

LEARN SQL FROM SCRATCH:

CAPSTONE ON FIRST- & LAST-TOUCH ATTRIBUTION

By: Courteney Zboyan



OBJECTIVE: ANALYZE COOLSHIRTS' MARKETING CAMPAIGNS

1. Get familiar with the company:

- How many campaigns and sources does CoolTShirts use? How are they related? What pages are on their website?

2. What is the user journey:

- How many first and last touches is each campaign responsible for?
- How many visitors make a purchase?
- How many last touches *on the purchase page* is each campaign responsible for?
- What is the typical user journey?

3. Optimize the campaign budget:

- CoolTShirts can re-invest in 5 campaigns, which should be chosen?

A QUICK PEEK AT THE DATA:

Query Results				
page_name	timestamp	user_id	utm_campaign	utm_source
1 - landing_page	2018-01-24 03:12:16	10006	getting-to-know-cool-tshirts	nytimes
2 - shopping_cart	2018-01-24 04:04:16	10006	getting-to-know-cool-tshirts	nytimes
3 - checkout	2018-01-25 23:10:16	10006	weekly-newsletter	email
1 - landing_page	2018-01-25 20:32:02	10030	ten-crazy-cool-tshirts-facts	buzzfeed
2 - shopping_cart	2018-01-25 23:05:02	10030	ten-crazy-cool-tshirts-facts	buzzfeed
3 - checkout	2018-01-28 13:26:02	10030	retargetting-campaign	email
4 - purchase	2018-01-28 13:38:02	10030	retargetting-campaign	email
1 - landing_page	2018-01-05 18:31:17	10045	getting-to-know-cool-tshirts	nytimes
2 - shopping_cart	2018-01-05 21:16:17	10045	getting-to-know-cool-tshirts	nytimes
3 - checkout	2018-01-09 03:05:17	10045	retargetting-ad	facebook

- At left is a sampling of CoolTShirts' database.
- The column names are page_name, timestamp, user_id, utm_campaign, and utm-source.

```
1 SELECT *
2 FROM page_visits
3 LIMIT 10;
4
```

I. GET FAMILIAR WITH THE COMPANY:

HOW MANY CAMPAIGNS AND SOURCES DOES COOLTSHIRTS USE?
HOW ARE THEY RELATED?

At right, the number of campaigns and sources that CoolTShirts uses can be seen.

CoolTShirts uses 8 different marketing campaigns on 6 different sources. Note that it is also easy enough to manually count how many.

The campaign is the advertisement used in hopes to get people to buy a T-Shirt. The source is the platform (or website) that the campaign is run on. Multiple sources can run different campaigns and vice versa.

Query Results	
Source	Campaign
6	8
utm_source	utm_campaign
nytimes	getting-to-know-cool-tshirts
email	weekly-newsletter
buzzfeed	ten-crazy-cool-tshirts-facts
email	retargetting-campaign
facebook	retargetting-ad
medium	interview-with-cool-tshirts-founder
google	paid-search
google	cool-tshirts-search

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

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

```
SELECT DISTINCT utm_source, utm_campaign  
FROM page_visits;
```

I. GET FAMILIAR WITH THE COMPANY:

WHAT PAGES ARE ON THEIR WEBSITE?

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

```
SELECT DISTINCT page_name AS Page  
FROM page_visits;
```

At left, are the 4 different pages of the CoolITShirts website. They are in order the landing_page, the shopping_cart, the checkout, and the purchase page.

