

# Projet de Segmentation Client

Architecture, Optimisation et Déploiement d'une Application d'Analyse RFM



# Synthèse du Projet

## Architecture Robuste et Modulaire

Un code source organisé, facile à maintenir et à étendre, séparant la logique métier de la présentation.

---

## Performance Optimisée

Un temps de démarrage réduit de **~60s à ~2s** grâce à une stratégie de pré-calculation efficace, garantissant une excellente expérience utilisateur.

---

## Reproductibilité Assurée

Des processus d'installation et de déploiement clairs et scriptés, permettant à quiconque de lancer ou de contribuer au projet facilement.

# Le Défi du Retail : Des Données Brutes aux Insights Actionnables

Les entreprises de retail collectent un volume massif de données transactionnelles. Cependant, sans les bons outils, ces données restent une ressource inexploitée, incapable de répondre aux questions stratégiques fondamentales.



Qui sont nos clients les plus précieux et comment les fidéliser ?  
(Nos 'Champions')

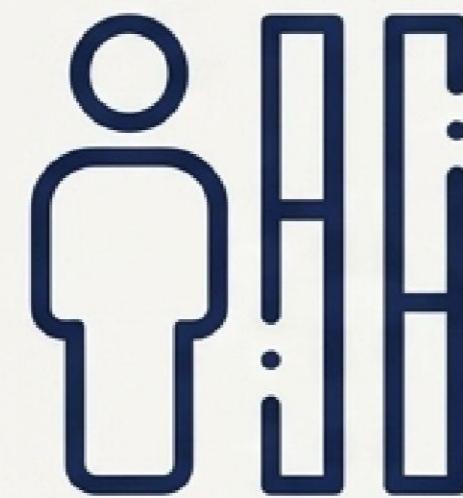


Quels sont les clients importants qui risquent de nous quitter ? (Nos clients 'À Ne Pas Perdre')



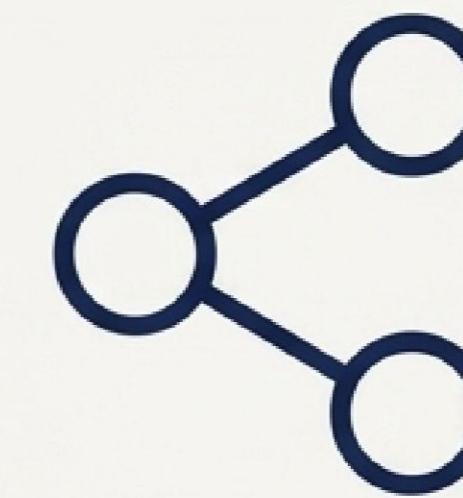
Comment augmenter la valeur de chaque transaction par des offres personnalisées ?  
(Le cross-sell)

# La Puissance de Deux Méthodologies Éprouvées



## Analyse RFM

Pour quantifier la valeur et l'engagement de chaque client en se basant sur trois dimensions clés : la **Récence**, la **Fréquence** et le **Montant** de leurs achats.



## Algorithme Apriori

Pour découvrir les affinités cachées entre les produits dans les paniers d'achat, permettant des recommandations de cross-sell intelligentes et pertinentes.

# Pilier 1 : La Segmentation Client via l'Analyse RFM



## Récence (R)

Mesure le nombre de jours écoulés depuis le dernier achat d'un client. Un score élevé indique une activité récente.



## Fréquence (F)

Compte le nombre total de commandes passées par un client. Un score élevé signifie des achats réguliers.

