WSCUBE TECH

E-COMMERCE SALES ANALYSIS

CAPSTONE PROJECT

DATA ANALYTICS WITH ADVANCED SQL | POWER BI | MYSQL | DAX | VISUALIZATION

Presnted by Nidhi Bharatkar

ABOUT PROJECT

As a Data Analyst at Target Sales Company, I focus on analyzing and extracting actionable insights from a wide range of datasets, including orders, customers, products, regions, and operational metrics.

My primary objective is to enhance the company's strategic decision-making process by uncovering critical trends in sales performance, customer behavior, and operational efficiency.

This project involves advanced SQL data extraction techniques and the creation of interactive dashboards using Power BI/Tableau, offering a comprehensive view of business performance to drive data-informed decisions.

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
4. Identify customers with the highest average order value (AOV).
5. Analyze delivery performance by calculating the average delivery time by region.

PROBLEM STATEMENT FOR SQL

6. Identify regions or products with the highest canceled rates.

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.

Phase 1:
SQL Analysis Advanced Queries

SALES PERFORMANCE ANALYSIS

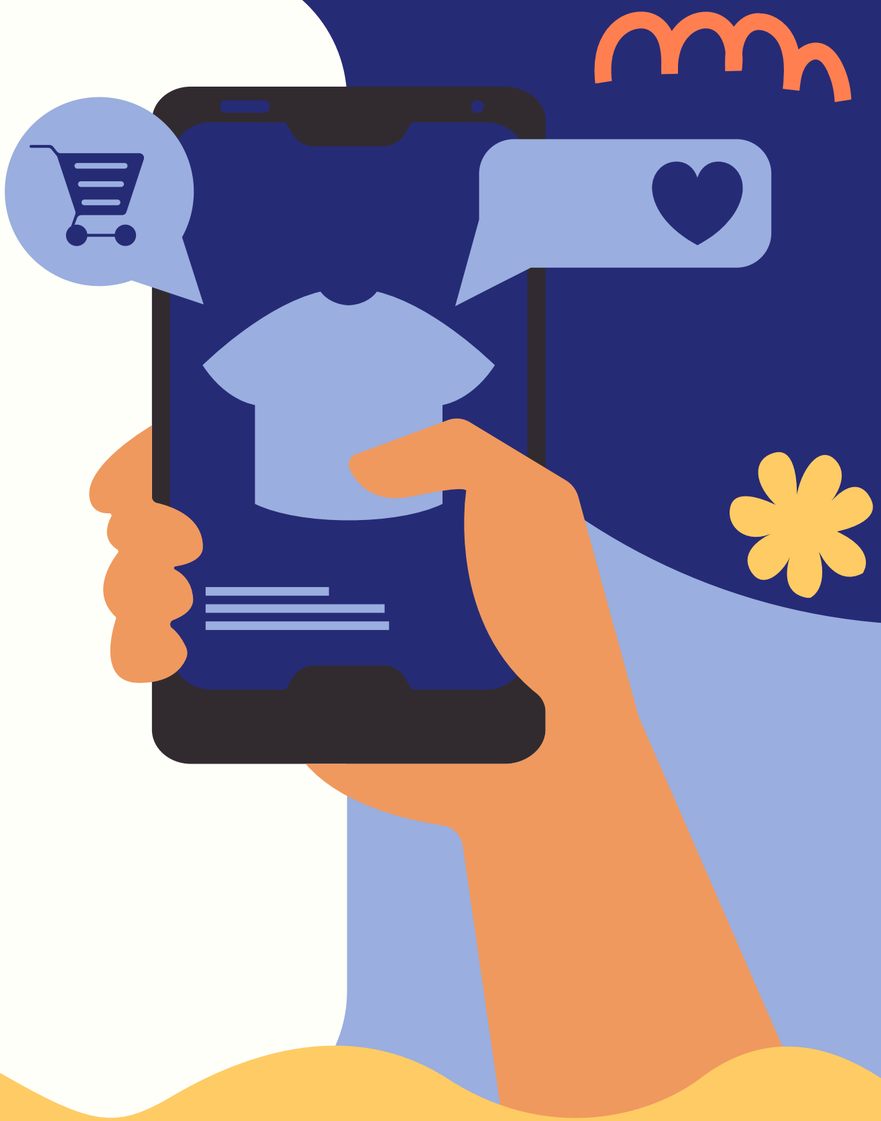
Image hosted by WittySparks | Image source FreePik

