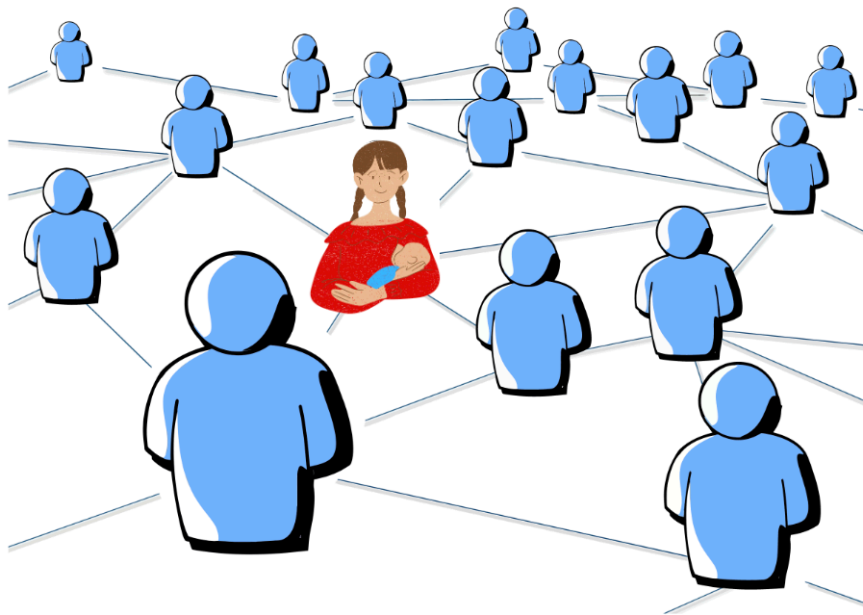




rijksuniversiteit
groningen

Persoonlijke netwerktypen van Nederlandse vrouwen:

Een onderzoek naar de samenhang met haar kinderwens



Sanne Boer | S4460006 | s.boer.2@student.rug.nl

Masterscriptie Sociologie | Rijksuniversiteit Groningen

Begeleider: Dr. G. Stulp | Referent: Dr. M. Huisman

Aantal woorden (exclusief bijlagen): 21.170

Januari 2023 – Juli 2023

Engelse titel: *“Personal network types of Dutch woman:
Exploring the relationship with Fertility Desires”*

Samenvatting

Sinds 1970 krijgen Nederlandse vrouwen minder kinderen dan het reproductiegetal, wat problematisch is voor een evenwichtige leeftijdsverdeling in een gezonde bevolking (CBS, 2021). Deze afname wordt beïnvloed door zowel de persoonlijke omstandigheden van de vrouw (Balbo et al., 2013) als door de nationale ontwikkelingen (Balbo et al., 2013). In deze scriptie wordt echter naar een derde niveau gekeken, namelijk de invloed van sociale contacten van de vrouw op haar kinderwens. Het doel is enerzijds om de grote hoeveelheid aan netwerkdata over Nederlandse vrouwen systematisch te ordenen en anderzijds om een bijdrage te leveren aan de kennis over de invloed van sociale contacten op fertilitateitswensen. Op basis van voorgaand onderzoek over het ordenen van grote hoeveelheden data, theorieën over sociaal kapitaal en mechanismen van invloed op fertilitateitskeuzes worden de methode gekozen en de hypothesen opgesteld.

De netwerkdata is afkomstig uit de studie ‘Sociale Relaties en Kinderkeuzes’ van het LISS-panel. Het betreft een representatieve steekproef van 1332 vrouwen, waarvan 758 vrouwen de studie met succes afronden. Na correctie van de data blijven 706 vrouwen over en deze vrouwen zijn goed vergelijkbaar met de vrouwen die de studie niet hebben afgerond (Stulp, 2020). Om deze netwerkdata systematisch te ordenen is gebruik gemaakt van de op subgroep gebaseerde clustermethode van Vacca (2020). Daarnaast wordt een tweede model gemaakt; deze neemt de eerdere structurele kenmerken mee evenals de compositie kenmerken over welke sociale relaties een vrouw heeft.

De toepassing van de op subgroep gebaseerde clustermethode met structurele kenmerken leidt tot de meest optimale indeling in zes netwerktypen. De netwerktypen die worden geïdentificeerd zijn: strongly bi-factional, weakly bi-factional, strongly tri-factional, weakly tri-factional, sparse tri-factional en dense multi-factional. Er wordt geen verschil gevonden tussen de vrouwen met verschillende netwerktypen in de kinderwens en de ideale gezinsgrootte die zij hebben. Daarna bleek dat het tweede model, met zowel structurele als compositie kenmerken, complexer te zijn maar de indeling in netwerktypen niet nauwkeuriger te kunnen uitvoeren. Hieruit blijkt dat de op subgroep gebaseerde methode de informatie die uit deze compositie kenmerken komt goed weet te vangen.

Een relatief klein aandeel van Nederlandse vrouwen heeft een persoonlijk netwerk met één subgroep. Oftewel, het grootste deel heeft verschillende sociale kringen waarin zij zich bevindt. Dit verschilt aanzienlijk van de netwerktypen die worden geïdentificeerd door Vacca (2020). Dit lijkt te maken te hebben met de nationale context. Toekomstig onderzoek zou in kunnen gaan op de invloed van de nationale context op de structurele netwerktypen, het verschil in uitkomst als fertilitateitskeuzes worden onderzocht in plaats van fertilitateitswensen en de meerwaarde van andere compositie kenmerken.

Sleutelwoorden: kinderwens, persoonlijke netwerkstudie, clustermethode, subgroepen, structurele kenmerken, compositie kenmerken, LISS-panel.

Inhoudsopgave

| | |
|--|-----------|
| Samenvatting | 2 |
| 1. Inleiding | 5 |
| <i>1.1 Introductie</i> | 5 |
| 1.1.1 Onderzoek naar persoonlijke netwerken | 6 |
| <i>1.2 Probleemstelling</i> | 7 |
| <i>1.3 De huidige studie</i> | 8 |
| 1.3.1 De methode | 8 |
| 1.3.2 Wetenschappelijke relevantie | 9 |
| 1.3.3 Sociologische relevantie | 10 |
| 2. Theoretisch kader | 11 |
| <i>2.1 Ordenen van persoonlijke netwerkdata</i> | 11 |
| 2.1.1 Meerwaarde van een netwerktypologie | 11 |
| 2.1.2 Soorten structurele netwerktypologieën | 13 |
| <i>2.2 Invloed van sociale contacten op fertiliteitswensen</i> | 17 |
| 2.2.1 Mechanismen van invloed door sociale contacten | 17 |
| 2.2.2 De netwerk compositie en sociaal kapitaal | 18 |
| 2.2.3 De netwerkstructuur en sociaal kapitaal | 19 |
| 2.2.4 Meerwaarde compositie van het netwerk | 21 |
| <i>2.3 Huidige studie</i> | 22 |
| 3. Data & Methoden | 24 |
| <i>3.1 Beschrijving dataset</i> | 24 |
| 3.1.1 Het LISS-panel: Sociale relaties en kinderkeuzes | 24 |
| 3.1.2 De dataset | 25 |
| <i>3.2 Operationalisering centrale begrippen</i> | 26 |
| 3.2.1 De uitkomstmaten | 26 |
| 3.2.2 Beschrijvende statistieken van Nederlandse vrouwen | 26 |
| 3.2.3 Beschrijvende statistieken voor clusters | 27 |
| 3.2.4 Variabelen voor de clusteranalyse | 28 |
| <i>3.3 Methode van analyse</i> | 29 |
| 3.3.1 Beschrijving methode van Vacca (2020) | 29 |
| 3.3.2 Beschrijving methode van huidige studie | 31 |

| | |
|---|-----------|
| 4. Resultaten | 32 |
| 4.1 <i>Beschrijvende statistieken</i> | 32 |
| 4.1.1 Univariate verdeling structurele kenmerken | 32 |
| 4.1.2 K-medoid analyse, clusters en medoids | 33 |
| 4.1.3 Bivariate verdeling structurele kenmerken per netwerktype | 35 |
| 4.2 <i>Definiëren van de netwerktypologie</i> | 37 |
| 4.2.1 Vergelijking met netwerktypen Vacca (2020) | 37 |
| 4.2.2 Netwerktypologie Nederlandse vrouwen | 39 |
| 4.3 <i>Multivariate analyses</i> | 40 |
| 4.3.1 Kenmerken van de vrouwen in elk netwerktype | 40 |
| 4.3.2 Relatie tussen structurele netwerktypen en fertiliteitswensen | 42 |
| 4.4 <i>Toevoeging van compositie kenmerken</i> | 45 |
| 4.4.1 Meerwaarde van compositie kenmerken | 45 |
| 4.4.2 Compositie binnen structurele netwerktypen | 46 |
| 5. Conclusie | 50 |
| 5.1 <i>Structurele netwerktypen Nederlandse vrouwen</i> | 50 |
| 5.2 <i>Samenhang tussen structurele netwerktypen en fertiliteitswensen</i> | 51 |
| 5.3 <i>Meerwaarde compositie kenmerken</i> | 52 |
| 6. Discussie | 55 |
| 6.1 <i>Interpretatie van resultaten</i> | 55 |
| 6.2 <i>Methodologische beperkingen</i> | 57 |
| 6.3 <i>Wetenschappelijke bijdrage en vervolgonderzoek</i> | 58 |
| Referenties | 60 |
| Bijlagen | 65 |
| <i>Bijlage I: Identificatie van netwerktypen</i> | 65 |
| <i>Bijlage II: Uitwerking typologie structurele en compositie kenmerken</i> | 67 |
| <i>Bijlage III: Assumpties uitgevoerde toetsen</i> | 70 |
| <i>Bijlage IV: R-code voor analyses</i> | 76 |

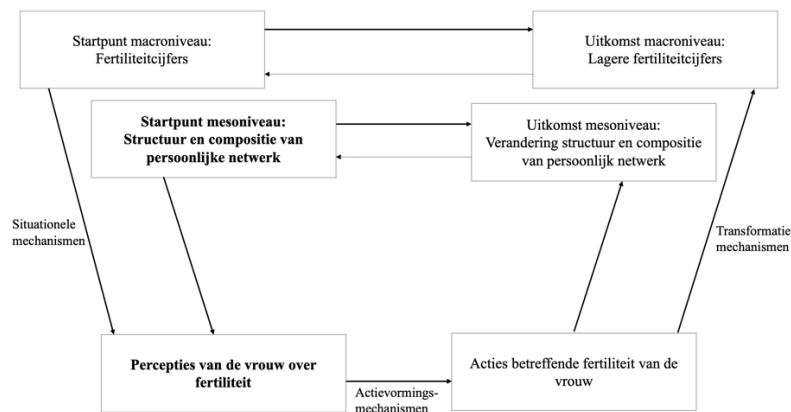
1. Inleiding

1.1 Introductie

In een groot deel van Europa is sprake van een trend waarin vrouwen steeds minder kinderen krijgen (EuroStat, 2023). Dit geldt ook voor Nederland en deze daling zorgt voor een andere leeftijdsverdeling. Dat er sprake is van een evenwichtige leeftijdsverdeling is belangrijk om hoge vergrijzingcijfers en een krimpende beroepsbevolking te voorkomen. Zo kregen Nederlandse vrouwen in 1950 nog gemiddeld 3,1 kinderen, maar ongeveer zeventig jaar later is dit gemiddelde gedaald naar 1,6 kinderen (CBS, 2022). Hiermee ligt het gemiddeld aantal kinderen aanzienlijk lager dan het reproductiegetal van 2,1. Het reproductiegetal is het gemiddeld aantal kinderen dat per vrouw geboren moet worden om de leeftijdsverdeling binnen een bevolking gelijk te houden (Craig, 1994). Het aantal kinderen dat wordt geboren is één van de factoren die de demografie beïnvloedt, maar ook sterfte- en migratiecijfers spelen hierbij een rol. Naast het aantal kinderen dat per vrouw wordt geboren verandert ook de leeftijd waarop vrouwen kinderen krijgen. In de periode van 1950 tot 2021 is de gemiddelde leeftijd van 26,4 jaar gestegen naar 30,3 jaar (CBS, 2022). Deze trends spelen zich af in een tijd waarin Nederland een geavanceerde samenleving wordt. Met betrekking tot fertiliteit houdt dit in dat het gebruik van anticonceptiemiddelen de norm is waardoor de keuzes rondom ouderschap grotendeels actief worden genomen.

Toch ligt de behoefte om een kind te krijgen in Nederland nog altijd hoog. Ongeveer 91% van de Nederlandse vrouwen die in 2020 de leeftijd van 50 jaar hebben bereikt heeft kinderen of wilde kinderen (CBS, 2021). Tegelijkertijd spelen de fertiliteitstrends. De verklaring voor deze fertiliteitstrends worden op verschillende niveau's onderzocht, namelijk: het macro-, meso- en micro niveau. Een veel gebruikt verklaringsmodel in de sociologie is de 'Coleman Boat'. Dit verklaringsmodel visualiseert hoe de verschillende niveaus van invloed zijn op elkaar. Om te verduidelijken hoe deze invloed plaatsvindt zijn verklarende mechanismen toegevoegd tussen de verschillende niveaus (Hedström & Ylikoski, 2010). In het onderzoek naar fertiliteit richt het microniveau zich op het koppel of de individuele man of vrouw. Op dit niveau wordt onderzocht wat de relatie is tussen levensloop omstandigheden, zoals opleiding, werk en partnerschap, en fertiliteitskeuzes (Balbo et al., 2013). Daarboven bevindt zich het macroniveau. Op dit niveau wordt gekeken naar de samenleving waarin de individuen zijn ingebed en hoe deze omstandigheden de fertiliteitskeuzes beïnvloedt. De belangrijkste factoren in dit onderzoeksgebied zijn economische trends en beleidsmaatregelen (Balbo et al., 2013). Het derde, en meest recent toegevoegde, niveau is het mesoniveau. Dit niveau beschrijft de effecten van sociale netwerken in de macro-micro-macro link, het netwerkperspectief (McCarty et al., 2019). Dit perspectief veronderstelt dat individuen

onderdeel zijn van een sociaal netwerk en dat het netwerk invloed heeft op de keuzes die individuen hierbinnen maken (Granovetter, 1985). Oftewel, individuen zijn ingebed in een sociaal systeem, waardoor er een onderlinge afhankelijkheid bestaat tussen de individuen. Het onderzoek binnen het netwerkperspectief richt zich op de invloed van sociale netwerken (Borgatti et al., 2009; Burt, 1998; Granovetter, 1985; Wiertz & de Graaf, 2022). Ongeveer dertig jaar geleden werd voor het eerst beargumenteerd dat de omgang met sociale contacten invloed heeft op de persoonlijke beslissingen met betrekking tot het krijgen van kinderen (Bongaarts & Watkins, 1996; Montgomery & Caterline, 1996). Deze scriptie is geschreven binnen het netwerkperspectief, met als centrale onderwerp de invloed van sociale contacten op de kinderwens van de vrouw. Om dit te verduidelijken is de Coleman Boat toegevoegd in Figuur 1.



Figuur 1: Coleman Boat inclusief mesoniveau fertilitcijfers vrouw, gebaseerd op figuur 24.3 Wiertz & de Graaf (2022)

1.1.1 Onderzoek naar persoonlijke netwerken

Binnen de sociale netwerkanalyse kan een onderscheid worden gemaakt in twee soorten analyses, namelijk een hele netwerk studie en de persoonlijke netwerk studie. Een hele netwerkstudie wordt vaak gedaan in een sociale context, bijvoorbeeld binnen een bedrijf, een marktonderzoek of een criminele organisatie. Daarentegen gaat het in een persoonlijke netwerkstudie om een bepaalde doelgroep, bijvoorbeeld Turkse migranten of eerste klas middelbare scholieren. In dit onderzoek zal sprake zijn van een persoonlijke netwerkstudie met als doelgroep Nederlandse vrouwen.

De individuen in een persoonlijke netwerkstudie worden in de netwerkliteratuur aangeduid als *ego*. Van deze individuen worden de persoonlijke netwerken in kaart gebracht, dit is het netwerk rondom één persoon (McCarty et al., 2019). De sociale contacten van deze persoon zijn de *alters*. Het persoonlijke netwerk van ego bestaat uit de relaties tussen de ego en alters en uit de relaties tussen de alters onderling. Een persoonlijk netwerk heeft verschillende kenmerken, zoals de structuur en de compositie. De structuur van een persoonlijk netwerk bestaat uit de onderlinge relaties tussen de

sociale contacten van ego, de *alter-alter* relaties. Zo kan er sprake zijn van een verbonden netwerk of van een gesegmenteerd netwerk. Hierbij gaat het om de hoeveelheid aan sociale relaties die er zijn onder de alters. Daarnaast is de compositie van het persoonlijke netwerk gericht op de netwerksamenstelling, ook wel kenmerken van de relatie tussen *ego-alter* of kenmerken van de *alter* zelf. Bijvoorbeeld een persoonlijk netwerk met een meerderheid aan familieleden of hoogopgeleiden.

Een uitdaging van persoonlijke netwerken is om de grote hoeveelheid aan data op de juiste manier te analyseren (Gianella & Fisher, 2016). In een persoonlijke netwerkstudie is vaak sprake van verschillende typen data, namelijk data over demografische kenmerken van *ego*, data over de relaties tussen *ego* en *alters*, data over demografische kenmerken van *alters* en data over onderlinge *alter-alter* relaties. Daarnaast is deze data beschikbaar voor meerdere persoonlijke netwerken. Dit vormt samen een grote hoeveelheid aan data. Zo is er bij een ongericht netwerk met 20 *alters* sprake van 190 mogelijke alter-alter relaties. Bij een steekproef van 300 individuen resulteert dit in 57.000 mogelijke relaties die onderzocht moeten worden. Dit voorbeeld geeft weer dat alleen de data over *alter-alter* relaties veel data oplevert, hierbij zijn nog niet de ander eerder genoemde typen data meegenomen.

Een manier om grote hoeveelheden aan netwerkdata te analyseren is het uitvoeren van een clusteranalyse. De clusteranalyse is een methode om de netwerken van individuen te groeperen op basis van kenmerken van het netwerk met als doel het identificeren van verschillende clusters (Bidart et al., 2018; Giannella & Fisher, 2016; Vacca, 2020). Deze verschillende clusters geven aan hoeveel verschillende netwerktypen worden gevonden in de data, uit hoeveel persoonlijke netwerken deze clusters bestaan en een indeling van individuen in clusters. Op basis van deze clusters kunnen verschillende typen netwerken worden geïdentificeerd; gezamenlijk vormen deze netwerktypen de netwerktypologie. Als er sprake is van een goede clustering zullen mensen met vergelijkbare persoonlijke netwerken in hetzelfde cluster vallen. Deze indeling in clusters vereenvoudigd de netwerkdata, waardoor patronen in de persoonlijke netwerken zichtbaar worden (Gianella & Fisher, 2016; Jain, 2010; Vacca, 2020).

1.2 Probleemstelling

Het doel van deze scriptie is tweeledig. Allereerst is het doel om de grote hoeveelheid data over het netwerk van Nederlandse vrouwen systematisch te ordenen. Daarnaast is het doel ook om een bijdrage te leveren aan de kennis over de invloed van sociale contacten op fertiliteitswensen.

Om deze doelen te behalen zal een netwerktypologie worden gemaakt op basis van structurele kenmerken van het netwerk. Waarna onderzocht wordt wat de samenhang is tussen verschillende netwerktypen en fertiliteitswensen. Daarna zal gekeken worden of de toevoeging van compositie

kenmerken aan de clustermethode een betere groepering oplevert en wat dit mogelijk verandert aan de samenhang tussen verschillende netwerktypen en fertiliteitswensen.

De vraag die centraal staat in de scriptie luidt als volgt:

‘In hoeverre kunnen de persoonlijke netwerken van vrouwen worden opgedeeld in zinvolle netwerktypen? En hoe hangen deze netwerktypen samen met de fertiliteitswensen van een vrouw?’

Deze vraag wordt onderzocht door antwoord te geven op de volgende deelvragen. De eerste, derde en vierde deelvraag zijn beschrijvende deelvragen. De andere twee deelvragen zijn verklarend.

1. Welke netwerktypen kunnen worden geïdentificeerd bij Nederlandse vrouwen tussen de 18 en 40 jaar, gebruikmakend van structurele kenmerken van het netwerk?
2. In hoeverre hangen de verschillende netwerktypen, op basis van structurele netwerkkkenmerken, samen met de fertiliteitswensen van een vrouw?
3. Welke netwerktypen kunnen worden geïdentificeerd bij Nederlandse vrouwen tussen de 18 en 40 jaar, gebruikmakend van zowel structurele kenmerken als de compositie kenmerken van het netwerk?
4. In welke mate is het relevant voor de identificatie van netwerktypen om de compositie kenmerken toe te voegen aan de clustermethode op basis van structurele kenmerken?
5. In hoeverre hangen de verschillende netwerktypen, op basis van structurele kenmerken en compositie kenmerken van het netwerk, samen met de fertiliteitswensen van een vrouw?

1.3 De huidige studie

In deze paragraaf zal ik uiteenzetten welke kennis dit onderzoek zal toevoegen aan de bestaande kennis, waarom dit van belang is en hoe dit beoogd wordt te behalen. Dit is opgedeeld in drie verschillende onderdelen, namelijk de gebruikte methode, de wetenschappelijke relevantie en de sociologische relevantie.

1.3.1 De methode

In dit onderzoek wordt gebruik gemaakt van de op subgroep gebaseerde clustermethode uit de studie van Vacca (2020). In een uitgebreide studie naar persoonlijke netwerken was zijn doel om de meest accurate maten te vinden om persoonlijke netwerken te clusteren. Hiervoor maakt Vacca (2020) gebruik van samenhangende subgroepen om de clusters te identificeren. Een samenhangende subgroep bestaat uit minstens drie personen die beter met elkaar zijn verbonden dan met de rest van het netwerk. Voor deze clusteranalyse gebruikt Vacca (2020) drie maten, namelijk het aantal subgroepen met ten minste drie alters, het aantal subgroepen met één of twee alters (hierna: isolaten

en dyads) en de modulariteitsscore van het netwerk. Hierbij is de modulariteitsscore een score die aangeeft in hoeverre de indeling in subgroepen geschikt is voor het netwerk (Newman & Girvan, 2004). Op basis van deze drie maten wordt een k-medoid analyse uitgevoerd, waardoor de meest representatieve datapunten voor elk cluster bekend worden. Om te valideren of deze methode effectief is, maakt hij gebruik van zes verschillende datasets van persoonlijke netwerken. De resultaten laten zien dat het gebruik van subgroepen effectief is in het detecteren van de variantie in persoonlijke netwerkstructuren in diverse datasets (Vacca, 2020). Hij weet met slechts drie maten de structuur van het persoonlijke netwerk te vatten. Deze drie maten staan dan ook vast voor elk onderzoek dat gebruik maakt van deze methode. Dit maakt dat de onderzoeker niet zelf statistische maten kiest, waardoor de invloed van de onderzoeker wordt verkleind en de vergelijking van de resultaten wordt vereenvoudigd.

In deze scriptie zal eerst alleen gebruik maken van structurele kenmerken om de netwerktypologie te maken. Daarna zal de compositie van het netwerk worden meegenomen in het creëren van de typologie. Een combinatie van deze twee kenmerken biedt kansen voor het vangen van meer informatie (Chunaev, 2020). Daarbij zijn zowel de structurele kenmerken als de compositie kenmerken relevant voor de persoonlijke uitkomsten van een individu (Vacca, 2020).

1.3.2 Wetenschappelijke relevantie

De methode van Vacca (2020) is nog niet in veel onderzoeken gebruikt. Tot op heden zijn er, naast de studie van Vacca (2020), drie andere onderzoeken verschenen die gebruik maken van deze methode. De steekproefgrootte en het aantal alter-alter relaties in deze studies zijn aanzienlijk kleiner dan de data die gebruikt wordt in dit onderzoek. In de studie van Vacca (2020) worden 6 datasets vergeleken, de steekproefgroottes hier lopen van 119 participanten tot 366 participanten. Daarbij is de variatie in aantal alter-alter relaties zeer groot, namelijk 5 tot 45 alters. De andere studies hadden een steekproefgrootte en aantal alter-alter relaties van 119 personen met 30 alter-alter relaties (Vacca et al., 2022), 30 personen met 20 alters (Fanfan et al., 2021) en 69 personen met 45 alters (Maya-Jariego & Holgado-Ramos, 2022). Daarentegen is dit onderzoek gebaseerd op data uit het LISS-panel (Langlopende Internet Studies voor de Sociale wetenschappen panel), beheerd door CentERdata van Tilburg University. De data uit het LISS-panel is door gerandomiseerde steekproef tot stand gekomen en is representatief voor de Nederlandse bevolking. Binnen dit panel is een studie gedaan onder de vrouwen tussen de 18 en 40 jaar genaamd ‘Sociale Relaties en Kinderkeuzes’. De steekproef bestaat uit 1332 vrouwen, waarvan 758 de studie met succes afronden. De vrouwen die de studie niet begonnen of afronden waren over het algemeen goed vergelijkbaar met de vrouwen die dit wel deden (Stulp, 2020). Deze steekproef levert een grote hoeveelheid netwerkdata op over een geavanceerde samenleving. Het is interessant om te onderzoeken in hoeverre de methode van Vacca

(2020) in staat is de grote hoeveelheid aan netwerkdata te clusteren en wat de overeenkomsten zijn in resultaten met zijn onderzoek.

1.3.3 Sociologische relevantie

Inmiddels tien jaar geleden publiceerden Balbo, Billari en Mills (2013) een overzichtsartikel over het onderzoek naar vruchtbaarheid in geavanceerde samenlevingen. Zij concludeerden dat er nog grootschalig kwantitatief onderzoek ontbreekt wat zich richt op het effect van sociale interactie op de fertiliteitskeuzes en -wensen van mensen. Nadien is gebleken dat sociale relaties, zoals broers en zussen (Buyukkececi et al., 2020), vrienden (Loïs & Arránz Becker, 2014) en collega's (Pink et al., 2014), invloed hebben op de fertiliteitskeuzes die iemand maakt. Er is alleen weinig bekend over de invloed van de structuur van het persoonlijke netwerk van de vrouw op haar fertiliteitskeuzes en -wensen. Daarbij is wel bekend welke sociale relaties invloedrijk zijn, maar niet hoe de totale samenstelling van relaties invloed heeft. Dit onderzoek wil een bijdrage leveren aan het in kaart brengen van de persoonlijke netwerken van vrouwen, op zowel structurele als compositie kenmerken. Daarnaast is dit onderzoek erop gericht de samenhang te onderzoeken tussen het hebben van een bepaald netwerktype en de kinderwens van een vrouw.

2. Theoretisch kader

In dit hoofdstuk zullen twee onderwerpen worden uitgediept, namelijk allereerst hoe persoonlijke netwerkdata systematisch kan worden geordend en ten tweede de invloed van sociale contacten op fertiliteitswensen en -keuzes.

2.1 Ordenen van persoonlijke netwerkdata

De data over het persoonlijke netwerk van een persoon bestaat uit de relaties die iemand heeft met mensen om zich heen. Daarbij kan de data ook informatie bevatten over de onderlinge contacten tussen de relaties van een persoon, de alter-alter relaties. Dit levert een grote hoeveelheid aan data op die op een juiste manier geanalyseerd dienen te worden (Gianella & Fisher, 2016). Een methode om deze grote hoeveelheden aan data te verwerken is de clusteranalyse.

2.1.1 Meerwaarde van een netwerktypologie

Persoonlijke netwerken zijn op basis van kenmerken van het netwerk in te delen in verschillende netwerktypen. Deze indeling in netwerktypen is om verschillende redenen relevant. Allereerst maakt een netwerktypologie het mogelijk om gegevens van grote aantallen persoonlijke netwerken samen te vatten (Gianella & Fisher, 2016; Jain, 2010; Vacca, 2020). Door gegevens samen te vatten wordt zichtbaar welke soort typen vaker voorkomen dan anderen en welke typen uitzonderlijk zijn. Dit maakt het mogelijk om verschillende netwerktypen met elkaar te vergelijken (Maya-Jariego, 2021; Vacca, 2020). Ten tweede wordt de onderliggende structuur zichtbaar waardoor patronen in de data kunnen worden herkend (Gianella & Fisher, 2016; Jain, 2010; Vacca, 2020). Deze patronen bieden inzicht in de sociale context van persoonlijke gegevens en laten zien waar persoonlijke netwerken afwijken. Ten derde is het relevant voor een groter begrip van persoonlijke netwerken in het algemeen en de vergelijking tussen verschillende populaties (Vacca, 2020). Ook kunnen verschillen tussen clusters worden getoetst, bijvoorbeeld het verschil in leeftijd waarop een vrouw haar eerste kind krijgt (Bidart et al., 2018; Robins 2015; Vacca, 2020). Als laatste legt het proces waarmee de netwerktypologie wordt gecreëerd ook bloot welke netwerkkenmerken een groot of klein aandeel hebben in de totstandkoming (Maya-Jariego, 2021; Vacca, 2020). Hiermee kan de invloed van theoretisch relevante variabelen worden onderzocht.

Een netwerktypologie wordt meestal gemaakt aan de hand van een clustermethode. Deze methode wordt gebruikt om netwerkdata te groeperen; het indelen van netwerken in verschillende clusters. Het doel van deze methode is om netwerken binnen hetzelfde cluster te groeperen op basis van gelijkheid, terwijl netwerken in verschillende clusters van elkaar dienen te verschillen. Er zijn in de

sociologie een aantal onderzoekers die deze methode gebruiken om persoonlijke netwerken in te delen en persoonlijke uitkomsten te verklaren. Een paar relevante onderzoeken worden gepresenteerd.

Het onderzoek door Lubbers en collega's (2007) clustert persoonlijke netwerken van migranten om de relatie tussen het netwerk en de eigen etnische identiteit te onderzoeken. Ze gebruiken zowel structurele als compositie kenmerken, zoals de netwerkdichtheid, het aantal subgroepen en het percentage familieleden. De *k-means* clusteranalyse wordt gebruikt om vijf verschillende groepen te identificeren, zoals het 'verbonden familie netwerk'. Op die manier kunnen ze laten zien dat migranten met een sterk verbonden netwerk zonder veel lokale contacten zich vaak identificeren met het land van herkomst.

Bidart en collega's (2018) maakten gebruik van enkel structurele kenmerken om een netwerktypologie te creëren voor persoonlijke netwerken van Franse jongeren. Hiervoor maakte ze gebruik van visuele herkenning voor een eerste indeling, om vervolgens vier structurele indicatoren (dichtheid, centraliteit, modulariteit en de diameter) te raadplegen voor de definitieve verdeling. Op basis van de *discriminantanalyse* definieerde zij zes clusters. Zij zijn erin geslaagd om correlaties vast te leggen tussen netwerktypen en individuele attributen (Bidart et al., 2018). Zo hadden jongeren met een gesegmenteerd netwerk vaak een hogere opleiding.

In het onderzoek van Maya-Jariego (2021) was het doel om een netwerktypologie te maken van de algemene bevolking van de Spaanse gemeente Alcala de Guadaira. Dit heeft hij gedaan door middel van enkel structurele kenmerken, namelijk *centralization*, *cliques* en *components*. Dit resulteerde in vier categorieën persoonlijke netwerken, namelijk een dicht netwerk, tussenliggend netwerk, geclusterd netwerk en gefragmenteerd netwerk. Deze typologie is vergelijkbaar met het eerdere onderzoek van Bidart en collega's (2018).

Op het werk van Bidart en collega's (2018) voortbordurend publiceerde Vacca (2020) het artikel waarin hij de *op subgroep gebaseerde clustermethode* introduceerde. Hierin maakt Vacca (2020) gebruik van drie variabelen die gezamenlijk de samenhang tussen de alters in het netwerk samenvatten en toetst deze methode op zes datasets. Deze datasets verschillen aanzienlijk in zowel het aantal alter-alter relaties die bekend zijn als in de context van de studie. Desondanks is de methode in staat om op basis van drie variabelen de verschillende data te clusteren.

In het onderzoek van Fanfan en collega's (2023) wordt gebruik gemaakt van de op subgroep gebaseerde clustermethode van Vacca (2020). Zij onderzoeken de verhalen en persoonlijke netwerken van de snelgroeijende Latino/a-bevolking op het platteland. Hiervoor gebruikte zij kwalitatieve interviews en persoonlijke netwerkanalyse. Visualisaties van geclusterde persoonlijke netwerken worden gebruikt om de relatie tussen ego en alter te kwalificeren, zoals familie, school of andere contexten. Ze trekken verschillende conclusies, waaronder dat de sociale context van de Latino/a's

op het platteland voornamelijk plaatsvindt op school en dat de meeste adolescenten een bi-factional netwerktype hebben waarbij de scheiding vaak tussen school en familieleden ligt.

2.1.2 Soorten structurele netwerktypologieën

In onderzoeken naar persoonlijke netwerken zijn er een aantal terugkerende structurele typologieën. Vaak worden de typologieën gebaseerd op verschillende structurele kenmerken (Bidart et al., 2018; Lubbers et al., 2007; Maya-Jariego, 2021). De structuur van persoonlijke netwerken kan, net zoals in gehele netwerk studies, statistisch gemeten worden aan belangrijke maatstaven zoals: *density*, *degree centrality*, *cliques* en *components* (McCarty, 2002). In de onderzoeken naar persoonlijke netwerken wordt vaak de dichtheid van het netwerk gebruikt om de typologie aan te duiden. Ter illustratie het onderzoek van Bidart en collega's (2018). Hierin zijn de netwerkgroepen: *regular dense*, *centered dense*, *centered star*, *segmented*, *pearl collar* en *dispersed*. In de *regular dense*-groep is elke alter relatief goed verbonden met elke andere alter (Bidart et al., 2018). Daartegenover is in de *dispersed*-groep sprake van weinige verbindingen tussen alters.

De twee maatstaven *cliques* en *components* berusten beide op het concept van subgroepen. Het onderscheid tussen *cliques* en *components* ligt in het feit dat *cliques* vereisen dat alle alters in een subgroep volledig verbonden zijn, terwijl bij *components* sprake is van een subgroep als alters direct of indirect verbonden zijn (McCarty, 2002). Daarnaast bestaat ook het concept gemeenschappen, dit zijn subgroepen binnen een netwerk met een hogere netwerkdichtheid dan de rest van het netwerk (Girvan & Newman, 2002). Het concept van gemeenschappen is theoretisch interessant om meerdere redenen. Ten eerste biedt het inzicht in verschillende sociale groepen binnen het persoonlijke netwerk en tevens hoe deze groepen zijn verdeeld (Girvan & Newman, 2002). De sociologische context kan worden toegevoegd door het kwalificeren van relaties, zoals een gemeenschap van familieleden of schoolvrienden (Vacca, 2020). Ten tweede is het op deze manier mogelijk om analyses te doen op basis van de gemeenschapsstructuur (Girvan & Newman, 2002). Door individuen toe te wijzen aan een specifiek cluster op basis van de structuur van hun netwerk, wordt een indeling gemaakt waardoor de analyse vereenvoudigd wordt.

De methode van Vacca (2020) weet een combinatie te maken tussen een theoretisch relevante analyse van subgroepen, een methode die in meerdere contexten een duidelijk onderscheid weet te maken in netwerktypen en de invloed van de keuzes van een onderzoeker te verkleinen. Om die reden is ervoor gekozen deze op subgroep gebaseerde methode van Vacca (2020) te gebruiken om de persoonlijke netwerken van Nederlandse vrouwen te clusteren. Er zijn een aantal recente onderzoeken waarin de netwerktypologie enkel gemaakt wordt op basis van structurele kenmerken. Hieruit komt naar voren

dat structurele kenmerken grotendeels samenhangen met de sociale context en de compositie van het netwerk (Bidart et al., 2018; Gianella & Fisher, 2016; Maya-Jariego, 2021). Oftewel, een groot gedeelte van de informatie om een netwerktypologie te maken kan worden gehaald uit de structurele kenmerken van het netwerk.

Om de vergelijking met compositie kenmerken te kunnen maken zal in deze scriptie na de analyse op basis van structurele kenmerken ook de compositie kenmerken worden toegevoegd aan het model. Dit is vergelijkbaar met het artikel van Fanfan en collega's (2021). Echter, in deze scriptie zal tweemaal een netwerktypologie worden gemaakt waarna een vergelijking zal volgen over welke typologie het best past bij de persoonlijke netwerkdata van de Nederlandse vrouwen. Oftewel, welk model het beste in staat is om vergelijkbare persoonlijke netwerken in hetzelfde cluster te plaatsen.

