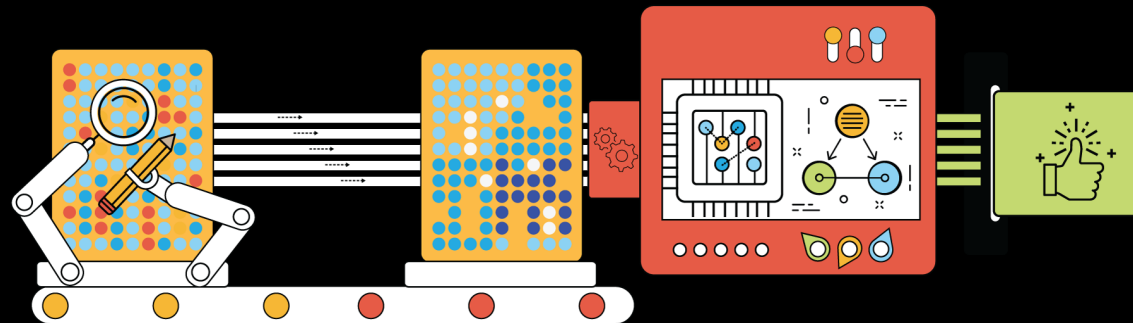


Studenten nóg beter begeleiden
met een datagedreven aanpak!



5 mbo-instellingen onderzoeken samen met DUO hoe data kan worden ingezet om de studieloopbaan van studenten te verbeteren en wat de toegevoegde waarde is om dit met een datagedreven aanpak te doen.

5 mbo-instellingen:



Noorderpoort



DUO:



Dienst Uitvoering Onderwijs
Ministerie van Onderwijs, Cultuur en
Wetenschap

Projectorganisatie:



Agenda

1 Inzichten

2 Pilot Noorderpoort

3 Pilot ROC Friese Poort

4 Pilot ROC van Twente

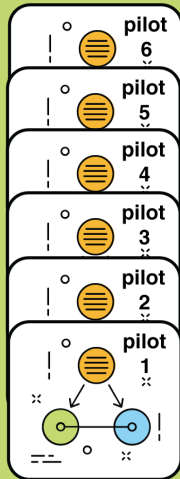
5 Pilot Summa College

6 Pilot Arcus Leeuwenborgh

7 Werkwijze

8 Vervolg

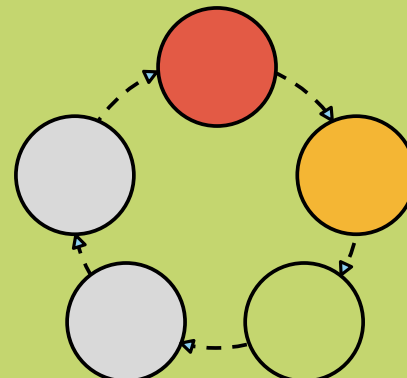
3 resultaten



6 pilots

..waarbinnen onderzoek is
uitgevoerd

levert inhoudelijke resultaten op!



Slimme werkwijze

..met praktisch materiaal
waar het hele veld gebruik van kan
maken!



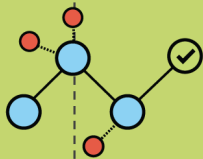
Blik in de toekomst

..van datagedreven aanpak in het
MBO!

Belangrijkste inzichten



Datagedreven onderzoek voegt waarde toe



Analyses leggen verbanden bloot die kunnen helpen bij ondersteunen van leerloopbaan



Kennishniveau binnen scholen (nog) onvoldoende om zelf analyses uit te voeren en ruwe resultaten te interpreteren



Eerste stap is gezet!

Noorderpoort

Welke kenmerken uit de vooropleiding van de student houden verband met studievertraging?

Succes van studenten staat voorop bij Noorderpoort. Als duidelijk is welke kenmerken kunnen leiden tot studievertraging dan kunnen we studenten nog beter begeleiden om die vertraging tegen te gaan. Elke vertraging, switch en 'onrust' in het opleidingstraject is ten nadele van het succes en geeft een verhoogde kans op uitval.

“We beginnen klein”

Experimenteren, meedoen en klein beginnen: leren van elkaars ervaringen

Mooie kans om met DUO te verdiepen wat mogelijk is op het gebied van data science

We verwachten niet meteen het verschil te maken, maar zetten eerste stappen in datagedreven werken

Belangrijkste lessons learned:

Redeneren vanuit de praktijk voegt waarde toe: bijv. nominale studieduur t.o.v. geprogrammeerde studieduur

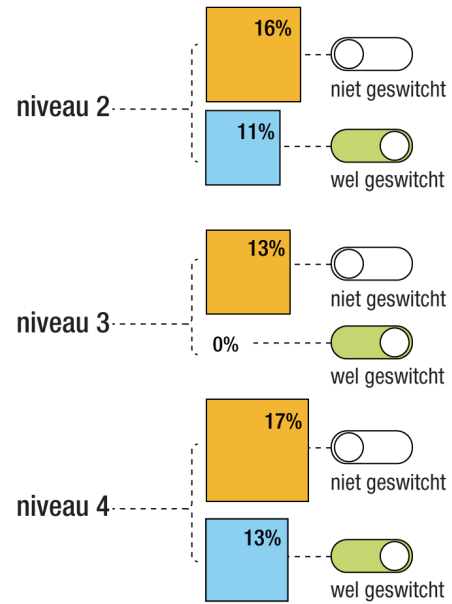
Houd de onderzoeksvraag en definities scherp of stel bedachtzaam bij

