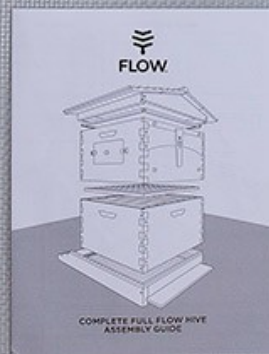


CODEBOEK VU ANALYTICS

TOELICHTING OPBOUW EN DOCUMENTATIE
DATASET VOOR ONDERZOEKERS



13-5-2020, V3

INHOUDSOPGAVE

Leeswijzer

Met inzet van VU Analytics wil de VU evidence-based, actiegerichte inzichten ontwikkelen voor verbetering van instroom, doorstroom en uitstroom van studenten ten gunste van begeleiding, beleidsvorming en wetenschappelijk onderzoek.

Dit codeboek is het basisdocument dat een onderzoeker kan raadplegen voor extra informatie bij het gebruik van de VU Analytics data. Dit codeboek geeft een overzicht van de dataextractie en -integratie, zodat deze data ook voor onderzoekers buiten het VU Analytics team gebruikt kan worden.

Colofon

Theo Bakker, Teamleider VU Analytics, Student- & Onderwijszaken, VU Analytics, t.c.bakker@vu.nl

Foto voorzijde: <http://www.thefloshow.com/flow-hive-beehive-opening/>

Inhoud	Pagina
1. Databronnen en totstandkoming van de dataset	3
2. Samenstelling van de dataset	6
I. Soorten data	6
II. Selectie	7
III. Inschrijvingen	8
IV. Besluiten	9
V. Variabele namen	10
VI. Studentnummers	11
3. Bestandsformaten	12
4. Citatie van de dataset	13
5. Categorieën	14
 Bijlagen	
1. Versiegeschiedenis	45

1. DATABRONNEN EN TOTSTANDKOMING VAN DE DATASET

VU Analytics maakt gebruik van diverse databronnen die afkomstig zijn van verschillende instanties. Hieronder vindt u een overzicht van deze instanties.

CBS	Het Centraal bureau voor de Statistiek publiceert openbare datasets die VU Analytics combineert met de studentendata als achtergrondinformatie.
C&M	Communicatie en Marketing afdeling van de VU levert informatie over de introductie en voorlichting van nieuwe studenten. Daarnaast ontvangt VU Analytics data vanuit Selligent, het CRM systeem van C&M.
DUO	Dienst Uitvoering Onderwijs biedt online datasets aan. Wanneer deze sets niet online staan kan een specifiek verzoek worden gedaan voor een bepaalde dataset.
IO	Het International Office van de VU levert data over de internationale studenten die via het IO zijn aangemeld.
GIS-VU	De onderzoeksgroep GIS-VU (Geographical Information Sciences) heeft eenmalig data geleverd over afstanden tussen middelbare scholen en de VU.
MIVU (vanaf 2019 VU data)	Mijn Informatie VU (MIVU) is een project dat gegevens onttrekt aan verschillende bronsystemen en deze informatie combineert. Dit zijn interne bronnen zoals studenten- en personeelsinformatiesystemen en ook externe bronnen zoals 1cijferHO en NSE scores.
NOA	NOA is een psychologisch advies- en onderzoeksbureau dat nauw samenwerkt met de sectie Sociale- & Organisationspsychologie van de VU. NOA verzorgt o.a. de matchingsvragenlijst voor de VU.
OKP	Het VU team Onderwijskwaliteit en procesmanagement levert een deel van de gegevens over de Pre University College aan VU Analytics.

1. DATABRONNEN EN TOTSTANDKOMING VAN DE DATASET

OTIR	OTIR is het team Onderwijsevaluaties, Toetskwaliteit en Institutional Research. Dit team verwerkt en levert de gegevens van de NSE aan VU Analytics. De Nationale Studenten Enquête (NSE) is een grootschalig jaarlijks landelijk onderzoek waarin studenten hun mening geven over hun opleiding.
SOZ	De afdeling Student en Onderwijzaken beheert het studentinformatiesysteem SAP Supplier Lifecycle Management. Hierin wordt alle administratie van de VU Studenten bijgehouden, zoals bijvoorbeeld inschrijvingen en resultaten.
Taalcentrum-VU	Het Taalcentrum-VU is een zelfstandig communicatiebureau waarmee de VU nauw samenwerkt. Het Taalloket verzorgt VU-brede taaltoetsen en bijspijkercursussen voor Nederlands en Engels.

1. DATABRONNEN EN TOTSTANDKOMING VAN DE DATASET

De brondata wordt verzameld, gecombineerd, opgeschoond en getransformeerd en omgezet naar nieuwe datasets. Een onderzoeker krijgt een subset van deze data.

Brondata

Databestanden zijn gepseudonimiseerd en op een beveiligde netwerkschijf gezet.



Bronnen

- DUO
- CBS
- MIVU / SAP SLM
- Selligent
- NOA
- NSE
- Alumni
- Etc.

Datamodel & datamanipulatie

De data wordt vervolgens in tabellen omgezet en verder bewerkt



Modellering

- Tabellen per sleutel/niveau

Manipulatie

- Transformeren, schonen, filteren
- Start met inschrijvingstabel, toevoegen van additionele informatie.

