



**rijksuniversiteit
groningen**

**faculteit gedrags- en
maatschappijwetenschappen**



Verschillen in Friesland

**Segmentatie in de Friese samenleving naar verdeling van economisch-, sociaal-, cultureel
en persoonskapitaal**

Naam: Marten Hospes
Student nr.: 2778777
Begeleider: dr. R.J.J. Wielers
Referent: prof. dr. M.A.J. van Duijn
Startdatum scriptie: 20 februari 2019
Einddatum scriptie: 15 juli 2021

Samenvatting

In onderzoek naar ongelijkheid wordt er in toenemende mate gekeken naar meerdere vormen van kapitaal om segmentatie in de samenleving te verklaren. Economisch kapitaal was al een belangrijke bron van segmentatie, maar sociaal-, cultureel- en persoonskapitaal tellen ook steeds meer mee als dimensie waarop een samenleving segmenteert. Voortbouwend op onderzoek van het SCP (2014) naar segmentatie in de Nederlandse samenleving, wordt in dit onderzoek een verkennend onderzoek gedaan naar de segmentatie in de Friese samenleving op basis van economisch-, sociaal-, cultureel en persoonskapitaal. Friesland onderscheidt zich van de rest van Nederland door een andere demografische samenstelling en de ‘Friese paradox’; hoewel de inwoners van Friesland gemiddeld een lagere sociaaleconomische status hebben, scoren ze wel hoger dan gemiddeld op welzijnsaspecten.

De segmentatie is onderzocht met een *Latent-Classes-Analyses* (LCA) met data van Panel Fryslân, een representatief panel in beheer van het Fries Sociaal Planbureau (FSP), verzameld onder volwassen inwoners van de provincie Friesland. De onderzoekspopulatie bedraagt 1137 respondenten, de vragenlijsten zijn op drie verschillende momenten afgenomen.

Uit de LCA blijkt dat er vier segmenten zijn te onderscheiden in Friesland: precariaat, socialen, rijk gepensioneerden en werkenden. Het precariaat is de groep die het minst van alle vier de vormen van kapitaal tot zijn beschikking heeft en daarmee lijkt het de groep die ‘het niet heeft’. De socialen hebben opvallend veel sociaal en cultureel kapitaal, ervaren hun eigen gezondheid als goed, maar op economisch gebied hebben ze relatief weinig hulpbronnen. De rijk gepensioneerden onderscheiden zich vooral op het gebied van economisch kapitaal. Van de andere vormen van kapitaal hebben ze weinig, vooral sociaal kapitaal. Tot slot is er de groep werkenden. Deze groep bestaat vooral uit werkenden en heeft de beschikking over veel persoons-, sociaal- en economisch kapitaal. Daarmee lijkt het de groep die ‘het heeft’.

Inhoud

Samenvatting.....	2
Inhoud	3
Inleiding	4
Relevantie.....	6
Theoretisch kader.....	7
Economisch kapitaal.....	7
Sociaal kapitaal	8
Cultureel kapitaal	9
Persoonskapitaal	10
Segmentatie in de Nederlandse samenleving.....	11
Methodologie	14
Analyse-opzet.....	14
Panel Fryslân.....	15
Werving	15
Respons	15
Onderzoekspopulatie	16
Latente Klassen Analyse (LCA).....	17
Wegen	20
Operationalisatie.....	21
Resultaten	29
Aantal klassen.....	29
Beschrijving van de klassen	32
Latente klassen analyse met gewogen data.....	33
Vergelijking tussen ongewogen en gewogen data.....	34
Invloed van onderliggende dimensies.....	34
Conclusie	36
Discussie.....	37
Literatuur.....	39
Bijlage	45

Inleiding

Ongelijkheid heeft nadelige gevolgen. Het zorgt voor tegenstellingen in de maatschappij en die tegenstellingen kunnen leiden tot maatschappelijke frustratie en gebrek aan sociale samenhang. Ongelijkheid in de samenleving komt voor uit de beschikking over hulpbronnen. In vroegere ongelijkheidsonderzoeken richtten onderzoekers zich vooral op hulpbronnen op economisch vlak, maar tegenwoordig wordt er naar meerdere vormen van hulpbronnen gekeken om segmentatie in de samenleving te onderscheiden, omdat economische verschillen niet meer voldoende zijn om ongelijkheid te verklaren (Putnam, 2015). Voorbeelden van een benaderingswijze op verschillende dimensies zijn de brede welvaartsmonitor van het CBS (Monitor Brede Welvaart CBS, 2018) en het segmentatie-onderzoek van het SCP (2014).

Een gevolg van het onderscheiden van meerdere hulpbronnen is dat de ongelijkheid in de samenleving met meerdere invalshoeken wordt bekeken. De hulpbronnen tezamen worden kapitaal genoemd. Er worden verschillende vormen van kapitaal onderscheiden. Economisch kapitaal betreft de verdeling van inkomen en vermogen. Sociaal kapitaal gaat over het voordeel wat mensen halen uit sociale normen. Cultureel kapitaal brengt groepsidentificatie en kan de toegang tot specifieke groepen gemakkelijker maken. En tot slot hebben mensen persoonlijk kapitaal, wanneer ze voordelen kunnen halen uit hun fysieke en mentale gesteldheid. De verdeling van deze verschillende vormen van kapitaal ligt ten grondslag aan de segmentatie van de samenleving.

Het Sociaal en Cultureel Planbureau heeft segmentatie op basis van economisch-, sociaal-, cultureel en persoonskapitaal voor de Nederlandse samenleving onderzocht (SCP, 2014). Dit heeft geleid tot 6 segmenten in de Nederlandse samenleving: de gevestigde orde, jongere kansrijken, werkende middengroep, comfortabel gepensioneerden, onzekere werkenden en precariaat. De twee groepen die de meeste tegenstellingen vertonen zijn de gevestigde orde en het precariaat. De gevestigde orde is het segment met beschikking over alle vier vormen van kapitaal. Het precariaat heeft van geen van deze vormen van kapitaal veel. De andere vier segmenten vallen hier tussen in. Dat betekent echter niet dat er een duidelijke rangorde valt aan te geven. Jongere kansrijken scoren bijvoorbeeld relatief hoog op sociaal kapitaal, terwijl de werkende middenklasse juist relatief hoog scoort op economisch kapitaal.

Voor Friesland is dit mogelijk anders. De Friese samenleving toont verschillen met de Nederlandse samenleving als geheel. Ten eerste zijn er demografische verschillen. De Friese bevolking is ouder dan de rest van Nederland. Friesland telt veel ouderen en de gemiddelde

leeftijd ligt hoog. Daarbij komt dat er relatief veel jongeren vertrekken uit de provincie. Ook telt Friesland relatief veel laagopgeleiden, meer dan het gemiddelde in de Nederlandse samenleving. Deze demografische verschillen hebben mogelijk invloed op de segmentatie van Friesland, waardoor die anders is als in de Nederlandse samenleving als geheel (FSP, 2019).

Ten tweede is er sprake van de ‘Friese Paradox’ (FSP, 2019). De Friese Paradox beschrijft het paradoxale verband tussen welvaart en welzijn in de provincie Friesland. De inwoners van Friesland hebben gemiddeld een lagere sociaaleconomische status (SES) dan de inwoners in de rest van Nederland; het gemiddelde opleidingsniveau en inkomen is lager en de werkloosheid hoger. Op basis van deze aspecten zou men verwachten dat Friezen lager scoren op welzijnsaspecten als gezondheid, geluk en tevredenheid (FSP, 2019). Echter, Friezen scoren juist relatief hoog op welzijnsaspecten. De inwoners van Friesland hebben de hoogste ervaren gezondheid van heel Nederland, ondanks een vergelijkbare leefstijl qua gezondheid als overige inwoners van Nederland en de hogere zorgvraag van de inwoners van Friesland (FSP, 2019). Daarnaast is de sociale samenhang in Friesland ook hoog, wat een positieve indicator van sociaal kapitaal is (Kawachi & Berkman, 2000). De Friese Paradox lijkt te suggereren dat de verbanden die op nationaal niveau gelden met economisch-, sociaal-, cultureel- en persoonskapitaal, niet gelden voor Friesland.

Vanwege de demografische samenstelling van Friesland en vanwege de Friese Paradox is het interessant om te kijken welke verschillen er in Friesland zijn in economisch-, sociaal-, cultureel- en persoonskapitaal, en hoe dat tot segmentatie leidt. Dit onderzoek heeft als doel om een eerste aanzet te geven om de sociale segmentatie van Friesland in kaart te brengen. Daartoe wordt bestaande data van het FSP gebruikt.

Hoewel dit onderzoek voortbouwt op het onderzoek van het SCP (2014), naar segmentatie in de Nederlandse samenleving, kan het onderzoek van SCP niet herhaald worden in Friesland. Dat komt omdat de dataset van het FSP niet de benodigde informatie bevat om het onderzoek te herhalen. Dit heeft tot gevolg dat een strakke vergelijking in de segmentatie tussen Nederland en Friesland niet mogelijk is.

Wel is een vergelijking mogelijk van de hulpbronnen die aan de basis van de segmentatie liggen, dus de mate waarin bijvoorbeeld inkomen of opleidingsniveau het beeld bepaalt. Hoewel de uiteindelijke segmenten dus niet strikt te vergelijken zijn, is de mate van invloed van specifieke hulpbronnen op de segmentatie dat wel. De vraag die ik wil beantwoorden is welke van vorm van kapitaal, of van de onderliggende dimensies, het meeste invloed heeft op de segmentatie in de Friese samenleving en hoe zich dat verhoudt tot de segmentatie in de Nederlandse samenleving.

Dit leidt tot de volgende probleemstelling: Welke groepen zijn er in Friesland te onderscheiden op basis van economisch-, sociaal-, cultureel- en persoonskapitaal en welke vorm van kapitaal of onderliggende dimensie heeft het meest invloed op de segmentatie en hoe verhoudt zich dat tot de rest van Nederland? Deze probleemstelling bestaat dus uit drie onderdelen. Als eerste wordt de segmentatie in Friesland onderzocht, als tweede wordt gekeken hoe de segmentatie tot stand komt en als derde wordt onderzocht hoe zich dat verhoudt tot de rest van Nederland.

Relevantie

De relevantie van dit onderzoek is te vinden in het ongelijkheidsvraagstuk, één van de centrale thema's van sociologie. Dit onderzoek draagt bij aan het ongelijkheidsvraagstuk door de segmentatie van de samenleving te onderzoeken aan de hand van meerdere dimensies. De verschuiving in ongelijkheidsonderzoek van economische dimensies naar meerdere dimensies is door verschillende onderzoekers al onderzocht (bv: Ganzeboom, 1988; Monitor Brede Welvaart SCP, 2018; Murray, 2012). Een belangrijke bijdrage komt van Putnam (2015) die onderzocht dat economische verschillen niet voldoende zijn om ongelijkheid in de maatschappij te verklaren en dat onder andere inbedding in gemeenschappen invloed heeft op de ongelijkheid. Een andere belangrijke bijdrage komt van Bourdieu (1986) die stelde dat de omgeving waarin je opgroeit, bepalend is voor de sociale klasse waarin je terechtkomt en daarmee wat de kansen in het leven zijn. Daarnaast is onderzocht hoe economisch-, sociaal-, cultureel- en persoonskapitaal tot segmentatie leiden (Savage et al. 2013; SCP, 2014). Het is interessant om de segmentatie van Friesland te onderzoeken, omdat de demografische samenstelling anders is ten opzichte van Nederland en er sprake is van 'de Friese Paradox'. Het vermoeden bestaat dat de verbanden tussen economisch-, sociaal-, cultureel- en persoonskapitaal anders lopen dan in de rest van Nederland en segmentatie daardoor op een andere manier tot stand komt.

Theoretisch kader

In het theoretisch kader worden economisch-, sociaal-, cultureel- en persoonskapitaal toegelicht. Voor elke vorm van kapitaal wordt toegelicht wat de vorm van kapitaal betekent en hoe dat tot segmentatie kan leiden. Daarna wordt onderzocht in hoeverre desbetreffende vorm van kapitaal invloed heeft in de Friese samenleving en hoe zich dat verhoudt tot Nederland. Tot slot wordt de segmentatie in de Nederlandse samenleving besproken volgens het onderzoek van het SCP over Verschil in Nederland (2014).

Kapitaal is een hulpmiddel waarvan meer kan worden gemaakt middels investeringen (Van Doorne-Huiskes, 2015). De vermeerdering van kapitaal wordt vaak gelinkt aan economische doeleinden, maar er zijn ook andere vormen van kapitaal (Savage et al., 2013). Zo worden ook sociaal-, cultureel- en persoonskapitaal onderscheiden. Sociaal kapitaal gaat over hulpbronnen waar men toegang tot krijgt via sociale relaties. Cultureel kapitaal bestaat uit gebruiken die gebonden zijn aan een sociale klasse (Bourdieu, 1986). Persoonskapitaal zijn lichaamseigene kenmerken op zowel fysiek als psychisch vlak waar men voordelen uit kan halen (SCP, 2014).

Economisch kapitaal

Economisch kapitaal zijn alle geldelijke middelen die een individu tot zijn of haar beschikking heeft en waaraan hij winst kan ontlene (Parcel & Meuller, 1983). Bij economisch kapitaal zijn geldelijke kapitaalmiddelen de hulpbron die het individu in kan zetten om zijn doel te bereiken. De geldelijke middelen die een individu tot zijn beschikking heeft, kunnen zowel afkomstig zijn uit arbeid als uit rendement op vermogen (SCP, 2014). Het rendement op vermogen is de laatste jaren meer onder de aandacht gekomen vanwege de toenemende ongelijkheid in vermogen (Piketty, 2015). Ook in Nederland is het vermogen ongelijk verdeeld (Van Bavel, 2014). De 5% meest vermogende huishoudens bezit in Nederland 52,5% van het totale vermogen (Reuten, 2018). Het gemiddelde vermogensaandeel van de 5% meest vermogende huishoudens is in OECD-landen is 38,2% (Reuten, 2018). Dit geeft te kennen dat de vermogensverdeling in Nederland scheef is.

Naast inkomen uit vermogen, is er ook inkomen uit arbeid. Het inkomen uit arbeid is in de loop van de jaren steeds meer afhankelijk geworden van de gevolgde opleiding (SCP, 2014). De *human capital theory* van Becker (Sweetland, 1996; Olaniyan & Okemakinde, 2008) verklaart deze toenemende afhankelijkheid van opleiding. Volgens de *human capital theory* is opleiding een economisch goed. Een betere opleiding zorgt voor een hogere

arbeidsproductiviteit. Aangezien mensen kunnen investeren in hun opleiding, wordt het gezien als kapitaal. De opleiding betaalt zich uit in een hoger inkomen. Een opleiding levert dus economisch rendement (Olaniyan & Okeminda, 2008).

Onderzoek in Nederland heeft doen blijken dat opleiding steeds belangrijker wordt om economische verschillen te verklaren (Ganzeboom & Ultee, 1996). Dit wordt ook wel de overgang van *ascription* naar *achievement* genoemd (Ganzeboom & Treiman, 2007). Mensen die een hoge opleiding hebben afgerond, verdienen bijna dubbel zoveel als mensen die een lage opleiding hebben afgerond (CBS, 2011).

De verdeling van het economisch kapitaal binnen de Friese samenleving verschilt met de Nederlandse samenleving. In Friesland is het opleidingsniveau gemiddeld lager dan in de rest van Nederland (Monitor Onderwijs FSP, 2018). Dit komt ook terug in het inkomen. Na de provincie Groningen is de provincie Friesland de provincie met het laagste inkomen (Monitor Inkomen FSP, 2018). Over de vermogensverdeling in Friesland is weinig bekend. Het is daarom lastig om aan te geven hoe de positie van vermogen in Friesland is ten opzichte van Nederland. Op basis van inkomen en opleiding kan er gesteld worden dat er minder economisch kapitaal is in Friesland ten opzichte van de rest van Nederland.

Sociaal kapitaal

Relaties binnen sociale structuren vormen het sociaal kapitaal. Sociale structuren zijn zowel persoonlijke relaties, sociale netwerken als gemeenschappen. Sociaal kapitaal faciliteert acties van mensen binnen sociale structuren. Door relaties en inbedding in sociale structuren komen er hulpbronnen beschikbaar die anders niet beschikbaar waren (Coleman, 1988; Lin, 1999; Engbersen, 2003; Putnam, 2015).

Sociaal kapitaal kan verdeeld worden in cognitief en structureel sociaal kapitaal. Het cognitieve aspect van sociaal kapitaal gaat om de perceptie van sociale steun, wederkerigheid en vertrouwen (Harpham, Grant & Thomas, 2002). Doordat een individu ingebed is in sociale structuren, voelt een individu zich gesteund. Dit helpt onder andere bij het verwerken van moeilijke levensgebeurtenissen. Daarnaast blijkt uit empirisch onderzoek dat dat er een sterk verband is tussen gezondheid, sociale netwerken en sociale steun. Mensen met meer sociale steun zijn gezonder (Campbell, Wood & Kelly, 1999)

Naast cognitief sociaal kapitaal is er ook structureel sociaal kapitaal. Structureel sociaal kapitaal gaat over de omgang en de intensiteit van sociale netwerken. Het structurele aspect komt vooral naar voren bij de toegang tot hulpbronnen, zoals informatiestromen. Het

sociale netwerk waarin men verkeert en de positie die men daarin aanhangt heeft, zijn bepalend voor de toegang tot informatiestromen.

Ten opzichte van de Nederlandse samenleving, lijkt het sociaal kapitaal in de Friese samenleving hoog. Een indicator voor een hoge mate van sociaal kapitaal is de sociale samenhang (Putnam, 1995). Hoewel deze indicator niet de relaties binnen sociale structuren meet, duiden de aanwezigheid van deze twee indicatoren op een hoog sociaal kapitaal in de samenleving. Voor Friesland geldt dat de sociale samenhang sterk is (FSP, 2019). Het aantal vrijwilligers is het hoogste van heel Nederland (Monitor Sociale Samenhang FSP, 2018). Dit zou er op kunnen duiden dat sociaal kapitaal in de Friese samenleving groter is dan in de Nederlandse samenleving. netwerk

Cultureel kapitaal

Cultureel kapitaal gaat over normen, gedeelde interesses, taalgebruik en consumptie die gebonden zijn aan een sociale klasse (Bourdieu, 1986). Het cultureel kapitaal van een individu (en gemeenschap) wordt gevormd in de habitus. De habitus zijn gesocialiseerde normen en gedragsneigingen die gedrag en denkwijzen sturen. Daardoor heeft de habitus invloed op de acties en keuzes van een individu (Bourdieu, 1986).

Habitus zijn gesocialiseerde normen. Dat betekent dat de habitus, en dus het cultureel kapitaal, wordt gevormd door interactie in een groep of gemeenschap (Pinxten & Lievens, 2014). Dit betekent dat een individu gevormd wordt door sociale normen, bepaald door de gemeenschap; ergo, hoe een persoon zich gedraagt, is afhankelijk van de omgeving waarin hij of zij opgroeit. Daarmee is cultureel kapitaal belangrijk bij het ontstaan en doorgeven van maatschappelijke verschillen.

Cultureel kapitaal zorgt voor verschillen tussen individuen vanwege de gemeenschap waar het individu in leeft. In een gemeenschap wordt het cultureel kapitaal van een individu gevormd. Aangezien individuen meestal verweven zijn in meerdere gemeenschappen, zorgt dit ervoor dat het cultureel kapitaal verschilt per individu. Dat betekent dat het cultureel kapitaal per individu uniek is (Bourdieu, 1986). Binnen gemeenschappen zijn er grote overeenkomsten in het cultureel kapitaal. Mensen die een verschillend cultureel kapitaal hebben, zijn verschillend gesocialiseerd en hebben daardoor verschillende normen. Hierdoor wordt het lastig om met elkaar om te gaan. Dit zorgt voor verschillen in de maatschappij.

