

Learn SQL from Scratch - Capstone

First- and last-touch attribution

Get familiar with the company

Campaign	Source
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
8	6

How many campaigns and sources does CoolTShirts use and how are they related?

- There are eight different campaigns spread across six sources.
- A source is a medium through which customers are targetted.
- A campaign incorporates the specific message the company wants to spread.
- From the data it is clear that a source can have multiple campaigns (email, google).
- None of the campaigns are currently run through multiple sources, although this is theoretically possible.

```
19 -- 1.
20 -- Number of distinct campaigns.
21 SELECT COUNT(DISTINCT utm_campaign) as campaign_count FROM page_visits;
22 -- Number of distinct sources
23 SELECT COUNT(DISTINCT utm_source) as source_count FROM page_visits;
24 -- Number of distinct combinations
25 SELECT DISTINCT utm_campaign, utm_source FROM page_visits;
```

Getting familiar: Pages on the website

Pages
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase
Page count: 4

What pages are on their website?

- There are four distinct pages represented in the database that presumably handle the complete purchase flow.
- The more static pages (e.g. a contact form, company history) are not listed in the database.

```
29  -- 2.  
30  -- Pages on the website.  
31  SELECT DISTINCT page_name FROM page_visits;
```

User journey: first touches per campaign

Campaign	Source	#
interview-with-cool-tshirts-founder	medium	622
getting-to-know-cool-tshirts	nytimes	612
ten-crazy-cool-tshirts-facts	buzzfeed	576
cool-tshirts-search	google	169

How many first touches is each campaign responsible for?

- The first entry of a potential customer to the website is determined by finding the **minimal timestamp** (first record in time) and joining this timestamp and user with the related campaign and source. The list of all users and related campaigns is then **grouped by the campaign and source** to get the total of first touches.
- Only four campaigns are responsible for the recorded first touches.

```
31 -- 3.
32 -- First touches per campaign.
33 WITH first_touches AS (
34   SELECT user_id, MIN(timestamp) as 'first_touch_at'
35     FROM page_visits
36    GROUP BY user_id
37 )
38 SELECT pv.utm_campaign, pv.utm_source, COUNT(pv.utm_campaign) as
39   'first_touch_count'
40 FROM first_touches ft
41 JOIN page_visits pv
42   ON ft.user_id = pv.user_id
43  AND ft.first_touch_at = pv.timestamp
44 GROUP BY pv.utm_campaign, pv.utm_source
45 ORDER BY first_touch_count DESC
46 ;
```

User journey: last touches per campaign

Campaign	Source	#
weekly-newsletter	email	447
retargeting-ad	facebook	443
retargeting-campaign	email	245
getting-to-know-cool-tshirts	nytimes	232
ten-crazy-cool-tshirts-facts	buzzfeed	190
interview-with-cool-tshirts-founder	medium	184
paid-search	google	178
cool-tshirts-search	google	60
	Total	1979

```
47 -- 4.
48 -- Last touches per campaign
49 WITH last_touches AS (
50   SELECT user_id, MAX(timestamp) as 'last_touch_at'
51   FROM page_visits
52   GROUP BY user_id
53 )
54 SELECT pv.utm_campaign, pv.utm_source, COUNT(pv.utm_campaign) as
55   'last_touch_count'
56 FROM last_touches lt
57 JOIN page_visits pv
58 ON lt.user_id = pv.user_id
59 AND lt.last_touch_at = pv.timestamp
60 GROUP BY pv.utm_campaign, pv.utm_source
61 ORDER BY last_touch_count DESC
62 ;
```

How many last touches is each campaign responsible for?

- The process for the last touches is the same as first touches, but now the initial list is created by getting the maximum timestamp (i.e. the last page visit) per user.
- The retargeting campaigns create the most last_touches. The campaigns that are responsible for the first touches are not responsible for many last touches (compare with previous slide).

User journey: actual purchases

```
65 -- 5.
66 -- Visitors that make a purchase
67 SELECT COUNT(DISTINCT user_id) as 'purchase_users'
68 FROM page_visits
69 WHERE page_name = '4 - purchase';
```

How many visitors make a purchase?

