

Case Study: Optimizing Sales and Order Fulfilment for an E-commerce Business

1. Title: Leveraging Data to Boost Sales and Customer Retention in E-commerce

2. Introduction:

The e-commerce company faced challenges in understanding sales performance, managing order fulfilment, and identifying trends in customer behaviour. By analysing sales, fulfilment, and inventory data, the goal was to uncover actionable insights to improve business operations, increase revenue, and optimize the customer experience.

3. Objective:

The primary objectives of this data analysis project were:

- To assess overall sales performance over a specific period.
- To evaluate fulfilment efficiency across various sales channels.
- To analyse customer ordering behaviour and geographic trends.
- To identify top-selling products, revenue-generating categories, and order fulfilment issues.
- To offer actionable recommendations based on data insights.

4. Methodology:

1. Data Source:

The dataset is derived from Amazon e-commerce transactions, including details like order IDs, fulfilment status, sales amounts, shipping locations, and SKUs.

2. Tools Used:

- **SQL** and **BigQuery** were used for data extraction, transformation, and analysis.
- **Google Looker Studio** was used to generate data visualizations for easier interpretation.

5. Data Loading into BigQuery

- **Create a Google Cloud Project** and set up a **BigQuery Dataset**

6. Data Cleaning and Transformation (ETL)

Clean the dataset for missing values, inconsistent data, and prepare for analysis.

1. Renaming columns:

-- Since hyphen is recognised by SQL, renaming columns

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `promotion-ids` TO promotion_ids;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `ship-city` TO ship_city;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `ship-state` TO ship_country;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `ship-postal-code` TO ship_postal_code;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `ship-country` TO ship_country;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `fulfilled-by` TO fulfilled_by;
```

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `Order ID` TO Order_ID
```

2. Identify missing values:

```
SELECT *  
FROM `new-project-439116.Ecommerce_project.1`  
WHERE promotion_ids IS NULL  
AND fulfilled_by IS NULL  
AND Amount IS NULL;
```

3. Replace null values in certain columns:

```
UPDATE `new-project-439116.Ecommerce_project.1`  
SET currency = '-1',  
Amount = -1,  
promotion_ids = '-1',  
fulfilled_by = '-1'  
  
WHERE currency IS NULL  
AND Amount IS NULL  
AND promotion_ids IS NULL  
AND fulfilled_by IS NULL
```

7. Questions Addressed:

The key business questions addressed during the analysis included:

- What is the total sales revenue for the year 2022?
- How many orders were successfully fulfilled, and how many were returned or canceled?
- Which sales channels generate the most revenue?
- Which product categories have the highest revenue and order volume?
- Are there geographic trends in order volume?
- What is the order turnover rate for different products?

8. Analysis & Insights:

1. **Total Sales Revenue in 2022:** The total sales revenue generated in 2022 amounted to **INR 78,586,959.30**.

```
SELECT SUM(Amount) AS TotalSalesRevenue
FROM `new-project-439116.Ecommerce_project.1`
WHERE Date BETWEEN '2022-01-01' AND '2023-01-01';
```

QUERY:

https://docs.google.com/spreadsheets/d/1ImBtjuKOWuCPMhpXs8MTErIsQ7TV_6F0_bNIQfNco8/edit?gid=281216554#gid=281216554

2. **Fulfillment Status:** Analyzing the order status, it was found that **92%** of orders were successfully fulfilled, while **8%** were returned or canceled.

- How does the fulfillment status vary across different sales channels?

```
SELECT promotion_ids, Status, COUNT(*) AS OrderCount
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY promotion_ids, Status;
```

QUERY:

<https://docs.google.com/spreadsheets/d/1UInn5QAmwBbB61yuNALgwdO8v3l8PsouBaE-OtN0-h8/edit?gid=1718569770#gid=1718569770>

```
SELECT Status, COUNT(*) AS OrderCount
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY Status;
```

QUERY:

https://docs.google.com/spreadsheets/d/1C_nq6uMErdNn1gNytX22yr2dh0gGBRapHmyy6u9NfTQ/edit?gid=1897607255#gid=1897607255

