



Capstone: Attribution

Learn SQL from Scratch

Naveed Javaid

30th July 2018

Background & Introduction

CoolTShirts, an innovative apparel shop, is running a series of marketing campaigns. In this capstone, I will be helping them answer questions around the following areas:

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?

- The success of ad campaigns run by CoolTShirts in driving traffic to their website can be measured using UTM parameters. These parameters capture when and how a user finds the site (using unique links containing these parameters in their ads)
- In the database *page_visits*, which records the clicks of these unique links:
 - *utm_source* identifies which touchpoint / channel sent the traffic (e.g. google, email, or facebook)
 - *utm_campaign* identifies the specific ad or email blast (e.g. retargeting-ad or weekly-newsletter)
- CoolTShirts used 8 campaigns and 6 sources in total

Unique Campaigns	Unique Sources
8	6

SQL Code Used

```
SELECT COUNT(DISTINCT utm_campaign) AS 'Unique  
Campaigns',  
COUNT(DISTINCT utm_source) AS 'Unique Sources'  
FROM page_visits;
```

1.2 Which source is used for each campaign?

- I have also identified which source is used for each campaign, as per the table on the right
- The sources used by CoolTShirts for their campaigns are; the New York Times, email, BuzzFeed, Facebook, medium and Google

SQL Code Used

```
SELECT DISTINCT utm_campaign AS 'Campaign',  
utm_source AS 'Source'  
FROM page_visits;
```

Campaign	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

1.3 What pages are on the CoolTShirts website?

- The CoolTShirts website consists of four pages; the landing, shopping cart, checkout and purchase pages

SQL Code Used

```
SELECT DISTINCT page_name AS 'Page Name'  
FROM page_visits;
```

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?

- When measuring the effectiveness of ad campaigns, we can look at it in two ways:
 - By looking at the user's *first touch* (the campaign / source of their first encounter with CoolTShirts). This is a good way of knowing how visitors initially discover the site
 - Or by looking at the user's *last touch* (the campaign / source of their final / most recent encounter with CoolTShirts). This is a good way of knowing how visitors are drawn back to the site, especially for making a final purchase
- In the table below, I have established how many 'first touches' to the website each campaign is responsible for – the interview with the founder of CoolTShirts led to the most first touches

Campaign	Source	# First Touches
interview-with-cool-tshirts-founder	medium	622
getting-to-know-cool-tshirts	nytimes	612
ten-crazy-cool-tshirts-facts	buzzfeed	576
cool-tshirts-search	google	169

SQL Code Used

```
WITH first_touch AS (  
    SELECT user_id,  
           MIN(timestamp) AS first_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
ft_attrib AS (  
    SELECT ft.user_id,  
           ft.first_touch_at,  
           pv.utm_source,  
           pv.utm_campaign  
    FROM first_touch AS ft  
    JOIN page_visits AS pv  
      ON ft.user_id = pv.user_id  
      AND ft.first_touch_at = pv.timestamp  
)  
SELECT ft_attrib.utm_campaign AS 'Campaign',  
       ft_attrib.utm_source AS 'Source',  
       COUNT(*) AS '# First Touches'  
FROM ft_attrib  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```


2.2 How many last touches is each campaign responsible for?

- In addition to looking at first touches, I have also established the number of 'last touches' to the website each campaign is responsible for. Note the last touch here does not necessarily lead to a purchase, which I explore later on
- The list of campaigns here is complete as it includes those that are targeted at users who have previously interacted with the website (including the weekly newsletter, Facebook retargeting ad, email retargeting campaign, and paid search)
- The weekly newsletter is the campaign which has led to the highest number of last touches, at nearly 450 – with the Facebook retargeting ad close behind

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

SQL Code Used

```
WITH last_touch AS (  
    SELECT user_id,  
           MAX(timestamp) AS last_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
lt_attrib 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 lt_attrib.utm_campaign AS 'Campaign',  
       lt_attrib.utm_source AS 'Source',  
       COUNT(*) AS '# Last Touches'  
FROM lt_attrib  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

2.3 How many visitors make a purchase?

- The total number of visitors who make a purchase on the CoolTShirts site is 361

