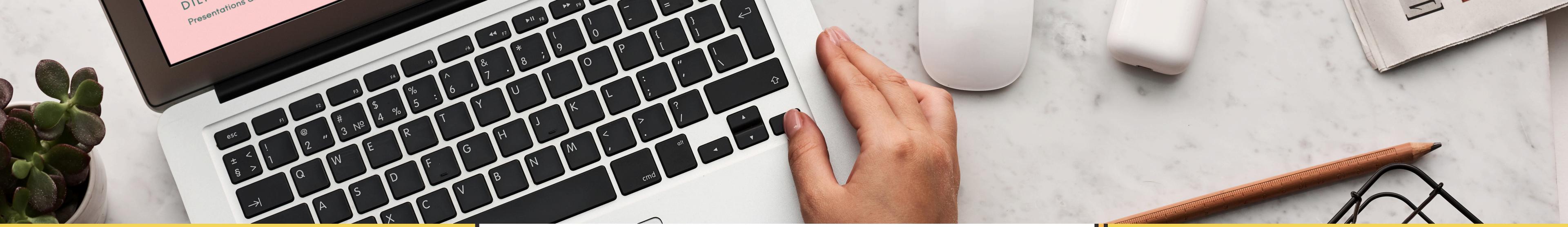




# **DATA-DRIVEN GROWTH OPTIMIZATION FOR MERCADO LIVRE E-COMMERCE PLATFORM**

by Ayan Ghosh



# BACKGROUND

Background of the Project:

Mercado Livre is a leading e-commerce platform in Latin America, connecting buyers, sellers, and logistics partners. To stay competitive, the platform aims to optimize its operations by gaining insights into customer behavior, seller performance, product trends, and fulfillment efficiency.

By analyzing data from various sources, Mercado Livre seeks to identify growth opportunities, improve customer experiences, and streamline supply chain operations. The goal is to leverage data analytics for informed decision-making, enhance inventory management, and forecast future sales trends to drive business growth and improve overall marketplace efficiency.



# PROJECT GOALS

---

Goal 1

SQL Analysis: Extracting and analyzing key performance metrics to identify factors affecting delays and operational inefficiencies.

Goal 2

Interactive Dashboard: Creating a dynamic dashboard in Power BI/Tableau to visualize performance insights.

Goal 3

Predictive Modeling: Building a machine learning model to predict flight delays and assess its performance.

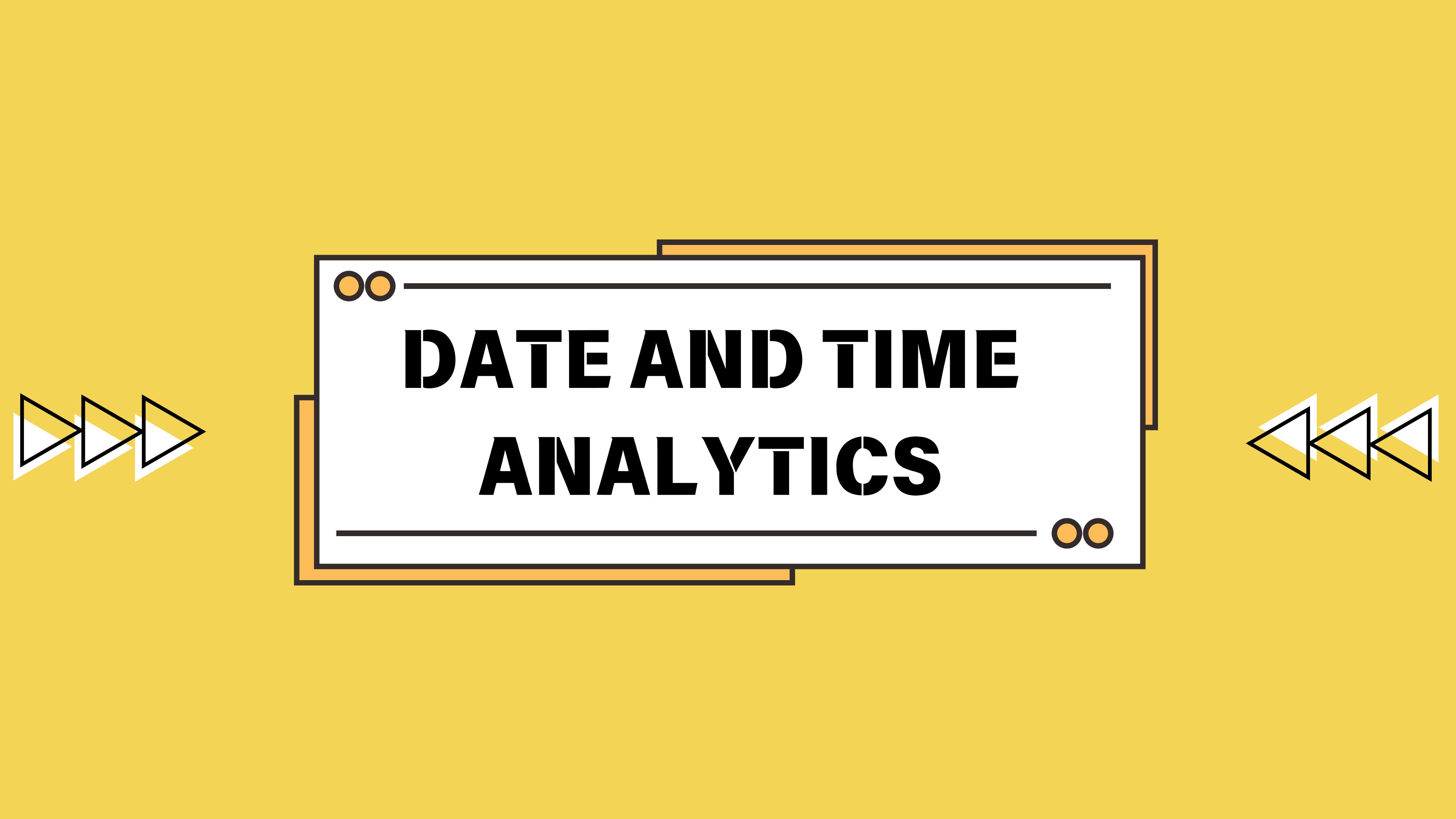
Goal 4

Actionable Insights: Providing recommendations to reduce delays, optimize operations, and enhance customer experience.

⋮  
⋮  
⋮  
⋮  
⋮



# **EXPLORING PATTERNS AND TRENDS WITH HIGH-LEVEL SQL TECHNIQUES**



# DATE AND TIME ANALYTICS

# 1. Monthly Revenue Growth Rate: For each month, calculate total revenue and month-over-month growth percentage.

```
1 •  SELECT
2      DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m') AS month,
3      ROUND(SUM(oi.price), 2) AS total_revenue,
4      ROUND(
5          (SUM(oi.price) - LAG(SUM(oi.price)) OVER (ORDER BY DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m'))))
6          / LAG(SUM(oi.price)) OVER (ORDER BY DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m')) * 100,
7          2
8      ) AS mom_growth_percentage
9  FROM
10     orders_dataset o
11  JOIN
12     order_items_dataset oi ON o.order_id = oi.order_id
13  GROUP BY
14      DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m')
15  ORDER BY
16      month;
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

	month	total_revenue	mom_growth_percentage
▶	2016-09	267.36	HULL
	2016-10	49507.66	18417.23
	2016-12	10.9	-99.98
	2017-01	120312.87	1103687.83
	2017-02	247303.02	105.55
	2017-03	374344.3	51.37
	2017-04	359927.23	-3.85
	2017-05	506071.14	40.6

## 2. Trailing 3-Month Moving Average of Orders: For each month, calculate the 3-month moving average of the number of orders.

```
21 • Ⓛ WITH monthly_orders AS (
22     SELECT
23         DATE_FORMAT(order_purchase_timestamp, '%Y-%m') AS month,
24         COUNT(order_id) AS total_orders
25     FROM
26         orders_dataset
27     GROUP BY
28         DATE_FORMAT(order_purchase_timestamp, '%Y-%m')
29 )
30     SELECT
31         month,
32         total_orders,
33         ROUND(
34             Ⓛ AVG(total_orders) OVER (
35                 ORDER BY month
36                 ROWS BETWEEN 2 PRECEDING AND CURRENT ROW
37             ), 2
38         ) AS trailing_3_month_avg
39     FROM
40         monthly_orders
41     ORDER BY
42         month;
```

Result Grid | Filter Rows:  | Export: | Wrap Cell Content:

	month	total_orders	trailing_3_month_avg
▶	2016-09	4	4.00
	2016-10	324	164.00
	2016-12	1	109.67
	2017-01	800	375.00
	2017-02	1780	860.33
	2017-03	2682	1754.00

### 3. Yearly Average Order Value (AOV): Trend

Compute average order value (total revenue / total orders) per year.

```
46 • SELECT
47     YEAR(o.order_purchase_timestamp) AS year,
48     ROUND(SUM(oi.price) / COUNT(DISTINCT o.order_id), 2) AS average_order_value
49 FROM
50     orders_dataset o
51 JOIN
52     order_items_dataset oi ON o.order_id = oi.order_id
53 GROUP BY
54     YEAR(o.order_purchase_timestamp)
55 ORDER BY
56     year;
57
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

year	average_order_value
2016	159.57
2017	138.09
2018	137.35

7.Time to First Purchase: For customers with multiple orders, calculate the average number of days between account creation and first purchase.

```
145 • WITH customer_orders AS (
146     SELECT
147         c.customer_unique_id,
148         MIN(o.order_purchase_timestamp) AS first_purchase_date,
149         COUNT(DISTINCT o.order_id) AS total_orders,
150         MIN(o.order_purchase_timestamp) AS account_creation_date
151     FROM
152         customers_dataset c
153     JOIN
154         orders_dataset o ON c.customer_id = o.customer_id
155     GROUP BY
156         c.customer_unique_id
157 ),
```

```
159 • repeat_customers AS (
160     SELECT
161         customer_unique_id,
162         DATEDIFF(first_purchase_date, account_creation_date) AS days_to_first_purchase
163     FROM
164         customer_orders
165     WHERE
166         total_orders > 1
167 )
168
169     SELECT
170         ROUND(AVG(days_to_first_purchase), 2) AS avg_days_to_first_purchase
171     FROM
172         repeat_customers;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:
avg_days_to_first_purchase ▶ 0.00			

# CUSTOMER INSIGHTS

#### 4.Customer Lifetime Value (CLTV) Approximation: Calculate total revenue generated by each customer. Sort top 10 highest lifetime values.

```
58 • SELECT
59     c.customer_unique_id,
60     ROUND(SUM(oi.price), 2) AS total_revenue
61 FROM
62     customers_dataset c
63 JOIN
64     orders_dataset o ON c.customer_id = o.customer_id
65 JOIN
66     order_items_dataset oi ON o.order_id = oi.order_id
67 GROUP BY
68     c.customer_unique_id
69 ORDER BY
70     total_revenue DESC
71 LIMIT 10;
72
```

Result Grid | Filter Rows: \_\_\_\_\_ | Export: | Wrap Cell Content: | Fetch rows:

	customer_unique_id	total_revenue
▶	0a0a92112bd4c708ca5fde585afaa872	13440
	da122df9eeddfedc1dc1f5349a1a690c	7388
	763c8b1c9c68a0229c42c9fc6f662b93	7160
	dc4802a71eae9be1dd28f5d788ceb526	6735
	459bef486812aa25204be022145caa62	6729
	ff4159h92r40phe40454e3e6a7c35ed6	6499

## 8. Review Sentiment vs Delivery Time: Compare average delivery days for orders with 1-2 stars vs 4-5 stars.

