Centraal Bureau
voor de Statistiek
Microdata Services

Documentatie
Nationaal Cohortonderzoek Onderwijs (NCO)

Datum: september 2020

Bronvermelding

Publicatie van uitkomsten geschiedt door de onderzoeksinstelling of de opdrachtgever op eigen titel. Verwijzing naar het Centraal Bureau voor de Statistiek (CBS) betreft uitsluitend het gebruik van de niet–openbare microdata. Deze microdata zijn onder *bepaalde voorwaarden* voor statistisch en wetenschappelijk onderzoek toegankelijk. Voor nadere informatie *microdata@cbs.nl*. Dat wordt als volgt geformuleerd:

"Resultaten [gedeeltelijk] gebaseerd op eigen berekeningen [*naam onderzoeksinstelling, c.q. opdrachtgever*] op basis van niet-openbare microdata van het Centraal Bureau voor de Statistiek en Dienst Uitvoering Onderwijs, betreffende Nationaal Cohort Onderzoek."

Engelse versie

"Results based on calculations by [*name of research institution or commissioning party*] using non-public microdata from Statistics Netherlands."

"Under *certain conditions*, these microdata are accessible for statistical and scientific research. For further information: *microdata@cbs.nl*."

Inhoud

Bronvermelding 2
Inleiding 4
Toestemming aanvragen 4
Gebruik van de bestanden 4
Beschikbare bestanden 5
Planning updates 5
Bestandsopbouw en toelichting 6
Cohorten 6
Primair onderwijs startcohort (NCOPORUG) 6
Primair onderwijs uitstroomcohort (NCOPOUITRUG) 6
Voortgezet onderwijs startcohort (NCOVORUG) 7
Onderwijsgegevens 9
Primair onderwijs (NCOPO) 9
Voortgezet onderwijs (NCOVO) 10
Speciaal (voortgezet) onderwijs (NCOSO) 14
Voortgezet algemeen volwassenenonderwijs (NCOVAVO) 14
Middelbaar beroepsonderwijs (NCOMBO) 16
Hoger beroepsonderwijs (NCOHBO) 18
Wetenschappelijk onderwijs (NCOWO) 19
Schoolgegevens 21
Schoolgegevens: primair onderwijs (NCOBRINPO) 21
Schoolgegevens: voortgezet onderwijs (NCOBRINVO) 22
Achtergrondgegevens 24
Achtergrondgegevens: stabiel (NCOACHTERGRONDSTABIEL) 24
Achtergrondgegevens: variabel (NCOACHTERGRONDVARIABEL) 24

Inleiding

Het team Microdata Services van het CBS stelt, onder *bepaalde voorwaarden*, niet–openbare microdata (geanonimiseerde data op persoons-, bedrijfs- en adresniveau) toegankelijk voor statistisch en wetenschappelijk onderzoek. Dit rapport beschrijft de inhoud en structuur van NCO bestanden.

De in deze documentatie beschreven microdatabestanden zijn door het Centraal Bureau voor de Statistiek (CBS) samengesteld voor het Nationaal Cohortonderzoek Onderwijs (NCO). Het NCO is geïnitieerd en wordt bekostigd door het Nationaal Regieorgaan Onderwijsonderzoek (NRO). Voor deze bestanden zijn bestaande registerdata uit de sociaal-statistische bestanden (SSB) van het CBS en bestaande onderwijsregistraties van het Basisregister Onderwijs (BRON) van Dienst Uitvoering Onderwijs (DUO) gecombineerd.

Toestemming aanvragen

Voor het gebruik van dit bestand, is zowel toestemming van het NRO als van het CBS nodig.
Toestemming van NRO kan aangevraagd worden via:
https://nationaalcohortonderzoek.nl/onderzoek/ .

Gebruik van de bestanden

Voor de NCO bestanden is een tool ontwikkeld waarin op eenvoudige wijze bestanden gekoppeld kunnen worden en ook longitudinale bestanden kunnen worden samengesteld. In de tool kan gekozen worden voor het aanmaken van een bestand, maar ook alleen een SPSS syntax. Onderzoekers kunnen indien gewenst de syntax zelf aanpassen.

De NCO tool is beschikbaar via Remote Access (RA), de “2020_Handleiding NCOTool.docx” is voor RA gebruikers beschikbaar in de map K:\Utilities\Toelichtingen\NCO en in de map G:\MAATWERK\NCOV3.

Beschikbare bestanden

Type bestand	Beschikbaar
Primair onderwijs startcohort (NCOPORUG)	Vanaf 2010
Primair onderwijs uitstroomcohort (NCOPOUITRUG)	Vanaf 2010
Voortgezet onderwijs startcohort (NCOVORUG)	Vanaf 2007
Deelname en toets gegevens primair onderwijs (NCOPO)	Vanaf 2010
Deelname en examen gegevens voortgezet onderwijs (NCOVO)	Vanaf 2007
Speciaal onderwijs (NCOSO)	Vanaf 2008
Volwassenenonderwijs NCOVAVO)	Vanaf 2010
Middelbaar beroepsonderwijs (NCOMBO)	Vanaf 2010
Hoger beroepsonderwijs (NCOHBO)	Vanaf 2010
Wetenschappelijk onderwijs (NCOWO)	Vanaf 2010
Scholen primair onderwijs (NCOBRINPO)	Vanaf 2010
Scholen voortgezet onderwijs (NCOBRINVO)	Vanaf 2007
Stabiele achtergrondgegevens (NCOACHTERGRONDSTABIEL)	Cumulatief bestand
Variabele achtergrondgegevens (NCOACHTERGRONDVARIABEL)	Vanaf 2007

Planning updates

Actualisaties vinden elk jaar eind september plaats.

- De cohortbestanden worden eind september aangevuld met zij-instromers. Voor het PO-cohort geldt dat het daarbij gaat om zeven jaren waarin leerlingen het cohort kunnen instromen. Samen met het jaar waarin het cohort start is dat acht jaar, de nominale duur van het basisonderwijs. Voor het VO gaat het om het jaar waarin het cohort start plus vijf jaren waarin leerlingen het cohort kunnen instromen. Samen zes jaar, de nominale duur van het vwo.
- Voor de deelnamegegevens zijn de cijfers van het laatste jaar voorlopig. Eind september worden die bijgesteld van voorlopig naar definitief en worden nieuwe voorlopige cijfers toegevoegd.
- Voor de toets- en examengegevens worden nieuwe definitieve cijfers toegevoegd.
- Van de scholen worden nieuwe definitieve gegevens toegevoegd.
- De stabiele achtergrondkenmerken worden aangevuld met gegevens van nieuwe cohorten en van zij-instromers.
- Voor de variabele achtergrondkenmerken worden de gegevens bijgesteld van VRL0 naar VRL1 en die van VRL1 naar definitief (DEF). Ook worden nieuwe voorlopige cijfers (VRL0) toegevoegd.

Bestandsopbouw en toelichting

Cohorten

Primair onderwijs startcohort (NCOPORUG)
Leerlingen in het (speciaal) basisonderwijs

Bestandsindeling NCOPORUG

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	STARTJAAR Schooljaar waarin leerling is ingestroomd in cohort	F4

Toelichting:
Deze bestanden bevatten gegevens van de leerlingen die in een bepaald jaar in het eerste jaar van het primair onderwijs zitten, jaarlijks aangevuld met de zij-instromers. De bestanden zijn bedoeld als ruggengraat en bevatten alleen de koppelsleutel (RINPERSOON en RINPERSOONS) en het schooljaar waarin de leerling het cohort is ingestroomd. Leerlingen komen in meerdere cohorten voor als zij zijn blijven zitten of als zij versnellen. In incidentele gevallen komt het voor dat leerlingen sterk lijken te vertragen. Zo zijn er bijvoorbeeld enkele leerlingen die in 2010 in groep 8 zitten en in 2011 in groep 2. Waarschijnlijk betreft het hier administratieve fouten. Hiervoor is niet gecorrigeerd. Blijven zij in groep 1 zitten dan is het startjaar in beide cohorten gelijk aan het geselecteerde cohort. In alle andere situaties geldt dat in het cohort waarin ze gestart zijn, het startjaar overeenkomt met het cohortjaar, en dat ze in het cohort waarin ze instromen een later startjaar hebben.

