

Segmentez des clients d'un site e-commerce

Soutenance du 26 juin 2023

Yann Pham-Van

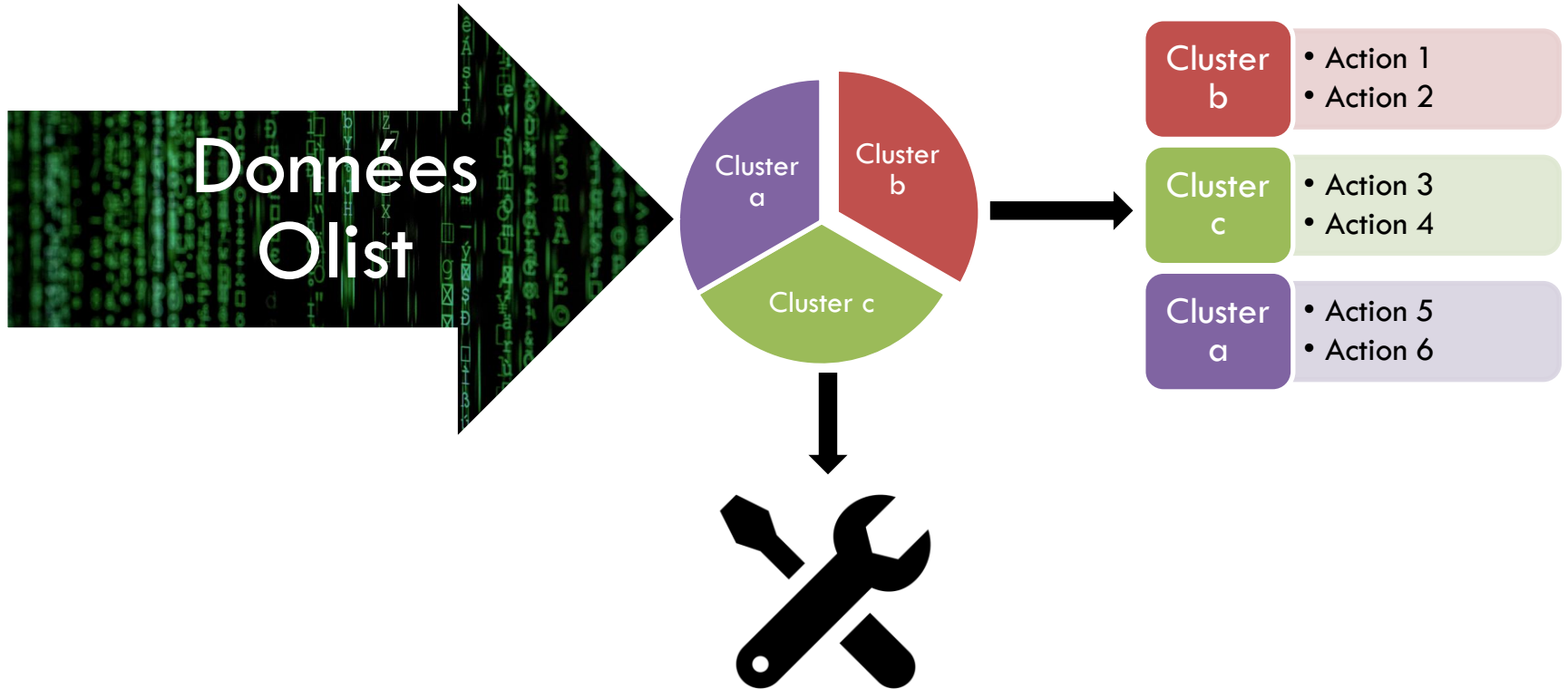
Segmentez des clients d'un site e-commerce

- Problématique
- Exploration
- Essais de modélisation
- Modèle final
- Maintenance du modèle
- Propositions

Segmentez des clients d'un site e-commerce

- **Problématique**
- Exploration
- Essais de modélisation
- Modèle final
- Maintenance du modèle
- Propositions

Problématique



Segmentez des clients d'un site e-commerce

- Problématique
- **Exploration**
- Essais de modélisation
- Modèle final
- Maintenance du modèle
- Propositions

