Segmenter des clients d’un site e-commerce

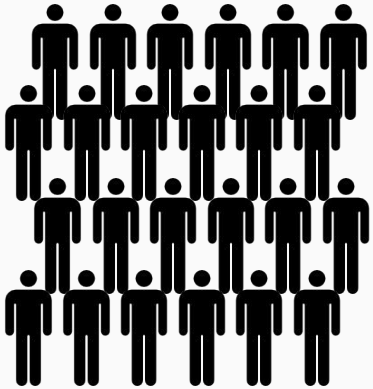
22/10/2021 - Parcours Data Scientist

Sébastien Bourgeois

Sommaire

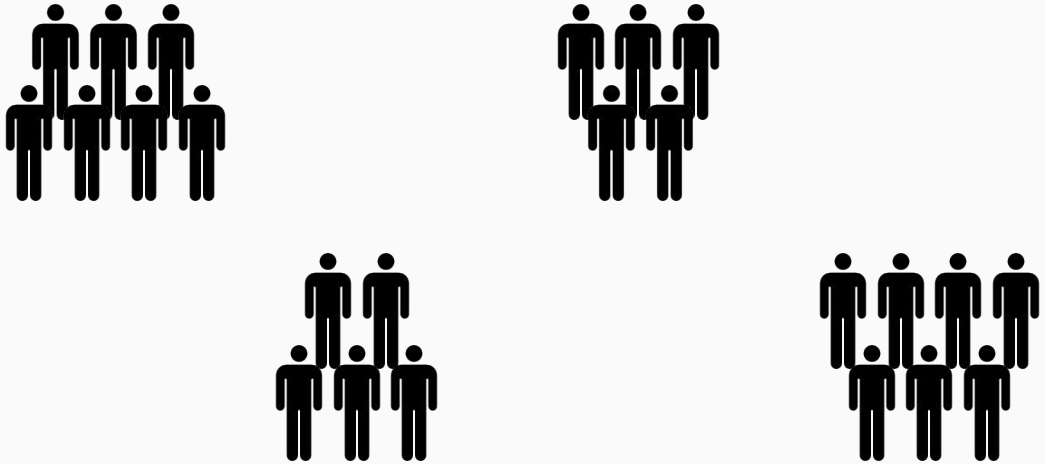
1. Problématique

Problématique



1. Problématique

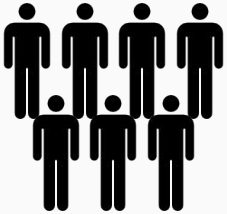
Problématique



Les icônes de cette slide ont été conçues par Freepik du site Flaticon.com

1. Problématique

Problématique



Les icônes de cette slide ont été conçues par Freepik, Alfredo Hernandez et Pixel perfect du site Flaticon.com

1. Problématique

Interprétation

Identification de groupes de clients similaires

Problème de clustering

1. Problématique

Pistes de recherche envisagées

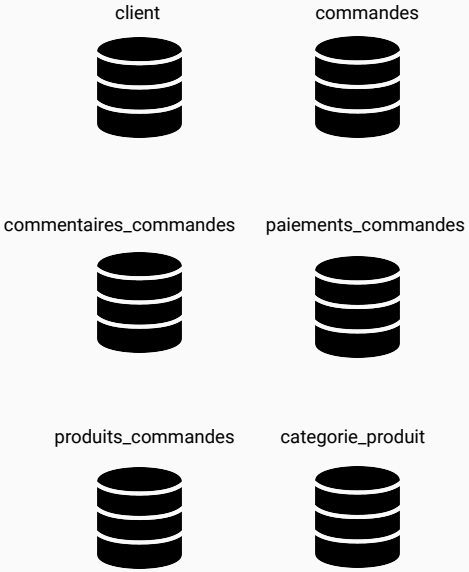
Segmentation RFM

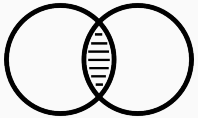
k-means

k-prototypes

2. Nettoyage & exploration

Nettoyage des données



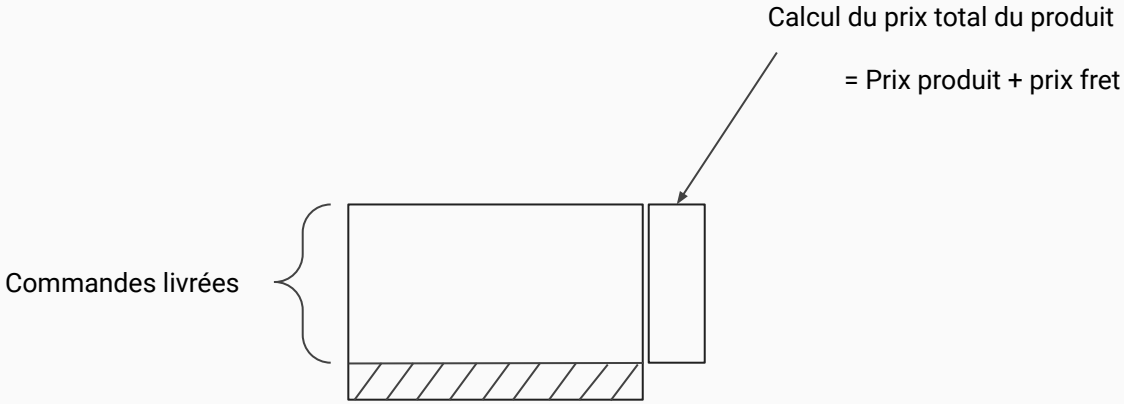


Produits commandés par les clients

Les icônes de cette slide ont été conçues par Stockio et Freepik du site Flaticon.com

2. Nettoyage & exploration

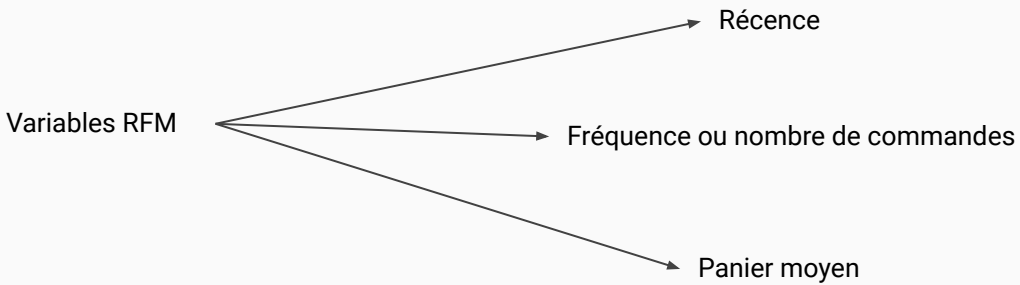
Nettoyage des données



2. Nettoyage & exploration

Features engineering

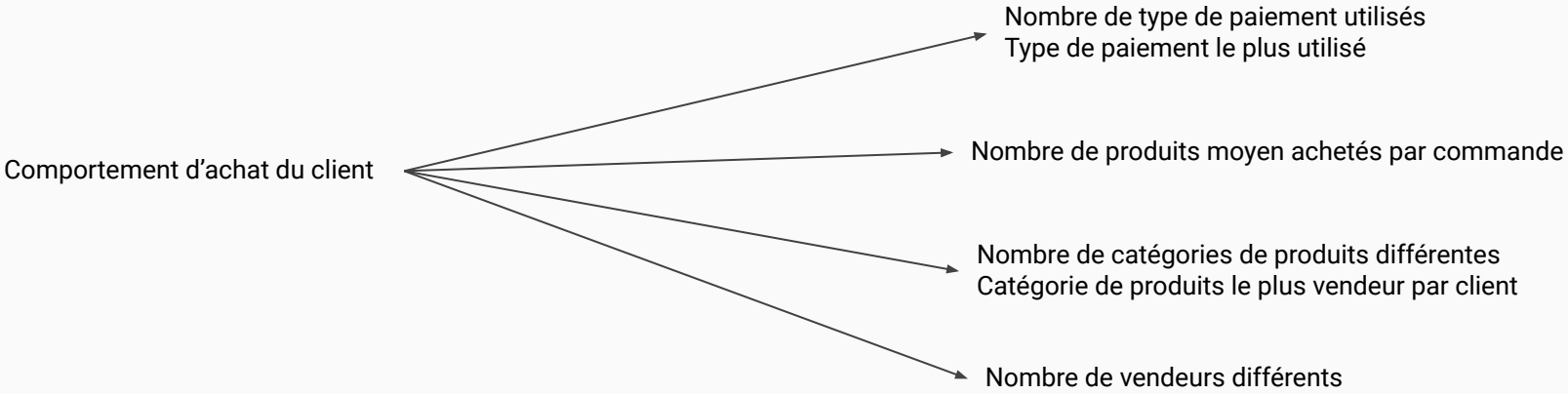
Création de variables orientées marketing



2. Nettoyage & exploration

Features engineering

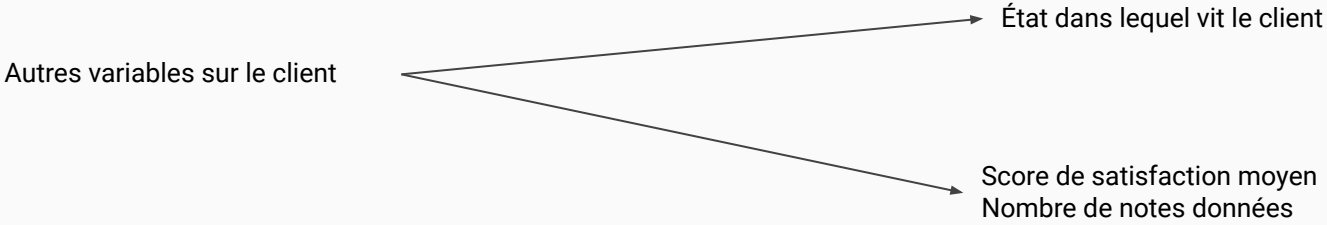
Création de variables orientées marketing



