Niloofar seyf

Major: Master of Medical Engineering (Bioelectric)

Group 2

About this dataset

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide.
Heart failure is a common event caused by CVDs and this dataset contains 12 features that can be used to predict mortality by heart failure.

Most cardiovascular diseases can be prevented by addressing behavioral risk factors such as tobacco use, unhealthy diet and obesity, physical inactivity and harmful use of alcohol using population-wide strategies.

People with cardiovascular disease or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidemia or already established disease) need early detection and management wherein a machine learning model can be of great help.

Dataset features:

Age

Smoking

If the patient smokes (0 = No, 1 = Yes)

anaemia

Decrease of red blood cells or hemoglobin (Boolean)

creatinine phosphokinase

Level of the CPK enzyme in the blood (mcg/L)

diabetes

If the patient has diabetes (Boolean)

ejection fraction

Percentage of blood leaving the heart at each contraction (percentage)

high_blood_pressure

If the patient has hypertension (boolean)

platelets

Platelets in the blood (kiloplatelets/mL)

serum_creatinine

Level of serum creatinine in the blood (mg/dL)

serum_sodium

Level of serum sodium in the blood (mEq/L)

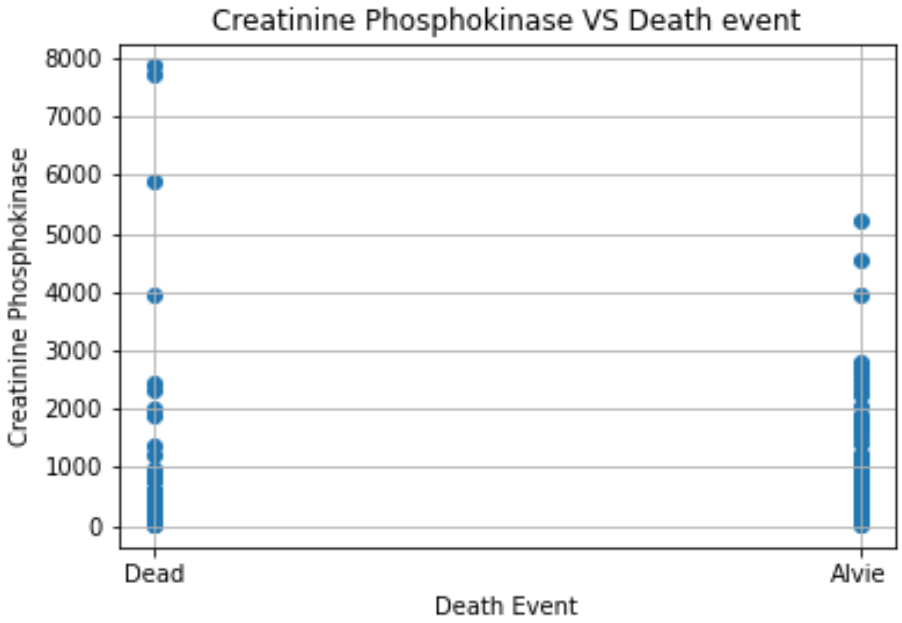
sex

Woman or man (binary)

