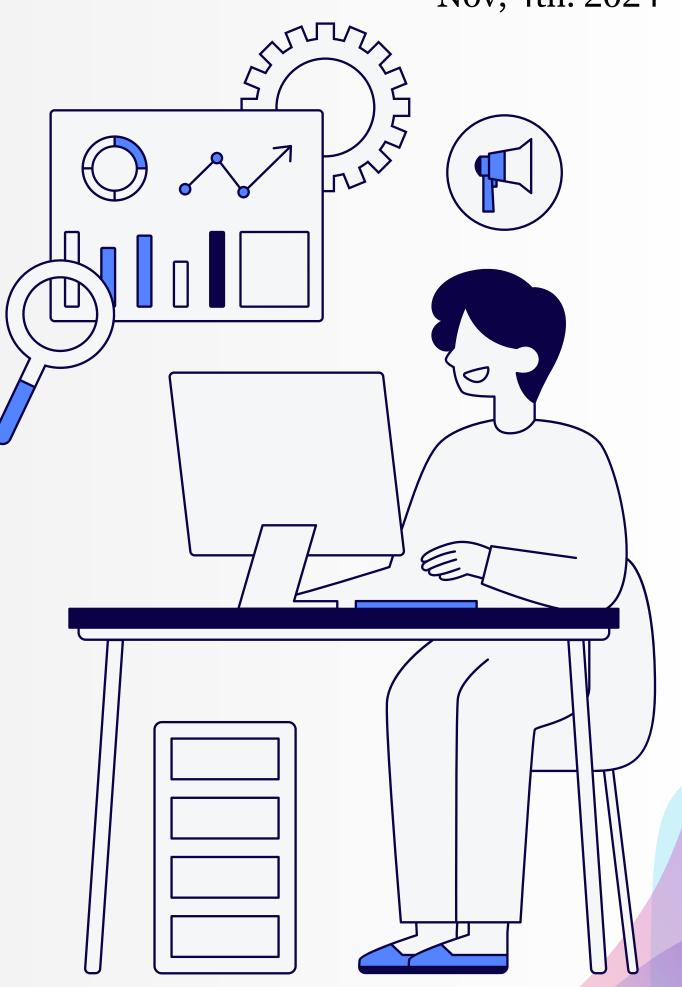
E-COMERCE SALES ANALYSIS

Data-Driven Insights and Recommendations

By: Fatma Mohamed Abdel Fattah



DATA PROFILE

This is a Brazilian ecommerce public dataset of orders made at **Olist Store.** The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil.

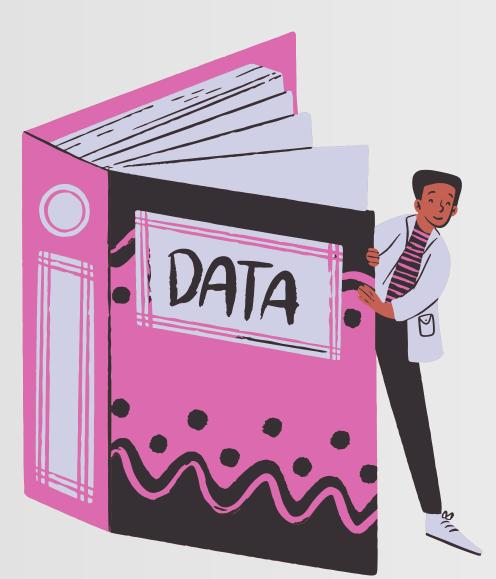


DATA TABLES

- Customers Dataset: Customer details and location; unique and repeat customer IDs.
- Sellers Dataset: Seller info and location.
- Orders Dataset: Core dataset; links all other tables, with order status, timestamps, and delivery dates.
- Order Items Dataset: Item details within orders; links products, sellers, and pricing.
- Payments Dataset: Payment methods, values, and installment details.
- Products Dataset: Product names and unique IDs.

DATA PROFILE

This is a Brazilian ecommerce public dataset of orders made at **Olist Store.** The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil.



KEY COLUMNS

- customer_id, order_id, seller_id, product_id: Link tables for customer, order, product, and seller details.
- price, freight_value: Cost and shipping for each item.
- order_status, timestamps: Order lifecycle and timing info.

