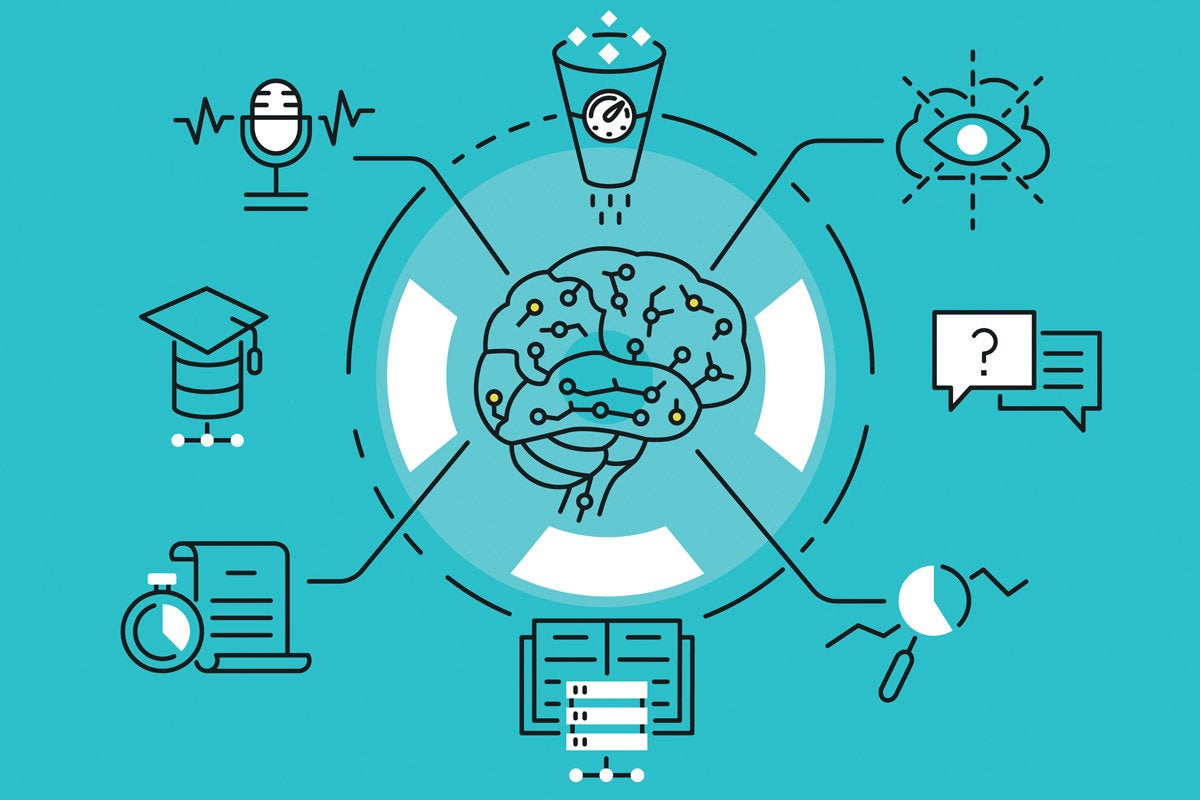
**Université Abdelmalek Essaâdi**

**Faculté des Sciences et Techniques-Tanger**

**Département Génie Informatique**

[**Machine Learning (Master SIBD)**](https://classroom.google.com/c/Njc0NzE2MzQxNjA5)

