Labwork 3:   
- Select two datasets (recommending datasets with labels)  
1. Kmeans  
- Apply K-Means clustering on both datasets.  
- Evaluate clustering performance  
- Vary the number of clusters k  
- Comment on how performance changes according to the k values  
  
2. K-means on subspace  
- Apply PCA/SVD/T-SNE/or UMAP to reduce dimensionality:  
- Run K-Means and AHC on the reduced data  
- Compare performance with original data  
- Visualize clusters in reduced space  
  
3. Repeat above questions with one of the clustering for the two selected datasets: DBSCAN, SOM, or

Bài tập thực hành 3:

- Chọn hai tập dữ liệu (khuyến nghị tập dữ liệu có nhãn)

1. Kmeans

- Áp dụng phân cụm K-Means trên cả hai tập dữ liệu.

- Đánh giá hiệu suất phân cụm

- Thay đổi số cụm k

- Nhận xét về cách hiệu suất thay đổi theo giá trị k

2. K-Means trên không gian con

- Áp dụng PCA/SVD/T-SNE/hoặc UMAP để giảm chiều:

- Chạy K-Means và AHC trên dữ liệu đã thu gọn

- So sánh hiệu suất với dữ liệu gốc

- Hình dung các cụm trong không gian thu gọn

3. Lặp lại các câu hỏi trên với một trong các phân cụm cho hai tập dữ liệu đã chọn: DBSCAN, SOM hoặc