

L'analyse du commerce électronique brésilien

Les données

Il s'agit d'un ensemble de données publiques de commerce électronique brésilien des commandes passées sur **Olist Store**. L'ensemble de données contient des informations sur 100 000 commandes de 2016 à 2018 effectuées sur plusieurs marchés au Brésil.

Ses fonctionnalités permettent de visualiser une commande à partir de plusieurs dimensions: du statut de la commande, du prix, du paiement et des performances de transport à l'emplacement du client, aux attributs du produit et enfin aux avis rédigés par les clients.

Il y a également un ensemble de données de géolocalisation qui relie les codes postaux brésiliens aux coordonnées latitude / longitude.

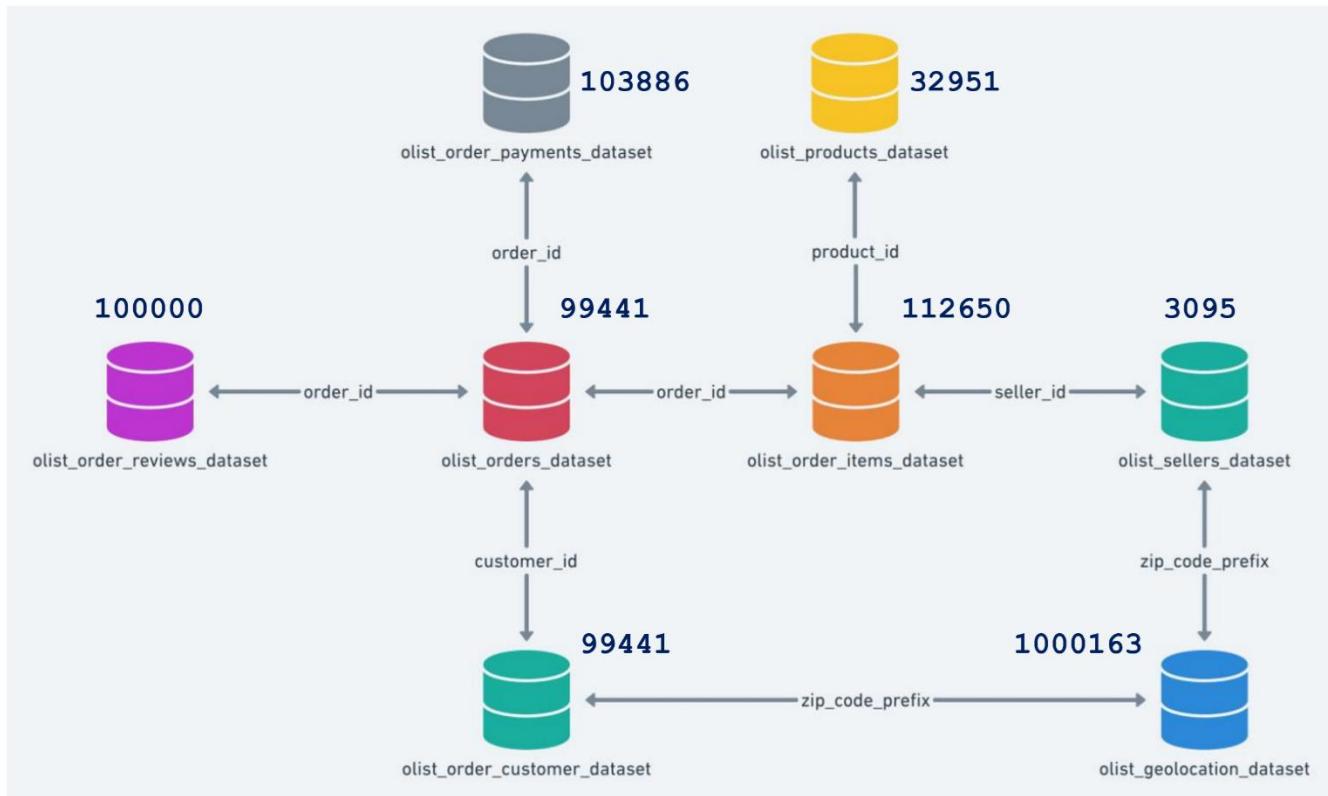
Olist connecte les petites entreprises de tout le Brésil aux canaux sans tracas et avec un seul contrat. Ces marchands peuvent vendre leurs produits via l'Olist Store et les expédier directement aux clients en utilisant les partenaires logistiques Olist.