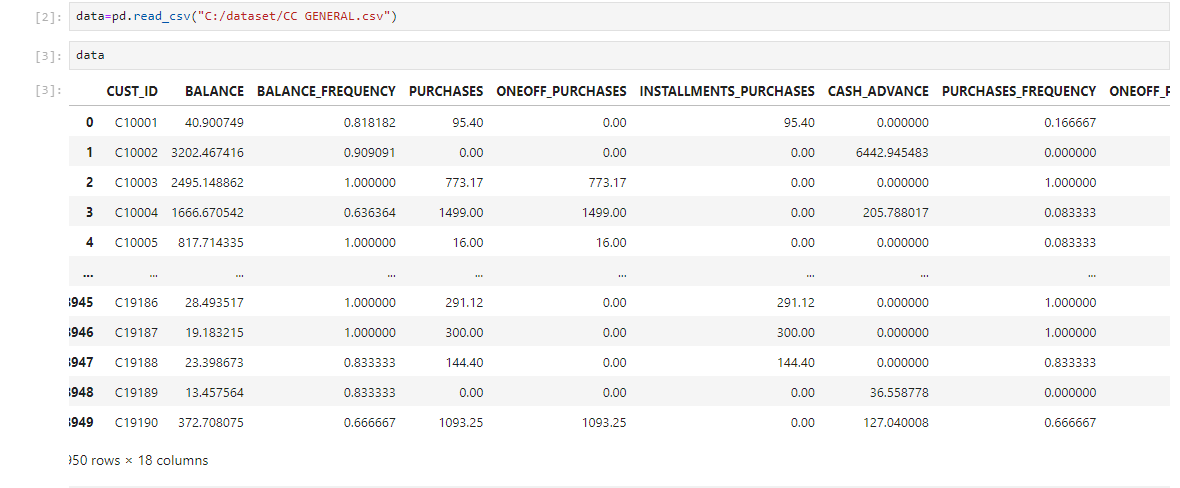
**Atelier 3**

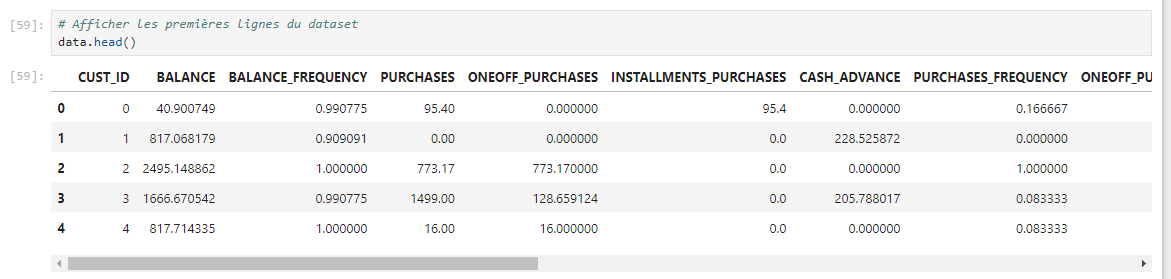
Réalisé Par : Encadre Par :

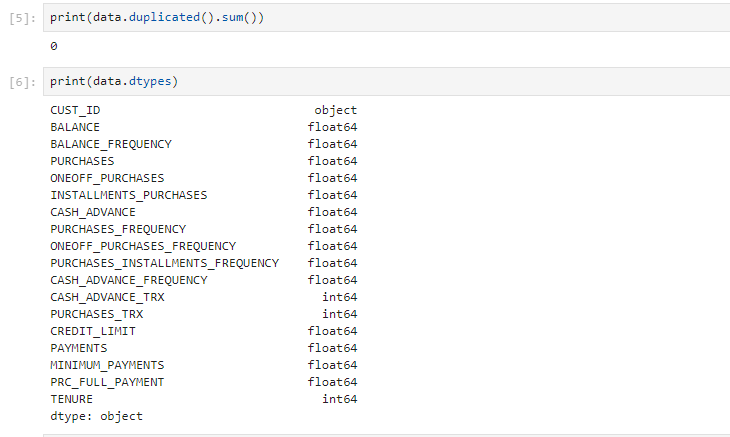
Yossra safi chetouan Pr . EL AACHAk LOTFI

