# Brazilian E-Commerce Public by Olist

# Agenda

Introduction

Import Data

Transformation Stage
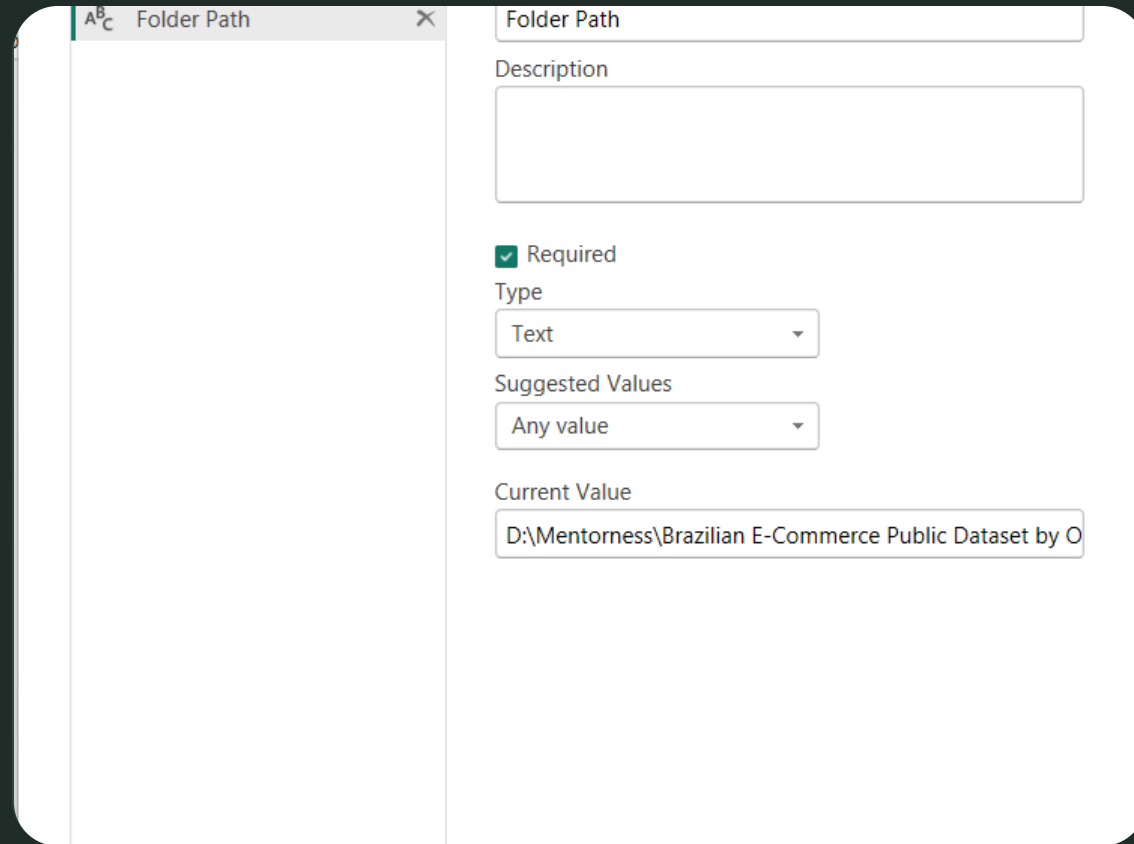
Data Modeling

Visualization

# Introduction

- 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.
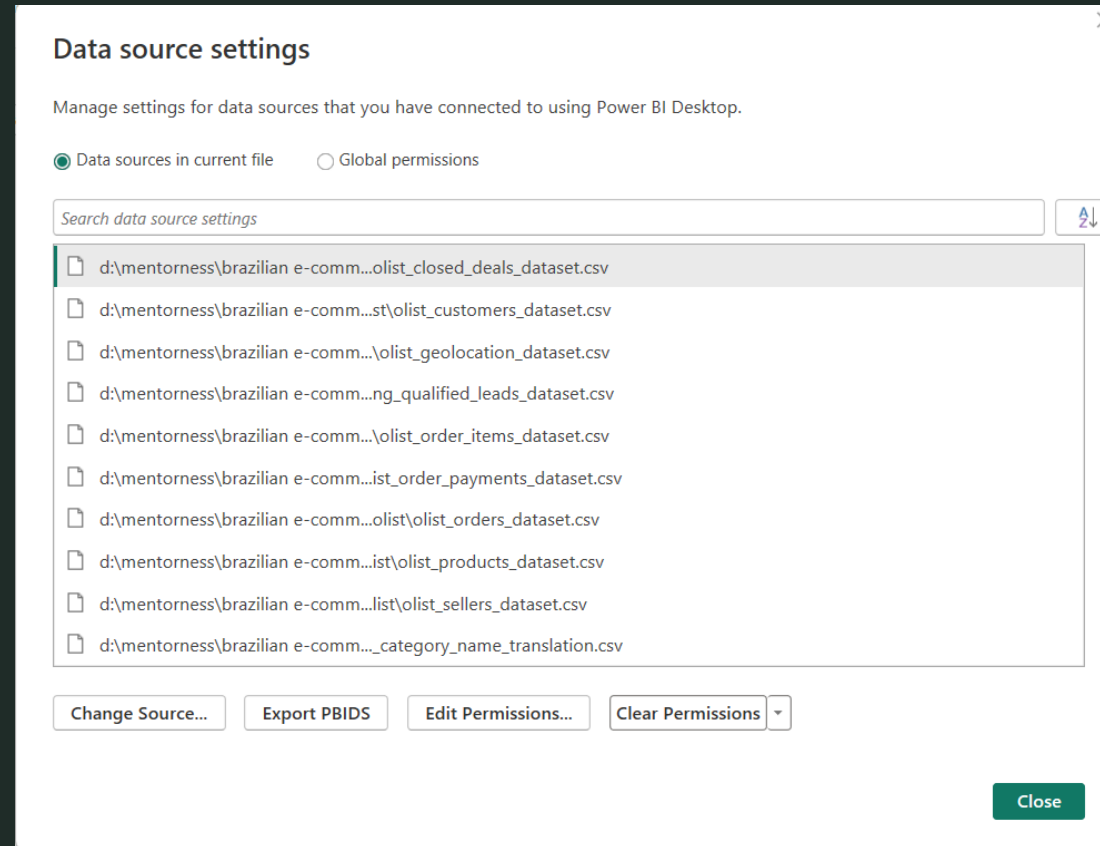


OLIST
ECOMMERCE

# Import Data

ABC Folder Path   ✕

Folder Path

Description

☑ Required

Type

Text ▾

Suggested Values

Any value ▾

Current Value

D:\Mentorness\Brazilian E-Commerce Public Dataset by O

# Make a Parameter to Put The File Path of the Data

# Import All Files from The Source of the Parameters

# Transformation Stage

# Change Type of Date Format

# Replace Null values with (Unkown,0) Values

# Create Order Status and Payment Table

# Merge Tables (Customers ,Order Items) with Orders

# Merge Tables (Order Status) with Orders and Payment With (Order Payment)

## Merge

Select a table and matching columns to create a merged table.

### Orders

| order_id | customer_id | order_status | order_purchase | order_a... |
|----------|-------------|--------------|----------------|-----------|
| e481f51cbdc54678b7cc49136f2d6af7 | 9ef432eb6251297304e76186b10a928d | delivered | 10/2/2017 | |
| e481f51cbdc54678b7cc49136f2d6af7 | 9ef432eb6251297304e76186b10a928d | delivered | 10/2/2017 | |
| e481f51cbdc54678b7cc49136f2d6af7 | 9ef432eb6251297304e76186b10a928d | delivered | 10/2/2017 | |
| 00010242fe8c5a6d1ba2dd792cb16214 | 3ce436f183e68e07877b285a838db11a | delivered | 9/13/2017 | |

**Orders Status** ▾

| Status_ID | order_status |
|-----------|--------------|
| 1 | delivered |
| 2 | invoiced |
| 3 | shipped |
| 4 | canceled |
| 5 | processing |

**Join Kind**

Left Outer (all from first, matching from second) ▾

☐ Use fuzzy matching to perform the merge

▷ Fuzzy matching options

[ OK ]  [ Cancel ]

## Merge

Select a table and matching columns to create a merged table.