Een belangrijk onderdeel van de omgeving waarin het individu opgroeit is de taal (Shwayli & Barnes, 2018). De manier waarop groeperingen en gemeenschappen met elkaar communiceren is kenmerkend voor een groep. Taal heeft op twee manieren invloed op

cultureel kapitaal. Ten eerste geeft een goede taalbeheersing voordelen op de arbeidsmarkt. Wanneer men op de arbeidsmarkt de heersende taal spreekt vindt, kan men zich sneller aanpassen en vindt men makkelijker aansluiting (Shwayli & Barnes, 2018). Ten tweede draagt taal bij aan het groepsgevoel. Wanneer men allemaal dezelfde taal spreekt, creëert dat een band. Het versterkt de identificatie als groep (Shwayli & Barnes, 2018).

In Friesland is de tweede rijkstaal van Nederland, het Fries, de meest gebruikte taal. Het Fries is een belangrijk onderdeel van de Friese gemeenschap en is voor ongeveer 60% van de Friezen de moedertaal (“[Taal in Nederland]”, z.d.). Een mogelijk gevolg zou kunnen zijn dat er segmentatie ontstaat door de beheersing, of juist het niet beheersen van de Friese taal.

Persoonskapitaal

Van de vier besproken vormen van kapitaal, is persoonskapitaal de recentste vorm van kapitaal als dimensie waarlangs onderzoek naar segmentatie plaatsvindt (SCP, 2014).

Persoonskapitaal bestaat uit kenmerken van een persoon die gebonden zijn aan zijn of haar lichaam, kort gezegd: lichaamseigene kenmerken. Die lichaamseigene kenmerken stellen individuen in staat om doelen te bereiken of voordelen uit die kenmerken te halen.

Persoonskapitaal is te verdelen in drie dimensies: fysiek, mentaal en esthetisch. Fysiek persoonskapitaal is de staat van iemands lichamelijke gezondheid, mentaal persoonskapitaal gaat over de geestelijke gesteldheid en esthetisch persoonskapitaal gaat over uiterlijke kenmerken zoals aantrekkelijkheid en uitstraling.

Het fysiek persoonskapitaal is ongelijk verdeeld in de samenleving. De reden daarvan is dat jongeren over het algemeen gezonder zijn dan ouderen (SCP, 2012). Een indicator van fysiek persoonskapitaal is zonder lichamelijke beperkingen een trap op te lopen. Negen van de tien jongeren kunnen trap lopen zonder lichamelijke beperkingen en van de 75-plussers is dat vier op de 10 (SCP, 2014). Dit geeft aan dat jongeren meer fysiek persoonskapitaal hebben dan ouderen.

De verdeling van mentaal persoonskapitaal is minder makkelijk te duiden dan de verdeling van fysiek persoonskapitaal. Er is in verhouding meer onderzoek gedaan naar de negatieve gevolgen van psychologische vermogens, zoals stoornissen, dan over positieve gevolgen van psychologische vermogens (SCP, 2014; De Graaf et al, 2012). Zaken die volgens het SCP (2014) onder psychologische vermogens vallen zijn zelfvertrouwen, zelfbeeld en veerkracht. Veenhoven definieert mentaal persoonskapitaal als het psychologische vermogen dat mensen bezitten om iets te bereiken en zich aan te passen en

staande te houden in hun omgeving (Veenhoven, 2000). Een individu met een hoge mate van mentaal persoonskapitaal is beter in staat om zijn doelen te bereiken. Het hoge mentale persoonskapitaal is een hulpbron die hem/haar voordeel geeft.

Esthetisch persoonskapitaal behelst zowel fysieke kenmerken, zoals schoonheid en slankheid, als mentale kenmerken zoals uitstraling. De verdeling van esthetisch kapitaal is niet eenvoudig te bepalen, omdat esthetische kenmerken vaak een subjectief karakter hebben. Het SCP (2014) noemt in het onderzoek 'verschil in Nederland' verschillende onderzoeken die te kennen geven dat mensen die als aantrekkelijk worden ervaren, betere kansen hebben op de arbeidsmarkt en ook daadwerkelijk meer verdienen.

Het objectieve gedeelte van esthetisch persoonskapitaal is te meten met het BMI, een maat op basis van iemands lengte en gewicht. Een lage BMI duidt op ondergewicht en een hoge BMI op overgewicht. Op het gebied van (over)gewicht vinden Brunello en D'Hombres (2007) een verband tussen BMI en hoogte van lonen. In een studie met data afkomstig uit meerdere Europese landen, hebben de onderzoekers aangetoond dat een 10% toename van de BMI de lonen met 3,3% doet dalen voor mannen en 1,9% voor vrouwen.

Voor fysiek persoonlijk kapitaal geldt dat er voor Friesland bekend is dat de inwoners van Friesland hoog scoren op ervaren gezondheid. De relatie tussen de ervaren gezondheid en de ware gezondheid is positief. Wanneer mensen hun gezondheid als goed ervaren, duidt dat in de meeste gevallen ook op een goede gezondheid (Eskelinen, Kohvakka, Merisalo, Hurri & Wagar, 1991)

Van de volwassen inwoners van Nederland beoordeelt 76% hun lichamelijke gezondheid als (zeer) goed en voor Friesland is dat 79%: het hoogst van alle provincies. Er is geen gemiddelde BMI score bekend van de inwoners van Friesland, wel is het aantal mensen met overgewicht bekend: 50%. Dit percentage is gelijk aan de rest van Nederland (www.volksgezondheid.info). Er kan dus gesteld worden dat de inwoners van Friesland het goed doen op het gebied van ervaren gezondheid. Over andere indicatoren van persoonskapitaal is weinig bekend, waardoor het moeilijk is om een vergelijking te maken tussen Friesland en de rest van Nederland.

Segmentatie in de Nederlandse samenleving

In deze paragraaf wordt de segmentatie in de Nederlandse samenleving beschreven op basis van de verdeling van economisch-, sociaal-, cultureel- en persoonskapitaal volgens het SCP (2014). Zoals aangegeven in de inleiding, is dit onderzoek geïnspireerd op het onderzoek van

het SCP (2014) 'Verschil in Nederland'. Vanwege de verschillende data is het onderzoek van het SCP niet één op één herhaalbaar in Friesland en daarom is een directe vergelijking niet mogelijk. Dat neemt niet weg dat het wel interessant is om de segmentatie van de Nederlandse samenleving te beschrijven en te analyseren welke vormen van kapitaal of onderliggende dimensies invloed hebben op de segmentatie.

In de Nederlandse samenleving zijn er zes segmenten te onderscheiden op basis van de verdeling van economisch-, sociaal-, cultureel- en persoonskapitaal. De twee segmenten die het meest van elkaar verschillen zijn de gevestigde bovenlaag en het precariaat. De andere vier segmenten vallen er tussenin.

De gevestigde bovenlaag scoort hoog op economisch-, sociaal-, cultureel- en persoonskapitaal. Zo heeft deze groep een hoog inkomen, is hoogopgeleid en beschikt het over een groot sociaal netwerk. De gevestigde bovenlaag heeft een aandeel van 15% in de Nederlandse samenleving. Het tegenovergestelde is het precariaat. Dit segment heeft zowel weinig economisch-, sociaal-, cultureel- als persoonskapitaal. Ook dit segment heeft een aandeel van 15% in de Nederlandse samenleving.

Naast de gevestigde orde, beschikken ook jongere kansrijken over veel kapitaal. Dit segment bestaat voornamelijk uit tamelijk jonge mensen die net zoveel sociaal-, cultureel- en persoonskapitaal hebben als de gevestigde orde. Echter, de jongere kansrijken lopen achter met economisch kapitaal ten opzichte van de gevestigde orde. De jongere kansrijken hebben een aandeel van 13% in de samenleving en zijn daarmee het kleinste segment in de Nederlandse samenleving volgens het SCP.

De werkende middengroep is het grootste segment in de Nederlandse samenleving met een aandeel van 27% van de Nederlandse samenleving. In verhouding tot de jonge kansrijken, heeft de werkende middengroep minder sociaal- en cultureel kapitaal, maar wel meer economisch kapitaal. Deze groep omvat relatief veel mensen in loondienst.

De comfortabel gepensioneerden vertonen veel overeenkomsten met de werkende middengroep op het gebied van economisch-, sociaal- en cultureel kapitaal. Ze onderscheiden zich vooral door de leeftijd. Het segment bestaat vooral uit ouderen en gepensioneerden. Vanwege de gevorderde leeftijd is deze groep in staat geweest om vermogen op te bouwen. Door het hoge vermogen kunnen de comfortabel gepensioneerden zich een tamelijk luxueuze leefstijl veroorloven. Deze groep vormt 17% van de Nederlandse samenleving.

De onzekere werkenden is een segment die qua leeftijdsopbouw en aantal werkenden overeenkomsten toont met de werkende middengroep. Echter, deze groep is beter te vergelijken met het precariaat. Ze vertonen op het gebied van kapitaal veel overeenkomsten

met het precariaat, ze scoren echter net wat hoger. Deze groep heeft veel werkenden met een tijdelijke baan en het zelfvertrouwen is het laagst van alle segmenten. Het aandeel van de onzekere werkenden in de Nederlandse samenleving is 14%.

Tot slot is er het segment precariaat. Dit segment vertoont veel overeenkomsten met de onzekere werkenden, maar scoort op bijna alle dimensies net wat minder. Ze hebben van alle segmenten de laagste vormen van kapitaal. Ze blijven op alle fronten achter en hebben de meeste onzekerheid. Het precariaat heeft een aandeel van 15% in de samenleving. De naam precariaat is afkomstig uit het onderzoek van Savage et al. (2013) en is letterlijk overgenomen door het SCP.

Eerder ongelijkheidsonderzoek richtte zich vooral op economische verschillen en nog steeds zorgt economisch kapitaal voor sterke verschillen, maar door de toevoeging van andere dimensies wordt de segmentatie wel specifieker. Economisch kapitaal geeft vooral een tweedeling. Zo hebben drie segmenten redelijk veel economisch kapitaal en drie groepen weinig economisch kapitaal. De gevestigde bovenlaag, werkende middengroep en comfortabel gepensioneerden hebben veel economisch kapitaal. De jonge kansrijken, onzekere werkenden en het precariaat hebben weinig economisch kapitaal.

Het economisch kapitaal alleen is niet voldoende om segmentatie te verklaren en geeft niet een volledig beeld. Zo hebben de jonge kansrijken veel overeenkomsten met de gevestigde orde, echter missen de jonge kansrijken het economisch kapitaal. Voor de werkende middenklasse geldt dat ze sociaal- en cultureel kapitaal missen ten opzichte van de gevestigde orde, terwijl ze wel bijna gelijk zijn qua economisch kapitaal.

Opvallend is de overeenkomst tussen sociaal en cultureel kapitaal. Binnen segmenten zijn weinig verschillen tussen deze twee vormen van kapitaal; wanneer een segment veel sociaal kapitaal heeft, heeft het ook veel cultureel kapitaal en visa versa.

Persoonskapitaal lijkt weinig te doen op de segmentatie van de Nederlandse samenleving. Het enige wat persoonlijk kapitaal lijkt te doen, is sterk gekoppeld aan leeftijd: jongere kansrijken hebben een hoge mate van gezondheid waardoor het persoonlijk kapitaal relatief hoog is. Hoewel leeftijd geen onderdeel is van één van de vormen van kapitaal, is de variatie binnen leeftijdsgroepen opvallend. Zo zijn ouderen vaker onderdeel van comfortabel gepensioneerden en het precariaat en zijn er vooral jongeren bij de jonge kansrijken en weinig bij de gevestigde bovenlaag. Dit geeft te kennen dat leeftijd en de mate van kapitaal afhankelijk van elkaar zijn.

Methodologie

In de methodologie wordt uitgelegd hoe de segmentatie in Friesland onderzocht wordt. Allereerst wordt de analyse-opzet gegeven. Vervolgens wordt in de uitleg gegeven hoe Panel Fryslân tot stand is gekomen. Na de onderzoeksverantwoording van Panel Fryslân wordt uitgelegd hoe een Latente Klassen Analyse (LCA) in zijn werk gaat. Tot slot wordt in de operationalisatie de variabelen behandeld waarmee de LCA wordt uitgevoerd.

Analyse-opzet

Dit onderzoek naar segmentatie in Friesland is geïnspireerd op het onderzoek van het SCP naar segmentatie in de Nederlandse samenleving (2014). Het SCP onderzoekt de segmentatie met een Latente-Klassen-Analyse (*Latent-Class- Analysis* LCA). Ook in dit onderzoek wordt de segmentatie onderzocht met een LCA.

Een LCA zoekt naar latente klassen op basis van manifeste variabelen. Daarmee is LCA een methode voor indelen van respondenten in klassen waarbij op voorhand niet vaststaat hoeveel klassen er zijn en wat de vorm is van de klassen (Vermunt & Magidson, 2002). Het indelen van de respondenten kan plaatsvinden op één of meerdere variabelen. De LCA zoekt naar gelijkenissen tussen de respondenten op basis van de variabelen en geeft per variabele aan hoe deze verdeeld is over de klassen (Vermunt & Magidson, 2002). Verdere informatie over de LCA is te vinden in de paragraaf ‘Latente Klassen Analyse’.

Om segmentatie in Friesland zo goed als mogelijk te onderzoeken, dient de onderzoekspopulatie een afspiegeling te zijn van de bevolking van Friesland, want hoe meer de respondenten een afspiegeling zijn van Friesland, hoe beter de uitkomsten van dit onderzoek van toepassing zijn op Friesland. In Panel Fryslân is dit niet het geval. Om die reden moet er gecorrigeerd worden voor de scheve representativiteit. Om te corrigeren wordt normaal een weegfactor toegevoegd in de analyse, maar in dit onderzoek gaat dat anders. In de paragraaf ‘Wegen’ wordt hier uitvoerig op ingegaan.

De LCA zoekt dus naar latente klassen op basis van manifeste variabelen. Hierdoor is een LCA uitermate geschikt om de segmentatie in de Friese samenleving te onderzoeken op basis van de gegevens die verzameld zijn met Panel Fryslân. Met behulp van de LCA kan duidelijk worden hoe Friesland is gesegmenteerd.

Nadat de segmentatie in Friesland is beschreven, wordt de segmentatie in Friesland vergeleken met de segmentatie in Nederland. Echter, een probleem is dat de data in dit onderzoek teveel verschilt met de data die het SCP gebruikt in haar segmentatieonderzoek.

Wel is het mogelijk om te vergelijken welke vorm van kapitaal of onderliggende dimensie het meest invloed heeft op de segmentatie in zowel Nederland als Friesland. Dus daarom wordt in dit onderzoek gekeken naar welke vorm van kapitaal of onderliggende dimensie, sterk of zwakke invloed heeft op de vorming van de klassen in de LCA.

Panel Fryslân

De segmentatie in Friesland op basis van economisch-, sociaal-, cultureel- en persoonskapitaal, wordt onderzocht met data van het FSP van hun eigen panel 'Panel Fryslân'. Panel Fryslân is een onderdeel van het FSP en bestaat uit een groep volwassen inwoners van Friesland. In zijn huidige samenstelling bestaat het panel sinds eind 2016 (FSP, 2017).

Werving

In augustus 2016 is het FSP gestart met de werving van het panel. Allereerst zijn er gegevens opgevraagd van 40.000 inwoners van Friesland, gelijk verdeeld over de 24 gemeenten van Friesland. Deze 40.000 zijn uitgezocht met een aselechte steekproef. Dit is 8% van het aantal volwassen inwoners van Friesland (510.000). Het aantal uitnodigingen is berekend op het gewenste minimum aantal leden voor het panel. Wegens afronding naar boven door gemeenten is het uiteindelijk aangeschreven aantal personen 40.059.

De geselecteerde inwoners van Friesland hebben allereerst een aankondigingsbrief gekregen, inclusief een folder waarin het doel en nut van het panel is uitgelegd. Twee weken na de aankondigingsbrief is er een herinneringsbrief verstuurd. Zie de onderzoeksverantwoording op de site van het FSP voor de inhoud van de brieven ([FSP, 2017](#)).

Respons

Van de 40.059 aangeschreven personen, hebben 3750 mensen zich aangemeld bij Panel Fryslân en de intake-vragenlijst ingevuld. Het percentage respons komt daarmee op 9,4%. Ter analyse van de non-respons zijn in de intake-vragenlijst vragen opgenomen waarmee de doelpopulatie en de panelleden met elkaar vergeleken kunnen worden.

De representativiteit van het panel is gemeten aan de hand van leeftijd, geslacht en regio. In tabel 1, staan de metingen. Uit de metingen blijkt dat de verdeling naar regio redelijk gelijk is aan de verdeling van de bevolking binnen Friesland. De verdeling van geslacht is niet

gelijk aan de verdeling binnen de bevolking van Friesland, in het panel zijn mannen oververtegenwoordigd ten opzichte van vrouwen, een mogelijke verklaring hiervoor is de manier van uitnodigen. Een deel van de uitnodigingsbrieven is geadresseerd per huishouden het vermoeden bestaat dat de naam van de man gebruikt is om aan te melden voor Panel Fryslân. Dit is echter een vermoeden en een bewijs is hier niet voor te leveren. De verdeling qua leeftijd is ook ongelijk. In de eerste groep, de groep 18 tot 40 jarigen, is de respons laag. Het aandeel van deze groep is in het panel veel lager dan dat het is in Friesland. Ook de groep 40 tot 50 is ondervertegenwoordigd, maar het verschil is wel minder groot dan bij de eerste groep. Vanaf 50 jaar ontstaat er oververtegenwoordiging. Vooral de groep 60 tot 70 jarigen is sterk oververtegenwoordigd. Dit is een veelgehoord probleem bij het werven van panels (FSP, 2017).

Onderzoekspopulatie

Om segmentatie te onderzoeken wordt er gebruik gemaakt van drie vragenlijsten en informatie van de intakevragenlijsten die zijn afgenomen bij Panel Fryslân. Informatie over economisch- en cultureel kapitaal komt uit de intakevragenlijst en de overige informatie uit de vragenlijsten over participatie (2017), zorg (2018) en duurzaamheid (2018). Niet alle panelleden hebben alle vragenlijsten ingevuld. De panelleden die de betreffende vragenlijsten hebben ingevuld, zijn geselecteerd. Dit levert een totaal op van 1137 respondenten. Dat betekent dat 2613 (3750-1137) panelleden de betreffende vragenlijsten niet compleet hebben ingevuld, hetzij halverwege afgehaakt, hetzij geen eens begonnen aan de vragenlijst.

In tabel 1 staat een beschrijving van de respondenten die in dit onderzoek worden gebruikt. Het blijkt dat met name het aantal jongeren omlaag is gegaan en ouderen omhoog. Qua panelleden is er al sprake van een oververtegenwoordiging van ouderen en in dit onderzoek is de oververtegenwoordiging van ouderen nog groter geworden. De verhouding tussen man en vrouw is wel redelijk gelijk gebleven in dit onderzoek ten opzichte van het panelbestand.

Tabel 1: Verdeling geslacht, regio en leeftijd van panel en doelpopulatie (in procenten).

	Friesland	Panel	Respondenten
Geslacht			
Man	49,8%	58,8%	61,1%
Vrouw	50,2%	41,2%	38,%

Regio			
Noord-Friesland	50,4%	51,6%	
Zuidwest-Friesland	20,8%	21,1%	
Zuidoost-Friesland	28,8%	27,3%	
leeftijd			
18-40	30,6%	12,8%	5,2%
40-50	17,7%	13,3%	8,6%
50-60	18,3%	22,7%	24,7%
60-70	16,7%	30%	32,5%
70+	16,6%	21,2%	28,9%

Latente Klassen Analyse (LCA)

Zoals aangegeven in de analyse-opzet, is latente klassen analyse (*Latent Class Analysis LCA*) een manier om latente groepen te zoeken op basis van manifeste variabelen. LCA is een statistische manier om niet-zichtbare overeenkomsten tussen respondenten bloot te leggen op basis van bestaande variabelen. Overeenkomsten tussen respondenten worden zichtbaar gemaakt door de respondenten toe te wijzen aan klassen op basis van individuele kenmerken. Dat maakt de LCA een geschikte manier om segmentatie te onderzoeken.

