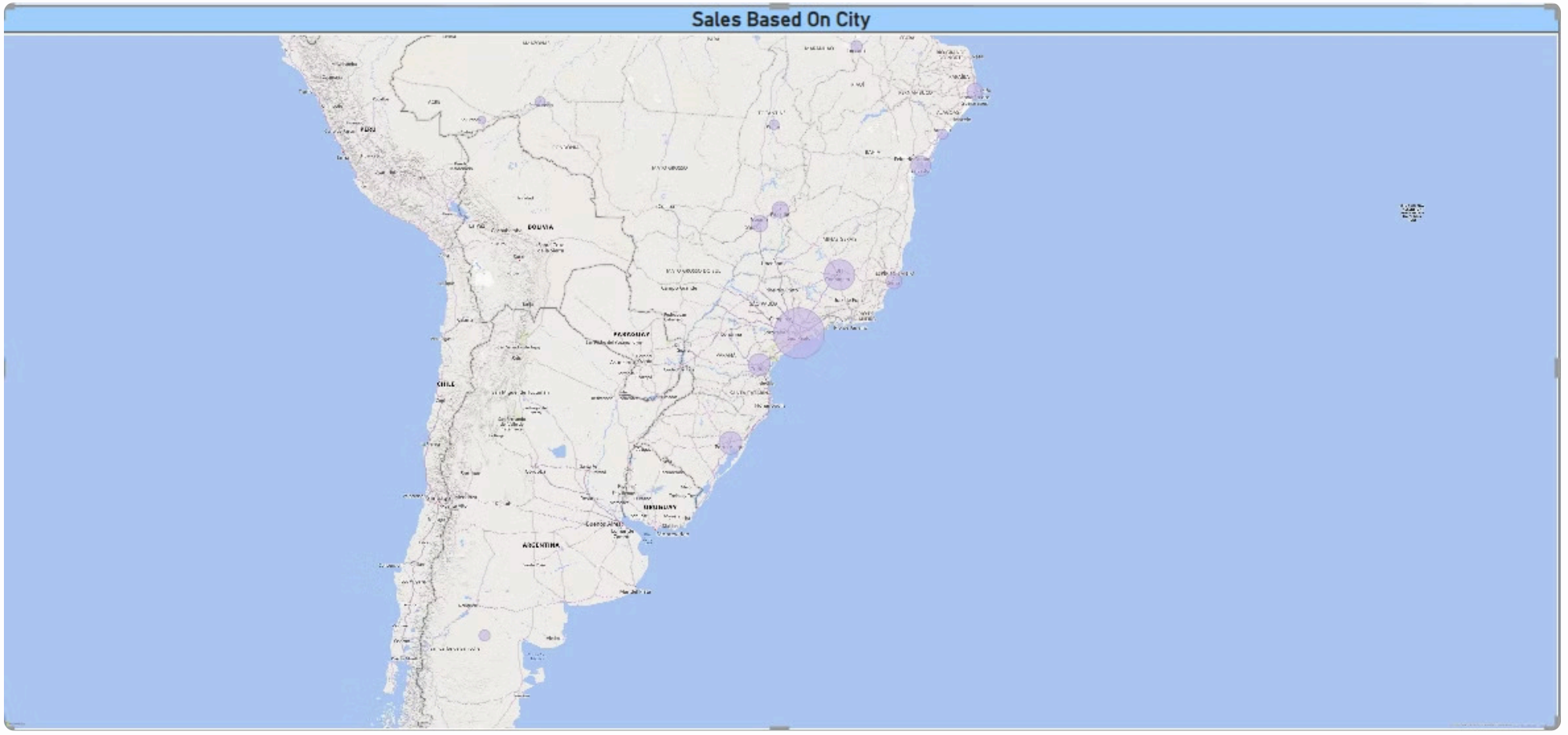
Data Analytics Mastery: Target Sales Company Capstone Project

Welcome to this challenging capstone project designed to elevate your data analysis skills. You'll dive deep into Target Sales Company's data, using SQL, Power BI/Tableau, and Python to uncover valuable insight

E-Commerce Sales and Customer Insights Analysis: SQL Deep Dive

1. Sales Performance Analysis: Write a query to calculate total sales revenue per category, sub-category, and region. Identify the top 5 best-selling products by both sales revenue and quantity sold.
2. Customer Insights: Find the most loyal customers by calculating their purchase frequency and total spend. Identify customers with the highest average order value (AOV).
3. Operational Efficiency: Analyze delivery performance by calculating the average delivery time by region. Identify regions or products with the highest return rates.
4. Date and Time Analytics: Write a query to find the monthly sales trend for the last two years. Analyze the seasonality of sales to identify peak months.
5. Advanced SQL Queries: Use window functions to rank products based on their sales within each category. Calculate month-to-date (MTD) and year-to-date (YTD) sales metrics.

