



PROJET 5: SEGMENTEZ DES CLIENTS D'UN SITE E-COMMERCE

Soutenu par:
Kourouma Sekouba Aissatou

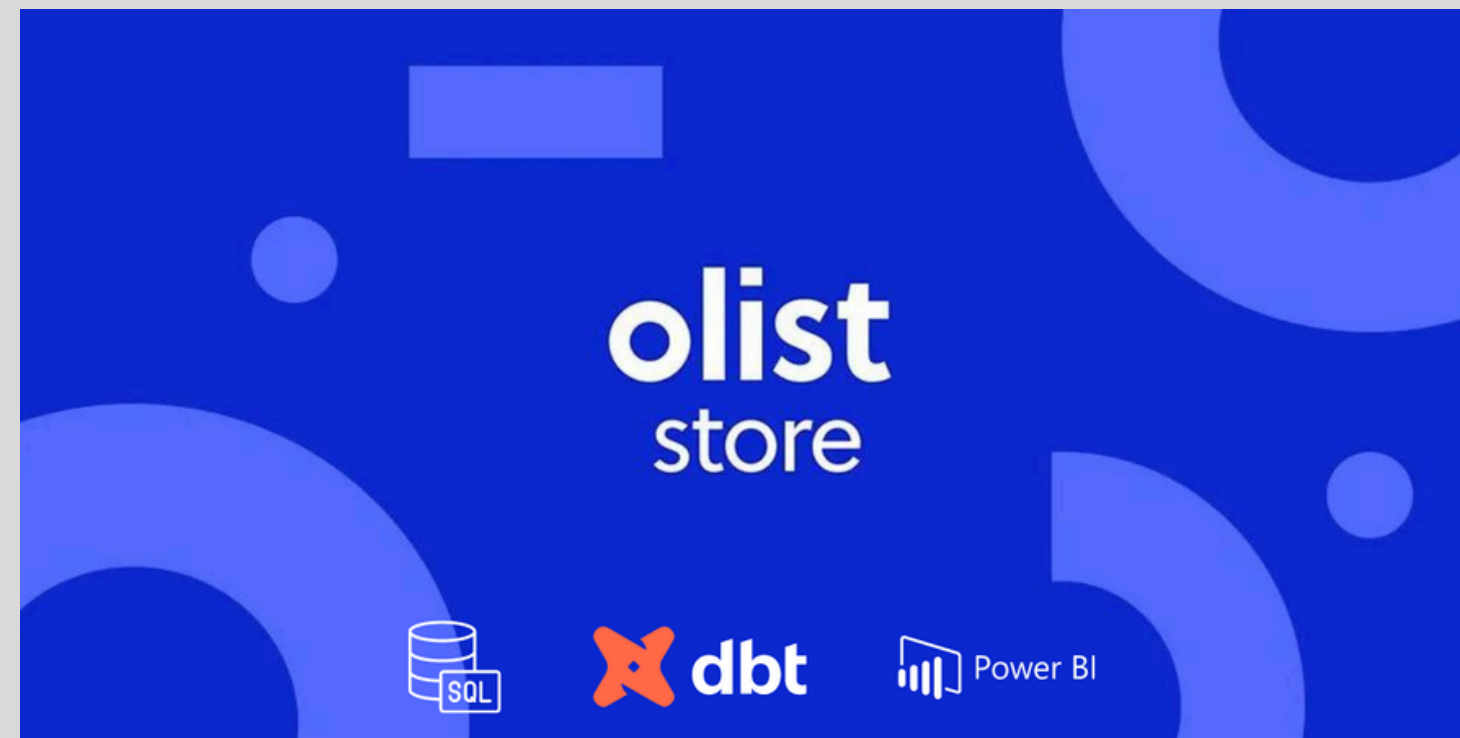


SOMMAIRE

1. MISSION
2. PRESENTATION DU JEU DE DONNÉE
3. ANALYSE EXPLORATOIRE
4. CREATION DU FICHIER CLIENT
5. ACP
6. ALGORITHME K-MEANS
7. CLUSTERING HIÉRARCHIQUE
8. ALGORITHME DBSCAN
9. TESTER UN CLUSTERRING K-MEANS AVEC LES VARIABLE QUALITATIVES
10. ETUDE DE LA STABILITÉ TEMPORAIRE
11. CONCLUSION

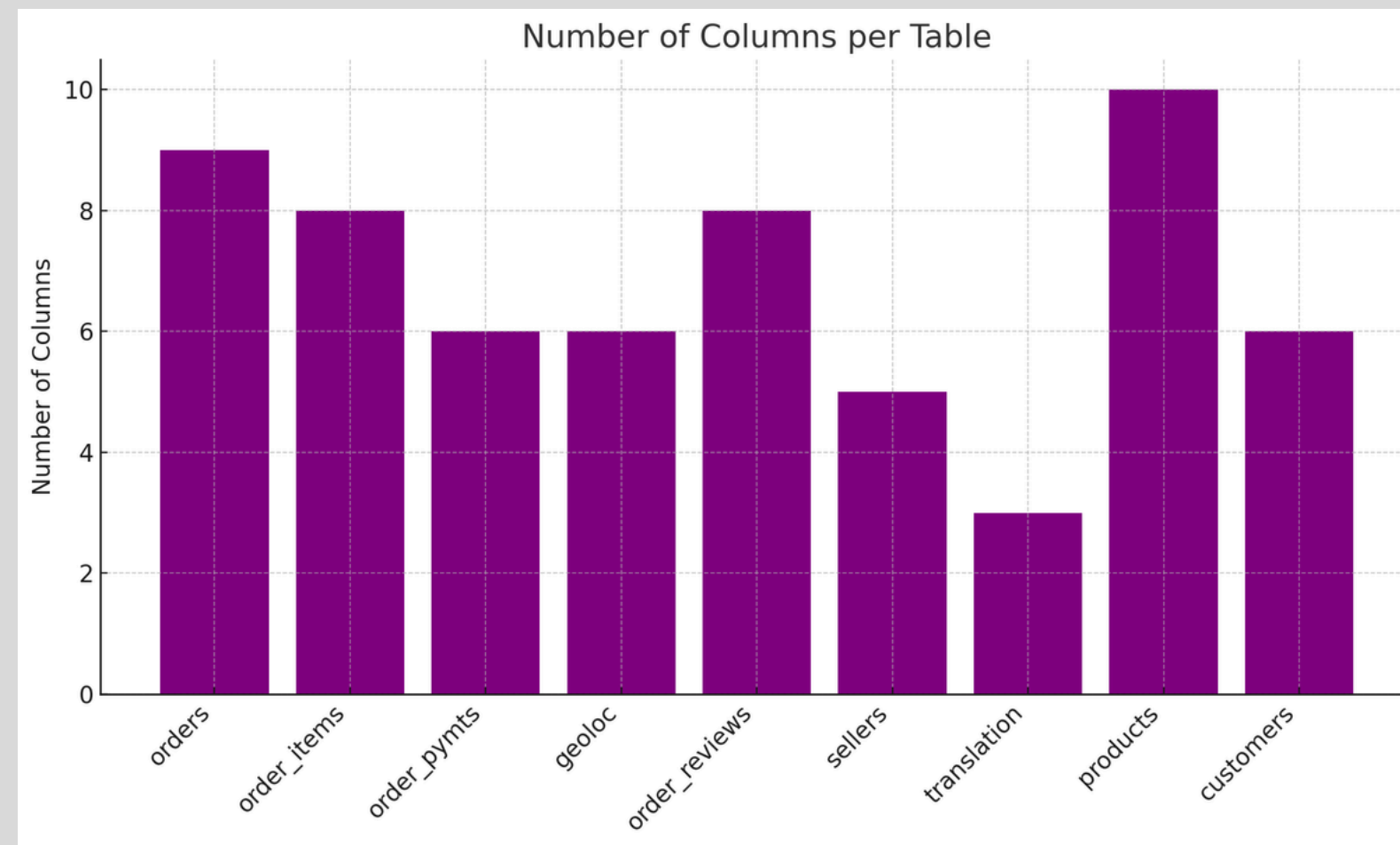
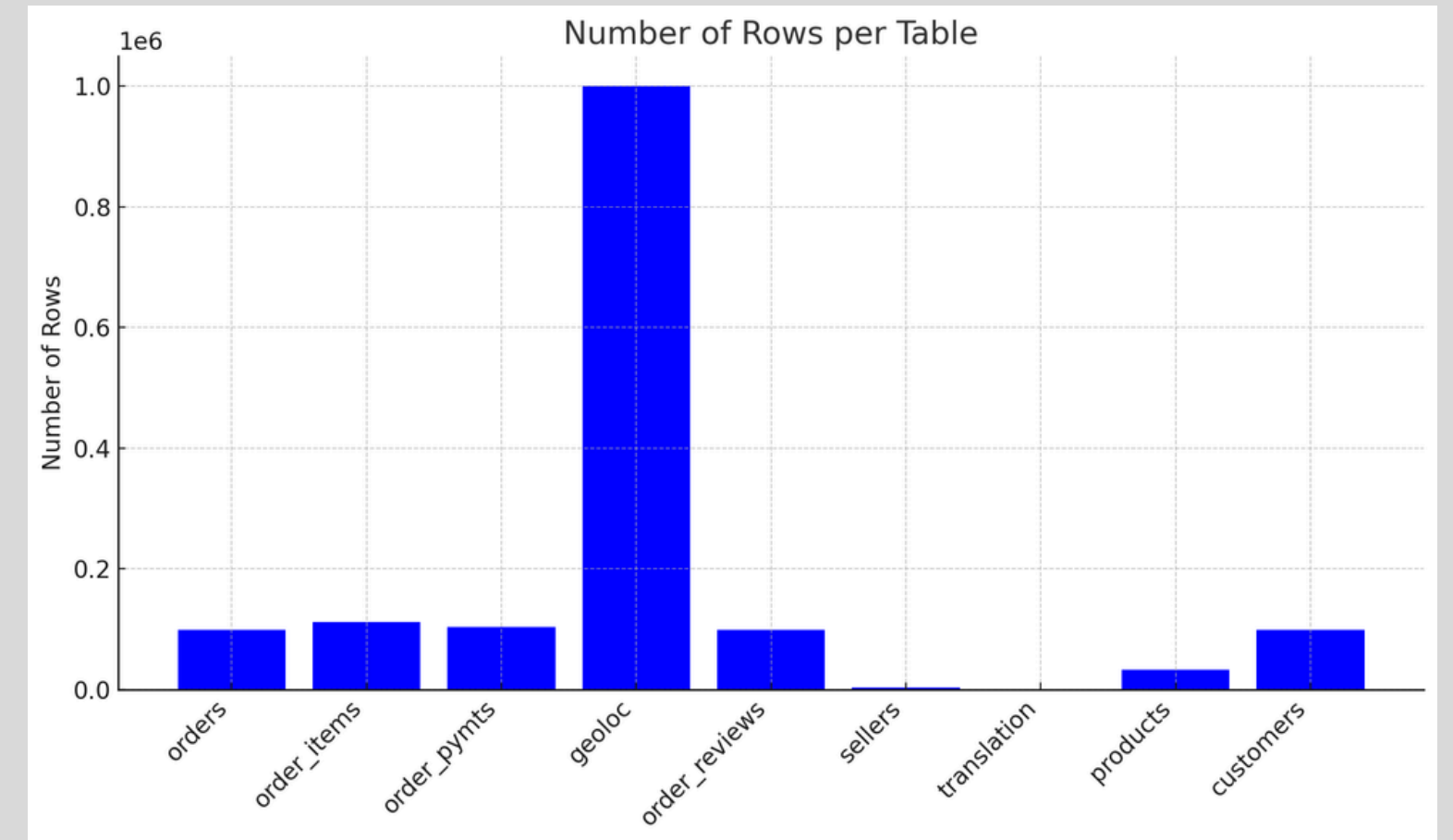
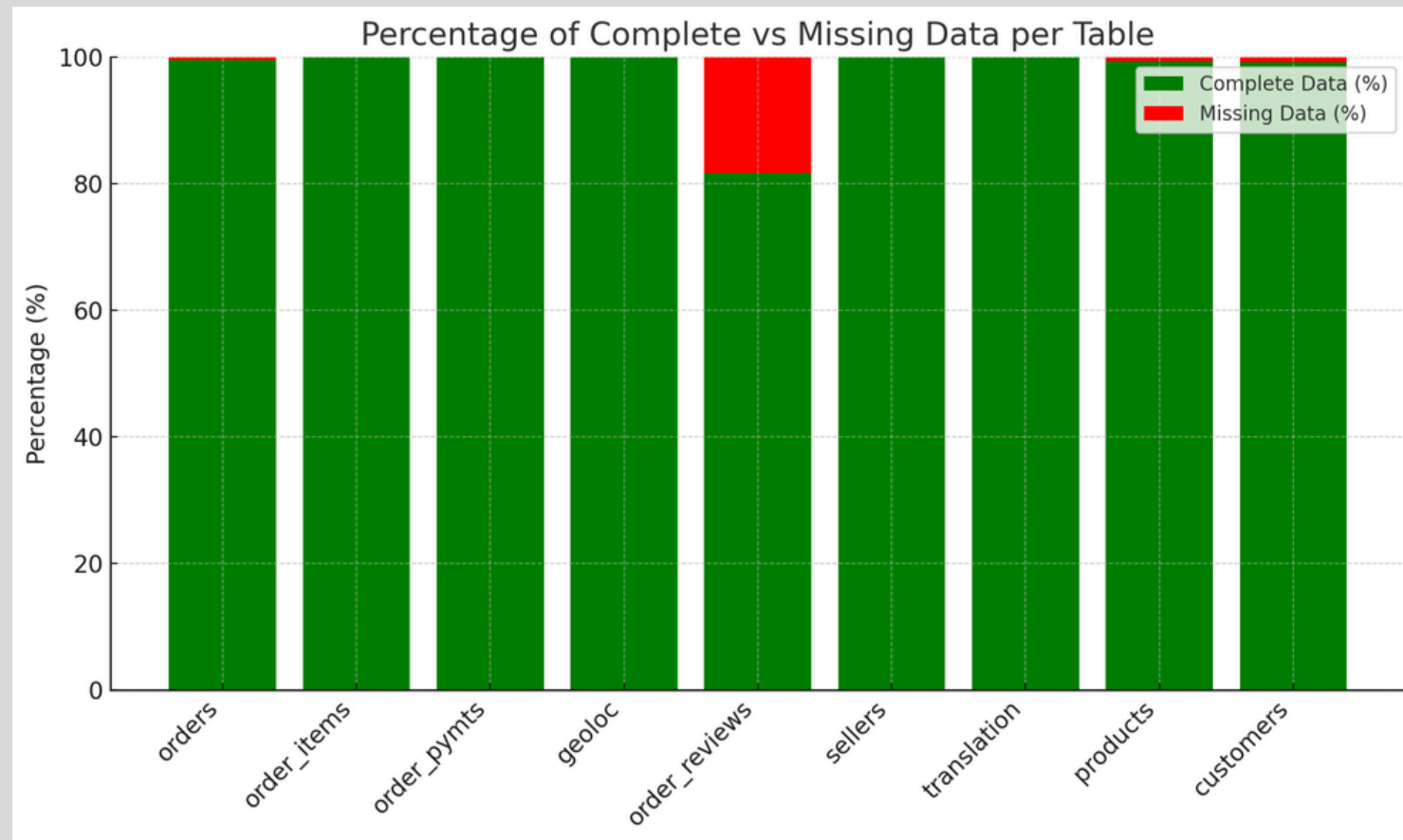
1. MISSION

