

-- 1. Doanh thu thay đổi như thế nào theo thời gian, và đâu là các tháng cao điểm cho doanh số bán hàng?

```
SELECT EXTRACT(MONTH FROM order_timestamp) AS "month",
       SUM(unit_price + freight_value) AS revenue
FROM gold.fact_orders
GROUP BY EXTRACT(MONTH FROM order_timestamp)
ORDER BY revenue DESC;
```

	month numeric	revenue numeric
1	5	1562899.13
2	8	1533292.16
3	7	1493165.65
4	3	1451056.90
5	4	1436275.16
6	6	1398412.40
7	2	1160436.45
8	1	1130397.35
9	11	1055581.23
10	12	798345.98
11	10	745005.12
12	9	645161.61

-- 2. Danh mục sản phẩm nào đóng góp nhiều nhất vào tổng doanh thu?

```
SELECT dp.category_name_english,
       SUM(fo.unit_price + fo.freight_value) AS revenue
FROM gold.fact_orders fo
JOIN gold.dim_products dp ON fo.product_key = dp.product_key
GROUP BY dp.category_name_english
ORDER BY revenue DESC
LIMIT 1;
```

	category_name_english character varying (100)	revenue numeric
1	health_beauty	1359900.38

-- 3. Sản phẩm nào bán chạy nhất (theo số lượng và doanh thu) trong từng danh mục?

-- Theo số lượng

```
SELECT dp.category_name_english, dp.product_id,
       COUNT(*) AS quantity_sold
FROM gold.fact_orders fo
JOIN gold.dim_products dp ON fo.product_key = dp.product_key
GROUP BY dp.category_name_english, dp.product_id
ORDER BY dp.category_name_english, quantity_sold DESC;
```

-- Theo doanh thu

```
SELECT dp.category_name_english, dp.product_id,
       SUM(fo.unit_price + fo.freight_value) AS revenue
FROM gold.fact_orders fo
JOIN gold.dim_products dp ON fo.product_key = dp.product_key
GROUP BY dp.category_name_english, dp.product_id
ORDER BY dp.category_name_english, revenue DESC;
```

	category_name_english character varying (100)	quantity_sold bigint
1	bed_bath_table	9324
2	health_beauty	8804
3	sports_leisure	7687
4	computers_accessories	6665
5	furniture_decor	6334
6	housewares	5823
7	watches_gifts	5604
8	telephony	4189

	category_name_english character varying (100)	revenue numeric
1	health_beauty	1359900.38
2	watches_gifts	1263716.81
3	bed_bath_table	1069958.35
4	sports_leisure	1064037.43
5	computers_accessories	904875.69
6	furniture_decor	735326.65
7	cool_stuff	694483.50
8	housewares	685182.29

-- 4. Có sự khác biệt nào về doanh thu giữa các khu vực địa lý ?

```
SELECT dc.city,
       SUM(fo.unit_price + fo.freight_value) AS revenue
FROM gold.fact_orders fo
JOIN gold.dim_customers dc ON fo.customer_key = dc.customer_key
GROUP BY dc.city
ORDER BY revenue DESC;
```

	city character varying (100)	revenue numeric
1	Sao Paulo	1962519.10
2	Rio De Janeiro	1043579.91
3	Belo Horizonte	383500.24
4	Brasilia	285342.93
5	Curitiba	219787.30
6	Porto Alegre	195824.16
7	Salvador	192890.63
8	Campinas	192394.12

-- 5. Phương thức thanh toán nào được khách hàng ưa chuộng nhất, và điều này ảnh hưởng đến giá trị đơn hàng trung bình (AOV)

```
-- Phương thức thanh toán ưa chuộng
SELECT dp.payment_type,
       COUNT(*) AS transaction_count
FROM gold.fact_orders fo
JOIN gold.dim_payments dp ON fo.payment_key = dp.payment_key
GROUP BY dp.payment_type
ORDER BY transaction_count DESC;
```

-- AOV theo phương thức thanh toán

```
SELECT dp.payment_type,
       AVG(fo.unit_price + fo.freight_value) AS aov
FROM gold.fact_orders fo
JOIN gold.dim_payments dp ON fo.payment_key = dp.payment_key
GROUP BY dp.payment_type
ORDER BY aov DESC;
```

	payment_type character varying (50)	transaction_count bigint
1	credit_card	71632
2	boleto	19166
3	voucher	3667
4	debit_card	1472

	payment_type character varying (50)	aov numeric
1	credit_card	153.7501919533169533
2	debit_card	132.8398505434782609
3	boleto	128.0294980694980695
4	voucher	124.1079656394873193

