Wetenschappelijke Verantwoording Doelgroepenanalyse Laaggeletterden

Colofon

Titel	Wetenschappelijke Verantwoording Doelgroepenanalyse Laaggeletterden
Auteur	Tim Huijts, Ineke Bijlsma en Rolf van der Velden
Versie	1.0
Datum	30-1-2020
Project	Doelgroepenanalyse Laaggeletterdheid

Inhoudsopgave

Inleiding **1**

1 Doel van de clusteranalyse **2**

2 Opzet van de clusteranalyse **3**

3 Typering van de doelgroepen op basis van de clusters **11**

4 Geschatte omvang van de doelgroepen op basis van de clusters **13**

5 Schatting op gemeente- en arbeidsmarktregio-niveau **17**

6 Slotoverweging **19**

Bronnen **20**

Bijlage 1 – Aanvullende kenmerken van de doelgroepen laaggeletterden **21**

Bijlage 2 – Tabellen met geschatte omvang van de doelgroepen **23**

Inleiding

Er zijn veel rapporten over kwalitatief en kwantitatief onderzoek naar laaggeletterden (waaronder PIAAC-onderzoek, CBS-onderzoek) beschikbaar (zie o.a. Bijlsma et al., 2016; Bijlsma & Van der Velden, 2019; Buisman et al., 2013; Stichting Lezen & Schrijven, 2018). Deze rapporten bieden de gemeenten echter onvoldoende houvast om gericht beleid mee te maken voor het aanpakken van de laaggeletterdheid. De groep laaggeletterden is namelijk een zeer diverse groep, van ouderen met slechte schoolervaringen die nauwelijks technisch kunnen lezen en schrijven en de digitale wereld vermijden, tot jonge ouders die eerder van school zijn gegaan, hun smartphone veelvuldig gebruiken en tegen taalniveau 2F aanzitten. Om laaggeletterdheid aan te pakken, willen gemeenten de verschillende groepen kunnen identificeren. Een verdere definitie en segmentatie van verschillende subgroepen laaggeletterden is daarom noodzakelijk.

Om gemeenten hierbij te helpen heeft de overheid aan ECBO, ROA en Etil de opdracht gegeven om een doelgroepenanalyse uit te voeren (de Doelgroepenanalyse Laaggeletterdheid). De doelgroepenanalyse is een kennisproduct dat gemeenten inzicht en handvatten biedt bij het maken van beargumenteerde beleidskeuzes voor specifieke (sub)groepen laaggeletterden en de cursus- of trajecttypen die hier het beste bij passen. Het doel is hierbij niet om de effectiviteit van bestaand beleid in beeld te brengen.

In deze wetenschappelijke verantwoording beschrijven we eerst hoe de typering van doelgroepen voor de Doelgroepenanalyse Laaggeletterdheid tot stand gekomen is. Deze typering is gebaseerd op een clusteranalyse van laaggeletterden in de PIAAC-data, aangevuld met CBS-registerdata. De presentatie van de typering van de doelgroepen wordt voorafgegaan door een korte bespreking van het doel van de clusteranalyse en een beschrijving van de opzet van de clusteranalyse. Vervolgens presenteren we de geschatte omvang van de doelgroepen. Tot slot bespreken we hoe we op gemeente- en arbeidsmarktregio-niveau geschat hebben hoe groot het totale percentage laaggeletterden is en wat de omvang van de diverse doelgroepen laaggeletterden is. We eindigen de rapportage met een slotoverweging waarin de belangrijkste overwegingen, randvoorwaarden en beperkingen die bij deze doelgroepenanalyse een rol gespeeld hebben nog eens kort op een rij gezet worden.

De schattingen van het percentage laaggeletterden, de omvang van de doelgroepen per gemeente en arbeidsmarktregio en een verdere profilering van de doelgroepen zijn verwerkt in een online tool (www.basisvaardighedeninzicht.nl). In een aparte handreiking bieden we verdere informatie over het gebruik van deze tool.

1 Doel van de clusteranalyse

Het doel van de clusteranalyse is om een beperkt en hanteerbaar aantal subgroepen te identificeren die 'bruikbaar' zijn voor het beleid. In eerdere analyses is een beschrijving gegeven van de omvang van doelgroepen per gemeente, maar dat is toen gebeurd op basis van afzonderlijke kenmerken (bijvoorbeeld de leeftijdsopbouw naar gemeente). Deze informatie geeft voor gemeenten wel inzicht in de aantallen, maar biedt nog onvoldoende handvatten om gericht beleid op te maken om de laaggeletterdheid aan te pakken. Aanvullende en meer concretere informatie over de verschillende subgroepen laaggeletterden is zeer wenselijk. We gebruiken clusteranalyse om onderscheid te maken in de verschillende subgroepen laaggeletterden. Clusteranalyse kan laten zien of er binnen de groep laaggeletterden subgroepen voor komen die zich van elkaar onderscheiden.

2 Opzet van de clusteranalyse

We maken een typering van de doelgroepen op basis van een aantal objectieve kenmerken. Hiervoor maken we gebruik van PIAAC-data, aangevuld met de CBS-registerdata. We richten ons op de bevolking tussen 16 en 65 jaar, omdat de PIAAC-data alleen betrekking heeft op deze groep. Het is daardoor niet mogelijk om mensen van ouder dan 65 jaar in onze analyse te betrekken. We beperken ons tot kenmerken die we ook in de registerdata kunnen vinden omdat dit nodig is om de omvang van doelgroepen ook per gemeente en arbeidsmarktregio te kunnen vaststellen, zodat gemeenten deze informatie daadwerkelijk kunnen gebruiken in hun beleid.

De volgende acht objectieve kenmerken zijn meegenomen in de clusteranalyse:

- geslacht (2 categorieën: man of vrouw);
- leeftijd (3 categorieën: 16 tot 30 jaar; 30 tot 50 jaar; of 50 tot en met 65 jaar);
- migratieachtergrond (3 categorieën: Nederlandse achtergrond; 1e generatie migrant; of 2e generatie migrant; hierbij wordt zowel in de PIAAC-data als de CBS-data de CBS-definitie van de migratieachtergrond gebruikt: iemand met een migratieachtergrond is een persoon van wie ten minste één ouder in het buitenland is geboren. Er wordt onderscheid gemaakt tussen personen die zelf in het buitenland zijn geboren (de 1e generatie) en personen die in Nederland zijn geboren (de 2e generatie));
- partner en/of kinderen (4 categorieën: geen partner en geen kinderen; wel partner, geen kinderen; wel kinderen, geen partner; of zowel partner en kinderen);
- opleidingsniveau (2 categorieën: HBO of hogere opleiding versus alle andere opleidingen);
- arbeidsmarktstatus (3 categorieën: werkend; werkloos; of niet actief op de arbeidsmarkt; bij de laatste groep gaat het onder meer om mensen die arbeidsongeschikt zijn, niet werkzoekend zijn, een voltijds opleiding volgen, of met vervroegd pensioen zijn);
- bijstandsuitkering (2 categorieën: wel of geen bijstandsuitkering);
- schulden (2 categorieën: wel of niet deelnemend aan een schuldsaneringstraject, op basis van CBS-bestanden over de Wet Schuldsanering Natuurlijke Personen; informatie over schulden voor mensen die wel schulden hebben, maar niet deelnemen aan een schuldsaneringstraject is niet beschikbaar in de PIAAC-data of de CBS-registerdata).