Créer une segmentation client exploitée par l'équipe Marketing pour personnaliser les campagnes et améliorer l'engagement.



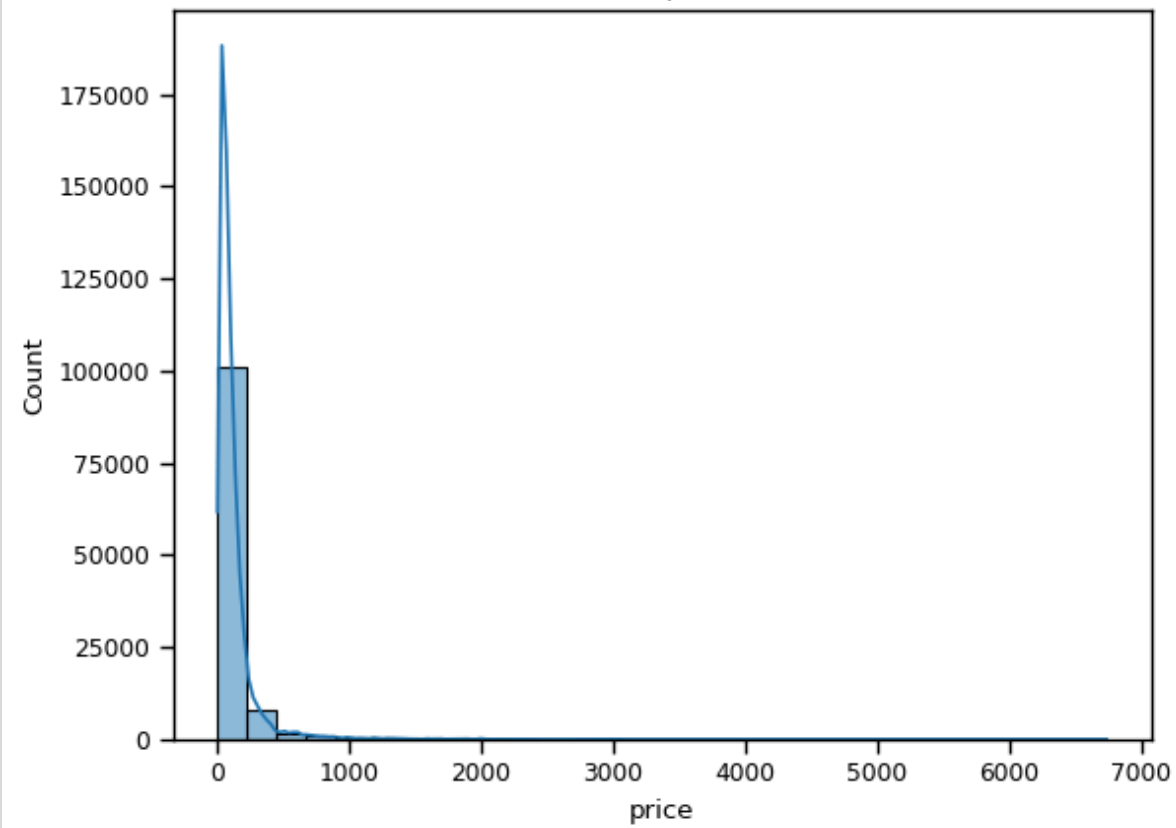
- Comment assurer que la segmentation client permettra à l'équipe Marketing de cibler efficacement chaque segment avec des offres personnalisées tout en optimisant les ressources
- Comment intégrer efficacement les variables qualitatives, comme la satisfaction, dans la segmentation en complément des données quantitatives ?
- Comment créer une segmentation dynamique qui reflète l'évolution des comportements clients et reste pertinente à long terme, avec une fréquence de mise à jour adaptée ?