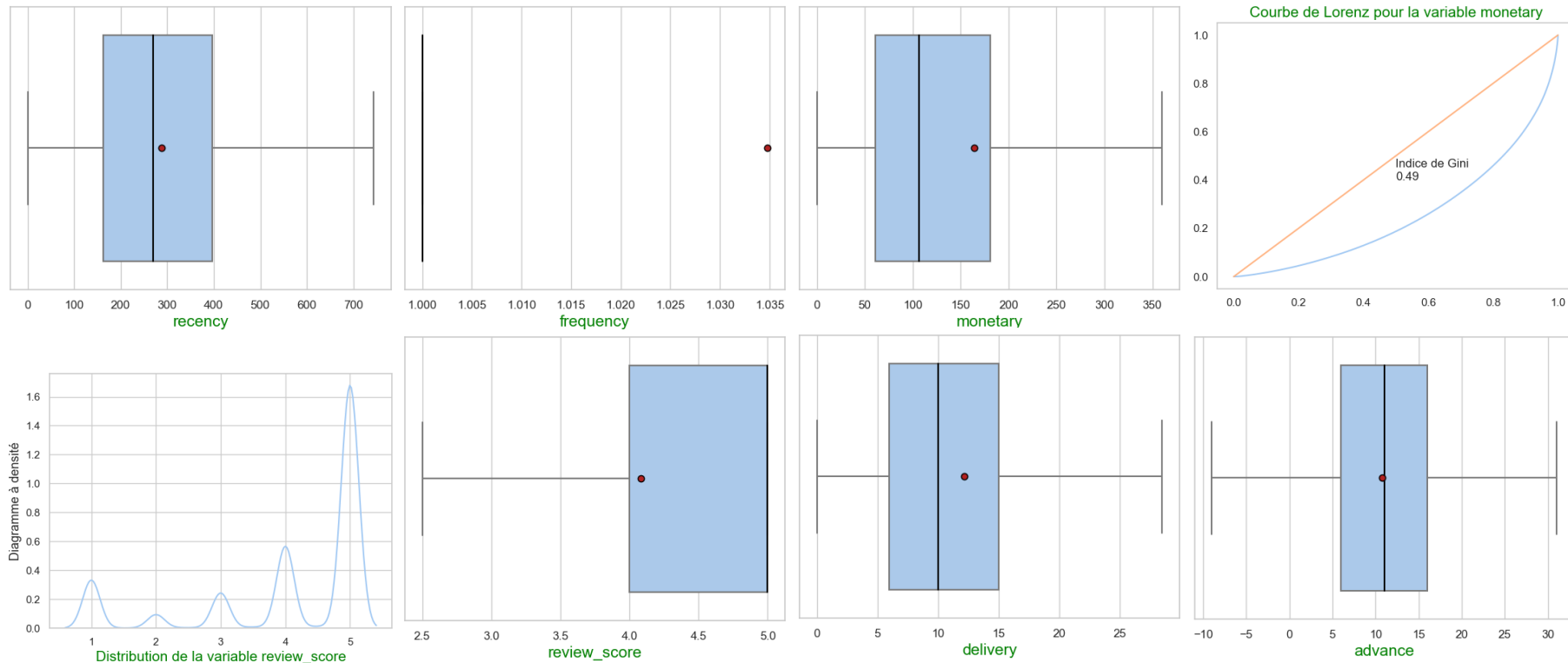
Exploration > nettoyage initial

Opération	Nombre de lignes	Nombre de variables
Fusion des datasets en un seul dataframe	119 143	40
Suppression catégories en portugais	119 143	39
Formatage des dates	119 143	39
Suppression doublons dus aux moyens de paiement supplémentaires	114 002	39
Suppression doublons dus aux échéances de paiement multiples	113 335	39

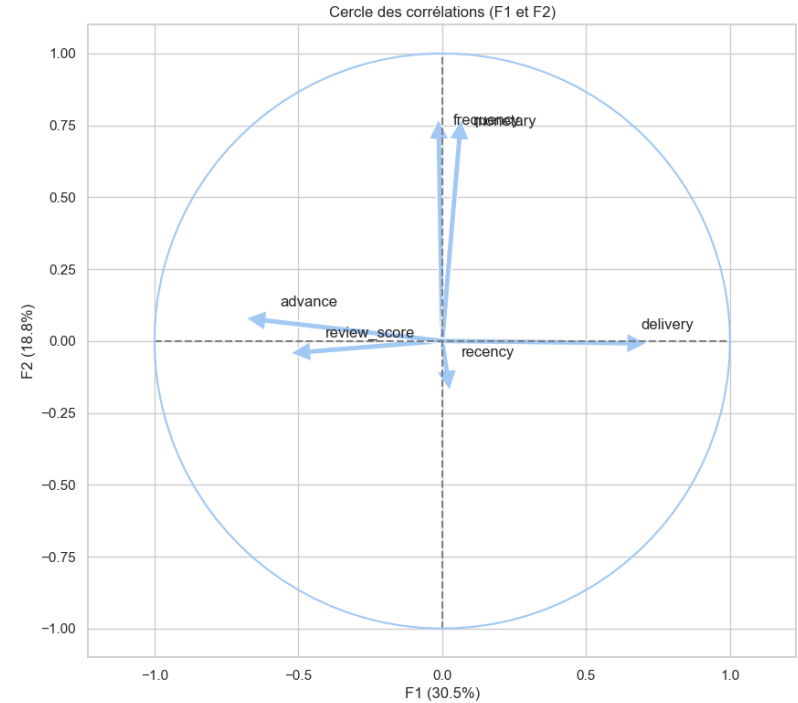
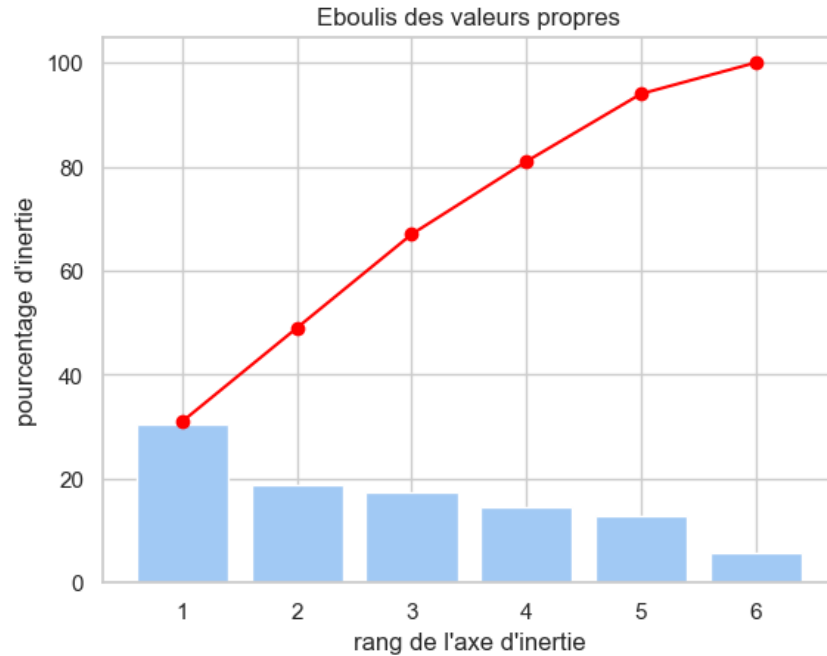
Exploration > feature engineering