Uitkomsten leveren soms juist vervolgvragen op i.p.v. antwoorden

Resultaten

Switchen

% studenten dat langer over de opleiding doet na wel/niet switchen

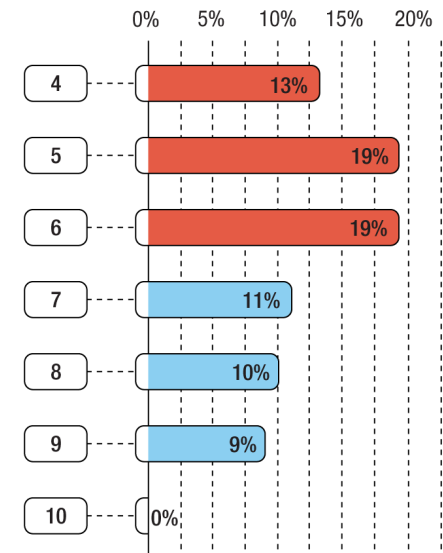


Kenmerken die een relatie hebben met vertraagde studieduur

Cijfer Engels op het VO

de invloed van het gemiddelde cijfer op het aandeel studenten dat langer over de opleiding doet

Cijfer Engels VO: % studenten langer bezig dan de geprogrammeerde studieduur:



Afwezigheid

gemiddeld aantal uren afwezig per geprogrammeerd studiejaar

Afwezigheid van studenten die **niet** langer over hun opleiding doen:

2%



0

Afwezigheid van studenten die **wel** langer over hun opleiding doen:

1%

46%



1-50



23%

omslagpunt!

32%



51-100



34%

11%



101-150



21%

5%



151-200



10%

5%



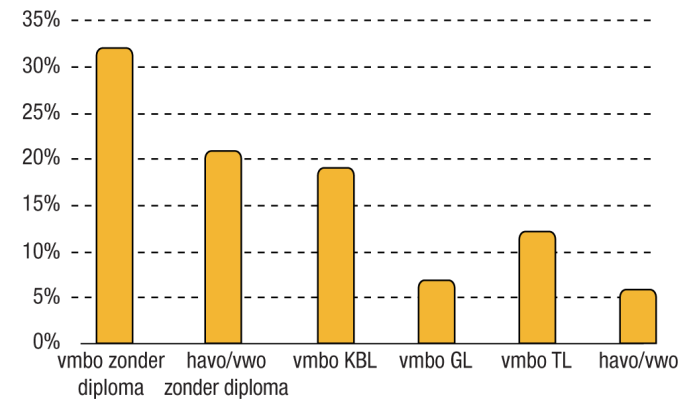
> 200



11%

Passende plaatsing

aansluiting van de vooropleiding op MBO niveau 4

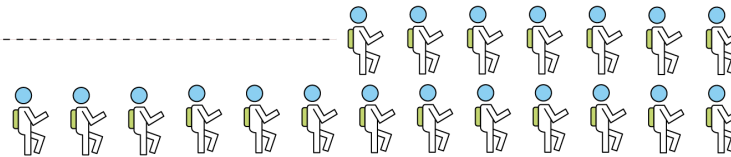


* % studenten dat langer doet over de opleiding t.o.v. de geprogrameerde studieduur

Populatie & onderzoek

Voor de analyse is gekeken naar

1967 studenten



12 onafhankelijke kenmerken

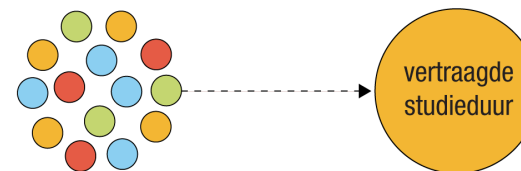
zijn onderzocht



Er is

1 afhankelijke variabele

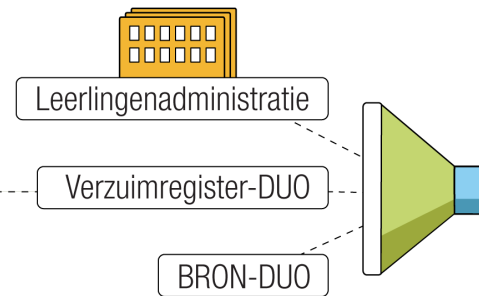
onderzocht



Er zijn

3 bronnen

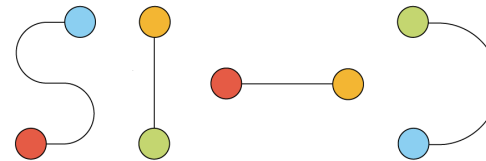
gebruikt



Er zijn

4 significante verbanden

aangetoond



Hoe ziet onze instroompopulatie eruit en wat maakt dat studenten in het eerste jaar switchen of uitvallen?

Om wendbaar, herkenbaar en zichtbaar te kunnen zijn, hebben we goede data nodig om de juiste keuzes te maken. We doen nu veel op ervaring en gevoel, maar kunnen aannames niet altijd toetsen of onderbouwen.

Data van onze toeleverende VO-scholen en de instroomdata kan gebruikt worden om de startende student in het eerste jaar een grotere kans op succes te bieden en beter te faciliteren.

“Meerdere disciplines in de wagen”

Onderzoeksvraag gebaseerd op een concrete vraag uit het onderwijs

De juiste persoon aan het stuur en diverse disciplines in de wagen

Belangrijkste lessons learned:

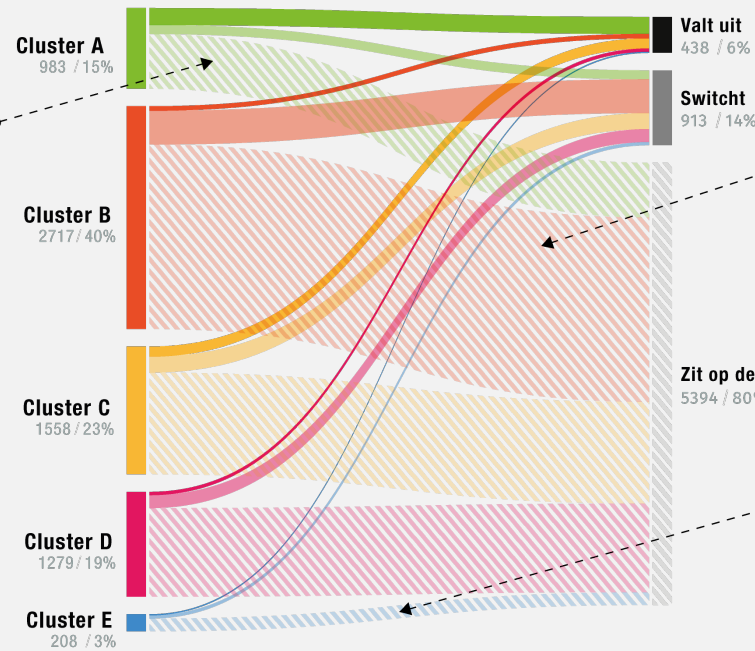
Houd het concreet en pakbaar

Niet alle benodigde data al in goede kwaliteit beschikbaar

Inzicht in welke data op welke manier vastgelegd zou moeten worden om onderzoek in de toekomst mogelijk te maken

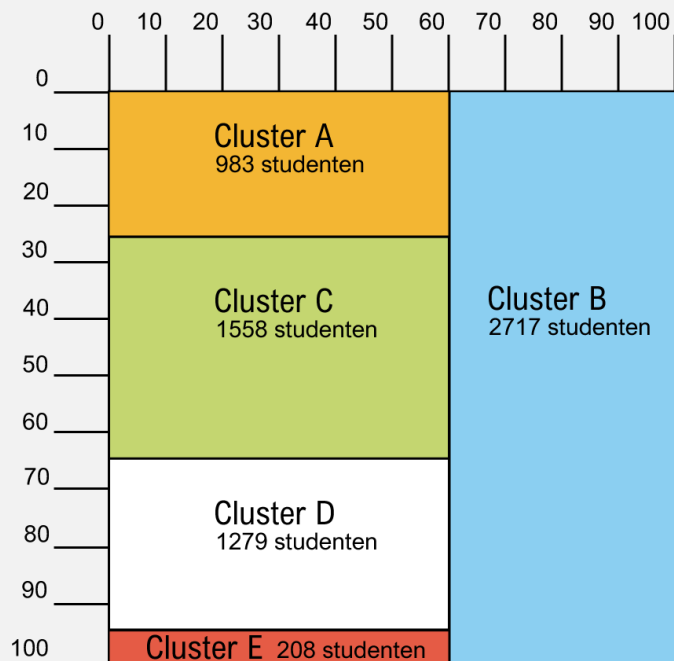
Clusters

Cluster A doet het minder goed. Het cluster bevat de meeste switchers en uitvallers. Dit cluster bestaat uit zij-instromers. De gemiddelde leeftijd is 19,5 jaar en de gemiddelde reisafstand naar de instelling is 6 km verder dan die van studenten in cluster B.



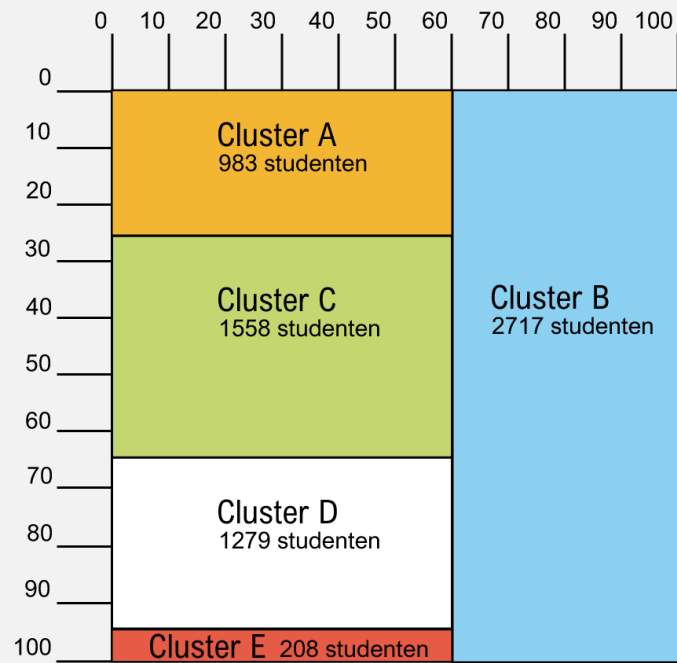
Cluster B is de grootste groep en doet het ten opzichte van de rest goed. Het bevat weinig uitvallers en een gemiddeld aantal switchers. Het bestaat uit studenten die doorstromen uit het VO naar een niveau 4 opleiding. Deze studenten zijn gemiddeld 16,4 jaar en hebben een korte reisafstand naar de instelling.

In **Cluster E** zitten de studenten waarvan bekend is dat ze een open dag hebben bezocht. De kans om uit te vallen of te switchen is voor deze groep niet opvallend anders. We weten niet of dit daadwerkelijk zo is, of komt door de beperkte hoeveelheid data die hierover beschikbaar is.



