



First- and Last-Touch Attribution with CoolTShirts.com

Learn SQL from Scratch

Julia Clasen

2018/10/13

Table of Contents



1. Get familiar with CoolTShirts
2. What is the user journey?
3. Optimize the campaign budget

1. Get familiar with the company

1.1 How many campaigns and sources does CoolTShirts use and how are they related? explain the difference between `utm_campaign` and `utm_source`

`utm_source` ➤ Identifies the source of traffic. Shows which site sent the traffic (i.e., google, newsletter, or an ad in social networks like facebook)

`utm_campaign` ➤ A specific promotion or a strategic campaign. Identifies the specific ad or email blast (i.e., "winter-sale" or "newsletter-october")

```
1 SELECT COUNT(DISTINCT utm_campaign) AS 'number
   campaigns' FROM page_visits;
2 SELECT COUNT(DISTINCT utm_source) AS 'number
   sources' FROM page_visits;
3 SELECT DISTINCT utm_campaign, utm_source FROM
   page_visits;
```

from 1

number campaigns

8

from 2

number sources

6

from 3

utm_campaign	utm_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

1.2 What pages are on the CoolTShirts website?

- To find the distinct values of the page_name column we use a SELECT DISTINCT query
- There are four different sections on the CoolTShirts website

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

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

2. What is the user journey?

2.1 How many first touches is each campaign responsible for?

- The query defines two temporary tables: the set of **first touches** by the **minimal timestamp** and ft_attr as the same set with source and campaign columns added from 'page visits'. Joined on user_id and time_stamp
- Group by 'source' and 'campaign' order by the value of first touches in a descending order

```
1  WITH first_touch AS (SELECT user_id,
2     MIN(timestamp) as first_touch_at FROM page_visits
3     GROUP BY user_id),
4  ft_attr AS (SELECT ft.user_id,
5     ft.first_touch_at,
6     pv.utm_source,
7     pv.utm_campaign
8  FROM first_touch ft
9  JOIN page_visits pv
10     ON ft.user_id = pv.user_id
11     AND ft.first_touch_at = pv.timestamp)
12  SELECT ft_attr.utm_source AS 'source',
13     ft_attr.utm_campaign AS 'campaign',
14     COUNT(*)
15  FROM ft_attr
16  GROUP BY 1, 2
17  ORDER BY 3 DESC;
```

- The *"interview-with-cool-tshirts-founder"* campaign aroused the greatest interest of the user and caused the most first touches
- only 4 of 8 campaigns are responsible for all first touches

source	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

2.2 How many last touches is each campaign responsible for?

- The query defines two temporary tables: the set of **last touches** by the **maximal timestamp** and lt_attr as the same set with source and campaign columns added from page visits. Joined on user_id and time_stamp
- Group by 'source' and 'campaign' order by the value of last touches in a descending order
- The “*weekly-newsletter*” via email as well as the “*retargeting-ad*” on facebook attract the highest attention of the users and encourage 447 respectively 443 users to re-visit the website
- All eight campaigns are drivers for last touches

```
1  WITH last_touch AS (SELECT user_id, MAX(timestamp)
2  as last_touch_at FROM page_visits GROUP BY
3  user_id),
4  lt_attr AS (SELECT lt.user_id,
5  lt.last_touch_at,
6  pv.utm_source,
7  pv.utm_campaign
8  FROM last_touch lt
9  JOIN page_visits pv
10  ON lt.user_id = pv.user_id
11  AND lt.last_touch_at = pv.timestamp)
12  SELECT lt_attr.utm_source AS 'source',
13  lt_attr.utm_campaign AS 'campaign',
14  COUNT(*)
15  FROM lt_attr
16  GROUP BY 1, 2
17  ORDER BY 3 DESC;
```

source	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

2.3 How many visitors make a purchase?

- To count the visitors who make a purchase → count the distinct user_id's who visited the page named '4 – purchase'
- 361 visitors make a purchase and 1979 users visit the website → 18 % of all visitors make a purchase ($361/1979*100=18.24\%$)

```
1 SELECT COUNT(DISTINCT user_id) AS 'users_purchase'
   FROM page_visits WHERE page_name = '4 - purchase';
2 SELECT COUNT(DISTINCT user_id) AS 'all_users' FROM
   page_visits;
```