- Construction d'un nouveau dataframe pour les 96 019 clients uniques avec 6 variables
 - Récence
 - Fréquence
 - Montant
 - Review Score moyen
 - Délai moyen de livraison
 - Avance de livraison moyenne
- Imputation des valeurs manquantes
 - review_score : 0,7%
 - delivery : 2,8%
 - Advance : 2,8%

Exploration > analyses univariées



Exploration > ACP

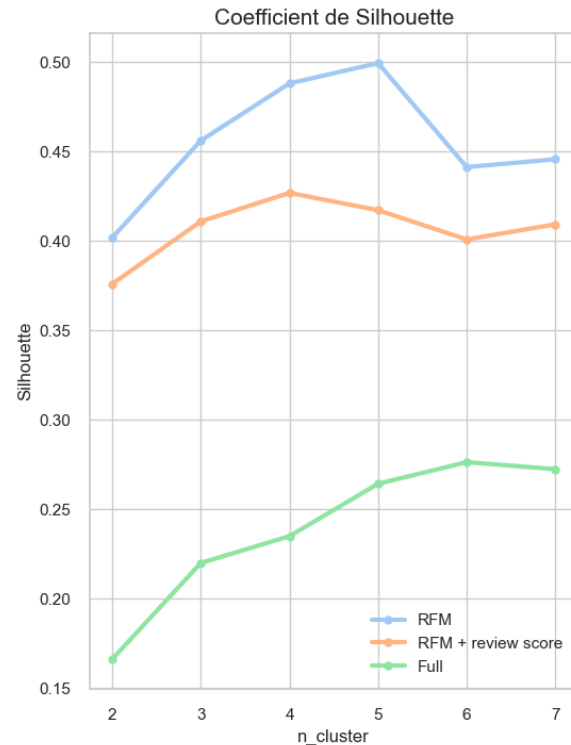
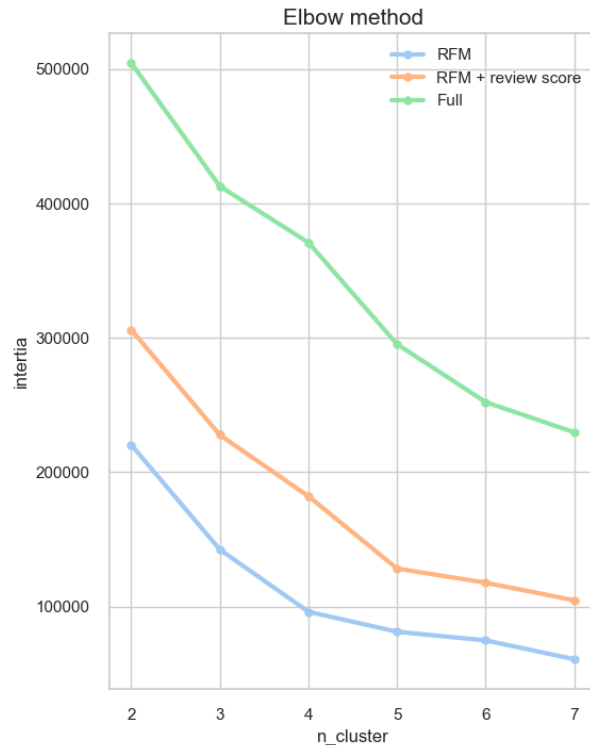


