

[←](#) Mon parcours PROJET VALIDÉ

Segmentez des clients d'un site e-commerce

MISSION

COURS

RESSOURCES

ÉVALUATION



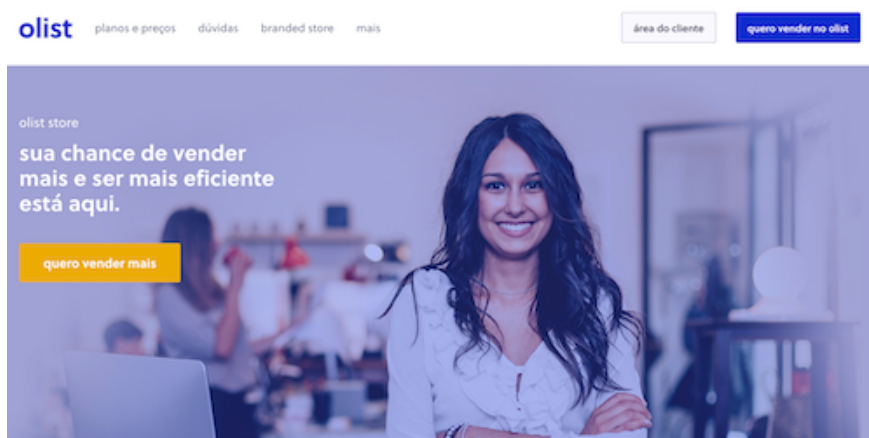
70 heures

Mis à jour le mercredi 1 décembre 2021



Si vous avez commencé ce projet avant le 01/12/2021, vous avez débuté votre travail sur **ce projet archivé**. Vous pouvez continuer l'existant ou recommencer avec ce nouveau projet, dont les modifications s'alignent avec les retours étudiants.

Vous êtes consultant pour [Olist](#), une entreprise brésilienne qui propose une solution de vente sur les marketplaces en ligne.



La page d'accueil du site Olist

Olist souhaite que vous fournissiez à ses équipes d'e-commerce une **segmentation des clients** qu'elles pourront utiliser au quotidien pour leurs campagnes de communication.

Votre objectif est de **comprendre les différents types d'utilisateurs** grâce à leur comportement et à leurs données personnelles.

Vous devrez **fournir à l'équipe marketing une description actionable** de votre segmentation et de sa logique sous-jacente pour une utilisation optimale, ainsi qu'une **proposition de contrat de maintenance** basée sur une analyse de la stabilité des segments au cours du temps.

Les données

Pour cette mission, Olist vous fournit une [base de données](#) anonymisée comportant des informations sur l'historique de commandes, les produits achetés, les commentaires de satisfaction, et la localisation des

clients depuis janvier 2017.

Votre mission

Votre mission est d'aider les équipes d'Olist à comprendre les différents types d'utilisateurs. Vous utiliserez donc des méthodes non supervisées pour regrouper des clients de profils similaires. Ces catégories pourront être utilisées par l'équipe Marketing pour mieux communiquer.

Vous créez donc un notebook et démarrez votre travail d'analyse exploratoire.



Après quelques premières analyses, vous vous rendez compte qu'Olist ne vous a pas fourni beaucoup de données ; vous enquêtez donc auprès de l'entreprise pour obtenir quelques informations complémentaires, et vérifier que vous avez bien compris la mission. Voici sa réponse.

Bonjour,

Pour des raisons de confidentialité, nous ne pouvons pas vous fournir beaucoup de données à ce stade. Ensuite, en raison de ressources limitées, nous avons dû vous fournir l'ensemble des données, alors que seule une partie va vous intéresser. Nos dashboards internes nous indiquent en effet que seuls 3 % des clients du fichier de données partagé avec vous ont réalisé plusieurs commandes.

Nous sommes confiants sur le fait que les données à disposition suffiront pour réaliser un premier clustering. Cela a déjà été fait par d'autres prestataires par le passé, avec encore moins de données.

La segmentation proposée doit être exploitable et facile d'utilisation par notre équipe Marketing. Elle doit au minimum pouvoir différencier les bons et moins bons clients en termes de commandes et de satisfaction. Nous attendons bien sûr une segmentation sur l'ensemble des clients.

Dans un deuxième temps, une fois le modèle de segmentation choisi, nous souhaiterions que vous nous fassiez une recommandation de fréquence à laquelle la segmentation doit être mise à jour pour rester pertinente, afin de pouvoir effectuer un devis de contrat de maintenance.

Pour information, le code fourni doit respecter la convention PEP8, pour être utilisable par Olist.

Bien à vous,

Juan, de l'équipe Marketing

Vous réalisez donc trois autres documents afin de préparer une réponse à Olist :

- un notebook avec des essais des différentes approches de modélisation ;

- un notebook de simulation pour déterminer la fréquence nécessaire de mise à jour du modèle de segmentation, afin que celui-ci reste pertinent ;
- une présentation pour un collègue afin d'obtenir ses retours sur votre approche.

Livrables

- Un **notebook de l'analyse exploratoire** (non cleané, pour comprendre votre démarche).
- Un **notebook** (ou code commenté au choix) **d'essais** des différentes approches de modélisation (non cleané, pour comprendre votre démarche).
- Un **notebook de simulation pour déterminer la fréquence nécessaire de mise à jour** du modèle de segmentation.
- Un **support de présentation** pour présenter votre travail à un collègue.



Pour faciliter votre passage devant le jury, déposez sur la plateforme, dans un dossier nommé "*P[numéroduprojet]_nom_prenom*", tous les livrables du projet. Chaque livrable doit être nommé avec le numéro du projet et selon l'ordre dans lequel il apparaît, par exemple "*P5_01_notebookanalyse*", "*P5_02_notebookessais*", et ainsi de suite.

Soutenance

Pendant la soutenance, l'évaluateur jouera le rôle d'un collègue à qui vous présentez votre travail, avant de le présenter à une personne chez Olist. Vous lui présenterez l'ensemble de votre travail.

- **Présentation (20 minutes)**
 - Présentation de la problématique, du cleaning effectué, du feature engineering et de l'exploration – 5 min.
 - Présentation des différentes pistes de modélisation effectuées et du modèle final sélectionné – 10 min.
 - Présentation de la simulation pour définir le délai de maintenance du modèle (contrat de maintenance) – 5 min.
- **Discussion (5 minutes)**
 - L'évaluateur vous challengera sur vos choix.
- **Débriefing (5 minutes)**
 - À la fin de la soutenance, vous pourrez débriefer ensemble.

Compétences évaluées



Mettre en place le modèle d'apprentissage non supervisé adapté au problème métier



Transformer les variables pertinentes d'un modèle d'apprentissage non supervisé



Adapter les hyperparamètres d'un algorithme non supervisé afin de l'améliorer



Évaluer les performances d'un modèle d'apprentissage non supervisé

OPENCLASSROOMS

Qui sommes-nous ?


Alternance

Financements

Expérience de formation


Forum

Blog 

Presse 

OPPORTUNITÉS

Nous rejoindre 

Devenir mentor 

Devenir coach carrière 

AIDE



FAQ

POUR LES ENTREPRISES

Former et recruter

EN PLUS

Boutique 

Mentions légales

Conditions générales d'utilisation

Politique de protection des données personnelles

Cookies

Accessibilité



Français



