



Centraal Bureau
voor de Statistiek

Rapport

Culturele veranderingen en SCP Leefsituatie Index (CV/SLI) 2019- 2022

Onderzoeksdocumentatie n kwaliteitsanalyse

Moniek Coumans
Josien Knops
Marko Budé
Ingeborgh Hofman-Thomassen

Projectnummer 304447/304783
Divisie Dataverzameling (DVZ)/Sector Arbeid inkomen en Leefsituatie (SAL)
3 oktober 2022

kennisgeving De in dit rapport weergegeven opvattingen zijn die van de auteurs en komen niet noodzakelijk overeen met het beleid van het Centraal Bureau voor de Statistiek.

Inhoudsopgave

1.	Inleiding	5
1.1	Doel van dit document	5
2.	Doel en achtergrond geschiedenis van het onderzoek	6
2.1	Doel	6
2.2	Achtergrond van het onderzoek	6
3.	Dataverzameling	7
3.1	Steekproef	7
3.2	Vragenlijstontwikkeling	13
3.3	Opleiding en Instructie	15
3.4	Benaderingsstrategie	15
3.5	Veldwerk	17
4.	Verwerken (data processing).....	22
4.1	Inleiding	22
4.2	Invoer	22
4.3	Controleren en uniformeren waarneemdata	22
4.4	Typeren	23
4.5	Verrijken	23
4.6	Afleiden	24
4.7	Correcties: gaafmaken, imputeren	25
5.	Non-responsanalyse	26
5.1	Afbakenen statistische respons en nonresponsanalyse	26
5.2	Steekproef- en responsverantwoording	26
5.3	Selectiviteit van de respons naar achtergrondkenmerken	29
5.4	Weging	36
6.	Plausibiliteitsanalyse	37
7.	Weging	38
7.1	Wegen	38
8.	Conclusies kwaliteitsanalyse.....	40
9.	Opgeleverde Producten	41
9.1	Microbestanden	41
9.2	Statistische beveiliging	41
10.	Referenties	42

11. Bijlage I – Algemene beschrijving dataprocessing 43

1. Inleiding

1.1 Doel van dit document

In dit document wordt verslag gedaan van het onderzoek Culturele Veranderingen (CV) en de SCP Leefsituatie Index (SLI), editie 2019 – 2022. De voorbereiding, dataverzameling en verwerking van de gegevens werden 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.

In hoofdstuk 2 wordt het doel, achtergrond en geschiedenis gemeld. Daarna wordt in hoofdstuk 3 uitgebreid ingegaan op de dataverzameling. Hoofdstuk 4 gaat over de uitzonderingen in de verwerkingsstraat die voor dit onderzoek zijn toegepast. In de bijlage staat de standaard verwerkingsstraat gemeld.

Tevens worden in deze notitie wordt op enkele kwaliteitsaspecten van het onderzoek ingegaan. In hoofdstuk 5 komt allereerst de non-responsanalyse aan de orde. Deze bestaat uit een korte beschrijving van de getrokken steekproef, de verantwoording van de uitgezette steekproef, de in het veld geconstateerde kaderfouten, en de gerealiseerde responsaantallen. Daarnaast wordt nagegaan in hoeverre er sprake is van selectiviteit in de respons naar achtergrondkenmerken, onder eer met behulp van zogenoemde R-indicatoren.

Daarna wordt in hoofdstuk 6 naar de plausibiliteit gekeken, door de gegevens met eerdere edities van dit onderzoek te vergelijken. Tot slot volgen enkele conclusies die uit deze analyses getrokken kunnen worden (hoofdstuk 7).

2. Doel en achtergrond geschiedenis van het onderzoek

2.1 Doel

De onderzoeken Culturele Veranderingen in Nederland (CV) en SCP Leefsituatie Index (SLI) worden jaarlijks gecombineerd uitgevoerd in samenwerking met het Sociaal Cultureel Planbureau (SCP) en het Centraal Bureau voor de Statistiek (CBS).

Doel van CV is vaststellen welke opvattingen “de Nederlander” heeft over samenleving en cultuur en hoe deze in de loop van de tijd zijn veranderd. Doel van CV is het ontwikkelen van een indicator om de sociale leefsituatie van Nederland in kaart te brengen en het monitoren van de ontwikkeling van de sociale leefsituatie van de Nederlandse bevolking. De onderzoeken liggen aan de basis van veel onderzoek uit het werkprogramma van het SCP. Door het CBS zullen de data vooral worden gebruikt ten behoeve van het speerpunt Sociale Samenhang en Welzijn.

2.2 Achtergrond van het onderzoek

Deze onderzoeken werden tot de jaarwisseling 2016/17 tweejaarlijks in zes maanden gecombineerd uitgevoerd in samenwerking tussen het Sociaal Cultureel Planbureau (SCP) en het CBS. SCP heeft in 2017 aangegeven dat er een wens ligt om de Tweede Kamer in de jaarlijkse begrotingscyclus te voorzien van informatie rondom het bredere welvaartsbegrip. Het gaat dan om informatie die gepresenteerd wordt in de Sociale Staat van Nederland, waarvoor CV/SLI een van de kernonderzoeken is. Om die reden heeft SCP aangegeven het onderzoek doorlopend te maken, in de jaren 2017 en 2018. Voor het eerst zijn hierover cijfers gepubliceerd in september 2018. Ook in 2019 heeft het onderzoek doorlopend kunnen plaatsvinden. Vanaf 2020 is dit vanwege de COVID-19-pandemie niet meer het geval. Het veldwerk is n.a.v. de lockdowns in 2020 en 2021 tweemaal uitgesteld en liep door tot en met april 2022. De editie 2020-2022 is de laatste editie van dit onderzoek.

3. Dataverzameling

De veldwerkperioden van CV/SLI werd uitgevoerd in twee waarneemperioden. De eerste waarneemperiode van CV werd uitgevoerd van 2 januari 2019 tot en met 31 december 2019. Het veldwerk van SLI liep van 2 januari 2019 tot en met 31 januari 2020.

In de initiële planning liep de waarneming van 2 januari 2020 tot en met 31 december 2020. Het veldwerk van de schriftelijke vervolgvragenlijst SLI liep van 2 januari 2020 tot en met 31 januari 2021. Door de maatregelen vanwege COVID-19 is deze planning aangepast en liep het veldwerk in januari 2020 tot en met half maart 2020. Het veldwerk is half maart 2020 vanwege de eerste lockdown stopgezet. Het veldwerk is hervat begin december 2020 maar is in de tweede helft van december weer stopgezet door een tweede lockdown. Vervolgens is de veldwerkperiode weer hervat in september, november, december 2021 en in januari tot en met maart 2022. Het veldwerk van SLI liep nog door tot en met april 2022.

3.1 Steekproef

3.1.1 Doelpopulatie en steekproefkader

De doelpopulatie van CV bestond uit de in Nederland woonachtige personen van 16 jaar of ouder; die van SLI uit in Nederland woonachtige personen van 18 jaar of ouder. De institutionele bevolking werd niet geënquêteerd, en van steekproefpersonen die de Nederlandse taal niet of onvoldoende beheersten, werden geen interviews afgenomen.

Voor CV zijn steekproeven getrokken uit het steekproefkader van personen dat door CBS werd gebruikt voor onderzoeken met waarneming startend in 2019, 2020, 2021 en 2022. Deze kaders zijn afgeleid uit de Basisregistratie Personen (BRP).

3.1.2 Steekproefontwerp

De steekproeven voor CV zijn getrokken volgens een gestratificeerd tweetrapsontwerp. In de eerste trap zijn per provincie systematisch gemeenten geselecteerd met kansen evenredig aan hun inwoneraantallen; tevens werd voor elke geselecteerde gemeente het aantal te trekken personen berekend. Het in beginsel gewenste aantal te trekken personen per gemeente was voor iedere gemeente gelijk aan één; voor zelfselecterende gemeenten was dit aantal gelijk aan het product van de steekproeffractie en het betreffende inwoneraantal. In de tweede trap zijn enkelvoudig aselekt steekproeven getrokken van personen van 16 jaar of ouder op de eerste dag van de maand van waarneming in de geselecteerde gemeenten, met omvang per gemeente zoals vastgesteld in de eerste trap. Zo had iedere persoon in de doelpopulatie bij benadering dezelfde kans om in de steekproef terecht te komen; de steekproeven waren dus bij benadering zelfwegend.

Voor SLI zijn geen additionele steekproeven getrokken: respondenten bij CV die op het moment van enquêtering 18 jaar of ouder waren, zijn gevraagd de SLI-vragenlijst in te vullen.

De doelstelling, de steekproeftrekking, -bewerking en gerealiseerde responsen worden voor de waarneemperioden 2019 en 2020 tot en met 2022 apart beschreven in de onderstaande paragrafen.

Waarneemperiode januari 2019 – december 2019

Doelstelling

Doelstelling was het behalen van 1.000 SLI-responsen, en om deze respons in verwachting te behalen was een uitzet van 2.169 personen nodig. Om met grotere waarschijnlijkheid aan de responseis te voldoen, is de CV-uitzet bepaald op het kleinste veelvoud van twaalf waarvoor de linker grens van het 95%-betrouwbaarheidsinterval van de verwachte SLI-responsomvang groter was dan 1000; deze omvang bedroeg 2.280 personen, dit is een CV-uitzet van 190 per maand.

De steekproefomvang voor CV 2019 is vastgesteld onder de aannames dat 55,2% van de benaderde personen respondent voor CV, dat 91,5% van de CV-respondenten 18 jaar of ouder zijn en SLI-deelname toezeggen en dat 91,3% van de SLI-toezeggers daadwerkelijk respondeert. Het te hanteren SLI-responspercentage op de CV-uitzet bedroeg hiermee 46,1%.

Steekproeftrekking en -bewerking

Na steekproeftrekking zijn die personen uit de steekproef verwijderd wier adresgegevens onvolledig of onbruikbaar waren, en die woonden op een adres dat ofwel in de voorgaande twaalf maanden al in een CBS-steekproef voorkwam, ofwel behoorde tot die van de institutionele bevolking, ofwel buiten het capi-waarnemingsgebied van CBS lag. In verband met uitval bij die zogeheten screening worden doorgaans steekproeven getrokken met omvang van 1,13 maal de gewenste uitzet. Vanwege de kleine omvang van de maandelijks uit te zetten steekproeven is ditmaal een screeningsfactor van 1,2 gehanteerd (i.e. te trekken waren maandelijks 228 of 229 elementen). De na screening resterende steekproefpersonen zijn tot slot systematisch uitgedund tot een uitzet van 190. In tabel 3.1.3.1 is de screeningsuitval per steekproefportie gegeven.

Tabel 3.1.3.1: Gerealiseerde steekproef- en screeningsaantallen per steekproefportie

periode	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	Σ
getrokken steekproef	228	228	228	228	228	228	228	228	228	228	229	229	2738
screeningsuitval	10	10	19	17	12	14	13	13	16	21	12	16	173
als % van getrokken	4,4	4,4	8,3	7,5	5,3	6,1	5,7	5,7	7	9,2	5,3	7	6,3
w.v. in personenstprf. beschikbaar voor uitzet	9	7	11	11	10	9	13	10	8	12	11	10	121
steekproefreductie	28	28	19	21	26	24	25	25	22	17	27	23	285
capi-uitzet CV	190	190	190	190	190	190	190	190	190	190	190	190	2280

Voor de twaalf waarneemporties zijn bij elkaar 2.738 personen in een steekproef getrokken. Van hen zijn er 173, dit is 6,3%, tijdens het screeningsproces uitgevallen; de meesten omdat iemand op hun adres in de afgelopen twaalf maanden al voor een CBS-onderzoek benaderd was (121 personen, ofwel 4,4% van de getrokken steekproef). Bij de systematische uitdunningen tot de gewenste uit te zetten aantallen zijn in totaal 285 personen uit de steekproeven verwijderd. Zie tabel 3.1.3.2 voor een uitsplitsing van de redenen van screeningsuitval per maandportie.

Tabel 3.1.3.2 Resultaten screening per waarneemperiode, CV 2019

	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec	totaal
Getrokken	228	228	228	228	228	228	228	228	228	228	229	229	2.738
Niet beschikbaar voor uitzet, waarvan:	10	10	19	17	12	14	13	13	16	21	12	16	173
- in eerdere CBS steekproef	9	7	11	11	10	9	13	10	8	12	11	10	121
- institutionele bevolking						1							1
- met onbekende straatnaam/NAW-gegevens	1	3	8	6	2	3		3	8	9	1	5	49
- buiten waarneming SDV (CAPI)						1						1	2
Beschikbaar voor uitzet	218	218	209	211	216	214	215	215	212	207	217	213	2.565
Uitdunning	28	28	19	21	26	24	25	25	22	17	27	23	285
Uitgezet *), waarvan:	190	190	190	190	190	190	190	190	190	190	190	190	2.280
- met telefoonnummer, waarvan:	44	36	40	42	33	38	23	32	176	166	163	171	964
- vast nummer	40	34	40	42	33	37	23	32	136	120	106	108	751
- mobiel nummer	4	2	0	0	0	1	0	0	40	46	57	63	213
- zonder telefoonnummer	146	154	150	148	157	152	167	158	14	24	27	19	1.316

*)vanaf september 2019 wijziging in leveranciers van telefoonnummers

Waarneemperioden januari 2020 tot en met maart 2022

Doelstelling

Doelstelling was het behalen van 1.000 SLI-responsen, en om deze respons in verwachting te behalen was een uitzet van 2.175 personen nodig. Om met grotere waarschijnlijkheid aan de responseis te voldoen, is de CV-uitzet bepaald op het kleinste veelvoud van twaalf waarvoor de linker grens van het 95%-betrouwbaarheidsinterval van de verwachte SLI-responsomvang groter was dan 1000; deze omvang bedroeg 2.280 personen, i.e. een CV-uitzet van 190 per maand. Het steekproefontwerp is later gewijzigd n.a.v. de tegenvallende respons als gevolg van de COVID-19-pandemie. Vooral de respons voor CV bleek sinds het begin van de COVID-19-pandemie veel lager uit te vallen dan daarvoor. Uiteindelijk zijn er 3.490 personen benaderd voor CV.

De steekproefomvang voor CV is oorspronkelijk vastgesteld onder de aannames dat 55,2% van de benaderde personen respondeert voor CV, dat 92,6% van de CV-respondenten 18 jaar of ouder zijn én SLI-deelname toezeggen, en dat 91,3% van de SLI-toezeggers daadwerkelijk respondeert, waarvan 1,5% nog uitvalt i.v.m. slechte kwaliteit van de ingeleverde vragenlijst. Het te hanteren SLI-responspercentage op de CV-uitzet bedroeg hiermee 46,0%.

Steekproeftrekking en -bewerking

Na steekproeftrekking zijn die personen uit de steekproef verwijderd wier adresgegevens onvolledig of onbruikbaar waren, en die woonden op een adres dat ofwel in de voorgaande twaalf maanden al in een CBS-steekproef voorkwam, ofwel behoorde tot die van de institutionele bevolking, ofwel buiten het capi-waarnemingsgebied van CBS lag. In verband met uitval bij die zogeheten screening worden doorgaans steekproeven getrokken met omvang van 1,10 maal de gewenste uitzet. Vanwege de kleine omvang van de maandelijks uit te zetten steekproeven is in 2020 en 2021 net als in 2019 een screeningsfactor van 1,2 gehanteerd (i.e. te trekken waren maandelijks 228 of 229 elementen). De na screening resterende steekproefpersonen zijn tot slot systematisch uitgedund tot uitzetten genoemd

in de tabellen hieronder. In de tabellen 3.1.3.3 tot en met 3.1.3.5 is de screeningsuitval per steekproefportie gegeven. Merk op dat in 2022 de uitdunning hoger lag, omdat pas na de trekkingen bekend werd dat de uitzet lager werd dan daarvoor gedacht.

