

Analyse de données centre commercial

...

Découverte de clusters parmi vos clients

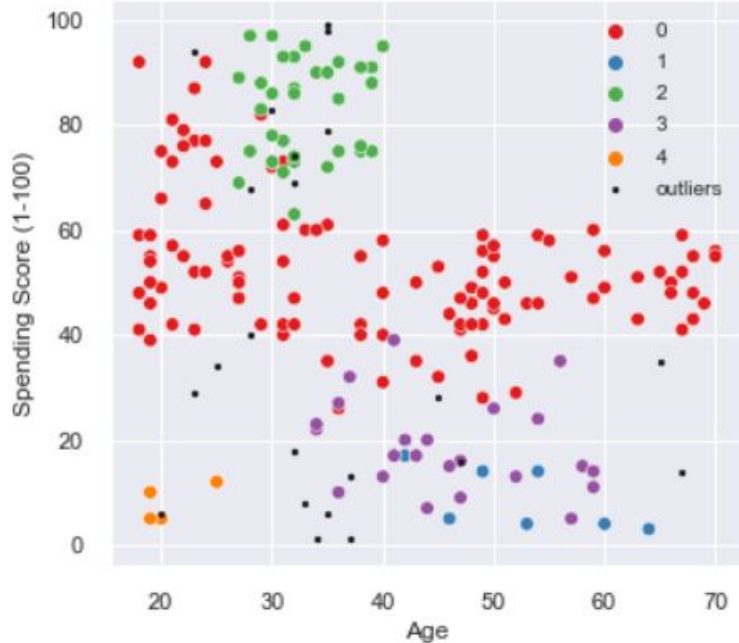
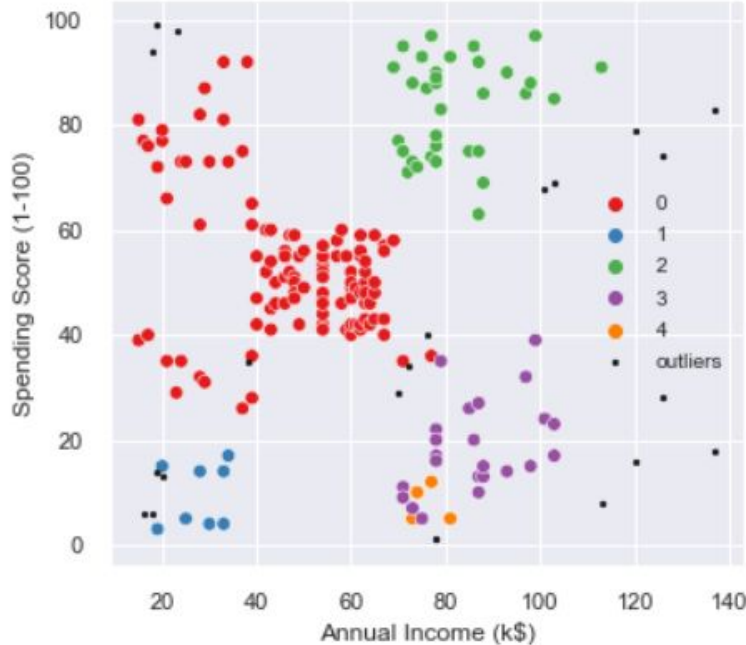
Analyse des données

Dataset de consommateurs:

- 5 features :
 - Identifiants
 - Âge
 - Revenu Annuelle
 - Fréquence d'Achat
 - Sex
- 200 personnes 112 Femmes et 88 Hommes

