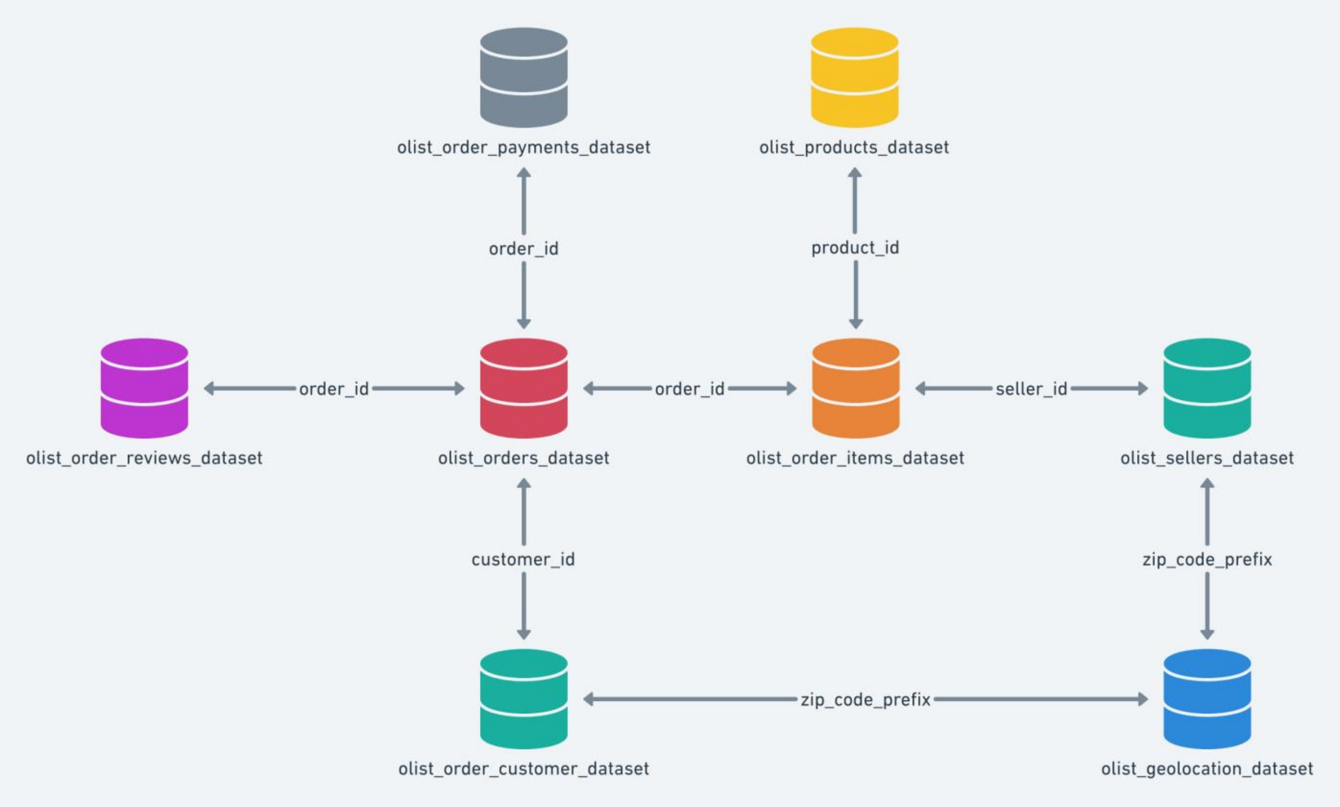
ETUDE CLIENTS OLIST

- Objectifs:
  - Comprendre les différents types de clients
  - Fournir une description actionnable pour l’équipe marketing
  - Proposition de contrat de maintenance

ANALYSE DES DONNÉES FOURNIES



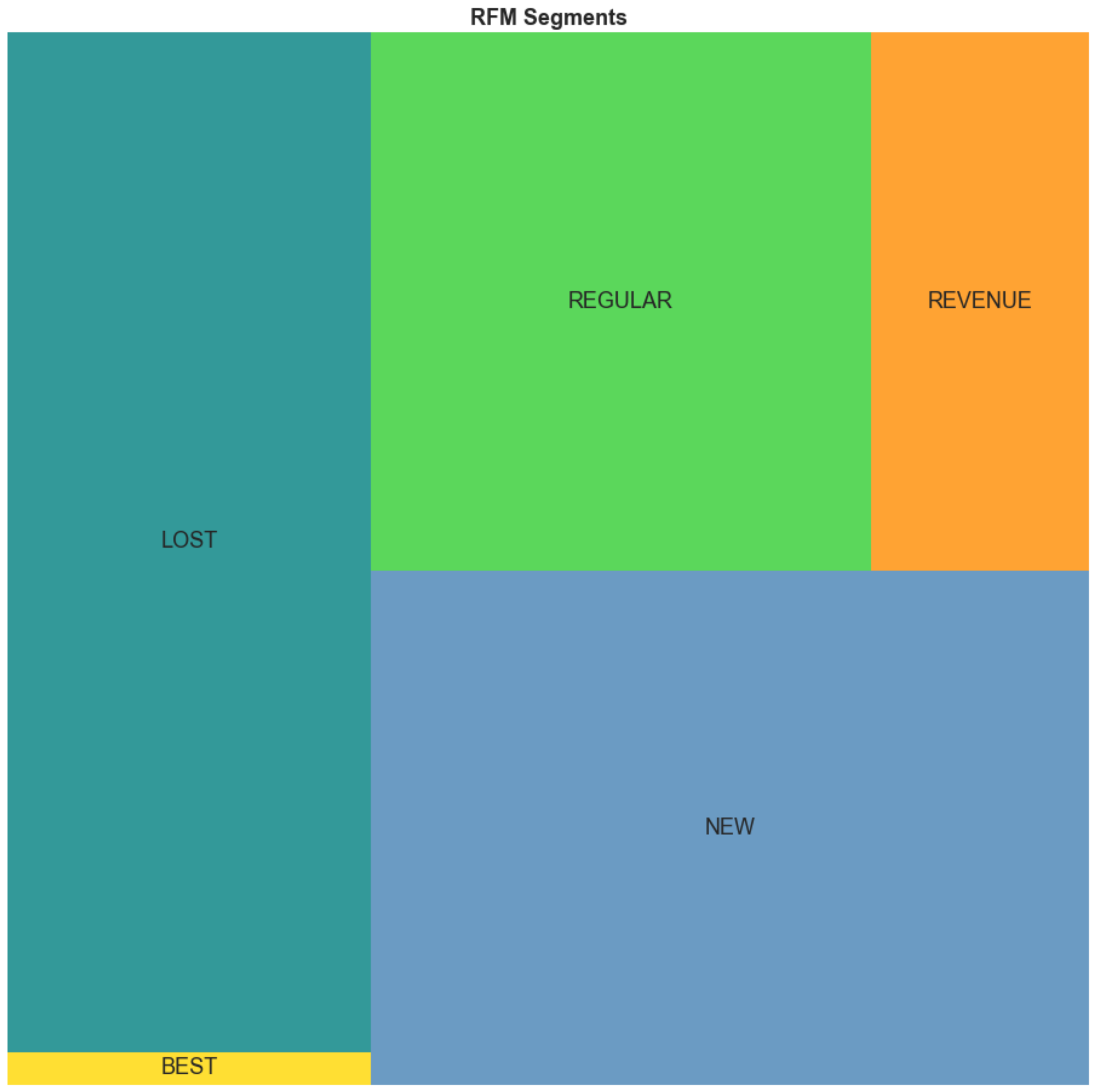
- Jonction des tables entre elles
- Nettoyage des données
- Feature engeneering

- Reste 32 variables portant sur 94925 transactions

SEGMENTATION SUPERVISÉE

- Création de 3 variables:
  - Recency: nombre de jours depuis la date de la dernière transaction
  - Frequence: nombre de fois où le client à commandé
  - Monetary: somme totale dépensée sur la période étudiée
- Attribution de notes
- Regroupement en types de comportements semblable

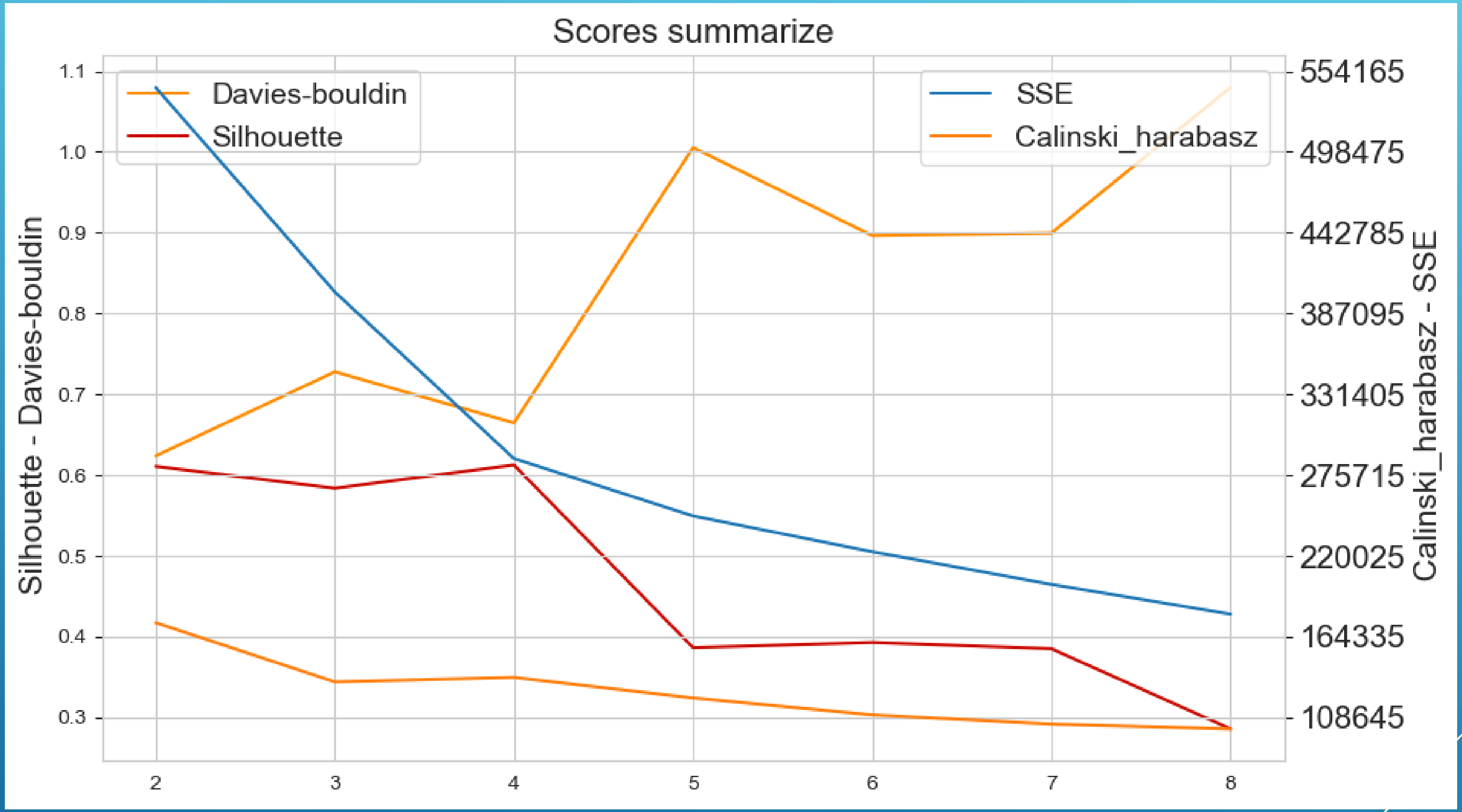
segments	RecencyMean	FrequencyMean	MonetaryMean	GroupSize
Lost	153.621776	1.000000	164.583805	12445
New	30.482447	1.000000	161.069825	12505
Regular	97.187439	1.103630	103.144367	9283
Revenues	94.825771	1.000000	349.283537	3989
Valuables	30.959900	2.157895	333.162531	399



