



SQL Capstone:

First-Touch and Last-Touch Attribution

Learn SQL from Scratch

Andrea Nelson

8/12/2018

Table of Contents

1. Get familiar with CoolTShirts
2. User Journey
3. Optimization of the Campaign Budget

1. Get Familiar with CoolTShirts

1.1 Campaigns and Sources

CoolTShirts has a total of eight campaigns (utm_campaign) and six sources (utm_source). The campaigns and sources are used in tandem to maximize their marketing. The first touch could be through any of the sources, but focused campaigns can remind a potential purchase of an abandoned shopping cart.

- Campaigns: marketing advertisements used to pique interest in their product by using staff interviews, listing facts about their company/product, and introducing themselves by way of media.
- Sources: outlets used by the company to maximize their marketing campaigns. Some are internal, and others are external paid sources that have a wide audience.

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

1.2 Pages on the CoolTShirts Website

CoolTShirts consists of four distinct pages. This is derived from selecting distinct pages from the table (code to the right).

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

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

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

2. User Journey

2.1 How many first touches is each campaign responsible for?

The First-Touch for each campaign can be shown for this table by using the SQL code to the right. Based on the results, you can see that interview with the CoolTShirts founder campaign through Medium had the highest impact, followed closely by the introduction feature posted with the New York Times.

First-Touch Source	First-Touch Campaign	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

```
WITH first_touch AS (  
  SELECT user_id, MIN(timestamp) as  
  'first_touch_at'  
  FROM page_visits  
  GROUP BY user_id),  
ft_camp 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_camp.utm_source AS 'First-Touch Source',  
ft_camp.utm_campaign AS 'First-Touch Campaign',  
COUNT(*) AS 'Count'  
FROM ft_camp  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

2.2 How many last touches is each campaign responsible for?

The Last-Touch for each campaign is shown to the right with an algorithm similar to the First-Touch code. Results show the emailed company newsletter with the highest count, with the retargeting ad through Facebook as a close second. Natural search from people using Google has the lowest count, which was also the case with the First-Touch results.

Last-Touch Source	Last_Touch Campaign	Count
email	weekly-newsletter	447
facebook	retargeting-ad	443
email	retargeting-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

```
WITH last_touch AS (  
  SELECT user_id, MAX(timestamp) as  
  'last_touch_at'  
  FROM page_visits  
  GROUP BY user_id),  
lt_camp AS(  
  SELECT lt.user_id, lt.last_touch_at,  
  pv.utm_source, pv.utm_campaign,  
  pv.page_name  
  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_camp.utm_source AS 'Last-  
Touch Source', lt_camp.utm_campaign AS  
'Last_Touch Campaign', COUNT(*) AS  
'Count'  
FROM lt_camp  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```


2.3 How many visitors make a purchase?

The website for CoolTShirts is set up in four pages (stages). The table to the right shows the total count for visitors to each page. While the landing page boast 2,000 visitors, we end up with 18% of the total visitors making a purchase with a count of 361. The major bounce happens on the checkout page. The purchase count is by not means paltry, but a customer survey campaign that asks directed questions regarding the website could help CoolTShirts find out what is causing in the significant drop-off in the checkout stage.

Page Name	Count	Drop-Off % From Total
1 - landing_page	2000	100%
2 - shopping_cart	1900	95%
3 - checkout	1431	72%
4 - purchase	361	18%

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

The code to the right gives us the below table. While similar to the Last-Touch process, use of the WHERE function allows us to hone in on the purchases we can attribute to each campaign. Based on the results, it shows that the CoolTShirts email campaigns bring in the majority of their online sales.

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

```
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),  
lt_camp AS(  
  SELECT lt.user_id, lt.last_touch_at,  
  pv.utm_source, pv.utm_campaign,  
  pv.page_name  
  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_camp.utm_source AS 'Last-  
Touch Purchase Source',  
lt_camp.utm_campaign AS 'Last-Touch  
Purchase Campaign', COUNT(*) AS  
'Purchase Count'  
FROM lt_camp  
  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

2.5 What is the typical user journey?