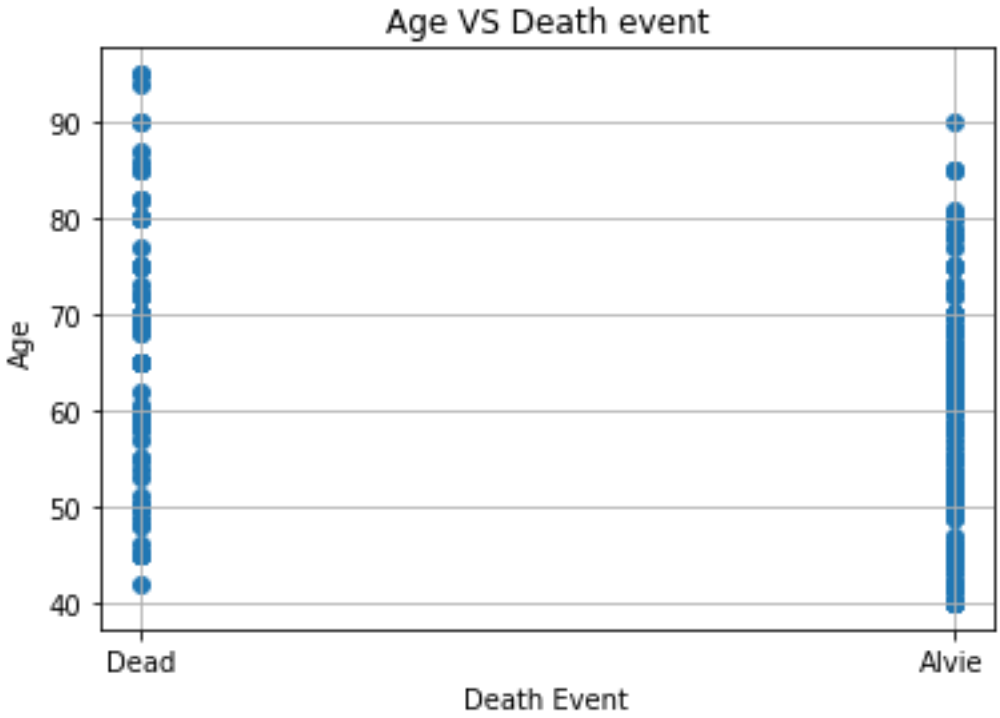
```
data = pd.read_csv("heart_failure_clinical_records_dataset.csv")
data
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelets	serum_creatinine	serum_sodium	sex	smoking	tim
0	75.0	0	582	0	20	1	265000.00	1.9	130	1	0	
1	55.0	0	7861	0	38	0	263358.03	1.1	136	1	0	
2	65.0	0	146	0	20	0	162000.00	1.3	129	1	1	
3	50.0	1	111	0	20	0	210000.00	1.9	137	1	0	
4	65.0	1	160	1	20	0	327000.00	2.7	116	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	
294	62.0	0	61	1	38	1	155000.00	1.1	143	1	1	27
295	55.0	0	1820	0	38	0	270000.00	1.2	139	0	0	27
296	45.0	0	2060	1	60	0	742000.00	0.8	138	0	0	27
297	45.0	0	2413	0	38	0	140000.00	1.4	140	1	1	28
