code|cademy

First-and Last-Touch Attribution Capstone
Learn SQL from Scratch
Ted Flavion
07/17/2018

Table of Contents

1. Get familiar with CoolTShirts

- 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?

Ted Flavion – Learn SQL from Scratch Capstone

1. Get familiar with CoolTShirts

1.1 Get Familiar with CoolTShirts

How many campaigns and sources does CoolTShirts use and how are they related?
- The utm_source is which site/source the campaign is sent to. Some of the examples from CoolTShirts include New York Times (nytimes), Buzzfeed, Google, Facebook, Medium and Email.
- The utm_campaign is the specific ad or email blast that those customers see when they are on that source.
- Some sources do have multiple campaigns running at the same time.

```
1   SELECT COUNT (DISTINCT utm_campaign) AS 'Campaign Count'
2   FROM page_visits;
3
4   SELECT COUNT (DISTINCT utm_source) AS 'Source Count'
5   FROM page_visits;
6
7   SELECT DISTINCT utm_campaign AS Campaigns,
8       utm_source AS Sources
9   FROM page_visits;
```

Query Results	
Campaign Count	
8	
Source Count	
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

4

1.2 What pages are on their website?

By using the DISTINCT page_names from the page_visits tables queary, finding the names of the pages on CoolTShirts is very simple.

This query may not be showing all of the actual pages on the site (no product pages), we can assume these are the pages that are currently showing the campaigns for CoolTShirts.

```
1  SELECT DISTINCT page_name AS 'Page Names'
2  FROM page_visits;
```

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

Ted Flavion – Learn SQL from Scratch Capstone

2. What is the User Journey?

2.1 What is the User Journey?

How many first touches is each campaign responsible for?
- 1979 unique visitors saw 4 campaigns as first touch. Only 4 of the 8 campaigns currently have first touch data.
- By using MIN timestamps, we are able to find the first touch campaigns. These campaigns may have made the customer visit CoolTShirts site.

```
     test.sqlite
1    WITH first_touch AS (
2        SELECT user_id,
3            MIN(timestamp) as first_touch_at
4        FROM page_visits
5        GROUP BY user_id),
6
7    ft_attr AS (
8        SELECT ft.user_id,
9            ft.first_touch_at,
10           pv.utm_source,
11           pv.utm_campaign
12       FROM first_touch ft
13       JOIN page_visits pv
14         ON ft.user_id = pv.user_id
15         AND ft.first_touch_at = pv.timestamp
16   )
17   SELECT ft_attr.utm_source AS Source,
18          ft_attr.utm_campaign AS Campaign,
19          COUNT(*)
20   FROM ft_attr
21   GROUP BY 1, 2
22   ORDER BY 3 DESC;
```

Query Results

Source	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

Database Schema

page_visits		5692 rows
page_name		TEXT
timestamp		TEXT
user_id		INTEGER
utm_campaign		TEXT
utm_source		TEXT

2.2 What is the User Journey?

How many last touches is each campaign responsible for?

- 8 of the 8 campaigns currently have last touch data. Only 4 have first touch data.
- By using MAX timestamps, we are able to find the last touch campaigns. These campaigns may explain the customers LAST visit to CoolTShirts site.
- Of the 4 campaigns running with first touch data, only Google has more LAST touch data (Buzzfeed, Medium & Nytimes all have significantly less last touches than first touches.

```
     test.sqlite
1
2    WITH last_touch AS (
3        SELECT user_id,
4            MAX(timestamp) as last_touch_at
5        FROM page_visits
6        GROUP BY user_id),
7
8    ft_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 ft_attr.utm_source AS Source,
19        ft_attr.utm_campaign AS Campaign,
20        COUNT(*) AS Count
21    FROM ft_attr
22    GROUP BY 1, 2
23    ORDER BY 3 DESC;
```

Query Results

Source	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

Database Schema

page_visits		5692 rows
page_name		TEXT
timestamp		TEXT
user_id		INTEGER
utm_campaign		TEXT
utm_source		TEXT

2.3 What is the User Journey?

How many visitors make a purchase?
- 361 visits have turned into purchases
- Conversion rate of 18% (361/1979)

```sql
SELECT COUNT (DISTINCT user_id) AS 'Customers Purchases'
FROM page_visits
WHERE page_name = '4 - purchase';
```

Query Results	
Customers Purchases	
361	
Database Schema	
page_visits	5692 rows
page_name	TEXT
timestamp	TEXT
user_id	INTEGER
utm_campaign	TEXT
utm_source	TEXT