Cluster A	
Gemiddelde leeftijd	19,5
Open dag bezocht?	Nee
Gemiddelde reisafstand	19 km
Waar komen ze vandaan?	100% uit MBO
Niveau van de opleiding	3,1(zowel 1,2,3 als 4)
Cluster B	
Gemiddelde leeftijd	16,4
Open dag bezocht?	Nee
Gemiddelde reisafstand	12,5 km
Waar komen ze vandaan?	100% uit het VO
Niveau van de opleiding	100% niveau 4

Resultaten



Cluster A	
Hoe groot is de kans dat...	
... ze uitvallen?	sterk hoger
... ze goed zitten?	39%

Cluster B	
Hoe groot is de kans dat...	
... ze uitvallen?	sterk lager
... ze goed zitten?	56%

Populatie & onderzoek

Voor de analyse is gekeken naar

6745 studenten



17 onafhankelijke kenmerken

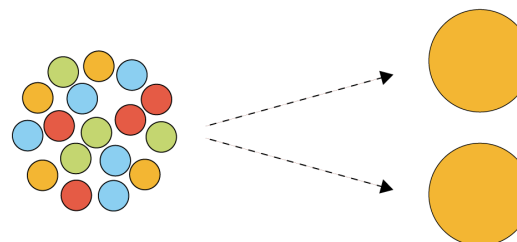
zijn onderzocht



Er zijn

2 afhankelijke variabelen

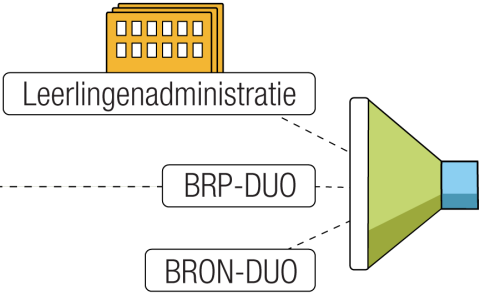
onderzocht



Er zijn

3 bronnen

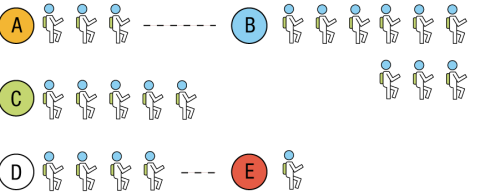
gebruikt



Het model heeft

5 clusters van studenten

opgeleverd



Welke kenmerken voorspellen of een student bij inschrijving gebruik maakt van het Studiefonds van ROC van Twente of het OCW-fonds voor de aanschaf van leermiddelen?

Nieuw ingeschreven studenten kunnen een 'valse start' hebben, bijvoorbeeld als zij op de eerste lesdag niet kunnen beschikken over de benodigde leermiddelen. Kostbare onderwijstijd gaat daardoor verloren en de student start met een achterstand. Mbo-instellingen zijn beter in staat op zulke zaken te anticiperen wanneer ze meer gebruik maken van student gerelateerde gegevens. Er valt zowel beleidsmatig als qua uitvoering het nodige 'te winnen' wanneer mbo-instellingen meer datagedreven werken.

“Stappen zetten met datagedreven onderzoek ”

Hoe mooi is het als we met data studenten nog beter kunnen helpen hun studieloopbaan succesvol te doorlopen

Het is tijd om de volgende stap te zetten om datagedreven onderzoek op een hoger plan te brengen.
Daarvoor is overleg en samenwerking nodig.

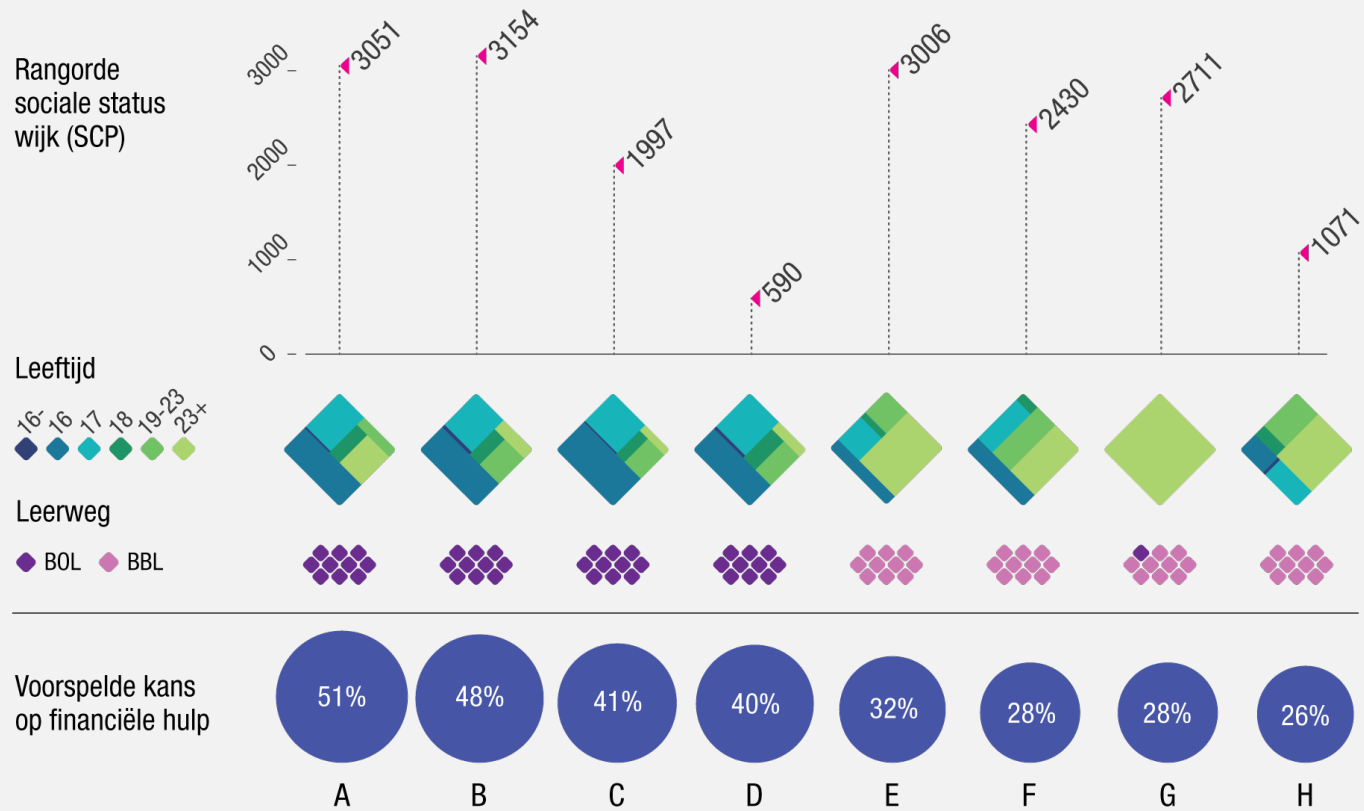
Belangrijkste lessons learned:

Alles valt of staat bij de kwaliteit van data

Bij het personaliseren van onderwijs goed nadenken over hoe we omgaan met de privacy van studenten

Datagedreven werken in de vorm van monitoren en voorspellen vraagt om nieuwe expertise en bedrijfsfuncties
en om datageletterdheid bij docenten

Resultaten

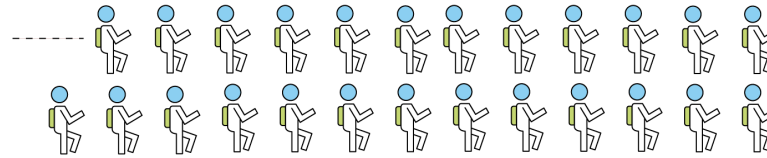




Populatie & onderzoek

Voor de analyse is gekeken naar

2526 studenten



44 onafhankelijke kenmerken

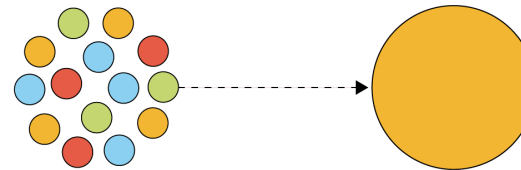
zijn onderzocht



Er is

1 afhankelijke variabele

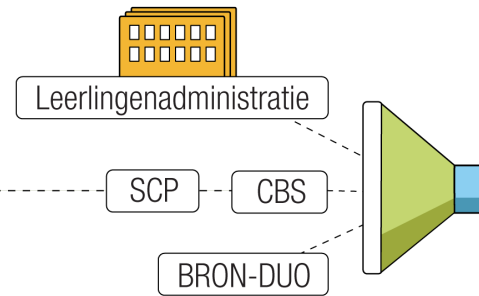
onderzocht



Er zijn

4 bronnen

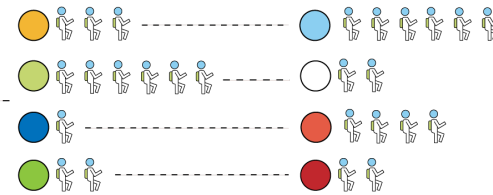
gebruikt



Het model heeft

8 clusters van studenten

opgeleverd



Welke kenmerken voorspellen uitval in het eerste studiejaar en welke verschillen of overeenkomsten kunnen we hierin herkennen tussen studenten die zich wel of niet in een kwetsbare positie bevinden?

Summa College streeft er naar om iedere student op zijn of haar bestemming te krijgen door het behalen van een diploma met doorstroom- en/of arbeidsmarktperspectief. Niet iedereen bereikt die bestemming; sommige studenten vallen in het eerste jaar al uit. Summa College wil meer weten over hoe dat komt en in hoeverre er verschillen zijn tussen studenten die zich niet of juist wel in een kwetsbare positie bevinden.

“Bundel de krachten”

Trek samen op, interpretatie van informatie is cruciaal

Een onderwijs-gedreven aanpak zorgt voor de meeste toegevoegde waarde

Naast inhoudelijke resultaten levert het samen uitvoeren van het onderzoek waardevolle inzichten over het proces op

Belangrijkste lessons learned:

De kwaliteit van data en de periode waarover data beschikbaar is, zijn bepalend voor het resultaat

