	2,897	1	,089		
Fisher's Exact Test				,121	,078
Linear-by-Linear Association	2,769	1	,096		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,20.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,047	,096
	Cramer's V	,047	,096
N of Valid Cases		1231	

\*Geslacht met seksuele oriëntatie\*

## gender\_final What is your gender?\* NEWsex What do you think you are?

### Crosstab

Count

		NEWsex What do you think you are?		Total
		1,00 Heterosexual	2,00 LHB+	
gender_final What is your gender?	1 Male	579	12	591
	2 Female	609	31	640
Total		1188	43	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	7,214 <sup>a</sup>	1	,007		
Continuity Correction <sup>b</sup>	6,404	1	,011		
Likelihood Ratio	7,499	1	,006		
Fisher's Exact Test				,008	,005
Linear-by-Linear Association	7,208	1	,007		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 20,64.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,077	,007
	Cramer's V	,077	,007
N of Valid Cases		1231	

\*Geslacht met alcohol drinken\*

### gender\_final What is your gender? \* ooit\_alcohol

#### Crosstab

Count

		ooit_alcohol		Total
		1,00 No	2,00 Yes	
gender_final What is your gender?	1 Male	134	457	591
	2 Female	155	485	640
Total		289	942	1231

#### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,408 <sup>a</sup>	1	,523		
Continuity Correction <sup>b</sup>	,327	1	,567		
Likelihood Ratio	,409	1	,523		
Fisher's Exact Test				,545	,284
Linear-by-Linear Association	,408	1	,523		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 138,75.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,018	,523
	Cramer's V	,018	,523
N of Valid Cases		1231	

\*Geslacht met bingedrinken\*

### gender\_final What is your gender? \* ooit\_bingedrink

#### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
gender_final What is your gender?	1 Male	299	292	591
	2 Female	368	272	640
Total		667	564	1231

#### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5,906 <sup>a</sup>	1	,015		
Continuity Correction <sup>b</sup>	5,631	1	,018		
Likelihood Ratio	5,909	1	,015		
Fisher's Exact Test				,016	,009
Linear-by-Linear Association	5,901	1	,015		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 270,77.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,069	,015
	Cramer's V	,069	,015
N of Valid Cases		1231	

\*Geslacht met tabak roken\*

## gender\_final What is your gender? \* ooit\_tabak

### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
gender_final What is your gender?	1 Male	341	250	591
	2 Female	402	238	640
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3,358 <sup>a</sup>	1	,067		
Continuity Correction <sup>b</sup>	3,148	1	,076		
Likelihood Ratio	3,358	1	,067		
Fisher's Exact Test				,071	,038
Linear-by-Linear Association	3,355	1	,067		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 234,29.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,052	,067
	Cramer's V	,052	,067
N of Valid Cases		1231	



\*Geslacht met wiet roken\*

## gender\_final What is your gender? \* ooit\_wiet

### Crosstab

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	
gender_final What is your gender?	1 Male	481	110	591
	2 Female	540	100	640
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,938 <sup>a</sup>	1	,164		
Continuity Correction <sup>b</sup>	1,733	1	,188		
Likelihood Ratio	1,937	1	,164		
Fisher's Exact Test				,173	,094
Linear-by-Linear Association	1,937	1	,164		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 100,82.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,040	,164
	Cramer's V	,040	,164
N of Valid Cases		1231	

\*Migratieachtergrond van de vader van de respondent met de overige variabelen\*

CROSSTABS

/TABLES=migratie\_vader BY migratie\_respondent migratie\_moeder NEWsex  
ooit\_alcohol ooit\_tabak ooit\_bingedrink ooit\_wiet

```

/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Migratieachtergrond van de vader van de respondent met migratieachtergrond van de moeder van de respondent\*

## migratie\_vader \* migratie\_moeder

### Crosstab

Count

		migratie_moeder		Total
		1,00 No migration background	2,00 Non-western migration background	
migratie_vader	1,00 No migration background	1141	14	1155
	2,00 Non-western migration background	28	48	76
Total		1169	62	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	572,098 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	559,220	1	,000		
Likelihood Ratio	239,970	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	571,634	1	,000		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,83.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,682	,000

Cramer's V	,682	,000
N of Valid Cases	1231	

\*Migratieachtergrond van de vader van de respondent met migratieachtergrond van de respondent\*

## migratie\_vader\* migratie\_respondent

### Crosstab

Count

		migratie_respondent		Total
		1,00 No migration background	2,00 Non-western migration background	
migratie_vader	1,00 No migration background	1150	5	1155
	2,00 Non-western migration background	66	10	76
Total		1216	15	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	95,928 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	85,647	1	,000		
Likelihood Ratio	38,455	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	95,850	1	,000		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is ,93.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,279	,000
	Cramer's V	,279	,000

N of Valid Cases	1231
------------------	------

\*Migratieachtergrond van de vader van de respondent met seksuele oriëntatie\*

## migratie\_vader\* NEWsex What do you think you are?

### Crosstab

Count

		NEWsex What do you think you are?		Total
		1,00 Heterosexual	2,00 LHB+	
migratie_vader	1,00 No migration background	1114	41	1155
	2,00 Non-western migration background	74	2	76
Total		1188	43	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,178 <sup>a</sup>	1	,673		
Continuity Correction <sup>b</sup>	,010	1	,920		
Likelihood Ratio	,193	1	,660		
Fisher's Exact Test				1,000	,497
Linear-by-Linear Association	,178	1	,673		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 2,65.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,012	,673
	Cramer's V	,012	,673
N of Valid Cases		1231	

\*Migratieachtergrond van de vader van de respondent met alcohol drinken\*

## migratie\_vader \* ooit\_alcohol

### Crosstab

Count

		Ooit_alcohol		Total
		1,00 No	2,00 Yes	
migratie_vader	1,00 No migration background	255	900	1155
	2,00 Non-western migration background	34	42	76
Total		289	942	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	20,379 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	19,137	1	,000		
Likelihood Ratio	17,780	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	20,362	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 17,84.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,129	,000
	Cramer's V	,129	,000
N of Valid Cases		1231	

\*Migratieachtergrond van de vader van de respondent met bingedrinken\*

## migratie\_vader \* ooit\_bingedrink

### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
migratie_vader	1,00 No migration background	618	537	1155
	2,00 Non-western migration background	49	27	76
Total		667	564	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3,455 <sup>a</sup>	1	,063		
Continuity Correction <sup>b</sup>	3,027	1	,082		
Likelihood Ratio	3,517	1	,061		
Fisher's Exact Test				,074	,040
Linear-by-Linear Association	3,452	1	,063		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 34,82.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,053	,063
	Cramer's V	,053	,063
N of Valid Cases		1231	

\*Migratieachtergrond van de vader van de respondent met tabak roken\*

**migratie\_vader \* ooit\_tabak**

### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
migratie_vader	1,00 No migration background	706	449	1155
	2,00 Non-western migration background	37	39	76
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,613 <sup>a</sup>	1	,032		
Continuity Correction <sup>b</sup>	4,108	1	,043		
Likelihood Ratio	4,512	1	,034		
Fisher's Exact Test				,039	,022
Linear-by-Linear Association	4,609	1	,032		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 30,13.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,061	,032
	Cramer's V	,061	,032
N of Valid Cases		1231	

\*Migratieachtergrond van de vader van de respondent met wiet roken\*

### migratie\_vader \* ooit\_wiet

### Crosstab

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	

migratie_vader	1,00 No migration background	963	192	1155
	2,00 Non-western migration background	58	18	76
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,513 <sup>a</sup>	1	,113		
Continuity Correction <sup>b</sup>	2,038	1	,153		
Likelihood Ratio	2,312	1	,128		
Fisher's Exact Test				,116	,080
Linear-by-Linear Association	2,511	1	,113		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 12,97.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,045	,113
	Cramer's V	,045	,113
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met de overige variabelen\*

CROSSTABS

```

/TABLES=migratie_moeder BY migratie_respondent NEWsex ooit_alcohol
ooit_bingedrink ooit_tabak ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Migratieachtergrond van de moeder van de respondent met migratieachtergrond van de respondent\*

**migratie\_moeder \* migratie\_respondent**



### Crosstab

Count

		migratie_respondent		Total
		1,00 No migration background	2,00 Non-western migration background	
migratie_moeder	1,00 No migration background	1165	4	1169
	2,00 Non-western migration background	51	11	62
Total		1216	15	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	148,090 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	133,987	1	,000		
Likelihood Ratio	50,671	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	147,970	1	,000		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is ,76.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,347	,000
	Cramer's V	,347	,000
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met seksuele oriëntatie\*

migratie\_moeder \* NEWsex What do you think you are?

### Crosstab

Count

		NEWsex What do you think you are?		Total
		1,00 Heterosexual	2,00 LHB+	
migratie_moeder	1,00 No migration background	1130	39	1169
	2,00 Non-western migration background	58	4	62
Total		1188	43	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1,695 <sup>a</sup>	1	,193		
Continuity Correction <sup>b</sup>	,897	1	,344		
Likelihood Ratio	1,383	1	,240		
Fisher's Exact Test				,271	,167
Linear-by-Linear Association	1,694	1	,193		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 2,17.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,037	,193
	Cramer's V	,037	,193
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met alcohol drinken\*

### migratie\_moeder \* ooit\_alcohol

### Crosstab

Count

ooit_alcohol	Total
--------------	-------

		1,00 No	2,00 Yes	
migratie_moeder	1,00 No migration background	259	910	1169
	2,00 Non-western migration background	30	32	62
Total		289	942	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	22,551 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	21,114	1	,000		
Likelihood Ratio	19,340	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	22,532	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 14,56.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,135	,000
	Cramer's V	,135	,000
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met bingedrinken\*

## migratie\_moeder \* ooit\_bingedrink

### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
migratie_moeder	1,00 No migration background	626	543	1169
	2,00 Non-western migration background	41	21	62

Total		667	564	1231
-------	--	-----	-----	------

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3,753 <sup>a</sup>	1	,053		
Continuity Correction <sup>b</sup>	3,263	1	,071		
Likelihood Ratio	3,838	1	,050		
Fisher's Exact Test				,066	,034
Linear-by-Linear Association	3,750	1	,053		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 28,41.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-,055	,053
	Cramer's V	,055	,053
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met tabak roken\*

## migratie\_moeder \* ooit\_tabak

### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
migratie_moeder	1,00 No migration background	711	458	1169
	2,00 Non-western migration background	32	30	62
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,086 <sup>a</sup>	1	,149		
Continuity Correction <sup>b</sup>	1,719	1	,190		
Likelihood Ratio	2,048	1	,152		
Fisher's Exact Test				,182	,096
Linear-by-Linear Association	2,085	1	,149		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 24,58.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,041	,149
	Cramer's V	,041	,149
N of Valid Cases		1231	

