



**rijksuniversiteit
 groningen**

**faculteit gedrags- en
 maatschappijwetenschappen**

**Nieuwsgebruik en Politieke Vertegenwoordiging:
 De invloed van sociale en traditionele media op het gevoel van vertegenwoordiging**

Redmar Huijgen – S5066638

r.v.huijgen@student.rug.nl

Faculteit Gedrags- en Maatschappijwetenschappen, Rijksuniversiteit Groningen

SOBA313A.2024-2025.1: Bachelorwerkstuk

M. Roos, MSc

3 juni 2025

Abstract

De democratie valt of staat bij het vertrouwen van burgers in hun politieke vertegenwoordiging. Een grote groep mensen voelt zich echter niet vertegenwoordigd en vindt dat er niet genoeg wordt opgekomen voor de belangen van mensen zoals zij (Miltenburg et al., 2024). De belangen van burgers en de problemen die zij belangrijk vinden, hangen erg samen met de media aandacht die deze onderwerpen krijgen (Den Ridder et al., 2023). Vanwege het toenemende gebruik van sociale media als nieuwsbron en het afnemende gebruik van traditionele media als nieuwsbron (Tump, 2024), is het belangrijk om te begrijpen hoe en of nieuwsconsumptie via verschillende mediavormen samenhangt met het gevoel van politieke vertegenwoordiging. In dit onderzoek is onderzocht of de manier waarop mensen het nieuws volgen, op sociale media of traditionele media, invloed heeft op het gevoel van politieke vertegenwoordiging en of het volgen van nieuws op de ene mediavorm invloed heeft op de relatie tussen de andere mediavorm en politieke vertegenwoordiging. Hierbij werd verwacht dat het volgen van nieuws op sociale media een negatieve invloed zou hebben op het gevoel van politieke vertegenwoordiging en het volgen van nieuws op traditionele media een positieve invloed. Daarnaast werd verwacht dat de positieve invloed van het volgen van nieuws op traditionele media de negatieve invloed van het volgen van nieuws op sociale media zou verzwakken. Dit werd onderzocht met een hiërarchische logistische regressie met data uit 2023 van het LISS panel. De resultaten bevestigden de verwachtingen dat het volgen van nieuws op sociale media een negatieve invloed heeft op het gevoel van politieke vertegenwoordiging en het volgen van nieuws op traditionele media een positieve invloed. De verwachting dat de positieve invloed van het volgen van nieuws op traditionele media de negatieve invloed van het volgen van nieuws op sociale media zou verzwakken werd niet ondersteund. Het onderzoek biedt echter wel een aantal aanknopingspunten voor vervolgonderzoek, zoals het meten van frequentie en intensiteit van nieuwsconsumptie. Daarnaast kunnen de conclusies meegenomen worden in het ontwerp van wet- en regelgeving en burgerschapsonderwijs ten aanzien van media, alsook in opvoedkundige overwegingen.

Inhoudsopgave

Inleiding	1
Theoretisch kader	4
Het volgen van nieuws op sociale media	4
Het volgen van nieuws op Traditionele media	5
Het volgen van nieuws op sociale media en traditionele media	6
Controle variabelen	7
Methoden	9
Beschrijving databestand en onderzoeksdesign	9
Operationalisaties	10
Je wel of niet vertegenwoordigd voelen door de politiek	10
Het volgen van nieuws op sociale media	11
Het volgen van nieuws op traditionele media	11
Opleidingsniveau	11
Leeftijd	11
Geslacht	11
De analyse-opzet	12
Resultaten	13
Beschrijvende analyses	13
Modelevaluatie	17
Hypothesetoetsing	18
Conclusie en discussie	21
Literatuur	24
Bijlage 1	29
Bijlage 2	55
Bijlage 3	73

Inleiding

De democratie valt of staat bij het vertrouwen van burgers dat hun belangen door de vertegenwoordigers die zij kiezen behartigd worden. In Nederland voelen verschillende groepen zich echter niet gehoord door de politiek; jongeren (Akagündüz, 2025), vrouwen (Universiteit van Amsterdam, 2023), mensen met een migratieachtergrond (Universiteit van Amsterdam, 2024) en mensen wonend buiten de randstad (Van Den Heuvel, 2024). Uit een rapport van Miltenburg et al. (2024) blijkt zelfs dat 57% van alle Nederlanders vindt dat er niet genoeg opgekomen wordt voor belangen van mensen zoals zij. Daarnaast geeft 47% aan te denken dat ze geen enkele invloed hebben op de politiek. Dit is een probleem, want volgens Dagevos en Vermeulen (2024) is het gevoel vertegenwoordigd te worden door de politiek erg belangrijk voor het functioneren van een representatieve democratie. Het vertrouwen in de politiek en de legitimiteit van de overheid kunnen onder druk komen te staan als mensen zich niet vertegenwoordigd voelen.

Volgens Linde en Peters (2020) zijn mensen eerder geneigd het politieke systeem positief te beoordelen als het beleid in lijn is met hun belangen. Een overheid die handelt in overeenstemming met de belangen van een meerderheid van de burgers kan rekenen op publieke steun en wordt als legitiem gezien. Legitimiteit van de overheid is belangrijk, omdat een overheid zonder legitimiteit alleen kan handelen in overeenstemming met wat alle burgers willen, dan wel dwang moet gebruiken om haar wil op te leggen (Andeweg & Aarts 2017).

Politieke vertegenwoordiging is dus erg belangrijk voor het functioneren van de representatieve democratie, maar wat is politieke vertegenwoordiging eigenlijk? Pitkin (1967) onderscheidt vier verschillende vormen van politieke vertegenwoordiging en in dit onderzoek staat de vorm 'inhoudelijke vertegenwoordiging' centraal. Bij inhoudelijke vertegenwoordiging gaat het erom dat de acties die de vertegenwoordiger onderneemt in het politieke proces – zoals wet- en regelgeving vaststellen, beleid formuleren en debatteren – de belangen van de burger vertegenwoordigen (Pitkin, 1967). In het artikel van De Mulder (2022) wordt beschreven dat mensen zich vertegenwoordigd voelen als de vertegenwoordiger luistert naar de burger, weet wat de burger wil, acties onderneemt die in het belang zijn van de burger en daarin ook slaagt.

De belangen van burgers en de problemen die zij belangrijk vinden, hangen volgens Den Ridder et al. (2023) erg samen met de media aandacht die deze onderwerpen krijgen. In de huidige samenleving speelt social media een steeds grotere rol in de manier waarop het nieuws wordt gevolgd; in een onderzoek van het Commissariaat voor de Media blijkt dat 78% van de jongeren sociale media gebruiken om het nieuws te volgen (Schut et al., 2024).

Het commissariaat maakt zich zorgen over de toename van het gebruik van sociale media en de daarmee gepaard gaande afname in het gebruik van andere vormen van media (Tump, 2024). De traditionele media heeft namelijk een belangrijke informerende rol, die in een democratie, en met name in een representatieve democratie, cruciaal is. Het is namelijk van groot belang dat burgers goed geïnformeerd zijn over politieke vraagstukken en standpunten van de partijen en politici die zich verkiesbaar stellen, zodat zij daadwerkelijk in staat zijn te stemmen op de volksvertegenwoordigers die hun belangen het best behartigen. Het is maar de vraag of sociale media deze informerende rol adequaat kan overnemen, gezien het aanbod van nieuws op sociale media gebaseerd is op algoritmes en de betrouwbaarheid minder goed gewaarborgd is (Tump, 2024).

Vanwege het toenemende gebruik van sociale media als nieuwsbron en het grote aantal burgers dat aangeeft zich niet vertegenwoordigd te voelen door de politiek, is het belangrijk om te onderzoeken of en hoe deze twee ontwikkelingen met elkaar samenhangen. Daarnaast wordt er gekeken naar de rol van traditionele media als nieuwsbron. Onder traditionele media vallen bijvoorbeeld televisie, radio en de krant. Nieuws op traditionele media wordt over het algemeen geproduceerd door professionele journalisten, daarom vallen nieuwswebsites zoals NU.nl of NOS.nl ook onder deze vorm (Gillis, 2024).

Er zijn al verschillende onderzoeken gedaan naar politieke vertegenwoordiging. Burgers die zich vertegenwoordigd voelen door de politiek hebben het gevoel dat de overheid luistert en handelt in hun belang, oftewel; dat de overheid responsief is (De Mulder, 2022). Deze burgers zijn meer tevreden over de overheid, zien de overheid als legitiem, en tonen vertrouwen in de politiek (Linde & Peters, 2020). Tegelijk wijst verschillend onderzoek uit dat niet elke groep zich even vertegenwoordigd voelt. Zo ervaren mensen met een migratieachtergrond minder inhoudelijke vertegenwoordiging (Dagevos & Vermeulen, 2024). Volgens Miltenburg et al. (2024) voelen ook jongeren, lager opgeleide en mensen met een laag inkomen zich minder vertegenwoordigd en ook zijn vrouwen volgens Mügge et al. (2025) niet voldoende vertegenwoordigd. Ook bij nieuwsconsumptie bestaan duidelijke demografische verschillen. Jonge mensen zijn meer geneigd om nieuws via sociale media tot zich te nemen en ouderen, mannen en hoogopgeleiden consumeren meer nieuws dan jongeren, vrouwen en laagopgeleiden (Klopfenstein Frei et al., 2024; Haugsgjerd & Karlsen, 2024).

Nieuwsmedia hebben de belangrijke rol om burgers te informeren over politieke vraagstukken en de motieven van vertegenwoordigers (Burkens et al., 2017; Cappella & Jamieson, 1997). Ligt in de manier waarop mensen nieuws consumeren een verklaring voor de mate waarin zij zich vertegenwoordigd voelen door de politiek? De relatie tussen politieke vertegenwoordiging en het volgen van nieuws op sociale media en traditionele media is nog onderbelicht in de wetenschappelijke literatuur. In de onderzoeken waarin de relatie tussen

het volgen van nieuws op sociale en traditionele media en politiek vertrouwen en tevredenheid over de overheid wel aan bod komt, wordt er niet gekeken of er een eventueel moderatie-effect kan zijn tussen de twee soorten nieuwsmedia. Dit terwijl mensen uiteraard nieuws kunnen consumeren via zowel sociale als traditionele media. Dit onderzoek draagt daarom bij aan de literatuur door te onderzoeken in hoeverre het gebruik van traditionele media het verband tussen sociale media en politieke vertegenwoordiging kan versterken of afzwakken.

De vraag waar dit onderzoek zich daarom mee bezig houdt, is: *'In welke mate heeft het volgen van nieuws op sociale en traditionele media invloed op het gevoel van politieke vertegenwoordiging, en in hoeverre beïnvloedt het volgen van nieuws op de ene mediavorm het verband tussen nieuwsgebruik op de andere mediavorm en dit gevoel?'*

Theoretisch kader

Het volgen van nieuws op sociale media

Om te onderzoeken wat de relatie tussen het volgen van nieuws op sociale media en politieke vertegenwoordiging is, is het belangrijk om te kijken naar de aard van social-mediagebruik en de aard van het volgen van nieuws op sociale media. Iemand die gebruik maakt van sociale media heeft een grotere kans om zonder actief op zoek te gaan naar nieuws er toch aan blootgesteld te worden. Nieuws is over het algemeen negatief en negatief nieuws wordt vaker gedeeld op sociale media. Daarnaast is het algoritme van sociale media platforms erg bepalend voor de hoeveelheid nieuws die iemand te zien krijgt en welk nieuws iemand te zien krijgt.

Bij zowel traditionele media als bij sociale media is er de mogelijkheid dat je wordt blootgesteld aan nieuws en dus nieuws consumeert ook al ben je niet actief op zoek naar nieuws (Park & Kaye, 2020). Bij traditionele media kan je bijvoorbeeld tijdens het televisie kijken het einde of het begin van een nieuwsuitzending meekrijgen na of voordat het entertainmentprogramma dat je wilt kijken eindigt of begint. Toch is met het brede gebruik van sociale media de kans op incidentele blootstelling aan nieuws groter geworden, omdat gebruikers daar eenvoudig nieuws tegenkomen via gedeelde content in hun netwerk en het verspreiden van nieuws gemakkelijk is door 'like'- en 'share'-functies (Park & Kaye, 2020).

Ook het algoritme wat actief is op sociale media platforms kan voor deze incidentele blootstelling aan nieuws zorgen. De algoritmes verschillen per platform, maar het is duidelijk dat de content die iemand te zien krijgt verschilt per persoon omdat de algoritmes die actief zijn op de verschillende platforms content selecteren die bij jou voorkeuren passen (Nowacki, 2024). Als je een aantal keer wat nieuwsartikelen hebt gelezen op social media platforms krijg je steeds meer nieuws te zien ook al ben je er niet per se naar op zoek. Het algoritme dat actief is op sociale media heeft dus invloed op de hoeveelheid nieuws dat iemand te zien krijgt.

Daarnaast heeft het algoritme invloed op wat voor nieuws je ziet en wat voor nieuws je niet ziet. Als je bijvoorbeeld veel content over klimaatverandering bekijkt of veel accounts volgt die content plaatsen of delen over klimaatverandering dan zal je meer content te zien krijgen die iets met dit onderwerp te maken hebben, terwijl onderwerpen die jou minder interesseren minder vaak geselecteerd worden door het algoritme omdat deze onderwerpen niet bij jouw voorkeuren passen (Nowacki, 2024). Zodoende personaliseren algoritmes op sociale media het nieuws dat mensen te zien krijgen.

Over het algemeen is nieuws negatief en mensen consumeren negatief nieuws vaker en reageren er sterker op (Soroka & McAdams, 2015; Zhang et al., 2024). Dit wordt de

negativity bias genoemd. Maar bij sociale media wordt de blootstelling aan dit negatieve nieuws extra versterkt omdat negatief nieuws vaker wordt gedeeld en zich daardoor verder verspreid dan nieuws dat niet negatief is (Knutson et al., 2024). Daarnaast heeft het nieuws op sociale media vaker een antipolitieke lading en is het vaker kritisch over het huidige politieke systeem (Ceron, 2015). Mensen die het nieuws niet volgen op sociale media krijgen dus minder negatief en kritisch nieuws met een antipolitieke lading te zien over onderwerpen die zij belangrijk vinden.

Dit personaliserende karakter van nieuws op sociale media en dit negatieve sentiment heeft invloed op hoe mensen kijken naar politici en hun beleid. Volgens Cappella en Jamieson (1997) leren burgers over de motieven en het handelen van politici grotendeels via de media. De wijze waarop media politiek nieuws presenteren kan ertoe leiden dat politici en hun beleid voornamelijk worden gezien als in eigenbelang handelend, in plaats van in belang van de burger (Cappella & Jamieson, 1997). Wanneer berichtgeving over politiek, politici en beleid overwegend negatief is en op deze wijze wordt gepresenteerd, kan dit leiden tot een verminderd vertrouwen in politici en het ontstaan van cynisme over hun intenties. Deze 'media/video malaise'-theorie geldt ook bij het volgen van nieuws op sociale media en je wel of niet vertegenwoordigd voelen door de politiek. Het personaliserende karakter van nieuws op sociale media en dit negatieve sentiment zorgt ervoor dat dit cynisme wordt geactiveerd bij onderwerpen die zij belangrijk vinden, waardoor het gevoel ontstaat dat politici niet handelen in de belangen van mensen zoals zij.

Voorgaande leidt tot de hypothese: *Mensen die het nieuws volgen op sociale media hebben een minder grote kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op sociale media.*

Het volgen van nieuws op Traditionele media

In dit onderzoek wordt daarnaast onderzocht wat de invloed is van het volgen van nieuws op traditionele media op het wel of niet vertegenwoordigd voelen door de politiek. De 'media/video malaise'-theorie stelt dat de negatieve aard van de media en dus ook traditionele media het vertrouwen in de politiek negatief beïnvloedt. Tegelijkertijd wijzen verschillende onderzoeken er ook op dat traditionele media een positieve invloed kan hebben op politiek vertrouwen en tevredenheid over de overheid. Ceron (2015) stelt dat het nieuws op traditionele media vaker een positiever beeld schetst dan nieuws op sociale media, waar het voornamelijk negatief is. De traditionele media verspreidt meer positief en systeembevestigend nieuws over het huidige politieke systeem, terwijl het nieuws op sociale media meer ruimte geeft voor antipolitieke perspectieven (Ceron, 2015). Zoals hierboven al uitgelegd, leren burgers over de motieven en het handelen van politici grotendeels via de media en kan de wijze waarop media politiek nieuws presenteert invloed hebben op hoe

politici en hun beleid worden gezien (Cappella & Jamieson, 1997). Omdat mensen, politici en hun beleid beoordelen met behulp van nieuws vanuit de media hebben deze positieve geluiden van traditionele media over de politiek, politici en hun beleid een positieve invloed op het vertrouwen in en de tevredenheid over de overheid en politieke responsiviteit (Ceron, 2015; Ceron & Memoli, 2016).

Tevredenheid over de overheid hangt nauw samen met het gevoel vertegenwoordigd te worden door de politiek. Wanneer de overheid keuzes maakt die niet overeenkomen met de belangen van de burgers, zullen mensen minder tevreden zijn met de overheid (Ceron & Memoli, 2016). Mensen die het nieuws volgen op traditionele media horen meer positieve geluiden over de overheid en de politiek, waardoor ze het gevoel krijgen dat er in de politiek meer naar hen geluisterd wordt en dat de politiek succesvol handelt in naam van hun belangen. Mensen die het nieuws niet volgen op traditionele media krijgen deze positieve geluiden en systeembevestigende geluiden niet of in mindere mate te horen waardoor ze minder het gevoel krijgen dat er in de politiek naar hun geluisterd wordt en succesvol wordt gehandeld in naam van hun belangen.

Dit leidt tot de volgende hypothese: *Mensen die het nieuws volgen op traditionele media hebben een grotere kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op traditionele media.*

Het volgen van nieuws op sociale media en traditionele media

In dit onderzoek wordt daarnaast nog gekeken naar welke rol het volgen van nieuws op traditionele media heeft bij de relatie tussen het volgen van nieuws op sociale media en je wel of niet vertegenwoordigd voelen door de politiek. Mensen die het nieuws volgen op traditionele media horen meer positieve geluiden over de overheid en de politiek, terwijl mensen die het nieuws volgen op sociale media meer negatieve geluiden horen over de overheid en de politiek (Ceron, 2015). In het onderzoek van Besalú and Pont-Sorribes (2021) wordt de geloofwaardigheid van nieuws op sociale media en nieuws op traditionele media besproken. Hieruit blijkt dat mensen het nieuws op traditionele media als meer betrouwbaar zien dan het nieuws op sociale media en dat ze met een meer voorzichtige blik kijken naar het nieuws op sociale media. Daarnaast worden traditionele media vaak gezien als een controlerende macht binnen de democratie, die corruptie aan het licht brengt en politici ter verantwoording roept (Camaj, 2012). Als mensen dus nieuws op sociale media en traditionele media volgen zullen ze het positievere nieuws van de traditionele media meer betrouwbaar vinden dan het nieuws op sociale media. De verwachting is dat deze positieve geluiden over de overheid en de politiek in de traditionele media, de voornamelijk negatieve geluiden die komen vanuit de sociale media verzwakken/nuanceren.

Dit leidt tot de hypothese: *De negatieve invloed van het volgen van nieuws op sociale media op je wel of niet vertegenwoordigd voelen door de politiek wordt verzwakt als iemand het nieuws ook volgt via de traditionele media.*

Controle variabelen

Ten slotte wordt de rol van opleidingsniveau, leeftijd en geslacht onderzocht. Schakel en Van Der Pas (2020) onderzoeken de rol van opleidingsniveau bij politieke vertegenwoordiging. Hoogopgeleiden participeren meer in het politieke proces en hebben daardoor meer invloed op het beleidsmakingsproces (Schakel & Van Der Pas, 2020). Het onderzoek van Haugsgjerd en Karlsen (2024) stelt dat lager opgeleide mensen minder nieuws consumeren wat er bijvoorbeeld voor zou kunnen zorgen dat het mechanisme van de 'media/video malaise'-theorie minder van toepassing is bij lager opgeleide mensen. Het is daarom belangrijk om te controleren voor opleidingsniveau omdat er een verschil kan zitten in de mate waarin hoogopgeleide mensen zich vertegenwoordigd voelen door de politiek en de mate waarin laagopgeleide mensen zich vertegenwoordigd voelen, maar ook in of ze het nieuws volgen op sociale media en/of traditionele media.

Het is ook belangrijk om te controleren voor leeftijd. Stockemer en Sundström (2023) stellen dat jonge mensen ondervertegenwoordigd worden in het politieke proces en dat er veel meer oudere mensen zijn die hieraan deelnemen. Er is daarnaast groeiend bewijs dat jonge politieke vertegenwoordigers jonge mensen inhoudelijk beter vertegenwoordigen dan oudere politieke vertegenwoordigers (Stockemer & Sundström, 2023). Daarnaast wordt in het artikel van Klopfenstein Frei et al. (2024) gesteld dat jongeren nieuws voornamelijk volgen op sociale media en in het rapport van Schut et al. (2024) wordt dit bevestigd. Daarnaast consumeren ze ook minder nieuws (Haugsgjerd & Karlsen, 2024). Omdat jonge mensen zich dus waarschijnlijk minder vertegenwoordigd voelen door de politiek en steeds meer gebruik maken van sociale media als nieuwsbron in plaats van traditionele media is het belangrijk om leeftijd mee te nemen in het model.

Als laatste wordt er gecontroleerd voor geslacht. In Kenworthy en Malami (1999) staat dat vrouwen over het algemeen minder parlementaire zetels hebben dan mannen. In Nederland hadden vrouwen in 1999 ongeveer 30% van de zetels en op dit moment is dat ongeveer 40% (*Vrouwen in De Tweede Kamer*, n.d.). In een Amerikaans onderzoek van Vega en Firestone (1995) staat dat als de aanwezigheid van vrouwen groeit in het congres het dan aannemelijk is dat de inhoudelijke vertegenwoordiging van vrouwen ook toeneemt. De vraag is natuurlijk of dit in Nederland ook het geval is. In het onderzoek van Haugsgjerd en Karlsen (2024) wordt ook gekeken naar gender en nieuwsconsumptie en hierbij geldt dat vrouwen minder nieuws consumeren dan mannen. Omdat er een verschil kan zijn tussen de

mate waarin mannen en vrouwen zich politiek vertegenwoordigd voelen en of ze het nieuws consumeren op sociale media en/of traditionele media, is het belangrijk om voor geslacht te controleren.

Methoden

In dit hoofdstuk worden de concepten die centraal staan in dit onderzoek geoperationaliseerd, de data en de dataverzameling wordt toegelicht en de opzet van de analyse wordt uitgelegd.

Beschrijving databestand en onderzoeksdesign

In dit onderzoek wordt er gebruik gemaakt van de dataset van het LISS¹ panel, een longitudinaal onderzoekproject voor wetenschappelijk, beleidsgericht of sociaal relevant onderzoek (*Methodology - LISS Panel*, 2024). Het LISS Panel heeft in 2007 met behulp van het populatieregister van het CBS (Centraal bureau voor de statistiek) 10.000 willekeurige huishoudens benaderd (*Methodology - LISS Panel*, 2024). Om ervoor te zorgen dat de initiële steekproef een goede representatie is van de Nederlandse bevolking die de Nederlandse taal beheerst, is de steekproef vergeleken met de nationale data (Knoef & de Vos, 2009). Het LISS panel heeft in eerste instantie een brief gestuurd naar de huishoudens en vervolgens gebeld en/of is op bezoek geweest om te vragen of ze mee willen doen aan het panel (*Methodology - LISS Panel*, 2024). Uiteindelijk hebben ongeveer 5.000 huishoudens mee gedaan met het panel, bestaande uit ongeveer 7.500 individuen van 16 jaar en ouder (*How It Works - LISS Panel*, 2023).

