



First- and Last-Touch Attribution with CoolTShirts.com

Capstone Project - Learn SQL from Scratch

Mariusz Wolski

2018.09.24

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?

- CTS started **eight campaigns** at **six sources** to increase website visits and purchases. The campaigns are listed in the table below – `utm_campaigns`. Every campaign is related to a certain source (`utm_source`), e.g. the „getting-to-know-cool-tshirts“ campaign is running at NY Times, but it is not at buzzfeed. The same is with „...search“ campaigns, they are running through google.
- **UTM parameters** capture when and how a user finds the site. Site owners use special links containing UTM parameters in their ads, blog posts, and other sources. When a user clicks one, a row is added to a database describing their page visit. Web developers, marketers, and data analysts use that information to improve their sources (sometimes called channels or touchpoints) online. If an ad campaign drives a lot of visits to their site, then they know that source is working!

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 the CTS website?

To get to know which section of the website is visited, they have following names:

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?

- „**First touch**“ is the first time a user was exposed to CoolTshirts.com and it is attributed (related) to a certain source

COUNT(*)	utm_campaign	utm_source
622	interview-with-cool-tshirts-founder	medium
612	getting-to-know-cool-tshirts	nytimes
576	ten-crazy-cool-tshirts-facts	buzzfeed
169	cool-tshirts-search	google

```
WITH first_touch AS (  
    SELECT user_id,  
           MIN(timestamp) as first_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
ft_utm AS (  
    SELECT ft.user_id,  
           ft.first_touch_at,  
           pv.utm_source,  
           pv.utm_campaign  
    FROM first_touch ft  
    JOIN page_visits pv  
    ON ft.user_id = pv.user_id  
    AND ft.first_touch_at = pv.timestamp)  
SELECT COUNT(*),  
       ft_utm.utm_campaign,  
       ft_utm.utm_source  
FROM ft_utm  
GROUP BY 2  
ORDER BY 1 DESC;
```

2.2 How many last touches is each campaign responsible for?

- „**Last touch**” is the last time a user was on CoolTshirts.com and as well as first touch, it is attributed (related) to a certain source

COUNT(*)	utm_campaign	utm_source
447	weekly-newsletter	email
443	retargeting-ad	facebook
245	retargeting-campaign	email
232	getting-to-know-cool-tshirts	nytimes
190	ten-crazy-cool-tshirts-facts	buzzfeed
184	interview-with-cool-tshirts-founder	medium
178	paid-search	google
60	cool-tshirts-search	google

```
WITH last_touch AS (  
    SELECT user_id,  
           MAX(timestamp) as last_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
lt_utm AS (  
    SELECT lt.user_id,  
           lt.last_touch_at,  
           pv.utm_source,  
           pv.utm_campaign  
    FROM last_touch AS 'lt'  
    JOIN page_visits AS 'pv'  
      ON lt.user_id = pv.user_id  
     AND lt.last_touch_at = pv.timestamp)  
SELECT COUNT(*),  
       lt_utm.utm_campaign,  
       lt_utm.utm_source  
FROM lt_utm  
GROUP BY 2  
ORDER BY 1 DESC;
```


2.3 How many visitors make a purchase?

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

The total number of users/visitors who make a purchase is **361** (all campaigns together).
With a total number of **1979 users**, it is a **18%** rate of who make a purchase.

