

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Projet 4 : Segmentez des clients d'un site e-commerce

Claire Gayral
Juillet 2021

Introduction

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- Données dans 9 tables, avec 3 échelles :
 - Client
 - Commande
 - Produit
- Objectif : Segmentation des clients

olist

from <https://www.kaggle.com/olistbr/brazilian-e-commerce>

Introduction - Plan

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

1 La mission

2 Analyse exploratoire

- Echelle Produit
- Echelle Commande
- Echelles Client

3 Modèles de segmentation

- Recency Frequency Moneraty value
- Table des commandes
- Choix du modèle

4 Conclusion

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

1 La mission

2 Analyse exploratoire

3 Modèles de ségmentation

4 Conclusion

Les données

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Les données d'entrée :

- 8 tables anonymisées avec des jointure possibles
- 2017-2018
- En portugais

Les données de sortie :

- Segmentation des clients
- Interprétation des types d'utilisateurs
- Solution actionnable par l'équipe marketing
- Proposition contrat de maintenance

Les données

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

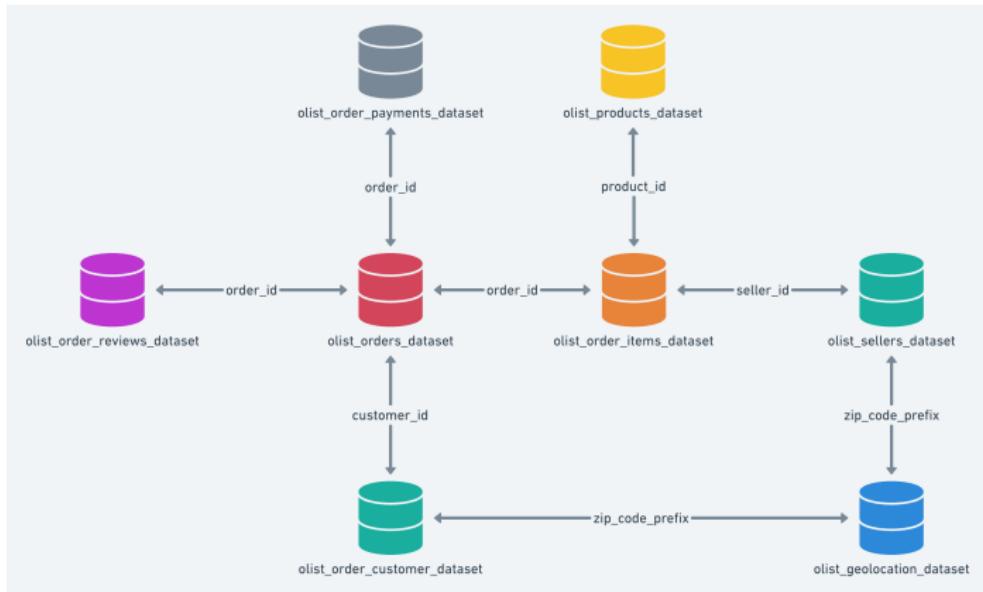
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



1 La mission

2 Analyse exploratoire

- Echelle Produit
- Echelle Commande
- Echelles Client

3 Modèles de segmentation

4 Conclusion

Exploration des tables

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

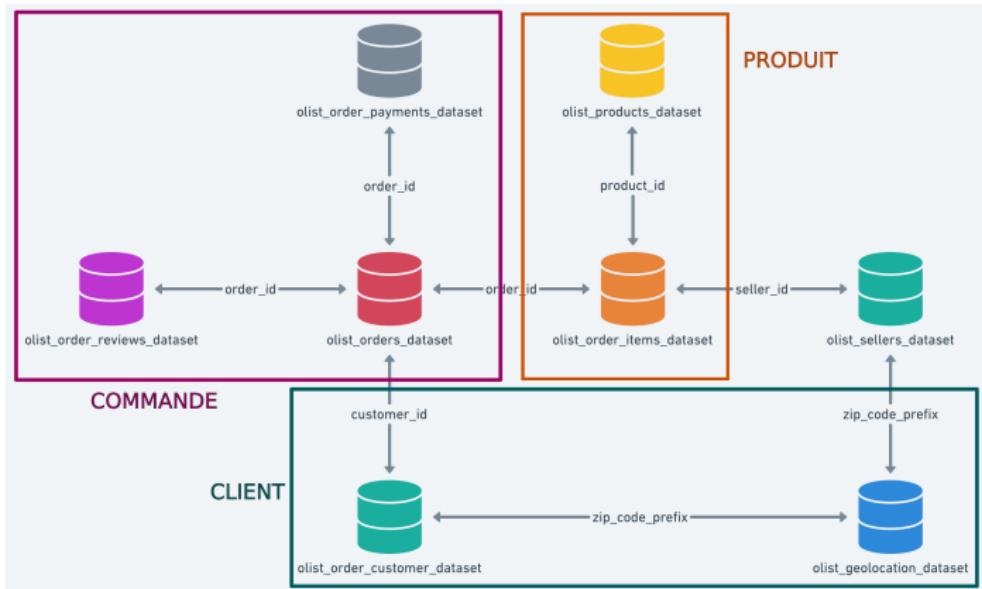
Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Exploration des tables - Produits

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Échelle Produit

Échelle Commande

Echelles Client

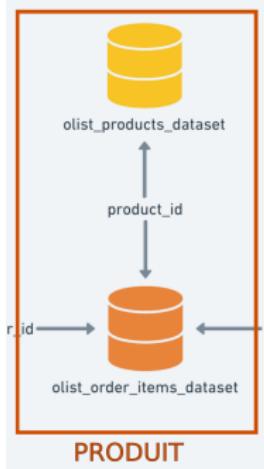
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



- **Table *products* :**

- Description de produit
- Dimension de produit
- Catégorie de produit

- **Table *order items* :**

- Nombre de produit commandé
- Prix des produits
- Fréquence d'achat

Produit - Réduction de dimension

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

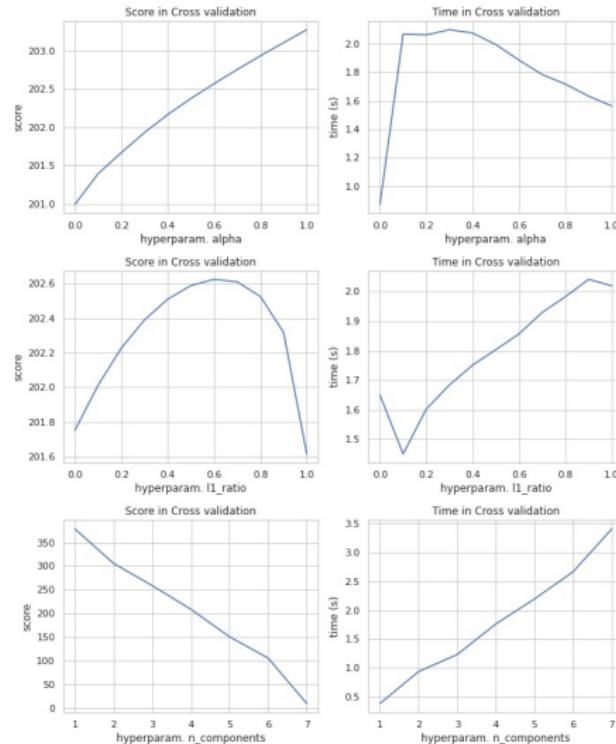
Modèles de segmentation

Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion



Produit - Réduction de dimension

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

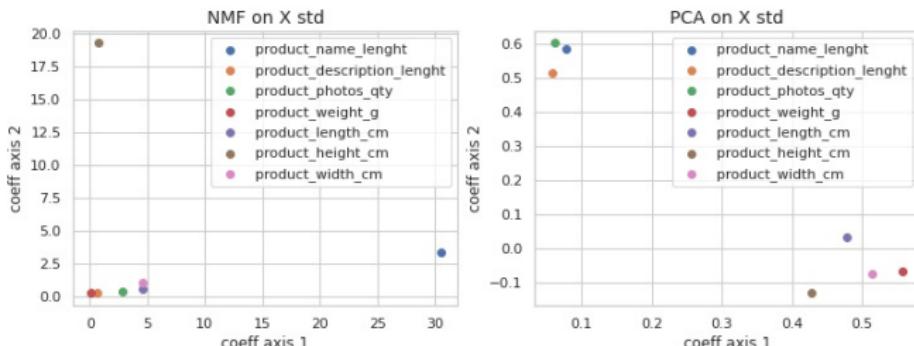
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Produits - nombre de produit commandé