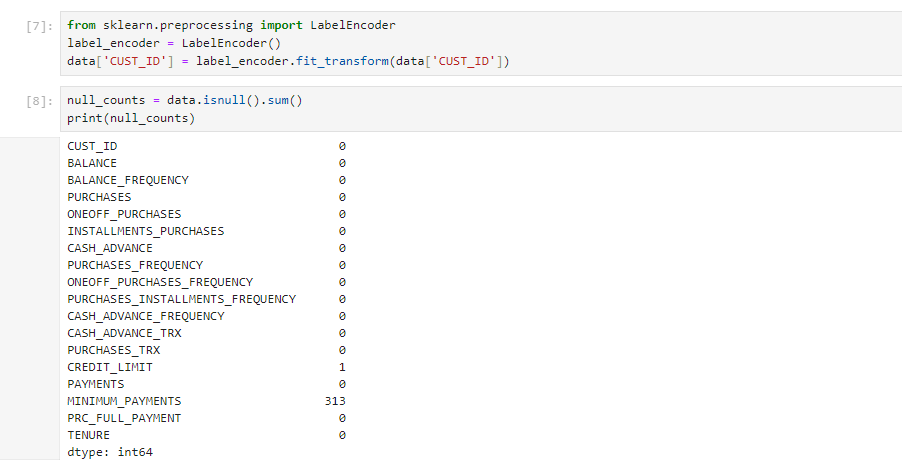
**Partie 1 (Data Visualisation):**

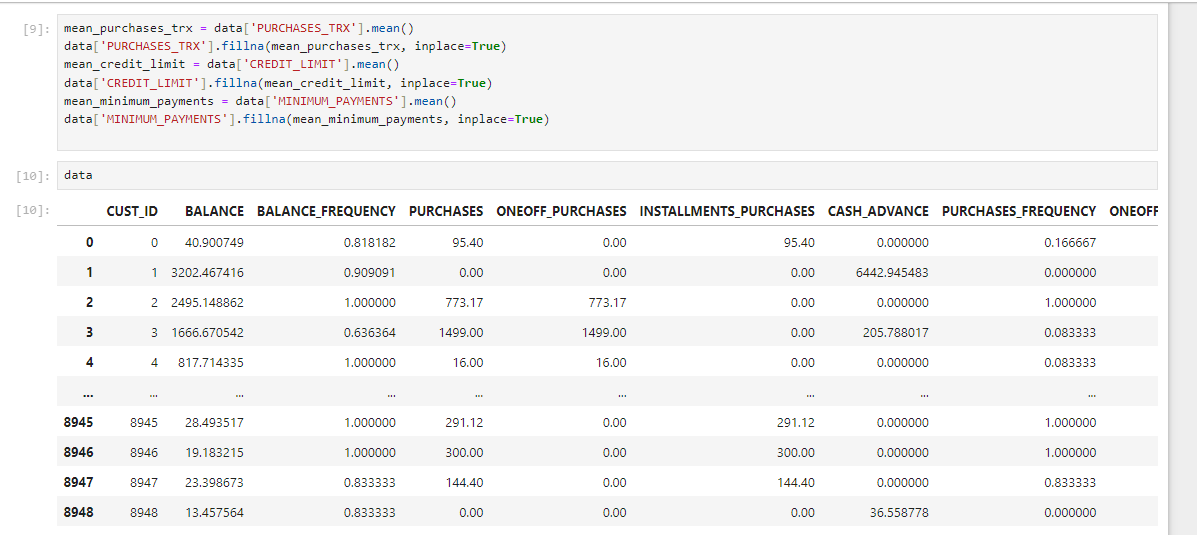
1. En utilisant pandas essayer d’explorer les données du Data set.



