

# CONTEXTE

Dans le cadre d'une analyse interne menée pour OLIST, ce rapport vise à fournir une lecture synthétique et opérationnelle des performances de la plateforme. L'étude s'appuie sur des données transactionnelles (commandes, paiements, produits, vendeurs, livraisons, et avis clients) afin d'identifier les atouts et les risques du modèle économique de l'entreprise.

L'ANALYSE ABOUTIT À DES RECOMMANDATIONS DESTINÉES À SOUTENIR LES DÉCISIONS STRATÉGIQUES ET OPÉRATIONNELS.

## L'entreprise OLIST

OLIST est une entreprise brésilienne spécialisée dans l'e-commerce fondée en 2015. Son activité consiste à connecter des vendeurs à des marketplaces, tout en proposant une simplification des canaux de vente via un ensemble de services permettant de gérer plus efficacement la mise en ligne des offres, le traitement des commandes ainsi que le pilotage de l'activité.

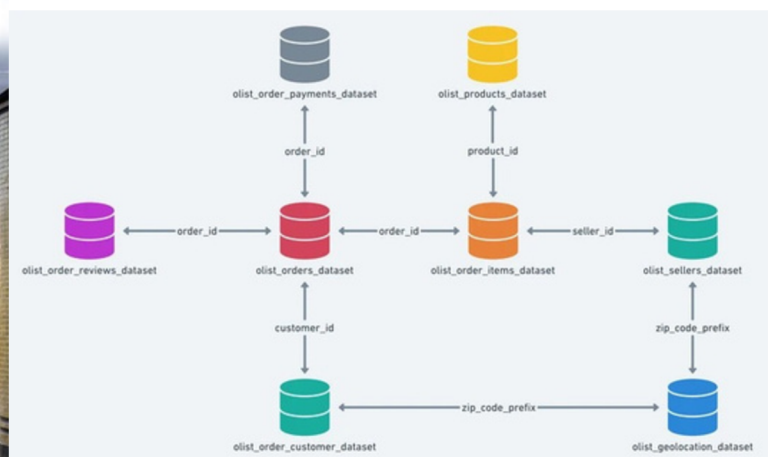
Le modèle économique repose donc sur une logique "plateforme" comparable aux grands acteurs du e-commerce comme Amazon : l'entreprise va centraliser les flux, réduire les frictions (prise en charge de la livraison à la demande du vendeur), afin d'améliorer à la fois la performance commerciale et la qualité de service délivrée aux clients.

Ces dernières années, OLIST a accéléré sa croissance via plusieurs acquisitions, et a atteint le statut de licorne après une levée de fonds annoncée le 15 décembre 2021 (181 M\$), avec une valorisation communiquée autour de 1,5 milliards de dollars

## Source des données

Les données mobilisées dans le cadre de ce rapport sont issues de la plateforme Kaggle via le Brazilian E-Commerce Public Dataset, by Olist. Ce jeu de données met à disposition un échantillon anonymisé d'activité e-commerce, couvrant environ **100 000 commandes** réalisées au **Brésil** sur la période **2016-2018**, provenant de plusieurs marketplaces.

Le jeu de données est fourni sous plusieurs fichiers CSV organisée comme une base relationnelle, permettant de reconstituer le parcours d'achat et de croiser des analyses de performance, de logistique et de satisfaction client.



# DESCRIPTION DES DONNÉES

La base relationnelle est composée de **9 tables**, avec un total de **52 champs**.

Les jointures reposent principalement sur les identifiants de commande :

- *order\_id*, la clé principale reliant les commandes aux tables de produits, de paiement et d'avis client ;
- *customer\_id*, relie la table de commande aux clients, *seller\_id* aux vendeurs ;
- *product\_id* relie les produits commandés à la table de produit ;
- *product\_category\_name*, permet de rattacher les catégories de produits.