\*Migratieachtergrond van de moeder van de respondent met wiet roken\*

### migratie\_moeder \* ooit\_wiet

#### Crosstab

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	
migratie_moeder	1,00 No migration background	975	194	1169
	2,00 Non-western migration background	46	16	62
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3,531 <sup>a</sup>	1	,060		
Continuity Correction <sup>b</sup>	2,910	1	,088		
Likelihood Ratio	3,172	1	,075		
Fisher's Exact Test				,081	,049
Linear-by-Linear Association	3,528	1	,060		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 10,58.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,054	,060
	Cramer's V	,054	,060
N of Valid Cases		1231	

\*Migratieachtergrond van de respondent met de overige variabelen\*

CROSSTABS

```

/TABLES=migratie_respondent BY NEWsex ooit_alcohol ooit_bingedrink
ooit_tabak ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Migratieachtergrond van de respondent met seksuele oriëntatie\*

### migratie\_respondent \* NEWsex What do you think you are?

#### Crosstab

Count

		NEWsex What do you think you are?		Total
		1,00 Heterosexual	2,00 LHB+	
migratie_respondent	1,00 No migration background	1175	41	1216

	2,00 Non-western migration background	13	2	15
Total		1188	43	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,362 <sup>a</sup>	1	,037		
Continuity Correction <sup>b</sup>	1,907	1	,167		
Likelihood Ratio	2,615	1	,106		
Fisher's Exact Test				,094	,094
Linear-by-Linear Association	4,358	1	,037		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is ,52.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,060	,037
	Cramer's V	,060	,037
N of Valid Cases		1231	

\*Migratieachtergrond van de respondent met alcohol drinken\*

## migratie\_respondent \* ooit\_alcohol

### Crosstab

Count

		ooit_alcohol		Total
		1,00 No	2,00 Yes	
migratie_respondent	1,00 No migration background	286	930	1216

	2,00 Non-western migration background	3	12	15
Total		289	942	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,102 <sup>a</sup>	1	,749		
Continuity Correction <sup>b</sup>	,000	1	,989		
Likelihood Ratio	,106	1	,745		
Fisher's Exact Test				1,000	,517
Linear-by-Linear Association	,102	1	,749		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,52.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,009	,749
	Cramer's V	,009	,749
N of Valid Cases		1231	

\*migratie achtergrond van de respondent met de bingedrinken\*

## migratie\_respondent \* ooit\_bingedrink

### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
migratie_respondent	1,00 No migration background	662	554	1216



	2,00 Non-western migration background	5	10	15
Total		667	564	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2,659 <sup>a</sup>	1	,103		
Continuity Correction <sup>b</sup>	1,877	1	,171		
Likelihood Ratio	2,675	1	,102		
Fisher's Exact Test				,122	,085
Linear-by-Linear Association	2,657	1	,103		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 6,87.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,046	,103
	Cramer's V	,046	,103
N of Valid Cases		1231	

\*Migratieachtergrond van de respondent met tabak roken\*

## migratie\_respondent \* ooit\_tabak

### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
migratie_respondent	1,00 No migration background	738	478	1216

2,00 Non-western migration background	5	10	15
Total	743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4,635 <sup>a</sup>	1	,031		
Continuity Correction <sup>b</sup>	3,562	1	,059		
Likelihood Ratio	4,515	1	,034		
Fisher's Exact Test				,036	,031
Linear-by-Linear Association	4,631	1	,031		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,95.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,061	,031
	Cramer's V	,061	,031
N of Valid Cases		1231	

\*Migratieachtergrond van de respondent met wiet roken\*

## migratie\_respondent \* ooit\_wiet

### Crosstab

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	
migratie_respondent	1,00 No migration background	1009	207	1216

2,00 Non-western migration background	12	3	15
Total	1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	,093 <sup>a</sup>	1	,761		
Continuity Correction <sup>b</sup>	,000	1	1,000		
Likelihood Ratio	,089	1	,765		
Fisher's Exact Test				,730	,485
Linear-by-Linear Association	,093	1	,761		
N of Valid Cases	1231				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 2,56.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,009	,761
	Cramer's V	,009	,761
N of Valid Cases		1231	

\*Seksuele oriëntatie met de overige variabelen\*

CROSSTABS

```

/TABLES=NEWsex BY ooit_alcohol ooit_bingedrink ooit_tabak ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Seksuele oriëntatie met alcohol drinken\*

**NEWsex What do you think you are? \* ooit\_alcohol**

### Crosstab

Count

		ooit_alcohol		Total
		1,00 No	2,00 Yes	
NEWsex What do you think you are?	1,00 Heterosexual	287	901	1188
	2,00 LHB+	2	41	43
Total		289	942	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8,790 <sup>a</sup>	1	,003		
Continuity Correction <sup>b</sup>	7,738	1	,005		
Likelihood Ratio	11,865	1	,001		
Fisher's Exact Test				,001	,001
Linear-by-Linear Association	8,783	1	,003		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 10,10.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,085	,003
	Cramer's V	,085	,003
N of Valid Cases		1231	

\*Seksuele oriëntatie met bingedrinken\*

**NEWsex What do you think you are? \* ooit\_bingedrink**

### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
NEWsex What do you think you are?	1,00 Heterosexual	655	533	1188
	2,00 LHB+	12	31	43
Total		667	564	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	12,393 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	11,320	1	,001		
Likelihood Ratio	12,615	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	12,382	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 19,70.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,100	,000
	Cramer's V	,100	,000
N of Valid Cases		1231	

\*Seksuele oriëntatie met tabak roken\*

### NEWsex What do you think you are? \* ooit\_tabak

### Crosstab

Count

ooit_tabak	Total
------------	-------

		1,00 No	2,00 Yes	
NEWsex What do you think you are?	1,00 Heterosexual	729	459	1188
	2,00 LHB+	14	29	43
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	14,391 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	13,212	1	,000		
Likelihood Ratio	14,040	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	14,379	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 17,05.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,108	,000
	Cramer's V	,108	,000
N of Valid Cases		1231	

\*Seksuele oriëntatie met wiet roken\*

### NEWsex What do you think you are? \* ooit\_wiet

#### Crosstab

Count

	ooit_wiet		Total
	1,00 No	2,00 Yes	
1,00 Heterosexual	992	196	1188

NEWsex What do you think you are?	2,00 LHB+	29	14	43
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	7,564 <sup>a</sup>	1	,006		
Continuity Correction <sup>b</sup>	6,472	1	,011		
Likelihood Ratio	6,366	1	,012		
Fisher's Exact Test				,011	,009
Linear-by-Linear Association	7,558	1	,006		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 7,34.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,078	,006
	Cramer's V	,078	,006
N of Valid Cases		1231	

\*Alcohol drinken met de overige variabelen\*

CROSSTABS

```

/TABLES=drinknew2 BY ooit_bingedrink ooit_tabak ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

\*Alcohol drinken met bingedrinken\*

**ooit\_alcohol \* ooit\_bingedrink**

### Crosstab

Count

		ooit_bingedrink		Total
		1,00 No	2,00 Yes	
ooit_alcohol	1,00 No	287	2	289
	2,00 Yes	380	562	942
Total		667	564	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	309,769 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	307,399	1	,000		
Likelihood Ratio	403,517	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	309,518	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 132,41.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,502	,000
	Cramer's V	,502	,000
N of Valid Cases		1231	

\*Alcohol drinken met tabak roken\*

**ooit\_alcohol \* ooit\_tabak**



### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
ooit_alcohol	1,00 No	264	25	289
	2,00 Yes	479	463	942
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	151,605 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	149,917	1	,000		
Likelihood Ratio	177,554	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	151,482	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 114,57.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,351	,000
	Cramer's V	,351	,000
N of Valid Cases		1231	

\*Alcohol drinken met wiet roken\*

### ooit\_alcohol \* ooit\_wiet

### Crosstab

Count

		ooit_wiet	Total
--	--	-----------	-------

		1,00 No	2,00 Yes	
ooit_alcohol	1,00 No	287	2	289
	2,00 Yes	734	208	942
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	71,504 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	70,000	1	,000		
Likelihood Ratio	106,210	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	71,446	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 49,30.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,241	,000
	Cramer's V	,241	,000
N of Valid Cases		1231	

\*Bingedrinken met de overige variabelen\*

CROSSTABS

/TABLES=ooit\_bingedrink BY ooit\_tabak ooit\_wiet

/FORMAT=AVALUE TABLES

/STATISTICS=CHISQ PHI

/CELLS=COUNT

/COUNT ROUND CELL.

\*Bingedrinken met tabak roken\*

**ooit\_bingedrink \* ooit\_tabak**

### Crosstab

Count

		ooit_tabak		Total
		1,00 No	2,00 Yes	
ooit_bingedrink	1,00 No	558	109	667
	2,00 Yes	185	379	564
Total		743	488	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	330,333 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	328,211	1	,000		
Likelihood Ratio	345,540	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	330,064	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 223,58.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,518	,000
	Cramer's V	,518	,000
N of Valid Cases		1231	

\*Bingedrinken met wiet roken\*

**ooit\_bingedrink \* ooit\_wiet**

### Crosstab

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	
ooit_bingedrink	1,00 No	651	16	667
	2,00 Yes	370	194	564
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	221,143 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	218,887	1	,000		
Likelihood Ratio	247,704	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	220,963	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 96,21.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,424	,000
	Cramer's V	,424	,000
N of Valid Cases		1231	

\*Tabak roken met wiet roken\*

CROSSTABS

```

/TABLES=ooit_tabak BY ooit_wiet
/FORMAT=AVALUE TABLES
/STATISTICS=CHISQ PHI
/CELLS=COUNT
/COUNT ROUND CELL.

```

### Crosstabs

### ooit\_tabak \* ooit\_wiet Crosstabulation

Count

		ooit_wiet		Total
		1,00 No	2,00 Yes	
ooit_tabak	1,00 No	737	6	743
	2,00 Yes	284	204	488
Total		1021	210	1231

### Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	349,864 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	346,973	1	,000		
Likelihood Ratio	391,588	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	349,580	1	,000		
N of Valid Cases	1231				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 83,25.

b. Computed only for a 2x2 table

### Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	,533	,000
	Cramer's V	,533	,000
N of Valid Cases		1231	

## Regressieanalyses

Tot slot zijn de logistische regressie analyses uitgevoerd zoals beschreven in het analyseplan. Na het uitvoeren van de analyses vielen een aantal veranderingen van hellingen op en zijn er extra analyses uitgevoerd om te kijken waar deze veranderingen in hellingen vandaan komen. Hieronder staan de syntax en de output van alle uitgevoerde analyses.

### Output regressieanalyses

\*Lineaire regressie van seksuele oriëntatie op uitgaansgedrag\*

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Schaal_uitgaan
  /METHOD=ENTER gender_final age_w1 migratie_respondent migratie_vader mi-
  gratie_moeder NEWsex.
```

## Regression

**Variables Entered/Removed<sup>a</sup>**

Model	Variables Entered	Variables Remo- ved	Method
1	NEWsex What do you think you are?, migratie_vader, gender_final What is your gender?, age_w1 What is your age?, migratie_respondent, migratie_moeder <sup>b</sup>	.	Enter

a. Dependent Variable: Schaal\_uitgaan

b. All requested variables entered.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,277 <sup>a</sup>	,077	,072	1,05880

a. Predictors: (Constant), NEWsex What do you think you are?, migratie\_vader, gender\_final What is your gender?, age\_w1 What is your age?, migratie\_respondent, migratie\_moeder

