# CDS503 Lab 7

**Lab Exercise**

Form a small group with a maximum of 2 people. Say you own a mall and want to understand the spending behaviors of different customer segments so you can provide the marketing team with insights on how to plan the marketing strategy. You will use clustering to obtain spending insights of different customer segments.

Load the customers dataset (customers.csv). The dataset contains 5 attributes:

- CustomerID: Customer ID
- Gender: Customer gender (male or female)
- Age: Customer age
- AnnualIncome: Annual income (thousand dollars)
- SpendingScore: A score assigned to the customer based on customer behavior and purchasing data (ranges from 1 – 100)

1) Pick two attributes. You must include SpendingScore as one of the attributes. You can choose either Gender, Age or AnnualIncome as the second attribute. For example, if you choose Gender, you will uncover clusters that represent customer segments based on Gender and SpendingScore.
2) Run k-means clustering based on the two selected attributes. Experiment with different number of clusters. Report the number of clusters that make the most sense to help you interpret customer segments. Describe each cluster based on the data points and centroids in the cluster (can try visualizing the clusters).
3) Run agglomerative clustering (hierarchical clustering) based on the two selected attributes. You can try to change different parameters such as the affinity and linkage parameters to obtain the number of clusters that make the most sense to help you interpret customer segments. Use the dendrogram to help you determine if you are obtaining good or bad clusters. Report the final parameters you set and the number of clusters that best represents customer segments. Describe each cluster based on the data points in the cluster (can try visualizing the clusters).
4) Compare the clusters from (2) and (3). Which clustering technique (k-means clustering or hierarchical clustering) generates more informative insights on spending behaviors of different customer segments at the mall? Justify your choice.