



**Centraal Bureau voor de Statistiek**

Centrum voor Beleidsstatistiek

---

**Documentatierapport Voortgezet Onderwijs Cohort  
Leerlingen (VOCL) 1999V5**

*Datum:* 21 oktober 2010

## **Bronvermelding**

Publicatie van uitkomsten geschiedt door het onderzoeksbureau of de opdrachtgever op eigen titel. Verwijzing naar het CBS betreft uitsluitend het gebruik van de microdatabestanden. Dat wordt als volgt geformuleerd:

“Eigen berekening [naam onderzoeksbureau, c.q. opdrachtgever] op basis van bij het CBS beschikbaar gestelde microdatabestanden betreffende Voortgezet Onderwijs Cohort Leerlingen (VOCL) 1999, welke afkomstig zijn van het GION die deze microdatabestanden heeft geleverd aan het CBS in het kader van een samenwerkingsverband.”

## **Versie 5**

101102 VOCL hoofdbestand 1999V5

090508 VOCL bovenbouw 1999V5

090506 VOCL derdejaar 1999V5

090506 VOCL vierdejaar 1999V5

101125 VOCL intelligentietoets 1999V5

101129 VOCL vmbo schoolverlaters 1999V5

101026 VOCL havo-vwo schoolverlaters 1999V5

101029 GBA-selectiebestand VOCL 1999V5

Naast de bovenstaande bestanden, is er nog het bestand ‘Schoolelementen’. In dit bestand wordt jaarlijks een tal variabelen uitgevraagd, dat aan het bestand wordt toegevoegd. De naam van dit bestand is ‘jjmdd VOCL Schoolelementen ‘99-‘jjVx. Voor dit bestand wordt geen nieuw documentatierapport opgesteld. Het bestand wordt niet apart in dit rapport beschreven. Dit bestand is de jaarlijkse update en wordt ieder jaar aangevuld.

---

De gebruiker dient rekening te houden met de volgende *do's & don'ts*:

- Uitsluitend de combinatie van RIN met SRTNUM 'R' of 'S' identificeert een persoon. Zie hoofdstuk 3 voor een beschrijving.
- Het bestand '090506 VOCL derdejaar 1999V5' bevat variabelen met cijfers voor vakken. De cijfers zijn uitgedrukt in behaalde puntenaantallen (maximum =100 punten). Daarnaast komen bij variabelen (ld17\_a, ld17\_b, ld17\_c, ld17\_e, ld17\_f) negatieve waarden (-1, -2, -3, -4 en -5) voor. Om praktische redenen zijn de negatieve waarden niet gelabeld. Het betreft gehercodeerde waarden van resp. o, v, g, n, en daarnaast nog +. Ontbrekende cijfers als '0' zijn ook niet gelabeld. Zie bijlage 5: '101201 Bijlage 5 VOCL 1999 Derdejaar technisch rapport'.
- Het bestand '090506 VOCL vierdejaar 1999V5' bevat variabelen die betrekking hebben op nederlands en wiskunde. Naast de 4 antwoordmogelijkheden (1, 2, 3, 4) komen er ook gemiddelde scores (1,5 2,5 of 3,5) voor. Leerlingen hebben in deze gevallen meerdere antwoordmogelijkheden gekozen. Hetzelfde verschijnsel doet zich voor in het bestand '101129 VOCL vmbo schoolverlaters 1999V5'. Deze waarden zijn niet gelabeld. Voor uitleg zie bijlage 6: '101201 Bijlage 6 VOCL 1999 Vierdejaar technisch rapport'. Verder komt de score '999' voor. Dit zijn ontbrekende waarden.
- Het bestand '101022 VOCL havo-vwo schoolverlaters 1999V5' bevat op sommige variabelen ongelabelde waarden (13, 34 etc.) met lage scores. Het gaat hier ook om gecombineerde waarden (13: leerlingen hebben zowel het 1<sup>e</sup> als 3<sup>e</sup> of 34: leerlingen hebben zowel het 3<sup>e</sup> als 4<sup>e</sup> antwoord omcirkeld).
- In de jaarlijkse update VOCL schoolelementen '99-'jjV1 worden steeds de volgende variabelen toegevoegd : schjajj; brinjversl ; vertdjj; afdlbojj; exajj101; exajj102; exajj103; exajj104; exajj105; exajj106; exajj107; exajj108; vestnrjj; isatjj; oplfsjj; oplrjj; registjj; enqjj; datinsjj; exresjj; klasjj; ondeljj; vertcjj; iltjj; crebojj; oplvmjj.
- In bijlage(n) genoemde bestandsnamen kunnen afwijken van de namen van de microdatabestanden zoals genoemd onder 'versie 5'. De reden is dat deze bijlagen verwijzen naar de oorspronkelijke, nog niet voor on site en remote access gedocumenteerde, bestanden.
- Er zijn pdf-bestanden van de originele toetsen en vragenlijsten van VOCL 1999, die bij de microdatabestanden horen, als bijlage beschikbaar.