LCA werkt als volgt: Op basis van overeenkomsten in de variabelen deelt LCA respondenten in klassen. Respondenten die overeenkomsten met elkaar vertonen worden toegewezen aan dezelfde klasse. Het toewijzen van een persoon aan een klasse gebeurt op basis van de kans van die respondent om bij een klasse te horen. Voor elke klasse wordt per respondent een kans gegeven dat die respondent bij betreffende klasse hoort. Het vormen van klassen krijgt dus betekenis door de verhouding van de klasse ten opzichte van andere klassen.

De LCA begint met twee klassen, daarna onderscheidt het een derde klasse, vervolgens een vierde enzovoort. In principe kan een LCA doorgaan totdat elke klasse nog maar uit twee respondenten bestaat, maar dat is niet interessant. De LCA stopt wanneer de onderzoeker aangeeft dat het juiste aantal klassen is bereikt en dat is allereerst het doel: het aantal klassen te vinden wat het beste bij de data past.

Het juiste aantal groepen gebeurt op basis van het statistische criterium de *Bayesian Information Criterion* (BIC) (Neath & Cavanaugh, 2012; SCP, 2014) en op basis van theoretische interpretatie. Op de statistische manier wordt het aantal groepen bepaald door de

BIC-waarde. Met het toevoegen van een nieuwe groep, wordt de BIC-waarde opnieuw berekend. Als de BIC-waarde daalt door het toevoegen van een extra groep in de LCA, dan geeft een extra groep een betere fit aan de data. Wanneer de BIC-waarde niet meer daalt, dan is statistisch gezien het optimale aantal groepen bereikt. De BIC-waarde is vergelijkbaar met de aangepaste R^2 (R^2_{adj}), een bekende maat voor modelfit in regressieanalyse. De R^2_{adj} daalt wanneer er een variabele wordt toegevoegd die geen bijdrage levert aan de fit van het model. Zo is het ook met de BIC, maar dan andersom; die stijgt wanneer het toevoegen van een extra klasse niets toevoegt aan de modelfit.

Alleen afgaan op de BIC-score is echter onvoldoende. Als het gegeven aantal klassen inhoudelijk weinig betekenis hebben, dan dient er theoretisch bepaald te worden welk aantal klassen het beste bij de data passen. De theoretische interpretatie van de klassen is dus ook van belang.

Wanneer het juiste aantal klassen is bepaald, kunnen de klassen beschreven worden. Het beschrijven van de klassen gebeurt op basis van de kenmerken van de personen in de groepen. De LCA geeft de output van de verdeling van de variabelen op de klassen. Het gaat hierbij om een relatieve score, want hoe hoog of laag een groep scoort, is altijd afhankelijk van de samenstelling van de andere groepen.

Een LCA kan met behulp van verschillende statistische analyse programma's worden uitgevoerd. In dit onderzoek wordt de LCA uitgevoerd met behulp van het statistische programma R (R-core team, 2017). Om een LCA uit te voeren is er gekozen voor het pakket *Polytomous Latent Class Analysis* (poLCA) (Drew & Linzer, 2011).

In onderstaande figuur is een afbeelding te zien van output van een LCA met poLCA. Het doel van dit figuur is om een beeld te schetsen van de werking van de poLCA. De output is afkomstig uit dit onderzoek. In de output is gegeven voor een LCA met drie groepen. In de output is de verdeling te vinden van de variabele inkomen, de geschatte groeps grootte in de populatie en de BIC-waarde. De overige cijfers zijn niet relevant voor dit onderzoek. In de volgende alinea wordt uitgelegd hoe de output gelezen wordt in dit onderzoek.

```

$Ink laag midden hoog weetniet rum
      Pr(1) Pr(2) Pr(3) Pr(4)
class 1: 0.5339 0.1574 0.0356 0.2731
class 2: 0.0631 0.1207 0.6071 0.2091
class 3: 0.1346 0.1312 0.5461 0.1881

Estimated class population shares
0.3295 0.3157 0.3548

Predicted class memberships (by modal posterior prob.)
0.3251 0.3175 0.3573

=====
Fit for 3 latent classes:
=====

number of observations: 1055
number of estimated parameters: 98
residual degrees of freedom: 957
maximum log-likelihood: -10829.06

AIC(3): 21854.13
BIC(3): 22340.33
G^2(3): 7195.106 (Likelihood ratio/deviance statistic)
X^2(3): 720381.8 (Chi-square goodness of fit)

```

Figuur 1: Output van poLCA. LCA heeft 1 variabele en 3 groepen. Variabele (inkomen) heeft vier antwoordcategorieën

De output begint met de verdeling van de variabele inkomen. Deze variabele heeft vier categorieën: Pr(1) laag inkomen; Pr(2) middeninkomen; Pr(3) hoog inkomen; Pr(4) weet niet/wil niet zeggen. Voor elke groep is de verdeling op de vier categorieën te zien. Klasse 1 heeft de hoogste score op laag inkomen. Klasse 2 heeft lage scores op laag en midden inkomen, maar hoog op hoog inkomen. Klasse 3 heeft net wat hogere scores dan groep 2, hoewel de verschillen klein zijn. Uit deze output kan geconcludeerd worden dat de eerste klasse het laagste inkomen heeft, de tweede het hoogst en de derde klasse staat in het midden, maar dichterbij de tweede dan de eerste klasse.

Onder de verdeling van de klassen staat de geschatte klassengrootte. Daaruit blijkt dat eerste klasse een geschatte groeps grootte heeft van 33%, klasse 2 een geschatte klassengrootte van 31% en groep 3 een geschatte groeps grootte heeft van 36%. Dit kan gezien als groepen met ongeveer een gelijke omvang.

Eerder is aangegeven dat de BIC-waarde een belangrijk criterium is om het aantal klassen te bepalen. Zolang de BIC-waarde daalt, heeft het zin om meer klassen te produceren en wanneer de BIC-waarde stijgt, zou het juiste aantal klassen bereikt zijn. De BIC-waarde in dit model is 22340,33.

Voordat de LCA uitgevoerd kan worden, is het nodig dat sommige variabelen bewerkt worden om te voldoen aan de voorwaarden om een LCA met de poLCA uit te voeren. Zo maakt de poLCA gebruik van categorische variabelen. Bij het maken van schalen moet er dus opgepast worden dat ze niet berekend worden op gemiddelde, iets wat wel gebruikelijk is bij het vormen van schalen.

Daarnaast dient de interpretatie van de variabelen ook ‘eenvoudig’ te zijn. In figuur 1 is een voorbeeld gegeven met één variabele en drie groepen. Om segmentatie in Friesland te onderzoeken worden er meer variabelen gebruikt en zal het aantal groepen misschien meer zijn dan drie. Dit maakt het interpreteren van de LCA meer complex, aangezien de interpretatie van de LCA een beschrijvende analyse is; hoe meer variabelen en groepen, hoe moeilijker de analyse. Concreet betekent dit dat een variabelen niet teveel categorieën moet hebben.

Wegen

Het doel van wegen van de data is om de steekproef representatief te maken ten opzichte van de doelpopulatie, in dit geval de inwoners van Friesland. Wegen heeft als principe om respondenten die ondervertegenwoordigd zijn in de steekproef, zwaarder mee te laten wegen en respondenten die oververtegenwoordigd, minder zwaar te laten meetellen. Het doel hiervan is om uitspraken te doen met de steekproef over de doelpopulatie: inwoners van Friesland.

In een normale gang van zaken wordt een weegfactor toegevoegd aan de analyse, om te compenseren voor de scheve verdeling van de steekproef ten opzichte van de doelpopulatie. Echter, bij de poLCA is dat niet mogelijk, dus wordt er op een andere manier gecorrigeerd voor de scheve verdeling van Panel Fryslân ten opzichte van Friesland.

Hoewel het dus niet mogelijk is om een weegfactor toe te voegen aan de LCA, is er wel een alternatieve manier. In plaats van het toevoegen van een weegfactor tijdens de LCA, wordt de dataset aangepast. Daarmee vindt de weging plaats vóór de uitvoering van de LCA. Om de respondenten in de steekproef te ‘wegen’ maakt het FSP gebruik van cijfers over geslacht, leeftijd en opleiding van de inwoners van Friesland. In tabel 1 staat de betreffende verdeling van de inwoners van Friesland ten opzichte van de steekproef.

Het wegen van de respondenten is wordt gedaan door elke respondent te vermenigvuldigen met haar eigen weegfactor. De originele weegfactor is een getal wat eindigt op decimalen. Dit getal is teruggebracht tot een getal met twee decimalen. Dit levert een

weegfactor tussen 0,21 en 12,97 met een gemiddelde van 1,00. De weegfactor wordt met 100 vermenigvuldigd, waardoor er een weegfactor ontstaat tussen 21 en 1297. De reden om voor 100 te kiezen is de keuze tussen informatiebehoud en werkbare dataset. Vermenigvuldigen met 10 zorgt voor informatieverlies wegens afronding en vermenigvuldigen met 1000 geeft een dermate grote dataset waarmee de software moeite heeft om mee te rekenen.

Na het vermenigvuldigen van de weegfactor wordt elke respondent vermenigvuldigd met de bijbehorende weegfactor, zodat de dataset wordt vergroot in verhouding tot de weegfactor. Dit betekent dat de dataset extra veel rijen krijgt. Dus een respondent met de originele weegfactor 0,21 telt 21 keer mee in de gewogen dataset, een respondent met de originele weegfactor telt 1297 keer mee en de gemiddelde respondent telt 100 keer mee.

Door deze manier van wegen ontstaat er een dataset met 111.662 respondenten. Hoewel deze dataset dezelfde verhoudingen heeft als met een weegfactor, is de gewogen dataset niet geschikt om het aantal groepen te bepalen, want het aantal groepen hangt af van de informatie in de dataset en die informatie is verhonderdvoudigd. De gewogen data wordt daarom enkel gebruikt om de groepen te vergelijken met de originele data. Dit betekent dat er geen modelselectie plaatsvindt met de gewogen data, maar dat het aantal klassen wordt aangehouden wat ook wordt gevonden in de modelselectie met de originele data.

Operationalisatie

In de operationalisatie worden de variabelen beschreven waarmee de vier vormen van kapitaal gemeten worden. De variabelen zijn per vorm van kapitaal behandeld. Per vorm van kapitaal wordt ook een beschrijving gegeven van de verdeling van de betreffende variabelen.

Aangezien dit onderzoek voortbouwt op het onderzoek van het SCP, is in eerste instantie gepoogd om dezelfde variabele te gebruiken als dat het SCP doet. Echter, een gelijke variabele is vaak niet aanwezig in de dataset. In dat geval wordt zo goed als mogelijk een gelijkwaardige variabele gebruikt. Per variabele en vorm van kapitaal wordt dat beschreven in de operationalisatie.

Economisch kapitaal

Het SCP meet het economisch kapitaal met de indicatoren opleiding, arbeidssituatie en inkomen. Deze drie indicatoren zijn ook gemeten door het FSP. Alle drie de vragen om economisch kapitaal te meten zijn door het FSP gecodeerd naar drie groepen die sterk op elkaar lijken. Echter, de drie groepen zijn niet met elkaar vergelijkbaar. De categorieën van arbeidssituatie hebben een andere betekenis dan de categorieën van opleiding. Bijvoorbeeld:

hoogopgeleid, wat staat voor veel economisch kapitaal, is niet vergelijkbaar met de status van werkend.

Hoogst voltooide opleiding

Het opleidingsniveau van de respondent is in twee vragenlijsten gevraagd. De reden daarvan was de onduidelijke vraagstelling in de intakevragenlijst. Door het FSP zijn de twee variabelen waarmee het opleidingsniveau gemeten wordt, samengevoegd tot één variabele. Het FSP werkt met zeven opleidingsniveaus die zijn samengevoegd in drie categorieën: laag, midden en hoog. In de laagopgeleide categorie zitten de volgende opleidingsniveaus: basisonderwijs; praktijkonderwijs; vmbo, leerwegondersteunend onderwijs (lwo) of voortgezet speciaal onderwijs (VSO). (inclusief ulo, mulo, lts, huishoudschool, lhno en vbo). In de midden categorie zitten: havo, vwo (inclusief vierjarige mms (VMS), hbs, 5 jarige mms); mbo opleiding. De laatste twee niveaus zijn de hoogopgeleiden en dat zijn: hbo-opleiding; universitaire opleiding. Deze indeling is gebaseerd op de Standaard Onderwijsindeling (CBS, 2016). Naast dat de Standaard Onderwijsindeling drie opleidingsniveaus aangeeft, komen drie categorieën in plaats van zeven de interpretatie van de resultaten ten goede.

Arbeidssituatie

De arbeidssituatie van de respondent is in de intakevragenlijst gevraagd: ‘welke van onderstaande situaties past het best bij mij’. De respondent had de keuze uit de volgende 8 categorieën: 1) huisvrouw / huisman / anders. 2) arbeidsongeschikt. 3) werkloos / werkzoekend / bijstand. 4) gepensioneerd of VUT. 5) studerend / schoolgaand. 6) werkzaam bij commercieel bedrijf. 7) werkzaam bij de overheid. 8) zelfstandiger ondernemer.

Deze 8 categorieën zijn door het FSP gecodeerd naar een variabele met drie categorieën: werkend; niet-werkend; pensioen. Werkend bestaat uit de arbeidssituaties: werkzaam bij commercieel bedrijf; werkzaam bij overheid; zelfstandig ondernemer. Niet werkend bestaat uit de arbeidssituaties: huisvrouw/huisman/anders; arbeidsongeschikt; werkloos/werkzoekend/bijstand; studerend/schoolgaand. Gepensioneerd bestaat uit de arbeidssituatie: gepensioneerd of VUT.

Bruto jaarinkomen van huishouden

Het inkomen van de respondent is in de intakevragenlijst gemeten als het bruto jaarinkomen van het gezin. De respondent had de keuze uit de volgende 11 categorieën: < 14.100; 14.100 -

< 19.200; 19.200 - < 26.300; 26.300 - < 29.500; 29.500 - < 36.500; 36.500 - < 43.500; 43.500 - < 57.600; 57.600 - < 73.300; 73.300 - < 87.100; >- 87.100; weet ik niet/ wil ik niet zeggen. Het jaarinkomen van de respondent is gecodeerd naar een variabele van 4 categorieën: beneden modaal; modaal; boven modaal en weet niet/wil niet zeggen. Het modale inkomen in Nederland was in 2017 37.000 euro (CPB, 2017).

Tabel 2: Verdeling van economisch kapitaal. Verdeling van variabelen gegeven in percentages (N=1134; ongewogen)

<i>Variabele</i>	<i>Categorie</i>			
	<i>Werkend</i>	<i>Niet-werkend</i>	<i>Gepensioneerd</i>	
<i>Arbeidssituatie</i>	45,1%	13,5%	41,4%	
<i>Opleiding</i>	<i>Laag</i>	<i>Midden</i>	<i>Hoog</i>	
	15,4%	32,4%	52,2%	
<i>Inkomen</i>	<i>Beneden modaal</i>	<i>Modaal</i>	<i>Boven modaal</i>	<i>Weet niet/wil niet zeggen</i>
	24,1%	13,2%	39,9%	22,7%

In tabel 2 staat de verdeling van de drie variabelen die het economisch kapitaal van de respondenten meet. Uit de arbeidssituatie blijkt dat het grootste gedeelte van de respondenten werkend is. Daarnaast zijn er relatief veel hoge inkomens en boven modale inkomens.

Sociaal kapitaal

Het SCP meet sociaal kapitaal met twee indicatoren: sociale steun en instrumentele steun. Met de data van het FSP is het niet mogelijk om instrumentele steun te meten, maar wel kan sociale steun gemeten worden. Sociale steun wordt gemeten aan de hand van drie schaalvariabelen waarmee de kwantiteit en kwaliteit van sociale contacten wordt gemeten. Door het missen van instrumentale steun wordt sociaal kapitaal niet volledig gemeten.

Kwantiteit van sociale steun

Kwantiteit van sociale steun wordt gemeten met twee schaalvariabelen. In de eerste schaalvariabele wordt gevraagd naar de hoeveelheid fysieke contacten met familie, burens en

vrienden. In de tweede schaalvariabele wordt gevraagd naar de hoeveelheid digitale contacten met familie, burens en vrienden.

De hoeveelheid fysieke contacten is gevraagd als: ‘Hoe vaak ontmoet u de volgende personen?’. Hierbij had de respondent keuze uit de volgende antwoordmogelijkheden: (6) (bijna) dagelijks; (5) 1x per week of vaker; (4) 2 tot 3x per week; (3) 1 keer per maand; (2) minder dan 1x per maand; (1) zelden en of nooit; (0) niet van toepassing. Dit is ingevuld voor drie groepen sociale contacten: familieleden; burens en goede vrienden; kennissen of goede kennissen. Dat betekent dat er drie items zijn. De categorie ‘niet van toepassing’ is gecodeerd naar missende waarden.

De hoeveelheid digitale contacten is gevraagd als: ‘Hoe vaak heeft u contact via internet en/of mobiele telefoon met de volgende personen?’. Hierbij had de respondent keuze uit de volgende antwoordmogelijkheden: (bijna) dagelijks; 1x per week of vaker; 2 tot 3x per week; 1 keer per maand; minder dan 1x per maand; zelden en of nooit; niet van toepassing. Ook voor deze variabele is dit ingevuld voor drie groepen sociale contacten: familieleden; burens en goede vrienden; kennissen of goede kennissen. Dat betekent dat er drie items zijn. De categorie ‘niet van toepassing’ is gecodeerd naar missende waarden.

Deze twee variabelen hebben dezelfde antwoordmogelijkheden variabelen. Daarom worden ze beide op dezelfde manier bewerkt, het blijven echter wel twee aparte variabelen waarbij de één fysiek contact (Fysiekcontact_schaal) meet en de andere digitaal contact (Digitaalcontact_schaal). Va elke variabele apart zijn de items bij elkaar opgeteld. Dit leidt bij beide variabelen tot een scoreverloop van 3 tot 18. Vervolgens zijn beide variabelen opgedeeld in mate van contact. Hierbij is de indeling gevolgd die het SCP ook heeft gevolgd (2014). Dit is op de volgende manier gedaan:

Tabel 3: Nieuwe indeling van schaal kwantiteit van sociale contacten (geldt voor fysiek en digitaal)

<i>Oude waarde</i>	<i>Nieuwe waarde</i>
0 – 9	1 zeer weinig contact
10 – 12	2 weinig contact
13 – 15	3 geregeld contact
16 – 17	4 veel contact
18	5 zeer veel contact

Door het categoriseren duidt een hoge score op een hoge kwantiteit van sociale contacten. Echter, door het categoriseren is het niet meer mogelijk om precies aan te geven hoeveel sociale contacten iemand heeft en met wie die sociale contacten zijn. Bij het interpreteren van de resultaten dient daar rekening mee worden gehouden.

Kwaliteit van sociale steun

Kwaliteit van sociale contacten is gemeten met 5 items. De items zijn een Likertschaal van zeer oneens tot zeer eens. De respondent kreeg de volgende stellingen: 1) er zijn mensen met wie ik echt goed kan praten; 2) er zijn mensen bij wie ik terecht kan; 3) er zijn mensen die me echt begrijpen; 4) ik maak deel uit van een groep vrienden; 5) mijn sociale contacten zijn oppervlakkig.

