codecademy

First- and Last-Touch Attribution
CoolTShirts.com

Learn SQL from Scratch

Marilyn Smith

29 June 2018

Table of Contents

1. Get to know CoolTShirts.com

1.1 About CoolTShirts.com

If it's T-shaped and cool, CoolTShirts.com probably carries it. However, CoolTShirts.com needs to analyze their sales data and optimize their marketing so their customer reach can keep pace with current trends.

The table we will be working with is called **page_visits** and below is a sample of the data – just the first four results with no sorting or filtering applied. We notice that the "**page_name**" column begins with a number, probably to denote the intended order customers would visit them in. Each visit has a timestamp, and each user is designated a numerical user ID. The **utm_campaign** and **utm_source** describe how the user arrived to CoolTShirts.com

page_name	timestamp	user_id	utm_campaign	utm_source
1 – landing_page	2018-01-24 03:12:16	10006	getting-to-know-cool-tshirts	nytimes
2 – shopping_cart	2018-01-24 04:04:16	10006	getting-to-know-cool-tshirts	nytimes
3 – checkout	2018-01-25 23:10:16	10006	weekly-newsletter	email
1 – landing_page	2018-01-25 20:32:02	10030	ten-crazy-cool-tshirt-facts	buzzfeed

1.2 Campaigns and Sources

To start analyzing the data, we will need to query for information about CoolTShirts.com's campaigns and sources. This will tell more about how the customer makes their first visit. First, we will query for the number of distinct campaigns. We start with the basic idea of a **SELECT** query selecting from **page_visits**. Instead of selecting all columns ("*****") we want only the **utm_campaign** column. Also, instead of showing the data in the columns, we just want to see how many different campaigns appear, without counting duplicates. **DISTINCT** eliminates the duplicates. Enclosing that part of the statement in **COUNT(…)** results in the number of campaigns

There are 8 campaigns

Count(DISTINCT utm_campaign)
8

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

1.2 Campaigns and Sources

Next, we will do the same thing for the column **utm_source**....

This query is based on the query used to count the number of campaigns. We can edit the query so it is counting **utm_source** instead of **utm_campaign**.

There are 6 sources

COUNT(DISTINCT utm_source)
6

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

1.2 Campaigns and Sources

Finally, we will see how the campaigns and sources relate to each other. The query selects the two columns, **utm_source**, and **utm_campaign** and gets the present combinations of source and campaign. Adding **DISTINCT** eliminates duplicate entries.

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

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

The resulting table shows that not every campaign was accessed by every source. We can notice that Google related to search-related campaigns, Facebook related to an ad, and news and blogging sites related to articles and a top-10 list. This makes sense.

2. User Journey

2.1 Pages

We will get a list of pages using a **SELECT** query that searches for **DISTINCT** entries in the **page_name** column.

There are four. Their names begin with numbers and this makes them easy to arrange sequentially, as that order appears to proceed through the steps a customer would take to make a purchase.

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

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

2.2 First touches

A first touch is the timestamp of the first time a customer visited a page.

We will query to see how many first touches each campaign is responsible for. First, we create a temporary table, **first_touch**, with a **WITH** statement. The temporary table will be simply joined with the **page_visits** table. Finally we add a **SELECT** statement that counts the first-touches and displays them grouped by **utm_campaign**.

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

COUNT(ft.first_touch_at)	utm_campaign
169	cool-tshirts-search
612	getting-to-know-cool-tshirts
622	interview-with-col-tshirts-founder
576	ten-crazy-cool-tshirts-facts

2.3 Last touches

A last touch is the timestamp of the last time a customer visited a page. It is similar to a first, but it is the most recent timestamp instead of the oldest, e.g. the first timestamp. The query to find the last touch is similar. We change all references to “first touch” to reflect the change and, most importantly, change **MIN(timestamp)** to **MAX(timestamp)** which retrieves the most recent one instead of the first one.