Cities Where Our Company Operates



Company Product Categories

product category

- Agro Industria e Comercio
- Art
- Arts and Crafts
- audio
- automotive
- babies
- Bags Accessories
- bed table bath
- Blu Ray DVDs
- Casa Construcao
- cds music dvds
- Christmas articles
- cine photo
- CITTE AND UPHACK FURNITURE
- climatization
- computer accessories
- CONSTRUCTION SECURITY TOOLS
- Construction Tools Construction
- Construction Tools Garden
- Construction Tools Illumination
- Construction Tools Tools
- Cool Stuff
- Drink foods
- drinks
- ELECTRICES 2
- electronics
- electrostile
- Fashion Bags and Accessories
- Fashion Calcados

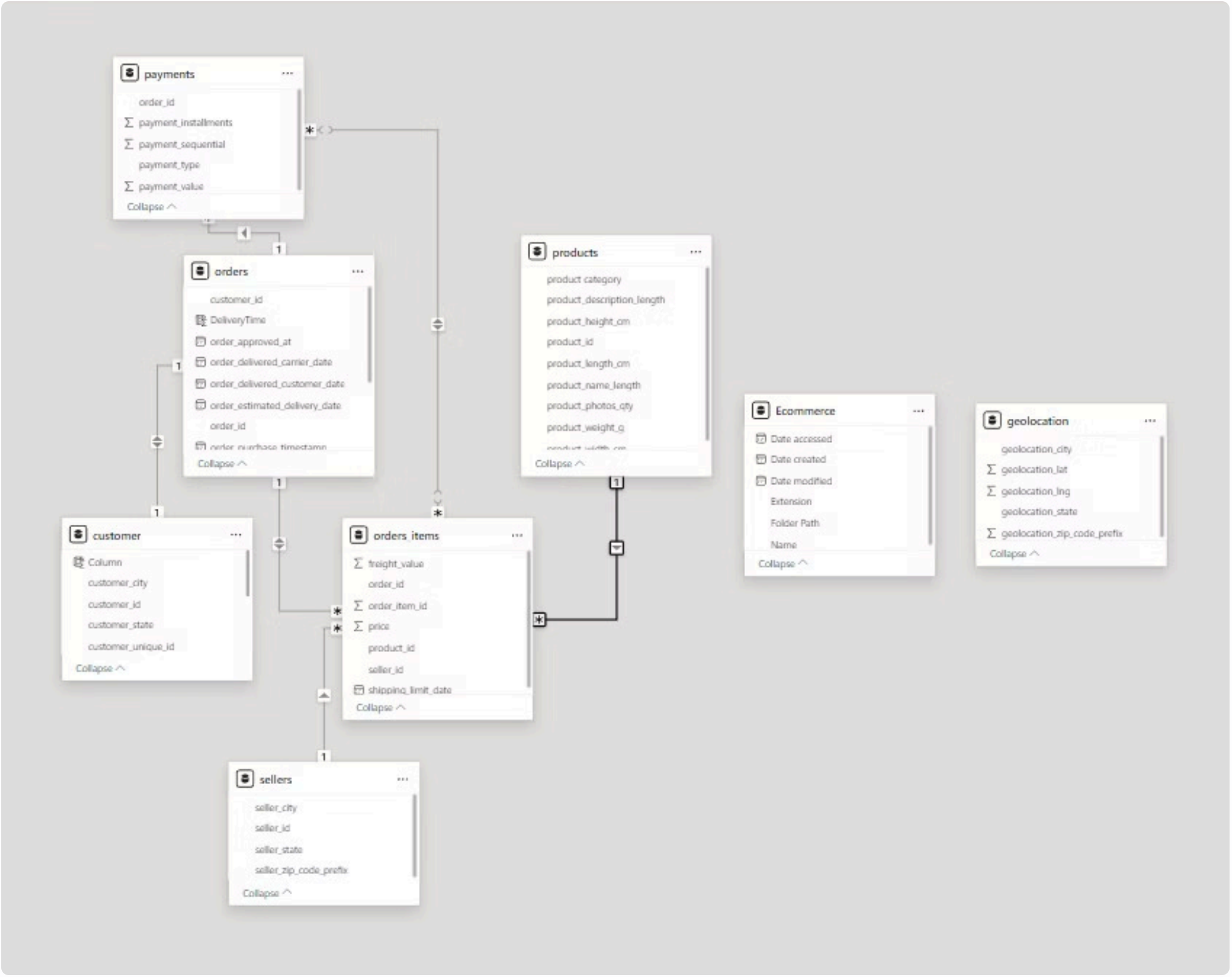
product category

- General Interest Books
- HEALTH BEAUTY
- home appliances
- House comfort
- House Comfort 2
- HOUSE PASTALS OVEN AND CAFE
- housewares
- Hygiene diapers
- IMAGE IMPORT TABLETS
- Imported books
- Industry Commerce and Business
- insurance and services
- Kitchen portable and food coach
