

# *Final Project On Data Analysis Using Power BI*

## 1. Dataset Overview

**Source:** Brazilian E-Commerce Public Dataset by Olist (2016-2018).

**Orders:** 100,000 orders from multiple marketplaces.

### **Key Files:**

- **Customers:** Customer details (unique IDs, ZIP codes, geographic info).
- **Geolocation:** Geographic data linked to ZIP code prefixes.
- **Order Items:** Product details within each order.
- **Orders:** Comprehensive order information (status, timestamps).
- **Products:** Product details (ID, categories, dimensions).
- **Sellers:** Seller information (unique seller IDs, ZIP codes).
- **Product Category Translation:** Translations for product categories to English.
- **Order Payments:** Payment details for each order.

## 2.Data Cleaning and Transformation

I conducted thorough data cleaning and transformation to ensure accuracy, consistency, and usability across all datasets. This involved:

- **Removing Duplicates:** Eliminated duplicate records to maintain unique entries.
- **Handling Missing Values:** Addressed missing values by either filling them with appropriate substitutes or excluding incomplete records.
- **Standardizing Formats:** Ensured uniformity in data formats, such as dates and currency, for consistency in analysis.
- **Normalizing Data:** Adjusted variables to be on a comparable scale for better analysis.
- **Creating Calculated Fields:** Developed new fields, such as total revenue per order, to provide deeper insights.
- **Structuring Data:** Organized the data into appropriate formats, ensuring that relationships between tables were clearly defined for effective modeling.

## 3.Data Model Overview: Fact Table and Dimension Tables

### Fact Table

- Order Items Fact Table

### Dimension Tables

- Customer Dimension Table
- Geolocation Dimension Table
- Orders Dimension Table
- Products Dimension Table
- Sellers Dimension Table
- Product Category Translation Dimension Table
- Order Payments Dimension Table

## 4.Key Business Questions

### ☐ **Customer Insights**

What is the average number of orders placed per customer?

Which geographic regions have the highest customer concentration, and how does this impact overall sales?

### ☐ **Seller Performance**

Who are the top-performing sellers in terms of revenue, and what products contribute most to their success?

How does seller performance vary across different geographic regions?

### ☐ **Product Analysis**

Which product categories generate the highest sales, and what trends can be observed over time?

Are there specific products with high return rates that require further investigation?

### ☐ **Order Trends**

What is the average order value, and how does it vary by customer demographics or geographic location?

During which months do order volumes peak, and what factors contribute to these trends?

### ☐ **Payment Preferences**

What are the most common payment methods used by customers, and do preferences vary by region?

Is there a correlation between the number of payment installments and the total order value?

### ☐ **Geolocation Insights**

Which areas experience the fastest delivery times, and are there regions that require improvement in shipping efficiency?

How do shipping costs vary by geographic location, and how do they affect overall sales?

## 5. Insights and Recommendations for Business Growth

- **Insights:**

**1.Customer Engagement:** Strong potential for customer loyalty, but significant drop-off after the first purchase suggests a need for improved retention strategies.

**2.Seller Performance:** Notable variability in seller performance; high-performing sellers offer unique products, highlighting the importance of differentiation.

**3.Product Trends:** Electronics and home goods generate the highest revenue, with seasonal fluctuations during holidays.

**4.Geographic Disparities:** Sales are concentrated in urban areas, indicating an opportunity for targeted marketing in underserved rural regions.

**5.Payment Preferences:** Varied payment methods based on location; customers in some areas prefer flexible installment payments.

- **Recommendations:**

**1.Enhance Loyalty Programs:** Implement rewards for repeat customers to improve retention and increase order frequency.

**2.Empower Top Sellers:** Provide marketing support and collaboration opportunities to high-performing sellers.

**3.Optimize Product Offerings:** Expand inventory in high-demand categories and refresh product lines based on feedback.

**4.Targeted Marketing:** Develop region-specific marketing campaigns to boost brand awareness in underserved areas.

**5.Flexible Payment Options:** Introduce diverse payment methods, including installment plans, to cater to customer preferences.

**6.Improve Logistics:** Streamline shipping processes for faster delivery and consider partnerships with local couriers.

## 6.DAX Functions Implementation

### 1. Added Columns

- **TotalRevenue:** I added a column to calculate the total revenue for each order by summing the pricing and freight value.
- **PaymentMethodCategory:** I created a column that categorizes the payment methods into specific groups, such as "Card," "Alternative," and "Other," based on the type of payment used.
- **DeliveryDuration:** I introduced a column to calculate the delivery duration in days by determining the difference between the order purchase timestamp and the order delivery date.

### 2. Added Measures

- **TotalRevenueMeasure:** I created a measure that calculates the total revenue across all orders by summing the total revenue column.
- **AverageOrderValue:** I established a measure to calculate the average order value by averaging the total revenue.

- **TotalOrders:** I added a measure to count the distinct number of orders, providing insights into the total orders placed.
- **TotalCustomers:** I implemented a measure to count the distinct number of customers, helping to understand the customer base.
- **RevenueByPaymentMethod:** I created a measure that calculates the total revenue segmented by payment method, allowing for analysis of revenue contributions by different payment types.

## 7.visualizations

These visualizations provide a clear and comprehensive view of Olist's performance in sales, customer reach, and revenue sources, helping guide strategic decisions.

- **Yearly Trends:** Charts that show Total Price plus Freight over the years and separate values for Total Price and Freight by year.
- **Product Categories:** Top 10 and Bottom 10 product categories based on total price.
- **Key Metrics Cards:** Quick summaries of total price, freight, revenue, orders, distinct customer locations, weekend/weekday counts, total products, customers, and sellers.
- **Gauges:** Indicators for delivered orders and average order value, showing progress toward targets.
- **KPIs:** Monthly order counts compared to targets for each month.
- **Order Patterns:** Pie chart showing the count of order states (e.g., delivered, canceled) by weekend or weekday.
- **Geographic Distribution:** A map chart of orders by city and state, showing geographic trends.
- **City-Based Spending:** Clustered column chart displaying average price and payment values by customer city.
- **Product Revenue Contribution:** Stacked column chart of total revenue by product category.
- **Total Orders Over Time:** Line chart showing total orders by year to identify growth trends.
- **Revenue by Payment Type:** Pie chart showing how revenue is distributed across different payment methods.