Het is belangrijk om te benadrukken dat er bij een clusteranalyse altijd gezocht wordt naar clusters die enerzijds voldoende detail bevatten om inhoudelijk betekenisvol te zijn, maar die anderzijds ook statistisch daadwerkelijk van elkaar verschillen en omvangrijk genoeg zijn om op gemeenteniveau op een robuuste manier aantallen te kunnen schatten. We moeten daarom het aantal kenmerken beperkt houden en voor elk kenmerk focussen op een klein aantal categorieën die inhoudelijk het belangrijkst zijn. Zo zou voor arbeidsmarktstatus een verdere verfijning in categorieën mogelijk zijn (bijvoorbeeld door fulltime en parttime werkenden en verschillende niet-actieve groepen te onderscheiden, of door migranten uit verschillende herkomstlanden te onderscheiden), maar omdat dit de statistische kwaliteit van de clusters verslechtert zonder inhoudelijk veel op te leveren, zien we hier in deze analyse van af. Voor het opleidingsniveau geldt daarnaast dat een verder

onderscheid binnen de groep niet-hoogopgeleiden niet mogelijk is, omdat we op basis van registerdata alleen goed onderscheid kunnen maken tussen hoogopgeleiden en niet-hoogopgeleiden.

PIAAC is een internationaal survey in opdracht van de OESO dat in 39 landen uitgevoerd is. Ook Nederland heeft deelgenomen aan dit survey (zie o.a. Buisman et al., 2013 voor informatie over de Nederlandse PIAAC-data). Voor dit survey zijn ruim 5000 mensen tussen 16 en 65 jaar oud in Nederland thuis geïnterviewd. In de PIAAC-data worden kernvaardigheden van volwassenen gemeten, waaronder ook geletterdheid. Het gaat hierbij niet slechts om taalvaardigheid, maar om kernvaardigheden die van essentieel belang zijn voor het gebruiken, begrijpen en analyseren van informatie die we in het dagelijks leven en op het werk gebruiken (Buisman et al., 2013). Deze vaardigheden worden gemeten op een schaal van 0 tot 500. In lijn met eerdere rapporten over laaggeletterdheid op basis van de PIAAC-data (zie bijv. Bijlsma & Van der Velden, 2019) classificeren we een score van 225 of lager als laaggeletterd (volgens deze definitie is 11.9% van de bevolking tussen 16 en 65 jaar (1.8 miljoen mensen) laaggeletterd). Voor de clusteranalyse selecteren we daarom mensen die in de PIAAC-data een geletterdheidsscore van 225 of lager hebben. We hanteren in de clusteranalyse alleen de smalle definitie van geletterdheid (namelijk zonder laaggecijferdheid of lage digitale vaardigheden mee te rekenen).

We voeren twee aparte clusteranalyses uit, één voor de hoofdgroep met een geletterdheidsscore tussen 200 en 225 (deze groep noemen we vanaf hier 'mild laaggeletterd'), en één voor de groep met een geletterdheidsscore lager dan 200 (vanaf hier 'zeer laaggeletterd'). We maken het onderscheid tussen mild laaggeletterd en zeer laaggeletterd omdat we ervan uitgaan dat het voor mild laaggeletterden haalbaarder is om een hoger niveau van geletterdheid te bereiken dan voor mensen die zeer laaggeletterd zijn. We trekken de grens tussen mild laaggeletterd en zeer laaggeletterd bij een geletterdheidsscore van 200, omdat het verschil in een geletterdheidsscore van 200 en 225 overeenkomt met het verschil in geletterdheid tussen twee opleidingsniveaus (bijv. het verschil in de gemiddelde geletterdheidsscore tussen mensen met mbo en hbo bedraagt ook ongeveer 25 punten). Het streven naar een verbetering in geletterdheid die zo groot is dat mensen als het ware meer dan een volledig opleidingsniveau vooruitgaan in hun geletterdheid is weinig realistisch. In deze zin sluit dit onderscheid aan op discussies rond leerbaarheid; meer directe metingen om leerbaarheid van de laaggeletterden in deze doelgroepenanalyse vast te stellen zijn niet beschikbaar in de PIAAC-data en de CBS-registerdata. We berekenen daarnaast ook voor elk van de doelgroepen het percentage mild laaggeletterden en het percentage zeer laaggeletterden, om zo ook binnen de doelgroepen enig zicht te geven op verschillen in leerbaarheid. Het bleek helaas ook niet mogelijk om registerdata over lichte verstandelijke beperkingen als indicatie voor leerbaarheid in het onderzoek te betrekken. Het CBS beschikt over registerdata over lichte verstandelijke beperkingen, maar dit betreft alleen mensen die als licht verstandelijk beperkt geregistreerd zijn (in registraties over arbeidsongeschiktheidsuitkeringen, de wet sociale werkvoorziening en CIZ-indicaties voor langdurige zorg). Het CBS schat in dat hierdoor minder dan 10% van de mensen met een licht verstandelijke beperking vertegenwoordigd is in de registerdata, en dat deze groep erg selectief is. Dit zou ertoe leiden dat we in sommige doelgroepen het percentage met licht verstandelijke beperkingen zouden overschatten en in andere doelgroepen juist onderschatten.

Aanvankelijk was het plan om voor alle analyses in de Doelgroepenanalyse Laaggeletterdheid zowel de smalle definitie (nl. alleen geletterdheid) als de brede definitie van laaggeletterdheid te gebruiken (nl. geletterdheid, gecijferdheid én digitale vaardigheden). Voor de schattingen van het percentage laaggeletterden per gemeente en per arbeidsmarktregio (zie Hoofdstuk 5) is dit uiteindelijk ook zo gebeurd. Voor deze stap van de analyse gebruiken we schattingen op basis van zowel de brede definitie als de smalle definitie (geletterdheid). Laaggecijferdheid is hierbij op dezelfde manier gedefinieerd als laaggeletterdheid (een gecijferdheidsscore van 225 of lager in PIAAC). We spreken ook van lage digitale vaardigheden indien mensen zeggen nooit een computer te gebruiken, of indien ze een zeer basale test in het gebruik van een computer in PIAAC niet succesvol afgerond hebben (informatie over specifieke vaardigheden in het gebruik van bijvoorbeeld tablets en smartphones is in de PIAAC-data helaas niet voorhanden).
Voor het vaststellen van de doelgroepen is echter aan het begin van het project in overleg met de opdrachtgever besloten om slechts de smalle definitie van geletterdheid te gebruiken. Een eerste reden hiervoor is dat de doelgroepen zo beter aansluiten op cijfers uit eerdere onderzoeken rond laaggeletterdheid, die vooral betrekking hebben op geletterdheid in de smalle definitie. Verder wilden wij zoals hierboven al omschreven een onderscheid kunnen maken tussen een groep mild laaggeletterden en een groep zeer laaggeletterden, om het concept van leerbaarheid zoveel mogelijk mee te kunnen nemen in dit project. Het maken van een dergelijk onderscheid was op basis van de brede definitie moeilijk geweest, vanwege het verschil in de manier waarop geletterdheid, gecijferdheid en digitale vaardigheden precies gemeten zijn in de PIAAC-data. Tot slot moet ook bedacht worden dat geletterdheid, gecijferdheid en digitale vaardigheden sterk samenhangen, en dat het daarom aannemelijk is dat een analyse op basis van de brede definitie van laaggeletterdheid sterk vergelijkbare doelgroepen zou opleveren.

