



First- and Last-Touch Attribution with CoolTShirts.com

Learn SQL from Scratch

Joanie Gentile

January 23, 2019

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 Get familiar with the company

1. 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`

CoolTshirts uses 8 different campaigns and 6 different sources. A source is the place (website, article, email, search engine) where a prospect interacts with a campaign, usually driving that prospect back to the company's website. The campaign is the specific message or type of message the company is running to try and engage with prospects. A company can run the same campaign across a variety of sources and/or different campaigns on a specific set of sources.

```
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;
```

Query Results	
COUNT(DISTINCT utm_campaign)	8
COUNT(DISTINCT utm_source)	6
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

1.2 Get familiar with the company

2. What pages are on their website?

The pages on their website can be found by writing a SELECT DISTINCT query for page-name. The pages are:

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

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

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

2. What is the user journey

2.1 What is the user journey?

1. How many first touches is each campaign responsible for?

This can be done by finding the timestamps with the MIN date for each user and then grouping that MIN timestamp by the utm_campaign. In order to have all of this data in one table, we must join the first_touch temporary table with the page_visits table.

```
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;
```

Query Results

user_id	first_touch_at	utm_source	utm_campaign	COUNT(utm_campaign)
99990	2018-01-13 23:30:09	medium	interview-with-cool-tshirts-founder	622
99933	2018-01-25 00:04:39	nytimes	getting-to-know-cool-tshirts	612
99765	2018-01-04 05:59:46	buzzfeed	ten-crazy-cool-tshirts-facts	576
99684	2018-01-13 13:20:49	google	cool-tshirts-search	169

2.2 What is the user journey?

2. How many last touches is each campaign responsible for?

This query is very similar to the previous first touch query except that we switch out the MIN date for the MAX date. We also need to switch out all the first_touch values for last_touch.

```
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;
```

Query Results

user_id	last_touch_at	utm_source	utm_campaign	COUNT(utm_campaign)
99933	2018-01-26 06:18:39	email	weekly-newsletter	447
99928	2018-01-24 05:26:09	facebook	retargetting-ad	443
99990	2018-01-16 11:35:09	email	retargetting-campaign	245
99589	2018-01-15 04:55:43	nytimes	getting-to-know-cool-tshirts	232
99765	2018-01-04 05:59:47	buzzfeed	ten-crazy-cool-tshirts-facts	190
99838	2018-01-02 07:40:34	medium	interview-with-cool-tshirts-founder	184
98840	2018-01-10 04:58:48	google	paid-search	178
99344	2018-01-18 21:36:32	google	cool-tshirts-search	60

2.3 What is the user journey?

3. How many visitors make a purchase?

To find the number of visitors who made a purchase, all we need to do is count the distinct number of user who hit the page "4 – purchase".

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

Query Results	
	COUNT(DISTINCT user_id)
	361

2.4 What is the user journey?

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

To find the number of last touches on the purchase page, we need to combine the previous two queries. We use the last touch query from question 2.2, and add in the WHERE clause from question 2.3.

Query Results

user_id	last_touch_at	utm_source	utm_campaign	COUNT(utm_campaign)
99933	2018-01-26 06:18:39	email	weekly-newsletter	115
99897	2018-01-06 09:41:19	facebook	retargetting-ad	113
99285	2018-01-24 09:00:58	email	retargetting-campaign	54
94567	2018-01-19 16:37:58	google	paid-search	52
92172	2018-01-16 15:15:29	nytimes	getting-to-know-cool-tshirts	9
98651	2018-01-15 04:17:36	buzzfeed	ten-crazy-cool-tshirts-facts	9
83547	2018-01-10 18:20:21	medium	interview-with-cool-tshirts-founder	7
95650	2018-01-18 00:25:00	google	cool-tshirts-search	2

