

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
Male	58.0	1.0	0.0	Yes	Private	Urban	87.96	39.2	never smoked	0.0
Female	70.0	0.0	0.0	Yes	Private	Rural	69.04	35.9	formerly smoked	0.0
Female	52.0	0.0	0.0	Yes	Private	Urban	77.59	17.7	formerly smoked	0.0
Female	75.0	0.0	1.0	Yes	Self-employed	Rural	243.53	27.0	never smoked	0.0
Female	32.0	0.0	0.0	Yes	Private	Rural	77.67	32.3	smokes	0.0
Female	10.0	0.0	0.0	No	children	Urban	58.64	20.4	never smoked	0.0
Female	56.0	0.0	0.0	Yes	Govt_job	Urban	213.61	55.4	formerly smoked	0.0
Female	82.0	1.0	0.0	Yes	Private	Urban	91.94	28.9	formerly smoked	0.0
Male	40.0	0.0	0.0	Yes	Private	Urban	99.16	33.2	never smoked	0.0
Female	82.0	0.0	0.0	Yes	Private	Urban	79.48	20.6	never smoked	0.0
	Male Female Female Female Female Female Female Female Female Male	Male 58.0 Female 70.0 Female 52.0 Female 75.0 Female 32.0 Female 10.0	Male 58.0 1.0 Female 70.0 0.0 Female 52.0 0.0 Female 75.0 0.0 Female 32.0 0.0 Female 10.0 0.0 Female 56.0 0.0 Female 82.0 1.0 Male 40.0 0.0	Male 58.0 1.0 0.0 Female 70.0 0.0 0.0 Female 52.0 0.0 0.0 Female 75.0 0.0 1.0 Female 32.0 0.0 0.0 Female 10.0 0.0 0.0 Female 56.0 0.0 0.0 Female 82.0 1.0 0.0 Male 40.0 0.0 0.0	Male 58.0 1.0 0.0 Yes Female 70.0 0.0 0.0 Yes Female 52.0 0.0 0.0 Yes Female 75.0 0.0 1.0 Yes Female 32.0 0.0 0.0 Yes Female 10.0 0.0 0.0 No Female 56.0 0.0 0.0 Yes Female 82.0 1.0 0.0 Yes Male 40.0 0.0 0.0 Yes	Male 58.0 1.0 0.0 Yes Private Female 70.0 0.0 0.0 Yes Private Female 52.0 0.0 0.0 Yes Private Female 75.0 0.0 1.0 Yes Self-employed Female 32.0 0.0 0.0 Yes Private Female 10.0 0.0 0.0 No children Female 56.0 0.0 0.0 Yes Govt_job Female 82.0 1.0 0.0 Yes Private Male 40.0 0.0 0.0 Yes Private	Male 58.0 1.0 0.0 Yes Private Urban Female 70.0 0.0 0.0 Yes Private Rural Female 52.0 0.0 0.0 Yes Private Urban Female 75.0 0.0 1.0 Yes Self-employed Rural Female 32.0 0.0 0.0 Yes Private Rural Female 10.0 0.0 0.0 No children Urban Female 56.0 0.0 0.0 Yes Govt_job Urban Female 82.0 1.0 0.0 Yes Private Urban Male 40.0 0.0 0.0 Yes Private Urban	Male 58.0 1.0 0.0 Yes Private Urban 87.96 Female 70.0 0.0 0.0 Yes Private Rural 69.04 Female 52.0 0.0 0.0 Yes Private Urban 77.59 Female 75.0 0.0 1.0 Yes Self-employed Rural 243.53 Female 32.0 0.0 0.0 Yes Private Rural 77.67 Female 10.0 0.0 0.0 No children Urban 58.64 Female 56.0 0.0 0.0 Yes Govt_job Urban 213.61 Female 82.0 1.0 0.0 Yes Private Urban 91.94 Male 40.0 0.0 0.0 Yes Private Urban 99.16	Male 58.0 1.0 0.0 Yes Private Urban 87.96 39.2 Female 70.0 0.0 0.0 Yes Private Rural 69.04 35.9 Female 52.0 0.0 0.0 Yes Private Urban 77.59 17.7 Female 75.0 0.0 1.0 Yes Self-employed Rural 243.53 27.0 Female 32.0 0.0 0.0 Yes Private Rural 77.67 32.3 Female 10.0 0.0 0.0 No children Urban 58.64 20.4 Female 56.0 0.0 0.0 Yes Govt_job Urban 213.61 55.4 Female 82.0 1.0 0.0 Yes Private Urban 91.94 28.9 Male 40.0	Male 58.0 1.0 0.0 Yes Private Urban 87.96 39.2 never smoked Female 70.0 0.0 0.0 Yes Private Rural 69.04 35.9 formerly smoked Female 52.0 0.0 0.0 Yes Private Urban 77.59 17.7 formerly smoked Female 75.0 0.0 1.0 Yes Self-employed Rural 243.53 27.0 never smoked Female 32.0 0.0 0.0 Yes Private Rural 77.67 32.3 smokes

Data Cleaning

Removing irrelevant data:

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

How?