users_purchase
361
all_users
1979

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

- Similar to point 2.2 the query defines a temporary table of **last touches**. But here the set of **last touches** is restricted to users who make a purchase. This restriction is done by the additional WHERE clause.
- Again, the campaigns **'weekly-newsletter'** and **'retargetting-ad'** generate most traffic on the website

```
1 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),
2 lt_attr AS (SELECT lt.user_id, lt.last_touch_at,
  pv.utm_source, pv.utm_campaign FROM last_touch lt
3 JOIN page_visits pv ON lt.user_id = pv.user_id AND
  lt.last_touch_at = pv.timestamp)
4 SELECT lt_attr.utm_source AS 'source',
  lt_attr.utm_campaign AS 'campaign', COUNT(*)
5 FROM lt_attr
6 GROUP BY 1, 2 ORDER BY 3 DESC;
```

source	campaign	COUNT(*)
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-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

2.5 What is the typical user journey?

- For the typical user journey it is also interesting to observe the visited pages from CoolTShirts website
- Running the same queries as in 2.1 and 2.2 but adding 'page_name' results in:

1st touch

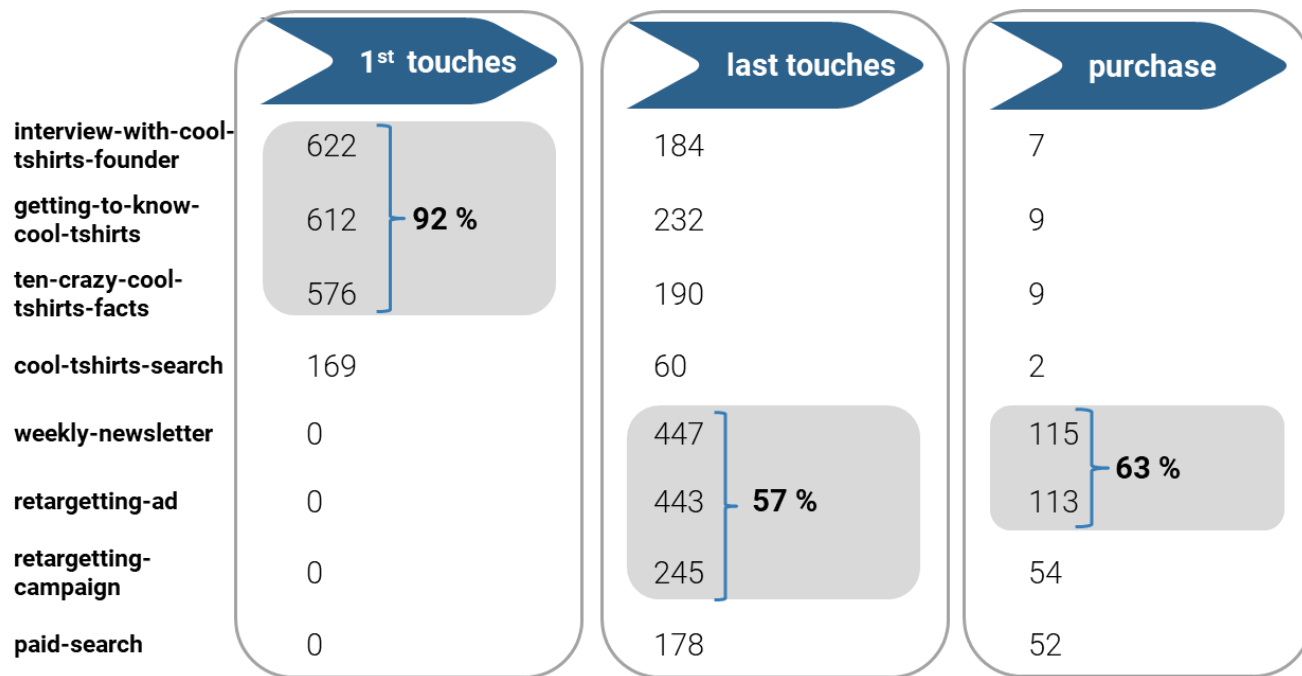
```
1  WITH first_touch AS (SELECT user_id,
2     MIN(timestamp) as first_touch_at FROM
3     page_visits GROUP BY user_id),
4  ft_attr AS (SELECT ft.user_id,
5     ft.first_touch_at,
6     pv.utm_source,
7     pv.utm_campaign,
8     pv.page_name
9  FROM first_touch ft
10 JOIN page_visits pv
11   ON ft.user_id = pv.user_id
12   AND ft.first_touch_at = pv.timestamp)
13 SELECT ft_attr.utm_source AS 'source',
14        ft_attr.utm_campaign AS 'campaign',
15        ft_attr.page_name AS 'page name', COUNT(*)
16 FROM ft_attr
17 GROUP BY 1, 2, 3 ORDER BY 4 DESC;
```

