



Basisvaardigheden in zicht

Centraal Bureau voor de Statistiek

Team Beleidsstatistiek

24 oktober 2025

Introductie

Het hebben van beperkte basisvaardigheden houdt in dat mensen moeite hebben met lezen, schrijven en/of rekenen. Dat bemoeilijkt het zelfstandig functioneren in de maatschappij, het realiseren van eigen doelen en de zelfontplooiing.¹

De arbeidsmarkt doet een sterk beroep op deze vaardigheden. Werkgevers stellen hoge eisen en door toenemende digitalisering en automatisering lopen mensen bij wie deze vaardigheden beperkt zijn steeds meer risico op (langdurige) werkloosheid en uitsluiting van de arbeidsmarkt.

In meer dan 30 landen worden eens in de tien jaar basisvaardigheden gemeten in het zogeheten *Programme for the International Assessment of Adult Competencies* (PIAAC). Dit wordt internationaal gecoördineerd door de OESO en in Nederland uitgevoerd door het Kohnstamm Instituut. In dit internationale onderzoek worden de competenties van volwassenen gemeten, met als doel de vaardigheden op het gebied van lezen, rekenen en het oplossen van problemen in technologierijke omgevingen van de bevolking te beoordelen. In het onderzoek wordt dit gemeten aan de hand van een taaltest, een rekentest en een test rondom het probleemoplossend vermogen. Op basis van de scores van de taal- en rekentest wordt bepaald of iemand in minimaal één van de categorieën moeite heeft met deze basisvaardigheden.

In december 2024 zijn data van een nieuwe editie van het PIAAC-onderzoek beschikbaar gekomen.² Deze publicatie is op nationaal niveau en vergelijkt Nederland met andere landen. Het ministerie van OCW heeft aan het CBS gevraagd verdiepende analyses te doen op basis van deze data en daarmee een schatting te maken van het aandeel mensen met beperkte basisvaardigheden op het regionale schaalniveau niveau van arbeidsmarktregio, gemeente en waar mogelijk wijken.

Het CBS heeft de PIAAC-data op persoonsniveau gekoppeld aan registerdata die bij het CBS over deze personen aanwezig zijn. Middels de methode kleine domeinschatters (KDS) is vervolgens een schatting gemaakt van het aandeel mensen met beperkte basisvaardigheden per arbeidsmarktregio, gemeente en waar mogelijk wijken.

Deze analyse is eerder uitgevoerd voor het PIAAC-onderzoek uit 2012. Daar is toen het dashboard Geletterdheid in Zicht uit voortgekomen, waarvan het dashboard Basisvaardigheden

¹ Expertisepunt Basisvaardigheden, <https://basisvaardigheden.nl/themas/thema-basis-over-basisvaardigheden>.

² <https://piaaonderzoek.nl>.

in Zicht de opvolger is. De huidige analyse op de PIAAC-data wijkt op een aantal onderdelen af van de methodiek van de eerdere analyse. Zo wordt er nu gekeken naar lage vaardigheden in taal en/of rekenen, waar voorheen alleen naar taalvaardigheden werd gekeken. Verder worden er nu registerdata gebruikt om aan te koppelen, waar dat eerder enquêtedata waren. Ook zijn er in het model andere keuzes gemaakt om beter om te gaan met steekproefafwijkingen, waardoor op meer gedetailleerd regioniveau schattingen kunnen worden gedaan. Ten slotte beslaan de PIAAC-data nu personen van 16 tot en met 75 jaar en was dit in de eerdere editie 16 tot en met 65 jaar. Door deze verschillen zijn de cijfers niet één-op-één te vergelijken met de resultaten van het voorgaande onderzoek.

Daarnaast is ervoor gekozen om voor specifieke doelgroepen te belichten op welke manier zij onderdeel zijn van de laaggeletterdheidspopulatie, door middel van een kruising tussen de kans dat iemand in de groep lage basisvaardigheden bevat en de grootte van de groep binnen de regio. Deze relatieve verdeling wordt getoond in een aparte tabel.

