# code cademy

*Learn SQL from Scratch*

# First- and Last-Touch Attribution
## with CoolTShirts

Jill R. Foley
4 June 2019

# Background

CoolTShirts, an innovative apparel shop, sells shirts of all kinds — as long as they are T-shaped and cool!

Recently, CoolTShirts (CTS) started a few marketing campaigns to increase website visits and purchases. Using touch attribution, they'd like to map their customers' journey — from initial visit to purchase — and optimize their marketing campaigns by answering these questions:

1. **Get to know CoolTShirts**
   a. How many campaigns and sources does CoolTShirts use and how are they related?
   b. What pages are on their website?

# Background (cont'd)

2. **What is the user journey?**

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

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

    c. How many visitors make a purchase?

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

    e. What is the typical user journey?

3. **Optimize the campaign budget**

    a. CoolTShirts can reinvest in 5 campaigns. Which should they pick and why?

# 1. Get To Know CoolTShirts

# 1.1 Get to Know CoolTShirts

| Campaigns | Sources |
|---|---|
| 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 |

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

- Using SELECT COUNT(DISTINCT column_name) we can determine CTS is running **eight** distinct campaigns in **six** distinct sources.

| CTS Campaigns |
|---|
| 8 |

| CTS Sources |
|---|
| 6 |

- CTS is running two email campaigns and two Google search campaigns.

- The utm_campaign is the reference name of a specific advertising message or tactic that has a common goal – in this case to promote CTS and encourage prospective customers to by from CoolTShirts.

- The utm_source is the method used for that advertising. For example, CTS sends a campaign called "weekly newsletter" using "email," and they post a "retargeting ad" on "Facebook."

- The cart at right shows the relationship of each campaign to its source.

```sql
1   -- How many distinct campaigns does CTS use?
2   SELECT COUNT(DISTINCT utm_campaign) AS 'CTS Campaigns'
3   FROM page_visits;
4
5   -- How many distinct sources does CTS use?
6   SELECT COUNT(DISTINCT utm_source) AS 'CTS Sources'
7   FROM page_visits;
8
9   -- How are campaigns related to sources?
10  SELECT DISTINCT utm_campaign AS 'Campaigns',
11      utm_source AS 'Sources'
12  FROM page_visits;
13
```

# 1.2 Get to Know CoolTShirts (cont'd)

**What pages are on CoolTShirt's website?**

- Using SELECT DISTINCT we can determine CTS has four pages on their website:
    1. Landing page
    2. Shopping Cart
    3. Checkout
    4. Purchase

| Webpages |
| --- |
| 1 - landing_page |
| 2 - shopping_cart |
| 3 - checkout |
| 4 - purchase |

```sql
1    -- What pages are on the CoolTShirts website?
2    SELECT DISTINCT page_name AS 'Webpages'
3    FROM page_visits;
```

# 2. What is the User Journey?

# 2.1 What is the User Journey?

**How many FIRST touches is each campaign responsible for?**

- First-touch attribution is a good way of knowing how visitors initially discover a website. We say that those visits are *attributed* to the ad campaign.

- To find first-touch attribution, we first create a temporary table to return the earliest timestamp using MIN(timestamp) for each visitor from the page_visits table. We then join the page_visits table and the first_touch table on user_id and timestamp to determine the source and campaign for each first-touch. Finally, we group by campaign and source, and count the number of first-touches for each.

| 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 |

```sql
--How many first 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 AS 'Source',
       ft_attr.utm_campaign AS 'Campaign',
       COUNT(*) AS '# First Touches'
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

# 2.2 What is the User Journey? (cont'd)

## How many LAST touches is each campaign responsible for?

- Last-touch attribution is a good way of knowing how visitors are drawn back to a website, especially for making a final purchase.

- We find last-touch attribution in a similar way as first-touch attribution. Instead of returning the MIN(timestamp), though, we return the MAX(timestamp), meaning we find the last timestamp a prospective customer visits a page.

| 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 |

```sql
--How many last touches is each campaign responsible for?
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 AS 'Source',
      lt_attr.utm_campaign AS 'Campaign',
      COUNT(*) AS '# Last Touches'
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

# 2.3 What is the User Journey? (cont'd)

**How many visitors make a purchase?**

- To determine the number of visitors who make a purchase from CoolTShirts, we use COUNT(DISTINCT user_id) to count the unique user_ids for every person who visits the webpage called '4 – purchase.'