VISUALIZATION: <https://lookerstudio.google.com/s/nRfNN9eyxs4>

3. **Sales Channel Insights:** The highest revenue was generated through **IN Core Free Shipping 2015/04/08 23-48-5-108** with a total revenue of **INR 31,815,155.00**.

```
SELECT promotion_ids, SUM(Amount) AS TotalRevenue
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY promotion_ids
ORDER BY TotalRevenue DESC
LIMIT 1;
```

Row	promotion_ids	TotalRevenue
1	IN Core Free Shipping 2015/04/08 23-48-5-108	31815155.0

4. **Top Product Categories:** The highest sales volume and revenue were generated by the **SET** category.

```
SELECT Category, SUM(Amount) AS TotalRevenue, COUNT(*) AS TotalOrders
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY Category
ORDER BY TotalRevenue DESC;
```

QUERY:

<https://docs.google.com/spreadsheets/d/1uyZVjsg29LhipollzNGhiqbRBpoXdWgwDAGI5UQO2gM/edit?gid=1389562928#gid=1389562928>

VISUALIZATION: <https://lookerstudio.google.com/s/hYhQv9cCSiQ>

- **How do product sizes affect sales performance?**

```
SELECT Size, SUM(Amount) AS TotalRevenue, COUNT(*) AS TotalOrders
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY Size
ORDER BY TotalRevenue DESC;
```

QUERY:

<https://docs.google.com/spreadsheets/d/12H3dNzJg5t3taCKU1gWB4maMm8njqaEhsaOjHFC-7ik/edit?gid=1017740885#gid=1017740885>

VISUALIZATION: <https://lookerstudio.google.com/s/jvOMHirHYHs>

5. **Geographic Trends:** The majority of orders were shipped to **Bengaluru**, with the highest order count in the country.

```
SELECT Category, AVG(Amount) AS AverageOrderValue
FROM `new-project-439116.Ecommerce_project.1`
GROUP BY Category
ORDER BY AverageOrderValue DESC;
```

QUERY:

https://docs.google.com/spreadsheets/d/1aOcEnjINn1PHPnct3ZQsmkBkUu0QNs4yfupqN0_sgiw/edit?gid=1160085017#gid=1160085017

- **What are the most common shipping locations (city, state, country)?**

```
ALTER TABLE `new-project-439116.Ecommerce_project.1`  
RENAME COLUMN `ship_country` TO ship_state;
```

```
SELECT ship_city, ship_state, COUNT(*) AS OrderCount  
FROM `new-project-439116.Ecommerce_project.1`  
GROUP BY ship_city, ship_state  
ORDER BY OrderCount DESC  
LIMIT 100;
```

<https://docs.google.com/spreadsheets/d/1UE9DEmPf0fBg5hEpHniwhmascMAidhowJbOID4kNVJM/edit?gid=1704342884#gid=1704342884>

<https://lookerstudio.google.com/s/IQ0Q2K012E8>

- Are there trends in order volume based on geographic location?

```
SELECT ship_state, DATE(Date) AS OrderDate, COUNT(*) AS OrderCount  
FROM `new-project-439116.Ecommerce_project.1`  
GROUP BY ship_state, OrderDate  
ORDER BY OrderCount DESC;
```

QUERY:

<https://docs.google.com/spreadsheets/d/1Ld-2CRsEaE7ZS4-zSloYobSXgp0cJWNCIkYnrEVivc/edit?gid=2047262135#gid=2047262135>

VISUALIZATION: https://lookerstudio.google.com/s/rBYL1g_7Rys

6. SKU and Inventory Management:

- Which SKUs are overstocked or understocked based on order patterns?

```
SELECT SKU, SUM(Qty) AS TotalSold  
FROM `new-project-439116.Ecommerce_project.1`  
GROUP BY SKU  
HAVING TotalSold < 50 OR TotalSold > 500  
ORDER BY TotalSold DESC;
```

QUERY:

<https://docs.google.com/spreadsheets/d/16QIFO9CiLAtpaUAHtZZZKvYxOVa-VTaHm7PY6zhsTDw/edit?gid=715526707#gid=715526707>

