

Prevalentiemonitor Huiselijk Geweld en Seksueel Geweld 2020

Onderzoeksdocumentatie

Auteurs:

Björn Janssen

Geert Gommans

Elke Moons

Carin Reep

Marleen Wingen

Inhoudsopgave

1.	Inleiding.....	4
	1.1 Doel van dit document	4
2.	Doel en achtergrond van het onderzoek	5
3.	Dataverzameling (waarnemen)	6
	3.1 Steekproefkader en doelpopulatie	6
	3.2 Benaderingsstrategie	7
	3.3 Vragenlijstontwikkeling	10
	3.4 Resultaten veldwerk	13
4.	Verwerken (dataprocessing)	16
	4.1 Inleiding	16
	4.2 Invoerdata	17
	4.3 Verrijken	17
	4.5 Afleiden	18
5.	Responsanalyse	20
	5.1 Selectiviteit van de respons	20
6.	Andere kwaliteitsindicatoren	24
	6.1 Geen antwoord op vragen	24
	6.2 Veldwerk tijdens corona-uitbraak	24
7.	Plausibiliteitsanalyse	26
	7.1 Verdeling kernindicatoren	26
	7.1 Vergelijking met ander onderzoek	27
8.	Weging	31
	8.1 Inleiding	31
	8.2 Weegmodel	31
	8.4 Resultaten van de weging	33
9.	Opgeleverde Producten.....	36
	10.1 Microbestanden	36
	10.2 Statistische beveiliging	36
10.	Referenties.....	37
11.	Bijlagen.....	37

Bijlage I: Aanschrijfbrieven en folder	37
Bijlage II: Weging	43

1. Inleiding

1.1 Doel van dit document

In dit document wordt verslag gedaan van de Prevalentiemonitor Huiselijk Geweld en Seksueel Geweld 2020. De voorbereiding, dataverzameling en verwerking van de gegevens zijn gedocumenteerd. Dit document dient als naslagwerk voor onderzoekers die meer inzicht wensen in de achtergronden van het onderzoek en in het tot stand komen van de analysebestanden. Daarnaast is in dit document een responsanalyse en plausibiliteitsanalyse opgenomen, plus een hoofdstuk over het weegmodel.

2. Doel en achtergrond van het onderzoek

De Prevalentiemonitor Huiselijk Geweld en Seksueel Geweld betreft een nieuw opgezet tweejaarlijkse monitor van de jaarprevalentie van slachtofferschap van huiselijk geweld, seksuele intimidatie en seksueel geweld onder de Nederlandse bevolking van 16 jaar of ouder. De doelstelling van het project is:

- Inzicht krijgen in de aard en omvang, gemeten over de periode van een jaar, van slachtofferschap van huiselijk geweld, seksuele intimidatie en seksueel geweld onder de Nederlandse bevolking van 16 jaar en ouder; en
- Inzicht krijgen in de trend van de jaarprevalentie van huiselijk geweld, seksuele intimidatie en -seksueel geweld.

Daarnaast dient de PHGSG antwoord te geven op de volgende vragen:

- Wat zijn de door slachtoffers ervaren gevolgen van huiselijk geweld, seksuele intimidatie en seksueel geweld?
- In hoeverre spreken slachtoffers van huiselijk geweld, seksuele intimidatie en seksueel geweld met anderen over hun ervaringen?
- In hoeverre worden huiselijk geweld, seksuele intimidatie en seksueel geweld door slachtoffers zelf of door anderen gemeld bij de politie?
- Bestaan er verschillen in de aard en omvang van slachtofferschap van huiselijk geweld, seksuele intimidatie en seksueel geweld tussen verschillende subgroepen (geslacht, leeftijdsgroepen, migratieachtergrond, opleidingsniveau, en andere relevante subgroepen)?

Om deze vraagstellingen te kunnen beantwoorden heeft het WODC op verzoek van JenV en VWS het CBS gevraagd om tweejaarlijks een vragenlijst af te nemen onder een representatieve steekproef uit de Nederlandse bevolking van 16 jaar en ouder.

De afgesproken looptijd van het onderzoek is 3 onderzoeksjaren (2020, 2022, 2024).

Deze onderzoeksdocumentatie beschrijft het eerste onderzoeksjaar. De voortgang van dit onderzoek wordt gemonitord door een begeleidingscommissie met experts door heel Nederland.

3. Dataverzameling (waarnemen)

3.1 Steekproefkader en doelpopulatie

De doelpopulatie bestaat uit alle in Nederland woonachtige personen van 16 jaar of ouder bij start van het veldwerk, die geregistreerd zijn als ingezetene in de Basisregistratie Personen (BRP) en die deel uitmaken van particuliere huishoudens. De institutionele bevolking, dat zijn personen in inrichtingen, instellingen of tehuizen (IIT), worden voor dit onderzoek niet benaderd. Er wordt een personensteekproef getrokken uit het steekproefkader dat het CBS gebruikt voor steekproeven met de persoon als trekkingseenheid.

1. Verhuisingsstrategie

Hoewel steekproeven zo kort mogelijk voor aanvang van een nieuwe waarneemperiode worden getrokken, is niet te voorkomen dat het adres in de Basisregistratie Personen (BRP) niet juist blijkt te zijn, omdat de steekproefpersoon is verhuisd. Tijdens de waarneemperiode worden geen verhuizers herbenaderd.

2. Steekproefontwerp

In het steekproefontwerp is rekening gehouden met een aselechte steekproef. In het algemeen betekent dit dat er een zelfwegende tweetrapssteekproef wordt getrokken, waarbij in de eerste trap gemeenten worden geselecteerd met kansen evenredig aan de omvang, en in de tweede trap een vast, vooraf bepaald aantal personen. Dit aantal wordt de clusteromvang genoemd. Omdat de waarneming via internet (cawi) gaat, kan de clusteromvang gelijk aan 1 worden gekozen. Vanwege de grote steekproefomvang van 100.000 personen vereenvoudigt dit ontwerp dan tot een steekproef waarbij wordt gestratificeerd naar gemeente. Binnen iedere gemeente wordt een enkelvoudig aselechte steekproef van personen getrokken. Het aantal te trekken personen per gemeente wordt zodanig bepaald dat de steekproeffractie in alle gemeenten hetzelfde is.

3. Steekproefbewerking

Aansluitend op de steekproeftrekking worden de steekproeven bewerkt. De steekproefbewerking begint met het opvragen van de meest recente NAW-gegevens van steekproefpersonen bij de BRP. Vervolgens worden uit de getrokken steekproef bepaalde personen verwijderd. Het proces waarin dit gebeurt, wordt screening genoemd. Personen die tijdens de screening uit de steekproef verwijderd worden, zijn personen van wie het adres gedurende de afgelopen twaalf maanden al eerder in een CBS-steekproef voorkwam¹, personen van wie het adres behoort tot de adressen van de institutionele bevolking en personen van wie geen bruikbaar adres wordt verkregen bij de BRP. Indien meerdere personen woonachtig op hetzelfde adres getrokken worden, blijft na screening maximaal één persoon woonachtig op dit adres over. De overige personen woonachtig op het adres vallen tijdens de screening uit, omdat hun adres dan gedurende de afgelopen twaalf maanden al eerder in een CBS-steekproef voorkwam. Het niet verkrijgen van een bruikbaar adres kan veroorzaakt worden doordat personen bij de BRP hebben aangegeven hun gegevens niet te willen delen met derden. Een andere reden voor het niet verkrijgen van een bruikbaar adres kan zijn dat de

¹ Onderzoeken met een heel specifieke doelpopulatie worden hierbij niet meegenomen omdat dit anders zou zorgen voor verwijdering van een selecte groep respondenten. Op deze manier wordt er voor gezorgd dat de verwijdering zoveel als mogelijk aselekt is.

betreffende persoon inmiddels is overleden. Om na screening nog genoeg personen in de steekproef over te houden, zijn de getrokken steekproeven groter dan de uit te zetten steekproeven.

De na screening overgebleven personen worden eerst gesorteerd op postcode, huisnummer, huisletter, toevoeging en aanduiding. Vervolgens wordt systematisch uitgedund hetgeen betekent dat er met een vooraf bepaalde stapgrootte door de steekproef wordt “gestapt”. De sortering zorgt er samen met het systematisch uitdunnen voor dat de steekproef landelijk gespreid blijft en niet bijvoorbeeld alleen steekproefpersonen in Amsterdam worden verwijderd. De na uitdunning overgebleven personen worden benaderd en verdeeld over twee steekproefporties.

Op de adressen van de personen die worden benaderd worden de komende twaalf maanden geen personen benaderd voor andere CBS-onderzoeken.

In tabel 1 wordt een samenvatting gegeven van het steekproefproces voor de totale steekproef.

Tabel 1. Verantwoording steekproefbewerking totale steekproef

	Aantal
Getrokken	111.500
Niet beschikbaar voor uitzet, waarvan:	8.519
- in eerdere CBS steekproef	6.467
- institutionele bevolking	17
- met onbekende straatnaam/NAW-gegevens	2.035
Beschikbaar voor uitzet	102.981
Uitdunning	2.981
Uitgezet	100.000

3.2 Benaderingsstrategie

Bij het ontwerp van het onderzoeksdesign voor jaargang 2020 is uitgegaan van de navolgende randvoorwaarden (uit het projectvoorstel, opgesteld in november 2019):

1. Het veldwerk wordt uitgevoerd in maart en april 2020.
2. Het onderzoek wordt uitgevoerd via internetwaarneming (Cawi). Steekproefpersonen ontvangen een aanschrijfbrief met het verzoek om de vragenlijst via internet in te vullen. Non-respondenten ontvangen maximaal twee rappelbrieven. Er zijn geen vervolgmodes (waarneming met enquêteurs).
3. De aanschrijf- en rappelbrieven worden verstuurd met 72-uurs post. In de aanschrijfbrief is de naam van de respondent vermeld, maar in de brief staat duidelijk dat de resultaten anoniem verwerkt worden.
4. In de aanschrijf- en rappelbrieven wordt vermeld dat het onderzoek wordt uitgevoerd in samenwerking met het WODC, waarbij het CBS de verzamelde gegevens aanvult met enkele reeds bij het CBS aanwezige achtergrondkenmerken. Er wordt een formulering voor de “informed consent” in de brieven opgenomen, waarbij in algemene zin wordt opgenomen welke variabelen het betreft. Het WODC wordt betrokken bij de inhoudelijke paragraaf in de brieven.
5. Er wordt een maatwerkfolder toegevoegd aan de aanschrijfbrieven (zie bijlage I).
6. Er worden geen incentives ingezet.
7. Er wordt een personensteekproef getrokken.

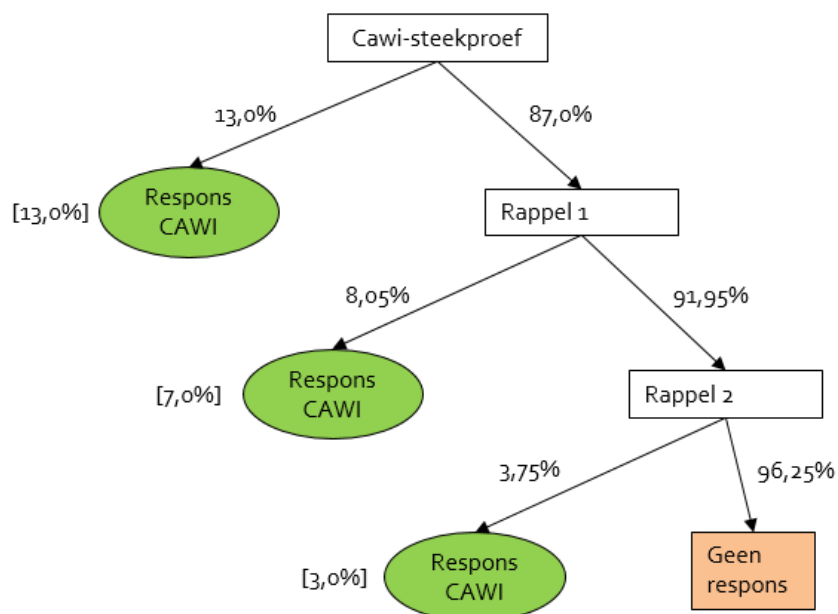
8. De steekproef wordt in twee gelijke porties uitgezet.
9. Er wordt naar verwachting een respons van 23% gerealiseerd. Er is sprake van een inspanningsverplichting.
10. Bij een afgesproken steekproefomvang van 100.000 personen worden daarom naar verwachting 23.000 responsen gerealiseerd.
11. De duur van de vragenlijst bedraagt gemiddeld 20 minuten.
12. Op de CBS-website komt informatie over dit onderzoek voor respondenten. Tevens zijn medewerkers van de CBS-helpdesk geïnstrueerd, zodat zij vragen van respondenten kunnen beantwoorden en hen op weg kunnen helpen waar nodig.
13. Personen die aangeschreven worden kunnen op werkdagen tenminste van 9.00 tot 17.00 uur met vragen terecht bij het Contact Center Inbound (CCI) van het CBS.