1 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;
```

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

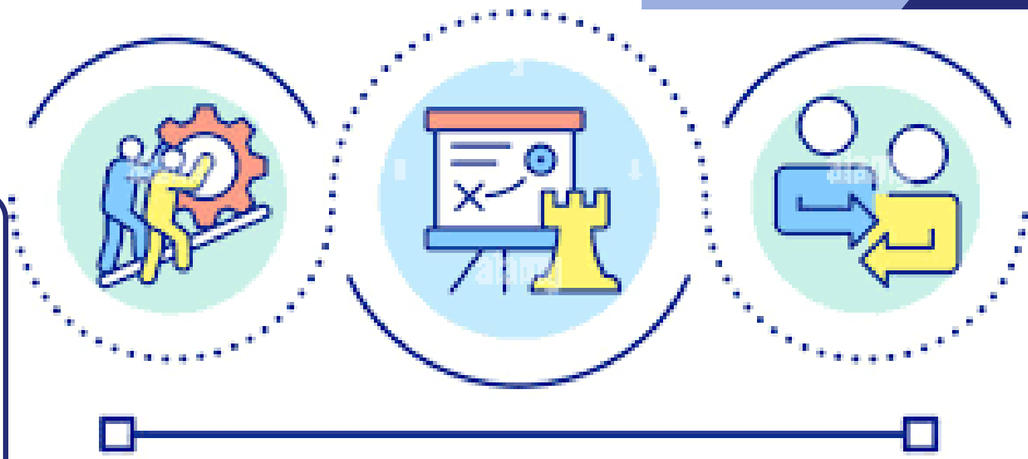
```sql
SELECT
    p.product_id,
    ROUND(SUM(oi.price), 2) AS revenue,
    SUM(p.product_photos_qty) AS total_quantity_sold
FROM
    order_items AS oi
        JOIN
    products AS p ON oi.product_id = p.product_id
        JOIN
    orders AS o ON o.order_id = oi.order_id
WHERE
    o.order_status != 'canceled'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



CUSTOMER INSIGHTS

3 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;
```



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

