**What is Python?**

**Python** is a **high-level**, **interpreted**, and **general-purpose programming language** known for its **simple and readable syntax**. It was created by **Guido van Rossum** and first released in **1991**.

Python is very popular among beginners and professionals because it is:

* Easy to learn and write
* Versatile (used in many fields)
* Open-source and has a large community
* Rich in libraries and frameworks

# Uses of Python

Python is used in a wide range of areas:

1. **Artificial Intelligence & Machine Learning**
   * Libraries: TensorFlow, PyTorch, Scikit-learn
   * Used to build models that can learn and make predictions
2. **Web Development**
   * Frameworks: Django, Flask, FastAPI
   * Used to create dynamic websites and REST APIs
3. **Data Science & Data Analysis**
   * Libraries: Pandas, NumPy, Matplotlib, Seaborn
   * Used to analyze, visualize, and manipulate data
4. **Automation & Scripting**
   * Automate boring or repetitive tasks
   * Example: Web scraping, file management, email sending
5. **Desktop Applications**
   * Libraries: Tkinter, PyQt
   * Used to create GUI-based applications
6. **Game Development**
   * Libraries: Pygame
   * Used to build simple 2D games

**What is NumPy?**

NumPy (short for Numerical Python) is a powerful Python library used for numerical and scientific computing. It provides support for:

* + Multidimensional arrays (ndarrays)
  + Mathematical functions
  + Linear algebra
  + Random number generation

NumPy is fast and efficient, and it’s the foundation of many other libraries like Pandas, TensorFlow, and Scikit-learn.

# Key Features of NumPy

* Powerful **N-dimensional arrays**
* Very **fast** operations (written in C under the hood)
* Supports **vectorized operations** (no need for loops)
* Tools for working with **matrices, Fourier transforms, and statistics**

**What is Pandas?**

**Pandas** is a powerful **open-source Python library** used for **data analysis and data manipulation**. It is built on top of **NumPy** and makes working with **structured data** (like tables or spreadsheets) easy and efficient.

# Core Data Structures in Pandas

1. **Series** – 1D labeled array (like a single column)
2. **DataFrame** – 2D table with rows and columns (like an Excel sheet)

# Key Features of Pandas

|  |  |
| --- | --- |
| **Feature** | **Description** |
| **DataFrame Support** | Handles labeled 2D data like SQL tables or Excel files |
| **Data Cleaning** | Handle missing data, filter, replace, or fill values |
| **Data Aggregation** | Grouping, summarizing, counting, mean, etc. |
| **Data Analysis** | Perform statistical operations and data exploration |
| **File Handling** | Easily read/write from CSV, Excel, JSON, SQL, etc. |

**Feature Description**

**Time Series Handling** Built-in support for date/time functions and indexing

**Data Visualization** Integrates well with libraries like Matplotlib and Seaborn for plotting **Fast Operations** Built on top of NumPy, optimized for performance

**What is Matplotlib?**

**Matplotlib** is a Python library used for **data visualization**. It lets you create **line charts, bar charts, scatter plots**, and many other types of graphs.

# Key Features of Matplotlib

**Feature Description**

Plotting Variety Line, bar, scatter, pie, histogram, etc.

Customization Control over colors, labels, legends, etc. Subplots Support Create multiple plots in one figure

Exporting Save plots as PNG, PDF, SVG, etc. Annotating Graphs Add labels, arrows, and custom styling

**What is Seaborn?**

**Seaborn** is a Python visualization library built **on top of Matplotlib**. It provides a **high-level interface** for making beautiful and informative statistical graphics.

# Key Features of Seaborn

**Feature Description**

Easy Plotting One-liner plots like histograms and box plots

Statistical Graphs Built-in support for correlation and regression

Beautiful Themes Pre-set styles and color palettes Works with Pandas Supports DataFrames directly Complex Plots Made Easy Heatmaps, pairplots, violin plots, etc.

**What is SciPy?**

**SciPy** (Scientific Python) is a Python library used for **scientific and technical computing**. It builds on NumPy and adds advanced features.

# Key Features of SciPy

**Feature Description**

Advanced Math Integration, differentiation, optimization Statistics Probability distributions, statistical tests Linear Algebra Matrix operations, eigenvalues, etc.

Signal & Image Processing Tools for filtering, image analysis, etc.