Ted Flavion – Learn SQL from Scratch Capstone

2.4 What is the User Journey?

How many last touches on the **purchase page** is each campaign responsible for?
- Email campaigns have the most purchases by customers, as both email campaigns have had 169 of the 361 purchases.
- Print campaigns seem to have the least amount of conversions tied to them.

```
1   WITH last_touch AS (
2       SELECT user_id,
3           MAX(timestamp) AS last_touch_at
4       FROM page_visits
5       WHERE page_name = '4 - purchase'
6       GROUP BY user_id),
7
8   ft_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
19  SELECT ft_attr.utm_source AS Source,
20          ft_attr.utm_campaign AS Campaign,
21          COUNT(*) AS COUNT
22  FROM ft_attr
23  GROUP BY 1, 2
24  ORDER BY 3 DESC;
```

Query Results

Source	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

Database Schema

page_visits	5692 rows
page_name	TEXT
timestamp	TEXT
user_id	INTEGER
utm_campaign	TEXT
utm_source	TEXT

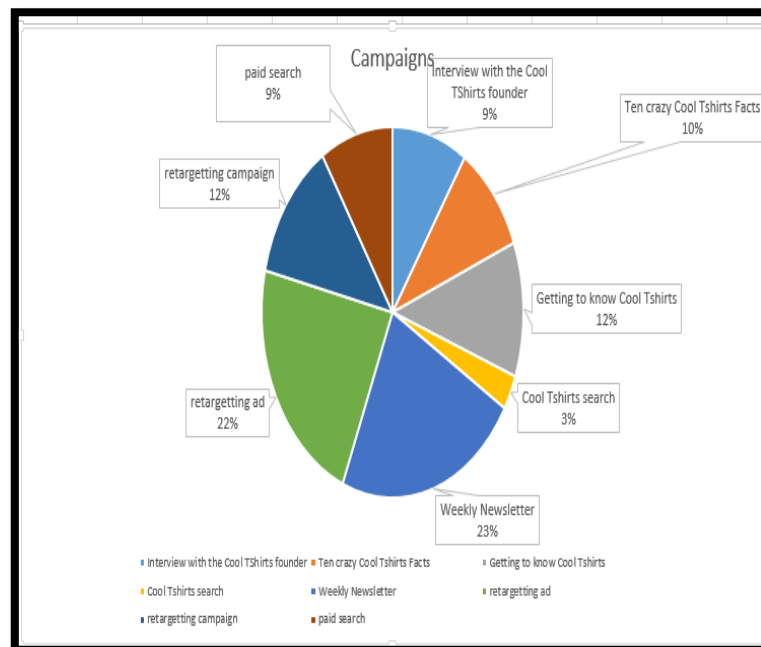# 2.5 What is the Typical User Journey?

First Touch Campaigns:
- 91% of users come from one of three campaigns:
    - Ten Crazy Cool TShirt facts
  - Interview with Cool TShirts Founder
    - Getting to Know Cool TShirts
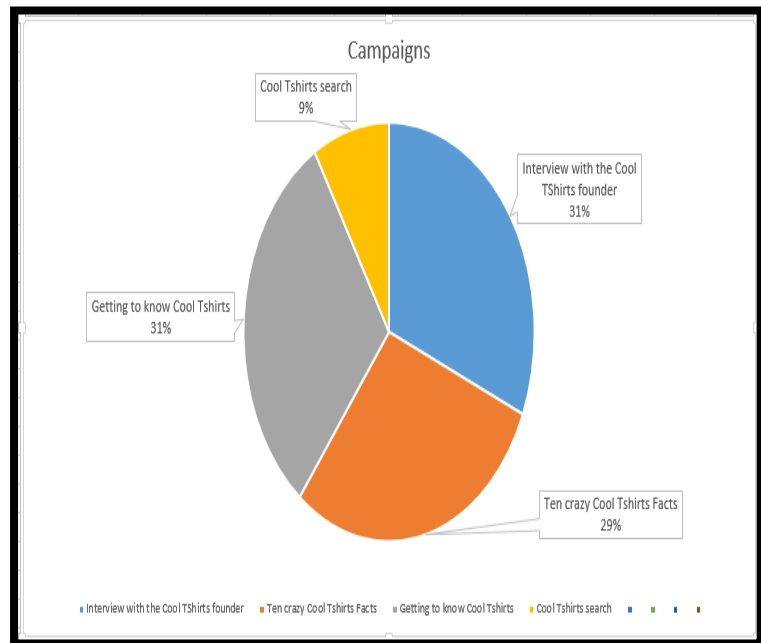
Last Touch Campaigns:
- First touch campaigns only account for 31% of all last touch campaign visits.
- Retargeting ad's, weekly newsletters and retargeting campaign have the best results when getting customers back to the site.

