Machine Learning Project- 01

HEART DISEASE PREDICTIONPROJECT

[Group Project]

Details of Group Members:

| Name: | GANESH WAGHMARE | PRANJAL YADAV |
| --- | --- | --- |
| College: | TERNA ENGINEERING COLLEGE | TERNA ENGINEERING COLLEGE |
| Email: | ganeshwaghmare@ternaengg.ac.in | pranjalyadav@ternaengg.ac.in |
| Phone No.: | 8459164229 | 9833553644 |

**Index**

| **Sr.No** | **Contents** | **Page** |
| --- | --- | --- |
| 1 | Overview | 2 |
| 2 | Proposed Methodology | 2 |
| 3 | Visualization | 4 |
| 4 | Implementation | 10 |
| 5 | Conclusion | 13 |

**Overview**

Machine Learning is used across many spheres around the world. The healthcare industry is no exception. Machine Learning can play an essential role in predicting presence/absence of Locomotor disorders, Heart diseases and more. Such information, if predicted well in advance, can provide important insights to doctors who can then adapt their diagnosis and treatment per patient basis.

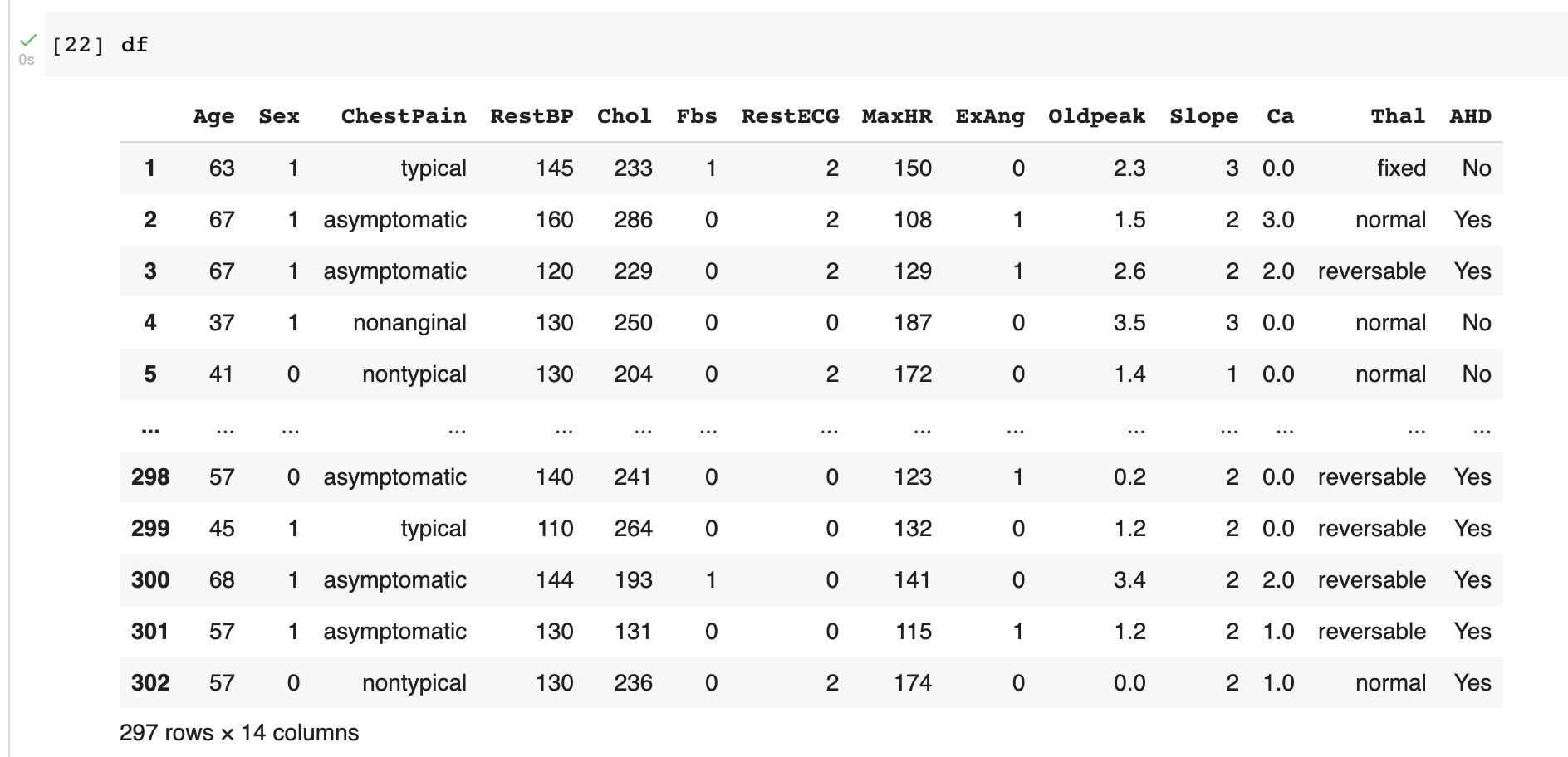
An estimate by the World Health Organization, that over 17.9 million deaths occur every year worldwide because of cardiovascular disease, and of these deaths, 80% are because of coronary artery disease and cerebral stroke. The vast number of deaths is common amongst low and middle-income countries. Many predisposing factors such as personal and professional habits and genetic predisposition accounts for heart disease. Various habitual risk factors such as smoking, overuse of alcohol and caffeine, stress, and physical inactivity along with other physiological factors like obesity, hypertension, high blood cholesterol, and pre-existing heart conditions are predisposing factors for heart disease. The efficient and accurate and early medical diagnosis of heart disease plays a crucial role in taking preventive measures to prevent death.