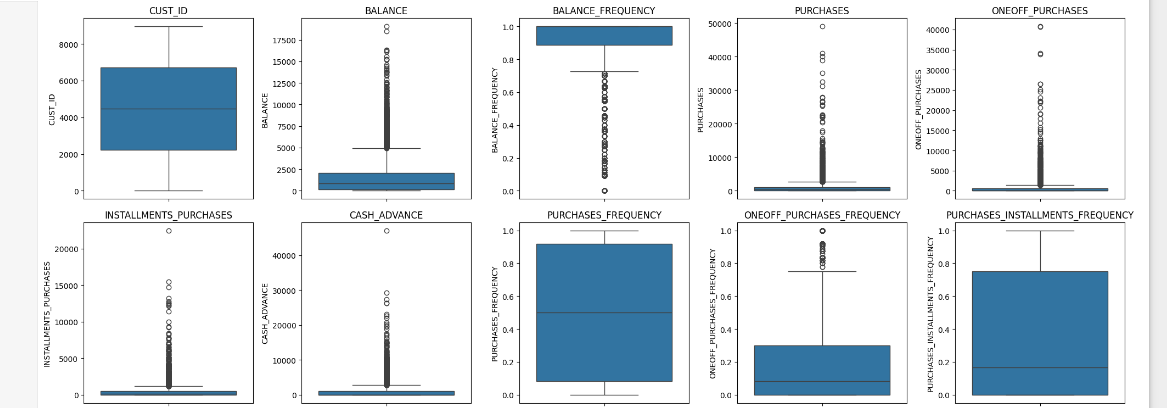


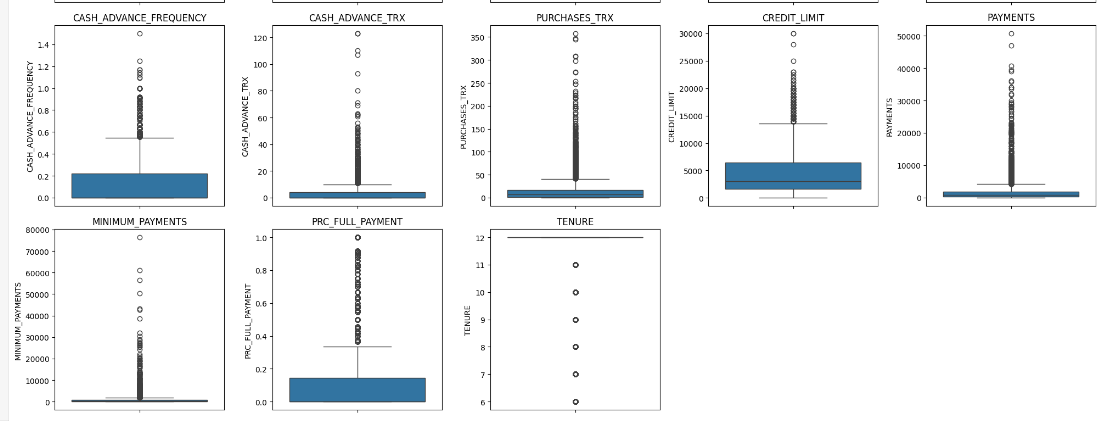




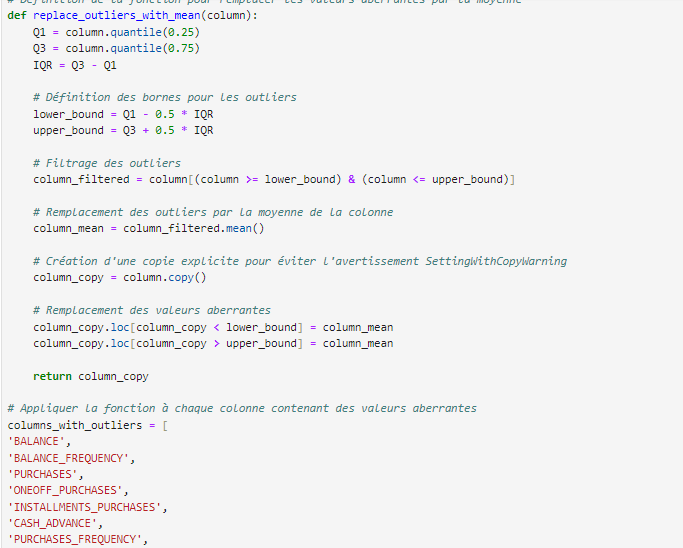


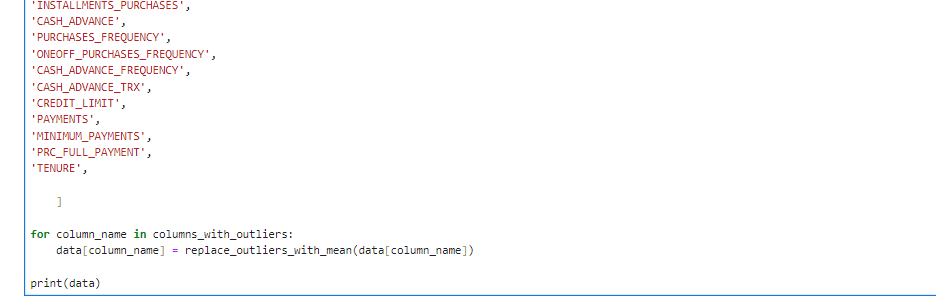
Afficher les valeurs aberrantes



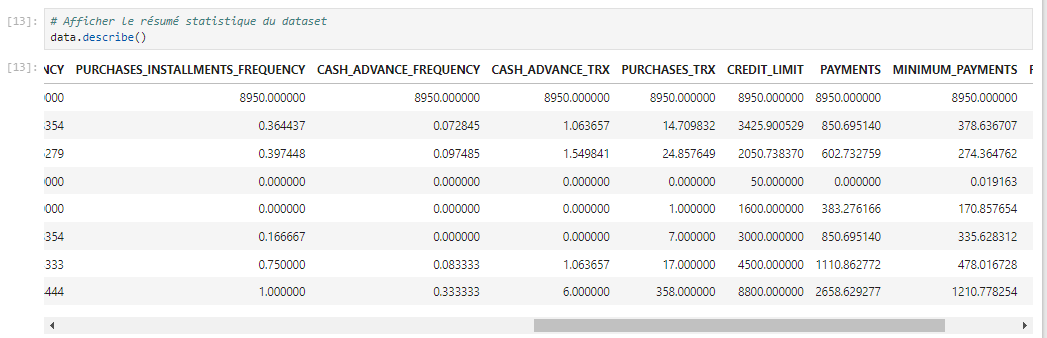


Traitement les valeurs aberrantes