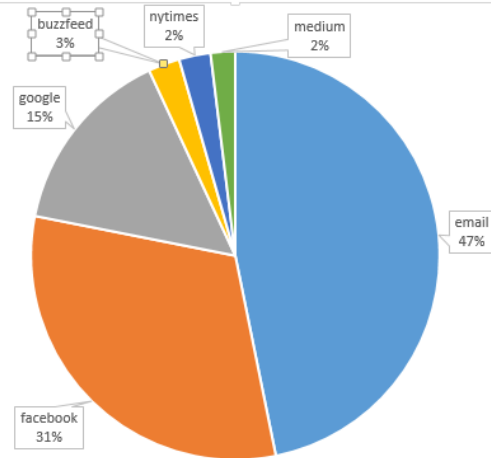# 2.5 What is the Typical User Journey?

Purchases by Campaign:
- Retargeting campaigns along with the weekly newsletter provide most conversions from visitors.
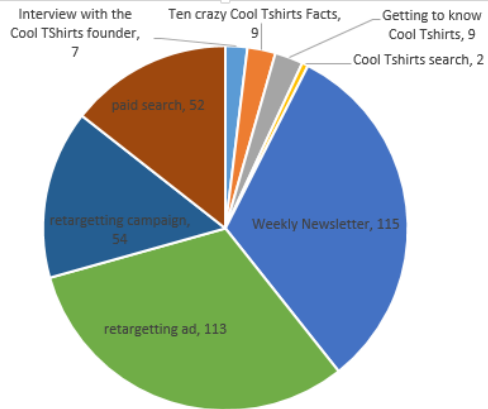
Purchases by Source:
- Email, Facebook and Google are the best sources for campaigns to be converted into orders.

12

3. Optimize the Campaign Budget

3.1 Overall Campaign Breakdown

For a better look at how the current campaigns are performing, take a look at the data below. After exploring this data, we can make our recommendations to Cool TShirts on which campaigns we should re-invest in.

Campaign	1st Touches	% of 1st Touches	Last Touches	% of Last Touches	Purchases	% of Purchases
Interview with founder (medium)	622	31%	184	9%	7	2%
Getting to know Cool Tshirts (nytimes)	612	31%	232	12%	9	2%
Ten crazy facts (buzzfeed)	576	29%	190	10%	9	2%
Cool Tshirts search (google)	169	9%	60	3%	2	1%
Weekly Newsletter (email)	0	0%	447	23%	115	32%
Retargeting Ad (facebook)	0	0%	443	22%	113	31%
Retargeting Campaign (email)	0	0%	245	12%	54	15%
Paid Search (google)	0	0%	178	9%	52	14%

Ted Flavion – Learn SQL from Scratch Capstone

3.2 Re-invest Recommendations

By looking at the data, the % of last touches for each campaign are good, however, the % of purchases made per campaign does not necessarily reflect the amount of touches for that campaign. Most of our session require a 2nd campaign (last touch) before they turn into customers:
 - Print ads (Buzzfeed, NY Times, Medium) are getting good touches, however, these touches are not turning into purchases (only 6% of purchases come from these campaigns yet they account for 91% of the first-touch results for our visitors).
- Retargeting campaigns may not have any 1st touch results, however, these campaigns are converting orders, accounting for 46% of all purchases. Cool TShirts should continue with these ads ad converting visits to sales is the goal.
- There are currently two paid Google campaigns, the Cool TShirt search and the Paid Search. However, the paid search campaign is converting sessions into orders at a better clip. I recommend Cool TShirts keep their current paid search as well as look into the Cool TShirt's seach campaign. Increasing the SEO of the site may also help with customers finding Cool TShirts through the Google Search campaign.
- The weekly newsletter is also showing to be a good source of order conversions as 32% of purchases come from this campaign. With many of our prospects needing a second campaign to become customers, sending this weekly email out more frequently may help increase sales.

Below are the 5 campaigns I recommend for Cool TShirts to reinvest in.

Campaign	1st Touches	% of 1st Touches	Last Touches	% of Last Touches	Purchases	% of Purchases
Cool Tshirts search (google)	169	9%	60	3%	2	1%
Weekly Newsletter (email)	0	0%	447	23%	115	32%
Retargeting Ad (facebook)	0	0%	443	22%	113	31%
Retargeting Campaign (email)	0	0%	245	12%	54	15%
Paid Search (google)	0	0%	178	9%	52	14%