EEO 388
Foundations of Machine Learning
Fall 2024
Faculty: Vibha Mane

Exercise Set 1b
Data Exploration

1   Task

1. Perform following data visualization with the **Raisin** data set.

   - Scatter Plots
   - Histograms

   Note that you want to use the target variable **Class** as hue for the function `sns.pairplot()`.
   What do you infer? For example, is the data numerical or categorical? How many samples
   in each class? Are the classes separated in any of the features? Are there any outliers?

2. Repeat the above data visualization task with the **Deep Space** data set. Note that this
   data set has no class labels. What do you infer?

2   Exercise Set Submission

The Exercise Set should be submitted in the provided link on Brightspace. It should include the
following:

1. The Python source code, with one file for each task.

2. A brief report describing what you did, including some results and plots.