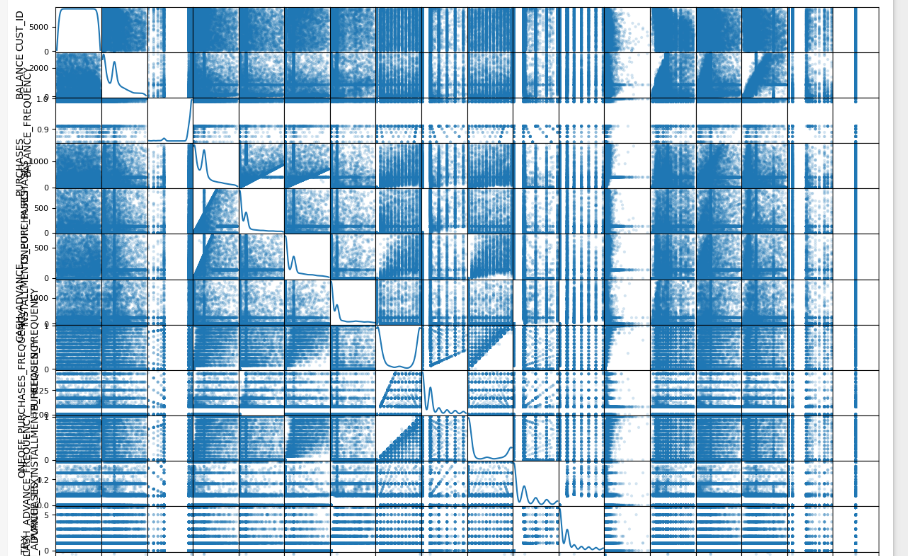




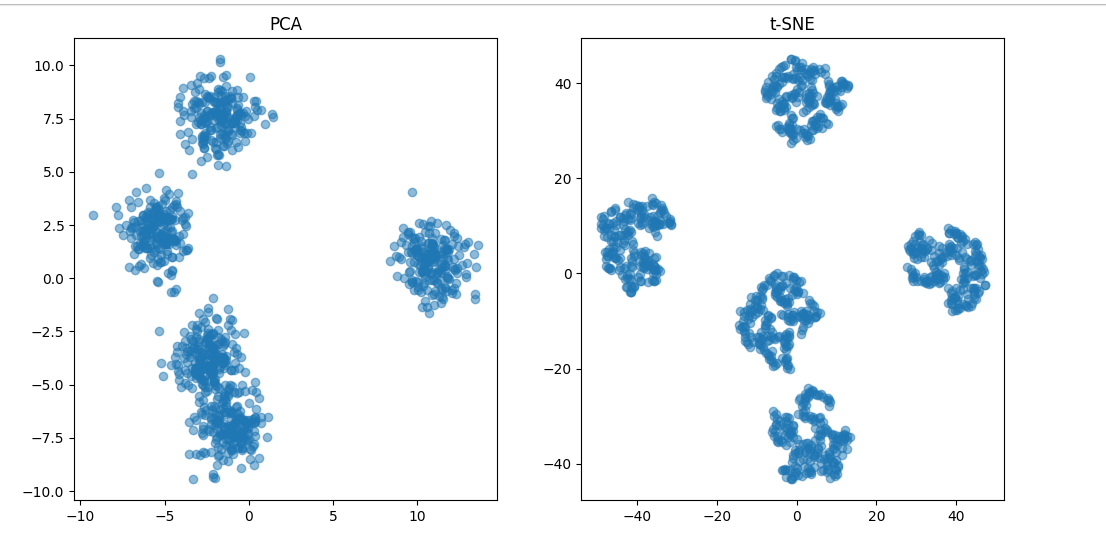
1. Afficher le résumer statistique du Data



1. Afficher les nuages des points du data set selon les propriétés « Features » en utilisant matplotlib et pandas « scatter\_matrix ».



