Attribution Queries

SQL From Scratch

Gerry Furlong

1/6/2019

I. Get to know the company CoolTShirts

    I. How many campaigns and sources does CoolTShirs use and how are they related?

    II. What pages are on their website?

II. What is the users journey?

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

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

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

    IV. What is the typical user journey?

III. Optimize the campaign budget

    I. CoolTShirts can re-invest in 5 campaigns. Which would they pick and why?

Attribution Queries
Table of Contents

- *Get familiar with CoolTShirts*
  - *What is the user journey?*
  - *Optimize campaign budget*

Get Familiar with CoolTShirts

- How many campaigns and sources does CoolTShirs use and how are they related?
  - A **source** is a communication method that uses a website to attract a customer to a campaign. The websites used for this purpose are listed on the right. Sources may run multiple campaigns and change to target an audience. The source field in this db is utm_source

  - A **campaign** is a specific method to elicit a specific response from possible purchasers. The same campaign can be run over multiple sources. In this case a campaign is specific to a source. The campaign field in this db is utm_campaign

```
test.sqlite

1    --Count Campaigns
2    Select Count(Distinct utm_campaign) AS
     'Campaign Count'
3    From page_visits;
4    --Count Source
5    Select Count(Distinct utm_sources) AS 'Source
     Count'
6    From page_visits;
7    --List Campaign and Sources
8    Select Distinct utm campaign As 'Campaign',
     utm_source AS 'Source'
9    From page_visits;
10
```

Query Results	Query Results
Campaign Count	Source Count
8	6

Query Results	
Campaign	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

Get Familiar with CoolTShirts

- What pages are on their website?

  - A simple distinct statement is used to identify the pages on the website
  - Each page is part of the user journey once they have selected a campaign from a source.
  - Presumably the pages are the steps or clicks the user processes through to purchase product(s).

Query Results

Pages
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

test.sqlite

```
1   -- Find the unique pages on website
2   Select Distinct page_name As 'Pages'
3   From page_visits;
4
5
```

What is the users journey?

- How many first touches is each campaign responsible for?
  - The first touch is found by identifying the MIN timestamp or the first time the user touched a campaign.
  - Creating the temp table ft_attr and joining with page_visits gives us the join between the user and first touch to the source and campaign. Then we counted the number of first touches and sources.
  - Noted is that from the 8 campaigns only 4 are responsible for the first touch.

Query Results

Source	Campaign	First Touch Count
medium	interview-with-cool-tshirts-founder	622
nytimes	getting-to-know-cool-tshirts	612
buzzfeed	ten-crazy-cool-tshirts-facts	576
google	cool-tshirts-search	169

```sql
--Locate first touch
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id),
--Create temp table, join to page_visits --
to locate source and campaign and
--count
ft_attr AS (
    SELECT ft.user_id,
        ft.first_touch_at,
        pv.utm_source,
        pv.utm_campaign
    FROM first_touch ft
    JOIN page_visits pv
        ON ft.user_id = pv.user_id
        AND ft.first_touch_at = pv.timestamp)
Select ft_attr.utm_source AS 'Source',
        ft_attr.utm_campaign AS 'Campaign',
        Count(*) As ' First Touch Count'
From ft_attr
Group by 1, 2
Order by 3 DESC;
```

test.sqlite

What is the users journey? – cont'd

- How many last touches is each campaign responsible for?
  - Last touch is found by the Max Timestamp and user.
  - Using the same concept as first touch, creating a temp table lt_attr and joining with the page_visits we can count the last touches by campaign.
  - Noted is those sites that drove first touches are low on the list for the last touches.

Query Results		
Source	Campaign	Last Touch Count
email	weekly-newsletter	447
facebook	retargetting-ad	443
email	retargetting-campaign	245
nytimes	getting-to-know-cool-tshirts	232
buzzfeed	ten-crazy-cool-tshirts-facts	190
medium	interview-with-cool-tshirts-founder	184
google	paid-search	178
google	cool-tshirts-search	60

test.sqlite

```sql
1  --locate all last touches
2  With last_touch AS(
3    Select user_id, Max(timestamp) AS
       last_touch_at
4    from page_visits
5    Group by user_id),
```

```sql
--Create second temp table. This will Join
the --last_touches_at to the utm_campaign and
--utm_source on the user_id and timestamp.
lt_attr AS(
Select lt.user_id
        lt.last_touch_at
        pv.utm_source
        pv.utm_campaign
From last_touch ft
Join page_visits pv
  On lt.user_id = pv.user_id
  And lt.last_touch_at = pv.timestamp)
Select lt_attr.utm_source AS
'Source', utm_campaign AS 'Campaign'
Count(*) AS 'Last Touch Count'
```

What is the users journey? – cont'd

- How many visitors made a purchase?
  - Page_name '4 – purchase' identifies when a user has made a purchase.
  - Using Count(Distinct user_id) enables us to count the users. The Where clause filters the users to those that made a purchase.
  - Visitors convert to customers at the rate of 18% (361/1979)

Query Results

Count(Distinct user_id)

361

test.sqlite

```
1   --Number of users that made a purchase
2   Select Count(Distinct user_id)
3   From page_visits
4   Where page_name = '4 - purchase';
```

What is the users journey? – cont'd

- How many last touches on the purchase page is each campaign responsible for
  - We identify the last touch by using the Max(timestamp) and filtering this group to the purchase page. We can then identify the campaign(s) most responsible for converting a visitor to a customer. This helps the company evaluate a campaign's level of success.
  - We see the same trend as prior, where the first touch sources decrease when looking through the lens of purchase.