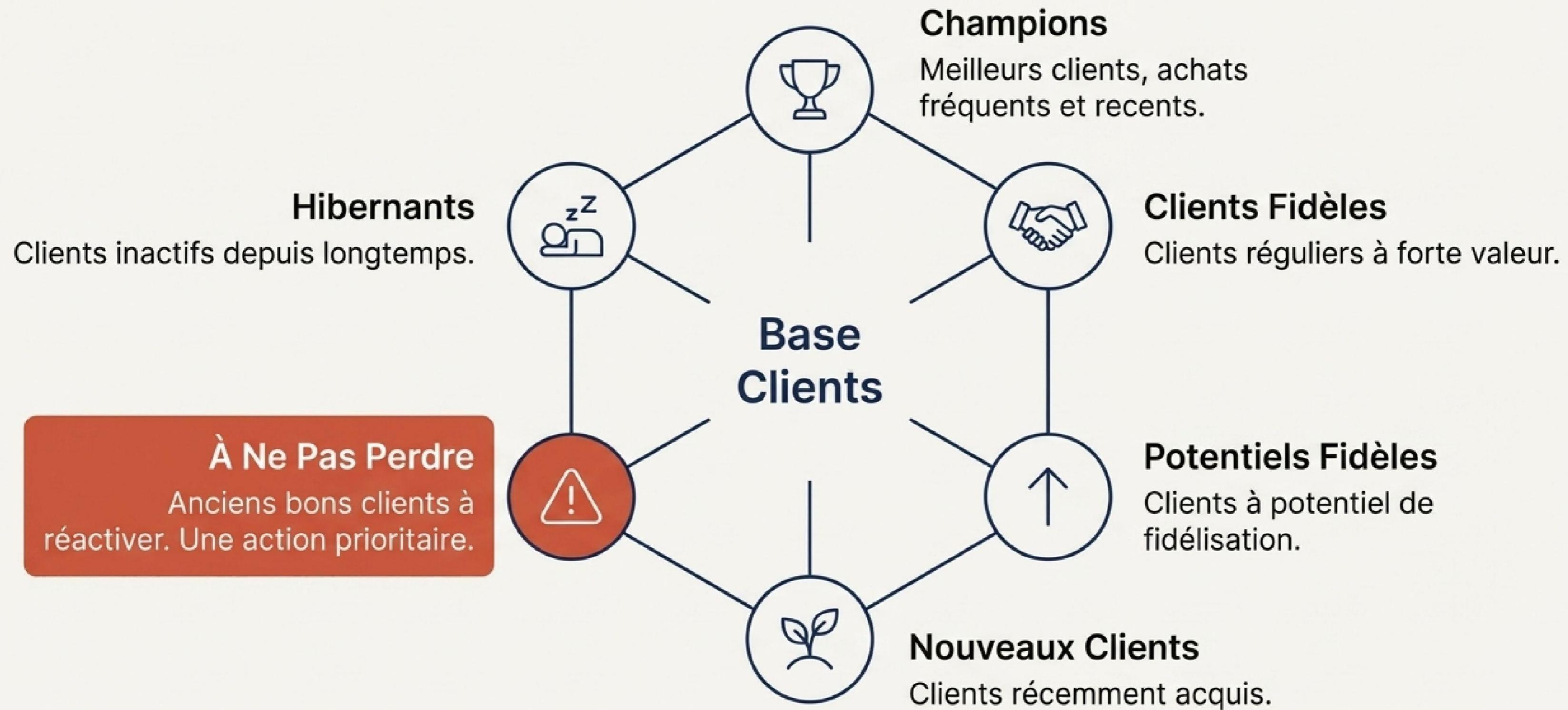


## Montant (M)

Calcule la valeur totale des achats d'un client. Un score élevé représente une forte contribution au chiffre d'affaires.

