

# Segmentez des clients d'un site e-commerce

Parcours Data Scientist  
Soutenance de Projet 5  
02 déc 2021



# 01

## CONTEXTE

Segmentation des clients pour une e-commerce

# 02

## DONNÉES

Les données des commandes, ses analyses et ses transformations

# 03

## MODÈLES

Constructions des modèles

# 04

## Stabilité

Analyse de la stabilité des segments



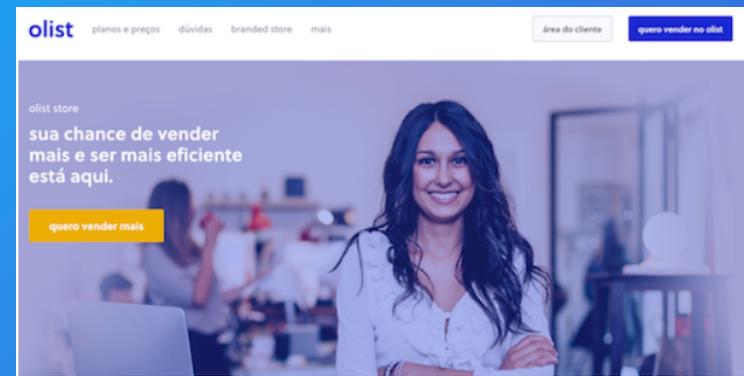
# 01. CONTEXTE

Segmentation des clients pour  
une e-commerce

# CONTEXTE

Une segmentation des clients pour leurs campagnes de communication de quotidien.

Une proposition de contrat de maintenance.



