codecademy

Capstone: Attribution

Learn SQL from Scratch
Jorge Valdez
8/30/2018

Table of Contents

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

Get Familiar with CoolTShirts

1.1 CoolTShirts Campaigns

```
SELECT COUNT(DISTINCT utm_campaign)
FROM page_visits;

SELECT COUNT(DISTINCT utm_source)
FROM page_visits;

SELECT DISTINCT utm_campaign,
               utm_source
FROM page_visits;
```

- How many campaigns and sources does CoolTShirts use?
- 8 distinct campaigns are being utilized by CoolTShirts.
- 6 distinct sources are used in conjunction with the campaigns to show which touchpoint was used to send the traffic.

- Which source is used for each campaign?
- The following chart shows the relationship between the campaign and the source that was used to attract customers to the CoolTShirts site.

Campaigns	Campaigns (cont'd)	Sources	Sources (cont'd)
getting-to-know-cool-tshirts	retargetting-ad	nytimes	facebook
weekly-newsletter	interview-with-cool-tshirts-founder	email	medium
ten-crazy-cool-tshirts-facts	paid-search	buzzfeed	google
retargetting-campaign	cool-tshirts-search	email	google

1.2 CoolTShirts Campaigns

- What pages are on their website?
- The following pages are listed in the chart below:

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

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

What is the User Journey?

2.1 User Journey

The results from the following query shows how many first touches each campaign is responsible for in the chart below.

As you can see the interview with the CoolTShirts founder was the most popular campaign to attract visitors to the site.

Initial Visit 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

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

2.2 User Journey

The results from the following query shows how many last touches each campaign is responsible for in the chart below.

As you can see the weekly newsletter which is sent via email next to Facebook is popular for drawing visitors back to the site.

Return Visit Source	Campaign	Count
email	weekly-newsletter	447
facebook	retargetting-ad	443
email	retargetting-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_attr AS (
  SELECT lt.user_id,
        lt.last_touch_at,
        pv.utm_source,
        pv.utm_campaign
  FROM last_touch lt
  JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
       lt_attr.utm_campaign,
       COUNT(*)
FROM lt_attr
GROUP BY utm_campaign, utm_source
ORDER BY COUNT(*) DESC;
```

2.3 User Journey

- How many visitors make a purchase?
- 361 distinct visitors made purchases.

Page name	Distinct Users
4 – Purchase	361

```
SELECT page_name, COUNT (DISTINCT user_id)
FROM page_visits
WHERE page_name = '4 - purchase';
```

2.4 User Journey

The results from the following query shows how many last touches on the purchase page each campaign is responsible for in the chart below.

As you can see the weekly newsletter next to Facebook is popular for drawing visitors back to the site and making a purchase.

Return Visit Source	Campaign	Purchases
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-campaign	54
google	paid-search	52
nytimes	getting-to-know-cool-tshirts	9
buzzfeed	ten-crazy-cool-tshirts-facts	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_attr AS (
  SELECT lt.user_id,
        lt.last_touch_at,
        pv.utm_source,
        pv.utm_campaign,
        pv.page_name
  FROM last_touch lt
  JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
      lt_attr.utm_campaign,
      COUNT(*)
FROM lt_attr
GROUP BY utm_campaign, utm_source
ORDER BY COUNT(*) DESC;
```

2.5 User Journey

Typical User Journey
- The interview with the founder of CoolTShirts along with articles from Buzzfeed and NYTimes were responsible for the intial amount of traffic towards the website.
- Emails from their weekly newsletter as well as Facebook ads brought return visits to the site which also led towards the highest amount of traffic that resulted in purchases made.

Optimizing the Campaign Budget

3.1 Optimizing the Campaign Budget

Campaign Re-Investment
- The first touch attribution is a good way to show how visitors initially find out about the company and visit the website.
- Based on this, it would be necessary to continue re-invest in interviews with the Media, NYTimes and Buzzfeed. This has resulted in the most initial traffic to the website.
- Sign-ups for the weekly newsletter follow the initial site visit which shows to be a popular campaign for return visits as well as visits ending with purchases.
- The Facebook ad campaign has also resulted in return traffic with users also committing towards purchases.

Initial Visit Source	Campaign	Count
medium	interview-with-cool-tshirts-founder	622
nytimes	getting-to-know-cool-tshirts	612
buzzfeed	ten-crazy-cool-tshirts-facts	576

Return Visit Source	Campaign	Purchases
email	weekly-newsletter	115
facebook	retargetting-ad	113