- La Cuisine
- Market Place
- musical instruments
- party articles
- PC Gamer
- PCs
- perfumery
- pet Shop
- Room Furniture
- SIGNALIZATION AND SAFETY
- song
- sport leisure
- stationary store
- technical books
- telephony
- toys
- Watches present

product category

- Cool Stuff
- Drink foods
- drinks
- ELECTRICES 2
- electronics
- electrostile
- Fashion Bags and Accessories
- Fashion Calcados
- Fashion Children's Clothing
- Fashion Men's Clothing
- Fashion Sport
- Fashion Underwear and Beach Fashion
- Fashion Women's Clothing
- fixed telephony
- flowers
- foods
- Furniture
- Furniture Decoration
- Furniture Kitchen Service Area Dinner and Garden
- Furniture office
- Games consoles
- Garden tools
- General Interest Books
- HEALTH BEAUTY
- home appliances
- House comfort
- House Comfort 2
- HOUSE PASTALS OVEN AND CAFE
- housewares
- Hygiene diapers

Company Data Model



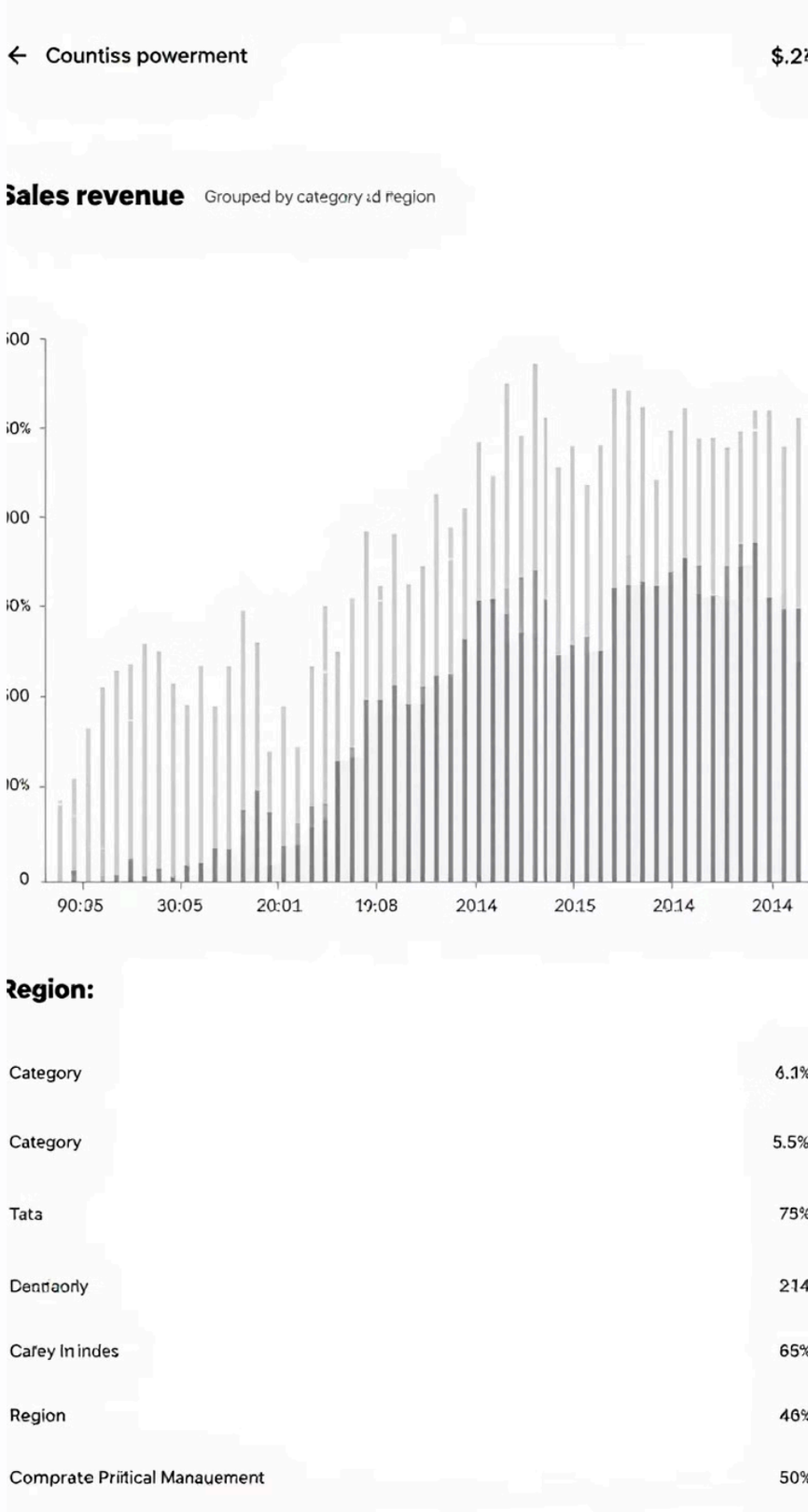
Query: Total Sales Revenue by Category, Sub-Category, and Region

