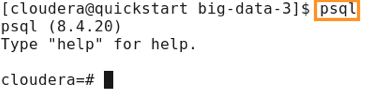
By this end of this activity, you will be able to:

1. View table and column definitions, and perform SQL queries in the Postgres shell
2. Query the contents of SQL tables
3. Filter table rows and columns
4. Combine two tables by joining on a column

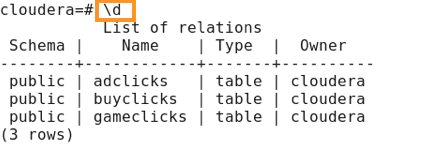
Step 1. **Open a terminal window and start Postgres shell.** Open a terminal window by clicking on the square black box on the top left of the screen.



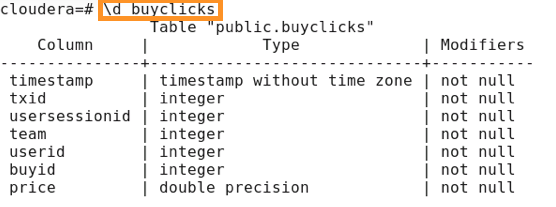
Next, start the Postgres shell by running *psql:*



Step 2. **View table and column definitions.** We can list the tables in the database with the *\d* command:



The database contains three tables: *adclicks, buyclicks,*and *gameclicks.* We can see the column definitions of the *buyclicks* table by running *\d buyclicks:*



This shows that the *buyclicks*table has seven columns, and what each column name and data type is.

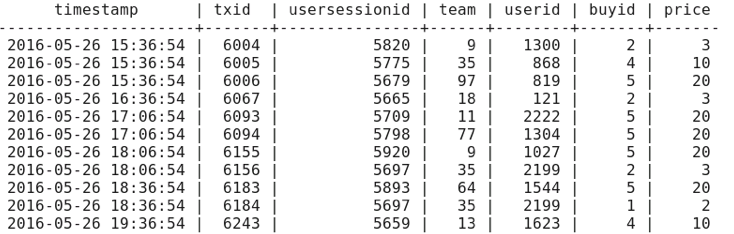
Step 3. **Query table.** We can run the following command to view the contents of the *buyclicks* table:

1

select \* from buyclicks;

The *select \** means we want to query all the columns, and *from buyclicks* denotes which table to query. Note that all query commands in the Postgres shell must end with a semi-colon.

The result of the query is:



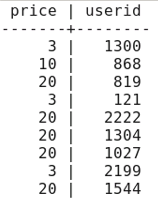
You can hit *<space>* to scroll down, and *q*to quit.

Step 4. **Filter rows and columns.** We can query only the *price* and *userid*columns with the following command:

1

select price, userid from buyclicks;

The result of this query is:

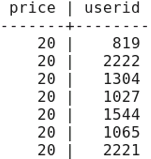


We can also query rows that match a specific criteria. For example, the following command queries only rows with a price greater than 10:

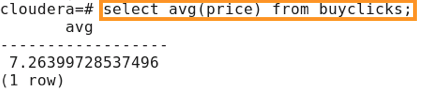
1

select price, userid from buyclicks where price > 10;

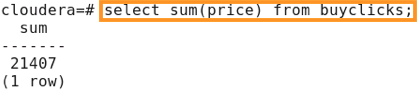
The result is:



Step 5. **Perform aggregate operations.** The SQL language provides many aggregate operations. We can calculate the average price:

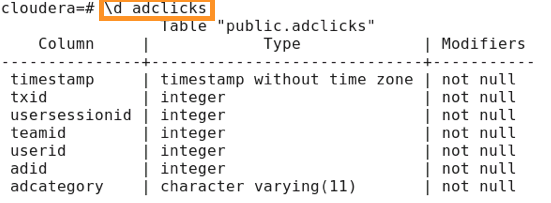


We can also calculate the total price:



The complete list of aggegrate functions for Postgres 8.4 (the version installed on the Cloudera VM) can be found here: <https://www.postgresql.org/docs/8.4/static/functions-aggregate.html>

Step 6. **Combine two tables.** We combine the contents of two tables by matching or joining on a single column. If we look at the definition of the *adclicks* table:



We see that *adclicks* also has a column named *userid.* The following query combines the *adclicks* and *buyclicks* tables on the *userid* column in both tables:

1

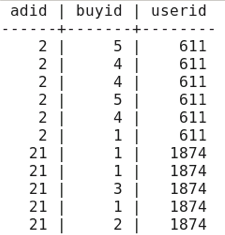
2

select adid, buyid, adclicks.userid

from adclicks join buyclicks on adclicks.userid = buyclicks.userid;

This query shows the columns *adid* and *userid* from the *adclicks* table, and the *buyid*column from the *buyclicks* table. The *from adclicks join buyclicks* denotes that we want to combine these two tables, and *on adclicks.userid = buyclicks.userid* denotes which two columns to use when the tables are combined.

The result of the query is:



Enter *\q* to quit the Postgres shell.