```
177 • SELECT
178   CASE
179     WHEN r.review_score BETWEEN 1 AND 2 THEN '1-2 Stars'
180     WHEN r.review_score BETWEEN 4 AND 5 THEN '4-5 Stars'
181   END AS review_group,
182   ROUND(AVG(DATEDIFF(o.order_delivered_customer_date, o.order_estimated_delivery_date)), 2) AS avg_delivery_days
183 FROM
184   order_reviews_dataset r
185 JOIN
186   orders_dataset o ON r.order_id = o.order_id
187 WHERE
188   r.review_score IN (1, 2, 3, 4, 5)
189   AND o.order_delivered_customer_date IS NOT NULL
190   AND o.order_estimated_delivery_date IS NOT NULL
191 GROUP BY
192   review_group;
```

Result Grid | Filter Rows: \_\_\_\_\_ | Export: | Wrap Cell Content:

review_group	avg_delivery_days
4-5 Stars	-13.14
1-2 Stars	-5.15
NULL	-10.77

# 10.State-Wise Revenue Per Customer: For each state, compute average revenue per customer.

```
2  SELECT
3      c.customer_state,
4      COUNT(DISTINCT c.customer_unique_id) AS unique_customers,
5      ROUND(SUM(p.payment_value), 2) AS total_payment,
6      ROUND(SUM(p.payment_value) / COUNT(DISTINCT c.customer_unique_id), 2) AS avg_payment_per_customer
7  FROM
8      customers_dataset c
9  JOIN
10     orders_dataset o ON c.customer_id = o.customer_id
11    JOIN
12        order_payments_dataset p ON o.order_id = p.order_id
13  GROUP BY
14      c.customer_state
15  ORDER BY
16      avg_payment_per_customer DESC;
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

	customer_state	unique_customers	total_payment	avg_payment_per_customer
▶	PB	519	141545.72	272.73
	AC	77	19680.62	255.59
	RO	240	60866.2	253.61
	AP	67	16262.8	242.73
	AL	401	96962.06	241.8
	CA	210	210000.0	1000.0



# SALES PERFORMANCE ANALYSIS

# 5.Top Product Categories by Growth: Identify top 5 categories with the highest year-over-year revenue growth.

```
77 • WITH category_revenue_by_year AS (
78     SELECT
79         YEAR(o.order_purchase_timestamp) AS year,
80         pct.product_category_name_english AS category,
81         ROUND(SUM(oi.price), 2) AS total_revenue
82     FROM
83         order_items_dataset oi
84     JOIN
85         orders_dataset o ON oi.order_id = o.order_id
86     JOIN
87         products_dataset p ON oi.product_id = p.product_id
88     LEFT JOIN
89         product_category_name_translation pct ON p.product_category_name = pct.product_category_name
90     GROUP BY
91         YEAR(o.order_purchase_timestamp),
92         pct.product_category_name_english
93 ),
94
95     category_growth AS (
96         SELECT
97             curr.category,
98             curr.year,
99             curr.total_revenue,
100            prev.total_revenue AS prev_revenue,
101            ROUND(
102                (curr.total_revenue - prev.total_revenue) / NULLIF(prev.total_revenue, 0) * 100, 2
103            ) AS yoy_growth_percentage
104         FROM
105             category_revenue_by_year curr
106         LEFT JOIN
107             category_revenue_by_year prev
108             ON curr.category = prev.category AND curr.year = prev.year + 1
109     )
110
111     SELECT
112         category,
113         year,
114         yoy_growth_percentage
115     FROM
116         category_growth
117     WHERE
118         yoy_growth_percentage IS NOT NULL
119     ORDER BY
120         yoy_growth_percentage DESC
121     LIMIT 5;
```

111     SELECT  
112         category,  
113         year,  
114         yoy\_growth\_percentage  
115     FROM  
116         category\_growth  
117     WHERE  
118         yoy\_growth\_percentage IS NOT NULL  
119     ORDER BY  
120         yoy\_growth\_percentage DESC  
121     LIMIT 5;  
122  
123  
124

Result Grid | Filter Rows: Export: Wrap Cell Content:

	category	year	yoy_growth_percentage
▶	bed_bath_table	2017	103960.72
	fashion_shoes	2017	50037.08
	cool_stuff	2017	35329.4
	fashion_male_clothing	2017	33920.28
	computers_accessories	2017	28848.25

## 6. Seller Fulfillment Efficiency: For each seller, calculate the average difference between estimated and actual delivery date. Highlight the most efficient ones.

```
126 •   SELECT
127       oi.seller_id,
128       ROUND(AVG(DATEDIFF(o.order_estimated_delivery_date, o.order_delivered_customer_date)), 2)
129       AS avg_delivery_difference_days
130   FROM
131       order_items_dataset oi
132   JOIN
133       orders_dataset o ON oi.order_id = o.order_id
134   WHERE
135       o.order_delivered_customer_date IS NOT NULL
136       AND o.order_estimated_delivery_date IS NOT NULL
137   GROUP BY
138       oi.seller_id
139   ORDER BY
140       avg_delivery_difference_days DESC;
141
142
```

Result Grid | Filter Rows:  Export: Wrap Cell Content: Fetch rows:

seller_id	avg_delivery_difference_days
933446e9a59dece7ae9175103820ca8f	66.00
0b09101900100c0e9d312861fad5a1b9	61.00
fa5fdc4e4bb6bd1009ad0e4ac4096562	58.00
58e4b302b54937e55a678c4d15111da4	48.00
fffff564a4f9085cd26170f4732393726	48.00
432c67055-nacd1fd6b0h5d678766a71	48.00

# 9.Payment Method Impact: Analyze which payment methods lead to the highest average review scores.

```
3 • WITH payment_review_data AS (
4     SELECT
5         order_payments_dataset.payment_type,
6         order_reviews_dataset.review_score
7     FROM
8         order_payments_dataset
9     JOIN
10        order_reviews_dataset
11    ON order_payments_dataset.order_id = order_reviews_dataset.order_id
12 )
13
14     SELECT
15         payment_type,
16         ROUND(AVG(review_score), 2) AS average_review_score,
17         COUNT(*) AS total_reviews
18     FROM
19         payment_review_data
20     GROUP BY
21         payment_type
22     ORDER BY
23         total_reviews desc ;
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

payment_type	average_review_score	total_reviews
credit_card	4.09	76600
boleto	4.09	19762
voucher	4.00	5783
debit_card	4.17	1529
not_defined	1.67	3

# 11.Repeat Purchase Analysis: What percentage of customers placed more than one order?

```
20  SELECT
21      ROUND(100.0 * COUNT(*) / (SELECT COUNT(DISTINCT customer_id) FROM orders_dataset), 2)
22      AS repeat_customer_percentage
23  FROM (
24      SELECT customer_id
25      FROM orders_dataset
26      GROUP BY customer_id
27      HAVING COUNT(order_id) > 1
28  ) AS repeat_customers;
```

29

30

Result Grid | Filter Rows:  Export: Wrap Cell Content:

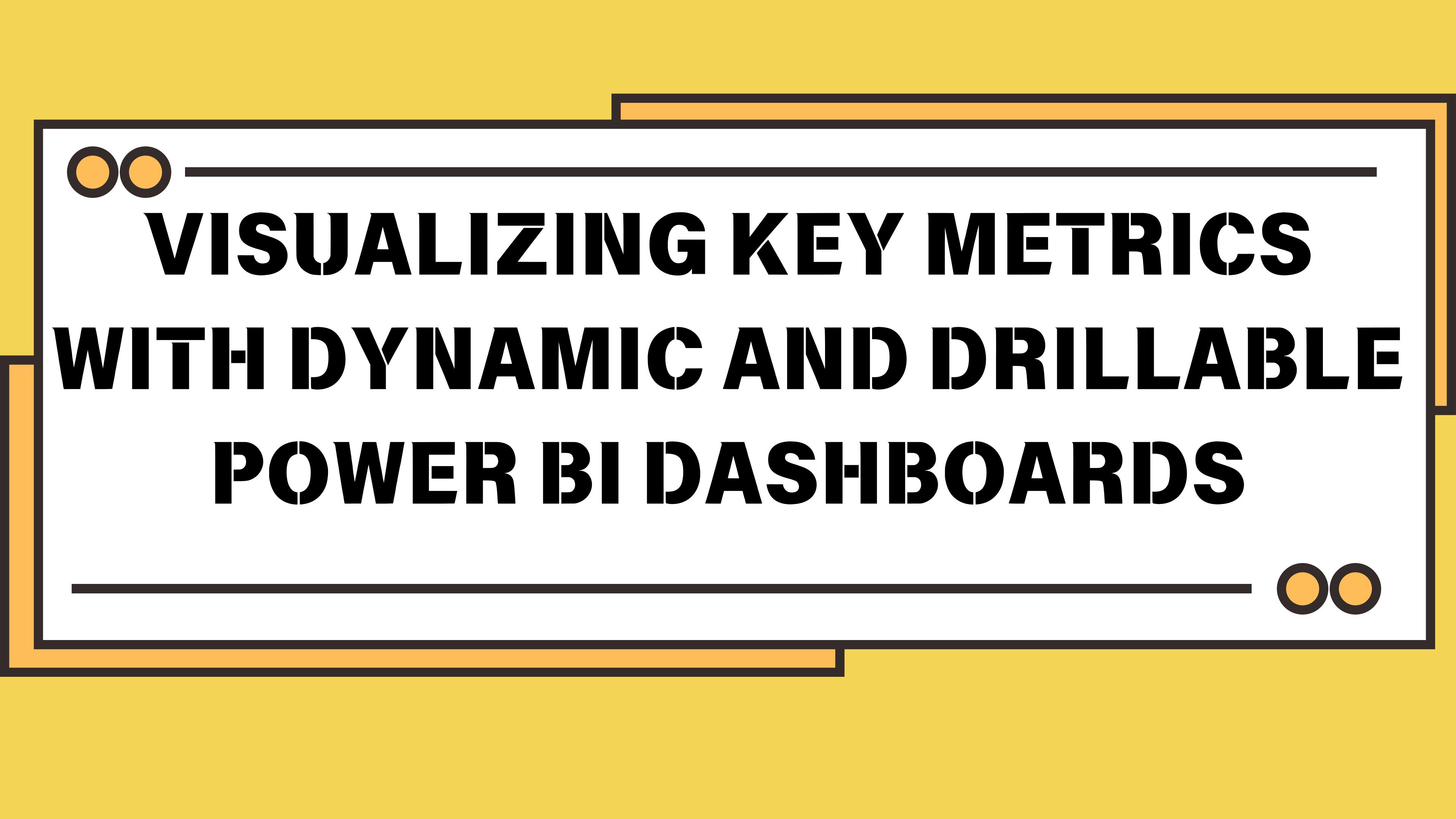
	repeat_customer_percentage
▶	0.00

12.Top SKUs by Profit Contribution: Calculate profit per SKU (price - freight + volume assumption), then list top 10 contributors.