7. Customer Behavior:

- What trends can be identified in customer ordering behavior over time? Are there peak periods for order placement based on historical data?

```
SELECT DATE(Date) AS OrderDate, COUNT(*) AS TotalOrders  
FROM `new-project-439116.Ecommerce_project.1`  
GROUP BY OrderDate  
ORDER BY TotalOrders DESC;
```

QUERY:

<https://docs.google.com/spreadsheets/d/1Wlp-MXYEemoJI5OOV7LeSkpJYBWCmNePF70upxGPPDMs/edit?gid=312607226#gid=312607226>

VISUALIZATION: https://lookerstudio.google.com/s/tRWnrgn4z_8

8. **Highest Sales Achievement:** The highest sales recorded to date amounted to INR 5,584 for Order ID 402-5239558-4281957, placed on April 5, 2022, in the "Set" category.

SELECT ORDER_ID, Amount **AS** HighestSalesAmount, **Date**, Category, Fulfilment
FROM `new-project-439116.Ecommerce_project.1`
WHERE Amount = (**Select MAX** (Amount) **FROM** `new-project-439116.Ecommerce_project.1`)

ORDER_ID	HighestSalesAmount	Date	Category	Fulfilment
402-5239558-4281957	5,584	2022-04-05	Set	Amazon

9. VISUALIZATIONS:

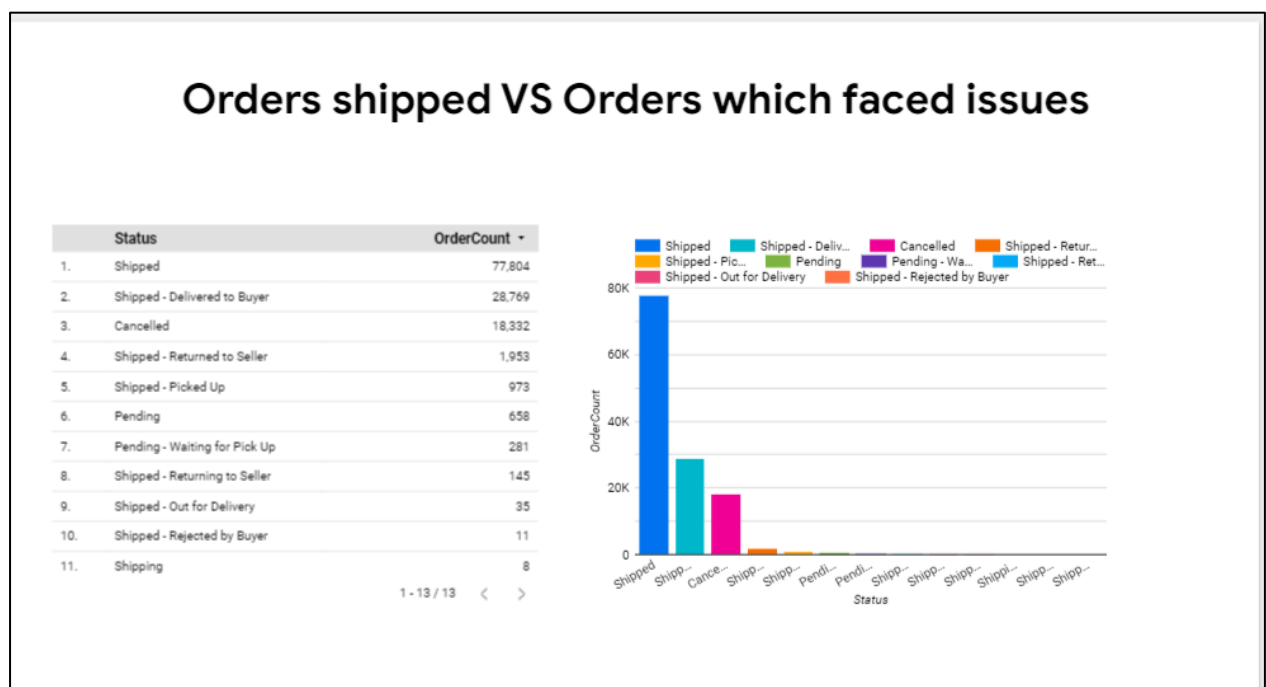


Figure1. showing the orders shipped vs the orders which faced issues.

