



Attribution Capstone

Learn SQL from Scratch

Andy Thom, andythom27

April 28, 2019

Table of Contents

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

1. How Many Campaigns and Sources does CoolTShirts Use?

How Many Distinct Campaigns?

The first question we were asked was to determine the amount of distinct campaigns that CoolTShirts is currently running:

At this time CoolTShirts has 8 active campaigns.

campaigns
8

```
--Sql:
```

```
SELECT COUNT(DISTINCT utm_campaign) AS 'campaigns'  
FROM page_visits;
```

How Many Distinct Sources?

Second we were asked to count the amount of sources that the campaigns are coming from:

There were a total of 6 sources.

```
SELECT COUNT(DISTINCT utm_source) AS 'sources'  
FROM page_visits;
```

sources
6

How are the campaigns and sources related?

The last step in this first question we were asked to show which sources were running which campaigns:

You can see below that the table shows that most campaigns are run on individual sources, although Google is running two at this time.

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

--Sql:

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

2. What Pages are on the CoolTShirts Website?

Find the Different Pages on the CoolTShirts WebSite?

We should know what pages are available to customers so we can track where they land and when they buy:

Currently, CoolTShirts has four available pages.

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

--Sql:

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


3. How Many First Touches is Each Campaign Responsible For?

Where are Consumers First Clicking on the CoolTShirts Link?

CoolTShirts Executives asked which campaigns were successful and where they are coming from.

As you can see, there are only four campaigns that are currently being clicked on.

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 (  
  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  
  FROM first_touch ft  
  JOIN page_visits pv  
    ON ft.user_id = pv.user_id  
    AND ft.first_touch_at = pv.timestamp),  
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(*) AS 'count'  
FROM ft_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

4. How Many Last Touches is Each Campaign Responsible For?

Where are Consumers Last Clicking on the CoolTShirts Link?

CoolTShirts Executives asked which campaigns were successful and where they ended up.

As you can see, all 8 campaigns came in as a last place that consumers accessed the CoolTShirts site.

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 (  
  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  
  FROM last_touch lt  
  JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
    AND lt.last_touch_at = pv.timestamp),  
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 AS 'source',  
       lt_attr.utm_campaign AS 'campaign',  
       COUNT(*) AS 'count'  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```

5. How Many Unique Visitors Make it to the Purchase Page?

How Many Distinct Users Make it to the Purchase Page?

Since we see all of this traffic coming into the site, Executives wanted to know how many unique consumers make it to the final purchase page:

As it stands now, we have 361 unique users that made it through the site to purchase their items. This doesn't necessarily mean that they didn't purchase multiple times.

completed_purchases
361

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

6. How Many Last Touches on the Purchase Page is Each Campaign Responsible For?

Which Campaigns Got the Consumer to the Purchase Page?

Finally, which campaigns are working and getting the consumer to the purchase page to buy some CoolTShirts?

source	campaign	count
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-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

```
WITH last_touch AS (  
  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  
  FROM last_touch lt  
  JOIN page_visits pv  
    ON lt.user_id = pv.user_id  
   AND lt.last_touch_at = pv.timestamp),  
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 AS 'source',  
       lt_attr.utm_campaign AS 'campaign',  
       COUNT(*) AS 'count'  
FROM lt_attr  
GROUP BY 1, 2  
ORDER BY 3 DESC;
```