2. Nettoyage & exploration

Features engineering

Création de variables orientées marketing



2. Nettoyage & exploration

Exploration

top_category
bed_bath_table
NULL
NULL
telephony
baby
auto
health_beauty
auto

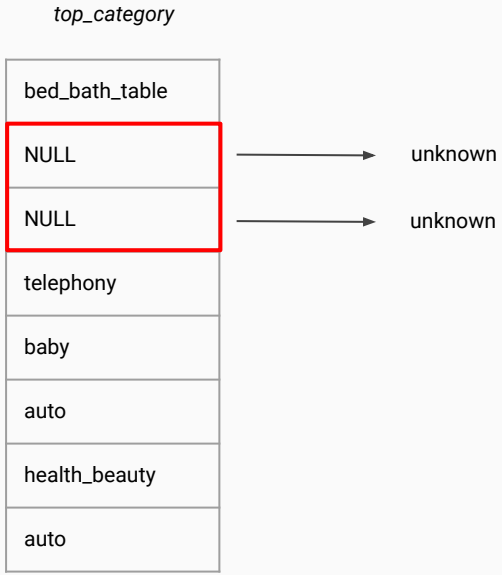
2. Nettoyage & exploration

Exploration

top_category
bed_bath_table
NULL
NULL
telephony
baby
auto
health_beauty
auto

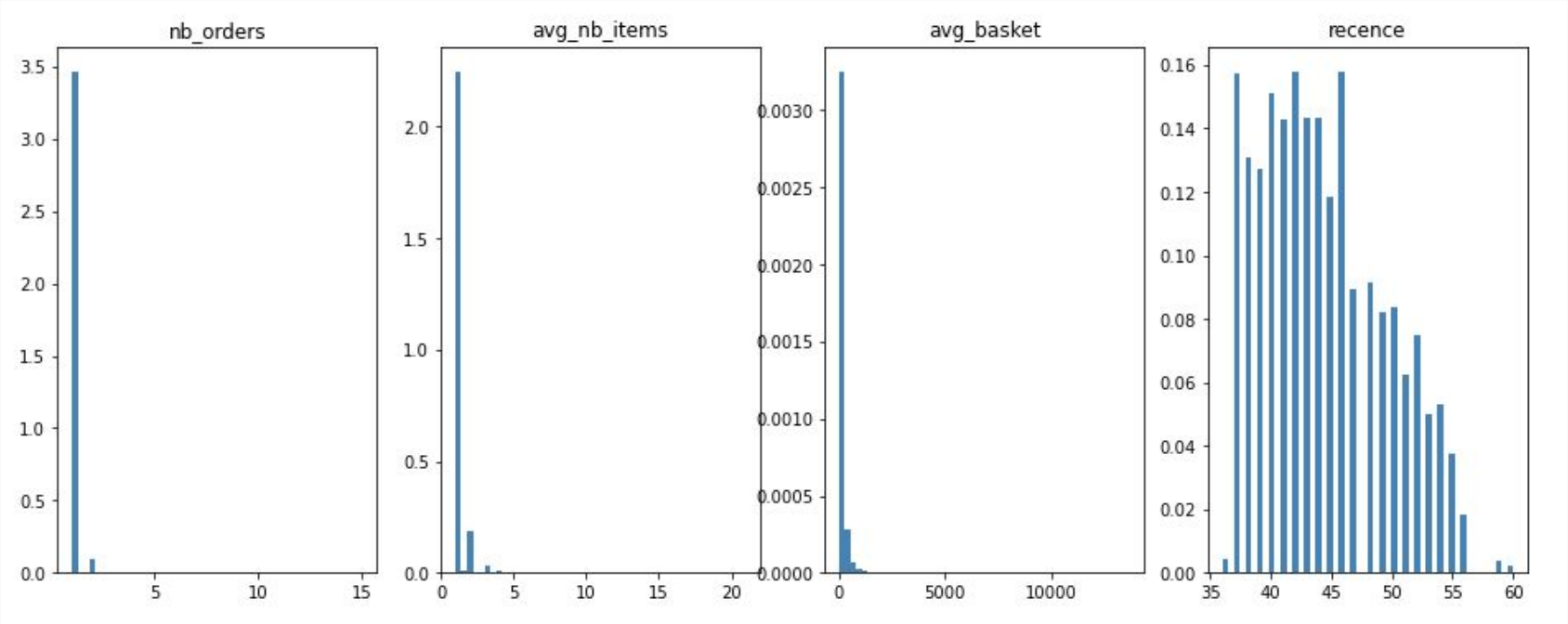
2. Nettoyage & exploration

Exploration



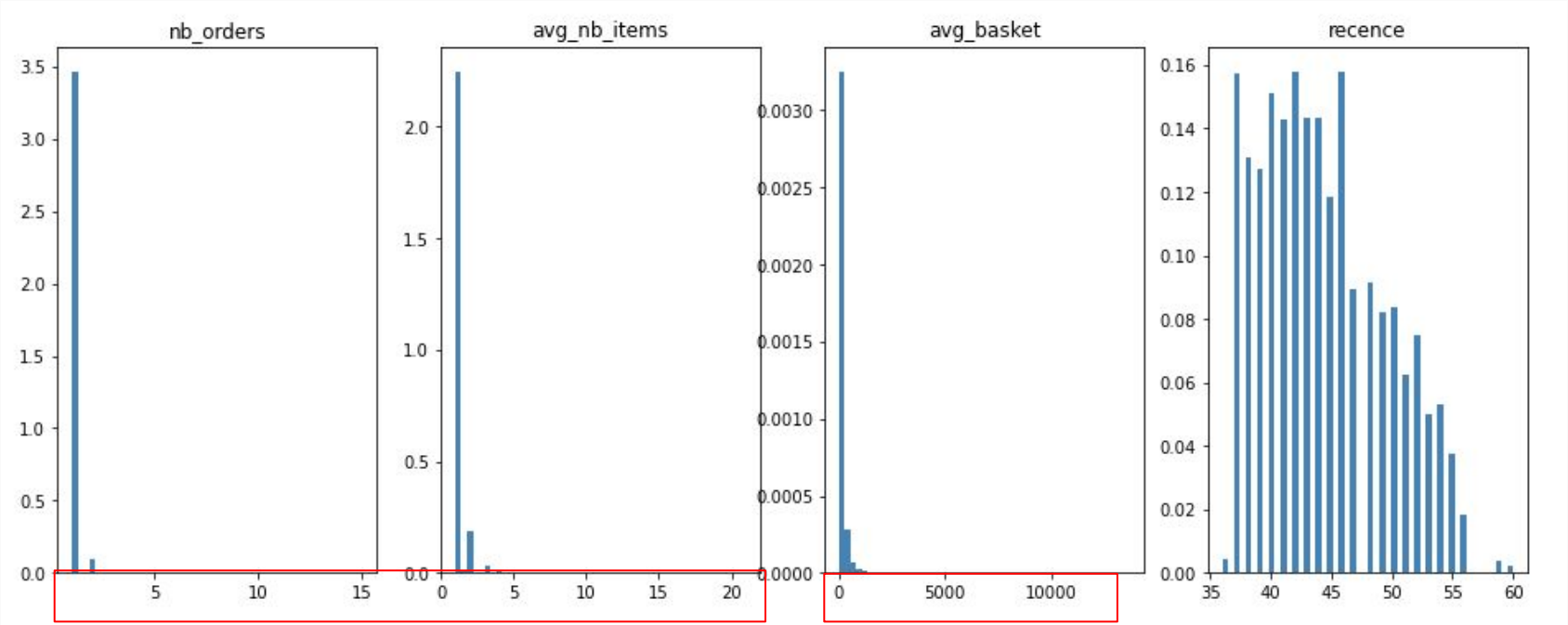
2. Nettoyage & exploration

Exploration



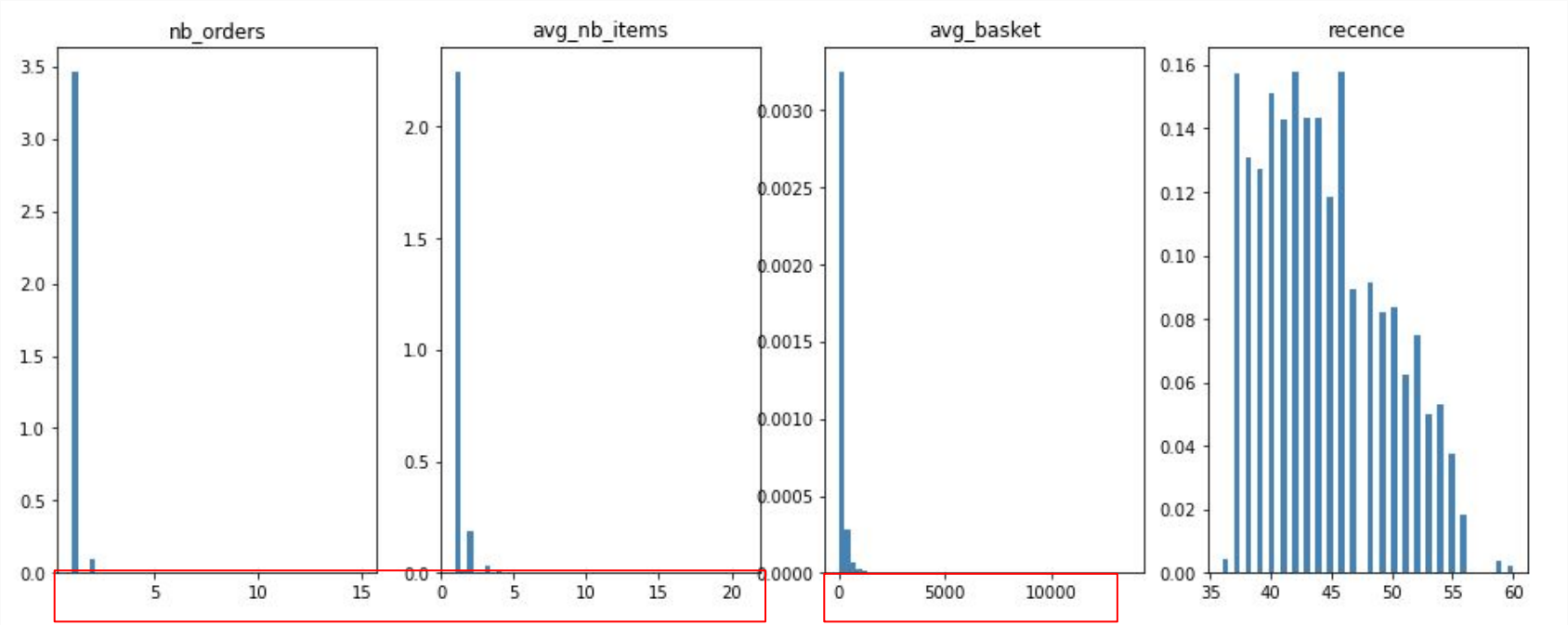
2. Nettoyage & exploration

Exploration



2. Nettoyage & exploration

Exploration