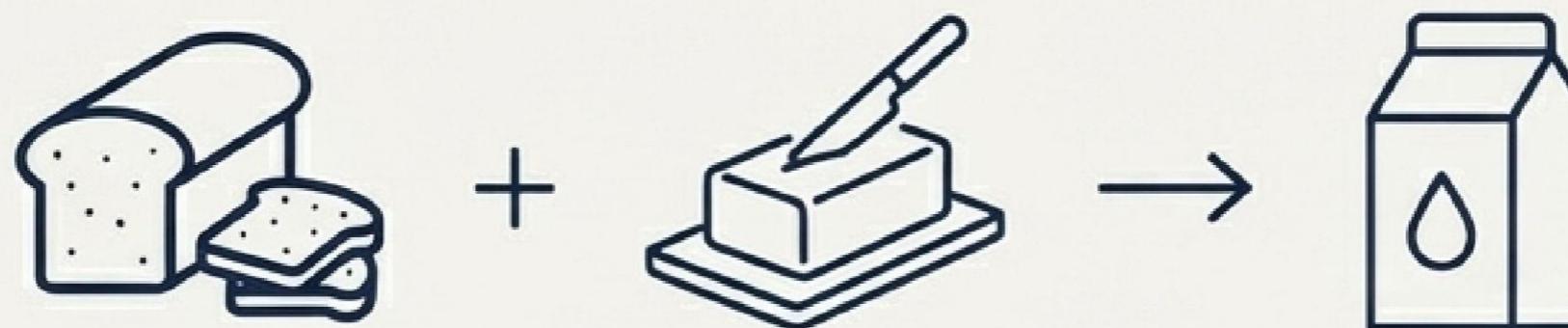


# Le Résultat : 6 Segments Clients Stratégiques



# Pilier 2 : La Recommandation via l'Analyse de Panier (Apriori)

L'analyse de panier identifie les produits fréquemment achetés ensemble.  
C'est le moteur de nos recommandations de cross-sell.



## Support

Popularité d'un ensemble d'articles.  
Indique la fréquence d'apparition  
dans toutes les transactions.

## Confiance

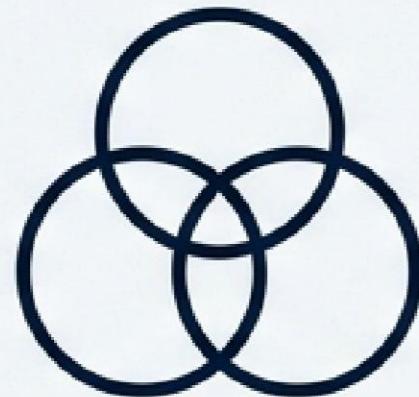
Probabilité d'acheter le produit Y si  
le produit X est déjà dans le panier.  
Mesure la fiabilité de la règle.

## Lift (Indice de pertinence)

Notre indicateur clé. Un Lift > 1  
signifie que l'association entre les  
produits est plus forte que le simple  
hasard. C'est la garantie d'une  
recommandation pertinente et non  
d'une simple coïncidence.

# Une application Streamlit pour transformer les données brutes en insights clients

L'application offre une interface interactive pour explorer les segments de clientèle et les habitudes d'achat, conçue pour être à la fois performante et facile à déployer.



## Segmentation RFM

Analyse les clients selon leur **Récence, Fréquence et Montant** d'achat pour identifier les segments les plus porteurs.



## Analyse du Panier d'Achat

Découvre les associations de produits fréquemment achetés ensemble pour optimiser les stratégies de vente croisée.



