



Cardiovascular Disease



What is Cardiovascular Disease?

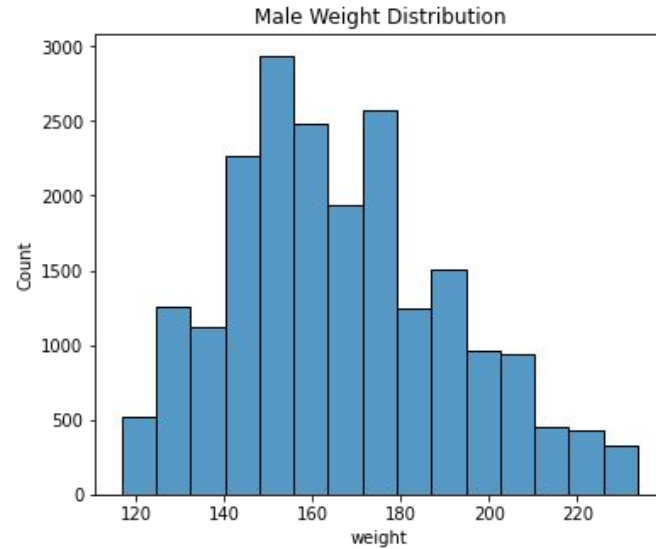
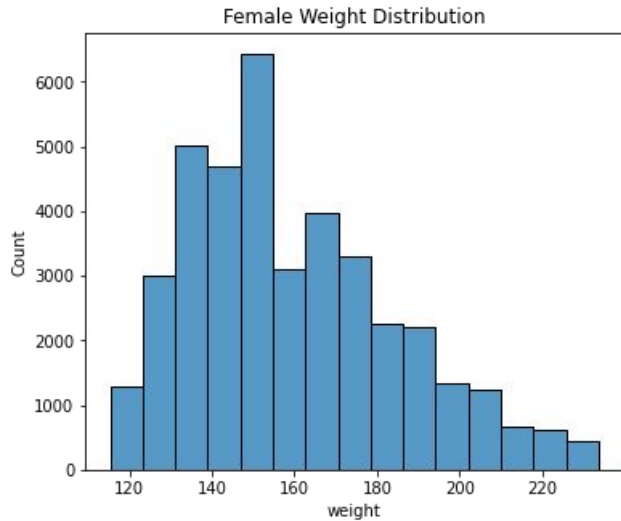
- The leading cause of death for men, women, and people of most racial and ethnic groups in the United States.
- Many factors contribute: weight, activity level, diet, genetics
- It is preventable with a healthy lifestyle



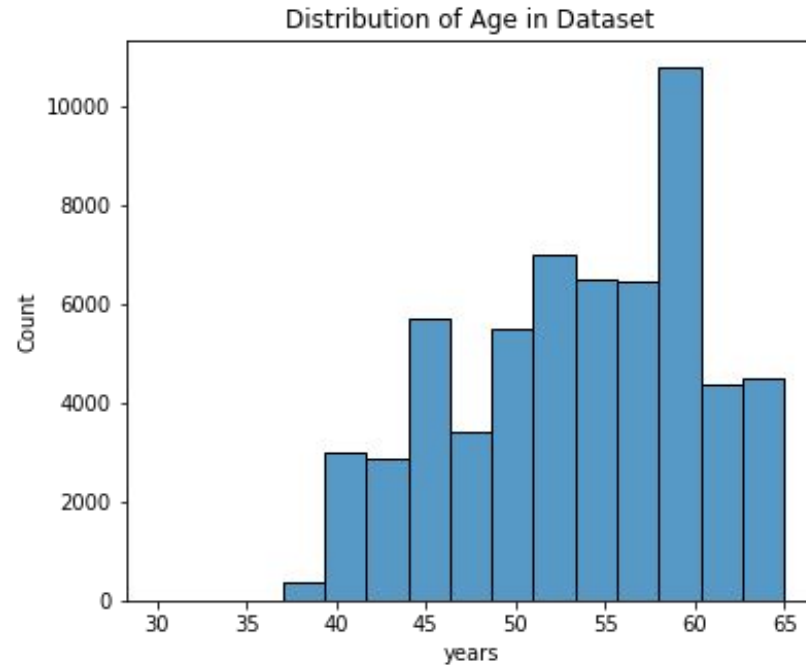
Data Set

- Source: Kaggle
- Size: 70k instances
- 13 features
- Continuous: Age, gender, blood pressure,BMI

