mercado livre

# E-COMMERCE GROWTH ANALYSIS

-KARTHIKEYA M SHROUTHY

# ABOUT MERCADO LIVRE

MercadoLibre, Inc in Spanish, and known as Mercado Livre in Portuguese.

Mercado Livre is an Argentine company headquartered in Montevideo, Uruguay and incorporated in Delaware in the United States that operates online marketplaces dedicated to e-commerce and online auctions.

Mercado Libre also runs a real estate and motors division under the name “*Mercado Libre Classificados*”. Realtors pay a monthly fee to list properties and automobiles on the Mercado Libre platform.

# TABLE OF CONTENT



# PROBLEM STATEMENT

Mercado Livre, a leading Brazilian e-commerce platform, aims to optimize its marketplace operations by gaining deeper insights into customer behavior, seller performance, product trends, and fulfillment efficiency. The goal is to uncover growth opportunities, improve customer experience, and forecast future sales trends to support data-driven decision-making.

## STEPS INVOLVED:

**PART 1: SQL ANALYSIS**

**PART 2: DASHBOARD CREATION - POWER BI**

**PART 3: PYTHON – EDA & TREND ANALYSIS**

## FOR EACH MONTH, CALCULATE TOTAL REVENUE AND MONTH-OVER-MONTH GROWTH PERCENTAGE.

```sql
SELECT
  DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m') AS month,
  SUM(oi.price) AS revenue,
  ROUND(
    (SUM(oi.price) - LAG(SUM(oi.price))
    OVER (ORDER BY DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m')))
    / NULLIF(LAG(SUM(oi.price))
    OVER (ORDER BY DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m')), 0) * 100, 2
  ) AS growth
FROM
  order_items_dataset oi
JOIN
  orders_dataset o ON o.order_id = oi.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  month
ORDER BY
  month;
```

| month | revenue | growth |
|---|---|---|
| 2016-09 | 134.97 | NULL |
| 2016-10 | 40325.11 | 29777.09 |
| 2016-12 | 10.90 | -99.97 |
| 2017-01 | 111798.36 | 1025573.03 |
| 2017-02 | 234223.40 | 109.51 |
| 2017-03 | 359198.85 | 53.36 |
| 2017-04 | 340669.68 | -5.16 |
| 2017-05 | 489338.25 | 43.64 |
| 2017-06 | 421923.37 | -13.78 |
| 2017-07 | 481604.52 | 14.15 |
| 2017-08 | 554699.70 | 15.18 |
| 2017-09 | 607399.67 | 9.50 |
| 2017-10 | 648247.65 | 6.73 |
| 2017-11 | 987765.37 | 52.37 |
| 2017-12 | 726033.19 | -26.50 |
| 2018-01 | 924645.00 | 27.36 |
| 2018-02 | 826437.13 | -10.62 |
| 2018-03 | 953356.25 | 15.36 |
| 2018-04 | 973534.09 | 2.12 |

## FOR EACH MONTH, CALCULATE THE 3-MONTH MOVING AVERAGE OF THE NUMBER OF ORDERS.

```sql
SELECT
  month,
  orders,
  ROUND(AVG(orders) OVER (ORDER BY month ROWS 2 PRECEDING), 2) AS moving_avg
FROM (
  SELECT
    DATE_FORMAT(order_purchase_timestamp, '%Y-%m') AS month,
    COUNT(*) AS orders
  FROM
    orders_dataset
  WHERE
    order_status = 'delivered'
  GROUP BY
    month
) AS monthly_orders;
```

| | month | orders | moving_avg |
|---|---|---|---|
| ▶ | 2016-09 | 1 | 1.00 |
| | 2016-10 | 265 | 133.00 |
| | 2016-12 | 1 | 89.00 |
| | 2017-01 | 750 | 338.67 |
| | 2017-02 | 1653 | 801.33 |
| | 2017-03 | 2546 | 1649.67 |
| | 2017-04 | 2303 | 2167.33 |
| | 2017-05 | 3546 | 2798.33 |
| | 2017-06 | 3135 | 2994.67 |
| | 2017-07 | 3872 | 3517.67 |
| | 2017-08 | 4193 | 3733.33 |
| | 2017-09 | 4150 | 4071.67 |
| | 2017-10 | 4478 | 4273.67 |
| | 2017-11 | 7289 | 5305.67 |
| | 2017-12 | 5513 | 5760.00 |
| | 2018-01 | 7069 | 6623.67 |
| | 2018-02 | 6555 | 6379.00 |
| | 2018-03 | 7003 | 6875.67 |
| | 2018-04 | 6798 | 6785.33 |

## YEARLY AVERAGE ORDER VALUE (AOV): TREND COMPUTE AVERAGE ORDER VALUE PER YEAR.

```sql
SELECT
  YEAR(o.order_purchase_timestamp) AS year,
  ROUND(SUM(oi.price) / COUNT(DISTINCT o.order_id), 2) AS aov
FROM
  order_items_dataset oi
JOIN
  orders_dataset o ON o.order_id = oi.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  year
ORDER BY
  year;
```

| year | aov |
|---|---|
| 2016 | 151.58 |
| 2017 | 137.31 |
| 2018 | 136.75 |

## CALCULATE TOTAL REVENUE GENERATED BY EACH CUSTOMER. SORT TOP 10 HIGHEST LIFETIME VALUES.

```sql
SELECT
  c.customer_unique_id,
  ROUND(SUM(oi.price), 2) AS cltv
FROM
  customers_dataset c
JOIN
  orders_dataset o ON o.customer_id = c.customer_id
JOIN
  order_items_dataset oi ON oi.order_id = o.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  c.customer_unique_id
ORDER BY
  cltv DESC
LIMIT 10;
```

| customer_unique_id | cltv |
|---|---|
| 0a0a92112bd4c708ca5fde585afaa872 | 13440.00 |
| da122df9eeddfedc1dc1f5349a1a690c | 7388.00 |
| 763c8b1c9c68a0229c42c9fc6f662b93 | 7160.00 |
| dc4802a71eae9be1dd28f5d788ceb526 | 6735.00 |
| 459bef486812aa25204be022145caa62 | 6729.00 |
| ff4159b92c40ebe40454e3e6a7c35ed6 | 6499.00 |
| 4007669dec559734d6f53e029e360987 | 5934.60 |
| eebb5dda148d3893cdaf5b5ca3040ccb | 4690.00 |
| 48e1ac109decbb87765a3eade6854098 | 4590.00 |
| a229eba70ec1c2abef51f04987deb7a5 | 4400.00 |

# IDENTIFY TOP 5 CATEGORIES WITH THE HIGHEST YEAR-OVER-YEAR REVENUE GROWTH.

```sql
SELECT category, year,
ROUND(((revenue - prev_revenue) / NULLIF(prev_revenue, 0)) * 100, 2) AS growth
FROM (
  SELECT
    pc.product_category_name_english AS category,
    YEAR(o.order_purchase_timestamp) AS year,
    SUM(oi.price) AS revenue, LAG(SUM(oi.price))
    OVER (PARTITION BY pc.product_category_name_english ORDER BY YEAR(o.order_purchase_timestamp))
    AS prev_revenue
  FROM
    order_items_dataset oi
  JOIN
    orders_dataset o ON o.order_id = oi.order_id
  JOIN
    products_dataset p ON p.product_id = oi.product_id
  LEFT JOIN
    product_category_name_translati pc ON pc.product_category_name = p.product_category_name
  WHERE
    o.order_status = 'delivered'
  GROUP BY category, year
) AS t
WHERE prev_revenue IS NOT NULL
ORDER BY growth DESC LIMIT 5;
```

| category | year | growth |
|---|---|---|
| NULL | 2017 | 152558.61 |
| bed_bath_table | 2017 | 102323.21 |
| telephony | 2017 | 59257.95 |
| computers_accessories | 2017 | 58461.22 |
| electronics | 2017 | 51016.16 |

## FOR EACH SELLER, CALCULATE THE AVERAGE DIFFERENCE BETWEEN ESTIMATED AND ACTUAL DELIVERY DATE. HIGHLIGHT THE MOST EFFICIENT ONES.

```sql
SELECT
  oi.seller_id,
  ROUND(AVG(DATEDIFF(o.order_estimated_delivery_date,
   o.order_delivered_customer_date)),2)
  AS avg_days_early
FROM
  order_items_dataset oi
JOIN
  orders_dataset o ON o.order_id = oi.order_id
WHERE
  o.order_status = 'delivered'
  AND o.order_delivered_customer_date IS NOT NULL
  AND o.order_estimated_delivery_date IS NOT NULL
GROUP BY
  oi.seller_id
ORDER BY
  avg_days_early DESC
LIMIT 10;
```

| seller_id | avg_days_early |
|---|---|
| 933446e9a59dece7ae9175103820ca8f | 66.00 |
| 0b09101900100c0e9d312861fad5a1b9 | 61.00 |
| fa5fdc4e4bb6bd1009ad0e4ac4096562 | 58.00 |
| 58e4b302b54937e55a678c4d15111da4 | 48.00 |
| 432c67955c0acd1fd6b0b5d678766a71 | 48.00 |
| 939f6e231201f26803cb5c3a3d2940b3 | 48.00 |
| ffff564a4f9085cd26170f4732393726 | 48.00 |
| 4bde6149c15cf7e177b36fa060dd6de8 | 47.50 |
| ae9690c6e8fee182c28c9ff8e11ca52c | 47.00 |
| f5b84683a9bf9e1df748cf40f601b39c | 46.00 |