- **361** Visitors to CoolTShirts.com made a purchase

| Visitors Who Purchase |
|:---:|
| 361 |

```sql
1  --How many visitors make a purchase?
2  SELECT COUNT(DISTINCT user_id) AS 'Visitors Who Purchase'
3  FROM page_visits
4  WHERE page_name = '4 - purchase';
```

# 2.4 What is the User Journey? (cont'd)

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

- We start with our last-touch attribution query used to determine # of last-touches per campaign (Section 2.2).

- We then add a WHERE clause to last_touch temporary table — Line 6 of the query at right. This addition limits our results to only include last-touches that can be attributed to the '4 – purchase' page.

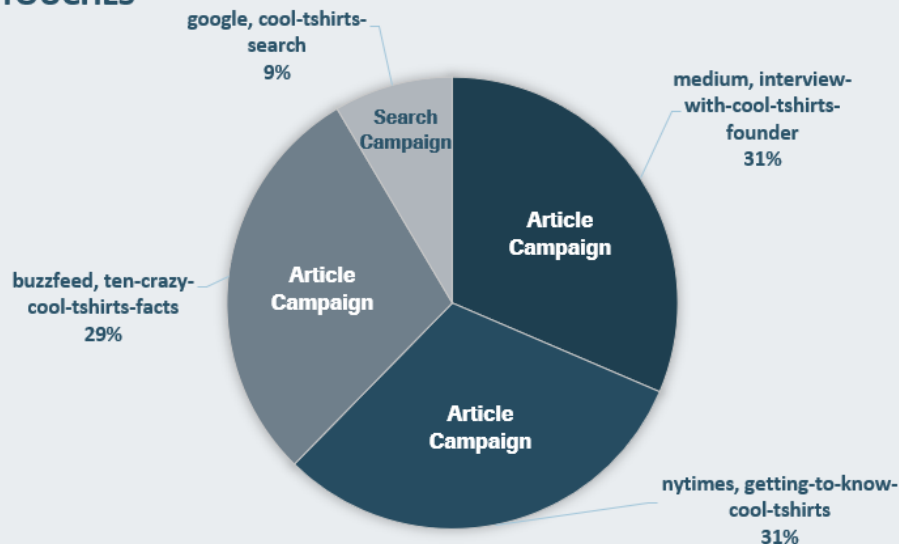| Source | Campaign | # Last Touches |
|--------|----------|----------------|
| email | weekly-newsletter | 115 |
| facebook | retargetting-ad | 113 |
| email | retargetting-campaign | 54 |
| google | paid-search | 52 |
| buzzfeed | ten-crazy-cool-tshirts-facts | 9 |
| nytimes | getting-to-know-cool-tshirts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

```sql
1   --How many last touches on the purchase page is each campaign
    responsible for?
2   WITH last_touch AS (
3       SELECT user_id,
4           MAX(timestamp) as last_touch_at
5       FROM page_visits
6       WHERE page_name = '4 - purchase'
7       GROUP BY user_id),
8   lt_attr AS (
9     SELECT lt.user_id,
10          lt.last_touch_at,
11          pv.utm_source,
12          pv.utm_campaign
13    FROM last_touch lt
14    JOIN page_visits pv
15      ON lt.user_id = pv.user_id
16      AND lt.last_touch_at = pv.timestamp
17  )
18  SELECT lt_attr.utm_source AS 'Source',
19         lt_attr.utm_campaign AS 'Campaign',
20         COUNT(*) AS '# Last Touches'
21  FROM lt_attr
22  GROUP BY 1, 2
23  ORDER BY 3 DESC;
```

# 2.5 What is the User Journey? (cont'd)

**What is the TYPICAL user journey?**

- The majority (91%) of users first visit CTS' website from an "Article" Campaign:

  1. *Ten Crazy Cool T-Shirt Facts*

  2. *Interview with CoolTShirts Founder*

  3. *Getting to Know CoolTShirts*

- The remaining 9% of first-touches are from organic Google searches

- Of the 8 campaigns CTS is running, only 4 of them attribute to first-touches

**FIRST-TOUCHES**
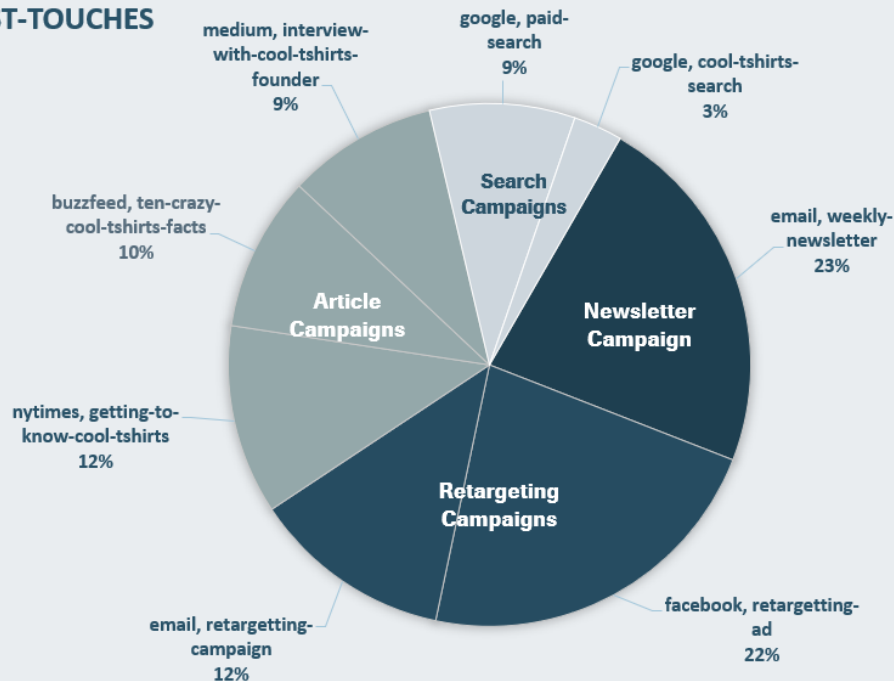


google, cool-tshirts-search 9%

medium, interview-with-cool-tshirts-founder 31%

Search Campaign

Article Campaign

buzzfeed, ten-crazy-cool-tshirts-facts 29%

Article Campaign

Article Campaign

nytimes, getting-to-know-cool-tshirts 31%

# 2.6 What is the User Journey? (cont'd)

**What is the TYPICAL user journey?**

- The majority (65%) of last-touches come from two types of campaigns:

  1. *Article Campaigns (31%)*
  2. *Retargeting Campaigns (34%)*

- Additionally, the Weekly Email Newsletter contributes to 23% of last touches

  o Because there are no first-touches from newsletter, we can assume everyone on the newsletter list is either an existing customers or has already visited from a search or article campaign and signed up to the list.

- Search campaigns only contribute to 11% of last-touch web traffic

**LAST-TOUCHES**

medium, interview-with-cool-tshirts-founder
9%

google, paid-search
9%

google, cool-tshirts-search
3%

buzzfeed, ten-crazy-cool-tshirts-facts
10%

Search Campaigns

email, weekly-newsletter
23%

Article Campaigns

Newsletter Campaign

nytimes, getting-to-know-cool-tshirts
12%

Retargeting Campaigns

email, retargetting-campaign
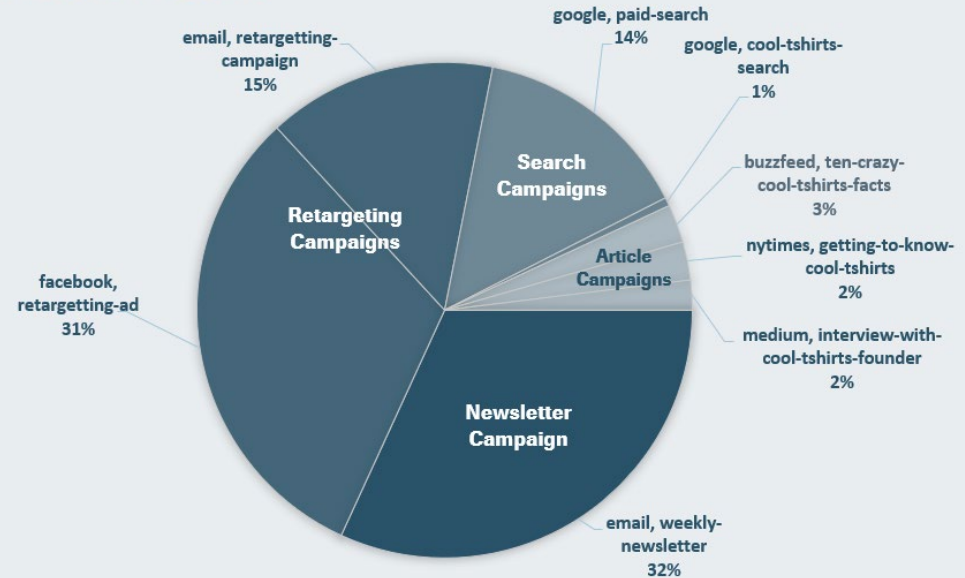12%

facebook, retargetting-ad
22%

# 2.7 What is the User Journey? (cont'd)

## What is the TYPICAL user journey?

- Looking specifically at what pages lead to a purchase (were visited last before purchase), we see that very few of the "Article" campaigns (7%) lead directly to a purchase even though they drove the majority of first-touch website traffic.