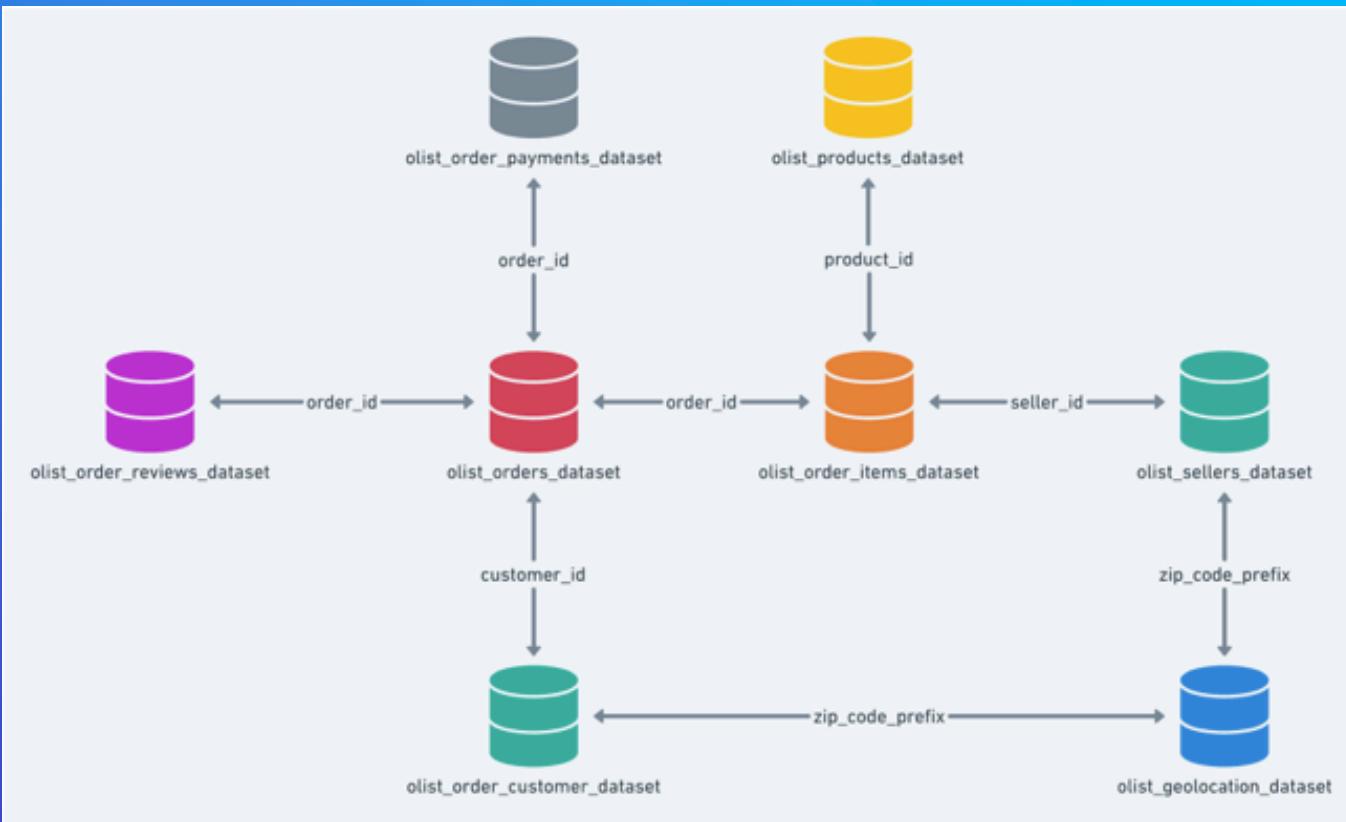
**Olist store**  
E-Commerce



## 02. DONNÉES

Les données des commandes, ses analyses et ses transformations

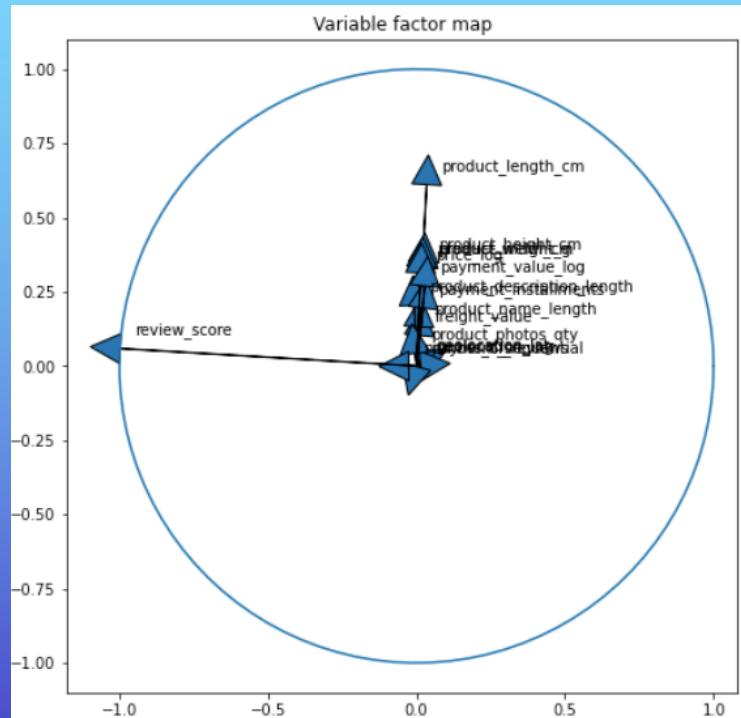
# Database Schema



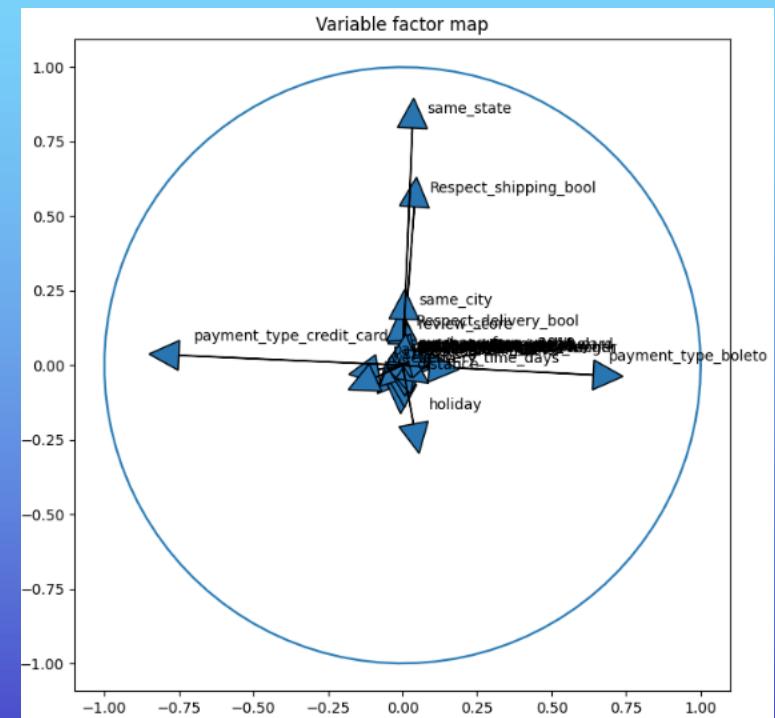
# Preprocessing & Feature Engineering

- Regrouper les catégories produits de 74 à 20
- Passage au log
- Transformation de 'latitude' et 'longitude' en distance
- Freight\_value to price ratio
- Delai Livraison
- Respect la Date de Livraison Estimée
- Volume d'Article
- Ratio de 'freight\_value' à Volume d'Article
- Nombre de produits par achat
- Fréquence d'Achat d'un Client par Mois
- Même Ville Entre l'Acheteur et le Vendeur
- Même Etat Entre l'Acheteur et le Vendeur
- Week-end VS. Semaine
- Vacances

