



# Capstone Attribution Queries

Learn SQL from Scratch

Caroline Larisch

16<sup>th</sup> October 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?

CoolTShirts uses:

- 8 campaigns
- 6 sources

Difference between utm\_campaign & utm\_source

- utm\_campaigns lists all the campaigns (advertisements or email blasts) that CoolTShirts launched to promote their online shop
- utm\_source shows on which website & campaign link the online shopper clicked to be forwarded to the online shop
- See how they are linked in the below table

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

```
SELECT  
COUNT(DISTINCT utm_campaign)  
FROM page_visits;
```

```
SELECT COUNT(DISTINCT utm_source)  
FROM page_visits;
```

```
SELECT DISTINCT UTM_campaign, utm_source  
FROM page_visits;
```

## 1.2 What pages are on their website?

CoolTShirts website has four pages:

- 1 - landing page
- 2 - shopping cart
- 3 - checkout
- 4 - purchase

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

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

## **2. What is the user journey?**

## 2.1 How many first touches is each campaign responsible for?

The first touches for CoolTShirts are generated by 4 campaigns:

- 1 – interview with cool t-shirts founder
- 2 – getting to know cool tshirts
- 3 – ten crazy cool t-shirts facts
- 4 – cool tshirts search
- The total number of all first touches is 1,979

ft_attr.utm_source	ft_attr.utm_campaign	COUNT(*)
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

```
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,  
       ft_attr.utm_campaign,  
       COUNT(*)  
FROM ft_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

## 2.2 How many last touches is each campaign responsible for?

The last touches for CoolTShirts are generated by all 8 campaigns:

- Please, see below columns lt\_attr.utm\_campaign & COUNT(\*) from the table below

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
email	weekly-newsletter	447
facebook	retargetting-ad	443
email	retargetting-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

```
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  
  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,  
       lt_attr.utm_campaign,  
       COUNT(*)  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```



## 2.3 How many visitors make a purchase?

The page name – 4 – purchase indicates how many customers shopped at CoolTShirts :

- 361 users have purchased at CoolTShirts

page_name	COUNT(DISTINCT user_id)
4 - purchase	361

```
SELECT page_name,  
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?

The last touches on the purchase page for CoolTShirts are generated by all 8 campaigns:

- Please, see below columns lt\_attr.utm\_campaign & COUNT(\*) from the table below

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
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

```
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  
    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,  
       lt_attr.utm_campaign,  
       COUNT(*)  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

## 2.5 What is the typical user journey?

Customers get in touch with CoolTShirts website through the following campaigns on the following channels (in brackets)

- 1 – interview with cool t-shirts founder (medium)
- 2 – getting to know cool tshirts (nytimes)
- 3 – ten crazy cool t-shirts facts (buzzfeed)
- 4 – cool tshirts search (google)

The most successful campaigns leading to completed purchases are

- 1 – weekly newsletter (email) – 23%
- 2 – retargeting-ad (facebook) – 22%

These are the remaining campaigns generating purchases:

- 3 – retargeting campaign (email)
- 4 – getting to know cool t-shirts (nytimes)
- 5 – ten crazy cool tshirts facts (buzzfeed)
- 6 – interview with cool t-shirts founder (medium)
- 7 – paid-search (google)
- 8 – cool tshirts search (google)

We can see that only 4 campaigns lead to first touches with the website, however when it comes to finalized purchases other campaigns than for the first touch play a role. Actually the top score campaigns responsible for first touch perform really low when it comes to generating purchases.

The purchases are essentially reached most successfully through a weekly reminder newsletter and and a retargeting advertisement on facebook.

### **3. Optimize the campaign budget**

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

Reinvest in the following campaigns

- 1 – weekly newsletter (email) – 23%
- 2 – retargeting-ad (facebook) – 22%
- 3 - Getting to know cool t-shirts(nytimes)
- 4 – Ten crazy cool tshirts facts (buzzfeed)
- 5 – Interview with cool-tshirtsfounder (nytimes)

We should invest in these campaigns since they generate the most purchases, however it is important to notice that especially the last two campaigns also score very high on generating the first interaction with the website, hence creating awareness for the brand which is essential for completing purchases based on 'reminder' newsletters or advertisement like here (retargeting – ad).

I decided against putting the email reminder campaign on the list even though it causes over 12% purchases since the other reminder type campaigns reach individually almost double the amount of first touches for the website.

The interview with cool t-shirts founder performs poorly compared to other campaigns when it comes to purchases, yet it generates the highest amount of first touch which is important for brand awareness.