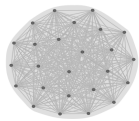
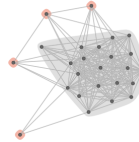
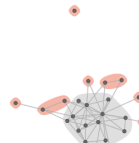
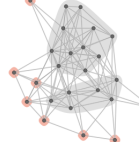
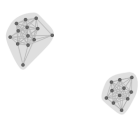
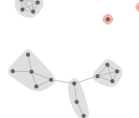
In het onderzoek van Vacca (2020) is het concept van samenhangende gemeenschappen gebruikt om een typologie te maken. Hij kon op basis van statistische maten zes verschillende datasets effectief indelen. Binnen deze zes verschillende datasets zijn achttien verschillende typen netwerken te onderscheiden. Er komen gemiddeld ongeveer zeven typen netwerken voor binnen één dataset. Een voorbeeld van de netwerktypen die binnen één van de zes datasets voorkomen zijn: *closed community*, *dense core-periphery*, *sparse core-periphery*, *weakly bi-factional*, *strong bi-factional* en *multi-factional*. Een *closed community* is in dit geval een netwerk waarin maar één gemeenschap is. Dit is vergelijkbaar met de *regular-dense*-groep uit het onderzoek van Bidart en collega's (2018). Daarentegen bestaat het *weakly bi-factional* netwerktype uit twee subgroepen die onderling sterk verbonden zijn, waarbij ook veel losse sociale contacten worden gevonden. Bij een *dense multi-factional netwerktype* is er sprake van meerdere goed gescheiden subgroepen.

In de typologie van Vacca (2020) worden drie soorten onderscheid gemaakt in de benaming van de netwerktypen, namelijk *dense-sparse*, *weakly-strongly* en het aantal subgroepen. Het onderscheid *dense* of *sparse* geeft aan of er sprake is van duidelijke centrale en dichte kernen (*dense*) of dat er verschillende subgroepen zonder duidelijke kern zijn (*sparse*). Het tweede onderscheid tussen *weakly-strongly* verwijst naar de dichtheid en de scheiding van groepen. In sommige netwerken is er sprake van een hoge dichtheid en zijn er veel banden tussen verschillende groepen (*weakly*) en in de andere netwerken is sprake van een duidelijke scheiding tussen netwerken en dit gaat vaak gepaard met een hoge modulariteit (*strongly*). Daarnaast heeft elk netwerk ook een modulariteitsscore. Een hoge score geeft aan dat er sprake is van een hoge mate van verbondenheid binnen de subgroepen, maar dat de verbondenheid tussen de subgroepen laag is. Als laatste wordt met de isolaten en dyads aangegeven in hoeverre sprake is van een verbonden netwerk of dat er veel losse mensen aanwezig zijn.

In het onderzoek van Vacca (2020) was er een redelijke overlap in typen netwerken tussen de zes datasets. De meest voorkomende netwerktypen waren: *closed community*, *dense core-periphery*,

sparse core-periphery, *weakly bi-factional*, *strongly bi-factional* en *sparse multi-factional*. Hieronder is Tabel 1 toegevoegd die meer duidelijkheid geeft over elk van deze netwerktypen. Naast deze netwerktypen worden ook andere netwerktypen gevonden, de indeling die Vacca (2020) hanteert is gepresenteerd in Tabel 2. Het doel van Vacca (2020) was om op basis van enkel kenmerken gericht op subgroepen de typologie te creëren en dit is gelukt. Deze methode zal in deze scriptie worden gebruikt.

Tabel 1: Omschrijving meestvoorkomende netwerktypen Vacca (2020).

| Netwerktype | Omschrijving | Schematisch plaatje* |
|------------------------|---|---|
| Closed community | In dit netwerktype is sprake van 1 subgroep waarbij alle alters goed met elkaar verbonden zijn. |  |
| Dense core-periphery | In dit netwerktype is sprake van 1 goed verbonden subgroep met daarnaast enkele niet goed verbonden alters. |  |
| Sparse core-periphery | In dit netwerktype is er sprake van 1 goed verbonden subgroep met daarnaast meerdere niet goed verbonden alters. |  |
| Weakly bi-factional | In dit netwerktype is er sprake van 2 subgroepen die intern goed verbonden zijn. Daarbij bestaan er ook relaties tussen deze twee gemeenschappen. |  |
| Strongly bi-factional | In dit netwerktype is er sprake van 2 subgroepen die intern goed verbonden zijn. Tussen de subgroepen bestaan er (nauwelijks) geen relaties. |  |
| Sparse multi-factional | In dit netwerktype is er sprake van meerdere subgroepen met drie of meer alters en daarnaast zijn er ook meerdere alters alleen of in een dyad. |  |

*Vacca (2020), Figuur 1: "Representative personal networks from each of the seven types of the subgroup-based typology (T1) in the ECU dataset." Vacca (2020) gebruikt het figuur om de netwerken uit één van de datasets te presenteren. Hier gebruikt ter voorbeeld van een mogelijk netwerk binnen dit netwerktype.

Hypothese 1: De netwerktypen van vrouwen zijn vergelijkbaar met de netwerkstructuren uit het onderzoek van Vacca (2020).

Om op basis van de resultaten de juiste benaming van Vacca (2020) toe te passen op de structurele kenmerken is een samenvatting gemaakt van zijn interpretatie (Tabel 2). Het is goed om te vermelden dat er geen grenswaarden worden gesteld. Daarbij zijn er verschillende kenmerken die leiden tot dezelfde benaming. Oftewel, in sommige gevallen is de interpretatie van de onderzoeker nodig als er meerdere netwerktypen zijn die onder dezelfde groep vallen. Ter illustratie; in een netwerktypologie komen twee netwerktypen met twee subgroepen waarvan één gemiddeld 3 isolaten of dyads heeft en de ander gemiddeld 5, dan zal de eerste worden gedefinieerd als strongly bi-factional en de ander als weakly bi-factional. Als laatste wordt de modulariteitscore ook meegenomen, maar deze is wel relatief. Bij een groter aantal subgroepen zal de modulariteitscore in de meeste gevallen hoger zijn. Het is dan ten opzichte van de andere netwerktypen met hetzelfde aantal subgroepen.

In het onderzoek van Vacca (2020) komen ook de groepen ‘large core-periphery’ en ‘small core-periphery’ voor. Deze twee groepen zijn in Tabel 2 weggelaten. Beide lijken te vallen onder 1 subgroep met ongeveer 5 isolaten, waar het onderscheid wordt gemaakt op de modulariteitscore. Echter, Vacca (2020) geeft hier geen toelichting over.

Tabel 2: Samenvatting drie structurele kenmerken en benaming netwerktypen

| | 0 isolaten of dyads | ± 5 isolaten of dyads | ± 10 isolaten of dyads | ±15 isolaten of dyads | ±20 isolaten of dyads |
|----------------------|-----------------------------------|-----------------------------------|-------------------------------------|------------------------------------|------------------------------------|
| 0 subgroepen | – | – | – | – | disconnected M = hoog |
| 1 subgroep | closed community M* = 0 | dense core-periphery M = laag | sparse core-periphery M = middel | sparse core-periphery M = hoog | dual core-periphery M = hoogst |
| 2 subgroepen | strongly bi-factional M = laag | strongly bi-factional M = laag | weakly bi-factional M = hoog | weakly bi-factional M = hoog | weakly bi-factional M = hoogst |
| 3 subgroepen | tri-factional M = laag | dense multi-factional M = laag | sparse multi-factional M = hoog | sparse multi-factional M = hoog | sparse multi-factional M = hoog |
| 4 of meer subgroepen | dense multi-factional M = laag | dense multi-factional M = laag | sparse multi-factional M = hoog | sparse multi-factional M = hoog | sparse multi-factional M = hoog |

* M = Relatieve modulariteitswaarde

2.2 Invloed van sociale contacten op fertiliteitswensen

De sociale contacten van een individu vormen samen het persoonlijke netwerk. Deze sociale contacten hebben invloed op de individuele uitkomsten van een persoon (Granovetter, 1985). In het onderzoek naar deze invloed binnen de fertiliteitskeuzes en -wensen is een driedeling te maken, namelijk welke mechanismen van invloed zijn aanwezig, welke personen hebben invloed en wat is de invloed van de structuur van het persoonlijke netwerk. De eerste twee onderdelen zijn uitvoerig onderzocht en de resultaten hiervan zullen worden gepresenteerd. Daarentegen is de invloed van de structuur nog onderbelicht, hier zullen hypothesen over worden opgesteld.

2.2.1 Mechanismen van invloed door sociale contacten

De vier meest onderzochte mechanismen van invloed door sociale contacten op fertiliteitskeuzes en -wensen zijn: sociaal leren, sociale druk, sociale besmetting en sociale steun (Bernardi & Klärner, 2014).

Er is sprake van *sociaal leren* als vrouwen leren van ervaringen van relaties en deze informatie meenemen in de besluitvorming. Een vrouw leert van de ervaring van bijvoorbeeld een vriendin of collega die kinderen krijgt en hierdoor zal zij zelf ook eerder kinderen krijgen (Balbo & Barban, 2014; Pink et al., 2014). Vrouwen zullen namelijk eerder overgaan op kinderen wanneer ze de positieve invloed van kinderen zien op het leven van sociale relaties en dat deze verandering opweegt tegen de kosten (Loïs, 2013).

Het tweede mechanisme, *sociale besmetting*, geeft aan dat een vrouw het gedrag of een idee overgedragen krijgt van een sociaal contact (Bernardi & Klärner, 2014). Hierbij is het niet nodig dat de vrouw bewust gaat handelen vanuit deze opgedane kennis. Oftewel, het is een meer onbewuste manier van gedragsverandering, waarbij enkel een emotie of een gedraging een positieve indruk kan achterlaten bij een vrouw. Als vrouwen omgaan met mensen die zelf ook kinderen hebben, dan zullen zij zelf ook eerder overgaan tot het moederschap (Loïs, 2013; Loïs & Arránz Becker, 2014).

Sociale druk is een mechanisme waarbij vrouwen conformeren aan de sociale norm en vanuit de overtuiging het juiste te doen overgaan tot bepaalde fertiliteitskeuzes. Een vrouw die afkomstig is uit een groot gezin of een netwerk heeft met veel familieleden gaat sneller over tot het moederschap en krijgt gemiddeld meer kinderen (Kotte & Ludwig, 2011; Mathews & Sear, 2013). Ook kunnen vrouwen bijvoorbeeld minder kinderen krijgen of het krijgen van kinderen uitstellen omdat ze weten dat sociale relaties hier een mening over hebben (Dommermuth et al., 2011). Andersom kunnen vrouwen ook eerder kinderen krijgen op het moment dat vrienden tot deze beslissing over gaan en het wordt gezien als de norm (Loïs, 2013; Loïs & Arránz Becker, 2014). Daarbij hoeft het niet om actieve druk te gaan; mensen kunnen eerder overgaan op het ouderschap als ze weten dat familie en vrienden hier positief tegenover staan (Dommermuth et al., 2011).

Als laatste is *sociale steun* het mechanisme waarmee wordt aangegeven dat vrouwen sociale contacten hebben waarmee ze goederen of diensten verruilen (Bernardi & Klärner, 2014). Voorbeelden hiervan zijn ouders die kunnen bijdragen aan het klaarmaken van de kinderkamer en wederzijdse opasuren met vrienden. Een vrouw is zich bewust van haar sociale contacten en neemt bewust de middelen die haar ter beschikking staan door deze contacten in overweging bij het nemen van fertiliteitskeuzes (Bühler & Philipov, 2005; Dommermuth et al., 2011, Loïs & Arránz Becker, 2014). Hierbij zijn vooral familierelaties belangrijk; vrouwen waarvan het overgrote gedeelte uit naaste familieleden bestaat ervaart namelijk veel sociale steun ten aanzien van het krijgen van kinderen (Loïs, 2016). Samenvattend, er zijn veel verschillende mechanismen (tegelijkertijd) actief die de fertiliteitskeuzes en -wensen van een vrouw beïnvloeden.

2.2.2 De netwerk compositie en sociaal kapitaal

De compositie van het persoonlijke netwerk richt zich op wat voor relaties iemand heeft en welke kenmerken de sociale contacten hebben. Als we ervan uitgaan dat de compositie van het persoonlijke netwerk belangrijk is voor de persoonlijke uitkomsten, dan veronderstel je dat bepaalde sociale contacten meer invloed hebben dan anderen. De meeste onderzoeken die zich bezighouden met welke sociale contacten invloedrijk zijn op de fertiliteitskeuzes en -wensen richten zich op één bepaalde groep, bijvoorbeeld de ouders van een vrouw.

Zo blijkt dat de ouders van een vrouw invloed hebben op haar fertiliteitskeuzes doordat ze mogelijk substantieel veel sociale steun bieden (Bühler & Philipov, 2005; Kavas & Jong, 2020; Keim et al., 2013). Daarnaast hebben ook broers en zussen invloed, want resultaten wijzen uit dat mensen vaker kinderen krijgen binnen een korte tijdsspanne nadat een broer of zus kinderen heeft gekregen (Balbo & Mills, 2011; Buyukkececi et al., 2020; Kotte & Ludwig, 2011, Lyngstad & Prskawetz, 2010). Daarbij zijn niet alleen familieleden invloedrijk, maar ook andere sociale contacten zoals vrienden (Balbo & Barban, 2014; Kotte & Ludwig, 2011) en collega's (Buyukkececi et al., 2020; Ciliberto et al., 2016; Pink et al., 2014) beïnvloeden fertiliteitskeuzes en -wensen.

Loïs (2016) wist op basis van een clusteranalyse vier soorten relatienetwerken te onderscheiden en de samenhang met overgang naar het ouderschap te onderzoeken. De resultaten uit dit onderzoek wijzen erop dat netwerken waarin veel naaste familieleden aanwezig zijn samenhangt met een snellere overgang naar het ouderschap. Zijn verklaring hiervoor is dat de sociale invloeden van de naaste familieleden sterker zijn en de vrouw kan rekenen op hogere sociale steun, meer sociale druk ervaart en eerder sociaal besmet raakt.

Andere compositie kenmerken zijn ook relevant voor de fertiliteitskeuzes en -wensen van een vrouw. Zo kiezen vrouwen die veel sociale contacten hebben met kinderen zelf ook eerder voor het ouderschap (Loïs, 2013). Bijvoorbeeld een vriendengroep waar al veel vrienden met kinderen zijn.

Daarbij maken vrouwen met meer familieleden in haar netwerk ook eerder de stap naar het ouderschap (Mathews & Sear, 2013; Balbo & Mills, 2011).

Een belangrijke theorie in de sociologie is dat het verschil in sociaal kapitaal in een persoonlijk netwerk invloed heeft op de uitkomsten voor het individu. Het sociaal kapitaal verwijst naar het kapitaal dat verworven wordt door het netwerk (Bourdieu, 1986; Coleman 1988; Lin, 2001). Dit zijn middelen waarvan een individu gebruikmaakt of weet dat er in de toekomst gebruik van gemaakt kan worden (Loïs & Arránz Becker, 2014). Het verschil in sociaal kapitaal zorgt voor een verschil in toegang tot informatie, invloed van sociale contacten, vertrouwen van derden door persoonlijke relaties, versterkte erkenning door derden en sociale controle (Lin, 2001). Oftewel, het sociaal kapitaal van een persoon beïnvloedt de wensen en keuzes die een individu heeft.

De compositie van een persoonlijk netwerk heeft invloed op het niveau van sociaal kapitaal dat een individu bezit, aangezien diverse mensen verschillende mogelijkheden met zich meebrengen. Om dit te illustreren: twee voorbeelden. Het eerste voorbeeld gaat in op de kenmerken van alters. Als alters veel kenmerken bezitten waarvan vrouwen mogelijk gebruik kunnen maken, dan zou dit mogelijk een positieve impact hebben op de fertiliteitswensen die zij heeft. Als zij veel vrienden heeft die al kinderen hebben, dan is de kans groter dat zij zelf ook sneller kinderen krijgt (Loïs, 2013). Het tweede voorbeeld gaat over de relaties die zij heeft. Bestaat haar netwerk uit veel naaste familie, dan is het mogelijk dat zij veel sociale steun ontvangt en leert van de ervaringen van haar dierbaren, waardoor ze andere fertiliteitskeuzes maakt. Daartegenover kan het ook zijn dat als zij relaties heeft met veel verschillende mensen, dat er dan verschillende informatie bij haar komt en dit andere kansen biedt waardoor haar wensen en keuzes veranderen (Burt, 1992).

2.2.3 De netwerkstructuur en sociaal kapitaal

Er is minder bekend over de invloeden van de structuur van persoonlijke netwerken op fertiliteitswensen en -keuzes. In het fertiliteits- en verwantschapsonderzoek dat tot nu toe is uitgevoerd wordt de netwerkdichtheid als leidende maat genomen. Dit is een maat die aangeeft hoeveel sociale verbindingen er tussen alters aanwezig zijn in het persoonlijke netwerk ten opzichte van de in totaal mogelijke verbindingen. Zo is onderzocht dat een hogere netwerkdichtheid een lager gebruik van anticonceptiemiddelen bij huishoudens laat zien dan een lagere netwerkdichtheid (Kohler et al., 2001). Een ander onderzoek laat zien dat een dicht verwantschapsnetwerk bijdraagt aan het kunnen praten met verwanten over het krijgen van kinderen (Stulp & Barrett, 2021). Daarentegen vonden zij ook dat voor een vriendschapsnetwerk het tegenovergestelde geldt, namelijk een dicht netwerk zorgt er in dat geval voor dat mensen minder goed met vrienden kunnen praten over het krijgen van kinderen (Stulp & Barrett, 2021). Er heerst een algeheel idee dat de structuur van het

netwerk invloed heeft op fertiliteitskeuzes, maar tot nu toe is enkel netwerkdichtheid hierin meegenomen.

Een bepaalde persoonlijke netwerkstructuur biedt een bepaalde mate van sociaal kapitaal omdat het beïnvloedt op welke manier het individu zijn of haar netwerk inzet (Vacca, 2020). Dit geeft een verschil in voordelen en beperkingen voor het individu. Zeer verbonden persoonlijke netwerken zijn netwerken waarin de alters veel onderlinge verbanden hebben. Het gaat dan om een gesloten netwerk. Dit wordt ook wel *bonding social capital* genoemd (Coleman, 1988). Deze vorm van sociaal kapitaal gaat gepaard met een hoge mate van vertrouwen, wederzijdse verplichtingen, hoge mate van steun en samenwerking en gedeelde sociale normen (Coleman, 1988). Daartegenover staan netwerken waarin de sociale contacten onderling niet goed verbonden zijn, *bridging social capital* (Coleman, 1988). Kenmerkend voor dit soort netwerken is dat er meerdere groepen en eenlingen bestaan die onderling niet met elkaar zijn verbonden. Dit heeft andere voordelen en beperkingen voor het individu dan bij een verbonden netwerk. De ideeën van *weak ties* en *structural holes* hangen samen met deze vorm van sociaal kapitaal. In een persoonlijke netwerk waarbij de *alters* niet goed met elkaar zijn verbonden heeft het individu toegang tot verschillende groepen mensen die verschillende soorten informatie en kansen bieden (Burt, 1992). Daarnaast is het waarschijnlijk dat het persoonlijke netwerk van *alters* meer verschillende individuen bevat dan bij een sterk verbonden netwerk. Oftewel, *ego* heeft meer toegang tot *weak ties* en dit biedt tevens meer informatie en kansen (Granovetter, 1973).

Een verbonden persoonlijk netwerk heeft als kenmerk dat de onderlinge *alter-alter* relaties sterk zijn en (bijna) iedereen met elkaar is verbonden. In deze studie gaat het om een contact-netwerk. Oftewel, de *ego* is gevraagd naar welke *alters* onderling contact met elkaar hebben. De voordelen die een verbonden netwerk biedt is een hoge mate van sociale steun vanuit sociale contacten (Wellman & Wortley, 1990). Hier is dan sprake van het mechanisme *sociale steun*. Daarnaast is in een verbonden netwerk ook een meer beperkte toegang tot informatie, hebben contacten meer invloed en heerst er een hogere mate van sociale controle (Coleman, 1988). Hierdoor zijn de mechanismen van *sociale besmetting* en *sociale controle* mogelijk effectiever, waardoor de vrouw ook eerder overgaat tot het moederschap. Hierop gebaseerd verwacht ik dat een vrouw met een verbonden netwerk een hogere kinderwens heeft.

Een ander veelvoorkomend persoonlijk netwerk is een *bi-factional* netwerk (Bidart et al., 2018; Vacca, 2020), een persoonlijk netwerk met twee goed verbonden gemeenschappen die onderling slecht verbonden zijn. Doordat er maar sprake is van twee verschillende gemeenschappen, wordt het voordeel van een goed verbonden netwerk waarschijnlijk minder. Oftewel, er zal waarschijnlijk nog altijd sprake zijn van een hoge mate van sociale steun, beperkte toegang tot

informatie en sociale controle. Echter, de verwachting is dat dit in mindere mate voorkomt waardoor een vrouw een lagere kinderwens heeft.

Een minder verbonden netwerk zorgt ervoor dat de vrouw toegang heeft tot een divers aanbod aan informatie, gebaseerd op het idee van *structural holes* en *weak ties*. Hierdoor is het mogelijk dat op basis van het mechanisme *sociaal leren* de vrouw in aanraking komt met meer vrouwen die positief zijn over het moederschap en zelf ook eerder hiertoe overgaan. Daarnaast zou ze ook het tegenovergestelde kunnen leren, namelijk de voordelen van het uitstellen van het moederschap of negatieve ervaringen met het moederschap. Ten tweede is de verwachting ook dat de invloedsprocessen minder goed werken als er sprake is van verschillende contacten (Loïs, 2016). Daarnaast is het in een minder verbonden netwerk aannemelijk dat een vrouw op minder sociale steun kan rekenen van haar netwerk. Gebaseerd op de mechanismen *sociale besmetting*, *sociale controle* en *sociale steun* verwacht ik dan ook dat vrouwen met een minder verbonden netwerk een lagere kinderwens hebben.

Hypothese 2: De persoonlijke netwerkstructuur hangt samen met de fertiliteitswensen van een vrouw.

- H2a: Een vrouw met een verbonden netwerktype heeft een hoge kinderwens.
- H2b: Een vrouw met een bi-factional netwerktype heeft een hoge kinderwens, maar minder hoog dan een vrouw met een verbonden netwerktype.
- H2c: Een vrouw met een niet-verbonden netwerktype heeft de laagste kinderwens.

2.2.4 Meerwaarde compositie van het netwerk

Aangezien op basis van enkel de structuur niet de volledige sociale context kan worden geschetst, wordt er in dit onderzoek tevens de compositie kenmerken over relaties toegevoegd (Chunaev, 2020). Vacca (2020) geeft aan dat met de structurele kenmerken een goede typologie kan worden gecreëerd waar de sociale context later aan toegevoegd kan worden. De interpretatie van de onderzoeker geeft in dat geval de sociale context. In dit onderzoek zal de sociale context zowel door mij zelf worden toegevoegd aan de structurele netwerktypen, maar er zal ook een nieuwe analyse worden gedaan. Het tweede model zal toetsen of de toevoeging van de compositie kenmerken, betreffende de sociale relaties, van het netwerk in staat is om een verbeterde netwerktypologie te maken. Om dit model te maken zal een verdeling worden gemaakt in vier groepen: naaste familieleden, andere familieleden, vrienden en kennissen en collega's. Op deze manier is het mogelijk voor elke vrouw het aantal per categorie mee te nemen in de clusteranalyse. Hierbij zullen de structurele kenmerken even zwaar wegen als de compositie kenmerken. De verwachting is dat de compositie kenmerken niet volledig samenvallen met de structurele kenmerken en dat ondanks dat het model complexer wordt, het tevens de persoonlijke netwerken beter weet in te delen. Ter illustratie; mocht het model verbeteren dan

kunnen netwerktypen gecreëerd worden zoals *dense family core-periphery*, vergelijkbaar met Lubbers en collega's (2007).

In voorgaand onderzoek is onderzocht welke sociale contacten invloed hebben op de fertiliteitskeuzes en -wensen van vrouwen. De meeste onderzoeken richten zich op één sociale relatie en de invloed op fertiliteitswensen en -keuzes, bijvoorbeeld broers en zussen. De verwachting is dat door de toevoeging van de compositie kenmerken een nieuwe verdeling kan worden gemaakt in de netwerktypen en dat clustering sterker samenhangt met fertiliteitswensen. Oftewel, de netwerktypen zullen naar verwachting sterker verschillen in de fertiliteitswensen. Op die manier kan gekeken worden of vrouwen met een *dense family core-periphery* netwerktype een hogere kinderwens hebben dan vrouwen met een *sparse work-related core-periphery* netwerktype.

Hypothese 3: De toevoeging van de compositie van het persoonlijke netwerk kan een betere typologie creëren dan enkel op basis van structurele kenmerken mogelijk is.

- H3a: De compositie van het netwerk biedt extra informatie om de persoonlijke netwerken van vrouwen beter in te delen in een juist netwerktype.
- H3b: De samenhang tussen de netwerktypen en de fertiliteitswensen van vrouwen neemt toe.

2.3 Huidige studie

Deze studie heeft als doel om de netwerken van Nederlandse vrouwen in kaart te brengen en de samenhang tussen type netwerken en fertiliteitswensen te onderzoeken. Dit zal worden gedaan door het uitvoeren van een clusteranalyse. Deze clusters worden allereerst gemaakt aan de hand van de structurele kenmerken en daarna aan de hand van zowel structurele als compositie kenmerken. Voor de clustering op basis van de structurele kenmerken wordt de methode van Vacca (2020) toegepast.

In deze studie zijn de volgende hypothesen geformuleerd:

Hypothese 1: De netwerktypen van vrouwen zijn vergelijkbaar met de netwerkstructuren uit het onderzoek van Vacca (2020).

Hypothese 2: De persoonlijke netwerkstructuur hangt samen met de fertiliteitswensen van een vrouw.

- H2a: Een vrouw met een verbonden netwerktype heeft een hoge kinderwens.
- H2b: Een vrouw met een bi-factional netwerktype heeft een hoge kinderwens, maar minder hoog dan een vrouw met een verbonden netwerktype.
- H2c: Een vrouw met een niet-verbonden netwerktype heeft de laagste kinderwens.

Hypothese 3: De toevoeging van de compositie van het persoonlijke netwerk kan een betere typologie creëren dan enkel op basis van structurele kenmerken mogelijk is.

- H3a: De compositie van het netwerk biedt extra informatie om de persoonlijke netwerken van vrouwen beter in te delen in een juist netwerktype.
- H3b: De samenhang tussen de netwerktypen en de fertiliteitswensen van vrouwen neemt toe.

3. Data & Methodes

In dit hoofdstuk zal aan bod komen op welke manier dit onderzoek is uitgevoerd en welke data hiervoor is gebruikt. Allereerst zal een beschrijving worden gegeven van de dataset. Daarna zal worden ingegaan op de operationalisering van centrale begrippen. Als laatste zal de methode van analyse worden gepresenteerd.

3.1 Beschrijving dataset

3.1.1 Het LISS-panel: Sociale relaties en kinderkeuzes

In deze scriptie wordt gebruikgemaakt van gegevens van het LISS-panel (Longitudinal Internet Studies for the Social Sciences) dat wordt beheerd door Centerdata (Universiteit van Tilburg, Nederland). Het LISS-panel is een representatieve steekproef van ongeveer 5.000 Nederlandse huishoudens met ongeveer 7.500 individuen. Deze steekproef is getrokken uit het bevolkingsregister van het Centraal Bureau voor de Statistiek. De personen in de steekproef vullen maandelijks vragenlijsten in en deze gegevens zijn verkrijgbaar voor wetenschappers en beleidsmakers.

Het LISS-panel geeft onderzoekers de mogelijkheid om gerichte data te verzamelen voor eigen onderzoek. Dr. G. Stulp heeft zijn eigen studie toegevoegd aan het panel genaamd ‘Sociale Relaties en Kinderkeuzes’. Deze studie is erop gericht te onderzoeken wat de sociale invloeden zijn op fertiliteitswensen. Alle vrouwen uit het LISS-panel tussen de 18 en 40 jaar zijn uitgenodigd om deel te nemen aan deze studie. Dit betrof een steekproef van 1332 vrouwen in de tijd tussen 20 februari 2018 en 27 maart 2018. In totaal voltooiden 758 vrouwen de enquête, wat resulteerde in een responspercentage van 56,9% (Stulp, 2020). Deze data zal gebruikt worden voor deze scriptie.

De vragen uit de Sociale Relaties en Kinderkeuzes-studie zijn op te delen in vier categorieën. Het eerste gedeelte richt zich op het verkrijgen van informatie over de persoonlijke situatie van de vrouw en haar fertiliteitswensen en -resultaten. In dit gedeelte komt onder andere naar voren of iemand een partner heeft, of iemand momenteel kinderen heeft en of er een toekomstige kindwens is.

Als tweede is de vrouw gevraagd om 25 namen van sociale relaties te noemen waarmee ze afgelopen jaar contact heeft gehad. De volledige vraag is: ‘Noem alstublieft 25 namen van personen die 18 jaar of ouder zijn en met wie u in het afgelopen jaar contact hebt gehad. Dit contact kan bestaan uit persoonlijke contact, maar ook contact via de telefoon, internet of e-mail. U kent deze personen en deze personen kennen u ook van naam en gezicht (denk aan vrienden, familie, kennissen, et cetera). U zou contact op kunnen nemen met deze personen wanneer dat zou moeten. Noem in ieder geval de naam van uw partner als u die hebt’. Dit wordt ook wel een naamgenerator genoemd. Waarna de

vrouw vragen kreeg over de persoonlijke informatie van deze sociale relaties, zoals het geslacht, wat voor relatie het is, hoe hecht deze relatie is en of deze persoon zelf kinderen heeft.

In het derde gedeelte van de enquête richtten de vragen zich op de onderlinge relaties tussen de sociale relaties van de vrouw. Dit wordt ook wel een edge interpreter genoemd. Concreet is de vrouwen de volgende vraag gesteld: ‘Met wie heeft PERSOON X contact? Met contact bedoelen we alle vormen van contact, zoals persoonlijk contact, contact via (mobiele) telefoon, post, email, sms, en andere manieren van online en offline communicatie’.

Als laatst kregen de vrouwen drie stellingen die in kaart kunnen brengen wat de mate van sociale druk is die de vrouwen ervaren en het plezier dat zij denken dat mensen met kinderen ervaren. Ter illustratie één stelling: ‘Mijn ouders/verzorgers vinden dat ik (meer) kinderen zou moeten krijgen’.

De volledige omschrijving van het onderzoek naar Sociale Relaties en Kinderkeuzes is terug te vinden op: <https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/EZCDOA>. Hier staat onder andere het codeboek van de studie, de correcties die zijn gedaan in de dataset en een omschrijving hoe de Graphical Ego-centered Network Survey Interface (GENSI) is gebruikt.

3.1.2 De dataset

Het verkrijgen van netwerkdata is erg intensief voor de respondenten van het onderzoek. In het onderzoek naar Sociale Relaties en Kinderkeuzes is de vrouwen gevraagd 25 alters te noemen. Elk netwerk geeft inzicht in de aanwezigheid of afwezigheid van een relatie tussen de alters. Per netwerk gaat het om 300 (mogelijke) relaties waar de respondent iets over dient in te vullen. Na analyse van invulsnelheid en de kwaliteit van antwoorden concludeert Stulp (2021) dat de resultaten erop wijzen dat de respondenten goed in staat zijn geweest om accuraat te antwoorden.

Alle vrouwen die onderdeel zijn van het LISS-panel in de leeftijdscategorie 18 tot en met 40 jaar ($N = 1332$) hebben een uitnodiging gekregen voor deze enquête. Hiervan hebben 758 vrouwen de enquête ingevuld. Dit komt neer op een response-rate van 56,9%. De vrouwen die het onderzoek hebben ingevuld lijken erg op de vrouwen die wel zijn uitgenodigd, maar het onderzoek niet hebben voldaan (Stulp, 2021). Na correctie is de netwerkdata van 706 Nederlandse vrouwen tussen de 18 en 40 jaar overgebleven. Vrouwen zijn uit de dataset verwijderd als ze minder dan 25 alters hebben genoemd, als ze geen alter-alter relaties hebben aangegeven, als ze de vragenlijst hebben ingevuld op hun iPad of telefoon, er fouten zaten in genoemde alter relaties en als een groot aantal van de verplichte alter attributen niet is aangegeven.

3.2 Operationalisering centrale begrippen

In deze paragraaf zullen de variabelen worden gepresenteerd die worden gebruikt om het model te maken. Bij elke variabele wordt vermeld hoe deze tot stand is gekomen en waarom er voor deze variabele is gekozen.

3.2.1 De uitkomstmaten

In deze scriptie wordt gekeken naar de samenhang tussen netwerktypen en fertiliteitswensen. De uitkomstmaat is daarom de fertiliteitswens van een vrouw. Er zijn twee variabelen die onafhankelijk van elkaar de fertiliteitswens van een vrouw vaststellen. Dit zijn ‘kinderwens’ en ‘ideale gezinsgrootte’.

De eerste uitkomstmaat is de kinderwens van de vrouw. Aan elke vrouw is de volgende vraag gesteld: ‘Denkt u in de toekomst (nog meer) kinderen te krijgen?’. Zij hadden de volgende antwoordmogelijkheden: absoluut niet (1), waarschijnlijk niet (2), weet ik niet (3), waarschijnlijk wel (4) en absoluut wel (5). Hiermee kan vastgesteld worden hoe groot de kindwens van de vrouw is; naarmate het cijfer toeneemt wordt de kindwens groter.

De tweede uitkomstmaat is de ideale gezinsgrootte van de vrouw. De vraag die aan de vrouwen is gesteld is: ‘Hoeveel kinderen zou u in totaal graag willen hebben? Dit is inclusief de X kinderen die u nu al hebt’. Hiervoor waren de volgende antwoordmogelijkheden mogelijk: 0 tot en met 10, meer dan 10, en weet ik niet. In totaal hebben 68 vrouwen aangegeven dat zij niet weten hoeveel kinderen zij graag zouden willen, dit is als missende waarde gerapporteerd. Ook hiervoor geldt dat een hoger cijfer refereert aan een grotere kindwens.

3.2.2 Beschrijvende statistieken van Nederlandse vrouwen

Om een volledig beeld te krijgen van de samenstelling van de steekproef zullen een aantal beschrijvende statistieken worden gegeven. De eerste vijf variabelen zullen niet meegenomen worden in het model, maar worden meegenomen om inzicht te geven van de samenstelling.

Voor elke beschrijvende statistiek geldt dat deze wordt gegeven over zowel de gehele steekproef als per netwerktype. De eerste beschrijvende statistiek is de leeftijd van de vrouw. De steekproef bestaat uit Nederlandse vrouwen tussen de 18 en 40 jaar. De vrouw is gevraagd naar haar huidige leeftijd en hiervan zal een gemiddelde en standaarddeviatie worden gegeven. De tweede beschrijvende statistiek is hoeveel procent van de vrouwen een partner heeft. De vrouwen konden hier antwoorden met ‘ja’ of ‘nee’. Hierbij wordt geen onderscheid gemaakt in het geslacht van de partner. Als derde wordt beschreven hoeveel procent van de vrouwen een biologisch kind, stiefkind, adoptiekind of pleegkind

heeft, deze variabele wordt heeft kind genoemd. Dit is een vraag waar enkel 'ja' of 'nee' op is geantwoord. Voor deze statistiek wordt ook een percentage gegeven. Daarnaast wordt ook beschreven hoeveel kinderen de vrouwen gemiddeld hebben. Hier zijn biologische kinderen, stiefkinderen, adoptiekinderen en pleegkinderen meegenomen. Op deze vraag hebben de vrouwen een antwoord tussen de 0 en 10 of meer dan 10 kunnen geven. Hierbij is het antwoord 'meer dan 10' niet voorkomen. De laatste beschrijvende variabele gaat over hoeveel biologische kinderen de vrouwen gemiddeld hebben. Hiervoor gelden dezelfde antwoordcategorieën als bij 'hoeveel kinderen'. Daarnaast hebben de vrouwen die aangaven geen kinderen te hebben automatisch een missende waarde op het aantal biologische kinderen. Deze missende waarde is omgezet naar 0 biologische kinderen. Ook hier wordt een gemiddelde en standaarddeviatie gegeven. Samenvattend, de beschrijvende statistieken zullen bestaan uit de leeftijd, het hebben van een partner, het hebben van kinderen, het aantal kinderen en het aantal biologische kinderen.

3.2.3 Beschrijvende statistieken voor clusters

Er zijn verschillende maten die aangeven wat de kenmerken van een cluster zijn. In deze studie zullen gemiddelde dissimilariteit, maximale dissimilariteit, scheiding en diameter worden gebruikt om de clusters te omschrijven (Kaufman & Rousseeuw, 1990). Deze maten worden hieronder beschreven. De k-medoid analyse gebruikt het meest representatieve datapunt uit een cluster om het cluster te vertegenwoordigen, dit wordt de medoid genoemd. In vergelijking met de k-means analyse wordt niet het gemiddelde van het cluster als referentiepunt genomen, maar een daadwerkelijk netwerk van een vrouw die het meest representatief is voor dat cluster. De medoid is het datapunt waarbij de afstand tussen alle andere punten het kleinst is. In deze scriptie wordt de Euclidische afstand als afstandsmaat gebruikt, vergelijkbaar met Vacca (2020).

