

What is the framing question of your analysis, or the purpose of the model/system you plan to build?

Knowing the causes of cardiovascular diseases using weight, smoking, age and cholesterol factors

Who benefits from exploring this question or building this model/system?

The health field, individuals and community awareness

What dataset(s) do you plan to use, and how will you obtain the data?

Cardio vascular Disease From Kaggle

<https://www.kaggle.com/sulianova/cardiovascular-disease-dataset>

What is an individual sample/unit of analysis in this project? What characteristics/features do you expect to work with?

Age

gender

cholesterol

smoke

weight

Activ

Ap-hi

If modeling, what will you predict as your target?

Percentage of smokers infected with the disease?

How many Smokers and Non-Smokers are there then?

Do men suffer from the disease more than women?

Is there a Relationship between smoke and Disease?

Is there a connection between Age and Disease?

Is there a relationship between physical activity and disease?