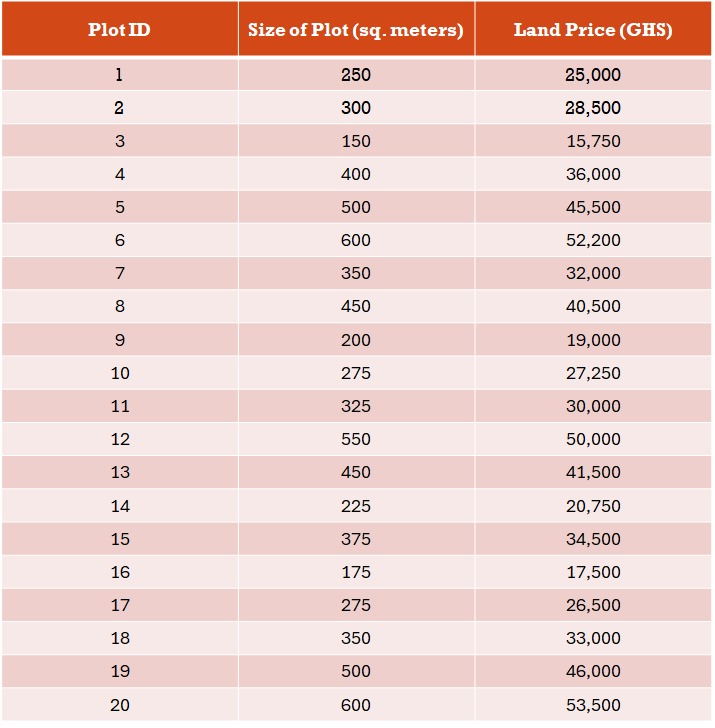
**Assignment 1: Pandas and Matplotlib for Data Analysis**

**Total Marks: 60 marks**

This assignment is designed to help you get comfortable with using **Pandas** for data manipulation and **Matplotlib** for visualizing data. The tasks will require you to create and manipulate DataFrames, as well as visualize the data with various types of plots.

**Question 1 (20 marks)**

1. Using the data set provided below create a **Pandas DataFrame**. The dataset contains information about land size (in square meters) and corresponding prices (in cedis) [10 marks]



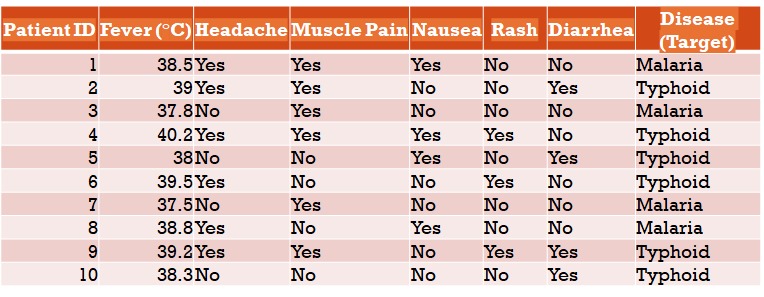
1. Create a scatter plot using Matplotlib library to visualize the dataset [10 marks]

**Question 2 (40 marks)**

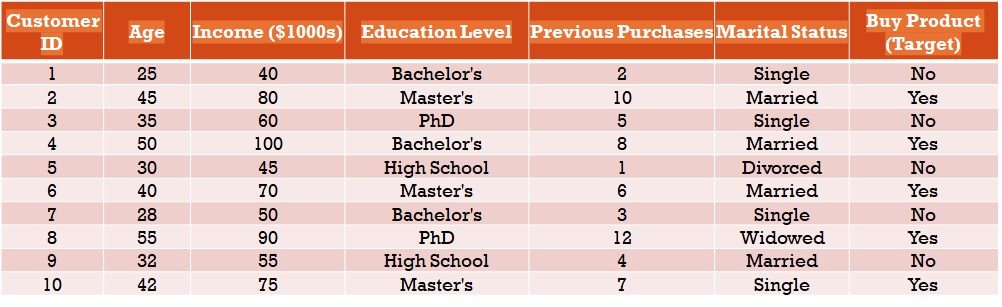
You are given two datasets:

* **Fig 2**: Dataset of hospital patients
* **Fig 3**: Dataset of store customers

1. Create a classification plot for both datasets. Use Matplotlib or any suitable classification plot function to visualize the data. [15 marks]
2. In your classification plot, identify which features are most significant in classifying the data. Explain your rationale for choosing the important features. [15 marks]
3. Evaluate whether it is better to use all the features or only some features when performing the classification. Explain your reasoning in detail. [10 marks]



*Fig 2: Dataset of hospital patients*



*Fig 3: Dataset store customers*

**Submission Instruction**

1. Save a Jupyter Notebook containing your answers – both code and text.
2. Save a Python (.py) version of your Jupyter notebook.
3. Zip the two files, and save them using the correct format, stated in the assignment rubric.
4. Submit only the zip file.

Note:

* Marks obtained will be awarded based on **correctness of code**, **quality of visualizations**, **clarity of explanations**, **creativity** and **insights**.
* Helpful resources on using Matplotlib and Pandas have been made available in on the course materials page under the “Getting Started” section.
* **The use of any AI tool is strictly prohibited for this assignment.**