Resultaten

Vervolgens worden er nieuwe sets gemaakt voor verschillende doeleinden



Analysesets

Gecombineerde data, één regel per inschrijving. Resultaten met één regel per vak

Verrijkt 1CHO bestand

Een imitatie van het 1CHO bestand voor gebruik in het najaar

Dashboards

Op basis van subsets van de analysesets

Extracten voor wetenschappers

Op basis van subsets van de analysesets

2. SAMENSTELLING VAN DE DATASET SOORTEN DATA

Als basis voor verschillende soorten analyses, beleidsontwikkeling en wetenschappelijk onderzoek zijn de volgende soorten data verzameld:



Geslacht
Leeftijd op 1 oktober
Land van herkomst VO

Demografie

Vooropleiding & aansluiting

Profielkeuze
Vooropleiding en onderwijsinstelling
Tussenjaren / Jaren sinds diploma
Verblijfsjaren
Voorlichting en introductie
Matching



Stroominformatie
Geografische spreiding (GIS)
Kwaliteit scholen
Achtergrond scholen

Instroom

Studiesucces & tevredenheid

Eindexamencijfers
Uitslagen taaltoets VU
Studieprestaties aan de VU
Honours en Cum Laude
Studenttevredenheid
Uitval en diplomarendement
Alumnidata



Bijzondere persoonsgegevens

Bijzondere persoonsgegevens zijn in principe uitgesloten. Gebruik van bijzondere persoonsgegevens kan alleen met toestemming van een student. Als toestemming is verleend zijn deze variabelen beschikbaar. De beschrijving van de categorieën etniciteit en functiebeperking zijn pas dan van toepassing.

2. SAMENSTELLING VAN DE DATASET **SELECTIE**

De analyseset is opgebouwd met één rij voor elke student per studiejaar. Daarnaast is er een selectie gemaakt van inschrijvingen die meegenomen zijn in de analyseset. Hierbij zijn de landelijke 1CHO bestanden leidend*.

1. **Alle studenten per studiejaar**

Inschrijvingen EOI & herinschrijvingen
per studiefase
+ geaggregeerde EC
+ succesvariabelen
per studiejaar

2. **Selectie**

Opleidingsvorm = 1, 2 of 3
Inschrijvingsvorm = S
Opleidingsvorm = Voltijd + Deeltijd
Excl. UvA studenten en AUC

Toelichting

EOI: Indicatie eerstejaars actuele opl.-instelling; Indicatie die aangeeft of een student voor het eerst in het aangegeven studiejaar in de betreffende actuele opleiding bij de VU staat ingeschreven.

S: Student,

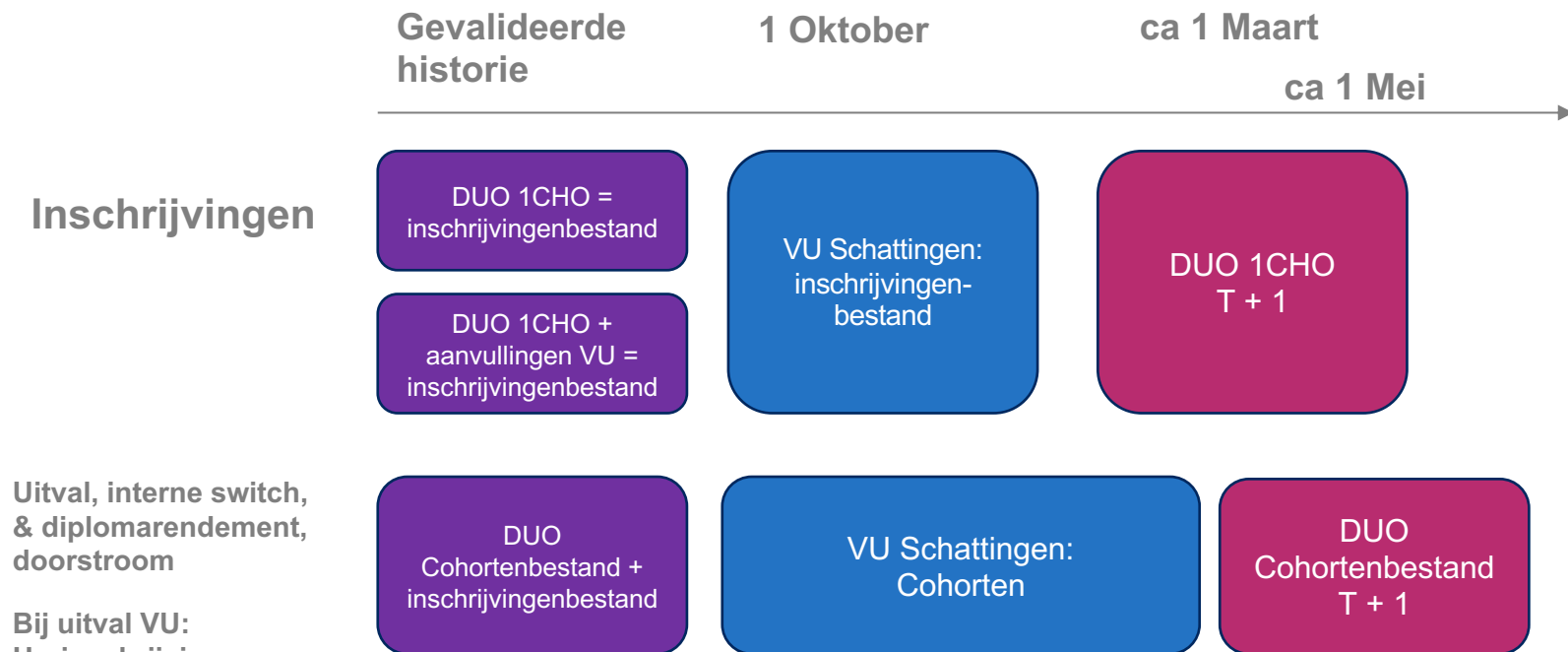
Opleidingsvorm: 1 = Voltijd, 2 = deeltijd, 3 = Duaal.