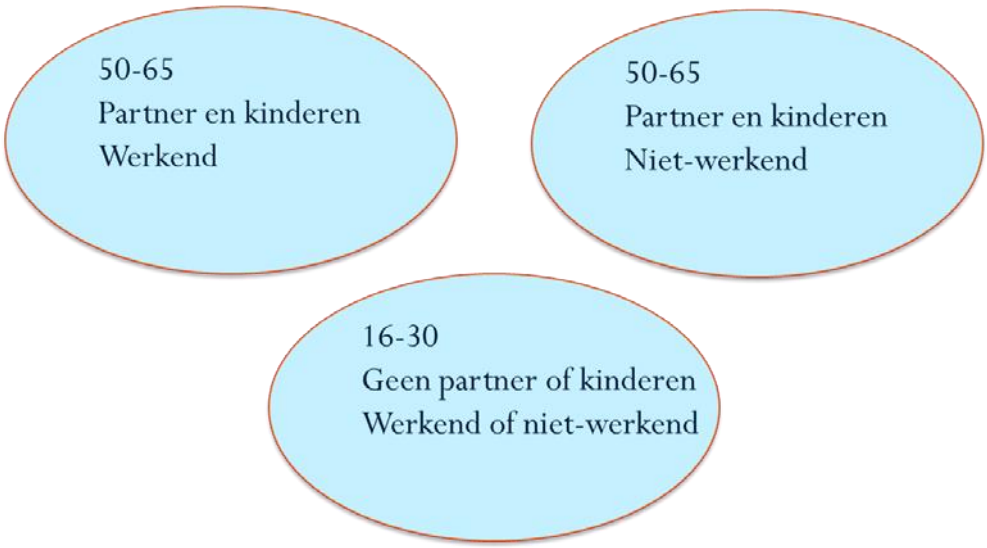
We voeren de clusteranalyse uit met het statistische softwarepakket SPSS 25. We doen de clusteranalyse apart voor de groep mild laaggeletterden en voor de groep zeer laaggeletterden, om de mogelijkheid open te laten dat we voor beide groepen verschillende clusters zouden kunnen vinden. We gebruiken een 'Two-step cluster analysis' (Norusis, 2007; Norusis, 2011) omdat deze vorm van clusteranalyse zowel geschikt is voor continue variabelen als voor categorische variabelen. Dit type clusteranalyse gebruikt de average silhouette-methode om de statistische kwaliteit van de clusteranalyse te bepalen. De average silhouette-waarde geeft daarbij aan hoe goed de clusters zich van elkaar onderscheiden, waarbij een hogere waarde duidt op een duidelijker onderscheid tussen de clusters. Een average silhouette-waarde van hoger dan 0.5 wordt als 'goed' beschouwd. Een te lage average silhouette-waarde geeft aan dat de clusters niet voldoende van elkaar verschillen. In de context van onze doelgroepenanalyse zou dit een probleem zijn: het zou betekenen dat de subgroepen zich niet echt van elkaar onderscheiden en dat een substantieel aantal laaggeletterden in meerdere subgroepen tegelijk kan vallen. Als teveel mensen in meerdere subgroepen tegelijk vallen, ondermijnt dit de betekenis van de subgroepen. Bijvoorbeeld: als drie subgroepen voor een aanzienlijk deel uit dezelfde mensen bestaan, kunnen we dan nog wel spreken van verschillende subgroepen? Daarom zijn de resultaten van de clusteranalyse voor ons alleen echt bruikbaar als onze clusters zich duidelijk van elkaar onderscheiden (zoals een average silhouette-waarde van hoger dan 0.5 zou suggereren). Naast de average silhouette-waarde voor de statistische kwaliteit van de clusters geeft de clusteranalyse ook voor elk kenmerk dat wordt meegenomen in de

clusteranalyse aan hoe belangrijk dit kenmerk is voor de clusters, met importance-scores op een schaal van 0 tot 1. Als een kenmerk een 1 scoort betekent dit dat dit kenmerk heel sterk onderscheidend is (met andere woorden, dit kenmerk bepaalt het sterkst tot welk cluster iemand behoort). Als een kenmerk een 0 scoort betekent dit dat dit kenmerk niet onderscheidend is (met andere woorden, dit kenmerk is helemaal niet van invloed op tot welk cluster iemand behoort).

We voeren de clusteranalyse eerst uit voor de groep mild laaggeletterden (we herhalen de analyse later voor de groep zeer laaggeletterden). In een eerste stap van de clusteranalyse zijn alle acht objectieve kenmerken meegenomen. Dit leverde een clusteranalyse op met een onvoldoende statistische kwaliteit (de average silhouette-waarde is niet hoger dan 0.5), wat erop wijst dat de gevonden clusters niet duidelijk genoeg van elkaar verschillen. De clusteranalyse laat verder zien dat vooral geslacht en het deelnemen aan schuldsanering niet belangrijk genoeg zijn (importance-scores van bijna 0). Dit suggereert dat geslacht en het deelnemen aan schuldsanering niet bepalen tot welk cluster mensen behoren; met andere woorden, deze kenmerken zijn niet onderscheidend bij het bepalen van de subgroepen.

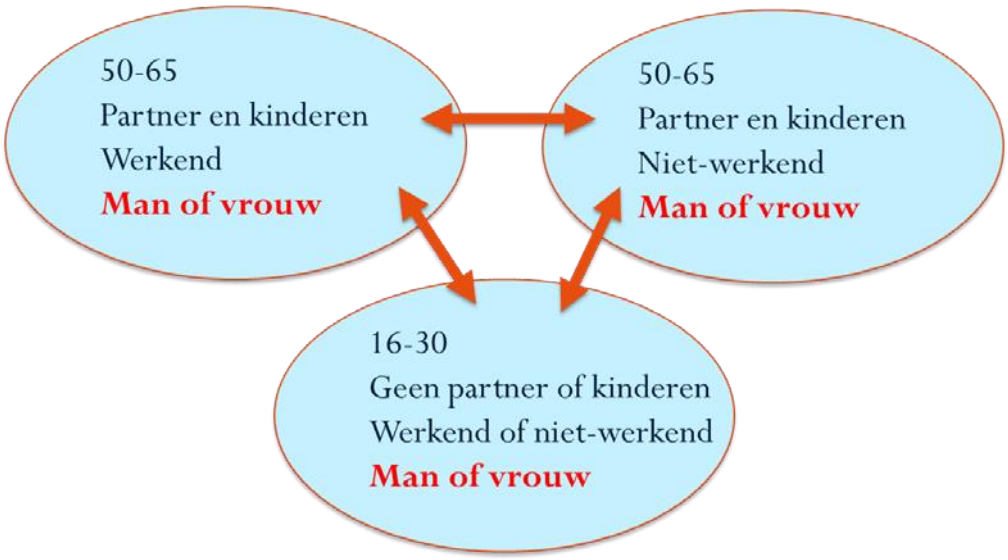
Een manier om de statistische kwaliteit van de clusteranalyse te verbeteren, en zo dus subgroepen te krijgen die wél duidelijk genoeg van elkaar verschillen, is het verwijderen uit de analyse van kenmerken die niet belangrijk genoeg zijn. Afbeeldingen 1, 2 en 3 hieronder illustreren aan de hand van een voorbeeld waarom dit nodig is. Stel dat we drie clusters hebben die zich duidelijk van elkaar onderscheiden, zoals weergegeven in Afbeelding 1. Twee clusters bevatten mensen uit dezelfde leeftijdsgroep (50 tot en met 65 jaar oud) die zowel een partner als kinderen hebben, maar het derde cluster bestaat uit mensen tussen 16 en 30 jaar oud die geen partner en kinderen hebben. Dit derde cluster bevat zowel werkenden als niet-werkenden, terwijl het eerste cluster alleen werkenden kent en het tweede cluster alleen niet-werkenden. In dit geval verschillen de clusters dan ook volledig van elkaar. Het is onmogelijk voor een individu om tot meer dan één van deze clusters tegelijk te behoren. Er is daarom geen overlap tussen de clusters.

Afbeelding 1: Voorbeeld van drie clusters die duidelijk van elkaar verschillen.