Using machine learning, it detects hidden patterns in the input dataset to build models. It makes accurate predictions for new datasets. The dataset is cleaned and missing values are filled. The model uses the new input data to predict heart disease and is then tested for accuracy. In its most basic sense, machine learning uses programmed algorithms that learn and optimize their operations by analyzing input data to make predictions within an acceptable range. With the feeding of new data, these algorithms tend to make more accurate predictions.

**Proposed Methodology**

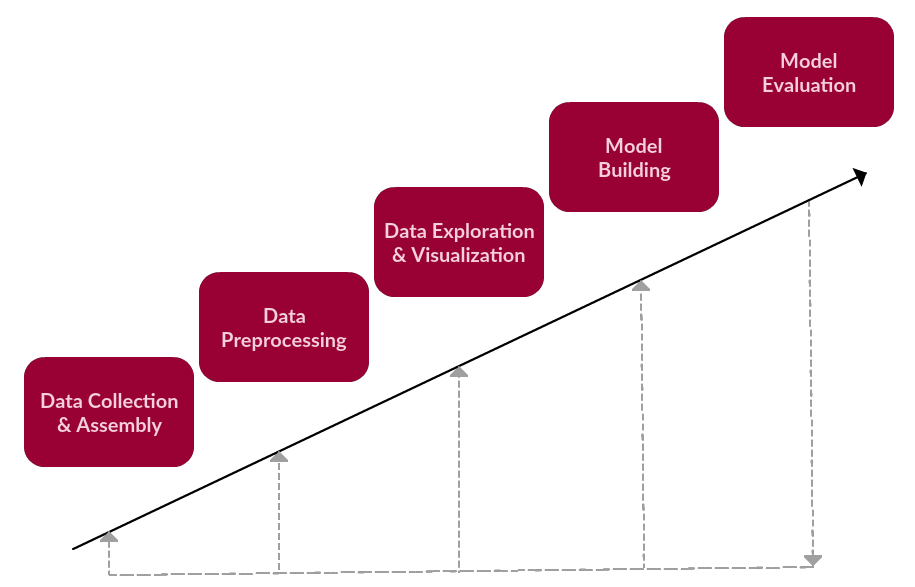
1. **Dataset:**

We have used a dataset from the Machine learning repository. It comprises a real dataset of 300 examples of data with 14 various attributes (13 predictors; 1 class) like blood pressure, type of chest pain, electrocardiogram result, etc.