Tabel 3.1.3.3: Gerealiseerde steekproef- en screeningsaantallen per steekproefportie (2020)

periode	2001	2002	2003	2004	2012	Σ
getrokken steekproef	228	229	229	229	228	1143
screeningsuitval	16	8	16	14	12	66
als % van getrokken	7	3,5	7	6,1	5,3	5,8
w.v. in personenstkprf.	10	8	12	10	5	45
beschikbaar voor uitzet	212	221	213	215	216	1077
steekproefreductie	22	31	23	25	26	127
capi-uitzet CV	190	190	190	190	190	950

Tabel 3.1.3.4: Gerealiseerde steekproef- en screeningsaantallen per steekproefportie (2021)

periode	2109	2111	2112	Σ
getrokken steekproef	390	390	780	1560
screeningsuitval	20	12	24	56
als % van getrokken	5,1	3,1	3,1	3,6
w.v. in personenstkprf.	6	4	8	18
beschikbaar voor uitzet	370	378	756	1504
steekproefreductie	45	53	106	204
capi-uitzet CV	325	325	650	1300

Tabel 3.1.3.5: Gerealiseerde steekproef- en screeningsaantallen per steekproefportie (2022)

periode	2201	2202	2203	Σ
getrokken steekproef	908	528	316	1752
screeningsuitval	29	24	10	63
als % van getrokken	3,2	4,5	3,2	3,6
w.v. in personenstkprf.	15	6	1	22
beschikbaar voor uitzet	879	504	306	1689
steekproefreductie	233	130	86	449
capi-uitzet CV	646	374	220	1240

In de tabellen 3.1.3.6 tot en met 3.1.3.8 is de uitsplitsing weergegeven van de redenen van screeningsuitval per maandportie per jaar. In de maanden waarin een kruis is weergegeven in de tabel heeft er geen face-to-face waarneming plaatsgevonden.

Tabel 3.1.3.6 Resultaten screening per waarneemperiode, CV 2020

	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec	totaal
Getrokken	228	229	229	229	X	X	X	X	X	X	X	228	1.143
Niet beschikbaar voor uitzet, waarvan:	16	8	16	14								12	66
- in eerdere CBS steekproef	10	8	12	11								5	46
- institutionele bevolking	1												1
- met onbekende straatnaam/NAW-gegevens	4		4	2								6	16
- buiten waarneming SDV (CAPI)	1			1								1	3
Beschikbaar voor uitzet	212	221	213	215								216	1077
Uitdunning	22	31	23	25								26	127
Uitgezet ¹⁾ , waarvan:	190	190	190	190								190	950
- met telefoonnummer, waarvan:	167	160	161	163								169	820
- vast nummer	103	106	105	103								120	537
- mobiel nummer	64	54	56	60								49	283
- zonder telefoonnummer	23	30	29	27								21	130

Tabel 3.1.3.7 Resultaten screening per waarneemperiode, CV 2021

	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec	totaal
Getrokken	X	X	X	X	X	X	X	X	390	X	390	780	1.560
Niet beschikbaar voor uitzet, waarvan:									20		12	24	56
- in eerdere CBS steekproef									6		4	8	18
- institutionele bevolking											1		1
- met onbekende straatnaam/NAW-gegevens									12		6	13	31
- buiten waarneming SDV (CAPI)									2		1	3	6
Beschikbaar voor uitzet									370		378	756	1504
Uitdunning									45		53	106	204
Uitgezet ¹⁾ , waarvan:									325		325	650	1300
- met telefoonnummer, waarvan:									274		271	539	1084
- vast nummer									196		199	395	790
- mobiel nummer									78		72	144	294
- zonder telefoonnummer									51		54	111	216

Tabel 3.1.3.8 Resultaten screening per waarneemperiode, CV 2022

	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec	totaal
Getrokken	908	528	316	X	X	X	X	X	X	X	X	X	1.752
Niet beschikbaar voor uitzet, waarvan:	29	24	10										63
- in eerdere CBS steekproef	15	15	1										31
- institutionele bevolking	1		1										2
- met onbekende straatnaam/NAW-gegevens	12	8	8										28
- buiten waarneming SDV (CAPI)	1	1											2
Beschikbaar voor uitzet	879	504	306										1689
Uitdunning	233	130	86										449
Uitgezet ¹⁾ , waarvan:	646	374	220										1240
- met telefoonnummer, waarvan:	553	319	186										1058
- vast nummer	393	237	126										756
- mobiel nummer	160	82	60										302
- zonder telefoonnummer	93	55	34										182

3.1.3 Verhuizingsstrategie

Conform de nota Verhuizingen Strategie (Goris G, mei 2012) worden steekproef-personen die verhuisd zijn niet herbenaderd.

3.2 Vragenlijstontwikkeling

Voor de vragenlijst van CV 2019/2020 is de vragenlijst van CV 2017/2018 als uitgangspunt genomen. Ten opzichte van 2017/2018 zijn 2 hele vraagblokken verwijderd, te weten de blokken 'Nederlandse identiteit' (14 vragen) en 'Milieu' (15 vragen). Het vraagblok 'Rol van de Overheid' (10 vragen) is nieuw toegevoegd in de vragenlijst. Daarnaast zijn ook nog enkele nieuwe vragen toegevoegd in de blokken 'Politiek en Democratie' en 'Misdad' (11 vragen).

Ten behoeve van opleiding is in de vragenlijst in kleur de wijzigingen ten opzichte van het vorig jaar weergegeven. Daarnaast zijn een aantal instructiepunten aangeduid die relevant kunnen zijn voor de handleiding. Sinds vorig jaar is voor dit onderzoek een omnimode-schema opgesteld, hetgeen betekent dat de interviewers moeten werken volgens de interviewerregels die gelden voor alle interviewmodes (bijvoorbeeld zowel persoonlijk/face tot face of telefonisch).

In tegenstelling tot editie 2017/2018 was voor edities 2019 en 2020-2022 geen sprake meer van een split design, waarbij een bepaald deel van de respondenten een net iets andere vraagstelling krijgt voorgelegd dan het andere deel. Met andere woorden: in de edities 2019 en 2020-2022 dit jaar kregen alle respondenten exact dezelfde vragen.

Het uitgangspunt was dat de vragenlijstduur van CV 2019/2020-2022 korter werd dan die van CV 2017/2018 (werkelijke duur: circa 40,5). De geschatte gemiddelde vragenlijstduur voor 2019/2020 was met 39 minuten iets korter.

Overzicht van vraagblokken:

- Blok 1: Maatschappelijke positie
- Blok 2: Eigen leven, Welvaart en Sociale zekerheid
- Blok 3: Geloof
- Blok 4: Vertrouwen en samenleving
- Blok 5: Politiek en Democratie
- Blok 6: Rol van de Overheid
- Blok 7: Misdad
- Blok 8: Verlenen van hulp en Morele kwesties
- Blok 9: Positie van de vrouw
- Blok 10: Positie van allochtonen
- Blok 11: Werving SLI
- Blok 12: Ras

Kwaliteit van Leven (SLI), PAPI-vragenlijst

Voor de schriftelijke vragenlijst van 'SCP Onderzoek naar Kwaliteit van Leven' 2019/2020-2022 is net als bij CV, de vragenlijst van 2017/2018 als uitgangspunt genomen. Dit betreft wel een herontwerp, waarbij vooral veel vragen geschrapt zijn t.o.v. een jaar geleden (29 vragen geschrapt, 10 vragen toegevoegd).

Op de voorkant van de vragenlijst kon de interviewer net als voorgaande jaren de interviewdatum en het interviewnummer handmatig noteren. In 2019/2020-2022 werd daar nog een datumveld aan toegevoegd, namelijk de datum van het ophaalbezoek. Daarnaast plakten de interviewers een sticker met barcode op de voorkant waarin het unieke identificatienummer (WE_ID) verwerkt zat. Het geslacht en de geboortedatum werd door de respondent zelf ingevuld op de eerste pagina van de vragenlijst.

Op verzoek van SCP werd net als in de vorige editie een foto/plaatje op de voorkant van de vragenlijst gedrukt, om zo de aantrekkelijkheid voor het invullen van de vragenlijst te vergroten.

De koppeling tussen SLI en CV vond plaats door middel van een zogenaamd 11-nummer. Dit 6-cijferige nummer + volglletter werd op de voorkant van de SLI vragenlijst gedrukt. De interviewer toetste dit nummer in bij de wervingsvraag voor SLI in de CV-vragenlijst.

Er waren twee varianten van de papieren vragenlijst (variant A en B), die alleen op het laatste blok (bestaande uit 2 open vragen) van elkaar verschilden. Variant A bevat het blok "Een goede ouder" en Variant B bevat het blok "Rol van de Overheid". De interviewer bepaalde zelf wie welke variant krijgt, mits deze random werden verdeeld over respondenten.

Aangezien de vragenlijsten qua inhoud volledig met elkaar overeen komen met uitzondering van de laatste 2 open vragen zijn de papieren vragenlijsten vertoetst in één Cadi-invoermachine. Het is efficiënter om met één invoermachine te werken, zodat er maar één invoerbestand en één datamodel verwerkt hoeft te worden. Op de voorpagina van de schriftelijke vragenlijst stond een 'A' of 'B' (verwijzend naar de variant) en in de invoermachine is een variabele 'Variant' toegevoegd waar dit kenmerk in vertoetst kan worden en op basis waarvan de laatste vragen (die afwijkend zijn voor beide versies) geroute worden.

De geschatte gemiddelde face-to-face (capi) vragenlijstduur in 2019 bedroeg maximaal 39 minuten. Tabel 3.2.1 laat zien dat de gemiddelde gerealiseerde face-to-face (capi) vragenlijstduur in 2019 met 40 minuten, 1 minuut hoger lag dan verwacht.

Tabel 3.2.1 Vragenlijstduur in minuten CV, per veldwerkperiode en totaal (=gemiddelde)

	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912	totaal
Vragenlijstduur CAPI in minuten	39,7	40,1	38,9	39,7	39,2	38,5	40,6	40,9	41,2	40,3	39,6	41,2	40,0
n*	118	109	101	116	102	100	108	108	102	97	91	91	1243
Max. aantal minuten	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0

In de veldwerkperioden januari 2020 tot en met maart 2022 is door de Covid-19 maatregelen het veldwerk twee keer stilgelegd en weer hervat toen dit door de versoepeling van de Covid-19 maatregelen weer mogelijk was. Daardoor is over deze perioden de gemiddelde gerealiseerde vragenlijstduur berekend over 3 verschillende veldwerkperioden.

De tabellen 3.2.2 en 3.2.3 laten zien dat de gemiddelde gerealiseerde vragenlijstduur voor de veldwerkperiode januari 2020 tot en met april 2020 met 41,5 minuten, 1,5 minuut langer dan verwacht was en voor de periode december 2020 tot en met maart 2023 was de vragenlijstduur met 41,8 minuten, 1,8 minuut langer dan verwacht.

Tabel 3.2.2 Vragenlijstduur in minuten CV, per veldwerkperiode en totaal (=gemiddelde)

	202001	202002	202003	202004	Totaal
Vragenlijstduur CAPI in minuten	39,3	41,7	43,1	41,8	41,5
n*	95	96	45	70	306
Max. aantal minuten	39,0	39,0	39,0	39,0	39,0

Tabel 3.2.3 Vragenlijstduur in minuten CV, per veldwerkperiode en totaal (=gemiddelde)

	202012	202109	202111	202112	202201	202202	202203	totaal
Vragenlijstduur CAPI in minuten	43,0	40,7	44,9	41,3	41,9	40,5	40,4	41,8
n*	46	130	141	139	228	139	88	911
Max. aantal minuten	39,0	39,0	39,0	39,0	39,0	39,0	39,0	39,0

* n = het aantal respondenten waarop de gemiddelde vragenlijstduur is gebaseerd. Dit aantal is kleiner dan het aantal respondenten omdat gecorrigeerd is voor uitschieters.

3.3 Opleiding en Instructie

De opleiding en instructie voor CV/SLI is in 2019 gegeven aan de interviewers voor de jaren 2019, 2020, 2021 en 2022. Dit vond plaats door middel van e-learning, waarbij de bestaande e-learning module werd aangepast en opnieuw gebruikt. Interviewers konden via de leeromgeving waar de e-learning module staat een cursus over de organisatie van het onderzoek en de inhoud van de vragenlijst volgen. Daarnaast konden er vragen worden gesteld. In de e-learning module was een toets opgenomen die door interviewers gemaakt moest worden. Om ingezet te kunnen worden voor CV was met goed gevolg afleggen van de toets een noodzakelijke voorwaarde.

Naast de instructie via e-learning werd extra aandacht besteed aan het onderzoek door de regiomanagers in de werkoverleggen met de interviewers. De regiomanagers worden hiervoor apart geïnstrueerd door de opleiders.

Vanwege Covid-19 werden de instructies, conform de maatregelen van de overheid, bijgesteld. Interviews werden alleen persoonlijk afgenomen als aan de dan geldende Corona maatregelen voldaan kon worden. Lukte dit niet dan werd na werving van een telefoonnummer aan-huis, het interview telefonisch afgenomen.

3.4 Benaderingsstrategie

Alle steekproefpersonen voor CV werden aan-huis benaderd. De benaderstrategie werd uitgevoerd, conform de benaderstrategie beschreven in de nota "Benaderstrategie: instructie voor het benaderen van respondenten door interviewers van het CBS, oktober 2012". Aanvullend op de gangbare benaderingsstrategie beschikte de interviewer voor het onderzoek CV-SLI over een onderzoeksspecifieke folder waarin het hoe en waarom van het onderzoek werd toegelicht. De interviewer verstuurde de aanschrijfbrief, de onderzoeksspecifieke folder en de onconditionele incentive van 5 euro aan de respondent.

In de waarneemperiode januari 2019 tot en met december 2019 werden maandelijks 190 adressen uitgezet. Over de gehele waarneemperioden dienden 1.000 responsen gerealiseerd te worden.

De oorspronkelijke waarneemperiode van het veldwerk in 2020 zou plaatsvinden van januari 2020 tot en met december 2020. In deze waarneemperiode zou elke maand 190 adressen worden uitgezet. Uiteindelijk zijn in de maanden januari, februari, maart en april 2020 iedere maand 190 adressen uitgezet. In de maanden januari, februari en 1^e helft van maart 2020 is de oorspronkelijke benaderstrategie toegepast. Vanaf de tweede helft van maart 2020 kon door de lockdown geen face-to-face waarneming plaatsvinden en daardoor heeft er in de tweede helft van maart 2020 geen waarneming plaatsgevonden. In de maand april 2020 is als alternatief voor de face-to-face waarneming, de telefonische waarneming ingezet. Daarna werd de waarneming stilgelegd vanwege de Coronapandemie. Het veldwerk werd hervat begin december 2020, maar vanwege een tweede lockdown wederom stilgelegd in de tweede helft van december 2020. Het veldwerk werd hervat van september 2021 tot en met maart 2022. In oktober 2021 vond geen waarneming plaats. De uitzet per maand voor waarneemperiode september 2021 tot en met maart 2022 is weergegeven in tabel 3.4.1.

Tabel 3.4.1: Verwacht aantal uit te zetten adressen per waarneemperiode

Waarneemperiode	aantallen adressen
202109	325
202111	325
202112	650
202201	650
202202	375
202203	220
Totaal	2545

Nadat de vragenlijst voor CV was afgenomen werd vervolgens aan respondenten van CV van 18 jaar of ouder de vraag gesteld of ze bereid waren de schriftelijke vragenlijst voor SLI in te vullen. Respondenten die daartoe bereid waren ontvingen een incentive ter waarde van 5 euro. Ter verantwoording van de 5 euro werd een cadeauverantwoording door de steekproefpersonen ondertekend en teruggegeven aan de interviewer.