Door de jaren heen heeft het panel meerdere aselechte steekproeven uitgevoerd om ervoor te zorgen dat het aantal huishoudens en het aantal individuen ongeveer gelijk zou blijven. Deze aselechte steekproeven werden afgewisseld met gestratificeerde steekproeven om ervoor te zorgen dat het panel een representatieve weergave is van de Nederlandse bevolking die de Nederlandse taal beheerst (*Methodology - LISS Panel*, 2024). Om er voor te zorgen dat iedereen mee kan doen aan het panel wordt apparatuur meegeleverd, zodat mensen zonder een computer ook meegenomen kunnen worden in de steekproeven (*Methodology - LISS Panel*, 2024).

In dit onderzoek wordt data gebruikt uit 2023. In april en mei van dat jaar is eenmalig een vragenlijst over 'Democratie in Nederland' voorgelegd aan 6.552 huishoudleden (Verheijen, 2024). 5.134 huishoudleden hebben gereageerd, 1.418 hebben niet gereageerd. 5.067 hebben de vragenlijst compleet ingevuld. Daarmee komt de respons grotendeels overeen met de gemiddelde maandelijkse respons van het LISS-panel, die tussen de 50% en 80% ligt (*Methodology - LISS Panel*, 2024). Er is een vragenlijst uit de LISS Core studie gebruikt genaamd 'Politiek en waarden'. Deze vragenlijst is in drie delen afgenomen, waarvan één vraag uit deel 1 van de vragenlijst is

¹ LISS staat voor Langlopende Internet Studies voor de Sociale wetenschappen.

gebruikt in dit onderzoek (Elshout, 2024). Dit deel is in december 2023 (herhaalt in januari 2024) voorgelegd aan 7.312 mensen waarvan 6.303 de vragenlijst volledig hebben ingevuld.

Ten slotte is er data uit de vragenlijst van de huishoudbox gebruikt. Dit is een vragenlijst die voorafgaand aan het deelnemen aan het LISS panel moet worden ingevuld door één iemand van het huishouden en deels gaat over de kenmerken van het huishouden en deels over de kenmerken van de individuele huishoudleden (Elshout, 2022).

In de oorspronkelijke dataset zitten 12.059 individuen, maar omdat de vragenlijsten ingevuld zijn door mensen binnen hetzelfde huishouden zijn de waarnemingen niet allemaal onafhankelijk. Om ervoor te zorgen dat de waarnemingen onafhankelijk van elkaar zijn, wordt er willekeurig één iemand van het huishouden geselecteerd, waardoor er uiteindelijk 5.188 respondenten overblijven. Daarnaast zijn respondenten die geen antwoord hebben gegeven op één van de vragen die gebruikt worden in dit onderzoek er ook uitgefilterd. De dataset die wordt gebruikt voor de analyses bestaat zodoende uit 3.207 respondenten², waarin geen missende waarden voorkomen. Hoe deze bovengenoemde filters zijn toegepast is te zien in bijlage 1.

Operationalisaties

Je wel of niet vertegenwoordigd voelen door de politiek

De vraag die in de vragenlijst gesteld werd om de mate waarin iemand zich vertegenwoordigd voelt door de politiek te meten, luidt als volgt: “Vindt u dat er in de politiek genoeg wordt opgekomen voor de belangen van mensen zoals u?” Op basis van de literatuur van Pitkin (1967) is deze vraag gekozen. Deze vraag past goed bij inhoudelijke vertegenwoordiging, waarbij iemand zich vertegenwoordigd voelt door de politiek als hij meent dat zijn belangen zichtbaar zijn in het politieke proces. De variabele was oorspronkelijk een variabele met 6 scores, maar omdat de assumpties van lineaire regressie werden geschonden (bijlage 3) is de variabele gehercodeerd tot een binaire variabele. De scores 1 ‘Veel te weinig’ en 2 ‘Te weinig’ zijn samengevoegd tot 0 ‘Niet’ en de scores 3 ‘Precies genoeg’, 4 ‘Te veel’ en 5 ‘Veel te veel’ zijn samengevoegd tot 1 ‘Wel’. Een geringe hoeveelheid respondenten had een score van 4 of 5 waardoor deze scores zijn samengevoegd met de 3. Omdat de variabele is gehercodeerd tot een binaire variabele is het niet meer de mate waarin iemand zich vertegenwoordigd voelt door de politiek, maar je wel of niet vertegenwoordigd voelen door de politiek. De vier aspecten benoemd in het artikel van De Mulder (2022) die belangrijk zijn voor mensen om zich vertegenwoordigd te voelen door een vertegenwoordiger passen ook niet helemaal bij de antwoordcategorieën ‘Te veel’ en ‘Veel te veel’. Je kunt je namelijk afvragen of het mogelijk is voor een vertegenwoordiger

² De missende waarden worden uitgebreid besproken in bijlage 2.

om té veel te luisteren naar de burgers of té succesvol te handelen in het belang van de burger. De score 6 'Weet ik niet' is veranderd naar een missende waarden omdat deze waarde op zichzelf niks zegt over of iemand zich vertegenwoordigd voelt door de politiek of niet.

Het volgen van nieuws op sociale media

Het volgen van nieuws op sociale media is gemeten met de vraag: "Hoe volgt u het nieuws?" Hierbij was één antwoordoptie 'via social media (Twitter, LinkedIn, YouTube, Instagram, Facebook, etc.)' Iemand die een 0 scoort op deze variabele volgt het nieuws niet op sociale media en iemand met een score van 1 volgt het nieuws wel op sociale media.

Het volgen van nieuws op traditionele media

Het volgen van nieuws op traditionele media is gemeten met de vraag: "Hoe volgt u het nieuws?" Hierbij waren verschillende antwoordopties mogelijk waarvan voor deze variabele drie zijn geselecteerd. Ten eerste 'op televisie en/of radio', ten tweede 'via een gratis internetsite of app; zoals Nu.nl, NOS.nl, RTLnieuws.nl of Nieuws.nl' en ten slotte 'in een betaald dagblad (digitaal, app en/of op papier)'. Deze drie items zijn samengevoegd tot één variabele om weer te geven wie het nieuws volgt op traditionele media. Deze items zijn niet samengevoegd om één concept te vormen, maar vallen allemaal onder traditionele media, waardoor een betrouwbaarheidsanalyse niet nodig is. Als iemand een 1 scoorde op één van de afzonderlijke items heeft deze persoon ook een score van 1 op de variabele 'het volgen van nieuws op traditionele media' en als iemand op geen van de afzonderlijke items een 1 scoorde heeft deze persoon een score van 0 op de variabele 'het volgen van nieuws op traditionele media'.

Opleidingsniveau

De respondenten zijn gevraagd naar hun opleidingsniveau. De variabele heeft 6 mogelijke scores namelijk 1 = 'basisonderwijs', 2 = 'vmbo', 3 = 'havo/vwo', 4 = 'mbo', 5 = 'hbo' en 6 = 'wo'. Hoe hoger de score, hoe hoger het opleidingsniveau.

Leeftijd

In dit onderzoek wordt er ook gecontroleerd voor leeftijd. De leeftijd is gemeten in jaren.

Geslacht

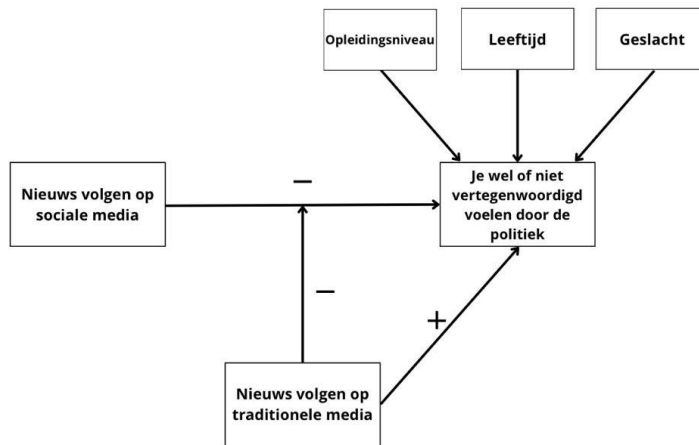
De respondenten zijn gevraagd naar hun geslacht. De oorspronkelijke variabele had drie antwoordmogelijkheden namelijk 1 'man', 2 'vrouw' en 3 'anders'. Slechts een klein aantal respondenten (22) heeft dit antwoord gegeven, en na het willekeurig selecteren van

één iemand per huishouden waren er nog maar negen respondenten met een score van 3 op geslacht. Na het wegfilteren van de missende waarden zou er nog maar één respondent met een score van 3 zijn, waardoor ervoor gekozen is om deze respondent niet mee te nemen in de analyse. De uiteindelijke variabele heeft de scores 0 'man' en 1 'vrouw'.

De analyse-opzet

In dit onderzoek worden vier logistische regressie modellen geschat. In model 1 wordt de uitkomst variabele 'je wel of niet vertegenwoordigd voelen door de politiek' voorspeld door de controle variabelen 'opleidingsniveau', 'leeftijd' en 'geslacht'. In model 2 wordt de variabele 'het volgen van nieuws op sociale media' toegevoegd en wordt de eerste hypothese getoetst namelijk: *Mensen die het nieuws volgen op sociale media hebben een minder grote kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op sociale media*. Daarna wordt in model 3 de variabele 'het volgen van nieuws op traditionele media' toegevoegd waarbij de tweede hypothese wordt getoetst namelijk: *Mensen die het nieuws volgen op traditionele media hebben een grotere kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op traditionele media*. Ten slotte wordt in model 4 de interactie tussen deze laatste twee variabelen toegevoegd. In dit laatste model wordt de derde hypothese getoetst namelijk: *De negatieve invloed van het volgen van nieuws op sociale media op je wel of niet vertegenwoordigd voelen door de politiek wordt verzwakt als iemand het nieuws ook volgt via de traditionele media*.

Daarnaast worden de modelfit, multicollineariteit en de assumpties die horen bij logistische regressie geëvalueerd en wordt modeldiagnostiek uitgevoerd. Daarbij wordt ten eerste gekeken naar de fit van het model door te kijken naar de hosmer-lemeshow toets, de LR-toets en classificatietabellen. Ten tweede wordt er gekeken naar multicollineariteit door te kijken naar de VIF-waarden van alle variabelen in het model. Ten derde wordt er gekeken naar de assumptie onafhankelijke observaties die hoort bij logistische regressie. Het is hierbij belangrijk dat de steekproef aselekt is en dat de respondenten onafhankelijk van elkaar zijn. Ten slotte wordt er gekeken naar uitbijters en invloedrijke punten. Hierbij wordt gekeken naar welke respondenten hoge of extreme waarden hebben op het gebied van *leverage* en *DFBetas*. Er wordt gekeken wat dit voor respondenten zijn en eventueel wordt het volledige model opnieuw geschat, waarbij deze respondenten worden weggelaten. Dan wordt onderzocht of dit een aanzienlijke invloed heeft op de regressiecoëfficiënten en de algehele modelfit en of de conclusies die uit het model worden getrokken veranderen na het verwijderen van deze invloedrijke respondenten.



Figuur 1 Conceptueel model

Resultaten

Beschrijvende analyses

Ten eerst is er aandacht voor de beschrijvende statistieken van de variabelen die zijn meegenomen in de analyse in tabel 1. Bij 'je wel of niet vertegenwoordigd voelen door de

politiek' valt op dat bijna twee derde van de respondenten zich niet vertegenwoordigd voelt terwijl iets meer dan één derde zich wel vertegenwoordigd voelt door de politiek (Niet = 64,76%; Wel = 35,24%). Ongeveer driekwart van de respondenten volgt het nieuws niet op sociale media en bijna een kwart van de respondenten volgt het nieuws wel op sociale media (Nee = 75,93%; Ja = 24,07%). Bij het volgen van nieuws op traditionele media is deze verdeling compleet anders, waarbij bijna iedereen het nieuws volgt op traditionele media en maar een klein deel het nieuws niet op traditionele media volgt (Nee = 5,89%; Ja = 94,11%). Het grootste deel van de respondenten voelt zich dus niet vertegenwoordigd door de politiek, volgt het nieuws niet op sociale media en volgt het nieuws wel op traditionele media.

Opleidingsniveau heeft een links-scheve verdeling waarbij het grootste deel van de respondenten het opleidingsniveau hbo heeft en het gemiddelde valt binnen het opleidingsniveau mbo ($M = 3,96$; $SD = 1,46$). De verdeling van opleidingsniveau is redelijk vergelijkbaar met de verdeling in de populatie³ (Centraal Bureau voor de Statistiek, 2024). Leeftijd heeft ook een links-scheve verdeling waarbij een gemiddeld persoon bijna 58 jaar is ($M = 57,83$; $SD = 17,06$). In de populatie³ is er meer sprake van een tweetoppige verdeling. Daarmee verschilt de verdeling van leeftijd in de populatie dus ook enigszins van de verdeling in de steekproef (Centraal Bureau voor de Statistiek, n.d.-a). De oudere respondenten zijn hierdoor oververtegenwoordigd in de steekproef ten opzichte van de respondenten van 16 tot ongeveer 55 jaar. Hierdoor moet er voorzichtiger gedaan worden met het generaliseren van de resultaten naar de populatie. Ten slotte zijn er ongeveer net zoveel mannelijke als vrouwelijke respondenten, maar er zitten iets meer vrouwen in de steekproef (Man = 49,67%; Vrouw = 50,33%). Deze verdeling is te vergelijken met de verdeling in de populatie³ (Centraal Bureau voor de Statistiek, n.d.-c).

Tabel 1: Beschrijving van de in de analyse opgenomen variabelen: gemiddelde (standaarddeviatie), minimum- en maximumwaarde, Mediaan, 1^e en 3^e kwartiel en totaal aantal respondenten

Variabele	Gemiddelde (Standaarddeviatie) ^a	Minimum	Maximum	Mediaan	1 ^e kwartiel	3 ^e kwartiel	N totaal
Je wel of niet vertegenwoordigd voelen door de politiek (Niet = 0; Wel = 1)	Niet = 64,76%; Wel = 35,24%	0	1				3207

³ Hoe de verdelingen van de variabelen zich verhouden tot de populatie wordt nader toegelicht in bijlage 2

Nieuws volgen op sociale media (Nee = 0; Ja = 1)	Nee = 75,93% Ja = 24,07%	0	1					3207
Nieuws volgen op traditionele media (Nee = 0; Ja = 1)	Nee = 5,89%; Ja = 94,11%	0	1					3207
Opleidingsniveau	3,96(1,46)	1	6	4	3	5		3207
Leeftijd	57,83(17,06)	16	95	61	45	71		3207
Geslacht (Man = 0; Vrouw = 1)	Man = 49,67%; Vrouw = 50,33%	0	1					3207

^a Bij nominale variabelen is de frequentieverdeling vermeld in percentages.

In tabel 2 is te zien hoe de verschillende variabelen met elkaar samenhangen. Hierin vallen een aantal dingen op. Je wel of niet vertegenwoordigd voelen door de politiek hangt met geen enkele variabele echt sterk samen, maar toch zijn er wel een aantal significante resultaten. Zo is er een positief zwakke significante samenhang tussen je wel of niet vertegenwoordigd voelen door de politiek en het volgen van nieuws op traditionele media ($r = 0,04$; $p = 0,026$). Op het eerste gezicht lijkt het erop dat mensen die het nieuws volgen op traditionele media zich vaker wel vertegenwoordigd voelen door de politiek dan mensen die het nieuws niet volgen op traditionele media of andersom. Dit komt overeen met de tweede hypothese. De positieve samenhang tussen je wel of niet vertegenwoordigd voelen door politiek en opleidingsniveau is wat sterker en ook significant ($r = 0,21$; $p = <0,01$). Dit is ook zichtbaar in tabel 3 waar te zien is dat mensen die zich wel vertegenwoordigd voelen door de politiek een hoger gemiddeld opleidingsniveau hebben dan mensen die zich niet vertegenwoordigd voelen door de politiek (Niet; $M = 3,73$ en Wel; $M = 4,38$). Ten slotte is er een zwakke significante negatieve samenhang tussen je wel of niet vertegenwoordigd voelen door de politiek en leeftijd ($r = -0,06$; $p = <0,01$). Oudere mensen voelen zich minder vaak wel vertegenwoordigd dan jongere mensen, hetgeen niet in lijn is met het artikel van Stockemer en Sundström (2023). Je wel of niet vertegenwoordigd voelen door de politiek hangt zwak negatief samen met het volgen van nieuws op sociale media en geslacht, maar deze resultaten zijn niet significant. Er valt op dat er een zwak maar significante negatieve samenhang is tussen het volgen van nieuws op sociale media en het volgen van nieuws op traditionele media ($r = -0,05$; $p = <0,01$). Mensen die het nieuws volgen op sociale media volgen dus minder vaak het nieuws op traditionele media en andersom. Dit komt overeen met het rapport van het Commissariaat voor de Media dat meldt dat het toenemende gebruik van sociale media samengaat met het afnemende gebruik van traditionele media (Tump, 2024). Ten slotte is het opvallend dat leeftijd een wat sterkere significante samenhang heeft met het volgen van nieuws op sociale en traditionele media. De middelmatige samenhang tussen leeftijd en het volgen van nieuws op sociale media is negatief ($r = -0,34$; $p = <0,01$). Dit komt overeen met de literatuur die zegt dat vooral jongeren nieuws volgen op sociale

media (Klopfenstein Frei et al., 2024; Schut et al., 2024). Ook hangt leeftijd zwak tot middelmatig samen met het volgen van nieuws op traditionele media, alleen is de samenhang hiertussen positief ($r = 0,22$; $p = <0,01$). Dit betekent dat oudere mensen vaker het nieuws op traditionele media volgen.

Tabel 2: Samenhangende maten van alle variabelen die zijn opgenomen in de analyse ^a

	Je wel of niet vertegenwoordigd voelen door de politiek (Niet = 0; Wel = 1)	Nieuws volgen op sociale media (Nee = 0; Ja = 1)	Nieuws volgen op traditionele media (Nee = 0; Ja = 1)	Opleidingsniveau	Leeftijd	Geslacht (Man = 0; Vrouw = 1)
Je wel of niet vertegenwoordigd voelen door de politiek (Nee = 0; Wel = 1)	-					
Nieuws volgen op sociale media (Nee = 0; Ja = 1)	-0,03 ^a	-				
Nieuws volgen op traditionele media (Nee = 0; Ja = 1)	0,04 ^{a*}	-0,05 ^{a**}	-			
Opleidingsniveau	0,21 ^{b**}	0,04 ^{b*}	0,07 ^{b**}	-		
Leeftijd	-0,06 ^{b**}	-0,34 ^{b**}	0,22 ^{b**}	-0,20 ^{b**}	-	
Geslacht (Man = 0; Vrouw = 1)	-0,02 ^a	0,01 ^a	-0,09 ^{a**}	-0,04 ^{b**}	-0,08 ^{b**}	-

^a zijn Phi-coëfficiënt waarden, ^b zijn Pearson correlatie waarden

* significant bij $p < 0,05$, ** significant bij $p < 0,01$; tweezijdige toets.

Tabel 3: De relatie tussen *je wel of niet vertegenwoordigd voelen door de politiek* en de andere variabelen (N=3207)

	Je wel of niet vertegenwoordigd voelen door de politiek	<i>M</i>	<i>SD</i>	<i>T-waarde</i>	<i>Df</i>
Nieuws volgen op sociale media	Niet	0,25	0,43	1.65	3205
	Wel	0,22	0,42		
Nieuws volgen op traditionele media	Niet	0,93	0,25	-2,41*	2665,6
	Wel	0,95	0,21		
Opleidingsniveau	Niet	3,73	1,44	-12,14**	3205
	Wel	4,38	1,41		
Leeftijd	Niet	58,62	16,70	3,48**	2216,1

	Wel	56,39	17,60		
Geslacht	Niet	0,51	0,50	0,86	3205
	Wel	0,49	0,50		

* significant bij $p < 0,05$, ** significant bij $p < 0,01$; tweezijdige toets.

Modelevaluatie

Om de hypothesen te toetsen zijn er vier logistische modellen geschat. In deze paragraaf wordt gekeken naar de kwaliteit van de modellen en dus hoe goed de modellen zijn in het verklaren van de variantie van je wel of niet vertegenwoordigd voelen door de politiek. Daarnaast worden de assumpties van logistische regressie, multicollineariteit en de uitbijters en invloedrijke punten kort besproken. In bijlage 3 worden de assumpties, multicollineariteit en de uitbijters en invloedrijke punten nader toegelicht.

In tabel 4 is te zien dat de modellen steeds significant minder fouten maken dan het lege model en de vorige modellen (Model 1; $X^2 = 147,970$; $df = 3$; $p = < 0,01$ en Model 2; $X^2 = 7,592$; $df = 1$; $p = < 0,01$ en Model 3; $X^2 = 4,043$; $df = 1$; $p = 0,04$). Het toevoegen van de controlevariabelen in model 1 en het toevoegen van het volgen van nieuws op sociale media in model 2 en het volgen van nieuws op traditionele media in model 3 zorgen er dus voor dat er significant minder fouten worden gemaakt en helpen dus significant mee met het voorspellen van je wel of niet vertegenwoordigd voelen door de politiek. Alleen het laatste model voorspelt niet een significant groter deel van de variantie van je wel of niet vertegenwoordigd voelen door de politiek (Model 4; $X^2 = 3,475$; $df = 1$; $p = 0,06$). Het interactie-effect helpt dus niet significant mee met het voorspellen van je wel of niet vertegenwoordigd voelen door de politiek. De hosmer-lemeshow toetsen en de classificatietabellen die te zien zijn in bijlage 2 vertellen een andere verhaal. Zo zijn de hosmer-lemeshow toetsen allemaal significant, wat erop duidt dat de verwachte waardes niet gelijk zijn aan de geobserveerde waardes. Het model heeft dus geen goede fit. Daarnaast laten de classificatietabellen zien dat het percentage goede voorspellingen wel bij elk model verbetert, maar wel minimaal. Zo voorspelde het lege model 64,8% goed en het laatste model 66,5% goed. Dat het percentage goede voorspellingen niet heel erg verbeterd komt waarschijnlijk door de scheve verdeling van je wel of niet vertegenwoordigd voelen door de politiek (tabel 1; Niet = 64,76,%; Wel = 35,24%). Dus hoewel het volgen van nieuws op sociale en traditionele media individueel significant helpen bij het verklaren van je wel of niet vertegenwoordigd voelen door de politiek, lijkt het gehele model dus niet heel erg goed te zijn in het verklaren van de variantie in je wel of niet vertegenwoordigd voelen door de politiek. Dit geeft onzekerheid over hoe betrouwbaar de gevonden effecten zijn, waardoor er voorzichtig gedaan moet worden met het trekken van conclusies over de sterkte van deze verbanden.

Bij logistische regressie is het belangrijk dat de waarnemingen onafhankelijk van elkaar zijn. Aan deze assumptie wordt voldaan vanwege de manier waarop de data verzameld is en door willekeurig één iemand per huishouden te selecteren. In tabel 4 is te zien dat er geen grote VIF-waarden zijn, wat er op duidt dat er geen sprake is van multicollineariteit behalve bij het volgen van nieuws op sociale media en het interactie-effect. Het lijkt erop dat het interactie-effect en het volgen van nieuws op sociale media sterk met elkaar samenhangen en hierdoor is het dus moeilijk om het effect van de afzonderlijke variabelen te bepalen. In bijlage 3 zijn de VIF-waarden ook berekend bij model 3 en hier is geen sprake van hoge multicollineariteit. Ten slotte is er gekeken naar de uitbijters en invloedrijke punten. Er zijn wel een redelijk aantal respondenten met hoge *leverage* waarden en extreme *DFBetas* waarden. Deze respondenten hebben geldige scores en zijn voornamelijk respondenten die het nieuws niet volgen op traditionele media. Deze groep is al erg klein waardoor er veel statistische power verloren gaat na het wegfilteren van deze respondenten.