DATA IMPORTING AND TRANSFORMING

DATA IMPORT

- Tool: Power BI
- File Type: Microsoft Excel Comma Separated Values File (.csv)



DATA TRANSFORMATION

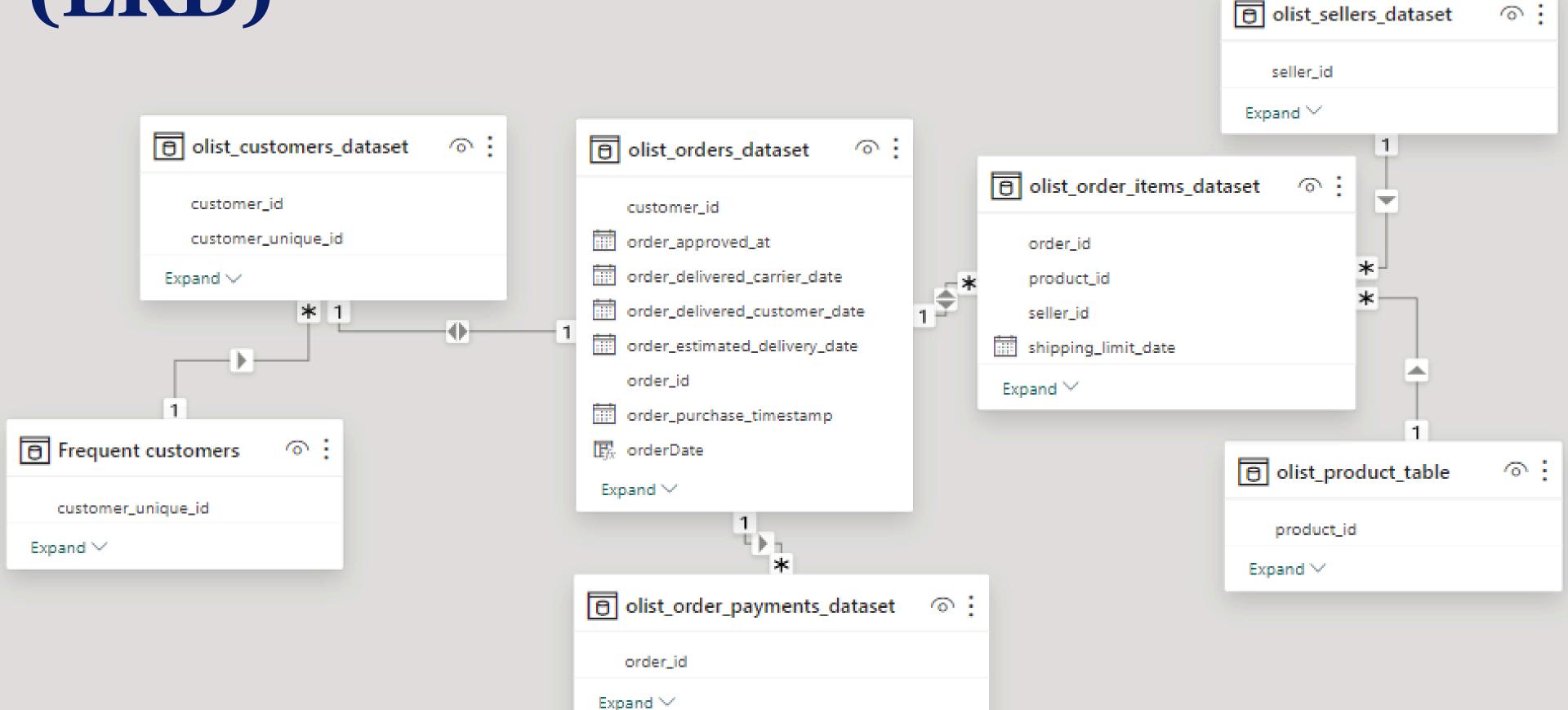
- Handled some **column mis-identification**, using *First* row as header function
- Dropped **irrelevant** columns i.e. seller_zip_code and customer_zip_code

DATA CLEANING

- **Merged** the *Products_dataset* with the *Product_category_translation*dataset on the product_id. Then, **dropped** the Brazilian

