

Problem Statement - Pima Indians Diabetes Analysis

Submission type	:	File Upload
Due Date	:	May 30, 9:00 AM
Total Marks	:	40
Available from	:	May 19, 10:30 AM
Your Marks	:	40/40

Description

Scoring guide (Rubric) - Pima Indian Diabetes Analysis Evaluated

Criteria	Ratings	Points
Q 1. Import the necessary libraries and briefly explain the use of each library	Areas performed well: Good job on importing the necessary libraries Explained the use of each library	3/3
Q 2. Read the given dataset	Areas performed well: Good job on reading the given dataset	1/1
Q3. Show the last 10 records of the dataset. How many columns are there?	Areas performed well: Good job on showing the last 10 records of the dataset. Good job on displaying columns	1/1
Q4. Show the first 10 records of the dataset	Areas performed well: Good job on showing the first 10 records of the dataset	1/1
Q5. What do you understand by the dimension of the dataset? Find the dimension of the `pima` dataframe.	Areas performed well: Good job on providing your understanding on dimension of the dataset Good job on finding the dimension of the pima data frame.	1/1

Criteria	Ratings	Points
Q6. What do you understand by the size of the dataset? Find the size of the `pima` dataframe.	Areas performed well: Good job on providing your understanding on size of the dataset Good job on finding the size of the pima dataframe.	1/1
Q7. What are the data types of all the variables in the data set?	Areas performed well: Good job displaying the data types of all the variables in the data set	2/2
Q8. What do you mean by missing values? Are there any missing values in the `pima` dataframe?	Areas performed well: Good job on finding out the missing values in the dataset	2/2
Q9. What does summary statistics of data represents? Find the summary statistics for all variables except 'Outcome' in the `pima` data? Take one column/variable from the output table and explain all the statistical measures.	Areas performed well: Good job on providing the thoughts on summary statistics Good job on finding the summary statistics for all variables except 'Outcome' in the pima data Explained the statistical measures for one of the columns from the output table.	3/3
Q 10. Plot the distribution plot for the variable 'BloodPressure'. Write detailed observations from the plot.	Areas performed well Good job on plotting the distribution plot for the variable 'BloodPressure'. Good job on writing the detailed observations from the plot.	2/2
Q 11. What is the 'BMI' for the person having the highest 'Glucose'?	Areas performed well: Good job on finding out the 'BMI' for the person having the highest 'Glucose'?	1/1

Criteria	Ratings	Points
Q 12. Q 12.1 What is the mean of the variable 'BMI'? 12.2 What is the median of the variable 'BMI'? 12.3 What is the mode of the variable 'BMI'? 12.4 Are the three measures of central tendency equal?	Areas performed well: Checked the mean of the variable 'BMI'? Checked median of the variable 'BMI'? Checked mode of the variable 'BMI'? Commented on whether these three measures of central tendency are equal or not.	3/3
Q 13. How many women's 'Glucose' level is above the mean level of 'Glucose'?	Areas performed well: Checked how many women's 'Glucose' level is above the mean level of 'Glucose'?	1/1
Q 14. How many entries (women) have their 'BloodPressure' equal to the median of 'BloodPressure' and their 'BMI' less than the median of 'BMI'?	Areas performed well: Checked How many women have their 'BloodPressure' equal to the median of 'BloodPressure' and their 'BMI' less than the median of 'BMI'?	2/2
Q 15. Below is the pairplot of variables 'Glucose', 'SkinThickness' and 'DiabetesPedigreeFunction'. Write your observations from the plot.	Areas performed well: Created the pairplot for variables 'Glucose', 'SkinThickness' and 'DiabetesPedigreeFunction'. Good job on providing the observations from the plot.	4/4
Q 16. Plot the scatterplot between 'Glucose' and 'Insulin'. Write your observations from the plot.	Areas performed well: Good job on plotting the scatterplot between 'Glucose' and 'Insulin'. Good job on providing the observations from the plot	2/2
Q 17. Plot the boxplot for the 'Age' variable. Are there outliers?	Areas performed well: Plotted the scatterplot between 'Glucose' and 'Insulin'. Good job on providing the observations from the plot.	2/2

Criteria	Ratings	Points
Q 18. Plot histograms for variable Age to understand the number of women in different Age groups given that they have diabetes or not. Explain both histograms and compare them.	Areas performed well: Plotted histograms for variable Age to understand the number of women in different Age groups given that they have diabetes or not. Explained both histograms and compared them.	3/3
Q 19. What is Inter Quartile Range of all the variables? Why is it used? Which plot visualizes the same?	Areas performed well: Explained what is inter Quartile Range of all the variables? Explained why is it used? Explained which plot visualizes the same?	2/2
Q 20. Find and visualize the the correlation matrix. Write your observations from the plot.	Areas performed well: Good job on finding and visualizing the correlation matrix. Good job writing your observations from the plot.	3/3
Points		40/40

Submitted Assignment



Notebook+Template+-+Pima+Indians+Diabetes+Analysis-1.html



Submitted at Mon, May 30

Comments:

[+ Add comments](#)

Previous

Next

Proprietary content. ©Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.