-- 6. Tỷ lệ khách hàng quay lại khác nhau như thế nào giữa các khu vực địa lý?

```
SELECT dc.city,
       COUNT(DISTINCT fo.customer_key) AS returning_customers
FROM gold.fact_orders fo
JOIN gold.dim_customers dc ON fo.customer_key = dc.customer_key
GROUP BY dc.city
HAVING COUNT(DISTINCT fo.order_id) > 1
ORDER BY returning_customers DESC;
```

	city character varying (100)	returning_customers bigint
1	Sao Paulo	15407
2	Rio De Janeiro	6851
3	Belo Horizonte	2744
4	Brasilia	1863
5	Curitiba	1497
6	Campinas	1429
7	Porto Alegre	1371
8	Salvador	1235
9	Guarulhos	1177
10	Sao Bernardo Do Campo	928
11	Niteroi	845

-- 8. Sản phẩm nào có tỷ lệ đánh giá thấp, và điều này ảnh hưởng đến doanh thu như thế nào?

```
SELECT dp.category_name_english,
       AVG(dr.review_score) AS avg_rating,
       SUM(fo.unit_price + fo.freight_value) AS revenue
FROM gold.fact_orders fo
JOIN gold.dim_products dp ON fo.product_key = dp.product_key
JOIN gold.dim_reviews dr ON fo.review_key = dr.review_key
GROUP BY dp.category_name_english
HAVING AVG(dr.review_score) < 3
ORDER BY avg_rating ASC;
```

	category_name_english character varying (100)	avg_rating numeric	revenue numeric
1	security_and_services	2.5000000000000000	324.51

-- 9. Người bán nào có hiệu suất tốt nhất dựa trên doanh thu và đánh giá khách hàng?

```
SELECT ds.seller_id, ds.city,
       SUM(fo.unit_price + fo.freight_value) AS revenue,
       AVG(dr.review_score) AS avg_rating
FROM gold.fact_orders fo
JOIN gold.dim_sellers ds ON fo.seller_key = ds.seller_key
JOIN gold.dim_reviews dr ON fo.review_key = dr.review_key
GROUP BY ds.seller_id, ds.city
ORDER BY revenue DESC, avg_rating DESC
LIMIT 1;
```

	seller_id character varying (50)	city character varying (100)	revenue numeric	avg_rating numeric
1	4869f7a5dfa277a7dca6462dcf3b52b2	Guariba	241823.42	4.1470852017937220

-- 16. Người bán nào có tỷ lệ đánh giá thấp, và họ cần cải thiện ở điểm nào để tăng doanh số?

```
SELECT ds.seller_id, ds.city,
       AVG(dr.review_score) AS avg_rating
FROM gold.fact_orders fo
JOIN gold.dim_sellers ds ON fo.seller_key = ds.seller_key
JOIN gold.dim_reviews dr ON fo.review_key = dr.review_key
GROUP BY ds.seller_id, ds.city
HAVING AVG(dr.review_score) < 3
ORDER BY avg_rating ASC;
```

	seller_id character varying (50)	city character varying (100)	avg_rating numeric
1	7a073a40a9d3f1fee8bc45670ea1e909	Londrina	1.00000000000000000000
2	9bf11dfc0bec77e5a23028043c3c5a8f	Contagem	1.00000000000000000000
3	1444c08e64d55fb3c25f0f09c07ffc2	Lauro De Freitas	1.00000000000000000000
4	67e43d802fde8cfd3f9580124f8167d1	Guarulhos	1.00000000000000000000
5	9591fc341b1bfb7ef561e2968ec6e011	Sao Paulo	1.00000000000000000000
6	61f159ef6da2d441951d2c0efa719362	Serra	1.00000000000000000000
7	9c57bc0cfad5ee62d35d3f1ce4593a1	Curitiba	1.00000000000000000000
8	9599519be538b98748162a2b482489...	Sao Paulo	1.00000000000000000000
9	51a04a8a6bdc23deccc82b0b80742cf	Braganca Paulista	1.00000000000000000000
10	e1a210d482714ce337763a19aef94ba4	Cascavel	1.00000000000000000000
11	c542f5a8c64d973f30e9df8d06f70b4a	Sao Paulo	1.00000000000000000000

Total rows: 284 of 284 Query complete 00:00:04.899 Ln 162, Col 1

-- 17. Sản phẩm nào có tỷ lệ hủy đơn hàng cao nhất, và nguyên nhân là gì?

```
SELECT dp.category_name_english,
       COUNT(*) AS canceled_orders
FROM gold.fact_orders fo
JOIN gold.dim_products dp ON fo.product_key = dp.product_key
WHERE fo.order_status = 'canceled'
GROUP BY dp.category_name_english
ORDER BY canceled_orders DESC;
```

	category_name_english character varying (100)	canceled_orders bigint
1	sports_leisure	47
2	housewares	37
3	health_beauty	36
4	computers_accessories	35
5	toys	31
6	furniture_decor	24
7	auto	24
8	watches_gifts	20
9	bed_bath_table	18
10	telephony	16
11	[null]	16