Heart disease prediction

About the dataset :

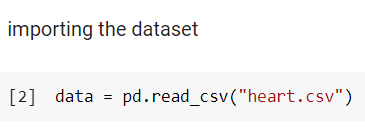
Code and their explanation

1. For the first step, we have to import the libraries which are NumPy for array, Pandas for the dataset, and matplotlib and seaborn are used for the data visualization.

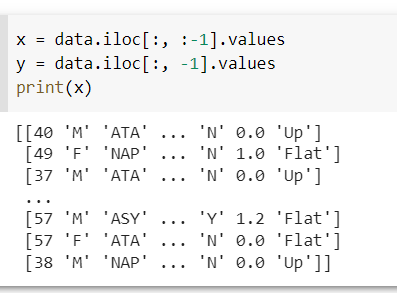
Graphical user interface, text, application, chat or text message

Description automatically generated

1. Now for the second step, we have to import the dataset using the Pandas library



1. Then we are going to split the data as in x and y. x as an independent variable and y as the dependent variable.



Background pattern

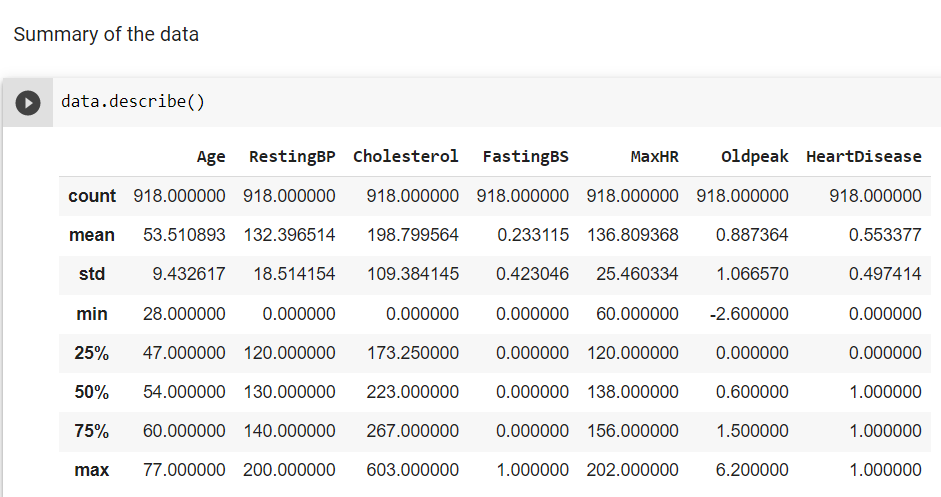
Description automatically generated

1. Now we have to check if there are any missing values are present or not. For this we have taken the info of the data then we see that the total number of range index and the number of rows in a particular column are the same then there are no missing values otherwise we will have to remove them.



We can see that number in the range index and the number of rows in particular columns is the same which means there are no missing values.

1. Now we have seen the summary of the data for all the numerical variables.



DATA VISUALIZATION

1. We are going to visualize the data based on their age.

Chart, histogram

Description automatically generated

In the above figure, we can clearly that age is the factor that is affected by heart diseases. 40-60 years of age were highly affected by heart diseases. After the age of 60 years, the level is decreasing. 20-40 years of people are less affected by heart diseases but as the age is increasing the risk of having a heart disease is also increasing. People whose age is more than 60 years of age have fewer heart diseases than others. People who have their aged between 0-20 have almost negligible heart diseases.

1. Now we are going to visualize based on sex how males and females affect heart diseases.

Chart, bar chart

Description automatically generated

In the above figure, we have seen that in males, 267 people do not have heart disease and 143 women are those who do not have any heart disease. There are 458 males are those which have heart disease, and 50 women are those which have heart disease.

1. Now we are going to visualize the slope for how much the peak exercise.

Chart, bar chart

Description automatically generated

It tells us that the people who have up the slope are very fewer heart diseases than the people who do not have any heart diseases. While a person who has a flat slope has a very higher risk for heart disease. When there is down the slope they have a slightly higher risk of having heart diseases.