

Online Retail Customer Segmentation and Analytics

Project Objective

This project analyzes an online retail transactional dataset to uncover customer and sales insights. Using **SQL** for data cleaning and analysis, **Python** for advanced RFM-based segmentation, and **Power BI** for interactive dashboards, the goal is to identify key customer groups and business patterns that can inform **targeted marketing strategies** and **improve retention**.

Key Business Insights

1. Overall Business Performance

- **Total Sales Revenue: ₹2,202,530.08**
- **Total Orders Processed: 6,287**
- **Total Quantity Sold: 1,271,333 items**

➡ The business is generating consistent sales volume with strong order activity.

2. Monthly Sales Trends

| Month (YY-MM) | Revenue (₹) |
|---------------|-------------|
| 2010-12 | 554,604.02 |
| 2011-01 | 475,074.38 |
| 2011-02 | 436,546.15 |
| 2011-03 | 579,964.61 |
| 2011-04 | 156,340.92 |

📊 **Insight:** Sales peaked in **March 2011 (₹579,964)**, while **April 2011 (₹156,340)** shows a sharp decline. This could indicate **seasonal demand patterns** or **inventory/operational issues**.

3. Best-Selling Products by Revenue

Top 5 products driving revenue:

1. Regency Cake Stand 3-Tier – ₹52,680.05
2. White Hanging Heart T-Light Holder – ₹35,024.20
3. Jumbo Bag Red Retrosport – ₹23,232.31
4. Postage – ₹18,982.31
5. Chilli Lights – ₹16,994.46

➡ A mix of **home décor and lifestyle items** dominate sales. These products should be prioritized in **marketing campaigns**.

4. Customer Behavior – Repeat vs One-Time

- Repeat Customers: 2,235 (96%)
- One-Time Customers: 84 (4%)

📌 **Insight:** The business has a **very loyal customer base**. Focusing on **retention strategies** and loyalty programs can maximize long-term revenue.

5. Product Returns

- Highest returns observed in **March (27,252 items)** and **December (25,979 items)**.
- Return values align closely with **sales peaks**, suggesting **higher sales volume naturally increases returns**.

➡ **Action:** Investigate **return reasons** to improve product quality and reduce reverse logistics costs.

6. Sales by Day of the Week

- Highest sales occur on **Thursdays and Fridays**, suggesting customers shop more **towards the weekend**.
- **Monday and Sunday sales are relatively lower**.

→ Marketing campaigns and promotions should be **scheduled mid-to-late week**.

7. VIP Customers (Top Spenders)

Top customers by total spending:

- Customer 14646 – ₹81,567.62
- Customer 18102 – ₹45,279.41
- Customer 12415 – ₹38,024.04
- ...

→ These VIPs can be **targeted for exclusive loyalty rewards**, premium offers, and early product launches.

8. Unique Products per Order

- Some invoices include **up to 271 different items** in a single order.
 - Indicates a segment of **bulk buyers/wholesalers** who should be treated differently from individual customers.
-

9. Sales by Hour of the Day


| Hour | Revenue (₹) |
|-------|-------------|
| 9 AM | 201,797.66 |
| 10 AM | 310,649.29 |
| 11 AM | 300,891.01 |
| 12 PM | 374,300.54 |
| 1 PM | 291,686.72 |

 **Insight:** Sales are **highest between 10 AM – 1 PM**, peaking at **12 PM (₹374,300.54)**.

→ Marketing efforts such as **flash sales and email campaigns** should be **scheduled in the morning to midday**.

Conclusion

- **Strong repeat customer base (96%)** is the backbone of revenue.
- **Top 10 products and VIP customers** generate a large portion of sales.
- **Midday and weekend-focused sales campaigns** could further boost revenue.
- **Returns in peak months** need closer monitoring to optimize operations.

 **Business Impact:** By leveraging segmentation, the company can build **targeted loyalty programs**, **reduce returns**, and **align promotions with customer behavior**, directly increasing profitability and retention.