1. **Data Pre-Processing**:

A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.



1. **Algorithm used:**

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

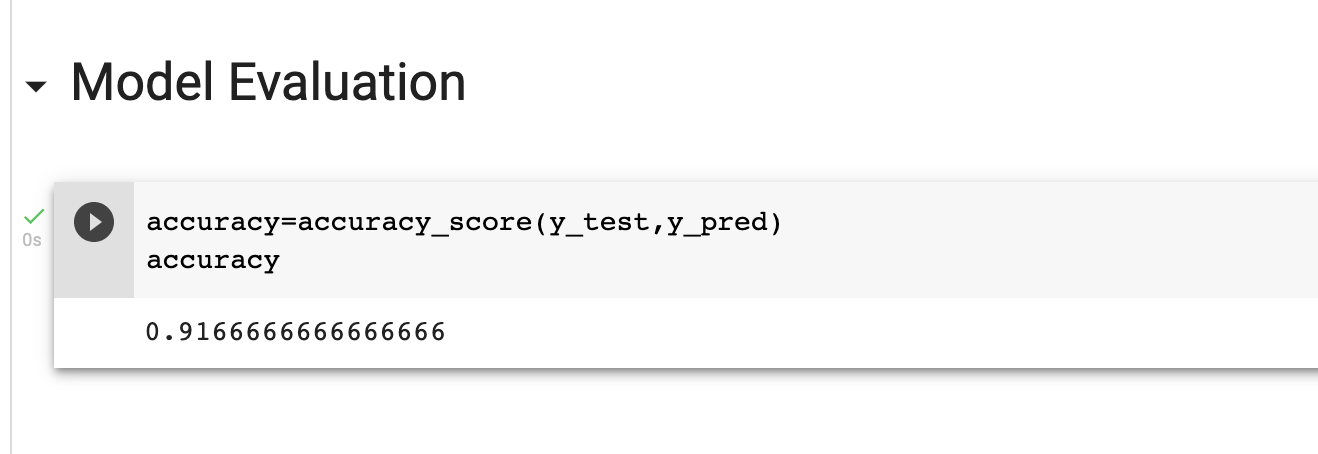
In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).

1. **Model Building:**

The model building by dividing the dataset into training set and test set.We have logistic Regression model to achieve high accuracy.