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

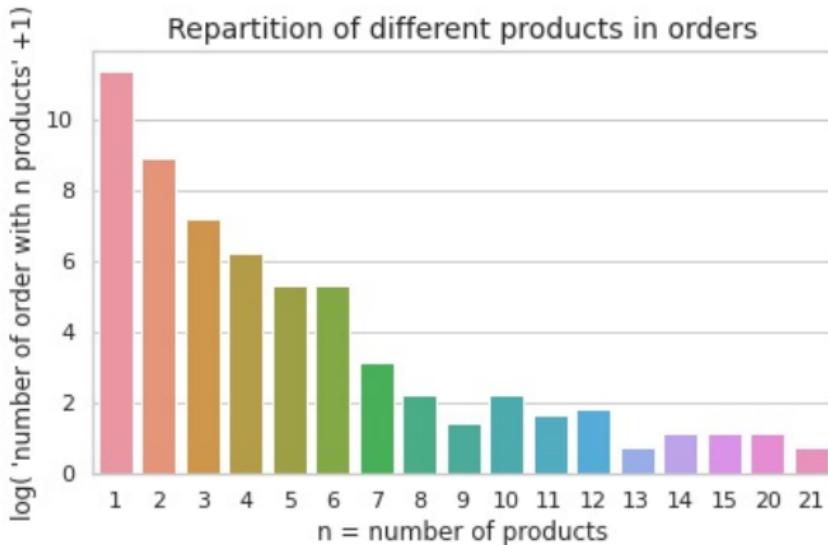
Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Produits - nombre d'apparition dans commandes

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

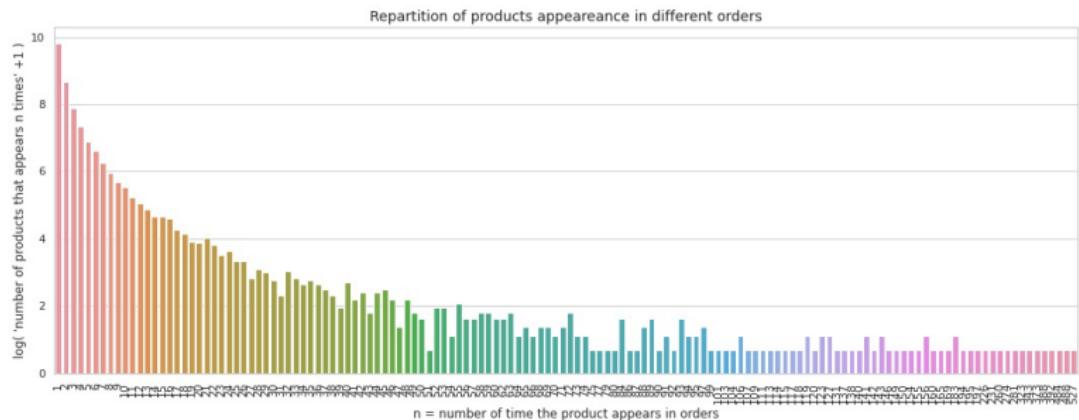
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Produits - Réduction nb catégories

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- Tentative clustering non supervisé
- Séparation par textmining et regroupement si premier mot commun : de 73 à 49 catégories
- Recherche information dans les variables numériques pour faire du target encoding : tentative de sélection des catégories avec une régression lasso sur les variables d'ACP

Pas de structure révélant les catégories mise en évidence.

Produits - Extraction des variables

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

```
1 my_products.dtypes
```

product_id	object
product_category_name	category
product_dimension	float64
product_description	float64
product_freq_buy	float64
product_flag_ordered_alone	int64
product_flag_never_ordered	int64
product_price	float64
product_freight_value	float64
dtype: object	

Exploration des tables - Commandes

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

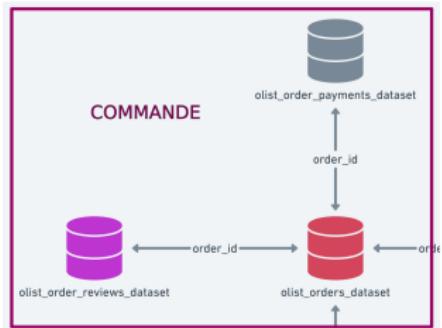
Modèles de segmentation

Recency Frequency Moneraty value

Table des commandes

Choix du modèle

Conclusion



- Table *orders* :

- Status de commande
- Dates d'achat et de livraison

- Table *orders payments* :

- Étalement des paiements
- Type de paiement

- Table *order reviews* (très creuse):

- Score de satisfaction
- Commentaire sur la satisfaction

Commandes - Status de commande

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

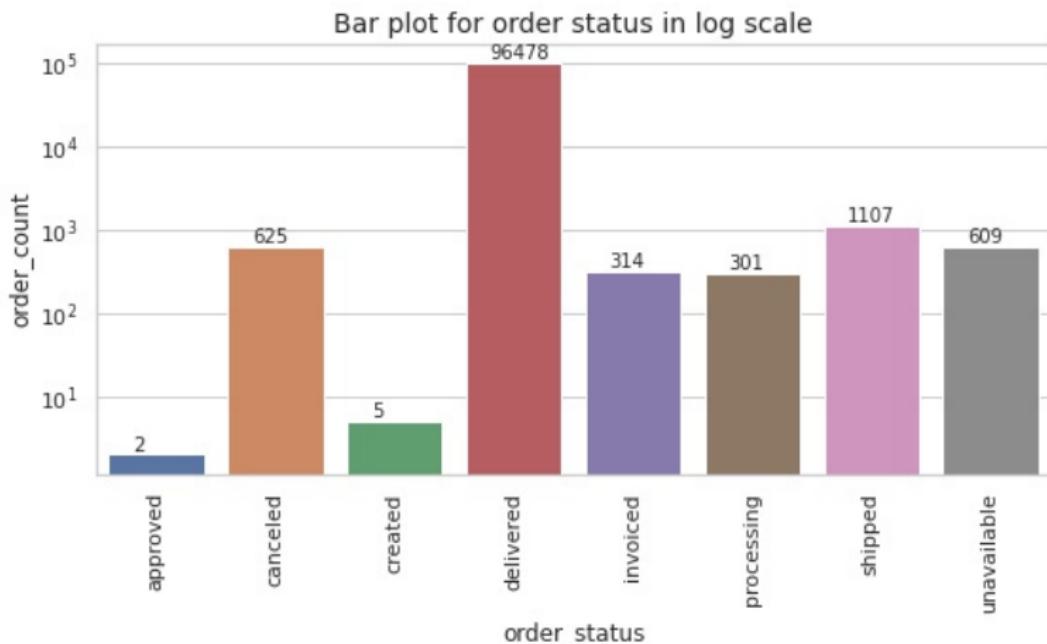
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Commandes - Dates

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

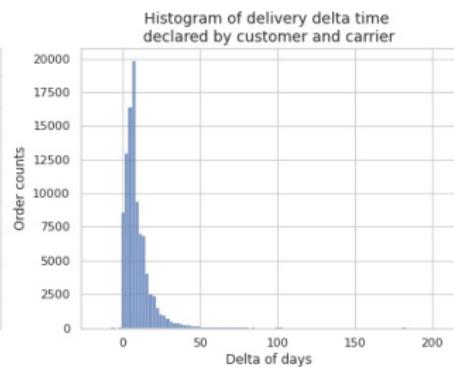
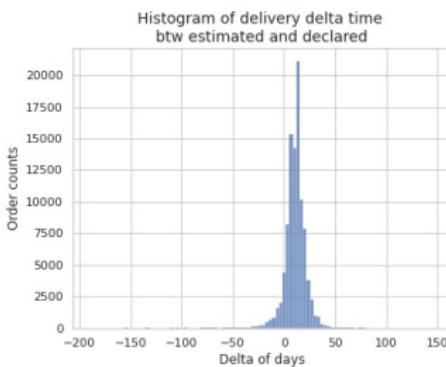
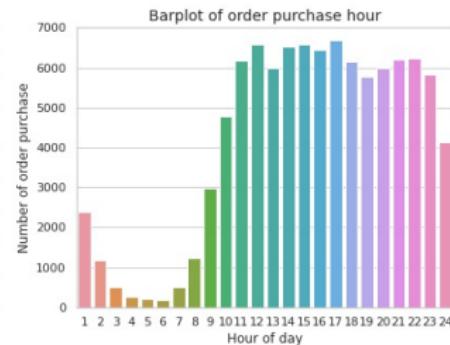
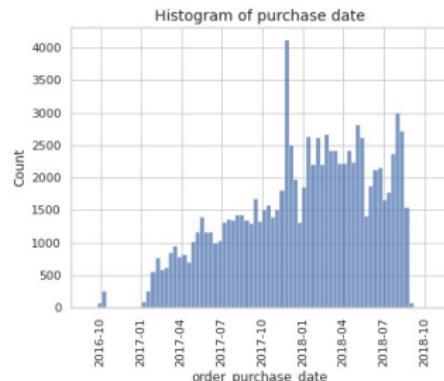
Modèles de ségmentation

Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion



Commandes - Paiements

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Échelle Produit

Échelle Commande

Échelles Client

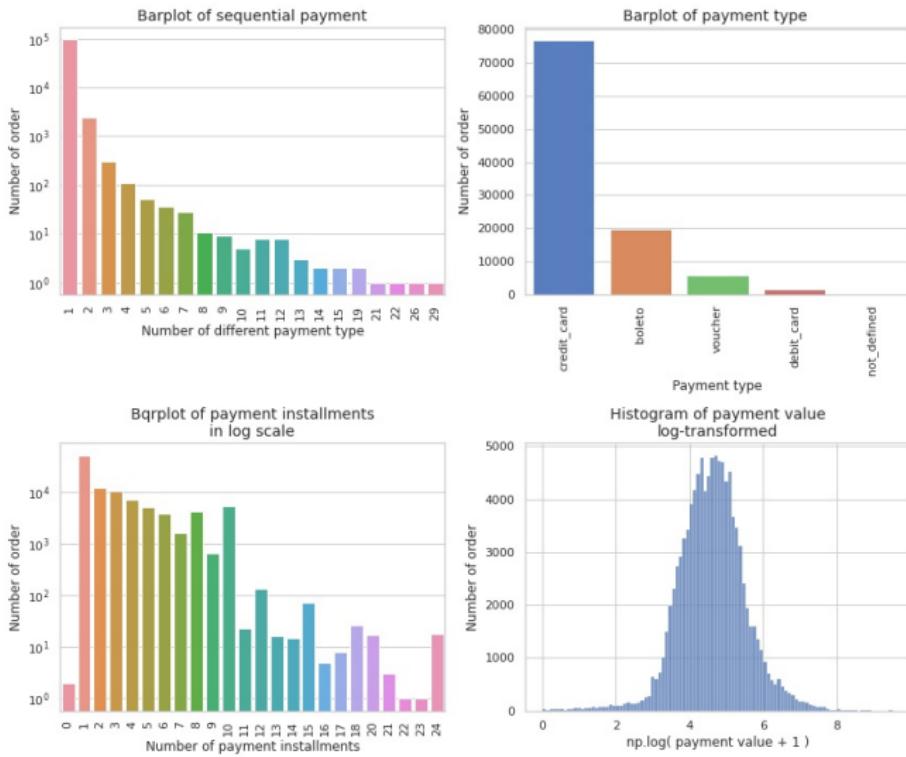
Modèles de segmentation

Recency Frequency
Monetary value

Table des commandes

Choix du modèle

Conclusion



Commandes - Extraction des variables

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

```
1 my_orders.dtypes
```

order_id		object
customer_id		object
order_status		category
order_purchase_date		datetime64[ns]
order_purchase_time		object
order_dt_estimated_declared		float64
order_dt_decalared_customer_carrier		float64
price		float64
freight_value		float64
payment_installments		float64
payment_type		category
nb_payment_type		float64
count_change_scores		int64
review_score		float64
count_review		int64
count_comment_title		int64
count_comment_message		int64
dtype:	object	

Commandes - Prétraitements statistiques

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- **Vectorisation des variables catégorielles**
- Passage à l'échelle logarithmique pour certaines variables
- Retrait des valeurs statistiquement aberrantes
- Inférence des variables manquantes

"One Hot Encoder"

- date → 4 saisons
- heure → matin, après-midi, soir, nuit
- autres variables : type de paiement, ...

Commandes - Prétraitements statistiques

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Échelle Produit

Échelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

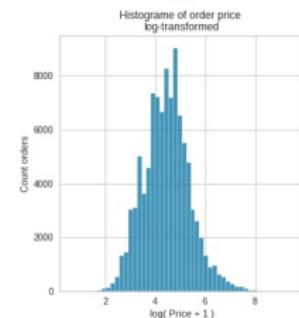
Table des
commandes

Choix du modèle

Conclusion

Exemple avec 'price':

- Vectorisation des variables catégorielles
- **Passage à l'échelle logarithmique pour certaines variables**
- Retrait des valeurs statistiquement aberrantes
- Inférence des variables manquantes



Commandes - Prétraitements statistiques

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- Vectorisation des variables catégorielles
- Passage à l'échelle logarithmique pour certaines variables
- **Retrait des valeurs statistiquement aberrantes**
- Inférence des variables manquantes

Pour chaque variable :

- Si $val \leq Q_{0.05} - 1.5 \times IR$, alors outlier inf et $val \leftarrow Q_{0.05} - 1.5 \times IR$
- Si $val \geq Q_{0.95} + 1.5 \times IR$, alors outlier sup et $val \leftarrow Q_{0.95} + 1.5 \times IR$
- Avec $IR = Q_{0.95} - Q_{0.05}$

Commandes - Prétraitements statistiques

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Échelle Produit

Échelle Commande

Échelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

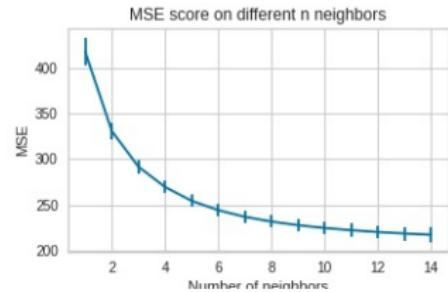
Table des
commandes

Choix du modèle

Conclusion

- Vectorisation des variables catégorielles
- Passage à l'échelle logarithmique pour certaines variables
- Retrait des valeurs statistiquement aberrantes
- **Inférence des variables manquantes**

Par KNN Impute :



$$n_neighbors = 6$$

Exploration des tables - Client

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Échelle Produit

Échelle Commande

Echelles Client

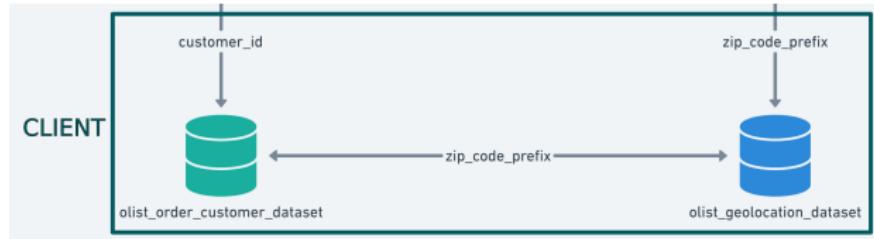
Modèles de segmentation

Recency Frequency
Moneraty value

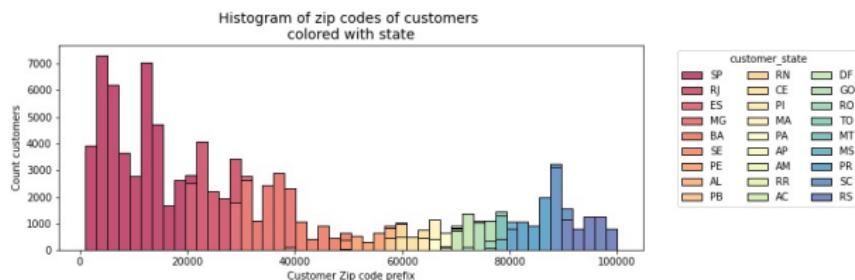
Table des commandes

Choix du modèle

Conclusion



- Table *customers* :
 - Signature unique des consommateurs
 - Information géographique - Zip Code



Client - Analyse RFM

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Échelle Produit

Échelle Commande

Échelles Client

Modèles de
ségmentation

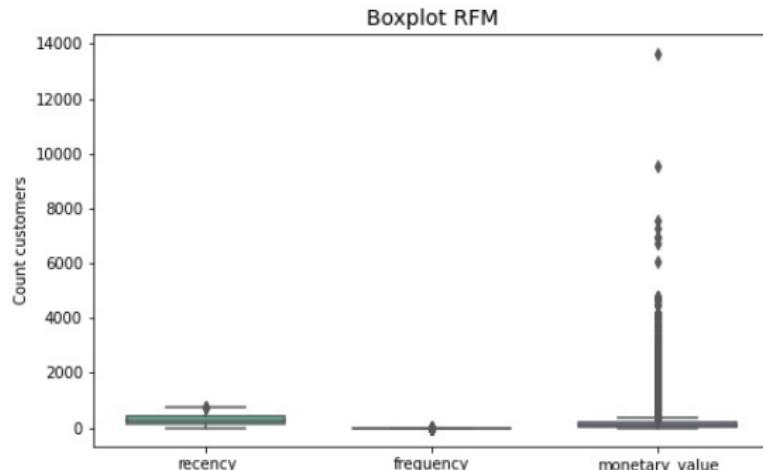
Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- R = Recency :
Quelle est la date du dernier achat pour ce client ?
- M = Frequency :
Quelle est la fréquence d'achat de ce client ?
- F = Monetary Value :
Combien ce client a-t-il dépensé sur le total de ses commandes ?



La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

1 La mission

2 Analyse exploratoire

3 Modèles de segmentation

- Recency Frequency Moneraty value
- Table des commandes
- Choix du modèle

4 Conclusion

Quel type de modèle ?

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Modèles : Clustering non supervisé

- K-means
- DBSCAN
- Classification hierarchique
- Mélange de Gaussiennes

Visualisation : Réduction de dimension non supervisée

- PCA
- NMF
- tSNE

Sur quel type de données ?

