



Centraal Bureau
voor de Statistiek

Rapport

MBO werkgeversonderzoek 2017

Onderzoeksdocumentatie en kwaliteitsanalyse

Sociaal-economische en Ruimtelijke Statistieken | Dataverzameling
Maart 2018

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
1.2	Definities, acroniemen en afkortingen	5
2.	Doel, achtergrond en geschiedenis van het onderzoek.....	6
2.1	Doel	6
2.2	Geschiedenis van het onderzoek	6
3.	Dataverzameling	7
3.1	Randvoorwaarden	7
3.2	Benaderingsstrategie	7
3.3	Communicatiestrategie	8
3.4	Vragenlijstontwikkeling	9
3.5	Steekproef	10
3.6	Veldwerk	11
4.	Verwerken (data processing)	14
4.1	Inleiding	14
4.2	Typeren	14
4.3	Verrijken	15
4.4	Afleiden	15
5.	Non-responsanalyse	16
5.1	Afbakenen statistische respons en nonresponsanalyse	16
5.2	Selectiviteit van de respons naar achtergrondkenmerken	16
5.3	Weging	17
6.	Andere kwaliteitsindicatoren.....	19
6.1	Missings, geen antwoord op vragen en tellingen	19
6.2	Inconsistenties binnen de vragenlijst en afgeleide variabelen	19
6.3	Afwijkingen achtergrondgegevens en registerdata	19
6.4	Belangrijke opmerkingen van respondenten	20
7.	Plausibiliteitsanalyse	22
7.1	Vergelijking oudere jaren met dit onderzoek	22
7.2	Vergelijking schoolverlatersonderzoek	24
8.	Conclusie kwaliteitsanalyse	26
8.1	Conclusies	26
8.2	Aanbevelingen	27
9.	Opgeleverde Producten	28

9.1	Microbestanden	28
9.2	Statistische beveiliging	28
10.	Referenties	29
11.	Bijlage I – Algemene beschrijving dataprocessing	30

1. Inleiding

1.1 Doel van dit document

In dit document wordt verslag gedaan van het mbo werkgeversonderzoek 2017. De voorbereiding, dataverzameling en verwerking van de gegevens worden beschreven. Daarnaast dient dit document als naslagwerk voor onderzoekers die meer inzicht wensen in de achtergronden van het onderzoek en in het tot stand komen van de analysebestanden. Hoofdstuk 2 beschrijft het doel, achtergrond en geschiedenis van het onderzoek. Vervolgens gaat hoofdstuk 3 in op de dataverzameling en komen in hoofdstuk 4 de uitzonderingen in de verwerkingsstraat die voor dit onderzoek zijn toegepast aan de orde. De standaard verwerkingsstraat is als bijlage toegevoegd.

Tevens gaat dit document in op enkele kwaliteitsaspecten van het onderzoek. In hoofdstuk 5 komt de non-responsanalyse aan de orde. Deze bestaat uit een korte beschrijving van de respons naar achtergrondkenmerken. Vervolgens gaat hoofdstuk 6 in op andere kwaliteitscriteria, zoals gegevens over het aantal keer dat geen antwoord is gegeven en eventuele inconsistenties, en hoofdstuk 7 op de plausibiliteit. Tot slot volgen enkele conclusies en aanbevelingen (hoofdstuk 8).

1.2 Definities, acroniemen en afkortingen

Begrip	Omschrijving
ABR	Algemeen Bedrijven Register
BE	Bedrijfseenheid
BRC	Beroepenclassificatie ROA CBS
CAWI	Computer Assisted Web Interviewing
CBS	Centraal Bureau voor de Statistiek
CCI	Contact Centre Inbound
DVS	CBS Vragenlijsten en Steekproeven
ISCO	International Standard Classification of Occupations
MWO	Mbo werkgeversonderzoek
OCW	Ministerie van Onderwijs, Cultuur en Wetenschap
RA	Remote Access
SAL	Sector Arbeid, Inkomen en Leefsituatie
SBB	Samenwerkingsorganisatie Beroepsopleiding Bedrijfsleven
SBI	Standaard Bedrijfsindeling
SER	Sociaal-economische en ruimtelijke statistieken
SVO	Schoolverlatersonderzoek

2. Doel, achtergrond en geschiedenis van het onderzoek

2.1 Doel

Het doel van het onderzoek is het leveren van informatie over de tevredenheid van werkgevers over de aansluiting van de door mbo'ers gevolgde opleiding en de functie die ze binnen het bedrijf uitoefenen. Het onderzoek wordt uitgevoerd onder bedrijven waar één of meer recent gediplomeerde mbo'ers werkzaam zijn. Deze personen moeten minimaal zes maanden in dienst zijn.

Het onderzoek wordt uitgevoerd in opdracht van Samenwerkingsorganisatie Beroepsopleiding Bedrijfsleven (SBB).

2.2 Geschiedenis van het onderzoek

In 2015 is een motie ingediend om een indicator voor de tevredenheid van werkgevers met de aansluiting van het onderwijs op de arbeidsmarkt op te nemen in de indicatorentabel van de rijksbegroting. SBB heeft in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW) in 2016 een eerste onderzoek uitgevoerd naar de tevredenheid van werkgevers over de aansluiting van opleidingen op de arbeidsmarkt. [1] Toen bestond de onderzoekspopulatie uit de door SBB erkende leerbedrijven. In deze tweede editie van het onderzoek zijn niet alleen de leerbedrijven benaderd maar ook de niet-bedrijven.

In een vooronderzoek [2] is nagegaan of gebruik gemaakt zou kunnen worden van één steekproefkader. Hiervoor is de door SBB aangeleverde informatie over leerbedrijven gekoppeld aan het Algemeen Bedrijven Register (ABR) van het CBS. Conclusie was dat het koppelrendement voldoende is en dat voor het MWO gebruik gemaakt zou kunnen worden van één steekproefkader op basis van het ABR, waarin de door SBB erkende leerbedrijven worden gevraagd. Daarnaast is onderzocht of door de koppeling van de onderwijsnummerbestanden en de Polisadministratie het mogelijk zou zijn om alleen de bedrijven te selecteren waar recent gediplomeerde mbo'ers in dienst zijn. Conclusie daarvan was dat dit mogelijk is en dat voor het MWO de steekproefomvang beperkt zou kunnen worden door bedrijven die buiten de onderzoekspopulatie vallen uit te sluiten.

3. Dataverzameling

3.1 Randvoorwaarden

Bij het ontwerp van het onderzoeksdesign is uitgegaan van de volgende randvoorwaarden. De projectovereenkomst mbo werkgeversonderzoek 2017 [3] en de bijbehorende kostenindicatie MWO 2017 [4] zijn hiervoor als uitgangspunt gebruikt.

1. Het databestand moet in maart 2018 via Remote Access voor analyse beschikbaar zijn. De waarneming vindt plaats in de periode van 9 oktober tot en met 30 november 2017.
2. Het onderzoek moet gegevens opleveren over de volgende doelgroepen:
 - a. door SBB erkende leerbedrijven;
 - b. niet-leerbedrijven.

De doelpopulatie bestaat uit bedrijven in Nederland waar één of meer recent gediplomeerde mbo'ers werkzaam zijn. Recent gediplomeerd betekent dat ze in studiejaar 2014/'15 of 2015/'16 hun diploma hebben behaald. Deze mbo'ers moeten op het moment van enquêteren minimaal zes maanden in dienst zijn bij het bedrijf.
3. De bedrijven worden via internet (cawi) benaderd.
4. Het aantal te behalen responsen is 3 000, waarvan 1 500 leerbedrijven en 1 500 niet-leerbedrijven. Hierbij gaat het om bedrijfseenheden (BE), zoals afgebakend in het ABR.
5. Voor de te trekken steekproef wordt uitgegaan van een respons van 40 procent. Verwachting is dat de respons onder leerbedrijven en niet-bedrijven gelijk zal zijn. De getrokken omvang is daarom voor beide groepen ook gelijk, 3.750 leerbedrijven en 3.750 niet-leerbedrijven. Bij het trekken van de steekproef wordt naar SBI gestratificeerd. Er wordt niet naar grootteklasse gestratificeerd.
6. Om te corrigeren voor non-respons zal een weegmodel worden ontwikkeld.
7. Voor de vragenlijst wordt aangesloten bij de vragenlijst zoals die eerder door SBB is gebruikt.
8. In de vragenlijst wordt de standaardvraagstelling van het CBS voor het beroep dat door recent gediplomeerde mbo'ers wordt uitgeoefend opgenomen. De open antwoorden worden getypeerd op 2 digits van de BRC.
9. De gemiddelde duur voor het volledig invullen van de internetvragenlijst bedraagt naar verwachting 5 minuten.
10. Aan ieder te benaderen bedrijf wordt een aanschrijfbrief verstuurd met daarin het verzoek om deel te nemen aan het onderzoek. De aanschrijfbrief wordt door CBS opgesteld, waarbij input van SBB gevraagd wordt.
11. SBB ontvangt gedurende de veldwerkperiode maandelijks een voortgangsoverzicht met daarin het aantal uitgezette steekproefelementen en de respons per groep en totaal.