De respondenten die "ja" antwoorden op de wervingsvraag krijgen een papieren (papi) SLI vragenlijst. De interviewer maakt direct een afspraak voor een ophaalbezoek. De SLI vragenlijsten worden dus altijd door de interviewer opgehaald en geretourneerd naar het CBS.

Om de interviewers te informeren welke SLI vragenlijsten nog niet als respons zijn ontvangen, werden rappels gedraaid over de waarneemperiode 2019 conform de data, opgenomen in tabel 3.4.1. De data voor de rappels in waarneemperiode september 2021 tot en met maart 2022 zijn opgenomen in tabel 3.4.2. In de laatste twee waarneemperioden werden rappels gedraaid over de totale waarneemperioden.

Tabel 3.4.1: Gegevens over het draaien en versturen van de SLI rappels naar Interviewers

Waarneemperiode	Selectiedatum rappel
201901	29-jan-19
201901	12-feb-19
201902	26-feb-19
201902	12-mrt-19
201903	26-mrt-19
201903	9-apr-19

201904	23-apr-19
201904	7-mei-19
201905	28-mei-19
201905	11-jun-19
201906	25-jun-19
201906	9-jul-19
201907	30-jul-19
201907	13-aug-19
201908	27-aug-19
201908	10-sep-19
201909	1-okt-19
201909	15-okt-19
201910	29-okt-19
201910	12-nov-19
201911	26-nov-19
201911	10-dec-19
201912	7-jan-20
201912	14-jan-20
201901 t/m 201912	21-jan-20
201901 t/m 201912	28-jan-20

Tabel 3.4.2: Gegevens over het draaien en versturen van de SLI rappels naar Interviewers

Waarneemperiode	Selectiedatum rappel
202109	28-sep-21
202109	12-okt-21
202111	23-nov-21
202111	7-dec-21
202112	28-dec-21
202112	11-jan-22
202201	25-jan-22
202201	8-febr-22
202202	22-2-2022
202202	8-3-2022
202203	22-3-2022
202203	5-4-2022
202109 t/m 202203	12-apr-22

3.5 Veldwerk

3.5.1 Samenvatting aanpassingen onderzoeksopzet

- In het jaar 2020 zou het CV veldwerk plaatsvinden van januari tot en met december 2020. Vanwege de COVID-19-pandemie kon de waarneming niet geheel in de beoogde maanden uitgevoerd worden. Het CV veldwerk vond plaats in de maanden januari, februari en in de

eerste helft van maart 2020. Door de eerste lockdown kon er in de laatste helft van maart 2020 geen waarneming plaatsvinden.

- Vanwege de pandemie werd ook de gebruikelijke face-to-face waarneming aangepast. Interviews werden alleen persoonlijk afgenomen als aan de dan geldende Corona maatregelen voldaan kon worden, lukte dit niet dan werd na werving van een telefoonnummer aan-huis, het interview telefonisch afgenomen. In april 2020 is voor CV de telefonische waarneming ingezet als alternatieve waarneming.
- Het CV veldwerk is in de eerste helft van december 2020 hervat, maar is bij de tweede lockdown in de laatste helft van december 2020 opnieuw stilgelegd.
- Het CV veldwerk is opnieuw hervat vanaf september 2021 tot en met maart 2022 met een uitloop naar de maand april 2022 voor het ophalen van de papieren vragenlijsten.
- Indien door Covid-19 het niet mogelijk was om het interview persoonlijk af te nemen, is de mogelijkheid geboden om het interview middels beeldbellen af te nemen. Deze optie is door 27 interviewers aangeboden vanaf de derde week van januari 2022.
- Aan 27 interviewers zijn extra instructiebijeenkomsten en instructies gegeven rondom het aanbieden van beeldbellen aan respondenten (zie verder 3.5.2).

3.5.2 Beeldbellen

Naar aanleiding van de lockdown door Covid-19 is er voor CV/SLI gekeken naar een alternatieve waarneming indien geen face-to-face waarneming mogelijk was. Er is een pilot beeldbellen gestart met als doel om ervaringen uit het veld in kaart te brengen en respons binnen te halen tijdens de lockdown voor het CV/SLI onderzoek. Hiervoor zijn voor 27 interviewers extra instructiebijeenkomsten georganiseerd en ook extra instructiematerialen gemaakt. De pilot beeldbellen is van start gegaan in week 3 2022. In dezelfde week werd de lockdown echter beëindigd en mochten interviewers weer naar binnen om face-to-face waarneming af te nemen. Naar de pilot interviewers is gecommuniceerd dat face-to-face waarneming de voorkeur heeft, maar dat beeldbellen als alternatief aangeboden mocht worden als een face-to-face waarneming niet op een veilige manier kon plaatsvinden.

Om meer inzicht te krijgen in zowel de beleving van de respondent als van de interviewer ten opzichte van beeldbellen, werd de interviewers gevraagd om een formulier in te vullen met een aantal vragen na afloop van het bezoek. Van de 34 bruikbare formulieren zijn 26 formulieren ontvangen met beeldbellen als respons. Zowel respondenten als interviewers stonden positief tegenover dit alternatief. Op basis van de geringe resultaten is het nog te vroeg om hier conclusies uit te trekken, daar zal meer onderzoek naar beeldbellen voor nodig zijn.

3.5.3 Respons

Maandelijks, na afsluiting van de maandportie werd op de 4de werkdag van de maand voortgangsoverzichten geleverd met daarin de CV en SLI veldwerkresultaten per maand en totaal.

In de tabellen 3.5.3.1 tot en met 3.5.3.4 zijn de eindresultaten weergegeven (in aantallen en in percentages) van de onderzoeken CV en SLI over de perioden januari 2019 tot en met december 2019. Voor de periode januari 2019 tot en met december 2019 werd gestreefd naar 1.000 SLI responsen.

Tabel 3.5.3.1 Responsaantallen CV naar veldwerkperiode en totaal, 2019

	Veldwerkperiode												
	Totaal	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
Uitgezette steekproef	2280	190	190	190	190	190	190	190	190	190	190	190	190
Geen eindresultaat	0	0	0	0	0	0	0	0	0	0	0	0	0
Kaderfouten	40	2	4	2	6	0	5	6	1	1	1	6	6
Adres onvindbaar/bestaat niet	1	0	1	0	0	0	0	0	0	0	0	0	0
In aanbouw/afgebroken/sloopwoning	2	1	0	0	1	0	0	0	0	0	0	0	0
Geen woonadres	1	0	0	0	0	0	0	1	0	0	0	0	0
Leegstaand/tijdelijk onbewoond	8	0	0	1	1	0	2	1	0	0	0	1	2
OP overleden	1	0	0	0	0	0	0	0	0	0	1	0	0
OP verhuisd naar buitenland	16	1	2	1	1	0	2	2	0	1	0	3	3
Niemand komt in aanmerking als OP	0	0	0	0	0	0	0	0	0	0	0	0	0
OP onbekend	4	0	1	0	2	0	1	0	0	0	0	0	0
Instelling/tehuis	7	0	0	0	1	0	0	2	1	0	0	2	1
Non-respons	913	64	69	82	61	78	78	70	74	77	86	87	87
OP verhuisd in Nederland	76	4	4	8	5	3	8	4	13	7	5	11	4
Taalbarrière	106	3	6	11	11	11	7	7	11	9	9	10	11
Weigering	429	38	31	31	28	38	31	35	30	42	49	40	36
Geen gelegenheid / niet in staat	160	7	7	22	8	18	21	8	13	12	13	14	17
Geen contact met huishouden	102	8	10	9	8	8	7	14	5	5	10	12	6
Overig	40	4	11	1	1	0	4	2	2	2	0	0	13
Respons	1327	124	117	106	123	112	107	114	115	112	103	97	97
Volledig	1297	123	113	105	122	106	104	112	113	108	101	95	95
Afgebroken	30	1	4	1	1	6	3	2	2	4	2	2	2
Werving KvL	1175	114	107	98	114	98	93	97	101	92	91	84	86

Tabel 3.5.3.2 Responspercentages CV naar veldwerkperiode en totaal, 2019

	Veldwerkperiode												
	Totaal	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
% Kaderfouten	1,8	1,1	2,1	1,1	3,2	0,0	2,6	3,2	0,5	0,5	0,5	3,2	3,2
% Adres onvindbaar	0,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
% In aanbouw, afgebroken, sloopwoning	0,1	0,5	0,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
% Geen woonadres	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,0	0,0	0,0	0,0	0,0
% Leegstaand	0,4	0,0	0,0	0,5	0,5	0,0	1,1	0,5	0,0	0,0	0,0	0,5	1,1
% O.P. overleden	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,0	0,0
% O.P. verhuisd naar het buitenland	0,7	0,5	1,1	0,5	0,5	0,0	1,1	1,1	0,0	0,5	0,0	1,6	1,6
% Niemand komt in aanmerking als O.P.	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
% O.P. onbekend/onvindbaar	0,2	0,0	0,5	0,0	1,1	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0
% Naar instelling/tehuis binnen Nederland	0,3	0,0	0,0	0,0	0,5	0,0	0,0	1,1	0,5	0,0	0,0	1,1	0,5
% Adm. nonrespons capi	40,0	33,7	36,3	43,2	32,1	41,1	41,1	36,8	38,9	40,5	45,3	45,8	45,8
% O.P. verhuisd	3,3	2,1	2,1	4,2	2,6	1,6	4,2	2,1	6,8	3,7	2,6	5,8	2,1
% Anders (bv. storting centrale)	1,8	2,1	5,8	0,5	0,5	0,0	2,1	1,1	1,1	1,1	0,0	0,0	6,8
% Taalbarrière	4,6	1,6	3,2	5,8	5,8	5,8	3,7	3,7	5,8	4,7	4,7	5,3	5,8
% Weigering	18,8	20,0	16,3	16,3	14,7	20,0	16,3	18,4	15,8	22,1	25,8	21,1	18,9
% Geen gelegenheid	7,0	3,7	3,7	11,6	4,2	9,5	11,1	4,2	6,8	6,3	6,8	7,4	8,9
% Geen contact	4,5	4,2	5,3	4,7	4,2	4,2	3,7	7,4	2,6	2,6	5,3	6,3	3,2
% Respons	58,2	65,3	61,6	55,8	64,7	58,9	56,3	60,0	60,5	58,9	54,2	51,1	51,1
% Volledig	56,9	64,7	59,5	55,3	64,2	55,8	54,7	58,9	59,5	56,8	53,2	50,0	50,0
% Afgebroken	1,3	0,5	2,1	0,5	0,5	3,2	1,6	1,1	1,1	2,1	1,1	1,1	1,1
Target %	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2	55,2
% wervingsrespons KvL	90,6	92,7	94,7	93,3	93,4	92,5	89,4	86,6	89,4	85,2	90,1	88,4	90,5
Target % werving KvL	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5	91,5

Tabel 3.5.3.3 Responsaantallen SLI naar veldwerkperiode en totaal, 2019

	Veldwerkperiode												
	Totaal	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
Toezegging schriftelijke vragenlijst	1179	114	107	98	114	98	94	97	103	93	91	84	86
Aantal geweigerd	2	0	0	0	1	0	0	0	0	1	0	0	0
Aantal respons	1082	103	98	92	101	91	88	87	96	87	82	75	82
Aantal nog niet terug ontvangen	95	11	9	6	12	7	6	10	7	5	9	9	4
Target	1000	83	83	83	83	83	83	83	83	83	83	83	83

Tabel 3.5.3.4 Responspercentages SLI naar veldwerkperiode en totaal, 2019

	Veldwerkperiode												
	Totaal	201901	201902	201903	201904	201905	201906	201907	201908	201909	201910	201911	201912
Toezegging schriftelijke vragenlijst	1179	114	107	98	114	98	94	97	103	93	91	84	86
% Geweigerd	0,2	0,0	0,0	0,0	0,9	0,0	0,0	0,0	0,0	1,1	0,0	0,0	0,0
% Respons	91,8	90,4	91,6	93,9	88,6	92,9	93,6	89,7	93,2	93,5	90,1	89,3	95,3
% Nog niet terug ontvangen	8,1	9,6	8,4	6,1	10,5	7,1	6,4	10,3	6,8	5,4	9,9	10,7	4,7
Target %	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3	91,3

In de tabellen 3.5.3.5 tot en met 3.5.3.8 zijn de eindresultaten weergegeven (in aantallen en in percentages) van de onderzoeken CV en SLI over de perioden januari 2020 tot en met april 2020 en december 2020 tot en met maart 2022. Voor eerdergenoemde periode was eveneens het streven het behalen van 1000 SLI responsen.

Tabel 3.5.3.5 Responspercentages CV, editie 2019 naar veldwerkperiode en totaal, 2020-2022

	Veldwerkperiode											
	Totaal	202001	202002	202003	202004	202012	202109	202111	202112	202201	202202	202203
Uitgezette steekproef	3490	190	190	190	190	190	325	325	650	646	374	220
Geen eindresultaat	30	0	0	0	0	0	0	8	12	7	0	3
Kaderfouten	53	4	6	2	1	1	9	4	6	10	6	4
Adres onvindbaar/bestaat niet	6	0	0	1	0	0	0	1	0	2	1	1
In aanbouw/afgebroken/sloopwoning	3	0	0	0	0	0	1	0	2	0	0	0
Geen woonadres	1	0	0	0	0	0	0	1	0	0	0	0
Leegstaand/tijdelijk onbewoond	8	2	0	0	0	0	4	0	1	1	0	0
OP overleden	4	0	0	1	1	0	0	1	0	0	0	1
OP verhuisd naar buitenland	18	1	4	0	0	0	3	0	2	4	3	1
Niemand komt in aanmerking als OP	0	0	0	0	0	0	0	0	0	0	0	0
OP onbekend	4	0	0	0	0	0	0	1	0	2	1	0
Instelling/tehuis	9	1	2	0	0	1	1	0	1	1	1	1
Non-respons	2101	85	80	140	114	141	174	163	479	387	219	119
OP verhuisd in Nederland	92	9	2	2	1	3	12	8	11	21	17	6
Taalbarrière	130	11	6	3	3	7	11	12	27	23	17	10
Weigering	684	44	45	21	16	25	73	74	86	151	101	48
Geen gelegenheid / niet in staat	213	12	17	6	7	6	20	24	18	49	28	26
Geen contact met huishouden	194	8	5	0	33	3	33	11	12	56	20	13
Overig	419	1	5	108	54	80	20	17	57	48	17	12
Onveilige situatie	369	0	0	0	0	17	5	17	268	39	19	4
Respons	1305	101	104	48	75	48	142	150	152	242	149	94
Volledig	1288	100	103	48	72	48	138	148	149	241	147	94
Afgebroken	17	1	1	0	3	0	4	2	3	1	2	0
Werving KvL	1161	91	94	45	65	45	116	140	133	218	135	79

Tabel 3.5.3.6 Responspercentages CV, editie 2020-2022 naar veldwerkperiode en totaal

