Learn SQL From Scratch: First/Last Touch Capstone

Justine Chong | December 7, 2018

Table of Contents

Get Familiar With the Company: CoolTShirts (CTS)

What is the user journey?

Optimize the Campaign Budget

Please note that all queries will be on the right hand side of the slide and included in a separate SQL file

Getting to know the company

How many campaigns & sources does CTS use?

8 Campaigns: cool-tshirts-search, getting-to-know-cooltshits, interview-with-cool-tshirts-founder, paid-search, retargetting-ad, retargetting-campaign, ten-crazy-cool-tshirts-facts, weekly-newsletter

6 Sources: Buzzfeed, E-Mail, Facebook, Google, Medium, NY Times

I used 2 SQL queries to find the total number of campaigns, sources and their names.

Campaign Query

```
SELECT utm_campaign,
COUNT(utm_campaign)
FROM page_visits
GROUP BY utm_campaign;
```

Source Query

```
SELECT utm_source,
COUNT(utm_source)
FROM page_visits
GROUP BY utm_source;
```
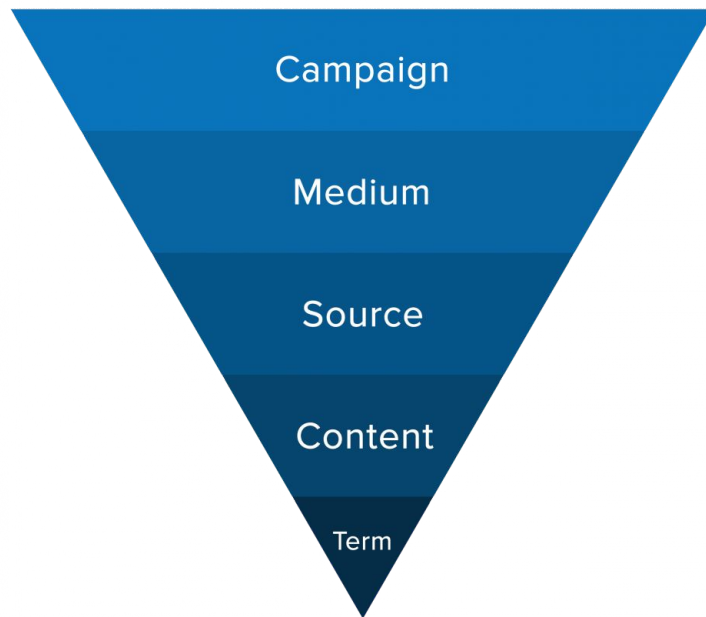
How are utm_campaign & utm source related?

Utm_campaign and utm_source work together to inform the company where their traffic/purchases are coming from.

A "utm_campaign" references a particular article, piece of content, or ad/marketing campaign.

While "utm_source" refers to where a user came into contact with the "campaign" and references the multiple channels that a campaign can be distributed through.
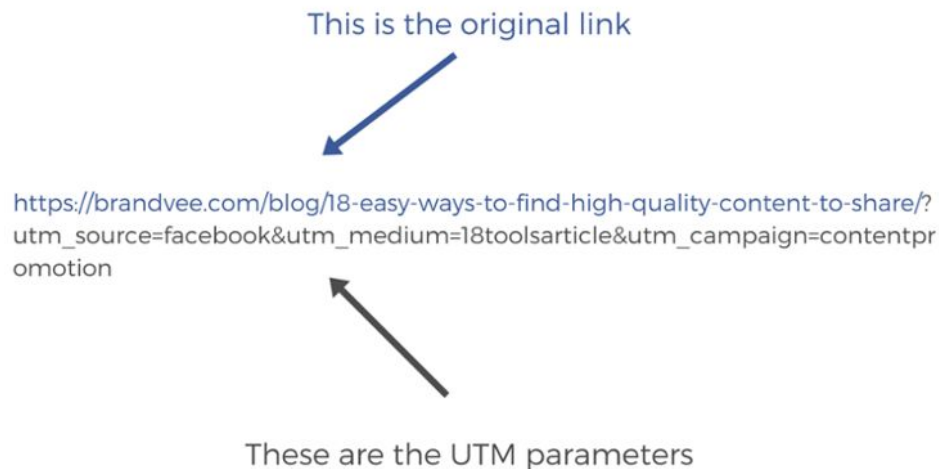
Hierarchy of UTM Parameters

How are utm_campaign and UTM_source different?

The UTM Parameters for campaigns and sources are related, but inherently different.

UTM_Campaign is more general and relates to a specific piece of content, marketing push or "campaign."

