

CBS Microdata als goudmijn voor sociaal-wetenschappelijk onderzoek

CBS Microdata: wat houdt dit in en wat gaan wij onderzoeken?

Het Centraal Bureau voor de Statistiek (CBS) kent sinds 1999 het Stelsel van Sociaal-Statistische Bestanden (CBS, 2025a). Dit is een verzameling onderling koppelbare gegevensbestanden op persoons-, bedrijfs- en adresniveau. Deze databestanden zijn afkomstig uit registratiesystemen van overheidsinstellingen, zoals gemeenten en de Belastingdienst. Veel bestanden bestrijken de voltallige populatie van Nederland en ze kunnen doorgaans gekoppeld worden aan steekproeven, zoals de Enquête Beroepsbevolking of het Woononderzoek Nederland.

Sinds 2004 kunnen ook onderzoekers die niet aan het CBS verbonden zijn, tegen betaling en onder strikte voorwaarden, deze databestanden gebruiken. De gepseudonimiseerde gegevens blijven fysiek op het CBS en worden online in een CBS werkomgeving geanalyseerd. Naar deze gegevens wordt verwezen als de CBS Microdata. Hiervan hebben inmiddels veel organisaties gebruikgemaakt. Aanvankelijk kon dat alleen op terminals die rechtstreeks op het CBS-systeem waren aangesloten, maar tegenwoordig kan dat op vrijwel elke laptop die verbonden is met het internet.

Het bestaan van deze faciliteit is best opmerkelijk, want behalve in Zweden, Noorwegen, Finland en Denemarken zijn er geen andere landen in de wereld waar onderzoekers toegang kunnen krijgen tot microdata op populatieniveau. De faciliteit biedt Nederlandse onderzoekers dus mogelijkheden die weinig onderzoekers elders tot hun beschikking hebben.

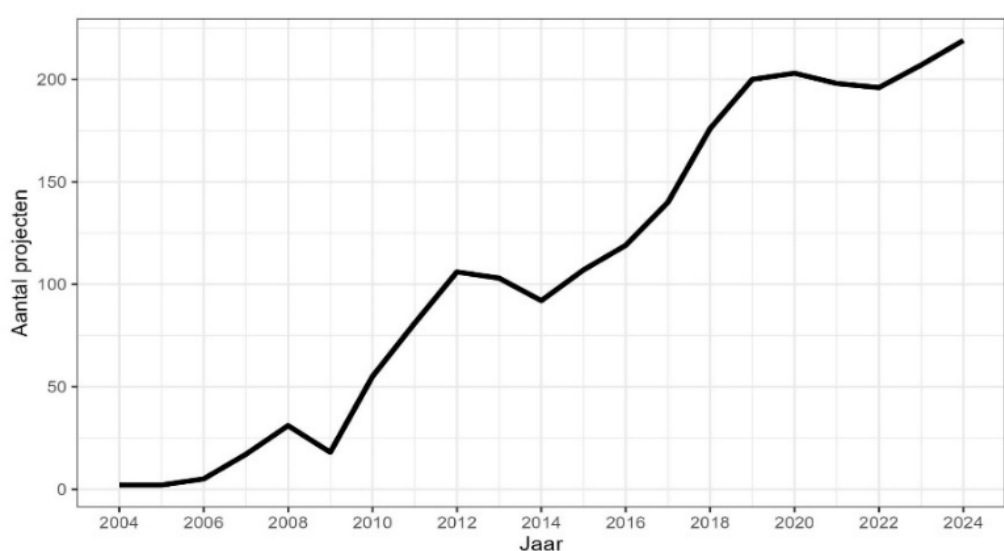
Hoewel het CBS op haar website in detail documenteert welke organisaties van de CBS Microdata gebruikmaken, in welke projecten ze dat doen en welke gegevensbestanden daarbij gebruikt worden (CBS, 2025b), bestaan er, voor zover ons bekend, geen publicaties die het gebruik van de CBS Microdata beschrijven. In deze korte bijdrage beantwoorden we de volgende vragen:

- Welke trends zien we in het gebruik van de CBS Microdata?
- Welke onderzoeksinstellingen gebruiken de CBS Microdata het meest?
- En uiteraard: Wat is de link tussen de CBS Microdata en Aat Liefbroer?

Trends in het gebruik van de CBS Microdata

Op de website van het CBS wordt bijgehouden welke projecten op basis van de CBS Microdata gestart zijn sinds 2004. Op basis van deze informatie hebben we Figuur 1 gemaakt. Die laat zien dat er bijna ieder jaar meer projecten gestart zijn dan in het jaar ervoor. In 2024 zijn er meer dan 200 projecten gestart. De groei verloopt niet helemaal regelmatig, maar kan vanaf 2006 goed benaderd worden door een rechte lijn. Vanaf 2018 ligt de groei echter wat onder het gemiddelde dat door deze lineaire benadering wordt gesuggereerd. Het beeld is dus niet eenduidig.

