# Capstone Project 3

Learn SQL from Scratch

Mikael Madsen
Date: 27/6 -2018

# Table of Contents

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

# 1. Get familiar with CoolTshirts

# Question 1.a
## How many campaigns and sources does CoolTShirts use and how are they related?

CoolTshirts uses **8** campaigns and **6** sources in order to attract visitors and hopefully increase sales.

The difference between a campaign and source is that the Utm campaign identifies the specific ad or email blast that is shown to the visitor and the utm source identifies which site sent the traffic. If an ad campaign drives a lot of visitors to their site, then they know that source is working.

The table to our right shows how a utm_source is realated to each utm_campaign. Email and Google has been used twice as a touchpoint.  It is through the source 'facebook' that the retargetting-ad is exposed or shown to the potential visitor.

| Query Results | |
|---|---|
| **utm_campaign** | **utm_source** |
| cool-tshirts-search | google |
| getting-to-know-cool-tshirts | nytimes |
| interview-with-cool-tshirts-founder | medium |
| paid-search | google |
| retargetting-ad | facebook |
| retargetting-campaign | email |
| ten-crazy-cool-tshirts-facts | buzzfeed |
| weekly-newsletter | email |

# Question 1.b
## What pages are on their website?

There are 4 distinct values in Page_name. Using the following SQL query will illustrate these 4 pages that are on the website. There is a landing_page, a shopping_cart page, a checkout page and a purchase page.

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

# 2. What is the user journey?

# Question 2. a
## How many first touches is each campaign responsible for?

The owner of the website can with help from utm parameters capture when and how the user finds the site. So every time a consumer is expsoed to a website for the first or last time they know which source they should attribute the traffic.

With help from 'With' we could combine our first calculated query, that helped me find 'first touch' for each user_id (consumer) with our table page_visits. The resut of the query isshown in the table beneath. The campaign "interview-with-cool-tshirts-founder" is responsible for most 'first touches' which was 622.

There are 8 campaigns and only 4 stands for first touch attribution. 3 of those are all very close and by far dominating the first touch position. All 3 are really good to create attention and makes people want to go visit the website. So a lot of traffic is created with the use of these campaigns and that is very positive.

| 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 |

# Question 2.b
## How many last touches is each campaign responsible for?

Very similar to the 'first_touch' query, what I did change was timestamp in the first query from 'min' to 'max', which is the exposure to the website that led to a purchase.

The campaign "interview-with-cool-tshirts-founder" with the source 'Medium' was number one in first_touches but when it comes to purchase and last_touches they are 6. Which says somethings that there are other more succesfull campaigns and sources that lead to a final purchase.

| Query Results | | |
|---|---|---|
| **Source** | **Campaign** | **Last touches** |
| 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 (Select user_id,
max(timestamp) as 'last_touch_at'
from page_visits
group by 1),

lt_atrr as (Select lt.user_id,
lt.last_touch_at, pv.utm_source,
pv.utm_campaign
from last_touch as 'lt'
left join page_visits as 'pv'
on lt.user_id = pv.user_id and
lt.last_touch_at = pv.timestamp)

select lt_atrr.utm_source as 'Source',
lt_atrr.utm_campaign as 'Campaign',
count(*) as 'Last touches'
from lt_atrr
group by 2
order by 3 desc;
```

# Question 2.c

## How many visitors make a purchase?

In the query result table, we can see that there are 361 out of the 1979 distinct users who make a purchase. Thats 18.2 %.

| Query Results | |
| --- | --- |
| page_name | count(distinct user_id) |
| 1 - landing_page | 1979 |
| 2 - shopping_cart | 1881 |
| 3 - checkout | 1431 |
| 4 - purchase | 361 |

```
Select page_name, count(distinct user_id)
from page_visits
group by page_name;
```

# Question 2.d

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

We knew before we calculated the query that 361 distinct users had made a purchase. To calculate the query result in the table beneath, we had to add an additional 'where' clause in our last_touch query to get results that only was about the purchase page. The "weekly-newsletter" and "retargeting-campaign were close in the top.

| Query Results | | |
|---|---|---|
| **Source** | **Campaign** | **Last touch on pp** |
| 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 |

```sql
With last_touch as (Select user_id,
max(timestamp) as 'Last_touch_at'
from page_visits
where page_name = '4 - purchase'
group by 1),

lt_attr as (Select lt.user_id,
lt.last_touch_at, pv.utm_source,
pv.utm_campaign
from last_touch as 'lt'
join page_visits as '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 touch on pp'
from lt_attr
group by 2
order by 3 desc;
```

# Question 2.e

**What is the typical user journey?**

The users journey, from initial visit to purchase. The way I see it, the most common journey in this case is that the initial contact to the website comes from one source/touchpoint, but the last touch and purchase comes from another source. The visitors goes away and then returns back through a new source.

So, the owners have to make decision on which source, first or last attribution, they should count on in order to improve the marketing and online presence to increase sales

# 3. Optimize the campaign budget

# Question 3.a

CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

We want to reinvest in 5 campaigns and we got to make sure that that investment is the best way to increase sales. Like it was said in the lesson it is up to the company to find out which source they should attribute to a purchase, which can be a super tough decision.

The first_touch or the last_touch. First touch attribution are good to see where and how visitors initially discover a website and I think that is very important because that can lead us to retarget the vistors. On the other hand we have Last touch attribution, which is good to see how vistitors are drawn back to it. Also very important. I have decided to go with a mix between campaigns that are strong in first_touch and last_touch and purchase.

My top 5 list for re-investment continues on the next slide.

# Re-invest in these 5 campaigns:

- **Paid search**: PS is not contributing in the initially first touch face and there are 6 campaigns that are more successful in generate a last touch. But we got to take into consideration that PS is number 4 on final purchase and has the highest conversation rate with 29%.

- **weekly newsletter**: WN is not existing in first touch but number one in last touch and purchase. I want to continue making this a strong one since it is clearly working.

- **interview with cool tshirtts founder** is number one in first touch attribution and last touch attribution and purchase not very impressive. But is is a great place for visitors to be exposed to the company and I want to make sure that it will continue to be that.

- **retargeting ad**: RA is not existing in first touch but second in last touch and purchase. I want to continue making this campaign still brings home sale.

- **getting to know cool tshirts** - just under number one in first touch but not impressive on the purchase page. I wanto to continue making this a strong entry source to the website and it has shown great promise on last_touch. With a little boost, it might be a strong campaign for purchase as well.