2. PRESENTATION DU JEU DE DONNÉE

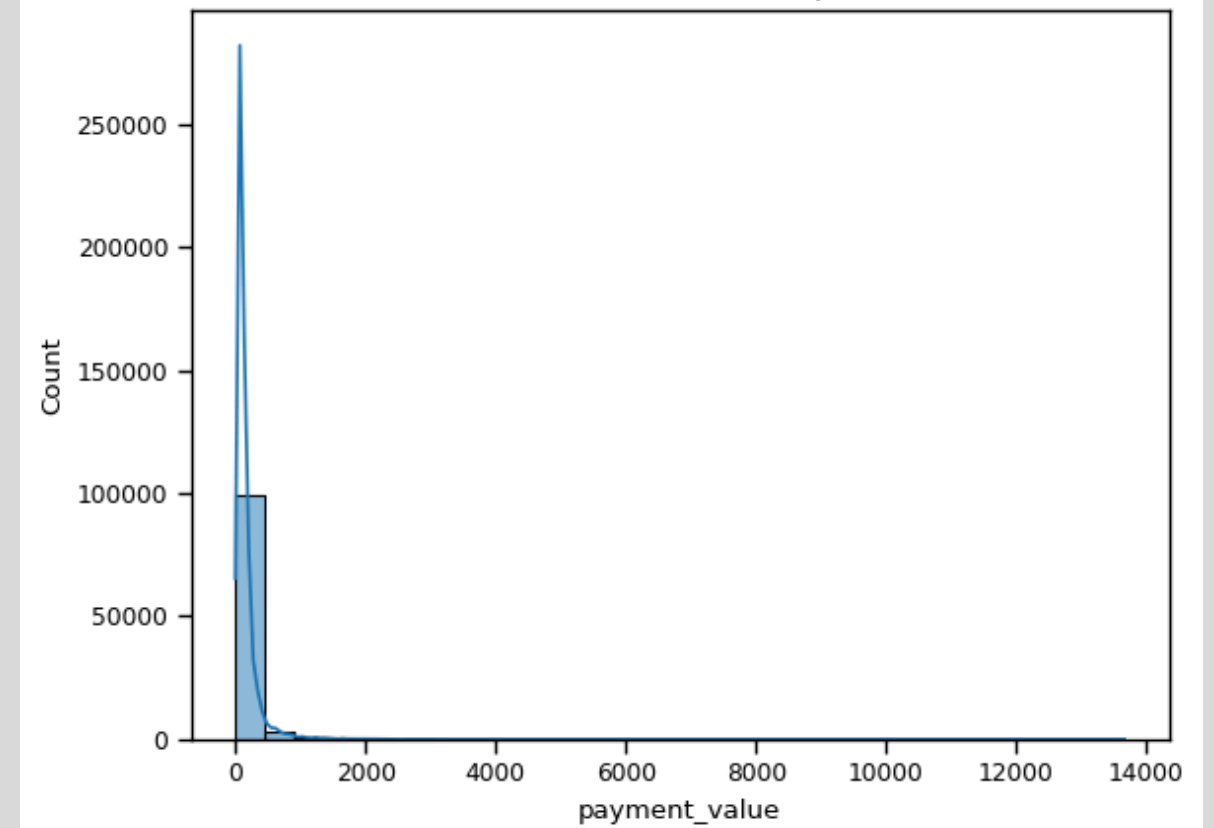


3. ANALYSE EXPLORATOIRE

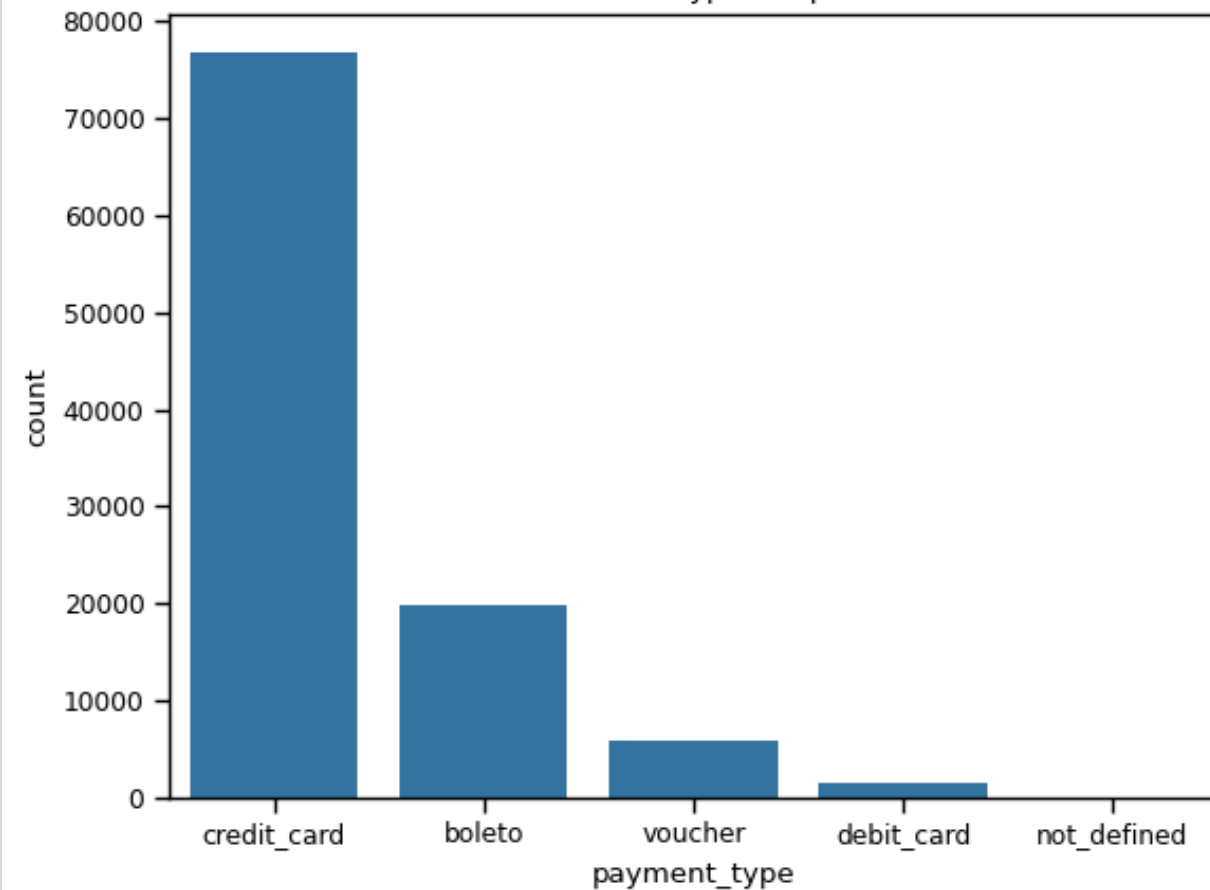
Distribution des prix des articles



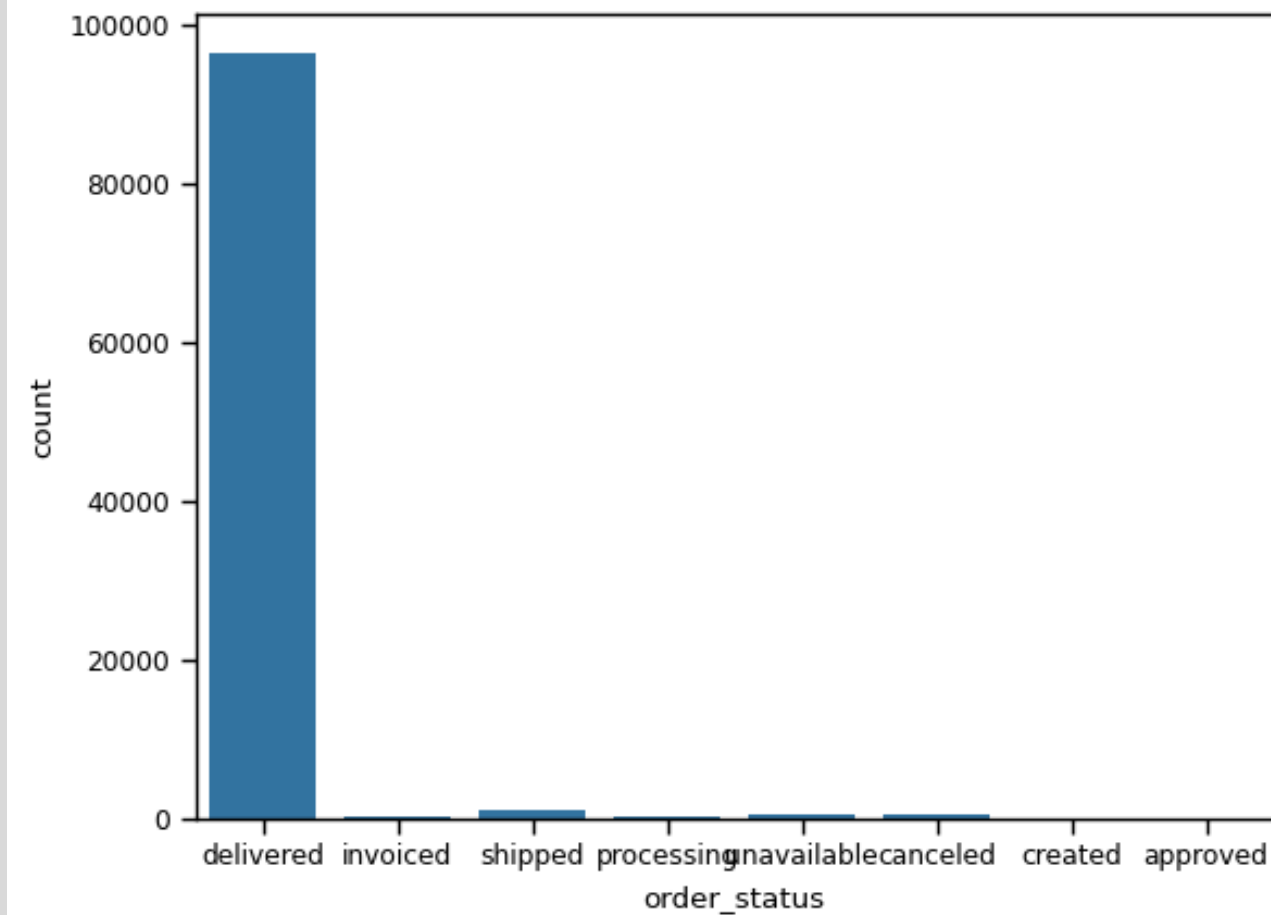
Distribution des montants des paiements

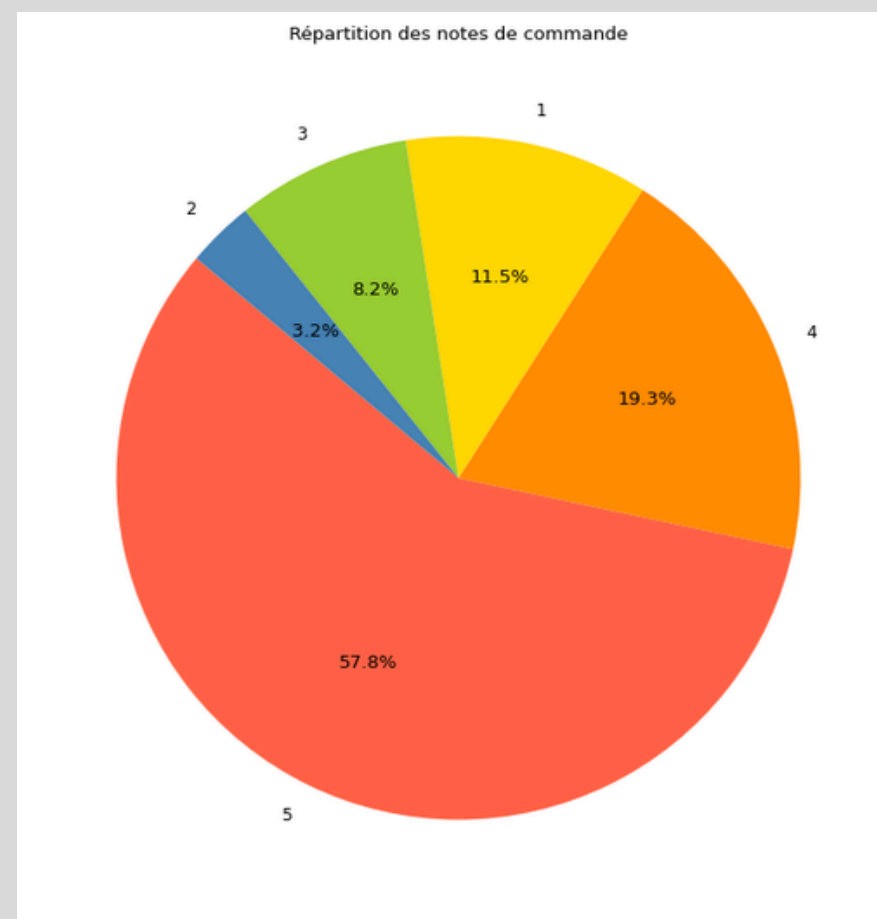
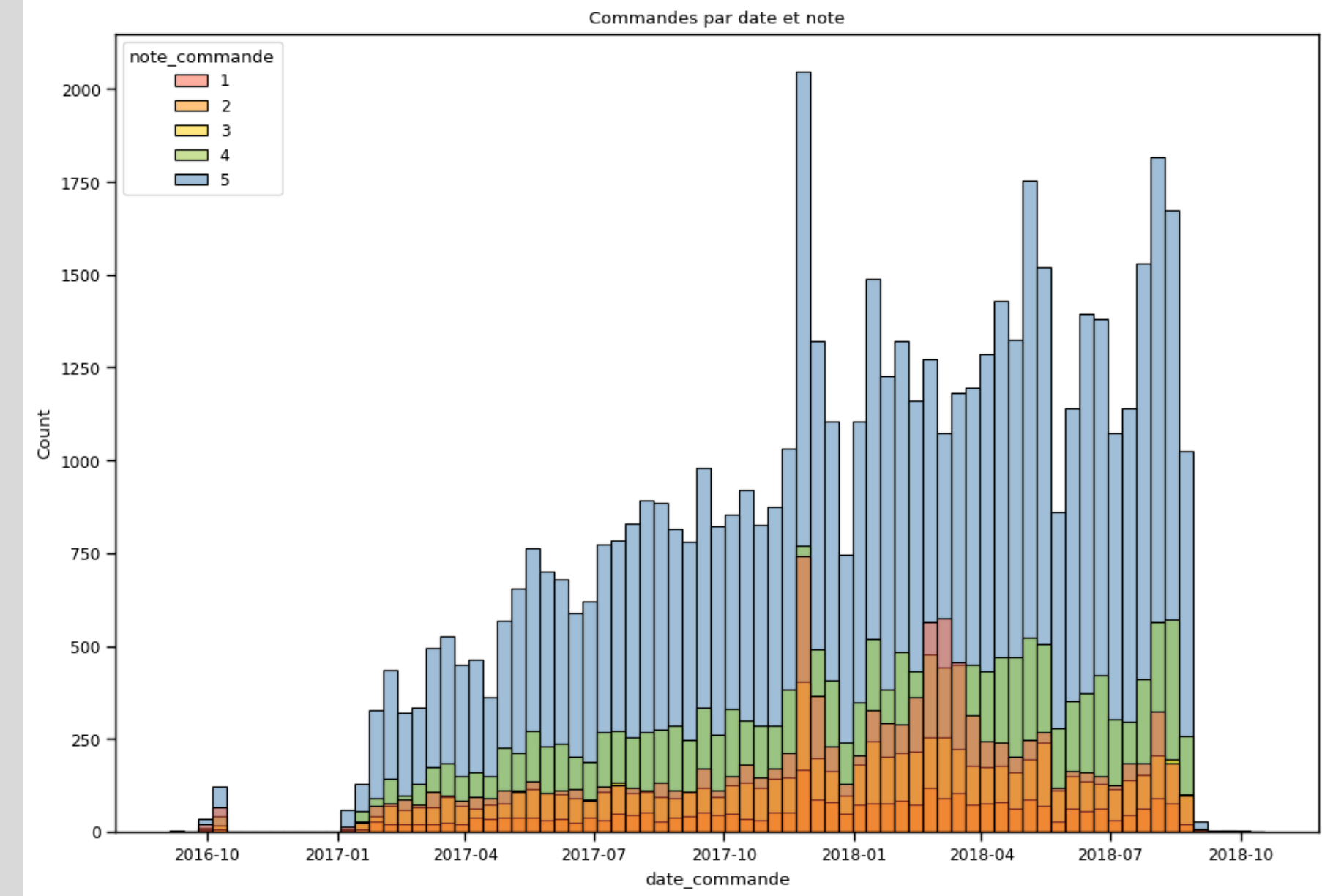
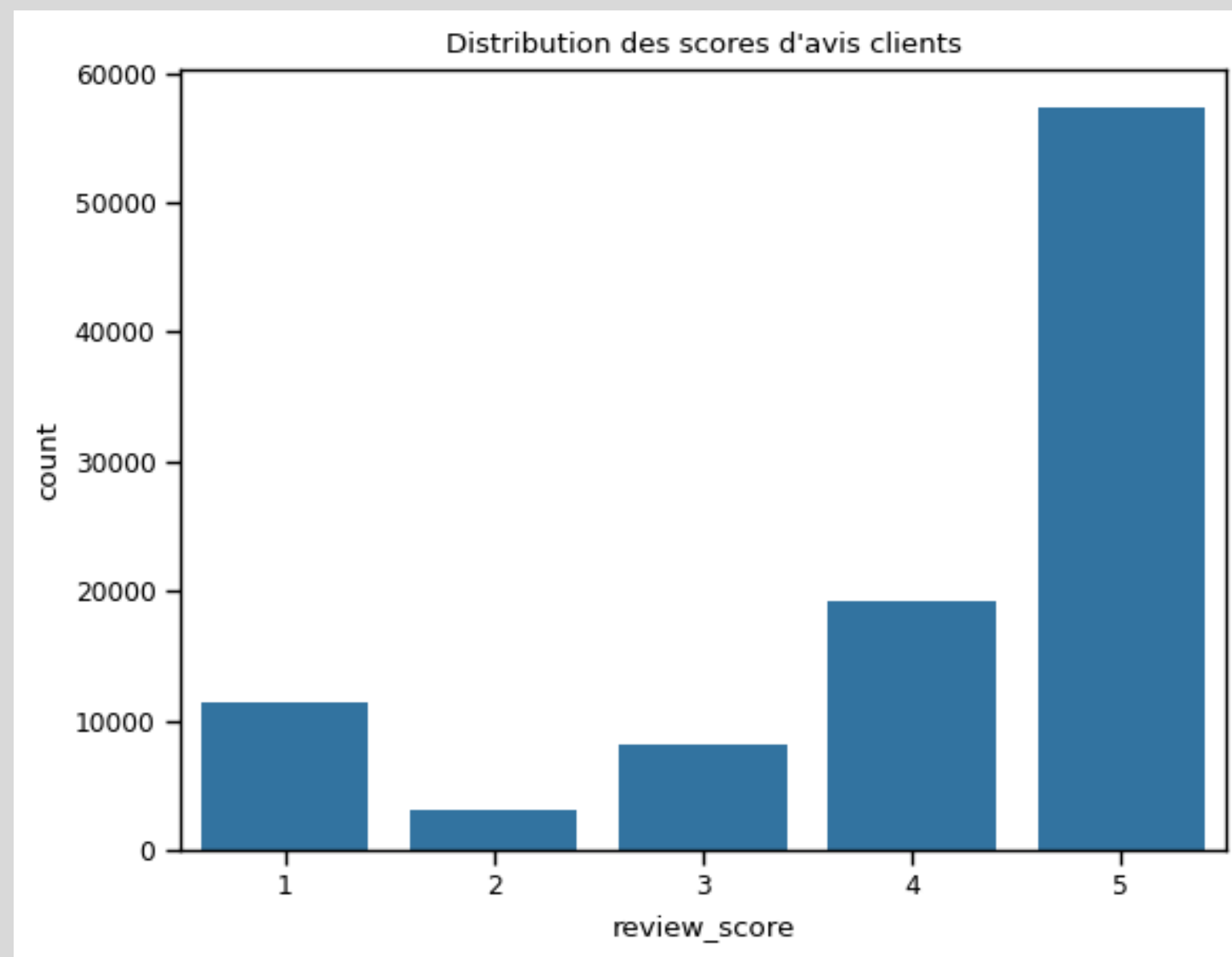


