

## Leiden grootschalige datasets tot saaie artikelen?

Wie in Berlijn op zondag fruit wil kopen kan terecht bij een *bakal* (waarschijnlijk van het Arabische *baqala* dat groenteman betekent, met dank aan Annelies Moors en Thijl Sunier), zoals zo'n winkel in Turkije heet, een kleine groentewinkel annex minisupermarkt die zelfs op zondagavond open is. In de kleine winkel die ik bezocht, zat de vrouw van het Turkse echtpaar achter de kassa, was de man bezig met de voorraad en speelde hun zoontje met een autootje. Er kwamen enkele Turkse kennissen langs om even te kletsen. Het sociale leven van het gezin bestond op zondagavond uit samen in de winkel werken. In mijn grootstedelijke buurt valt op zondag in zulke winkeltjes hetzelfde sociale leven waar te nemen. Die allochtone winkeliers zijn bovendien, in vergelijking met de alledaagse Amsterdamse omgangsvormen, onwaarschijnlijk vriendelijk. Hier zien we een andere vorm van sociaal leven dan gebruikelijk is in Nederland en een klantgerichte vriendelijkheid die bijna was verdwenen. Het is te hopen dat die Turkse en Marokkaanse winkeliers niet alle Nederlandse normen en waarden zullen overnemen want anders zijn we onze vriendelijke nieuwe middenstanders weer kwijt.

Wie over dit onderwerp wil lezen kan terecht bij talrijke studies over etnisch ondernemerschap, veelal geschreven door antropologen. Meestal beginnen die studies met observaties zoals bovenstaande stijlloefening en dan zal de publicatie uitwaaiëren naar beschouwingen over etnische winkels als economische *niche* voor migranten, transnationale netwerken of schrijven de auteurs over de betekenis van het sociale leven in andere culturen waar werk en privé minder gescheiden lijken te zijn dan in de huidige Nederlandse samenleving. Het is een stijl die streeft naar levensechtheid. Dat levert dikwijls boeiende literatuur op, maar een strenge systematiek ontbreekt, want hypothesen worden zelden expliciet geformuleerd of getoetst. Deze stijl wordt zeker ingegeven door de kleine dataset, maar dat een kleinschalige dataset ook anders geanalyseerd kan worden, hebben Miles en Huberman (1994) in hun boek *Qualitative Data Analysis* laten zien. Maar antropologen zijn zeer aan hun stijl van de gevalstudie, die overigens al zo'n zeventig jaar oud is, gehecht. Door geavanceerd kwantitatieve onderzoekers wordt die stijl soms 'journalistiek' genoemd en die van henzelf 'wetenschappelijk', maar zo simpel ligt dat niet. Beide groepen laten zich leiden door de heersende normen binnen hun eigen peergroup en die is modegevoelig. Bij kwantitatief onderzoek is de stijl steeds sterker gestructureerd geraakt, zo leert een blik op artikelen in *Mens en Maatschappij* van dertig jaar geleden. Die huidige opzet van veel sociaal-wetenschappelijke artikelen kan positief gezien worden als een wetenschappelijke evolutie, maar is even zo goed te begrijpen als een verlangen om op een 'echte' wetenschap als psychologie of economie te lijken. Die gestructureerde stijl wordt door buitenstaanders vaak onleesbaar genoemd. En om zulke publicaties aan een breder publiek uit te leggen zijn weer geheel andere schrijfvvaardigheden nodig, waar de meeste wetenschappers niet over beschikken of geen zin in hebben.

Is het nu zo dat alleen bij een beperkt aantal onderzoekseenheden zo'n boeiende stijl mogelijk is? Een bepaalde coëfficiënt uit een multi-niveau-analyse op een grote dataset kan erg interessant zijn maar valt immers moeilijk te beschrijven in een 'levensechte' stijl. Het lijkt erop dat de schrijfstijl van een wetenschappelijk artikel mede bepaald wordt door de omvang van de gebruikte dataset. Ik heb grote bewondering voor de ijver en het technische vernuft waarmee die grootschalige analyses worden uitgevoerd, maar de lectuur is soms ietwat teleurstellend, omdat interessante variabelen in het bestand helaas blijken te ontbreken of de helft van de onderzoekseenheden moet worden verwijderd, zoals de schrijvers trouwhartig melden. De analyses leveren meestal zeer globale conclusies op, zonder een uitgebreidere interpretatie.

Enige tijd geleden merkte de Groningse socioloog Rafael Wittek op dat de Nederlandse sociologen steeds meer op geavanceerde wetenschappelijke boekhouders beginnen te lijken (*Facta*, april 2004). Het was een overduidelijke boutade. Boekhouders zijn immers saai en vallen niet bepaald op als nieuwsgierige ontdekkingsreizigers, zoals Wittek opmerkte. Saaiheid is echter een slecht criterium want het gaat eigenlijk om goed of slecht schrijven. En wie denkt dat saaiheid kenmerkend is voor de kwantitatieve sociaal-wetenschappen moet eens een blik werpen in een gerenommeerd internationaal tijdschrift voor geschiedenis vol pietluttige onderwerpjes. Iets saai noemen betekent vaak dat het onderwerp bij de lezer geen belangstelling wekt maar dat is het gevolg van een zich steeds verder specialiserende wetenschapsbeoefening. Maar saai kan ook betekenen dat het artikel niet boeiend is geschreven en eindigt zonder prikkelende conclusies. Het vereist immers een grote schrijfvaardigheid om die kwantitatieve analyses leesbaar te rapporteren en dat wordt vaak onderschat, zodat zelfs een onderwerp verprutst kan worden door een onbeholpen stijl.

Maar het is toch niet helemaal alleen een kwestie van stijl. In bijvoorbeeld CBS-publicaties zijn de etnische ondernemers moeilijk terug te vinden. Hoe grootschalig moet een dataset eigenlijk zijn om een kleinschalig verschijnsel waar te kunnen nemen? Misschien blijken etnische ondernemers een veel te kleine groep voor de beroepsstatistieken. Dat zou kunnen, maar het gebrek aan aandacht in de statistieken is misschien vooral een kwestie van andere vragen stellen. Iets meer avontuurlijkheid kan geen kwaad, daar geef ik Wittek gelijk in. Wetenschappelijke boekhouders moeten zeker doorgaan met hun werk, op voorwaarde dat ze beter op de kleine postjes letten. Of anders gezegd: als ze erin slagen om ook eens een klein onderwerpje te analyseren dat in hun grote datasets te vinden is en zulke uitzonderingen niet wegpoezen in hercoderingen en onverklaarde variantie. Over zo'n analyse mogen ze als tegenhanger van de antropologische casestudies best een streng kwantitatief artikel publiceren en geplaatst in een bredere context, als het dan maar boeiend is geschreven.

H. Schijf