



Synthetische Data

Gebruikersdag MBO Digitaal

Lotte Pater- DUO Informatieproducten-
lotte.pater@duo.nl



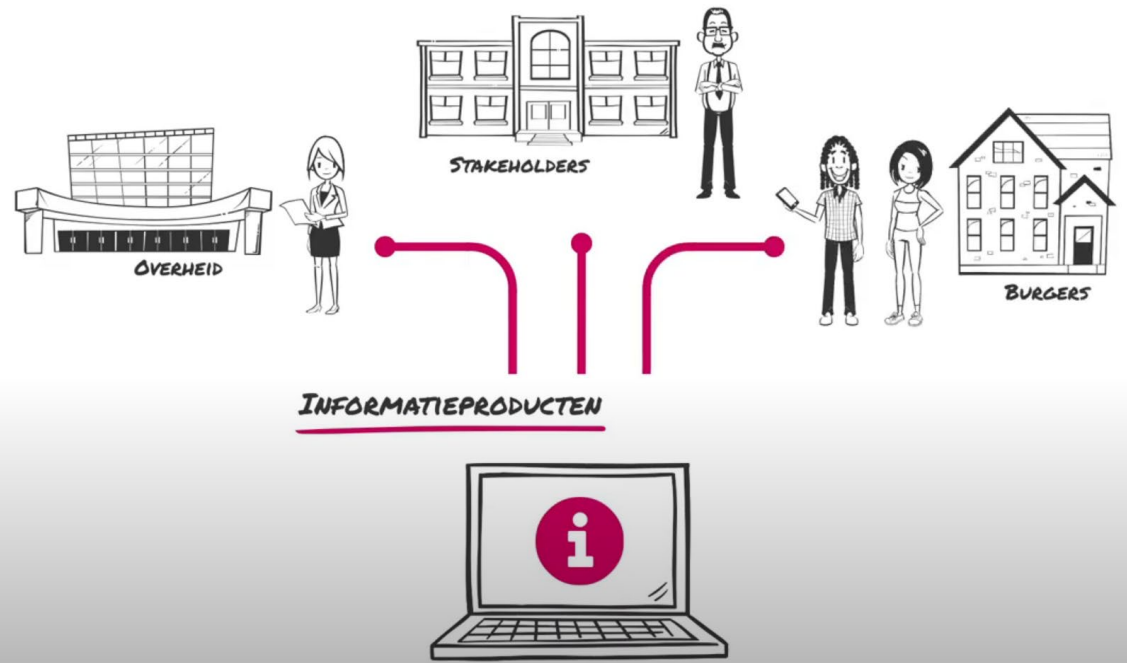
**Gedoe met data delen en
privacy: herkenbaar?**