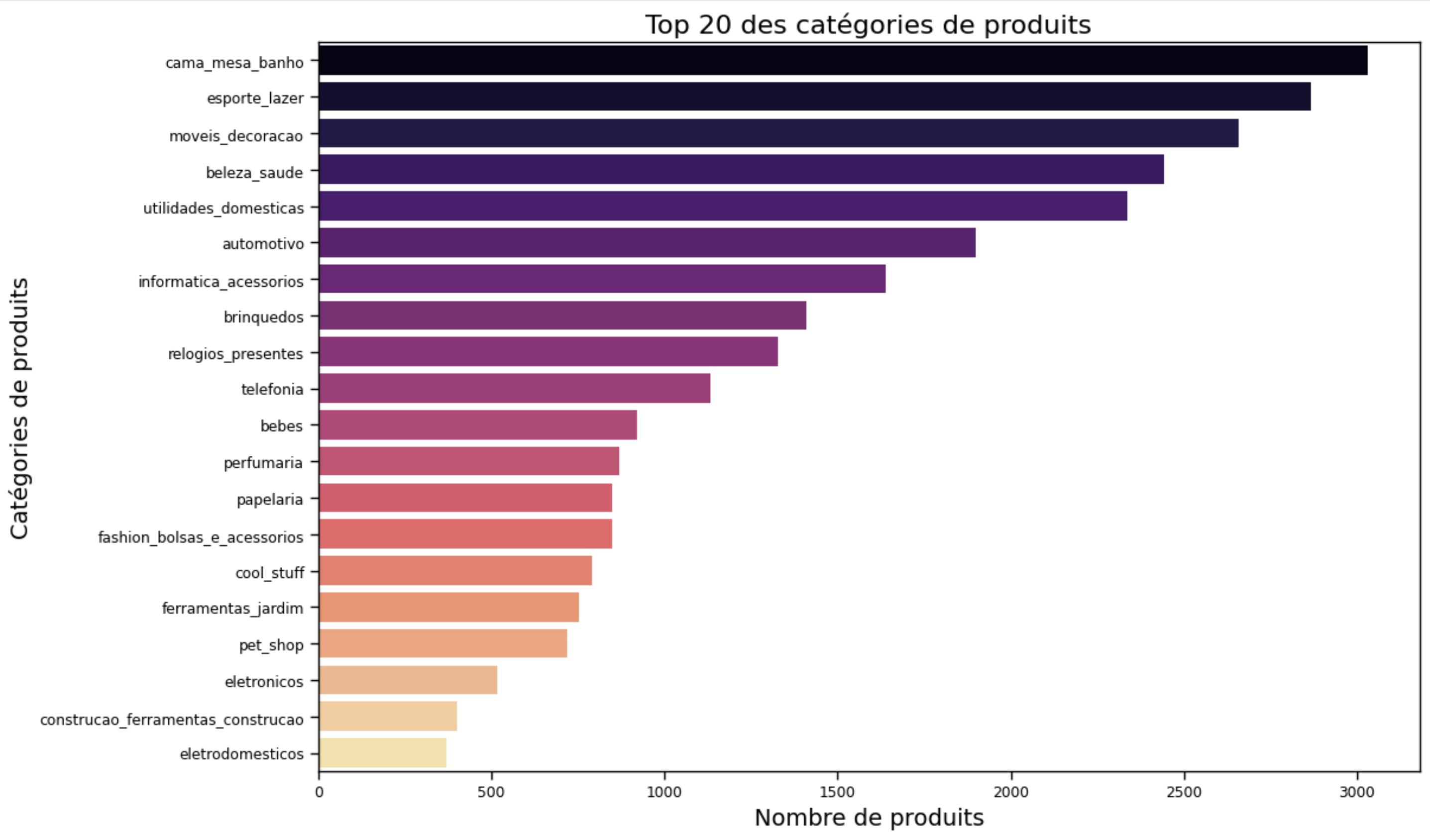
Distribution des types de paiement



Distribution des états de commande





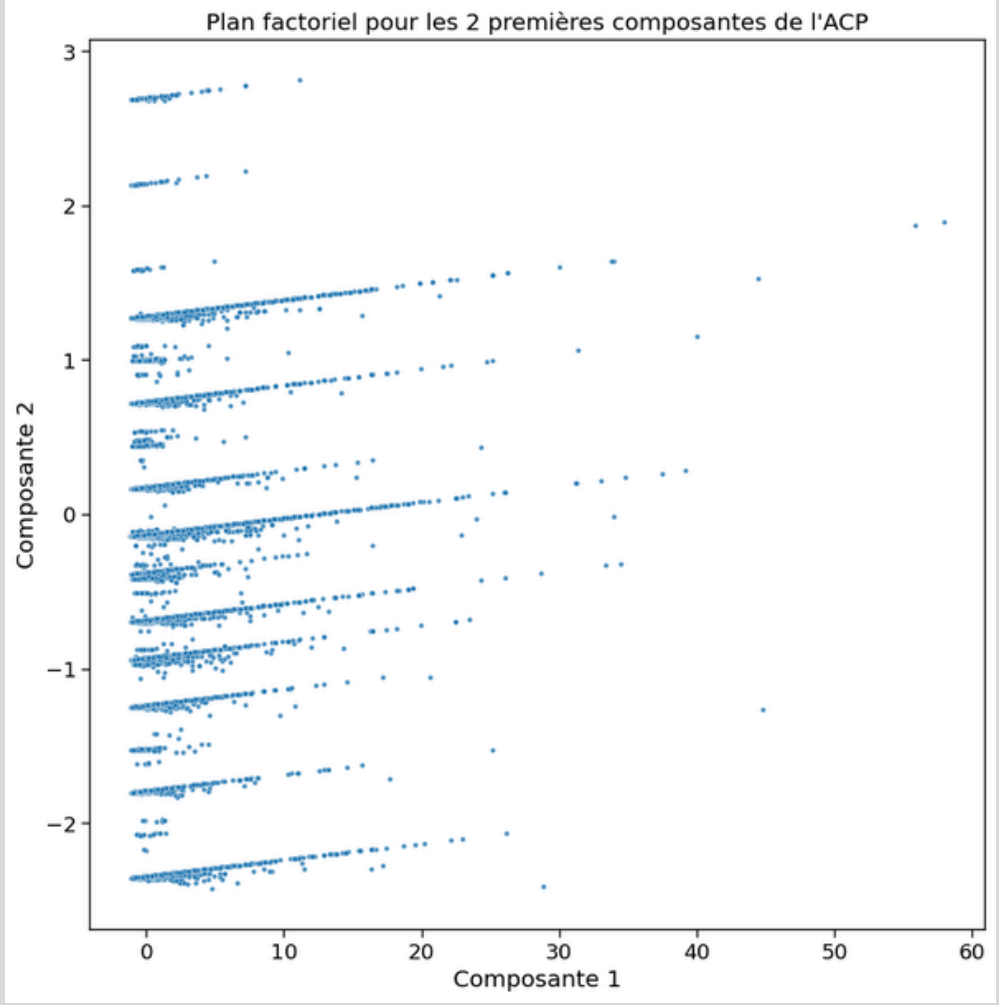
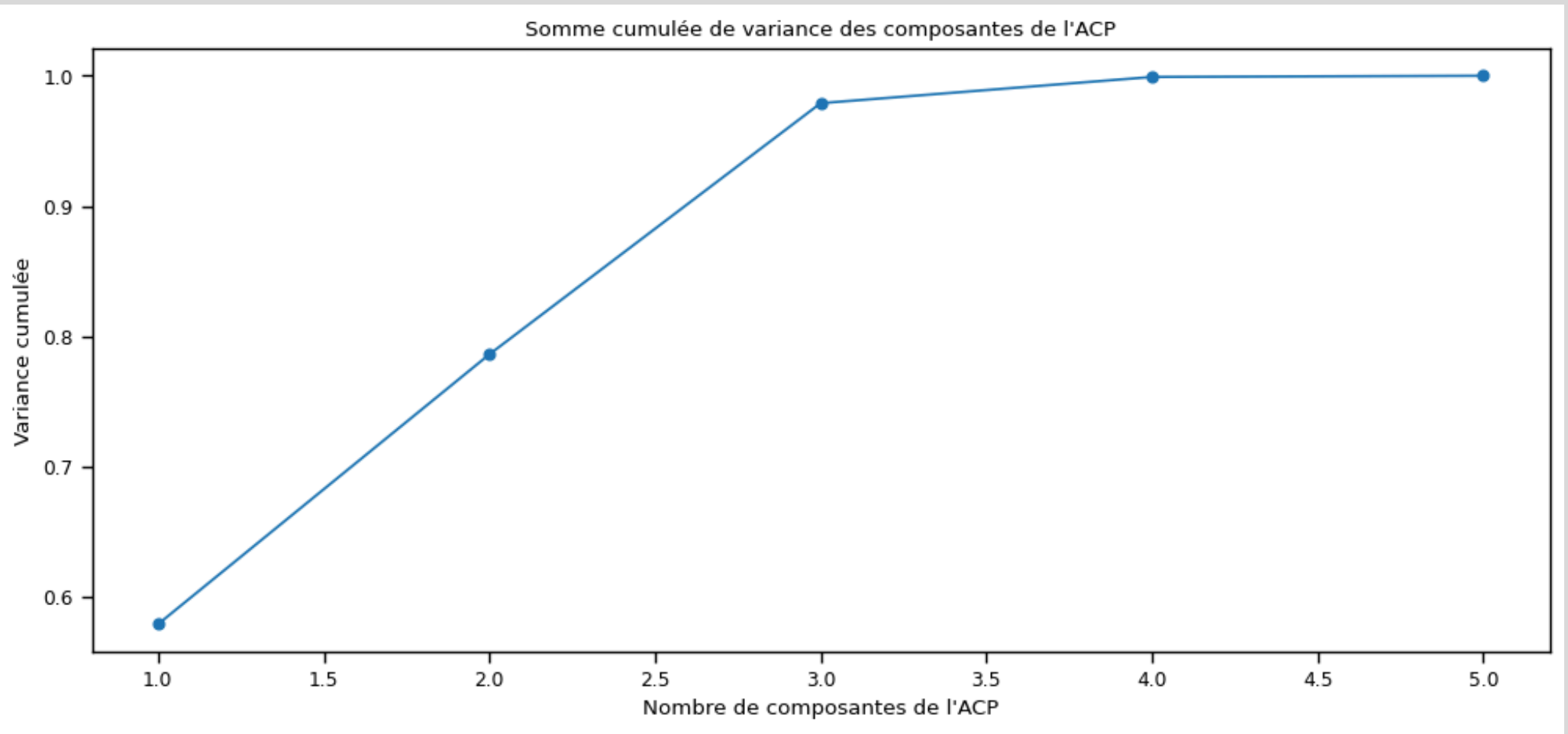
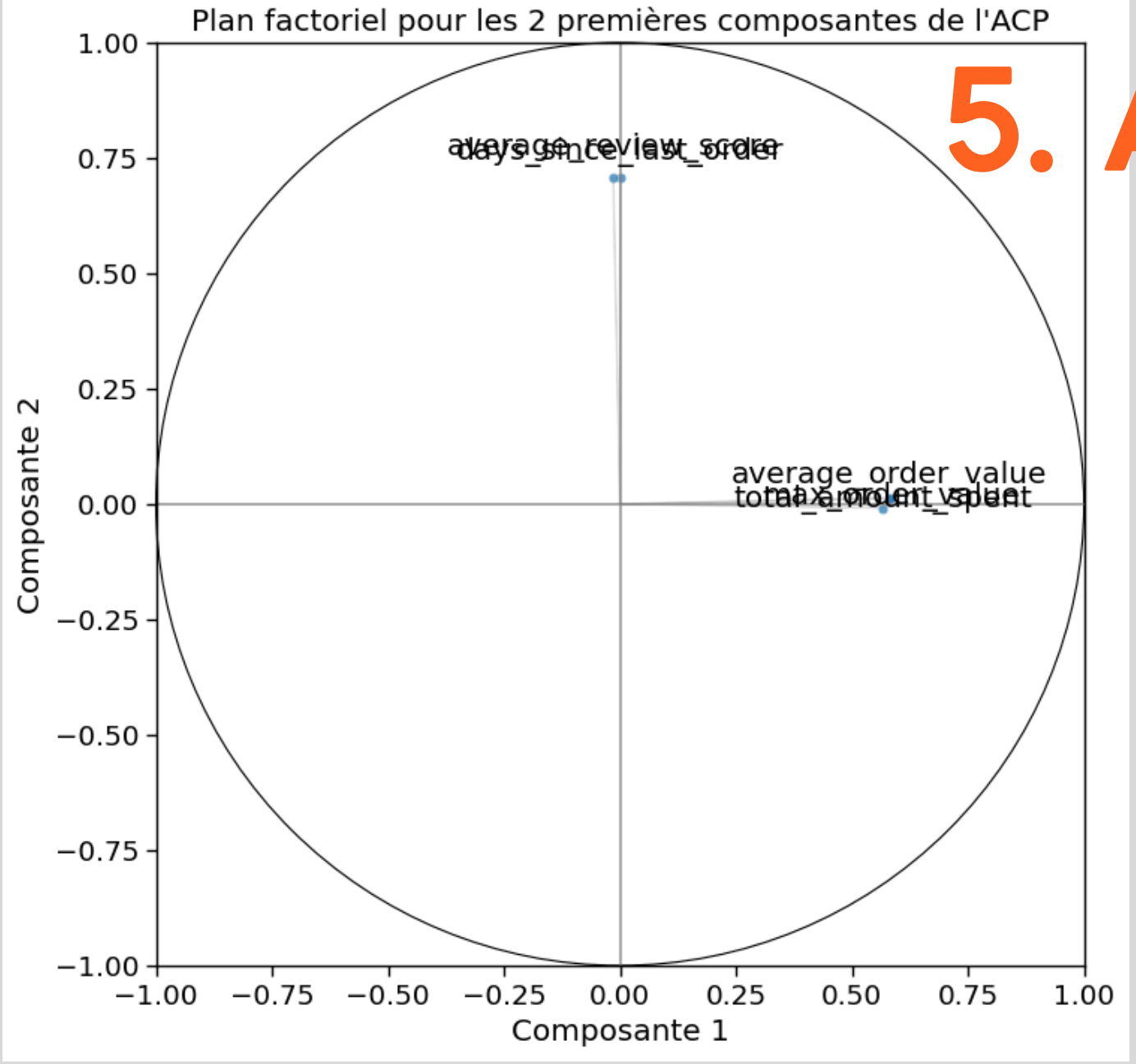


