

# Assignment 4 (50 Points)

**Due Date: 11/14/2024**

## Marketing to Frequent Fliers.

The file *EastWestAirlinesCluster.csv* contains information on 3999 passengers who belong to an airline's frequent flier program. For each passenger, the data include information on their mileage history and on different ways they accrued or spent miles in the last year. The goal is to try to identify clusters of passengers that have similar characteristics for the purpose of targeting different segments for different types of mileage offers.

- a.** Apply hierarchical clustering with Euclidean distance and Ward's method. Make sure to normalize the data first. How many clusters appear?
- b.** What would happen if the data were not normalized?
- c.** Compare the cluster centroid to characterize the different clusters, and try to give each cluster a label.
- d.** To check the stability of the clusters, remove a random 5% of the data (by taking a random sample of 95% of the records), and repeat the analysis. Does the same picture emerge?
- e.** Use *k*-means clustering with the number of clusters that you found above. Does the same picture emerge?
- f.** Which clusters would you target for offers, and what types of offers would you target to customers in that cluster?