codecademy

Attribution Queries

Analyze Real Data with SQL

Leshego Ledwaba

4 August 2020

Table of Contents

CoolTShirts sells shirts of all kinds, as long as they are T-shaped and cool. Recently, CTS started a few marketing campaigns to increase website visits and purchases. Questions below seek to determine performance of campaigns.

1. How many campaigns and sources does CoolTShirts use? Which source is used for each campaign?

Three queries are needed: One for the number of distinct campaigns. One for the number of distinct sources. One to find how they are related.

- CoolTShirts has 4 campaigns
- CoolTShirts has 4 source websites for traffic
- How the campaigns and sources are related: each campaign is dedicated to one source website

Number of campaign	Number of sources
4	4

utm_campaign	utm_source
cool-tshirts-search	google
getting-to-know-cool-tshirts	nytimes
interview-with-cool-tshirts-founder	medium
ten-crazy-cool-tshirts-facts	buzzfeed

```
WITH first_touch AS (
SELECT user_id,
MIN(timestamp) as first_touch_at
FROM page_visits
GROUP BY user_id)
SELECT
COUNT (DISTINCT pv.utm_campaign) AS 'Number of
campaign',
COUNT (DISTINCT pv.utm_source) AS 'Number of sources'
FROM first_touch ft
JOIN page_visits pv
ON ft.user_id = pv.user_id
AND ft.first_touch_at = pv.timestamp;

WITH first_touch AS (
SELECT user_id,
MIN(timestamp) as first_touch_at
FROM page_visits
GROUP BY user_id)
SELECT
pv.utm_campaign,
pv.utm_source
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;
```

2. What pages are on the CoolTShirts website?

CoolTShirts has 4 pages on website

Number of pages on CoolTshirts website
4

```
SELECT COUNT (DISTINCT page_name) AS 'Number of pages on CoolTshirts website'
FROM page_visits;
```

3. How many first touches is each campaign responsible for?

Each of the campaigns is responsible for the following:

- Cool-tshirts-search campaign is responsible for 169 attributions
- Getting-to-know-cool-tshirts campaign is responsible for 612 attributions
- Interview-with-cool-tshirts-founder campaign is responsible for 622 attributions
- Ten-crazy-cool-tshirts-facts campaign is responsible for 576 attributions

utm_campaign	Number of first touches
cool-tshirts-search	169
getting-to-know-cool-tshirts	612
interview-with-cool-tshirts-founder	622
ten-crazy-cool-tshirts-facts	576

```
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 (ft.first_touch_at) AS 'Number of 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;
```

4. How many last touches is each campaign responsible for?

Each of the campaigns is responsible for the following:

- **Cool-tshirts-search** campaign is responsible for **60** attributions
- **Getting-to-know-cool-tshirts** campaign is responsible for **232** attributions
- **Interview-with-cool-tshirts-founder** campaign is responsible for **184** attributions
- **Paid-search** campaign is responsible for **178** attributions
- **Re-targetting ad** is responsible for **443** attributions
- **Re-targetting campaign** is responsible for **245** attributions
- **Ten-crazy-cool-tshirts-facts** campaign is responsible for **190** attributions
- **Weekly-newsletter** campaign is responsible for **447** attributions

utm_campaign	Number of last touches
cool-tshirts-search	60
getting-to-know-cool-tshirts	232
interview-with-cool-tshirts-founder	184
paid-search	178
retargetting-ad	443
retargetting-campaign	245
ten-crazy-cool-tshirts-facts	190
weekly-newsletter	447

```
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 (lt.last_touch_at) AS 'Number of 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;
```

5. How many visitors make a purchase?

361 visitors made a purchase

page_name	Number of users purchase
1 - landing_page	1979
2 - shopping_cart	1881
3 - checkout	1431
4 - purchase	361

```
SELECT page_name,
COUNT(DISTINCT user_id) AS 'Number of users purchase'
FROM page_visits
GROUP BY 1;
```

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

See in the table below the number of last touches that each campaign can be attributed to.

page_name	utm_campaign	Number of last touches
4 - purchase	cool-tshirts-search	2
4 - purchase	getting-to-know-cool-tshirts	9
4 - purchase	interview-with-cool-tshirts-founder	7
4 - purchase	paid-search	52
4 - purchase	retargetting-ad	113
4 - purchase	retargetting-campaign	54
4 - purchase	ten-crazy-cool-tshirts-facts	9
4 - purchase	weekly-newsletter	115

```
-- You can put your query here
WITH last_touch AS (
SELECT user_id,
MAX(timestamp) as last_touch_at
FROM page_visits
GROUP BY user_id)
SELECT
page_name,
pv.utm_campaign,
COUNT (lt.last_touch_at) AS 'Number of 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 2;
```

7. CoolTShirts can re-invest in 5 campaigns. Given your findings in the project, which should they pick and why?

CoolTShirts should invest in campaigns that have resulted in high volumes of last touch which occured on the purchase page, because these maximise sales. Furthermore, since retargettting ads and campaigns are the next dominant drivers for purchase, and they require first touch data to work, CoolTShirts should also include campaigns with the highest first touch volumes - say top two first touch campaigns.

Campaigns that meet these conditions (in order of priority) are:

- Weekly newsletter - with highest number of last touches (purchase), i.e. 115
- Retargetting ad - with second highest number of last touches (purchase), i.e. 113
- Retargetting campaign - with third highest number of last touches (purchase), i.e. 54
- Interview with cooltshirts founder – with the highest number of of first touch .e. 622 (necessary for successful retargetting),
- Getting to know cooltshirts – with the second highest number of first touch i.e. 612 (necessary for successful retargetting)

page_name	utm_campaign	Number of last touches (purchase)	No. of first touches (all)
4 - purchase	cool-tshirts-search	2	169
4 - purchase	getting-to-know-cool-tshirts	9	612
4 - purchase	interview-with-cool-tshirts-founder	7	622
4 - purchase	paid-search	52	--
4 - purchase	retargetting-ad	113	--
4 - purchase	retargetting-campaign	54	--
4 - purchase	ten-crazy-cool-tshirts-facts	9	576
4 - purchase	weekly-newsletter	115	--

```
-- You can put your query here
WITH last_touch AS (
SELECT user_id,
MAX(timestamp) as last_touch_at
FROM page_visits
GROUP BY user_id)
SELECT
page_name,
pv.utm_campaign,
COUNT (lt.last_touch_at) AS 'Number of 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 2;
```