4. CREATION DU FICHER CLIENT

- frequency
- days_since_last_order
- first_order_date
- total_amount_spent
- average_order_value
- max_order_value
- average_freight_value
- most_frequent_payment_type
- average_review_score

	customer_unique_id	frequency	total_amount_spent	average_review_score	days_since_last_order	first_order_date	average_order_value	max_order_value	average_freight_value	most_frequent_payment
0	0000366f3b9a7992bf8c76cfd3221e2	1	129.90	5.0	6	2018-05-10 10:56:27	129.90	129.90	12.00	cred
1	0000b849f77a49e4a4ce2b2a4ca5be3f	1	18.90	4.0	6	2018-05-07 11:11:27	18.90	18.90	8.29	cred
2	0000f46a3911fa3c0805444483337064	1	69.00	3.0	7	2017-03-10 21:05:03	69.00	69.00	17.22	cred
3	0000f6ccb0745a6a4b88665a16c9f078	1	25.99	4.0	7	2017-10-12 20:29:41	25.99	25.99	17.63	cred
4	0004aac84e0df4da2b147fca70cf8255	1	180.00	5.0	7	2017-11-14 19:45:42	180.00	180.00	16.89	cred

5. ACP



6. ALGORITHME K-MEANS

6.1 Trouver l'optimum du nombre de clusters

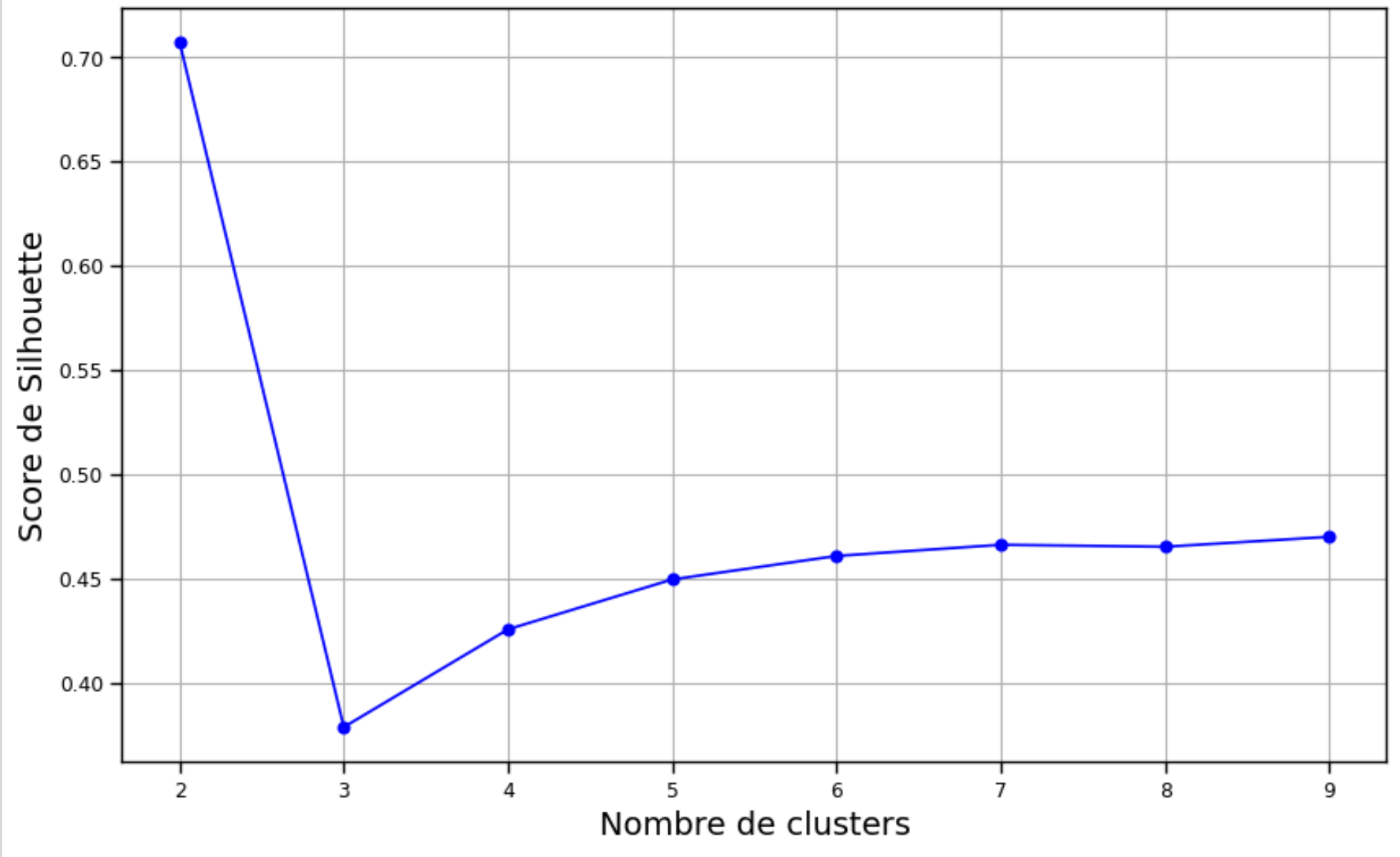
- calcul des Scores de Silhouette
- l'Indice de Davies-Bouldin