3.2 Benaderingsstrategie

Steekproefbedrijven kunnen via internet (cawi) deelnemen aan het onderzoek. De volgende strategie is daarbij gehanteerd:

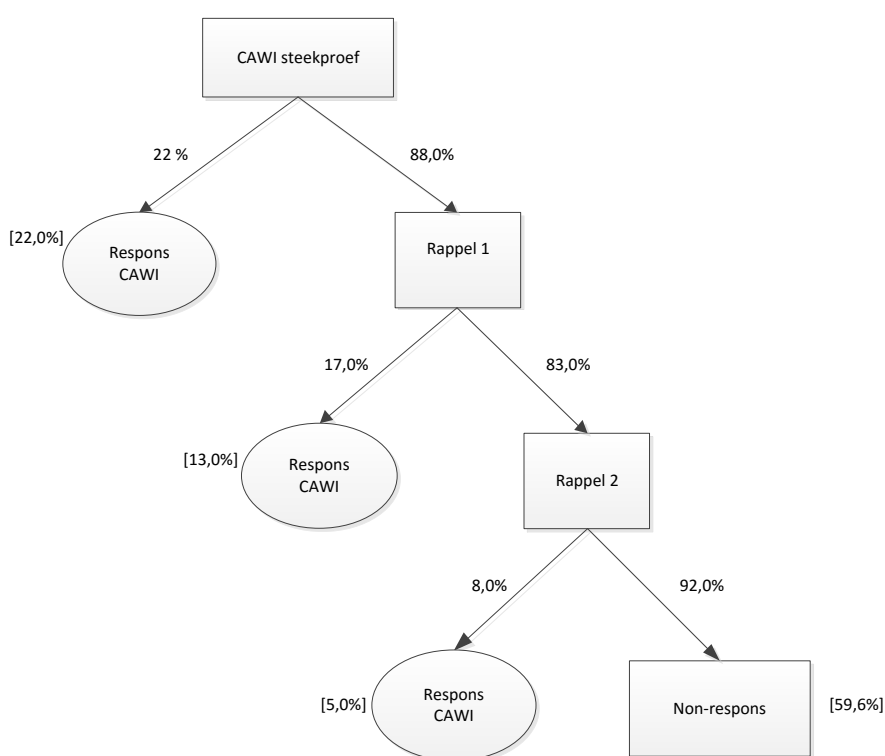
- Steekproefbedrijven ontvangen een aanschrijfbrief met daarin het verzoek om via internet deel te nemen aan het onderzoek en de bijbehorende inloggegevens. Daarnaast wordt in de brief ingegaan op het doel van het onderzoek. Voor de leerbedrijven is de bij SBB bekende contactpersoon aangeschreven, voor de niet-leerbedrijven wordt de standaard contactpersoon uit het ABR aangeschreven.
- Vier weken daarna wordt een eerste rappelbrief verstuurd met daarin opnieuw het verzoek om via internet te reageren. Deze brief wordt alleen verstuurd aan bedrijven waarvan geen

respons is ontvangen via internet en die niet via het Contact Centre Inbound (CCI) hebben aangegeven niet te willen of kunnen deelnemen aan het onderzoek.

- Drie weken na de eerste rappelbrief wordt een tweede rappelbrief verstuurd met opnieuw het verzoek om via internet te reageren. Deze brief wordt alleen verstuurd aan bedrijven waarvan nog geen respons is ontvangen via internet en die niet via het CCI hebben aangegeven niet te willen of kunnen deelnemen aan het onderzoek. In de brief staat een uiterste datum tot wanneer de vragenlijst kan worden ingevuld.

In figuur 3.2.1 wordt de benaderstrategie schematisch weergegeven. Per benaderingsstap is de geschatte responskans vermeld, waarbij zowel de responskans is vermeld uitgaande van de betreffende deelsteekproef (zonder vierkante haken) als van de uitgezette cawi-steekproef (tussen vierkante haken).

3.2.1: Benaderstrategie mbo werkgeversonderzoek 2017



3.3 Communicatiestrategie

Voor MWO is een communicatiestrategie opgesteld. Binnen het CBS is het team Berichtgevers Contact (DBC) van de sector Productie Services verantwoordelijk hiervoor. In de communicatie met bedrijven is voor dit onderzoek gebruik gemaakt van brieven en de CBS website.

Brieven

Beide groepen ontvangen dezelfde aanschrijfbrief en rappelbrieven. In de brieven wordt een link naar de vragenlijst met de bijbehorende inloggegevens opgenomen. De algemene URL die op de brieven komt te staan en die de berichtgevers/respondenten moeten intypen is 'antwoord.cbs.nl'. Deze URL komt automatisch op de loginpagina '<https://antwoord.cbs.nl/login/...>' terecht.

Wanneer de berichtgever de juiste gebruikersnaam en wachtwoord invullen vindt een automatische verwijzing naar de vragenlijstomgeving plaats. De gebruiker krijgt dan een andere URL te zien. Deze URL zal dan beginnen met: '<https://onderzoek.cbs.nl/...>'.

In de brief wordt een standaardtekst opgenomen over hoe het CBS omgaat met privacy en de geleverde gegevens.

Website

Voor het onderzoek is een onderzoekspagina op de CSB-website ingericht (www.cbs.nl/mbowerkgeversonderzoek). Voor deze onderzoekspagina zijn de volgende uitgangspunten gehanteerd:

- moet iets extra's bieden naast de brief;
- bevat een link naar de inlogpagina van de vragenlijst;
- moet enthousiasmeren om aan het onderzoek mee te doen.

3.4 Vragenlijstontwikkeling

Uitgangspunten

Voor MWO wordt een elektronische vragenlijst gebruikt voor cawi. Binnen het CBS is het team Vragenlijsten en Steekproeven (DVS) van de sector Waarnemen verantwoordelijk hiervoor. Het vragenlijstontwerp wordt door middel van een review en vragenlijstschemas door de ontwikkelaar afgestemd met SBB. De vragenlijst wordt in Blaise geprogrammeerd, de door CBS ontwikkelde software voor elektronische vragenlijsten.

Voor de vragenlijst zal waar mogelijk worden aangesloten bij de vragenlijst zoals die eerder door SBB is uitgevoerd. Daarnaast dient rekening gehouden te worden met de werkwijze van CBS. Daarbij is het relevant om te noemen dat:

1. Geen informatie wordt uitgevraagd die al via registers beschikbaar is.
2. Voor informatie over beroep zal zoveel mogelijk wordt gewerkt met standaardvraagstellingen. Dit vanwege het geautomatiseerde typeerproces.

Het CBS is verantwoordelijk voor de vragenlijsten. Er vindt echter intensieve afstemming met SBB plaats over de vragenlijst.

Wijzigingen in de vragenlijst

In het steekproefkader zijn de leerbedrijven gevraagd. Omdat niet alle leerbedrijven gekoppeld konden worden met het ABR is de kans aanwezig dat er onder de groep niet-bedrijven toch ook leerbedrijven zitten. Om in beeld te brengen om hoeveel bedrijven het gaat, is in de vragenlijst aan de niet-leerbedrijven de vraag gesteld worden of zij een erkend leerbedrijf zijn. De leerbedrijven krijgen deze vraag niet.

Er is een vraag toegevoegd over het geslacht van de recent gediplomeerde mbo'er. Deze informatie wordt in de vraagteksten gebruikt. In geval van een mannelijke werknemer wordt in de vraagteksten gesproken over hij en zijn, in geval van een vrouwelijke werknemer over zij en haar.

De open vraag naar functie is vervangen door de standaardvraag die het CBS voor beroep hanteert. Daarbij zijn aanvullende vragen opgenomen die als hulpinformatie gebruikt worden voor het typeren van beroep.

De vraag naar bedrijfssector en bedrijfsgrootte is verwijderd, deze informatie is via registers al beschikbaar.

Inhoud vragenlijst

In de vragenlijst zijn de volgende blokken opgenomen:

1. Recent aangenomen mbo'ers
2. Over de recent gediplomeerde mbo'ers
3. Tevredenheid recent gediplomeerde mbo'ers
4. Afsluiting

Testprocedure

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

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

3.5 Steekproef

Binnen het CBS is het team Vragenlijsten en Steekproeven (DVS) van de sector Waarnemen verantwoordelijk hiervoor.

3.5.1 Steekproefkader

Om te komen tot één steekproefkader zijn op basis van de onderwijsnummerbestanden de personen geselecteerd die in studiejaar 2014/'15 of in studiejaar 2015/'16 een mbo diploma hebben behaald¹. Daarbij zijn de personen die geen onderwijs volgden, maar slechts ingeschreven stonden voor het doen van een examen, de extraneï, buiten beschouwing gelaten. Personen die in beide jaren een diploma hebben behaald, worden maar één keer meegeteld, in totaal ging het om 275 duizend unieke personen.

Vervolgens is voor deze groep recent gediplomeerden nagegaan of ze werknemer zijn en zo ja, bij welk bedrijf. Hiervoor is op persoonsniveau een koppeling gemaakt met de Polisadministratie. Voor deze analyse is uitgegaan van de maand april, dat wil zeggen dat personen de hele maand april werkzaam moesten zijn. Van de 275 duizend recent geslaagde mbo'ers waren er 213 duizend die volgens de Polisadministratie gedurende de hele maand april werkzaam waren. In totaal ging het daarbij om ruim 57 duizend bedrijfseenheden. Hiervan bleven er 55,8 duizend over op het peilmoment van de steekproeftrekking.

¹ De gegevens over de gediplomeerden in het studiejaar 2016/'17 komen pas in het voorjaar 2018 beschikbaar en kunnen dus niet mee worden genomen in de selectie.