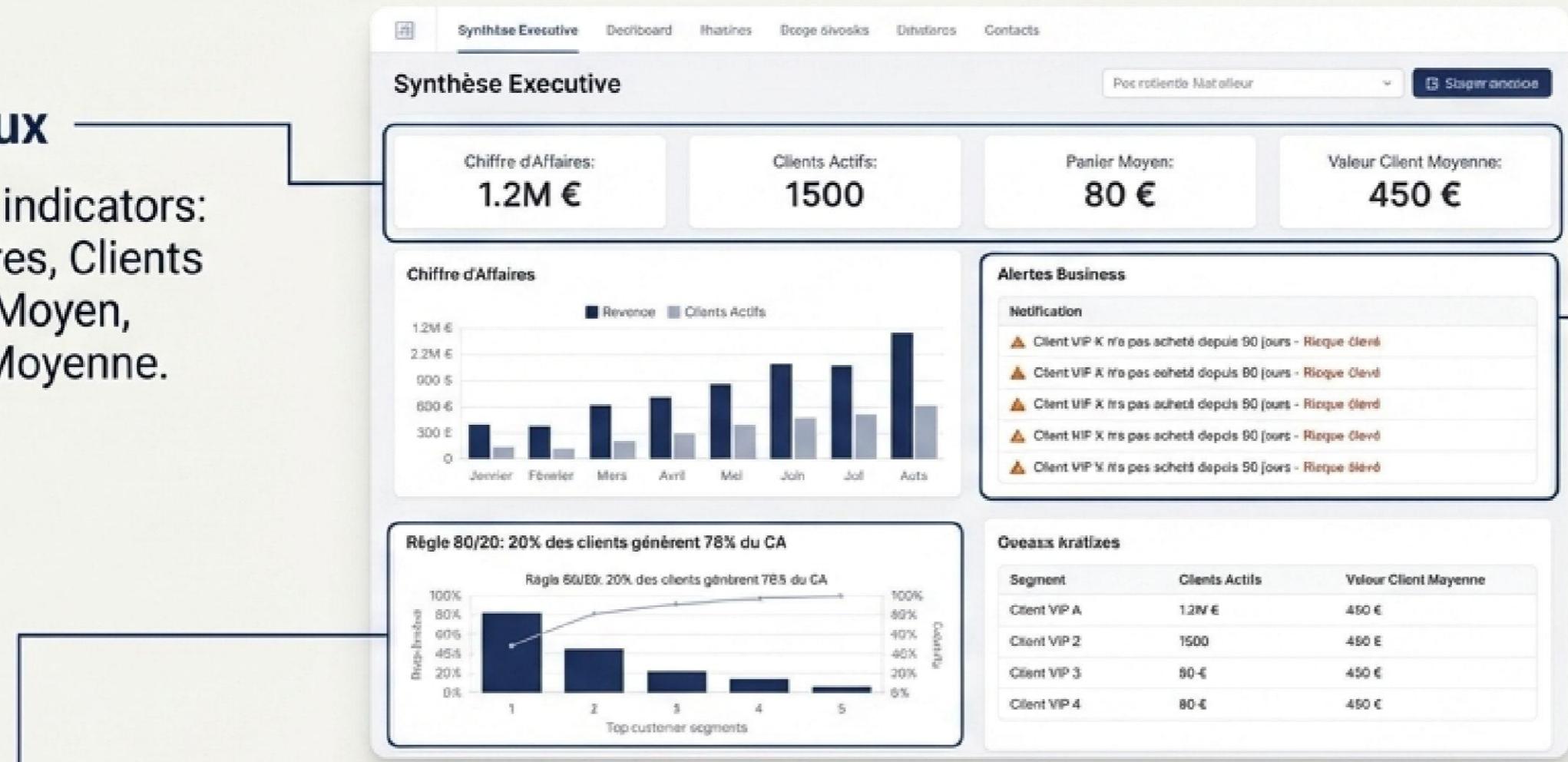
## Déploiement Optimisé

Une architecture pensée pour un déploiement rapide et un temps de démarrage quasi-instantané.

# Vue 1 : La Synthèse Exécutive – Le Pouls de l'Activité

## KPIs Globaux

Highlights the indicators:  
Chiffre d'Affaires, Clients  
Actifs, Panier Moyen,  
Valeur Client Moyenne.