QUERY 1

```sql
SELECT
    p.`product category` AS category,
    s.seller_state AS region,
    SUM(oi.price) AS total_sales_revenue
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
JOIN sellers s ON oi.seller_id = s.seller_id
GROUP BY p.`product category`, s.seller_state;
```

OUTPUT

category	region	total_sales_revenue
sport leisure	SP	610096.6100000154
Cool Stuff	SP	445749.44999999437
Furniture Decoration	SP	498629.4100000255
HEALTH BEAUTY	SP	697858.5000000027
housewares	PR	35795.02000000007
Fashion Underwear and Beach Fashion	SP	9421.949999999986
Watches present	SP	971086.6000000155
Furniture Decoration	PR	134750.9399999983
bed table bath	SP	909463.04000005
perfumery	SP	227363.6099999961



Top 5 Best-Selling Products: Revenue vs. Quantity

By Revenue

QUERY 2

```
SELECT
    p.`product category` AS category,
    p.product_id,
    SUM(oi.price) AS total_sales_revenue,
    SUM(oi.order_item_id) AS total_quantity_sold
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
GROUP BY p.product_id, p.`product category`
ORDER BY total_sales_revenue DESC
LIMIT 5;
```

By Quantity

OUTPUT

category	product_id	total_sales_revenue	total_quantity_sold
HEALTH BEAUTY	bb50f2e236e5eea0100680137654686c	63885	215
HEALTH BEAUTY	6cdd53843498f92890544667809f1595	54730.200000000106	164
PCs	d6160fb7873f184099d9bc95e30376af	48899.34	35
computer accessories	d1c427060a0f73f6b889a5c7c61f2ac4	47214.50999999998	369
bed table bath	99a4788cb24856965c36a24e339b6058	43025.56000000037	542