	Veldwerkperiode											
	Totaal	202001	202002	202003	202004	202012	202109	202111	202112	202201	202202	202203
% Kaderfouten	1,5	2,1	3,2	1,1	0,5	0,5	2,8	1,2	0,9	1,5	1,6	1,8
% Adres onvindbaar	0,2	0,0	0,0	0,5	0,0	0,0	0,0	0,3	0,0	0,3	0,3	0,5
% In aanbouw, afgebroken, sloopwoning	0,1	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,3	0,0	0,0	0,0
% Geen woonadres	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,0
% Leegstaand	0,2	1,1	0,0	0,0	0,0	0,0	1,2	0,0	0,2	0,2	0,0	0,0
% O.P. overleden	0,1	0,0	0,0	0,5	0,5	0,0	0,0	0,3	0,0	0,0	0,0	0,5
% O.P. verhuisd naar het buitenland	0,5	0,5	2,1	0,0	0,0	0,0	0,9	0,0	0,3	0,6	0,8	0,5
% Niemand komt in aanmerking als O.P.	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
% O.P. onbekend/onvindbaar	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,3	0,0	0,3	0,3	0,0
% Naar instelling/tehuis binnen Nederland	0,3	0,5	1,1	0,0	0,0	0,5	0,3	0,0	0,2	0,2	0,3	0,5
% Adm. nonrespons capi	60,2	44,7	42,1	73,7	60,0	74,2	53,5	50,2	73,7	59,9	58,6	54,1
% O.P. verhuisd	2,6	4,7	1,1	1,1	0,5	1,6	3,7	2,5	1,7	3,3	4,5	2,7
% Overig	3,7	0,5	2,6	56,8	28,4	42,1	6,2	5,2	8,8	7,4	4,5	5,5
% Onveilige situatie	19,6	0,0	0,0	0,0	0,0	8,9	1,5	5,2	41,2	6,0	5,1	1,8
% Taalbarrière	6,1	5,8	3,2	1,6	1,6	3,7	3,4	3,7	4,2	3,6	4,5	4,5
% Weigering	5,6	23,2	23,7	11,1	8,4	13,2	22,5	22,8	13,2	23,4	27,0	21,8
% Geen gelegenheid	12,0	6,3	8,9	3,2	3,7	3,2	6,2	7,4	2,8	7,6	7,5	11,8
% Geen contact	10,6	4,2	2,6	0,0	17,4	1,6	10,2	3,4	1,8	8,7	5,3	5,9
% Respons	37,4	53,2	54,7	25,3	39,5	25,3	43,7	46,2	23,4	37,5	39,8	42,7
% Volledig	36,9	52,6	54,2	25,3	37,9	25,3	42,5	45,5	22,9	37,3	39,3	42,7
% Afgebroken	0,5	0,5	0,5	0,0	1,6	0,0	1,2	0,6	0,5	0,2	0,5	0,0
Target %	54,5	55,2	55,2	55,2	55,2	54,3	54,3	54,3	54,3	54,3	54,3	54,3
% wervingsrespons KvL	89,0	91,0	91,3	93,8	90,3	93,8	84,1	94,6	89,3	90,5	91,8	84,0
Target % werving KvL	91,8	92,6	92,6	92,6	92,6	91,6	91,6	91,6	91,6	91,6	91,6	91,6

Tabel 3.5.3.7 Responsaantallen SLI naar veldwerkperiode en totaal, 2020-2022

	Veldwerkperiode											
	Totaal	202001	202002	202003	202004	202012	202109	202111	202112	202201	202202	202203
Toezegging schriftelijke vragenlijst	1167	91	94	46	66	45	116	138	134	221	135	81
Aantal geweigerd	0	0	0	0	0	0	0	0	0	0	0	0
Aantal respons	1055	84	85	38	55	40	104	125	127	200	123	74
Aantal nog niet terug ontvangen	111	7	9	8	11	5	11	13	7	21	12	7
Target	1000	83	83	83	83	84	144	144	288	286	166	97

Tabel 3.5.3.8 Responspercentages SLI naar veldwerkperiode en totaal, 2020-2022

	Veldwerkperiode											
	Totaal	202001	202002	202003	202004	202012	202109	202111	202112	202201	202202	202203
Toezegging schriftelijke vragenlijst	1167	91	94	46	66	45	116	138	134	221	135	81
% Geweigerd	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
% Respons	90,4	92,3	90,4	82,6	83,3	88,9	89,7	90,6	94,8	90,5	91,1	91,4
% Nog niet terug ontvangen	9,5	7,7	9,6	17,4	16,7	11,1	9,5	9,4	5,2	9,5	8,9	8,6
Target %	89,6	91,3	91,3	91,3	91,3	89,1	89,1	89,1	89,1	89,1	89,1	89,1

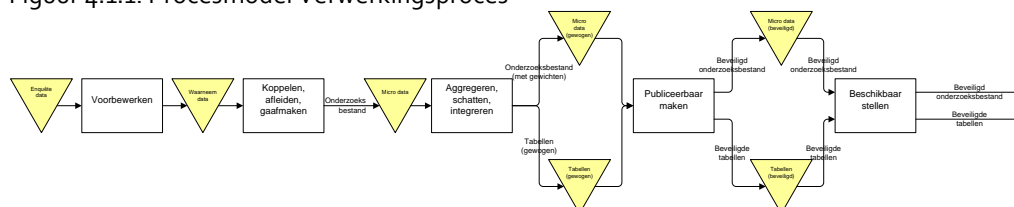
4. Verwerken (data processing)

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 plausibele statistische informatie te kunnen ontleen. De verwerking van de data is zoveel mogelijk geautomatiseerd. Het verwerken is uitgevoerd door het team Dynamisering Verwerking, sector Arbeid, Inkomen en Leefsituatie (SAL) van de hoofddirectie Sociaaleconomische en Ruimtelijke Statistieken (SER).

4.1 Inleiding

Onderstaande figuur geeft het verwerkingsproces van de onderzoeksdata in hoofdlijnen aan. De vijf deelprocessen (weergegeven in de vierkanten blokken) worden in de bijlage in meer detail beschreven.

Figuur 4.1.1: Procesmodel Verwerkingsproces



Hieronder wordt de verwerking van CV en SLI 2017-2018 beschreven en wordt aandacht besteed aan de keuzes die hierin gemaakt zijn.

4.2 Invoer

CV is een CAPI onderzoek, dat betekent dat de respondent thuis bezocht wordt door een interviewer en dat de enquête wordt afgenomen met gebruik van een computergestuurde vragenlijst. Dat betekent dat range- en routingcontroles tijdens het veldwerk worden uitgevoerd. En dat de data geen controles en correcties behoeven.

SLI is een PAPI onderzoek, de respondent vult thuis zelf een papieren vragenlijst in. Deze wordt vervolgens na binnenkomst bij het CBS in de computer ingevoerd bij de hoofddirectie Dataverzameling (DVZ). Hierbij worden van alle afzonderlijke antwoordmogelijkheden van elke vraag vastgelegd of de respondent die heeft aangekruist of niet. Dat betekent dat in de verwerking door SAL per vraag een definitief antwoord wordt bepaald en dat de range- en routecontroles worden uitgevoerd.

4.3 Controleren en uniformeren waarneemdata

In de computergestuurde vragenlijsten komen variabelen voor die nodig zijn voor het vraaggesprek, niet alle variabelen zijn nodig voor de verwerking. In een eerste stap wordt een selectie gemaakt van variabelen uit de vragenlijst en van respondenten. Alleen CV respondenten met een volledige respons worden verwerkt. Een SLI respons dient gekoppeld te kunnen worden aan een CV respons.

Volgens DVZ zijn er 1.297 respondenten voor CV 2019 en 1.288 respondenten voor CV 2020-2022. Door SAL zijn respectievelijk 1.293 en 1.281 respondenten verwerkt. Volgens DVZ zijn er 1.082

respondenten voor SLI 2019 en 1.055 respondenten voor SLI 2020-2022. Door SAL zijn respectievelijk 1.058 en 1.041 respondenten verwerkt.

De SLI vragenlijst kent twee varianten. In variant A komen vragen voor over wat een goede ouder wel en niet doet. In variant B zijn er vragen over de bemoeienis van de overheid met gezondheid en duurzaamheid.

In 2019 komen beide varianten exact gelijk voor, elk met 529 respondenten (50%). In 2020-2022 komen er 543 respondenten (52%) met variant A en 498 (48%) respondenten met variant B voor.

4.4 Typeren

In de CV vragenlijst komen de 'standaard' vragen voor naar Onderwijs (actueel en gevolgd onderwijs, inclusief in het buitenland). Deze vragen worden door een typeersysteem gehaald. In plaats van de vragen over onderwijs uit de vragenlijst, worden alleen de typeercodes opgeleverd van het hoogst behaalde onderwijs, ingedikt in twee variabelen (Afl_SOI3_HB en Afl_SOI5_HB).

Tabel 4.4.1 Publicatie-indeling SOI 2006 hoogst behaalde opleiding

Publicatie-indeling SOI 2006 hoogst behaalde opleiding	
Afl_SOI5_HB aggregatie 2	Afl_SOI3_HB aggregatie 1
10 basisonderwijs	1 lager
20 VMBO, MBO 1, AVO onderbouw	1 lager
21 w.o. AVO onderbouw	1 lager
31 w.o. MBO 2 en 3	2 middelbaar
32 w.o. MBO 4	2 middelbaar
33 w.o. HAVO, VWO	2 middelbaar
40 Hbo, WO Bachelor	3 hoger
50 WO Master, Doctor	3 hoger
99 onbekend	9 onbekend

4.5 Verrijken

Voor de verrijking van het onderzoek worden uit het Stelsel van Sociaal-Statistische Bestanden variabelen opgevraagd van personen van 16 jaar en ouder woonachtig in Nederland. Voor editie 2019 is als peildatum 01-01-2019 gebruikt en voor editie 2020-2022 als peildatum 01-01-2021. Het bestand met deze variabelen wordt zowel gebruikt voor de verrijking als voor de weging (voor het aanmaken van de randtotalen, gebruikt bij de weging).

Tabel 4.5.1 Verrijkingsvariabelen CV/SLI, editie 2019 en 2020-2022

Variabele	Omschrijving
Persoonskenmerken	
GBAAANTALoudersBuitENL AND	Aantal ouders in het buitenland geboren
GBAGENERATIE	Generatie
ETNGROEPKORT ₁	herkomstgroepering cbs omgecodeerd voor statline-publicatie 3-deling incl. onbekend
LFT	leeftijd in jaren op peilmoment
VRLGBAGESLACHT	Geslacht
BURKLASSE ₄	burgerlijke staat in 4 klassen
GBABURGERLIJKESTAATNW	GBA burgerlijke staat in 7 klassen
Regionale indelingen	
GEMJJJJ	Gemeentecode van de woongemeente op peildatum
STEDGEM	stedelijkheid van gemeente van de woning op peilmoment
PROVINCIE	provincie van de woning op peilmoment
LANDSDEEL	landsdeel van de woning op peilmoment
Huishouden samenstelling	
TYPHH	Type huishouden op peildatum
PLHH	Plaats in het huishouden op peildatum
VRLAANTALPERSHH	Aantal personen in huishouden
HUISHOUDNR	Nummer huishouden op verrijksdatum
Inkomen (zo recent mogelijk)	
INHPOPIV	populatieafbakening huishoudens integrale inkomens- en vermogensstatistiek
INHPRIMINKH	Primair inkomen van het huishouden
INHBRUTINKH	Bruto inkomen van het huishouden
INHBESTINKH	Besteedbaar inkomen van het huishouden
INHGESTINKH	Gestandaardiseerd besteedbaar inkomen van het huishouden
INHBIHJ	Belangrijkste inkomensbron van het huishouden
INPPERSPRIM	Persoonlijk primair inkomen
INPPERSBRUT	Persoonlijk bruto inkomen
INPPERSINK	Persoonlijk primair inkomen
INPSECJ	sociaaleconomische categorie op jaarbasis
INHP100HGEST	percentielgroepen gestandaardiseerd inkomen particuliere huishoudens
CWI	UWV bestand met ingeschreven werkzoekenden
CWI_INGESCHREVEN	De persoon die op peildatum ingeschreven, is deze dan werkend of niet werkend

4.6 Afleiden

Meerkeuzevragen uit de CV vragenlijst worden 'gedichotomiseerd'. Dat wil zeggen dat er variabelen worden gemaakt per antwoordmogelijkheid en dat per variabele wordt aangegeven of de respondent deze antwoordmogelijkheid heeft genoemd of niet (Ja, Nee). Alleen de 'gedichotomiseerde' variabelen worden opgeleverd, niet de meerkeuzevraag. Het gaat om slechts één meerkeuzevraag:

Tabel 4.6.1 Meerkeuzevraag over religie

Geloof	VAR198[1..7]
	VAR198_RK
	VAR198_NH
	VAR198_Gereform
	VAR198_Hindoe
	VAR198_Islam
	VAR198_Boeddhis
	VAR198_Ander

Een aantal variabelen worden omgecodeerd of ingedikt. Het betreft vooral verrijkingsvariabelen. Sommige hebben alleen een andere naam gekregen, die aansluit bij eerdere onderzoeksjaren. Slechts een enkele variabele is een afleiding gebaseerd op meerdere invoervariabelen.

Tabel 4.6.2 Hercoderingen, indikkingen en afleidingen CV/SLI, 2019 en 2020-2022

SSB	Naam afleiding
VRLGBAGESLACHT	ag_Geslacht
LFT	GBALft, LFT10, LFT11
BURKLASSE4	BurgStat
VRLGBABURGERLIJKESTAATNW	BurgStat7, ag_BurgStat
ETNGROEPKORT1	ag_Herkomst, Herkomst3
VRLGBAAANTALOUDESBUITENLAND	Afl_Generatie, ag_Generatie
VRLGBAGENERATIE	Afl_Generatie, ag_Generatie
VRLAANTALPERSHH	HHGrootte
VRLPLHH	ag_PlaatsHH
VRLTYPHH	TypHH, ag_TypeHH
GEMJJJJ	Gemcode, ProvPlus
LANDSDEEL	Landsdl
PROVINCIE	Provincie, Prov_12, ProvPlus
STEDGEM	StedGem
INHPOPIIV	IndHINK
INHBRUTINKH	BVRBRUTINKH, CATBVRBRUTINKH
INHPRIMINKH	BVRPRIMINKH, CATBVRPRIMINKH
INHBESTINKH	BVRBESTINKH, CATBVRBESTINKH
INHGESTINKH	BVRGESTINKH, CATBVRGESTINKH
INHBBIHJ	
INPPERSBRUT, INPPERSPRIM, INPPERSINK	PINK
INPPERSBRUT	PERSBRUT, CATPERSBRUT
INPPERSPRIM	PERSPRIM, CATPERSPRIM
INPPERSINK	PERSINK, CATPERSINK
INPSECJ	
INHP100HGEST	BVRPERCGESTINKH, Huishoudinkomen*, Kwartielen*
INGESCHREVEN	CWI_Ingeschreven, CWI2

4.7 Correcties: gaafmaken, imputeren

Het controleren en eventueel corrigeren van enquêtedata is alleen nodig voor het PAPI deel. In de vragenlijst staan weliswaar instructies voor het invullen, maar respondenten houden zich daar niet altijd aan. Zo kruisen ze meer dan een antwoordmogelijkheid aan waar dat niet de bedoeling is of beantwoorden ze vragen die niet op de route liggen. In de verwerking wordt een random keuze gemaakt bij teveel gegeven antwoorden. Bij route problemen wordt per vragen set een werkwijze gekozen en worden de bijbehorende vragen gecorrigeerd.

5. Non-responsanalyse

5.1 Afbakenen statistische respons en nonresponsanalyse

In dit hoofdstuk wordt verslag gedaan van de responsanalyse op het onderzoek Culturele Veranderingen (CV) en de SCP Leefsituatie Index (SLI), die gecombineerd werden uitgevoerd. De gegevens worden gepresenteerd per editie, editie 2019 die liep van januari tot en met december 2019, en editie 2020, die liep van januari 2020 tot en met april 2022. CV werd face-to-face afgenomen (CAPI) bij respondenten van 16 jaar en ouder. Aan respondenten van CV van 18 jaar of ouder werd gevraagd of ze bereid waren om een schriftelijke vervolgvragenlijst (SLI) in te vullen. Respondenten die bereid waren om deel te nemen aan het schriftelijke vervolgonderzoek ontvingen een incentive ter waarde van vijf euro.