Op basis van het startjaar kunnen de leerlingen geselecteerd worden die oorspronkelijk in het betreffende cohort zijn gestart. Doordat de zij-instromers ook zijn opgenomen kan door koppeling met de onderwijskenmerken een dwarsdoorsnede gegeven worden van alle leerlingen in een bepaald leerjaar en schooljaar.

Primair onderwijs uitstroomcohort (NCOPOUITRUG)
Leerlingen in groep 8 van het (speciaal) basisonderwijs.

Bestandsindeling NCOPOUITRUG

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9

Toelichting:
Deze bestanden bevatten gegevens van de leerlingen die in een bepaald jaar in groep acht van het primair onderwijs zitten. De bestanden bevatten alleen een koppelsleutel (RINPERSOON en RINPERSOONS) en zijn specifiek bedoeld voor analyses betreffende de overgang van het primair

onderwijs naar het voortgezet onderwijs. De informatie over het primair onderwijs is vanaf 2010 beschikbaar. Dit betekent dat de leerlingen die in dat schooljaar gestart zijn, in principe voor het eerst in 2018 naar het voortgezet onderwijs gaan. Analyses betreffende de overgang primair onderwijs naar voortgezet onderwijs zijn op basis van de huidige PO startcohorten dus voor het eerst pas in 2019 mogelijk aan de hand van de dan nog voorlopige VO-gegevens. Vandaar dat de uitstroomcohorten zijn samengesteld.

Als gesproken wordt over uitstroom worden de leerlingen in groep 8 bedoeld. Het kan voorkomen dat leerlingen in groep 8 blijven zitten. In dat geval stromen ze niet uit naar het voortgezet onderwijs, en zijn ze niet in het volgend jaar in het voortgezet onderwijs terug te vinden, maar pas het jaar daarop. Omdat ze groep 8 twee keer doen, maken ze deel uit van twee uitstroomcohorten.

Voortgezet onderwijs startcohort (NCOVORUG)
Leerlingen in het voortgezet onderwijs

Bestandindeling NCOVORUG

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	STARTJAAR Schooljaar waarin leerling is ingestroomd in cohort	F4
4	VMBOB Of leerling een vmbo-b-diploma behaald heeft	A1
5	VMBOB_JAAR Schooljaar waarin leerling een vmbo-b-diploma behaald heeft	F4
6	VMBOK Of leerling een vmbo-k-diploma behaald heeft	A1
7	VMBOK_JAAR Schooljaar waarin leerling een vmbo-k-diploma behaald heeft	F4
8	VMBOGT Of leerling een vmbo-g/t-diploma behaald heeft	A1
9	VMBOGT_JAAR Schooljaar waarin leerling een vmbo-g/t-diploma behaald heeft	F4
10	HAVO Of leerling een havo-diploma behaald heeft	A1
11	HAVO_JAAR Schooljaar waarin leerling een havo-diploma behaald heeft	F4
12	VWO Of leerling een vwo-diploma behaald heeft	A1
13	VWO_JAAR Schooljaar waarin leerling een vwo-diploma behaald heeft	F4

Toelichting:
Deze bestanden bevatten gegevens van de leerlingen die in een bepaald jaar in het eerste jaar van het voortgezet onderwijs zitten, jaarlijks aangevuld met de zij-instromers. De bestanden zijn

bedoeld als ruggengraat en bevatten de koppelsleutel (RINPERSOON en RINPERSOONS) en het schooljaar waarin de leerling het cohort is ingestroomd.
Daarnaast bevatten deze VO startcohorten voor elke schoolsoort een diplomavariabele en een variabele voor het jaar waarin de leerling dat diploma heeft behaald. Voor leerlingen die na het behalen van hun diploma nog een opleiding in het voortgezet onderwijs zijn gaan doen, en ook daar een diploma voor behalen, is voor beide opleidingen aangegeven dat er een diploma is behaald en wordt ook het jaar waarin dat diploma is behaald weergegeven.
Deze leerlingen die nog een opleiding in het voortgezet onderwijs doen, komen in meerdere cohorten voor. Dit geldt ook voor de leerlingen die blijven zitten of versnellen. Blijven zij in leerjaar 1 zitten, dan is het startjaar in beide cohorten gelijk aan het geselecteerde cohort. In alle andere situaties geldt dat in het cohort waarin ze gestart zijn het startjaar overeenkomt met het cohortjaar, en dat ze in het cohort waarin ze instromen een later startjaar hebben.

Op basis van het startjaar kunnen de leerlingen geselecteerd worden die oorspronkelijk in het betreffende cohort zijn gestart. Doordat de zij-instromers ook zijn opgenomen, kan door koppeling met de onderwijskenmerken een dwarsdoorsnede gegeven worden van alle leerlingen in een bepaald leerjaar en schooljaar.

Onderwijsgegevens

Primair onderwijs (NCOPO)
Inschrijvings- en toetsgegevens van alle leerlingen in het (speciaal) basisonderwijs op persoonsniveau

Bestandsnaam: NCOPODEF of NCOPOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	WPOADVIESVO Advies over het te volgen aansluitend voortgezet onderwijs	A2
4	WPOBRIN_versl Code van de onderwijsinstelling van de inschrijving in het (speciaal) basisonderwijs - versleuteld	A4
5	WPOBRINVEST Het onderwijsvestigingsnummer van de vestiging waar de leerling het (speciaal) basisonderwijs volgt	A2
6	WPOCODEEINDTOETS Het soort eindtoets dat de leerling in het (speciaal) basisonderwijs heeft afgelegd.	A2
7	WPOGEWICHT Het gewicht van een leerling in het basisonderwijs volgens de gewichtenregeling.	A4
8	WPOLEERJAAR Het leerjaar in het (speciaal) basisonderwijs waarin de betreffende leerling zit.	A2
9	WPOTYPEPO Het soort primair onderwijs dat de leerling in het (speciaal) basisonderwijs volgt.	A3
10	WPOUITSLAGEINDTOETS De standaardscore op de door de leerling afgelegde eindtoets in het (speciaal) basisonderwijs.	A3
11	CITOPERCENTIELTAAL Percentielscore taal op de CITO-toets	F3
12	CITOPERCENTIELREKENENWISKUNDE Percentielscore rekenen op de CITO-toets	F3
13	WPOADVIESHERZ Herzien leerkrachtadvies over het volgen van aansluitend voortgezet onderwijs	A2
14	WPOREKENEN Toetsonderdeel rekenen	A2
15	WPOTAALLV Toetsonderdeel taal - leesvaardigheid	A2
16	WPOTAALTV Toetsonderdeel taal - taalverzorging	A2

Toelichting:
De NCOPO-bestanden bevatten de gegevens van leerlingen die in minimaal een van de PO-cohorten voorkomen. De cohorten vormen de ruggengraat van het NCO waar op basis van de koppelsleutel (RINPERSOON en RINPERSOONS) de inschrijvings- en toetsgegevens van leerlingen in het (speciaal) basisonderwijs aangekoppeld kunnen worden.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de toetsgegevens. De NCOPODEF bestanden bevatten definitieve inschrijvings- en toetsgegevens, het NCOPOVRL bestand bevat voorlopige inschrijvingsgegevens en geen toetsgegevens. De toets (en advies-)gegevens betreffen de variabelen WPOADVIESVO, WPOCODEEINDTOETS, WPOUITSLAGEINDTOETS, CITOPERCENTIELTAAL, CITOPERCENTIELREKENENWISKUNDE, WPOADVIESHERZ, WPOREKENEN, WPOTAALLV en WPOTAALTV; deze zijn in het voorlopige bestand dus niet gevuld.

De variabele CITOPERCENTIELTAAL is ook in NCOPO2017DEF en NCOPO2018DEF niet gevuld, omdat de percentielscore niet berekend kon worden in verband met afwijkende afname van de CITO toets de betreffende schooljaren.

De variabele WPOGEWICHT is vanaf 2019 niet meer gevuld.

Als een leerling in jaar JJJJ is gestart in het primair onderwijs, dan zijn de inschrijvings- en deelnamegegevens ook vanaf dat jaar beschikbaar. Zie voor meer informatie K:\onderwijs*INSCHRWPOTAB* en *CITOTAB*.

