



CoolTShirts Marketing Analysis

Touch Attribution Queries

Learn SQL from Scratch Capstone

James G. Seltzer

2019.05.02

Introduction to Presentation Content

1) Get Familiar With CoolTShirts

- Current Campaign Strategies and Sources

2) What Is The User Journey?

- Campaign Effectiveness and Supporting Data

3) Optimize The Campaign Budget

- Recommendations For Re-Investment

1. Get Familiar With CoolTShirts

1.1 CoolTShirts Campaigns & Sources

To better understand how CoolTShirts (CTS) attracts customers to their website we must analyze information from their marketing efforts. This is possible through UTM parameters! These parameters are contained in the links from various ad campaigns that, when clicked, capture and store details of their customers' online visits in the CTS database.

The CTS database table “page_visits” contains several columns that can be used to store and retrieve data secured from these UTM parameters. Two of which are the “utm_campaign” and “utm_source” columns.

- The column “utm_campaign” is designed to identify the specific ad or email blast that led a particular user to the site. CTS has *eight* unique campaigns.
- The column “utm_source” serves to identify which site directed the traffic. It contains *six* different sources.
- Visits to CTS from these sources (also referred to as touchpoints) can then be attributed to those ad campaigns!

The CTS database not only maintains which ads and blog posts lead customers to their site but it also tracks user activity and navigation throughout the CTS site.

- The column “page_name” describes the title of the section of the page that CTS users visit. The CTS website has *four* distinct pages that can be seen in the table to the right.

CTS Pages

1 - landing_page

2 - shopping_cart

3 - checkout

4 - purchase

1.2 CoolTShirts Campaigns & Sources Continued

The four SQL queries below provide us with information about CTS. The first query supplies us with the four pages of the CTS website (see previous slide). The second and third queries supply us with the number of distinct campaigns (8) and sources (6). The last query shows us the relationship between these sources and campaigns and is illustrated in the table to the right.

```
SELECT DISTINCT page_name AS 'CTS Pages'
FROM page_visits;

SELECT COUNT(DISTINCT utm_campaign) AS 'CTS
Campaigns'
FROM page_visits;

SELECT COUNT(DISTINCT utm_source) AS 'CTS
Sources'
FROM page_visits;

SELECT DISTINCT utm_source AS 'Source',
utm_campaign AS 'Campaign'
FROM page_visits;
```

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

2. What Is The User Journey?

2.1 First Touches Per Campaign

The query to the right returns the results below, indicating how many *first* touches (first time the user is exposed to the CTS site) each campaign is responsible for. From this information we can determine that the campaign “interview-with-cool-tshirts-founder” appears to be the most successful in netting a customers initial visit. The last two commands in the query present the data in a meaningful way:

The “GROUP BY 1, 2” statement arranges the requested data (sources and campaigns) into the first two columns. The “ORDER BY 3 DESC” statement sorts the requested count by the highest number of first touches in descending order. The “cool-tshirts-search” appears to have the lowest number of 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

```
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 AS  
'Source',  
           ft_attr.utm_campaign AS  
'Campaign',  
           COUNT(*)  
    FROM ft_attr  
    GROUP BY 1, 2  
    ORDER BY 3 DESC;
```

2.2 Last Touches Per Campaign

The query on the right provides us with the results below, representing how many *last* touches (exposure to CTS that led to a purchase) each campaign is responsible for.

- From evaluating this data we can see that the “weekly-newsletter” campaign has the highest number of last touches while the “cool-tshirts-search” campaign has the fewest.

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

```
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,  
           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  
AS 'Source',  
           lt_attr.utm_campaign  
AS 'Campaign',  
           COUNT(*)  
    FROM lt_attr  
    GROUP BY 1, 2  
    ORDER BY 3 DESC;
```


2.3 Purchase Data & The Typical User Journey

With the below query we are able to determine how many visitors to CTS actually commit to making a purchase! *The query returns a count of 361 users.*

- This is possible by counting the number of individual users (user_id) that visited the purchase page (page_name = '4 – purchase').
- **A user only reaches this page after a purchase has been made*

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

By adding an additional “where” clause to our previous subquery for last touches (slide 8) we can find out how many last touches on the purchase page each campaign is responsible for.