Segmentez des clients d'un site e-commerce

- Problématique
- Exploration
- **Essais de modélisation**
- Modèle final
- Maintenance du modèle
- Propositions

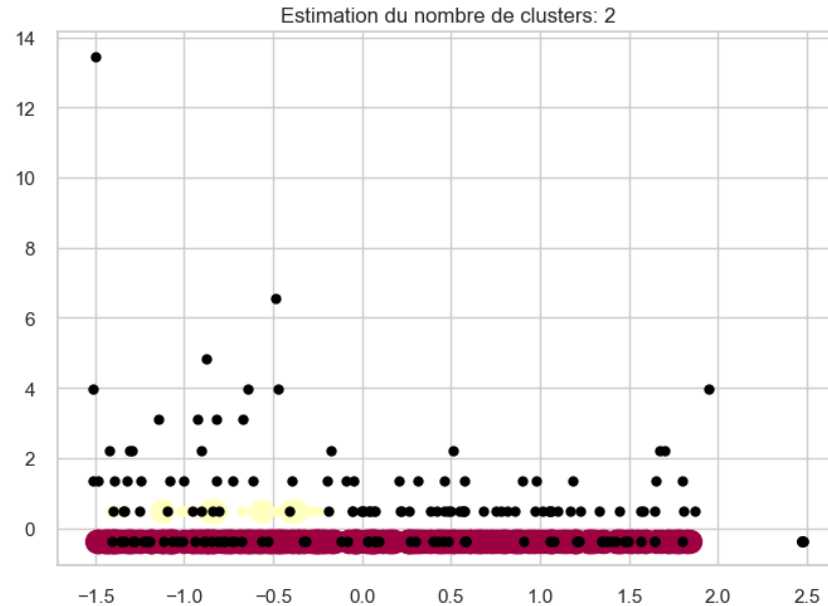
Essais de modélisation > dataframes

- RFM
- RFM + review_score
- RFM + review_score + delivery + advance



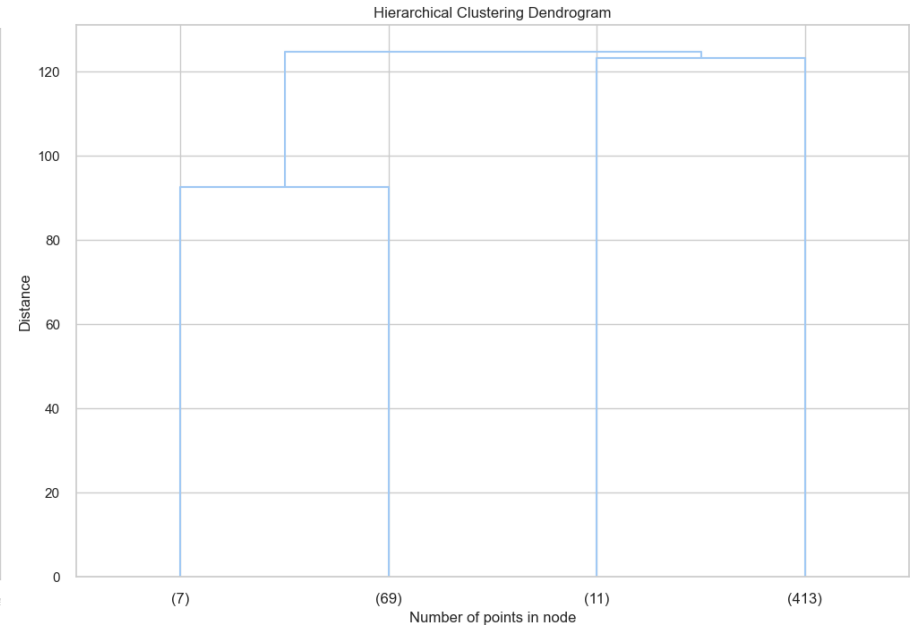
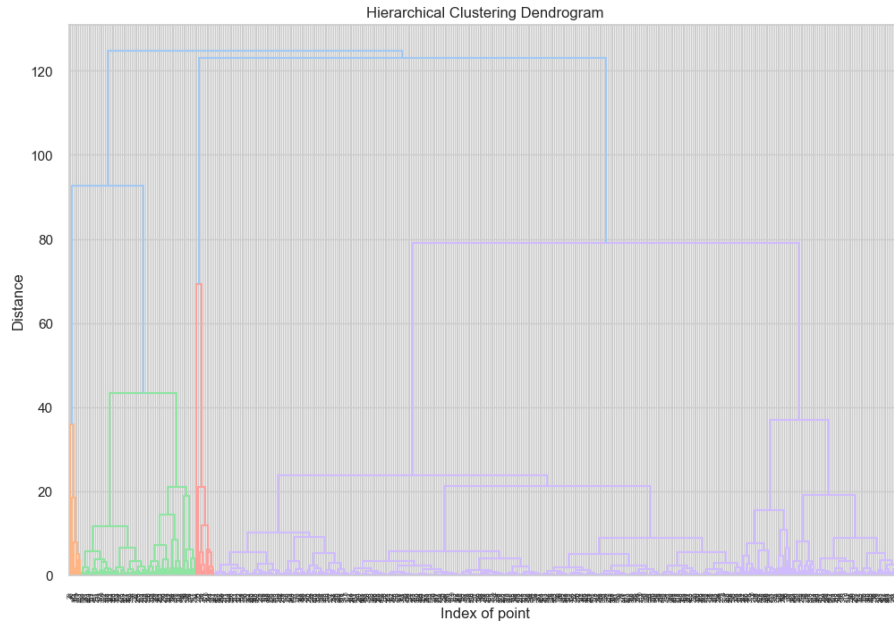
Essais de modélisation > algorithmes testés

