

FETCH SQL QUERIES

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

I don't see the values of the status column as accepted or rejected , I wrote a query grouped by status to get the result

```
SELECT a.STATUS
       ,sum(a.totalSpent) AS TotalSpent
FROM (
    SELECT DISTINCT _id
           ,rewardsReceiptStatus AS STATUS
           ,totalSpent
    FROM dbo.receipts_new1
    ) a
GROUP BY a.STATUS
ORDER BY sum(a.totalSpent) DESC
```

OUTPUT:

STATUS	TotalSpent
FINISHED	41882.53
FLAGGED	8300.78
REJECTED	1656.15
PENDING	1373.59
SUBMITTED	NULL

4. Which considering total number of items purchased from receipts with rewardreceiptstatus of 'Accepted' or 'rejected' which is greater?

I don't see the values of the status column as accepted or rejected , I wrote a query grouped by status to get the result.

Since there is a lot of duplicate data, I decided to get the distinct receipt id with the number of items in each id and then sum the total count per each status to get the total itemcount.

```

SELECT a.STATUS
      ,sum(a.purchasedItemCount)
FROM (
      SELECT DISTINCT _id
            ,rewardsReceiptStatus AS STATUS
            ,purchasedItemCount
      FROM dbo.receipts_new1
    ) a
GROUP BY a.STATUS
ORDER BY sum(a.purchasedItemCount) DESC

```

OUTPUT:

STATUS	Count
FINISHED	8184
FLAGGED	1014
REJECTED	173
PENDING	0
SUBMITTED	0

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

Here, I considered past 6 months from the maximum creation date , as the data is outdated

```

SELECT TOP 1 r.rewardsReceiptItemList_brandcode as brandcode
      ,SUM(r.TotalSpent)as TotalSpent
FROM (
      SELECT DISTINCT r.rewardsReceiptItemList_brandcode
            ,r.userid
            ,r.TotalSpent
      FROM dbo.receipts_new1 r
      JOIN dbo.users_new u ON r.userid = u._id
      WHERE u.createddate >= dateadd(m, - 6, '2021-02-12 14:11:06')
            AND r.rewardsReceiptItemList_brandcode IS NOT NULL
    ) r
GROUP BY r.rewardsReceiptItemList_brandcode
ORDER BY SUM(r.TotalSpent) DESC

```

OUTPUT:

brandcode	TotalSpent
BEN AND JERRYS	20040.77

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

The same logic as above question 5 is taken into consideration for past 6 months. Considering each receipt item as a transaction

```
SELECT TOP 1 r.rewardsReceiptItemList_brandcode as brandcode
, COUNT(r.userid) as Transactions
FROM dbo.receipts_new1 r
JOIN dbo.users_new u ON r.userid = u._id
WHERE u.createddate >= dateadd(m, - 6, '2021-02-12 14:11:06')
AND r.rewardsReceiptItemList_brandcode IS NOT NULL
GROUP BY r.rewardsReceiptItemList_brandcode
ORDER BY COUNT(r.userid) DESC
```

OUTPUT:

brandcode	Transactions
HY-VEE	291

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

Since data is outdated, I took the max of date scanned to be the most recent month. But still the data is not accurate and missing data for brands, hence giving the query which gives results for above results.

```
SELECT a.brandcode
, count(a._id)
FROM (
    SELECT DISTINCT _id
    , rewardsReceiptItemList_brandcode AS brandcode
    , dateScanned
    FROM dbo.receipts_new1
    WHERE rewardsReceiptItemList_brandcode IS NOT NULL
    AND month(datescanned) = (
        SELECT month(max(datescanned))
        FROM dbo.receipts_new1
    )
    AND year(datescanned) = (
        SELECT year(max(datescanned))
        FROM dbo.receipts_new1
    )
) a
GROUP BY a.brandcode
ORDER BY count(a._id) DESC
```

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

To get previous month , I subtracted one month from the most recent one.

```
SELECT a.brandcode
      ,count(a._id)
FROM (
      SELECT DISTINCT _id
            ,rewardsReceiptItemList_brandcode AS brandcode
            ,dateScanned
      FROM dbo.receipts_new1
      WHERE rewardsReceiptItemList_brandcode IS NOT NULL
            AND month(datescanned) = (
                  SELECT month(max(datescanned)-1)
                  FROM dbo.receipts_new1
            )
            AND year(datescanned) = (
                  SELECT year(max(datescanned)-1)
                  FROM dbo.receipts_new1
            )
      ) a
GROUP BY a.brandcode
ORDER BY count(a._id) DESC
```