

```
select count(distinct userId) from users;
# 212
select count(distinct userId) from receipts
where userId not in
(select userId from users);
# 117
```

# There are 495 rows in the raw data table. After dropping duplicates, there are only 212 users in the users table. There are 117 records of distinct users in receipts table but these users have no information in the users table. The users data should be collected and completed in the future.

```
select count(*)
from brands
where brandCode is null or brandCode = '';
# 269
```

There are lots of null brandCode records in the brands table.

```
select month(scanDate) as month, count(*) as count
from receipts
group by 1;
```

	month	count
▶	10	2
	11	6
	1	622
	2	451
	3	38

# In the data provided, the records of receipts are not balanced among months. The information from last year and Mar 2021 is not enough for analysis so the data should be updated in time.