Products or styles with the highest sales volumes and revenues

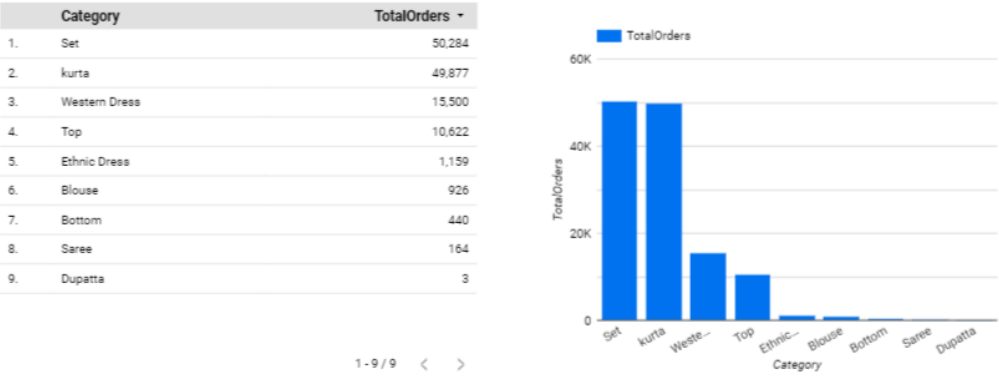


Figure 2. Showing the category vs the highest orders generated

Revenue generated by Total orders based on size

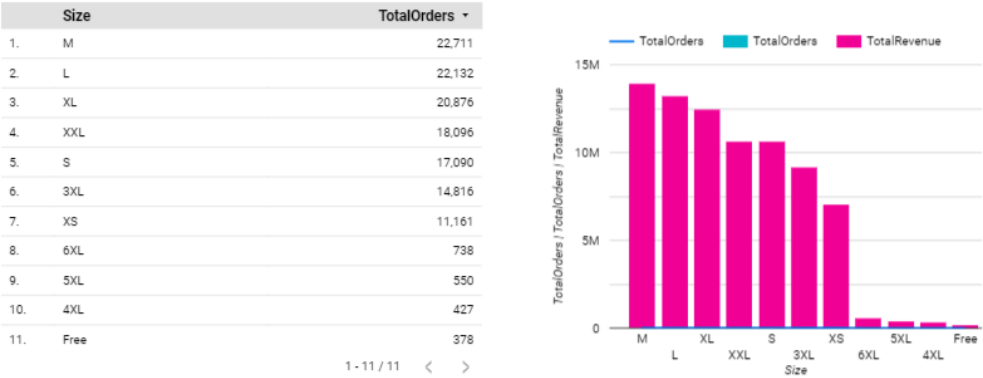


Figure 3. Showing the Size vs total orders generated based on size

Most common shipping locations

	ship_city	OrderCount
1.	BENGALURU	11,217
2.	HYDERABAD	8,065
3.	MUMBAI	6,126
4.	NEW DELHI	5,785
5.	CHENNAI	5,421
6.	PUNE	3,857
7.	KOLKATA	2,381
8.	GURUGRAM	1,868
9.	THANE	1,701
10.	LUCKNOW	1,458
11.	NOIDA	1,419

1 - 50 / 100 < >

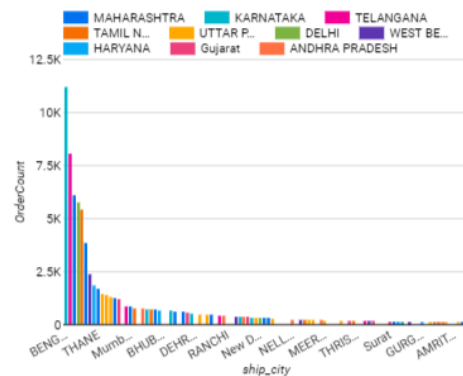


Figure 4. Showing the highest ship_city based on the OrderCount

Trends in order volume based on geographic location

	ship_state	OrderCount
1.	MAHARASHTRA	22,260
2.	KARNATAKA	17,326
3.	TAMIL NADU	11,483
4.	TELANGANA	11,330
5.	UTTAR PRADESH	10,638
6.	DELHI	6,782
7.	KERALA	6,585
8.	WEST BENGAL	5,963
9.	ANDHRA PRADESH	5,430
10.	Gujarat	4,489
11.	HARYANA	4,415

1 - 50 / 70 < >

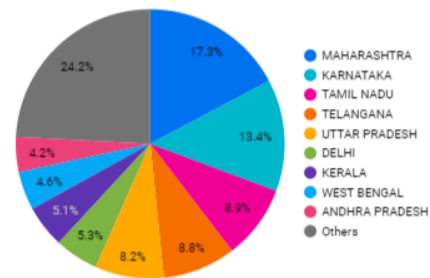


Figure 5. Showing the OrderCount generated based on the ship_state

Identifying overstocked or understocked SKUs

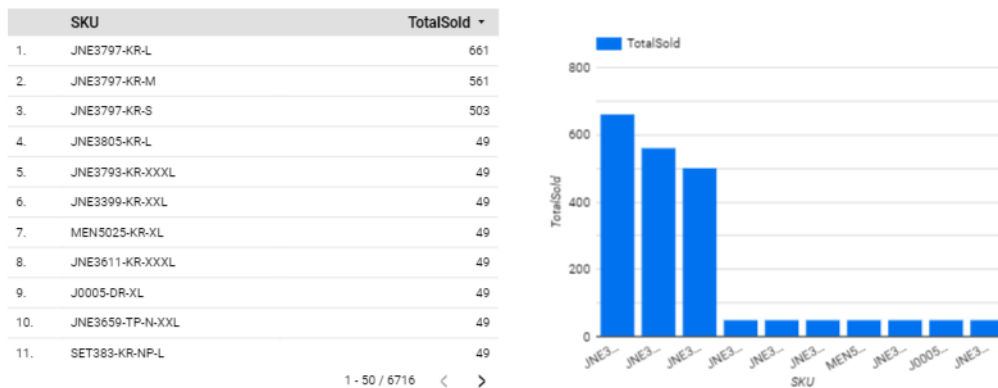


Figure 6. Showing the Totalsold based on the SKU

Trends in customer ordering behavior over time

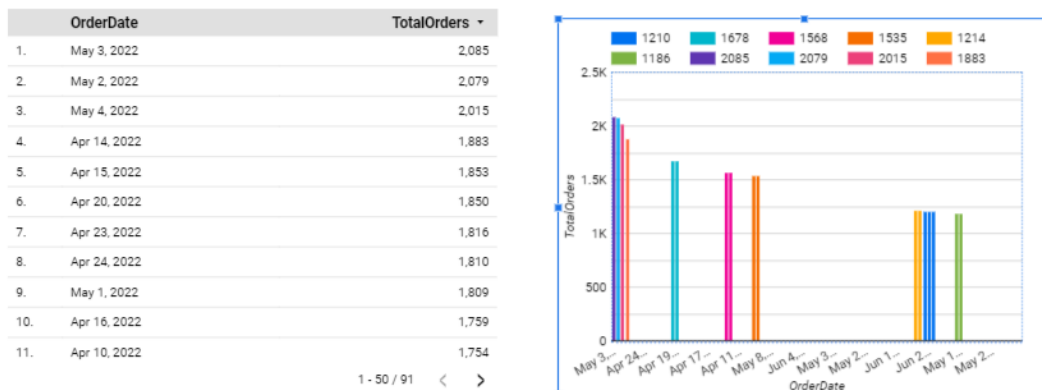


Figure 7. Showing the TotalOrders based on the OrderDate

10. Recommendations:

Based on the analysis, the following data-driven recommendations were made:

1. **Improve Fulfillment Efficiency:** Focus on reducing the return and cancellation rate by improving product descriptions and customer support.
2. **Optimize Sales Channels:** Bengaluru generates the highest revenue, but other channels should be optimized for better performance.
3. **Target Top-Selling Categories:** Increase inventory and promotions for top-selling categories such as Sets.
4. **Enhance Geographic Marketing:** Focus marketing efforts on regions with high order volumes like Maharashtra

11. Conclusion:

This data analysis provided insights into the company's sales performance, fulfillment efficiency, and customer behavior. By addressing identified trends and issues, the company can implement strategies to boost revenue, optimize order fulfillment, and enhance customer satisfaction.
