



Cool T-shirts Capstone

Learn SQL from Scratch

Ryan Snyder

June 15 - 2018

Table of Contents

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

1. Get Familiar with the Cool T-Shirts

1.1 Know the Cool T-Shirts

- **Q: How many campaigns and sources does CoolTShirts use** and how are they related?
- **A:** We count 8 distinct Campaigns and 8 unique Sources using two separate Queries, with aliases for the Column Headers.

Campaigns
8
Sources
6

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

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

1.2 Know the Cool T-Shirts

- **Q:** How many campaigns and sources does CoolTShirts use and **how are they related?**
- **A:** We start by selecting the distinct values from the Campaign Column. We select this column first because it has a greater number of distinct values than the Sources column, that way we don't omit any rows. In the same statement we also select the Sources column. We add aliases to both columns for a comprehensive table showing us a list of all Campaigns right next to their corresponding Sources.

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

```
SELECT DISTINCT utm_campaign AS All_Campaigns,  
utm_source AS Sources FROM page_visits;
```

1.3 Know the Cool T-Shirts

- **Q: What pages are on their website?**
- **A:** Here we simply return the distinct values from the `page_name` column. There are four pages: Landing page, Shopping Cart, Checkout, and Purchase.

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

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

2. What is the User Journey?

2.1 The User Journey

- **Q: How many first touches is each campaign responsible for?**
- **A:** Some campaigns are “re-targetting” campaigns which only apply to users who have already visited the site. That being said, only 4 campaigns result in ‘first touch’ traffic. The most traffic comes from the “interview with cool tshirts founder” at 622. “Getting to know cool tshirts” is a close second at 612, followed closely again by “ten crazy cool tshirts facts” at 576. The google search is a distant last place at 169 accounting for roughly 8.5% of all first touch traffic.

Source	Campaign	First_Touches
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(*) AS First_Touches  
  FROM ft_attr GROUP BY 1, 2  
 ORDER BY 3 DESC;
```


2.2 The User Journey

- **Q: How many last touches is each campaign responsible for?**
- **A:** This Query will return rows for all of the campaigns. The weekly newsletter is the winner here at 447. The Facebook re-targeting ad is nearly tied at 443. These two campaigns are almost twice as effective as the next runner up.

Source	Campaign	Last_Touches
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(*) AS Last_Touches  
  FROM lt_attr GROUP BY 1, 2  
 ORDER BY 3 DESC;
```

2.3 The User Journey

- **Q: How many visitors make a purchase?**
- **A:** Counting the number of distinct users to touch the purchase page we get the answer: 361 This may only determine the number of customers, not the total number of sales as some customers may have made more than one purchase. I added a second query counting all 'touches' on the purchase page shows us there was in fact one transaction per distinct user ID.

Customers
361
Sales
361

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

2.4 The User Journey

- **Q: How many last touches *on the purchase page* is each campaign responsible for?**
- **A:** After adding a Count and a Where statement to the last touch attribution query, we get the answer. Once again, the email weekly newsletter and the facebook re-targetting ad are dominating here at 114, and 112. The next runner up is less than half as effective with 53.

Source	Campaign	Purchase_page_Last_Touches
email	weekly-newsletter	114
facebook	retargetting-ad	112
email	retargetting-campaign	53
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_purchase AS (  
  SELECT user_id, MAX(timestamp) AS last_touch_at  
    FROM page_visits GROUP BY user_id),  
  
lp_attr AS (  
  SELECT lp.user_id, lp.last_touch_at, pv.utm_source,  
         pv.utm_campaign, pv.page_name  
    FROM last_purchase lp JOIN page_visits pv  
   ON lp.user_id = pv.user_id  
  AND lp.last_touch_at = pv.timestamp)  
SELECT lp_attr.utm_source AS Source,  
  
       lp_attr.utm_campaign AS Campaign,  
       COUNT(*) AS Purchase_page_Last_Touches  
  FROM lp_attr  
 WHERE lp_attr.page_name = '4 - purchase'  
 GROUP BY 1, 2  
 ORDER BY 3 DESC;
```

2.5 The User Journey

- **Q: What is the typical user journey?**
- **A:** Most users arrive to the site for the first time via one of these Campaigns/sources: “Interview with cool tshirts founder” - Medium, “Getting to know cool tshirts” - Nytimes, or “ten crazy cool tshirts facts” - Buzzfeed. The total number of first touches greatly exceeds the total number of sales (361) which would indicate that few users make a purchase on their first visit. Repeat visits can be largely attributed to the top performers in our last touch attribution query: Email weekly-newsletter, and Facebook retargeting ad. These also land in top 2 spots for our purchase page last touch query. This serves as further evidence that most purchases are made by users who have already visited the site at least once before. The Purchase page last touches column in the final query sums up to 358. Knowing that total sales are 361, this tells us that only 3 users continued to navigate through the site after making a purchase.

Source	Campaign	First_Touches
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
Source	Campaign	Last_Touches
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
Source	Campaign	Purchase_page_Last_Touches
email	weekly-newsletter	114
facebook	retargeting-ad	112
email	retargeting-campaign	53
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

3. Optimize the Campaign budget!

3 Optimize the budget

- **Q: CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?**
- **A:** I was always taught that getting your customers to show up was the most important factor in capturing sales. That being said, the top 3 first touch campaigns (medium, nytimes, buzzfeed) seem like an obvious choice to start. With this web store we also see that the top 2 re-targeting campaigns (email weekly newsletter, and facebook re-targeting ad) have directly led to the most repeat traffic, as well as the most sales as we see in the Purchase page last touch column. So those would be my high five. You could make a strong argument to go with the third top performing re-targeting ad but since it shares a source with the weekly newsletter (email) I would suggest it might be a little redundant, and less valuable than the buzzfeed article as it generates a lot of first touch traffic. That's where the business begins, you've got to get those customers in the door first.

Source	Campaign	First_Touches
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
Source	Campaign	Last_Touches
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
Source	Campaign	Purchase_page_Last_Touches
email	weekly-newsletter	114
facebook	retargeting-ad	112
email	retargeting-campaign	53
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