De eerste twee maten die de kenmerken van het cluster beschrijven zijn de gemiddelde dissimilariteit en maximale dissimilariteit. De gemiddelde dissimilariteit is de gemiddelde afstand tussen de punten in het cluster en de medoid. Dit is een maat die aangeeft hoe vergelijkbaar de punten zijn. Daarnaast geeft de maximale dissimilariteit de grootste afstand tussen een punt in het cluster en de medoid. Deze maat geeft daardoor aan hoe verschillend de punten in het cluster zijn. De derde maat die wordt gebruikt om de clustering te omschrijven is de scheiding. Deze maat geeft aan wat de minimale afstand is tussen een punt van het cluster en een punt van een ander cluster. Als de scheiding hoog is, dan zullen er duidelijke verschillen zijn tussen de clusters. Als laatste geeft de diameter aan wat de grootste afstand is tussen twee punten in het cluster. Hiermee geeft het de randen van het cluster aan.

3.2.4 Variabelen voor de clusteranalyse

Naast de uitkomstmaten zijn er andere variabelen die gebruikt worden om het model te maken. In paragraaf 3.3 ga ik uitgebreid in op de manier van analyse, maar ik zal nu uiteenzetten welke variabelen hiervoor worden gebruikt.

De eerste variabele is de modulariteitsscore van het netwerk. De modulariteitsscore wordt berekend door gebruik te maken van het Girvan-Newman-algoritme. Deze maat geeft aan in hoeverre de indeling in subgroepen geschikt is voor het netwerk (Newman & Girvan, 2004). Voor de clusteranalyse is het noodzakelijk dat variabelen worden gestandaardiseerd, zodat elke variabele een gemiddelde heeft van 0 en een standaarddeviatie van 1. Er is gestandaardiseerd door de functie `scale()` in R te gebruiken. De modulariteitsscore vanuit de GN-algoritme zal daarom voor elk netwerk worden gestandaardiseerd. Op deze manier weegt elke variabele in de clusteranalyse even zwaar en kan de data vergeleken worden ondanks dat het niet op dezelfde manier is gemeten.

De tweede variabele geeft het aantal subgroepen aan. Allereerst worden op basis van de GN-algoritme communities gemaakt. Elke community die drie of meer alters heeft wordt gedefinieerd als een subgroep. De variabele geeft aan hoeveel van deze subgroepen voorkomen in het netwerk. De laatste variabele geeft het aantal isolaten en dyads van het persoonlijke netwerk. Dit zijn alle communities uit de GN-algoritme die uit één of twee alters bestaan. Zowel het aantal subgroepen als het aantal isolaten en dyads wordt op dezelfde manier gestandaardiseerd als de modulariteitsscore.

In dit onderzoek worden ook de compositie kenmerken gericht op sociale relaties meegenomen. De variabele die hieruit voortkomt heet relatie. Aan elke vrouw is gevraagd per alter aan te geven welke relatie zij heeft met deze persoon. De antwoordmogelijkheden zijn als volgt: Dit is mijn partner (1), vader/moeder (2), broer/zus (3), ander familielid (bijvoorbeeld oom/tante, neef/nicht) (4), familielid van partner (5), kennis/vriend(in) van partner (6), van de basisschool (7), van de middelbare school (8), studeren (9), via werk (10), via sociale activiteit (sport, hobby, kerk) (11), via een gezamenlijke kennis/vriend(in) (12), buurtgenoot (13), via een andere manier, namelijk... (14).

Om deze variabele te gebruiken zijn vier categorieën gemaakt, namelijk: naaste familieleden (1), andere familieleden (2), vrienden en kennissen (3) en collega's (4). Bij de eerste groep zijn de partner, vader/moeder en broer/zus meegenomen. De tweede groep bestaat uit andere familieleden (bijvoorbeeld oom/tante, neef/nicht) en de familieleden van partner. De derde groep bevat de meeste groepen, namelijk kennis/vriend(in) van partner, van de basisschool, van de middelbare school, studeren, via sociale activiteit (sport, hobby, kerk), via een gezamenlijke kennis/vriend(in), buurtgenoot, via een andere manier, namelijk... Als laatste bevat de vierde categorie enkel via werk. Voor elk netwerktype zal beschreven worden hoeveel aantallen er gemiddeld voorkomen per

categorie. Daarnaast zal deze variabelen worden toegevoegd in het tweede model voor de clusteranalyse. Voor elke vrouw geldt dat zij 25 sociale relaties heeft gedefinieerd. Drie van de vier categorieën zullen daarom als gestandaardiseerde waarde worden meegenomen, dit zijn het aantal naaste familieleden, het aantal andere familieleden en het aantal collega's. Ook hier is op dezelfde manier gestandaardiseerd als bij de voorgaande variabelen.

Samenvattend, er zullen vier variabelen worden gebruikt om de persoonlijke netwerken van de Nederlandse vrouwen in deze steekproef te clusteren. De drie structurele kenmerken zijn de modulariteitscore, het aantal subgroepen en het aantal isolaten en dyads (Vacca, 2020). Het vierde kenmerk bestaat uit het aantal sociale relaties per categorie, dit worden de compositie kenmerken genoemd.

3.3 Methode van analyse

In deze paragraaf zal het analyseplan worden gepresenteerd. Allereerst komt de methode naar voren die Vacca (2020) gebruikt. Daarna zal aan bod komen hoe de compositie kenmerken hieraan worden toegevoegd.

3.3.1 Beschrijving methode van Vacca (2020)

De methode die Vacca (2020) beschrijft wordt de 'op subgroep gebaseerde methode' genoemd. Deze methode bestaat uit drie onderdelen, namelijk de identificatie van subgroepen, het samenvatten van de subgroep structuur en een k-medoid clusteranalyse. Deze drie onderdelen worden in deze volgorde uitgevoerd:

1) Identificatie van subgroepen

De identificatie van subgroepen wordt gedaan door middel van de Girvan-Newman-algoritme (GN-algoritme), dit wordt een *community detection algorithm* genoemd (Newman & Girvan, 2004). Dit algoritme maakt gebruik van de *edge-betweenness* als basis. Deze maat berekent hoeveel knooppunten nodig zijn om de kortste route om elk paar knooppunten te verbinden. Oftewel, in een persoonlijk netwerk van 25 knooppunten zijn er 300 ($25 \times 24 / 2$) paren van knooppunten en van elke daadwerkelijke verbinding kan een *edge-betweenness* worden uitgerekend. De verbindingen met een hoge *edge-betweenness* score zijn interessant voor het detecteren van subgroepen, doordat deze verbinding door veel andere knooppunten gebruikt moet worden om de kortste route af te leggen. In de GN-algoritme wordt telkens de verbinding met de hoogste *edge-betweenness* score verwijderd, waarna opnieuw de *edge-betweenness* score voor de reeds bestaande verbindingen wordt berekend. Dit wordt uitgevoerd totdat elke community niet meer met elkaar verbonden is (Newman & Girvan,

2004). Als deze identificatie heeft plaatsgevonden dan kunnen de resultaten van het aantal subgroepen, het aantal isolaten en dyads en de modulariteitsscore worden opgevraagd.

2) Samenvatten van de subgroep structuur

Voor het samenvatten van de subgroepstructuur worden drie statistische maten gebruikt, namelijk de modulariteit, het aantal isolaten en dyads, en het aantal subgroepen met drie of meer alters.

De eerste statistische maat, modulariteit, is ontworpen om de kwaliteit van de indeling van het netwerk te kunnen testen (Newman & Girvan, 2004). Deze maat vergelijkt de daadwerkelijke banden in subgroepen in het netwerk met willekeurige banden en vertaalt dit naar een modulariteitsscore tussen de 0 en de 1. Als deze score hoog is dan is er een hoge mate van verbondenheid binnen de subgroepen en een lage mate van verbondenheid tussen de subgroepen. Een lage score houdt in dat de indeling in subgroepen niet optimaal is en er geen hoge verbondenheid is binnen de subgroepen.

De andere twee statistische maten richten zich op de subgroep verdeling. De tweede maat geeft het aantal isolaten en dyads aan. Oftewel, degene die niet in contact staan met anderen of twee personen die enkel met elkaar in contact staan. Als er sprake is van veel isolaten en dyads dan kun je spreken van een versnipperd netwerk. Als laatste geeft de derde maat hoeveel subgroepen er zijn met drie of meer personen. De reden dat Vacca (2020) een subgroep heeft gebaseerd op drie individuen is de sociologische definitie van een sociale groep (Simmel, 1908).

3) K-medoid clusteranalyse

Met de K-medoid clusteranalyse worden de persoonlijke netwerken in k -clusters ingedeeld, met andere woorden in structurele netwerktypen. Alle clusters krijgen één waarde in de gegevens, de medoid. Dit is het meest representatieve datapunt binnen het cluster. De medoid is de waarde van dit cluster, waarbij het gemiddelde verschil tot elk van de persoonlijke netwerken in het cluster is geminimaliseerd. Op deze manier is het mogelijk om elk persoonlijke netwerk aan één van de clusters toe te voegen. Om het meest geschikte aantal clusters te vinden heeft Vacca (2020) deze procedure 19 keer uitgevoerd, namelijk met 2 - 20 clusters. Op basis van twee statistische maten *silhouette* en *Akaike Information Criterion* (AIC) wordt gekeken wat het meest optimale aantal clusters is. De combinatie van deze twee maten moet de doorslag geven voor het aantal clusters.

- *Silhouette*: Deze maat laat zien in hoeverre een persoonlijke netwerk vergelijkbaar is met het cluster waarin het is geplaatst ten opzichte van de andere clusters. Dit is een waarde tussen de -1 en 1. Een hoge waarde geeft aan dat een persoonlijk netwerk goed overeenkomt met het cluster waarin het is geplaatst.
- *AIC*: Deze maat geeft aan hoeveel informatie er verloren gaat bij het gebruiken van een bepaald aantal clusters (Akaike, 1974). Het is een foutenmaat, vergelijkbaar met de Sum of Squared Errors. De schatting wordt gemaakt door de voorspellingsfouten te vergelijken met

de andere k-medoid modellen. Het is daarom mogelijk om zowel het aantal clusters als het aantal variabelen met elkaar te vergelijken. Hierbij staat een lage score voor een goed model; het model maakt dan de minste fouten. In een visualisatie van de AIC is te zien dat de voorspellingsfouten eerst aanzienlijk afnemen en daarna is een ‘elleboog’ te zien, de waarden nemen nog wel af maar deze afname is klein. Deze ‘elleboog’-waarde is het meest gunstig voor het aantal clusters.

3.3.2 Beschrijving methode van huidige studie

In de huidige studie wordt de methode van Vacca (2020) volledig overgenomen zoals hierboven beschreven. Op basis hiervan zal ook hypothese 1 en 2 worden beantwoord. Daarna zal de analyse opnieuw worden uitgevoerd, maar wordt in onderdeel 2 tevens de compositie van het netwerk toegevoegd. Het hernieuwde onderdeel ziet er dan als volgt uit:

4) Samenhang met fertiliteitswensen

Na de clustermethode van Vacca (2020) zal de samenhang worden gezocht tussen de netwerktypen en de fertiliteitswensen van de vrouw. Om te toetsen of de vrouwen verschillen op ‘kinderwens’ en ‘ideale gezinsgrootte’ zal een Kruskal Wallis toets worden uitgevoerd.

5) Clustering met compositie kenmerken

Naast de drie eerder genoemde statistische maten wordt de compositie aan het model toegevoegd. Er is ervoor gekozen om één onderdeel van de compositie van het netwerk mee te nemen, namelijk de sociale relatie tussen *ego* en *alter*. Waarna opnieuw onderdeel 3 wordt uitgevoerd. Op basis van de combinatie tussen de maten Silhouette en AIC is te zien of de netwerktypologie is verbeterd. Hierbij zal de waarde van de AIC doorslaggevend zijn, omdat deze maat er tevens op gericht is te onderzoeken of een model significant beter is geworden door het toevoegen van een extra variabele. Hierna zal de samenhang met fertiliteitswensen opnieuw worden onderzocht.

4. Resultaten

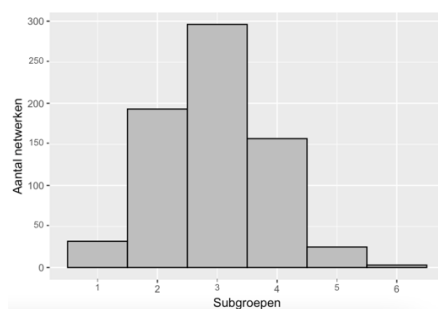
In dit hoofdstuk worden de resultaten van de clusteranalyse besproken. Het hoofdstuk is opgebouwd uit vier onderdelen. De eerste paragraaf geeft de beschrijvende statistieken van de drie structurele kenmerken, de uitgevoerde k-medoid analyse, de clusters en medoids en de bivariate verdelingen van de structurele kenmerken per netwerktype. In de tweede paragraaf wordt een vergelijking gemaakt met de netwerktypen van Vacca (2020) en wordt de netwerktypologie voor deze studie vastgesteld. Daarna worden in paragraaf drie de resultaten van de multivariate analyses gepresenteerd. Hier wordt gekeken naar de samenhang tussen de structurele netwerktypen en de kindwens. Als laatste worden de resultaten gepresenteerd van de clustermethode met zowel compositie als structurele kenmerken en wordt een vergelijking gemaakt tussen deze twee kenmerken.

4.1 Beschrijvende statistieken

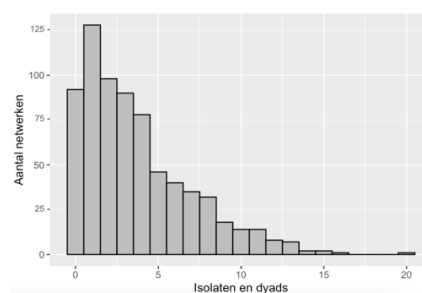
In deze paragraaf worden verschillende univariate en bivariate statistieken gepresenteerd.

4.1.1 Univariate verdeling structurele kenmerken

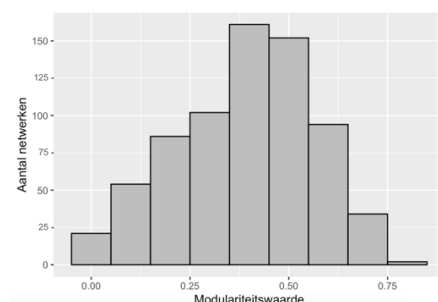
In de methode van Vacca (2020) worden drie structurele kenmerken gebruikt. In deze subparagraaf zullen we de verdeling in de dataset bespreken en daarna ingaan op de verschillen tussen de netwerktypen.



Figuur 2: Verdeling aantal subgroepen



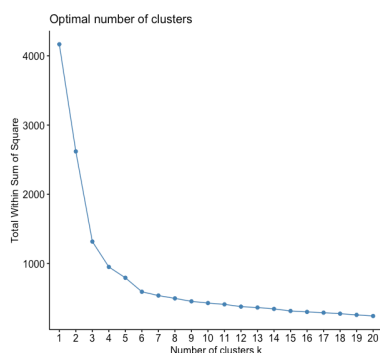
Figuur 3: Verdeling aantal isolaten en dyads



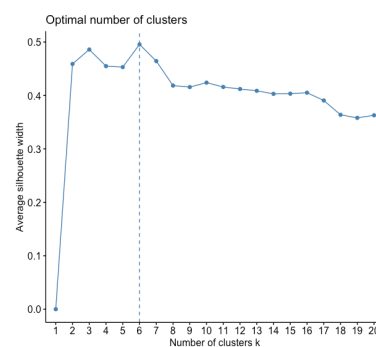
Figuur 4: Verdeling van modulariteitscores

De univariate verdeling van het aantal subgroepen laat zien dat de meeste vrouwen 2, 3 of 4 subgroepen in hun persoonlijke netwerk hebben (Figuur 2). Daarentegen worden er weinig vrouwen gevonden die 1 subgroep hebben, of waarbij 5 of 6 subgroepen aanwezig zijn. Ook blijkt dat de verdeling van het aantal isolaten en dyads rechtsscheef is (Figuur 3). Dit betekent dat de meeste vrouwen weinig isolaten en dyads in hun netwerk hebben. Oftewel, veel alters zijn onderdeel van een subgroep. In het onderzoek van Vacca (2020) wordt onderscheid gemaakt in netwerktypen op basis van een verschil in aantal isolaten en dyads met een veelvoud van 5 (Tabel 2). Doordat in dit onderzoek aanzienlijk minder isolaten en dyads worden gevonden zal de indeling in netwerktypen mogelijk lastig zijn. Dit komt doordat de scheiding tussen netwerktypen minder duidelijk is. Als laatste is de modulariteitscore symmetrisch verdeeld (Figuur 4). Dit wijst erop dat er enerzijds netwerken zijn die duidelijke subgroepen hebben met een hoge interne samenhang, namelijk een hoge modulariteitscore en anderzijds zijn er ook netwerken zonder duidelijke subgroepen, dit wordt gekenmerkt door samenhang tussen de subgroepen.

4.1.2 K-medoid analyse, clusters en medoids



Figuur 5: AIC-score, structurele kenmerken



Figuur 6: Silhouette-score, structurele kenmerken

Op basis van de modulariteitscore, het aantal subgroepen, en het aantal isolaten en dyads is de k-medoid analyse uitgevoerd. De AIC-score laat twee belangrijke punten zien, namelijk bij drie en zes clusters (Figuur 5). Tot de drie clusters dalen de voorspellingsfouten sterk, waarna de daling afzwakt. Echter, de daling wordt na de zes clusters pas vlak, waardoor dit gezien wordt als het ‘elleboog-punt’. Bij de vergelijking in silhouette-scores geven tevens drie en zes clusters de hoogste scores (Figuur 6). Hiervoor geldt dat zes clusters net hoger scoort. De AIC-score voor zes clusters is lager dan 500 en de silhouette-score is 0,50. De volledige uitwerking van deze identificatie is te vinden in Bijlage 1. De typologie op basis van structurele kenmerken bestaat dan ook uit zes persoonlijke netwerktypen.

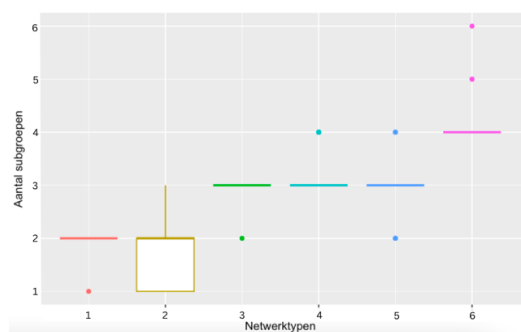
Tabel 3: Beschrijvende statistieken van netwerktypen ($n = 706$)

| | Netwerktype 1 ($n = 106$) | Netwerktype 2 ($n = 113$) | Netwerktype 3 ($n = 117$) | Netwerktype 4 ($n = 128$) | Netwerktype 5 ($n = 88$) | Netwerktype 6 ($n = 154$) |
|-------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|-------------------------------|--------------------------------|
| Gemiddelde dissimilariteit | 0,79 | 1,09 | 0,50 | 0,70 | 0,98 | 0,81 |
| Maximale dissimilariteit | 2,14 | 4,05 | 1,21 | 1,52 | 2,44 | 2,36 |
| Diameter | 3,15 | 5,67 | 2,07 | 2,39 | 3,15 | 3,22 |
| Scheiding | 0,10 | 0,10 | 0,09 | 0,09 | 0,14 | 0,30 |

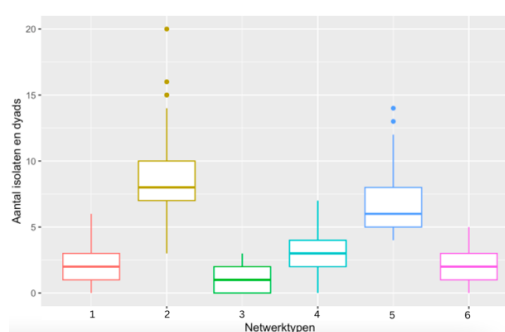
De clusters hebben verschillende kenmerken, namelijk de gemiddelde dissimilariteit, de maximale dissimilariteit, de diameter en de scheiding. De lage dissimilariteit van netwerktype 3 geeft aan dat de punten in dit cluster het meest vergelijkbaar zijn, want de gemiddelde afstand tot de medoid is het kleinst ($gem\ diss_3 = 0,50$). Dit gaat gepaard met de laagste maximale dissimilariteit ($max\ diss_3 = 1,21$). Oftewel, in dit cluster is de afstand het kleinst tussen de medoid en het meest afwijkende punt. Daarnaast zijn de scores op gemiddelde dissimilariteit en maximale dissimilariteit van netwerktypen 1, 4 en 6 vergelijkbaar. Deze netwerktypen hebben een gemiddelde dissimilariteitscore tussen de 0,70 en 0,81, wat beduidend hoger is dan netwerktype 3. De interne samenhang in deze typen zal daarom minder zijn, waardoor de netwerken binnen deze clusters verder van elkaar afwijken, op de drie structurele kenmerken, dan in netwerktype 3. Als laatste hebben netwerktypen 2 en 5 beide een hogere gemiddelde dissimilariteitscores, waarbij netwerktype 2 ook een aanzienlijk hogere maximale dissimilariteitscore heeft ($max\ diss_2 = 4,05$). Dit wijst erop dat hier het grootste verschil bestaat tussen de medoid en het meest afwijkende punt; de netwerken in netwerktype 2 zullen grotere verschillen laten zien in structurele kenmerken dan in de andere clusters.

De diameter van een cluster geeft aan wat de grootste afstand is tussen twee punten in een cluster. Ook hier laat netwerktype 2 de grootste omvang zien ($diameter_2 = 5,67$). Dit wijst erop dat de structurele kenmerken van de netwerken in netwerktype 2 het meest van elkaar verschillen. Dit is het laagst in netwerktype 3, waardoor de netwerken hier het meest op elkaar zullen lijken ($diameter_3 = 2,07$). De resultaten over de scheiding laten erg kleine aantallen zien. Dit betekent dat de afstanden van punten tussen clusters niet groot zijn. Dit wijst erop dat er mogelijk overlap zit tussen netwerktypen en dat het onderscheid tussen clusters soms niet duidelijk is ($scheiding = 0,09 - 0,30$).

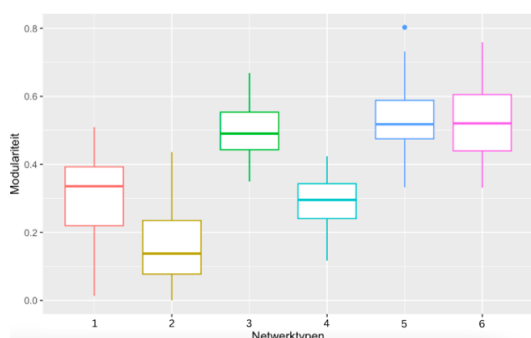
4.1.3 Bivariate verdeling structurele kenmerken per netwerktype



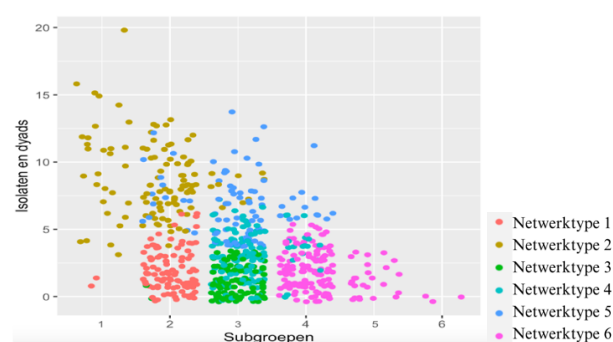
Figuur 7: Verdeling aantal subgroepen



Figuur 8: Verdeling aantal isolaten en dyads.



Figuur 9: Verdeling van modulariteitsscores



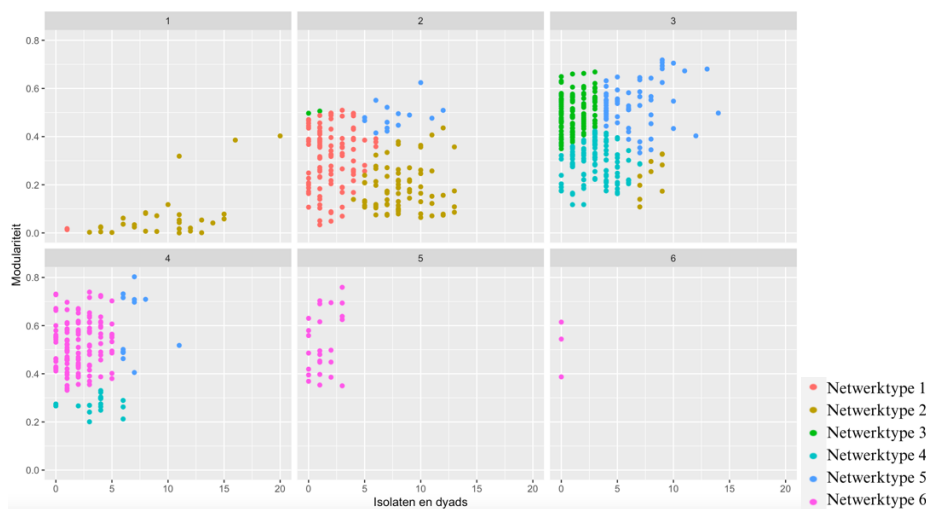
Figuur 10: Verdeling in netwerktypen op basis van subgroepen en isolaten en dyads

Tabel 4: Beschrijving van structurele kenmerken per netwerktype ($n = 706$)

| | Netwerktype 1 ($n = 106$) | Netwerktype 2 ($n = 113$) | Netwerktype 3 ($n = 117$) | Netwerktype 4 ($n = 128$) | Netwerktype 5 ($n = 88$) | Netwerktype 6 ($n = 154$) |
|---|-----------------------------|-----------------------------|-----------------------------|-----------------------------|----------------------------|-----------------------------|
| Modulariteit medoid | 0,33 | 0,16 | 0,49 | 0,31 | 0,53 | 0,51 |
| Aantal subgroepen medoid | 2 | 2 | 3 | 3 | 3 | 4 |
| Aantal isolaten en dyads medoid | 2 | 8 | 1 | 3 | 6 | 2 |
| Gemiddelde en sd modulariteit | 0,31 (0,13) | 0,16 (0,11) | 0,50 (0,07) | 0,29 (0,07) | 0,54 (0,10) | 0,52 (0,10) |
| Gemiddelde en sd aantal subgroepen | 1,98 (0,14) | 1,82 (0,57) | 2,98 (0,13) | 3,14 (0,35) | 2,99 (0,56) | 4,20 (0,44) |
| Gemiddelde en sd aantal isolaten en dyads | 1,97 (1,53) | 8,66 (2,85) | 1,20 (1,07) | 3,10 (1,70) | 6,75 (2,32) | 1,92 (1,46) |

De verdeling van het aantal subgroepen laat zien dat in bijna elk netwerktype sprake is van een duidelijk onderscheid in aantal subgroepen, op enkele outliers na (Figuur 7). Zo is er in netwerktype 3

(groene kleur) sprake van drie subgroepen (Figuur 7 en Figuur 10). Daarentegen is een grotere spreiding gevonden bij netwerktype 2, een groot deel van de netwerken heeft één subgroep en een ander deel twee subgroepen (Figuur 7 en 10). Dit verklaart ook dat de beschrijvende statistieken over de clusters de grootste afwijkingen lieten zien in netwerktype 2 (Tabel 3). Daarnaast laat de verdeling van de isolaten en dyads zien dat het aantal isolaten en dyads in de meeste netwerken relatief laag is (Figuur 8). De spreiding is het grootst in netwerktype 2, waar een netwerk wordt gevonden met 20 isolaten en dyads. Dit is erg hoog, aangezien er in totaal 25 alters worden opgegeven. Als laatst laten de modulariteitscores zien dat netwerktypen met minder subgroepen ook een gemiddeld lagere modulariteitscore hebben, namelijk netwerktype 1 en 2. Opvallend is dat netwerktype 1 en 3 hoger scoren dan andere netwerktypen met evenveel subgroepen, dit wijst op een hoge verbondenheid van alters binnen de subgroepen en een goede scheiding tussen de subgroepen.



Figuur 11: Verdeling van modulariteit en isolaten en dyads per aantal subgroepen



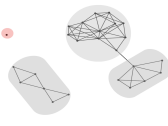
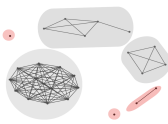
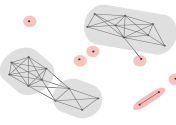
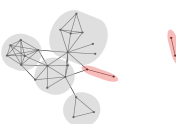
De zes spreidingsdiagrammen vertegenwoordigen elk een groep met een bepaald aantal subgroepen, waarbij de kleuren de netwerktypen aangeven (Figuur 11). De resultaten laten het onderscheid zien tussen de verschillende netwerktypen, bijvoorbeeld binnen de categorie 4 subgroepen laat netwerktype 4 een duidelijk lagere modulariteitscore zien en heeft netwerktype 5 meer isolaten en dyads (Figuur 11, spreidingsdiagram 4). Deze resultaten laten zien dat het aantal subgroepen bepalend is voor de modulariteitscore, want in de eerste spreidingsdiagram zijn enkel lage scores gevonden en deze scores nemen toe naarmate het aantal subgroepen toeneemt. Echter, vanaf vier subgroepen wordt dit minder (Figuur 11). Daarnaast is de samenhang tussen het aantal isolaten en dyads en de modulariteit veel minder sterk. Dit is te zien doordat er geen stijging lijkt te zijn binnen elk diagram als het aantal isolaten en dyads toeneemt. Deze resultaten zijn vergelijkbaar met het onderzoek van Vacca (2020).

4.2 Definiëren van de netwerktypologie

In deze paragraaf wordt gekeken naar de verschillende netwerktypen. Op basis van de structurele kenmerken zal een vergelijking worden gemaakt met de veelvoorkomende netwerktypen uit het onderzoek van Vacca (2020). Daarnaast zal de typologie uit dit onderzoek worden gepresenteerd.

4.2.1 Vergelijking met netwerktypen Vacca (2020)

Tabel 5: Beschrijving gevonden netwerktypen

| Netwerk- type | Omschrijving van de medoid* | Schematisch plaatje** | Vergelijkbaar netwerktype van Vacca (2020) | Classificatie binnen huidige studie |
|------------------|---|--|--|---|
| 1 | In het schematische plaatje is te zien dat de medoid bestaat uit 2 subgroepen. Daarnaast bestaan er enkele isolaten en dyads. Daarnaast is te zien dat er een redelijk aantal relaties zijn tussen de subgroepen. |  | Strongly bi-factional | Strongly bi-factional |
| 2 | Dit netwerktype heeft 2 subgroepen en meerdere isolaten en dyads. Er zijn veel onderlinge relaties tussen de subgroepen, maar ook met de isolaten en dyads. |  | Weakly bi-factional | Weakly bi-factional |
| 3 | In dit netwerktype is sprake van 3 duidelijk te scheiden subgroepen. Daarnaast is er sprake van een gering aantal isolaten en dyads. Er zijn niet veel onderlinge relaties tussen de subgroepen. |  | Tri-factional | Strongly tri-factional |
| 4 | In dit netwerktype zijn er ook 3 duidelijke subgroepen. Daarnaast zijn er meerdere subgroepen met 1 of 2 alters. |  | Dense multi-factional | Weakly tri-factional |
| 5 | In dit netwerktype zijn er tevens 3 duidelijke subgroepen met 3 of meer alters. Daarnaast zijn er meerdere isolaten en dyads. Onderling bestaan er een redelijk aantal banden. |  | Sparse multi-factional | Sparse tri-factional |
| 6 | In dit netwerktype zijn er 4 subgroepen. Daarbij zijn er ook isolaten en dyads. Er zijn een aantal banden onderling aanwezig, maar de onderlinge banden zijn sterk. |  | Dense multi-factional | Dense multi-factional |

* Medoid het meest representatieve datapunt van het cluster; niet alle netwerken in het cluster komen exact overeen.

** Grijs: subgroepen met meer dan drie alters, rood: isolaten en dyads.

De eerste hypothese in dit onderzoek is dat de netwerktypen van vrouwen vergelijkbaar zijn met de netwerkstructuren uit het onderzoek van Vacca (2020). In Tabel 5 onder de kolom ‘vergelijkbaar netwerktype in Vacca (2020)’ wordt aangegeven hoe dit netwerktype zou zijn gekwalificeerd in zijn

studie. Dit is gedaan aan de hand van Tabel 2. Hieruit blijkt dat er vier netwerktypen worden geïdentificeerd waarbij de vrouwen drie of meer subgroepen in hun persoonlijke netwerk hebben. Dit maakt dat de indeling van Vacca (2020) niet toereikend is en dat het dense multi-factional netwerktype tweemaal voorkomt. Vergelijkend met de veelvoorkomende netwerktypen van Vacca (2020) komen alleen het strongly bi-factional, weakly bi-factional en sparse multi-factional netwerktype overeen.

De eerste drie van de zes veelvoorkomende netwerktypen van Vacca (2020) zijn gebaseerd op het hebben van één subgroep in het persoonlijke netwerk, namelijk: closed community, dense core-periphery en sparse core-periphery. Er is niet een netwerktype te onderscheiden waarbij de medoid één subgroep heeft, waardoor deze drie veelvoorkomende netwerktypen niet voorkomen in deze studie. Daarentegen zijn er in deze studie 32 vrouwen gevonden met slechts één subgroep in het netwerk. Opvallend is dat het netwerktype closed community niet voorkomt in de data; er is geen vrouw waarbij alle alters goed met elkaar zijn verbonden. De 32 vrouwen met één subgroep zullen waarschijnlijk een netwerktype hebben die het beste is te beschrijven als een *dense core-periphery* of *sparse core periphery*. Twee van deze vrouwen vallen onder netwerktype 1, zij hebben één subgroep met drie of meer alters en daarnaast nog één isolaat of dyad. De andere vrouwen vallen onder netwerktype 2, zij hebben meerdere isolaten en dyads; variërend van drie tot twintig. Vermoedelijk door het geringe aantal vrouwen met enkel één subgroep is er geen apart netwerktype binnen de netwerktypologie van deze studie.

Daarnaast komt zowel *weakly bi-factional* als *strongly bi-factional* vaak voor in de verschillende studies van Vacca (2020). In deze studie is dit patroon ook terug te zien bij netwerktype 1 en netwerktype 2. In beide gevallen is in het merendeel van de netwerkdata sprake van twee subgroepen, waarbij de vrouwen in netwerktype 2 meer isolaten en dyads hebben. Daarbij hebben deze vrouwen ook een lagere modulariteitsscore. Het eerste netwerktype kan dan ook worden gekwalificeerd als *strongly bi-factional* en het tweede netwerktype als *weakly bi-factional*.

In de huidige studie hebben netwerktype 3 tot en met 6 drie of meer subgroepen. Op basis van de studie van Vacca (2020) zou het betekenen dat één wordt gekwalificeerd als tri-factional en de rest als multi-factional. Hij maakt namelijk geen onderscheid in verschillende tri-factional netwerktypen. Hierdoor zouden de andere twee netwerktypen met drie subgroepen vallen onder de multi-factional netwerktypen, waardoor het dense multi-factional netwerktype tweemaal voorkomt. In de eigen typologie zal verder onderscheid worden gemaakt tussen deze netwerktypen.

Er is een redelijke overlap tussen veelvoorkomende netwerken bij Vacca (2020) en de gevonden netwerken in dit onderzoek. Echter, er waren te weinig vrouwen waarbij maar één subgroep werd gevonden om hier een apart netwerktype van te maken. Hierdoor worden de netwerktypen *close community*, *dense core-periphery* en *sparse core-periphery* niet geïdentificeerd in deze

netwerktypologie. Daarentegen zijn de andere drie veelvoorkomende netwerktypen wel gevonden op basis van de k-medoid analyse. Dit zijn *strongly bi-factional*, *weakly bi-factional* en *sparse multi-factional*. Doordat maar de helft van de netwerktypen overeenkomen kan de eerste hypothese niet worden aangenomen.

4.2.2 Netwerktypologie Nederlandse vrouwen

Zoals blijkt uit subparagraaf 4.2.1 zijn de verschillende netwerktypen die Vacca (2020) aandraagt niet toereikend om de netwerken in dit onderzoek te typeren. Daarom zal nu een eigen indeling worden gemaakt, met als uitgangspunt de onderverdeling die Vacca (2020) maakt. Het onderscheid bestaat uit *dense-sparse*, *weakly-strongly* en aantal subgroepen met drie of meer.

Het eerste netwerktype wordt gekwalificeerd als strongly bi-factional. Dit netwerktype wordt gekenmerkt door twee subgroepen met daarnaast soms een aantal isolaten en dyads. Daarbij is de modulariteitsscore redelijk hoog, dit wijst erop dat er sprake is van een hoge samenhang binnen de subgroepen en een goede scheiding tussen de subgroepen ($mod_1 = 0,31$). In het tweede netwerktype zijn tevens twee subgroepen geïdentificeerd, ondanks de 30 vrouwen die maar één subgroep hebben. Alleen het hoge aantal isolaten en dyads en daarnaast een lage modulariteitsscore wijzen erop dat deze twee subgroepen niet dominant aanwezig zijn en dat ze onderling ook veel banden hebben ($mod_2 = 0,16$). Hierdoor komt het overeen met het weakly bi-factional netwerktype.

In het huidige onderzoek zijn er in totaal vier netwerktypen met drie of meer subgroepen, dit is veel in vergelijking met de andere datasets van Vacca (2020). Hierdoor wordt er in de indeling van Vacca (2020) niet voldoende onderscheid gemaakt voor de huidige typologie (Tabel 2 en Tabel 5). Om die reden is ervoor gekozen om het tri-factional netwerktype onder te verdelen op dezelfde manier als bij de bi-factional netwerktypen. Allereerst heeft netwerktype 3 een hoge modulariteitsscore en weinig isolaten en dyads in het netwerk zitten; dit maakt dat het netwerktype wordt gekwalificeerd als strongly tri-factional ($mod_3 = 0,50$). Ten tweede laat netwerktype 4 meerdere isolaten en dyads zien in combinatie met een lage modulariteitsscore ($mod_4 = 0,29$). Dit wijst erop dat de drie subgroepen in dit netwerktype onderling nog sterk verbonden zijn. Hierdoor kan dit netwerk gekwalificeerd worden als een weakly tri-factional netwerktype. Het vijfde netwerktype heeft ook drie subgroepen, met nog meer isolaten en dyads. Echter, de modulariteitsscore is redelijk hoog, waardoor de indeling in subgroepen goed de structuur van het netwerk kan tonen ($mod_5 = 0,54$). Alleen doordat er ook sprake is van meerdere subgroepen is deze modulariteit ook hoger. Om die reden is ervoor gekozen om dit netwerktype te kwalificeren als sparse tri-factional.

Als laatste komt netwerktype zes overeen met het dense multi-factional netwerktype van Vacca (2020). Er is sprake van een hoge modulariteitsscore ($mod_6 = 0,51$) en er zijn redelijk wat isolaten en dyads aanwezig.

Samenvattend, de typologie in dit onderzoek bestaat uit zes netwerktypen. Deze zijn gedefinieerd als: strongly bi-factional, weakly bi-factional, strongly tri-factional, weakly tri-factional, sparse tri-factional en dense multi-factional (Tabel 5). Deze benamingen zullen nu leidend zijn in dit stuk.

4.3 Multivariate analyses

In deze paragraaf zullen allereerst de kenmerken van de vrouwen worden besproken die er per netwerktype worden gevonden. Daarna zal worden ingegaan op de samenhang tussen de geïdentificeerde netwerktypen en de fertiliteitswensen van de vrouwen.

4.3.1 Kenmerken van de vrouwen in elk netwerktype

Tabel 6: Beschrijvende statistieken kenmerken vrouwen ($n = 706$): gemiddelde (standaarddeviatie) of percentage.

| | Alle netwerktypen | Strongly bi-factional ($n_1 = 106$) | Weakly bi-factional ($n_2 = 113$) | Strongly tri-factional ($n_3 = 117$) | Weakly tri-factional ($n_4 = 128$) | Sparse tri-factional ($n_5 = 88$) | Dense multi-factional ($n_6 = 154$) |
|-----------------------------|-------------------------|---------------------------------------|-------------------------------------|--|--------------------------------------|-------------------------------------|---------------------------------------|
| Leeftijd | 29,13 (6,49) | 29,01 (6,54) | 30,07 (6,63) | 27,57 (6,81) | 29,59 (6,29) | 30,42 (6,09) | 28,58 (6,27) |
| Partner | 72,80% ja 27,20% nee | 77,36% ja 22,64% nee | 71,68% ja 38,32% nee | 70,08% ja 29,92% nee | 78,12% ja 21,88% nee | 61,36% ja 38,64% nee | 74,68% ja 25,32% nee |
| Kinderen | 35,41% ja 64,59% nee | 37,74% ja 62,26% nee | 37,17% ja 62,83% nee | 25,64% ja 74,35% nee | 45,31% ja 54,69% nee | 36,26% ja 63,74% nee | 31,17% ja 68,83% nee |
| Aantal kinderen | 0,65 (1,04) | 0,69 (1,01) | 0,70 (1,02) | 0,44 (0,81) | 0,91 (1,29) | 0,66 (1,08) | 0,53 (0,92) |
| Aantal biologische kinderen | 0,60 (0,99) | 0,63 (1,30) | 0,63 (0,92) | 0,38 (0,99) | 0,87 (1,15) | 0,65 (1,06) | 0,50 (0,89) |

De vrouwen in dit onderzoek zijn gemiddeld 29,13 jaar oud ($sd = 6,49$). Hierbij zijn de vrouwen met het sparse tri-factional netwerktype ($gem_5 = 30,42$, $sd_5 = 6,09$) gemiddeld 2,85 jaar ouder dan de vrouwen met het strongly tri-factional netwerktype ($gem_3 = 27,57$, $sd_3 = 6,81$), dit is het grootste gemiddelde verschil. De vrouwen met het strongly tri-factional netwerktype zijn significant jonger dan de vrouwen met een weakly bi-factional, weakly tri-factional en een sparse tri-factional netwerktype ($t_2(227) = -2,82$, $p < 0,01$; $t_4(234) = 2,40$, $p = 0,02$; $t_5(197) = -3,15$, $p < 0,01$). De overige netwerktypen verschillen niet significant van elkaar.

Van de Nederlandse vrouwen in deze steekproef heeft 72,8% een partner. Op grond van de chi-kwadraat toets voor onafhankelijke groepen is vast te stellen dat er verschillen zijn in het hebben van een partner tussen de netwerktypen, maar dat deze verschillen niet significant zijn ($X^2(5) = 9,54, p = 0,09$). Als de gestandaardiseerde residuen worden bekeken is te zien dat vrouwen met een sparse tri-factional netwerktype het minst vaak een partner hebben (61,4%, $z = 2,05$) en zijn gemiddeld de meeste partners bij het weakly tri-factional netwerktype gevonden (78,1%, $z = 0,71$).

Van alle vrouwen heeft 35,4% een kind en de verschillen tussen de netwerktypen zijn significant ($X^2(5) = 12,02, p = 0,03$). Waarbij vrouwen met het weakly tri-factional netwerktype gemiddeld de meeste kans hebben om moeder te zijn (45,3%, $z = 1,88$) en vrouwen met het strongly tri-factional netwerktype de minste kans (25,6%, $z = -1,78$).

Op basis van de Kruskal Wallis toets is gebleken dat het aantal kinderen per vrouw significant verschilt per netwerktype ($H(5) = 13,58, p < 0,02$). In deze steekproef hebben de Nederlandse vrouwen tussen de 18 en 40 jaar gemiddeld 0,65 kind, maar de spreiding is groot ($sd = 1,04$). Vrouwen met het weakly tri-factional netwerktype hebben gemiddeld de meeste kinderen ($gem_4 = 0,91, sd_4 = 1,29$) en vrouwen met strongly tri-factional netwerktype de minste kinderen ($gem_3 = 0,44, sd_3 = 0,81$), dit verschil is significant op basis van de Mann-Witney U toets als post-hoc toets met Bonferroni correctie ($p = 0,02$). Als een vrouw kinderen heeft, dan werd haar gevraagd hoeveel van deze kinderen een biologisch kind van haar was. Ook hiervoor geldt dat vrouwen met een weakly tri-factional netwerktype gemiddeld meer biologische kinderen hebben ($gem_4 = 0,87, sd_4 = 1,15$) en vrouwen een strongly tri-factional netwerktype minder ($gem_3 = 0,38, sd_3 = 0,99$). Ook deze verschillen zijn significant ($H(5) = 17,12, p < 0,01$).

Concluderend, de vrouwen met een strongly tri-factional netwerktype zijn gemiddeld het jongst, hebben het minst vaak kinderen en ook het aantal (biologische) kinderen dat zij hebben is het laagst. In vergelijking zijn de vrouwen uit het dense multi-factional netwerktype gemiddeld jonger dan de rest, hebben minder vaak kinderen en ook minder (biologische kinderen). De vrouwen uit beide netwerktypen hebben echter wel gemiddeld even vaak een partner als de andere netwerktypen. De kenmerken van de vrouwen uit netwerktype strongly bi-factional, weakly bi-factional en sparse tri-factional zijn erg vergelijkbaar en ze scoren alles rond het gemiddelde. Het grote verschil tussen deze netwerktypen is dat vrouwen met een strongly bi-factional netwerktype veel vaker een partner hebben dan vrouwen met een sparse tri-factional netwerktype. Als laatste hebben vrouwen uit het weakly tri-factional netwerktype het vaakst een partner, het vaakst kinderen en ook de meeste (biologische) kinderen.

4.3.2 Relatie tussen structurele netwerktypen en fertiliteitswensen

Op basis van de literatuur is hypothese 2 opgesteld, waarin verwacht wordt dat de persoonlijke netwerkstructuur samenhangt met de fertiliteitswensen van een vrouw. De verwachting is dat als een vrouw een meer verbonden persoonlijk netwerk heeft, er ook sprake is van een hoge kinderwens. Om dit te onderzoeken zijn voor kinderwens twee verschillende variabelen gebruikt, namelijk of de vrouw in de toekomst nog kinderen wil en wat zij de ideale gezinsgrootte vindt. Deze worden apart besproken.

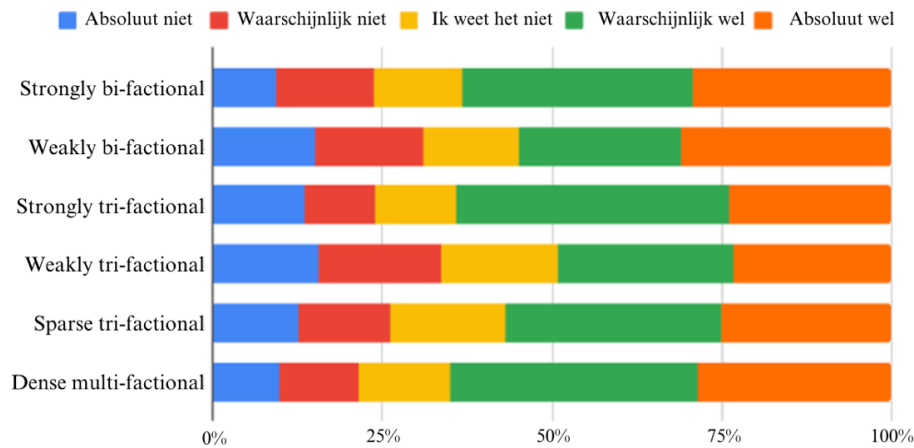
Om hypothese twee te toetsen is het van belang om vast te stellen wat wordt gezien als meer verbonden netwerk binnen de huidige typologie. Allereerst is een meer verbonden netwerk als er minder subgroepen zijn. Daarom wordt een strongly bi-factional netwerk als meer verbonden gezien dan een strongly tri-factional netwerktype. Daarnaast is een hoge verbondenheid ook dat er veel banden zijn tussen de alters van eigen subgroepen, maar er relatief weinig banden zijn tussen de verschillende subgroepen. Oftewel, een strongly bi-factional netwerktype wordt gezien als meer verbonden dan een weakly bi-factional netwerktype. Als laatst wordt een grote hoeveelheid aan isolaten en dyads ook gezien als een minder goed verbonden netwerk.

Concreet wordt verwacht dat vrouwen met het strongly bi-factional netwerktype, het strongly tri-factional netwerktype en het dense multi-factional netwerktype een grotere kinderwens hebben dan vrouwen met een weakly bi-factional netwerktype en het weakly tri-factional netwerktype. Daarnaast wordt verwacht dat de vrouwen met het strongly bi-factional netwerktype een hogere kinderwens hebben dan vrouwen met het strongly tri-factional netwerktype en het dense multi-factional netwerktype.

Relatie tussen structurele netwerktypen en kinderwens

Tabel 7: Verdeling van kinderwens per structureel netwerktype ($n = 706$)

| | Gemiddelde (sd) | Mediaan | Minimaal | Maximaal | 1e kwartiel | 3e kwartiel |
|--------------------------|-----------------|---------|----------|----------|-------------|-------------|
| Kinderwens (alle) | 3,47 (1,35) | 4,00 | 1,00 | 5,00 | 2,00 | 5,00 |
| Strongly bi-factional | 3,59 (1,30) | 4,00 | 1,00 | 5,00 | 2,00 | 5,00 |
| Weakly bi-factional | 3,40 (1,45) | 4,00 | 1,00 | 5,00 | 2,00 | 5,00 |
| Strong tri-factional | 3,50 (1,33) | 4,00 | 1,00 | 5,00 | 3,00 | 4,00 |
| Weakly tri-factional | 3,23 (1,40) | 3,00 | 1,00 | 5,00 | 2,00 | 4,00 |
| Sparse tri-factional | 3,43 (1,34) | 4,00 | 1,00 | 5,00 | 2,00 | 4,00 |
| Dense multi-factional | 3,62 (1,28) | 4,00 | 1,00 | 5,00 | 3,00 | 5,00 |



Figuur 12: Staafdiagram van Kinderwens

De mediaan van kinderwens is 4,00, dit betekent dat de meeste vrouwen hebben aangegeven in de toekomst nog wel kinderen te willen. Van alle vrouwen denkt 59,1% dat ze in de toekomst nog wel een kind zal krijgen. Echter, vrouwen met een weakly tri-factional netwerktype geven in meer dan 50% van de gevallen aan dat ze geen kinderen meer willen of het niet te weten ($mediaan_4 = 3,00$; Figuur 12). Vrouwen met dit netwerktypen hebben ook gemiddeld de hoogste leeftijd en het hoogste percentage kinderen. Daartegenover zijn in de netwerktypen dense multi-factional (65,0%) en strongly tri-factional (64,8%) gemiddeld de meeste vrouwen gevonden die nog een kinderwens hebben. Tegelijkertijd zijn dit ook de twee netwerktypen met gemiddeld de jongste leeftijd en gemiddeld het minst aantal kinderen.

De verwachting was dat de verbonden netwerken hoog zouden scoren op kinderwens. Ondanks dat de verschillen niet groot zijn, scoren het strongly bi-factional, strongly tri-factional en dense multi-factional netwerktype hoger dan het weakly bi-factional, weakly tri-factional en het sparse tri-factional netwerktype. Om te toetsen of deze verschillen significant zijn is gebruik gemaakt van de Kruskal Wallis toets. Uit deze toets blijkt dat er geen significante verschillen bestaan in kinderwens tussen de verschillende netwerktypen ($H(5) = 6,43, p = 0,27$). De effectgrootte laat zien dat in totaal 1% van de variantie die wordt gevonden in de kinderwens verklaard kan worden door de structurele netwerktypen ($\eta^2 = 0,01$).

De verwachting in hypothese 2 was dat een vrouw met een meer verbonden netwerk een hogere kinderwens zou hebben. Aan de hand van deze resultaten kan deze hypothese **niet** worden bevestigd. De resultaten laten geen significant effect zien tussen de verschillende netwerktypen in kinderwens. Het grootste verschil wordt gevonden tussen vrouwen met een weakly tri-factional netwerktype en vrouwen met een dense multi-factional netwerktype (Figuur 12), dit is alleen niet significant op basis van de post hoc toets Wilcoxon Withney U met Bonferroni correctie ($p = 0,31$).

Relatie tussen structurele netwerktypen en ideale gezinsgrootteTabel 8: Verdeling van ideale gezinsgrootte $n = 637$).

| | Gemiddelde (sd) | Mediaan | Minimaal | Maximaal | 1e kwartiel | 3e kwartiel |
|------------------------------------|-----------------|---------|----------|----------|-------------|-------------|
| Ideale Gezinsgrootte (alle) | 2,32 (1,01) | 2,00 | 0,00 | 8,00 | 2,00 | 3,00 |
| Strongly bi-factional | 2,44 (0,92) | 2,00 | 0,00 | 4,00 | 2,00 | 3,00 |
| Weakly bi-factional | 2,34 (1,06) | 2,00 | 0,00 | 5,00 | 2,00 | 3,00 |
| Strong tri-factional | 2,15 (1,00) | 2,00 | 0,00 | 5,00 | 2,00 | 3,00 |
| Weakly tri-factional | 2,35 (1,14) | 2,00 | 0,00 | 8,00 | 2,00 | 3,00 |
| Sparse tri-factional | 2,39 (1,16) | 2,00 | 0,00 | 6,00 | 2,00 | 3,00 |
| Dense multi-factional | 2,29 (0,82) | 2,00 | 0,00 | 4,00 | 2,00 | 3,00 |

De ideale gezinsgrootte bestaat gemiddeld uit 2,32 kinderen (Tabel 8). Waarbij in ieder geval 50% van alle vrouwen aangeeft 2 of 3 kinderen te willen (Tabel 8). Dit is aanzienlijk hoger dan het gemeten gemiddelde van 1,6 daadwerkelijk geboren kinderen per gezin (CBS, 2020).

Vrouwen met een strongly bi-factional netwerktype vinden gemiddeld dat de ideale gezinsgrootte uit 2,44 kinderen bestaat ($sd_I = 0,92$). Dit is het hoogst gemeten gemiddelde. Daarentegen vinden vrouwen met een strongly tri-factional dat de ideale gezinsgrootte uit 2,15 kinderen bestaat. Dit is gemiddeld het laagst. De resultaten zijn niet hoger bij de verbonden netwerktypen in vergelijking met de niet-verbonden netwerktypen. Om te kijken of de verschillen in ideale gezinsgrootte tussen de netwerktypen significant zijn wordt opnieuw de Kruskal Wallis toets gedaan. Deze toets laat zien dat er geen significante verschillen zijn in de ideale gezinsgrootte tussen de verschillende netwerktypen ($H(5) = 5,46, p = 0,36$). Van alle variatie tussen de vrouwen in ideale gezinsgrootte is 1% te verklaren uit de verschillende netwerktypen ($\eta^2 = 0,01$). Het verschil tussen vrouwen met een strongly bi-factional netwerktype en een strongly tri-factional netwerktype is niet groot genoeg om een significant resultaat op de post hoc toets op te leveren ($p = 0,38$).

Op basis van deze resultaten kan de verwachting dat vrouwen met een meer verbonden netwerk een hogere kinderwens hebben **niet** worden bevestigd. De resultaten laten wel zien dat vrouwen met verschillende netwerktypen gelijkgestemd zijn over de ideale gezinsgrootte.

4.4 Toevoeging van compositie kenmerken

Op basis van de derde hypothese wordt verwacht dat door de toevoeging van de compositie van het persoonlijke netwerk extra informatie wordt verkregen, wat resulteert in een sterkere typologie dan enkel gebaseerd op structurele kenmerken. Allereerst zal worden besproken wat de resultaten zijn van de clustering op basis van beide kenmerken. Daarna zullen de resultaten van de samenhang tussen de gevonden netwerktypen en de fertiliteitswensen worden besproken. Als laatste wordt gekeken naar de verdeling van de compositie kenmerken binnen de structurele netwerktypen.

4.4.1 Meerwaarde van compositie kenmerken

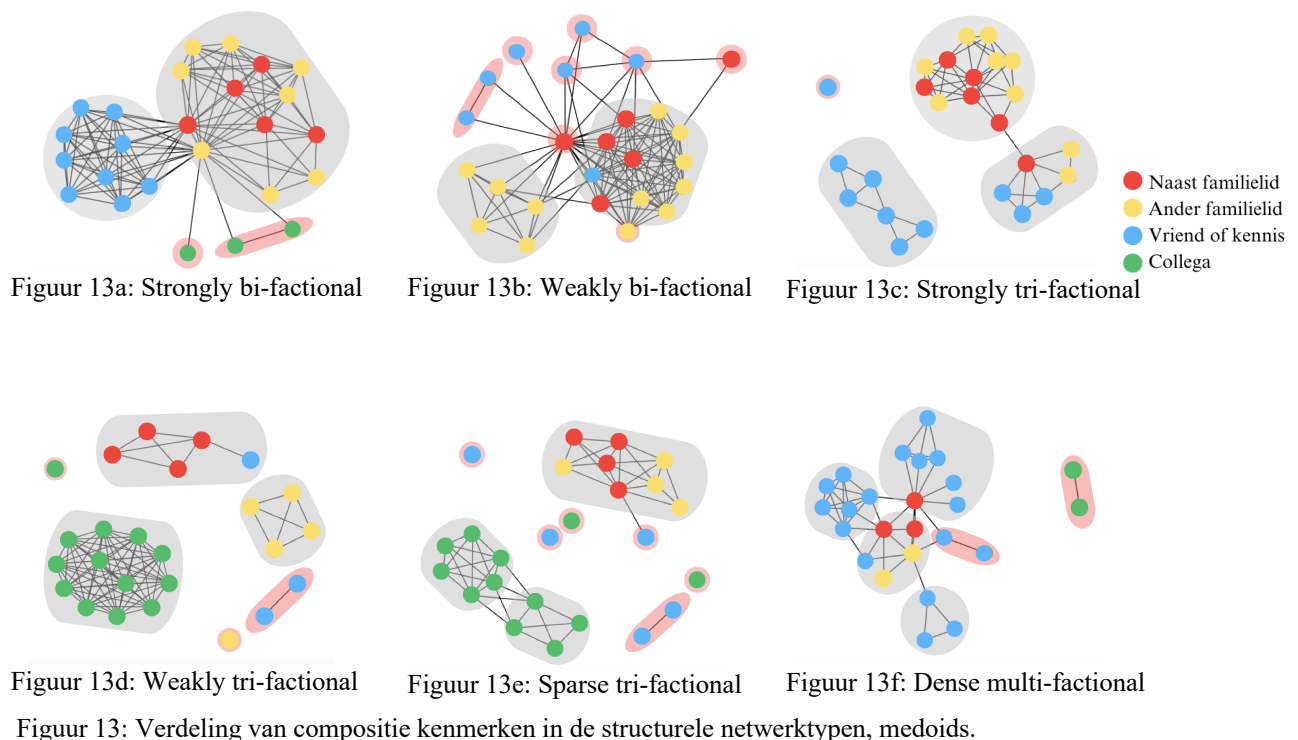
In Bijlage 1 is uitgebreid besproken hoe de clustering op basis van structurele kenmerken en op basis van structurele en compositie kenmerken heeft plaatsgevonden. De meest efficiënte verdeling van de netwerken op basis van beide soorten kenmerken geeft twee netwerktypen.

Om te bepalen of de clustering beter verloopt met inachtneming van de compositie kenmerken zijn de scores van de Silhouette en AIC vergeleken. De Silhouette-score neemt af als de compositie kenmerken worden toegevoegd, op basis van enkel structurele kenmerken heeft de typologie een Silhouette-score van 0,50 ten opzichte van een Silhouette-score van 0,23 wanneer beide kenmerken worden gebruikt. Een hogere score geeft aan dat elk datapunt beter geplaatst kan worden in de juiste clusters. Deze resultaten laten zien dat het model met enkel structurele kenmerken beter in staat is om de netwerken van de vrouwen in de juiste clusters te plaatsen. Daarnaast heeft het model met alleen structurele kenmerken een AIC-score van onder de 500, tegenover een AIC-score van ongeveer 3300 in het tweede model. Deze score is gebaseerd op het idee van maximum-likelihood; hierbij worden de complexiteit van het model en de nauwkeurigheid ervan tegen elkaar afgewogen. Een lagere score geeft aan dat het model beter bij de gegevens past, waardoor ook in dit geval het model met alleen de structurele kenmerken meer geschikt is ($AIC_1 < 500$; $AIC_2 < 3300$).

Om deze reden wordt de derde hypothese die in dit onderzoek is opgesteld **verworpen**. De clustering is beter uit te voeren door alleen de structurele kenmerken mee te nemen. Dit is in overeenstemming met wat Vacca (2020) laat zien in zijn onderzoek. Dit maakt het overbodig om een typologie te maken op zowel structurele als compositie kenmerken en de samenhang met fertiliteitswensen te onderzoeken. Uiteraard is het wel uitgevoerd; de uitwerking is te vinden in Bijlage 2. Hier zijn **geen** significante resultaten gevonden.

4.4.2 Compositie binnen structurele netwerktypen

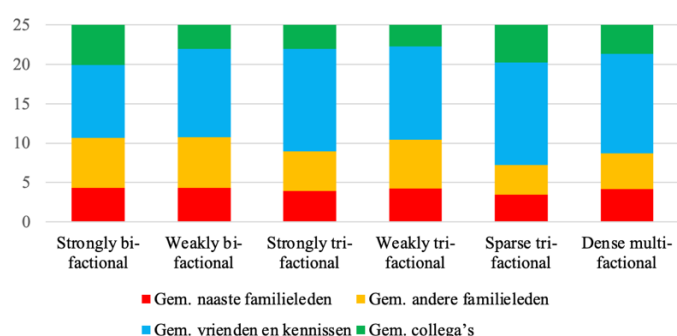
Zoals uit de bovenste resultaten blijkt, verslechtert het model als de compositie van het netwerk wordt toegevoegd aan de clustermethode. Een van de redenen dat dit zou kunnen gebeuren is dat de structurele kenmerken zorgen voor dezelfde verdeling als de compositie kenmerken. Oftewel, de subgroepen hangen samen met de compositie kenmerken. Op die manier wordt het model complexer zonder dat er extra variantie wordt verklaard. Daarom is ervoor gekozen om per structureel netwerktype te kijken naar de compositie van het netwerk.



Om een beeld te schetsen van de verdeling van de compositie kenmerken binnen de structurele netwerktypen zijn de medoids gevisualiseerd, waarbij de grijze vlakken subgroepen tonen en de rode vlakken de isolaten en dyads (Figuur 13). Hieruit blijkt dat de subgroepen voornamelijk bestaan uit één of twee compositie kenmerken. De combinatie tussen naaste familieleden en andere familieleden wordt hierbij vaak gevonden. Deze visualisaties laten zien dat de structurele kenmerken redelijk goed in staat zijn om de informatie te vangen die met het compositie kenmerk over de relatie tot de alter wordt gegeven. Om die reden zal model 2 met drie kenmerken meer aanzienlijk complexer zijn, maar hiermee niet in staat zijn om de netwerken beter in te delen.

Tabel 9: Compositie van de structurele netwerktypen; gemiddelden (standaarddeviaties)

| | Strongly bi-factional | Weakly bi-factional | Strongly tri-factional | Weakly tri-factional | Sparse tri-factional | Dense multi-factional |
|----------------------------|-----------------------|---------------------|------------------------|----------------------|----------------------|-----------------------|
| Gem. naaste familieleden | 4,37 (2,03) | 4,34 (1,98) | 3,95 (1,74) | 4,26 (1,92) | 3,45 (1,70) | 4,16 (1,84) |
| Gem. andere familieleden | 6,29 (5,00) | 6,41 (4,58) | 5,03 (3,31) | 6,21 (4,05) | 3,75 (2,69) | 4,54 (3,62) |
| Gem. vrienden en kennissen | 9,26 (6,07) | 11,19 (4,83) | 12,96 (4,92) | 11,82 (4,52) | 13,02 (4,85) | 12,64 (4,60) |
| Gem. collega's | 5,07 (5,56) | 3,06 (3,32) | 3,04 (3,78) | 2,71 (3,02) | 4,77 (3,86) | 3,66 (3,11) |



Figuur 14: Verdeling compositie van structurele netwerken op basis van gemiddelde aantallen

De resultaten laten zien dat vrouwen met een strongly bi-factional, weakly bi-factional en weakly tri-factional gemiddeld ongeveer evenveel familieleden als vrienden en kennissen in het persoonlijke netwerk hebben (Tabel 9; Figuur 14). Daarentegen hebben de vrouwen in de andere netwerken gemiddeld aanzienlijk meer vrienden en kennissen dan familieleden. Daarbij valt op dat het aantal collega's in elk netwerktype ongeveer gelijk is, maar dat vrouwen met het strongly bi-factional netwerktype meer collega's in hun netwerk hebben. Om de gemiddelde verschillen te toetsen is een ANOVA uitgevoerd.

Uit de ANOVA toets blijkt dat er een significant verschil zit in het gemiddelde aantal naaste familieleden tussen de netwerktypen ($F(5,700) = 3,26, p < 0,01$). Daarom is de post-hoc analyse Tukey uitgevoerd om te kijken welke verschillen significant blijken en daarnaast is de effectgrootte gemeten met Cohen's D. Hieruit blijkt dat de vrouwen met het sparse tri-factional netwerktype significant minder naaste familieleden in hun netwerk dan vrouwen met een strongly bi-factional netwerktype ($T_{5-1} = -0,91, p < 0,01, d = 0,49$), weakly bi-factional ($T_{5-2} = -0,88, p = 0,01, d = 0,47$) en weakly tri-factional ($T_{5-4} = -0,80, p = 0,02, d = 0,43$). De effectgrootte wijst op een medium effect, waardoor geconcludeerd kan worden dat het verschil dat wordt gevonden tussen deze groepen

relevant is. Het grootste gemiddelde verschil is 0,92, wat inhoudt dat ze gemiddeld nog niet 1 persoon van de 25 sociale contacten schelen in naaste familieleden.

Als de twee groepen familieleden samen genomen worden zijn de netwerktypen strongly bi-factional en weakly bi-factional zeer vergelijkbaar. Zij hebben gemiddeld ongeveer 10 - 11 familieleden in hun netwerk, terwijl vrouwen met een sparse tri-factional netwerktype in hun netwerk ongeveer 7 familieleden hebben. Dit is een aanzienlijk verschil.

Op dezelfde manier is naar de andere drie soorten relaties gekeken en ook binnen deze groepen worden significante verschillen gevonden; het aantal andere familieleden per netwerktype ($F(5,700) = 8,31, p < 0,01$), het aantal vrienden en kennissen ($F(5,700) = 9,11, p < 0,01$) en het aantal collega's ($F(5,700) = 7,17, p < 0,01$).

Zowel vrouwen met een sparse tri-factional netwerktype ($gem_5 = 3,75, sd_5 = 2,69$) als het dense multi-factional netwerktype ($gem_6 = 4,54, sd_6 = 3,62$) hebben een laag aantal andere familieleden in hun netwerk. Zij verschillen significant van de drie netwerktypen die ook hoog een hoog aantal naaste familieleden in hun netwerk hadden, namelijk: strongly bi-factional netwerktype ($T_{5-1} = -2,51, p < 0,01, d = 0,62$; $T_{6-1} = -1,75, p < 0,01, d = 0,41$), weakly bi-factional netwerktype ($T_{5-2} = -2,66, p < 0,01, d = 0,69$; $T_{6-2} = -1,82, p < 0,01, d = 0,46$), en het weakly tri-factional netwerktype ($T_{5-4} = -2,46, p < 0,01, d = 0,69$; $T_{6-4} = -1,66, p < 0,01, d = 0,43$). Deze effectgroottes wijzen op een medium tot sterk effect, waardoor de verschillen in aantal andere familieleden aanzienlijk is. Hierbij is het grootste gemiddelde verschil gevonden tussen de vrouwen met het sparse tri-factional netwerktype en de vrouwen met het weakly bi-factional netwerktype, namelijk 2,66. Dit een redelijk verschil, helemaal wanneer wordt meegenomen dat deze netwerktypen ook ongeveer één persoon verschillen in naaste familieleden. Oftewel, de vrouwen met de netwerktypen strongly bi-factional en weakly bi-factional hebben gemiddeld meer familieleden in hun netwerk dan dat vrouwen met een sparse tri-factional netwerktype hebben.

Aangezien alle vrouwen in totaal 25 alters opgeven, is het voorspelbaar dat de verschillen die bij de eerste twee categorieën zijn gevonden ook terug te zien zijn in de laatste twee categorieën. Zo hebben vrouwen met een met een sparse tri-factional netwerktype gemiddeld veel vrienden en kennissen in hun netwerk ($gem_5 = 13,02, sd_5 = 4,85$). In verhouding bestaat het netwerk van vrouwen met een strongly bi-factional netwerktype ($gem_1 = 9,26, sd_1 = 6,07$) en een weakly bi-factional netwerktype ($gem_2 = 11,19, sd_2 = 4,83$) gemiddeld uit minder vrienden en kennissen. Het verschil tussen sparse tri-factional en strongly bi-factional is aanzienlijk; het gemiddelde verschil is namelijk ongeveer 4 personen ($T_{5-1} = 3,76, p < 0,01, d = 0,68$). Vrouwen met een strongly bi-factional netwerktype hebben significant minder vrienden en kennissen in hun netwerk zitten dan elk ander netwerktype ($p < 0,01 - p = 0,05$). Alle andere netwerktypen laten geen significante verschillen zien.

Als laatste wordt een vergelijking gemaakt tussen het aantal collega's in het netwerk van de vrouwen. Vrouwen met een strongly bi-factional netwerktype ($gem_1 = 5,07$, $sd_1 = 5,56$) en vrouwen met een sparse tri-factional netwerktype ($gem_5 = 4,77$, $sd_5 = 3,86$) laten de hoogste gemiddelden zien. Vrouwen met dit netwerktype hebben gemiddeld ongeveer 2 collega's meer in hun netwerk als de andere netwerktypen. Deze verschillen zijn significant met het netwerk weakly bi-factional ($T_{1-2} = -1,40$, $p < 0,01$, $d = 0,44$; $T_{5-2} = 1,71$, $p = 0,02$, $d = 0,48$), strongly tri-factional ($T_{1-3} = -2,01$, $p < 0,01$, $d = 0,43$; $T_{5-3} = 1,71$, $p = 0,02$, $d = 0,45$), weakly tri-factional ($T_{1-4} = -2,36$, $p < 0,01$, $d = 0,54$; $T_{5-4} = 2,06$, $p < 0,01$, $d = 0,32$) en dense multi-factional ($T_{1-6} = -1,40$, $p = 0,04$, $d = 0,33$).

De spreiding in het aantal sociale relaties binnen een bepaalde categorie is voor elk netwerktype hoog. In veel gevallen is de standaarddeviatie vrijwel even groot als het gemiddelde. Dit wijst op veel spreiding binnen een netwerktype. Oftewel, vrouwen binnen een structureel netwerktype verschillen erg op het compositie kenmerk. Hierdoor moet voorzichtig worden omgegaan met de conclusies die hieruit worden getrokken.

5. Conclusie

In dit hoofdstuk zal de onderzoeksvraag worden beantwoord aan de hand van de opgestelde deelvragen en hypothesen. De onderzoeksvraag die centraal stond in dit onderzoek was als volgt: *“In hoeverre kunnen persoonlijke netwerken van vrouwen worden opgedeeld in zinvolle netwerktypen? En hoe hangen deze netwerktypen samen met de fertiliteitswensen van een vrouw?”*

Om antwoord te geven op deze vraag is gebruikgemaakt van de op subgroep gebaseerde clustermethode van Vacca (2020). Allereerst heeft de identificatie van subgroepen plaatsgevonden aan de hand van het Girvan-Newman-algoritme (Girvan & Newman, 2002). Daarna is de subgroepstructuur samengevat aan de hand van drie structurele kenmerken, namelijk de modulariteitscore, het aantal subgroepen met drie of meer alters en het aantal isolaten en dyads. Op basis van deze drie kenmerken is de k-medoid clusteranalyse uitgevoerd. Om het verschil in fertiliteitswensen tussen de netwerktypen te onderzoeken is de Kruskal Wallis toets gedaan. Daarnaast is opnieuw een typologie gemaakt, waarbij ook compositie kenmerken zijn gebruikt. De compositie kenmerken waren het aantal naaste familieleden, het aantal andere familieleden, het aantal vrienden en kennissen en het aantal collega's.

5.1 Structurele netwerktypen Nederlandse vrouwen

In het onderzoek van Vacca (2020) kwamen zes structurele netwerktypen het vaakst voor, namelijk: closed community, dense core-periphery, sparse core-periphery, weakly bi-factional, strongly bi-factional en dense multi-factional. De eerste deelvraag ging over welke netwerktypen geïdentificeerd kunnen worden bij Nederlandse vrouwen; de eerste hypothese stelde dan ook dat de structurele netwerktypen overeen zouden komen met de veelvoorkomende netwerktypen van Vacca (2020). Uit de op subgroep gebaseerde clustermethode kwam naar voren dat de meest geschikte typologie bestaat uit zes netwerktypen. Dit is gebaseerd op twee testscores, namelijk de AIC-score (<500) en de Silhouette-score (0,50). De lage AIC-score geeft aan dat het model goed past bij de data; de complexiteit van het model is evenredig aan de nauwkeurigheid ervan. Oftewel, het model is in staat om de variantie in de gegevens goed te verklaren. Hierdoor kunnen de netwerken op basis van deze kenmerken goed ingedeeld worden in verschillende clusters. Daarbij geeft de Silhouette-score van 0,50 aan dat de netwerktypen redelijk goed te scheiden zijn. Anders verwoord, de netwerken binnen een cluster lijken meer op elkaar dan netwerken tussen de clusters.