1. Appliquer les deux techniques PCA et Tsne sur les features du Dataset, que ce que vous constatez.

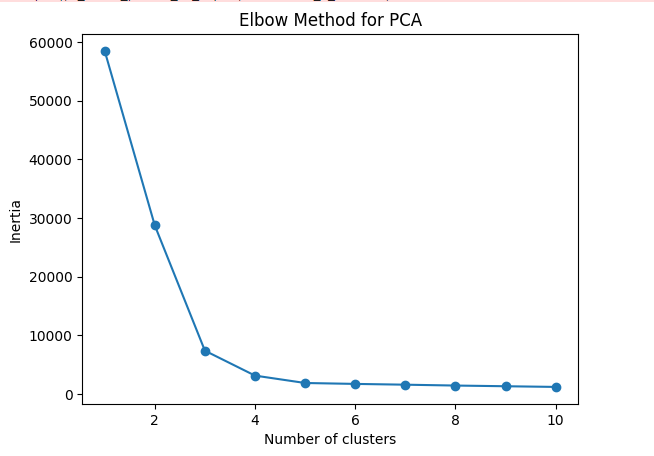


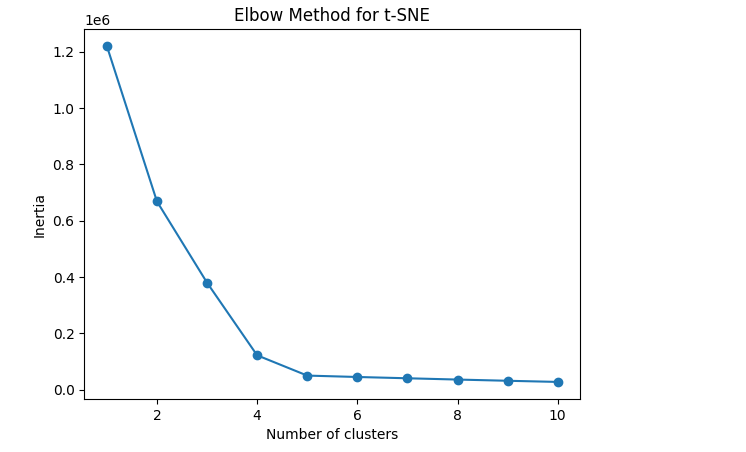
**Partie 2 (Clustering ):**

1. Essayer de construire les modèles de clustering en utilisant Kmeans (avec les nouvelles

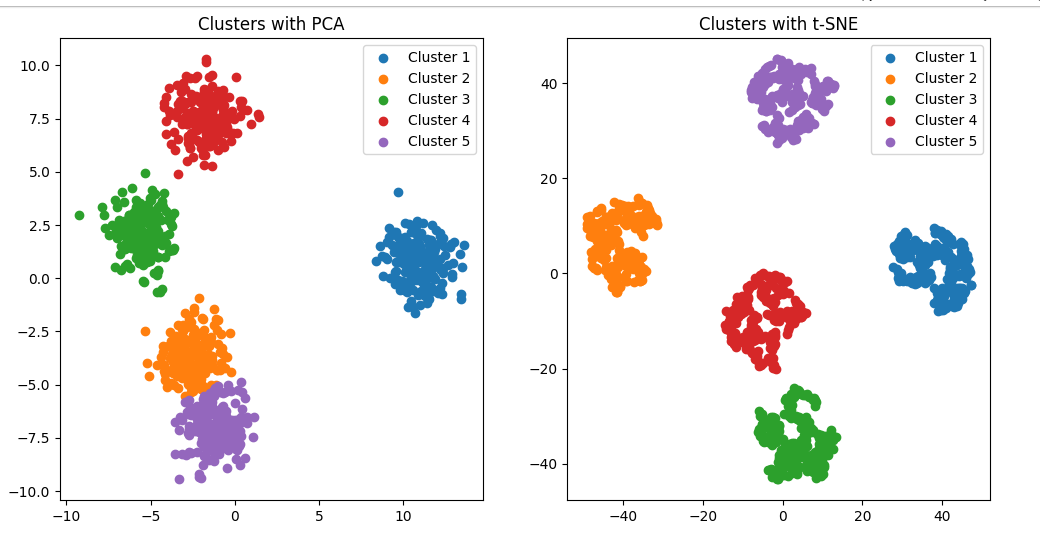
features (Un modèle basè sur PCA et l’autre sur Tsne) Question 4 de la partie 1.

2. Définir le K nécessaire pour les deux modèles en utilisant la méthode d’Elbow.





3. Présenter les clusters obtenues dans un graphe en utilisant matplotlib.



1. Refaire la même chose en utilisant l’algorithme fuzzy cmeans « il faut utiliser la bibliothèque skfuzzy» DBSCAN , EM ,Hierarchical clustering .