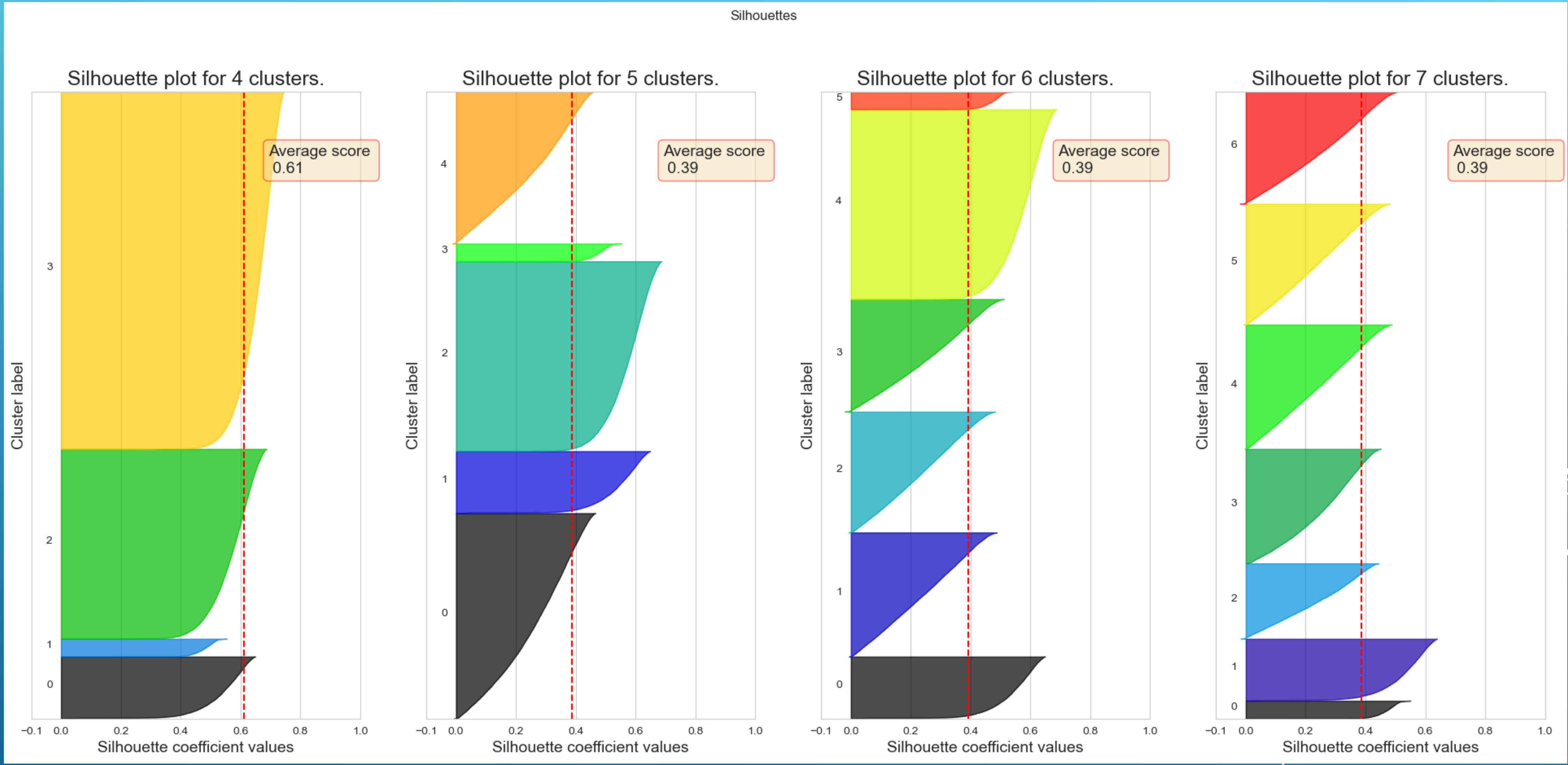
SEGMENTATION NON SUPERVISÉE

- Process
  - Agrégation des données
    - Recency
    - Frequency
    - Monetary
    - Review_score
    - Delivery
  - Etudes de la qualité et de l'interprétabilité du clustering
  - Sur le partitionnement retenu
    - Etude du client type pour chaque segment
    - Actions marketing envisageables

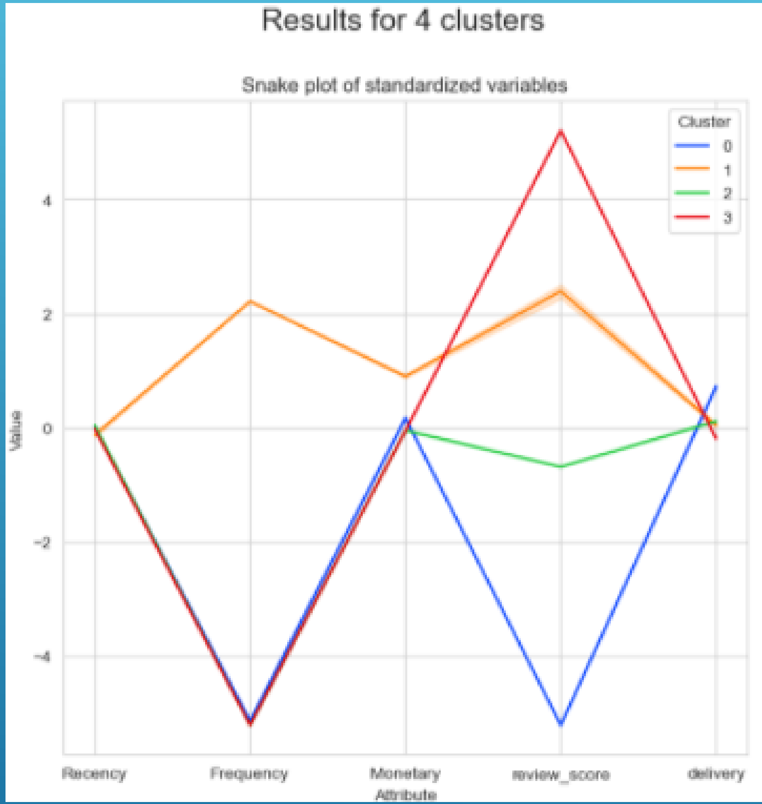
DÉTERMINATION DU NOMBRE OPTIMAL DE CLUSTERS

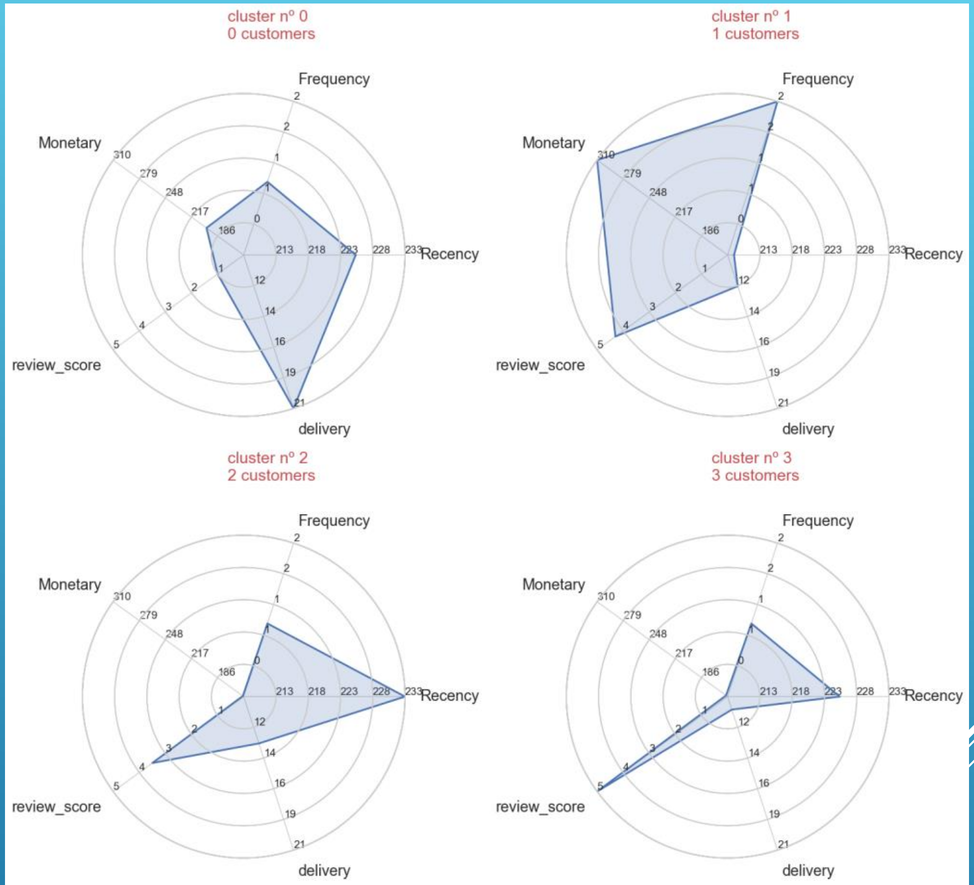


DÉTERMINATION DU NOMBRE OPTIMAL DE CLUSTERS

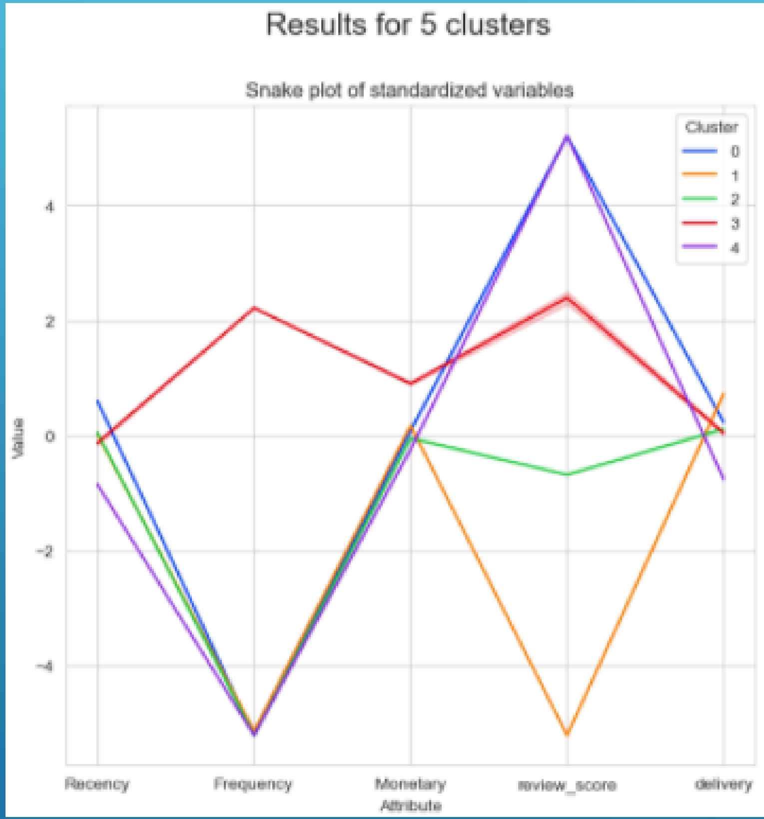


4 Segments

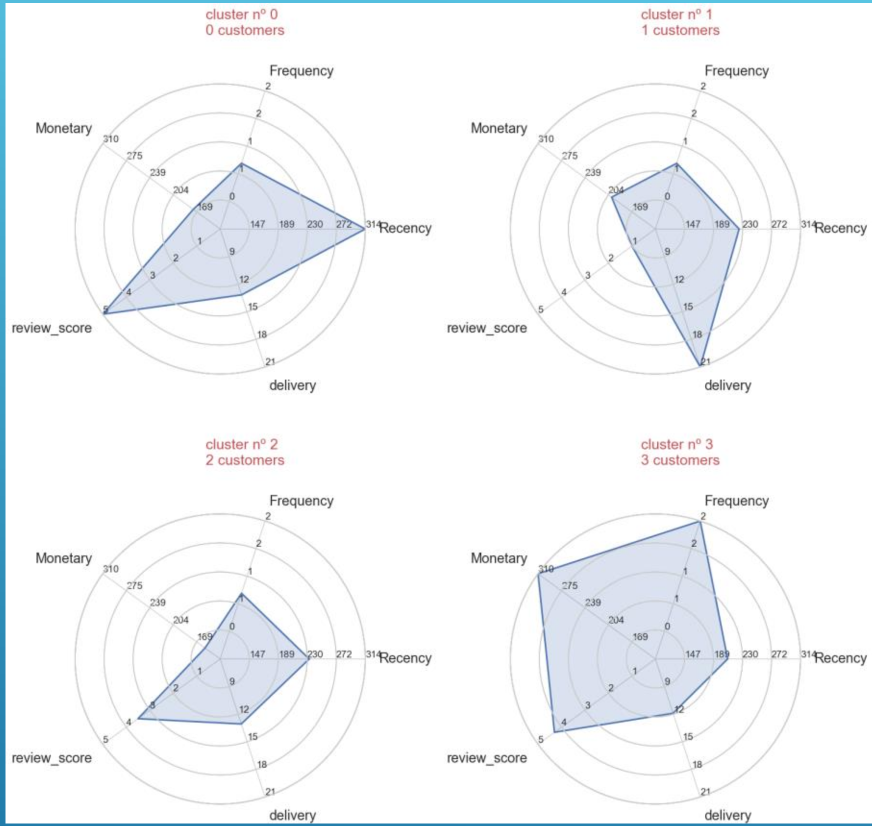

Results for 4 clusters
Snake plot of standardized variables
Cluster
0
1
2
3
Value
Recency
Frequency
Monetary
review_score
delivery
Attribute


cluster nº 0
0 customers
cluster nº 1
1 customers
cluster nº 2
2 customers
cluster nº 3
3 customers
Frequency
Monetary
Recency
review_score
delivery

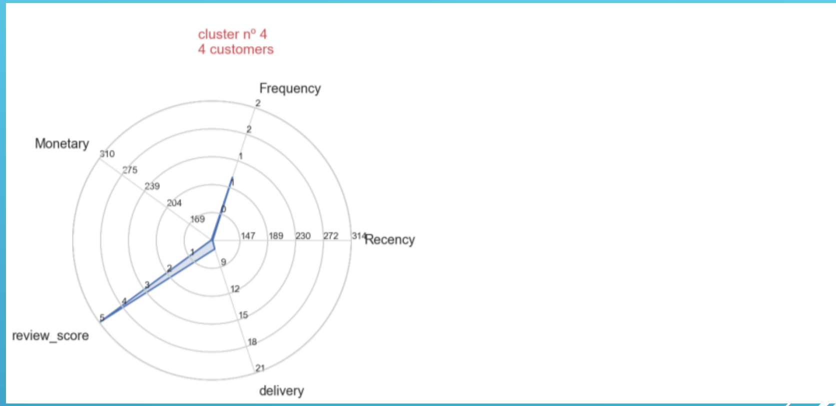
5 Segments



Results for 5 clusters
Snake plot of standardized variables
Cluster
0
1
2
3
4
Value
Recency
Frequency
Monetary
review_score
delivery
Attribute

