



Attribution Queries

Learn SQL from Scratch

Kaitlyn Williams

March 20, 2019

Table of Contents

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

Get familiar with CoolTshirts

How many campaigns does CoolTshirts use?

The CoolTshirts has 8 distinct campaigns.

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

Query Results	
count(distinct utm_campaign)	
8	

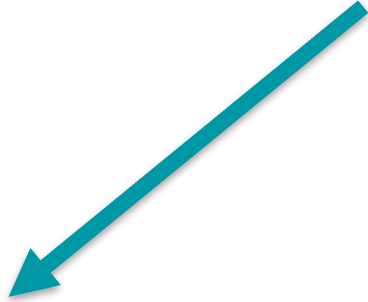
How many sources does CoolTshirts use?

The CoolTshirts has 6 distinct sources.

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

Query Results
count(distinct utm_source)
6

Campaigns and Sources



The `utm_campaign` column lists the campaign that attributed to the first or last touch. The campaign is the way that a person was attracted to their website. The 8 distinct various campaigns are; weekly-newsletter, retargeting-ad, retargeting-campaign, getting-to-know-cool-tshirts, ten-crazy-cool-tshirts-facts, interview-with-cool-tshirts-founder, paid-search, and cool-tshirts-search.



The `utm_source` column lists the source type of the campaign that attributed to the first or last touch. The source is where the campaign was found. The 6 distinct various sources are; email, facebook, google, nytimes, buzzfeed, and medium.

Campaigns and Sources

The `utm_campaign` and `utm_source` are related because the `utm_source` is where a specific `utm_campaign` is found that attributes to the first or last touch. For example, the retargeting-ad campaign was found on the source facebook. The `utm_source` and `utm_campaign` are intertwined, for every `utm_source` there is a `utm_campaign`.

Which source is used for each campaign?

The source nytimes is used for the getting-to-know-cool-tshirts campaign. The source email is used for the weekly-newsletter campaign and the retargeting-campaign campaign. The source buzzfeed is used for the ten-crazy-cool-tshirts-facts campaign. The source facebook is used for the retargeting-ad campaign. The source medium is used for the interview-with-cool-tshirts-founder campaign. The source google is used for the paid-search campaign and the cool-tshirts-search campaign.

```
1 SELECT DISTINCT utm_campaign,  
2     utm_source  
3 FROM page_visits;
```

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

What pages are on the CoolTshirts website?

The CoolTshirts website has 4 distinct pages; 1-landing_page, 2-shopping_cart, 3-checkout, and 4-purchase.

```
1  SELECT DISTINCT page_name
2  FROM page_visits;
```

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

What is the user journey?

How many first touches is each campaign responsible for?

```
1 WITH first_touch AS (  
2     SELECT user_id,  
3         MIN(timestamp) as first_touch_at  
4     FROM page_visits  
5     GROUP BY user_id)  
6 SELECT ft.user_id,  
7     ft.first_touch_at,  
8     pv.utm_source,  
9     pv.utm_campaign,  
10    COUNT(utm_campaign)  
11 FROM first_touch ft  
12 JOIN page_visits pv  
13     ON ft.user_id = pv.user_id  
14     AND ft.first_touch_at = pv.timestamp  
15 GROUP BY utm_campaign  
16 ORDER BY 5 DESC;
```

Of the first touches the campaign interview-with-cool-tshirts-founder contributes 622, the campaign getting-to-know-cool-tshirts contributes 612, the campaign ten-crazy-cool-tshirts-facts contributes 576, the campaign cool-tshirts-search contributes 169, and the other four campaigns don't contribute to any of the first touches.

