

# Feedback labo 5

## Data Foundations

De feedback op volgende slides is klassikaal en gebaseerd op jullie code.

Het doel is om jullie te voorzien van tastbare feedback zodat jullie hiermee aan de slag kunnen.

## Dataset verkennen

- Gebrek aan foutafhandeling en/of output bij ontbrekende waarden in meerdere inzendingen.
- Suggestie: gebruik steeds een correlatiematrix om kenmerken met hoge correlatie te reduceren.

# Feature scaling

- Inconsistent gebruik van `StandardScaler` tussen trainings- en testdata.
- Sommige studenten gebruikten verschillende scalers voor visualisaties en modeltraining.
- Aanbeveling: houd één scaler consistent door het hele proces.

# Visualisaties

- Scatterplots missen vaak duidelijke labels en legenda's.
- Tip: gebruik consistente kenmerken tussen training en visualisatie, of geef duidelijk de beperkingen aan.

## Evaluatie van resultaten

- Dubbele berekeningen van precisie, recall en F1-score zonder verdere analyse.
- Confusion matrix wordt vaak weergegeven zonder toelichting op de resultaten of legenda's.
- Suggestie: integreer prestatiestatistieken in een samenvatting van de modelprestaties.

# Conclusie

- Focus op consistentie in dataverwerking en visualisaties.
- Introduceer validaties en foutafhandeling in alle stappen van het proces.
- Geef meer visuele uitleg bij gebruikte methoden en visualisaties om interpretatie te verbeteren.