## Indicateurs Stratégiques

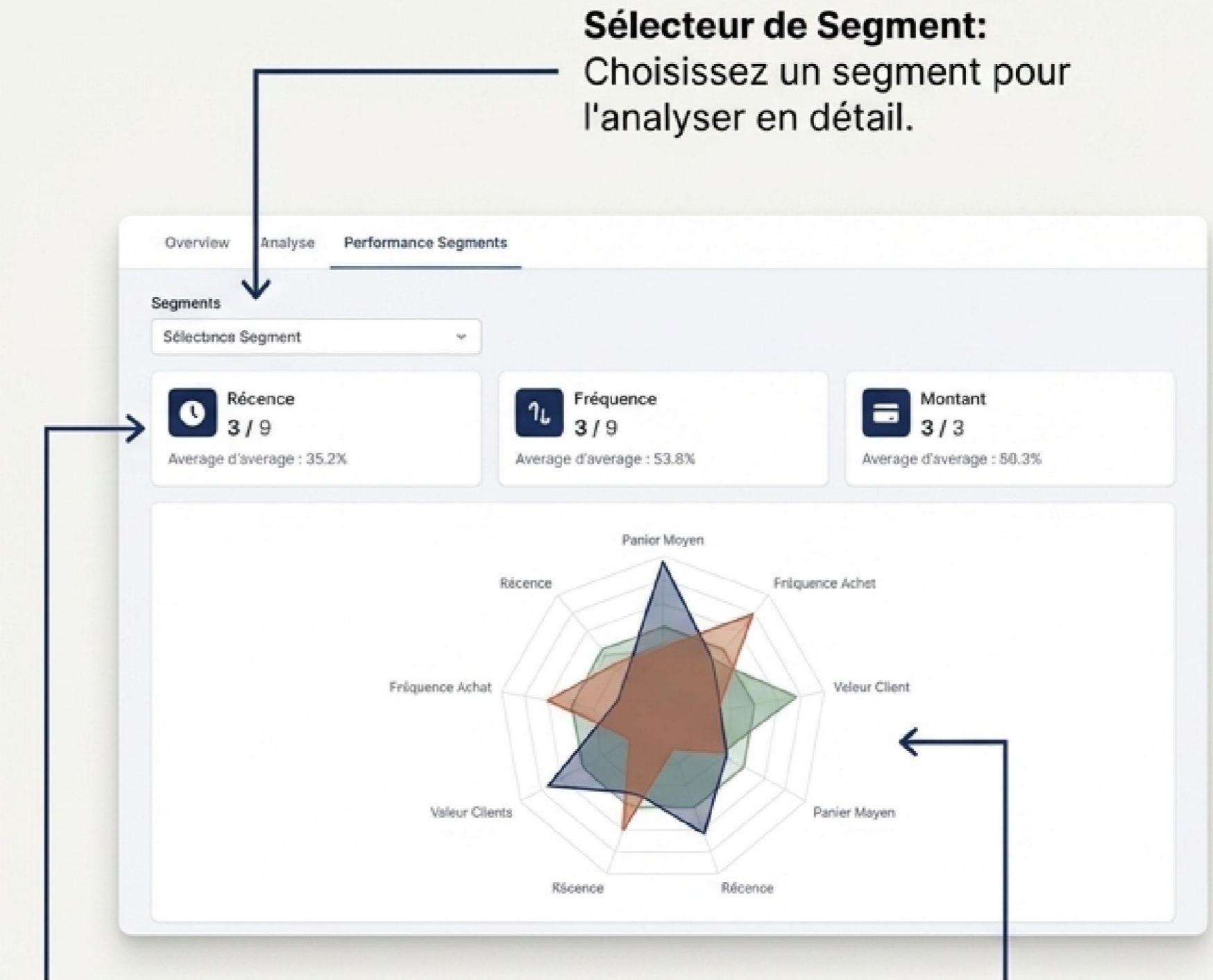
**Règle 80/20 :** Quel pourcentage de clients génère 80% du CA ?

## Alertes Business

**\*\*Clients haute valeur à risque\*\*:** Identifie les clients importants qui n'ont pas acheté récemment, nécessitant une action immédiate.

