

Stroke Data Set

Group D

Aya Abouelela

Amel Khirreddine

Kwaku Asamoah Gyimah

Arlizze Faye R. Ongchua

Supervisor: Prof. Elnaz Gholipiour

Goal of the project: cleaning the data and identifying the contributing factors that lead to having a stroke.

	gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0	Male	58.0	1.0	0.0	Yes	Private	Urban	87.96	39.2	never smoked	0.0
1	Female	70.0	0.0	0.0	Yes	Private	Rural	69.04	35.9	formerly smoked	0.0
2	Female	52.0	0.0	0.0	Yes	Private	Urban	77.59	17.7	formerly smoked	0.0
3	Female	75.0	0.0	1.0	Yes	Self-employed	Rural	243.53	27.0	never smoked	0.0
4	Female	32.0	0.0	0.0	Yes	Private	Rural	77.67	32.3	smokes	0.0
...
29060	Female	10.0	0.0	0.0	No	children	Urban	58.64	20.4	never smoked	0.0
29061	Female	56.0	0.0	0.0	Yes	Govt_job	Urban	213.61	55.4	formerly smoked	0.0
29062	Female	82.0	1.0	0.0	Yes	Private	Urban	91.94	28.9	formerly smoked	0.0
29063	Male	40.0	0.0	0.0	Yes	Private	Urban	99.16	33.2	never smoked	0.0
29064	Female	82.0	0.0	0.0	Yes	Private	Urban	79.48	20.6	never smoked	0.0

Data Cleaning

- Removing irrelevant data:

e.g. children which have a very rare chance of getting a stroke

How?

