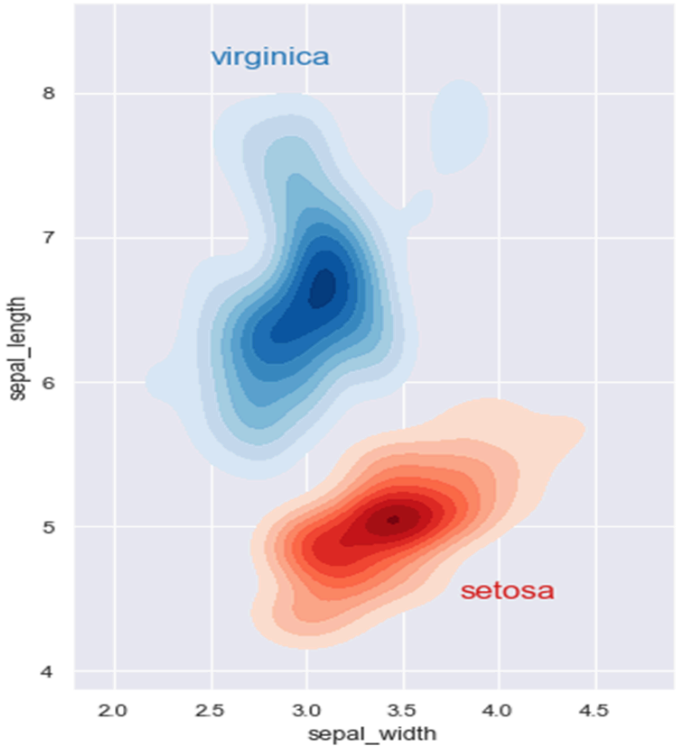
Computer Science and Engineering Department, SVNIT, Surat
M.Tech.- I DS (Semester - 1)
Foundations of Data Science
(CSDS111)

Lab Assignment: 6

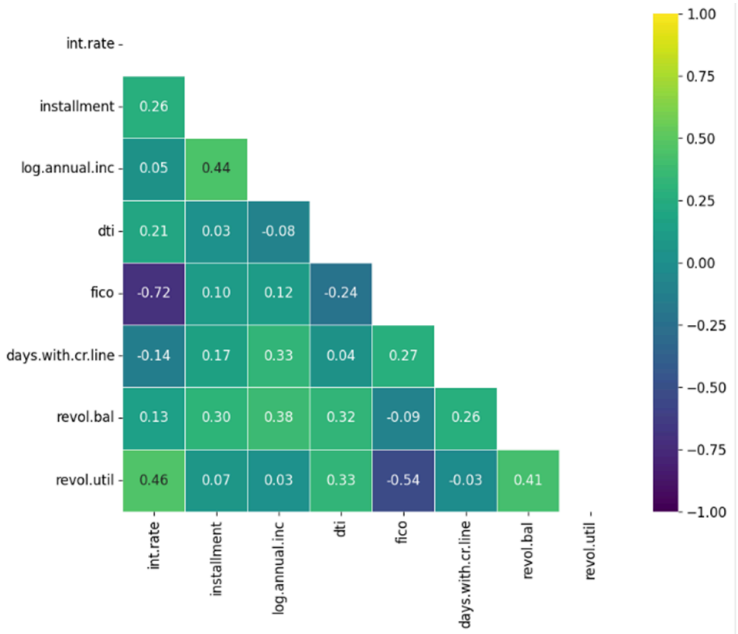
Data Processing:

1) Download any dataset with numeric values to plot the histogram as per the values given by the user:
    a) Read the x-axis values from the user which are the values where edges of the histogram's bars are lying.
    b) Read Number of bins to display
    c) Read the bar color and edge color to display
    d) Read the transparency value of bar color to display
    e) Read colors of each bar to display

    [Do some data wrangling to get the data in good shape to plot, if required.]
2) Using your random number generator program, generate age for 100 persons and display the age like barcode.
3) Prepare the contour graph for numeric data as shown in the following:



4) Prepare the heatmap as shown in the following:



5) Download the Earthquake data of India from Kaggle and for a time you are interested in (the data starts from 2018). Do some data wrangling to get the data in good shape to plot, if required.
    a) Use the Colour-code of the points based on the depth of the earthquake.
    b) Use the size of the point/bubble to represent the magnitude of the earthquake.
    c) Label the visualization for easy understanding.
    d) Report the result analysis of your visualization.
6) Use the diabetes data set from UCI and Pima Indians Diabetes data set for performing the following:
    a. Using built in functions explore both the datasets.
    b. Perform Bivariate analysis: Linear and logistic regression modeling
    c. Perform Multiple Regression analysis
    d. Compare the results of the above analysis for both the datasets.

- Data Repositories
    UCI Repository
    Kaggle Dataset
        Earthquake Data India
    NASDAQ Dataset
    Google Public Dataset