Voortgezet onderwijs (NCOVO)
Inschrijvings- en examengegevens van alle vo-leerlingen op persoonsniveau

Bestandsindeling: NCOVODEF en NCOVOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	BRIN_versl Code van de onderwijsinstelling van de inschrijving in het voortgezet onderwijs - versleuteld	A4
4	VOBRINVEST Het onderwijsvestigingsnummer van de vestiging waar de leerling in het voortgezet onderwijs ingeschreven staat.	A2
5	ONDERWIJSSOORTVO Onderwijssoort	A2
6	VOLEERJAAR Het leerjaar in het voortgezet onderwijs waarin de betreffende leerling zit.	A1
7	ILTCODE Ilt-code	A4
8	SECTORVO	A1

Nr	Variabele en omschrijving	Form.
	Vmbo-sector	
9	PROFIELVO	A2
	Profiel binnen havo/vwo	
10	EXAMUITSLAGVO	A1
	De uitslag van het examen van een leerling in het voortgezet onderwijs.	
11	BRINEXAMVO_versl	A4
	Code van de onderwijsinstelling waar de leerling in het voortgezet onderwijs het examen heeft afgelegd - versleuteld	
12	BRINVESTEXAMVO	A2
	Het onderwijsvestigingsnummer van de vestiging waar de leerling in het voortgezet onderwijs het examen heeft afgelegd.	
13	ONDERWIJSSOORTVODIPL	A2
	Onderwijssoort van het diploma	
14	SECTORVODIPL	A1
	Vmbo-sector van het diploma	
15	PROFIELVODIPL	A2
	Profiel binnen havo/vwo van het diploma	
16	CEGEM	F2.1
	Gemiddelde CE-cijfer	
17	SEGEM	F2.1
	Gemiddelde SE-cijfer	
18	CYFLYSTGEM	F2.1
	Gemiddelde cijferlijst	
19	NL_SE	F2.1
	Nederlands, cijfer schoolexamen	
20	NL_CE	F2.1
	Nederlands, cijfer centraal schriftelijk examen	
21	ENG_SE	F2.1
	Engels, cijfer schoolexamen	
22	ENG_CE	F2.1
	Engels, cijfer centraal schriftelijk examen	
23	DUI_SE	F2.1
	Duits, cijfer schoolexamen	
24	DUI_CE	F2.1
	Duits, cijfer centraal schriftelijk examen	
25	FR_SE	F2.1
	Frans, cijfer schoolexamen	
26	FR_CE	F2.1
	Frans, cijfer centraal schriftelijk examen	
27	AARD_SE	F2.1
	Aardrijkskunde, cijfer schoolexamen	
28	AARD_CE	F2.1
	Aardrijkskunde, cijfer centraal schriftelijk examen	
29	MAATSCHII_SE	F2.1
	Maatschappijleer II, cijfer schoolexamen	
30	MAATSCHII_CE	F2.1
	Maatschappijleer II, cijfer centraal schriftelijk examen	

Nr	*Variabele en omschrijving*	*Form.*
31	GESCHSTAAT_SE Geschiedenis en staatsinrichting, cijfer schoolexamen	F2.1
32	GESCHSTAAT_CE Geschiedenis en staatsinrichting, cijfer centraal schriftelijk examen	F2.1
33	ECO_SE Economie, cijfer schoolexamen	F2.1
34	ECO_CE Economie, cijfer centraal schriftelijk examen	F2.1
35	WISK_SE Wiskunde, cijfer schoolexamen	F2.1
36	WISK_CE Wiskunde, cijfer centraal schriftelijk examen	F2.1
37	NATSCH1_SE Natuur- en scheikunde I, cijfer schoolexamen	F2.1
38	NATSCH1_CE Natuur- en scheikunde I, cijfer centraal schriftelijk examen	F2.1
39	BIOL_SE Biologie, cijfer schoolexamen	F2.1
40	BIOL_CE Biologie, cijfer centraal schriftelijk examen	F2.1
41	GESCH_SE Geschiedenis, cijfer schoolexamen	F2.1
42	GESCH_CE Geschiedenis, cijfer centraal schriftelijk examen	F2.1
43	WISKA_SE Wiskunde A, cijfer schoolexamen	F2.1
44	WISKA_CE Wiskunde A, cijfer centraal schriftelijk examen	F2.1
45	WISKB_SE Wiskunde B, cijfer schoolexamen	F2.1
46	WISKB_CE Wiskunde B, cijfer centraal schriftelijk examen	F2.1
47	WISKC_SE Wiskunde C, cijfer schoolexamen	F2.1
48	WISKC_CE Wiskunde C, cijfer centraal schriftelijk examen	F2.1
49	NAT_SE Natuurkunde, cijfer schoolexamen	F2.1
50	NAT_CE Natuurkunde, cijfer centraal schriftelijk examen	F2.1
51	SCHEIK_SE Scheikunde, cijfer schoolexamen	F2.1
52	SCHEIK_CE Scheikunde, cijfer centraal schriftelijk examen	F2.1
53	LAT_SE Latijn, cijfer schoolexamen	F2.1
54	LAT_CE	F2.1

Nr	Variabele en omschrijving	Form.
	Latijn, cijfer centraal schriftelijk examen	
55	GR_SE	F2.1
	Grieks, cijfer schoolexamen	
56	GR_CE	F2.1
	Grieks, cijfer centraal schriftelijk examen	
57	MAATSCHK_SE	F2.1
	Maatschappijkunde, cijfer schoolexamen	
58	MAATSCHK_CE	F2.1
	Maatschappijkunde, cijfer centraal schriftelijk examen	

Toelichting:
De NCOVO-bestanden bevatten de gegevens van leerlingen die in minimaal een van de cohorten voorkomen. De cohorten vormen de ruggengraat van het NCO waar op basis van de koppelsleutel (RINPERSOON en RINPERSOONS) de inschrijvings- en examengegevens van leerlingen in het voortgezet onderwijs aangekoppeld kunnen worden. Bij de cohorten gaat het vooral om de VO-cohorten. Vanaf schooljaar 2018/'19 kan echter voor het eerst ook informatie aan de PO-cohorten gekoppeld worden. Dan zitten de leerlingen uit het PO-cohort 2010 grotendeels in de eerste klas van het voortgezet onderwijs.

Voor de inschrijvingen is geselecteerd op HOOFDINSCHR = '01' (hoofdinschrijving op 1 oktober in VO) of '11' (hoofdinschrijving, niet op 1 oktober in VO). Hierdoor zijn ook de voor- en na - inschrijvingen meegenomen en dus niet alleen de leerlingen die op 1 oktober in het vo stonden ingeschreven. De variabele HOOFDINSCHR is afkomstig uit het bestand *ONDERWIJSINSCHRTAB*.

De examengegevens (CEGEM tot en met MAATSCHK_CE) zijn samengesteld uit de recent ontwikkelde *EXAMVOVAKTAB*. De informatie die in het NCO is opgenomen is echter niet direct uit dit bestand te herleiden. Zo zijn in dat bestand de vakken het uitgangspunt en komt een leerling meerdere keren in het bestand voor, afhankelijk van het aantal vakken waarin examen is gedaan. Voor het NCO vormt de leerling het uitgangspunt en zijn voor de verschillende vakken afzonderlijke variabelen opgenomen.

De variabele CEGEM bevat het gemiddelde van alle CE-cijfers van alle vakken waarin ook een SE wordt afgenomen. De variabele SEGEM bevat het gemiddelde van alle SE-cijfers van alle vakken waarin ook een CE wordt afgenomen. De variabele CYFLYSTGEM bevat het eindlijstgemiddelde van alle vakken waarvoor een cijfer beschikbaar is. Daarnaast zijn in het bestand alle voor de verschillende schoolsoorten verplichte vakken en verplichte sectorvakken opgenomen met het cijfer voor het schoolexamen en voor het centraal examen. Deze informatie is alleen beschikbaar voor leerlingen die in een gegeven jaar voorkomen als examenkandidaat in het bestand *EXAMVOTAB*. Vanaf NCOVODEF2017 is het examenvak Maatschappijkunde toegevoegd.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de examengegevens. De NCOVODEF bestanden bevatten definitieve inschrijvingsgegevens en voor het laatste jaar voorlopige examengegevens. Deze voorlopige examengegevens worden echter als definitief beschouwd. Het NCOVOVRL bestand bevat voorlopige inschrijvingsgegevens en geen examengegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB*, *EXAMVOTAB* en *EXAMVOVAKTAB*.

Speciaal (voortgezet) onderwijs (NCOSO)
Inschrijvingsgegevens van leerlingen in het speciaal (voortgezet) onderwijs op persoonsniveau

Bestandsindeling: NCOSODEF en NCOSOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	WECTYPEPO Het soort onderwijs dat de leerling in het speciaal (voortgezet) onderwijs volgt	A3

Toelichting:
In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen.