Doel DUOP: meerwaarde bieden met onderwijsgegevens

Eis DUOP: privacy waarborgen

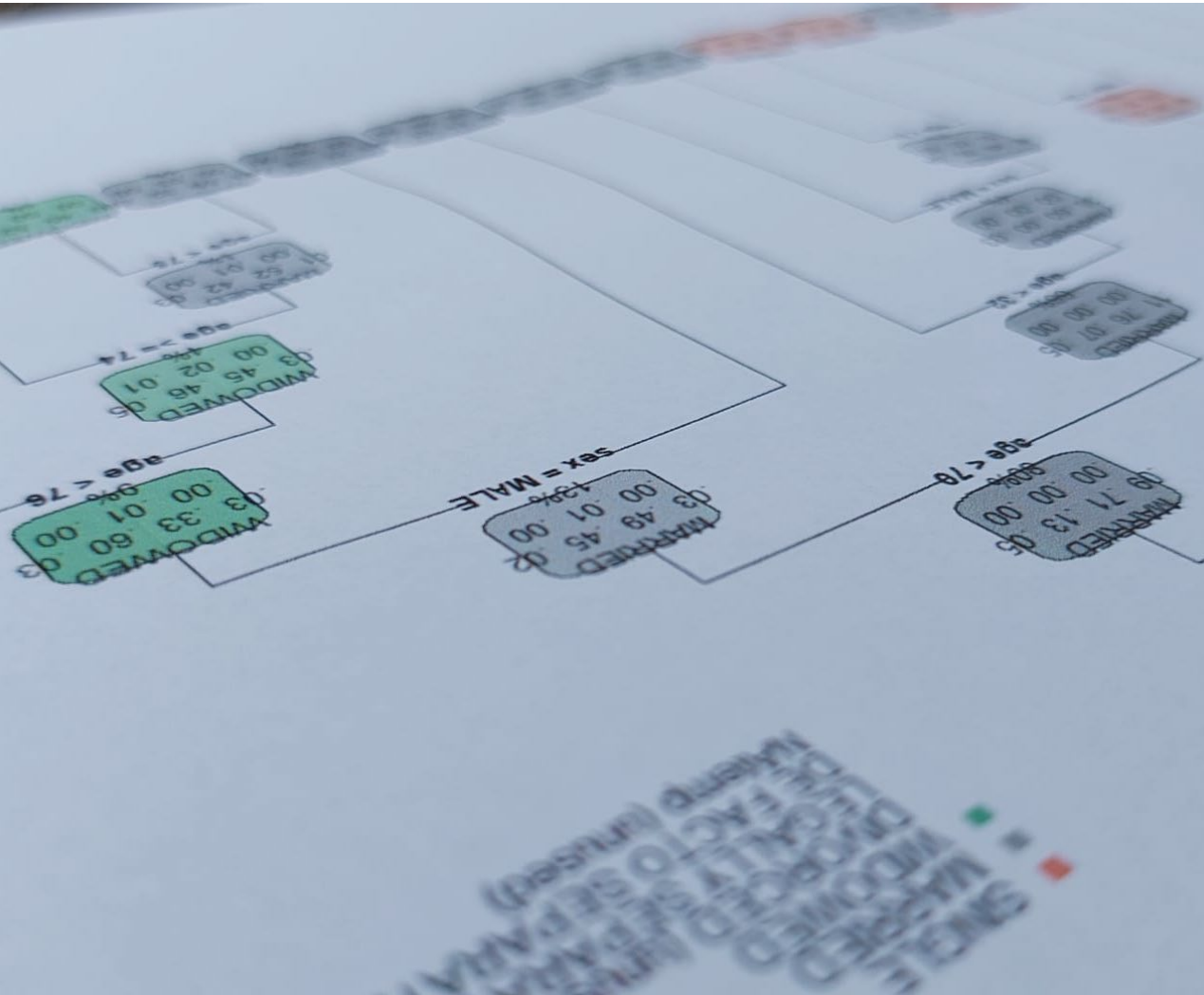
- > Intern: wachttijden en procedures om data te gebruiken
- > Samenwerkingspartners & stakeholders: mogen data meestal niet gebruiken
- > Externen en burgers: weinig toegankelijke data





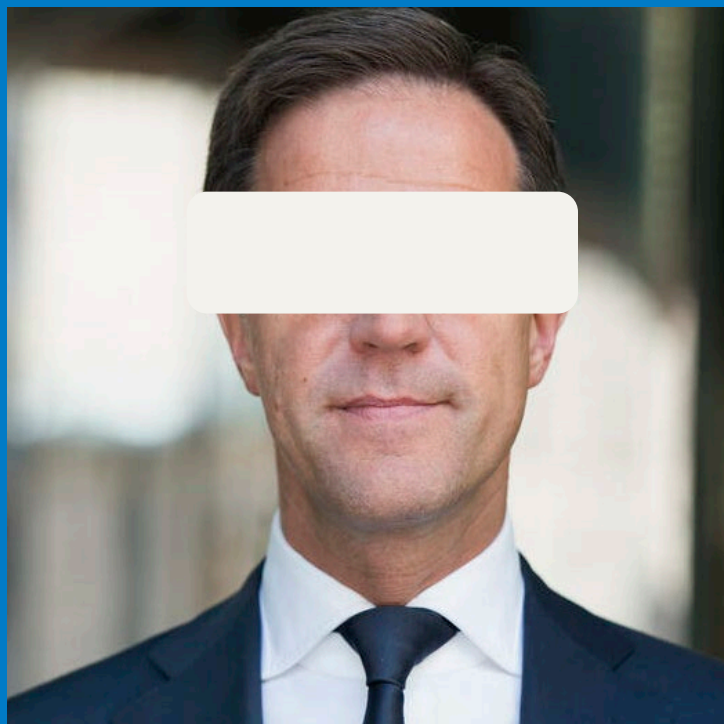
Synthetisch Data

- > Idee:
 - Privacy: individueel niveau
 - Informatie: structuurniveau
- > Synthetische data:
 - Nep op individueel niveau
 - AVG niet van toepassing
 - Zelfde conclusies op structuurniveau





