

Uitgebreide technische verantwoording SIC 2018

Eindredactie: Studiekeuze123, 2018

V4 2018 07 04

Inhoud

STUDIE IN CIJFERS – DEFINITIES, BEWERKINGEN EN BEREKENINGEN	1
Aggregatieniveaus	2
NSE: STUDENTTEVREDENHEID EN CONTACTTIJD EERSTEJAARS (HBO)	3
Studenttevredenheid	3
Ervaren contacttijd eerstejaars (hbo)	4
NSE: CONTACTTIJD EERSTEJAARS (WO): AANLEVERING DOOR UNIVERSITEITEN	6
1CIJFERHO: AANTAL EERSTEJAARS, DOORSTROOM NAAR TWEEDE JAAR, BEHAALDE DIPLOMA'S EN DOORSTUDEREN	7
Aantal eerstejaars	7
Doorstroom naar 2 ^e jaar	8
Diploma's	9
Doorstuderen	11
HBO-MONITOR: WERKLOOSHEID EN BAAN OP NIVEAU	13
Werkloosheid	13
Baan op niveau	14

STUDIE IN CIJFERS – DEFINITIES, BEWERKINGEN EN BEREKENINGEN

Dit document beschrijft in detail hoe de gegevens in Studie in Cijfers 2018 tot stand zijn gekomen. De indeling van deze toelichting is gebaseerd op de gebruikte bronnen: Nationale Studenten Enquête, aanlevering gegevens voor contacttijd door universiteiten, het 1cijferHO-bestand, en de HBO-monitor.

Studie in Cijfers omvat alle door de overheid bekostigde hbo- en wo-bacheloropleidingen in Nederland. Verder geldt dat bij unieke opleidingen de landelijke cijfers niet worden getoond omdat beide identiek zijn.

Aggregatieniveaus

In Studie in Cijfers worden gegevens uit het 1cijferHO-bestand berekend op verschillende aggregatieniveaus. Deze zijn gebaseerd op de variabelen Brinnummer, Brinvolnummer, Opleidingscode zoals deze in het Centraal Register Opleidingen in het Hoger Onderwijs (croho) voorkomen, en die ook in de diverse bronbestanden worden gehanteerd (echter niet altijd onder dezelfde namen). Hieronder volgt een omschrijving van de aggregatieniveaus die gebruikt worden, en hoe deze zich verhouden tot de genoemde variabelen in 1cijferHO en in croho.

Opleiding: Een Opleiding is in de definitie van Studiekeuze123 een specifiek curriculum dat aan een specifieke vestiging van een specifieke instelling wordt verzorgd. Een opleiding staat in principe gelijk aan een combinatie van Instelling, Vestiging en een opleidingscode in 1cijferHO. Uitzonderingen zijn Joint Degrees: hoewel deze door meerdere instellingen in samenwerking verzorgd worden, worden deze in SiC als één Opleiding gezien.

Studie: Een Studie is bij Studiekeuze123 een verzameling van een of meer sterk gerelateerde Opleidingen (zie bovenstaande definitie). Bij een Studie gaat het dus altijd om landelijke cijfers. In het algemeen staat een Studie gelijk aan een opleidingscode zoals die in croho of 1cijferHO voorkomt. In enkele gevallen worden opleidingscodes gegroepeerd tot één Studie. Bijvoorbeeld: in croho hebben Scheikunde en Chemistry verschillende opleidingscodes, maar voor Studiekeuze123 vormen zij samen één Studie.

Instelling x Opleiding: Dit is in zekere zin een niveau dat tussen Studie en Opleiding in zit. Het gaat hier om een opleiding aan een instelling *zonder onderscheid naar vestiging*. Met uitzondering van aantal eerstejaars worden de gegevens uit 1cijferHO op dit niveau getoond in plaats van op het niveau Opleiding. De reden hiervoor is dat binnen 1cijferHO per regel op dit niveau is aangegeven of het een op de peildatum geldige hoofdschrijving betreft. Dit is een belangrijk gegeven dat niet bekend is op het niveau Opleiding zoals hierboven gedefinieerd. In het geval van Joint Degrees wordt altijd de Instelling van de penvoerder van de Joint Degree gebruikt, ook als in 1cijferHO een ander brinnummer staat. Uit de NSE wordt contacttijd op dit niveau weergegeven.

Enkele voorbeelden van de implicaties van deze definities:

- 'Accountancy aan de Hogeschool InHolland Alkmaar' is een **Opleiding**. Het aantal eerstejaars wordt getoond op dit niveau. 'Hbo bachelor Accountancy' is een **Studie**. Alle landelijke cijfers worden op dit niveau getoond. Doorstroom naar het 2^e jaar, percentage behaalde diploma's binnen 5 jaar, percentage doorstudeerders en contacttijd worden weergegeven voor de hbo bachelor Accountancy aan InHolland (dus niet uitgesplitst per vestiging – dit is het niveau **Instelling x Opleiding**).
- Bij de cijfers voor Chemistry aan de Radboud Universiteit worden de landelijke cijfers getoond van alle opleiding Chemistry *en Scheikunde* (deze vormen samen één **Studie**).
- De Joint Degree Physics And Astronomy wordt verzorgd door de UvA en VU. De UvA is penvoerder voor deze opleiding. Alle eerstejaarsstudenten aan deze **Opleiding**, ongeacht waar ze ingeschreven staan worden daarom meegeteld bij de cijfers van de UvA.