In deze non-responsanalyse is allereerst gekeken naar hoeveel respondenten getrokken zijn in de steekproef voor CV, hoeveel van hen daadwerkelijk uitgezet zijn in veld, en hoeveel daarvan tot de respons respectievelijk non-respons gerekend kunnen worden.

Tot de non-respons worden hier gerekend degenen:

- die deelname weigerden;
- die behoren tot de administratieve non-respons, waaronder taalbarrière; verhuisd binnen Nederland en on- (volledig)bewerkt retour;
- waarvan het interview voortijdig afgerond werd;
- geen gelegenheid hadden voor het interview tijdens de veldwerkperiode;
- waar geen contact mee gelegd kon worden.

Tot de respons wordt gerekend de uitzet minus kaderfouten en afgevallen records bij verwerking. Er worden dus op basis van bepaalde verwerkingskeuzes nog records verwijderd. Dit is de statistische respons en deze wijkt af van de reeds gepresenteerde respons in hoofdstuk 3, waar het gaat om de administratieve respons. Daarin worden alle administratief/technisch correct binnengekomen responsen / uitzet minus kaderfouten als respons gerekend.

5.2 Steekproef- en responsverantwoording

Omdat uit ervaring bekend is dat er meer steekproefpersonen getrokken moeten worden dan daadwerkelijk benaderd kunnen worden, zijn er vooraf meer steekproefpersonen getrokken.

Voor CV, editie 2019 werden 2738 steekproefpersonen getrokken. Daarvan waren er 173 niet beschikbaar voor uitzet, omdat ze in de 12 maanden voorafgaand aan het onderzoek al eerder aan CBS-onderzoek hadden deelgenomen, hun adresgegevens onjuist waren, ze in een instelling verbleven of dat ze in een gebied woonden dat buiten de waarneming viel. Uiteindelijk bleken er 2565 adressen beschikbaar voor uitzet. Na een steekproefreductie, werden er daarvan 2280 daadwerkelijk uitgezet in het veld (zie Tabel 5.2.1).

Voor CV, editie 2020-2022 werden er in totaal, dus verdeeld over de drie veldwerkperiodes (zie ook hoofdstuk 3), 4455 steekproefpersonen getrokken. Daarvan waren er 185 niet beschikbaar voor uitzet. Na steekproefreductie (780) werden er 3490 adressen uitgezet.

Tabel 5.2.1: Steekproefverantwoording CV en SLI, editie 2019 en 2020-2022

	2019	2020-2022
Getrokken steekproef	2738	4455
<u>Onbruikbare steekproef, waarvan:</u>	173	185

- In eerdere CBS-steekproef	121	95
- Institutionele bevolking	1	4
- Met onbekende straatnaam	49	75
- Buiten waarneming SDV (CAPI)	2	11
Bruikbare steekproef	2565	4270
Zonder telefoonnummer	1316	528
Met telefoonnummer, waarvan:	964	2962
- Vast telefoonnummer	751	2083
- Mobiel telefoonnummer	213	879
Steekproefreductie	285	780
Uitgezet	2280	3490

Tabel 5.2.2 laat zien dat voor CV, zonder de kaderfouten (respondent overleden, verhuisd naar buitenland, onbekend of in een instelling), er in 2019 2240 personen overblijven, waarvan er 1297 daadwerkelijk deelgenomen hebben aan het onderzoek. Voor editie 2020-2022 bleven er rekening houdend met de kaderfouten (83 steekproefelementen), 3407 personen over. Daarvan hebben er 1281 aan het onderzoek deelgenomen.

Tabel 5.2.2: Respons naam CV en SLI, editie 2019 en 2020-2022

	Totaal 2019	Totaal 2020-2022
<i>Uitgezette steekproef</i>	2280	3490
<i>Kaderfouten</i>	40	83
w.v. OP overleden	1	4
w.v. OP verhuisd naar buitenland	16	18
w.v. OP onbekend/onvindbaar	4	4
w.v. naar instelling/ tehuis binnen Nederland	7	9
w.v. leegstand/tijdelijk onbewoond	8	8
w.v. geen woonadres/adres onvindbaar	4	7
w.v. in aanbouw/afgebroken/sloopwoning		3
w.v. niet uitgezet CV		30
<i>Nettosteekproef (=uitgezette steekproef excl. kaderfouten en uitval bij verwerking)</i>	2240	3407
Respons (uit overzicht DVZ)	1297	1288
Respons (bruikbaar voor verwerking)	1297	1281
Non-respons	947	2126
w.v. volledig onbewerkt retour	21	215
w.v. gedeeltelijk onbewerkt retour	16	201
w.v. OP verhuisd binnen Nederland	76	92
w.v. weigering	429	684
w.v. geen gelegenheid tijdens veldwerkperiode	117	134
w.v. lang niet in staat (zie, te zwak e.d.)	43	79
w.v. taalbarrière	106	130
w.v. Geen contact met huishouden	102	194

w.v. afgebroken interview	30	17
w.v. onveilige situatie	1	369
w.v. vragenlijst binnen zonder eindresultaat	2	
w.v. non-respons als gevolg van mixed mode		3
w.v. niet beschikbaar voor verwerking	4	7
w.v. geen internet, telefonisch benaderen		1
SLI		
Doelpopulatie SLI 18 jaar of ouder	2143	3299
Toestemming deelname SLI	1173	1167
Vragenlijsten als geretourneerd afgeboekt in DVZ overzicht	1082	1055
Niet bruikbaar/non-respons voor verwerking	45	21
Niet afgeboekt als respons in responsoverzicht DVZ, wel verwerkt als respons	20	7
Respons SLI	1057	1041

Tabel 5.2.3. geeft een overzicht van de uiteindelijke, statistische respons. Dit zijn de records die na verwerking over zijn gebleven. Volgens DVZ waren dat er 1.297 respondenten voor CV, editie 2019 en 1.288 respondenten voor CV, editie 2020-2022. Door SAL zijn respectievelijk 1.293 en 1.281 respondenten verwerkt. De responspercentages voor editie 2019 en 2020-2022 liggen daarmee respectievelijk op 57,9 en 37,6.

Volgens DVZ zijn er 1.082 respondenten voor SLI 2019 en 1.055 respondenten voor SLI 2020-2022. Door SAL zijn respectievelijk 1.058 en 1.041 respondenten verwerkt. Omdat in editie 2019 1 vragenlijst achteraf toch niet geschikt bleek als respons voor verwerking, kwam de uiteindelijke respons op 1057 uit. Het responspercentage van SLI ligt daarmee op 49,3 procent.

Tabel 5.2.3: Respons CV en SLI, 2019 en 2020-2022

	2019		2020-2022	
	CV CAPI	SLI PAPI	CV CAPI	SLI PAPI
	abs.	abs.	abs.	abs.
Netto steekproef	2240	2143	3407	3299
Non-respons	1297	1086	1281	2258
Respons	943	1057	2126	1041
	%	%	%	%
Responspercentage totaal	57,9	49,3	37,6	31,6

5.3 Selectiviteit van de respons naar achtergrondkenmerken

Voor de analyse van de responscijfers is een databestand samengesteld met informatie bevat over een aantal kenmerken voor alle steekproefpersonen. Voor de non-responsanalyse van CV is gepercenteerd op de netto steekproef, dat wil zeggen de uitgezette steekproef minus kaderfouten en eventuele uitval bij verwerking.

Het overall responspercentage van CV waarmee de respons van de verschillende bevolkingsgroepen vergeleken kan worden ligt op 57,9 procent in 2019 en 37,6 in 2020-2022. Voor 2019 was dat nagenoeg gelijk aan het percentage in 2017/2018 (58 procent). Voor 2020-2022 lag het responspercentage lager. Dit hangt deels samen met de coronapandemie en gerelateerde maatregelen, waardoor de waarneming tot twee maal toe werd stopgezet en weer opgestart. Daarbij werden uiteindelijk meer steekproefelementen uitgezet om in elk geval het beoogde aantal SLI-respondenten (van 1000) te behalen.

Voor SLI ligt dit percentage waarmee de respons van de groepen vergeleken kan worden in 2019 op met 49,3 procent en in 2020-2022 op 31,6. Net als in voorgaande jaren is de respons van CV in beide edities hoger dan die van het schriftelijke vervolg. Tabel 3.2.1 geeft het responspercentage per subgroep. Dit geeft aan welke groepen er beter of juist slechter responderen dan gemiddeld.

Tabel 5.3.1: Respons CV en SLI 2019, naar achtergrondkenmerken (%)¹⁾

	2019		2020-2022	
	CV	SLI	CV	SLI
Responspercentage totaal	57,9	49,3	37,6	31,6
<i>Geslacht</i>				
man	57,5	48,8	38,8	32,2
vrouw	58,6	49,8	36,9	31,3
<i>Leeftijd</i>				
16-24 jaar ²⁾	55,4	41,2	32,9	23,9
25-34 jaar	51,7	42,3	31,4	25,4
35-44 jaar	56,7	47,3	34,9	29,3
45-54 jaar	55,9	46,8	37,9	30,9
55-64 jaar	56,8	51,8	39,6	34,6
65-74 jaar	70,6	63,4	49,8	45,1
75 jaar of ouder	62,3	51,9	38,5	30,8
<i>Burgerlijke staat</i>				
gehuwd (incl. partnerschap)	61,7	53,9	41,8	35,6
gescheiden	56,4	44,1	42,4	36,4
verweduwd	64,2	56,0	41,8	35,9
nooit gehuwd geweest	53,1	43,2	30,9	24,5
<i>Herkomst</i>				
Nederlands	63,4	54,7	41,2	34,8
Migratieachtergrond, westers	42,5	36,0	29,7	25,9
Migratieachtergrond, niet westers	39,3	27,0	23,1	16,9

	2019		2020-2022	
	CV	SLI	CV	SLI
<i>Huishoudensinkomen</i>				
Laag (deciel t/m 3)	49,2	41,5	32,9	26,6
Midden (deciel 4 t/m 7)	61,8	52,5	38,7	32,9
Hoog (deciel 8+)	62,9	53,7	41,3	35,3
<i>Landsdeel</i>				
Noord-Nederland	62,1	55,0	41,4	32,7
Oost-Nederland	61,2	49,9	38,2	32,4
West-Nederland	52,8	43,1	34,3	28,0
Zuid-Nederland	64,4	59,2	43,4	39,1
<i>Stedelijkheid van gemeente</i>				
zeer sterk stedelijk	49,9	42,3	29,6	23,9
sterk stedelijk	60,3	50,2	38,8	33,6
matig stedelijk	58,6	50,6	38,4	32,5
weinig stedelijk	61,6	52,2	43,0	36,3
niet stedelijk	63,9	57,4	45,3	36,7

1) In 2019 waren voor 12 personen de achtergrondkenmerken in deze tabel niet bekend. In 2020-2022 waren dat 39 personen.

2) Bij SLI is de populatie 18 jaar of ouder

De respons was in beide edities van CV, net als in voorgaande edities, selectief naar bevolkings- en regionale kenmerken (Tabel 5.3.1). Mannen doen in beide edities evenveel mee aan CV als vrouwen; er is geen significant verschil. Net als in voorgaande edities namen 65 tot 74-jarigen het vaakst deel van alle leeftijdsgroepen. Voor CV, editie 2019 valt op dat de jongeren van 16 tot 25 jaar relatief in 2019 wat beter deelnamen dan in voorgaande edities. In 2020-2022 was dat weer minder ten opzichte van de leeftijdsgroepen vanaf 55 jaar. Verder nemen in beide edities Nederlanders zonder migratieachtergrond het vaakst deel, en mensen met een niet-westerse migratieachtergrond het minst vaak. Degenen in de laagste inkomensgroep en inwoners van de sterkst verstedelijkte gebieden en oost- en west-Nederland nemen minder vaak deel dan mensen in de hogere inkomensgroepen, minder verstedelijkte gemeentes en noord- en zuid-Nederland. Migratieachtergrond differentieert net als bij voorgaande edities sterk naar de respons; 39,3 procent van de mensen met een niet-westerse achtergrond en 42,5 procent van de mensen met een westerse achtergrond hebben deelgenomen, tegenover 63,4 procent van degenen met een Nederlandse achtergrond.

Voor SLI geldt, voor beide edities, grofweg eenzelfde beeld als bij CV. Tussen mannen en vrouwen is er geen significant verschil naar deelname. De verschillen naar de overige achtergrondkenmerken zijn allemaal significant. De leeftijdsgroepen vanaf 55 jaar namen relatief goed deel ten opzichte van de jongste twee leeftijdscategorieën (18 tot 24 en 25 tot 34 jaar). 65-tot 74-jarigen nemen net als bij CV het vaakst deel aan SLI. Bij herkomst is de respons van de mensen met een niet-westerse migratieachtergrond met 16,9 procent fors lager dan die van westerse allochtonen (25,9 procent) en vooral autochtone Nederlanders (34,8 procent). De laagste inkomensgroep respondeert ook hier aanmerkelijk minder dan de hogere inkomensgroepen. Respondenten in West-Nederland hebben wederom de laagste respons en de inwoners van zeer sterk verstedelijkte gebieden nemen ook beduidend minder deel aan SLI dan de inwoners van andere landsdelen en minder verstedelijkte gebieden.

Tabel 5.3.2 geeft de verdeling naar achtergrondkenmerken voor de populatie, de uitgezette steekproef en de respons. Ook dit geeft een beeld van de selectiviteit naar achtergrondkenmerken. In de steekproeftrekking wordt rekening gehouden met mogelijke kaderfouten, waardoor de verdelingen in de getrokken steekproef en de uiteindelijk uitgezette steekproef nauwelijks verschillen.