Bovenstaande randvoorwaarden hebben geleid tot de volgende benaderstrategie:

- Steekproefpersonen ontvangen een aanschrijfbrief met daarin het verzoek om via internet deel te nemen aan het onderzoek en de bijbehorende inloggegevens. Daarnaast wordt in de brief ingegaan op het doel van het onderzoek. Bij de aanschrijfbrief wordt een maatwerkfolder bijgevoegd.
- Drie weken na de aanschrijfbrief wordt aan steekproefpersonen een eerste rappelbrief verstuurd met daarin opnieuw het verzoek om via internet deel te nemen aan het onderzoek. Deze brief wordt alleen verstuurd aan steekproefpersonen waarvan geen respons is ontvangen via internet en die niet via het Contact Centre Inbound (CCI) hebben aangegeven niet te willen of kunnen deelnemen aan het onderzoek.
- Zes weken na de aanschrijfbrief wordt aan steekproefpersonen een tweede rappelbrief verstuurd met opnieuw het verzoek om via internet deel te nemen aan het onderzoek. Ook deze brief wordt alleen verstuurd aan steekproefpersonen waarvan nog geen respons is ontvangen via internet en die niet via CCI hebben aangegeven niet te willen of kunnen deelnemen aan het onderzoek. In de brief staat een uiterste datum tot wanneer de vragenlijst kan worden ingevuld.

Op basis van deze benaderingsstrategie werd een respons van 23% en daarmee 23.000 responsen verwacht. In figuur 1 wordt de benaderstrategie per benaderstap weergegeven. Per benaderstap is de geschatte responskans vermeld, waarbij zowel de responskans is vermeld uitgaande van de betreffende deelsteekproef die daadwerkelijk de aanschrijfbrieven, eerste rappelbrieven of twee rappelbrieven heeft gekregen (de kans weergegeven zonder vierkante haken) als van de totaal uitgezette cawi-steekproef (de kans weergegeven tussen vierkante haken). Zo reageert naar verwachting 8,05% van de steekproefpersonen die een eerste rappelbrief heeft gehad. Op de totaal uitgezette cawi-steekproef (de personen die de aanschrijfbrief hebben gehad) is dit 7,0%. De responskans op de totaal uitgezette cawi-steekproef wordt uitgerekend door de kansen tussen de vierkante haken op te tellen.

Figuur 1. Benaderstrategie



De steekproef werd in twee porties uitgezet. De uitzending van de brieven gebeurde conform tabel 2. Deze tabel bevat per portie de datums waarop de aanschrijf- en rappelbrieven werden verstuurd en aankwamen, de momenten waarop non-respondenten werden geselecteerd die een eerste of tweede rappelbrief moesten krijgen, de momenten waarop het drukproces startte en de geschatte steekproefaantallen. De tijd tussen de selectie van de elementen die moeten worden gerappelleerd en het daadwerkelijk versturen van het rappel werd benut om de brieven te drukken, adresseren en verzendklaar te maken. De aanschrijf- en rappelbrieven werden, voor zover mogelijk, op dinsdagen verstuurd met 72-uurs post, zodat steekproefpersonen deze voor of in het weekend zouden ontvangen. Alle datums liggen in 2020.

Tabel 2. Gegevens over de cawi-porties

Brief	Verzending		Aankomst		Post	Aantal
	Dag	Datum	Dag	Datum		
p1 aanschrijf	dinsdag	03-mrt	vrijdag	06-mrt	72-uurs	50.000
p1 rappel 1	dinsdag	24-mrt	vrijdag	27-mrt	72-uurs	43.573
p1 rappel 2	dinsdag	14-apr	vrijdag	17-apr	72-uurs	40.067
p2 aanschrijf	dinsdag	10-mrt	vrijdag	13-mrt	72-uurs	50.000
p2 rappel 1	dinsdag	31-mrt	vrijdag	03-apr	72-uurs	43.573
p2 rappel 2	dinsdag	21-apr	vrijdag	24-apr	72-uurs	40.067

In totaal zouden 100.000 aanschrijfbrieven worden verstuurd, en werd verwacht dat 87.146 eerste rappelbrieven en 80.134 tweede rappelbrieven zouden moeten worden verstuurd. In werkelijkheid zijn 87.926 eerste rappelbrieven gestuurd en 76.736 tweede rappelbrieven.

In tabel 3 wordt per portie de afsluitdatum opgenomen die tevens in de tweede rappelbrief is opgenomen.

Tabel 3. Afsluiting internetwaarneming per cawi-portie

Cawi-portie	Afsluitdatum internetwaarneming = datum in tweede rappel
p1	26-apr 23:59
p2	03-mei 23:59

3.3 Vragenlijstontwikkeling

Er werd een elektronische vragenlijst gebruikt voor de internetwaarneming. Het vragenlijstontwerp is door middel van een review en vragenlijstschema's² door de ontwikkelaar afgestemd met Rutgers en het WODC. De vragenlijst is in Blaise geprogrammeerd, de door het CBS ontwikkelde software voor elektronische vragenlijsten.

De vragenlijst die ziet op huiselijk geweld en stalking is ontwikkeld door het WODC. Dat is gebeurd op basis van bestaande vragenlijsten huiselijk geweld en het concept van de vragenlijst is voorgelegd aan een groep experts op het gebied van huiselijk geweld. De vragenlijst die ziet op seksuele intimidatie en geweld is in opdracht van het WODC ontwikkeld door Rutgers.

De inhoud van de vragenlijst wordt voorafgaande aan iedere jaargang in onderling overleg tussen het CBS en WODC vastgesteld op basis van de vastgestelde informatiebehoefte. De bestaande vragenlijst is hierbij het uitgangspunt, waarbij bekeken wordt welke wijzigingen nodig of gewenst zijn. Bij de herhaalmetingen kan maximaal 20% van de vragenlijst gewijzigd worden. Eventuele wijzigingen in de vragenlijst geeft WODC uiterlijk door in september in het jaar voorafgaand aan het veldwerk.

Geplande versus gerealiseerd Vragenlijstduur

De gemiddelde duur voor het volledig invullen van de internetvragenlijst werd geschat op gemiddeld 20 minuten. Uiteindelijk is dit 9,5 minuten geworden. Alleen volledig ingevulde vragenlijsten zijn meegenomen in de tijdsduurberekening (dus exclusief afbrekers). In tabel 4 is te zien dat de vragenlijst het langst duurt voor respondenten die aangeven dat ze slachtoffer zijn geweest en het minst lang voor respondenten die aangeven geen slachtoffer te zijn geweest.

² Schematische weergave van een vragenlijst waarbij de volgorde van vragen, routingvoorwaarden, tekstimputaties en afleidingen worden weergegeven, onderverdeeld in vragenblokken.

Tabel 4. Gemiddelde invulduur PHGSG 2020

	Invulduur (min)
Totaal	9,5
Afgelopen 5 jaar verbale agressie in huiselijke kring meegemaakt	
Nee	8,8
Ja	10,3
Weigert	10,4
Afgelopen 5 jaar fysiek geweld in huiselijke kring meegemaakt	
Nee	9,3
Ja	12,9
Weigert	10,7
Afgelopen 5 jaar dwingende controle in huiselijke kring meegemaakt	
Nee	9,1
Ja	13,0
Weigert	11,3
Afgelopen 5 jaar online seksuele intimidatie meegemaakt	
Nee	9,2
Ja	12,5
Weigert	10,2
Afgelopen 5 jaar niet-fysieke seksuele intimidatie meegemaakt (offline)	
Nee	9,1
Ja	12,5
Weigert	10,9
Afgelopen 5 jaar fysiek seksueel geweld meegemaakt	
Nee	9,2
Ja	12,8
Weigert	11,0

Inhoud vragenlijst

De vragenlijst gaat over persoonlijke veiligheid en meer in het bijzonder over niet-seksuele en seksuele grensoverschrijding en geweld. De vragen hebben betrekking op ervaringen van verbale en lichamelijke agressie en ongewenste seksuele ervaringen zowel binnen als buiten de huiselijke kring.

De vragenlijst bestaat uit de volgende inhoudelijke blokken (en vragen c.q. gebeurtenissen):

<i>Blok</i>	<i>Voorbeelden van vragen c.q. gebeurtenissen</i>
Meningsverschillen en verbale agressie	- iemand beledigde u en/of vloekte tegen u - iemand heeft met opzet spullen van u kapot gemaakt - iemand kleinerde of vernederde u
Lichamelijke agressie	- iemand dreigde een mes of ander wapen tegen u te gebruiken - iemand heeft u geschopt - iemand heeft geprobeerd om u te verstikken of te wurgen
Coercive control	- iemand hield u de hele tijd weg bij uw familie en/of vrienden - iemand bepaalde voor u waaraan u uw geld mocht uitgeven - iemand controleerde regelmatig of de hele tijd uw post, telefoontjes, berichten, e-mails, of sociale media
Stalking	- Een ex-partner belde u regelmatig, terwijl u dit niet wilde (inclusief opgehangen, voicemail en berichten) - Een ex-partner bespioneerde u bijvoorbeeld met een verborgen camera, opname apparatuur of andere technologie - Een ex-partner is uw auto of huis binnen gegaan zonder uw toestemming, of heeft dit geprobeerd
Hands-off online	- iemand maakte online seksueel kwetsende opmerkingen of grapjes - iemand bleef online aandringen op seks met hem/haar, terwijl u dat niet wilde - iemand maakte online een naaktfoto of seksfilmpje van u, terwijl u dat niet wilde
Hands-off offline	- iemand bleef aandringen op een date, terwijl u dat niet wilde - iemand staarde naar u op een seksuele manier, terwijl u dat niet wilde - iemand maakte een naaktfoto of seksfilmpje van u, terwijl u dat niet wilde
Hands-on	- iemand zoende u, terwijl u dat niet wilde - Orale seks (seks met de mond), terwijl u dat niet wilde - iemand dwong u om seks met iemand anders te hebben voor geld en/of goederen

In de vragenlijst wordt aan respondenten gevraagd of ze in de afgelopen 5 jaar slachtoffer zijn geweest van 1 of meer van bovenstaande gebeurtenissen (zie voorbeelden in de rechterkolom van de tabel). En zo ja, hoe vaak dit in de afgelopen 12 maanden is voorgekomen en wie in dat

geval/ die gevallen de pleger was. Deze vragen worden per gebeurtenis gesteld. Eventuele gevolgen van het slachtofferschap en het praten met anderen (inclusief de politie en Veilig Thuis) over het slachtofferschap zijn ook onderwerpen die aan bod komen in de vragenlijst. Ook wordt er gevraagd naar ervaringen voor het 12^{de} levensjaar. Deze vragen worden aan het einde van elk blok over alle gebeurtenissen in dat blok tezamen gesteld.

In het blok 'Achtergrondgegevens' wordt vastgesteld of de respondent ooit een relatie heeft gehad. Indien de respondent aangeeft nooit een relatie te hebben gehad, wordt het blok over stalking overgeslagen. De vragenlijst eindigt met de vraag of de respondent zelf weleens huiselijk of seksueel geweld heeft gepleegd.

In verband met de gevoeligheid van de vragenlijst, is er besloten om geen incentive te verloten. Op de ontvangstbevestigingspagina (dit is nadat de respondent de vragenlijst heeft verzonden) zijn de contactgegevens van een aantal hulpinstanties (Veilig Thuis, Centrum Seksueel Geweld, Slachtofferhulp Nederland) opgenomen.

Testprocedure

Respondenten krijgen niet alle vragen uit het onderzoek voorgelegd, omdat de vragenlijst zo gebouwd wordt dat respondenten alleen de vragen krijgen die voor hen relevant zijn. Soms wordt de respondent ook om vragen heen geleid, omdat uit eerdere antwoorden zijn antwoord op bepaalde vragen al kan worden afgeleid.

De in de computergestuurde vragenlijsten opgenomen routing kan complex zijn. Het is van belang de routing goed te controleren om te voorkomen dat tijdens de analysefase blijkt dat essentiële vragen ten onrechte niet zijn gesteld. Er wordt binnen het CBS een standaard testmethode toegepast, waarmee kan worden nagegaan of de in het ontwerp van de vragenlijst vastgelegde routing op een juiste wijze is geprogrammeerd. De gehanteerde testmethode bestaat eruit dat op basis van de te testen vragenlijst een testopzet opgesteld wordt. In de testopzet wordt aangegeven welke vragenblokken op welke manier getest gaan worden. Vervolgens worden voor de te testen vragenblokken testprotocollen ontwikkeld. Deze beschrijven stap voor stap welke gegevens in de vragenlijst ingevoerd dienen te worden. De vragenlijstontwikkelaars testen aan de hand van deze protocollen de elektronische vragenlijst en rapporteren via een standaardformulier hun bevindingen aan de programmeur. Het testen is een iteratief proces, dat meestal na drie iteraties kan worden afgesloten. Het WODC en aanvragers van J&V en VWS hebben ook mee getest.

