

FIRST-AND-LAST-TOUCH ATTRIBUTION

Learn SQL from Scratch

Steven Feyerer

6/28/2018 – 8/7/2018

TABLE OF CONTENTS

1. Get familiar with the company.
 - 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`.
 - How many pages are on the CoolTShirts website?
2. What is the user journey?
 - How many first touches is each campaign responsible for?
 - How many 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?
 - What is the typical user journey?
3. Optimize the campaign budget.
 - CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

GET FAMILIAR WITH COOLTSHIRTS CAMPAIGNS AND SOURCES

- A campaign is a specific ad or email blast, such as a newsletter or sale ad, that attracts a customer to the site
- A source is the site that sent the traffic
- Each source can have multiple campaigns, as shown in the table
- For CoolTShirts, its 8 campaigns are used across 6 sources, with google and email each utilizing 2 different campaigns. Each other source uses a unique campaign.
- Although it is easily possible to use the same campaign across multiple sources, that does not happen in this situation.

Campaign Count	Source Count
8	6
Campaign	Source
getting-to-know-cool-tshirts	nytimes
weekly-newsletter	email
ten-crazy-cool-tshirts-facts	buzzfeed
retargeting-campaign	email
retargeting-ad	facebook
interview-with-cool-tshirts-founder	medium
paid-search	google
cool-tshirts-search	google

```
1  -- Get a count of campaigns and sources
2  SELECT COUNT(DISTINCT utm_campaign) AS 'Campaign Count',
3         COUNT(DISTINCT utm_source) AS 'Source Count'
4  FROM page_visits;
5
6  -- Determine the relationship between the campaigns and sources
7  SELECT DISTINCT utm_campaign AS Campaign,
8                 utm_source AS Source
9  FROM page_visits;
```

GET FAMILIAR WITH COOLTSHIRTS CONTINUED

- What pages are on the CoolTShirts website?
 - Use a SELECT DISTINCT query to find all Page Names while verifying the validity of the query so as not to include duplicates in the result.
 - Most likely, there are more than four pages on the site. However, these were the four that were listed in the dataset.
 - Renaming the result using an AS clause helps with readability.

Page Names	
1 -	landing_page
2 -	shopping_cart
3 -	checkout
4 -	purchase

```
9  -- Find the distinct page names on the CoolTShirts website
10 SELECT DISTINCT page_name AS 'Page Names'
11 FROM page_visits;
```

WHAT IS THE USER JOURNEY?

FIRST TOUCH ATTRIBUTION CODING

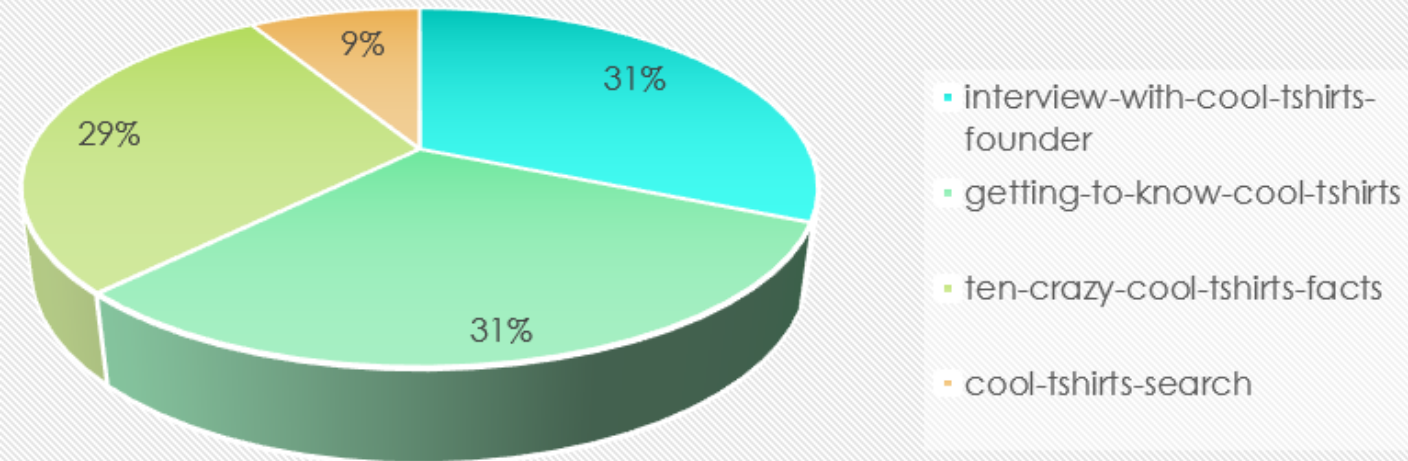
- By using the MIN(timestamp) query and grouping by user id, we determine each prospect's initial visit to the site.
- Next we create a joined table that shows the association of each campaign and source with each prospect's initial visit.
- Finally, we are able to count the number of times each campaign resulted in a prospect's initial visit to the website.

```
12 -- Find the campaign and source responsible for each user's first touch
13 -- First we create a temp table that finds first touches
14 -- and groups them by user id
15 WITH first_touch AS (
16     SELECT user_id,
17            MIN(timestamp) AS first_touch_at
18     FROM page_visits
19     GROUP BY user_id),
20 -- Then we create a second temp table joining campaigns
21 -- and sources from the page_visits table on user_id and timestamp
22 ft_attr AS (
23     SELECT ft.user_id,
24            ft.first_touch_at,
25            pv.utm_campaign,
26            pv.utm_source
27     FROM first_touch ft
28     JOIN page_visits pv
29         ON ft.user_id = pv.user_id
30         AND ft.first_touch_at = pv.timestamp)
31 -- Now we select and count the number of times each campaign and source
32 -- is responsible for a first touch
33 SELECT ft_attr.utm_campaign AS 'Campaign',
34        ft_attr.utm_source AS 'Source',
35        COUNT(*) AS 'Count'
36 FROM ft_attr
37 GROUP BY 1, 2
38 ORDER BY 3 DESC;
```