Le tableau ci dessous présente le **dictionnaire des variables** du jeu de donnée :

<b>Identifiant (STRING/VARCHAR)</b>	order_id, customer_id, customer_unique_id, product_id, seller_id, review_id	0015a82c2db 000af6aaaf3a e2ecb0532
<b>Catégorielle (STRING/VARCHAR)</b>	order_status, payment_type, product_category_name, customer_city, customer_state, seller_city, seller_state, geolocation_city, geolocation_state	delivered / credit_card / SP
<b>Texte long (TEXT)</b>	review_comment_title, review_comment_message	"Producto excelente"
<b>DATETIME/TIME STAMP</b>	order_purchase_timestamp, order_approved_at, order_delivered_carrier_date, order_delivered_customer_date, order_estimated_delivery_date, shipping_limit_date, review_creation_date, review_answer_timestamp	2017-03-02 10:20:12
<b>Integer (INT)</b>	payment_sequential, payment_installments, review_score, product_name_lenght, product_description_lenght, product_photos_qty, product_weight_g, product_lenght_cm, product_height_cm, product_width_cm	450
<b>Zip Code (INT)</b>	customer_zip_code_prefix, seller_zip_code_prefix, geolocation_zip_code_prefix	13056
<b>Montants et coordonnées (FLOAT)</b>	price, freight_value, payment_value / geolocation_lat, geolocation_lng	79.90 / -23.550352

Figure 01. Dictionnaire des variables.



# ANALYSE EXPLORATOIRE

L'analyse exploratoire tient pour objectif de construire un **profil d'OLIST** à partir d'une sélection d'indicateurs clés. L'intérêt étant de caractériser l'ampleur de l'activité et l'étendue du marché (volume de commande, diversité de l'offre,...). Cette étape a également permis de cadrer les principaux points de vigilance liés au jeu de données, tels que la **présence de valeurs manquantes** (attributs produits, avis vide,...) et les **différences de granularité** (commande et ligne d'achat) entre tables. La recrudescence de données manquantes à partir du 31/08/2018, conduit également à limiter l'analyse aux dates antérieures.

## Profil OLIST

Clients : 99 421  
Vendeurs : 3 095  
États servis : 27

Produits (catalogue) : 32 951  
Catégories de produits : 73  
Produits vendu : 112 649  
Nb moyen de produit par commande : 1,133

Commandes : 99 421  
Commandes moyennes par mois : 4 323  
Commandes moyennes par client : 1.035  
Taux de clients récurrents : 3.119%

Moyens de paiement (types) : 5  
Paiements (enregistrements) : 103 866  
Paiements par commande : 1,045

Note moyenne : 4.09  
% Retard sur commande : 8.11%

En revanche, la consommation semble majoritairement ponctuelle : la moyenne de commande par client (1,035) et le taux de client récurrents (3,119%) suggèrent un modèle davantage porté par la **recherche de nouveau client** et le **volume** que par la fidélisation.

Sur le plan transactionnel, l'entreprise permet cinq types de paiement et le ratio de **1,405 paiement par commande** indiquent que le processus est globalement "simpliste", les paiements se faisant majoritairement en une seule fois.

Enfin, la satisfaction des client apparait globalement positive avec une **note moyenne de 4,09/5**, sachant toutefois, que **8,11%** des **commandes** enregistrent un **retard de livraison**. Soit un niveau suffisant pour impacter la perception de certains services (livraisons, vendeurs, ...) et justifier un approfondissement dans l'analyse.

Figure 02. Tableau récapitulatif des indicateurs clés

**OLIST apparait comme une plateforme e-commerce nationale avec un catalogue large, et une satisfaction globalement bonne.**

Les indicateurs mettent en évidence qu'OLIST opère comme une plateforme e-commerce d'envergure nationale, avec **99 421 clients**, **3 095 vendeurs** et une couverture de **27 États**, soit une activité étendue à l'échelle du territoire.

