1. **What is the business problem?**
   Without knowing that a person has Heart Disease.

2. **Who are the intended stakeholders, and why is this problem relevant to them?**
   Doctors - it will help the doctor's decision classifying heart disease thru data values.|

   Patients - thru Patients symptoms we can tell if there have heart problem and go to the doctor for second opinion.

3. **Where are the datasets available from?**
   There are four available at
   https://archive.ics.uci.edu/ml/datasets/heart+disease

4. **Which one do you like the most?**
   Patients Classification of Heart Disease

5. **What type of data science approach would you use?**
   Supervised Binary Classification

6. **How many rows and how many columns does t h e dataset have?**
   It consist of 303 individuals row data and there are effectively 14 columns so, t here are effectively 14 columns.

   1. Age(age) : the age of the individual.

   2. Sex(sex) : the gender of the individual using the following format :
      1 = male 0 = female.

   3. Chest-Pain type (cp): the type of chest-pain experienced by the individual using the following format :
      1 = typical angina
      2 = atypical angina
      3 = non - anginal pain
      4 = asymptotic

   4. Resting Blood Pressure(trestbpd) : displays the resting blood pressure value of an individual in mmHg (unit)

   5. Serum Cholesterol(chol): displays the serum cholesterol in mg/dl (Unit)

   6. Fasting Blood Sugar(fbs): compares the fasting blood sugar value of an individual with 120mg/dl.
      If fasting blood sugar > 120mg/dl then : 1 (True) else : 0 (False)

7. Resting ECG(restecg) : 0 = normal 1 = having ST-T wave abnormality 2 = left ventricular hyperthrophy

8. Max heart rate achieved(thalach): displays the max heart rate achieved by an individual.

9. Exercise induced angina(exang):
   1 = yes
   0 = no
10. ST depression induced by exercise relative to rest(oldpeak): displays the value which is integer or float.

11. Peak exercise ST segment(slope) :
    1 = upsloping
    2 = flat
    3 = downsloping

12. Number of major vessels (0-3) colored by flourosopy(ca): displays the value as integer or float.

13. Thal : displays the thalassemia(thal) : 3 = normal 6 = fixed defect 7 = reversable defect

14. Diagnosis of heart disease(num) : Displays whether the individual is suffering from heart disease or not : 0 = absence 1,2,3,4 = present.

7. Is this project similar to the "Big Mountain Resort"? Yes, no, why?

No because BMR is asking ‘How Much’ and it’s a supervised regression approach while this project is asking ‘Do patients have heart disease’ answerable by Yes or No, and this project will use Supervised Binary Classification.