

## Activité 5 : catégorisation K-means

---

### Travail à Faire

On désire réaliser une catégorisation des individus de la base "pima-indians-diabetes.csv" par l'algorithme k-means.

#### 1- Préparation des données

Pour pouvoir afficher les individus dans un plan (2D), ainsi que leurs classes d'appartenance on fera les restrictions suivantes :

- Pour la description des individus, on se restreindra à deux attributs.
- Pour le nombre des individus, on se restreindra à 100 individus.

a- Construire le nouvel ensemble de données (100 individus et 2 attributs).

b- Effectuer la normalisation des données afin que valeurs des deux attributs soient dans un intervalle [0-1].

c- Afficher dans un graphe les individus tel que le premier attribut représente les abscisses et le deuxième attribut représente les ordonnées.

2- Catégorisation : Effectuer la catégorisation par la méthode des k-moyennes puis afficher le résultat de la catégorisation pour des valeurs de  $k=2, 3$  et  $4$ .

3- Conclure.