3.5.1. Afleiding steekproefkader

Recent gediplomeerde mbo'ers	275 000
w.v. werkzaam als werknemer	213 000
Bedrijven met recent gediplomeerde mbo'ers	57 000
Bedrijven met recent gediplomeerde mbo'ers in finale steekproefkader	55 800
w.v. erkende leerbedrijven	35 500
w.v. niet-leerbedrijven	20 300



De aangekoppelde variabelen uit het ABR zijn:

- SbiGecoördineerd (5 digits)
- GkSbsGecoördineerd (grootteklasse, 2 digits)

Zie de verantwoording van het vooronderzoek voor een uitgebreide beschrijving van de gemaakte koppeling en de afbakening van het steekproefkader. [2]

3.5.2 Steekproefontwerp

In het steekproefontwerp is gestratificeerd naar twee groepen: leerbedrijven en niet-leerbedrijven. Voor beide groepen is de responseis 1 500 eenheden. Er is ook gestratificeerd naar SBI2 (de eerste twee digits van SBI); alle SBI2-strata hebben dezelfde steekproeffractie gekregen. De stratificatie naar SBI2 is dus niet gedaan om bepaalde SBI's over of onder te vertegenwoordigen in de steekproef, maar juist om zo'n over- of ondervertegenwoordiging te voorkomen. Er is niet gestratificeerd naar grootteklasse van de bedrijven, omdat dit in combinatie met de stratificatie naar SBI2 tot veel kleine strata zou leiden. De kans om in de steekproef te komen (steekproeffractie) was dus verschillend voor leerbedrijven en niet-leerbedrijven, maar was gelijk voor alle leerbedrijven en gelijk voor alle niet-leerbedrijven.

Gezien de responseis en het verwachte responspercentage van 40 procent bedroeg de initiële steekproef 3 747 leerbedrijven en 3 753 niet-leerbedrijven.

3.6 Veldwerk

3.6.1 Aanpassingen onderzoeksopzet

Uit analyse van de eerste resultaten bleek dat bijna de helft van de bedrijven aangaf geen recent gediplomeerde mbo'er(s) in dienst te hebben, waarmee ze buiten de doelpopulatie vielen. Ook bleek dat een aanzienlijk deel van de niet-leerbedrijven in de vragenlijst zelf aangaven dat ze wel een door SBB erkend leerbedrijf zijn. Er zijn daarom de volgende maatregelen getroffen:

Brieven

In de rappelbrief is expliciet benoemd waarom de bedrijven worden aangeschreven. Namelijk dat op basis van de bij het CBS beschikbare registraties geconcludeerd wordt dat er een recent gediplomeerde mbo'er bij het bedrijf in dienst is.

Vragenlijst

In de vragenlijst is de vraag of bedrijven in de afgelopen twee jaar een mbo'er hebben aangenomen iets aangepast en wordt genoemd dat ze volgens de CBS registraties een recent gediplomeerde mbo'er in dienst hebben.

Steekproef

Om aan de responseis te kunnen voldoen, is een tweede steekproef getrokken, waarbij er rekening mee is gehouden dat

- niet-bedrijven relatief vaak aangeven dat ze toch een erkend leerbedrijf zijn;
- de respons onder niet-bedrijven lager is dan onder leerbedrijven;

- niet-leerbedrijven vaker dan leerbedrijven aangeven dat er geen recent gediplomeerde mbo'er in dienst is.

Deze tweede steekproef bedroeg in totaal 15 001 bedrijven, waarvan 4 998 leerbedrijven en 10 003 niet-leerbedrijven. De totale steekproef bedroeg daarmee in totaal 22 501 elementen.

3.5.2 Aantallen, te benaderen steekproefaantallen en steekproeffracties per groep

Groep	N	n	f
Leerbedrijf	35 479	8 745	0,246
Niet-leerbedrijf	20 341	13 756	0,676
Totaal	55 820	22 501	0,403

3.6.2 Veldwerkresultaten

In deze paragraaf worden de gerealiseerde responsresultaten gepresenteerd. Omdat alleen bedrijven die een recent gediplomeerde mbo'er in dienst hebben tot de doelpopulatie behoren, wordt onderscheid gemaakt tussen responsen en bruikbare responsen.

Gerealiseerde respons totaal

Tabel 3.6.2.1 geeft de gerealiseerde responsresultaten weer voor de totale steekproef en voor beide steekproefporties.

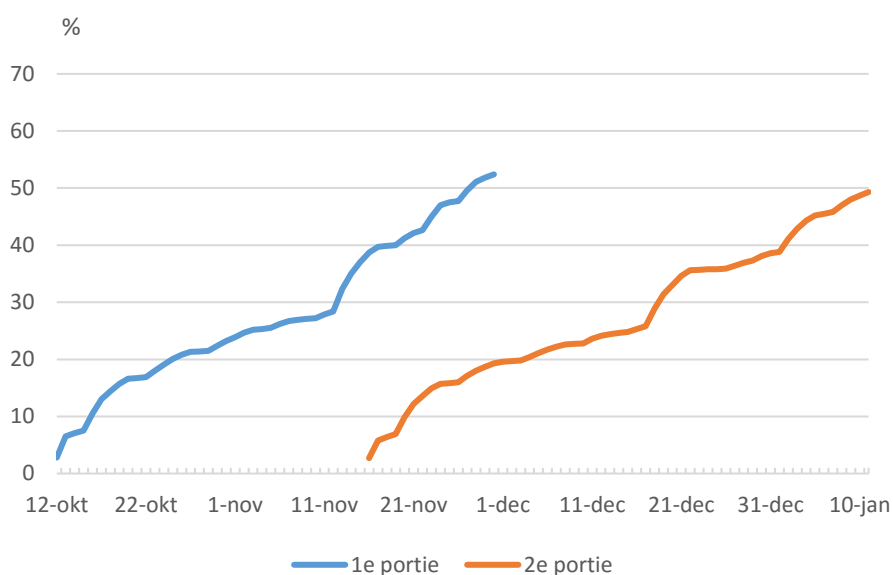
3.6.2.1 Responsresultaten naar steekproefportie en totaal

	Totaal		1 ^e portie		2 ^e portie	
Uitgezette steekproef	22 501		7 500		15 001	
Geen respons	10 721	45,6%	3 425	45,7%	7 296	48,6%
Non-respons	395	1,8%	141	1,9%	254	1,7%
Respons	11 385	50,6 %	3 934	52,5 %	7 451	49,7 %
Bruikbare respons	5 753	25,6%	1 297	17,3%	4 456	29,7%

De totale respons bedraagt 50,6 procent. Dit is ruim 10 procentpunten hoger dan de verwachte respons. De respons in de eerste portie is met 52,5 procent is iets hoger dan in de tweede portie (49,7 procent).

De totale bruikbare respons is 25,6 procent. Dit betekent dat van de bedrijven die gerepsondeerd hebben, ongeveer de helft één of meerdere recent gediplomeerde mbo'ers in dienst heeft. Dit aandeel is in de eerste portie beduidend kleiner dan in de tweede portie. Zo is de bruikbare respons in de eerste portie 17,3 procent, wat er op neer komt dat een derde van de bedrijven die gerepsondeerd hebben één of meerdere mbo'ers in dienst hebben. In de tweede portie is de bruikbare respons 29,7 procent en heeft dus 60 procent van de bedrijven die gerepsondeerd hebben één of meerdere mbo'ers in dienst. Verwachting is dat dit verschil samenhangt met de aanpassingen in de brief en de vragenlijst.

3.6.2.2 Responsontwikkeling (%) in de tijd (dagen) per steekproefportie



Het veldwerk van de eerste portie is gestart op 9 oktober 2017, toen is de aanschrijfbrief verstuurd, en eindigde op 30 november 2017. De rappelbrieven zijn verstuurd op 8 en 20 november. De eerste responsen zijn ontvangen op 12 oktober, wat komt doordat de brieven met de 72-uurspost verstuurd zijn. De tweede portie is op 14 november 2017 gestart en eindigde op 12 januari 2018. De rappelbrieven zijn verstuurd op 14 en 28 december.

Gerealiseerde respons voor leerbedrijven en niet-leerbedrijven

Naast het totale responspercentage en het responspercentage per steekproefportie, worden in tabel 3.6.2.2 de gerealiseerde responsaantallen en -percentages voor de leer- en niet-leerbedrijven gepresenteerd.

3.6.2.2 Responsresultaten naar steekproefportie en totaal

	Totaal		Leerbedrijven		Niet-leerbedrijven	
Uitgezette steekproef	22 501		8 745		13 756	
Geen respons	10 721	45,6%	3 942	45,1%	6 779	49,3%
Non-respons	395	1,8%	143	1,6%	252	1,8%
Respons	11 385	50,6 %	4 660	53,3%	6 725	48,9%
Bruikbare respons	5 753	25,6%	2 722	31,1%	3 031	22,0%

In totaal zijn de responsen van 2 722 leerbedrijven en 3 031 niet-leerbedrijven bruikbaar. Dit betekent dat de beoogde responseis van 1 500 leerbedrijven en 1 500 niet-leerbedrijven ruimschoots is gehaald.

4. Verwerken (data processing)

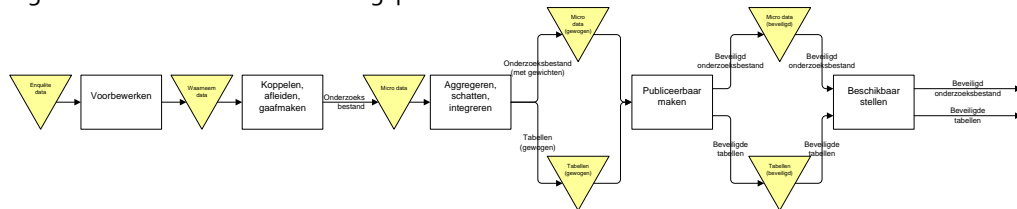
De term dataprocessing wordt gebruikt om de werkzaamheden samen te vatten die nodig zijn om aan de antwoorden die respondenten geven op de vragen in de vragenlijst plausibele statistische informatie te kunnen ontleen. De verwerking van de data is zoveel mogelijk geautomatiseerd.