Hypothesetoetsing

De eerste hypothese “*Mensen die het nieuws volgen op sociale media hebben een minder grote kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op sociale media*” wordt getoetst bij model 2. In tabel 4 bij model 2 is te zien dat het volgen van nieuws op sociale media een negatieve invloed heeft op je wel of niet vertegenwoordigd voelen door de politiek (Model 2; $b = 0,348$; $Odds-ratio = 0,771$; $p = <0,01$). Dit betekent dat als iemand nieuws volgt op sociale media, dit ervoor zorgt dat er een minder grote kans is dat hij of zij zich vertegenwoordigd voelt door de politiek, gecontroleerd voor opleidingsniveau, leeftijd en geslacht. Omdat het volgen van nieuws op sociale media significant negatief samenhangt met leeftijd en significant positief samenhangt met opleidingsniveau worden de kansen om je wel vertegenwoordigd te voelen door de politiek vergeleken tussen twee vrouwen van 20 met een score van 5 op opleidingsniveau. In tabel 4 bij model 2 heeft de vrouw die het nieuws niet volgt op sociale media een kans van 48,0% dat zij zich vertegenwoordigd voelt door de politiek terwijl de vrouw die het nieuws wel volgt op sociale media een kans van 41,6% heeft om zich vertegenwoordigd te voelen door de politiek. De invloed van het volgen van nieuws op sociale media op je wel of niet vertegenwoordigd voelen door de politiek is wel redelijk groot, want het verschil tussen iemand die het nieuws volgt op sociale media en iemand die het nieuws niet volgt op sociale media is dus ongeveer 6,4%-punt als de controle variabelen gelijk blijven. Dit verschil is voor verschillende combinaties van achtergronden uitgerekend en wordt niet heel veel groter of kleiner. Het effect van nieuws volgen op sociale media is significant en dit betekent dat dit effect niet alleen toevallig voorkomt in de steekproef. De hypothese dat mensen die het

nieuws volgen op sociale media een minder grote kans hebben om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op sociale media wordt dus ondersteund.

De tweede hypothese *“Mensen die het nieuws volgen op traditionele media hebben een grotere kans om zich vertegenwoordigd te voelen door de politiek dan mensen die het nieuws niet volgen op traditionele media”* wordt getoetst bij model 3. In tabel 4 bij model 3 staat dat het volgen van nieuws op traditionele media een positieve invloed heeft op je wel of niet vertegenwoordigd voelen door de politiek (Model 3; $b = 0,348$; $Odds-ratio = 1,417$; $p = 0,048$). Iemand die het nieuws volgt op traditionele media heeft een grotere kans om zich vertegenwoordigd te voelen door de politiek. Bij dit model worden de kansen van twee mannen van 70 die het nieuws niet volgen op sociale media met een score van 6 op opleidingsniveau met elkaar vergeleken, omdat leeftijd en opleidingsniveau significant positief samenhangen met het volgen van nieuws op traditionele media en geslacht en het volgen van nieuws op sociale media significant negatief samenhangen met het volgen van nieuws op traditionele media. In tabel 4 bij model 3 heeft de man die het nieuws niet volgt op traditionele media en niet volgt op sociale media een kans van 41,9% om zich vertegenwoordigd te voelen door de politiek. De man die het nieuws alleen volgt op traditionele media heeft een kans van 50,4% om zich vertegenwoordigd te voelen door de politiek. Het volgen van nieuws op traditionele media heeft een redelijk grote invloed op je wel of niet vertegenwoordigd voelen door de politiek. Ook hier geldt dat dit verschil voor verschillende combinaties is uitgerekend, waarbij het verschil niet heel veel kleiner of groter wordt. Het verschil tussen iemand die het nieuws niet volgt op traditionele media en iemand die het nieuws wel volgt op traditionele media is dus ongeveer 8,5%-punt als de controle variabelen gelijk blijven. Ook het positieve effect van nieuws volgen op traditionele media is significant. Daarmee wordt ook de tweede hypothese ondersteund.

De derde hypothese *“De negatieve invloed van het volgen van nieuws op sociale media op je wel of niet vertegenwoordigd voelen door de politiek wordt verzwakt als iemand het nieuws ook volgt via de traditionele media”* wordt getoetst bij model 4. In tabel 4 bij model 4 is te zien dat het interactie-effect een negatieve invloed heeft op je wel of niet vertegenwoordigd voelen door de politiek (Model 4; $b = -0,678$; $Odds-ratio = 0,508$; $p = 0,06$). In dit model is ook te zien dat de invloed van het volgen van nieuws op sociale media niet meer negatief is, maar juist positief is ook al is het niet meer significant (Model 4; $b = 0,366$; $Odds-ratio = 1,433$; $p = 0,29$). In tabel 4 bij model 4 heeft een man van 58 met een score van 4 op opleidingsniveau die het nieuws volgt op sociale media maar niet op traditionele media een kans van 33,0% om zich vertegenwoordigd te voelen door de politiek. De man die het nieuws alleen volgt op traditionele media heeft een kans van 37,7% om zich vertegenwoordigd te voelen door de politiek (Model 4; $b = 0,572$; $Odds-ratio = 1,771$; $p =$

<0,01). De man die het nieuws volgt op sociale en traditionele media heeft een kans van 30,7% om zich vertegenwoordigd te voelen door de politiek. Dit ondersteunt de hypothese niet en laat zelfs een tegenovergesteld effect zien. Iemand die het nieuws volgt op sociale media en op traditionele media heeft een minder grote kans om zich vertegenwoordigd te voelen door de politiek dan iemand die het nieuws alleen volgt op sociale media. Het toevoegen van het interactie-effect zorgt niet voor een significante vermindering in het aantal fouten en is zelf ook niet significant. Er kan dus niet met voldoende zekerheid gezegd worden dat er een interactie-effect is. De derde hypothese wordt niet ondersteund en kan dus worden verworpen.

Ten slotte wordt er nog kort gekeken naar de controlevariabelen opleidingsniveau, leeftijd en geslacht. De invloed van opleidingsniveau op je wel of niet vertegenwoordigd voelen door de politiek is significant en vrij groot (Model 1; $b = 0,313$; *Odds-ratio* = 1,368; $p = <0,01$). Dit betekent dat hoe hoger het opleidingsniveau van iemand is, hoe groter de kans om zich vertegenwoordigd te voelen door de politiek. De invloed van opleidingsniveau verandert door de modellen heen niet heel erg en blijft significant. Leeftijd heeft in model 1 een niet significante negatieve invloed op je wel of niet vertegenwoordigd voelen door de politiek (Model 1; $b = -0,003$; *Odds-ratio* = 0,997; $p = 0,19$). Dit betekent dat hoe ouder iemand wordt hoe minder groot de kans is dat hij of zij zich vertegenwoordigd voelt door politiek. Leeftijd heeft een vrij grote schaal waardoor dit effect redelijk groot kan zijn en vanaf model 2 is dit effect significant wat betekent dat dit gevonden effect waarschijnlijk niet op toeval berust. Zo heeft bij model 2 een man van 80 met een score van 4 op opleidingsniveau die het nieuws niet volgt op sociale media een kans van 34,4% om zich vertegenwoordigd te voelen door de politiek terwijl een man van 20 die hetzelfde opleidingsniveau heeft en het nieuws op dezelfde manier volgt, een kans van 41,4% heeft. Dit is toch een verschil van 7%-punt terwijl het verschil tussen een man van 80 en een man van 79 0,1%-punt van elkaar verschillen. Geslacht heeft in model 1 een vrij klein niet significant negatief effect (Model 1; $b = -0,042$; *Odds-ratio* = 0,959; $p = 0,58$). Een vrouw heeft dus een minder grote kans om zich vertegenwoordigd te voelen door de politiek dan een man, maar dit effect is vrij klein en in geen enkel model significant.

	Model 1		Model 2		Model 3		Model 4		VIF
	<i>b/Odds-ratio</i>	<i>SE</i>	<i>b/Odds-ratio</i>	<i>SE</i>	<i>b/Odds-ratio</i>	<i>SE</i>	<i>b/Odds-ratio</i>	<i>SE</i>	
Intercept	- 1,691**/0,184	0,201	- 1,495**/0,224	0,213	- 1,743**/0,175	0,249	- 1,957**/0,141	0,278	
Opleidingsniveau	0,313**/1,368	0,028	0,312**/1,366	0,028	0,306**/1,358	0,028	0,308**/1,360	0,028	1,044

Leeftijd	-0,003/0,997	0,002	-0,005*/0,995	0,002	-0,006*/0,994	0,002	-0,006*/0,994	0,002	1,234
Geslacht (Man = 0; Vrouw = 1)	-0,042/0,959	0,076	-0,046/0,955	0,076	-0,035/0,966	0,076	-0,030/0,970	0,076	1,016
Nieuws volgen op sociale media (Nee = 0; Ja = 1)			-0,260*/0,771	0,095	-	0,095	0,366/1,443	0,348	15,148
					0,266**/0,766				
Nieuws volgen op traditionele media (Nee = 0; Ja = 1)					0,348*/1,417	0,176	0,572**/1,771	0,218	1,622
Interactie-effect sociale media en traditionele media							-0,678/0,508	0,360	15,272
<i>Deviance (-2LL)</i>	4014,030		4006,438		4002,395		3998,920		
<i>X² Toets (LR)</i>	147,970**	(df = 3)	7,592**	(df = 1)	4,043*	(df = 1)	3,475	(df = 1)	
<i>Hosmer-Lemeshowtoets</i>	36,197**	(df = 8)	28,468**	(df = 8)	21,153**	(df = 8)	18,121*	(df = 8)	
<i>N</i>	3207		3207		3207		3207		

Tabel 4: Resultaten van sequentiële logistische regressieanalyse met je wel of niet vertegenwoordigd voelen door de politiek als afhankelijke, nieuws volgen op sociale media onafhankelijke variabele en nieuws volgen op traditionele media modererende variabele.

* significant bij $p < 0,05$, ** significant bij $p < 0,01$; tweezijdige toets.

Conclusie en discussie

In dit onderzoek is met behulp van data uit het LISS panel gekeken naar de relatie tussen het volgen van nieuws op sociale en traditionele media en het gevoel van politieke vertegenwoordiging. De volgende onderzoeksvraag werd gesteld: *In welke mate heeft het volgen van nieuws op sociale en traditionele media invloed op het gevoel van politieke vertegenwoordiging, en in hoeverre beïnvloedt het volgen van nieuws op de ene mediavorm het verband tussen nieuwsgebruik op de andere mediavorm en dit gevoel?* Hierbij werd verwacht dat het volgen van nieuws op sociale media een negatieve invloed heeft op het

gevoel van politieke vertegenwoordiging en het volgen van nieuws op traditionele media een positieve invloed heeft op het gevoel van politieke vertegenwoordiging en het negatieve effect van het volgen van nieuws op sociale media verzwakt. Ten eerste laten de resultaten zien dat het volgen van nieuws op sociale media inderdaad een significante negatieve invloed heeft op het gevoel van politieke vertegenwoordiging en dat het volgen van nieuws op traditionele media een significante positieve invloed heeft. Dit sluit aan bij soortgelijk onderzoek van Ceron (2015) en Ceron en Memoli (2016) waar het volgen van nieuws op sociale media een negatieve invloed heeft op politiek vertrouwen en tevredenheid met de overheid terwijl het volgen van nieuws op traditionele media juist een positieve invloed heeft. Hoewel deze onderzoeken niet specifiek focussen op politieke vertegenwoordiging, hangen politiek vertrouwen en tevredenheid met de overheid nauw samen met het gevoel van politieke vertegenwoordiging (Ceron, 2015; Ceron & Memoli, 2016). Ten tweede werd er geen significant interactie-effect gevonden. Er kan dus niet met genoeg zekerheid geconcludeerd worden dat het volgen van nieuws op de ene mediavorm invloed heeft op de relatie tussen het gevoel van politieke vertegenwoordiging en de andere mediavorm.

Ook al worden de eerste twee verwachtingen ondersteund met een significant resultaat er moet toch voorzichtig gedaan worden met het generaliseren naar de populatie en het interpreteren van de sterkte van de gevonden effecten. Ten eerste was de uiteindelijke steekproef niet volledig representatief voor de populatie. Respondenten tussen de 16 en 55 jaar waren ondervertegenwoordigd, terwijl respondenten van 55 jaar en ouder juist oververtegenwoordigd waren. Hoewel dit niet automatisch betekent dat de gevonden effecten niet generaliseerbaar zijn, kan de scheve leeftijdsverdeling wel invloed hebben gehad op de resultaten. Als de steekproef representatiever was geweest, zou het aandeel jongere respondenten hoger zijn, wat waarschijnlijk had geleid tot een grotere groep die nieuws volgt op sociale media en een kleinere groep die uitsluitend nieuws volgt op traditionele media. Aangezien leeftijd negatief samenhangt met het volgen van nieuws op sociale media en positief met het volgen van nieuws op traditionele media, is het aannemelijk dat een meer representatieve steekproef meer variatie had opgeleverd in het mediagebruik. In de huidige steekproef is het aantal respondenten dat wel sociale media maar geen traditionele media gebruikt zeer beperkt, wat de statistische power voor het detecteren van een interactie-effect aanzienlijk beperkt. Het is daarom mogelijk dat er wel een significante interactie-effect was gevonden als de steekproef beter representatief was. Ten tweede blijkt uit de significante Hosmer-Lemeshow-toetsen en de beperkte verbeteringen in de classificatietabellen dat de modellen geen goede fit hebben. Hoewel het volgen van nieuws op sociale en traditionele media afzonderlijk een significante bijdrage levert aan de verklaring van het gevoel van politieke vertegenwoordiging, blijken de modellen als geheel slechts in beperkte mate in staat om dit gevoel goed te voorspellen. Daardoor is het moeilijk in te

schatten hoe sterk de invloed van het volgen van nieuws op sociale media en traditionele media op het gevoel van politieke vertegenwoordiging werkelijk is.

Daarnaast is het belangrijk om te kijken naar de operationalisaties van het gevoel van politieke vertegenwoordiging en het volgen van nieuws op sociale media en traditionele media in dit onderzoek. Om je wel of niet vertegenwoordigd voelen door de politiek te meten, werd in dit onderzoek de vraag gesteld: 'Vindt u dat er in de politiek genoeg wordt opgekomen voor de belangen van mensen zoals u?' Hierop kon onder andere de antwoorden 'te veel' en 'veel te veel' gegeven worden, maar als je het hebt over de 4 aspecten die volgens De Mulder (2022) belangrijk zijn voor mensen om zich vertegenwoordigd te voelen door de politiek kan je je afvragen of een vertegenwoordiger te veel naar je kan luisteren of kan handelen in naam van jouw belangen. Daarnaast geeft De Mulder (2022) aan dat iemand zich vertegenwoordigd kan voelen door sommige politici, maar niet door allen. De huidige vraagstelling houdt geen rekening met deze nuance. Het is dus mogelijk dat respondenten die zich niet vertegenwoordigd voelen door alle politici een negatief antwoord geven, terwijl zij zich wel degelijk door sommige vertegenwoordigers vertegenwoordigd voelen. Dat de antwoordopties in de vragenlijst niet goed passen bij het concept van politieke vertegenwoordiging en de vraagstelling enige nuance mist kan de resultaten hebben vertekend en de validiteit hebben aangetast.

De operationalisatie van het volgen van nieuws op sociale media en traditionele media zorgt voor verlies van informatie. Er wordt namelijk bij de operationalisatie alleen rekening gehouden met of iemand het nieuws volgt op sociale media of traditionele media, zonder bijvoorbeeld te differentiëren in frequentie. De verklaringsmechanismen die in dit onderzoek worden onderzocht, zouden namelijk kunnen verschillen tussen iemand die incidenteel nieuws tegenkomt op sociale media en iemand die dagelijks actief nieuws consumeert op sociale media. Dit beperkte onderscheid in mediagebruik kan de verklarende kracht van de modellen verzwakken en kan ervoor zorgen dat de sterkte van de gevonden effecten onderschat is.

Dit onderzoek richt zich op de invloed van nieuwsmedia op het gevoel van politieke vertegenwoordiging. Hoewel eerdere onderzoeken al aantoonde dat het volgen van nieuws op sociale media en traditionele media invloed kan hebben op het vertrouwen in en tevredenheid met de overheid, was er nog weinig bekend over de relatie tussen het volgen van nieuws op deze mediavormen en het gevoel vertegenwoordigd te worden. Dit onderzoek is uniek, omdat het bekijkt of het volgen van nieuws via de ene mediavorm invloed heeft op hoe het gebruik van een andere mediavorm samenhangt met het gevoel van politieke vertegenwoordiging. Daarnaast draagt dit onderzoek bij aan het maatschappelijke debat over sociale media en traditionele media. De conclusies kunnen meegenomen worden in het

ontwerp van wet- en regelgeving en burgerschapsonderwijs ten aanzien van media, alsook in opvoedkundige overwegingen.

Dit onderzoek geeft inzicht in de sterkte en de manier waarop het volgen van nieuws op sociale en traditionele media invloed heeft op het gevoel van politieke vertegenwoordiging. Er kan echter niet met zekerheid worden vastgesteld of de in het onderzoek aangehaalde verklaringsmechanismen daadwerkelijk verantwoordelijk zijn voor de gevonden effecten. Kwalitatief onderzoek zou meer inzicht kunnen geven in hoe mensen het volgen van nieuws op sociale en traditionele media ervaren en hoe dat hun perceptie over politici en hun beleid kan beïnvloeden. Daarnaast vormen de hierboven genoemde beperkingen een goed startpunt voor vervolgonderzoek, omdat ze inhoudelijke en methodologische aspecten blootleggen die verder uitgewerkt kunnen worden.

Literatuur

Akagündüz, A. (2025, maart 13). *Opinie: Voor een functionerende democratie moeten we jongeren een toekomst én een heden bieden*. *De Volkskrant*. Geraadpleegd op 15 mei 2025, van <https://www.volkskrant.nl/columns-opinie/opinie-voor-een-functionerende-democratie-moeten-we-jongeren-een-toekomst-en-een-heden-bieden~be4e8489/>

Andeweg, R. B., & Aarts, K. (2017). Studying political legitimacy. In *Oxford University Press eBooks*. <https://doi.org/10.1093/oso/9780198793717.003.0011>

Besalú, R., & Pont-Sorribes, C. (2021). Credibility of Digital Political News in Spain: Comparison between Traditional Media and Social Media. *Social Sciences*, 10(5), 170. <https://doi.org/10.3390/socsci10050170>

Burkens, M. C., Kummeling, H. R. B. M., Vermeulen, B. P., & Widdershoven, R. J. G. M. (2017). *Beginnelsen van de democratische rechtsstaat : inleiding tot de grondslagen van het Nederlandse staats- en bestuursrecht* (8e druk). Wolters Kluwer.

Camaj, L. (2012). The media's role in fighting corruption. *The International Journal of Press/Politics*, 18(1), 21–42. <https://doi.org/10.1177/1940161212462741>

Cappella, J. N., & Jamieson, K. H. (1997). *Spiral of cynicism : the press and the public good*. Oxford University Press. Geraadpleegd op 18 mei 2025, van <http://site.ebrary.com/id/10086932>

Centraal Bureau voor de Statistiek. (n.d.-a). *Bevolkingspiramide*. Centraal Bureau Voor De Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide>

Centraal Bureau voor de Statistiek. (n.d.-b). [Leeftijdsopbouw Nederland 2023]. Geraadpleegd op 25 april 2025, van <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingspiramide>

Centraal Bureau voor de Statistiek. (n.d.-c). *Mannen en vrouwen*. Centraal Bureau Voor De Statistiek. <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/mannen-en-vrouwen>

Centraal Bureau voor de Statistiek. (2024, oktober 8). *Nederlanders hebben steeds vaker hbo of universiteit gevolgd*. Centraal Bureau Voor De Statistiek. <https://www.cbs.nl/nl-nl/nieuws/2024/41/nederlanders-hebben-steeds-vaker-hbo-of-universiteit-gevolgd>

Ceron, A. (2015). Internet, News, and Political Trust: The Difference Between Social Media and Online Media Outlets. *Journal of Computer-Mediated Communication*, 20(5), 487–503. <https://doi.org/10.1111/jcc4.12129>

Ceron, A., & Memoli, V. (2016). Flames and Debates: Do Social Media Affect Satisfaction with Democracy? *Social Indicators Research : An International and Interdisciplinary Journal for Quality-of-Life Measurement*, 126(1), 225–240. <https://doi.org/10.1007/s11205-015-0893-x>

Dagevos, J., & Vermeulen, F. (2024). *Is de politiek er voor iedereen?* Sociaal en Cultureel Planbureau. Geraadpleegd op 01 maart 2025, van <https://www.scp.nl/publicaties/publicaties/2024/01/25/is-de-politiek-er-voor-iedereen>

De Mulder, A. (2022). Making sense of citizens' sense of being represented. a novel conceptualisation and measure of feeling represented. *Representation*, 59(4), 633–657. <https://doi.org/10.1080/00344893.2022.2095662>

Den Ridder, J., van 't Hul, L., & Broere, J. (2024). Burgerperspectieven 2024 bericht 1. Sociaal en Cultureel Planbureau. Geraadpleegd op 27 februari 2025, van <https://www.scp.nl/publicaties/publicaties/2024/04/25/burgerperspectieven-2024-bericht-1>

Elshout, S. (2022, september 1). *Codeboek – Achtergrondvariabelen LISS Panel, 1 september 2022* [Codeboek]. CentERdata. <https://www.dataarchive.lissdata.nl/study-units/view/322>

Elshout, S. (2024, mei). *Codeboek – Politics and Values – LISS Core Study* [Codeboek]. CentERdata. <https://www.dataarchive.lissdata.nl/study-units/view/1478>

Gillis, A. S. (2024, december 5). *Traditional media vs. new media: Differences, pros and cons*. WhatIs. Geraadpleegd op 11 maart 2025, van <https://www.techtarget.com/whatis/feature/Traditional-media-vs-new-media-Differences-pros-and-cons>

Haugsgjerd, A., & Karlsen, R. (2024). Election Campaigns, News Consumption Gaps, and Social Media: Equalizing Political News Use When It Matters? *International Journal of Press/Politics*, 29(2). <https://doi.org/10.1177/19401612221112014>

How it works - LISS panel. (2023, november 13). LISS Panel. Geraadpleegd op 25 maart 2025, van <https://www.lissdata.nl/how-it-works>

Kenworthy, L., & Malami, M. (1999). Gender Inequality in Political Representation: A Worldwide Comparative Analysis. *Social Forces*, 78(1), 235–268. <https://doi.org/10.2307/3005796>

Klopfenstein Frei, N., Wyss, V., Gnach, A., & Weber, W. (2024). “It’s a matter of age”: Four dimensions of youths’ news consumption. *Journalism*, 25(1), 100–121. <https://doi.org/10.1177/14648849221123385>

Knoef, M., & de Vos, K. (2009). *The representativeness of LISS, an online probability panel* (Versie 2). CentERdata. Geraadpleegd op 26 maart 2025, van https://www.lissdata.nl/app/uploads/sites/4/2023/10/3.-paper_knoef_devos_website.pdf

Knutson, B., Hsu, T. W., Ko, M., & Tsai, J. L. (2024). News source bias and sentiment on social media. *PLoS ONE*, 19(10). <https://doi.org/10.1371/journal.pone.0305148>

Lijphart, A. (1997). Unequal Participation: Democracy’s Unresolved Dilemma. *The American Political Science Review*, 91(1), 1–14. <https://doi.org/10.2307/2952255>

Linde, J., & Peters, Y. (2020). Responsiveness, support, and responsibility: How democratic responsiveness facilitates responsible government. *Party Politics*, 26(3), 291–304. <https://doi.org/10.1177/1354068818763986>

Methodology - LISS panel. (2024, juli 22). LISS Panel. Geraadpleegd op 25 maart 2025, van <https://www.lissdata.nl/methodology>

Miltenburg, E., Geurkink, B., den Ridder, J., Kunst, S., & Sipma, T. (2024). *Burgerperspectieven 2024 bericht 2. Sociaal en Cultureel Planbureau*. Geraadpleegd op 27 februari 2025, van <https://www.scp.nl/publicaties/publicaties/2024/09/03/burgerperspectieven-2024-bericht-2>

Mügge, L., Runderkamp, Z., & Spierings, N. (2025). Descriptive representation of marginalized groups in Dutch politics. In *Oxford University Press eBooks* (pp. 495–512). <https://doi.org/10.1093/oxfordhb/9780198875499.013.29>

Nowacki, L. (2024). The old king is dead, long live the algorithmic king – the decline of Facebook and the rise of TikTok – comparative study of algorithmic design of social media platforms. *International Social Science Journal*, 74(254), 1325–1338. <https://doi-org.proxy-ub.rug.nl/10.1111/issj.12513>

Park, C. S., & Kaye, B. K. (2020). What's This? Incidental Exposure to News on Social Media, News-Finds-Me Perception, News Efficacy, and News Consumption. *Mass Communication & Society*, 23(2), 157–180. <https://doi.org/10.1080/15205436.2019.1702216>

Pitkin, H. F. (1967). *The concept of representation*. New York, NY: Atherton Press.

