# Truemeds Assignment

**Que (1.)** Write an SQL query to calculate the quarter on quarter retention for the last one year & visualise this in the same format below:

**Sol.**

```
query = """
WITH CustomerOrders AS (
    SELECT
        customer_id,
        Order_Year,
        Order_Quarter
    FROM orders
    WHERE "Event Time" >= date('now', '-1 year')
),
RetentionData AS (
    SELECT
        c1.customer_id,
        c1.Order_Year,
        c1.Order_Quarter AS Current_Quarter,
        COUNT(DISTINCT c1.customer_id) AS Current_Customers,
        COUNT(DISTINCT c2.customer_id) AS Previous_Customers
    FROM CustomerOrders c1
    LEFT JOIN CustomerOrders c2
        ON c1.customer_id = c2.customer_id
        AND c2.Order_Year = c1.Order_Year
        AND c2.Order_Quarter = c1.Order_Quarter - 1
    GROUP BY c1.Order_Year, c1.Order_Quarter
),
QoQ_Customer AS (
    SELECT
        Order_Year,
        SUM(CASE WHEN Current_Quarter = 1 THEN Current_Customers END) AS Q1_Customers,
        SUM(CASE WHEN Current_Quarter = 2 THEN Current_Customers END) AS Q2_Customers,
        SUM(CASE WHEN Current_Quarter = 3 THEN Current_Customers END) AS Q3_Customers,
```

```sql
        SUM(CASE WHEN Current_Quarter = 4 THEN Current_Customers END) AS Q4_Customers,

        SUM(CASE WHEN Current_Quarter = 1 THEN Previous_Customers END) AS Q1_Previous_Customers,

        SUM(CASE WHEN Current_Quarter = 2 THEN Previous_Customers END) AS Q2_Previous_Customers,

        SUM(CASE WHEN Current_Quarter = 3 THEN Previous_Customers END) AS Q3_Previous_Customers,

        SUM(CASE WHEN Current_Quarter = 4 THEN Previous_Customers END) AS Q4_Previous_Customers
    FROM RetentionData
    GROUP BY Order_Year
    ORDER BY Order_Year
)
SELECT
    Order_Year,
    COALESCE(Q1_Customers * 100 / NULLIF(Q1_Previous_Customers, 0), 0) AS Q1_Retention,
    COALESCE(Q2_Customers * 100 / NULLIF(Q2_Previous_Customers, 0), 0) AS Q2_Retention,
    COALESCE(Q3_Customers * 100 / NULLIF(Q3_Previous_Customers, 0), 0) AS Q3_Retention,
    COALESCE(Q4_Customers * 100 / NULLIF(Q4_Previous_Customers, 0), 0) AS Q4_Retention
FROM QoQ_Customer;
"""
df_retention = pd.read_sql(query, conn)
print(df_retention)
```

Output Format:

```
   Order_Year  Q1_Retention  Q2_Retention  Q3_Retention  Q4_Retention
0        2024             0             0             0             0
```

Q2.Business analysis & insight generation:

Find the dataset attached, and the analysis query below:

Dataset below for an e-pharma company with a specific focus on generic medicines. Whenever a customer has an intent to buy medicines, the company via its advisors asks the customers to *switch* to the trusted generic medicines which are comparatively cheaper for the same composition and efficacy. Naturally, customers who are value seekers are more likely to opt for this switch.

(1) **Which media source has the biggest delta between install time & event time? What is the average time from install to the 3 events? Given the three events and funnel shared, can you provide a reasoning for this delay from install?**

**Sol:**

(A) **Rocketship** media source has the biggest delta between install time & event time.

