code|cademy

First- and Last-Touch Attribution project

Learn SQL from Scratch

YANNI ZENG

07-16-2018
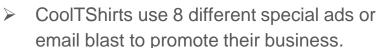
Table of Contents

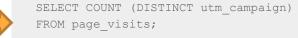# 1. Get familiar with the company

Campaigns AND Sources

- How many campaigns CoolTShirts use?
- CoolTShirts use 8 different special ads or email blast to promote their business.

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

- How many sources does CoolTShirts use? there are 6 sites sent the customer to visit the website.

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

How campaigns and sources does CoolTShirts use related ?

The sources is identifies which site sent the traffic.
The campaigns is Identifies the specific ad or email blast.

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

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

What pages are on the CoolTShirts website?

There are 4 page names: landing_page, shopping_cart, check_out, purchase.

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

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

2. What is the user journey?

How many first touches is each campaign responsible for?

Interview-with-Cool-Tshirts-founder is a great campaign for this business, which responsible for 622 times initially discover for the website.

ft_attr.utm_source	ft_attr.utm_campaign	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

```
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 AS 'ft'
  JOIN page_visits AS 'pv'
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
)
SELECT ft_attr.utm_source,
       ft_attr.utm_campaign,
       COUNT(*)
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

How many last touches is each campaign responsible for?

```
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
  FROM last_touch AS 'lt'
  JOIN page_visits AS 'pv'
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
        lt_attr.utm_campaign,
        COUNT(*)
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

t_attr.utm_source	lt_attr.utm_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

How many visitors make a purchase?

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

page_name	COUNT (*)
4 - purchase	361

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

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-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

```
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 AS  'lt'
  JOIN page_visits AS'pv'
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
       lt_attr.utm_campaign,
       COUNT(*)
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

CoolTShirt Website Customer Journey

Search	Shopping	Purchase
How are prospects finding the website? (Responsible campaigns)	The Customers drawn back to the website (Responsible Campaign)	Making the final purchase (Responsible Campaign)
interview-with-cool-tshirts-founder	weekly-newsletter	weekly-newsletter
getting-to-know-cool-tshirts	retargetting-ad	retargetting-ad
ten-crazy-cool-tshirts-facts	retargetting-campaign	retargetting-campaign
cool-tshirts-search	getting-to-know-cool-tshirts	paid-search
	ten-crazy-cool-tshirts-facts	ten-crazy-cool-tshirts-facts
	interview-with-cool-tshirts-founder	getting-to-know-cool-tshirts
	paid-search	interview-with-cool-tshirts-founder
	cool-tshirts-search	cool-tshirts-search

3. Optimize the campaign budget

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

For CoolTShirts, I think company can keeps top two campaigns to increase product and brand's visibility and awareness, those are **interview-with-cool-tshirts-founder** and **getting-to-know-cool-tshirts**. Another two could be most profitable campaigns, these are **weekly-newsletter** and **retargetting-ad**.

For the last one, company could choose either ten-crazy-cool-tshirts-facts or retargeting-campaign. Ten-crazy-cool-tshirts-facts has such high fist-touch-rate, but with the profit-rate only (9/190) 4.74%
On the other hand, the retargetting-campaign has profit-rate is (54/245) 22%,but it has low fist touch rate.
The final choice depends on CoolTShirts business stage:
If CoolTShirts still growing, I would recommend re-invest ten-crazy-cool-tshirts-facts promote it's reputation and increase sales.
If CoolTShirts is a mature company, they could pick most profitable retargetting-campaign.