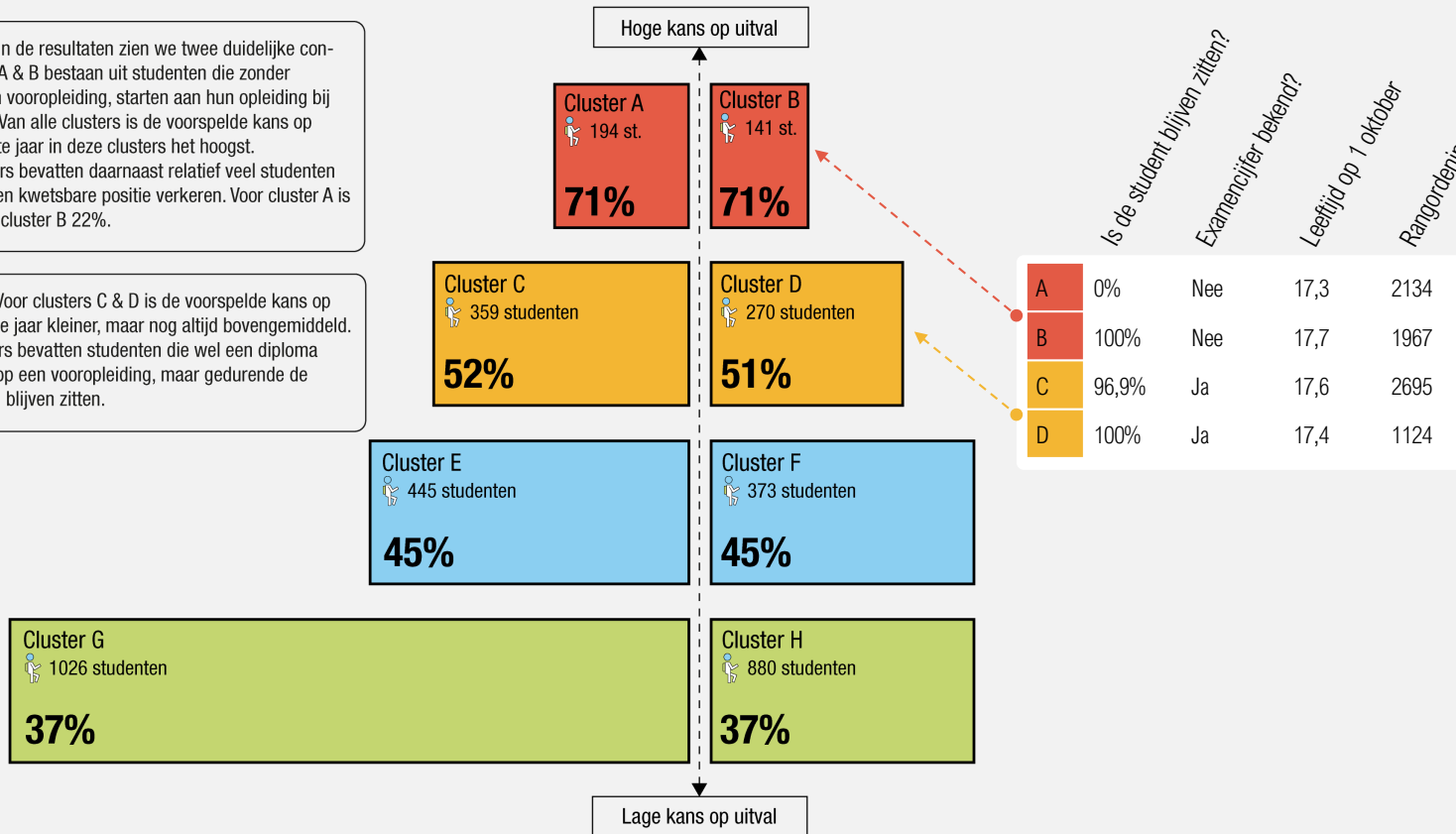
Een werkgroep met verschillende expertises voorkomt dat de samenhang ontbreekt

Focus je niet te veel op het mogelijke resultaat, maar juist op het formuleren van een goede onderzoeksvraag

Resultaten

Clusters A & B: In de resultaten zien we twee duidelijke conclusies. Clusters A & B bestaan uit studenten die zonder diploma voor hun vooropleiding, starten aan hun opleiding bij Summa College. Van alle clusters is de voorspelde kans op uitval in het eerste jaar in deze clusters het hoogst. Deze twee clusters bevatten daarnaast relatief veel studenten die mogelijk in een kwetsbare positie verkeren. Voor cluster A is dat 24% en voor cluster B 22%.

Clusters C & D: Voor clusters C & D is de voorspelde kans op uitval in het eerste jaar kleiner, maar nog altijd bovengemiddeld. Deze twee clusters bevatten studenten die wel een diploma hebben gehaald op een vooropleiding, maar gedurende de vooropleiding zijn blijven zitten.



Hoge kans op uitval

Cluster A
194 st.
71%

Cluster B
141 st.
71%

Cluster C
359 studenten
52%

Cluster D
270 studenten
51%

Cluster E
445 studenten
45%

Cluster F
373 studenten
45%

Cluster G
1026 studenten
37%

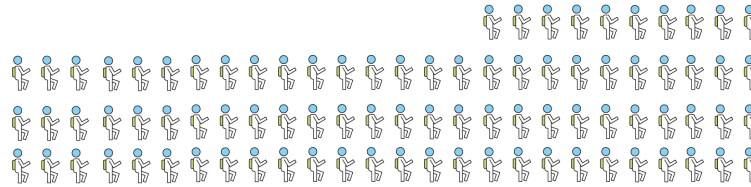
Cluster H
880 studenten
37%

Lage kans op uitval

Populatie & onderzoek

Voor de analyse is gekeken naar

8430 studenten



28 onafhankelijke kenmerken

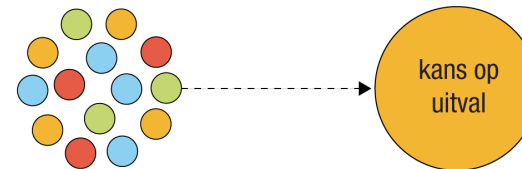
zijn onderzocht



Er is

1 afhankelijke variabele

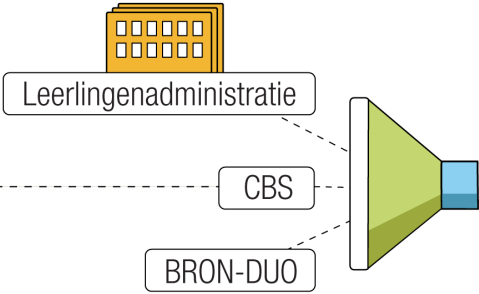
onderzocht



Er zijn

3 bronnen

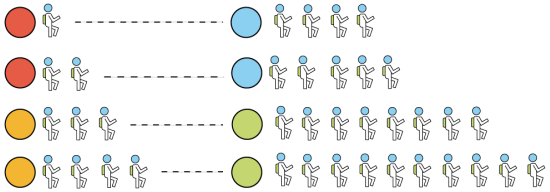
gebruikt



Het model heeft

8 clusters van studenten

opgeleverd



Welke data(bronnen) zijn op welke manier beschikbaar (te maken) en te koppelen met de eigen data om inzicht te krijgen in het studiesucces van een mbo niveau 4 student op het hbo?

Inzicht in het studiesucces van mbo-studenten op het hbo is noodzakelijk om goed in te kunnen schatten of en op welke wijze er op het mbo extra aandacht moet worden besteed aan succesvolle doorstroom. Data uit verschillende bronnen kan dat inzicht bieden. Arcus Leeuwenborgh richt zich op het inventariseren en naar boven halen van die data zodat in de toekomst datagedreven onderzoek kan plaatsvinden.