3.4 Resultaten veldwerk

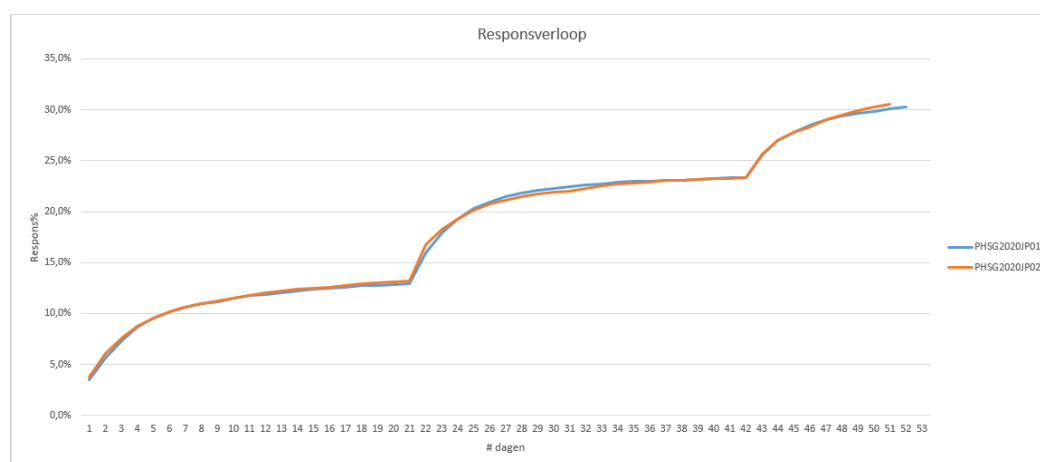
1. Respons

Het WODC heeft wekelijks een voortgangsoverzicht ontvangen. In tabel 5 is het definitieve responsoverzicht opgenomen. Met 30,5% is het eindresultaat 7,5 procentpunten boven verwachting. Er zijn 30.459 responsen gerealiseerd, 7.459 meer dan verwacht.

Tabel 5. Definitief responsoverzicht PHGSG 2020

	Aantal	%
Uitgezette steekproef	100.000	100,0
Non respons		
Geen reactie	67.519	67,5
Non respons	2.022	2,0
Verhuisd	452	0,5
Weigering	1.018	1,0
Geen tijd	85	0,1
Benadergegevens niet correct	8	0,0
Overleden	78	0,1
Taalbarrière	35	0,0
Niet in staat deel te nemen	155	0,2
Overig	191	0,2
Respons	30.459	30,5

Figuur 2 geeft het responsverloop in de tijd per steekproefportie. Het verloop is voor beide porties nagenoeg identiek. Het effect van de rappelbrieven is duidelijk aanwezig.

Figuur 2. Responsverloop PHGSG 2020

Tabel 6 geeft per device het percentage van de steekproefpersonen dat heeft ingelogd maar uiteindelijk de vragenlijst heeft afgebroken. Dit aandeel is het hoogst op smartphones. . Nagenoeg alle afbrekingen vinden plaats bij de start van de vragenlijst, dus er is geen duidelijke vraag of aanwijsbaar onderwerp in de vragenlijst waar dit gebeurt. Enkele achtergrondkenmerken van de afbrekers zijn bekeken maar er zijn geen duidelijke conclusies te verbinden aan een specifieke groep.

Tabel 6. Afgebroken interviews naar device

	Ingelogd	Afgebroken	% afgebroken
Computer	23.508	1.980	8,4
Smartphone	6.557	941	14,4
Tablet	3.394	357	10,5
Overig	321	43	13,4

2 Realisatie

Het veldwerk is volledig conform planning uitgevoerd. Desondanks zijn er twee bijzonderheden op te merken:

1. Tegen het einde van het veldwerk is op basis van een melding van een respondent geconstateerd dat brieven zijn uitgestuurd zonder voorletters van de potentiële respondent. Onderzoek heeft uitgewezen dat dit het geval is geweest voor alle uitgestuurde rappelbrieven door een fout in de inrichting van het enquête administratie systeem Phoenix. Alhoewel het vreemd overkomt op de brieven, heeft dit de datakwaliteit niet beïnvloed. In de vragenlijst zelf wordt na inloggen duidelijk aangegeven voor wie de vragenlijst bedoeld is. Zowel de voorletters, de achternaam als de geboortedatum worden getoond.
2. Het veldwerk was net een week gestart toen de effecten van het coronavirus zich in volle omvang manifesteerden in Nederland (rond 13 maart) en het uitdrukkelijk verzoek uitging van de overheid om zoveel als mogelijk thuis te werken. De invloed op het responsverloop lijkt in zoverre aanwezig dat het gerealiseerde responsresultaat beduidend hoger is dan vooraf werd verwacht. Opvallend is dat tussen de eerste steekproefportie waarvoor het invullen op 6 maart kon worden gestart en de tweede steekproefportie waarvoor het invullen op 13 maart kon worden gestart geen verschil in responsverloop zichtbaar is. Dit geeft enige indicatie dat de invloed van (de maatregelen omtrent) het coronavirus op het responsverloop toch meevalt. Met zekerheid valt dit natuurlijk niet te zeggen.

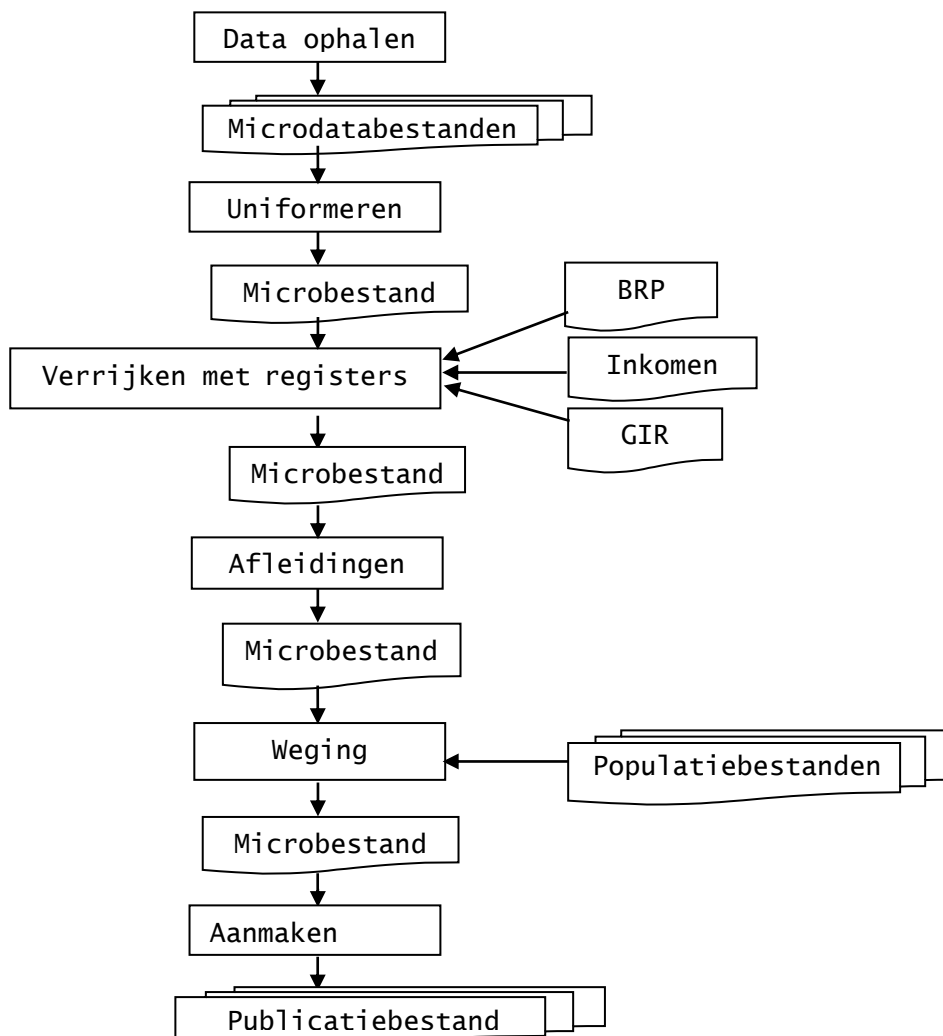
4. Verwerken (dataprocessing)

De term dataprocessing wordt gebruikt om de werkzaamheden samen te vatten die nodig zijn om aan de antwoorden die respondenten geven op de vragen in de vragenlijst plausible statistische informatie te kunnen ontlelen. De verwerking van de data is zoveel mogelijk geautomatiseerd. Het verwerken is uitgevoerd door statistische afdeling, sector SAL/SDV van de Divisie SER.

4.1 Inleiding

Onderstaande figuur geeft het verwerkingsproces van de onderzoeksdata in hoofdlijnen aan. De deelprocessen worden in de volgende paragrafen in meer detail beschreven.

Figuur 1. Procesmodel Verwerkingsproces



4.2 Invoerdata

De data zijn afkomstig van enquêtes uitgezet via het internet (CAWI) en worden elektronisch opgeslagen in het enquêtesysteem van het CBS. De range- en routingcontrole zijn vastgelegd in de vragenlijstschema's en zijn vervolgens in de elektronische vragenlijsten ingebouwd. Deze controles worden automatisch uitgevoerd tijdens het afnemen van de enquête.

Rangecontroles zijn o.a.:

- formaat van variabele,
- variabele heeft juiste naam,
- waarde bereik van variabele conform afspraak,
- dubbele sleutels.

Routingcontroles zijn o.a.:

- indien een blok niet op de route ligt, mogen de variabelen ook geen waarde hebben,
- indien een blok wel op de route ligt moeten deze in principe (tenzij veld niet verplicht) ook een waarde hebben.

Er worden ook variabelen opgeleverd die lange teksten kunnen bevatten. Die worden gecontroleerd op inhoud die mogelijk identificerend zijn en indien nodig aangepast.

4.3 Verrijken

Na het samenvoegen van alle enquêtedata in één bestand, is het bestand verrijkt met registerinformatie. Deze informatie is afkomstig uit de Gemeentelijke Basisadministratie (GBA), het GIR (gebiedsindelingen register) en uit de inkomensgegevens.

De GBA-gegevens van 2020 zijn gebruikt voor verrijking. Hierbij is zo recent mogelijke informatie gebruikt, gebaseerd op de enquêtedatum. De meest recente inkomensgegevens die voor het CBS beschikbaar zijn, zijn de (voorlopige) inkomensgegevens van 2018. Regionale gegevens zijn verrijkt op enquêtedatum. Hierbij is de codelijst gebruikt die vastgesteld is op 1 januari 2020.

In de volgende tabel is weergegeven welke registervariabelen zijn toegevoegd. De tweede kolom bevat de variabelenamen zoals deze voorkomen in het geleverde SPSS-bestand. In de derde kolom is vermeld waar de variabele vandaan komt.

Tabel 1. Overzicht verrijkingsvariabelen

Variabele omschrijving	Naam variabele bestand	Bron
Landsdeel	LANDSDEEL	GIR 2020
Provincie	PROV	GIR 2020
Stedelijkheid	STEDGEM	GIR 2020
Postcode (cijfers)	Postcode_num	GIR 2020
Type huishouden	TypeHH	GBA 2020
herkomstgroepering	ETNGROEPKORT1	GBA 2020
Gestandaardiseerd inkomen particuliere huishoudens in kwintielen	Afl_GestinkH5	Inkomen 2018
Indeling vermogen particuliere huishoudens (in kwintielen)	Afl_VermogenH5	Inkomen 2018
Indeling welvaart (in kwintielen)	Afl_Welvaart5	Inkomen 2018
sociaaleconomische categorie op jaarbasis	INPSECJ	Inkomen 2018

4.5 Afleiden

Naast registervariabelen zijn er ook afgeleide variabelen toegevoegd. Deze worden toegevoegd voor analyse en publicatie doeleinden, of als deze variabelen in de weging (zie H8) gebruikt worden. Ook zijn meerkeuzevragen omgezet naar dummy variabelen (dichotomiseren).

In onderstaande tabel is aangegeven welke afgeleide variabelen gemaakt zijn.