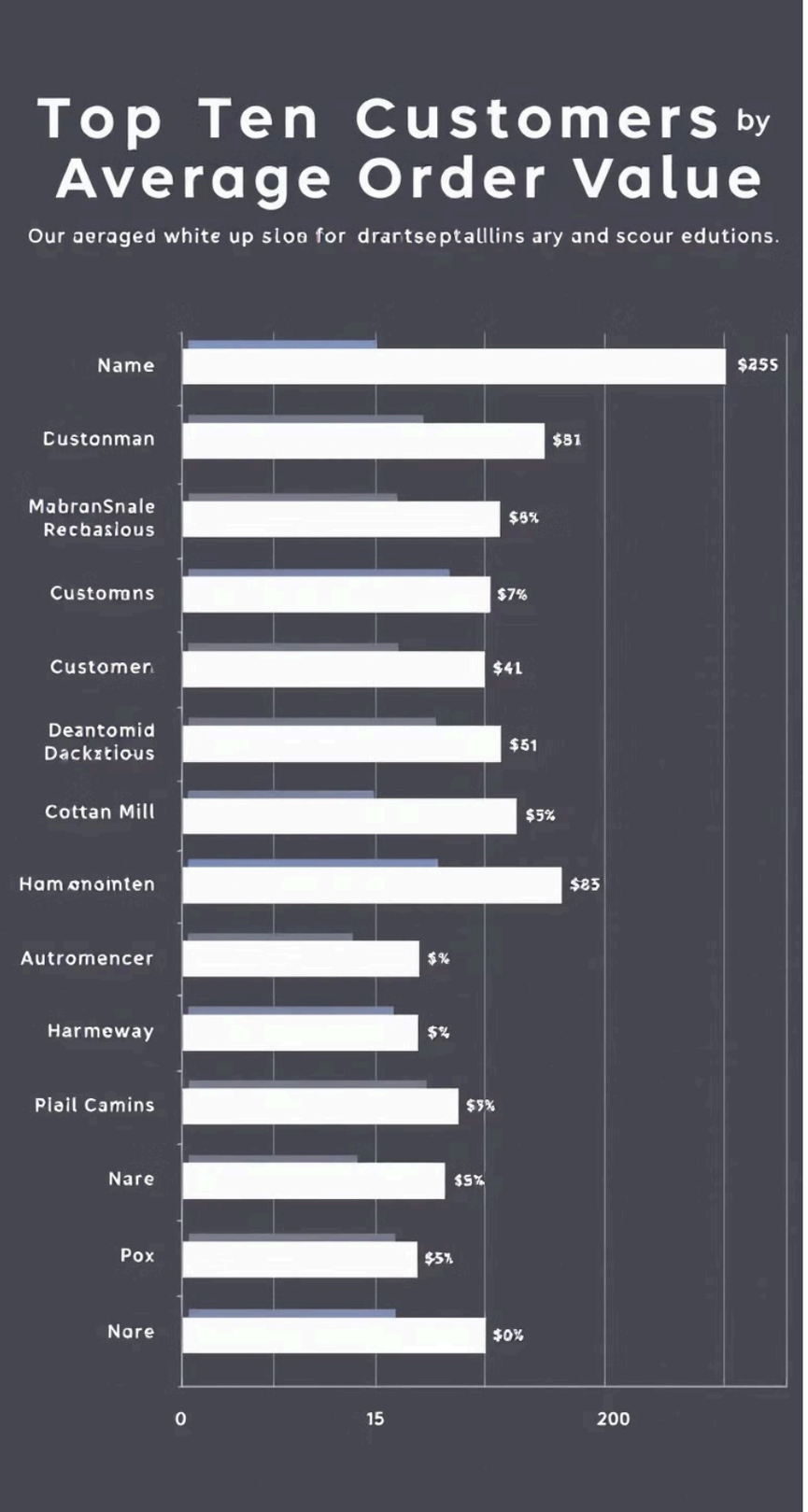
Identifying Most Loyal Customers

QUERY 3

```sql
SELECT
    c.customer_unique_id AS customer,
    COUNT(o.order_id) AS purchase_frequency,
    SUM(oi.price) AS total_spend
FROM customers c
JOIN orders o ON c.customer_id = o.customer_id
JOIN order_items oi ON o.order_id = oi.order_id
GROUP BY c.customer_unique_id
ORDER BY total_spend DESC, purchase_frequency DESC;
```

OUTPUT

customer	purchase_frequency	total_spend
0a0a92112bd4c708ca5fde585afaa872	8	13440
da122df9eeddfedc1dc1f5349a1a690c	2	7388
763c8b1c9c68a0229c42c9fc6f662b93	4	7160
dc4802a71eae9be1dd28f5d788ceb526	1	6735
459bef486812aa25204be022145caa62	1	6729
ff4159b92c40ebe40454e3e6a7c35ed6	1	6499
4007669dec559734d6f53e029e360987	6	5934.6
eebb5dda148d3893cdaf5b5ca3040ccb	1	4690
5d0a2980b292d049061542014e8960bf	2	4599.9
48e1ac109decbb87765a3eade6854098	1	4590
a229eba70ec1c2abef51f04987deb7a5	2	4400



