HEART DISEASE DIAGNOSTIC ANALYSIS

- By Dhanshri Manusmare

PROJECT DETAILS

|	
Project Title	Heart Disease Diagnostic Analysis
Technologies	Business Intelligence
Domain	Healthcare
Project Difficulty Level	Advanced
Programming Language Used	Python
Tools Used	Jupyter Notebook, MS Office, MS PowerBI

OBJECTIVE

- The aim of project is Analysis Heart Disease changes based on given features. To achieve this we will use dataset which contains 14 features or attributes from 303 patients.

PROBLEM STATEMENT

- Health is real wealth in the pandemic time we all realized the brute effects of covid-19 on all irrespective of any status. You are required to analyze this health and medical data for better future preparation.

- A dataset is formed by taking into consideration some of the information of 303 individuals

ARCHITECTURE
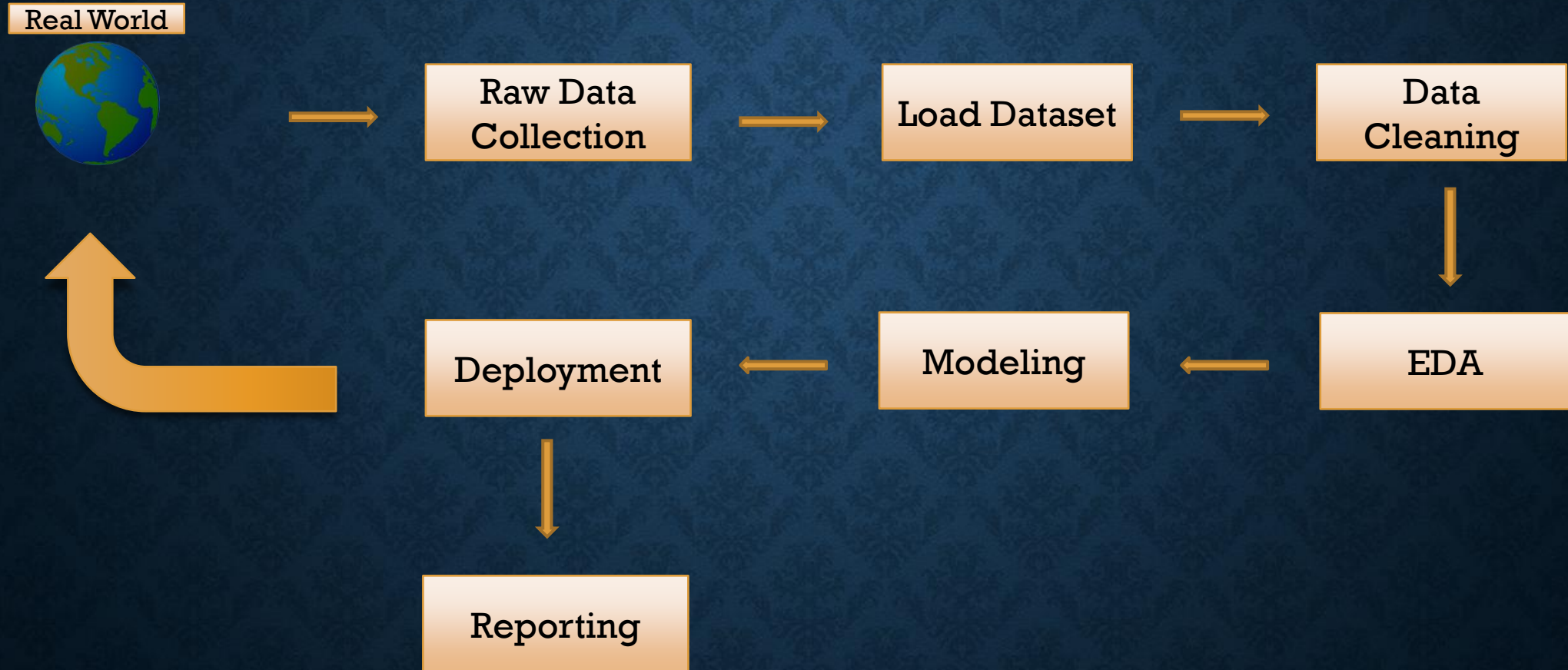
DATASET INFORMATION

- Age : The person's age in years

- sex: The person's sex (1 = male, 0 = female)