NB. Studenten van de VU van Joint degrees vallen binnen scope

**zie volgende pagina voor extra uitleg.*

2.SAMENSTELLING VAN DE DATASET **INSCHRIJVINGEN**

Vanuit gevalideerde, historische bestanden zijn inschrijvingen en rendementen 'gestapeld' opgebouwd, afhankelijk van het moment in het academisch jaar.



Uitval, interne switch, & diplomarendement, doorstroom

Bij uitval VU:
Herinschrijving en succes in HO elders.
Bij diploma VU:
doorstroom en succes binnen VU en HO elders.

Inrichtingsprincipes opbouw dataset

- De basis voor inschrijvingen en rendementen wordt gevormd door 2 bestanden van DUO: het 1CHO-Inschrijvingen en 1CHO-Cohortenbestand.
- Zolang deze bestanden niet door DUO geupdate worden vormt het schattingenbestand de basis voor aanvullingen op de bestaande DUO bestanden.

2. SAMENSTELLING VAN DE DATASET **BESLUITEN**

Besluiten over discrepanties tussen het 1CHO inschrijvingenbestand en cohortenbestand

- Er bestaat discrepantie tussen het inschrijvingenbestand en het cohortenbestand. Er is besloten uit dat het cohortenbestand leidend is, hierdoor wordt 1,7% van de rijen van het de inschrijvingenbestand verwijderd. Er wordt gewerkt aan een nieuwe oplossing.

Uitschrijvingen voor 1 oktober zijn meegenomen in de analyseset

Dit betreft volledige inschrijvingen op 1 september, maar deze studenten schrijven zich vóór 1 oktober uit. Dit kan verschillende redenen hebben zoals:

- Student was al niet meer van plan om deze studie te volgen, en komt nooit opdagen.
- Student vindt binnen een maand de studie niets, en valt dus echt uit.
- Student loot zich in, om feitelijk pas een jaar later pas echt aan de opleiding te studeren.
- Student wordt later toch geplaatst bij een andere opleiding (evt. bij de VU zelf) en studeert aan deze opleiding verder.

Besloten is deze studenten wel mee te nemen in de analyseset om de volgende redenen:

- Tussen de verschillende categorieën kunnen de data geen sluitend onderscheid maken.
- Als we deze studenten niet zouden meenemen in de analyses, kunnen we ook geen uitspraak over deze groep kunnen doen ten aanzien van uitval.
- Beredeneerd vanuit begeleiding, is het juist belangrijk om deze studenten zo snel mogelijk te kunnen begeleiden – juist omdat ze zo snel uit de opleiding vallen (en met enige waarschijnlijkheid advies kunnen gebruiken bij het vinden van een geschikte opleiding).

2. SAMENSTELLING VAN DE DATASET **VARIABELE NAMEN**

Variabelen hebben de volgende naamconventies:

- Elke variabele wordt voorafgegaan door een prefix van 3 hoofdletters (zie hoofdstuk categorieën voor een toelichting).
- Elke naam is leesbaar in begrijpelijk Nederlands; afzonderlijke woorden worden gesplitst door een _.
- Daar waar continue variabelen zijn omgezet naar categorieën, wordt het suffix _cat gebruikt.
- P staat voor Periode = Onderwijsperiode. Een jaar heeft 7 onderwijsperiodes: 1 t/m 6 zijn de reguliere onderwijsperiodes en 7 is de herkansingsperiode in de zomer.

2. SAMENSTELLING VAN DE DATASET STUDENTNUMMERS

Voor studentnummers geldt het volgende:

- Op basis van studentnummers zijn verschillende bestanden aan elkaar gekoppeld binnen het project.
- Studentnummers zijn vervolgens versleuteld. De encryptie is zo uitgevoerd dat deze ook bij meerdere inschrijvingen en in de resultatentabel gelijk is. Een studentnummer ziet er daardoor bijvoorbeeld zo uit:
739e3b71d0baf3d3648dfc51b41e333d0078ed260f640b92066d4d21c3cd9d7746ccb
- De encryptie is voor het projectteam niet terug te draaien, maar kan wel herhaald worden over verschillende periodes.
- Alle overige sleutels die mogelijk gebruikt zouden kunnen worden voor identificatie komen niet voor in de geleverde dataset: voorletters, voornaam, achternaam, geboortedatum, e-mail adres, sofinummer, onderwijsnummer, VUnetID, SAP studentnummer, Selligent nr.

3. BESTANDSFORMATEN

VU Analytics maakt gebruik van verschillende soorten bestandsformaten:

Codeboek	.pdf
Dashboard	.twb (Tableau Workbook)
Dataset	.Rds
Kwaliteitsrapport subset	.html
Standaard bestanden	.csv

4. CITATIE VAN DE DATASET

De gewenste manier van verwijzing/citatie naar een geleverde dataset is:

VU Analytics, afdeling Student- & Onderwijszaken Vrije Universiteit Amsterdam (jaartal). [*Naam DATASET*]. Geleverd op aanvraag op [datum].


In-text citatie:

VU Analytics (jaartal) of (VU Analytics, jaartal).

