



Capstone: Attribution

Learn SQL from Scratch

Marije van Schie

08/08/2018

Hi there,

My name is Marije van Schie. I am from The Netherlands, Europe. I work as a strategic program manager for a SaaS company called Blue10. Blue10 offers a Cloud solution for the digital processing of booking documents, such as procurement invoices, sales invoices, receipts, etc. The online workflow can be used to have an invoice inspected and approved digitally by one or more employees. Once an invoice has gone through the entire process in Blue 10, the document can easily be accessed in the digital archive and is retrievable from any location.

As a strategic program manager I am responsible for different strategic programs within the company. Before the course I have played with data quite a bit in data visualization tools (like Power BI) and Excel, but always relied on others to do the SQL. I really liked the course and learned a lot of it.



*English isn't my first language,
so please excuse any mistakes.*



Table of Contents

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

* The queries used to find that data
are all included in the .sql file

1. Get familiar with CoolTShirts

1.1 Traffic CoolTShirts

CoolTShirts sells shirts of all kinds, as long as they are T-shaped and cool. They started **8 marketing campaigns** to increase website visits and purchases. CoolTShirts uses **6 sources** from where the traffic is sent to their website. A campaign is a specific program that is designed to market the product. Each campaign can be launched via a different channel. The specific channel that is used to drive a potential customer to a website is called the 'source'. Sources can be used for more than one campaign.

| Query Results | |
|-------------------------------------|------------|
| 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 |

Source used for each campaign

| Query Results |
|----------------------------|
| Number of distinct sources |
| 6 |

| Query Results |
|------------------------------|
| Number of distinct campaigns |
| 8 |

1.2 Website

There are **4 pages** on the website of CoolTShirts, shown in the graph. The goal of CoolTShirts must be to increase the conversion rate from landingpage to purchase. Using touch attribution, they can map their customers' journey: from initial visit to purchase. A customer journey is the route your customer takes to complete your desired conversion action. CoolTShirts can use that information to optimize their marketing campaigns.

| Query Results |
|-------------------|
| page_name |
| 1 - landing_page |
| 2 - shopping_cart |
| 3 - checkout |
| 4 - purchase |



2. What is the user journey?

2.1 First touches

The following campaigns are responsible for first touches:

| Query Results | | |
|---------------|-------------------------------------|---------------|
| 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 |

This means that, for example, 622 visitors came for the first time to CoolTShirts.com due to the campaign interview-with-cool-tshirts-founder

```
WITH first_touch AS (  
    SELECT user_id,  
           MIN(timestamp) as first_touch_at  
    FROM page_visits  
    GROUP BY user_id  
)  
ft_alloc AS (  
    SELECT first_touch.user_id,  
           first_touch.first_touch_at,  
           page_visits.utm_source,  
           page_visits.utm_campaign  
    FROM first_touch  
    JOIN page_visits  
      ON first_touch.user_id = page_visits.user_id  
     AND first_touch.first_touch_at =  
        page_visits.timestamp  
)  
SELECT  
    ft_alloc.utm_source AS Source,  
    ft_alloc.utm_campaign AS Campaign,  
    COUNT(*) AS 'First Touches'  
FROM ft_alloc  
GROUP BY 2  
ORDER BY 3 DESC;
```


2.1 Last touches

The following campaigns are responsible for last touches:

| Query Results | | |
|---------------|-------------------------------------|--------------|
| 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 |

This means that, for example, 447 visitors has the last time been at CoolTShirts.com due to the campaign weekly-newsletter.

```
WITH last_touch AS (  
    SELECT user_id,  
           MAX(timestamp) as last_touch_at  
    FROM page_visits  
    GROUP BY user_id  
)  
lt_alloc AS (  
    SELECT last_touch.user_id,  
           last_touch.last_touch_at,  
           page_visits.utm_source,  
           page_visits.utm_campaign  
    FROM last_touch  
    JOIN page_visits  
      ON last_touch.user_id = page_visits.user_id  
     AND last_touch.last_touch_at =  
        page_visits.timestamp  
)  
SELECT  
    lt_alloc.utm_source AS Source,  
    lt_alloc.utm_campaign AS Campaign,  
    COUNT(*) AS 'Last Touches'  
FROM lt_alloc  
GROUP BY 2  
ORDER BY 3 DESC;
```

1.3 Purchases

361 unique visitors made a purchase on CoolTShirts.

Each purchase can be attributed to a certain campaign. Below you can see how many purchases each campaign is responsible for. Weekly-newsletter was most succesful.

| Query Results | | |
|---------------|-------------------------------------|--------------|
| Source | Campaign | Last Touches |
| 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 |

361

Query Results

Unique visitors purchase page

361

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

```
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
),
lt_alloc AS (
  SELECT last_touch.user_id,
    last_touch.last_touch_at,
    page_visits.utm_source,
    page_visits.utm_campaign
  FROM last_touch
  JOIN page_visits
    ON last_touch.user_id =
    page_visits.user_id
    AND last_touch.last_touch_at =
    page_visits.timestamp
)
SELECT
  lt_alloc.utm_source AS Source,
  lt_alloc.utm_campaign AS Campaign,
  COUNT(*) AS 'Last Touches'
FROM lt_alloc
GROUP BY 2
ORDER BY 3 DESC;
```

1.4 Typical user journey



Top 3 of:

- Campains responsible for the most first touches:

1. Interview-with-cool-tshirts-founder
2. Getting-to-know-cool-tshirts
3. Ten-crazy-cool-tshirts-facts

- Campains responsible for the most last touches:

1. Weekly-newsletter
2. Retargeting-ad
3. Retargeting-campaign



Conversion rate
visitor to purchase:
18,25%



| Query Results |
|-----------------------------|
| Unique visitors CoolTShirts |
| 1979 |

| Query Results |
|-------------------------------|
| Unique visitors purchase page |
| 361 |

```
SELECT COUNT (DISTINCT user_id)  
AS 'Unique visitors CoolTShirts'  
FROM page_visits;
```

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

3. Optimize the campaign budget

3.1 Advice re-invest

- **First-touch attribution** only considers the first source for each customer. This is a good way of knowing how visitors initially discover a website.
- **Last-touch attribution** only considers the last source for each customer. This is a good way of knowing how visitors are drawn back to a website, especially for making a final purchase.

There are a lot of different ways to assign value to campaigns and touch points. Is the first touch (like clicking a facebook ad) more important than the e-mail that generated the actual purchase? Or, should credit get divided up equally across each touchpoint? There are different ways to look at marketing attribution.

That being said, both first-touch attribution and last-touch attribution are very important.

If they want to focus on more traffic to their website, the advice will be to re-invest in the following campaigns:

| Query Results | | |
|---------------|-------------------------------------|---------------|
| 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 |

If they want to focus on more purchases, the advice will be to re-invest in the following campaigns

| Query Results | | |
|---------------|----------------------|--------------|
| Source | Campaign | Last Touches |
| email | weekly-newsletter | 115 |
| facebook | retargeting-ad | 113 |
| email | retargeting-campaign | 54 |
| google | paid-search | 52 |

Best advice is probably to do both. *'No sale ever gets made if customers doesn't know you exist'*. And *'There will be no revenue if no one made the purchase'*