

```
In [4]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

```
In [8]: data = pd.read_csv("C:\\Users\\shashi\\Desktop\\python project\\heart failure.csv")
data.head()
```

```
Out[8]:
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure	platelet
0	75.0	0	582	0	20	1	265000.0
1	55.0	0	7861	0	38	0	263358.0
2	65.0	0	146	0	20	0	162000.0
3	50.0	1	111	0	20	0	210000.0
4	65.0	1	160	1	20	0	327000.0

```
In [9]: #describing the data
data.describe()
```

```
Out[9]:
```

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pre
count	299.000000	299.000000	299.000000	299.000000	299.000000	299.000000
mean	60.833893	0.431438	581.839465	0.418060	38.083612	0.391271
std	11.894809	0.496107	970.287881	0.494067	11.834841	0.486590
min	40.000000	0.000000	23.000000	0.000000	14.000000	0.000000
25%	51.000000	0.000000	116.500000	0.000000	30.000000	0.000000
50%	60.000000	0.000000	250.000000	0.000000	38.000000	0.000000
75%	70.000000	1.000000	582.000000	1.000000	45.000000	1.000000
max	95.000000	1.000000	7861.000000	1.000000	80.000000	1.000000

```
In [10]: #Exploratory data analysis

len_live = len(data['DEATH_EVENT'][data['DEATH_EVENT'] == 0])
len_death = len(data['DEATH_EVENT'][data['DEATH_EVENT'] == 1])

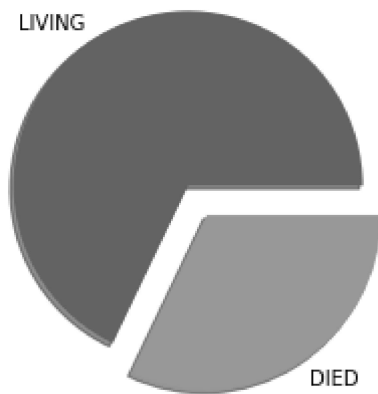
arr = np.array([len_live, len_death])
labels = ['LIVING', 'DIED']

print(f'Total number of Living case:- {len_live}')
print(f'Total number of Death case:- {len_death}')
```

```
plt.pie(arr, labels = labels, explode=[0.2,0.0], shadow = True)
plt.show()
```