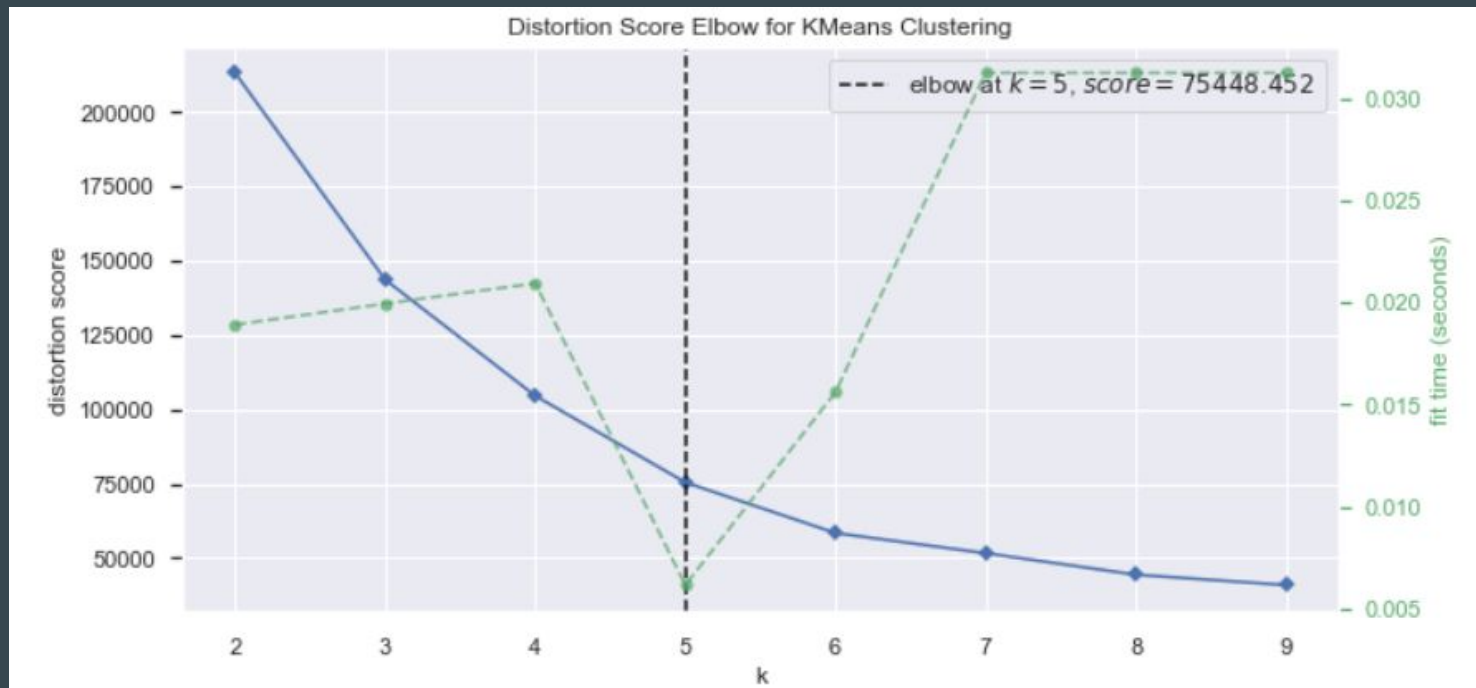
DBscan

Technique non retenue pour le clustering car séparation des différents groupes pas assez marquée.
Mais permet de détecter les clients atypiques de votre magasin.