Une fois qu'un client achète le produit sur Olist Store, un vendeur est averti de traiter cette commande. Une fois que le client reçoit le produit ou que la date de livraison estimée est due, le client reçoit une enquête de satisfaction par e-mail dans laquelle il peut donner une note pour l'expérience d'achat et écrire quelques commentaires.

The screenshot shows a product page for a Motorola Moto G6 Play smartphone. The main image displays the phone from three angles: front view, back view, and a side view. Below the main image is a thumbnail gallery showing various angles of the phone. To the right of the image, the product title is listed: "Smartphone Motorola Moto G6 Play Dual Chip Android Oreo - 8.0 Tela 5.7" Octa-Core 1.4 GHz 32GB 4G Câmera 13MP - Índigo". Below the title, there is a rating of 4 stars from 215 reviews. A promotional offer is shown: "pegue na loja hoje!" (Get it at the store today!) with a note that it's subject to price change. A "ver lojas" button is also present. The main price is listed as R\$ 1.299,00. Payment options include 10x of R\$ 129,90 with interest or R\$ 1.069,90 with interest. There is also an option for R\$ 975,00 with interest. A note indicates that there are only 5 units left in stock. A "comprar" (buy) button is prominently displayed. At the bottom, there is a note that the product is sold by a partner store and information about payment methods.

Les données sont divisées en plusieurs ensembles de données pour une meilleure compréhension et une meilleure organisation. Veuillez-vous référer au schéma de données suivant lorsque vous travaillez avec:

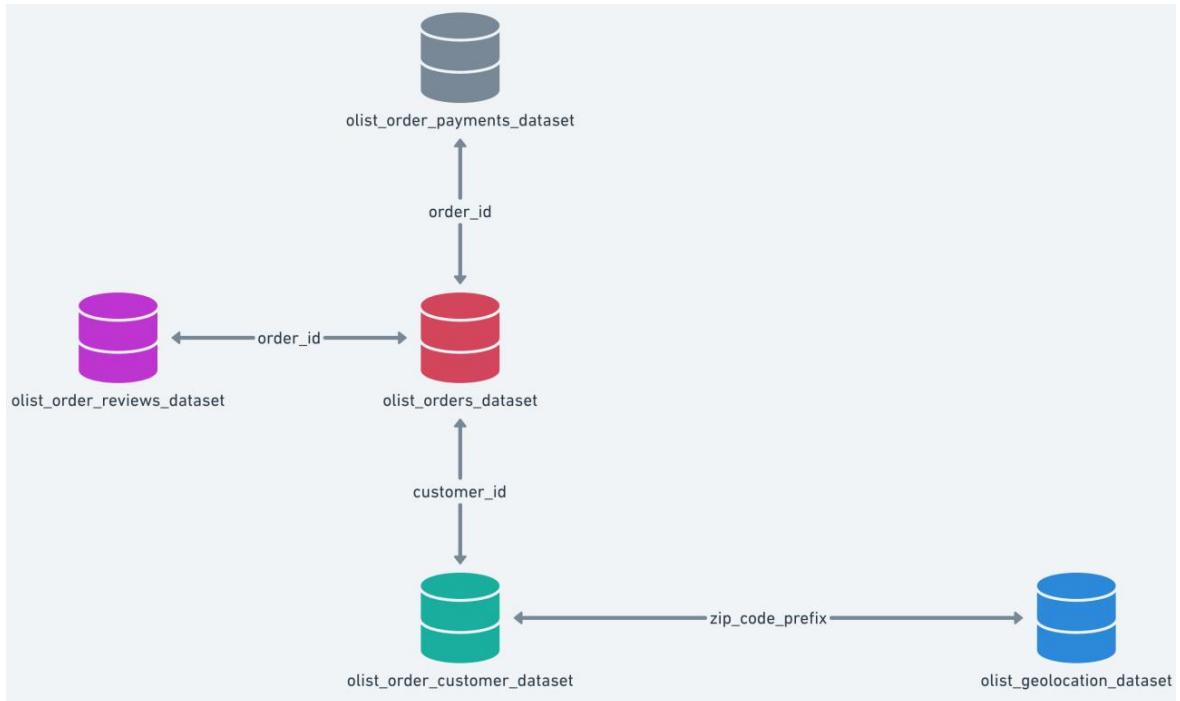
<https://www.kaggle.com/olistbr/brazilian-e-commerce>



