CASE-STUDY STYLE SQL QUESTIONS

1. If Olist wants to improve customer satisfaction, which operational metrics should be analyzed?

Customer satisfaction in Olist is mainly driven by:

- **Delivery performance**
- **Order issues (late delivery)**
- **Review scores**
- **Seller performance**

Metric 1: Average Delivery Time vs Review Score

```
SELECT
  ROUND(
    DATEDIFF(
      o.order_delivered_customer_date,
      o.order_purchase_timestamp
    )
  ) AS delivery_days,
  AVG(r.review_score) AS avg_review_score,
  COUNT(*) AS total_orders
FROM orders o
JOIN order_reviews r ON o.order_id = r.order_id
WHERE o.order_status = 'delivered'
GROUP BY delivery_days
ORDER BY delivery_days desc limit 10;
```

Result Grid | Filter Rows:

delivery_days	avg_review_score	total_orders
208	2.0000	1
196	1.0000	1
195	2.5000	2
194	4.0000	1
191	1.0000	1
190	1.0000	1
188	3.0000	2
187	5.0000	1
186	4.0000	1
183	5.0000	1

Metric 2: Late Delivery Rate

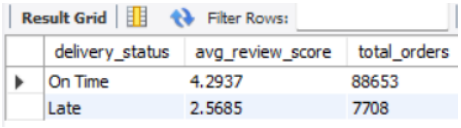
```
SELECT
  ROUND(
    COUNT(CASE
      WHEN o.order_delivered_customer_date > o.order_estimated_delivery_date
      THEN 1 END) * 100.0 / COUNT(*),
    2
  ) AS late_delivery_percentage
FROM orders o
WHERE o.order_status = 'delivered';
```

Result Grid | Filter R

late_delivery_percentage
8.11

Metric 3: Review Score by Delivery Status

```
SELECT
   CASE
      WHEN o.order_delivered_customer_date <= o.order_estimated_delivery_date
         THEN 'On Time'
      ELSE 'Late'
   END AS delivery_status,
   AVG(r.review_score) AS avg_review_score,
   COUNT(*) AS total_orders
FROM orders o
JOIN order_reviews r ON o.order_id = r.order_id
WHERE o.order_status = 'delivered'
GROUP BY delivery_status;
```

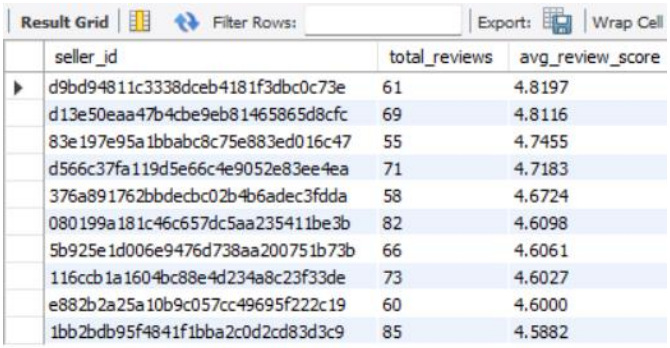


Result Grid | Filter Rows:

delivery_status	avg_review_score	total_orders
On Time	4.2937	88653
Late	2.5685	7708

Metric 4: Seller-Level Customer Satisfaction

```
SELECT
   oi.seller_id,
   COUNT(r.review_id) AS total_reviews,
   AVG(r.review_score) AS avg_review_score
FROM order_items oi
JOIN order_reviews r ON oi.order_id = r.order_id
GROUP BY oi.seller_id
HAVING COUNT(r.review_id) >= 50
ORDER BY avg_review_score desc limit 10;
```



Result Grid | Filter Rows: | Export: | Wrap Cell

seller_id	total_reviews	avg_review_score
d9bd94811c3338dceb4181f3dbc0c73e	61	4.8197
d13e50eaa47b4cbe9eb81465865d8cfc	69	4.8116
83e197e95a1bbabc8c75e883ed016c47	55	4.7455
d566c37fa119d5e66c4e9052e83ee4ea	71	4.7183
376a891762bbdecbc02b4b6adec3fdda	58	4.6724
080199a181c46c657dc5aa235411be3b	82	4.6098
5b925e1d006e9476d738aa200751b73b	66	4.6061
116ccb1a1604bc88e4d234a8c23f33de	73	4.6027
e882b2a25a10b9c057cc49695f222c19	60	4.6000
1bb2bdb95f4841f1bba2c0d2cd83d3c9	85	4.5882

📌Summary: To improve customer satisfaction, it's important to analyse delivery time, late delivery rate, review scores, and seller-level performance, as these metrics directly affect customer experience and repeat purchases.

