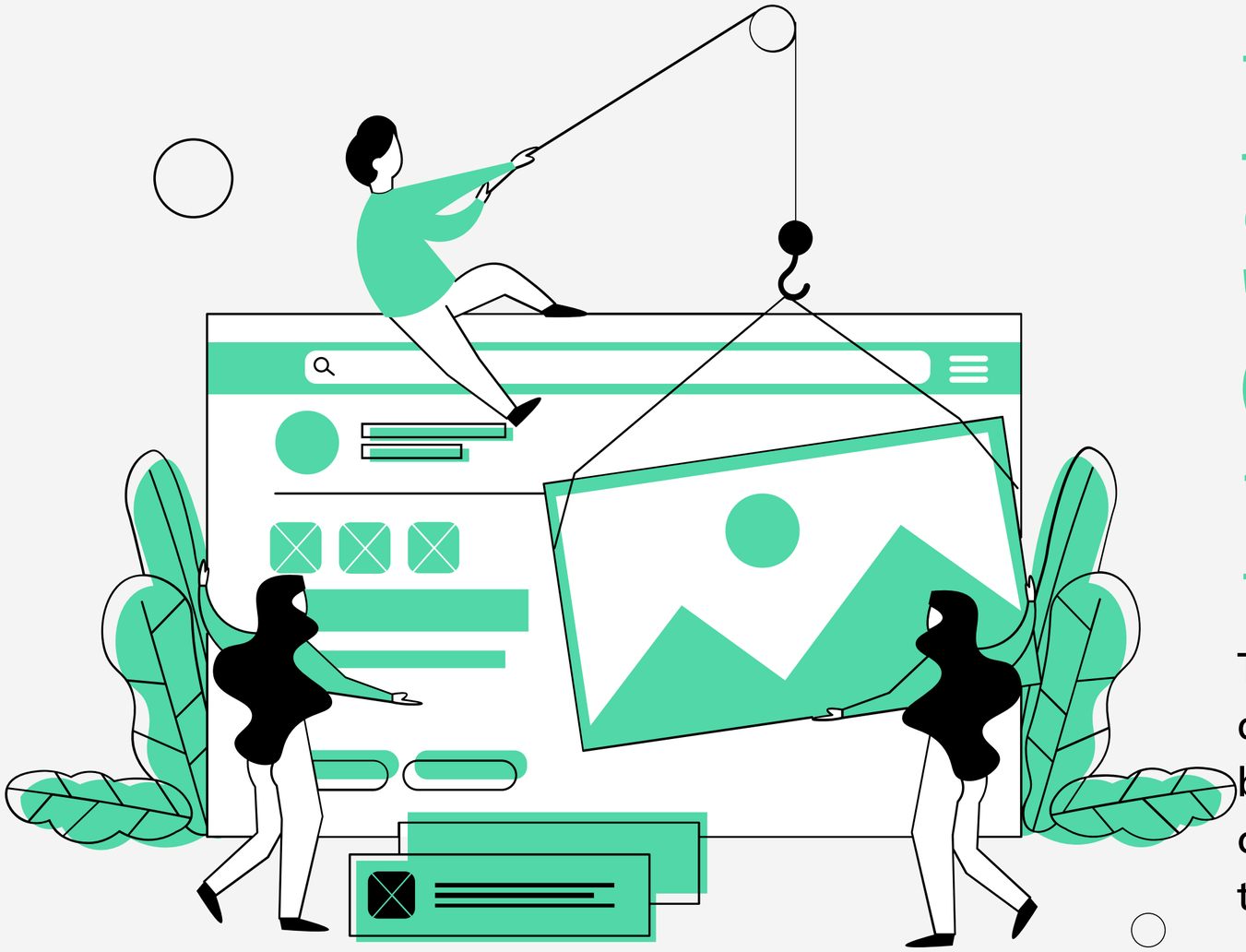
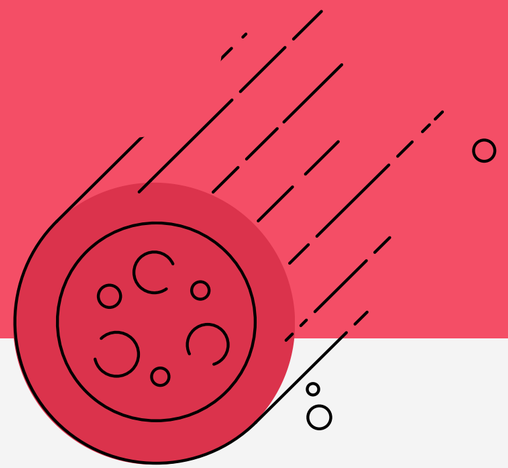


# E-COMMERCE SALES CAPSTONE PROJECT

To perform deep data analysis using advanced MySQL queries and extract actionable insights to answer business questions related to sales performance, customer behavior, operational efficiency, and time trends.

