Likelihood of Stroke Predictions

...

Main Goal

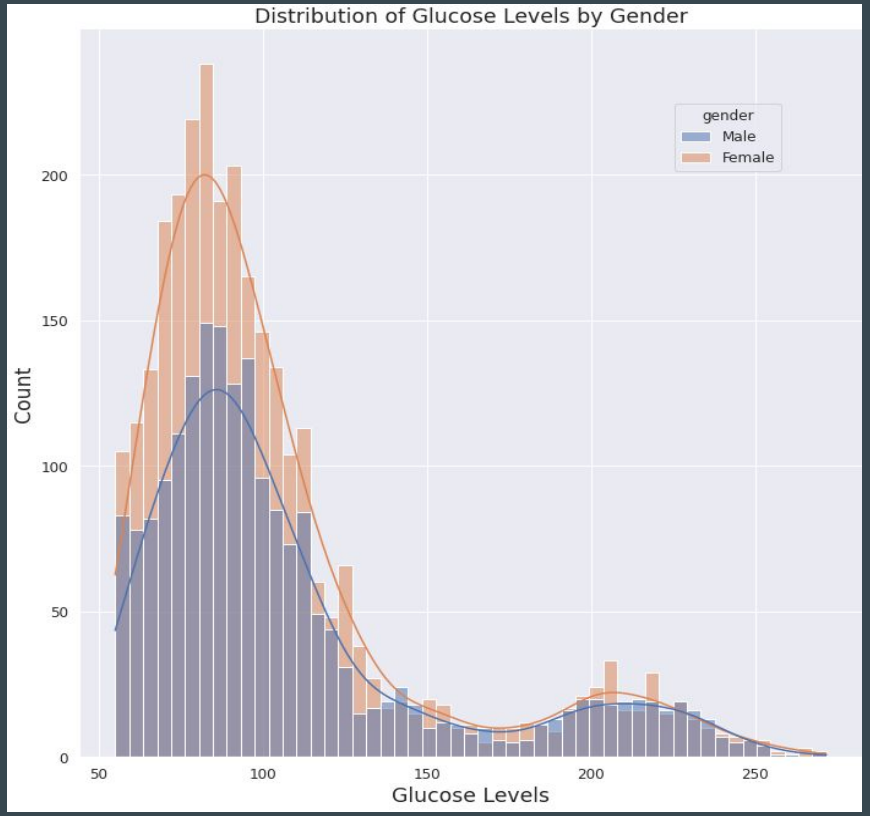
- The main goal of this project is to predict the likelihood of a patient having a stroke.
- In hopes that we can help and guide patients that are at risk of having a stroke.

Background on Data

- 12 features and 5510 observations.
- Our features consist of Id, Gender, Age, Hypertension, Heart Disease, Ever Married, Work Type, Residence Type, Average Glucose Level, Bmi, Smoking Status, and Stroke.
- Our data consist of patients from the ages starting as early as infancy all the way to 80 years of age
- Fun fact Stroke is the second leading cause of death worldwide. Stroke also attributes to disabilities as well.