Beide tabellen zijn als volgt opgebouwd: Er zijn vier verschillende regionale niveaus: nationaal (land), arbeidsmarktregio, gemeente en wijk. Per gebied is de regiocodering en de regionaam gegeven; bij wijken staat ook de bijbehorende gemeentenaam genoemd in de vorm van “gemeentenaam: wijknaam”. De volgende drie kolommen gaan over de groepsindelingen van de doelgroepen: één kolom maakt een uitsplitsing naar NT1 en NT2³, een tweede kolom gaat over doelgroepen in de populatie zoals uitsplitsingen naar leeftijd of werkstatus.. Als laatste volgen de resultaten met het geschatte aandeel laaggeletterden en deze uitgedrukt in aantallen personen (afgerond op 100-tallen). Daarbij gaat het om de relevante populatie van 16- tot 75-jarigen.

Technische toelichting

Populatie

De steekproefpopulatie uit PIAAC bestaat uit alle respondenten die officieel onderdeel zijn van het PIAAC-onderzoek. Dit zijn mensen die tussen september 2022 en augustus 2023 vallen in de leeftijdsgroep van 16 tot 75 jaar en waarvan de wijk kon worden afgeleid.

De populatie uit de bevolkingsregisters bestaat uit degenen die op 31 december 2022 woonachtig waren in Nederland, waarvan de wijk kon worden afgeleid en die op dat moment een leeftijd hadden tussen de 16 en 75 jaar. Deze selectie is genomen om zo goed mogelijk aan te sluiten bij de PIAAC-populatie.

Methode

In dit onderzoek worden schattingen gemaakt van de aanwezigheid van beperkte basisvaardigheden per arbeidsmarktregio, gemeente en wijk. Dat gebeurt op basis van het landelijke PIAAC-onderzoek. Landelijk zijn er ongeveer 4250 waarnemingen. Zou je alleen de respondenten uit PIAAC gebruiken om tot een gewogen gemiddelde te komen voor een

³ Een uitleg over wat NT1 en NT2 inhouden staat onder het kopje *Doelgroepen* in de paragraaf *Technische toelichting*.

arbeidsmarktregio, gemeente of wijk, dan zou de onzekerheid van dit gemiddelde voor de meeste regio's te groot zijn om tot bruikbare cijfers te leiden. Dit komt simpelweg doordat er in de meeste regio's te weinig waarnemingen zijn om een goede schatting te maken van de aanwezigheid van beperkte basisvaardigheden.

In dit onderzoek wordt daarom gebruik gemaakt van de methode kleine domeinschatters (KDS). Dit is een statistische methode om bij beperkte data schattingen te maken die je met alleen enquêtedata niet zou kunnen maken. Je maakt daarbij gebruik van verbanden die in de enquêtedata aanwezig zijn.

In het geval van vaardigheden kunnen we kijken naar persoonskenmerken die sterk samenhangen met de aanwezigheid beperkte basisvaardigheden. Deze hebben een hoge voorspelkracht: ofwel omdat ze sterk samenhangen met het hebben van basisvaardigheden (zoals het hoogst behaalde opleidingsniveau), ofwel omdat ze samenhangen met het verlies ervan (zoals een hogere leeftijd). Deze verbanden zijn niet afhankelijk van een regio en gelden voor heel Nederland. Maar er zijn uiteraard wel verschillen tussen regio's qua kenmerken van de personen die er wonen. In sommige regio's woont bijvoorbeeld een groter aandeel oudere personen dan in andere regio's.

Deze achtergrondkenmerken zijn voor alle personen in Nederland in de registerdata bekend. Hierdoor kunnen we een modelmatige schatting maken van het verwachte percentage personen met beperkte basisvaardigheden. Vervolgens kunnen we de directe schattingen op regioniveau (waarin enkel de PIAAC-deelnemers zijn meegenomen) bijsturen om zo de onzekerheid per schatting te verminderen. Dit bijsturen gebeurt door het gewogen gemiddelde te nemen van de directe schatting en de modelmatige schatting. De weging varieert per regio. Hoe onzekerder de directe schatting is, hoe meer gewicht de modelmatige schatting zal krijgen.

Door het bijsturen en de onzekerheid te reduceren kunnen we toch iets zeggen over de mate van beperkte vaardigheden voor de meeste regio's.

Risico's en interpretatie

Het KDS-model kent een paar risico's. Zo heeft het model de neiging om extreme waardes te onderschatten en dichter bij het gemiddelde te gaan zitten. Ook wordt er aangenomen dat de verbanden tussen persoonskenmerken en de aanwezigheid van basisvaardigheden geen regionaal karakter hebben. Dat wil zeggen dat er in de modelmatige schatting vanuit wordt gegaan dat bijvoorbeeld een persoon in Heerlen en een persoon in Groningen met identieke kenmerken dezelfde kans hebben om beperkte basisvaardigheden te hebben. Dit kan tot verkeerde schattingen leiden als er onvoldoende rekening wordt gehouden met verbanden binnen een regio.

Bij de interpretatie van de cijfers is belangrijk om te beseffen dat het schattingen zijn. Daarom hoeft een verschil van een paar procentpunten niet te betekenen dat er in realiteit tussen deze regio's een verschil is in basisvaardigheden.

De cijfers zeggen niets over het beleid dat door een gemeente gevoerd wordt. Verschillen tussen regio's worden veroorzaakt door demografische verschillen. Dat betekent dat in een regio met een hoog percentage de kans op beperkte basisvaardigheden meer aanwezig is dan in andere regio's, maar dat wil niet zeggen dat de basisvaardigheden in die regio ook daadwerkelijk beperkt zijn. Als bijvoorbeeld in een wijk veel aandacht besteed wordt aan basisvaardigheden en dit een positief effect heeft, zal het cijfer dat wij presenteren voor deze wijk een overschatting van de beperkte basisvaardigheden zijn, omdat wij de effecten van dit beleid niet meenemen.

Variabelen en model

Deze paragraaf gaat dieper in op de variabelen en het model en geeft een verantwoording van de methode. Hiervoor is enige statistische kennis vereist.

De kenmerken die zijn meegenomen in de KDS-analyse wijken iets af van de kenmerken die in voorgaande edities van het onderzoek zijn meegenomen. De reden hiervoor is dat in vorige edities gebruik werd gemaakt van de Enquête Beroepsbevolking (EBB) om achtergrondkenmerken per regio te bepalen. Nu is ervoor gekozen om gebruik te maken van registerdata. In tegenstelling tot de data uit de EBB, hebben we bij registerdata de beschikking over data van alle personen. Daarom zijn deze betrouwbaarder. Daaruit zijn de volgende kenmerken uiteindelijk gebruikt als achtergrondkenmerk in het model:

- leeftijd (naar groepen van 10 jaar)
- hoogst behaalde opleidingsniveau⁴
- Sociaaleconomische status zijnde in opleiding
- Inkomen (in 10 decielgroepen; hiervoor zijn de waarnemingen geordend van laag naar hoog en verdeeld in 10 categorieën met elk even veel waarnemingen)
- Migratieachtergrond (in Nederland geboren, migrant Europees, migrant niet-Europees, kind van migrant Europees, kind van migrant niet-Europees)
- Leeftijd waarop de persoon in Nederland is gekomen, in het geval het een migrant is (in klassen 0 tot 6 jaar, 6 tot 12 jaar, 12 tot 18 jaar, 18 tot 27 jaar, 27 jaar en ouder)
- Aantal jaren in Nederland (in klassen van 3 jaar tot maximaal 15 jaar)

Ook wordt een random effect per arbeidsmarktregio, gemeente en wijk aan het model toegevoegd om eventuele afwijkingen op te vangen waar de kenmerken in het model geen rekening mee houden.

Dit model schat uiteindelijk de voorspellingen op wijkniveau. Daarbij wordt gebruik gemaakt van de directe resultaten, de verbanden van het model zoals ze hierboven zijn geschat, en de registerdata op regionaal niveau.⁵

⁴ Opgedeeld in de volgende categorieën: 1) basisonderwijs, 2) vmbo, havo-, vwo-onderbouw, mbo1, 3) havo, vwo, mbo, 4) hbo-, wo-bachelor, 5) hbo-, wo-master, doctor, 6) onbekend.

