**Heart Failure Prediction**

Project number : B2

1. Alisha Chandel (20BT30002)
2. Ankita (20BT30003)
3. Sachin Itkan (20BT30025)

**Problem Definition:**

Based on the given dataset of patients of cardiovascular diseases or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidaemia or already established disease), predict the likeliness of heart disease using machine learning algorithms.

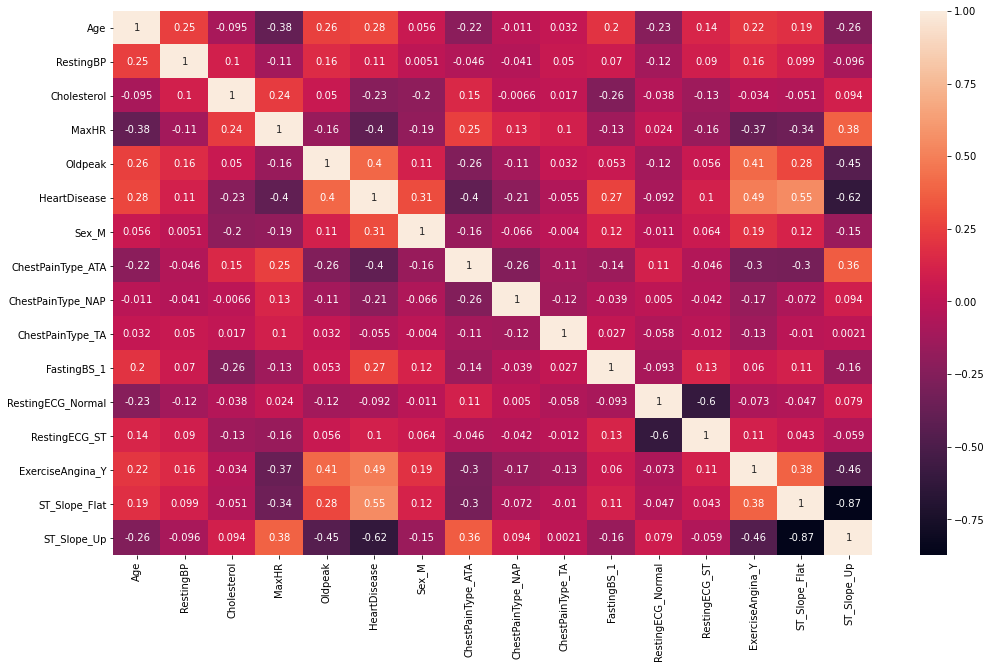
**Analysis of Problem:**

Heart Disease:1 Normal:0

## Heart Disease vs Normal:



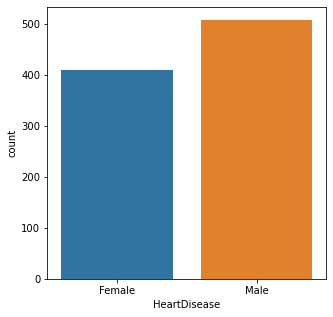
## Correlation Matrix



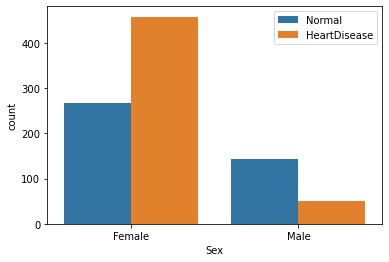
Highly Positively correlated features are: Age, RestingBP , Cholestrol , MaxHR, ChestTypePain, RestingECG, ExcerciseAngina

Highly Negatively Correlated Features are: ST\_Slope, OldPeak

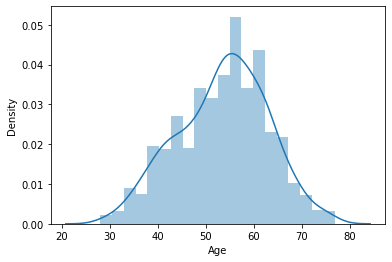
## Heart Disease as per Sex:



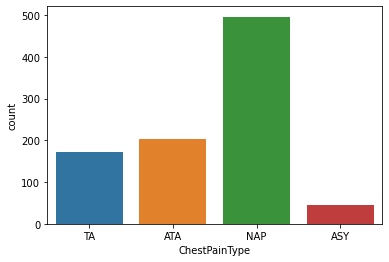
## Gender Distribution:



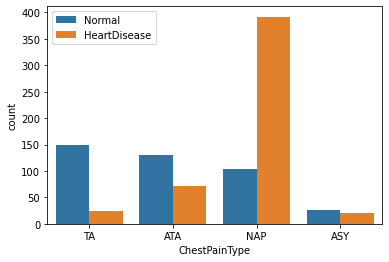
## Age Distribution



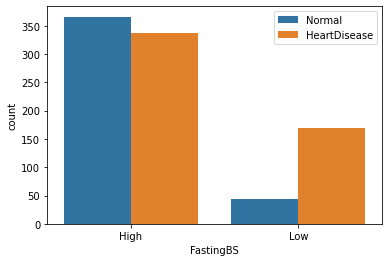
## Chest Pain Distribution:



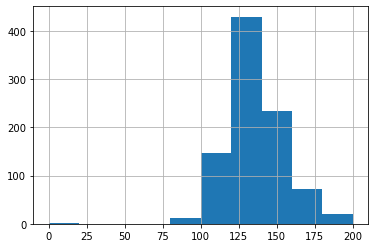
## ChestPain Distribution as per Heart Disease:



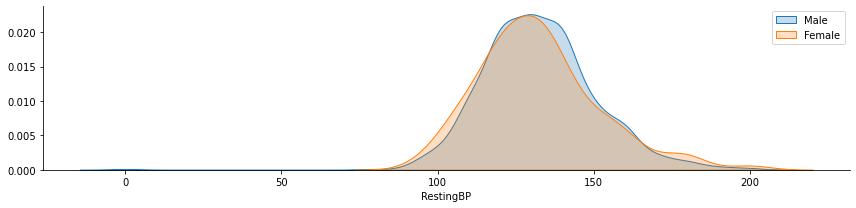
## Fasting BloodSugar vs Heart Disease:



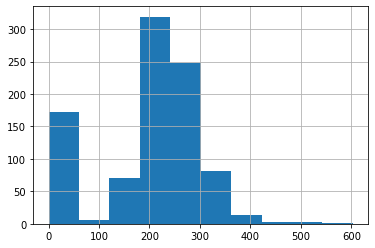
## Blood sugar as per HeartDisease:



## Blood pressure as per sex:



## Serum cholestrol distribution:



After applying different ML Algorithms for classification:

Linear ML Algorithms:

1. Logistic Regression
2. Support Vector Machine(SVM)
3. K-Nearest Neighbors

Non-Linear Algorithms:

1. Decision Tree Classifier
2. Random Forest Classifier
3. Gradient Boost Classifier
4. XGBoost Classifier

Result:

Checking Accuracy of different algorithms, we found that Random Forest Classifier classifies best.

1:Heart Disease

0:Normal

Accuracy=0.907608695652174