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

test.sqlite

```sql
1   --Identify last touch and filter to the
2   |- purchase page, join to
3   --page_visits to link to utm_source,
4   -- utm_campaign by creating temp table
5   WITH last_touch AS(
6   Select user_id, Max(timestamp) AS
    last_touch_at
7   From page_visits
8   Where page_name = '4 - purchase'
9   Group by user_id),
10  lt_attr AS (
11  Select lt.user_id,
12      lt.last_touch_at,
13      pv.utm_source,
14      pv.utm_campaign
15  From last_touch lt
16  Join page_visits pv
17  On lt.user_id = pv.user_id
18  And lt.last_touch_at = pv.timestamp)
18  --From temp table lt_attr we count
19  --last touches attributed to the source
    --and campaign and order
20  Select
21    lt_attr.utm_source As 'Source',
22    lt_attr.utm_campaign As 'Campaign',
23    Count(*) As 'Last Touch Count'
24  From lt_attr
25  Group by 1,2
26  Order by 3 DESC;
```

Typical User's Journey
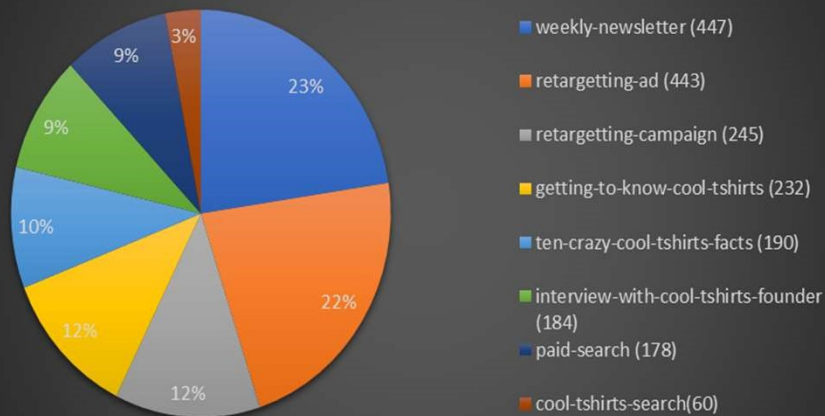
First Touch by Campaign

- interview-with-cool-tshirts-founder (622)
- getting-to-know-cool-tshirts (612)
- ten-crazy-cool-tshirts-facts (576)
- cool-tshirts-search (169)

Majority of users come to the website via 3 campaigns
All of which are informational in nature. (91%)

Last Touch w Purchase

- weekly-newsletter (115)
- retargetting-ad (113)
- retargetting-campaign (54)
- paid-search (52)
- ten-crazy-cool-tshirts-facts (9)
- getting-to-know-cool-tshirts (9)
- interview-with-cool-tshirts-founder (7)
- cool-tshirts-search (2)

Last Touch by Campaign

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

- Last touches by *reminder campaigns* such as email newsletter and re-targeting are *57%*
  - *Information campaigns* returned users at a rate of *31%* but did not result in strong sales.

- *Purchases* were more likely when the last touch was a *reminder campaign.*
  - Within the reminder campaigns 25% of last touch campaign became purchasers with those being 78% of total sales.
  - Conversion rate from last touch to purchase is 18%.

Campaign Re-Investment

Campaign	First Touch	% First Touch	Last Touch	% Last Touch	Purchases	% Purchases
interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%
getting-to-know-cool-tshirts	612	31%	232	12%	9	2%
ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	2%
cool-tshirts-search	169	9%	60	3%	2	1%
weekly-newsletter			447	23%	115	32%
retargetting-ad			443	22%	113	31%
retargetting-campaign			245	12%	54	15%
paid-search			178	9%	52	14%
Total	1979		1979		361	

Customers that arrive at CoolTShirts website do so by the top 3 campaigns as highlighted in green. These campaigns bring in 91% of the total.

Customers that purchase do so primarily through re-targeting and the newsletter. These 3 campaigns contribute to 78% of the sales.

- **Five campaigns can be funded. Recommendation is to fund:**
- *Interview with Founder*
- *Getting to Know CoolTShirts*
- *Ten Crazy CoolTShirt facts*
- *Re-targeting Ad*
- *Re-targeting Campaign*
  - *While the newsletter is successful it should be discontinued. Re-targeting is a very good revenue generator. These 2 campaigns generate 46% of the revenue. As one of the re-targeting sources is email, it may actually be a redundant method. Focusimg on increasing email capture during the process could maximize the re-targeting ads and compensate for the loss of the newsletter. Consider increased use of new sources for re-targeting ads. (see following slides)*

Campaign Re-Investment – cont'd

Query Results

Page Name	User Visits
4 - purchase	361
3 - checkout	1431

Query Results

Source	Users Visits
nytimes	747
email	696
buzzfeed	648
medium	625
facebook	445
google	339

- Additional Notes
  - Checkout to Purchase data further supports re-targeting as an area of further opportunity. 25% go from checkout to purchase. Enhancing re-targeting may increase this conversion rate.

  - Evaluate the sources used and the number of visits. Findings indicate that some sources are underutilized. Opportunities for re-targeting ads in the NY Times, Buzzfeed, and Medium may exist due to the general high volume visits to these sites by CoolTShirt customers.

  - The informational campaigns generate a large amount of first touches but yield little in actual purchase, accounting for 6% of purchases. You may consider expanding the number of sources for first touch. Adding Facebook to the first touch campaigns may be useful as the re-targeting ads generate 31% of the revenue.