

The kind of data you'd like to work with and the field you're interested in (finance, healthcare surveys, etc.)

- Health Care (*Heart failure*)
- Health Care (*Diabetes*)
- Health Care (*Alzheimer's disease*)
- Health Care (*Stroke*)

The questions you'll ask of the data

- Will I get this disease?
- What is the probability of me getting the disease?
- How accurate is this model in predicting the disease?

Possible source for the data

- Kaggle
 - <https://www.kaggle.com/datasets/nimapourmoradi/heart-failure-clinical-records>
 - <https://www.kaggle.com/datasets/kaushil268/disease-prediction-using-machine-learning?select=Training.csv>

-
- Get data source(s)
 - Store data in a SQLite database
 - SQL query to print out data needed for ML model
 - Run supervised model
 - Confirm accuracy (eg >90%)
 - — Run test Data or input data? —
 - Visuals about the dataset (3 visuals)