COUNT(ft.first_touch_at)	utm_campaign
60	cool-tshirts-search
232	getting-to-know-cool-tshirts
184	interview-with-col-tshirts-founder
178	paid-search
443	retargetting-ad
245	retargetting-campaign
190	ten-crazy-cool-tshirts-facts
447	weekly-newsletter

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

2.4 Purchases

If we want to find out how many customers make a purchase, we can write a **SELECT** query. Instead of selecting all rows, we select **user_id**. Adding **COUNT(…)** and **DISTINCT** counts the entries of each username, ignoring duplicates. To retrieve this data for only the purchase page, we add a **WHERE** statement that specifies **'4 – purchase'**

There are 361 users who have made a purchase.

COUNT(DISTINCT user_id)
361

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

2.4 Purchases

We know how many customers made purchases. Now how did those customers get to the site?
We will take the last touch query and add to it the **WHERE** statement from the previous query that counted how many customers made purchases.

COUNT(lt.last_touch_at)	utm_campaign
2	cool-tshirts-search
9	getting-to-know-cool-tshirts
7	interview-with-cool-tshirts-founder
52	paid-search
112	retargetting-ad
53	retargetting-campaign
9	ten-crazy-cool-tshirts-facts
114	weekly-newsletter

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

3. Campaign Optimization

3.1 Reinvestment

Top five campaigns

1. Weekly newsletter – email to customers who have subscribed by entering their email during a purchase or directly signing up for the newsletter

2. Retargetting ad – ad that appears to customers after they have visited and left the page

3. Retargetting campaign – other ways of retargeting include cart recovery, upselling and cross-selling, and special targeting of first-time visitors, frequent visitors, and demographic and geographic targeting [1]

4. Paid search – a marketer can pay to have their website populated at the top of search results

5. Blogs – articles written on third party sites such as Buzzfeed

COUNT(lt.last_touch_at)	utm_campaign
114	weekly-newsletter
112	retargetting-ad
53	retargetting-campaign
52	paid-search
9	getting-to-know-cool-tshirts
9	ten-crazy-cool-tshirts-facts
7	interview-with-cool-tshirts-founder
2	cool-tshirts-search

[1] https://www.shopify.com/enterprise/62605443-how-to-execute-an-effective-ad-retargeting-campaign-that-works

3.1 Reinvestment

The top three campaigns – the weekly newsletter, the retargeting ad, and the retargeting campaign - are most relevant to returning customers, both those who have signed up or purchased before, and those who have just visited the site before at some point.

The third campaign, the retargeting campaign, has the potential to reach new and returning customers. For example, a new customer can receive a discount, and people who are frequent return customers may receive promotions that can include discounts, free gifts, and buy-one-get-one offers.

The fourth campaign, paid search, can reach new and existing customers through search engines.

Finally, third party blogs' content is up to the creators, but there is the potential for sponsored content on traditional blogs and social media like Instagram and YouTube where many users review and model products.

COUNT(lt.last_touch_at)	utm_campaign
114	weekly-newsletter
112	retargetting-ad
53	retargetting-campaign
52	paid-search
9	getting-to-know-cool-tshirts
9	ten-crazy-cool-tshirts-facts
7	interview-with-cool-tshirts-founder
2	cool-tshirts-search

3.1 Reinvestment

Since most of the purchases can be attributed to existing customers, and the purchases related to campaigns that primarily reached new customers have initiated far fewer sales, I propose the following plan:

1. Maintain existing customers and moderately increase investment in weekly newsletters, retargeting ads, and retargeting campaign. These are campaigns that are already working well and should be maintained and expanded to keep up with customer base expansion
2. Moderately increase investment in paid search to attract more new customers. This campaign is already working well but has potential for improvement
3. Substantially increase investment in creating strategies related to marketing on blogs and social networks to attract new customers. This campaign has attracted some customers to purchase but is likely not being utilized to its full potential

COUNT(lt.last_touch_at)	utm_campaign
114	weekly-newsletter
112	retargetting-ad
53	retargetting-campaign
52	paid-search
9	getting-to-know-cool-tshirts
9	ten-crazy-cool-tshirts-facts
7	interview-with-cool-tshirts-founder
2	cool-tshirts-search