Het verwerken is uitgevoerd door statistische afdeling dynamisering en verwerking, sector van de Divisie Sociaal-economische en Ruimtelijke Statistiek.

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 x: Procesmodel Verwerkingsproces



Hieronder wordt de verwerking van het mbo werkgeversonderzoek 2017 beschreven en wordt aandacht besteed aan de keuzes die hierin gemaakt zijn.

4.2 Typeren

Om vast te stellen welk onderwijsniveau en –richting respondenten behaald hebben, in welke bedrijfstak zij werken en welk beroep zij uitoefenen, heeft CBS een standaard verwerkingsproces ontwikkeld om bedrijf, beroep en onderwijs te typeren. Dit gebeurt deels automatisch en deels handmatig, en sluit aan op de standaard vraagstellingen van het CBS voor bedrijf, beroep en onderwijs.

Voor het mbo werkgeversonderzoek wordt een steekproef getrokken op basis van het ABR. Dit betekent dat bekend is tot welke bedrijfstak het bedrijf behoort en dat deze informatie niet meer uitgevraagd en getypeerd hoeft te worden. SBB heeft aangegeven te willen weten wat het beroep is van de recent gediplomeerde mbo’er over wie het in de vragenlijst gaat. Hiervoor is in de vragenlijst het standaardblok voor beroep opgenomen. Het typeren van beroep gebeurt deels automatisch en deels handmatig. Het typeerproces maakt gebruik van de informatie die vragenlijst over beroep is verzameld en daarnaast ook van de SBI. Dit resulteert in een ISCO 2008 code, waarvan tevens een BRC 2014 code wordt afgeleid. Voor het mbo werkgeversonderzoek is getypeerd op 2 digits, dat wil zeggen dat de kwaliteit voldoende is voor de ISCO major groups en de BRC beroepsklassen.

In de vragenlijst is ook opgenomen welk mbo niveau de recent gediplomeerde mbo’er heeft gevolgd. De open vraag voor onderwijsrichting wordt daarbij niet gesteld, waardoor het niet nodig is ook onderwijs te typeren.

4.3 Verrijken

Doordat in het ABR aan ieder bedrijf een uniek nummer is toegekend, het zogeheten BEID, is het mogelijk om aan de gegevens uit de enquête informatie uit andere bronnen (registraties) te koppelen. Voor het mbo werkgeversonderzoek is het bestand alleen verrijkt met gegevens uit het ABR die nodig waren voor de steekproef en de weging.

4.4 Afleiden

Het CBS werkt in het verwerkingsproces met afleidingen. Het is efficiënter om in het verwerkingsproces vast te leggen hoe een bepaalde afbakening er uit moet zien zodat niet ieder voor zich dat hoeft te bepalen. Gevolg daarvan is ook dat cijfers eenvoudig te reproduceren zijn en gebruikers van het bestand ook dezelfde afbakening gebruiken en daarmee tot dezelfde uitkomsten.

Voor MWO is een procesmatige afleiding gemaakt om de afbakening van de statistische respons te bepalen (MWOAflRespons) en tien afleidingen voor publicatievariabelen.

- | | |
|------------------------------|---|
| - MWOAflHalfJaarLanger | Is mbo'er een half jaar of langer in dienst? |
| - MWOAFLSector | In welke sector en deelsector is hij / zij opgeleid? |
| - MWOAFLGK3Cat | Grootteklasse van het bedrijf in 3 categorieën |
| - MWOAFLsbiz2008Aggr4 | Bedrijfsclassificatie 2008 (4-deling) |
| - MWOAFLsbiz2008Aggr10 | Bedrijfsclassificatie 2008 (10-deling) |
| - MWOAFLsbiz2008Aggr21 | Bedrijfsclassificatie 2008 (21-deling) |
| - MWOAFLsbiz2008Aggr38 | Bedrijfsclassificatie 2008 (38-deling) |
| - MWOAFLbrc2014beroepsklasse | Beroepenindeling ROA-CBS 2014, 2 digits:
beroepsklasse |
| - MWOAFLisco08majorgroup | Beroep ISCO-08 major groups |
| - MWOAFLISCO08beroepsniveau | Beroepsniveau: geeft de complexiteit en omvang van
taken weer die bij een beroep horen |

5. Non-responsanalyse

5.1 Afbakenen statistische respons en nonresponsanalyse

De statistische respons bestaat uit bedrijven van wie de vragenlijsten administratief en technisch correct binnengekomen zijn. In paragraaf 3.5 is het steekproefkader en de steekproef beschreven en in paragraaf 3.6 de gerealiseerde respons. In totaal hebben 11 385 bedrijven de vragenlijst ingevuld, wat neerkomt op een aandeel van 50,6 procent. Een deel van de responsen blijkt uiteindelijk niet tot de doelgroep te behoren. Van de 11 385 ontvangen vragenlijsten zijn er 5 753 bruikbaar.

Voor de non-responsanalyse wordt de statistische respons beschreven en waar relevant ook de bruikbare respons. Het responspercentage voor de analyses in deze paragraaf is gedefinieerd als de statistische respons gedeeld door de uitgezette steekproef. Voor het percentage bruikbare respons wordt het aandeel van de uitgezette steekproef gegeven.

5.2 Selectiviteit van de respons naar achtergrondkenmerken

Doordat verschillende groepen personen in de getrokken steekproef meer of juist minder vaak responderen, kan de representativiteit van de steekproef verminderd worden. In deze paragraaf wordt de respons, en daarmee ook de non-respons, naar achtergrondkenmerken beschreven.

5.2.1 Respons mbo werkgeversonderzoek 2017 naar achtergrondkenmerken

	Uitgezette steekproef	Respons	Bruikbare respons
		<i>% van de uitgezette steekproef</i>	
Totaal	22 501	50,6	25,6
Steekproefportie			
Eerste portie	7 500	52,5	17,3
Tweede portie	15 001	49,7	29,7
Onderzoeksgroep			
Leerbedrijf	8 745	53,3	31,1
Niet-leerbedrijf	13 756	48,9	22,0
Bedrijfsklasse			
A: Landbouw, bosbouw en visserij	942	44,6	20,9
B: Delfstoffenwinning	9	77,8	66,7
C: Industrie	1 951	60,3	30,5
D: Energievoorziening	15	46,7	6,7
E: Waterbedrijven en afvalbeheer	63	76,2	44,4
F: Bouwnijverheid	1 541	49,3	24,9
G: Handel	5 285	54,1	25,8
H: Vervoer en opslag	870	62,6	28,4
I: Horeca	4 058	36,1	15,9
J: Informatie en communicatie	660	55,2	30,2
K: Financiële dienstverlening	257	43,2	22,6
L: Verhuur van en handel in onroerend goed	225	48,4	25,3

M: Specialistische zakelijke dienstverlening	1 366	56,0	32,7
N: Verhuur van overige zakelijke dienstverlening	1 768	59,1	25,2
O: Openbaar bestuur en overheidsdiensten	89	46,1	25,8
P: Onderwijs	309	49,8	28,8
Q: Gezondheids- en welzijnzorg	1 580	51,5	37,7
R: Cultuur, sport en recreatie	611	50,9	22,7
S: Overige dienstverlening	902	42,7	25,6
Grootteklasse			
1-9 werknemers	12 310	42,2	21,5
10-99 werknemers	8 604	60,8	29,6
>=100 werknemers	1 587	60,7	35,2

Uit tabel 5.2.1 blijkt dat de verschillen tussen de steekproefporties en ook tussen de onderzoeksgroepen klein zijn. De respons van de bedrijfsklassen verschilt wel, waarbij echter ook rekening gehouden moet worden met de omvang van de uitgezette steekproef. De bedrijfsklasse 'Delfstoffenwinning' kent met 77,8 procent de hoogste respons, maar daarbij gaat het om een klein aantal bedrijven. De respons onder de bedrijfsklasse 'I: Horeca' is met 36,1 procent het laagst. Dit zijn veelal bedrijven met minder dan 10 werknemers in dienst. De respons onder de kleine bedrijven ligt beduidend lager dan onder de grotere bedrijven. De bruikbare respons is voor deze bedrijfsklasse echter ook laag. Wat hier kan meespelen is dat in de horeca veel studenten werkzaam zijn. Op basis van registerinformatie is bepaald dat bij een bedrijf een recent afgestudeerde mbo'er werkzaam is. Het kan daarbij echter gaan om iemand die een vervolgopleiding doet en die een baan heeft naast zijn studie. In dat geval zal een werkgever wellicht niet weten dat het om een recent afgestudeerde mbo'er gaat en betreft het waarschijnlijk ook geen functie waar mbo niveau voor vereist is.

5.3 Weging

Om te corrigeren voor de selectiviteit in de respons en te voorkomen dat de doelvariabelen worden vertekend, worden de uitkomsten van het mbo werkgeversonderzoek gewogen. Hiervoor zijn insluitgewichten, correctiegewichten en eindgewichten ontwikkeld.

Het insluitgewicht d_i voor leerbedrijf i wordt als volgt berekend:

$$d_i = \frac{N_{\text{leerbedrijven}}}{n_{\text{leerbedrijven}}} \times \frac{n_{\text{leerbedrijven}}}{r_{\text{leerbedrijven}}} = \frac{N_{\text{leerbedrijven}}}{r_{\text{leerbedrijven}}}$$

met populatie-omvang N , steekproefomvang n en responsomvang r .