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

How many last touches is each campaign responsible for?

```
1 WITH last_touch as(
2   SELECT user_id,
3     max(timestamp) as last_touch_at
4   FROM page_visits
5   GROUP BY user_id)
6 SELECT lt.user_id,
7    lt.last_touch_at,
8    pv.utm_source,
9    pv.utm_campaign,
10   COUNT(utm_campaign)
11 FROM last_touch lt
12 JOIN page_visits pv
13   ON lt.user_id = pv.user_id
14   AND lt.last_touch_at = pv.timestamp
15 GROUP BY utm_campaign
16 ORDER BY 5 DESC;
```

Of the last touches the campaign weekly-newsletter contributes 447, the campaign retargetting-ad contributes 443, the retargetting-campaign contributes 245, the campaign getting-to-know-cool-tshirts contributes 232, the campaign ten-crazy-cool-tshirts-facts contributes 190, the interview-with-cool-tshirts-founder contributes 184, the campaign paid-search contributes 178, and the campaign cool-tshirts-search contributes 60.

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

How many visitors make a purchase?

361 of the visitors who visited the CoolTshirts page made a purchase.

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

Query Results
count(distinct user_id)
361

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

```
1 WITH last_touch as(
2   SELECT user_id,
3     max(timestamp) as last_touch_at
4   FROM page_visits
5   WHERE page_name = '4 - purchase'
6   GROUP BY user_id)
7 SELECT lt.user_id,
8   lt.last_touch_at,
9   pv.utm_source,
10  pv.utm_campaign,
11  COUNT(utm_campaign)
12 FROM last_touch lt
13 JOIN page_visits pv
14   ON lt.user_id = pv.user_id
15   AND lt.last_touch_at = pv.timestamp
16 GROUP BY utm_campaign|
17 ORDER BY 5 DESC;
```

Of the last touches on the purchase page the campaign weekly-newsletter contributes 115, the campaign retargeting-ad contributes 113, the campaign retargeting-campaign contributes 54, the campaign paid-search contributes 52, the campaign getting-to-know-cool-tshirts contributes 9, the campaign ten-crazy-cool-tshirts-facts contributes 9, the campaign interview-with-cool-tshirts-founder contributes 7, and the campaign cool-tshirts-search contributes 2.

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	retargeting-ad	113	
99285	2018-01-24 09:00:58	email	retargeting-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	

What is the typical user journey?

Based on the queries done for the first and last touch attributes, the typical user discovers the CoolTshirts page through three main campaigns; the interview-with-cool-tshirts-founder, the getting-to-know-cool-tshirts and ten-crazy-cool-tshirts-facts. These three campaigns use the sources; medium, nytimes, and buzzfeed, respectively. The typical user last accesses the CoolTshirts page through several campaigns but the two most popular are the weekly-newsletter and the retargeting-ad. These two campaigns use the source email and facebook, respectively. These two campaigns also contribute to the most last touches on the purchase page. The typical user will visit the CoolTshirts page but will most likely not make a purchase since the amount of last touches on the purchase page is much lower than the total amount of last touches on any of the pages.

Optimize the campaign budget

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

The 5 campaigns that the CoolTshirts should re-invest in are; weekly-newsletter, retargetting-ad, interview-with-cool-tshirts-founder, getting-to-know-cool-tshirts, and ten-crazy-cool-tshirts-facts.

They should re-invest in the weekly-newsletter campaign because that campaign brought in the most last touches (447) and the most last touches on the purchase page (115). This campaign contributes to a lot of purchases on the page. They should re-invest in the retargetting-ad because that campaign brought in the second most last touches (443) and the second most last touches on the purchase page (113). This campaign also contributes to a lot of purchases on the page. They should re-invest in the interview-with-cool-tshirts-founder because that campaign brought in the most first touches (622) and several last touches (184). This campaign will continue to draw people to their page. They should re-invest in the getting-to-know-cool-tshirts campaign because it brought in the second most first touches (612) and several last touches (232). This campaign will also continue to draw several people to their page. The last campaign they should re-invest in is ten-crazy-cool-tshirts-facts because this campaign brought in the third most first touches (576) and several last touches (190). This campaign will also continue to draw people to their page.

Re-investing in the campaigns that brought in a lot of last touches on the purchase page will continue to lead to more purchases on the CoolTshirts page. Re-investing in the campaigns that brought in a lot of first touches will help to continue to draw people to the CoolTshirts page and in turn lead to more people making purchases on their page.