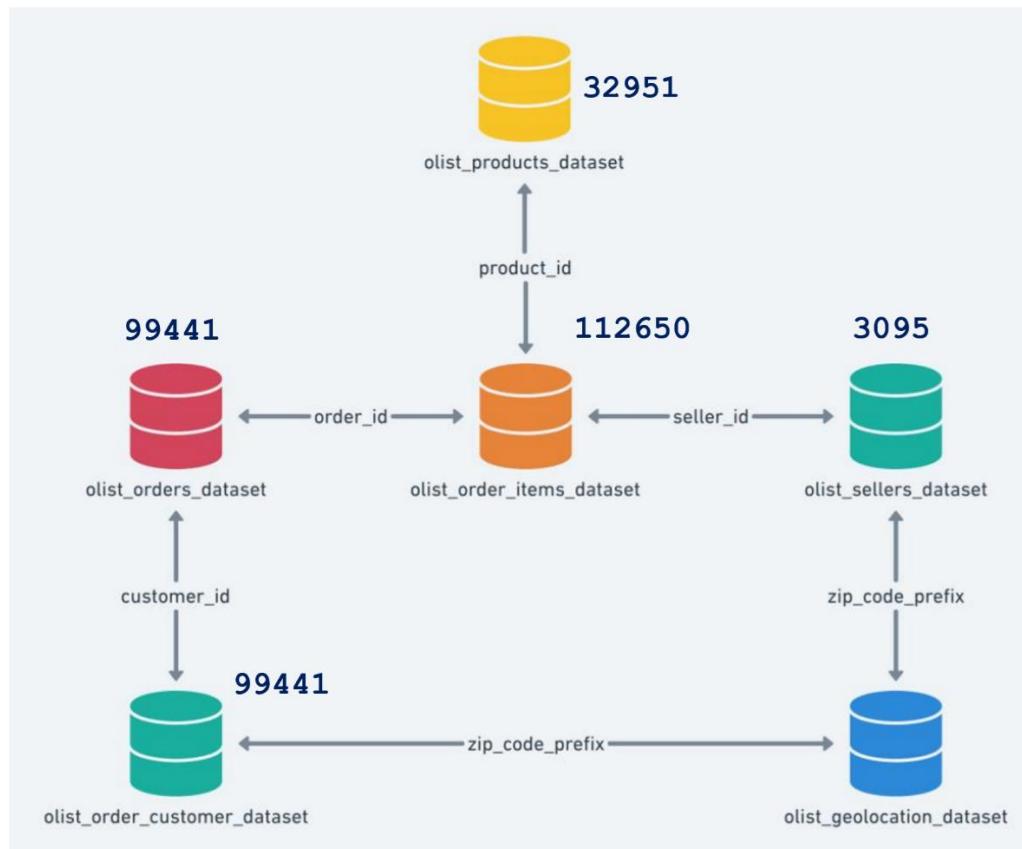
Les données de l'analyse

L'analyse est faite à partir des trois dataFrames consolidés :

