**🏷️ Brazilian E‑Commerce Public Dataset by Olist**

**What it includes:**

* ~100K orders (2016–2018)
* Related tables: customers, geolocation, orders, order\_items, order\_payments, products, sellers, category\_translation, etc.
* Rich relational structure for **data modeling** [kaggle.com+3kaggle.com+3kaggle.com+3](https://www.kaggle.com/datasets/saurabhshahane/ecommerce-text-classification?utm_source=chatgpt.com)[kaggle.com](https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce?utm_source=chatgpt.com)[kaggle.com](https://www.kaggle.com/datasets/abdelrahmanalimo/e-commerce-dataset?utm_source=chatgpt.com)

This covers *multiple tables* that mirror a real-world online retail platform.

**✅ Why this dataset is ideal:**

1. **Relational Tables**  
   Build star/snowflake schemas in PostgreSQL + Power BI.
2. **Complex Cleaning Scenarios**
   * Date gaps/time zones
   * Multiple shipping events
   * Payment splits, cancellations, refunded orders
3. **Insight Use Cases**
   * Customer lifetime value
   * Seller lead-time and performance
   * Payment method trends
   * Geographic sales distribution
4. **Massive but manageable size**  
   Enough volume to be realistic (~100k orders, ~660k items) without being overwhelming.

**📌 Suggested Project Plan**

**1. Define Business Scenario**

e.g., “Help Olist optimize seller performance and customer satisfaction by analyzing delivery times, payment methods, and geographic behavior.”

**2. KPIs to Compute & Visualize**

* ✅ Total Revenue, Total Orders, Avg. Order Value
* ✅ Order delivery time (shipping limit → delivery actual)
* ✅ Customer Lifetime Value
* ✅ Seller On-time Order Rate
* ✅ Payment Method Distribution
* ✅ Geographic Sales by State/City
* ✅ Cancellation & Refund Rates by Seller

**3. Data Cleaning (Power Query)**

* Fix date formats & missing dates
* Manage null geolocations or bad address data
* Standardize categorical fields (payment types, statuses)

**4. Pre-Aggregation Using SQL (PostgreSQL)**

* Calculate delivery times per order (delivery\_date - purchase\_date)
* Use **window functions**:
  + Compute cumulative spend per customer
  + Rank sellers by average delivery time or sales volume

**5. Data Modeling in Power BI**

* Import from PostgreSQL
* Create relationships: orders ↔ customers, orders ↔ order\_payments, order\_items ↔ products ↔ category

**6. Visualizations & Dashboard Suggestions**

* Overview page with core KPIs
* Seller performance dashboard
* Customer cohort & retention analysis
* Map-based sales insights (Brazil states map)

**🎯 What You Should Do Next:**

1. **Download Dataset**  
   🛒 Go to “Brazilian E‑Commerce Public Dataset by Olist” on Kaggle.
2. **Tell Me When You Have It**  
   I’ll help you:
   * Define step-by-step **ETL and cleaning flows**
   * Write SQL window function queries
   * Map out DAX measures for your Power BI dashboard
   * Draft a compelling project **README** for your GitHub résumé

**About Dataset**

**Brazilian E-Commerce Public Dataset by Olist**

Welcome! This is a Brazilian ecommerce public dataset of orders made at [Olist Store](http://www.olist.com/). The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil. Its features allows viewing an order from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers. We also released a geolocation dataset that relates Brazilian zip codes to lat/lng coordinates.

This is real commercial data, it has been anonymised, and references to the companies and partners in the review text have been replaced with the names of Game of Thrones great houses.

**Join it With the Marketing Funnel by Olist**

We have also released a [Marketing Funnel Dataset](https://www.kaggle.com/olistbr/marketing-funnel-olist/home). You may join both datasets and see an order from Marketing perspective now!

**Instructions on joining are available on this**[**Kernel**](https://www.kaggle.com/andresionek/joining-marketing-funnel-with-brazilian-e-commerce)**.**

**Context**

This dataset was generously provided by Olist, the largest department store in Brazilian marketplaces. Olist connects small businesses from all over Brazil to channels without hassle and with a single contract. Those merchants are able to sell their products through the Olist Store and ship them directly to the customers using Olist logistics partners. See more on our website: [www.olist.com](https://www.olist.com/)

After a customer purchases the product from Olist Store a seller gets notified to fulfill that order. Once the customer receives the product, or the estimated delivery date is due, the customer gets a satisfaction survey by email where he can give a note for the purchase experience and write down some comments.

**Attention**

1. An order might have multiple items.
2. Each item might be fulfilled by a distinct seller.
3. All text identifying stores and partners where replaced by the names of Game of Thrones great houses.