



# Capstone: Attribution

Learn SQL from Scratch

Kerry Abass

14<sup>th</sup> July 2018

# Table of Contents

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

# **1. Get familiar with CoolTShirts**

# 1.1 Campaigns and Sources

CoolTShirts receives visitors to its website from 6 unique sources and 8 different campaigns. Utm\_source identifies which external websites sent the traffic to the website and utm\_campaign describes the marketing campaign designed to drive traffic.

Most campaigns targeted traffic from different sources, with the exception of the 'paid search' campaign and 'cool-tshirts-search' which both drove visitors from google, and 'weekly-newsletter' and 'retargeting-campaign' which were both through emails.

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

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

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

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

## 1.2 CoolTShirts pages

CoolTShirts has 4 different page names.

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

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

## **2. CoolTShirts user journey**

## 2.1 First touches

The table below shows how many first touches each campaign is responsible for. 'interview-with-cool-tshirts-founder' drives the most first touches and 4 campaigns do not drive any first touch traffic at all.

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

## 2.2 Last touches

The table below shows how many last touches each campaign is responsible for. 'weekly-newsletter' drives the most last touches and 'cool-tshirts-search' drives the least. All campaigns drive last touch traffic.

The supporting .sql document demonstrates an alternative query for calculating Last Touch.

utm_campaign	last_touch_count
weekly-newsletter	447
retargetting-ad	443
retargetting-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 1)  
SELECT utm_campaign,  
       COUNT (last_touch_at) as last_touch_count  
FROM last_touch lt  
JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
    AND lt.last_touch_at = pv.timestamp  
GROUP BY 1  
ORDER BY 2 DESC;
```



## 2.3 Purchases

361 visitors to the CoolTShirts website made a purchase.

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

## 2.4 Last touches on the purchase page

The table below shows how many last touches on the purchase page each campaign is responsible for. All campaigns are responsible for at least some last touches on the purchase page.

utm_campaign	last_touch_count_purchased
weekly-newsletter	115
retargetting-ad	113
retargetting-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_purchased AS (  
  SELECT user_id,  
  
         MAX(timestamp) as last_touch_at  
  FROM page_visits  
  WHERE page_name = '4 - purchase'  
  GROUP BY 1)  
SELECT utm_campaign,  
       COUNT (last_touch_at) as last_touch_count_purchased  
FROM last_touch_purchased lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
GROUP BY 1  
ORDER BY 2 DESC;
```

## 2.5 Typical User Journey

The tables below shows the typical user journey is to initially visit the site via a referral from a website called 'medium' targeted by the 'interview-with-cool-tshirts-founder' campaign. The last touchpoint is driven by an email from the 'weekly-newsletter' campaign, which also leads to the most purchases.

The supporting evidence can be found on the following page.

utm_campaign	Most_common_first_touch
interview-with-cool-tshirts-founder	622

utm_campaign	Most_common_last_touch
weekly-newsletter	447

utm_campaign	Most_last_touch_on_purchase_page
weekly-newsletter	115

## 2.5 Typical User Journey

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

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

## 2.5 Typical User Journey

```
WITH last_touch_purchased AS (  
    SELECT user_id,  
           MAX(timestamp) as last_touch_at  
    FROM page_visits  
    WHERE page_name = '4 - purchase'  
    GROUP BY 1)  
SELECT utm_campaign,  
       COUNT (last_touch_at) as most_last_touch_on_purchase_page  
FROM last_touch_purchased lt  
JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
    AND lt.last_touch_at = pv.timestamp  
GROUP BY 1  
ORDER BY 2 DESC  
LIMIT 1;
```

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

## 3.1 Optimise the campaign budget

CoolTShirts can re-invest in 5 campaigns. I recommend they re-invest in:

1. **weekly-newsletter**, as it generated the greatest number of last touchpoints and last touches on the purchase page
2. **retargeting-ad**, as it generated the second greatest number of last touchpoints and last touches on the purchase page
3. **retargeting-campaign**, as it generated the third greatest number of last touchpoints last touches on the purchase page
4. **interview-with-cool-tshirts-founder**, as it generated the greatest number of first time visits to the website
5. **getting-to-know-cool-tshirts**, as it generated the second greatest number of first time visits to the website