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	113,926	6	18,988	16,937	,000 <sup>b</sup>
	Residual	1372,164	1224	1,121		
	Total	1486,089	1230			

a. Dependent Variable: Schaal\_uitgaan

b. Predictors: (Constant), NEWsex What do you think you are?, migratie\_vader, gender\_final What is your gender?, age\_w1 What is your age?, migratie\_respondent, migratie\_moeder

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.
		B	Std. Error			
1	(Constant)	-1,346	,426		-3,157	,002
	gender_final What is your gender?	-,136	,061	-,062	-2,242	,025
	age_w1 What is your age?	,181	,020	,256	9,260	,000
	migratie_respondent	,379	,295	,038	1,284	,199
	migratie_vader	,490	,172	,107	2,844	,005
	migratie_moeder	-,382	,194	-,076	-1,968	,049
	NEWsex What do you think you are?	,144	,166	,024	,867	,386

a. Dependent Variable: Schaal\_uitgaan

\*Logistische regressie analyses voor ieder type middelengebruik\*

\*Alcohol drinken

```
LOGISTIC REGRESSION VARIABLES ooit_alcohol
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
  gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
```

## Logistic Regression

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1231	100,0
	Missing Cases	0	,0
	Total	1231	100,0
Unselected Cases		0	,0
Total		1231	100,0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
1,00 No	0
2,00 Yes	1

### Block 0: Beginning Block

Classification Table<sup>a,b</sup>

	Observed	Predicted		Percentage Correct
		ooit_alcohol		
		1,00 No	2,00 Yes	
Step 0	ooit_alcohol	1,00 No	2,00 Yes	
		0	289	,0
		2,00 Yes	942	100,0
Overall Percentage				76,5

a. Constant is included in the model.

b. The cut value is ,500

### Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 0	Constant	1,182	,067	308,756	1	,000	3,260



### Variables not in the Equation

		Score	df	Sig.	
Step 0	Variables	What is your age?	111,411	1	,000
		What is your gender?	,408	1	,523
		migratie_respondent	,102	1	,749
		migratie_vader	20,379	1	,000
		migratie_moeder	22,551	1	,000
Overall Statistics		142,275	5	,000	

### Block 1: Method = Enter

#### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	148,924	5	,000
	Block	148,924	5	,000
	Model	148,924	5	,000

#### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1192,803 <sup>a</sup>	,114	,172

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

#### Classification Table<sup>a</sup>

		Predicted		Percentage Correct
		ooit_alcohol 1,00 No	ooit_alcohol 2,00 Yes	
Step 1	ooit_alcohol	1,00 No	250	13,5
		2,00 Yes	916	97,2
Overall Percentage				77,6

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,532	,052	104,807	1	,000	1,702
	What is your gender?	-,137	,144	,900	1	,343	,872
	migratie_respondent	1,275	,750	2,887	1	,089	3,577
	migratie_vader	-,423	,360	1,380	1	,240	,655
	migratie_moeder	-1,343	,403	11,089	1	,001	,261
	Constant	-5,822	1,038	31,444	1	,000	,003

a. Variable(s) entered on step 1: What is your age?, What is your gender?, migratie\_respondent, migratie\_vader, migratie\_moeder.

## Block 2: Method = Enter

### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	7,566	1	,006
	Block	7,566	1	,006
	Model	156,490	6	,000

### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1185,237 <sup>a</sup>	,119	,180

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

### Classification Table<sup>a</sup>

	Observed	Predicted		Percentage Correct	
		1,00 No	2,00 Yes		
Step 1	ooit_alcohol	1,00 No	41	248	14,2
		2,00 Yes	25	917	97,3
Overall Percentage					77,8

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,521	,052	100,663	1	,000	1,683
	What is your gender?	-,160	,145	1,218	1	,270	,852
	migratie_respondent	1,290	,763	2,859	1	,091	3,634
	migratie_vader	-,367	,368	,993	1	,319	,693
	migratie_moeder	-1,438	,416	11,976	1	,001	,237
	What do you think you are?	1,640	,746	4,836	1	,028	5,154
	Constant	-7,279	1,264	33,153	1	,000	,001

a. Variable(s) entered on step 1: What do you think you are?.

### Block 3: Method = Enter

#### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	64,876	1	,000
	Block	64,876	1	,000
	Model	221,366	7	,000

#### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1120,360 <sup>a</sup>	,165	,248

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

#### Classification Table<sup>a</sup>

	Observed	Predicted		Percentage Correct
		1,00 No	2,00 Yes	
Step 1	ooit_alcohol	1,00 No	2,00 Yes	
		61	228	21,1
		2,00 Yes		
		35	907	96,3
	Overall Percentage			78,6

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,434	,054	65,320	1	,000	1,544
	What is your gender?	-,123	,149	,688	1	,407	,884
	migratie_respondent	1,029	,791	1,691	1	,193	2,798
	migratie_vader	-,764	,399	3,658	1	,056	,466
	migratie_moeder	-1,274	,446	8,147	1	,004	,280
	What do you think you are?	1,592	,760	4,387	1	,036	4,911
	Schaal_uitgaan	,807	,122	43,486	1	,000	2,241
	Constant	-6,754	1,284	27,674	1	,000	,001

a. Variable(s) entered on step 1: Schaal\_uitgaan.

\*Bingedrinken\*

```
LOGISTIC REGRESSION VARIABLES ooit_bingedrink
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
  gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
```

## Logistic Regression

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1231	100,0
	Missing Cases	0	,0
	Total	1231	100,0
Unselected Cases		0	,0
Total		1231	100,0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
1,00 No	0
2,00 Yes	1

## Block 0: Beginning Block

**Classification Table<sup>a,b</sup>**

	Observed	Predicted		Percentage Correct	
		ooit_bingedrink 1,00 No	2,00 Yes		
Step 0	ooit_bingedrink	1,00 No	667	0	100,0
		2,00 Yes	564	0	,0
	Overall Percentage				54,2

a. Constant is included in the model.

b. The cut value is ,500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-,168	,057	8,598	1	,003	,846

**Variables not in the Equation**

		Score	df	Sig.	
Step 0	Variables	What is your age?	191,410	1	,000
		What is your gender?	5,906	1	,015
		migratie_respondent	2,659	1	,103
		migratie_vader	3,455	1	,063
		migratie_moeder	3,753	1	,053
	Overall Statistics	207,711	5	,000	

## Block 1: Method = Enter

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	224,034	5	,000
	Block	224,034	5	,000
	Model	224,034	5	,000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1473,866 <sup>a</sup>	,166	,222

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5,474	8	,706

### Contingency Table for Hosmer and Lemeshow Test

		ooit_bingedrink = 1,00 No		ooit_bingedrink = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	107	107,094	19	18,906	126
	2	94	95,774	29	27,226	123
	3	95	92,019	33	35,981	128
	4	87	81,029	37	42,971	124
	5	70	71,463	53	51,537	123
	6	61	63,077	63	60,923	124
	7	45	53,692	79	70,308	124
	8	44	45,013	80	78,987	124
	9	42	37,225	84	88,775	126
	10	22	20,614	87	88,386	109

### Classification Table<sup>a</sup>

	Observed	Predicted		Percentage Correct
		ooit_bingedrink 1,00 No	ooit_bingedrink 2,00 Yes	
Step 1	ooit_bingedrink 1,00 No	493	174	73,9
	ooit_bingedrink 2,00 Yes	210	354	62,8
	Overall Percentage			68,8

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,592	,045	170,142	1	,000	1,807
	What is your gender?	-,375	,126	8,806	1	,003	,687
	migratie_respondent	1,544	,680	5,163	1	,023	4,683
	migratie_vader	-,101	,371	,074	1	,786	,904
	migratie_moeder	-,930	,424	4,811	1	,028	,394
	Constant	-8,979	,947	89,872	1	,000	,000

a. Variable(s) entered on step 1: What is your age?, What is your gender?, migratie\_respondent, migratie\_vader, migratie\_moeder.

## Block 2: Method = Enter

### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	7,462	1	,006
	Block	7,462	1	,006
	Model	231,496	6	,000

### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1466,404 <sup>a</sup>	,171	,229

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4,647	8	,795

### Contingency Table for Hosmer and Lemeshow Test

		ooit_bingedrink = 1,00 No		ooit_bingedrink = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	106	107,449	20	18,551	126
	2	102	99,891	26	28,109	128
	3	87	85,223	31	32,777	118
	4	85	80,423	37	41,577	122
	5	70	70,523	50	49,477	120
	6	61	63,805	63	60,195	124
	7	46	54,420	78	69,580	124
	8	44	45,170	79	77,830	123
	9	41	37,489	84	87,511	125
	10	25	22,608	96	98,392	121

**Classification Table<sup>a</sup>**

	Observed	Predicted		Percentage Correct	
		1,00 No	2,00 Yes		
Step 1	ooit_bingedrink	1,00 No	493	174	73,9
		2,00 Yes	203	361	64,0
Overall Percentage					69,4

a. The cut value is ,500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,585	,045	165,629	1	,000	1,794
	What is your gender?	-,404	,127	10,086	1	,001	,668
	migratie_respondent	1,525	,694	4,831	1	,028	4,594
	migratie_vader	-,055	,373	,022	1	,883	,947
	migratie_moeder	-1,001	,426	5,506	1	,019	,368
	What do you think you are?	,962	,368	6,849	1	,009	2,617
	Constant	-9,781	1,010	93,866	1	,000	,000

a. Variable(s) entered on step 1: What do you think you are?.

### Block 3: Method = Enter

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	212,435	1	,000
	Block	212,435	1	,000
	Model	443,931	7	,000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1253,969 <sup>a</sup>	,303	,405

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

**Hosmer and Lemeshow Test**

Step	Chi-square	df	Sig.
1	8,126	8	,421



### Contingency Table for Hosmer and Lemeshow Test

		ooit_bingedrink = 1,00 No		ooit_bingedrink = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	115	112,137	9	11,863	124
	2	104	106,328	22	19,672	126
	3	106	98,484	19	26,516	125
	4	90	88,031	32	33,969	122
	5	72	78,963	51	44,037	123
	6	66	66,794	57	56,206	123
	7	51	53,395	72	69,605	123
	8	33	37,874	90	85,126	123
	9	23	19,937	100	103,063	123
	10	7	5,058	112	113,942	119

### Classification Table<sup>a</sup>

		Predicted		Percentage Correct	
		ooit_bingedrink			
		1,00 No	2,00 Yes		
Step 1	ooit_bingedrink				
		1,00 No	550	117	82,5
		2,00 Yes	185	379	67,2
Overall Percentage				75,5	

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,494	,049	102,837	1	,000	1,639
	What is your gender?	-,406	,140	8,374	1	,004	,666
	migratie_respondent	1,285	,753	2,913	1	,088	3,616
	migratie_vader	-,717	,423	2,866	1	,090	,488
	migratie_moeder	-,741	,481	2,373	1	,123	,477
	What do you think you are?	,941	,390	5,829	1	,016	2,562
	Schaal_uitgaan	1,125	,095	139,086	1	,000	3,079
	Constant	-9,661	1,064	82,391	1	,000	,000

a. Variable(s) entered on step 1: Schaal\_uitgaan.