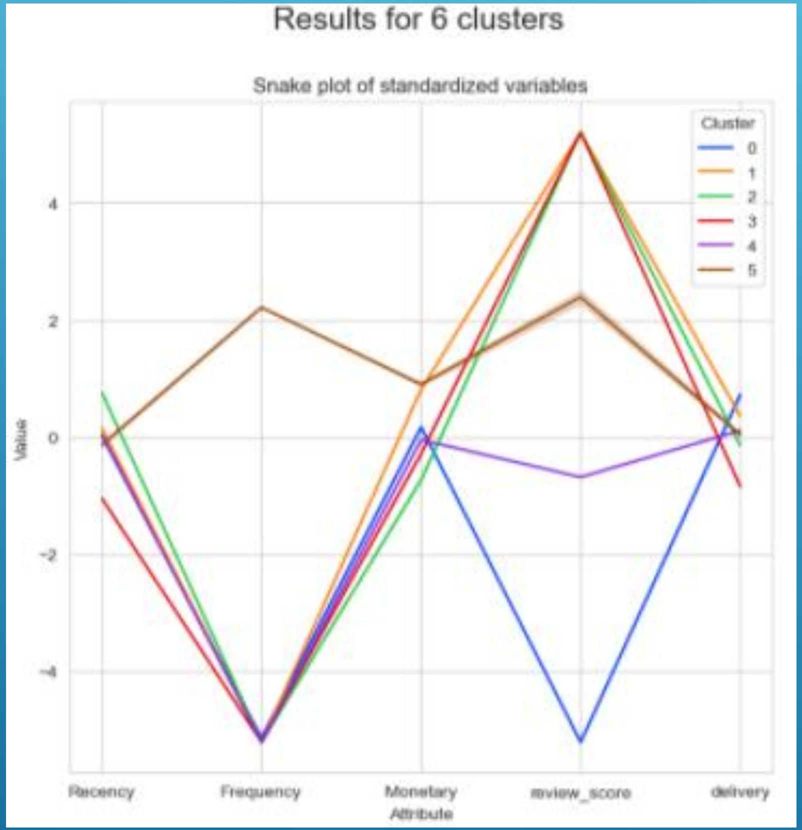


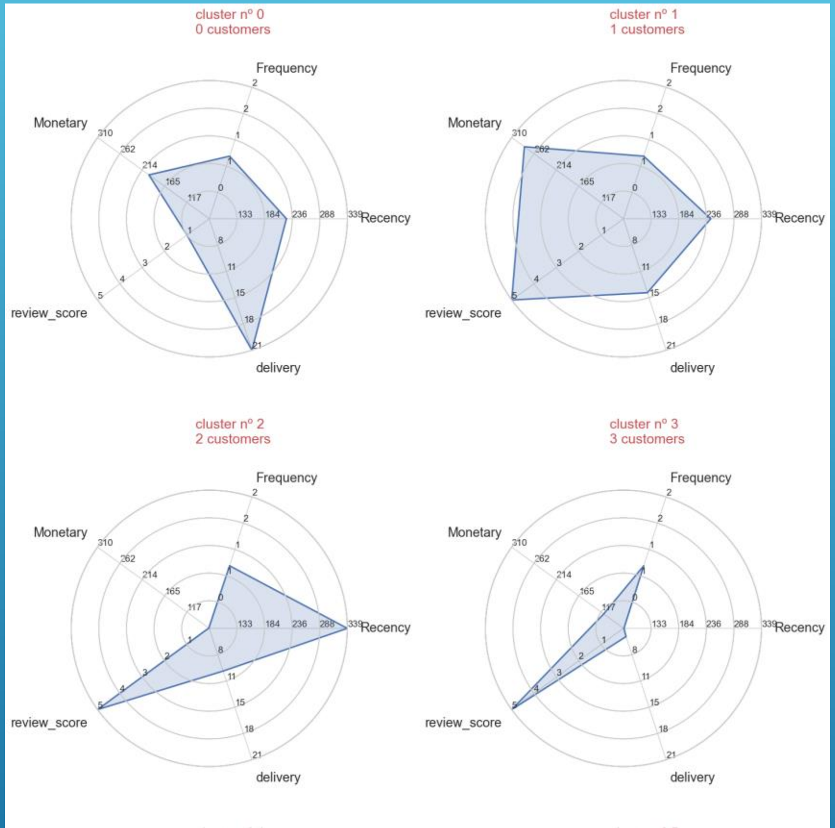
cluster nº 0
0 customers
cluster nº 1
1 customers
cluster nº 2
2 customers
cluster nº 3
3 customers
Frequency
Monetary
Recency
review_score
delivery

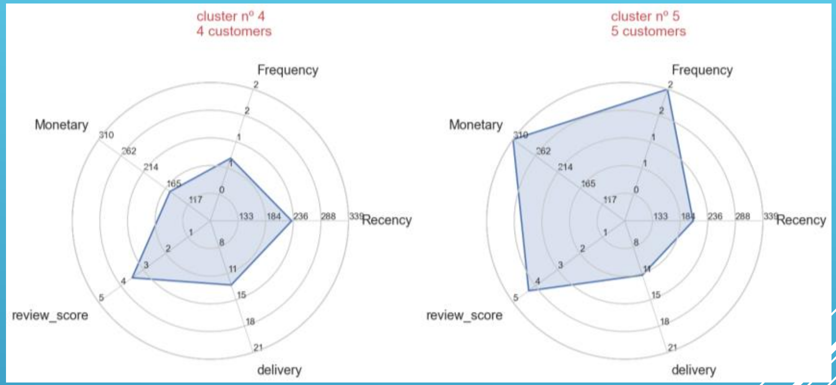


cluster nº 4
4 customers
Frequency
Monetary
Recency
review_score
delivery

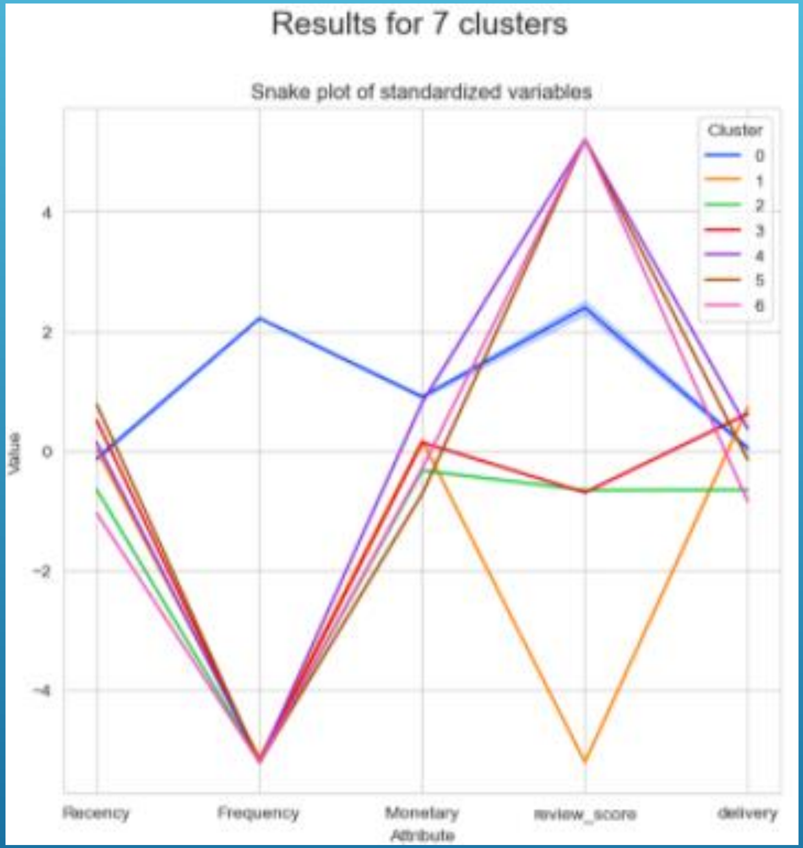
6 Segments

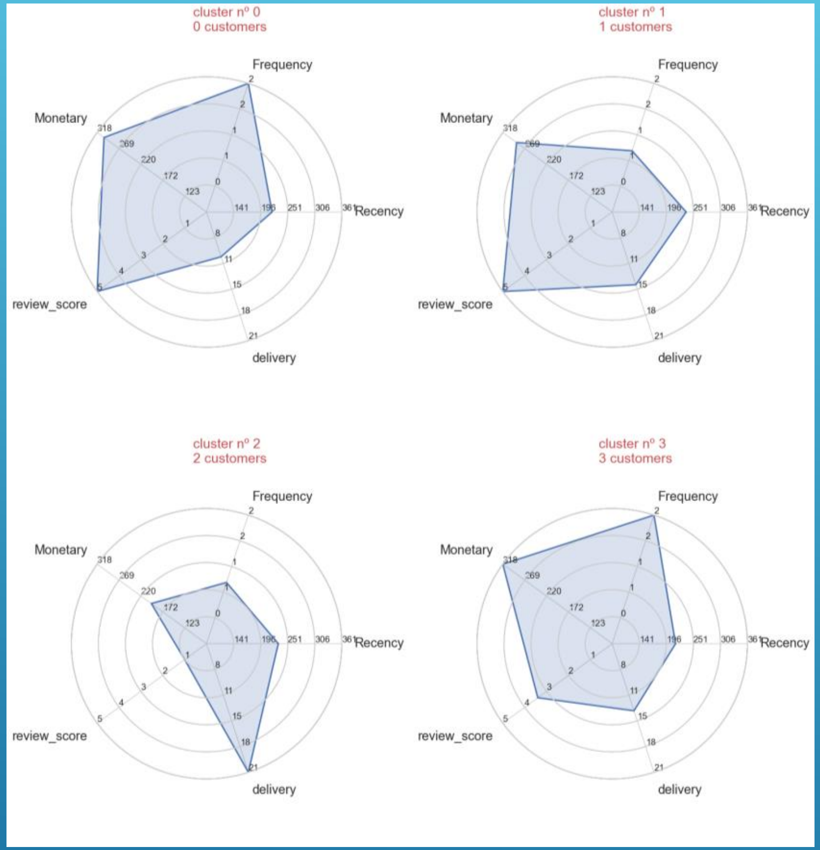

Results for 6 clusters
Snake plot of standardized variables
Cluster
0
1
2
3
4
5
Value
Recency
Frequency
Monetary
review_score
delivery
Attribute

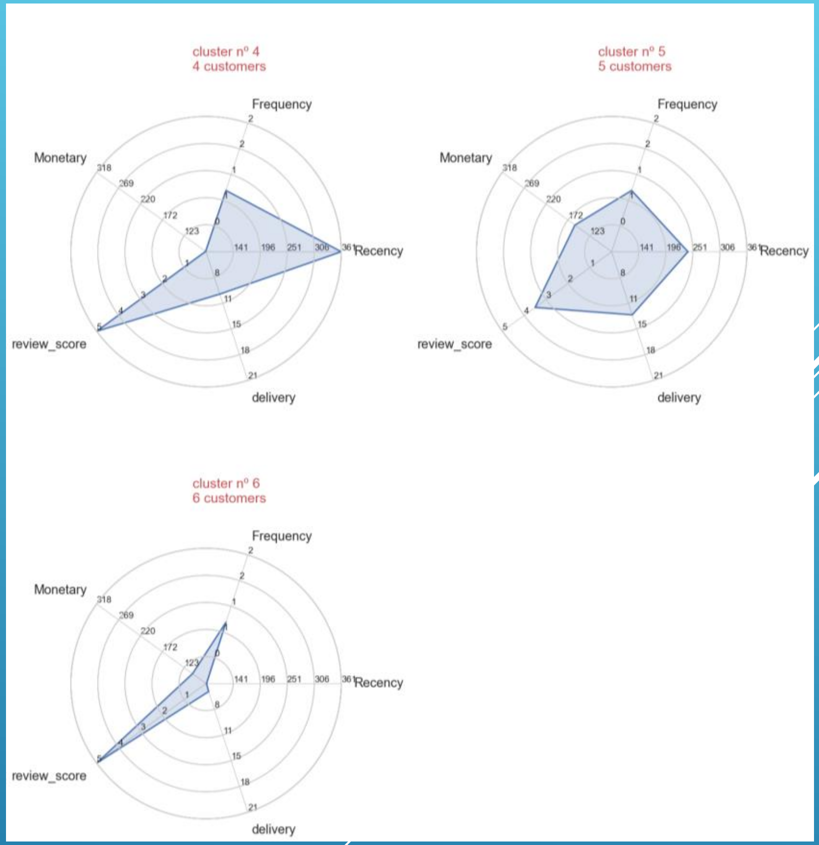

cluster n° 0
0 customers
cluster n° 1
1 customers
cluster n° 2
2 customers
cluster n° 3
3 customers
Frequency
Monetary
Recency
review_score
delivery


cluster n° 4
4 customers
cluster n° 5
5 customers
Frequency
Monetary
Recency
review_score
delivery

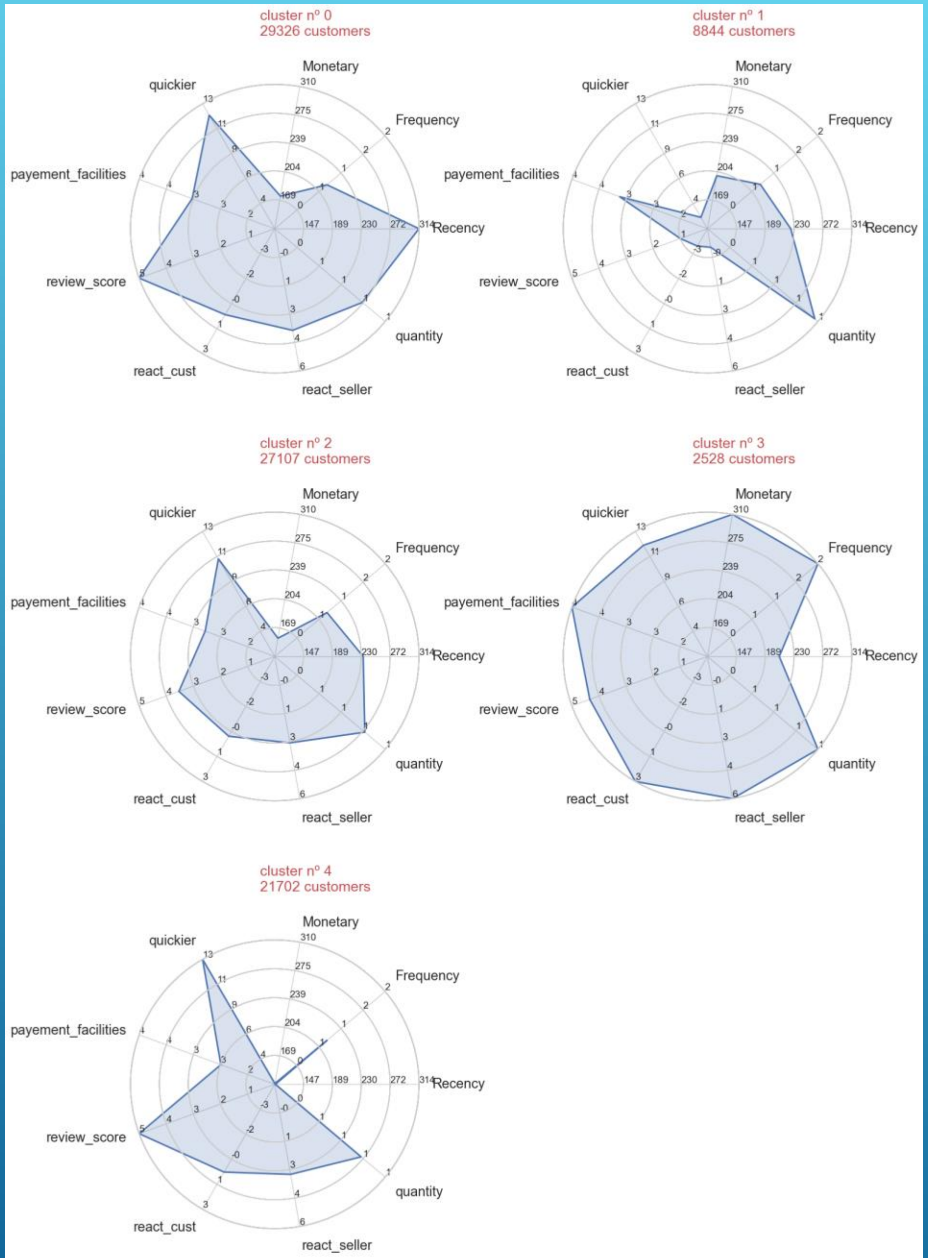
7 Segments


Results for 7 clusters
Snake plot of standardized variables
Cluster
0
1
2
3
4
5
6
Value
4
2
0
-2
-4
Recency
Frequency
Monetary
review_score
delivery
Attribute


cluster nº 0
0 customers
cluster nº 1
1 customers
cluster nº 2
2 customers
cluster nº 3
3 customers
Frequency
Monetary
Recency
review_score
delivery


cluster nº 4
4 customers
cluster nº 5
5 customers
cluster nº 6
6 customers
Frequency
Monetary
Recency
review_score
delivery

- Monetary: panier moyen
- Frequency : nombres de commandes
- Recency : nombre de jours écoulés depuis la dernière commande
- Quantity: nombre articles par commande
- react_seller: jours écoulés entre post de la reiew customer et post de la review seller
- react_cust: jours écoulés entre la réception de la commande et le post de la review
- review_score: note de satisfaction
- payement facilities: nombre de payements échelonnés
- quickier: nombre de jours d'avance de la livraison par rapport à la date estimée



Actions marketing pour 5 clusters

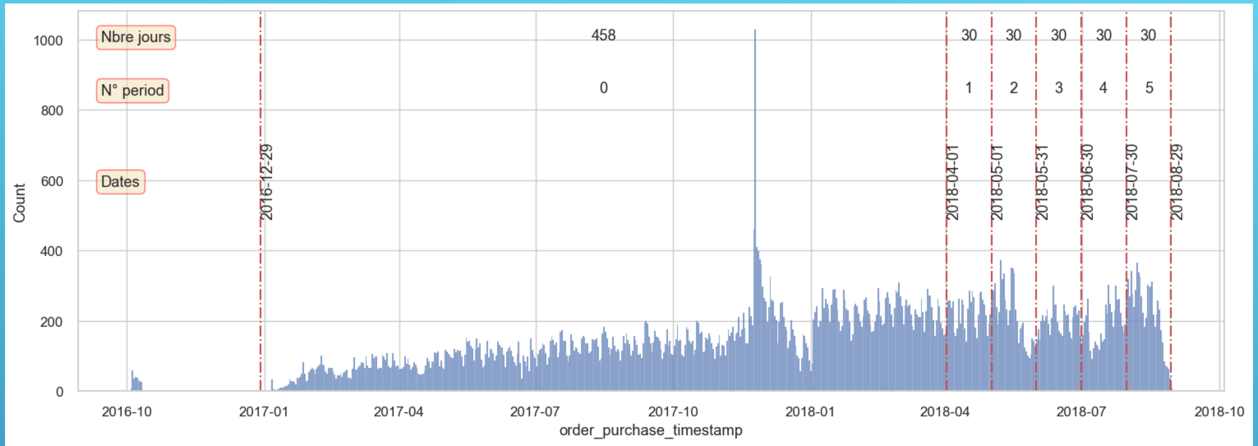
* Cluster 0: Clients perdus, on peut tenter de les relancer

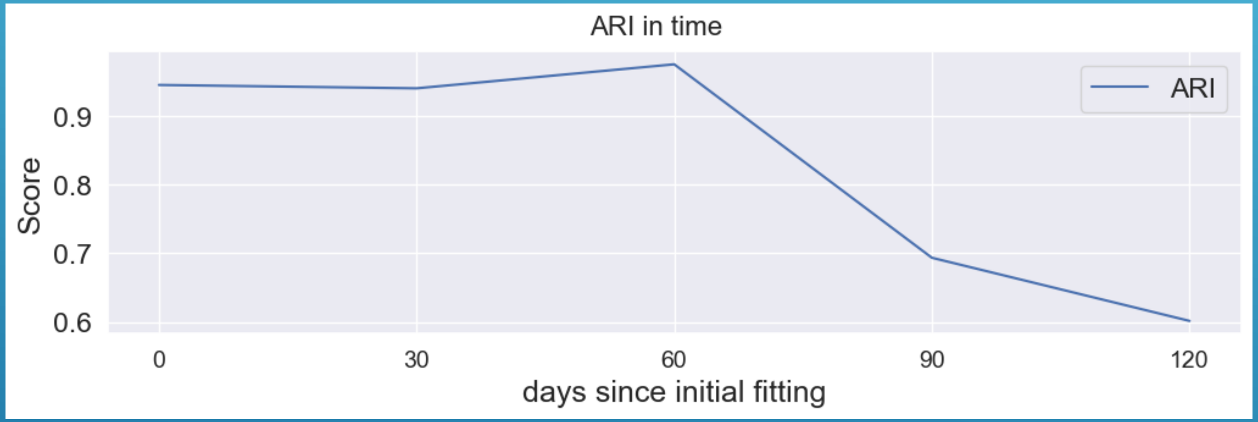
* Cluster 1: Clients ayant eu un problème avec la livraison à contacter pour enquête UX, proposer un service de livraison express/premium

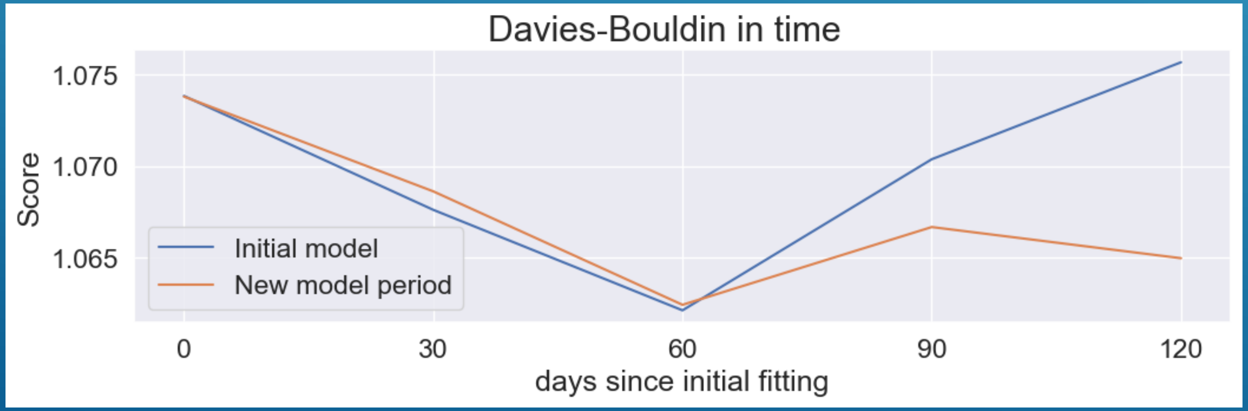
* Cluster 2: Plutôt déçus par la qualité, action marketing de type remboursement ou offre satisfait ou remboursé

* Cluster 3 : Ce sont les clients les plus importants, mais ils sont en perte de vitesse, il faut les relancer en leur offrant des bons d'achats

* Cluster 4: Nouveaux clients à fidéliser







Stabilité du modèle

Les courbes s'infléchissent à partir de 60 jours et les scores se dégradent après 90 jours.

Je suggère une maintenance tous les deux mois afin de pouvoir suivre au mieux l'évolution de la clientèle suite au actions marketing envisagées

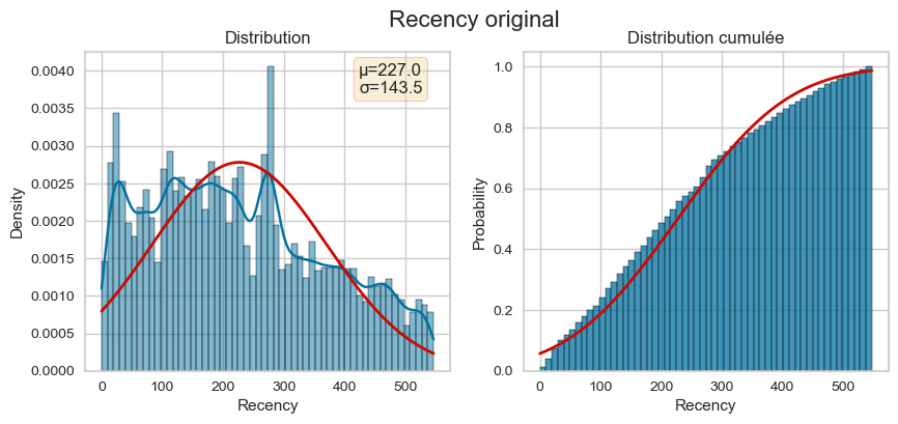
Conclusion

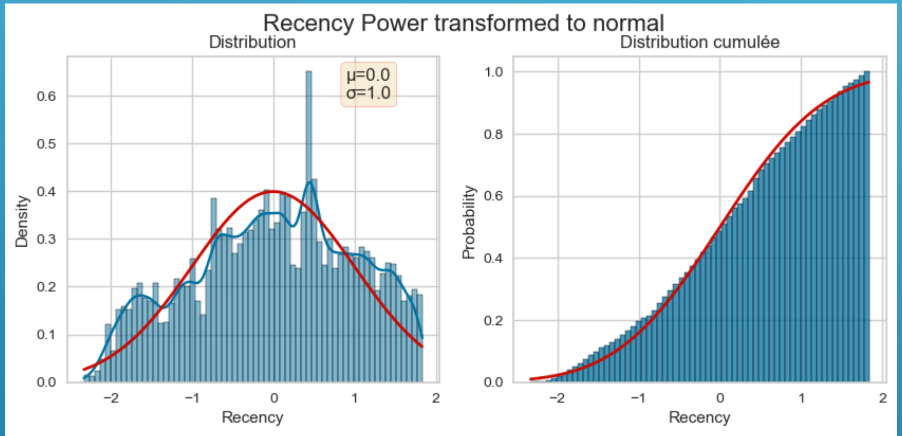
On a montré que la clientèle de Olist pouvait être partitionnée de manière non supervisée.

