1. Description about the ICP:

This is ICP is about mounting Google drive to Google Colab notebook to read the data using Pandas library as well as performing any 3 analysis tasks on the given data.

1. Objective

The objective of this ICP is how to use Google Colab Notebook and upload your source code on Github repository. In addition, to do a python practice on the provided data.

1. Design/ implementation (include screenshots with explanations)

In the first step, I imported the required libraries and mounted Google drive to Google Colab. Also, I’ve used .head function to print out the first 6 rows.

Table

Description automatically generated

Task1: I used .describe to show the basic statistic details of the data such as Min, Max, and standard division values in each column. Since the IDs can’t be sum, and the Unnamed: 32 column contains null values, I removed the two columns from the data.

Table

Description automatically generated

Task 2: I utilized the seaborn library to plot the (perimeter\_mean) column. As shown in the screenshot below the density of the mean size of the core tumor. The highest density of the perimeter\_mean was between the 75-85.

Graphical user interface, text, application, email

Description automatically generated

Task3: I counted all the number who has breast tissues and diagnosed as (M = malignant, B = benign). As we can see there are about 210 who has diagnosed as malignant and the rest are benign.

Chart, bar chart

Description automatically generated

1. Video link:

I’ve explained my source code to the instructor during the class time.