⁵ Via *mcmcscsae*: <https://cran.r-project.org/web/packages/mcmcscsae/index.html>.

De modelvoorspellingen op wijkniveau worden vervolgens samengenomen naar het niveau van gemeente en arbeidsmarktregio. Deze tellen samen op tot het landelijk gemiddelde.

Doelgroepen

Op het landelijke niveau en op het niveau van arbeidsmarktregio's en gemeenten worden ook uitsplitsingen gemaakt naar zogenaamde doelgroepen. Het idee daarachter is dat voor beleidsmakers niet alleen de uiteindelijke uitkomst belangrijk is, maar ook welke subpopulatie de meest gevoelige positie inneemt en welke groepen eventueel makkelijker te bereiken zijn dan anderen. Daarbij is gekozen voor de volgende uitsplitsing:

- Onder NT1 wordt hier verstaan personen die in Nederland geboren zijn en kinderen van migranten; onder NT2 wordt verstaan migranten.⁶
- De doelgroepen zijn opgesplitst naar leeftijd (16-27, 28-65, 66-75); werkstatus (Werkend, in opleiding, anders, getypeerd naar de sociaaleconomische status van een persoon), type huishouden (paar met/zonder thuiswonend kind) en de combinatie van de laatste twee met de leeftijdsdoelgroep.

Deze cijfers zijn berekend door het geschatte aandeel personen met lage basisvaardigheden binnen de betreffende doelgroep te vermenigvuldigen met het aandeel van deze doelgroep binnen de regio. Het idee van deze extrapolatie is dat het meer duiding geeft voor beleidsmakers wáár in de regio zich de problemen bevinden, en hoe groot deze groepen ruwweg zullen zijn. Deze doelgroepen zijn niet alleen gekozen omdat ze een verband hebben met de verwachte basisvaardigheden, maar ook met de potentiële bereikbaarheid van deze groepen.

Bescherming van persoonsgegevens

Door gebruik te maken van de KDS-methodiek worden schattingen altijd modelmatig bijgevuld, zodat er geen directe persoonsgegevens afleidbaar zijn uit de resultaten. Voor alle arbeidsmarktregio's en gemeenten kunnen schattingen worden gemaakt. Maar voor wijken moeten bijna alle schattingen modelmatig gebeuren. Daarom moet de wijk wel voldoende groot zijn om een betrouwbare schatting te kunnen doen. Voor wijken met minder dan 7 500 inwoners is daarom geen schatting gemaakt.

Privacy

Privacy is een groot goed. Ook als je niks te verbergen hebt, heb je heel wat te beschermen. Het CBS is het Statistisch Bureau van Nederland dat onafhankelijk onderzoek uitvoert. Het CBS werkt bij elk onderzoek met strenge eisen om data op een veilige manier te verwerven, te verwerken en te publiceren en is transparant over de manier van werken en de methodieken.

Het CBS verzamelt gegevens van natuurlijke personen, bedrijven en instellingen. Dit is wettelijk vastgelegd in de CBS-wet en de Algemene Verordening Gegevensbescherming (AVG).

⁶ NT1 en NT2 staat voor respectievelijk Nederlands als eerste taal en Nederlands als tweede taal.

Identificerende persoonskenmerken worden na ontvangst direct gepseudonimiseerd. Hierdoor kan het onderzoek alleen worden uitgevoerd op gegevens met een pseudosleutel. Bij publicatie zorgt het CBS er bovendien voor dat natuurlijke personen of bedrijven niet herkenbaar of herleidbaar zijn. Ook hanteert het CBS diverse maatregelen tegen diefstal, verlies of misbruik van persoonsgegevens. Het CBS levert geen herkenbare gegevens aan derden, ook niet aan andere overheidsinstellingen. Wel kunnen sommige (wetenschappelijke) instellingen onder strenge voorwaarden toegang krijgen tot gegevens met pseudosleutel op persoons- of bedrijfsniveau. Dit noemen we microdata. Voor meer informatie, zie onze website: www.cbs.nl/privacy.