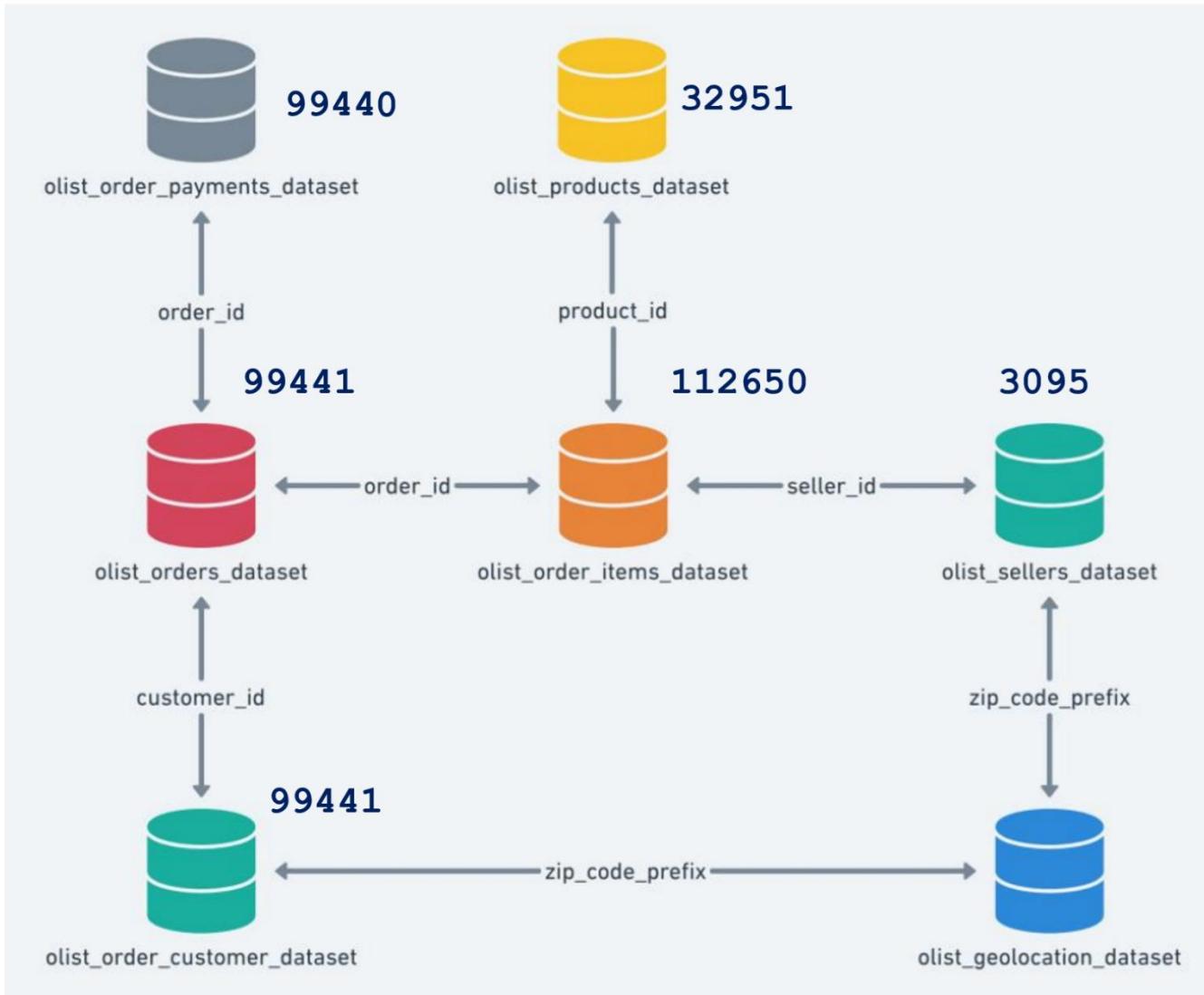
1. Analyse des ventes et avis clients contenant les données suivantes :