2. WHAT IS THE USER JOURNEY: HOW MANY FIRST AND LAST TOUCHES IS EACH CAMPAIGN RESPONSIBLE FOR?

```
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,
    ft_attr.utm_campaign,
    COUNT(*)
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

ft_attr.utm_source	ft_attr.utm_campaign	COUNT(*)
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

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
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 user_id),
lt_attr AS (
    SELECT lt.user_id,
        lt.last_touch_at,
        pv.utm_source,
        pv.utm_campaign
    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,
    lt_attr.utm_campaign,
    COUNT(*)
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

As seen above, the campaigns responsible for the most first-touches is ‘interview-with-cool-tshirts-founder,’ ‘getting-to-know-cool-tshirts,’ and ‘ten-crazy-cool-tshirts’ are not far behind with 622, 612, and 576 first-touches respectively. The best last-touch campaigns are ‘weekly-newsletter’ and ‘retargetting-ad’ with 447 and 443 last-touches respectively. All other first- or last-touch campaigns are less effective in generating page visits. The 4 campaigns for first-touch ranked 4th, 5th, 6th, and 8th for last-touch.

2. WHAT IS THE USER JOURNEY: HOW MANY VISITORS MAKE A PURCHASE?

```
SELECT COUNT(DISTINCT user_id) AS 'Purchasers'  
FROM page_visits  
WHERE page_name = '4 - purchase';  
  
SELECT COUNT(DISTINCT user_id) AS 'To Check-out'  
FROM page_visits  
WHERE page_name = '3 - checkout';  
  
SELECT COUNT(DISTINCT user_id) AS 'To Cart'  
FROM page_visits  
WHERE page_name = '2 - shopping_cart';  
  
SELECT COUNT(DISTINCT user_id) AS 'Visit Site'  
FROM page_visits  
WHERE page_name = '1 - landing_page';
```

Purchasers	361
To Check-out	1431
To Cart	1881
Visit Site	1979

At left, it can be seen that 361 visitors make a purchase.

I also slightly adjusted the code to see how many people visited the other pages as well.

Taking 361/1431, we see that ~25% purchase if they reach the checkout stage.

Overall, 361/1979 or 18.24% of people purchase something if they visit CoolTShirts.

2. WHAT IS THE USER JOURNEY: HOW MANY LAST TOUCHES ON THE PURCHASE PAGE IS EACH CAMPAIGN RESPONSIBLE FOR?

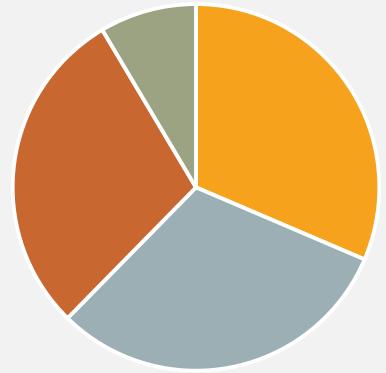
```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) AS last_touch_at
    FROM page_visits
    WHERE page_name = '4 - purchase'
    GROUP BY user_id),
lt_attr AS (
    SELECT lt.user_id,
        lt.last_touch_at,
        pv.utm_source,
        pv.utm_campaign
    FROM last_touch lt
    JOIN page_visits pv
        ON lt.user_id = pv.user_id
        AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
    lt_attr.utm_campaign,
    COUNT(*)
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-campaign	54
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

At above, it can be seen that the campaigns responsible for the most last-touches on the purchase page are ‘weekly-newsletter’ and ‘retargetting-ad’ with 115 and 113 last-touches respectively. All others are less effective in generating last-touches.

2. WHAT IS THE USER JOURNEY: WHAT IS THE TYPICAL USER JOURNEY?

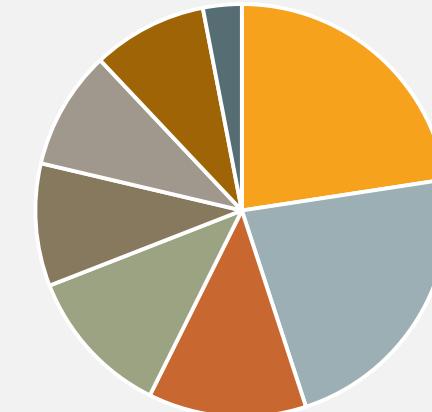
First-Touches



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

Above, it can be seen that most or 91% of users are attracted to the site by one of 'interview-with-cool-tshirts-founder,' 'getting-to-know-cool-tshirts,' and 'ten-crazy-cool-tshirts-facts' campaigns.