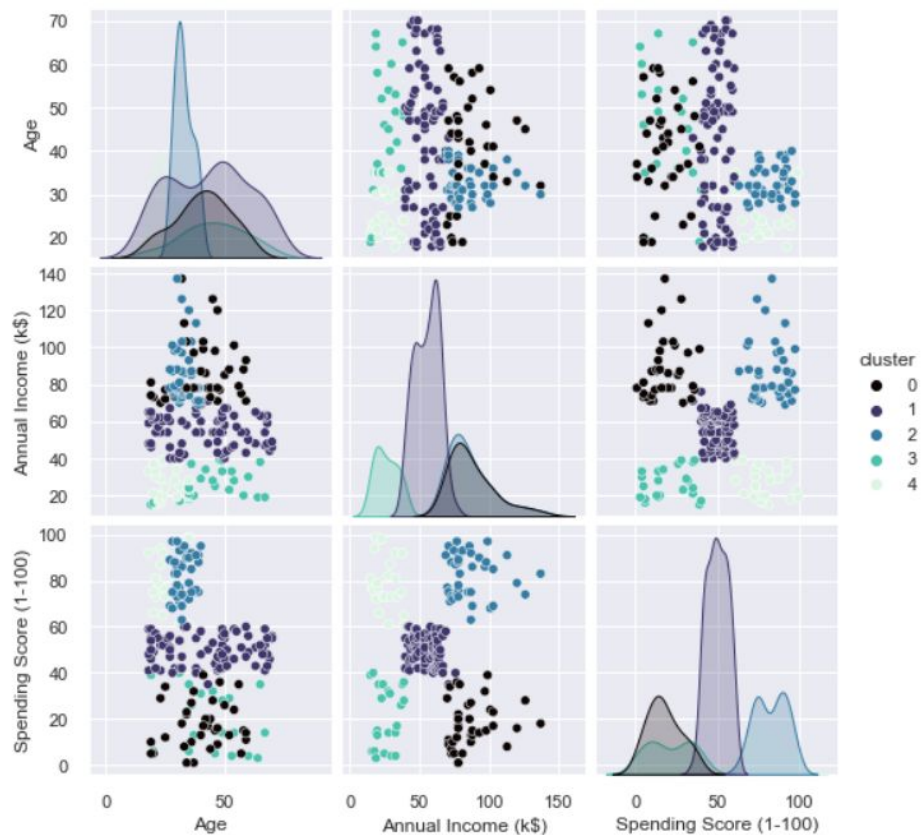


Trouver le nombre de clusters

5 clusters trouvés d'après la méthode du coude

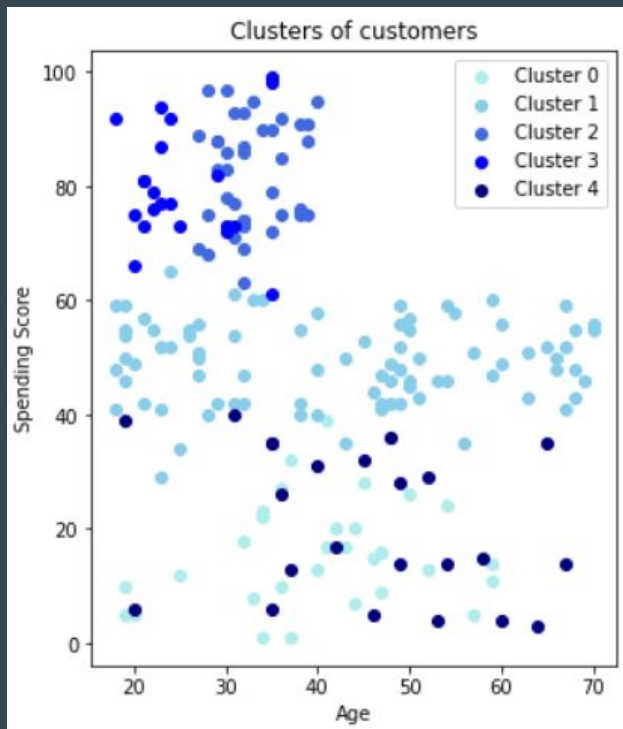
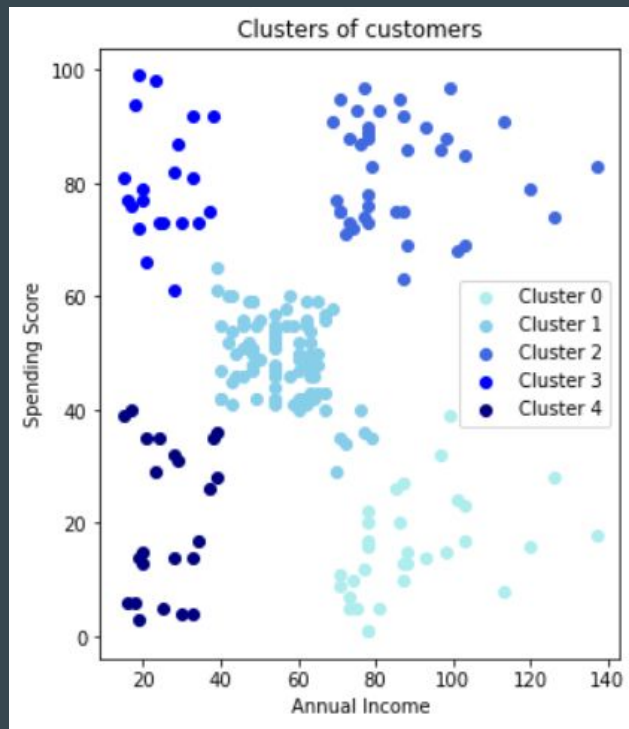