UTM_Sources are more granular and identifies where traffic came from whether it's through a specific website, type of ad, email, or newsletter. Campaigns can also be run over multiple sources.

This is the original link

https://brandvee.com/blog/18-easy-ways-to-find-high-quality-content-to-share/?utm_source=facebook&utm_medium=18toolsarticle&utm_campaign=contentpromotion

These are the UTM parameters

What Pages does CTS have on their website?

CoolTShirts has four pages on their website:

- 1-Landing Page
- 2-Shopping Cart
- 3-Checkout
- 4-Purchase

I found this out using the query on the right hand side of the page.

Page Count Query

```
SELECT DISTINCT page_name
FROM page_visits;
```

What is the user Journey

How many F. touches is each campaign responsible for?

I used the query on the right hand of the page to find out how many first touches each campaign was responsible for:

interview-with-cooltshirts-founder	622
getting-to-know-cooltshirts	612
10-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

First Touch SQL Query:

```
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign, COUNT(first_touch_at)
FROM first_touch ft
JOIN page_visits pv
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
    GROUP BY utm_campaign
    ORDER BY count(first_touch_at) DESC;
```

How many L. touches is each campaign responsible for?

weekly-newsletters	447
retargetting-ad	443
retargetting-campaign	245
getting-to-know-cooltshirts	232
10-crazy-cool-tshirts-facts	190
interview-with-cooltshirts-founder	184
paid-search	178
cooltshirts-search	60

I used the query below to find how many Last Touches each campaign was responsible for.

Last Touch Campaign Query:

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign, COUNT(last_touch_at)
FROM last_touch lt
JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
    GROUP BY utm_campaign
    ORDER BY count(last_touch_at) DESC;
```

How many visitors make a purchase?

361 users out of 1979 made a purchase through CoolTshirts.

I found the total number of users and narrowed that down to specific users who successfully made it to Page 4 - Purchase.

How Many Visitors Made It To the Purchase Page:

SELECT COUNT(DISTINCT user_id)
FROM page_visits
WHERE page_name LIKE '%4%';

Total Visitor Query:
SELECT COUNT(DISTINCT user_id)
FROM page_visits;

How many L.touches on the purchase page is each campaign responsible for?

weekly-newsletter	114
retargetting-ad	112
retargetting-campaign	53
paid-search	52
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirtsfacts	9
interview-with-cool-tshirts-dounder	7
cool-tshirts-search	2

Using the last touch query that was written before, I limited the results to only include results that hit on "4-Purchase" Page
----
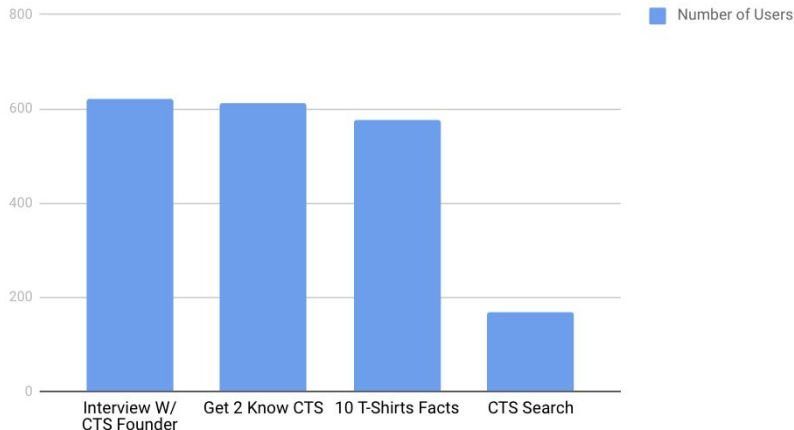**Purchase Last Touch Queries**
```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign, COUNT(last_touch_at)
FROM last_touch lt
JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
    WHERE page_name LIKE '%4%'
    GROUP BY utm_campaign
    ORDER BY count(last_touch_at) DESC;
```

What is the typical user journey?

The typical user journey starts with a first touch, matriculation through interim pages(shopping cart or checkout), and then ends on a last touch. The last touch doesn't always have to end in purchase.

Taking the data from our query on "First Touches" we can see the most popular introduction into the website:

First Touches By Campaign

Since campaigns can be run over multiple sources, we can take a deeper look into the two most successful Campaigns: Interview w/ the CTS Founder and Get To Know CTS to see where their users came from.

After running the query below I found out that all users from the interview campaign came from Medium and all users from the "Get to Know" campaign came from the New York Times

Finding Out Where the Campaign Sources Came From
```
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign, utm_source, COUNT(first_touch_at)
FROM first_touch ft
JOIN page_visits pv
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
WHERE utm_campaign = 'getting-to-know-cool-tshirts'
OR utm_campaign = 'interview-with-cool-tshirts-founder'
GROUP BY utm_source
ORDER BY count(first_touch_at) DESC;
```

What is a typical user journey? cont'd

To make sure that I didn't format the Query incorrectly, I took a look at each distinct combination of utm_campaign and utm_source.

Which gave me the following combination of campaigns and sources.

SQL Query:

SELECT DISTINCT utm_campaign, utm_source
FROM page_visits;

getting-to-know-cool-tshirts	nytimes
weekly-newsletter	email
10-crazy-tshirts-facts	buzzfeed
retargetting-campaign	email
retargetting-ad	facebook
interview-with-paid-t-shirts-founder	medium
paid-search	google
cool-tshirts-search	google

What is a typical user journey? Cont'd

I then realized that a "First Touch" doesn't always necessarily have to be the "landing page." So I ran a query and found which campaigns led to the highest views of the landing page.

interview-with-cool-tshirts-founder	625
getting-to-know-cool-tshirts	617
ten-crazy-cool-tshirts-facts	587
cool-tshirts-search	171

Landing Page Query:

```
SELECT page_name, utm_campaign,
COUNT(utm_campaign), utm_source
FROM page_visits
WHERE page_name LIKE '%1%'
GROUP BY utm_campaign
ORDER BY COUNT(utm_campaign) DESC;
```

What is a typical user journey? Cont'd

I also wanted to find out how many people got to the Shopping Cart Page from a specific campaign and ran the query to the right.

getting-to-know-cool-tshirts	682
ten-crazy-cool-tshirts-facts	570
interview-with-cool-tshirts-founder	515
cool-tshirts-search	133

"Shopping Cart Query":

```
SELECT page_name, utm_campaign,
COUNT(utm_campaign), utm_source
FROM page_visits
WHERE page_name LIKE '%2%'
GROUP BY utm_campaign
ORDER BY COUNT(utm_campaign) DESC;
```

What is a typical user journey? cont'd

Following the same curiosity from the previous slide – I wanted to see how many people made it to the "Check Out Screen"

"Check Out Query":

```
SELECT page_name, utm_campaign,
COUNT(utm_campaign), utm_source
FROM page_visits
WHERE page_name LIKE '%3%'
GROUP BY utm_campaign
ORDER BY COUNT(utm_campaign) DESC;
```

weekly-newsletter	450
retargetting-ad	445
retargetting-campaign	246
paid-search	179
getting-to-know-cool-tshirts	41
ten-crazy-cool-tshirts-facts	32
interview-with-cool-tshirts-founder	31
cool-tshirts-search	7

WHAT IS A TYPICAL USER JOURNEY? CONT'D

Finally, I wanted to see how many users successfully got to the purchase screen and what campaign led to the most purchases.

"Purchase Query":

```
SELECT page_name, utm_campaign,
COUNT(utm_campaign), utm_source
FROM page_visits
WHERE page_name LIKE '%4%'
GROUP BY utm_campaign
ORDER BY COUNT(utm_campaign) DESC;
```

weekly-newsletter	115
retargetting-ad	113
retargetting-campaign	54
paid-search	54
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	7
cool-tshirts-search	2

What is the typical user journey?

Knowing the overall flow and what campaigns led to the most views of pages 1,2,3 & 4. I ran the query that found first touches and modified it to look for the last touch of by campaign.

Taking the data from our query on "Last Touches" we can see when users had the last action on the site via a specific campaign.

Finding Out Where the Campaign Sources Came From
```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign, utm_source, COUNT(last_touch_at)
FROM last_touch lt
JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
GROUP BY utm_source
ORDER BY count(last_touch_at) DESC;
```

retargetting-campaign	692
retargetting-ad	443
cool-tshirts-search	238
getting-to-know-cool-tshirts	232
ten-crazy-cool-tshirts-facts	190
interview-with-cool-tshirts-founder	182

What campaigns should CTS Re-invest in?

Taking a look at the tables created from the SQL queries run, it looks like "content marketing" pieces like Interviews w/ the founder, Get to Know CTS, and 10 CTS Facts were the most successful in drawing in new leads.

However, taking a look into how many of those leads actually converted and purchased something, it was the Weekly Newsletter and the Retargeting Campaign that were most successful.

I would advise Cool T-shirts to reinvest in the following campaigns:
- Interview with CTS Founder
- Get to Know CTS
- 10 T-Shirt Facts
- The Weekly Newsletter
- Retargeting Campaign

In order to bring in the most leads and capitalize on the Newsletter and Retargeting Campaign's higher purchase rates.