Het insluitgewicht d_j voor niet-leerbedrijf j wordt als volgt berekend:

$$d_j = \frac{N_{\text{niet-leerbedrijven}}}{n_{\text{niet-leerbedrijven}}} \times \frac{n_{\text{niet-leerbedrijven}}}{r_{\text{niet-leerbedrijven}}} = \frac{N_{\text{niet-leerbedrijven}}}{r_{\text{niet-leerbedrijven}}}$$

Invullen van de juiste aantallen leidt tot een insluitgewicht van 7,61 voor leerbedrijven en een insluitgewicht van 3,02 voor niet-leerbedrijven. Door insluitgewichten op deze manier te berekenen, tellen de insluitgewichten van alle berichtgevers samen op tot de populatie-omvang. Merk op dat de reciproke van het insluitgewicht geïnterpreteerd kan worden als de kans dat de berichtgever heeft gerepondeerd.

Met correctiegewichten wordt gecorrigeerd voor selectiviteit in respons. Eindgewichten zijn het product van insluit- en correctiegewichten. Deze gewichten zijn berekend aan de hand van een weegmodel bestaande uit één term: een kruising van groep (leerbedrijven vs. niet-leerbedrijven), SBI en grootteklasse. Groep is in deze kruising opgenomen omdat ernaar is gestratificeerd in het steekproefontwerp; SBI en grootteklasse zijn in deze kruising opgenomen omdat de respons mogelijk selectief is naar deze kenmerken en omdat deze kenmerken mogelijk samenhangen met de belangrijkste doelvariabelen van het onderzoek.

SBI is ingedikt tot branche (19 categorieën) en grootteklasse is ingedikt tot drie categorieën, zie tabel 5.2.1. Deze indikkingen leidden tot $2 \times 19 \times 3 = 114$ cellen. Hiervan bevatten 2 cellen geen eenheden die tot de populatie behoorden. Daarnaast bevatten meerdere cellen geen of slechts enkele eenheden in de respons. Deze cellen zijn zoveel als mogelijk samengevoegd met cellen met dezelfde groep en branche en andere grootteklasse, met als doel om in iedere cel minimaal 10 responsen te krijgen. Dit resulteerde in 85 cellen. Niet in al deze cellen zijn minimaal 10 responsen behaald. Desondanks is er voor gekozen om geen cellen samen te voegen met verschillende groepen of branches, omdat de eindgewichten per groep of branche dan niet meer zouden optellen tot de populatie-omvang per groep of branche.

De correctiegewichten variëren van 0,55 tot 2,76; het gemiddelde correctiegewicht is 1. Met de berekende eindgewichten wordt de respons in elk van de 85 cellen opgehoogd naar de populatie-omvang in de betreffende cel. Het minimale eindgewicht is 1,67 en het maximale eindgewicht is 21. De som van de eindgewichten is gelijk aan de populatie-omvang. Dat geldt niet alleen voor het totaal, maar ook per groep en per branche. Voor grootteklasse geldt dat niet, omdat cellen met verschillende grootteklassen zijn samengevoegd.

6. Andere kwaliteitsindicatoren

Behalve door naar de respons en de verdeling van de respons naar achtergrondkenmerken te kijken, kan ook op andere manieren een indruk gekregen worden van de kwaliteit van het onderzoek. In deze paragraaf wordt in beeld gebracht of er vragen zijn waarbij relatief vaak missings of de antwoordcategorieën “weet niet” of “weigert” voorkomen. Daarnaast wordt gekeken of bepaalde antwoordpatronen voorkomen of dat er inconsistenties zijn binnen de inhoud of de structuur van de vragenlijst. Ten slotte is het van belang dat er wordt gekeken naar opmerkingen van de respondenten.

6.1 Missings, geen antwoord op vragen en tellingen

De default van de vragenlijst is dat er geen missings en geen “weet niet” of “weigert” zijn toegestaan. Bij de vragen over de opleiding die een recent gediplomeerde heeft gevolgd, is dit wel toegestaan om te voorkomen dat respondenten die het antwoord op een vraag niet weten een willekeurig of sociaal wenselijk antwoord geven.


Bijna 20 procent respondenten geeft aan niet te weten welk mbo niveau de bij hem of haar werkzame mbo’er heeft behaald. Gaat het om de gevolgde leerweg, dat wil zeggen of het een bol of bbl opleiding betrof, dan geeft bijna 30 procent van de respondenten aan dit niet te weten. En voor de gevolgde sector gaat het om 23 procent. Verondersteld zou kunnen worden dat het vooral voor grote bedrijven lastig is om dat van alle werknemers te weten of in te schatten. Een kruising met de grootteklasse van het bedrijf leert echter dat het vooral de bedrijven zijn met minder dan 10 werknemers die aangeven het niet te weten.

6.2 Inconsistenties binnen de vragenlijst en afgeleide variabelen

Hoewel de vragenlijst uitgebreid wordt getest, kan het voorkomen dat er toch inconsistenties binnen de vragenlijst voorkomen of in de afleidingen die op basis van de data worden gemaakt. Ook kan het bijvoorbeeld voorkomen dat er overlap is tussen vragen, wat tot ergernis kan leiden bij respondenten. Door hier op terug te kijken kan de vragenlijst worden verbeterd voor de toekomst. Er zijn in de vragenlijst van het mbo werkgeversonderzoek geen inconsistenties geconstateerd.

6.3 Afwijkingen achtergrondgegevens en registerdata

6.3.1 Leerbedrijven

De informatie die SBB heeft aangeleverd over leerbedrijven is gekoppeld aan het ABR. Deze koppeling was niet volledig en vandaar dat niet-leerbedrijven de vraag wordt gesteld of ze een door SBB erkend leerbedrijf zijn. Op basis van de koppeling met de informatie van SBB zijn er 4.660 leerbedrijven die hebben gerespondeerd. Voor de vraag naar leerbedrijven blijven zij buiten beschouwing. Er zijn 6.725 niet-leerbedrijven die hebben gerespondeerd, hiervan geven er 2.753 zelf aan dat ze een door SBB erkend leerbedrijf zijn 

6.3.2 Recent aangenomen

Bij het afbakenen van het steekproefkader is gebruik gemaakt van informatie uit de onderwijsnummerbestanden en de Polisadministratie. Achterliggende gedachte was dat op die manier alleen bedrijven aangeschreven zouden worden die tot de onderzoekspopulatie behoren. Omdat de afbakening van de bedrijven dus voorafgaand aan het veldwerk heeft plaatsgevonden, is de vraag *Heeft uw bedrijf de afgelopen 2 jaar een mbo’er aangenomen die in*

2015 of 2016 het diploma heeft behaald? vooral bedoeld als controlevraag. Verwachting was dat vanwege administratieve ruis of wellicht recent ontslag een kleine groep respondenten deze vraag met nee zouden beantwoorden. In werkelijkheid blijkt de helft van de bedrijven deze vraag met nee te beantwoorden. In de eerste steekproefportie ging het zelfs om twee derde van de bedrijven. Voor de tweede steekproefportie is in de brief en de vragenlijst expliciet benoemd dat het bedrijf volgens de bij het CBS bekende gegevens een recent gediplomeerde mbo'er in dienst heeft. Het aandeel die deze vraag met nee beantwoordde was in deze twee portie weliswaar beduidend kleiner, maar nog steeds 40 procent.

Bedrijven met 100 werknemers of meer geven vaker aan dat ze een recent gediplomeerde mbo'er hebben aangenomen dan bedrijven met minder dan 100 werknemers. En leerbedrijven doen dit vaker dan niet-leerbedrijven.

Van de bedrijfsklassen is het de gezondheids- en welzijnszorg die met 73 procent relatief vaak aangeven dat ze inderdaad een of meer recent gediplomeerde mbo'ers in dienst hebben.

Het is niet precies duidelijk waarom bedrijven zo vaak aangeven dat ze geen recent gediplomeerde mbo'er in dienst hebben terwijl dat volgens de registraties wel het geval zou moeten zijn. Uit de reacties van respondenten bleek wel dat ze het moeilijk vinden om te achterhalen om wie het gaat. En zeker als het gaat om grote bedrijven of bedrijven met meerdere vestigingen, heeft men daar ook wellicht niet altijd goed zicht op.

6.4 Belangrijke opmerkingen van respondenten

Opmerkingen van respondenten kunnen inzicht geven in de manier waarop respondenten de vragenlijst hebben ingevuld en de begrijpelijkheid van vraagstellingen. Respondenten hebben in de vragenlijst van het mbo werkgeversonderzoek twee keer de mogelijkheid om een toelichting te geven.

6.4.1 Tevredenheid aansluiting opleiding beroep

Meer dan de helft van de bedrijven is tevreden of zeer tevreden over de aansluiting van de door de mbo'er gevolgde opleiding en het beroep dat ze binnen het bedrijf uitoefenen. Degenen die (zeer) tevreden zijn, maken minder van de mogelijkheid gebruik om hun antwoord toe te lichten dan degenen die ontevreden zijn. Zo gaat het bij degenen die (zeer) tevreden zijn om nog geen 10 procent. Uit hun antwoorden blijkt dat het van de persoon en de onderwijsinstelling afhangt hoe goed de aansluiting is. Ook wordt aangegeven dat de aansluiting van de bbl-opleidingen naar tevredenheid is, maar dat dit minder geldt voor de bol-opleidingen. En dat de opleiding een goede basis is, maar dat het echte leren in de praktijk plaatsvindt.

Bijna 600 respondenten geven aan dat ze ontevreden of zeer ontevreden zijn over de aansluiting van de door de mbo'er gevolgde opleiding en het beroep dat ze binnen het bedrijf uitoefenen. Van deze personen die (zeer) ontevreden zijn, heeft de helft van de mogelijkheid gebruik gemaakt om hun antwoord toe te lichten. Uit deze toelichtingen komt onder meer naar voren dat er geen specifieke opleidingen zijn voor een aantal beroepen, maar ook dat het niveau van de opleiding te laag is en dat de opleidingen te kort zijn waardoor de studenten niet de mogelijkheid hebben om het vak echt te kunnen. Ook sluit de inhoud van de opleidingen niet goed aan bij de praktijk.