WHAT IS THE USER JOURNEY?

FIRST TOUCH ATTRIBUTION ANALYSIS

First Touch Attribution



Campaign	Source	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

- Of the eight campaigns used by CoolTShirts, only four actually result in a first touch to the website.
- Of the four campaigns with first touch attribution, a cool-tshirts-search through Google is by far the least effective at generating new prospects, accounting for only 9% of first touches, less than 1/3 of the next lowest campaign.

WHAT IS THE USER JOURNEY?

LAST TOUCH ATTRIBUTION CODING

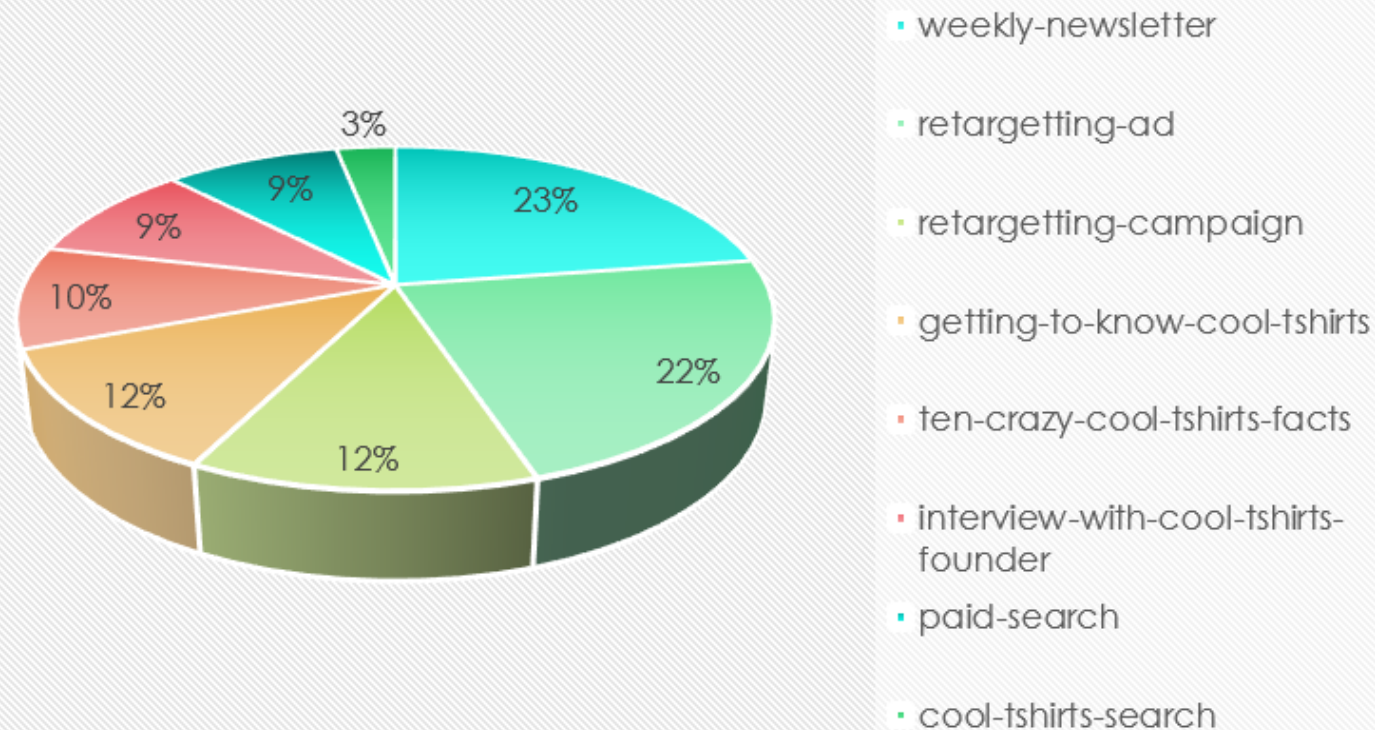
- By using the MAX(timestamp) query and grouping by user id, we determine each prospect's final visit to the site.
- Next we create a joined table that shows the association of each campaign and source with each prospect's final visit.
- Finally, we are able to count the number of times each campaign resulted in a prospect's final visit to the website.

```
39 -- Find the campaign and source responsible for each user's last touch
40 -- First we create a temp table that finds last touches
41 -- and groups them by user_id
42 WITH last_touch AS (
43     SELECT user_id,
44            MAX(timestamp) AS last_touch_at
45     FROM page_visits
46     GROUP BY user_id),
47 -- Then we create a second temp table joining the campaigns and sources
48 -- from the page_visits table on user_id and timestamp
49 lt_attr AS (
50     SELECT lt.user_id,
51            lt.last_touch_at,
52            pv.utm_campaign,
53            pv.utm_source
54     FROM last_touch lt
55     JOIN page_visits pv
56         ON lt.user_id = pv.user_id
57        AND lt.last_touch_at = pv.timestamp)
58 -- Now we select and count the number of times each campaign and source
59 -- is responsible for a last touch
60 SELECT lt_attr.utm_campaign AS 'Campaign',
61        lt_attr.utm_source AS 'Source',
62        COUNT(*) AS 'Count'
63 FROM lt_attr
64 GROUP BY 1, 2
65 ORDER BY 3 DESC;
```

