



Attribution Queries

Learn SQL from Scratch

Verena Stecher

04/06/2019

Table of Contents

1. Get familiar with CoolTShirts
2. What is the user journey?
3. Optimize the campaign budget

1. Get familiar with CoolTShirts

1.1 How many campaigns and sources does CoolTShirts use and how are they related?

Be sure to explain the difference between 'utm_campaign' and 'utm_source'.

- CoolTShirts runs 8 strategic campaigns ('utm_campaign').
- CoolTShirts uses 6 distinct sources which sent the traffic ('utm_source').
- Please refer to the table on the right hand side to see the relation between the different campaigns ('utm_campaign') and sources ('utm_source').

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

1.2 What pages are on their website?

There are four distinct values ('page_name's) on their website:

- Please refer to the table on the right hand side for a list of the four different pages on the CoolTShirts website.

page_name
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

2. What is the
user journey?

2.1 How many first touches is each campaign responsible for?

Please refer to the table below to extract how many first touches ('COUNT(utm_campaign)') each campaign ('utm_campaign') is responsible for.

utm_campaign	COUNT(utm_campaign)
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

```
WITH first_touch AS (  
    SELECT user_id, MIN(timestamp) as  
    first_touch_at  
    FROM page_visits  
    GROUP BY user_id)  
SELECT ft.user_id, ft.first_touch_at,  
pv.utm_source, pv.utm_campaign,  
COUNT(utm_campaign)  
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 5 DESC;
```

2.2 How many last touches is each campaign responsible for?

Please refer to the table below to extract how many last touches ('COUNT(utm_campaign)') each campaign ('utm_campaign') is responsible for.

utm_campaign	COUNT(utm_campaign)
weekly-newsletter	447
retargeting-ad	443
retargeting-campaign	245
getting-to-know-cool-tshirts	232
ten-crazy-cool-tshirts-facts	190
interview-with-cool-tshirts-founder	184
paid-search	178
cool-tshirts-search	60

```
WITH last_touch AS (  
    SELECT user_id, MAX(timestamp) as  
    last_touch_at  
    FROM page_visits  
    GROUP BY user_id)  
SELECT lt.user_id, lt.last_touch_at,  
pv.utm_source, pv.utm_campaign,  
COUNT(utm_campaign)  
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 5 DESC;
```


2.3 How many visitors make a purchase?

361 visitors made a purchase.

COUNT(DISTINCT user_id)
361

```
SELECT COUNT(DISTINCT user_id)
FROM page_visits
WHERE page_name = '4 - purchase';
```

2.4 How many last touches *on the purchase page* is each campaign responsible for?

Please refer to the table below to extract how many last touches on the purchase page ('COUNT(utm_campaign)') each campaign ('utm_campaign') is responsible for.

utm_campaign	COUNT(utm_campaign)
weekly-newsletter	115
retargeting-ad	113
retargeting-campaign	54
paid-search	52
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	7
cool-tshirts-search	2

```
WITH last_touch AS (  
    SELECT user_id, MAX(timestamp) as  
    last_touch_at  
    FROM page_visits  
    WHERE page_name = '4 - purchase'  
    GROUP BY user_id)  
SELECT lt.user_id, lt.last_touch_at,  
pv.utm_source, pv.utm_campaign,  
COUNT(utm_campaign)  
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 5 DESC;
```

2.5 What is the typical user journey?

A typical user journey looks as follows:

First touches

- 91.5% of total first touches are generated through campaigns run on publishing platforms (medium), newspaper (nytimes) and digital media (buzzfeed). However, only 7.5% of the traffic from those sources results in a purchase.
- Most people that end up purchasing a product are not initially landing on the CoolTShirts page through the listed marketing campaigns.

Last touches

- 45% of last touches are generated through weekly newsletters (email) and retargeting ads (facebook). 25.6% of these last touches continued to purchases completed.
- 24% of last touches are generated through retargeting campaigns (email) and paid searches (google). 22.2% of these last touches continued to purchases completed.

Purchases

- 63.2% of total purchases are generated through campaigns run on e-mail (newsletter) and facebook (retargeting ad).
- 29% of total purchases are generated through retargeting campaigns (email) and paid searches (google).

2.5 What is the typical user journey?

Overview table:

utm_campaign	utm_source	first-touches	last-touches	purchases completed
weekly-newsletter	email		447	115
retargeting-ad	facebook		443	113
retargeting-campaign	email		245	54
paid-search	google		232	52
getting-to-know-cool-tshirts	nytimes	612	190	9
ten-crazy-cool-tshirts-facts	buzzfeed	576	184	9
interview-with-cool-tshirts-founder	medium	622	178	7
cool-tshirts-search	google	169	60	2

3. Optimize the campaign budget

3.1 CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

Variables to consider

This question is dependent on two key variables:

1. What are the matrixes that CoolTShirts has set to meet with these campaigns? These matrixes could vary e.g. brand awareness, purchases, amount of total sales.
2. A second interesting variable to take into consideration is the costs for each of this campaigns as the cost effectiveness of running these campaigns could be a deciding factor as well i.e. the 'retargetting' campaign (email) will have different costs compared to a 'paid-search' (google).

Assumptions

Ad 1. I am deciding to choose the key matrix of purchases completed.

Ad 2. As there is no information provided with regards to costs per campaign I do not include it in the option appraisal.

I am deciding to choose the key matrix of purchases completed.

3.1 CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

Decision

In order to drive purchases I recommend CoolTShirts to re-invest in the following 5 campaigns:

utm_campaign	utm_source	first-touches	last-touches	purchases completed
weekly-newsletter	email		447	115
retargetting-ad	facebook		443	113
retargetting-campaign	email		245	54
paid-search	google		232	52
getting-to-know-cool-tshirts	nytimes	612	190	9

Rationale

As the matrix of the marketing campaigns is to drive sales I am recommending to re-invest in the 5 campaigns that have generated 95% of total completed purchases.