=> Standardisation des variables numériques

2. Nettoyage & exploration

Exploration

1 achat

1 produit

1 commentaire

1 vendeur

Bonne satisfaction

Les icônes de cette slide ont été conçues par Freepik du site Flaticon.com

2. Nettoyage & exploration

Exploration

1 achat

1 produit

1 commentaire

1 vendeur

Bonne satisfaction

En moyenne

Les icônes de cette slide ont été conçues par Freepik du site Flaticon.com

2. Nettoyage & exploration

Exploration

1 achat

1 produit

1 commentaire

1 vendeur

Bonne satisfaction

75% des clients

Les icônes de cette slide ont été conçues par Freepik du site Flaticon.com

2. Nettoyage & exploration

Exploration

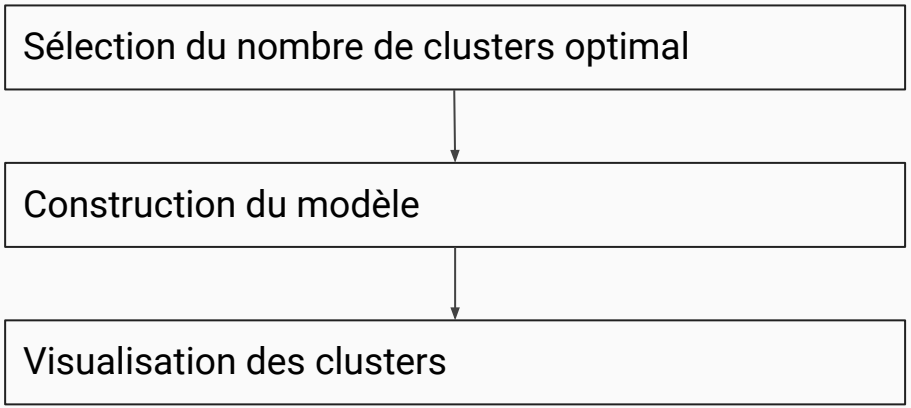
