Analysis of the risk to make stroke

Azzah Ahmed

01

Introduction

02

Data cleaning

03

Exploratory Data Analysis

04

Modeling

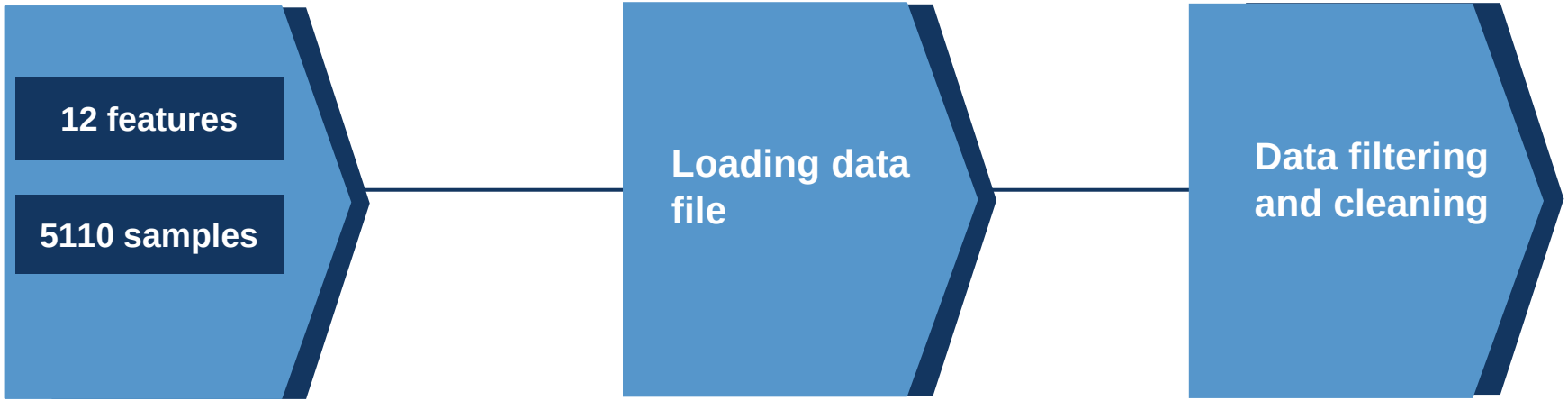
Introduction

- In this project we look for the factors that increase the risk of strokes and analysis them, like gender, Hypertension, age, Heart disease, smoking .

Dataset

From

kaggle



id	gender	age	hypertension	Heart disease	married	Work type	Residence type	Glucose level	bmi	smoking	stroke
9046	1	67.0	0	1	1	2	1	228.69	36.6	1	1
31112	1	80.0	0	1	1	2	0	105.92	32.5	2	1
60182	0	49.0	0	0	1	2	1	171.23	34.4	3	1

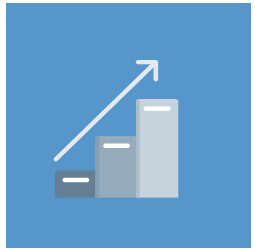
QUESTIONS

2

Does gender or age affect stroke?