5. CATEGORIEËN

De data van VU Analytics is onderverdeeld in 28 categorieën. Elke categorie bevat verschillende variabelen die bij deze categorie horen. Op de volgende pagina's wordt elke categorie toegelicht met een beschrijving. Daarnaast wordt de populatie waarover de variabelen binnen de categorie bekend zijn gegeven. Tenslotte wordt informatie verstrekt over de databronnen en de samenstelling van de data.

In veel gevallen is de data beschikbaar vanaf cohort 2010 omdat in bij de conversie naar SAP SLM de inschrijvingen zijn meegenomen vanaf 2010 (zie [pagina 7](#)). In hoofdstuk 1 van dit codeboek kan meer uitleg gevonden worden over instanties waarvan de databronnen afkomstig zijn.

 Als de categorie bijzondere persoonsgegevens bevat staat dit icoon rechts bovenin de hoek. Deze worden alleen na toestemming van een student verstrekt.



4. CATEGORIEËN

	Categorie	Prefix	Toelichting
1	Alumni	ALU_	Alumnigegevens
2	BSA	BSA_	Studieadvies
3	Bijspijkercursus	BSC_	Bijspijkercursus taaltoets
4	Buitenland	BUI_	Buitenland
5	CBS	CBS_	Centraal Bureau voor de Statistiek
6	Demografie	DEM_	Demografie
7	Etniciteit	ETN_	Etniciteit (bijzonder persoonsgegeven)
8	Experiment	EXP_	Experiment VU Analytics
9	Functiebeperking	FBP_	Functiebeperking (bijzonder persoonsgegeven)
10	Honoursprogramma	HNP_	Honoursprogramma
11	Inschrijvingen	INS_	Inschrijvingen
12	Introductieweek	INT_	Introductieweek
13	Matchingactiviteit	MAC_	Matchingactiviteit
14	Matchingsvragenlijst	MVR_	Matchingsvragenlijst

4. CATEGORIEËN

	Categorie	Prefix	Toelichting
15	Middelbareschool	VOP_	Middelbare school
16	NSE	NSE_	Nationale Studenten Enquête
17	Opleidingen	OPL_	Opleidingen
18	Oriëntatie	ORI_	Oriëntatie
19	Outliers	OUT_	Outliers
20	Premasterassessment	PMA_	Premaster-assessment
21	Pre University College	PUC_	Pre University college
22	Resultaten	RES_	Resultaten
23	Schoolbezoeken	SBZ_	Schoolbezoeken
24	Succesvariabelen	SUC_	Succesvariabelen: uitval, diploma, etc
25	Taaltoets	TTS_	Taaltoets Nederlands of Engels
26	Uitwisseling	UIT_	Uitwisseling
27	Voorzieningen	VRZ_	Voorzieningen (bijzonder persoonsgegevens)

Beschrijving	De alumni data bevat twee delen: data vanuit de VU Alumni Monitor en data vanuit de Nationale Alumni Enquête (NAE). De VU Alumni Monitor is een enquête van de VU met het doel om inzicht te krijgen in beroepsrelevante competenties, carrièrepatronen en de maatschappelijke betrokkenheid van VU alumni. De NAE is een landelijke enquête van de VSNU.
Populatie	VU Alumni van masteropleidingen.
Bijzonderheden	<ul style="list-style-type: none">• Een cohort wordt 6 en 11 jaar na afstuderen gevraagd de enquête in te vullen.• Alleen de NAE informatie van examencohort 2013 is beschikbaar.• De VU Alumni monitor betreft enkel een enquête uit 2015 waarbij de examencohorten 2003 en 2008 bevraagd zijn.
Bronnen	VU Alumni Monitor en NAE
Levering	De data zijn geleverd door C&M.

BSA – BINDEND STUDIE ADVIES

Beschrijving	Deze categorie bevat variabelen m.b.t. BSA-status van de student: het tussentijdse advies, het eindejaarsadvies, het behalen van de kwalitatieve eis door de student en het academische jaar van de beoordeling.
Populatie	Eerstejaars bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf cohort 2012; vanaf toen is het BSA VU-breed ingevoerd.• Studenten die zich uitschrijven voor 1 februari hebben geen BSA eindejaarsadvies.• Studenten die een joint degree volgen kunnen zich zowel bij de UvA als VU inschrijven, daardoor is er niet altijd een BSA advies bij de VU geregistreerd.
Bronnen	SAP-SLM
Levering	De data zijn geleverd door SOZ.

BSC – BIJSPIJKERCURSUS NEDERLANDS

Beschrijving	<p>Deze categorie bevat variabelen over de bijspijkerkursus van de Nederlandse taaltoets.</p> <p>Hiervan is bekend: het jaar van deelname, of een student de cursus heeft afgerond en het resultaat van de student.</p>
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data is beschikbaar vanaf cohort 2010.• De bijspijkerkursus is verplicht voor studenten die op de Taaltoets Nederlands de categorie <i>laag</i> gescoord hebben. De cursus wordt aangeboden door het Taalcentrum-VU. Er zijn twee cursusmomenten, oktober-december en februari-maart. De cursus bestaat uit 6 bijeenkomsten, 1 keer per week.
Bron	Taalcentrum-VU
Levering	De data zijn geleverd door Taalcentrum-VU.

Beschrijving	Deze categorie bevat variabelen met informatie over studenten met een buitenlandse vooropleiding die bij de VU komen studeren en waarvan de aanmelding verloopt via het International Office (IO).
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2015; toen is het IO begonnen met registreren.• Deze categorie bevat niet alle studenten die als buitenlandse student kunnen worden aangemerkt, maar enkel de studenten die via het IO zijn aangemeld. Studenten uit Suriname vallen hier bijvoorbeeld buiten.• Omdat buitenlandse diploma's niet kunnen worden gecontroleerd via registers van DUO, worden deze diploma's door het International Office geverifieerd.
Bronnen	Database IO, Excel IO, VN-Landen
Levering	De data zijn geleverd door het IO.