São Paulo

Rio de Janeiro

Les icônes de cette slide ont été conçues par Freepik du site Flaticon.com

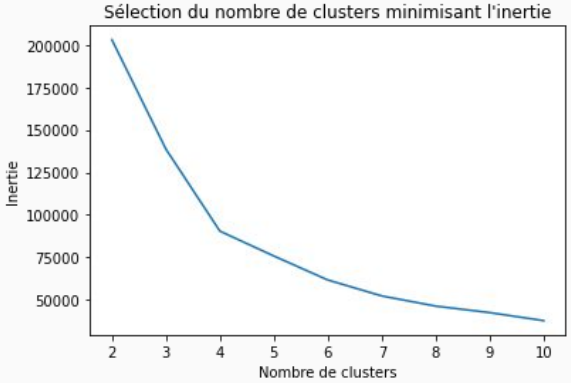
3. Pistes de segmentation des clients

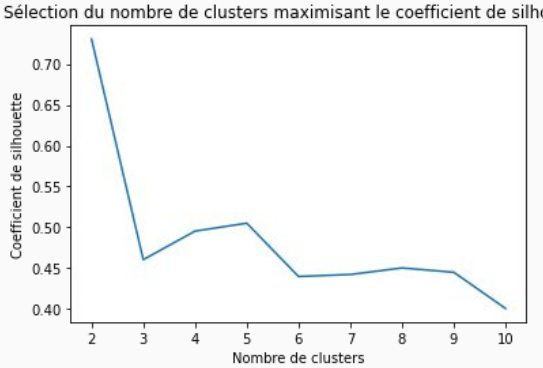
Méthodologie

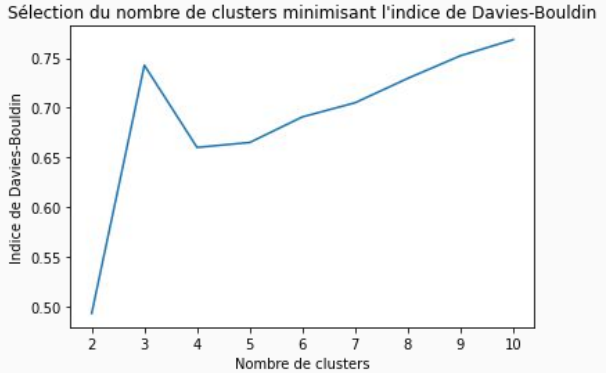


3. Pistes de segmentation des clients

Baseline : segmentation RFM

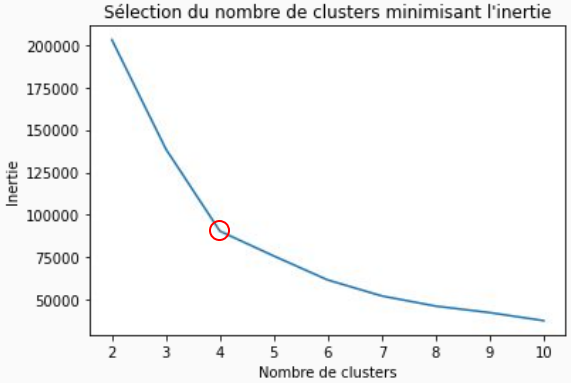


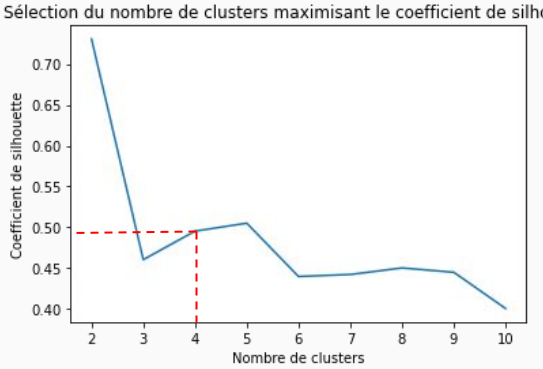


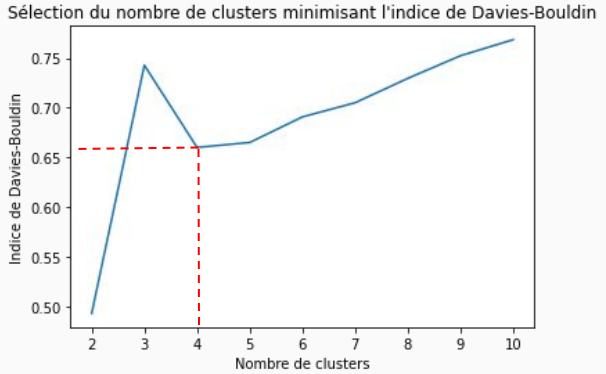


3. Pistes de segmentation des clients

Baseline : segmentation RFM



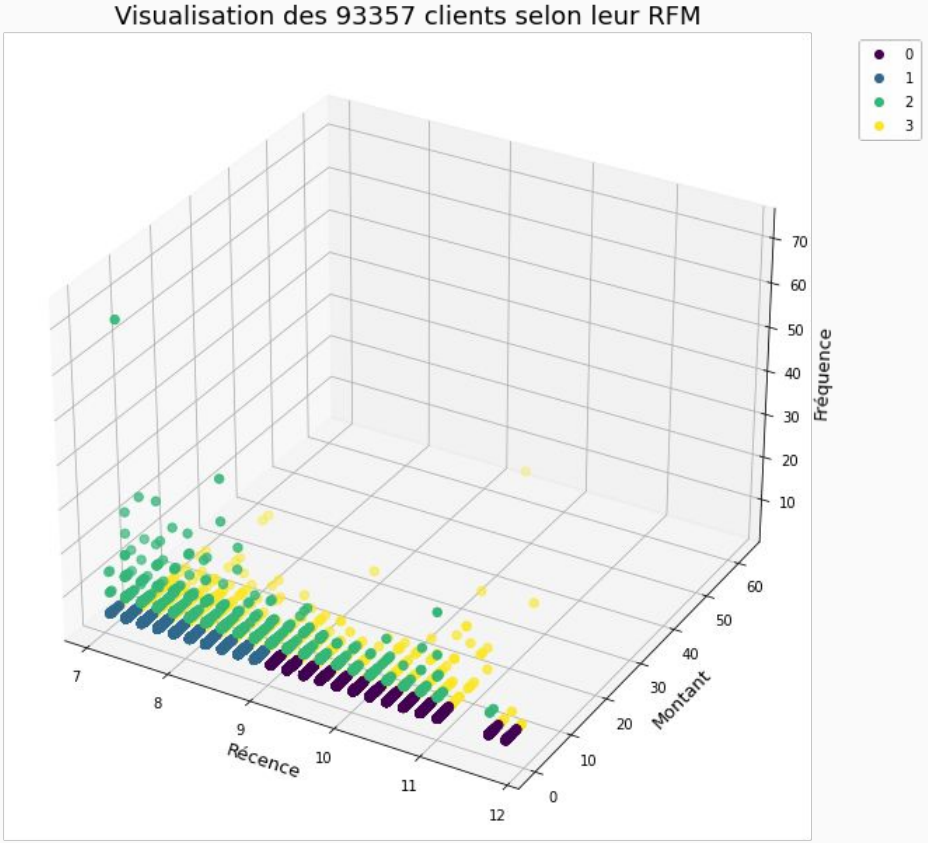




k-means avec 4 clusters avec les variables RFM