Omdat het alleen gaat om leerlingen die instromen vanuit het (speciaal) basisonderwijs of het voortgezet onderwijs, is het op basis van deze bestanden niet mogelijk om een overzicht te maken van alle leerlingen in het speciaal (voortgezet) onderwijs.

De NCOSODEF bestanden bevatten definitieve inschrijvingsgegevens, het NCOSOVRL bestand bevat voorlopige inschrijvingsgegevens. Zie voor meer informatie K:\onderwijs*INSCHRWECTAB*.

Voortgezet algemeen volwassenenonderwijs (NCOVAVO)
Inschrijvings- en diplomagegevens van vavo-leerlingen op persoonsniveau

Bestandsindeling NCOVAVODEF en NCOVAVOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	ONDERWIJSSOORTVAVO Onderwijssoort	A1
4	ONDERWIJSSOORTVAVODIPL Onderwijssoort van het diploma	A1

Toelichting:
Voor de inschrijvingen is geselecteerd op HOOFDINSCHR = ‘07’ (hoofdinschrijving op 1 oktober in VE) of ‘17’ (hoofdinschrijving, niet op 1 oktober in VE). Hierdoor zijn ook de voor- en na-inschrijvingen meegenomen en dus niet alleen de leerlingen die op 1 oktober in het vavo stonden ingeschreven. Uitgesloten zijn leerlingen die binnen het vavo Educatie volgen (TYPEONDERWIJS <> ‘42’). Educatie omvat cursussen taal, rekenen en digitale vaardigheden. De variabele HOOFDINSCHR is afkomstig uit *ONDERWIJSINSCHRTAB*.

In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen.

In de bestanden kunnen leerlingen voorkomen die wel een vavo-diploma hebben, maar geen vavo-inschrijving. Dat kan als leerlingen in hetzelfde jaar waarin ze een vavo-diploma hebben behaald een inschrijving in het mbo of hoger onderwijs hadden. Maar het komt ook voor dat leerlingen een vavo-diploma halen zonder vavo-inschrijving.

Omdat het gaat om leerlingen die instromen vanuit het (speciaal) basisonderwijs of het voortgezet onderwijs, is het op basis van deze bestanden niet mogelijk om een overzicht te maken van alle leerlingen in het vavo.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de diplomagegevens. De NCOVAVODEF bestanden bevatten definitieve inschrijvingsgegevens en voor het laatste jaar voorlopige diplomagegevens. Deze voorlopige diplomagegevens worden echter als definitief beschouwd. Het NCOVAVOVRL bestand bevat voorlopige inschrijvingsgegevens en geen diplomagegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB* en *DIPLOMAVAVOTAB*.

Middelbaar beroepsonderwijs (NCOMBO)
Inschrijvings- en diplomagegevens van mbo-studenten op persoonsniveau

Bestandsindeling NCOMBODEF en NCOMBOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	BRINMBO_versl Code van de onderwijsinstelling van de inschrijving in het mbo - versleuteld	A4
4	TYPEONDERWIJSMBO Het type onderwijs waarbinnen de inschrijving valt	A2
5	CREBOCODE Opleidingscode van de inschrijving	A5
6	ISCEDF2013BROADFIELDNLMBO Opleidingsrichting van de inschrijving	A2
7	NIVEAUMBO Opleidingsniveau van de inschrijving	A1
8	SECTORMBO Sector van de inschrijving	A1
9	INSCHRAANTALMBO Aantal mbo-inschrijvingen	F7
10	BRINDIPL_versl Code van de onderwijsinstelling van het diploma in het mbo - versleuteld	A4
11	TYPEMBODPL Opleidingsvorm van het diploma	A1
12	CREBOCODEDIPL Opleidingscode van het diploma	A5
13	ISCEDF2013BROADFIELDNLMBODIPL Opleidingsrichting van het diploma	A2
14	NIVEAUMBODIPL Niveau van het diploma	A1
15	SECTORMBODIPL Sector van het diploma	A1
16	DIPLAANTALMBO Aantal behaalde mbo-diploma's	F7

Toelichting:
Voor de hoofdinschrijving is geselecteerd op HOOFDINSCHR = ‘02’ (hoofdinschrijving op 1 oktober in MBO), ‘12’ (hoofdinschrijving, niet op 1 oktober in MBO). Hierdoor zijn ook de voor- en na-inschrijvingen meegenomen en dus niet alleen de leerlingen die op 1 oktober in het mbo stonden ingeschreven. De variabele HOOFDINSCHR is afkomstig uit *ONDERWIJSINSCHRTAB*.

Voor het hoofddiploma is geselecteerd op SELECTIEHOOFDDIPLMBO = ‘1’, afkomstig uit *DIPLOMAMBOTAB*.

Met behulp van de variabele INSCHRAANTALMBO is weergegeven voor hoeveel (neven-)opleidingen binnen het mbo een persoon staat ingeschreven. Hiervoor zijn niet alleen de hoofdinschrijving, maar ook de overige inschrijvingen (HOOFDINSCHR = '99') in combinatie met ONDERWIJSSOORT = 'MBO' (uit OPLEIDINGSNRREF) meegenomen. Met behulp van de variabele DIPLAANTALMBO is weergegeven voor hoeveel (neven-)opleidingen binnen het mbo een persoon een diploma heeft behaald in het betreffende studiejaar. Hiervoor zijn naast het hoofddiploma ook alle nevendiploma's meegenomen uit *DIPLOMAMBOTAB*.

In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen.

In de bestanden kunnen leerlingen voorkomen die wel een mbo-diploma, maar geen mbo-inschrijving hebben. Dat kan als leerlingen in hetzelfde jaar waarin ze een mbo-diploma hebben behaald een inschrijving in het hoger onderwijs hadden. Maar het komt ook voor dat leerlingen een mbo-diploma halen zonder mbo-inschrijving.

Omdat het gaat om leerlingen die instromen vanuit het (speciaal) basisonderwijs of het voortgezet onderwijs, is het op basis van deze bestanden niet mogelijk om een overzicht te maken van alle leerlingen in het mbo.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de diplomagegevens. De NCOMBODEF bestanden bevatten definitieve inschrijvingsgegevens en voor het laatste jaar voorlopige diplomagegevens. Deze voorlopige diplomagegevens worden echter als definitief beschouwd. Het NCOMBOVRL bestand bevat voorlopige inschrijvingsgegevens en geen diplomagegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB* en *DIPLOMAMBOTAB*.

Hoger beroepsonderwijs (NCOHBO)
Inschrijvings- en diplomagegevens van hbo-studenten op persoonsniveau

Bestandsindeling NCOHBODEF en NCOHBOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	BRINHBO_versl Code van de onderwijsinstelling van de inschrijving in het hbo - versleuteld	A4
4	TYPEONDERWIJSHBO Het type onderwijs waarbinnen de inschrijving valt	A2
5	CROHOCODEHBO Opleidingscode hbo	A5
6	CROHOONDHBO Croho-onderdeel van de inschrijving in het hbo	A1
7	INSCHRAANTALHBO Aantal hbo-inschrijvingen	F7
8	SOORTDIPLSOORTHOHBO Soort diploma behaald in het hbo	A2
9	CROHOCODEHBODIPL Opleidingscode van het diploma	A5
10	CROHOONDHBODIPL Croho-onderdeel van het diploma	A1
11	DIPLAANTALHBO Aantal behaalde hbo-diploma's	F7

Toelichting:
Voor de hoofdinschrijving is geselecteerd op HOOFDINSCHR = '03' (hoofdinschrijving op 1 oktober in HO én HBO), 13' (hoofdinschrijving, niet op 1 oktober in HO én HBO), '05' (hoofdinschrijving in HBO op 1 oktober) en '15' (hoofdinschrijving, niet op 1 oktober in HBO). Hierdoor zijn ook voor- en na-inschrijvingen meegenomen en dus niet alleen de leerlingen die op 1 oktober in het hbo stonden ingeschreven. De variabele HOOFDINSCHR is afkomstig uit *ONDERWIJSINSCHRTAB*.

