Capstone Project: Driving Insights for Target Sales Company
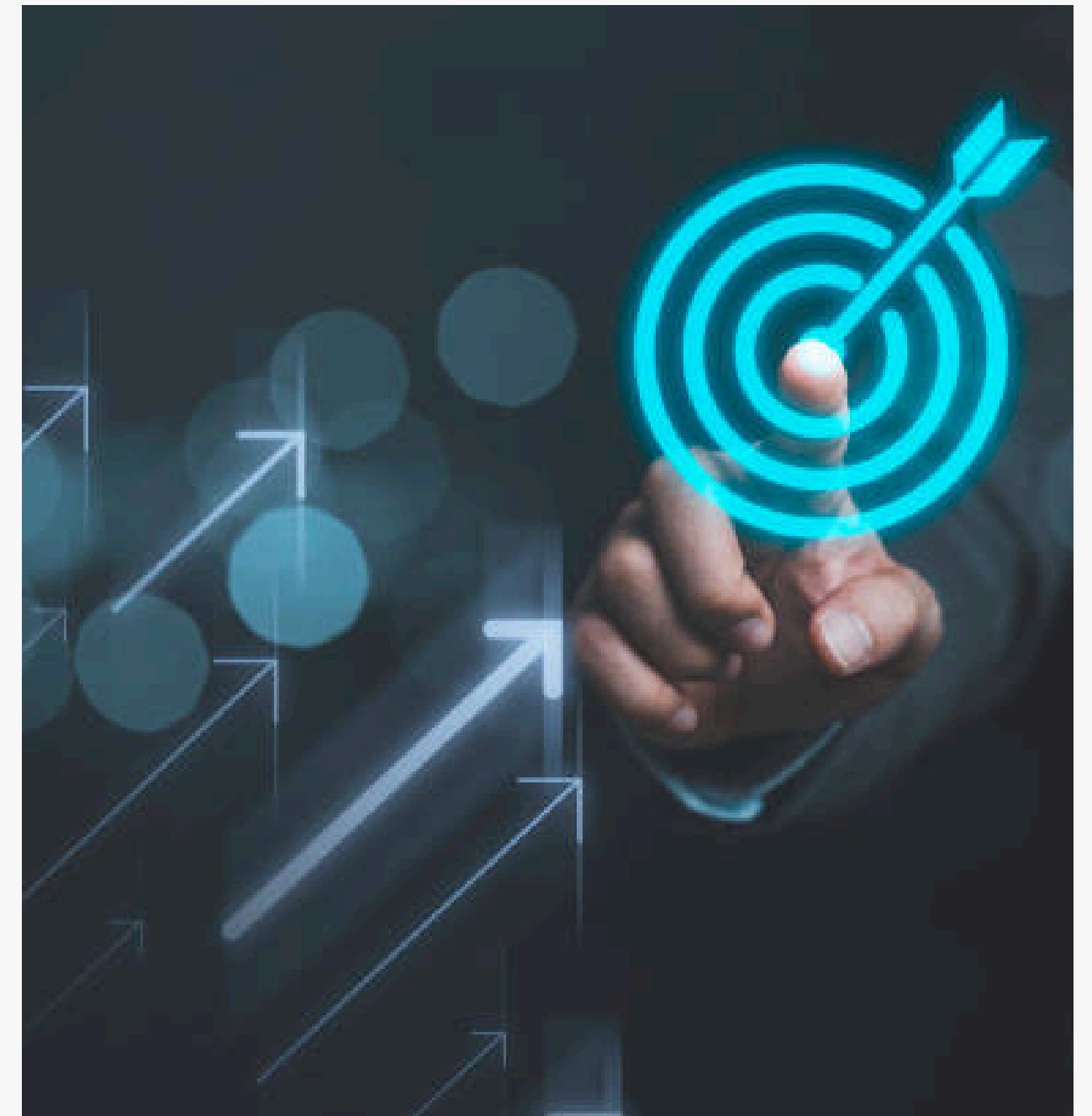
Exploring Advanced Analytics and Strategic Decision-Making