### Order Payments

| order_id | payment_sequential | payment_type | payment_installments | payment_valu... |
|----------|--------------------|--------------|----------------------|-----------------|
| b81ef226f3fe1789b1e8b2acac839d17 | 1 | credit_card | 8 | 9 |
| a9810da82917af2d9aefd1278f1dcfa0 | 1 | credit_card | 1 | 2 |
| 25e8ea4e93396b6fa0d3dd708e76c1bd | 1 | credit_card | 1 | 6 |
| ba78997921bbcdc1373bb41e913ab953 | 1 | credit_card | 8 | 10 |

**Payments** ▾

| Payment_ID | payment_type |
|------------|--------------|
| 1 | credit_card |
| 2 | boleto |
| 3 | voucher |
| 4 | debit_card |
| 5 | not_defined |

**Join Kind**

Left Outer (all from first, matching from second) ▾

☐ Use fuzzy matching to perform the merge

▷ Fuzzy matching options

[ OK ]  [ Cancel ]

# Data Modeling

# Snowflake Schema

- Orders as Fact Table

- With Dimension

- Sellers relation With Close Deal

- Customer

- Payment

- Order Status
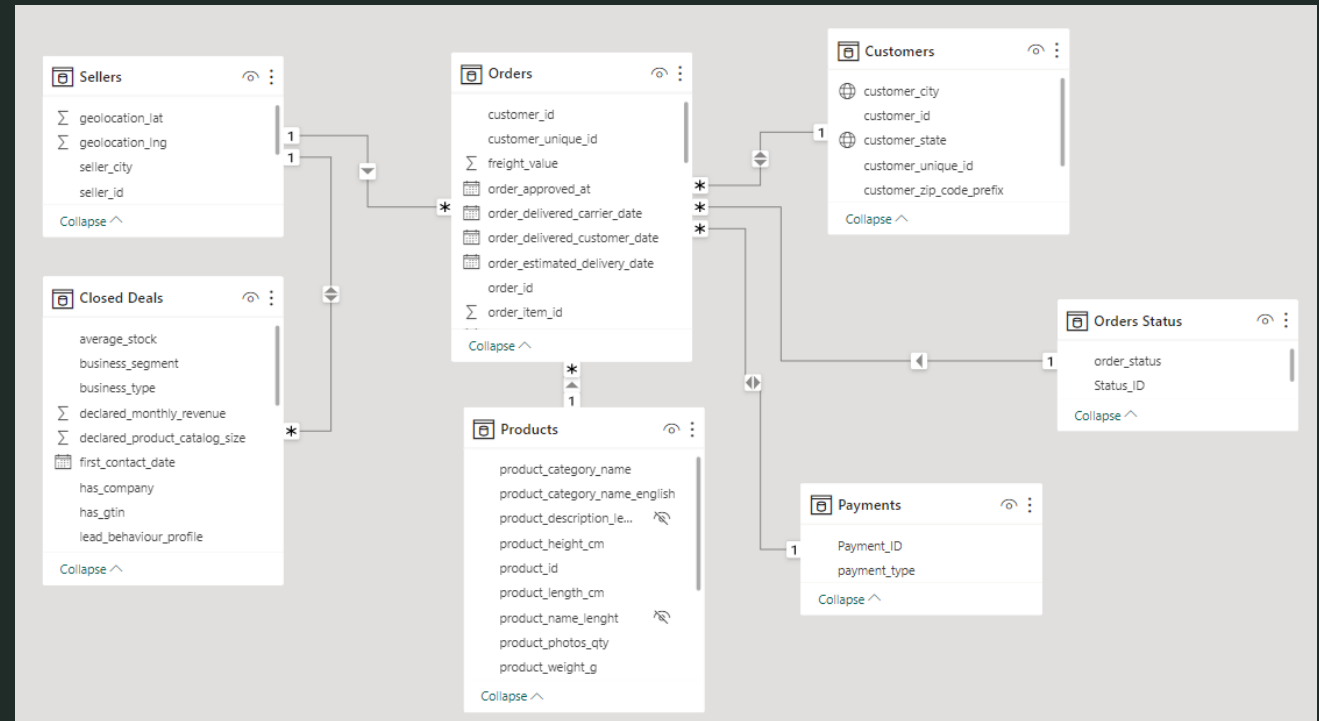
# Visualization

# Creating Measures

AVG
- Average Orders per Quarter
- Average Orders per Sellers
- AverageDealSize
- AVG_Freight
- AVG_MonthlyRevenue
- AVG_OrderItem
- Avg_Price