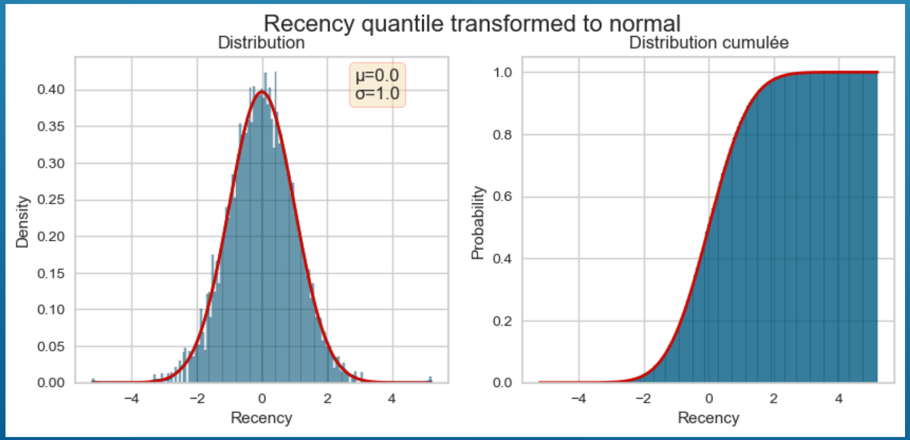
Que les profils des clients types permettent d'envisager des actions distinctes afin de mener une campagne marketing