**Presented By Jagruti Jadhav**

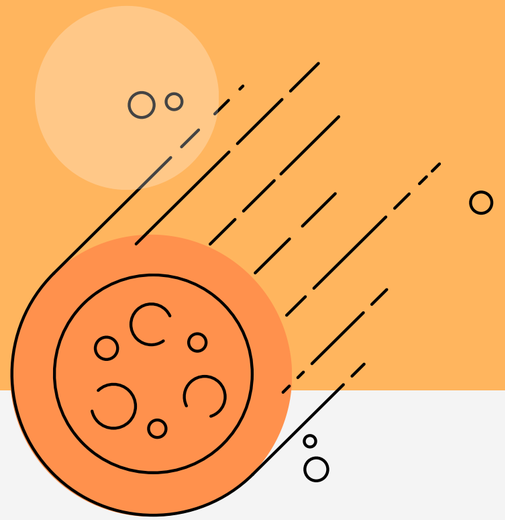


# Problem Statement for SQL

1. Write a query to calculate total sales revenue per category, sub-category, and region.

2. Identify the top 5 best-selling products by both sales revenue and quantity sold.

3. Find the most loyal customers by calculating their purchase frequency and total spend



# Problem Statement for SQL

4 Identify customers with the highest average order value (AOV).

5 Analyze delivery performance by calculating the average delivery time by region.

6 Identify regions or products with the highest canceled rates.

# Problem Statement for SQL

7 Write a query to find the monthly sales trend for the last two years.

8 Analyze the seasonality of sales to identify peak months.

9 Use window functions to rank products based on their sales within each category

10 Calculate month-to-date (MTD) and year-to-date (YTD) sales metrics.

# Sales Performance Analysis:

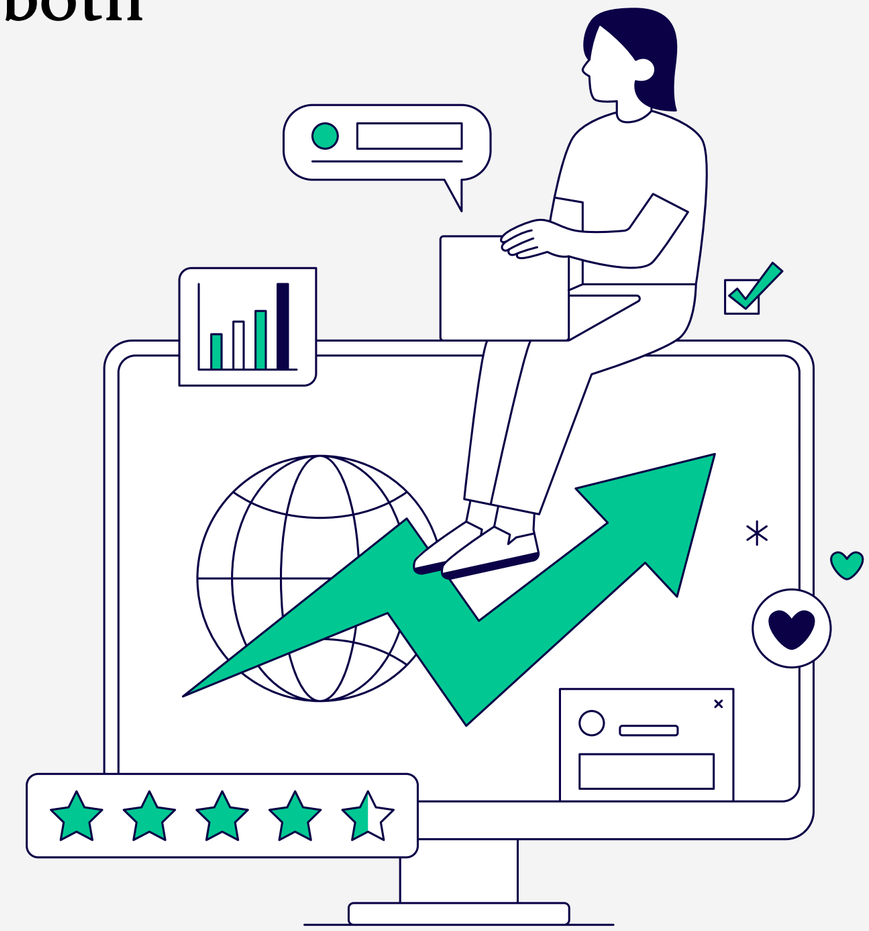
Write a query to calculate total sales revenue per category, sub-category, and region.

```sql
SELECT
    p.product_category AS category,
    p.product_id as sub_category,
    s.seller_state AS region,
    ROUND(SUM(oi.price), 2) AS revenue
FROM
    order_items AS oi
        JOIN
    products AS p ON oi.product_id = p.product_id
        JOIN
    orders AS o ON o.order_id = oi.order_id
        JOIN
    sellers AS s ON s.seller_id = oi.seller_id
WHERE
    o.order_status != 'canceled'
GROUP BY p.product_category, p.product_id, s.seller_state
ORDER BY revenue DESC;
```

# Sales Performance Analysis:

Identify the top 5 best-selling products by both sales revenue and quantity sold.

```sql
SELECT
    p.product_id,
    ROUND(SUM(oi.price), 2) AS revenue,
    SUM(p.product_photos_qty) AS total_quantity_sold
FROM
    order_items AS oi
        JOIN
    products AS p ON oi.product_id = p.product_id
        JOIN
    orders AS o ON o.order_id = oi.order_id
WHERE
    o.order_status != 'canceled'
GROUP BY p.product_id
ORDER BY revenue DESC , total_quantity_sold DESC
LIMIT 5;
```



# Customer Insights

Identify customers with the highest average order value (AOV).

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

# Operational Efficiency (By Region):

Analyze delivery performance by calculating the average delivery time by region

```sql
SELECT
    s.seller_state AS region,
    ROUND(AVG(DATEDIFF(o.order_estimated_delivery_date,
                       o.order_purchase_timestamp))) AS average_delivery_time
FROM
    order_items AS oi
        JOIN
    orders AS o ON oi.order_id = o.order_id
        JOIN
    sellers AS s ON s.seller_id = oi.seller_id
GROUP BY s.seller_state;
```

# Operational Efficiency (By Region):

## Identify regions or products with the highest canceled rates

```sql
SELECT A.regions, A.total_order, B.total_return, ((B.total_return/A.total_order)*100) as return_rates
FROM
(SELECT s.seller_state AS regions, count(o.order_id) AS total_order
FROM orders AS o
JOIN order_items AS oi
ON o.order_id = oi.order_id
JOIN sellers AS s
ON oi.seller_id = s.seller_id
GROUP BY s.seller_state) AS A
JOIN
(SELECT s.seller_state AS regions, count(o.order_id) AS total_return
FROM orders AS o
JOIN order_items AS oi
ON o.order_id = oi.order_id
JOIN sellers AS s
ON oi.seller_id = s.seller_id
WHERE o.order_status = "canceled" OR o.order_status = "unavailable"
GROUP BY s.seller_state) AS B
ON A.regions = B.regions
ORDER BY return_rates DESC;
```

# Date and Time Analytics:

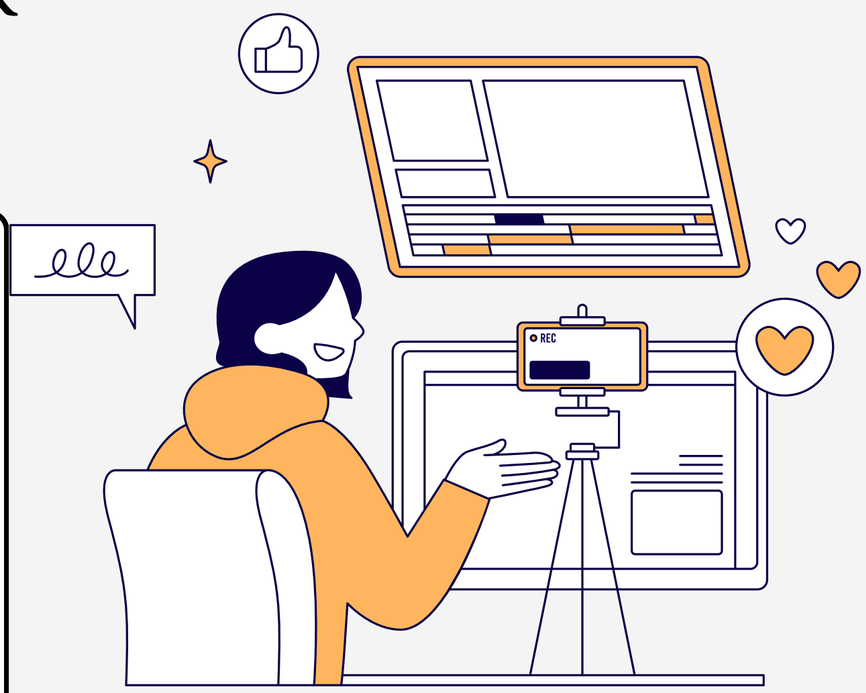
**Write a query to find the monthly sales trend for the last two years.**

```sql
SELECT
    EXTRACT(YEAR FROM o.order_purchase_timestamp) AS year,
    EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
    SUM(oi.price)
FROM
    orders AS o
        JOIN
    order_items AS oi ON o.order_id = oi.order_id
WHERE
    EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2017
        OR EXTRACT(YEAR FROM o.order_purchase_timestamp) = 2018
GROUP BY EXTRACT(YEAR FROM o.order_purchase_timestamp) ,
EXTRACT(MONTH FROM o.order_purchase_timestamp)
ORDER BY year DESC , month DESC;
```