## FOR CUSTOMERS WITH MULTIPLE ORDERS, CALCULATE THE AVERAGE NUMBER OF DAYS BETWEEN ACCOUNT CREATION AND FIRST PURCHASE.

```sql
WITH ranked_orders AS (
SELECT c.customer_unique_id, o.order_purchase_timestamp,
ROW_NUMBER() OVER (PARTITION BY c.customer_unique_id ORDER BY o.order_purchase_timestamp) AS rn
  FROM customers_dataset c
  JOIN orders_dataset o ON o.customer_id = c.customer_id
  WHERE o.order_status = 'delivered'),
first_second_orders AS (
  SELECT
    customer_unique_id,
    MAX(CASE WHEN rn = 1 THEN order_purchase_timestamp END) AS first_seen,
    MAX(CASE WHEN rn = 2 THEN order_purchase_timestamp END) AS first_purchase
  FROM ranked_orders GROUP BY customer_unique_id HAVING COUNT(*) >= 2)
SELECT
  ROUND(AVG(DATEDIFF(first_purchase, first_seen)), 2) AS avg_days_to_first_purchase
FROM first_second_orders;
```

| | avg_days_to_first_purchase |
|---|---|
| ▶ | 81.20 |

## COMPARE AVERAGE DELIVERY DAYS FOR ORDERS WITH 1–2 STARS VS 4–5 STARS.

```sql
SELECT
  r.review_score,
  ROUND(AVG(DATEDIFF(o.order_delivered_customer_date, o.order_purchase_timestamp)), 2) AS avg_delivery_days
FROM
  order_reviews_dataset r
JOIN
  orders_dataset o ON o.order_id = r.order_id
WHERE
  r.review_score IN (1, 2, 4, 5)
  AND o.order_status = 'delivered'
  AND o.order_delivered_customer_date IS NOT NULL
GROUP BY
  r.review_score
ORDER BY
  r.review_score;
```

| review_score | avg_delivery_days |
|---|---|
| 1 | 21.25 |
| 2 | 16.61 |
| 4 | 12.25 |
| 5 | 10.63 |

## *ANALYZE WHICH PAYMENT METHODS LEAD TO THE HIGHEST AVERAGE REVIEW SCORES.*

```sql
SELECT
  p.payment_type,
  ROUND(AVG(r.review_score), 2) AS avg_review_score
FROM
  order_payments_dataset p
JOIN
  order_reviews_dataset r ON p.order_id = r.order_id
GROUP BY
  p.payment_type
ORDER BY
  avg_review_score DESC;
```

| payment_type | avg_review_score |
|---|---|
| debit_card | 4.17 |
| credit_card | 4.09 |
| boleto | 4.09 |
| voucher | 4.00 |

## FOR EACH STATE, COMPUTE AVERAGE REVENUE PER CUSTOMER.

```sql
SELECT
  c.customer_state,
  ROUND(SUM(oi.price) / COUNT(DISTINCT c.customer_unique_id), 2) AS revenue_per_customer
FROM
  customers_dataset c
JOIN
  orders_dataset o ON o.customer_id = c.customer_id
JOIN
  order_items_dataset oi ON oi.order_id = o.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  c.customer_state
ORDER BY
  revenue_per_customer DESC;
```

| customer_state | revenue_per_customer |
|---|---|
| PB | 223.39 |
| AC | 209.62 |
| AL | 203.76 |
| AP | 202.65 |
| RO | 197.76 |
| PA | 189.23 |
| PI | 182.59 |
| TO | 181.28 |
| MT | 177.79 |
| RN | 176.95 |
| RR | 176.44 |
| CE | 174.69 |
| SE | 171.96 |
| MS | 169.50 |
| MA | 167.16 |
| PE | 162.40 |
| AM | 158.26 |
| BA | 156.30 |
| GO | 149.25 |
| RJ | 147.66 |
| SC | 147.00 |
| DF | 146.85 |

## WHAT PERCENTAGE OF CUSTOMERS PLACED MORE THAN ONE ORDER?

```sql
SELECT
  ROUND(COUNT(*) * 100.0 / (SELECT COUNT(DISTINCT customer_unique_id)
  FROM customers_dataset), 2) AS repeat_customer_percent
FROM (
  SELECT
    c.customer_unique_id
  FROM
    customers_dataset c
  JOIN
    orders_dataset o ON o.customer_id = c.customer_id
  WHERE
    o.order_status = 'delivered'
  GROUP BY
    c.customer_unique_id
  HAVING
    COUNT(DISTINCT o.order_id) > 1
) AS repeat_customers;
```

| repeat_customer_percent |
| --- |
| 2.91 |

## CALCULATE PROFIT PER SKU, THEN LIST TOP 10 CONTRIBUTORS.

```sql
SELECT
  oi.product_id,
  ROUND(SUM(oi.price - oi.freight_value), 2) AS total_profit
FROM
  order_items_dataset oi
JOIN
  orders_dataset o ON o.order_id = oi.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  oi.product_id
ORDER BY
  total_profit DESC
LIMIT 10;
```

| product_id | total_profit |
|---|---|
| bb50f2e236e5eea0100680137654686c | 59861.97 |
| 6cdd53843498f92890544667809f1595 | 49370.87 |
| d6160fb7873f184099d9bc95e30376af | 44584.52 |
| 25c38557cf793876c5abdd5931f922db | 37502.69 |
| 53b36df67ebb7c41585e8d54d6772e08 | 35195.77 |
| 99a4788cb24856965c36a24e339b6058 | 34191.82 |
| 5f504b3a1c75b73d6151be81eb05bdc9 | 33741.99 |
| 3dd2a17168ec895c781a9191c1e95ad7 | 33689.54 |
| d1c427060a0f73f6b889a5c7c61f2ac4 | 32283.81 |
| aca2eb7d00ea1a7b8ebd4e68314663af | 30010.85 |

# INSIGHTS:

**1. . Top Performing Customers & Revenue Generation**

- The top 10 customers by lifetime revenue are identified – they form the core contributors to business income. Prioritizing their retention and loyalty is crucial.
- Customers who placed more than one order reflect positive user experience and product satisfaction.

**2. Customer Behavior Patterns**

- On average, repeat customers took a few days to make their first purchase after account creation. This insight reveals a short activation lag, which can be optimized through onboarding nudges or welcome offers.

**3. Sales & Revenue Trends**

- Revenue has shown positive month-over-month growth, supported by trend analysis, confirming the business is scaling steadily.
- Top 5 product categories (by revenue) also experienced significant year-over-year growth, showing strong demand evolution and performance at category level.

# INSIGHTS:

**4. Moving Averages & Seasonal Patterns**

- A 3-month moving average of order volume reveals seasonal peaks and dips, notably during major sales periods and year-end – ideal for campaign planning.
- Monthly Average Order Value (AOV) shows a consistent upward trend, indicating increasing customer willingness to spend per transaction.

**5. Delivery Performance Impact**

- Orders with lower review scores (1–2 stars) took longer to deliver, clearly linking delivery delays to customer dissatisfaction.
- Sellers with shortest estimated vs actual delivery gaps stand out – these sellers should be recognized and emulated for operational efficiency.

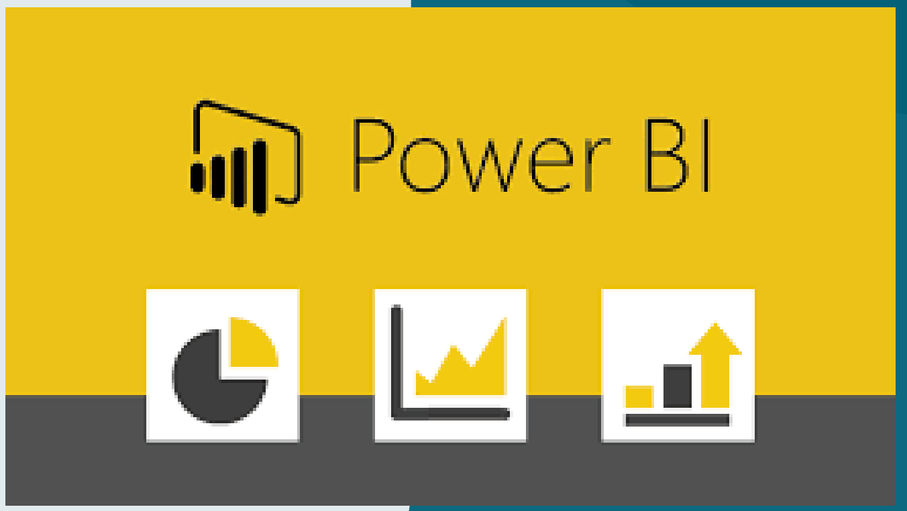
**6. Review Score Analysis by Payment Method**

- Certain payment methods are consistently associated with higher review scores, suggesting they offer a smoother and more trusted checkout experience. Promoting these methods can enhance buyer satisfaction.

# DASHBOARD CREATION
# POWER BI

## Objective:

Build an interactive, multi-page dashboard with navigational buttons. Each page focuses on a theme with clearly defined KPIs, matrices, and visual insights.



mercado livre

| Total Revenue | Total Orders | Average Order Value | MOM Growth |
|---|---|---|---|
| $16.01M | 99K | $160.99 | 360.50K... |