```
3 •   SELECT
4       product_id,
5           ROUND(SUM(price - freight_value), 1) AS total_profit
6   FROM
7       order_items_dataset
8   GROUP BY
9       product_id
10  ORDER BY
11      total_profit DESC
12  LIMIT 10;
```

Result Grid | Filter Rows:  Export: Wrap Cell Content:

	product_id	total_profit
▶	bb50f2e236e5eea0100680137654686c	60163.9
	6cdd53843498f92890544667809f1595	50366.4
	d6160fb7873f184099d9bc95e30376af	47472.5
	25c38557cf793876c5abdd5931f922db	37502.7
	53b36df67ebb7c41585e8d54d6772e08	35408.9
	99a4788cb24856965c36a24e339b6058	34979.5
	3dd2a17168ec895c781a9191c1e95ad7	33953
	5f504b3a1c75b73d6151be81eb05bdc9	33742
	d1c427060a0f73f6b889a5c7c61f2ac4	33453
	aca2eb7d00ea1a7b8ebd4e68314663af	30397



# **VISUALIZING KEY METRICS WITH DYNAMIC AND DRILLABLE POWER BI DASHBOARDS**

# PAGE 1: EXECUTIVE SALES OVERVIEW:-

Total Revenue

R\$15.74M

Total orders

98.2K

Average Order Value

R\$160.24

Active Customer

98.2K

Total Sellers

3095

Repeat Customer

0%



**Customer Insights**

**Product Performance**

**Seller Analytics**

**Fulfillment Metrics**

**Payment & Financial Insights**

Seller\_id

All

Customer\_id

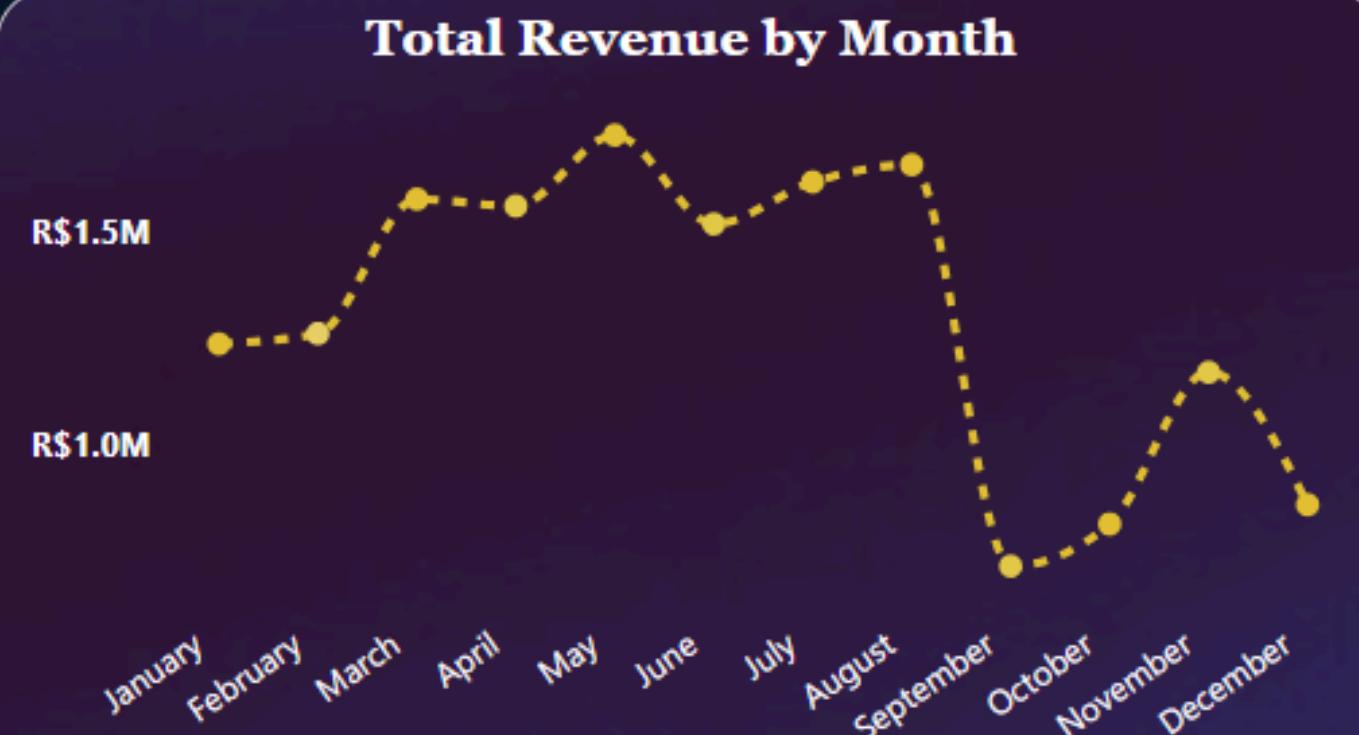
All

Monthna...

(Blank)

April

December

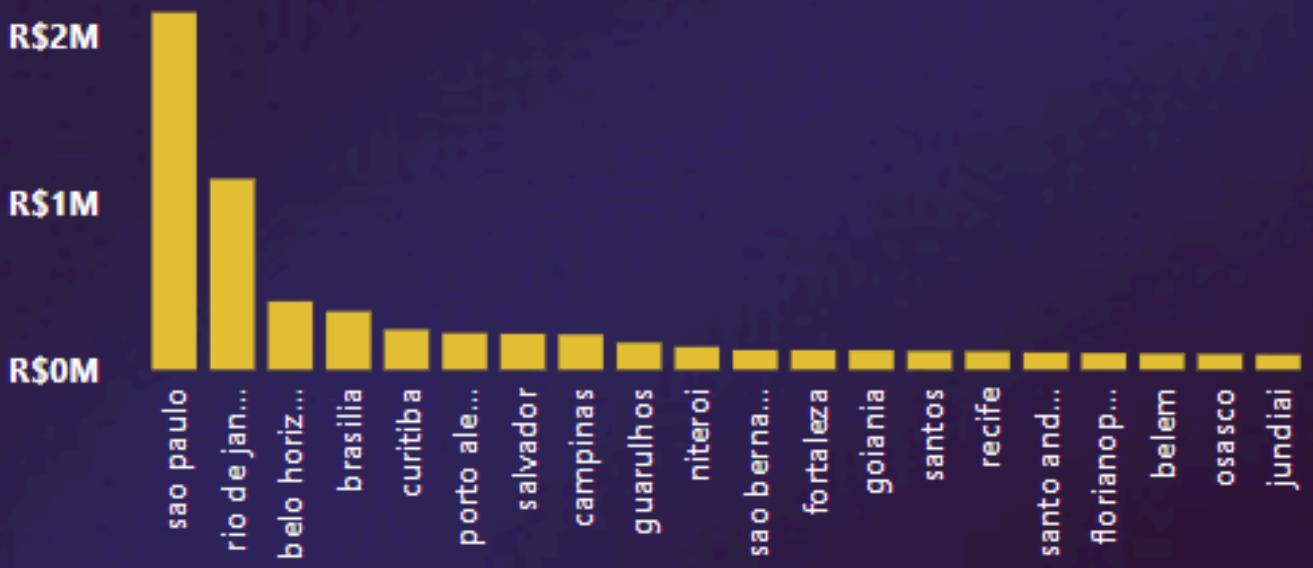


### MOM Growth Rate

Year	Month	MOM%GROTH
2016	September	0.0%
2016	October	18261.2%
2016	November	-100.0%
2016	December	0.0%
2017	January	697878.9%
2017	February	107.1%
2017	March	50.1%
2017	April	-4.6%
2017	May	43.6%
2017	June	-14.3%
2017	July	15.8%
Total		0.0%

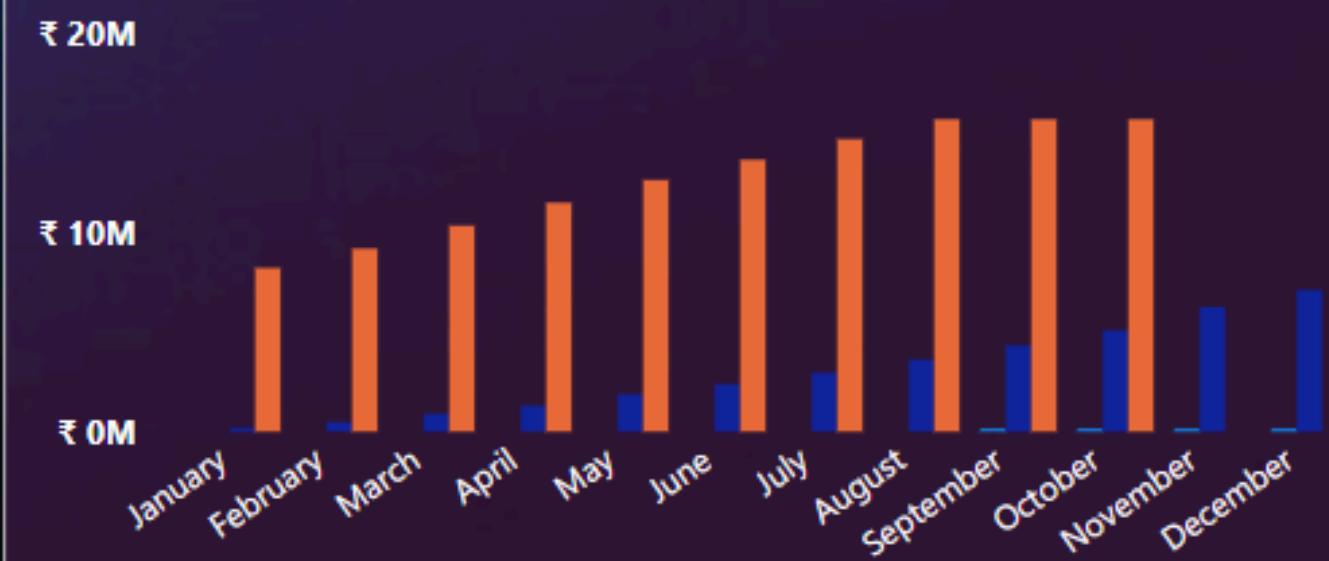
Date	Running_total
17-01-2017	₹ 77,770.53
18-01-2017	₹ 82,453.91
19-01-2017	₹ 86,594.07
20-01-2017	₹ 91,250.12
21-01-2017	₹ 94,344.89
22-01-2017	₹ 98,486.76
23-01-2017	₹ 1,06,040.65
24-01-2017	₹ 1,12,824.18
25-01-2017	₹ 1,23,164.39
26-01-2017	₹ 1,38,211.8
27-01-2017	₹ 1,50,967.95
28-01-2017	₹ 1,56,850.96
Total	₹ 1,57,35,527.03

### Revenue By City

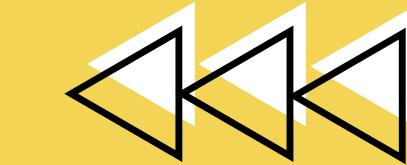
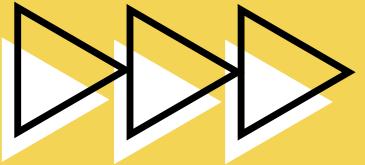


### Running\_total by Month and year

year ● 2016 ● 2017 ● 2018



# PAGE 1: EXECUTIVE SALES OVERVIEW



## Key Findings:

Revenue: R\$15.74M total. Strong mid-year peak, sharp October drop.

Customers: 98.2K active, but 0% repeat customer rate is a major concern.

Orders: 98.2K total, suggesting many single purchases.

Average Order Value: R\$160.24.

Growth: Significant revenue growth from 2016 to 2017.

Location: Sao Paulo dominates revenue.

MoM Growth: Highly volatile, extreme percentages suggest data review needed.

## Top Actions:

Immediately investigate and address the 0% repeat customer rate.

Analyze the October revenue drop.

Leverage the strong Sao Paulo market.

Review the accuracy of MoM growth data.



# PAGE 2: CUSTOMER INSIGHTS:-

Total Customers

98.2K

Average Orders per Customer

1.00

Customer Retention Rate

0%

New Customers by Year

- Date Year
- 2016
- 2017
- 2018



Executive Sales Overview