# Date and Time Analytics:

Analyze the seasonality of sales to identify peak months.

```sql
SELECT
    EXTRACT(MONTH FROM o.order_purchase_timestamp) AS month,
    SUM(oi.price) AS total_sales,
    AVG(oi.price) AS avg_sales,
    COUNT(DISTINCT EXTRACT(YEAR FROM o.order_purchase_timestamp)) AS total_year
FROM
    orders AS o
        JOIN
    order_items AS oi ON o.order_id = oi.order_id
GROUP BY EXTRACT(MONTH FROM o.order_purchase_timestamp)
HAVING COUNT(DISTINCT EXTRACT(YEAR FROM o.order_purchase_timestamp)) >= 2
ORDER BY total_sales DESC;
```



# Advanced SQL Queries:

Use window functions to rank products based on their sales within each category.

```sql
WITH a AS (SELECT p.product_category, round(sum(oi.price),2) AS sales
FROM order_items AS oi
JOIN products AS p
ON oi.product_id = p.product_id
GROUP BY p.product_category)

SELECT *,
RANK() OVER(ORDER BY sales DESC) AS rn
FROM a;
```

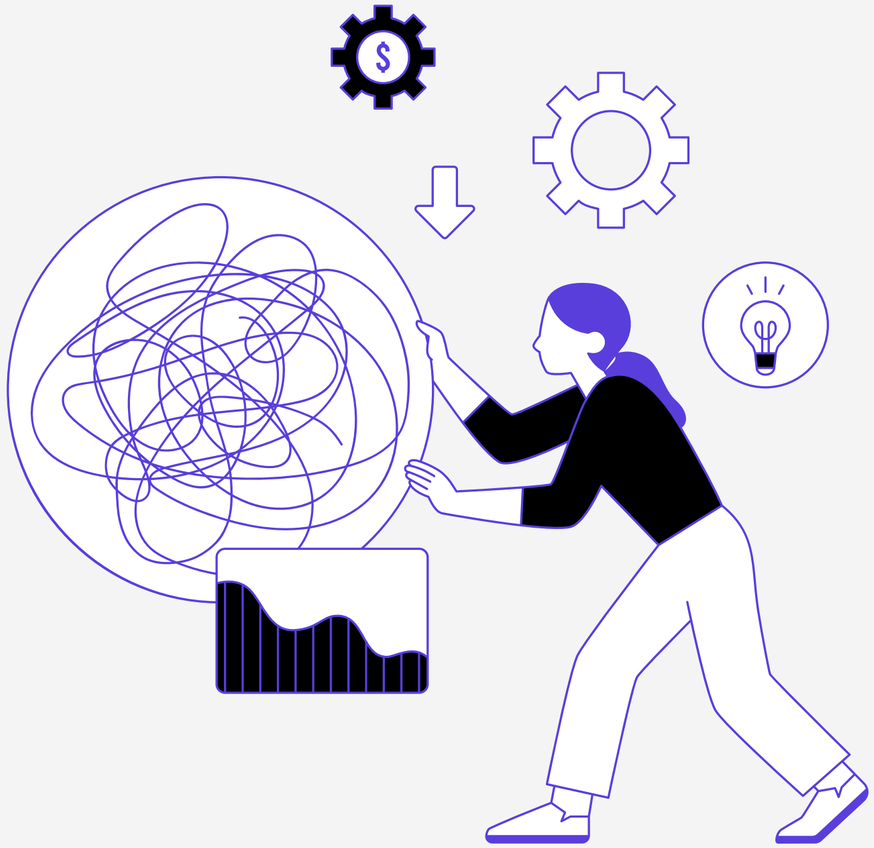
# Advanced SQL Queries

## Calculate month-to-date (MTD) and year-to-date (YTD) sales metrics.

```sql
# Calculate month-to-date (MTD) and year-to-date (YTD) sales metrics.

-- The delivered_customer_date has multiple enteries for a particular date,
-- SO we will make use of Rows parameter with (unbounded precedings and current row) as YTD_Frame

select
    oi.order_id,
    date(o.order_delivered_customer_date) as dates ,
    year(o.order_delivered_customer_date) as years ,
    month(o.order_delivered_customer_date) as months ,
    oi.price,
    sum(oi.price) over(partition by year(o.order_delivered_customer_date)
                       order by date(o.order_delivered_customer_date)) as YTD,
    sum(oi.price) over(partition by year(o.order_delivered_customer_date)
                       order by date(o.order_delivered_customer_date)
    Rows between unbounded preceding and current row) as YTD_Frame,
    sum(oi.price) over(partition by month(o.order_delivered_customer_date)
                       order by date(o.order_delivered_customer_date)) as MTD,
    sum(oi.price) over(partition by month(o.order_delivered_customer_date)
                       order by date(o.order_delivered_customer_date)
    Rows between unbounded preceding and current row) as MTD_Frame
from order_items oi
join orders o
on o.order_id = oi.order_id
where date(o.order_delivered_customer_date) is not null;
```