# Analyses PCA – MinMax Scaler

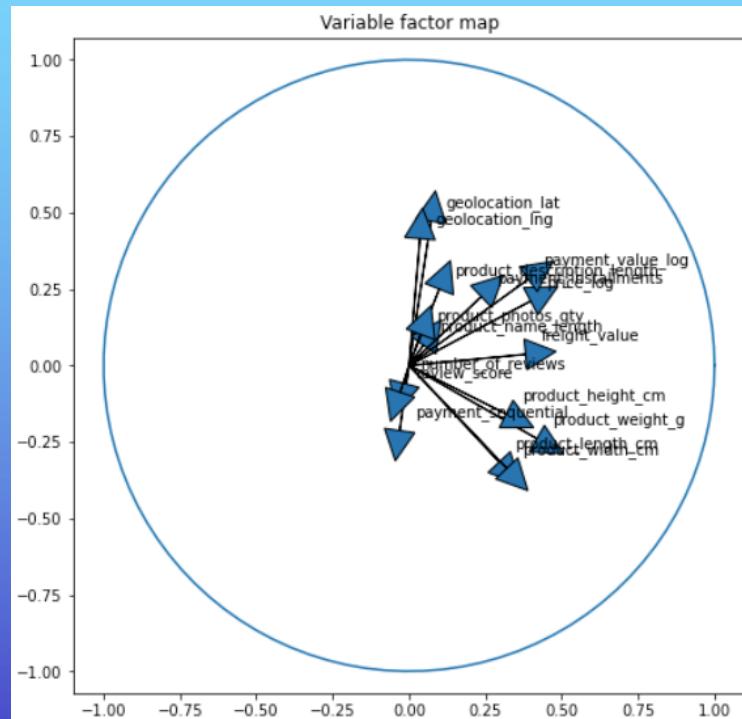


AVANT

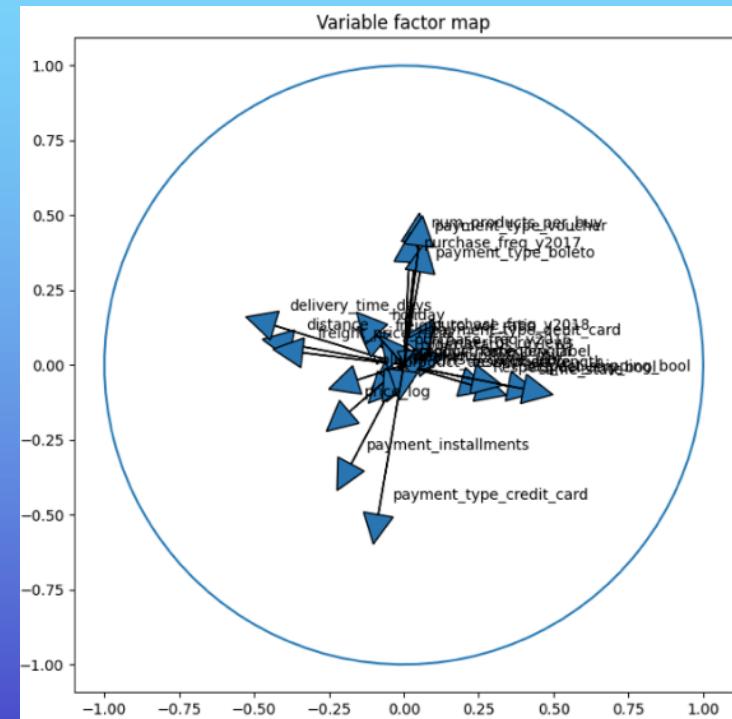


APRÈS

# Analyses PCA – Standard Scaler



AVANT



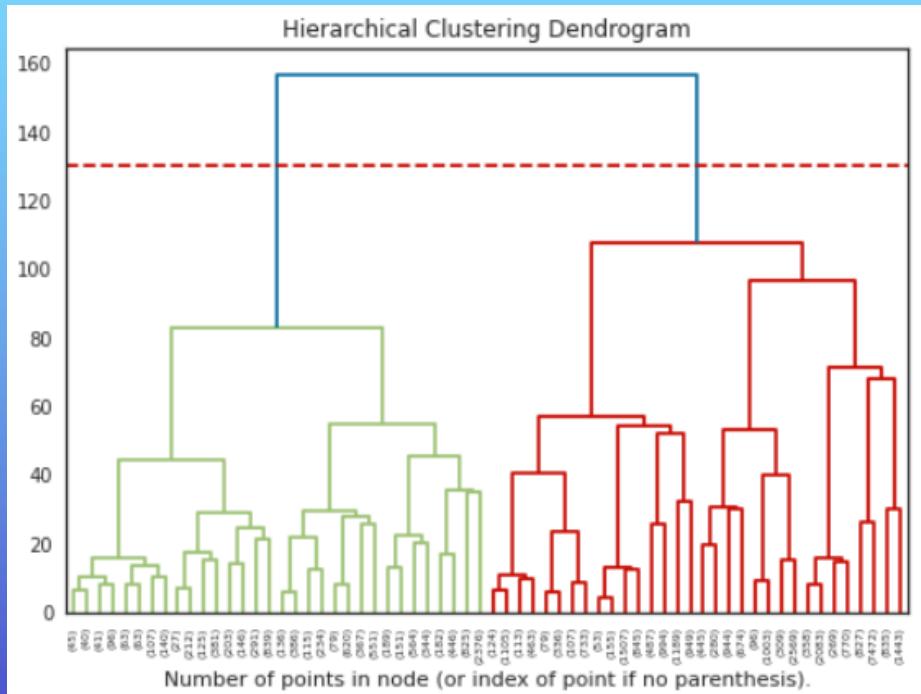
APRÈS



