



Learn SQL from Scratch

Capstone Project: Attribution


Bom Lee
June 19, 2018



Table of Contents

- Getting to know CoolTShirts
- The user journey through CoolTShirts
- Optimizing their campaign budget

Getting to know CoolTShirts


- 
- CoolTShirts is an apparel shop that is currently running 8 campaigns (also known as specific ad or email blasts): CoolTShirts Search, Getting to Know CoolTShirts, Interview with CoolTShirts Founder, Paid Search, Retargeting Ad, Retargeting Campaign, Ten Crazy CoolTShirts Facts, Weekly Newsletter

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

The query above returns the number of distinct, or unique, campaigns, which is 8

```
SELECT DISTINCT utm_campaign
FROM page_visits;
```

The query above returns the names of the distinct campaigns


- 
- These campaigns are running in 6 sources (also called channels or touchpoints): BuzzFeed, Email, Facebook, Google, Medium, New York Times
 - Each campaign is running in one source, but sources can have more than one affiliated campaign

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

The query above returns the number of distinct, or unique, sources, which is 6

```
SELECT utm_campaign, utm_source
FROM page_visits
GROUP BY utm_campaign;
```

The query above returns the distinct campaigns with their corresponding sources

- 
- Users ideally make their way from the initial page of CoolTShirts site where they can browse shirts to the final page where they complete their purchase
 - The 4 page names of the site are: Landing Page, Shopping Cart, Checkout, Purchase

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

The query above returns the distinct values for page names on the CoolTShirts site


The user journey through CoolTShirts

- Some campaigns are more successful at drawing unique visitors to the site, known as first touches
- The result-set from the query on the right shows the campaign “Interview with CoolTShirts Founder” was the most successful at first touches with 622 visits from unique visitors:

utm_campaign	count(*)
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

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


The query above returns the number of first touches per unique user grouped by the campaign which drove them first to CoolTShirts

- 
- Some campaigns are more successful at driving visitors back to the site, known as last touches
 - The result-set from the query on the right shows “Weekly Newsletter” was the most successful at last touches:

utm_campaign	count(*)
weekly-newsletter	447
retargeting-ad	443
retargeting-campaign	245
getting-to-know-cool-tshirts	232
ten-crazy-cool-tshirts-facts	190
interview-with-cool-tshirts-founder	184
paid-search	178
cool-tshirts-search	60


```
WITH last_touch AS (  
  SELECT user_id,  
         MAX(timestamp) AS last_touch_at  
  FROM page_visits  
  GROUP BY user_id)  
SELECT pv.utm_campaign, COUNT(*)  
FROM last_touch lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
GROUP BY utm_campaign  
ORDER BY COUNT(*) DESC;
```

The query above returns the number of last touches per unique user grouped by the campaign which drove them back to CoolTShirts

- 
- Of all the unique visitors to CoolTShirts, 361 complete the journey from first touch to purchase

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

The query above returns the number of unique visitors who completed a purchase on the CoolTShirts site (this may mask the number of repeat visitors who returned on more than one occasion to complete a purchase)


- 
- Of all the campaigns, “Weekly Newsletter” has the highest number of last touches to the purchase page, with “Retargeting Ad” a close second
 - The result-set from the query on the right shows “Weekly Newsletter” was the most successful at last touches to the purchase page:

utm_campaign	count(*)
weekly-newsletter	115
retargetting-ad	113
retargetting-campaign	54
paid-search	52
getting-to-know-cool-tshirts	9
ten-crazy-cool-tshirts-facts	9
interview-with-cool-tshirts-founder	7
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)  
SELECT pv.utm_campaign, count(*)  
FROM last_touch lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
GROUP BY utm_campaign  
ORDER BY COUNT(*) DESC;
```

The query above returns the number of unique visitors who completed a purchase on the CoolTShirts site (this may mask the number of repeat visitors who returned on more than one occasion to complete a purchase)

Optimizing their campaign budget

- 
- When it's time for CoolTShirts to re-invest in campaigns, they can maximize their return by selecting the following 5:
 - Interview with CoolTShirts Founder (for driving the highest number of first touches for unique visitors)
 - Getting to Know CoolTShirts (for driving the second highest number of first touches for unique visitors)
 - Ten Crazy CoolTShirts Facts (for driving the third highest number of first touches for unique visitors)
 - Weekly Newsletter (for driving the highest number of last touches to the purchase page)
 - Retargeting Ad (for driving the second highest number of last touches to the purchase page)

**Thank you for the opportunity
to expand my skill set!**