(B) The average time from install to the 3 events is **62.69 hours (3761.69 minutes).**

```
Average Time from Install to Each Event:
        Event Name  Install_to_Event_Hours
0  app_order_placed              48.125448
1      box_verified              63.551308
2    order_delivered             113.394379
```

Cumulative Average Time from Install to All Events Together: 62.69 hours

Cumulative Average Time from Install to All Events Together: 3761.69 minutes

(C) The delay from install to event completion in your funnel, consider the following factors:

- **User Onboarding & Engagement Time** – Users may take time to explore the app before engaging in key events as they may want to feel comfortable navigating the app first. If onboarding is complex or requires manual input, delays can increase.
- **Promotional Campaign** – Different acquisition channels attract users with varying intent. Promotions can influence when user decide to place orders. Organic users might take longer to convert than paid users who install with a clear purpose.
- **Discounts or special offers:** Many customers are waiting for discounts or special offers, they may delay purchasing until they feel they are getting the best deal, many customers also look for other purchasing options to guarantee their purchase.
- **Customer Support Interaction:** Many customers contact customer support for inquiries about product, services etc before finalising their purchases. This process delay the order placement.
- **Technical factors or user interface:** Slow loading times, glitches, complex user interface slow down their action in the app. Additionally slow internet connectivity and problem during payment processing or unavailability of COD also delay the process.
- **Decision-Making:** Customer always read reviews and check price in other platform also delivery and platform fee etc kind of decision increase time between installing the app and placing orders.

**(2)** **What is the most revenue driving channel? Put a case forward for where you can accurately visualise the revenue driven vs quality factors ( 'core customers' who accepts the 'switch', 'does not use coupon' can be considered as quality metrics)**