Schakel, W., & Van Der Pas, D. (2021). Degrees of influence: Educational inequality in policy representation. *European Journal of Political Research*, 60(2), 418–437. <https://doi.org/10.1111/1475-6765.12405>

Schut, K., Costera Meijer, I., Lauf, E. (2024). *Jongeren, nieuws en sociale media. Een blik op de toekomst van het nieuws*. Commissariaat voor de Media. Geraadpleegd op 25 februari 2025, van <https://www.cvdm.nl/mediamonitor/mediamonitor-jongeren-nieuws-en-sociale-media/>

Soroka, S., & McAdams, S. (2015). News, Politics, and Negativity. *Political Communication*, 32(1), 1–22. <https://doi.org/10.1080/10584609.2014.881942>

Stockemer, D., & Sundström, A. (2025). Age Inequalities in Political Representation: A Review Article. *Government and Opposition*, 60(1), 271–288. <https://doi.org/10.1017/gov.2023.11>

Tump, Y. (2024, december 12). *Mediamonitor 2024: Verschuiving in mediagebruik naar internationale platformen vormt risico voor geïnformeerdheid van burgers*. Commissariaat voor de Media. Geraadpleegd op 25 februari 2025, van

<https://www.cvdm.nl/nieuws/mediamonitor-2024-verschuiving-in-mediagebruik-naar-internationale-platformen-vormt-risico-voor-geinformeerdheid-van-burgers/>

Universiteit van Amsterdam. (2023, oktober 23). *'Er zijn nog steeds te weinig vrouwen in de politiek.'* Universiteit Van Amsterdam. Geraadpleegd op 15 mei 2025, van <https://www.uva.nl/shared-content/faculteiten/nl/faculteit-der-maatschappij-en-gedragwetenschappen/nieuws/2022/02/verkiezingen-in-nederland-nog-lang-niet-genoege-vrouwen-in-de-politiek.html>

Universiteit van Amsterdam. (2024, januari 25). *Mensen met een migratieachtergrond voelen zich politiek niet goed vertegenwoordigd.* Universiteit Van Amsterdam. Geraadpleegd op 15 mei 2025, van <https://www.uva.nl/content/nieuws/nieuwsberichten/2024/01/mensen-met-een-migratieachtergrond-voelen-zich-politiek-niet-goed-vertegenwoordigd.html>

Van Den Heuvel, M. (2024, juli 23). *Van Randstad naar Regio: de verschuiving in de parlementaire vertegenwoordiging - Bulletineke Justitia.* Bulletineke Justitia. Geraadpleegd op 14 mei 2025, van <https://www.bjnimmegen.nl/artikelen/van-randstad-naar-regio-de-verschuiving-in-de-parlementaire-vertegenwoordiging/>

Vega, A., & Firestone, J. M. (1995). The Effects of Gender on Congressional Behavior and the Substantive Representation of Women. *Legislative Studies Quarterly*, 20(2), 213–222. <https://doi.org/10.2307/440448>

Verheijen, M. (2024, februari). *Codeboek – Democratic innovations* [Codeboek]. CentERdata. <https://www.dataarchive.lisssdata.nl/study-units/view/1456>

Vrouwen in de Tweede Kamer. (n.d.). Parlement.com. Geraadpleegd op 28 mei 2025, van https://www.parlement.com/id/vh8lnhrre0zv/vrouwen_in_de_tweede_kamer

Zhang, M., Wu, H., Huang, Y., Han, R., Fu, X., Yuan, Z., & Liang, S. (2024). Negative news headlines are more attractive: negativity bias in online news reading and sharing. *Current Psychology : A Journal for Diverse Perspectives on Diverse Psychological Issues*, 43(38), 30156–30169. <https://doi.org/10.1007/s12144-024-06646-6>

Bijlage 1

In bijlage 1 staan de beschrijvende statistieken van de gebruikte variabelen en wordt de syntax en de output van de onbewerkte en de bewerkte variabelen weergegeven.

Daarnaast is de dataset bewerkt door willekeurig één iemand per huishouden te selecteren en de respondenten die op één van de bewerkte variabelen missende waarden heeft er ook uit te filteren. Dit eerste is gedaan om de assumptie van onafhankelijke waarnemingen niet te schenden. Na het willekeurig selecteren van één iemand per huishouden waren er nog 5188 respondenten over. Om er voor te zorgen dat het onderzoek reproduceerbaar is wordt er een seed gebruikt. Daarna werden alle respondenten die op één van de variabelen een missende waarde hadden eruit gefilterd waardoor er nog 3207 respondent overblijven. Hieronder is de syntax te zien van deze filters. De beschrijvende statistieken van de onbewerkte variabelen worden op twee manieren weergegeven, namelijk één waarbij willekeurig één iemand van een huishouden wordt geselecteerd en één waarbij er geen enkel filter actief is. De beschrijvende statistieken van de bewerkte variabelen worden weergegeven waarbij willekeurig iemand per huishouden is geselecteerd en de missende waarden weggefilterd zijn.

R script:

```
> library(tidyverse)
> library(haven)
> library(skimr)
> library(vcd)
> library(rms)
> library(glmtoolbox)
> library(MASS)
> library(qqplotr)
> library(gmodels)
> library(psych)
> #Ik begin met het maken van verschillende datasets aanmaken met verschill
ende selecties en filters
> #Oorsponkelijke dataset
> LISS <- read_sav("Merged data.sav")
> #LISS2 bevat alleen de benodigde variabelen en is nog zonder filter.
```



```

> LISS2 <- LISS %>%
+   select(leeftijd, geslacht, oplcat, cv24p002, cv24p003, cv24p166, xg23a0
23, cv24p335, nohouse_encr, nomem_encr)
> #seed instellen voor reproduceerbaarheid
> set.seed(5066638)
> #LISS3 hierbij wordt 1 iemand gekozen per huishouden.
> LISS3 <- LISS2 %>%
+   group_by(nohouse_encr) %>%
+   slice_sample(n = 1) %>%
+   ungroup()
>
> #LISS4 hierbij worden ook nog alle missende waarden eruit gefilterd. Hier
door worden alleen respondenten meegenomen die op alle variabelen een antwo
ord heeft gegeven
> LISS4 <- LISS2 %>%
+   filter(complete.cases(.)) %>%
+   group_by(nohouse_encr) %>%
+   slice_sample(n = 1) %>%
+   ungroup()

```

Je wel of niet vertegenwoordigd voelen door de politiek

Oorspronkelijke variabele:

De mate waarin iemand zich vertegenwoordigd voelt door de politiek wordt gemeten met de vraag: Vindt u dat er in de politiek genoeg wordt opgekomen voor de belangen van mensen zoals u? Deze variabele (xg23a023) had 6 mogelijke antwoord opties namelijk 1 = Veel te weinig, 2 = Te weinig, 3 = Precies genoeg, 4 = Te veel, 5 = Veel te veel en 6 = Weet ik niet. Hieronder zijn de beschrijvende statistieken en de daarbij behorende syntax van de oorspronkelijke variabele te zien.

R script oorspronkelijke variabele:

```

> #Mate waarin iemand zich vertegenwoordigd voelt door de politiek
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele xg23a023
> attributes(LISS2$xg23a023)
$label
[1] "In your opinion, do politicians adequately represent the interests of
people like you?"

$format.spss
[1] "F1.0"

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

$labels
Far from enough      Not enough  Exactly enough      Too much
      1              2              3              4
Far too much        I don't know
      5              6

> table(LISS2$xg23a023, useNA = "always")

  1    2    3    4    5    6 <NA>
651 2376 1574   79    8  444 6927
> skim(LISS2$xg23a023)
— Data Summary —
Name                               Values
LISS2$xg23a023

```

Number of rows 12059
Number of columns 1

Column type frequency:
numeric 1

Group variables None

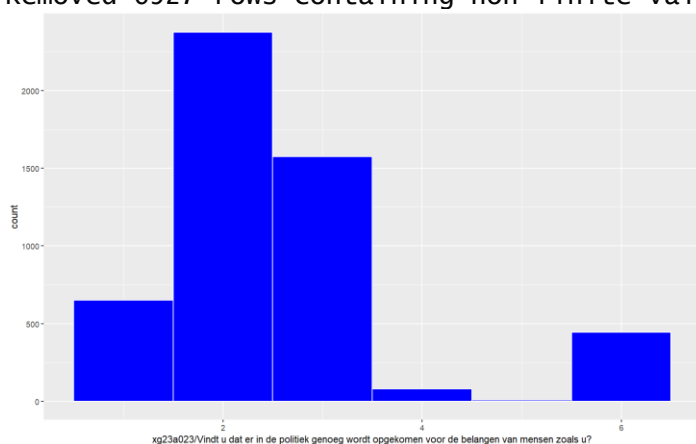
— Variable type: numeric —

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	6927	0.426	2.56	1.26	1	2	2	3	6	

```
> ggplot(data = LISS2, mapping = aes(x = xg23a023))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("xg23a023/Vindt u dat er in de politiek genoeg wordt opgekomen voor de belangen van mensen zoals u?")
```

Warning message:

Removed 6927 rows containing non-finite values (`stat\_bin()`).



R script willekeurig iemand per huishouden:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de oorspronkelijke variabele xg23a023 waarbij 1 iemand per huishouden is geselecteerd
```

```
> attributes(LISS3$xg23a023)
```

\$label

[1] "In your opinion, do politicians adequately represent the interests of people like you?"

\$format.spss

[1] "F1.0"

\$labels

Far from enough	Not enough	Exactly enough	Too much
1	2	3	4
Far too much	I don't know		
5	6		

\$class

[1] "haven_labelled" "vctrs_vctr" "double"

```
> table(LISS3$xg23a023, useNA = "always")
```

1	2	3	4	5	6	<NA>
334	1268	823	35	4	228	2496

```
> skim(LISS3$xg23a023)
```

— Data Summary —

	values
Name	LISS3\$xg23a023
Number of rows	5188

Number of columns	1
Column type frequency:	
numeric	1
Group variables	None

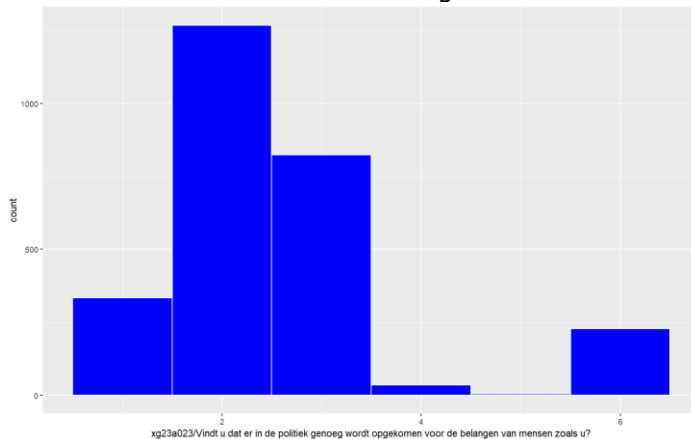
— variable type: numeric —

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	2496	0.519	2.55	1.24	1	2	2	3	6	

```
> ggplot(data = LISS3, mapping = aes(x = xg23a023))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("xg23a023/Vindt u dat er in de politiek genoeg wordt opgekomen voor de belangen van mensen zoals u?")
```

Warning message:

Removed 2496 rows containing non-finite values (`stat\_bin()`).



Bewerkingen:

De variabele is op twee manier bewerkt. In eerste instantie is de score 6 veranderd naar een missende waarde. Deze score is eruit gehaald omdat deze score het analyseren moeilijker maakt en 'Weet ik niet' nietszeggend is. De variabele is daarna gecodeerd tot een binaire variabele, omdat er niet aan de assumpties van lineaire regressie werd voldaan is er gekozen om een binaire logistische regressie te doen. De figuren die zijn gebruikt om de assumpties van lineaire logistische regressie te controleren zijn te zien in bijlage 3. De binaire variabele heeft 2 scores namelijk 0 = Niet 1 = Wel, waarbij de oorspronkelijke scores van 1 en 2 zijn samengevoegd tot de nieuwe score 0 en de oorspronkelijke scores van 3,4 en 5 zijn samengevoegd tot de score 1. Zo is de variabele veranderd van de mate waarin iemand zich vertegenwoordigd voelt door de politiek tot je wel of niet vertegenwoordigd voelen door de politiek.

R script bewerkingen:

```
> #Hercoderen politieke vertegenwoordiging weghalen 6 = weet ik niet
> LISS2 <- LISS2 %>%
+   mutate(Gevoel_Politieke_vertegenwoordiging = case_when(
+     xg23a023 == 1 ~ 1,
+     xg23a023 == 2 ~ 2,
```

```

+   xg23a023 == 3 ~ 3,
+   xg23a023 == 4 ~ 4,
+   xg23a023 == 5 ~ 5,
+   xg23a023 == 6 ~ NA_real_
+ ))
>
> #Gevoel politieke vertegenwoordiging een dummy maken voor logistische regressie.
> LISS2 <- LISS2 %>%
+   mutate(Gevoel_Politieke_vertegenwoordiging_dummy = case_when(
+     xg23a023 %in% c(1, 2) ~ 0,
+     xg23a023 %in% c(3, 4, 5) ~ 1,
+     xg23a023 == 6 ~ NA_real_
+   ))

```

Uiteindelijke variabele:

De uiteindelijke variabele (Gevoel_Politieke_vertegenwoordiging_dummy) heeft dus 2 scores waarbij bijna twee derde zich niet vertegenwoordigd voelt door de politiek en iets meer dan één derde zich wel vertegenwoordigd voelt door de politiek (0 = 65%; 1 = 34%).

R script uiteindelijke variabele:

```

> #Hieronder staan de attributes, frequentieverdeling en histogram van de uiteindelijke variabele je wel of niet vertegenwoordigd voelen door de politiek waarbij 1 iemand per huishouden is geselecteerd en de missende waarden zijn weggefilterd
> attributes(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy)
NULL
> table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, useNA = "always")
 0      1 <NA>
2077 1130      0
> skim(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy)
— Data Summary —

```

Name	Values
LISS4\$Gevoel_Politieke_vertegenwoordiging_dummy	LISS4\$Gevoel_Politieke_vertegenwoordiging_dummy
Number of rows	3207
Number of columns	1

```

Column type frequency:
  numeric      1
Group variables:
  None
— Variable type: numeric —

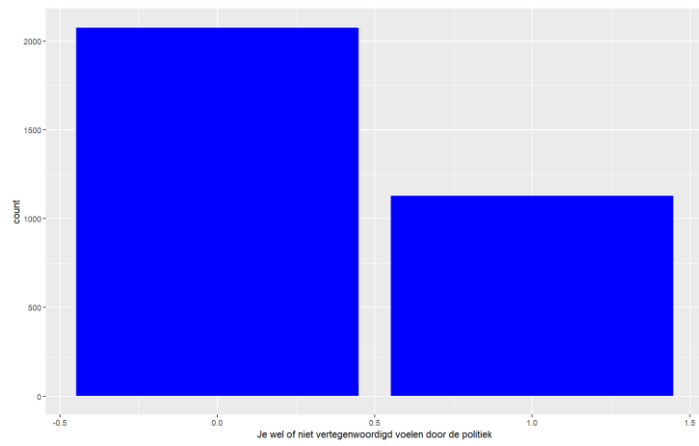
```

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
data	0	1	0.352	0.478	0	0	0	1	1	█

```

> ggplot(data = LISS4, mapping = aes(x = Gevoel_Politieke_vertegenwoordiging_dummy)) +
+   geom_bar(fill = "blue", colour = "white") +
+   xlab("Je wel of niet vertegenwoordigd voelen door de politiek")

```



Het volgen van nieuws op sociale media

Oorspronkelijke variabele:

Het volgen van nieuws op sociale media wordt gemeten met de vraag: Hoe volgt u het nieuws? Hierbij waren verschillende antwoordopties waarvan via social media (Twitter, LinkedIn, YouTube, Instagram, Facebook, etc.) daar één van was (cv24p335). Iemand die een 0 scoort op deze variabele volgt het nieuws niet via social media en iemand met een score van 1 volgt het nieuws wel via social media.

R script oorspronkelijke variabele:

```
> #Nieuws volgen op sociale media
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele cv24p335
> attributes(LISS2$cv24p335)
$label
[1] "Do you follow the news: More than one answer is possible. via social m
edia (Twitter, LinkedIn, YouTube, Instagram, Facebook, etc.)"

$format.spss
[1] "F2.0"

$display_width
[1] 4

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

$labels
I dont know      no      yes
      -9         0         1

> table(LISS2$cv24p335, useNA = "always")
 0    1 <NA>
4369 1889 5801
> skim(LISS2$cv24p335)
— Data Summary —
Name          LISS2$cv24p335
Number of rows 12059
Number of columns 1
```

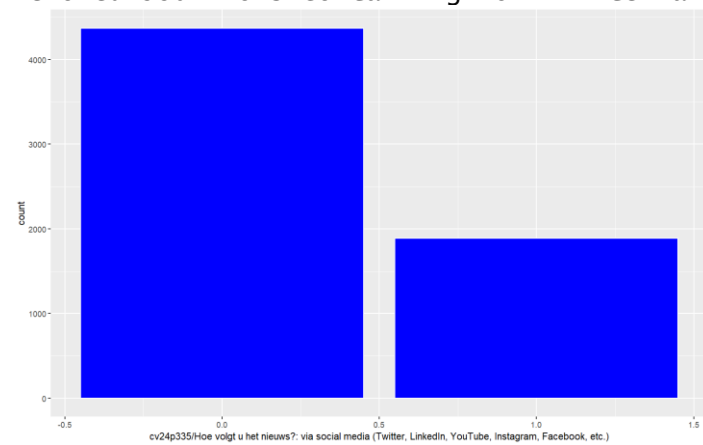
Column type frequency:	
numeric	1

Group variables	
None	

Variable type: numeric

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	5801	0.519	0.302	0.459	0	0	0	1	1	

```
> ggplot(data = LISS2, mapping = aes(x = cv24p335))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("cv24p335/Hoe volgt u het nieuws?: via social media (Twitter, Link
edIn, YouTube, Instagram, Facebook, etc.)")
Warning message:
Removed 5801 rows containing non-finite values (`stat_count()`).
```



R script willekeurig iemand per huishouden:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele cv24p335 waarbij 1 iemand per huishouden is gesele
cteed
> attributes(LISS3$cv24p335)
$label
[1] "Do you follow the news: More than one answer is possible. via social m
edia (Twitter, LinkedIn, YouTube, Instagram, Facebook, etc.)"

$format.spss
[1] "F2.0"

$display_width
[1] 4

$labels
I dont know      no      yes
      -9         0         1

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

> table(LISS3$cv24p335, useNA = "always")

 0    1 <NA>
2321 916 1951
> skim(LISS3$cv24p335)
— Data Summary —
Name      LISS3$cv24p335
Number of rows 5188
Number of columns 1
```

Column type frequency:
numeric

1

Group variables

None

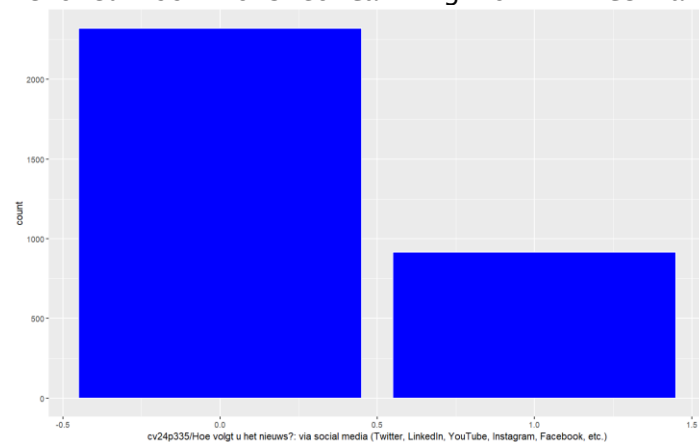
— Variable type: numeric —

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	1951	0.624	0.283	0.451	0	0	0	1	1	

```
> ggplot(data = LISS3, mapping = aes(x = cv24p335))+  
+   geom_bar(fill = "blue", colour = "white")+  
+   xlab("cv24p335/Hoe volgt u het nieuws?: via social media (Twitter, Link  
edIn, YouTube, Instagram, Facebook, etc.)")
```

Warning message:

Removed 1951 rows containing non-finite values (`stat\_count()`).



R script bewerkingen:

Deze variabele is niet bewerkt.

Uiteindelijke variabele:

De uiteindelijke variabele heeft ook 2 scores waarbij het grootste deel het nieuws niet volgt op sociale media. 76% volgt het nieuws niet op sociale media en 24% volgt het nieuws wel op sociale media.