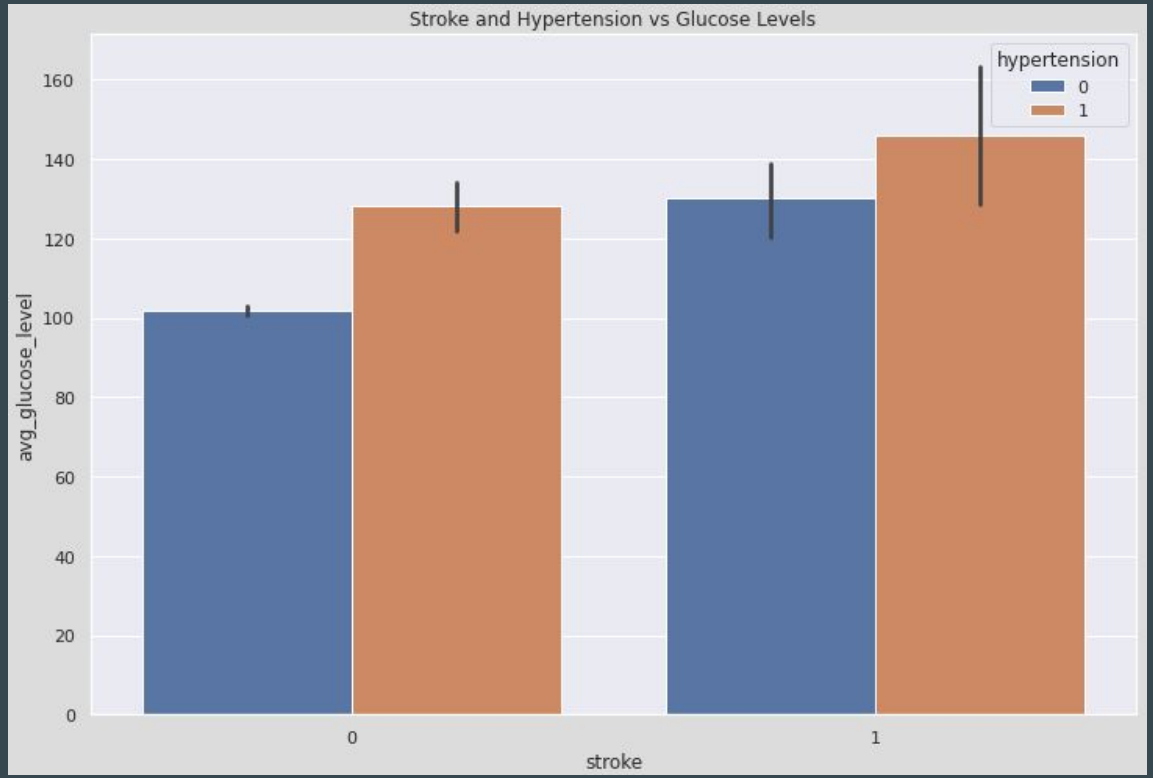
Distribution of Glucose Levels

- Glucose levels are evenly distributed between gender.
- Have more female entries than male entries
- We have some outliers for both genders



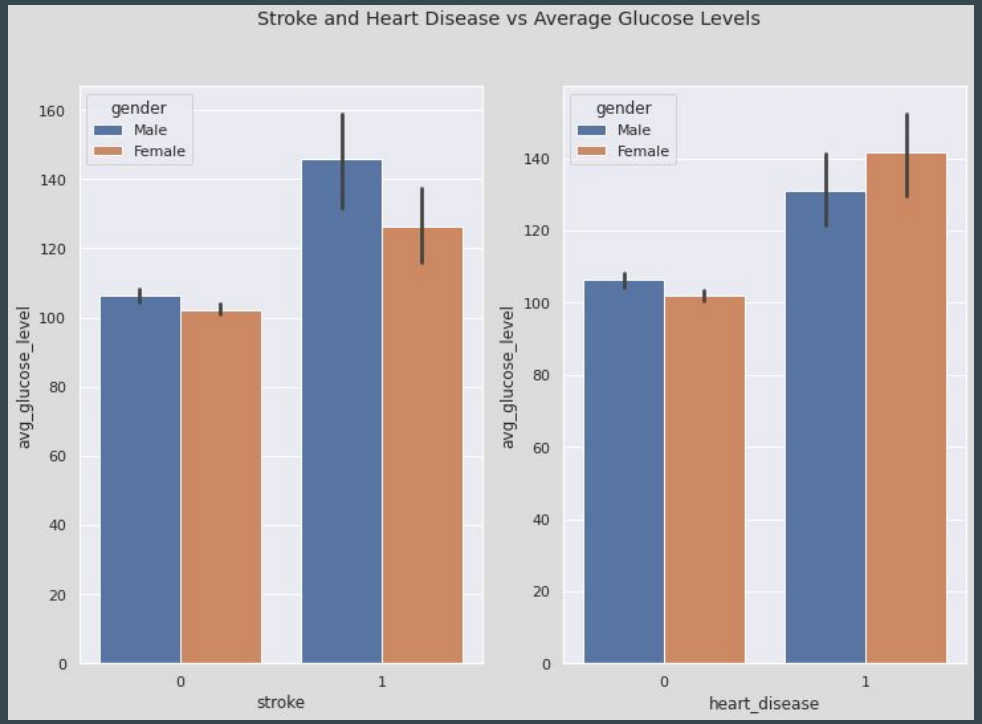
Stroke and Hypertension vs Glucose Levels

- Patients with hypertension tend to have higher glucose levels regardless if they have had a stroke or not.
- Those who have had a stroke and do not have hypertension have similar glucose levels to those who haven't had a stroke but do have hypertension.



Stroke and Heart Disease vs. Glucose Levels

- Patients who have had a stroke tend to have higher glucose levels.
- Higher glucose levels can be an indicator for heart disease as well.
- We can see the “at risk” glucose levels for heart disease and stroke by gender.



Production Model

- For our production model I chose a random forest model.
- Our testing accuracy score for our model was 87%.
- Recall Score was 87%. How our model did labeling patients who weren't at risk for a stroke.
- Precision Score was 87%. How well our model did labeling patients who were at risk for stroke.

Final Recommendations

- A model with higher accuracy would be preferred.
- A higher recall accuracy also would improve our model, so we have the least amount of false negatives as possible.
- With an accuracy of 87%, I believe this model wouldn't be great for predicting patients being at risk for a stroke.
- I believe BMI isn't an accurate metric to measure body fat content. As it doesn't take into account muscle mass, bone density, or body fat distribution.