Een vergelijking van de gevonden netwerktypen en de meest voorkomende netwerktypen van Vacca (2020) laat zien dat de structurele netwerktypen redelijk overeenkomen. Echter, de structurele netwerktypen closed community, dense core-periphery, sparse core-periphery, zoals gedefinieerd

door Vacca (2020), zijn alle drie gebaseerd op de voorwaarde dat er maar één subgroep is met drie of meer alters. In de steekproef van het huidige onderzoek zijn in totaal 32 netwerken aanwezig die maar één subgroep met drie of meer alters hebben. Dit is een redelijk aantal, maar deze verschillen blijkbaar niet voldoende van andere netwerken met meerdere subgroepen om een apart netwerktype te vormen. Hierdoor zijn drie van de zes meest voorkomende netwerktypen uit het onderzoek van Vacca (2020) niet geïdentificeerd als een apart netwerktype in dit onderzoek. Daarentegen zijn wel beide bi-factional netwerktypen geïdentificeerd, het eerste netwerktype in de typologie van deze studie is daarom strongly bi-factional en de tweede weakly bi-factional. Daarbij worden in de huidige studie vier netwerktypen geïdentificeerd met drie subgroepen of meer. Dit is bij Vacca (2020) niet voorgekomen en om die reden was het nodig om een verder onderscheid te maken in de meervoudige partities. Het tri-factional netwerktype van Vacca (2020) is dan ook zelf ingedeeld, resulterend in strongly tri-factional, weakly tri-factional en sparse tri-factional. Het laatste netwerktype dat wordt geïdentificeerd is het dense multi-factional netwerktype. Vergelijkend met de meest voorkomende netwerktypen van Vacca (2020) zijn twee netwerktypen gelijk, namelijk het strongly bi-factional en weakly bi-factional netwerktype. Daarbij is het sparse tri-factional netwerktype vergelijkbaar met het sparse multi-factional netwerktype van Vacca (2020), maar door de onderverdeling van de tri-factional netwerktypen is de naam veranderd. Naast de netwerktypen die gebaseerd zijn op één subgroep zijn de netwerktypen in het huidige onderzoek **redelijk vergelijkbaar** met de netwerktypen die gevonden zijn in het onderzoek van Vacca (2020).

De gevonden netwerktypen laten zien dat de Nederlandse vrouwen vaak meerdere subgroepen hebben die een centrale rol spelen in hun leven. Een subgroep vertegenwoordigt een groep mensen die elkaar kennen en met elkaar interacteren, een sociale context. Een verscheidenheid aan subgroepen laat zien dat de vrouw zich bevindt in verschillende sociale contexten.

5.2 Samenhang tussen structurele netwerktypen en fertiliteitswensen

In dit onderzoek is ook onderzocht in hoeverre de verschillende structurele netwerktypen samenhangen met de fertiliteitswensen van een vrouw. Binnen het fertiliteitsonderzoek is de invloed van de structuur van het persoonlijke netwerk relatief nieuw. Toch is wel gebleken dat structurele kenmerken van het persoonlijke netwerk kunnen bijdragen aan onderwerpen rondom fertiliteit, zoals het praten over het krijgen van kinderen (Stulp & Barrett, 2021) en het gebruik van anticonceptiemiddelen (Kohler et al., 2021).

Op basis van de theorie over sociaal kapitaal werd verwacht dat een verbonden persoonlijk netwerk leidt tot een hogere kinderwens. De verwachting was dat vrouwen met sterk verbonden subgroepen in hun persoonlijke netwerk kunnen rekenen op een hoge mate van sociale steun en dat de mechanismen van sociale besmetting en sociale controle mogelijk effectiever zijn. Daarbij was ook de

verwachting deze mechanismen sterker zijn wanneer er sprake was van een lager aantal subgroepen. De geïdentificeerde netwerktypen die werden gekenmerkt door sterk verbonden subgroepen waren het strongly bi-factional netwerk, het strongly tri-factional netwerk en het dense multi-factional netwerk. Daarentegen hebben de netwerktypen weakly bi-factional netwerktype, weakly tri-factional en sparse tri-factional netwerktype ook subgroepen, maar de samenhang tussen deze subgroepen is groter en de onderlinge verbondenheid kleiner. Om die reden werd verwacht dat de mechanismen van sociale besmetting, sociale controle en sociale steun minder sterk zouden zijn en hierdoor de kinderwens van de vrouw minder zouden beïnvloeden.

De resultaten toonde aan dat de indeling van persoonlijke netwerken van Nederlandse vrouwen op basis van subgroepen niet kan verklaren wat voor kinderwens zij hebben en wat hun ideale gezinsgrootte is. De trend waarin vrouwen steeds minder kinderen krijgen in Nederland valt daarom niet te verklaren aan de hand van de structuur van het persoonlijke netwerk van de vrouw. Bijna alle vrouwen geven aan waarschijnlijk in de toekomst (nog) wel een kind te willen. Een uitzondering hierop zijn de vrouwen met het weakly tri-factional netwerktype, hier geeft het merendeel aan dat zij geen kinderwens meer hebben of niet weten of ze in de toekomst nog een kind willen. Dit zijn echter ook de vrouwen die gemiddeld het oudst zijn en het vaakst al kinderen hebben. Hier lijkt dan ook sprake te zijn van cofounder variabelen; de leeftijd en het wel of niet hebben van kinderen beïnvloedt mogelijk zowel de netwerkstructuur als de kinderwens van de vrouw. Daarnaast wordt het vaakst aangegeven dat de ideale gezinsgrootte voor de Nederlandse vrouw uit twee kinderen bestaat. De gemiddeld jonge vrouwen uit het strongly tri-factional netwerktype geven aan gemiddeld de minste kinderen te willen. Verwacht werd dat vrouwen met dit netwerktype hoog zouden scoren, aangezien het ten opzichte van de andere netwerktypen relatief goed verbonden is. Er zijn geen aanwijzingen dat de indeling op basis van subgroepen leidt tot een beter begrip van de kinderwens van Nederlandse vrouwen.

5.3 Meerwaarde compositie kenmerken

Om te kijken naar de meerwaarde van compositie kenmerken zijn drie deelvragen opgesteld. In deze deelvragen ging het om het identificeren van netwerktypen gebaseerd op structurele en compositie kenmerken, de relevantie hiervan voor de identificatie van netwerktypen en de gevonden samenhang met fertiliteitswensen. Verwacht werd dat de toevoeging van de compositie van het persoonlijke netwerk een betere typologie zou creëren gebaseerd op diverse literatuur. De literatuur laat zien dat verschillende relaties invloed hebben op de fertiliteitskeuzes en -wensen, namelijk ouders (Bühler & Philipov, 2005; Kavas & Jong, 2020; Keim et al., 2013), broers en zussen (Balbo & Mills, 2011; Buyukkececi et al., 2020; Kotte & Ludwig, 2011, Lyngstad & Prskawetz, 2010), vrienden (Balbo & Barban, 2014; Kotte & Ludwig, 2011) en collega's (Buyukkiececi et al., 2020; Ciliberto et al., 2016;

Pink et al., 2014). Deze literatuur ondersteunt de uitspraken dat sociale relaties invloedrijk zijn betreffende fertiliteit (Bongaarts & Watkins, 1996; Granovetter, 1985; Montgomery & Caterline, 1996). Deze onderzoeken hebben allemaal de focus gelegd op één soort sociale relatie, verwacht werd dat het maken van onderscheid tussen deze sociale relaties mogelijk meer inzicht zou brengen in welke relaties meer invloedrijk zijn en ook sociale context zou brengen.

Na de toevoeging van de compositie kenmerken aan de analyse bleek het model te verslechteren ten opzichte van het model met alleen de structurele kenmerken (AIC 1 < 500; AIC 2 = 3300; Silhouette1 = 0,50 en Silhouette2 = 0,23). Deze scores laten zien dat het model complexer is geworden, maar het heeft niet een verbeterde indeling kunnen maken in verschillende clusters. Dit laat zien dat de op subgroep gebaseerde methode van Vacca (2020) niet aangevuld hoeft te worden met informatie over de samenstelling van de sociale relaties. Voor de volledigheid is een nieuwe typologie gemaakt en de samenhang met fertiliteitswensen gemaakt. Deze typologie bestaat uit twee netwerktypen, deze vinden de hoogste overeenkomst met het strongly tri-factional netwerktype en het weakly bi-factional netwerktype. In deze clusters was de verscheidenheid aan sociale relaties hoog. Dit houdt in dat de compositiematen zich niet makkelijk laten reduceren tot een bepaald clusters. Lubbers en collega's (2007) wisten bijvoorbeeld een *dense family network* te onderscheiden. Doordat de samenstelling in sociale relaties tussen deze twee clusters niet dermate verschilde was dit niet mogelijk in deze studie.

Ook het onderscheid in deze twee netwerktypen laat geen verschil in kinderwens van de vrouw zien. De verwachting uit hypothese 3 was dat de samenhang tussen de netwerktypen op basis van structurele en compositie kenmerken sterker werd met de kinderwens van de vrouw, dit kan **niet** worden bevestigd.

Een van de redenen dat het model verslechtert wanneer de compositie kenmerken worden meegenomen is mogelijk dat de op subgroep gebaseerde methode de verschillende sociale kringen al weet te vinden. De subgroepen in het persoonlijke netwerk leggen namelijk groepen vast die sterk met elkaar verbonden zijn. Tegelijkertijd is het waarschijnlijk dat bijvoorbeeld de naaste familieleden verbonden zijn in één groep evenals de collega's. De verdeling in sociale relaties geeft hierdoor mogelijk vergelijkbare informatie, waardoor het model wel complexer wordt maar dat het er niet in slaagt de netwerken beter in te delen. De resultaten suggereren dat de structurele kenmerken van het netwerk een deel van de compositie vangt. Dit is vergelijkbaar met de uitkomst dat structurele kenmerken samenhangen met de sociale context en compositie van het netwerk (Bidart et al., 2018; Gianella & Fisher, 2016; Maya-Jariego, 2021). Echter, hiervoor zal verder onderzoek moeten worden gedaan.

Concluderend, er zijn afwijkende patronen te ontdekken in de persoonlijke netwerken van vrouwen gebaseerd op structurele kenmerken. Aan de hand van de op de subgroep gebaseerde clustermethode van Vacca (2020) zijn zes structurele netwerktypen geïdentificeerd voor de netwerktypologie. Deze netwerktypen kwamen redelijk overeen met de veelvoorkomende netwerktypen uit het onderzoek van Vacca (2020). Daarbij was er geen samenhang te ontdekken tussen de structuur van het persoonlijke netwerk van een vrouw en haar kindwens. Ten derde maakt de toevoeging van het compositie kenmerk dat aangeeft welke sociale relaties de vrouw heeft het model voor de clustering niet beter.

6. Discussie

Dit onderzoek had twee doelstellingen; enerzijds het systematisch ordenen van de persoonlijke netwerkdata van Nederlandse vrouwen en anderzijds bijdragen aan de kennis over de invloed van sociale relaties op de kinderwens van vrouwen. In dit hoofdstuk wordt gekeken in hoeverre deze doelstellingen zijn behaald. Dit wordt gedaan aan de hand van drie subparagrafen, waarin allereerst de resultaten worden geïnterpreteerd en worden vergeleken met voorgaand onderzoek. Daarnaast worden de methodologische beperkingen van dit onderzoek besproken. Als laatste zal de wetenschappelijke bijdrage van dit onderzoek aan bod komen en mogelijk vervolgonderzoek worden gepresenteerd.

6.1 Interpretatie van resultaten

De resultaten lieten zien dat de gevonden netwerktypen redelijk vergelijkbaar waren met de veelvoorkomende netwerktypen van Vacca (2020), maar een netwerktype gebaseerd op één subgroep komt in het huidige onderzoek niet voor. Dit is interessant aangezien in Vacca (2020) drie veelvoorkomende netwerktypen gebaseerd zijn op één subgroep. Dit verschil is mogelijk te verklaren door het verschil in context bij de gekozen doelgroepen. In de studie van Vacca (2020) wordt het niet volledig duidelijk welke datasets zijn gebruikt, wel is bekend dat er een variatie is in relatiedata (tussen de 5 en 45 alters) en wat de studiecontext is. In de studies zijn tussen 30 en 45 alters meegenomen, maar bij sommige studies is een random sample getrokken voor alter-alter relaties. De grootte van het aantal alters is hoger dan in deze studie, maar in de alter-alter ties zijn in sommige studies kleiner. Daarnaast is de studiecontext ook zeer verschillend, namelijk in twee dataset gaat het over migratie ($n = 385$, $n = 119$), twee gingen over rampen ($n = 293$, $n = 264$) en twee over gezondheid (levensloop ($n = 366$) en raciale verschillen ($n = 254$)). Vooralsnog zijn er drie andere onderzoeken die gebruik maken van de methode van Vacca (2020). In het eerste onderzoek dat is gebaseerd op de op subgroep gebaseerde clustermethode worden vijf netwerktypen gevonden van Roemeense Roma-Migranten in Frankrijk ($n = 119$), deze komen alle vijf voor in de veelvoorkomende netwerktypen en alle drie de netwerktypen op basis van één subgroep zijn aanwezig (Vacca et al., 2021). Het tweede onderzoek identificeert ook vijf netwerktypen van Latino/a-adolescenten op het platteland ($n = 30$) waarvan twee, namelijk fragmented en hybrid, niet voorkomen in de meest voorkomende netwerktypen (Fanfan et al., 2021). Wel worden hier zowel een closed community netwerktype als een core-periphery netwerktype gevonden. Als laatste identificeren Maya-Jariego en Holgado-Ramos (2022) drie netwerktypen van laatstejaars middelbare school scholieren in de provincie Sevilla ($n = 69$). In deze netwerktypen was er geen netwerktype dat maar één subgroep vond, elk netwerktype was bi-factional of tri-factional. Dit is een vergelijkbaar resultaat met het huidige onderzoek. Een vergelijking van deze studies geeft aan dat de context en doelgroep

van deze onderzoeken sterk verschillen. Zoals terugkwam in de inleiding is Nederland een geavanceerde samenleving en daarbij is er sprake van een individualistische cultuur. (SCP, 2019) Daarentegen zijn de andere studies uitgevoerd in een collectivistischere nationale context, waardoor hier waarschijnlijk ook vaker één centrale subgroep (Vacca, 2020). De nationale context lijkt dan ook te herleiden te zijn uit het netwerk van de Nederlandse vrouwen uit deze steekproef. De resultaten laten zien dat de meeste Nederlandse vrouwen meerdere subgroepen hebben in hun netwerk. Daarbij worden in elk structureel netwerktype gemiddeld meer vrienden en kennissen gevonden dan familieleden.

De resultaten laten ook zien dat de kinderwens van een vrouw niet samenhangt met de op subgroep gebaseerde structurele kenmerken van haar persoonlijke netwerk. Er werden wel wat verschillen in kinderwens gevonden, maar deze waren niet significant en de effectgrootte van de indeling was ook zeer klein. Daarbij waren de kenmerken van de netwerktypen met een hogere kinderwens ook dat ze gemiddeld het jongst waren en de gemiddeld de minste familieleden in hun netwerk hadden. Oftewel, de structurele kenmerken van een netwerk lijken wel degelijk de sociale context van het individu te vangen (Vacca, 2020).

Visualisaties van de structurele netwerktypen met de verdeling van de sociale relaties laten zien dat een subgroep vaak één of twee soorten sociale relaties bevatten. Ook had elk netwerktype een andere verdeling in het aantal naaste familieleden, andere familieleden, vrienden en kennissen en collega's. De visualisatie van het netwerk van de medoid van het strongly bi-factional netwerktype laat bijvoorbeeld een goed verbonden familie subgroep zien en een goed verbonden vrienden en kennissen subgroep. Een groter aantal familieleden en een groter aantal vrienden en kennissen kan al wijzen op twee subgroepen, terwijl een verdeling in drie groepen, namelijk ook met collega's wijst meer op een tri-factional netwerktype. Dit is goed te zien bij de visualisatie van het strongly tri-factional netwerktype (Figuur 13c). In die visualisatie van het persoonlijke netwerk van de medoid is een subgroep te zien van met haar familieleden, haar partner en zijn familie en vrienden, en haar eigen vrienden. De sociale kringen worden gevangen door de op subgroep gebaseerde methode van Vacca (2020), waardoor de informatie over de sociale relaties weinig weet toe te voegen. De resultaten spraken dan ook tegen de derde hypothese, waarin werd verwacht dat de toevoeging van de compositie van het persoonlijke kenmerk een betere typologie kan creëren dan enkel op basis van structurele kenmerken mogelijk is. Hierbij had meer in overweging genomen moeten worden dat gemeenschap op basis van structurele kenmerken mogelijk vergelijkbaar is met het compositie kenmerk over de relaties.

6.2 Methodologische beperkingen

Er zijn een aantal methodologische beperkingen waar op teruggekomen moet te worden. Allereerst de verdeling van kinderwens en ideale gezinsgrootte. Zoals in de inleiding al terugkomt is de kinderwens in Nederland nog altijd hoog. Uit de verdeling van de variabele *kinderwens* blijkt dan ook dat de Nederlandse vrouwen veel gelijkenissen vertonen. In totaal geeft maar 26,5% van alle vrouwen aan geen kinderwens meer te hebben. Het is hierbij mogelijk dat deze vrouwen wel een kinderwens hadden, maar dat deze al in vervulling is gegaan. Dit lijkt zo het geval te zijn bij vrouwen met een weakly tri-factional netwerktype, zij zijn gemiddeld het oudste, hebben de meeste kinderen en hebben daarnaast ook de laagste kinderwens. Hetzelfde geldt, in mindere mate, voor de verdeling van ideale gezinsgrootte. Zo geeft namelijk 45,8% aan dat de ideale gezinsgrootte uit 2 kinderen bestaat en 24,5% uit 3 kinderen. De variantie is daardoor zeer klein, waardoor mogelijk ook de verschillen per netwerktype niet zichtbaar zijn.

Als de resultaten van de huidige silhouette-score worden vergeleken met de criteria die Kaufman & Rousseeuw (1990) hebben gesteld, dan is de structuur die wordt gevonden binnen de clusters niet sterk genoeg. Oftewel, de netwerken van verschillende clusters hebben overeenkomsten. Kaufman & Rousseeuw (1990) geven vier categorieën, waarbij de scheidingslijn tussen categorie twee (zwakke structuur in data) en categorie drie (redelijke structuur in data) een silhouette-score van 0,50 is; vanaf een silhouette-score van 0,51 wordt gesproken over een redelijke structuur. Het gaat om subjectieve criteria en de gevonden silhouette-score ligt precies op de scheidingslijn, waardoor er alsnog voor is gekozen om de huidige typologie te behouden en geen nieuwe methode te gebruiken. Echter, dit resultaat laat wel zien dat de structuren van netwerken binnen een cluster kunnen verschillen van het meest representatieve punt, de medoid. Daarnaast is het ook mogelijk dat dit leidt tot minder verschillen in kinderwens of ideale gezinsgrootte. Uit het artikel van Vacca (2020) is niet op te maken welke silhouette-score hij heeft gekregen in de zes datasets, deze kunnen daarom ook niet worden vergeleken. Hierdoor is het lastig vast te stellen wat de oorzaak is van de lage silhouette-score; mogelijk levert de grote hoeveelheid aan netwerkdata een minder goed resultaat.

Een derde methodologische beperking in het onderzoek is het onderzoeken van invloedsprocessen met cross-sectionele data in een persoonlijke netwerkstudie. Allereerst is dit problematisch aangezien in persoonlijke netwerkstudie niet alle potentiële alters in kaart worden gebracht, enkel een selectieve groep (Steglich et al., 2010). Voor het vastleggen van een invloedsproces is het namelijk noodzakelijk om het gehele netwerk onderzoeken. Daarnaast is er in dit onderzoek sprake van cross-sectionele data. Een verandering ten opzichte van een ander tijdsmoment is hierdoor niet aanwezig en daarmee kan geen verandering in de data plaatsvinden. Hierdoor is het niet mogelijk om invloed vast te leggen (Steglich et al., 2010). De resultaten van het huidige onderzoek laten geen samenhang zien tussen de persoonlijke netwerktypen en de kinderwens van een

vrouw. Echter, dit resultaat is mogelijk niet gevonden doordat de invloedsprocessen niet op één tijdstip gemeten kunnen worden. Daarnaast had een mogelijk resultaat ook kunnen wijzen op een selectieproces. Dit had niet kunnen worden uitgesloten op basis van de huidige data.

Een laatste beperking van dit onderzoek is dat de indeling in netwerktypen niet aan grenswaarden gebonden is. De methode gepresenteerd door Vacca (2020) geeft richtlijnen voor het categoriseren van de netwerktypen, deze zijn gepresenteerd in Tabel 2. Hij presenteert echter geen grenswaarden voor modulariteitscores en aantal isolaten en dyads van bepaalde netwerktypen. In dit onderzoek was het daarom lastig te bepalen of Vacca (2020) bepaalde netwerktypen geïnterpreteerd zou hebben als ‘tri-factional’ of ‘multi-factional’. Doordat hij een onderscheid maakt in tri-factional en multi-factional is er in dit onderzoek voor gekozen om tri-factional aan te nemen bij drie groepen en multi-factional bij vier groepen. Hierdoor komt bijvoorbeeld het sparse tri-factional netwerktype goed overeen met het sparse multi-factional netwerktype van Vacca (2020). Geconcludeerd kan worden dat een deel van de categorisering toe komt aan de interpretatie van de onderzoeker. Dit levert mogelijke inconsistenties op tussen onderzoeken die dezelfde methode gebruiken. Het resultaat hiervan is enerzijds dat vergelijkbare netwerktypen niet als zodanig worden gekenmerkt en anderzijds dat er mogelijk cruciale verschillen zitten tussen dezelfde netwerktypen in andere onderzoeken.

6.3 Wetenschappelijke bijdrage en vervolgonderzoek

De twee doelstellingen voor het onderzoek waren het systematisch ordenen van de persoonlijke netwerkdata van Nederlandse vrouwen en de samenhang tussen de netwerktypen en fertiliteitswensen onderzoeken. De resultaten van de clustermethode op basis van structurele kenmerken geven een aantal opvallende resultaten, namelijk enerzijds zijn er weinig vrouwen met maar één subgroep in hun netwerk en anderzijds worden er veel vrouwen gevonden die veel verschillende subgroepen hebben. Dit verschilt met de typologieën die zijn opgesteld door Vacca (2020); hier werden verschillende netwerktypen geïdentificeerd met één subgroep en aanzienlijk minder met verschillende subgroepen. Het zou interessant zijn om verder te onderzoeken of netwerktypen binnen verschillende contexten van elkaar verschillen en welke context relevant is voor de verdeling in subgroepen, bijvoorbeeld het eerder aangedragen onderscheid van individualistische en collectivistische cultuur.

Ten tweede is er geen samenhang gevonden tussen de structurele netwerktypen en de fertiliteitswensen van de Nederlandse vrouw. De resultaten wijzen erop dat Nederlandse vrouwen zeer vergelijkbare kinderwensen hebben. De jongere vrouwen met gemiddeld minder kinderen uit netwerktype dense multi-factional en strongly tri-factional lieten een hogere kinderwens zien, tegenover de vrouwen met het weakly tri-factional netwerktype die gemiddeld de oudste waren en gemiddeld al de meeste kinderen hadden. Deze resultaten laten zien dat de structurele kenmerken in staat zijn om een deel van de sociale context en individuele kenmerken te vangen en verschillen weet

te vinden in bijvoorbeeld leeftijd en aantal kinderen. Echter, deze verschillen worden niet gevonden bij de fertiliteitswensen. Andere onderzoeken die zich richten op fertiliteit gebruiken vaak de variabele ‘leeftijd waarop iemand eerste kind kreeg’. Voor toekomstig onderzoek zou gekeken kunnen worden naar het verschil in resultaten tussen kinderwens en fertiliteitskeuzes.

Als laatste wist de context van sociale relaties het model met alleen structurele kenmerken niet te verbeteren. Het model werd complexer, er waren namelijk zes variabelen nodig in plaats van drie. Daarbij wisten deze variabelen niet de nauwkeurigheid van de clustering te verbeteren. Om die reden is er gekeken naar de overlap tussen de informatie uit de subgroepen en de compositie kenmerken gericht op sociale relaties. Hieruit komt naar voren dat informatie over de subgroepen al een groot onderscheid maakt tussen de verschillende sociale relaties. Hierdoor is het te verklaren dat het model slechter werd. Dit betekent echter niet dat alle compositie kenmerken niet relevant zijn in combinatie met de structurele kenmerken, maar de compositie kenmerken gericht op de relatie tot de alter lijken goed gevangen te worden door de subgroep partitie. Vervolgonderzoek zal moeten uitwijzen welke compositie kenmerken relevant zijn om toe te voegen aan de structurele kenmerken.

Referenties

- Akaike, H. (1974). A new look at the statistical model identification. *Ieee Transactions on Automatic Control*, 19(6). <https://doi.org/10.1109/TAC.1974.1100705>.
- Balbo, N., & Barban, N. (2014). Does fertility behavior spread among friends? *American Sociological Review*, 79(3), 412–431. <https://doi.org/10.1177/0003122414531596>.
- Balbo, N., Billari, F. C., & Mills, M. (2013). Fertility in advanced societies: a review of research. *European Journal of Population*, 29(1), 1–38. <https://doi.org/10.1007/s10680-012-9277-y>.
- Balbo, N., & Mills, M. (2011). The influence of the family network on the realisation of fertility intentions. *Vienna Yearbook of Population Research*, 9, 179–205. <https://doi.org/10.1553/populationyearbook2011s179>.
- Bernardi, L., & Klärner A. (2014). Social networks and fertility. *Demographic Research*, 30, 641–670. <https://doi.org/10.4054/DemRes.2014.30.22>.
- Bidart, C., Degenne, A., & Grossetti, M. (2018). Personal networks typologies: a structural approach. *Social Networks*, 54, 1–11. <https://doi.org/10.1016/j.socnet.2017.11.003>.
- Bongaarts, J., & Watkins, S. C. (1996). Social interactions and contemporary fertility transitions. *Population and Development Review*, 22(4), 639–682. <https://doi.org/10.2307/2137804>.
- Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895. <https://doi.org/10.1126/science.1165821>.
- Bourdieu, P. (1986). The Forms of Capital. In J.G. Richardson (ed.) *Handbook of Theory and Research for the Sociology of Education*. (pp. 241-258). New York: Greenwood Press.
- Bühler C., & Philipov, D. (2005). Social capital related to fertility: Theoretical foundations and empirical evidence from Bulgaria. *Vienna Yearbook of Population Research*, 3(1), 53–81. <https://doi.org/10.1553/populationyearbook2005s53>.
- Burt, R.S. (1992). *Structural holes: The social structure of competition*. Harvard University Press.
- Burt, R. S. (1998). The gender of social capital. *Rationality and Society*, 10(1), 5–46. <https://doi.org/10.1177/104346398010001001>.
- Buyukkececi, Z., Leopold, T., van Gaalen, R., & Engelhardt, H. (2020). Family, firms, and fertility: a study of social interaction effects. *Demography*, 57(1), 243–266. <https://doi.org/10.1007/s13524-019-00841-y>.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 94-120. <https://doi.org/10.1086/228943>.
- Centraal Bureau voor de Statistiek. (2022). *Geboorte; kerncijfers*. Geraadpleegd op 1 maart 2023 van <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37422ned/table>.
- Centraal Bureau voor de Statistiek. (2021). *Kinderen krijgen*. Geraadpleegd op 16 februari 2023 van <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/levensloop/kinderen-krijge>.

- Centraal Bureau voor de Statistiek. (2020). *Kerncijfers geboorte; waarneming en prognose, geboortegeneratie*. Geraadpleegd op 16 februari 2023 van <https://opendata.cbs.nl/#/CBS/nl/dataset/84878NED/line?dl=4BB03&ts=1675676955599>.
- Chunaev, P. (2020). Community detection in node-attributed social networks: A survey. *Computer Science Review*, 37. <https://doi.org/10.1016/j.cosrev.2020.100286>.
- Ciliberto, F., Miller, A. R., Nielsen, H. S., & Simonsen, M. (2016). Playing the fertility game at work: an equilibrium model of peer effects. *International Economic Review*, 57(3), 827–856.
- Craig, J. (1994). Replacement level fertility and future population growth. *Population Trends*, 78, 20–22.
- Dommermuth, L., Klobas, J., & Lappegård, T. (2011). Now or later? the theory of planned behavior and timing of fertility intentions. *Advances in Life Course Research*, 16(1), 42–53. <https://doi.org/10.1016/j.alcr.2011.01.002>.
- EuroStat. (2023). *Key figures on the EU in the world*. Geraadpleegd op 15 juli 2023 van <https://ec.europa.eu/eurostat/documents/15216629/16118334/KS-EX-23-001-EN-N.pdf/d4413940-6ef7-2fa8-d6f1-a60cdc4b89f3?version=1.0&t=1676459907834>.
- Fanfan, D., D’Ingeo, D., Vacca, R., & Stacciarini, J.-M. R. (2023). Navigating disconnected social circles: Experiences and personal networks of Latino/a adolescents in the U.S. rural south. *Youth & Society*, 55(3), 469–493. <https://doi.org/10.1177/0044118X211059206>.
- Giannella, E., & Fischer, C. S. (2016). An inductive typology of egocentric networks. *Social Networks*, 47, 15–23. <https://doi.org/10.1016/j.socnet.2016.02.003>.
- Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821–7826. <https://doi.org/10.1073/pnas.122653799>.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380. <https://doi.org/10.1086/225469>.
- Granovetter, M. S. (1985). Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510. <https://doi.org/10.1086/228311>.
- Hedström P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology*, 36, 49–67. <https://doi.org/10.1146/annurev.soc.012809.102632>.
- Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons.
- Kavas, S., & Jong, J. (2020). Exploring the mechanisms through which social ties affect fertility decisions in Turkey. *Journal of Marriage and Family*, 82(4), 1250–1269. <https://doi.org/10.1111/jomf.12668>.

- Keim, S., Klärner, A., & Bernardi, L. (2013) Tie strength and family formation: Which personal relationships are influential? *Personal Relationships*, 20(3), 462–478.
<https://doi.org/10.1111/j.1475-6811.2012.01418.x>.
- Kohler, H. P., Behrman, J. R., & Watkins, S. C. (2001). The density of social networks and fertility decisions: Evidence from south Nyanza district, Kenya. *Demography*, 38(1), 43–58.
<https://doi.org/10.1353/dem.2001.0005>.
- Kotte, M., & Ludwig, V. (2011). Intergenerational transmission of fertility intentions and behaviour in germany: the role of contagion. *Vienna Yearbook of Population Research*, 9, 207–226.
<https://doi.org/10.1553/populationyearbook2011s207>.
- Lin. (2001). *Social capital: A theory of social structure and action*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511815447>.
- Lois, D. (2013). Zur Erklärung von sozialer Ansteckung beim Übergang zur Elternschaft. *Kolner Zeitschrift für Soziologie und Sozialpsychologie*, 65(3), 397-422.
<https://doi.org/10.1007/s11577-013-0213-x>.
- Lois, D. (2016). Types of social networks and the transition to parenthood. *Demographic Research*, 34, 657–688. <https://doi.org/10.4054/DemRes.2016.34.23>.
- Lois, D., & Arránz Becker O. (2014). Is fertility contagious? Using panel data to disentangle mechanisms of social network influences on fertility decisions. *Advances in Life Course Research*, 21, 123–134. <https://doi.org/10.1016/j.alcr.2013.10.001>.
- Lubbers, M.J., Molina, J.L., & McCarty, C. (2007). Personal networks and ethnic identifications. *International Sociology*, 22(6), 721-741. <https://doi.org/10.1177/0268580907082255>.
- Lyngstad, T. H., & Prskawetz, A. (2010). Do siblings' fertility decisions influence each other? *Demography*, 47(4), 923–934. <https://doi.org/10.1007/BF03213733>.
- Marqués-Sánchez, P., Pinto-Carral, A., Fernández-Villa, T., Vázquez-Casares, A., Liébana-Presa, C., & Benítez-Andrades, J. A. (2021). Identification of cohesive subgroups in a university hall of residence during the covid-19 pandemic using a social network analysis approach. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-01390-4>.
- Mathews, P., & Sear, R. (2013). Does the kin orientation of a British woman's social network influence her entry into motherhood? *Demographic Research*, 28, 313–340.
<https://doi.org/10.4054/DemRes.2013.28.11>.
- Maya-Jariego, I. (2021). Building a structural typology of personal networks: individual differences in the cohesion of interpersonal environment. *Social Networks*, 64, 173–180.
<https://doi.org/10.1016/j.socnet.2020.09.006>.
- Maya-Jariego, I., & Holgado-Ramos, D. (2022). Relationships in context and contexts of relationship: a normative transition at the end of secondary education. *Social Network Analysis and Mining*, 12(1). <https://doi.org/10.1007/s13278-022-00962-0>.

- McCarty, C. (2002). Structure in personal networks. *Journal of Social Structure*, 3(1), 1-11.
- McCarty, C., Lubbers, M. J., Vacca, R., & Luís M.J. (2019). *Conducting personal network research: A practical guide*. Guilford Press.
- Montgomery, M. R., & Casterline, J. B. (1996). Social learning, social influence, and new models of fertility. *Population and Development Review*, 22, 151–175. <https://doi.org/10.2307/2808010>.
- Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review. E*, 69(2), 026113-1–026113-15. <https://doi-org.proxy-ub.rug.nl/10.1103/PhysRevE.69.026113>.
- Pink, S., Leopold, T., & Engelhardt, H. (2014). Fertility and social interaction at the workplace: Does childbearing spread among colleagues? *Advances in Life Course Research*, 21, 113–122. <https://doi.org/10.1016/j.alcr.2013.12.001>.
- Robins, G. L. (2015). *Doing social network research: network-based research design for social scientists*. SAGE Publications. <https://doi.org/10.4135/9781473916753>.
- Sociaal en Cultureel Planbureau. (2019). *Nederlandse waarden en normen internationaal vergeleken: Hoofdstuk 9 van Denkend aan Nederland*. Geraadpleegd op 16 juli 2023 van <https://www.scp.nl/publicaties/monitors/2019/06/26/denkend-aan-nederland>.
- Steglich, C., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: separating selection from influence. *Sociological Methodology*, 40, 329–393. <https://doi-org.proxy-ub.rug.nl/10.1111/j.1467-9531.2010.01225.x>.
- Stulp, G. (2021). Collecting Large Personal Networks in a Representative Sample of Dutch Women. *Social Networks*, 64, 63-71. <https://doi.org/10.1016/j.socnet.2020.07.012>.
- Stulp, G. (2020). Methods and Materials of the Social networks and fertility survey (Sociale realties en kinderkeuzes). <https://doi.org/10.34894/EZCDOA>.
- Stulp, G., & Barrett, L. (2021). Do Data from Large Personal Networks Support Cultural Evolutionary Ideas about Kin and Fertility? *Social Sciences*, 10(5), <https://doi.org/10.3390/socsci10050177>.
- Vacca, R., Cañarte, D., & Vitale, T. (2022). Beyond ethnic solidarity: the diversity and specialisation of social ties in a stigmatised migrant minority. *Journal of Ethnic and Migration Studies*, 48(13), 3113–3141. <https://doi.org/10.1080/1369183X.2021.1903305>.
- Vacca, R. (2020). Structure in personal networks: constructing and comparing typologies. *Network Science*, 8(2), 142–167. <https://doi.org/10.1017/nws.2019.29>.
- Wasserman, S., & Faust, K. (1994). Social network Analysis in Social and Behavioral Science. In *Social network analysis: methods and applications* (pp. 3-27). New York: Cambridge University Press.

- Wellman, B., & Wortley, S. (1990). Different strokes from different folks: community ties and social support. *American Journal of Sociology*, 96(3), 558–588. <https://doi-org.proxy-ub.rug.nl/10.1086/229572>.
- Wiertz, D., & de Graaf, N.D. (2022). The climate crisis: what sociology can contribute. In K. Gërxhani, N.D. de Graaf, & W. Raub (eds.), *Handbook of Sociological Science: Contributions to Rigorous Sociology* (pp. 475-492). <https://doi.org/10.4337/9781789909432>.

Bijlagen

Bijlage I: Identificatie van netwerktypen

Structurele kenmerken

Allereerst heeft de identificatie van subgroepen plaatsgevonden. Dit is gebeurd door middel van het GN-algoritme te gebruiken in het programma R. In deze methode wordt voor elke edge de edge betweenness uitgerekend en degene met de hoogste score wordt verwijderd. Dit gaat door tot alle edges zijn verwijderd. Daarna wordt bekeken welke subgroepen hieruit naar voren komen.

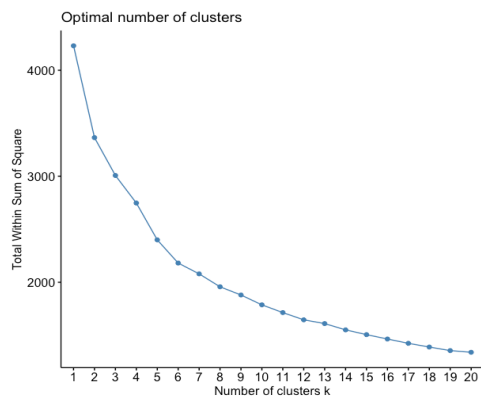
Belangrijke R codes hierbij zijn `edge.betweenness` en `edge.betweenness.community` uit `igraph`.