Product Performance

Seller Analytics

Fulfillment Metrics

Payment & Financial Insights

Seller\_id

All

Customer\_id

All

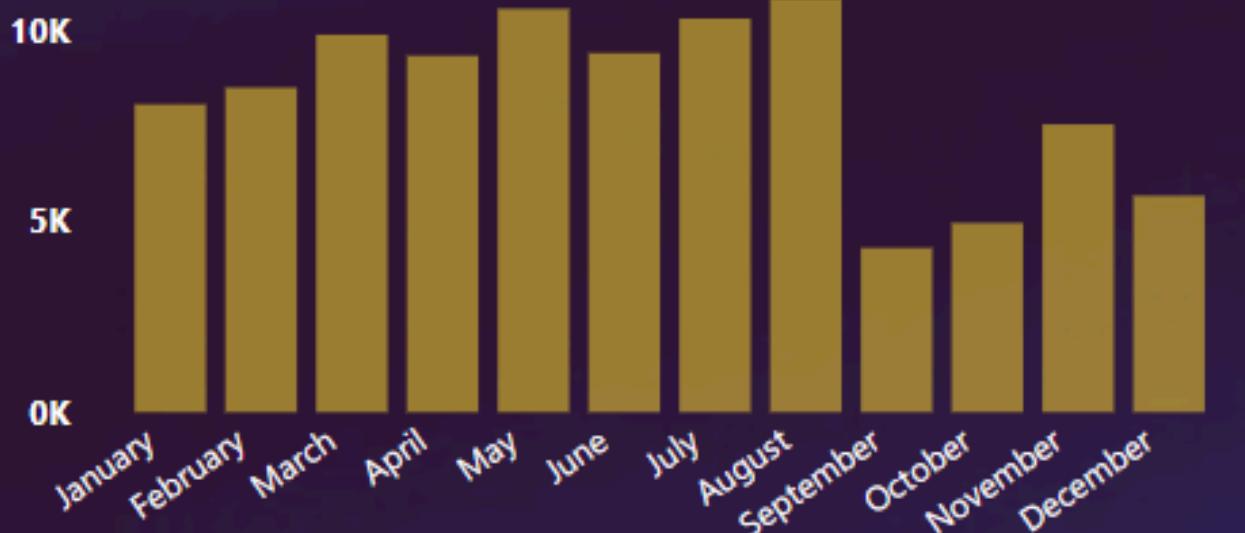
Monthna...

(Blank) 04-09-2016

April 17-10-2018

Distinct

New vs Returning Customers (monthly)



Total Orders by orders hours.



Top States by Number of Customers



Customer Locations



Average Review Score by States

Review ... ● 1 ● 2 ● 3 ● 4 ● 5



# PAGE 2: CUSTOMER INSIGHTS

## Key Findings:

- Acquisition: 98.2K total customers, peaked in 2017.
- Engagement: Very low engagement with 1.00 average order per customer.
- Retention: Critical issue: 0% customer retention rate.
- Location: Most customers in São Paulo (SP), but geographically spread.
- Sentiment: Review scores vary by state, RJ has higher average.

## Top Priority:

- Immediately investigate and fix the 0% customer retention rate.  
Understand why customers aren't returning.

# PAGE 3: PRODUCT PERFORMANCE:-



**Executive Sales Overview**

**Customer Insights**

**Seller Analytics**

**Fulfillment Metrics**

**Payment & Financial Insights**

**Seller\_id**  
All

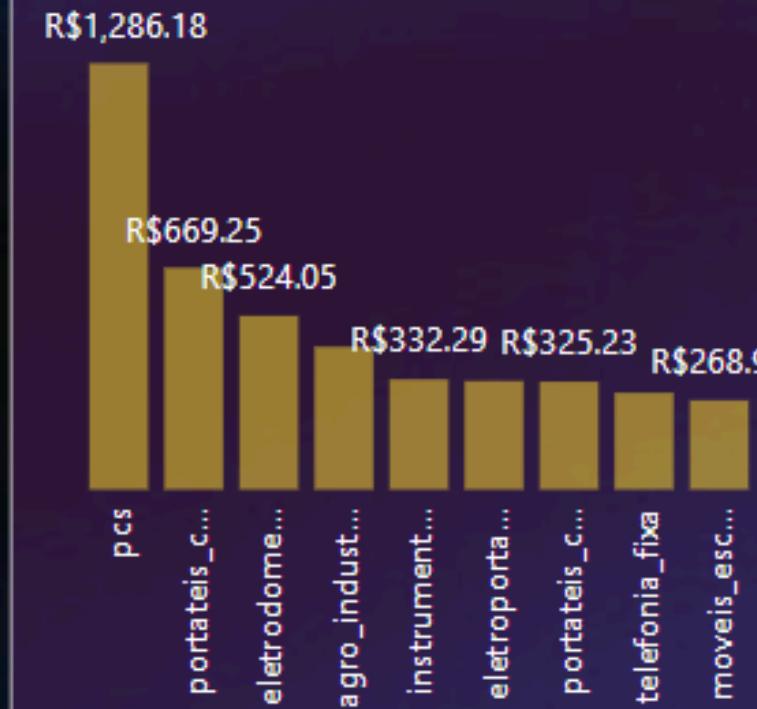
**Customer\_id**  
All

**Monthna...**  
 (Blank)  
 April  
 August  
  
**Date**  
04-09-2016  
17-10-2018  
  
**Total** 100.00%

## Top 10 Product Categories by Revenue



## Category-wise Average Order Value



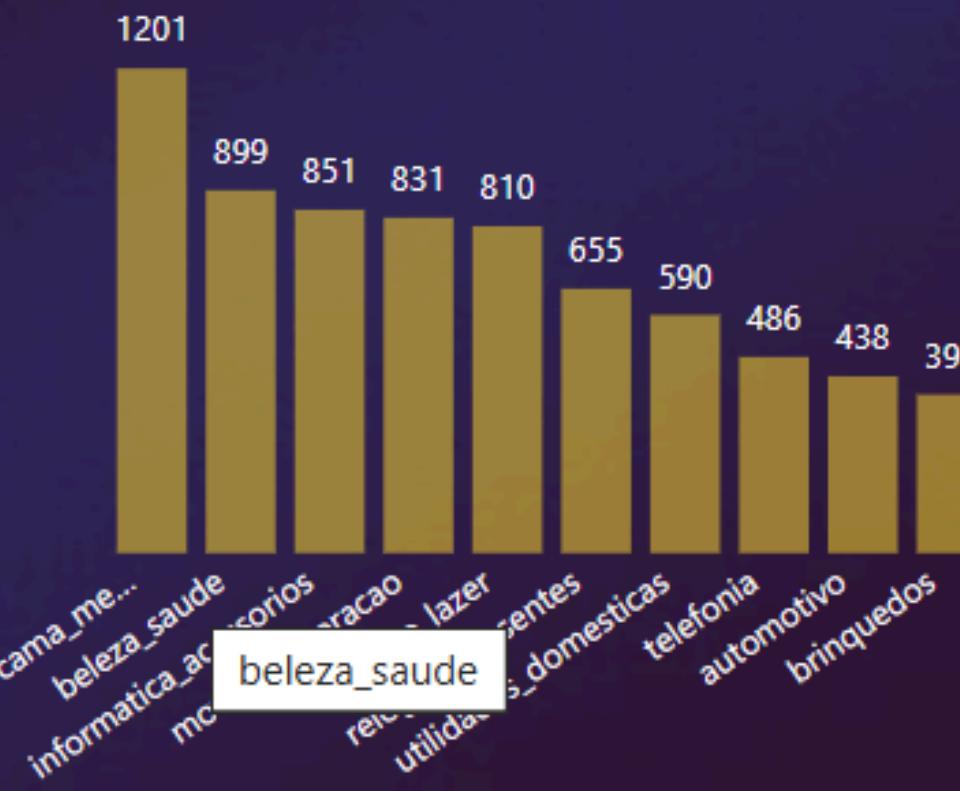
## Order Volume by Product Category



## % Share of Each Category

product_category_name	% Share of Category
cama_mesa_banh...	9.87%
beleza_saude	8.58%
esporte_lazer	7.67%
moveis_decorac...	7.40%
informatica_acess...	6.95%
utilidades_domes...	6.18%
relogios_pres...	5.32%
telefonia	4.03%
<b>Total</b>	<b>100.00%</b>

## Bottom 10 SKUs by Review Score(1)

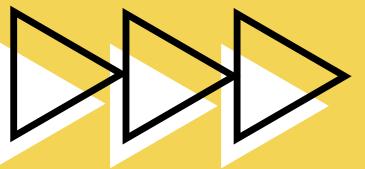


## Review Score by State

Review Score ● 1 ● 2 ● 3 ● 4 ● 5



# PAGE 3: PRODUCT PERFORMANCE

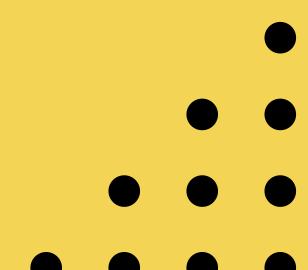
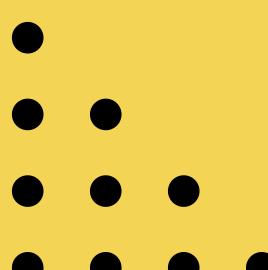


## Key Findings:

- Top Revenue: Beauty & Health, Watches & Gifts, Bed/Bath/Table.
- High Value: Laptops have the highest average order value.
- High Volume: Bed/Bath/Table has the most orders.
- Low Reviews: Some top categories (Bed/Bath/Table, Beauty & Health) have poorly rated SKUs. Bed/Bath/Table generally has better state-wise reviews.

## Top Actions:

- Investigate and improve the low-rated SKUs, especially in top categories.
- Understand the drivers of high average order value for Laptops.
- Leverage the popularity of top revenue and volume categories.



# PAGE 4: SELLER ANALYTICS:-

Total Sellers  
**3095**



**Executive Sales Overview** ...

**Customer Insights**

**Product Performance**

**Fulfillment Metrics**

**Payment & Financial Insights**

**Seller\_id**

All

**Customer\_id**

All

**Monthna...**

(Blank)

April

August

## Top 10 Sellers id by Revenue



## Order Count per Seller



seller_id	Average of Delivery Time (Days)
