# CS 457 - Homework Assignment 10: Clustering

## Due Date: Monday, November 13 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 `McDonaldsMenu.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 Category and Item for clustering purpose.
2. You may drop column Serving Size. If not, then you need to clean the values and get oz or g values and then remove the units.
3. Scale (normalize) the data before clustering.
4. Generate WSS (Knee Plot) to find optimal number of clusters. You can also plot Silhouette Coefficient plot to support your answer.
5. Perform K-Means clustering again with optimal number of clusters.
6. Visualize the clustering results.
7. Use cluster labels to your original dataset to create cluster specific data frames.
8. Generate at least 3 insights from your clustering results to show unique and interesting features for each cluster. Discuss your findings.
9. 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.