Met de kwaliteit van sociale contacten is het de bedoeling om te meten in hoeverre een respondent sociale steun heeft van zijn sociale contacten. Een hoge score dient aan te geven dat de respondent veel sociale contacten heeft. Het item ‘mijn sociale contacten zijn oppervlakkig’ staat in negatieve richting en daarom is dat item gespiegeld. De schaalscore van kwaliteit van sociale steun toont voldoende betrouwbaarheid om samen te voegen ($\alpha = 0,79$). De schaal is gevormd door de items bij elkaar op te tellen. Dit leidt tot een range van 5 tot 25. Net als kwantiteit van sociale contacten, wordt ook kwaliteit van sociale contacten gecategoriseerd. De nieuwe variabele (Kwalcontact_schaal) heeft een scoreverloop van 1 tot en met 5, waarbij een hogere categorie staat voor betere kwaliteit van sociale contacten.

Tabel 4: Verdeling kwantiteit en kwaliteit van sociale contacten. Verdeling van categorieën wordt gegeven in percentages. Categorieën zijn van laag (1) naar hoog (5). (N=1134; ongewogen)

variabele	categorie				
	1	2	3	4	5
Fysieke contacten	14,4%	32,5%	45,1%	7,9%	,2%
Digitale contacten	32,4%	39,9%	22%	5,4%	,4%
Kwaliteit sociale contacten	1,1%	79,5%	97,5%	52,8%	22,4%

In tabel 4 staan de variabelen van sociaal kapitaal. De range van alle drie de variabelen loopt van 1 tot 5. Het gemiddelde van fysieke contacten is hoger dan het gemiddelde van digitale contacten. Dat betekent dat de respondenten meer fysieke contacten hebben dan digitale contacten.

Cultureel kapitaal

Het SCP meet cultureel kapitaal met drie indicatoren: Engelse taal en communicatie; smaken, voorkeuren en culturele kennis; symbolische aspecten. Met de data die beschikbaar is bij het FSP is het mogelijk om cultureel kapitaal te meten aan de hand van taalbeheersing. In dit onderzoek wordt gekozen voor Friese taalbeheersing. Friesland heeft een eigen taal die door 60% van de Friezen wordt gesproken. Daarom wordt de beheersing van de Friese taal gebruikt om het cultureel kapitaal te meten. Opmerking hierbij is wel dat taal een onderdeel is van cultureel kapitaal en hierdoor het cultureel kapitaal van de inwoners van Friesland niet volledig gemeten wordt.

Beheersing van taal is als volgt gevraagd: Spreekt u Fries? De antwoordcategorieën waren: helemaal niet; matig; redelijk; goed; vloeiend. Deze variabele hoeft niet bewerkt te worden om er analyses mee te doen.

Tabel 5 beheersing van Friese taal (N=1134; ongewogen)

<i>Variabele</i>	<i>Categorie</i>				
	<i>Helemaal niet</i>	<i>Matig</i>	<i>Redelijk</i>	<i>Goed</i>	<i>Vloeiend</i>
Fries sprekend	14,6%	14,0%	8,3%	7,3%	55,7%

In tabel 5 wordt beschreven in welke mate de respondenten de Friese taal beheersen. Het grootste gedeelte van de respondenten spreekt de Friese taal. Een relatief groot gedeelte geeft aan de Friese taal helemaal niet tot matig te beheersen. Dit is in overeenstemming met het gegeven dat 60% van de inwoners van Friesland Fries spreekt (“[Taal in Nederland]”, z.d.).

Persoonskapitaal

Het SCP meet persoonskapitaal met drie indicatoren: fysiek-, mentaal- en esthetisch persoonskapitaal. Het FSP heeft geen data over esthetisch persoonskapitaal, maar wel is het mogelijk om fysiek en mentaal kapitaal te meten.

Fysiek kapitaal

Het fysieke persoonskapitaal is gemeten met de vraag: hoe ervaart u over het algemeen uw gezondheid? De antwoordmogelijkheden waren: zeer slecht; slecht; niet slecht en niet goed; goed; zeer goed. Deze vraag heeft geen missende waarden en hoeft niet bewerkt te worden.

Mentaal kapitaal

Het SCP meet het mentale persoonskapitaal met het zelfvertrouwen van de respondent. Dit is echter niet gemeten door het FSP en daarom wordt er voor de Pearlin mastery scale (1978) gekozen als proxy van zelfvertrouwen. Deze schaal meet in hoeverre iemand het gevoel heeft controle over eigen leven te hebben. In dit onderzoek wordt deze variabele dan ook ‘zelfregie’ genoemd. Voor mentaal persoonskapitaal geldt dat zelfregie niet optimaal is om mentaal persoonskapitaal te meten. Vanuit de literatuur wordt persoonlijk kapitaal gemeten als mate van zelfvertrouwen (CBS, 2016; SCP, 2014; Tafarodi & Swann, 2001; Hupert et al., 2009). Er bestaan vermoedens dat zelfvertrouwen sterk afhangt van de persoon en minder afhankelijk is van externe factoren. Vanuit die redenering is zelfvertrouwen geschikter om mentaal persoonskapitaal te meten dan zelfregie, omdat controle over eigen leven meer kan afhangen van externe factoren.

De Pearlin mastery scale bestaat uit 7 items met antwoordmogelijkheden op een Likertschaal van zeer oneens tot zeer eens. 1) Ik heb weinig controle over de dingen die me overkomen; 2) Sommige van mijn problemen kan ik met geen mogelijkheid oplossen; 3) Er is weinig dat ik kan doen om belangrijke dingen in mijn leven te veranderen; 4) Ik voel me vaak hulpeloos bij het omgaan met de problemen van het leven; 5) Soms voel ik dat ik een speelbal van het leven ben; 6) Ik kan ongeveer alles als ik mijn zinnen erop gezet heb; 7) Wat in de toekomst gebeurt, hangt voor het grootste deel van mezelf af.

De eerste vijf items staan in negatieve richting. Dat wil zeggen: een hoge score betekent dat een individu weinig controle ervaart over zijn leven. Om controle over eigen leven te meten, dienen antwoorden in positieve richting te staan. Daarom zijn de eerste 5 items gespiegeld. Om een schaal te creëren, zijn de items samengevoegd (Cronbach's $\alpha = 0,81$).

Net als de andere drie schalen, zijn ook deze items bij elkaar opgeteld. Dit geeft een scoreverloop van 7 t/m 35. Vervolgens is dit gecategoriseerd tot een schaal (Zelfregie_schaal) van 1 t/m 7, waar 1 staat voor weinig mentaal kapitaal en 7 voor veel mentaal kapitaal. Het categoriseren van de data is gedaan om de interpretatie van de resultaten te vereenvoudigen.

Tabel 6: Verdeling van ervaren gezondheid en zelfregie. Verdeling van categorieën wordt gegeven in percentages. Categorieën zijn van laag (1) naar hoog (6) (N=1134; ongewogen)

<i>variabele</i>	<i>categorie</i>					
	1	2	3	4	5	6
Ervaren gezondheid	0,2%	3,6%	18,2%	66,5%	11,6	
Zelfregie	0,2%	3,9%	16,2%	38,4%	34,7	6,7%

Resultaten

Zoals vermeld in de Latente Klassen Analyse, wordt in de resultaten wordt eerst gezocht naar het aantal klassen dat het best bij de data past. Daarna worden die klassen beschreven aan de hand van de output van de LCA. Vervolgens zal datzelfde aantal klassen onderzocht worden met een LCA met de gewogen data. Tot slot wordt geanalyseerd welke kapitaalvormen of onderliggende dimensie, het meest invloed hebben op de segmentatie in Friesland.

Aantal klassen

Het aantal klassen in de LCA wordt zowel statistisch als theoretisch bepaald. Op de statistische wijze wordt met de BIC-waarde onderzocht welk aantal klassen het beste is om de segmentatie in de Friese samenleving te verklaren. Vervolgens wordt beredeneerd of dit aantal klassen bij de data past.

Met de BIC-waarde wordt gezocht naar het aantal klassen met de laagste BIC-waarde. In tabel 7 staan de BIC-waarden van twee tot en met vijf klassen. Het blijkt dat de een LCA met drie klassen de laagste BIC-waarde heeft. Deze waarde stijgt bij vier klassen en stijgt sterker bij een LCA met vijf klassen. Dit geeft te kennen dat een LCA met drie klassen statistisch gezien het beste de data past. Echter, theoretisch gezien hoeft dit nog niet het beste te zijn.

Tabel 7 BIC-waarde van 2, 3, 4 & 5 klassen

	<i>BIC-waarde</i>
<i>Klassen</i>	
2	22350
3	22340
4	22451
6	22595

Bij de LCA met drie klassen ontstaat er een duidelijk onderscheid tussen drie klassen. Er ontstaat een groep die het minst kapitaalkrachtig is van alle drie de groepen, met een opvallend hoog aantal gepensioneerden. Een tweede groep heeft vooral economisch kapitaal en ook deze groep bevat veel gepensioneerden. Tot slot ontstaat er een groep met veel persoons- en sociaal kapitaal. Deze groep bestaat vooral uit werkenden.

De LCA met drie klassen wordt opgesplitst naar een LCA met vier klassen om te kijken of er een uitsplitsing wordt gemaakt wat betreft economisch kapitaal. In de LCA met vier klassen blijft er één klasse gelijk, dat is de klasse die het minst kapitaalkrachtig is. De andere klassen nemen andere vormen aan. Zo ontstaat er een klasse die veel sociaal- en cultureel kapitaal heeft, een klasse met vooral veel economisch kapitaal en veel gepensioneerden en tot slot ontstaat er een klasse met veel persoons, sociaal en economisch kapitaal.

Ten opzichte van de drie klassen, heeft de LCA met vier klassen een groter onderscheid tussen economisch kapitaal en zorgt voor meer specifieke verschillen tussen de klassen. Een LCA met vijf klassen is volgens de BIC-waarde niet optimaal en ook theoretisch gezien is dit geen goed aantal. Wanneer er nog een klasse wordt toegevoegd aan de LCA wordt het onderscheid tussen gepensioneerden verder opgerekt. Echter, gepensioneerden zijn oververtegenwoordigd in data en daarom zou de vijfde groep niet representatief voor Friesland zijn. In tabel 8 is de LCA uitgewerkt voor drie en vier klassen. Aan de hand van deze tabel zijn groepen beschreven in de volgende paragraaf.

Tabel 8 LCA voor drie klassen en vier klassen. In de rijen staan de variabelen en de in de kolommen de klassen.

Variabele		Klasse						
		1	2	3	1	2	3	4
zelfregie	1	,00	,00	,00	,00	,00	,00	,00
	2	,09	,02	,01	,09	,03	,02	,01
	3	,30	,18	,04	,24	,25	,23	,00
	4	,44	,39	,04	,45	,46	,39	,29
	5	,16	,35	,50	,19	,21	,31	,58
	6	,01	,06	,12	,01	,05	,05	,13
gezondheid	1	,01	,00	,00	,01	,00	,00	,00
	2	,07	,03	,02	,07	,03	,03	,01
	3	,25	,20	,11	,24	,16	,21	,13
	4	,65	,67	,69	,64	,79	,68	,62
	5	,04	,09	,18	,06	,02	,08	,24
Kwaliteit	1	,02	,02	,00	,03	,00	,02	,01
	2	,12	,12	,00	,14	,02	,14	,00
Van contacten	3	,45	,46	,10	,50	,21	,50	,12
	4	,38	,32	,56	,31	,60	,31	,50

	5	,04	,09	,33	,02	,16	,03	,38
Digitale	1	,45	,53	,08	,60	,04	,54	,15
Sociale	2	,35	,43	,40	,33	,36	,42	,45
Contacten	3	,14	,04	,43	,07	,44	,04	,34
	4	,07	,00	,08	,01	,17	,00	,05
	5	,00	,00	,01	,00	,00	,00	,01
Fysieke	1	,16	,29	,12	,21	,00	,31	,06
Sociale	2	,32	,44	,23	,42	,10	,44	,30
contacten	3	,43	,26	,64	,33	,65	,24	,59
	4	,10	,01	,12	,04	,25	,01	,06
	5	,00	,00	,00	,00	,00	,00	,00
Friessprekend	1	,13	,19	,12	,12	,12	,21	,14
	2	,09	,20	,14	,11	,06	,09	,12
	3	,04	,09	,12	,04	,08	,09	,12
	4	,10	,05	,07	,13	,03	,04	,09
	5	,64	,48	,54	,61	,71	,49	,47
opleiding	1	,40	,03	,06	,36	,28	,03	,00
	2	,16	,21	,35	,37	,44	,22	,30
	3	,20	,76	,59	,26	,28	,75	,70
arbeidsstatus	1	,21	,49	,63	,20	,36	,48	,69
	2	,28	,02	,11	,28	,20	,01	,07
	3	,51	,49	,26	,52	,44	,51	,24
inkomen	1	,53	,06	,13	,57	,35	,00	,09
	2	,16	,12	,13	,17	,14	,13	,11
	3	,04	,61	,55	,04	,19	,63	,66
	4	,27	,21	,19	,22	,33	,24	,14
Geschatte klassengrote		33%	32%	35%	26%	20%	24%	30%

Beschrijving van de klassen

Klasse 1:

De eerste klasse beschikt over weinig sociaal-, persoons- en economisch kapitaal. De beheersing van de Friese taal is goed. Op de overige variabelen heeft deze klasse een lage score. Zo is deze klasse het laagst opgeleid en heeft het de laagste inkomens. Omdat deze klasse niet over veel hulpbronnen beschikt, wordt deze klasse het 'Precariaat' genoemd. De naam precariaat is afkomstig van het SCP (2014). De vergelijkbare klasse in hun onderzoek noemen zij ook precariaat en vertoont veel overeenkomsten met de klasse in dit onderzoek. Het precariaat bestaat vooral uit werklozen en ouderen en heeft een aandeel van 26,8% in Friesland.

Klasse 2:

De tweede klasse heeft opvallend hoge scores op sociaal- en cultureel kapitaal. Deze klasse heeft de beste beheersing van de Friese taal en op het gebied van sociaal kapitaal zijn ze een uitschieter. Het persoonskapitaal wordt van deze klasse vooral gevormd door de hoge ervaren gezondheid. Het inkomen is laag en opleiding gemiddeld. Omdat deze klasse een uitschieter hebben op sociaal kapitaal, noem ik deze klasse: Socialen. Het aandeel van de socialen in de Friese samenleving is 19,5%.

Klasse 3:

De derde klasse heeft vooral de beschikking over economisch kapitaal, maar weinig van de andere vormen van kapitaal. Deze klasse bestaat hoofdzakelijk uit gepensioneerden, maar ook relatief veel werkenden. Opvallend aan deze klasse is de lage mate van sociaal kapitaal, net als het precariaat, dit doet vermoeden dat vooral ouderen weinig sociaal kapitaal hebben. De derde groep is het hoogst opgeleid van alle klassen. Vanwege het hoge aandeel gepensioneerden en hun economisch kapitaal, noem ik deze groep: rijk gepensioneerden. Het aandeel van rijk gepensioneerden in de Friese samenleving is 24%.

Klasse 4:

De vierde klasse heeft beschikking over veel persoons-, sociaal- en economisch kapitaal, kortom: dit lijkt de klasse die 'het heeft'. Deze klasse heeft het hoogste inkomen en bestaat vooral uit werkenden. Deze klasse heeft de hoogste scores op persoonskapitaal, zowel ervaren gezondheid als zelfregie. Qua beheersing van het Fries heeft deze klasse de laagste score.

Deze klasse heeft op meerdere vlakken een uitschieter, maar vanwege het hoge aandeel werkenden, noem ik deze klasse: werkenden. Het aandeel werkenden in Friesland is 30%

Latente klassen analyse met gewogen data

Uit de LCA met de originele data bleek dat vier klassen het beste bij de data passen. De LCA met gewogen data wordt daarom ook met vier klassen uitgevoerd. Dit is de enige manier waarop de gewone LCA met gewogen data kan worden vergeleken.

Klasse 1: Groep 1 is een groep met weinig economisch, sociaal en persoonskapitaal. Het belangrijkste kenmerk van deze klasse is het hoge aantal gepensioneerden. Deze klasse heeft sterke overeenkomsten met het precariaat. Daarnaast heeft klasse 1 een hoge beheersing van de Friese taal. Op de overige variabelen zijn de scores laag. Het sociaal kapitaal van deze groep wordt vooral gevormd door het aantal fysieke contacten en op persoonskapitaal heeft deze groep samen met klasse 3 de laagste scores. Het meest onderscheidt deze klasse zich op economisch kapitaal: het aantal laagopgeleiden is bij deze klasse opvallend hoog en het inkomen is het laagst van alle vier de klassen. Klasse 1 heeft een omvang van 32,1% in de onderzoekspopulatie.

Klasse 2: Klasse 2 onderscheidt zich het meest op sociaal en persoonskapitaal. Zo heeft klasse 2 een hoge mate van zelfregie en heeft het de hoogste ervaren gezondheid. Op het gebied van sociaal kapitaal heeft groep 2 vooral hoge kwaliteit van contacten en veel fysieke contacten. Een belangrijk kenmerk van klasse 2 is het hoge aantal werkenden, samen met klasse 4 heeft deze klasse verreweg het hoogste aantal werkenden. Het opleidingsniveau is vooral gemiddeld en die gemiddelde score geldt ook voor het inkomen. Klasse 2 heeft een aandeel van 41% in de onderzoekspopulatie.

Klasse 3: Klasse 3 is lastig te duiden door de vreemde verhoudingen. Het opvallendste is het kleine aandeel van deze klasse in de LCA: 1,3%. Het persoonskapitaal van klasse 3 is laag, net als het sociaal kapitaal. De beheersing van de Friese taal is middelmatig. Het economisch kapitaal van klasse 3 is opvallend. Zo is de klasse hoogopgeleid en bijna helemaal werkloos. Het inkomen van klasse 3 is verdeeld over laag inkomen en hoog inkomen.

Klasse 4: klasse 4 is een redelijk kapitaalkrachtige groep. Het persoonskapitaal van deze groep komt sterk overeen met het persoonskapitaal van klasse 2. Het sociaal kapitaal van deze

klasse is verdeeld; de kwaliteit van contacten is hoog, maar de kwantiteit van contacten is laag. De beheersing van de Friese taal is laag. Het economisch kapitaal van klasse 4 is het hoogst van alle vier de klassen. Klasse 4 is vooral hoogopgeleid en bestaat vooral uit werkenden. Deze werkenden hebben het hoogste inkomen van alle vier de groepen. Groep 4 heeft een omvang van 24.7% in de onderzoekspopulatie.

Vergelijking tussen ongewogen en gewogen data

De vergelijking tussen de beide LCA's is vooral om inzicht te geven hoe de klassen veranderen wanneer de respondenten in de dataset een betere afspiegeling zijn van Friesland. De conclusies die voortkomen uit de gewogen LCA kunnen niet worden doorgetrokken naar Friesland, maar is dus vooral om inzicht te bieden in veranderingen die optreden wanneer de LCA gedaan wordt met data wat wel een afspiegeling is van Friesland.

Tussen de ongewogen en gewogen LCA zijn enkele opvallende verschillen. Het grootste verschil is de grootte van klasse 3 in de gewogen LCA. Door het aandeel van 1,3% in de LCA is deze klasse niet te vergelijken met de ongewogen LCA. Qua kenmerken is deze groep ook lastig te plaatsen. Door het wegvallen van een klasse, is er geen klasse meer over met vooral gepensioneerden met een hoog inkomen en hoge opleiding. Er kan gesteld worden dat de LCA met vier klassen niet optimaal is met de gewogen data.

De meeste gepensioneerden zijn in de gewogen LCA te vinden in klasse 1; de klasse die sterk overeenkomt met het precariaat van de originele data. Door het wegen van de data blijkt dus dat gepensioneerden grotendeels wegvallen. Dit is geen verassing, aangezien gepensioneerden sterk oververtegenwoordigd zijn in de data.

In de LCA met gewogen data blijven er nog twee klassen over: 2 en 4. Deze twee klassen komen in de gewone LCA sterk overeen met de socialen en de werkenden met de originele data. Met het wegen van de data, is het onderscheid tussen de twee klassen veranderd. Met de gewogen data scoort klasse 2 hoog op sociaal- en persoonskapitaal en klasse 4 hoog op economisch kapitaal. In de gewone LCA is het vooral de klasse werkenden die het meest kapitaalkrchtig is, op zowel sociaal,- persoons- en economisch kapitaal.