Performance Fast calculations built in C/Fortran

**What is Scikit-learn (sklearn)?**

**Scikit-learn** is a powerful **machine learning library** in Python. It provides simple tools for **classification, regression, clustering**, and more.

# Key Features of Scikit-learn

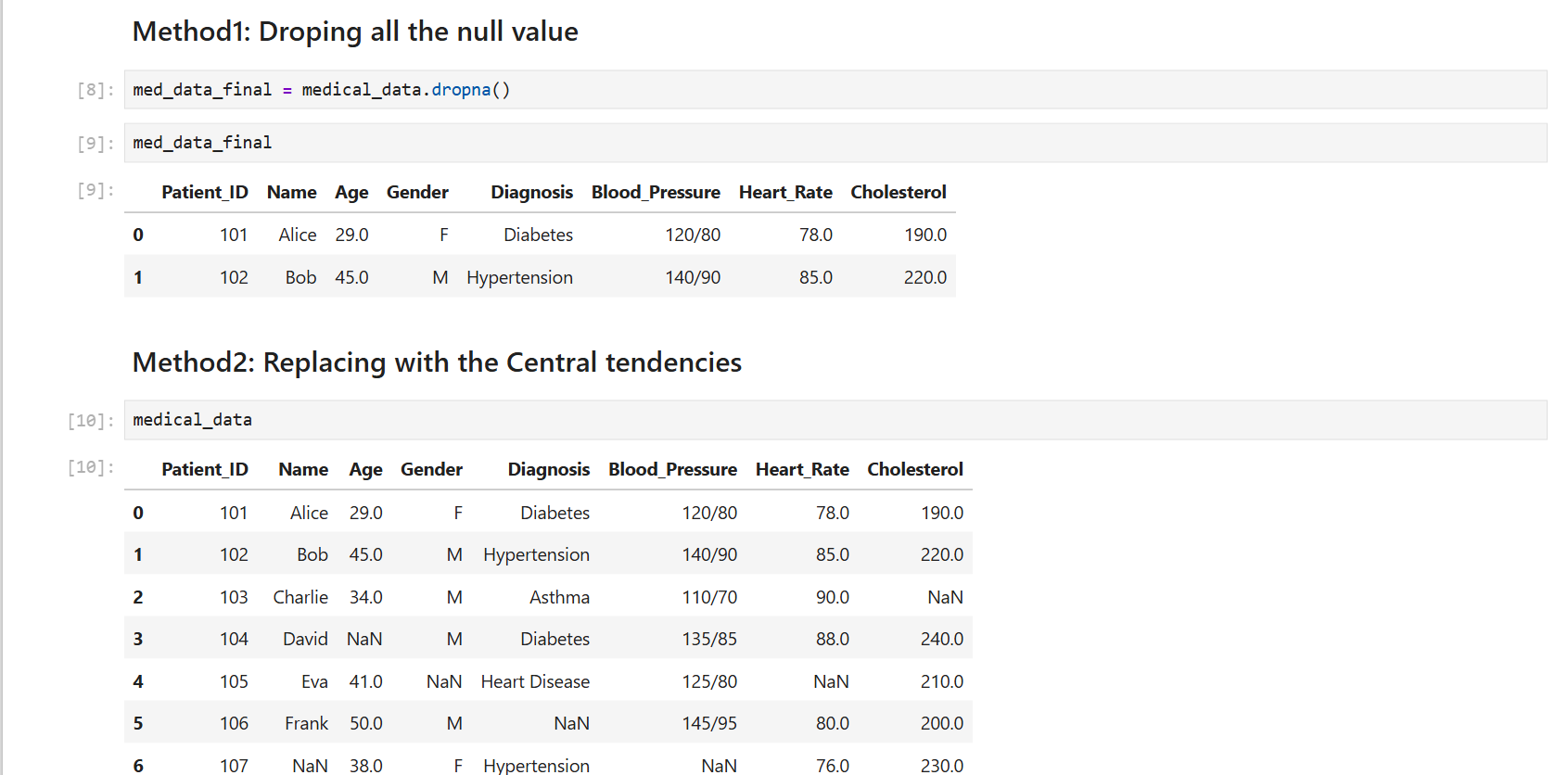
|  |  |
| --- | --- |
| **Feature** | **Description** |
| Machine Learning | Classification, regression, clustering |
| Model Tools | Training, testing, model evaluation |
| Preprocessing | Feature scaling, encoding, normalization |
| Model Selection | Cross-validation, hyperparameter tuning |
| Pipeline Support | Combine multiple steps into one workflow |