In de tweede stap zijn de drie variabelen gemaakt die gebruikt worden voor de k-medoid analyse. Dit zijn de modulariteit, subgroepen met 1 of 2 leden en subgroepen met 3 of meer leden. Elk van deze variabelen is gestandaardiseerd voor gebruik in de k-medoid analyse. Op basis van de Within Cluster Sum of Squares (AIC) en de Silhouette score is bepaald dat de meest geschikte typologie bestaat uit 6 netwerktypen. De AIC-score ligt onder de 500 (Figuur 5) en de Silhouette-score is 0,50 (Figuur 6). Als de resultaten worden bekeken dan zou het aannemen van 3 netwerktypen een iets lagere Silhouette-score opleveren en minder variantie verklaren, maar het levert wel een minder complex model op. Echter, hier is niet voor gekozen.

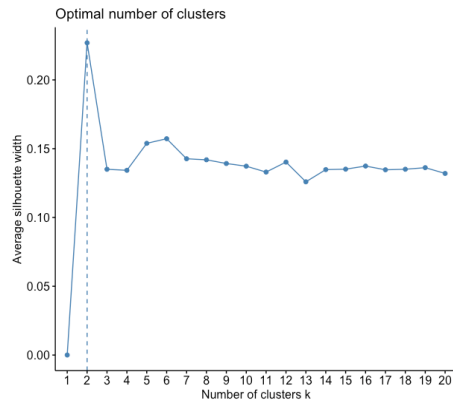
Om de k-medoid analyse uit te voeren is de Partitioning Around Medoids (PAM) functie in R gebruikt. Deze functie is gemaakt om de k medoid analyse uit te voeren. Om de k medoid analyse uit te voeren is de gestandaardiseerde dataset gebruikt, is aangegeven dat er zes clusters gemaakt dienen te worden en wordt de afstandsmaat euclidean gebruikt. Dit is conform Vacca (2020).

Structurele en compositie kenmerken

Op basis van de theorie werd verwacht dat de clustering beter zou verlopen als het compositie kenmerk 'relatie' meegenomen werd in de analyse. Om de compositie te vatten heb ik de volgende variabelen aangemaakt: aantal naaste familieleden, aantal andere familieleden, aantal vrienden en kennissen en aantal collega's. Aangezien elk netwerk uit 25 alters bestaat, is de variabele met het aantal vrienden en kennissen niet meegenomen in de analyse. Deze variabelen zijn gestandaardiseerd.



Figuur 15: AIC-score, beide soorten kenmerken



Figuur 16: Silhouette-score, beide soorten kenmerken

Van alle categorieën is een nieuwe variabele gemaakt. Elke variabele geeft aan hoeveel van deze relatie bestaan in het netwerk van de vrouw. Gezamenlijk vormen deze vier variabelen altijd 25 alters. Daarom is ervoor gekozen om drie van de vier variabelen te gebruiken in de analyse. Deze zijn eerst gestandaardiseerd. De totale within sum of squares is aanzienlijk toegenomen door het toevoegen van deze variabelen; wat duidt op een minder goed model. Het aantal clusters dat ideaal is voor dit model is lastig af te lezen uit Figuur 15 en Figuur 16. Het ‘elleboog’ punt en de meest optimale Silhouette score komen namelijk niet overeen. In dit geval is ervoor gekozen om de Silhouette-score leidend te laten zijn, want deze daalt aanzienlijk en klimt niet meer op. Hierdoor zou de ideale clustering op basis van deze kenmerken bestaan uit twee clusters, met een AIC van ongeveer 3300 en een Silhouette-score van 0,23.

Bijlage II: Uitwerking typologie structurele en compositie kenmerken

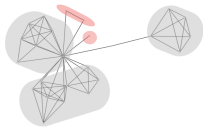
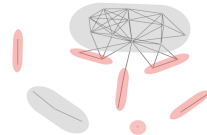
Het tweede model wordt niet beter dan het eerste model. De compositie kenmerken zijn niet in staat om de clustering te verbeteren ten opzichte van enkel de structurele kenmerken. Voor de volledigheid is echter wel gekeken hoe de typologie eruit had gezien en of er samenhang is met de fertiliteitswensen van de Nederlandse vrouwen. De typologie zou bestaan uit twee netwerktypen. Om verwarring te voorkomen is doorgeteld vanaf zes.

Beschrijvende statistieken

Tabel 10: Beschrijvende statistieken van netwerktypen structurele en compositie kenmerken

| | Netwerktype 7 (<i>n</i> = 489) | Netwerktype 8 (<i>n</i> = 217) |
|---|---------------------------------|---------------------------------|
| Gemiddelde dissimilariteit | 2,03 | 2,23 |
| Maximale dissimilariteit | 6,29 | 5,21 |
| Diameter | 9,00 | 8,55 |
| Scheiding | 0,53 | 0,53 |
| Modulariteit | 0,46 | 0,21 |
| Aantal subgroepen medoid | 3 | 2 |
| Aantal isolaten en dyads medoid | 2 | 6 |
| Gemiddelde en sd modulariteit | 0,46 (0,12) | 0,21 (0,13) |
| Gemiddelde en sd aantal subgroepen | 3,31 (0,78) | 2,10 (0,63) |
| Gemiddelde en sd aantal isolaten en dyads | 2,41 (2,15) | 6,62 (3,55) |
| Gemiddeld aantal naaste familieleden | 4,04 (1,81) | 4,28 (2,03) |
| Gemiddeld aantal andere familieleden | 5,04 (3,73) | 6,17 (4,62) |
| Gemiddeld aantal vrienden en kennissen | 11,96 (5,07) | 11,61 (5,15) |
| Gemiddeld aantal collega's | 3,95 (3,99) | 2,94 (3,52) |
| Percentage naaste familieleden | 16,2% | 17,1% |
| Percentage andere familieleden | 20,2% | 24,7% |
| Percentage vrienden en kennissen | 47,8% | 46,4% |
| Percentage collega's | 15,8% | 11,8% |

Tabel 11: Netwerktypen structurele en compositie kenmerken ($n = 706$)

| Netwerk type | Omschrijving | Schematisch plaatje* | Vergelijkbaar netwerktype van Vacca (2020) |
|--------------|---|--|--|
| 7 | In dit netwerktype is er sprake van 3 subgroepen en 2 isolaten en dyads. Daarnaast zijn er ook nog veel relaties tussen de groepen. |  | Sparse multi-factional |
| 8 | Dit netwerktype heeft 2 subgroepen en daarnaast meerdere isolaten en dyads. |  | Weakly bi-factional |

*Grijs: subgroepen met meer dan drie alters, rood: subgroepen met één of twee alters.

De twee netwerktypen verschillen aardig in grootte; netwerktype 7 bestaat namelijk uit 489 vrouwen en netwerktype 8 uit 217. Ook de dissimilariteitscijfers liggen beduidend hoger dan in het model met enkel de structurele kenmerken. Dit is te verklaren doordat er minder clusters zijn, waardoor de verschillen binnen het cluster hoger zullen liggen. Netwerktype 7 heeft drie subgroepen met drie of meer alters, een relatief hoge modulariteitscore en twee subgroepen met één of twee alters. Dit zou in de typologie van Vacca (2020) worden gekwalificeerd als tri-factional of sparse multi-factional. In vergelijking met de eerdere typologie gemaakt in dit onderzoek komt het overeen met het netwerktype strongly tri-factional. Daarentegen heeft netwerktype 8 maar twee subgroepen die bestaan uit drie of meer subgroepen, een zeer lage modulariteit en veel subgroepen met één of twee alters. Om die reden komt dit netwerktype overeen met het weakly bi-factional netwerktype.

Voor Nederlandse vrouwen geldt dat ze gemiddeld 4,11 naaste familieleden in hun netwerk hebben zitten. Dit gemiddelde ligt hoger voor vrouwen met netwerktype 8 ($gem_8 = 4,28, sd_8 = 2,03$). Daarbij hebben vrouwen met netwerktype 8 ook gemiddeld meer andere familieleden in hun netwerk zitten ($gem_8 = 6,17, sd_8 = 4,62$) ten opzichte van vrouwen met netwerktype 7 ($gem_7 = 5,04, sd_7 = 3,73$). Het verschil in andere familieleden is zelf significant ($t(346,89) = -3,17, p < 0,01$). Daartegenover bestaat het netwerk van vrouwen met netwerktype 7 voor een groter deel uit vrienden en kennissen ($gem_7 = 11,96, sd_7 = 5,07; gem_8 = 11,61, sd_8 = 5,15$). Daarbij gaan vrouwen met netwerktype 7 ook significant vaker om met collega's ($t(464,88) = 3,37, p < 0,01$).

Aan de hand van de structurele kenmerken is vastgesteld dat netwerktype 7 overeenkomt met het strongly tri-factional netwerktype en netwerktype 8 met het weakly bi-factional netwerktype. Aangezien de compositie kenmerken ook zijn meegenomen in dit model, hoort dit ook in de naam van het netwerktype voor te komen. Netwerktype 8 zou om die reden gekwalificeerd kunnen worden als

family-oriented weakly tri-factional. Echter, de verschillen zijn dermate klein dat ik dit niet gerechtvaardigd vind.

Relatie met fertiliteitswensen

Allereerst wordt er gekeken naar het verschil in kinderswens. Hiervoor geldt dat 60% van de vrouwen met een strongly tri-factional netwerktype een kinderswens heeft tegenover 58% van de vrouwen met een weakly bi-factional netwerktype. Uit de Kruskal Wallis toets blijkt dat er geen significante verschillen bestaan in kinderswens tussen deze twee netwerktypen ($H(1) = 0,00$, $p = 0,99$). De effectgrootte is kleiner dan 0,00; er wordt geen variantie verklaard in de kinderswens van vrouwen door ze in te delen in netwerktypen gebaseerd op zowel structurele als compositie kenmerken.

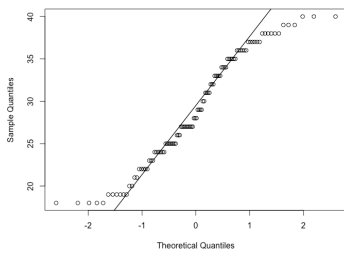
Dezelfde toets is uitgevoerd om te kijken of deze twee netwerktypen verschillende ideale gezinsgroottes hebben. In het strongly tri-factional netwerktype vinden vrouwen een gezinsgrootte van gemiddeld 2,34 het meest ideaal. Daartegenover zien vrouwen met het weakly bi-factional netwerktype gemiddeld 2,31 kinderen als ideaal. Dit verschil is niet significant ($H(1) = 1,00$, $p = 0,32$).

Bijlage III: Assumpties uitgevoerde toetsen

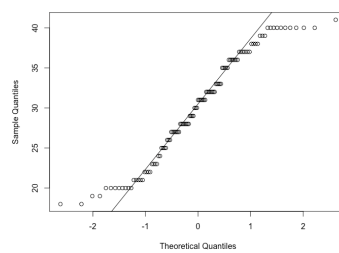
ANOVA, Kruskal Wallis toets en Chi-kwadraat toets voor beschrijvende statistieken

In subparagraaf 4.1.3 wordt ANOVA gebruikt om de verschillen tussen de netwerktypen te onderzoeken in leeftijd, het aantal kinderen en het aantal biologische kinderen. De assumpties voor deze toets zijn normaliteit, homogeniteit van varianties en onafhankelijkheid. Aangezien voor de kenmerken van de vrouwen is gekeken naar het verschil tussen netwerktypen, zijn deze kenmerken elk een afhankelijke variabele. Ze zullen om de beurt worden besproken.

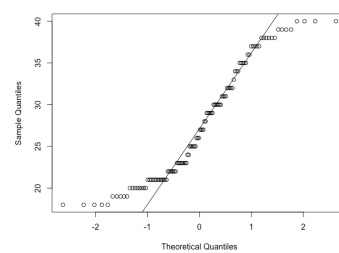
Leeftijd



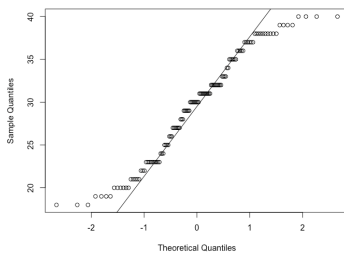
Figuur 17a: strongly bi-factional



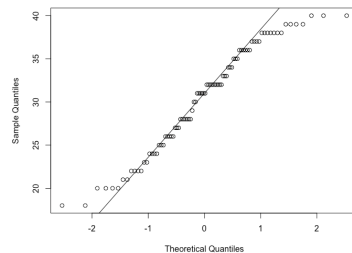
Figuur 17b: Weakly bi-factional



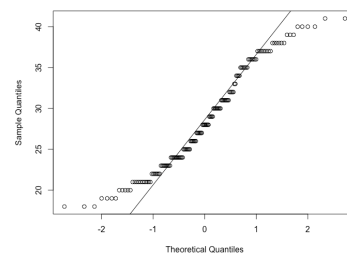
Figuur 17c: Strongly tri-factional



Figuur 17d: Weakly tri-factional



Figuur 17e: Sparse tri-factional



Figuur 17f: Dense multi-factional

Figuur 17: QQ-plots van leeftijdsverdeling per netwerktype

Grotendeels liggen alle punten in de QQ-plot op de lijn, dit duidt op een normaalverdeling. Daarbij heeft elke kop of staart welk wat punten die niet op de lijn vallen, maar dit zal geen problemen opleveren voor de analyse. Op basis van de Levine test is vastgesteld dat er sprake is van homogeniteit van de varianties ($F(5,699) = 0,91, p = 0,47$). Als laatst zijn de groepen onafhankelijk van elkaar.

Aantal kinderen en aantal biologische kinderen

Er wordt niet voldaan aan de assumpties voor de ANOVA-toets. Om de resultaten toch te kunnen vergelijken is ervoor gekozen om hier een Kruskal Wallis toets uit te voeren. De assumptie hiervoor is

dat het moet gaan om een ordinale of continue variabele. In beide gevallen is sprake van een continue ratio variabele.

Hebben van kinderen

Het hebben van kinderen is een binaire variabele, namelijk wel kinderen of geen kinderen. Dit wordt gezien als een ordinale afhankelijke variabele met twee groepen. Om die reden is ervoor gekozen om een chi-kwadraat toets te doen voor onafhankelijke groepen. De assumpties hiervoor is dat het om een categorische variabele gaat, dat de observaties onafhankelijk zijn en dat de groepsgrootte binnen elke groep groter dient te zijn dan 5. Hier wordt aan voldaan.

Hebben van een partner

De omschrijving die is gegeven bij ‘hebben van kinderen’ is vergelijkbaar met het hebben van een partner. Het gaat om een binaire variabele, namelijk wel een partner of geen partner. Ook hier wordt aan de assumpties voldaan en zal een chi-kwadraattoets worden uitgevoerd.

Kruskal Wallis toets

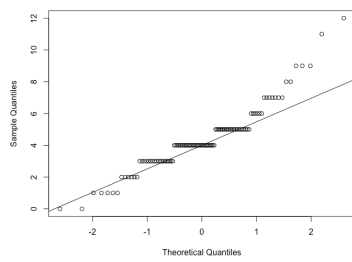
De enige voorwaarde voor de non-parametrische Kruskal Wallis toets is dat de afhankelijke variabele ordinaal of continu is. In dit geval hebben we twee afhankelijke variabele, namelijk kinderwens en ideale gezinsgrootte. Kinderwens is een ordinale variabele met vijf categorieën. Daarnaast is ideale gezinsgrootte een continue variabele, namelijk een ratio variabele.

Aangezien beide variabele meerdere groepen kennen, waardoor verschillende toetsen tegelijk plaatsvinden is gekozen om de post-hoc analyse Bonferroni te gebruiken. Deze toets corrigeert het significantieniveau voor het aantal toetsen dat wordt gedaan.

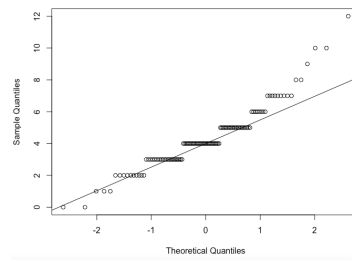
ANOVA relatie tussen structurele netwerktypen en compositie kenmerken

Om het verschil in aantal leden van een bepaalde categorie in relatie te testen is de ANOVA-toets gebruikt. De assumpties die gelden voor een ANOVA toets zijn normaliteit, homogeniteit van variantie en onafhankelijkheid. Zoals eerder vastgesteld is er sprake van onafhankelijke groepen. De andere twee assumpties zullen hieronder worden uitgewerkt.

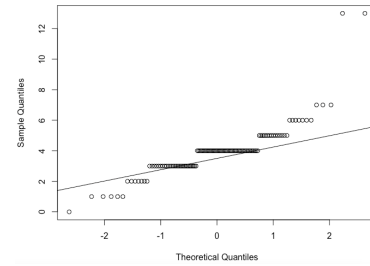
Naaste familieleden



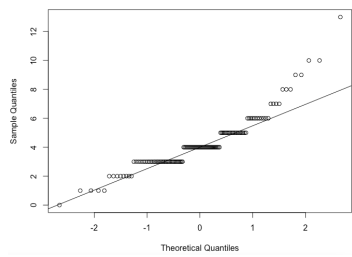
Figuur 18a: strongly bi-factional



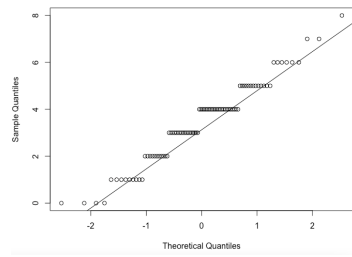
Figuur 18b: Weakly bi-factional



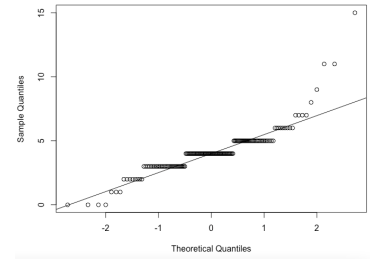
Figuur 18c: Strongly tri-factional



Figuur 18d: Weakly tri-factional



Figuur 18e: Sparse tri-factional

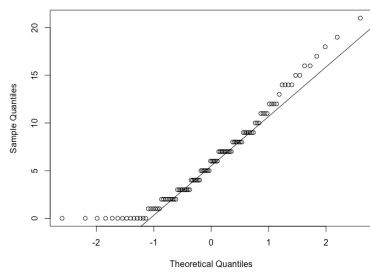


Figuur 18f: Dense multi-factional

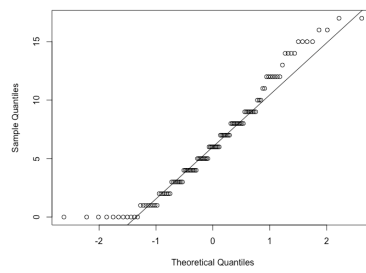
Figuur 18: QQ-plots van naaste familieleden verdeling per netwerktype

Zoals aan de QQ-plots te zien is, is er in de meeste gevallen sprake van een enigszins rechtsscheve verdeling. Alleen het vijfde netwerktype, Figuur 18e, lijkt hier geen last van te hebben. Kijkend naar de y-as hebben zij ook relatief weinig grote aantallen familieleden. Echter, de ANOVA is een robuuste toets en bijna alle groepen hebben meer dan 100 respondenten. Daarbij is er geen significant verschil in variantie ($F(5,700) = 1,53, p = 0,18$).

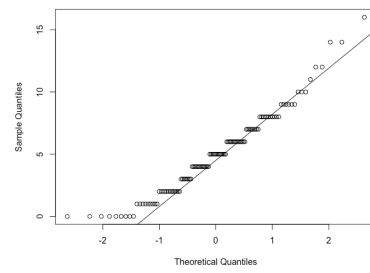
Andere familieleden



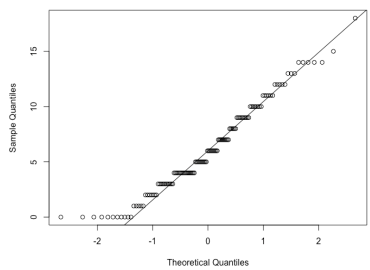
Figuur 19a: strongly bi-factional



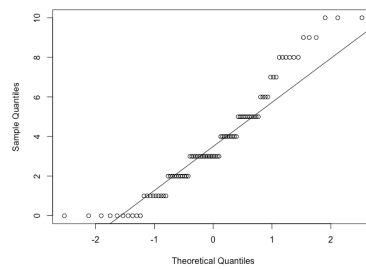
Figuur 19b: Weakly bi-factional



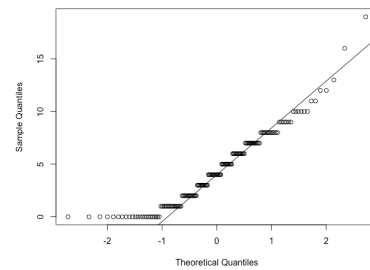
Figuur 19c: Strongly tri-factional



Figuur 19d: Weakly tri-factional



Figuur 19e: Sparse tri-factional

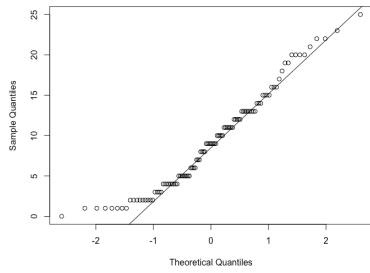


Figuur 19f: Dense multi-factional

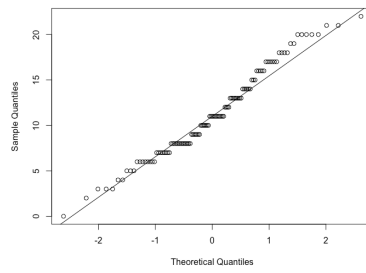
Figuur 19: QQ-plots van andere familieleden verdeling per netwerktype

In Figuur 19 is te zien dat in de verdeling van andere familieleden bij elk netwerktype sprake is van een verdeling met staarten. Oftewel, er worden redelijk kleine aantallen en redelijk grote aantallen gevonden. De grootste verschillen zijn te vinden bij het sparse tri-factional netwerktype. Wel wordt een significant verschil gevonden in de variantie tussen de netwerktypen ($F(5,700) = 9,42, p < 0,01$). Ook hier geldt dat de groepen nog altijd groot zijn, waardoor deze schendingen niet erg zijn.

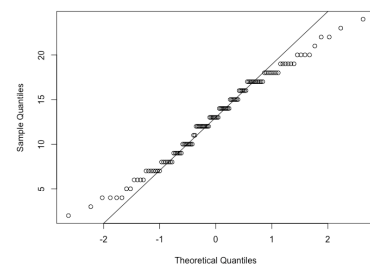
Vrienden en kennissen



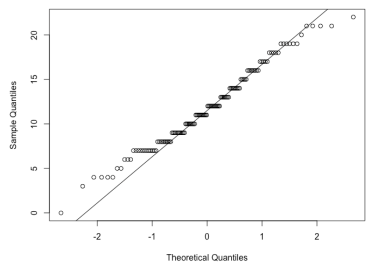
Figuur 20a: strongly bi-factional



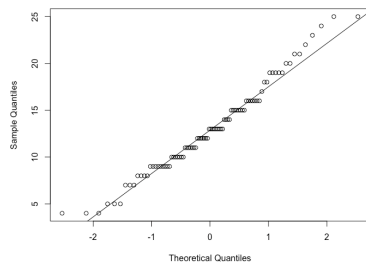
Figuur 20b: Weakly bi-factional



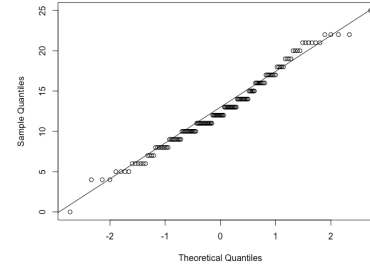
Figuur 20c: Strongly tri-factional



Figuur 20d: Weakly tri-factional



Figuur 20e: Sparse tri-factional

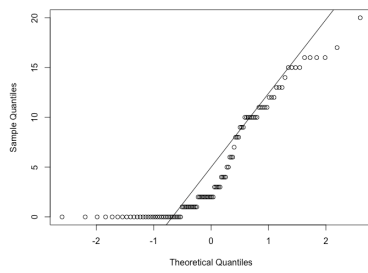


Figuur 20f: Dense multi-factional

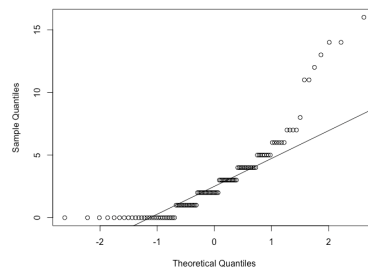
Figuur 20: QQ-plots van vrienden en kennissen verdeling per netwerktype

De verdeling van de vrienden en kennissen is in alle netwerktypen normaal. De grootste afwijking is te vinden bij het strongly bi-factional netwerktype waar de verdeling mogelijk lichtelijk rechtsscheef is. Wel wordt een significant verschil gevonden in de variantie tussen de netwerktypen ($F(5,700) = 3,43, p < 0,01$). Ook hier geldt dat de groepen nog altijd groot zijn, waardoor deze schendingen niet erg zijn.

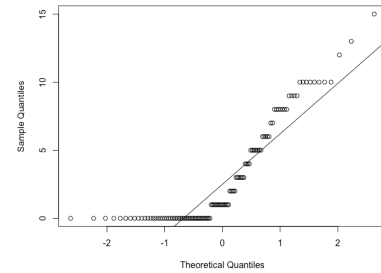
Collega's



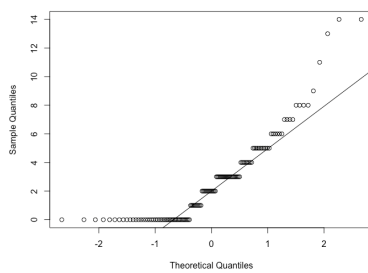
Figuur 21a: strongly bi-factional



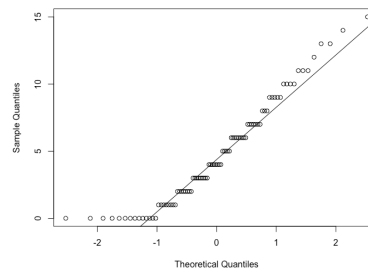
Figuur 21b: Weakly bi-factional



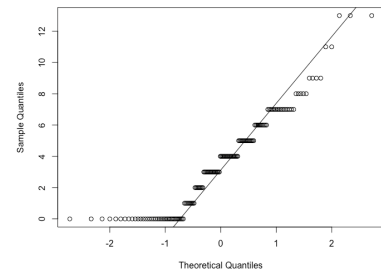
Figuur 21c: Strongly tri-factional



Figuur 21d: Weakly tri-factional



Figuur 21e: Sparse tri-factional



Figuur 21f: Dense multi-factional

Figuur 21: QQ-plots van collega's verdeling per netwerktype

De verdeling van collega's is niet normaal verdeeld. In elk netwerktype geven veel vrouwen aan dat ze geen collega's hebben in het netwerk. Daartegenover zijn er ook veel vrouwen die aangeven veel collega's in het netwerk te hebben. Oftewel, de verdelingen hebben in alle gevallen aanwezige staarten. Daarbij is ook een significant verschil gevonden in de variantie tussen de netwerktypen ($F(5,700) = 9,50, p < 0,01$). Ook hier geldt dat de groepen nog altijd groot zijn, waardoor deze schendingen niet erg zijn.

Kurkus Wallis toets

Aangezien in een aantal van bovenstaande gevallen de assumpties worden geschonden, is ter controle de non-parametrische Kurkus Wallis toets uitgevoerd. Hier worden geen afwijkende resultaten gevonden van de ANOVA-toets. Anders verwoord, dezelfde netwerktypen verschillen significant van elkaar op de verdeling van relaties. Om die reden is ervoor gekozen om de ANOVA te rapporteren.

Bijlage IV: R-code voor analyses

Om de grootte van het document te beperken is gebruik gemaakt van twee kolommen om de informatie toe te voegen. In deze bijlage wordt de volledige code gegeven die is gebruikt voor dit onderzoek. Allereerst wordt de data klaargemaakt voor gebruik. Daarna worden de variabelen gemaakt die nodig zijn voor de analyse. Ten derde volgt de clusteranalyse. Daarna volgen de beschrijvende statistieken en de bivariate analyses. Daarna wordt ingegaan op de samenhang tussen de structurele netwerktypen en de fertiliteitswensen. Als laatste worden de compositie kenmerken toegevoegd en worden de voorgaande stappen herhaald.

Bij het lezen van de code is het belangrijk dat de namen van de clusters en de netwerktypen niet overeenkomen. Dit kan voor verwarring zorgen. Netwerktipe 1 komt overeen met cluster 1, netwerktipe 2 met cluster 6, netwerktipe 3 met cluster 3, netwerktipe 4 met cluster 2, netwerktipe 5 met cluster 4 en netwerktipe 6 met cluster 5.

```

### Benodigde pakketten R
#### packages
# if (!require("remotes"))
install.packages("remotes")
# remotes::install_github("gertstulp/FertNet")
# install.packages("tidygraph")
# if (!require("ggraph"))
install.packages("ggraph")

library(FertNet)
library(tidygraph)
library(ggraph)
library(purrr)
library(igraph)
library(dplyr)
library(car)
library(factoextra)
library(cluster)

### 3.1.2 De verkregen dataset
#### load-in
data <- produce_data(tidygraph_col = TRUE,
                    background_vars = TRUE)

#### data selectie 1
data_sel <- data |>
  filter(!is.na(edgelist) & !is.na(tidygraph) &
!is.na(alter_attr))

#### data selectie 2
exclude <- c(
  # reported fewer than 25 alters, or mentioned
  # same person twice
  "alters reported < 25",
  "alters reported < 25 & >= 10 missing on
  alter attributes",
  "Listed child below 18 & alters reported < 25
  & all alter-ties missing & >= 10 missing on
  alter attributes",
  "25 alters too many",
  "removed duplicate alter",
  # reported no existing ties between alters
  "all alter-ties missing",
  "all alter-ties missing & >= 10 missing on
  alter attributes",
  # did survey on phone/ipad against explicit
  instructions
  "survey on phone/ipad",
  "survey on phone/ipad & >= 10 missing on
  alter attributes",
  "probably on phone",
  "survey on phone/ipad & all alter-ties missing
  & >= 10 missing on alter attributes",
  # reported problematic alter relations
  "problematic alter relations & >= 10 missing
  on alter attributes",
  # more than 10 missing values on 'mandatory'
  alter attribute questions
  ">= 10 missing on alter attributes"
)

```

```

data_sel <- data_sel |>
  filter(!(notes %in% exclude))

## Selectie data: nieuwe dataset, voor GN-
analyse
data_sel_GN <- data_sel |>
  select(age, partner, has_children,
num_children, num_children_biol, childwish,
  ideal_family_size, alter_attr, edgelist,
tidygraph)

### 3.2 Operationalering begrippen
## Afhankelijke variabelen: Childwish
numeriek en ideal_family_size (don't know ->
NA)

data_sel_GN$ideal_family_size[data_sel_GN$
ideal_family_size == "I don't know"] <- NA

data_sel_GN <- data_sel_GN |>
  mutate(
  ideal_family_size_num =
as.numeric(ifelse(is.na(data_sel_GN$ideal_fa
mily_size), NA,
data_sel_GN$ideal_family_size)))

data_sel_GN <- data_sel_GN |>
  mutate(
  childwish_num = case_when(
childwish == "Absolutely so" ~ 5,
childwish == "Probably so" ~ 4,
childwish == "I don't know" ~ 3,
childwish == "Probably not" ~ 2,
childwish == "Absolutely not" ~ 1))

## Andere variabelen (beschrijvende stat):age,
has_children, has_partner num_children,
num_children_biol, & primary_relation_a
# has_children naar 0 (geen kinderen) en 1
(wel kinderen)
data_sel_GN <- data_sel_GN |>
  mutate(has_children_num =
case_when(has_children == "Yes" ~ 1,
  has_children == "No" ~ 0))

# num_children NA naar 0, anderen blijven
gelijk
data_sel_GN$num_children[is.na(data_sel_G
N$num_children)] <- 0

# num_children_biol NA naar 0, anderen
blijven gelijk

```

```

data_sel_GN$num_children_biol[is.na(data_se
l_GN$num_children_biol)] <- 0

# has_partner naar 0 (geen partner) en 1 (wel
partner)
data_sel_GN <- data_sel_GN |>
  mutate(partner_num =
case_when(partner == "Yes" ~ 1,
  partner == "No" ~ 0))

# primary_relation_a into 4 categories
data_sel_GN <- data_sel_GN |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1, TRUE ~
0))),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1, TRUE ~ 0))),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1, TRUE ~ 0))),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school", "Primary
school", "Social activity", "College", "Mutual
acquaintance", "Neighbourhood", "Other") ~
1, TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law", "Partner's friends", "High-school",
"Work", "Primary school", "Social activity",
"College", "Mutual acquaintance",
"Neighbourhood", "Other") ~ 0, TRUE ~ 1))))

data_sel_GN <- data_sel_GN |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

### 4.1.2 Uitvoering analyse Vacca (2020)
### Stap 1:GN algorithm

g <- data_sel_GN$tidygraph
betw <- lapply(g, function(g) betw <-
edge.betweenness(g))
ebc <- lapply(g, function(g) ebc <-
edge.betweenness.community(g))

```

```

### STAP 2: De subgroepstructuur
samenvatten

data_sel_GN <- data_sel_GN |>
  mutate(
    community_detection = map(tidygraph,
  function(x)
  igraph::cluster_edge_betweenness(x, directed
  = FALSE) ),
    community_sizes =
  map(community_detection, function(x)
  c(table(x$membership))),
    comm_1or2 = map_dbl(community_sizes,
  function(x) sum(x <= 2)),
    comm_3orhigher =
  map_dbl(community_sizes, function(x) sum(x
  >= 3)),
    modularity =
  map_dbl(community_detection, function(x)
  igraph::modularity(x))
  )

### STAP 3: K-medoids clustering ~ PAM
data_clus <- data_sel_GN |>
  select(modularity, comm_1or2,
  comm_3orhigher)

data_clus <- data_clus |>
  mutate(
    comm_1or2 = unlist(comm_1or2),
    comm_3orhigher = unlist(comm_3orhigher)
  )

## standaardiseren van data
data_clus_stan <- scale(data_clus)

## Elleboog methode: Within Cluster Sum of
Squares (AIC) = Figuur 5
fviz_nbclust(data_clus_stan, pam, k=20,
  method = "wss")

## Silhouette: Optimaal aantal clusters =
Figuur 6
fviz_nbclust(data_clus_stan, pam, k=20,
  method = "silhouette")

## Optimal clusters: 6, Vacca gebruikt
euclidean
pamC6 <- pam(data_clus_stan, 6, metric =
"euclidean", stand = FALSE)

#Medoids info

```

```

pamC6
pamC6$silinfo
pamC6$medoids
pamC6$id.med #cases: 577, 187, 551, 485,
530, 155
pamC6$objective
pamC6$isolation
pamC6$clusinfo # Tabel 3

## Cluster plot
fviz_cluster(pamC6)
fviz_cluster(pamC6, geom = "point",
  show.clust.cent = TRUE)

### Elk netwerktype een aparte dataset
# netwerktype moderate bi-factional
(netwerktype 1)
data_GN1 <- data_sel_GN |>
  filter(clusters == 1)

# netwerktype weakly tri-factional
(netwerktype 4)
data_GN2 <- data_sel_GN |>
  filter(clusters == 2)

# netwerktype strongly tri-factional
(netwerktype 3)
data_GN3 <- data_sel_GN |>
  filter(clusters == 3)

# netwerktype sparse tri-factional
(netwerktype 5)
data_GN4 <- data_sel_GN |>
  filter(clusters == 4)

# netwerktype dense multi-factional
(netwerktype 6)
data_GN5 <- data_sel_GN |>
  filter(clusters == 5)

# netwerktype weakly bi-factional
(netwerktype 2)
data_GN6 <- data_sel_GN |>
  filter(clusters == 6)

data_sel_GN <- data_sel_GN |>
  mutate(netwerktypen = case_when(
    num_clusters == "1" ~ 1,
    num_clusters == "2" ~ 4,
    num_clusters == "3" ~ 3,
    num_clusters == "4" ~ 5,
    num_clusters == "5" ~ 6,

