



Capstone: Attribution

Learn SQL from Scratch

Grace Clapperton-Rees

27 June 2018

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?

Campaigns are the different marketing messages used to promote CoolTshirts. These campaigns are communicated via different *Sources*.

To calculate the number of campaigns and sources CoolTShirts uses you need to count the unique number of each. To show their relationship you need to list the unique combinations of campaigns and sources.

CoolTshirts have 8 different campaigns that are communicated through 6 different sources.

```
/* Question 1 - 1 */
/* Counts the number of unique campaigns */
SELECT COUNT(DISTINCT utm_campaign) AS "Number of distinct campaigns"
FROM page_visits;

/* Counts the number of unique sources */
SELECT COUNT(DISTINCT utm_source) AS "Number of distinct sources"
FROM page_visits;

/* Lists the campaigns per source */
SELECT DISTINCT utm_source AS "Source",
               utm_campaign AS "Campaign"
FROM page_visits;
```

Query Results	
Number of distinct campaigns	
8	
Number of distinct sources	
6	
Source	Campaign
nytimes	getting-to-know-cool-tshirts
email	weekly-newsletter
buzzfeed	ten-crazy-cool-tshirts-facts
email	retargetting-campaign
facebook	retargetting-ad
medium	interview-with-cool-tshirts-founder
google	paid-search
google	cool-tshirts-search

1.2 What pages are on CoolTShirts website?

To find what pages are on the CoolTShirts website you need to list all the unique page names.

There are four main pages on CoolTShirts website.

```
/* Question 1 - 2 */  
/* Lists the pages on their website */  
SELECT DISTINCT page_name AS "Pages"  
FROM page_visits;
```

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

2. User Journey

2.1 How many first touches is each campaign responsible for?

To ascertain which campaign first brought a user to the CoolTShirts website you need identify the users minimum date and look at the corresponding campaign.

The 'Interview with CoolTShirts founder' on the Medium brought in the most new users.

The weekly newsletter, retargeting campaign, retargeting ad and paid search campaigns didn't bring in any new users to the website. This makes sense as these four campaigns are mainly targeting users who have already visited the site to come back and make a purchase.

Source	Campaign	Number of first touches
medium	interview-with-cool-tshirts-founder	622
nytimes	getting-to-know-cool-tshirts	612
buzzfeed	ten-crazy-cool-tshirts-facts	576
google	cool-tshirts-search	169

```
/* Question 2 - 1 */
/* Counts the number of first touches per source and campaign by:
-creating a temporary table that records the first touches by user id
-creating a temporary table that joins the temporary first touches table
with the page visits table
-listing and counting the occurrences where first touch has a source and a
campaign*/
WITH first_touch AS (
    SELECT user_id,
           MIN(timestamp) AS first_touch_at
    FROM page_visits
    GROUP BY user_id),
ft_attr 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 ft_attr.utm_source AS "Source",
       ft_attr.utm_campaign AS "Campaign",
       COUNT (*) AS "Number of first touches"
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

2.2 How many last touches is each campaign responsible for?

To ascertain which campaign last brought a user to the CoolTShirts website you need identify the user's maximum date and look at the corresponding campaign.

All campaigns have resulted in last touches, unlike first touches. The campaigns that were not present in the first touches appear to be resulting in the most amount of last touches (with the exception of the paid search campaign).

Source	Campaign	Number of last touches
email	weekly-newsletter	447
facebook	retargeting-ad	443
email	retargeting-campaign	245
nytimes	getting-to-know-cool-tshirts	232
buzzfeed	ten-crazy-cool-tshirts-facts	190
medium	interview-with-cool-tshirts-founder	184
google	paid-search	178
google	cool-tshirts-search	60

```
/* Question 2 - 2 */
/* Counts the number of last touches per source and campaign by:
-creating a temporary table that records the last touches by user id
-creating a temporary table that joins the temporary last touches table
with the page visits table
-listing and counting the occurrences where last touch has a source and a
campaign*/