**1.Write a Python program to calculate the Mean, Median, and Standard Deviation of a dataset using a NumPy array.**



**2)Write a Python program to read a CSV file, clean the data by handling missing values, and perform basic data analysis using the Pandas library.**



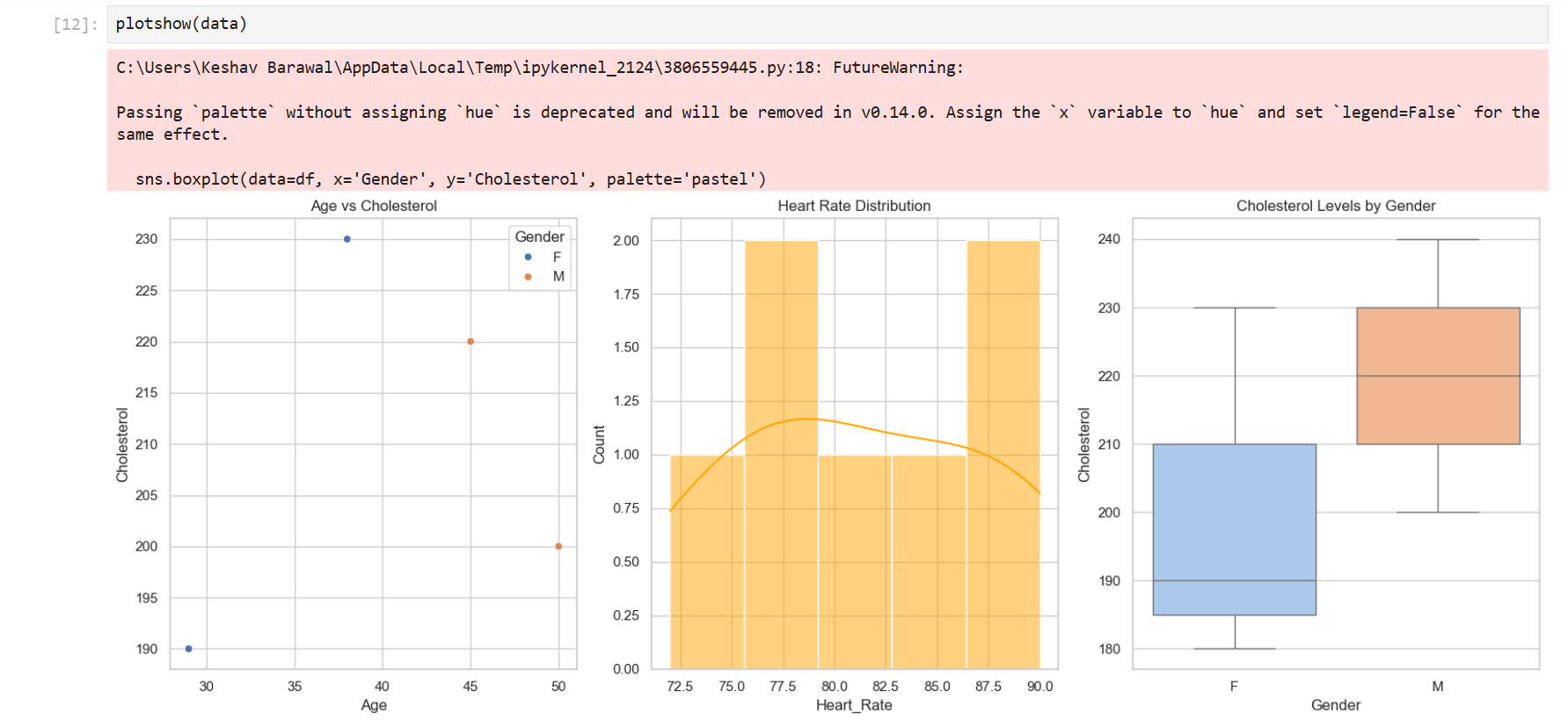




**3.) Create a Python program to plot a scatter plot, histogram, and boxplot using Matplotlib and Seaborn.**



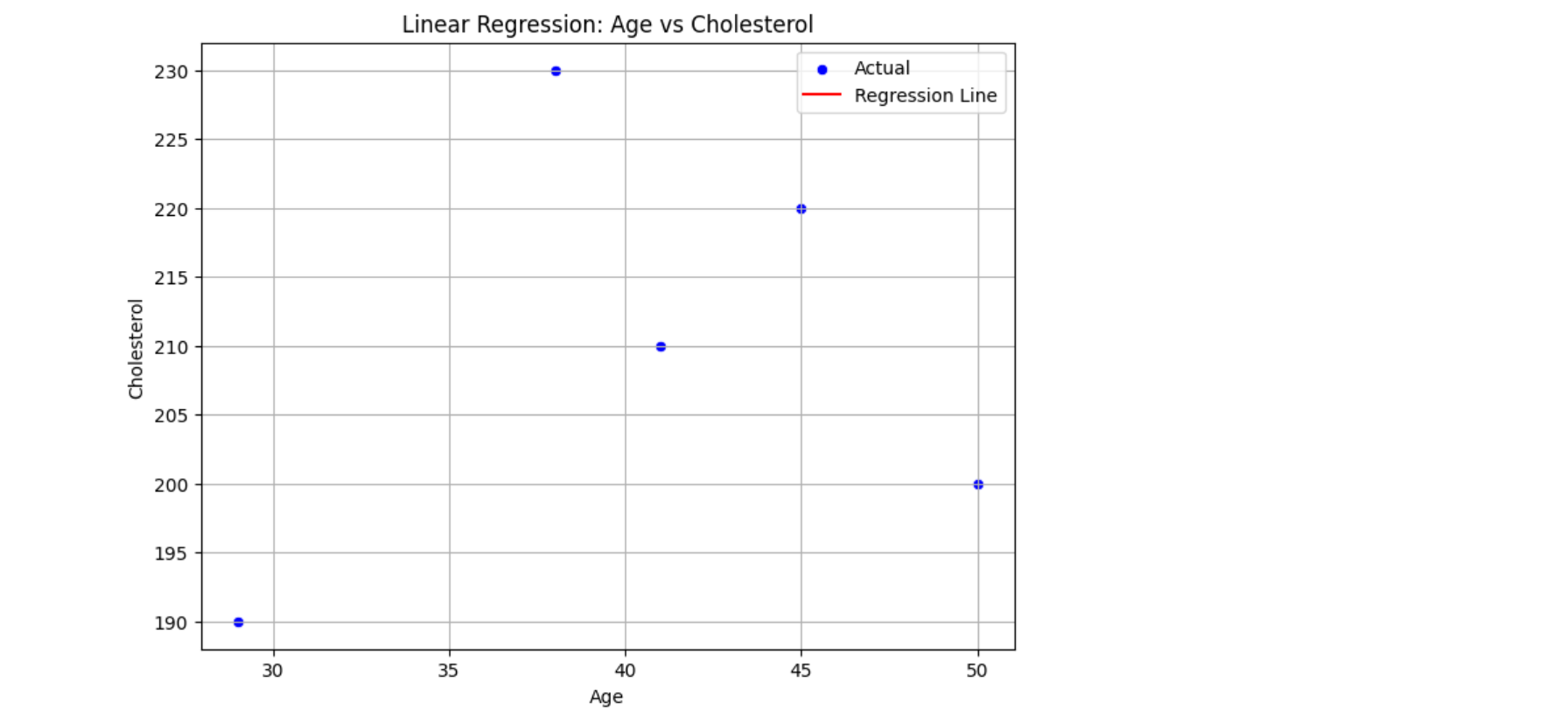




4) Write Python Programme to implement linear regression.