Presented By:Pankaj Kumar

Project Goal

Conduct an in-depth data analysis using advanced MySQL queries and Power BI to uncover valuable insights that drive informed decision-making for the company. The focus is on analyzing sales performance, understanding customer behavior, improving operational efficiency, and identifying key time-based trends. The goal is to optimize sales strategies, refine customer targeting, and streamline operations while presenting these insights through dynamic, interactive Power BI dashboards to support strategic decisions.
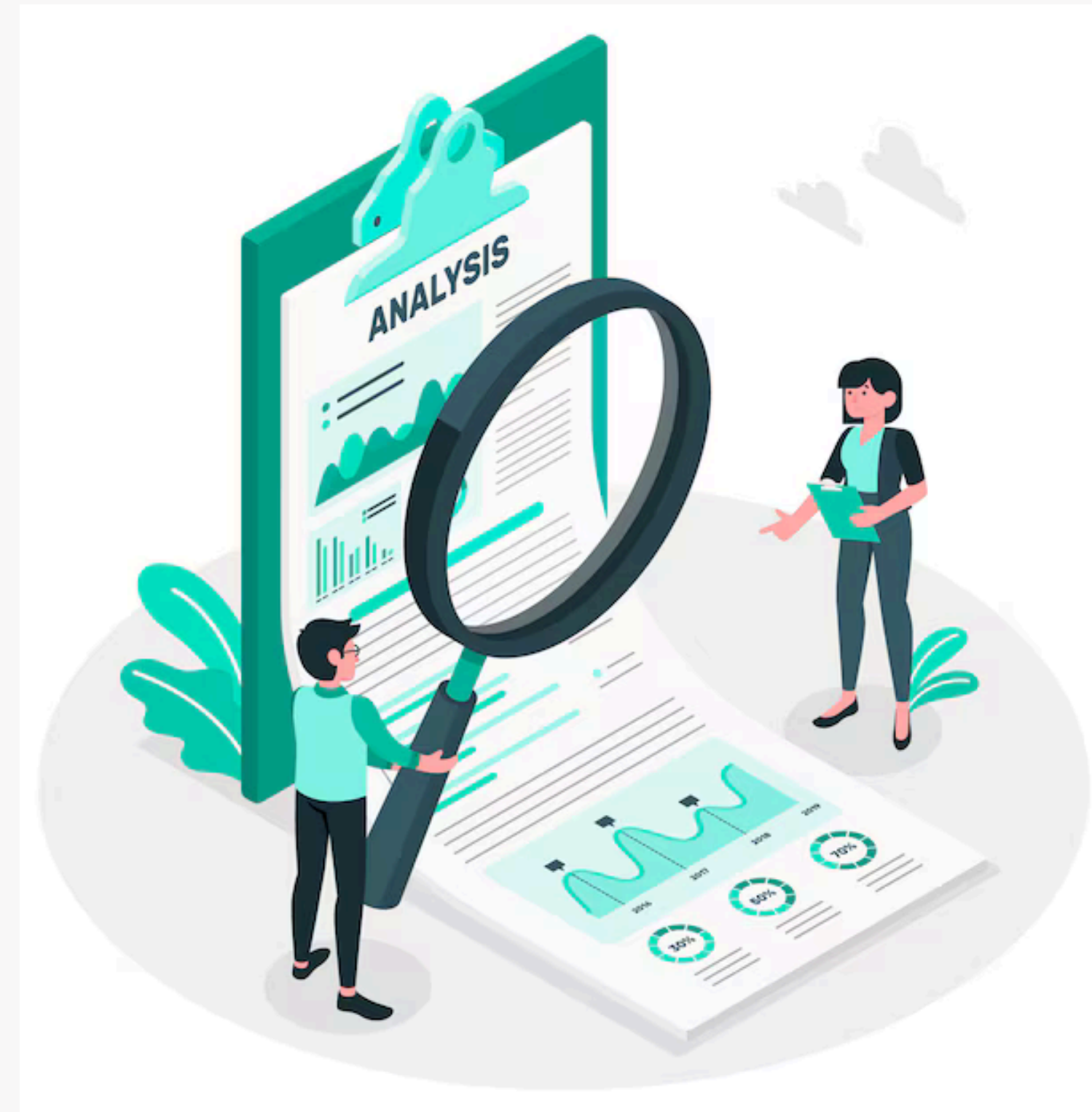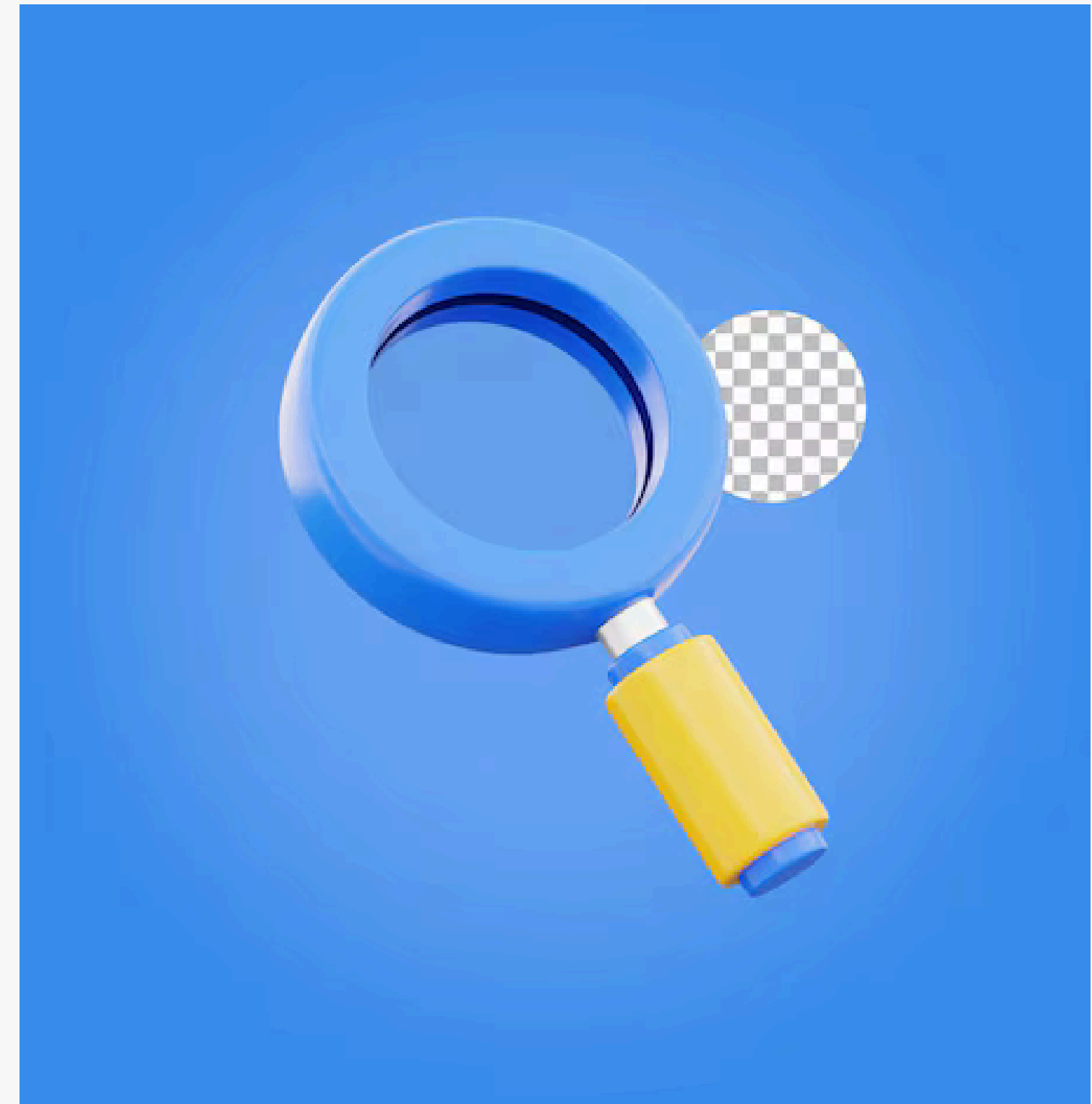
Dataset Overview

- **Orders: Transaction details, timestamps, etc.**
- **Customers: Demographics, preferences.**
- **Products: Categories, pricing, performance.**
- **Regions: Geographical insights.**
- **Operational Metrics: Efficiency, costs, etc.**

Tool used:

MySQL

+

Power BI

Key Phases of the Capstone Project:

- **Phase 1: Advanced SQL Queries**
  - **Data extraction and transformations.**
- **Phase 2: Visualization with Power BI/Tableau**
  - **Interactive dashboards and insights.**

Phase 1

Advance analytics using Mysql Queries

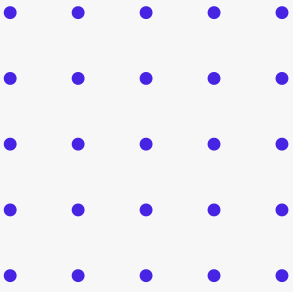## Key Metrics: Query to calculate total sales revenue per category

```sql
SELECT
    p.product_category AS category,
    round(SUM(pay.payment_value),2) AS total_sales_revenue
FROM
    Products AS p
JOIN
    Order_items AS oi ON p.product_id = oi.product_id
JOIN
    Orders AS o ON oi.order_id = o.order_id
JOIN
    Payments AS pay ON o.order_id = pay.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY
    p.product_category
ORDER BY
    total_sales_revenue DESC;
```

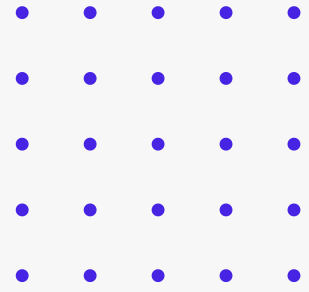| category	total_sales_revenue
bed table bath	1692714.28
HEALTH BEAUTY	1620684.04
computer accessories	1549372.59
Furniture Decoration	1394466.93
Watches present	1387362.45
sport leisure	1349446.93
housewares	1069787.97
automotive	833745.67

