



# Capstone Project: First- and Last-Touch Attribution with CoolTShirts.com

Learn SQL from Scratch

Catherine Heywood

July 2018

# Table of Contents

- Get familiar with the company
  - How many campaigns and sources does CoolTShirts use and how are they related? Be sure to explain the difference between `utm_campaign` and `utm_source`.
  - What pages are on their website?
- What is the user journey?
  - How many first touches is each campaign responsible for?
  - How many last touches is each campaign responsible for?
  - How many visitors make a purchase?
  - How many last touches on the purchase page is each campaign responsible for?
  - What is the typical user journey?
- Optimize the campaign budget
  - CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

# **1. Get Familiar with CoolTShirts**

# 1.1 Get Familiar with CoolTShirts

## How many campaigns and sources does CoolTShirts use and how are they related?

- CoolTshirts had **6 total sources** (utm\_source). CoolTshirts' sources identify which touchpoint sent the traffic, such as Google, email, or Facebook.
- CoolTshirts has **8 total campaigns** (utm\_campaign). These campaigns identify the specific ad or email blast that is used to see where site's traffic is coming from.

```
----- Count Total Sources -----  
SELECT COUNT(DISTINCT utm_source) AS 'Total_Source'  
FROM page_visits;
```

```
----- Count Total Campaigns -----  
SELECT COUNT(DISTINCT utm_campaign) AS 'Total_Campaign'  
FROM page_visits;
```

```
----- How They are Related -----  
SELECT DISTINCT utm_source AS 'Sources',  
               utm_campaign AS 'Campaigns'  
FROM page_visits;
```

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

## 1.2 Get Familiar with CoolTShirts continued...

### What pages are on the Cool Tshirts website?

There are **4 total pages** on the CoolTshirts website:

1. Landing page
2. Shopping cart
3. Checkout page
4. Purchase page

Page Names
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

----- Distinct Page Names -----

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

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

## 2.1 What is the user journey?

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

- CoolTShirts customers' first touch was divided into **four sources** and **four campaigns**.
- The campaign 'Interview with CoolTShirts Founder', for the source, Medium, had the most first touches of all the campaigns (n=622), followed by the campaign, 'Getting to Know CoolTShirts', (NY Times, n=612)

Source	Campaign	First Touch 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

```
----- First Touch -----  
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 'First Touch Count'  
FROM ft_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

## 2.2 What is the user journey? (continued)

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

- CoolTShirts customers' last touch was divided into **six sources** and **eight campaigns**.
- The campaign, 'Weekly Newsletter', (source: email) had the most last touches of all the campaigns (n=447), followed by the campaign, 'Retargetting Campaign', (source: Facebook, n=443)

source	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

```
----- Last Touch -----  
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 count  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```



## 2.3 What is the user journey? (continued)

### How many visitors make a purchase?

- Out of the 2000 people that went to the landing page, 95% (1900) of those customers went to the shopping cart page.
- 76% (1431) of those 2000 people that went to the landing page continued to the checkout page.
- **Only 18% of the original 2000 customers (n=361) of those that went to the landing page went to the purchase page.**

page	count
1 – landing page	2000
2 – shopping cart	1900
3- checkout	1431
4 - purchase	361

---- Number of visitors that made a purchase ----

```
SELECT page_name as page,  
       COUNT(*) as count  
FROM page_visits  
WHERE page_name = '4 - purchase'  
GROUP BY page;
```

---- Number of visitors per page ----

```
SELECT page_name as page,  
       COUNT(*) as count  
FROM page_visits  
GROUP BY page;
```

## 2.4 What is the user journey? (continued)

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

- The weekly newsletter (from email) and the retargeting ad (from Facebook) are the campaigns that are responsible for the most last touches on the purchase page.
- The “10 Crazy CoolTShirts Facts” (from BuzzFeed), “Getting To Know CoolTShirts” (from the New York Times), the interview with the CoolTShirts founder (from Medium), and a CoolTShirts search (Google) are responsible for 7.5% of last touches on the purchase page combined.

source	campaign	count
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

----Last touch on purchase page for each 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 count  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

## 2.5 What is the user journey? (continued)

### What is the typical user journey?

To get a visual representation of the “typical user journey,” I ordered the page\_visits table by user\_id (limited to 75) to examine a cross section of users’ shopping experiences.

Based on the table, nearly all of the customers go from the landing page to the shopping cart, which typically are using the same campaign.

There is both a time-gap and most often, a campaign change, when the customer gets to the checkout page. This could be due to many reasons: price comparison between other websites, style comparisons between other sites, or forgetting about shopping cart due to another type of distraction.

Most customers don’t end up purchasing a product. However, based on the snapshot table, those who do end up buying a t-shirt do so shortly after arriving to the checkout page (checkout page and purchase page share the same campaign targeting most of the time).

---- Snapshot of typical user journey ----

```
SELECT *  
FROM page_visits  
ORDER BY user_id ASC  
LIMIT 75;
```

page	count
1 – landing page	2000
2 – shopping cart	1900
3- checkout	1431
4 - purchase	361

# **3. Optimizing the Campaign Budget**

# 3.1 The 5 Campaigns that CoolTShirts Should Re-Invest In

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

If only 5 campaigns can be re-invested in, I believe that the following campaigns should be reinvested in:

### 1. Weekly newsletter

Rationale: The weekly newsletter boasts the highest share of last touches in general and on the purchase page, which is converting potential customers into actual customers. Since this is a proven successful campaign, boosting this with additional investments will likely increase total purchases.

### 2. Retargeting ad

Rationale: The retargeting ad has almost the same amount of last touches in general and on the purchase page as the weekly newsletter, so the same logic applies to this – keep reinvesting in what is already working.

### 3. Retargeting campaign

Rationale: While there is a considerable drop off in last touch attribution between the weekly newsletter/retargeting ad and the retargeting campaign, the retargeting campaign has shown some potential (third highest last touch) and also benefits from having the same source as the weekly newsletter (email), which could maximize the effectiveness of the re-investment – potentially resulting in greater efficiency and lower costs for distribution to potential customers.

## 3.2 The 5 Campaigns that CoolTShirts Should Re-Invest In (ctd.)

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

If only 5 campaigns can be re-invested in, I believe that the following campaigns should be reinvested in:

#### **4. Getting to Know CoolTShirts**

Rationale: Getting to Know CoolTShirts ranks fourth in last touch attribution (right after the retargeting campaign), so we know it has some degree of sticking power. But more importantly, it has the second highest number of first touches, which means it is excellent at drawing potential customers in, thereby increase revenue upside, which makes it well worth reinvesting in.

#### **5. 10 Crazy CoolTShirts Facts**

Rationale: Similar to the Getting To Know CoolTShirts campaign, 10 Crazy CoolTShirts Facts has shown some success with first and last touch attribution (more so the former than the latter). While this has led to few last touches from the purchase page, most of the value in investing in this campaign comes from its strength in first touch – getting people onto the landing page in the first place is invaluable.

**Thank You!**