- The total number of purchases is 361.
- This is 18.2% of to total (1979, see previous slide).

User journey: purchase last touches per campaign

Campaign	Source	#
weekly-newsletter	email	115
retargeting-ad	facebook	113
retargeting-campaign	email	54
paid-search	google	52
getting-to-know-cool-tshirts	nytimes	9
ten-crazy-cool-tshirts-facts	buzzfeed	9
interview-with-cool-tshirts-founder	medium	7
cool-tshirts-search	google	2

```
69 -- 6.
70 -- Last touches on purchase page per campaign
71 WITH last_touches AS (
72   SELECT user_id, MAX(timestamp) as 'last_touch_at'
73   FROM page_visits
74   WHERE page_name = '4 - purchase'
75   GROUP BY user_id
76 )
77 SELECT pv.utm_campaign, pv.utm_source, COUNT(pv.utm_campaign) as
78   'last_touch_count'
79 FROM last_touches lt
80 JOIN page_visits pv
81 ON lt.user_id = pv.user_id
82 AND lt.last_touch_at = pv.timestamp
83 GROUP BY pv.utm_campaign, pv.utm_source
84 ORDER BY last_touch_count DESC
85 ;
```

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

- The previous last touch query is modified to add the purchase page restriction to the initial list of last touches per user.
- Again, the **retargeting campaigns** are largely responsible for the last touches on the purchase page, although paid search has a decent contribution as well.

The typical user journey

First touches

Campaign	Source	#
interview-with-cool-tshirts-founder	medium	622 (31%)
getting-to-know-cool-tshirts	nytimes	612 (31%)
ten-crazy-cool-tshirts-facts	buzzfeed	576 (29%)
cool-tshirts-search	google	169 (9%)
	Total	1979

Last touches

Campaign	Source	#
weekly-newsletter	email	447 (23%)
retargeting-ad	facebook	443 (22%)
retargeting-campaign	email	245 (12%)
getting-to-know-cool-tshirts	nytimes	232 (12%)
ten-crazy-cool-tshirts-facts	buzzfeed	190 (10%)
interview-with-cool-tshirts-founder	medium	184 (9%)
paid-search	google	178 (9%)
cool-tshirts-search	google	60 (3%)
	Total	1979

Last touches on purchase page

Campaign	Source	#
weekly-newsletter	email	115 (32%)
retargeting-ad	facebook	113 (31%)
retargeting-campaign	email	54 (15%)
paid-search	google	52 (14%)
getting-to-know-cool-tshirts	nytimes	9 (3%)
ten-crazy-cool-tshirts-facts	buzzfeed	9 (3%)
interview-with-cool-tshirts-founder	medium	7 (2%)
cool-tshirts-search	google	2 (1%)
	Total	361

When we summarize the data in the previous slides and add percentages a 'typical user journey' can easily be extracted.

1. Campaigns with a general overview (interview, getting to know...) of the company get people to the website (91% first touches).
2. These campaigns are terrible at producing purchases (9% including google search, see last touches on purchase page).
3. Retargeting campaigns (including a newsletter) produce the most purchases (78%).

Optimizing the budget

From the data and numbers presented in the previous slides it is clear that there are **two different types of campaigns** that are important to generate revenue. The first is targetted at **generating interest** in the company and its products. The second type consists of the **retargetting campaigns** that get users back to the website to finalize their purchase.

In the first category, there are three campaigns that excel at generating interest, and should be continued.

- interview-with-cool-tshirts-founder
- getting-to-know-cool-tshirts
- ten-crazy-cool-tshirts-facts.

In the second category there are two campaigns that push customers to make a purchase.

- weekly-newsletter
- retargetting-ad

These five campaigns should be the ones taken into consideration for reinvestment.

Optional strategy - combine email campaigns

As a side note, the third best performing retargetting campaign is the also an 'email' campaign. This is likely to target the same audience as the weekly-newsletter campaign. When these two campaigns are combined there is only one large source of customer traffic left, which is the paid-search through google. Instead of five, it might be worthwhile to invest in this campaign as well, giving a total of **six campaigns** (including a combined retargetting email campaign) worthy of reinvestment.