codecademy

Capstone:Attribution Queries

Simon Deng

Date

Table of Contents

1. Get familiar with CoolTShirts

1.1 How many campaigns & sources does CoolTShirts use?

The first query show that there were 8 distinct campaigns whilst the second query showed that there were 6 distinct sources.

The third query returned the below table, which demonstrates the relationship between each source with each campaign.

Query Results	
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

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

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

(3) SELECT DISTINCT utm_source,utm_campaign
FROM page_visits;
```

1.2 What pages are on the CoolTShirts website?

There are four unique pages on the CoolTShrts website:

- The landing page
- The hopping cart page
- The checkout page;
- And the purchase page

```
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 How many first touches is each campaign responsible for?

After running the query on the right, the campaign that was responsible for the highest number of first touches was interview-with-cool-tshirts-founder, followed by getting-to-know-cool-tshirts, ten-crazy-cool-tshirts-facts and cool-tshirts-search.

utm_campaign	first_touch_count
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,
    COUNT (*) AS first_touch_count
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_source
ORDER BY first_touch_count DESC;
```

2.2 How many last touches is each campaign responsible for?

The campaign that was responsible for the largest number of last_touches was retargetting-campaign, followed by retargetting-ad, cool-tshirts-search, getting-to-know-cool-tshirts, ten-crazy-cool-tshirts-facts and interview-with-cool-tshirts-founder.

utm_campaign	last_touch_count
retargetting-campaign	692
retargetting-ad	443
cool-tshirts-search	238
getting-to-know-cool-tshirts	232
ten-crazy-cool-tshirts-facts	190
interview-with-cool-tshirts-founder	184

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT pv.utm_campaign,
    COUNT (*) AS last_touch_count
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_source
ORDER BY last_touch_count DESC;
```

2.3 How many visitors make a purchase?

A COUNT function was used in this query to count the number of web pages in the table, however the returning result was filtered by the WHERE clause so that only the total number of the ‘purchase’ webpage would be produced. This total number was 358.

Number of purchases made
358

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT COUNT (page_name) AS 'Number of purchases
made'
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';
```

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

The query from 2.3 was modified so that it would return a column for utm_campaigns as well as the number of purchases made. In order for the query to work properly, a GROUP BY statement was used to group together the different utm_campaigns. Lastly, the 'number of purchases made' column was ordered in a descending order so that that it would place the utm_campaign responsible the most number of purchases at the top.

utm_campaign	Number of purchases made
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,
COUNT (page_name) AS 'Number of purchases made'
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

2.5 What is the typical user journey?

Looking at 2.1, we can see the beginning of the user journey. The utm_campaign responsible for the largest number of first touches was interview-with-cool-tshirts-founder, which was responsible for 622 first touches.

Also, using the query on the right, we can map out what the user journey was like on the CoolTShirts website. In the table below we can see that the majority of people that clicked into the landing page eventually navigated to the shopping cart however a significant proportion of users dropped out in the checkout and purchase pages.

Lastly, looking at 2.4, we can see that of those who made it to the purchase page, the utm campaign 'weekly-newsletter' was responsible for the largest proportion of the users.

```
SELECT page_name,
COUNT (page_name) AS 'number of visits'
FROM page_visits
GROUP BY 1;
```

Page_name	Number of visits
1 - landing_page	2000
2 - shopping_cart	1900
3 - checkout	1431
4 - purchase	361

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

If CoolTShirts is able to re-invest in 5 campaigns, the campaigns that I would recommend investing in the following campaigns:

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

These are the top three campaigns which drove first touches to the CoolTShirts website. Together they made up 91% of total first touches. Focusing on campaigns that will help to maximize first touches will give the website the exposure it needs to sell its products.

The last two campaigns I would recommend re-investing money into are:

- retargetting-campaign
- retargetting-ad

Whilst there were 6 different campaigns that contributed to last touches for the CoolShirts website, these two campaigns alone were responsible for 57% of last touches. Focusing efforts on investing into these two campaigns will help to maximise CoolTShirts's chances in selling its products.