



Marketing Attribution – CoolTShirts

Analyze Data with SQL
Athena Mascarenhas
September 14th, 2024

Table of Contents

1. Project Description
2. Company Information
3. User Journey
4. Conclusion

1. Project Description

1. Project Description

CoolTShirts sells shirts of all kinds, as long as they are T-shaped and cool. Recently, CTS started a few marketing campaigns to increase website visits and purchases. Using market attribution, they'd like to map their customers' journey: from initial visit to purchase. They can use that information to optimize their marketing campaigns. This project will guide you through some of that process. We will use the `page_visits` table (information below).

page_visits	
A table describing each time a user visits the CoolTShirts website	
Column	Description
<code>user_id</code>	A unique identifier for each visitor to a page
<code>timestamp</code>	The time at which the visitor came to the page
<code>page_name</code>	The title of the section of the page that was visited
<code>utm_source</code>	Identifies which site sent the traffic (i.e., google, newsletter, or facebook_ad)
<code>utm_campaign</code>	Identifies the specific ad or email blast (i.e., june-21-newsletter or memorial-day-sale)

2. Company Information

2.1 Company Information

CoolTShirts, an innovative apparel shop, is running a bunch of marketing campaigns.

Here is the snippets of the data:

```
SELECT * FROM page_visits LIMIT 10;
```

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	retargeting-campaign	email
4 - purchase	2018-01-28 13:38:02	10030	retargeting-campaign	email
1 - landing_page	2018-01-05 18:31:17	10045	getting-to-know-cool-tshirts	nytimes
2 - shopping_cart	2018-01-05 21:16:17	10045	getting-to-know-cool-tshirts	nytimes
3 - checkout	2018-01-09 03:05:17	10045	retargeting-ad	facebook

2.2 Company Information: Distinct Campaigns and Sources

- There are 8 unique ad campaigns:

```
SELECT DISTINCT utm_campaign  
from page_visits;
```

utm_campaign
getting-to-know-cool-tshirts
weekly-newsletter
ten-crazy-cool-tshirts-facts
retargeting-campaign
retargeting-ad
interview-with-cool-tshirts-founder
paid-search
cool-tshirts-search

- There are 6 unique sources:

```
SELECT DISTINCT utm_source  
from page_visits;
```

utm_source
nytimes
email
buzzfeed
facebook
medium
google

2.2 Company Information: Distinct Campaigns and Sources

- There are 8 distinct ad campaigns on 6 different sources
 - Email was the source of 2 different ad campaigns
 - Google was the source of 2 different ad campaigns
 - All other sources have one unique ad campaign associated with that source.

```
SELECT DISTINCT utm_campaign, utm_source  
from page_visits;
```

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

2.3 Company Information: Pages

- There are 4 pages on the CoolTShirts website
:

```
SELECT DISTINCT page_name  
from page_visits;
```

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

3. User Journey

3.1 User Journey: First-Touch campaigns

- Users initially discover CoolTShirts through 4 out of the 8 sources/campaigns (first-touches).
- Medium's interview-with-cool-tshirts-founder ad campaign, Nytimes' getting-to-know-cool-tshirts ad campaign and BuzzFeed's ten-crazy-cool-tshirts-facts bring in most of the users.
- Google's cool-tshirts-search ad campaign brings in the least number of users.

utm_source	utm_campaign	users	proportion_of_users
medium	interview-with-cool-tshirts-founder	622	0.31
nytimes	getting-to-know-cool-tshirts	612	0.31
buzzfeed	ten-crazy-cool-tshirts-facts	576	0.29
google	cool-tshirts-search	169	0.09

```
WITH first_touch AS (  
    SELECT user_id,  
           MIN(timestamp) as first_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
total_users AS (  
    SELECT COUNT(DISTINCT user_id) AS total_users from  
    page_visits),  
ft_counts AS (  
    SELECT pv.utm_source,  
           pv.utm_campaign,  
           total_users,  
           COUNT(*) as users  
    FROM first_touch ft  
    JOIN page_visits pv  
      ON ft.user_id = pv.user_id  
      AND ft.first_touch_at = pv.timestamp  
    CROSS JOIN total_users tu  
    GROUP BY 1,2  
    ORDER BY 3 DESC)  
SELECT utm_source,  
       utm_campaign,  
       users,  
       ROUND(1.0 * users/total_users,2) as  
       proportion_of_users  
FROM ft_counts  
ORDER BY 3 DESC;
```

3.2 User Journey: Last-Touch campaigns

- All 8 sources/campaigns result in the last-touches.
- Email's weekly-newsletter and facebook's retargeting-ad results in the largest number of last-touches.
- Google's cool-tshirts-search ad campaign results in the lowest number of last-touches (3%).
- All other campaigns fall in between.