WITH last_touch AS (
    SELECT user_id,
           MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id),
lt_attr AS (
    SELECT lt.user_id,
           lt.last_touch_at,
           pv.utm_source,
           pv.utm_campaign,
           pv.page_name
    FROM last_touch lt
    JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source AS "Source",
       lt_attr.utm_campaign AS "Campaign",
       COUNT (*) AS "Number of last touches"
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```


2.3 How many visitors make a purchase?

This can be calculated by counting the number of unique users who visited the purchase page on CoolTShirts's website.

Of the 1,979 users who visited CoolTShirts (calculated by adding together all the first touches from slide 2.1), 361 users made a purchase. This is equivalent to saying that when a user visits CoolTShirts there is an 18.24% chance that they'll make a purchase.

Numer of visitors who made a purchase
361

```
/* Question 2 - 3 */  
/* Counts the number of unique purchases made by users*/  
SELECT COUNT (DISTINCT user_id) AS "Numer of visitors who made a purchase"  
FROM page_visits  
WHERE page_name = '4 - purchase';
```

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

To ascertain which campaign brought a user to the CoolTShirts website purchase page (i.e. they made a purchase) you need identify the user's maximum date and look at the corresponding campaign (as per slide 2.2) and then limit the search to only the last touches that were on the purchase page.

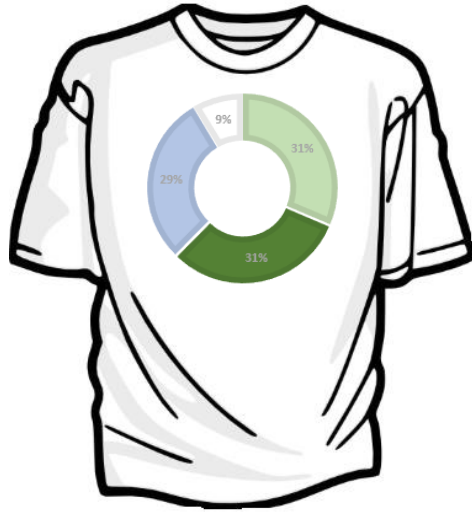
The campaigns aimed at re-targeting users who have already visited the site before are clear winners for getting users across the purchase line!