Tabel 5.3.2 Verdelingen populatie, netto steekproef en respons CV en SLI, 2019

	CV			SLI		
	Populatie (16+) 1 jan. 2019	Netto steekproef	Respons	Populatie (18+) 1 jan. 2019	Netto steekproef	Respons
	%	%	%	%	%	%
<i>Geslacht</i>						
man	49,4	49,5	49,0	49,4	49,6	49,1
vrouw	50,6	50,5	51,0	50,6	50,4	50,9
<i>leeftijd</i>						
16-17 jaar	2,1	3,1	3,3	-	-	-
18-19 jaar	3,1	2,8	3,0	3,1	2,9	2,6
20-24 jaar	7,7	6,7	5,6	7,8	6,9	5,6
25-29 jaar	7,9	8,1	7,3	8,1	8,4	7,5
30-34 jaar	7,6	7,0	6,2	7,8	7,2	5,9
35-39 jaar	7,3	7,1	6,6	7,5	7,3	6,3
40-49 jaar	16,1	17,3	16,6	16,4	17,9	16,7
50-59 jaar	17,7	17,9	18,1	18,1	18,4	19,8
60-69 jaar	14,8	14,4	15,8	15,1	14,8	17,5
70-79 jaar	10,7	10,4	12,4	10,9	10,7	13,0
80 jaar of ouder	5,1	5,3	5,2	5,2	5,5	5,2
<i>Burgerlijke staat</i>						
Gehuwd (incl. Partnerschap)	47,3	48,4	51,4	48,7	50,3	55,0
gescheiden	9,4	8,4	8,2	9,7	8,8	7,9
verweduwd	5,6	4,9	5,4	5,8	5,1	5,8
nooit gehuwd geweest	37,7	38,3	35,0	35,9	35,8	31,4

Vervolg Tabel 5.3.2 Verdelingen populatie, netto steekproef en respons CV en SLI, 2019

	CV			SLI		
	Populatie (16+)	Netto steekproef	Respons	Populatie (18+)	Netto steekproef	Respons
	1 jan. 2019			1 jan. 2019		
<i>Herkomst</i>						
Nederlands	76,9	76,5	83,5	77,0	76,7	85,1
Migratieachtergr, westers	10,7	11,4	8,3	10,8	11,5	8,4
Migratieachtergr, niet- westers	12,4	12,1	8,2	12,2	11,8	6,4
<i>Huishoudensinkomen</i>						
Laag (deciel t/m 3)	33,1	32,7	27,7	33,2	32,7	27,5
Midden (deciel 4 t/m 7)	30,5	31,4	33,4	30,4	31,4	33,4
Hoog (deciel 8+)	36,4	35,9	38,9	36,4	35,9	39,1
<i>Landsdeel</i>						
Noord-Nederland	10,0	10,2	10,9	10,0	10,2	11,4
Oost-Nederland	20,9	21,3	22,4	20,8	21,0	21,3
West-Nederland	47,5	46,5	42,3	47,6	46,6	40,8
Zuid-Nederland	21,6	22,0	24,4	21,6	22,2	26,6
<i>Stedelijkheid van gemeente</i>						
zeer sterk stedelijk	23,9	24,2	20,8	24,6	24,4	20,9
sterk stedelijk	30,9	30,5	31,7	30,4	30,5	31,0
matig stedelijk	15,4	15,6	15,8	15,5	15,5	15,9
weinig stedelijk	21,7	22,1	23,4	21,2	22,1	23,4
niet stedelijk	8,0	7,6	8,3	8,3	7,6	8,8

Tabel 5.3.3 Verdelingen populatie, netto steekproef en respons CV en SLI (%), 2020-2022

	CV			SLI		
	Populatie (16+) 1 jan. 2021	Netto steekproef	Respons	Populatie (18+) 1 jan. 2021	Netto steekproef	Respons
man	49,5	49,5	50,7	49,4	49,5	50,2
vrouw	50,5	50,5	49,3	50,6	50,5	49,8
<i>leeftijd</i>						
16-17 jaar	2,8	2,5	2,5	-	-	-
18-19 jaar	2,0	2,4	1,9	1,7	2,4	1,6
20-24 jaar	7,8	7,3	5,8	8	7,4	5,3
25-29 jaar	7,9	8,2	5,9	8,2	8,4	5,4
30-34 jaar	7,9	7	6,8	8,1	7,2	7,2
35-39 jaar	7,3	7,3	6,8	7,6	7,4	6,9
40-49 jaar	15,1	15,1	14,3	15,6	15,5	14,7
50-59 jaar	17,9	19,3	19,9	18,5	19,8	20
60-69 jaar	14,9	14,5	16,8	15,4	14,9	19,2
70-79 jaar	11,2	11,6	14,8	11,5	11,8	15,8
80 jaar of ouder	5,2	5	4,6	5,3	5,1	4,1
<i>Burgerlijke staat</i>						
Gehuwd	46,4	47,8	52,8	47,7	49,4	55,2
gescheiden	9,4	9,5	10,7	9,7	9,8	11,3
verweduwd	5,5	5,5	6,0	5,7	5,6	6,4
ongehuwd	38,6	37,2	30,5	36,9	35,2	27,1
<i>Migratieachtergrond</i>						
Nederlands	75,8	77,6	84,5	75,9	77,8	85,3
Westers	11,1	10,2	8,0	11,2	10,2	8,3
Niet- westers	13,1	12,2	7,5	12,9	12	6,4
<i>Huishoudensinkomen</i>						
Laag (decil t/m 3)	32,7	31,1	27,1	32,8	31,4	26,3
Midden (decil 4 t/m 7)	30,7	33,0	33,8	30,5	32,8	34
Hoog (decil 8+)	36,7	35,8	39,1	36,7	35,8	39,8
<i>Landsdeel</i>						
Noord-Nederland	10,0	10,4	11,4	10,0	10,3	10,6
Oost-Nederland	20,9	21,2	21,4	20,9	21,2	21,6
West-Nederland	47,7	46,9	42,5	47,7	47,2	41,5
Zuid-Nederland	21,4	21,5	24,6	21,4	21,3	26,3
<i>Stedelijkheid gemeente</i>						
zeer sterk stedelijk	26,0	25,5	19,9	26,3	25,7	19,3
sterk stedelijk	30,1	29,4	30,1	30,0	29,3	31
matig stedelijk	14,5	15,5	15,7	14,9	15,6	15,9
weinig stedelijk	21,3	22,1	25,1	20,8	21,9	25
niet stedelijk	8,2	7,6	9,1	7,9	7,6	8,8

Het ligt voor de hand dat de groepen met een hogere respons, ook vaak iets oververtegenwoordigd zijn ten opzichte van de netto steekproef en de populatie. Dit is bij beide edities inderdaad het geval (zie ook de tabellen 5.2.2 en 5.2.3). Het gaat bij CV bijvoorbeeld om personen in de leeftijdsgroepen tussen 50 en 80 jaar, gehuwden, mensen met een hoog huishoudensinkomen en autochtonen. Bij SLI is deze oververtegenwoordiging nog duidelijker te zien.

Groepen met een lage respons zijn vaak juist ondervertegenwoordigd. Voorbeelden daarvan bij CV zijn mensen tussen 20 en 29 jaar, de ongehuwden, degenen met een laag huishoudensinkomen, inwoners van zeer sterk verstedelijkte gebieden en mensen met een niet-westerse migratieachtergrond. Ook bij SLI gaat het vooral om groepen als de ongehuwden, de inwoners van West-Nederland en de zeer sterk verstedelijkte gebieden, mensen met het laagste inkomen of een niet-westerse migratieachtergrond.

R-indicatoren

Om de selectiviteit van de steekproef na te gaan zijn ook de zogenoemde R-indicatoren berekend, die de representativiteit van de opgenomen kenmerken weergeeft (Schouten en Cobben, 2007). Deze indicator kan een waarde tussen 0 en 1 aannemen, waarbij een score van 1 betekent dat er geen sprake van selectiviteit is en de steekproef dus volledig representatief.

In 2019 was voor CV de overall R-indicator 0,765. Dit is een fractie hoger dan in 2018, toen was de indicator 0,758. Toch is de indicator nog steeds wat lager dan in voorgaande jaren, toen deze varieerde van 0,787 in 2012 tot 0,80 in 2016. Dit wijst erop dat de respons in 2019 selectiever was dan voorheen.

Naast R-indicatoren zijn ook partiële r-indicatoren berekend. Deze indicatoren ontbinden de variatie in responskansen in componenten die kunnen worden toegeschreven aan variabelen of categorieën. Conditionele partiële R-indicatoren meten de bijdrage van een variabele, rekening houdend met de samenhang met andere variabelen. In lijn met de hierboven beschreven resultaten lieten bij CV 2019 migratieachtergrond, leeftijd en gestandaardiseerd besteedbaar huishoudinkomen de hoogste partiële R-indicatoren zien. Dit betekent dat deze kenmerken het meest selectief waren. Voor de overige kenmerken is de selectiviteit van de respons niet significant. De coëfficiënten van de selectieve kenmerken (hier niet gepresenteerd), die iets zeggen over de categorieën die het meest verantwoordelijk zijn voor de selectie, lieten een vergelijkbaar beeld zien aan de eerder gepresenteerde resultaten aan het begin van deze paragraaf (zie ook de eerdere rapportage over de responsanalyse van CV in 2019; Coumans, 2020).

SLI 2019 liet een overall r-indicator van 0,749 zien, wat erop wijst dat, net als in voorgaande jaren, het schriftelijke vervolgdeel van dit onderzoek wat selectiever was dan het face-to-face deel. Ook de conditionele partiële R-indicatoren bevestigden dat beeld. Voor meer specifieke informatie over de non-responsanalyse CV en SLI in 2019 wordt verwezen naar de eerdere rapportage (Coumans, 2020) hierover en Tabel 5.3.4.

Tabel 5.3.4: R-indicatoren en conditionele partiële R-indicatoren voor CV en SLI, 2019

R-indicator	CV		Signif.	SLI		Signif.
	R-indicator	95%-BI		R-indicator	95%-BI	
	0,765	0,725-0,805		0,749	0,710-0,789	
	Conditionele partiële R-indicator			Conditionele partiële R-indicator		
Geslacht	0,008	- 0,027 - 0,043	NS	0,006	- 0,030 - 0,048	NS
Leeftijd ¹⁾	0,037	0,017 - 0,058	*	0,025	0,016 - 0,059	*
Migratie-achtergrond	0,050	0,030 - 0,071	*	0,056	0,033 - 0,072	*
Burg. staat	0,011	- 0,011 - 0,032	NS	0,026	-0,009 - 0,034	NS
Landsdeel	0,011	- 0,043- 0,064	NS	0,021	0,017 - 0,060	*
Stedelijkheid gemeente	0,015	- 0,007 - 0,036	NS	0,018	-0,009 - 0,035	NS
Gest.huishoud -inkomen	0,034	0,014 - 0,055	*	0,047	0,005 - 0,048	*

1) De indeling in leeftijdscategorieën is ingedikt tot 7 categorieën.

In 2020-2021 was de R-Indicator van CV weer hoger, namelijk 0,819. Kennelijk is de respons in het algemeen wat minder selectief dan in voorgaande jaren. Wel blijkt, gekeken naar de partiële r-indicatoren dat op geslacht na alle onderzochte achtergrondkenmerken (significant) selectief zijn. Het meest selectief was de respons naar migratieachtergrond, leeftijd en burgerlijke staat, maar ook naar landsdeel, verstedelijking van de woongemeente en huishoudensinkomen is er enige selectiviteit. Voor huishoudensinkomen was deze net significant. De coëfficiënten van de selectieve kenmerken (hier niet gepresenteerd), die iets zeggen over de categorieën die het meest verantwoordelijk zijn voor de selectie, laten een vergelijkbaar beeld zien dan de eerder gepresenteerde resultaten in deze paragraaf. Het gaat daarbij vooral om de 76 tot 75-jarigen, mensen met een Nederlandse achtergrond. Zij repondeerden beter, zo bleek al eerder in deze paragraaf. Mensen in de jongste leeftijdsgroepen 35 jaar, met een migratieachtergrond, ongehuwden, mensen in het westen van het land en in zeer sterk verstedelijkte gemeentes en de laagste inkomensgroep nemen juist relatief minder deel aan het onderzoek (zie ook Tabel 5.3.5).

Voor SLI 2020-2022 was de R-indicator een fractie lager dan CV, namelijk 0,813. De partiële indicatoren laten een vergelijkbaar beeld zien als bij CV, namelijk dat met name migratieachtergrond, landsdeel en leeftijd een selectieve respons laten zien. Ook hier zijn op geslacht na alle andere indicatoren significant. De partiële indicator voor inkomen is hier relatief iets hoger dan bij CV. Dit betekent dat in het schriftelijke vervolgdeel de respons wat selectiever is geworden naar inkomen. Specifiek voor de categorieën van de achtergrondvariabelen (niet gepresenteerd) blijkt het wederom te gaan om de 65 tot 75 jarigen en de mensen zonder migratieachtergrond die de minste selectieve respons laten zien, terwijl het de ongehuwden, mensen met een migratieachtergrond, de laagste twee leeftijdscategorieën (18 tot 25 en 15 tot 35 jaar), mensen in het westen des lands en de laagste inkomensgroep de hoogste coëfficiënten lieten zien. Dit stemt overeen met de categorieën waarvan eerder in deze paragraaf van werd vastgesteld dat zij het minst deelnamen.

Tabel 5.3.5: R-indicatoren en conditionele partiële R-indicatoren voor CV en SLI, 2020-2022

R-indicator	CV		SLI			
		95%-BI		95%-BI		
	0,819	0,787-0,851	0,813	0,782-0,845		
	Conditionele partiële R-indicator		Signif.	Conditionele partiële R-indicator	Signif.	
Geslacht	0,009	- 0,010 - 0,029	ns	0,006	- 0,021 - 0,033	ns
Leeftijd ¹⁾	0,028	0,011 - 0,045	*	0,034	0,017 - 0,051	*
Migratieachtergrond	0,032	0,017 - 0,047	*	0,030	0,016 - 0,044	*
Burg. staat	0,023	0,006 - 0,039	*	0,020	0,005 - 0,036	*
Landsdeel	0,019	0,002- 0,036	*	0,026	0,009 - 0,043	*
Stedelijkheid gemeente	0,020	0,003 - 0,036	*	0,017	0,001 - 0,033	*
Gest.huishoud.inkomen	0,016	0,000 - 0,033	*	0,018	0,002 - 0,034	*

1) De indeling in leeftijdscategorieën is ingedikt tot 7 categorieën.

5.4 Weging

Om te corrigeren voor de selectiviteit in de respons, wordt een weging toegepast. De weging voor CV en SLI is gebaseerd op de volgende variabelen in het weegmodel: geslacht, leeftijd (in categorieën), burgerlijke staat, provincie, mate van verstedelijking van de woongemeente, migratieachtergrond, type huishouden, als werkzoekend geregistreerd bij CWI en huishoudensinkomen. Zie voor verdere uitleg over de weging hoofdstuk 7 van dit rapport.

6. Plausibiliteitsanalyse

Er is net als in voorgaande jaren een beperkte analyse op de plausibiliteit van de resultaten uitgevoerd. De uitkomsten voor enkele inhoudelijke variabelen van CV en van SLI zijn voor het opleveren van de eindbestanden vergeleken met die van eerdere edities van 2014 en 2016. Ook zijn deze inhoudelijke variabelen met elkaar vergeleken voor en na weging. Deze analyse leverde geen inplausibele resultaten op.

De beoordeling van de plausibiliteit wordt verder overgelaten aan de inhoudelijk deskundigen bij het SCP.

7. Weging

7.1 Wegen

Om te corrigeren voor de selectiviteit in de respons, wordt een weging toegepast. Tegelijkertijd vindt er ook een ophoging plaats naar de populatie-aantallen van CV (16+) en SLI (18+).

Eerst wordt het CV deel gewogen. Er wordt een startgewicht gebruikt, dat bestaat uit N/n (voor 2019: 14.100.500/1.293; voor 2020-2022: 14.321.412/1.281).

De randtotalen voor de weging zijn berekend uit hetzelfde bestand waarmee het onderzoek verrijkt is, namelijk een selectie van variabelen en personen van 16 jaar en ouder in Nederland woonachtig op 01-01-2019 respectievelijk 01-01-2021.

De weging is uitgevoerd met (combinaties van) variabelen zoals in onderstaand overzicht.

De weging levert dan per persoon een Correctiegewicht en een Eindgewicht op. Het Startgewicht en het Correctiegewicht leveren tesamen het Eindgewicht op. Het Eindgewicht wordt Weegpop_CV genoemd en geeft zowel de weging als de ophoging weer, telt dus op tot het populatie-aantal (NL 16+ op 01-01-2019 respectievelijk 01-01-2021). Er wordt een factor Weeg_CV gemaakt, door het Eindgewicht te corrigeren met (n/som Eindgewichten). Dat geeft alleen de weging weer en telt op tot het aantal respondenten.

Vervolgens wordt voor de SLI respondenten de weging uitgevoerd. Als Startgewicht wordt per SLI respondent zijn CV Eindgewicht gebruikt gecorrigeerd met een factor, die weer bestaat uit de som van de CV Eindgewichten van alle SLI respondenten, gedeeld door het aantal SLI respondenten (voor 2019: 13.689.869/1.058 en voor 2020-2022: 13.918.236/1.041). De weging wordt verder op een vergelijkbare manier uitgevoerd als bij CV en levert een Weegpop_SLI (NL 18+ op respectievelijk 01-01-2019 en 01-01-2021) en een Weeg_SLI op.