0015a82c2db000af	10
6aaaf3ae2ecb0532	13
001cca7ae9ae17fb1	16
caed9dfb1094831	4
002100f778ceb843	14
1b7a1020ff7ab48f	11
003554e2dce176b5	9
555353e4f3555ac8	12
004c9cd9d87a3c30	
c522c48c4fc07416	
00720abe85ba0859	
807595bbf045a33b	
00ab3eff1b5192e5f	
<b>Average delivery date</b>	

## Seller Efficiency (Delivery vs Estimation)

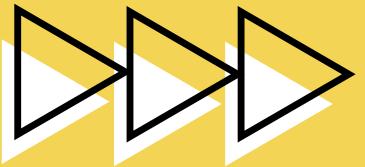
seller_id	deliver actual days	estimated delivery days
05e107217c726636	4	3
2fd44b75b2cd4cc4		
06a2c3af7b3aee5d	3	3
69171b0e14f0ee87		
0885AAF116795758	3	3
dfeb5f1032487bcd		
0beb82cea103a4a	1	3
020c95d43fd7d754		
0beb82cea103a4a	2	3
020c95d43fd7d754		
116ccb1a1604bc88	4	3
e4d234a8c23f33de		
1354d51653f64534	2	3
9064725ed204e85e		



## Revenue per Seller by Product Category

seller_id	product_cat	Total Revenue
7c67e1448b00f6e9	moveis_escritorio	R\$2,19,595.93
69d365cea6b010ab	relogios_premium	R\$2,18,530.99
4869f7a5dfa277a7d	sentences	R\$2,00,830.41
ca6462dcf3b52b2	relogios_premium	
fa1c13f2614d7b5c4	sentences	
749cbc52fecda94	camas_mesa_banho	
4a3ca9315b744ce9f	camas_mesa_banho	R\$1,93,268.33
8e9374361493884	camas_mesa_banho	
da8622b14eb17ae2	camas_mesa_banho	R\$1,74,270.23
831f4ac5b9dab84a	camas_mesa_banho	
53243585a1d6dc26	pcss	R\$1,73,995.95
43021fd1853d8905	relogios_premium	
7e93a43ef30c4f03f	relogios_premium	R\$1,71,743.86
<b>Total</b>		<b>R\$1,57,35,52</b>
		7.03

# PAGE 4: SELLER ANALYTICS



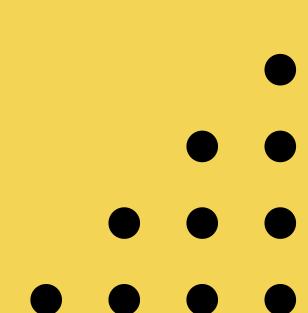
## Key Findings:

- Top Sellers: Revenue and order volume are concentrated among a few top sellers.
- Delivery: Average delivery times vary significantly (4-16 days), with an overall average of 7 days. Most sampled sellers meet delivery estimates.
- Location: Sellers are concentrated in southeastern Brazil.
- Category Revenue: Performance varies by seller and product category.



## Top Actions:

- Improve delivery times for slower sellers.
  - Learn from top-performing sellers.
  - Consider geographical expansion of the seller base.
- •  
•  
•  
•



# PAGE 5: OPERATIONAL & FULFILLMENT METRICS:-

Most of the orders are delivered around

**16.23**



**Executive Sales Overview**

**Customer Insights**

**Product Performance**

**Seller Analytics**

**Payment & Financial Insights**

**Seller\_id**

**Customer\_id**

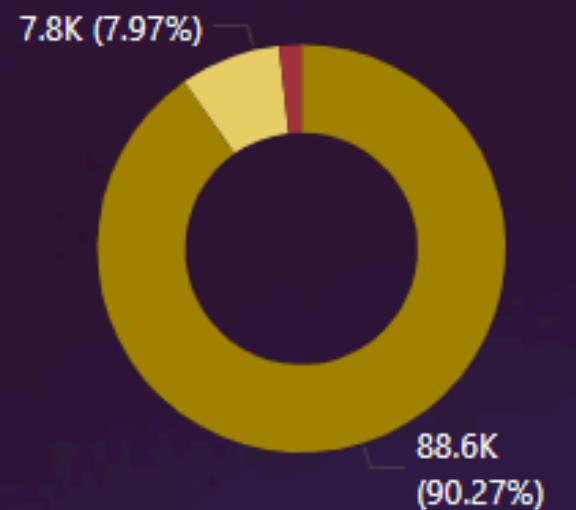
**Monthna...**

 (Blank)
  April
  August
 

**Date**



## On-time Delivery %



Average Delivery Time (Days)

**12**

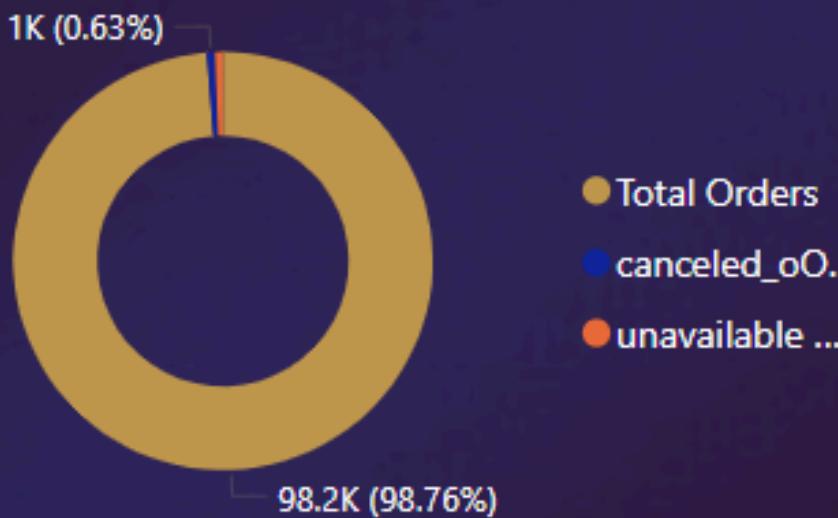
Average Freight Cost per Category

pcs	R\$48.45
eletrodomesti...	R\$44.54
moveis_colch...	R\$42.91
moveis_cozin...	R\$42.70
moveis_quarto	R\$42.50
moveis_escrit...	R\$40.55
portateis_casa...	R\$36.16
moveis_sala	R\$35.72
sinalizacao_e...	R\$32.70
industria_com...	R\$29.42
malas_acessor...	R\$27.88

Freight Cost per Order

order_id	freight_value
00010242fe8c5a6d	R\$13.29
1ba2dd792cb16214	
00018f77f2f0320c5	R\$19.93
57190d7a144bdd3	
000229ec398224ef	R\$17.87
6ca0657da4fc703e	