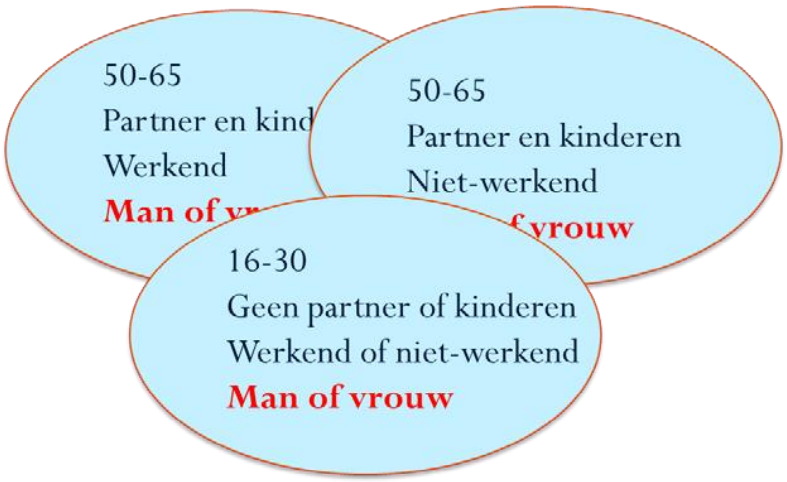
Wat gebeurt er nu als we een kenmerk toevoegen dat voor elk van de clusters hetzelfde is? Dit wordt geïllustreerd in Afbeelding 2. Stel dat we geslacht toevoegen aan de clusteranalyse en vinden dat zich in elk cluster zowel mannen als vrouwen bevinden (er is geen cluster dat (vrijwel) helemaal alleen uit vrouwen of alleen uit mannen bestaat; het kenmerk geslacht zou in dit geval dan ook weinig belangrijk zijn, en een lage importance-score hebben). Op dit punt onderscheiden de clusters zich dan ook niet duidelijk van elkaar. Het gevolg is dat de clusters meer op elkaar gaan lijken: de clusters worden hierdoor minder makkelijk van elkaar te onderscheiden.

Afbeelding 2: Na toevoeging van een kenmerk (geslacht) dat voor alle clusters hetzelfde is gaan de clusters meer op elkaar lijken.



Dit leidt ertoe (zoals geïllustreerd in Afbeelding 3) dat bij het schatten van de clusters de statistische kwaliteit lager wordt: de clusters gaan meer op elkaar lijken en verschillen daardoor minder duidelijk van elkaar. Dat is problematisch (op deze manier kunnen we immers geen duidelijke subgroepen vaststellen), maar ook onnodig. Omdat het voor de clusters blijkbaar toch niet uitmaakt of mensen man of vrouw zijn (geslacht is niet belangrijk) heeft het toch al geen toegevoegde waarde om geslacht mee te nemen in de clusteranalyse. De oplossing is dan ook om de clusteranalyse nog eens uit te voeren, maar dan zonder de kenmerken die niet belangrijk zijn, zodat deze kenmerken de statistische kwaliteit niet langer negatief kunnen beïnvloeden.

Afbeelding 3: Het gevolg van de toevoeging van geslacht aan de clusteranalyse is dat de clusters zich niet duidelijk meer van elkaar onderscheiden.



In de tweede stap van onze clusteranalyse verwijderen we daarom de kenmerken geslacht en het deelnemen aan een schuldsaneringstraject uit de analyse (als we de kenmerken één voor één verwijderen in plaats van allebei tegelijk leidt dit niet tot andere resultaten); in deze stap van de clusteranalyse worden dus alleen leeftijd, migratieachtergrond, partner en/of kinderen, opleidingsniveau, arbeidsmarktstatus en het hebben van een bijstandsuitkering meegenomen. Deze clusteranalyse is echter nog steeds van een onvoldoende statistische kwaliteit (de average silhouette-waarde is ook hier niet hoger dan 0.5). In deze stap blijken het hebben van een bijstandsuitkering en het opleidingsniveau niet belangrijk genoeg te zijn. Net als geslacht en het deelnemen aan een schuldsaneringstraject zijn deze kenmerken te weinig onderscheidend en bepalen ze vrijwel niet in welk cluster mensen terecht komen.

In een derde stap verwijderen we daarom ook het hebben van een bijstandsuitkering en het opleidingsniveau uit de analyse (ook hier hebben we ook bekeken wat er gebeurt als we de kenmerken één voor één verwijderen in plaats van allebei tegelijk, en ook nu leidt dit niet tot andere resultaten). Dit levert een clusteranalyse op die wél voldoende statistische kwaliteit heeft (de average silhouette-waarde is 0.6). De clusters die uit deze analyse naar voren komen zijn dus onderscheidend genoeg en verschillen voldoende van elkaar om echt afzonderlijke subgroepen te vormen. Geen van de vier overgebleven kenmerken (leeftijd, migratieachtergrond, partner en/of kinderen, en arbeidsmarktstatus) heeft een lage importance-score in deze analyse, wat betekent dat alle vier de kenmerken ertoe bijdragen dat de clusters duidelijk van elkaar verschillen. Als we vervolgens de vier kenmerken voor de zekerheid één voor één een keer verwijderen en de clusteranalyse herhalen, zien we dat dit de statistische kwaliteit niet nog verder verbetert. Er is dus geen reden om het aantal kenmerken in de clusteranalyse nog verder terug te brengen. We besluiten daarom dat we deze derde stap van de clusteranalyse als de definitieve clusteranalyse beschouwen.

Hoewel geslacht, opleidingsniveau, het hebben van een bijstandsuitkering en het deelnemen aan een schuldsaneringstraject dus niet meegenomen worden in de definitieve clusteranalyse betekent dit niet dat deze kenmerken niet van belang zijn in de laaggeletterdheidsproblematiek. We weten uit eerder onderzoek dat deze factoren gerelateerd zijn aan laaggeletterdheid. Daarom brengen we na het vaststellen van de clusters voor elk van de clusters in kaart hoe de man-vrouwverdeling eruit ziet, welke percentages deelnemen aan een schuldsaneringstraject en een bijstandsuitkering ontvangen, en welke opleidingsniveaus mensen binnen elk cluster bereikt hebben.

Tot slot hebben we de clusteranalyse herhaald voor de groep zeer laaggeletterden. Dit levert dezelfde clusters op, met één uitzondering: in de groep mild laaggeletterden vallen mensen van 30 jaar en ouder met een migratieachtergrond in één cluster, terwijl dit cluster voor de groep zeer laaggeletterden uiteenvalt in drie subclusters, waarbij de migratieachtergrond het voornaamste gemeenschappelijke kenmerk is (nl. oudere migranten met partner en kinderen; migranten tussen 30 en 50 jaar met partner en kinderen; en singles met migratie-achtergrond van 30 jaar en ouder). Omdat we ons in onze hoofdanalyse vooral richten op de groep mild laaggeletterden,en omdat het splitsen in meerdere subgroepen met een migratieachtergrond slechts voor een beperkt aantal gemeenten relevant is, hebben we besloten om bij het vaststellen van de doelgroepen vast te houden aan één cluster waarbij de migratieachtergrond het verbindende element is.

We brengen echter wel in kaart hoe dit cluster samengesteld is op het gebied van leeftijd, het hebben van een partner en kinderen, en de arbeidsmarktstatus. Het is strikt genomen mogelijk dat er zich onder de 1e generatie-groep in de data ook inburgeringsplichtige mensen bevinden. Het gaat daarbij echter om dusdanig kleine aantallen dat het wel of niet uitsluiten van deze groep in onze analyses voor onze schattingen geen noemenswaardig verschil zou maken (zo waren er volgens gegevens van het CBS op 1 september 2019 in totaal 18.552 inburgeringsplichtige personen in Nederland). Het is om deze reden ook moeilijk om het aantal laaggeletterde inburgeraars op gemeenteniveau in onze doelgroepen op een betrouwbare manier te schatten.

3 Typering van de doelgroepen op basis van de clusters

De definitieve clusteranalyse levert uiteindelijk 6 clusters op (en één restgroep van mensen die niet in een van de 6 clusters passen), die als basis gebruikt kunnen worden voor de typering van de doelgroepen.

