BAIS 3020 - Computational Thinking
Lab Activity 3: Heart Attack Analysis

The purpose of this assignment is to analyze medical data, given to you in the file hearts.csv. This is a synthetic dataset regarding heart attacks. In this assignment you will learn how to perform basic analysis with Python and Pandas library and generate graphics of the analysis results using Matplotlib. There is no input from the user for this app. Instead, your app will read the input file and analyze the data according to the following instructions.

Use Spyder IDE to complete this assignment. Start by creating a folder on your local computer called LabActivity3. Then download the data file called heart.csv from this ICON Assignment. Place the file in the folder you created. Next, create a file called **main.py** and place it in the LabActivity3 folder. The file should have a main function which is called according to the value in Python's __name__ variable. The main function should start by displaying the heading and four items below in blue. These are the factors that are to be considered in the analysis of the data by the app. Notice that each of the four items specify what is considered and why it is being considered.

Heart Attack Analysis

These are the factors related to heart attacks that are considered in this data:

1. Heart Attack Percentage is calculated based on all data to see if the data is skewed.
2. Gender Percent is used to distinguish between the heart attack rates in men and women.
3. EKG is considered, as it detects heart problems that can forecast chance of future heart attacks.
4. Exercise is considered to see if likelihood of heart attack increases when physically active.

The main function should call a separate function for four statements listed above. All of the functions should be implemented in a separate module called **utilities.py**. Create this file and place it also in the LabActivity3 folder. Name the functions as follows:

- getHeart()
- getGender()
- getEKG()
- getExercise()

Read the data file into a data frame using the Pandas library at the top of the utilities.py file, as module level code. This way the data frame is available to all functions. Then proceed with answering each of the four statements above in their own function. Each function should provide results as both text output and plot(s). Simply print the text outputs but save the plots as png files.

Text Output:

Below lists the precise text output and format that you should display for each statement. The lines in blue are for your reference and your app does not need to display them again; just proceed with the output in the order listed.

1. Heart Attack Percentage is calculated based on all data to see if the data is skewed.
The percent of no heart attacks in the data: 45.54%
The percent of heart attacks in the data: 54.46%

2. Gender Percent is used to distinguish between the heart attack rates in men and women.
Male heart attacks: 43.64%
Female heart attacks: 56.36%

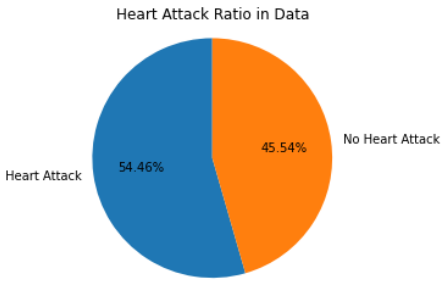
3. EKG is considered, as it detects heart problems that can forecast chance of future heart attacks.
EKG type:
    0 - EKG was normal
    1 - EKG had an abnormality with ST-T Wave
    2 - EKG had an abnormality in the Left ventricular
Resting EKG Type 0 had this percent of heart attack: 46.26%
Resting EKG Type 1 had this percent of heart attack: 63.16%
Resting EKG Type 2 had this percent of heart attack: 25.00%

4. Exercise is considered to see if likelihood of heart attack increases when physically active.
Percent who Exercised and had heart attack: 23.23%
Percent with No Exercise and had heart attack: 69.61%

Plot Output:
Below are the precise plots that each function needs to create and save. You create these plots in the utilities.py function from above. Again, the statements in blue are for your reference only. Scope of each analysis and the names of the png files are given above each plot.

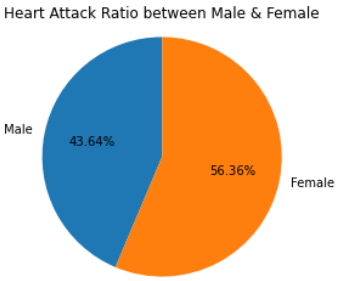
1. Heart Attack Percentage is calculated based on all data to see if the data is skewed.

HeartAttack.png – Considers all patient data.



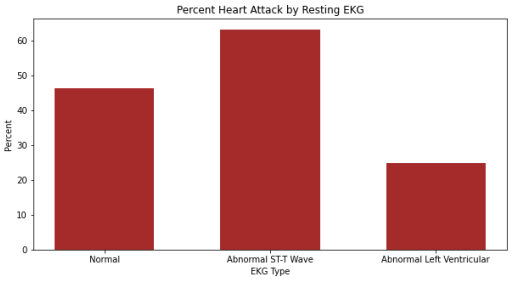
2. Gender Percent is used to distinguish between the heart attack rates in men and women.

Gender.png – Considers patients with heart attack only.



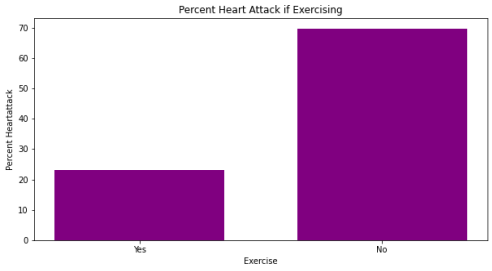
3. EKG is considered, as it detects heart problems that can forecast chance of future heart attacks.

EKG.png – Considers patients with heart attack.



4. Exercise is considered to see if likelihood of heart attack increases when physically active.

Exercise.png – Considers patients with heart attack.



Submission:
Zip the entire LabActivity3 folder, which should contain both python code files, the data file and all png image files. Upload the zipped file to the ICON Assignment dropbox.

Explanation of Data:
Each row in the data file hearts.csv represents a different patient's examination results at a single point in time. The file contains the following fields:

Column	*Description*	*Values*
age	age	
sex	gender	1: female, 0: male
cp	chest pain	0: typical chest pain, 1: atypical chest pain, 2: any pain that is not chest pain, 3: asymptomatic
fbs	fasting blood sugar	1: patient has high blood sugar, 0: regular blood sugar
restecg	resting EKG	0: EKG if normal, 1: EKG had an abnormality with ST-T Wave, 2: EKG had an abnormality in the Left ventricular
exng	exercise history	1: patient exercises regularly, 0: no regular exercise
attack	heart attack history	1: patient has had a heart attack, 0: no heart attack history