

Second: Write queries that directly answer predetermined questions from a business stakeholder

***Note** - I am using Redshift SQL

1.) What are the top 5 brands by receipts scanned for most recent month?

```
WITH RecentMonth AS (  
    SELECT year, month  
    FROM dim_Date  
    ORDER BY year DESC, month DESC  
    LIMIT 1  
)  
  
SELECT  
    b.brand_name, COUNT(DISTINCT r.receipt_id) AS receipt_count  
FROM fact_Receipts r  
    JOIN dim_Date d ON r.date_scanned = d.date  
    JOIN fact_ReceiptItems ri ON r.receipt_id = ri.receipt_id  
    JOIN dim_Brands b ON ri.barcode = b.barcode  
    JOIN RecentMonth rm ON d.year = rm.year AND d.month = rm.month  
WHERE ri.deleted <> 'TRUE'  
GROUP BY b.brand_name  
ORDER BY receipt_count DESC  
LIMIT 5
```

2.) How does the ranking of the top 5 brands by receipts scanned for the recent month compare to the ranking for the previous month?

```
WITH RecentMonths AS (  
    SELECT year, month  
    FROM dim_Date  
    ORDER BY year DESC, month DESC  
    LIMIT 2  
)  
  
SELECT  
    d.year, d.month, b.brand_name,  
    COUNT(DISTINCT r.receipt_id) AS receipt_count,
```

```

RANK() OVER (PARTITION BY d.year, d.month ORDER BY
COUNT(DISTINCT r.receipt_id) DESC) AS rank
FROM fact_Receipts r
    JOIN dim_Date d ON r.date_scanned = d.date
    JOIN fact_ReceiptItems ri ON r.receipt_id = ri.receipt_id
    JOIN dim_Brands b ON ri.barcode = b.barcode
WHERE (d.year, d.month) IN (SELECT year, month FROM RecentMonths)
GROUP BY d.year, d.month, b.brand_name
HAVING RANK() OVER (PARTITION BY d.year, d.month ORDER BY
COUNT(DISTINCT r.receipt_id) DESC) <= 5
ORDER BY d.year DESC, d.month DESC, rank

```

3.) When considering *average spend* from receipts with 'rewardsReceiptStatus' of 'Accepted' or 'Rejected', which is greater?

```

SELECT
    reward_receipt_status, AVG(total_spent) AS avg_spend
FROM fact_Receipts
WHERE reward_receipt_status IN ('Accepted', 'Rejected')
AND description <> 'DELETED ITEM'
GROUP BY reward_receipt_status
ORDER BY avg_spend DESC

```

4.) When considering *total number of items purchased* from receipts with 'rewardsReceiptStatus' of 'Accepted' or 'Rejected', which is greater?

***Note** - The question could mean either the total quantity of items purchased or the number of different items. To avoid confusion, I've provided both solutions. The first solution calculates the total quantity of items and the second solution counts the number of unique items.

1.)

```

SELECT
    r.reward_receipt_status, SUM(ri.quantity_purchased) AS
total_items_purchased
FROM fact_Receipts r
    JOIN fact_ReceiptItems ri ON r.receipt_id = ri.receipt_id
WHERE r.reward_receipt_status IN ('Accepted', 'Rejected')
AND ri.deleted <> 'TRUE'
GROUP BY r.reward_receipt_status

```

ORDER BY total_items_purchased DESC

2.)

```
SELECT
    reward_receipt_status, SUM(purchased_item_count) AS
total_items_purchased
FROM fact_Receipts
WHERE reward_receipt_status IN ('Accepted', 'Rejected')
AND description <> 'DELETED ITEM'
GROUP BY reward_receipt_status
ORDER BY total_items_purchased DESC
```

5.) Which brand has the most *spend* among users who were created within the past 6 months?

```
WITH RecentUsers AS (
    SELECT user_id
    FROM dim_Users
    WHERE create_date >= DATEADD(MONTH, -6, (SELECT MAX(create_date)
FROM dim_Users))
)

SELECT
    b.brand_name,
    SUM(r.total_spent) AS total_spend
FROM fact_Receipts r
    JOIN RecentUsers u ON r.user_id = u.user_id
    JOIN fact_ReceiptItems ri ON r.receipt_id = ri.receipt_id
    JOIN dim_Brands b ON ri.barcode = b.barcode
WHERE ri.deleted <> 'TRUE'
GROUP BY b.brand_name
ORDER BY total_spend DESC
LIMIT 1
```

6.) Which brand has the most *transactions* among users who were created within the past 6 months?

```
WITH RecentUsers AS (
    SELECT user_id
    FROM dim_Users
```

```
WHERE create_date >= DATEADD(MONTH, -6, (SELECT MAX(create_date)
FROM dim_Users))
)
```

```
SELECT
    b.brand_name, COUNT(DISTINCT r.receipt_id) AS total_transactions
FROM fact_Receipts r
    JOIN RecentUsers u ON r.user_id = u.user_id
    JOIN fact_ReceiptItems ri ON r.receipt_id = ri.receipt_id
    JOIN dim_Brands b ON ri.barcode = b.barcode
WHERE ri.deleted <> 'TRUE'
GROUP BY b.brand_name
ORDER BY total_transactions DESC
LIMIT 1
```