Remove `work_type = 'children'`

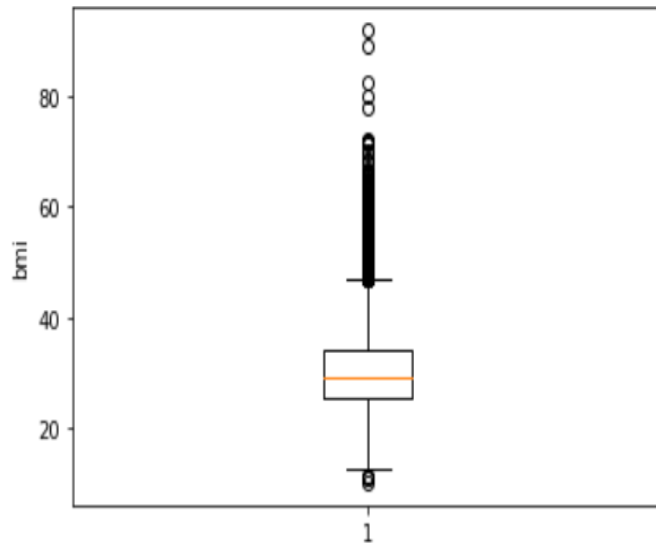
Remove `age <= 17`

- No duplicated data

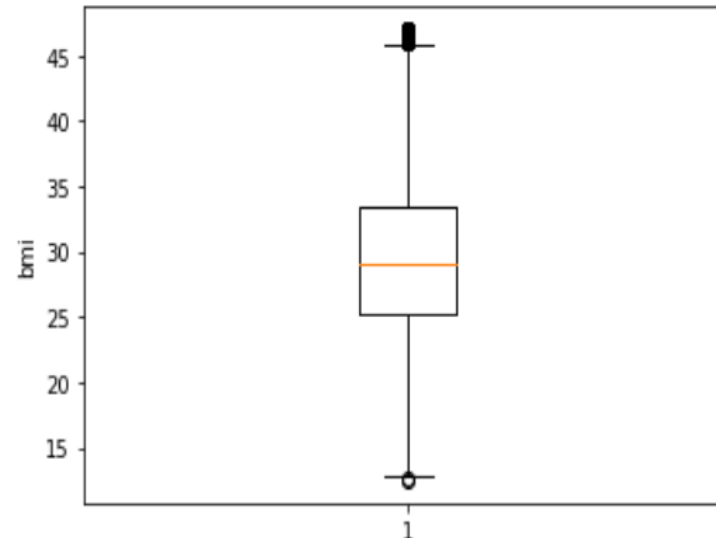
Data Cleaning

- Removing outliers: BMI

$10 \leq \text{bmi} \leq 92$

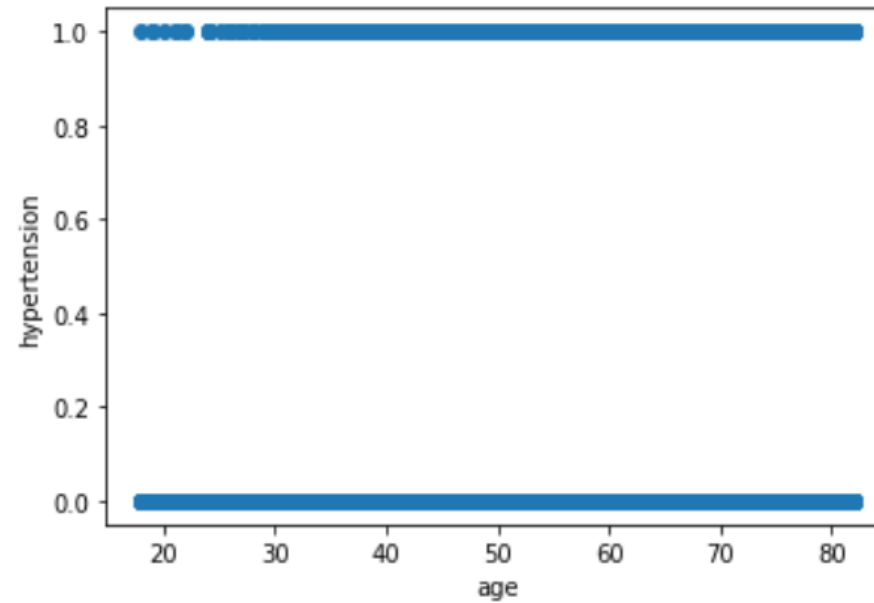
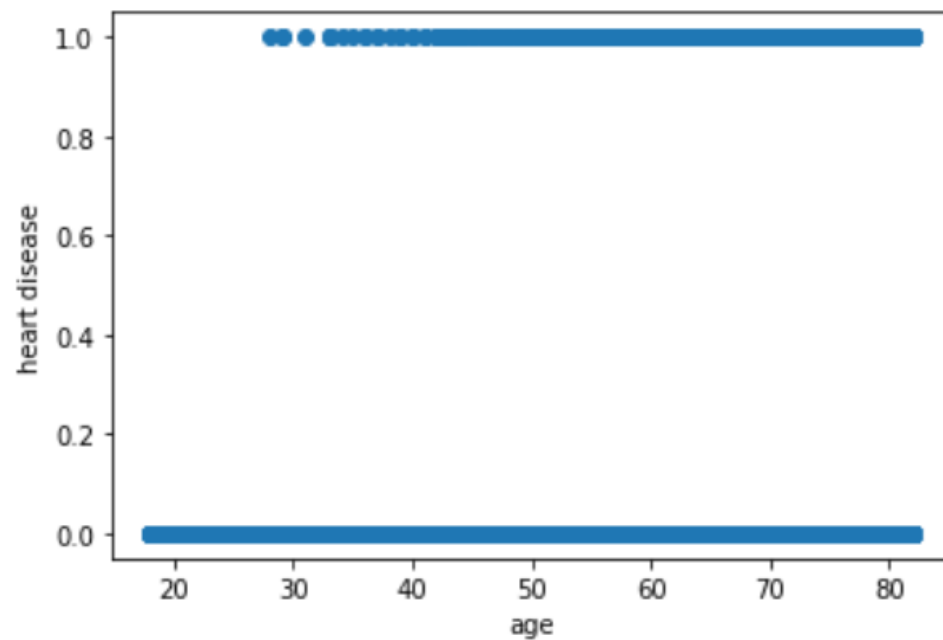