Invloed van onderliggende dimensies

De invloed van economisch, sociaal, cultureel en persoonskapitaal of onderliggende dimensie wordt beoordeeld door de verdeling van de klassen op de variabelen te analyseren. Dit wordt

gedaan met de originele, ongewogen data. Deze scores staan in tabel 8. Wanneer de klassen sterk verdeeld zijn over een variabele, dat wil zeggen: de scores van de klasse binnen de variabele liggen sterk uiteen, dan wil dat zeggen dat de variabele sterke invloed heeft op de LCA. Andersom geldt ook dat wanneer de scores van de klassen zijn geconcentreerd rond één punt bij de variabele, dan is er weinig onderscheid tussen de klassen en heeft het dus weinig invloed op de LCA.

De invloed van persoonskapitaal op de groepsvorming in de LCA wordt vooral veroorzaakt door zelfregie en weinig door de ervaren gezondheid. De waarschijnlijkheidsscores voor ervaren gezondheid zijn geconcentreerd richting de maximale score, en dus een goede ervaren gezondheid. Zelfregie daarentegen heeft wel spreiding in het scoreverloop en heeft derhalve duidelijk verschillen tussen de klassen.

Het sociaal kapitaal bestaat uit kwantiteit en kwaliteit van sociale contacten. Van de kwantiteit van sociale contacten heeft fysiek contact het meeste spreiding, hoewel beide variabelen geconcentreerd zijn richting minimale kwantiteit van sociale contacten. De kwaliteit van sociale contacten heeft meer spreiding en duidelijk is ook het verschil tussen de klassen te zien.

Met de beheersing van de Friese taal wordt cultureel kapitaal gemeten. Het scoreverloop op de beheersing van de Friese taal is weinig verspreid. Alle vier de klassen hebben een redelijk gelijke score op de beheersing van de Friese taal. Het cultureel kapitaal lijkt weinig invloed te hebben op de LCA.

Economisch kapitaal lijkt de meeste invloed te hebben op de vorming van de klassen. Economisch kapitaal wordt gemeten met het inkomen, arbeidssituatie en opleidingsniveau. Bij alle drie de variabelen is duidelijk te zien welke klasse het beste scoort en welke klasse het minste en ook is er spreiding over de gehele variabele. Geen één van de variabelen springt eruit wat betreft de invloed op de groepsvorming.

Conclusie

Dit onderzoek geeft een aanzet tot het antwoord op de vraag hoe Friesland is gesegmenteerd en welke vormen van kapitaal of onderliggende dimensies daar het meest invloed op hebben. Het doel van dit onderzoek was om een eerste aanzet te geven tot het onderzoeken van de segmentatie in de Friese samenleving op basis van economisch-, sociaal-, cultureel- en persoonskapitaal én hoe zicht dat verhoudt tot de landelijke situatie. Het onderzoek naar segmentatie in Friesland is geïnspireerd op onderzoek van het SCP naar segmentatie in de Nederlandse samenleving (SCP, 2014). De opdrachtgever van dit onderzoek is het Fries Sociaal Planbureau (FSP).

Volgens dit onderzoek is Friesland gesegmenteerd in vier segmenten: het precariaat, socialen, rijk gepensioneerden en werken. Het precariaat (26,8%) is het segment die het minst kapitaalkrchtig is, met als uitzondering de beheersing van de Friese taal. Een relatief groot deel van het precariaat is met pensioen of is werkloos. De socialen (19,5%) hebben een opvallend veel sociaal- en cultureel kapitaal. Het inkomen van de socialen is redelijk laag en opleiding is gemiddeld. De rijk gepensioneerden (24%) hebben beschikking over een hoge mate van economisch kapitaal met vooral veel gepensioneerden. Opvallend aan deze klasse is de lage mate van sociaal- en persoonskapitaal. De werkenden (30%) is het segment met het meest economisch- en persoonskapitaal. Het kenmerkende aan dit segment is het hoge aantal werkenden.

Met de manier waarop de data gewogen is, is het mogelijk om inzicht te krijgen in de veranderingen die er optreden in de segmentatie wanneer de data een betere afspiegeling is van Friesland. In de LCA met gewogen data valt het segment rijk gepensioneerden weg. De gewogen data heeft ook invloed op de andere drie segmenten. Er is een segment wat sterk overeenkomt met het precariaat uit de originele LCA. Opvallend aan die klasse is het hoge aantal gepensioneerden. Van de andere twee segmenten heeft één segment veel sociaal- en persoonskapitaal en de andere klasse heeft vooral veel economisch kapitaal. Wanneer de data op een officiële manier wordt gewogen, is het aannemelijk dat gepensioneerden grotendeels wegvallen in de segmentatie. Dit is weinig verrassend, aangezien ouderen oververtegenwoordigd zijn in de steekproef.

Een deel van de hoofdvraag is het beantwoorden van de vraag welke vorm van kapitaal het meeste invloed heeft op de segmentatie in Friesland en hoe dat zich verhoudt tot Nederland. De invloed van economisch kapitaal is het grootst; alle drie de dimensies die het economisch kapitaal vormen dragen sterk bij aan de segmentatie. Daarnaast zijn de mate van

zelfregie en kwaliteit van contacten belangrijke dimensies om de segmentatie in Friesland te verklaren.

In zowel in Nederland als geheel als in Friesland heeft het economisch kapitaal veel invloed en kan de eerste indeling in segmentatie worden gemaakt op basis van economisch kapitaal. Van de ‘nieuwe’ dimensies waarop segmentatie plaatsvindt vallen persoonskapitaal en cultureel kapitaal op als verschillen tussen Nederland en Friesland. In Friesland heeft het persoonskapitaal meer invloed, vooral via controle over het eigen leven. Het cultureel kapitaal heeft in Nederland een grotere invloed dan in Friesland en ook heeft het een sterke samenhang met het sociaal kapitaal. In de Nederlandse segmentatie geldt dat wanneer het sociaal kapitaal hoog is, het cultureel kapitaal ook en visa versa.

Discussie

Segmentatie in Friesland is onderzocht met een Latente Klassen Analyse (LCA). De LCA is uitgevoerd met de poLCA (Drew & Linzer, 2011). Hoewel dit pakket op een eenvoudige manier een LCA weergeeft, zijn er wel enkele beperkingen aan verbonden. Zo heeft de poLCA geen mogelijkheid om een weegfactor toe te voegen in de analyse. Dit heeft tot gevolg dat de segmenten die gevonden worden in de LCA niet optimaal van toepassing zijn op Friesland.

Een andere beperking aan de poLCA is het gemis aan parameters. Om een completer beeld te krijgen van de segmenten, zou het handig zijn om de segmenten te kunnen onderzoeken in relatie met andere variabelen. Op deze manier wordt het beeld van de segmenten completer. Er zijn statistische programma's die wel die mogelijkheid hebben. De poLCA kan dat niet.

Voor verder onderzoek naar segmentatie in de Friese samenleving is het een mogelijkheid om te kijken naar andere software. Het SCP gebruikt de software Mplus (www.statmodel.com) en andere aanbevolen software is LatentGold (www.statisticalinnovations.com). Deze software is ontwikkeld door hoogleraar prof. dr. J. Vermunt; een Nederlandse hoogleraar verbonden aan de Tilburg University. Hij doet veel onderzoek naar latente klassen analyse (Vermunt & Magidson, 2002; Vermunt, 2004).

De segmentatie in de Friese samenleving is onderzocht met data van Panel Fryslân, een panel van het FSP (FSP, 2017). Dit panel is geworven onder de bevolking van provincie Friesland. De data is geschikt om uitspraken te doen over Friesland. Echter, niet alle vier de vormen van kapitaal worden optimaal gemeten.

Op het gebied van economisch kapitaal missen er gegevens over het vermogen van huishoudens. Dit heeft mogelijk invloed op de vorming van klassen en dan vooral klassen met ouderen. Ouderen hebben meestal een afbetaalde koopwoning. Dat betekent dat ze een relatief groot vermogen hebben en hun financiële situatie niet van een dermate laag niveau is, zoals naar voren komt in dit onderzoek met de groep ‘arme gepensioneerden’.

Het sociaal kapitaal wordt breed gemeten met de kwaliteit en kwantiteit van sociale contacten. Een uitbreiding op sociaal kapitaal zou zijn om het instrumentele netwerk uit te vragen. Dit kan door te vragen naar beroepen van de mensen waarmee men in contact staat. door het kennen van invloedrijke mensen mee te nemen als indicator van sociaal kapitaal. In onderzoek naar segmentatie in de Britse samenleving is dit gedaan (Savage et al. 2013).

Het cultureel kapitaal is gemeten aan de hand van de Friese taalbeheersing. Echter, taal is maar één van de onderdelen van cultureel kapitaal. In andere onderzoeken wordt sociaal kapitaal onder meer gemeten als luxe consumptie, bezit van kunst en/of muziekinstrumenten en aantal keer op vakantie (Eijck & Kraaykamp, 2009).

Tot slot is er persoonskapitaal. In dit onderzoek is dat gemeten als ervaren gezondheid en controle over het eigen leven. Dit dient als een proxy voor de mate van zelfvertrouwen. Het zou een goed idee zijn om persoonskapitaal uit te breiden met minder subjectieve indicators van persoonskapitaal, zoals het BMI. Het SCP heeft het BMI gemeten door te vragen naar lengte en gewicht (SCP, 2014).

Dit onderzoek dient als verkennend onderzoek naar de segmentatie in de Friesland. Met de beschikbare data die voorhanden zijn bij het Fries Sociaal Planbureau is de eerste aanzet geleverd. Dit heeft tot een onderscheid van vier segmenten geleid. Door de meetinstrumenten uit te breiden kunnen de vormen van kapitaal completer worden uitgevraagd. Hierdoor bevatten de uitkomsten over de Friese samenleving minder ruis, waardoor de uitkomsten nog beter van toepassing zijn op de Friese samenleving.

Literatuur

- Bavel van, B. (2014). Vermogensongelijkheid in Nederland. De vergeten dimensie. In M. Kremer, M. Bovens, E. Schrijvers & R. Went (Eds.), *hoe ongelijk is Nederland* (pp. 79 – 100). Amsterdam University Press.
- Bourdieu, P. (1986). *The forms of capital handbook of theory and research for the sociology of education*. New York: Greenwald Press.
- Brunello, G., & d’Hombres, B. (2007). Does body weight affect wages?: Evidence from Europe. *Economics & Human Biology*, 5(1), 1-19. Doi: 10.1016/j.ehb.2006.11.002
- Burt, R. S., (2017). Structural holes versus network closure as social capital. *Social Capital*, 31-56.
- Centraal Bureau voor de Statistiek. (2016). Standaard Onderwijsindeling, 2016. Geraadpleegd van <https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs-en-beroepen/standaard-onderwijsindeling--soi--/standaard-onderwijsindeling-2016>
- Centraal Bureau voor de Statistiek. (2018). Inkomen naar leeftijd, branche en opleiding, 2007 – 2015. Geraadpleegd van <https://www.cbs.nl/nl-nl/maatwerk/2018/18/inkomen-naar-leeftijd-branche-en-opleiding-2007-2015>
- Centraal Bureau voor de Statistiek. (2018). Monitor Brede Welvaart 2018. Geraadpleegd van [file://fs1/sh-fsp/home/martenhospes/Documents/Downloads/Eindversie%20Monitor%20Brede%20welvaart%202018%20%20web_%20\(1\).pdf](file://fs1/sh-fsp/home/martenhospes/Documents/Downloads/Eindversie%20Monitor%20Brede%20welvaart%202018%20%20web_%20(1).pdf)
- Coleman, J. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, 94, 95-120.

- CPB (2017). Macro Economische Verkenning 2017. Geraadpleegd op <https://www.cpb.nl/publicatie/macro-economische-verkenning-mev-2017#docid-158785>
- Graaf de, R. M., Ten Have, C., Van Gool, Van Dorsselaer, S., (2012). Prevalence of mental disorders, and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence study. *Social psychiatry and psychiatric epidemiology*, 47, 203-213. doi: 10.1007/s00127-010-0334-8
- Doorne-Huiskes van, A. (2015). Verschil in Nederland: een mooie studie. *Mens & Maatschappij*, 90, 199-211. doi: 10.1557/MEM2015.2.DOOR
- Drew A. L., & Jeffrey B. L. (2011). polCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*, 42(10), 1-29. Geraadpleegd van <http://www.jstatsoft.org/v42/i10/>.
- Van Eijck, K., & Kraaykamp, G. (2009). De intergenerationele reproductie van cultureel kapitaal in belichaamde, geïnstitutionaliseerde en geobjectiveerde vorm. *Mens En Maatschappij*, 84(2), 177-206. doi:10.5117/MEM2009.2.EIJC
- Engbersen, G. B. M. (2003). De armoede van sociaal kapitaal. *Economische Statistische Berichten*, 88, geraadpleegd van <http://hdl.handle.net/11245/1.223337>
- Eskelinen, L., Kohvakka, A., Merisalo, T., Hurri, H., & Wagar, G. (1911). Relationship between the self-assessment and clinical assessment of health status and work ability. *Scandinavian Journal of Work, Environment & Health*, 17, 40-47.
- Fries Sociaal Planbureau. (2017). Onderzoeksverantwoording Panel Fryslân. Geraadpleegd van https://www.fsp.nl/wp-content/uploads/2019/01/onderzoeksverantwoording_panel_fryslan_0.pdf
- Fries Sociaal Planbureau. (2018). Monitor sociale samenhang. Geraadpleegd van <https://www.fsp.nl/monitoren/sociale-samenhang/>

- Fries Sociaal Planbureau. (2018). Monitor onderwijs. Geraadpleegd van <https://www.fsp.nl/monitoren/onderwijs/>
- Fries Sociaal Planbureau. (2018). Leven in Fryslân; editie 2019. Geraadpleegd op https://www.fsp.nl/wp-content/uploads/2019/06/FSP_Leven_In_Fryslan_TOTAAL-los-DEF.pdf
- Ganzeboom, H. B. G., & Ultee, W. C. (1996). De sociale segmentatie van Nederland in 2015. Den Haag: Sdu uitgevers
- Ganzeboom, H. B. G. (1988). Leefstijlen in Nederland. Geraadpleegd van: [file://fs1/sh-fsp/home/martenhospes/Documents/Downloads/Leefstijlen%20in%20Nederland%20\(1\).pdf](file://fs1/sh-fsp/home/martenhospes/Documents/Downloads/Leefstijlen%20in%20Nederland%20(1).pdf)
- Ganzeboom, H. B. G., & Treiman, D. J. (2007). Ascription and achievement in occupational attainment in comparative perspective. *Tijdschrift onbekend*
- Harpham, T., Grant, E., & Thomas, E. (2002). Measuring social capital within health surveys: key issues. *Health policy and planning*, 17(1), 106-111.
- Huppert, F. A., Marks, N., Clark, J., Siegrist, A., & Vitters, J. (2009). Measuring well-being across Europe: Description of the ess well-being module and preliminary findings. *Social Indicators Research*, 91(3), 301-315.
- Kawachi, I., & Berkman, L. I. (2000). Social cohesion, social capital, and health. In L. F. Berkman & I. Kawachi. *Social epidemiology*. 174-190. New York: Oxford University Press
- Lin, N. (1999). Building a network theory of social capital. *Connections*, 22, 28-51.
- Monitor Inkomen. (2018). Geraadpleegd van <https://www.fsp.nl/monitoren/inkomen/>
- Monitor Gezondheid. (2019). Geraadpleegd van <https://www.fsp.nl/monitoren/gezondheid/>

- Monitor Onderwijs. (2018). Geraadpleegd van <https://www.fsp.nl/monitoren/onderwijs/>
- Monitor Sociale Samenhang. (2018). Geraadpleegd van <https://www.fsp.nl/monitoren/sociale-samenhang/>
- Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: background, derivation and applications. *WIREs Comput Stat*, 4, 188-203. Doi 10.1002/wics.199
- Olaniyan, D. A., & Okemakinde, T. (2008). Human capital theory: Implications for educational development. *Pakistan Journal of Social Sciences*, 5(5), 479-483.
- Parcel, T. L., & Meuller, C. W. (1983). Occupational differentiation, prestige, and socioeconomic status. *Work and Occupations*, 10(1), 49-80. doi: 10.1177/0730888483010001004
- Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of health and social behavior*, 19(1), 2-21.
- Piketty, T. (2015). About capital in the twenty-first century. *American Economic Review*, 105(5), 48-53. Doi 10.1257/aer.p20151060
- Pinxten, W., & Lievens, J. (2014). The importance of economic, social and cultural capital in understanding health inequalities: using a Bourdieu-based approach in research on physical and mental health perceptions. *Sociology of health & illness*, 36(7), 1095-1110. Doi 10.1111/1467-9566.12154
- Pennewaard, M. (2018, december 18). De Friese paradox: gelukkig met minder. Leeuwarder Courant. Geraadpleegd op <https://www.lc.nl/friesland/De-Friese-paradox-gelukkig-met-minder-23921030.html>
- Putnam, R. D. (2015). *Our kids: The American Dream In Crisis*. VS: Simon & Schuster

Putnam, R. D. (1995). Tuning in, Tuning out: the strange disappearance of social capital in America. *Political Science & Politics*, 28 (4), 664-684.

Rijksuniversiteit Groningen. (z.d.). Sociologie. Geraadpleegd van

<https://www.rug.nl/gmw/sociology/>

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Geraadpleegd van: <https://www.R-project.org/>

Reuten, G. (2018). De Nederlandse vermogensverdeling in internationaal perspectief.

TPEdigitaal, 12, 1-8. Geraadpleegd bij

<https://www.tpedigitaal.nl/sites/default/files/bestand/De-Nederlandse-vermogensverdeling-in-internationaal-perspectief.pdf>

Savage et al. (2013). A new model of social class? Findings from the BBC's great British survey experiment. *Sociology*, 47(2), 219- 250. doi: 10.1177/0038038513481128

Sociaal & Cultureel Planbureau. (2014). Verschil in Nederland. Geraadpleegd van

https://www.scp.nl/Publicaties/Alle_publicaties/Publicaties_2014/Verschil_in_Nederland

Strategieeraad Rijksbreed. (2013). Rijksbrede trendverkenning. Geraadpleegd van

<https://kennisopenbaarbestuur.nl/media/103181/rijksbrede-trendverkenning.pdf>

Sweetland, S. R. (1996). Human capital theory: foundations of a field of inquiry. *Review of Educational Research*, 66 (3), 341-359. Doi 10.3102/00346543066003341

Taal in Nederland. (z.d.). Geraadpleegd van <http://taal.phileon.nl/fries.php>

Tafarodi, R. W., & Swann Jr, W. B. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences*, 31(5), 653-673. Doi 10.1016/S0191-8869(00)00169-0

Veenhoven, R. (2000). The four qualities of life. Ordering concepts and measures of the good life. *Journal of Happiness Studies*, 1 (1), 1-39.

Vermunt, J. K. (2004). Toepassingen van latent klasse analyse in social wetenschappelijk onderzoek. *Sociale Wetenschappen* 47(1), 2-14.

Vermunt J. K., & Magidson, J. (2002). Latent Cluster Analysis. In J. Hagnaars & a. McCutcheon (Eds.), *Applied latent class analysis*. (pp. 89-106). Cambridge: Cambridge University Press.

Bijlage

Leeswijzer

In de bijlages staat de uitgebreide analyse van de LCA, een correlatietabel en de vragenlijst die het FSP gebruikt bij PanelFryslân. De bijlage begint met de correlaties tussen de gebruikte variabelen, plus het geslacht en leeftijd van de panelleden. Vervolgens wordt de LCA behandeld. De latente klassen die benoemd worden in de resultaten, worden weergegeven in de bijlage. In het laatste gedeelte van de bijlage staat de vragenlijst. Dit betreft de introductievragenlijst en nog enkele vragen uit verschillende vragenlijsten.