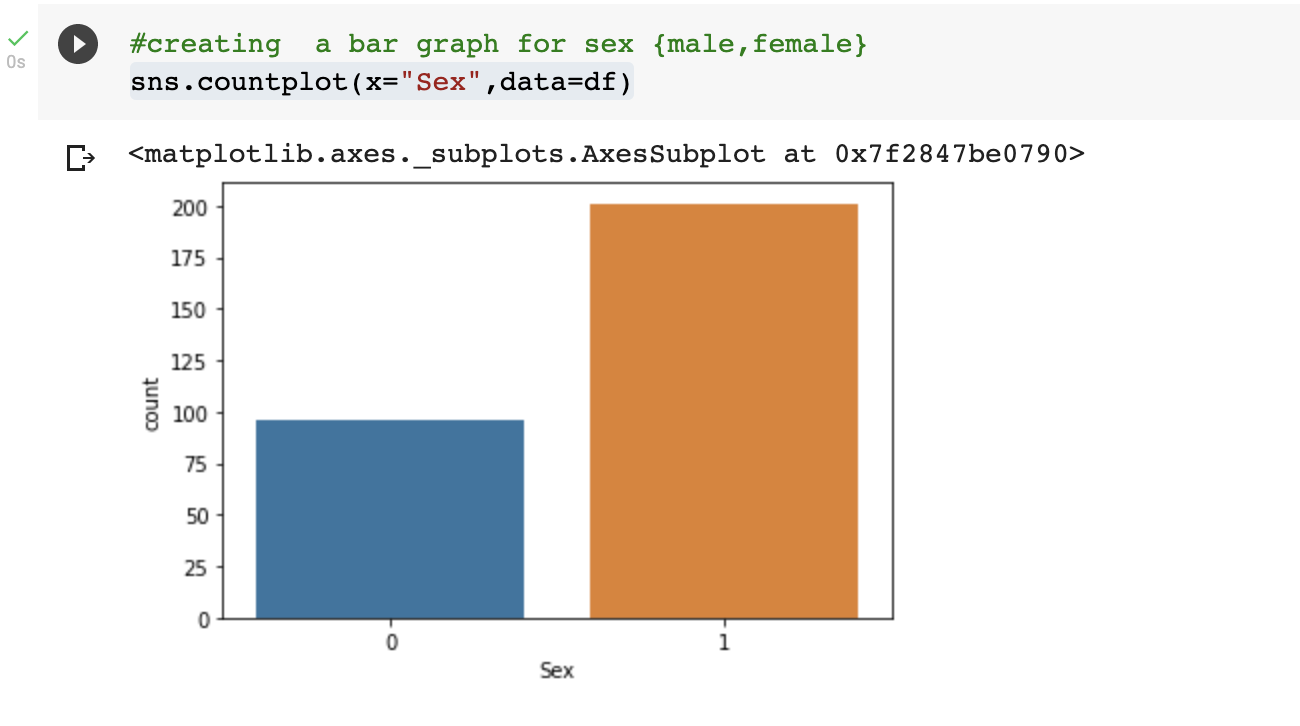
1. **Model Evaluation**:

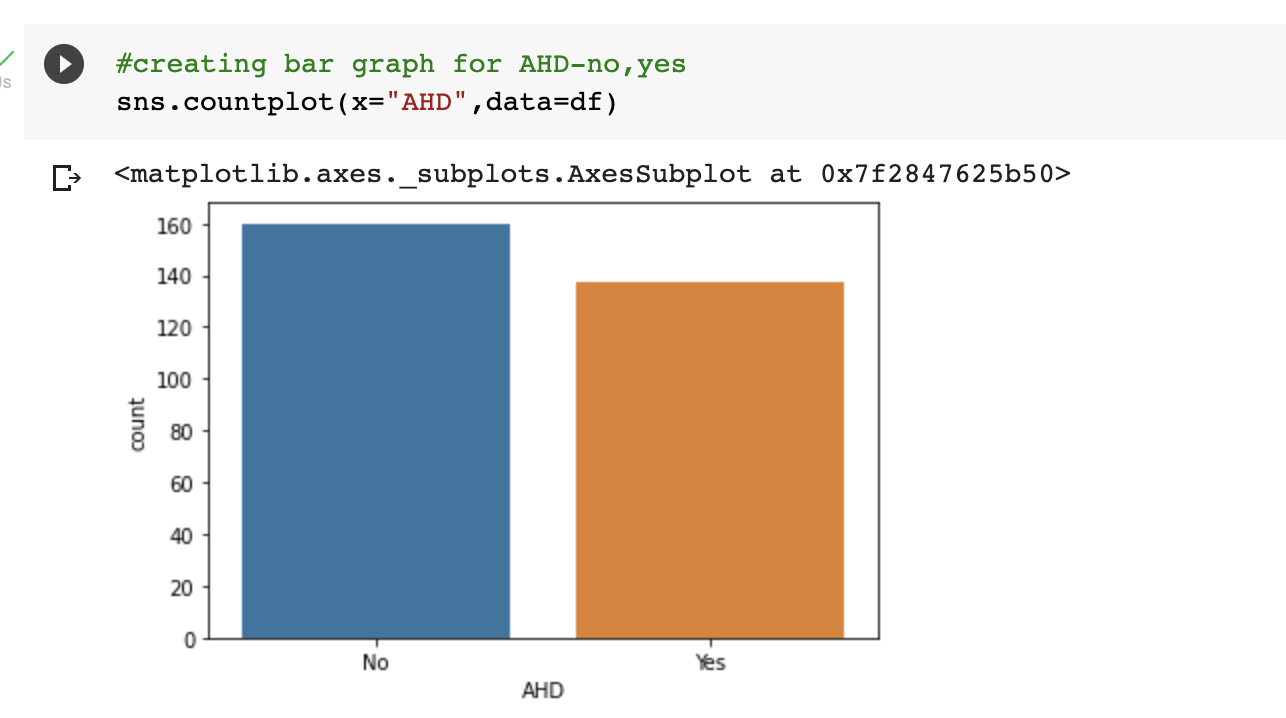
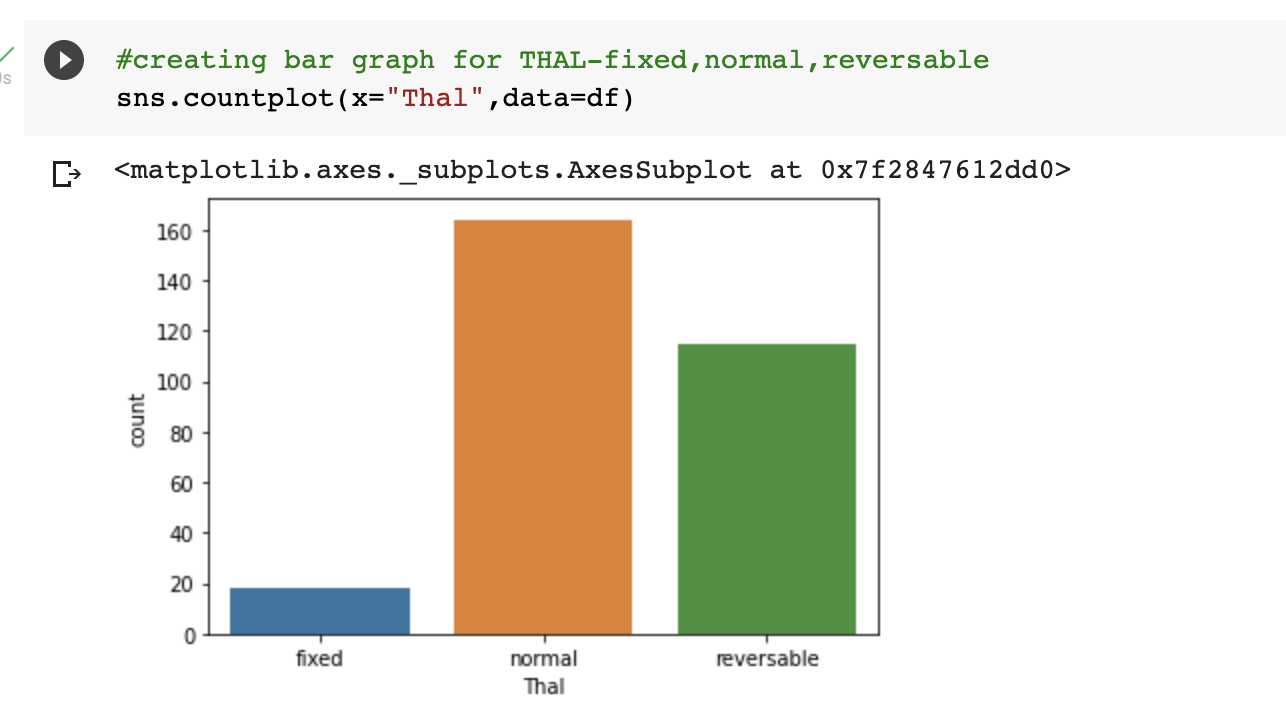