WHAT IS THE USER JOURNEY?

LAST TOUCH ATTRIBUTION ANALYSIS

Last Touch Attribution



Campaign	Source	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

- Retargeting campaigns through email and Facebook are the most effective at generating last touches
- Interesting observation: the four campaigns responsible for generating initial visits to the site are also four of the five **least effective** campaigns for generating final visits to the site
- Getting-to-know-cool-tshirts is the second most effective campaign for first touches, and is still in the top half for generating last touches, indicating a successful overall campaign

WHAT IS THE USER JOURNEY?

HOW MANY VISITORS MAKE A PURCHASE?

- An important part of understanding the success of a marketing campaign is how many users actually make a purchase.
 - By counting only DISTINCT user ids from our page visits table, we eliminate the risk of over counting those customers who make multiple purchases.
 - By using a WHERE clause in our query we are able to narrow down the scope of our search to those visitors who actually make a purchase.

Customers Making a Purchase	
361	
<pre>66 -- Find the number of visitors who make a purchase 67 SELECT COUNT (DISTINCT user_id) AS 'Customers Making a Purchase' 68 FROM page_visits 69 WHERE page_name = '4 - purchase';</pre>	

- 361 of the 1,979 users who visited the site made a purchase.
 - This is a purchase rate of roughly 18% of visitors

WHAT IS THE USER JOURNEY?