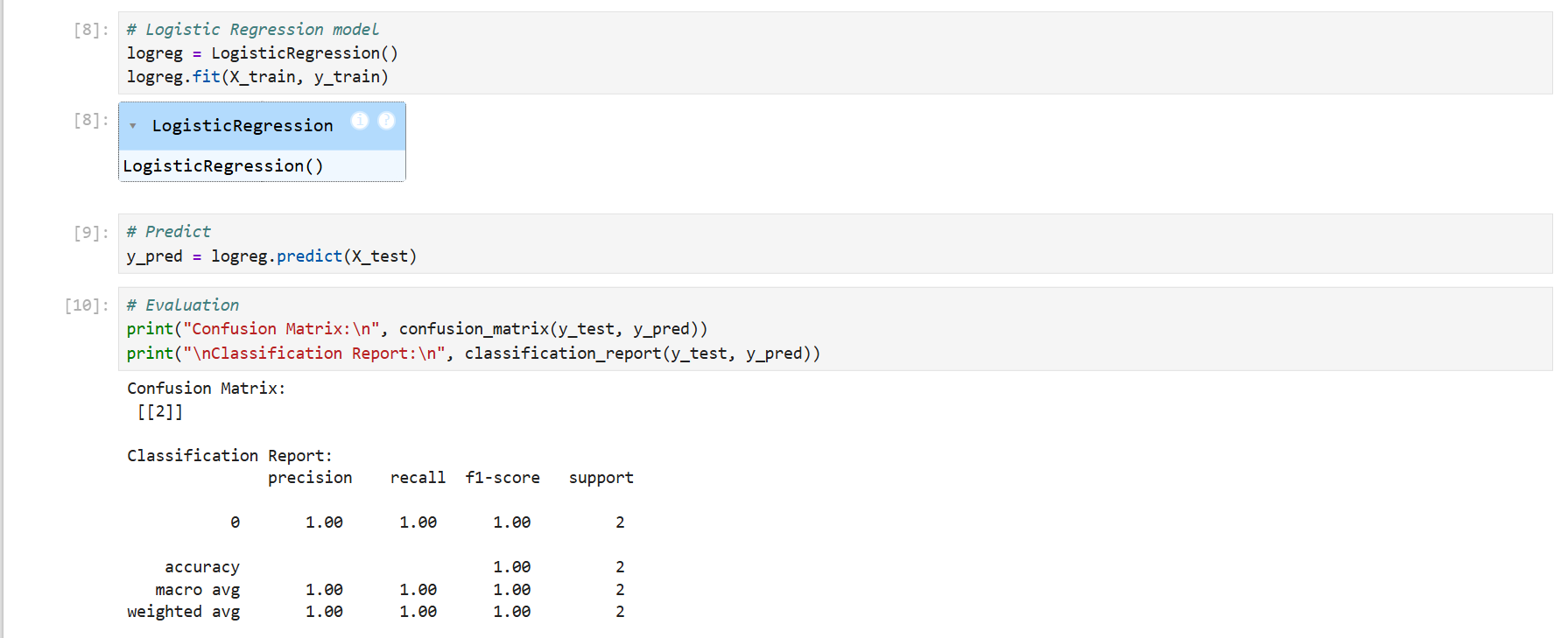


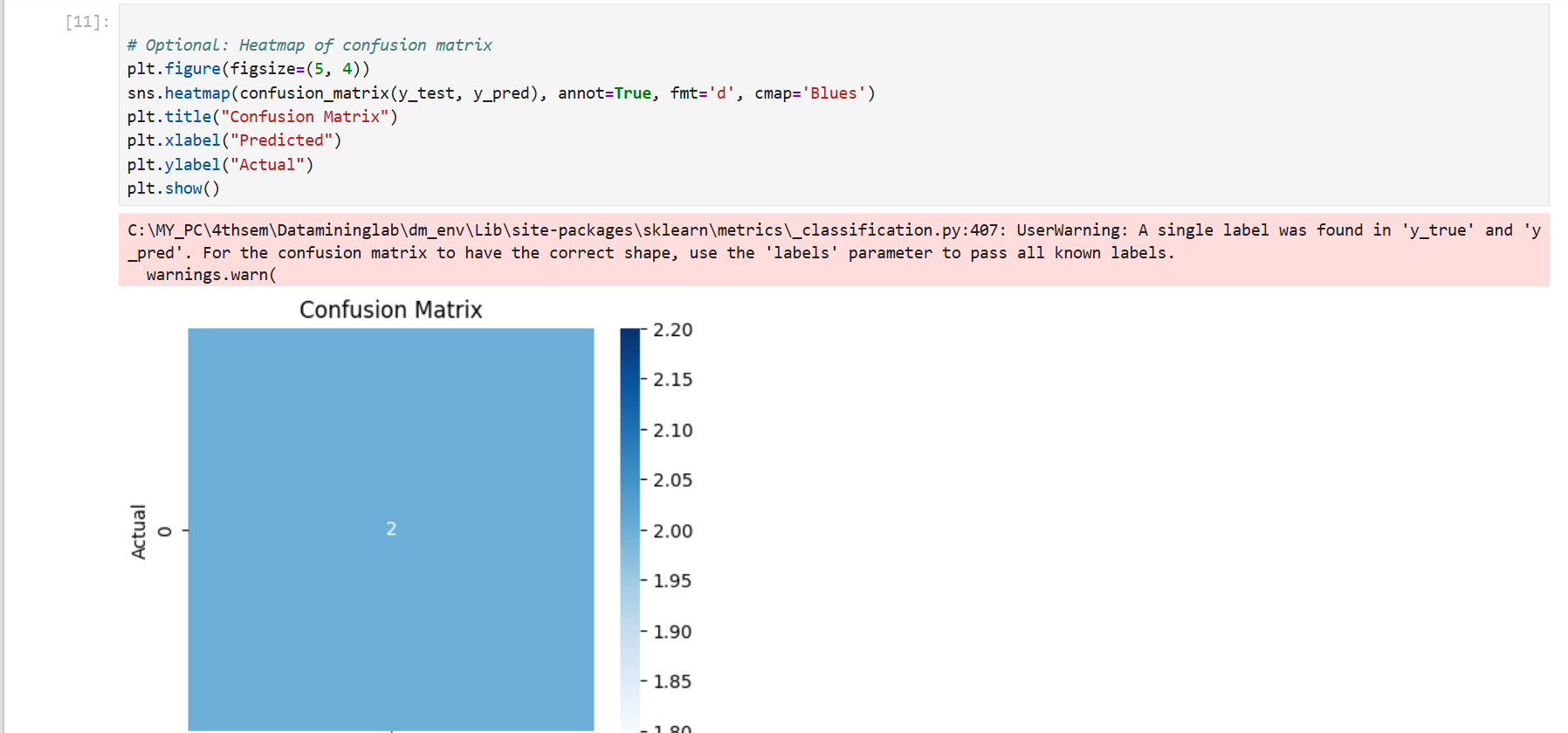




5) Write a python programme to implement logistic regression.