```

```

    num_clusters == "6" ~ 2,
  ))

### Tabel 4
## Gemiddelde en standaarddeviatie aantal
subgroepen en dyads per netwerktype.
# Netwerktipe 1
mean(data_GN1$comm_3orhigher)
sd(data_GN1$comm_3orhigher)
mean(data_GN1$comm_1or2)
sd(data_GN1$comm_1or2)
mean(data_GN1$modularity)
sd(data_GN1$modularity)

# Netwerktipe 2
mean(data_GN6$comm_3orhigher)
sd(data_GN6$comm_3orhigher)
mean(data_GN6$comm_1or2)
sd(data_GN6$comm_1or2)
mean(data_GN6$modularity)
sd(data_GN6$modularity)

# Netwerktipe 3
mean(data_GN3$comm_3orhigher)
sd(data_GN3$comm_3orhigher)
mean(data_GN3$comm_1or2)
sd(data_GN3$comm_1or2)
mean(data_GN3$modularity)
sd (data_GN3$modularity)

#z-Netwerktipe 4
mean(data_GN2$comm_3orhigher)
sd(data_GN2$comm_3orhigher)
mean(data_GN2$comm_1or2)
sd(data_GN2$comm_1or2)
mean(data_GN2$modularity)
sd(data_GN2$modularity)

# Netwerktipe 5
mean(data_GN4$comm_3orhigher)
sd(data_GN4$comm_3orhigher)
mean(data_GN4$comm_1or2)
sd(data_GN4$comm_1or2)
mean(data_GN4$modularity)
sd(data_GN4$modularity)

# Netwerktipe 6
mean(data_GN5$comm_3orhigher)
sd(data_GN5$comm_3orhigher)
mean(data_GN5$comm_1or2)
sd(data_GN5$comm_1or2)
mean(data_GN5$modularity)
sd(data_GN5$modularity)

# Modularity-scores
data_sel_GN$modularity[[577]]
data_sel_GN$modularity[[155]]
data_sel_GN$modularity[[551]]
data_sel_GN$modularity[[187]]
data_sel_GN$modularity[[485]]
data_sel_GN$modularity[[530]]

### Figuren 4.1
frequency <- data_sel_GN |>
  group_by(comm_3orhigher, comm_1or2) |>
  summarise(freq = n())

color_paLETTE1 <- c('1', '2','3','4','5','6')

modulariteitswaarde <-
c(data_sel_GN$modularity)

# Figuur 2
ggplot(data_sel_GN, aes(x =
comm_3orhigher)) +
  geom_histogram(binwidth = 1, fill = "grey",
color = "black") +
  labs(x = "Subgroepen", y = "Aantal
netwerken")

#Figuur 3
ggplot(data_sel_GN, aes(x = modularity)) +
  geom_histogram(binwidth = 0.1, fill =
"grey", color = "black") +
  labs(x = "Modulariteitswaarde", y = "Aantal
netwerken")

#Figuur 4
ggplot(data_sel_GN, aes(x = modularity)) +
  geom_histogram(binwidth = 0.1, fill =
"grey", color = "black") +
  labs(x = "Modulariteitswaarde", y = "Aantal
netwerken")

# Figuur 10
ggplot(data_sel_GN, aes(x =
comm_3orhigher, y = comm_1or2, color =
color_paLETTE1[netwerktypen])) +
  geom_jitter() +
  labs(title = "Scatterplot aantal netwerken",
x = "Subgroepen", y = "Isolaten en
dyads")
theme(legend.position = "none")

```

#Figuur 9

```
ggplot(data_sel_GN, aes(x = netwerktypen, y
= modularity, color =
color_paLETTE1[netwerktypen])) +
  geom_boxplot() +
  labs(title = "Boxplot van verdeling van
modulariteit",
x = "Netwerktypen", y = "Modulariteit")
+
  theme(legend.position = "none")
```

Figuur 7

```
ggplot(data_sel_GN, aes(x = netwerktypen, y
= comm_3orhigher, color =
color_paLETTE1[netwerktypen])) +
  geom_boxplot() +
  labs(title = "Boxplot van verdeling aantal
subgroepen",
x = "Netwerktypen", y = "Aantal
subgroepen") +
  theme(legend.position = "none")
```

Figuur 8

```
ggplot(data_sel_GN, aes(x = netwerktypen, y
= comm_1or2, color =
color_paLETTE1[netwerktypen])) +
  geom_boxplot() +
  labs(title = "Boxplot van verdeling van
isolaten en dyads",
x = "Netwerktypen", y = "Aantal isolaten
en dyads") +
  theme(legend.position = "none")
```

Figuur 11

```
ggplot(data_sel_GN, aes(x = comm_1or2, y =
modularity, color =
color_paLETTE1[netwerktypen])) +
  geom_point() +
  labs(title = "Scatterplot verdeling
modulariteit",
x = "Isolaten en dyads", y =
"Modulariteit") +
  facet_wrap(~ comm_3orhigher)
```

4.2.1 Visualisatie van 6 medoids

```
# first: ID-medoid: 557, netwerktype 1
ggraph(data_sel_GN$tidygraph[[577]], layout
= "kk") +
  geom_edge_link(colour = "red") +
  geom_node_point(colour = "black") +
  geom_node_text(aes(label = names_a),
colour = "black") +
```

```
labs(colour = NULL) +
  theme_graph()
```

```
g3 <- data_sel_GN$tidygraph[[557]]
plot(g3)
betw3 <- edge.betweenness(g3)
ebc3 <- edge.betweenness.community(g3)
plot(ebc3, g3)
```

```
# second: ID-medoid: 187, netwerktype 4
ggraph(data_sel_GN$tidygraph[[187]], layout
= "kk") +
  geom_edge_link(colour = "orange") +
  geom_node_point(colour = "black") +
  labs(colour = NULL) +
  theme_graph()
```

```
g4 <- data_sel_GN$tidygraph[[187]]
plot(g4)
betw4 <- edge.betweenness(g4)
ebc4 <- edge.betweenness.community(g4)
plot(ebc4, g4)
```

```
# third: ID-medoid: 551, netwerktype 3
ggraph(data_sel_GN$tidygraph[[551]], layout
= "kk") +
  geom_edge_link(colour = "green") +
  geom_node_point(colour = "black") +
  labs(colour = NULL) +
  theme_graph()
```

```
g5 <- data_sel_GN$tidygraph[[551]]
plot(g5)
betw5 <- edge.betweenness(g5)
ebc5 <- edge.betweenness.community(g5)
plot(ebc5, g5)
```

```
# fourth: ID-medoid: 485, netwerktype 5
ggraph(data_sel_GN$tidygraph[[485]], layout
= "kk") +
  geom_edge_link(colour = "light blue") +
  geom_node_point(colour = "black") +
  labs(colour = NULL) +
  theme_graph()
```

```
g6 <- data_sel_GN$tidygraph[[485]]
plot(g6)
betw6 <- edge.betweenness(g6)
ebc6 <- edge.betweenness.community(g6)
plot(ebc6, g6)
```

```
# fifth: ID-medoid: 530, netwerktype 6
```



```
ggraph(data_sel_GN$tidygraph[[530]], layout
= "kk") +
  geom_edge_link(colour = "dark blue") +
  geom_node_point(colour = "black") +
  labs(colour = NULL) +
  theme_graph()
```

```
g7 <- data_sel_GN$tidygraph[[530]]
plot(g7)
betw7 <- edge.betweenness(g7)
ebc7 <- edge.betweenness.community(g7)
plot(ebc7, g7)
```

```
# sixth: ID-medoid: 155, netwerktype 2
ggraph(data_sel_GN$tidygraph[[155]], layout
= "kk") +
  geom_edge_link(colour = "pink") +
  geom_node_point(colour = "black") +
  geom_node_text(aes(label = names_a,
colour = "black")) +
  labs(colour = NULL) +
  theme_graph()
```

```
g8 <- data_sel_GN$tidygraph[[155]]
plot(g8)
betw8 <- edge.betweenness(g8)
ebc8 <- edge.betweenness.community(g8)
plot(ebc8, g8)
```

```
### Variabele clusters toevoegen dataset
data_sel_GN <- data_sel_GN |>
  mutate(
    clusters = pamC6$clustering)
```

Visualisaties voor Figuur 13

```
ggraph(data_sel_GN$tidygraph[[577]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

```
ggraph(data_sel_GN$tidygraph[[155]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

```
ggraph(data_sel_GN$tidygraph[[551]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

```
ggraph(data_sel_GN$tidygraph[[187]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

```
ggraph(data_sel_GN$tidygraph[[485]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

```
ggraph(data_sel_GN$tidygraph[[530]], layout
= "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "black") +
  geom_node_point(aes(colour =
primary_relation_a), size = 7) +
  labs(colour = NULL) +
  theme_graph()
```

Verdeling van ideale gezinsgrootte, leeftijd

ideale gezinsgrootte

```
summary(data_sel_GN$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN1$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN2$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN3$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN4$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN5$ideal_family_size_num,
na.rm = TRUE)
```

```
summary(data_GN6$ideal_family_size_num,
na.rm = TRUE)

sum(data_sel_GN$ideal_family_size_num ==
"0", na.rm = TRUE) #39 = 5,5%
sum(data_sel_GN$ideal_family_size_num ==
"1", na.rm = TRUE) #36 = 5,1%
sum(data_sel_GN$ideal_family_size_num ==
"2", na.rm = TRUE) #323 = 45,8%
sum(data_sel_GN$ideal_family_size_num ==
"3", na.rm = TRUE) #173 = 24,5%
sum(data_sel_GN$ideal_family_size_num ==
"4", na.rm = TRUE) #56 = 7,9%
sum(data_sel_GN$ideal_family_size_num ==
"5", na.rm = TRUE) #8 = 1,1%
sum(data_sel_GN$ideal_family_size_num ==
"6", na.rm = TRUE) #1 = 0,1%
sum(data_sel_GN$ideal_family_size_num ==
"7", na.rm = TRUE) #0
sum(data_sel_GN$ideal_family_size_num ==
"8", na.rm = TRUE) #1 = 0,1%

sd(data_sel_GN$ideal_family_size_num,
na.rm = TRUE)
sd(data_GN1$ideal_family_size_num, na.rm =
TRUE)
sd(data_GN2$ideal_family_size_num, na.rm =
TRUE)
sd(data_GN3$ideal_family_size_num, na.rm =
TRUE)
sd(data_GN4$ideal_family_size_num, na.rm =
TRUE)
sd(data_GN5$ideal_family_size_num, na.rm =
TRUE)
sd(data_GN6$ideal_family_size_num, na.rm =
TRUE)

hist(data_sel_GN$ideal_family_size_num,
breaks = 8, main = "Ideal Family Size", xlab =
"Children")
hist(data_GN1$ideal_family_size_num, breaks
= 4, col = "black", main = "Ideal Family Size
cluster 1")
hist(data_GN2$ideal_family_size_num, breaks
= 8, col = "black", main = "Ideal Family Size
cluster 2")
hist(data_GN3$ideal_family_size_num, breaks
= 5, col = "black", main = "Ideal Family Size
cluster 3")
hist(data_GN4$ideal_family_size_num, breaks
= 6, col = "black", main = "Ideal Family Size
cluster 4")
```

```
hist(data_GN5$ideal_family_size_num, breaks
= 4, col = "black", main = "Ideal Family Size
cluster 5")
hist(data_GN6$ideal_family_size_num, breaks
= 5, col = "black", main = "Ideal Family Size
cluster 6")
```

leeftijd

```
summary(data_sel_GN$age, na.rm = TRUE)
summary(data_GN1$age, na.rm = TRUE)
summary(data_GN2$age, na.rm = TRUE)
summary(data_GN3$age, na.rm = TRUE)
summary(data_GN4$age, na.rm = TRUE)
summary(data_GN5$age, na.rm = TRUE)
summary(data_GN6$age, na.rm = TRUE)
```

```
sd(data_sel_GN$age, na.rm = TRUE)
sd(data_GN1$age, na.rm = TRUE)
sd(data_GN2$age, na.rm = TRUE)
sd(data_GN3$age, na.rm = TRUE)
sd(data_GN4$age, na.rm = TRUE)
sd(data_GN5$age, na.rm = TRUE)
sd(data_GN6$age, na.rm = TRUE)
```

```
hist(data_sel_GN$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling")
hist(data_GN1$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 1",
xlim = c(18,40), xlab = "Leeftijd")
hist(data_GN2$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 2",
xlim = c(18,40), xlab = "Leeftijd")
hist(data_GN3$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 3",
xlim = c(18,40), xlab = "Leeftijd")
hist(data_GN4$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 4",
xlim = c(18,40), xlab = "Leeftijd")
hist(data_GN5$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 5",
xlim = c(18,40), xlab = "Leeftijd")
hist(data_GN6$age, breaks = 23, col =
"black", main = "Leeftijdsverdeling cluster 6",
xlim = c(18,40), xlab = "Leeftijd")
```

Verdeling van groepen, partner,

```
has_children, num_children,
num_childre_biol, childwish,
ideal_family_size_num
```

```
# alle clusters gezamenlijk
mean(data_sel_GN$age, na.rm = TRUE)
sd(data_sel_GN$age, na.rm = TRUE)
```

```

mean(data_sel_GN$partner_num, na.rm =
TRUE)
mean(data_sel_GN$has_children_num, na.rm
= TRUE)
mean(data_sel_GN$num_children, na.rm =
TRUE)
sd(data_sel_GN$num_children, na.rm =
TRUE)
mean(data_sel_GN$num_children_biol, na.rm
= TRUE)
sd(data_sel_GN$num_children_biol, na.rm =
TRUE)
mean(data_sel_GN$childwish_num, na.rm =
TRUE)
sd(data_sel_GN$childwish_num, na.rm =
TRUE)
mean(data_sel_GN$ideal_family_size_num,
na.rm = TRUE)
sd(data_sel_GN$ideal_family_size_num,
na.rm = TRUE)

```

```

sum(data_sel_GN$partner_num == "0") #192
= 27,2%
sum(data_sel_GN$partner_num == "1") #514
= 72,8%
hist(data_sel_GN$partner_num, breaks = 2,
main = "Gender partner", xlim = c(0,1), xlab =
"Gender")

```

```

sum(data_sel_GN$has_children_num == "0")
#456 = 64,6%
sum(data_sel_GN$has_children_num == "1")
#250 = 35,4%
hist(data_sel_GN$has_children_num, breaks =
2, main = "Has Children", xlim = c(0,1), xlab
= "Has Children")

```

```

sum(data_sel_GN$num_children == "0") #456
= 64,6%
sum(data_sel_GN$num_children == "1") #99
= 14,0%
sum(data_sel_GN$num_children == "2") #109
= 15,4%
sum(data_sel_GN$num_children == "3") #32
= 4,5%
sum(data_sel_GN$num_children == "4") #7 =
1,0%
sum(data_sel_GN$num_children == "5") #1 =
0,1%
sum(data_sel_GN$num_children == "6") #1 =
0,1%
sum(data_sel_GN$num_children == "7") #0 =
0%

```

```

sum(data_sel_GN$num_children == "8") #1 =
0,1%
hist(data_sel_GN$num_children, breaks = 8,
main = "Number of Children", xlab =
"Children")
summary(data_sel_GN$num_children)
sd(data_sel_GN$num_children)

```

```

sum(data_sel_GN$num_children_biol == "0",
na.rm = TRUE) #14 = 2,0% (of total)
sum(data_sel_GN$num_children_biol == "1",
na.rm = TRUE) #97 = 13,7%
sum(data_sel_GN$num_children_biol == "2",
na.rm = TRUE) #99 = 14,0%
sum(data_sel_GN$num_children_biol == "3",
na.rm = TRUE) #31 = 4,4%
sum(data_sel_GN$num_children_biol == "4",
na.rm = TRUE) #7 = 1,0%
sum(data_sel_GN$num_children_biol == "5",
na.rm = TRUE) #1 = 0,1%
sum(data_sel_GN$num_children_biol == "6",
na.rm = TRUE) #1 = 0,1%
sum(data_sel_GN$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_sel_GN$num_children_biol == "8",
na.rm = TRUE) #0 = 0%
hist(data_sel_GN$num_children_biol, breaks
= 5, main = "Number of Biological Children",
xlab = "Children")
summary(data_sel_GN$num_children_biol,
na.rm = TRUE)
sd(data_sel_GN$num_children_biol, na.rm =
TRUE)

```

```

sum(data_sel_GN$childwish_num == "1")
#89 = 12,6%
sum(data_sel_GN$childwish_num == "2")
#98 = 13,9%
sum(data_sel_GN$childwish_num == "3")
#102 = 14,4%
sum(data_sel_GN$childwish_num == "4")
#227 = 32,2%
sum(data_sel_GN$childwish_num == "5")
#190 = 26,9%
hist(data_sel_GN$childwish_num, breaks = 6,
main = "Childwish", xlab = "Childwish")
summary(data_sel_GN$childwish_num)
summary(data_GN1$childwish_num, na.rm =
TRUE)
summary(data_GN2$childwish_num, na.rm =
TRUE)
summary(data_GN3$childwish_num, na.rm =
TRUE)

```

```
summary(data_GN4$childwish_num, na.rm = TRUE)
summary(data_GN5$childwish_num, na.rm = TRUE)
summary(data_GN6$childwish_num, na.rm = TRUE)
sd(data_sel_GN$childwish_num)
```

```
# Netwerktipe 1: N = 106 GN1
sum(data_GN1$partner_num == "0") #24 = 22,6%
sum(data_GN1$partner_num == "1") #82 = 77,4%
hist(data_GN1$partner_num, breaks = 2, main = "Gender partner", xlim = c(0,1), xlab = "Gender")
```

```
sum(data_GN1$has_children_num == "0") #66 = 62,3%
sum(data_GN1$has_children_num == "1") #40 = 37,7%
hist(data_GN1$has_children_num, breaks = 2, main = "Has Children", xlim = c(0,1), xlab = "Has Children")
```

```
sum(data_GN1$num_children == "0") #66 = 62,3%
sum(data_GN1$num_children == "1") #15 = 14,2%
sum(data_GN1$num_children == "2") #18 = 17,0%
sum(data_GN1$num_children == "3") #6 = 5,7%
sum(data_GN1$num_children == "4") #1 = 0,9%
sum(data_GN1$num_children == "5") #0 = 0%
sum(data_GN1$num_children == "6") #0 = 0%
sum(data_GN1$num_children == "7") #0 = 0%
sum(data_GN1$num_children == "8") #0 = 0%
hist(data_GN1$num_children, breaks = 5, main = "Number of Children", xlab = "Children")
summary(data_GN1$num_children)
sd(data_GN1$num_children)
```

```
sum(data_GN1$num_children_biol == "0", na.rm = TRUE) #4 = 3,8% (of total)
sum(data_GN1$num_children_biol == "1", na.rm = TRUE) #13 = 12,3%
```

```
sum(data_GN1$num_children_biol == "2", na.rm = TRUE) #16 = 15,1%
sum(data_GN1$num_children_biol == "3", na.rm = TRUE) #6 = 5,7%
sum(data_GN1$num_children_biol == "4", na.rm = TRUE) #1 = 0,9%
hist(data_GN1$num_children_biol, breaks = 5, main = "Number of Biological Children", xlab = "Children")
summary(data_GN1$num_children_biol, na.rm = TRUE)
sd(data_GN2$num_children_biol, na.rm = TRUE)
```

```
sum(data_GN1$childwish_num == "1") #10 = 9,4%
sum(data_GN1$childwish_num == "2") #15 = 14,2%
sum(data_GN1$childwish_num == "3") #14 = 13,2%
sum(data_GN1$childwish_num == "4") #36 = 34,0%
sum(data_GN1$childwish_num == "5") #31 = 29,2%
hist(data_GN1$childwish_num, breaks = 5, main = "Childwish", xlab = "Childwish")
summary(data_GN1$childwish_num)
sd(data_GN1$childwish_num)
```

```
sum(data_GN1$ideal_family_size_num == "0", na.rm = TRUE) #4 = 3,8%
sum(data_GN1$ideal_family_size_num == "1", na.rm = TRUE) #4 = 3,8%
sum(data_GN1$ideal_family_size_num == "2", na.rm = TRUE) #47 = 44,3%
sum(data_GN1$ideal_family_size_num == "3", na.rm = TRUE) #29 = 27,4%
sum(data_GN1$ideal_family_size_num == "4", na.rm = TRUE) #13 = 12,3%
sum(data_GN1$ideal_family_size_num == "5", na.rm = TRUE) #0 = 12,3%
sum(data_GN1$ideal_family_size_num == "6", na.rm = TRUE) #0 = 15,1%
sum(data_GN1$ideal_family_size_num == "7", na.rm = TRUE) #0 = 5,7%
sum(data_GN1$ideal_family_size_num == "8", na.rm = TRUE) #0 = 0,9%
```

```
# Netwerktipe 4: N = 128 GN2
sum(data_GN2$partner_num == "0") #28 = 21,9%
sum(data_GN2$partner_num == "1") #100 = 78,1%
```

```

sum(data_GN2$has_children_num == "0")
#70 = 54,7%
sum(data_GN2$has_children_num == "1")
#58 = 45,3%

sum(data_GN2$num_children == "0") #70 =
56,7%
sum(data_GN2$num_children == "1") #21 =
16,4%
sum(data_GN2$num_children == "2") #25 =
19,5%
sum(data_GN2$num_children == "3") #7 =
5,4%
sum(data_GN2$num_children == "4") #3 =
2,3%
sum(data_GN2$num_children == "5") #1 =
0,8%
sum(data_GN2$num_children == "6") #0 =
0%
sum(data_GN2$num_children == "7") #0 =
0%
sum(data_GN2$num_children == "8") #1 =
0,8%
summary(data_GN2$num_children)
sd(data_GN2$num_children)

sum(data_GN2$num_children_biol == "0",
na.rm = TRUE) #0 = 0%
sum(data_GN2$num_children_biol == "1",
na.rm = TRUE) #22 = 17,2%
sum(data_GN2$num_children_biol == "2",
na.rm = TRUE) #24 = 18,8%
sum(data_GN2$num_children_biol == "3",
na.rm = TRUE) #8 = 6,3%
sum(data_GN2$num_children_biol == "4",
na.rm = TRUE) #3 = 2,3%
sum(data_GN2$num_children_biol == "5",
na.rm = TRUE) #1 = 0,8%
sum(data_GN2$num_children_biol == "6",
na.rm = TRUE) #0 = 0%
sum(data_GN2$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_GN2$num_children_biol == "8",
na.rm = TRUE) #0 = 0,%
summary(data_GN2$num_children_biol,
na.rm = TRUE)
sd(data_GN2$num_children_biol, na.rm =
TRUE)

sum(data_GN2$childwish_num == "1") #20 =
15,6%

```

```

sum(data_GN2$childwish_num == "2") #23 =
18,0%
sum(data_GN2$childwish_num == "3") #22 =
17,2%
sum(data_GN2$childwish_num == "4") #33 =
25,8%
sum(data_GN2$childwish_num == "5") #30 =
23,4%
summary(data_GN2$childwish_num)
sd(data_GN2$childwish_num)

sum(data_GN2$ideal_family_size_num ==
"0", na.rm = TRUE) #7 = 5,5%
sum(data_GN2$ideal_family_size_num ==
"1", na.rm = TRUE) #7 = 5,5%
sum(data_GN2$ideal_family_size_num ==
"2", na.rm = TRUE) #57 = 44,5%
sum(data_GN2$ideal_family_size_num ==
"3", na.rm = TRUE) #30 = 23,4%
sum(data_GN2$ideal_family_size_num ==
"4", na.rm = TRUE) #8 = 6,3%
sum(data_GN2$ideal_family_size_num ==
"5", na.rm = TRUE) #3 = 2,3%
sum(data_GN2$ideal_family_size_num ==
"6", na.rm = TRUE) #0 = 0%
sum(data_GN2$ideal_family_size_num ==
"7", na.rm = TRUE) #0 = 0%
sum(data_GN2$ideal_family_size_num ==
"8", na.rm = TRUE) #1 = 0,8%

hist(data_GN2$partner_num, breaks = 2, main
= "Gender partner cluster 2", xlim = c(0,1),
xlab = "Gender")
hist(data_GN2$has_children_num, breaks = 2,
main = "Has Children cluster 2", xlim = c(0,1),
xlab = "Has Children")
hist(data_GN2$num_children, breaks = 5,
main = "Number of Children cluster 2", xlab =
"Children")
hist(data_GN2$num_children_biol, breaks =
5, main = "Number of Biological Children
cluster 2", xlab = "Children")
hist(data_GN2$childwish_num, breaks = 5,
main = "Childwish cluster 2", xlab =
"Childwish")

# N = 117 GN3 ~ netwerktype 3
sum(data_GN3$partner_num == "0") #35 =
29,9%
sum(data_GN3$partner_num == "1") #82 =
70,1%

```

```

sum(data_GN3$has_children_num == "0")
#87 = 74,3%
sum(data_GN3$has_children_num == "1")
#30 = 25,7%

sum(data_GN3$num_children == "0") #87 =
74,4%
sum(data_GN3$num_children == "1") #12 =
10,3%
sum(data_GN3$num_children == "2") #15 =
12,8%
sum(data_GN3$num_children == "3") #3 =
2,6%
sum(data_GN3$num_children == "4") #0 =
0%
sum(data_GN3$num_children == "5") #0 =
0%
sum(data_GN3$num_children == "6") #0 =
0%
sum(data_GN3$num_children == "7") #0 =
0%
sum(data_GN3$num_children == "8") #0 =
0%
summary(data_GN3$num_children)
sd(data_GN3$num_children)

sum(data_GN3$num_children_biol == "0",
na.rm = TRUE) #4 = 3,4%
sum(data_GN3$num_children_biol == "1",
na.rm = TRUE) #11 = 9,4%
sum(data_GN3$num_children_biol == "2",
na.rm = TRUE) #12 = 10,3%
sum(data_GN3$num_children_biol == "3",
na.rm = TRUE) #3 = 2,6%
sum(data_GN3$num_children_biol == "4",
na.rm = TRUE) #0 = 0%
sum(data_GN3$num_children_biol == "5",
na.rm = TRUE) #0 = 0%
sum(data_GN3$num_children_biol == "6",
na.rm = TRUE) #0 = 0%
sum(data_GN3$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_GN3$num_children_biol == "8",
na.rm = TRUE) #0 = 0%
summary(data_GN3$num_children_biol,
na.rm = TRUE)
sd(data_GN3$num_children_biol, na.rm =
TRUE)

sum(data_GN3$childwish_num == "1") #16 =
13,7%
sum(data_GN3$childwish_num == "2") #12 =
10,3%

```

```

sum(data_GN3$childwish_num == "3") #14 =
12,0%%
sum(data_GN3$childwish_num == "4") #47 =
40,2%
sum(data_GN3$childwish_num == "5") #28 =
23,9%
summary(data_GN3$childwish_num)
sd(data_GN3$childwish_num)

sum(data_GN3$ideal_family_size_num ==
"0", na.rm = TRUE) #10 = 8,5%
sum(data_GN3$ideal_family_size_num ==
"1", na.rm = TRUE) #4 = 3,4%
sum(data_GN3$ideal_family_size_num ==
"2", na.rm = TRUE) #61 = 52,1%
sum(data_GN3$ideal_family_size_num ==
"3", na.rm = TRUE) #21 = 17,9%
sum(data_GN3$ideal_family_size_num ==
"4", na.rm = TRUE) #8 = 6,8%
sum(data_GN3$ideal_family_size_num ==
"5", na.rm = TRUE) #1 = 0,8%
sum(data_GN3$ideal_family_size_num ==
"6", na.rm = TRUE) #0 = 0%
sum(data_GN3$ideal_family_size_num ==
"7", na.rm = TRUE) #0 = 0%
sum(data_GN3$ideal_family_size_num ==
"8", na.rm = TRUE) #0 = 0%

hist(data_GN3$partner_num, breaks = 2, main
= "Gender partner cluster 3", xlim = c(0,1),
xlab = "Gender")
hist(data_GN3$has_children_num, breaks = 2,
main = "Has Children cluster 3", xlim = c(0,1),
xlab = "Has Children")
hist(data_GN3$num_children, breaks = 5,
main = "Number of Children cluster 3", xlab =
"Children")
hist(data_GN3$num_children_biol, breaks =
5, main = "Number of Biological Children
cluster 3", xlab = "Children")
hist(data_GN3$childwish_num, breaks = 5,
main = "Childwish cluster 3", xlab =
"Childwish")

# N = 88 GN4 ~ netwerktype 5
sum(data_GN4$partner_num == "0") #34 =
38,6%
sum(data_GN4$partner_num == "1") #54 =
61,4%

sum(data_GN4$has_children_num == "0")
#56 = 63,6%

```

```

sum(data_GN4$has_children_num == "1")
#32 = 36,4%

sum(data_GN4$num_children == "0") #56 =
63,6%
sum(data_GN4$num_children == "1") #15 =
17,0%
sum(data_GN4$num_children == "2") #11 =
12,5%
sum(data_GN4$num_children == "3") #5 =
5,7%
sum(data_GN4$num_children == "4") #1 =
0,9%
sum(data_GN4$num_children == "5") #0 =
0%
sum(data_GN4$num_children == "6") #1 =
1,1%
sum(data_GN4$num_children == "7") #0 =
0%
sum(data_GN4$num_children == "8") #0 =
0%
summary(data_GN4$num_children)
sd(data_GN4$num_children)

sum(data_GN4$num_children_biol == "0",
na.rm = TRUE) #0 = 0%
sum(data_GN4$num_children_biol == "1",
na.rm = TRUE) #15 = 17,0%
sum(data_GN4$num_children_biol == "2",
na.rm = TRUE) #12 = 13,6%
sum(data_GN4$num_children_biol == "3",
na.rm = TRUE) #4 = 4,5%
sum(data_GN4$num_children_biol == "4",
na.rm = TRUE) #0 = 0%
sum(data_GN4$num_children_biol == "5",
na.rm = TRUE) #0 = 0%
sum(data_GN4$num_children_biol == "6",
na.rm = TRUE) #1 = 1,1%
sum(data_GN4$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_GN4$num_children_biol == "8",
na.rm = TRUE) #0 = 0%
summary(data_GN4$num_children_biol,
na.rm = TRUE)
sd(data_GN4$num_children_biol, na.rm =
TRUE)

sum(data_GN4$childwish_num == "1") #11 =
12,5%
sum(data_GN4$childwish_num == "2") #12 =
13,6%
sum(data_GN4$childwish_num == "3") #15 =
17,0%

```

```

sum(data_GN4$childwish_num == "4") #28 =
31,8%
sum(data_GN4$childwish_num == "5") #22 =
25,0%
summary(data_GN4$childwish_num)
sd(data_GN4$childwish_num)

sum(data_GN4$ideal_family_size_num ==
"0", na.rm = TRUE) #4 = 4,5%
sum(data_GN4$ideal_family_size_num ==
"1", na.rm = TRUE) #8 = 9,1%
sum(data_GN4$ideal_family_size_num ==
"2", na.rm = TRUE) #35 = 39,8%
sum(data_GN4$ideal_family_size_num ==
"3", na.rm = TRUE) #19 = 21,6%
sum(data_GN4$ideal_family_size_num ==
"4", na.rm = TRUE) #7 = 8,0%
sum(data_GN4$ideal_family_size_num ==
"5", na.rm = TRUE) #3 = 3,4%
sum(data_GN4$ideal_family_size_num ==
"6", na.rm = TRUE) #1 = 1,1%
sum(data_GN4$ideal_family_size_num ==
"7", na.rm = TRUE) #0 = 0%
sum(data_GN4$ideal_family_size_num ==
"8", na.rm = TRUE) #0 = 0%

hist(data_GN4$partner_num, breaks = 2, main
= "Gender partner cluster 4", xlim = c(0,1),
xlab = "Gender")
hist(data_GN4$has_children_num, breaks = 2,
main = "Has Children cluster 4", xlim = c(0,1),
xlab = "Has Children")
hist(data_GN4$num_children, breaks = 5,
main = "Number of Children cluster 4", xlab =
"Children")
hist(data_GN4$num_children_biol, breaks =
5, main = "Number of Biological Children
cluster 4", xlab = "Children")
hist(data_GN4$childwish_num, breaks = 5,
main = "Childwish cluster 4", xlab =
"Childwish")

# N = 154 GN5 ~ netwerktype 6
sum(data_GN5$partner_num == "0") #39 =
25,3%
sum(data_GN5$partner_num == "1") #115 =
74,7%

sum(data_GN5$has_children_num == "0")
#106 = 68,8%
sum(data_GN5$has_children_num == "1")
#48 = 31,2%

```

```

sum(data_GN5$num_children == "0") #106 =
68,8%
sum(data_GN5$num_children == "1") #23 =
14,9%
sum(data_GN5$num_children == "2") #18 =
11,7%
sum(data_GN5$num_children == "3") #5 =
3,2%
sum(data_GN5$num_children == "4") #2 =
1,3%
sum(data_GN5$num_children == "5") #0 =
0%
sum(data_GN5$num_children == "6") #0 =
0%
sum(data_GN5$num_children == "7") #0 =
0%
sum(data_GN5$num_children == "8") #0 =
0%
summary(data_GN5$num_children)
sd(data_GN5$num_children)

```

```

sum(data_GN5$num_children_biol == "0",
na.rm = TRUE) #2 = 1,3%
sum(data_GN5$num_children_biol == "1",
na.rm = TRUE) #23 = 14,9%
sum(data_GN5$num_children_biol == "2",
na.rm = TRUE) #17 = 11,0%
sum(data_GN5$num_children_biol == "3",
na.rm = TRUE) #4 = 2,6%
sum(data_GN5$num_children_biol == "4",
na.rm = TRUE) #2 = 1,3%
sum(data_GN5$num_children_biol == "5",
na.rm = TRUE) #0 = 0%
sum(data_GN5$num_children_biol == "6",
na.rm = TRUE) #0 = 0%
sum(data_GN5$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_GN5$num_children_biol == "8",
na.rm = TRUE) #0 = 0%
summary(data_GN5$num_children_biol,
na.rm = TRUE)
sd(data_GN5$num_children_biol, na.rm =
TRUE)

```

```

sum(data_GN5$childwish_num == "1") #15 =
9,7%
sum(data_GN5$childwish_num == "2") #18 =
11,7%
sum(data_GN5$childwish_num == "3") #21 =
13,6%
sum(data_GN5$childwish_num == "4") #56 =
36,4%

```

```

sum(data_GN5$childwish_num == "5") #44 =
28,6%
summary(data_GN5$childwish_num)
sd(data_GN5$childwish_num)

```

```

hist(data_GN5$partner_num, breaks = 2, main
= "Gender partner cluster 5", xlim = c(0,1),
xlab = "Gender")
hist(data_GN5$has_children_num, breaks = 2,
main = "Has Children cluster 5", xlim = c(0,1),
xlab = "Has Children")
hist(data_GN5$num_children, breaks = 5,
main = "Number of Children cluster 5", xlab =
"Children")
hist(data_GN5$num_children_biol, breaks =
5, main = "Number of Biological Children
cluster 5", xlab = "Children")
hist(data_GN5$childwish_num, breaks = 5,
main = "Childwish cluster 5", xlab =
"Childwish")

```

```

# N = 113 GN6 ~ netwerktype 2
sum(data_GN6$partner_num == "0") #32 =
28,3%
sum(data_GN6$partner_num == "1") #81 =
71,7%

```

```

sum(data_GN6$has_children_num == "0")
#71 = 62,8%
sum(data_GN6$has_children_num == "1")
#42 = 37,2%

```

```

sum(data_GN6$num_children == "0") #71 =
62,8%
sum(data_GN6$num_children == "1") #13 =
11,5%
sum(data_GN6$num_children == "2") #22 =
19,5%
sum(data_GN6$num_children == "3") #6 =
5,3%
sum(data_GN6$num_children == "4") #1 =
0,9%
sum(data_GN6$num_children == "5") #0 =
0%
sum(data_GN6$num_children == "6") #0 =
0%
sum(data_GN6$num_children == "7") #0 =
0%
sum(data_GN6$num_children == "8") #0 =
0%
summary(data_GN6$num_children)
sd(data_GN6$num_children)

```



```

sum(data_GN6$num_children_biol == "0",
na.rm = TRUE) #4 = 3,5%
sum(data_GN6$num_children_biol == "1",
na.rm = TRUE) #13 = 11,5%
sum(data_GN6$num_children_biol == "2",
na.rm = TRUE) #18 = 15,9%
sum(data_GN6$num_children_biol == "3",
na.rm = TRUE) #6 = 5,3%
sum(data_GN6$num_children_biol == "4",
na.rm = TRUE) #1 = 0,9%
sum(data_GN6$num_children_biol == "5",
na.rm = TRUE) #0 = 0%
sum(data_GN6$num_children_biol == "6",
na.rm = TRUE) #0 = 0%
sum(data_GN6$num_children_biol == "7",
na.rm = TRUE) #0 = 0%
sum(data_GN6$num_children_biol == "8",
na.rm = TRUE) #0 = 0%
summary(data_GN6$num_children_biol,
na.rm = TRUE)
sd(data_GN6$num_children_biol, na.rm =
TRUE)

sum(data_GN6$childwish_num == "1") #17 =
15,0%
sum(data_GN6$childwish_num == "2") #18 =
15,9%
sum(data_GN6$childwish_num == "3") #16 =
14,2%
sum(data_GN6$childwish_num == "4") #27 =
23,9%
sum(data_GN6$childwish_num == "5") #35 =
31,0%
summary(data_GN6$childwish_num)
sd(data_GN6$childwish_num)

hist(data_GN6$partner_num, breaks = 2, main
= "Gender partner cluster 6", xlim = c(0,1),
xlab = "Gender")
hist(data_GN6$has_children_num, breaks = 2,
main = "Has Children cluster 6", xlim = c(0,1),
xlab = "Has Children")
hist(data_GN6$num_children, breaks = 5,
main = "Number of Children cluster 6", xlab =
"Children")
hist(data_GN6$num_children_biol, breaks =
5, main = "Number of Biological Children
cluster 6", xlab = "Children")
hist(data_GN6$childwish_num, breaks = 5,
main = "Childwish cluster 6", xlab =
"Childwish")

```