NSE: STUDENTTEVREDENHEID EN CONTACTTIJD EERSTEJAARS (HBO)

De *Nationale Studenten Enquête (NSE)* wordt jaarlijks onder auspiciën van Studiekeuze123 uitgevoerd. Het is een tevredenheidsmeting die op vele aspecten peilt hoe studenten van hogescholen en universiteiten denken over hun studie en de instelling waar zij studeren.

Twee gegevens uit de NSE gebruiken we voor Studie in Cijfers: de 'algemene tevredenheid' en de hoeveelheid contacturen per week in het hbo. Contacttijd in het wo wordt niet aan de hand van de NSE-resultaten bepaald, maar wordt aangeleverd door de instelling zelf.

Studenttevredenheid

De enquêtevraag die we gebruiken om iets over tevredenheid van studenten te zeggen luidt: '*Geef alsjeblieft hieronder aan hoe tevreden je bent over: Je studie in het algemeen*'. De antwoordmogelijkheden zijn 1=zeer ontevreden t/m 5 =zeer tevreden (tussenvolgende categorieën zijn niet gedefinieerd) of 6 = n.v.t.

Studie in Cijfers gaat over bachelor voltijds studenten. Verder zijn in de berekeningen de premasters en postmasters buiten beschouwing gelaten. We selecteren dus eerst uit alle ingevulde NSE vragenlijsten alleen de bachelor voltijders. We kijken alleen naar de meest recente editie van de NSE, op het moment van schrijven is dat 2018. Voor instellingen die om het jaar meedoen nemen we de data uit 2017. Iedereen die 'n.v.t.' invulde of de vraag anderszins niet heeft gekregen (een 'missing') wordt buiten beschouwing gelaten, we tellen alleen de antwoorden tussen 1 en 5 mee.

Vervolgens maken we berekeningen op 2 niveaus: per Opleiding en per Studie (zie de definitie van deze twee aggregatieniveaus hierboven).

We berekenen drie soorten gegevens op basis van de studenttevredenheidsvraag:

- De *gemiddelde tevredenheid*. Dit gemiddelde zal dus tussen 1 en 5 uitkomen en wordt in getallen (1 cijfer achter de komma) weergegeven. Het gaat om een *gewogen* gemiddelde. Gewogen wil zeggen: er is een weegfactor gebruikt die in het bestand zit. De weegfactor zorgt ervoor dat studenten van opleidingen met relatief weinig respons zwaarder meetellen in het berekende gemiddelde dan studenten van opleidingen met relatief veel respons. De weegfactor in dit bestand corrigeert op opleiding (instelling x opleiding x vestiging). Op opleidingsniveau (instelling x opleiding x vestiging) heeft die weging dus géén invloed (want dat is het niveau van de gemiddelden en dan zijn alle weegfactoren voor alle respondenten hetzelfde. Dit geldt niet bij joint degrees want daar worden opleidingen samengenomen). Landelijk kan er wel invloed zijn van die weging.
- De *verdeling (in percentages) van antwoorden 1 (zeer ontevreden) t/m 5 (zeer tevreden)*. Hiertoe pakken we de 'gewogen aantallen' 1tjes, 2tjes, 3tjes, 4tjes en 5jes. Deze delen we steeds door het totale gewogen aantal (wat per definitie de optelsom is van die 1tjes t/m 5tjes). Dit leidt tot vijf percentages, die in een staafdiagram worden weergegeven.
- *Het aantal studenten* waarop deze gegevens zijn gebaseerd (ook wel *n* genoemd). Het gaat hier om het *ongewogen* aantal, wat dus kan afwijken van het gewogen aantal dat we voor de berekening van de verdeling in percentages gebruiken.

Als we zodoende resultaten hebben, per opleidingslocatie en landelijk, per studie, bepalen we welke resultaten we gaan tonen. Er zijn twee weergaveregels,

1. de eerste is simpel: bij *minder dan 5 respondenten* geven we het resultaat niet weer (het gaat hier om *ongewogen* aantallen).
2. De tweede is complexer en heeft met de verdeling van antwoorden te maken. Stel, een opleiding heeft een gemiddelde tevredenheid van 3. Maar dit gemiddelde bestaat uit twee even grote groepen: een groep die een 1 geeft, en een groep die een 5 geeft. In dat voorbeeld is de *spreiding* rond het gemiddelde heel groot. Een andere opleiding heeft ook een gemiddelde van 3, maar in dit geval heeft iedereen '3' ingevuld. De spreiding is daar heel klein (zelfs nul). Hoe kleiner de spreiding én hoe groter het aantal respondenten, hoe betrouwbaarder het getal dat je toont. Die betrouwbaarheid wordt gekwantificeerd met een betrouwbaarheidsinterval rond het gemiddelde. De regel is: het 95% betrouwbaarheidsinterval mag niet breder dan 0,5 zijn. Bij de berekening van het betrouwbaarheidsinterval wordt een *eindigheidscorrectie* toegepast, die corrigeert voor de grootte van de populatie waarvan de respondenten deel uitmaken. Zie de berekeningswijze hieronder:

Voor de tevredenheid op opleidingsniveau- en landelijk niveau wordt de marge met eindigheidscorrectie als volgt berekend: $MARGE = (1,96 * (Standaarddeviatie / \sqrt{n_{ongewogen}})) * (\sqrt{(N - n_{ongewogen}) / (N - 1)})$. Hierbij is $n_{ongewogen}$ het aantal respondenten en N de netto populatie waarvan de respondenten deel uitmaken. Als $MARGE < 0,5$ is (en $n \geq 5$) wordt de uitkomst getoond.

Ervaren contacttijd eerstejaars (hbo)

In het hbo geven we een indicatie van de hoeveelheid contacturen die een student in het eerste jaar heeft, volgens de eerstejaars zelf. In het wo wordt contacttijd uit een andere bron gehaald, zie het volgende hoofdstuk.

De vraag uit de vragenlijst die we hiervoor gebruiken, is als volgt gesteld:

"De volgende vragen gaan over contacttijd. Dat is het aantal klokuren per week waarin je geprogrammeerd contact hebt met een onderwijsgevende (docent, tutor, e.d.). Niet tot contacttijd moet je activiteiten rekenen als stages, werkplekleren, scripties en afstudeeronderzoek. Geef aan welke antwoordcategorie voor jou van toepassing is ten aanzien van contacttijd binnen je opleiding. In deze onderwijsperiode wordt mij de volgende hoeveelheid contacttijd aangeboden:

- minder dan 6 klokuren per week,
- 6 tot 12 klokuren per week
- 12 tot 18 klokuren per week,
- 18 tot 24 klokuren per week,
- 24 tot 30 klokuren per week,
- 30 of meer klokuren per week,
- Nvt

We maken een selectie van studenten die we meenemen in de berekeningen: het gaat om *voltijds bachelors*, en uitsluitend om *eerstejaars*. Tevens worden alleen de data van de reguliere studenten gebruikt (post- en pre-masters studenten worden niet meegenomen in de berekening).

We geven voor contacttijd het meest gegeven antwoord weer op het niveau instelling x opleiding en daarnaast op het landelijk niveau (Studie). Dit wordt ook wel de *modus* genoemd (het gaat hier dus niet om het gemiddelde).

Voorbeeld: als er 100 respondenten zijn, en 75 zeggen 'minder dan 6 klokuren per week' en 25 '30 of meer klokuren per week', dan zou het gemiddelde ergens daartussen liggen (de antwoordcategorieën zijn niet heel precies dus je zou kunnen discussiëren over waar dat precies ligt. Maar de modus is dan 'minder dan 6 klokuren per week': het vaakst gegeven antwoord. Indien er twee antwoordcategorieën even vaak zijn gekozen dan worden ze beide getoond, door de antwoordopties uit te breiden, bijvoorbeeld als 6 tot 12 en 12 tot 18 even vaak zijn gekozen dan wordt 6 tot 18 klokuren per week getoond. Dit wordt alleen gedaan als de antwoordcategorieën aansluitend zijn. Indien er drie modi zijn die even vaak zijn gekozen, dan wordt de middelste getoond, maar alleen als ze aansluitend zijn.

NSE: CONTACTTIJD EERSTEJAARS (WO): AANLEVERING DOOR UNIVERSITEITEN

In het wo is besloten om contacttijd niet te berekenen op basis van wat eerstejaars studenten zeggen, maar op basis van wat de universiteiten aangeven (meestal overgenomen uit de studiegids van de opleiding).

De universiteiten hebben zelf d.m.v. een tool op de website studiekeuzeinformatie.nl dit gegeven aangeleverd in uren, per opleiding (dus opleidingscode x instelling) voor eerstejaars voltijd bachelor studenten. Vervolgens zijn deze uren omgezet in de volgende categorieën:

- minder dan 6 klokuren per week,
- 6 tot 12 klokuren per week
- 12 tot 18 klokuren per week,
- 18 tot 24 klokuren per week,
- 24 tot 30 klokuren per week,
- 30 of meer klokuren per week,

1CIJFERHO: AANTAL EERSTEJAARS, DOORSTROOM NAAR TWEEDE JAAR, BEHAALDE DIPLOMA'S EN DOORSTUDEREN

1cijferHO is het bestand met álle inschrijvingen in het Nederlandse hoger onderwijs, voornamelijk van rijksbekostigde instellingen. Instellingen voor particulier onderwijs (zoals LOI, NTI) geven wel erkende diploma's maar hun studenten worden niet centraal ingeschreven; daarmee zijn deze niet vertegenwoordigd in 1cijferHO. Ook studenten in het buitenland zitten niet in dit bestand. Jaarlijks stelt DUO dit bestand in het voorjaar (deels) ter beschikking aan instellingen, en voor onderzoeksdoeleinden aan sommige onderzoekers, waaronder Studiekeuze123. De peildatum is 1 oktober (hoofregel: wie ingeschreven is op die datum telt bij dat jaar mee). Het bestand is niet openbaar.

