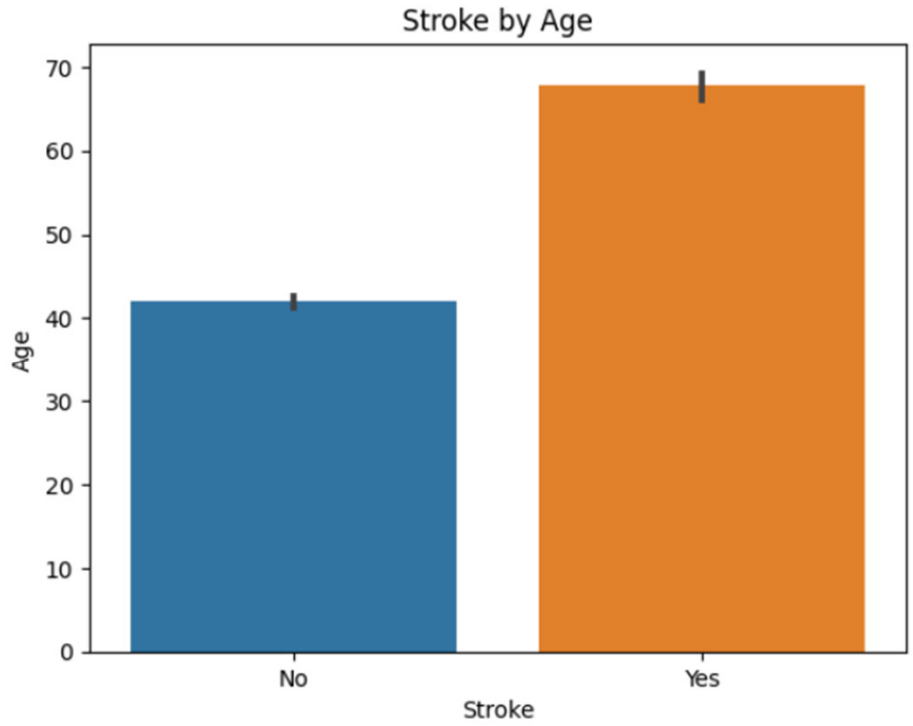
Predicting Stroke

Data from Kaggle

Business Problem

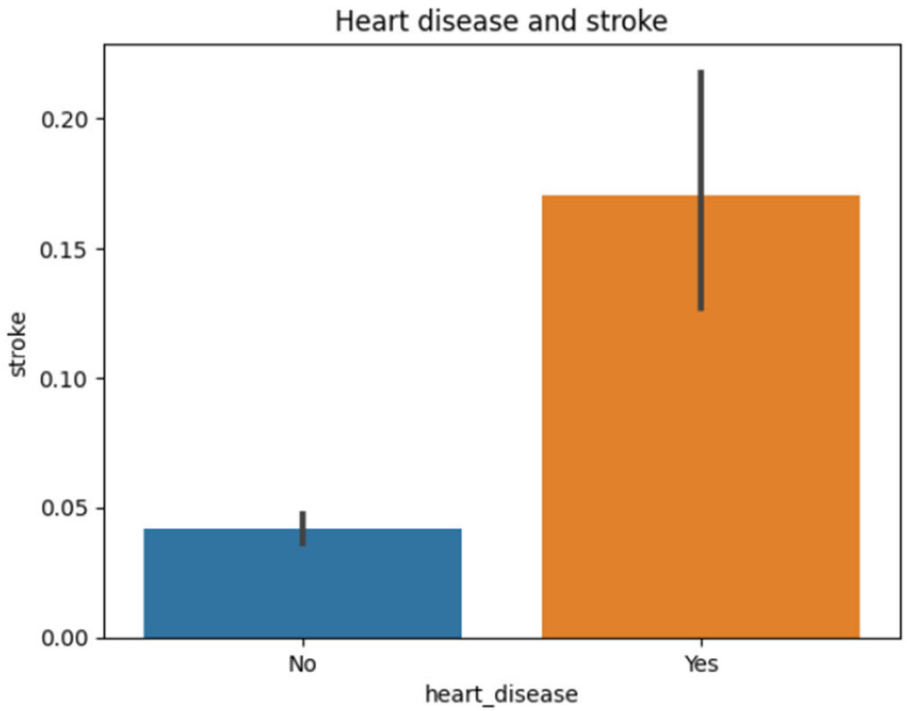
- The dataset is used to predict the likelihood of a patient getting a stroked based on gender, age, various diseases, and smoking status.
- Stakeholders: doctors, hospital, and patients, insurance company

Predicting the relationship of strokes and age



Patients over 40 are more likely to have a stroke

Prediction of relationship for heart disease and stroke



Patients with heart disease are at higher risk of a stroke

Strengths and limitations of the model

- False Negative means patients who are prone to a stroke will not be identified. This is bad for the stakeholders as strokes can lead to deaths and losing the trust of doctors for not being able to accurately predict the strokes.
- False positives are not as serious as the false negative. It is better to predict someone with the potential to have a stroke than to not predict it. However, it may mean unnecessary costs to the patient.

- RECOMMENDATION: Reduce false negatives as much as possible to make sure patients prone to strokes have a higher chance of being flagged.

References

- URL for data: https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset