HW #2
Due: December 13, 2019

Student ID: _____ Name: _____

A survey from 100 patients was conducted to collect their views (*no stomach upset, no side effect, stop pain, work quickly, keep me awake, limited relief*) for several pain relief medications. The `pain-relief.csv` is the resulting dataset.

1. Cluster the data using Ward's method and decide the number of clusters (*k*) by observing the resulting dendrogram.
2. Construct 6 panels of boxplots where each panel associated with each original variable consists of *k* boxplots. Hence, a boxplot represents the distribution of an original variable within a cluster.
3. Interpret the clustering results using your answer to Question #2.
4. Perform *k*-means clustering using the number of clusters (*k*) you decided in Question #1.
5. How many patients are consistently clustered together based on the results from Ward's method and *k*-means clustering?
6. Perform PCA based on the given dataset. Report the eigenvalues, eigenvectors, and PC scores of first 10 observations.
7. How many PCs would you like to use? Explain why you chose that number.
8. Visualize the data using two PCs and mark the Ward's clustering result on it.
9. Interpret the Ward's clustering result using your PCA result and the plot from Question #8.
10. Compare your answers to Question #9 and #3.

✓ Your homework must be turned in as a hard copy by the due. Do not upload your homework on the icampus. Please make sure that NO late homework is accepted.