



# Capstone project

Learn SQL from Scratch

Tomasz Bulawa

3.07.2018

# Table of Contents

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

# 1. Getting to know CoolTShirts

# 1.1 Campaigns and sources

Here are distinct campaigns and sources that are used by CoolTShirts.com

- There are 8 distinct campaigns
- Also there are 6 sources
- On the right we can see which campaign relates to which source
- The difference between source and campaign is that source is the medium used to present content to clients, and campaign is form in which that contet is presented

```
select count (distinct utm_campaign)
from page_visits
;
select count (distinct utm_source)
from page_visits
;
select distinct utm_campaign, utm_source
from page_visits
;
```

## 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

## 1.2 Sites used.

Bottom table shows which sites are on CoolTShirts.com.

```
select distinct page_name
from page_visits
;
```

### Query Results

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

## 2. User journey

## 2.1 How many first touches is each campaign responsible for?

Table below shows how many first touches each campaign is responsible for. (On the right is query used to get that table).

utm_campaign	count (first_touch.first_touch_at)
cool-tshirts-search	169
getting-to-know-cool-tshirts	612
interview-with-cool-tshirts-founder	622
paid-search	0
retargetting-ad	0
retargetting-campaign	0
ten-crazy-cool-tshirts-facts	576
weekly-newsletter	0

```
WITH first_touch AS (  
  SELECT user_id,  
         min(timestamp) as first_touch_at  
  FROM page_visits  
  GROUP BY user_id)  
SELECT utm_campaign, count  
      (first_touch.first_touch_at)  
FROM page_visits  
left join first_touch  
on page_visits.user_id = first_touch.user_id  
and first_touch.first_touch_at = page_visits.timestamp  
group by page_visits.utm_campaign;
```

## 2.2 How many last touches each campaign is responsible for

Table below shows for how many last touches each campaign is responsible for.

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

```
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 (distinct  
ft.last_touch_at) as 'Last_Touch_Amount'  
FROM last_touch ft  
JOIN page_visits pv  
  ON ft.user_id = pv.user_id  
  AND ft.last_touch_at = pv.timestamp  
group by pv.utm_campaign;
```



## 2.3 How many users visited page 4 - purchase

Table below shows how many users visited page 4 - purchase

```
select count (distinct user_id) as 'Visits_on_4'  
from page_visits  
where page_name = '4 - purchase'  
;
```

Visits_on_4
361

## 2.4 How many last touches on the purchase page is each campaign responsible for?

Below is the table showing how many last touches on the purchase page is each campaign responsible for.

utm_campaign	Last_Touch_Amount
cool-tshirts-search	2
getting-to-know-cool-tshirts	9
interview-with-cool-tshirts-founder	7
paid-search	52
retargeting-ad	112
retargeting-campaign	53
ten-crazy-cool-tshirts-facts	9
weekly-newsletter	114

```
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 (distinct  
lt.last_touch_at) as 'Last_Touch_Amount'  
FROM last_touch lt  
JOIN page_visits pv  
  ON lt.user_id = pv.user_id  
  AND lt.last_touch_at = pv.timestamp  
where pv.page_name = '4 - purchase'  
group by pv.utm_campaign;
```

## 2.5 Typical user journey

With simple count of visits on each page we can determine that typical user journey starts at Landing\_Page and stops at Checkout. That's why retargeting campaigns are very important. More on that in conclusion.

Landing_page
2000
Shopping_cart
1900
Checkout
1431
Purchase
361

```
select count (user_id) as 'Landing_page'
from page_visits
where page_name = '1 - landing_page'
;
select count (user_id) as 'Shopping_cart'
from page_visits
where page_name = '2 - shopping_cart'
;
select count (user_id) as 'Checkout'
from page_visits
where page_name = '3 - checkout';
select count (user_id) as 'Purchase'
from page_visits
where page_name = '4 - purchase';
```

# 3. Conclusion

## 3.1 Conclusion, reinvesting in campaigns

Taking into consideration data that we acquired querying database, we can draw up conclusion that the campaigns with most return are: Getting To Know Cool T-Shirts, Interview with the founder and Ten crazy cool tshirts facts because they provided us with most of the first touches (Influx of new potential clients). Also retargetting ads and weekly newsletter are important and worth reinvesting into because they provided us with much more sales than any other campaigns.

Where the first three campaigns are responsible for new clientele there the two other are responsible for closing the deals.

utm_campaign	count (first_touch.first_touch_at)	Last_Touch_Amount	Purchase
cool-tshirts-search	169	60	2
getting-to-know-cool-tshirts	612	232	9
interview-with-cool-tshirts-founder	622	184	7
paid-search	0	178	52
retargetting-ad	0	443	112
retargetting-campaign	0	245	53
ten-crazy-cool-tshirts-facts	576	190	9
weekly-newsletter	0	447	114