Key Metrics: Query to calculate total sales revenue per sub-category

```sql
SELECT
  CASE
    WHEN p.product_category IN ('perfumery', 'HEALTH BEAUTY') THEN 'Luxury & Beauty'
    WHEN p.product_category IN ('Art', 'stationary store') THEN 'Art & Stationery'
    WHEN p.product_category IN ('sport leisure', 'Garden tools') THEN 'Sports & Outdoors'
    WHEN p.product_category = 'babies' THEN 'Baby Products'
    WHEN p.product_category IN ('housewares', 'bed table bath') THEN 'Home Essentials'
    WHEN p.product_category = 'musical instruments' THEN 'Instruments'
    WHEN p.product_category = 'Cool Stuff' THEN 'Gadgets'
    WHEN p.product_category IN ('Furniture Decoration', 'Furniture office') THEN 'Furniture'
    WHEN p.product_category = 'home appliances' THEN 'Appliances'
    WHEN p.product_category = 'toys' THEN 'Kids Toys'
    WHEN p.product_category = 'CONSTRUCTION SECURITY TOOLS' THEN 'Construction Tools'
    WHEN p.product_category IN ('computer accessories', 'PCs') THEN 'Computing'
    WHEN p.product_category = 'Bags Accessories' THEN 'Bags & Accessories'
    WHEN p.product_category = 'automotive' THEN 'Automotive'
    WHEN p.product_category IN ('electronics', 'electrostile') THEN 'Electronics'
    WHEN p.product_category = 'Fashion Calcados' THEN 'Footwear'
    WHEN p.product_category = 'telephony' THEN 'Telecommunication'
    WHEN p.product_category = 'Fashion Bags and Accessories' THEN 'Fashion Accessories'
    WHEN p.product_category = 'Casa Construcao' THEN 'Home Construction'
    WHEN p.product_category = 'Watches present' THEN 'Watches & Gifts'
    WHEN p.product_category = 'Construction Tools Construction' THEN 'Construction'
    WHEN p.product_category = 'pet Shop' THEN 'Pet Supplies'
    WHEN p.product_category = 'Agro Industria e Comercio' THEN 'Agriculture & Commerce'
    ELSE 'Other'
  END AS sub_category,
  round(SUM(pay.payment_value),2) AS total_sales_revenue
FROM
  Products AS p
JOIN
  Order_items AS oi ON p.product_id = oi.product_id
JOIN
  Orders AS o ON oi.order_id = o.order_id
JOIN
  Payments AS pay ON o.order_id = pay.order_id
WHERE
  o.order_status = 'delivered'
GROUP BY
  CASE
    WHEN p.product_category IN ('perfumery', 'HEALTH BEAUTY') THEN 'Luxury & Beauty'
    WHEN p.product_category IN ('Art', 'stationary store') THEN 'Art & Stationery'
    WHEN p.product_category IN ('sport leisure', 'Garden tools') THEN 'Sports & Outdoors'
    WHEN p.product_category = 'babies' THEN 'Baby Products'
    WHEN p.product_category IN ('housewares', 'bed table bath') THEN 'Home Essentials'
    WHEN p.product_category = 'musical instruments' THEN 'Instruments'
    WHEN p.product_category = 'Cool Stuff' THEN 'Gadgets'
    WHEN p.product_category IN ('Furniture Decoration', 'Furniture office') THEN 'Furniture'
    WHEN p.product_category = 'home appliances' THEN 'Appliances'
    WHEN p.product_category = 'toys' THEN 'Kids Toys'
    WHEN p.product_category = 'CONSTRUCTION SECURITY TOOLS' THEN 'Construction Tools'
    WHEN p.product_category IN ('computer accessories', 'PCs') THEN 'Computing'
    WHEN p.product_category = 'Bags Accessories' THEN 'Bags & Accessories'
    WHEN p.product_category = 'automotive' THEN 'Automotive'
    WHEN p.product_category IN ('electronics', 'electrostile') THEN 'Electronics'
    WHEN p.product_category = 'Fashion Calcados' THEN 'Footwear'
    WHEN p.product_category = 'telephony' THEN 'Telecommunication'
    WHEN p.product_category = 'Fashion Bags and Accessories' THEN 'Fashion Accessories'
    WHEN p.product_category = 'Casa Construcao' THEN 'Home Construction'
    WHEN p.product_category = 'Watches present' THEN 'Watches & Gifts'
    WHEN p.product_category = 'Construction Tools Construction' THEN 'Construction'
    WHEN p.product_category = 'pet Shop' THEN 'Pet Supplies'
    WHEN p.product_category = 'Agro Industria e Comercio' THEN 'Agriculture & Commerce'
    ELSE 'Other'
  END
ORDER BY
  total_sales_revenue DESC;
```

sub_category	total_sales_revenue
Home Essentials	2762502.25
Sports & Outdoors	2160061.86
Luxury & Beauty	2116947.64
Furniture	2030591.81
Other	1915121.09
Computing	1824044.47
Watches & Gifts	1387362.45
Automotive	833745.67

Key Metrics:Query to calculate total sales revenue per region

```sql
SELECT
    c.customer_state AS region,
    round(SUM(pay.payment_value),2) AS total_sales_revenue
FROM
    Orders AS o
JOIN
    Payments AS pay ON o.order_id = pay.order_id
JOIN
    Customers AS c ON o.customer_id = c.customer_id
WHERE
    o.order_status = 'delivered'
GROUP BY
    c.customer_state
ORDER BY
    total_sales_revenue DESC;
```

region	total_sales_revenue
SP	5770266.19
RJ	2055690.45
MG	1819277.61
RS	861802.4
PR	781919.55
SC	595208.4
BA	591270.6
DF	346146.17

Key Metrics: top 5 best-selling products by sales revenue

```sql
SELECT
    p.product_id,
    p.product_category,
    round(SUM(oi.price),2) AS total_revenue
FROM
    order_items oi
JOIN
    products p ON oi.product_id = p.product_id
GROUP BY
    p.product_id, p.product_category
ORDER BY
    total_revenue DESC
LIMIT 5;
```

| product_id	product_category	total_revenue
bb50f2e236e5eea0100680137654686c	HEALTH BEAUTY	63885
6cdd53843498f92890544667809f1595	HEALTH BEAUTY	54730.2
d6160fb7873f184099d9bc95e30376af	PCs	48899.34
d1c427060a0f73f6b889a5c7c61f2ac4	computer accessories	47214.51
99a4788cb24856965c36a24e339b6058	bed table bath	43025.56

Key Metrics: top 5 best-selling products by quantity sold
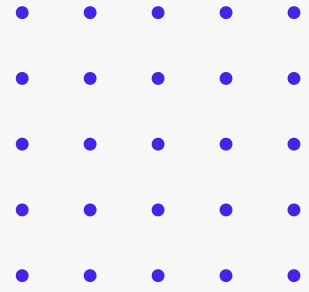
```sql
SELECT
    p.product_id,
    p.product_category,
    COUNT(oi.product_id) AS quantity_sold
FROM
    order_items oi
JOIN
    products p ON oi.product_id = p.product_id
GROUP BY
    p.product_id, p.product_category
ORDER BY
    quantity_sold DESC
LIMIT 5;
```

product_id	product_category	quantity_sold
aca2eb7d00ea1a7b8ebd4e68314663af	Furniture Decoration	527
99a4788cb24856965c36a24e339b6058	bed table bath	488
422879e10f46682990de24d770e7f83d	Garden tools	484
389d119b48cf3043d311335e499d9c6b	Garden tools	392
368c6c730842d78016ad823897a372db	Garden tools	388