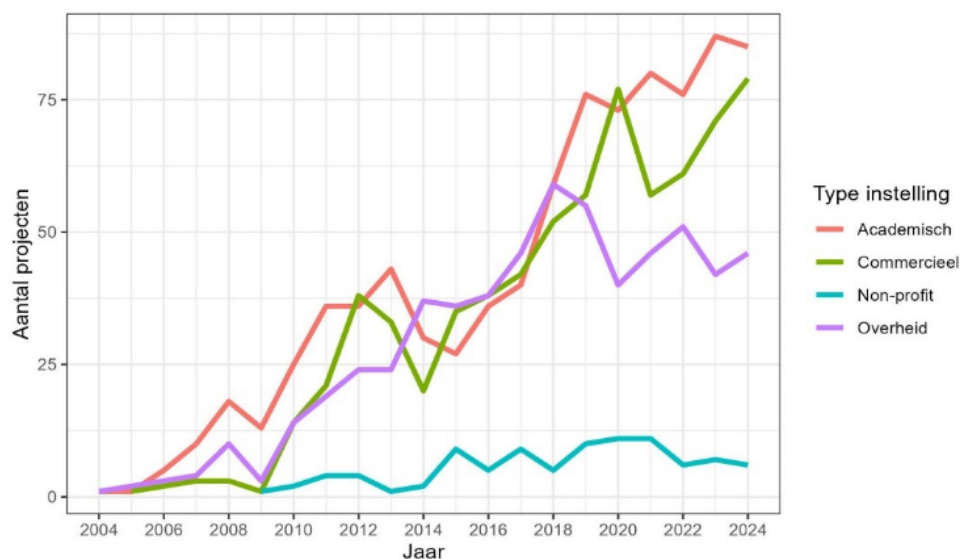
Figuur 1. Ontwikkeling van aantallen projecten in Nederland



Type onderzoeksinstellingen

In totaal is door 386 onderzoeksinstellingen tenminste één project gestart op basis van de CBS Microdata, waaronder 289 onderzoeksinstellingen uit Nederland en 97 onderzoeksinstellingen uit het buitenland. Laatsgenoemde groep is echter voor slechts 5% van het totale aantal projecten verantwoordelijk. In Figuur 2 beschrijven we de ontwikkeling van het aantal projecten naar type onderzoeksinstelling in Nederland. Academische instellingen, commerciële instellingen en overheden hebben lange tijd in toenemende mate en bijna evenveel gebruikgemaakt van de CBS microdata. Sinds 2018 is de benutting door overheden gestabiliseerd. Non-profitinstellingen hebben minder vaak gebruikgemaakt van de data dan academische instellingen, commerciële instellingen en overheden.

Figuur 2. Ontwikkeling van het aantal projecten naar type onderzoeksinstelling in Nederland



CBS Microdata en Aat Liefbroer

Aat maakte meermaals gebruik van de CBS Microdata. Zo was hij, samen met Wim Bernasco en Jan Rouwendal, promotor van Janique Kroese die van deze data gebruikmaakte voor haar onderzoek naar het verband tussen het opgroeien in een eenoudergezin en jeugdcriminaliteit (Kroese et al., 2021; 2024). Ook in verschillende andere projecten waarbij Aat betrokken was, werd gebruikgemaakt van de CBS Microdata (Van Wijk et al., 2025; Vermeulen, 2023a; 2023b). Daarnaast heeft Aat ook gewerkt met het Noorse equivalent van de CBS Microdata (bijv. in Straiton et al., 2024). De opkomst van de CBS Microdata voltrok zich in de laatste 20 jaren van de loopbaan van Aat. Hij heeft deze goudmijn niet leeggeroofd, maar haar wel gevonden en verstandig gebruikt.

Referenties

CBS (2025a). *Stelsel van Sociaal-statistische Bestanden (SSB)*.

CBS (2025b). *Instellingen en projecten*.

Kroese, J., Bernasco, W., Liefbroer, A.C. & Rouwendal, J. (2021). Single-parent families and adolescent crime: Unpacking the role of parental separation, parental decease, and being born to a single-parent family. *Journal of Developmental and Life-Course Criminology*, 7, 596–622.

Kroese, J., Bernasco, W., Liefbroer, A.C. & Rouwendal, J. (2024). The anticipatory, short-term, and long-term effects of parental separation and parental death on adolescent delinquency. *Journal of Developmental and Life-Course Criminology*, 10, 288–308.

Vermeulen, W. R. J., Zoutewelle-Terovan, M., Kooiman, N. & Liefbroer, A. C. (2023a). Comparing marriage and cohabitation dissolution risks across municipalities: Which characteristics matter? *Population, Space and Place*, 29(5).

Vermeulen, W. R. J., Zoutewelle-Terovan, M., Kooiman, N. & Liefbroer, A. C. (2023b). Religion and union dissolution: Effects of couple and municipal religiosity on divorce and separation. *Demographic Research*, 49, 513–542.

Wijk, D. van, Valk, H. A. G. de & Liefbroer, A. C. (2025). Who becomes a parent while jobless? On the importance of limited opportunities and the normalization of disadvantage. *European Sociological Review*, 1-16.

Straiton, M., Reneflot, A., Hynek, K. A., Liefbroer, A. C. & Hauge, L. J. (2022). Mental disorder and subsequent marital separation among migrant and non-migrant women. *Health Care for Women International*, 44(9), 1073–1091.