\*Tabak roken\*

```
LOGISTIC REGRESSION VARIABLES ooit_tabak
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
  gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
```

## Logistic Regression

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1231	100,0
	Missing Cases	0	,0
	Total	1231	100,0
Unselected Cases		0	,0
Total		1231	100,0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
1,00 No	0
2,00 Yes	1

## Block 0: Beginning Block

### Classification Table<sup>a,b</sup>

		Predicted		Percentage Cor- rect
		ooit_tabak 1,00 No	2,00 Yes	
Step 0	ooit_tabak	743	0	100,0
		1,00 No	488	0
	2,00 Yes			
	Overall Percentage			60,4

a. Constant is included in the model.

b. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-,420	,058	52,052	1	,000	,657

### Variables not in the Equation

		Score	df	Sig.	
Step 0	Variables	What is your age?	76,380	1	,000
		What is your gender?	3,358	1	,067
		migratie_respondent	4,635	1	,031
		migratie_vader	4,613	1	,032
		migratie_moeder	2,086	1	,149
Overall Statistics		87,676	5	,000	

### Block 1: Method = Enter

#### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	90,059	5	,000
	Block	90,059	5	,000
	Model	90,059	5	,000

#### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1563,262 <sup>a</sup>	,071	,095

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

#### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	16,012	8	,042

### Contingency Table for Hosmer and Lemeshow Test

		ooit_tabak = 1,00 No		ooit_tabak = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	98	100,324	28	25,676	126
	2	99	90,495	22	30,505	121
	3	90	91,601	39	37,399	129
	4	88	86,246	41	42,754	129
	5	61	74,828	58	44,172	119
	6	77	71,865	45	50,135	122
	7	65	67,722	59	56,278	124
	8	58	63,590	68	62,410	126
	9	66	55,928	56	66,072	122
	10	41	40,402	72	72,598	113

### Classification Table<sup>a</sup>

		Predicted		Percentage Correct	
		ooit_tabak			
		1,00 No	2,00 Yes		
Step 1	ooit_tabak	1,00 No	616	127	82,9
		2,00 Yes	335	153	31,4
	Overall Percentage				62,5

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,349	,041	73,144	1	,000	1,417
	What is your gender?	-,256	,121	4,444	1	,035	,774
	migratie_respondent	,859	,621	1,916	1	,166	2,361
	migratie_vader	,681	,341	4,002	1	,045	1,977
	migratie_moeder	-,312	,386	,655	1	,418	,732
	Constant	-6,555	,876	55,972	1	,000	,001

a. Variable(s) entered on step 1: What is your age?, What is your gender?, migratie\_respondent, migratie\_vader, migratie\_moeder.

## Block 2: Method = Enter

### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10,533	1	,001
	Block	10,533	1	,001
	Model	100,592	6	,000

### Model Summary

Step	-2 Log likelihood	Cox & Snell R	Nagelkerke R
		Square	Square
1	1552,729 <sup>a</sup>	,078	,106

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	16,254	8	,039

### Contingency Table for Hosmer and Lemeshow Test

		ooit_tabak = 1,00 No		ooit_tabak = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	93	95,431	26	23,569	119
	2	105	94,230	20	30,770	125
	3	85	88,903	39	35,097	124
	4	80	83,322	43	39,678	123
	5	74	80,863	53	46,137	127
	6	81	71,895	40	49,105	121
	7	64	69,885	63	57,115	127
	8	61	64,070	65	61,930	126
	9	66	55,953	57	67,047	123
	10	34	38,448	82	77,552	116

**Classification Table<sup>a</sup>**

	Observed	Predicted		Percentage Correct
		ooit_tabak 1,00 No	ooit_tabak 2,00 Yes	
Step 1	ooit_tabak	1,00 No	2,00 Yes	
		627	116	84,4
		331	157	32,2
	Overall Percentage			63,7

a. The cut value is ,500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,340	,041	69,277	1	,000	1,405
	What is your gender?	-,288	,122	5,564	1	,018	,750
	migratie_respondent	,783	,631	1,540	1	,215	2,188
	migratie_vader	,742	,344	4,652	1	,031	2,100
	migratie_moeder	-,380	,389	,952	1	,329	,684
	What do you think you are?	1,067	,340	9,821	1	,002	2,906
	Constant	-7,402	,926	63,950	1	,000	,001

a. Variable(s) entered on step 1: What do you think you are?.

### Block 3: Method = Enter

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	116,866	1	,000
	Block	116,866	1	,000
	Model	217,458	7	,000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1435,863 <sup>a</sup>	,162	,219

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7,685	8	,465

### Contingency Table for Hosmer and Lemeshow Test

		ooit_tabak = 1,00 No		ooit_tabak = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	106	100,959	15	20,041	121
	2	97	98,851	27	25,149	124
	3	99	95,460	26	29,540	125
	4	83	89,541	40	33,459	123
	5	85	84,260	38	38,740	123
	6	86	78,417	37	44,583	123
	7	66	69,758	57	53,242	123
	8	56	59,180	67	63,820	123
	9	41	44,911	82	78,089	123
	10	24	21,664	99	101,336	123

### Classification Table<sup>a</sup>

	Observed	Predicted		Percentage Correct
		ooit_tabak 1,00 No	ooit_tabak 2,00 Yes	
Step 1	ooit_tabak 1,00 No	641	102	86,3
	ooit_tabak 2,00 Yes	263	225	46,1
	Overall Percentage			70,3

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,244	,043	32,088	1	,000	1,277
	What is your gender?	-,229	,129	3,182	1	,074	,795
	migratie_respondent	,515	,658	,614	1	,433	1,674
	migratie_vader	,467	,366	1,629	1	,202	1,595
	migratie_moeder	-,135	,413	,107	1	,743	,873
	What do you think you are?	1,045	,350	8,896	1	,003	2,842
	Schaal_uitgaan	,676	,070	93,318	1	,000	1,966
	Constant	-6,944	,949	53,529	1	,000	,001

a. Variable(s) entered on step 1: Schaal\_uitgaan.

\*Wiet roken\*

```
LOGISTIC REGRESSION VARIABLES ooit_wiet
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
  gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
```

## Logistic Regression

### Case Processing Summary

Unweighted Cases <sup>a</sup>		N	Percent
Selected Cases	Included in Analysis	1231	100,0
	Missing Cases	0	,0
	Total	1231	100,0
Unselected Cases		0	,0
Total		1231	100,0

a. If weight is in effect, see classification table for the total number of cases.

### Dependent Variable Encoding

Original Value	Internal Value
1,00 No	0
2,00 Yes	1

## Block 0: Beginning Block

### Classification Table<sup>a,b</sup>

		Predicted			Percentage Cor- rect
		ooit_wiet			
Observed		1,00 No	2,00 Yes		
Step 0	ooit_wiet	1,00 No	1021	0	100,0
		2,00 Yes	210	0	,0
Overall Percentage					82,9

a. Constant is included in the model.

b. The cut value is ,500



### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1,581	,076	435,599	1	,000	,206

### Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	What is your age?	94,811	1	,000
		What is your gender?	1,938	1	,164
		migratie_respondent	,093	1	,761
		migratie_vader	2,513	1	,113
		migratie_moeder	3,531	1	,060
	Overall Statistics		101,125	5	,000

## Block 1: Method = Enter

### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	104,003	5	,000
	Block	104,003	5	,000
	Model	104,003	5	,000

### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1020,701 <sup>a</sup>	,081	,135

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	16,696	8	,033

### Contingency Table for Hosmer and Lemeshow Test

		ooit_wiet = 1,00 No		ooit_wiet = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	122	118,660	2	5,340	124
	2	118	116,348	6	7,652	124
	3	117	115,820	9	10,180	126
	4	117	111,993	8	13,007	125
	5	103	108,653	22	16,347	125
	6	97	100,610	23	19,390	120
	7	93	96,176	27	23,824	120
	8	86	93,694	37	29,306	123
	9	85	88,304	39	35,696	124
	10	83	70,743	37	49,257	120

### Classification Table<sup>a</sup>

		Predicted		Percentage Correct	
		ooit_wiet			
		1,00 No	2,00 Yes		
Step 1	ooit_wiet	1,00 No	1014	7	99,3
		2,00 Yes	202	8	3,8
	Overall Percentage				83,0

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,517	,056	85,355	1	,000	1,676
	What is your gender?	-,242	,159	2,321	1	,128	,785
	migratie_respondent	-,690	,752	,842	1	,359	,502
	migratie_vader	,379	,452	,702	1	,402	1,460
	migratie_moeder	,336	,486	,476	1	,490	1,399
	Constant	-9,225	1,119	68,000	1	,000	,000

a. Variable(s) entered on step 1: What is your age?, What is your gender?, migratie\_respondent, migratie\_vader, migratie\_moeder.

## Block 2: Method = Enter

### Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	4,324	1	,038
	Block	4,324	1	,038
	Model	108,328	6	,000

### Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	1016,376 <sup>a</sup>	,084	,141

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	14,979	8	,060

### Contingency Table for Hosmer and Lemeshow Test

		ooit_wiet = 1,00 No		ooit_wiet = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	117	114,068	2	4,932	119
	2	118	116,638	6	7,362	124
	3	115	112,525	7	9,475	122
	4	119	114,262	8	12,738	127
	5	96	103,087	22	14,913	118
	6	102	105,326	23	19,674	125
	7	99	101,293	27	24,707	126
	8	86	93,046	36	28,954	122
	9	86	87,469	37	35,531	123
	10	83	73,285	42	51,715	125

**Classification Table<sup>a</sup>**

	Observed	Predicted		Percentage Correct	
		1,00 No	2,00 Yes		
Step 1	ooit_wiet	1,00 No	1013	8	99,2
		2,00 Yes	202	8	3,8
Overall Percentage					82,9

a. The cut value is ,500

**Variables in the Equation**

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,513	,056	83,451	1	,000	1,671
	What is your gender?	-,275	,160	2,938	1	,087	,760
	migratie_respondent	-,762	,744	1,051	1	,305	,467
	migratie_vader	,435	,446	,954	1	,329	1,546
	migratie_moeder	,282	,481	,345	1	,557	1,326
	What do you think you are?	,754	,349	4,675	1	,031	2,125
	Constant	-9,848	1,155	72,707	1	,000	,000

a. Variable(s) entered on step 1: What do you think you are?.