De 6 clusters zijn als volgt te typeren:

A Oudere, werkend, met Nederlandse achtergrond, en zowel partner én kinderen

Mensen in deze groep zijn allemaal 50 tot en met 65 jaar oud, en hebben allemaal zowel een partner én kinderen. Uit eerder onderzoek (Bijlsma & Van der Velden, 2019) weten we dat werkende laaggeletterden relatief vaak werkzaam zijn in facility management en reiniging, vervaardiging van kleding, schoenen, metaal en andere goederen, landbouw, en horeca. Het gaat daarbij relatief vaak om mensen met een beroep als schoonmaker, productiemachinebediener, assemblagemedewerker of hulpkracht in bouw, industrie, horeca en landbouw.

B Oudere, niet-actief, met Nederlandse achtergrond, en zowel partner én kinderen

Identiek aan de eerste groep, maar niet actief op de arbeidsmarkt in plaats van werkend. Verdere analyse suggereert dat deze groep vooral bestaat uit mensen die met (vervroegd) pensioen of arbeidsongeschikt zijn. Op deze groep is mogelijk de 'use it or lose it'-redenering van toepassing (vaardigheden verslechteren indien ze niet gebruikt worden, zoals bijvoorbeeld leesvaardigheid in de werkomgeving).

C Tussen 30 en 50 jaar oud, werkend, met Nederlandse achtergrond, divers qua partners en kinderen

Deze groep onderscheidt zich niet als het gaat om partners en kinderen; mensen met en zonder partners en kinderen komen proportioneel voor in deze groep. Ook in dit cluster kan bijvoorbeeld gedacht worden aan werkenden in sectoren zoals facility management en reiniging, vervaardiging van kleding, schoenen, metaal en andere goederen, landbouw, en horeca, en om beroepen als schoonmaker, productiemachinebediener, assemblagemedewerker of hulpkracht in bouw, industrie, horeca en landbouw.

D Oudere, zowel werkend als niet-werkend, met Nederlandse achtergrond, en ófwel een partner, ófwel kinderen

De mensen in deze groep zijn vrijwel allemaal 50 tot en met 65 jaar oud, en zijn vooral singles met kinderen (een kleiner aandeel is kinderloos met partner). Mensen in deze groep zijn relatief iets vaker werkloos en niet-actief, maar de groep bevat ook werkenden.

E Jongere, zowel werkend als niet-werkend, zowel met Nederlandse achtergrond als met migratieachtergrond, en zonder partner en kinderen

Deze groep bestaat zowel uit jongeren (16 tot 30 jaar oud) met een Nederlandse achtergrond als jongeren met een migratie-achtergrond. Het gaat deels (47.6%) om jongeren die nog met een opleiding bezig zijn, bijvoorbeeld op mbo-niveau, al zijn er ook werkenden in deze groep.

F Mensen met migratieachtergrond, 30 jaar en ouder, zowel werkend als niet-werkend, en met partner, met kinderen of met beide

Mensen in deze groep hebben allen een migratieachtergrond (m.n. 1e generatie). Mensen in deze groep zijn allemaal 30 jaar en ouder (een combinatie van mensen tussen 30 en 50 jaar oud en mensen van 50 tot en met 65 jaar oud), en hebben meestal zowel een partner en kinderen (een kleinere groep heeft partner *of* kinderen). Dit cluster onderscheidt zich niet duidelijk van andere clusters op het gebied van arbeidsmarktstatus.

G Divers: mensen die niet in één van de 6 clusters (A t/m F) vallen

Tot slot blijft er een groep mensen over die enerzijds niet bij één van de 6 clusters passen, maar die anderzijds ook onderling te verschillend zijn om samen nog een extra zevende cluster te kunnen vormen. Hierbij gaat het bijvoorbeeld om 2e generatie migranten, jongeren met kinderen, 1e generatie migranten van 30 jaar en ouder die zowel geen partner en geen kinderen hebben, en oudere werknemers die zowel geen partner als geen kinderen hebben.

Het is hierbij belangrijk om op te merken dat de clusters niet overlappen en dat het dus niet mogelijk is voor individuen om zich in meer dan één cluster te bevinden.

Op 7 november 2019 hebben we de uitkomsten van de clusteranalyse voorgelegd aan de adviesgroep en de begeleidingscommissie. Er is daarbij besloten om de hierboven genoemde doelgroepen vast te stellen en te gebruiken als basis voor de verdere uitwerking van de profielen van de doelgroepen in de volgende stappen van het project.

Zoals al vermeld in Hoofdstuk 2 hebben we na het vaststellen van de doelgroepen op verzoek van de adviesgroep en de begeleidingscommissie ook voor elk van de doelgroepen laaggeletterden in kaart gebracht hoe de man-vrouwverdeling eruit ziet, welke percentages deelnemen aan een schuldsaneringstraject en een bijstandsuitkering ontvangen, welke opleidingsniveaus mensen bereikt hebben en hoe de percentages mild laaggeletterden en zeer laaggeletterden zich verhouden in de diverse doelgroepen. Voor cluster F brengen we daarnaast ook nog in kaart hoe dit cluster samengesteld is op het gebied van leeftijd, het hebben van een partner en kinderen, en de arbeidsmarktstatus. Deze aanvullende informatie wordt weergegeven in Bijlage 1.

4 Geschatte omvang van de doelgroepen op basis van de clusters

Op basis van de 6 clusters kunnen we vervolgens op landelijk niveau (voor de bevolking tussen 16 en 65 jaar, op basis van de PIAAC-data) een schatting maken van de omvang van de doelgroepen. We doen dat niet alleen voor de groep laaggeletterden, maar ook voor de bevolking als geheel. We kunnen op deze manier voor elke doelgroep vaststellen hoe groot de groep ongeveer is (ongeacht of mensen laaggeletterd zijn), en welk percentage in de doelgroep laaggeletterd is. We gaan hierbij (net als bij het vaststellen van de clusters) uit van de smalle definitie van laaggeletterdheid.

Om deze schattingen te kunnen maken, moet eerst voor elk van de doelgroepen worden bepaald welke kenmerken onderscheidend zijn. Bij doelgroep F hebben we bijvoorbeeld gezien dat het cluster zich niet duidelijk onderscheidt van andere clusters op het gebied van arbeidsmarktstatus. De verdeling van werkenden, werklozen en niet-actieven is in deze doelgroep vrijwel hetzelfde als in de totale groep laaggeletterden. Bij het schatten van de omvang van doelgroep F nemen we de arbeidsmarkstatus daarom niet mee (mensen in deze doelgroep kunnen zich immers in alle arbeidsmarktstatuscategorieën bevinden) en wordt de doelgroep gedefinieerd als de groep mensen die (a) tussen 30 en 50 jaar oud is, of 50 tot en met 65 jaar oud, (b) een partner, kinderen of beide heeft én (c) 1e generatie migrant is. In Tabel 1 is weergegeven aan welke kenmerken elke doelgroep voldoet. Kenmerken die niet onderscheidend zijn (zoals de arbeidsmarktstatus voor doelgroep F) worden hier als 'mix' omschreven.