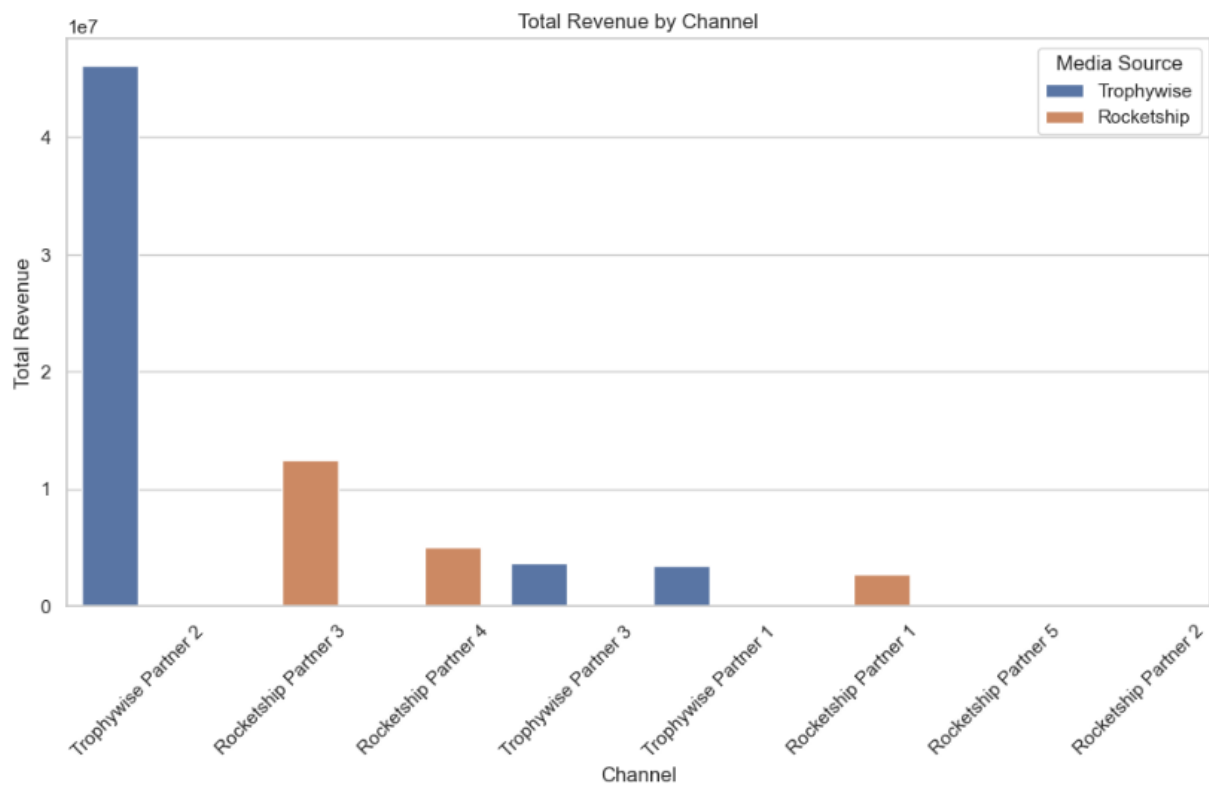**Sol.** **Trophywise Partner 2** is the most revenue driving channel
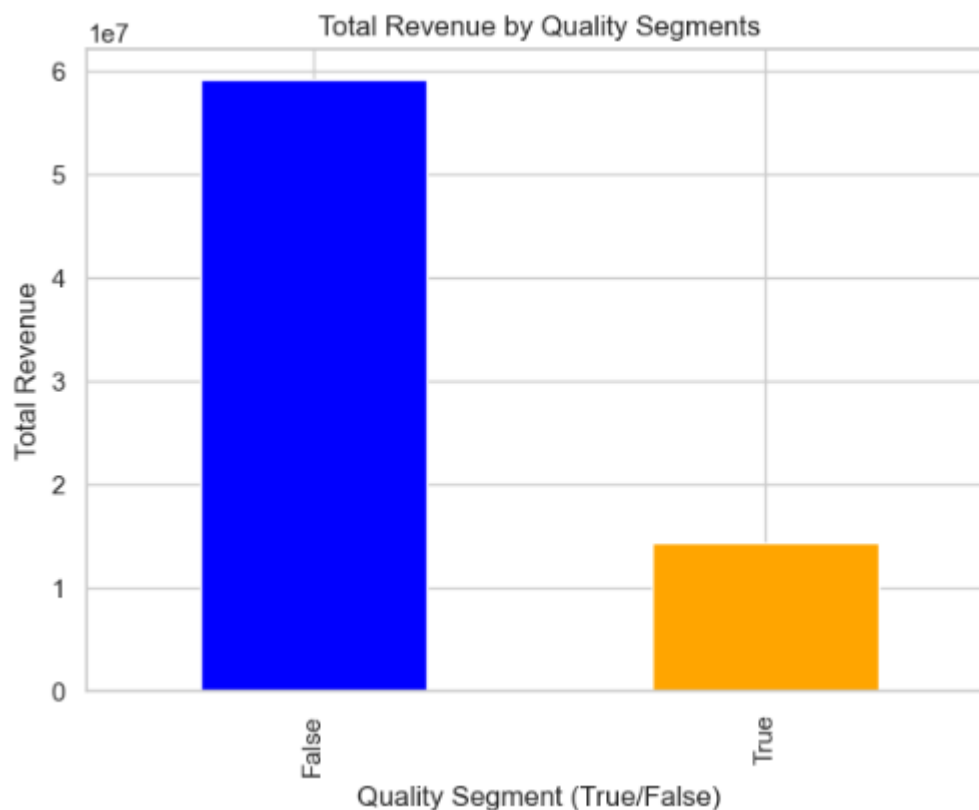


**Fig: Total Revenue per Channel**



**Fig: Total Revenue per quality segment**
**True**: customer is a core customer, accepts the switch and does not use coupon

**(3) Location level analysis: Given the current dataset, which are the top locations in terms of coupon usage, switch medicines, & gets max core customers)**

