Stroke Analysis

21/6/2021

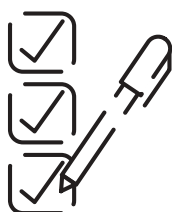## BRIEF

The project "Stroke Analysis" includes the following:

1- Cleaning & tidying of the dataset "healthcare-dataset-stroke-data.csv".
2- Analysis & visualization of the different features illustrating their impact on strokes.
3- A binary logistic model that predicts a stroke based on 3 clinical features: "Age", "Average glucose level", "Body Mass Index (BMI)".

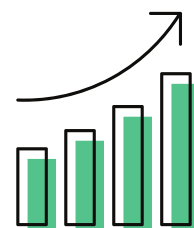|	id	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	9046	Male	67.0	0	1	Yes	Private	Urban	228.69	36.6	formerly smoked	1
1	51676	Female	61.0	0	0	Yes	Self-employed	Rural	202.21	NaN	never smoked	1
2	31112	Male	80.0	0	1	Yes	Private	Rural	105.92	32.5	never smoked	1
3	60182	Female	49.0	0	0	Yes	Private	Urban	171.23	34.4	smokes	1
4	1665	Female	79.0	1	0	Yes	Self-employed	Rural	174.12	24.0	never smoked	1

1 Dataset

5110 observations

12 attributes
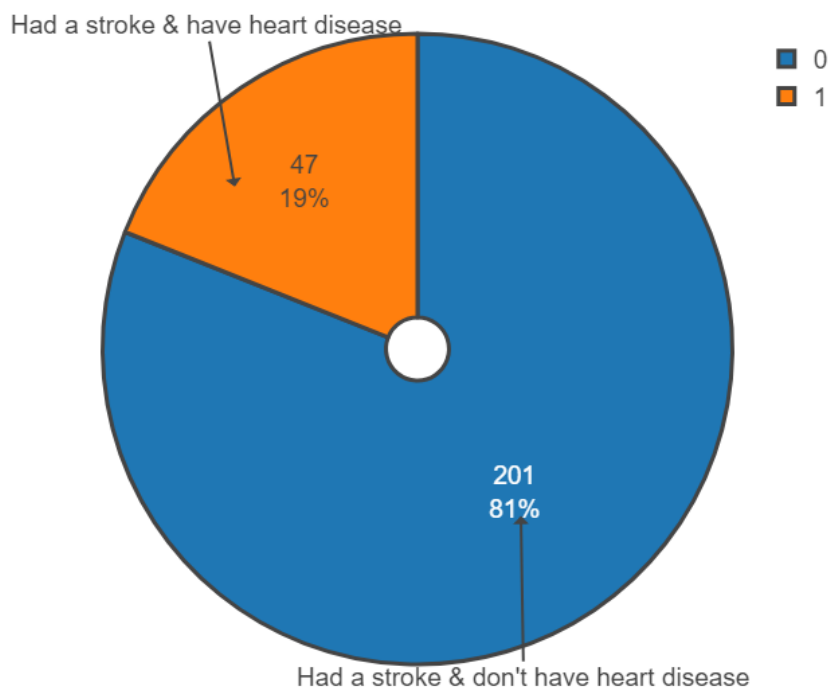
Charts

Stroke according to age category

Adult
35.1%

Senior
64.1%

Child
0.806%

Youth
0%

- Senior
- Adult
- Child
- Youth

Stroke Vs Gender

Male
43.5%

Female
56.5%

- Female
- Male
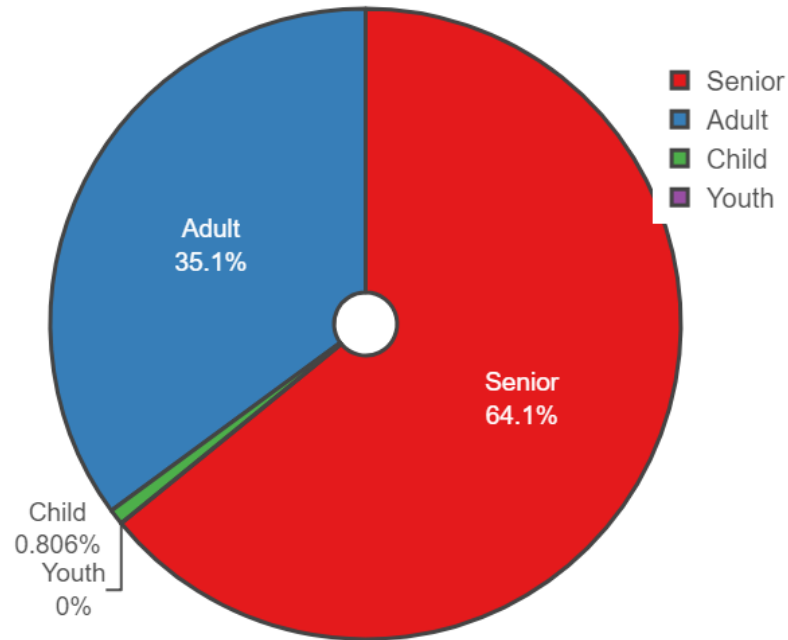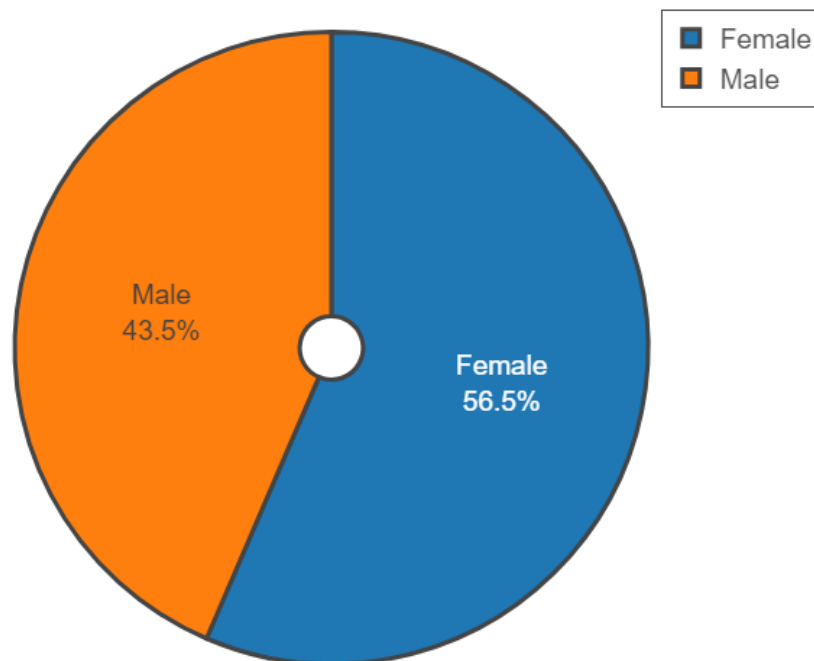
Charts

