code cademy

Learn SQL from Scratch
Sarah Russell Cheung
August 27, 2018 – October 22, 2018

TABLE OF CONTENTS

1. Get familiar with the company.

- How many campaigns and sources does CoolTShirts use and how are they related? Be sure to explain the difference between `utm_campaign` and `utm_source`.
- What pages are on their website?

2. What is the user journey?

- How many first touches is each campaign responsible for?
- How many 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 they pick and why?

1. GET FAMILIAR WITH COOLTSHIRTS

1. How many campaigns and sources does CoolTshirts use? Which
    source is used for each campaign?

*   A source is used to identify which site sent the traffic, such as Google,
    Facebook, Buzzfeed, The New York Times etc.

*   A campaign identifies the specific add or email blast

QUERY RESULTS CONT'D

```sql
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;
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
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;
```

Campaign Count	Source Count
8	6

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

1.1 GET FAMILIAR WITH COOLTSHIRTS CONT'D

What pages are on the CoolTShirts website?

- Page names are the title of the section of the page that was visited

- You can find which pages are on the website by finding the DISTINCT values of the page_name FROM page_visits

QUERY RESULTS CONT'D

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

```sql
SELECT COUNT(DISTINCT page_name)
FROM page_visits;
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id)
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;
```

2. WHAT IS THE USER JOURNEY?

How many first touches is
each campaign responsible
for?

- Finding the first touch of
  a campaign can simply
  be found by using
  timestamps with MIN
  date for each user and
  grouping by the user_id

```sql
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_campaign,
COUNT(*)
FROM ft_attr
GROUP BY utm_campaign
ORDER BY 2 DESC;
```

Query Results	
ft_attr.utm_campaign	**COUNT(*)**
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576
cool-tshirts-search	169

2. WHAT IS THE USER JOURNEY CONT'D

How many last touches is each campaign responsible for?

- Finding the last touch of a campaign can simply be found by using timestamps with MAX date for each user and grouping by the user_id

QUERY RESULTS CONT'D

Query Results	
lt_attr.utm_campaign	COUNT(*)
weekly-newsletter	447
retargetting-ad	443
retargetting-campaign	245
getting-to-know-cool-tshirts	232
ten-crazy-cool-tshirts-facts	190
interview-with-cool-tshirts-founder	184
paid-search	178
cool-tshirts-search	60

```sql
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_campaign,
COUNT(*)
FROM lt_attr
GROUP BY utm_campaign
ORDER BY 2 DESC;
```

2. WHAT IS THE USER JOURNEY CONT'D

How many visitors make a purchase?

- By using a COUNT(DISTINCT user_id) command you can see that 1,979 visitors make a purchase

```
1   SELECT COUNT(DISTINCT user_id) AS 'Visitors
    that Purchase'
2   FROM page_visits
3   WHERE page_name - '4 - purchase';
```

Query Results
Visitors that Purchase
1979

2. WHAT IS THE USER JOURNEY CONT'D

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

- Finding the last touch of a campaign can simply be found by using timestamps with MAX date for each user and grouping by the user_id

```sql
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_utm 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 COUNT(*),
    lt_utm.utm_campaign,
       lt_utm.utm_source
FROM lt_utm
GROUP BY 2
ORDER BY 1 DESC;
```

QUERY RESULTS CONT'D

COUNT(*)	lt_utm.utm_campaign	lt_utm.utm_source
115	weekly-newsletter	email
113	retargetting-ad	facebook
54	retargetting-campaign	email
52	paid-search	google
9	getting-to-know-cool-tshirts	nytimes
9	ten-crazy-cool-tshirts-facts	buzzfeed
7	interview-with-cool-tshirts-founder	medium
2	cool-tshirts-search	google

Query Results

2. WHAT IS THE USER JOURNEY CONT'D

What is the typical user journey?

- Most users' first touch is typically through Getting to know cool t-shirts, Interview with cool t-shirts founder, and Ten crazy cool t-shirts facts
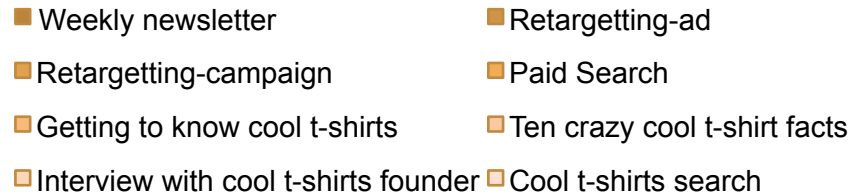
First Touches by Campaign

Getting to know cool t-shirts 31%

Interview with cool t-shirts founder 31%

Cool tshirts search 9%

Ten crazy cool tshirts facts 29%
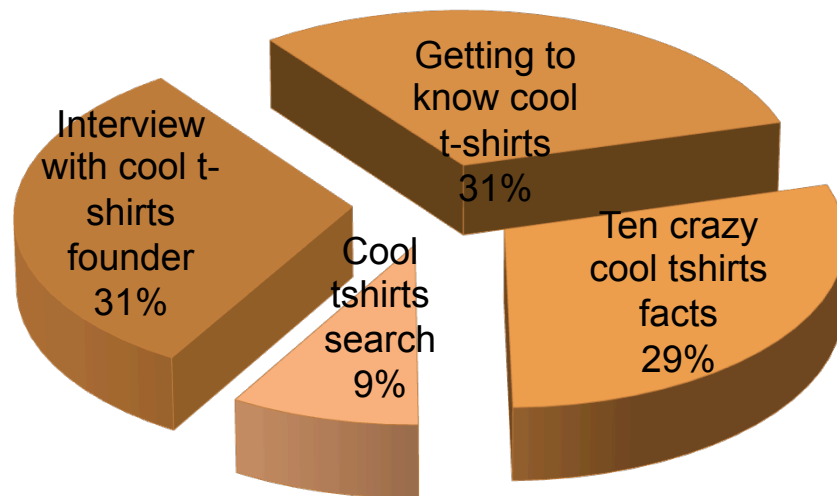
2. WHAT IS THE TYPICAL USER JOURNEY CONT'D

- In the results you can see that most customers went through different campaigns than the four listed from the first touches.

- The most success was through second campaigns to get the customers to make a final purchase

Last Touches by Campaign

- Weekly newsletter
- Retargetting-ad
- Retargetting-campaign
- Paid Search
- Getting to know cool t-shirts
- Ten crazy cool t-shirt facts
- Interview with cool t-shirts founder
- Cool t-shirts search

31%
15%
32%
14%
1%
2% 2% 3%

3. OPTIMIZING THE CAMPAIGN BUDGET

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

  - From the data we can see that the 3 first touch campaigns (Getting to know cool t-shirts, Interview with cool t-shirts founder, and Ten crazy cool t-shirt facts) returned a success rate of 31%, 31%, and 29% respectively.

  - As seen from the last touch results, the 2 campaigns (Weekly newsletter and Retargetting-ad) returned a high success rate of 32% and 31%.

  - Thus, CoolTshirts should make a re-investment with the 5 listed above campaigns. The campaigns brought traffic and 3 remained from the first to last touch story.