R script uiteindelijke variabele:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de u  
iteindelijke variabele nieuws volgen op sociale media waarbij 1 iemand per  
huishouden is geselecteerd en de missende waarden zijn weggefilterd  
> attributes(LISS4$cv24p335)
```

```
$label
```

```
[1] "Do you follow the news: More than one answer is possible. via social m  
edia (Twitter, LinkedIn, YouTube, Instagram, Facebook, etc.)"
```

```
$format.spss
```

```
[1] "F2.0"
```

```
$display_width
```

```
[1] 4
```

```
$labels
```

```
I dont know  
-9
```

```
no  
0
```

```
yes  
1
```

```

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

> table(LISS4$cv24p335, useNA = "always")
 0    1 <NA>
2435 772    0
> skim(LISS4$cv24p335)
— Data Summary —
Name          LISS4$cv24p335
Number of rows 3207
Number of columns 1

Column type frequency:
  numeric      1

Group variables: None

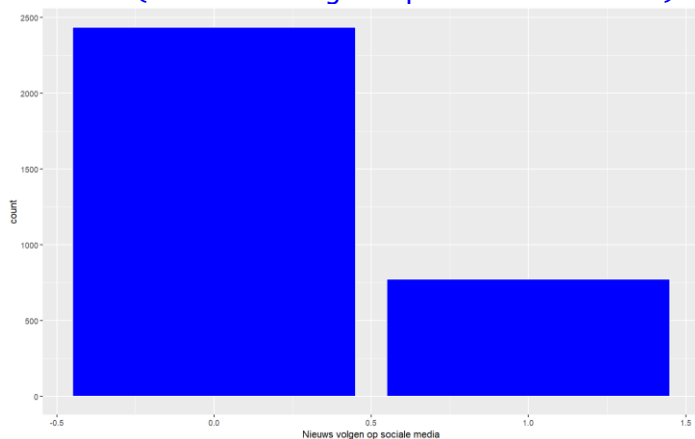
— variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 hist
1 data          0             1 0.241 0.428 0  0  0  0  0  1

```

```

> ggplot(data = LISS4, mapping = aes(x = cv24p335))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("Nieuws volgen op sociale media")

```



Het volgen van nieuws op traditionele media

Oorspronkelijke variabelen:

Het volgen van nieuws op traditionele media wordt gemeten met de vraag: Hoe volgt u het nieuws? Hierbij waren verschillende antwoordopties waarvan voor deze variabele 3 zijn geselecteerd. Op televisie en/of radio (cv24p002), via een gratis internetsite of app; zoals Nu.nl, NOS.nl, RTLnieuws.nl of Nieuws.nl (cv24p003) en in een betaald dagblad (digitaal, app en/of op papier) (cv24p166). Iemand die een 0 scoort op een van deze variabelen volgt het nieuws wel op het desbetreffende kanaal en iemand met een score van 1 volgt het nieuws wel op het desbetreffende kanaal.

R script oorspronkelijke variabelen:


```
> #Nieuws volgen op traditionele media
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabelen cv24p002, cv24p003 en cv24p166
> attributes(LISS2$cv24p002)
$label
[1] "Do you follow the news: More than one answer is possible. on televisio
n and/or radio"
```

```
$format.spss
```

```
[1] "F2.0"
```

```
$display_width
```

```
[1] 4
```

```
$class
```

```
[1] "haven_labelled" "vctrs_vctr"      "double"
```

```
$labels
```

```
I dont know      no      yes
      -9          0          1
```

```
> table(LISS2$cv24p002, useNA = "always")
```

```
  0    1 <NA>
1249 5009 5801
```

```
> skim(LISS2$cv24p002)
```

```
— Data Summary —
```

Name	Values
LISS2\$cv24p002	
Number of rows	12059
Number of columns	1

Column type frequency:	
numeric	1

Group variables	
None	

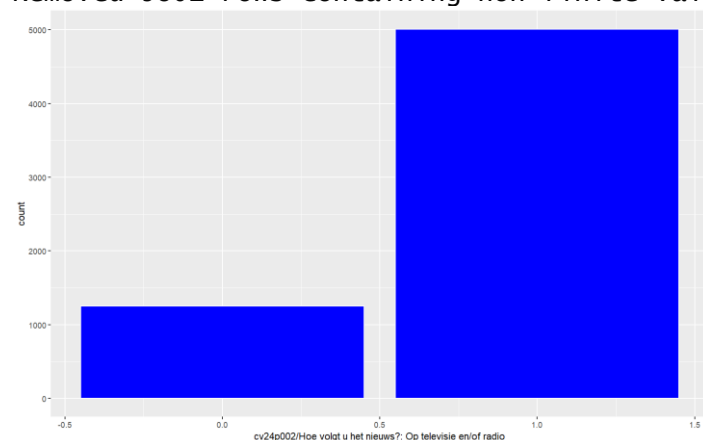
```
— variable type: numeric —
```

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	5801	0.519	0.800	0.400	0	1	1	1	1	___

```
> ggplot(data = LISS2, mapping = aes(x = cv24p002))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("cv24p002/Hoe volgt u het nieuws?: Op televisie en/of radio")
```

Warning message:

Removed 5801 rows containing non-finite values (`stat\_count()`).



```
> attributes(LISS2$cv24p003)
```

```
$label
[1] "Do you follow the news: More than one answer is possible. via a free i
nternet site or app; such as Nu.nl, NOS.nl, RTLnieuws.nl or Nieuws.nl"
```

```
$format.spss
[1] "F2.0"
```

```
$display_width
[1] 4
```

```
$class
[1] "haven_labelled" "vctrs_vctr"      "double"
```

```
$labels
I dont know      no      yes
      -9          0          1
```

```
> table(LISS2$cv24p003, useNA = "always")
```

```
  0    1 <NA>
2680 3578 5801
```

```
> skim(LISS2$cv24p003)
```

```
— Data Summary —————
Name                Values
Number of rows      LISS2$cv24p003
Number of columns    12059
Number of columns    1
```

```
Column type frequency:
  numeric              1
```

```
Group variables      None
```

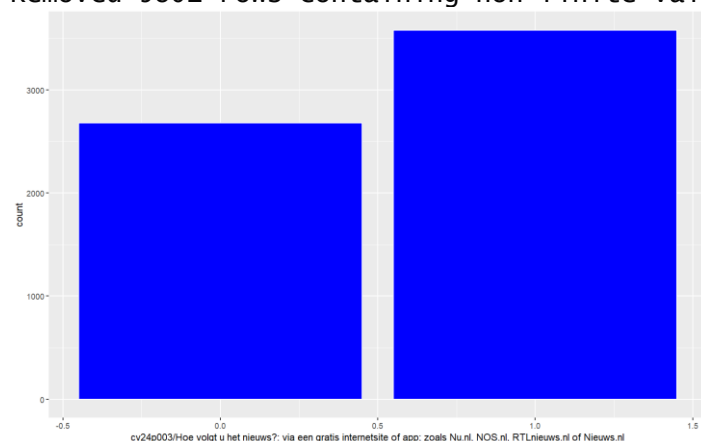
```
— Variable type: numeric —————
```

```
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 his
t
1 data          5801          0.519 0.572 0.495 0  0  1  1  1 1
```

```
> ggplot(data = LISS2, mapping = aes(x = cv24p003))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("cv24p003/Hoe volgt u het nieuws?: via een gratis internetsite of
app; zoals Nu.nl, NOS.nl, RTLnieuws.nl of Nieuws.nl")
```

Warning message:

Removed 5801 rows containing non-finite values (`stat\_count()`).



```
> attributes(LISS2$cv24p166)
```

```
$label
[1] "Do you follow the news: More than one answer is possible. in a paid da
ily newspaper (digital, app and/or on paper)"
```

```

$format.spss
[1] "F2.0"

$display_width
[1] 4

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

$labels
I dont know      no      yes
      -9          0        1

> table(LISS2$cv24p166, useNA = "always")

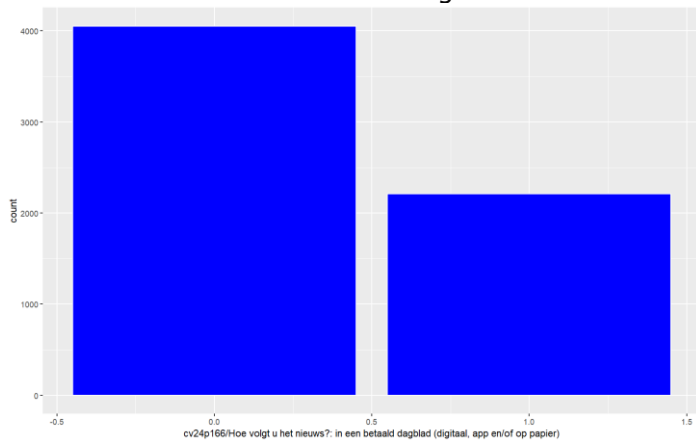
  0    1 <NA>
4050 2208 5801
> skim(LISS2$cv24p166)
— Data Summary —
Name                LISS2$cv24p166
Number of rows      12059
Number of columns    1

Column type frequency:
  numeric            1

Group variables      None

— Variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 histogram
1 data              5801          0.519 0.353 0.478 0  0  0  1  1 1
Warning message:
Removed 5801 rows containing non-finite values (`stat_count()`).

```



R script willekeurig iemand per huishouden:

```

> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabelen cv24p002, cv24p003 en cv24p166 waarbij 1 iemand p
er huishouden is geselecteerd
> attributes(LISS2$cv24p002)
$label
[1] "Do you follow the news: More than one answer is possible. on televisio
n and/or radio"

```

```

$format.spss
[1] "F2.0"

$display_width
[1] 4

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

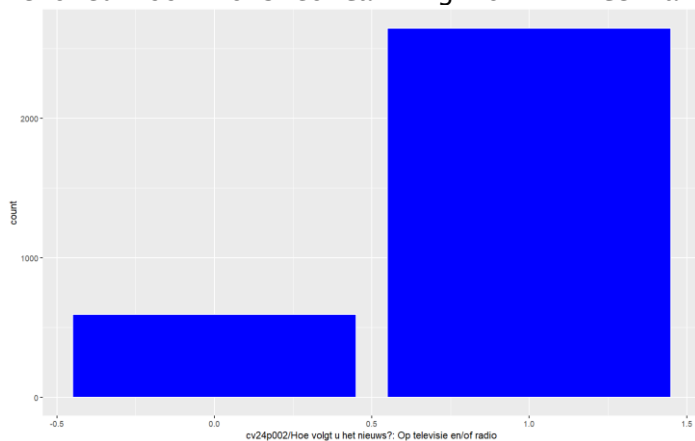
$labels
I dont know      no      yes
      -9          0        1

> table(LISS3$cv24p002, useNA = "always")

  0    1 <NA>
592 2645 1951
> skim(LISS3$cv24p002)
— Data Summary —
Name                               Values
Number of rows                     5188
Number of columns                   1
Column type frequency:
  numeric                           1
Group variables                      None
— Variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 his
t
1 data                1951          0.624 0.817 0.387 0  1  1  1  1  1

```

Warning message:
Removed 1951 rows containing non-finite values (`stat\_count()`).



```

> attributes(LISS2$cv24p003)
$label
[1] "Do you follow the news: More than one answer is possible. via a free i
nternet site or app; such as Nu.nl, NOS.nl, RTLnieuws.nl or Nieuws.nl"

$format.spss
[1] "F2.0"

$display_width
[1] 4

```

```
$class
[1] "haven_labelled" "vctrs_vctr"      "double"
```

```
$labels
I dont know      no      yes
      -9          0          1
```

```
> table(LISS3$cv24p003, useNA = "always")
```

```
  0    1 <NA>
1411 1826 1951
```

```
> skim(LISS3$cv24p003)
```

```
— Data Summary —
Name      LISS3$cv24p003
Number of rows      5188
Number of columns    1
```

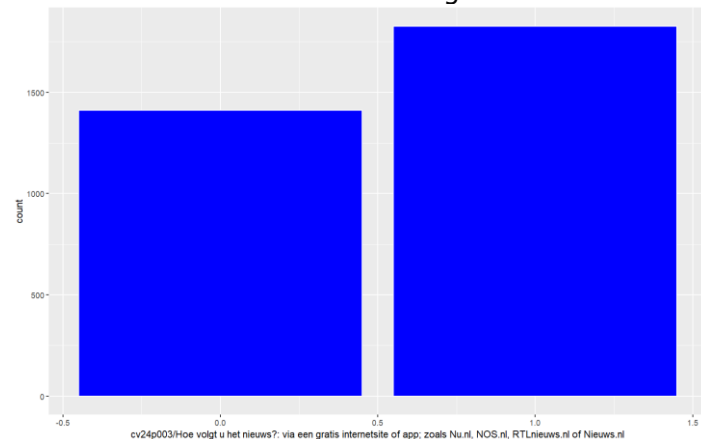
```
Column type frequency:
  numeric      1
```

```
Group variables      None
```

```
— Variable type: numeric —
```

```
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 histogram
1 data          1951          0.624 0.564 0.496 0  0  1  1  1  1
```

```
> ggplot(data = LISS3, mapping = aes(x = cv24p003))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("cv24p003/Hoe volgt u het nieuws?: via een gratis internetsite of
app; zoals Nu.nl, NOS.nl, RTLnieuws.nl of Nieuws.nl")
Warning message:
Removed 1951 rows containing non-finite values (`stat_count()`).
```



```
> attributes(LISS3$cv24p166)
```

```
$label
[1] "Do you follow the news: More than one answer is possible. in a paid daily newspaper (digital, app and/or on paper)"
```

```
$format.spss
[1] "F2.0"
```

```
$display_width
[1] 4
```

```
$labels
I dont know      no      yes
      -9          0          1
```

```

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

> table(LISS3$cv24p166, useNA = "always")

 0      1 <NA>
2023 1214 1951
> skim(LISS3$cv24p166)
— Data Summary —
Name      LISS3$cv24p166
Number of rows      5188
Number of columns    1

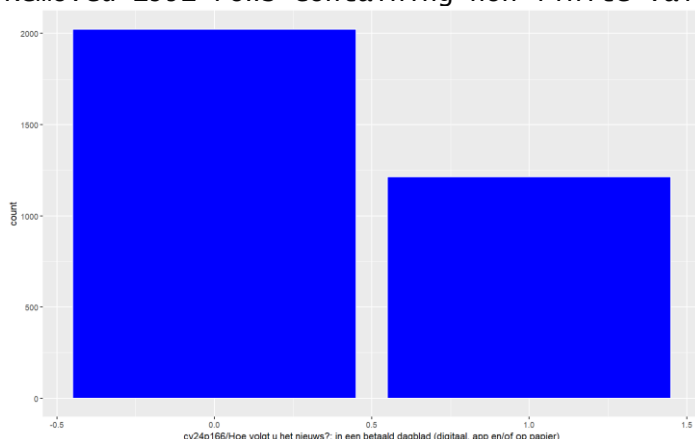
Column type frequency:
  numeric      1

Group variables      None

— Variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 hist
1 data          1951          0.624 0.375 0.484 0 0 0 1 1

```

Warning message:
Removed 1951 rows containing non-finite values (`stat\_count()`).



R script bewerkingen:

Deze 3 items zijn samengevoegd tot één variabele om weer te geven wie het nieuws volgt op traditionele media.

```

> #Samenvoegen van de verschillende variabelen van traditionele media
> LISS2 <- LISS2 %>%
+   mutate(traditionele_media_non_dummy = cv24p002 + cv24p003 + cv24p166)
> LISS2 <- LISS2 %>%
+   mutate(traditionele_media_dummy = case_when(
+     traditionele_media_non_dummy == 0 ~ 0,
+     traditionele_media_non_dummy >= 1 ~ 1
+   ))

```

Uiteindelijke variabele:

De uiteindelijke variabele (Traditionele_media_dummy) heeft 2 scores. Als iemand een 1 scoorde op één van de afzonderlijke items heeft deze persoon ook een score van 1 op de nieuwe variabele en als iemand op geen één van de afzonderlijke items een 1 scoorde heeft deze persoon een score van 0 op de nieuwe variabele. Veruit het grootste deel volgt het nieuws wel op traditionele media en maar een klein deel volgt het nieuws niet op traditionele media (0 = 6% ; 1 = 94%).

R script uiteindelijke variabele:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de u
uiteindelijke variabele nieuws volgen op traditionele media waarbij 1 iemand
per huishouden is geselecteerd en de missende waarden zijn weggefilterd
> attributes(LISS4$Traditionele_media_dummy)
NULL
> table(LISS4$Traditionele_media_dummy, useNA = "always")
```

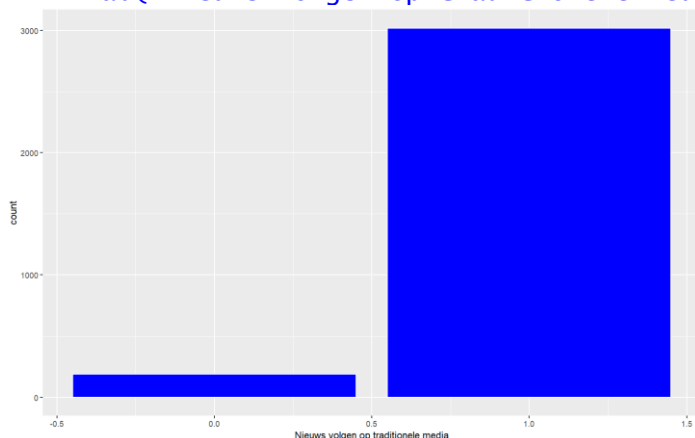
```
  0    1 <NA>
189 3018    0
> skim(LISS4$Traditionele_media_dummy)
— Data Summary —
Name                               Values
Number of rows                     3207
Number of columns                   1

Column type frequency:
  numeric                           1

Group variables                     None

— variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 his
t
1 data          0          1 0.941 0.236 0  1  1  1  1  1 __
```

```
> ggplot(data = LISS4, mapping = aes(x = Traditionele_media_dummy))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("Nieuws volgen op traditionele media")
```



Opleidingsniveau

Oorspronkelijke variabele:

Het opleidingsniveau wordt vastgesteld met de vraag: Opleiding in CBS-categorieën (oplcats). Deze variabele heeft 6 mogelijke scores namelijk 1 = basisonderwijs, 2 = vmbo, 3 = havo/vwo, 4 = mbo 5 = hbo en 6 = wo.

R script oorspronkelijke variabele:

```
> #Opleidingsniveau
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
oorspronkelijke variabele oplcat
> attributes(LISS2$oplcats)
$label
[1] "Level of education in categories"

$format.spss
[1] "F1.0"

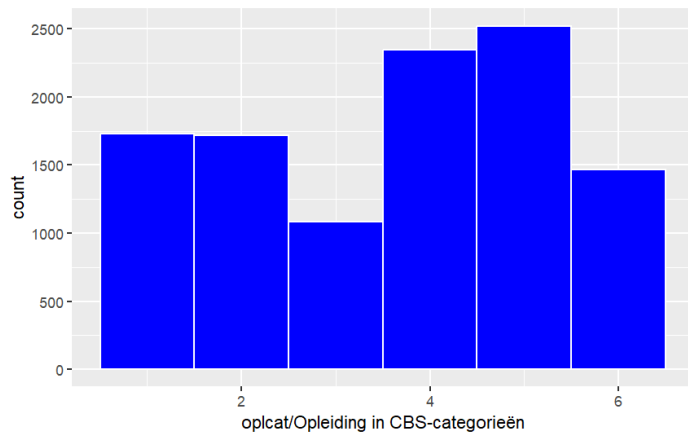
$class
[1] "haven_labelled" "vctrs_vctr"      "double"

$labels
primary school
1
junior high school)          vmbo (intermediate secondary education, US:
2
havo/vwo (higher secondary education/preparatory university education, US:
senior high school)
3
US: junior college)          mbo (intermediate vocational education,
4
ation, US: college)          hbo (higher vocational educ
5
wo (university)
6

> table(LISS2$oplcats, useNA = "always")
  1    2    3    4    5    6 <NA>
1728 1720 1082 2344 2523 1465 1197
> skim(LISS2$oplcats)
— Data Summary —
Name          Values
Number of rows LISS2$oplcats
Number of columns 1
Column type frequency:
  numeric      1
Group variables None
— variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 hist
1 data          1197          0.901 3.61 1.67 1 2 4 5 6
```



```
> ggplot(data = LISS2, mapping = aes(x = oplcat))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("oplcat/Opleiding in CBS-categorieën")
Warning message:
Removed 1197 rows containing non-finite values (`stat_bin()`)
```



R script willekeurig iemand per huishouden:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele oplcat waarbij 1 iemand per huishouden is geselect
eerd
> table(LISS3$oplcat, useNA = "always")
```

```
  1    2    3    4    5    6 <NA>
561  878  482 1078 1276  747  166
```

```
> skim(LISS3$oplcat)
```

— Data Summary —

Name	LISS3\$oplcat
Number of rows	5188
Number of columns	1

Column type frequency:	
numeric	1

Group variables	None
-----------------	------

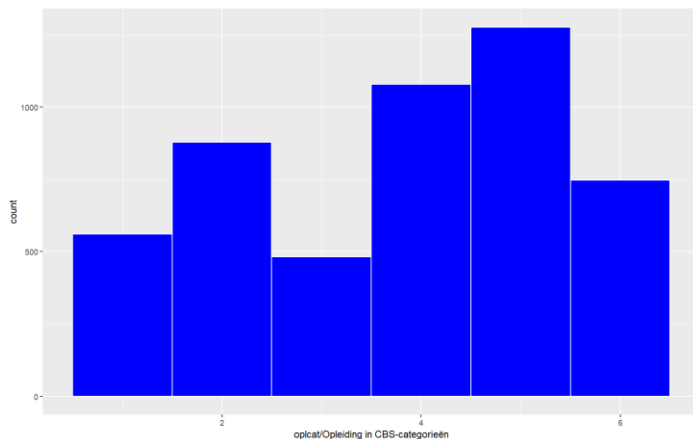
— Variable type: numeric —

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1 data	166	0.968	3.77	1.61	1	2	4	5	6	

```
> ggplot(data = LISS3, mapping = aes(x = oplcat))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("oplcat/Opleiding in CBS-categorieën")
```

Warning message:

Removed 166 rows containing non-finite values (`stat\_bin()`).



R script bewerkingen:

Deze variabele is niet bewerkt.

Uiteindelijke variabele:

De uiteindelijke variabele (oplcatt) is qua verdeling niet heel anders dan de oorspronkelijke variabele. Er is nog steeds te zien dat de grootste groep mensen een score van 5 heeft en dus het opleidingsniveau hbo heeft. Daarnaast is te zien er een kleiner deel is dat een score heeft van 1. Dit komt omdat iedereen die jonger is dan 16 eruit gefilterd is, waardoor een groot deel van deze groep niet meer in de steekproef zit.

R script uiteindelijke variabele:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de uiteindelijke variabele opleidingsniveau waarbij 1 iemand per huishouden is geselecteerd en de missende waardes zijn weggefilterd
> table(LISS4$oplcatt, useNA = "always")
```

```
  1    2    3    4    5    6 <NA>
161  588  302  777  925  454    0
```

```
> skim(LISS4$oplcatt)
```

```
— Data Summary —
```

Name	LISS4\$oplcatt
Number of rows	3207
Number of columns	1

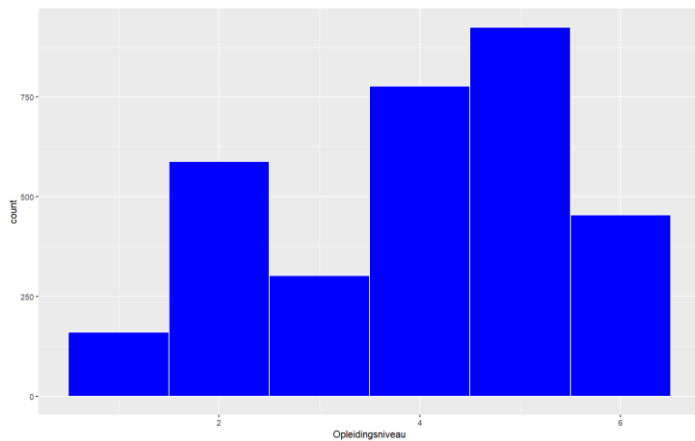
Column type frequency:	
numeric	1

Group variables	None
-----------------	------

```
— variable type: numeric —
```

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	0	1	3.96	1.46	1	3	4	5	6	

```
> ggplot(data = LISS4, mapping = aes(x = oplcatt))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("Opleidingsniveau")
```



Leeftijd

Oorspronkelijke variabele:

De variabele leeftijd wordt gemeten in jaren (leeftijd).