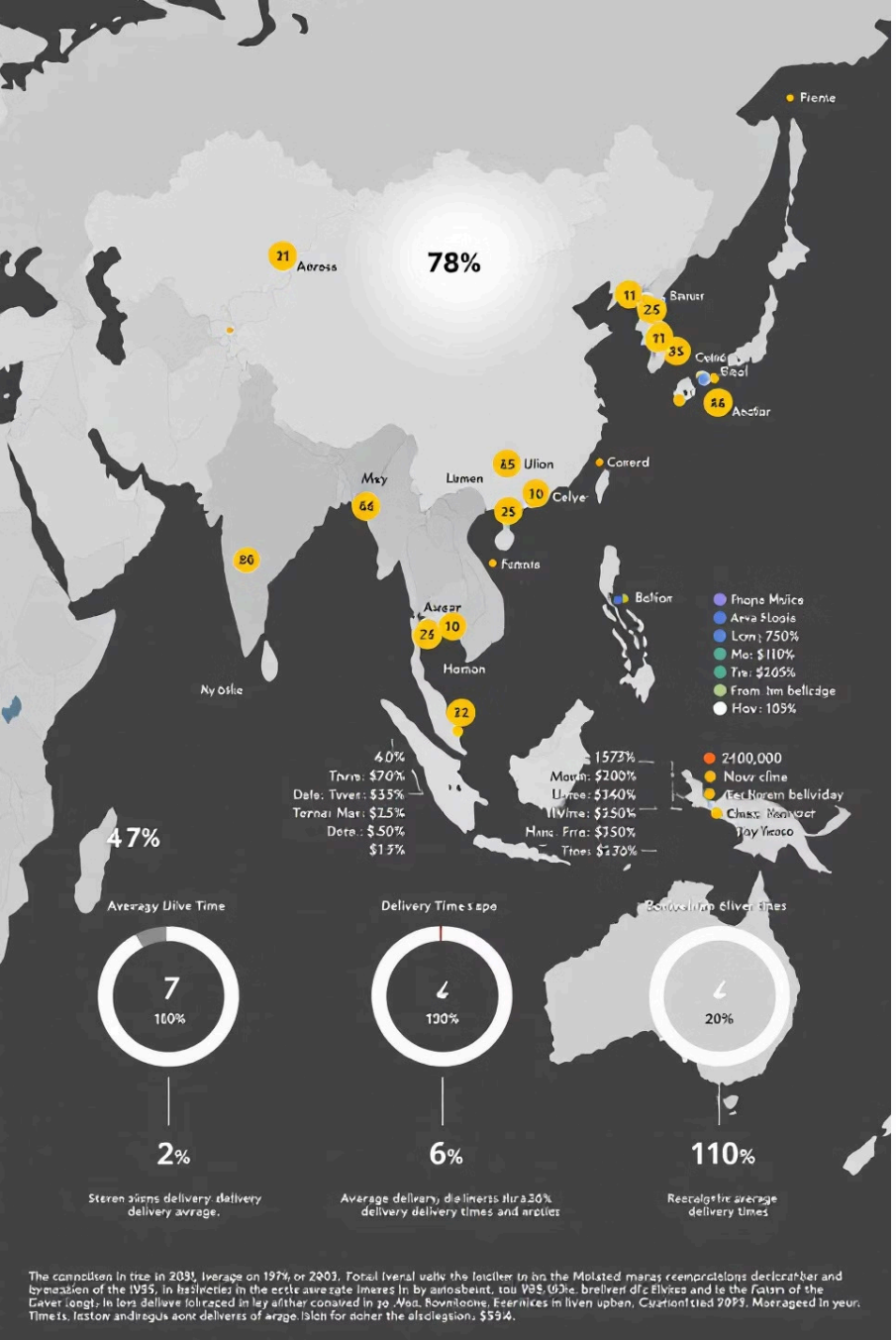
Customers with Highest Average Order Value (AOV)

QUERY 4

```sql
SELECT
    c.customer_unique_id AS customer,
    AVG(p.payment_value) AS average_order_value
FROM customers c
JOIN orders o ON c.customer_id = o.customer_id
JOIN payments p ON o.order_id = p.order_id
GROUP BY c.customer_unique_id
ORDER BY average_order_value DESC
LIMIT 10;
```

OUTPUT

customer	average_order_value
0a0a92112bd4c708ca5fde585afaa872	13664.08
763c8b1c9c68a0229c42c9fc6f662b93	7274.88
dc4802a71eae9be1dd28f5d788ceb526	6929.31
459bef486812aa25204be022145caa62	6922.21
ff4159b92c40ebe40454e3e6a7c35ed6	6726.66
4007669dec559734d6f53e029e360987	6081.54
5d0a2980b292d049061542014e8960bf	4809.44
eebb5dda148d3893cdaf5b5ca3040ccb	4764.34
48e1ac109decbb87765a3eade6854098	4681.78
edde2314c6c30e864a128ac95d6b2112	4513.32



Analyzing Delivery Performance by Region

```sql
SELECT
    s.seller_state AS region,
    AVG(TIMESTAMPDIFF(DAY, o.order_delivered_carrier_date, o.order_delivered_customer_date)) AS avg_delivery_time
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
JOIN sellers s ON oi.seller_id = s.seller_id
WHERE o.order_delivered_customer_date IS NOT NULL
GROUP BY s.seller_state;
```

region	avg_delivery_time
SP	8.5263
MG	9.0885
DF	9.0430
SC	10.0160
RJ	8.6944
PR	9.4088
RS	7.6187
RN	8.1071
GO	9.7323
BA	9.9359
MT	11.0000