## Inhoudsopgave

<i>Bronvermelding</i> .....	2
<i>Versie 5</i> .....	3
<b>1. Inleiding</b> .....	7
<b>2. Toelichting op de samenstelling van de bestanden</b> .....	9
<b>Microdatabestand</b> .....	9
Recordbeschrijving .....	9
Populatieafbakening en steekproef.....	9
Opbouw VOCL 1999 .....	9
Het bestand met schoolelementen .....	10
Gegevens op Statline.....	10
<b>GBA-selectiebestand</b> .....	10
<b>3. Bestandsopbouw en toelichting</b> .....	11
<b>Bestandsopbouw van het microdatabestand</b> .....	11
Kwaliteitsaanduiding .....	11
<b>Toelichting op de variabelen</b> .....	11
SRTNUM en RIN .....	11



## 1. Inleiding

Bij het Centrum voor Beleidsstatistiek (CvB) worden microdatabestanden van het Centraal Bureau voor de Statistiek (CBS) beschikbaar gesteld voor gebruik door externe onderzoekers. Dit documentatierapport beschrijft de inhoud en structuur van de microdatabestanden

- '101102 VOCL hoofdbestand 1999V5'
- '090508 VOCL bovenbouw 1999V5'
- '090506 VOCL derdejaar 1999V5'
- '090506 VOCL vierdejaar 1999V5'
- '101125 VOCL intelligentietoets 1999V5'
- '101129 VOCL vmbo schoolverlaters 1999V5'
- '101026 VOCL havo-vwo schoolverlaters 1999V5'

Van alle personen die in het hier beschreven microdatabestand met SRTNUM 'R' voorkomen, zijn gegevens uit de Gemeentelijke basisadministratie (GBA) geselecteerd. Deze selectie kunt u vinden in het bestand:

- '101029 GBA-selectiebestand VOCL 1999V5'

Het complete GBA-bestand is in een apart [documentatierapport](#) beschreven. Voor Remote Access en On-site gebruikers is het documentatierapport over de GBA te vinden op \2-Metadata\GBA.

Hoofdstuk 2 beschrijft in het algemeen de microdatabestanden zoals populatieafbakening, methodologische bijzonderheden, kwaliteit en herkomst van de gegevens. In hoofdstuk 3 wordt de bestandsindeling per microdatabestand gegeven. Daarnaast wordt per bestand per variabele meer informatie gegeven. Voor de categoriale variabelen worden alle mogelijke scores en hun betekenis opgesomd.

Bij het rapport zijn de volgende bijlagen beschikbaar:

- '101124 Bijlage 1a Frequentietellingen VOCL Bovenbouw 1999V5'
- '101124 Bijlage 1b Frequentietellingen VOCL Hoofdbestand 1999V5'
- '101124 Bijlage 1c Frequentietellingen VOCL Intelligentietoets 1999V5'
- '101124 Bijlage 1d Frequentietellingen VOCL Derdejaar 1999V5'
- '101124 Bijlage 1e Frequentietellingen VOCL Vierdejaar 1999V5'
- '101124 Bijlage 1f Frequentietellingen VOCL Schoolverlaters Havo-Vwo 1999V5'
- '101124 Bijlage 1g Frequentietellingen VOCL Schoolverlaters Vmbo 1999V5'

In deze bijlagen worden van de (semi-) continue variabelen de frequenties van de eerste en laatste vijf waarden gegeven.

Met deze frequentietellingen kan tevens een beeld verkregen worden over de vullingsgraad van de verschillende variabelen dat van belang kan zijn omdat niet altijd alle variabelen (volledig) gevuld zijn.