### Block 3: Method = Enter

**Omnibus Tests of Model Coefficients**

		Chi-square	df	Sig.
Step 1	Step	97,431	1	,000
	Block	97,431	1	,000
	Model	205,758	7	,000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	918,945 <sup>a</sup>	,154	,257

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

### Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9,233	8	,323

### Contingency Table for Hosmer and Lemeshow Test

		ooit_wiet = 1,00 No		ooit_wiet = 2,00 Yes		Total
		Observed	Expected	Observed	Expected	
Step 1	1	124	120,353	0	3,647	124
	2	120	117,768	3	5,232	123
	3	118	116,053	5	6,947	123
	4	111	114,869	13	9,131	124
	5	110	111,144	13	11,856	123
	6	106	107,205	17	15,795	123
	7	101	102,403	22	20,597	123
	8	91	95,240	32	27,760	123
	9	88	83,004	35	39,996	123
	10	52	52,961	70	69,039	122

### Classification Table<sup>a</sup>

		Predicted		Percentage Correct
		ooit_wiet		
Observed		1,00 No	2,00 Yes	
Step 1	ooit_wiet 1,00 No	996	25	97,6
	ooit_wiet 2,00 Yes	166	44	21,0
Overall Percentage				84,5

a. The cut value is ,500

### Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 <sup>a</sup>	What is your age?	,433	,060	52,118	1	,000	1,542
	What is your gender?	-,162	,170	,901	1	,343	,851
	migratie_respondent	-,927	,732	1,604	1	,205	,396
	migratie_vader	,173	,457	,144	1	,704	1,189
	migratie_moeder	,556	,493	1,269	1	,260	1,743
	What do you think you are?	,737	,366	4,045	1	,044	2,089
	Schaal_uitgaan	,648	,069	88,791	1	,000	1,913
	Constant	-9,959	1,213	67,439	1	,000	,000

a. Variable(s) entered on step 1: Schaal\_uitgaan.

## Bijlage 3 Assumptiecontrole, uitbijters en multicollineariteit

### Assumptiecontrole

Regressieanalyses moeten voldoen aan bepaalde assumpties. Voor alle modellen in de analyse geldt dat er sprake moet zijn van onafhankelijke waarnemingen, dus voor zowel de logistische als de lineaire regressieanalyse. Dit houdt in dat er geen samenhang bestaat tussen de geobserveerde scores van de respondenten. In het geval van mijn steekproef zijn er vier middelbare scholen benaderd, binnen deze scholen zijn alle scholieren benaderd om de vragenlijst in te vullen. Scholieren binnen een bepaalde school of klas zijn niet volledig onafhankelijk van elkaar. Zo worden ze bijvoorbeeld beïnvloed door dezelfde docent. Deze assumptie wordt dus geschonden. Een schending van deze assumptie zorgt ervoor dat de geschatte standaardfouten kleiner zijn dan ze in werkelijkheid zijn, omdat de steekproefgrootte overschat wordt. Hierdoor lijkt het alsof de regressiecoëfficiënten zuiver worden geschat, maar dit is niet het geval. De uitkomsten van de toetsen zijn hierdoor ook onjuist en dit kan invloed hebben op de generaliseerbaarheid van de uitkomsten, oftewel ik kan niet met zekerheid zeggen dat de gevonden effecten gelden voor de gehele populatie van Nederlandse jongeren.

Voor model 3, oftewel de lineaire regressieanalyse, kijk ik ook naar lineariteit, homoscedasticiteit en een normale verdeling van de residuen. Te beginnen met lineariteit. Deze assumptie houdt in dat er een lineair verband moet zijn tussen de afhankelijke variabele en de onafhankelijke variabele(n). Er wordt aan de assumptie voldaan als er geen systematische afwijkingen van de nullijn zijn en voor iedere waarde van de onafhankelijke variabele het gemiddelde van de residuen 0 is. Een niet-lineair verband kan leiden tot verkeerde conclusies als gevolg van een slechte fit van het model, aangezien een lineair model minder goede schattingen kan doen voor een niet-lineair verband. Ten eerste kijk ik naar het residual plot, zie Figuur 3. Hier is te zien dat er veel waarden boven de nullijn liggen, die ervoor zorgen dat het gemiddelde hoger is dan 0. Bij de partial plots kan hetzelfde geconstateerd worden, zie Figuur 4. Deze assumptie wordt dus geschonden.

Voor homoscedasticiteit kijk ik ook naar de residual plot in Figuur 3. Deze assumptie houdt in dat er een gelijke variantie van de residuen is voor iedere waarde van de onafhankelijke variabele, dus er mogen geen systematische afwijkingen in de spreiding rondom de nullijn zijn. Een schending van deze assumptie leidt tot minder zuivere schattingen, waardoor de toetsen en dus de conclusies onjuist zijn. Ik zie in de residual plot dat er in het midden meer spreiding is dan aan de uiteindes, dus de spreiding is niet constant. De assumptie wordt dus geschonden.

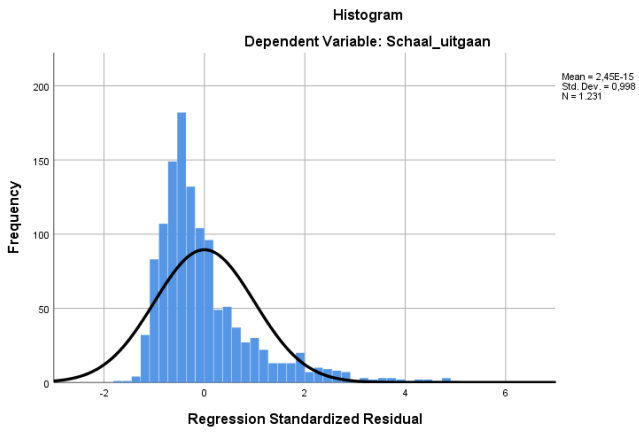
Tot slot de assumptie van een normale verdeling van de residuen. Voor deze assumptie kijk ik naar het histogram en de PP-plot, zie Figuur 1 en 2. In het histogram is een duidelijke piek te zien en een rechtsscheve verdeling. Rechtsscheef betekent dat de verdeling aan de rechterkant langer is. De verdeling is niet symmetrisch en loopt niet zoals de zwarte curve. In de PP-plot is te zien dat de punten rondom de lijn een S-vorm volgen in plaats van een rechte lijn zoals de lijn in het figuur. Een S-vorm betekent dat er een probleem is met de welving van de verdeling, oftewel het is te gepiekt. Deze pieken zijn duidelijk terug te zien in het histogram. Ik concludeer dat er niet wordt voldaan aan de assumptie van een normale verdeling van de residuen. Hierdoor zijn de uitkomsten van de toetsen minder betrouwbaar. De uitgevoerde toetsen gaan uit van een normale verdeling en een verdeling die niet normaal verdeeld is leidt daardoor tot minder betrouwbare uitkomsten van de toetsen en dus conclusies.

Concluderend worden alle assumpties geschonden. De toetsen zijn hierdoor mogelijk onbetrouwbaar en daarom is het goed om strenger te toetsen, om de kans van het ten onrechte aannemen van een hypothese te verkleinen. Daarom zal ik gebruikmaken van een significantieniveau van 0,01 in plaats van 0,05 voor het toetsen van de hellingen in de lineaire regressie. Op die manier zal er minder snel onterecht een hypothese verworpen worden. De logistische regressies gaan niet uit van lineariteit, homoscedasticiteit en een normale verdeling van de residuen, waardoor deze analyses robuuster zijn tegen schendingen van deze assumpties. Daarom heb ik ervoor gekozen om geen lager significantieniveau te hanteren voor de logistische regressie.

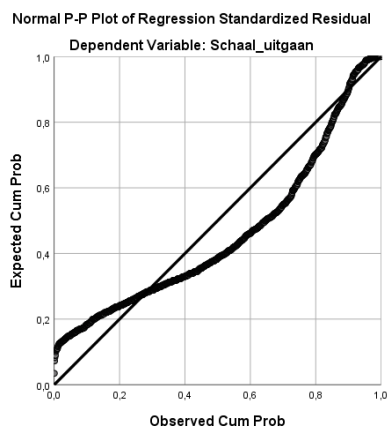
## Output van de assumptiecontrole

\*Opvragen van een residual plot, histogram en PP-plot van de residuen voor de lineaire regressie analyse\*

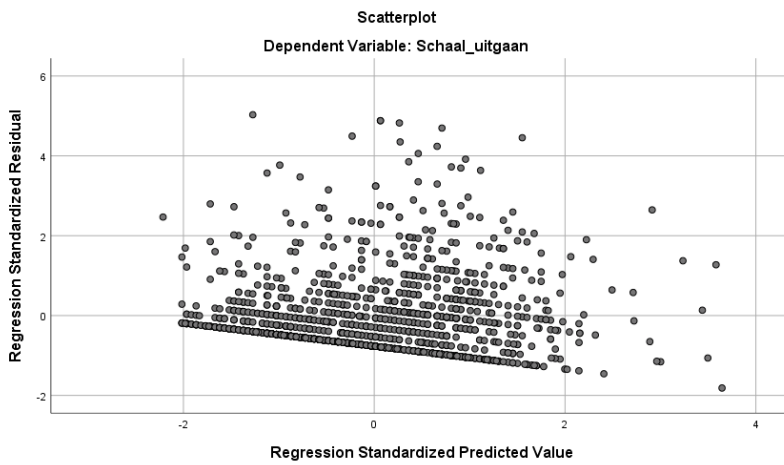
```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Schaal_uitgaan
  /METHOD=ENTER gender_final age_w1 migratie_respondent migratie_vader mi-
gratie_moeder NEWsex
  /PARTIALPLOT ALL
  /SCATTERPLOT=(*ZRESID ,*ZPRED)
  /RESIDUALS HISTOGRAM(ZRESID) NORMPROB(ZRESID)
```