Key Metrics:Query to find loyal customers
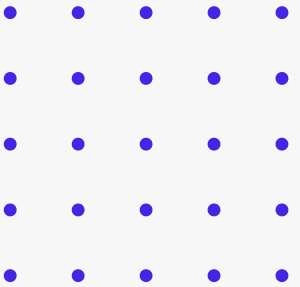
```sql
SELECT
    c.customer_id,
    c.customer_unique_id,
    COUNT(o.order_id) AS purchase_frequency,
    round(SUM(p.payment_value),2) AS total_spend
FROM
    customers c
JOIN
    orders o ON c.customer_id = o.customer_id
JOIN
    payments p ON o.order_id = p.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY
    c.customer_id, c.customer_unique_id
ORDER BY
    total_spend DESC, purchase_frequency DESC
LIMIT 10;
```

customer_id	customer_unique_id	purchase_frequency	total_spend
1617b1357756262bfa56ab541c47bc16	0a0a92112bd4c708ca5fde585afaa872	1	13664.08
ec5b2ba62e574342386871631fafd3fc	763c8b1c9c68a0229c42c9fc6f662b93	1	7274.88
c6e2731c5b391845f6800c97401a43a9	dc4802a71eae9be1dd28f5d788ceb526	1	6929.31
f48d464a0baaea338cb25f816991ab1f	459bef486812aa25204be022145caa62	1	6922.21
3fd6777bbce08a352fddd04e4a7cc8f6	ff4159b92c40ebe40454e3e6a7c35ed6	1	6726.66
05455dfa7cd02f13d132aa7a6a9729c6	4007669dec559734d6f53e029e360987	1	6081.54
df55c14d1476a9a3467f131269c2477f	da122df9eeddfedc1dc1f5349a1a690c	1	4950.34
24bbf5fd2f2e1b359ee7de94defc4a15	eebb5dda148d3893cdaf5b5ca3040ccb	1	4764.34

Key Metrics:Identify customers with the highest average order value (AOV)

```sql
SELECT
    c.customer_id,
    c.customer_unique_id,
    COUNT(o.order_id) AS total_orders,
    SUM(p.payment_value) AS total_spend,
    ROUND(SUM(p.payment_value) / COUNT(o.order_id), 2) AS average_order_value
FROM
    customers c
JOIN
    orders o ON c.customer_id = o.customer_id
JOIN
    payments p ON o.order_id = p.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY
    c.customer_id, c.customer_unique_id
HAVING
    COUNT(o.order_id) > 1
ORDER BY
    average_order_value DESC
LIMIT 10;
```

customer_id	customer_unique_id	total_orders	total_spend	average_order_value
926b6a6fb8b6081e00b335edaf578d35	edf81e1f3070b9dac83ec83dacdbb9bc	2	4194.7598876953125	2097.38
eb7a157e8da9c488cd4ddc48711f1097	58483a1c055dfb600f57c5b867174542	2	4034.4400024414062	2017.22
46bb3c0b1a65c8399d0363cefbcc4f37	7a96eb0a685f5c19b7dad29fc802aa64	2	3297.39990234375	1648.7
9cf0a858f5f153406bc333860eb23e22	c7794313e2cd53472b116a751350f132	2	2323.25	1161.62
e69714d6b494f7279512327ebcc01889	cad51aa032a0834a45c27586bce31d1c	2	2150.8099975585938	1075.4
d5c70e33b26d11965027895a8e8cb748	5e5a9e721cb83187f5893a8c7109faee	2	2066.3399658203125	1033.17
466dcf5fcdc7c7a8a4bfebc7ce59b281	a22ee6bd65e350c2d1a0d9a12c2e465f	2	2014.989990234375	1007.49
0928c1a893356a97b2191a1e812c9914	3a205021b542565ebf170c0ec55783d9	2	1942.5999755859375	971.3

Key Metrics: Analyze delivery performance by calculating the average delivery time by region

```sql
SELECT
    c.customer_state AS region,
    ROUND(AVG(DATEDIFF(o.order_delivered_customer_date,
o.order_purchase_timestamp)), 2) AS avg_delivery_time
FROM
    orders o
JOIN
    customers c ON o.customer_id = c.customer_id
WHERE
    o.order_status = 'delivered'
    AND o.order_delivered_customer_date IS NOT NULL
GROUP BY
    c.customer_state
ORDER BY
    avg_delivery_time ASC;
```

region	avg_delivery_time
RR	29.34
AP	27.18
AM	26.36
AL	24.50
PA	23.73
MA	21.51
SE	21.46
CE	21.20

Key Metrics: products with the highest canceled rates
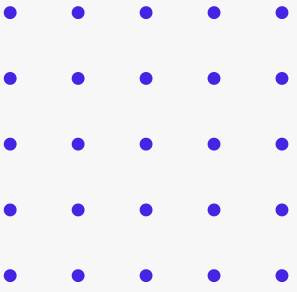
```sql
WITH ProductTotals AS (
  SELECT
    oi.product_id,
    COUNT(*) AS total_orders
  FROM order_items oi
  JOIN orders o ON oi.order_id = o.order_id
  GROUP BY oi.product_id
)
SELECT
  pt.product_id,
  COUNT(o.order_status) AS canceled_orders,
  (COUNT(o.order_status) * 1.0 / pt.total_orders)  AS cancellation_rate
FROM ProductTotals pt
JOIN order_items oi ON pt.product_id = oi.product_id
JOIN orders o ON oi.order_id = o.order_id
WHERE o.order_status = 'canceled'
GROUP BY pt.product_id, pt.total_orders
ORDER BY cancellation_rate DESC;
```

product_id	canceled_orders	cancellation_rate
d062241f27854330b57851b64d08d472	3	1.00000
b23cca4f37391d03408aed65983bdc40	1	1.00000
f52a0f70e54976873a4a7402349c6105	1	1.00000
5016388edcd982069aa2d39a8a878da7	1	1.00000
5548040d5b1cc71c61c03d2f1d2e6ede	1	1.00000
7d4f932dd4fa16256c20de073329eacc	5	1.00000
2a66698f3c9f8708a0505b5ab2cb380d	1	1.00000
5321c283dcbf309eac57110f5e306cf0	1	1.00000

Key Metrics: Query to find the monthly sales trend for the last two years

```sql
SELECT
    DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m') AS month_year,
    round(SUM(oi.price),2) AS total_sales
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
WHERE o.order_status = 'delivered'
GROUP BY DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m')
ORDER BY month_year desc;
```

month_year	total_sales
2018-08	838576.64
2018-07	867953.46
2018-06	856077.86
2018-05	977544.69
2018-04	973534.09
2018-03	953356.25
2018-02	826437.13
2018-01	924645

Key Metrics:Analyze the seasonality of sales to identify peak months

```sql
SELECT
    MONTH(o.order_purchase_timestamp) AS month,
    round(SUM(oi.price),2) AS total_sales
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
WHERE o.order_status = 'delivered'
GROUP BY MONTH(o.order_purchase_timestamp)
ORDER BY total_sales DESC;
```

month	total_sales
5	1466882.94
8	1393276.34
7	1349557.98
4	1314203.77
3	1312555.1
6	1278001.23
2	1060660.53
1	1036443.36