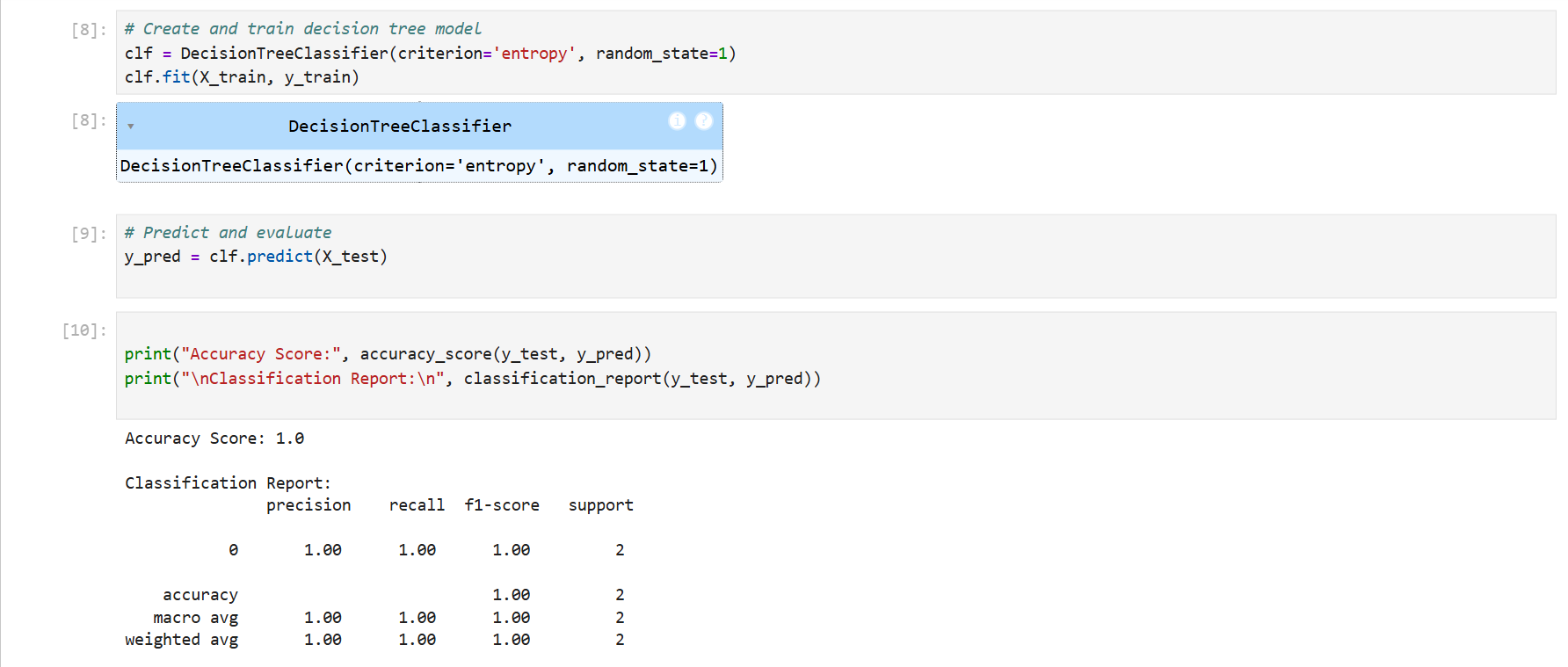


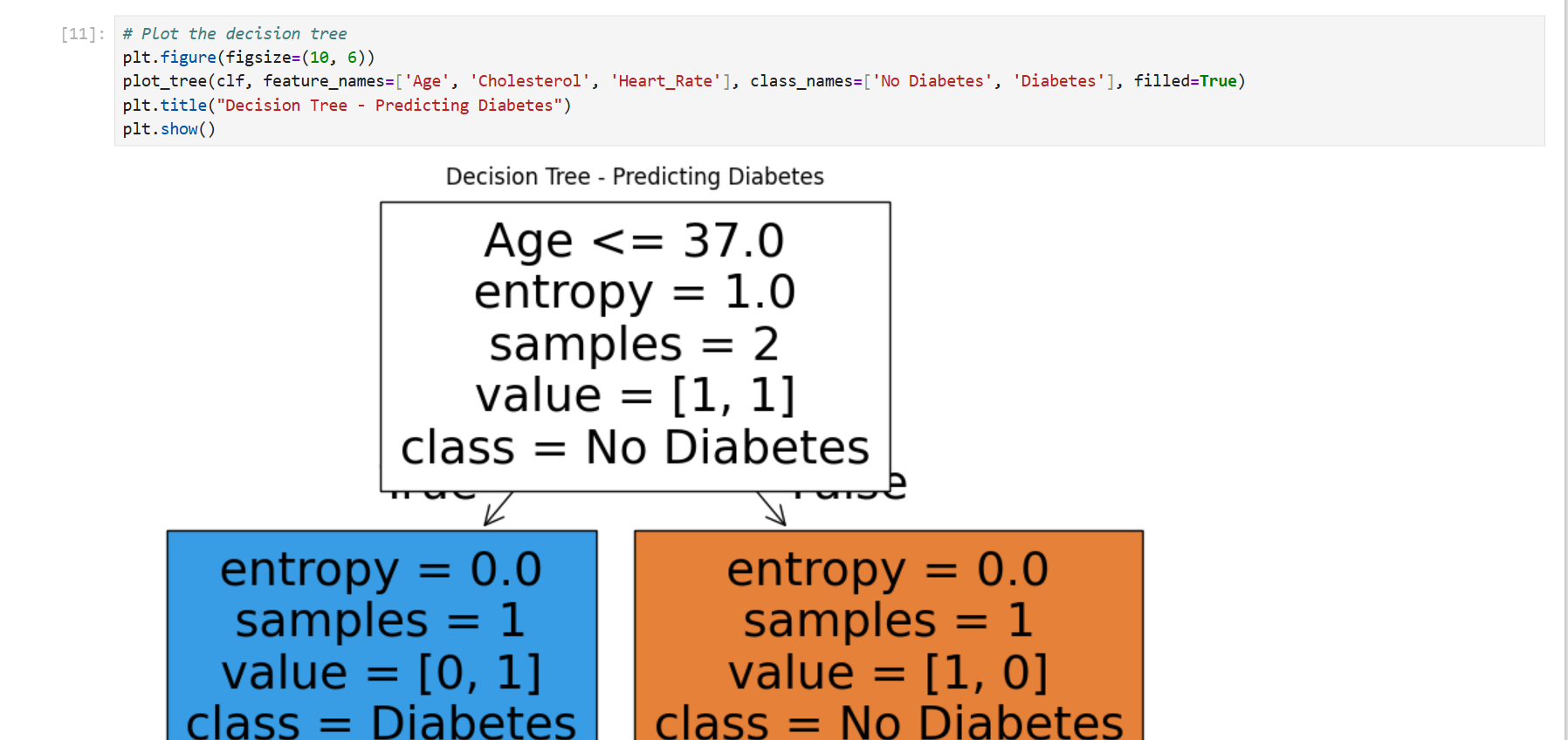




6) Write python programme to implement decision tree.







7) Write python programme to