

**Subject :- Programming In Python**

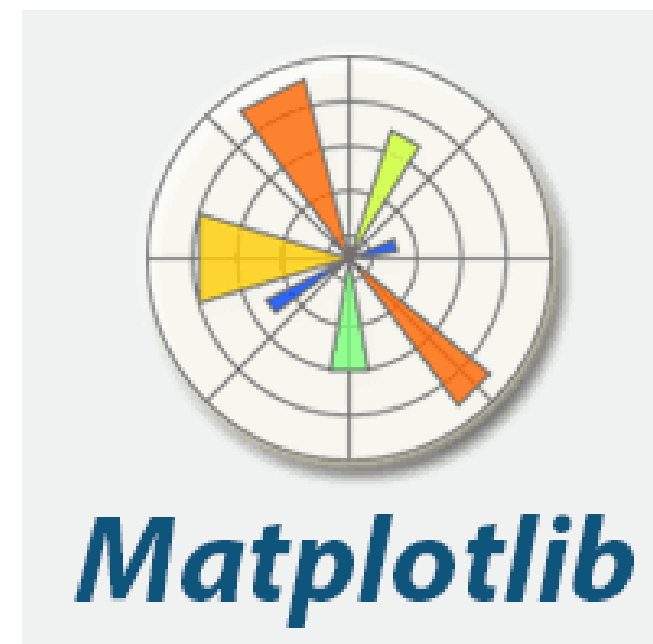
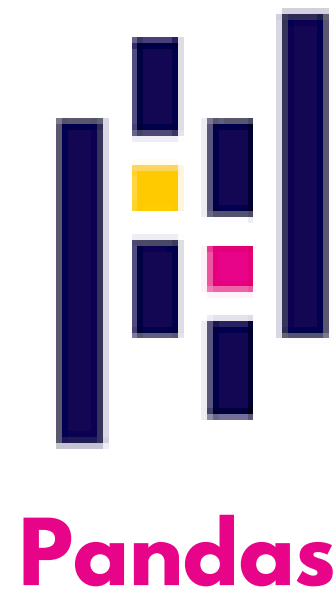
**Guide Name:- Sanket Shah**

**Div :- C2 - 62**

# **ADVANCED LIBRARIES IN PYTHON**

Presented By Janmejay Pandya (2301031800059)

4th Semester



# INTRODUCTION

Data Analysis using Advanced Python Libraries

**Objective:** Analyze a heart disease dataset and extract meaningful insights using advanced Python libraries like Pandas, NumPy, Matplotlib, and Seaborn

**Importance:**

- Understand the significance of data analysis in healthcare.
- Learn how Python libraries simplify complex data handling.

# UNDERSTANDING THE PROBLEM

## PROBLEM STATEMENT:

Analyze the dataset to gain insights into factors contributing to heart disease.

## GOALS:

- Identify trends and relationships between features.
- Visualize data for better understanding.

# OVERVIEW OF DATASET

## MARKET TRENDS

- Source: [heart.csv](#)
- Key Features: Age, sex, chest pain type, cholesterol, fasting blood sugar, resting ECG, maximum heart rate, exercise-induced angina, etc.
- Target Variable: Presence of heart disease (binary classification).



# LIBRARIES USED IN THE PROJECT

## **NUMPY**

**For numerical operations (e.g., array handling).**

## **PANDAS**

**For data manipulation, including loading and exploring the dataset**

## **MATPLOTLIB**

**For creating basic plots and charts.**

## **SEABORN**

**For advanced statistical visualizations.**

# DATA LOADING AND INITIAL EXPLORATION

## EXPLANATION

This code loads the dataset and displays the first few rows for an initial overview.

```
import pandas as pd
df = pd.read_csv('heart.csv')
print(df.head())
```

## CONCISE SUMMARY

```
df.info()
```

- The total number of rows and columns.
- Column names and data types (e.g., int, float, object).
- Non-null count of each column (helps identify missing data).
- Memory usage of the dataset.

```
print(df.columns.values)
```

List all the columns to understand the dataset's available features

# OUTCOMES OF THE DATASET

"1.how many people have heart disease and how many people doesn't have heart disease?",

"2. people of which sex has most heart disease?",

"3. people of which sex has which type of chest pain most?",

"4. people with which chest pain are most prone to have heart disease?",

"5. What is the average age of people with heart disease versus those without?",

"6. Which 'thal' (thalassemia) type is most associated with heart disease presence?",

"7. Do people with 'exang' (exercise-induced angina) have a higher likelihood of heart disease?",