last touch

```
1  WITH last_touch AS (SELECT user_id,
2     MAX(timestamp) as last_touch_at FROM
3     page_visits GROUP BY user_id),
4  lt_attr AS (SELECT lt.user_id,
5     lt.last_touch_at,
6     pv.utm_source,
7     pv.utm_campaign,
8     pv.page_name
9  FROM last_touch lt
10 JOIN page_visits pv
11   ON lt.user_id = pv.user_id
12   AND lt.last_touch_at =
13   pv.timestamp)
14 SELECT lt_attr.utm_source AS 'source',
15        lt_attr.utm_campaign AS 'campaign',
16        lt_attr.page_name AS 'page name', COUNT(*)
17 FROM lt_attr
18 GROUP BY 1, 2, 3 ORDER BY 4 DESC;
```

2.5 What is the typical user journey?

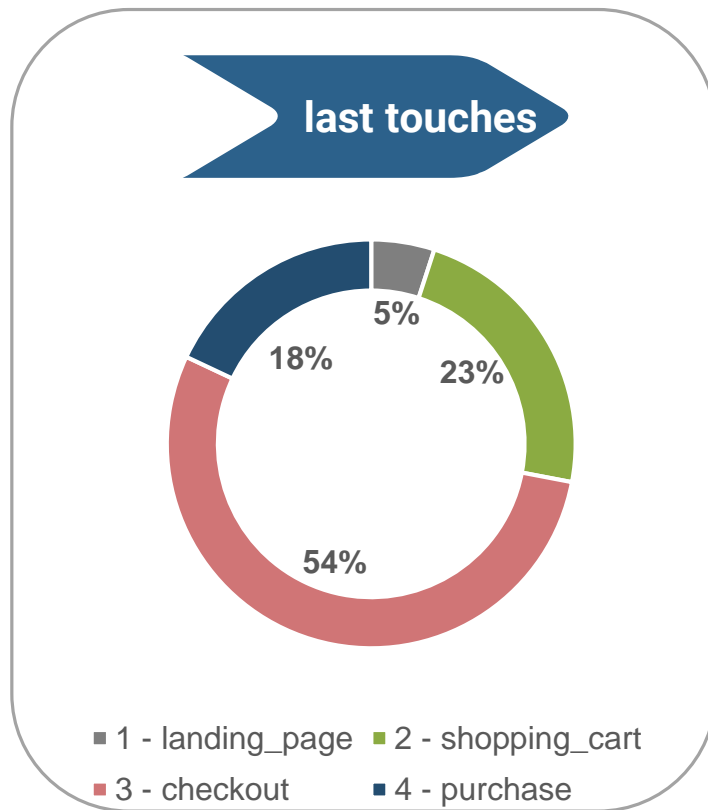
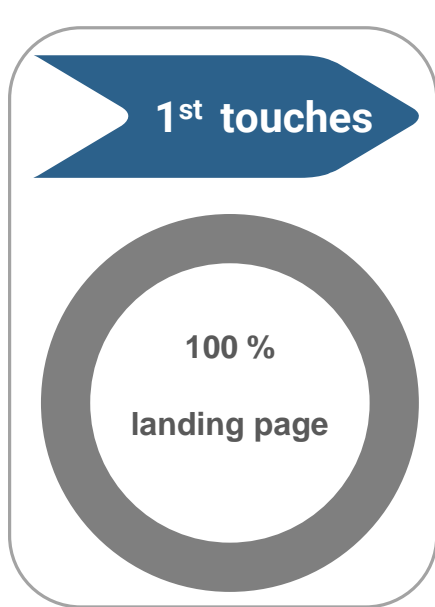
customers' journey: from initial visit to purchase → which campaigns drive most traffic and purchases?



- About 92% of the 1st touches come from 'story' campaigns; these campaigns are the strongest lever to increase website visits and lead to 31% (606) of last touches, but cause only 6% (25) of the purchases
- Most of the prospects need a reminder. So, 57% of the last touches come through newsletter and retargeting campaigns and lead to 63 % of the purchases

2.5 What is the typical user journey?

customers' journey: from initial visit to purchase → breakdown of the visited pages



- All visitors are directed to the landing page with their first touch, but do not directly convert to the purchase page
- The majority of visitors (54%) check out by their last visit, 23% of the last touches are directed to the shopping cart and 18% of last touches visit the purchase page

3. Optimize the campaign budget

3.1 CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

- The story telling campaigns (**interview-with-cool-tshirts-founder**, **getting-to-know-cool-tshirts**, **ten-crazy-cool-tshirts-facts**) are most important to acquire website visitors (92% of first touches)
- In order to remain the visitors' interest, the reminder campaigns (**weekly-newsletter**, **retargeting-ad**, **retargeting-campaign**) are the most effective ones. The **weekly-newsletter** and the **retargeting-ad** campaigns have the highest conversion rate to purchases (63% of purchases)
- Recommendation is to re-invest into **interview-with-cool-tshirts-founder**, **getting-to-know-cool-tshirts**, **ten-crazy-cool-tshirts-facts** to acquire a relevant amount of website visitors and to continue with the **weekly-newsletter** and the **retargeting-ad** campaigns to generate the most purchases

Type	Campaign	1st (#)	1st (%)	Last (#)	Last (%)	Pur (#)	Pur (%)	Comment
Story / Content	interview-with-cool-tshirts-founder	622	31,4	184	9,3	7	1,9	acquisition Campaign
Story / Content	getting-to-know-cool-tshirts	612	30,9	232	11,7	9	2,5	acquisition Campaign
Story / Content	ten-crazy-cool-tshirts-facts	576	29,1	190	9,6	9	2,5	acquisition Campaign
Active Search	cool-tshirts-search	169	8,5	60	3,0	2	0,6	low conversion rate
"Reminder"	weekly-newsletter	0	0,0	447	22,6	115	31,9	Conversion Campaign: Free Channel with 32% Conversion
"Reminder"	retargeting-ad	0	0,0	443	22,4	113	31,3	Conversion Campaign: Paid Channel with 31% Conversion
"Reminder"	retargeting-campaign	0	0,0	245	12,4	54	15,0	
Active Search (Paid)	paid-search	0	0,0	178	9,0	52	14,4	
Total #		1979		1979		361		
Total %						18,2		

BACKUP

2.5 What is the typical user journey?

- The query from 2.5 gives the following results:

source	campaign	page name	COUNT(*)
email	weekly-newsletter	3 - checkout	333
facebook	retargeting-ad	3 - checkout	331
email	retargeting-campaign	3 - checkout	192
nytimes	getting-to-know-cool-tshirts	2 - shopping_cart	153
medium	interview-with-cool-tshirts-founder	2 - shopping_cart	132
buzzfeed	ten-crazy-cool-tshirts-facts	2 - shopping_cart	126
google	paid-search	3 - checkout	126
email	weekly-newsletter	4 - purchase	114
facebook	retargeting-ad	4 - purchase	112
email	retargeting-campaign	4 - purchase	53
google	paid-search	4 - purchase	52
google	cool-tshirts-search	2 - shopping_cart	46
nytimes	getting-to-know-cool-tshirts	1 - landing_page	38
buzzfeed	ten-crazy-cool-tshirts-facts	1 - landing_page	32
nytimes	getting-to-know-cool-tshirts	3 - checkout	32
medium	interview-with-cool-tshirts-founder	3 - checkout	24
buzzfeed	ten-crazy-cool-tshirts-facts	3 - checkout	23
medium	interview-with-cool-tshirts-founder	1 - landing_page	21
buzzfeed	ten-crazy-cool-tshirts-facts	4 - purchase	9
nytimes	getting-to-know-cool-tshirts	4 - purchase	9
google	cool-tshirts-search	1 - landing_page	7
medium	interview-with-cool-tshirts-founder	4 - purchase	7
google	cool-tshirts-search	3 - checkout	5
google	cool-tshirts-search	4 - purchase	2