$12.5 \leq \text{bmi} \leq 47$



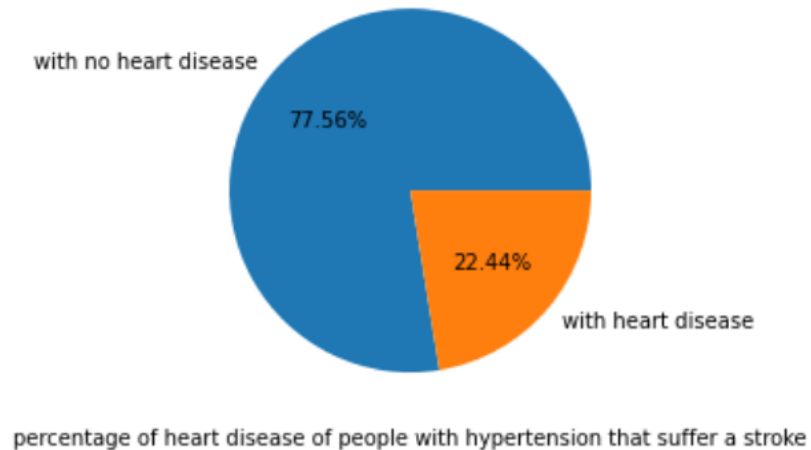
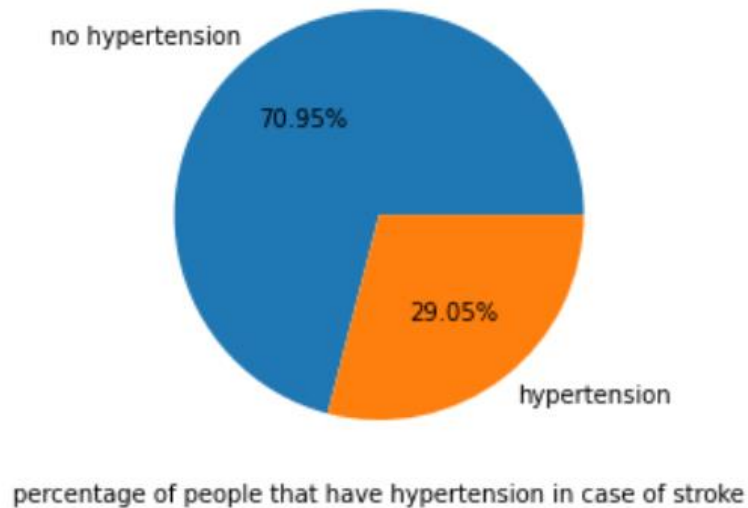
Analysis of the data

