**Lab 5: Eindopdracht**

**Analyse van de evolutie van een vriendschapsnetwerk met behulp van sociale netwerkanalyse in R**

1. **Doel van de opdracht**

Deze eindopdracht is gericht op het toepassen van theoretische en methodologische kennis omtrent sociale netwerkanalyse (SNA) binnen een toegepaste, computationele context (R). Centraal staat het analyseren van een longitudinaal vriendschapsnetwerk aan de hand van een reeks structurele netwerkindicatoren, met bijzondere aandacht voor de globale netwerkstructuur, de positie van (centrale) actoren en de evolutie van onderliggende sociale relaties doorheen de tijd.

Doorheen deze opdracht wordt aldus de mogelijkheid geboden om (het merendeel van) de praktische kennis die de afgelopen sessies aan bod is gekomen, zelfstandig toe te passen op een realistische dataset. De nadruk ligt niet alleen op het uitvoeren van relevante netwerkberekeningen, maar ook op het interpreteren van de resultaten, en het onderbouwen van inzichten met behulp van visuele en statistische technieken beschikbaar binnen het R-ecosysteem.

Het doel van deze eindopdracht is om een zo volledig mogelijke analyse aan te leveren van een vriendschapsnetwerk, gemeten en geobserveerd op twee momenten doorheen de tijd. Hierbij ligt de focus voornamelijk op het analyseren van de globale structuur van het netwerk en het analyseren van de belangrijkste maten en concepten per analyseniveau, zoals uiteengezet doorheen de cursus. Deze analyses zullen jullie dienen te verwerken in een kort en bondig rapport, waarin jullie antwoord geven op de door jullie geformuleerde onderzoeksvragen.

1. **Dataset**

Voor deze opdracht wordt gebruikgemaakt van een subset van de ‘RECENS Wired into Each Other’ dataset. Deze dataset bevat sociale netwerkdata omtrent vriendschapsnetwerken verzameld onder Hongaarse middelbare scholieren tussen 2010 en 2013 (Vörös et al., 2022). De originele dataset omvat gegevens van 1.767 leerlingen uit 44 klassen, verspreid over zeven scholen, waarbij sociale relaties en individuele gedragskenmerken werden geregistreerd op vier tijdstippen.

In dit opzicht werden meer dan 30 verschillende soorten relaties en percepties gemeten, waaronder vriendschap, affectie, pesten, sociale rollen en gedragingen. Daarnaast zijn uitgebreide individuele kenmerken en klasinformatie verzameld, wat de dataset bij uitstek geschikt maakt voor het bestuderen van netwerkdynamieken en co-evolutie van relaties en gedrag.

Voor de huidige analyse wordt een beperkte subset gebruikt, bestaande uit de eerste twee meetmomenten (wave 1 en wave 2) van de vriendschapsnetwerken binnen één enkele geselecteerde klas **(code 4200 uit de RECENS data folder)**. De netwerkdata betreffen de nominaties van vriendschapsbanden tussen leerlingen, waarbij elke leerling gevraagd werd zijn of haar vrienden binnen de klas aan te duiden over de verschillende tijdstippen heen.

Daarnaast worden ook twee individuele attributen meegenomen:

* Gender (biologische sekse van de leerling),
* Drinkgedrag (frequentie van alcoholconsumptie), zoals gerapporteerd door de leerlingen zelf.

1. **Structuur**

De bedoeling is dat jullie de analyses en interpretatie van de analyses uitschrijven in een RMarkdown rapport onder de vorm van een notebook. Dit ‘rapport’ moet een helder, overzicht bieden van de door jullie uitgevoerde stappen en de centrale bevindingen die voortvloeien uit de analyses. Jullie zijn vrij in het bepalen van de structuur van het rapport, maar ik zet jullie alvast even op weg met enkele aspecten die zeker aan bod dienen te komen:

* 1. Inleiding
     + Beschrijving van de datasets die worden gebruikt
     + Beschrijving van de centrale onderzoeksvragen in relatie tot de data die beschikbaar zijn
     + Beknopt overzicht van de structuur van het rapport
  2. Inlezen, opkuisen en voorbereiden van de data **(tip: zet de netwerkdata om naar binair formaat)**
     + Documenteer waarom je bepaalde packages gebruikt
     + Documenteer hoe je data inleest in R
     + Documenteer de belangrijkste stappen die je hebt ondernomen om de data te inspecteren
     + Documenteer de belangrijkste stappen die je hebt ondernomen om de data op te kuisen
     + Documenteer de belangrijkste stappen die je hebt ondernomen om de data voor te bereiden in functie van verdere analyse
  3. Analyse en rapportage van de resultaten: dit omvat minstens:
     + Initiële visualisatie van het netwerk op tijdstip 1 en tijdstip 2 (wat valt reeds op?)
     + Initiële definitie van het netwerk op tijdstip 1 en tijdstip 2
     + Analyse op *node niveau* en interpretatie van de resultaten over de twee waves heen
     + Analyse op *dyad niveau* en interpretatie van de resultaten over de twee waves heen
     + Analyse op *triad niveau* en interpretatie van de resultaten over de twee waves heen
     + Analyse op *tussen-niveau* en interpretatie van de resultaten over de twee waves heen
     + Analyse op *netwerk niveau* en interpretatie van de resultaten over de twee waves heen
  4. Beknopte discussie van de centrale resultaten en reflectie omtrent mogelijk vervolg analyses
     + Vrij te structureren

1. **Praktische informatie**

Hierbij nog enkele praktische richtlijnen:

* Dit rapport mag kort en bondig zijn, maar documenteer wel alle benodigde stappen in de analyse en interpreteer de resultaten waar nodig
* Werk in RMarkdown en zorg ervoor dat de volledige analyse voldoende inzichtelijk is
* Gebruik voldoende commentaar in de codeblokken om de analyse begrijpelijk te maken voor een externe lezer zonder voorkennis van de dataset alsook om aan te tonen wat een bepaald codeblok/lijn doet
* Het RMarkdown-bestand kan geëxporteerd worden naar verschillende output-formaten. Voor de cursus is het wenselijk om het rapport als PDF-document in te leveren. Alternatief kan het rapport ook worden geëxporteerd als HTML of Word-document, afhankelijk van persoonlijke voorkeur en technische mogelijkheden, maar PDF-formaat heeft de voorkeur

Veel sucess!