Key Metrics: window functions to rank products based on their sales within each category
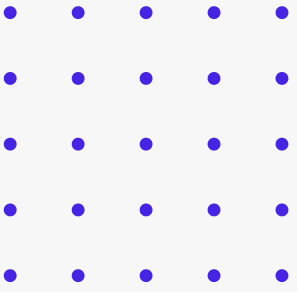
```sql
WITH ranked_products AS (
  SELECT
    p.product_category,
    p.product_id,
    SUM(pay.payment_value) AS total_sales,
    RANK() OVER (PARTITION BY p.product_category ORDER BY
SUM(pay.payment_value) DESC) AS rank_within_category
  FROM
    ecommerce.order_items oi
  JOIN
    ecommerce.products p ON oi.product_id = p.product_id
  JOIN
    ecommerce.orders o ON oi.order_id = o.order_id
  JOIN
    ecommerce.payments pay ON o.order_id = pay.order_id
  WHERE
    o.order_status = 'delivered'
  GROUP BY
    p.product_category, p.product_id
)
SELECT
  product_category,
  product_id,
  total_sales,
  rank_within_category
FROM
  ranked_products
ORDER BY
  product_category, rank_within_category;
```

product_id	total_sales	rank_within_category
5a848e4ab52fd5445cdc07aab1c40e48	27240.810455322266	1
b1d207586fca400a2370d50a9ba1da98	10986.669967651367	2
eed5cbd74fac3bd79b7c7ec95fa7507d	8200.239990234375	3
82758517d6c9cc65c509781f1e14e9f2	7466.7998046875	4
76d1a1a9d21ab677a61c3ae34b1b352f	6154.129974365234	5
ad88641611c35ebd59ecda07a9f17099	5304.989974975586	6
3b60d513e90300a4e9833e5cda1f1d61	5023.989929199219	7
4c50dcc50f1512f46096d6ef0142c4a9	4042.739990234375	8

Target Sale Analysis

Revenue

Customers

Operational

product_subcategory
All

Month
All

Year
All

Region
All

Total_revenue
13.59M

Average Order Value
136.68

Total_revenue by Region

- Southeast: 8.9M
- South: 2.0M
- Northeast: 1.5M
- Center-West: 0.9M
- North: 0.3M
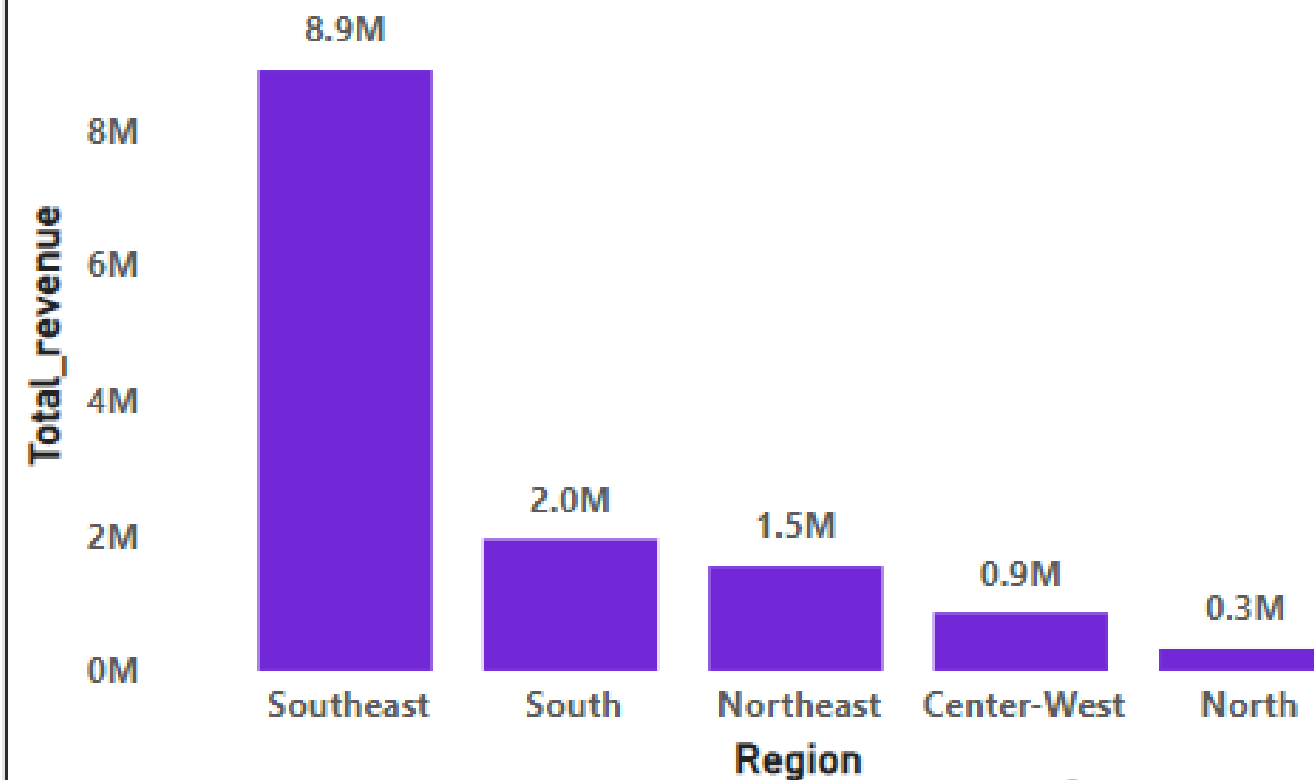
Top 10 product category by revenue
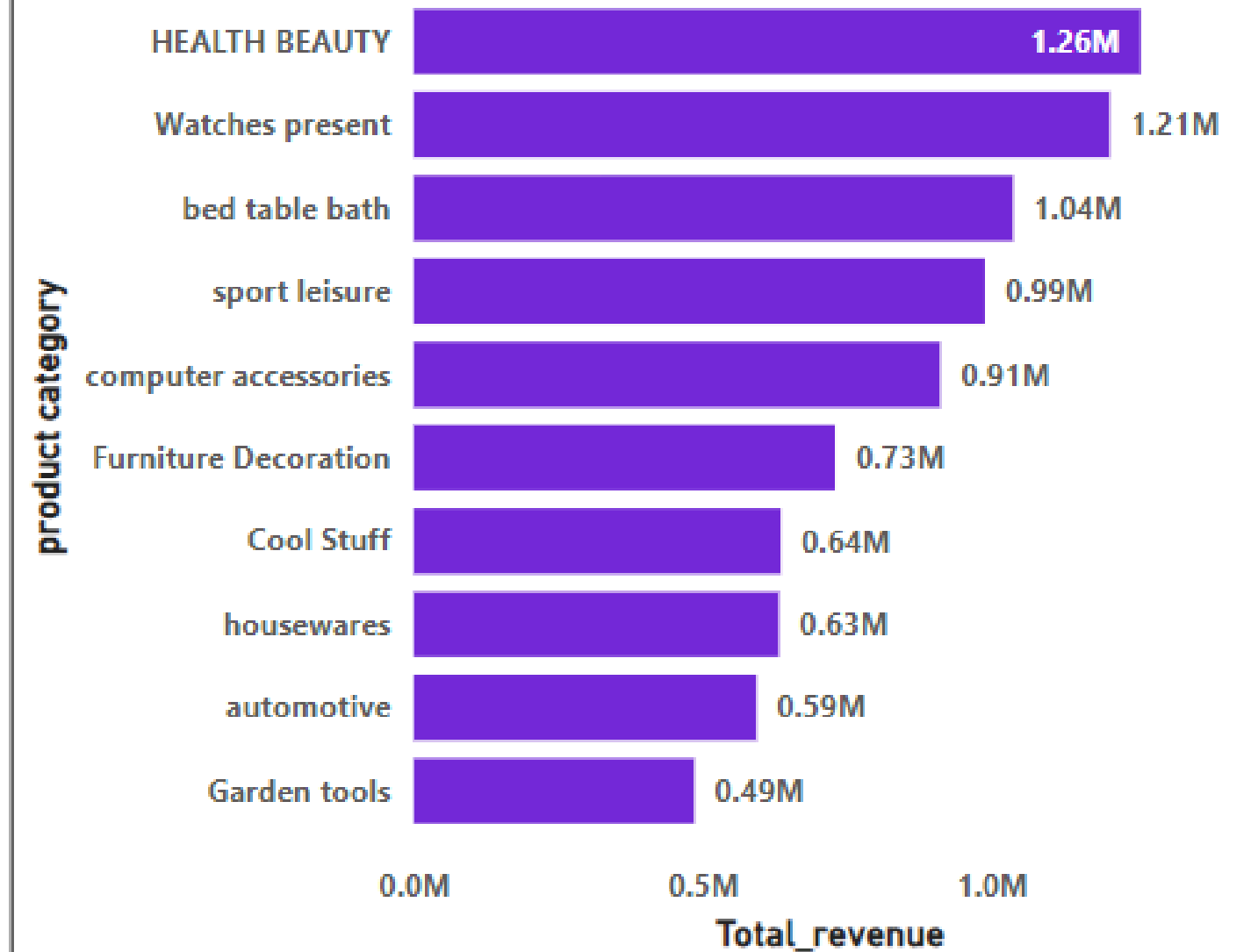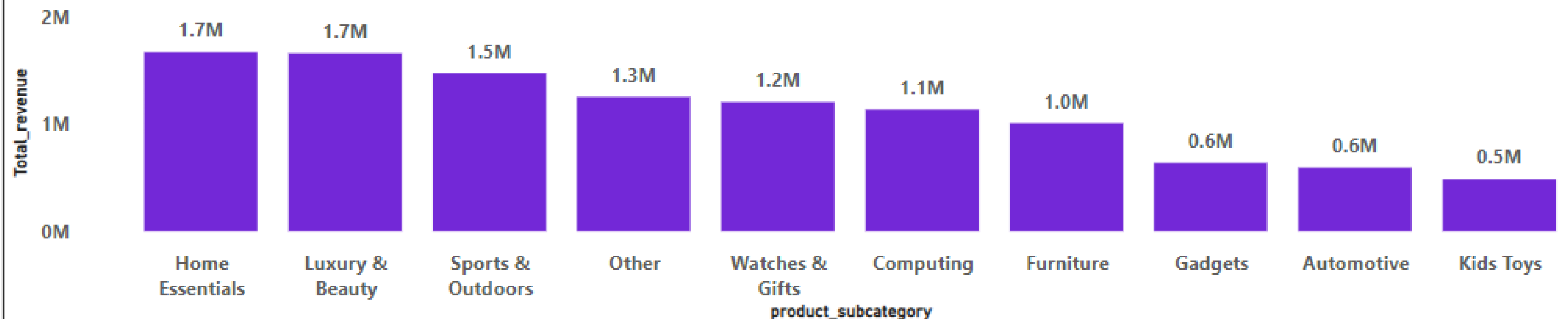
product category	Total_revenue
HEALTH BEAUTY	1.26M
Watches present	1.21M
bed table bath	1.04M
sport leisure	0.99M
computer accessories	0.91M
Furniture Decoration	0.73M
Cool Stuff	0.64M
housewares	0.63M
automotive	0.59M
Garden tools	0.49M