Beschrijving	Deze data bevat publieke informatie over de leefomgeving van de student op basis van postcode van de middelbare school. Variabelen zijn het percentage huishoudens met een hoog inkomen op wijkniveau, percentage eenpersoonshuishoudens op wijkniveau, gemiddelde huishoudensgrootte op wijkniveau en de stedelijkheid van de wijk.
Populatie	Alle inwoners van Nederland
Bijzonderheden	<ul style="list-style-type: none">• Deze data is gekoppeld op adres van middelbare school.• Beschikbaar voor de kalenderjaren 2011 tot 2013, de data is gekoppeld aan het jaar dat het dichtste bij het jaar van de gekoppelde data ligt (bijvoorbeeld: data uit 2010 wordt gekoppeld aan 2011, data uit 2018 aan 2013).
Bronnen	CBS
Levering	De data zijn gedownload van de website van het CBS.

DEM - DEMOGRAFIE

Beschrijving	Deze categorie bevat variabelen over demografie. Hieronder valt data over de leeftijd, het geslacht en de nationaliteit(en) van de student. Ook bevat deze data de overlijdensdatum van overleden studenten.
Populatie	Bachelor-, premaster- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf 2010 tot heden (m.u.v. data overledenen).• De data over overledenen zijn beschikbaar voor 2010 – 2016.
Bronnen	MIVU (via 1CHO), SAP
Levering	De data is geleverd door SOZ en MIVU.

Beschrijving	Deze categorie bevat variabelen met informatie over de etniciteit van de student en diens ouders.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf cohort 2010 en worden alleen met toestemming van de student verstrekt.• Deze informatie valt onder de wet bijzondere persoonsgegevens.• De data wordt vanaf 2019 niet meer geleverd i.v.m. de AVG.• De informatie over etniciteit wordt meegeleverd door DUO en is afkomstig uit de Gemeentelijke Basis Administratie (GBA).
Bronnen	MIVU (via 1CHO)
Levering	De data zijn geleverd door MIVU.

EXP - EXPERIMENT

Beschrijving De categorie experiment bevat variabelen van een matchingspilot (experiment) van VU Analytics. De deelnemende studenten hadden zich hier vrijwillig voor aangemeld. De beschikbare gegevens zijn de aanmelding van de student voor deelname aan de pilot en de datum waarop de student zich heeft afgemeld.

Populatie Bachelorstudenten

Bijzonderheden

- Aan de matchingspilot namen 12 bacheloropleidingen deel*.
- De pilot vond plaats in 2016-2017.

Bronnen VU Analytics

Levering -

*B Aarde en Economie, B Aardwetenschappen, B Biologie, B Biomedische Wetenschappen, B Gezondheid en Leven, B Gezondheidswetenschappen, B Bedrijfskunde, B Econometrie en Operationele research, B Economie en Bedrijfseconomie, B Bewegingswetenschappen, B Pedagogische Wetenschappen, B Psychologie.

Beschrijving	De categorie functiebeperking bevat variabelen over studenten met een functiebeperking.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• Deze data valt onder de wet bijzondere persoonsgegevens en worden alleen met toestemming van de student verstrekt.• Er is informatie over 7 soorten functiebeperkingen: ADD/ADHD, Autisme Spectrum Stoornis (ASS), chronische ziekten, Dyslexie, fysieke beperkingen, psychologische beperkingen en overige beperkingen (zoals taalachterstand of doofheid). Een student kan meer dan 1 soort functiebeperking hebben.• De informatie over functiebeperkingen is door de studenten met een functiebeperking zelf aangeleverd om toegang te kunnen krijgen tot verschillende voorzieningen van de universiteit. De studenten met een functiebeperking zijn door een BIG geregistreerde arts gediagnostiseerd en hebben hiervan bewijs geleverd aan de universiteit. Dit bewijs is gecontroleerd door studentpsychologen op het moment van aanmelding.
Bron	SAP
Levering	Deze data zijn geleverd door SOZ.

HNP - HONOURSPROGRAMMA

Beschrijving	De categorie honoursprogramma bevat variabelen over de loopbaan van de student binnen het honoursprogramma. Van deze studenten is bekend wanneer ze zijn gestart met het programma en wanneer een student begonnen, bezig, geslaagd of uitgestroomd is.
Populatie	Bachelor- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf het startjaar van het honoursprogramma, academisch jaar 2009/2010.• Het honoursprogramma in de master is enkel voor de studie M Geneeskunde.
Bronnen	MIVU
Levering	De data zijn geleverd door MIVU.

INS - INSCHRIJVINGEN

Beschrijving	De categorie inschrijvingen bevat de informatie over de inschrijving(en) van een student. Variabelen in deze categorie zijn bijvoorbeeld de opleidingsnaam, het studentnummer, de diplomadatum, de gevolgde minoren en het aantal verblijfsjaren in het hoger onderwijs.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf cohort 2010.• Deze data zijn gebaseerd op jaarlijkse 1CHO bestanden en SAP bestanden. De levering van deze bestanden is in mei, waardoor de data uit het huidige jaar afkomstig is uit SAP. Deze data worden maandelijks geüpdatet (zie p.8 voor meer informatie).• Als een variabele niet in de maandelijks bestanden aanwezig is, maar wel in 1CHO (bijv. vooropleiding voor HO), dan wordt deze variabele opgevuld met informatie vanuit eerdere inschrijvingen.
Bronnen	MIVU
Levering	De data zijn geleverd door MIVU.

