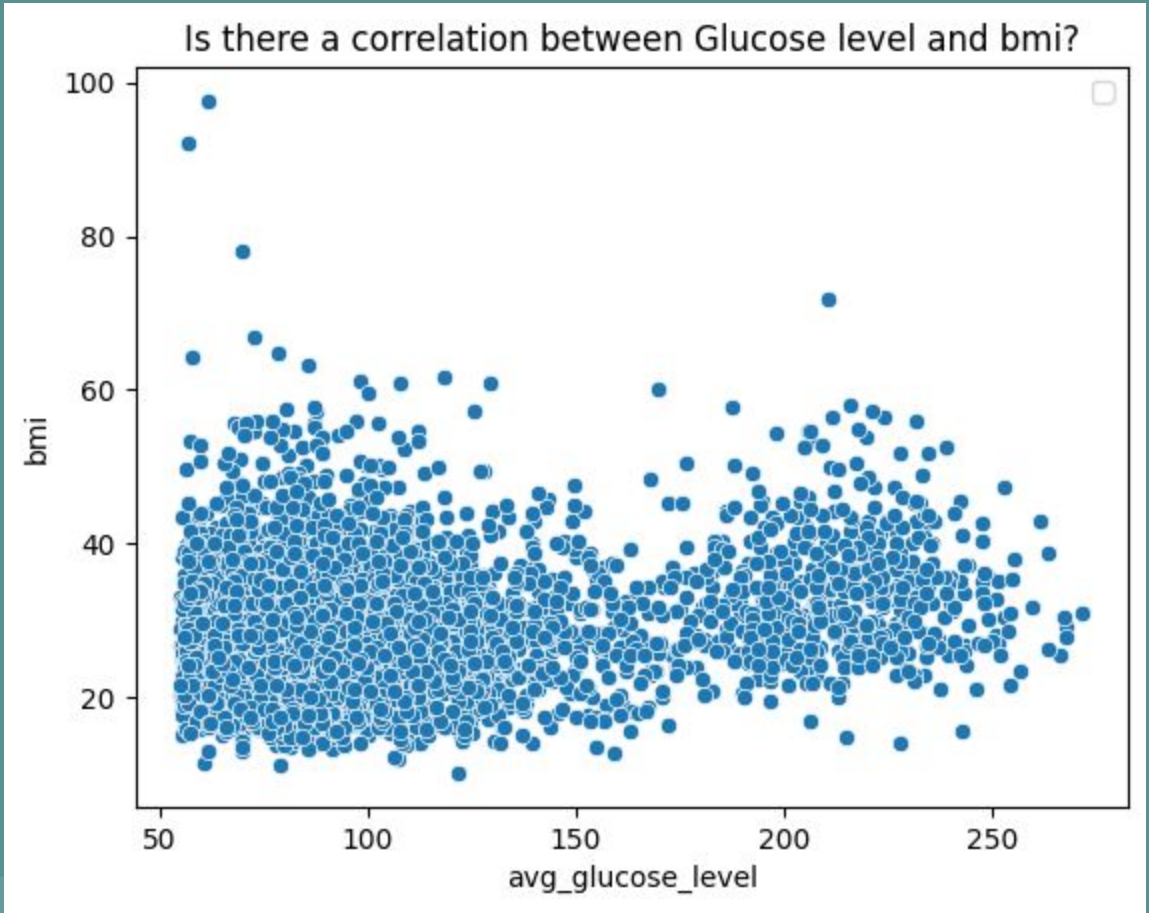
Stroke Predictions

By:Nour Tafa

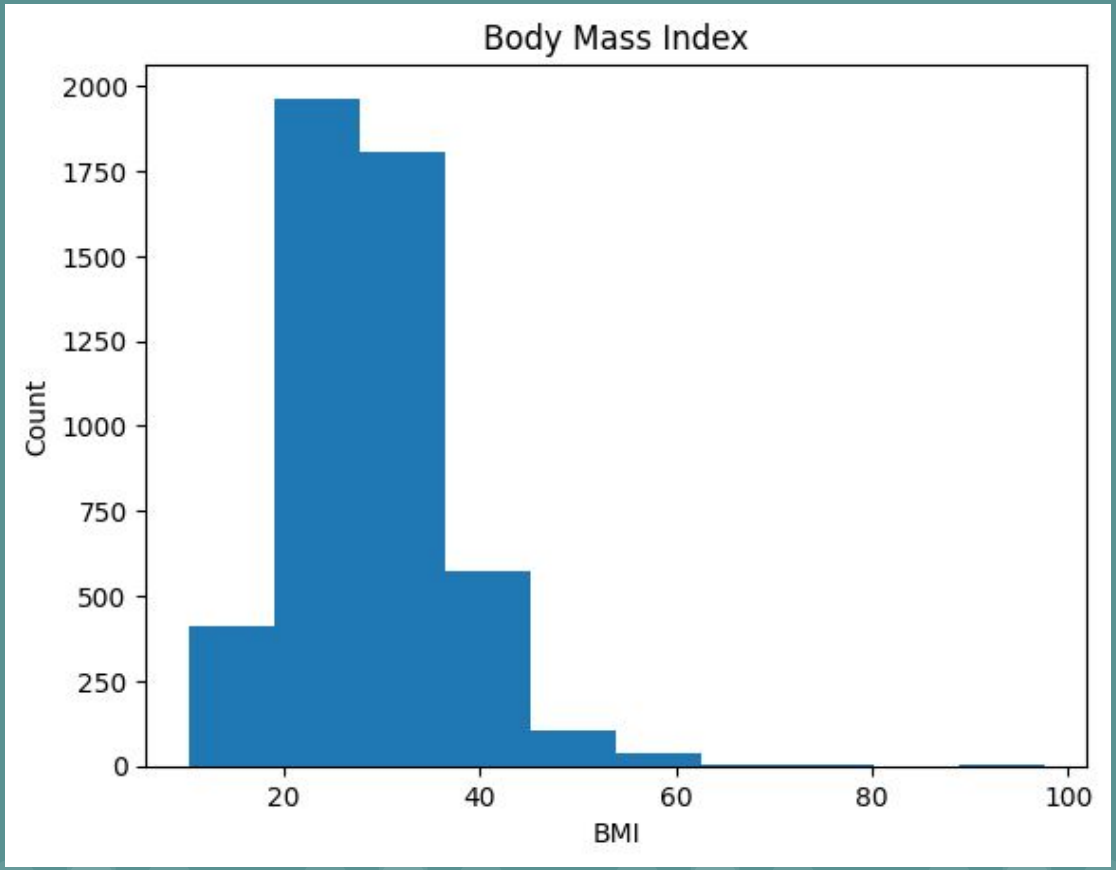
Stakeholders:

1. Healthcare Professionals
2. Patients
3. Researchers and Academics
4. Public Health Organizations
5. Insurance Companies
6. Government Agencies

The problem I am addressing is using machine learning to predict the likelihood of stroke in individuals. By analyzing factors like age, gender, heart diseases, and smoking status, the model aims to identify those at higher risk. This can help healthcare professionals take preventive measures and provide targeted interventions to reduce the incidence and impact of strokes, improving patient outcomes and reducing the burden on healthcare systems.



Scatterplot: This plot illustrates the relationship between BMI (Body Mass Index) and average glucose level. It visualizes how these two factors are interconnected



Histogram: This histogram reveals that the majority of patients have a high BMI (Body Mass Index)

After evaluating and optimizing multiple models to predict the likelihood of stroke in patients, it was found that the XGBoost and PCA Random Forest models performed the best. These models achieved the highest macro average scores, indicating their overall effectiveness in predicting stroke risk. Additionally, they demonstrated the lowest number of false positive and false negative results, suggesting their ability to accurately identify individuals at risk of experiencing a stroke.

Note: To further enhance the performance of the models and improve their scores, collecting more data from a larger pool of patients would be beneficial. Increasing the dataset size can provide the models with a more comprehensive and diverse range of examples to learn from.

Conclusion:

Factors that lead to strokes are interconnected and correlated, enabling predictive models to identify high-risk individuals based on age, gender, presence of heart diseases, and smoking status.

Each person at risk of stroke can benefit from a personalized prevention plan tailored to address their specific leading factor, allowing healthcare professionals to implement targeted interventions and reduce the likelihood and impact of strokes.