Total number of Living case:- 203

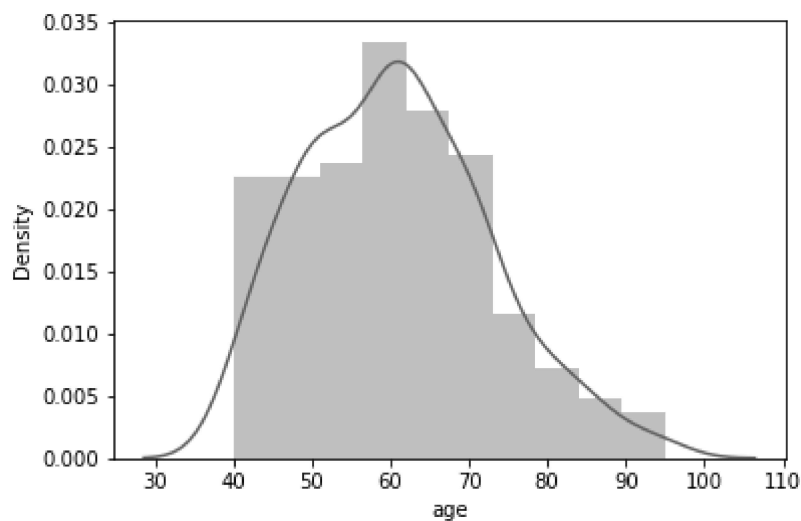
Total number of Death case:- 96



In [11]: *##Seeing the distribution of age*

```
sns.distplot(data['age'])
```

Out[11]: <AxesSubplot:xlabel='age', ylabel='Density'>



In [28]: *## Selecting columns that are above age 50 and seeing died or not*

```
age_above_50_not_died = data['DEATH_EVENT'][data.age >=50][data.DEATH_EVENT == 0]
age_above_50_died = data['DEATH_EVENT'][data.age >= 50][data.DEATH_EVENT == 1]

len_died = len(age_above_50_died)
len_not_died = len(age_above_50_not_died)

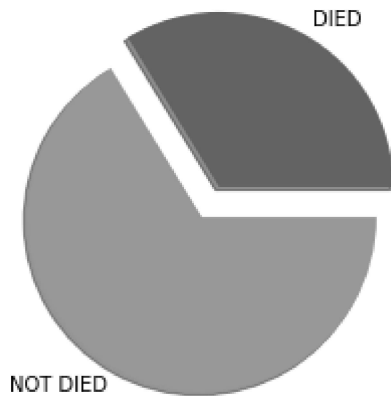
arr1 = np.array([len_died, len_not_died])
labels = ['DIED', 'NOT DIED']

print(f'Total number of Died:- {len_died}')
print(f'Total number of Not Died:- {len_not_died}')
```

```
plt.pie(arr1, labels=labels, explode = [0.2, 0.0], shadow= True)
plt.show()
```

Total number of Died:- 85

Total number of Not Died:- 167



```
In [29]: patient_nhave_diabetes_0 = data['DEATH_EVENT'][data.diabetes == 0][data.DEATH_EVENT == 0]
patient_have_diabetes_1 = data['DEATH_EVENT'][data.diabetes == 1][data.DEATH_EVENT == 1]

len_d_died = len(patient_have_diabetes_1)
len_d_alive = len(patient_nhave_diabetes_0)

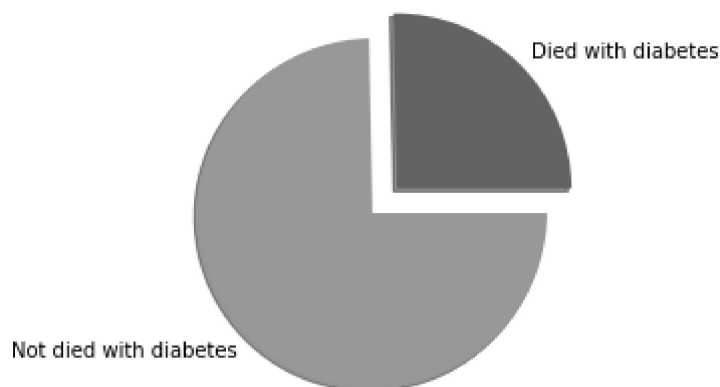
arr2 = np.array([len_d_died, len_d_alive])
labels = ['Died with diabetes', 'Not died with diabetes']

print(f'Total number of Died with diabetes:- {len_d_died}')
print(f'Total number of Not died with diabetes: {len_d_alive}')

plt.pie(arr2, labels=labels, explode = [0.2,0.0], shadow = True)
plt.show()
```