Gepseudonimiseerd versus Synthetisch





Gepseudonimiseerd versus Synthetisch





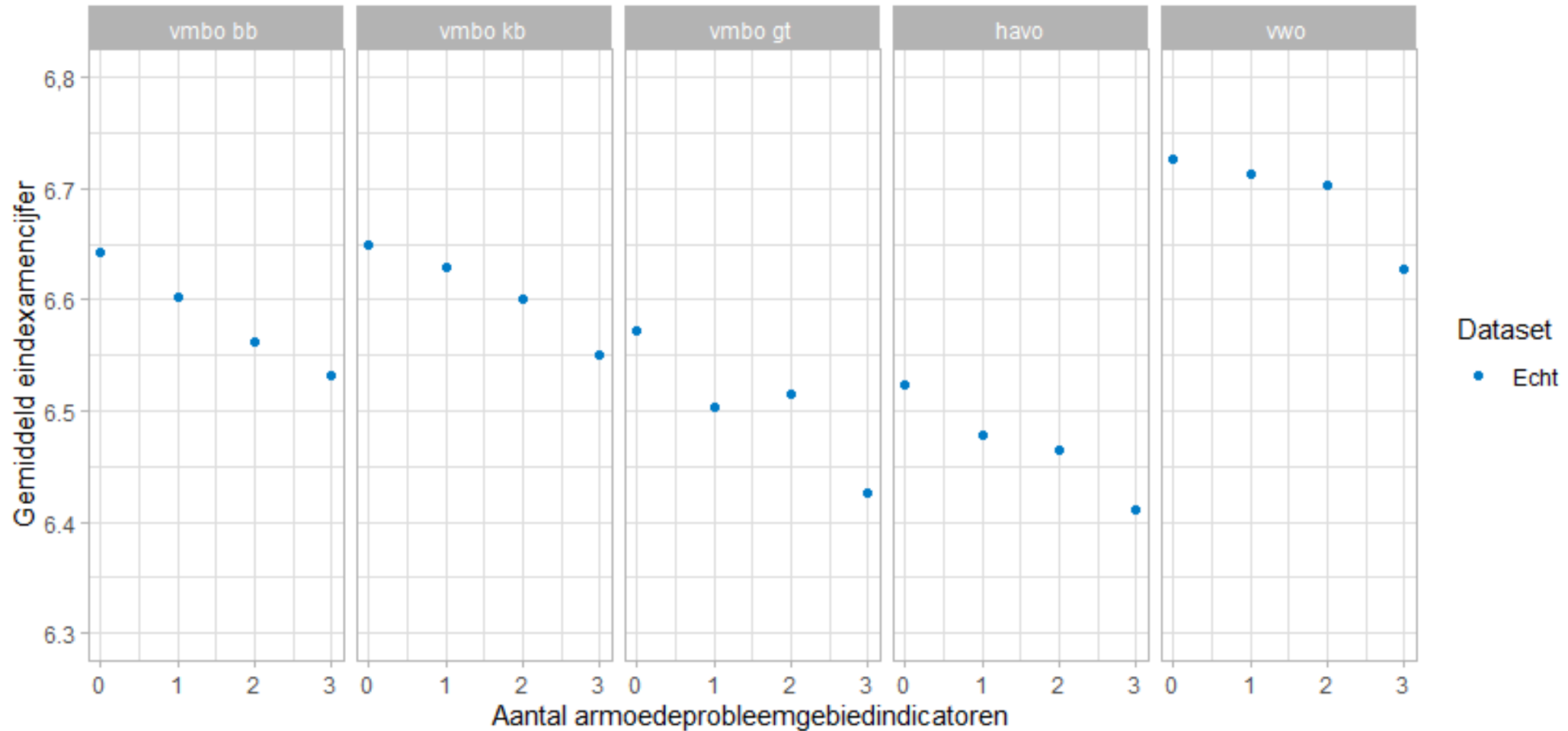
VB: Vergelijking echte en synthetische data

- Levering personeelsgegevens
 - In de echte data: BAPO komt voor bij geboortejaren '56'65, sommigen meer dan anderen
 - In de synthetische data: zelfde verhoudingen, maar een beetje 'overflow' op andere jaren
- Analyse gebruiker: voldoende voor doel.

Report OORSPRONKELIJK		Report SYNTHETISCH	
geboortejahr	BAPO FTE	geboortejahr	BAPO FTE
1938	0	1938	0
1939	0	1939	0
1940	0	1940	0
1941	0	1941	0
1942	0	1942	0
1943	0	1943	0
1944	0	1944	0
1945	0	1945	0
1946	0	1946	0
1947	0	1947	0,1
1948	0	1948	0,1
1949	0	1949	0,1
1950	0	1950	0,3
1951	0	1951	0
1952	0	1952	0,4
1953	0	1953	0,1
1954	0	1954	1,0
1955	0	1955	1,3
1956	47,4	1956	50,9
1957	89,8	1957	83,9
1958	118,7	1958	114,0
1959	119,4	1959	113,5
1960	98,4	1960	97,7
1961	82,5	1961	86,1
1962	70,8	1962	65,7
1963	39,4	1963	35,9
1964	31,0	1964	34,4
1965	17,0	1965	16,1
1966	0	1966	0,1

