

Diabetes data has been provided

- 1.)Print out the columns in the data and find descriptive stats for the data
- 2.)Find out the percentage of missing values in the columns
- 3.)Plot a pie plot for our target variable to understand the distribution
- 4.)Plot KDE for all our numerical data points against our target variable
- 5.)Create a new column using 2 or more existing columns
- 6.)Create a correlation heatmap
- 7.)Create Boxen plots and violin plots for all our numerical columns on the same graph/y_axis(hint: use scaling)
- 8.)Plot CDF plots for all our numerical columns against both our target classes
- 9.)Grid search and build Knn,svm models with highest cv accuracy giving hyper parameters and find out the best params for each.

Please write a short description of the plots and an insight that you've gotten from these plots as well below each visualization

You can use seaborn/plotly as well