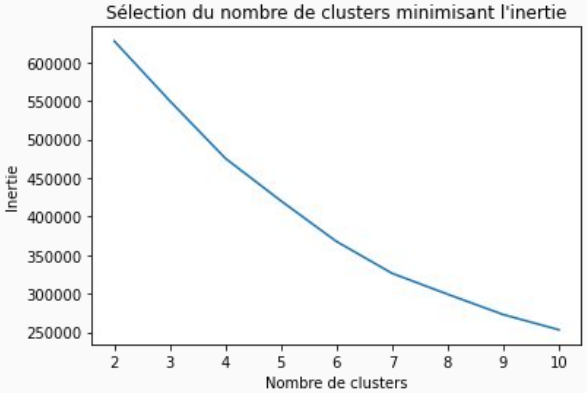
3. Pistes de segmentation des clients

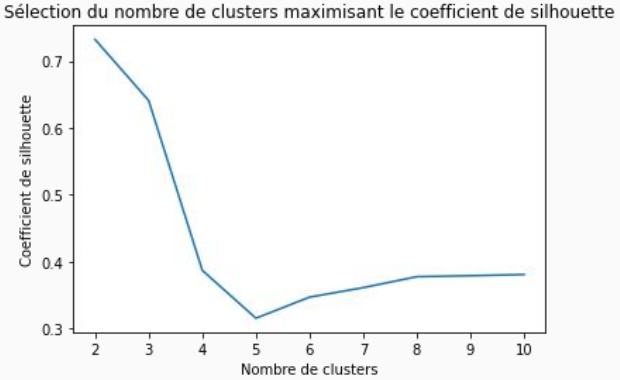
Baseline : segmentation RFM

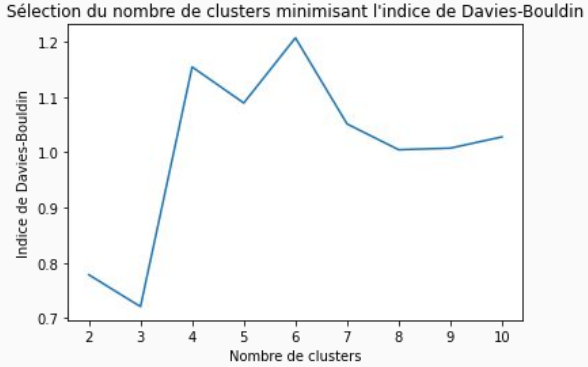


3. Pistes de segmentation des clients

k-means

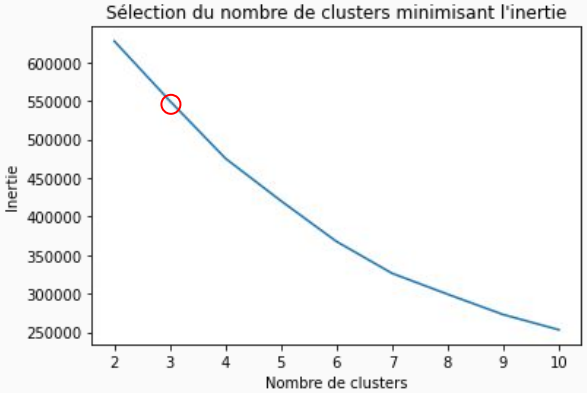


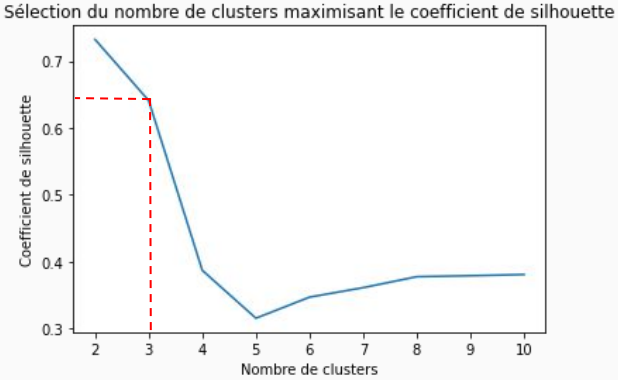


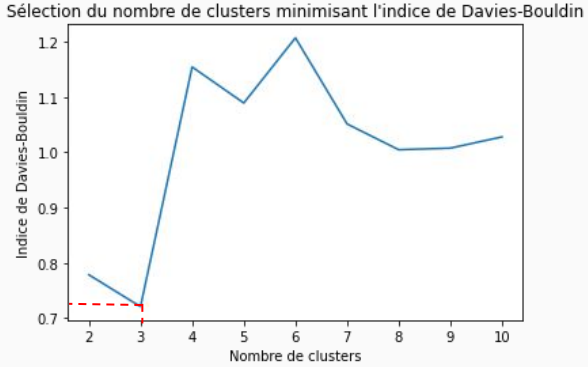


3. Pistes de segmentation des clients

k-means



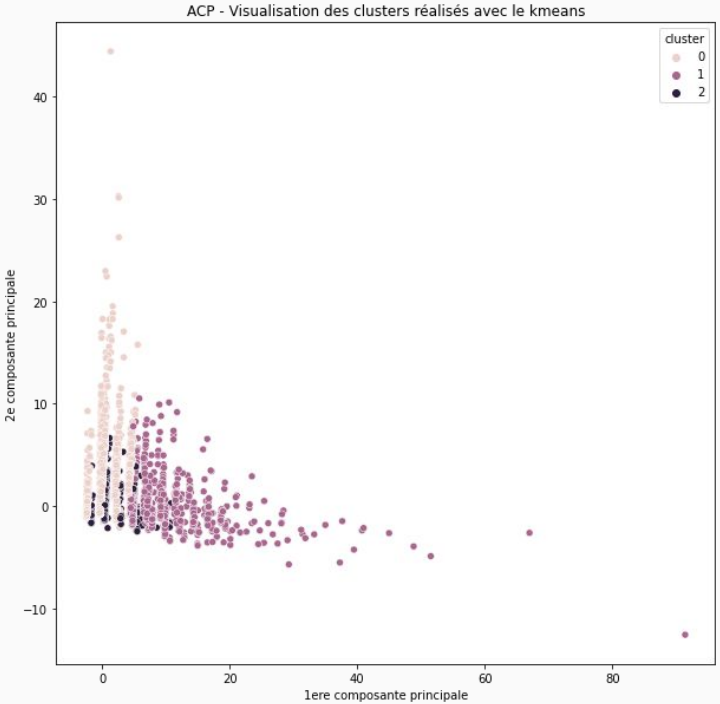


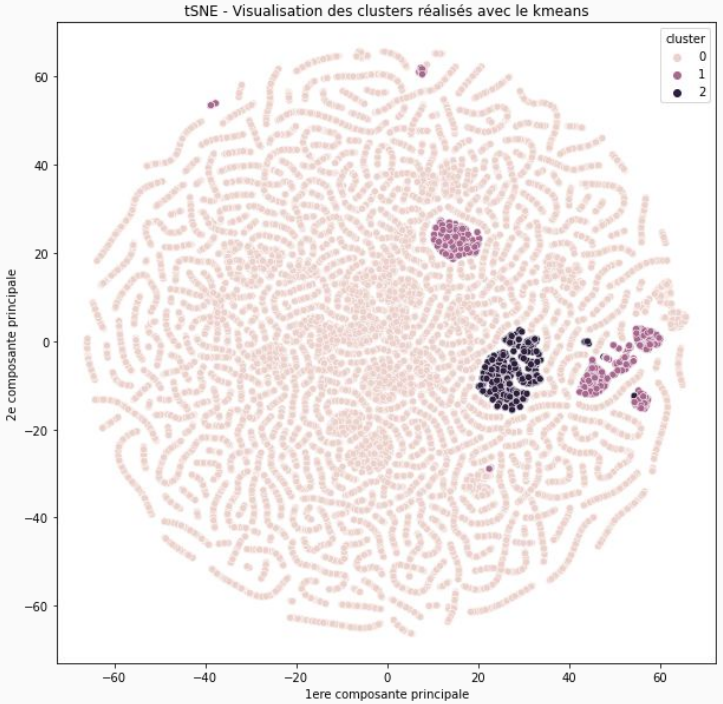


k-means avec 3 clusters sur toutes les variables numériques

3. Pistes de segmentation des clients

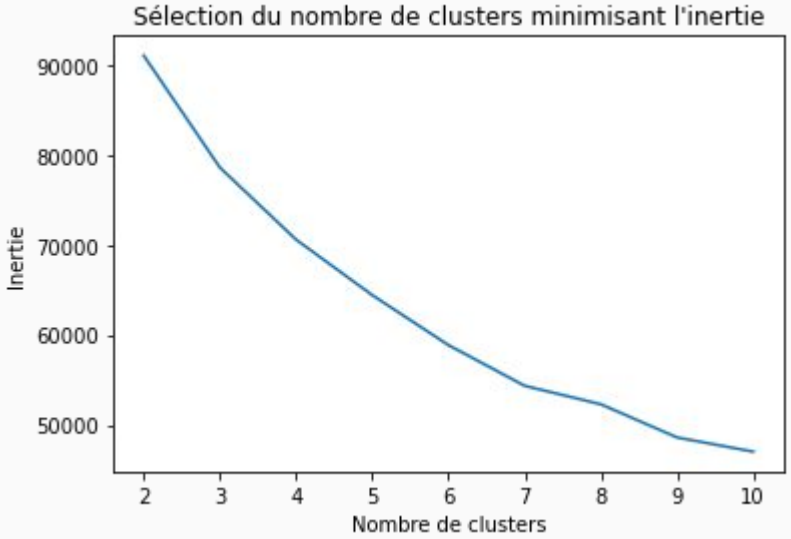
k-means





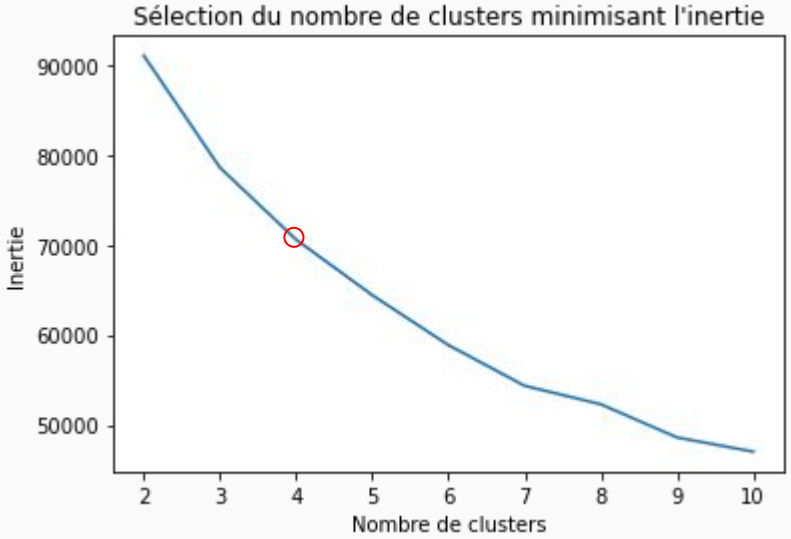