Identifying High Return Rate Regions and Products

1

```sql
SELECT
    p.`product category` AS category,
    s.seller_state AS region,
    COUNT(CASE WHEN o.order_status = 'returned' THEN 1 END) AS return_count,
    COUNT(o.order_id) AS total_orders,
    (COUNT(CASE WHEN o.order_status = 'returned' THEN 1 END) / COUNT(o.order_id)) * 100 AS return_rate
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
JOIN products p ON oi.product_id = p.product_id
JOIN sellers s ON oi.seller_id = s.seller_id
GROUP BY p.`product category`, s.seller_state
ORDER BY return_rate DESC;
```

2

OUTPUT

category	region	return_count	total_orders	return_rate
housewares	SP	0	4952	0.0000
HEALTH BEAUTY	SP	0	6404	0.0000
babies	SP	0	2164	0.0000
Watches present	SP	0	4830	0.0000
Furniture Decoration	SP	0	6256	0.0000
perfumery	SP	0	2287	0.0000
bed table bath	SP	0	9818	0.0000
musical instruments	SP	0	411	0.0000
pet Shop	MG	0	410	0.0000
computer accessories	MG	0	1575	0.0000
flowers	SP	0	32	0.0000

Analyzing Sales Seasonality

1

QUERY

```sql
SELECT
    MONTH(o.order_purchase_timestamp) AS month,
    SUM(oi.price) AS total_sales
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
GROUP BY month
ORDER BY total_sales DESC;
```

2

OUTPUT

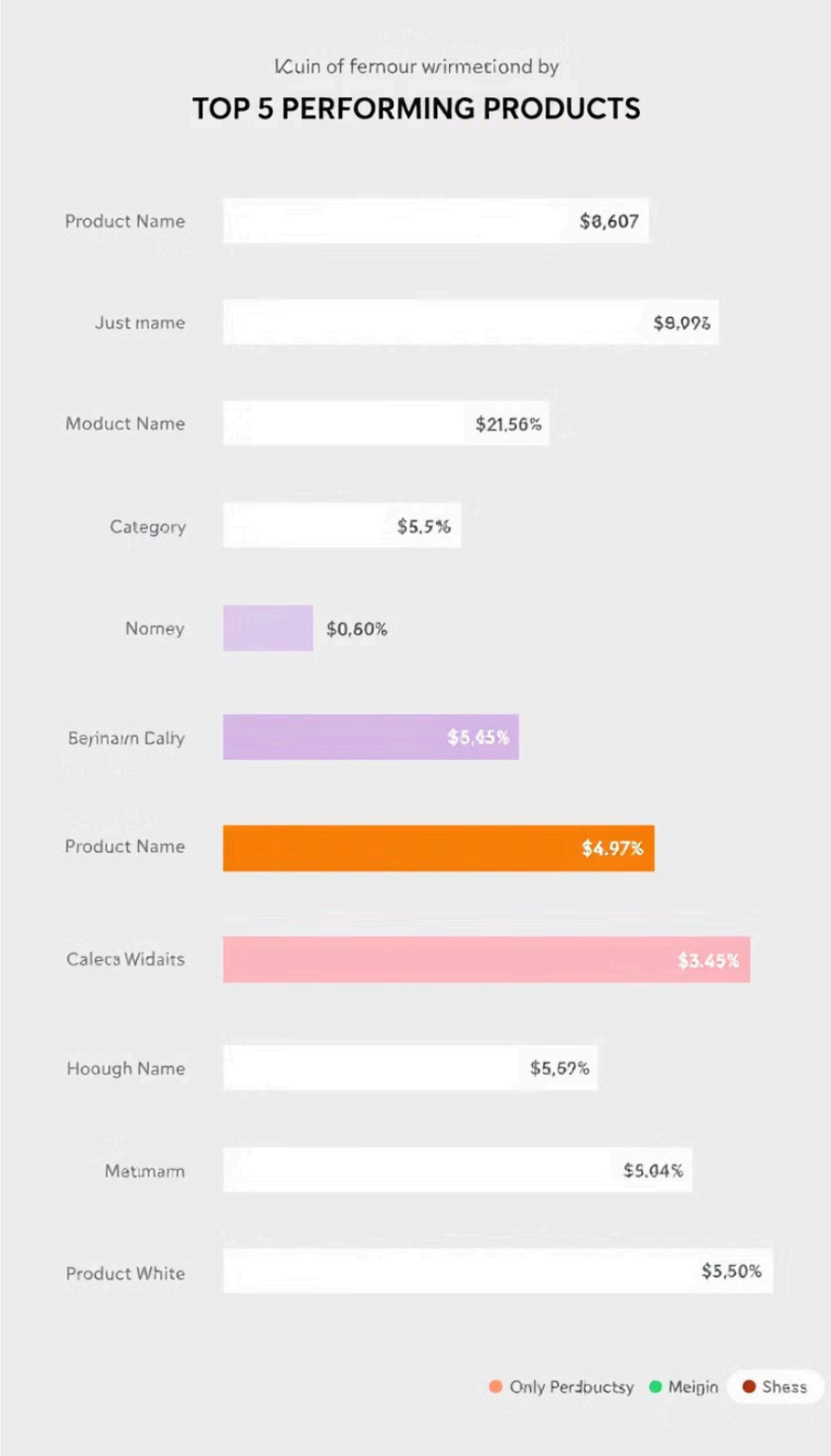
month	total_sales
5	1502588.819999901
8	1428658.0099999176
7	1393538.699999928
3	1357557.7399999497
4	1356574.979999948
6	1298162.9099999587
2	1091481.7300000244
1	1070343.2300000393
11	1010271.3700000389
12	743925.0700000194
10	713727.0900000224

