

LEARN SQL FROM SCRATCH

ELEANOR SHEERS
JANUARY 27TH, 2019

OPTION 3: FIRST- AND LAST-TOUCH ATTRIBUTION WITH COOLTSHIRTS.COM

GETTING FAMILIAR WITH THE COMPANY

1. How many campaigns and sources does CoolTShirts use and how are they related?
2. What pages are on their website?

WHAT IS THE USER JOURNEY

1. How many first touches is each campaign responsible for?
2. How many last touches is each campaign responsible for?
3. How many visitors make a purchase?
4. How many last touches on the purchase page is each campaign responsible for?
5. What is the typical user journey?

OPTIMIZE THE CAMPAIGN BUDGET

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

1. GETTING FAMILIAR WITH THE COMPANY

1.1 GETTING FAMILIAR WITH THE COMPANY

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

Using `SELECT COUNT (DISTINCT)` and `SELECT DISTINCT` queries, we can find that the CoolTShirts company uses **8 types of Distinct Campaigns** and **6 types of Distinct Sources**. These two variables are related in that each campaign solely comes from 1 source, while each source may have more than 1 campaign. As we can see in the “Campaigns/Sources” table, email and google are the only sources with more than one campaign for CoolTShirts with 2 campaigns each.

- UTM parameters capture information regarding how users discover a certain website. Owners of a website embed special links containing UTM parameters in various forms of sources. Data is added upon a user clicking on a source embedded with a UTM parameter, representing visits attributed to a certain parameters.
- UTM sources (`utm_source`) indicate which type of medium sent the traffic, such as NY Times, Email, Buzfeed, etc.
- UTM campaigns (`utm_campaign`) indicates which marketing campaign sent the traffic, such as a Paid Search, a Cool TShirts Search, the Weekly Newsletter, etc.

```
SELECT COUNT(DISTINCT utm_campaign) AS 'Distinct Campaigns'
FROM page_visits;

SELECT COUNT(DISTINCT utm_source) AS 'Distinct Sources'
FROM page_visits;

SELECT DISTINCT utm_campaign AS 'Campaigns', utm_source AS 'Sources'
FROM page_visits;
```

Distinct Campaigns

8

Distinct Sources

6

Campaigns	Sources
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

1.2 GETTING FAMILIAR WITH THE COMPANY

What pages are on their website?

Ad campaigns such as those used by Cool TShirts use page names (page_names) to indicate the title of the section of the page that was visited by the traffic. To discover the page names for Cool TShirt's campaigns, we should use a **SELECT DISTINCT** query to isolate the distinct page names.

Using this query, we discover **four distinct pages**:

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

```
SELECT DISTINCT page_name AS 'Distinct Pages'  
FROM page_visits;
```

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

2. WHAT IS THE USER JOURNEY

2.1 WHAT IS THE USER JOURNEY

How many first touches is each campaign responsible for?

Using **first touch attribution** using the **minimum time stamp**, we can discover which **UTM Campaign** each customer initially discovered Cool TShirts through. Using this method, we find that the “Interview with Cool TShirts Founder” campaign produced 622 first touches, “Getting to Know Cool TShirts” produced 612 first touches, “Ten Crazy Cool TShirts Facts” produced 576 first touches, and “Cool TShirts Search” produced 169 first touches - which is also represented in the following table:

utm_campaign	COUNT(utm_campaign)
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 user_id)  
SELECT ft.user_id,  
       ft.first_touch_at,  
       pv.utm_source,  
       pv.utm_campaign,  
       COUNT(utm_campaign)  
FROM first_touch ft  
JOIN page_visits pv  
    ON ft.user_id = pv.user_id  
   AND ft.first_touch_at = pv.timestamp  
GROUP BY utm_campaign  
ORDER BY 5 DESC;
```

2.2 WHAT IS THE USER JOURNEY

How many last touches is each campaign responsible for?

Using **last touch attribution** using the **maximum timestamp**, we can discover which **UTM Campaign** each potential customer last visited Cool TShirts through. Using this method, we find that following campaigns as (utm_campaign) produced the corresponding amount of customer's last touches (COUNT(utm_campaign)):

utm_campaign	COUNT(utm_campaign)
weekly-newsletter	447
retargeting-ad	443
retargeting-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 user_id)  
SELECT lt.user_id,  
       lt.last_touch_at,  
       pv.utm_source,  
       pv.utm_campaign,  
       COUNT(utm_campaign)  
FROM last_touch lt  
JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
    AND lt.last_touch_at = pv.timestamp  
GROUP BY utm_campaign  
ORDER BY 5 DESC;
```


2.3 WHAT IS THE USER JOURNEY

How many visitors make a purchase?