Alle 'reguliere' Nederlandse studenten zijn in dit bestand opgenomen (wel wordt ter bescherming van privacy het burgerservicenummer door DUO versleuteld alvorens het bestand aan Studiekeuze123 ter beschikking te stellen). Om terug te kunnen rekenen (bijvoorbeeld bepalen wie na x jaar een diploma heeft behaald) krijgt Studiekeuze123 een 1cijferHO-bestand met alle inschrijvingen tot 11 jaar terug. In het meest recente geval: 2006 t/m 2017.

In 1cijferHO komen diverse velden voor die Brinnummers en opleidingscodes bevatten. Wij gebruiken hiervan BRIN_ACTUEEL en OPL_ACT_EQ. Dit zijn geactualiseerde versies van brinnummer en opleidingscode (opleidingscode wordt alleen geactualiseerd voor het hbo).

Aantal eerstejaars

Hieronder wordt weergegeven op welke manier het aantal eerstejaars op het opleiding niveau wordt berekend. Daarnaast wordt de berekening op studie niveau uiteengezet.

Stel vast wat het meest recente jaar is in het 1cijferHO bestand. Noem dit MEESTRECENT.

- 1) Selecteer uit het 1cijferHO bestand alle eerstejaars voltijd bachelorstudenten in het meest recente jaar: `IND_EJ_OPLINS = 1 AND OPLVORM = 1 AND TYPE_HO = 'ba' AND JAAR = MEESTRECENT1`.
- 2) Verwijder voor wo-opleidingen studenten die als hoogste vooropleiding een hbo bachelor hebben. Het gaat hier in de meeste gevallen om premasters, niet om 'echte' eerstejaars. Selecteer daarom alleen regels waarvoor geldt: `SOORT_HO='hbo' OR (SOORT_HO='wo') AND (VROPL_HO<800 OR VROPL_HO>899)2`.
- 3) Verwijder studenten die al voor aanvang van het studiejaar zijn gestart met hun studie (voor het grootste deel februari-instroom). Deze studenten staan als eerstejaars geregistreerd maar zijn in feite al in het voorgaande studiejaar gestart. Deze zijn te herkennen aan de variabele `EERSTE_JR_OPLINS` (eerste jaar van inschrijving aan de opleiding/instelling). Selecteer alleen studenten waarvoor geldt `EERSTE_JR_OPLINS=MEESTRECENT`.

¹ `IND_EJ_OPLINS` staat voor 'indicatie eerstejaars aan de opleiding/instelling', `OPLVORM = 1` betekent 'opleidingsvorm = voltijd', `TYPE_HO = 'ba'` betekent dat het om een bacheloropleiding gaat.

² `VROPL_HO` staat voor 'vooropleiding in het hoger onderwijs', en de waarden van 800 t/m 899 staan voor een hbo bachelor.

- 4) Selecteer alleen hoofdinschrijvingen binnen het domein Opleiding x Instelling (we gebruiken dit niveau omdat het gegeven 'hoofdinschrijving' niet bekend is op het niveau Opleiding x Instelling x Vestiging). Hiermee verwijderen wij dubbele inschrijvingen en uitwisselingstudenten. Selecteer regels waarvoor geldt: $SINS_OPLINS^1 = '1' \text{ OR } SINS_OPLINS = 'A'$.
- 5) Voor Joint Degrees moet stap 4 worden aangescherpt: als een student switcht tussen instellingen die gezamenlijk een opleiding verzorgen, dan willen we ze aan de nieuwe instelling niet als eerstejaars meetellen. Bepaal daarom aan de hand van croho of het om een joint degree gaat. Maak een variabele aan JOINTDGR die in dat geval 1 is, anders 0. Bij Joint Degrees geldt de extra voorwaarde dat het binnen het domein *Opleiding* om een eerstejaars inschrijving moet gaan (let op: dit gaat over het domein opleiding zoals gedefinieerd binnen 1cijferHO, niet om het niveau Opleiding volgens de definitie die binnen de studiekeuzedatabase gehanteerd wordt). Selecteer daarom de regels waarvoor geldt: $JOINTDGR = 0 \text{ OR } (JOINTDGR = 1 \text{ AND } IND_EJ_OPL = 1^2)$
- 6) Tel per Opleiding het aantal resterende regels in het bestand. Het landelijke cijfer geeft het gemiddelde van alle Opleidingen die onder dezelfde Studie vallen.

Doorstroom naar 2^e jaar

Dit cijfer laat zien welk percentage van de eerstejaarsstudenten het volgende jaar weer staat ingeschreven aan dezelfde opleiding en aan dezelfde instelling. Het wordt op twee aggregatieniveaus getoond: op Opleiding x Instelling- en op Studieniveau (dwz. landelijk).

NB: Studie in Cijfers gaat over voltijdstudenten, dus alleen degenen die een jaar later nog steeds in de *voltijdsvariant* van dezelfde opleiding aan dezelfde instelling is ingeschreven wordt meegenomen.