Remove work_type = 'children'

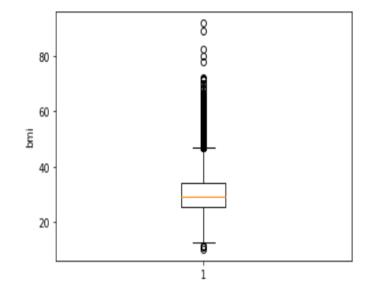
Remove age <= 17

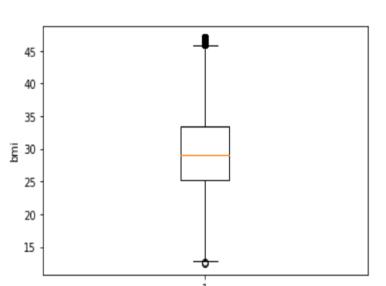
No duplicated data

Data Cleaning

Removing outliers: BMI







BMI CHART



		120	130	140	150	160	170	180	190	200	210	220	230	240	250
Height in feet and inches	4.6	29		34											60
	4.8			31	34									54	56
	4.10					34	36				44		48		52
	5.0														49
	5.2						32				38			44	46
	5.4							31		34	36				43
	5.6								31		34				40
	5.8									30		34			38
	5.10														36
	6.0											30			34
	6.2												30		32
	6.4														30
	6.6														29
	6.8	13	14	15	17	18	19	20	21	22	23	24	25	26	28



Jnderweigh

Being underweight could be a sign you'n not eating enough or you may be ill. If you're underweight, a GP can help.



Norma

Keep up the good work! For tips on maintaining a healthy weight, check out the food and diet and fitness sections.



Overwei

The best way to lose weight if you're overweight is through a combination of diet and exercise.

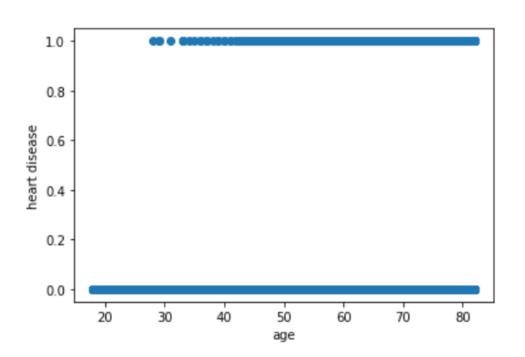


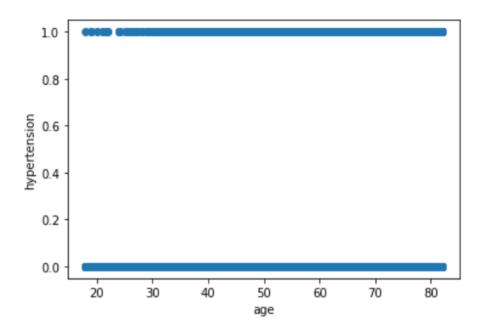
Oher

The best way to lose weight if you're obese is through a combination of diet and exercise,

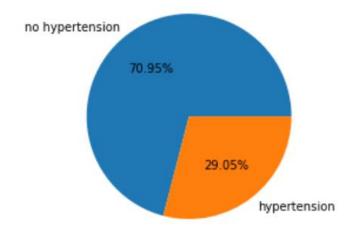
Analysis of the data

Checking for realistic correlations

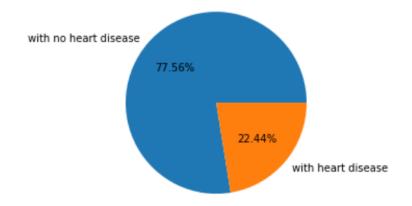




Does having hypertension w/o heart disease increase the chances of getting a stroke?

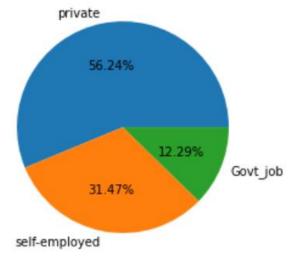


percentage of people that have hypertension in case of stroke

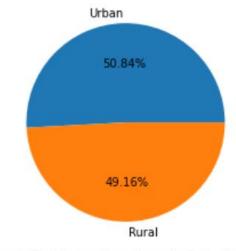


percentage of heart disease of people with hypertension that suffer a stroke

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

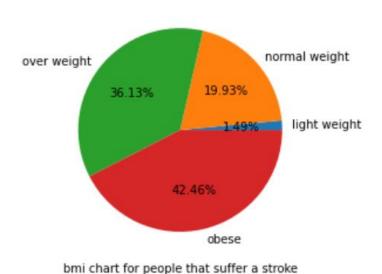


percentage of job types of people that have a stroke



percentage of residence type of people that suffer a stroke

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



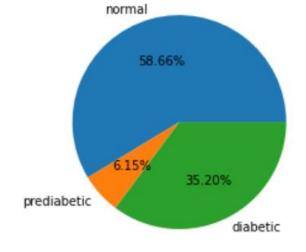
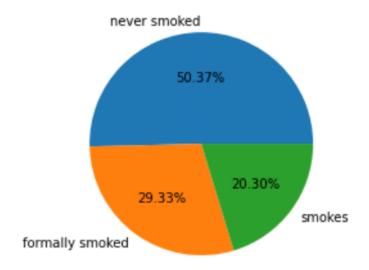


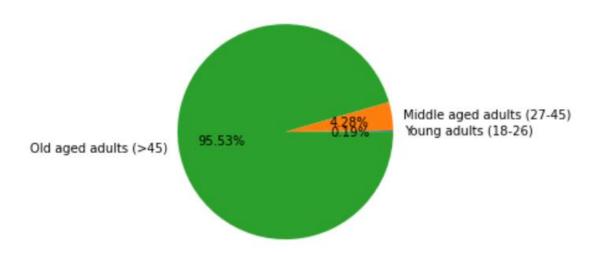
chart of the average glucose levels for people with a stroke

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

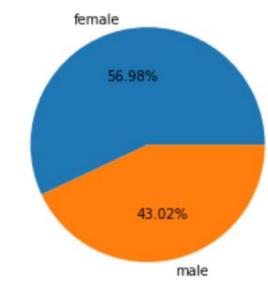


percentage of smoking status of people with hypertension that suffer a stroke

• 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.