Count
- No_BussinessType
- No_CloseDeal
- No_Customer
- No_HasCompany
- No_HasGitn
- No_LeadType
- No_OrderItem
- No_Orders
- No_OrderStatus
- No_Product
- No_ProductCategory
- No_Sellers

Sum
- SalesbyYear2017
- SalesbyYear2018
- SalesGrowth
- Total_Cost
- Total_Freight
- TotalMonthlyRevenue

# Sales Performance

- sales performance dashboard involves gathering and presenting key metrics and KPIs related to sales activities and outcomes

# Orders Management

- An orders management dashboard is crucial for monitoring and optimizing the order fulfillment process

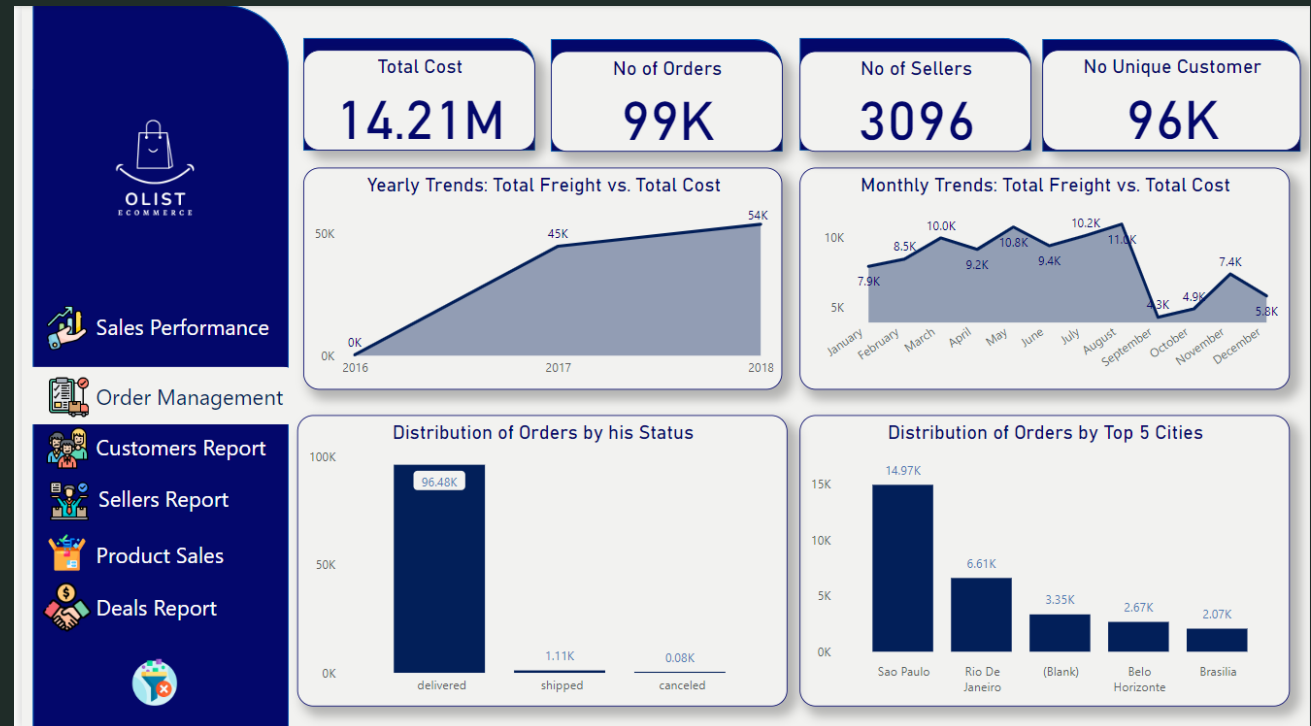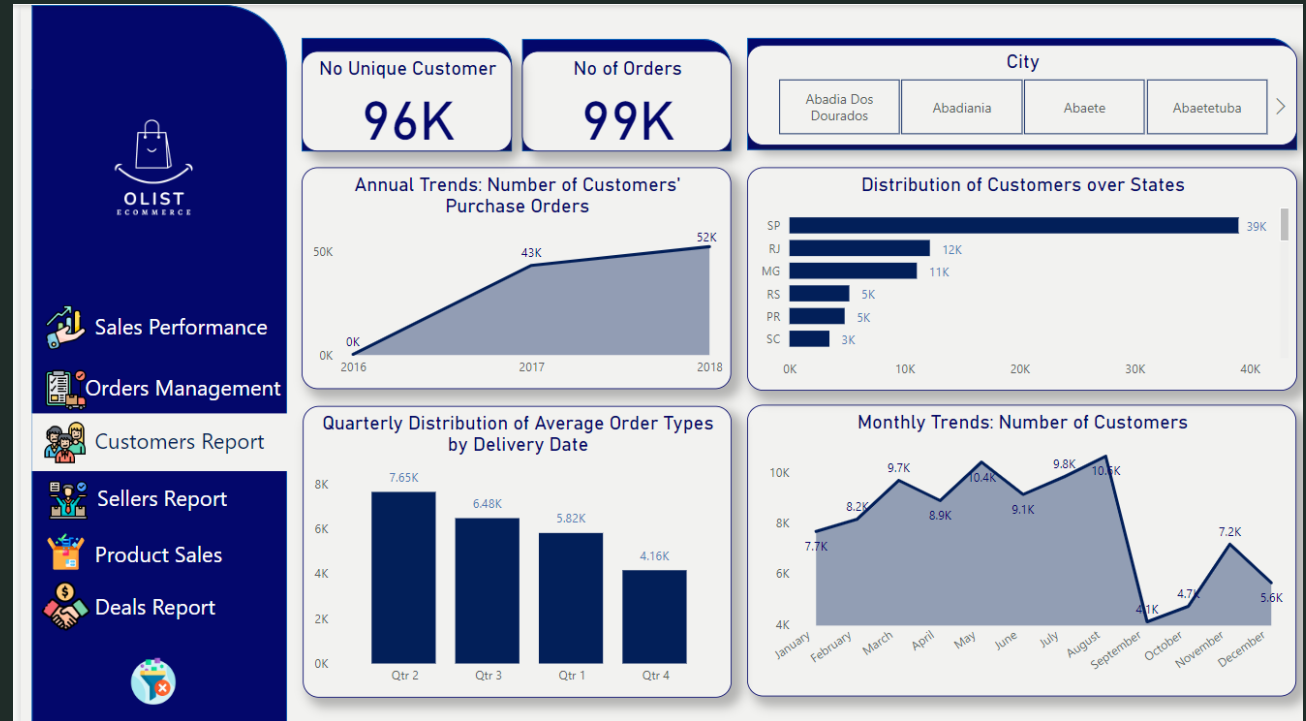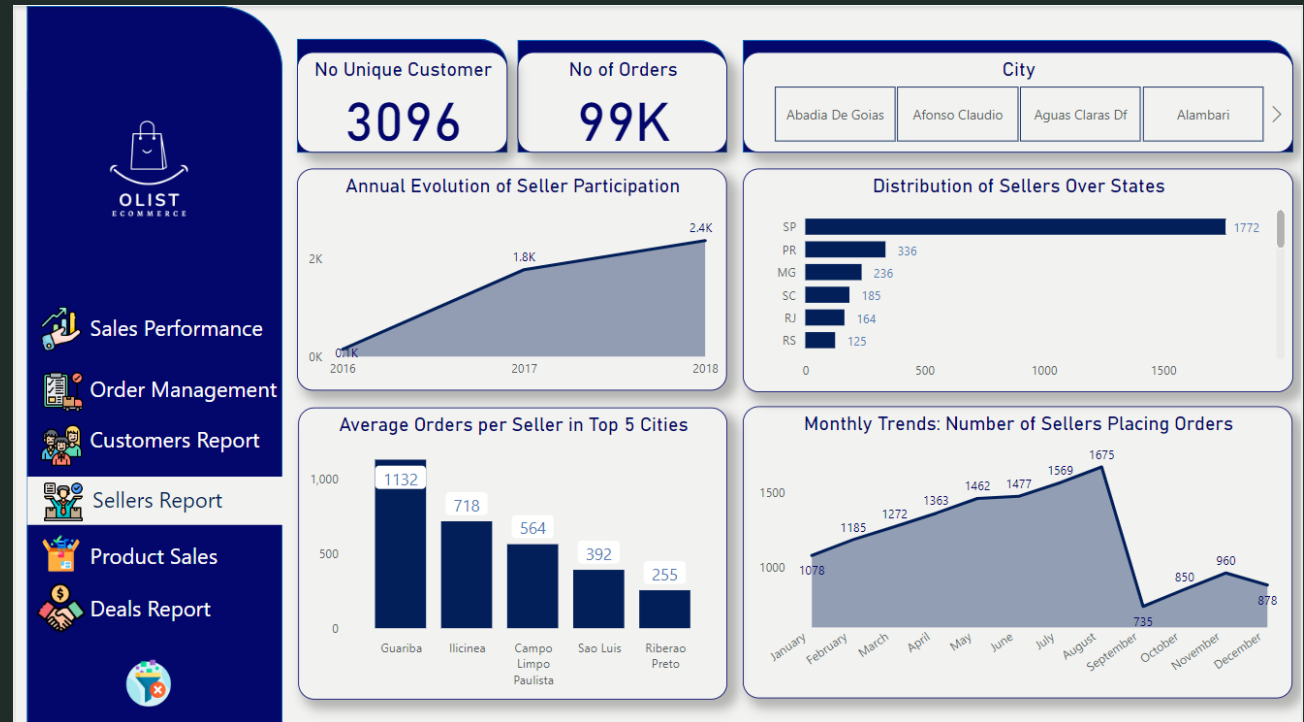# Customers Report