6.2 STATISTIQUE DES CLUSTERS

- La répartition des cluster
- La description des cluster

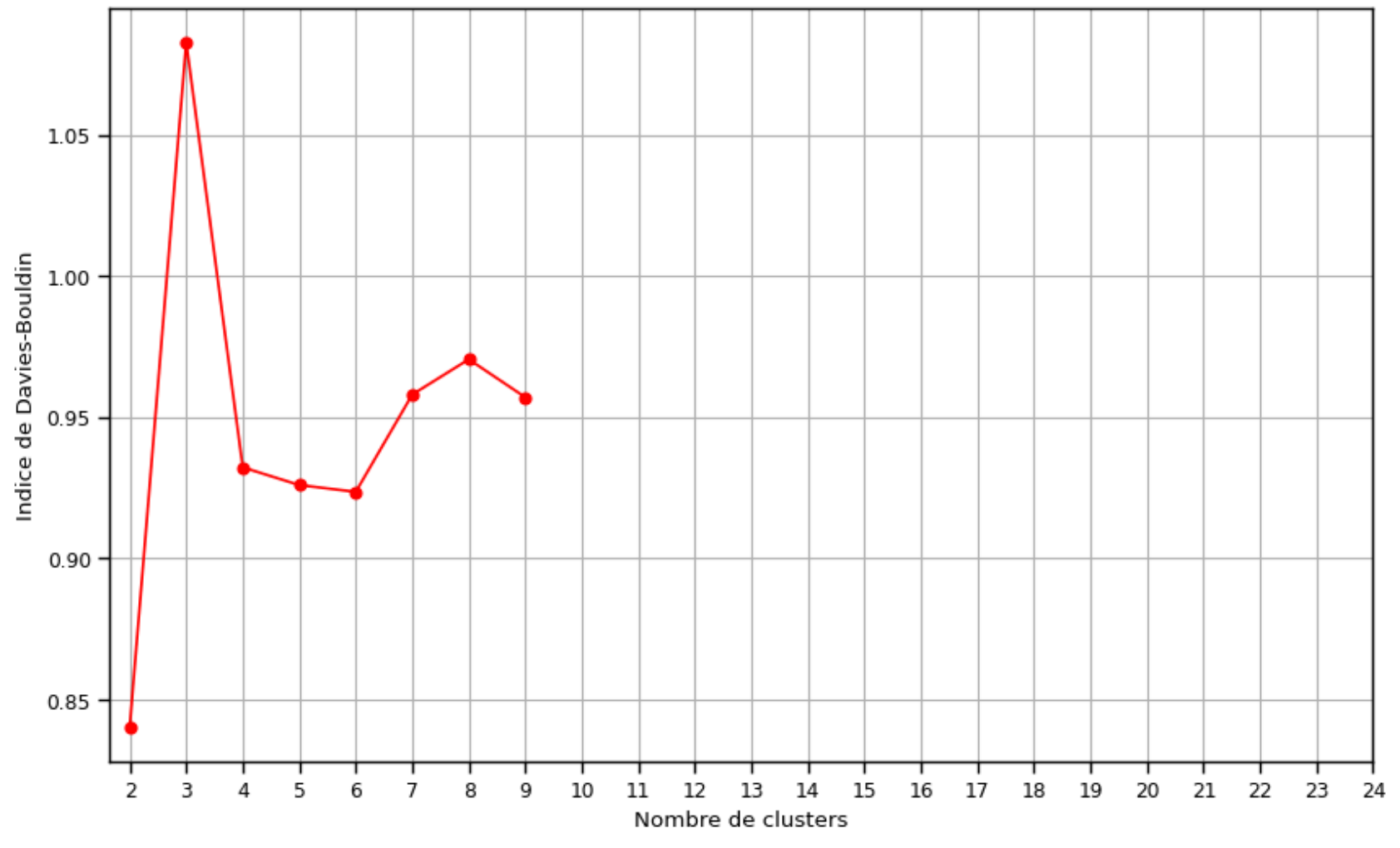
SCORES DE SILHOUETTE

Score de Silhouette en fonction du nombre de clusters



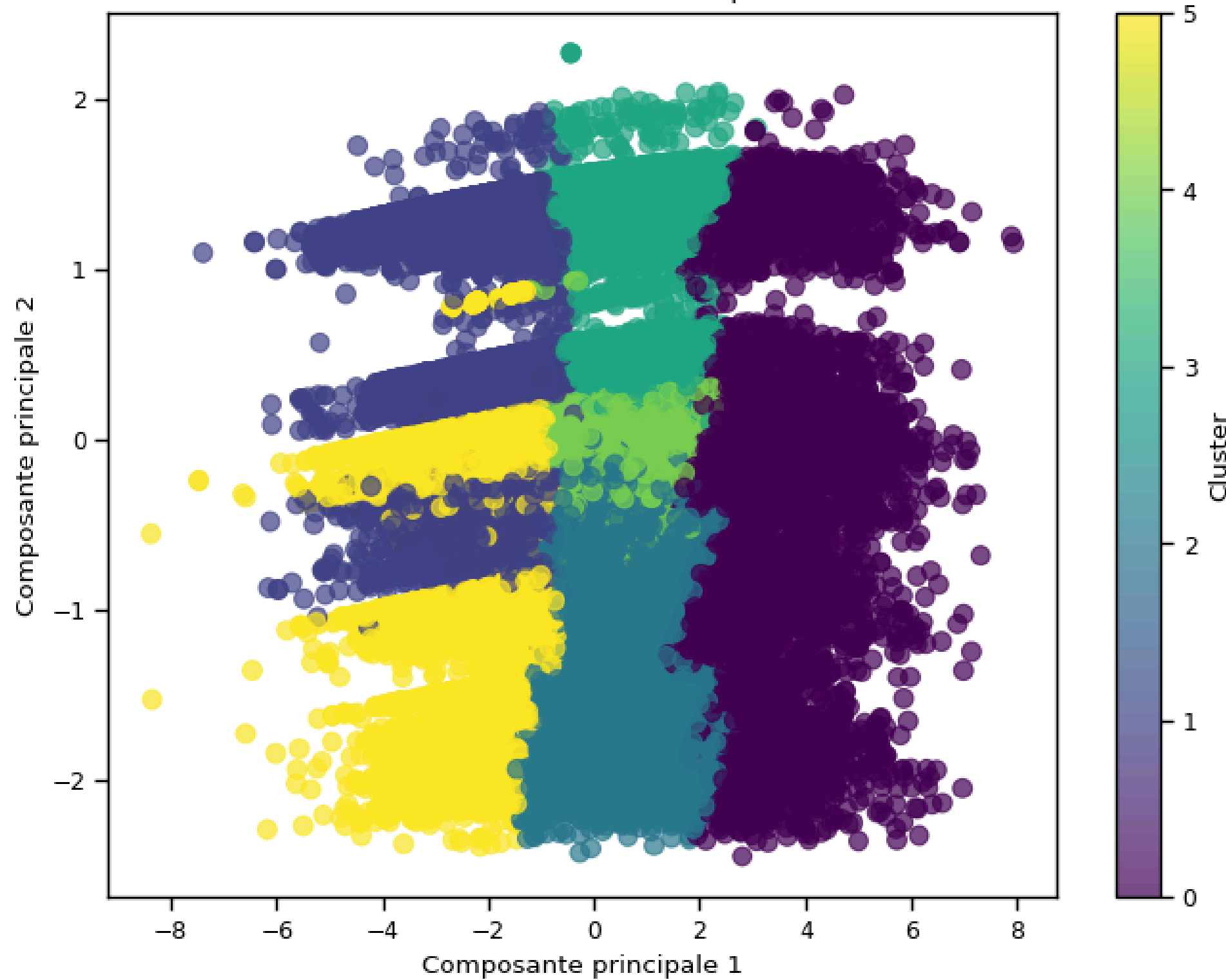
INDICE DE DAVIES-BOULDIN

Indice de Davies-Bouldin en fonction du nombre de clusters

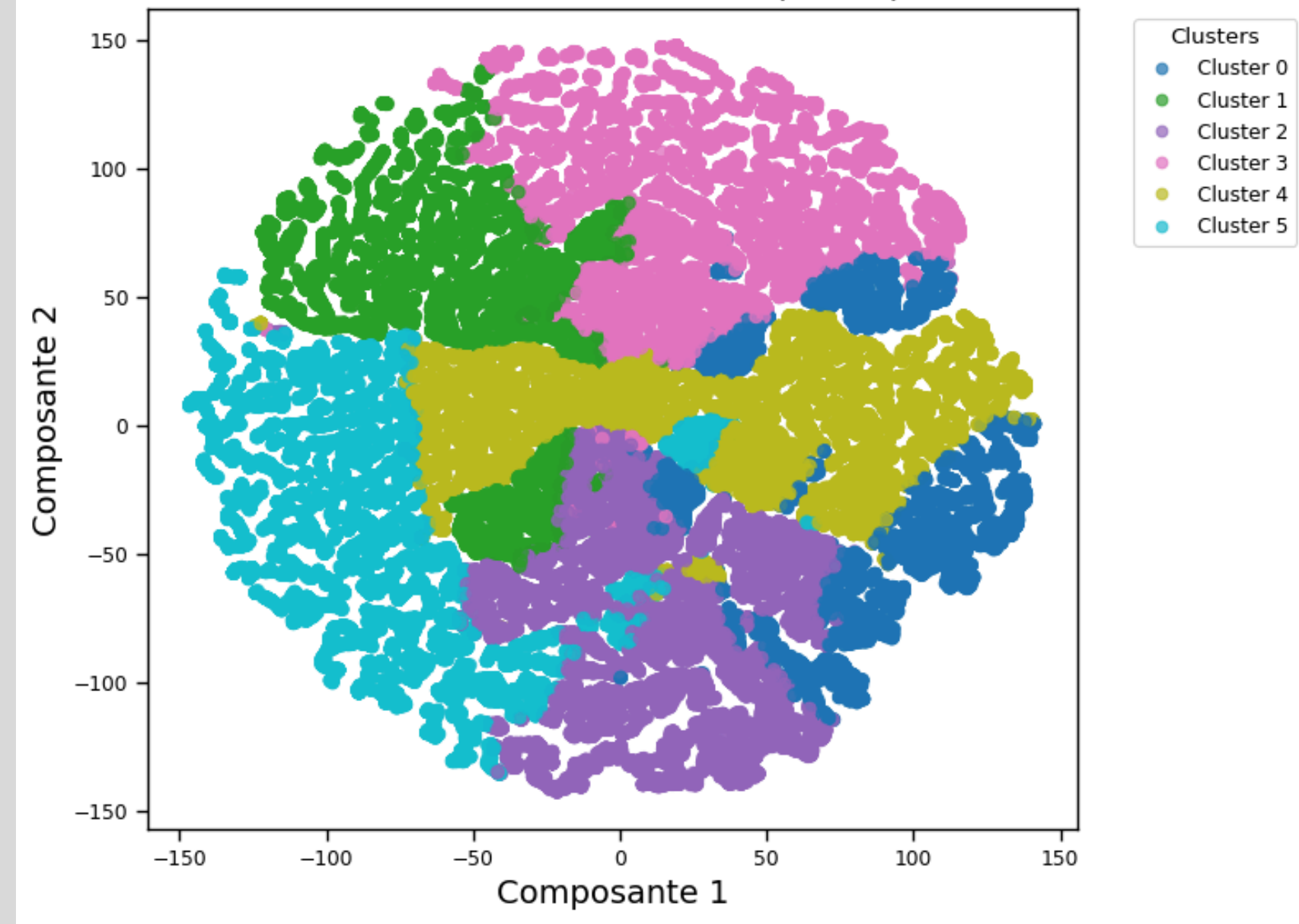