# Vue 2 : La Performance des Segments – L'ADN de Chaque Groupe

Cette vue permet une analyse approfondie de chaque segment. En sélectionnant un groupe, les utilisateurs peuvent visualiser son profil RFM moyen, métriques de performance et ses caractéristiques uniques.



**Profil RFM Moyen:**  
Visualisez l'ADN du segment : Récence, Fréquence, Montant.

**Graphique Radar:**  
Comparez instantanément le comportement du segment par rapport à la moyenne.

# Vue 3 : La Matrice d'Actions Prioritaires – De l'Analyse à la Décision

Le tableau de bord ne se contente pas de présenter des données ; il prescrit des actions. La matrice d'actions fournit une feuille de route stratégique pour chaque segment.

Tableau de bord   Données par segments   **Actions Prioritaires**   Critiques similaires

**Actions Prioritaires**  
Chaque segments

Segment	Priorité	Action	Tactique
👑 Champions	Haute	❤️ Fidéliser	Offre exclusive
⚠️ A Ne Pas Perdre	Critique	⌚ Réactiver	Campagne de réengagement
⚠️ A Ne Pas Perdre	Critique	⌚ Réactiver	Campagne de réengagement
⚠️ A Ne Pas Perdre	Critique	⌚ Réactiver	Campagne de maxtiver
⚠️ A Ne Pas Perdre	Critique	⌚ Actions les réactiver	Campagne de maxtiver
⚠️ A Ne Pas Perdre	Critique	❤️ Fidéliser	Offre exclusive
⚠️ A Ne Pas Perdre	Critique	⌚ Réactiver	Campagne de réengagement
👑 Champions	Critique	⌚ Actions les réactiver	Offre exclusive

**Clients Haute Valeur à Risque**

Client	Risk score	Recommandé actions
Clients firmas	38	📋 Liste nominative des clients haute valeur à risque
Khal Werten	36	📋 Offre vxaloise meumare pour ms mmainten
Coant Slimeti	33	⌚ Réactive campagne de réengagement
Jenathen Domard	33	⌚ Réactive campagne de réengagement
Client Lammers	26	📋 Désactiver connie pour une intervention immédiate

Segment	Priorité	Action	Tactique
👑 Champions	Haute	❤️ Fidéliser	Offre exclusive
⚠️ A Ne Pas Perdre	Critique	⌚ Réactiver	Campagne de réengagement

Une liste nominative des **clients haute valeur à risque** est également fournie pour une intervention immédiate.

# Vue 4 : Le Client 360 – La Personnalisation à l'Échelle Individuelle

The screenshot shows the 'Client 360' dashboard interface. At the top left is a user icon. The main header reads 'Client 360'. Below it, a profile box for 'Jean Dupont Client Fidèle' shows a photo of a man, the title 'Customer profile', and an 'RFM score' section with values: Récence 9/10, Fréquence 8/10, and Valeur Totale 9/10. A sidebar on the left includes icons for home, search, filters, and more. The central area has a section titled 'Historique d'Achats' (Purchase History) with a table of recent purchases:

Date	Ordre n°	Mémoirs	Total
01/11/2023	00200274300	Mennmo A	290,00 €
01/11/2023	00200298293	Clien. Fidèle	70,00 €
16/11/2023	00200770100	Rienmmo A	131,00 €

At the bottom is a section titled 'Recommandations Produit Personnalisées' (Personalized Product Recommendations) enclosed in an orange box. It lists six product suggestions with their respective lift and confidence scores:

- Produit A - Lift: 1.5, Confiance: 85%
- Produit A - Lift: 1.5, Confiance: 85%
- Produit A - Lift: 1.5, Confiance: 85%
- Produit B - Lift: 2.1, Confiance: 80%
- Produit B - Lift: 2.1, Confiance: 80%
- Produit B - Lift: 2.1, Confiance: 80%
- Produit D - Lift: 2.1, Confiance: 90%
- Produit F - Lift: 2.1, Confiance: 90%
- Produit D - Lift: 1.8, Confiance: 80%

## Profil Client

Highlights the customer's segment (e.g., "Client Fidèle"), their score RFM (Récence, Fréquence, Valeur Totale).