- A customer report dashboard provides valuable insights into customer behavior, preferences, and interactions with the business
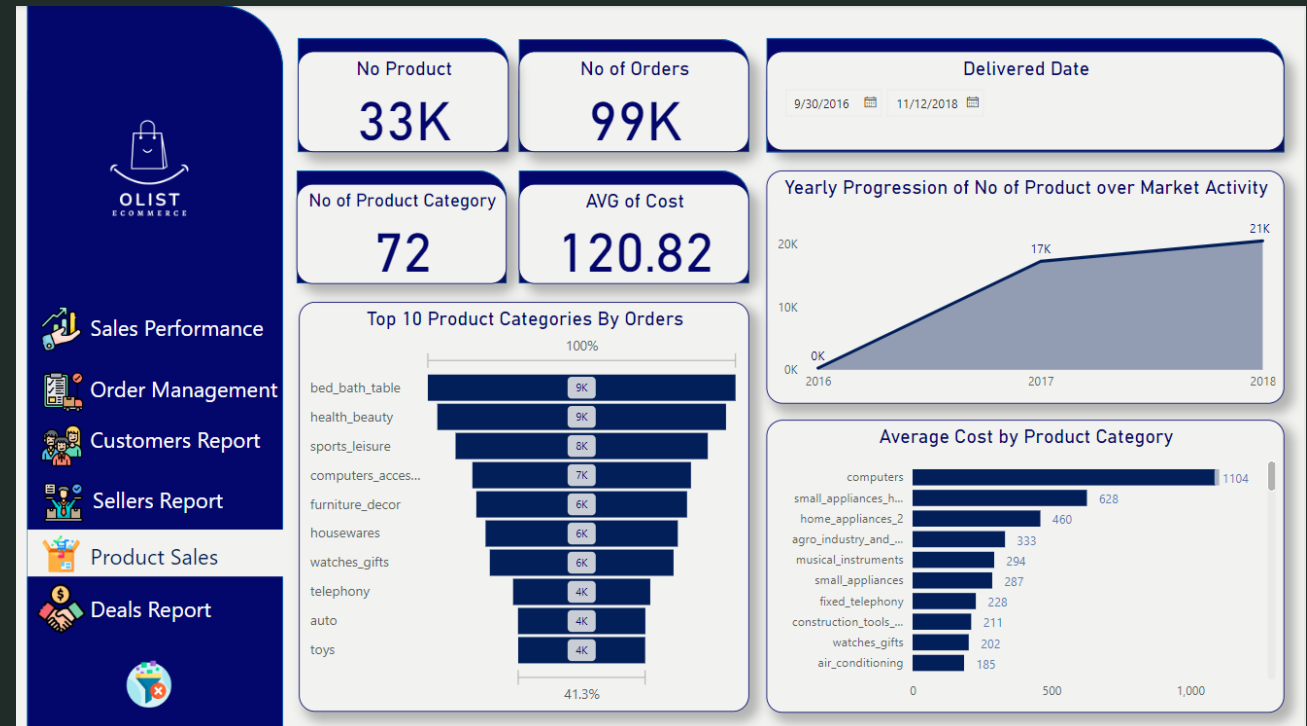
# Sellers Report