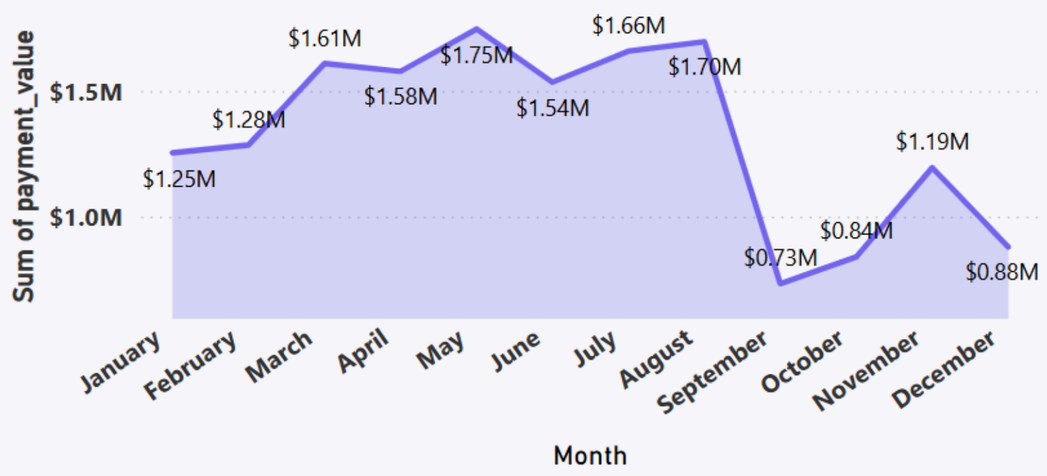


# E-COMMERCE TARGET SALES POWER BI DASHBOARD & INSIGHTS

Presenting the Dashboard for the Target Sales Company

# Sales Performance

## SUM OF PAYMENT VALUE BY MONTH




## TOTAL SALES

$16.01M

Total Sales Revenue

## AVERAGE VALUE PER ORDER

$154.1003...

