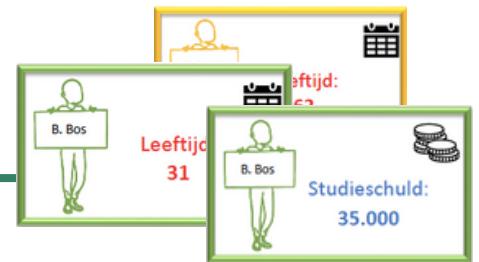


Instructie data dialogen voor de gespreksleider

Deel 1: De database game



In het spel krijgen deelnemers een set met groene kaarten met daarop informatie over fictieve personen die wel werk hebben gevonden, en een set met gele kaarten met informatie over fictieve personen die nog geen werk hebben gevonden.

In iedere set zit informatie van tien fictieve personen. Van elke fictieve persoon zitten er 10 kaarten in waar verschillende kenmerken van die persoon op staan genoteerd.

Bijvoorbeeld: van de persoon B. Bos staat er op de kaartjes informatie over leeftijd, opleidingsniveau, migratieachtergrond, studieschuld, sector, aantal ziekenhuisbezoeken, aantal maanden in de uitkering, aantal interventies, postcode, en geslacht.

Deze kenmerken zijn in verschillende kleuren gedrukt, om aan te geven dat gegevens in een dataset uit verschillende bronnen afkomstig kunnen zijn. Dus alle informatie in het **rood** komen uit dezelfde bron, alle informatie in het **zwart** komt uit dezelfde bron, en alle informatie in het **blauw** komt uit dezelfde bron.

De deelnemers krijgen daarna twee opdrachten. Bespreek de reflectie vragen bij de opdrachten.

Opdracht 1: Pseudonimiseren

Opdracht: pseudonimiseer deze dataset.

Deelnemers vervangen op elk van hun kaarten de fictieve naam door een nummer om de notie van pseudonimiseren te begrijpen. Om de deelnemers een hint te geven over wat pseudonimiseren precies is en hoe ze dat zouden kunnen doen zijn er in reeks twee alvast een paar personen gepseudonimiseerd.



Vragen bij deze opdracht:

Wat is pseudonimiseren?

Pseudonimiseren is een procedure waarbij wordt gezorgd dat de informatie in een dataset niet meer direct herleidbaar is tot een persoon. Er wordt een pseudoniem (cijfer in dit geval) gegeven aan informatie (proces is met een sleutel onomkeerbaar). (Anonimiseren is een procedure waarbij ervoor wordt gezorgd dat de informatie in een dataset helemaal niet meer te herleiden is tot een persoon. Onomkeerbaar proces).

Waarom zou je een database willen pseudonimiseren?

Niet herkenbaar dus je kunt mensen niet voortrekken of achterstellen.

Zijn mensen ook echt onherkenbaar als je pseudonimiseert?

Waarschijnlijk niet, ze zijn alleen niet meer DIRECT herleidbaar. Er zit bijvoorbeeld maar 1 persoon in waar bij sector staat 'motorvoertuigen'. Als medewerker van de gemeente weet je dan waarschijnlijk meteen over wie het gaat.

Opdracht 2: Datakoppeling, klantprofielen, voorspellende statistiek

Opdracht 2a:

Leg de kaarten op een logische manier bij elkaar.

Opdracht 2b:

Ga zodra de kaarten op een logische manier gesorteerd zijn op zoek naar verbanden tussen de blauwe en de groene set. Bespreek de overeenkomsten en verschillen tussen de personen die wel een baan hebben gevonden en de personen die geen baan hebben gevonden.

Vragen bij deze opdracht:

1. Waarom heb je de kaarten zo neergelegd als ze nu liggen?

Realisatie dat kaarten op verschillende manieren gesorteerd kunnen worden, afhankelijk van waar je de nadruk op legt. Als je iets over een individueel persoon wilt weten sorteert je ze waarschijnlijk anders dan wanneer je iets over bepaalde kenmerken wilt weten.

2. Wat zijn de verbanden?

Let op, verband is niet altijd een causale relatie maar misschien alleen een correlatie.

3. Wat hebben de personen binnen een groep met elkaar gemeen?

4. Op welke manier verschillen de personen van de ene groep met de personen van de andere groep?

5. Wat valt op als je kijkt naar wie wel of niet een baan heeft gevonden?

6. Kan je op basis van deze informatie voorspellen wat voor mensen (met welke kenmerken) heel snel een baan vinden en welke niet?

(Voorspellende statistiek)

Niet altijd. Let op verband is geen causale relatie maar misschien alleen correlatie.

7. De kenmerken komen uit verschillende bronnen. Is het wel of niet wenselijk om informatie uit verschillende bronnen te koppelen voor cliënten? (Datakoppeling)

8. Wat zeggen data nu eigenlijk wel of niet?

Voorbeeld: Het aantal interventies zegt niet zoveel over hoe goed die hebben gewerkt. Zijn veel interventies goed of niet? Dat weten we eigenlijk niet.