DBSCAN avec pré-clustering de 500 centroïdes obtenus avec KMeans



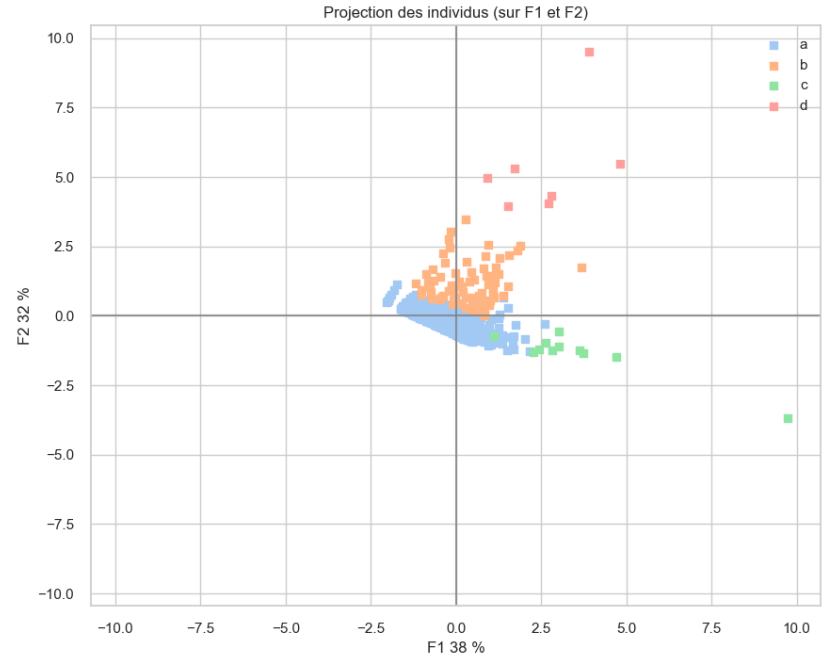
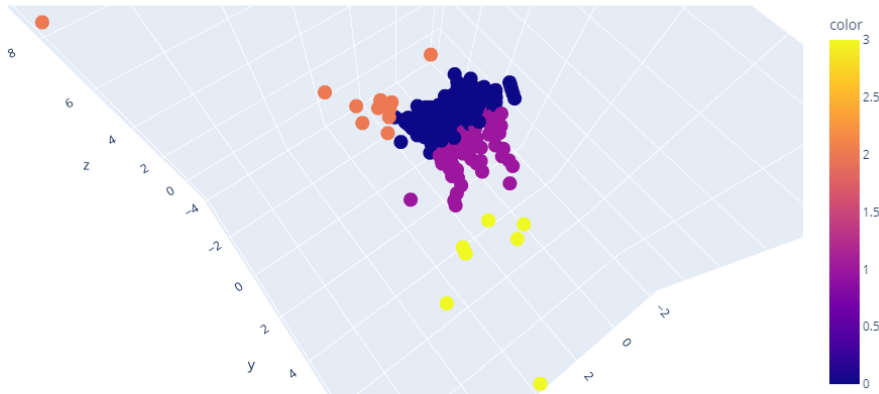
Essais de modélisation > algorithmes testés

CAH avec pré-clustering de 500 centroïdes obtenus avec KMeans



Essais de modélisation > algorithmes testés

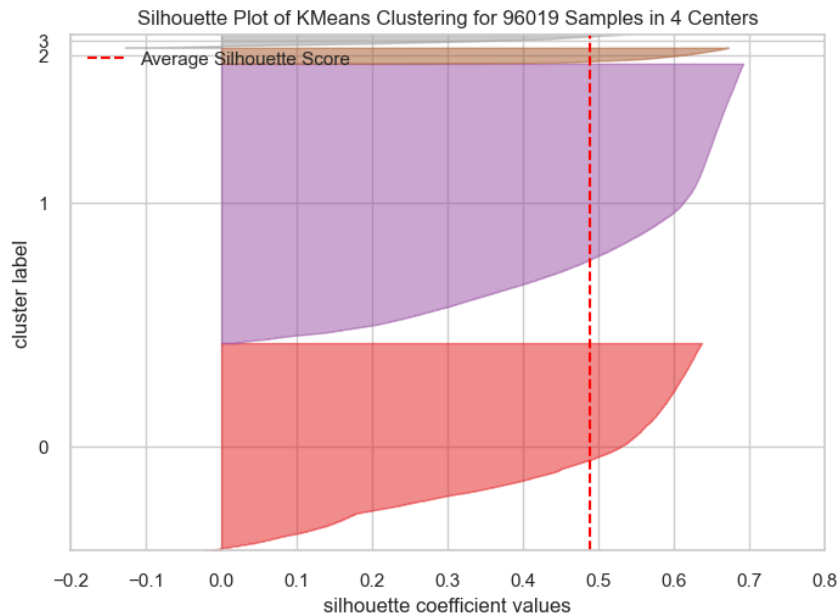
CAH avec pré-clustering de 500 centroïdes obtenus avec KMeans