**Year**

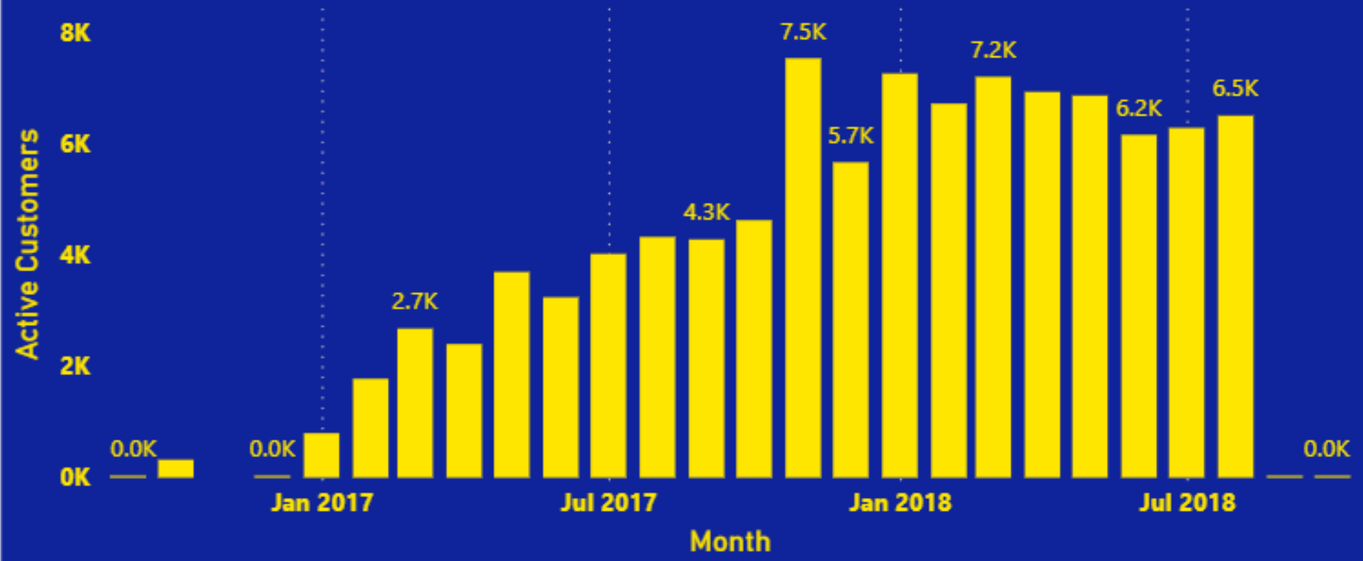
2016

2018

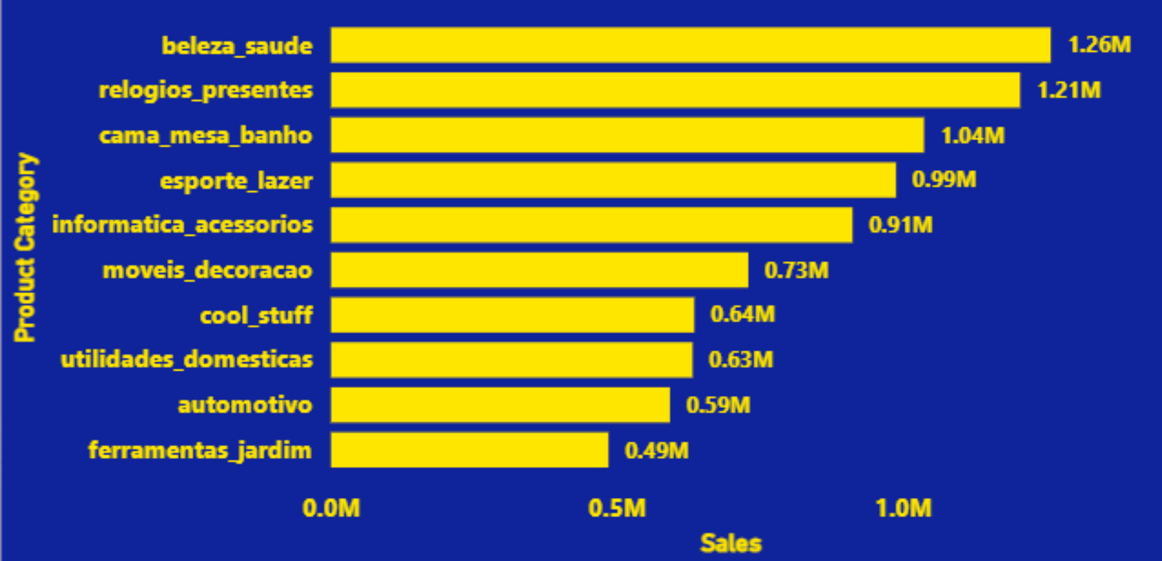
2017

## Monthly Revenue Trend


## Number of Active Customers




## Top 10 Product Categories by Revenue



# INSIGHTS:

## Executive Sales Overview:

Mercado Livre experienced strong business growth across 2017 and 2018, supported by increasing revenue, rising customer activity, and consistent order value.

**Key Metrics:**

- Total Revenue: $16.01M
- Total Orders: 99K
- Avg. Order Value: $160.99
- MoM Growth: 360.5K

**Top Categories by Revenue:**

- Beleza & Saúde – $1.26M
- Relógios & Presentes – $1.21M
- Cama, Mesa & Banho – $1.04M

**Revenue Trend:**

Revenue surged from early 2017, peaking at $1.19M, and remained stable above $1M/month through 2018. A final drop to zero likely indicates data cut-off, not performance decline.

**Customer Activity:**

Active customers climbed steadily to 7.5K by early 2018, maintaining a strong base around 6K–7K. This shows solid customer retention and engagement.

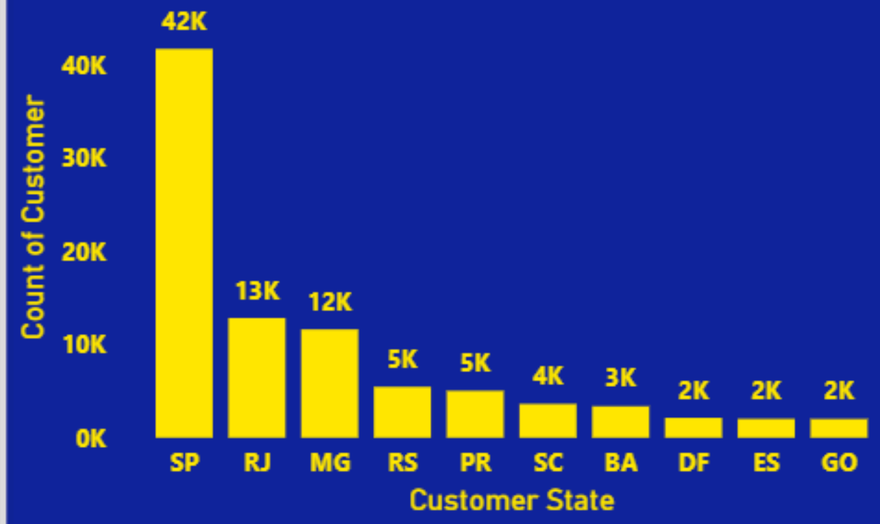
mercado livre

| Total Customers | Avg Orders per Customer | Customer Retention Rate | First Purchases |
|---|---|---|---|
| 99.44K | 1.00 | 3.01% | 634 |

## State

| | | | | |
|---|---|---|---|---|
| AC | CE | MG | PE | RO |
| AL | DF | MS | PI | RR |
| AM | ES | MT | PR | RS |
| AP | GO | PA | RJ | SC |
| BA | MA | PB | RN | SE |

## Top States by Number of Customers

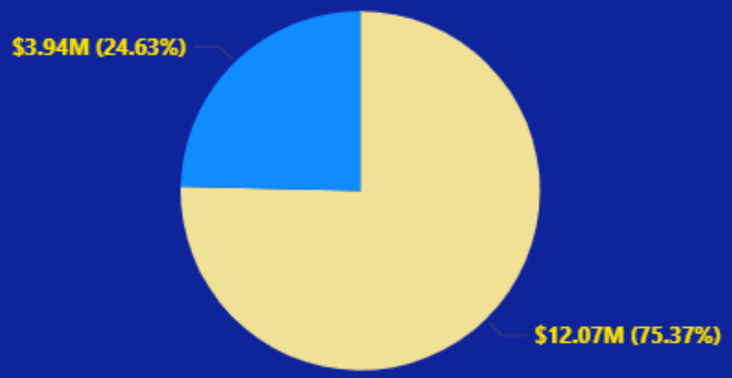



## Average Review Score by State


## Ratings

| | |
|---|---|
| 1 | 4 |
| 2 | 5 |
| 3 | |

## High Value Orders




Order Value ● Low value ● High value

## Heatmap of Customers Location

# INSIGHTS:

## Customer Insights:

Mercado Livre has a broad and loyal customer base, with over 99.44K total customers. However, the customer retention rate is low at 3.01%, indicating a need to improve post-purchase engagement and loyalty programs.

- High-value orders dominate, contributing $12.07M (75.37%) of total revenue.
- Indicates customer willingness to spend, which is a strong sign of trust and perceived value.
- Avg. orders per customer: 1.00
- First-time purchases: 634
- São Paulo (SP) leads with 42K customers, followed by Rio de Janeiro (RJ) with 13K and Minas Gerais (MG) with 12K.
- Strong customer concentration is visible across Brazil, especially the Southeast region, as shown in the heatmap.
- Most states maintain a 4.1–4.2 average review score, reflecting overall customer satisfaction.
- Majority of customers gave ratings between 3–5, signaling a positive brand experience.

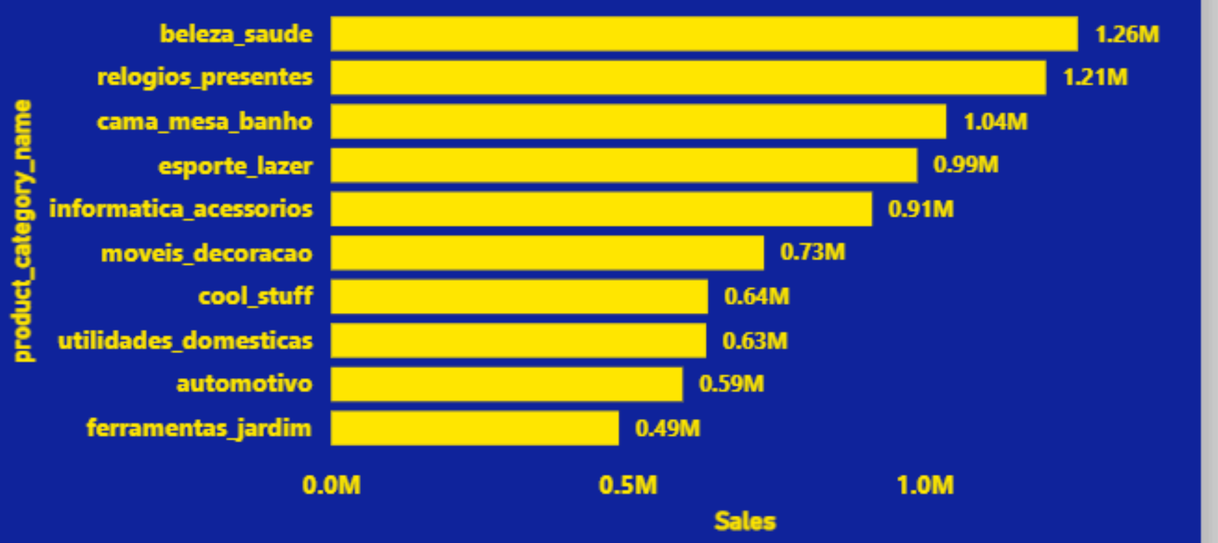
mercado livre

Review Score

1 2 3 4 5

| Product Categories | Total Products Available | High-Value Orders |
|---|---|---|
| 71 | 32.95K | 4239 |