- 1) Selecteer alle eerstejaars in het jaar vóór het meest recente ($JAAR = MEESTRECENT - 1$). We gebruiken dezelfde selectie als voor eerstejaars, maar dan voor een jaar eerder. Daar voegen we nog een selectie aan toe: als iemand in het jaar van eerste inschrijving direct het diploma behaalt telt deze persoon niet mee in de teller en noemer bij het berekenen van het doorstroompercentage. Selecteer daarom alleen de regels waarvoor geldt: $SDIP_HO > 4 \text{ OR } SYSMIS(SDIP_HO)^2$
- 2) Tel per Instelling x Opleiding en per Studie het aantal unieke personen ($OMNUMMER_HO$) in het eerstejaarscohort. Dit zijn de noemers in de berekening van het percentage.
- 3) Bepaal van elke combinatie van persoon, instelling en opleiding of dezelfde combinatie van deze drie variabelen een jaar later ook in het bestand voorkomt³. Dat betekent: bestaat er een regel waarvoor geldt: $JAAR = MEESTRECENT \text{ AND } OPLVORM = 1 \text{ AND } INSCHRVRM = 'S'$.

¹ SINS_Oplins staat voor: soort inschrijving binnen domein 'actuele opleiding en instelling'. Geeft aan of het om een hoofdinschrijving of neveninschrijving gaat in dit domein, of dat de inschrijving niet actief is op de peildatum (bijv. in het geval van februari-instroom). Mogelijke waarden: 1 t/m 9, A t/m E. Hoofdinschrijving is 1 of A voor respectievelijk bekostigde en aangewezen instellingen.

² SDIP_HO staat voor 'soort diploma in het hoger onderwijs' en is gevuld als de student in het jaar van inschrijving een diploma behaalt. Een waarde van 4 of kleiner betekent dat een bachelordiploma is behaald.

³ 'De combinatie van persoon, instelling en opleiding' is in het hbo iets anders geïmplementeerd dan in het wo. In het hbo gebruiken we hiervoor de combinatie van OMNUMMER, ACTUELE_INS en

- 4) Tel per Instelling x Opleiding en per Studie het aantal unieke personen (OMNUMMER_HO) waarvoor dit het geval is. Dit zijn de tellers in de berekening van het percentage.
- 5) Koppel op beide niveaus de tellers en de noemers op basis van Instelling x Opleiding of Studie, deel de teller door de noemer en vermenigvuldig dit getal met 100. Dit is het percentage doorstromers.

Het percentage Doorstromers wordt alleen getoond als de noemer in de berekening groter is dan 10.

Diploma's

Er worden twee gegevens getoond met betrekking tot behaalde diploma's:

- a) Het percentage studenten dat binnen de nominale studieduur plus 1 jaar het diploma behaalt van dezelfde opleiding en aan dezelfde instelling als waaraan ze zijn begonnen.
- b) Het percentage studenten dat binnen de nominale studieduur plus 1 jaar een diploma behaalt van hetzelfde niveau als de opleiding waaraan ze zijn begonnen, ongeacht welk diploma, of aan welke instelling.

'Nominale studieduur plus 1 jaar' betekent voor wo-bachelors 4 jaar en voor hbo-bachelors 5 jaar.

Diploma's die bij b) worden meegeteld als 'geldig' zijn:

- in het hbo: alle bachelors, alle masters, alle postinitiële masters, alle doctoralen, alle beroepsfase/wo-postmaster. Anders gezegd: alle hogeronderwijsdiploma's behalve propedeuses en associate degree.
- In het wo: wo-bachelor, alle masters, alle postinitiële masters, alle doctoralen, alle beroepsfase/wo-postmaster (afstroom naar hbo bachelor wordt dus *niet* meegerekend).

- 1) De cijfers onder a) en b) hierboven bepalen we op basis van een eerstejaarscohort (op dezelfde manier gecorrigeerd als bij de doorstroom naar jaar 2: verwijder degenen die in het 1^e jaar al direct een diploma behalen en degenen die waarschijnlijk een premaster volgen). Om uitspraken te kunnen doen over diploma's moeten we afgeronde studie jaren beschouwen. Van het nog lopende jaar weten we immers nog niet wie gaat afstuderen. Daarom kijken we naar diploma's behaald tot het jaar vóór het meest recente. Vanaf dat jaar rekenen wij vijf jaar (hbo) en vier jaar (wo) terug om een cohort te selecteren. Voor hbo selecteren we daarom op JAAR = MEESTRECENT-5, en voor wo op JAAR = MEESTRECENT-4. Voor het overige is de selectie gelijk aan de selectie voor het cohort bij *Doorstroom naar 2^e jaar*.

OPL_ACT_EQ, in het wo gebruiken we de combinatie van OMNUMMER, ACTUELE_INS en OPL_HIS_EQ (of, als deze niet bestaat: OPL_ACT_EQ). De reden hiervoor ligt in het feit dat de variabelen OPL_ACT_EQ en OPL_HIS_EQ in 1cijferHO op een andere manier gevuld zijn voor hbo dan voor wo.

