codecademy

First- and Last-Touch Attribution project

Learn SQL from Scratch
Elena Repyakh
Jan 15, 2019 – Mar 12, 2019

Table of Contents

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

1. Get familiar with the company

1.1 How many campaigns and sources does CoolTShirts use and how are they related?

The query results show us that:
- CoolTShirts uses 8 campaigns within 6 sources.
- Such sources as "email" and "google" are used twice as often as others (please refer to the table below).

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

```
-- campaigns query

SELECT COUNT(DISTINCT utm_campaign) AS count_campaign
FROM page_visits;

-- sources query

SELECT COUNT(DISTINCT utm_source) AS count_source
FROM page_visits;

-- campaigns and sources relation query

SELECT DISTINCT utm_campaign,
          utm_source
FROM page_visits
ORDER BY 2;
```

1.2 What pages are on CoolTShirts website?

To track the visitor's journey, CoolTShirts uses 4 pages on their website.
Page names are:

- 1 – landing_page;
- 2 – shopping_cart;
- 3 – checkout;
- 4 – purchase.

```
-- pages query

SELECT DISTINCT page_name
FROM page_visits;
```

2. What is the user journey?

2.1 How many first touches is each campaign responsible for?

First-touch attribution query results tell us that:
- visitors reacted just to 4 out of 8 campaigns to visit the website for the first time;
- majority of visitors initially discovered the CoolTShirts website with the help of "interview-with-cool-tshirts-founder" campaign via "medium" source with a number of 622 first touches,
- least popular campaign "cool-tshirts-search" was launched on "google" source and had just 169 first touches.

The table below shows us how many first touches each campaign is responsible for:

campaign	source	first_touch_count
interview-with-cool-tshirts-founder	medium	622
getting-to-know-cool-tshirts	nytimes	612
ten-crazy-cool-tshirts-facts	buzzfeed	576
cool-tshirts-search	google	169

```
-- first touch query

WITH first_touch AS (
    SELECT user_id,
           MIN(timestamp) AS first_touch_at
    FROM page_visits
    GROUP BY user_id),
ft_attr AS (
    SELECT ft.user_id,
           ft.first_touch_at,
           pv.utm_source,
           pv.utm_campaign
    FROM first_touch AS ft
    JOIN page_visits AS pv
        ON ft.user_id = pv.user_id
        AND ft.first_touch_at = pv.timestamp)
    SELECT ft_attr.utm_campaign AS campaign,
           ft_attr.utm_source AS source,
           COUNT(*) AS first_touch_count
    FROM ft_attr
    GROUP BY 1
    ORDER BY 3 DESC;
```

2.2 How many last touches is each campaign responsible for?

With the help of the results from the last-touch attribution query we can see:
- Majority of visitors were drawn back to the website by "weekly-newsletter" campaign via "email" source with 447 last touches.
- "Cool-tshirts-search" campaign via "google" source achieved to get just 60 returns.

The table below shows us how many last touches each campaign is responsible for:

campaign	source	last_touch_count
weekly-newsletter	email	447
retargetting-ad	facebook	443
retargetting-campaign	email	245
getting-to-know-cool-tshirts	nytimes	232
ten-crazy-cool-tshirts-facts	buzzfeed	190
interview-with-cool-tshirts-founder	medium	184
paid-search	google	178
cool-tshirts-search	google	60

```
--last touch query

WITH last_touch AS (
    SELECT user_id,
           MAX(timestamp) AS last_touch_at
    FROM page_visits
    GROUP BY user_id),
lt_attr AS (
    SELECT lt.user_id,
           lt.last_touch_at,
           pv.utm_source,
           pv.utm_campaign
    FROM last_touch AS lt
    JOIN page_visits AS pv
        ON lt.user_id = pv.user_id
        AND lt.last_touch_at = pv.timestamp)
    SELECT lt_attr.utm_campaign AS campaign,
           lt_attr.utm_source AS source,
           COUNT(*) AS last_touch_count
    FROM lt_attr
    GROUP BY 1
    ORDER BY 3 DESC;
```

2.2 How many last touches is each campaign responsible for?

Moreover, we can notice that "email" source was the most successful in achieving the goal to return customers back with the result of 692 last touches with the help of two campaigns: "weekly-newsletter" and "retargetting-campaign".

Also, "google" source managed to initiate 238 returns with the help of two campaigns: "paid-search" and "cool-tshirts-search". Thus, we can make a conclusion that as a source, "google" is more successful then "nytimes", "buzzfeed" or "medium" ones in terms of returning customers back to the CoolTShirts website.

campaign	source	last_touch_count
weekly-newsletter	email	447
retargetting-ad	facebook	443
retargetting-campaign	email	245
getting-to-know-cool-tshirts	nytimes	232
ten-crazy-cool-tshirts-facts	buzzfeed	190
interview-with-cool-tshirts-founder	medium	184
paid-search	google	178
cool-tshirts-search	google	60

2.3 How many visitors make a purchase?

Page named "4 - purchase" catches the fact of a final purchase on CoolTShirt.com.
With the help of the query we counted a number of visitors performed a purchase which is 361.