1

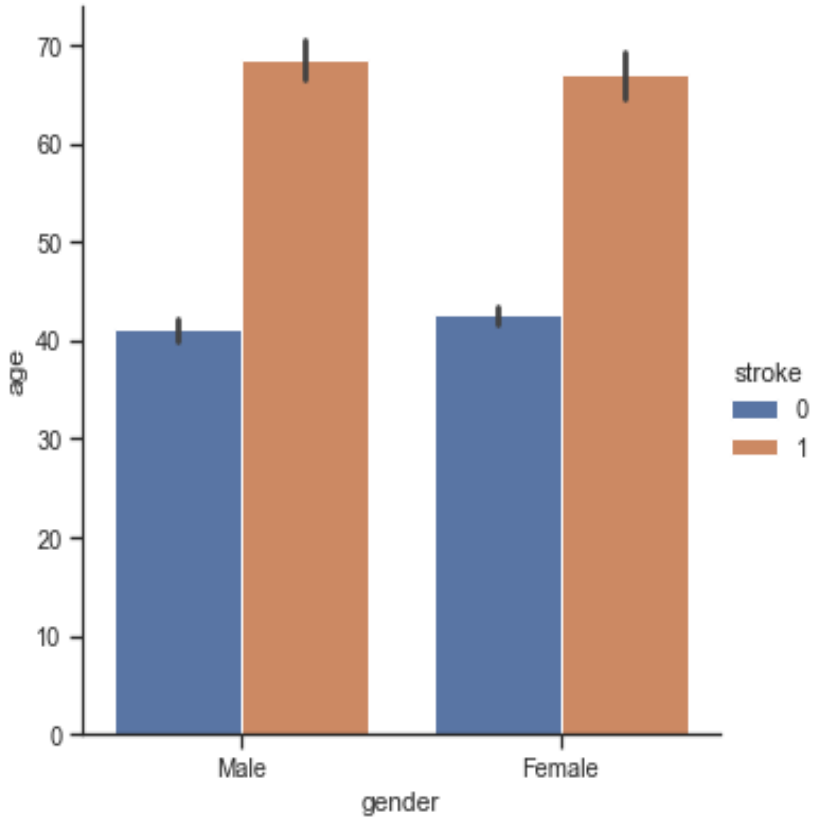
Does the person have a stroke or not?