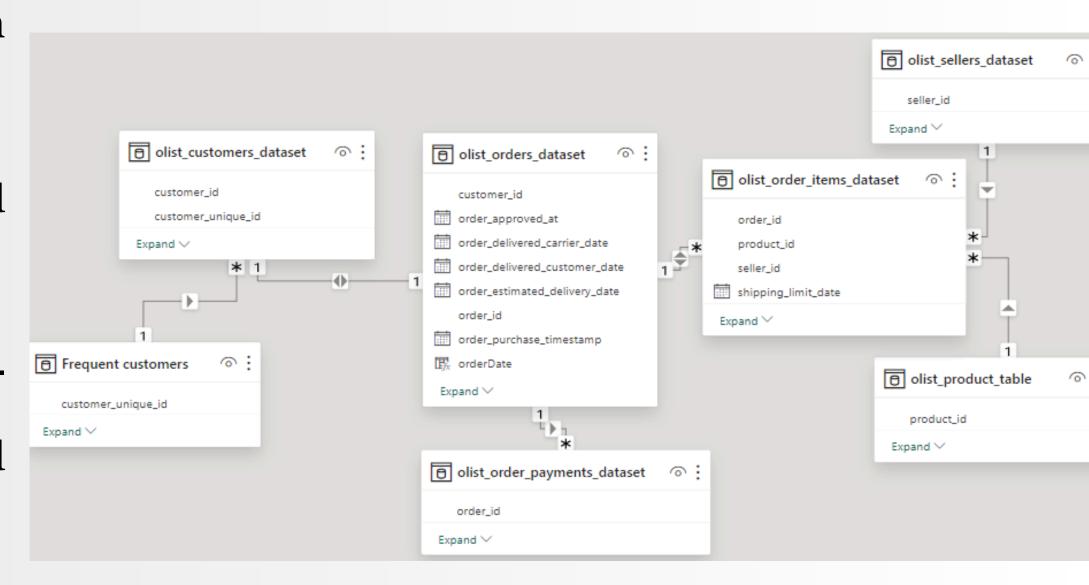
 category_name; reducing redundancy
- **Dropped** the *geolocation dataset*; since it is out of my scope of analysis
- Created an "OrderDate" column to ensure Date consistency of the orders
- **Created** a Summarization table for Frequent customers with more than **1** visit to my <u>Olist Store</u>

DATA MODELLING (ERD)



DATA MODELLING (ERD)

- Structure: SnowFlake schema with the "Orders" table as the Fact table
- Connecting the <u>customers</u>, <u>payments</u>, and <u>order items</u> as the **Dimensional** Tables
- The "order items" table has two subdimensional tables; <u>sellers</u> and <u>products</u>
- customers has one sub-dimensional table; <u>Frequent Customers</u>



BUSINESS QUESTIONS

- What are the total sales?
- What are the sales metrics for different time periods? (YTD, QTD, MTD)
- Which states and cities have the highest sales?
- How do sales vary by date?
- What percentage of orders were <u>delivered successfully</u>? And how many of them were <u>delivered late</u>?
- How many product categories are in my store and their average price?

BUSINESS QUESTIONS

- How many customers visited my E-Store? And How many of them were <u>Loyal customers</u> who returned?
- How do my customers prefer to pay?
- Which product categories are the best sellers?
- What are the sales figures for each product category?
- How many sellers are there, and what are their sales figures?
- What is the average delivery duration compared to the estimated delivery duration?



Overview

Sales Performance

E-commerce Sales Dashboard

Go to Detailed Analysis

27

States

Filters

YEAR

2016

2017

2018

MONTH

January

February

March

April

May

WEEKDAY

Friday

Monday

\$16M **Total Sales**

Delivered Orders Pct

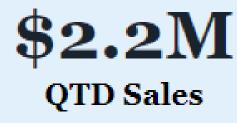
Top Products Sales 17% -23% **Products** 19% health beauty watches gifts bed_bath_table sports leisure computers_ac... 21% 20%