L'offre apparait dense et diversifiée, avec un catalogue de **31 951 produits** répartis en **73 catégories**. La dynamique des commande est soutenue avec en moyenne un rythme de **4 323 commandes par mois**.

## CADRAGE

Après la vue d'ensemble de l'activité, l'étude est approfondie afin de passer d'une vision descriptive de l'activité à une lecture orientée performance économique.

Il s'agit d'identifier les moteurs de performance, les vendeurs qui y sont à l'origine ainsi que leurs clientèle, puis d'évaluer la satisfaction client.

# LES DÉTERMINANTS DU CHIFFRE D'AFFAIRES

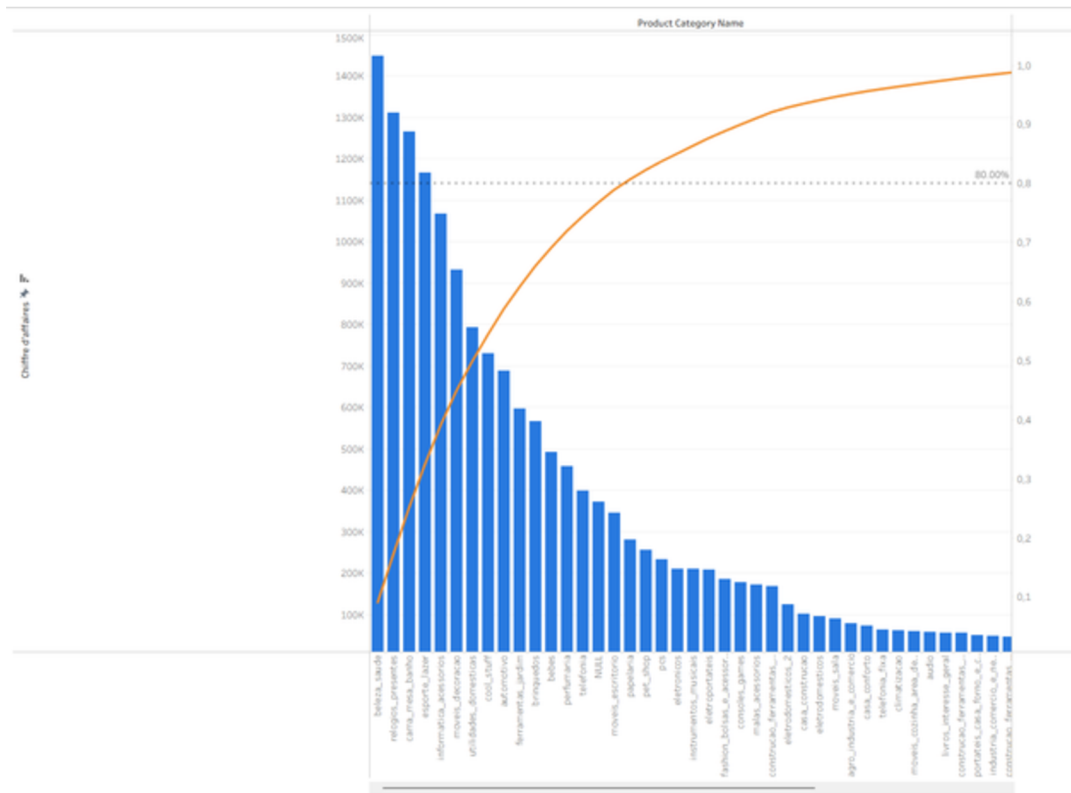


Figure 02. Analyse de Pareto du chiffre d'affaires.

## Le chiffre d'affaires est générée par quelques catégories de produits

L'analyse de type Pareto, met en évidence les catégories de produits qui produisent le plus de chiffre d'affaires (barres bleues), tout en comparant la part cumulée du chiffre d'affaires par catégorie (courbe orange).

Notons également, la présence d'une catégorie NULL correspond aux produits non catégorisés dans la base. Ce point constitue un enjeu de qualité de données: il peut biaiser l'analyse par catégorie.