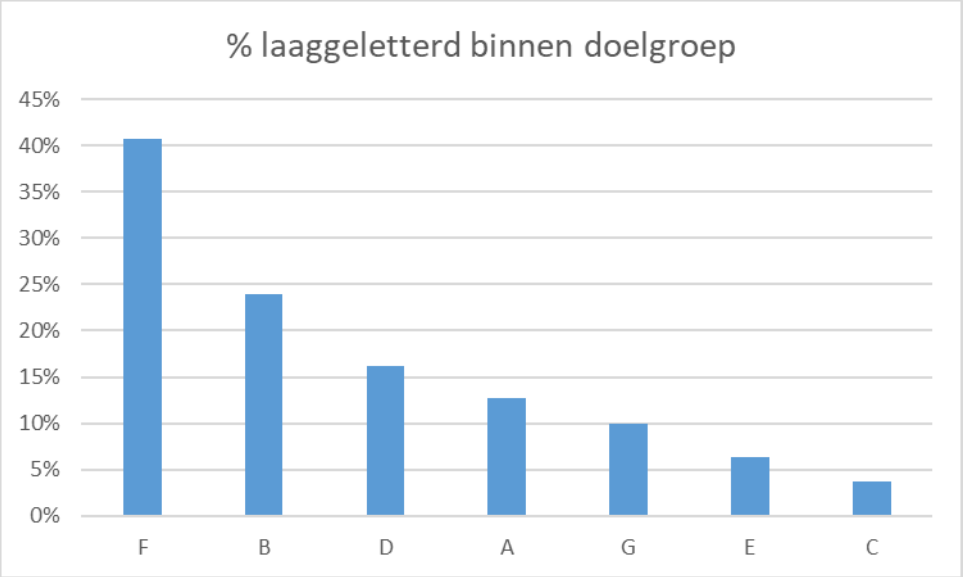
Tabel 1: Kenmerken van de doelgroepen

Doelgroep	Leeftijd	Partner en/of kinderen	Migratie-achtergrond	Arbeidsmarkt-status
A	50-65	Partner *en* kinderen	Niet-migrant	Werkend
B	50-65	Partner *en* kinderen	Niet-migrant	Niet-actief
C	30-50	Mix	Niet-migrant	Werkend
D	50-65	Partner *of* kinderen	Niet-migrant	Mix
E	16-30	Geen van beide	Mix	Mix
F	30-50 en 50-65	Partner *en/of* kinderen	1e generatie migrant	Mix
G	Mix	Mix	Mix	Mix

In Figuren 1, 2 en 3 wordt op verschillende manieren de geschatte omvang van de doelgroepen op basis van de clusters weergegeven. Hierbij definiëren we laaggeletterdheid opnieuw met een geletterdheidsscore van 225 of lager in de PIAAC-data. In Bijlage 2 tonen we daarnaast een tabel met meer precieze percentages voor de schattingen van de omvang van de doelgroepen, en tabellen met extra schattingen voor mild laaggeletterden (met een geletterdheidsscore tussen 200 en 225) en voor zeer laaggeletterden (met een geletterdheidsscore van lager dan 200).

Figuur 1 laat zien welk percentage van de bevolking binnen elke doelgroep laaggeletterd is. Dit dient ertoe om te kunnen zien in hoeverre laaggeletterdheid in elk van de doelgroepen veelvoorkomend is. We zien dat vooral doelgroep F een hoog percentage laaggeletterden kent (40.7%), gevolgd door doelgroepen B (23.9%), D (16.2%) en A (12.7%). In doelgroepen E en C ligt het percentage laaggeletterden erg laag (6.4% en 3.7%). Laaggeletterdheid komt dus relatief vaak voor binnen de groep mensen met een migratieachtergrond die 30 jaar en ouder zijn, terwijl het relatief erg weinig voorkomt binnen de groep werkenden tussen 30 en 50 jaar oud met een Nederlandse achtergrond.

Figuur 1: Percentage laaggeletterden binnen elk van de doelgroepen.

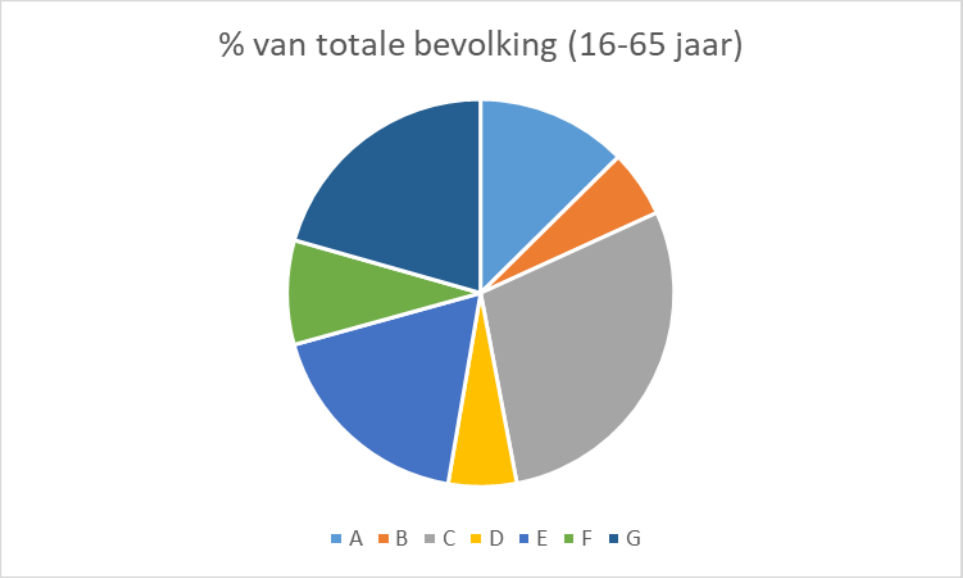


In Figuur 2 presenteren we welk percentage van alle laaggeletterden binnen elk van de doelgroepen valt. Dit laat zien hoe de totale groep laaggeletterden uiteenvalt in de verschillende doelgroepen. We zien dat 30.8% van alle laaggeletterden binnen doelgroep F valt en dat deze doelgroep daarmee de grootste groep vormt binnen de totale groep laaggeletterden. De andere doelgroepen verschillen niet heel sterk in grootte, met een aandeel in de totale groep laaggeletterden dat varieert van 8.1% (doelgroep D) tot 14.0% (doelgroep A). Verder blijkt in totaal 16.1% van de laaggeletterden in de 'Divers'-groep te vallen (groep G in de figuur). Dit betekent dat een overgrote meerderheid van 83.9% van de totale groep laaggeletterden wel te classificeren valt binnen één van de zes doelgroepen. Kortom, met onze doelgroepen ondervangen we grotendeels de totale groep laaggeletterden, en (op de grootste doelgroep met mensen met een migratieachtergrond die 30 jaar en ouder zijn na) ontlopen de doelgroepen elkaar maar weinig qua omvang.

Figuur 2: Percentage van alle laaggeletterden per doelgroep.

Tot slot laat Figuur 3 zien hoe de doelgroepen zich verhouden tot de totale bevolking tussen 16 en 65 jaar. Meer specifiek geeft dit weer welk aandeel de verschillende doelgroepen hebben in de bevolking als geheel. Dit kan nuttige informatie zijn om beter in beeld te krijgen om hoeveel mensen het bij elk van de doelgroepen gaat, en om de omvang van de doelgroepen zo iets meer context te geven. Als het percentage laaggeletterden in een doelgroep laag lijkt (zoals bijvoorbeeld geldt voor doelgroep C, zie Figuur 1) maakt het namelijk veel uit of het aantal mensen in deze doelgroep hoog of laag is. Uit Figuur 3 leiden we af dat deze extra analyse inderdaad belangrijk is. We zien namelijk dat de doelgroepen waar laaggeletterdheid in Figuur 1 nauwelijks een rol leek te spelen (doelgroepen C en E) wel een groot deel van de totale bevolking vormen (respectievelijk 28.8% en 18.0% van alle mensen tussen 16 en 65 jaar oud). Voor doelgroep F geldt juist het omgekeerde: waar het aandeel laaggeletterden binnen deze groep het hoogste is, is het aandeel van deze doelgroep in de totale bevolking juist klein (8.7% valt binnen deze doelgroep). Met andere woorden, waar sommige doelgroepen klein in omvang zijn, maar wel voor een aanzienlijk deel uit laaggeletterden bestaan (zoals dus de mensen met een migratieachtergrond van 30 jaar en ouder uit doelgroep F), zijn andere doelgroepen groter in omvang maar slechts voor een klein deel bestaand uit laaggeletterden (zoals de werkenden tussen 30 en 50 jaar met een Nederlandse achtergrond in doelgroep C).