Distribution of Weight in Data Set

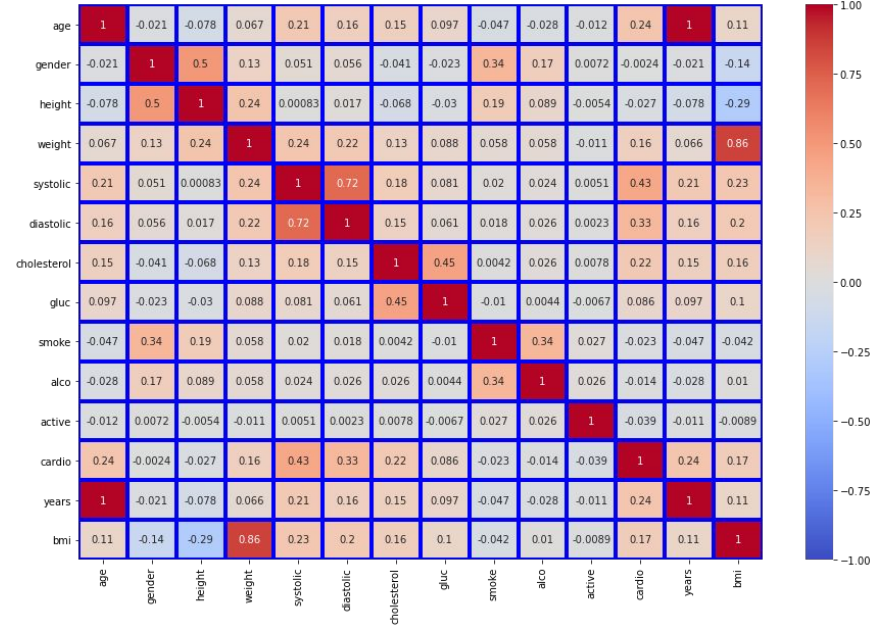


Distribution of Age in Data Set



Correlations:

- Systolic Pressure
- Age - Cholesterol Level
- Weight
- Diastolic Pressure
- Sedentary Lifestyle (Not Active)





Recommendations

- Eat healthy- follow the plate method to increase fruits and vegetables
- Physically active- at least 150 minutes per week recommended by the CDC
- Manage stress- helps in blood pressure
- Don't smoke
- Be more mindful to stay healthy as you get older

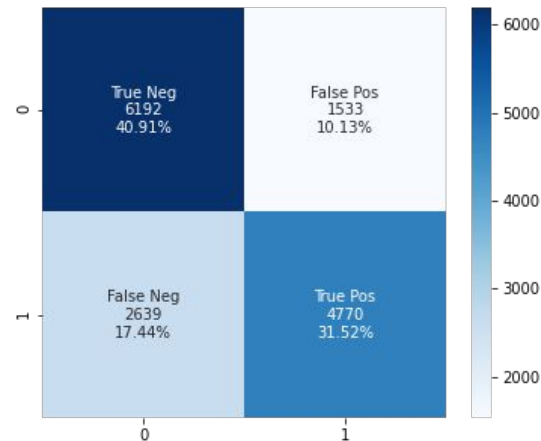
Models

The best performing models with AUC-ROC scores at .72 - .73 were:

- Logistic Regression
- Decision Trees
- SVM
- Random Forest (see confusion matrix)

The worst performing models were:

- Naive Bayes (AUC-ROC at .71)
- Ada Boost (AUC-ROC at .63)





Future Work

BMI- let's change this measure

More details on certain variables, i.e. effect of smoking on HDL and LDL numbers

Curious how COVID19 pandemic has impacted on cardiovascular disease risk



Thank You