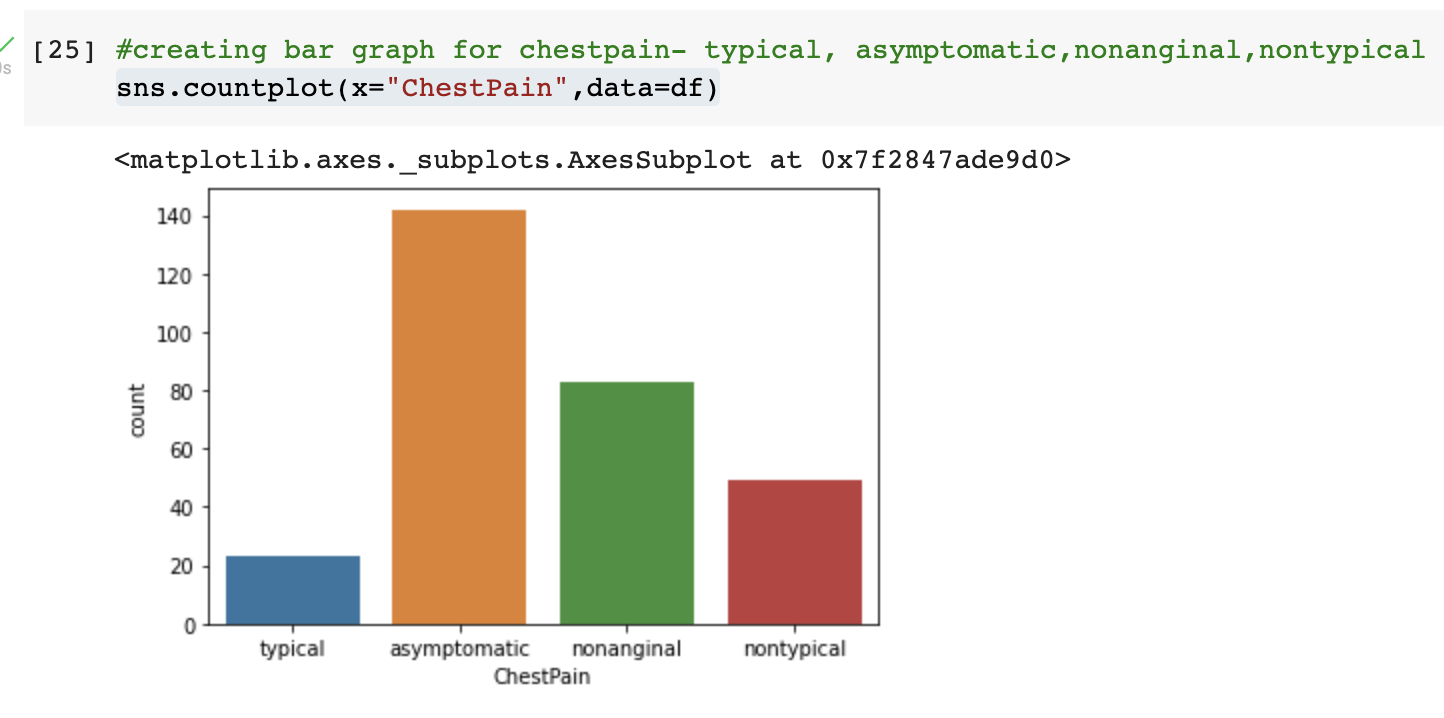
Our model gives accuracy of 91%.