5.6.1 Weegmodel, editie 2019 en 2020-2022

CV 16+	SLI 18+
geslacht_x_lft11	geslacht_x_lft10
geslacht_x_burgstat	geslacht_x_burgstat
geslacht_x_prov	geslacht_x_prov
stedelijkheid_x_lft11	stedelijkheid_x_lft10
herkomst3	herkomst3
typehh	typehh
cwiz	cwiz
huishoudinkomen	huishoudinkomen

5.6.2 Verslagbestand weging

	CV 16+		SLI 18+	
	2019	2020-2022	2019	2020-2022
Sample size	1 293	1 281	1 058	1 041
Correction weights information				
Minimum weight	0.12	0.17	0.51	0.21
Median	0.948	0.934	0.962	0.955
First quartile	0.733	0.685	0.841	0.852
Third quartile	1.226	1.248	1.095	1.085
Maximum weight	2.679	3.026	1.753	2.128
Standard deviation	0.374	0.435	0.189	0.202
Number of negative weights	0	0	0	0
Final weights information				
Minimum weight	1 301	1 915	2 380	1 469
Median	10 334	10 436	11 604	11 955
First quartile	7 992	7 654	8 691	8 556
Third quartile	13 369	13 956	15 689	16 486
Maximum weight	29 219	33 831	43 508	52 449
Standard deviation	4 083	4 865	6 060	7 131
Number of negative weights	0	0	0	0

8. Conclusies kwaliteitsanalyse

Het responspercentage in 2019 was voor CV 57,9 procent en voor SLI 49,3 procent. In de periode 2020-2022, de periode van de Coronapandemie en bijbehorende lockdowns en andere maatregelen, lagen deze percentages met respectievelijk 37,6 en 31,6 procent beduidend lager.

Toch kan op basis van de kwaliteitsanalyse, en dan met name de responsanalyse en bijbehorende R-indicatoren, geconcludeerd worden dat de representativiteit van de steekproef voldoende is. De geconstateerde scheefheid op de variabelen, zoals leeftijd, migratieachtergrond, landsdeel en inkomen kan worden gecorrigeerd met de weging. Daar zijn dezelfde variabelen in opgenomen.

9. Opgeleverde Producten

9.1 Microbestanden

De volgende eindbestanden zijn geleverd:

Naam bestand	Data	Meta	Records	Variabelen
CV_SLI_2019_SCP_2020309	*.sav	*.sps	1293	396
VNL_KvL_2020_2022_SCP	*.sav	*.sps	1281	397

9.2 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 CCF-bestand. De inhoud van dat bestand zal nog met het SCP.

10. Referenties

1. Budé, M. Steekproefontwerp Culturele Veranderingen en SCP Leefsituatie Index 2019/2021, CBS Heerlen, september 2018 t/m 2020;
2. Knops, J. Onderzoeksdesign Culturele Veranderingen en SCP Leefsituatie Index 2019/2020, , oktober 2018 t/m 2021;
3. Budé, M. Steekproefverantwoording Culturele Veranderingen en SCP Leefsituatie Index 2019, maart 2020;
4. Budé, M. Steekproefverantwoording Culturele Veranderingen en SCP Leefsituatie Index 2020, 2021,2022, augustus 2020;
5. Coumans, A.M. Responsanalyse Culturele Veranderingen/ SCP Leefsituatie Index (CV/SLI), juni 2020.
6. Cobben en Schouten (2008). An empirical validation of R-indicators. CBS, Voorburg/Heerlen.

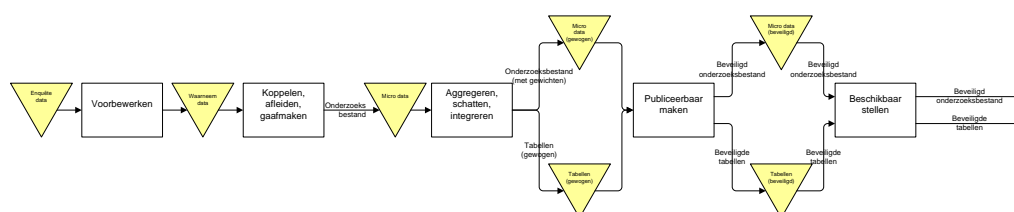
11. Bijlage I – Algemene beschrijving dataprocessing

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, plausibele statistische informatie te kunnen ontleen. De verwerking van de data is zoveel mogelijk geautomatiseerd. Het verwerken is uitgevoerd door de Sector Arbeid, Inkomen en Leefsituatie (SAL) van de Divisie Sociaal-economische en ruimtelijke statistieken (SER) van het CBS.

Inleiding

Het onderstaande figuur 1 geeft het verwerkingsproces van de onderzoeksdata in hoofdlijnen aan. De vijf deelprocessen (weergegeven in de vierkanten blokken) worden in hierna in detail beschreven.

Figuur 1: Procesmodel Verwerkingsproces



Voorbewerken: Van enquêtedata naar waarneemdata

Voordat enquêtedata geschikt is om te verwerken zijn er een aantal activiteiten nodig als voorbereiding. Kort gezegd gaat het erom de juiste data (qua cases en variabelen) in het juiste formaat beschikbaar te maken. Dit resulteert in waarneemdata. De waarneemdata is de grondstof voor het 'echte' verwerken.

Binnen het subproces "Voorbewerken" kunnen de volgende processtappen worden onderkend:

1.1. Controleren enquêtedata: range en routing controle

Met de range- en routingcontrole wordt gekeken of het enquêtebestand voldoet aan de eisen voor de verdere verwerking van het bestand.

Het enquêtebestand wordt hier gedefinieerd als:

- alle data uit een vragenlijst;
- van een of meerdere respondenten;
- van een specifieke mode.

De range- en routingcontrole is specifiek:

- per onderzoek;
- per mode;
- per versie van de vragenlijst.

De eisen van de controles zijn overwegend gebaseerd op de definitie van de vragenlijsten zoals vastgelegd in de gecompileerde vragenlijsten en op het ontwerp van het uniforme datamodel.

Range controles zijn o.a.:

- formaat van variabele;
- variabele heeft juiste naam;
- waardebereik van variabele conform afspraak;

- dubbele sleutels.

Routing controles 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.

1.2 Uniformeren

De recordstructuur van enquêtedata kan mode specifiek zijn. Dit komt omdat de vragenlijst mode specifiek is. Daarmee is ook de metadata mode specifiek.

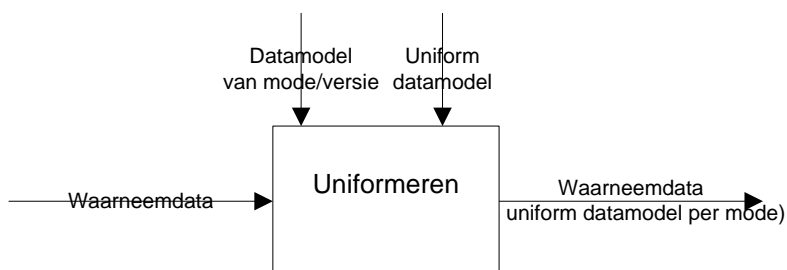
Het kan ook voorkomen dat gedurende het onderzoek er voor een mode meerdere versies van een vragenlijst worden gebruikt (eventueel ook met een verschil in variabelen). Dus ook per versie kunnen er verschillen qua recordstructuur zijn.

Doel bij uniformeren is om tot één uniforme recordstructuur te komen. De uniforme recordstructuur is onderzoekspecifiek en geldt in principe voor de duur van het hele onderzoek. Het bepalen van de uniforme recordstructuur¹ en het beschrijven van de bijbehorende metadata is een ontwerpactiviteit. Dit betekent dat de metadata van het uniforme datamodel vooraf gedefinieerd is; dus voordat het uniformeren daadwerkelijk plaats vindt. Bij het uniformeren van de data hoeft er in principe dus geen metadata meer te worden aangepast; als je de data in het uniforme model zet is de metadata automatisch correct.

Pre-conditie:

- het mode-specifieke (en vragenlijstversie specifieke) datamodel moet bekend zijn;
- het uniforme datamodel moet bekend zijn.

Figuur 2: Procesmodel Uniformeren

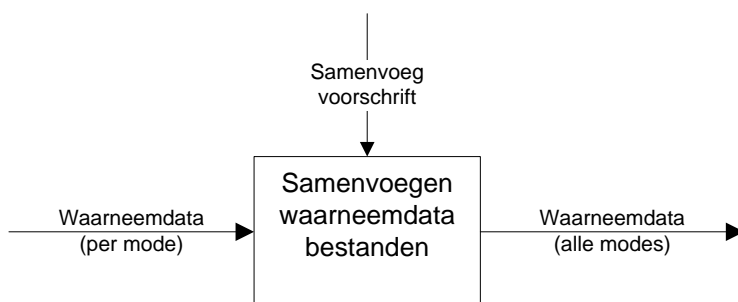


1.3 Controleren enquêtedata: range en routing controle

De waarneemdatabestanden van de diverse modes worden samengevoegd tot één fysiek bestand. Per case moet wel duidelijk blijven uit welke mode het record komt.

¹ en trouwens ook de mode-specifieke datamodellen

Figuur 3: Procesmodel Samenvoegen waarneemdata



2. Koppelen, afleiden en gaafmaken

Bij het koppelen, afleiden en gaafmaken wordt de data verrijkt met o.a. data uit de steekproef, registerdata en andere bronnen. Tevens wordt de respons afgebakend, vindt imputeren en gaafmaken plaats en worden variabelen afgeleid (inclusief coderen). Dit resulteert in een onderzoeksbestand.

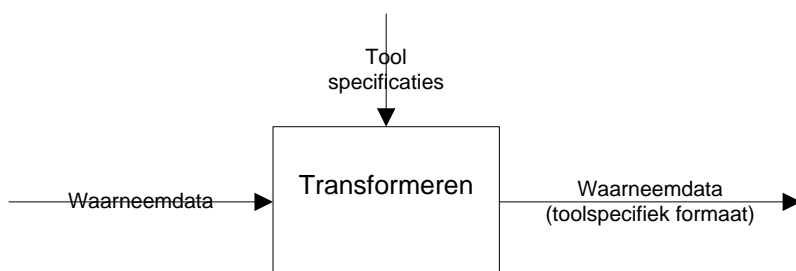
2.1 Transformeren

Doel van deze procesactiviteit is de waarneemdata transformeren naar een vorm die gebruikt kan worden bij de vervolg activiteiten. De waarneemdata is in een bepaald technisch formaat, in dit geval Blaise. Bij de vervolgstappen van het verwerkingsproces wordt SPSS als verwerkingstool gebruikt. Daartoe moet het technische formaat van de data worden aangepast (van Blaise via ASCII naar SPSS).

T.b.v. SPSS is transformeren bijvoorbeeld:

- Dichitomiseren (men werkt in het onderzoeksbestand niet met meervoudige antwoorden).
- Labels aanbrengen: Op basis van de metadata uit de vragenlijst worden variabelen en value labels gegenereerd, die worden gecombineerd met dit SPSS-systeembestand. Deze labels vormen de beschrijving van de data.
- Imputatie routing: Respondenten hoeven in de vragenlijst alleen die vragen te beantwoorden die op hun situatie van toepassing zijn. Vragen die door de respondent niet beantwoord hoefden te worden, gaan als blanco naar ASCII en krijgen vervolgens in SPSS de waarde SYSMIS. Dit betekent dat het SPSS-systeem de waarden als "n.v.t." beschouwd bij de uitvoering van statistische analyses.

Figuur 4: Procesmodel Transformeren



2.2 Verrinnen

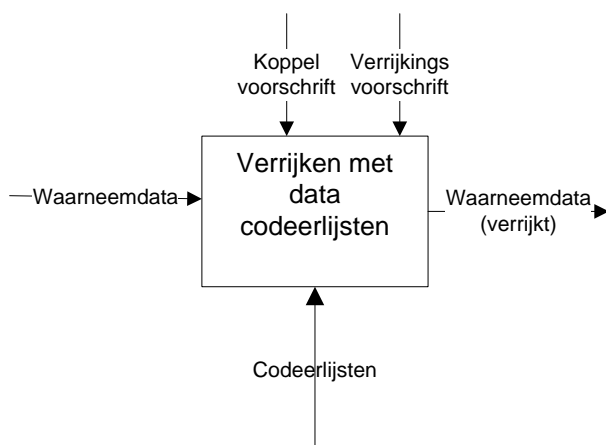
Binnen de verwerkingsprocessen gebeurt “het verrinnen” alleen t.b.v. het koppelen met registerdata en dus niet om met geanonimiseerde data in het verwerkingsproces te werken. Om waarneemdata te kunnen verrijken met registerdata dient de waarneemdata eerst verrind te worden. Dit betekent dat voor iedere persoon in de waarneemdata een betekenisloze identificerende variabele wordt bepaald (genaamd “RINPersoon”). Dit nummer is gebaseerd op data uit de GBA.

De te koppelen waarneemdata wordt daartoe geleverd aan CBK. Deze koppelen de data aan het Centrale Koppelbestand Personen (CKP). Sector CBK levert de verrinde data vervolgens terug aan het verwerkingsproces. Voor het leggen van een koppeling zijn het Burger Service Nummer (BSN) en/of de combinatie van geboortedatum, geslacht en adres nodig. Een geslaagde koppeling betekent in concreto dat aan het originele record RINPersoon en RINPersoonVolgNr uit het CKP worden toegevoegd. Hiermee is de desbetreffende persoon in het CKP identificeerbaar. Naast de CKP persoonidentificatie wordt nog informatie over de koppeling aan het record toegevoegd.

2.3 Verrijken met data codeerlijsten

De waarneemdata wordt hier gekoppeld aan verschillende codeerlijsten. Hier wordt bijvoorbeeld op basis van de viercijferige postcode de gemeentecode bepaald. Op basis van de gemeentecode worden de bovengemeentelijke regionale indelingen bepaald.

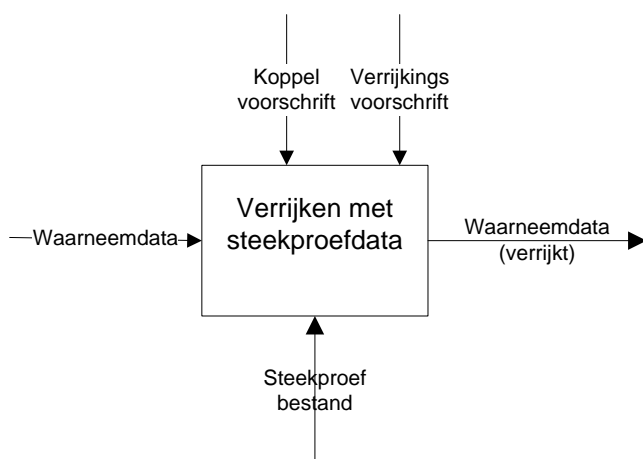
Figuur 5 : Procesmodel Verrijken met data codeerlijsten



2.4 Verrijken met steekproefdata

T.b.v. non-respons analyses kan de waarneemdata worden verrijkt met de complete steekproef met voor elk element de voor uitdunning van de adressensteekproef gebruikte variabelen (alle adresgegevens bijvoorbeeld), het startgewicht en een eindresultaat (bijvoorbeeld: uitgedunde GBA65plus, niet uitgezet door regiomanager, geen woonadres, leegstand, niemand thuis, taalbarrière, weigering, enzovoorts). Op deze manier kan een betere (uitgebreidere) non-respons analyse naar allerlei achtergrondkenmerken gemaakt worden. Bij het koppelen kan het voorkomen dat er een steekproefeenheid is waarvoor (nog) geen waarneemdata is. En tevens waarneemdata waarvoor geen steekproefeenheid is. In het laatste geval is blijkbaar een verkeerde respondent bevroegd.