Tabel 2. Overzicht afgeleide variabelen

Variabele omschrijving	Naam variabele bestand	Overige opmerking
Enquetedatum	Afl_peildatum	-
Burgerlijke staat OP	Afl_BurgSt	Op enquête datum
Geslacht OP	Afl_Geslacht	Op enquête datum
Leeftijd onderzoekspersoon (OP)	Afl_Leeftijd	Op enquête datum
Herkomst in drie categoriën (achtergrondbestand/indikking)	Afl_Herkomst3	Onbekenden in categorie 1.
Type huishouden (SSB/indikking) in 7 categorieën	Afl_Typehuishouden	Onbekenden in categorie 7.
Afgelopen 5 jaar Verbale Agressie meegemaakt	Afl_VerbalAgres5Jr	Voor analyse
Afgelopen 5 jaar Lichamelijke Agressie meegemaakt	Afl_LichAgres5Jr	Voor analyse
Afgelopen 5 jaar Coercive Control meegemaakt	Afl_CoCo5Jr	Voor analyse
Afgelopen 5 jaar ongewenst Seksueel Gedrag Online meegemaakt	Afl_Online5Jr	Voor analyse
Afgelopen 5 jaar ongewenst Seksueel Gedrag zonder aanraking meegemaakt (offline)	Afl_Offline5Jr	Voor analyse
Afgelopen 5 jaar ongewenst Seksueel Gedrag met aanraking meegemaakt	Afl_HandsOn5Jr	Voor analyse
Afgelopen 12 maanden Verbale Agressie meegemaakt	Afl_VerbalAgres12M	Voor analyse
Afgelopen 12 maanden Lichamelijke Agressie meegemaakt	Afl_LichAgres12M	Voor analyse
Afgelopen 12 maanden Coercive Control meegemaakt	Afl_CoCo12M	Voor analyse
Afgelopen 12 maanden ongewenst Seksueel Gedrag Online meegemaakt	Afl_Online12M	Voor analyse
Afgelopen 12 maanden ongewenst Seksueel Gedrag zonder aanraking meegemaakt (offline)	Afl_Offline12M	Voor analyse
Afgelopen 12 maanden ongewenst Seksueel Gedrag met aanraking meegemaakt	Afl_HandsOn12M	Voor analyse
Afgelopen 12 maanden zelf pleger geweest	Afl_Dader	Voor analyse
Eindgewicht weging	Gewicht	Om te corrigeren voor scheve respons en op te hogen naar doelpopulatie

Voor de afleiding van sommige variabelen zijn wel voor missende waarden inhoudelijke beslissingen gemaakt waar deze bij te voegen.

Bijvoorbeeld:

Om te bepalen of iemand de afgelopen 5 jaar verbale agressie heeft meegemaakt (Afl_VerbalAgres5Jr) wordt er gekeken naar alle gebeurtenissen die onder verbale agressie vallen. Als er op minimaal 1 van deze gebeurtenissen positief wordt geantwoord is de persoon slachtoffer van fysieke geweld. Als op geen enkele gebeurtenis positief is geantwoord, maar er

op een of meerdere van deze gebeurtenissen geen antwoord is gegeven is het onbekend of deze persoon slachtoffer is geweest. Dit onbekende percentage wordt apart gepubliceerd.

Verder zijn de meerkeuzevragen omgezet naar dummy variabelen (dichotomiseren). Als een vraag niet op de route ligt en dus niet gesteld hoefde te worden, zijn de gedichotomiseerde variabelen leeg.

In de vragenlijst kwam het geregeld voor dat de respondent meerdere antwoorden kon geven. Per vraagvariabele levert dit voor het op te leveren SPSS-bestand dus een aantal bestandsvariabelen op, namelijk zoveel als er antwoordcategorieën zijn. Omdat de vraag vaak ook nog betrekking had op de vraag door wie dit gedaan was of op welke situatie dit betrekking had, werd het aantal variabelen met dezelfde naam nog groter; het enige onderscheid tussen de variabelen was een tellertje achteraan de variabelenaam. Daarom is ervoor gekozen om de variabelen opnieuw in de delen met variabelenamen die de inhoud duidelijk weergeven. In de variabelenaam is te zien uit welk vragenblok het komt (eerste 3 letters), welke antwoordcategorie het betreft (middenstuk) en of het om de afgelopen 5 jaar, afgelopen 12 maanden of een wie- of wat-vraag betreft (laatste stuk van variabelenaam). De verschillende stukken van de variabelenaam worden onderscheiden door een laag streepje (_).

Hieronder enkele voorbeelden:

VeA_Bel_afg5Jr

Blok Verbale agressie, vraag naar beledigen, in de afgelopen 5 jaar

LiA_FrPijn_12M

Blok Lichamelijke agressie, frequentievraag naar dreiging lichamelijke pijn, afgelopen 12 maanden

Ctl_Weg_Wie01

Blok Coercive Control, vraag naar wie het slachtoffer weg hield bij familie/vrienden, pleger 1 (van maximaal 12 plegers)

Sta_Gev1

Blok Stalking, vraag naar gevolgen van stalking, gevolg 1 (van maximaal 8 gevolgen)

5. Responsanalyse

Om de kwaliteit van de behaalde respons op de Prevalentiemonitor Huiselijk Geweld en Seksueel Geweld 2020 (PHGSG) te bepalen, is een (non-)responsanalyse uitgevoerd. Hiermee kan worden vastgesteld of de respons een representatieve weerspiegeling is van de doelpopulatie van het onderzoek. In hoofdstuk 3 zijn de steekproef en de behaalde strenge respons van de dataverzameling al beschreven. In dit hoofdstuk wordt bekeken in hoeverre bevolkingsgroepen verschillen in hun respons, en in hoeverre de verdeling van respondenten over deze bevolkingsgroepen overeenkomt met de verdeling van de groepen in de populatie. Dit verschaft inzicht in de mate van selectiviteit.

5.1 Selectiviteit van de respons

Zoals ook beschreven in hoofdstuk 3, paragraaf 1 en 4 zijn er 111500 steekproefpersonen getrokken waarvan waren er 102981 bruikbaar. Uiteindelijk zijn er 100000 personen daadwerkelijk benaderd. Hiervan hebben 30459 personen deelgenomen aan het onderzoek. Dit komt neer op een respons van 30,5 procent.

Voor de analyse van de responscijfers is gebruik gemaakt van achtergrondinformatie die beschikbaar is voor alle steekproefpersonen. Deze informatie is hoofdzakelijk afkomstig uit de Basis Registratie Personen (BRP), zoals bijvoorbeeld geslacht, leeftijd, migratieachtergrond en type huishouden. Ook is er regionale informatie beschikbaar, zoals stedelijkheid van de woongemeente. Voor elk steekproefpersoon is ook inkomensinformatie (van het huishouden) beschikbaar.

De respons is selectief naar bevolkings- en regionale kenmerken (Tabel 1). Voor geslacht is alleen de indeling man/vrouw mogelijk omdat momenteel nog niet met een andere categorie wordt gewerkt in de vragenlijst en ook niet bij de weging. In de toekomst komen hier andere mogelijkheden voor. De categorieën beschreven in tabel 1 zijn afgestemd op het model van de weging en zijn niet altijd gelijk aan de indeling gebruikt bij de publicatie van deze cijfers (bijvoorbeeld de indeling van leeftijd wijkt af omdat op het moment van weging de leeftijdsindeling voor publicatie nog niet bekend was). De hieronder beschreven verschillen tussen categorieën van deze kenmerken zijn allemaal significant. De achtergrondkenmerken zijn afgestemd met het model voor de weging en de indeling in leeftijd is ook overgenomen uit het model voor de weging (zie hoofdstuk 8).

De jongere leeftijdscategorieën nemen minder vaak deel dan de oudere leeftijdsgroepen. De 65- tot 73-jarigen hebben met 46 procent juist relatief vaak deelgenomen aan het onderzoek. Gehuwden nemen, met 37 procent, vaker deel dan gescheiden en verweduwde steekproefpersonen en vooral degenen die nooit gehuwd zijn geweest. Van deze laatste groep heeft 24 procent deelgenomen.

Paren zonder kinderen hebben ook relatief vaker meegedaan aan het onderzoek dan paren met kinderen, alleenstaanden en eenouderhuishoudens.

Vooral bij migratieachtergrond is een sterk verschil te zien in de respons; slechts 16 procent van de mensen met een niet-westerse migratieachtergrond heeft deelgenomen, tegenover 27 procent van de mensen met een westerse achtergrond en 33 procent van de mensen met een Nederlandse achtergrond. De respons neemt toe naarmate het gestandaardiseerd besteedbaar inkomen van het huishouden hoger is.

Het beeld voor de respons naar achtergrondkenmerken stemt overeen met dat van andere CBS-onderzoeken.

Tabel 1. Respons naar achtergrondkenmerken, PHGSG 2020

	Netto st.pr.(abs.)	Respons (abs.)	Respons
<i>Geslacht</i>			
Man	49572	15125	30,5%
Vrouw	50428	15334	30,4%
<i>Leeftijd</i>			
16 – 23 jaar	11475	8607	25,0%
24 – 33 jaar	15536	12098	22,1%
34 – 48 jaar	21976	16328	25,7%
49 – 63 jaar	26984	17896	33,7%
64 - 73 jaar	12912	6915	46,4%
74 jaar of ouder	11117	7697	30,8%
<i>Burgerlijke staat</i>			
Gehuwd (incl. partnerschap)	46864	17085	36,5%
Gescheiden	9344	2601	27,8%
Verweduwd	5844	1498	25,6%
Nooit gehuwd geweest	37948	9275	24,4%
<i>Provincie</i>			
Groningen	3376	1046	31,0%
Friesland	3772	1102	29,2%
Drenthe	2827	950	33,6%
Overijssel	6710	2114	31,5%
Flevoland	2420	637	26,3%
Gelderland	11973	3883	32,4%
Utrecht	7754	2439	31,5%
Noord-Holland	16592	4680	28,2%
Zuid-Holland	21231	5944	28,0%
Zeeland	2216	687	31,0%
Noord-Brabant	14716	4900	33,3%
Limburg	6413	2077	32,4%
<i>Type huishouden</i>			
Eenpersoonshuishouden	21895	5392	24,6%
Niet-gehuwd paar zonder kinderen	8432	2579	30,6%
Gehuwd paar zonder kinderen	24208	10486	43,3%
Niet-gehuwd paar met kinderen	7329	1863	25,4%
Gehuwd paar met kinderen	30079	8508	28,3%
Eenouderhuishouden	7268	1484	20,4%

Vervolg Tabel 1

Overig huishouden	746	134	18,0%
<i>Achtergrond</i>			
Nederlandse achtergrond	76731	25516	33,3%
Westerse migratieachtergrond	10910	2922	26,8%
Niet-westerse migratieachtergrond	12359	2021	16,4%
<i>Huishoudensinkomen*</i>			
Eerste (laagste) kwintiel	13882	2764	19,9%
Tweede kwintiel	16968	4212	24,8%
Derde kwintiel	19835	6120	30,9%
Vierde kwintiel	22360	7749	34,7%
Vijfde (hoogste) kwintiel	23951	9055	37,8%

*n = 3004: onbekend

Aangezien de achtergrondkenmerken ook bekend zijn voor degenen die niet gerespondeerd hebben, kan de mate van selectiviteit ook vastgesteld worden door een vergelijking met de totale steekproef. Ook kan de verdeling vergeleken worden met de populatieverdeling (16 jaar of ouder). In Tabel 2 worden de populatieverdeling, de verdeling van de steekproef en de verdeling van de respons weergegeven.

Het ligt voor de hand dat de groepen met een hogere respons op de enquête relatief gezien ook vaak oververtegenwoordigd zijn ten opzichte van de steekproef en de populatie. Het gaat dan om personen tussen 65 en 73 jaar, gehuwden (zonder kinderen), mensen met een Nederlandse achtergrond en mensen met een hoog huishoudensinkomen. Groepen met een lage respons zijn vaak ondervertegenwoordigd; in dit geval zijn dat de jongeren, ongehuwden, mensen met een niet-westerse migratieachtergrond en degenen in de laagste twee inkomenskwartielen. Ook dit beeld is vergelijkbaar met andere CBS-onderzoeken en daarmee is PHGSG niet anders dan andere onderzoeken qua selectiviteit. De weging (hoofdstuk 8) corrigeert zo goed mogelijk voor deze verschillen in selectiviteit waardoor je mag stellen dat de onderzoekspopulatie van PHGSG representatief is voor de doelpopulatie.