298	50.0	0	196	0	45	0	395000.00	1.6	136	1	1	28

299 rows × 13 columns

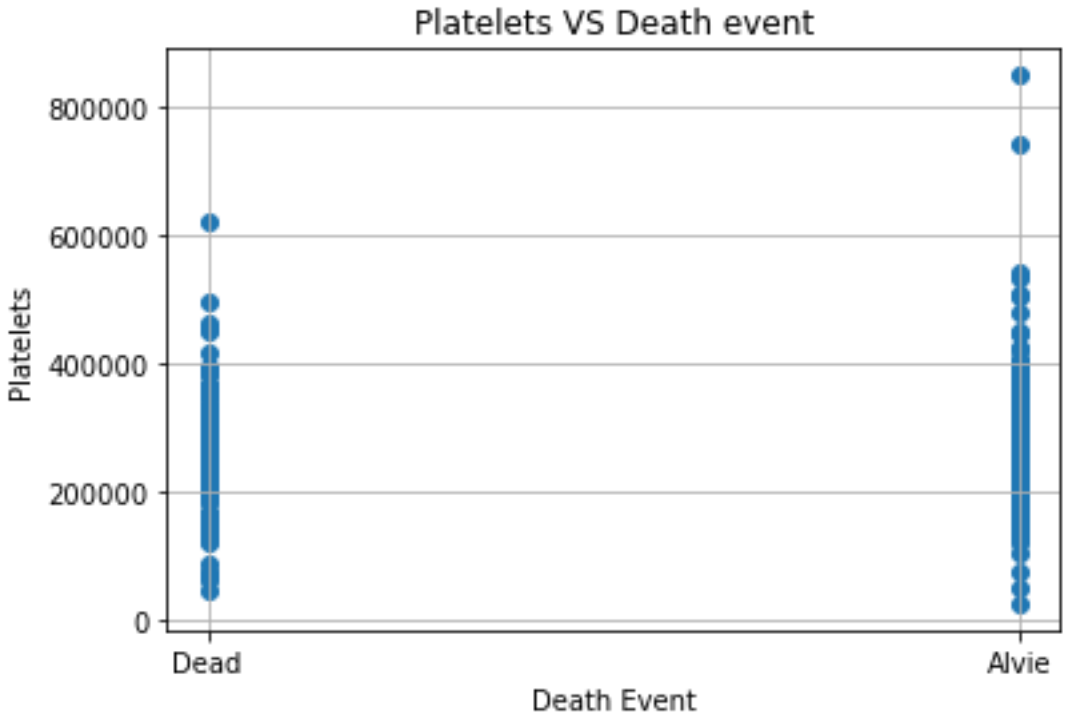


This plot, shows that death rate in subjects with Creatinine phosphokinase above 6000 is more than the people who lived. In other hand it is good for people to keep their Creatinine phosphokinase below 6000 to stay alive.

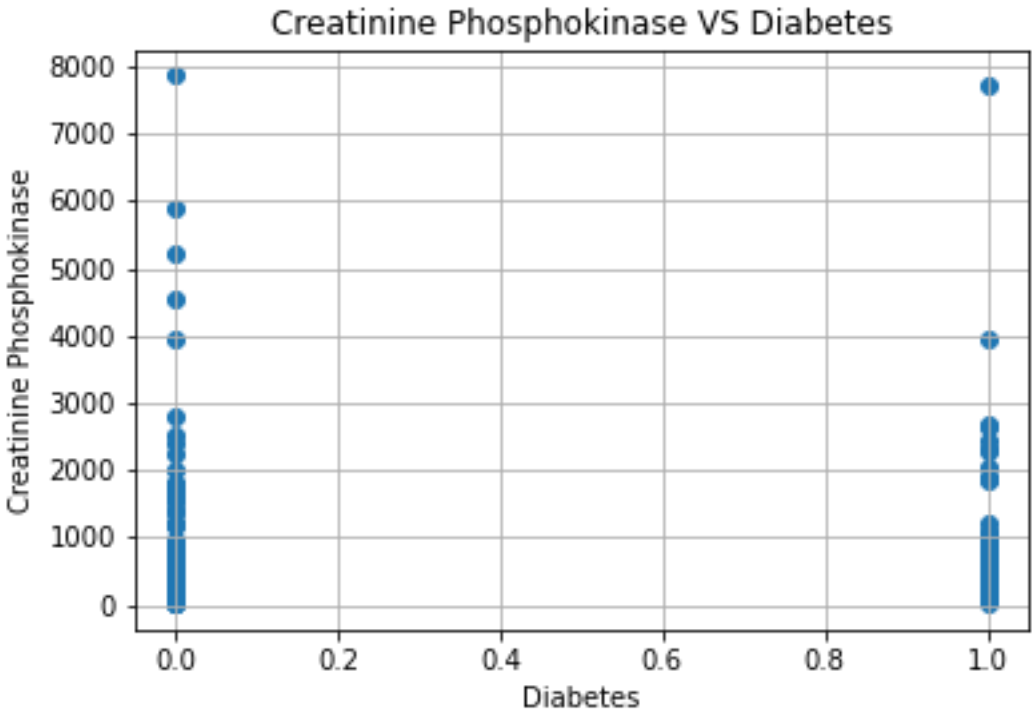


Death rate increased in people at age above 90

.

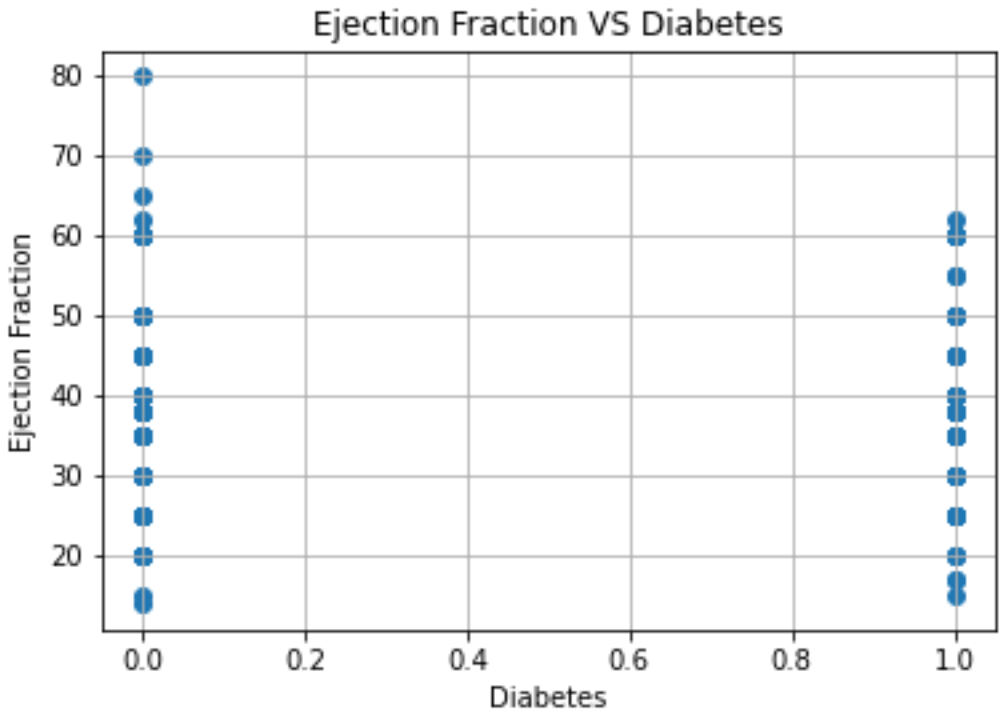


Some people who lived, had platelets more than 7000.



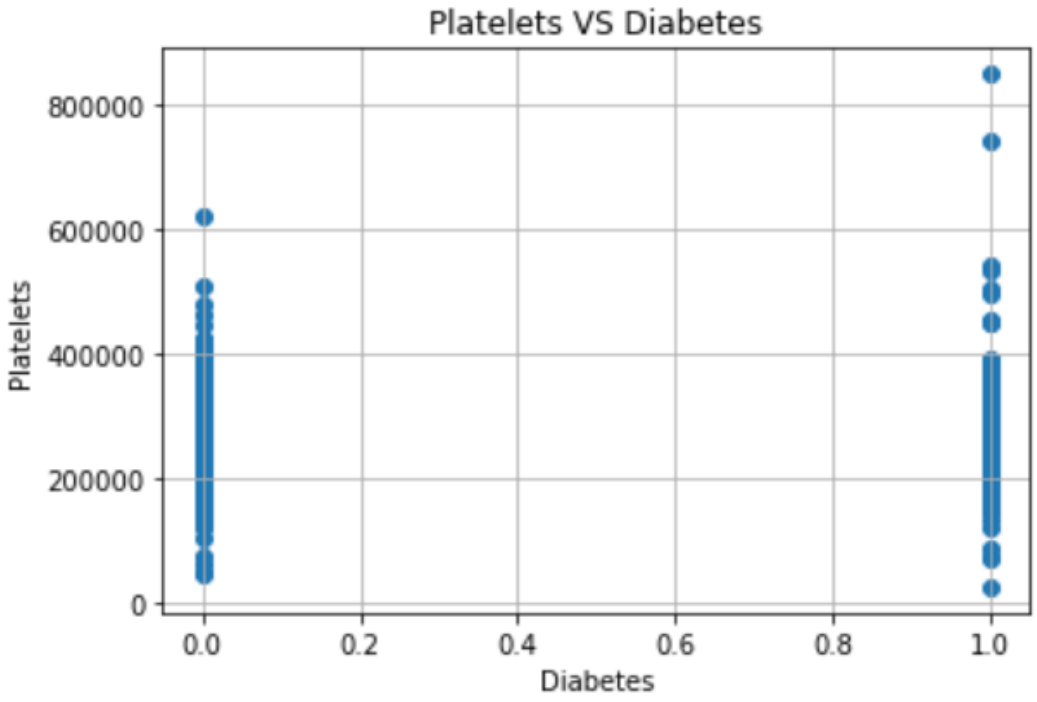
Diabetes: Yes:1, No: 0

