CS/CE 457/464 - Homework Assignment 8: Clustering

Due Date: Monday, November 4 at 11:59 pm

Purpose:
Demonstrate understating of Clustering technique for grouping and segmentation

Points: 100

Deliverables: Submit ipynb code file along with your answer

- Review IDMA Book Chapter 10 Clustering
- Use the dataset `baskin_robbins_icecream.csv`
- Perform clustering analysis on the following questions**.** Make sure to include interpretation of each result including visualizations to support your answer

1. Need to clean the data. Remove columns **Flavour** and **Category** for clustering purpose.
2. Scale (normalize) the data before clustering.
3. Generate WSS (Knee Plot) to find optimal number of clusters. You can also plot Silhouette Coefficient plot to support your answer.
4. Perform K-Means clustering again with optimal number of clusters.
5. Visualize the clustering results.
6. Use cluster labels to create cluster specific data frames from original dataset.
7. Generate at least 3 insights from your clustering results to show unique and interesting features for each cluster. Discuss your findings.
8. Can you suggest a unique name for each cluster? For example if a cluster is showing high Sugar level, then you can call it a "SugarLovers" cluster.