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Plusieurs échelles :

- Client : Utiliser les axes RFM
 - 3 dimensions
 - 96096 clients
- Commande : Utiliser le feature engineering de l'analyse exploratoire
 - Grosse hypothèse : Perte de mémoire des clients
 - 35 variables : Beaucoup plus d'information, grande dimension
 - 99441 commandes : jeu de données à pein plus gros

RFM - Projection 3D et 2D

Projet 4 :
Segmentez des clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

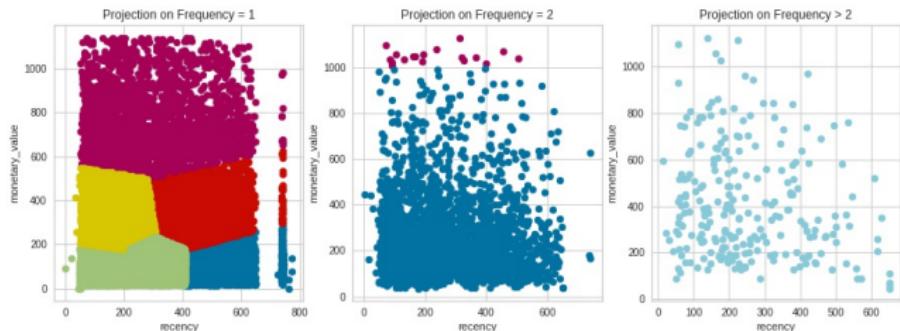
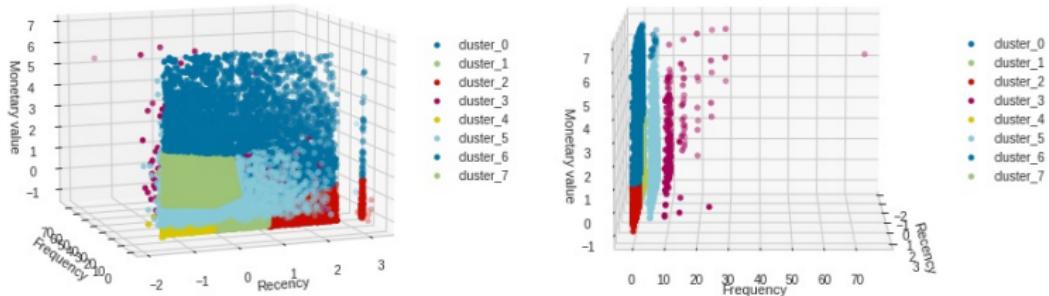
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

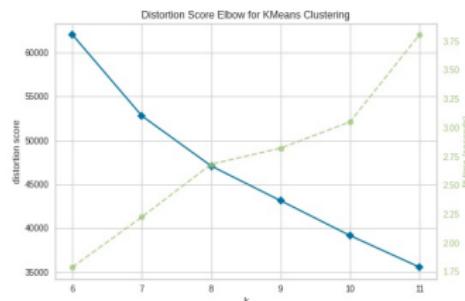
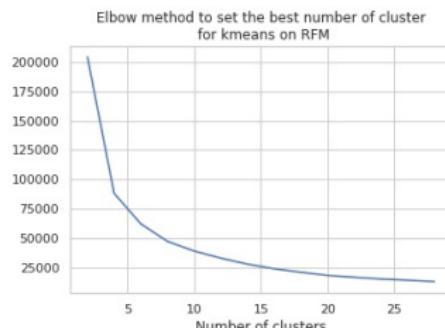
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

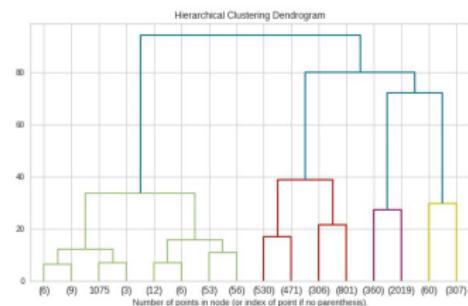
Conclusion

kmeans :



avec Yellowbrick

Clustering hiérarchique



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

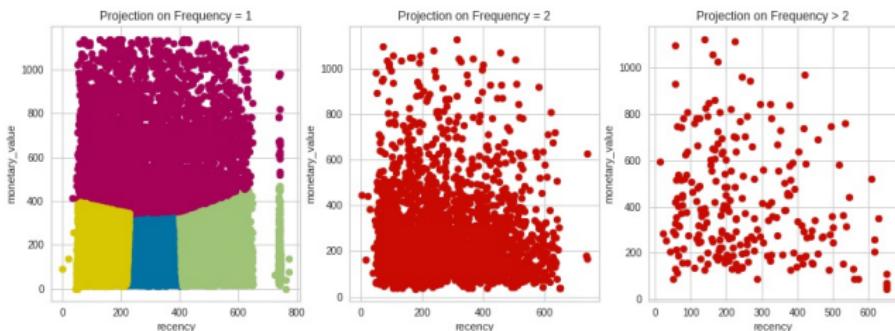
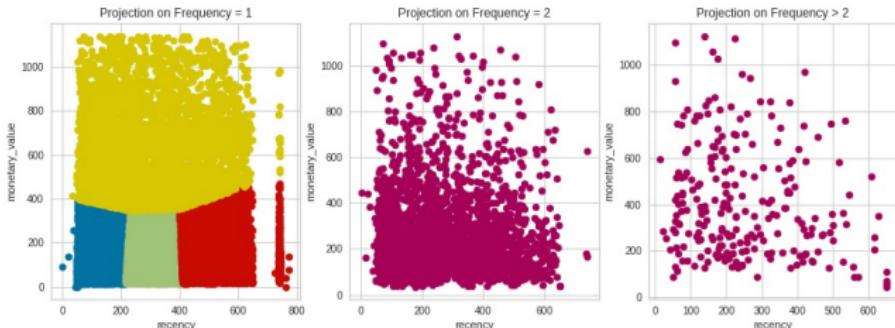
Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Avec 5 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

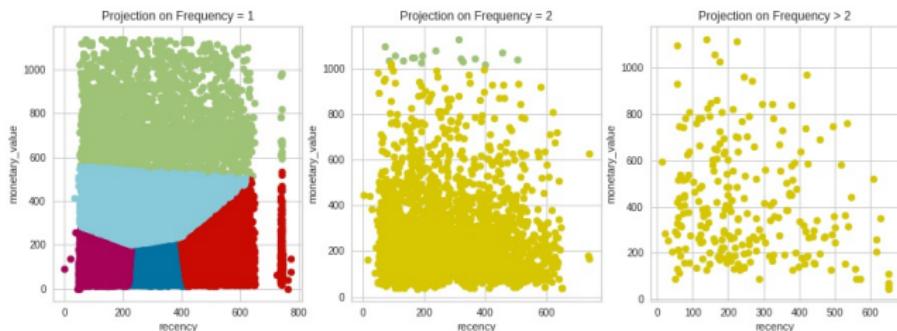
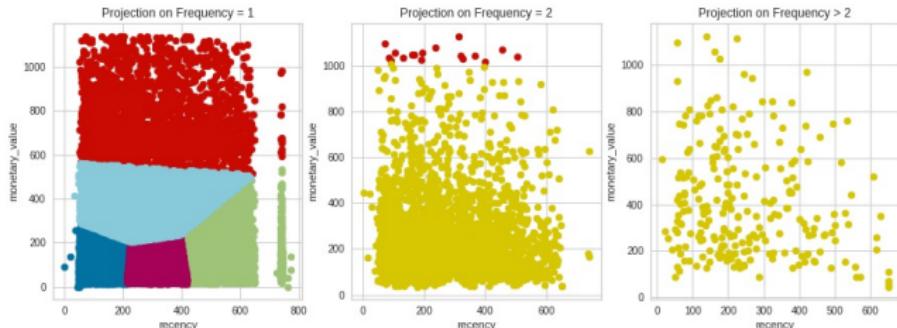
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 6 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

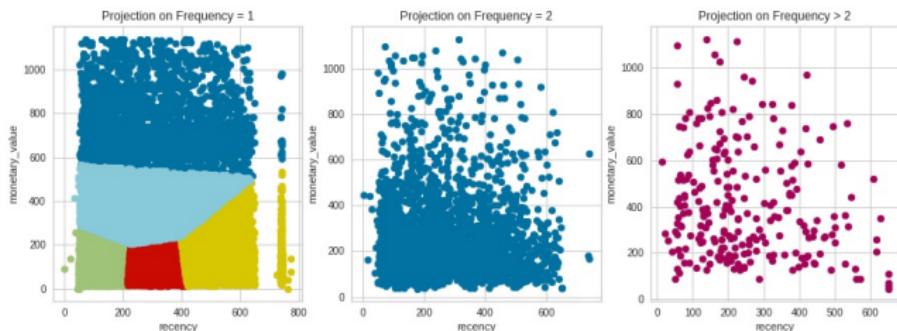
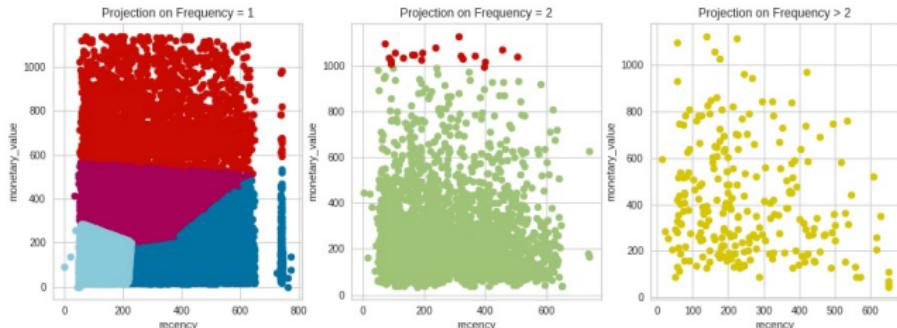
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 7 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