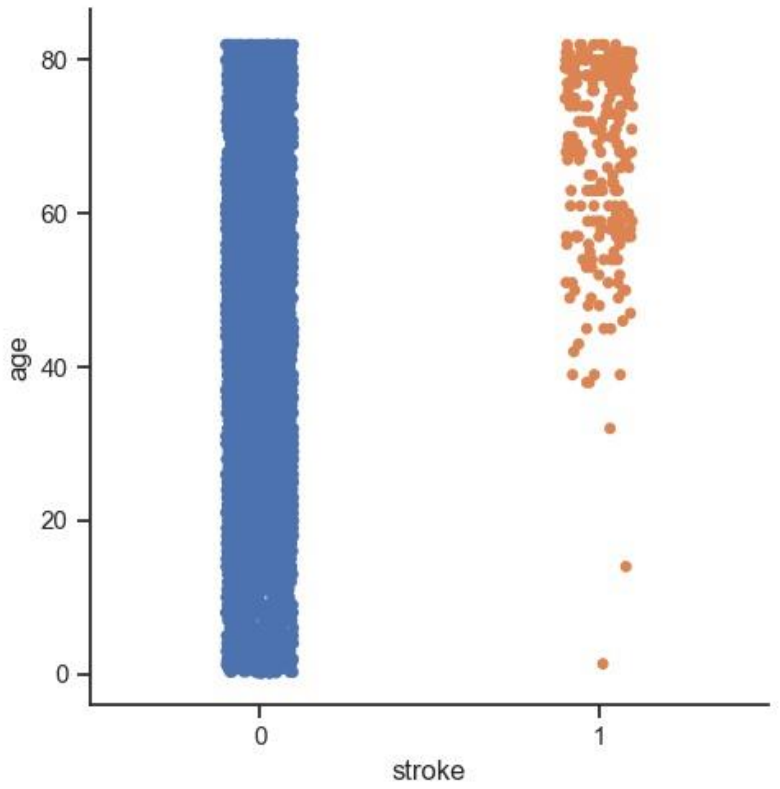


3

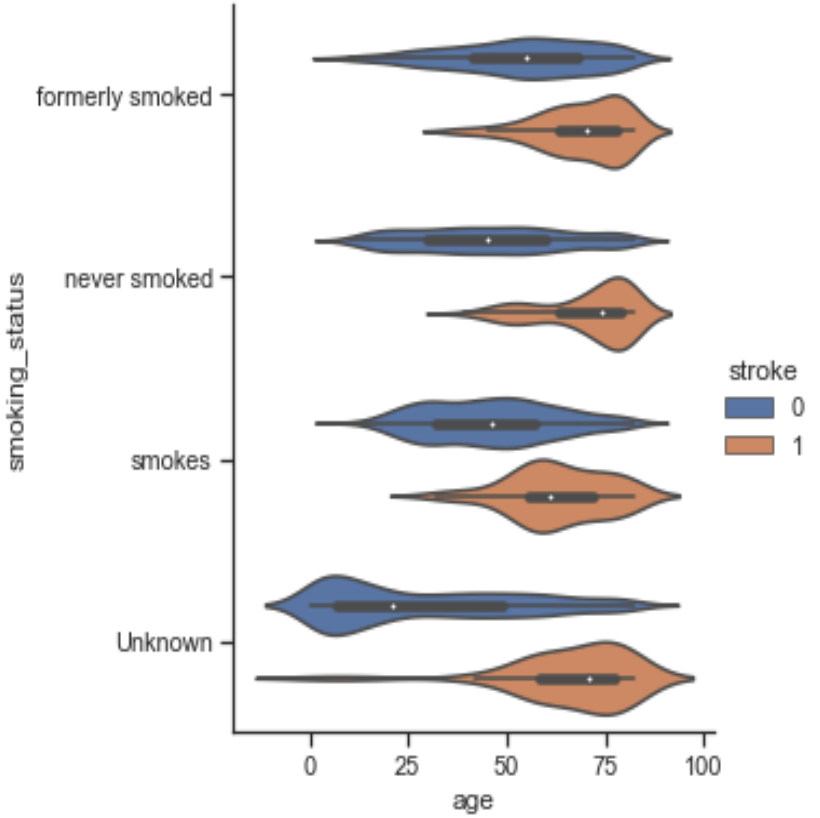
What is the most effective factor on stroke?

