Renata Batista

GitHub.: <https://github.com/renatab-code>

Tutor.: Georgi Dimchev

26. May 2025

***Project Title.:***

**“Unlocking Brazilian E-Commerce:**

**Customer Behavior, Delivery Patterns & Sales Trends”**

**Project Objective:**

To build an **interactive Tableau Public dashboard**that visually narrates insights uncovered through **advanced exploratory analysis** conducted in Python, using the Olist Brazilian E-Commerce dataset. The dashboard will be structured to support business decision-making related to **customer satisfaction, delivery logistics**, and**sales optimization.**

**Key Research Questions & Hypotheses:**

**1. Customer & Order Behavior**

* **Q1:** What are the dominant customer segments based on location, order frequency, and product categories?
* **Q2:** Can we cluster customers to uncover distinct behavioral patterns (e.g., high-value vs. frequent buyers)?

**2. Sales & Category Trends**

* **Q3:** Which product categories generate the most revenue and how do they perform over time?
* **Q4:** How does seasonality affect sales volumes across different product categories?

**3. Delivery Performance**

* **Q5:** What are the regional differences in delivery time, and how do they impact customer review scores?
* **Q6:** Can delivery delays be predicted based on product, seller, and order features?

**4. Review & Sentiment Insights**

* **Q7:** What factors contribute to negative review scores?
* **Q8:** Is there a correlation between review scores and delivery time or payment method?

***Data Source***

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

**Source:** [Kaggle - Brazilian E-Commerce Public Dataset by Olist](https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce?resource=download)

***Summary:***This dataset is provided by Olist, a Brazilian e-commerce company that operates a marketplace platform allowing small businesses to sell products through major online retail channels. The dataset contains detailed information about orders made on the Olist store from 2016 to 2018.

It includes a comprehensive set of interrelated CSV files capturing the complete sales lifecycle, including customer data, order details, product information, payment transactions, seller profiles, and customer reviews.

***Files Included:***

* olist\_orders\_dataset.csv – order status, purchase timestamp, delivery details
* olist\_customers\_dataset.csv – customer location and unique identifiers
* olist\_order\_items\_dataset.csv – product details per order
* olist\_products\_dataset.csv – product attributes like category, name length, etc.
* olist\_sellers\_dataset.csv – seller location and IDs
* olist\_order\_payments\_dataset.csv – payment method and installments
* olist\_order\_reviews\_dataset.csv – customer review scores and comments
* product\_category\_name\_translation.csv – English translation of product category names

***Use Case:***The dataset is ideal for a wide range of e-commerce analytics tasks such as customer segmentation, delivery performance analysis, product recommendation systems, seller performance evaluation, and sentiment analysis from customer reviews. It serves as a rich resource for both descriptive and predictive data modeling.

***Size and Format:***The data is provided in CSV format and totals approximately 10 files, with over 100,000 orders and multiple data points across customers, products, and reviews.

***License:***The dataset is publicly available under the terms set by the Kaggle platform and the Olist company.