6.4.2 Opmerkingen over de vragenlijst

Aan het einde van de vragenlijst hebben de respondenten de mogelijkheid om nog opmerkingen te maken over de vragenlijst zelf. Dit geldt zowel voor degenen die een recent gediplomeerde mbo'er in dienst hebben en de hele vragenlijst doorlopen hebben, als degenen die aan hebben

gegeven geen recent gediplomeerde mbo'er in dienst te hebben en die de vragenlijst niet hebben hoeven invullen.

Van degenen die hebben aangegeven dat ze een recent gediplomeerde mbo'er in dienst hebben, heeft 8 procent een opmerking gemaakt. Daarbij gaat het om procesmatige opmerkingen, namelijk dat de contactpersoon bijvoorbeeld is gewijzigd, dat ze geen interesse hebben in het invullen van de vragenlijst of dat ze de vragenlijst niet zinvol vinden. Maar ook dat het lastig is om de vragenlijst in te vullen omdat dit soort gegevens niet worden bijgehouden. Het aandeel respondenten dat een opmerking heeft gemaakt is vrijwel gelijk voor beide steekproefporties en ook voor de onderzoeksgroepen.

Bijna 5 duizend respondenten hebben aangegeven geen recent gediplomeerde mbo'er in dienst te hebben. Zij hebben de vragenlijst dus niet hoeven in te vullen. Van deze groep heeft 12 procent een opmerking gemaakt. Door deze groep wordt bijvoorbeeld aangegeven dat het bedrijf slechts één persoon in dienst heeft of dat het een holding betreft zonder personeel. Of dat ze wel een recent gediplomeerde mbo'er in dienst hebben, maar dat die al langer in dienst is dan twee jaar. En ook wordt aangegeven dat de aangeschreven persoon niet werkzaam is in het bedrijf. Het aandeel respondenten dat een opmerking heeft gemaakt, is vrijwel gelijk onder de beide onderzoeksgroepen (leerbedrijven en niet-bedrijven) maar verschilt tussen de steekproefporties. Van de respondenten in de eerste portie heeft 8 procent van de mogelijkheid gebruik gemaakt een opmerking te maken, van degenen in de tweede portie is dat 15 procent.

7. Plausibiliteitsanalyse

Om na te gaan of de uitkomsten van een onderzoek plausibel zijn, kan een vergelijking gemaakt met de data van eerdere edities, en indien beschikbaar andere databronnen. Voor het mbo werkgeversonderzoek is voor de leerbedrijven informatie beschikbaar uit 2016. In onderstaande wordt een vergelijking met dat jaar gemaakt. Hierbij gaat het alleen om de vragen over het oordeel van de bedrijven. Nagegaan wordt of de resultaten over de twee onderzoeksjaren vergelijkbaar zijn of dat de resultaten verschillen. Een vergelijking voor andere onderwerpen en voor de niet-leerbedrijven kan door het CBS niet gemaakt worden.

Voor deze plausibiliteitsanalyse zijn de belangrijkste doelvariabelen bepaald, waarbij het zowel kan gaan om vraagvariabelen als om afgeleide variabelen. Voor de analyses worden gewogen cijfers gebruikt omdat het gaat om de uitkomsten van het onderzoek. En zijn alleen de respondenten geselecteerd die recent een mbo'er hebben aangenomen. Er wordt een vergelijking naar onderzoeksgroep (leerbedrijven en niet-leerbedrijven) gemaakt. Het gaat hierbij om de afbakening zoals voorafgaand aan het veldwerk is gemaakt en niet die op basis van de informatie uit de vragenlijst.

7.1 Vergelijking oudere jaren met dit onderzoek

Het CBS beschikt voor 2016 alleen over de resultaten over de beoordeling van leerbedrijven over de recent gediplomeerde mbo'ers. In onderstaande tabellen worden die uitkomsten vergeleken met de uitkomsten op basis van het mbo werkgeversonderzoek 2017. Ter vergelijking zijn de resultaten van de niet-leerbedrijven ook gepresenteerd.

In totaal gaat het om zeven aspecten. Voor alle aspecten is het patroon voor de jaren 2016 en 2017 vergelijkbaar. De meeste werkgevers beoordelen de aspecten met een goed, de reken- en taalvaardigheden worden het vaakst met een slecht of zeer slecht beoordeeld. Dit geldt zowel in 2016 als in 2017, en zowel voor de leerbedrijven als de niet-leerbedrijven.

7.1.1 Beoordeling vakkennis

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	1	.	0,5	0,7
Slecht	4	.	2,9	2,3
Niet goed/niet slecht	18	.	19,7	21,8
Goed	66	.	71,2	67,6
Zeer goed	10	.	5,7	7,6

7.1.2 Beoordeling beroepsvaardigheden

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	0	.	0,3	0,6
Slecht	4	.	2,4	2,1
Niet goed/niet slecht	20	.	17,5	20,1
Goed	67	.	71,6	68,1
Zeer goed	9	.	8,3	9,1

7.1.3 Beoordeling beroepshouding

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	1	.	0,3	0,8
Slecht	3	.	2,4	2,9
Niet goed/niet slecht	15	.	15,5	17,0
Goed	67	.	66,4	63,4
Zeer goed	14	.	15,4	15,9

7.1.4 Beoordeling taal- en rekenvaardigheden

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	1	.	0,4	1,1
Slecht	6	.	5,8	5,9
Niet goed/niet slecht	31	.	29,0	29,1
Goed	58	.	59,6	57,1
Zeer goed	4	.	5,2	6,8

7.1.5 Beoordeling omgang met collega's en leidinggevenden

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	0	.	0,3	0,7
Slecht	1	.	1,5	1,5
Niet goed/niet slecht	10	.	10,0	13,0
Goed	72	.	69,4	66,2
Zeer goed	16	.	18,8	18,6

7.1.6 Beoordeling bereidheid zich in te zetten voor het bedrijf

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	0	.	0,3	1,0
Slecht	2	.	1,6	1,5
Niet goed/niet slecht	11	.	9,3	12,0
Goed	64	.	60,9	57,2
Zeer goed	23	.	27,9	28,3

7.1.7 Beoordeling vermogen om zicht in het vak te ontwikkelen en bij te leren

	2016		2017	
	Leerbedrijf	Niet-leerbedrijf	Leerbedrijf	Niet-leerbedrijf
Zeer slecht	0	.	0,2	0,9
Slecht	3	.	2,6	3,0
Niet goed/niet slecht	16	.	15,1	16,8
Goed	65	.	64,1	60,5
Zeer goed	16	.	18,1	18,8

7.2 Vergelijking schoolverlatersonderzoek

Omdat het gaat om kenmerken van recent gediplomeerde mbo, is daar waar relevant een vergelijking gemaakt met het mbo schoolverlatersonderzoek 2017. Die vergelijking is niet helemaal zuiver omdat het in het mbo werkgeversonderzoek gaat om het werkgeversperspectief en werkgevers bovendien, in geval zij meerdere recent gediplomeerde mbo'ers in dienst hebben, slechts één van die mbo'ers voor ogen moeten houden. Doelpopulatie van het schoolverlatersonderzoek zijn alle mbo'ers die recent hun diploma hebben behaald. Voor de vergelijkbaarheid is in het schoolverlatersonderzoek ook een selectie gemaakt van de personen die een baan hebben als werknemer.

Tabel 7.2.1 laat voor het niveau en de leerweg van de gevolgde opleiding zien dat de uitkomsten van beide bronnen niet goed vergelijkbaar zijn. Dat heeft onder meer te maken met het grote aandeel onbekenden op zowel niveau als leerweg. Vooral voor niveau 2 is het verschil groot tussen het mbo werkgeversonderzoek en het schoolverlatersonderzoek en datzelfde geldt voor de bol opleiding. Worden deze onbekenden buiten beschouwing gelaten dan worden de verschillen voor niveau 4 en de bbl groter.

Conclusie is dat het op basis van het mbo werkgeversonderzoek niet mogelijk is om in beeld te brengen welk niveau en welke leerweg werkzame recent gediplomeerde mbo'ers hebben gevolgd. Wel is deze informatie relevant als achtergrondinformatie als het gaat om de tevredenheid over de aansluiting van de opleiding en de functie op de arbeidsmarkt.

7.2.1 mbo niveau en leerweg

	MWO 2017	SVO 2017
	%	
<i>Mbo niveau</i>		
Niveau 1	2,1	3,5
Niveau 2	11,8	22,7
Niveau 3	27,6	31,9
Niveau 4	41,7	41,9
Onbekend	16,8	
<i>Mbo leerweg</i>		
BOL	40,2	62,6
BBL	34,6	37,4
Onbekend	25,2	


Eenzelfde analyse als voor het behaalde mbo niveau en de gevolgde leerweg is ook gedaan voor het beroep dat recent gediplomeerde mbo'ers uitoefenen. De uitkomsten daarvan staan in tabel 7.2.2. De verschillen tussen beide bronnen zijn voor de meeste beroepsklassen klein. Alleen in de zorg en welzijn beroepen zijn de verschillen relatief groot.