3. Pistes de segmentation des clients

k-prototypes



3. Pistes de segmentation des clients

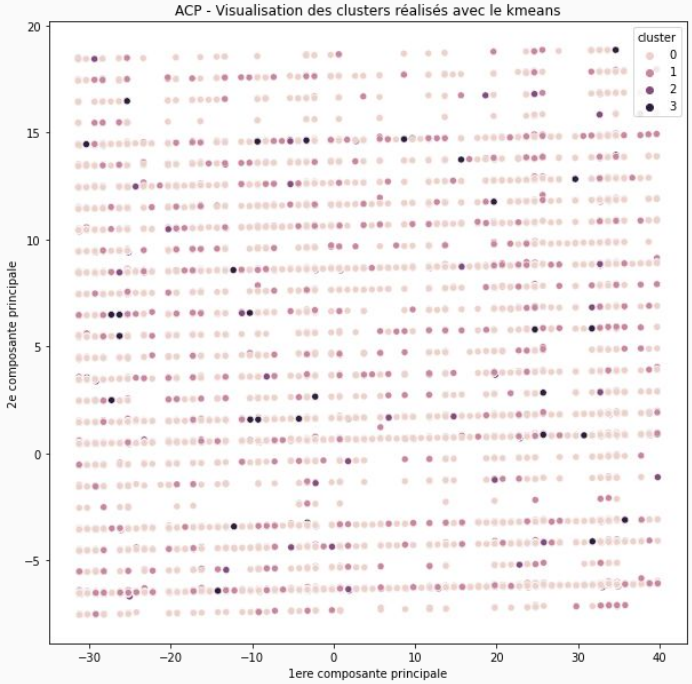
k-prototypes

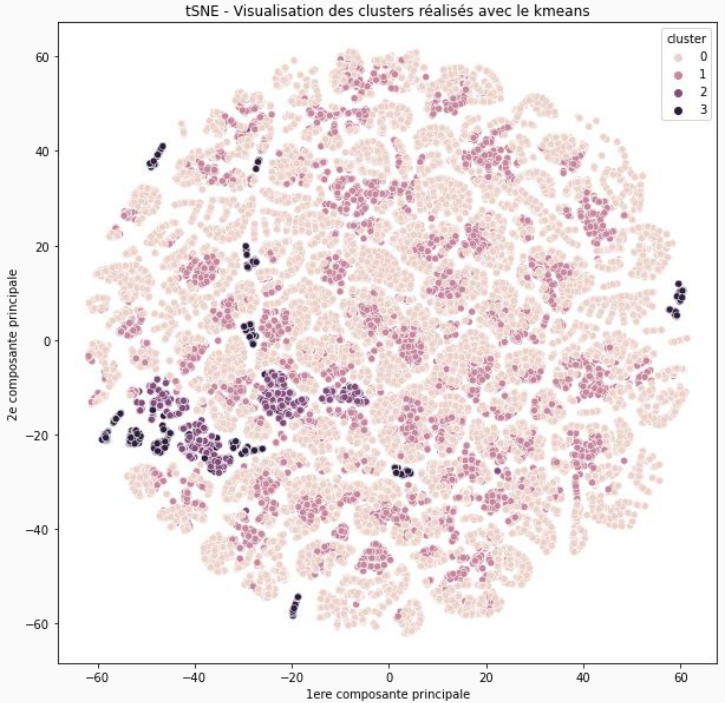


k-prototypes avec 4 clusters sur toutes les variables

3. Pistes de segmentation des clients

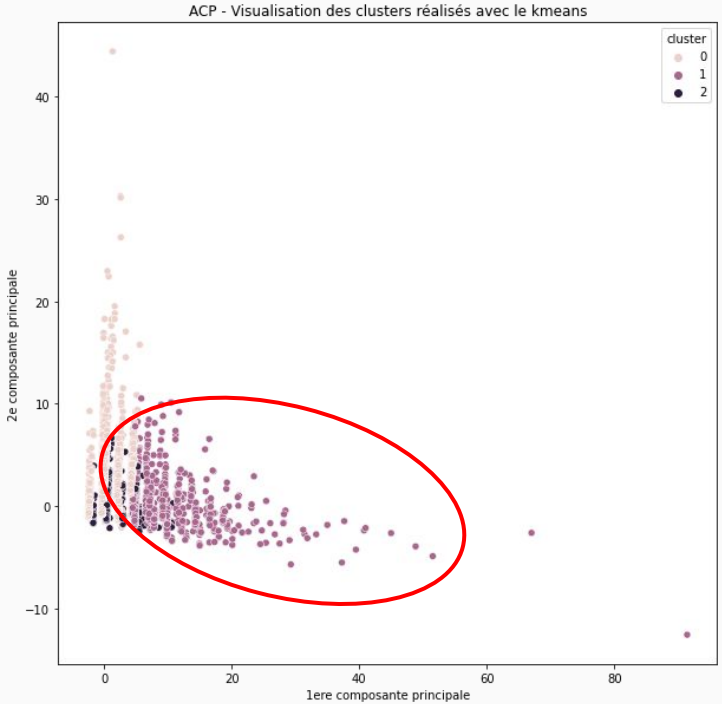
k-prototypes

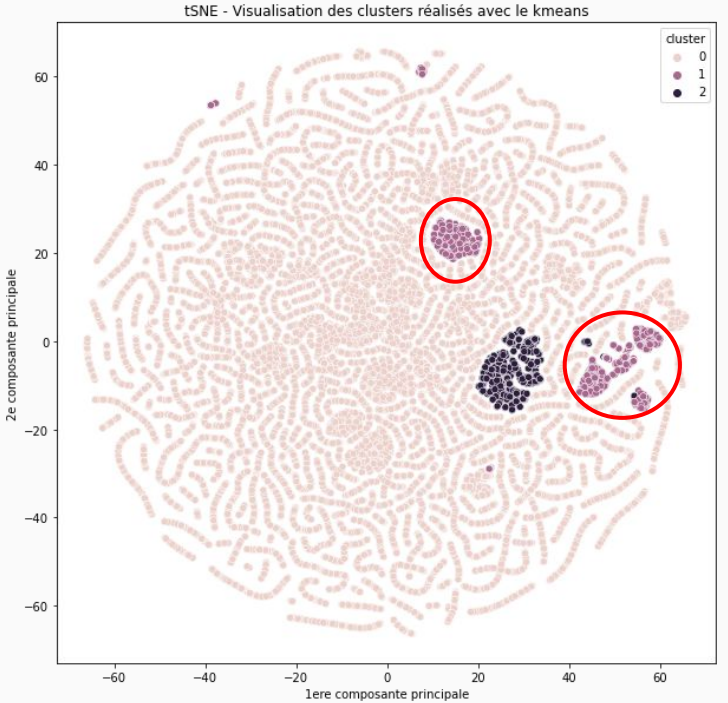




4. Modèle final sélectionné

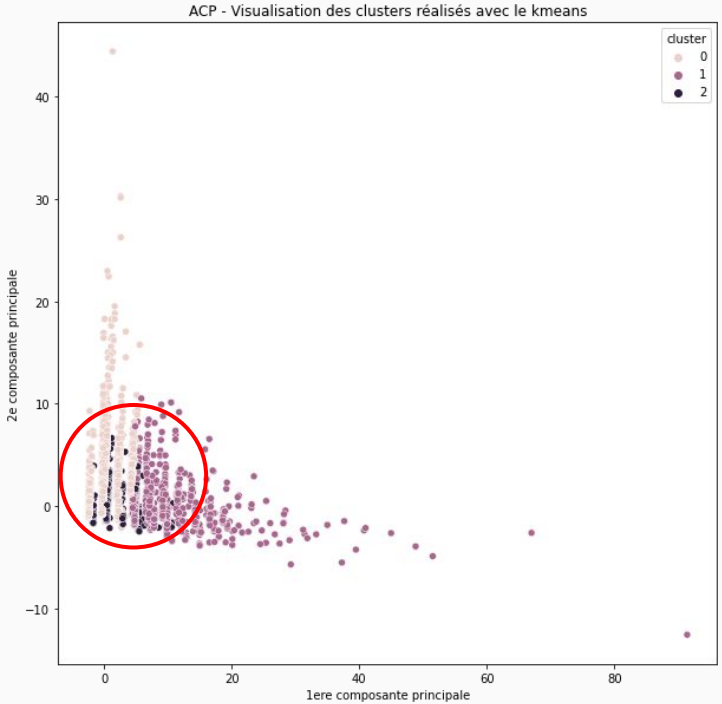
k-means sur les variables numériques

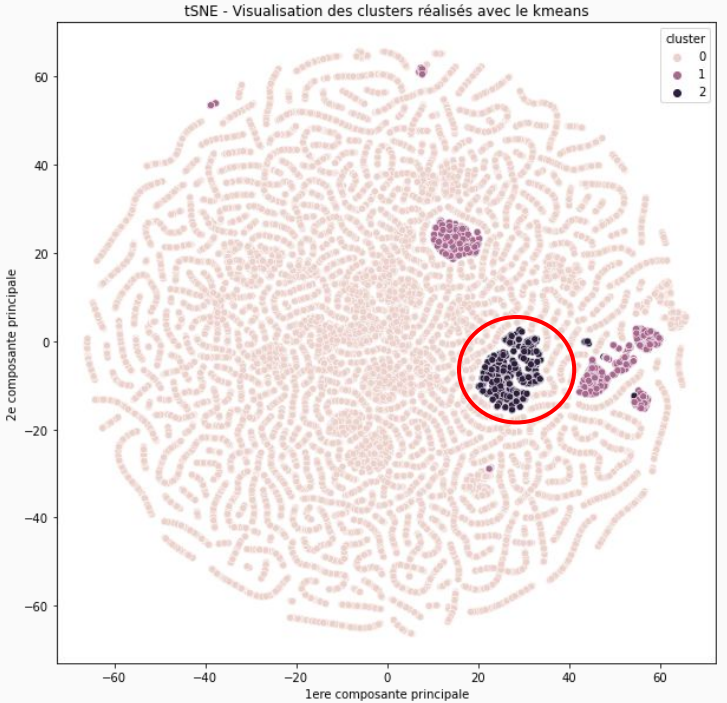




4. Modèle final sélectionné

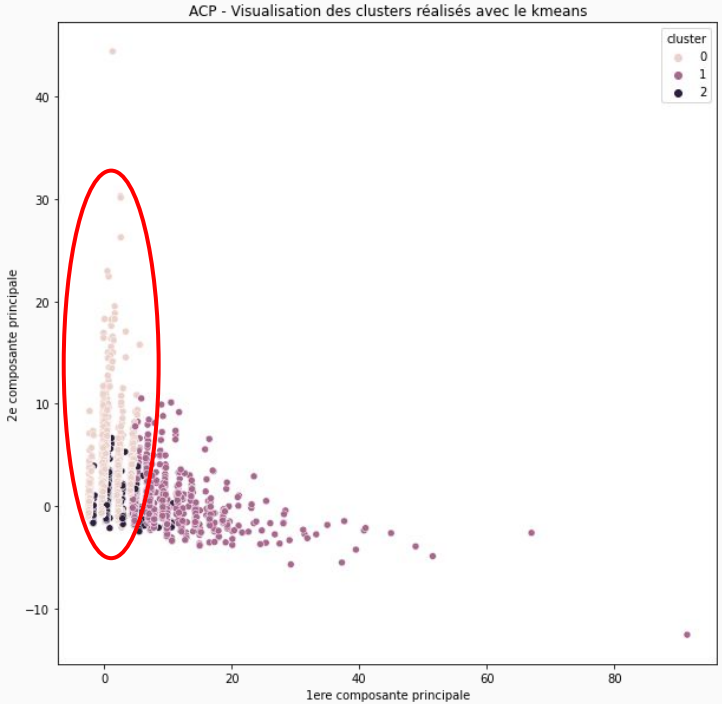
k-means sur les variables numériques

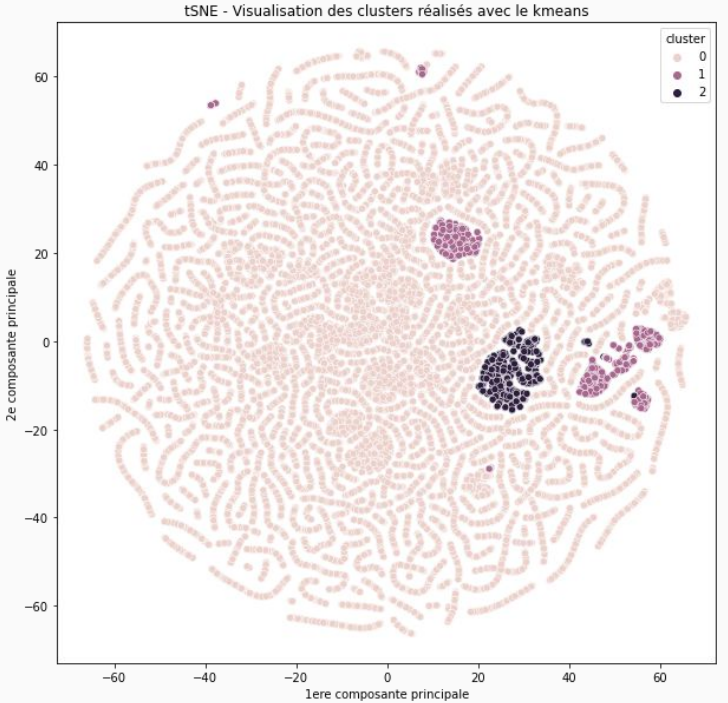




