



CoolTShirts: Attribution Queries

Learn SQL from Scratch

OLENA ARNAUTOVA | 2019.05.06

Table of Contents

1. Get familiar with CoolTShirts.

- 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?

2. 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?

3. Optimize the campaign budget

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

1. GET FAMILIAR WITH CoolTShirts

1. GET FAMILIAR WITH CoolTShirts

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*.

- With UTM parameters, you can tag your links to change the effectiveness of the campaigns and identify the best ways to increase the number of visitors to your website.
- *utm_campaign* is used to identify which campaign the promotion is associated with.
- *utm_source* is used to show which site the visitors are coming from (email, social networks, websites etc.).

```
1  -- Get number of the campaigns
2  SELECT
3    COUNT(DISTINCT utm_campaign) AS 'Number of campaigns'
4  FROM page_visits;
5
6  -- Get number of the sources
7  SELECT
8    COUNT(DISTINCT utm_source) AS 'Number of sources'
9  FROM page_visits;
10
11 -- Relationship between the campaigns and the sources
12 SELECT
13   DISTINCT utm_campaign AS 'Campaigns',
14   utm_source AS 'Sources'
15 FROM page_visits
16 ORDER BY utm_campaign;
17
```

Number of campaigns

8

Number of sources

6

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

1. GET FAMILIAR WITH CoolTShirts

What pages are on the company's website?

- There is a statement DISTINCT to find what pages are on CoolTShirt's website. This statement is used to return unique values in the output and removes all duplicate values.
- The output shows there are 4 kind of pages on the CoolTShirt's website.

```
-- Get names of the website's pages

SELECT DISTINCT page_name AS 'Names of pages'
FROM page_visits;
```

Names of pages
1 – landing_page
2 – shopping_cart
3 – checkout
4 – purchase

2. WHAT IS THE USER JOURNEY?

2. WHAT IS THE USER JOURNEY?

How many first touches is each campaign responsible for?

- User journey shows a user flow through the website.
- Using the time at which the user came to the page (column “timestamp”) and a function MIN() it is possible to set the time of first touches and count number of first touches in each campaign.

Source	Campaign	Number
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

```
-- Count first touches for each campaign
-- Create temporary table first_touch to find all first touches for every user

WITH first_touch AS (
    SELECT user_id, MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id),

-- Create temporary table ft_attr to add extra info (source and campaign) by
-- joining 2 tables: first_touch and page_visits

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
)

-- Count first touches for each campaign

SELECT ft_attr.utm_source AS 'Source',
       ft_attr.utm_campaign AS 'Campaign',
       COUNT(*) AS 'Number'
FROM ft_attr
GROUP BY ft_attr.utm_source, ft_attr.utm_campaign
ORDER BY COUNT(*) DESC;
```

2. WHAT IS THE USER JOURNEY?

How many last touches is each campaign responsible for?

- Using the time at which the user came to the page (column “timestamp”) and a function MAX() it is possible to set the time of last touches and count number of last touches in each campaign.

Source	Campaign	Number
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

```
-- Count last touches for each campaign
-- Create temporary table last_touch to find all last touches for every user

WITH last_touch AS (
  SELECT user_id,
         MAX(timestamp) AS last_touch_at
  FROM page_visits
  GROUP BY user_id),

-- Create temporary table lt_attr to add extra info (source and campaign) by
-- joining 2 tables: last_touch and page_visits

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)

-- Count last touches for each campaign

SELECT lt_attr.utm_source AS 'Source',
       lt_attr.utm_campaign AS 'Campaign',
       COUNT(*) AS 'Number'
FROM lt_attr
GROUP BY lt_attr.utm_source, lt_attr.utm_campaign
ORDER BY COUNT(*) DESC;
```


2. WHAT IS THE USER JOURNEY?

How many visitors make a purchase?

- To count how many users make a purchase simply select all DISTINCT users that visited page "4 – purchase".

```
-- Using COUNT and DISTINCT to count Visitors that made a purchase

SELECT COUNT(DISTINCT user_id) AS 'Visitors that made a purchase'
FROM page_visits
WHERE page_name = '4 - purchase';
```

Visitors that made a purchase
361

2. WHAT IS THE USER JOURNEY?

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

- Using the time at which the user came to the page (column "timestamp"), a function MAX() and simply restrict of visiting only page "4-purchase" by add WHERE clause it is possible to set the time of last touches and count number of last touches on the purchase page for each campaign.