7.2.2 Beroepsklasse mbo'ers

	MWO 2017	SVO 2017
	%	
Pedagogische beroepen	4,7	3,8
Creatieve en taalkundige beroepen	1,3	1,0
Commerciële beroepen	10,2	13,1
Bedrijfseconomische en administratieve beroepen	14,0	9,3
Managers	0,9	1,3
Openbaar bestuur, veiligheid en juridische beroepen	0,9	2,8



Technische beroepen	23, 	19,5
ICT beroepen	3,5	2,5
Agrarische beroepen	3,3	2,0
Zorg en welzijn beroepen	11,7	25,3
Dienstverlenende beroepen	17,1	11,9
Transport en logistieke beroepen	4,9	6,3
Overige beroepen	4,1	1,2

8. Conclusie kwaliteitsanalyse


In dit document is het mbo werkgeversonderzoek 2017 beschreven. Voor zowel de voorbereidingen en het veldwerk als voor de kwaliteit van de uitkomsten worden de belangrijkste conclusies beschreven en aanbevelingen gedaan. 

8.1 Conclusies

Op basis van onderliggende onderzoeksverantwoording worden de volgende conclusies getrokken:

- De responstargets zijn gerealiseerd. Om die te halen was een aanvullende steekproeftrekking nodig. Het gerealiseerde responspercentage is 50,6%, ruim 10 procentpunten hoger dan het verwachte responspercentage. In totaal is 25,6% bruikbare respons.
- De targets voor zowel de leerbedrijven als de niet-leerbedrijven zijn gehaald, bij de leerbedrijven gaat het om 4 660 bedrijven, waarvan 2 722 tot de bruikbare respons behoren. Voor de niet-leerbedrijven gaat het om 6 725 bedrijven, waarvan er 3 031 tot de bruikbare respons behoren.
- In de vragenlijst is vanaf de tweede steekproefportie aangegeven dat volgens de informatie van het CBS een recent gediplomeerde mbo'er in dienst is bij het bedrijf. Dit leverde veel vragen op bij de CCI-helpdesks  werkgevers vroegen zich af welke mbo'er bedoeld werd. Bij meerdere mbo'ers vroeg men zich bovendien af of de vragenlijst voor alle mbo'ers ingevuld moest worden.
- De aangeschreven contactpersonen blijken niet altijd de juiste personen te zijn om de vragenlijst in te vullen . Dit de opmerkingen van respondenten bleek informatie soms verouderd, en ook bleek uit de reactie die binnenkwamen bij de CCI-helpdesk dat de contactpersonen weinig zicht hadden op de werkzaamheden van de recent gediplomeerde mbo'er.
- Bijna de helft van de bedrijven in de doelpopulatie heeft aangegeven geen recent gediplomeerde mbo'er in dienst te hebben. De afbakening in de steekproef blijkt daarmee niet nauwkeurig genoeg te zijn. Verklaring kan zijn dat mbo'ers intussen het bedrijf hebben verlaten. Maar ook kunnen het mbo'ers zijn die een vervolgopleiding doen en een bijbaan hebben of mbo'ers die al langer dan twee jaar in dienst waren. Wat hierbij kan meespelen is dat in de intro van de vragenlijst is opgenomen dat er eerst vragen worden gesteld om te achterhalen of het bedrijf een recent gediplomeerde mbo'er in dienst heeft genomen. En als dit zo is, dat men dan de vervolgvragen voorgelegd krijgt. In geval men niet weet waar in het bedrijf een mbo'er in dienst is en dit uitzoeken wellicht veel tijd gaat kosten, kan het zijn dat men geneigd is om nee aan te vinken.
- Onder de bedrijven die aangeven geen recent gediplomeerde in dienst te hebben, blijken veel holdings te zitten die geen personeel in dienst hebben.
- Gezien het grote aandeel 'weet niet' scores hebben respondenten moeite met het beantwoorden van de vragen over de kenmerken van de gevolgde opleiding.
- De analyses laten voor de variabelen over het oordeel van de werkgever over aansluiting van de opleiding en de functie veelal plausible uitkomsten zien. Een vergelijking met het schoolverlatersonderzoek 2017 laat zien dat de resultaten voor de gevolgde opleiding geen dwarsdoorsnede laten zien van alle recent gediplomeerde mbo'ers die als werknemer werkzaam zijn.

8.2 Aanbevelingen

Gezien bovenstaande conclusies worden de volgende aanbevelingen gedaan 

- Aanpassen van de afbakening van het steekproefkader: gediplomeerde mbo'ers die nog een vervolgopleiding doen, moeten buiten beschouwing worden gelaten, net als degenen die al langer dan twee jaar in dienst zijn.
- Communicatie verbeteren: in de brief en de vragenlijst moet duidelijker uitgelegd worden dat het gaat om één mbo'er die de werkgever voor ogen moet houden. Ook kan in de brief worden aangegeven dat het aan het bedrijf is wie de vragenlijst invult en dat dit ook een direct betrokken leidinggevende van de recent gediplomeerde kan zijn.
- Aanpassingen in de vragenlijst: overwegen om toe te voegen wat werkgevers van mbo-opleidingen verwachten en of ze zelf nog opleidingen en cursussen aanbieden.

9. Opgeleverde Producten

9.1 Microbestanden

Voor SBB is het definitieve publicatie bestand via Remote Access (RA) ter beschikking gesteld. Bij dit bestand is een codeboek beschikbaar, conform de standaard werkwijze van het CBS.

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, nieuwsberichten, 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.

10. Referenties

1. SBB, Opzet en uitvoering onderzoek 'Motie Straus, verkregen via <https://www.s-bb.nl> op 31 maart 2018
2. Mooren, F. van der, Boutorat, A. en Paulissen, R. (2017). Verantwoording Vooronderzoek mbo werkgeversonderzoek, 12 september 2017, CBS: externe nota.
3. Projectovereenkomst mbo werkgeversonderzoek 2017, versie 25-09-2017;
4. Braams, N., Hoksbergen, M. en Kok, S. de (2017). Kostenindicatie MBO aansluiting op arbeidsmarkt 2017, versie 11-08-2017. CBS: externe nota.

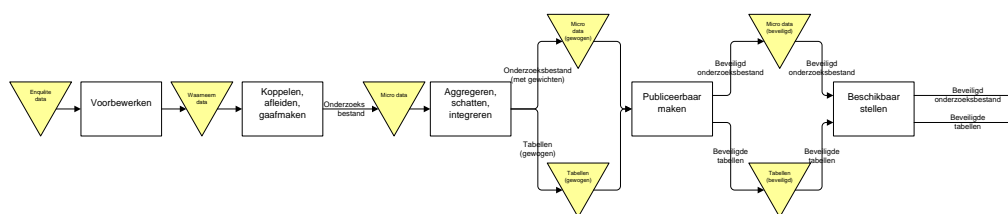
11. Bijlage I – Algemene beschrijving dataprocessing

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.

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



1. Voorbewerken: Van enquête data naar waarneem data

Voordat enquête data 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 waarneem data. De waarneem data is de grondstof voor het 'echte' verwerken.

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

1.1. Controleren enquête data: 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

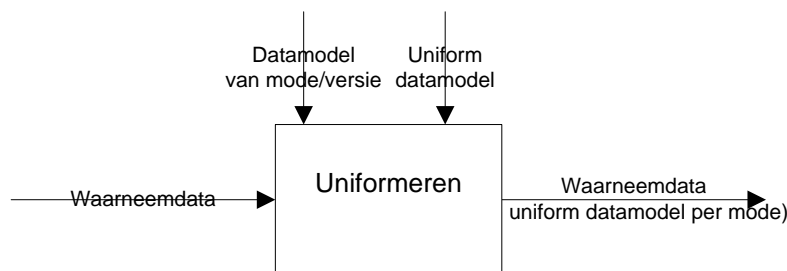
De recordstructuur van enquêtedata kan mode specifiek zijn. Dit komt omdat de vragenlijst mode specifiek is. Daarmee is ook de metadata mode specifiek. Het kan ook voorkomen dat gedurende het onderzoek er voor een mode meerdere versies van een vragenlijst worden gebruikt (eventueel ook met een verschil 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 recordstructuur² 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

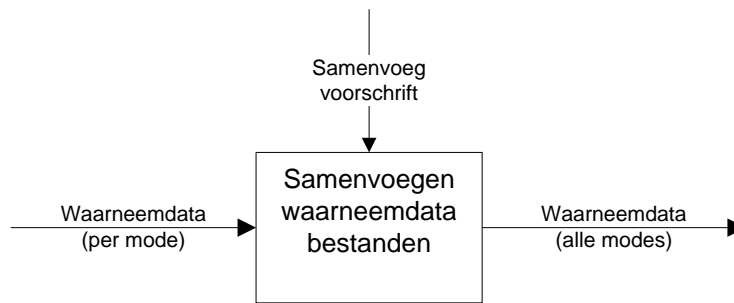


1.3 Controleren enquêtedata: range en routing controle

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

² 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.