2. **Which categories should Olist prioritise for marketing based on revenue and growth?**

Marketing should focus on:

- **High revenue categories**
- **Emerging categories with growth momentum**
- **Fast-growing categories**

Metric 1: Revenue by Category

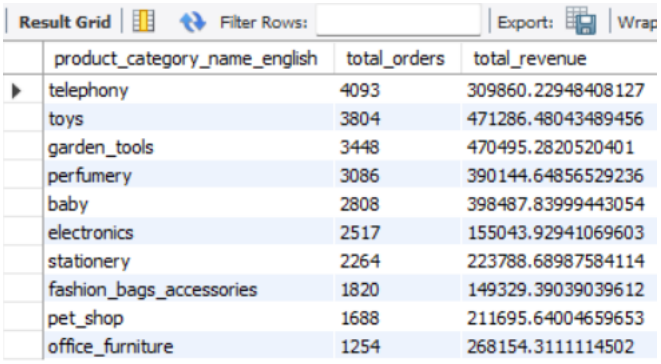
```
SELECT
   p.product_category_name_english,
   SUM(oi.price) AS total_revenue
FROM order_items oi
JOIN products p ON oi.product_id = p.product_id
JOIN orders o ON oi.order_id = o.order_id
WHERE o.order_status = 'delivered'
GROUP BY p.product_category_name_english
ORDER BY total_revenue DESC limit 10;
```

Result Grid | Filter Rows: | Expo

product_category_name_english	total_revenue
health_beauty	1233131.7208693027
watches_gifts	1166176.9777069092
bed_bath_table	1023434.7600488663
sports_leisure	954852.5489358902
computers_accessories	888724.6074113846
furniture_decor	711927.6923160553
housewares	615628.690759182
cool_stuff	610204.101855278
auto	578966.6507818699
toys	471286.48043489456

Metric 2: High Growth but Medium Revenue Categories (Hidden Opportunities)

```
SELECT
   p.product_category_name_english,
   COUNT(DISTINCT o.order_id) AS total_orders,
   SUM(oi.price) AS total_revenue
FROM orders o
JOIN order_items oi ON o.order_id = oi.order_id
JOIN products p ON oi.product_id = p.product_id
WHERE o.order_status = 'delivered'
GROUP BY p.product_category_name_english
HAVING total_revenue BETWEEN 100000 AND 500000
ORDER BY total_orders DESC limit 10;
```

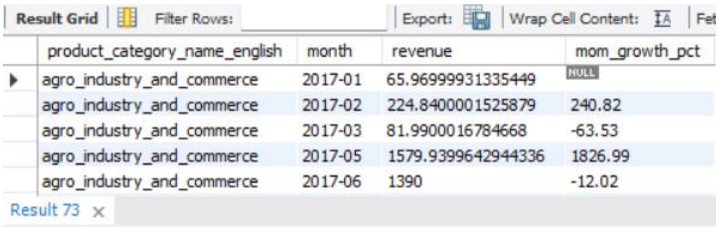


Result Grid | Filter Rows: | Export: | Wrap

product_category_name_english	total_orders	total_revenue
telephony	4093	309860.22948408127
toys	3804	471286.48043489456
garden_tools	3448	470495.2820520401
perfumery	3086	390144.64856529236
baby	2808	398487.83999443054
electronics	2517	155043.92941069603
stationery	2264	223788.68987584114
fashion_bags_accessories	1820	149329.39039039612
pet_shop	1688	211695.64004659653
office_furniture	1254	268154.3111114502

Metric 3: Monthly Revenue Growth by Category

```
WITH monthly_category_revenue AS (
   SELECT
      p.product_category_name_english,
      DATE_FORMAT(o.order_purchase_timestamp, '%Y-%m') AS month,
      SUM(oi.price) AS revenue
   FROM orders o
   JOIN order_items oi ON o.order_id = oi.order_id
   JOIN products p ON oi.product_id = p.product_id
   WHERE o.order_status = 'delivered'
   GROUP BY p.product_category_name_english, month
)
SELECT
   product_category_name_english,
   month,
   revenue,
   ROUND(
      (revenue - LAG(revenue) OVER (
         PARTITION BY product_category_name_english
         ORDER BY month
      )) / LAG(revenue) OVER (
         PARTITION BY product_category_name_english
         ORDER BY month
      ) * 100,
      2
   ) AS mom_growth_pct
FROM monthly_category_revenue;
```



Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fe

product_category_name_english	month	revenue	mom_growth_pct
agro_industry_and_commerce	2017-01	65.96999931335449	NULL
agro_industry_and_commerce	2017-02	224.8400001525879	240.82
agro_industry_and_commerce	2017-03	81.9900016784668	-63.53
agro_industry_and_commerce	2017-05	1579.9399642944336	1826.99
agro_industry_and_commerce	2017-06	1390	-12.02

Result 73

📌Summary: To prioritise categories for marketing, it's important to analyse category-level revenue, month-over-month growth, and order volume to identify high-performing and fast-growing categories with strong ROI potential.

Soumi Mukherjee