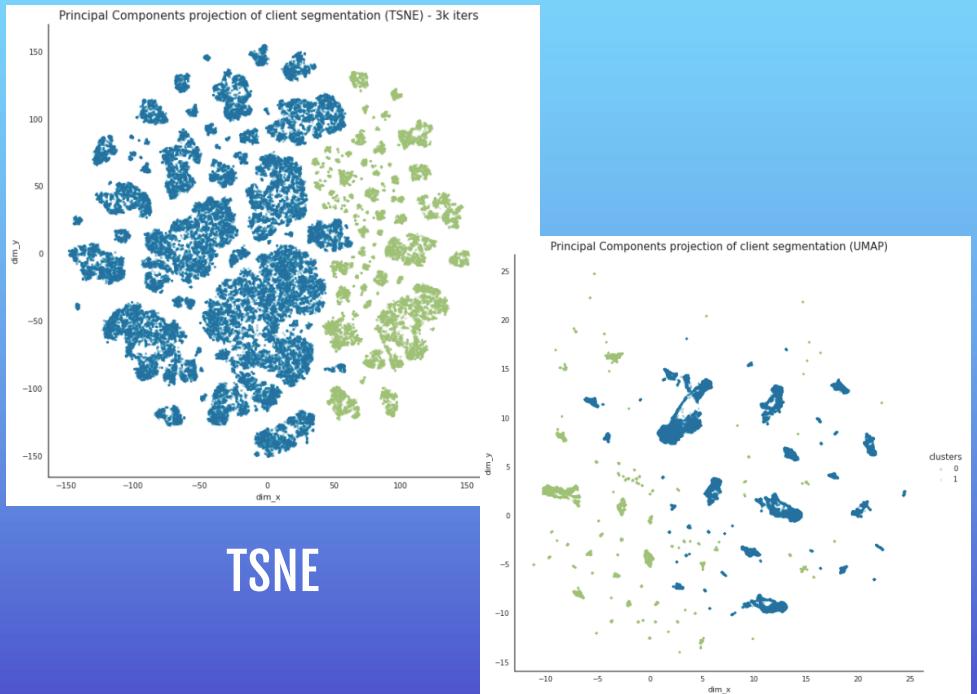
## 03. MODÈLES

Constructions des modèles et  
analyse de la stabilité des  
segments

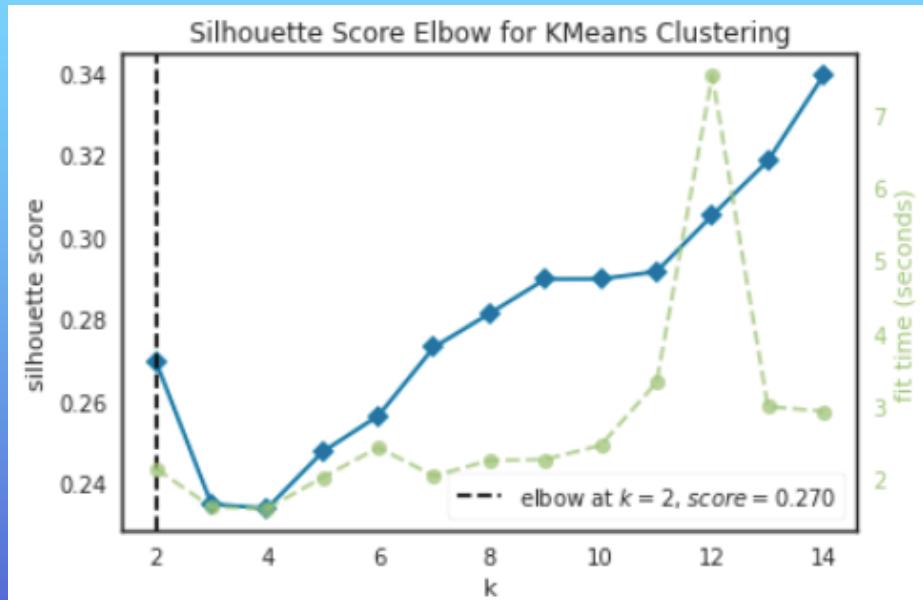
# CLUSTERING HIÉRARCHIQUE



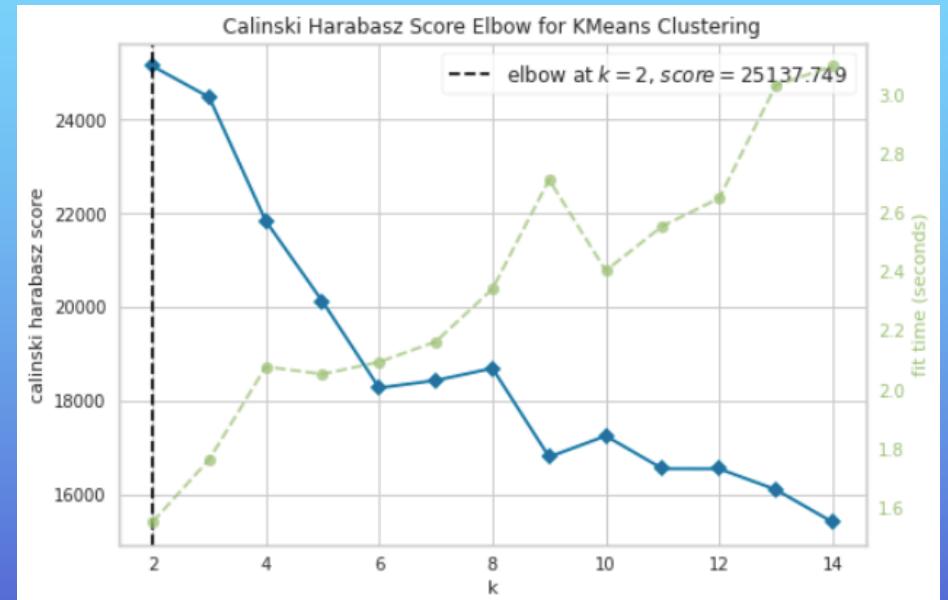
AGGLOMERATIVE CLUSTERING