00024acbcdf0a6da	R\$12.79
a1e931b038114c75	
00042b26cf59d7ce	R\$18.14
69dfabb4e55b4fd9	
00048cc3ae777c65	R\$12.69
dbb7d2a0634bc1ea	
00054e8431b9d767	R\$11.85

## Cancelled vs Completed Orders



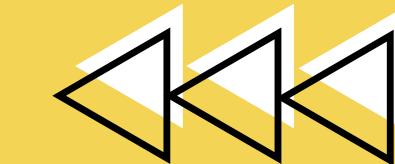
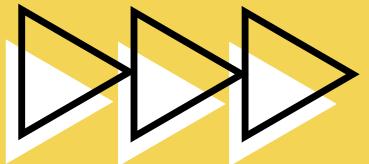
Most Delayed Orders (Top 10)

order_id	deliver actual days
ca07593549f1816d26a572e06dc1eab6	210
1b3190b2dfa9d789e1f14c05b647a14a	208
440d0d17af552815d15a9e41abe49359	196
285ab9426d6982034523a855f55a885e	195
2fb597c2f772eca01b1f5c561bf6cc7b	195
0f4519c5f1c541ddec9f21b3bdd533a	194
47b40429ed8cce3aee9199792275433f	191
2fe324feb907e3ea3f2aa9650869fa5	190
2d7561026d542c8dbd8f0daeadf67a43	188
c27815f7e3dd0b926b58552628481575	188

State-wise Avg. Shipping Duration (Days)

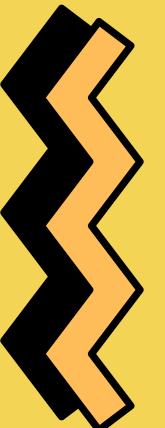
RR	29
AP	27
AM	26
AL	25
PA	24
MA	22
SE	21
CE	21

# PAGE 5: OPERATIONAL & FULFILLMENT METRICS



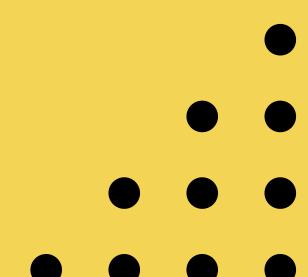
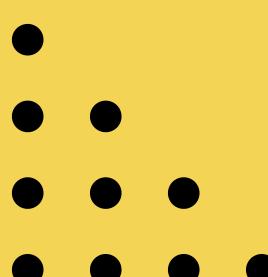
## Key Findings:

- Delivery: Average delivery time is 12 days, with 90% on time.
- Freight: Costs vary by category, with "pcs" being highest.
- Cancellations: Very low (under 1%).
- Delays: Top 10 orders have extreme delays (188-210 days).
- Shipping Time by State: Varies significantly, with RS being the longest (29 days).



## Top Actions:

- Reduce the overall average delivery time.
- Immediately address the extremely delayed orders.
- Optimize shipping times in states with long durations (e.g., RS).



# PAGE 6: PAYMENT & FINANCIAL INSIGHTS:-

Cancel Revenue

R\$ 108.03K

Total Cancel Orders

625

Refund Rate

0.0%

Total Unavailable Orders

609



Executive Sales Overview

Customer Insights

Product Performance

Seller Analytics

Fulfillment Metrics

**Seller\_id**

**Customer\_id**

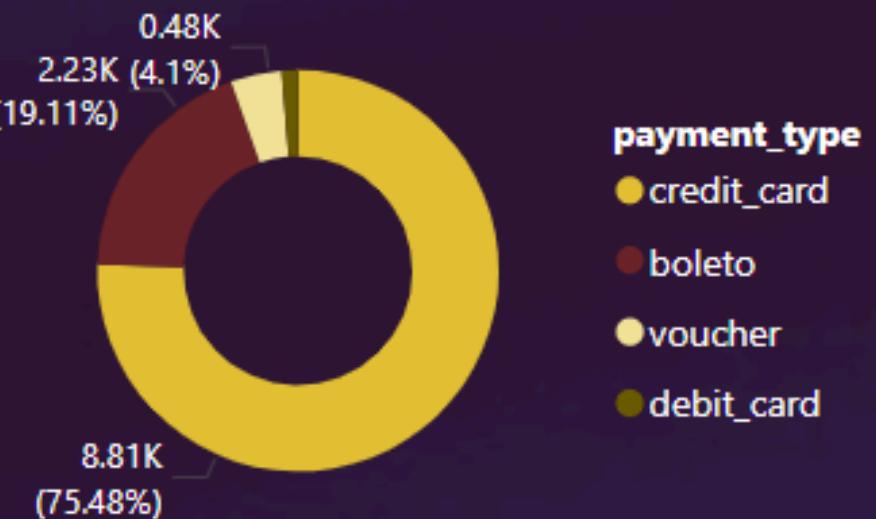
**Monthname**

 (Blank)
  April
  August
   

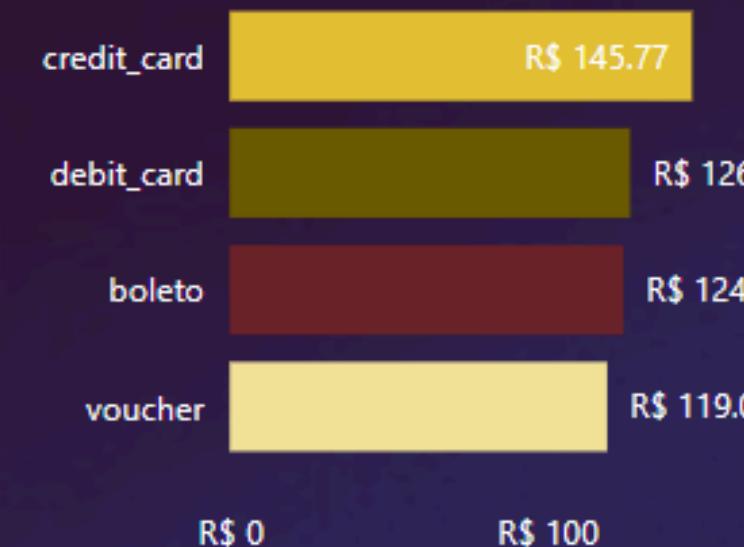
**Date**



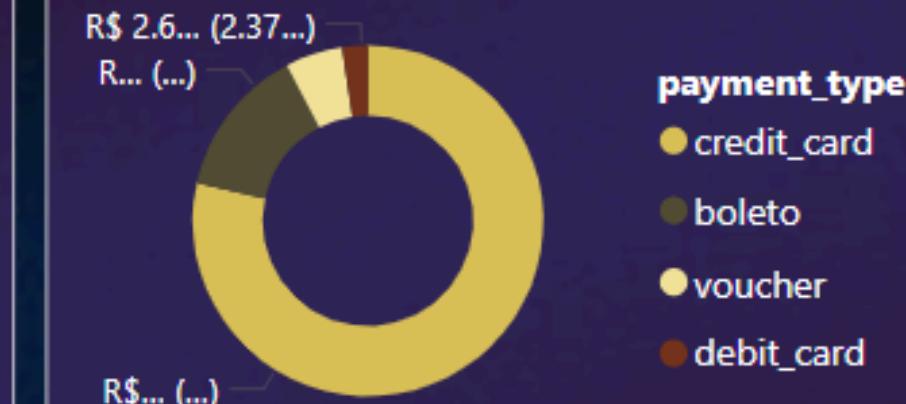
## Payment Methods Distribution



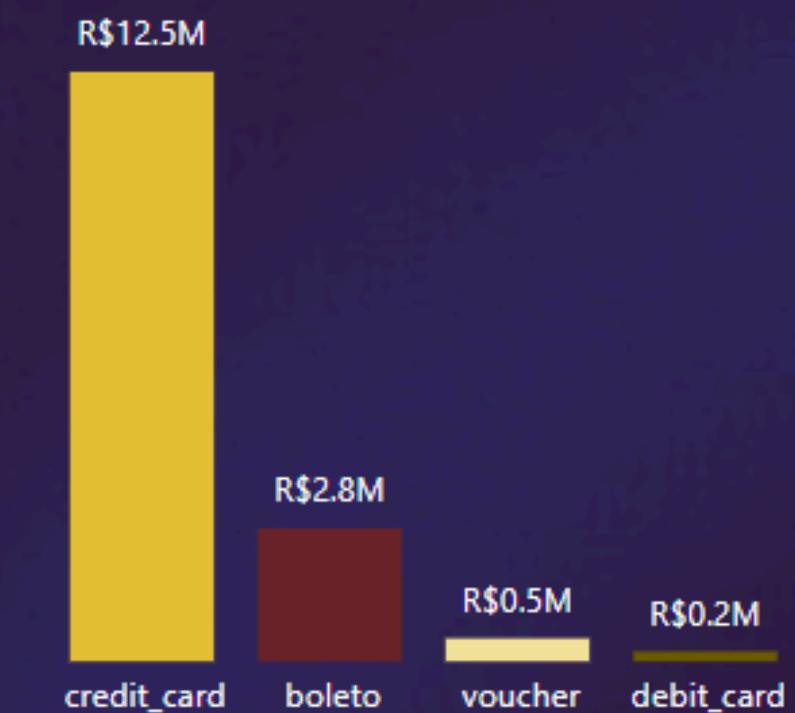
## Avg Payment Value per Method



## Cancel Payment Method



## Revenue by Payment Method

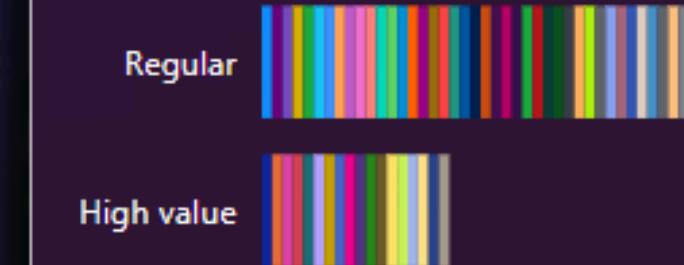


## Month-on-Month Change in Transaction Volume

Year	Month	MOM%GROTH	Total Orders
2016	September	0.0%	2.0
2016	October	18261.2%	290.0
2016	November	-100.0%	
2016	December	0.0%	1.0
2017	January	697878.9%	787.0
2017	February	107.1%	1718.0
2017	March	50.1%	2617.0
2017	April	-4.6%	2377.0
2017	May	43.6%	3640.0
2017	June	-14.3%	3205.0
2017	July	15.8%	3946.0
<b>Total</b>		<b>0.0%</b>	<b>98199.0</b>

## High-Value Orders Breakdown

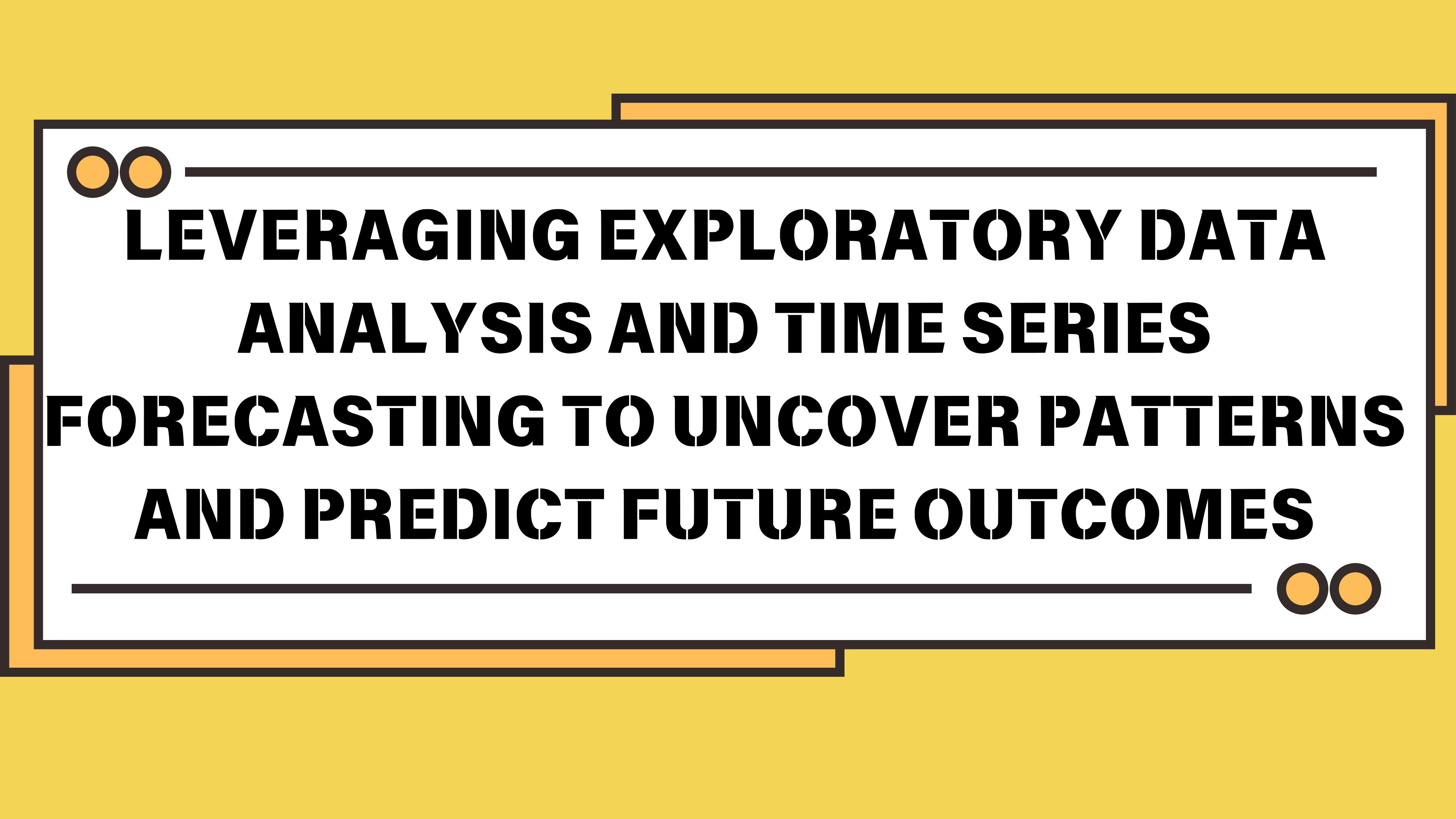
order\_id ● 00010242fe8... ● 00018f77f2f... ▾



Count of product\_category\_name

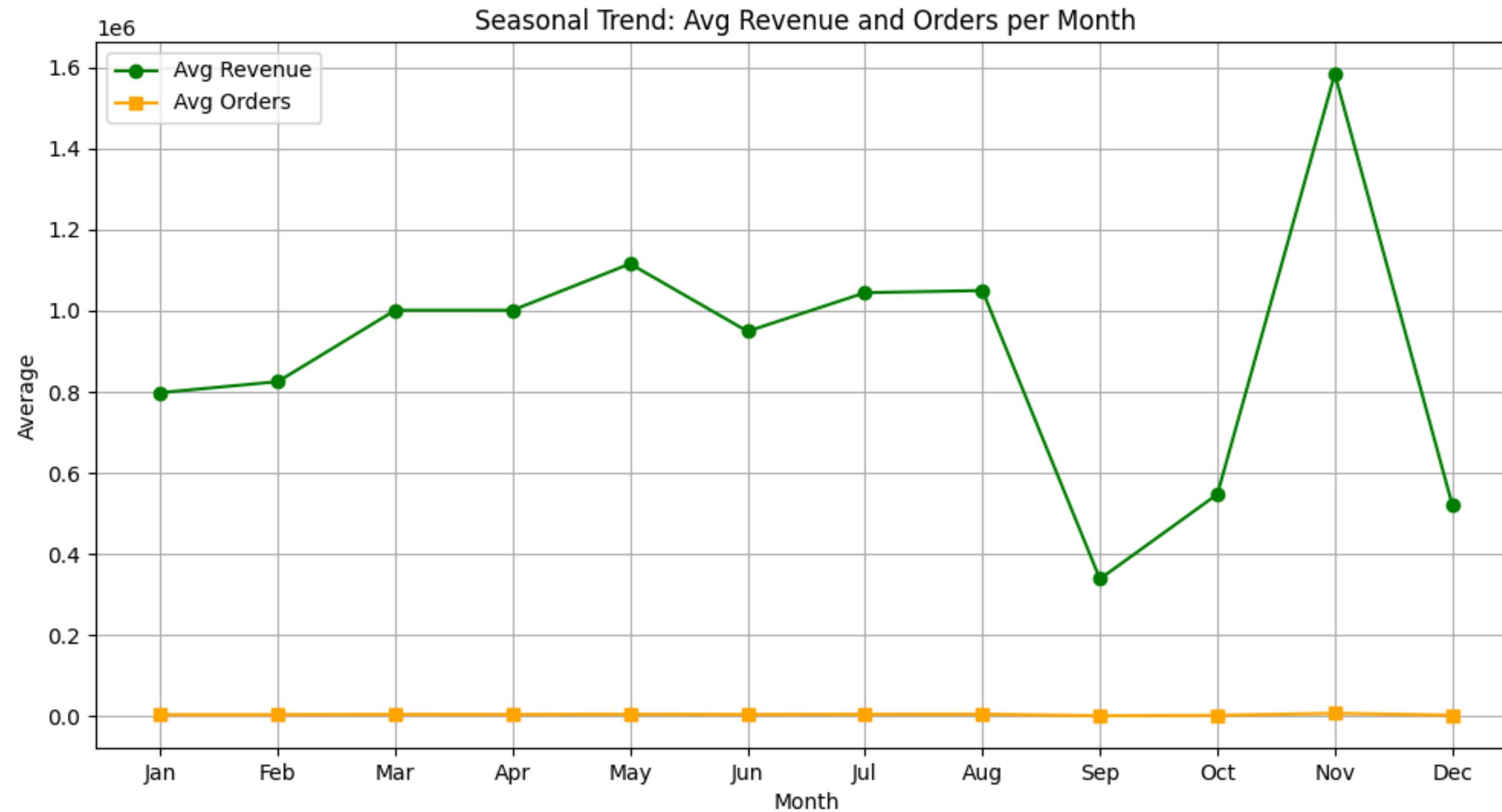
# PAGE 6: PAYMENT & FINANCIAL INSIGHTS

- Cancellations: R\$108K revenue lost from 625 cancelled orders.
  - Refunds: Unusually low 0% refund rate.
  - Unavailable: 609 orders unavailable.