- Checking for realistic correlations



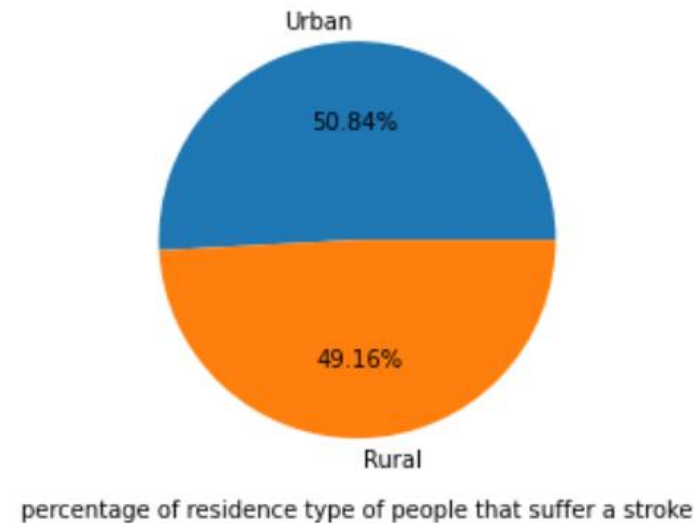
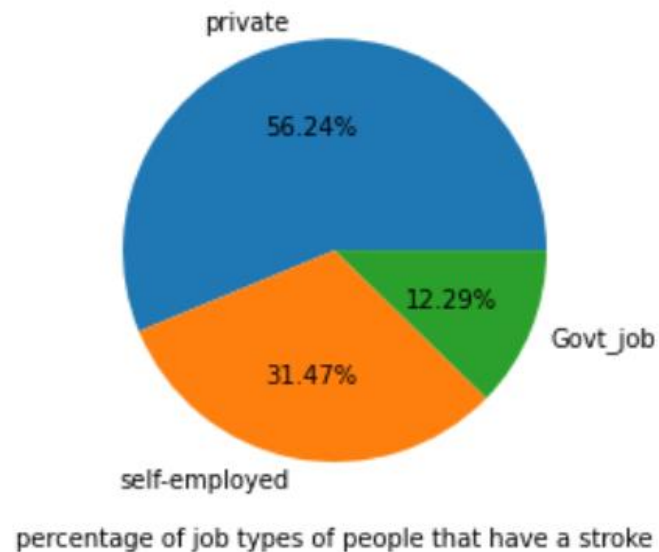
Findings:

- Does having hypertension w/o heart disease increase the chances of getting a stroke?



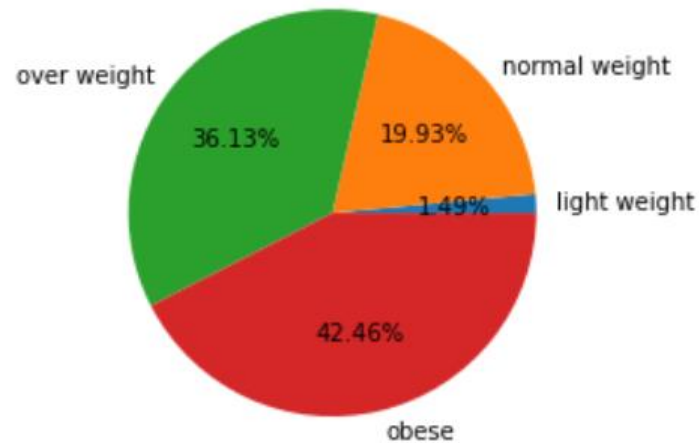
Findings:

- Are the work type and living environment contributing factors to getting a stroke?



Findings:

- Does having diabetes/ being overweight increase the chance of getting a stroke?



bmi chart for people that suffer a stroke

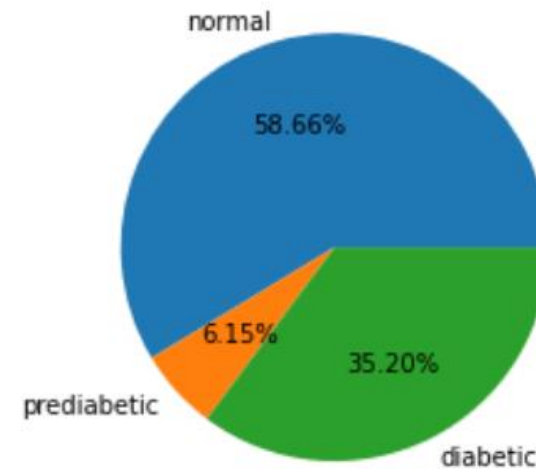
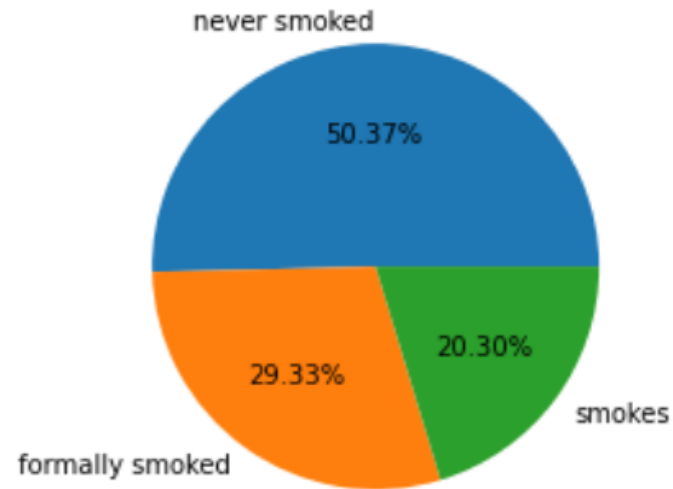