*Figuur 1: Histogram van de residuen*



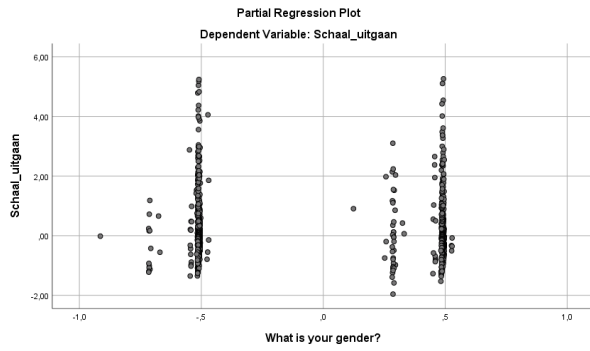
*Figuur 2: PP-plot van de residuen*



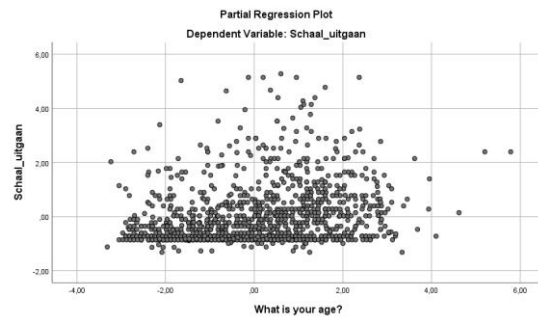
*Figuur 3: Residual plot*



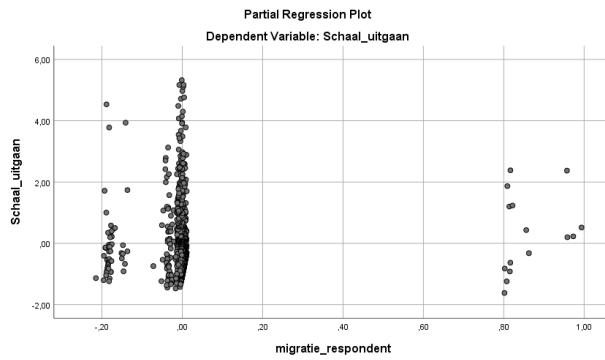
A



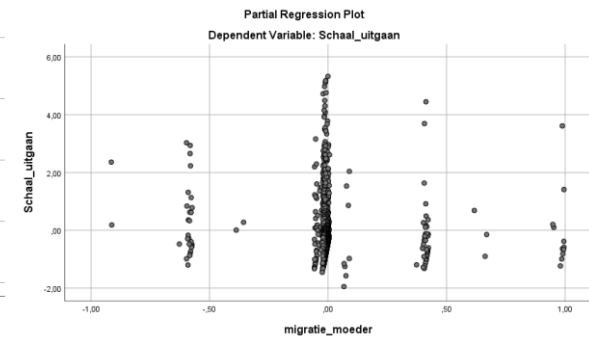
B



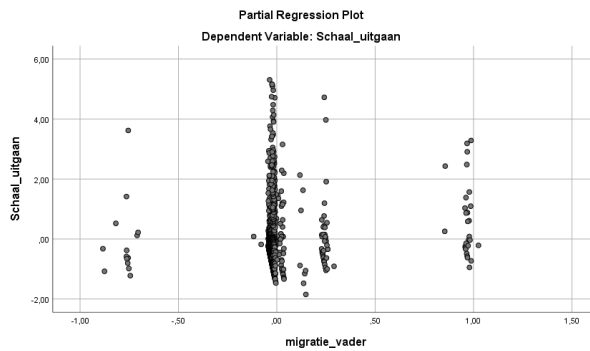
C



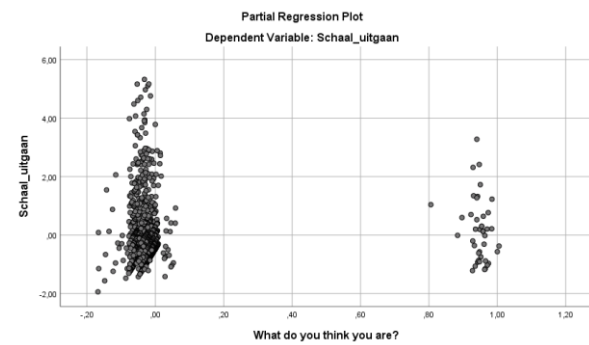
D



E



F



Figuur 4: Partial plots van uitgangsgedrag met de onafhankelijke variabelen. A) geslacht, B) leeftijd, C) migratieachtergrond van de respondent, D) migratieachtergrond van de moeder van de respondent, E) migratieachtergrond van de vader van de respondent en F) seksuele oriëntatie

## Invloedrijke punten en uitbijters

Voor de invloedrijke punten en uitbijters van de lineaire regressie op uitgaansgedrag kijk ik naar de gestandaardiseerde residuen, Cook's distance, leverage en DFFIT. Voor het constateren van grote residuen heb ik gekeken naar de residual plot in Figuur 3. Een waarde wordt gezien als een uitbijter als het een hoger residu heeft dan 3 of een lager residu dan -3. Een hoog residu betekent dat er een groot verschil zit tussen de geobserveerde waarde en de geschatte waarde en dat het model dus niet goed is in het schatten van waarden. In de residual plot is te zien dat er een aantal observaties een residu hoger dan 3 hebben; volgens het residu zijn 23 observaties een uitbijter.

Een volgende uitbijtermaat die onderzocht kan worden is de leverage. De leverage onderzoekt of er mogelijk uitbijters in de x-richting zijn. De leverage waarde onderzoekt hoe ver een bepaald punt van het gemiddelde van de onafhankelijke variabele(n) aflight. Een hoge leverage waarde stelt dat het punt veel invloed heeft op de geschatte hellingen, omdat het ver afwijkt. Er is mogelijk sprake van een probleem als de leverage waarde hoger is dan  $2 \times p/n$  of  $3 \times p/n$ .  $p$  betekent het aantal geschatte parameters, oftewel de constante en alle hellingen.  $n$  is het aantal observaties in de dataset, oftewel 1231. Een waarde is een uitbijter in deze analyse als de leverage waarde hoger is dan  $2 \times 7/1231 = 0,01$  of  $3 \times 7/1231 = 0,02$ . Er zijn 130 observaties die boven de eerste grens leggen en 129 observaties die boven de tweede grens liggen. In Tabel 6 staan de meest afwijkende observaties.

Naast de leverage en residuen, kijk ik naar de Cook's distance. Dit is het product van het residu en de leverage, dus als zowel het residu als de leverage hoog is, is de Cook's distance ook hoog. Voor Cook's distance geldt dat een observatie een problematische uitbijter is als het een waarde heeft boven de 1. Verder geldt ook de vuistregel  $4/n$ . Voor deze analyse is dat  $4/1231 = 0,003$ . Als een observatie een Cook's distance heeft die hoger is dan 0,003, is het een mogelijk probleem. Geen enkele observatie heeft een Cook's distance die hoger is dan 1. Er zijn 75 observaties met een Cook's distance die voorbij 0,003 gaat. De meest afwijkende observaties staan in Tabel 1.

Tot slot kijk ik naar de DFFIT. Deze waarde geeft de verandering in de geschatte waarde van de afhankelijke variabele als de observatie in kwestie uit de analyse wordt gehaald. Een hoge positieve of lage negatieve DFFIT betekent dat de geschatte waarde van  $y$  veel verandert als de observatie uit de analyse wordt gehaald. De vuistregel voor de DFFIT die gebruikt wordt is  $2 \times$  de wortel van  $p/n$  of  $3 \times$  de wortel van  $p/n$ . Een observatie is volgens de DFFIT een uitbijter als het hoger (of lager) is dan  $2 \times$  de wortel van  $7/1231 = (-)0,15$  of  $3 \times$  de wortel van  $7/1231 = (-)0,23$ . Er zijn drie observaties die over de eerste grens gaan en geen enkele observatie die over de tweede grens gaat. De meest afwijkende observaties staan in Tabel 1.

Zoals in Tabel 1 te zien is, zijn niet alle observaties voor alle uitbijtermaten een uitbijter. Observatie 1992 heeft bijvoorbeeld een hoge leverage, maar geen hoge Cook's distance. Dit betekent dat het wel ver afwijkt in de x-richting, maar niet in zowel de x- als y-richting. De DFFIT is ook niet groot voor deze observatie. Case 2014 heeft voor alle maten een relatief hoge waarde, dus deze observatie is vergeleken met andere observaties invloedrijk.

*Tabel 1: Diagnostics voor invloedrijke punten van de analyse van uitgaansgedrag*

Case ID	Cooks distance	DFFIT	Leverage
Case 2773	0,01	0,11	0,10
Case 1993	<0,01	-0,02	0,10
Case 2014	0,06	0,22	0,10
Case 2747	0,04	0,16	0,07
Case 1198	0,08	0,14	0,03
Case 1863	0,06	0,10	0,02

Voor de logistische regressieanalyses kijk ik naar de leverage en DFBETA. Te beginnen met de invloedrijke punten en uitbijters van de analyse van alcohol drinken. De eerste uitbijtermaat die bekeken wordt is de leverage. De grenswaarden voor deze analyse zijn:  $2 \times 8 / 1231 = 0,01$  of  $3 \times 8 / 1231 = 0,02$ . Er gaan 121 observaties voorbij de eerste grenswaarde en 92 observaties voorbij de tweede grenswaarde. De 5 meest afwijkende observaties staan in Tabel 2.

De tweede uitbijtermaat is de DFBETA. Deze waarde geeft de verandering in de geschatte regressiecoëfficiënten weer als de observatie uit de analyse wordt gehaald. Deze waarde wordt voor iedere geschatte parameter uitgerekend. Echter, ik kijk enkel naar de DFBETA's van seksuele oriëntatie en uitgaansgedrag. Dit omdat deze variabelen het belangrijkste zijn voor dit onderzoek en centraal staan in de hypothesen. Als er veel invloed wordt uitgeoefend op deze variabelen, kan dit invloed hebben op de conclusies van de hypothesen. De vuistregel voor de DFBETA is  $2 / \text{de wortel van } n$  of  $3 / \text{de wortel van } n$ . Dus  $2 / \text{de wortel van } 1231 = 0,06$  of  $3 / \text{de wortel van } 1231 = 0,09$ . De DFBETA kan zowel positief als negatief zijn, dus een observatie is een uitbijter als het voorbij  $(-)0,06$  of  $(-)0,09$  gaat. Bij seksuele oriëntatie gaan vier observaties voorbij de eerste grens en drie voorbij de tweede. Voor uitgaansgedrag geldt dat een observatie voorbij de eerste grens gaat en geen voorbij de tweede. Ik kijk naar de vijf meest extreme observaties per parameter.

*Tabel 2: Diagnostics voor invloedrijke punten van de analyse van alcohol drinken*

Leverage	DFBETA Seksuele oriëntatie	DFBETA Uitgaansgedrag
0,16 (2354)	-0,56 (1425)	-0,08(1570)
0,16 (3118)	-0,421(2042)	-0,05 (2079)
0,15 (2042)	0,12 (1414)	-0,04 (1193)
0,13 (1014)	0,06 (2377)	-0,03 (2884)
0,12 (3147)	0,05 (1012)	-0,03 (2750)

Ten derde de invloedrijke punten en uitbijters voor de analyse van bingedrinken. Te beginnen met de leverage. De grenswaarden zijn voor deze analyse ook 0,01 en 0,02. Er zijn 117 observaties die voorbij de eerste grenswaarde gaan en 85 observaties voor de tweede grenswaarde. De meest afwijkende observaties staan in Tabel 3.

Voor de DfBeta gelden wederom de grenswaarden (-)0,06 en (-)0,09. De meest afwijkende observaties staan in Tabel 3. Bij seksuele oriëntatie gaan 16 observaties voorbij de eerste grens en zes voorbij de tweede. Bij uitgaansgedrag zijn er geen die voorbij de grenzen gaan.

*Tabel 3: Diagnostics voor invloedrijke punten van de analyse van bingedrinken*

Leverage	DFBETA Seksuele oriëntatie	DFBETA Uitgaansgedrag
0,17 (2192)	-0,11 (1861)	-0,04 (1570)
0,13 (2354)	-0,11 (1137)	-0,03 (2079)
0,12 (3118)	-0,11 (2889)	-0,03 (2127)
0,11 (2661)	-0,10 (1822)	-0,031(2154)
0,115(1014)	-0,09 (1425)	-0,03 (1790)

Ten vierde de analyse van invloedrijke punten en uitbijters voor tabak roken. Wat betreft de leverage wordt er wederom uitgegaan van de grenswaarden 0,01 en 0,02. Er gaan 131 observaties voorbij de eerste grens gegaan en 113 voorbij de tweede grens. De meest afwijkende observaties staan in Tabel 4.