  - Payments: Boleto is most popular, credit card has highest average value.
  - MoM Volume: Highly volatile transaction volume.
  - Top Actions:
    - Investigate the 0% refund rate.
    - Analyze reasons for cancelled and unavailable orders.
    - Review the accuracy of the volatile MoM transaction data.



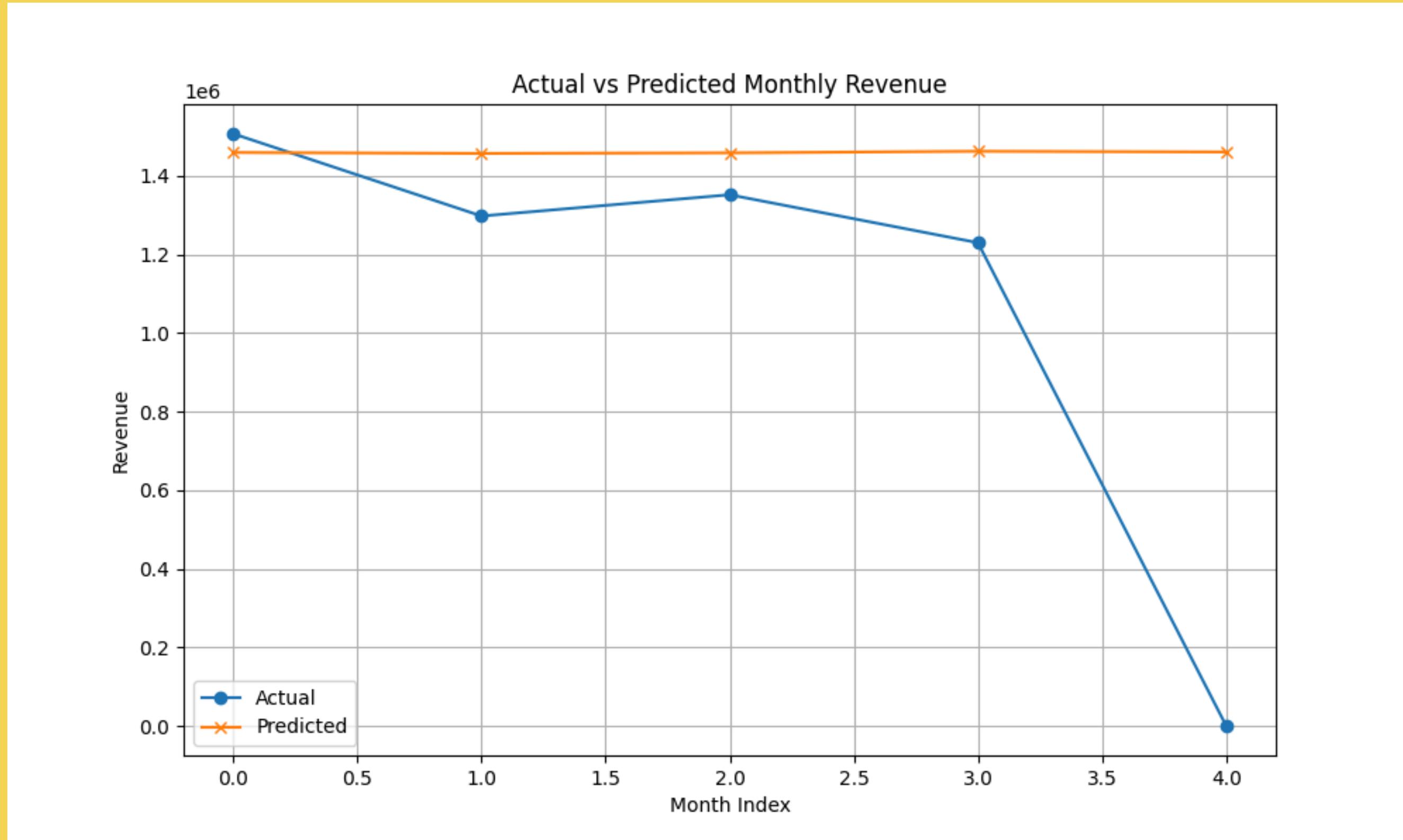
# **LEVERAGING EXPLORATORY DATA ANALYSIS AND TIME SERIES FORECASTING TO UNCOVER PATTERNS AND PREDICT FUTURE OUTCOMES**

# Plot monthly trends



## Identify peak/dip months

- Peak Revenue Month: November 2017 - ₹1583869.01
- Lowest Revenue Month: December 2016 - ₹19.62



## TOP & BOTTOM MONTHS REVENUE

### ↑ Top 3 Revenue Months:

	order_month	payment_value
13	2017-11-01	1583869.01
19	2018-05-01	1506974.84
18	2018-04-01	1496811.52

month revenue

2016-09-01 252.24  
2016-10-01 59090.48

### ▼ Bottom 3 Revenue Months:

	order_month	payment_value
2	2016-12-01	19.62
23	2018-09-01	166.46
0	2016-09-01	347.52

2016-12-01 19.62  
2017-01-01 138488.04  
2017-02-01 291908.01

## TOP & BOTTOM MONTHS REVENUE

↑ Top 3 Revenue Months:

	order_month	payment_value
13	2017-11-01	1583869.01
19	2018-05-01	1506974.84
18	2018-04-01	1496811.52

▼ Bottom 3 Revenue Months:

	order_month	payment_value
2	2016-12-01	19.62
23	2018-09-01	166.46
0	2016-09-01	347.52

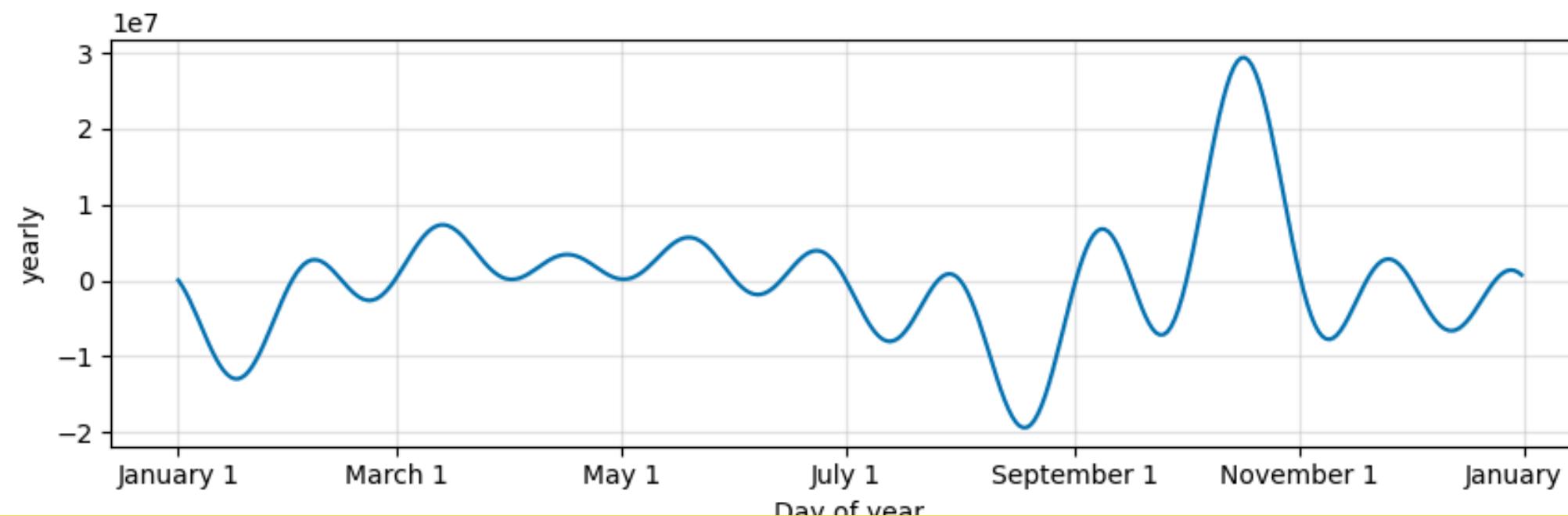
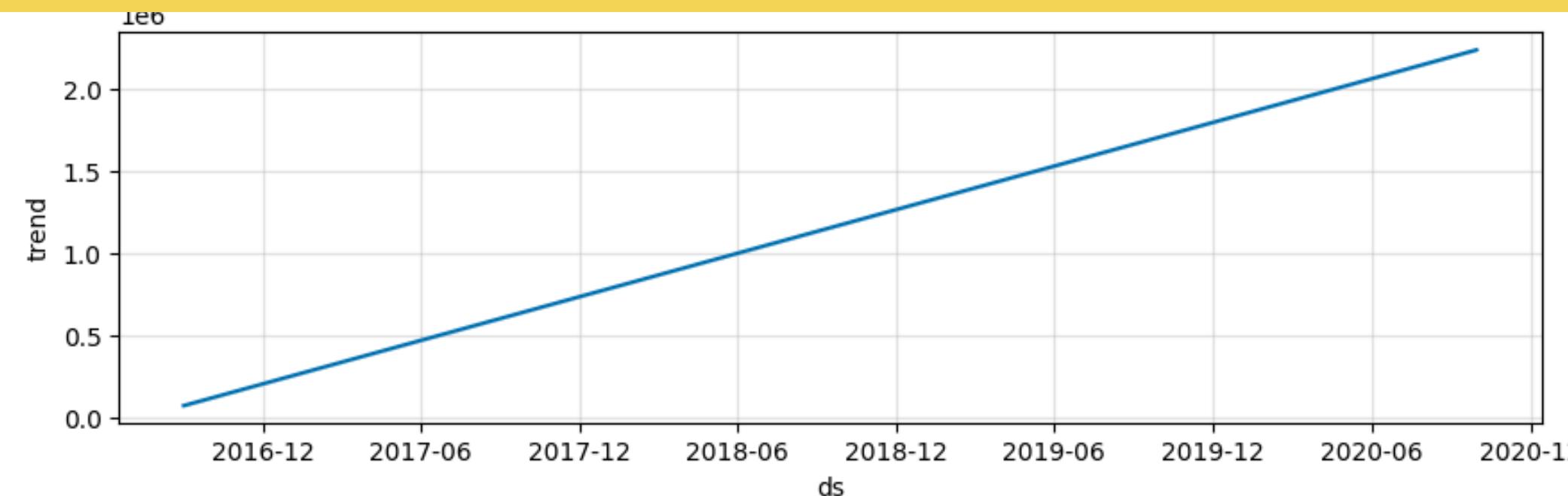
# EVALUATION MATRIX

## Evaluation Metrics:

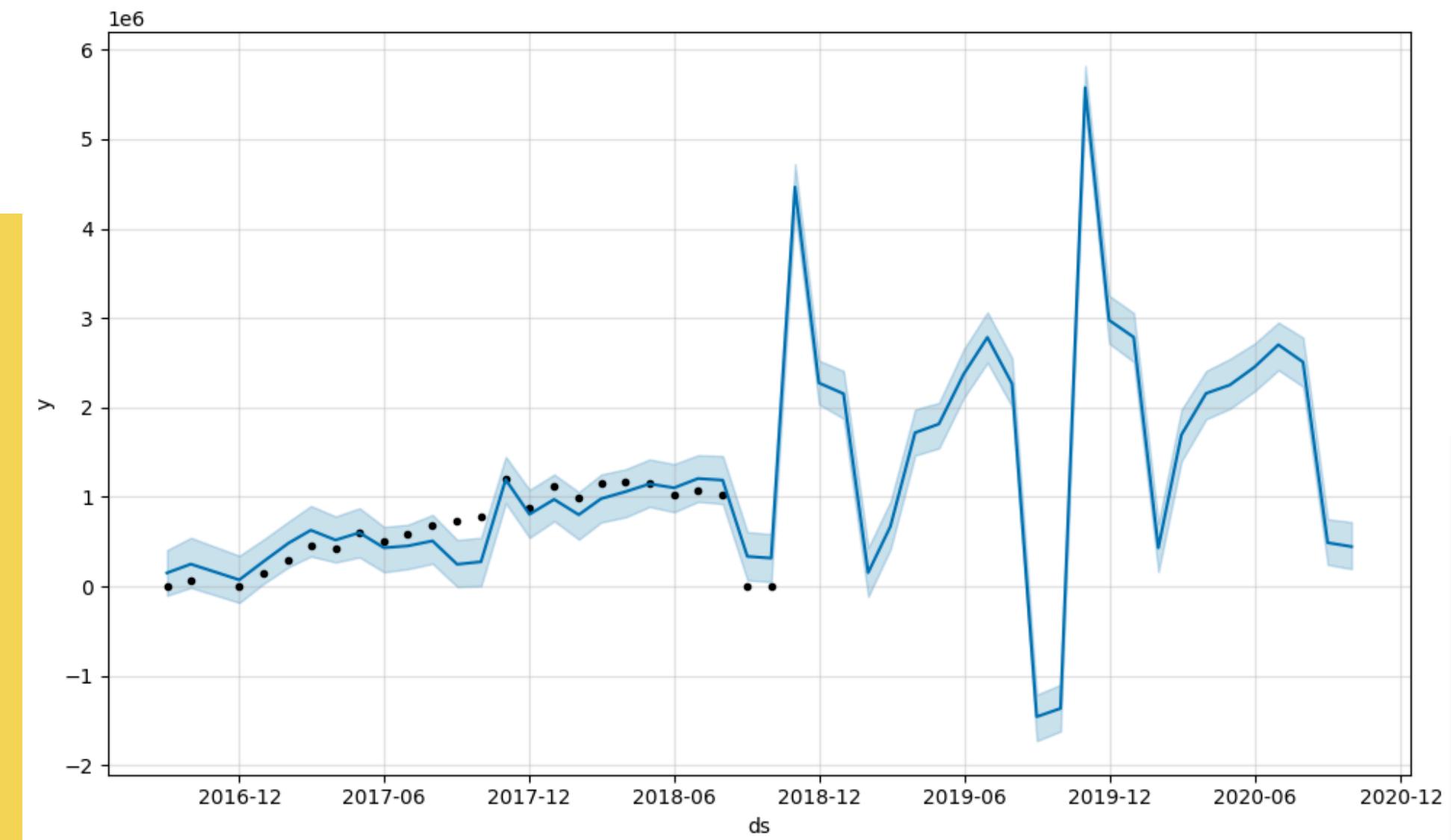
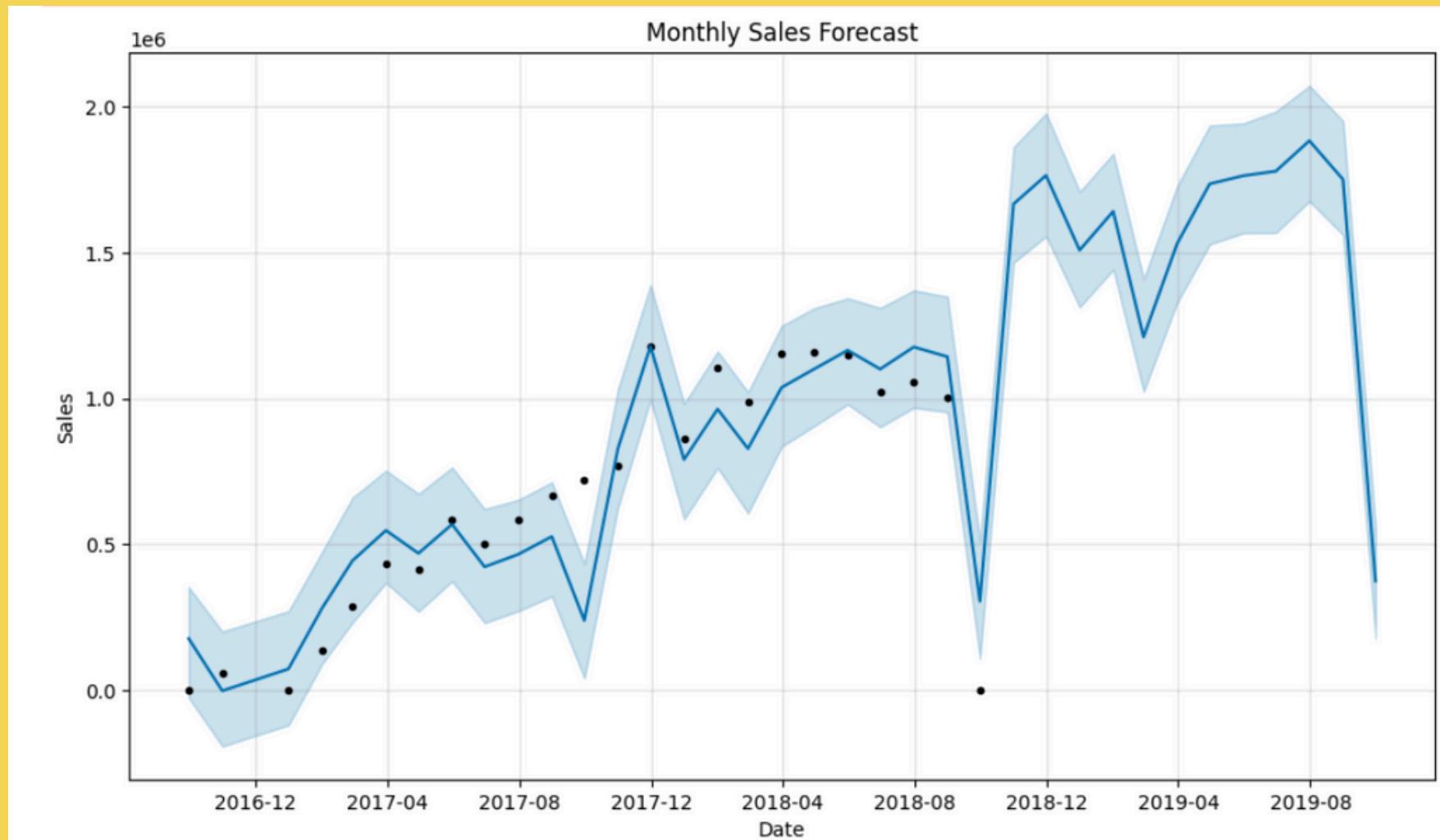
R<sup>2</sup> Score : -0.4921

RMSE : 667239.42

MAE : 401353.66



# MONTHLY SALE FORECAST



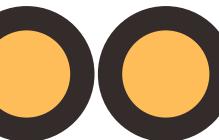
You provided several visualizations related to your sales data and a goal to perform EDA and time-series forecasting in Python. We observed:

1. Initial "Actual vs Predicted" plot: Showed a flat prediction line, indicating a poorly performing initial model.
- Subsequent Forecast Plot (likely Prophet): Displayed historical data, a more dynamic forecast capturing trend and yearly seasonality, and uncertainty intervals. This suggested a more sophisticated model was used.
- Top/Bottom Revenue Months: This was identified as descriptive EDA, providing insights into past performance.

#### Key Takeaways and Next Steps:

- You've likely moved beyond basic EDA and implemented a time-series forecasting model (potentially Prophet).
- The latest forecast plot shows the model capturing trend and seasonality, which is a positive step.
- Crucially, the next step is to rigorously evaluate the model's performance using a train-test split and appropriate evaluation metrics (MAE, RMSE, MAPE) on unseen data.
- Analyzing the trend and seasonality components separately (if your model allows) will provide further insights.
- Understanding the reasons behind significant past events (like the revenue spikes) and whether the model accurately predicts their recurrence is important.