Le Cluster formé

Visualisation des clusters K-Means avec k=6 après PowerTransformer

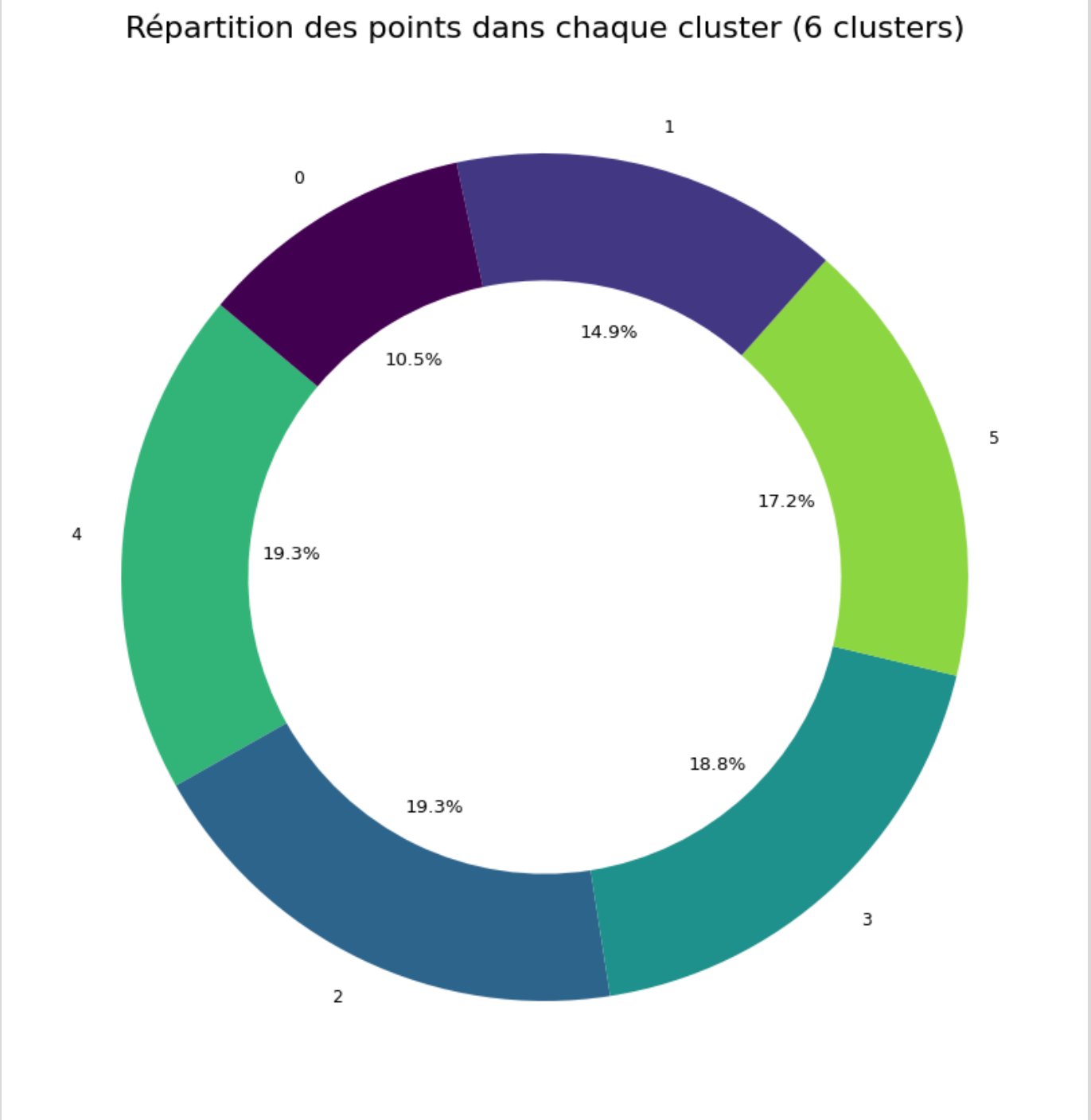
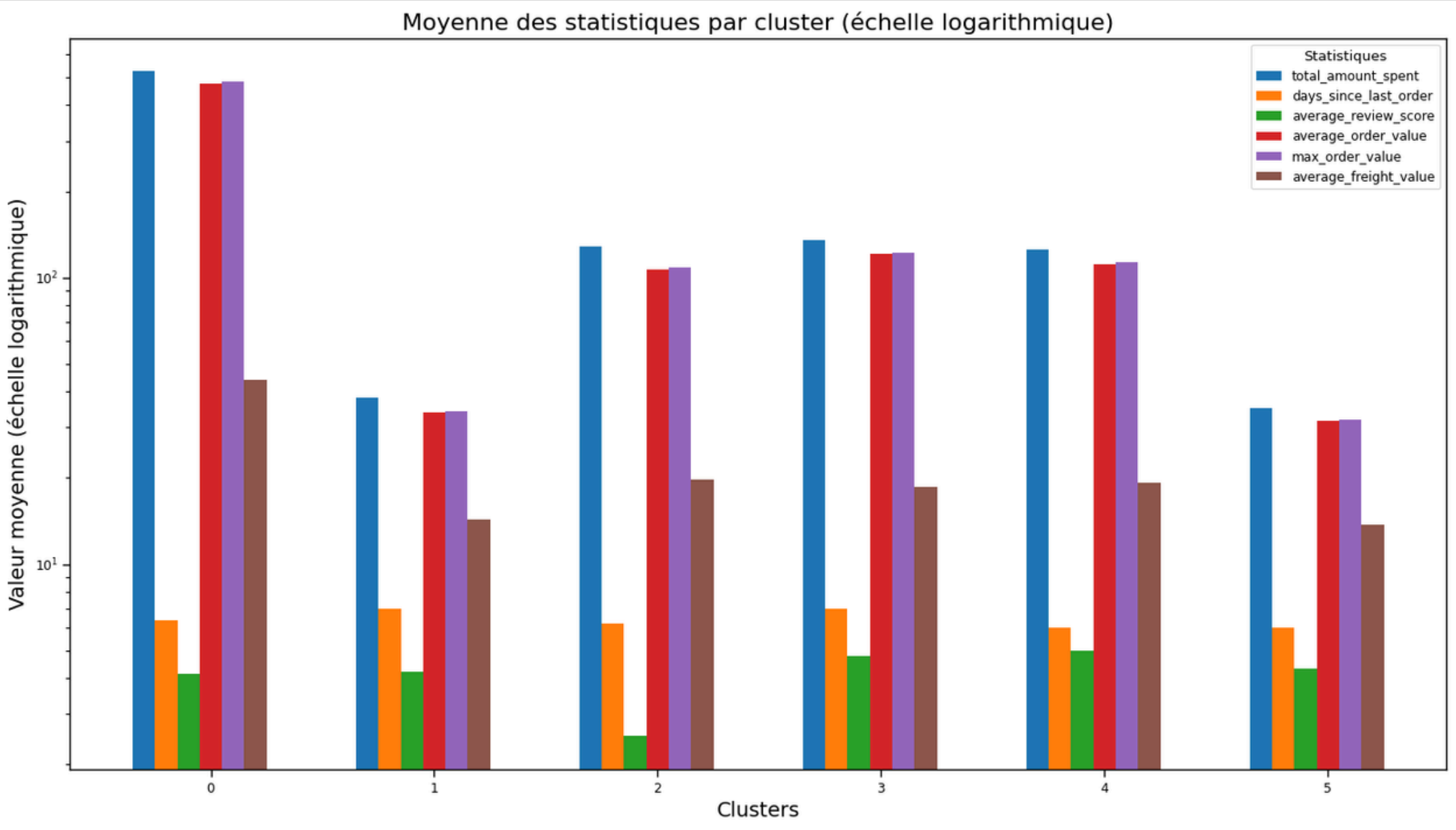


K-Means avec 6 clusters (t-SNE)

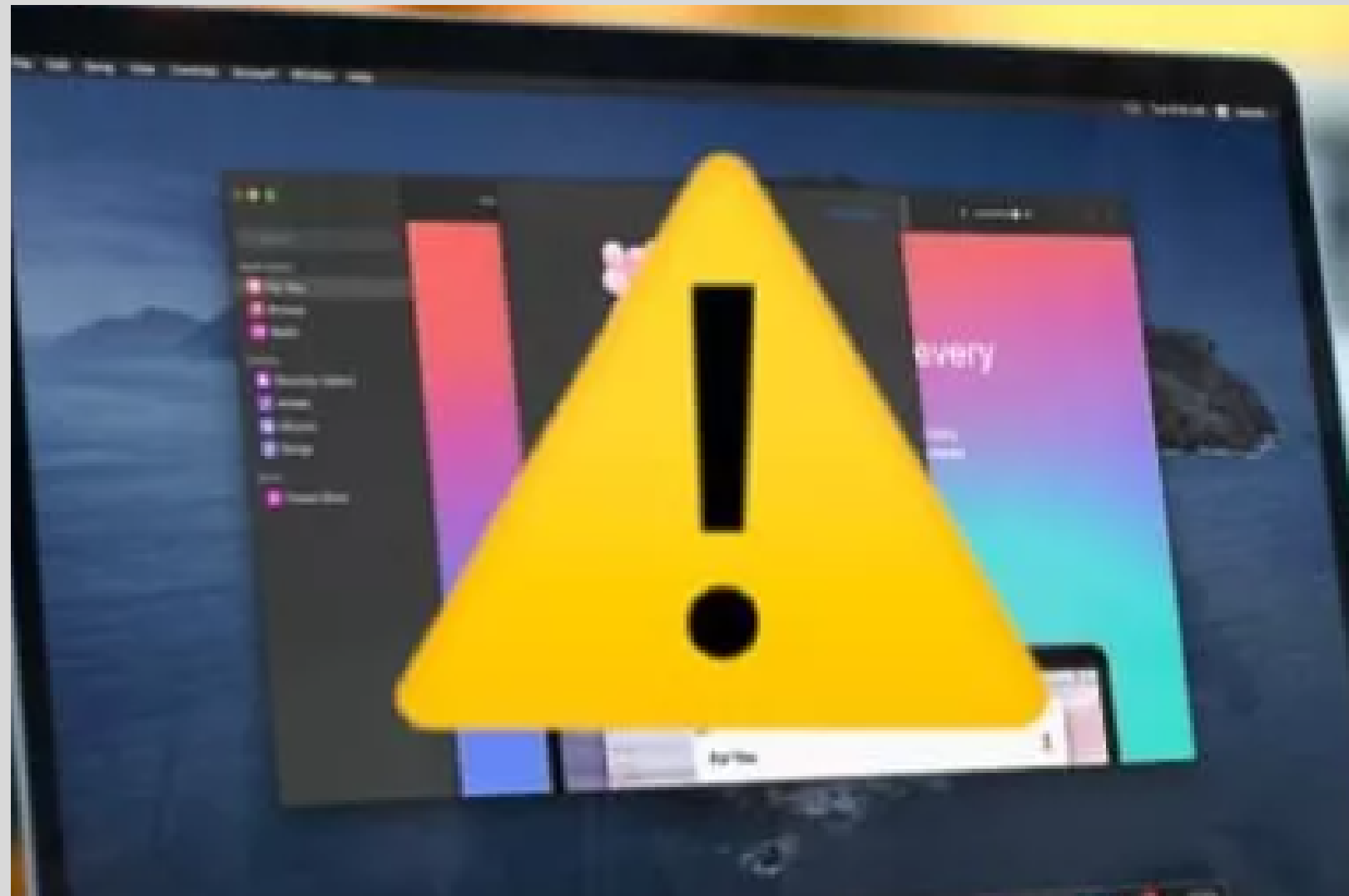


La répartition des cluster

cluster	
4	18273
2	18271
3	17799
5	16271
1	14114
0	9989
Name: count, dtype: int64	

