

Projet 5 : Segmentez les clients d'un site de e-commerce

Lancelot LECLERCQ

21 janvier 2022

Sommaire

1. Introduction

2. Analyse et transformation des données

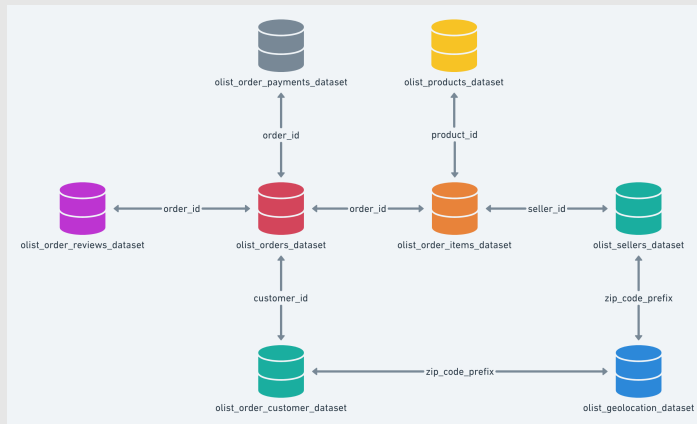
Introduction

Problématique

- Client : Olist, site de e-commerce
 - Souhaite effectuer une segmentation des clients
 - Comprendre les différents types d'utilisateurs
- Objectifs :
 - Fournir une description actionable de la segmentation et de sa logique pour une utilisation optimale
 - Faire une proposition de contrat de maintenance à partir de l'analyse de la stabilité de la classification au cours du temps

Jeu de données

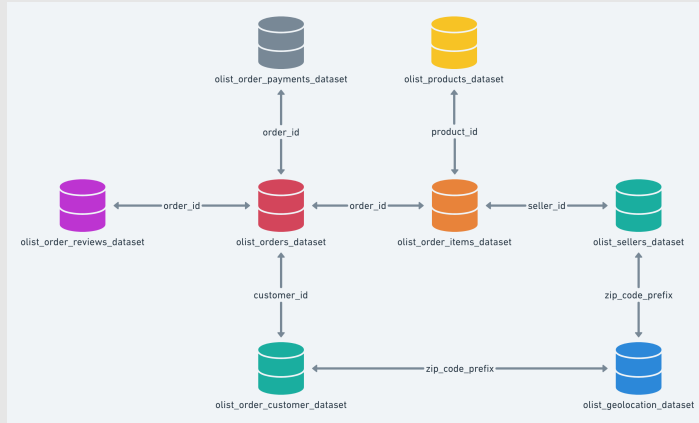
- Base de données anonymisée du site
- Données concernant :
 - les clients,
 - les vendeurs,
 - les commandes,
 - les produits vendus,
 - les commentaires et la satisfaction
- Données des années 2017 et 2018



Analyse et transformation des données

Analyse et transformation des données : structure de la base de données

- Fichier central contenant les commandes (olist_orders_dataset)
- Des identifiants permettent de rassembler les différents fichiers constituant la base de données
 - order_id relie les données de paiement, les produits commandés et les données de notation aux commandes,
 - customer_id relie les identifiants des clients aux commandes,
 - products_id relie les données concernant les produits aux produits commandés,
 - seller_id relie les données de vendeurs aux produits commandés,
 - zip_code_prefix relie les données de géolocalisation des acheteurs et des vendeurs
- Ces données vont nous permettre de calculer des variables plus intéressantes dans le cadre de cette segmentation des clients



Analyse et transformation des données : structure de la base de données

- Études des clients = nécessité de regrouper les données de commandes par client
- Permet de compter :
 - la date de la dernière commande
 - le prix moyen d'une commande d'un client
 - le nombre de commandes par client
 - le temps moyen entre deux commandes effectuées par un même client
 - la note moyenne données par un client
 - le nombre de produits par catégories de produits achetés par un clients
 - ⋮
- Utilisation des méthodes traditionnelles de marketing pour classifier les clients : classification RFM
 - Recency : la date à laquelle a été effectuée la dernière commande
 - Frequency : le nombre de commande
 - Monetary : le prix moyen d'une commandes d'un client