SIT720 MACHINE LEARNING, Trimester 2 2021

ASSESSMENT TASK 1 - Rubric

Questions	Excellent	Good	Fair	Unsatisfactory
Q1. *** Load the data from supplied data file.** *** Remove the observations/samples where the heart diseases are not diagnosed by the Cardiologists.** *** Print the data dimension before and after removing the observations/samples.**	Successfully completed all three tasks.	Successfully completed any two of three tasks.	Successfully completed only one of three tasks.	Fail to complete any given task.
Q2. Continue from question 1. *** Display the number of rows and their indices that have missing data in one or more cells.** *** Replace the missing data by the lowest value of the corresponding feature if it is a continuous variable.** *** In case of categorical variable, remove the sample.** *** Print the median values of all features before and after replacing missing data.**	Successfully completed all four tasks.	Successfully completed any three of the four tasks.	Successfully completed any two of the four tasks.	Completed one or no task successfully.

Q3. Continue from question 2. *** Is there any change in data type? If yes, convert them back to appropriate data types.** *** Print all variables with corresponding data type.**	Successfully completed both parts of the question.	Successfully completed one part and partially completed the other part of the question.	Successfully completed only one part of the question.	Failed to complete any tasks successfully.
Q4. Continue from question 3. Print the total numbers and ration of male and female patients who are at highest risk of heart disease.	Successfully completed the task.	Approach is correct, but the generated outputs are partially correct.	Approach is correct but failed to generate correct outputs.	Failed to understand the problem.
Q5. Continue from question 3. Is there any association between heart rate and severity of heart disease? Explain your results from given dataset.	Successfully showed the association with appropriate explanation.	Successfully showed the association but the explanation is partially correct.	Successfully showed the association with unacceptable/no explanation.	Failed to produce any acceptable output.
Q6. Continue from question 3. Print the average cholesterol level for different number of blocked blood vessels across gender. Please report the pattern found in the result, if any.	Successfully displayed the requested values with appropriate description of the pattern.	Successfully displayed the requested values with partial description of the pattern.	Successfully displayed the requested values with unacceptable/no description of the pattern.	Failed to produce any acceptable output.
Q7. Print the percentage of patients at risk of heart disease having abnormality in both ECG and blood sugar with asymptomatic chest pain.	Successfully displayed the requested values.		Approach for solving problem is correct but it does not generate the correct outputs.	Failed to understand the problem.

Q8. Calculate and print the average blood pressure of all observations with non-flat ST slopes of ECG.	Successfully displayed the requested values.		Approach for solving problem is correct but it does not generate the correct outputs.	Failed to understand the problem.
Q9. Create and print a dataframe of the heart rate, blood pressure and cholesterol levels for different age groups (based on 10 years interval).	Successfully displayed the requested values.		Approach for solving problem is correct but it does not generate the correct outputs.	Failed to understand the problem.
Q10. Continue from question 3. Find the average cholesterol level of across gender for each age group. Please explain the results.	Successfully displayed the requested values with appropriate explanation.	Successfully displayed the requested values with partial explanation.	Successfully displayed the requested values with unacceptable/no explanation.	Failed to produce any acceptable output.
Q11. Continue from question 3. Draw two scatter plots of cholesterol level, one against blood pressure and another against heart rate. Draw them in two subplots of the same plot.	Successfully plotted as requested.		Partially correct plotting.	Failed to plot the data.
Q12. Visualize the cholesterol level against number of blood vessel blocked for male and female using line plot. Explain the graph base on your observation.	Successfully plotted with appropriate explanation.	Successfully plotted with partial explanation.	Successfully plotted with unacceptable/no explanation.	Failed to plot the data.

Q13. Draw a group bar diagram of heart rate, blood pressure and total number of patients, based on age groups defined in question 9. Explain your observation from the graph.	Successfully plotted with appropriate explanation.	Successfully plotted with partial explanation.	Successfully plotted with unacceptable/no explanation.	Failed to plot the data.
Q14. Continue from question 9. *** Add two more columns named ['num_male_patients', 'num_female_patients'] and having values of the number of male and female patients affected by heart disease in each age group respectively.** *** Save the combined dataset to a csv file named 'age_group_stat.csv' in the same directory of your code file.**	Successfully completed both parts of the question.		Successfully completed only one part of the question.	Failed to complete any parts.
Q15. Continue from question 1. *** Replace all the rows where the 'state' is null with its immediate previous row.** *** Display and save the resultant dataset to a csv file named 'clean_data.csv' in the same directory of your code file.**	Successfully completed both parts of the question.		Successfully completed only one part of the question.	Failed to complete any parts.