



First/Last touch attribution project

Learn SQL from Scratch
Omar Mustafa

2018-07-13

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`.
- What pages are on their 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?

1. GETTING FAMILIAR WITH COOLTSHIRTS – PART 1

- 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`.
 - Sources are the touchpoints that traffic comes from. These sources are Facebook, Google, Email or Medium.
 - Campaigns are messages that attract certain people's attention to help introduce the website.
 - In a source, like Facebook, a retargeting ad is used. Or in emails, a weekly newsletter campaign helps keep the website in customers' mind.
 - There are six distinct Sources and Campaigns in CoolTShirts.

Number Of Sources
6
Number Of Campaigns
8

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

```
1  /* 1. Three queries will allow us to count and find how source
2     and campaigns are related */
3  -- Count the distinct campaigns
4  SELECT COUNT(DISTINCT utm_campaign) AS 'Number Of Campaigns'
5  FROM page_visits;
6  -- Count the distinct sources
7  SELECT COUNT(DISTINCT utm_source) AS 'Number Of Sources'
8  FROM page_visits;
9  -- find out relation between Source & Campaign
10 SELECT DISTINCT utm_source AS 'Source',
11                  utm_campaign AS 'Campaigns'
12 FROM page_visits;
```

1. GETTING FAMILIAR WITH COOLTSHIRTS – PART 2

- What pages are on CoolTShirts website?
 - The unique values of page_name column are found using DISTINCT query. This returns unique row values in the page_name column.
 - There are four page names. Landing page, shopping cart, checkout and purchase.

```
1  -- 2. get the distinct pages on the website
2  SELECT DISTINCT page_name AS 'CoolTShirts pages'
3  FROM page_visits;
```

Query Results
CoolTShirts pages
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

2. WHAT IS THE USER JOURNEY? – PART 1

- How many FIRST touches is each campaign responsible for?
 - A first touch shows how and when a potential customer is first introduced to the website through a source. And by Using MIN on the timestamp, we can know which source and campaign are responsible for introducing a new customer to the CoolTShirt website.
 - These four campaigns are the most responsible for high first touches for CoolTShirt website out of the eight campaigns.

Campaigns	Number of First Touches
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

```
1 -- 3. The first touches the website has per campaign
2 -- using WITH, a temporary table finds the first touches by each user_id
3 WITH first_touch AS (
4     SELECT user_id,
5            MIN(timestamp) as first_touch_at
6     FROM page_visits
7     GROUP BY user_id),
8 -- next table is used to apply an inner join on first_touch and page_visits tables,
9 -- this adds source and campaign to user_id and timestamp
10 ft_attr AS (
11     SELECT ft.user_id,
12            ft.first_touch_at,
13            pv.utm_source,
14            pv.utm_campaign
15     FROM first_touch AS ft
16     JOIN page_visits AS pv
17     ON ft.user_id = pv.user_id
18     AND ft.first_touch_at = pv.timestamp
19 )
20 -- counting and selecting each row that includes campaigns and ordering it
21 SELECT ft_attr.utm_campaign AS 'Campaigns',
22        COUNT(*) AS 'Number of First Touches'
23 FROM ft_attr
24 GROUP BY 1
25 ORDER BY 2 DESC;
```

2. WHAT IS THE USER JOURNEY? – PART 2

- How many LAST touches is each campaign responsible for?
 - A last touch shows what drove the last user to make a purchase or what was the last page they were on. By using MAX on timestamp, the query returns the maximum date for each user.
 - These eight campaigns represent the last touch for campaigns.

Campaigns	Number of Last Touches
weekly-newsletter	447
retargeting-ad	443
retargeting-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

```
1 -- 4. The last touches the website has per campaign
2 -- using WITH, a temporary table finds the last touches by each user_id
3 WITH last_touch AS (
4     SELECT user_id,
5            MAX(timestamp) as last_touch_at
6     FROM page_visits
7     GROUP BY user_id),
8 -- next table is used to apply an inner join on last_touch and page_visits tables,
9 -- this adds source and campaign to user_id and timestamp
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 AS lt
16     JOIN page_visits AS pv
17     ON lt.user_id = pv.user_id
18     AND lt.last_touch_at = pv.timestamp
19 )
20 -- counting and selecting each row that includes campaigns and ordering it
21 SELECT lt_attr.utm_campaign AS 'Campaigns',
22        COUNT(*) AS 'Number of Last Touches'
23 FROM lt_attr
24 GROUP BY 1
25 ORDER BY 2 DESC;
```

2. WHAT IS THE USER JOURNEY? – PART 3

- How many visitors make a purchase?
 - If we take the number of users who made a purchase and divide it from the total of last touches, we get the percentage of users who end up making a purchase, which is 18% ($361 / 1979 = 18\%$).
 - By adding a WHERE clause to the last touches query, we can only target the purchase page. While using the same MAX timestamp to find the last timestamp for users. The number of last touch effectiveness decrease from one campaign to another.