```sql
SELECT
    c.customer_unique_id,
    COUNT(o.order_id) AS total_orders,
    SUM(pay.payment_value) AS total_spend,
    SUM(pay.payment_value) / COUNT(o.order_id) AS average_order_value
FROM
    ecommerce.customers c
        JOIN
    ecommerce.orders o ON c.customer_id = o.customer_id
        JOIN
    ecommerce.order_items oi ON o.order_id = oi.order_id
        JOIN
    ecommerce.payments pay ON o.order_id = pay.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY c.customer_unique_id
HAVING total_orders > 1
ORDER BY average_order_value DESC
LIMIT 10;
```

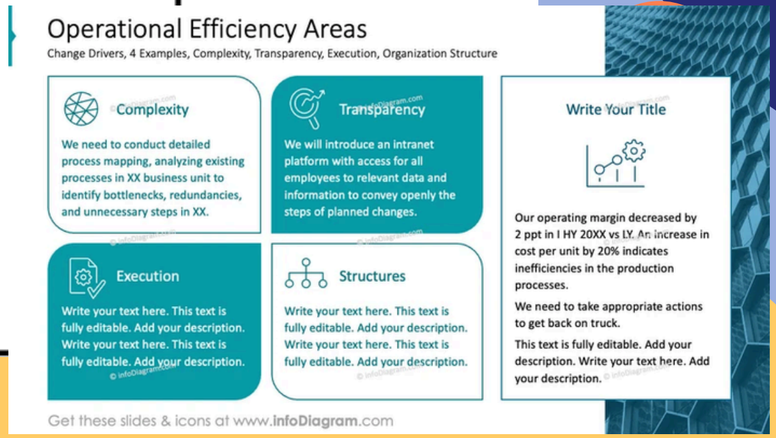
OPERATIONAL EFFICIENCY (BY REGION):

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

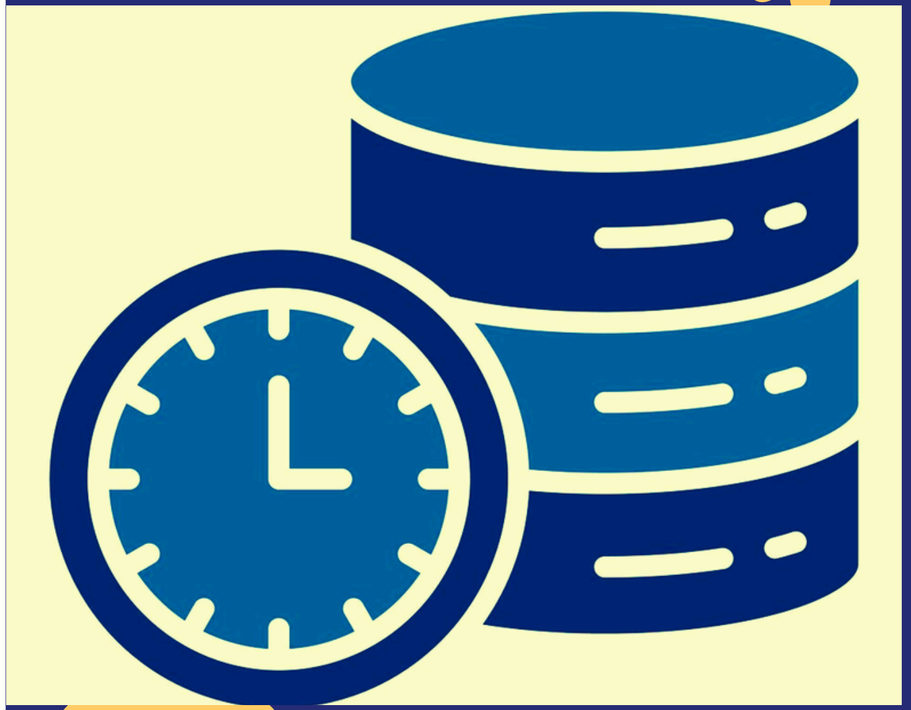
```sql
SELECT
    s.seller_state AS region,
    ROUND(AVG(DATEDIFF(o.order_estimated_delivery_date,
                       o.order_purchase_timestamp))) AS average_delivery_time
FROM
    order_items AS oi
        JOIN
    orders AS o ON oi.order_id = o.order_id
        JOIN
    sellers AS s ON s.seller_id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

6 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:



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

```sql
SELECT
    YEAR(o.order_purchase_timestamp) AS year,
    MONTH(o.order_purchase_timestamp) AS month,
    SUM(pay.payment_value) AS total_sales
FROM
    ecommerce.orders o
        JOIN
    ecommerce.order_items oi ON o.order_id = oi.order_id
        JOIN
    ecommerce.payments pay ON o.order_id = pay.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY YEAR(o.order_purchase_timestamp) , MONTH(o.order_purchase_timestamp)
ORDER BY year DESC , month DESC;
```

8 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:

9 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;
```

10 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;
```

Phase 2:
Data Visualization – Power BI

PowerBI

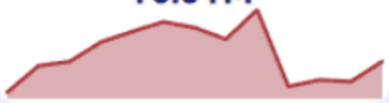
Presenting the Target Sales Company Dashboard – a dynamic and data-driven view into our sales performance, customer insights, and operational effectiveness.