Voor het hoofddiploma is geselecteerd op SOORTDIPLSOORTHO = '03' (hoofd-bachelor-diploma), '05' (hoofd-master-diploma) of '13' (hoofddiploma associate degree) en vervolgens ontdubbeld, waarbij het hoogste hoofddiploma voorgaat. Deze variabele SOORTDIPLSOORTHO is afkomstig uit *HDIPLOMAHBOTAB*.

Met behulp van de variabele INSCHRAANTALHBO is weergegeven voor hoeveel (neven-)opleidingen binnen het hbo een persoon staat ingeschreven. Hiervoor zijn niet alleen de hoofdinschrijving, maar ook de overige inschrijvingen (HOOFDINSCHR = '99') in combinatie met ONDERWIJSSOORT = 'HBO' (uit OPLEIDINGSNRREF) meegenomen. Met behulp van de variabele DIPLAANTALHBO is weergegeven voor hoeveel (neven-)opleidingen binnen het hbo een persoon een diploma heeft behaald in het betreffende studiejaar. Hiervoor zijn naast de hbo hoofd-

bachelor-diploma, hoofd-master diploma en hoofddiploma associate degree ook de betreffende neven hbo-diploma's meegenomen en wel via SOORTDIPLSOORTHO (= '03', '04', '05', '06', '13' en '14'), afkomstig uit *DIPLOMAHOTAB*.

In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen. Omdat het gaat om leerlingen die instromen vanuit het (speciaal) basisonderwijs of het voortgezet onderwijs, is het op basis van deze bestanden niet mogelijk om een overzicht te maken van alle leerlingen in het hbo. Zodoende missen ook de personen die vanuit het buitenland in Nederland komen studeren.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de diplomagegevens. De NCOHBODEF bestanden bevatten definitieve inschrijvingsgegevens en voor het laatste jaar voorlopige diplomagegevens. Deze voorlopige diplomagegevens worden echter als definitief beschouwd. Het NCOHBOVRL bestand bevat voorlopige inschrijvingsgegevens en geen diplomagegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB*, *HDIPLOMAHOTAB* en *DIPLOMAHOTAB*.

Wetenschappelijk onderwijs (NCOWO)
Inschrijvings- en diplomagegevens van wo-studenten op persoonsniveau

Bestandsindeling: NCOWODEF en NCOWOVRL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	BRINWO_versl Code van de onderwijsinstelling van de inschrijving in het wo - versleuteld	A4
4	TYPEONDERWIJSWO Het type onderwijs waarbinnen de inschrijving valt	A2
5	CROHOCODEWO Opleidingscode wo	A5
6	CROHOONDWO Croho-onderdeel van de inschrijving in het wo	A1
7	INSCHRAANTALWO Aantal wo-inschrijvingen	F7
8	SOORTDIPLSOORTHOWO Soort diploma behaald in het wo	A2
9	CROHOCODEWODIPL Opleidingscode van het diploma	A5
10	CROHOONDWODIPL Croho-onderdeel van het diploma	A1
11	DIPLAANTALWO Aantal behaalde wo-diploma's	F7

Toelichting:
Voor de hoofdinschrijving is geselecteerd op HOOFDINSCHR = '04' (hoofdinschrijving op 1 oktober in HO én WO), '06' (hoofdinschrijving op 1 oktober in WO), '14' (hoofdinschrijving, niet op 1 oktober in HO én WO) of '16' (hoofdinschrijving, niet op 1 oktober in WO). Hierdoor zijn ook de voor- en na-inschrijvingen meegenomen en dus niet alleen de leerlingen die op 1 oktober in het wo stonden ingeschreven. De variabele HOOFDINSCHR is afkomstig uit *ONDERWIJSINSCHRTAB*.

Voor het hoofddiploma is geselecteerd op SOORTDIPLSOORTHO = '03' (hoofd-bachelor-diploma), '05' (hoofd-master-diploma) of '09' (hoofddiploma beroepsfase/postmaster) en vervolgens ontdubbeld, waarbij het hoogste hoofddiploma voorgaat. Deze variabele SOORTDIPLSOORTHO is afkomstig uit *HDIPLOMAWOTAB*.

Met behulp van de variabele INSCHRAANTALWO is weergegeven voor hoeveel (neven-)opleidingen binnen het wo een persoon staat ingeschreven. Hiervoor zijn niet alleen de hoofdinschrijving, maar ook de overige inschrijvingen (HOOFDINSCHR = '99') in combinatie met ONDERWIJSSOORT = 'WO' (uit OPLEIDINGSNRREF) meegenomen. Met behulp van de variabele DIPLAANTALWO is weergegeven voor hoeveel (neven-)opleidingen binnen het wo een persoon een diploma heeft behaald in het betreffende studiejaar. Hiervoor zijn naast de wo hoofd-bachelor-diploma, hoofd-master diploma en hoofddiploma beroepsfase/postmaster ook de betreffende neven wo-diploma's meegenomen en wel via SOORTDIPLSOORTHO (= '03', '04', '05', '06', '09' en '10'), afkomstig uit *DIPLOMAHOTAB*.

In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen.

Omdat het gaat om leerlingen die instromen vanuit het (speciaal) basisonderwijs of het voortgezet onderwijs, is het op basis van deze bestanden niet mogelijk om een overzicht te maken van alle leerlingen in het wo. Zodoende missen ook de personen die vanuit het buitenland in Nederland komen studeren.

De inschrijvingsgegevens zijn een jaar eerder beschikbaar dan de diplomagegevens. De NCOWODEF bestanden bevatten definitieve inschrijvingsgegevens en voor het laatste jaar voorlopige diplomagegevens. Deze voorlopige diplomagegevens worden echter als definitief beschouwd. Het NCOWOVRL bestand bevat voorlopige inschrijvingsgegevens en geen diplomagegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB*, *HDIPLOMAWOTAB* en *DIPLOMAHOTAB*.

Schoolgegevens

Schoolgegevens: primair onderwijs (NCOBRINPO)
Schoolgegevens van alle leerlingen in het (speciaal) basisonderwijs op persoonsniveau

Bestandsindeling NCOBRINPO

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	WPOBRIN_versl Code van de onderwijsinstelling van de inschrijving in het po - versleuteld	A4
4	WPOBRINVEST Het onderwijsvestigingsnummer van de vestiging waar de leerling het (speciaal) basisonderwijs volgt	A2
5	POAANTALBRINBESTUUR Aantal scholen dat onder een bestuur valt	F7
6	POGEMBRINVEST Gemeentecode van de onderwijsvestiging	A4
7	POSTEDGEMBRINVEST Stedelijkheid van de gemeente waarin zich de onderwijsvestiging bevindt	A1
8	PODENOMINATIE Denominatie van de school	A3
9	POSCHOOLGROOTTE Aantal leerlingen aan de onderwijsvestiging	F7
10	PERCGEWICHT Percentage gewichtenleerlingen binnen de onderwijsvestiging	F3
11	PERCACHTERSTAND Percentage achterstandsleerlingen binnen de onderwijsvestiging	F3
12	GEMACHTERSTAND Percentage gewichtenleerlingen binnen de onderwijsvestiging	F3

Toelichting:
De NCOBRINPO-bestanden bevatten de schoolgegevens van leerlingen die in minimaal een van de PO-cohorten voorkomen. De cohorten vormen de ruggengraat van het NCO waar op basis van de koppelsleutel (RINPERSOON en RINPERSOONS) de schoolgegevens van leerlingen in het (speciaal) basisonderwijs aangekoppeld kunnen worden.

De schoolgegevens zijn op persoonsniveau in de bestanden opgenomen en kunnen indien gewenst door de onderzoeker zelf geaggregeerd worden op BRIN(VEST)-niveau.

De variabele POAANTALBRINBESTUUR is het geaggregeerde aantal scholen op basis van het NRBEVOEGDGEZAGWPOWEC, afkomstig uit *BRINWPOWECTAB*.

De variabele POSCHOOLGROOTTE betreft het geaggregeerde aantal leerlingen binnen de onderwijsvesting (WPOBRINVEST).

De variabele PERCGEWICHT is mede met behulp van de variabele WPOGEWICHT, afkomstig uit INSCHRWPOTAB gemaakt. Deze variabele is vanaf NCOBRINPO2019 niet meer beschikbaar is.

De variabelen PERCACHTERSTAND en GEMACHTERSTAND zijn beschikbaar vanaf NCOBRINPO2018. Deze variabelen zijn mede met behulp van gegevens uit de, op de CBS maatwerkpagina beschikbare bestanden "Achterstandsscores-scholen-JJJJ.xlsx" gemaakt. Daarnaast is van de ook in de RA beschikbare bestanden "Verwachte_onderwijsscores_JJJJ" gebruik gemaakt.