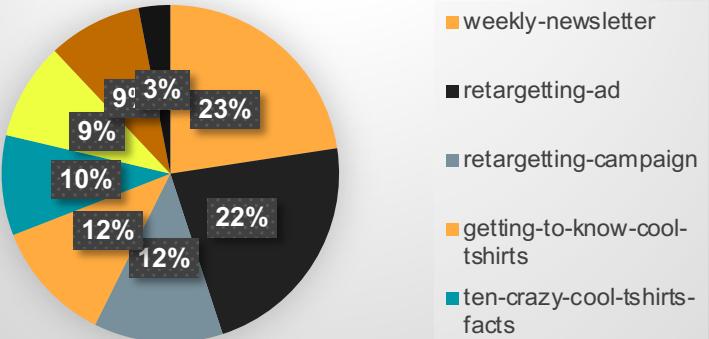
```
WITH last_touch AS (
  SELECT user_id,
    MAX(timestamp) as last_touch_at
  FROM page_visits
  WHERE page_name = '4 - purchase'
  GROUP BY user_id)
SELECT lt.user_id,
  lt.last_touch_at,
  pv.utm_source,
  pv.utm_campaign,
  COUNT(utm_campaign)
FROM last_touch lt
JOIN page_visits pv
  ON lt.user_id = 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?

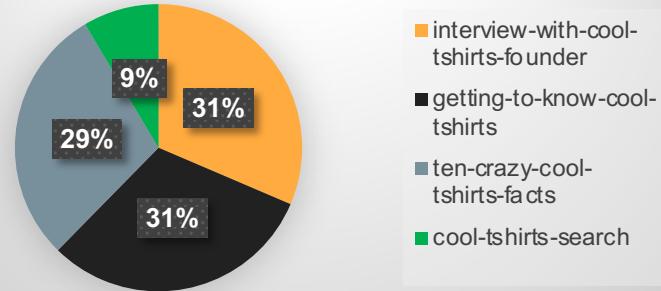
5. What is the typical user journey?

Most prospects find the CoolTShirts site for the first time via an article on Medium, NY Times, or Buzzfeed. However, these prospects almost never purchase right away. Instead, they are drawn back to the site for by an email newsletter or retargeting campaign or via a Facebook ad. These three campaigns are also the most likely to be the campaigns that push the prospect to a purchase.

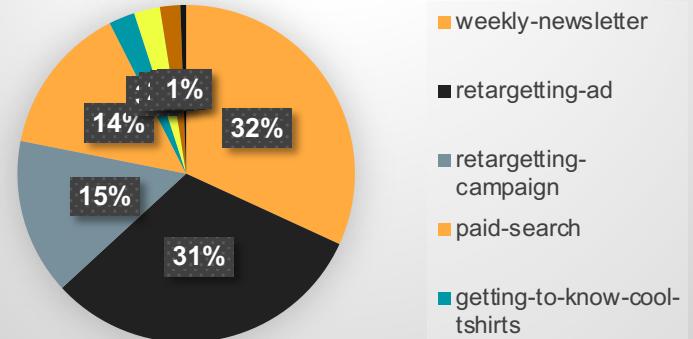
Last Touch by Campaign



First Touch by Campaign



Purchase by Campaign



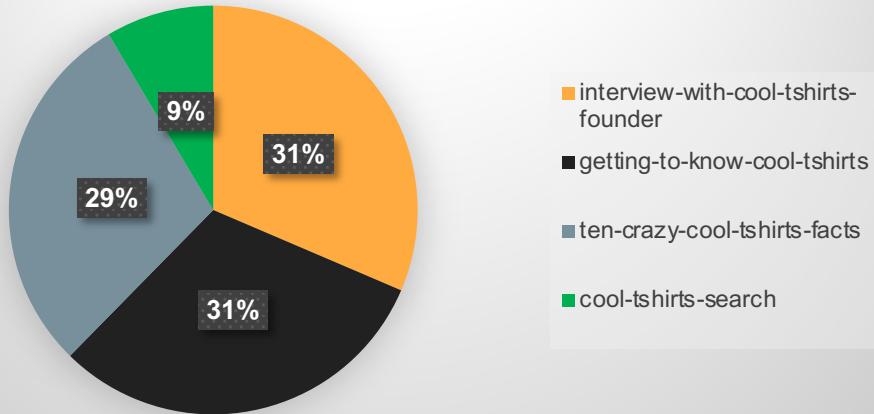
3. Optimize the campaign budget

3.1 Optimize the campaign budget

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

If CoolTShirts can only re-invest in 5 campaigns, they need to pick a mix of campaigns that get new users to their site and push returning prospects to purchase. For this reason, I would suggest they re-invest in the 3 top first touch campaigns (interview-with-cool-tshirts-founder, getting-to-know-cool-tshirts, and ten-crazy-cool-tshirts-facts) and the 2 top purchase campaigns (weekly-newsletter and retargetting-ad).

First Touch by Campaign



Purchase by Campaign