Méthode du K-means



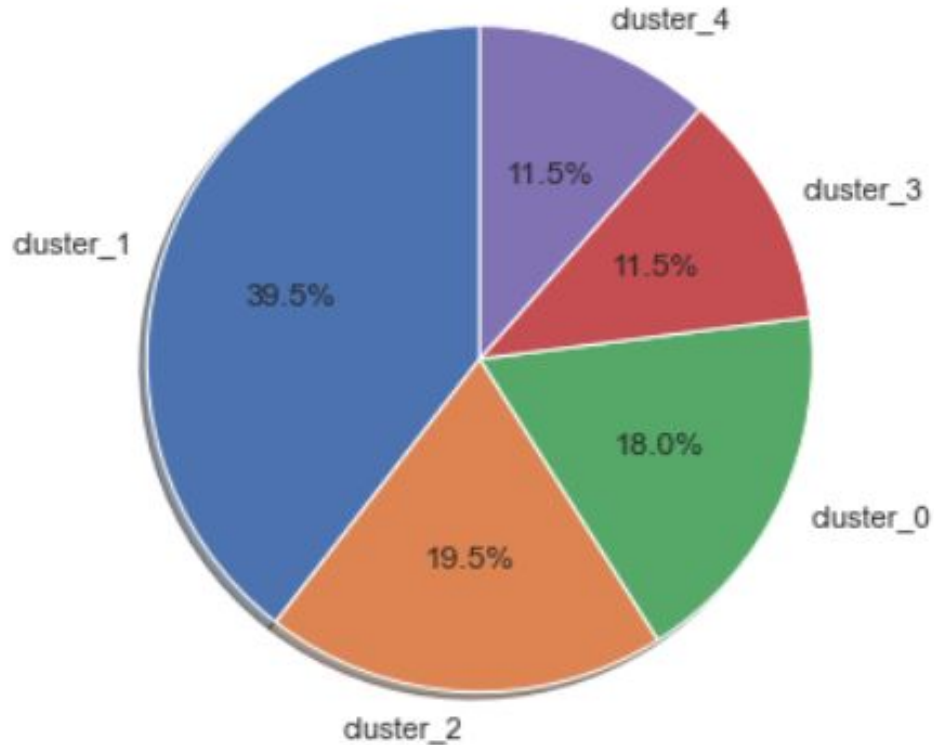
- Distribution des différents groupes de clients de votre magasin
- Découverte des éléments constitutifs de ces différents groupes (âge, salaire)
- score 0.44