8% \$8.8M YTD Sales

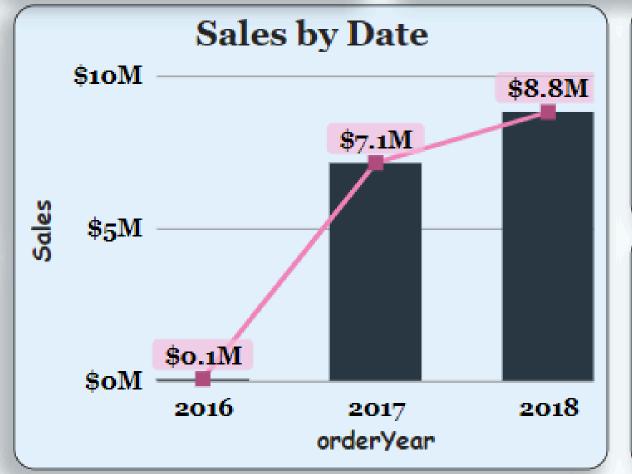
Oops! Late Deliveries

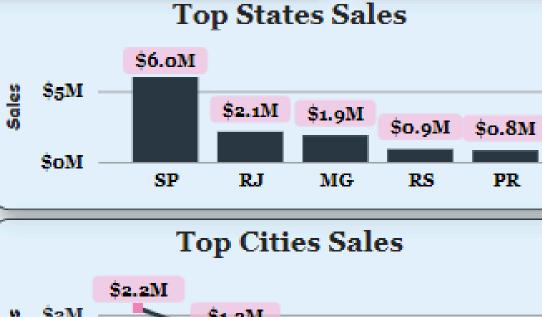
4119 Cities

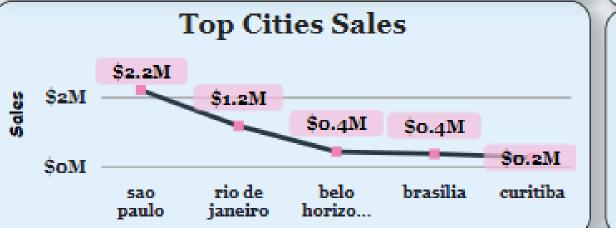
PR



MTD Sales





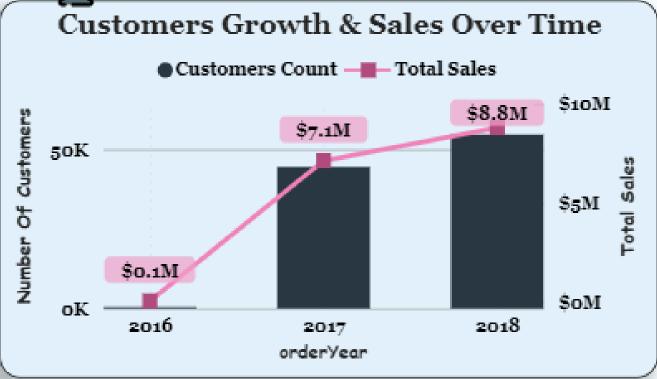




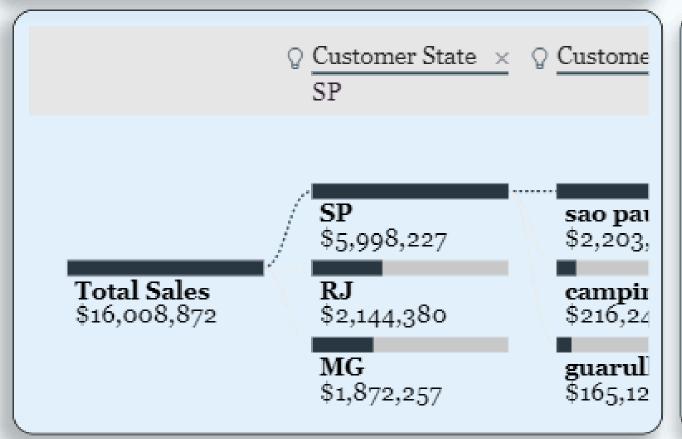
Customer-Wise Analysis

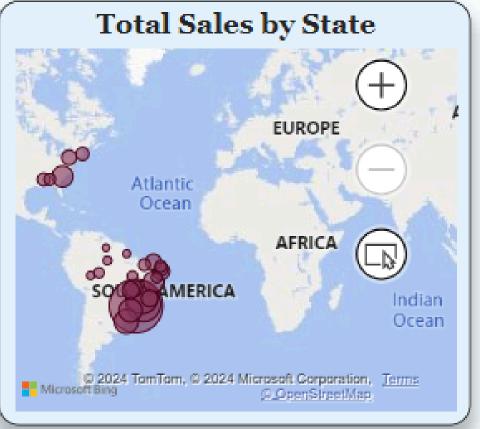


Customer and Sales Analysis by Time and Location







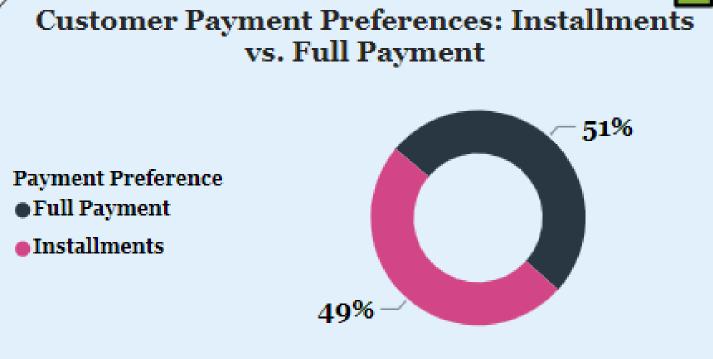


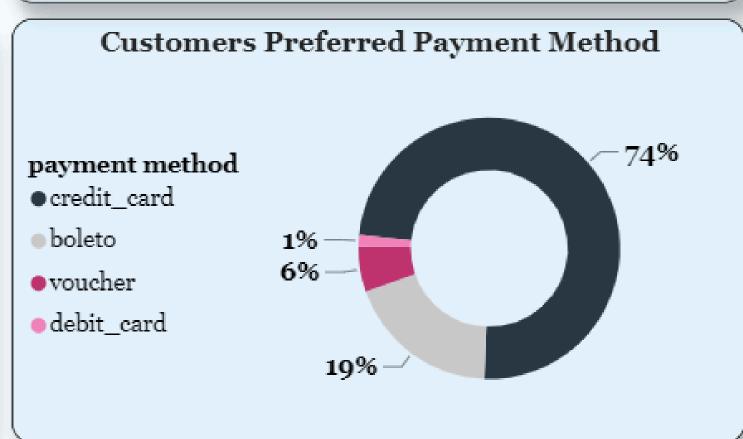














Product-Wise Analysis

72 Products

\$120.7
Average Product Price

BestSellers Product
Categories

bed_bath_table

health_beauty

sports_leisure

watches_gifts

All Products Sales in The E-Store

The Products in Our E-Store	Total Sales
health_beauty	\$1,448,730
watches_gifts	\$1,310,893
bed_bath_table	\$1,265,918
sports_leisure	\$1,166,061
computers_accessories	\$1,066,264
furniture_decor	\$932,340
housewares	\$793,239
cool_stuff	\$729,806
auto	\$688,628
garden_tools	\$596,759
tovs	\$566,025
Total	\$16,008,872

Sellers-Wise Analysis

YEAR

2016 2017

2018

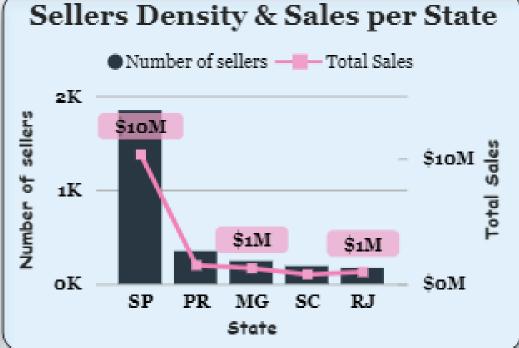
3K Sellers

Average Actual Delivery Duration Vs Average Estimated ...