Voor de DFBETA gelden wederom de grenswaarden (-)0,06 en (-)0,09. De meest afwijkende observaties staan in Tabel 4. Er gaan bij seksuele oriëntatie 14 observaties voorbij de eerste grens en drie voorbij de tweede. Bij uitgaansgedrag zijn er geen observaties die voorbij de grenzen gaan. De meest afwijkende observaties staan in Tabel 4.

*Tabel 4: Diagnostics voor invloedrijke punten van de analyse van tabak roken*

Leverage	DFBETA Seksuele oriëntatie	DFBETA Uitgaansgedrag
0,13 (2192)	-0,11(1777)	-0,02 (3053)
0,10 (2354)	-0,11 (1993)	-0,02(2472)
0,10 (3118)	-0,09(1515)	-0,02 (1570)
0,10 (1014)	-0,08(2659)	-0,02 (2079)
0,09 (2661)	-0,08(2203)	-0,01 (2417)

Als laatste een analyse van invloedrijke punten en uitbijters voor wiet roken. Voor de leverage geldt dat er 166 observaties voorbij 0,01 gaan en 93 voorbij de 0,02 gaan. De meest afwijkende observaties zijn weergegeven in Tabel 5.

Voor de DFBETA gelden wederom de grenswaarden (-)0,06 en (-)0,09. De meest afwijkende observaties staan in Tabel 5. Bij seksuele oriëntatie gaan 14 observaties voorbij de eerste grens en vijf voorbij de tweede. Er gaan geen observaties voorbij de grenzen bij uitgaansgedrag. De meest afwijkende observaties staan in Tabel 5.

*Tabel 5: Diagnostics voor invloedrijke punten van de analyse van wiet roken*

Leverage	DFBETA Seksuele oriëntatie	DFBETA Uitgaansgedrag
0,17 (2773)	-0,10 (1777)	-0,02 (1849)
0,13 (2030)	0,11(1012)	-0,02 (3053)
0,13 (1993)	0,11 (1486)	-0,01 (1570)
0,12 (3011)	0,10(1313)	-0,01 (2472)
0,12 (2014)	0,09 (2922)	-0,01 (2850)

## Syntax voor het controleren van invloedrijke punten en uitbijters

\*Opslaan van de cook's distance, leverage, DFFIT en de gestandaardiseerde residuen voor de lineaire regressie.\*

```
REGRESSION
  /MISSING LISTWISE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Schaal_uitgaan
  /METHOD=ENTER migratie_respondent migratie_vader migratie_moeder gen-
der_final age_w1 NEWsex
  /SAVE COOK LEVER ZRESID DFFIT.
```

\*Opslaan van de leverage en DFBETA voor de logistische regressies.\*  
\*Alcohol drinken\*

```
LOGISTIC REGRESSION VARIABLES ooit_alcohol
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
gratie_moeder /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
  /SAVE LEVER DFBETA.
```

\*Bingedrinken\*

```
LOGISTIC REGRESSION VARIABLES ooit_bingedrink
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
  /SAVE LEVER DFBETA.
```

\*Tabak roken\*

```
LOGISTIC REGRESSION VARIABLES ooit_tabak
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
  /SAVE LEVER DFBETA.
```

\*Wiet roken\*

```
LOGISTIC REGRESSION VARIABLES ooit_wiet
  /METHOD=ENTER age_w1 gender_final migratie_respondent migratie_vader mi-
gratie_moeder
  /METHOD=ENTER NEWsex
  /METHOD=ENTER Schaal_uitgaan
  /PRINT=GOODFIT
  /SAVE LEVER DFBETA.
```

## Analyses opnieuw uitvoeren zonder uitbijters

Iedere analyse heb ik opnieuw gedaan nadat de uitbijters per analyse verwijderd zijn. Voor iedere analyse bespreek ik wat de meest opmerkelijke veranderingen zijn.

Ten eerste de analyse van uitgaansgedrag. De meeste hellingen zijn weinig veranderd. De proportie verklaarde variantie is met 0,01 gestegen voor de  $R^2$ -adjusted, maar de  $R^2$  niet. Het model is dus amper beter in het verklaren van de variantie in uitgaansgedrag ( $R^2_{\text{met uitbijters}} = 0,08$ ;  $R^2_{\text{zonder uitbijters}} = 0,08$ ;  $R^2\text{-adjusted}_{\text{met uitbijters}} = 0,07$ ;  $R^2\text{-adjusted}_{\text{zonder uitbijters}} = 0,08$ ). Ook heb ik de assumpties nog een keer gecontroleerd door naar de residual plot, het histogram en de pp-plot van de residuen te kijken. Hieruit concludeer ik dat de assumpties nog steeds geschonden zijn.

Ten tweede de analyse van alcohol drinken zonder de uitbijters. Enkel model 2 heeft een betere fit gekregen in termen van de Hosmer-Lemeshowtoets, omdat de p-waarde gestegen is. Model 1 en 4 hebben een minder goede fit, want daar is de p-waarde gedaald door het verwijderen van de uitbijters (*Hosmer-Lemeshow*<sub>met uitbijters model 1</sub> = 9,32;  $p = 0,32$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 1</sub> = 9,75;  $p = 0,28$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 2</sub> = 5,62;  $p = 0,68$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 2</sub> = 4,78;  $p = 0,78$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 4</sub> = 3,37;  $p = 0,91$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 4</sub> = 5,69;  $p = 0,68$ ). Het blijkt dus dat voor model 1 en 4 de uitbijters wel zorgden voor een betere fit en dus het maken van betere schattingen. De meeste hellingen zijn weinig veranderd. Echter, de helling van seksuele oriëntatie is sterk toegenomen, in zowel model 2 als in model 4 ( $b_{\text{met uitbijters model 2}} = 1,64$ ;  $b_{\text{zonder uitbijters model 2}} = 19,52$ ;  $b_{\text{met uitbijters model 4}} = 1,59$ ;  $b_{\text{zonder uitbijters model 4}} = 19,22$ ). De verwijderde uitbijters hebben dus een grote invloed op het effect van seksuele oriëntatie op alcohol drinken.

Ten derde de analyse van bingedrinken zonder de uitbijters. Voor deze analyse zijn er geen opmerkelijke veranderingen; de hellingen zijn amper veranderd. De p-waardes voor de Hosmer-Lemeshow toets zijn gedaald, dus de fit van de modellen is minder goed door het verwijderen van de uitbijters (*Hosmer-Lemeshow*<sub>met uitbijters model 1</sub> = 5,47;  $p = 0,71$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 1</sub> = 8,88;  $p = 0,35$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 2</sub> = 5,47;  $p = 0,71$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 2</sub> = 9,94;  $p = 0,27$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 4</sub> = 8,13;  $p = 0,42$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 4</sub> = 8,22;  $p = 0,41$ ). De uitbijters hadden dus een positieve invloed op de fit en de accuraatheid van de schattingen van de modellen.

Ten vierde de analyse van tabak roken zonder uitbijters. Voor deze analyse zijn er ook geen opmerkelijke veranderingen; de hellingen zijn amper veranderd. De p-waarden voor de Hosmer-Lemeshow zijn voor alle modellen gestegen, dus de fit van de modellen is verbeterd door het verwijderen van de uitbijters (*Hosmer-Lemeshow*<sub>met uitbijters model 1</sub> = 16,01;  $p = 0,04$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 1</sub> = 12,79;  $p = 0,12$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 2</sub> = 16,25;  $p = 0,04$ ;

*Hosmer-Lemeshow*<sub>zonder uitbijters model 2</sub> = 14,56;  $p = 0,07$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 4</sub> = 7,69;  $p = 0,47$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 4</sub> = 5,59;  $p = 0,69$ ). De uitbijters leken dus een negatieve invloed te hebben op de fit van de modellen, waardoor ze minder accuraat konden schatten.

Tot slot de analyse van wiet roken zonder uitbijters. Model 1 en 2 hebben een betere fit zonder uitbijters, model 4 niet. Dit omdat de p-waarden voor de eerste twee modellen is gestegen en voor model 4 is gedaald als de uitbijters uit de analyse zijn (*Hosmer-Lemeshow*<sub>met uitbijters model 1</sub> = 16,70;  $p = 0,03$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 1</sub> = 12,18;  $p = 0,14$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 2</sub> = 14,98;  $p = 0,06$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 2</sub> = 10,20;  $p = 0,25$ ; *Hosmer-Lemeshow*<sub>met uitbijters model 4</sub> = 9,23;  $p = 0,32$ ; *Hosmer-Lemeshow*<sub>zonder uitbijters model 4</sub> = 11,62;  $p = 0,17$ ). Verder valt het op dat het effect van de migratieachtergrond van de vader van de respondent van positief naar negatief gaat in model 4 ( $b_{\text{met uitbijters model 4}} = 0,17$ ;  $b_{\text{zonder uitbijters model 4}} = -0,14$ ).

Concluderend verandert er voor de meeste analyses niet veel, behalve voor de helling van seksuele oriëntatie voor alcohol drinken en de helling van de migratieachtergrond van de vader van de respondent in model 4 van wiet roken. Het wisselt per model of ze een betere of slechtere fit krijgen, dus de uitbijters zorgen niet altijd voor een minder accurate schatting en een slechte fit. De uitbijters die de veranderingen veroorzaken zijn ondanks hun invloed wel geldige observaties en daarom worden ze wel in de analyse gehouden.