While several of the potential customers who follow a campaign are represented by a “last touch”, not all of these visitors actually make a purchase. We use **WHERE** to identify those users who made it all the way to the “**purchase**” page, indicating that 361 visitors made a purchase.

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

COUNT (DISTINCT user_id)

361

2.4 WHAT IS THE USER JOURNEY

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

To break down how many **last touches on the purchase page** the campaign is responsible for, we are able to combine the previous **WHERE** query with the **last_touch** query. By combining these queries, we are able to filter out all of the last touches that did not make it to the purchase page. We can find the amounts of last touches on the purchase page that is attributed to each campaign in the chart below:

utm_campaign	COUNT(utm_campaign)
weekly-newsletter	115
retargeting-ad	113
retargeting-campaign	54
getting-to-know-cool-tshirts	52
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	9
paid-search	7
cool-tshirts-search	2

```
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)  
SELECT lt.user_id,  
       lt.last_touch_at,  
       pv.utm_source,  
       pv.utm_campaign,  
       COUNT(utm_campaign)  
FROM last_touch lt  
JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
    AND lt.last_touch_at = pv.timestamp  
GROUP BY utm_campaign  
ORDER BY 5 DESC;
```

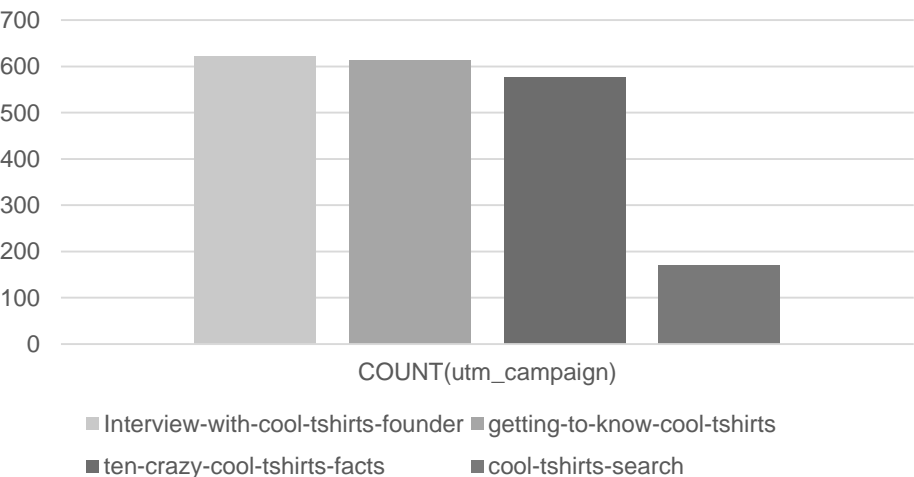
2.5 WHAT IS THE USER JOURNEY

What is the typical user journey?