INT - INTRODUCTIE

Beschrijving	Deze data bevat gegevens over deelname van de student aan de VU introductie. Er is informatie beschikbaar over of de student zich heeft aangemeld, of de student aanwezig was en of het een 5-daagse of 2-daagse introductie betrof.
Populatie	Bachelor-, premaster- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data is beschikbaar vanaf cohort 2010.• 2-daagse variant van de introductie is ingevoerd in 2017, deze bestaat naast de 5-daagse variant. Voor 2017 duurde de introductie altijd 5 dagen.• De kwaliteit van de data in de eerste jaren is minder goed omdat de registratie voor de introductie nog niet gekoppeld was aan studentnummers. Er is eenmalig een match o.b.v. emailadressen uitgevoerd om de data te koppelen.
Bronnen	Selligent
Levering	De data is afkomstig van C&M.

MAC - MATCHINGACTIVITEIT

Beschrijving	Deze data bevat informatie over de matchingactiviteit van de student. Er is informatie over of een student vrijstelling had voor de matchingactiviteit, of de student aanwezig was, de datum van de matchingactiviteit en het resultaat van de activiteit.
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2014.• Opleidingen bieden minimaal 1x per jaar een matchingsactiviteit aan.• Als een student voor 1 februari stopt en zich het jaar daarop inschrijft voor dezelfde opleiding hoeft de activiteit niet opnieuw bijgewoond te worden.• Vrijstelling voor matchingsactiviteiten kan verkregen worden door persoonlijke omstandigheden, door onderwijsverplichtingen op momenten van de activiteiten of als de student moet deelnemen aan een verplichte matchingsactiviteit bij een andere opleiding. Of vrijstelling van toepassing is wordt beoordeeld door de toelatingscommissie van de betreffende faculteit.
Bronnen	SAP
Levering	De data zijn geleverd door SOZ.

MVR - MATCHINGSVRAGENLIJST

Beschrijving	Deze data bevat informatie uit de matchingsvragenlijst van alle bacheloropleidingen aan de VU, dit is een onderdeel van studiekeuzeactiviteiten door VU-Matching. Dit is een VU-brede digitale vragenlijst en bestaat uit de onderdelen: oriëntatie, beeldvorming en keuze, academische vaardigheden; vakinhoudelijke interesse en de persoonlijke studiesituatie.
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2014.• Het invullen van de vragenlijst duurt 25 minuten en de student ontvangt hier een automatisch gegenereerde rapportage van.• Het invullen van de matchingsvragenlijst is (als de student geen vrijstelling heeft) een voorwaarde om toegelaten te worden tot het eerste jaar van een reguliere bacheloropleiding aan de VU.• Vrijstelling voor matchingsactiviteiten kan verkregen worden door persoonlijke omstandigheden, door onderwijsverplichtingen op momenten van de activiteiten of als de student moet deelnemen aan een verplichte matchingsactiviteit bij een andere opleiding. Dit wordt bepaald door de toelatingscommissie.
Bronnen	NOA
Levering	De data zijn geleverd door NOA.

VOP - VOOROPLEIDINGEN & MIDDELBARE SCHOOLCIJFERS

Beschrijving	<p>Deze categorie bevat de vooropleidingen en middelbare schoolcijfers van studenten.</p> <p>Daarnaast bevat de categorie ook informatie over de reistijden van het adres van de scholier naar de Nederlandse universiteiten. Verder bevat de categorie informatie over de middelbare school zoals gemiddelde cijfers van de examenkandidaten, aantal gezakte en geslaagde examenkandidaten en reisafstanden van de school naar de verschillende Nederlandse universiteiten.</p>
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2010.• Middelbare data zijn alleen beschikbaar voor studenten afkomstig van Nederlandse middelbare scholen.• Reisafstanden zijn berekend vanaf de hoofdvestiging van de middelbare school, dit is niet altijd de vestiging waar de student naar school ging. De afstanden zijn bepaald door GIS-VU, dit is gebaseerd op de kortste route met het OV in 2015 in de spits.• Studenten leveren zelf diploma's aan die worden gecontroleerd door de centrale studenten administratie.
Bron	SAP, DUO, GIS-VU
Levering	De data zijn geleverd door MIVU, DUO & GIS-VU.

NSE – NATIONALE STUDENTEN ENQUÊTE

Beschrijving	<p>De categorie NSE bevat informatie afkomstig uit de Nationale Studenten Enquête (NSE).</p> <p>Hierin worden studenten bevraagd over hun mening over de opleiding die ze volgen, zoals bijvoorbeeld de studielast, de inhoud van de opleiding, studieroosters, activiteiten buiten de studie zoals bijbanen etc.</p>
Populatie	Bachelor-, premaster- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf jaar 2010.• De NSE wordt jaarlijks afgenomen in februari/maart.
Bronnen	Studiekeuze-123
Levering	De data wordt verkregen via OTIR

Beschrijving	Deze data bevat opleidingskenmerken van alle opleidingen aan de VU. De kenmerken betreffen de matchingsactiviteit van een opleiding, of de opleiding een numerus fixus hanteert, de instructietaal, het BSA en de CROHO code.
Populatie	Bachelor-, premaster- en masteropleidingen
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2010.
Bronnen	Handmatig bijgehouden tabel, gemaakt door VU Analytics. Informatie nodig voor het samenstellen van deze tabel is opgevraagd bij beleidsmedewerkers van de VU.
Levering	-

ORI - ORIËNTATIE

Beschrijving	Deze data bevat informatie over de oriëntatie van studenten voor een opleiding zoals de eerste datum van oriëntatie, of de student brochures van de opleiding heeft bekeken en wanneer de student een oriëntatieactiviteit (open dag of meeloopdag) heeft gevolgd.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2010.• De data bevat oorspronkelijk iedereen die via C&M contact heeft met de VU in het studiekeuze- en oriëntatieproces; achteraf wordt de data gekoppeld op aanmeldingen waardoor deze data beschikbaar is in de analyseset.
Bronnen	Selligent
Levering	De data zijn geleverd door C&M.