Tabel 2. Verdelingen populatie, steekproef en respons, PHGSG 2020

	Populatie (16+)	Steekproef	Respons (ongewogen)
<i>Geslacht</i>			
Man	49,5%	49,6%	49,7%
Vrouw	50,5%	50,4%	50,3%
<i>Leeftijd</i>			
16 – 23 jaar	12,0%	11,5%	9,4%
24 – 33 jaar	15,6%	15,5%	11,3%
34 – 48 jaar	22,3%	22,0%	18,5%
49 – 63 jaar			
Vervolg Tabel 2	27,2%	27,0%	29,8%

64 - 73 jaar	12,5%	12,9%	19,7%
74 jaar of ouder	10,5%	11,1%	11,2%

Burgerlijke staat

Gehuwd (incl. partnerschap)	46,7%	46,9%	56,1%
Gescheiden	9,4%	9,3%	8,5%
Verweduwd	5,6%	5,8%	4,9%
Nooit gehuwd geweest	38,3%	37,9%	30,5%

Provincie

Groningen	3,4%	3,4%	3,4%
Friesland	3,7%	3,8%	3,6%
Drenthe	2,8%	2,8%	3,1%
Overijssel	6,6%	6,7%	6,9%
Flevoland	2,3%	2,4%	2,1%
Gelderland	11,9%	12,0%	12,7%
Utrecht	7,6%	7,8%	8,0%
Noord-Holland	16,6%	16,6%	15,4%
Zuid-Holland	21,2%	21,2%	19,5%
Zeeland	2,2%	2,2%	2,3%
Noord-Brabant	14,8%	14,7%	16,1%
Limburg	6,6%	6,4%	6,8%

Type huishouden

Eenpersoonshuishouden	21,7%	21,9%	17,7%
Niet-gehuwd paar zonder kinderen	8,4%	8,4%	8,5%
Gehuwd paar zonder kinderen	23,6%	24,2%	34,4%
Niet-gehuwd paar met kinderen	7,5%	7,3%	6,1%
Gehuwd paar met kinderen	30,7%	30,1%	27,9%
Eenouderhuishouden	7,4%	7,3%	4,9%
Overig huishouden	0,7%	0,7%	0,4%

Achtergrond

Nederlandse achtergrond	76,3%	76,7%	83,8%
Westerse migratieachtergrond	10,9%	10,9%	9,6%
Niet-westerse migratieachtergrond	12,8%	12,4%	6,6%

Huishoudensinkomen

Eerste (laagste) kwintiel	14,2%	13,9%	9,1%
Tweede kwintiel	16,7%	17,0%	13,8%
Derde kwintiel	19,7%	19,8%	20,1%
Vierde kwintiel	22,2%	22,4%	25,4%
Vijfde (hoogste) kwintiel	24,7%	24,0%	29,7%

6. Andere kwaliteitsindicatoren

Behalve door naar de respons en de verdeling van de respons naar achtergrondkenmerken te kijken, kan ook op andere manieren een indruk gekregen worden van de kwaliteit van het onderzoek, namelijk door te kijken naar het percentage respondenten dat geen antwoord geeft op vragen.

6.1 Geen antwoord op vragen

Er is een relatief klein deel, van 145 respondenten, dat voor elke gebeurtenis van de afgelopen 5 jaar geweigerd heeft antwoord te geven, inclusief de vragen naar slachtofferschap voor 12 jaar en de vraag naar of men zelf tegen iemand uit de huiselijke kring lichamelijk agressief, controlerend en/of intimiderend was, ongewenst seksueel gedrag liet zien of een ex-partner heeft gestalkt.

Deze respondenten zijn vaker man, hebben vaker een niet-westerse migratieachtergrond en/of een laag inkomen. Maar de aantallen zijn zo klein dat er geen conclusies aan verbonden mogen worden.

Deze respondenten zijn wel meegenomen in de analyses aangezien ze wel door de hele vragenlijst heen zijn gegaan (een voorwaarde om hem naar het CBS te kunnen verzenden) en ze ook meegenomen zijn in de weging.

6.2 Veldwerk tijdens corona-uitbraak

Eerder is gemeld dat het veldwerk net een week gestart was toen het coronavirus zich manifesteerde in Nederland en de 'intelligente lockdown' werd ingesteld, die bestond uit verschillende maatregelen, waaronder het uitdrukkelijk verzoek uitging van de overheid om zoveel mogelijk thuis te werken. De invloed op het responsverloop lijkt mee te vallen (zie hoofdstuk 3). De vraag is natuurlijk wel in hoeverre de maatregelen omtrent de corona-uitbraak het beantwoorden van vragen heeft beïnvloed en in hoeverre de maatregelen invloed hadden op de daadwerkelijke omvang van huiselijk geweld en seksueel geweld. De resultaten van het onderzoek van de respondenten die vóór 15 maart meededen wijken af van de resultaten van de degenen die na 15 maart meededen. Zo meldden de respondenten die de enquête na 15 maart invulden bij alle gebeurtenissen minder vaak slachtofferschap. Deze verschillen zijn significant, behalve bij stalking [<verwijzen naar maatwerktabellen>](#).

We kunnen niet direct concluderen dat de coronamaatregelen de prevalentiecijfers van huiselijk en seksueel geweld gedrukt hebben omdat er meer speelt. Het kan zijn dat slachtoffers geneigd zijn om eerder met een enquête mee te doen dan niet-slachtoffers en dat daardoor de prevalentiecijfers in het begin van de waarneemperiode hoger liggen. Dit is eerder bevonden in vergelijkbare enquêtes en dit staat dan los van de corona-uitbraak. Voor dit onderzoek geldt dat 71 procent van de respondenten de vragenlijst op of na 15 maart heeft ingevuld.

Als de corona-uitbraak er wel voor gezorgd heeft dat de prevalentiecijfers omlaag gingen is het niet duidelijk of er in het begin van de intelligente lockdown daadwerkelijk minder slachtoffers waren, dat dit het invulgedrag beïnvloedde of dat er een andere groep respondenten mee heeft gedaan die anders niet zou hebben meegedaan.

Daar de bevroegde periode een jaar betrof en de coronatijd een relatief klein deel daarvan was lijkt het waarschijnlijker dat het invulgedrag beïnvloed is. Deze aanname wordt versterkt doordat bij alle gebeurtenissen het percentage van de respondenten dat ná 15 maart geen

antwoord gaf op de prevalentievragen hoger ligt dan vóór 15 maart. Mogelijk was de respondent vaker niet alleen en keek er dus vaker iemand over de schouder mee.

Daarnaast kan het zijn dat een andere groep respondenten meedoet na 15 maart, doordat men meer tijd heeft gehad om een enquête in te vullen (dus meer niet-slachtoffers).

Anderzijds is er misschien werkelijk minder geweld geweest. Veel van de bevraagde voorvallen die buiten de huiselijke kring plaatsvinden kunnen tijdens een lockdown immers niet of nauwelijks voorkomen omdat potentiële plegers en slachtoffers elkaar niet kunnen/mogen ontmoeten. Verder was de aandacht wellicht gefocust op corona.

7. Plausibiliteitsanalyse

In deze paragraaf wordt ingegaan op de plausibiliteit van een aantal kernindicatoren van de Prevalentiemonitor Huiselijk Geweld en Seksueel Geweld. Allereerst worden de verdelingen ervan getoond, zowel voor als na weging. Daarna volgt een vergelijking van de bevindingen van PHGSG met bevindingen van voorgaand onderzoek over huiselijk en seksueel geweld’.

7.1 Verdeling kernindicatoren

Als kernindicatoren is gekeken naar het meemaken van verbale agressie, fysiek geweld en dwingende controle in de huiselijke kring in de afgelopen 5 jaar. Verder gaat het om niet-fysieke seksuele intimidatie, fysiek seksueel geweld en online seksuele intimidatie in de afgelopen 5 jaar.

Tabel 1 toont de uitkomsten van deze 5-jaars prevalenties voor en na weging. De weging zorgt ervoor dat bij alle vormen van geweld en intimidatie de prevalentie iets omhoog gaat. Mogelijk trekt de weging een ondervertegenwoordiging van mensen die slachtoffer zijn van geweld en intimidatie recht, maar het gaat niet om extreme verschillen.

Verdere analyses van doelvariabelen, ook naar achtergrondkenmerken, zijn beschreven in de inhoudelijke rapportage [<verwijzen>](#).

Tabel 1. Kernindicatoren ongewogen en gewogen, PHGSG 2020

	Ongewogen (%)	Gewogen (%)
Afgelopen 5 jaar verbale agressie in huiselijke kring meegemaakt		
Nee	54,0	51,7
Ja	43,6	45,6
Weigert	2,5	2,7
Afgelopen 5 jaar fysiek geweld in huiselijke kring meegemaakt		
Nee	92,7	91,3
Ja	5,8	6,9
Weigert	1,5	1,8
Afgelopen 5 jaar dwingende controle in huiselijke kring meegemaakt		
Nee	89,7	88,1
Ja	8,1	9,3
Weigert	2,2	2,6
Afgelopen 5 jaar online seksuele intimidatie meegemaakt		
Nee	90,4	88,9
Ja	7,8	8,9
Weigert	1,8	2,2
Afgelopen 5 jaar niet-fysieke seksuele intimidatie meegemaakt (offline)		
Nee	87,6	86,1
Ja	10,5	11,7
Weigert	1,9	2,2
Afgelopen 5 jaar fysiek seksueel geweld meegemaakt		
Nee	91,5	90,3
Ja	6,8	7,6
Weigert	1,7	2,1

7.1 Vergelijking met ander onderzoek

Om enig inzicht te kunnen krijgen in de plausibiliteit van een onderzoek kan een vergelijking worden gemaakt met ander onderzoek dat gedaan is over hetzelfde onderwerp. Het is voor het eerst dat PHGSG wordt uitgevoerd door het CBS dus er kan geen exacte vergelijking gemaakt worden met eerdere jaren.

Wel heeft het CBS eerder het onderzoek 'Persoonlijke veiligheid' (PV) middels papier en internet uitgevoerd, in opdracht van het WODC, onder bijna 7 duizend mannen en vrouwen van 18 jaar of ouder). Daarnaast zijn er andere externe onderzoeken geweest op het gebied van huiselijk en seksueel geweld: huiselijk geweld (HG), een online (panel)enquête onder ruim 6 duizend mannen en vrouwen van 18 jaar of ouder, in 2008 uitgevoerd door Intomart en daarnaast 'Violence against women', een face-to-face enquête in 2012 uitgevoerd in de EU

door FRA– European Union Agency for Fundamental Rights. Voor deze laatste zijn in Nederland 1500 vrouwen van 18-74 jaar geënquêteerd. Getracht is een vergelijking te maken tussen de uitkomsten van deze onderzoeken en PHGSG 2020.

Tabel 2 laat zien wat er in de verschillende onderzoeken wordt gevraagd en welke aspecten wel en niet, onder welke noemer, tot het officiële gepubliceerde cijfer worden gerekend. De vergelijking is alleen gericht op de prevalentie van lichamelijk-, en seksueel geweld en seksuele intimidatie, al dan niet in huiselijke kring. Cijfers over stalking en psychisch geweld uit voorgaande onderzoeken zijn niet vergelijkbaar met PHGSG. In PHGSG wordt alleen naar stalking door een ex-partner gevraagd. Door FRA is een cijfer voor stalking, ongeacht door wie, gepubliceerd, en in de onderzoeken uit 2008 en 2017 werd stalking niet goed uitgevraagd. Hetzelfde geldt voor ernstig psychisch geweld.

Tabel 2. Bevroagde items in de verschillende onderzoeken behorende bij de hoofdcategorieën lichamelijk en seksueel geweld/ ongewenst lichamelijk contact

	FRA	HG	PV	PHSG
		Lichamelijk geweld en stalking		
voortdurend in de gaten gehouden/ gestalkt		x		
		Physical violence	Lichamelijk geweld	Lichamelijke agressie
geslagen met een voorwerp	x	x	x	
gepoogd te verstikken, wurgen of branden	x	x	x	x
bedreigd met mes of wapen		x	x	x
verwond met mes of wapen	x	x	x	x
met de dood bedreigd				x
hoofd ergens tegenaan geslagen	x			
		Agressief gedrag		
bedreigd met lichamelijk pijn doen	?	x	x	x
bekogeld met voorwerp	x	x	x	x
geduwd, arm verdraaid, aan haar getrokken	x	x	x	x
geslagen, geschopt, gebeten of gestompt	x	x	x	x
andere vorm van lichamelijk geweld		x	x	x
+ 4 andere aspecten		x		
		Serious sexual harassment		Ongewenste seksuele ervaringen zonder lichamelijk contact
seksueel kwetsende opmerkingen of grapjes	x			x
ongewenst naaktfoto's of seksfilmpjes toegestuurd	x			x
ongewenst iemand naakt moeten zien/				
ongewenst iemand moeten zien masturberen	x			x
gedwongen naar porno te kijken	x			
beledigende seksueel getinte e-mails of SMS ontvangen	x			
+veel andere items, online of offline				x
				Ongewenst lichamelijk contact
ongewenst aangeraakt/ ongewenst gezoend	x			x
		Sexual violence	Seksueel geweld	Seksueel geweld
verkracht (oraal/ vaginaal/ anaal)	x	x	x	x
gepoogd seks op te dringen	x	x	x	
door (ex-)partner gedwongen tot seksuele handelingen		x	x	
gedwongen tot seksuele handelingen ¹	x			x
ander seksueel geweld		x	x	
ongewenst afgetrokken of gevingerd				x
geprostitueerd				x
afpersing met seksuele handelingen	x			

¹ bij PHSG wordt dit benaderd door de toevoeging ' >>Het kan zowel zijn dat iemand dit bij u deed of dat u dit bij een ander moest doen.<<' bij de vragen over verkrachting en seksuele handelingen

Uit tabel 2 blijkt dat de vragen over lichamelijk en seksueel geweld in alle onderzoeken ongeveer gelijk zijn gebleven, maar dat de aspecten niet allemaal onder dezelfde noemer vallen. Het belangrijkste en meest opvallende verschil tussen PHGSG en de eerder onderzoeken is dat bij PHGSG niet wordt gesproken over seksueel geweld, maar over 'ongewenst lichamelijk contact'. Hiertoe behoort ook 'ongewenst aanraken', hetgeen poging tot verkrachting kan zijn maar grotendeels lichtere vormen zal betreffen (die in FRA onder de noemer 'seksuele intimidatie' vallen, en in de andere onderzoeken niet voorkomen).