Sales Performance Dashboard

E-COMMERCE DASHBOARD

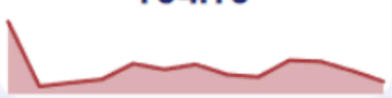
Total Revenue

16.01M



Average Order Values

154.10



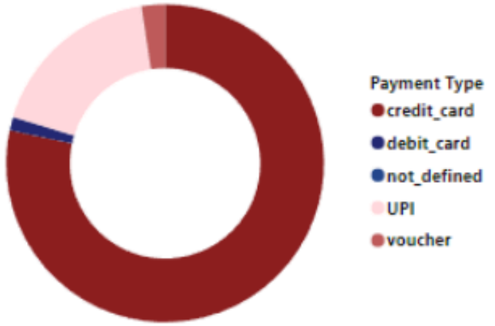
Total Orders

99K

Sum of Price By Product Category

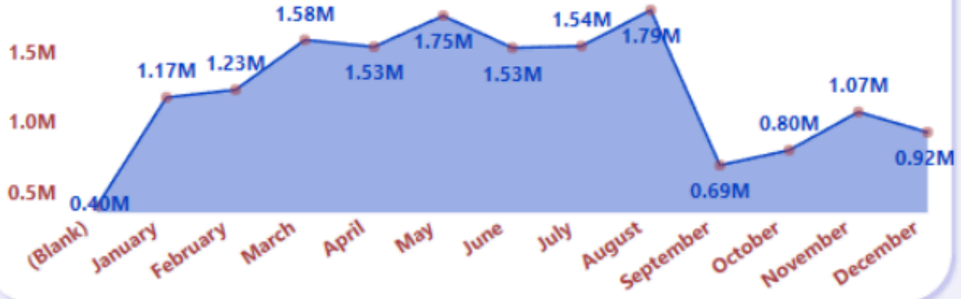
Product Category	Total Sales
Watches Present	13,10,893.45
Toys	5,66,925.14
Telephony	3,98,760.08
Total	1,60,08,872.14