- Cp : The chest pain experienced (Value 1: typical angina, Value 2: atypical angina, Value 3: non-anginal pain, Value 4: asymptomatic)

- Trestbps : The person's resting blood pressure (mm/Hg on admission to the hospital)

- Chol :   The person's cholesterol measurement in mg/dl

- Fbs : The person's fasting blood sugar (> 120 mg/dl, 1 = true; 0 = false)

- Restecg : Resting electrocardiographic measurement (0 = normal, 1 = having ST-T wave abnormality, 2 = showing probable or definite left ventricular hypertrophy by Estes' criteria)

- Thalach : The person's maximum heart rate achieved

- Exang : Exercise induced angina (1 = yes; 0 = no)

- Oldpeak : ST depression induced by exercise relative to rest

- Slope : the slope of the peak exercise ST segment (Value 2: upsloping, Value 1: flat, Value 0: downsloping)

- Ca : The number of major vessels (0-3)

- Thal : A blood disorder called thalassemia (3 = normal; 6 = fixed defect; 7 = reversable defectt)

- Num : Heart disease (0 = no, 1 = yes)

Age : Age is the most important risk factor in developing cardiovascular or heart diseases, with approximately a tripling of risk with each decade of life. Coronary fatty streaks can begin to form in adolescence. It is estimated that 82 percent of people who die of coronary heart disease are 65 and older. Simultaneously, the risk of stroke doubles every decade after age 55.

Sex : Men are at greater risk of heart disease than pre-menopausal women. Once past menopause, it has been argued that a woman's risk is similar to a man's although more recent data from the WHO and UN disputes this. If a female has diabetes, she is more likely to develop heart disease than a male with diabetes.

Resting blood pressure : Over time, high blood pressure can damage arteries that feed your heart. High blood pressure that occurs with other conditions, such as obesity, high cholesterol or diabetes, increases your risk even more.

Fasting blood sugar : Not producing enough of a hormone secreted by your pancreas (insulin) or not responding to insulin properly causes your body's blood sugar levels to rise, increasing your risk of heart attack.

Cholesterol : A high level of low-density lipoprotein (LDL) cholesterol (the "bad" cholesterol) is most likely to narrow arteries. A high level of triglycerides, a type of blood fat related to your diet, also ups your risk of heart attack. However, a high level of high-density lipoprotein (HDL) cholesterol (the "good" cholesterol) lowers your risk of heart attack.

Resting ECG : For people at low risk of cardiovascular disease, the USPSTF concludes with moderate certainty that the potential harms of screening with resting or exercise ECG equal or exceed the potential benefits. For people at intermediate to high risk, current evidence is insufficient to assess the balance of benefits and harms of screening.

Max heart rate : The increase in the cardiovascular risk, associated with the acceleration of heart rate, was comparable to the increase in risk observed with high blood pressure. It has been shown that an increase in heart rate by 10 beats per minute was associated with an increase in the risk of cardiac death by at least 20%, and this increase in the risk is similar to the one observed with an increase in systolic blood pressure by 10 mm hg.

ST depression : In unstable coronary artery disease, st-segment depression is associated with a 100% increase in the occurrence of three-vessel/left main disease and to an increased risk of subsequent cardiac events. In these patients an early invasive strategy substantially decreases death/myocardial infarction.