R script oorspronkelijke variabele:

```
> #Leeftijd
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele leeftijd
> attributes(LISS2$leeftijd)
$label
[1] "Age of the household member"
```

```
$format.spss
[1] "F2.0"
```

```
> table(LISS2$leeftijd, useNA = "always")
```

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
54	79	99	91	93	110	84	117	101	91	106	104	97	125	119
15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
118	120	136	149	144	143	139	161	157	136	147	148	108	147	159
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
145	151	179	157	160	143	164	156	124	158	138	113	120	151	150
45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
123	120	112	123	154	144	144	132	181	174	157	157	158	160	163
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
175	168	172	173	158	180	156	157	144	161	165	163	154	161	156
75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
136	154	122	89	93	74	62	51	65	48	32	29	20	25	11
90	91	92	93	94	95	96	97	98	103	106	110	119	123	<NA>
18	5	9	6	3	6	1	2	1	2	1	1	1	1	555

```
> skim(LISS2$leeftijd)
```

```
— Data Summary —
```

Name	values
Number of rows	LISS2\$leeftijd
Number of columns	12059

Column type frequency:	
numeric	1

Group variables	None
-----------------	------

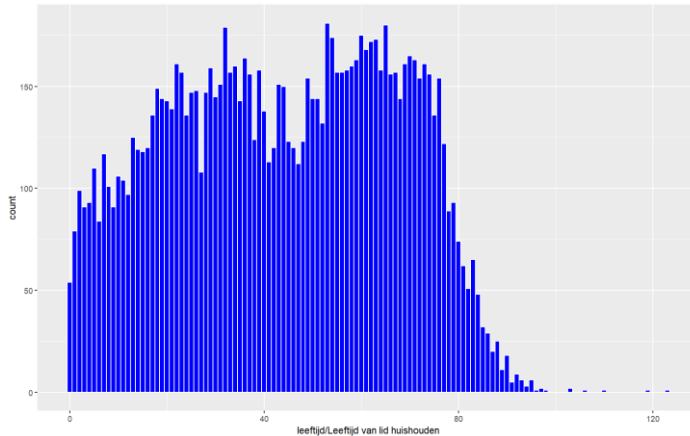
```
— variable type: numeric —
```

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
---------------	-----------	---------------	------	----	----	-----	-----	-----	------	------

```
1 data          555          0.954 43.8 23.1  0  25  44  63 123
```

```
> ggplot(data = LISS2, mapping = aes(x = leeftijd))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("leeftijd/Leefijd van lid huishouden")
```

Warning message:
Removed 555 rows containing non-finite values (`stat\_bin()`).



R script willekeurig iemand per huishouden:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele leeftijd waarbij 1 iemand per huishouden is gesele
cteerd
```

```
> attributes(LISS3$leeftijd)
$label
[1] "Age of the household member"
```

```
$format.spss
[1] "F2.0"
```

```
> table(LISS3$leeftijd, useNA = "always")
```

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
18	21	25	19	28	26	14	27	22	23	20	29	21	28	35
15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
27	25	33	50	55	53	48	69	67	58	67	76	54	70	92
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
73	81	89	91	78	70	66	75	52	63	59	50	46	65	61
45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
43	43	47	47	52	61	63	54	71	70	59	71	81	71	83
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74
88	85	100	104	73	93	88	89	76	97	94	109	90	100	97
75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
84	95	80	56	51	44	40	31	40	29	22	18	14	16	9
90	91	92	93	94	95	96	97	103	110	119	123	<NA>		
14	2	7	5	2	4	1	1	1	1	1	1	1		

```
> skim(LISS3$leeftijd)
```

```
— Data Summary —
Name          LISS3$leeftijd
Number of rows 5188
Number of columns 1
```

```
Column type frequency:
  numeric          1
```

```
Group variables      None
```

```
— variable type: numeric —
```

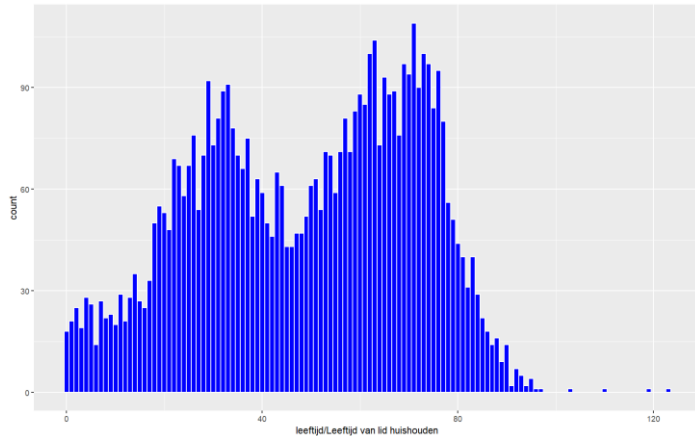
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
---------------	-----------	---------------	------	----	----	-----	-----	-----	------	------

```
1 data 1 1.00 49.0 22.2 0 31 52 68 123
```

```
> ggplot(data = LISS3, mapping = aes(x = leeftijd))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("leeftijd/Leef tijd van lid huishouden")
```

Warning message:

Removed 1 rows containing non-finite values (`stat\_bin()`).



R script bewerkingen:

De variabele leeftijd is niet bewerkt.

Uiteindelijke variabele:

De verdelingen van de uiteindelijke variabele (leeftijd) is erg veranderd vergeleken met de verdelingen van de oorspronkelijke variabele. De variabele had eerst een tweetoppige verdeling, maar is nu meer een links-scheve verdeling geworden. Het gemiddelde is van ongeveer 44 veranderd naar ongeveer 59. Dit is wel een redelijk groot verschil, wat deels komt door het wegfilteren van mensen onder 16 jaar, maar de eerste piek is ook min of meer weg en het lijkt er dus op dat mensen met een leeftijd van 16 tot ongeveer 50 vaker missende waarden hadden op 1 van de andere variabelen.

R script uiteindelijke variabele:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de u
iteindelijke variabele Leeftijd waarbij 1 iemand per huishouden is geselect
eerd en de missende waarden zijn weggefilterd
```

```
> attributes(LISS4$leeftijd)
```

```
$label
```

```
[1] "Age of the household member"
```

```
$format.spss
```

```
[1] "F2.0"
```

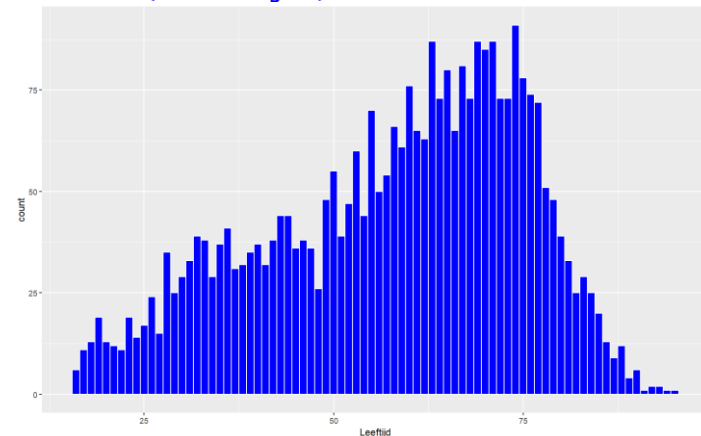
```
> table(LISS4$leeftijd, useNA = "always")
```

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
6	11	13	19	13	12	11	19	14	17	24	15	35	25	29
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
33	39	38	29	37	41	31	32	35	37	32	38	44	44	36
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
38	36	26	48	55	39	47	60	44	70	50	54	66	61	76
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75

```

65 63 87 73 80 65 81 73 87 85 87 73 73 91 78
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
74 72 51 48 39 33 25 29 25 20 13 9 12 4 6
91 92 93 94 95 <NA>
1 2 2 1 1 0
> skim(LISS4$leeftijd)
— Data Summary —
Name                               values
Number of rows                    LISS4$leeftijd
Number of columns                  3207
Column type frequency:
  numeric                          1
Group variables                    None
— Variable type: numeric —
skim_variable n_missing complete_rate mean sd p0 p25 p50 p75 p100 hist
1 data        0              1 57.8 17.1 16 45 61 71 95

```



Geslacht

Oorspronkelijke variabele:

De variabele geslacht (geslacht) heeft 3 mogelijke scores: 1 = man, 2 = vrouw en 3 = anders.

R script oorspronkelijke variabele:

```

> #Geslacht
> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele geslacht
> attributes(LISS2$geslacht)
$label
[1] "Gender"

$format.spss
[1] "F1.0"

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

```

```

$labels
  Male Female  Other
    1     2     3

> table(LISS2$geslacht, useNA = "always")

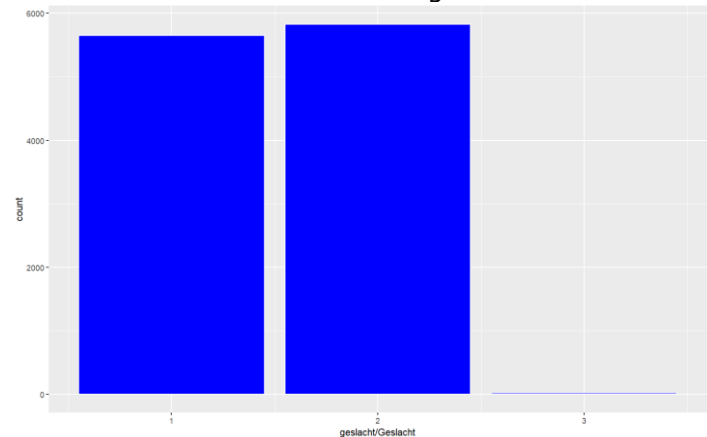
  1     2     3 <NA>
5654 5828  22  555
> skim(LISS2$geslacht)
— Data Summary —
Name      LISS2$geslacht
Number of rows 12059
Number of columns 1

Column type frequency:
  numeric      1

Group variables      None

— Variable type: numeric —
skim_variable n_missing complete_rate mean    sd p0 p25 p50 p75 p100 hist
1 data          555          0.954 1.51 0.504  1  1  2  2  3
Warning message:
Removed 555 rows containing non-finite values (`stat_count()`).

```



R script willekeurig iemand per huishouden:

```

> #Hieronder staan de attributes, frequentieverdeling en histogram van de o
orspronkelijke variabele geslacht waarbij 1 iemand per huishouden is gesele
cteed
> attributes(LISS3$geslacht)
$label
[1] "Gender"

$format.spss
[1] "F1.0"

$labels
  Male Female  Other
    1     2     3

$class
[1] "haven_labelled" "vctrs_vctr"      "double"

> table(LISS3$geslacht, useNA = "always")

```

```

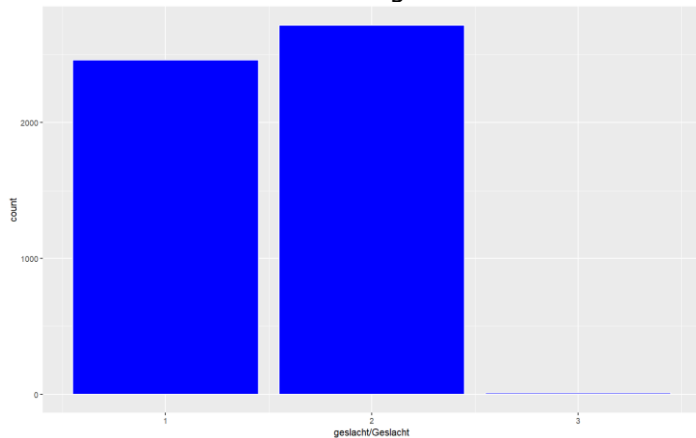
1      2      3 <NA>
2461 2717      9      1
> skim(LISS3$geslacht)
— Data Summary —
Name                               Values
Number of rows                     5188
Number of columns                   1
Column type frequency:
  numeric                           1
Group variables                     None
— Variable type: numeric —
skim_variable n_missing complete_rate mean    sd p0 p25 p50 p75 p100 hist
1 data        1          1.00 1.53 0.503  1  1  2  2  3  █

```

```

> ggplot(data = LISS3, mapping = aes(x = geslacht))+
+   geom_bar(fill = "blue", colour = "white")+
+   xlab("geslacht/Geslacht")
Warning message:
Removed 1 rows containing non-finite values (`stat_count()`).

```



R script bewerkingen:

De score 3 = anders is eruit gefilterd door er missende waarden van te maken. Dit is gedaan omdat het de analyse makkelijker maakt en er maar een kleine hoeveelheid respondenten deze score had (Oorspronkelijke variabele = 22; Willekeurig iemand per huishouden variabele = 9). De nieuwe scores zijn 0 = man en 1 = vrouw.

```

> #Geslacht hercoderen tot een dummy
> LISS2 <- LISS2 %>%
+   mutate(Geslacht_dummy = case_when(
+     geslacht == 1 ~ 0,
+     geslacht == 2 ~ 1,
+     geslacht == 3 ~ NA_real_
+   ))

```

Uiteindelijke variabele:

De uiteindelijke variabele (Geslacht_dummy) is niet heel erg veranderd van de oorspronkelijke variabelen. Er zijn ongeveer net zoveel mannen als vrouwen. Er zijn wel iets

meer vrouwen dan mannen namelijk 21 vrouwen meer. De score 3 = anders wordt niet meer meegerekend en is nu een missende waarde.

R script uiteindelijke variabele:

```
> #Hieronder staan de attributes, frequentieverdeling en histogram van de uiteindelijke variabele Geslacht waarbij 1 iemand per huishouden is geselecteerd en de missende waarden zijn weggefilterd
> attributes(LISS4$Geslacht_dummy)
NULL
> table(LISS4$Geslacht_dummy, useNA = "always")
```

```
  0    1 <NA>
1593 1614    0
```

```
> skim(LISS4$Geslacht_dummy)
```

— Data Summary —

	Values
Name	LISS4\$Geslacht_dummy
Number of rows	3207
Number of columns	1

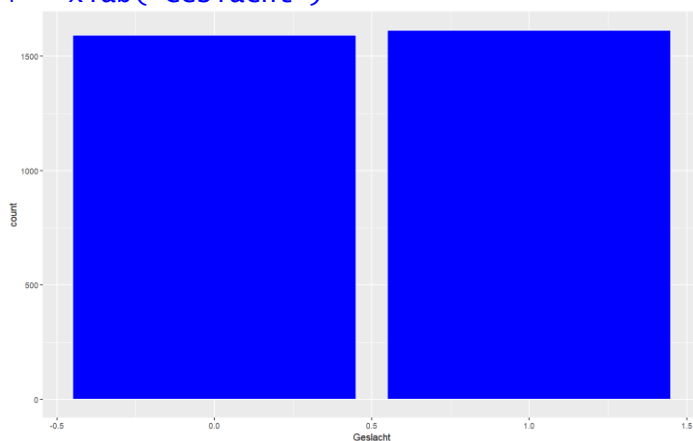
Column type frequency:	
numeric	1

Group variables	
	None

— Variable type: numeric —

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
1	data	0	1	0.503	0.500	0	0	1	1	1	█

```
> ggplot(data = LISS4, mapping = aes(x = Geslacht_dummy)) +
+   geom_bar(fill = "blue", colour = "white") +
+   xlab("Geslacht")
```



Bijlage 2

Univariate statistieken

```
> #Resultaten en bijlage 2 en 3
> #Univariate statistieken
> summary(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 0.0000  0.0000  0.3524  1.0000  1.0000
> prop.table(table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy))

      0      1
0.6476458 0.3523542
> summary(LISS4$cv24p335)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 0.0000  0.0000  0.2407  0.0000  1.0000
> prop.table(table(LISS4$cv24p335))

      0      1
0.7592766 0.2407234
> summary(LISS4$Traditionele_media_dummy)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 1.0000  1.0000  0.9411  1.0000  1.0000
> prop.table(table(LISS4$Traditionele_media_dummy))

      0      1
0.05893358 0.94106642
> summary(LISS4$oplcat)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.00   3.00   4.00   3.96   5.00   6.00
> sd(LISS4$oplcat)
[1] 1.460532
> summary(LISS4$leeftijd)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 16.00  45.00  61.00  57.83  71.00  95.00
> sd(LISS4$leeftijd)
[1] 17.06418
> summary(LISS4$Geslacht_dummy)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 0.0000  1.0000  0.5033  1.0000  1.0000
> prop.table(table(LISS4$Geslacht_dummy))

      0      1
0.4967259 0.5032741
```

Missende waarden analyse²

Hieronder in tabel 4 zijn de missende waarden per variabele te zien. In de methoden werd al gezegd dat de vragenlijst over ‘Democratie in Nederland’ is voorgelegd aan 6552 huishoudleden (Verheijen, 2024). Daarvan hebben 5134 huishoudleden gereageerd waardoor al een groot deel van de 12059 respondenten die in de oorspronkelijke dataset zitten niet mee doen. In het R script hieronder is te zien dat de proporties en daarmee de verdelingen van de variabelen niet heel veel veranderen na het wegfilteren van de missende waarden en het willekeurig selecteren van één iemand per huishouden behalve bij de variabelen het volgen van nieuws op sociale media, het volgen van nieuws op traditionele media en voornamelijk leeftijd. Ten eerste is te zien dat er in de oorspronkelijke steekproef iets meer respondenten waren die het nieuws wel volgen op sociale media en iets meer respondenten waren die het nieuws niet volgen op traditionele media. De proporties zijn niet heel erg veranderd maar verschillen wel ongeveer 5%. Ten slotte is de variabele leeftijd

veranderd. Alle respondenten die jonger zijn dan 16 zijn eruit gefilterd, maar er is ook duidelijk te zien dat er veel minder respondenten met de leeftijd van 16 tot ongeveer 50 zijn.

Tabel 4: Missende waarden analyse per variabele.

	Missing	N gefilterde dataset	N totaal
Je wel of niet vertegenwoordigd voelen door de politiek (Niet = 0; Wel = 1)	7371 (weet ik niet = 444)	3207	12059
Nieuws volgen op sociale media (Nee = 0; Ja = 1)	5801	3207	12059
Nieuws volgen op traditionele media (Nee = 0; Ja = 1)	5801	3207	12059
Opleidingsniveau	1197	3207	12059
Leeftijd	555	3207	12059
Geslacht (Man = 0; Vrouw = 1)	577 (anders = 22)	3207	12059

R script Missende waarden analyse:

```
> #Missende waarden
> table(LISS2$yg23a023, useNA = "always")

  1      2      3      4      5      6 <NA>
651 2376 1574   79      8  444 6927
> prop.table(table(LISS2$yg23a023))

      1      2      3      4      5      6
0.126851130 0.462977397 0.306703040 0.015393609 0.001558846 0.086515978
> prop.table(table(LISS2$Gevoel_Politieke_vertegenwoordiging_dummy))

      0      1
0.6456911 0.3543089
> table(LISS4$yg23a023, useNA = "always")

  1      2      3      4      5 <NA>
442 1635 1072   54      4      0
> prop.table(table(LISS4$yg23a023))

      1      2      3      4      5
0.137823511 0.509822264 0.334268787 0.016838167 0.001247272
> prop.table(table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy))

      0      1
0.6476458 0.3523542
> table(LISS2$cv24p335, useNA = "always")

  0      1 <NA>
4369 1889 5801
> prop.table(table(LISS2$cv24p335))

      0      1
0.6981464 0.3018536
> table(LISS4$cv24p335, useNA = "always")

  0      1 <NA>
2435  772      0
> prop.table(table(LISS4$cv24p335))

      0      1
```

```

0.7592766 0.2407234
> table(LISS2$Traditionele_media_dummy, useNA = "always")

  0    1 <NA>
578 5680 5801
> prop.table(table(LISS2$Traditionele_media_dummy))

  0    1
0.09236178 0.90763822
> table(LISS4$Traditionele_media_dummy, useNA = "always")

  0    1 <NA>
189 3018    0
> prop.table(table(LISS4$Traditionele_media_dummy))

  0    1
0.05893358 0.94106642
> table(LISS2$oplcat, useNA = "always")

  1    2    3    4    5    6 <NA>
1728 1720 1082 2344 2523 1465 1197
> prop.table(table(LISS2$oplcat))

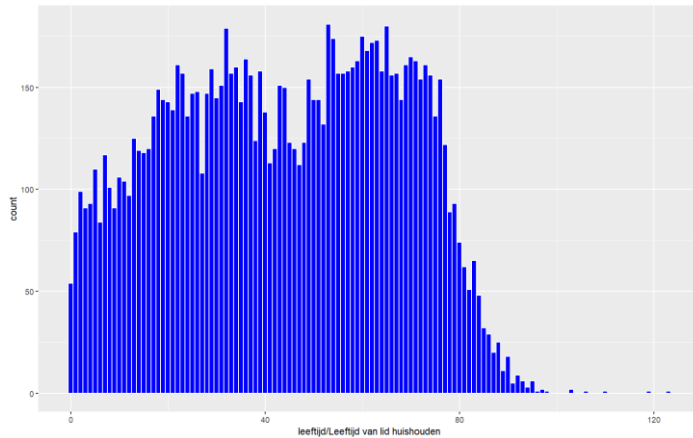
  1    2    3    4    5    6
0.15908672 0.15835021 0.09961333 0.21579820 0.23227767 0.13487387
> table(LISS4$oplcat, useNA = "always")

  1    2    3    4    5    6 <NA>
161  588  302  777  925  454    0
> prop.table(table(LISS4$oplcat))

  1    2    3    4    5    6
0.05020268 0.18334892 0.09416901 0.24228251 0.28843156 0.14156533
> table(LISS2$leeftijd, useNA = "always")

  0    1    2    3    4    5    6    7    8    9    10    11    12    13    14
 54   79   99   91   93  110   84  117  101   91  106  104   97  125  119
 15   16   17   18   19   20   21   22   23   24   25   26   27   28   29
118  120  136  149  144  143  139  161  157  136  147  148  108  147  159
 30   31   32   33   34   35   36   37   38   39   40   41   42   43   44
145  151  179  157  160  143  164  156  124  158  138  113  120  151  150
 45   46   47   48   49   50   51   52   53   54   55   56   57   58   59
123  120  112  123  154  144  144  132  181  174  157  157  158  160  163
 60   61   62   63   64   65   66   67   68   69   70   71   72   73   74
175  168  172  173  158  180  156  157  144  161  165  163  154  161  156
 75   76   77   78   79   80   81   82   83   84   85   86   87   88   89
136  154  122   89   93   74   62   51   65   48   32   29   20   25   11
 90   91   92   93   94   95   96   97   98  103  106  110  119  123 <NA>
 18    5    9    6    3    6    1    2    1    2    1    1    1    1  555
> ggplot(data = LISS2, mapping = aes(x = leeftijd))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("leeftijd/Leeftijd van lid huishouden")
Warning message:
Removed 555 rows containing non-finite values (`stat_bin()`).

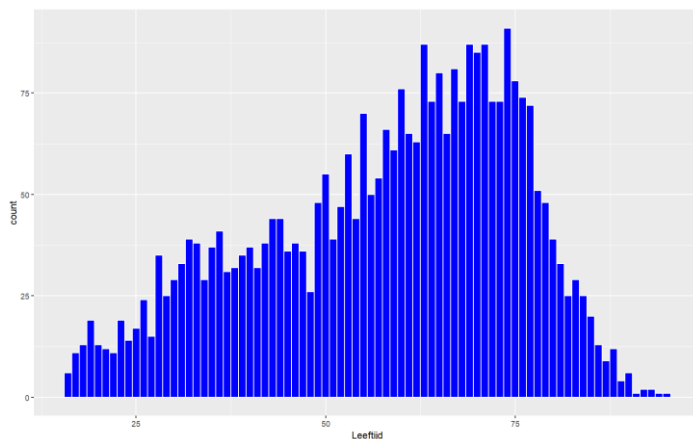
```



```
> table(LISS4$leeftijd, useNA = "always")
```

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
6	11	13	19	13	12	11	19	14	17	24	15	35	25	29
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
33	39	38	29	37	41	31	32	35	37	32	38	44	44	36
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
38	36	26	48	55	39	47	60	44	70	50	54	66	61	76
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
65	63	87	73	80	65	81	73	87	85	87	73	73	91	78
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
74	72	51	48	39	33	25	29	25	20	13	9	12	4	6
91	92	93	94	95	<NA>									
1	2	2	1	1	0									

```
> ggplot(data = LISS4, mapping = aes(x = leeftijd))+
+   geom_histogram(fill = "blue", colour = "white", binwidth = 1)+
+   xlab("Leeftijd")
```



```
> table(LISS2$geslacht, useNA = "always")
```

1	2	3	<NA>
5654	5828	22	555

```
> prop.table(table(LISS2$geslacht))
```

1	2	3
0.491481224	0.506606398	0.001912378

```
> table(LISS4$geslacht, useNA = "always")
```

1	2	<NA>
1593	1614	0

```
> prop.table(table(LISS4$geslacht))
```

1	2
0.4967259	0.5032741

Verhouding tot de populatie³

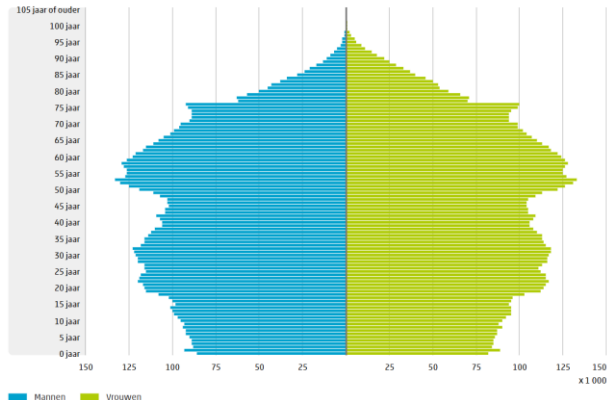
De verdeling van opleidingsniveau is vergelijkbaar met de verdeling in de populatie, maar het lijkt erop dat de respondenten vaker het opleidingsniveau hbo of wo hebben dan in de populatie het geval is. In de populatie heeft 36,4% hbo of wo gedaan terwijl dat in de steekproef 43,0% is (Centraal Bureau voor de Statistiek, 2024).