Please refer to the next slide to find out how many purchases each campaign had triggered.

visitors_purchased	page_name
361	4 - purchase

```
-- visitors purchased query

SELECT COUNT(DISTINCT user_id) AS
          visitors_purchased,
          page_name
FROM page_visits
WHERE page_name = '4 - purchase';
```

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

From the data in the table below, it is apparent that "weekly-newsletter" campaign (launched via email) and "retargeting-ad" campaign (launched via facebook) are responsible for the majority of sales with 115 and 113 purchases accordingly.
In oppose, such campaigns as "cool-tshirt-search" (via google) and "interview-with-cool-tshirts-founder" (via medium) did not show significant results with just 2 and 7 purchases accordingly.

purchase_count	utm_campaign	utm_source
115	weekly-newsletter	email
113	retargetting-ad	facebook
54	retargetting-campaign	email
52	paid-search	google
9	getting-to-know-cool-tshirts	nytimes
9	ten-crazy-cool-tshirts-facts	buzzfeed
7	interview-with-cool-tshirts-founder	medium
2	cool-tshirts-search	google

```
--last touches on the purchase page query

WITH last_touch AS(
    SELECT user_id,
           MAX(timestamp),
           page_name,
           utm_campaign,
           utm_source
    FROM page_visits
    WHERE page_name = '4 - purchase'
    GROUP BY user_id)
SELECT COUNT(*) AS purchase_count,
       last_touch.utm_campaign,
       last_touch.utm_source
FROM last_touch
GROUP BY 2
ORDER BY 1 DESC;
```

2.5 What is the typical user journey?

The typical user journey is shown on the picture below:

	WebSite Pages tracking user's journey					
	1 - landing page	1 - landing page	2-shopping cart	2-shopping cart	3-checkout	4-purchase
User's journey step	First visit	Browsing a selection of products	Adding products to shopping cart	Viewing a shopping cart	Taking decision to purchase	Making a purchase

In addition, based on previous results from queries:

1. The majority of users are attracted to the website ('first visit' step) by "interview-with-cool-tshirts-founder", "getting-to-know-cool-tshirts", "ten-crazy-cool-tshirts-facts" campaigns.

2. After the 'first visit' step, users can leave a webpage or keep the journey following further steps shown at the picture above. Also, users can leave a page at any of the further steps. However, a certain number of users can go through all of the further steps and make a final purchase.

3. Users who left the website without final purchase could be returned to the site. The most successful campaigns to return customers were: "weekly-newsletter", "retargetting-ad" and "retargetting-campaign".

4. "Weekly-newsletter" and "retargetting-ad" campaigns caused the major number of final purchases.

3. Optimize the campaign budget

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

Before making any conclusions about the future strategy for the CoolTShirts we need to calculate purchase rate as well as last touch rate.

Purchase rate will show us percent of final purchases among "returned visitors" per campaign. The last touch rate shows us percent of 'first touches' which led to the 'last touches' per campaign.

The table below represents combined results from SQL queries and math calculations.

Based on data in the table, there are:

1) Four companies with the purpose to initiate users to visit the website: "interview-with-cool-tshirts-founder", "getting-to-know-cool-tshirts", "ten-crazy-cool-tshirts-facts", "cool-tshirts-search".

2) Four companies with the purpose to return visitors to the website: weekly-newsletter, retargetting-ad, retargetting-campaign, paid-search.

campaign	source	first_touch_count	last_touch_count	last touch rate (%)	purchase_count	purchase rate (%)
interview-with-cool-tshirts-founder	medium	**622**	**184**	184/662*100=**27.80**	**7**	7/184*100=**3.80**
getting-to-know-cool-tshirts	nytimes	**612**	**232**	232/612*100=**37.91**	**9**	9/232*100=**3.88**
ten-crazy-cool-tshirts-facts	buzzfeed	**576**	**190**	190/576*100=**32.99**	**9**	9/190*100=**4.74**
cool-tshirts-search	google	**169**	**60**	60/169*100=**35.51**	**2**	60/169*100=**3.33**
weekly-newsletter	email	-	**447**	-	**115**	115/447*100=**25.73**
retargetting-ad	facebook	-	**443**	-	**113**	113/443*100=**25.51**
retargetting-campaign	email	-	**245**	-	**54**	54/245*100=**22.04**
paid-search	google	-	**178**	-	**52**	52/178*100=**29.21**

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

The following conclusions can be drawn from the purchase rate calculations:

1) In terms of attraction the potential customers to the Website, such campaigns as ten-crazy-cool-tshirts-facts (576 first touches led to 32.99% of last touch rate and then to 4.74% of final purchase rate) and getting-to-know-cool-tshirts (612 first touches led to 37.9 % of last touch rate and then to 3.88% of final purchase rate) have the highest purchase rates. While interview-with-cool-tshirts-founder and cool-tshirts-search campaigns led to 3.80% and 3.33% purchase rates with 622 and 169 first touches, 27.8 % and 32.99% of last touch rates correspondingly.

2) With regard to the return visitors to the Website, the most successful purchase rates have paid-search, weekly-newsletter and retargetting-ad campaigns with 29.21%, 25.73% and 25.51% appropriately.

3) Such campaigns as paid-search, weekly-newsletter and retargeting-ad are the definitely worthwhile reinvesting as they returned the majority of the "first touch visitors" back to the CoolTshirt website and caused the highest rates of final purchases. Paid-search is the leader with 29.21% of sales after returns, while weekly-newsletter and retargeting-ad goes with 25.73% and 25.51% accordingly.

campaign	campaign purpose	purchase rate (%)
interview-with-cool-tshirts-founder	Initiate first visit	3.80
getting-to-know-cool-tshirts	Initiate first visit	3.88
ten-crazy-cool-tshirts-facts	Initiate first visit	4.74
cool-tshirts-search	Initiate first visit	3.33
weekly-newsletter	Return visitors	25.73
retargetting-ad	Return visitors	25.51
retargetting-campaign	Return visitors	22.04
paid-search	Return visitors	29.21

For more sophisticated analysis more SQL queries are required to understand the relationship between campaigns and user's journey.

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

Taken together, the results indicate that the following campaigns are the best to reinvest in:

- ten-crazy-cool-tshirts-facts
- getting-to-know-cool-tshirts
- paid-search
- weekly-newsletter
- retargetting-ad