VISUALIZATIONS/DASHBOARD
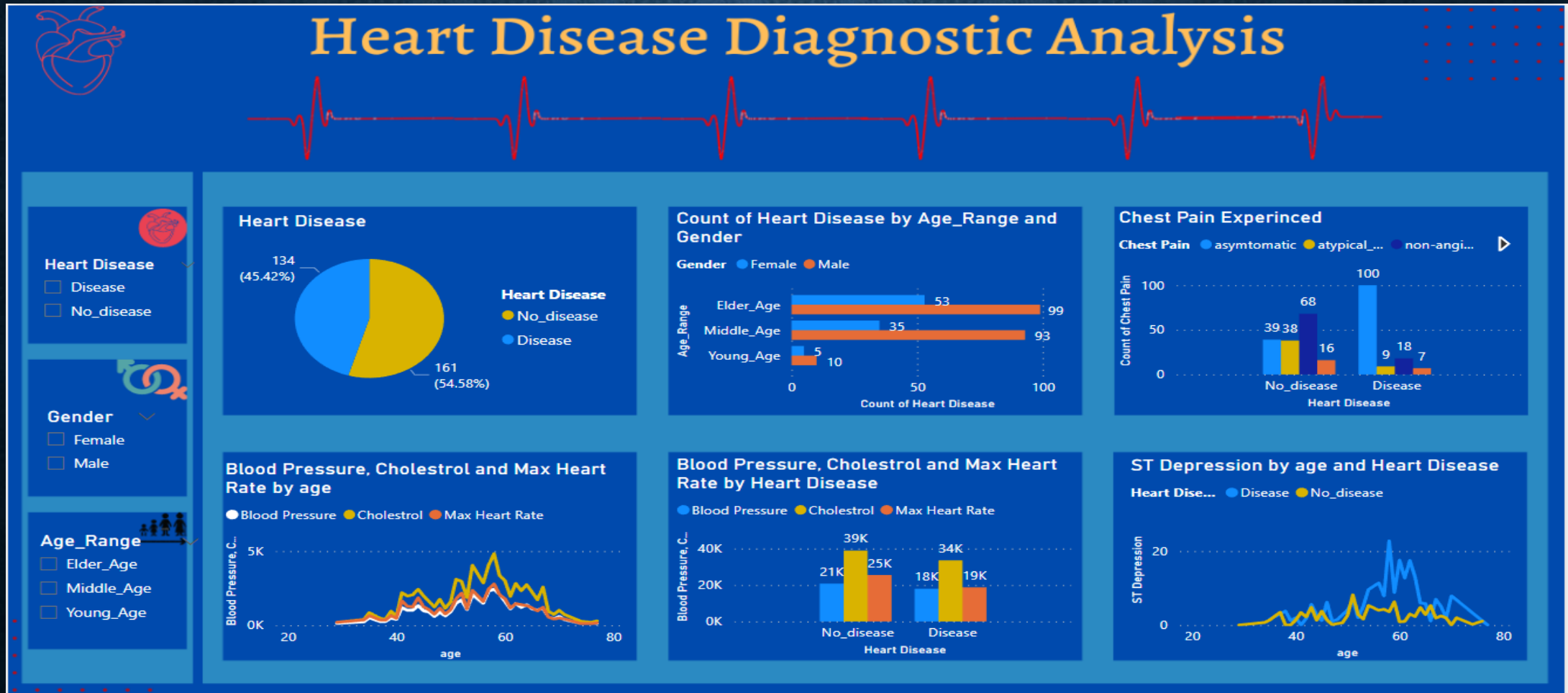
KPI

- Percentage of People Having Heart Disease

- Heart Disease Based On Age Distribution

- Heart Disease Based On Gender Distribution

- Chest Pain Experienced By People Suffering from Heart Disease

- Blood Pressure and Cholesterols Level and maximum heart rate played important role in heart diseases

CONCLUSIONS

1.From the overall population, 46% people having heart disease & 56% people having no heart disease

2.Elder Age People are most affected by Heart Disease & Middle Age People are mostly FREE from any kind of Disease

3.Males are more prone to Heart Disease than female

4.People having asymptomatic chest pain have a higher chance of heart disease

5.Elderly Age people have Higher number of Asymptomatic Pain

6.Higher Blood Pressure Level Have Chances Of Heart Disease

7.Blood Pressure increases between age of 50 to 60 and continue till 70

8.Cholestrol Level higher in female than male

9.Higher Cholestrol Level have chances of heart disease

10.High ST Depression have chances of heart disease

11.ST Depression mostly increases between the age group of 30-40

12.More Fasting Blood Sugar may have chances of heart disease

13.If patient suffer from Angina, he/she may be concerned that exercise will make their symptoms worse

Q & A

Q1) What's the source of data?

Ans) The Dataset was taken from iNeuron's Provided Project Description Document.
https://drive.google.com/drive/folders/165Pjmfb9W9PGy0rZjHEA22LW0Lt3Y-Q8

Q2) What was the type of data?

Ans) The data was the combination of numerical and Categorical values.

Q 3) What's the complete flow you followed in this Project?

Ans) Refer slide 5th for better Understanding

Q4) What techniques were you using for data?

Ans) -Removing unwanted attributes -Visualizing relation of independent variables with each other and output variables -Removing outliers -Cleaning data and imputing if null values are present. -Converting Numerical data into Categorical values.

Q 5) What were the libraries that you used in Python?

Ans) I used Pandas, NumPy and Matplotlib and Seaborn libraries in Pandas.

THANK YOU