- A seller's report dashboard is essential for monitoring the performance of individual sellers or sales teams within the organization
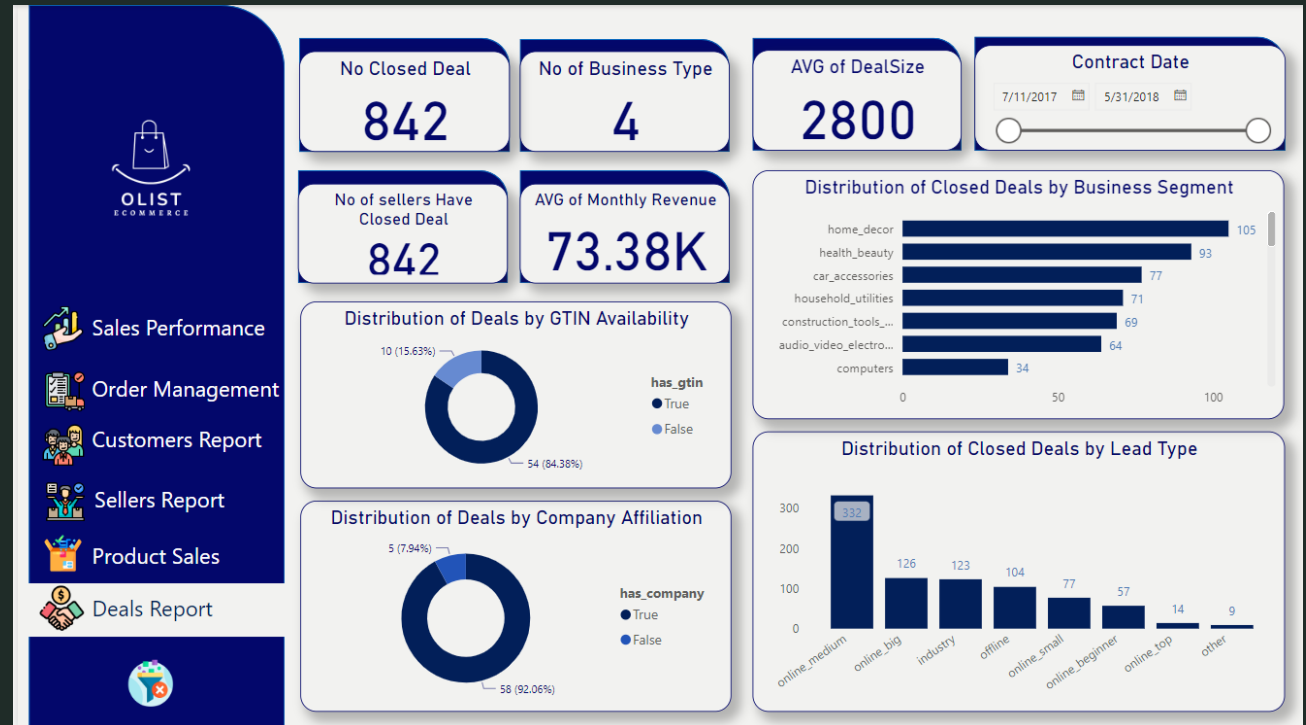
# Product Sales