Total number of Died with diabetes:- 40

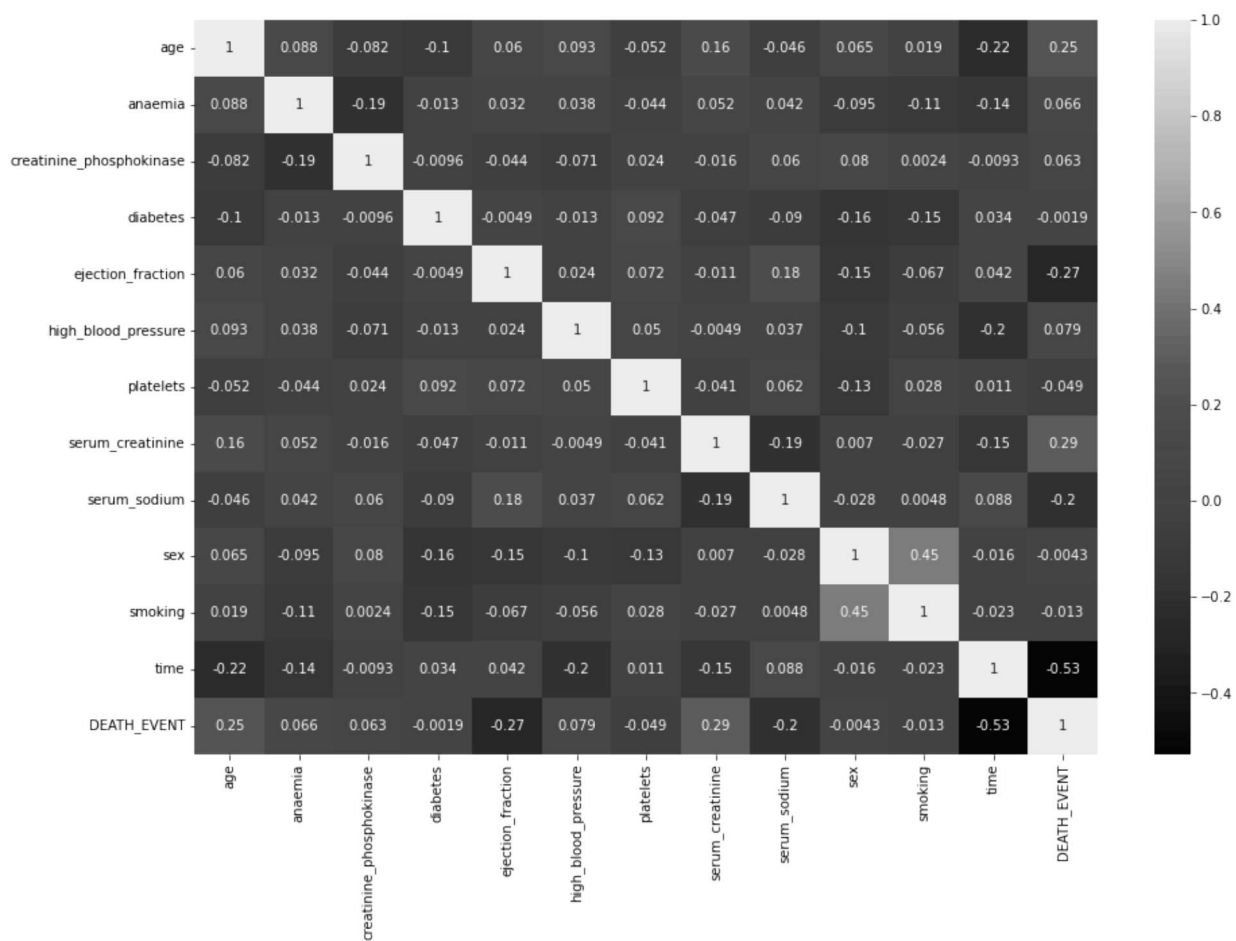
Total number of Not died with diabetes: 118



```
In [35]: ##Checking correlation of our variables
## -1 indicates a perfectly negative linear correlation between two variables
## 0 indicates no linear correlation between two variables
## 1 indicates a perfectly positive linear correlation between two variables

corr = data.corr()
plt.subplots(figsize=(15,10))
sns.heatmap(corr, annot=True)
```

Out[35]: <AxesSubplot:>



In [4]: data.corr().style.background_gradient(cmap='coolwarm')

Out[4]:

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction
age	1.000000	0.088006	-0.081584	-0.101012	0.060098
anaemia	0.088006	1.000000	-0.190741	-0.012729	0.031557
creatinine_phosphokinase	-0.081584	-0.190741	1.000000	-0.009639	-0.044080
diabetes	-0.101012	-0.012729	-0.009639	1.000000	-0.004850
ejection_fraction	0.060098	0.031557	-0.044080	-0.004850	1.000000
high_blood_pressure	0.093289	0.038182	-0.070590	-0.012732	0.024445
platelets	-0.052354	-0.043786	0.024463	0.092193	0.072177
serum_creatinine	0.159187	0.052174	-0.016408	-0.046975	-0.011302
serum_sodium	-0.045966	0.041882	0.059550	-0.089551	0.175902
sex	0.065430	-0.094769	0.079791	-0.157730	-0.148386
smoking	0.018668	-0.107290	0.002421	-0.147173	-0.067315
time	-0.224068	-0.141414	-0.009346	0.033726	0.041729
DEATH_EVENT	0.253729	0.066270	0.062728	-0.001943	-0.268603



In []:

In []: