ECOMMERCE SALES AND CUSTOMER BEHAVIOUR ANALYSIS

MySQL | Power BI | DAX | Visualization

-by Ayush Vardhan

Introduction

Target Sales Company manages a vast dataset containing detailed information on orders, customers, products, regions, and operational metrics. The project focuses on extracting meaningful insights from this data to support strategic decision-making. Using advanced SQL queries for data extraction and visualization tools like Power BI, key business questions were addressed, trends were identified, and actionable insights were presented effectively to enhance decision- making processes.

Sales Performance Analysis :

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

```sql
SELECT
    p.product_category,
    c.customer_state,
    SUM(oi.price) AS Total_sales_revenue
FROM
    order_items oi
        JOIN
    products p ON p.product_id = oi.product_id
        JOIN
    orders o ON o.order_id = oi.order_id
        JOIN
    customers c ON c.customer_id = o.customer_id
WHERE
    o.order_status = 'delivered'
GROUP BY p.product_category , c.customer_state;
```

	product_category	customer_state	Total_sales_revenue
▶	sport leisure	SP	3750448.6950540543
	bed table bath	BA	249495.2992630005
	bed table bath	RJ	1442517.9023265839
	HEALTH BEAUTY	RJ	1402363.499970436
	bed table bath	SP	4722380.698928833
	automotive	GO	113092.59939193726

Sales Performance Analysis :

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

```sql
with cte as(
select
    p.product_id,
    p.product_category,
    count(oi.price) as qty_sold,
    sum(oi.price) as sales_revenue,
dense_rank() over(order by sum(oi.price) desc , count(oi.price) desc) as RA
from products p
    join order_items oi
    on p.product_id = oi.product_id
group by p.product_id,p.product_category
order by qty_sold desc, sales_revenue desc)

Select cte.* from cte
where cte.RA <= 5
order by RA;
```

product_id	product_category	qty_sold	sales_revenue	RA
bb50f2e236e5eea0100680137654686c	HEALTH BEAUTY	390	127770	1
6cdd53843498f92890544667809f1595	HEALTH BEAUTY	312	109460.39855957031	2
d6160fb7873f184099d9bc95e30376af	PCs	70	97798.67993164062	3
d1c427060a0f73f6b889a5c7c61f2ac4	computer accessories	686	94429.02098083496	4
99a4788cb24856965c36a24e339b6058	bed table bath	976	86051.12139892578	5

Customer Insights :

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

PF : (Total no of orders/ distinct(total no of customers)

```sql
SELECT
    c.customer_id,
    (COUNT(o.order_id) / COUNT(DISTINCT (c.customer_id))) AS purchase_freq,
    SUM(oi.price) AS total_spend
FROM
    customers c
        JOIN
    orders o ON o.customer_id = c.customer_id
        JOIN
    order_items oi ON oi.order_id = o.order_id
WHERE
    o.order_delivered_customer_date BETWEEN '2016-10-11 13:46:32' AND '2018-10-17 13:22:46'
GROUP BY customer_id
ORDER BY purchase_freq DESC , total_spend DESC
LIMIT 10;
```

customer_id	purchase_freq	total_spend
fc3d1daec319d62d49bfb5e1f83123e9	210.0000	318.0000114440918
be1b70680b9f9694d8c70f41fa3dc92b	200.0000	20000
bd5d39761aa56689a265d95d8d32b8be	200.0000	19739.999389648438
10de381f8a8d23fff822753305f71cae	150.0000	9823.49967956543
adb32467ecc74b53576d9d13a5a55891	150.0000	7650
d5f2b3f597c7ccafbb5cac0bcc3d6024	140.0000	8260

Customer Insights :

Identify customers with the highest average order value (AOV)

AOV = sum(Total revenue)/count(Total no of orders)

customer_id	AOV
1617b1357756262bfa56ab541c47bc16	134400
ec5b2ba62e574342386871631fafd3fc	71600
c6e2731c5b391845f6800c97401a43a9	67350
f48d464a0baaea338cb25f816991ab1f	67290
3fd6777bbce08a352fddd04e4a7cc8f6	64990
05455dfa7cd02f13d132aa7a6a9729c6	59345.99853515625
df55c14d1476a9a3467f131269c2477f	47990

```sql
SELECT
    c.customer_id,
    (SUM(oi.price) / COUNT(DISTINCT (o.order_id))) AS AOV
FROM
    customers c
        JOIN
    orders o ON o.customer_id = c.customer_id
        JOIN
    order_items oi ON oi.order_id = o.order_id
WHERE
    o.order_status = 'delivered'
GROUP BY c.customer_id
ORDER BY AOV DESC
LIMIT 10;
```

Operational Efficiency:

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

customer_state	customer_city	avg_del_time_in_days
AC	manoel urbano	12.0000
AC	epitaciolandia	13.0000
AC	senador guiomard	13.5000
AC	rio branco	21.0000
AC	cruzeiro do sul	23.6667
AC	xapuri	24.5000
AC	porto acre	29.0000

```sql
SELECT
    c.customer_state,
    c.customer_city,
    AVG(DATEDIFF(o.order_delivered_customer_date,
            o.order_purchase_timestamp)) AS avg_del_time_in_days
FROM
    customers c
        JOIN
    orders o ON o.customer_id = c.customer_id
WHERE
    o.order_delivered_customer_date IS NOT NULL
GROUP BY c.customer_state , c.customer_city
ORDER BY c.customer_state ASC , avg_del_time_in_days asc;
```

Operational Efficiency:

Identify regions or products with the highest delivery rates.

Deliver_rates = (No. of Orders delivered/ Total no pf orders places) * 100

Region	Delivery_Rate
MS	99.0232
AC	98.9130
AM	98.7879
AP	98.7805
ES	98.6259
RN	98.4877
PR	98.4146

```sql
with cte as
    (SELECT
        c.customer_state as Region,
        count(oi.order_id) AS delivered_orders
    FROM customers c
    JOIN orders o ON c.customer_id = o.customer_id
    JOIN order_items oi ON oi.order_id = o.order_id
    WHERE
        o.order_status = 'delivered'
    GROUP BY c.customer_state),

cte_total_orders as
    (Select
        c.customer_state as Region,
        count(oi.order_id) as Total_orders
    FROM customers c
    JOIN orders o ON c.customer_id = o.customer_id
    JOIN order_items oi ON oi.order_id = o.order_id
    group by c.customer_state)

Select cte.Region,
((cte.delivered_orders/cte_total_orders.Total_orders)* 100) as Delivery_Rate
from cte join cte_total_orders
on cte.Region = cte_total_orders.Region
order by Delivery_Rate desc;
```

Date and Time Analytics :

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

years	months	Monthly_sales
2018	1	1702236.2421512604
2018	2	1499941.1213374138
2018	3	1784260.2030334473
2018	4	2223944.96062088
2018	5	2017203.1197574139
2018	6	2006464.8793053627
2018	7	1592857.6992111206
2018	8	2283583.0795092583
2018	9	22938.0802192688
2018	10	550.7999992370605
2017	1	67198.23999977112
2017	2	397818.579536438
2017	3	671772.4192199707
2017	4	537374.4592638016
2017	5	1040282.898481369
2017	6	861559.1602668762
2017	7	910001.980609417
2017	8	1071059.8613443375
2017	9	1160536.719256878

```sql
with cte as
(select
year(o.order_delivered_customer_date) as years,
month(o.order_delivered_customer_date) as months,
sum(oi.price) as Monthly_sales
from orders o
join order_items oi
on o.order_id = oi.order_id
group by years, months
order by years asc, months asc),

rank_cte as
(Select cte.years,cte.months,cte.Monthly_sales,
dense_rank() over(order by cte.years desc) as RA
from cte)

select rank_cte.years, rank_cte.months,
rank_cte.Monthly_sales from rank_cte
where rank_cte.RA <=2
order by rank_cte.RA asc;
```

Date and Time Analytics :

Analyze the seasonality of sales to identify peak months.

	years	months	Sales	RA
▶	2016	10	60668.25993347168	1
	2016	11	19988.359790802002	2
	2016	12	1517.7199821472168	3
	2017	12	1889985.8000946045	1
	2017	10	1315130.5392332077	2
	2017	11	1302915.041929245	3
	2018	8	2283583.0795092583	1
	2018	4	2223944.96062088	2
	2018	5	2017203.1197574139	3

```
with cte as
(select year(o.order_delivered_customer_date) years,
month(o.order_delivered_customer_date) as months,
sum(oi.price) as Sales
from orders o
join order_items oi
on oi.order_id = o.order_id
where year(o.order_delivered_customer_date) is not null
group by years, months),

cte_rank as
(Select cte.*,
dense_rank() over(partition by cte.years order by cte.Sales desc) as RA
from cte
order by cte.years asc)

Select cte_rank.* from cte_rank
where RA <= 3;
```

Advanced SQL Queries :

Analyze the seasonality of sales to identify peak months.

```sql
with cte as
(select p.product_category as category, p.product_id,
sum(oi.price) as sales
from products p
join order_items oi
on oi.product_id = p.product_id
where p.product_category is not null
group by  p.product_category, p.product_id
order by p.product_category asc, sales desc)

select cte.*, dense_rank()
 over(partition by cte.category order by cte.sales desc) as RA from cte;
```

	category	product_id	sales	RA
▶	Agro Industria e Comercio	11250b0d4b709fee92441c5f34122aed	18222	1
	Agro Industria e Comercio	423a6644f0aa529e8828ff1f91003690	16086	2
	Agro Industria e Comercio	672e757f331900b9deea127a2a7b79fd	13770	3
	Agro Industria e Comercio	c183fd5d2abf05873fa6e1014ed9e06c	11869.19970703125	4
	Agro Industria e Comercio	2b69866f22de8dad69c976771daba91c	5980	5
	Agro Industria e Comercio	c89226b8a795ae3d6bca9d90b20dbf04	5643	6
	Agro Industria e Comercio	5fb0955cb683eb6f65a1f613e502eef5	5440	7
	Agro Industria e Comercio	b7a60a397d4efd05c1b5d398fb9f9097	4798	8
	Agro Industria e Comercio	cd5df6a3db7a3d064a55afd08289d762	4720	9
	Agro Industria e Comercio	cd2f5c10e4e8dbc701f0bb68a09fdfe8	4398	10

Advanced SQL Queries :

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

```sql
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;
```

order_id	dates	years	months	price	YTD	YTD_Frame	MTD	MTD_Frame
ec7a019261fce44180373d45b442d78f	2017-01-11	2017	1	10.9	21.799999237060547	10.899999618530273	21.799999237060547	10.899999618530273
ec7a019261fce44180373d45b442d78f	2017-01-11	2017	1	10.9	21.799999237060547	21.799999237060547	21.799999237060547	21.799999237060547
e1fe072ef14b519af1f0a8ed997c1301	2017-01-12	2017	1	9.9	84.55999755859375	31.69999885559082	84.55999755859375	31.69999885559082
e1fe072ef14b519af1f0a8ed997c1301	2017-01-12	2017	1	9.9	84.55999755859375	41.599998474121094	84.55999755859375	41.599998474121094
bf44408de1d05eef70fb026ba64a90ea	2017-01-12	2017	1	10.49	84.55999755859375	52.08999824523926	84.55999755859375	52.08999824523926

Key Findings :

Regional Insights: States with high delivery rates correlate with higher customer satisfaction and repeat orders.

- **Top Performers:** The top 5 best-selling products drive a substantial percentage of overall revenue.

- **Customer Behavior:** High AOV customers are often also the most loyal.

- **Sales Trends:** Consistent monthly growth in sales revenue observed in the last two years.

- **Efficiency:** Early identification of peak months allows strategic inventory planning.

Sales Performance Analysis :

Year
(Blank) | 2016 | 2017 | 2018

Weekday or weekend
(Blank) | WeekDay | Weekend
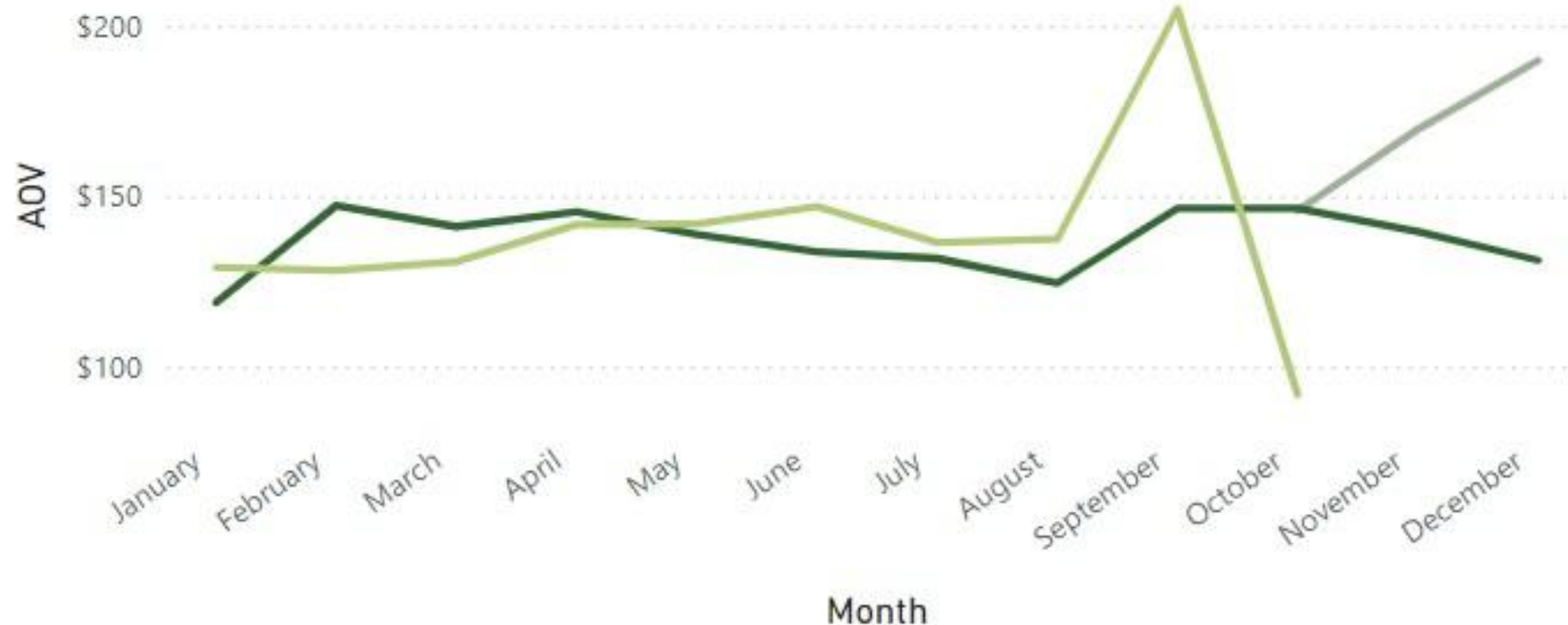
customer_state
All

product category
All

AOV by Month and Year

Year ● 2016 ● 2017 ● 2018
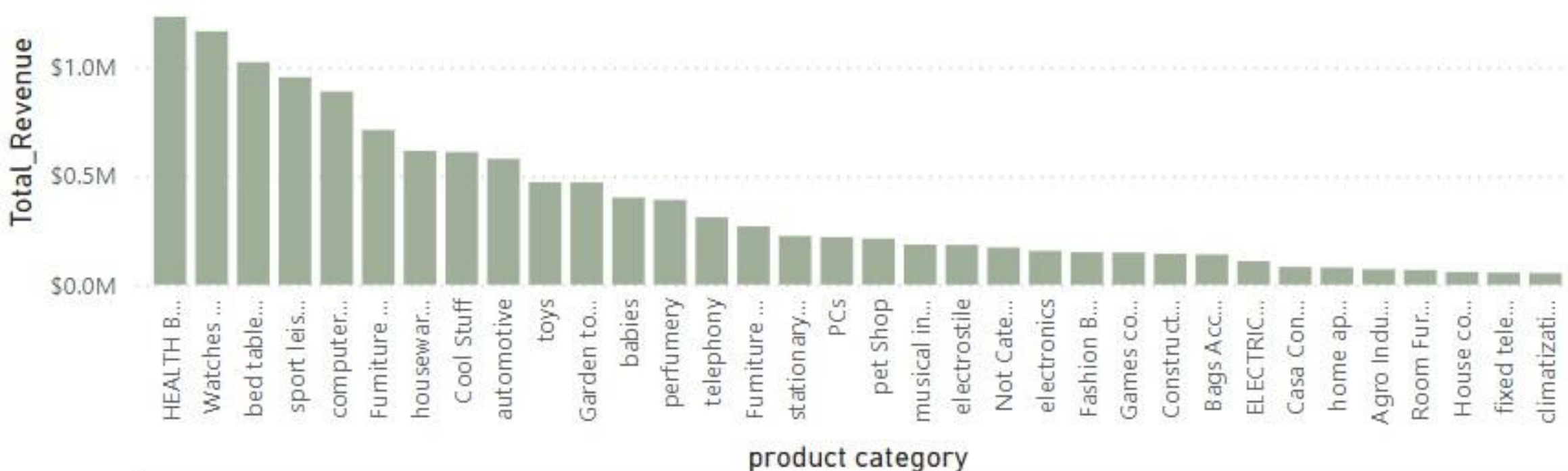

$13.22M
Total Revenue
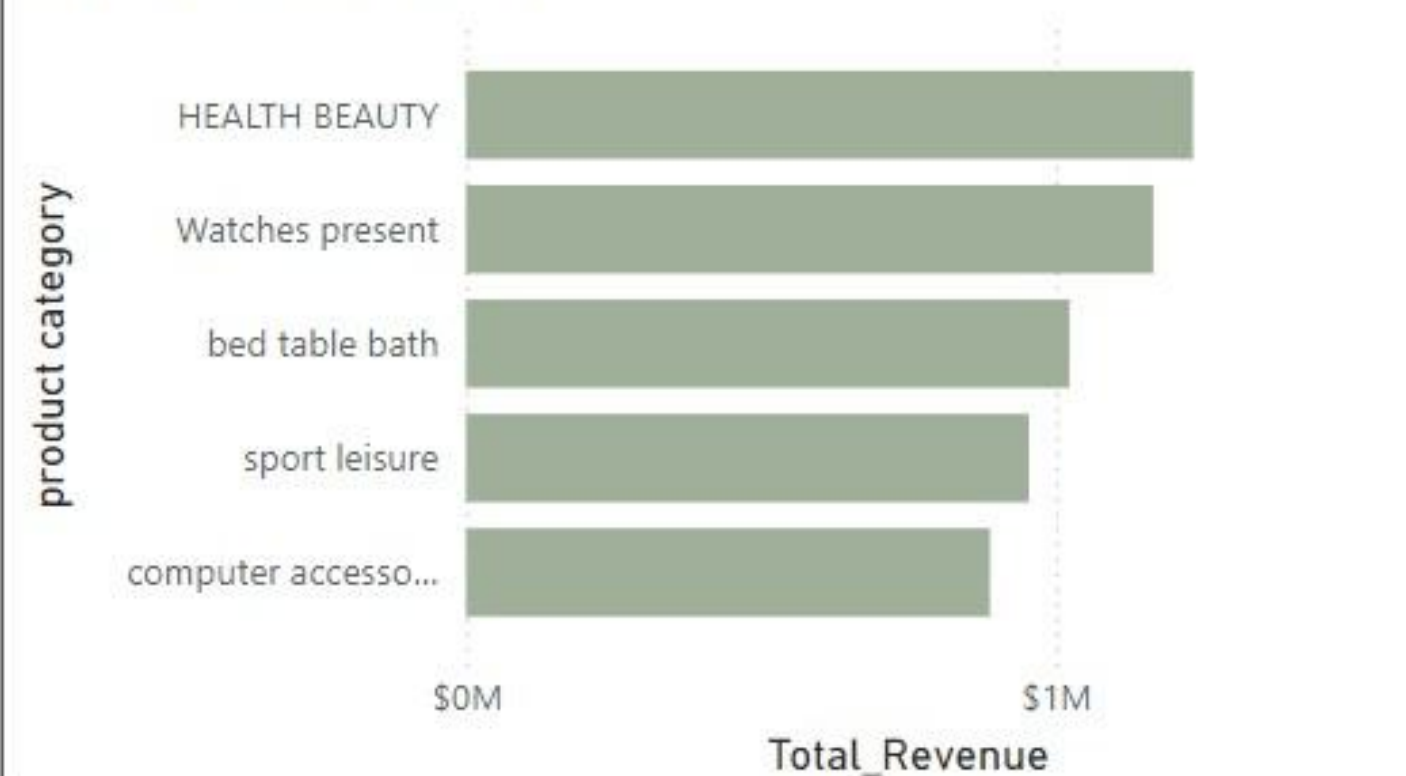
$137.04
AOV

HEALTH BEAUTY
Top Product Category

Total Revenue by Product Category

Top 5 products by Sales
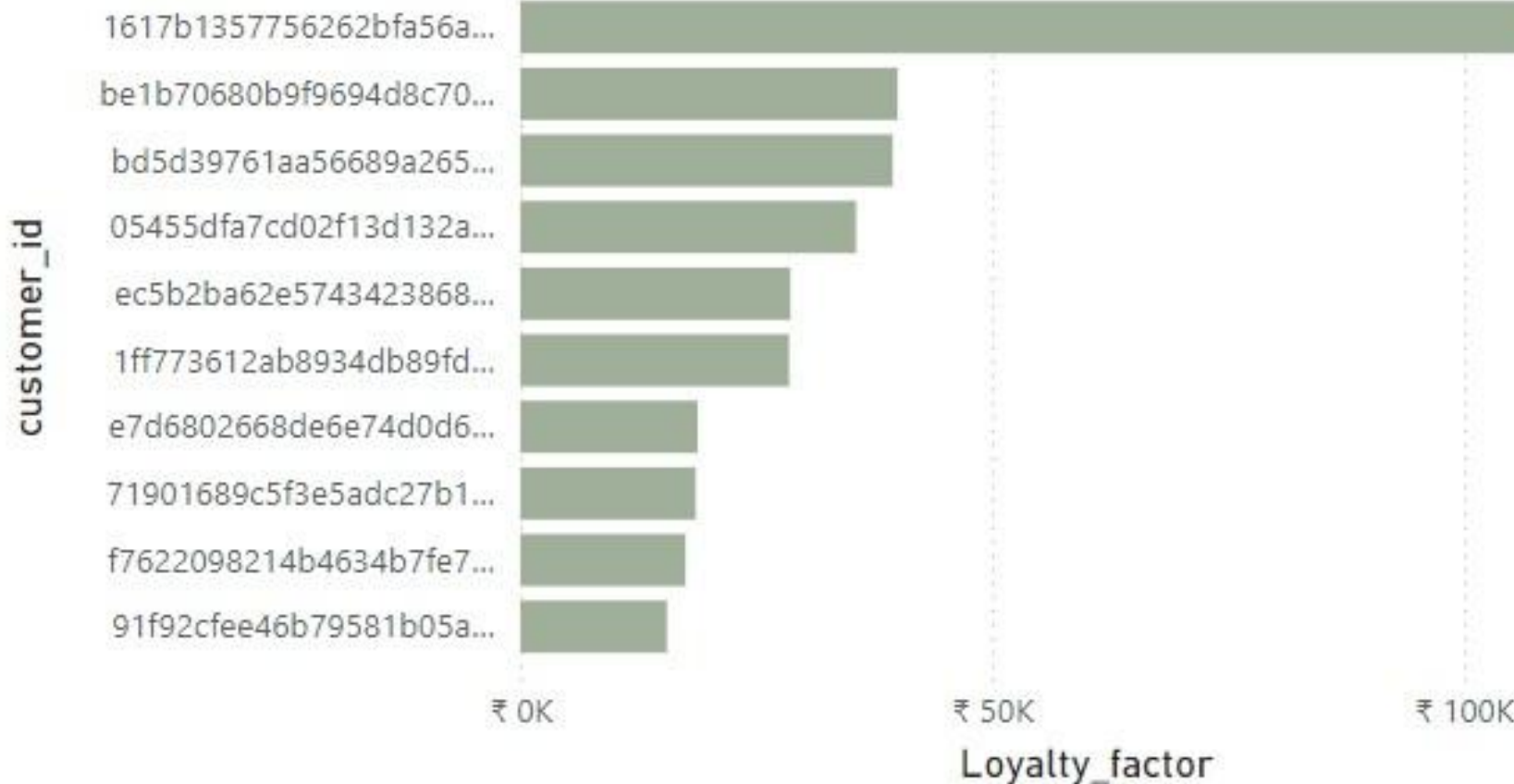
Customer Insights :

Year

(Blank)	2016	2017	2018

customer_state

All

product category

All

Loyalty_factor by customer_id

customer_id	
1617b1357756262bfa56a...	
be1b70680b9f9694d8c70...	
bd5d39761aa56689a265...	
05455dfa7cd02f13d132a...	
ec5b2ba62e5743423868...	
1ff773612ab8934db89fd...	
e7d6802668de6e74d0d6...	
71901689c5f3e5adc27b1...	
f7622098214b4634b7fe7...	
91f92cfee46b79581b05a...	

₹ 0K   ₹ 50K   ₹ 100K

Loyalty_factor

$13.59M
Spend_per_customer

113K
purchase_freq

Total_Revenue by customer_id

$10K

$5K

$0K

Total_Revenue

1617b135... ec5b2ba6... c6e2731c... f48d464 a... 3fd6777b... 05455dfa... df55c14d... 24bbf5fd2... 3d979689... cc803a2c4... 1afc82cd6... 35a4113c7... e9b0d0eb... c6695e3b... 926b6a6f... 3be2c536... 31e83c01f... eb7a157e... 19b32919... 66657bf1... addc91fdf... 7d03bf20f... 39d66580... e7c905bf4... 0b16ce67... a7ea318c... 46bb3c0b...

customer_id

customer_id	Spend_per_customer	purchase_freq	Loyalty_factor
bd5d39761aa56689a265d95d8d32b8be	$1,974	20	₹ 39,480
be1b70680b9f9694d8c70f41fa3dc92b	$2,000	20	₹ 40,000
91f92cfee46b79581b05aa974dd57ce5	$1,296	12	₹ 15,552
1ff773612ab8934db89fd5afa8afe506	$2,849.9	10	₹ 28,499
e7d6802668de6e74d0d6c56565bf2a24	$1,879.9	10	₹ 18,799
1617b1357756262bfa56ab541c47bc16	$13,440	8	₹ 1,07,520
05455dfa7cd02f13d132aa7a6a9729c6	$5,934.6	6	₹ 35,607.6
71901689c5f3e5adc27b1dd16b33f0b8	$3,099.75	6	₹ 18,598.5
f7622098214b4634b7fe7eee269b5426	$2,919.4	6	₹ 17,516.4
ec5b2ba62e574342386871631fafd3fc	$7,160	4	₹ 28,640

Regional Analysis :

Year

| (Blank) | 2016 | 2017 | 2018 |

customer_state

All

product category

All

Orders delivered by Customer state

customer_state
- SP
- RJ
- MG
- RS
- PR
- SC
- BA
- DF
- GO
- ES

0K
1K (0.96%) (0.28%)
2K (2.02%)
2K (2.14%)
4K (3.34%)
4K (3.72%)
6K (5.13%)
6K (5.57%)
13K (11.72%)
14K (12.83%)
46K (42.15%)
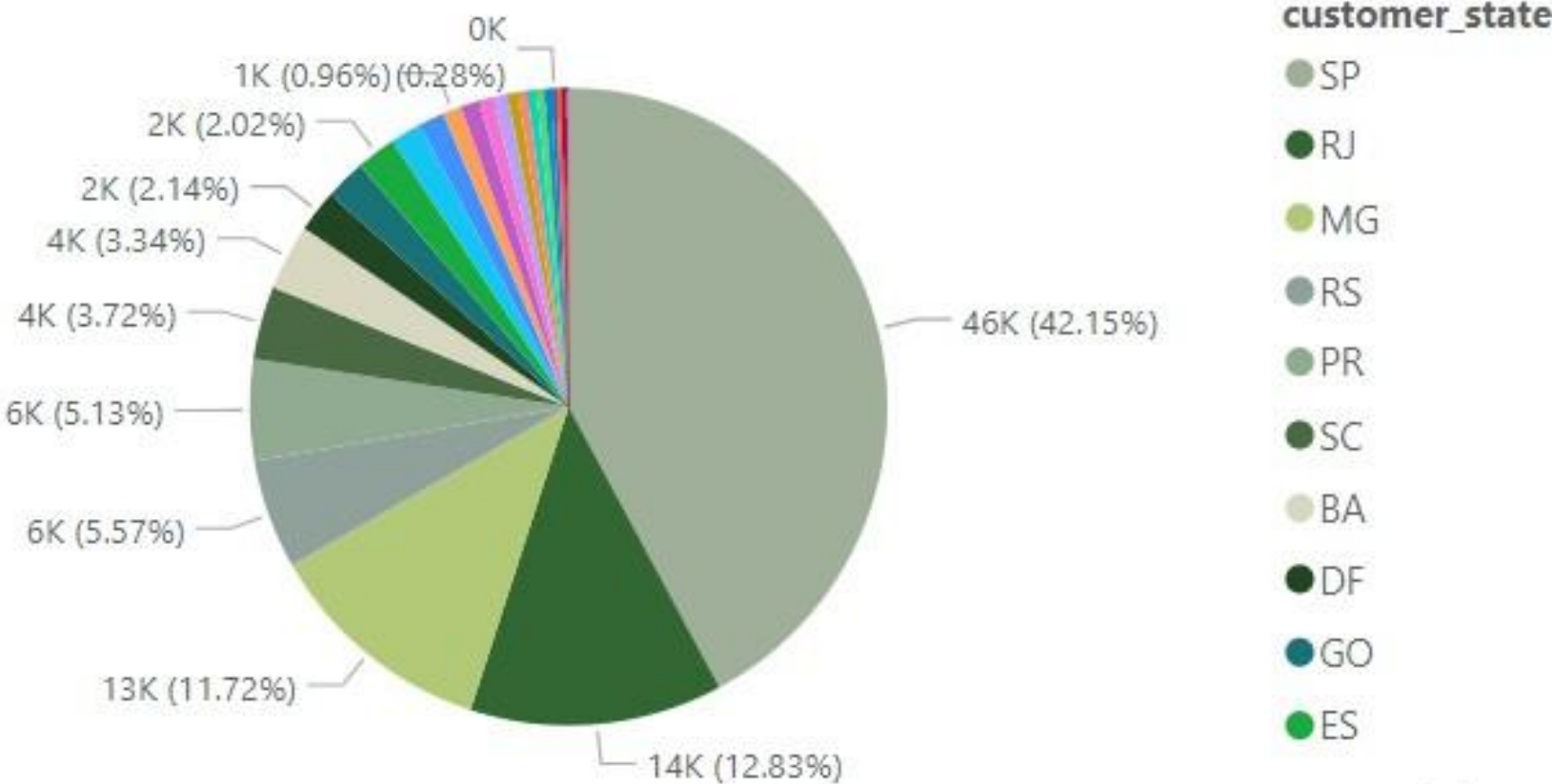
$13.22M
Total Revenue

12.50
Average_Delivery_time

Total Revenue by Region

Operational Metrics :

Year	**customer_state**
All ⌄	All ⌄

product category	products_cancelled
sport leisure	51
housewares	49
computer accessories	46
Furniture Decoration	36
HEALTH BEAUTY	36
toys	34
automotive	30
Watches present	21
babies	20
Garden tools	19
bed table bath	18
telephony	18
Cool Stuff	16
perfumery	15
Not Categorized	14
stationary store	12
electronics	11
musical instruments	11
Games consoles	10
Total	**542**

Canceled Product % by Product Category

PC Gamer 21.21%
Kitchen porta... 12.73%
Blu Ray DVDs 5.97%
CONSTRUCTION SECURITY TOO... 4.92%
Construction Tools Garden 3.21%
ELECTRICES 2 2.41%
fixed telephony 2.17%
Not Categoriz... 1.67%
Fashion Calc... 1.46%
Drin... 1.37%
Furni... 1.36%
housewares 1.34%
Const... 1.26%
babies 1.25%
computer accessor... 1.12%
Garden tools 0.83%
electronics 0.76%
Fashion Bags and ... 0.56%
foods 0.37%
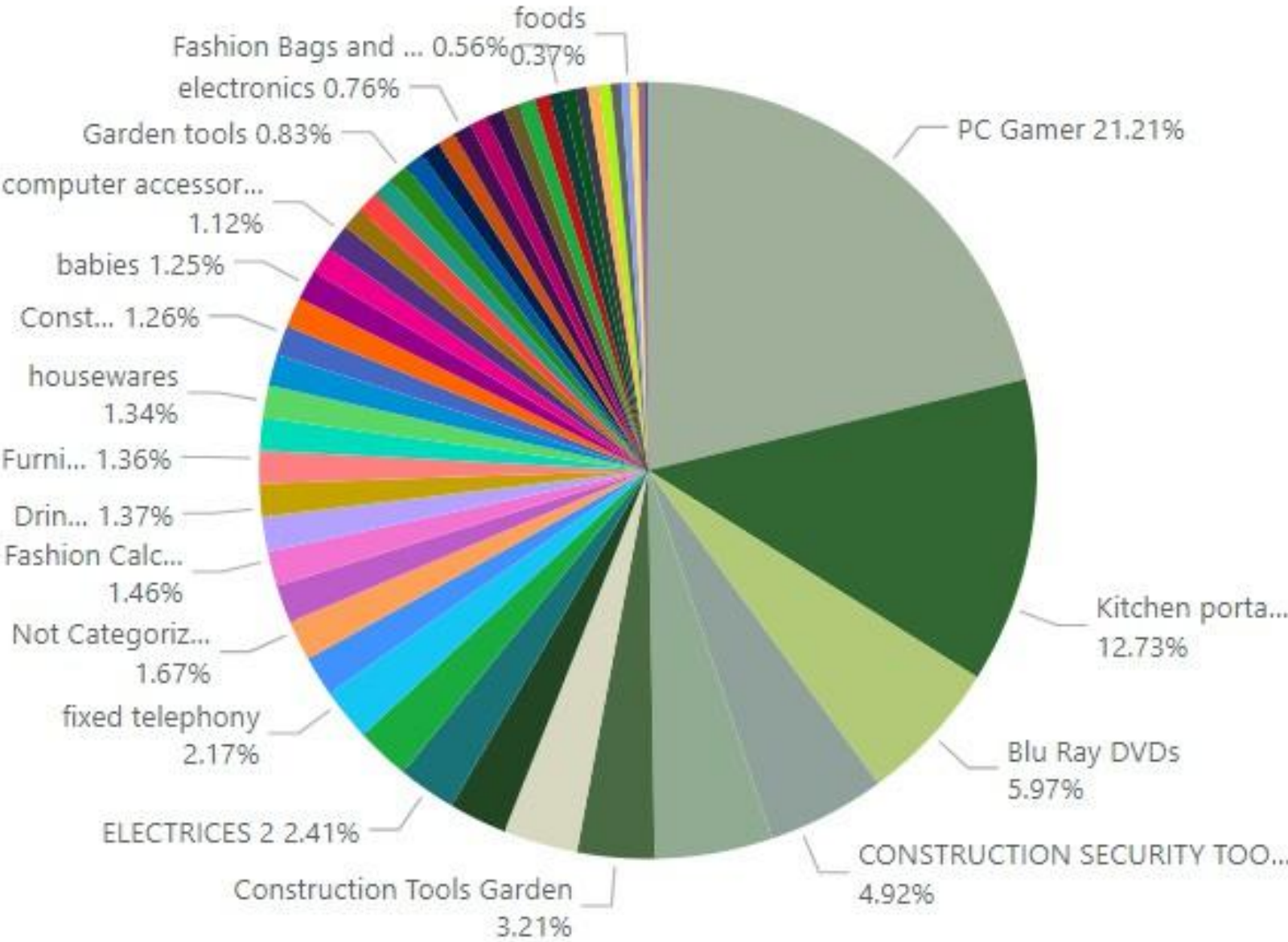
113K	Total_products_ordered
542	products_cancelled
12.50	Average_Delivery_time
210	Max_delivery_time
13	Min_delivery_time