Annexes

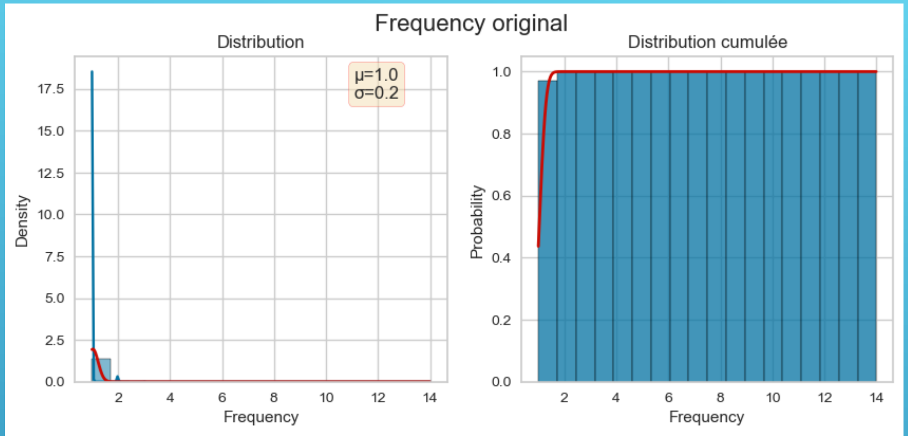
DISTRIBUTION DE RECENCY

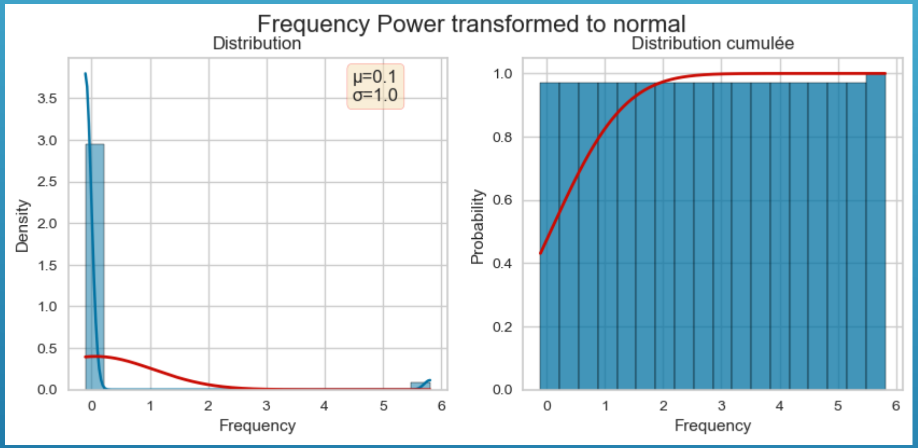


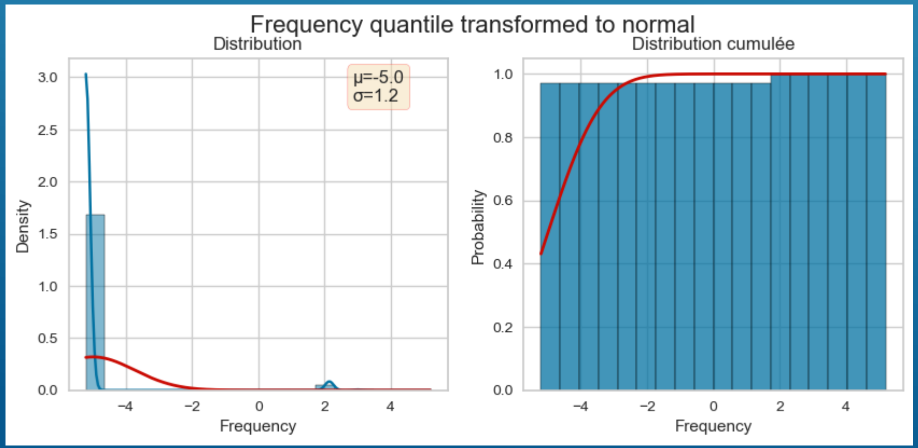




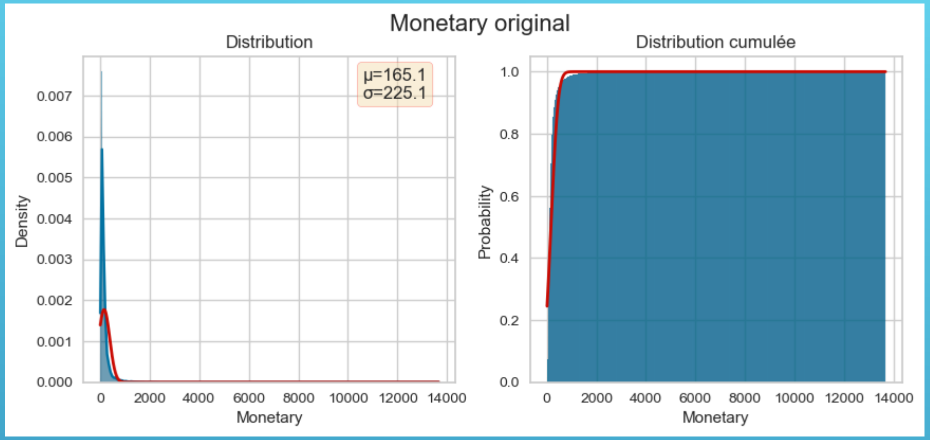
DISTRIBUTION DE FREQUENCY

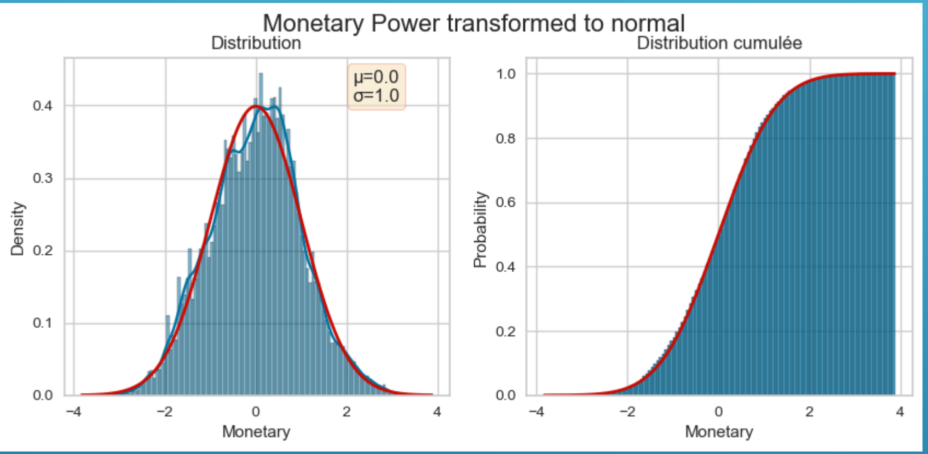


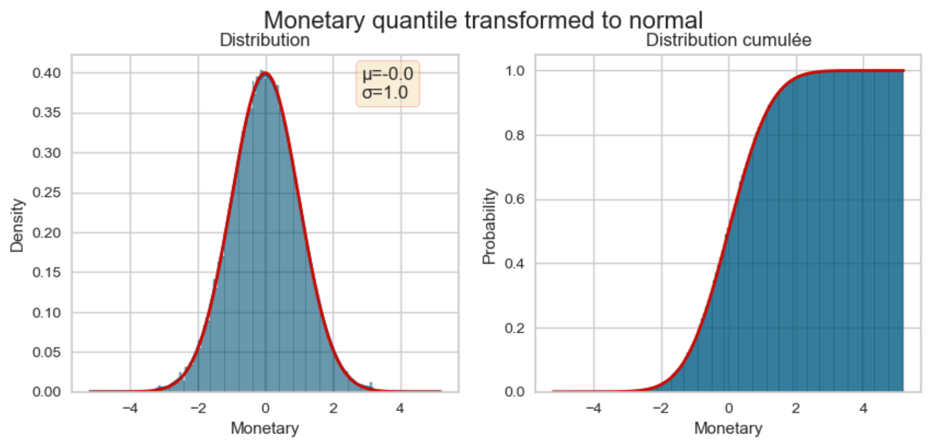


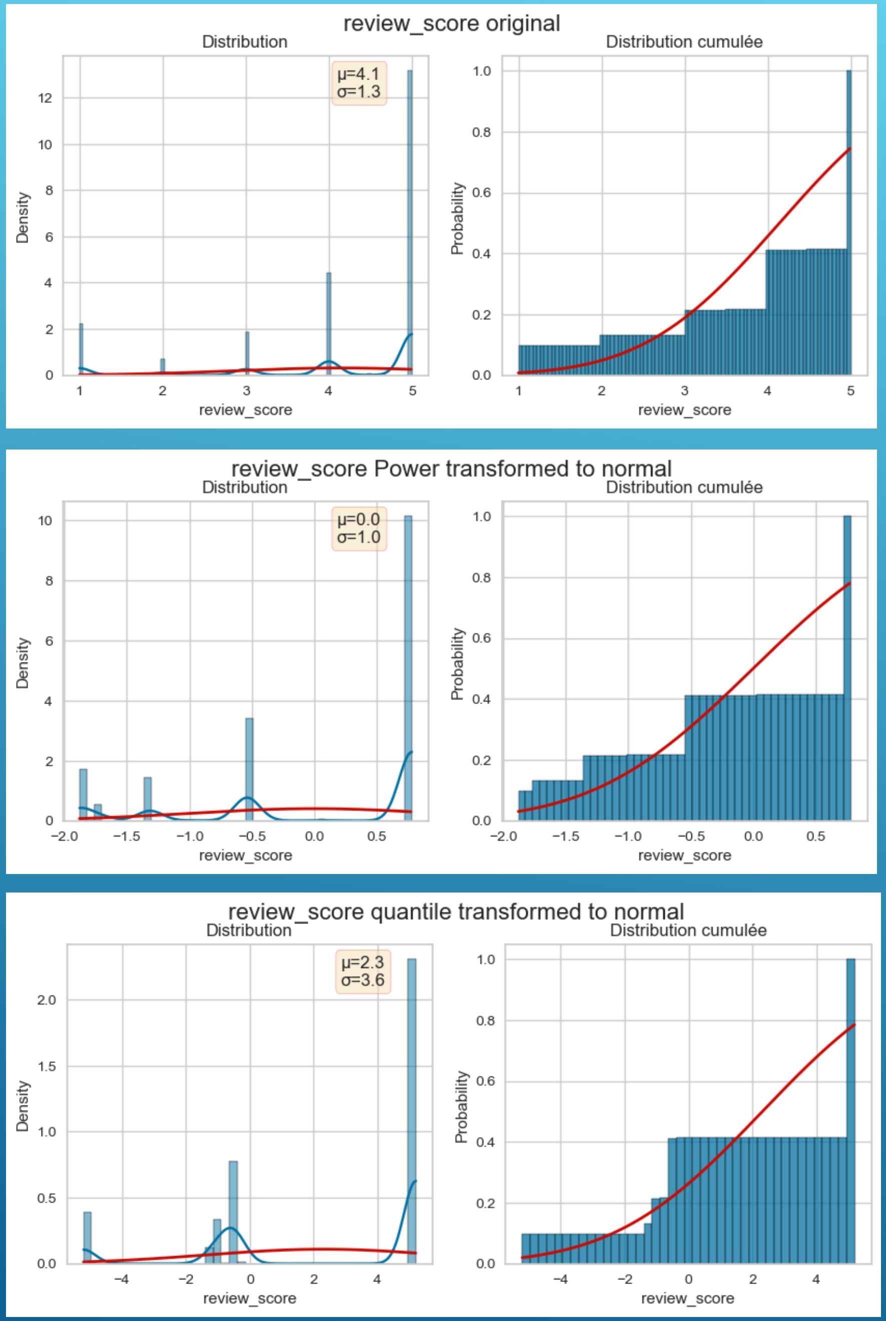


DISTRIBUTION DE MONETARY

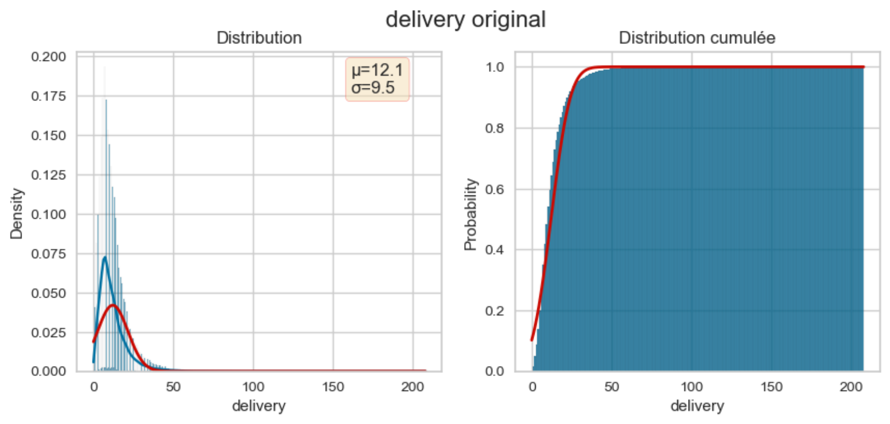


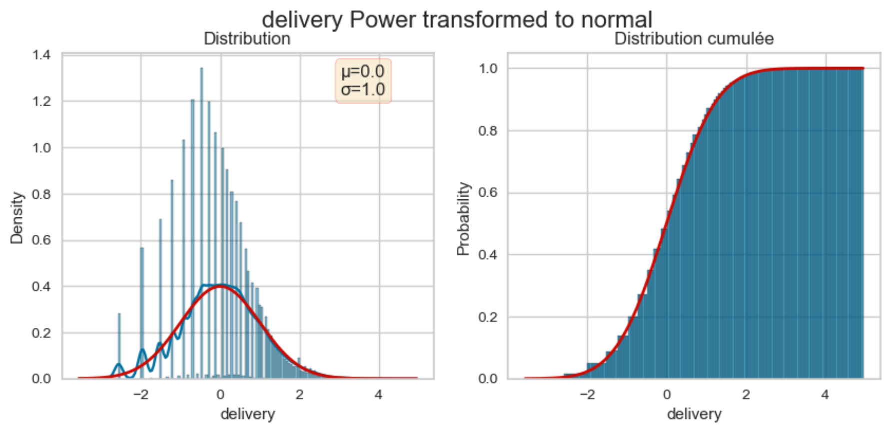


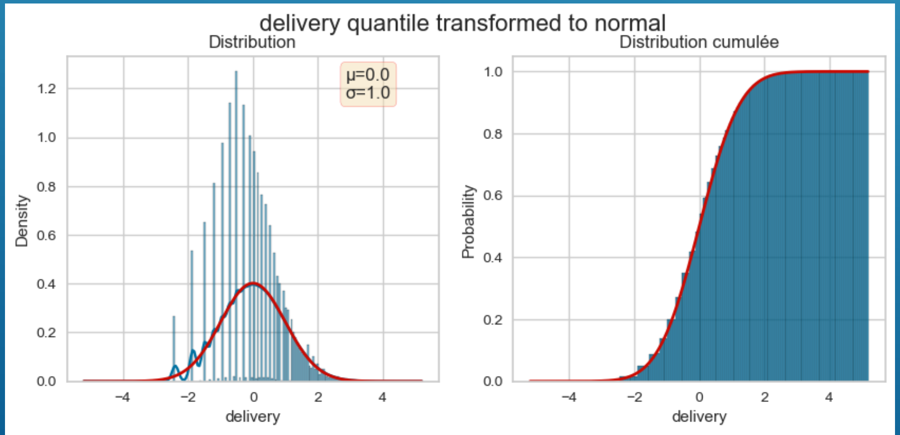




DISTRIBUTION DE REVIEW_SCORE





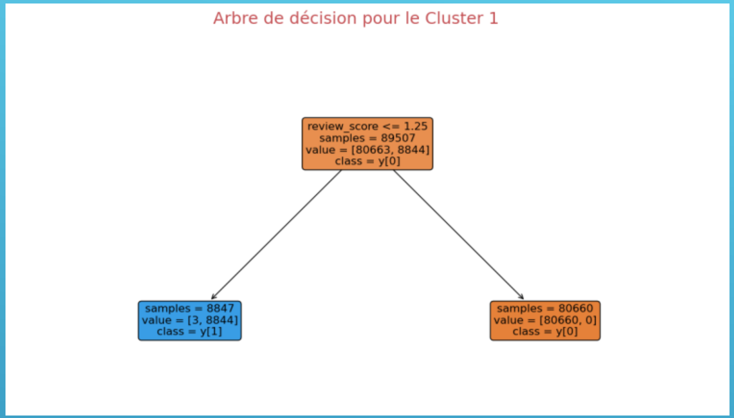


DISTRIBUTION DE DELIVERY

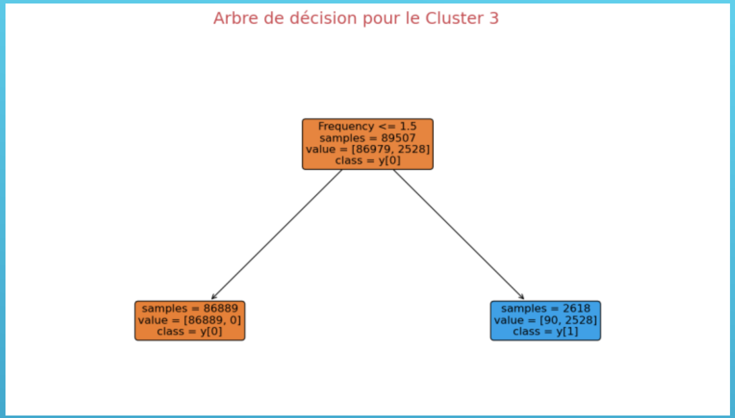
Règles de décision pour la constitution de 5 clusters

class_name	instance_count	rule_list
0	29326	[0.6590553644041288] (review_score > 4.099999904632568) and (Recency <= 158.5) and (delivery > 13.166666507720947) [0.7887961029923451] (review_score > 4.099999904632568) and (Recency > 158.5) and (delivery <= 5.75) and (Recency > 311.5) [0.9397736617457023] (review_score > 4.099999904632568) and (Recency > 158.5) and (delivery > 5.75) and (Frequency <= 1.5)
1	8844	[0.9996609020006781] (review_score <= 4.099999904632568) and (review_score <= 1.25)
2	27107	[1.0] (review_score <= 4.099999904632568) and (review_score > 1.25) and (Frequency <= 1.5)
3	2528	[1.0] (review_score <= 4.099999904632568) and (review_score > 1.25) and (Frequency > 1.5) [1.0] (review_score > 4.099999904632568) and (Recency <= 158.5) and (delivery <= 13.166666507720947) and (Frequency > 1.5) [1.0] (review_score > 4.099999904632568) and (Recency > 158.5) and (delivery > 5.75) and (Frequency > 1.5)
4	21702	[0.9597351467430207] (review_score > 4.099999904632568) and (Recency <= 158.5) and (delivery <= 13.166666507720947) and (Frequency <= 1.5) [0.8758998971546109] (review_score > 4.099999904632568) and (Recency > 158.5) and (delivery <= 5.75) and (Recency <= 311.5)

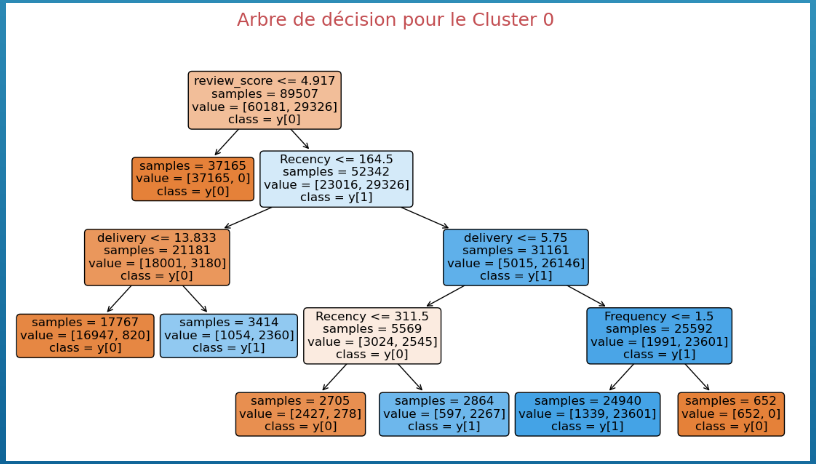
Arbres de décision pour la constitution de 5 clusters