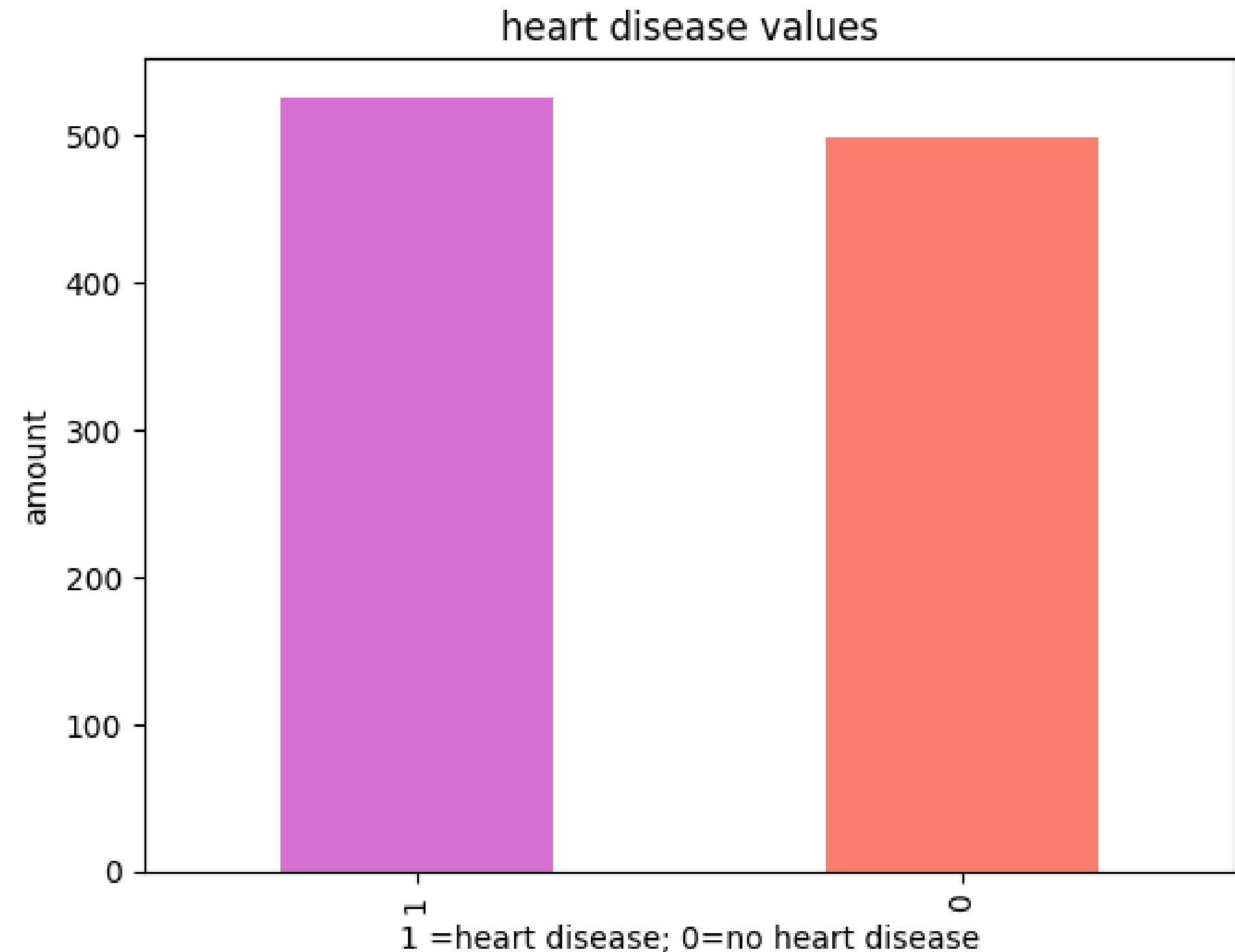
"8. How does 'oldpeak' (ST depression) vary between people with and without heart disease?"

# DEMONSTRATION OF DATASET

#1.how many people have heart disease and how many people doesn't have heart disease?

```
df.target.value_counts().plot(kind = 'bar',color=[  
    "orchid","salmon"])  
plt.title("heart disease values")  
plt.xlabel("1 =heart disease; 0=no heart disease")  
plt.ylabel("amount");
```

creates a bar plot using Matplotlib and Pandas to visualize the distribution of heart disease cases in the dataset.





