SENTONGO HAMZA

MASTER’S IN DATA SCIENCE

QUESTION 1

First, in order to have a good K-means classifier we need to have time series data in a 2d format(Shape: **(n\_samples, n\_timestamps)**).Instead of random x and y data points. We don’t include the z because we are dealing with unsupervised learning.

Precautions

-same length of time series

-Data should be normalized or standardized to avoid scale bias.

-We should use feature extraction (mean, variance, frequency components) should be applied before clustering.

We have already practiced the above in the previous exercises

QUESTION 2