Correlaties

In de tabel op de volgende pagina staan de correlaties tussen de variabelen. Deze correlaties staan niet in de resultatenparagraaf. Normaliter gebeurt dat wel, maar in het geval van een LCA zou dat niet een optimale bijdrage hebben aan het interpreteren van de segmenten. Een correlatietabel laat zien hoe variabelen met elkaar verbonden zijn en hoe zich tot elkaar verhouden. Bij een multiple regressieanalyse geeft het inzicht in hoe de variabelen hetzelfde voorspellen. Bij een LCA speelt dat echter niet. Correlaties zijn nodig om te bepalen hoeveel klassen de LCA moet hebben en het heeft ook geen invloed op de interpretatie van de gevonden klassen. Daarom is de keuze gemaakt om correlaties in de bijlage weer te geven.

Variabele	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Inkomen laag	-												
2. Inkomen midden	-,22*	-											
3. Inkomen hoog	-,46*	-,32*	-										
4. Inkomen onbekend	-,31*	-,21*	-,44*	-									
5. Opleiding laag	,23*	-,03	-,26*	,08*	-								
6. Opleiding Midden	,06*	,07	-,16*	,06	-,30*	-							
7. Opleiding hoog	-,22*	-,05	,32	-,11	-,45*	-,72*	-						
8. Werkend	-,21*	-,05	,22*	,00	-,19*	,04	,07	-					
9. Werkloos	,22*	-,01	-,21*	,03*	,08*	,07*	-,13*	-,36*	-				
10. Pensioen	,06*	,06	-,07*	-,02	,14*	-,12*	,02	-,76*	-,32	-			
11. Fysieke contacten	,02	,04	,00	-,06	,06	,04	-,08	,09	-,01	,03	-		
12. Digitale contacten	-,04	-,03	,01	,05	,00	,06	-,06	,02	-,01	-,08	,36*	-	
13. Kwaliteit contacten	-,09	-,04	,13*	-,04	-,11	-,04	,12*	-,18	-,06	-,14	,27*	,30*	
14. Fries	,07*	,02	-,08*	,01	,07*	,07	-,11*	,03	,07	-,07	,11*	-,02	
15. Ervaren gezondheid	-,13	-,06	,14*	,01	,06	-,01	,05	,12	-,16	-,01	,05	,04	
16. Zelfregie	-,17*	-,01	,20*	-,06	-,14*	-,05	,14*	,23*	-,10*	-,16*	,05	,11*	
17. Geslacht	,03	-,06	-,14*	,18*	,02	,03	,04	,11*	,17*	-,23	-,01	,17	
18. Leeftijd	,07*	,04	-,10*	,02	,17*	-,11*	-,11	-,59*	-,17*	,71*	,01	,01	

* significant bij $p < 0,05$.

De variabelen hoge opleiding en hoog inkomen vertonen een negatieve samenhang met beheersing van de Friese taal. Wanneer opleiding en inkomen hoger worden, daalt de beheersing van de Friese taal. Dit suggereert dat een hoge mate van economisch kapitaal gepaard gaat met een geringe beheersing van het Fries.

De variabelen hoge opleiding, hoog inkomen en werkend heeft een positieve correlatie met zelfregie. Niet-werkend en gepensioneerd vertoont juist een negatieve samenhang met zelfregie. Dit suggereert dat de werkende bevolking met een hoge mate van economisch kapitaal, een hoge mate van zelfregie ervaart.

De variabelen van kwantiteit van sociale contacten hebben lage correlaties met andere variabelen. De kwaliteit van sociale contacten heeft een positieve samenhang met opleiding en inkomen. Werkenden hebben een hogere kwaliteit van sociale contacten. Er is een relatief sterk verband tussen de kwaliteit van sociale contacten en de mate van zelfregie.

De beheersing van de Friese taal, heeft naast verbanden met opleiding en inkomen, ook nog een significant verband met fysiek contact. Dit verband is positief. Dat betekent dat naarmate de beheersing van de Friese taal stijgt, het aantal fysieke contacten omhoog gaat. Met andere variabelen heeft de beheersing van de Friese taal lage correlaties.

Persoonskapitaal wordt gemeten met ervaren gezondheid en zelfregie. Zelfregie is een variabele die veel opvallende correlaties heeft met andere variabelen. Dit is al benoemd in voorgaande beschrijving. Er kan gesteld worden dat een hoge mate van zelfregie, samengaat met veel fysieke contacten, werkend, hoge opleiding en hoog inkomen. De ervaren gezondheid is minder uitgesproken, maar het volgt bijna exact hetzelfde verband als zelfregie.

Latente Klassen Analyse

Leeswijzer

In deze bijlage wordt de LCA gedaan. Eerst wordt de LCA gedaan met de originele, ongewogen data en daarna wordt de LCA gedaan met de gewogen data. In deze bijlage wordt eerst gezocht naar het juiste aantal klassen met de BIC-waarde en vervolgens wordt de LCA gedaan met een bepaald aantal klassen. De LCA is als volgt beschreven: eerst wordt per variabele beschreven hoe de verschillende klassen daarop scoren en daarna worden de vorm en de grootte van de klassen beschreven.

R-script zonder gewogen data

De LCA wordt gedaan binnen R. Dit is gedaan met het volgende script: (tekst in het scriptie met een '#' staat voor uitleg)

```
# Allereerst worden de ontbrekende packages binnengehaald welke nodig zijn om
# een LCA te doen en vervolgens worden de packages geladen.

# Installeer ontbrekende packages
packageList <- c("foreign", "stats4", "poLCA")
new.packages <- packageList[!(packageList %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages)

# Laad alle packages
lapply(packageList, require, character.only = TRUE)

# Na het installeren van de benodigde packages worden de bestanden binnengehaald.

file1 <- "R:\\FSP\\PW\\00 stagiaires\\Marten
Hospes\\statistiek\\20190529mergedData_panelRI_w.sav"
dataset <- read.spss(file1, use.value.labels = TRUE, to.data.frame = TRUE)

# Er wordt gebruik gemaakt van de poLCA. Deze kan alleen met
# numerieke variabelen omgaan. voor het gemak worden alle variabelen numeriek
# gemaakt.

#variabelen numeric maken
dataset$Zelfregie_schaal.num <- as.numeric(dataset$Zelfregie_schaal)
dataset$Kwalcontact_schaal.num <- as.numeric(dataset$Kwalcontact_schaal)
dataset$Digitaalcontact_schaal.num <- as.numeric(dataset$Digitaalcontact_schaal)
dataset$Fysiekcontact_schaal.num <- as.numeric(dataset$Fysiekcontact_schaal)

dataset$gezondheid.num <- as.numeric(dataset$zorg_V004a)

dataset$Friessprekend.num <- as.numeric(dataset$duur_sprektFries)
dataset$werk_werkloos_pensioen.num <- as.numeric(dataset$duur_werk)
dataset$Ink_laag_midden_hoog_weetniet.num <- as.numeric(dataset$duur_ink_hh)
dataset$Opleiding.num <- as.numeric(dataset$duur_opleiding)
```

```
## LCA ##

# 1. testdraai met twee klassen.
lca.poLCA <- poLCA(
  cbind(Zelfregie_schaal.num,
        gezondheid.num,
        kwalicontact_schaal.num,
        Digitaalcontact_schaal.num,
        Fysiekcontact_schaal.num,
        Friessprekend.num,
        Opleiding.num,
        werk_werkloos_pensioen.num,
        Ink_laag_midden_hoog_weetniet.num)~1,
  data=dataset,
  nclass=2,
  nrep=1,
  na.rm=T,
  graphs=T,
  maxiter = 100000)

# We hebben nu een latente klassen analyse gedaan, waarmee een vast aantal
# klassen wordt gegeven. Het wordt echter interessanter als we een analyse doen
# waarbij wordt berekend wat het optimale model is. Dit wordt berekend
# met de Bayesian Information Criterion (BIC). Het model met de laagste BIC
# waarde, is het beste model.

# 2. model met laagste BIC-waarde

oldBIC = 100000
newBIC = 100000
modellen <- NULL
aantalklassen <- 2

while(newBIC<=oldBIC){
  modelnr <- aantalklassen - 1

  modellen[[modelnr]] <- poLCA(
    cbind(Zelfregie_schaal.num,
          gezondheid.num,
          kwalicontact_schaal.num,
          Digitaalcontact_schaal.num,
          Fysiekcontact_schaal.num,
          Friessprekend.num,
          Opleiding.num,
          werk_werkloos_pensioen.num,
          Ink_laag_midden_hoog_weetniet.num)~1,
    data=dataset,
    nclass=aantalklassen,
    nrep=1,
    na.rm=T,
    graphs=T,
    maxiter = 100000)

  aantalklassen = aantalklassen + 1
  oldBIC <- newBIC
  newBIC <- modellen[[modelnr]]$bic
}

```

```
# plot model met laagste BIC (één na laatste model)
plot(modellen[[length(modellen)-1]])

# plot BIC's van alle modellen
for (i in 1:length(modellen)){cat(paste0("Aantal clusters: ", i + 1, "\nBIC: ",
modellen[[i]]$bic, "\n\n")) }
```

Aantal klassen

Het aantal klassen wordt onder andere bepaald met de BIC-waarde. Het aantal klassen met de laagste BIC waarde geeft de beste fit bij de data. Met R is een script geschreven waardoor de poLCA naar het aantal klassen zoekt met de laagste BIC score. Wanneer het laagste getal is gevonden, stopt de berekening. Dit levert de volgende output op:

```
Aantal clusters: 2
BIC: 22349.8824350007
```

```
Aantal clusters: 3
BIC: 22340.3339986346
```

```
Aantal clusters: 4
BIC: 22445.9492918643
```

Uit de BIC-waarde blijkt dat bij drie klassen de BIC-waarde het laagst is. Op basis van de BIC-waarde kan er gesteld worden dat drie klassen het beste bij de data past. In de volgende output staan de uitkomsten van de LCA met de drie klassen.

```
Conditional item response (column) probabilities,
by outcome variable, for each class (row)
```

```
$zelfregie_schaal.num
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5) Pr(6)
class 1: 0.0029 0.0893 0.2962 0.4402 0.1601 0.0114
class 2: 0.0030 0.0160 0.1791 0.3947 0.3459 0.0614
class 3: 0.0000 0.0096 0.0385 0.3301 0.5053 0.1165
```

```
$gezondheid.num
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0058 0.0692 0.2448 0.6438 0.0364
class 2: 0.0000 0.0250 0.1972 0.6691 0.1086
class 3: 0.0000 0.0176 0.1128 0.6902 0.1794
```

```
$kwalcontact_schaal.num
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0166 0.1196 0.4455 0.3766 0.0416
class 2: 0.0164 0.1183 0.4570 0.3199 0.0884
class 3: 0.0047 0.0000 0.1014 0.5624 0.3315
```

```
$digitaalcontact_schaal.num
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.4458 0.3444 0.1434 0.0664 0.0000
```

```
class 2: 0.5265 0.4325 0.0410 0.0000 0.000
class 3: 0.0767 0.4013 0.4315 0.0826 0.008
```

\$Fysiekcontact_schaal.num

```
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1562 0.3204 0.4260 0.0974 0.0000
class 2: 0.2862 0.4413 0.2623 0.0102 0.0000
class 3: 0.0143 0.2261 0.6347 0.1222 0.0027
```

\$Friessprekend.num

```
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1328 0.0869 0.0360 0.1016 0.6427
class 2: 0.1885 0.1970 0.0857 0.0523 0.4765
class 3: 0.1230 0.1447 0.1227 0.0675 0.5421
```

\$Opleiding.num

```
      Pr(1) Pr(2) Pr(3)
class 1: 0.3968 0.4054 0.1978
class 2: 0.0259 0.2118 0.7624
class 3: 0.0547 0.3541 0.5912
```

\$Werk_werkloos_pensioen.num

```
      Pr(1) Pr(2) Pr(3)
class 1: 0.2071 0.2812 0.5117
class 2: 0.4893 0.0201 0.4906
class 3: 0.6306 0.1057 0.2637
```

\$Ink_laag_midden_hoog_weetniet.num

```
      Pr(1) Pr(2) Pr(3) Pr(4)
class 1: 0.5339 0.1574 0.0356 0.2731
class 2: 0.0631 0.1207 0.6071 0.2091
class 3: 0.1346 0.1312 0.5461 0.1881
```

Estimated class population shares

```
0.3295 0.3157 0.3548
```

Predicted class memberships (by modal posterior prob.)

```
0.3251 0.3175 0.3573
```

```
=====
Fit for 3 latent classes:
=====
```

```
number of observations: 1055
number of estimated parameters: 98
residual degrees of freedom: 957
maximum log-likelihood: -10829.06
```

```
AIC(3): 21854.13
```

```
BIC(3): 22340.33
```

```
G^2(3): 7195.106 (Likelihood ratio/deviance statistic)
```

```
X^2(3): 720381.8 (Chi-square goodness of fit)
```

Beschrijving van variabelen

Zelfregie: Op zelfregie scoort klasse 3 duidelijk hoger dan de andere groepen. Klasse 2 scoort net wat hoger dan klasse 1.

Gezondheid: Op het gebied van ervaren gezondheid zijn de verschillen niet zo groot als op zelfregie. Wel valt op dat klasse 1 het laagst scoort op zeer goede gezondheid, klasse 2 middelmatig en klasse 3 het hoogst.

Kwaliteit van contacten: Op de kwaliteit van contacten scoort klasse 3 het hoogst. Klasse 1 en 2 liggen dicht bij elkaar.

Kwantiteit van contacten: Op het gebied van digitale contacten scoren alle drie de klassen laag. Klasse 3 scoort het minst laag van de groepen. Het aantal fysieke contacten liggen net wat hoger voor alle drie de klassen. Net als bij digitale contacten scoort ook klasse 3 hier het hoogst. Echter, klasse 1 scoort bijna even hoog.

Beheersing van het Fries: klasse 1 heeft de beste taalbeheersing en klasse 2 het minste. Verder valt op dat de scores over het algemeen hoog liggen op het gebied van taalbeheersing.

Opleiding: Op het gebied van opleiding worden de verschillen groter. Klasse 2 is de hoogstopgeleide groep en groep 1 het laagstopgeleide groep. Klasse 3 zit daartussen in, maar ligt dicht bij klasse 2 dan klasse 1. Klasse 1 en 3 hebben beide veel midden opgeleiden.

Arbeidssituatie: Klasse 3 heeft het hoogst aantal werkenden, vlak gevolgd door klasse 2. Klasse 1 heeft zowel het hoogst aantal niet-werkenden als gepensioneerden. Klasse 2 heeft ook veel gepensioneerden.

Inkomen: Het inkomen staat hier in 4 categorieën beneden modaal – modaal – boven modaal – weet niet/wil niet zeggen. De categorie ‘weet niet/ wil niet zeggen’ is ingevuld door een groot aantal respondenten. In andere variabelen is deze categorie omgezet naar *missing values*. Bij de inkomensvariabele echter niet omdat inkomensvragen over het algemeen in Nederland weinig worden ingevuld. Een andere keuze zou zijn om de inkomensvraag niet mee te nemen in de LCA.

Klasse 2 en 3 hebben het hoogste inkomen, waarbij klasse 2 net wat hoger scoort. Klasse 1 heeft verreweg het laagste inkomen. Het verschil tussen klasse 1 en de andere klassen is zeer groot.

Beschrijving van groepen

Klasse 1: Klasse 1 heeft een aandeel van 33,0% in de onderzoekspopulatie. Afgaande op de scores op de variabelen valt af te leiden dat klasse 1 laag scoort op alle vormen van kapitaal. Op de beheersing van de Friese taal na, scoort klasse 1 op alle variabelen het laagst. Hieruit kan geconcludeerd worden dat klasse 1 de groep is met het minste economisch, sociaal, cultureel en persoonskapitaal. Opvallend aan klasse 1 is het hoge aantal gepensioneerden.

Klasse 2: klasse 2 heeft een aandeel van 31,6% in de onderzoekspopulatie. Daarmee is het de kleinste klasse in de onderzoekspopulatie. Klasse 2 scoort hoog op economisch kapitaal. Het verschil wordt vooral gemaakt door de hoge opleiding en een hoger inkomen dan klasse 1. Daarnaast telt klasse 2 veel gepensioneerden. Op persoonskapitaal en sociaal kapitaal ligt klasse 2 in het midden.

Klasse 3: klasse 3 heeft een aandeel van 35,5% in de onderzoekspopulatie. Op de variabelen die persoonskapitaal en sociaal kapitaal meten scoort klasse 3 het hoogst. Op de beheersing van de Friese taal het laagst en op het gebied van economisch kapitaal zit klasse 3 dicht bij klasse 2. Klasse 3 bestaat vooral uit werkenden.

Verdere zoektocht

Drie klassen geven de laagste BIC-waarde en daarom zou op statistische basis drie klassen het beste bij de onderzoekspopulatie passen. Echter, een extra klasse zou meer theoretische duiding kunnen geven aan de segmentatie. Om te onderzoeken hoe vier klassen worden gevormd, wordt een LCA met vier klassen gedaan.

Conditional item response (column) probabilities,
by outcome variable, for each class (row)

\$zelfregie_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)	Pr(6)
class 1:	0.0037	0.0943	0.2433	0.4526	0.1948	0.0113
class 2:	0.0000	0.0349	0.2456	0.4577	0.2127	0.0491
class 3:	0.0039	0.0154	0.2334	0.3855	0.3104	0.0514
class 4:	0.0000	0.0088	0.0000	0.2848	0.5755	0.1309

\$gezondheid.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0072	0.0743	0.2383	0.6242	0.0560
class 2:	0.0000	0.0332	0.1614	0.7870	0.0184
class 3:	0.0000	0.0278	0.2094	0.6828	0.0800
class 4:	0.0000	0.0142	0.1278	0.6185	0.2395

\$Kwalcontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0245	0.1444	0.4958	0.3131	0.0222
class 2:	0.0000	0.0223	0.2095	0.6072	0.1610
class 3:	0.0162	0.1436	0.5012	0.3058	0.0332
class 4:	0.0065	0.0000	0.1176	0.4986	0.3773

\$Digitaalcontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.5965	0.3275	0.0674	0.0086	0.0000
class 2:	0.0388	0.3577	0.4364	0.1671	0.0000
class 3:	0.5428	0.4186	0.0386	0.0000	0.0000
class 4:	0.1507	0.4503	0.3353	0.0543	0.0094

\$Fysiekcontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.2117	0.4209	0.3294	0.0379	0.0000
class 2:	0.0000	0.0982	0.6502	0.2516	0.0000
class 3:	0.3100	0.4420	0.2346	0.0134	0.0000
class 4:	0.0556	0.2950	0.5916	0.0547	0.0031

\$Friessprekend.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.1228	0.1127	0.0344	0.1246	0.6054
class 2:	0.1161	0.0622	0.0830	0.0328	0.7059
class 3:	0.2063	0.1755	0.0864	0.0424	0.4895
class 4:	0.1405	0.1928	0.1207	0.0814	0.4646

\$opleiding.num

	Pr(1)	Pr(2)	Pr(3)
class 1:	0.3639	0.3723	0.2638
class 2:	0.2816	0.4362	0.2822
class 3:	0.0329	0.2205	0.7466
class 4:	0.0000	0.2989	0.7011

\$werk_werkloos_pensioen.num

	Pr(1)	Pr(2)	Pr(3)
class 1:	0.2023	0.2758	0.5219
class 2:	0.3607	0.2037	0.4356
class 3:	0.4778	0.0112	0.5110
class 4:	0.6887	0.0718	0.2394

\$Ink_laag_midden_hoog_weetniet.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)
class 1:	0.5697	0.1681	0.0420	0.2202
class 2:	0.3512	0.1348	0.1851	0.3289
class 3:	0.0000	0.1342	0.6258	0.2401
class 4:	0.0846	0.1120	0.6604	0.1430

Estimated class population shares

0.2628 0.1945 0.2403 0.3025

Predicted class memberships (by modal posterior prob.)