For the people who had diabetes, none of them had Creatinine phosphokinase above 4000.



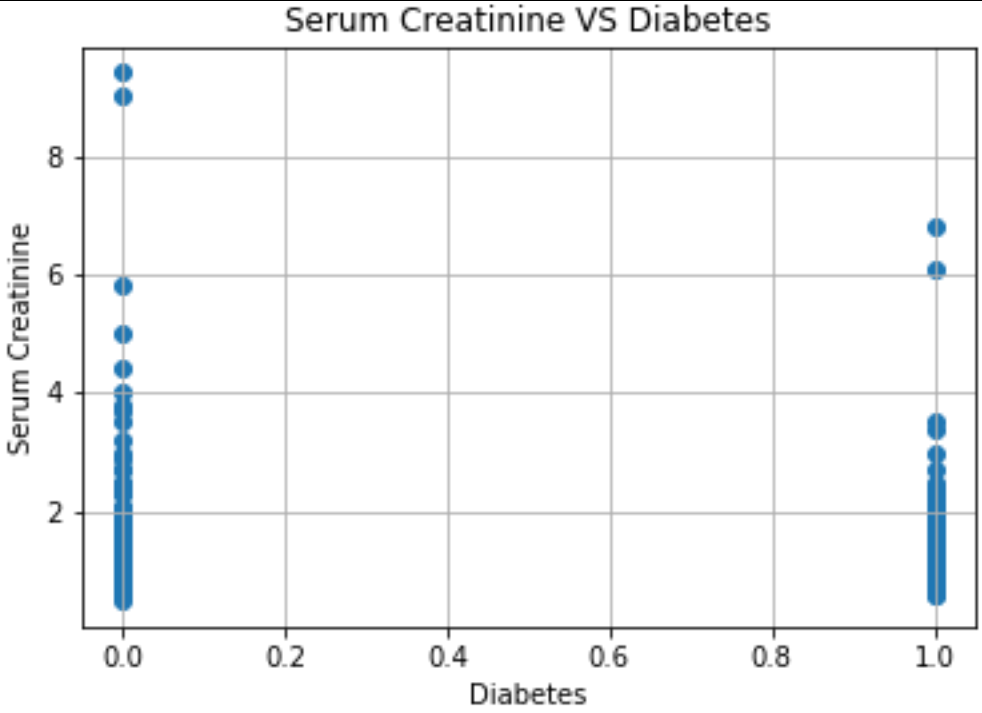
Diabetes: Yes:1, No: 0

For the people who had diabetes, ejection fraction was below 65.



Diabetes: Yes:1, No: 0

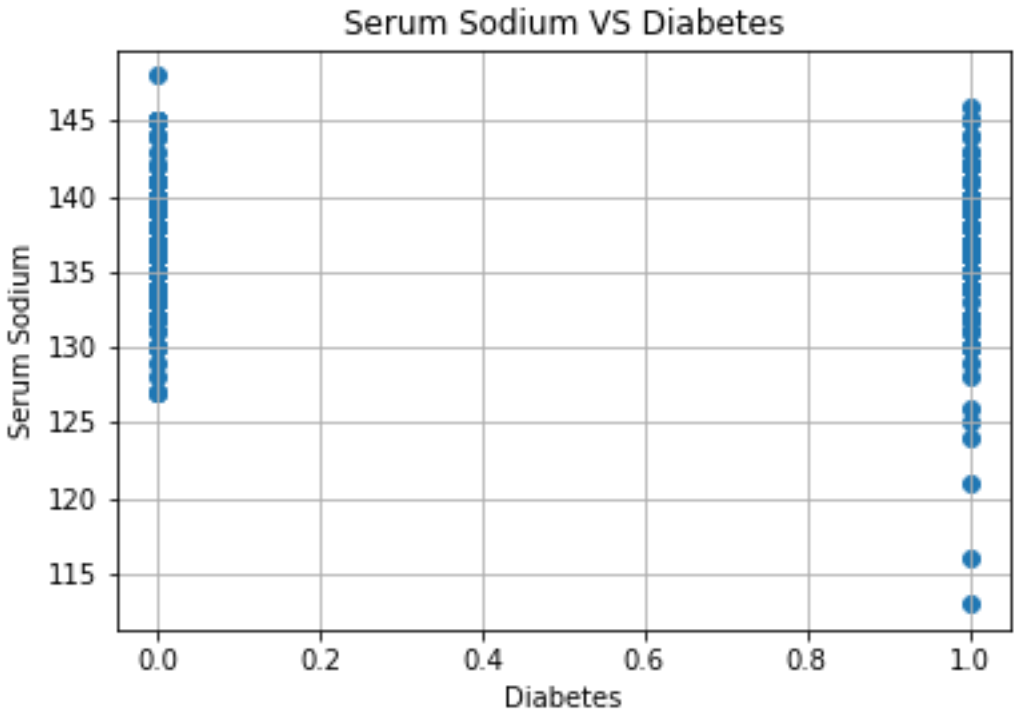
For the people who had diabetes, the platelets were above 7000.



Diabetes: Yes:1, No: 0

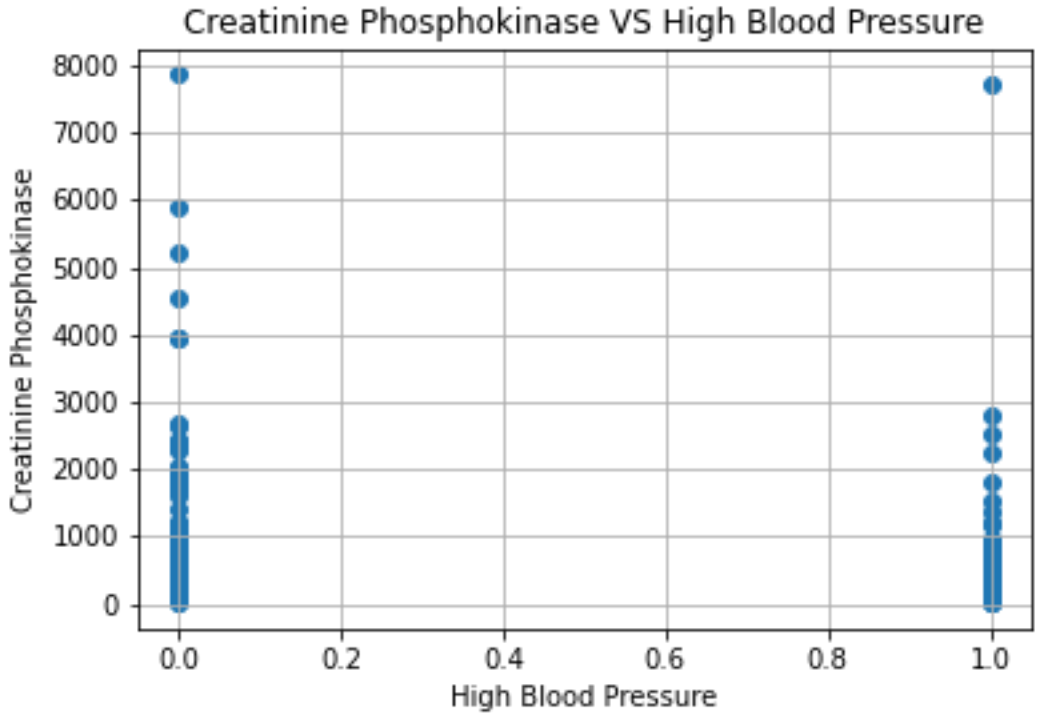
The people who had diabetes, the serum creatinine was above 6.

The people who didn't have diabetes, the serum creatinine was above 8.



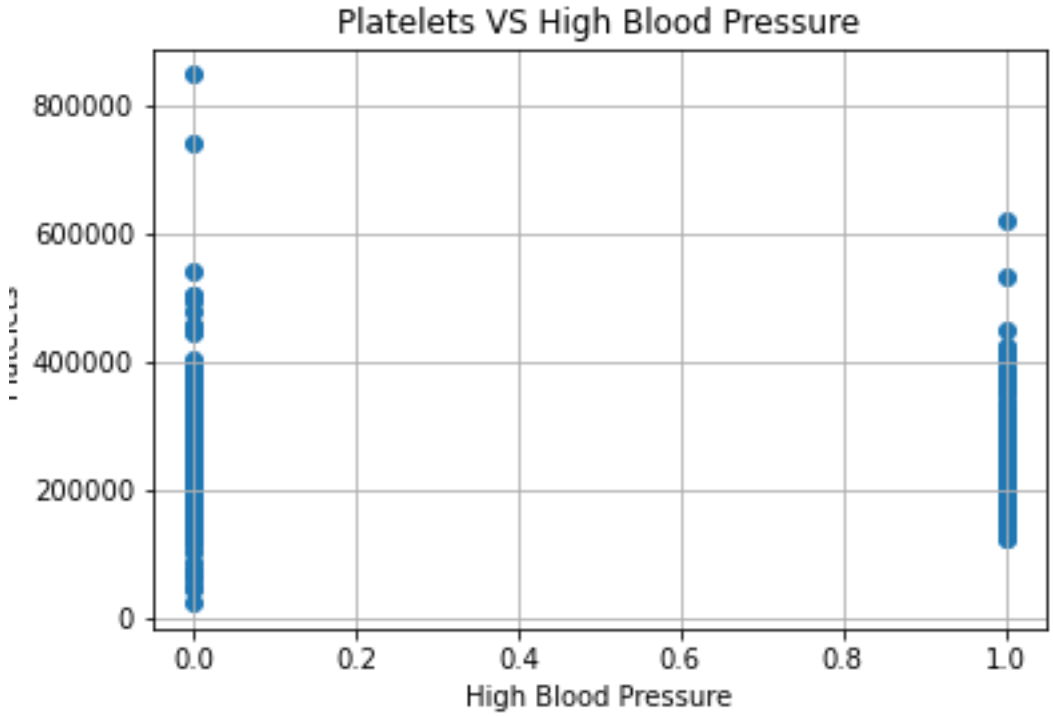
Diabetes: Yes:1, No: 0

For the people who didn't have diabetes, they didn't have serum sodium below 125.