PURCHASE ATTRIBUTION CODING

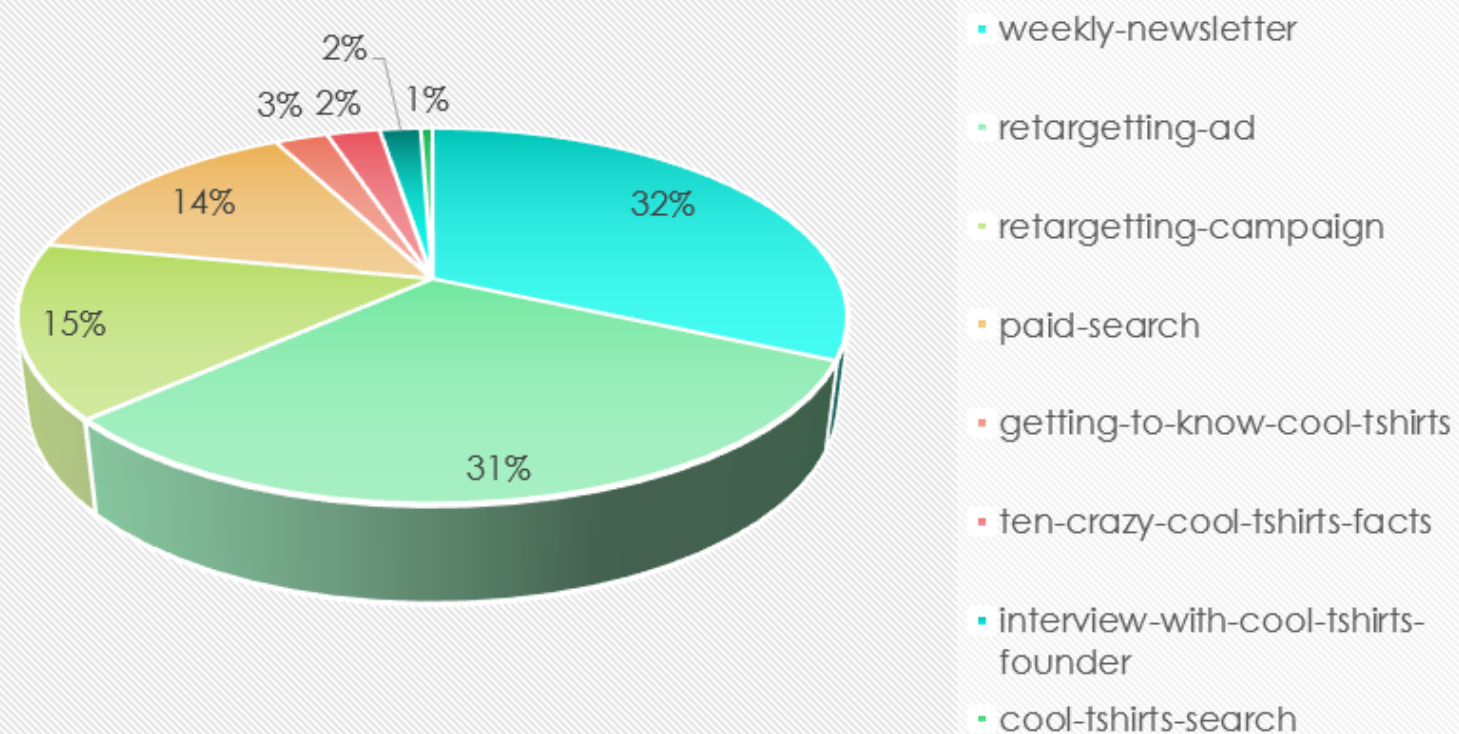
- In order to understand how each marketing campaign is doing, we need to find how many purchases each is responsible for.
 - Creating the temporary tables enables us to more easily look at only those visitors who make a purchase while still knowing which campaign is responsible for their purchasing visit

```
70 -- Find the number of LAST TOUCHES from each campaign
71 -- that result in a purchase
72 -- First we create a temp table to find last touches by user id
73 WITH last_touch AS (
74     SELECT user_id,
75            MAX(timestamp) AS last_touch_at
76     FROM page_visits
77     -- WHERE clause focuses query to only users that make a purchase
78     WHERE page_name = '4 - purchase'
79     GROUP BY 1),
80 -- Now we create a second temp table joining our first
81 -- with the campaign and source columns from the page_visits table
82 lt_attr AS (
83     SELECT lt.user_id,
84            lt.last_touch_at,
85            pv.utm_campaign,
86            pv.utm_source
87     FROM last_touch lt
88     JOIN page_visits pv
89     ON lt.user_id = pv.user_id
90     AND lt.last_touch_at = pv.timestamp)
91 -- Finally we determine the number of purchases each campaign
92 -- is responsible for
93 SELECT lt_attr.utm_campaign AS 'Campaign',
94        lt_attr.utm_source AS 'Source',
95        COUNT(*) AS 'Count'
96 FROM lt_attr
97 GROUP BY 1, 2
98 ORDER BY 3 DESC;
```

WHAT IS THE USER JOURNEY?

PURCHASE ATTRIBUTION ANALYSIS

Purchase Attribution



Campaign	Source	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

- The top three campaigns responsible for last touches to the site are also the top three campaigns responsible for purchases.
- The ONLY four campaigns responsible for first touches are the bottom four in terms of user purchases, accounting for only 8% of purchases.

WHAT IS THE USER JOURNEY?