# KMEANS

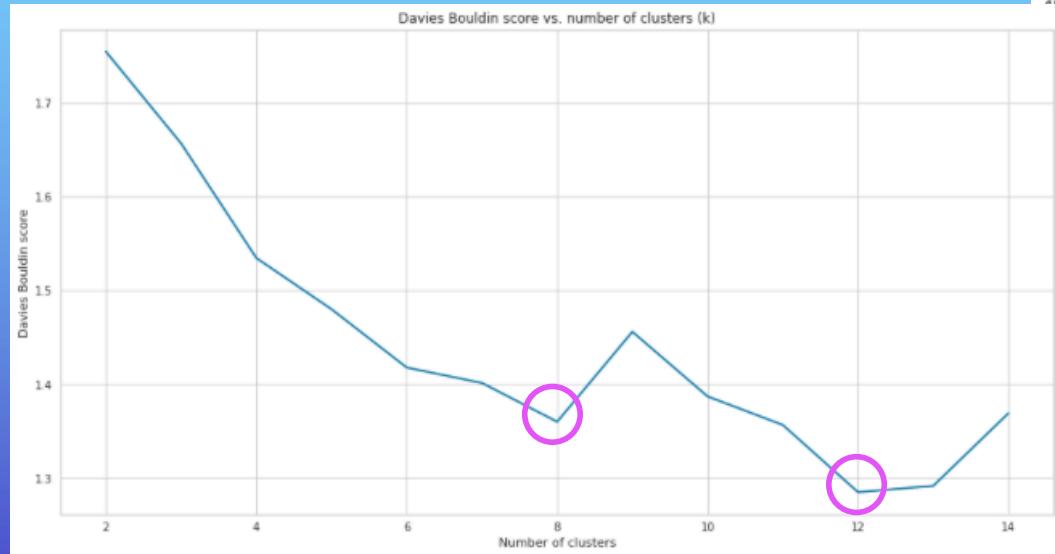


Silhouette

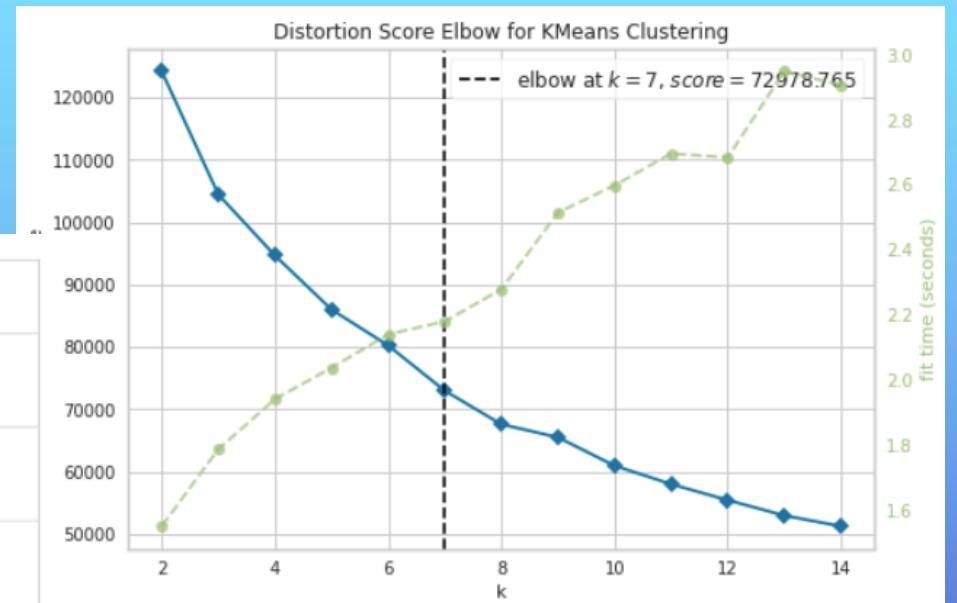


CALINSKI HARABASZ

# KMEANS

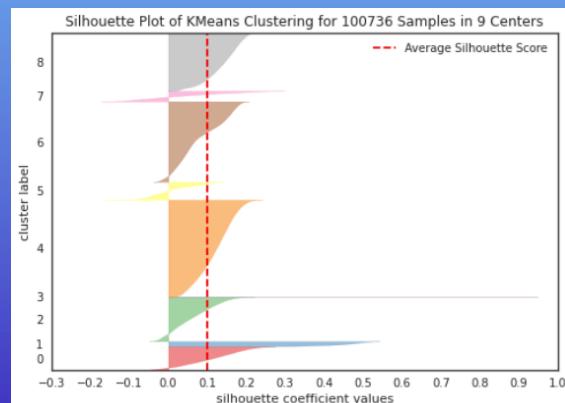
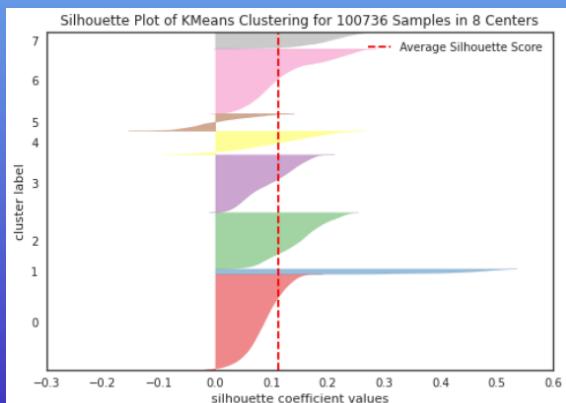
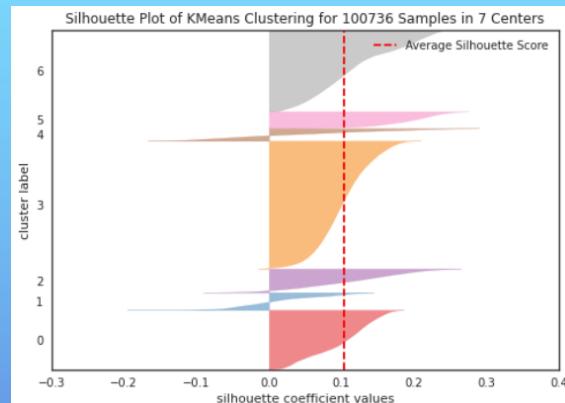