```

#####
#
## Assumpties toetsen Bijlage 3
#leeftijd
qqnorm(data_GN1$age)
qqline(data_GN1$age)
qqnorm(data_GN2$age)
qqline(data_GN2$age)
qqnorm(data_GN3$age)
qqline(data_GN3$age)
qqnorm(data_GN4$age)
qqline(data_GN4$age)
qqnorm(data_GN5$age)
qqline(data_GN5$age)
qqnorm(data_GN6$age)
qqline(data_GN6$age)

leveneTest(data_sel_GN$age,
data_sel_GN$num_clusters)

#kinderwens
qqnorm(data_GN1$rel_closefamily)
qqline(data_GN1$rel_closefamily)
qqnorm(data_GN2$rel_closefamily)
qqline(data_GN2$rel_closefamily)
qqnorm(data_GN3$rel_closefamily)
qqline(data_GN3$rel_closefamily)
qqnorm(data_GN4$rel_closefamily)
qqline(data_GN4$rel_closefamily)
qqnorm(data_GN5$rel_closefamily)
qqline(data_GN5$rel_closefamily)
qqnorm(data_GN6$rel_closefamily)
qqline(data_GN6$rel_closefamily)
leveneTest(data_sel_GN$rel_closefamily,
data_sel_GN$num_clusters)

qqnorm(data_GN1$rel_otherfamily)
qqline(data_GN1$rel_otherfamily)
qqnorm(data_GN2$rel_otherfamily)
qqline(data_GN2$rel_otherfamily)
qqnorm(data_GN3$rel_otherfamily)
qqline(data_GN3$rel_otherfamily)
qqnorm(data_GN4$rel_otherfamily)
qqline(data_GN4$rel_otherfamily)
qqnorm(data_GN5$rel_otherfamily)
qqline(data_GN5$rel_otherfamily)
qqnorm(data_GN6$rel_otherfamily)
qqline(data_GN6$rel_otherfamily)
leveneTest(data_sel_GN$rel_otherfamily,
data_sel_GN$num_clusters)

qqnorm(data_GN1$rel_friendsacquaintance)
qqline(data_GN1$rel_friendsacquaintance)

```

```
qqnorm(data_GN2$rel_friendsacquaintance)
qqline(data_GN2$rel_friendsacquaintance)
qqnorm(data_GN3$rel_friendsacquaintance)
qqline(data_GN3$rel_friendsacquaintance)
qqnorm(data_GN4$rel_friendsacquaintance)
qqline(data_GN4$rel_friendsacquaintance)
qqnorm(data_GN5$rel_friendsacquaintance)
qqline(data_GN5$rel_friendsacquaintance)
qqnorm(data_GN6$rel_friendsacquaintance)
qqline(data_GN6$rel_friendsacquaintance)
leveneTest(data_sel_GN$rel_friendsacquaintance, data_sel_GN$num_clusters)
```

```
qqnorm(data_GN1$rel_colleagues)
qqline(data_GN1$rel_colleagues)
qqnorm(data_GN2$rel_colleagues)
qqline(data_GN2$rel_colleagues)
qqnorm(data_GN3$rel_colleagues)
qqline(data_GN3$rel_colleagues)
qqnorm(data_GN4$rel_colleagues)
qqline(data_GN4$rel_colleagues)
qqnorm(data_GN5$rel_colleagues)
qqline(data_GN5$rel_colleagues)
qqnorm(data_GN6$rel_colleagues)
qqline(data_GN6$rel_colleagues)
leveneTest(data_sel_GN$rel_colleagues, data_sel_GN$num_clusters)
```

significante t-toetsen (4.1.3: beschrijvende statistieken kenmerken vrouwen)

```
# leeftijd
ANOVA_age <- aov(data_sel_GN$age ~ factor(data_sel_GN$cluster))
summary(ANOVA_age)
t.test(data_GN2$age, data_GN3$age, paired=FALSE) # t = 2,40, df = 234,6, p = 0,02
t.test(data_GN3$age, data_GN4$age, paired=FALSE) # t = -3,15, df = 197,6, p = <0,01
t.test(data_GN3$age, data_GN6$age, paired=FALSE) # t = -2,82, df = 227, p = <0,01
```

```
# aantal kinderen
KW_Stat3 <- kruskal.test(num_children ~ clusters, data = data_sel_GN)
KW_Stat3 <- kruskal.test(num_children ~ clusters, data = data_sel_GN)$statistic
Eta_squared3 <- KW_Stat3 / (nrow(data_sel_GN) - 1)
paste("Eta squared is", Eta_squared3)
```

```
pairwise.wilcox.test(data_sel_GN$num_children, data_sel_GN$clusters, p.adjust.method = "bonferroni")
```

```
# aantal biologische kinderen
KW_Stat4 <- kruskal.test(num_children_biol ~ clusters, data = data_sel_GN)
KW_Stat4 <- kruskal.test(num_children_biol ~ clusters, data = data_sel_GN)$statistic
Eta_squared4 <- KW_Stat4 / (nrow(data_sel_GN) - 1)
paste("Eta squared is", Eta_squared4)
```

```
pairwise.wilcox.test(data_sel_GN$num_children_biol, data_sel_GN$clusters, p.adjust.method = "bonferroni")
```

```
# Proportietoets voor partner en kind: voor elke groep opnieuw ingevoerd.
```

```
Chi_partner <-
chisq.test(data_sel_GN$partner_num, data_sel_GN$clusters)
Chi_partner$residuals
```

```
Chi_kind <-
chisq.test(data_sel_GN$has_children_num, data_sel_GN$clusters)
Chi_kind$residuals
```

```
prop.test(c(48, 58), c(154, 128), correct = FALSE)
```

```
prop.test(c(100, 54), c(128, 88), correct = FALSE)
```

Kruskal Wallis toets = Verschil in kinderwens + Tabel 5

```
KW_Stat <- kruskal.test(childwish ~ clusters, data = data_sel_GN)
KW_Stat <- kruskal.test(childwish ~ clusters, data = data_sel_GN)$statistic
Eta_squared <- KW_Stat / (nrow(data_sel_GN) - 1)
paste("Eta squared is", Eta_squared)
```

```
pairwise.wilcox.test(data_sel_GN$childwish_num, data_sel_GN$clusters, p.adjust.method = "bonferroni")
```

```

kruistabel_ideal_family <-
table(data_sel_GN$ideal_family_size,
data_sel_GN$netwerktypen)
print(kruistabel_ideal_family)
prop.table(kruistabel_ideal_family, 2)
kruskal.test(ideal_family_size_num ~
netwerktypen, data = data_sel_GN)

KW_Stat2<-
kruskal.test(ideal_family_size_num ~
netwerktypen, data = data_sel_GN)$statistic
Eta_squared2 <- KW_Stat /
(nrow(data_sel_GN) - 1)
paste("Eta squared is",Eta_squared2)

pairwise.wilcox.test(data_sel_GN$ideal_famil
y_size_num, data_sel_GN$netwerk_naam,
p.adjust.method = "bonferroni")

### STAP 2 Vernieuwd: Toevoegen van
cluster kenmerken relaties
data_sel_GN2 <- data_sel_GN |>
  select(age, partner, childwish_num,
ideal_family_size_num, has_children,
num_children, num_children_biol, childwish,
  ideal_family_size, alter_attr, edgelist,
tidygraph, has_children_num,
  relationEA2, comm_1or2,
comm_3orhigher, modularity, clusters,
partner_num, relationEA2_num)

data_sel_GN2 <- data_sel_GN2 |>
mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,
TRUE ~ 0)
)),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,
TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,
TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",

```

```

"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",
"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))

data_sel_GN2 <- data_sel_GN2 |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

## Nieuwe gestandaardiseerde dataset voor k-
medoid analyse 2.
data_clus_stan2 <- data_sel_GN2 |>
  select(modularity, comm_1or2,
comm_3orhigher, rel_closefamily,
rel_otherfamily, rel_colleagues)

data_clus_stan2 <- scale(data_clus_stan2)

## Elleboog methode: Within Cluster Sum of
Squares (AIC)
fviz_nbclust(data_clus_stan2, pam, k=20,
method = "wss") #Elbow method, Within
Cluster Sum of Squares(AIC)

## Silhouette: Optimal number of clusters
fviz_nbclust(data_clus_stan2, pam, k=20,
method = "silhouette")

## Optimaal aantal clusters: 2, Vacca uses
euclidean
pamC2 <-pam(data_clus_stan2, 2, metric =
"euclidean", stand = FALSE)

#Medoids info,clustering vector, objective
function, available components
pamC2
pamC2$silinfo
pamC2$clustering

```

```

pamC2$medoids
pamC2$id.med #cases: 124, 480
pamC2$objective
pamC2$clusinfo

## Cluster plot
fviz_cluster(pamC2)
fviz_cluster(pamC2, geom = "point",
show.clust.cent = TRUE)

## Visualisatie van clusters
ggraph(data_sel_GN2$tidygraph[[124]],
layout = "kk") +
  geom_edge_link(colour = "black") +
  geom_node_point(colour = "grey") +
  labs(colour = NULL) +
  theme_graph()

g10 <- data_sel_GN2$tidygraph[[124]]
plot(g10)
betw10 <- edge.betweenness(g10)
ebc10 <- edge.betweenness.community(g10)
plot(ebc10, g10)

ggraph(data_sel_GN2$tidygraph[[480]], layout
= "kk") +
  geom_edge_link(colour = "Black") +
  geom_node_point(colour = "grey") +
  labs(colour = NULL) +
  theme_graph()

g11 <- data_sel_GN2$tidygraph[[480]]
plot(g11)
betw11 <- edge.betweenness(g11)
ebc11 <- edge.betweenness.community(g11)
plot(ebc11, g11)

## Dataset per cluster
data_sel_GN2 <- data_sel_GN2 |>
  mutate(
    clusters = pamC2$clustering)

data_GN7 <- data_sel_GN2 |>
  filter(clusters == 1)

data_GN8 <- data_sel_GN2 |>
  filter(clusters == 2)

## Kruskal Willes toets
kruistabel_childwish2 <-
table(data_sel_GN2$childwish,
data_sel_GN2$clusters)

```

```

Volgorde <- c("absolutely not", "probably
not", "I don't know", "probably so",
"Absolutely so")
print(kruistabel_childwish2)
prop.table(kruistabel_childwish2, 2)
kruskal.test(childwish ~ clusters, data =
data_sel_GN2)

KW_Stat3 <- kruskal.test(childwish ~ clusters,
data = data_sel_GN2)$statistic
Eta_squared <- KW_Stat3 /
(nrow(data_sel_GN2) - 1)
paste("Eta squared is",Eta_squared)

pairwise.wilcox.test(data_sel_GN$childwish_
num, data_sel_GN$netwerk_naam,
p.adjust.method = "bonferroni")

kruistabel_ideal_family2 <-
table(data_sel_GN2$ideal_family_size,
data_sel_GN2$clusters)
print(kruistabel_ideal_family2)
prop.table(kruistabel_ideal_family2, 2)
kruskal.test(ideal_family_size_num ~ clusters,
data = data_sel_GN2)

KW_Stat2<-
kruskal.test(ideal_family_size_num ~
netwerktypen, data = data_sel_GN)$statistic
Eta_squared2 <- KW_Stat /
(nrow(data_sel_GN) - 1)
paste("Eta squared is",Eta_squared2)

pairwise.wilcox.test(data_sel_GN$ideal_famil
y_size_num, data_sel_GN$netwerk_naam,
p.adjust.method = "bonferroni")

## Boxplot, histogram & barplot voor
visualisatie en bekijken verdeling.
## visualisatie van beschrijvende
statistieken
boxplot(data_sel_GN$sage ~
data_sel_GN$netwerktypen,
  main = "Boxplots van Leeftijd per
Cluster",
  xlab = "Cluster",
  ylab = "Leeftijd")

barplot(table(data_sel_GN$partner), main =
"Staaftiagram van Partner")
barplot(table(data_sel_GN$partner,
data_sel_GN$netwerktypen),

```

```

    main = "Staafdiagram van Partner per
Cluster",
    xlab = "Partner",
    ylab = "Aantal",
    legend = TRUE)

barplot(table(data_sel_GN$has_children),
main = "Staafdiagram van Kinderen")
barplot(table(data_sel_GN$has_children,
data_sel_GN$netwerktypen),
    main = "Staafdiagram van Kinderen per
Cluster",
    xlab = "Kinderen",
    ylab = "Aantal",
    legend = TRUE)

hist(data_sel_GN$num_children,
    breaks = 8,
    main = "Histogram van Aantal Kinderen",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN1$num_children,
    breaks = 4,
    main = "Histogram aantal Kinderen,
netwerktype 1",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN2$num_children,
    breaks = 8,
    main = "Histogram aantal Kinderen,
netwerktype 4",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN3$num_children,
    breaks = 3,
    main = "Histogram aantal Kinderen,
netwerktype 3",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN4$num_children,
    breaks = 6,
    main = "Histogram aantal Kinderen,
netwerktype 5",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN5$num_children,
    breaks = 6,

```

```

    main = "Histogram aantal Kinderen,
netwerktype 6",
    xlab = "Kinderen",
    ylab = "Frequentie")

hist(data_GN6$num_children,
    breaks = 6,
    main = "Histogram aantal Kinderen,
netwerktype 2",
    xlab = "Kinderen",
    ylab = "Frequentie")

barplot(table(data_sel_GN$schildwish),
    main = "Staafdiagram van Kinderwens",
    xlab = "In de toekomst nog kinderen",
    ylab = "Aantal",
    col = c("dark blue"),
    legend = FALSE)

barplot(table(data_sel_GN$schildwish,
data_sel_GN$netwerktypen),
    main = "Staafdiagram van Kinderwens",
    xlab = "Netwerktype",
    ylab = "Aantal",
    col = c("yellow", "light green", "light
blue", "red", "pink"),
    legend = TRUE)

barplot(table(data_sel_GN$schildwish,
data_sel_GN$netwerk_naam),
    main = "Staafdiagram van Kinderwens",
    xlab = "Netwerktype",
    ylab = "Aantal",
    col = c("yellow", "light green", "light
blue", "red", "pink"),
    legend = TRUE)

barplot(table(data_sel_GN$ideal_family_size_
num),
    main = "Staafdiagram Ideale
FamiliEGrootte",
    xlab = "Aantal kinderen",
    ylab = "Aantal",
    col = c("dark blue"),
    legend = FALSE)

barplot(table(data_GN1$ideal_family_size_nu
m),
    main = "Staafdiagram Ideale
FamiliEGrootte, netwerktype 1",
    xlab = "Aantal kinderen",
    ylab = "Aantal vrouwen",
    col = c("dark blue"),

```

```

legend = FALSE)

barplot(table(data_GN2$ideal_family_size_num),
  main = "Staafdiagram Ideale
FamiliEGrooTte, netwerktype 4",
  xlab = "Aantal kinderen",
  ylab = "Aantal vrouwen",
  col = c("dark blue"),
  legend = FALSE)

barplot(table(data_GN3$ideal_family_size_num),
  main = "Staafdiagram Ideale
FamiliEGrooTte, netwerktype 3",
  xlab = "Aantal kinderen",
  ylab = "Aantal vrouwen",
  col = c("dark blue"),
  legend = FALSE)

barplot(table(data_GN4$ideal_family_size_num),
  main = "Staafdiagram Ideale
FamiliEGrooTte, netwerktype 5",
  xlab = "Aantal kinderen",
  ylab = "Aantal vrouwen",
  col = c("dark blue"),
  legend = FALSE)

barplot(table(data_GN5$ideal_family_size_num),
  main = "Staafdiagram Ideale
FamiliEGrooTte, netwerktype 6",
  xlab = "Aantal kinderen",
  ylab = "Aantal vrouwen",
  col = c("dark blue"),
  legend = FALSE)

barplot(table(data_GN6$ideal_family_size_num),
  main = "Staafdiagram Ideale
FamiliEGrooTte, netwerktype 2",
  xlab = "Aantal kinderen",
  ylab = "Aantal vrouwen",
  col = c("dark blue"),
  legend = FALSE)

barplot(table(data_sel_GN$childwish,
data_sel_GN$netwerktypen),
  main = "Staafdiagram van Kinderwens",
  xlab = "Netwerktype",
  ylab = "Aantal",

```

```

col = c("yellow", "green", "blue", "red",
"pink"),
  legend = FALSE)

boxplot(data_sel_GN$ideal_family_size_num)
boxplot(data_sel_GN$ideal_family_size_num
~ data_sel_GN$netwerktypen,
  main = "Boxplot Ideale Gezinsgrootte",
  xlab = "Netwerktype",
  ylab = "Aantal kinderen")

##Beschrijvende statistieken voor bijlage
mean(data_sel_GN2$rel_closefamily)

summary(data_GN7$ideal_family_size_num,
na.rm = TRUE)
summary(data_GN8$ideal_family_size_num,
na.rm = TRUE)

sum(data_GN7$partner_num == "0") #129
sum(data_GN7$partner_num == "1") #360
hist(data_GN7$partner_num, breaks = 2, main
= "Gender partner", xlim = c(0,1), xlab =
"Gender")

sum(data_GN7$has_children_num == "0")
#66 = 62,3%
sum(data_GN7$has_children_num == "1")
#40 = 37,7%
hist(data_GN7$has_children_num, breaks = 2,
main = "Has Children", xlim = c(0,1), xlab =
"Has Children")

hist(data_GN7$num_children, breaks = 8,
main = "Number of Children", xlab =
"Children")
summary(data_GN7$num_children)
sd(data_GN7$num_children)

hist(data_GN7$num_children_biol, breaks =
5, main = "Number of Biological Children",
xlab = "Children")
summary(data_GN7$num_children_biol,
na.rm = TRUE)
sd(data_GN7$num_children_biol, na.rm =
TRUE)

sum(data_GN7$childwish_num == "1")
sum(data_GN7$childwish_num == "2")
sum(data_GN7$childwish_num == "3")
sum(data_GN7$childwish_num == "4")
sum(data_GN7$childwish_num == "5")

```

```
hist(data_GN7$childwish_num, breaks = 6,
main = "Childwish", xlab = "Childwish")
summary(data_GN7$childwish_num)
sd(data_GN7$childwish_num)
```

```
sum(data_GN7$ideal_family_size_num ==
"0", na.rm = TRUE) #4 = 3,8%
sum(data_GN7$ideal_family_size_num ==
"1", na.rm = TRUE) #4 = 3,8%
sum(data_GN7$ideal_family_size_num ==
"2", na.rm = TRUE) #47 = 44,3%
sum(data_GN7$ideal_family_size_num ==
"3", na.rm = TRUE) #29 = 27,4%
sum(data_GN7$ideal_family_size_num ==
"4", na.rm = TRUE) #13 = 12,3%
sum(data_GN7$ideal_family_size_num ==
"5", na.rm = TRUE) #0 = 12,3%
sum(data_GN7$ideal_family_size_num ==
"6", na.rm = TRUE) #0 = 15,1%
sum(data_GN7$ideal_family_size_num ==
"7", na.rm = TRUE) #0 = 5,7%
sum(data_GN7$ideal_family_size_num ==
"8", na.rm = TRUE) #0 = 0,9%
hist(data_GN7$ideal_family_size_num, breaks
= 8, main = "Ideale gezinsgrootte", xlab =
"Ideale gezinsgrootte")
summary(data_GN7$ideal_family_size_num)
sd(data_GN7$ideal_family_size_num)
```

```
sum(data_GN8$partner_num == "0") #129
sum(data_GN8$partner_num == "1") #360
hist(data_GN8$partner_num, breaks = 2, main
= "Gender partner", xlim = c(0,1), xlab =
"Gender")
```

```
sum(data_GN8$has_children_num == "0")
#66 = 62,3%
sum(data_GN8$has_children_num == "1")
#40 = 37,7%
hist(data_GN8$has_children_num, breaks = 2,
main = "Has Children", xlim = c(0,1), xlab =
"Has Children")
```

```
hist(data_GN8$num_children, breaks = 8,
main = "Number of Children", xlab =
"Children")
summary(data_GN8$num_children)
sd(data_GN8$num_children)
```

```
hist(data_GN8$num_children_biol, breaks =
5, main = "Number of Biological Children",
xlab = "Children")
```

```
summary(data_GN8$num_children_biol,
na.rm = TRUE)
sd(data_GN8$num_children_biol, na.rm =
TRUE)
```

```
sum(data_GN8$childwish_num == "1")
sum(data_GN8$childwish_num == "2")
sum(data_GN8$childwish_num == "3")
sum(data_GN8$childwish_num == "4")
sum(data_GN8$childwish_num == "5")
hist(data_GN8$childwish_num, breaks = 6,
main = "Childwish", xlab = "Childwish")
summary(data_GN8$childwish_num)
sd(data_GN8$childwish_num)
```

```
sum(data_GN8$ideal_family_size_num ==
"0", na.rm = TRUE) #4 = 3,8%
sum(data_GN8$ideal_family_size_num ==
"1", na.rm = TRUE) #4 = 3,8%
sum(data_GN8$ideal_family_size_num ==
"2", na.rm = TRUE) #47 = 44,3%
sum(data_GN8$ideal_family_size_num ==
"3", na.rm = TRUE) #29 = 27,4%
sum(data_GN8$ideal_family_size_num ==
"4", na.rm = TRUE) #13 = 12,3%
sum(data_GN8$ideal_family_size_num ==
"5", na.rm = TRUE) #0 = 12,3%
sum(data_GN8$ideal_family_size_num ==
"6", na.rm = TRUE) #0 = 15,1%
sum(data_GN8$ideal_family_size_num ==
"7", na.rm = TRUE) #0 = 5,7%
sum(data_GN8$ideal_family_size_num ==
"8", na.rm = TRUE) #0 = 0,9%
hist(data_GN8$ideal_family_size_num, breaks
= 8, main = "Ideale gezinsgrootte", xlab =
"Ideale gezinsgrootte")
summary(data_GN8$ideal_family_size_num)
sd(data_GN8$ideal_family_size_num)
```

```
boxplot(data_sel_GN2$ideal_family_size_nu
m ~ data_sel_GN2$clusters,
main = "Boxplot Ideale Gezinsgrootte",
xlab = "Netwerkttype",
ylab = "Aantal kinderen")
```

```
barplot(table(data_sel_GN2$childwish,
data_sel_GN2$clusters),
main = "Staafdiagram van Kinderwens",
xlab = "Netwerkttype",
ylab = "Aantal",
col = c("yellow", "green", "blue", "red",
"pink"),
legend = FALSE)
```

Gemiddelden t-toets voor verschil in relaties

```
prop.test(c(4443, 2267),
          c(12225, 5425),
          correct = FALSE)

t.test(data_GN7$rel_closefamily,
       data_GN8$rel_closefamily, paired=FALSE)
t.test(data_GN7$rel_otherfamily,
       data_GN8$rel_otherfamily, paired=FALSE)
t.test(data_GN7$rel_friendsacquaintance,
       data_GN8$rel_friendsacquaintance,
       paired=FALSE)
t.test(data_GN7$rel_colleagues,
       data_GN8$rel_colleagues, paired=FALSE)

sum(data_GN7$rel_closefamily)
sum(data_GN7$rel_otherfamily)
sum(data_GN7$rel_friendsacquaintance)
sum(data_GN7$rel_colleagues)
sum(data_GN7$rel_closefamily,
     data_GN7$rel_otherfamily,
     data_GN7$rel_friendsacquaintance,
     data_GN7$rel_colleagues)

sum(data_GN8$rel_closefamily)
sum(data_GN8$rel_otherfamily)
sum(data_GN8$rel_friendsacquaintance)
sum(data_GN8$rel_colleagues)
sum(data_GN8$rel_closefamily,
     data_GN8$rel_otherfamily,
     data_GN8$rel_friendsacquaintance,
     data_GN8$rel_colleagues)

mean(data_GN7$modularity)
mean(data_GN7$comm_1or2)
mean(data_GN7$comm_3orhigher)
sd(data_GN7$modularity)
sd(data_GN7$comm_1or2)
sd(data_GN7$comm_3orhigher)

mean(data_GN8$modularity)
mean(data_GN8$comm_1or2)
mean(data_GN8$comm_3orhigher)
sd(data_GN8$modularity)
sd(data_GN8$comm_1or2)
sd(data_GN8$comm_3orhigher)
```

Compositie kenmerken toevoegen aan structurele netwerktypen

```
## Netwerktipe 1
```

```
data_GN1 <- data_GN1 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
  function(x)
    sum(case_when(x$primary_relation_a %in%
    c("Partner", "Parent", "Sibling") ~ 1,
    TRUE ~ 0) )),
    rel_otherfamily = map_dbl(alter_attr,
    function(x)
    sum(case_when(x$primary_relation_a %in%
    c("Kin", "In-law") ~ 1,
    TRUE ~ 0) )),
    rel_colleagues = map_dbl(alter_attr,
    function(x)
    sum(case_when(x$primary_relation_a %in%
    c("Work") ~ 1,
    TRUE ~ 0) )),
    rel_friends = map_dbl(alter_attr,
    function(x)
    sum(case_when(x$primary_relation_a %in%
    c("Partner's friends", "High-school",
    "Primary school", "Social activity", "College",
    "Mutual acquaintance", "Neighbourhood",
    "Other") ~ 1,
    TRUE ~ 0))),
    rel_missend = map_dbl(alter_attr,
    function(x)
    sum(case_when(x$primary_relation_a %in%
    c("Partner", "Parent", "Sibling", "Kin", "In-law",
    "Partner's friends", "High-school", "Work",
    "Primary school", "Social activity", "College",
    "Mutual acquaintance", "Neighbourhood",
    "Other") ~ 0, TRUE ~ 1))))

data_GN1 <- data_GN1 |>
  mutate(rel_friendsacquaintance = rel_friends
  + rel_missend)

mean(data_GN1$rel_closefamily)
mean(data_GN1$rel_otherfamily)
mean(data_GN1$rel_friendsacquaintance)
mean(data_GN1$rel_colleagues)
sd(data_GN1$rel_closefamily)
sd(data_GN1$rel_otherfamily)
```



```

sd(data_GN1$rel_friendsacquaintance)
sd(data_GN1$rel_colleagues)
sum(data_GN1$rel_closefamily)
sum(data_GN1$rel_otherfamily)
sum(data_GN1$rel_friendsacquaintance)
sum(data_GN1$rel_colleagues)

## Netwerktipe 2
data_GN6 <- data_GN6 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,
TRUE ~ 0) )),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,
TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,
TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",
"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))

data_GN6 <- data_GN6 |>

```

```

  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

mean(data_GN6$rel_closefamily)
mean(data_GN6$rel_otherfamily)
mean(data_GN6$rel_friendsacquaintance)
mean(data_GN6$rel_colleagues)
sd(data_GN6$rel_closefamily)
sd(data_GN6$rel_otherfamily)
sd(data_GN6$rel_friendsacquaintance)
sd(data_GN6$rel_colleagues)
sum(data_GN6$rel_closefamily)
sum(data_GN6$rel_otherfamily)
sum(data_GN6$rel_friendsacquaintance)
sum(data_GN6$rel_colleagues)

## Netwerktipe 3
data_GN3 <- data_GN3 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,
TRUE ~ 0) )),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,
TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,
TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",
"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",

```

```

"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))
data_GN3 <- data_GN3 |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

mean(data_GN3$rel_closefamily)
mean(data_GN3$rel_otherfamily)
mean(data_GN3$rel_friendsacquaintance)
mean(data_GN3$rel_colleagues)
sd(data_GN3$rel_closefamily)
sd(data_GN3$rel_otherfamily)
sd(data_GN3$rel_friendsacquaintance)
sd(data_GN3$rel_colleagues)
sum(data_GN3$rel_closefamily)
sum(data_GN3$rel_otherfamily)
sum(data_GN3$rel_friendsacquaintance)
sum(data_GN3$rel_colleagues)

## Netwerktipe 4
data_GN2 <- data_GN2 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,

TRUE ~ 0) )),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,

TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,

TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",

```

```

"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,

TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",

"Partner's friends", "High-school", "Work",

"Primary school", "Social activity", "College",

"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))

data_GN2 <- data_GN2 |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

mean(data_GN2$rel_closefamily)
mean(data_GN2$rel_otherfamily)
mean(data_GN2$rel_friendsacquaintance)
mean(data_GN2$rel_colleagues)
sd(data_GN2$rel_closefamily)
sd(data_GN2$rel_otherfamily)
sd(data_GN2$rel_friendsacquaintance)
sd(data_GN2$rel_colleagues)
sum(data_GN2$rel_closefamily)
sum(data_GN2$rel_otherfamily)
sum(data_GN2$rel_friendsacquaintance)
sum(data_GN2$rel_colleagues)

##Netwerktipe 5
data_GN4 <- data_GN4 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,

TRUE ~ 0) )),
  rel_otherfamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,

TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,

TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",

```

```

TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",
"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))
data_GN4 <- data_GN4 |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

mean(data_GN4$rel_closefamily)
mean(data_GN4$rel_otherfamily)
mean(data_GN4$rel_friendsacquaintance)
mean(data_GN4$rel_colleagues)
sd(data_GN4$rel_closefamily)
sd(data_GN4$rel_otherfamily)
sd(data_GN4$rel_friendsacquaintance)
sd(data_GN4$rel_colleagues)
sum(data_GN4$rel_closefamily)
sum(data_GN4$rel_otherfamily)
sum(data_GN4$rel_friendsacquaintance)
sum(data_GN4$rel_colleagues)

## Netwerktipe 6
data_GN5 <- data_GN5 |>
  mutate(rel_closefamily = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling") ~ 1,
TRUE ~ 0) )),
  rel_otherfamily = map_dbl(alter_attr,
function(x)

```

```

sum(case_when(x$primary_relation_a %in%
c("Kin", "In-law") ~ 1,
TRUE ~ 0) )),
  rel_colleagues = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Work") ~ 1,
TRUE ~ 0) )),
  rel_friends = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner's friends", "High-school",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 1,
TRUE ~ 0))),
  rel_missend = map_dbl(alter_attr,
function(x)
sum(case_when(x$primary_relation_a %in%
c("Partner", "Parent", "Sibling", "Kin", "In-
law",
"Partner's friends", "High-school", "Work",
"Primary school", "Social activity", "College",
"Mutual acquaintance", "Neighbourhood",
"Other") ~ 0, TRUE ~ 1))))
data_GN5 <- data_GN5 |>
  mutate(rel_friendsacquaintance = rel_friends
+ rel_missend)

mean(data_GN5$rel_closefamily)
mean(data_GN5$rel_otherfamily)
mean(data_GN5$rel_friendsacquaintance)
mean(data_GN5$rel_colleagues)
sd(data_GN5$rel_closefamily)
sd(data_GN5$rel_otherfamily)
sd(data_GN5$rel_friendsacquaintance)
sd(data_GN5$rel_colleagues)
sum(data_GN5$rel_closefamily)
sum(data_GN5$rel_otherfamily)
sum(data_GN5$rel_friendsacquaintance)
sum(data_GN5$rel_colleagues)

data_sel_GN2 <- data_sel_GN2 |>

```

```

mutate(netwerktypen =
recode(data_sel_GN$num_clusters,
      "1" = 1,
      "2" = 4,
      "3" = 3,
      "4" = 5,
      "5" = 6,
      "6" = 2))

ANOVA_relatie_closefamily <-
aov(data_sel_GN2$rel_closefamily ~
factor(data_sel_GN2$netwerktypen))
summary(ANOVA_relatie_closefamily)
TukeyCF <-
TukeyHSD(ANOVA_relatie_closefamily)
plot(TukeyHSD(ANOVA_relatie_closefamily)
)

library(lsr)
cohensD(data_GN4$rel_closefamily,
data_GN1$rel_closefamily) #0,49
cohensD(data_GN4$rel_closefamily,
data_GN6$rel_closefamily) #0,47
cohensD(data_GN4$rel_closefamily,
data_GN2$rel_closefamily) #0,43

KW_Stat5 <- kruskal.test(rel_closefamily ~
netwerktypen, data = data_sel_GN2)
KW_Stat5 <- kruskal.test(rel_closefamily ~
netwerktypen, data = data_sel_GN2)$statistic
Eta_squared5 <- KW_Stat5 /
(nrow(data_sel_GN2) - 1)
paste("Eta squared is",Eta_squared5)

pairwise.wilcox.test(data_sel_GN2$rel_closef
amily, data_sel_GN2$netwerktypen,
p.adjust.method = "bonferroni")

ANOVA_relatie_otherfamily <-
aov(data_sel_GN2$rel_otherfamily ~
factor(data_sel_GN2$netwerktypen))
summary(ANOVA_relatie_otherfamily)
TukeyHSD(ANOVA_relatie_otherfamily)
plot(TukeyHSD(ANOVA_relatie_otherfamily)
)

cohensD(data_GN4$rel_otherfamily,
data_GN1$rel_otherfamily) #0,62
cohensD(data_GN4$rel_otherfamily,
data_GN6$rel_otherfamily) #0,69
cohensD(data_GN4$rel_otherfamily,
data_GN2$rel_otherfamily) #0,69

```

```

cohensD(data_GN5$rel_otherfamily,
data_GN1$rel_otherfamily) #0,62
cohensD(data_GN5$rel_otherfamily,
data_GN6$rel_otherfamily) #0,69
cohensD(data_GN5$rel_otherfamily,
data_GN2$rel_otherfamily) #0,69

KW_Stat6 <- kruskal.test(rel_otherfamily ~
netwerktypen, data = data_sel_GN2)
KW_Stat6 <- kruskal.test(rel_otherfamily ~
netwerktypen, data = data_sel_GN2)$statistic
Eta_squared6 <- KW_Stat6 /
(nrow(data_sel_GN2) - 1)
paste("Eta squared is",Eta_squared6)

pairwise.wilcox.test(data_sel_GN2$rel_otherf
amily, data_sel_GN2$netwerktypen,
p.adjust.method = "bonferroni")

ANOVA_relatie_friends <-
aov(data_sel_GN2$rel_friendsacquaintance ~
factor(data_sel_GN2$netwerktypen))
summary(ANOVA_relatie_friends)
TukeyHSD(ANOVA_relatie_friends)
plot(TukeyHSD(ANOVA_relatie_friends))

KW_Stat7 <-
kruskal.test(rel_friendsacquaintance ~
netwerktypen, data = data_sel_GN2)
KW_Stat7 <-
kruskal.test(rel_friendsacquaintance ~
netwerktypen, data = data_sel_GN2)$statistic
Eta_squared7 <- KW_Stat7 /
(nrow(data_sel_GN2) - 1)
paste("Eta squared is",Eta_squared7)

pairwise.wilcox.test(data_sel_GN2$rel_friends
acquaintance, data_sel_GN2$netwerktypen,
p.adjust.method = "bonferroni")

ANOVA_relatie_colleagues <-
aov(data_sel_GN2$rel_colleagues ~
factor(data_sel_GN2$netwerktypen))
summary(ANOVA_relatie_colleagues)
TukeyHSD(ANOVA_relatie_colleagues)
plot(TukeyHSD(ANOVA_relatie_colleagues))

cohensD(data_GN1$rel_colleagues,
data_GN6$rel_colleagues) #0,44
cohensD(data_GN1$rel_colleagues,
data_GN3$rel_colleagues) #0,69
cohensD(data_GN1$rel_colleagues,
data_GN2$rel_colleagues) #0,69

```

```
cohensD(data_GN1$rel_colleagues,  
data_GN5$rel_colleagues) #0,69
```

```
cohensD(data_GN4$rel_colleagues,  
data_GN6$rel_colleagues) #0,69  
cohensD(data_GN4$rel_colleagues,  
data_GN3$rel_colleagues) #0,69  
cohensD(data_GN4$rel_colleagues,  
data_GN5$rel_colleagues) #0,69
```

```
KW_Stat8 <- kruskal.test(rel_colleagues ~  
netwerktypen, data = data_sel_GN2)  
KW_Stat8 <- kruskal.test(rel_colleagues ~  
netwerktypen, data = data_sel_GN2)$statistic  
Eta_squared8 <- KW_Stat8 /  
(nrow(data_sel_GN2) - 1)  
paste("Eta squared is",Eta_squared8)
```

```
pairwise.wilcox.test(data_sel_GN2$rel_collea  
gues, data_sel_GN2$netwerktypen,  
p.adjust.method = "bonferroni")
```