Last-Touches

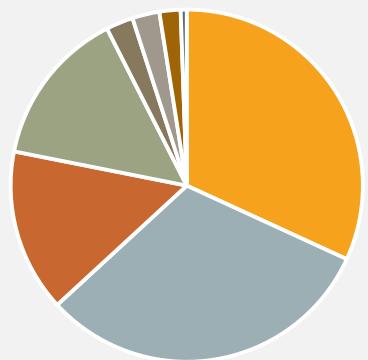


- weekly-newsletter
- retargetting-ad
- ten-crazy-cool-tshirts-facts
- getting-to-know-cool-tshirts
- interview-with-cool-tshirts-founder
- cool-tshirts-search

Above, it can be seen that a little over half or 55% of users see either 'weekly-newsletter,' 'retargetting-ad,' or ' as their last-touch campaign. The three main campaigns for first-touches make up 31% of last-touches.

2. WHAT IS THE USER JOURNEY: WHAT IS THE TYPICAL USER JOURNEY?

Sales



- weekly-newsletter
- retargetting-ad
- retargetting-campaign
- paid-search
- ten-crazy-cool-tshirt-facts
- getting-to-know-cool-tshirts
- interview-with-cool-tshirts-founder
- cool-tshirts-search

At left, it can be seen that 78% of purchases result from a user seeing one of the three following campaigns: 'weekly-newsletter,' 'retargetting-ad,' or 'retargetting-campaign.'

This information shows that 'interview-with-cool-tshirts-founder,' 'getting-to-know-cool-tshirts,' and 'ten-crazy-cool-tshirts-facts' are the main attracting campaigns while 'weekly-newsletter,' 'retargetting-ad,' or 'retargetting-campaign' are the main reminder or re-targeting campaigns.

3. OPTIMIZE THE CAMPAIGN BUDGET: COOLTSHIRTS CAN RE-INVEST IN 5 CAMPAIGNS, WHICH SHOULD BE CHOSEN?

ft_attr.utm_source	ft_attr.utm_campaign	COUNT(*)
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

lt_attr.utm_source	lt_attr.utm_campaign	COUNT(*)
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-campaign	54
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

CoolTShirts should invest in the following five campaigns: ‘interview-with-cool-tshirts-founder,’ ‘getting-to-know-cool-tshirts,’ ‘ten-crazy-cool-tshirts-facts,’ ‘weekly-newsletter’ and ‘retargeting-ad’.

The three attracting campaigns of ‘interview-with-cool-tshirts-founder,’ ‘getting-to-know-cool-tshirts,’ and ‘ten-crazy-cool-tshirts-facts’ are all very effective in procuring site visits from three different sources, so they should be kept.

The two reminder campaigns of ‘weekly-newsletter,’ and ‘retargeting-ad’ should be kept since they are the most effective in procuring purchases from different sources, so they should be kept.

3. OPTIMIZE THE CAMPAIGN BUDGET: COOLTSHIRTS CAN RE-INVEST IN 5 CAMPAIGNS, WHICH SHOULD BE CHOSEN?

Marketing Plan:

- Best attracting/reminder campaigns.
- Variety of sources/platforms.
- Consider expansion of various campaigns onto even more sources/platforms.

Chosen Campaigns:

‘interview-with-cool-tshirts-founder’
‘getting-to-know-cool-tshirts’
‘ten-crazy-cool-tshirts-facts’
‘weekly-newsletter’
‘retargeting-ad’

Deciding Considerations:

- I chose the five best overall campaigns for procuring initial site visits and the most purchases on the most sources.
- I feel that since women are more likely to be shopping, they are more likely to take the time to read through / respond to a ‘weekly-newsletter’ that came via email.
- Men would respond more to a ‘retargeting-campaign’ on their email.
- If a sixth could be chosen, ‘retargeting-campaign’ would be added back in.
- The google campaigns were not very effective.