**Sol.**
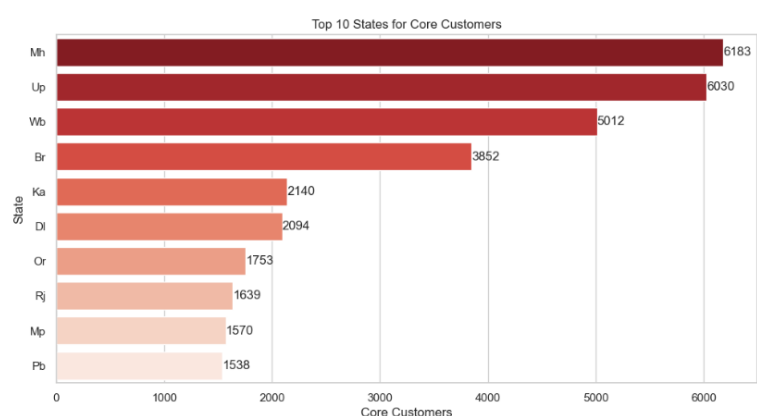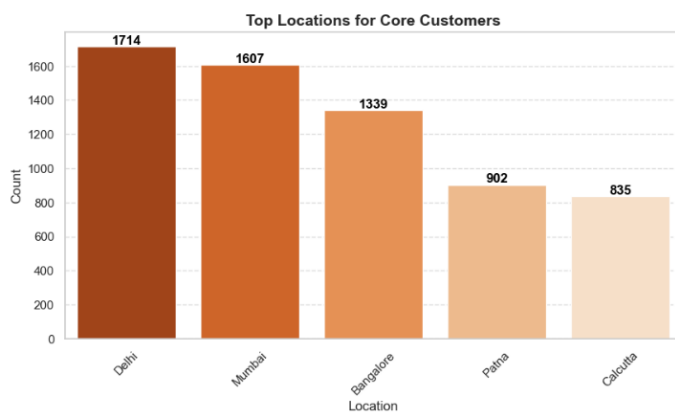
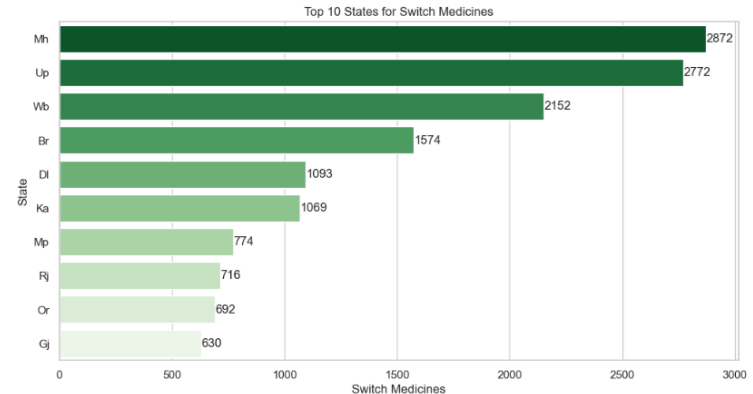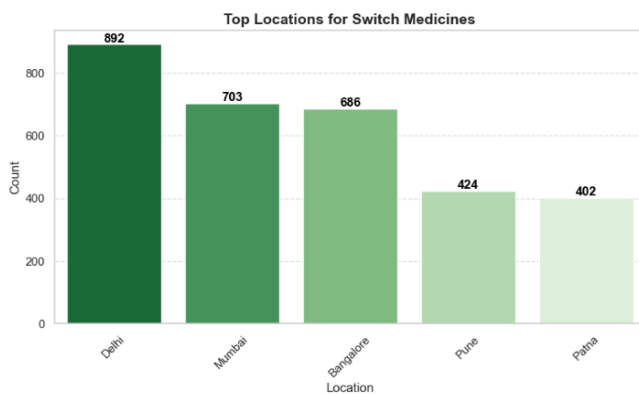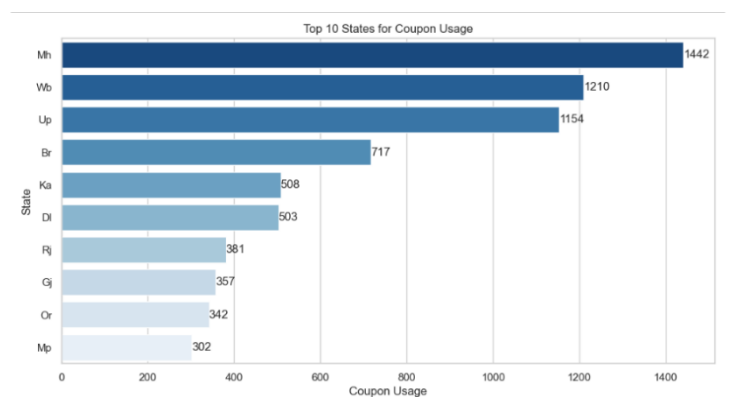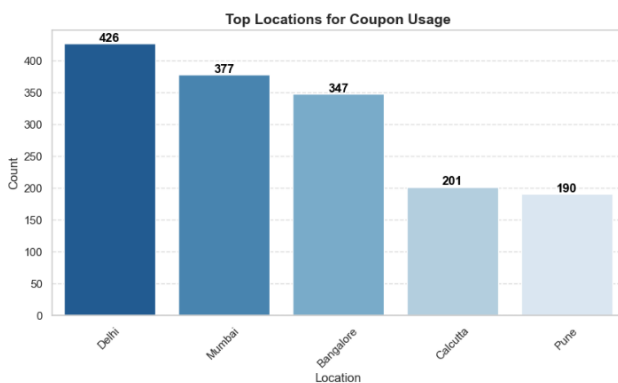The top location in terms of coupon usage is Delhi.