In de onderzoeken uit 2008 en 2017 werd de 'huiselijke kring' breder gedefinieerd dan in het onderhavige onderzoek; ook huisgenoten (anders dan familie) en huisvrienden werden hiertoe gerekend. In 2008 werd 2 à 3 procent van het huiselijke geweld door een huisgenoot of huisvriend gepleegd. In 2017 lag dat een stuk hoger, namelijk bij 10 procent van de vrouwelijke slachtoffers en 19 procent van de mannelijke slachtoffers. Bij een vergelijking met het PHGSG-cijfer zou dit niet zomaar 1 op 1 kunnen.

Voor nu moet de conclusie worden getrokken dat de landelijke enquête-onderzoeken die in het verleden zijn gedaan naar seksueel geweld of geweld in de huiselijke kring niet goed vergeleken kunnen worden met PHGSG. De onderzoeken lijken qua vraagstelling sterk op elkaar, maar de afbakening van de hoofdgebeurtenissen is in PHGSG iets anders. Daarnaast is het onderzoeksdesign niet hetzelfde, worden verschillende doelgroepen benaderd en zijn de onderzoeken in andere jaren uitgevoerd.

PHGSG zal om de twee jaar uitgevoerd worden. We moeten er vanuit gaan dat de cijfers de werkelijkheid goed benaderen. Het onderzoek zal hierdoor een betrouwbaar beeld geven van de ontwikkelingen door de tijd en een volgende uitvoering van het onderzoek kan dan vergeleken worden met deze versie uit 2020.

8. Weging

8.1 Inleiding

Om met de responsgegevens schattingen te kunnen maken van populatie-eigenschappen worden de responsgegevens gewogen. Dit wegen van de responsgegevens bestaat uit het berekenen van een insluitgewicht, π_i^{-1} , en twee correctiegewichten, $c_i^{(1)}$ en $c_i^{(2)}$, voor iedere persoon i die heeft gerespondeerd. Het insluitgewicht moet corrigeren voor de mogelijke verschillen in trekkansen die het steekproefontwerp kent. De twee correctiegewichten moeten corrigeren voor de mogelijk selectieve uitval bij het bewerken van de steekproef en de non-respons na het uitzetten. Het product van het insluitgewicht en de twee correctiegewichten is het eindgewicht,

$$w_i^{(end)} = \pi_i^{-1} \cdot c_i^{(1)} \cdot c_i^{(2)}.$$

Om de populatie-eigenschappen te schatten moeten de responsgegevens met dit eindgewicht gewogen worden. Het product van het insluitgewicht en het eerste correctiegewicht wordt ook wel het startgewicht genoemd,

$$w_i^{(start)} = \pi_i^{-1} \cdot c_i^{(1)}$$

Het insluitgewicht is gelijk aan de inverse van de trekkingskans, de kans die een persoon heeft om getrokken te worden in de steekproef. Het steekproefontwerp dat is gebruikt is zó, dat alle personen met gelijke kansen in de te benaderen steekproef terecht zijn gekomen. Deze kans is gelijk aan de steekproeffractie: aantal te benaderen personen in de steekproef n , gedeeld door de populatieomvang N . Gezien het insluitgewicht gelijk is aan de inverse van deze kans wordt de formule als volgt:

$$\pi_i^{-1} = \frac{N}{n}$$

Het eerste correctiegewicht, dat alleen corrigeert voor de totale uitval door non-respons, is gelijk aan de verhouding van het aantal personen in de steekproef en het aantal respondenten

$$c_i^{(1)} = \frac{n}{r}.$$

Door de formules van het insluitgewicht en het eerste correctiegewicht in te vullen voor het startgewicht wordt duidelijk dat dit alleen afhankelijk is van het populatietotaal N en de responsomvang r :

$$w_i^{(start)} = \pi_i^{-1} \cdot c_i^{(1)} = \frac{N}{n} \cdot \frac{n}{r} = \frac{N}{r}$$

8.2 Weegmodel

De populatie van de PHGSG bestaat uit inwoners in Nederland van 16 jaar en ouder. Bij de weging is gebruik gemaakt van de populatietotalen van 1 maart 2020. Het populatietotaal is op deze peildatum: 14.252.565. Verder is de steekproefomvang n gelijk aan 100.000 en de responsomvang r gelijk aan 30.459. Hiermee kunnen het insluitgewicht, het eerste correctiegewicht en het startgewicht worden uitgerekend. Het insluitgewicht is gelijk aan 143, het eerste correctiegewicht is gelijk aan 3,283 en het startgewicht is 467,93. Door te corrigeren voor de uitval door non-respons wordt ook de som van alle startgewichten gelijk aan het populatietotaal voor de weging, te weten 14.252.565, en daarmee wordt het gemiddelde correctiegewicht 1.

Het tweede correctiegewicht dat moet corrigeren voor de selectieve uitval door non-respons, wordt met de lineaire regressieschatter berekend. Hiervoor moet eerst een weegmodel worden opgesteld. Omdat de PHGSG een nieuw onderzoek is en er dus nog geen eerdere edities zijn om het weegmodel op te baseren, is ervoor gekozen het weegmodel van PHGSG 2020 te baseren op het weegmodel van de Persoonlijke Veiligheid (PV) [1]. Dit omdat het onderzoek PV qua thema (over de aard en omvang van slachtoffer- en daderschap van huiselijk geweld in NL) in de richting komt van het thema van de PHGSG (over de prevalentie van huiselijk geweld en seksueel geweld in NL). Gezien de onderzoeken niet precies gelijk zijn, is er een analyse gedaan van dit 'basisweegmodel'. Naar aanleiding daarvan zijn termen toegevoegd/verwijderd/uitgekruist.

Het weegmodel dat als basis is gebruikt voor de PHGSG is als volgt:

(*Geslacht* (2) x *Leeftijd* (6))

+ *Burgerlijke Staat* (4)

+ *Landsdeel* (4)

+ *Etnische groep* (3)

+ *Inkomen* (5).

De cijfers tussen haakjes geven het aantal categorieën per variabele weer welke terug zijn te vinden in Tabel 4. Aan de hand van dit model is gekeken naar andere hulpvariabelen die in aanmerking komen voor het weegmodel. Dit is gedaan door te kijken naar de totaalschattingen van deze hulpvariabelen met het hierboven genoemde weegmodel. Als deze totaalschattingen erg afwijken van de echte populatietotalen kan worden besloten om de hulpvariabelen op te nemen in het weegmodel. Dit corrigeert dan voor de afwijking van die hulpvariabele. Met een Chi-kwadraat toets kan dit worden bepaald. Zie hiervoor ook **Fout! Verwijzingsbron niet gevonden.** in de bijlage II, deze laat de Chi-kwadraat analyse voor het basismodel zien. Daarnaast wordt er ook gelet op wat het opnemen van deze variabelen doet met de minimale celvulling, de verdeling van de correctiegewichten, het design effect en schattingen van de belangrijkste doelvariabelen. Het opnemen van extra hulpvariabelen in het weegmodel leidt namelijk tot grotere spreiding van de gewichten en een groter designeffect. Kleine toenames hierin zijn acceptabel wanneer het een verkleining van de marges oplevert door sterke correlatie tussen doel- en hulp variabele. Het balanceren van deze kenmerken is wat het uiteindelijke weegmodel bepaalt.

Voor de PHGSG is naast de variabelen die in het voorgenoemde model zitten ook gekeken naar de hulpvariabelen *Provincie* (12), *Type huishouden* (7) en *Stedelijkheidsgraad* (5). Door de schattingen voor deze variabelen te toetsen aan de Chi-kwadraat toets is gebleken dat *Provincie* en *Type huishouden* goede kandidaten zijn om opgenomen te worden in het weegmodel. Gezien *Landsdeel* een indikking is van *Provincie* is besloten om deze te vervangen. Met *Provincie* in het weegmodel zijn de schattingen voor *Landsdeel* automatisch gewaarborgd en kloppend. Het weegmodel wordt nu als volgt:

(*Geslacht* (2) x *Leeftijd* (6))

+ *Burgerlijke Staat* (4)

+ *Provincie* (12)

+ *Etnische groep* (3)

+ *Inkomen* (5)

+ *Type huishouden* (7)

Gezien de celvulling van dit weegmodel aanzienlijk is, en de spreiding van de gewichten ook nog ruimte toelaat, is er gekeken of er nog verdere verbeteringen konden worden aangebracht. Vanwege het onderwerp van de PHGSG lijkt het een goed idee om termen uit te

kruisen met *Geslacht*. Het is aannemelijk dat *Geslacht* verklarend is voor een deel van de antwoorden in dit onderzoek, en daarnaast kan de statistische afdeling dan ook nog publiceren voor elke hulpvariabele uitgekruist met *Geslacht*. Behalve de variabele *Inkomen*, deze is gekruist met zowel *Geslacht* als *Leeftijd*, gezien het vermoeden bestaat dat deze termen veel samenhang hebben met de doelvariabelen. Dit leidde tot een weegmodel waarbij de celvulling voor de term (*Geslacht x Leeftijd x Inkomen*) te klein was. Daarom is er met een Chi-kwadraat toets gekeken naar de termen (*Geslacht x Inkomen*) en (*Leeftijd x Inkomen*) om te kijken of deze dan i.p.v. (*Geslacht x Leeftijd x Inkomen*) opgenomen moesten worden. Hieruit kwam naar voren kwam dat (*Leeftijd x Inkomen*) dusdanig scheef zat dat deze opgenomen werd in het weegmodel ter vervanging van (*Geslacht x Leeftijd x Inkomen*). Daarmee komt het uiteindelijke weegmodel er als volgt uit te zien:

PHGSG 2020
(Geslacht (2) x Leeftijd (6))
+ (Geslacht (2) x Burgerlijke Staat (4))
+ (Geslacht (2) x Provincie (12))
+ (Geslacht (2) x Etnische groep (3))
+ (Geslacht (2) x Type huishouden (7))
+ (Leeftijd (6) x Inkomen (5))

Tabel 1 Het weegmodel voor PHGSG 2020. Cijfers tussen haakjes geven het aantal categorieën voor de variabele weer

Verder zullen hieronder de resultaten van de analyse voor het gekozen model besproken worden.

8.4 Resultaten van de weging

Verdeling en design effect van de correctie gewichten per model

Voor het gekozen weegmodel kunnen de correctiegewichten $c_i^{(2)}$ worden uitgerekend. Dit gebeurt voor elk record in de respons en wordt berekend met de lineaire regressie schatter. Deze schatter berekent het tweede correctiegewicht zó, dat de met het product van het startgewicht en het tweede correctiegewicht gewogen som van iedere term uit het weegmodel gelijk is aan het populatietotaal van die term. In het bijzonder zorgt de lineaire regressie schatter ervoor dat de correctiegewichten zo bepaald worden dat de Euclidische afstand tussen start- en eindgewichten minimaal is. Voor het weegmodel uit Tabel 1 geldt het volgende:

	Startgewichten	2 ^e correctie gewichten	Eindgewichten
Negatief #	0	0	0
Minimum	467,9	0,283	132,5
1e kwart	467,9	0,697	326,1
Mediaan	467,9	0,943	441,0
Gemiddelde	467,9	1,000	467,9
3e kwart	467,9	1,235	578,0
Maximum	467,9	2,897	1355,7
Standaard deviatie	0	0,421	197,1
Design-effect	1	1,177	1,177
Minimale celvulling	62	62	62

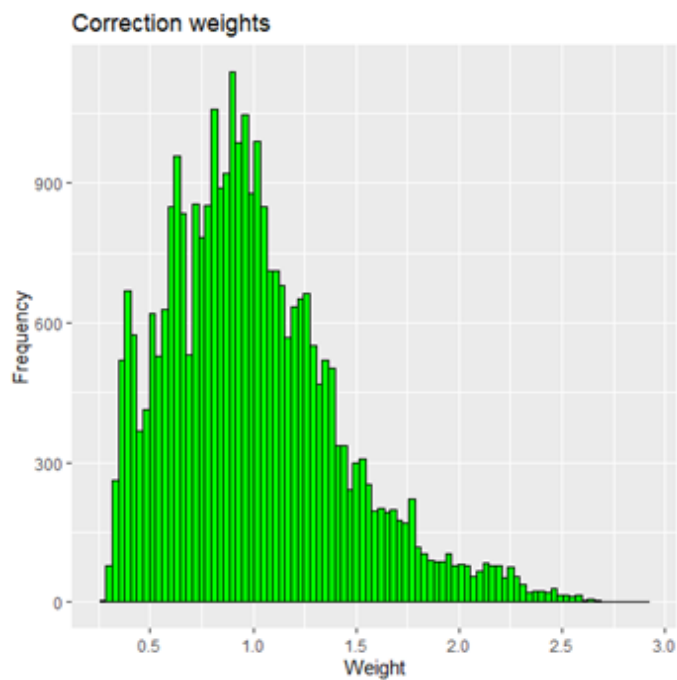
Tabel 2 Verdeling en design effect van start-, correctie- en eindgewichten

Negatieve eindgewichten, hoe klein ook, bemoeilijken het werken met gewogen responsgegevens. Om dit te voorkomen worden de gewichten dan begrensd.