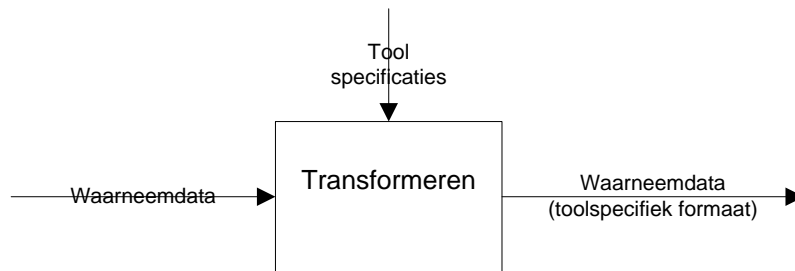
2.1 Transformeren

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 SYMIS. 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

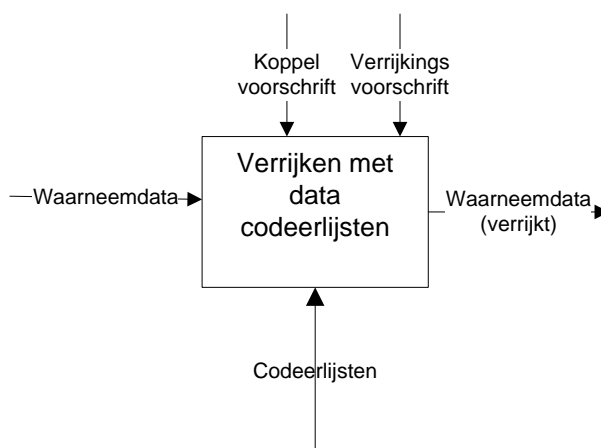
Binnen de verwerkingsprocessen gebeurt “het verrinnen” alleen t.b.v. het koppelen met registerdata en dus niet om met geanonimiseerde data in het verwerkingsproces te werken. Om waarneemdata te kunnen verrijken met registerdata dient de waarneemdata eerst verrind te worden. Dit betekent dat voor iedere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

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

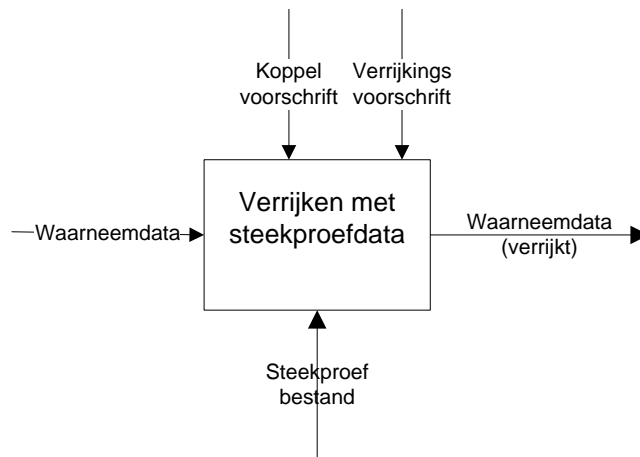


2.4 Verrijken met steekproefdata

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 bevestigd.

Figuur 6: Procesmodel Verrijken met steekproefdata

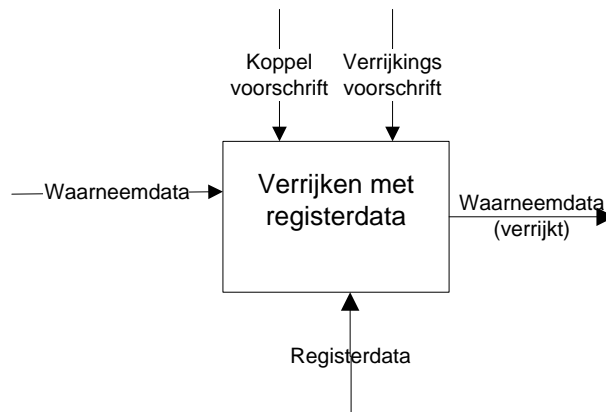


2.5 Verrijken met registerdata

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 (BRP), bron inkomen en hoogte inkomen (Polisadministratie), provincie, inschrijfduur (is een afleiding), werkend (UWV-Werkbedrijf).

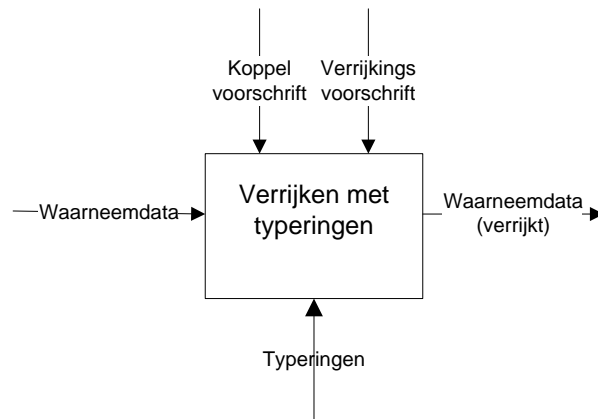
Figuur 7: Procesmodel Verrijken met registerdata



2.6 Verrijken met typeringen

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

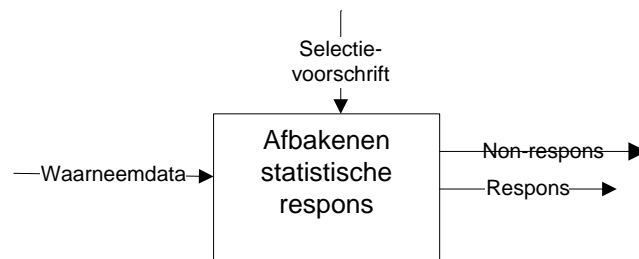
Figuur 8: Procesmodel Verrijken met typeringen



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.

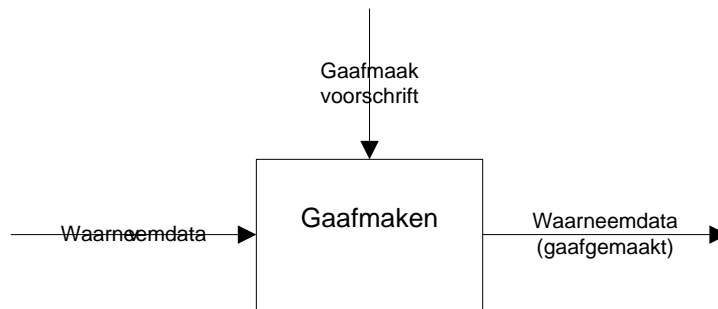
Figuur 9: Procesmodel Afbakenen statistische respons



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.

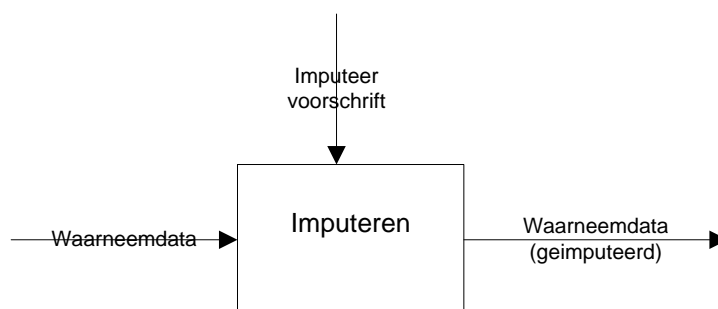
Er zijn een aantal manieren om met ontbrekende waarden om te gaan. Eén daarvan is imputeren van een geldige waarde voor de ontbrekende waarde in het databestand.

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

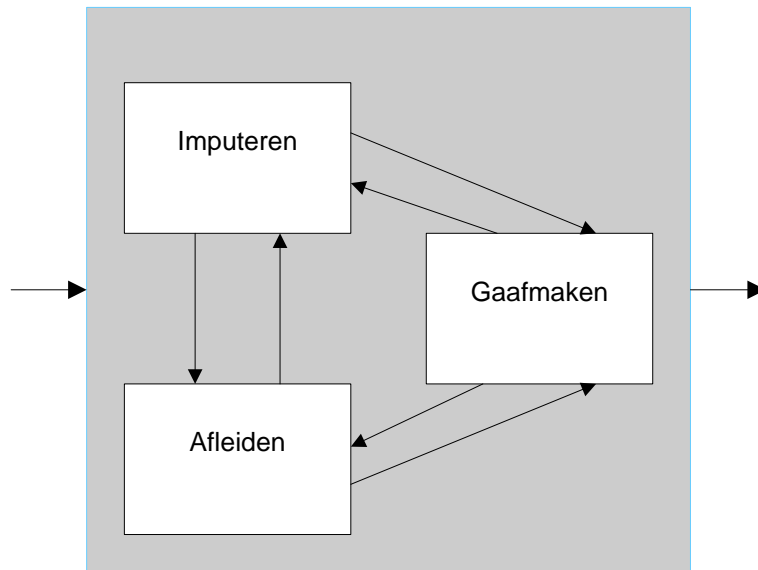
Met afleiden wordt hier bedoeld het creëren van nieuwe variabelen als functie van in het bestand reeds bestaande variabelen.

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.

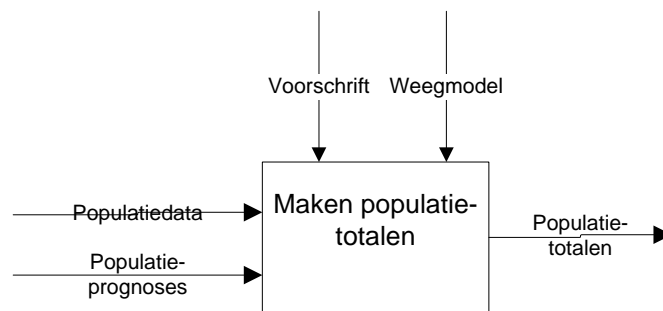
Figuur 12: Procesmodel Gaafmaken, imputeren en afleiden: Samenhang



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.

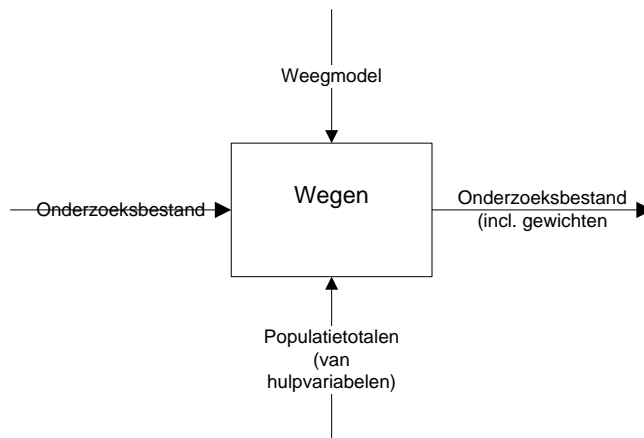
Figuur 13: Procesmodel maken populatietotalen



2.1.3 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.