**The top location in terms of coupon usage (State) is MH.**

**The top location in terms of switch medicines is Delhi.**

**The top location in terms of switch medicines (State) is MH.**

**The top location in terms of gets max core customers is Delhi.**

**The top location in terms of gets max core customers (State) is MH.**



Top Locations for Coupon Usage



Top 10 States for Coupon Usage



Top Locations for Switch Medicines



Top 10 States for Switch Medicines



Top Locations for Core Customers



Top 10 States for Core Customers

**(4)** If you were to advise the marketing team to double down on spending on such campaigns, which are the top campaigns to increase spending and why?

Sol.

| | Campaign_ID | Media_Source | total_events | total_revenue | campaign_start_date |
|---|---|---|---|---|---|
| 0 | 19769693693 | Trophywise | 44331 | 43061544 | 2024-02-01 00:13:54 |
| 1 | 238529013680040200 | Rocketship | 18616 | 14860120 | 2024-02-01 00:10:28 |
| 2 | 20835741561 | Trophywise | 11017 | 7962765 | 2024-02-01 00:18:00 |
| 3 | 1202036388903710000 | Rocketship | 1756 | 1832635 | 2024-02-01 00:15:25 |
| 4 | 1202030662375470000 | Rocketship | 3272 | 1752918 | 2024-02-01 00:25:09 |
| 5 | 20957689169 | Trophywise | 5 | 3559 | 2024-02-06 18:23:45 |
| 6 | 20871219798 | Trophywise | 3 | 1616 | 2024-02-09 20:12:51 |
| 7 | 20874970004 | Trophywise | 4 | 511 | 2024-02-12 14:55:16 |
| 8 | 23849096377440200 | Rocketship | 3 | 308 | 2024-02-22 11:57:05 |
| 9 | 20840388691 | Trophywise | 1 | 184 | 2024-02-16 15:42:14 |

**Note**: The first five entries represent the top 5 campaigns while the last five entries represent the 5 least successful campaigns.

- We can double down the spending on Campaign_ID=19769693693 as:
  This campaign generates highest revenue=Rs. 43061544.92 which shows strong customer engagement and conversion effectiveness. And this campaign attracts a large number of core customers (Total Customers= 12180), investing more in such campaigns can foster brand loyalty and repeat purchases.
  I also want to know about the reason behind campaign_id=23849096377440200 as it is a recently launched campaign or the motive seems promising, I also like to spend on this event.

_____

**(5)** Summarise key learning and business insight in brief.

Sol.

- The highest revenue come from **Maharashtra** state.
- The highest revenue come from **Delhi.**
- The highest revenue generating Media Source is **Trophywise(53182329).**
- The most successful campaign is **Campaign_ID=19769693693** which is running within**Trophywise.**
- Average discount per event: Rs 360.37
- Average Order Value: Rs 916.30
- Total Revenue: Rs 73504564.64
- Average Revenue Per Day: 2940182.59