implement Ensemble Techniques.



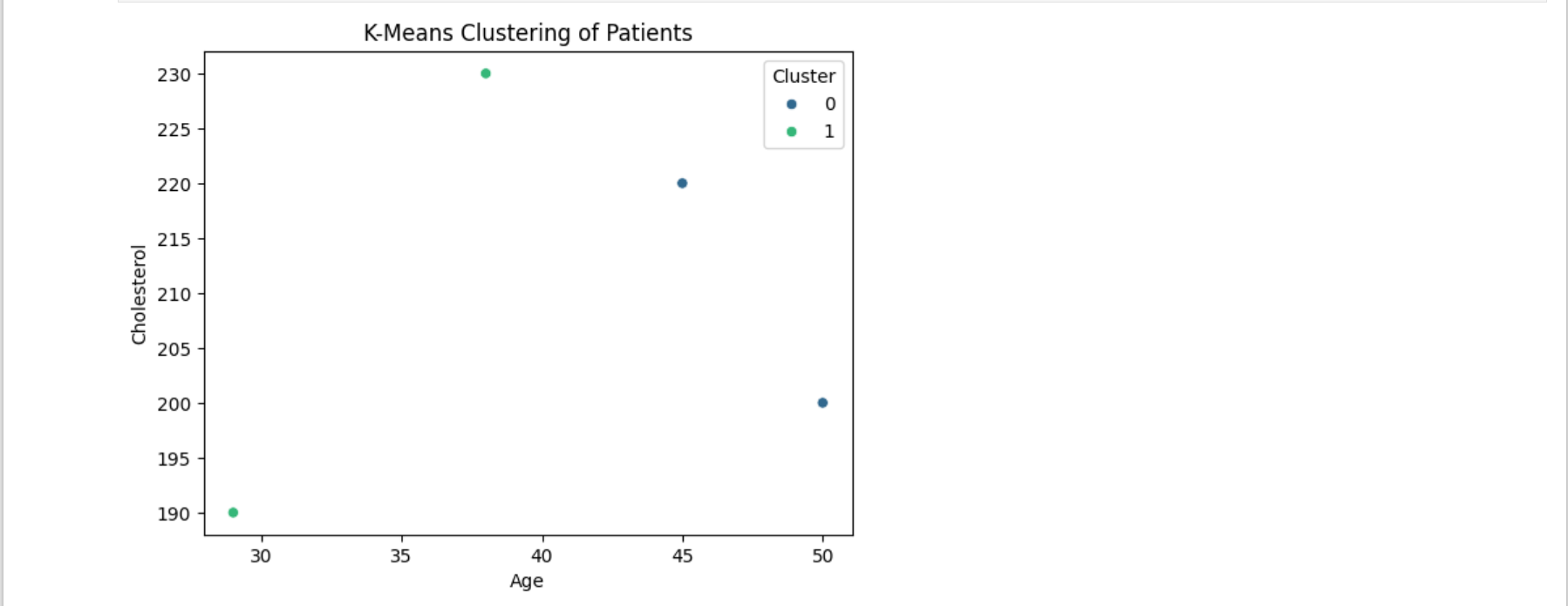




8) Write a python program to implement clustering.







9) Write a python programme to implement SVM.