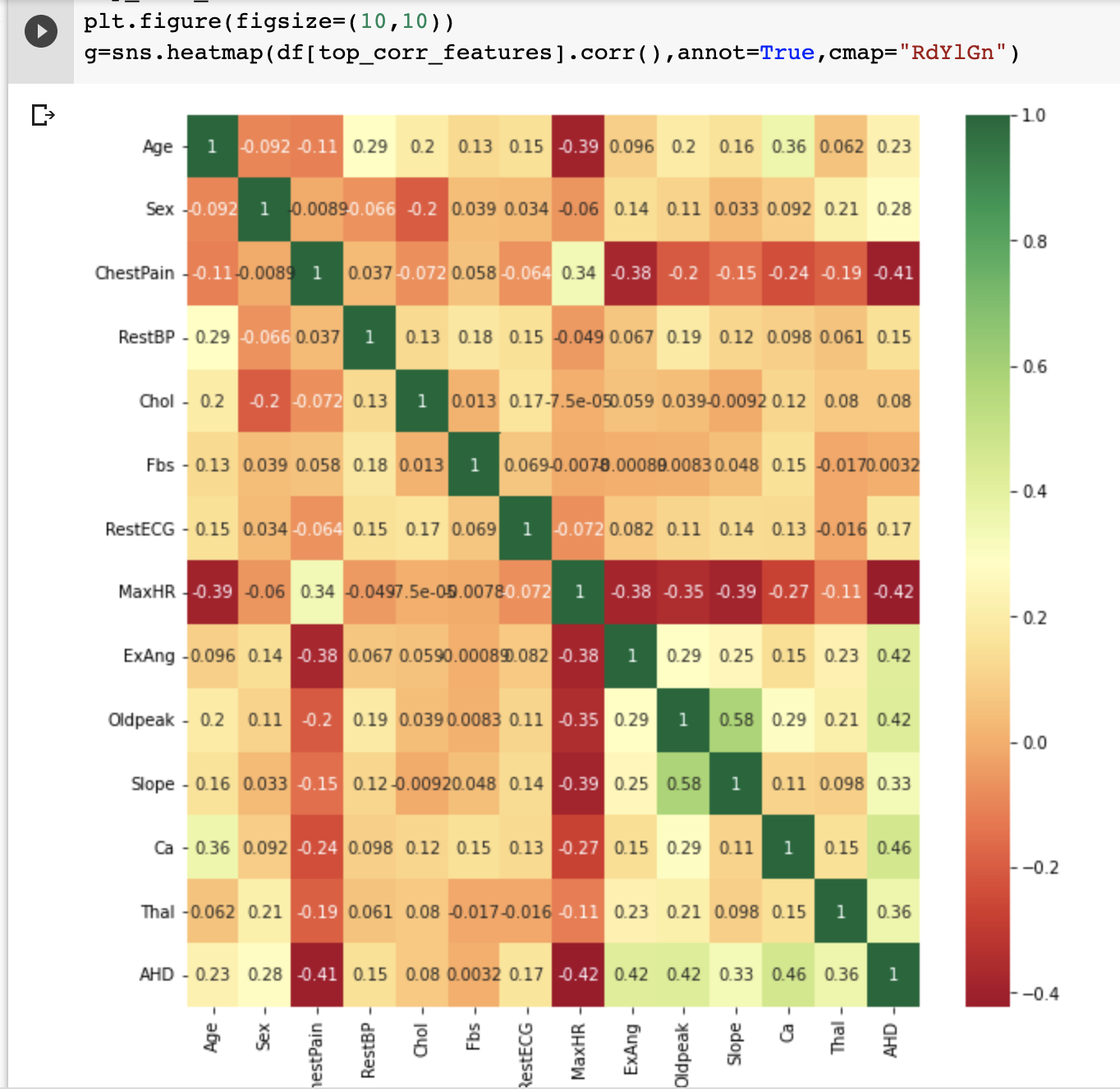


**Visualization**

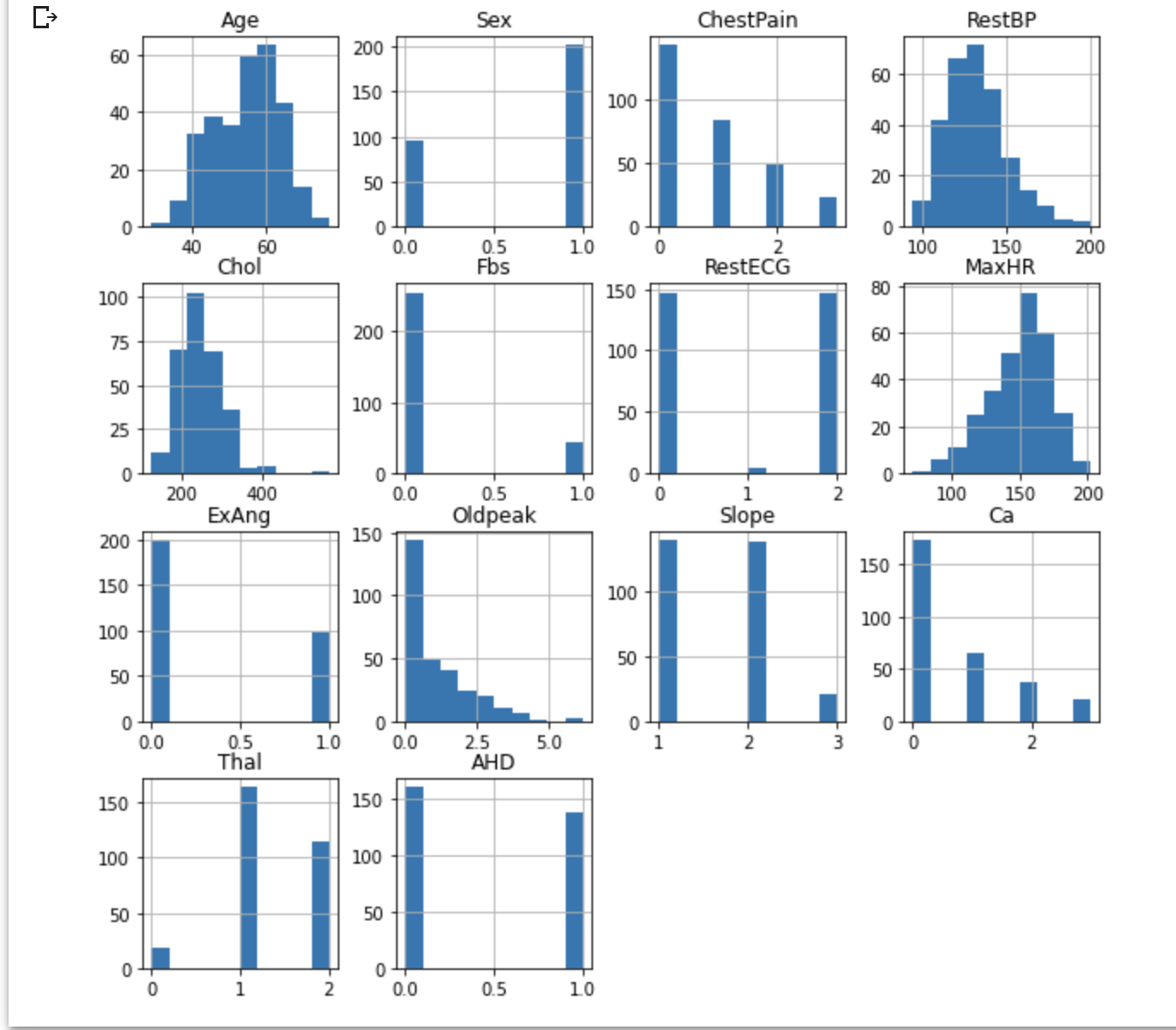
1. Comparing the risk of heart disease on different parameters.





Heatmap-To find all features all positively correlated or negatively correlated.

Histogram-



**Implementation**

| 1. |  |
| --- | --- |
| 2. |  |
| 3. |  |
| 4. |  |
| 5. |  |
| 6. |  |
| 7. |  |
|  |  |
|  |  |

**Conclusion**

* Men seem to be more susceptible to heart disease than women. Increase in Age, number of cigarettes smoked per day and systolic Blood Pressure also show increasing odds of having heart disease.
* The model predicted with 91% accuracy. The model is more specific than sensitive.