For each one of the following questions, write Python code in PyCharm.

- For each question, create a *new* Python file. Name each *lastname_firstname_hw#_q#.py* etc.
- Create a header in each file using *comments* to display your name and HW information. After that write your Python code.
  # Tommy Trojan
  # ITP 449 Fall 2020
  # HW8
  # Q1
- Create a Word document with screenshots of your output. Zip the python file and the Word document together. Submit on Blackboard.

Problem #1

Create a KNN model for diabetes prediction. The dataset is posted on Blackboard. Explore the factors in the dataset.

1. Create a DataFrame "diabetes_knn" to store the diabetes data and set option to display all columns without any restrictions on the number of columns displayed.
2. Determine the dimensions of the "diabetes_knn" dataframe.
3. Update the DataFrame to account for missing values.
4. Create the Feature Matrix and Target Vector.
5. *Standardize* the attributes of Feature Matrix (https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html)
6. Split the Feature Matrix and Target Vector into training and testing sets, reserving 30% of the data for testing. *random_state = 2020, stratify = y*
7. Develop a KNN based model and obtain KNN score (accuracy) for train and test data for k's values ranging between 1 to 15.
8. Plot a graph of train and test score and determine the best value of k.
9. Display the test score of the model with best value of k and print the confusion matrix for it.
10. Predict the Outcome for a person with *2 pregnancies, 150 glucose, 85 blood pressure, 22 skin thickness, 200 insulin, 30 BMI, 0.3 diabetes pedigree, 55 age.*