chart of the average glucose levels for people with a stroke

Findings:

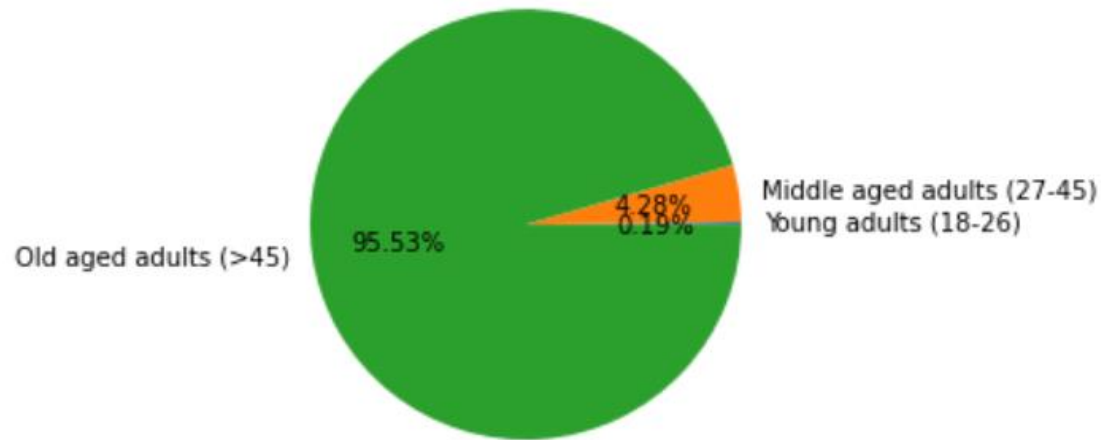
- Does smoking while having hypertension make you more susceptible to having a stroke?



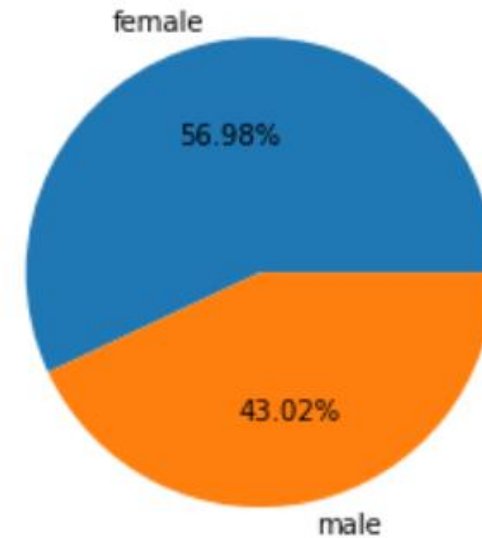
percentage of smoking status of people with hypertension that suffer a stroke

Findings:

- Which gender and age group more susceptible to having a stroke?



percentage of age groups of people with a stroke



percentage of gender for people with a stroke

Conclusion

- The data sample shows that the percentage of people that suffer a stroke is 1.8% of the full sample.
- Of those suffering a stroke, smoking, having hypertension and heart disease don't necessarily cause the stroke.
- The probability of having a stroke increases with age and with having a high BMI.
- It is not conclusive to state that females are more prone to having a stroke than men.
- Having a private job/being self-employed might contribute more to causing a stroke due to stress than having a government job.