De NCOBRINPODEF bestanden bevatten definitieve schoolgegevens, het NCOBRINPOVRL bestand bevat voorlopige schoolgegevens.

Als een leerling in jaar JJJJ is gestart in het primair onderwijs, dan zijn de po schoolgegevens ook vanaf dat jaar beschikbaar. Zie voor meer informatie ook K:\onderwijs*INSCHRWPOTAB* en *BRINWPOWECTAB*.

Schoolgegevens: voortgezet onderwijs (NCOBRINVO)
Schoolgegevens van alle leerlingen in het voortgezet onderwijs op persoonsniveau

Bestandsindeling NCOBRINVO

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	VOBRIN_versl Code van de onderwijsinstelling van de inschrijving in het voortgezet onderwijs - versleuteld	A4
4	VOBRINVEST Het onderwijsvestigingsnummer van de vestiging waar de leerling het voortgezet onderwijs volgt	A2
5	VOAANTALBRINBESTUUR Aantal scholen dat onder een bestuur valt	F7
6	VOGEMBRINVEST Gemeentecode van de onderwijsvestiging	A4
7	VOSTEDGEMBRINVEST Stedelijkheid van de gemeente waarin zich de onderwijsvestiging bevindt	A1
8	VODENOMINATIE Denominatie van de school	A3
9	VOSCHOOLGROOTTE Aantal leerlingen aan de onderwijsvestiging	F7
10	PERCAPC Percentage leerlingen binnen de onderwijsvestiging dat in een Armoede Probleem Cumulatiegebied woont	F3
11	AFDGROOTTEVMBOBK Aantal leerlingen dat vmbo-b of vmbo-k volgt vanaf leerjaar 3 aan de onderwijsvestiging	F5
12	AFDGROOTTEVMBOGT	F5

	Aantal leerlingen dat vmbo-g of vmbo-t volgt vanaf leerjaar 3 aan de onderwijsvestiging	
13	AFDGROOTTEHAVO Aantal leerlingen dat havo volgt vanaf leerjaar 3 aan de onderwijsvestiging	F5
14	AFDGROOTTEVWO Aantal leerlingen dat vwo volgt vanaf leerjaar 3 aan de onderwijsvestiging	F5
15	AFDGROOTTEOVERIG Aantal leerlingen in een gemengde klas of dat praktijkonderwijs volgt aan de onderwijsinstelling	F5
16	BRINBESTUUR Onderwijsinstellingsstructuur	A3
17	BRINVESTSTRUCTUUR Onderwijsvestigingsstructuur	A3

Toelichting:
De NCOBRINVO-bestanden bevatten de schoolgegevens van leerlingen die in minimaal een van de cohorten voorkomen. De cohorten vormen de ruggengraat van het NCO waar op basis van de koppelsleutel (RINPERSOON en RINPERSOONS) de schoolgegevens van leerlingen in het voortgezet onderwijs aangekoppeld kunnen worden. Bij de cohorten gaat het vooral om de VO-cohorten. Vanaf schooljaar 2018/'19 kan echter voor het eerst ook informatie aan de PO-cohorten gekoppeld worden. Dan zitten de leerlingen uit het PO-cohort 2010 grotendeels in de eerste klas van het voortgezet onderwijs.

De schoolgegevens zijn op persoonsniveau in de bestanden opgenomen en kunnen indien gewenst door de onderzoeker zelf geaggregeerd worden op BRIN(VEST)-niveau.

De variabele VOAANTALBRINBESTUUR is het geaggregeerde aantal scholen op basis van het NRBEVOEGDGEZAG, afkomstig uit *BRINTAB*.

De variabele VOSCHOOLGROOTTE betreft het geaggregeerde aantal leerlingen binnen de onderwijsvesting (VOBRINVEST).

De variabele PERCAPC is gemaakt met behulp van de variabele APC, zoals opgenomen in de NCOACHTERGRONDVARIABEL bestanden (zie hierna).

De NCOBRINVODEF bestanden bevatten definitieve schoolgegevens, het NCOBRINVOVRL bestand bevat voorlopige schoolgegevens. Zie voor meer informatie K:\onderwijs*ONDERWIJSINSCHRTAB* en *BRINTAB*.

Achtergrondgegevens

Achtergrondgegevens: stabiel (NCOACHTERGRONDSTABIEL)
Demografische gegevens op persoonsniveau, van alle personen uit de PO- en VO-cohorten

Bestandsindeling NCOACHTERGRONDSTABIEL

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	GESLACHT Geslacht	A1
4	GEBDATUM Geboortedatum	A8
5	HERKOMST Migratieachtergrond	A1
6	GENERATIE Generatie migratieachtergrond	A1

Toelichting:
Omdat het gaat om stabiele gegevens is het niet nodig om dit bestand voor elk jaar afzonderlijk te maken, maar volstaat één actueel bestand. Elk jaar, als er een nieuw cohort bij komt, wordt het stabiele achtergrondbestand aangevuld. Bij de variabele GEBDATUM gaat het vanuit privacyoverwegingen steeds om de eerste dag van de maand.

Zie voor meer informatie K:\bevolking*GBAPERSOONTAB*

Achtergrondgegevens: variabel (NCOACHTERGRONDVARIABEL)
Demografische-, regio-, inkomens- en opleidingsgegevens op persoonsniveau, van (de ouders van) alle personen uit de PO- en VO-cohorten

Bestandsindeling NCOACHTERGRONDVARIABELDEF, NCOACHTERGRONDVARIABELVRL1 en NCOACHTERGRONDVARIABELVRL0

Nr	*Variabele en omschrijving*	*Form.*
1	RINPERSOONS Soort rinpersoon	A1
2	RINPERSOON Samen met rinpersoons is dit de persoon	A9
3	OUDERLIJKESTRUCTUUR Geeft ouderlijke structuur aan op 31 december	A1
4	SAMENWONINGOUDERS Geeft aan of juridische ouders op 31 december samenwonen en zo niet, waarom niet	A1
5	AANTALKINDHH Aantal personen met plaats (positie) in het huishouden 'thuiswonend kind' op 31 december	F2
6	TYPHH Type huishouden op 31 december	A1

Nr	Variabele en omschrijving	Form.
7	WIJKBUURTCODEJJJJ Wijkbuurtcode op 31 december	A8
8	APC Geeft aan of vo-leerling in een Armoede Probleem Cumulatiegebied woont op 1 oktober	F1
9	SECJ Sociaaleconomische categorie in JJJJ	A2
10	SECJPA Sociaaleconomische categorie van de juridische vader in JJJJ	A2
11	SECJMA Sociaaleconomische categorie van de juridische moeder in JJJJ	A2
12	INKHP100HBESTPA Percentielgroep besteedbaar inkomen van het huishouden van de juridische vader in JJJJ	F3
13	INKHP100HBESTMA Percentielgroep besteedbaar inkomen van het huishouden van de juridische moeder in JJJJ	F3
14	VEHP100HVERMPA Percentielgroep vermogen van het huishouden van de juridische vader op 1 januari JJJJ+1 (t/m 2010) of JJJJ (vanaf 2011)	F3
15	VEHP100HVERMMA Percentielgroep vermogen van het huishouden van de juridische moeder op 1 januari JJJJ+1 (t/m 2010) of JJJJ (vanaf 2011)	F3
16	OPLNIVSOI2016AGG3HBPA Hoogst behaalde opleidingsniveau van de juridische vader op 1 oktober JJJJ; 8-deling	A3
17	OPLNIVSOI2016AGG3HBMA Hoogst behaalde opleidingsniveau van de juridische vader op 1 oktober JJJJ; 8-deling	A3
18	GEWICHTHOOGSTEOPLPA Ophooggewicht horend bij opleidingsniveau juridische vader; 1 bij registerrecords, <>1 bij ebb-records	DOT14.6
19	GEWICHTHOOGSTEOPLMA Ophooggewicht horend bij opleidingsniveau juridische moeder; 1 bij registerrecords, <>1 bij ebb-records	DOT14.6
20	OPLNIVBRONPA Bron waarin opleidingsgegeven juridische vader beschikbaar is	A3
21	OPLNIVBRONMA Bron waarin opleidingsgegeven juridische moeder beschikbaar is	A3

Toelichting:
In de bestanden zijn de gegevens opgenomen van leerlingen die in voorgaande jaren in het (speciaal) basisonderwijs of het voortgezet onderwijs hebben gezeten, en dus in een van de cohorten voorkomen.