0.2502 0.1791 0.2559 0.3147

=====
Fit for 4 latent classes:
=====

number of observations: 1055
number of estimated parameters: 131
residual degrees of freedom: 924
maximum log-likelihood: -10769.43

AIC(4): 21800.86
BIC(4): 22450.79
 $G^2(4)$: 7075.839 (Likelihood ratio/deviance statistic)
 $\chi^2(4)$: 492087.3 (Chi-square goodness of fit)

Beschrijving van de variabelen

Zelfregie: klasse 4 is de uitschieter op zelfregie. Deze klasse scoort veel hoger op zelfregie dan andere klassen. De andere klassen staan redelijk gelijk aan elkaar, waarbij klasse 1 het laagst scoort.

Gezondheid: klasse 4 scoort het hoogst op ervaren gezondheid. Klasse 1 scoort het laagst.

Kwaliteit van contacten: Op de kwaliteit van contacten zijn klasse 2 en 4 uitschieters. Op de hoogste kwaliteit van contacten scoort klasse 4 het hoogst. Op de één na hoogste kwaliteit van contacten is klasse 2 een uitschieter.

Kwantiteit van contacten: De kwantiteit van contacten wordt gevormd door digitale en fysieke contacten. Klasse 2 en 4 scoren samen het hoogst op het aantal digitale contacten. Op fysieke contacten is klasse 2 het hoogst en klasse 4 een klein beetje lager.

Beheersing van de Friese taal: klasse 2 scoort het hoogst op de beheersing van de Friese taal en klasse 3 en 4 het laagst.

Opleiding: klasse 3 en 4 zijn verreweg de hoogst opgeleide groepen van de onderzoekspopulatie. Klasse 3 scoort net wat hoger op de hoogste opleiding en klasse 4 scoort net wat hoger op de middelste opleidingscategorie.

Arbeidssituatie: De arbeidssituatie is verdeeld in werkend, niet-werkend en gepensioneerd. Klasse 4 scoort het hoogst op de werkenden gevolgd door groep 3. Klasse 1 en 2 scoren het hoogst op niet-werkenden. Qua gepensioneerden scoren klasse 1 en 3 het hoogst en 2 het laagst.

Inkomen: klasse 3 en 4 hebben het hoogste inkomen waarbij groep 4 net een beetje hoger inkomen heeft. Klasse 1 heeft verreweg het laagste inkomen.

Beschrijving van de klassen

Klasse 1: Net als in de LCA met 3 klassen, is ook in deze LCA klasse 1 de groep met weinig economisch, sociaal, cultureel en persoonskapitaal. Op alle vormen van kapitaal scoort klasse 1 het laagst. Het aandeel van klasse 1 in de onderzoekspopulatie is 26,3%. Opvallend aan deze klasse is het hoge aantal gepensioneerden.

Klasse 2: Op de variabelen die sociaal en cultureel kapitaal meten, scoort klasse 2 hoog. Dit staat in contrast tot de variabelen waarmee persoons en economisch kapitaal worden gemeten. Met een aandeel van 19,5% in de onderzoekspopulatie is klasse 2 wel de kleinste klasse.

Klasse 3: klasse 3 is een klasse met weinig persoons, cultureel en sociaal kapitaal. Klasse 3 scoort vooral hoog op economisch kapitaal. Deze klasse onderscheidt zich vooral door het hoge aantal gepensioneerden. Samen met klasse 1 heeft klasse 3 het hoogste aantal gepensioneerden en waar in klasse 1 arme gepensioneerden zitten, beschikken de gepensioneerden van klasse 3 over meer economisch kapitaal.

Klasse 4: Op de variabelen waarmee persoons, sociaal en economisch kapitaal wordt gemeten, scoort groep 4 hoog. Met economisch kapitaal vertoont klasse 4 overeenkomsten met klasse 3. Echter, klasse 4 bestaat vooral uit werkenden. Dat betekent dat klasse 4 de groep is over veel kapitaal beschikt en bestaat uit de werkende klasse.

Het aantal groepen

De LCA is nu gedaan met drie en vier klassen, waarbij drie klassen de laagste BIC-waarde geeft. Bij een LCA van drie klassen zijn economisch, sociaal, cultureel en persoonlijk kapitaal ongeveer als volgt verdeeld: er is een klasse met weinig kapitaal, een klasse die vooral uitstijgt op economisch kapitaal en een klasse die het meer heeft van sociaal en persoonlijk kapitaal. Wanneer er een klasse wordt toegevoegd de LCA liggen de verhoudingen enigszins anders. De klasse die erbij komt, manifesteert zich vooral op economisch kapitaal. Waar er bij drie klassen één absolute groep is met veel economisch kapitaal, is er bij vier klassen in de LCA nog een klasse bijgekomen met veel economisch kapitaal. De ene klasse bestaat dan uit werkenden en de andere klasse uit gepensioneerden. Dit geeft te kennen dat economisch kapitaal en dan vooral arbeidssituatie onderscheid maakt tussen klassen.

Het is interessant om een LCA met vijf klassen te doen zodat er onderzocht kan worden op welke kenmerken de klassen zich nog meer splitsen. In onderstaande output staat de LCA met 5 klassen.

Conditional item response (column) probabilities,
by outcome variable, for each class (row)

\$Zelfregie_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)	Pr(6)
class 1:	0.0054	0.0836	0.2009	0.5033	0.1880	0.0187
class 2:	0.0000	0.0228	0.1057	0.3791	0.4284	0.0641
class 3:	0.0037	0.0175	0.2661	0.3679	0.2987	0.0461
class 4:	0.0000	0.0072	0.0000	0.3233	0.5191	0.1504
class 5:	0.0000	0.0907	0.3735	0.3757	0.1601	0.0000

\$gezondheid.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0097	0.0904	0.2213	0.6391	0.0395
class 2:	0.0000	0.0268	0.1090	0.8156	0.0486
class 3:	0.0000	0.0300	0.2267	0.6802	0.0630
class 4:	0.0000	0.0107	0.1369	0.5660	0.2864
class 5:	0.0000	0.0346	0.2584	0.6458	0.0612

\$kwalicontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0338	0.1525	0.4937	0.3029	0.0171
class 2:	0.0000	0.0000	0.0881	0.6966	0.2154
class 3:	0.0159	0.1556	0.5118	0.3034	0.0133
class 4:	0.0085	0.0000	0.1664	0.4201	0.4050
class 5:	0.0000	0.0932	0.4449	0.3910	0.0709

\$Digitaalcontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.6642	0.3358	0.0000	0.0000	0.0000
class 2:	0.0245	0.3565	0.5134	0.0920	0.0136
class 3:	0.5467	0.4139	0.0394	0.0000	0.0000
class 4:	0.2052	0.4817	0.2579	0.0552	0.0000
class 5:	0.2459	0.3250	0.2778	0.1513	0.0000

\$Fysiekcontact_schaal.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.2570	0.4255	0.3175	0.0000	0.0000
class 2:	0.0000	0.1347	0.6560	0.2093	0.0000
class 3:	0.3026	0.4399	0.2400	0.0174	0.0000
class 4:	0.0934	0.3357	0.5404	0.0266	0.0039
class 5:	0.0374	0.2537	0.5082	0.2008	0.0000

\$Friessprekend.num

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0873	0.1232	0.0527	0.1203	0.6165
class 2:	0.0747	0.0873	0.1320	0.0530	0.6530
class 3:	0.2120	0.1568	0.0848	0.0438	0.5026
class 4:	0.1678	0.2260	0.1029	0.0772	0.4260
class 5:	0.2021	0.0706	0.0000	0.0853	0.6421

\$opleiding.num

	Pr(1)	Pr(2)	Pr(3)
class 1:	0.2558	0.3906	0.3537
class 2:	0.1008	0.4547	0.4445

```

class 3:  0.0395 0.2504 0.7102
class 4:  0.0000 0.2195 0.7805
class 5:  0.6414 0.3586 0.0000

$werk_werkloos_pensioen.num
      Pr(1) Pr(2) Pr(3)
class 1:  0.2101 0.3296 0.4603
class 2:  0.5395 0.1545 0.3060
class 3:  0.4791 0.0000 0.5209
class 4:  0.6541 0.0703 0.2756
class 5:  0.1875 0.1851 0.6274

$Ink_laag_midden_hoog_weetniet.num
      Pr(1) Pr(2) Pr(3) Pr(4)
class 1:  0.6038 0.1438 0.0836 0.1688
class 2:  0.2261 0.1875 0.3184 0.2680
class 3:  0.0000 0.1651 0.5714 0.2635
class 4:  0.0659 0.0680 0.7493 0.1167
class 5:  0.5103 0.1212 0.0000 0.3684

Estimated class population shares
  0.1954 0.2087 0.2287 0.2451 0.122

Predicted class memberships (by modal posterior prob.)
  0.1839 0.2009 0.2559 0.2483 0.1109

=====
Fit for 5 latent classes:
=====
number of observations: 1055
number of estimated parameters: 164
residual degrees of freedom: 891
maximum log-likelihood: -10726.69

AIC(5): 21781.38
BIC(5): 22595.04
GA2(5): 6990.362 (Likelihood ratio/deviance statistic)
XA2(5): 451856.6 (Chi-square goodness of fit)

```

Beschrijving van de variabelen

Zelfregie: op zelfregie is klasse 4 de uitschieter, gevolgd door klasse 2. De andere drie klassen liggen dicht bij elkaar en het is niet echt te zeggen welke klasse het laagst is.

Gezondheid: Op gezondheid is klasse 4 de uitschieter in de hoogst ervaren gezondheid. Op de één na hoogst ervaren gezondheid liggen de verschillen dicht bij elkaar.

Kwaliteit van sociale contacten: klasse 2 en 4 scoren het hoogst op de kwaliteit van contacten, waarvan klasse 4 net wat hoger scoort op de hoogste kwaliteit van contacten. Klasse 1 scoort het laagst op de kwaliteit van contacten, gevolgd door klasse 5.

Kwantiteit van sociale contacten: kwantiteit van sociale contacten wordt gevormd door de hoeveelheid digitale en fysieke contacten. Op digitale contacten scoort klasse 2 het hoogst gevolgd door klasse 4 en 5. Klasse 3 heeft weinig digitale contacten. Het aantal fysieke

contacten is het hoogst bij klasse 2, gevolgd door klasse 4 en 5. Hoewel het aantal fysieke contacten in het algemeen hoger ligt, kan er toch gesteld worden dat klasse 2 het meeste kwantiteit van sociale contacten heeft, gevolgd door klasse 4 en 5.

Beheersing van Friese taal: klasse 2 en 5 scoren het hoogst op de beheersing van de Friese taal, maar klasse 5 scoort ook het hoogst op de laagste beheersing van de Friese taal samen met klasse 3.

Opleiding: klasse 4 is het hoogst opgeleid en kort daarop groep 3. Ver daarachter komt klasse 2 en daarna 1. Klasse scoort 0 op de hoogst opleiding. Klasse 5 scoort wel verreweg het hoogst op de laagste opleiding.

Arbeidssituatie: Klasse 4 heeft het meeste werkenden, gevolgd door Klasse 2 en daarna groep 3. Klasse 1 heeft het hoogste aantal werklozen. Klasse 5 heeft het hoogste aantal gepensioneerden, gevolgd door klasse 3.

Inkomen: Verreweg het hoogste inkomen is te vinden bij klasse 4, gevolgd door klasse 3. Klasse 1 heeft verreweg het laagste inkomen en klasse 5 het één na laagste. Opvallend aan klasse 5 is de hoge score op het antwoord 'weet niet/wil niet zeggen'.

Beschrijving van de klassen

Klasse 1: Net als in de LCA met minder klassen, is ook in de LCA met 5 klassen deze groep met het minste kapitaal. Hoewel klasse 1 niet de laagst opgeleide klasse is, is het wel de klasse met het hoogst aantal werklozen en het laagste inkomen. Op het gebied van persoonskapitaal vertoont deze klasse overeenkomsten met klasse 5. Klasse 1 heeft een aandeel van 19,5% in de onderzoekspopulatie.

Klasse 2: Op het gebied van sociaal kapitaal is klasse 2 redelijk kapitaalkrchtig, net als klasse 4. Qua opleiding is klasse 2 een middelmatige groep. Wel heeft deze klasse veel werkenden. Klasse 2 heeft een aandeel van 20,9% in de onderzoekspopulatie.

Klasse 3: Klasse 3 is een middelmatige klasse. Op geen één variabele scoort ze het hoogst of het laagst. Er kan wel gesteld worden dat klasse 3 redelijk veel economisch kapitaal heeft. Alleen klasse 4 scoort hoger op economisch kapitaal. Klasse 3 heeft wel relatief veel gepensioneerden. Klasse 3 heeft een aandeel van 22,9 in de onderzoekspopulatie.

Klasse 4: klasse 4 is de klasse die het meest kapitaalkrchtig is. Deze klasse heeft het meeste persoonskapitaal en de hoogste kwaliteit van sociale contacten. Het cultureel kapitaal, gemeten met de beheersing van de Friese taal, is wel middelmatig van klasse 4. Op het gebied van economisch kapitaal is klasse 4 de hoogst opgeleide klasse, heeft het hoogste inkomen en heeft de meeste werkenden. Klasse 4 heeft een aandeel van 24,5% in de onderzoekspopulatie.

Klasse 5: Net als klasse 3, is ook klasse 5 een klasse met middelmatige scores op kapitaal. Het meest kenmerkende aan deze klasse is het hoge aantal gepensioneerden met een laag inkomen die allemaal een lage opleiding hebben gehad. Deze klasse heeft een laag aandeel in de onderzoekspopulatie, 12,2%.

Aantal klassen

Allereerst valt op de nieuwe klasse, 5, weinig persoonlijk en sociaal kapitaal heeft. Op cultureel kapitaal scoort het hoog. Wanneer er gekeken wordt naar economisch kapitaal, dan valt op dat klasse 5 uit gepensioneerden bestaat met een laag inkomen en een lage opleiding. Qua klassengrootte is klasse 5 de kleinste klasse.

Er kan gesteld worden dat de vijfde klasse geen goede toevoeging is om de verschillen in de onderzoekspopulatie bloot te stellen. De belangrijkste reden is dat de vijfde klasse weinig verklaart. Dat komt door de onderzoekspopulatie, die oververtegenwoordigd is door ouderen. In de LCA heeft dat gevolg dat er klassen ontstaan die niet representatief zijn voor Friesland. In het geval van de vijfde klasse met alleen maar arme ouderen, zou het kunnen kloppen dat die klasse bestaat in Friesland. Echter, omdat de data teveel uit verhouding staat met de Friese samenleving, wordt er voor gekozen om geen LCA met vijf klassen te doen. In plaats daarvan wordt er gekozen voor vier klassen om verschillen op basis van economisch, sociaal, cultureel en persoonskapitaal te duiden binnen de onderzoekspopulatie. Dit heeft als voordeel ten opzichte van drie klassen dat er wel onderscheidt wordt gemaakt tussen de werkende bevolking en de gepensioneerden.

Wegen van de data

De representativiteit wordt verhoogd door de dataset te wegen. De gewogen LCA wordt anders gedaan dan de ongewogen LCA. In eerste instantie zou het aantal klassen worden onderzocht op basis van de laagste BIC waarde, maar het probleem is dat de BIC-waarde blijft dalen, zelfs bij meer dan 20 klassen. Daarom is het R-script geschreven op een dermate

manier dat het zeven klassen geeft en daaruit kan men handmatig kiezen wat het beste te interpreteren is. In onderstaande output staat het script voor de LCA met de gewogen data.

R-script met gewogen data

```
# De LCA wordt uitgevoerd met de package poLCA. Dit is gedaan op basis van
# de volgende bronnen over LCA in R:

# - https://stats.idre.ucla.edu/mplus/dae/latent-class-analysis/
# - https://maksimrudnev.com/2016/12/28/latent-class-analysis-in-r/
# - https://statistics.ohlsen-web.de/latent-class-analysis-polca/
# - https://www.r-bloggers.com/latent-class-modeling-election-data/

# Allereerst worden de ontbrekende packages binnengehaald welke nodig zijn om
# een LCA te doen en vervolgens worden de packages geladen.

# Installeer ontbrekende packages
packageList <- c("foreign", "stats4", "poLCA")
new.packages <- packageList[!(packageList %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages)

# Laad alle packages
lapply(packageList, require, character.only = TRUE)

# Na het installeren van de benodigde packages worden de bestanden binnengehaald.
# vervolgens worden de bestanden gelabeld als dataset met bijbehorende naam.
# Bestandenlijst
file1 <- "R:\\FSP\\PW\\00 Stagiaires\\Marten Hospes\\statistiek\\20190529mergedData_panelRI_w.sav"
dataset <- read.spss(file1, use.value.labels = TRUE, to.data.frame = TRUE)
# bronbestand <- file.choose() . maar deze zal wel niet nodig meer zijn.

# Het databestand is nu ingeladen. Omdat de data niet representatief is voor Friesland,
# is het de bedoeling om de dataset te manipuleren. Dit gebeurt op basis van de weegfactor die is toegevoegd
# in dataset.
# Uitleg van wat hier onder gebeurt:
#
# Allereerst wordt dataset omgedoopt tot origineel. Dit wordt gedaan
# om later de dataset naar zijn origineel kunnen brengen.
#
# Na het benoemen wordt de weegfactor afgerond naar 2 decimalen en vermenigvuldigt met 100. Op deze
# manier ontstaat er een vol getal zonder decimalen (de originele weegfactor was een getal tussen 0 en 12).
#
# Vervolgens krijgt elke rij in de dataset een uniek nummer. Elke rij (oftewel: elke respondent)
# wordt daarna 'vermenigvuldigt' met de weegfactor. Dat betekent dat er veel grotere dataset ontstaat.
# Dus iemand die als originele weegfactor 1,21 heeft, wordt nu 121 meegenomen in de LCA.

origineel <- dataset
dataset$w_rounded <- round(dataset$w, digits=2) * 100
```

```

dataset$rijnummer <- seq.int(nrow(dataset))

duplicaten <- as.data.frame(rep(dataset$rijnummer,dataset$w_rouded))
names(duplicaten) <- "rijnummer"
dataset<- merge(duplicaten, dataset)

# We gaan nu kijken naar de
# losse variabelen. Er wordt gebruik gemaakt van de poLCA. Deze kan alleen met
# numerieke variabelen omgaan. voor het gemak worden alle variabelen numeriek
# gemaakt.

#variabelen numeric maken
dataset$Zelfregie_schaal.num <- as.numeric(dataset$Zelfregie_schaal)
dataset$Kwalcontact_schaal.num <- as.numeric(dataset$Kwalcontact_schaal)
dataset$Digitaalcontact_schaal.num <- as.numeric(dataset$Digitaalcontact_schaal)
dataset$Fysiekcontact_schaal.num <- as.numeric(dataset$Fysiekcontact_schaal)

dataset$gezondheid.num <- as.numeric(dataset$zorg_V004a)

dataset$Friessprekend.num <- as.numeric(dataset$duur_spreektFries)
dataset$Werk_werkloos_pensioen.num <- as.numeric(dataset$duur_Werk)
dataset$Ink_laag_midden_hoog_weetniet.num <- as.numeric(dataset$duur_ink_hh)
dataset$Opleiding.num <- as.numeric(dataset$duur_opleiding)

## LCA ##
# Nu kan de LCA gedraaid worden.
# De data voor de LCA is afkomstig van dataset 'lcaset'.
# Dit kan de gewogen dataset zijn of de ongewogen.
# met onderstaande kan dat aangepast worden.

lcaset<-dataset
#lcaset<-origineel

# 1. testdraai met twee klassen.
lca.poLCA <- poLCA(
  cbind(Zelfregie_schaal.num,
        gezondheid.num,
        Kwalcontact_schaal.num,
        Digitaalcontact_schaal.num,
        Fysiekcontact_schaal.num,
        Friessprekend.num,
        Opleiding.num,
        Werk_werkloos_pensioen.num,
        Ink_laag_midden_hoog_weetniet.num)~1,
  data=lcaset,
  nclass=8,
  nrep=1,
  na.rm=T,
  #graphs=T,
  maxiter = 100000)