In de populatie is er meer sprake van een tweetoppige verdeling en verschilt dus ook wel redelijk van de verdeling van leeftijd in de steekproef (Centraal Bureau voor de Statistiek, n.d.-a). In de figuren hieronder is te zien dat de verdeling in de populatie meer lijkt op de verdeling van de oorspronkelijke variabele (Centraal Bureau voor de Statistiek, n.d.-b). De oudere respondenten zijn hierdoor oververtegenwoordigd in de steekproef en een grote groep mensen van de leeftijden van 16 tot ongeveer 55 zijn ondervertegenwoordigd in de steekproef. Hierdoor moet er wat voorzichtiger gedaan met het generaliseren van de resultaten naar de populatie.

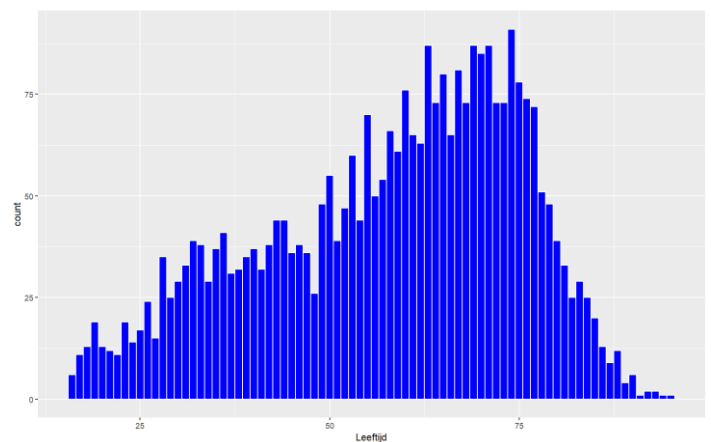
De verdeling van mannen en vrouwen in de populatie is erg vergelijkbaar met de verdeling in de steekproef. In 2022 waren er 8 745 468 mannen en 8 845 204 vrouwen in Nederland wat betekent dat 49,72% man is en 50,28% vrouw is. In de steekproef is 49,67% man en 50,33% vrouw (Centraal Bureau voor de Statistiek, n.d.-c).

Leeftijdsopbouw Nederland 2023

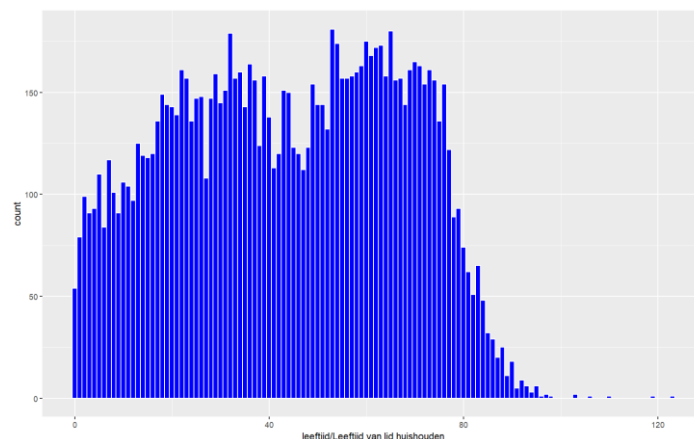
Totaal: 17 811 000 inwoners



Leeftijdsopbouw Nederland 2023



Verdeling Leeftijd gefilterde dataset



Verdeling Leeftijd oorspronkelijke dataset

R script verhouding tot de populatie:

```
> #Verhouding tot de populatie
> prop.table(table(LISS4$oplcat))
```

```

      1      2      3      4      5      6
0.05020268 0.18334892 0.09416901 0.24228251 0.28843156 0.14156533
> prop.table(table(LISS4$leeftijd))
```

```

      16      17      18      19      20
21 0.0018709074 0.0034299969 0.0040536327 0.0059245401 0.0040536327 0.00374181
48
      22      23      24      25      26
27 0.0034299969 0.0059245401 0.0043654506 0.0053009043 0.0074836296 0.00467726
85
      28      29      30      31      32
33 0.0109136264 0.0077954475 0.0090427191 0.0102899906 0.0121608980 0.01184908
01
      34      35      36      37      38
39 0.0090427191 0.0115372622 0.0127845338 0.0096663548 0.0099781727 0.01091362
64
      40      41      42      43      44
45 0.0115372622 0.0099781727 0.0118490801 0.0137199875 0.0137199875 0.01122544
43
      46      47      48      49      50
51 0.0118490801 0.0112254443 0.0081072654 0.0149672591 0.0171499844 0.01216089
80
      52      53      54      55      56
57 0.0146554412 0.0187090739 0.0137199875 0.0218272529 0.0155908949 0.01683816
65
      58      59      60      61      62
63 0.0205799813 0.0190208918 0.0236981603 0.0202681634 0.0196445276 0.02712815
72
      64      65      66      67      68
69 0.0227627066 0.0249454319 0.0202681634 0.0252572498 0.0227627066 0.02712815
72
      70      71      72      73      74
75 0.0265045214 0.0271281572 0.0227627066 0.0227627066 0.0283754287 0.02432179
61
      76      77      78      79      80
81 0.0230745245 0.0224508887 0.0159027128 0.0149672591 0.0121608980 0.01028999
06
      82      83      84      85      86
87 0.0077954475 0.0090427191 0.0077954475 0.0062363580 0.0040536327 0.00280636
11
      88      89      90      91      92
93 0.0037418148 0.0012472716 0.0018709074 0.0003118179 0.0006236358 0.00062363
58
      94      95
0.0003118179 0.0003118179
> prop.table(table(LISS4$Geslacht_dummy))
```

```

      0      1
```

0.4967259 0.5032741

Bivariate statistieken

```
> #Bivariate statistieken
> #Je wel of niet vertegenwoordigd voelen door de politiek
> tabely_x1 <- table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, LISS4
$cv24p335)
> phi(tabely_x1, digits = 3)
[1] -0.029
> chisq.test(tabely_x1)
```

Pearson's Chi-squared test with
Yates' continuity correction

```
data: tabely_x1
X-squared = 2.5635, df = 1,
p-value = 0.1094
```

```
> tabely_x2 <- table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, LISS4
$Traditionele_media_dummy)
> phi(tabely_x2, digits = 3)
[1] 0.04
> chisq.test(tabely_x2)
```

Pearson's Chi-squared test with
Yates' continuity correction

```
data: tabely_x2
X-squared = 4.8947, df = 1,
p-value = 0.02694
```

```
> cor.test(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, LISS4$oplcat)
```

Pearson's product-moment
correlation

```
data: LISS4$Gevoel_Politieke_vertegenwoordiging_dummy and LISS4$oplcat
t = 12.141, df = 3205, p-value
< 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1763520 0.2425362
sample estimates:
      cor
0.2096843
```

```
> cor.test(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, LISS4$leeftijd)
```

Pearson's product-moment
correlation

```
data: LISS4$Gevoel_Politieke_vertegenwoordiging_dummy and LISS4$leeftijd
t = -3.5368, df = 3205, p-value
= 0.0004107
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.09675603 -0.02780058
sample estimates:
      cor
-0.06235271
```

```
> tabely_c3 <- table(LISS4$Gevoel_Politieke_vertegenwoordiging_dummy, LISS4
$Geslacht_dummy)
> phi(tabely_c3, digits = 3)
[1] -0.015
> chisq.test(tabely_c3)
```

Pearson's Chi-squared test with


```

      Yates' continuity correction

data:  tabelx1_c3
X-squared = 0.68561, df = 1,
p-value = 0.4077

>
> #Nieuws volgen op sociale media
> tabelx1_x2 <- table(LISS4$cv24p335, LISS4$Traditionele_media_dummy)
> phi(tabelx1_x2, digits = 3)
[1] -0.051
> chisq.test(tabelx1_x2)

      Pearson's Chi-squared test with
      Yates' continuity correction

data:  tabelx1_x2
X-squared = 7.878, df = 1,
p-value = 0.005004

> cor.test(LISS4$cv24p335, LISS4$oplcat)

      Pearson's product-moment
      correlation

data:  LISS4$cv24p335 and LISS4$oplcat
t = 2.2869, df = 3205, p-value
= 0.02227
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.005758369 0.074869988
sample estimates:
      cor
0.04036245

> cor.test(LISS4$cv24p335, LISS4$leeftijd)

      Pearson's product-moment
      correlation

data:  LISS4$cv24p335 and LISS4$leeftijd
t = -20.283, df = 3205, p-value
< 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3675976 -0.3062396
sample estimates:
      cor
-0.3372767

> tabelx1_c3 <- table(LISS4$cv24p335, LISS4$Geslacht_dummy)
> phi(tabelx1_c3, digits = 3)
[1] 0.011
> chisq.test(tabelx1_c3)

      Pearson's Chi-squared test with
      Yates' continuity correction

data:  tabelx1_c3
X-squared = 0.33176, df = 1,
p-value = 0.5646

>
> #Nieuws volgen op traditionele media
> cor.test(LISS4$Traditionele_media_dummy, LISS4$oplcat)

      Pearson's product-moment
      correlation

data:  LISS4$Traditionele_media_dummy and LISS4$oplcat

```

```

t = 3.7273, df = 3205, p-value
= 0.0001969
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.0311554 0.1000812
sample estimates:
      cor
0.06569669

> cor.test(LISS4$Traditionele_media_dummy, LISS4$leeftijd)

      Pearson's product-moment
      correlation

data:  LISS4$Traditionele_media_dummy and LISS4$leeftijd
t = 12.997, df = 3205, p-value
< 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1906279 0.2563901
sample estimates:
      cor
0.2237637

> tabelx2_c3 <- table(LISS4$Traditionele_media_dummy, LISS4$Geslacht_dummy)
> phi(tabelx2_c3, digits = 3)
[1] -0.087
> chisq.test(tabelx2_c3)

      Pearson's Chi-squared test with
      Yates' continuity correction

data:  tabelx2_c3
X-squared = 23.582, df = 1,
p-value = 1.197e-06

>
> #Opleidingsniveau
> cor.test(LISS4$leeftijd, LISS4$oplcat)

      Pearson's product-moment
      correlation

data:  LISS4$leeftijd and LISS4$oplcat
t = -11.337, df = 3205, p-value
< 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.2294104 -0.1628520
sample estimates:
      cor
-0.1963574

> cor.test(LISS4$oplcat, LISS4$Geslacht_dummy)

      Pearson's product-moment
      correlation

data:  LISS4$oplcat and LISS4$Geslacht_dummy
t = -2.0463, df = 3205, p-value
= 0.04081
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.070645603 -0.001511559
sample estimates:
      cor
-0.0361218

>
> #Leetijd

```

```

> cor.test(LISS4$leeftijd, LISS4$Geslacht_dummy)

Pearson's product-moment
correlation

data: LISS4$leeftijd and LISS4$Geslacht_dummy
t = -4.4624, df = 3205, p-value
= 8.384e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.11288455 -0.04408722
sample estimates:
cor
-0.07857944

> #T-toetsen bij de afhankelijke variabele
> var.test(cv24p335 ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LIS
S4)

F test to compare two variances

data: cv24p335 by Gevoel_Politieke_vertegenwoordiging_dummy
F = 1.0783, num df = 2076, denom df = 1129, p-value = 0.1527
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
0.9724078 1.1937523
sample estimates:
ratio of variances
1.078259

> t.test(cv24p335 ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4
, var.equal=TRUE)

Two Sample t-test

data: cv24p335 by Gevoel_Politieke_vertegenwoordiging_dummy
t = 1.6445, df = 3205, p-value = 0.1002
alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
95 percent confidence interval:
-0.004996612 0.056968269
sample estimates:
mean in group 0 mean in group 1
0.2498796 0.2238938

> LISS4 %>%
+ group_by(Gevoel_Politieke_vertegenwoordiging_dummy) %>%
+ summarise(sd_cv24p335 = sd(cv24p335))
# A tibble: 2 x 2
  Gevoel_Politieke_vertegenwoordiging_dummy sd_cv24p335
  <dbl> <dbl>
1 0 0.433
2 1 0.417
> var.test(Traditionele_media_dummy ~ Gevoel_Politieke_vertegenwoordiging_d
ummy, data = LISS4)

F test to compare two variances

data: Traditionele_media_dummy by Gevoel_Politieke_vertegenwoordiging_dumm
y
F = 1.4028, num df = 2076, denom df = 1129, p-value =
2.191e-10
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
1.265126 1.553101
sample estimates:
ratio of variances
1.402842

```

```
> t.test(Traditionele_media_dummy ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4, var.equal=FALSE)
```

Welch Two Sample t-test

```
data: Traditionele_media_dummy by Gevoel_Politieke_vertegenwoordiging_dummy
t = -2.4085, df = 2665.6, p-value = 0.01609
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -0.03617902 -0.00370662
sample estimates:
mean in group 0 mean in group 1
 0.9340395      0.9539823
```

```
> LISS4 %>%
+   group_by(Gevoel_Politieke_vertegenwoordiging_dummy) %>%
+   summarise(sd_Traditionele_media_dummy = sd(Traditionele_media_dummy))
# A tibble: 2 x 2
  Gevoel_Politieke_vertegenwoordiging_dummy sd_Traditionele_media_d...1
  <dbl> <dbl>
1      0 0.248
2      1 0.210
# i abbreviated name: 'sd_Traditionele_media_dummy'
> var.test(oplcat ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4)
)
```

F test to compare two variances

```
data: oplcat by Gevoel_Politieke_vertegenwoordiging_dummy
F = 1.0514, num df = 2076, denom df = 1129, p-value = 0.3414
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.948214 1.164051
sample estimates:
ratio of variances
 1.051432
```

```
> t.test(oplcat ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4, var.equal=TRUE)
```

Two Sample t-test

```
data: oplcat by Gevoel_Politieke_vertegenwoordiging_dummy
t = -12.141, df = 3205, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -0.7445080 -0.5374704
sample estimates:
mean in group 0 mean in group 1
 3.734232      4.375221
```

```
> LISS4 %>%
+   group_by(Gevoel_Politieke_vertegenwoordiging_dummy) %>%
+   summarise(sd_oplcat = sd(oplcat))
# A tibble: 2 x 2
  Gevoel_Politieke_vertegenwoordiging_dummy sd_oplcat
  <dbl> <dbl>
1      0 1.44
2      1 1.41
> var.test(leeftijd ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4)
```

F test to compare two variances

```
data: leeftijd by Gevoel_Politieke_vertegenwoordiging_dummy
F = 0.89976, num df = 2076, denom df = 1129, p-value =
```

```

0.04188
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.8114330 0.9961355
sample estimates:
ratio of variances
 0.8997616

> t.test(leeftijd ~ Gevoel_Politieke_vertegenwoordiging_dummy, data = LISS4
, var.equal=FALSE)

welch Two Sample t-test

data: leeftijd by Gevoel_Politieke_vertegenwoordiging_dummy
t = 3.4821, df = 2216.1, p-value = 0.0005071
alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
95 percent confidence interval:
 0.972808 3.481134
sample estimates:
mean in group 0 mean in group 1
 58.61724      56.39027

> LISS4 %>%
+   group_by(Gevoel_Politieke_vertegenwoordiging_dummy) %>%
+   summarise(sd_leeftijd = sd(leeftijd))
# A tibble: 2 × 2
  Gevoel_Politieke_vertegenwoordiging_dummy sd_leeftijd
  <dbl> <dbl>
1      0      16.7
2      1      17.6

> var.test(Geslacht_dummy ~ Gevoel_Politieke_vertegenwoordiging_dummy, data =
LISS4)

F test to compare two variances

data: Geslacht_dummy by Gevoel_Politieke_vertegenwoordiging_dummy
F = 0.99948, num df = 2076, denom df = 1129, p-value = 0.9879
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.9013617 1.1065342
sample estimates:
ratio of variances
 0.9994795

> t.test(Geslacht_dummy ~ Gevoel_Politieke_vertegenwoordiging_dummy, data =
LISS4, var.equal=TRUE)

Two Sample t-test

data: Geslacht_dummy by Gevoel_Politieke_vertegenwoordiging_dummy
t = 0.86481, df = 3205, p-value = 0.3872
alternative hypothesis: true difference in means between group 0 and group
1 is not equal to 0
95 percent confidence interval:
-0.02025840 0.05223185
sample estimates:
mean in group 0 mean in group 1
 0.5089071      0.4929204

> LISS4 %>%
+   group_by(Gevoel_Politieke_vertegenwoordiging_dummy) %>%
+   summarise(sd_Geslacht_dummy = sd(Geslacht_dummy))
# A tibble: 2 × 2
  Gevoel_Politieke_vertegenwoordiging_dummy sd_Geslacht_dummy
  <dbl> <dbl>
1      0      0.500
2      1      0.500

```

Regressieanalyse

```
> #Modelevaluatie
> #Logistische regressie
> #Leeg model
> LR_model0 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ 1,
+                  family = binomial(link = "logit"),
+                  data = LISS4)
> summary(LR_model0)
```

```
Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ 1,
     family = binomial(link = "logit"), data = LISS4)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.60871	0.03697	-16.47	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4162 on 3206 degrees of freedom
Residual deviance: 4162 on 3206 degrees of freedom
AIC: 4164

Number of Fisher Scoring iterations: 4

```
> #Modelevaluatie
> #Logistische regressie
> #Leeg model
> LR_model0 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ 1,
+                  family = binomial(link = "logit"),
+                  data = LISS4)
> summary(LR_model0)
```

```
Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ 1,
     family = binomial(link = "logit"), data = LISS4)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.60871	0.03697	-16.47	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4162 on 3206 degrees of freedom
Residual deviance: 4162 on 3206 degrees of freedom
AIC: 4164

Number of Fisher Scoring iterations: 4

```
>
> #Model 1
> LR_model1 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd + Geslacht_dummy,
+                  family = binomial(link = "logit"),
+                  data = LISS4)
> summary(LR_model1)
```

```
Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat +
     leeftijd + Geslacht_dummy, family = binomial(link = "logit"),
     data = LISS4)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
--	----------	------------	---------	----------

```

(Intercept)    -1.691050    0.200506    -8.434    <2e-16 ***
oplcat         0.313039    0.027947    11.201    <2e-16 ***
leeftijd       -0.002950    0.002272    -1.298    0.194
Geslacht_dummy -0.041927    0.076064    -0.551    0.581
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(Dispersion parameter for binomial family taken to be 1)

```

Null deviance: 4162 on 3206 degrees of freedom
Residual deviance: 4014 on 3203 degrees of freedom
AIC: 4022

```

Number of Fisher Scoring iterations: 4

```

> exp(coef(LR_model1))
      (Intercept)      oplcat      leeftijd Geslacht_dummy
      0.1843260      1.3675742      0.9970544      0.9589396
> h1test(LR_model1)

```

The Hosmer-Lemeshow goodness-of-fit test

Group	Size	Observed	Expected
1	323	79	61.42051
2	325	65	70.66010
3	322	97	84.31368
4	320	105	104.80165
5	318	83	110.60469
6	320	94	119.83420
7	322	128	132.91342
8	322	137	136.82819
9	321	165	148.08161
10	314	177	160.54195

```

      Statistic = 36.19654
degrees of freedom = 8
      p-value = 1.6163e-05

```

```

> anova(LR_model0, LR_model1, test = "LRT")
Analysis of Deviance Table

```

```

Model 1: Gevoel_Politieke_vertegenwoordiging_dummy ~ 1
Model 2: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy

```

	Resid.	Df	Resid.	Dev	Df	Deviance	Pr(>Chi)
1	3206		4162				
2	3203		4014	3		147.97	< 2.2e-16 ***

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

>
> #Model 2
> LR_model2 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + lee
ftijd + Geslacht_dummy + cv24p335,
+               family = binomial (link = "logit"),
+               data = LISS4)
> summary(LR_model2)

```

```

Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat +
  leeftijd + Geslacht_dummy + cv24p335, family = binomial(link = "logit")
,
  data = LISS4)

```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.494984	0.212591	-7.032	2.03e-12 ***
oplcat	0.311594	0.027978	11.137	< 2e-16 ***
leeftijd	-0.005153	0.002413	-2.136	0.03270 *
Geslacht_dummy	-0.045829	0.076157	-0.602	0.54733
cv24p335	-0.259938	0.094958	-2.737	0.00619 **

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 4162.0  on 3206  degrees of freedom
Residual deviance: 4006.4  on 3202  degrees of freedom
AIC: 4016.4

Number of Fisher Scoring iterations: 4

> exp(coef(LR_model2))
      (Intercept)      oplcat      leeftijd Geslacht_dummy      cv24p335
      0.2242522      1.3656000      0.9948606      0.9552056      0.7710997
> h1test(LR_model2)

The Hosmer-Lemeshow goodness-of-fit test

Group Size Observed Expected
  1  321         71  59.06424
  2  322         75  70.98679
  3  320         92  83.09913
  4  322         94 100.76886
  5  323         96 113.42635
  6  321         88 120.52573
  7  317        132 129.61622
  8  320        142 137.79346
  9  323        157 149.89932
 10  318        183 164.81989

      Statistic = 28.46787
degrees of freedom = 8
      p-value = 0.00039299

> anova(LR_model1, LR_model2, test = "LRT")
Analysis of Deviance Table

Model 1: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy
Model 2: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy + cv24p335
Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1      3203      4014.0
2      3202      4006.4  1    7.5917 0.005864 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> #Model 3
> LR_model3 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + lee
ftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy,
+                  family = binomial(link = "logit"),
+                  data = LISS4)
> summary(LR_model3)

Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat +
      leeftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy,
      family = binomial(link = "logit"), data = LISS4)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.742798   0.248982  -7.000 2.57e-12 ***
oplcat         0.306271   0.028119  10.892 < 2e-16 ***
leeftijd      -0.006258   0.002477  -2.526  0.01152 *
Geslacht_dummy -0.034914   0.076380  -0.457  0.64759
cv24p335      -0.266220   0.095127  -2.799  0.00513 **
Traditionele_media_dummy 0.348337   0.176287   1.976  0.04816 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```


(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4162.0 on 3206 degrees of freedom
Residual deviance: 4002.4 on 3201 degrees of freedom
AIC: 4014.4

Number of Fisher Scoring iterations: 4