De variabelen en bijbehorende waarden worden beschreven in:

- ‘101201 Bijlage 2a Variabelen VOCL 1999V5’ en
- ‘101201 Bijlage 2b Waarden VOCL 1999V5’

Een omschrijving van de veranderingen en trendbreuken in de loop van de tijd wordt beschreven in:

- ‘070906 Bijlage 3 Veranderingen en trendbreuken’

Daarnaast zijn de volgende algemene bijlagen beschikbaar:

- ‘101201 Bijlage 4 VOCL 1999 Eerstejaar technisch rapport’
- ‘101201 Bijlage 5 VOCL 1999 Derdejaar technisch rapport’.
- ‘101201 Bijlage 6 VOCL 1999 Vierdejaar technisch rapport’



## **2. Toelichting op de samenstelling van de bestanden**

### **Microdatabestand**

#### ***Recordbeschrijving***

Een record in het microdatabestand bevat informatie over leerlingen die in het schooljaar 1999/2000 in de brugklas van het voortgezet onderwijs zaten. Een persoon kan niet meerdere records in het microdatabestand hebben. Het bestand bevat SRTNUM, RIN als sleutel(s) waarmee gekoppeld kan worden met andere bestanden.

#### ***Populatieafbakening en steekproef***

De populatie van het Voortgezet Onderwijs Cohort Leerlingen 1999 (VOCL 1999) bestaat uit alle leerlingen die in schooljaar 1999/2000 in de brugklas zaten. De gegevensverzameling is uitgevoerd door het CBS en het Gronings Instituut voor Onderzoek van Onderwijs, Opvoeding en Ontwikkeling (GION). In dit cohort zijn zo'n 20.000 leerlingen gevolgd gedurende hun schoolloopbaan. Deze leerlingen vormen een representatieve steekproef van de instroomlichting 1999/2000 (10%-steekproef). Het steekproefkader bestaat uit alle scholen voor voortgezet onderwijs in Nederland in 1999. Belangrijke doelen van VOCL 1999 waren de evaluatie van twee in 1999 doorgevoerde veranderingen in het voortgezet onderwijs: het VMBO en het tweede fase-onderwijs in HAVO/VWO. Dit cohort is de voortzetting van de reeks schoolloopbaanonderzoeken die het CBS sinds 1965 uitvoert. Eerdere cohortonderzoeken vonden plaats in 1965 (het zogenaamde 'van Heek'-cohort) en vervolgens in 1977, 1982, 1989, en in 1993. De cohorten 1989 en 1993 lopen nog: niet alle brugklassers uit die jaren hebben het onderwijs verlaten. Kenmerkend voor deze panelonderzoeken is dat na het verzamelen van een aantal basisgegevens van brugklassers en hun ouders bij de start van het voortgezet onderwijs, de cohortleerlingen gedurende minimaal 15 jaren in hun gang door het (voortgezet en hoger) onderwijs worden gevolgd. Ieder jaar wordt de positie in het onderwijs geregistreerd, er worden regelmatig toetsen afgenomen en gevraagd naar de beleving van de school. Daarnaast worden er vragen gesteld aan ouders en schooldirecties en docenten.

#### ***Opbouw VOCL 1999***

Het totale cohortonderzoek VOCL'99 bestaat uit de bestanden zoals opgesomd in de inleiding van dit rapport. Toelichtingen op deze bestanden zijn als (aparte) bijlagen toegevoegd.

### ***Het bestand met schoolelementen***

Dit is een afzonderlijk bestand waarin alle onderwijsposities zijn opgenomen tot en met schooljaar 2009/2010 (jmmdd VOCL Schoolelementen '99-'jjVx.sav). Dit bestand wordt jaarlijks aangevuld zolang de leerlingen onderwijs volgen. De aanvulling geschiedt sinds schooljaar 2003/2004 via koppeling met de onderwijsnummerbestanden van de IB-Groep. In het hoofdbestand zitten alleen onderwijsposities tot en met het schooljaar 2002/2003, toen er nog veldwerk bij de scholen plaatsvond.

### ***Gegevens op Statline***

Er zijn geen gegevens over VOCL op Statline aanwezig.

### **GBA-selectiebestand**

Het bijbehorende GBA-selectiebestand bevat persoons- en huishoudensgegevens van alle personen die in het microdatabestand voorkomen en ook geregistreerd staan in de GBA. De persoonsgegevens bestrijken een periode van 1 januari 1995 tot 1 januari 2010. De huishoudensgegevens worden door het CBS berekend voor de stand op 1 januari. In het GBA-selectiebestand zijn huishoudensgegevens beschikbaar van 1998 tot en met 2010.