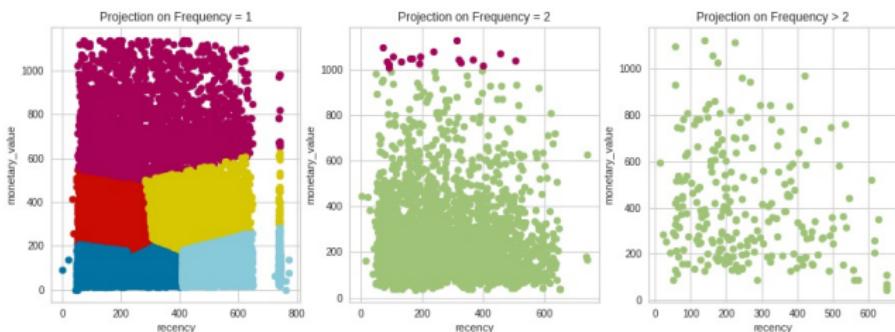
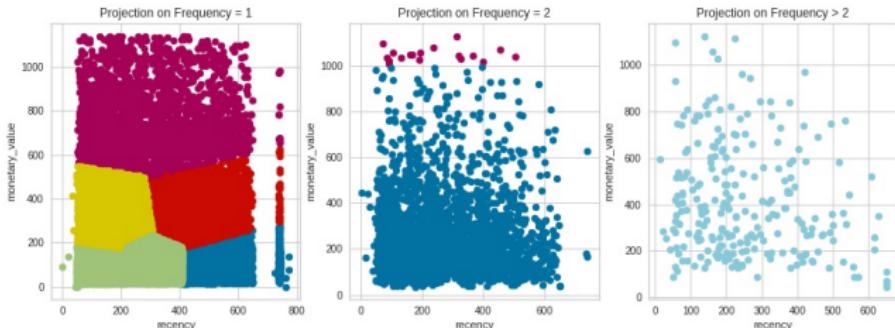
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 8 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

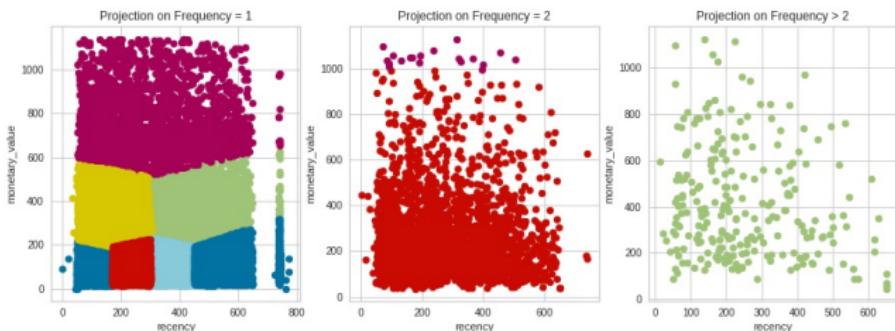
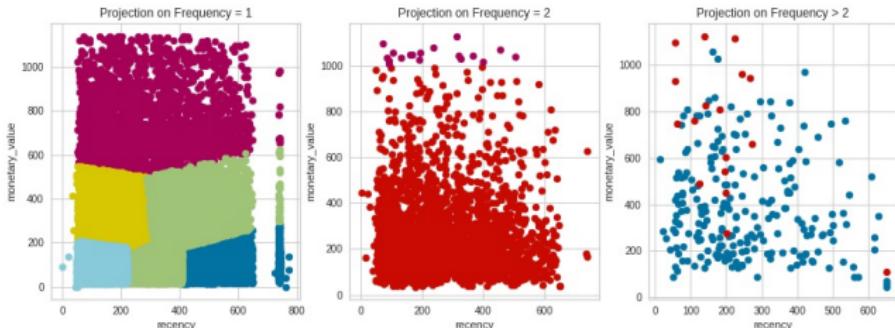
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 9 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

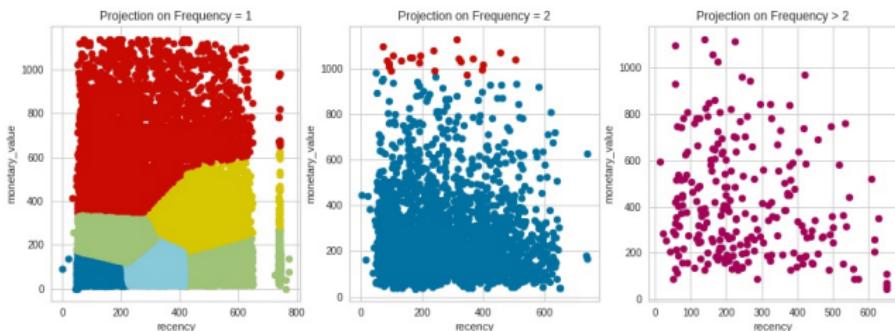
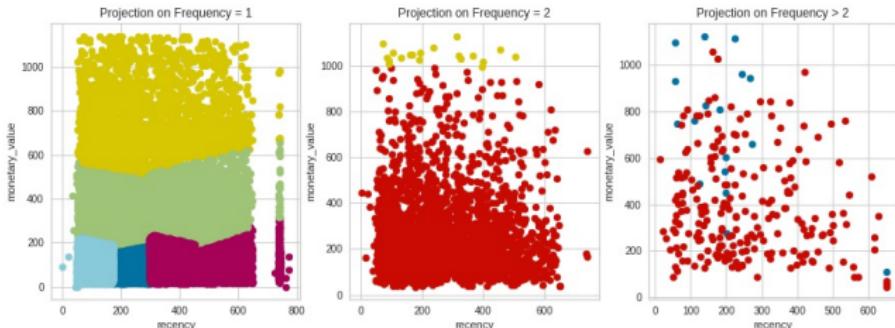
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 10 clusters



RFM - Kmeans et clustering hiérarchique

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

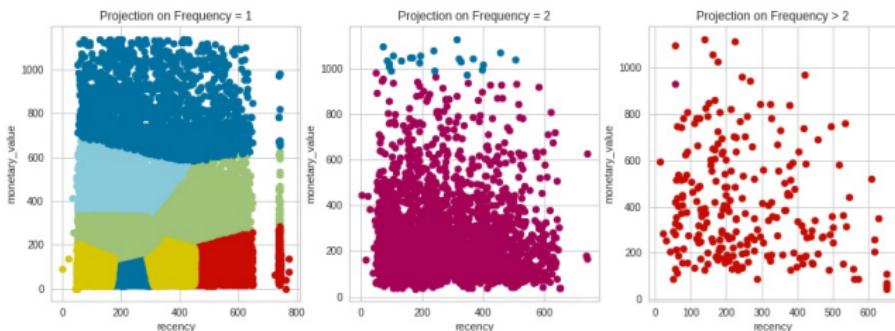
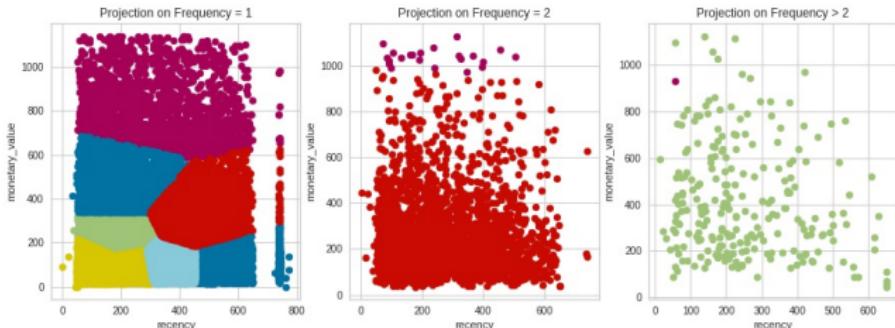
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Avec 11 clusters



RFM - Autres modèles

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de segmentation

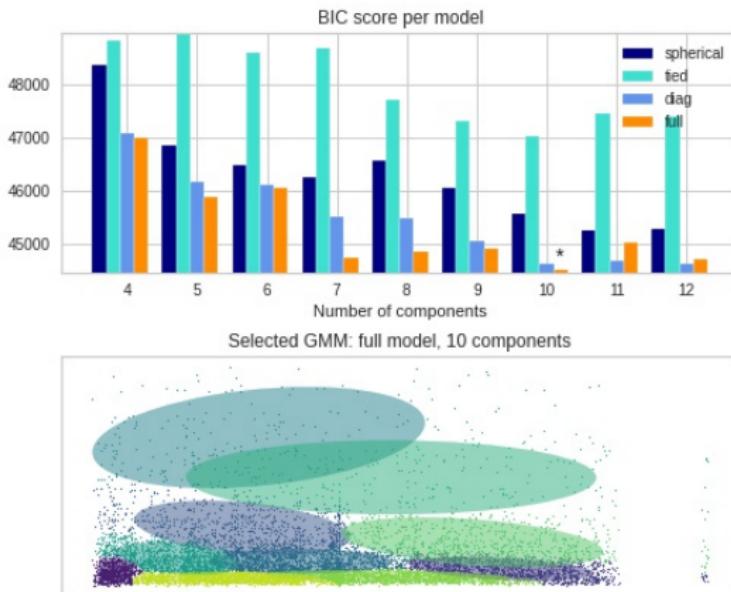
Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion

Gaussian mixture model :



RFM - Autres modèles

DBSCAN :