Top 10 product_subcategory by revenue

product_subcategory	Total_revenue
Home Essentials	1.7M
Luxury & Beauty	1.7M
Sports & Outdoors	1.5M
Other	1.3M
Watches & Gifts	1.2M
Computing	1.1M
Furniture	1.0M
Gadgets	0.6M
Automotive	0.6M
Kids Toys	0.5M

Target Sale Analysis

Revenue

Customers

Operational

product_subcategory
All

Month
All

Year
All

Region
All

## 132.96
average customer spend

99K
total customer

Total Spend by payment_type

Total Spend
10M
5M
0M

- credit_card: 10.6M
- UPI: 2.3M
- voucher: 0.4M
- debit_card: 0.2M

payment_type
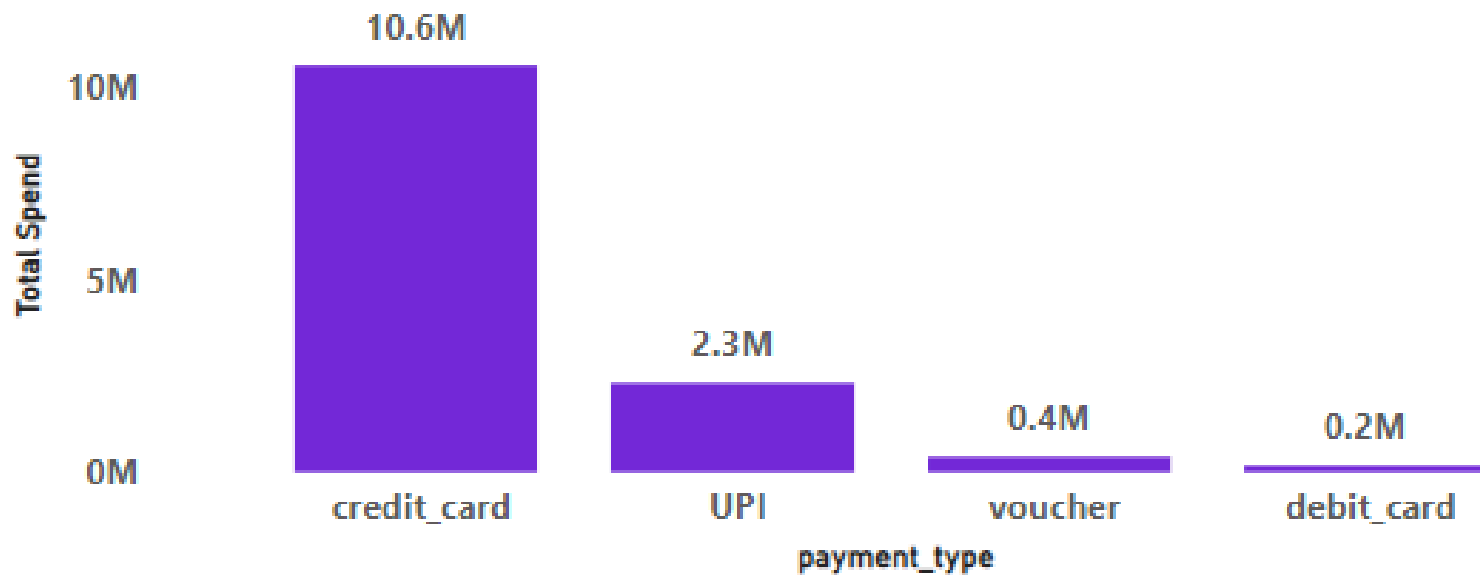
customer_id	Purchase Frequency	Total Spend	Customer Rank
1617b1357756262bfa56ab541c47bc16	1	13,440.00	1
ec5b2ba62e574342386871631fafd3fc	1	7,160.00	2
c6e2731c5b391845f6800c97401a43a9	1	6,735.00	3
f48d464a0baaea338cb25f816991ab1f	1	6,729.00	4
3fd6777bbce08a352fddd04e4a7cc8f6	1	6,499.00	5
05455dfa7cd02f13d132aa7a6a9729c6	1	5,934.60	6
df55c14d1476a9a3467f131269c2477f	1	4,799.00	7
24bbf5fd2f2e1b359ee7de94defc4a15	1	4,690.00	8
3d979689f636322c62418b6346b1c6d2	1	4,590.00	9
cc803a2c412833101651d3f90ca7de24	1	4,400.00	10
Total	**10**	**64,976.60**	**1**