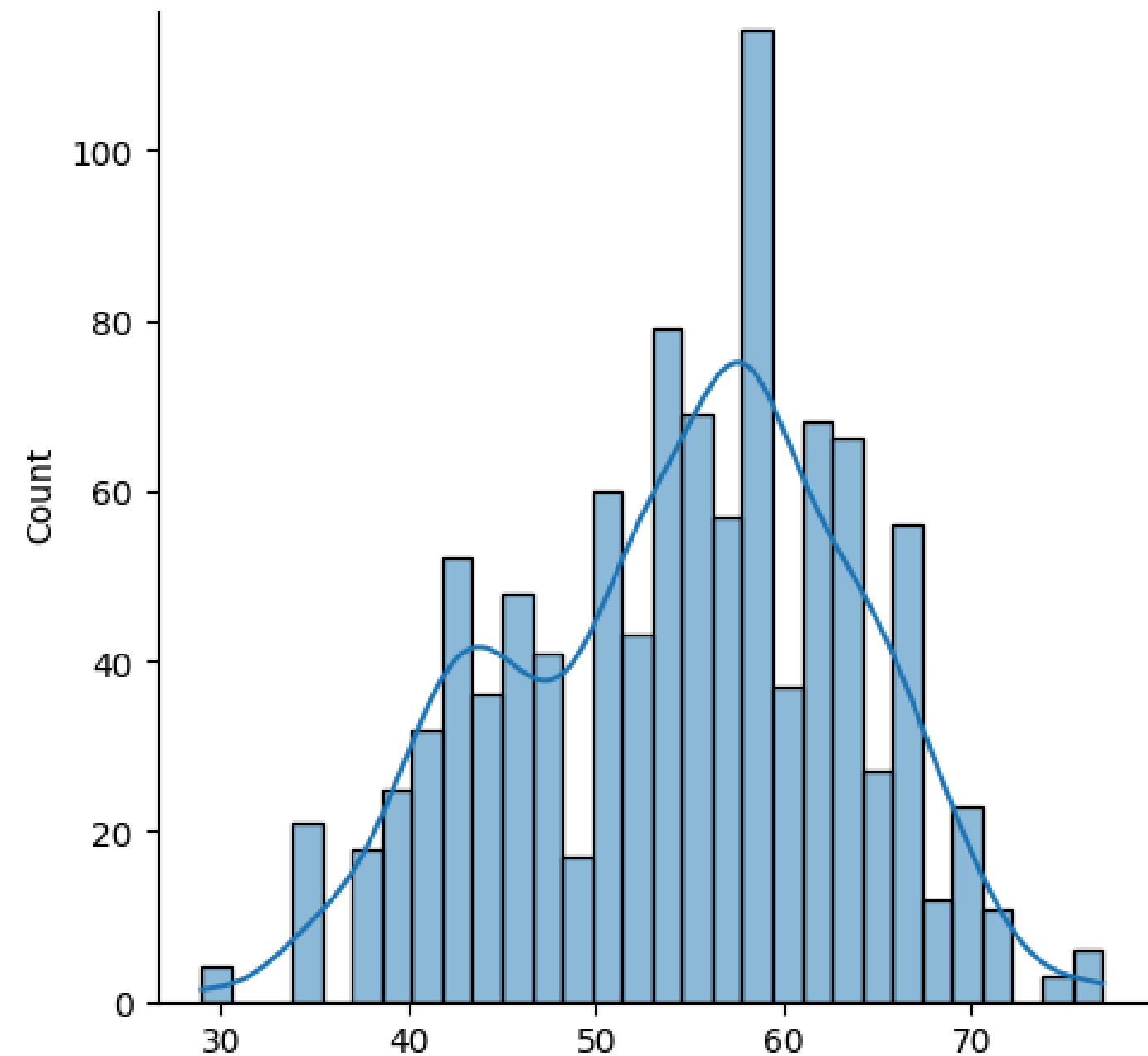
# ADVANCED VISUALIZATION WITH SEABORN

people in which age pain are most pron to have heart disease?

```
#create a distribution plot with normal diss. curve  
import seaborn as sns  
sns.displot(x = 'age', data = df, bins = 30, kde = True);
```

a distribution plot (histogram) with a normal distribution curve (KDE) overlay using Seaborn to visualize the distribution of the 'age' feature in the dataset.

#58-59 years old are most in dataset.



# KEY INSIGHTS FROM DATA ANALYSIS

## EXAMPLE INSIGHTS

- Age and cholesterol are important factors in determining heart disease risk.
- Patients with higher chest pain levels tend to have higher risks.
- The target variable is slightly imbalanced.



# FUTURE SCOPE

Apply machine learning algorithms to predict heart disease.

Implement additional data preprocessing techniques for improved analysis.

**GitHub Link:**

**<https://github.com/JanmejaySocet/Advanced-Libraries/tree/main>**

# CONCLUSION

## Summary:

Successfully analyzed the heart disease dataset.

Gained hands-on experience with advanced Python libraries.

Extracted meaningful insights through data visualizations.

**Key Learning:** Python libraries make data handling and visualization much easier and more efficient.



# THANK YOU

Presented By Janmejay Pandya  
(2301031800059)