## Historique d'Achats

Un aperçu complet de l'historique d'interaction.

## Recommandations Produit Personnalisées

Suggestions basées sur l'analyse de panier (Apriori), avec des scores de **Lift** et de **Confiance** pour garantir la pertinence.

# Synthèse de la Valeur Business



## Identification Précise de la Valeur Client

Le scoring RFM permet de quantifier et de hiérarchiser la valeur de chaque client pour l'entreprise.



## Détection des Opportunités et des Risques

Mettez en lumière les clients à fort potentiel de croissance et identifiez proactivement les clients de valeur sur le point de partir.



## Augmentation du Panier Moyen

Générez des revenus additionnels grâce à des recommandations de cross-sell intelligentes et personnalisées.



## Optimisation des Actions Marketing

Cessez les campagnes génériques. Allouez vos ressources efficacement en ciblant chaque segment avec la bonne action.

# Une structure de projet claire et conventionnelle.

```
customer_segmentation_project/
├── app.py          # Application Streamlit
├── requirements.txt # Dépendances Python
├── render.yaml     # Config. de déploiement
└── config/
    └── config.yaml  # Paramètres de configuration
├── data/
│   └── processed/  # Données pré-calculées
└── scripts/
    └── precompute.py # Script de pré-calculation
└── src/
    ├── data_preprocessing.py
    ├── rfm_analysis.py
    ├── basket_analysis.py
    ├── recommendations.py
    ├── metrics.py
    └── visualization.py
```

# Architecture et Stack Technique

## Structure du Projet

```
customer_segmentation_project/
├── app.py
├── scripts/
│   └── precompute.py
└── src/
    ├── rfm_analysis.py
    └── basket_analysis.py
└── data/processed/
└── requirements.txt
```

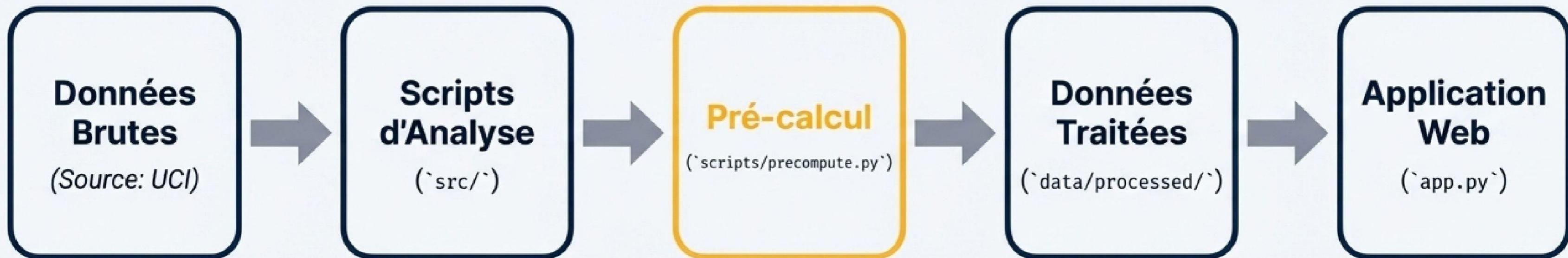
Le pré-calcul des données via `precompute.py` réduit le temps de démarrage de l'application de ~60s à ~2s.

## Stack Technique

Catégorie	Technologie
Frontend	Streamlit
Traitement de Données	Pandas
ML / Analytics	Scikit-learn, MLxtend (Apriori)
Visualisation	Plotly
Déploiement	Render

# Une architecture modulaire pour la clarté et la maintenance.

La structure du projet sépare clairement la gestion des données, la logique d'analyse et la couche de présentation. Cette séparation garantit la maintenabilité et la scalabilité du code.