- See the where clause added to the (last_touch) subquery and the result set to the right

```
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),
```

Source	Campaign	COUNT (*)
email	weekly-newsletter	115
facebook	retargeting-ad	113
email	retargeting-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

2.4 Purchase Data & The Typical User Journey Continued

To better assess when customers follow through in purchasing merchandise as opposed to navigating away from the CTS website, we can run SQL queries to research user progression through the pages of the site and ultimately measure the success of the various ad campaigns.

The query below produces five of the user_id's assigned to individuals who visited CTS and purchased merchandise off of the website.

- See the result set labeled “CTS Customers” to the right

```
SELECT DISTINCT user_id AS 'CTS Customers'  
FROM page_visits  
WHERE page_name = '4 - purchase'  
LIMIT 5;
```

CTS Customers
10030
10069
10162
10329
10354

2.5 Purchase Data & The Typical User Journey Continued

If we choose a user_id (10030 in example) from the table on the last slide and run the query below we can see that customer's visit history in the result set!

- After reviewing the information on several CTS customers' visit history we can determine that the typical user's journey consists of:

- 1) Appearing on the CTS "1 – landing_page"
- 2) Selecting merchandise on the "2 – shopping_cart" page
- 3) Reviewing and placing an order in "3 – checkout"
- 4) Being directed to the "4 – purchase" page

```
SELECT *  
FROM page_visits  
WHERE user_id = 10030;
```

page_name	timestamp	user_id	utm_campaign	utm_source
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

3. Optimize The Campaign Budget

3.1 CoolTShirts Budget Optimization

The queries from our previous slides present us with some very useful information. We can visualize and explore the schema of the CTS database table “page_visits”, we can identify and quantify users that made purchases to research their online journey, and more importantly we can evaluate the effectiveness of the CoolTShirts’ ad campaigns.

While our first-touch attribution results only consider the initial utm_source for each customer and demonstrate how the user discovered the website, our last-touch attribution results are limited to considering the last utm_source per customer but reveal insight on their final purchase.

Therefore both must be evaluated to make an informed decision on where to re-invest funds.

3.2 CoolTShirts Budget Optimization Continued

With the funds to re-invest in five campaigns CTS should focus their energy into the following:

- 1) The “**interview-with-cool-tshirts-founder**” campaign. Our results from the first touch research indicate that the most first touches (622) are attributed to this ad through the “medium” source.
- 2) The “**getting-to-know-cool-tshirts**” campaign in the “nytimes” source. This should be another easy decision as the campaign comes in a close second with marginally fewer first touches (612).
- 3) The “**ten-crazy-cool-tshirts**” campaign hosted by “buzzfeed” as it is also not far behind with 576 and is clearly competing for the lead whereas the “cool-tshirts-search” campaign (do not re-invest) that follows attributes a dismal 169 first touches.
- 4) The “**weekly-newsletter**” campaign through “email” which is responsible for the most (115) last touches on the purchase page.
- 5) The “**retargeting-ad**” on “facebook” as it attributes 113 last touches on the purchase page and is clearly worth the investment. These last two ad campaigns show the most promise in facilitating users to reach the purchase point.

** These three ad campaigns are proficient in creating both initial traffic to the CTS website and potential customers.*
**The two latter campaigns are also both responsible for the most last touches (447 & 443) independent of purchases made, by a large margin (see slide 8).*

The recommendations above are supported by the results and data highlighting customer acquisition and reacquisition efforts presented in the previous slides. If pursued, they would likely yield higher sales through the CTS website and a more cost-effective marketing approach by re-investing in the campaigns that have statistically proven to be the most favorable.

THE END

**Presentation & Notepad Queries uploaded to [JamesGSeltzer/Capstone](#)
Repository on GitHub**