- 2) Eerst berekenen we het percentage studenten dat binnen de nominale studieduur plus 1 jaar het diploma behaalt van dezelfde opleiding en aan dezelfde instelling als waaraan ze zijn begonnen (a). Hiervoor maken we een selectie van alle behaalde diploma's binnen vier of vijf jaar na het jaar waar een cohort begonnen is. De selectie ziet er als volgt uit:

JAAR>MEESTRECENT-5 AND JAAR<MEESTRECENT
AND SOORT_HO ='hbo' AND (SDIP_HO = 3 OR SDIP_HO = 4)

of

JAAR>MEESTRECENT-4 AND JAAR<MEESTRECENT
AND SOORT_HO ='wo' AND (SDIP_HO = 3 OR SDIP_HO = 4)

voor respectievelijk hbo en wo.¹

Voor elke combinatie van persoon, instelling en opleiding in het eerstejaarscohort wordt vastgesteld of deze ook voorkomt in de selectie van diploma's (*zie de voetnoot bij doorstroom naar het 2^e jaar*). Zo ja, dan is het diploma van de opleiding binnen de vastgestelde tijd behaald.

- 3) Wij bepalen vervolgens het percentage eerstejaars dat het diploma behaalt binnen x jaar op twee niveaus, namelijk Instelling x Opleiding en Studie. De noemer in de berekening van het percentage wordt bepaald door in alle regels in het eerstejaarscohort het aantal unieke waarden van OMNUMMER_HO te tellen. Om de teller te bepalen doen we hetzelfde, maar dan tellen wij alleen degenen die het diploma hebben behaald mee. Het percentage behaalde niveaus aan dezelfde opleiding en instelling bepalen we door op beide niveaus de teller te delen door de noemer en de uitkomst te vermenigvuldigen met 100.
- 4) Daarna berekenen we het percentage studenten dat binnen de nominale studieduur plus 1 jaar een diploma behaalt van hetzelfde niveau als de opleiding waaraan ze zijn begonnen, ongeacht welk diploma, of aan welke instelling (b). Voor hbo maken we een selectie van alle diploma's van het niveau hbo-bachelor of hoger (zie boven), behaald in de jaren na aanvang van het cohort:

JAAR>MEESTRECENT-5 AND JAAR <MEESTRECENT AND
SDIP_HO>2 AND (SDIP_HO<11 OR SDIP_HO=15 OR SDIP_HO=16)

Voor wo selecteren we iedereen die in de jaren na aanvang van het cohort een diploma behaald op wo-bachelorniveau of hoger:

JAAR>MEESTRECENT-4
AND JAAR <MEESTRECENT
AND ((SDIP_HO>4 AND SDIP_HO<11)
OR SDIP_HO=15 OR SDIP_HO=16
OR(SOORT_HO="wo" AND (SDIP_HO=3 OR SDIP_HO=4))²

¹ SDIP_HO staat voor 'soort diploma in het hoger onderwijs' een waarde van 3 of 4 geeft aan dat een bachelordiploma is behaald in het jaar van inschrijving.

² Deze selecties zijn uitwerkingen van de diploma's die meetellen bij b), zoals beschreven aan het begin van deze paragraaf.

Voor elke persoon (OMNUMMER_HO) in het eerstejaars cohort wordt vastgesteld of deze persoon ook voorkomt in deze bestanden van studenten die een diploma behaald hebben.

- 5) Wij bepalen vervolgens het percentage behaalde diploma's op twee niveaus, namelijk Instelling x Opleiding en Studie. De noemer in de berekening van het percentage wordt bepaald door in alle regels in het eerstejaarscohort het aantal unieke waarden van OMNUMMER_HO te tellen. Om de teller te bepalen doen we hetzelfde, maar dan tellen wij alleen degenen mee die in de bovenstaande selecties van gediplomeerden voorkomen. Het percentage behaalde diploma's op tenminste het niveau van de opleiding bepalen we door op beide niveaus de teller te delen door de noemer en de uitkomst te vermenigvuldigen met 100.
- 6) Als het cohort groter is dan 30 studenten, dan worden de percentages met betrekking tot behaalde diploma's zonder verdere bewerkingen getoond. Als het cohort uit 10 studenten of minder bestaat, dan worden er geen cijfers getoond. Als het cohort bestaat uit 11 t/m 30 studenten dan volgt nog een extra stap: dezelfde bewerkingen als hierboven beschreven worden dan ook uitgevoerd voor het cohort dat een jaar eerder begon (dus voor hbo en wo respectievelijk 6 en 5 jaar vóór het meest recente jaar) De tellers en noemers van beide cohorten worden dan bij elkaar opgeteld en op basis van die opgetelde cijfers worden de percentages berekend.

Deze regels voor het tonen/niet tonen van de percentages worden op beide aggregatieniveaus afzonderlijk toegepast.

Doorstuderen

Dit cijfer geeft weer welk percentage van de afgestudeerden van deze opleiding binnen twee jaar begint met een nieuwe opleiding in het hoger onderwijs.