Sales By Month

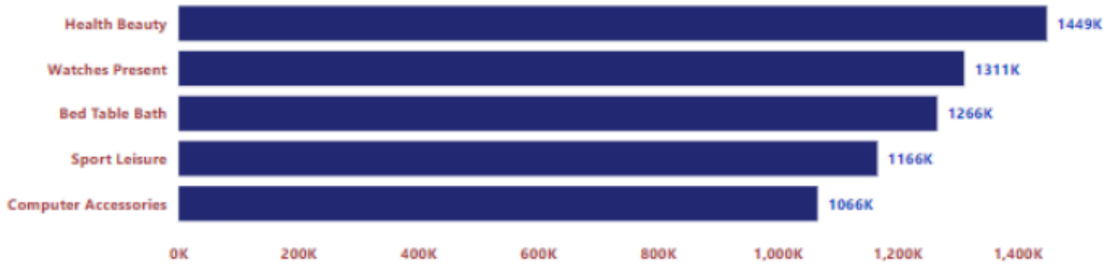


Sales By City Area

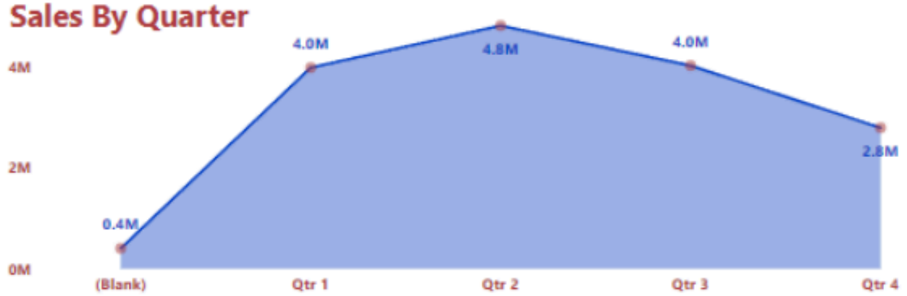
Sales By Month



Top 5 Product Category



Sales By Quarter



Customer Insights Dashboard

E-COMMERCE DASHBOARD

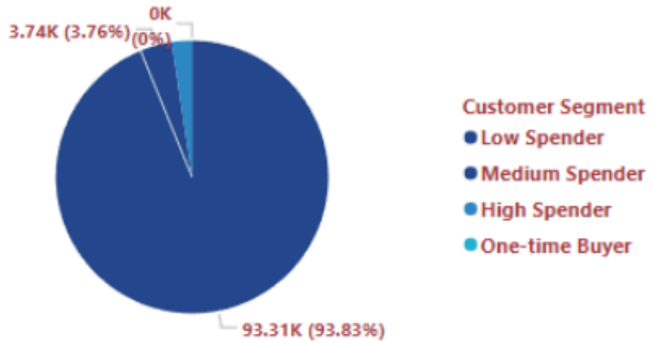
Total Customers

99K

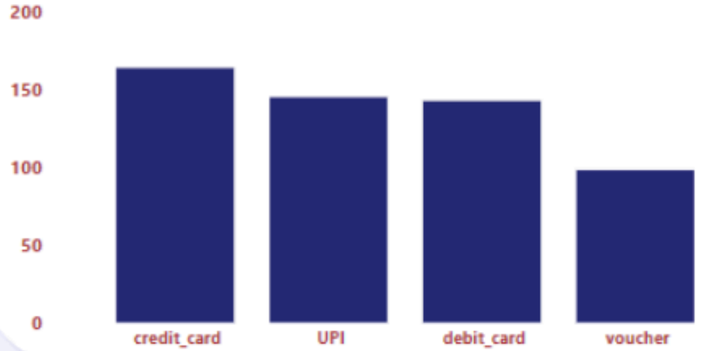
Average Customer Spend

160.99

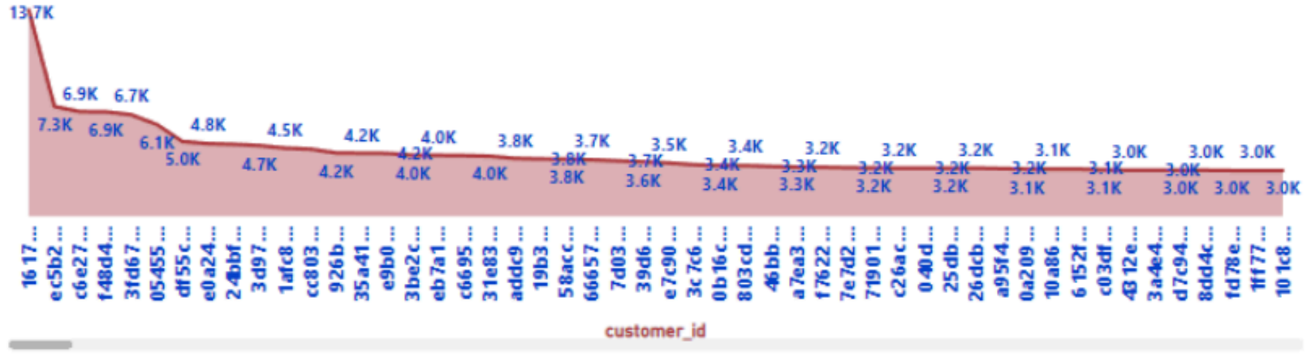
Count of Customer ID by Customer Segment



