codecademy

Capstone- Attribution

Learn SQL from Scratch
Daniel Vasquez
07/15/2008

Example Table of Contents

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

1. Get Familiar with Cool T-shirts!

1.1 Get Familiar with the Company, pt 1

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.

- Cool T-shirts uses 8 different campaigns to increase it's websites visits and purchases. Those campaigns are; Getting-to-know-coo-tshirt, weekly-newsletter, ten-crazy-cool-tshirts-facts, retargeting-campaign, retargeting-ad, interview-with-cool-tshirts-founder, paid-search, cool-tshirts-search.
- Similarly, Cool T-shirt also uses 6 different sources to attain attraction to its websites for purchases. Those sources are; NY times, email, buzzfeed, facebook, medium and Google.
- These two factors within Cool T-shirts are related because each campaign is led via different sources. These source are what we can call marketing touchpoints. A campaign can be attached to a source/marketing touchpoint where the customer first encounters Cool T-shirts, in what we would say is a touch attribution to the overall process of visiting Cool T-shirts and ultimately buying from.. See table in the following slide
- (See table on following slide for queries and table related to the questions above)
- A utm campaign is a code attached to the specific urls of campaigns or rather strategies or course of actions formulated to achieve a desired group of people to purchase from Cool T-shirts. Its designed to peak the interest of a specific group. In contrast, a url source is also a code used in a url, but the url is the place where they attain the visitor or customer. For example in this case, the url was found through facebook or google aka the source.

utm_campaign	utm_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 COUNT(DISTINCT(utm_campaign))
FROM page_visits;


SELECT COUNT(DISTINCT(utm_source))

FROM page_visits;


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

1.2 Getting Familiar with the company, pt 2

- What pages are on their website?

There are 4 pages in the company's website;

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 How many first touches is each campaign responsible for?

Each campaign is responsible for a different number of first touches, or first point at which the customer accessed Cool T-shirts (touch attribution). For example (as seen below), Cool T-shirts-search is responsible for 169, Getting-to-know-cool-t-shirts is responsible for 612, Interview-with-cool-shirts-founder is responsible for 622 and ten-crazy-cool-tshirt-facts is responsible for 576. The rest are 0 first touches.

user_id	first_touch_at	utm_source	utm_campaign	COUNT(*)
99684	2018-01-13 13:20:49	google	cool-tshirts-search	169
99933	2018-01-25 00:04:39	nytimes	getting-to-know-cool-tshirts	612
99990	2018-01-13 23:30:09	medium	interview-with-cool-tshirts-founder	622
99765	2018-01-04 05:59:46	buzzfeed	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 ft.user_id,
    ft.first_touch_at,
    pv.utm_source,
        pv.utm_campaign, 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 pv.utm_campaign;
```

2.2 How many last touches is each campaign responsible for?

Each campaign is responsible for a different number of first touches, or first point at which the customer accessed Cool T-shirts. For example (as seen in the table on the next slide), cool-tshirts-search is responsible for 60 last touches, getting-to-know-cool-tshirt is responsible for 232 last touches, interview-with-cool-tshirts-founders is responsible for 184, paid search is responsible for 178, retargeting-ad is responsible for 443, retargetting-campaign is responsible for 245, ten-crazy-cool-tshirt-facts is responsible for 190 last touches, and not but not least weekly newsletter is responsible for 447 last touches.