Source	Campaign	Number of last touches on the purchase page
email	weekly-newsletter	115
facebook	retargeting-ad	113
email	retargeting-campaign	54
google	paid-search	52
buzzfeed	ten-crazy-cool-tshirts-facts	9
nytimes	getting-to-know-cool-tshirts	9
medium	interview-with-cool-tshirts-founder	7
google	cool-tshirts-search	2

```
/* Question 2 - 4 */
/* Counts the number of last touches that occurred on the purchase page,
per source and campaign by:
-creating a temporary table that records the last touches that occurred on
the purchase page by user id
-creating a temporary table that joins the temporary last touches table
with the page visits table
-listing and counting the occurrences where last touch has a source and a
campaign*/
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_attr AS (
    SELECT lt.user_id,
           lt.last_touch_at,
           pv.utm_source,
           pv.utm_campaign,
           pv.page_name
    FROM last_touch lt
    JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source AS "Source",
       lt_attr.utm_campaign AS "Campaign",
       COUNT (*) AS "Number of last touches on the purchase page"
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

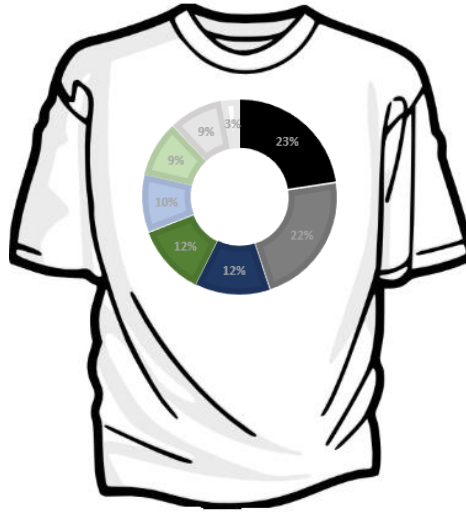
2.5 What is the typical user journey?

What campaigns help users initially discover CoolTshirts?



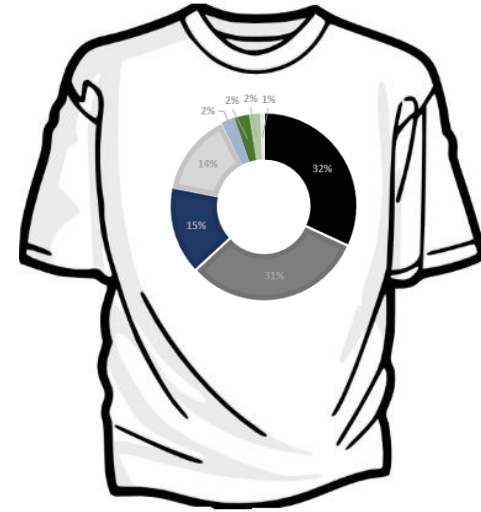
■ Weekly Newsletter
■ Ten Crazy Cool Tshirts Facts

What campaigns last brought a user to CoolTShirts?



■ Retargeting Ad
■ Retargeting Campaign
■ Interview with CoolTshirts Founder
■ Paid Search

What campaigns result in a user purchasing a CoolTShirt?



■ Getting to Know CoolTShirts
■ CoolTShirts Search

Typically users first visit the CoolTShirts website as a result of the three campaigns **Interview with CoolTShirts Founder**, **Getting to Know CoolTShirts** or **Ten Crazy Cool TShirts Facts**. However, these three campaigns aren't particularly good at getting users to make a purchase (in fact only 6% of user purchases are made solely from these three campaigns). For the remaining purchases, other campaigns are required to remind the users about the CoolTShirts website before they make a purchase. Typically users who made a purchase were successfully retargeted using the **Weekly Newsletter**, **Retargeting Ad**, **Retargeting Campaign** or **Paid Search**. 92% of purchases were as a result of one of those 4 campaigns reminding the user about CoolTShirts.

3. Optimise the Campaign Budget

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

As per slide 2.5 there are 7 campaigns that stand out as being the most successful. Three designed to initially get a user to visit CoolTShirts and four that retarget users to come back and make a purchase.

Successful Initial Campaigns		Successful Retargeting Campaigns	
Campaign	% of website users who found CoolTShirts from this campaign	Campaign	% of users who made a purchase as a result of this campaign
Interview with CoolTShirts Founder	31%	Weekly Newsletter	32%
Getting to Know CoolTShirts	31%	Retargeting Ad	31%
Ten Crazy CoolTShirts Facts	29%	Retargeting Campaign	15%
Total	91%	Paid Search	14%
		Total	92%
		Total (not including Retargeting Campaign and Paid Search)	63%

5 Campaigns to re-invest

1.

Interview with CoolTShirts

2.

Getting to Know CoolTShirts

3.

Ten Crazy CoolTShirts Facts

4.

Weekly Newsletter

5.

Retargeting Ad

In an ideal world we would maintain all 7 campaigns however, if we were to stop 2 it should be the Retargeting Campaign and Paid Search.

Successful campaigns that bring visitors to the website should remain as that's the first and most important step in finding paying customers. Each campaign listed in the Successful Initial Campaigns column are highly effective at attracting new users to CoolTShirts. Together they account for 91% of the users to the website.

Although all four campaigns in the Successful Retargeting Campaigns column are effective, by removing Retargeting Campaign and Paid Search you are still ensuring that the campaigns that account for the majority of successful purchases remain; Weekly Newsletter and Retargeting Ad. Together the two campaigns account for 63% of purchases.