Projet 4 :
Segmentez des clients d'un site e-commerce

Claire Gayral

La mission

Analyse exploratoire

Echelle Produit

Echelle Commande

Echelles Client

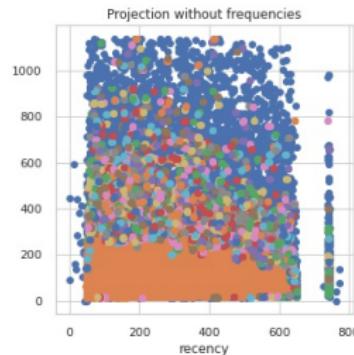
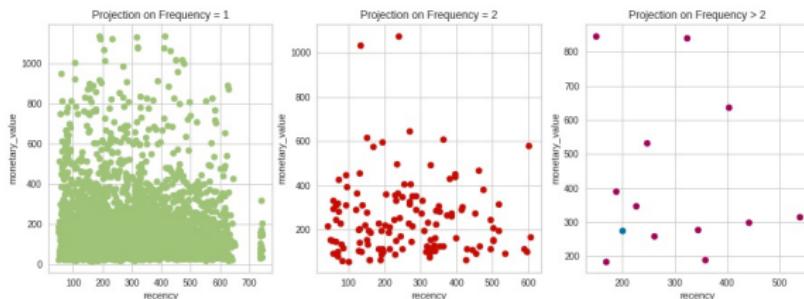
Modèles de segmentation

Recency Frequency
Moneraty value

Table des commandes

Choix du modèle

Conclusion



Commandes - Réduction de dimension

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

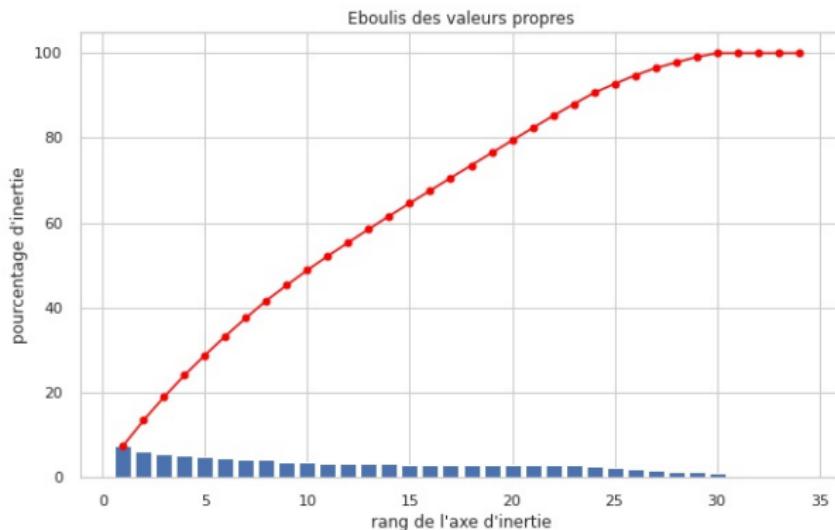
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Commandes - K-Means

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

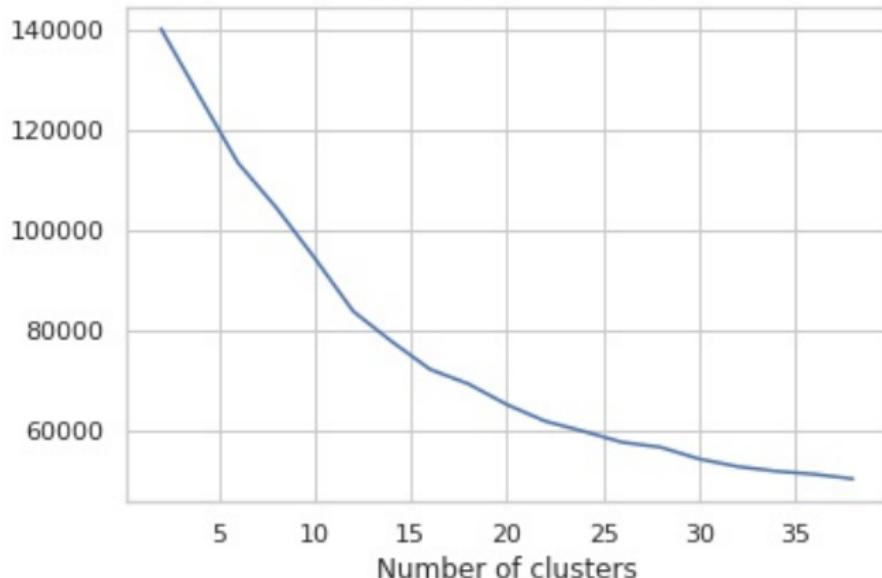
Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Elbow method to set the best number of cluster
for kmeans on orders



