Learn SQL by Calculating Customer Lifetime Value

# Part 1: Introduction

The purpose of this project is to learn fundamental SQL queries using a typical data analysis scenario. As I am a product manager at an e-commerce (or SaaS) startup, I calculate the Customer Lifetime Value (CLV) – how much revenue they bring to the company over their lifetime. We can use this critical metric to allow us to make stronger business decisions and customize our products effectively for each customer.

This project uses SQLite3 and the two tables of interest (“users” and “payments”) are read from bootstrap.sql taken from <http://bit.ly/sql-starter>. We conclude that a direct campaign and customers with user\_id of 8 and 3 generate the least revenue. Thus the company needs to focus on improving services offered to or produced through these elements.

# Part 2: Project

**Loading windows command prompt and running sqlite3.exe**

> cd C:/sqlite

> sqlite3.exe

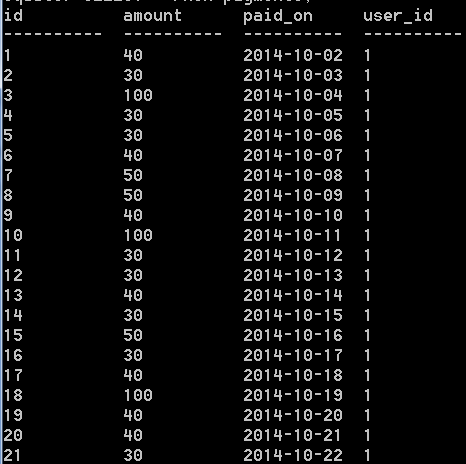
**Loading tables**

> .read bootstrap.sql

> .tables payments users

> SELECT \* FROM users;

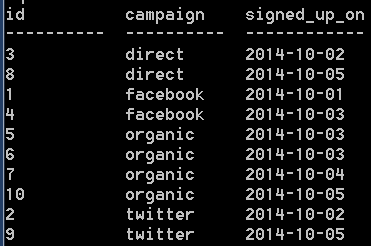
> SELECT \* FROM payments;





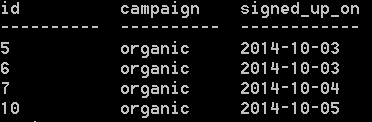
**Ordering by campaign in alphabetical order**

> SELECT \* FROM users ORDER BY campaign;



**Filtering with WHERE**

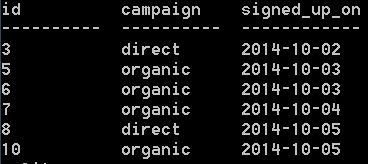
> SELECT \* FROM users WHERE campaign = ‘organic’;



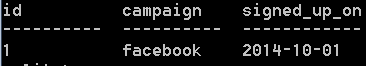
> SELECT \* FROM users WHERE campaign IN (‘facebook’, ‘twitter’);



> SELECT \* FROM users WHERE campaign NOT IN (‘facebook’, ‘twitter’);



> SELECT \* FROM users WHERE campaign IN (‘facebook’, ‘twitter’) AND signed\_up\_on = ‘2014-10-01’;



> SELECT \* FROM users WHERE campaign = ‘organic’ OR signed\_up\_on < ‘2014-10-04’;



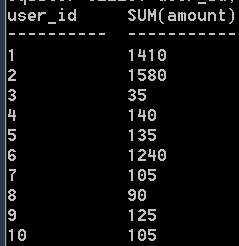
**Filtering and sorting**

> SELECT \* FROM users WHERE campaign in (‘facebook’, ‘twitter’) ORDER BY campaign;



**Group by: SQL’s PivotTable**

> SELECT user\_id, SUM(amount) FROM payments GROUP BY user\_id;



**JOIN: connecting multiple sources of information**

> SELECT user\_id, SUM(amount) as clv, campaign FROM users JOIN payments ON users.id = user\_id GROUP by user\_id ORDER BY clv DESC;

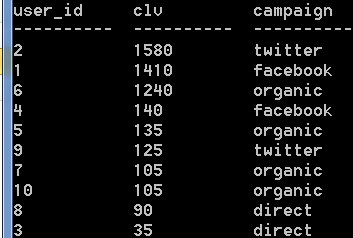


Figure : CLV per user

**The Other CLV: Campaign Lifetime Value**

> SELECT campaign, SUM(amount) as campaign\_value FROM users JOIN payments ON users.id = user\_id GROUP by campaign ORDER BY campaign\_value DESC;

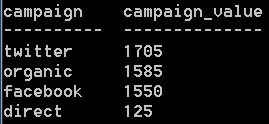


Figure : Value according to campaign method

# Part 3: Conclusion

I conclude from…

* figure 1 that customers with user\_id of 2 and 1 generate the most revenue over their lifetime. Whilst customers with user\_id of 8 and 3 generate the least revenue. Hence the company would do best to examine these customers’ preferences and produce more tailored products and services more favourable to them.
* figure 2 that twitter generates the most revenue, whilst direct the least. Hence the company would do best to examine why a direct campaign is not working successfully, and strive to rework it so as to be more profitable.