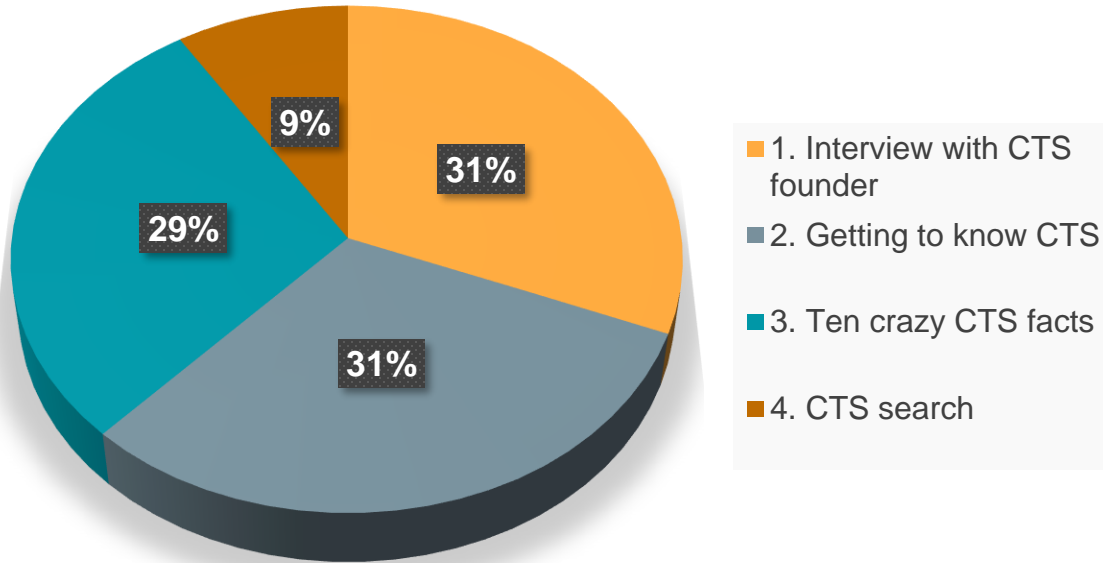
2.4 How many last touches on the purchase page is each campaign responsible for?

```
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),  
lt_utm AS (  
    SELECT lt.user_id,  
           lt.last_touch_at,  
           pv.utm_source,  
           pv.utm_campaign  
    FROM last_touch AS 'lt'  
    JOIN page_visits AS 'pv'  
      ON lt.user_id = pv.user_id  
      AND lt.last_touch_at = pv.timestamp)  
SELECT COUNT(*),  
       lt_utm.utm_campaign,  
       lt_utm.utm_source  
FROM lt_utm  
GROUP BY 2  
ORDER BY 1 DESC;
```

COUNT(*)	utm_campaign	utm_source
115	weekly-newsletter	email
113	retargetting-ad	facebook
54	retargetting-campaign	email
52	paid-search	google
9	getting-to-know-cool-tshirts	nytimes
9	ten-crazy-cool-tshirts-facts	buzzfeed
7	interview-with-cool-tshirts-founder	medium
2	cool-tshirts-search	google