Segmentez des clients d'un site e-commerce

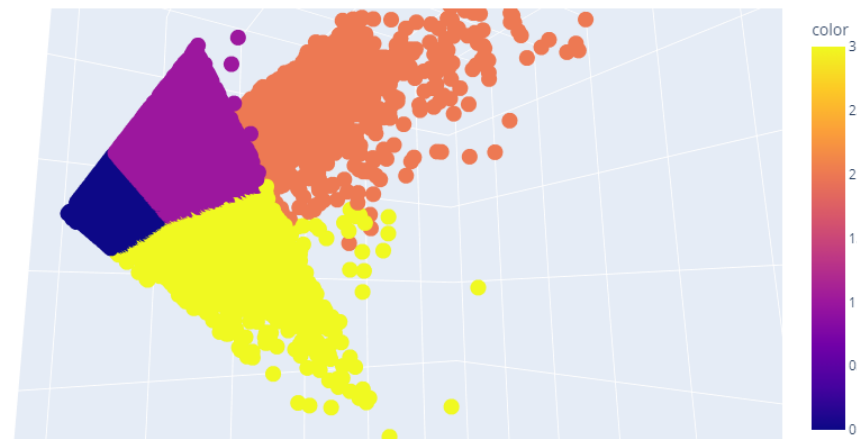
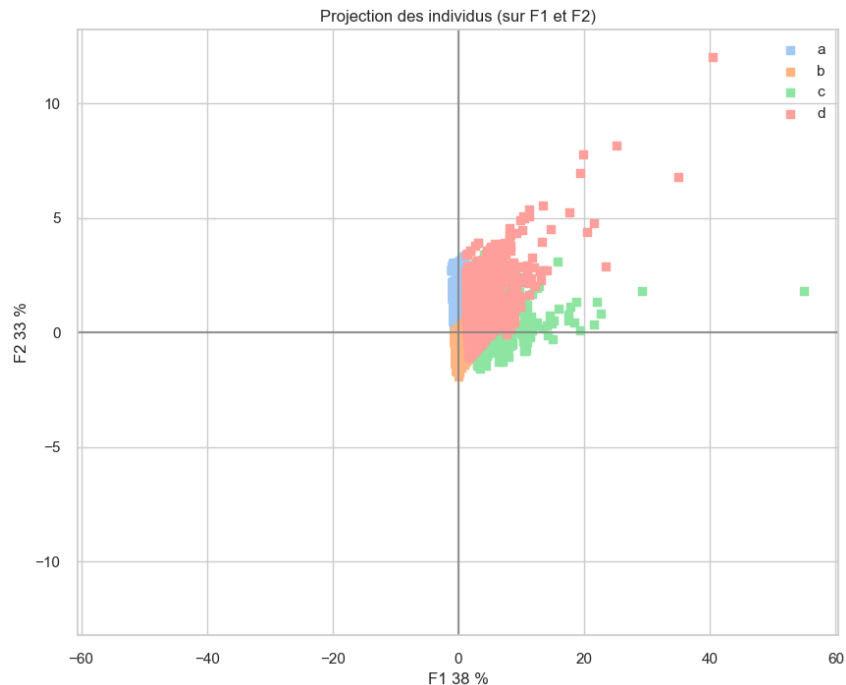
- Problématique
- Exploration
- Essais de modélisation
- **Modèle final**
- Maintenance du modèle
- Propositions