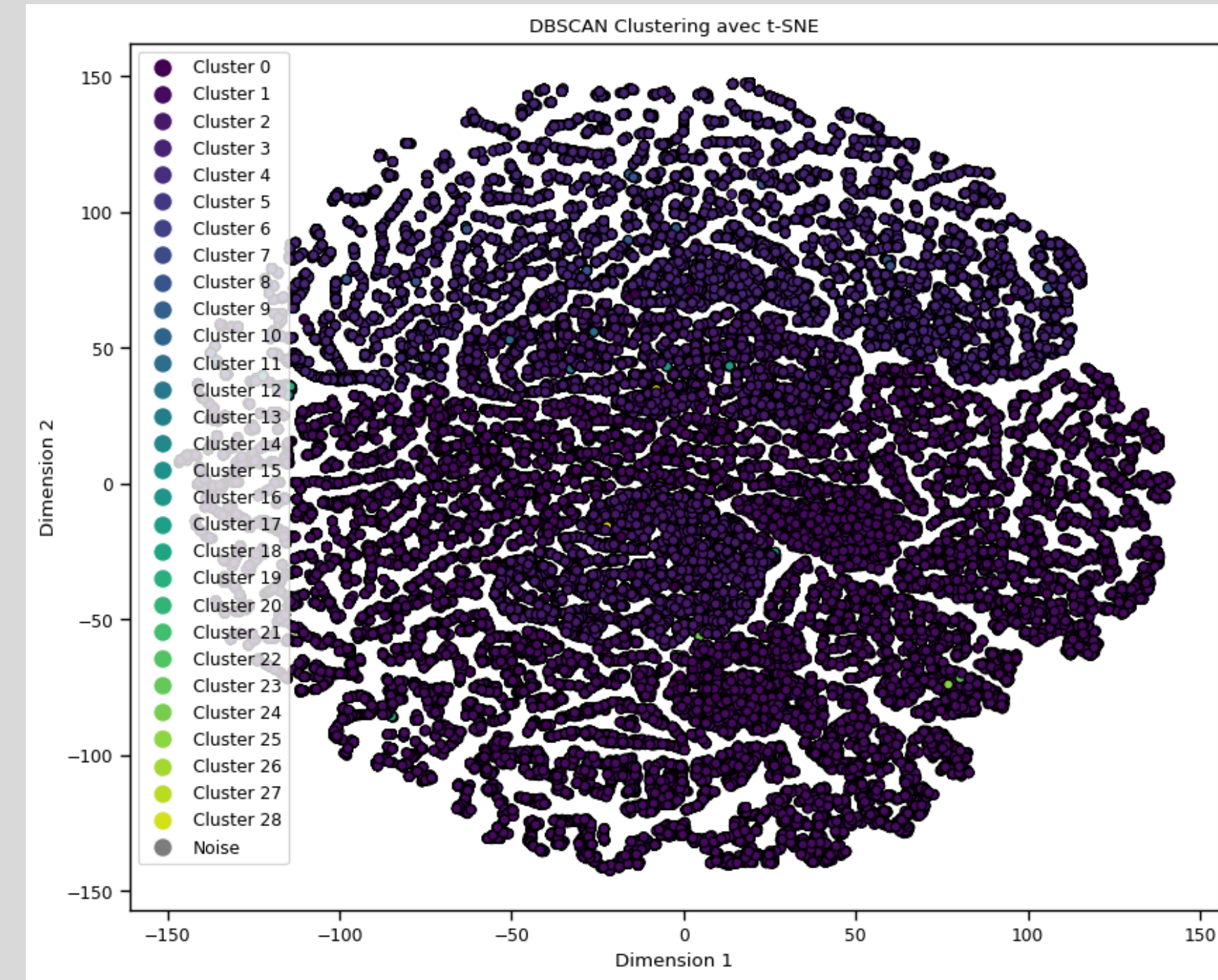
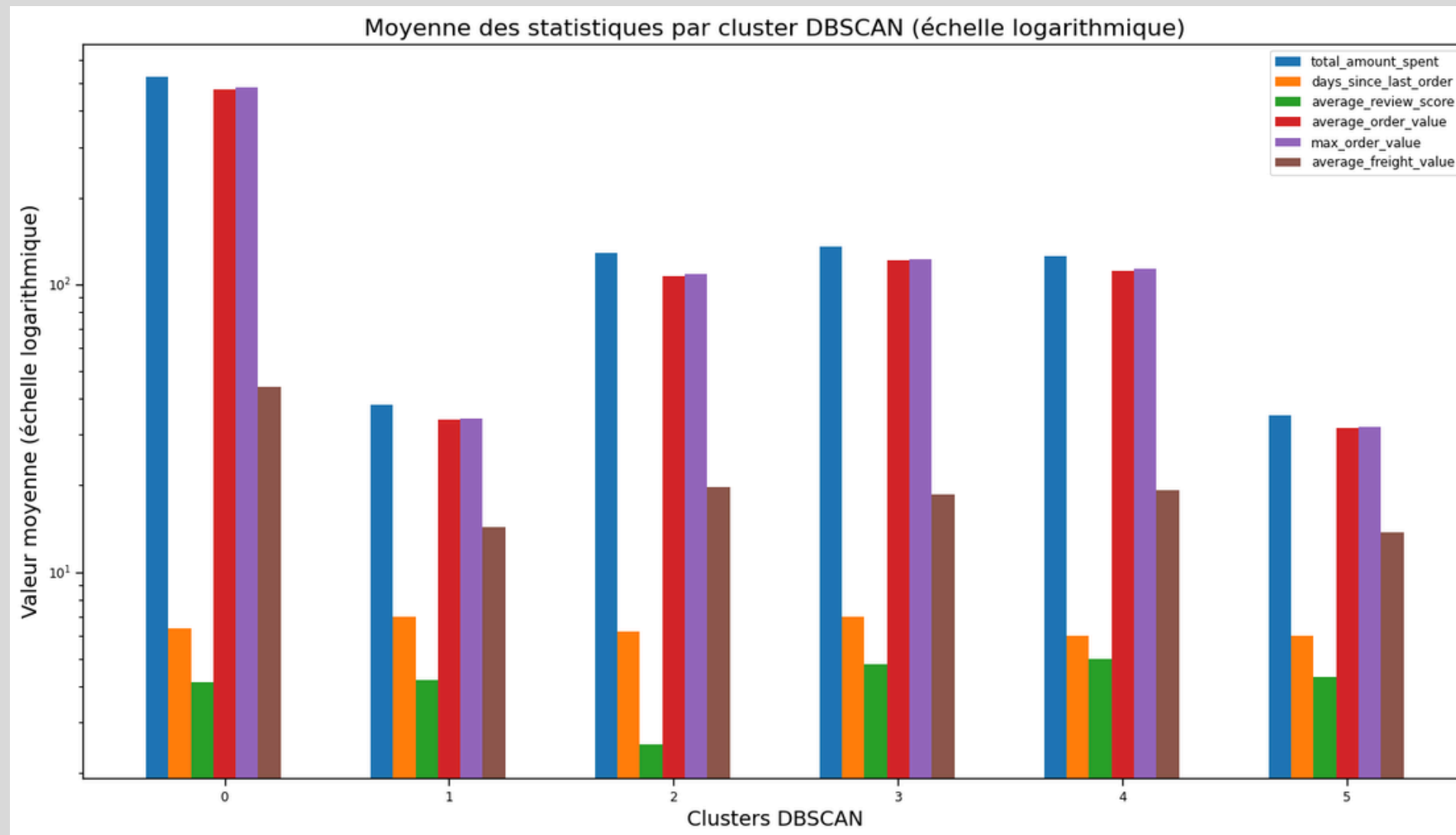


7. CLUSTERING HIÉRARCHIQUE



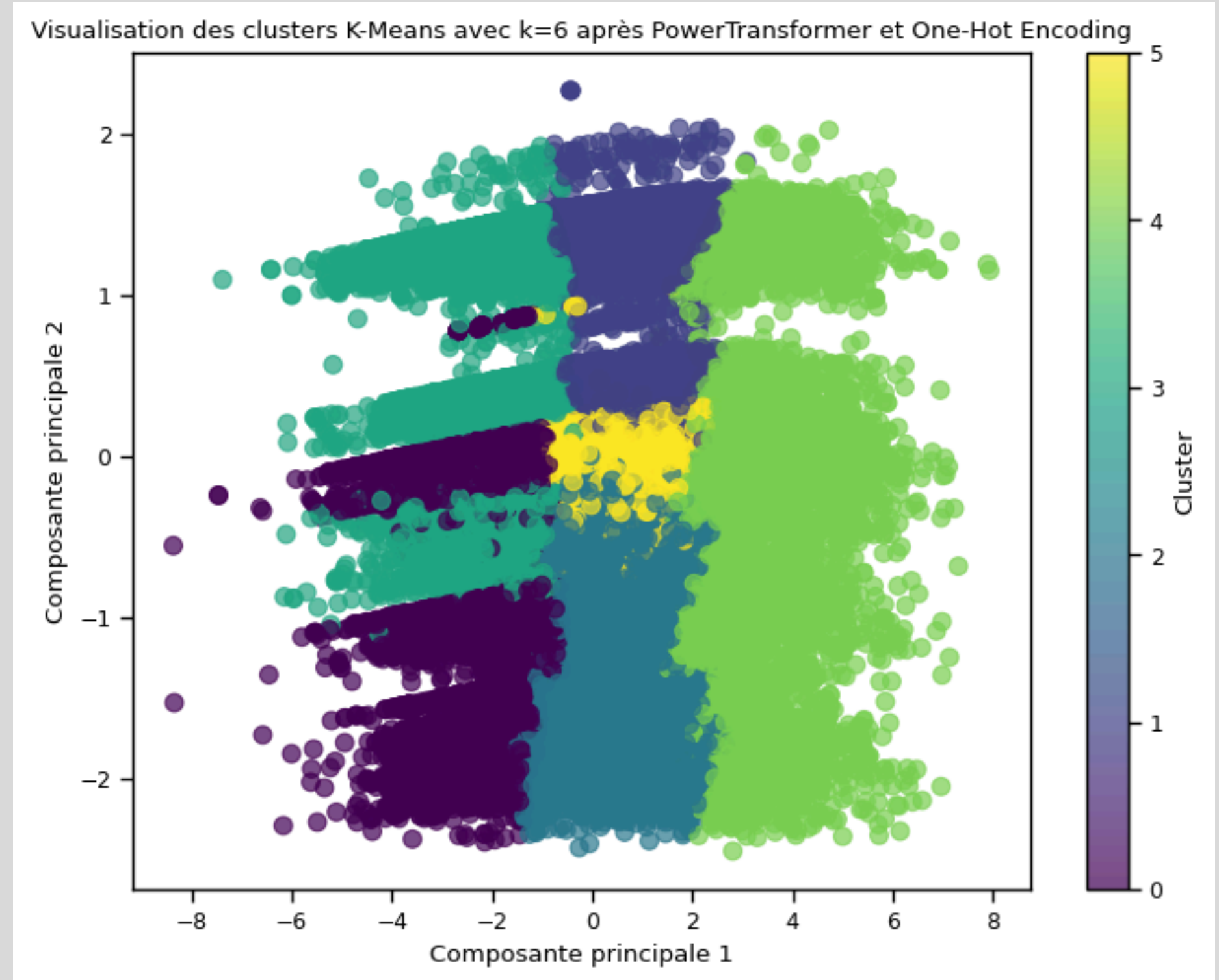
Non applicable en raison d'un problème de mémoire.

8. ALGORITHME DBSCAN

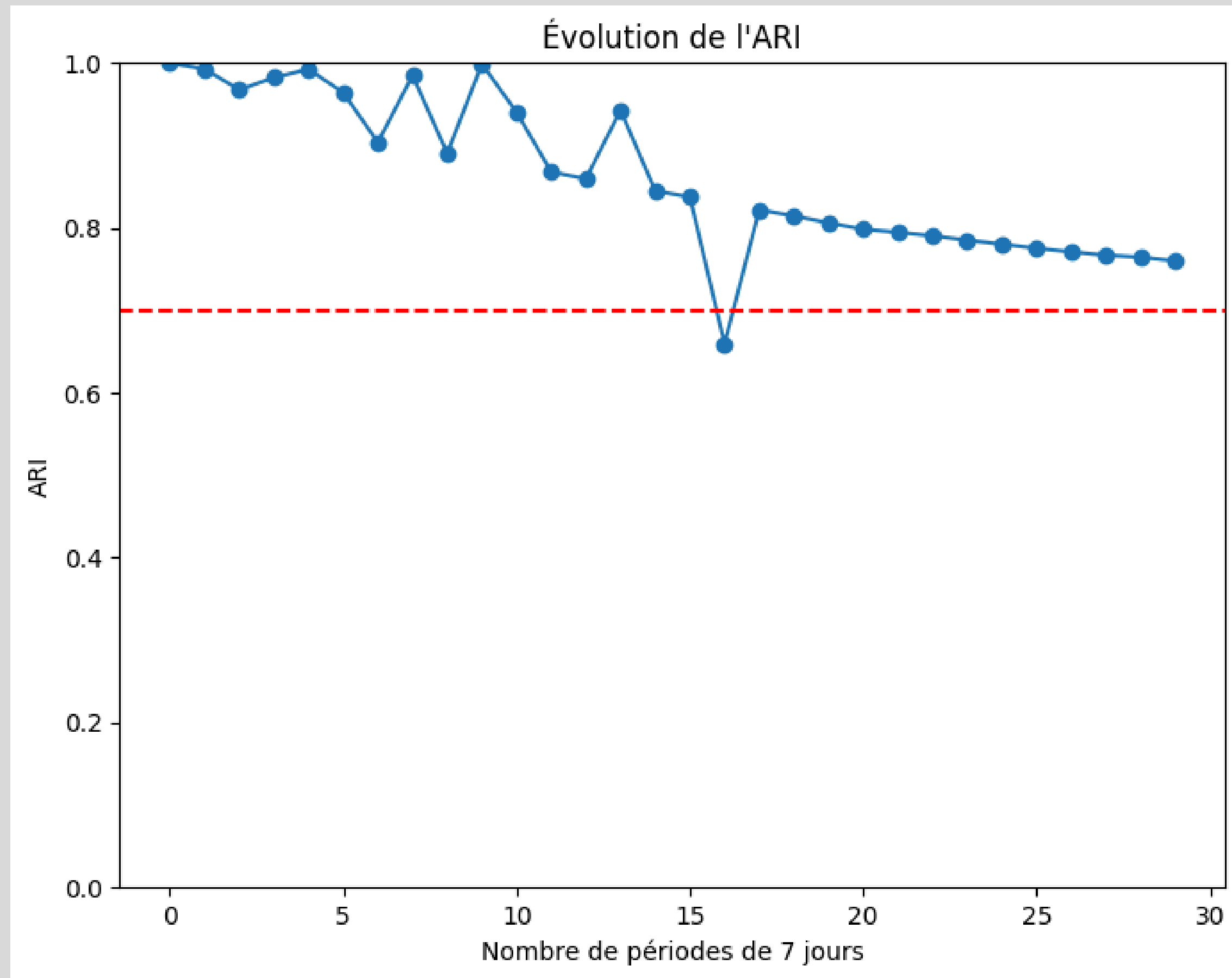


9. TESTER UN CLUSTERRING K-MEANS AVEC LES VARIABLE QUALITATIVES

```
Répartition des clusters  
2      18273  
5      18198  
1      17805  
0      16446  
3      14121  
4       9874  
Name: count, dtype: int64
```



ETUDE DE LA STABILITÉ TEMPORAIRE



CONCLUSION

- **SEGMENTATION IDENTIFIÉE**
- **IMPACT SUR L'ÉQUIPE MARKETING**
- **SUIVI DE LA SATISFACTION CLIENT**
- **FRÉQUENCE DE MISE À JOUR**