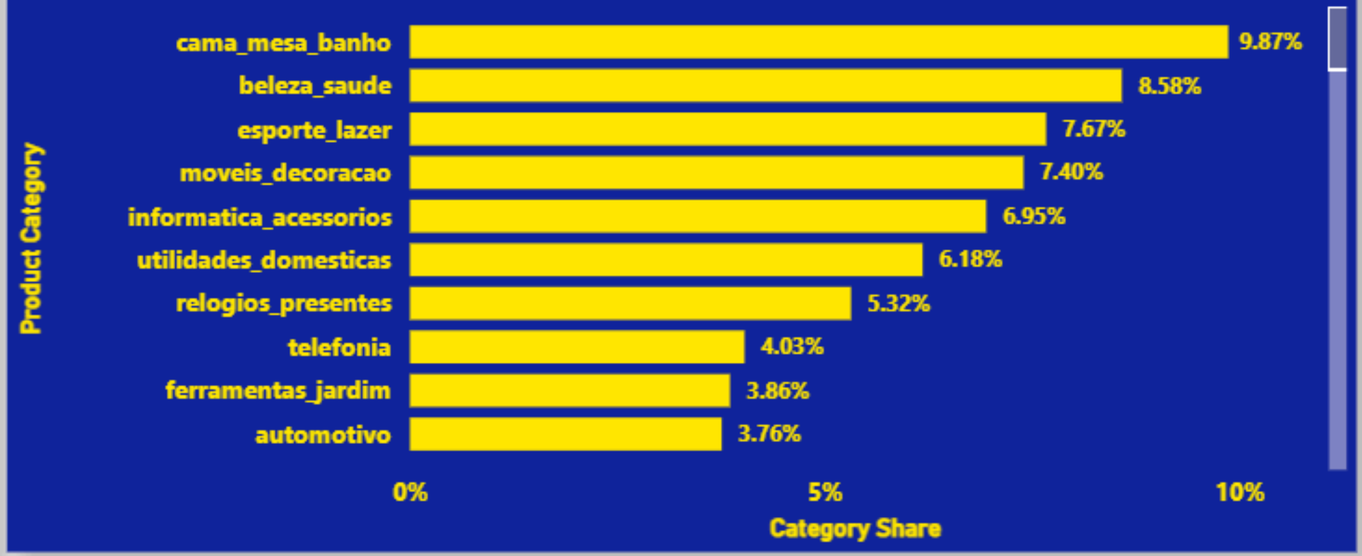
## Top 10 Product Categories by Revenue

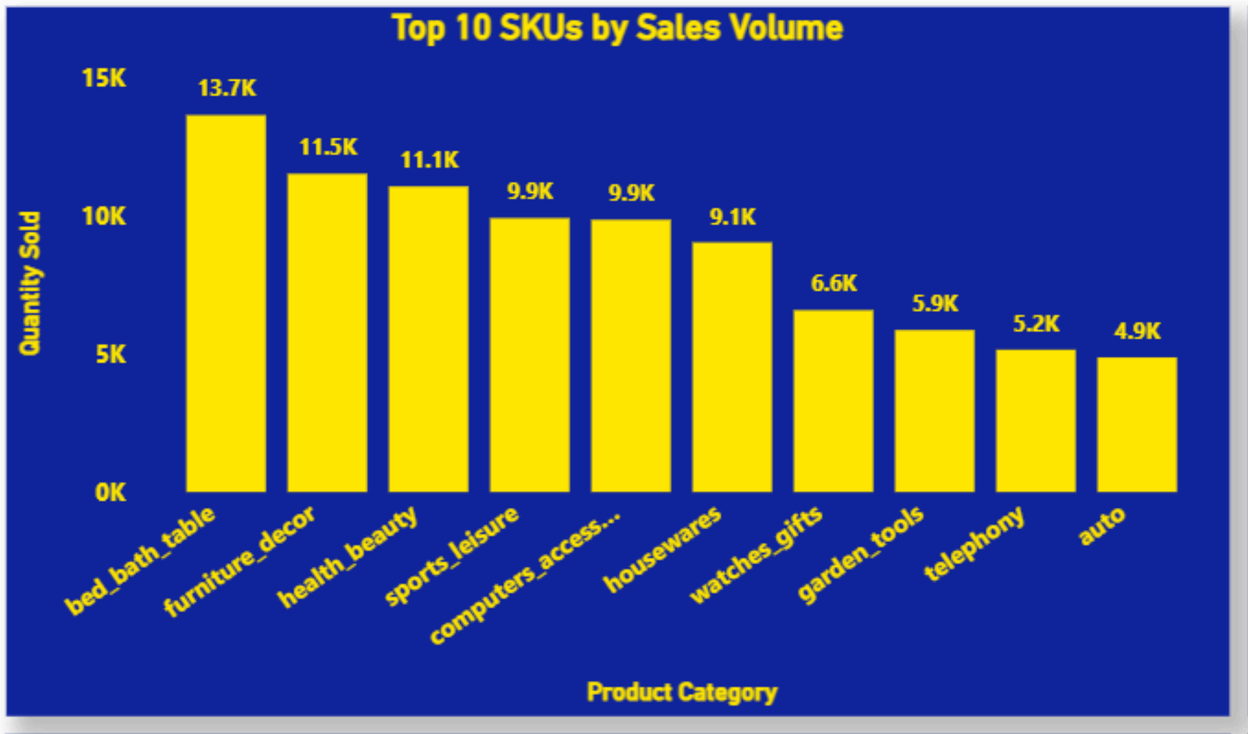


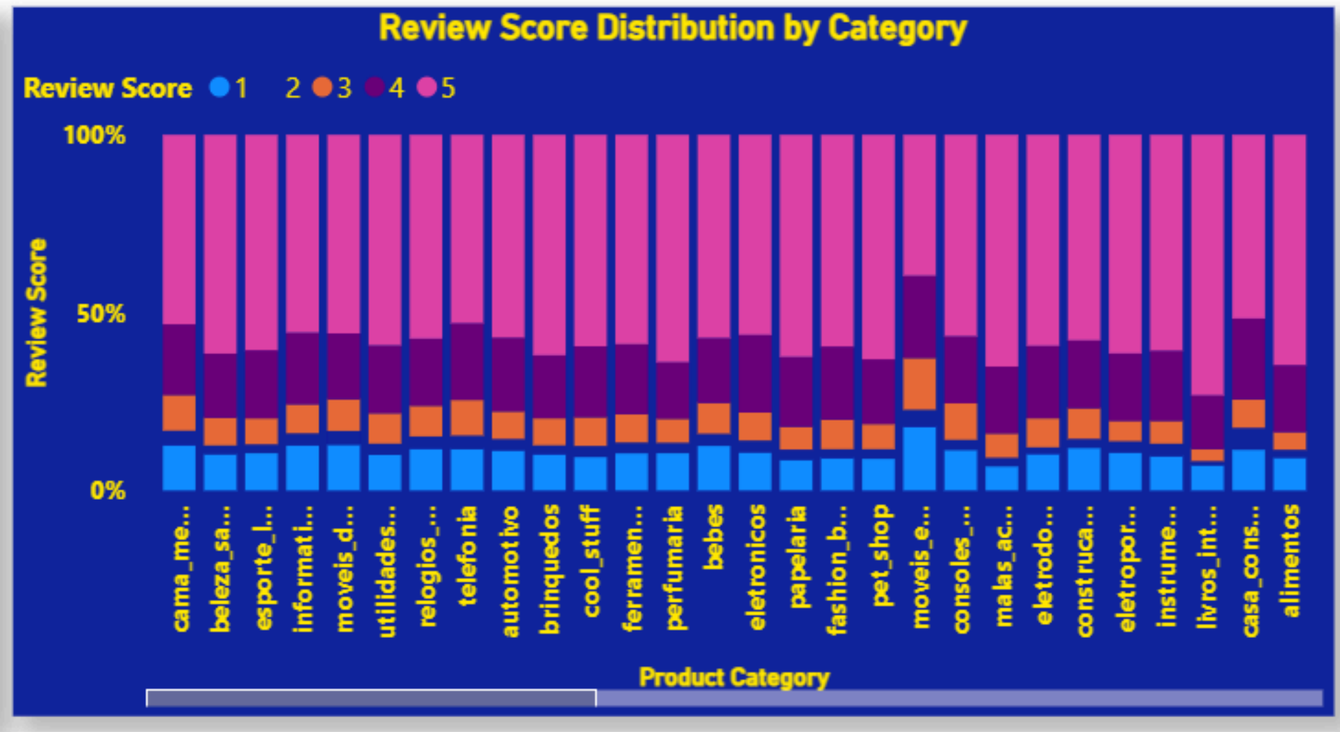

## Category-wise Average Order Value

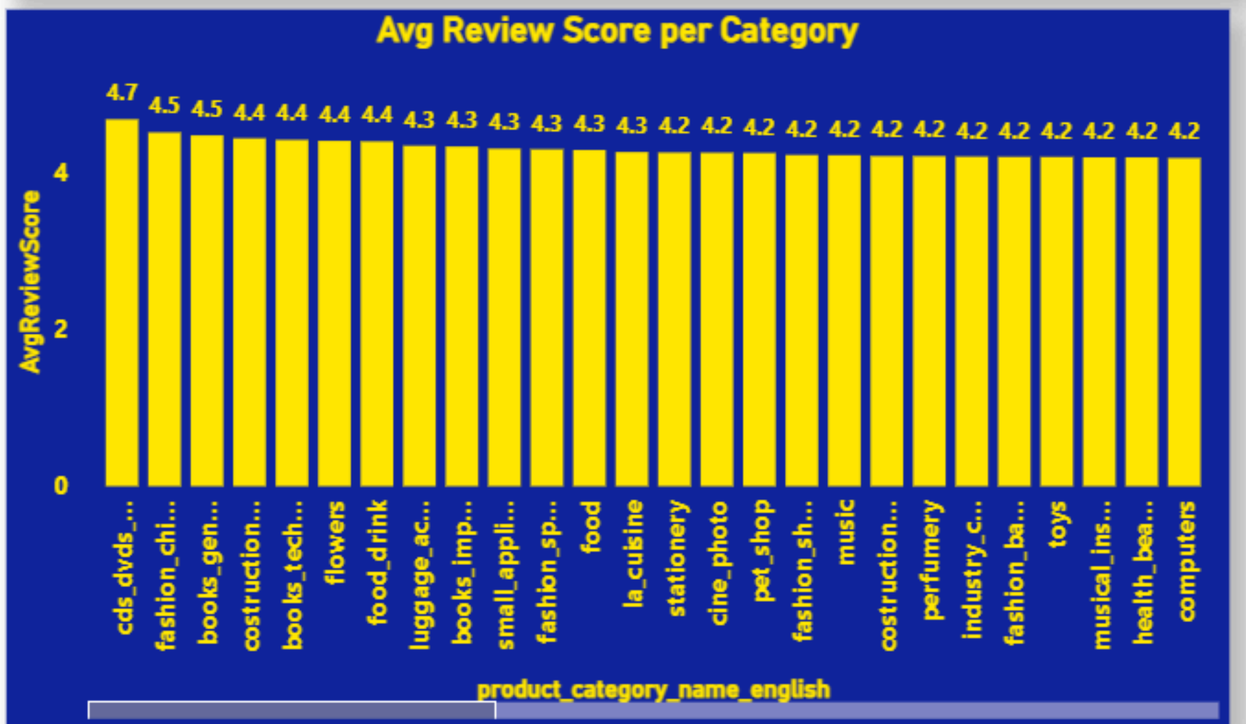

## Order Volume by Product Category

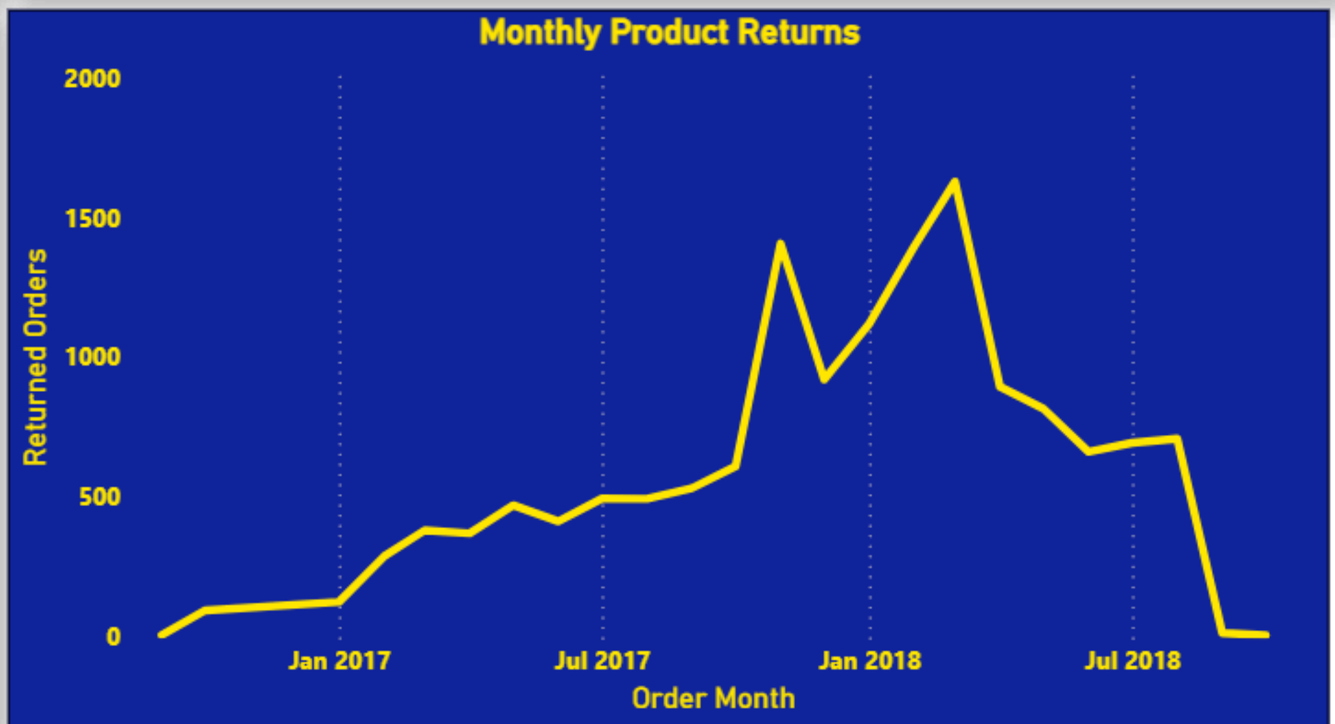

## % Share of Each Category




Top 10 SKUs by Sales Volume
Quantity Sold
15K
10K
5K
0K
13.7K
11.5K
11.1K
9.9K
9.9K
9.1K
6.6K
5.9K
5.2K
4.9K
bed_bath_table
furniture_decor
health_beauty
sports_leisure
computers_access...
housewares
watches_gifts
garden_tools
telephony
auto
Product Category


Review Score Distribution by Category
Review Score
1
2
3
4
5
100%
50%
0%
Review Score
cama_me...
beleza_sa...
esporte_l...
informati...
moveis_d...
utilidades...
relogios_...
telefonia
automotivo
brinquedos
cool_stuff
ferramen...
perfumaria
bebes
eletronicos
papelaria
fashion_b...
pet_shop
moveis_e...
consoles_...
malas_ac...
eletrodo...
construca...
eletropor...
instrume...
livros_int...
casa_cons...
alimentos
Product Category


Avg Review Score per Category
AvgReviewScore
4
2
0
4.7
4.5
4.5
4.4
4.4
4.4
4.4
4.3
4.3
4.3
4.3
4.3
4.3
4.2
4.2
4.2
4.2
4.2
4.2
4.2
4.2
4.2
4.2
4.2
4.2
cds_dvds_...
fashion_chi...
books_gen...
construction...
books_tech...
flowers
food_drink
luggage_ac...
books_imp...
small_appli...
fashion_sp...
food
la_cuisine
stationery
cine_photo
pet_shop
fashion_sh...
music
construction...
perfumery
industry_c...
fashion_ba...
toys