Average of payment_value

## Region Slicer

- [ ] Central-West
- [ ] North
- [ ] Northeast
- [ ] South

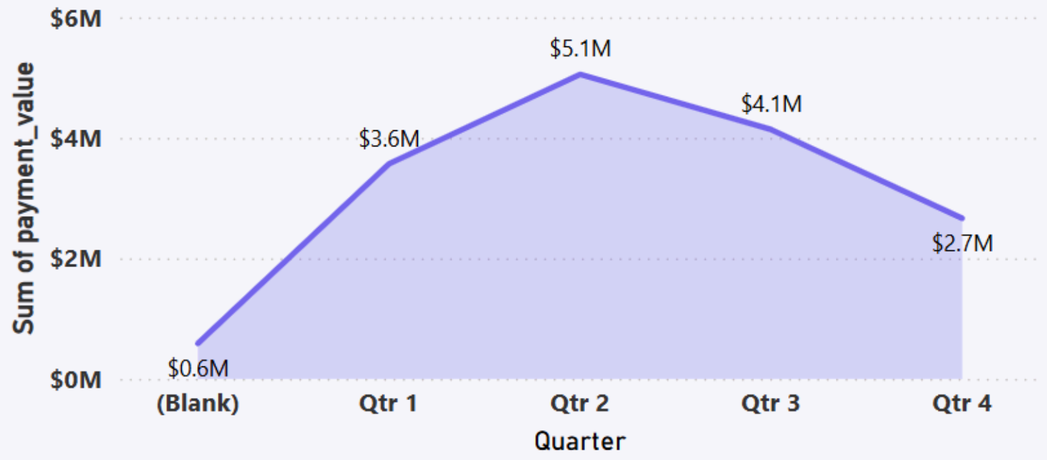
## Year, Quarter, Month, Day

- [ ] 2016
  - [ ] Qtr 1
  - [ ] Qtr 2

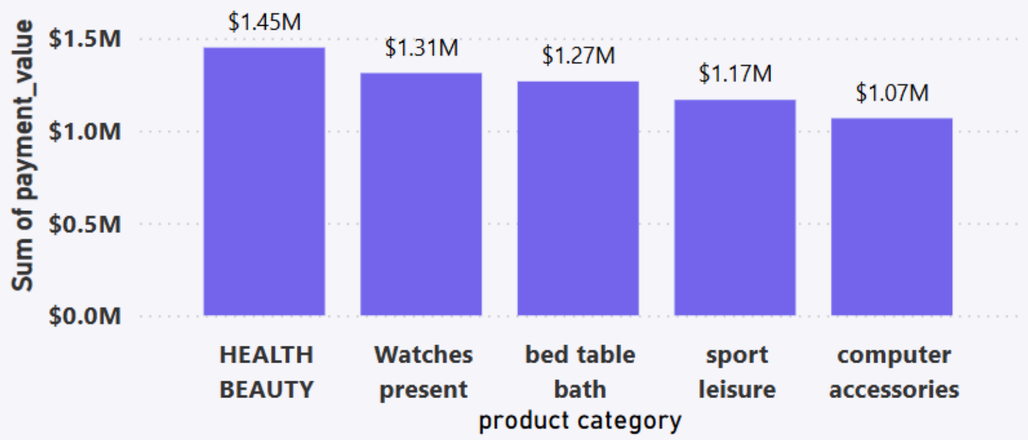
## Product Category Slicer

- [ ]
- [ ] Agro Industria e C...
- [ ] Art
- [ ] Arts and Crafts

## SUM OF PAYMENT VALUE BY QUARTERS

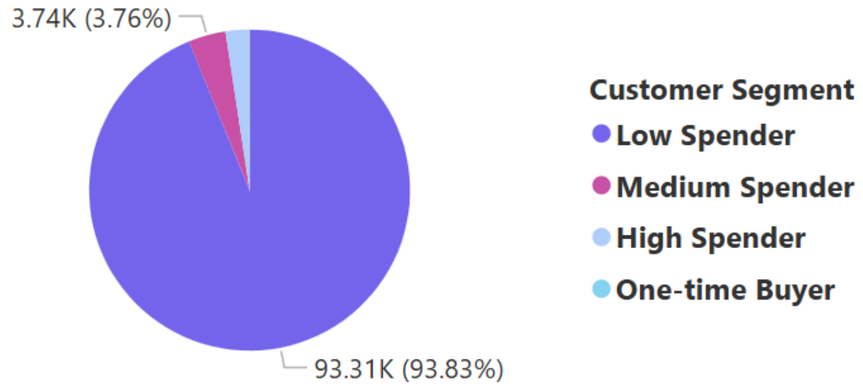



## TOP 5 BEST SELLING PRODUCTS



# Customer Insights

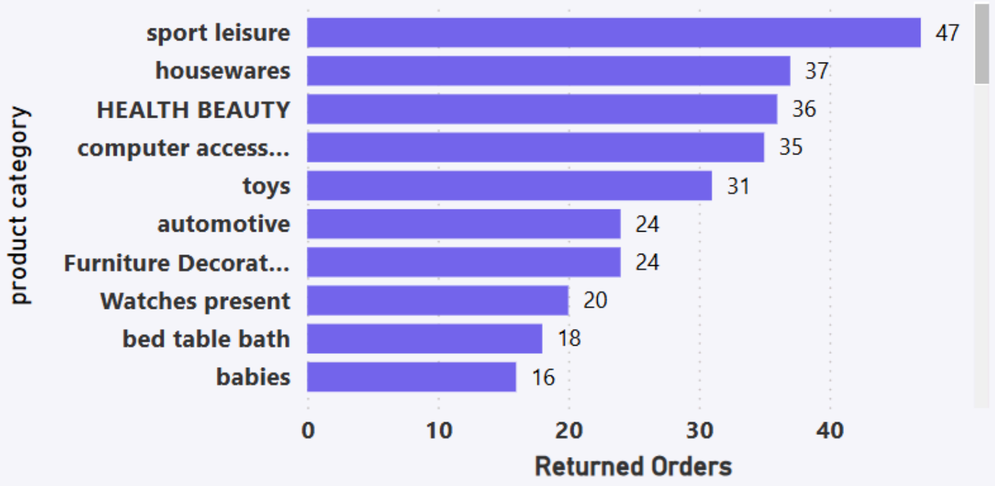
## COUNT OF CUSTOMER ID BY CUSTOMER SEGMENT



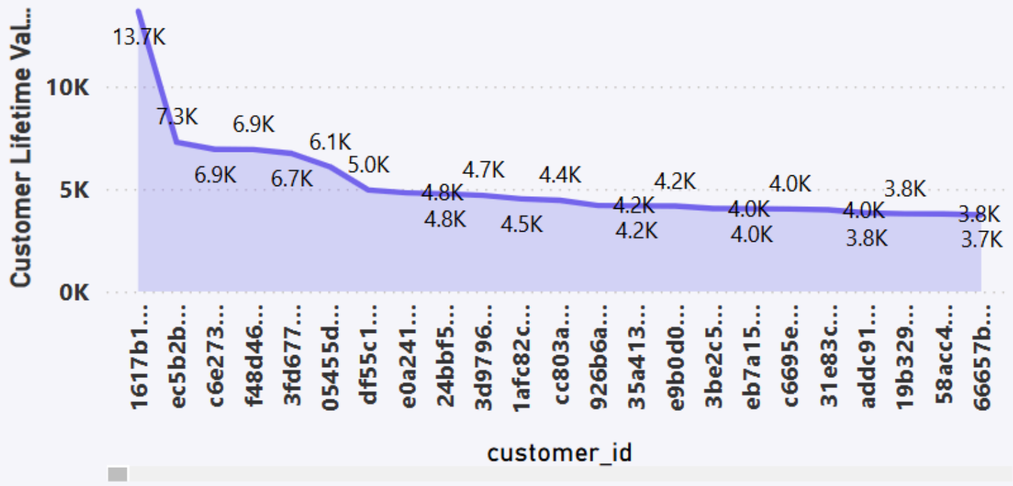

## EACH PRODUCT CATEGORY SALES

| product category | Sum of payment_value |
|---|---|
| Watches present | $13,10,893.450000005 |
| toys | $5,66,925.1400000002 |
| telephony | $3,98,760.08 |
| technical books | $23,627.89999999999 |
| stationary store | $2,81,098.9300000014 |
| sport leisure | $11,66,060.500000003 |
| song | $6,901.429999999999 |
| SIGNALIZATION AND SAFETY | $28,118.26000000001 |
| Room Furniture | $89,422.88000000011 |
| pet Shop | $2,55,330.2900000001 |

## RETURNED ORDERS BY PRODUCT CATEGORY

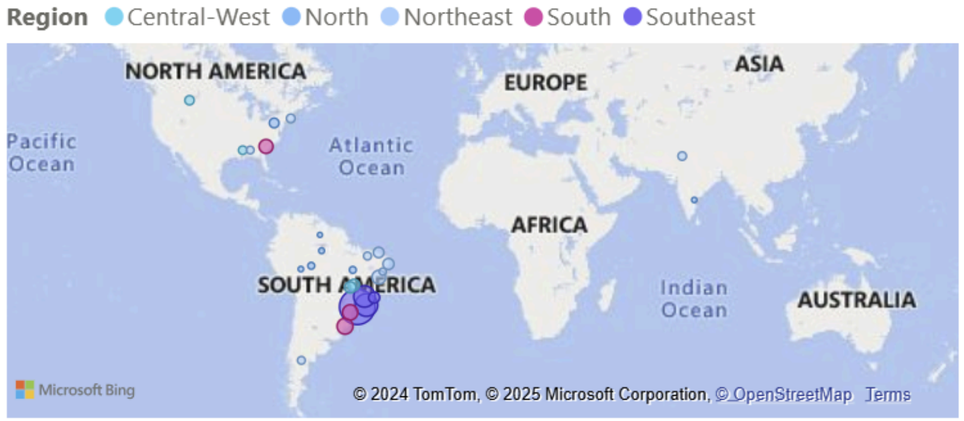



## CUSTOMER LIFETIME VALUE BY CUSTOMER ID

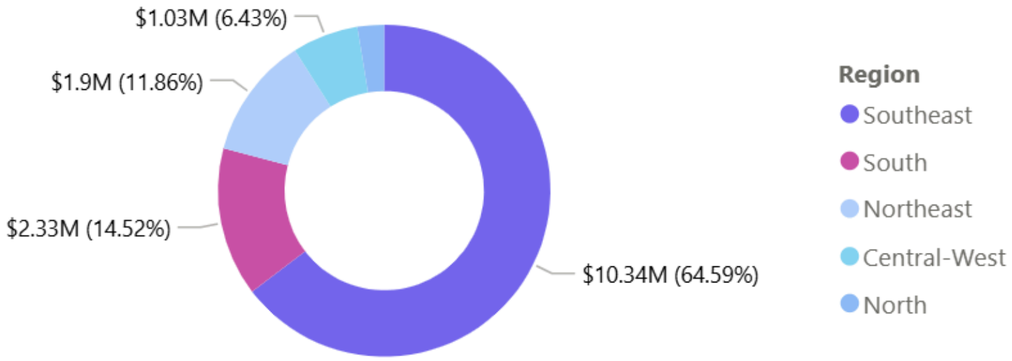


# Regional Analysis as per Operational Efficiency

## TOTAL SALES REVENUE BY CUSTOMER STATE AND REGION




## SUM OF PAYMENT VALUE BY REGION




## Region Slicer

- [ ] Central-West
- [ ] North
- [ ] Northeast
- [ ] South
- [ ] Southeast

## Year, Quarter, Month, Day

- [ ] 2016
  - [ ] Qtr 1
  - [ ] Qtr 2
  - [ ] Qtr 3
  - [ ] Qtr 4
- [ ] 2017
  - [ ] Qtr 1
  - [ ] Qtr 2
  - [ ] Qtr 3