OUT - OUTLIERS

Beschrijving	Deze data zijn gemaakt door het team van VU Analytics om te gebruiken bij het modelleren en bestaat uit variabelen die zijn gemaakt op basis van de bestaande data. Het doel hiervan is het reduceren van de variantie in de data. Een voorbeeld is een variabele die aangeeft of een student in periode 3 meer dan 36 ECTS heeft behaald.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2010.
Bronnen	Gemaakt door het VU Analytics team.
Levering	-

PMA – PREMASTERASSESSMENT

Beschrijving	Deze dataset bevat informatie over het VU premasterassessment. Het assessment wordt individueel per computer afgenomen en bevat maximaal 5 onderdelen: cognitieve capaciteitentest, tekstbegrip en leesvaardigheid Engelse taal, rekenkundige vaardigheden en wiskunde/statistiek, een wetenschappelijke casus en een taaltoets tekstopbouw. Vrijstelling voor een onderdeel is niet mogelijk.
Populatie	Premasterstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2010.• Het premasterassessment is alleen verplicht voor studenten met een Nederlandse hbo-bachelor. Het is niet verplicht voor studenten met een wo-diploma of een buitenlands diploma.• Het premasterassessment is verplicht en bindend voor toelating tot de premasters van alle faculteiten m.u.v. Rechtsgeleerdheid, Godsgeleerdheid, Bètawetenschappen en de School of Business and Economics.• De studenten ontvangen 4-6 weken na het assessment een rapportage per e-mail.
Bronnen	NOA
Levering	Deze data zijn geleverd door NOA.

PUC – PRE UNIVERSITY COLLEGE

Beschrijving	Deze categorie bevat variabelen over het Pre-university college (PUC) van de VU. Het PUC organiseert verschillende programma's voor leerlingen in de bovenbouw van het VWO: masterclasses, Be Prepared, Derde klasprofielkeuze traject en Profielwerkstukstarter. De dataset bevat informatie over deelnamedatum, deelname aan gevolgde programma's en het aantal keren dat een student heeft deelgenomen.
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• Deze data zijn beschikbaar vanaf 2016.
Bronnen	Selligent
Levering	Deze data zijn geleverd door OKP.

RES - RESULTATEN

Beschrijving	Deze categorie bevat variabelen over de behaalde resultaten van de student. Er zijn data van resultaten beschikbaar op verschillende aggregatieniveaus. Ten eerste is er data beschikbaar vanuit het administratiesysteem, de resultaten voor specifieke modules die een student heeft gevolgd. Daarnaast zijn er geaggregeerde variabelen per student per periode per opleiding die de student volgt beschikbaar. Tenslotte zijn er geaggregeerde variabelen beschikbaar per studiejaar en over de studiejaren. De geaggregeerde variabelen zijn gekoppeld aan de analyseset. Variabelen zijn bijvoorbeeld het aantal behaalde EC met of zonder vrijstellingen per periode, het gewogen gemiddelde van de student tot en met studiejaar 2 en het aantal no-shows in een periode. De resultatendata is ook afzonderlijk van de analyseset aan te vragen.
Populatie	Bachelor-, premaster- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf cohort 2010.
Bronnen	SAP
Levering	De data zijn geleverd door functioneel beheer SAP.

SBZ - SCHOOLBEZOEKEN

Beschrijving	Deze dataset bevat informatie over scholenbezoeken op verschillende data voor verschillende studies. Deze scholen zijn bezocht door (student)voorlichters om informatie te verstrekken aan scholieren over specifieke opleidingen.
Populatie	Studiekeizers voor bachelorstudies.
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf 09-10-2017. Vanaf deze datum zijn schoolbezoeken geregistreerd.
Bronnen	De data is verzameld door C&M.
Levering	Deze data zijn geleverd door C&M.

SUC - SUCCESVARIABELEN

Beschrijving	Deze categorie bevat variabelen die berekend zijn door VU Analytics. De succesvariabelen zijn berekend om inzicht te krijgen in de loopbaan van de student op de universiteit. Succesvariabelen zijn bijvoorbeeld uitval per opleiding, switchers, de studieduur en het diplomarendement van herinschrijvers.
Populatie	Bachelor-, premaster- en masterstudenten.
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf cohort 2010.• Rendementsvariabelen zijn onbekend in het lopende jaar omdat gegevens zoals het behalen van het diploma nog onbekend zijn. Deze variabelen zijn in dat geval opgevuld met NA.
Bronnen	MIVU
Levering	De data zijn geleverd door MIVU.