- 1) Bepaal wie twee jaar voor het meest recente jaar in 1cijferHO een bacheloropleiding heeft behaald van de voltijdsvariant van de opleiding. Neem alleen hoofdschrijvingen binnen het domein Opleiding/Instelling mee. Het gaat om inschrijvingen waarvoor geldt:

```
JAAR = (MEESTRECENT-2)
AND (SDIP_HO=3 OR SDIP_HO=4)
AND OPLVORM = 1
AND (SINS_OPLINS1= '1' OR SINS_OPLINS='A')
```

- 2) Bepaal of de ingeschrevenen in die groep in de twee daaropvolgende jaren staan ingeschreven in het hoger onderwijs (niet-actieve inschrijvingen en uitwisselaars uitgezonderd). Dat doen we door een selectie te maken van regels waarvoor geldt: JAAR > MEESTRECENT-2 AND SINS_HO='1' OR SINS_HO='A'. Vervolgens bepalen we van elke persoon (OMNUMMER_HO) in

¹ SINS_OPLINS staat voor: SINS_OPLINS: soort inschrijving binnen domein 'actuele opleiding en instelling'. Geeft aan of het om een hoofdschrijving of neveninschrijving gaat in dit domein, of dat de inschrijving niet actief is op de peildatum (bijv. in het geval van februari-instroom). Mogelijke waarden: 1 t/m 9, A t/m E. Hoofdschrijving is 1 of A voor respectievelijk bekostigde en aangewezen instellingen.

het afgestudeerdcohort of dezelfde persoon ook voorkomt in deze selectie. In dat geval telt diegene als doorstudeerder, onafhankelijk voor welke opleiding hij/zij heeft gekozen.

- 3) Het aantal doorstudeerders (aantal unieke waarden van OMNUMMER_HO) gedeeld door de grootte van het afgestudeerdcohort (aantal unieke waarden van OMNUMMER_HO), vermenigvuldigd met 100, is het doorstudeerpercentage. Dit wordt bepaald op de niveaus Instelling x Opleiding en Studie.

Het percentage Doorstudeerders wordt alleen getoond als de noemer in de berekening groter is dan 10.

HBO-MONITOR: WERKLOOSHEID EN BAAN OP NIVEAU

Het ROA (Researchcentrum Onderwijs en Arbeidsmarkt, Universiteit Maastricht) voert sinds 1991 jaarlijks een grootschalige enquête uit onder afgestudeerden van het hbo: de **HBO-monitor** (incl. een speciale variant voor het kunstonderwijs, de Kunstenmonitor).

Jaarlijks worden hbo-afgestudeerden hierin ongeveer 1,5 jaar na afstuderen bevraagd over hun opleiding, hun vervolgstappen (vervolgopleidingen, werk) en details hiervan. Dit geeft onder andere inzicht in de tevredenheid over de opleiding, arbeidsdeelname, salaris van pas afgestudeerden en werkloosheid onder afgestudeerden. Het bestand dat aan Studiekeuze123 ter beschikking wordt gesteld bestaat uit hbo-bachelor- en master-afgestudeerden (met diploma).

NB Een vergelijkbare enquête wordt in het wo gehouden (de NAE, Nationale Alumni Enquête, vroeger als WO-monitor door ROA uitgevoerd maar tegenwoordig onder auspiciën van de VSNU). Maar Studiekeuze123 maakt hiervan geen gebruik voor Studie in Cijfers. Studie in Cijfers gaat immers over bacheloropleidingen, en de NAE bevraagt wo-*master*-afgestudeerden. Bovendien is de uitstroom van wo-bachelors naar de arbeidsmarkt klein; verreweg de meesten studeren door in een master. Arbeidsmarktinformatie zal dus (op de site van studiekeuze123) bij de wo-*master*opleidingen getoond worden. Wel krijgen wo-bachelor-studiekiezers op de Studie in Cijfers-bijsluiter een link te zien naar de meest gekozen masteropleidingen na die specifieke bachelor, waarna ze kunnen doorklikken om (o.a.) die arbeidsmarktinformatie te zien.

Werkloosheid

Het eerste gegeven dat we tonen op basis van de HBO-monitor gaat over werkloosheid van afgestudeerden, anderhalf jaar na afstuderen. We presenteren alleen de landelijke gegevens voor een Studie, en splitsen niet uit naar instelling. Om een beeld te geven van werkloosheid na een specifieke studie laten we, waar mogelijk, regionale verschillen buiten beschouwing (zoals het soort bedrijven in een regio; die kunnen immers sterk verantwoordelijk zijn voor de omvang van werkloosheid onder afgestudeerden).

Het gaat hier dus om de werkloosheid *op het moment van bevragen, dat is anderhalf jaar na afstuderen*. Over werkloosheid in de tussentijd gaat het niet.

Eerst kiezen we de juiste doelgroep: Studie in Cijfers gaat over *bachelor voltijdsopleidingen*, dus we kiezen uit de HBO-monitor de afgestudeerden van *bacheloropleidingen* die de voltijds variant hebben doorlopen.