Zowel het microdatabestand als het GBA-selectiebestand is gesorteerd op respectievelijk SRTNUM en RIN. Met behulp van deze twee variabelen als koppelsleutel kan de gebruiker persoons- en huishoudensgegevens uit het GBA-selectiebestand toevoegen aan het microdatabestand, zodat deze gegevens in het onderzoek meegenomen kunnen worden.

Om RIN-nummers toe te kennen aan het microdatabestand is gebruik gemaakt van de meest recente versie van de GBA. Hierin is slechts een beperkt aantal variabelen aanwezig, daarom wordt dit bestand niet gebruikt voor het GBA-selectiebestand. Het GBA-selectiebestand maakt gebruik van gegevens tot 1 januari 2010. Hierdoor kunnen personen die na 31 december 2009 naar Nederland zijn geïmmigreerd, niet worden teruggevonden in het GBA-selectiebestand.

### **3. Bestandsopbouw en toelichting**

#### **Bestandsopbouw van het microdatabestand**

De bestandsopbouw van alle VOCL bestanden die betrekking hebben op 1999V5 staan in:

- '101201 Bijlage 2a Variabelen VOCL 1999V5' en
- '101201 Bijlage 2b Waarden VOCL 1999V5' .

#### ***Kwaliteitsaanduiding***

De meeste VOCL-variabelen cohort 1999 zijn gebruikt in analyses. Zie voor een (niet-uitputtende) publicatielijst '101201 Bijlage 4 VOCL 1999 Eerstejaar technisch rapport' .

Het grootste probleem is de uitval. Deze is sinds cohort SMVO1977 toegenomen. Tot schooljaar 2004/2005 is de jaarlijkse onderwijspositie opgevraagd bij de scholen. Indien een leerling verhuisde, is weliswaar gevraagd of de scholen het nieuwe adres van de leerling kenden, maar dat was lang niet altijd het geval. Deze leerling werd als 'vertrokken' gekenmerkt, maar is dus eigenlijk 'nonrespons'. In dezelfde categorie 'vertrokken' kwamen ook de schoolverlaters (opleiding beëindigd) terecht.

Vanaf schooljaar 2004/2005 is voor het VO en MBO overgestapt op koppeling via het SSB van het CBS met de integrale onderwijsnummerbestanden van de IB-Groep (DUO). Hoewel in het eerste jaar de bestanden nog niet volledig waren, impliceerde dit toch een verbetering in de gegevensverzameling in vergelijking met het veldwerk via de scholen. Leerlingen die als 'vertrokken' waren gekenmerkt, bleken 1 of 2 jaar later toch op een andere school te zitten. Deze trendbreuken zijn zichtbaar in het bestand met de jaarlijkse onderwijsposities. Zie ook '070906 Bijlage 3 Veranderingen en trendbreuken'.

#### **Toelichting op de variabelen**

##### ***SRTNUM en RIN***

Bronbestanden die aan het CBS geleverd worden, bevatten doorgaans als persoons-identificator het burgerservicenummer. Het burgerservicenummer wordt als zeer identificerend beschouwd en daarom vervangt het CBS dit nummer door een intern persoonsidentificatienummer, het betekenisloze RIN-nummer. Met dit RIN-nummer kan het bestand worden gekoppeld aan andere bestanden bij het CBS.

Het bronbestand van dit microdatabestand bevatte geen burgerservicenummer. Via andere identificerende gegevens dan het burgerservicenummer kon het RIN-nummer ook aan dit bestand worden toegevoegd. Dit is gebeurd door van ieder record in het bronbestand de MODNUMMERS op te zoeken. Het CBS heeft namelijk iedere leerling voorzien van een modulair nummer. Op deze manier kon het RIN-nummer van leerlingen in het bronbestand worden achterhaald.

Het microdatabestand is voorzien van de door het CBS toegevoegde variabelen soort identificatiecode (SRTNUM) en identificatienummer (RIN). Records in het microdatabestand die konden worden voorzien van een RIN-nummer hebben de waarde 'R' op SRTNUM gekregen. De overige records zijn niet teruggevonden in de GBA. Deze records hebben de waarde 'G' op SRTNUM gekregen, met een lege score in het veld RIN.

Uitsluitend de combinatie van SRTNUM 'R' en RIN identificeert een uniek persoon. Koppelen aan andere bestanden of aggregeren op persoonsniveau moet daarom alleen worden uitgevoerd op records met SRTNUM 'R'. Met behulp van SRTNUM en RIN kan de gebruiker gegevens uit het GBA-selectiebestand toevoegen aan de records in het microdatabestand met SRTNUM 'R'.