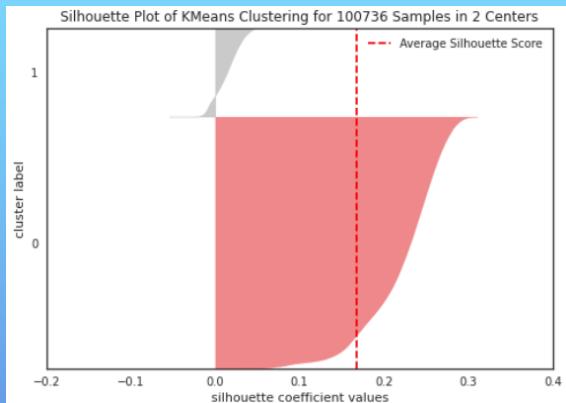


Davies Bouldin

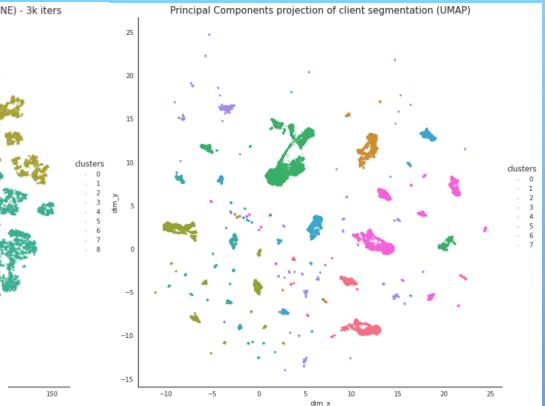
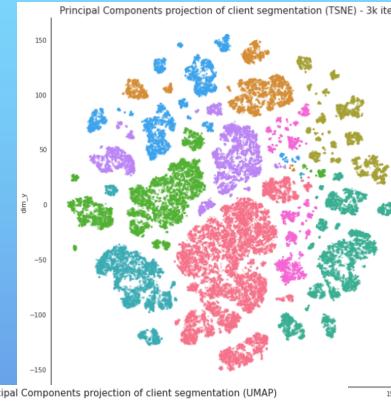
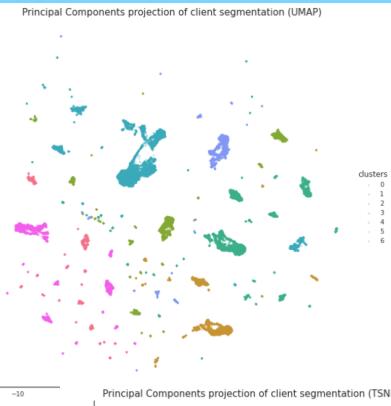
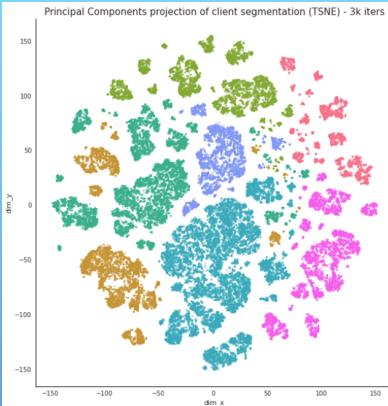