## Syntax voor het uitfilteren van de invloedrijke punten en uitbijters

\*Uitbijters eruit filteren voor uitgaansgedrag

```
COMPUTE filter_$ = (ID <> 2773 & ID <> 1993 & ID <> 2014 & ID <> 2747 & ID
<> 1198 & ID <> 1863).
VARIABLE LABELS filter_$ 'ID <> 2773 & ID <> 1993 & ID <> 2014 & ID <> 2747
& ID <> 1198 & ID <> 1863 (FILTER)'.
VALUE LABELS filter_$ 0 'Selected' 1 'Not Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

\*Uitbijters eruit filteren voor alcohol drinken\*

```
COMPUTE filter_$ = (ID <> 2354 & ID <> 3118 & ID <> 2042 & ID <> 1014 & ID
<> 3147 & ID <> 1425 & ID <> 1414 & ID <> 2377 & ID <> 1012 & ID <> 1570 &
ID <> 2079 & ID <> 1193 & ID <> 2884 & ID <> 2750).
VARIABLE LABELS filter_$ 'ID <> 2354 & ID <> 3118 & ID <> 2042 & ID <> 1014
& ID <> 3147 & ID <> 1425 & ID <> 1414 & ID <> 2377 & ID <> 1012 & ID <>
1570 & ID <> 2079 & ID <> 1193 & ID <> 2884 & ID <> 2750 (FILTER)'.
VALUE LABELS filter_$ 0 'Selected' 1 'Not Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```



\*Uitbijters eruit filteren voor bingedrinken\*

```
COMPUTE filter_$ = (ID <> 2192 & ID <> 2354 & ID <> 3118 & ID <> 2661 & ID
<> 1014 & ID <> 1861 & ID <> 1137 & ID <> 2889 & ID <> 1822 & ID <> 1425 &
ID <> 1570 & ID <> 2079 & ID <> 2127 & ID <> 2154 & ID <> 1790).
VARIABLE LABELS filter_$ 'ID <> 2192 & ID <> 2354 & ID <> 3118 & ID <> 2661
& ID <> 1014 & ID <> 1861 & ID <> 1137 & ID <> 2889 & ID <> 1822 & ID <>
1425 & ID <> 1570 & ID <> 2079 & ID <> 2127 & ID <> 2154 & ID <> 1790 (FIL-
TER)'.
VALUE LABELS filter_$ 0 'Selected' 1 'Not Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

\*Uitbijters eruit filteren voor tabak roken\*

```
COMPUTE filter_$ = (ID 2192 & ID 2354 & ID 3118 & ID 2661 & ID 1014 &
ID 1777 & ID 1993 & ID 1515 & ID 2659 & ID 2203 & ID 3053 & ID 2472
& ID 1570 & ID 2079 & ID 2417).
VARIABLE LABELS filter_$ 'ID <> 2192 & ID <> 2354 & ID <> 3118 & ID <> 2661
& ID <> 1014 & ID <> 1861 & ID <> 1137 & ID <> 2889 & ID <> 1822 & ID <>
1425 & ID <> 1570 & ID <> 2079 & ID <> 2127 & ID <> 2154 & ID <> 1790 (FIL-
TER)'.
VALUE LABELS filter_$ 0 'Selected' 1 'Not Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

\*Uitbijters eruit filteren voor wiet roken\*

```
COMPUTE filter_$ = (ID 2773 & ID 2030 & ID 1993 & ID 3011 & ID 2014 &
ID 1777 & ID 1012 & ID 1486 & ID 1313 & ID 2922 & ID 1849 & ID 3053
& ID 1570 & ID 2472 & ID 2850).
VARIABLE LABELS filter_$ 'ID <> 2773 & ID <> 2030 & ID <> 1993 & ID <> 3011
& ID <> 2014 & ID <> 1777 & ID <> 1012 & ID <> 1486 & ID <> 1313 & ID <>
2922 & ID <> 1849 & ID <> 3053 & ID <> 1570 & ID <> 2472 & ID <> 2850 (FIL-
TER)'.
VALUE LABELS filter_$ 0 'Selected' 1 'Not Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
```

## Multicollineariteit

Tot slot heb ik gekeken naar de multicollineariteit. Multicollineariteit betekent dat er een (te) grote samenhang is tussen de onafhankelijke variabelen, oftewel seksuele oriëntatie, uitgaansgedrag, geslacht, leeftijd, migratieachtergrond van de respondent, migratieachtergrond van de vader van de respondent en die van de moeder van de respondent voor model 4 en seksuele oriëntatie, leeftijd, geslacht, migratieachtergrond van de respondent, migratieachtergrond van de vader van de respondent en die van de moeder van de respondent voor model 3. Als er een grote VIF-score uitkomt, betekent dit een hoge multicollineariteit. In deze analyses wordt een VIF-score als hoog gezien, als het hoger is dan vier. Een hoge multicollineariteit zorgt voor grotere standaardfouten en dus minder zuivere schattingen van parameters. Daarnaast betekent hoge multicollineariteit dat er veel overlap zit tussen één of meerdere onafhankelijke variabelen, waardoor ze dezelfde variantie verklaren in de afhankelijke variabele. Dit zorgt ervoor dat de afzonderlijke onafhankelijke variabelen bij toevoeging weinig toevoegen aan de fit van het model. In Tabel 3 tot en met 7 staan de VIF-scores voor de eindmodellen gegeven.

In Tabel 3 is te zien dat er geen VIF-scores hoger zijn dan 4; er is geen sprake van hoge multicollineariteit. De hoogste VIF-score is bij de migratieachtergrond van de moeder van de respondent; deze variabele heeft dus de meeste samenhang met de andere variabelen. Er is geen sprake van multicollineariteit voor de lineaire regressie.

In Tabel 4 tot en met 7 zijn de VIF-scores hetzelfde, aangezien het om dezelfde variabelen gaat. In deze tabellen zijn geen hoge VIF-scores te zien. Wederom is de VIF-score van de migratieachtergrond van de moeder van de respondent het hoogst. Er kan geconcludeerd worden dat er geen sprake is van een hoge multicollineariteit en de schattingen van parameters niet veel beïnvloedt worden door te hoge samenhang tussen de onafhankelijke variabelen.

## Syntax voor het berekenen van de VIF-scores ter controle van multicollineariteit

\*Opvragen VIF-scores voor iedere analyse. Ik bereken het voor alleen alcohol drinken, omdat elke analyse van middelengebruik dezelfde variabelen in het model heeft en de VIF-scores dus hetzelfde zijn\*

\*Uitgaansgedrag\*

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT Schaal_uitgaan
  /METHOD=ENTER gender_final age_w1 migratie_respondent migratie_vader mi-
  gratie_moeder NEWsex
```

		Coefficients <sup>a</sup>					Collinearity Statistics	
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Tolerance	VIF
		B	Std. Error	Beta				
1	(Constant)	-,886	,461		-1,921	,055		
	gender_final What is your gender?	-,129	,058	-,061	-2,208	,027	,993	1,007
	age_w1 What is your age?	,176	,019	,261	9,390	,000	,991	1,009
	migratie_respondent	-,097	,346	-,008	-,280	,780	,866	1,154
	migratie_moeder	-,329	,192	-,068	-1,715	,087	,488	2,051
	migratie_vader	,491	,171	,111	2,877	,004	,514	1,944
	NEWsex What do you think you are?	,142	,173	,023	,821	,412	,986	1,014

a. Dependent Variable: Schaal\_uitgaan

\*Alcohol drinken\*

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT ooit_alcohol
  /METHOD=ENTER gender_final age_w1 migratie_respondent migratie_vader mi-
  gratie_moeder NEWsex Schaal_uitgaan
```

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	,669	,179		3,740	,000		
	gender_final What is your gender?	-,017	,023	-,020	-,733	,464	,989	1,011
	age_w1 What is your age?	,067	,008	,246	8,943	,000	,924	1,083
	migratie_respondent	,211	,134	,045	1,579	,115	,866	1,154
	migratie_moeder	-,230	,075	-,117	-3,085	,002	,486	2,056
	migratie_vader	-,130	,066	-,073	-1,969	,049	,511	1,957
	NEWsex What do you think you are?	,131	,067	,052	1,960	,050	,985	1,015
	Schaal_uitgaan On average, how much time do you spend on going out?	,079	,011	,194	7,054	,000	,923	1,083

a. Dependent Variable: ooit\_alcohol

## Bijlage 4: gebruik van AI-software

In dit onderzoek heb ik weinig gebruik gemaakt van AI. De enige taken waarvoor ik AI heb gebruikt, is het controleren van interpunctie en grammatica bij twijfel over de formulering van een zin. Bij ieder hoofdstuk waren er wel eens twijfels, dus voor ieder hoofdstuk is er minstens één keer AI ingeschakeld. Hieronder twee voorbeelden van vragen die gesteld zijn aan ChatGPT, inclusief het antwoord van ChatGPT.

### **Voorbeeld 1:**

*Eenmalig gebruik van een middel is voldoende om in dit onderzoek in de categorie 'ja' te worden geplaatst bij de variabelen van middelengebruik. Dus zolang respondenten niet aangeven dat ze nooit een middel hebben gebruikt terwijl dat wel het geval is, is het geen probleem. Is deze zin punctueel en grammaticaal correct?*

*ChatGPT:*

*Ja, de zin is zowel punctueel als grammaticaal correct. Het is helder en goed geformuleerd.*

### **Voorbeeld 2:**

*Onder mannen was er ook een verschil; 22% van de heteroseksuele mannen tegenover 29% van de biseksuele mannen en 55% van de homoseksuele mannen. Is de puntkomma correct?*

*ChatGPT:*

*Ja, de puntkomma is correct gebruikt in deze zin. Het scheidt de verschillende delen van de zin die gerelateerd zijn aan de percentages onder heteroseksuele mannen, biseksuele mannen en homoseksuele mannen.*

De accuraatheid van de output heb ik iedere keer op vergelijkbare wijze gewaarborgd. Ter illustratie van hoe ik dit gedaan heb, bespreek ik het tweede voorbeeld.

Ik twijfelde over of de puntkomma correct geplaatst was. Daarom heb ik de schrijfsyllabus geraadpleegd. De schrijfsyllabus is een handleiding over het schrijven van een sociaal wetenschappelijk onderzoeksartikel. In deze syllabus staat een specifieke paragraaf over puntkomma's; deze paragraaf heb ik geraadpleegd. Echter, na de raadpleging van de schrijfsyllabus twijfelde ik nog steeds. Deze twijfel bleef ook na raadpleging van het internet, maar de zin leek correct geformuleerd te zijn op basis van de gevonden regels en informatie. Ondanks dat heb ik AI ingeschakeld om bevestiging te vragen. Het bleek voldoende om de regels toe te passen die ik had gevonden in de schrijfsyllabus en op het internet, maar het was fijn om bevestiging te krijgen dat de

regels correct toegepast waren. In dit specifieke geval heb ik zowel het internet als de schrijfsyllabus geraadpleegd, maar in sommige gevallen heb ik enkel het internet gebruikt of de schrijfsyllabus. Zo heb ik enkel naar de regels over komma's in de schrijfsyllabus gekeken voor het eerste voorbeeld.

Door eerst zelf onderzoek te doen naar grammatica- en interpunctieregels, kon ik een indicatie verkrijgen of de geformuleerde zinnen correct waren. De verwachting dat de zinnen correct geformuleerd waren in combinatie met kennis van de regels van grammatica en interpunctie, zorgden ervoor dat de accuraatheid van de output gewaarborgd konden worden. Door eerst zelf onderzoek te doen, kon ik ook mijn academische integriteit waarborgen, aangezien ik eerst zelf informatie heb opgezocht en pas als laatste stap AI ingeschakeld heb ter bevestiging. Daarnaast heeft eerst zelf onderzoek doen ervoor gezorgd dat ik niet teveel cognitieve taken heb gedelegeerd naar AI; ik heb eerst zelf nagedacht en geprobeerd de zin goed te laten lopen. Pas in gevallen van twijfel heb ik AI ingeschakeld.

Ik heb niet aan ChatGPT gevraagd om stukken tekst te generen, enkel om zinnen te controleren. Daarom is het ook niet nodig geweest om zinnen van de output te parafraseren om mijn academische integriteit te waarborgen.