	Startgewichten	2 ^e correctie gewichten	Eindgewichten
Negatief #	0	0	0
Minimum	467,9	0,283	132,5
1e kwart	467,9	0,697	326,1
Mediaan	467,9	0,943	441,0
Gemiddelde	467,9	1,000	467,9
3e kwart	467,9	1,235	578,0
Maximum	467,9	2,897	1355,7
Standaard deviatie	0	0,421	197,1
Design-effect	1	1,177	1,177
Minimale celvulling	62	62	62

Tabel 2 met kenmerken van de start-, correctie- en eindgewichten laat echter zien dat het gekozen weegmodel niet tot negatieve gewichten leidt en geen begrenzing nodig is. Verder is het goed om te zien dat de eindgewichten allemaal boven de 1 liggen. Bij eindgewichten kleiner dan 1 telt de respons niet eens helemaal voor de respondent zelf mee, wat niet intuïtief is. De minimale celvulling voor het weegmodel uit Tabel 1 is 62 responsen in de categorie 'Vrouwelijk x Overige' van de variabele *Leeftijd x Type huishouden*. Dit is genoeg massa om met dit weegmodel door te gaan. Het minimale tweede correctiegewicht is 0,283 en het maximale 2,897.

Het design-effect is een maat voor de maximale toename van de marge op doelvariabelen van het onderzoek door de weging. Die is gedefinieerd als $1 + V_w^2$, waarbij V_w de variatiecoëfficiënt van de eindgewichten is. Als een doelvariabele en de achtergrondkenmerken van het weegmodel correleren neemt de marge weer af. Het is daarom belangrijk de spreiding van de eindgewichten beperkt te houden en te zoeken naar een weegmodel dat (sterk) correleert met de doelvariabelen. Voor het gekozen weegmodel is het design-effect 1,177. Die waarde ligt lager dan bij de laatste editie van PV (1,277) [1] maar hoger dan bij het onderzoek Cybercrime (1,084) [2] dat een vergelijkbare responsomvang had. De waarde is dus aannemelijk, in vergelijking met andere door het CBS uitgevoerde wegingen. De spreiding in eindgewichten wordt uitsluitend veroorzaakt door de correctie gewichten, gezien de startgewichten overal gelijk zijn. De spreiding van de tweede correctie gewichten is te zien in Figuur 1.



Figuur 1 Spreiding van de tweede correctiegewichten PHGSG 2020

Waardes van de doelvariabelen per model

Fout! Verwijzingsbron niet gevonden.1 in hoofdstuk 7 laat voor het gekozen weegmodel de opgehoogde waarden en verdelingen in de populatie van de belangrijkste doelvariabelen zien. Dit is bij de analyse van alle weegmodellen ook gedaan om te kijken of de waarden van deze variabelen erg verschillen per weegmodel. Hieruit kan worden afgeleid wat de impact van ieder model op deze variabelen is. Uit analyse van alle voorgestelde weegmodellen blijkt dat deze verschillen in opgehoogde doelvariabelen tussen de weegmodellen zeer klein zijn en dit dus weinig invloed heeft bij de keuze van het weegmodel.

9. Opgeleverde Producten

10.1 Microbestanden

Het volgende eindbestand is geleverd:

Naam bestand	Datum	Records	Variabelen
PHSG_2020_respons.sav	31-08-2020	30459	1778

10.2 Statistische beveiliging

De microbestanden voor de verschillende afnemers worden vaak nog statistisch beveiligd. De wijze van beveiliging kan verschillen.

Onder statistische beveiliging verstaan we hier het voorkómen dat er inhoudelijke conclusies over herkenbare eenheden kunnen worden getrokken op basis van gepubliceerd of anderszins beschikbaar gesteld CBS-materiaal. Uit de statistische publicaties van het CBS (StatLine-tabellen, web-artikelen, persberichten, wetenschappelijke artikelen) mogen zulke conclusies niet getrokken kunnen worden. Maar ook als het CBS microdata beschikbaar stelt voor wetenschappelijke analyse, moet deze grondregel van de statistiek overeind blijven.

Er is nog geen verdere statistische beveiliging toegepast. Dat zal gebeuren op het RA-bestand.

Externe personen kunnen op de volgende manieren toegang krijgen tot 2 verschillende databestanden.

- 1) Het RA-bestand. Dat is vergelijkbaar met het microbestand, en staat bij CvB. Daar kunnen instellingen met een instellingsmachtiging via RA (tegen betaling) op werken.
- 2) Het DANS-bestand. Dat is een ingedikt bestand (bv leeftijd in klassen, regio's ingedikt enz.). Dat staat bij DANS en hier kunnen instellingen en studenten onder bepaalde voorwaarden toegang toe krijgen.

10. Referenties

- [1] RapportWegingPersoonlijkeVeiligheid2017.docx, 2017
[2] CYCRI18 wege rapport v1.0.pdf Arthur Kierkels, 28 februari 2019

11. Bijlagen

Bijlage I: Aanschrijfbrieven en folder



Centraal Bureau voor de Statistiek
Postbus 24500 | 2490 HA Den Haag
Postbus 4481 | 6401 CZ Heerlen
www.cbs.nl

16 jaar
PHSG

1

correspondentienr. <correspondentienummer>
onderwerp CBS-Onderzoek
datum 15 augustus 2019

Beste <voorletters> <Naam>,

Veiligheid is belangrijk voor iedereen. Toch komen we weleens in aanraking met onveilige situaties. Hoe is dat bij jou? Maak jij persoonlijk weleens onveilige situaties mee of zelfs geweld? Bijvoorbeeld op straat, thuis of online?

Het Centraal Bureau voor de Statistiek (CBS) onderzoekt hoe vaak mensen in aanraking komen met onveilige persoonlijke situaties. Dit doen we samen met het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC). Deze gegevens zijn belangrijk voor preventie en hulpverlening.

Voor dit onderzoek vraagt het CBS een beperkt aantal personen om een vragenlijst in te vullen. Jij bent daar één van. Het is daarom belangrijk dat je aan dit onderzoek meedoet. Ook als je nog nooit iets vervelends hebt meegemaakt willen we dat graag weten. Je helpt ons daar enorm mee.

Hoe doe je mee?

Je kunt de vragenlijst invullen via internet.

Website: <https://antwoord.cbs.nl>
Gebruikersnaam: <Gebruikersnaam>
Wachtwoord: <Wachtwoord>

Je gegevens zijn veilig

Bij al onze onderzoeken zijn je gegevens veilig. Je gegevens worden anoniem verwerkt. Zo worden je antwoorden gescheiden van je naam en adres. Onderaan deze brief en in de folder lees je meer hierover.

Heb je vragen?

Bel ons gerust op (045) 570 64 00. Wij zijn bereikbaar van maandag tot en met vrijdag tussen 9.00 en 17.00 uur. Je kunt ook naar contactcenter@cbs.nl mailen. Of kijk op www.cbs.nl/veilig voor meer informatie.

U doet ons een groot plezier als u één van de komende dagen de vragenlijst invult.

Ik dank u alvast hartelijk voor uw tijd en medewerking.

Met vriendelijke groet,



Mevrouw dr. A. Boeijen
Hoofddirecteur Dataverzameling

Uw gegevens zijn bij al onze onderzoeken veilig. Dit is een verplichting van het CBS die in de wet is bepaald. Het CBS heeft verschillende maatregelen getroffen om uw gegevens te beschermen. Zo zorgt de wet ervoor dat uw gegevens alleen voor statistische doelen worden gebruikt. Geen enkele instelling kan toegang eisen tot de gegevens die het CBS verzamelt. In de statistische informatie van het CBS zijn persoonlijke gegevens nooit te herkennen.

Het CBS verzamelt zelf gegevens maar krijgt ook veel bestanden van andere instellingen. Hierin staan bijvoorbeeld gegevens over werk en scholing. Die informatie voegen we samen. Het CBS en WODC maken statistieken over de Nederlandse samenleving met die informatie. Zo werken we zo zuinig mogelijk.



Centraal Bureau voor de Statistiek
Postbus 24500 | 2490 HA Den Haag
Postbus 4481 | 6401 CZ Heerlen
www.cbs.nl

16 jaar
PHSG
Rappel 1

1

correspondentienr. <correspondentienummer>
onderwerp CBS-Onderzoek
datum 15 augustus 2019

Beste <voorletters> <Naam>,

Drie weken geleden heb ik je gevraagd om mee te doen aan een onderzoek van het Centraal Bureau voor de Statistiek (CBS) en het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC). Het onderzoek gaat over veiligheid. Maak jij persoonlijk weleens onveilige situaties mee of zelfs geweld? Bijvoorbeeld op straat, thuis of online?

Veel jongeren hebben al gereageerd. Jouw vragenlijst hebben wij helaas nog niet ontvangen. Ook als je nog nooit iets vervelends hebt meegemaakt willen we dat graag weten. Zo help je ons een beter beeld te krijgen van de veiligheid in Nederland.
Heb je de vragenlijst al ingevuld? Hartelijk dank! Je hoeft niet meer te reageren.

Hoe doe je mee?

Je kunt de vragenlijst invullen via internet.

Website:	https://antwoord.cbs.nl
Gebruikersnaam:	<Gebruikersnaam>
Wachtwoord:	<Wachtwoord>

Je gegevens zijn veilig

Bij al onze onderzoeken zijn je gegevens veilig. Je gegevens worden anoniem verwerkt. Zo worden je antwoorden gescheiden van je naam en adres. Onderaan deze brief en in de folder lees je meer hierover.

Heb je vragen?

Bel ons gerust op (045) 570 64 00. Wij zijn bereikbaar van maandag tot en met vrijdag tussen 9.00 en 17.00 uur. Je kunt ook naar contactcenter@cbs.nl mailen. Of kijk op www.cbs.nl/veilig voor meer informatie.



Centraal Bureau voor de Statistiek
Postbus 24500 | 2490 HA Den Haag
Postbus 4481 | 6401 CZ Heerlen
www.cbs.nl

16 jaar
PHSG
Rappel 2

3

correspondentiennr. <correspondentienummer>
onderwerp CBS-Onderzoek
datum 15 augustus 2019

Beste <voorletters> <Naam>,

Zes weken geleden heb ik je gevraagd om mee te doen aan een onderzoek van het Centraal Bureau voor de Statistiek (CBS) en het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC). Het onderzoek gaat over veiligheid. Maak jij persoonlijk weleens onveilige situaties mee of zelfs geweld? Bijvoorbeeld op straat, thuis of online?

Veel jongeren hebben al gereageerd. Jouw vragenlijst hebben wij helaas nog niet ontvangen. Slechts een klein aantal jongeren is gevraagd om mee te doen met dit onderzoek. Daarom is het extra belangrijk dat je meedoet. Ook als je nog nooit iets vreselijks hebt meegemaakt willen we dat graag weten. Pas dan weten we zeker dat de resultaten zo goed mogelijk zijn.

Je kunt de vragenlijst nog tot en met **1 september 2020** invullen. Heb je de vragenlijst al ingevuld? Hartelijk dank! Je hoeft niet meer te reageren.

Hoe doe je mee?

Je kunt de vragenlijst invullen via internet.

Website:	https://antwoord.cbs.nl
Gebruikersnaam:	<Gebruikersnaam>
Wachtwoord:	<Wachtwoord>

Je gegevens zijn veilig

Bij al onze onderzoeken zijn je gegevens veilig. Je gegevens worden anoniem verwerkt. Zo worden je antwoorden gescheiden van je naam en adres. Onderaan deze brief en in de folder lees je meer hierover.

Heb je vragen?

Bel ons gerust op (045) 570 64 00. Wij zijn bereikbaar van maandag tot en met vrijdag tussen 9.00 en 17.00 uur. Je kunt ook naar contactcenter@cbs.nl mailen. Of kijk op www.cbs.nl/veilig voor meer informatie.

U doet ons een groot plezier als u één van de komende dagen de vragenlijst invult.

Ik dank u alvast hartelijk voor uw tijd en medewerking.

Met vriendelijke groet,



Mevrouw dr. A. Boeijen
Hoofddirecteur Dataverzameling

Uw gegevens zijn bij al onze onderzoeken veilig. Dit is een verplichting van het CBS die in de wet is bepaald. Het CBS heeft verschillende maatregelen getroffen om uw gegevens te beschermen. Zo zorgt de wet ervoor dat uw gegevens alleen voor statistische doelen worden gebruikt. Geen enkele instelling kan toegang eisen tot de gegevens die het CBS verzamelt. In de statistische informatie van het CBS zijn persoonlijke gegevens nooit te herkennen.

