Insights of Categorical Analysis

Sex

- 207 of the patients are male, 96 are female.
- 68.32% of the patients are male and 31.68% are female.
- The number of male patients in the data set is 2.15 times the number of females.

CP (Chest Pain)

- Almost half of the patients in the dataset have **asymptomatic** chest pain.
  - Observations show us that **143** patients experienced this type of pain and the ratio to other types of chest pain is **47.19%**.
- Just over a quarter of patients have **atypical angina**.
  - Observations show us that **87** patients experienced this type of pain and the ratio to other types of chest pain is **28.71%**.
- The remaining two chest pain types are less than the other two. These types are **typical angina** and **non-anginal** pain.
  - The number of patients with typical angina is **50** and the ratio to all patients is **16.50%**.
  - Lastly, The number of patients with typical angina is **23** and the ratio to all patients is **7.59%**.

FBS (Fast Blood Sugar)

- If a patient's fasting blood sugar is lower than 120 mg/dl, the patient has no diabetes (0/False). On the other hand, if a patient's fasting blood sugar is higher than 120 mg/dl, the patient has diabetes (1/True).
  - Individuals with diabetes (fbs:1/True) in the data set are 45 people with a rate of 14.85%.
  - There are **258** people in the dataset who do not have diabetes (fbs:0/False), and their rate is **85.15%**.

---

restecg (Resting Electrocardiographic)

- The first thing that stands out is that the two values are almost equal and make up almost half of the data set. Another thing is that the third value is negligible.
  - The number of patients with normal resting electrocardiography results is **152**, with a rate of **50.17%**.
  - The percentage of patients with hypertrophy value is **48.51%** and their number is **147**.
  - The negligible result is having ST-T wave abnormality. This abnormality has only **4** patients and the rate is **1.32%**.

Exang (Exercise induced angina)

- According to the variable "exang," the pain caused by this angina is represented by a value of 1 if it occurs with any exercise and 0 if it does not.

- - Observations show us that there are 2.06 times less patients with exang than those without. -The number of "Exang" patients is **99** with a rate of **32.67%, those** who don't are **204** with a rate of **67.33%**.

---

slp (the slope of the peak exercise ST segment)

- The lowest value observed is in patients with a downsloping ST wavelength, with a number of **21** patients, and the rate is **6.93%**.
- Once again the remaining two variables are almost equal.
  - The number of patients with flat/stable ST wavelength and upward sloping ST wavelength is **142-140** patients and the rate is **46.86%-46.20%**, respectively.

---

caa (number of major vessels)

- This variable is the number of major vessels coloured by fluoroscopy.
  - More than half patients (175 patients - 57.76%) the number of major vessels is 0.
  - The second highest percentage (21.45%) is value major vessel number 1 with 65 patients.

---

thall (Thallium stress test)

- The thallium stress test is simply an imaging method that evaluates the amount of blood reaching the heart muscle and determines whether a person has coronary artery disease.
- There are three observation values in the description of this variable. However, the pie chart shows four values. Values 0, 1, 2 and 3.
- According to our research, the observation value of 0 is null. Therefore, in the next step, 0 observation values will be returned to null and filled with logical data.
- Observation shows that **54.79%** of **166** patients had a normal Thallium stress test result.
- **36.61%** (**117** patients) have a reversible defect Thallium stress test result.
- **5.94%** (**18** patients) have a fixed defect Thallium stress test result.

Output

- More than half of the patients, **54.46% (165 patients)**, have a heart attack risk. The remaining **45.54% (138 patients)** have no heart attack risk

Insights of Numerical Analysis

Age

- The vast majority of patients are between 50 and 60.
- Mean of age is somewhere between 54-55 and the median is most-likely 55.
- According to boxplot and histogram it looks like there is no outlier data in the age variable.

trtbps (resting blood pressure (in mm Hg on admission to the hospital))

- trtbps value in most patients is between 120-and 140
- Mean of trtbps is around 131 and median of trtbps very likely 130
- According to boxplot and histogram it seems that there may be two outlier data after 180 trtbps.

chol (serum cholesterol in mg/dl)

- The cholesterol of most patients is generally between 200 and 280.
- Both mean and median of chol is somewhere around 240.
- Highest values can be considered an outlier.

thalachh (maximum heart rate achieved)

- The maximum heart rate achieved in most patients is between 140-170
- Both mean and median of chol is approximately 150.
- The value under 80 can be considered an outlier.

oldpeak (ST depression induced by exercise relative to rest)

- Values of the vast majority of patients are between 0 and 1.5.
- Mean of oldpeak is like 1 and median of trtbps more less 0.75
- Values above 3 can be considered outliers.