Belangrijkste lessons learned:

Samen met experts een onderzoeksvraag opstellen creëert draagvlak en prioriteit in de organisatie

Mbo, hbo en DUO beschikken samen over een flinke set aan data die gebruikt kan worden

Hoe SMART-er de onderzoeksvraag, hoe gericht het onderzoek

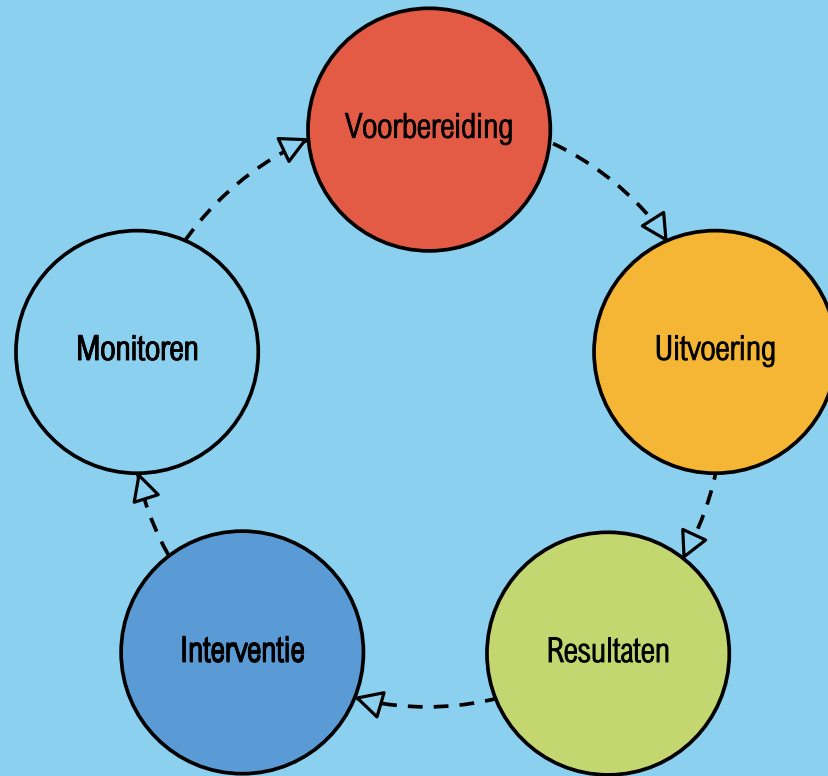
Niet alle bronnen die je nodig hebt, zijn openbaar of goed te koppelen

Betrek al bij de selectie van mogelijke variabelen een privacy specialist

Zorg dat de definitie van een kenmerk/variabele goed duidelijk is

Prioriteit binnen de organisatie is nodig om stappen te zetten

Werkwijze

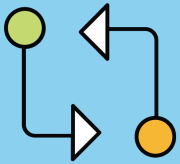
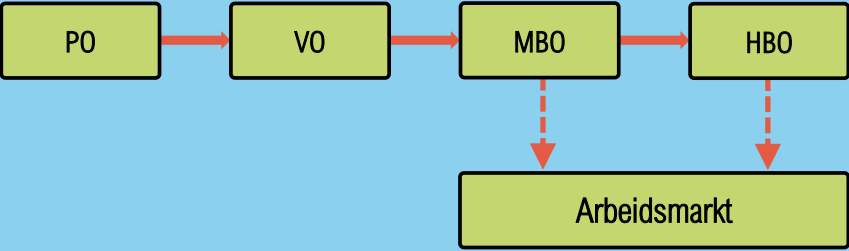




1 De juiste vraag stellen

Waarover gaat de vraag

Voorbeeldvraag: Welke kenmerken van de student, het onderwijs en de regio hebben een relatie met de uitval van mbo-studenten in het eerste jaar van hun hbo opleiding?



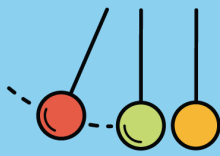
vergelijken



voorspellen



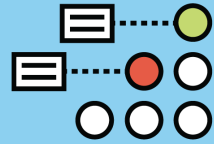
definiëren



verklaren



evalueren



beschrijven

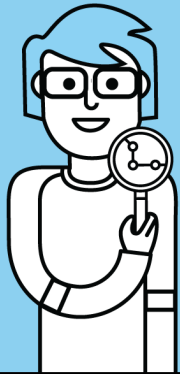
Vorbereiding

2

Wie betrek je bij het project



Onderwijs
expert



Data scientist



Informatie
manager



Bestuurder



Communicati
e
medewerker



Onderzoeker



Functionaris
gegevensbeschermi
ng



3 Privacy



Werken met data vereist zorgvuldige afwegingen ten aanzien van ethiek en privacy



In de pilotsetting is zo veel mogelijk aangesloten bij de kennis over wat wel en niet mag



Toepassing van de wetgeving op cases in de praktijk is nog niet zo zwart/wit



Binnen de pilots is de privacy van de student beschermd door o.a. rekening te houden met groepsgrootte, pseudonimisering, beveiligde omgeving van DUO