- A product sales report dashboard is crucial for tracking the performance of products or services offered by The business

# Deals Report

- A deals report dashboard is essential for tracking and managing sales opportunities or deals in the pipeline

# Summery

the Olist dataset provides a valuable resource for studying various aspects of e-commerce operations and consumer behavior in the Brazilian market, making it a valuable asset for researchers, analysts, and businesses in the e-commerce industry.
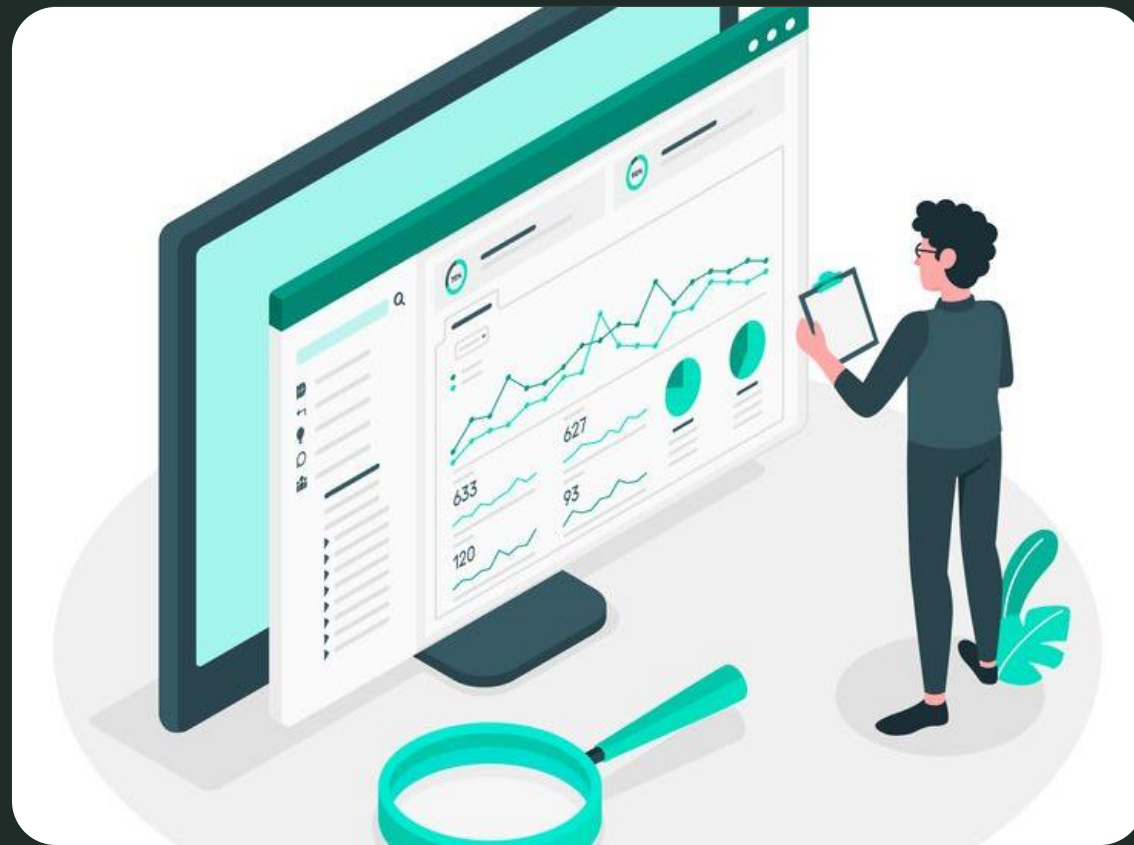
# Thank you

Marwa Ashraf

marwaashraf5814@gmail.com

https://github.com/Marwashraf

www.linkedin.com/in/marwashraf