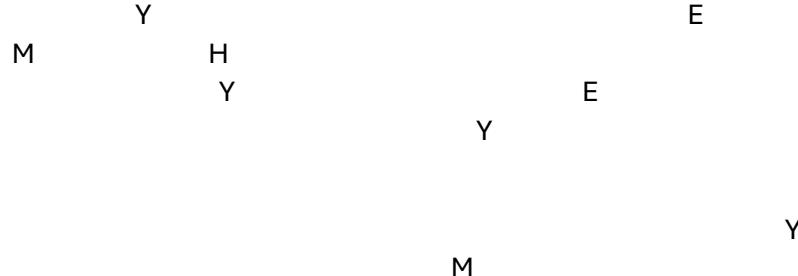


# Data Science Aufgabe 3:

## Umgang mit fehlenden Werten in einem Machine-Learning-Projekt

### Hintergrund



Ihr Ziel ist es einen Klassifikator zu entwickeln welcher vorhersagen kann, ob eine Person herzkrank ist oder nicht.

### Abgabe

- Y
- M
- E E

### Beschreibung der Aufgabe

E

#### 1. Datenvorbereitung

- **Daten laden:** D E E
- E E
- H
- **Datenvorbereitung:** E
- **Erste Visualisierung:** E

## 2. Analyse fehlender Werte

- Untersuchung des Ausmaßes und der Art der fehlenden Werte:

Y

- Visualisierung:

E

Y

## 3. Methoden zur Behandlung fehlender Werte

M H Y H

H D M

E

- Ignorieren fehlender Werte:

E Y

- Löschen fehlender Werte:

• M E

• M Y D

- Statistische Imputation:

• Y Y E

• Y Y Y

• Y Y

- Weitere Imputationsmethoden:

• Y Y

• Y Y

• M Y

E

- Fortgeschrittene Imputationstechniken:

• M	E
Y	

#### 4. Modelltraining und Evaluation

- Modelltraining: H
- Modellbewertung: H

#### 5. Vergleich und Diskussion der Ergebnisse

- Ergebnisvergleich: H
- Diskussion: E

#### AbgabefORMAT

- Jupyter-Notebook: E D
- Dokumentation: E
- Schlussfolgerung: E

#### Hinweise

-