

# CMPE 302 – K-Means Clustering Python Questions

Anıl Demirel

## Question 1 – Implementing K-Means from Scratch

Given a 2D dataset:

$$X = \{(1, 2), (1.5, 1.8), (5, 8), (8, 8), (1, 0.6), (9, 11)\}$$

- (a) Write a Python function that initializes two random centroids and performs **one iteration** of K-Means clustering (assignment and centroid update steps only).
- (b) Print the new centroid coordinates after one iteration.

## Question 2 – Elbow Method Interpretation

The following Python code performs clustering on a set of 2D points:

```
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

x = [4, 5, 10, 4, 3, 11, 14, 6, 10, 12]
y = [21, 19, 24, 17, 16, 25, 24, 22, 21, 21]
data = list(zip(x, y))

inertias = []
for i in range(1, 11):
    kmeans = KMeans(n_clusters=i)
    kmeans.fit(data)
    inertias.append(kmeans.inertia_)

plt.plot(range(1,11), inertias, marker='o')
plt.title('Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('Inertia')
plt.show()
```

- (a) What does the `inertia_` value represent in K-Means clustering?
- (b) Describe how the Elbow Method helps determine the optimal number of clusters  $K$ .