2.4 What is the typical user journey?

FIRST TOUCH BY CAMPAIGN

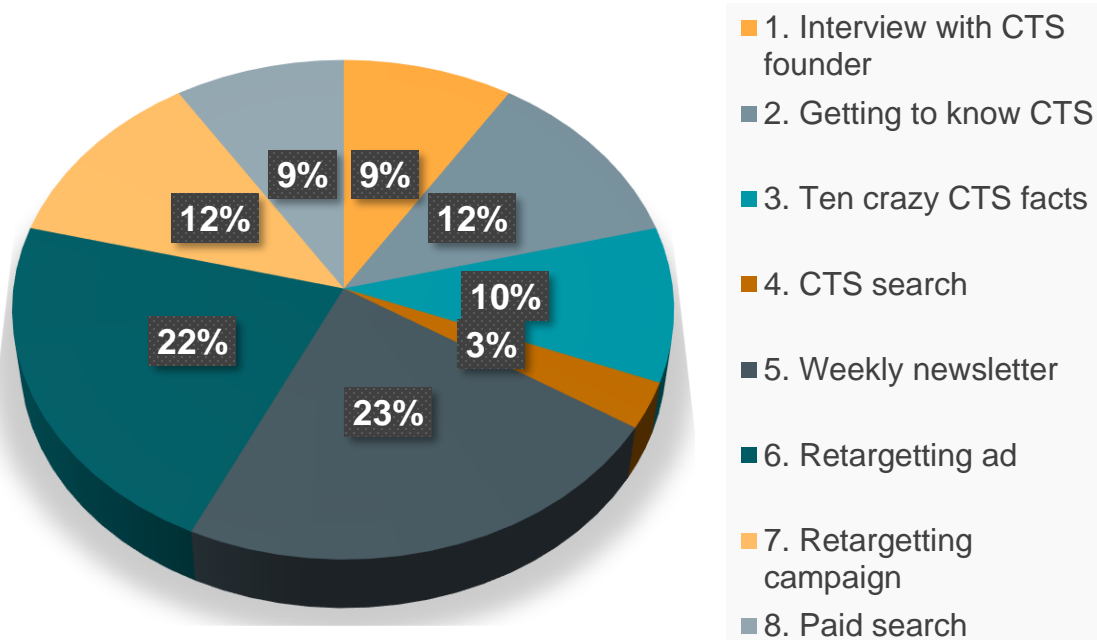


91% users getting first to the website through articles

- Interview with CTS founder
 - Getting to know CTS
 - Ten crazy facts about CTS
- And not through searches

2.4 What is the typical user journey?

LAST TOUCH BY CAMPAIGN

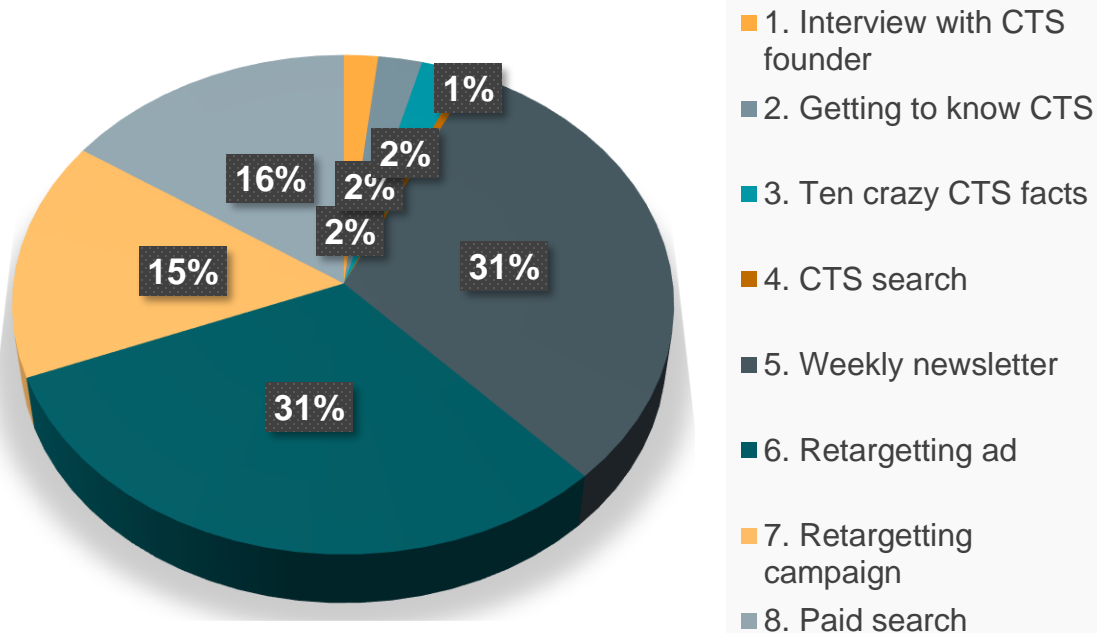


Most users need a 2nd campaign to get back to the website and eventually making a purchase.

For getting back the articles are not enough.

2.4 What is the typical user journey?

PURCHASE BY CAMPAIGN



78% of purchases were generated by

- Weekly newsletter
- Retargetting ad
- Paid search
- Retargetting campaign

3. Optimize the campaign budget

3.1 Which campaigns should be invested in?

First touch	utm_campaign
31%	interview-with-cool-tshirts-founder
31%	getting-to-know-cool-tshirts
29%	ten-crazy-cool-tshirts-facts

Purchase	utm_campaign
31%	Weekly newsletter
31%	Retargeting-ad

The tables above show in which 5 campaigns should be re-invested.

The articles should be continued, because of high user-attraction efficiency, together with two most valuable, leading-to-purchase campaigns