Modèle final > KMeans

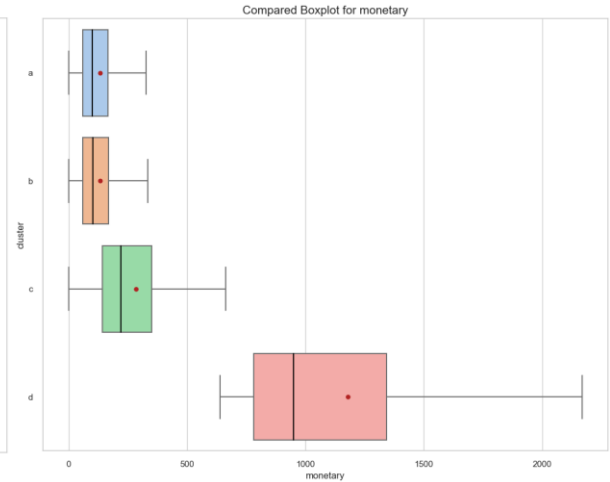
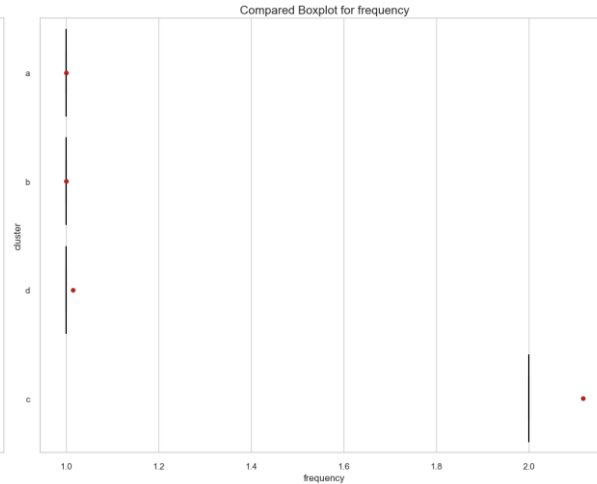
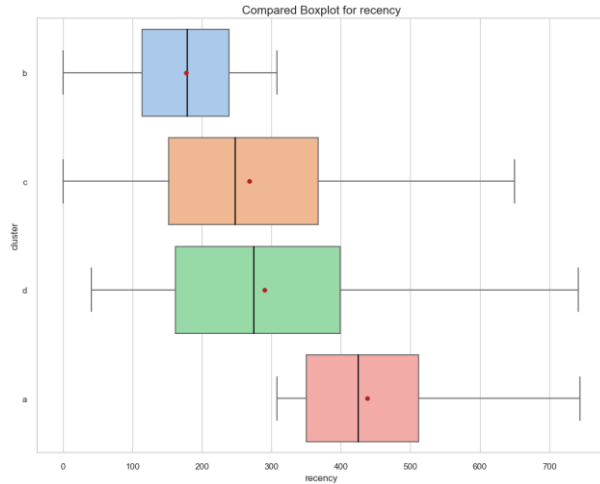


Cluster	Nombre clients	Part (%)
a	38623	40.2
b	51993	54.1
c	2961	3.1
d	2442	2.5

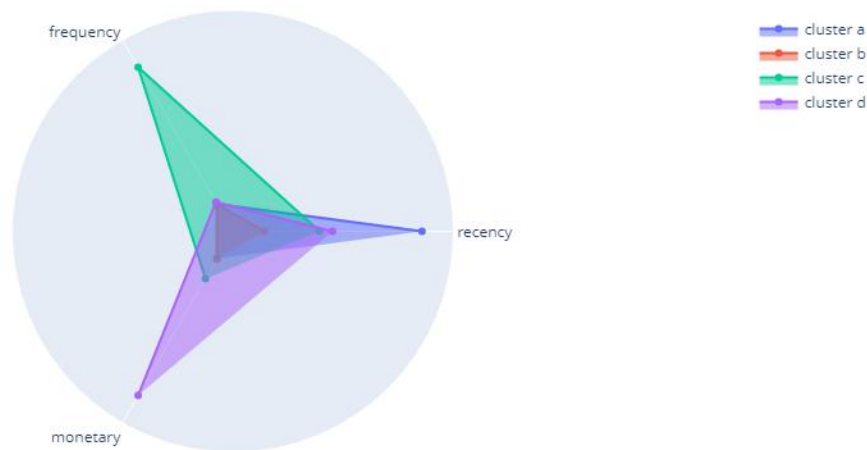
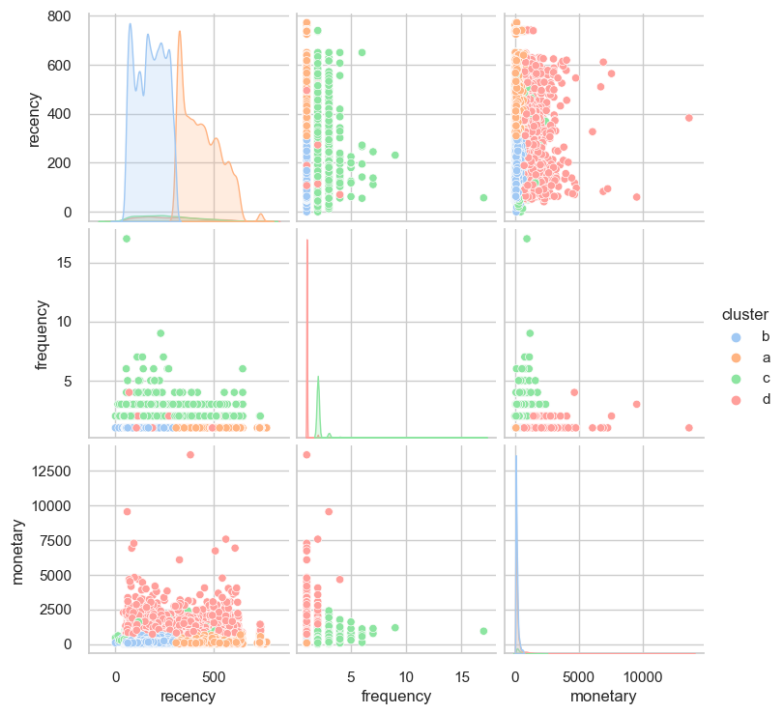
Modèle final > visualisations