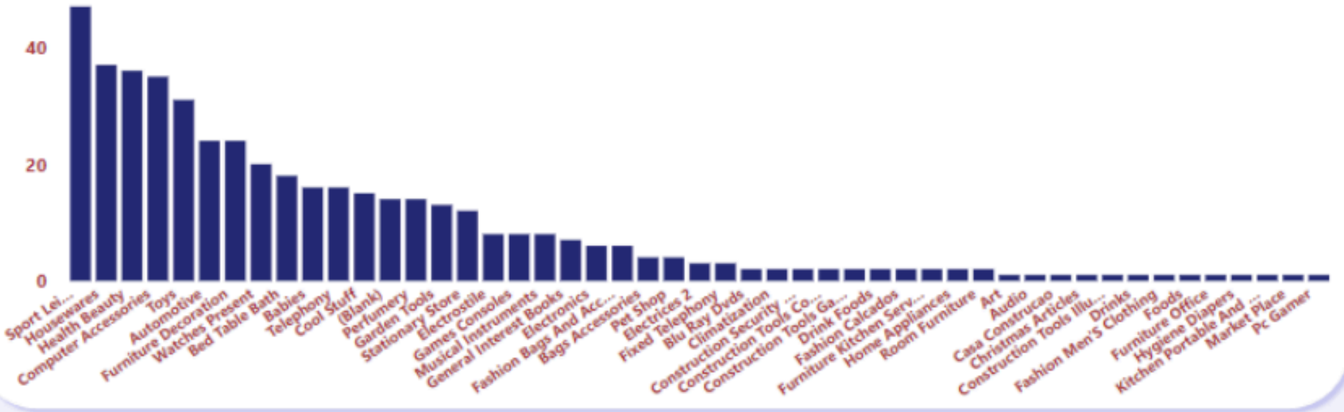
Average Customer Spend by Payment Type



Customer Lifetime Value by Customer Id



Returned Order by Product Category



Regional & Operational Trends Dashboard

E-COMMERCE DASHBOARD

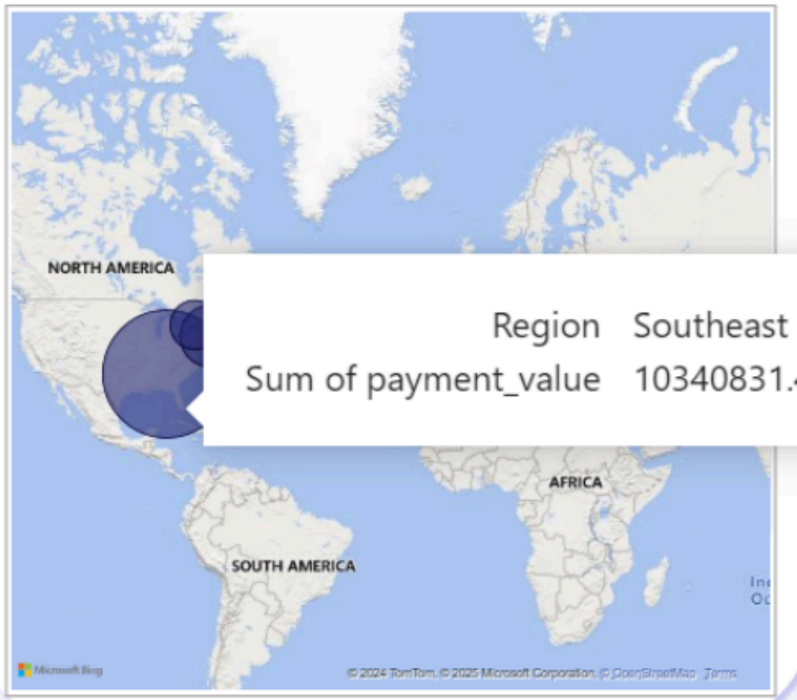
Average Delivery Time

12.50

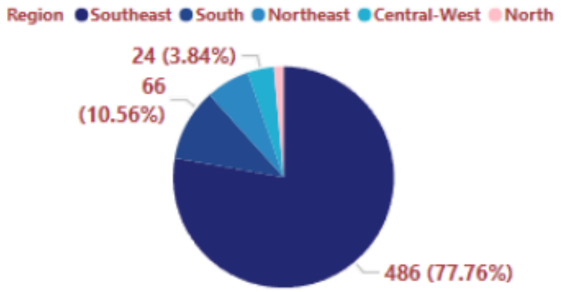
Return Order

625

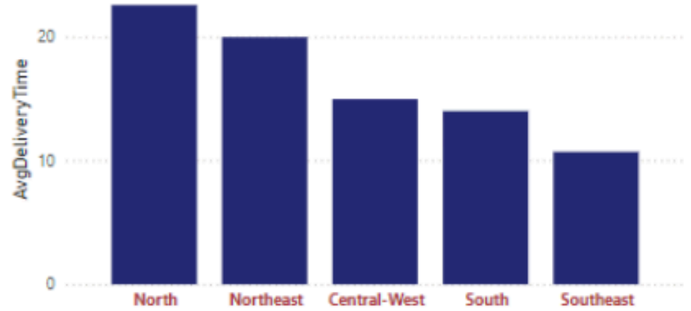
Total Sales by Region



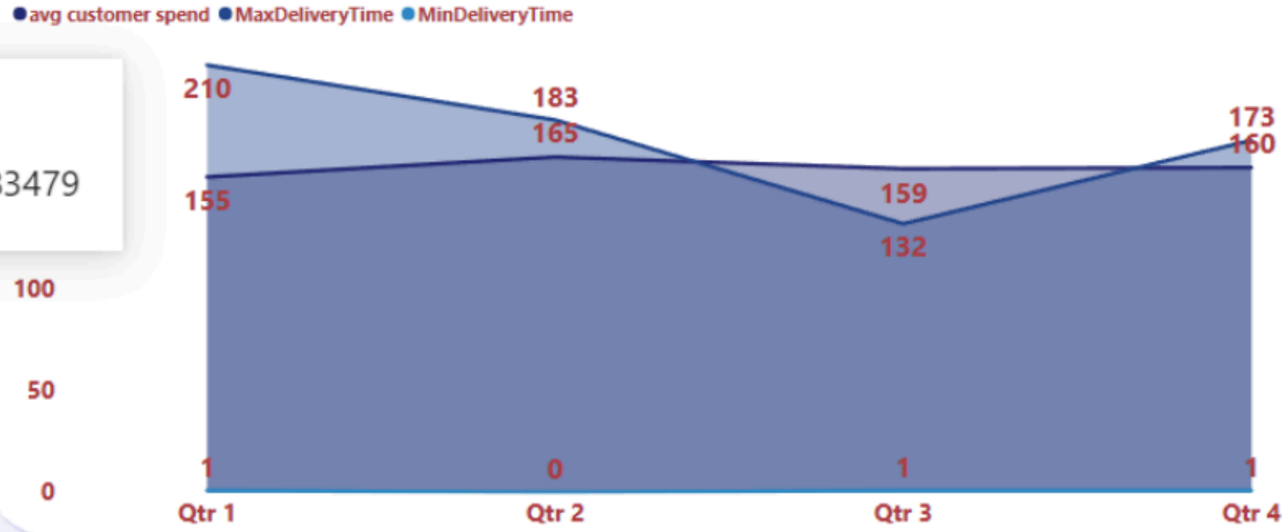
Returned Orders by Region



Average Delivery Time by Region



Average Customer Spend : Max & Min Delivery Time by Quarter



Phase 3:
E-CommerceSales Performance Insights Report

Executive Summary

As a Data Analyst at Target Sales Company, I analyzed sales data to uncover key insights and provide actionable recommendations. The analysis focused on sales performance, customer behavior, regional trends, and operational metrics, with findings presented through an interactive Power BI dashboard for easy stakeholder access.

Key focus areas include:

- **Sales Performance Trends**
- **Customer Lifetime Value (CLV) and Segmentation**
- **Regional Sales and Return Analysis**
- **Operational Metrics and Delivery Time**
- **Seasonality and Monthly Sales Trends**

1.SALES PERFORMANCE

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

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

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

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

2.CUSTOMER INSIGHTS

1. 93.83% of customers are low spenders, making small, infrequent purchases: The customer base is predominantly low-value, suggesting that most buyers are price-sensitive. Actionable Insight: Implement loyalty programs, bundle deals, and discount campaigns to increase customer lifetime value.

2.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.

3.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.

3.REGIONAL ANALYSIS

1.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)..

2.Returned Orders by Region: North America sees the highest returns, possibly due to operational issues or misalignment with customer expectations.

4.DELIVERY TIME TREND ANALYSIS

1.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.

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

3.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)

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

Conclusion

The E-Commerce Sales Dashboard provides valuable insights into sales performance, customer behavior, and operational efficiency. By addressing seasonal trends, customer engagement, and operational bottlenecks, the company can boost revenue, improve customer satisfaction, and enhance overall efficiency.

Key areas of focus include:

- Launching loyalty programs to retain high-value customers.
- Introducing targeted promotions during off-peak months.
- Improving delivery times in the North region to enhance customer satisfaction.
- Addressing return rates to reduce operational losses.

These recommendations, if implemented effectively, can help Target Sales Company drive growth, enhance operational efficiency, and improve customer loyalty

THANK YOU

GET IN TOUCH

LinkedIn : www.linkedin.com/in/nidhi-bharatkar-dataanalyst