Elbow Method

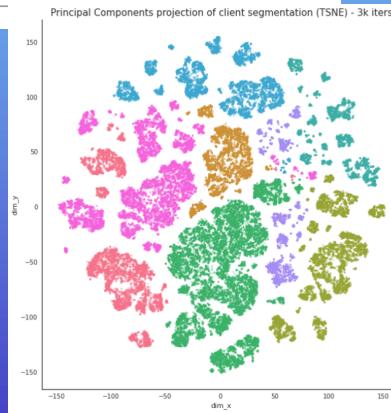
# KMEANS – SILHOUETTE VISUALIZER



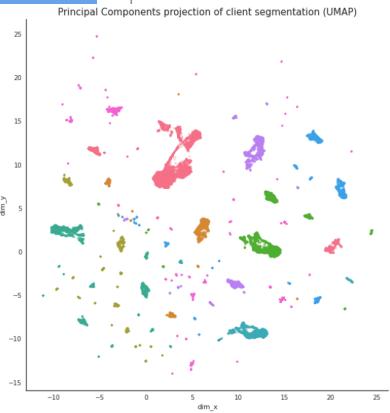
# VISUALISATIONS



$k = 7$

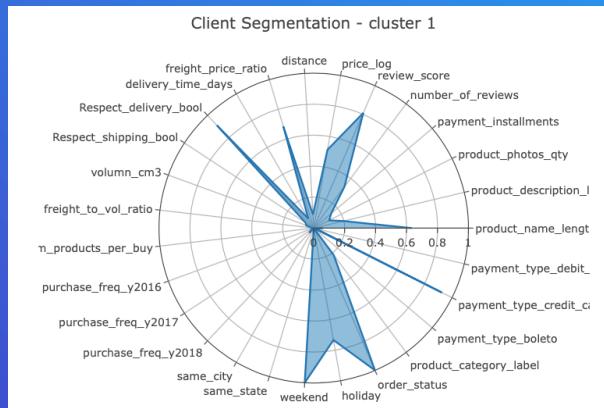


$k = 9$

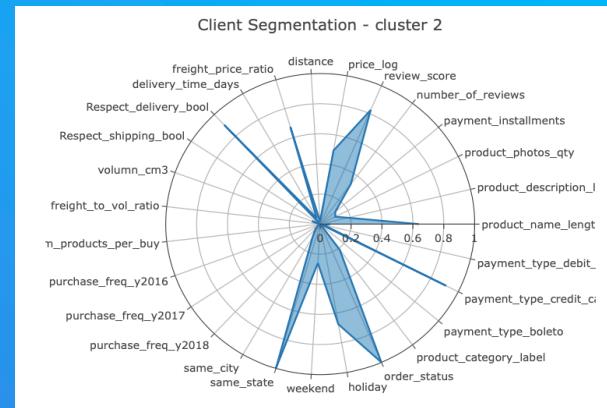


$k = 8$

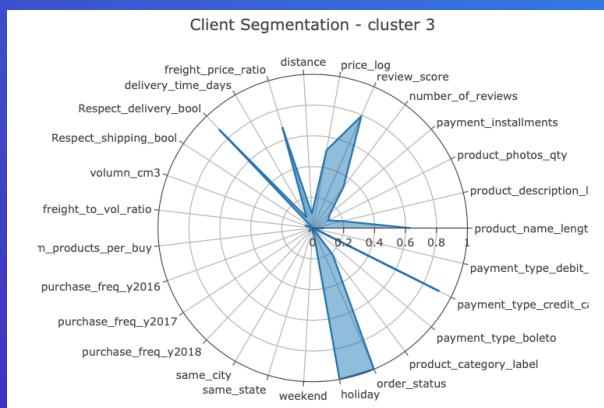
# Profils (Clusters = 7)



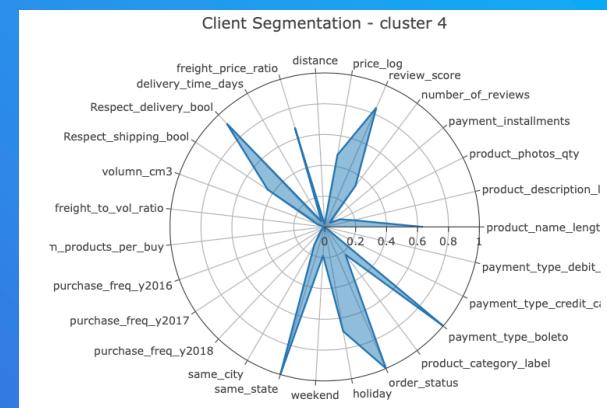
Weekend and  
occassional holiday  
user



Regional and  
occasional holiday  
user (credit card)

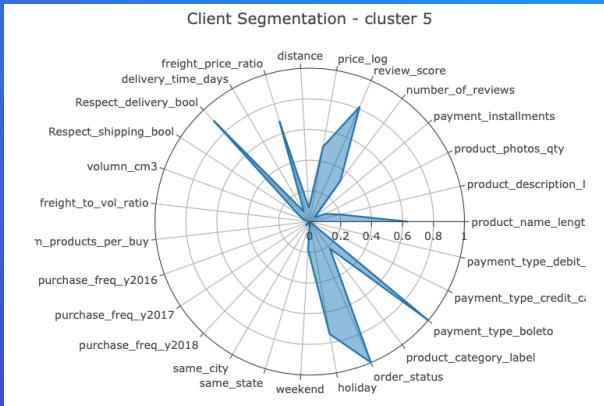


Holiday user  
(credit card)