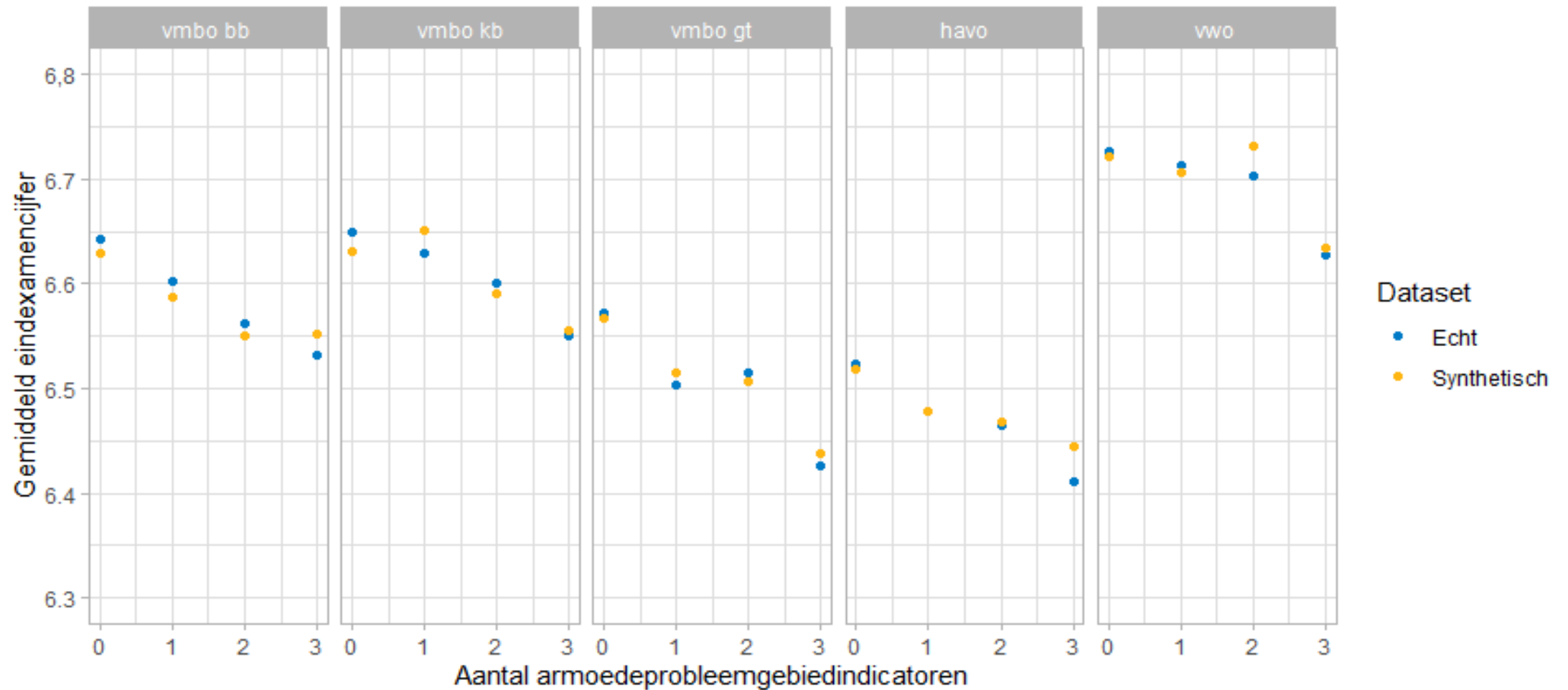


VB: Relatie eindexamencijfers en armoedeprobleemgebieden





VB: Relatie eindexamencijfers en armoedeprobleemgebieden





Synthetische Data aanvragen

WEL DOEN ALS...

- > De open data niet gedetailleerd genoeg is
- > Je geïnteresseerd bent in de voor en nageschiedenis van studenten
- > Je bereid bent mee te denken over de precieze vorm van de data

NIET DOEN ALS...

- > Exacte aantallen noodzakelijk zijn
- > Er geen tijd is voor analyse



Kennisnetwerk Synthetische Data

- > Voor organisaties die zelf werken aan synthetische data
- > Mail: kennisnetwerk@synthetischedata.nl
- > Of join de online community: www.synthetischedata.nl





Vragen?