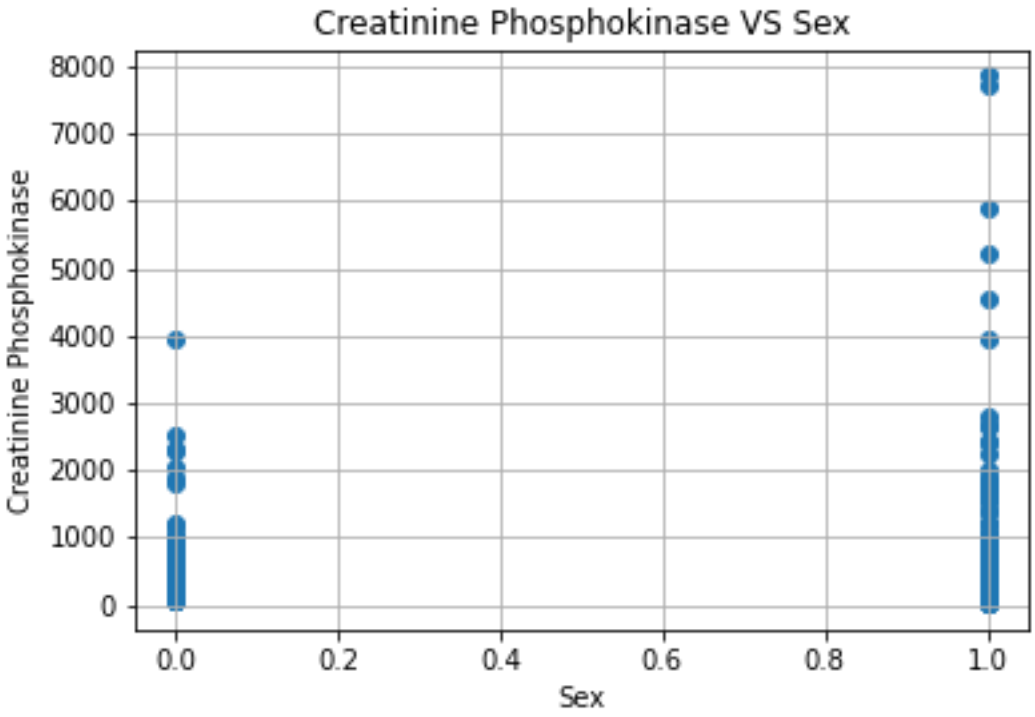
High blood pressure: 1: Yes, 0: No

The people who didn't have high blood pressure, the amount of Creatinine phosphokinase was above 3000.



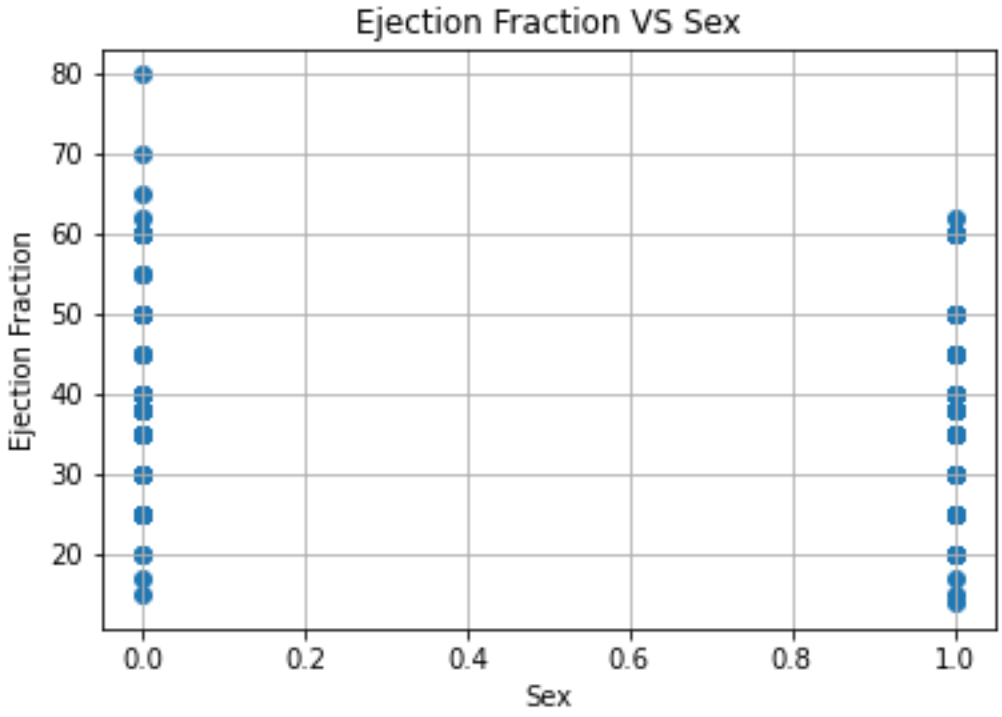
High blood pressure: 1: Yes, 0: No

The people who didn't have high blood pressure, the amount of platelets was above 70000.



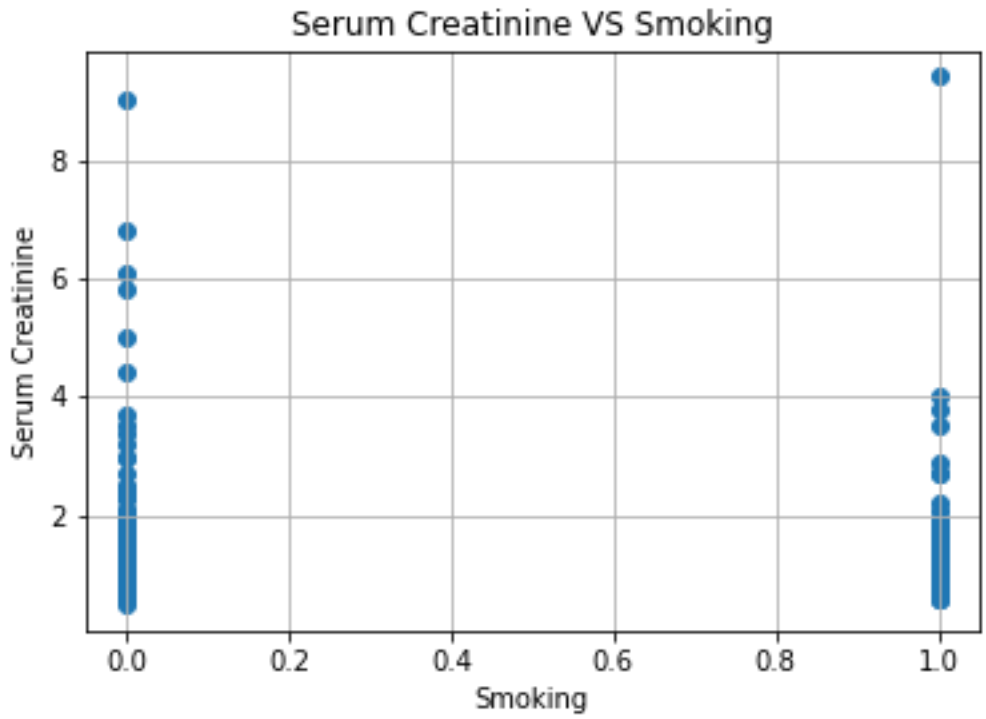
Sex: Female: 0, Male: 1

Women had Creatinine phosphokinase below 4001, but Men had phosphokinase above 4000.



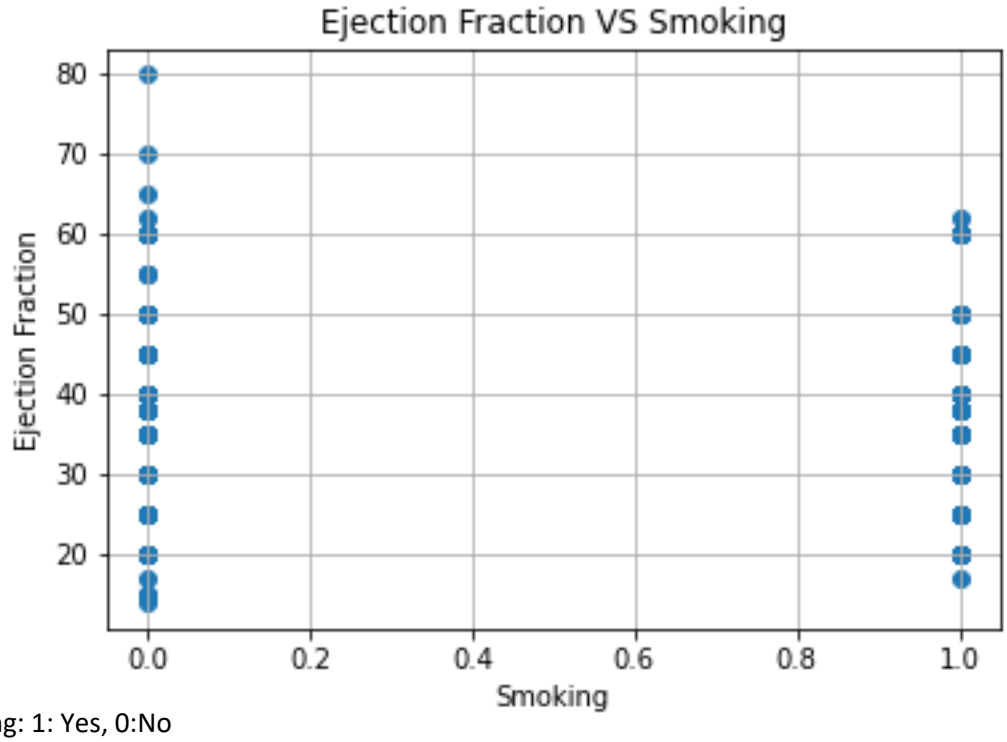
Sex: Female: 0, Male: 1

Women had ejection fraction above 65, but men had ejection fraction below 65.