TYPICAL USER JOURNEY

- The typical user journey begins with one of three 'article' campaigns.
 - interview-with-cool-tshirts-founder
 - getting-to-know-cool-tshirts
 - ten-crazy-cool-tshirts-facts
 - These articles account for 91% of first touches to the CoolTShirts website
- The majority of visitors do not make a purchase on the initial visit
 - Retargeting campaigns account for 78% of purchases.
 - Weekly newsletters through email account for 32% of purchases
 - Retargeting ads through Facebook account for 31% of purchases
 - Retargeting email campaigns account for 15 % of purchases
 - The four campaigns responsible for 100% of first touches only account for a total of 8% of purchases
 - This demonstrates the importance of finding a balance between the article campaigns and retargeting campaigns.

OPTIMIZE THE CAMPAIGN BUDGET OVERVIEW

CAMPAIGN	FIRST TOUCHES	% OF FIRST TOUCHES	LAST TOUCHES	% OF LAST TOUCHES	PURCHASES	% OF PURCHASES	TOUCH-TO-PURCHASE RATE
weekly-newsletter	0	0%	447	23%	115	32%	25.727%
retargeting-ad	0	0%	443	22%	113	31%	25.508%
retargeting-campaign	0	0%	245	12%	54	15%	22.041%
paid-search	0	0%	178	9%	52	14%	29.213%
getting-to-know-cool-tshirts	612	31%	232	12%	9	2%	1.066%
ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	2%	1.175%
interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%	0.868%
cool-tshirts-search	169	9%	60	3%	2	1%	0.873%

- Based on the table above:
 - the three 'article' campaigns are each responsible for a much higher percentage of first touches than all other campaigns.
 - the weekly-newsletter and retargeting-ad are responsible for a much higher percentage of last touches, as well as purchases, than all other campaigns.

OPTIMIZE THE CAMPAIGN BUDGET

TOUCH-TO-PURCHASE RATE

CAMPAIGN	FIRST TOUCHES	% OF FIRST TOUCHES	LAST TOUCHES	% OF LAST TOUCHES	PURCHASES	% OF PURCHASES	TOUCH-TO-PURCHASE RATE
weekly-newsletter	0	0%	447	23%	115	32%	25.727%
retargeting-ad	0	0%	443	22%	113	31%	25.508%
retargeting-campaign	0	0%	245	12%	54	15%	22.041%
paid-search	0	0%	178	9%	52	14%	29.213%
getting-to-know-cool-tshirts	612	31%	232	12%	9	2%	1.066%
ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	2%	1.175%
interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%	0.868%
cool-tshirts-search	169	9%	60	3%	2	1%	0.873%

- Although not part of the required data, I felt it useful to determine the percentage of touches from each campaign that resulted in a purchase. This results in a better understanding of the efficiency of each campaign in generating sales.
 - Using this metric, it is clear that the three retargeting campaigns, along with the paid-search, are most effective at going from visit to purchase.

OPTIMIZE THE CAMPAIGN BUDGET RECOMMENDATIONS

- **Recommend continuing:**

- getting-to-know-cool-tshirts and ten-crazy-cool-tshirts-facts campaigns based on first touch effectiveness
- Facebook retargeting-ad and email weekly-newsletter based on last touch effectiveness and 25% touch-to-purchase rate of each
- Google paid-search based on 29% touch-to-purchase rate, indicating close to 3 out of 10 users who go to the site based on this campaign make a purchase, a higher rate than any other campaign

- **Recommend discontinuing:**

- Google cool-tshirts-search based on low percentages across all categories
- Email retargeting-campaign based on redundancy with email weekly-newsletter, which has a higher level of effectiveness in both last touches and touch-to-purchase rate.
- interview-with-cool-tshirts-founder 'article' campaign
 - In spite of having the highest percentage of first touches of any campaign, it results in the third fewest last touches and has the lowest touch-to-purchase rate of any campaign

ADDITIONAL SUGGESTIONS

- It does not make sense to re-invest in both email campaigns in spite of reasonably high touch-to-purchase rates.
 - It would seem logical that either email-based campaign would be reasonably effective at generating last touches and sales.
 - However, using both would potentially overwhelm and frustrate potential customers, thereby discouraging them to actually return to the site to make a purchase.
- If more information was known about each campaign, such as cost and man-hours involved in maintaining the campaigns, the recommendations may change.
 - If the email newsletters, for instance, cost significantly more than the email retargeting ads to maintain, then it may be more cost effective to discontinue the newsletters and re-invest in the retargeting ads.
 - Also, if the cool-tshirts-search campaign through Google were relatively inexpensive compared to other campaigns, it may not be a bad campaign in which to re-invest because it does produce more first touches than half of the campaigns.