4. Modèle final sélectionné

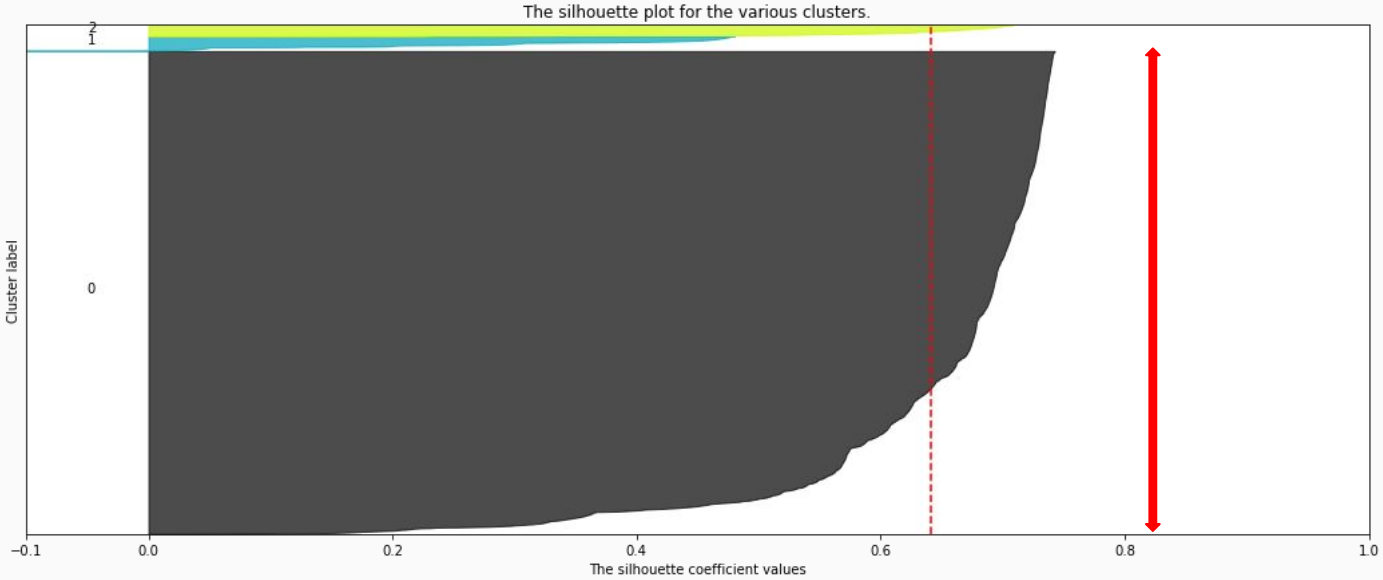
k-means sur les variables numériques





4. Modèle final sélectionné

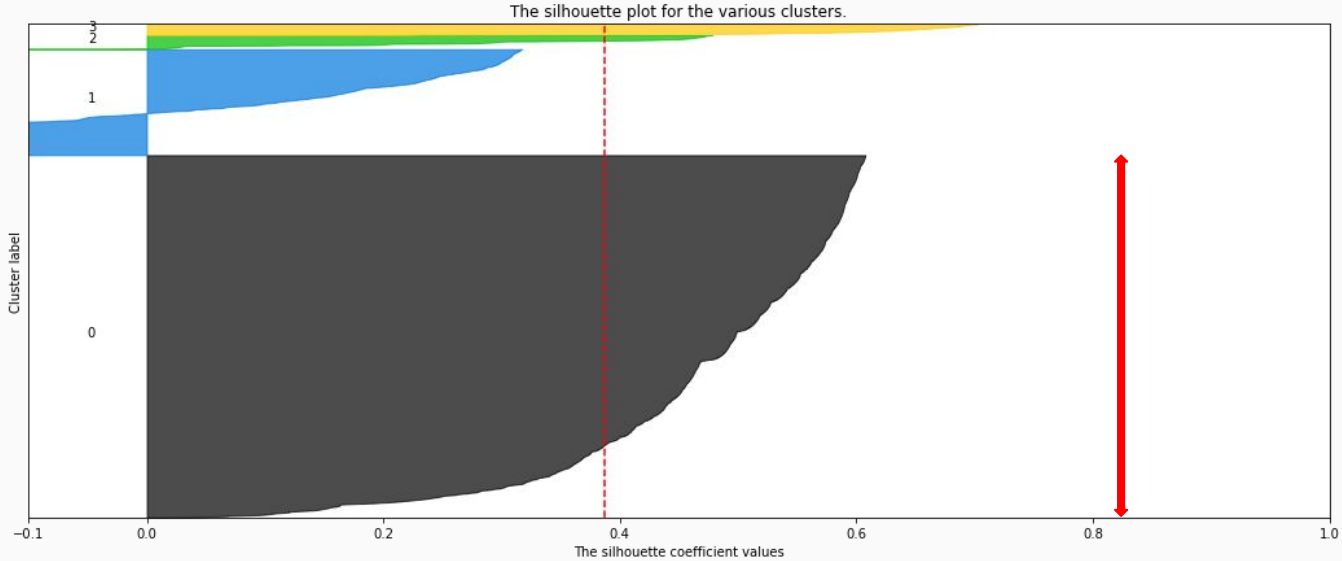
k-means sur les variables numériques



Surreprésentation

4. Modèle final sélectionné

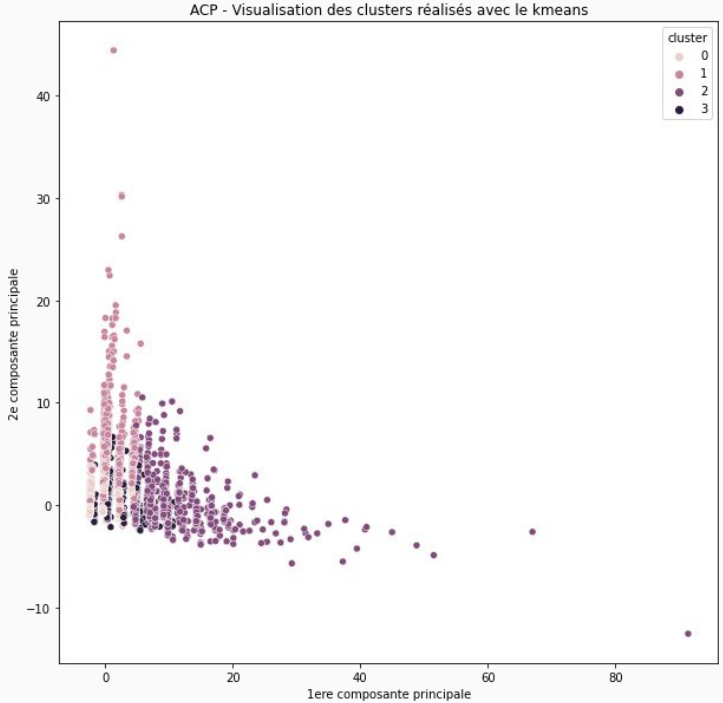
k-means sur les variables numériques

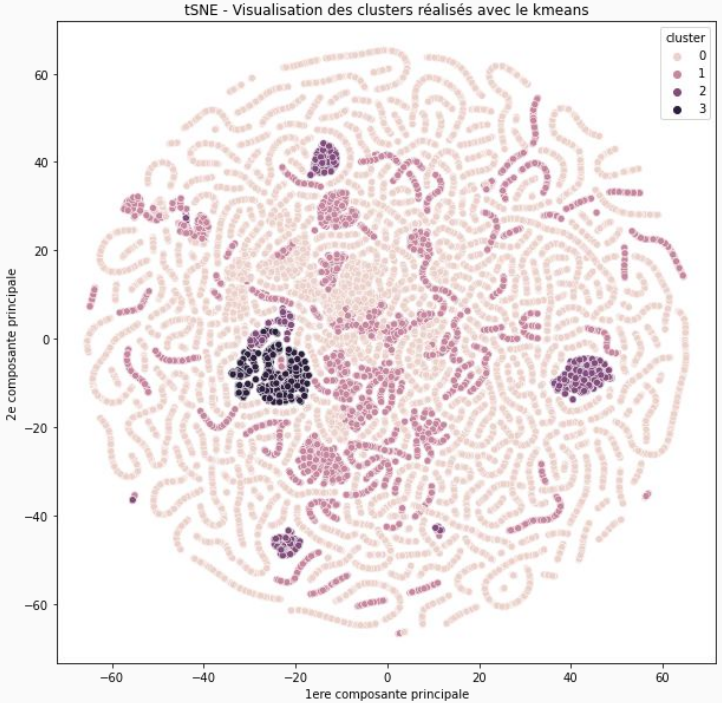


Meilleur équilibrage

4. Modèle final sélectionné

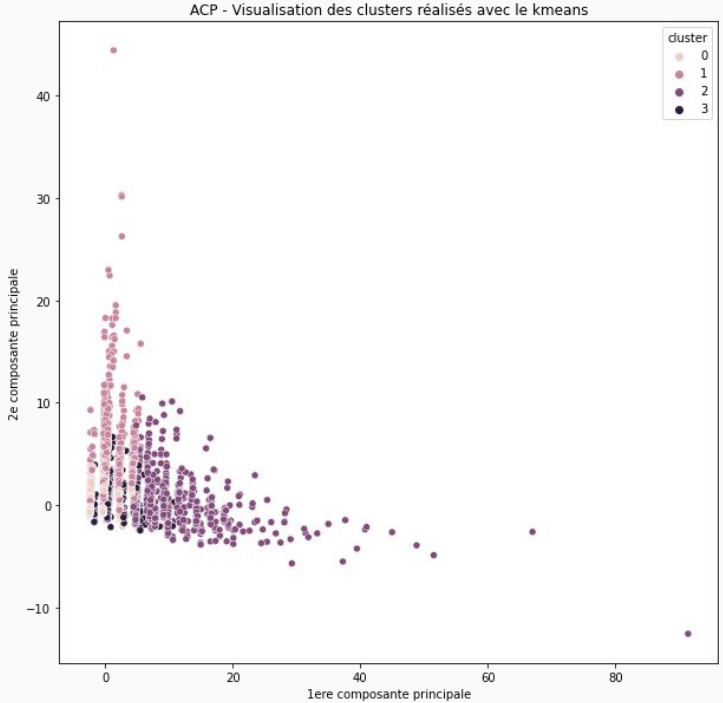
k-means sur les variables numériques

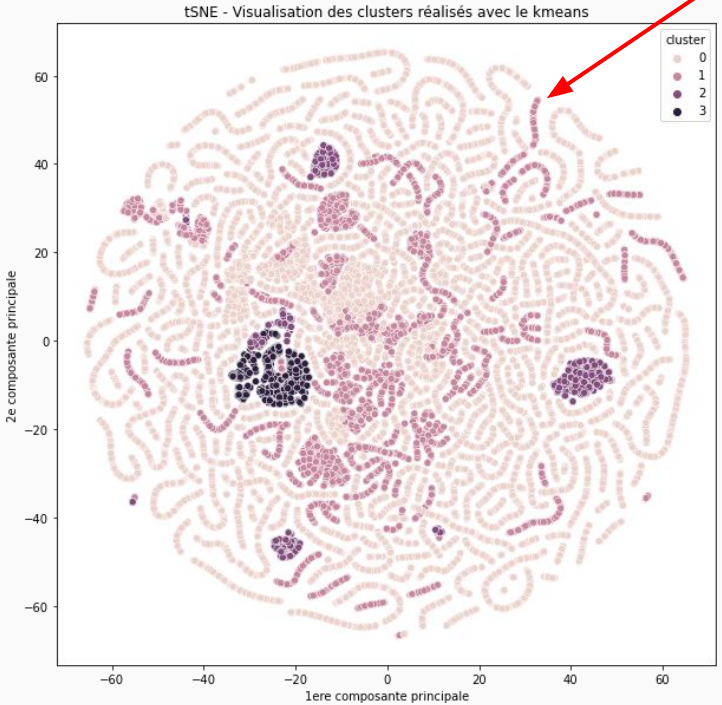




4. Modèle final sélectionné

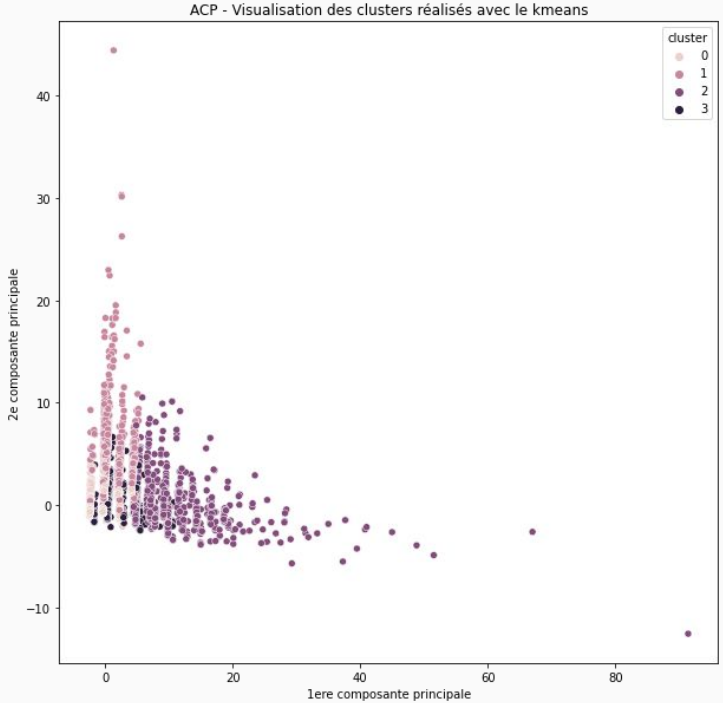
k-means sur les variables numériques