"4 - purchase" Users

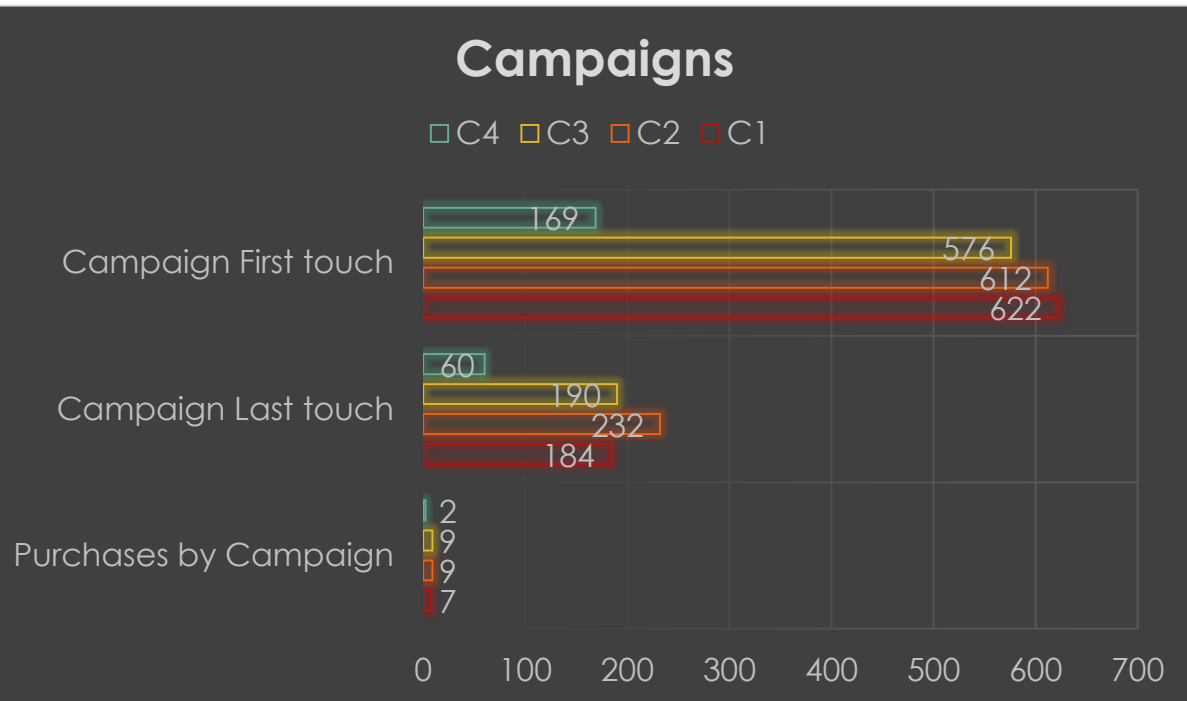
361

Campaigns	No. of Last Touches for purchase page
weekly-newsletter	115
retargeting-ad	113
retargeting-campaign	54
paid-search	52
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	7
cool-tshirts-search	2

```
1 -- 5. selecting counting only the distinct page_name which
  corresponds to: "4 - purchase"
2 SELECT COUNT(DISTINCT user_id) AS "4 - purchase" Users'
3 FROM page_visits
4 WHERE page_name = '4 - purchase';
```

```
1 -- 6. counting the last touches for each campaign in the purchase page
2 -- using WITH, a temporary table finds the last touches by each user_id
3 WITH last_touch AS (
4     SELECT user_id,
5         MAX(timestamp) as last_touch_at
6     FROM page_visits
7     WHERE page_name = '4 - purchase'
8     GROUP BY user_id),
9 -- next table is used to apply an inner join on last_touch and page_visits tables,
10 -- this adds source and campaign to user_id and timestamp
11 lt_attr AS (
12     SELECT lt.user_id,
13         lt.last_touch_at,
14         pv.utm_source,
15         pv.utm_campaign
16     FROM last_touch AS lt
17     JOIN page_visits AS pv
18     ON lt.user_id = pv.user_id
19     AND lt.last_touch_at = pv.timestamp
20 )
21 -- counting and selecting each row that includes campaigns and ordering it
22 SELECT lt_attr.utm_campaign AS 'Campaigns',
23     COUNT(*) AS 'No. of Last Touches for purchase page'
24 FROM lt_attr
25 GROUP BY 1
26 ORDER BY 2 DESC;
```

2. WHAT IS THE TYPICAL USER JOURNEY?



- Certain campaigns were picked to show the typical journey.
- C1: interview-with-cool-tshirt-founder
C2: getting-to-know-cool-tshirt
C3: ten-crazy-cool-tshirt-facts
C4: cool-tshirt-search
- From these campaigns, the number of customers who end up purchasing is very low for these campaigns.

3. OPTIMIZE THE CAMPAIGN BUDGET

- Which 5 campaigns CoolTShirt should re-invest in?
 - There are five top achieving campaigns which should be considered, out of those 5, there are 3 campaigns that generate high first touches which are:
 - Also, the Company should invest more in incorporating the 3 top achieving first touches with the following 2 last touch campaigns that landed customers on the purchase page:

Campaigns	Number of First Touches
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

Campaigns	No. of Last Touches for purchase page
weekly-newsletter	115
retargeting-ad	113
retargeting-campaign	54
paid-search	52