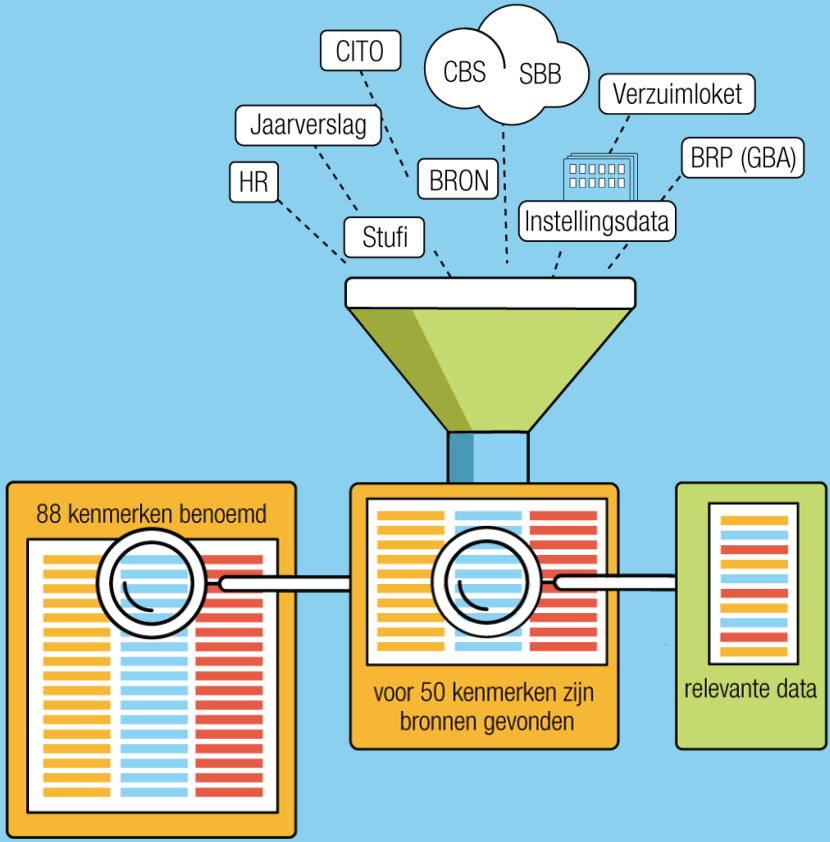
Omdat we het onderwerp serieus nemen, is er een **handreiking** uitgewerkt hoe met privacy en de AVG omgegaan kan worden in toekomstig datagedreven onderzoek door onderwijsinstellingen

Vorbereiding

4 Data verzamelen

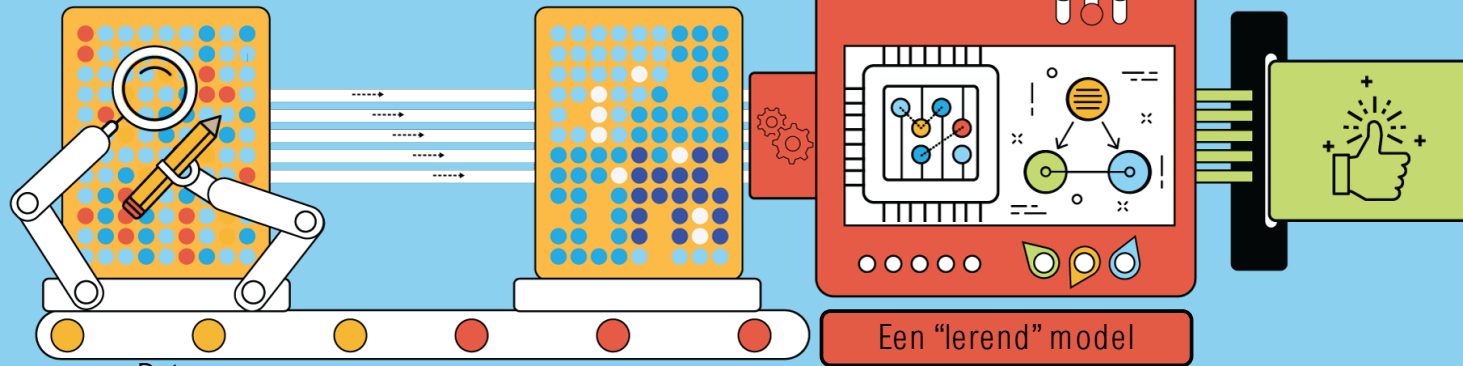
Data beschikbaar maken

Beschikbare data is de basis voor de analyse die antwoord gaat geven op de onderzoeksvraag:
beschikbaar krijgen van informatie binnen de instellingen
Op zoek naar externe relevante bronnen



Uitvoering

5 Data analyses



Data prepareren
Om data te kunnen gebruiken binnen een algoritme moet deze geschikt zijn of geschikt gemaakt worden.

Een "lerend" model

Een model moet worden getraind en getest. Afhankelijk van de parameters die je meegeeft functioneert een model beter of slechter. Daarbij zoek je altijd naar een model dat niet alleen op de huidige dataset goede resultaten behaalt maar ook op nieuwe data.

Resultaten

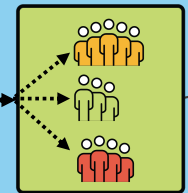
6 Variatie in resultaten

Afhankelijk van het soort onderzoek kunnen de resultaten ook uiteenlopen.

Bijvoorbeeld:

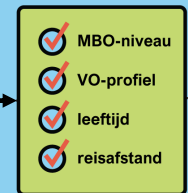
Een indeling in onderscheidbare groepen tbv beleidsvorming

Studenten



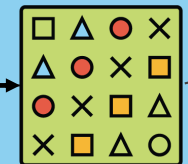
Een overzicht met kenmerken die een verband hebben met uitval in het eerste jaar

Uitval 1e jaar



Een overzicht van alle historische patronen op basis van reeds afgestudeerde studenten

Afgestud. studenten



Contact



www.datagedrevenonderzoekmbo.nl

Boudewijn – b.leunissen@vistacollege.nl

Elsa – elsa@hutspot.nl

Xaverius – xso.vanhien@summacollege.nl

Edwin – tm.vanderweit@noorderpoort.nl

Paulo – pmoekotte@rocvantwente.nl

Jaap-Jan – jaap-jan.bakker@duo.nl

Elisabeth – eleenstra@rocfriesepoort.nl