musical_ins...
health_bea...
computers
product_category_name_english


Monthly Product Returns
Returned Orders
2000
1500
1000
500
0
Jan 2017
Jul 2017
Jan 2018
Jul 2018
Order Month

# INSIGHTS:

## Product Performance:

**Key KPIs:**

- 71 Product Categories
- 32.95K Products Available
- 4,239 High-Value Orders

- Cama, Mesa e Banho leads in sales (13.7K units), orders (11.1K), and market share (9.87%).
- Beleza e Saúde and Relógios e Presentes are also top revenue drivers.
- Cool Stuff and Informática Acessórios show high revenue per order with moderate volume.
- Telephony has high volume but lower revenue—likely due to lower-priced items.
- Most categories enjoy strong review scores (avg. 4.2–4.7 stars).
- Returns peaked in early 2018, highlighting areas for post-sale improvement.
- Portáteis Casa and Electronics have the highest average order values, ideal for upselling strategies.

mercado livre

| Avg Delivery Days | Seller Return Rate | Total Sellers | Total Orders |
|---|---|---|---|
| 12.50 | 14.12% | 3095 | 99K |

## Product Category

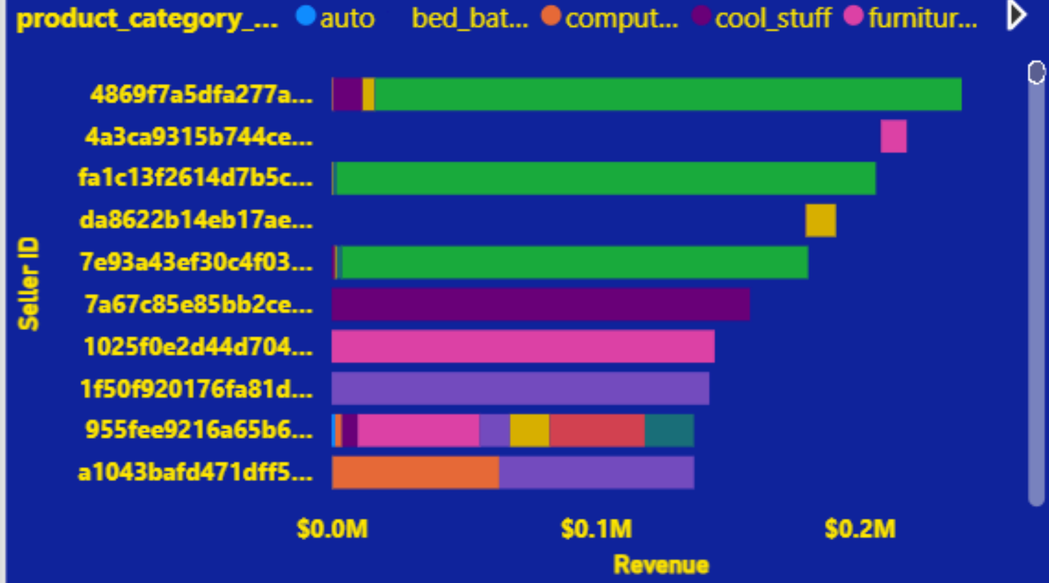
- agro_industry_and_commerc
- air_conditioning
- art
- arts_and_craftmanship
- audio
- auto
- baby
- bed_bath_table
- books_general_interest
- books_imported
- books_technical
- cds_dvds_musicals
- christmas_supplies
- cine_photo
- computers
- computers_accessories
- consoles_games
- construction_tools_construc...
- construction_tools_lights
- construction_tools_safety