Modèle final > analyse des clusters



Modèle final > an. bivariable, radarplot



Modèle final > segmentation des clients

Cluster	Caractéristiques	Appellation
A	Anciens clients ponctuels à faible pouvoir d'achat	Les perdus
B	Nouveaux clients à faible pouvoir d'achat	Les prometteurs
C	Clients récurrents à pouvoir d'achat modéré	Les fidèles
D	Clients occasionnels à haut pouvoir d'achat	Les risqués

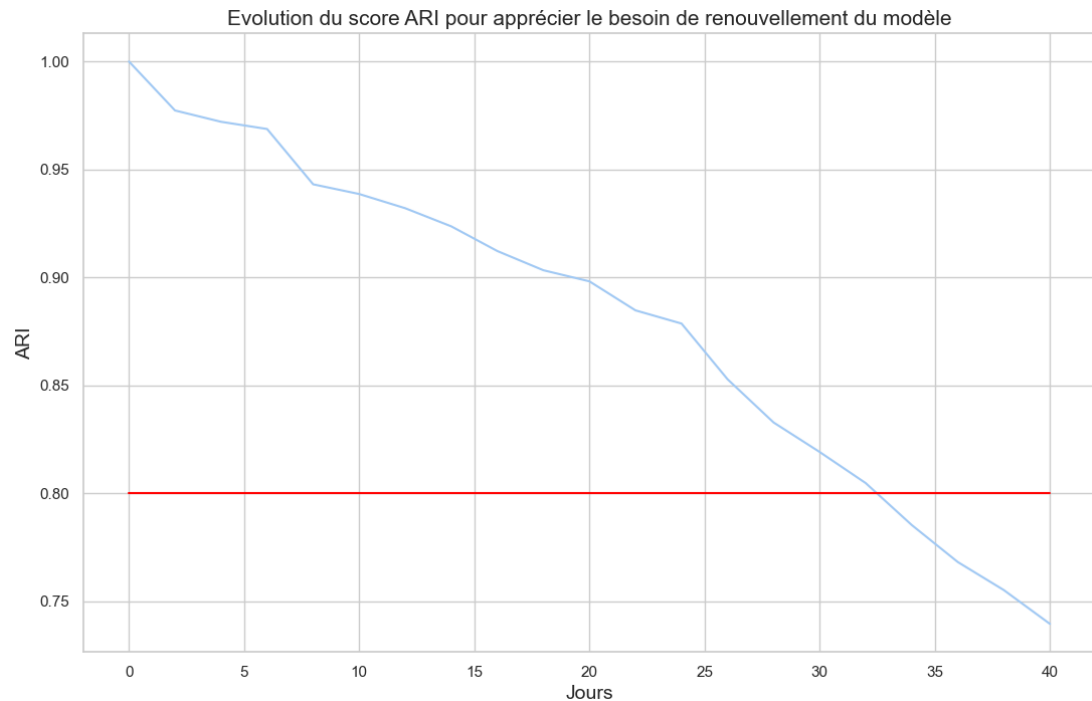
Segmentez des clients d'un site e-commerce

- Problématique
- Exploration
- Essais de modélisation
- Modèle final
- **Maintenance du modèle**
- Propositions

Maintenance du modèle > démarche

- Objectif
 - Déterminer la périodicité d'actualisation du modèle
- Étapes
 - Récupération du dataframe nettoyé, avant feature engineering
 - Création d'une ligne chronologique pour se soustraire aux contraintes de dates
 - Création de fonctions
 - Extraire des dataframes RFM périodiques
 - Instancier des scalers/modèles
 - Agglomérer les fonctions pour compiler les scores de comparaison entre scaler/modèle initial et ceux des périodes prolongées suivantes

Maintenance du modèle > conclusion



Segmentez des clients d'un site e-commerce

- Problématique
- Exploration
- Essais de modélisation
- Modèle final
- Maintenance du modèle
- **Propositions**

Propositions

Actions ciblées sur les clients

Segment de clients	Actions suggérées
Les perdus	<ul style="list-style-type: none">•raisons du départ•reconquête•réveil intérêt via campagne MKT adaptée
Les prometteurs	<ul style="list-style-type: none">•découverte des produits via réseaux sociaux•période d'essai offerte
Les fidèles	<ul style="list-style-type: none">•up-sell ciblés•communication via enquête de satisfaction•exclusivités, promotions
Les risqués	<ul style="list-style-type: none">•réveil du potentiel avec emails personnalisés•avantages, réductions sur prochains achats

Maintenance

32 jours

Avez-vous des questions ?