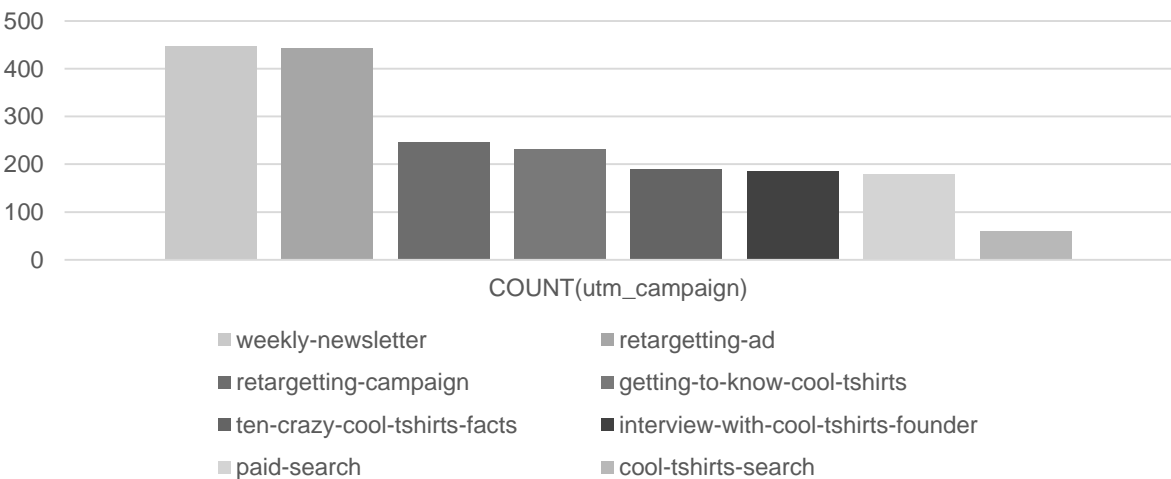
FIRST TOUCHES

91% of potential customers are initially drawn in by article campaigns from the New York Times (Getting to Know Cool TShirts), BuzzFeed (Ten Crazy Cool TShirts Facts), and Medium (Interview with Cool TShirts Founder), as seen in the “First Touches” chart below.