- Consider applying similar forecasting techniques to individual product categories for more granular predictions.

In essence, you've made progress in building a time-series forecasting model. The immediate focus should now be on quantitatively evaluating its accuracy and further interpreting the results (trend, seasonality, uncertainty) in the context of your business goals.



# **COMPREHENSIVE SALES ANALYSIS TO ACCELERATE E- COMMERCE GROWTH**



# Executive Summary

- This project uses data analytics to optimize Mercado Livre's operations and drive growth.
- The project involves SQL analysis, dashboard creation, and predictive modeling to provide actionable insights.
- Key findings include revenue growth, customer acquisition trends, and operational inefficiencies.
- Actionable steps are proposed to improve customer retention, delivery times, and overall performance.
- The goal is to empower Mercado Livre to make data-driven decisions and enhance its competitive edge.

## • Sales Performance Insights

- Total revenue was R\$15.74M, with a mid-year peak and an October drop.
- 
- Average Order Value: R\$160.24.
- 
- Significant revenue growth occurred from 2016 to 2017.
- 
- Month-over-month revenue growth was highly volatile.
- 
- Top revenue categories include Beauty & Health, Watches & Gifts, and Bed/Bath/Table.
- 
- Laptops have the highest average order value.

# Customer Insights

- Total customers: 98.2K, with a peak in 2017.
- 
- Very low customer engagement: 1.00 average order per customer.
- 
- Critical issue: 0% customer retention rate.
- 
- Most customers are in São Paulo (SP), but they are geographically spread.
- 
- Review scores vary by state.

⋮  
⋮⋮⋮

⋮⋮⋮

# Regional & Operational Insights

- Sao Paulo dominates revenue.
- 
- Delivery times vary significantly (4-16 days), with an average of 7 days.
- 
- Seller concentration in southeastern Brazil.
- 
- Average delivery time is 12 days, with 90% on time.
- 
- Shipping time varies significantly by state, with RS being the longest (29 days).
- 
- A small percentage of orders were canceled.
- 
- 
- 
-

# Recommendations for Stakeholders

- Investigate and address the 0% repeat customer rate.
- 
- Analyze the October revenue drop.
- 
- Improve low-rated SKUs, especially in top categories.
- 
- Optimize delivery times, especially for slower sellers and in states with longer durations.
- 
- Analyze reasons for canceled and unavailable orders.
- 
- Review the accuracy of volatile MoM data (revenue and transaction volume).
- 
- 
- 
-

# Conclusion

- The project provides data-driven insights to optimize Mercado Livre's sales, customer engagement, and operations.
  - Addressing the 0% customer retention rate is a top priority.
  - Improving delivery efficiency and SKU ratings are crucial for enhancing customer satisfaction and driving growth.
  - Further analysis and monitoring of volatile trends are recommended.



**GET IN TOUCH**

**WITH ME!**

**GITHUB LINKEDIN**