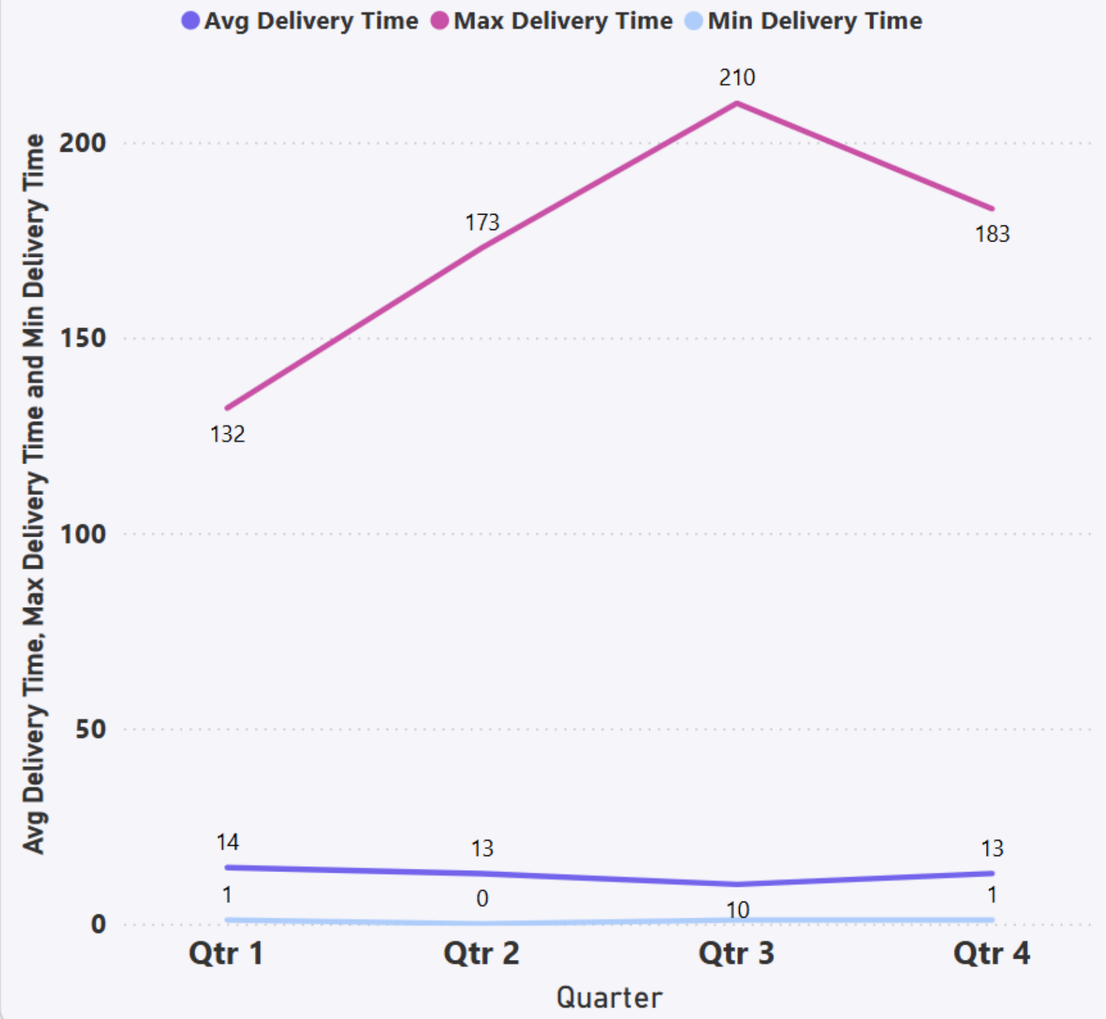
## Product Category Slicer

- [ ]
- [ ] Agro Industria ...
- [ ] Art
- [ ] Arts and Crafts
- [ ] audio
- [ ] automotive
- [ ] babies
- [ ] Bags Accessori...
- [ ] bed table bath

## RETURNED ORDER BY REGION

# Delivery Time Analysis

## Avg Delivery Time, Max Delivery Time and Min Delivery Time by Quarter



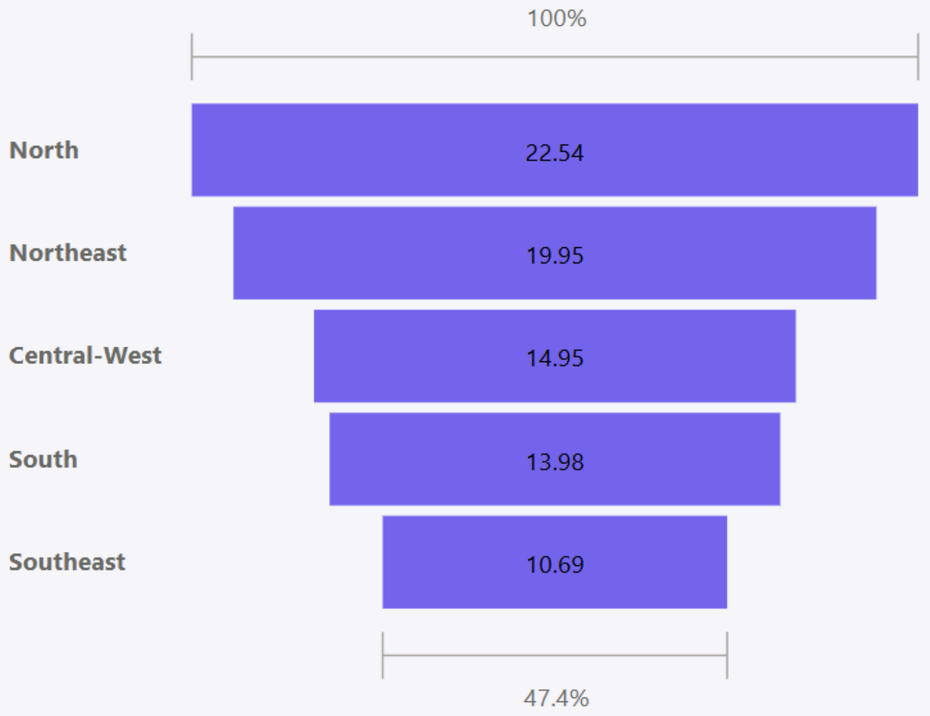

## Product Category Slicer

- [ ]
- [ ] Agro Industria e Comercio
- [ ] Art
- [ ] Arts and Crafts

## Year, Quarter, Month, Day

- [ ] 2016
  - [ ] Qtr 1
  - [ ] Qtr 2
  - [ ] Qtr 3

## AVERAGE DELIVERY TIME BY REGION



# Sales Performance

*Monthly Trends :*
Sales peaked in July (1.7M) but dropped significantly in Q4, with December being the lowest-performing month (0.88M).