Il semble qu'une poignée de catégories domine la génération de revenus, notamment :

- Beauté & santé,
- Montres & cadeaux,
- Linge de maison (lit, table, bain),
- Sport & loisirs,
- Informatique & accessoires,
- Meubles & décoration,
- et Articles ménagers.

Au niveau du seuil, 80% du chiffre d'affaires est principalement portée par 12-13 catégories de produits. Cela veut également dire que les 60 catégories restantes contribuent marginalement à la production du chiffre d'affaires total. D'un point de vue pilotage, cela implique de prioriser les actions sur ces catégories clés, et de suivre leur évolution.

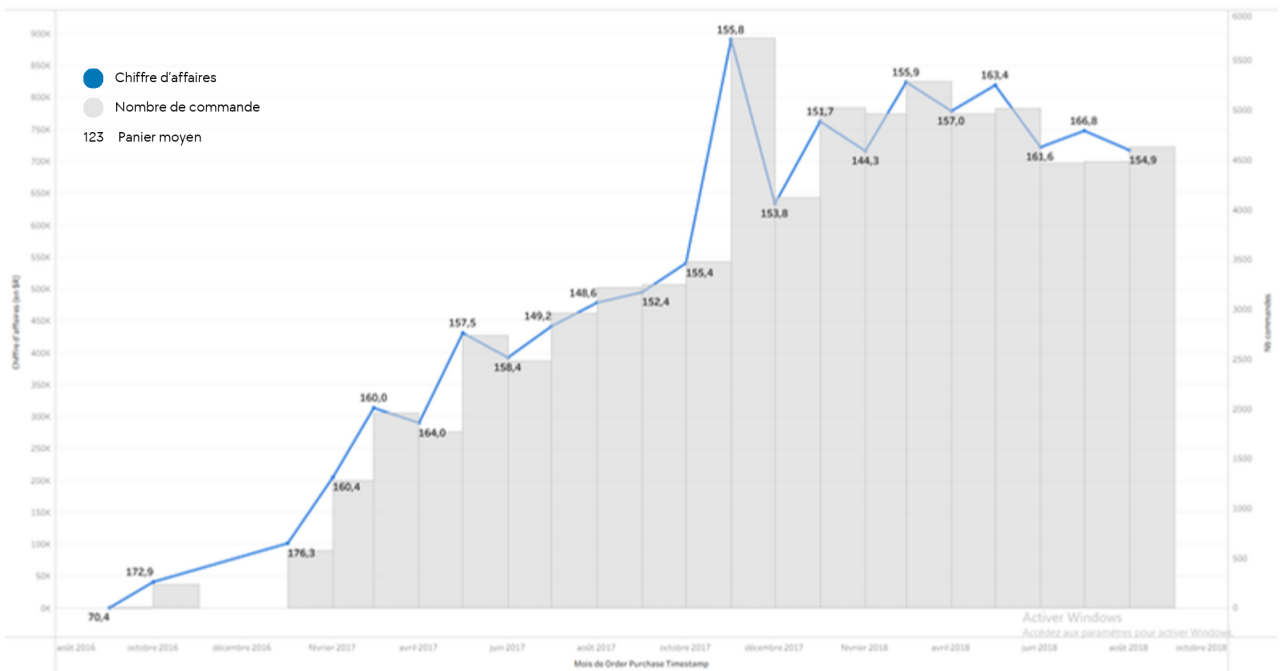


Figure 03. Evolution mensuelle du chiffre d'affaires, du panier moyen et du nombre de commandes.

### L'activité d'OLIST est principalement tirée par le volume de commandes

Le volume de commandes progresse fortement jusqu'à fin 2017, traduisant une forte dynamique de croissance de l'activité. Elle semble cependant marqué par une rupture autour de novembre 2017, où le nombre de commandes atteint un pic, avant un retour à un niveau normalisé, reprenant la trajectoire de la tendance haussière.