Figuur 6: Procesmodel Verrijken met steekproefdata

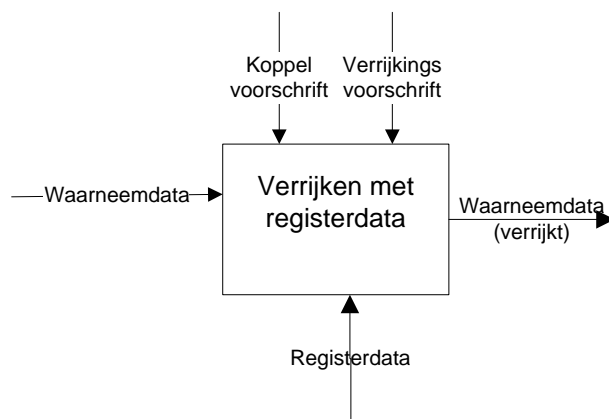


2.5 Verrijken met registerdata

Het koppelen met de registerdata gebeurt op basis van het RIN-nummer. Veel gebruikte registers zijn de GBA, de Polisadministratie en het UWV-Werkbedrijf.

De waarneemdata wordt uit de registers verrijkt met o.a. type huishouden, geboorteland persoon en van diens vader en moeder en afleidingen daarop (GBA), bron inkomen en hoogte inkomen (Polisadministratie), provincie, inschrijfduur (is een afleiding), werkend (UWV-Werkbedrijf). Het verrijken gebeurt niet alleen voor de OP maar ook voor alle vastgestelde personen in het huishouden. Als er een koppeling is, wordt de waarneemdata vervolgens verrijkt met een aantal variabelen uit de registers. Als er geen koppeling is blijven de betreffende register-variabelen leeg.

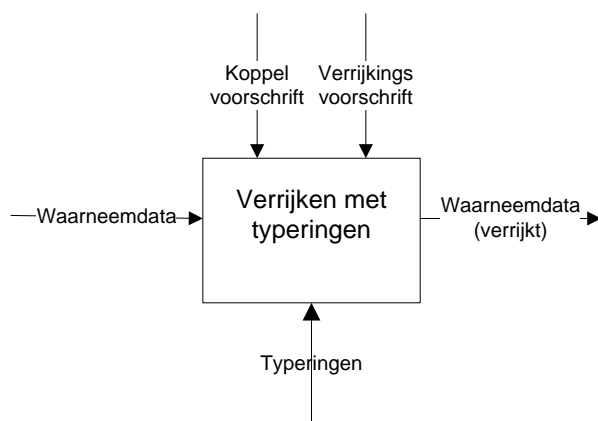
Figuur 7: Procesmodel Verrijken met registerdata



2.6 Verrijken met typering

Deze stap kan pas worden uitgevoerd als de typering beschikbaar zijn. Afhankelijk van het specifieke onderzoek heeft het typeren een bepaalde doorlooptijd, waardoor deze data (meestal) niet meteen beschikbaar is.

Figuur 8: Procesmodel Verrijken met typering

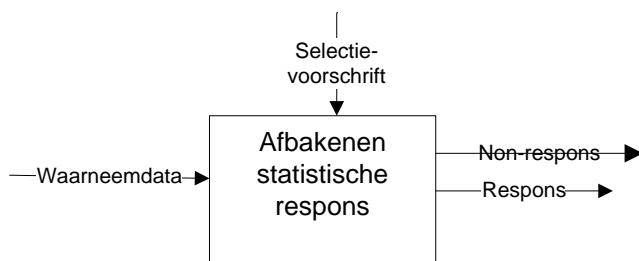


2.7 Afbakenen statistische respons

Alleen statistische respons wordt meegenomen in de verdere verwerking. Wat wel/niet tot respons behoort, staat in een voorschrift. In deze activiteit wordt o.b.v. het voorschrift de respons bepaald. Hóe de non-respons gescheiden wordt van de respons is vanuit logisch oogpunt niet relevant; dit kan bijvoorbeeld door de records fysiek van elkaar te scheiden, maar kan bv ook door met indicatoren te

werken. Voorwaarde is dat de non-respons op een gegeven moment beschikbaar is ten bate van non-respons analyses.

Figuur 9: Procesmodel Afbakenen statistische respons



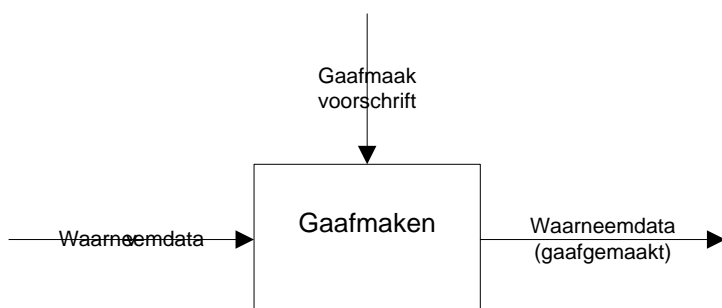
2.8 Gaafmaken (micro)

Gaafmaken is het opsporen en corrigeren van foutieve gegevens in de waarneemdata.

Bij micro gaafmaken vinden zowel de controles als de correcties plaats op microniveau.

Voorbeelden van voorkomende fouten zijn: het geboortjaar klopt niet of is onwaarschijnlijk, een respondent rapporteert in euro's in plaats van in duizenden euro's (of omgekeerd), of de winst van een bedrijf is niet gelijk aan het verschil tussen baten en lasten. Gaafmaken gebeurt op basis van, bij het ontwerp bepaalde, voorschriften.

Figuur 10: Procesmodel Gaafmaken (micro)



2.9 Imputeren

Imputeren is het bepalen en introduceren van een (nieuwe) waarde op een plaats waar een waarde ontbreekt of op 'onbekend' (ontbrekend) is gezet.

Bij enquêtes komt het voor dat respondenten op één of meer vragen geen antwoord geven, terwijl dit wel van ze wordt verlangd. Men spreekt dan van item-nonrespons (of partiële nonrespons) en van (ten onrechte) ontbrekende waarden (missing values). Redenen om een vraag niet te beantwoorden zijn het niet kunnen of willen geven van het antwoord. Op ingewikkelde of moeilijk te begrijpen vragen kan men vaak geen antwoord geven, op gevoelige vragen wil men het dikwijls niet. Ook bij registers kunnen gegevens ontbreken die het CBS wel had willen hebben.

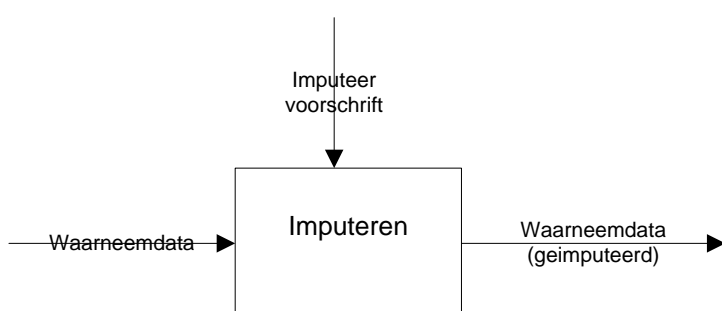
Er zijn een aantal manieren om met ontbrekende waarden om te gaan. Eén daarvan is imputeren van een geldige waarde voor de ontbrekende waarde in het databestand.

Een alternatief voor imputeren is om het achterwege te laten. De ontbrekende waarden blijven dan onbekend. Redenen om te imputeren, in plaats van het veld leeg te laten, zijn:

1. het verkrijgen van een 'volledig' (geheel gevuld) databestand;
2. verhoging van de kwaliteit van het micro-bestand en/of van de parameterschattingen.

We maken verder onderscheid tussen imputeren en afleiden. Bij het afleiden van variabelen worden nieuwe variabelen gecreëerd als functie van in het bestand reeds bestaande variabelen. Bij imputeren worden ontbrekende waarden op een bestaande variabele gecreëerd. Imputeren gebeurt op basis van, bij het ontwerp bepaalde, voorschriften. Steeds dient te worden vastgelegd dat een waarde is geïmputeerd.

Figuur 11: Procesmodel Imputeren



2.10 Afleiden

Met afleiden wordt hier bedoeld het creëren van nieuwe variabelen als functie van in het bestand reeds bestaande variabelen.

Coderen is ook een vorm van afleiden. Het coderen van een vraag is het (keuze)proces waarbij een beslissing wordt genomen om een antwoord te interpreteren in termen van een voor gedefinieerde verzameling mogelijke antwoorden. Een dergelijke keuze wordt, tijdens een interview of bij het invullen van een vragenformulier, vaak gedaan door respondenten, al of niet met hulp van een interviewer. Soms echter wordt deze keuze achteraf gedaan, op het CBS en zonder de aanwezigheid van een respondent, door codeurs of typeurs.

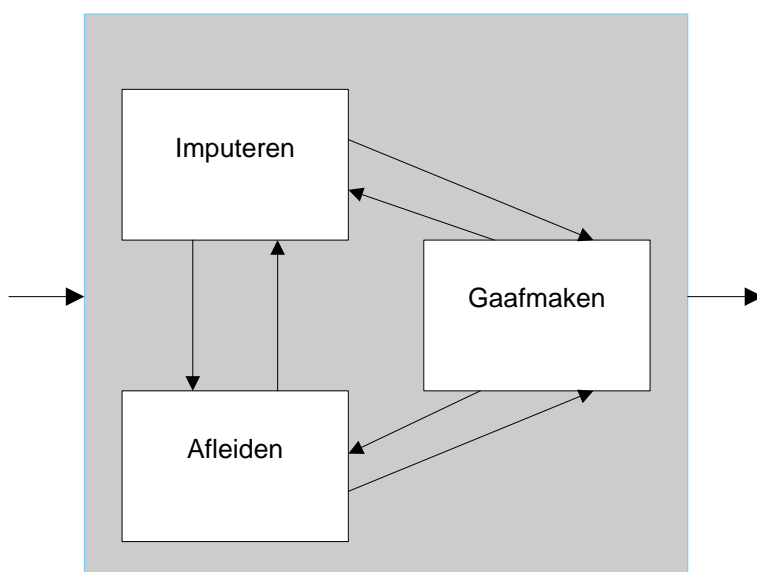
2.11 Gaafmaken, imputeren en afleiden: samenhang

De logische volgorde is eerst gaafmaken, dan imputeren en dan variabelen afleiden.

Echter, gaafmaken, imputeren en afleiden zijn geen activiteiten die voor een dataset en individuele case sequentieel verlopen. De activiteiten worden per variabele of set van variabele doorlopen, waarna de activiteiten voor andere variabelen worden doorlopen.

Eventueel kan de volgorde van de activiteiten ook anders zijn: bv eerst imputeren dan gaafmaken. Dit is afhankelijk van de specifieke regels die van toepassing zijn binnen het onderzoek.

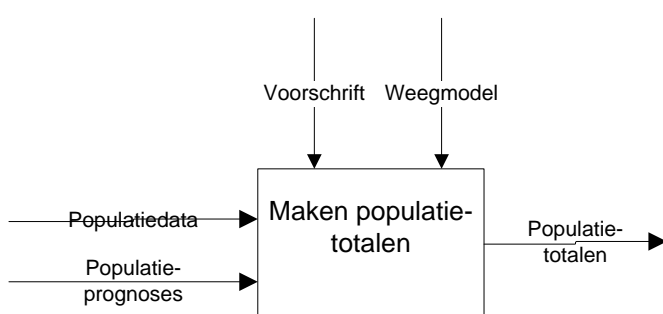
Figuur 12: Procesmodel Gaafmaken, imputeren en afleiden: Samenhang



2.12 Maken Populatietotalen

Voor het wegen zijn populatietotalen nodig. Het kan zijn dat de populatietotalen zelf geschat moeten worden aangezien de data bij de taakgroep Demografie niet altijd voldoende gedetailleerd en soms onvoldoende actueel zijn. Populatietotalen worden bepaald op het totale register, niet op de met register verrijkte waarneemdata.

Figuur 13: Procesmodel maken populatietotalen



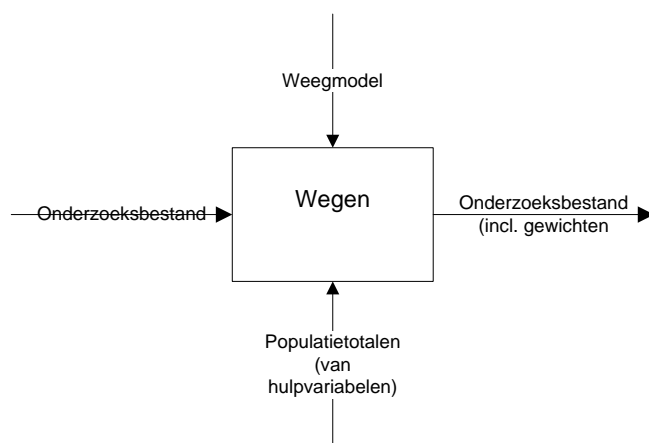
2.13 Wegen

Het wegen is de activiteit waarbij weegfactoren worden bepaald. In het weegmodel staat beschreven hoe het wegen moet plaatsvinden.

Bij het wegen wordt de verdeling van variabelen in de steekproef in overeenstemming gebracht met de verdeling daarvan in de populatie. Daartoe wordt aan iedere case een gewicht toegekend. Ten bate van het wegen zijn populatietotalen (hulpvariabelen) nodig op persoonsniveau. Tevens kan

gewogen worden naar verdelingen waaraan het responsproces idealiter moet voldoen, bijvoorbeeld: elke dag evenveel respons of een gelijk responspercentage per mode. De weging resulteert in één of meerdere ophoogfactoren (afhankelijk van het aantal entiteiten waarvoor gewogen wordt).

Figuur 14: Procesmodel Wegen



3. Publiceerbaar maken

Bij het publiceerbaar maken worden de tabellen en het gewogen onderzoeksbestand statistisch beveiligd. Binnen het subproces “publiceerbaar maken” kunnen de hieronder genoemde processtappen plaatsvinden.

3.1 Maken micro output

Op basis van het onderzoeksbestand wordt de output (microbestanden) voor bijvoorbeeld de externe klant, SAL-SAD/SIL, CvB, DANS en/of Eurostat gemaakt.

De output voor de diverse afnemers bevatten meestal slechts een deelverzameling van de variabelen uit het onderzoeksbestand. Deze deelverzamelingen voor de verschillende afnemers worden in deze stap gemaakt.

SAL-SAD/SIL krijgen meestal het gehele onderzoeksbestand ten behoeve van het maken van publicaties. Hiervoor hoeven geen extra stappen gezet te worden.

3.2 Statistisch beveiligen microdata

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

3.3 Statistisch beveiligen standaardtabellen

De tabellen (t.b.v. de Externe klant, Ministeries, Statline, CvB, DANS en/of Eurostat) dienen ook statistisch beveiligd te worden.

4. Beschikbaar stellen

De statistisch beveiligde producten worden vervolgens als output geleverd (beschikbaar gesteld) aan diverse belanghebbenden, waaronder externe opdrachtgevers, verschillende ministeries, Eurostat, DANS, Centrum voor Beleidsstatistiek. Binnen het subproces "Beschikbaar stellen" kunnen de hieronder genoemde processtappen plaatsvinden.

4.1 Leveren microbestanden

Dit is de activiteit waarbij microdata daadwerkelijk wordt geleverd. Aan welke partijen geleverd wordt is afhankelijk van het betreffende onderzoek.

Er worden voor het onderzoek CV twee bestanden opgeleverd: een responsbestand met data voor de ene vragenlijstvariant en een responsbestand met data voor de andere vragenlijstvariant.

4.2 Leveren standaardtabellen

Dit is de activiteit waarbij de tabellen daadwerkelijk worden geleverd. Aan welke partijen geleverd wordt is afhankelijk van het betreffende onderzoek.