*Quarterly Trends:*
Q2 generated the highest sales (5.1M), followed by Q3 (4.1M). Q4 had the lowest contribution (2.7M).

*Best-Selling Products:*
Top categories include Health Beauty (1.45M) and Watches Present (1.31M), highlighting potential focus areas for promotions.

*Average Order Value:*
The average order value stands at 154.10, indicating customer purchase behavior and potential upselling opportunities.

# Sales Customer Insights:

*Customer Segments:*

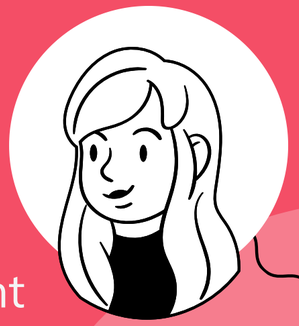
93.83% of customers are low spenders. This suggests a significant portion of the audience focuses on smaller transactions.

*Returned Orders:*

The "Sport Leisure" and "Housewares" categories have the highest returns, with 47 and 37 returns, respectively. Addressing product quality or return policies may help reduce these numbers.

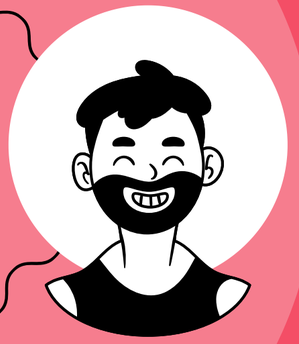
*Customer Lifetime Value (CLV):*

A small subset of customers drives higher lifetime value (e.g., the highest CLV is 13.7K), indicating potential for loyalty programs targeting this segment.



**Cassandra Romaine**

“I couldn't be happier with my online shopping experience”



**Idris Lettuce**

"The product exceeded my expectations".

**Vanessa Arugula**

"Carefully packaged, item came as described".

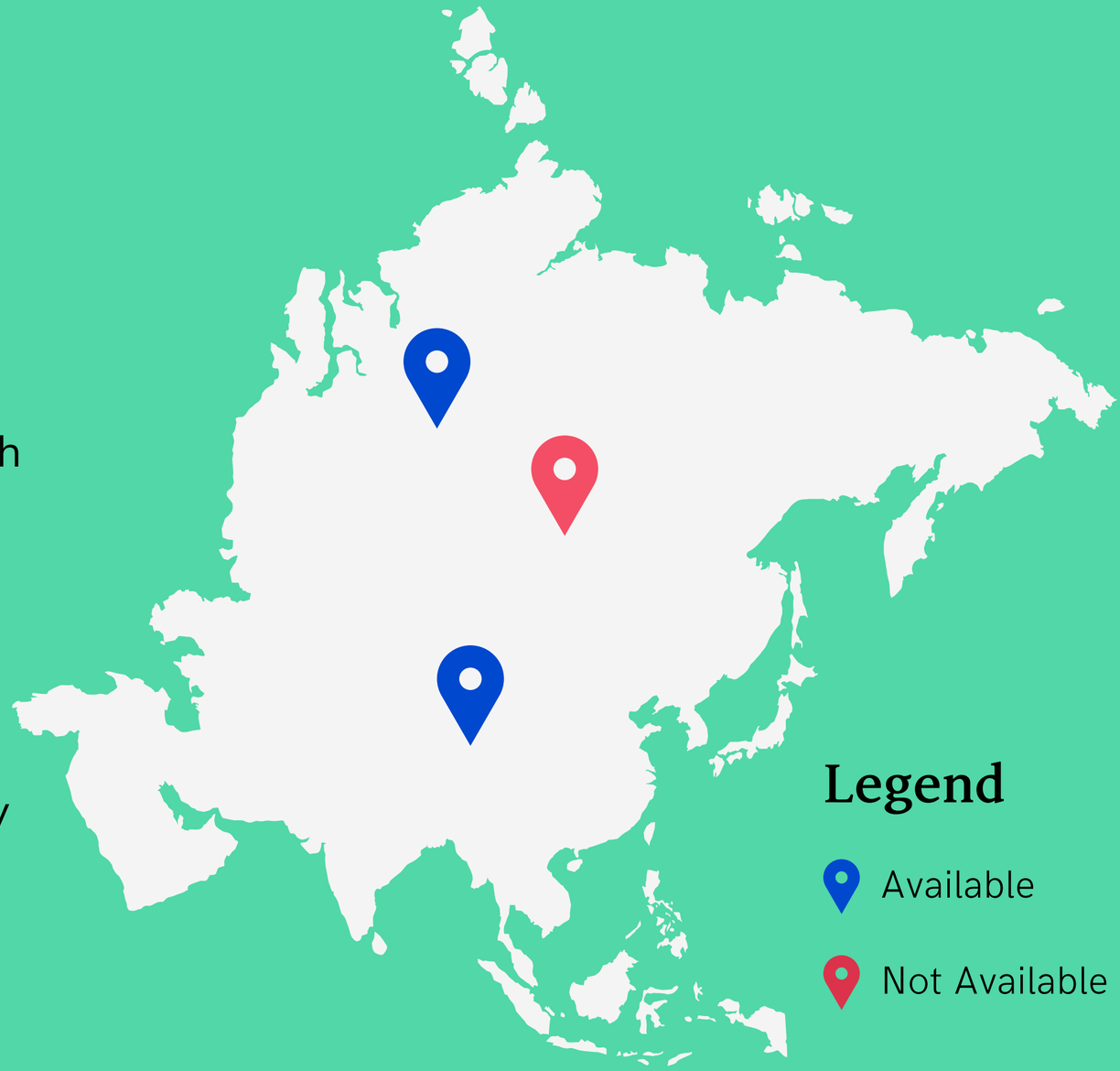
# Regional Analysis

*Sales by Region :*

The North region dominates with 64.59% of total payment value (10.34M). The Southeast and South regions contribute less significantly (6.43% and 11.86%, respectively).

*Returned Orders by Region:*

North America sees the highest returns, possibly due to operational issues or misalignment with customer expectations.



# Delivery Time Trend Analysis

*Seasonal Declines :*

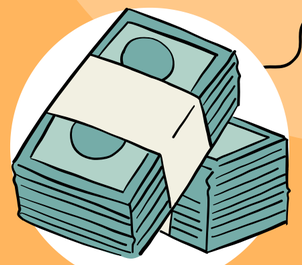
There’s a noticeable drop in sales from September onward, with a dramatic decrease in October. Strategic holiday promotions in Q4 could help mitigate this decline.

*Consistency :*

Despite the drop in Q4, Health Beauty products maintain strong sales, indicating steady demand even during off-peak times.