En parallèle, le panier moyen semble peu progresser, contraint dans un plage resserrée (entre 145 à 170). Cela semble alors indiquer que la croissance provient **surtout de l'augmentation du volume**, plutôt que d'une hausse durable de la valeur moyenne par commande.

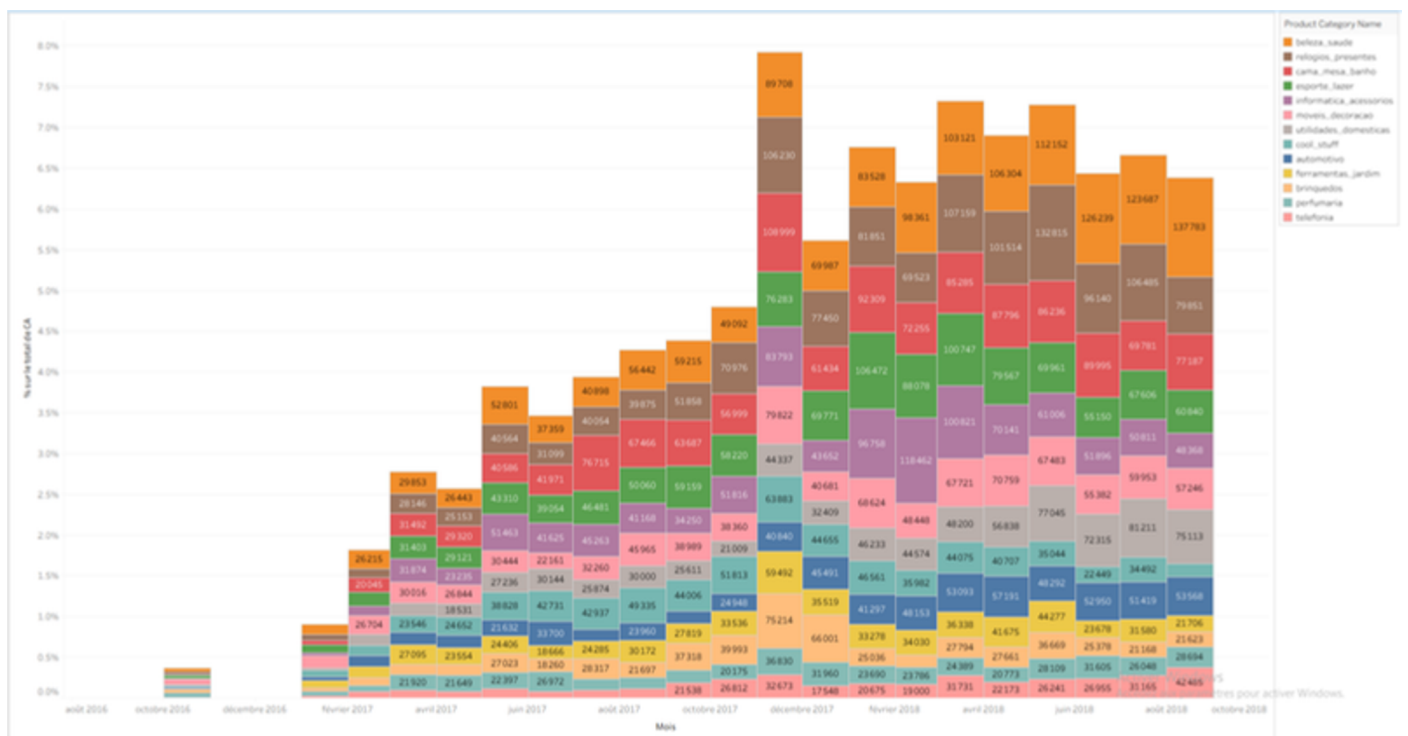


Figure 04. Evolution de la structure du chiffre d'affaires.