Source	Campaign	Number
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

```
-- Count last touches for the purchase page
-- Create temporary table last_touch to find all last touches for every user

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),

-- Create temporary table lt_attr to add extra info (source and campaign) by
-- joining 2 tables: last_touch and page_visits

lt_attr AS (
  SELECT lt.user_id,
    lt.last_touch_at,
    pv.utm_source,
    pv.utm_campaign
  FROM last_touch lt
  JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp)

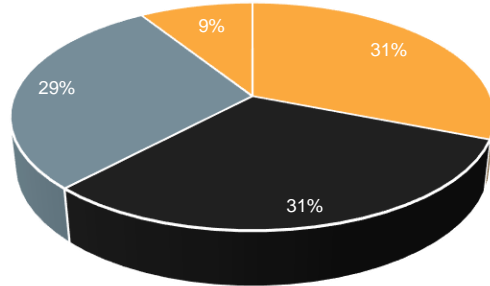
-- Count last touches for the purchase page
SELECT lt_attr.utm_source AS 'Source',
  lt_attr.utm_campaign AS 'Campaign',
  COUNT(*) AS 'Number'
FROM lt_attr
GROUP BY lt_attr.utm_source, lt_attr.utm_campaign
ORDER BY COUNT(*) DESC;
```

2. WHAT IS THE USER JOURNEY?

What is the typical user journey?

- A user journey is a process where a user goes through to achieve a goal with the company.
- According to the charts only 7% users from 91% made a purchase.
- 78% of purchases were made via retargetting adds or campaigns, weekly newsletters.

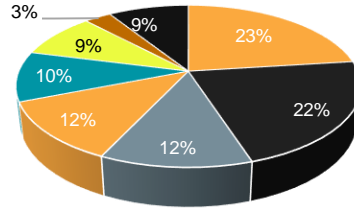
First touches by campaign



91%

- interview-with-cool-tshirts-founder
- getting-to-know-cool-tshirts
- ten-crazy-cool-tshirts-facts
- cool-tshirts-search

Last touches by campaign

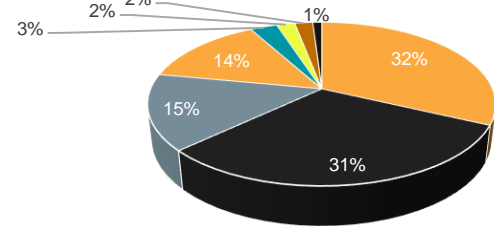


57%

- weekly-newsletter
- retargetting-ad
- retargetting-campaign
- getting-to-know-cool-tshirts
- ten-crazy-cool-tshirts-facts
- interview-with-cool-tshirts-founder
- cool-tshirts-search
- paid-search

31%

Purchases by campaign



78%

- weekly-newsletter
- retargetting-ad
- retargetting-campaign
- paid-search
- ten-crazy-cool-tshirts-facts
- getting-to-know-cool-tshirts
- interview-with-cool-tshirts-founder
- cool-tshirts-search

7%

3. OPTIMIZE THE CAMPAIGN BUDGET

3. OPTIMIZE THE CAMPAIGNE BUDGET

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

- Since three “Story” campaigns have a high percentage of the first touches it should be continued followed by an analysis of how to increase purchases in this group.
- In addition, there is a recommendation to continue also the campaigns with the highest percentage of purchases.

Campaign	1 st touches	1 st touches, %	Last touches	Last touches, %	Purchases	Purchases, %
interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%
getting-to-know-cool-tshirts	612	31%	232	12%	9	2%
ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	3%
cool-tshirts-search	169	9%	60	3%	2	1%
weekly-newsletter	0	0%	447	23%	115	32%
retargetting-ad	0	0%	443	22%	113	31%
retargetting-campaign	0	0%	245	12%	54	15%
paid-search	0	0%	178	9%	52	14%