



Attribution Queries

Learn SQL from Scratch

Jesús Ortega Ochoa

Dec 20, 2018

Example Table of Contents

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

1. Get familiar with CoolTShirts

1.1.1 Get familiar with CoolTShirts

The first part of the project asks for the campaigns and sources that Cool Tshirts uses.

- For the first query we select the total of different campaigns
(column 1. Total_Campaigns)
- The second query shows us the total of different sources.
(column 2. Total_Sources)

Total_Campaigns	Total_Sources
8	6

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

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

1.1.2 Get familiar with CoolTShirts

Continuing with the first part, we get the relation between campaigns and sources .

Total_Campaigns	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

```
SELECT DISTINCT utm_campaign as Campaign, utm_source  
as Source  
FROM page_visits;
```

1.2 Get familiar with CoolTShirts

What pages are on the CoolTshirts website?

- With a simple query using a DISTINCT statement for the page_name column, we get the desired result.

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

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

2. What is the user journey?

2.1 What is the user journey?

How many first touches is each campaign responsible for?

- We create a temporary table named 'first_touch' selecting the MIN(timestamp) for each user (grouping by user_id).
- Then we count the MIN(timestamp) from the 'first_touch' table for each campaign from the 'page_visits' table. In this case we join the two tables on the same user. Later, we group the result by campaign.

Campaign	First_Touches
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 pv.utm_campaign as Campaign,  
       COUNT(ft.first_touch_at) as First_Touches  
FROM first_touch ft  
JOIN page_visits pv  
  ON ft.user_id = pv.user_id  
 AND ft.first_touch_at = pv.timestamp  
GROUP BY 1  
ORDER BY 2 DESC;
```


2.2 What is the user journey?

How many last touches is each campaign responsible for?

- This case is exactly like the it was for the first touches, except for the temporary table, where we select the MAX(timestamp).

Campaign	Last_Touches
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 pv.utm_campaign as Campaign,  
       COUNT(lt.last_touch_at) as Last_Touches  
FROM last_touch lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
GROUP BY 1  
ORDER BY 2 DESC;
```

2.3 What is the user journey?

How many visitors make a purchase?

- In this case, with a simple query we count the different users whose page_name value is '4 – purchase', using a DISTINCT statement.

prchs_visitors
361

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

2.4 What is the user journey?

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

- This case is the same for the last touches adding a WHERE clause in order to get only the results for the 'page_name' column with '4 – purchase' values.

Campaign	prchs_lt
weekly-newsletter	114
retargetting-ad	112
retargetting-campaign	53
paid-search	52
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	7
cool-tshirts-search	2

```
WITH last_touch AS (  
  SELECT user_id,  
         MAX(timestamp) as last_touch_at  
  FROM page_visits  
  GROUP BY user_id)  
SELECT pv.utm_campaign as Campaign,  
       COUNT(lt.last_touch_at) as prchs_lt  
FROM last_touch lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
WHERE page_name = '4 – purchase'  
GROUP BY 1  
ORDER BY 2 DESC;
```

3. Optimize the campaign budget

3. Optimize the campaign budget

- In order to re-invest in 5 campaigns we consider the top 5 of total touches.
- Analyzing the results from the first and last touches CoolTShirts should pick the campaigns listed on the table with the highest total touches.
- Considering the campaigns with the highest first touches that lead to the page, have a great impact on costumers in order to discover the website and lately assure the return of the customer even if is back through another campaign.

Campaigns to Re-Invest	Total Touches
getting-to-know-cool-tshirts	844
interview-with-cool-tshirts-founder	806
ten-crazy-cool-tshirts-facts	766
weekly-newsletter	447
retargeting-ad	443