Figuur 3: Percentage van de totale bevolking (16-65 jaar) per doelgroep



5 Schatting op gemeente- en arbeidsmarktregio-niveau

Na het vaststellen van de doelgroepen maken we schattingen op gemeente- en arbeidsmarktregio-niveau. Dit doen we in twee stappen: 1) de schatting van het totale percentage laaggeletterden per gemeente en arbeidsmarktregio; en 2) de schatting van de omvang van de verschillende doelgroepen laaggeletterden per gemeente en arbeidsmarktregio.

Om de schattingen te kunnen berekenen, combineren we de PIAAC-surveydata met CBS-registerdata. Het combineren van de PIAAC-data en de CBS-registerdata is noodzakelijk voor de schattingen op gemeente- en arbeidsmarktregio-niveau. We hebben immers data nodig die een meting van geletterdheid bevatten, die metingen van achtergrondkenmerken bevatten die met de meting van geletterdheid in verband gebracht kunnen worden en die de mogelijkheid bieden om deze metingen te gebruiken om schattingen op gemeenteniveau te berekenen. De PIAAC-data zijn niet alleen representatief voor de Nederlandse bevolking tussen 16 en 65 jaar oud, maar ook één van de weinige representatieve databronnen met een goede meting van geletterdheid. Daarbij komt dat een overgrote meerderheid van de respondenten in de PIAAC-data toestemming gegeven heeft om hun gegevens te koppelen aan CBS-registerdata. Dit maakt het mogelijk om voor de PIAAC-respondenten vast te stellen in welke gemeente ze wonen en om deze informatie te gebruiken om de meting van geletterdheid in PIAAC te vertalen in schattingen voor alle gemeenten in Nederland. Het is dus de specifieke combinatie van de CBS-registerdata van alle gemeenten en de PIAAC-data met de meting van geletterdheid, en het unieke feit dat deze twee bronnen gekoppeld kunnen worden, dat maakt dat we alleen op basis van deze bestanden voor alle gemeenten in Nederland op een vergelijkbare manieren schattingen kunnen berekenen over laaggeletterdheid.

1) *Schatting van het totale percentage laaggeletterden per gemeente en arbeidsmarktregio.*

In deze stap berekenen we schattingen van het totale percentage laaggeletterden per gemeente en arbeidsmarktregio voor zowel de smalle (alleen geletterdheid) als de brede definitie (geletterdheid, gecijferdheid én digitale vaardigheden) van laaggeletterdheid. Ook geven we voor de smalle definitie van geletterdheid aparte schattingen voor het percentage mild laaggeletterden (geletterdheidsscore tussen 200 en 225 in PIAAC) en voor zeer laaggeletterden (geletterdheidsscore van lager dan 200 in PIAAC).

De schatting gaat als volgt in zijn werk:
Eerst maken we gebruik van de techniek van 'kleinedomeinschatters'. De techniek van kleinedomeinschatters wordt door het Centraal Bureau voor de Statistiek (CBS) en andere statistische bureaus gebruikt om voor eenheden met weinig data (bijvoorbeeld een regio of een sector) tot meer robuuste schattingen te komen. Zo wordt bijvoorbeeld de werkloosheid per gemeente geschat op basis van deze techniek. Door deze methode kunnen we er bijvoorbeeld rekening mee houden dat de PIAAC-data slechts een beperkt aantal respondenten per gemeente bevat. Wij hebben deze techniek eerder al met succes toegepast om de gemiddelde geletterdheid en percentages laaggeletterden per gemeente,

beroep en sector te schatten (zie ook Bijlsma et al., 2016, en Bijlsma & Van der Velden, 2019 voor eerdere toepassingen van precies dezelfde methode). De methode maakt gebruik van het feit dat geletterdheid in de populatie sterk samenhangt met kenmerken als opleidingsniveau, leeftijd en arbeidsmarktstatus. Deze kenmerken kunnen vrij gedetailleerd per gemeente worden geschat op basis van CBS-registerdata.
Vervolgens wordt deze informatie gebruikt om een 'synthetische' schatting van de geletterdheid per gemeente te maken. Naast deze synthetische schatting wordt de schatting op basis van de PIAAC-data berekend (de 'directe' schatter). Met beide schattingen samen komen we tot een overall schatting van de geletterdheid. De einduitkomst is dus een gewogen som van 'directe' en 'synthetische' schattingen.

De op deze manier geschatte percentages laaggeletterden per gemeente en arbeidsmarktregio worden verwerkt in de online tool voor gemeenten. Op deze manier kunnen gemeenten snel informatie verkrijgen over het percentage laaggeletterden in hun gemeente of arbeidsmarktregio en vergelijken hoe zich dit verhoudt tot de percentages in andere gemeenten of arbeidsmarktregio's. Het is hierbij wel erg belangrijk om te benadrukken dat het gaat om schattingen en dat schattingen altijd gepaard gaan met een onzekerheidsmarge. We adviseren gemeenten dan ook om zich bij het gebruik van de tool niet te veel te laten leiden door de precieze geschatte percentages laaggeletterden en vooral te kijken naar hoe de gemeente zich verhoudt tot andere gemeenten in het geschatte percentage laaggeletterden.

2) *Schatting van de omvang van de verschillende doelgroepen laaggeletterden per gemeente en arbeidsmarktregio.*

We berekenen op basis van recente registerdata (2017) per gemeente de omvang van de doelgroepen per gemeente en arbeidsmarktregio. We schatten vervolgens op basis van de landelijke gegevens over het percentage laaggeletterden per doelgroep hoe groot voor elk van de doelgroepen het percentage laaggeletterden per gemeente en arbeidsmarktregio is. Zoals al besproken in Hoofdstuk 2 doen we dit voor deze stap alleen voor de smalle definitie van laaggeletterdheid (nl. geletterdheid).

Uit deze analyse volgt een duidelijk en landelijk goed vergelijkbaar overzicht van hoe groot de verschillende doelgroepen laaggeletterden zijn per gemeente en arbeidsmarktregio. Dit overzicht en de verdere uitwerking van de profielen van de doelgroepen laaggeletterden (bijv. de samenstelling van de doelgroepen en informatie over hoe en waar de doelgroepen bereikt kunnen worden) worden verwerkt in de online tool. Ook bieden we een handreikingsdocument aan om gemeenten verder te ondersteunen in het gebruik van de online tool. Met deze informatie kunnen gemeenten hun beleid op het gebied van laaggeletterdheid afstemmen en verbeteren.

6 Slotoverweging

Tot slot zetten we de belangrijkste overwegingen, randvoorwaarden en beperkingen die bij deze doelgroepenanalyse een rol gespeeld hebben nog eens kort op een rij.

- Het doel van de clusteranalyse is om een beperkt en hanteerbaar aantal subgroepen te identificeren die 'bruikbaar' zijn voor het beleid.
- Dit houdt in dat we subgroepen identificeren die landelijk goed vergelijkbaar zijn en in principe te gebruiken zijn door alle gemeenten.
- Dit betekent dat we voor de clusteranalyse en subgroepen alleen kenmerken kunnen gebruiken die in zowel de PIAAC-data als CBS-registerdata voorhanden zijn.
- In deze analyse beperken we ons tot de Nederlandse bevolking tussen 16 en 65 jaar oud, omdat de PIAAC-data alleen op deze leeftijdsgroep betrekking hebben.
- Het gebruik van clusteranalyse heeft verder tot gevolg dat er een balans moet zijn tussen voldoende inhoudelijke betekenis, statistische kwaliteit van de clusteroplossing, en clusters die omvangrijk genoeg zijn.
- Dit leidt ertoe dat het aantal kenmerken dat we gebruiken voor het typeren van de doelgroepen beperkt moet blijven, en dat we voor elk kenmerk moeten focussen op een klein aantal categorieën die inhoudelijk het belangrijkst zijn.
- Na het vaststellen van de doelgroepen op basis van de clusteranalyse zijn de profielen van de subgroepen nog verder aangevuld met vindplaatsen, motieven en kenmerken van passend aanbod. Daarnaast zijn de profielen van de subgroepen nog aangevuld met informatie over de man/vrouwverdeling, het opleidingsniveau, het hebben van een bijstandsuitkering, het deelnemen aan een schuldsaneringstraject en het niveau van laaggeletterdheid. Op die manier krijgen gemeenten voor elke subgroep een beter overzicht van de samenstelling van de doelgroepen en van mogelijke vindplaatsen, motieven/behoeften van subgroepen en kenmerken van aanbod en mogelijke type aanbieders hiervan. Dit helpt om de bruikbaarheid van de doelgroepen voor het beleid verder te vergroten.
- Het is in principe mogelijk voor gemeenten om hierin met eigen data nog meer diepgang en detail te zoeken, maar dit reikt buiten de landelijke focus van onze analyse.