# La fondation du projet : des données transactionnelles réelles.

Le projet s'appuie sur le jeu de données 'Online Retail' de l'UCI Machine Learning Repository, qui contient des transactions réelles d'un e-commerce.

## UCI Online Retail Dataset

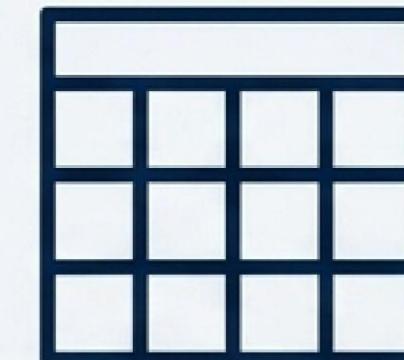
Un standard reconnu pour les projets d'analyse de comportement client et de segmentation.

``src/data_preprocessing.py``

Ce script est responsable du nettoyage des données, de la gestion des valeurs manquantes et de la création des features nécessaires pour les analyses RFM et de panier.



``data_preprocessing.py``



# Au cœur du projet : les moteurs d'analyse.

Deux scripts principaux encapsulent la logique métier pour segmenter les clients et analyser les paniers d'achat.

## Analyse RFM

`src/rfm\_analysis.py`

Calcule les scores de Récence, Fréquence et Monétaire pour chaque client, puis les assigne à des segments prédefinis (ex: 'Champions', 'Clients à risque').



Récence



Fréquence



Monétaire

## Analyse du Panier d'Achat

`src/basket\_analysis.py`

Utilise des algorithmes d'analyse d'association pour identifier les produits qui sont fréquemment achetés ensemble.



# L'optimisation clé : le pré-calcul pour un démarrage quasi-instantané.

## Problème

Le traitement complet des données au démarrage de l'application prend environ 60 secondes, ce qui nuit à l'expérience utilisateur.



**~60s**

Sans Pré-calcul

## Solution

Le script `scripts/precompute.py` exécute toutes les analyses lourdes en amont et sauvegarde les résultats (dataframes, segments, etc.) dans le répertoire `/data/processed/`.

## Impact

L'application `app.py` charge directement ces fichiers pré-calculés, réduisant le temps de démarrage à environ 2 secondes.



**~2s**

Avec Pré-calcul

# Installation locale en 4 étapes simples.

Suivez ces commandes pour cloner le projet et configurer l'environnement virtuel.

## 1 Cloner le dépôt

```
git clone https://github.com/Souley225/customer_segmentation_project.git  
cd customer_segmentation_project
```

## 3 Installer les dépendances

```
pip install -r requirements.txt
```

## 2 Créer et activer l'environnement virtuel

```
python -m venv .venv &&  
.venv\Scripts\activate
```

## 4 Lancer l'application

```
streamlit run app.py
```

# Un workflow de déploiement simple et efficace.

Le déploiement est conçu pour être direct. Après toute modification des données ou de la logique d'analyse, il suffit de régénérer les fichiers pré-calculés et de les pousser sur le dépôt.



# Explorez la Solution en Action

L'application est déployée et accessible en ligne.

L'application est hébergée sur le plan gratuit de Render. Le premier chargement peut prendre 30 à 60 secondes si le serveur est en veille.

Découvrez l'application en direct :

**customer-segmentation-project-591h.onrender.com**



Souley225/Customer\_segmentation\_project



/in/souleymane-diabate



# Code Source & Licence

## Dépôt GitHub

L'intégralité du code source, la documentation et les instructions sont disponibles sur GitHub.



[github.com/Souley225/customer\\_segmentation\\_project](https://github.com/Souley225/customer_segmentation_project)

## Licence

Ce projet est distribué sous la licence MIT, autorisant une large réutilisation à des fins commerciales et non commerciales.



Langage principal: Python (100.0%)

Stars: 1