```

```

# We hebben nu een latente klassen analyse gedaan, waarmee een vast aantal
# klassen wordt gegeven. Het wordt echter interessanter als we een analyse doen
# waarbij wordt berekend wat het optimale model is. Dit wordt berekend
# met de Bayesian Information Criterion (BIC). Het model met de laagste BIC
# waarde, is statistisch gezien het beste model.

# 2. model met laagste BIC-waarde

oldBIC = 100000
newBIC = 100000
modellen <- NULL
aantalKlassen <- 2

while(newBIC<=oldBIC){
  modelnr <- aantalKlassen - 1

  modellen[[modelnr]] <- polCA(
    cbind(Zelfregie_schaal.num,
          gezondheid.num,
          Kwalicontact_schaal.num,
          Digitaalcontact_schaal.num,
          Fysiekcontact_schaal.num,
          Friessprekend.num,
          Opleiding.num,
          Werk_werkloos_pensioen.num,
          Ink_laag_midden_hoog_weetniet.num
    )~1,
    data=lcaset,
    nclass=aantalKlassen,
    nrep=1,
    na.rm=T,
    graphs=T,
    maxiter = 100000)

  aantalKlassen = aantalKlassen + 1
  oldBIC <- newBIC
  newBIC <- modellen[[modelnr]]$bic
}

# plot model met laagste BIC (één na laatste model)
plot(modellen[[length(modellen)-1]])

# plot BIC's van alle modellen
for (i in 1:length(modellen)){cat(paste0("Aantal clusters: ", i + 1, "\nBIC: ", modellen[[i]]$bic, "\n\n")) }

# Zoeken naar het beste model op basis van de laagste BIC-waarde, heeft
# als nadeel dat stopt wanneer de laagste BIC is gevonden. Het zou echter
# mogelijk zijn dat er nog een model is met nog een lagere BIC waarde, maar dat dat
# model zich pas manifesteert bij 10 klassen. Daarom wordt de volgende analyse gedaan:

```



```
# 3. genereer serie modellen binnen cluster-range (Zit fout in)

maxaantalklassen<-7
modellen2 <- NULL

for (model in 2:maxaantalklassen){
  modelnr2 <- model - 1

  modellen2[[modelnr2]] <- poLCA(
    cbind(Zelfregie_schaal.num,
          gezondheid.num,
          Kwaliccontact_schaal.num,
          Digitaalcontact_schaal.num,
          Fysiekcontact_schaal.num,
          Friessprekend.num,
          Opleiding.num,
          Werk_werkloos_pensioen.num,
          Ink_laag_midden_hoog_weetniet.num)~1,
    data=lcaset,
    nclass=model,
    nrep=1,
    na.rm=T,
    maxiter = 100000)
}

for (model in 1:length(modellen2)){
  print(paste0("classes: ", model + 1, " BIC: ", modellen2[[model]]$bic))
}

modellen2[[3]]
```

LCA met 4 klassen

In de ongewogen LCA is er gekozen voor vier klassen om verschillen te verklaren. Dit lijkt voor de gewogen LCA ook een goede manier, omdat minder klassen ruim te interpreteren zijn. Dit zou tot gevolg hebben dat er geen relevante uitspraken kunnen worden gedaan over Friesland.

Om de verschillen te verklaren tussen ongewogen en gewogen data, wordt de gewogen LCA gedaan met vier klassen. Dit levert de volgende output op:

```
Conditional item response (column) probabilities,
by outcome variable, for each class (row)

$zelfregie_schaal.num
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5) Pr(6)
class 1: 0.0025 0.0720 0.2831 0.4185 0.2072 0.0167
```

```

class 2: 0.0000 0.0022 0.0668 0.2940 0.4996 0.1374
class 3: 0.0000 0.0000 0.4788 0.5212 0.0000 0.0000
class 4: 0.0007 0.0051 0.0874 0.2819 0.5432 0.0818

```

\$gezondheid.num

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.003 0.1013 0.2566 0.5908 0.0483
class 2: 0.000 0.0153 0.0654 0.7700 0.1493
class 3: 0.000 0.4526 0.0000 0.5474 0.0000
class 4: 0.000 0.0039 0.1260 0.6760 0.1941

```

\$kwalcontact_schaal.num

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0117 0.1361 0.4034 0.4294 0.0194
class 2: 0.0013 0.0000 0.1594 0.4720 0.3674
class 3: 0.0000 0.0000 1.0000 0.0000 0.0000
class 4: 0.0094 0.0764 0.2857 0.3962 0.2321

```

\$Digitaalcontact_schaal.num

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.4462 0.4078 0.1023 0.0437 0.0000
class 2: 0.0119 0.3921 0.4901 0.0998 0.0061
class 3: 0.0000 1.0000 0.0000 0.0000 0.0000
class 4: 0.3826 0.4794 0.1373 0.0007 0.0000

```

\$Fysiekcontact_schaal.num

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1642 0.4492 0.3283 0.0583 0.000
class 2: 0.0000 0.1107 0.7485 0.1408 0.000
class 3: 0.4259 0.5741 0.0000 0.0000 0.000
class 4: 0.2582 0.4483 0.2608 0.0118 0.021

```

\$Friessprekend.num

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1330 0.1169 0.0155 0.0937 0.6409
class 2: 0.0873 0.1507 0.1038 0.0966 0.5616
class 3: 0.0272 0.0000 0.9728 0.0000 0.0000
class 4: 0.1754 0.2544 0.0683 0.0629 0.4390

```

\$Opleiding.num

```

      Pr(1) Pr(2) Pr(3)
class 1: 0.2832 0.4346 0.2821
class 2: 0.0339 0.4940 0.4721
class 3: 0.0000 0.0000 1.0000
class 4: 0.0000 0.1489 0.8511

```

\$werk_werkloos_pensioen.num

```

      Pr(1) Pr(2) Pr(3)
class 1: 0.3205 0.2996 0.3799
class 2: 0.7352 0.1721 0.0927
class 3: 0.0000 0.9467 0.0533
class 4: 0.7920 0.0091 0.1989

```

\$Ink_laag_midden_hoog_weetniet.num

```

      Pr(1) Pr(2) Pr(3) Pr(4)
class 1: 0.4697 0.1676 0.0730 0.2897
class 2: 0.2775 0.1670 0.4075 0.1481
class 3: 0.4259 0.0267 0.5474 0.0000
class 4: 0.0186 0.0381 0.7659 0.1774

```

Estimated class population shares

```
0.3209 0.41 0.0126 0.2565
```

Predicted class memberships (by modal posterior prob.)
 0.3105 0.4302 0.0131 0.2463

```
=====
Fit for 4 latent classes:
=====
```

```
number of observations: 104835
number of estimated parameters: 131
residual degrees of freedom: 104704
maximum log-likelihood: -1044514
```

```
AIC(4): 2089291
BIC(4): 2090543
G^2(4): 764332.3 (Likelihood ratio/deviance statistic)
X^2(4): 73594615 (Chi-square goodness of fit)
```

Beschrijving van de variabelen

Zelfregie: Op de variabele zelfregie scoort klasse 4 het hoogst en klasse 1 en 3 het laagst.

Gezondheid: De ervaren gezondheid is lastig te bepalen. De scores vallen hoofdzakelijk op de goede gezondheid, de één na hoogste categorie. Klasse 4 scoort het hoogst op zeer goede gezondheid en klasse 2 het hoogst op goede gezondheid.

Kwaliteit van contacten: Op de kwaliteit van contacten is er vooral een uitschieter naar beneden: klasse 1. De andere drie groepen liggen dicht bij elkaar. Er kan voorzichtig gesteld worden dat klasse 4 hoger scoort op de kwaliteit van contacten.

Kwantiteit van contacten: Op de kwantiteit van contacten scoort klasse 1 het laagst en klasse 3 en 4 het hoogst.

Beheersing van het Fries: klasse 3 heeft de beste beheersing van de Friese taal en klasse 4 de minste beheersing.

Opleiding: klasse 1 heeft de hoogste score op lage opleiding. Op de hoge opleiding scoort klasse 2 verreweg het hoogst.

Arbeidssituatie: kijkend naar de arbeidssituatie valt op dat er vooral twee werkende klassen zijn en een gepensioneerde groep. Klasse 1 is de groep met de hoogste score op gepensioneerden. Klasse 4 heeft ook een relatief hoge score in verhouding tot klasse 2 en 3 op

gepensioneerden. Klasse 2 en 4 hebben de hoogste score op werkenden. Klasse 3 heeft de laagste score op werkenden

Inkomen: klasse 1 en 3 scoren het hoogst op het laagste inkomen. Op het hoogste inkomen scoort klasse 4 het hoogst, gevolgd door klasse 3.

Beschrijving van de groepen

Klasse 1: klasse 1 heeft van de vier klasen het minste economisch, sociaal, cultureel en persoonskapitaal. Klasse 1 heeft een omvang van 32,1% in de steekproef.

Klasse 2: klasse 2 heeft middelmatige scores op economisch, sociaal, cultureel en persoonskapitaal. Op geen van de variabelen heeft deze klasse de laagste score of de hoogste score. Klasse 2 valt vooral op door het hoge aantal werkenden en het grote aandeel in de onderzoekspopulatie: 41%.

Klasse 3: klasse 3 is lastig te duiden. Allereerst door de kleine omvang van 1,2%, maar ook door de scores op de variabelen. Zo scoort het zowel hoog op het lage inkomen als op het hoge inkomen. Klasse 3 bestaat uit louter hoogopgeleiden, maar heeft ook weer het hoogste aandeel op niet-werkenden. Klasse 3 lijkt een klasse te zijn die niet optimaal naar voren komt door het wegen van de data.

Groep 4: klasse 4 is de groep met het meeste economisch, sociaal en persoonlijk kapitaal. Deze klasse heeft, net als klasse 2, veel werkenden. Dit lijkt erop te duiden dat er twee klassen werkenden zijn, waarvan klasse 4 de klasse is met het meeste kapitaal.

Vragenlijst

CONCEPTVRAGENLIJST VERSCHILLEN IN FRYSLAN

Persoonskapitaal

1 fysiek kapitaal

- a. subjectieve gezondheid: ‘wat vindt u, over het algemeen genomen, van uw gezondheid’? (uitstekend; zeer goed; goed; matig; slecht)
- b. beperking bij het trappenlopen: ‘wordt u door uw gezondheid beperkt bij het trappenlopen’? (ja, ernstig beperkt; ja een beetje beperkt; nee)

2 mentaal kapitaal

- a. zelfvertrouwen: ‘ik heb veel zelfvertrouwen’ (tienpuntsschaal: 1=zeer mee oneens, 10=zeer mee eens)
- b. het beeld dat iemand van zichzelf heeft: ‘ik heb een negatief beeld van mijzelf’ (tienpuntsschaal: 1=zeer mee oneens, 10=zeer mee eens)
- c. depressiviteit: ‘heeft u in de afgelopen 12 maanden een periode gehad waarin u erg somber of depressief was, minstens 2 weken achter elkaar’? (ja, nee).

3 esthetisch kapitaal

- a. inschatting van het eigen uiterlijk: ‘ik vind mijn uiterlijk zoals het is precies goed’ (tienpuntsschaal: 1=zeer mee oneens, 10=zeer mee eens)
- b. inschatting van wat anderen van het uiterlijk vinden: ‘de meeste anderen vinden mij er goed uitzien’ (tienpuntsschaal: 1=dat vindt niemand, 10=dat vindt iedereen).

Economisch kapitaal

Ook bij het economisch kapitaal onderscheiden we vier subdimensies.

1. opleiding en vaardigheden: het opleidingsniveau in vier groepen

- a. Basisschool / lagere school niet afgemaakt (speciaal onderwijs telt ook mee ; Alleen basisschool / lagere school afgemaakt (speciaal onderwijs telt ook mee)
- b. lbo, vbo, leao, lts, ambachtsschool, huishoudschool, lhno, vmbo (basisberoepsgericht, kaderberoepsgericht, gemengd) afgemaakt; mulo, ulo, mavo, vmbo (theoretische leerweg); havo jaar 3-4; vwo jaar 3-5 afgemaakt; mbo niveau 1 (duur < 2 jaar) afgemaakt
- c. havo, mms, msvm afgemaakt; vwo, hbs, atheneum, gymnasium afgemaakt; mbo niveau 2 en 3 afgemaakt; kort mbo; leerlingwezen (duur 2-3 jaar); mbo niveau 4 afgemaakt (duur 4 jaar); mbo-plus voor havisten afgemaakt
- d. propedeuse wo afgemaakt of ou-certificaat; korte hbo-opleiding einddiploma (2 of 3 jaar); hbo associate degree; bachelor hbo afgemaakt; bv kweekschool, mo-akte a, heao, hts, pedagogische academie, sociale academie; bachelor of kandidaats universiteit afgemaakt; hbo: master's degree, tweede fase opleidingen; post hbo-opleidingen, pre-master onderwijs voor hbo; mo-akte B (eerstegraads onderwijsbevoegdheid); wo/universiteit: master's degree, tweede fase opleidingen; ingenieur, meester, doctorandus doctoraat/gepromoveerd.

2. arbeid

- a. de arbeidssituatie (in zes groepen)
 - i. niet werkend (nooit gewerkt; nu geen betaalde baan; niet op zoek)
 - ii. werkzoekend (nu geen betaalde baan; wél op zoek)
 - iii. pensioen (met pensioen of pre-pensioen, vut)
 - iv. student
 - v. loondienst (overheid of semi-overheid (bv. onderwijs, gezondheidszorg); commercieel bedrijf; niet op winst gerichte instelling)
 - vi. zelfstandige of freelancer.

3. Inkomen

- a. gestandaardiseerd besteedbaar huishoudensinkomen (door het CBS geleverde registervariabele, in percentielen en vervolgens in 6 groepen verdeeld: 0-9;10-29; 30-49; 50-89; 90-98; 99).

4. (financiële) vermogen

- a. vrij besteedbaar vermogen (door het CBS geleverde registervariabele, in 6 groepen verdeeld: negatief vermogen; klein vermogen (0-5.000 euro); middelgroot vermogen (5.000-50.000 euro; groot vermogen (50.000-500.000 euro); zeer groot vermogen (>500.000 euro).
- b. vermogen gerelateerd aan de woning (door het CBS geleverde registervariabele, in 5 groepen verdeeld: negatief vermogen; huurwoning; positief vermogen onderste 33%; positief vermogen middelste 33%; positief vermogen bovenste 33%).

Cultureel kapitaal

In het cultureel kapitaal onderscheiden we drie subdimensies.

1. taal en communicatie

- a. beheersing van de Engelse taal: “Beheerst u de Engelse taal”?
 - i. spreekt geen Engels (nee);
 - ii. beheerst het Engels een beetje (ja, een paar woorden (bijv. iets bestellen in een café met behulp van een reisgids); ja, genoeg om te praten over simpele en alledaagse zaken (bijv. de weg vragen);
 - iii. beheerst het Engels redelijk (ja, genoeg om zonder voorbereiding een gesprek te voeren over dagelijkse onderwerpen (bijv. over familie, hobby’s); ja, genoeg om de kern van een ingewikkelde tekst te begrijpen, en een spontaan gesprek te voeren met iemand die Engels als moedertaal heeft;
 - iv. beheerst het Engels goed (ja, genoeg om ingewikkelde en lange teksten helemaal te begrijpen, en mijzelf vloeiend uit te drukken in een gesprek met iemand die Engels als moedertaal heeft; ja, genoeg om zonder voorbereiding aan elk gesprek of discussie deel te nemen en een ingewikkeld betoog te houden);
 - v. Engels is mijn moedertaal.

b. ICT-vaardigheden

Dan volgen nu enkele vragen over uw computervaardigheden. Het gaat er alleen om of u dit zou KUNNEN doen en niet of u het daadwerkelijk doet. **Kunt u:**

- i. werken met een tekstverwerkingsprogramma (bv. Word)
- ii. een programma installeren op een computer
- iii. de beveiliging op een computer regelen (met bv. een virusscanner of een firewall).

Op basis van deze drie vragen is samengestelde indicator voor basale ICT-vaardigheden gemaakt, door de antwoorden op te tellen (minimale waarde van 0 – iemand beheerst geen van de vaardigheden – tot 3 – iemand beheerst alle vaardigheden).

2. smaken, voorkeuren en culturele kennis

a. leefstijl

- i. buitenlandse vakantie: “Naar welk land ging de verste vakantiereis die u in 2013 heeft gemaakt? *Een vakantie is een verblijf buiten uw eigen woning voor tenminste drie nachten achtereen.* (ik ben in 2013 niet op vakantie geweest; Nederland; buitenland, namelijk....)
- ii. duur uit eten: “Bent u in 2013 uit eten geweest in een restaurant waar u voor meer dan 100 euro per persoon heeft gegeten (dus exclusief drank)?” (nee; ja, 1 keer; ja 2-3 keer; ja, 3-4 keer; ja, 5-10 keer; ja, meer dan 10 keer)
- iii. bezoek aan ‘hogere’ cultuur: “Hoe vaak bent u in de afgelopen 12 maanden naar klassieke concerten, toneel of (kunst)musea geweest?” (1 keer per week of vaker; 1 keer per twee weken; 1 keer per maand; paar keer per jaar; zelden of nooit).

Vervolgens is voor leefstijl een samengestelde indicator gemaakt, met behulp van homogeniteitsanalyse. De eigenwaarde is 0,47 en de discriminatiemaat is voor buitenlandse vakantie 0,50; voor duur uit eten 0,36; en voor bezoek aan cultuur 0,54 (de discriminatiemaat is te vergelijken met factorladingen).

- a. 3. symbolische aspecten, zoals reputatie, titels of eretekens
 - a. hiervoor hebben we **geen meting** beschikbaar.

Sociaal kapitaal

Bij sociaal kapitaal onderscheiden we twee subdimensies:

1. sociale steun

- a. kwantiteit van de sociale steun: “Hoe vaak heeft u contact (d.w.z. ontmoetingen, telefonische en schriftelijke contacten, e-mail etc.) met familieleden (die niet bij u in huis wonen); vrienden, vriendinnen of echt goede kennissen; buren en buurtbewoners” (6=1 keer per week of vaker; 5=1 keer per 2 weken; 4=1 keer per maand; 3=minder dan 1 keer per maand; 2=zelden of nooit; 1=heb ik niet). Op basis van deze vragen is een samengestelde indicator gemaakt, waarbij de antwoorden zijn opgeteld en vervolgens in 5 groepen verdeeld: zeer weinig contact (score van maximaal 9); weinig contact (score 10-12); geregeld contact (score 13-15); veel contact (score 16-17); zeer veel contact (score 18).
- b. kwaliteit van de sociale steun: “Met hoeveel mensen (die niet bij u in huis wonen) besprak u in de afgelopen zes maanden belangrijke persoonlijke zaken? *Het gaat hier niet om mensen die beroepsmatig hulp aan u verlenen, bijv. een huisarts*” (met niemand, met 1 persoon, met 2 of 3 personen; met 4 of 5 personen; met meer dan 5 personen).

2. instrumentele steun

- a. de mate waarin iemand kapitaalkrachtige of invloedrijke mensen kent: “Heeft u momenteel vrienden, kennissen of familieleden met één van de volgende beroepen? *Het gaat erom dat u ze persoonlijk kent, buiten hun werk om*”. (een burgemeester of lid van de Tweede Kamer; een arts of advocaat; een directeur van een commercieel bedrijf met meer dan tien mensen in dienst; een hoge ambtenaar, bv een gemeentesecretaris of directeur op een ministerie; een professioneel musicus, artiest of schrijver) (ja, nee). Op basis hiervan is een samengestelde indicator gemaakt door het aantal keer dat iemand ‘ja’ zegt op te tellen.

