



Strokes



Background Research

The topic is strokes. Strokes kill many people per year, and many things, such as high blood pressure, are related to a person's risk for strokes. You can minimize your risk for a stroke, so knowing the things that increase your risk for a stroke are helpful.

Dataset

My dataset is [here](#). It is a slightly changed version of the original dataset, which I found [here](#). The first column, "no," is just to help keep track of the rows. The next column is for the gender of the people, and 2 is supposed to female while 1 is supposed to represent male. The next column is for the age of the people. The next column is to show if any of the people suffer from hypertension (high blood pressure), and 0 means no while 1 means yes. The next column is to show if the person has heart disease, and 0 means no while 1 means yes. The next column shows the work type. A private job is represented by 1, being self-employed is represented by 2, having a government job is represented by 3, staying at home is represented by 4, and never having a job is represented by 5. The next column is residence type. Living in an urban area is represented by 1 while living in a rural area is represented by a 2. The next column shows the person's average glucose levels, and the column after that shows their BMI (Body Mass Index). The next column shows if the ever smoked; if a person never smoked, it says 1, if a person formerly smoked, it says 2, and if a person smokes, then it says 3. The last column is if the person had a stroke or not, and 0 means no while 1 means yes.

Visualization

[Link](#)

I am plotting each of the columns with whether the person had a stroke or not. Most of the graphs show a trend.

Process

[Link](#)

I used numpy, pandas, matplotlib.pyplot, and seaborn.

Future Implications

The possible positive/negative impacts are that people can figure out how likely they are to get a stroke. Most of the graphs, which show up in the Collab when it is run, show a trend. The first graph, Age vs Stroke, shows that the higher one's age, the higher their stroke risk. The second graph, Gender vs Stroke, shows that women are more likely to have a stroke, though not by much. The third graph, Hypertension vs Stroke, shows that if one suffers from hypertension, then they are more likely to have a stroke. The fourth graph, Heart Disease vs Stroke, shows that if one suffers from heart disease, then they are more likely to have a stroke. The fifth graph, Work Type vs Stroke, shows that one is more likely to have a stroke if they have never had a job. The Residence Type vs Stroke graph shows that there is not a correlation between where someone lives to how likely they are to have a stroke. The seventh graph, Average Glucose vs Stroke, shows that the higher the glucose levels, the more likely it is for that person to have a stroke. The eighth graph, Smoking vs Stroke, shows that if someone smokes, then they are more likely to have a stroke. The ninth graph, BMI (Body Mass Index) vs Stroke, shows that the higher the person's BMI, the more likely it is for that person to have a stroke. People can improve their lifestyle to reduce their risk of stroke.

Reflect

Visualizing the data helped me see it all at once, and it was easier to analyze and find patterns in. This can help someone figure out if they are at a higher risk for a stroke, and what changes can they make to avoid a stroke.