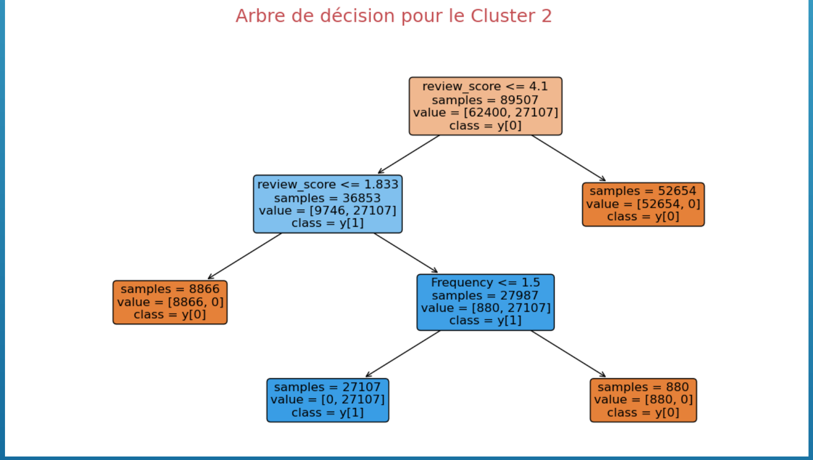
Arbre de décision pour le Cluster 1



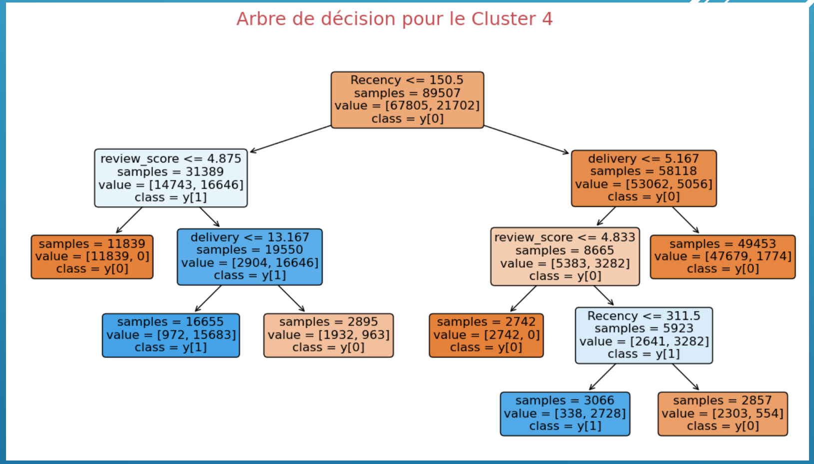
Arbre de décision pour le Cluster 3



Arbre de décision pour le Cluster 0

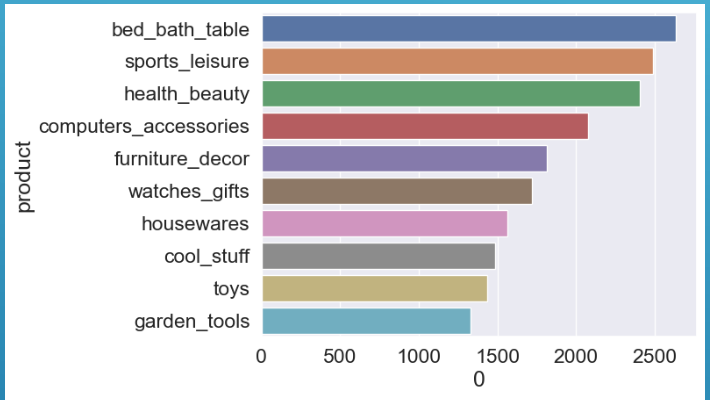


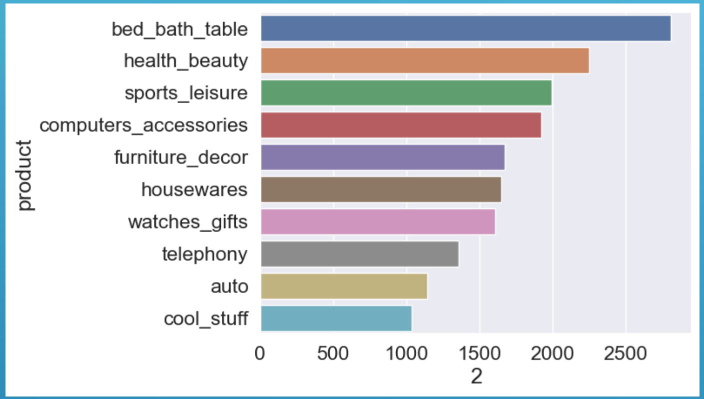
Arbre de décision pour le Cluster 2

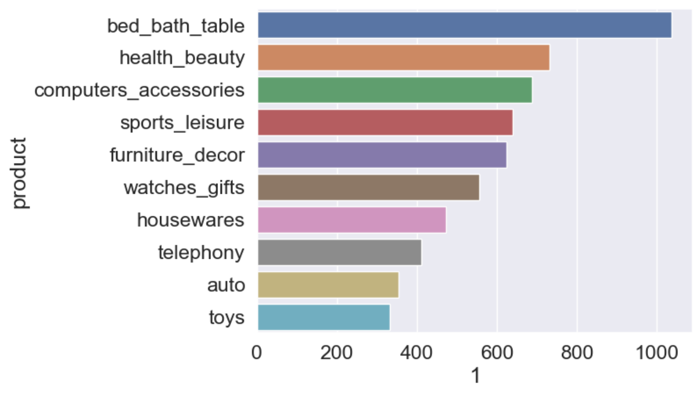


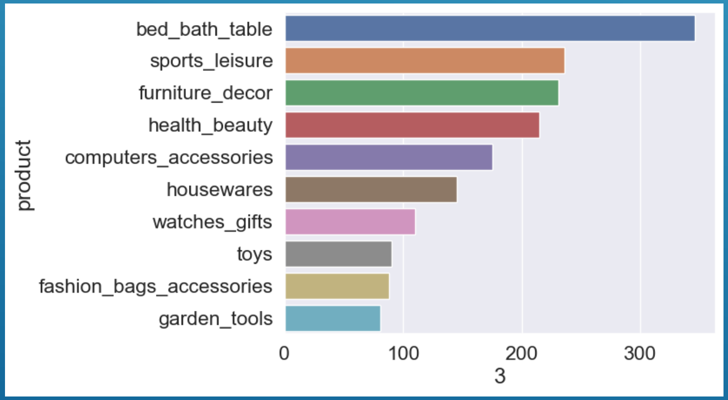
Arbre de décision pour le Cluster 4

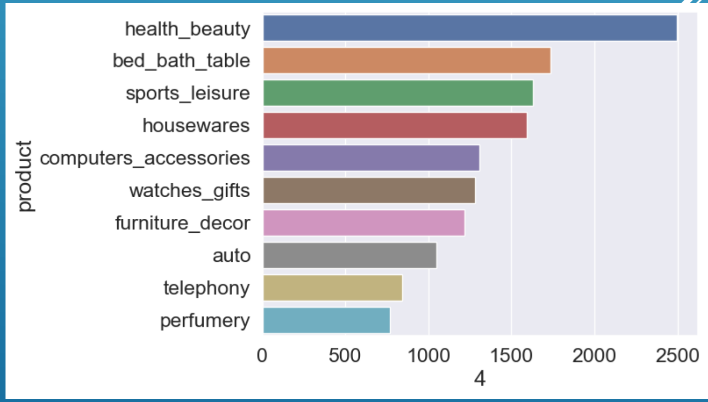
Catégories à mettre en avant pour les actions commerciales en fonction du segment visé