Ranking Products within Categories

```sql
SELECT
    p.`product category` AS category,
    p.product_id,
    SUM(oi.price) AS total_sales_revenue,
    RANK() OVER (PARTITION BY p.`product category` ORDER BY SUM(oi.price) DESC) AS rank_within_category
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
GROUP BY p.`product category`, p.product_id;
```

OUTPUT

product_id	total_sales_revenue	rank_within_category
5a848e4ab52fd5445cdc07aab1c40e48	24229.030000000075	1
eed5cbd74fac3bd79b7c7ec95fa7507d	9945	2
b1d207586fca400a2370d50a9ba1da98	7152	3
76d1a1a9d21ab677a61c3ae34b1b352f	5712.640000000003	4
ad88641611c35ebd59ecda07a9f17099	4515.330000000001	5
3b60d513e90300a4e9833e5cda1f1d61	4393.330000000001	6
4c50dcc50f1512f46096d6ef0142c4a9	3980	7
17823ffd2de8234f0e885a71109613a4	2969.8899999999994	8
0e030462875259ec0cb868f7ecf1fd5e	2740	9
b36f3c918c91478c4559160022d3f14e	2550	10
f58e45b16a42a325c144eb2c46a2bc57	2359.91	11



Calculating MTD and YTD Sales Metrics

QUERY

```
    daily_sales,
    SUM(daily_sales) OVER (
        PARTITION BY YEAR(order_date), MONTH(order_date)
    ) AS mtd_sales,
    SUM(daily_sales) OVER (
        PARTITION BY YEAR(order_date)
    ) AS ytd_sales
FROM (
    SELECT
        DATE(o.order_purchase_timestamp) AS order_date,
        SUM(oi.price) AS daily_sales
    FROM orders o
    JOIN order_items oi ON o.order_id = oi.order_id
    GROUP BY DATE(o.order_purchase_timestamp)
) daily_data;
```

OUTPUT

date	daily_sales	mtd_sales	ytd_sales
2016-09-15	134.97	267.36	49785.91999999998
2016-09-04	72.89	267.36	49785.91999999998
2016-09-05	59.5	267.36	49785.91999999998
2016-10-07	7228.049999999997	49507.659999999974	49785.91999999998
2016-10-04	9940.959999999986	49507.659999999974	49785.91999999998
2016-10-05	8343.249999999996	49507.659999999974	49785.91999999998
2016-10-09	3336.99	49507.659999999974	49785.91999999998
2016-10-06	7960.509999999994	49507.659999999974	49785.91999999998
2016-10-03	463.48	49507.659999999974	49785.91999999998
2016-10-10	3692.5699999999997	49507.659999999974	49785.91999999998
2016-10-08	8441.849999999997	49507.659999999974	49785.91999999998

E-commerce Sales and Customer Insights Analysis - Power BI

Tasks and Key Metrics to Visualize:

1. Sales Performance: Total Sales Revenue: Overall revenue generated. Average Order Value (AOV): Average revenue per order. Sales by Category and Sub-Category: Revenue breakdown by product categories. Top-Selling Products: Display the top 5 products by sales revenue.
2. Customer Insights: Customer Lifetime Value (CLV): Total revenue generated per customer. Top 10 Loyal Customers: Customers with the highest purchase frequency and spend. Customer Segments: Categorize customers based on purchasing behavior (e.g., high spenders, one-time buyers).
3. Regional Analysis: Revenue by Region: Comparison of sales by region. Return Rates by Region: Percentage of orders returned per region. Average Delivery Time by Region: Assess operational performance.
4. Operational Metrics: Delivery Time Analysis: Average, minimum, and maximum delivery times. Product Return Rates: Percentage of returned products across categories.
5. Time Trends: Monthly Sales Trends: Visualization of revenue trends over time. Seasonality Analysis: Highlight peak sales periods (e.g., festive months).