Customer by Customer_Segment

0.02K (0.02%)

98.64K (99.97%)

Customer_Segment
- Low Spenders
- Medium Spenders
- High Spenders
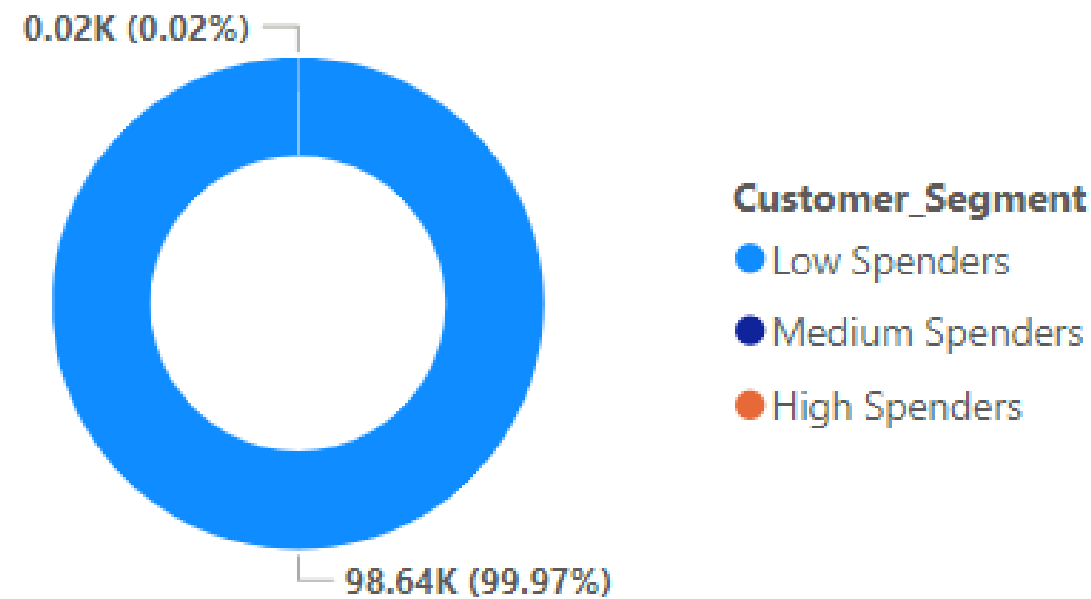
customer_id	CLV(customer lifetime values)
1617b1357756262bfa56ab541c47bc16	13,440.00
ec5b2ba62e574342386871631fafd3fc	7,160.00
c6e2731c5b391845f6800c97401a43a9	6,735.00
f48d464a0baaea338cb25f816991ab1f	6,729.00
3fd6777bbce08a352fddd04e4a7cc8f6	6,499.00
05455dfa7cd02f13d132aa7a6a9729c6	5,934.60
df55c14d1476a9a3467f131269c2477f	4,799.00
24bbf5fd2f2e1b359ee7de94defc4a15	4,690.00
3d979689f636322c62418b6346b1c6d2	4,590.00
cc803a2c412833101651d3f90ca7de24	4,400.00
1afc82cd60e303ef09b4ef9837c9505c	4,399.87
35a413c7ca3c69756cb75867d6311c0d	4,099.99
e9b0d0eb3015ef1c9ce6cf5b9dcbee9f	4,059.00
c6695e3b1e48680db36b487419fb0398	3,999.90
Total	**1,32,21,498.11**