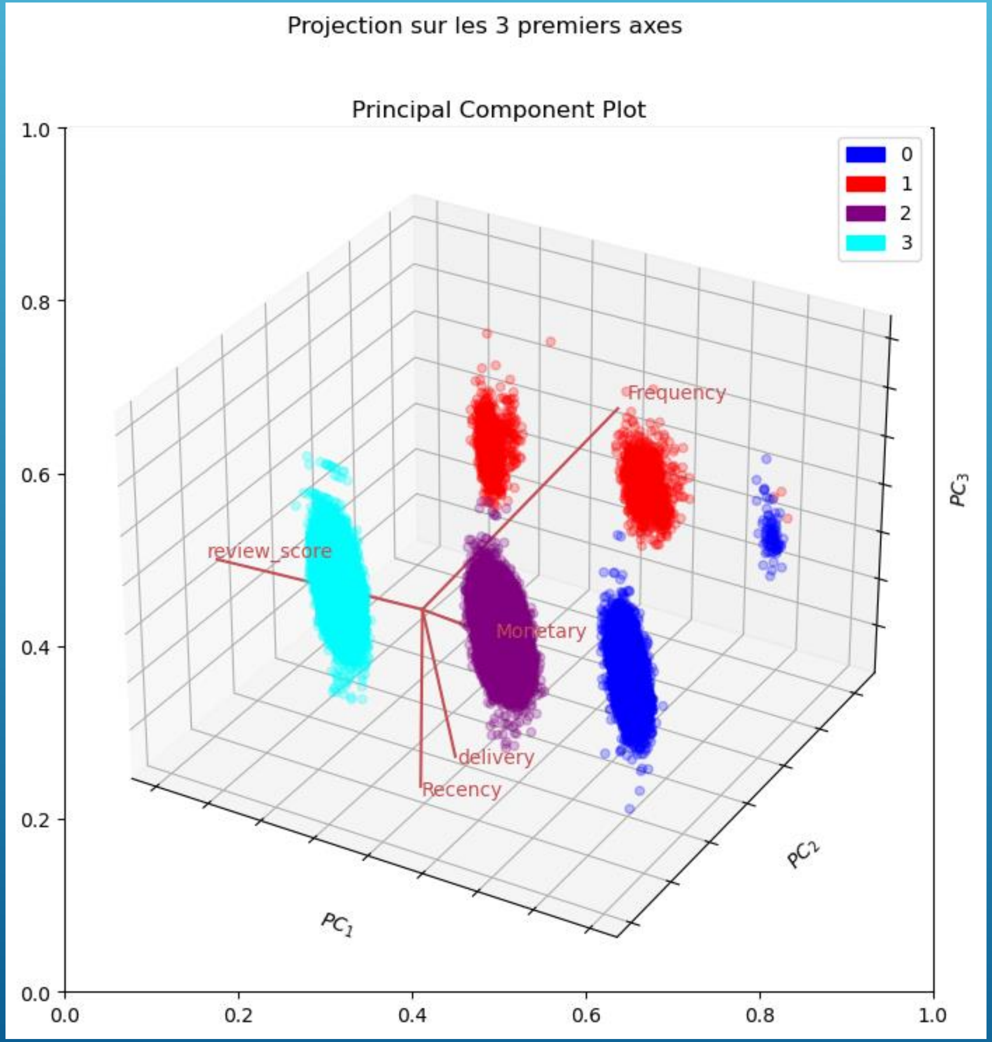




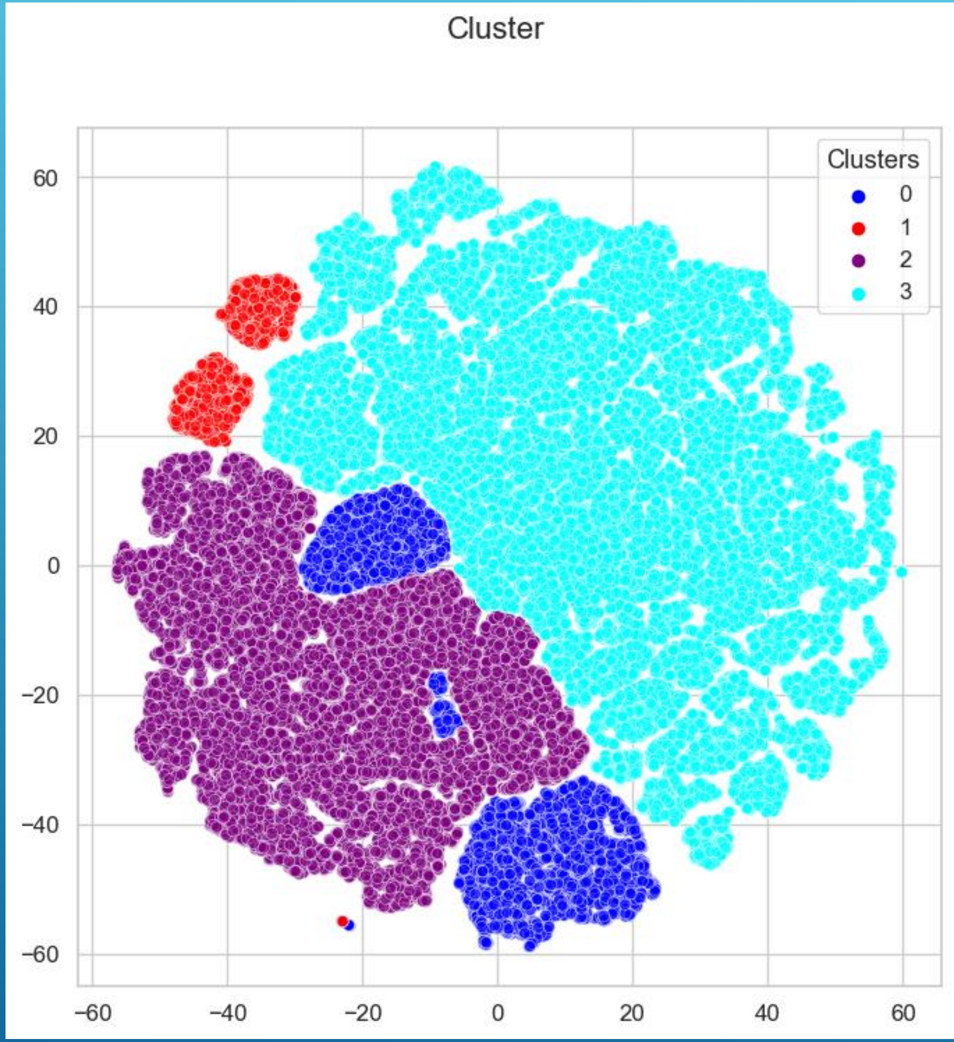


Visualisations 4 clusters

Réduction ACP

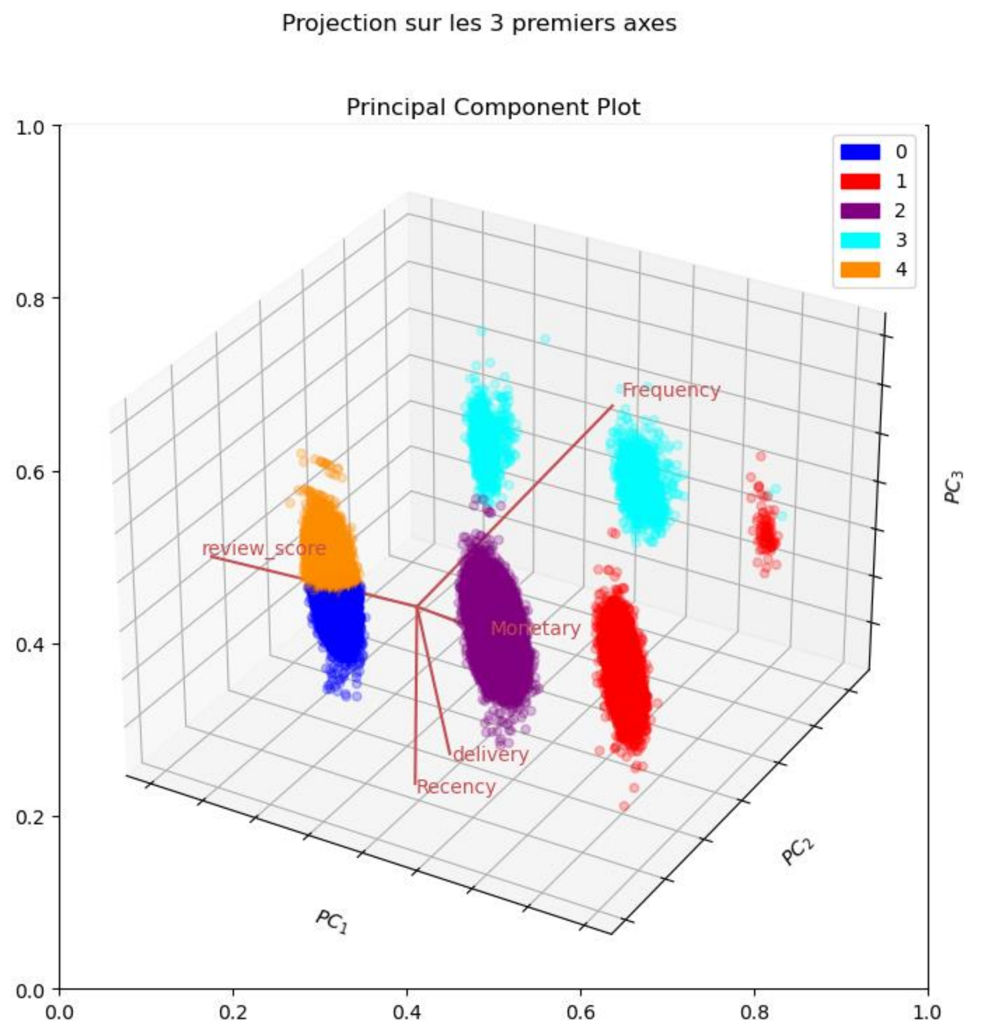


Réduction T-SNE

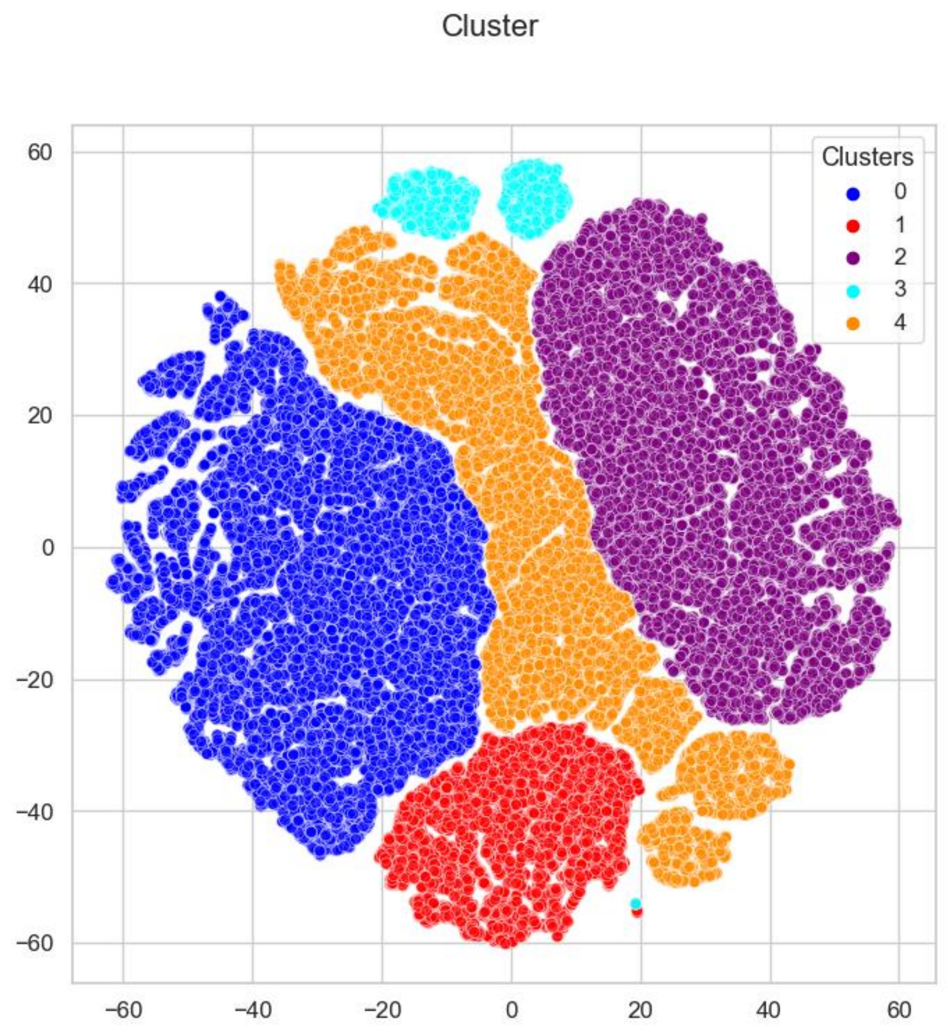


Visualisations 5 clusters

Réduction ACP

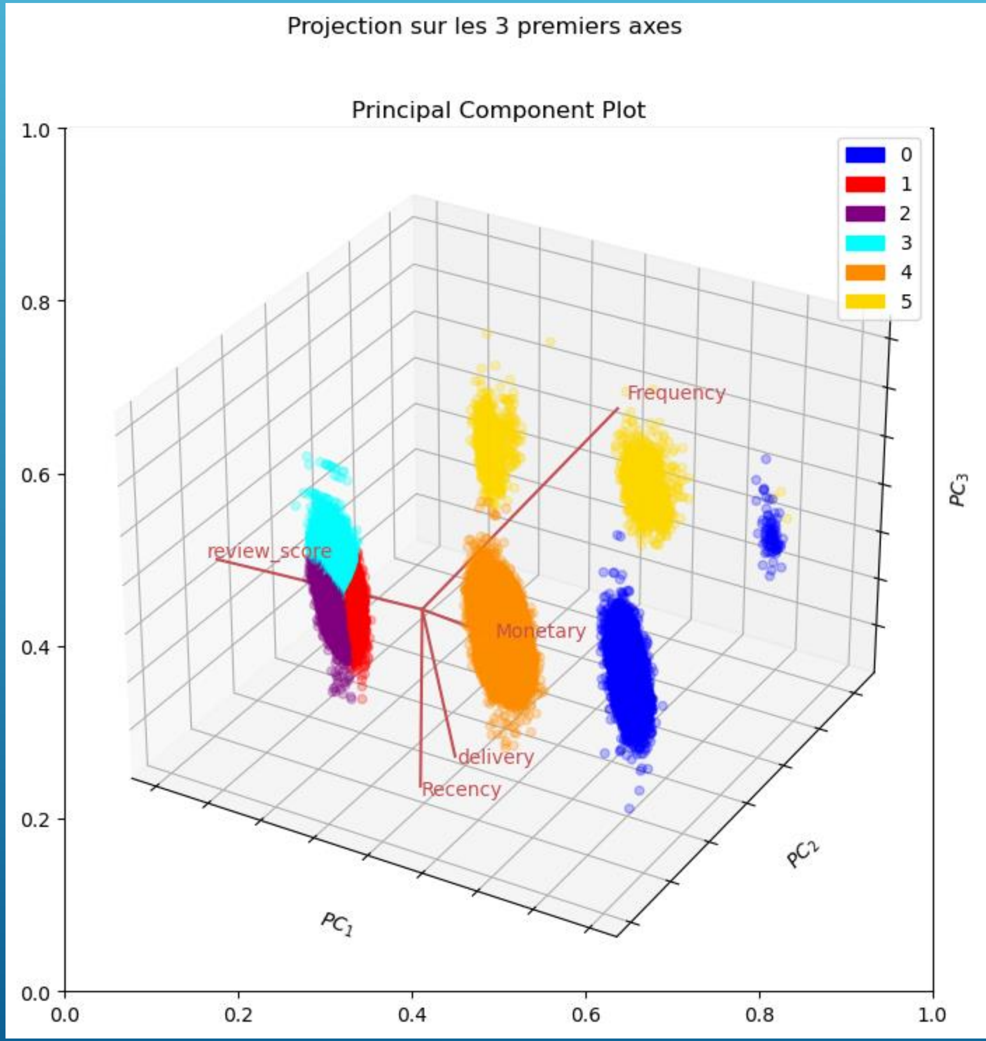


Réduction T-SNE

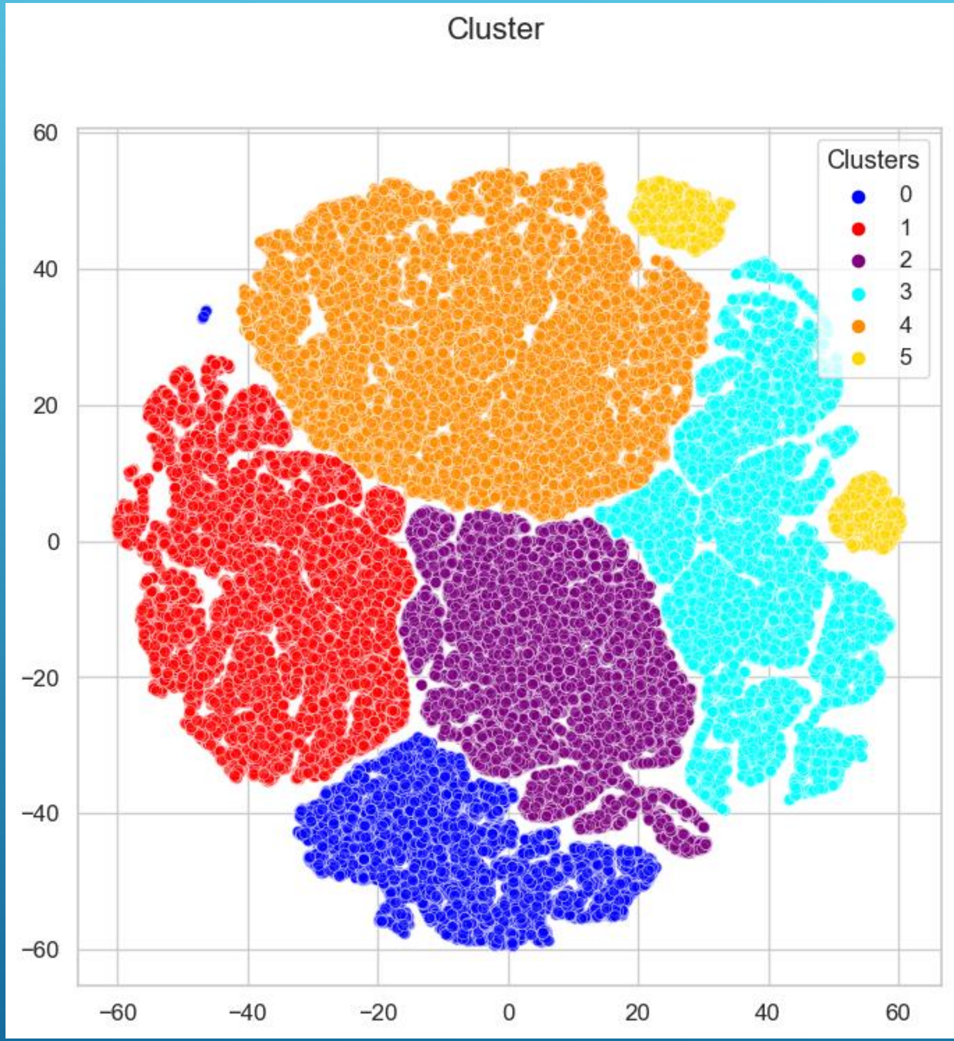


Visualisations 6 clusters

Réduction ACP

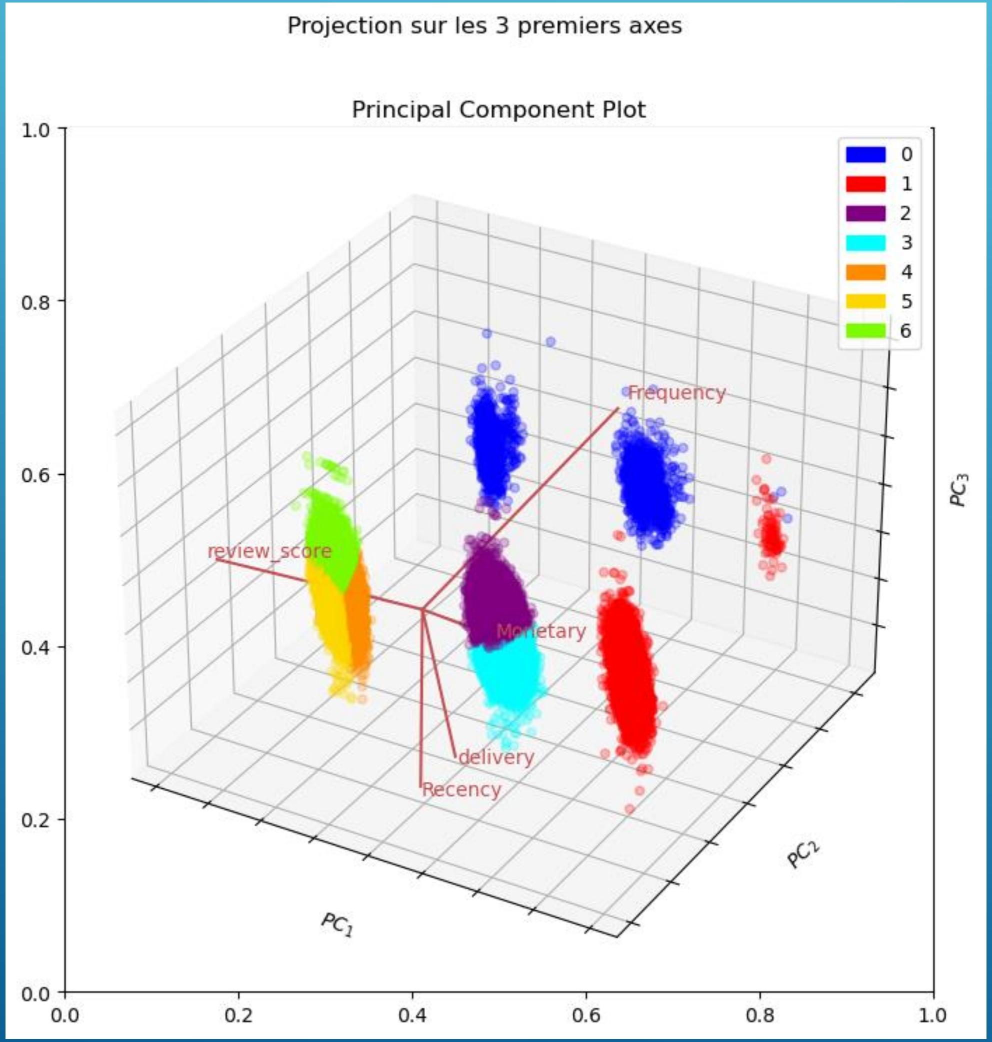


Réduction T-SNE



Visualisations 7 clusters

Réduction ACP



Réduction T-SNE