1. Consider dataset III and do the following steps:

   a. Read the file; how many classes exist in the dataset? Visualize data distribution in each class.
   b. Visualize necessary features and describe the dataset.
   c. Use the KNN method in order to fit the appropriate model. For validation, use the F1 score metric. Choose one k between 30-35, and do not change it during part c. For training the model: 1- Use all the possible features. 2- Use three arbitrary features. 3- Use one arbitrary feature. Which of these three options has the better result? Why? Write your conclusion in detail.
   d. Choose one of the part c options and find the best k. Use an arbitrary metric to find the best result.
   e. What will happen if the value of K is chosen very high?

   f. Use the Logistic Regression method and fit a model and report the results. (Use all the possible features for this part)
   g. Compare Logistic Regression and KNN method and write your own opinion.
   h. (Bonus point) Use Support Vector Machine and random forest methods using all possible features to fit a model.