2. Analyse des achats contenant les données suivantes :

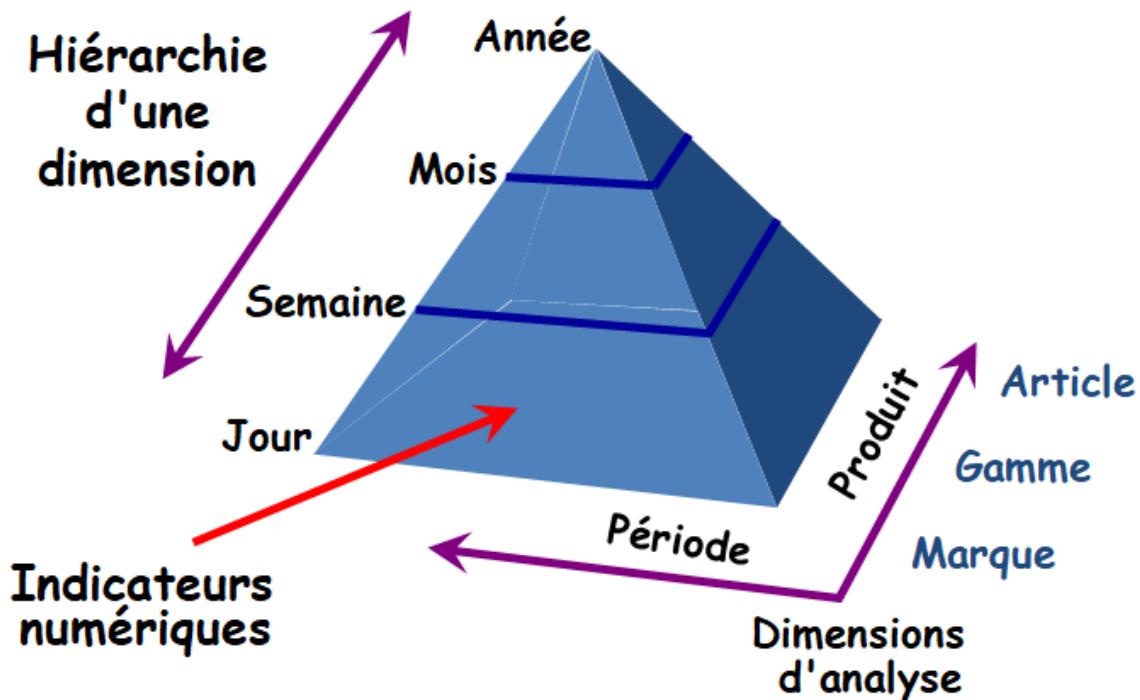


3. Analyse des achats (agrèges) et ventes contenant les données suivantes :



La constitution des deux modèles en étoile

Identifier pour chaque modèle les métriques, les dimensions et les hiérarchies pour chacune de dimension.



Voici un exemple de dimensions et les **hiérarchies** correspondantes :

annee		status	category_name
trimestre	annee_trimestre	order_id	10 product_id
mois	annee_mois	7 order_item_id	product_id
semaine	annee_semaine	state	
jour	annee_jour	city	
jour_semaine		zip_code	(lat,lng) - (lat_min,lat_max,lng_min,lng_max)
heure		customer_unique_id	
1 purchase_timestamp		customer_id	
annee		state	
trimestre	annee_trimestre	city	
mois	annee_mois	zip_code	
semaine	annee_semaine	customer_unique_id	(lat,lng) - (lat_min,lat_max,lng_min,lng_max)
jour	annee_jour	9 seller_id	
jour_semaine			
heure			
2 delivered_customer			
3 delivered_carrier			
4 estimated_delivery			
5 approved_at			

Voici un exemple des **métriques** du modèle :

orders - payments - reviews - customers - geolocation

value_boleto	int_boleto	score_1	comment_1	approuvee	creation_1	answer_1
value_credit_card	int_credit_card	score_2	comment_2	envoyee	creation_2	answer_2
value_debit_card	int_debit_card	score_3	comment_3	livree	creation_3	answer_3
value_not_defined	int_not_defined	score_4	comment_4	estimee	creation_4	answer_4
value_voucher	int_voucher	score_5	comment_5		creation_5	answer_5
value		score			creation	answer