De gegevens over het huishouden zijn afkomstig uit *GBAHUISHOUDENSBUS* en *GBASAMENWONERBUS*. De wijkbuurtcode is o.a. met behulp van *VSLGWBTAB* gemaakt.

De APC-postcodegebieden worden door het CBS afgeleid en aangeleverd aan DUO, waarna het deze worden gekoppeld aan de VO-leerlingen in het betreffende 1 cijferbestand. Voor meer informatie zie https://www.cbs.nl/nl-nl/maatwerk/2020/07/armoedeprobleem-cumulatie-gebieden-2017

Inkomensgegevens van de huishoudens zijn beschikbaar voor de juridische ouders die op een bepaald peilmoment tot de bevolking van Nederland behoren. Na revisie van de inkomensstatistiek (vanaf verslagjaar 2011) is die peildatum 1 januari JJJJ, vóór revisie 1 januari JJJJ + 1. De inkomensgegevens betreffen in beide gevallen die van het jaar JJJJ.

Dit geldt ook voor de SECJ-gegevens. Vanaf 2011 gaat het voor de vaststelling van de Nederlandse bevolking om het peilmoment 1 januari JJJJ en tot en met 2010 om het peilmoment 1 januari JJJJ + 1. Het betreft in beide gevallen de gegevens van jaar JJJJ.

Voor de vermogensgegevens geldt dat ook dat het peilmoment voor de vaststelling van de populatie verspringt; vóór revisie – dus tot en met 2010 - 1 januari JJJJ + 1 en daarna, vanaf 2011, 1 januari JJJJ. En vóór 2011 gaat om het vermogen op 1 januari JJJJ + 1 en vanaf 2011 om het vermogen op 1 januari JJJJ.

De inkomens-, secj- en vermogensgegevens zijn vanaf 2011 afkomstig uit *INPATAB*, *INHATAB* en *VEHTAB*. Vóór 2011 is (ook) gebruik gemaakt van *INTEGRAAL PERSOONLIJK INKOMEN* en *INTEGRAAL HUISHOUDENS INKOMEN*.

De **opleidingsniveaugegevens** van de juridische ouders zijn afkomstig uit *HOOGSTEOPLTAB*. Ze zijn gebaseerd op de variabele OPLNIVSOI2016AGG4HBMETNIRWO. De variabelen OPLNIVBRONPA- en MA zijn vanaf 2016 beschikbaar.

EXTRA INFORMATIE BIJ GEBRUIK VAN HOOGSTE OPLEIDINGSNIVEAU VAN DE OUDERS
De aanvullende informatie over deze gegevens en bijbehorende hulpmiddelen is nog in ontwikkeling, maar op dit moment is de volgende informatie beschikbaar:

1. Inleiding

Deze informatie heeft betrekking op het gebruik van de variabelen opleidingsniveau van moeder en van vader in het NCO. Deze variabelen zijn afkomstig uit het bestand HOOGSTEOPLJJJJTABVV. Dat bestand is een combinatie van integrale registerdata en steekproef-data (enquête EBB) en bevat gewichten om te wegen en op te hogen naar de gehele bevolking. Daarom zijn er bij gebruik van dit bestand verschillende aandachtspunten. Deze informatie is bedoeld als aanvulling op de informatie in het documentatierapport van HOOGSTEOPLJJJJTABVV en heeft betrekking op het gebruik van het opleidingsniveau van ouders.

We maken onderscheid tussen het opleidingsniveau van ouders als doelvariabele en het opleidingsniveau van ouders als achtergrondkenmerk bij een andere doelvariabele. Voorbeeld van het opleidingsniveau als doelvariabele: van alle 14-jarigen wordt bepaald hoeveel procent van de leerlingen laag opgeleide ouders heeft, hoeveel procent middelbaar opgeleide ouders en hoeveel procent hoog opgeleide ouders. Het opleidingsniveau zit dan alleen in de teller van het percentage.
Voorbeeld van het opleidingsniveau van ouders als achtergrondkenmerk bij een andere doelvariabele: van alle 14-jarigen wordt bepaald in welke onderwijssoort zij zitten en er wordt een vergelijking gemaakt van deze verdeling (waaronder bv. het aandeel op het vwo) tussen kinderen met laag

opgeleide ouders, kinderen met middelbaar opgeleide ouders en kinderen met hoog opgeleide ouders. Het opleidingsniveau zit dan in de teller én in de noemer van het percentage (bv. kinderen van laag opgeleide ouders die op vwo zitten in de teller en alle kinderen van laag opgeleide ouders in de noemer).

2. Ophooggewichten in HOOGSTEOPLTABVV

Omdat het bestand HOOGSTEOPLJJJJTABVV opleidingsinformatie bevat over een deel van de bevolking en dit deel niet representatief is voor de totale bevolking, heeft elke persoon een ophooggewicht. Registerwaarnemingen hebben een gewicht van 1; EBB-waarnemingen hebben een gewicht ongelijk aan 1 (in veruit de meeste gevallen is het groter dan 1). Dit gewicht dient te worden gebruikt bij het maken van cijfers met het bestand om te zorgen voor representativiteit. Dit geldt ook als het opleidingsniveau niet de doelvariabele is, maar achtergrondkenmerk bij een andere doelvariabele.

Bij het gebruik van bestanden van het NCO zijn de kinderen de populatie. We willen dan ook gewichten om op te hogen en te wegen naar de doelpopulatie kinderen van het betreffende onderzoek. Het gaat hier echter om het opleidingsniveau van de ouders, dus kunnen we het beste hun ophooggewichten gebruiken. Deze zijn echter niet zonder meer geschikt. De gewichten zijn namelijk gemaakt voor de (afzonderlijke) ouders, niet voor de kinderen van deze afzonderlijke ouders of 'ouderkoppels'. Omdat er steeds vaker onderzoek wordt gedaan naar het opleidingsniveau van ouders worden er momenteel binnen het CBS *intergenerationele gewichten* ontwikkeld.

Tot dat deze intergenerationele gewichten beschikbaar zijn voor onderzoekers, raden we de volgende methode aan. Dit zal in veel gevallen niet de beste methode zijn, maar het is wat minimaal nodig is voor de kwaliteit van de uitkomsten én kan door onderzoekers zelfstandig uitgevoerd worden. De methode bestaat uit twee stappen: (1) het bepalen van het opleidingsniveau en de startgewichten en vervolgens (2) het herschalen van de gewichten.

3. Bepalen van het opleidingsniveau en het startgewicht

We gaan uit van twee mogelijkheden. De eerste mogelijkheid is dat het in het onderzoek gaat het om het opleidingsniveau van moeder óf van vader (of van beide ouders apart). In die gevallen starten we met de ouders waarvoor het opleidingsniveau bekend is en hun gewicht. We kennen (per ouder) het opleidingsniveau en het gewicht toe aan het kind. De rest van deze paragraaf is dan niet relevant en er kan meteen worden doorgegaan met het herschalen.

De tweede mogelijkheid is dat het gaat om het hoogste opleidingsniveau van het ouderkoppel, dus als de ene ouder laag is opgeleid en de andere ouder middelbaar is opgeleid, is het hoogste opleidingsniveau van de ouders middelbaar opgeleid.
Het gaat in het NCO om de juridische ouders. Dat zijn er in het algemeen twee. Het kan dan voorkomen dat het opleidingsniveau van één van de ouders wel bekend is maar van de andere ouder niet. Er zijn dan twee mogelijkheden met beide voor- en nadelen: alleen het hoogste opleidingsniveau bepalen als het van beide ouders bekend is, of ook het opleidingsniveau bepalen als het maar van een van beide ouders bekend is. We raden de tweede optie aan. Deze heeft als grootste voordeel dat we dan het opleidingsniveau van meer ouders kunnen meenemen, waardoor de betrouwbaarheidsmarges kleiner

worden. Ook is het een minder selectieve groep dan bij de eerste optie, waardoor de representativiteit beter zal zijn. Uiteraard is het gemeten opleidingsniveau van sommige ouderkoppels dan te laag, namelijk in de gevallen dat het onbekende opleidingsniveau hoger is dan het bekende opleidingsniveau. Maar als we dat weten kunnen we daar rekening mee houden. We raden dan ook aan de betreffende gevolgen voor de uitkomsten te vermelden in rapportages:

- Als het opleidingsniveau van de ouders de doelvariabele is: de omvang van de groep kinderen met laag opgeleide ouders wordt enigszins overschat en de omvang van de groep kinderen met hoog opgeleide ouders wordt enigszins onderschat.
- Als het opleidingsniveau van de ouders het achtergrondkenmerk is bij een andere doelvariabele: eventuele verschillen in de doelvariabele tussen kinderen van lager opgeleide ouders en van de hoogst opgeleide ouders worden waarschijnlijk onderschat, omdat in de groepen met lager opleidingsniveau ook kinderen zitten waarvan de andere ouder hoger opgeleid is. Bijvoorbeeld: als blijkt dat kinderen van hoog opgeleide ouders vaker op het vwo zitten dan kinderen van laag opgeleide ouders, is dit verschil waarschijnlijk in werkelijkheid wat groter dan in de uitkomsten naar voren komt.

