

Opdracht

1. **Dataset kiezen:** Selecteer een dataset van een open data portal of via een API.
2. **Importeren:** Laad de dataset in Python (pandas) of Excel.
3. **Data cleaning:** Voer de volgende stappen uit en lever op:
 - Identificeer en behandel ontbrekende waarden.
 - Detecteer en behandel outliers.
 - Verwijder duplicaten.
 - Normaliseer één of twee kolommen.
4. **Data profiling:** Maak een overzicht met:
 - Statistische samenvatting (`describe()`).
 - Twee histogrammen voor numerieke kolommen.
 - Een correlatiematrix als heatmap.
5. **Documentatie:** Schrijf kort (300–500 woorden) over je reinigingsproces en uitdagingen.

Bundel alles in een verslag met de belangrijkste grafieken en een toelichting bij elke stap.

Revision #1

Created 13 June 2025 09:52:04 by Admin

Updated 13 June 2025 09:52:35 by Admin