Exploratory Data Analysis

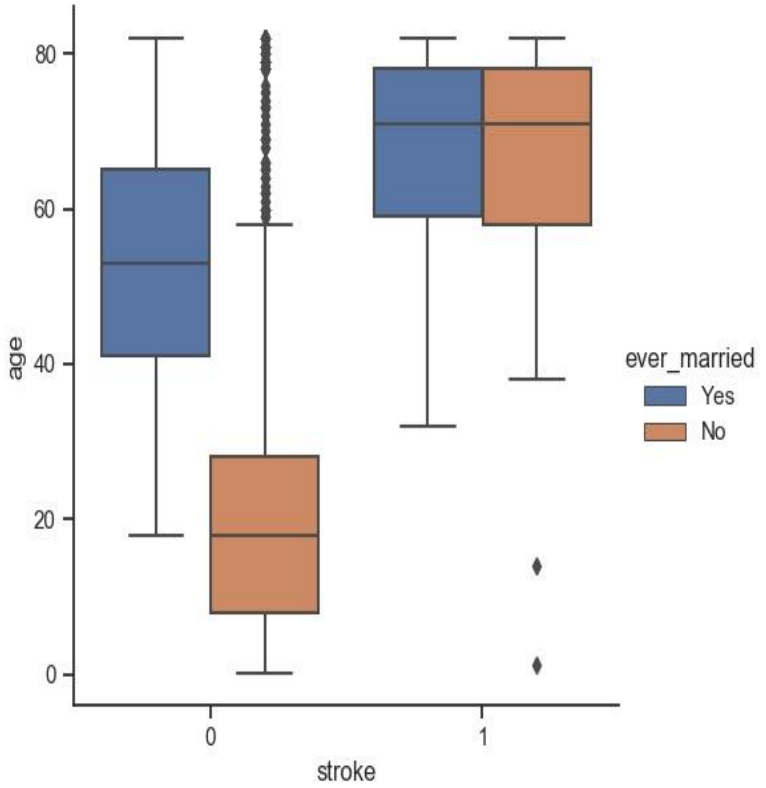




Exploratory Data Analysis



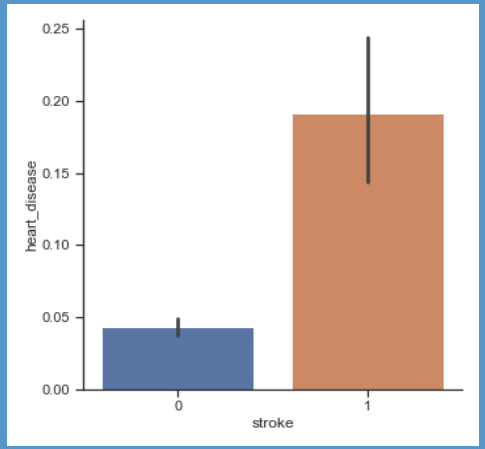
smoking, age, and stroke



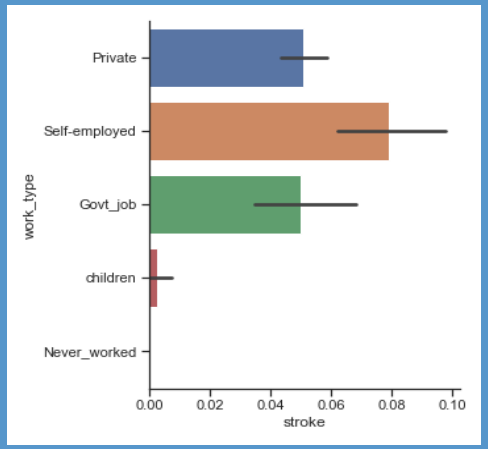
marriage, age, and stroke.

Does work type, hypertension , glucose level , bmi , have any relation with stroke?

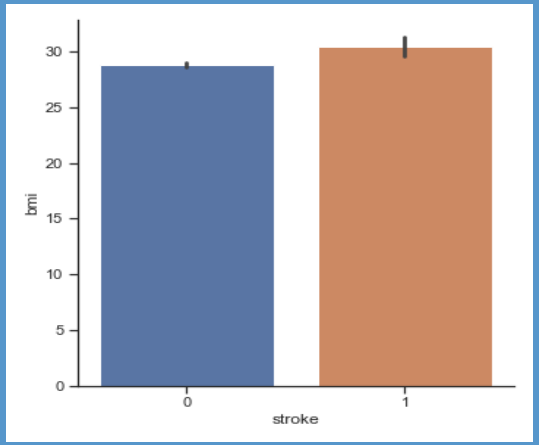
Heart disease& smoking



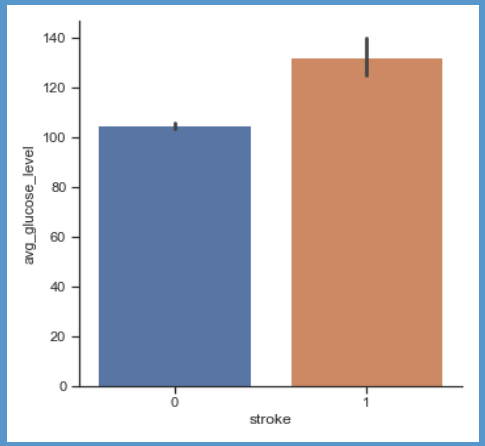
Work type & stroke



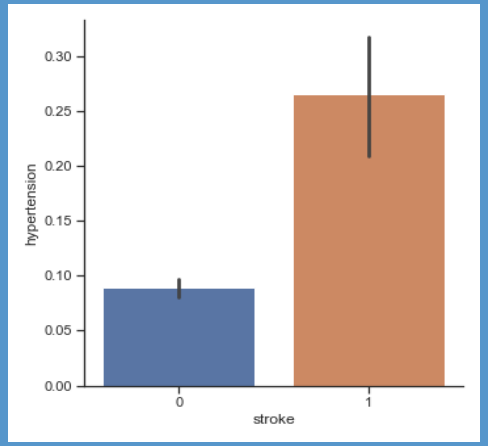
bmi& stroke



glucose level & stroke



hypertension & stroke



Models

I used Decision tree
&
Random Forest
Classifier

The table shows the comparison between the two models

	Model	Accuracy Score
1	Random Forest	0.946029
2	Decision tree	0.909369

The Random Forest Classifier accurate more than Decision tree

Thank You