Target Sale Analysis

Revenue

Customers

Operational

product_subcategory
All

Month
All

Year
All

Region
All

210
Max of Delivery Time(days)

12.50
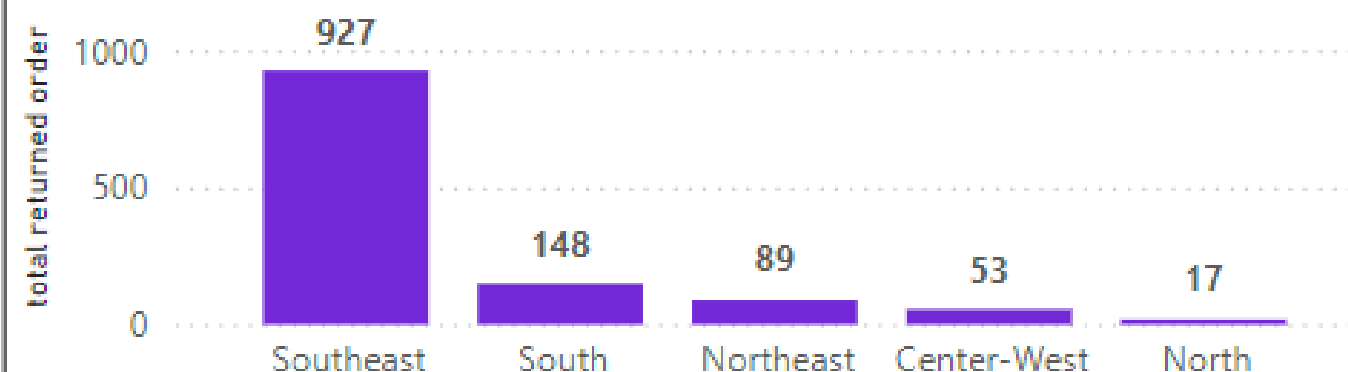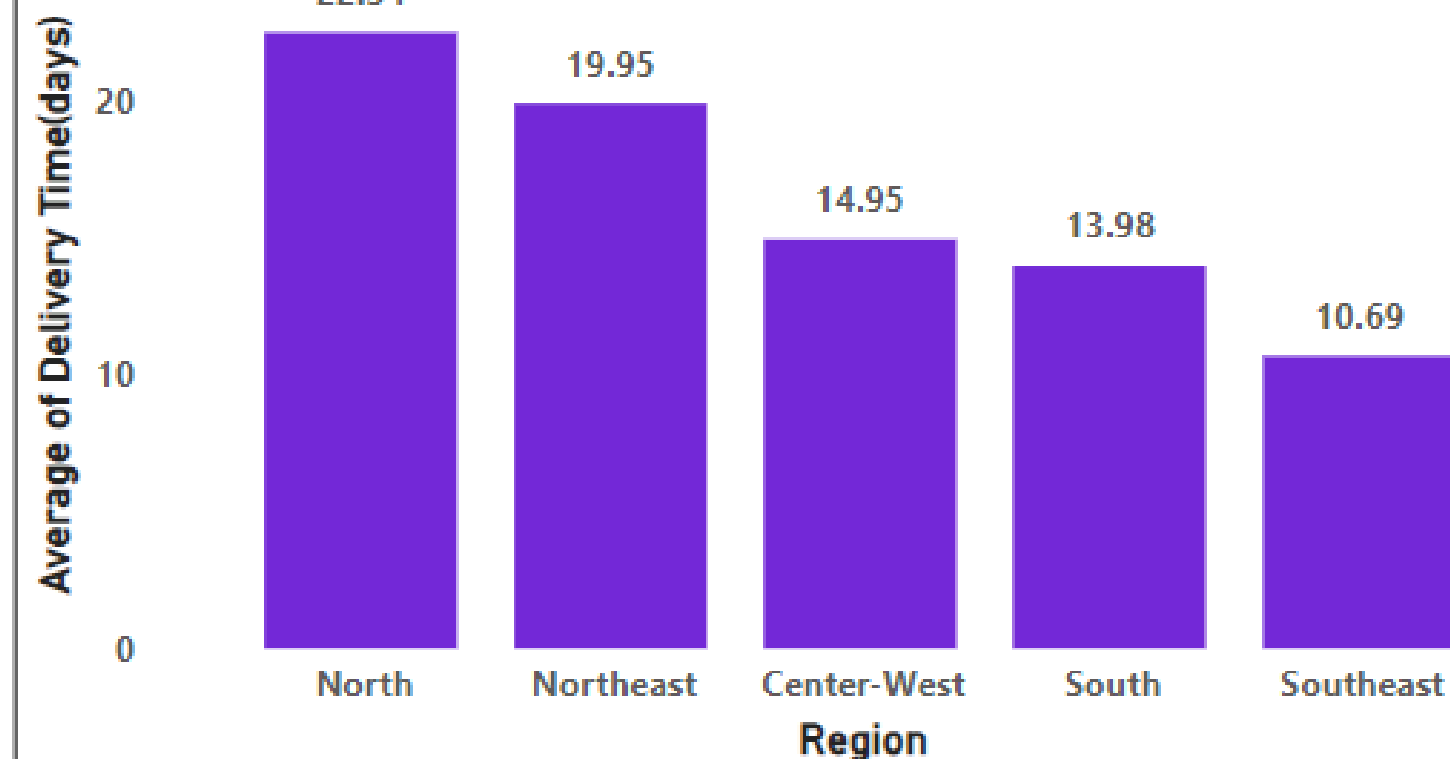average delivery time
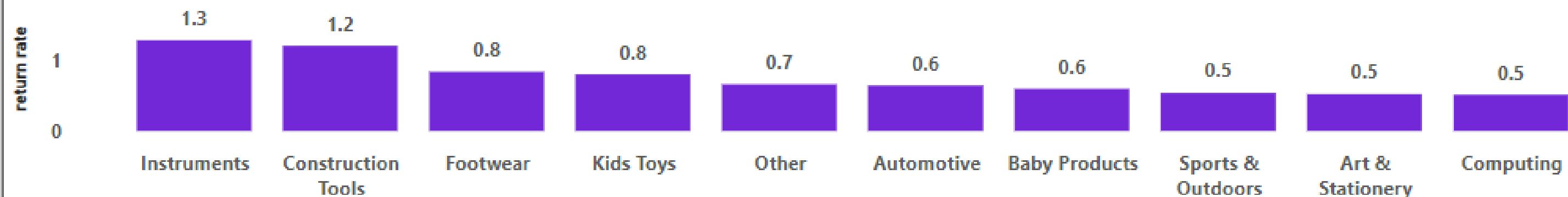
1234
Returned order

99K
total orders

0
Min of Delivery Time

total returned order by Region

927 Southeast
148 South
89 Northeast
53 Center-West
17 North

Average of Delivery Time(days) by Region

22.54 North
19.95 Northeast
14.95 Center-West
13.98 South
10.69 Southeast

Top 10 product_subcategory by return rate

1.3 Instruments
1.2 Construction Tools
0.8 Footwear
0.8 Kids Toys
0.7 Other
0.6 Automotive
0.6 Baby Products
0.5 Sports & Outdoors
0.5 Art & Stationery
0.5 Computing

Total_revenue by Year and Month

0.0M, 0.0M, 0.0M, 0.4M, 0.4M, 0.5M, 0.4M, 1.0M, 0.7M, 1.0M, 0.8M, 1.0M, 0.9M, 0.9M, 0.0M

Jan 2017   Jul 2017   Jan 2018   Jul 2018

Key Insights:

- **Revenue Insights:**
- **Total revenue: $13.59M**
- **The Southeast region contributed the highest revenue ($8.9M), while the North region contributed the least ($0.3M).**
- **Top revenue-generating product category: Health & Beauty ($1.26M).**
- **Top revenue-generating product subcategory: Home Essentials and Luxury & Beauty ($1.7M each).**
- **Customer Insights:**
- **Total customers: 99K.**
- **Average customer spend: $132.96.**
- **Majority of customers belong to the Low Spenders segment (99.97%).**
- **Top customer segment by spend: High Spenders contribute significantly more per transaction.**
- **Payment Insights:**
- **Most preferred payment method: Credit Card (10.6M).**
- **UPI transactions are significantly lower in comparison (2.3M).**
- **Operational Insights:**
- **Average delivery time: 12.5 days, with the North region experiencing the longest delays (22.54 days).**
- **Return orders: 1234 orders returned, with the Southeast leading the returns (927 returns).**
- **Top product subcategory with the highest return rate: Instruments (1.3%).**
- **Customer Lifetime Value (CLV):**
- **Highest CLV is $13,440, indicating a small set of high-value customers.**

Executive Summary

The sales analysis reveals that the Southeast region is the top revenue contributor, while the North region lags both in revenue and operational efficiency with the highest delivery delays. The Health & Beauty category generates the most revenue, while Instruments have the highest return rate. A significant portion of sales are made through credit card payments, and a small segment of high spenders accounts for a large portion of total revenue.

Recommendations:

- **Revenue Diversification:**
- **Focus marketing efforts on underperforming regions like the North and Center-West.**
- **Expand the Health & Beauty product category further due to its high revenue generation.**
- **Operational Efficiency:**
- **Address the long delivery times in the North region with potential logistic partnerships or process optimization.**
- **Investigate reasons behind the high return rate in the Southeast and for Instruments.**

- **Customer Retention:**
- **Develop targeted loyalty programs for high spenders and encourage repeat purchases.**
- **Increase awareness and discounts for medium spenders to move them up the value ladder.**
- **Payment Strategy:**
- **Promote UPI payments with cashback offers or discounts to diversify payment methods.**

Connect With me

www.linkedin.com/in/pankaj-dataanalyst

pankajkumardataanalyst@gmail.com

Github

Portfolio

Pankaj Kumar ✓

Quality Analyst | Excel, Power BI, Python, MySQL, Tableau | Aspiring Data Analyst & Data Analytics Mentee at WsCube Tech

Noida, Uttar Pradesh, India · **Contact info**

Data Analytics Project-Github 🔗