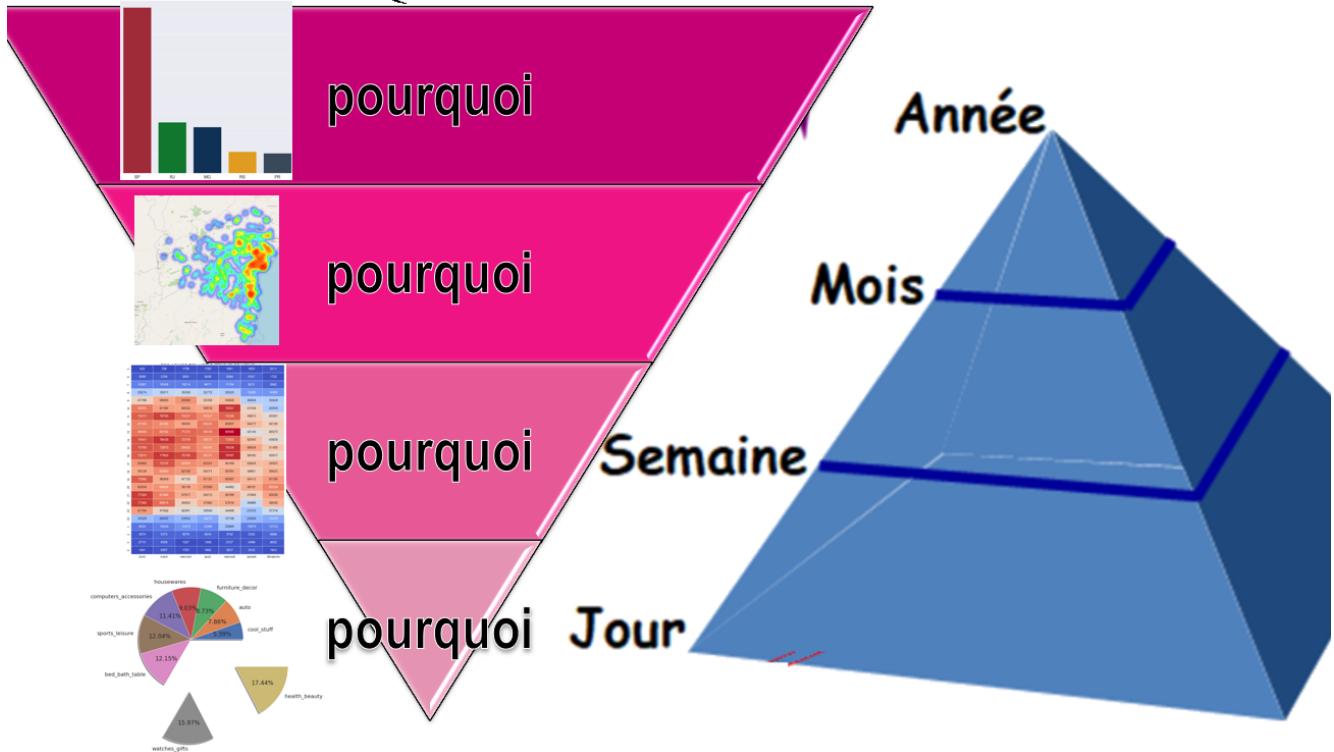
orders - items - products - sellers

price	weight_g	limit
freight_value	length_cm	
	height_cm	
	width_cm	
	name_lenght	
	description_lenght	
	photos_qty	

Attention : il s'agit d'un exemple avec une grande similitude avec vos données mais pas exploitable dans l'état.

L'analyse des deux modèles

Problème/Question/Recherche



Vous devez créer trois analyses pour chacune des étoiles.

Chaque analyse essaie de répondre (résoudre) à une question (problème, recherche ...) et doit prendre en compte au moins trois hiérarchies (vous devez commencer les rapports avec les données les plus agrégées avant de descendre dans les hiérarchies).

Chaque analyse doit avoir un minimum de 6-8 graphiques.

Une analyse est un document Notebook que vous sauvegardez en format html pour faciliter la lecture.

Le projet doit consolider les analyses dans une seule archive .zip ou .rar et vous devez me l'envoyer à l'adresse suivante « razvan@bizoi.fr » avant le 30/11/2025. Il est préférable de l'envoyer par <https://wetransfer.com> car la taille peut être conséquente.

Vous devez préparer une présentation de 15 à 20 minutes pour l'examen le 03/12/2025.