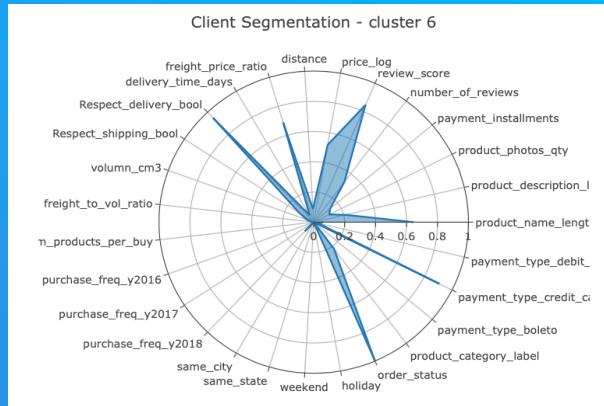


Regional and  
occasional holiday  
user (boleto)

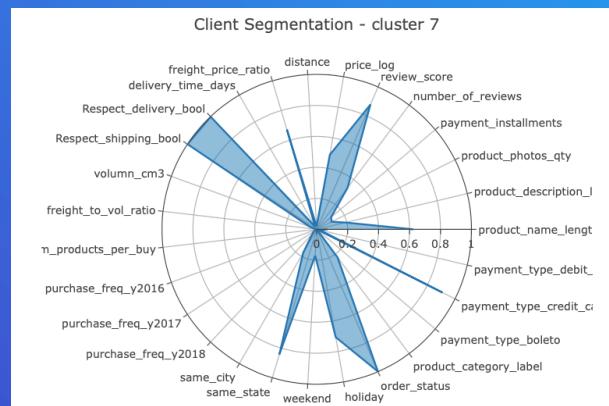
# Profils (Clusters = 7)



National and  
occassional holiday  
user (boleto)



National regular  
user



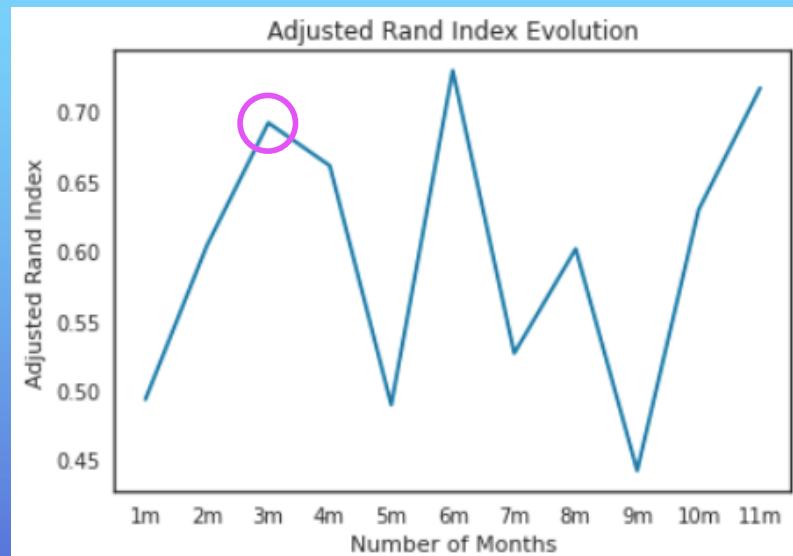
On-time user  
(occassional  
regional, occasional  
holiday and credit  
card)



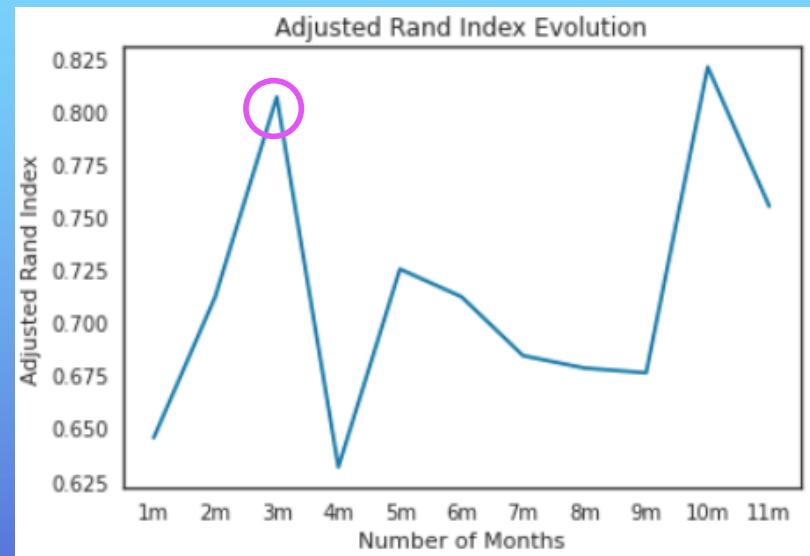
## 03. STABILITÉ

Analyse de la stabilité des segments

# Teste de la Stabilité



Tous les Clients



Clients Recurrents

# THANKS!

Do you have any questions?  
[channing.platevoet@gmail.com](mailto:channing.platevoet@gmail.com)  
+33 6 99 06 97 77

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

