**Basic SQL analysis**

**Connect to RDS**

mysql -h mlc-testcapstone.cyaielc9bmnf.us-east-1.rds.amazonaws.com -u student -p

STUDENT123

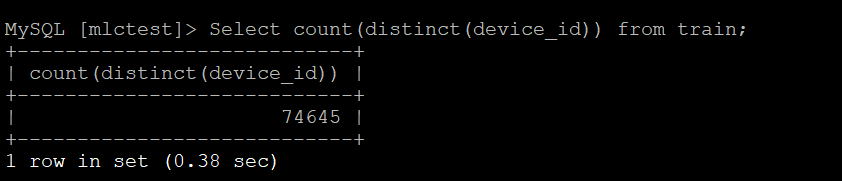
Use mlctest;

Show tables;

**Count of unique device ids in the train table**

**Query:** Select count(distinct(device\_id)) from train;

Identified 74645 unique device ids



**Check whether there are any duplicate device ids present in the brand\_device table. If yes, how many duplicates?**

**Query:** Select device\_id, count(1) cnt from brand\_device group by device\_id having cnt>1;

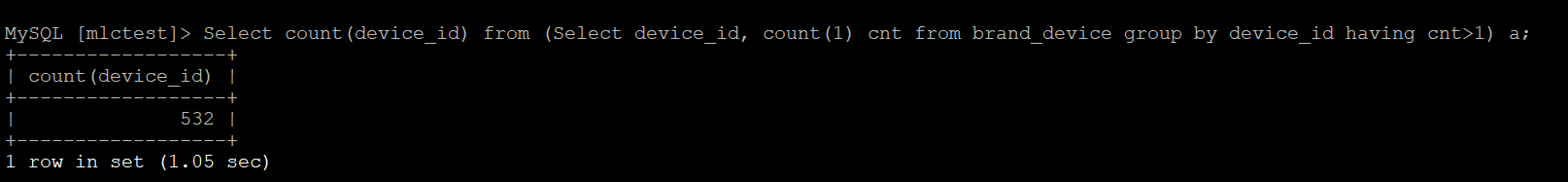
There are duplicates

A picture containing text

Description automatically generated

**Query:** Select count(device\_id) from (Select device\_id, count(1) cnt from brand\_device group by device\_id having cnt>1) a;

Identified 532 duplicates



**Number of unique phone brands from the brand\_device table**

**Query:** Select count(distinct(phone\_brand)) from brand\_device;

There are distinct 97 phone brands

Text

Description automatically generated

**Count of device ids where the latitude and longitude detail are zero, from the events table**

**Query:** Select count(device\_id) from events where latitude=0 and longitude=0;

There are 968675 rows with lat and long =0

Text

Description automatically generated