Attribution Project

Learn SQL from Scratch

By: Irene Anibogwu

December 6th 2018

Table of Contents

1. Get Familiar with 'CoolTshirts'

In this section we will discuss:

◆ The differences between a UTM Source and a UTM Campaign
◆ The types of campaigns and sources CoolTshirts uses to make their business profitable and how they are related?
◆ What pages are on the CoolTshirts Website?

1.1: UTM SOURCE VS UTM CAMPAIGN

- A UTM source is a type of communication such as an email, search engine or social media site or website on which a campaign is run on.

For example: A customer who is looking for 'cool t-shirts' types this search phrase on Google and finds our store through this medium and buys a shirt. We can then attribute the source as Google because that's how the customer came to know about the 'coolTshirts' website in the first place

- A UTM campaign is a specific marketing effort that targets a particular audience from a specific source.

For example: If we ran an ad titled 'Cool T-shirts that wont make you look fat' on Facebook and a user clicks that link, it would direct them to our website where they would be able to purchase these 'cool shirts'. This 'ad' has peaked the interest of users who have experienced this problem, found it funny or just had a genuine interest and in doing so, has drawn new faces to our website.

1.2: TYPES OF UTM CAMPAIGNS AND SOURCES

- In this project, CoolTshirts is currently utilizing 6 sources and 8 campaigns for their marketing efforts.

- To find the relation between the campaigns and sources. Select DISTINCT utm_campaign and utm_source from page_visits. Here we'll find that while we can have multiple sources, we can also run different campaigns on each of these sources

- For example, 'Paid Search' and 'Cool-tshirts-search' can be attributed to Google's platform and 'Weekly-newsletter' and 'Retargeting-campaign' can be linked to Emails as a Source

Here's what a typical query would look like:

```
SELECT DISTINCT utm_campaign AS
'Campaign', utm_source AS 'Source'
        FROM page_visits;
```

Campaigns	Sources
8	6

Campaigns	Sources
Getting-to-know-cool-t-shirts	NY times
Weekly-newsletter	Email
Ten-crazy-cool-tshirts- shirts	Buzzfeed
Retargeting- campaign	Email
Retargeting-ad	Facebook
Interview-with-cool-tshirts-founder	Medium
Paid-search	Google
Cool-tshirts-search	Google

1.3: What are the page names on Cool T-Shirts Website?

- Cool-Tshirts website page names can be found using the DISTINCT value of page_names from the page_visits table.

- SELECT DISTINCT page_name AS 'Page Name' FROM page_visits;

- Here we see that the site has 4 distinct names for each section that a potential customer goes through.

Page name
1- Landing Page
2- Shopping Cart
3- Checkout
4- Purchase

2. What is the user's journey?

In this section we will discuss:

◆ How many first touches and last touches is each campaign responsible for?

◆ How many visitors make a purchase?

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

2.1 How many first touches is each campaign responsible for?

- First touch can be defined as the first time a user comes in contact with a website through a particular campaign from a specific source

- To answer this question, we must evaluate the minimum timestamp at which each visitor enters the site. This can be found by utilizing the MIN () and Count () and JOIN functions

- Take a look on the right to see what a typical query would look like

```
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_campaign,
pv.utm_source
          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 'First Touch
Campaign',
ft_attr.utm_source AS 'First Touch Source',
COUNT (*) AS 'Total Count'
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

2.1 How many first touches is each campaign responsible for cont'd?

- To explain our first touch query further, we created a temporary table (which we called 'last _touch') to find the MIN(timestamp) then we performed a JOIN back to the main table 'Page_Visits'. We then ended with a COUNT function to calculate the total number of visitors
- As we can see with the results below, of **the total 1,979 visitors, only 4 out of the 8 campaigns** were responsible for the first touches on the 'CoolTshirts' website

First Touch Campaign	First Touch Source	Total Count
Interview-with-cool-tshirts-founder	Medium	622
Getting-to-know-cool-tshirts	NY Times	612
Ten-crazy-cool-tshirts-facts	Buzzfeed	576
Cool-tshirts-search	Google	169

2.2 How many last touches is each campaign responsible for?

- Last touch can be defined as the last time the user comes in contact with the site. This can happen either through purchase or dropping off at one of the 4 pages on the 'CoolTshirts' Website.
- Similar to our last query, we will find the last touch by utilizing the JOIN and COUNT () functions but this time we will be looking for the **MAX (timestamp)** instead of the MIN.
- Take a look on the right to see what a typical query of this calculation may look like

```
WITH last_touch AS
           (SELECT user_id, MAX
(timestamp) AS 'last_touch_at'
FROM page_visits
GROUP BY user_id)