Het CBS verzamelt zelf gegevens maar krijgt ook veel bestanden van andere instellingen. Hierin staan bijvoorbeeld gegevens over werk en scholing. Die informatie voegen we samen. Het CBS en WODC maken statistieken over de Nederlandse samenleving met die informatie. Zo werken we zo zuinig mogelijk.

Het CBS gebruikt uw informatie uitsluitend voor statistische doeleinden. De gegevens worden nooit aan derden verstrekt. Uw anonimiteit is strikt gewaarborgd.

Voor wat er feitelijk gebeurt



Niemand ziet uw gegevens

Bij het Centraal Bureau voor de Statistiek (CBS) is uw informatie veilig. Daar kunt u zeker van zijn. Dat is vastgelegd in een aantal wetten: de wet Algemene Verordening Gegevensbescherming (AVG), de CBS-wet en de Europese Statistiek wet. In deze wetten staat dat het CBS verplicht is om te zorgen dat de informatie in veilige handen is. Ook onze computersystemen moeten goed beveiligd zijn. Wat houdt dit in? Lees hieronder meer daarover.

Hoe komen we bij u terecht?

Uw naam en adres zijn gekozen uit de Basisregistratie Personen (BRP). Dit bevolkingsregister is een groot bestand van de gemeenten. Hierin staan de adressen en namen van alle Nederlanders. In het bevolkingsregister staat bijvoorbeeld ook uw geboortedatum. Uit dit grote bestand worden willekeurig een aantal personen gekozen. U bent dus toevallig uitgekozen om de vragen van het CBS te beantwoorden.

Wat gebeurt er met uw persoonlijke gegevens en antwoorden?

Al uw gegevens komen in een beveiligde omgeving terecht. Uw antwoorden en persoonlijke gegevens worden van elkaar gescheiden. We weten dus nooit wie welk antwoord heeft gegeven.

 Meer informatie of video zien? www.cbs.nl/privacy

Uw gegevens zijn veilig

Hoe het CBS omgaat met uw privacy



Hoe het CBS omgaat met uw privacy



Bijlage II: Weging

Chi-kwadraat toetsen voor ongebruikte hulpvariabelen in het basismodel

Per hulpvariabele wordt bekeken of deze met het gekozen weegmodel (te) scheef in de opgehoogde populatie zit. Hiervoor wordt enerzijds gekeken naar de verwachte verdeling van een enkelvoudig aselechte steekproef uit de doelpopulatie (zie **Fout! Verwijzingsbron niet gevonden.** kolom 'Verwachte verdeling werkelijke populatie') en anderzijds naar een enkelvoudig aselechte steekproef uit de opgehoogde populatie (zie **Fout! Verwijzingsbron niet gevonden.** kolom 'Verwachte verdeling na ophoging'). Beide steekproeven hebben een zelfde grootte als de responsomvang. De afwijking tussen de verdelingen is een hulpmiddel bij de beslissing om een hulpvariabele al dan niet in het weegmodel op te nemen, en wordt berekend met behulp van de Chi-kwadraat toets. Hierbij is een significantieniveau van 0,05 gehanteerd. Het voelt wellicht raar om niet te kijken naar wat de daadwerkelijke verdeling in de respons is, maar die wordt meegenomen in de verdeling die na wegen wordt verkregen. Het weegmodel wordt immers losgelaten op de respons waarmee de initiële verdeling in de respons al gedeeltelijk gecorrigeerd wordt.

Fout! Verwijzingsbron niet gevonden. laat zien dat bij het initieel gebruikte weegmodel, voor de variabelen *type huishouden* en *provincie* de overschrijdingskansen onder het significantieniveau van 0,05 liggen. Dat betekent dat, wanneer men een enkelvoudig aselechte steekproef uit de populatie trekt, de kans kleiner dan 5% is dat een steekproef een verwachte verdeling heeft zoals de daadwerkelijke verdeling. Daarmee wordt aangetoond dat deze variabelen naar verwachting dusdanig scheef in respons voorkomen, dat het goede kandidaten zijn om opgenomen te worden in het weegmodel. Daarmee wordt dan gecorrigeerd voor de scheefheid. Deze Chi-kwadraat toets is herhaald voor alle geanalyseerde weegmodellen, voor zowel losse variabelen als uitkruisingen. Dit leidde uiteindelijk tot het definitieve weegmodel

Type Huishouden	geschat	verdeling	verwacht in resp o.b.v. schatting		TYPHH	totaal	verdeling	Echte verwachting response	
1	3022292	21,2%	6459		1	3097459	21,7%	6620	3,99521931
2	1396719	9,8%	2985		2	1193271	8,4%	2550	63,33159405
3	3603464	25,3%	7701		3	3363369	23,6%	7188	34,18760321
4	1025179	7,2%	2191		4	1066430	7,5%	2279	3,54725060
5	4202751	29,5%	8982		5	4372712	30,7%	9345	14,68884019
6	915002	6,4%	1955		6	1054294	7,4%	2253	45,31614060
7	87159	0,6%	186		7	105030	0,7%	224	7,83083515
									172,89748311
									1,09E-34
Provincie	geschat				PROVINCIE	totaal			
20	503699	3,5%	1076		20	490085	3,4%	1047	0,78636290
21	494911	3,5%	1058		21	529562	3,7%	1132	5,18474079
22	425370	3,0%	909		22	404333	2,8%	864	2,22342952
23	962064	6,8%	2056		23	941279	6,6%	2012	0,95966290
24	301873	2,1%	645		24	333057	2,3%	712	6,88433481
25	1713346	12,0%	3662		25	1702947	11,9%	3639	0,13488408
26	1171512	8,2%	2504		26	1089939	7,6%	2329	12,13860385
27	2342659	16,4%	5006		27	2367482	16,6%	5060	0,56211098
28	2958824	20,8%	6323		28	3017601	21,2%	6449	2,49527439
29	318398	2,2%	680		29	316371	2,2%	676	0,02757781
30	2149437	15,1%	4594		30	2115491	14,8%	4521	1,14571106
31	910472	6,4%	1946		31	944418	6,6%	2018	2,70478799
									35,24748108
									0,00022555
Stedelijkheidsgraad	geschat				STEDGEM	totaal			
1	3569666	25,0%	7629		1	3598598	25,2%	7691	0,50113178
2	4356192	30,6%	9310		2	4330928	30,4%	9256	0,31312650
3	2210743	15,5%	4725		3	2186740	15,3%	4673	0,55694897
4	3028685	21,3%	6473		4	3037934	21,3%	6492	0,06036123
5	1087279	7,6%	2324		5	1098365	7,7%	2347	0,24156353
									1,67313199
									0,79559225

voor PHGSG.

Tabel 3 Chi-kwadraat analyse op (initieel) ongebruikte hulpvariabelen

Tabel 4 Variabelen die in de weging kunnen worden meegenomen

Variabele	Omschrijving
Geslacht	Geslacht van de persoon: man of vrouw
Leeftijd	Leeftijd van de persoon op 1 maart 2020, ingedeeld in 6 categorieën: 16 – 23 jaar, 24 – 33 jaar, 34 – 48 jaar, 49 – 63 jaar, 64 - 73 jaar, 74+.
Burgerlijke staat	Burgerlijke staat van de persoon op 1 maart 2020 ingedeeld in de categorieën gehuwd of partnerschap, gescheiden, verweduwd en ongehuwd.
Landsdeel	Regio waar de persoon op 1 maart 2020 staat ingeschreven, ingedeeld in de categorieën Noord Nederland (Groningen, Friesland, Drenthe), Oost Nederland (Overijssel, Flevoland, Gelderland), West Nederland (Utrecht, Noord-Holland, Zuid-Holland, Zeeland), Zuid Nederland (Noord-Brabant, Limburg)
Provincie	Provincie waar de persoon op 1 maart 2020 staat ingeschreven.
Etnische groep	Herkomst van de persoon, ingedeeld in de categorieën Nederlandse achtergrond, migratie-achtergrond niet-westers of migratie-achtergrond westers.
Type huishouden	Type huishouden waar de persoon lid van was op 1 maart 2020 ingedeeld in de categorieën: 1. Eenpersoonshuishouden, 2. Gehuwd paar met kinderen, 3. Ongehuwd paar met kinderen, 4. Gehuwd paar zonder kinderen, 5. Ongehuwd paar zonder kinderen, 6. Eenouderhuishouden, 7. overig huishouden.
Inkomen	Persoonlijk inkomen van de persoon in 2020 in percentielen, ingedeeld in de categorieën: Lager dan percentiel 20, Percentiel 20 tot 40, Percentiel 40 tot 60, Percentiel 60 tot 80, Percentiel 80 of hoger.

Tabel 5 De belangrijkste weegmodel die geanalyseerd zijn voor PHGSG

Weegmodel o.b.v. PV	Weegmodel + Provincie en Type hh	PHGSG 2020
(Geslacht (2) x Leeftijd (6)) + Burgerlijke Staat (4) + Landsdeel (4) + Etnische groep (3) + Inkomen (5)	(Geslacht (2) x Leeftijd (6)) + Burgerlijke Staat (4) + Provincie (12) + Etnische groep (3) + Inkomen (5) + Type huishouden (7)	(Geslacht (2) x Leeftijd (6)) + (Geslacht (2) x Burgerlijke Staat (4)) + (Geslacht (2) x Provincie (12)) + (Geslacht (2) x Etnische groep (3)) + (Geslacht (2) x Type huishouden (7)) + (Leeftijd (6) x Inkomen (5))

Randtotalen van de weging

De doelpopulatie voor PHGSG 2020 bestaat uit in Nederland woonachtige personen die 16 jaar of ouder waren op het moment van de peildatum (start van het veldwerk, 1 maart 2020) en deel uitmaken van particuliere huishoudens. De institutionele bevolking, dat zijn personen in inrichtingen, instellingen of tehuizen, behoort niet tot de doelpopulatie en is niet benaderd. Voor het wegen van de responsgegevens wordt deze doelpopulatie benaderd door een lijst te nemen van personen die op 1 maart 2020 in de Basisregistratie van Personen (BRP) staan ingeschreven en op dat moment zestien jaar of ouder zijn. Deze lijst van personen telt 14.252.565 personen.

Aan deze lijst van personen, het zogenaamde weegkader, zijn de waarden van de variabelen uit **Fout! Verwijzingsbron niet gevonden.** gekoppeld. Niet voor iedere persoon in het weegkader en niet voor iedere respondent is de waarde van iedere variabele bekend. Is de waarde van een variabele onbekend, dan is aan de personen met een onbekende waarde de waarde toegekend die het meest voorkomt in het weegkader, de modus. Tellen van alle personen in de lijst als functie van de waarden van de variabelen levert de randtotalen voor de weging. **Fout!**

Verwijzingsbron niet gevonden. toont de in de weging gebruikte randtotalen voor de individuele variabelen. De randtotalen voor de uitkruisingen zijn waar nodig ook berekend maar worden hier niet getoond.

Tabel 6 Randtotalen PHGSG 2020

Geslacht	N	Provincie	N
Man	7.044.390	Groningen	490.085
Vrouw	7.208.175	Friesland	529.562
		Drenthe	404.333
		Overijssel	941.279
		Flevoland	333.057
		Gelderland	1.702.947
		Utrecht	1.089.939
		Noord-Holland	2.367.482
		Zuid-Holland	3.017.601
		Zeeland	316.371
		Noord-Brabant	2.115.491
		Limburg	944.418
Leeftijd	N	Type huishouden	N
16 t/m 23	1.712.059	Eenpersoons huishouden	3.097.459
24 t/m 33	2.218.269	Ongehuwd paar zonder kinderen	1.193.271
34 t/m 48	3.177.071	Gehuwd paar zonder kinderen	3.363.369
49 t/m 63	3.666.142	Ongehuwd paar met kinderen	1.066.430
64 t/m 73	1.983.070	Gehuwd paar met kinderen	4.372.712
74+	1.495.954	Ouder in eenouderhuishouden	1.054.294
		In overig huishouden	105.030
Burgerlijke staat	N	Inkomen (kwintiel)	N
Gehuwd / Partnerschap	6.659.679	0 - 20	2.020.374
Gescheiden	1.337.015		
Verweduwd	791.208		
Ongehuwd	5.464.663		
Landsdeel	N		
Noord Nederland	1.423.980		
Oost Nederland	2.977.283		
West Nederland	6.791.393		
Zuid Nederland	3.059.909		
Etnische groep	N		

Nederlander	10.882.338	20 - 40	2.384.448
Westerse Allochtoon	1.552.654	40 - 60	2.812.418
Niet-Westerse Allochtoon	1.817.573	60 - 80	3.160.332
		80 - 100	3.874.993