utm_source	utm_campaign	users	proportion_of_users
email	weekly-newsletter	447	0.23
facebook	retargeting-ad	443	0.22
email	retargeting-campaign	245	0.12
nytimes	getting-to-know-cool-tshirts	232	0.12
buzzfeed	ten-crazy-cool-tshirts-facts	190	0.1
medium	interview-with-cool-tshirts-founder	184	0.09
google	paid-search	178	0.09
google	cool-tshirts-search	60	0.03

```
WITH last_touch AS (  
    SELECT user_id,  
           MAX(timestamp) as last_touch_at  
    FROM page_visits  
    GROUP BY user_id),  
total_users AS (  
    SELECT COUNT(DISTINCT user_id) AS total_users from  
           page_visits),  
lt_counts AS (  
    SELECT  
        pv.utm_source,  
        pv.utm_campaign,  
        tu.total_users,  
        COUNT(*) as users  
    FROM last_touch lt  
    JOIN page_visits pv  
        ON lt.user_id = pv.user_id  
        AND lt.last_touch_at = pv.timestamp  
    CROSS JOIN total_users tu  
    GROUP BY 1,2  
    ORDER BY 3 DESC)  
SELECT utm_source,  
       utm_campaign,  
       users,  
       ROUND(1.0 * users/total_users,2) as  
proportion_of_users  
FROM lt_counts  
ORDER BY 3 DESC;
```

3.3 User Journey: Purchases

```
--purchased users
WITH total_users AS (
  SELECT COUNT(DISTINCT user_id) AS total_users from
  page_visits),
purchased AS (
  SELECT COUNT(DISTINCT user_id) AS users_that_purchased
  from page_visits
  where page_name = '4 - purchase')
SELECT total_users,
       users_that_purchased,
       ROUND(1.0 * users_that_purchased/total_users,2) as
proportion_of_users_that_purchased
FROM purchased p
CROSS JOIN total_users;
```

total_users	users_that_purchased	proportion_of_users_that_purchased
1979	361	0.18

361 visitors (18%) made a purchase.

3.4 User Journey: Campaigns that led to purchases

- There were purchases made on all 8 sources/campaigns
- Email's weekly-newsletter and facebook's retargeting-ad had the highest number of purchases
- Google's cool-tshirts-search ad campaign has the lowest numbers
- All other campaigns fall in between.

utm_source	utm_campaign	users_purchased	proportion_of_users_purchased
email	weekly-newsletter	115	0.32
facebook	retargeting-ad	113	0.31
email	retargeting-campaign	54	0.15
google	paid-search	52	0.14
buzzfeed	ten-crazy-cool-tshirts-facts	9	0.02
nytimes	getting-to-know-cool-tshirts	9	0.02
medium	interview-with-cool-tshirts-founder	7	0.02
google	cool-tshirts-search	2	0.01

```
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),  
total_users AS (  
    SELECT COUNT(DISTINCT user_id) AS total_users  
    from    page_visits  
    where page_name = '4 - purchase'),  
lt_counts AS (  
    SELECT  
        pv.utm_source,  
        pv.utm_campaign,  
        tu.total_users,  
        COUNT(*) as users  
    FROM last_touch lt  
    JOIN page_visits pv  
        ON lt.user_id = pv.user_id  
        AND lt.last_touch_at = pv.timestamp  
    CROSS JOIN total_users tu  
        GROUP BY 1,2  
        ORDER BY 3 DESC)  
SELECT utm_source,  
       utm_campaign,  
       users as users_purchased,  
       ROUND(1.0 * users/total_users,2) as  
       proportion_of_users_purchased  
FROM lt_counts  
ORDER BY 3 DESC;
```

3.5 User Journey: Page Funnel

```
SELECT page_name,  
       COUNT(DISTINCT user_id) as users  
FROM page_visits  
GROUP BY 1;
```

utm_source	utm_campaign	percentage
1 - landing_page	1979	100%
2 - shopping_cart	1881	95%
3 - checkout	1431	76%
4 - purchase	361	25%

- 95% of users place items in their shopping cart
- 76% checkout, but only 25% purchase – Should be explored further

4. Conclusion

4. Conclusion

The five campaigns that CoolTShirts should re-invest in are the following:

- ❑ Email's weekly-newsletter:
 - Even though it is not one of the first-touch campaigns, it resulted in the highest number of purchases.
- ❑ Facebook's retargeting-ad:
 - Like the email's weekly-newsletter campaign, it is not one of the first-touch campaigns, but it resulted in a high number of purchases. The number of purchases were very close to the number of purchases from email's weekly-newsletter.
- ❑ Medium's interview-with-cool-tshirts-founder:
 - Brought in the highest number of users even though few people purchase items through the campaign.
- ❑ NyTimes's getting-to-know-cool-tshirts:
 - Close second to the medium, this campaign brought in the second highest number of users
- ❑ BuzzFeed's ten-crazy-cool-tshirts-facts:
 - Also, brought in a high number of users and worth re-investing in.

Medium, NyTimes and BuzzFeed are all first touch campaigns but are worth reinvesting in as it brings in users that can be targeted by the last-touch campaigns and could result in purchases

Further Analysis:

- Explore why 95% of visitors add to their shopping cart, but 76% checkout and even fewer buy