Total_Revenue by Month and Year

Year ● 2016 ● 2017 ● 2018

Year

All

customer_state

All

Monthname

All
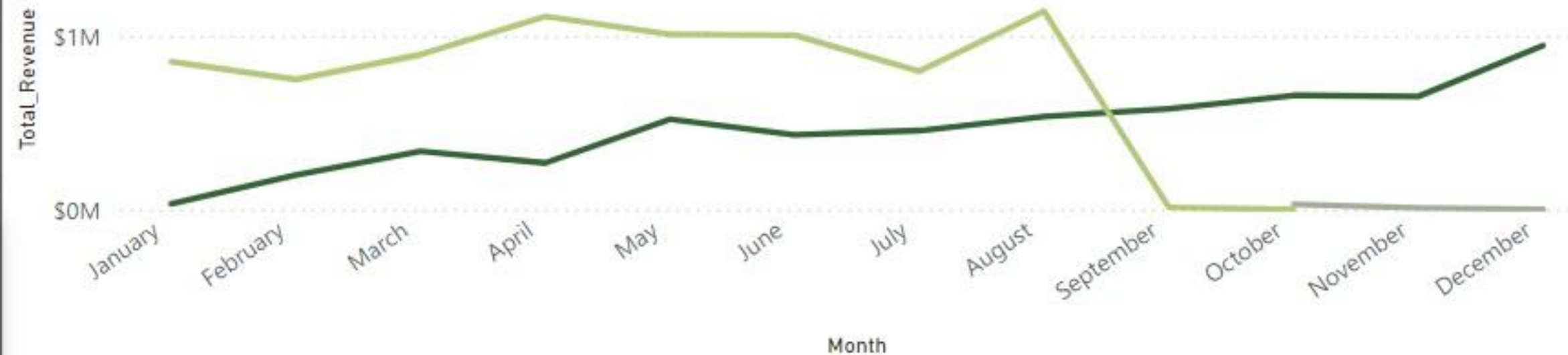
2.26
YOY_growth

0.00%
MOM_growth

Total_Revenue by Weekday_or_weekend

Year	Total_Revenue	MOM_growth
⊞		-100.00%
⊟ 2016		
January		-100.00%
February		-100.00%
March		-100.00%
April		-100.00%
May		-100.00%
June		-100.00%
July		-100.00%
August		-100.00%
September		-100.00%
October	$29,874.44	-100.00%
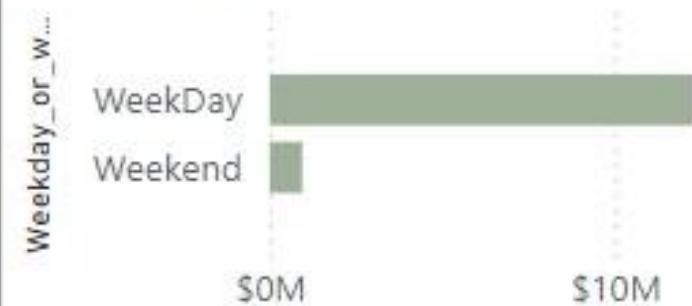
Year	Total_Revenue	YOY_growth
2017	$56,12,817.85	137.69
2018	$75,66,960.1	0.88
		-1.00
2016	$40,470.98	-1.00

Total_Revenue by Weekday
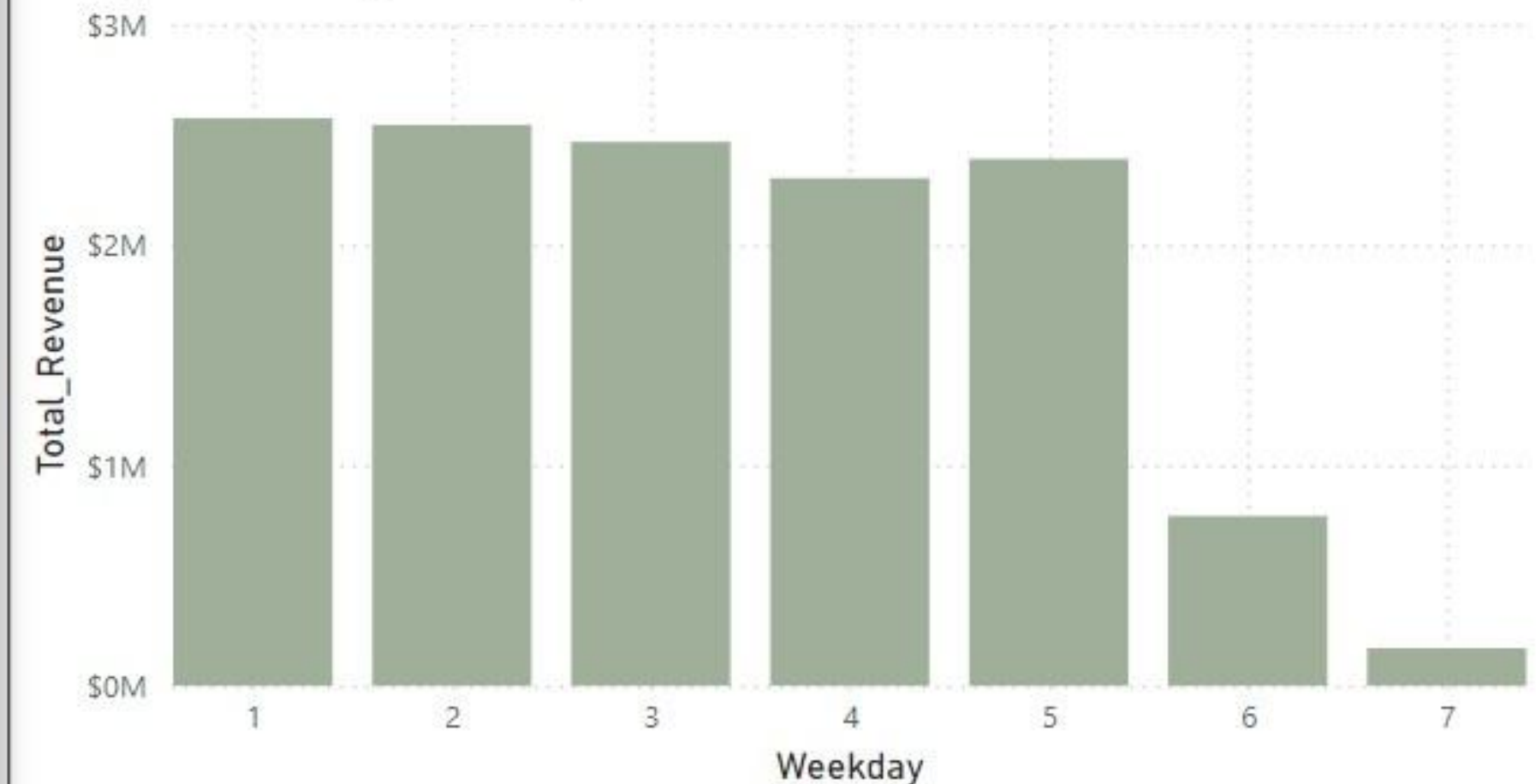
Key Findings

Revenue Insights:

Total revenue: $13.22M, with an Average Order Value (AOV) of $137.04.

Health and Beauty is the top product category, contributing the most to revenue.

Customer Insights:

- Loyal customers show high purchase frequency and contribute significantly to revenue.
- The top 10 customers were identified using the Loyalty Factor, combining purchase frequency and spending metrics.

Operational Efficiency:

- Average delivery time: 12.50 days, with delays in some regions.
- Peak sales months and weekday/weekend patterns identified for targeted planning.

Order Insights:

- Total products ordered: 113K, with 542 cancellations, highlighting areas for improvement in order handling.
- customers.
- Strategic seasonal marketing campaigns during peak months to maximize sales.