Bronnen

Bijlsma, I., van den Brakel, J., van der Velden, R. K. W., & Allen, J. P. (2016). *Regionale spreiding van geletterdheid in Nederland*. ROA External Reports.

Bijlsma, I. & van der Velden, R. (2019). S*preiding van laaggeletterdheid: Inzicht in taal- en rekenvaardigheden per beroep, sector en type werkzoekende.* (ROA External Reports). Den Haag: Stichting Lezen & Schrijven.

Buisman, M., Allen, J., Fouarge, D., Houtkoop, W., & van der Velden, R. (2013). PIAAC: *Kernvaardigheden voor werk en leven: Resultaten van de Nederlandse survey 2012.* Den Bosch: Ecbo.

Norusis, M. J. (2007). *SPSS 15.0 advanced statistical procedures companion.* Chicago, IL: Prentice Hall.

Norusis, M.J. (2011). *IBM SPSS Statistics 19 Procedures Companion.* Reading, MA: Addison Wesley.

Stichting Lezen en Schrijven (2018). *Feiten en cijfers laageletterdheid.* Den Haag: Stichting Lezen & Schrijven.
https://www.lezenenschrijven.nl/uploads/editor/2018_SLS_Literatuurstudie_FeitenCijfers_interactief_DEF.pdf

Bijlage 1 – Aanvullende kenmerken van de doelgroepen laaggeletterden

N.B. De onderstaande kenmerken zijn berekend op basis van de totale groep laaggeletterden (nl. iedereen met een geletterdheidsscore van <225) in de PIAAC-data). Dit betekent ook dat de hieronder gepresenteerde cijfers alleen betrekking hebben op laaggeletterden. We hanteren hiervoor wederom de smalle definitie van laaggeletterdheid (nl. alleen geletterdheid, en niet gecijferdheid of digitale vaardigheden). Voor de hoogst voltooide opleiding kunnen we hier vier opleidingsniveaus onderscheiden (en niet slechts twee zoals bij het uitvoeren van de clusteranalyse). Dit komt doordat we bij de berekening van deze aanvullende kenmerken geen gebruik gemaakt hebben van de registerdata, waarin dit soort meer gedetailleerde informatie over het opleidingsniveau niet beschikbaar is.

Tabel A1. Man/vrouwverdeling per doelgroep

Doelgroep	% man	% vrouw
A	43.7%	56.3%
B	33.3%	66.7%
C	55.5%	44.5%
D	43.9%	56.1%
E	53.4%	46.6%
F	46.3%	53.7%
G	56.2%	43.8%

Tabel A2. Hoogst voltooide opleiding per doelgroep

Doelgroep	% basisonderwijs	% vmbo	% mbo	% havo/vwo, hbo of wo
A	14.9%	47.5%	27.4%	10.2%
B	29.4%	52.9%	15.3%	2.4%
C	22.6%	38.1%	38.0%	1.3%
D	29.0%	40.6%	25.8%	4.6%
E	49.0%	23.9%	24.7%	2.4%
F	44.8%	18.1%	19.5%	17.6%
G	30.8%	39.0%	21.3%	8.9%

Tabel A3. Verdeling mild laaggeletterd (literacy score 200-225) en zeer laaggeletterd (literacy score <200) per doelgroep

Doelgroep	% mild laaggeletterd	% zeer laaggeletterd
A	62.2%	37.8%
B	59.3%	40.7%
C	73.1%	26.9%
D	57.0%	43.0%
E	51.9%	48.1%
F	39.9%	60.1%
G	41.9%	58.1%

Tabel A4. Percentage mensen dat deelnam aan schuldsaneringstraject tussen 2007 en 2011 en percentage mensen met bijstandsuitkering in de afgelopen maand, per doelgroep

Doelgroep	% in schuldsanering	% met bijstands-uitkering
A	0.0%	0.0%
B	0.0%	4.7%
C	1.7%	3.9%
D	0.0%	5.0%
E	0.0%	2.9%
F	4.0%	16.9%
G	4.0%	19.5%

Tot slot vermelden we nog aanvullende informatie die specifiek relevant is voor doelgroep F:

- Doelgroep F bestaat voor 58.9% uit mensen tussen 30-50 jaar en voor 41.1% uit mensen tussen 50-65 jaar.
- Doelgroep F bestaat voor 8.1% uit mensen met partner maar zonder kinderen, voor 24.1% uit mensen met kinderen maar zonder partner en voor 67.8% uit mensen met zowel een partner en kinderen.
- Doelgroep F bestaat voor 53.4% uit werkenden, voor 5.3% uit mensen die werkloos zijn en voor 41.3% uit mensen die niet-actief zijn. Bij de laatste groep gaat het onder meer om mensen die arbeidsongeschikt zijn, niet werkzoekend zijn, een voltijds opleiding volgen, of met vervroegd pensioen zijn.

Bijlage 2 – Tabellen met geschatte omvang van de doelgroepen

Tabel B1: Omvang van de doelgroepen (percentage binnen elke doelgroep dat laaggeletterd is, aandeel van elke doelgroep binnen alle laaggeletterden en aandeel van elke doelgroep binnen de totale bevolking tussen 16 en 65 jaar)

Doelgroep	% laaggeletterd binnen doelgroep	% van alle laaggeletterden	% van totale bevolking (16-65 jaar)
F	40.7%	30.8%	8.7%
B	23.9%	11.7%	5.6%
D	16.2%	8.1%	5.7%
A	12.7%	14.0%	12.6%
G	10.0%	16.1%	20.6%
E	6.4%	10.0%	18.0%
C	3.7%	9.3%	28.8%

Bron: PIAAC (2011-2012).

Tabel B2: Omvang van de doelgroepen (percentage binnen elke doelgroep dat laaggeletterd is en aandeel van elke doelgroep binnen alle laaggeletterden) – alleen voor mild laaggeletterden (laaggeletterdheidsscore 200-225)

Doelgroep	% laaggeletterd binnen doelgroep	% van alle laaggeletterden
F	16.3%	24.0%
B	14.2%	13.6%
D	9.3%	9.1%
A	7.9%	16.9%
G	4.2%	13.1%
E	3.3%	10.2%
C	2.7%	13.3%

Bron: PIAAC (2011-2012).

Tabel B3: Omvang van de doelgroepen (percentage binnen elke doelgroep dat laaggeletterd is en aandeel van elke doelgroep binnen alle laaggeletterden) – alleen voor zeer laaggeletterden (laaggeletterdheidsscore <200)

Doelgroep	% laaggeletterd binnen doelgroep	% van alle laaggeletterden
F	24.5%	38.0%
B	9.7%	12.3%
D	7.0%	9.8%
G	5.8%	19.1%
A	4.8%	10.9%
E	3.1%	7.2%
C	1.0%	5.2%

Bron: PIAAC (2011-2012).