code|cademy

Attribution Queries

Learn SQL from Scratch
Andrew Brown
June 4, 2019

Project Table of Contents

1. Campaign and Data Review
2. What Does the Data Tell Us?
3. Campaign Budget Proposal

1. Campaign and Data Review

Recent Campaigns Used to Drive New Sales

- Recently the Marketing department has implemented new marketing campaigns to bring in new sales to CoolTShirts.
- There was a total of eight campaigns used in conjunction with 6 different source websites.
- As people clicked on the links to these sources they where assigned a UTM parameter.
- These UTM parameters are used to capture how and when a customer gets to CoolTShirts.

Campaign	Source
Getting to Know CoolTShirts	New York Times
Interview with CoolTShirts Co-Founder	Medium
Ten Crazy Cool T-shirt Facts	Buzzfeed
Retargeting Ad	FaceBook
Retargeting Campaign	Email
Weekly Newsletter	Email
Paid Search	Google
CoolTShirts Search	Google

Data Tables Created by Website Visits

The UTM parameters are placed in a table where they can be tracked and categorized. These parameters along with additional data tell a story on how each person finds the website.
- There are two important data points that contain the UTM parameters, utm_campaign and utm_source
- utm_campaign identifies the specific ad or email that brought the customer to the website.
- utm_source identifies which site the customer came from.
- The table below shows how the data is sorted.

user_id	timestamp	page_name	utm_source	utm_campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign

Website Mapping

- To track how effective each campaign is, the customer is tracked by their location within the CoolTShirts website.
- By tracking the users with a unique ID, CoolTShirts can see how far into the t-shirt buying process the customer gets.
- The website is broken up into four categories that are in the table on the right.
- The table below shows how far the each customer gets in the buying procress for each of the campaigns and sources.

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

user_id	timestamp	page_name	utm_source	utm_campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign
ID number	Date	website location	source	campaign

2. What Does the Data Tell Us?

The Primary Source of the Website First Touches

To see how each campaign brought new customers to the CoolTShirts webpage the SQL on the right was run to collect the data
- The Medium interview closing beat out the New York Time's article for bringing in the most customer traffic.
- The good old fashion Google search brought up the rear for first touches.
- Articles from large publishers appear to be effective at driving traffic to CoolTShirts.

```
/* The following SQL query was used to identify first
touches. */

WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT ft.user_id,
    ft.first_touch_at,
    pv.utm_source,
        pv.utm_campaign,
        COUNT(utm_campaign)
FROM first_touch AS ft
JOIN page_visits AS pv
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
GROUP BY 4
ORDER BY 5 DESC;
```

utm_source	utm_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

The Primary Source of the Website Last Touches

The first time a Customer visits CoolTShirts sometimes doesn't lead to a purchase. So the query on the right was created to see what brought them back to the website.

- The weekly newsletter did the best job at having customers return to the website closely followed by FaceBook ads.
- The Google search was the least productive at bring users back, this is to be expected as it's passive last touch with no marketing engagement.

utm_source	utm_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

```
/* The following SQL query was used to identify last
touches. */


WITH last_touch AS (
    SELECT user_id,
        Max(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id)
SELECT lt.user_id,
    lt.last_touch_at,
    pv.utm_source,
        pv.utm_campaign,
        COUNT(utm_campaign)
FROM last_touch AS lt
JOIN page_visits AS pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
GROUP BY 4
ORDER BY 5 DESC;
```

The Total Number of Purchases From the Current Campaigns

The final page on the page names shows that the customer has made a purchase. Please see chart below.
- The query on the right was used to determine the total number of purchases made by checking the total number of DISTINCT user_id's to reach the page_name '4 – purchase'.
- The total purchase count from all the campaigns was 361.

```
/* The following SQL query was used to identify total
number of purchases. */

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

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

The Campaign With the Most Purchases From a Last Touch

The first time a Customer visits CoolTShirts sometimes doesn't lead to a purchase. So the query on the right was created to see what brought them back to the website to purchase a t-shirt.
- The weekly newsletter did the best job at having customers return to the website closely followed by FaceBook ads to purchase a t-shirt.
- The Google search was the least productive at bringing users back, this is to be expected as it's passive last touch with no marketing engagement.

```
/* The following SQL query was used to identify last
touches. */


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 lt.user_id,
    lt.last_touch_at,
    pv.utm_source,
        pv.utm_campaign,
        COUNT(utm_campaign)
FROM last_touch AS lt
JOIN page_visits AS pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
GROUP BY 4
ORDER BY 5 DESC;
```

utm_source	utm_campaign	count
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

The Typical User Experince When Visiting CoolTShirts

The typical user journey through CoolTShirts can be broken down by each step in the t-shirt buying process:
- The majority of customers that end up on the landing page come from Medium.com.
- A close second for customers ending up on the landing page is nytimes.com.
- The New York Times starts to pull away with most users from that source adding a t-shirt to the shopping cart.
- However the majority of customers that make it the checkout page come from the company newsletter and a targeted Facebook ad .
- The difference between the newsletter and Facebook is a small margin .
- This trend continues for the purchase page, where FaceBook and the Newsletter are getting the majority of customers to purchase a t-shirt .
- Interestingly of the large number of customers that come from Medium.com only 1% ended up buying t-shirts.

3. Campaign Budget Proposal

Reviewing Presented Material

The Charts below are the first touch, last touch, and purchase.
- The news source campaigns do an excellent job of bring customers to the website, however, they do not lead to many purchases.
- Once the Customer has visited the website social media ads and email give them the nudge to purchase a t-shirt.

utm_source	utm_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

utm_source	utm_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

utm_source	utm_campaign	count
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

Campaign Recommendations

- The upcoming budget only allows for 5 of the 8 campaigns to continue.
- The 5 recommendations can be found on the graph on the right.
- The Facebook re-targeting ads where effective in bringing customers back to the website after their intial visit. Bundled with the news source campaign it will be effective in bringing customers back.
- The Newsletter had the largest number of follow up purchases, so a continued effort is needed in order to promt customers to give their email address for direct advertising.
- The Google paid searchs (while having lower first and last touches) were 29% effective in getting a customer to buy a t-shirt.
- While Buzzfeed and the New York Times customer purchases where low, they pushed the largest amount of traffic to the website.
- The large number of visitors from the news sources allows some of the other campaigns to work their 'Magic'.

Source	Campaign
Facebook	Retargeting Ad
Email	Newsletter
Google	Paid Search
Buzzfeed	ten-crazy-cool-tshirts-facts
New York Times	getting-to-know-cool-tshirts

Questions?