```
> exp(coef(LR_model3))
              (Intercept)              oplcat              leeftijd
              0.1750299              1.3583507              0.9937612
              Geslacht_dummy              cv24p335 Traditionele_media_dummy
              0.9656882              0.7662707              1.4167100
```

```
> h1test(LR_model3)
```

The Hosmer-Lemeshow goodness-of-fit test

Group	Size	Observed	Expected
1	321	66	57.78322
2	320	80	71.07776
3	323	91	84.10180
4	321	88	99.53480
5	320	101	111.93702
6	322	96	121.11549
7	318	123	129.81023
8	322	142	138.52998
9	321	160	150.01495
10	319	183	166.09475

Statistic = 21.15264
degrees of freedom = 8
p-value = 0.0067528

```
> anova(LR_model2, LR_model3, test = "LRT")
```

Analysis of Deviance Table

Model 1: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy + cv24p335

Model 2: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy + cv24p335 + Traditionele_media_dummy

	Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	3202	4006.4			
2	3201	4002.4	1	4.0431	0.04435 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
>
```

```
> #Model 4
```

```
> LR_model4 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + lee  
ftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy + Social_Media  
XTraditionele_Media,
```

```
+ family = binomial (link = "logit"),  
+ data = LISS4)
```

```
> summary(LR_model4)
```

Call:

```
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat +  
leeftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy +  
Social_MediaXTraditionele_Media, family = binomial(link = "logit"),  
data = LISS4)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.956921	0.278399	-7.029	2.08e-12	***
oplcat	0.307749	0.028138	10.937	< 2e-16	***
leeftijd	-0.006385	0.002481	-2.574	0.01006	*
Geslacht_dummy	-0.030215	0.076446	-0.395	0.69266	
cv24p335	0.366441	0.348473	1.052	0.29300	
Traditionele_media_dummy	0.571730	0.218190	2.620	0.00878	**
Social_MediaXTraditionele_Media	-0.678251	0.360341	-1.882	0.05980	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4162.0 on 3206 degrees of freedom
Residual deviance: 3998.9 on 3200 degrees of freedom
AIC: 4012.9

Number of Fisher Scoring iterations: 4

```
> exp(coef(LR_model4))
              (Intercept)              oplcat
              0.1412928              1.3603593
              leeftijd              Geslacht_dummy
              0.9936355              0.9702367
              cv24p335      Traditionele_media_dummy
              1.4425910              1.7713292
Social_MediaXTraditionele_Media
              0.5075037
```

```
> h1test(LR_model4)
```

The Hosmer-Lemeshow goodness-of-fit test

Group	Size	Observed	Expected
1	322	68	57.46466
2	320	74	71.41335
3	322	90	83.34248
4	321	90	98.66154
5	317	99	110.87382
6	320	102	119.94736
7	321	121	131.14365
8	323	140	139.56923
9	320	157	149.75044
10	321	189	167.83347

Statistic = 18.12121
degrees of freedom = 8
p-value = 0.020336

```
> anova(LR_model3, LR_model4, test = "LRT")
```

Analysis of Deviance Table

Model 1: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy + cv24p335 + Traditionele_media_dummy

Model 2: Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat + leeftijd +
Geslacht_dummy + cv24p335 + Traditionele_media_dummy + Social_MediaXTraditionele_Media

	Resid. Df	Resid. Dev	Df	Deviance	Pr(>Chi)
1	3201	4002.4			
2	3200	3998.9	1	3.4748	0.06231 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
>
> # Functie voor classificatietabellen
> Class.table <- function(LOGMOD = NULL){
+   DATSET <- LOGMOD$data
+   DATSET <- mutate(DATSET,
+                     p_hat = predict(LOGMOD, type = "response"),
+                     y_hat = as.factor(ifelse(p_hat >= 0.5, 1, 0)))
+   DV <- LOGMOD$formula[[2]]
+   Class_tmp <- table(DATSET[[DV]], DATSET$y_hat)
+   Class_tab <- matrix(data = NA, nrow = 3, ncol = 3)
+   # Controle of alle cases in 1 groep geclassificeerd worden
+   C1 <- ifelse(Class_tmp[1]+Class_tmp[2] == length(DATSET[[DV]]), T, F)
+   # Controle of dit de 0 groep is (anders 1)
+   C2 <- ifelse(dimnames(Class_tmp)[2] == "0", T, F)
+   if(C1){
+     Class_tab[1,1] <- ifelse(C2, Class_tmp[1], 0)
+     Class_tab[1,2] <- ifelse(C2, 0, Class_tmp[1])
+   }
```

```

+   Class_tab[2,1] <- ifelse(c2, class_tmp[2], 0)
+   Class_tab[2,2] <- ifelse(c2, 0, class_tmp[2])
+ } else {
+   Class_tab[1:2,1:2] <- class_tmp
+ }
+
+   Class_tab[1,3] <- round(Class_tab[1,1] / (Class_tab[1,1] + Class_tab[1,
2])), 4)
+   Class_tab[2,3] <- round(Class_tab[2,2] / (Class_tab[2,1] + Class_tab[2,
2])), 4)
+   Class_tab[3,3] <- round((Class_tab[1,1] + Class_tab[2,2]) / length(DATS
ET[[DV]]), 4)
+   rownames(Class_tab) <- c("Obs0", "Obs1", "Tot")
+   colnames(Class_tab) <- c("Exp0", "Exp1", "Tot")
+
+   return(Class_tab)
+ }
>
> #Classificatietabellen
> Classt0 <- Class.table(LOGMOD = LR_model0)
> Classt1 <- Class.table(LOGMOD = LR_model1)
> Classt2 <- Class.table(LOGMOD = LR_model2)
> Classt3 <- Class.table(LOGMOD = LR_model3)
> Classt4 <- Class.table(LOGMOD = LR_model4)
>
> print(Classt0)
      Exp0 Exp1  Tot
Obs0 2077   0 1.0000
Obs1 1130   0 0.0000
Tot   NA   NA 0.6476
> print(Classt1)
      Exp0 Exp1  Tot
Obs0 1953  124 0.9403
Obs1  974  156 0.1381
Tot   NA   NA 0.6576
> print(Classt2)
      Exp0 Exp1  Tot
Obs0 1980   97 0.9533
Obs1  991  139 0.1230
Tot   NA   NA 0.6607
> print(Classt3)
      Exp0 Exp1  Tot
Obs0 1989   88 0.9576
Obs1  986  144 0.1274
Tot   NA   NA 0.6651
> print(Classt4)
      Exp0 Exp1  Tot
Obs0 1985   92 0.9557
Obs1  983  147 0.1301
Tot   NA   NA 0.6648

```

Bijlage 3

Assumpties

Bij lineaire regressie moet er aan vier assumpties worden voldaan. De observaties moeten onafhankelijk van elkaar zijn, er moet een lineair verband zijn (lineariteit), er moet sprake zijn van een constante conditionele standaarddeviatie (homoscedasticiteit) en de conditionele verdeling van de afhankelijke variabele moet een normale verdeling zijn (normaliteit). Hieronder wordt met behulp van een aantal figuren toegelicht of de assumpties geschonden worden of niet. Omdat één van de assumpties duidelijk geschonden wordt is er besloten om een logistische regressie te doen. Bij logistische regressie moeten de observaties onafhankelijk van elkaar zijn.

Onafhankelijke observaties:

Om te controleren of de observaties onafhankelijk van elkaar zijn moet er gekeken worden naar de manier waarop de data verzameld is. Het LISS panel maakt gebruik van aselechte steekproeven door willekeurig huishoudens te selecteren uit het populatieregister van het centraal bureau voor de statistiek (*Methodology - LISS Panel*, 2024). Een aselechte steekproef zorgt ervoor dat de observaties onafhankelijk van elkaar zijn. Deze aselechte steekproeven werden afgewisseld met gestratificeerde steekproeven waardoor moeilijk te bereiken groepen ook in het panel terecht konden komen (*Methodology - LISS Panel*, 2024). Deze manier van respondenten verzamelen kan ervoor zorgen dat de observaties niet helemaal meer onafhankelijk van elkaar zijn, maar omdat dit is gedaan zodat de steekproef representatief is voor de populatie en het steeds wordt afgewisseld met aselechte steekproeven hoeft dit geen probleem te zijn. Daarnaast zitten in de steekproef meerdere personen per huishouden wat kan betekenen dat de observaties van de personen binnen zo'n huishouden niet onafhankelijk van elkaar zijn is er gekozen om willekeurig één iemand per huishouden te selecteren. Er wordt dus aan deze assumptie voldaan vanwege de manier waarop de data verzameld is en door willekeurig één iemand per huishouden te selecteren.

R script onafhankelijke observaties:

Hieronder staat het filter wat is toegepast om willekeurig iemand te selecteren per huishouden. Er is ook een seed (5066638) geselecteerd zodat de resultaten reproduceerbaar zijn.

```
> #seed instellen voor reproduceerbaarheid
> set.seed(5066638)
> #LISS3 hierbij wordt 1 iemand gekozen per huishouden.
> LISS3 <- LISS2 %>%
+   group_by(nohouse_encr) %>%
+   slice_sample(n = 1) %>%
+   ungroup()
>
```

```
> #LISS4 hierbij worden ook nog alle missende waarden eruit gefilterd. Hier
door worden alleen respondenten meegenomen die op alle variabelen een antwo
ord heeft gegeven
> LISS4 <- LISS2 %>%
+   filter(complete.cases(.)) %>%
+   group_by(nohouse_encr) %>%
+   slice_sample(n = 1) %>%
+   ungroup()
```

Assumpties lineaire regressie

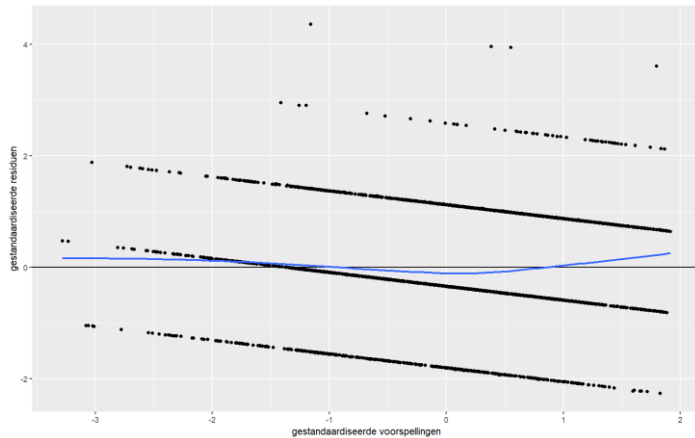
Om de assumpties van lineaire regressie te controleren is er eerst een lineair model geschat. Hieronder staat het r-script wat gebruikt is om deze assumpties te controleren.

R script assumpties lineaire regressie:

```
> #Assumptiecontrole
> #Lineaire regressie
> model_4_lineair <- (lm(Gevoel_Politieke_vertegenwoordiging ~ oplcat + lee
ftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy + Social_Media
XTraditionele_Media, data = LISS4))
>
>
> LISS4 <- mutate(LISS4,
+                 pred = predict(model_4_lineair),
+                 zpred = (pred - mean(pred))/sd(pred))
>
> LISS4 <- mutate(LISS4,
+                 sresid = studres(model_4_lineair))
> #Lineariteit en homoscedasticiteit
> LISS4 %>%
+   ggplot(mapping = aes(x= zpred, y= sresid)) +
+   geom_point()+
+   geom_hline(yintercept = 0)+
+   geom_smooth(method = "loess", se = F)+
+   xlab("gestandaardiseerde voorspellingen")+
+   ylab("gestandaardiseerde residuen")
+   `geom_smooth() using formula = 'y ~ x'`
>
> #Normale verdeling residuen
> ggplot(data = LISS4, mapping = aes(sample = sresid)) +
+   stat_pp_line(linewidth = 1.25, color = "black") +
+   stat_pp_point(size = 2, color = "blue") +
+   xlab("PP-plot")
>
> ggplot(data = LISS4, mapping = aes(x = sresid))+
+   geom_histogram(fill = "blue", colour = "white")+
+   xlab("gestandaardiseerde residuen")+
+   ylab("aantal")
+   `stat_bin() using `bins = 30`. Pick better value with `binwidth`.
```

Lineariteit:

Om aan de assumptie van lineariteit te voldoen moet het gemiddelde van de residuen gelijk zijn aan 0 voor elke set van waarden van de x-en. In het figuur hieronder is te zien dat deze assumptie een klein beetje wordt geschonden, want de lijn in het figuur is niet helemaal recht. Toch is deze schending niet zo groot dat lineaire regressie niet meer kan.

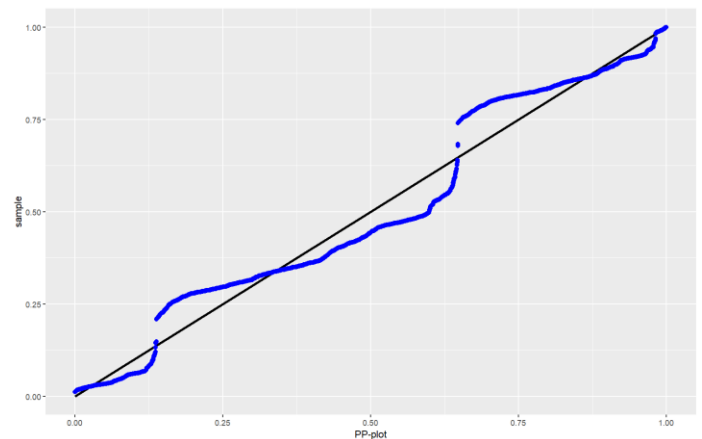
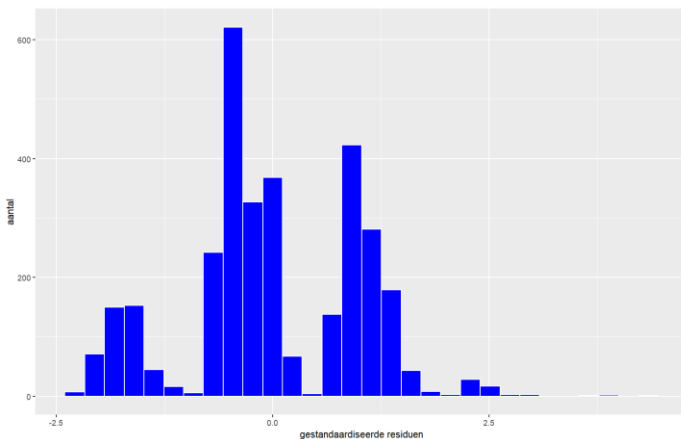


Homoscedasticiteit:

Om aan de assumptie van homoscedasticiteit te voldoen moet de mate waarin iemand zich vertegenwoordigd voelt door de politiek dezelfde conditionele variantie hebben voor elke set van waarden van de x -en (in de populatie). In het figuur hierboven is te zien dat deze assumptie ook niet echt wordt geschonden, er zijn wel een aantal uitbijters maar de spreiding van de punten veranderd niet systematisch.

Normaliteit:

Om aan deze assumptie te voldoen moet de conditionele verdeling van de mate waarin iemand zich vertegenwoordigd voelt door de politiek normaal zijn gegeven de x -en. In de figuren hieronder is te zien dat deze conditionele verdeling niet normaal is verdeeld. In het histogram zijn drie pieken te zien en in het pp-plot is te zien dat de punten niet netjes langs de lijn lopen. Deze assumptie wordt dus geschonden en hierdoor is er besloten om een logistische regressie te doen.



Multicollineariteit

Hieronder staat het R script met verschillende VIF waarden. VIF waarden groter dan 4 zijn mogelijk een probleem en hieronder bij model 4 zijn twee VIF waarden die groter zijn dan 4. De VIF waarden van het volgen van nieuws op sociale media (cv24p335) en het interactie-effect (Social_MediaXTraditionele_Media) zijn beide groter dan 4 en dit laat zien dat deze variabelen erg met elkaar correleren. In de frequentieverdelingen en het kruistabel eronder is te zien dat het interactie-effect qua scores ook nauwelijks afwijkt van het volgen van nieuws op sociale media. Bijna iedereen die het nieuws volgt op sociale media volgt het ook op traditionele media, want anders kunnen ze geen 1 scoren op het interactie-effect.

R script Multicollineariteit:

```
> #Multicollineariteit
> vif(LR_model4)
```

```
              oplcat              leeftijd
              1.044148              1.233502
Geslacht_dummy
              1.016467              cv24p335
Traditionele_media_dummy Social_MediaXTraditionele_Media
              1.621933              15.148079
              15.271518
```

```
> vif(LR_model3)
```

```
              oplcat              leeftijd              Geslacht_dummy
              1.043269              1.228026              1.015824
cv24p335 Traditionele_media_dummy
              1.127243              1.061305
```

```
> table(LISS4$cv24p335, useNA = "always")
```

```
  0    1 <NA>
2435 772    0
```

```
> table(LISS4$Social_MediaXTraditionele_Media, useNA = "always")
```

```
  0    1 <NA>
2497 710    0
```

```
> CrossTable(LISS4$cv24p335, LISS4$Social_MediaXTraditionele_Media)
```

Cell Contents			
			N
Chi-square	contribution		
	N / Row Total		
	N / Col Total		
	N / Table Total		

Total observations in Table: 3207

LISS4\$cv24p335	LISS4\$Social_MediaXTraditionele_Media		Row Total
	0	1	
0	2435	0	2435
	153.284	539.086	
	1.000	0.000	0.759
	0.975	0.000	
	0.759	0.000	
1	62	710	772
	483.481	1700.357	
	0.080	0.920	0.241
	0.025	1.000	
	0.019	0.221	
Column Total	2497	710	3207
	0.779	0.221	

Uitbijters en invloedrijke punten

Om de uitbijters en invloedrijke punten te vinden zijn er *leverage* waarden en DFBetas waarden uitgerekend die kijken of er invloedrijke punten zijn die invloed hebben op

de hellingen. De grenswaarde dat wordt aangehouden voor dit onderzoek bij *leverage* is $h > \frac{3 \times 7}{3207} = 0,0065$ en bij *DFBetas* is $DFBetas > \frac{3}{\sqrt{3207}} = 0,0530$. Hieronder staat het script voor het genereren van de leverage waarden en het filteren van de dataset. Wat direct opvalt bij de *leverage* is dat de respondenten die een hogere leverage waarde hebben dan de grenswaarde van 0,0065 allemaal een 0 score het volgen van nieuws op traditionele media. Na het wegfilteren van de respondenten die boven deze grenswaarde zijn er nog 3047 respondenten over. Er worden dus 160 respondenten weggehaald die allemaal het nieuws niet volgen op traditionele media. In het nieuwe model is het interactie-effect weg, want het heeft nu precies dezelfde verdeling als het volgen van nieuws op sociale media en de correlatie tussen deze twee variabelen is dan ook 1. Het interactie-effect is er dus eigenlijk uitgefilterd. De hosmer-lemeshow toets is nog steeds niet significant en het effect van traditionele media is niet meer significant maar wel sterker positief. Dit komt omdat er 160 respondenten uitgefilterd zijn die het nieuws niet volgen op traditionele media waardoor er nog maar 29 respondenten zijn die het nieuws niet volgen op traditionele media, hierdoor is er te weinig statistische power om een significant resultaat te vinden.

Als we kijken naar de *DFBetas* dan vallen er ook een aantal dingen op. De extreme *DFBetas* waarden bij opleidingsniveau zijn allemaal respondenten met een score van 1 op opleidingsniveau en een score van 1 op je wel of niet vertegenwoordigd voelen door de politiek, of te wel laag opgeleide mensen die zich wel vertegenwoordigd voelen door de politiek. De extreme *DFBetas* waarden bij leeftijd zijn voornamelijk wat jongere respondenten met een score van 1 op je wel of niet vertegenwoordigd voelen door de politiek, of te wel jonge mensen die zich wel vertegenwoordigd voelen door de politiek. Er zijn geen extreme *DFBetas* waarden bij geslacht en de extreme *DFBetas* bij Traditionele media en het interactie-effect zijn allemaal respondenten die het nieuws niet volgen op traditionele media. Deze respondenten met extremere *DFBetas* waarden overlappen heel erg met de respondenten met een extreme leverage waarde. De respondenten met hoge *leverage*- en extreme *DFBetas*-waarden respondenten zijn respondenten met geldige waarden en zorgen juist voor nuance in de gevonden resultaten, waardoor er besloten is om deze respondenten in de dataset te laten.

R script uitbijters en invloedrijke punten:

```
> #Outliers en invloedrijke punten
> LISS4 <- mutate(LISS4,
+               lever = hatvalues(LR_model4))
> LISS4 <- mutate(LISS4,
+               dfbetas = dfbetas(LR_model4))
> FilterLISS <- LISS4 %>%
+   filter(lever < (21/3207))
> FilterLR_model4 <- glm(Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat
+ leeftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy + Social_MediaXTraditionele_Media,
+                       family = binomial(link = "logit"),
+                       data = FilterLISS)
> summary(FilterLR_model4)
```

```
Call:
glm(formula = Gevoel_Politieke_vertegenwoordiging_dummy ~ oplcat +
  leeftijd + Geslacht_dummy + cv24p335 + Traditionele_media_dummy +
  Social_MediaXTraditionele_Media, family = binomial(link = "logit"),
  data = FilterLISS)
```

```
Coefficients: (1 not defined because of singularities)
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.346774    0.638358  -3.676 0.000237 ***
oplcat       0.316512    0.029125  10.867 < 2e-16 ***
leeftijd     -0.005356    0.002553  -2.098 0.035918 *
```



```

Geslacht_dummy          -0.023211    0.078129   -0.297  0.766398
cv24p335                -0.299963    0.098793   -3.036  0.002395 **
Traditionele_media_dummy  0.859007    0.616297    1.394  0.163372
Social_MediaXTraditionele_Media      NA          NA          NA          NA

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(Dispersion parameter for binomial family taken to be 1)

```

Null deviance: 3963.2  on 3046  degrees of freedom
Residual deviance: 3801.6  on 3041  degrees of freedom
AIC: 3813.6

```

Number of Fisher Scoring iterations: 4

```
> hltest(FilterLR_model4)
```

The Hosmer-Lemeshow goodness-of-fit test

Group	Size	Observed	Expected
1	304	62	53.08918
2	306	74	67.89037
3	305	84	78.93700
4	305	92	95.16996
5	310	94	110.15587
6	305	88	115.59566
7	305	127	126.63781
8	306	134	133.97529
9	305	152	143.53006
10	296	174	156.01881

```

Statistic = 22.72226
degrees of freedom = 8
p-value = 0.0037397

```

```
> table(LISS$Traditionele_media_dummy)
```

```

  0    1
189 3018

```

```
> table(FilterLISS$Traditionele_media_dummy)
```

```

  0    1
29 3018

```

```
> table(FilterLISS$cv24p335, useNA = "always")
```

```

  0    1 <NA>
2337  710    0

```

```
> table(FilterLISS$Social_MediaXTraditionele_Media, useNA = "always")
```

```

  0    1 <NA>
2337  710    0

```