First Touches



Last Touches



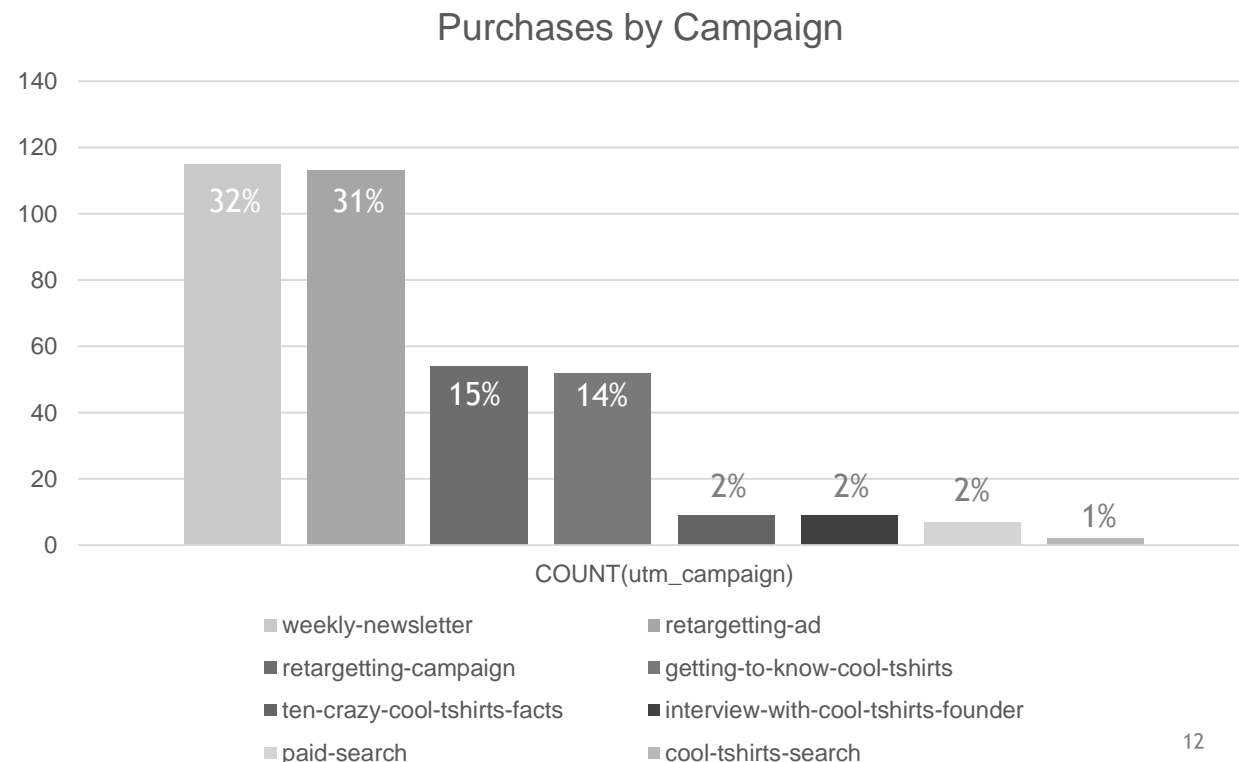
LAST TOUCHES

While first touch predominantly can be attributed to articles, last touches derive from a **more diverse variety of sources**, as seen in the chart above. In the majority, 35% of potential customers returned for their last touch via email campaigns (Weekly Newsletter, Retargetting Campaign). Social media was the second largest campaign contributing 22% of last touches through the Retargetting Ad. Articles contribute to a small but decent portion at 22% (Getting To Know Cool TShirts, Ten Crazy Cool TShirts Facts). Finally, Google produces the smallest portion of last touches at 11% (Paid Search, Cool TShirts Search).

2.5 WHAT IS THE USER JOURNEY

What is the typical user journey?

Across the 361 users who made a purchase, we can trace that a large amount (**115 purchasers or 47%**) did so via an **email campaign** (Weekly Newsletter, Retargetting Campaign). Social media comes in next at 31% representing the 113 purchasers coming from Facebook. News articles comprised 18%, and Google provided 3%.



3. OPTIMIZE THE CAMPAIGN BUDGET

3.1 GETTING FAMILIAR WITH THE COMPANY

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

Getting To Know
CoolTShirts

Interview With Cool
TShirts Founder

Weekly Newsletter

Ten Crazy Cool TShirts
Facts

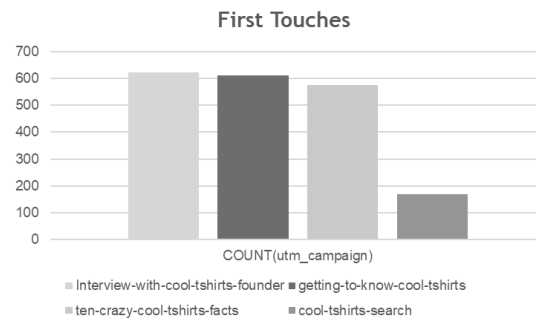
Retargeting Ad

3.2 GETTING FAMILIAR WITH THE COMPANY

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

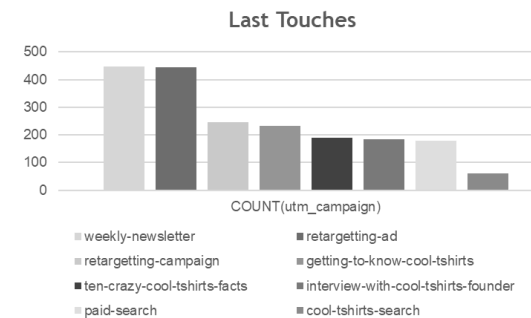
FIRST TOUCHES

Looking at the data from the First Touches, we find that three campaigns produced the predominant amount of First Touches: [Getting to Know Cool TShirts](#), [Ten Crazy Cool TShirts Facts](#), and [Interview With Cool TShirts Funder](#). There may also be additional benefit towards investing in these campaigns other than purchases, such as for effective general brand awareness.



LAST TOUCHES

For last touches, the campaigns attributing to the most last touches are [Weekly Newsletter](#) and [Retargeting Ad](#), both pulling more than 22% of last touches each. Together, they contribute to 50% of last touches, while they represent only 2/8 of the campaigns.



PURCHASES

Looking at the data from the purchases, of the 361 purchasers 63% came from the same two campaigns that reaped the highest amount of last touches, [Weekly Newsletter](#) and [Retargeting Ad](#). Therefore in addition to the three campaigns attributing to First Touches, these latter two campaigns yield significant return for Last Touches and thus Purchases.