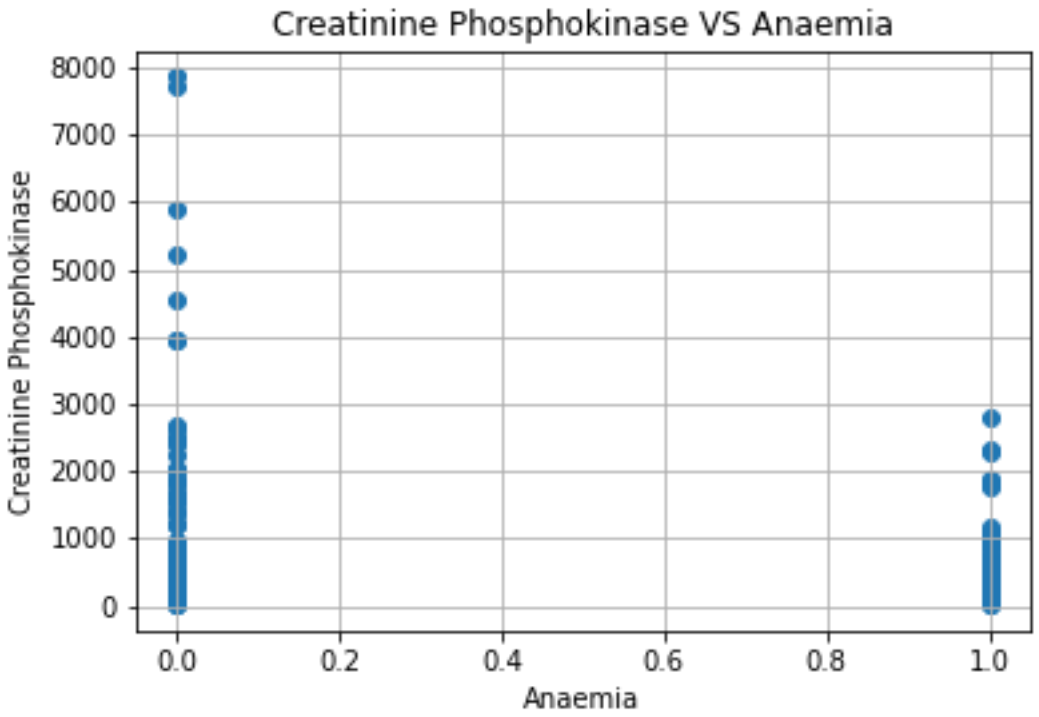
Smoking: 1: Yes, 0:No

Serum creatinine in people who smoke, is above 4.



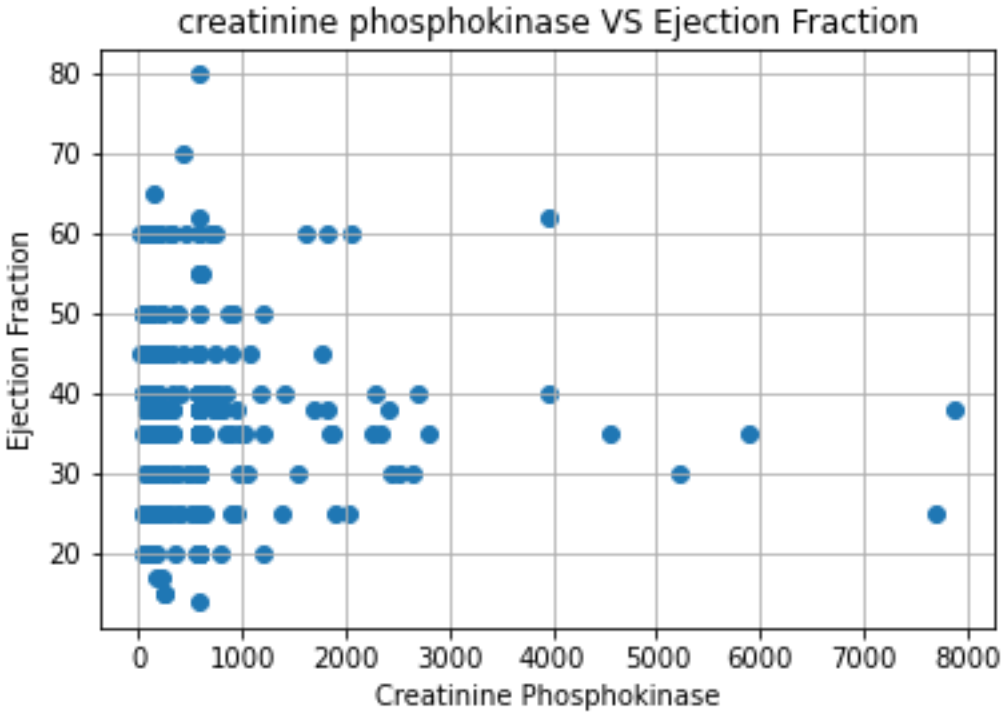
Smoking: 1: Yes, 0:No

Ejection fraction in some people who smoke, is above 65.

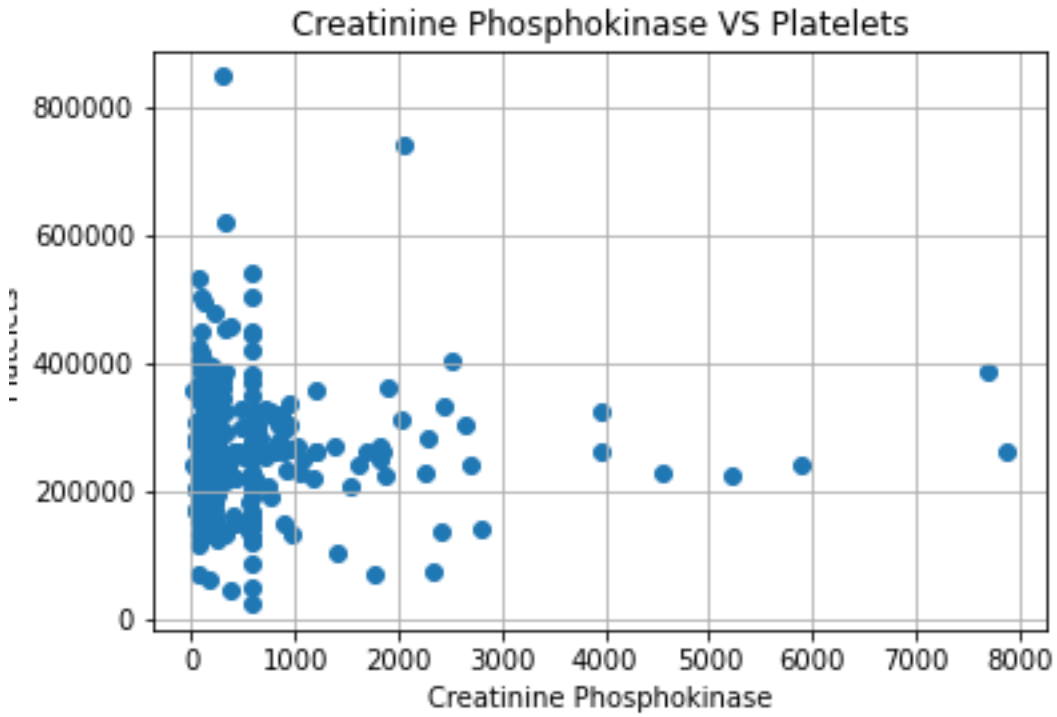


Anemia: 1: Yes, 0:No

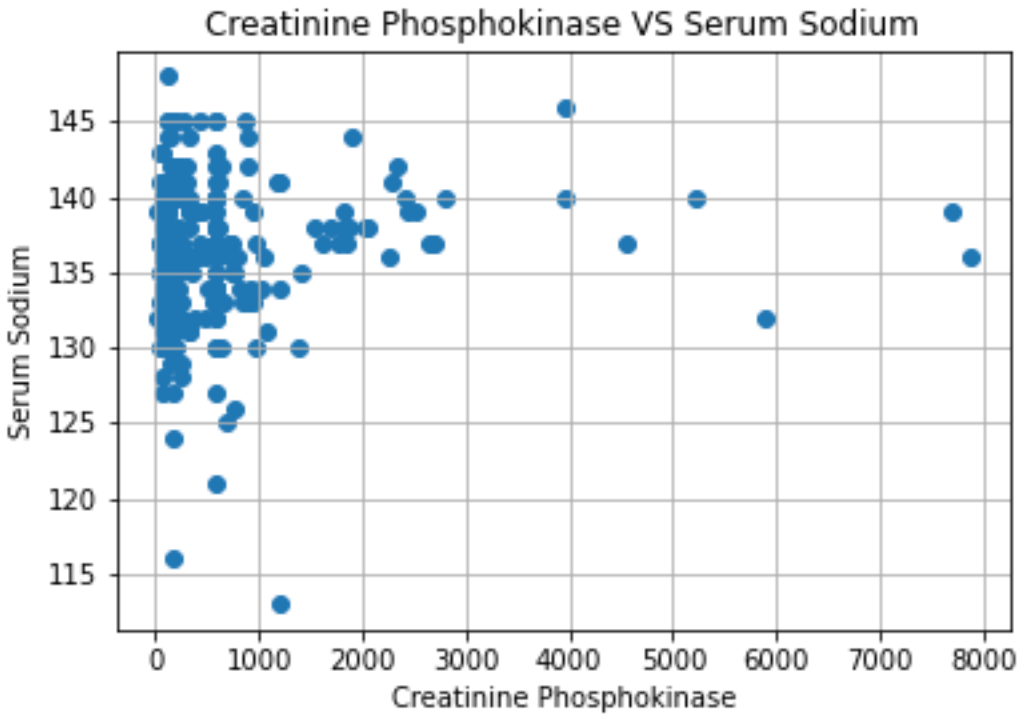
Creatinine phsphokinase in people who had anaemia is below 3000, but in people who didn't have anaemia, it is goes to 8000.



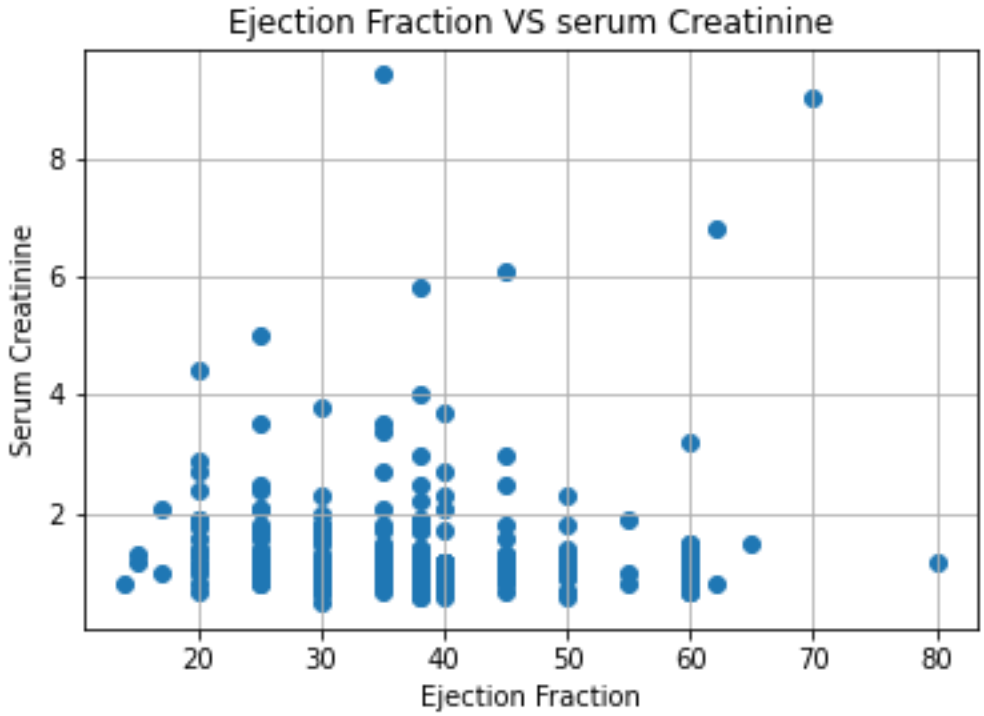
Creatinine phosphokinase increases with a sharp slope from 0 to 8000, between 0 to 40 of ejection fraction. between 40 to 80 it decreses with a sharp slope to almost 1000.