SELECT pv.utm_campaign AS 'Last Touch
Campaign', pv.utm_source AS 'Last Touch
Source',
COUNT (utm_campaign) AS 'Total Count'
FROM last_touch AS 'lt'
JOIN page_visits AS 'pv'
ON lt.user_id = pv.user_id AND
lt.last_touch_at = pv.timestamp
GROUP BY 1
ORDER BY 3 DESC;
```

2.2: Last touch campaign cont'd

- Evaluating the results from our query on the previous page, all 8 of our campaigns returned a last touch however, the most profitable were the **weekly news letters and the retargeting campaigns/ads**
- **It is also important to note that the 4 campaigns(highlighted in yellow) responsible for the most first touches, seem to not contribute as much to the last touches**
- **This tells us that those 4 campaigns are best for introducing users to the site but more needs to be done to keep them there**

Last Touch Campaign	Last Touch Source	Total Count
Weekly-newsletter	Email	447
Retargeting –ad	Facebook	443
Retargeting-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?

- To find this value, we would need to count the distinct user_id for where the user appears in the page name titled '4-purchase'.
- From the results, we find that of the 1,979 visitors, 361 made a purchase.

Query:
SELECT
COUNT (DISTINCT user_id) AS 'Customers'
FROM page_visits
WHERE page_name = '4 - purchase';

Customers
361

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

- At this point we know that 361 out of the 1,979 made a purchase on CoolTshirts website.
- This would mean, ratio wise that:
- 361/1979 = **18.24%** of people who visit the site made a   purchase
- So from the results to our query below, we can see that 4 campaigns (highlighted in yellow) are responsible for the most 'last touches' which led to a purchase.
- **It is also important to note that the campaigns that focused on retargeting/ reoccurring marketing can be credited for the most purchases as compared to the campaigns that focused on stories about the company**

A Query for this solution would look something like this:
```
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 pv.utm_campaign AS 'Purchase Campaign',
                 pv.utm_source AS 'Purchase Source',
                 COUNT (utm_campaign) AS 'Total Count'
FROM last_touch AS 'lt'
JOIN page_visits AS 'pv'
ON lt.user_id = pv.user_id AND
lt.last_touch_at = pv.timestamp
GROUP BY 1
ORDER BY 3 DESC;
```

Purchase Campaign	Purchase Count	Total Count
Weekly-newsletter	email	115
Retargeting-ad	Facebook	113
Retargeting-campaign	Email	54
Paid-search	Google	52
Getting-to-know-cool-tshirts	Nytimes	9
Ten-crazy-cool-tshirts-facts	Buzzfeed	9
Interview-with-cool-tshirts-founder	Medium	7
Cool-tshirts-search	Google	2

3. Optimizing the campaign budget

In this section we will discuss:

- Which 5 campaigns Cool-Tshirts should focus on and why?

3.1: Which 5 campaigns should 'coolTshirts' invest in?

Source	Campaign	First Touch Count Total = 1,979 users	First-Total Conversion	Last Touch Count	Purchase Count Total = 361 customers	Purchase/ customers Conversion
NY times	Getting-to-know-cool-t-shirts	612	31%	232	9	2.4%
Email	Weekly-newsletter	-	0%	447	115	32%
Buzzfeed	Ten-crazy-cool-tshirts- shirts	576	29%	190	9	2.4%
Email	Retargeting-campaign	-	0%	245	54	15%
Facebook	Retargeting-ad	-	0%	443	113	31.3%
Medium	Interview-with-cool-tshirts-founder	622	31%	184	7	1.9%
Google	Paid-search	-	0%	178	52	14.4%
Google	Cool-tshirts-search	169	8.5%	60	2	.5%

Conclusion: Optimizing the budget

- From our previous chart, we can see that the most profitable campaigns coming in at 1st and 2nd are **the weekly-newsletters through Email and the retargeting ads on Facebook.**
- Coming in 3rd and 4th would be the **retargeting campaign through Emails and paid searches on Google.**
- I believe it would be profitable to **reinvest in all 4 campaigns** as these campaigns peaks the consumers interest by bringing the product directly to them.
- As for the 5th choice my suggestion would be this:

  - Invest in a **retargeting campaign with any of the 3 story campaigns.** These campaigns were pivotal in drawing people's attention to the website however, more needs to be done to keep them there.

  - Seems like we are missing out on a completely different market by not investing in a retargeting effort with these campaigns and more specifically, on the NYTimes, Buzzfeed, and Medium platforms.

  - Speaking off of the numbers with the first touches these particular stories received, it's a reasonable conclusion that reinvesting in them would prove to be profitable.

  - To be more specific, we can begin this retargeting effort with the **'Interview-with-cool-Tshirts-founder' campaign on Medium's platform as it did receive the most first touches.**