code|cademy

First- and Last- Touch Attribution

Learn SQL from Scratch
Merlin Sanders
October 4, 2018

Table of Contents

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

1. Get Familiar with CoolTShirts

1.1 Get familiar with the company.

CoolTshirts tracks users activity by date/time a page was viewed. Each user is assigned a user-id and each visit is attributed to an advertising campaign and source.

page_name	timestamp	user_id	utm_campaign	utm_source
1 - landing_page	2018-01-24 03:12:16	10006	getting-to-know-cool-tshirts	nytimes
2 - shopping_cart	2018-01-24 04:04:16	10006	getting-to-know-cool-tshirts	nytimes
3 - checkout	2018-01-25 23:10:16	10006	weekly-newsletter	email
1 - landing_page	2018-01-25 20:32:02	10030	ten-crazy-cool-tshirts-facts	buzzfeed
2 - shopping_cart	2018-01-25 23:05:02	10030	ten-crazy-cool-tshirts-facts	buzzfeed
3 - checkout	2018-01-28 13:26:02	10030	retargetting-campaign	email
4 - purchase	2018-01-28 13:38:02	10030	retargetting-campaign	email

1.1 Advertising Campaigns

CoolTshirts has eight advertising campaigns across six sources.
- Some sources are used for multiple campaigns
- The utm_campaign is the advertising campaign attributed to the users visit.
- The utm_source is the advertising source attributed to the users visit.

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

```
/* Distinct Campaigns */
SELECT COUNT(DISTINCT utm_campaign)
FROM page_visits;

/* Distinct Sources */
SELECT COUNT(DISTINCT utm_source)
FROM page_visits;

/* Distinct Campaign/Source Combinations */
SELECT DISTINCT utm_campaign, utm_source
FROM page_visits
GROUP BY utm_campaign, utm_source;
```

1.2 Website Pages

The website is comprised of four pages.

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

```
/* Distinct Website Pages */
SELECT DISTINCT page_name
FROM page_visits;
```

2. What is the User Journey

2.1 First Touches

Four advertising campaigns account for users first touch.

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

```
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT pv.utm_campaign,
    COUNT(*) AS first_touch
FROM first_touch ft
JOIN page_visits pv
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
    GROUP BY 1
    ORDER BY 2 DESC;
```

2.2 Last Touches

Eight advertising campaigns account for users last touch.

utm_campaign	last_touch
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

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT pv.utm_campaign,
    COUNT(*) AS last_touch
FROM last_touch lt
JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
    GROUP BY 1
    ORDER BY 2 DESC;
```

2.3 Number of Purchases

How many visitors make a purchase?

num_purchase
361

```
SELECT COUNT(DISTINCT user_id) AS num_purchase
FROM page_visits
WHERE page_name = '4 - purchase'
GROUP BY page_name;
```

2.4 Last Touches on the Purchase Page

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

utm_campaign	last_touch
weekly-newsletter	115
retargetting-ad	113
retargetting-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

```
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)
SELECT pv.utm_campaign,
    COUNT(*) AS last_touch
FROM last_touch lt
JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
    GROUP BY 1
    ORDER BY 2 DESC;
```

2.5 Typical Journey

What is the typical user journey?

o The majority of visitors add items to the shopping cart
o Approximately seventy-five percent leave before checkout
o Twenty-five percent complete their purchase

```
SELECT page_name, COUNT(page_name)
FROM page_visits
GROUP BY 1 ORDER BY 2 DESC;
```

page_name	visits
1 - landing_page	2000
2 - shopping_cart	1900
3 - checkout	1431
4 - purchase	361

3. Optimize the Campaign Budget

3.1 What campaigns should CoolTshirts re-invest in?

The following campaigns should be the focus of CoolTshirts advertising. These campaigns account for the majority of page visits and purchases.

CoolTshirts should examine the checkout process. A large percentage of users abandon the checkout before completing their purchase.

utm_campaign
interview-with-cool-tshirts-founder
getting-to-know-cool-tshirts
ten-crazy-cool-tshirts-facts
weekly-newsletter
retargetting-ad