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

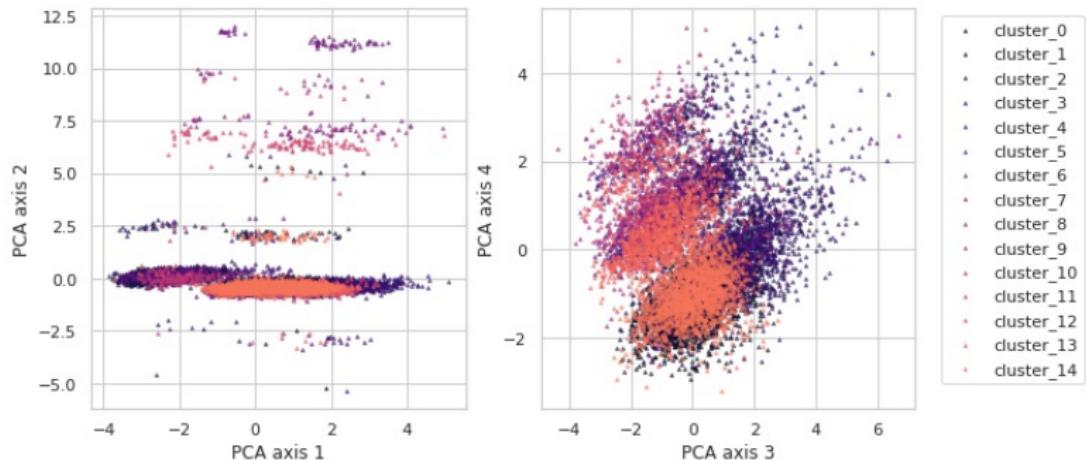
Table des
commandes

Choix du modèle

Conclusion

PCA - n_clusters = 15

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

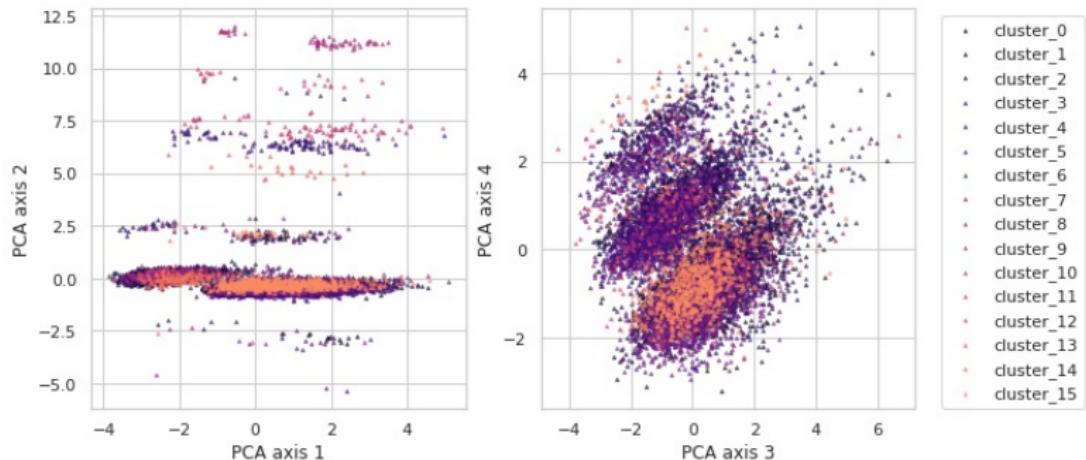
Table des
commandes

Choix du modèle

Conclusion

PCA - n_clusters = 16

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

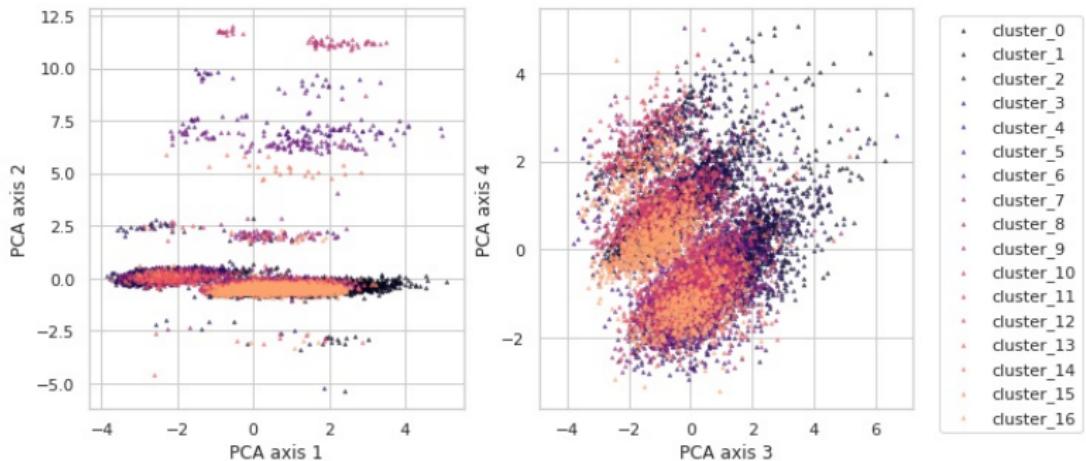
Table des
commandes

Choix du modèle

Conclusion

PCA - n_clusters = 17

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

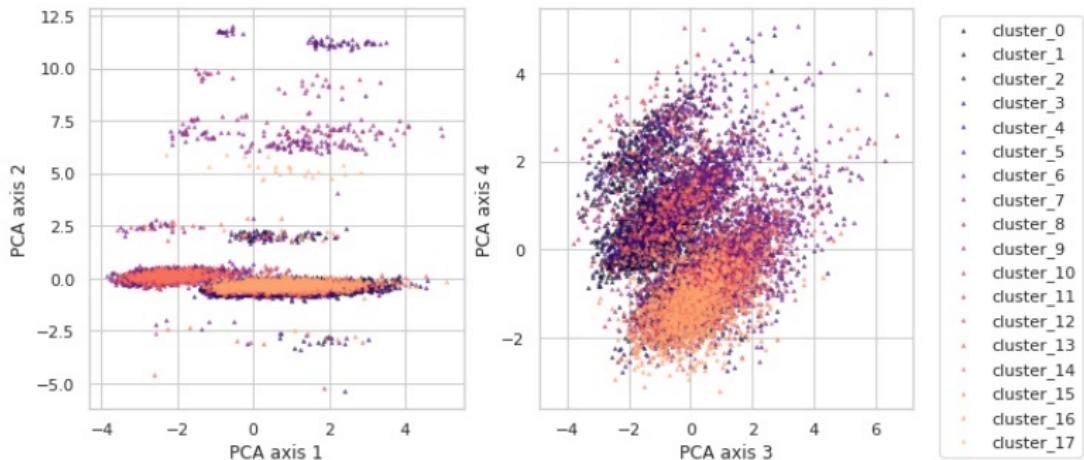
Table des
commandes

Choix du modèle

Conclusion

PCA - n_clusters = 18

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

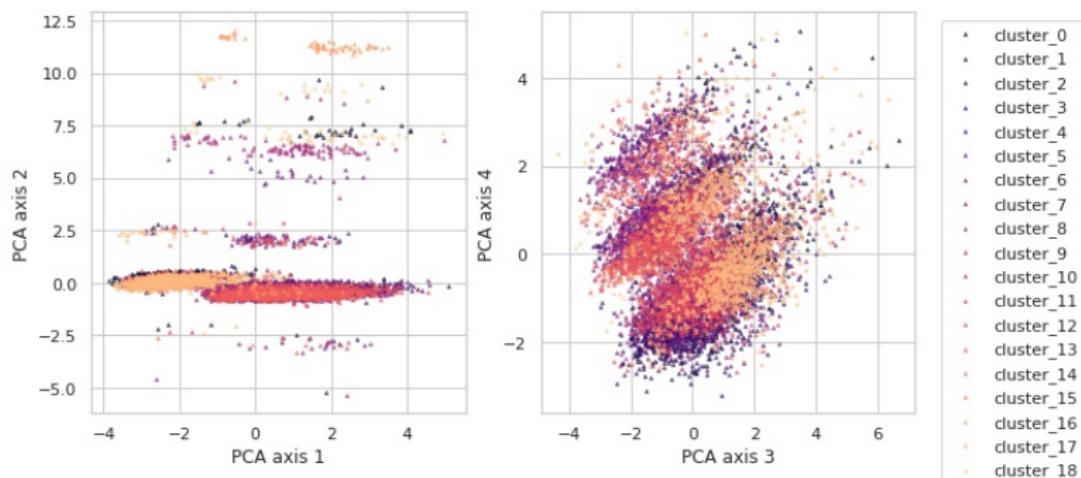
Table des
commandes

Choix du modèle

Conclusion

PCA - n_clusters = 19

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

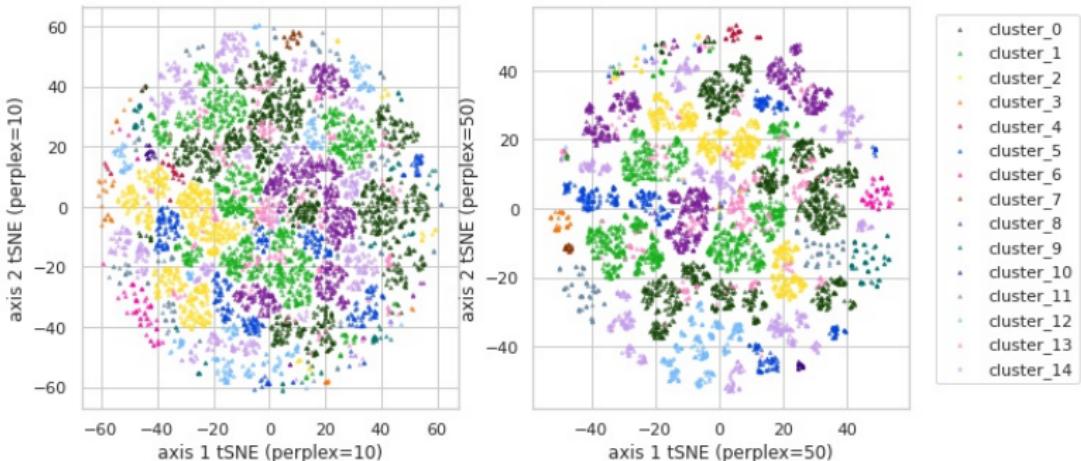
Table des
commandes

Choix du modèle

Conclusion

tSNE - n_clusters = 15

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

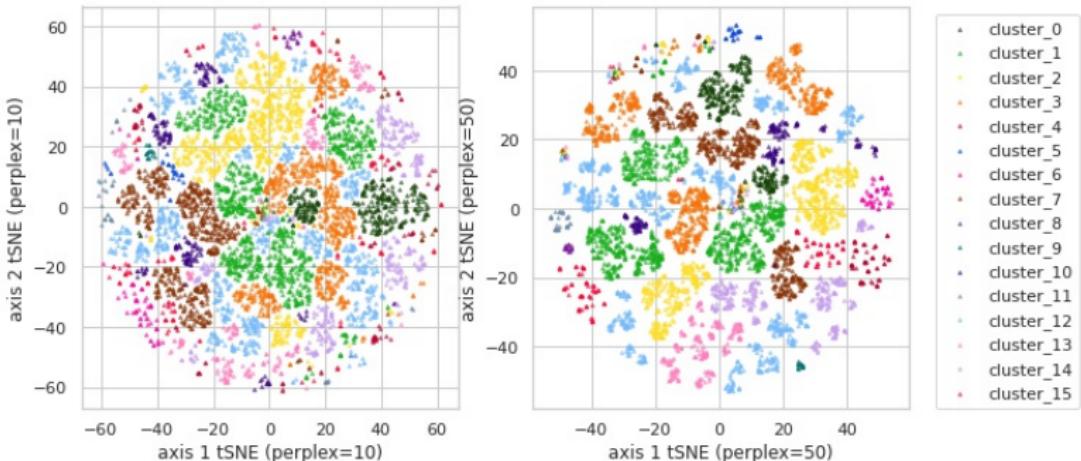
Table des
commandes

Choix du modèle

Conclusion

tSNE - n_clusters = 16

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

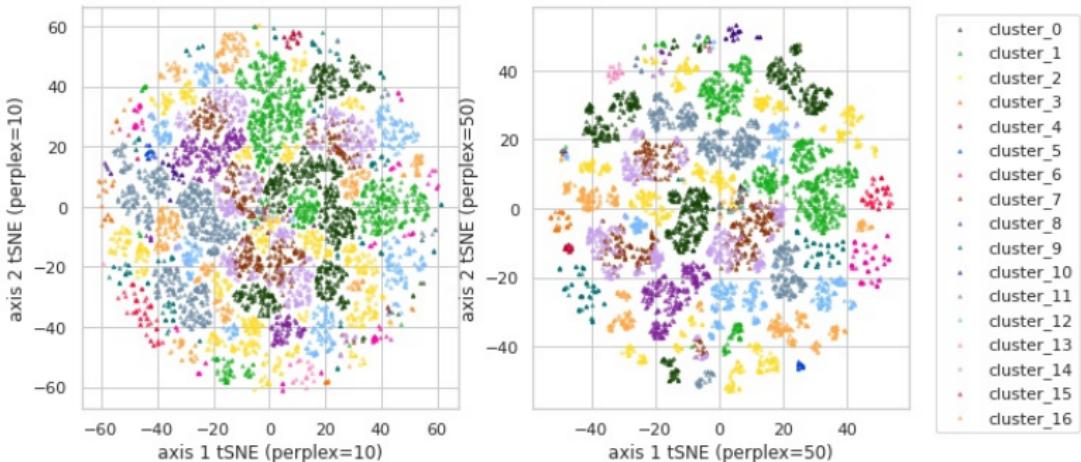
Table des
commandes

Choix du modèle

Conclusion

tSNE - n_clusters = 17

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

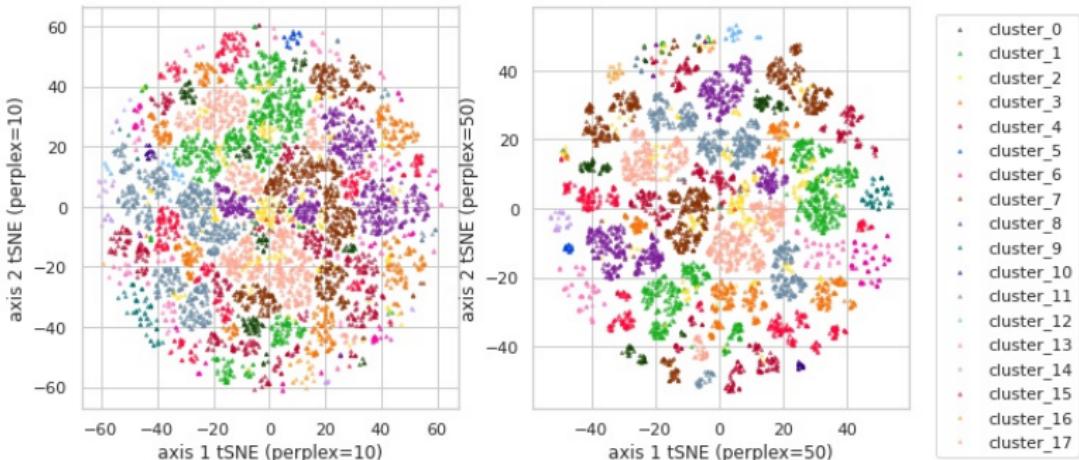
Table des
commandes

Choix du modèle

Conclusion

tSNE - n_clusters = 18

Clustering on projected PCA axes



Commandes - Visualisation du clustering par PCA

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

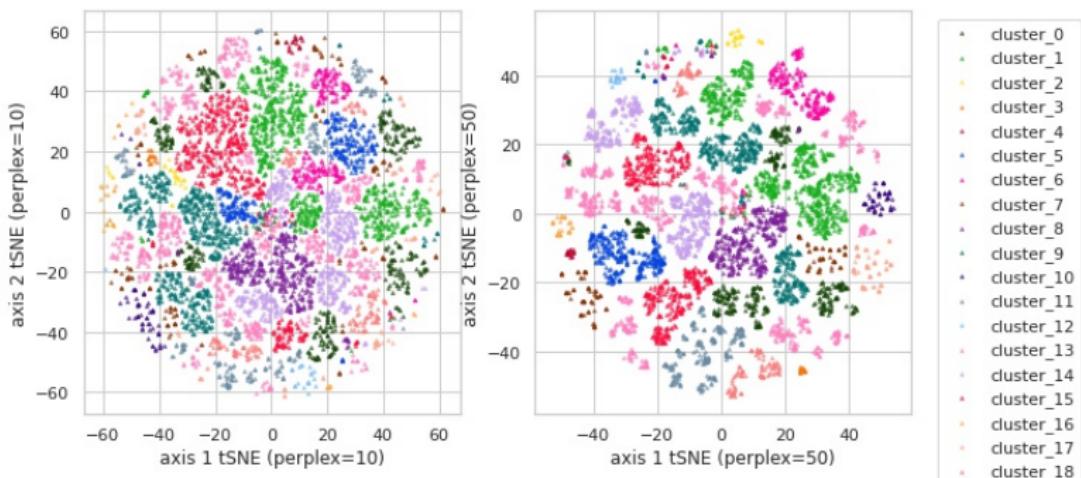
Table des
commandes

Choix du modèle

Conclusion

tSNE - n_clusters = 19

Clustering on projected PCA axes



Commandes - Pistes d'amélioration

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Échelle Produit

Échelle Commande

Échelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

- Utiliser une NMF plutôt qu'une ACP comme réduction de dimension
