

Questions:

What is the relation between line of work type and getting a stroke?

What is the relation between type of residence and getting a stroke?

Does a getting stroke depend on age?

Dose Having hypertension and being a smoker increase the chance of getting a stroke.

The benefit from exploring these questions :

To discover reasons behind the increased rate in strokes in people based on some variables.

Data Description:

The dataset is about the stroke, and I get this dataset from Kaggle website

(<https://www.kaggle.com/fedesoriano/stroke-prediction-dataset>).

Aggregate the average for stroke that depend on gender, married or not, work type, residence type, clean the missing values.

Features:

There are 12 columns in this dataset which are:

1. Id (unique number).
2. Age.
3. Gender.
4. Work type.
5. Ever married (Has the person ever been married).
6. Type of residence (to know the style of living of person).
7. Smoking status (to know if the person smoking or not).
8. Hypertension (to know if the person hit hypertension).
9. Avg_glucose_level (Average glucose level in blood).
10. Stroke (to know if the person hit stroke).
11. Heart disease (to know if the person hit heart disease).
12. BMI (Body Mass Index: for calculation using a person's height and weight by this formula:
 kg/m^2 where kg is a person's weight in kilograms and m² is their height in metres squared).

Tools:

Data analysis in pandas.

Python visualization libraries (matplotlib).

The MVP goal:

Explore and analyze the data.

find the average stroke rate by gender, smoking status, heart disease, type of residence, age, hypertension and ever married.