```
WITH first_touch AS (  
  SELECT user_id, MIN(timestamp) as  
    'first_touch_at'  
  FROM page_visits  
  GROUP BY user_id),  
ft_camp 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 COUNT(*) AS 'Count'  
FROM ft_camp;
```

Based on the total unique visitor count (2,000 shown on slide 2.3), we see that most of the visitors came to the site because of marketing campaigns. The total count for these campaigns is 1,979. We can find this by making a simple change to the SELECT portion of our First-Touch code (shown to the left). Or, simply add up the campaign counts with a calculator. From the Last-touch campaigns we see that 167 of the 361 total purchases had to be retargeted before making their purchases. From this we can gather that 46% of those purchasing are not making the purchase on their first visit. Looking back at the count by page, we can see that the typical journey is portrayed in the graphic below. The bounce comes happens most frequently at purchase.

Page Name		Count
1 - landing_page	↓	2000
2 - shopping_cart	↓	1900
3 - checkout	↓	1431
4 - purchase	×	361

3. Optimization of the Campaign Budget

3.1 Optimized by First-Touch

If we look solely at the counts on the First-Touch campaigns, we might think that putting all of the budget into sources like Medium and New York Times is the way to go. We see that 1,810 people responded to the articles that were placed with Medium, the New York Times, and BuzzFeed. However, we also see that the purchase count on each of these campaigns is low. We see each one hovering around 1%.

Source	Campaign	First-Touch Count	Purchase Count	Conversion %
medium	interview-with-cool-tshirts-founder	622	7	1.1 %
nytimes	getting-to-know-cool-tshirts	612	9	1.4 %
buzzfeed	ten-crazy-cool-tshirts-facts	576	9	1.5 %
google	cool-tshirts-search	169	2	1.1 %

3.2 Optimized by Last-Touch

By comparing our Last-Touch count to our purchase count, we are better able to see the which campaigns gave the highest conversion rate to CoolTShirts. If we focus on the percentages we see that the paid Google search campaign had the highest purchase percentage. If we look at combined purchase count by source, we see that marketing campaigns by email brought in the most sales. Also, when we combine the retargeting campaigns (Facebook and email sources), we see this count as a close second.

Source	Campaign	Last-Touch Count	Purchase Count	Conversion %
google	paid-search	178	52	29.0 %
email	weekly-newsletter	447	115	25.7 %
facebook	retargeting-ad	443	113	25.5 %
email	retargeting-campaign	245	54	22.0 %
buzzfeed	ten-crazy-cool-tshirts-facts	190	9	4.7 %
nytimes	getting-to-know-cool-tshirts	232	9	3.8 %
medium	interview-with-cool-tshirts-founder	184	7	3.8 %
google	cool-tshirts-search	60	2	3.3 %

3.4 Recommendations for Budget

Given the results, my recommendation would be to focus on paid searches with Google, and retargeting ads with Facebook. Focusing on these will give you the most bang for your buck. This does not mean you should not use other sources, especially since they can help get the word out about the company, but those sources (Buzzfeed, New York Times, etc.) should be secondary.

Source	Campaign	Last-Touch Count	Purchase Count	Conversion %
google	paid-search	178	52	29.0 %
email	weekly-newsletter	447	115	25.7 %
facebook	retargeting-ad	443	113	25.5 %
email	retargeting-campaign	245	54	22.0 %
buzzfeed	ten-crazy-cool-tshirts-facts	190	9	4.7 %
nytimes	getting-to-know-cool-tshirts	232	9	3.8 %
medium	interview-with-cool-tshirts-founder	184	7	3.8 %
google	cool-tshirts-search	60	2	3.3 %

3.4 End of SQL Capstone Analysis and Thank you

THANK YOU!

Thank you for taking the time to review my analysis of CoolTShirts (a fictional company).
I enjoyed working on this project.