- Utiliser d'autres méthodes de clustering : classification hierarchique, Gaussian Mixture Models
- Lancer le clustering directement dans l'espace réduit de tSNE

Meilleur modèle - Choix des données

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

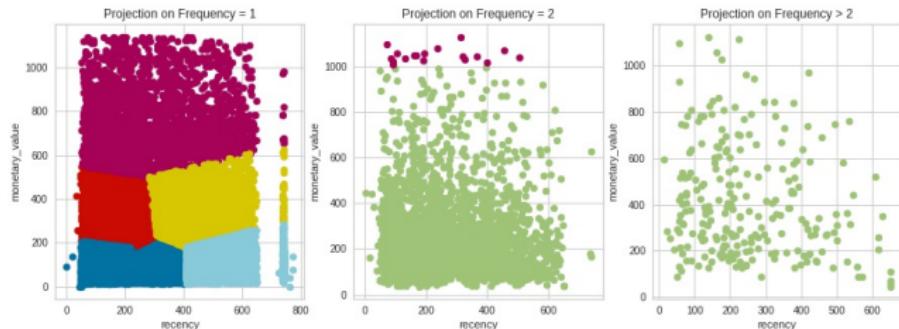
Table des
commandes

Choix du modèle

Conclusion

RFM - 8 clusters

- Consigne "facile d'utilisation pour l'équipe marketing"
- Séparation claire des clusters : représentation 3D
- Plus petit nombre de clusters : 8 VS 20
- Pas de pré-traitement



Meilleur modèle - Choix du modèle

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

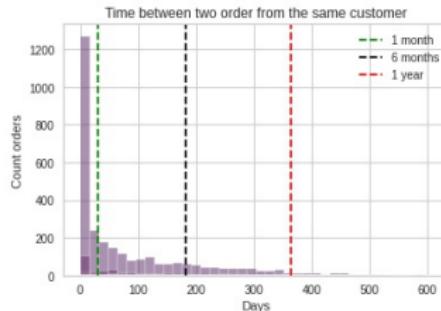
Table des
commandes

Choix du modèle

Conclusion

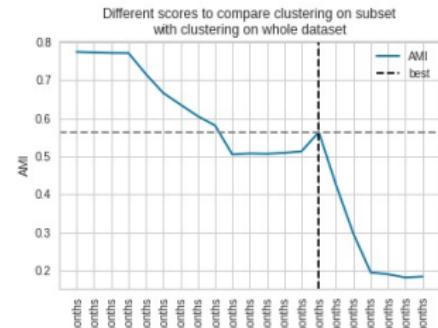
Prop 1 : classification hiérarchique

- Ajout online de nouveaux clients (coût mise en prod)
- Mise à jour bi-annuelle
- Stockage matrice distances



Prop 2 : K-Means

- Coût computation plus faible
- Mise à jour tous les 10 mois
- Performance $AMI = 58\%$



Meilleur modèle - Interprétation

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

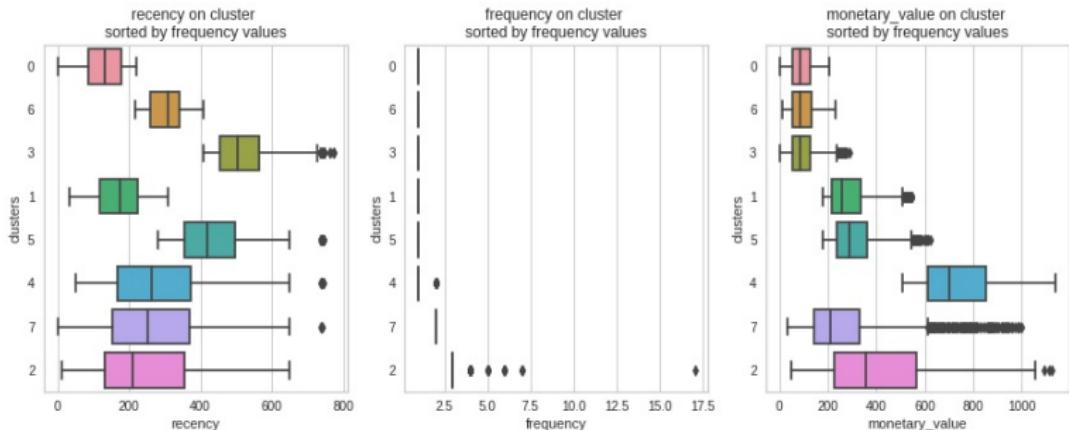
Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion



Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
ségmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

1 La mission

2 Analyse exploratoire

3 Modèles de ségmentation

4 Conclusion

Conclusion

Projet 4 :
Segmentez des
clients d'un site
e-commerce

Claire Gayral

La mission

Analyse
exploratoire

Echelle Produit

Echelle Commande

Echelles Client

Modèles de
segmentation

Recency Frequency
Moneraty value

Table des
commandes

Choix du modèle

Conclusion

Résumé

- Une analyse sur 3 échelles
- Deux façons de modéliser le problème
- Meilleur modèle = interprétabilité + facilité de calculs

Améliorations et suite :

- Finir de faire un script avec le meilleur modèle en PEP8
- Utiliser la table "orders" pour caractériser les clusters de clients

Merci pour votre écoute !