Business Stakeholder Queries

I ran the following queries:

1. What are the top 5 brands by receipts scanned for most recent month?
2. When considering average spend from receipts with 'rewardsReceiptStatus' of 'Accepted' or 'Rejected', which is greater?
3. When considering total number of items purchased from receipts with 'rewardsReceiptStatus' of 'Accepted' or 'Rejected', which is greater?

The Queries are run in SQL Server.

Query 1:

```
-- What are the top 5 brands by receipts scanned for most recent month?
select top 5
	i.brandCode,
	sum(ti.finalPrice) as spend
-- joined the table 'receipts' along with the table 'transactions' on the id
-- joined the items table with the transactions table on barcode
-- merged the transactions table with the transactionItems table
from transactionItems as ti
inner join (select month(max(dateScanned))-2 m, year(max(dateScanned)) y from receipts) as
recent on 1=1
inner join transactions as t on ti.transactionId = t.id
inner join receipts as r on t.receiptId = r.id and month(r.dateScanned) = recent.m and
year(r.dateScanned) = recent.y
inner join items as i on ti.barcode = i.barcode
-- group by the barcodes
group by i.brandCode
-- in order to get the top 5 brands, order by spen, spend variable is created by summing the
finalPrice from the transactionItems table
order by spend desc
```

Output:

	brandCode	spend
1	BEN AND JERRYS	2091.79000067711
2	CRACKER BARREL	1125.59997558594
3	HEMPLER'S	1102.30000257492
4	KNORR	706.110000252724
5	DORITOS	469.160000324249

Query 2:

```
-- When considering average spend from receipts with 'rewardsReceiptStatus' of 'Accepted' or
'Rejected', which is greater?

-- Display the average of the spend based on rejected and finished categories.
-- After performing the transformations, the data did not have accepted category, so I used the
finished category instead
select
	avg(case when lower(rewardsReceiptStatus) = 'finished' then isnull(spend,0) end) as
avg_spend_accepted,
```

```
	avg(case when lower(rewardsReceiptStatus) = 'rejected' then isnull(spend,0) end) as
avg_spend_rejected
from
(
	-- select the recipt id, rewardsReceiptStatus and sum the finalPrice as spend
	select
		t.receiptId,
		w.rewardsReceiptStatus,
		sum(l.finalPrice) as spend
	-- merge the transactionItems, transactions, items and rewards table using the common
fields
	from transactionItems as l
	inner join transactions as t on l.transactionId = t.id
	inner join items as i on l.barcode = i.barcode
	inner join rewards as w on t.receiptId = w.receiptId
	-- group by the receiptId and rewardReceiptStatus
	group by t.receiptId, w.rewardsReceiptStatus
) result
```

Output:

	avg_spend_accepted	avg_spend_rejected
1	168.644646847752	2.28999996185303

By running the query, I found that the average spend for rewardReceiptStatus = 'Finished' was greater than the rewardReceiptStatus='Rejected'

Query 3:

```
-- When considering total number of items purchased from receipts with 'rewardsReceiptStatus' of
'Accepted' or 'Rejected', which is greater?

-- Display the sum of the spend based on rejected and finished categories.
-- After performing the transformations, the data did not have accepted category, so I used the
finished category instead
select
	sum(case when lower(rewardsReceiptStatus) = 'finished' then isnull(items,0) end) as
total_items_accepted,
	sum(case when lower(rewardsReceiptStatus) = 'rejected' then isnull(items,0) end) as
total_items_rejected
from
(
	-- select receiptId, rewardsReceiptStatus and count of items
	-- Items are foudn by counting the distinct barcodes of the items
	select
		t.receiptId,
		r.rewardsReceiptStatus,
		count(distinct i.barcode) as items
	-- merge the transactionItems, transactions, items and rewards table using the common
fields
	from transactionItems as ti
	inner join transactions as t on ti.transactionId = t.id
	inner join items as i on ti.barcode = i.barcode
	inner join rewards as r on t.receiptId = r.receiptId
	-- grouping the entire data on the receiptId and rewardsReceiptStatus
	group by t.receiptId, r.rewardsReceiptStatus
) result
```

Output:

	total_items_accepted	total_items_rejected
1	1037	1

By running the query, I found that the total number of items purchases for rewardReceiptStatus = 'Finished' was greater than the rewardReceiptStatus='Rejected'