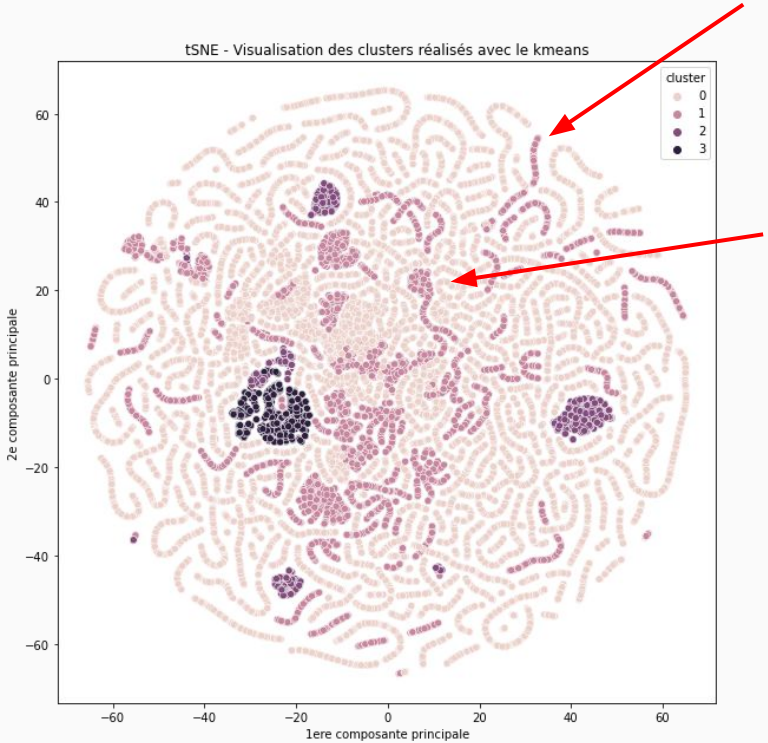




4. Modèle final sélectionné

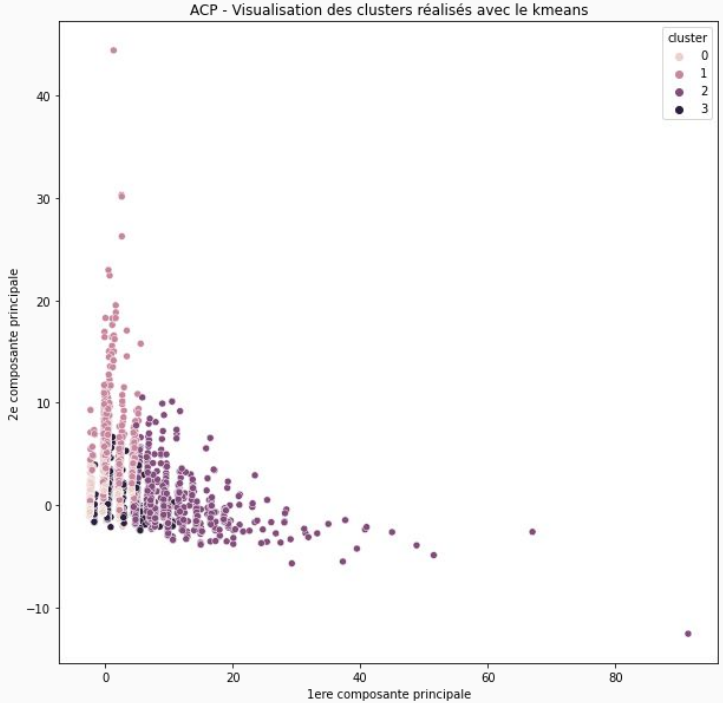
k-means sur les variables numériques

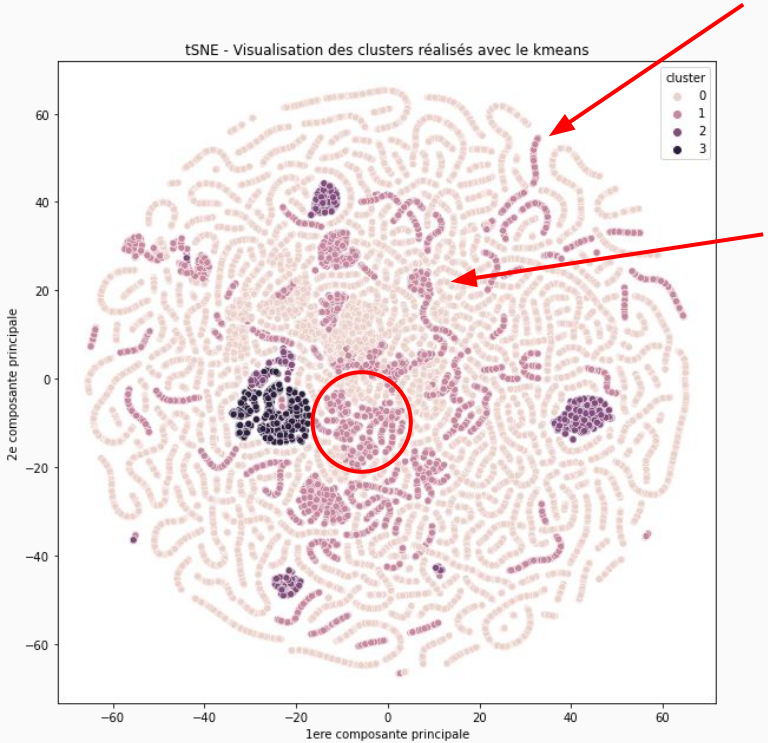




4. Modèle final sélectionné

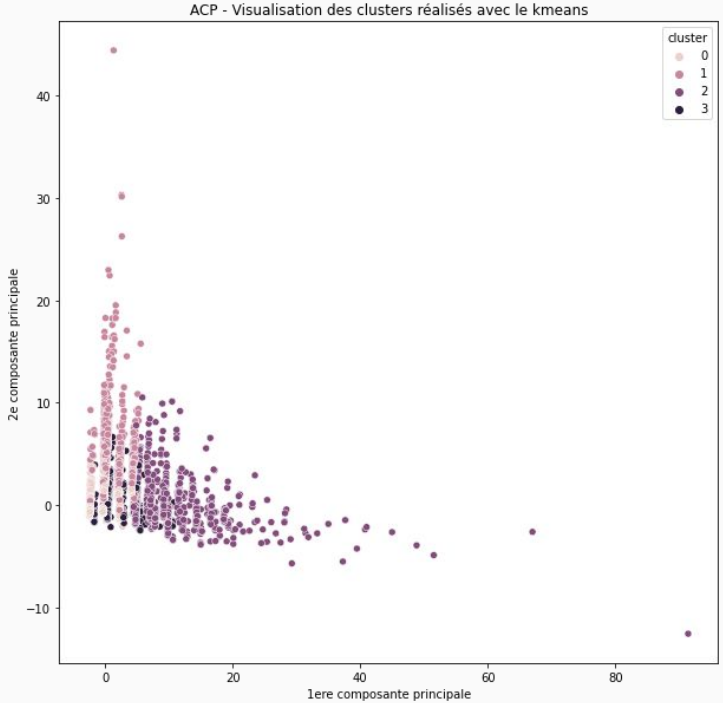
k-means sur les variables numériques

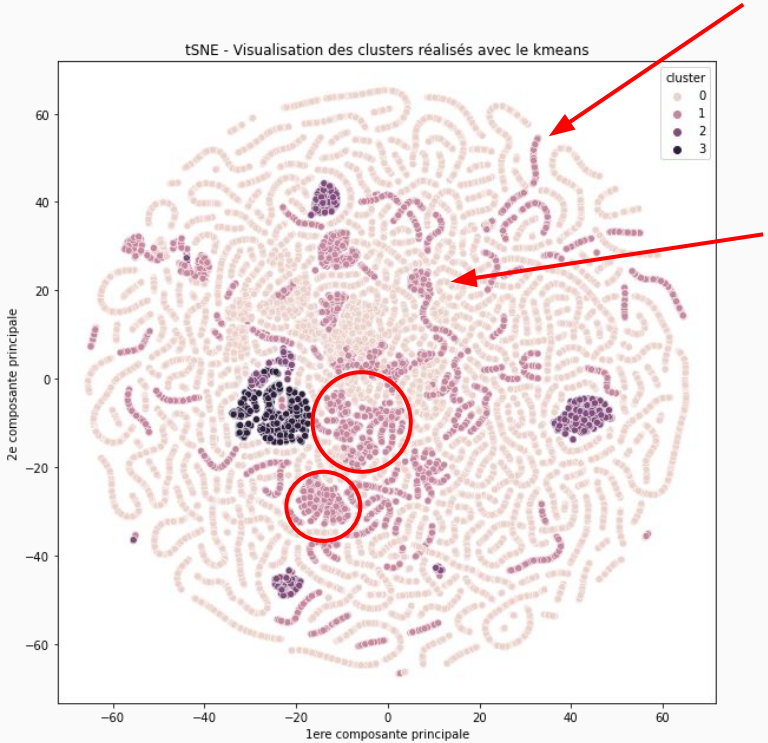




4. Modèle final sélectionné

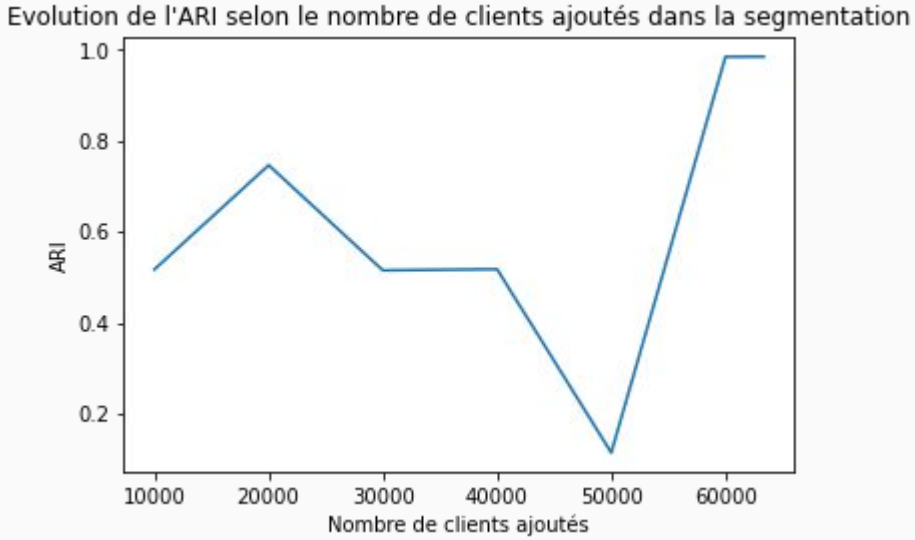
k-means sur les variables numériques





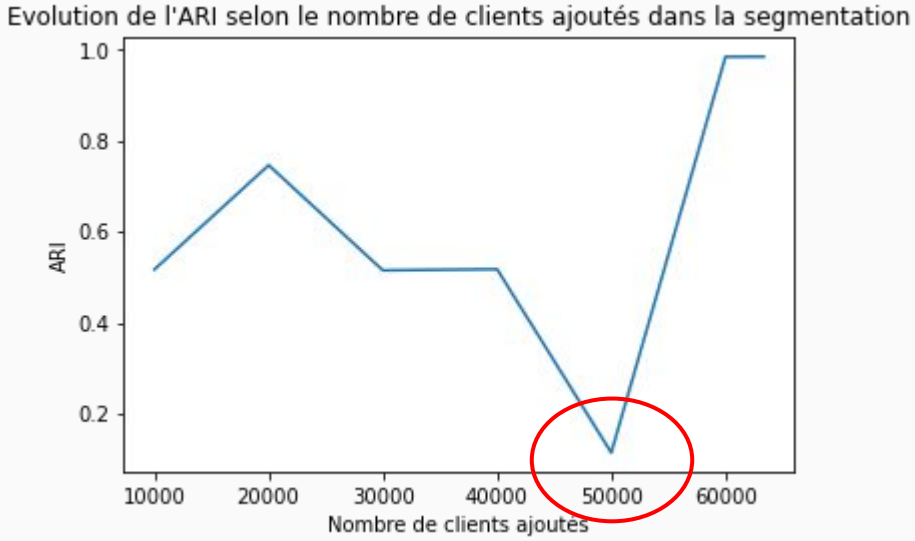
4. Modèle final sélectionné

k-means sur les variables numériques



4. Modèle final sélectionné

k-means sur les variables numériques



Questions/Réponses

Fin.