StratScratch 2

[StrataScratch - Primary Key Violation](https://platform.stratascratch.com/coding/2107-primary-key-violation?code_type=1)

**Primary Key Violation**

Interview Question Date: May 2022

AmazonEasyActive InterviewID 2107

30

Write a query to return all Customers (cust\_id) who are violating primary key constraints in the Customer Dimension (dim\_customer) i.e. those Customers who are present more than once in the Customer Dimension. For example if cust\_id 'C123' is present thrice then the query should return 'C123' | '3' as output.

Table: dim\_customer

Sql :

select cust\_id, count(\*) as customer\_count

from dim\_customer

group by cust\_id

having count(\*) >1;

## Output

View the output in a separate browser tab

| **cust\_id** | **customer\_count** |
| --- | --- |
| C276 | 2 |
| C281 | 2 |
| C274 | 3 |

Python :

# Import your libraries

import pandas as pd

# Start writing code

dim\_customer.head()

data1= dim\_customer.groupby(by = 'cust\_id',as\_index=False)['cust\_name'].count()

data2= pd.DataFrame(data1.loc[data1.cust\_name >1])

print(data2)

## Output

cust\_id cust\_name

1 C274 3

3 C276 2

8 C281 2