Sales Of The Top Sellers seller_id Total Sales 4869f7a5dfa277a7dca6462dcf3b52b2 \$253,369 4a3ca9315b744ce9f8e9374361493884 \$248,764 53243585a1d6dc2643021fd1853d8905 \$235,857 7c67e1448b00f6e969d365cea6b010ab \$241,168 fa1c13f2614d7b5c4749cbc52fecda94 \$205,834 Total \$1,183,407





OBSERVED INSIGHTS

- Total Sales: \$16M revenue.
- Sales by Year: **Grows** from 2016 to 2018.
- Successful Delivery: 97% rate.
- Late Deliveries: 8% of orders were delivered later than estimated.
- Delivery Duration: The average estimated duration was **24 days** but our E-Store could achieve an average of only **12 days**
- Sales by Region
 - Top States: **SP**, RJ, MG, RS, PR
 - Top Cities: Sao Paulo, Rio de Janeiro, Belo Horizonte, Brasilia, Curitiba

RECOMMENDATIONS

- With **8%** of orders delivered late, focus on **optimizing logistics** and supply chain processes to reduce delays.
 - Consider partnering with more reliable carriers or implementing better tracking

systems.

- States like SP, and RJ and cities like Sao Paulo and Rio de Janeiro, show strong sales.
 - Consider opening more distribution centers or local stores in these areas to capitalize on the demand.

CONCLUSION

In summary, our e-commerce platform is thriving with impressive sales figures and high on-time delivery rates. By focusing on our top-performing product categories and expanding in high-demand regions, we can continue to drive growth and enhance customer satisfaction.