Vervolgens berekenen we het gewogen (De weegfactor zorgt ervoor dat studenten van Studies met relatief weinig respons zwaarder meetellen in het berekende percentage dan studenten van Studies met relatief veel respons.) werkloosheidspercentage op basis van de beroepsbevolking en het aantal werklozen die ROA meet in de hbo-monitor. ROA stelt op basis van verschillende enquêtevragen (omschrijving van de huidige situatie, aantal uren betaald werk, werkzoekend of niet) vast wie werkzoekend is, werkloos is, of werk heeft. De beroepsbevolking is, conform de CBS definitie, gedefinieerd als: mensen die óf betaald werken (minimaal 1 uur per week) óf betaald werk zoeken (voor minimaal 1 uur per week), en die géén student zijn.

We onderscheiden daarbij de *werkzame beroepsbevolking* (dus de werkenden), en *werkloze beroepsbevolking* (die dus geen student zijn en wel naar betaald werk zoeken).

De som die Studiekeuze123 op het niveau per Studie (landelijke opleiding) maakt is: *werkloze beroepsbevolking* gedeeld door *totale beroepsbevolking*. Wij maken bij de berekening van het gemiddelde gebruik van de weefactor in het bestand.

Een fictief voorbeeld. Er zijn 50 afgestudeerden van een bepaalde studie in de HBO-monitor. 10 van hen studeren verder, 10 van hen worden huisman-/vrouw, 27 gaan werken en 3 zoeken werk. De totale beroepsbevolking voor deze studie is dan 30. En 3 van hen zijn werkloos. Het percentage werkloosheid is dan voor die studie 3 gedeeld door 30: 0,1 oftewel 10%. Het aantal werklozen wordt dus gepercenteerd op de beroepsbevolking.

Op de uitkomsten per Studie passen we een weergaveregel toe: als het totaal aantal respondenten dat we op dit niveau beschouwen (de totale beroepsbevolking voor die Studie dus) kleiner dan 20 is, geven we dit percentage niet weer.

Baan op niveau

Een ander arbeidsmarktgegeven dat we presenteren in het hbo op basis van de HBO-monitor, is werk 'op niveau'. Immers, werkloosheid kan laag zijn onder afgestudeerden, maar als die werkenden allemaal een laaggeschoold bijbaantje hebben in plaats van dat ze een 'echte' hbo-baan hebben, is dat relevante informatie voor studiekeizers. Net als werkloosheid presenteren we dit gegeven alleen per Studie (landelijk) en maken we geen onderscheid naar instelling.

De HBO-monitor vraagt aan *werkende* afgestudeerden onder meer welk opleidingsniveau vereist is voor hun huidige functie. Als dat het niveau van de gevolgde opleiding is, óf hoger, dan geldt dat als een baan op niveau.

We geven hier het gewogen percentage van de werkende beroepsbevolking weer dat op niveau werkt. NB: we selecteren hier de *werkende* afgestudeerden, dus we rekenen niet met dezelfde noemer als voor de berekening van het werkloosheidspercentage. Het kan dus bijvoorbeeld zo zijn dat een bepaalde studie 20% werkloosheid heeft, maar 90% van de werkenden een baan op niveau.

We maken de berekening als volgt:

1. Eerst kiezen we weer de juiste groep afgestudeerden uit de HBO-monitor. Net als bij werkloosheid gaat het hier om afgestudeerden van de *voltijds hbo-bacheloropleidingen*. Verder selecteren we hier alléén de *werkzame beroepsbevolking*, want we berekenen het percentage van *werkende* afgestudeerden t.o.v. de totale werkende beroepsbevolking.
2. De hbo-monitor bevat een variabele die is gemaakt op basis van de antwoorden op onderstaande enquêtevraag die aan werkende afgestudeerden is gesteld: *Welk opleidingsniveau werd door uw werkgever voor deze functie/werkzaamheden minimaal vereist? (Indien u als zelfstandige of freelancer werkzaam bent dan heeft deze vraag betrekking op het opleidingsniveau dat door uw opdrachtgever voor uw huidige, belangrijkste opdracht vereist werd.)*

Als het antwoord 'WO master', 'WO bachelor', 'HBO master' of 'HBO bachelor' voor hbo-bachelor afgestudeerden was, is dat gekwalificeerd als 'baan op niveau'. Andere niveaus (HBO Associate degree, HAVO/VWO, MBO, VMBO en lager) zijn gekwalificeerd als 'baan onder niveau'. NB:

deelnemers aan de hbo-monitor die deze vraag niet hebben beantwoord worden niet meegerekend in het bepalen van het percentage.

3. Met de waarde van de variabele uit de vorig stap maken we per Studie de volgende berekening: we delen het aantal mensen binnen de groep werkzame afgestudeerden die een baan op niveau hebben door de totale grootte van de werkzame beroepsbevolking. Dit levert een fractie op (tussen 0 en 1). Deze fractie wordt vermenigvuldigd met honderd en levert na afronden op hele getallen per Studie het percentage 'baan op niveau' op.

Op de uitkomsten per studie passen we weergaveregel toe: als minder 20 respondenten deel uitmaken van de werkzame beroepsbevolking, geven we dit percentage niet weer.