

```
/*
```

```
select * from [dbo].[user];  
select * from [dbo].[transaction];  
select * from [dbo].[products];
```

```
*/
```

```
-----  
-- What are the top 5 brands by receipts scanned among users 21 and over?  
-----
```

```
select top 5 p.brand, count(t.receipt_id) as total_receipts -- selecting top 5 brands  
from [dbo].[transaction] t  
left outer join [dbo].[user] u on t.user_id = u.id -- left outer join to select records  
from transaction  
left outer join [dbo].[products] p on t.barcode = p.barcode  
where datediff(year, birth_date, getdate()) >= 21 -- calculating the age of users  
      and t.receipt_id is not null  
      and p.brand is not null  
      -- there are some data where Brand is not listed for the products (removing those  
for non null values)  
group by p.brand  
order by total_receipts desc, p.brand; -- ordering by Brand as well to be more in  
sequential result
```

```
/* result
```

```
-- commented top 7 here just in case but query is for top 5
```

```
brand          total_receipts
DOVE            3
NERDS CANDY     3
COCA-COLA       2
GREAT VALUE     2
HERSHEY'S       2
MEIJER          2
SOUR PATCH KIDS 2
```

```
*/
```

```
--
13 select top 5 p.brand, count(t.receipt_id) as total_receipts -- selecting top 5 brands
14 from [dbo].[transaction] t
15 left outer join [dbo].[user] u on t.user_id = u.id -- left outer join to select records from transaction
16 left outer join [dbo].[products] p on t.barcode = p.barcode
17 where datediff(year, birth_date, getdate()) >= 21 -- calculating the age of users
18 and t.receipt_id is not null
19 and p.brand is not null
20 -- there are some data where Brand is not listed for the products (removing those for non null values)
21 group by p.brand
22 order by total_receipts desc, p.brand; -- ordering by Brand as well to be more in sequential result
23
```

Results Messages

	brand	total_receipts
1	DOVE	3
2	NERDS CANDY	3
3	COCA-COLA	2
4	GREAT VALUE	2
5	HERSHEY'S	2

---

-- What are the top 5 brands by sales among users that have had their account for at least six months?

---

```
select top 5 p.brand, sum(t.final_sale) as total_sales -- selecting top 5
from [dbo].[transaction] t
left outer join [dbo].[user] u on t.user_id = u.id -- left outer join to select records
from transaction
left outer join [dbo].[products] p on t.barcode = p.barcode
    where datediff(month, u.created_date, getdate()) >= 6 -- calculating at least six
months from created_Month
    and u.created_date is not null
    and p.brand is not null -- there are some data where Brand is not listed for the
products(removing those for non-null values)
group by p.brand
order by total_sales desc, p.brand;
```

/\* result

-- commented top 7 here just in case but query is for top 5

brand	total_sales
CVS	72
DOVE	31
TRIDENT	24
COORS LIGHT	17
TRESEMMÉ	15
PEPPERIDGE FARM	12
GREAT VALUE	10

\*/

```
45
46 select top 5 p.brand, sum(t.final_sale) as total_sales -- selecting top 5
47 from [dbo].[transaction] t
48 left outer join [dbo].[user] u on t.user_id = u.id -- left outer join to select records from transaction
49 left outer join [dbo].[products] p on t.barcode = p.barcode
50 ... where datediff(month, u.created_date, getdate()) >= 6 -- calculating at least six months from created_Month
51 ... and u.created_date is not null
52 ... and p.brand is not null -- there are some data where Brand is not listed for the products (removing those for printing non-null values)
53 group by p.brand
54 order by total_sales desc, p.brand;
55
```

Results Messages

	brand	total_sales
1	CVS	72
2	DOVE	31
3	TRIDENT	24
4	COORS LIGHT	17
5	TRESEMMÉ	15

-----  
-- What is the percentage of sales in the Health & Wellness category by generation?  
-----

```
select age_group.generation,
sum(case when p.category_1 like '%health & wellness%' then t.final_sale else 0 end) *
100.0 / sum(t.final_sale)
-- calculating and printing percentage sales for category_1 is having 'health &
wellness'
as percentage_sales
```

```

from (
    select
        id as user_id,
        case -- differentiating age groups and giving them some fancy names
            when datediff(year, birth_date, getdate()) >= 57 then 'boomer'
            when datediff(year, birth_date, getdate()) between 41 and 56 then 'gen x'
            when datediff(year, birth_date, getdate()) between 25 and 40 then
'millennial'
            when datediff(year, birth_date, getdate()) between 10 and 24 then 'gen z'
            else 'unknown'
        end as generation
    from dbo.[user]
) as age_group
left outer join dbo.[transaction] t on age_group.user_id = t.user_id
left outer join dbo.products p on t.barcode = p.barcode
where t.final_sale is not null
group by age_group.generation
order by percentage_sales desc;

```

```
/*
```

```
result
```

generation	percentage_sales
boomer	34.351145
gen x	25.388601
millennial	13.888888
unknown	0.000000

-- the result showing - generation of %health & wellness% / total sale of that generation  
-- This tells us what portion of each generation's spending goes to the health & wellness category

\*/

```
78 select age_group.generation,
79 sum(case when p.category_1 like '%health & wellness%' then t.final_sale else 0 end) * 100.0 / sum(t.final_sale)
80 -- calculating and printing percentage sales for category_1 is having 'health & wellness'
81 as percentage_sales
82 from (
83     select
84         id as user_id,
85         case -- differentiating age groups and giving them some fancy names
86             when datediff(year, birth_date, getdate()) >= 57 then 'boomer'
87             when datediff(year, birth_date, getdate()) between 41 and 56 then 'gen x'
88             when datediff(year, birth_date, getdate()) between 25 and 40 then 'millennial'
89             when datediff(year, birth_date, getdate()) between 10 and 24 then 'gen z'
90             else 'unknown'
91         end as generation
92     from dbo.[user]
93 ) as age_group
94 left outer join dbo.[transaction] t on age_group.user_id = t.user_id
95 left outer join dbo.products p on t.barcode = p.barcode
96 where t.final_sale is not null
97 group by age_group.generation
98 order by percentage_sales desc;
99
```

Results Messages

	generation	percentage_sales
1	boomer	34.351145
2	gen x	25.388601
3	millennial	13.888888
4	unknown	0.000000