Stroke Vs Heart disease

Had a stroke & have heart disease

47
19%

0
1

201
81%

Had a stroke & don't have heart disease

Stroke Vs Marital status

Had a stroke & not married

No
29
11.7%

Yes
No

Yes
219
88.3%

Had a stroke & married
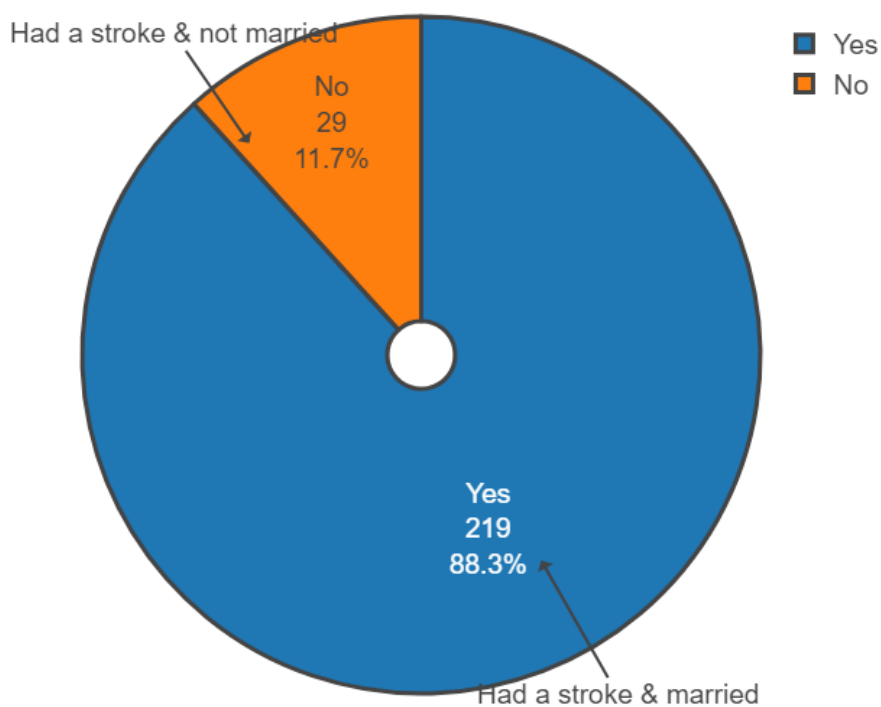
Charts

Stroke Vs Hypertension

Had a stroke and hypertensive

66
26.6%

182
73.4%

Had a stroke & not hypertensive

- 0
- 1
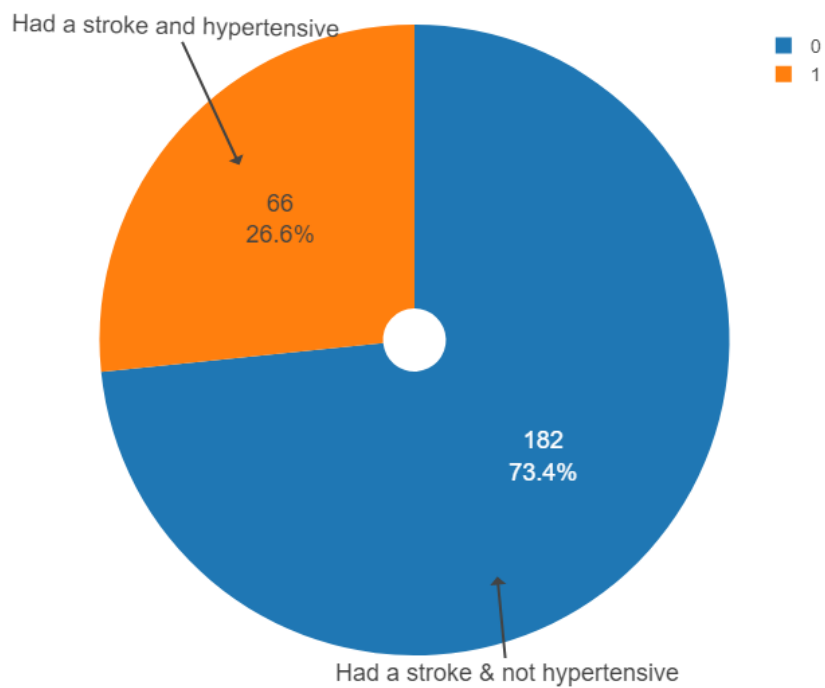
Stroke Vs Residence type

- Urban
- Rural

Rural
113
45.6%

Urban
135
54.4%
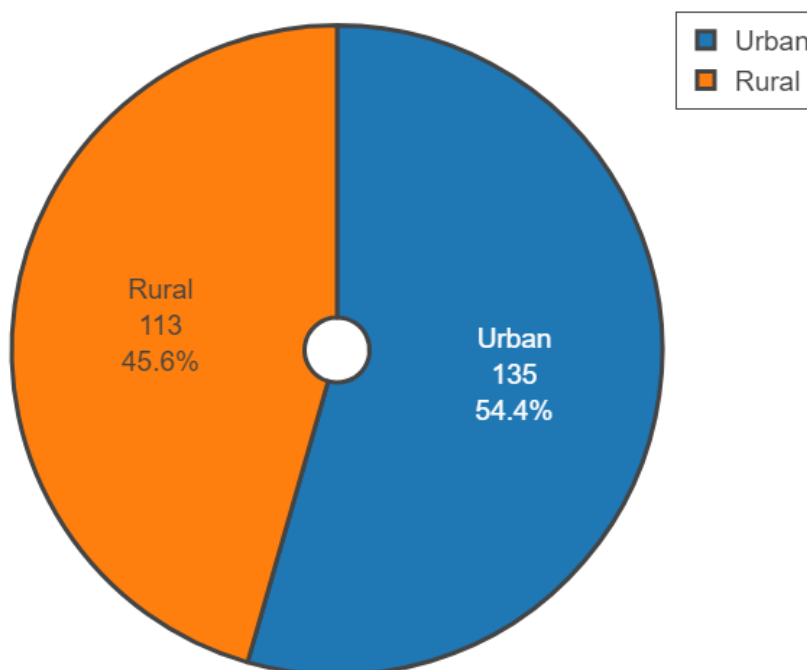
Charts

Stroke Vs Smoking status

- never smoked
- formerly smoked
- Unknown
- smokes

formerly smoked
70
28.2%

never smoked
89
35.9%

Unknown
47
19%

smokes
42
16.9%

Stroke Vs Work type

Work type

Never_worked

Govt_job

Self-employed

Private

0   1

Stroke