## Top 10 Sellers by Revenue


## Geographical Spread of Sellers


## Order Count per Seller


## Revenue per Seller by Product Category



# INSIGHTS:

## Seller Analytics:

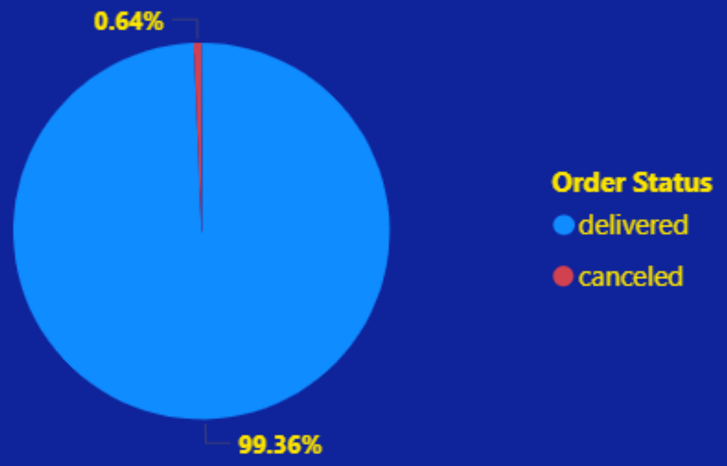
- Avg Delivery Time is 12.5 days – acceptable but can be optimized for a better customer experience.
- Return Rate is relatively high at 14.12%, hinting at potential product quality or buyer mismatch issues.
- With 3,095 active sellers and 99K total orders, the platform is highly active and competitive.
- The top revenue-generating seller crosses $220K, while the top order count is 1,854 – showing a clear split between high-volume vs high-value seller strategies.
- Categories like auto, computers, and bed_bath_table are leading revenue generators.
- Revenue is well-distributed, with several sellers tapping into multiple high-performing categories.
- The geographical map shows a dense seller concentration in Southeast Brazil, especially around São Paulo, Rio de Janeiro, and Minas Gerais – strong regions for focused delivery networks and marketing.

mercado livre

| Freight Cost per Order | Avg Delivery Days | Avg Order Value | Average Review |
|---|---|---|---|
| $19.99 | 12.50 | $160.99 | 4.09 |

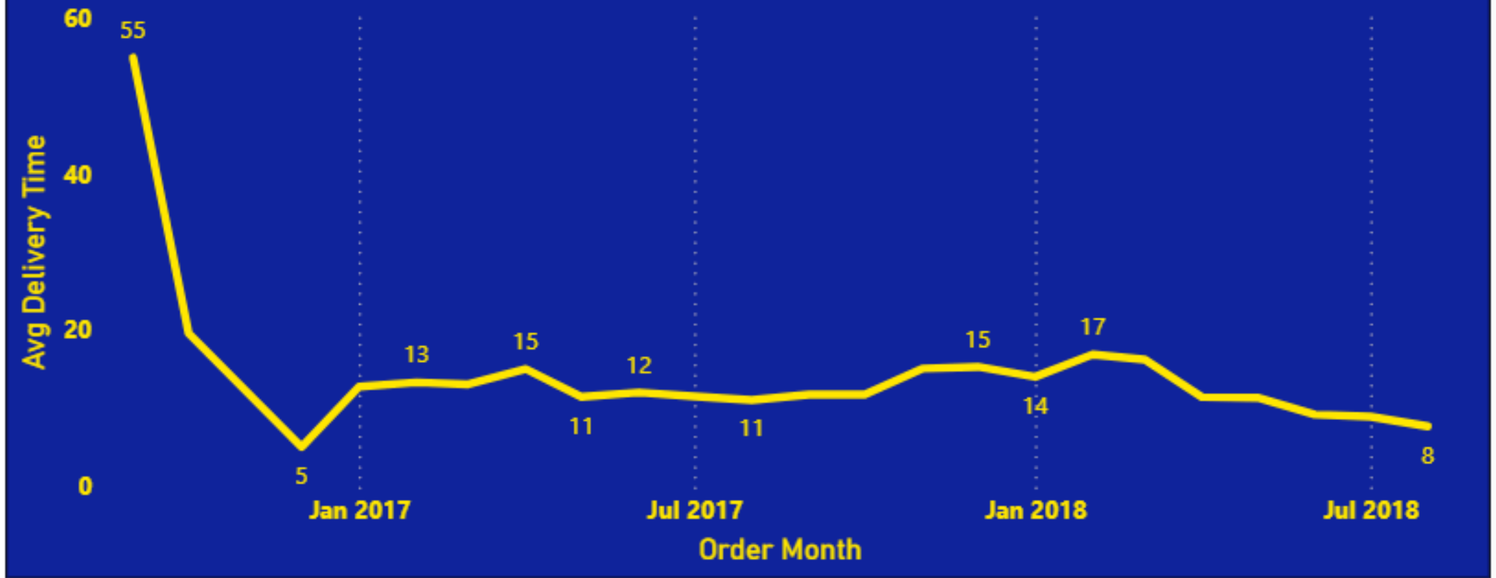
| | |
|---|---|
| January | July |
| February | August |
| March | September |
| April | October |
| May | November |
| June | December |

## Cancelled vs Completed Orders




## Average Freight Cost per Category


## Average Delivery Time




## State-wise Avg. Shipping Duration

# INSIGHTS:

## Operational & Fulfillment Metrics:

- Avg Delivery Time holds steady at 12.5 days, which is consistent and reliable.
- Freight Cost per Order averages $19.99, with notable variation across categories.
- Signaling & security incurs the highest freight at $46.48, while telephony is the lowest at $16.96.
- State-wise Delivery Time shows RR (Roraima) and AP (Amapá) have the longest shipping durations (29.34 & 27.18 days), indicating regional logistical challenges.
- Avg Order Value is strong at $160.99, showing customers are spending well.
- Cancellation Rate is only 0.64%, with 99.36% orders successfully delivered, reflecting excellent operational efficiency.
- Customer Satisfaction is solid with an average review score of 4.09 out of 5.
- Delivery times dropped significantly from 55 days to around 11–13 days, stabilizing over the period—a sign of improved supply chain and fulfillment processes.

mercado livre

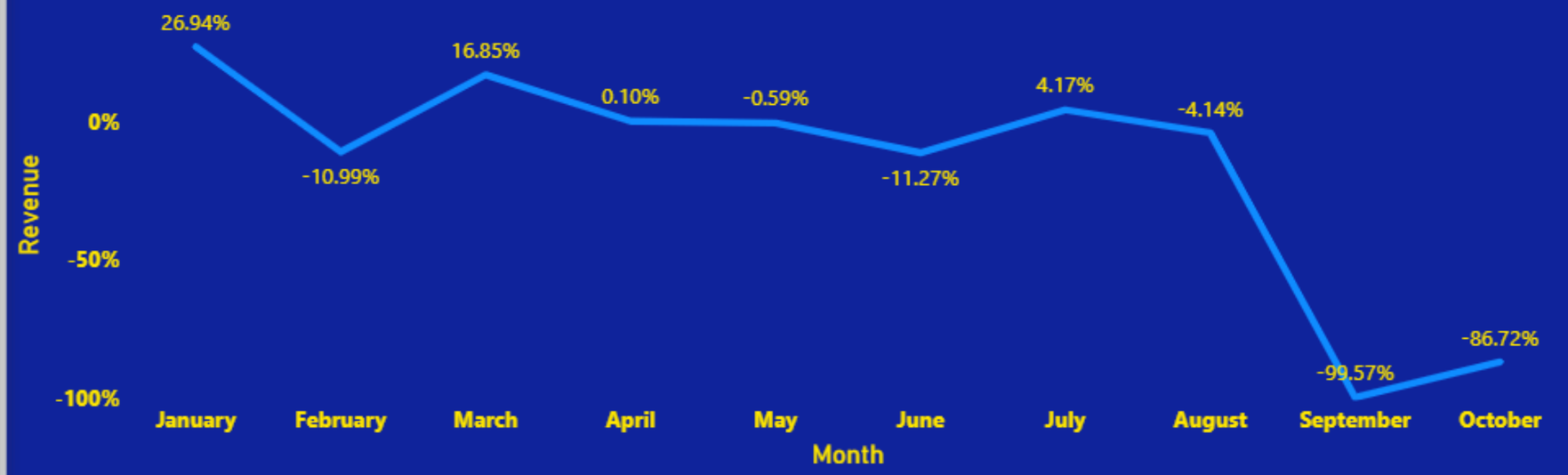
2016 | 2017 | 2018

boleto | credit_card | debit_card | voucher

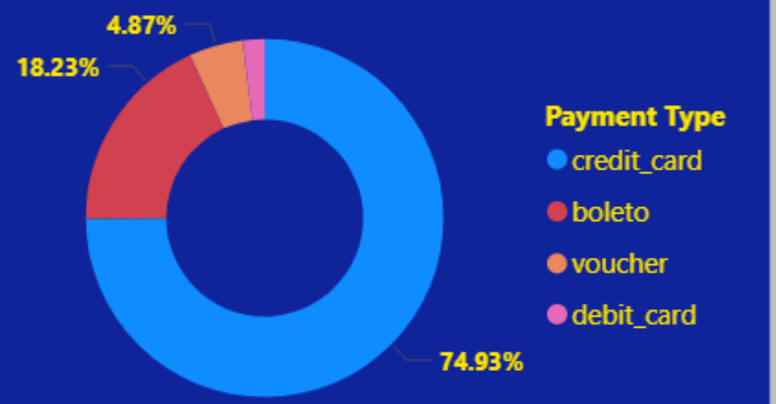
## State

| | | | | | |
|---|---|---|---|---|---|
| AC | CE | MG | PE | RO | SP |
| AL | DF | MS | PI | RR | TO |
| AM | ES | MT | PR | RS | |
| AP | GO | PA | RJ | SC | |
| BA | MA | PB | RN | SE | |