Overzicht van verschillen en overeenkomsten:

Hieronder vindt u een aantal verschillen en overeenkomsten tussen Reeks 1 en Reeks 2. Bij elk verschil/overeenkomst wordt er een aanname gedaan. Deze aannames hoeven niet te kloppen maar zijn bedoeld om te bespreken tijdens de discussie met de deelnemers.

Opleidingsniveau:

Reeks 1 zitten veel meer hoger opgeleiden (9 van de 10 WO/HBO)
Reeks 2 meer lager opgeleiden (3 uit 10 WO/HBO)

Aanname: hoger opgeleiden (reeks 1) vinden sneller een baan

Postcode:

Reeks 1: relatief veel in 3582/3583,
Reeks 2: relatief veel in 3531/3532/3533/3534.

Wonen blijkbaar in verschillende delen van de stad.

Aanname: Je postcode heeft invloed op of je werk kan vinden. Goede/slechte wijk? Werkgelegenheid in wijk?

Leeftijd:

Gemiddelde leeftijd:
Reeks 1: 39,5
Reeks 2: 50

Aanname: Hoe ouder je bent, des te moeilijker het is om een baan te vinden.

Sector:

Reeks 1:
Reeks 2:

Aanname: Hier zijn geen opvallende verschillen of overeenkomsten.

Migratie-achtergrond:

Reeks 1: 8 uit 10 Nederlands,
Reeks 2: 3 uit 10 Nederlands.

Aanname: Als je een Nederlandse achtergrond hebt vindt je sneller een baan.

Ziekenhuisbezoeken:

Reeks 1: 15
Reeks 2: 49

Aanname: Hoe gezonder je bent, des te makkelijker het is om een baan te vinden.

Studieschuld:

Reeks 1: gemiddeld 7600
Reeks 2: gemiddeld 4600

Aanname: Hoe meer studieschuld, des te makkelijker is het om een baan te vinden (hangt samen met hogere opleiding)

Aantal maanden in uitkering:

Reeks 1: gemiddeld 13,5,
Reeks 2: gemiddeld: 11,9

Aanname: Mensen die een baan gevonden hebben zitten langer in een uitkering.

Dit lijkt niet samen te hangen met leeftijd. Oudere mensen zitten blijkbaar niet langer in een uitkering.

Aantal interventies*:

Reeks 1: 38 interventies, gemiddeld 3, 8
Reeks 2: 31, gemiddeld 3,1

Aanname: Hoe meer interventies, hoe sneller je aan een baan komt.

* Interventies zijn activiteiten gericht op het bespoedigen van terugkeer van de werknemer in eigen of passend werk binnen of buiten de werkgever, of ter preventie van ziekteverzuim.

Deel 2: De datadialoog

In het tweede deel wordt aan de hand van een aantal vragen en stellingen in de groep gediscussieerd over ervaringen, opvattingen, kansen, risico's en voorwaarden voor data-gedreven werken.

Vragen

1. Heeft u ervaring met data-gedreven werken?

4. Onder welke voorwaarde is data-gedreven werken volgens u mogelijk?

2. Welke kansen ziet u voor data-gedreven werken?

5. Heeft u er vertrouwen in data-gedreven werken door de gemeente?

3. Welke risico's ziet u voor data-gedreven werken?

6. Waar bent u bang voor als de gemeente data-gedreven werkt?



Stellingen

Data-gedreven werken biedt *kansen*

- om beter inzicht te krijgen in de situatie van de cliënt/bepaalde groepen cliënten.
- om beter passende re-integratiedienstverlening aan te bieden aan een client/bepaalde groepen cliënten.
- om de kwaliteit van de re-integratie dienstverlening te verbeteren voor de client/bepaalde groepen cliënten.

Het *risico* van data-gedreven werken is

- dat onjuiste conclusie worden getrokken op basis van de data
- dat op basis van de data/een klantprofiel 'automatisch' een beslissing wordt gemaakt over de re-integratiedienstverlening aan een cliënt/bepaalde groepen cliënten.

Een *voorwaarde* van data-gedreven werken is

- dat er helderheid moet zijn over de inhoud en het doel van de geregistreerd gegevens
- dat transparant is wie, wanneer, en met welk doel toegang krijgt tot welke informatie
- dat cliënten inzicht krijgen in welke profielen er zijn en hoe beslissingen op basis van een profiel tot stand komen
- dat een cliënt moet kunnen zeggen: "Dit wil ik niet, of ben ik niet".

Ik heb er *vertrouwen* in

- dat de uitkeringsinstantie op een verantwoorde manier met mijn gegevens en die van andere cliënten omgaat
- dat de gemeente dankzij data-gedreven werken mij en andere cliënten een beter passend aanbod kan bieden

Ik ben *bang* dat

- Data-gedreven werken leidt tot ongelijke kansen tussen (groepen) mensen
- ik of bepaalde groepen cliënten onzichtbaar blijven en daardoor geen passende dienstverlening krijgen