Beschrijving	Deze categorie bevat variabelen over de Nederlandse taaltoets. De taaltoets is verplicht en afhankelijk van de opleiding wordt de Nederlandse of Engelse taaltoets afgenomen. De taaltoets wordt in september afgenomen, als deze is gemist is er een tweede gelegenheid in november. De toets kan niet herkanst worden. Als de uitslag van de toets in de categorie <i>laag</i> valt, dan is de student verplicht om de bijspijkercursus te volgen.
Populatie	Bachelorstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data voor de Nederlandse taaltoets zijn beschikbaar vanaf cohort 2010.• De data voor de Engelse taaltoets zijn beschikbaar vanaf cohort 2018, voor 2018 werd alleen de Nederlandse taaltoets afgenomen.• Studenten met dyslexie, een functiebeperking of chronische ziekte die een aanvraag voor voorzieningen toegekend hebben gekregen, mogen van deze voorzieningen gebruik maken tijdens de taaltoets.
Bronnen	Taalcentrum-VU
Samenstelling	De data zijn geleverd door Taalcentrum-VU.

Beschrijving	<p>Deze dataset bevat informatie over de internationale uitwisseling van VU-studenten.</p> <p>Variabelen in deze dataset zijn bijvoorbeeld het land waarnaar de student op uitwisseling gaat, de GPA score van de student voorafgaand aan de uitwisseling en de motivatie van de student.</p>
Populatie	Bachelor- en masterstudenten
Bijzonderheden	<ul style="list-style-type: none">• De data is beschikbaar vanaf academisch jaar 2014-2015; toen is het IO begonnen met registreren van deze gegevens.
Bronnen	IO
Levering	De data zijn geleverd door het International Office van de VU.

VRZ - VOORZIENINGEN

Beschrijving	De voorzieningen categorie bevat gegevens over de voorzieningen die zijn toegewezen aan studenten met een functiebeperking. Voorbeelden hiervan zijn toegang tot minder validen ingangen en extra tijd bij een tentamen.
Populatie	Bachelor-, premaster- en master studenten.
Bijzonderheden	<ul style="list-style-type: none">• De data zijn beschikbaar vanaf 2014.• Deze categorie bevat bijzondere persoonsgegevens.• Als een student heeft een voorziening aangevraagd wordt daarna is besloten of de aanvraag wel/niet gehonoreerd wordt. De examencommissie of het faculteitsbestuur beslist over aanvragen m.b.t. tentamens en onderwijs voorzieningen, Centrum voor Studie en Loopbaan (CSL) beslist over aanvragen m.b.t. financiële ondersteuning en centrale voorzieningen. In de data zijn alleen gehonoreerde aanvragen opgenomen.
Bron	SAP
Levering	De data zijn geleverd door SAP functioneel beheer.

BIJLAGE 1. VERSIEGESCHIEDENIS

V0.9.1 – Totstandkoming 1^e versie

V0.9.2 – Uitbreiding 1^e versie

V0.9.3 – Verwerking feedback

V0.9.4 – Conceptversie voor verspreiding

V0.9.5 – Data categorie verwijderd vanwege C.O.P.

V1 – Eerste versie na review van gebruikers (A. de Wit en R.W. van Niekerk)

V2 – Extra symbool toegevoegd voor bijzondere persoonsgegevens

V3 – CC variant

VOORWAARDEN VOOR GEBRUIK VAN DEZE PUBLICATIE (1/2)



Deze uitgave deelt de VU met externen onder de Creative Commons licentie:
Naamsvermelding-NietCommercieel-GelijkDelen.



Dit is de vereenvoudigde (human-readable) versie van de volledige [licentie](#) en geen vervanging van de volledige licentie. [Vrijwaring](#).

Je bent vrij om:

- **het werk te delen** — te kopiëren, te verspreiden en door te geven via elk medium of bestandsformaat
- **het werk te bewerken** — te remixen, te veranderen en afgeleide werken te maken
- De licentiegever kan deze toestemming niet intrekken zolang aan de licentievoorwaarden voldaan wordt.

Onder de volgende voorwaarden:

- **Naamsvermelding** — De gebruiker dient de maker van het werk te [vermelden](#), een link naar de licentie te plaatsen en [aan te geven of het werk veranderd is](#). Je mag dat op redelijke wijze doen, maar niet zodanig dat de indruk gewekt wordt dat de licentiegever instemt met je werk of je gebruik van het werk.
- **NietCommercieel** — Je mag het werk niet gebruiken voor [commerciële doeleinden](#).
- **GelijkDelen** — Als je het werk hebt geremixt, veranderd, of op het werk hebt voortgebouwd, moet je het veranderde materiaal verspreiden onder [dezelfde licentie](#) als het originele werk.
- **Geen aanvullende restricties** — Je mag geen juridische voorwaarden of [technologische voorzieningen](#) toepassen die anderen er juridisch in beperken om iets te doen wat de licentie toestaat.

VOORWAARDEN VOOR GEBRUIK VAN DEZE PUBLICATIE (2/2)



(Vervolg)



Let op:

Voor elementen van het materiaal die zich in het publieke domein bevinden, en voor vormen van gebruik die worden toegestaan via een [uitzondering of beperking](#) in de Auteurswet, hoeft u niet aan de voorwaarden van de licentie te houden.

Er worden geen garanties afgegeven. Het is mogelijk dat de licentie u niet alle gebruiksvrijheden geeft die nodig zijn voor het beoogde gebruik. Bijvoorbeeld, andere rechten zoals [publiciteits-, privacy- en morele rechten](#) kunnen het gebruik van een werk beperken.

De volledige versie van de licentie op deze publicatie is van toepassing. Zie <https://creativecommons.org/licenses/by-nc-sa/4.0/legalcode.nl>