Méthode Agglomerative clustering

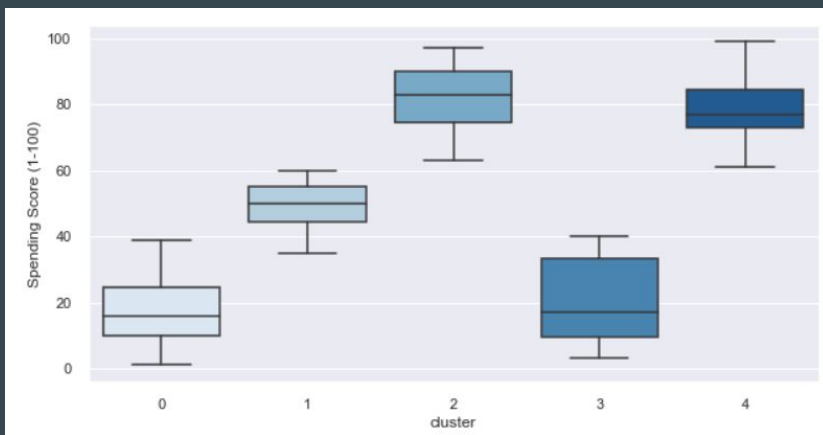
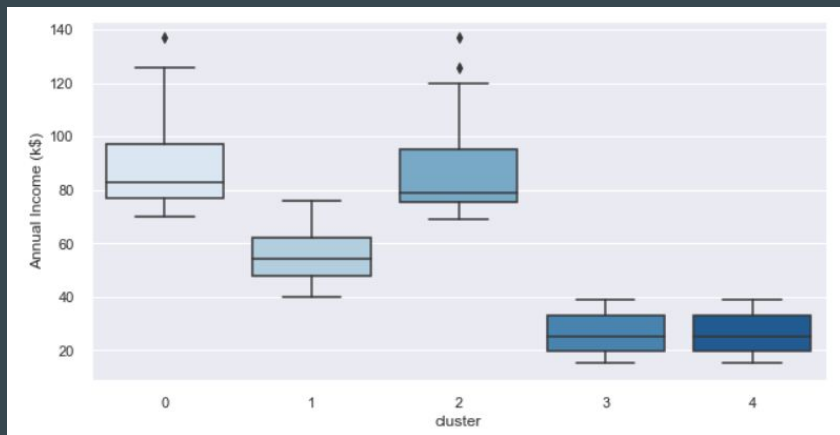
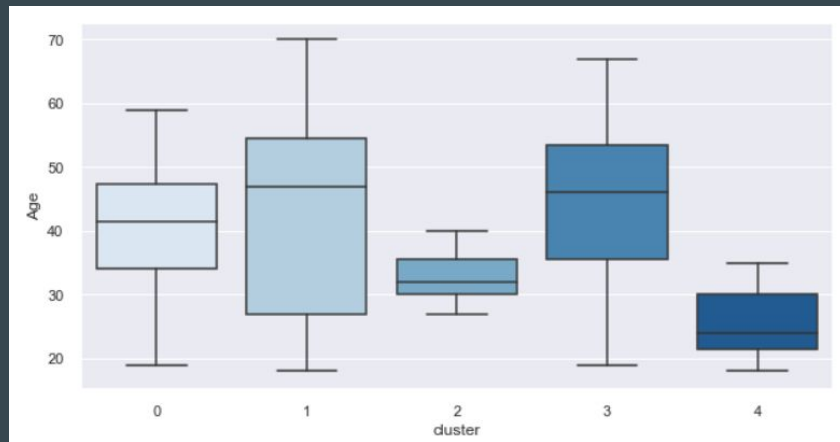


- Très proche du K-means
- Clusters différents bien séparés
- Confirme le choix du nombre de clusters

Répartitions des Cluster



Répartition des clusters par Âge/Salaire/Dépenses



Typologie des clusters

- **Cluster 1** : Environ 40 ans / Revenu annuel élevé / Dépenses basses
- **Cluster 2** : Environ 50 ans / Revenu annuel moyen / Dépenses moyennes
- **Cluster 3** : Environ 30 ans / Revenu annuel élevé / Dépenses très hautes.
- **Cluster 4** : Entre 40 et 50 ans / Revenu annuel très faible / Dépenses très basses
- **Cluster 5** : Environ 20 ans / Revenu annuel très faible / Dépenses très élevées

Gaussien cluster

1. j'ai standardisé mes données, et j'ai déjà encodé la colonne sex(0,1)
2. j'ai appliqué la gaussienne

`gm = GaussianMixture(n_components = 2, n_init = 10, max_iter = 100).`

3. j'ai eu des résultats

```
Entrée [28]: # Clusters Labels  
gm.predict(X)
```

```
Out[28]: array([1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1,  
                0, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0,  
                0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1,  
                0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 0, 0,  
                0, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1,  
                1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1,  
                0, 0, 1, 0, 0, 1, 1, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0,  
                0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0,  
                1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1,  
                1, 1], dtype=int64)
```