Visitors Who Make A Purchase
361

SQL Code Used

```
SELECT COUNT(DISTINCT user_id) AS '# Visitors Who Make 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?

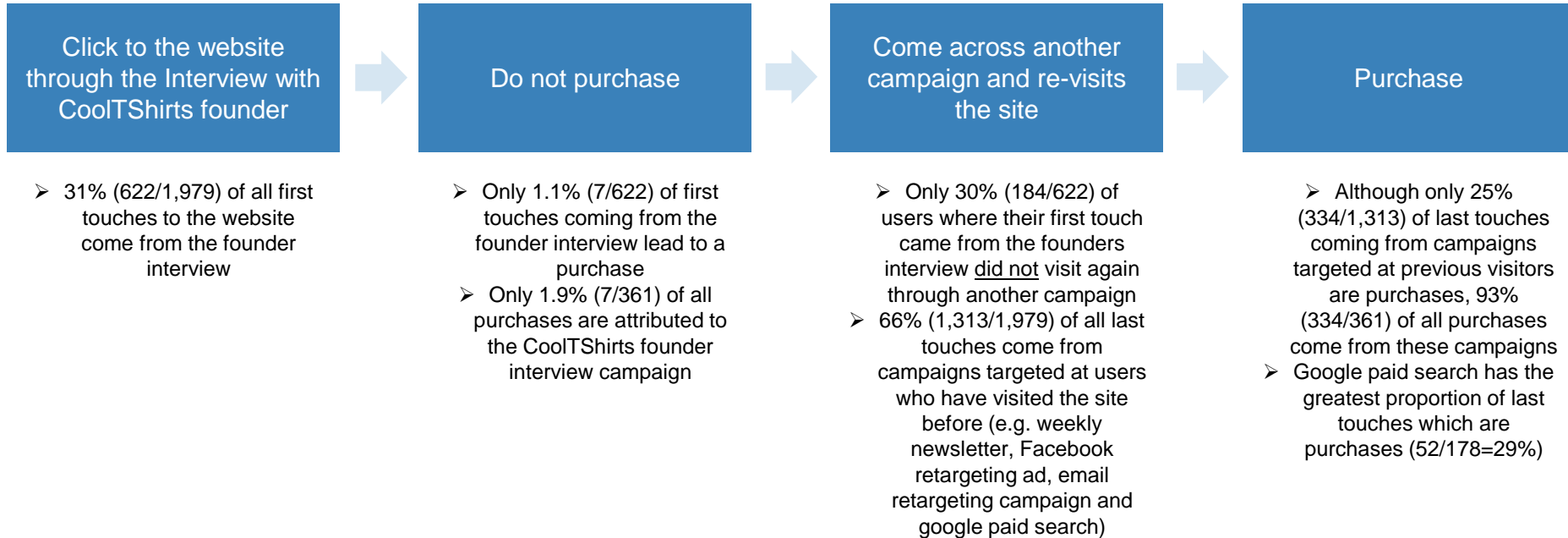
- I have also determined the effectiveness of campaigns in leading to last touches to CoolTShirts' site which are on the purchase page; and thus mean a sale has been achieved
- This is done simply by adding a WHERE clause for page_name = '4 - purchase' to the code establishing how many last touches is each campaign responsible for (slide 9)
- Of the 361 purchases made on the site, 334 (93%) come from campaigns targeted at users who have previously interacted with the website
- Google paid search has the greatest proportion of last touches which are purchases (52/178 = 29%)

Campaign	Source	# Purchase Last Touches
weekly-newsletter	email	115
retargeting-ad	facebook	113
retargeting-campaign	email	54
paid-search	google	52
getting-to-know-cool-tshirts	nytimes	9
ten-crazy-cool-tshirts-facts	buzzfeed	9
interview-with-cool-tshirts-founder	medium	7
cool-tshirts-search	google	2

SQL Code Used

```
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_attrib AS (  
    SELECT lt.user_id,  
           lt.last_touch_at,  
           pv.utm_source,  
           pv.utm_campaign,  
           pv.page_name  
    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 lt_attrib.utm_campaign AS 'Campaign',  
       lt_attrib.utm_source AS 'Source',  
       COUNT(*) AS '# Purchase Last Touches'  
FROM lt_attrib  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

2.5 What is the typical user journey?



3. Optimise the campaign budget

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

- The five campaigns CoolTShirts should reinvest in are:
 1. Interview with the CoolTShirts co-founder – this is the strongest driver of traffic to the site, and feeds the top of the purchase funnel. 31% of all first touches to the website came from the founder interview, and although it did not directly lead to a high number of purchases, it is driving awareness of the brand and creates opportunities for purchase conversion further down the line with the more effective campaigns. It should be a focus for the business to produce content similar to this, and release through the same channels
 2. Weekly Newsletter (email), 3. Re-targeting ad (Facebook) and 4. Re-targeting campaign (email) – 66% of all last touches come from campaigns targeted at users who have visited the site before, which include these three campaigns (as well as google paid search). In addition, 93% of all purchases come from these campaigns. These are proving effective tools to both continually engage with, and convert, previous visitors to the site
 5. Google paid search – in addition to the reasoning outlined above, google paid search has the greatest proportion of last touches which are purchases, at 29%, proving this to be a particularly effective campaign at converting users to purchase
- To give a complete view on the return on investment of each of the campaigns, an idea of the costs for each campaign would be required, as well as the basket size / value of the purchases coming from the different campaigns