```
WITH last_touch AS (
  SELECT user_id,
    MAX(timestamp) as last_touch_at
FROM page_visits
GROUP BY user_id)
SELECT lt.user_id, lt.last_touch_at,
pv.utm_source, pv.utm_campaign,
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 pv.campaign;
```

2.2 How many last touches is each campaign responsible for?

user_id	last_touch_at	utm_source	utm_campaign	COUNT(pv.utm_campaign)
99344	2018-01-18 21:36:32	google	cool-tshirts-search	60
99589	2018-01-15 04:55:43	nytimes	getting-to-know-cool-tshirts	232
99838	2018-01-02 07:40:34	medium	interview-with-cool-tshirts-founder	184
98840	2018-01-10 04:58:48	google	paid-search	178
99928	2018-01-24 05:26:09	facebook	retargetting-ad	443
99990	2018-01-16 11:35:09	email	retargetting-campaign	245
99765	2018-01-04 05:59:47	buzzfeed	ten-crazy-cool-tshirts-facts	190
99933	2018-01-26 06:18:39	email	weekly-newsletter	447

2.3 How many visitors make a purchase?

At a every point of the Cool T-shirts purchase process, visitors visit different pages of the website starting with the Landing Page and ending with an actual purchase. However, only some make it all the way to the purchase page and complete their purchase. Using a query, we can determine that 361 visitors made a purchase as detailed in the table below.

```
SELECT COUNT(DISTINCT(user_id)),
   MAX(timestamp) as 'last_touch_at',
page_name
FROM page_visits
GROUP BY page_name;
```

COUNT(DISTINCT(user_id))	last_touch_at	page_name
1979	2018-01-30 23:34:11	1 - landing_page
1881	2018-01-31 03:49:25	2 - shopping_cart
1431	2018-02-04 10:57:47	3 - checkout
361	2018-02-04 11:09:47	4 - purchase

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

Using the last touch query, grouped by user_id and taking the count for utm_campaigns (see query to the right), we can determine the count of last touches for each campaign. As detailed in the table below you can see that each campaign is responsible for 9, 169, 113, 54, 7 and 9 last touches respective to the campaigns listed below.

```
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)
  SELECT lt.user_id, lt.last_touch_at,
pv.utm_source, pv.utm_campaign,
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 pv.utm_campaign;
```

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

user_id	last_touch_at	utm_source	utm_campaign	COUNT(*)
95650	2018-01-18 00:25:00	google	cool-tshirts-search	2
92172	2018-01-16 15:15:29	nytimes	getting-to-know-cool-tshirts	9
83547	2018-01-10 18:20:21	medium	interview-with-cool-tshirts-founder	7
94567	2018-01-19 16:37:58	google	paid-search	52
99897	2018-01-06 09:41:19	facebook	retargetting-ad	113
99285	2018-01-24 09:00:58	email	retargetting-campaign	54
98651	2018-01-15 04:17:36	buzzfeed	ten-crazy-cool-tshirts-facts	9
99933	2018-01-26 06:18:39	email	weekly-newsletter	115

2.5 What is the typical user journey?

It appears from the data and the tables retrieved from the queries, that a majority of the user journey initiates through the Inteview-with-the-cool-tshirts founder and Getting-to-know-cool-tshirts Campaigns via sources Medium and NY Times respectively. As they continue through each of the website pages making their way to the purchase page, more visitors begin to withdraw. However, the weekly newsletter and the retargeting ad is able to draw back in a significant number of visitors , 447 and 443 respectively via email and facebook, leading them all the way to the purchase page.

3. Optimize the Campaign Budget

3.1 Example Cool TShirts can re-invest in 5 campaigns. Which should they pick and why?

Since Cool Tshirts want to consider both first touches and last touches in their user journey to determine which campaigns initially attract the most amount of users (hoping a number of the largest sample makes it to purchase) and to determine which campaigns push customers to actually complete the purchase, the following campaigns should be reinvested in:

1. Interview with cool tshirts founder - this brought in the highest number of initial visitors to their landing page (622)
2. Getting to know cool tshirts - This brough in the second highest number of intial visitors (612)
3. Weekly newsletter - this campaign brought in the largest number of final visitors or last touch customers to the company's web page (447). Additionally, it also led to highest number of purchases from the last touch count (115).
4. Retargeting ad - this brought the second highest number of final visitors to the company's website (443) and it led to the second highest number of final purchases from last touches (113).
5. Retargeting Campaign - Lastly this campaign brought the third largest number of final visitors to the company's website (245) and it also led to the third highest number of final purchases from last touches (54).

*All information can be see through queries and tables in previous slides.