“I love this product. What a lifesaver!” - Corey Gates

“This product changes the game for real." - Dylan Walter



“Why have I not heard about this before?" - Erin Grey

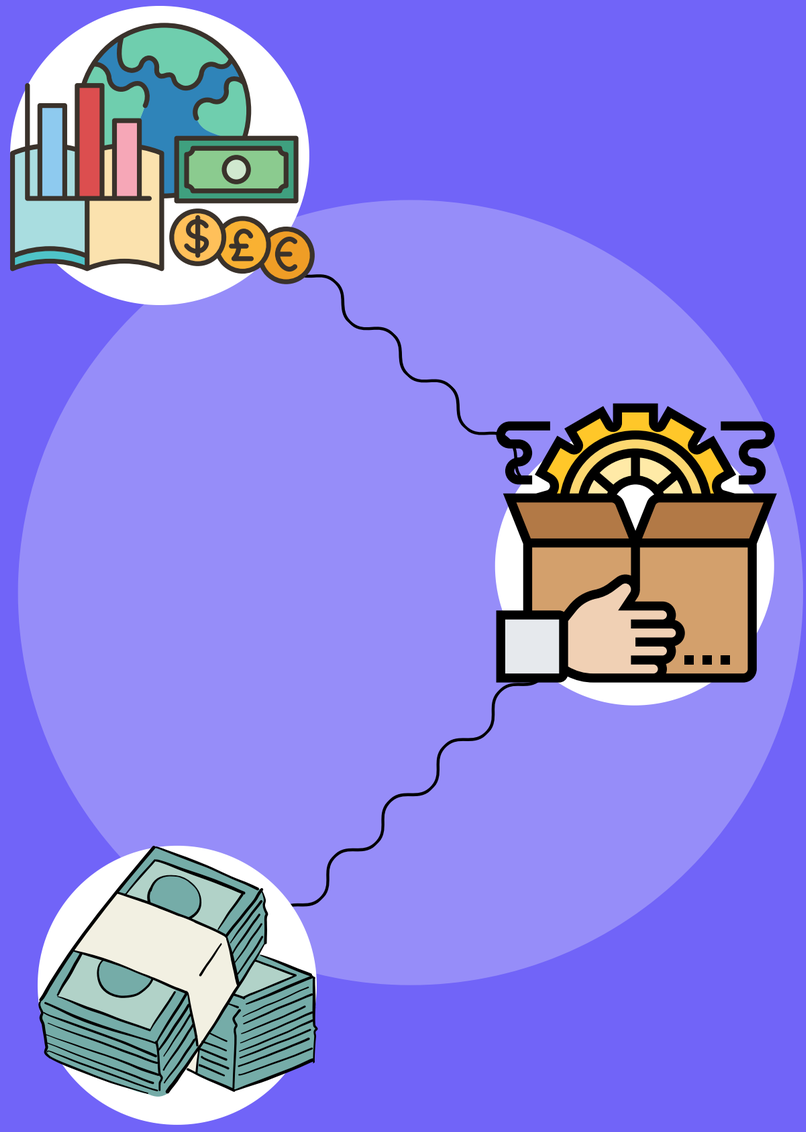
# Delivery Time Trend Analysis

*Delivery Times :*

The North region has the longest average delivery time (22.5 days), while the Southeast region performs best with the shortest average delivery time (10.7 days).

*Trends by Quarter :*

Average delivery time increased sharply in Q2 (210 days max), pointing to potential logistical inefficiencies during this period.



# Call to Action

As a data enthusiast, I am passionate about turning raw data into actionable insights that drive business growth. My focus lies in identifying trends, solving problems, and making data-backed decisions.

As a proactive learner with a keen eye for detail, I am eager to apply my analytical skills to real-world challenges while continuously expanding my knowledge to add value to any organization.

Let's connect and explore how I can bring fresh perspectives to your team!

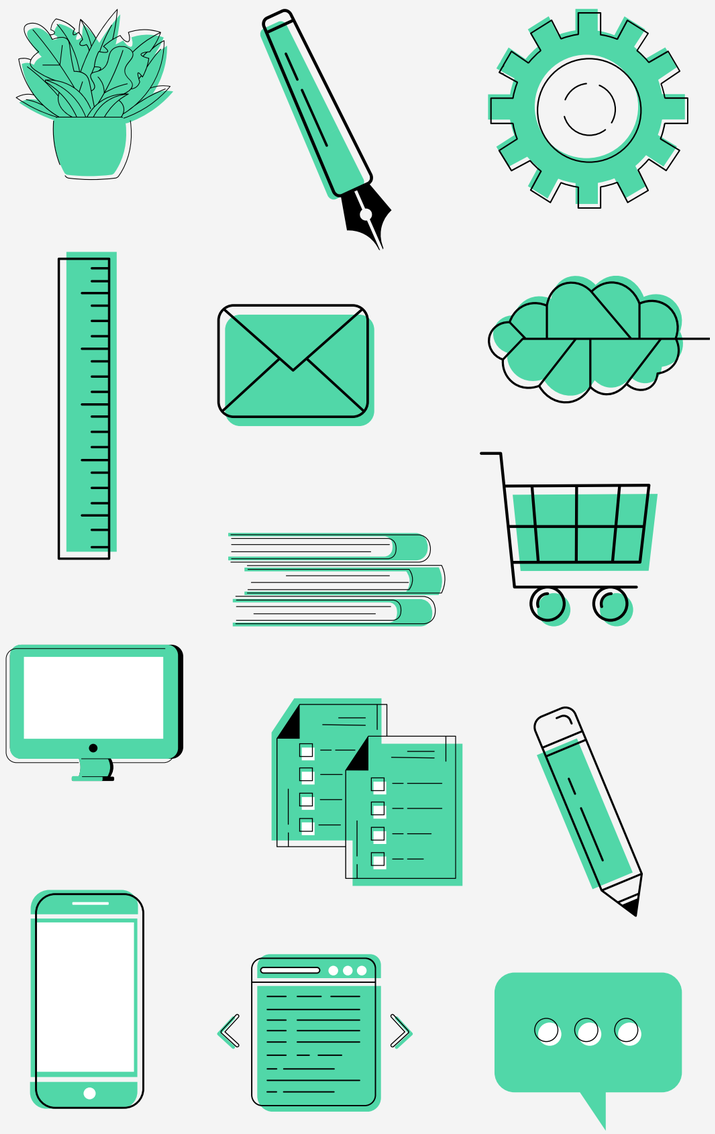
# Contact Me!


8928382552

jagrutijadhav21@gmail.com

https://github.com/jagruti217

https://www.linkedin.com/in/jagruti-jadhav-0275192a8

# THANK YOU FOR WATCH-ING