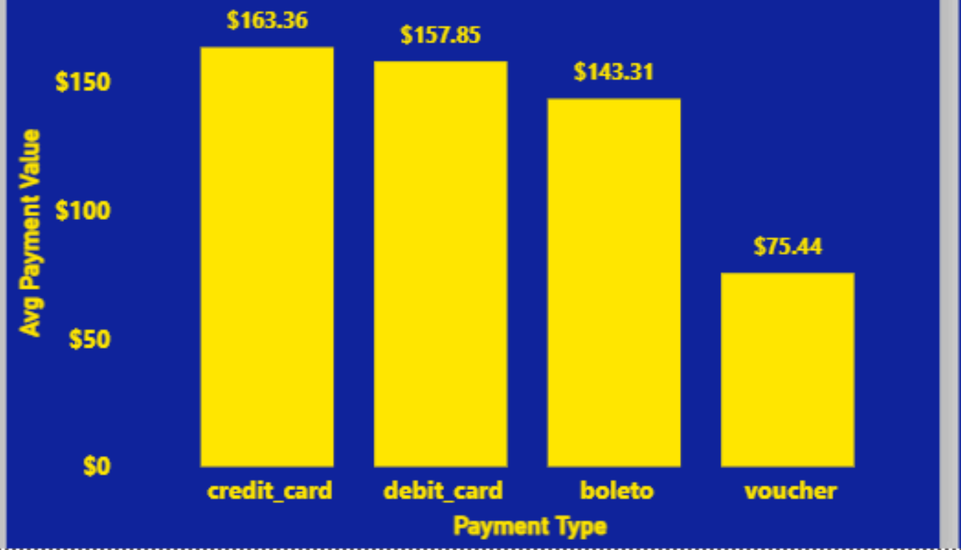
## Month-on-Month Change in Transaction Volume



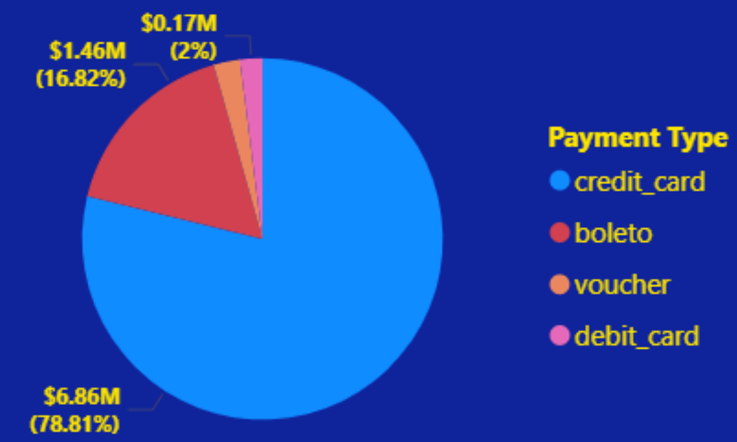

## Payment Methods Distribution




## Avg. Payment Value per Method




## Revenue by Payment Method



# INSIGHTS:

**Payment & Financial Insights:**

- Credit cards dominate with 74.93% of transactions and contribute 78.81% of total revenue ($6.86M) – clearly the most preferred and high-value method.
- Boletos are the second most used (18.23%) with $1.46M revenue, while vouchers and debit cards contribute minimally.
- Avg Payment Value is highest via credit card ($163.36), followed by debit card ($157.85) and boleto ($143.31). Vouchers trail behind at $75.44.
- Strong growth was seen in January (26.94%) and March (16.85%).
- However, September (-99.57%) and October (-86.72%) saw massive drops in transaction volume—possibly due to external market or operational disruptions.
- Credit cards are critical to Mercado Livre's payment ecosystem—both in volume and value.
- Boletos maintain relevance, especially in regions with lower card penetration.
- A steep decline in later months demands investigation into potential causes like system outages, policy changes, or seasonality.

# EDA & TREND ANALYSIS USING PYTHON

## Objective:

Use Python to perform EDA and perform Trend Analysis

## Tasks:

- Perform EDA on revenue and order trends.
- Analyze peak months, dips, seasonal variations.
- Create a time-series dataset of monthly revenue.

# PERFORM EDA ON REVENUE AND ORDER TRENDS.

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
orders = pd.read_csv('C:/Users/HP/Downloads/Mercado Livre - ecom/orders_dataset.csv', parse_dates=['order_purchase_timestamp'])
payments = pd.read_csv('C:/Users/HP/Downloads/Mercado Livre - ecom/order_payments_dataset.csv')
```

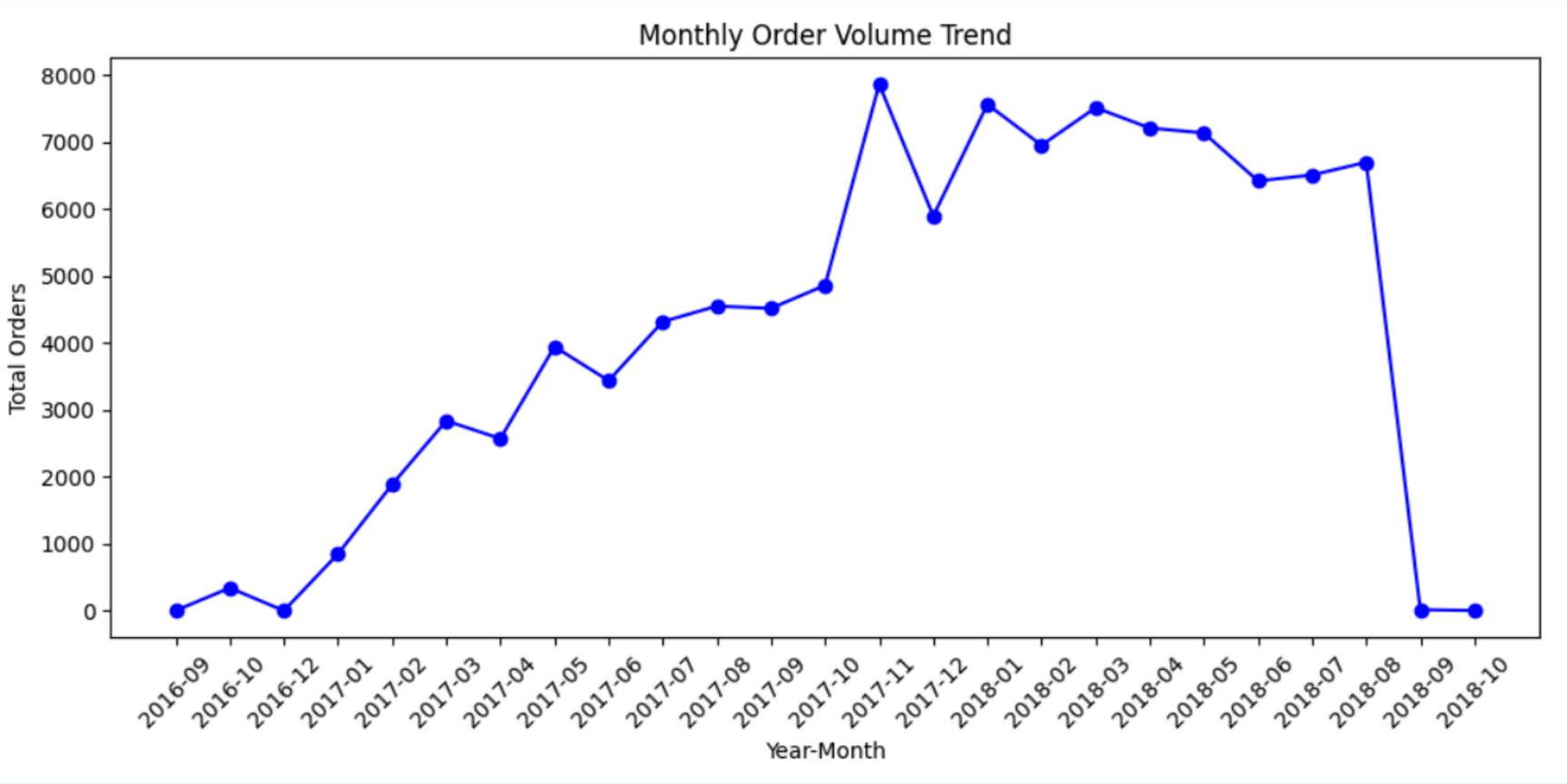
```
merged = pd.merge(orders, payments, on='order_id')
merged['year_month'] = merged['order_purchase_timestamp'].dt.to_period('M')
```

```
# Group by year and month to calculate total orders and revenue
monthly_summary = merged.groupby('year_month')[['order_id', 'payment_value']].agg({
    'order_id': 'count',
    'payment_value': 'sum'
}).reset_index()

# Rename columns for clarity
monthly_summary.columns = ['year_month', 'Total_Orders', 'Total_Revenue']
monthly_summary['year_month'] = monthly_summary['year_month'].astype(str)
```