Deze twee effecten zijn minder sterk naarmate de kinderen jonger zijn en de ouders dus (gemiddeld) ook, waardoor van meer ouders het opleidingsniveau bekend is.

Vervolgens is het advies de startgewichten als volgt te bepalen:

1) Als niveau van een van beide ouders bekend: → gewicht van die ouder gebruiken.
2) Als niveau van beide ouders bekend en ongelijk: → gewicht gebruiken van de ouder met het hoogste niveau.
3) Als niveau van beide ouders bekend en gelijk:
   a) en beide uit register: → gewicht = 1.
   b) en één uit register en één uit EEB: → gewicht = 1.
   c) en beide uit EBB: → gewicht = gemiddelde v.d. twee gewichten.

Zowel het op deze manier bepaalde hoogste opleidingsniveau van de ouders als het startgewicht worden toegekend aan de kinderen. We spreken daarom in het vervolg van kinderen met een gewicht van 1 en kinderen met een gewicht ongelijk aan 1.

4. Herschalen van de gewichten

Als de startgewichten zijn bepaald, dan adviseren we deze te herschalen. We doen dit omdat de gewichten van de ouders niet zonder meer geschikt zijn voor de kinderen van afzonderlijke ouders of van ouderkoppels. Daarnaast is er in de meeste onderzoeken sprake van een selectieve populatie op basis van kenmerken die niet zijn meegenomen in het weegmodel van HOOGSTEOPLTABVV (dit is niet specifiek voor het werken met het opleidingsniveau van ouders). Herschalen kan er bij helpen om deze selectiviteit kleiner te maken.

Bij het herschalen worden alléén de gewichten van de EBB-waarnemingen aangepast. Hierbij wordt gebruik gemaakt van de bekende randtotalen in de te maken tabel. Bijkomend voordeel is, dat na wegen de aantallen in de verschillende opleidingsniveau-categorieën ook precies optellen tot deze randtotalen.

Het herschalen gaat, per tabel, als volgt:

Bepaal welke subgroepen je kunt onderscheiden (binnen de populatie van de tabel die je gaat maken) waarvan je het werkelijke totale aantal weet; neem alleen de subgroepen die je ook opneemt in je tabel. Programmeer de volgende stappen in het statistische pakket dat je gebruikt:

a) Bepaal per subgroep de omvang van de restpopulatie, dus de totale groep kinderen minus het aantal kinderen met een gewicht van 1 (ouders zijn register-waarneming).
b) Bepaal per subgroep het aantal kinderen waar je op uitkomt als je alleen de kinderen meeneemt met een gewicht ongelijk aan 1 (ouders zijn EBB-waarneming) en je het gewicht gebruikt.
c) Deel per subgroep a door b (de uitkomst kan ook kleiner dan 1 zijn).
d) Vermenigvuldig per subgroep alle gewichten ongelijk aan 1 met de uitkomst van c.

Voorbeeld: in de te maken tabel wordt van alle 14-jarigen in Nederland, per regio, getoond in welke onderwijssoort zij zitten, en dit wordt gegeven voor kinderen met respectievelijk laag, middelbaar en hoog opgeleide ouders (opleidingsniveau is achtergrondkenmerk voor andere doelvariabele). Er zijn 10 regio's en 7 onderwijssoorten; dan zijn er 70 subgroepen (kinderen in een bepaalde onderwijssoort in een bepaalde regio) waarvan we het werkelijke aantal weten.
Stel dat in één van de subgroepen 1000 kinderen zitten. Van 700 van deze kinderen hebben we het opleidingsniveau van de ouders met een gewicht van 1 (ouders zijn register-waarneming); van 50 kinderen hebben we het opleidingsniveau van de ouders met een gewicht ongelijk aan 1 (ouders zijn EBB-waarneming) en van 250 kinderen hebben we géén opleidingsniveau van de ouders. De 700 kinderen met een gewicht van 1 tellen voor zichzelf. De restpopulatie bestaat uit 1000 – 700 = 300 kinderen (a). De 50 kinderen met een gewicht ongelijk aan 1 tellen voor de restpopulatie (dus voor zichzelf én de kinderen waarvan we geen opleidingsniveau van de ouders hebben).
Als we nu alleen de 50 kinderen meenemen met een gewicht ongelijk aan 1 en we hun aantal bepalen na ophoging met dit gewicht, dan komen we bijvoorbeeld uit op 275 kinderen (b) (gemiddelde gewicht is 5,5). Als we dan a (300) delen door b (275) dan komt daar 1,09 uit (c). Als we vervolgens de gewichten ongelijk aan 1 allemaal vermenigvuldigen met 1,09 (gewichten worden dan gemiddeld 6,0) dan komen we na wegen van die 50 kinderen wel op 300 kinderen uit. En met de (700) kinderen met een gewicht van 1 erbij komen we dan voor de totale subgroep uit op 1000 kinderen, precies het werkelijke aantal. Hetzelfde wordt gedaan voor de andere 69 subgroepen.

Als de gewichten op deze manier zijn herschaald kan de tabel worden samengesteld.

5. Bepalen van de betrouwbaarheidsmarges

Omdat het bestand HOOGSTEOPLJJJJTABVV niet integraal is, hebben cijfers die met dit bestand worden samengesteld betrouwbaarheidsmarges. Het is raadzaam om deze marges te bepalen om een indicatie te krijgen van de kwaliteit van de uitkomsten, en daar rekening mee te houden bij het interpreteren van de cijfers. Voor aantallen en percentages die betrekking hebben op het opleidingsniveau van ouders kunnen de marges worden bepaald met de bij het NCO beschikbare 'Benaderingstool'. Ook daarin wordt er een onderscheid gemaakt tussen het opleidingsniveau als doelvariabele ('noemer bekend'), en het opleidingsniveau als achtergrondkenmerk bij een andere doelvariabele ('noemer geschat'). De tool is vooral bedoeld voor het maken van tabellen. Let extra goed op met het doel van (regressie-)analyses; neem in dat geval bij voorkeur contact op.
Zie verder de instructie op het voorblad van de tool. Er zijn ook ingevulde voorbeelden beschikbaar. Het betreft de excel bestanden 'RelMargeHOOGSTEOPL_v0.9.3.xlsx' en 'RelMargeHOOGSTEOPL_v0.9.3_voorbeeld2a/2b/3a.xlsx'.

Een VRL0-bestand bevat APC-gebied (definitief gegeven), andere achtergrondgegevens (voorlopige gegevens), maar nog géén sec-, inkomens-, vermogens- en opleidingsniveau-gegevens van de ouders.
Een VRL1-bestand bevat APC-gebied (definitief), (bijgestelde) andere achtergrondgegevens én voorlopige sec-, inkomens-, vermogens- en opleidingsniveau-gegevens van de ouders.
Een DEF-bestand bevat APC-gebied (definitief), andere achtergrondgegevens (definitief) én definitieve sec-, inkomens-, vermogens- en opleidingsniveau-gegevens van de ouders.

Zie voor meer informatie:
K:\bevolking*GBAHUISHOUDENSBUS* en *GBASAMENWONERBUS*
K:\bouwenwonen*VSLGWBTAB*
K:\inkomenbestedingen*INPATAB*, *INHATAB*, *VEHTAB*, *INTEGRAAL PERSOONLIJK INKOMEN* en *INTEGRAAL HUISHOUDENS INKOMEN*
en K:\onderwijs*HOOGSTEOPLTAB*.