- Retargeting campaigns lead to the majority of purchases (45%)

- The weekly email newsletter also contributed to a large percentage of purchases (32%)

- We can deduce that the majority of purchases on CTS' website is not a direct first-touch to purchase journey.

- Instead, the *typical* user journey looks something like this:

**PURCHASE LAST-TOUCHES**



email, retargetting-campaign 15%

google, paid-search 14%

google, cool-tshirts-search 1%

buzzfeed, ten-crazy-cool-tshirts-facts 3%

nytimes, getting-to-know-cool-tshirts 2%

medium, interview-with-cool-tshirts-founder 2%

Search Campaigns

Retargeting Campaigns

Article Campaigns

facebook, retargetting-ad 31%

Newsletter Campaign

email, weekly-newsletter 32%

| [Prospective] customer reads an article about CTS | → | Customer clicks article directly to website (or does a Google search for CTS) | → | Customer browses T-Shirts and signs up for the newsletter (no purchase) | → | Customer sees a retargeting ad or receives the newsletter | → | Customer visits CTS site again and makes a purchase |

# 3. Optimize the Campaign Budget

# 3.1 Optimize the Campaign Budget

**CoolTShirts can reinvest in 5 campaigns. Which should they pick and why?**

- facebook-retargetting-ad (retargeting campaign)
- email-retargetting-campaign (retargeting campaign)
- ten-crazy-cool-tshirts-facts (article campaign)
- getting-to-know-cool-tshirts (article campaign)
- interview-with-cool-tshirts-founder (article campaign)

| Source | Campaign | First Touch | % First Touches | Last Touch | % Last Touches | Purchase | % Purchases |
|---|---|---|---|---|---|---|---|
| email | weekly-newsletter | 0 | 0% | 447 | 23% | 115 | 32% |
| facebook | **retargetting-ad** | 0 | 0% | 443 | 22% | 113 | 31% |
| email | **retargetting-campaign** | 0 | 0% | 245 | 12% | 54 | 15% |
| google | paid-search | 0 | 0% | 178 | 9% | 52 | 14% |
| google | cool-tshirts-search | 169 | 9% | 60 | 3% | 2 | 1% |
| buzzfeed | **ten-crazy-cool-tshirts-facts** | 576 | 29% | 190 | 10% | 9 | 3% |
| nytimes | **getting-to-know-cool-tshirts** | 612 | 31% | 232 | 12% | 9 | 2% |
| medium | **interview-with-cool-tshirts-founder** | 622 | 31% | 184 | 9% | 7 | 2% |

# 3.2 Optimize the Campaign Budget

**Reinvestment Rationale:**

- While the "Article" campaigns are not likely to lead to a purchase directly after website first-touch, they are essential for brand awareness.

- As we saw in the *typical* customer journey, customers are most likely to purchase after receiving a retargeting ad or the weekly email newsletter that reminds them of CTS and the product(s) they likely viewed during a prior website visit.

- Therefore, the "Articles and the "Rebranding" campaigns are most critical to sales by raising brand awareness and redirecting prospects back to CTS' website.

- The email newsletter is also an important campaign to redirect customers back to the site, but as an email marketer myself, I would argue that it is already doing it's job. Email is a fairly cheap medium, and once you have a newsletter established, there isn't much to reinvest in. I believe that money would be better spent continuing to raise brand awareness and retargeting customers who do not make a purchase on their first visit.

  - One might argue that CTS could buy email lists to increase their first-touch attribution from the email newsletter, but purchasing lists is not an email marketing best practice. Luckily CTS is not purchasing lists (and potentially emailing people who have not opted-in) since they do not have any first-touches from email.

  - By continuing to raise awareness of the brand and directing traffic from other sources, like Buzzfeed and NYTimes articles, to CTS' site, the email list will grow organically, helping to increase sales from the newsletter.

- CoolTShirts might also want to consider posting articles about CTS and retargeting ads on other web and social sites besides Buzzfeed, NYTimes and Facebook due to their proven effectiveness.