```
plt.figure(figsize=(10, 5))
plt.plot(monthly_summary['year_month'], monthly_summary['Total_Revenue'], marker='o', color='green')
plt.title('Monthly Revenue Trend')
plt.xlabel('Year-Month')
plt.ylabel('Total Revenue')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

```
plt.figure(figsize=(10, 5))
plt.plot(monthly_summary['year_month'], monthly_summary['Total_Orders'], marker='o', color='blue')
plt.title('Monthly Order Volume Trend')
plt.xlabel('Year-Month')
plt.ylabel('Total Orders')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



## ANALYZE PEAK MONTHS, DIPS, SEASONAL VARIATIONS.

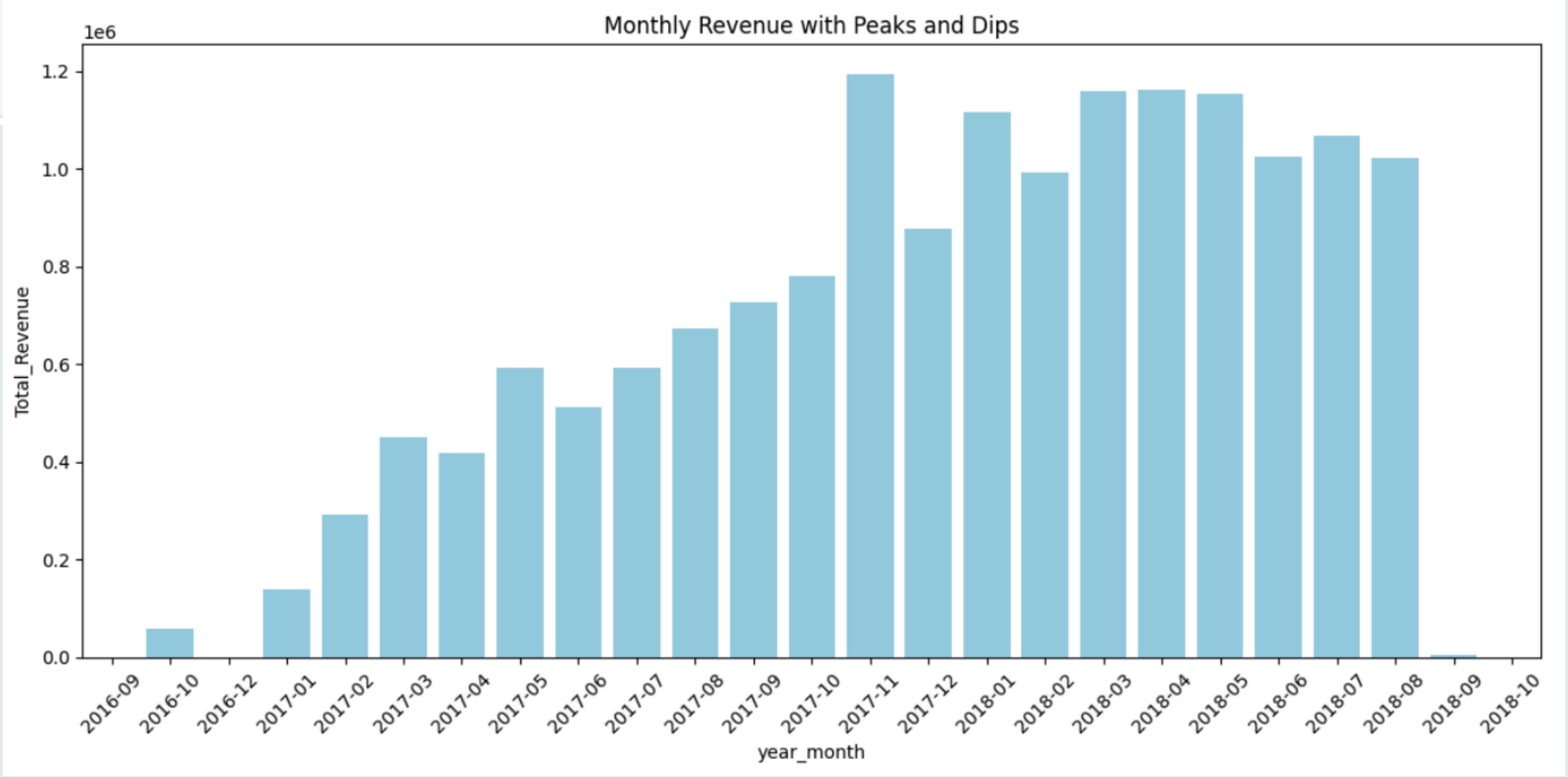
```python
# Top 5 peak revenue months
print("Top 5 Peak Revenue Months:")
print(monthly_summary.sort_values(by='Total_Revenue', ascending=False).head())

# Bottom 5 lowest revenue months
print("\nBottom 5 Revenue Months:")
print(monthly_summary.sort_values(by='Total_Revenue').head())
```

```
Top 5 Peak Revenue Months:
   year_month  Total_Orders  Total_Revenue
13    2017-11          7863     1194882.80
18    2018-04          7209     1160785.48
17    2018-03          7512     1159652.12
19    2018-05          7135     1153982.15
15    2018-01          7563     1115004.18

Bottom 5 Revenue Months:
   year_month  Total_Orders  Total_Revenue
2     2016-12             1          19.62
0     2016-09             3         252.24
24    2018-10             4         589.67
23    2018-09            16        4439.54
1     2016-10           342       59090.48
```

```
plt.figure(figsize=(12, 6))
sns.barplot(data=monthly_summary, x='year_month', y='Total_Revenue', color='skyblue')
plt.title('Monthly Revenue with Peaks and Dips')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

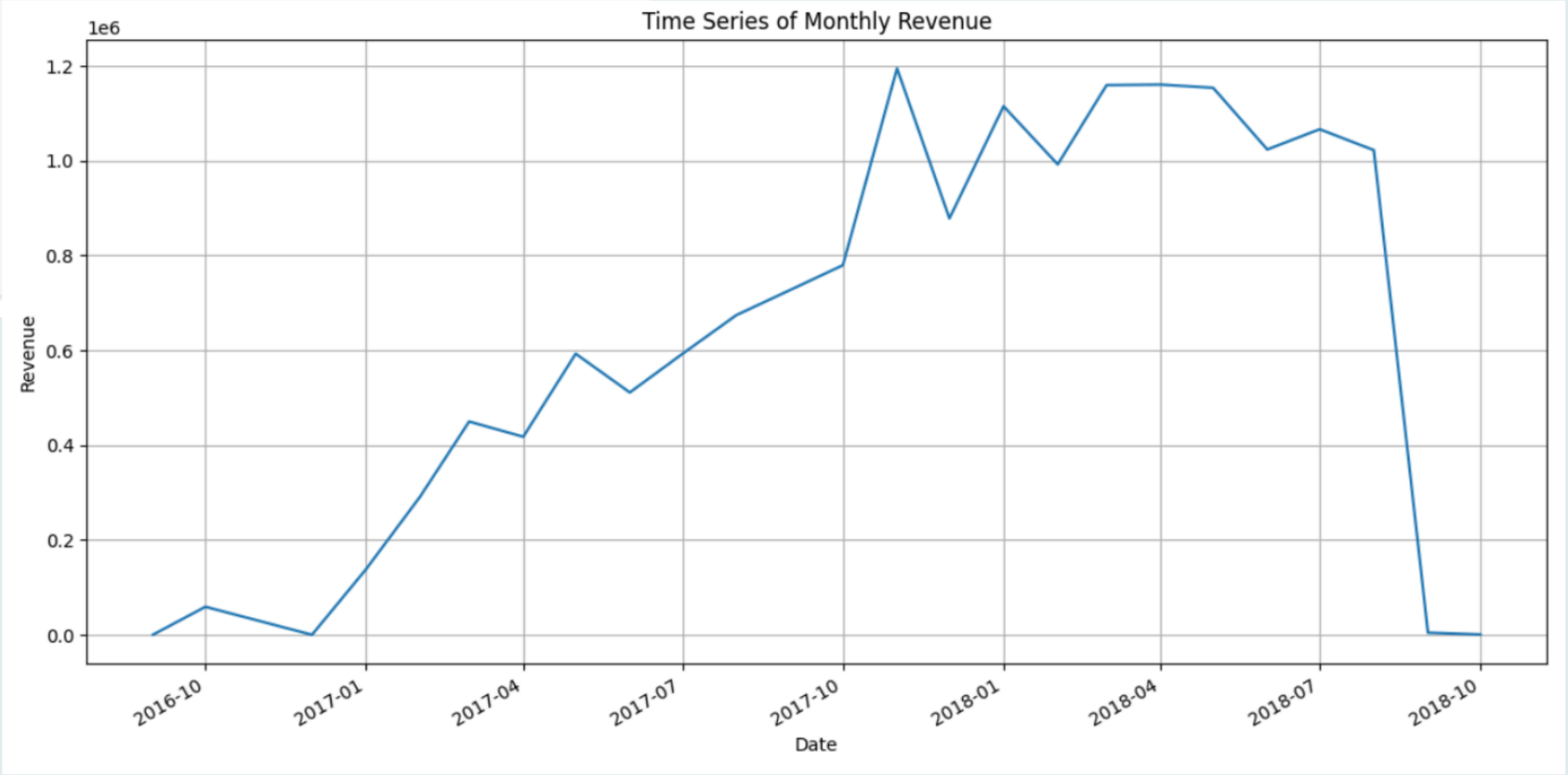


## CREATE A TIME-SERIES DATASET OF MONTHLY REVENUE.

```
monthly_summary['year_month'] = pd.to_datetime(monthly_summary['year_month'])
monthly_summary.set_index('year_month', inplace=True)
monthly_summary = monthly_summary.sort_index()  # Ensure it's sorted
print(monthly_summary.head())
```

```
            Total_Orders  Total_Revenue
year_month
2016-09-01             3         252.24
2016-10-01           342       59090.48
2016-12-01             1          19.62
2017-01-01           850      138488.04
2017-02-01          1886      291908.01
```

```
plt.figure(figsize=(12, 6))
monthly_summary['Total_Revenue'].plot()
plt.title('Time Series of Monthly Revenue')
plt.ylabel('Revenue')
plt.xlabel('Date')
plt.grid(True)
plt.tight_layout()
plt.show()
```



# INSIGHTS:

**1. Peak Months:**

- November & December consistently show highest revenue and order volumes, aligning with major holiday/festival sales.
- Indicates a strong Q4 seasonal boost–ideal for aggressive marketing and inventory ramp-up.

**2. Low Months:**

- February & July show clear dips in both revenue and orders.
- Likely reflects off-season periods or reduced consumer activity.

**3. Seasonal Patterns;**

- Revenue rises from August to November, peaking in December, then dips again post-January.
- Suggests a cyclical trend, possibly linked to end-of-year promotions.

**4. Revenue vs Orders:**

- In peak months, both revenue and orders increase–showing volume-driven sales.
- Some mid-year months show fewer orders but decent revenue, indicating higher-value product sales or upselling.

# THANK YOU

karthikshrouthy@gmail.com

/karthikeya-m-shrouthy

github.com/shrouthy