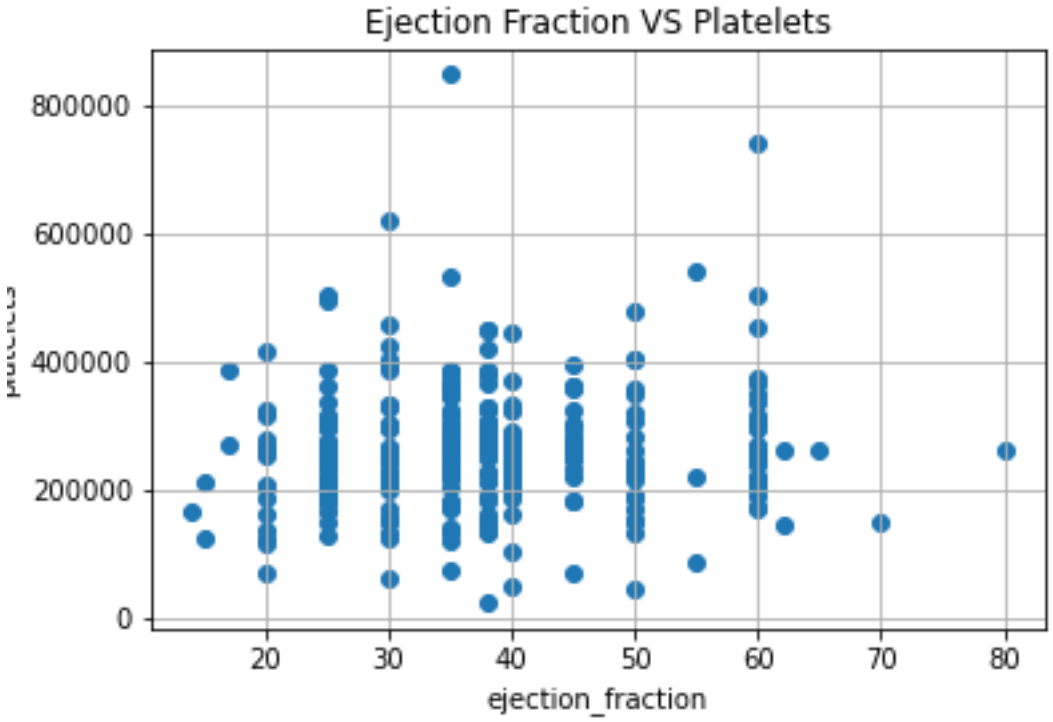
In range 0-400000 of platelets, the creatinine phosphokinase increases to 8000 with a sharp slope then it decreases to range 0-1000 in range 400000-800000 of platelets.



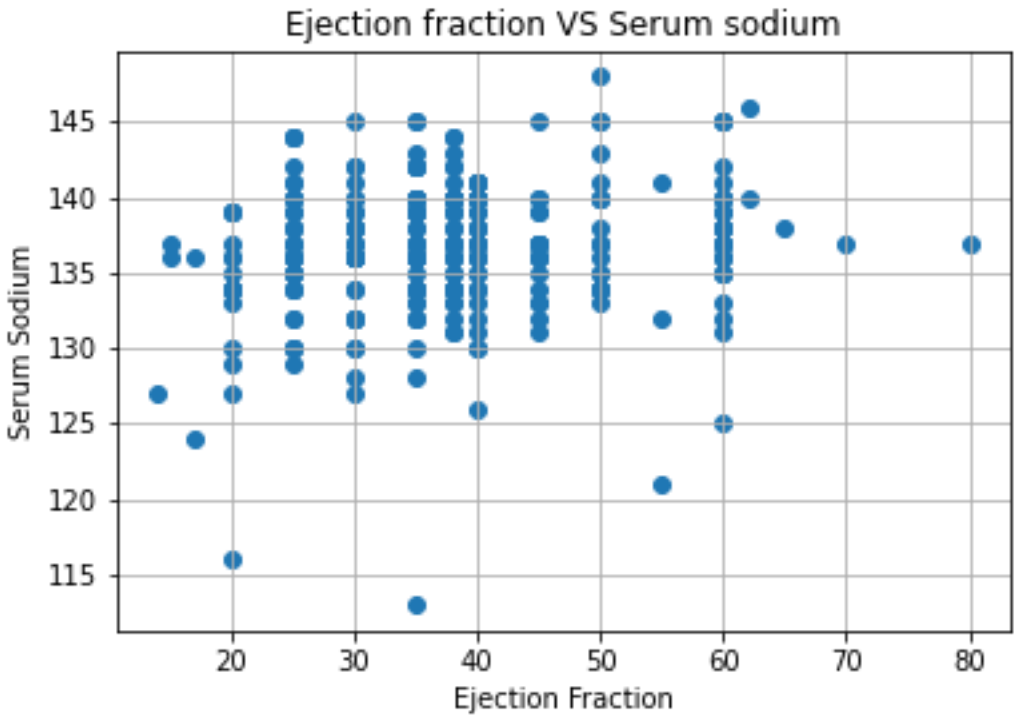
In range 0-130 of serum sodium, the amount of of creatinine phosphokinase is almost steady and between 0-1200. In range 130-140 of serum sodium, creatinine phosphokinase increases up to 8000. Above 140 of serum sodium, creatinine phosphokinase ecreases to 0 with a sharp slope.



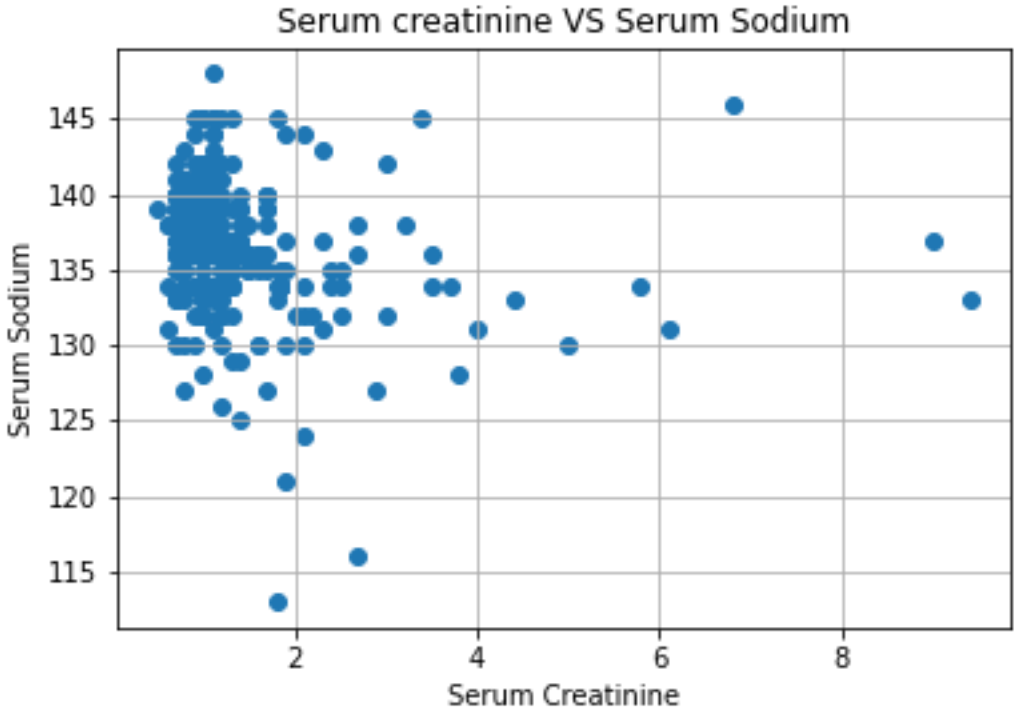
In range between 0-70 of ejection fraction, serum creatinine increases with a soft slope up to 9. From 70 to 80 of ejection fraction, serum creatinine decreases with a sharp slope to 1.



In range of 10-30 of ejection fraction, platelets increase up to 600000. in range 30-60 of ejection fraction platelets are almost steady in range 400000-600000. above amount 60 of ejection fraction platelets decreases to almost 200000.



In range 20-60 of ejection fraction, serum sodium is in range 125-145.



In range of 115-123 of serum sodium, serum creatinine increases with a sharp slope up to 6. From 135 to 145 of serum sodium, serum creatinine decreases with a sharp slope to almost 2.