CoolTShirts

First- and Last-Touch Attribution Project

2 July 2018

Table of contents

1. An Overview of CoolTShirts
2. What is the user journey?
3. Optimizing the campaign budget

1. An Overview of CoolTShirts

1.1 Who Is CoolTShirts?

- CoolTShirts is an online apparel store specialising in cool, funky T-shirts

- CoolTShirts has started several marketing campaigns to increase its website traffic volumes and purchases

- It would like to optimize its campaigns to get the best improvement in both volumes and translation to purchases for the funds available

1.1 CoolTShirts UTM parameters

CoolTShirts has developed a database of UTM parameters that captures when and how users find the CoolTShirts site. The database is an SQL table called page_visits, with 5692 records and five columns (schema tabled below). The columns are:

- user_id. A unique identifier for each visitor to a page.
- timestamp. The time at which the visitor came to the page.
- page_name. The title of the section of the page that was visited.
- utm_source. Identifies which site sent the traffic (e.g. facebook, google, nytimes).
- utm_campaign. Identifies the specific ad or email blast marketing campaign (e.g. weekly-newsletter, cool-tshirts-search)

Column name	Data type
user_id	INTEGER
timestamp	TEXT
page_name	TEXT
utm_source	TEXT
utm_campaign	TEXT

1.2 Understanding the data (1)

- An inspection of CoolTShirt's page_visits table reveals the following:
  - CoolTShirts uses eight distinct campaigns
  - It uses six distinct sources (or touchpoints)
- The table below shows the relationship between CoolTShirts' marketing campaigns and its sources. A source can run multiple campaigns. For example, email runs the weekly-newsletter and retargetting-campaign campaigns.

Marketing Campaign	Source
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 does CoolTShirts use?
SELECT COUNT(DISTINCT utm_campaign)
FROM page_visits;

How many sources does CoolTShirts use?
SELECT COUNT(DISTINCT utm_source)
FROM page_visits;

--How are CoolTShirts campaigns and sources related?
SELECT DISTINCT utm_campaign AS 'Marketing Campaign',
  utm_source AS 'Source'
FROM page_visits;
```

1.2 Understanding the data (2)

- The page_visits table shows that there are four distinct pages on the CoolTShirts website. The pages are listed below.

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

```
--What pages are on the CoolTShirts website?
SELECT DISTINCT page_name AS 'Page Name'
FROM page_visits;
```

2. What is the user journey?

2.1 How many <u>first</u> touches is each campaign responsible for?

- The table below shows the number of first touches by marketing campaign (and associated source). There are a total of 1,979 first touches
- Only four of the eight campaigns are responsible for first touches
- The top three of these campaigns are responsible for well over 91% of first touches.

Source	Marketing 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

```
--How many first touches is each campaign responsible for?
--The source for each campaign is also identified in this query.

--Create a temporary table (first_touch) that identifies the timestamp
--for each user's first touch.
WITH first_touch AS (
    SELECT user_id,
           MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id),
--Create a temporary table (ft_attribute) that combines source,
--campaign from page_visits table with user_id and first_touch_at from
--first_touch,joined on user_id and timestamp.  This table enables the
--source to be identified in the query results.
ft_attribute AS (
    SELECT ft.user_id,
           ft.first_touch_at,
           pv.utm_source,
           v.utm_campaign
    FROM first_touch ft
    JOIN page_visits pv
      ON ft.user_id = pv.user_id
     AND ft.first_touch_at = pv.timestamp
)
--Count the rows where the first touch is associated with a source and
--campaign. Order from highest to lowest count.
SELECT ft_attribute.utm_source AS 'Source',
       ft_attribute.utm_campaign AS 'Marketing Campaign',
       COUNT(*) AS 'Count'
FROM ft_attribute
GROUP BY 1, 2
ORDER BY 3 DESC;
```

2.2 How many <u>last</u> touches is each campaign responsible for?

- The table below shows the number of last touches by marketing campaign (and associated source)
- Two of the campaigns are responsible for many (45%) of the last touches (weekly-newsletter, retargetting-ad)
- Note that search-based campaigns using google are responsible for relatively few last touches
- The campaigns responsible for first touches fall in the middle of the pack.

Source	Marketing 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

```
--How many last touches is each campaign responsible for?
--The utm_source for each campaign is also identified in this query.

--Create a temporary table (last_touch) that identifies the timestamp
--for each user's last touch.
WITH last_touch AS (
    SELECT user_id,
           MAX(timestamp) AS 'last_touch_at'
      FROM page_visits
     GROUP BY user_id),
--Create a temporary table (lt_attribute) that combines source,
--campaign from page_visits table with user_id and last_touch_at from
--last_touch,joined on user_id and timestamp.  This table enables the
--source to be identified in the query results.
lt_attribute 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
)
--Count the rows where the last touch is associated with a source and
--campaign. Order from highest to lowest count.
SELECT lt_attribute.utm_source AS 'Source',
       lt_attribute.utm_campaign AS 'Marketing Campaign',
       COUNT(*) AS 'Count'
FROM lt_attribute
GROUP BY 1, 2
ORDER BY 3 DESC;
```

2.3 How many visitors make a purchase?

- For the sample of data contained in page_visits, 361 visitors make a purchase. Given that there are 1,979 initial touches, this implies a conversion rate of 18%
- The table below identifies the campaigns that drive these last touches where purchases are made
- The top two campaigns are responsible for 63% of the 361 purchases. The top four campaigns are responsible for 93% of purchases.

Source	Marketing 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

```
--How many visitors make a purchase?
--Count the distinct users who visited the page named 4 - purchase.
SELECT COUNT(DISTINCT user_id) AS 'Visitors that make a purchase'
FROM page_visits
WHERE page_name ='4 - purchase';


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

--Create a temporary table (last_touch) that identifies the timestamp
--for each user's last touch.  This table now filters by last touches where
--the page was '4 - purchase'
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),
--Create a temporary table (lt_attribute) that combines source,
--campaign from page_visits table with user_id and last_touch_at from
--last_touch,joined on user_id and timestamp.
lt_attribute AS (
  SELECT lt.user_id,
         lt.last_touch_at,
         pv.utm_source,
         pv.utm_campaign,
         pv.page_name
  FROM last_touch lt
  JOIN page_visits pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
--Count the rows where the last touch is associated with a source and
--campaign. Order from highest to lowest count.
SELECT lt_attribute.utm_source AS 'Source',
       lt_attribute.utm_campaign AS 'Marketing Campaign',
       COUNT(*) AS 'Count'
FROM lt_attribute
GROUP BY 1, 2
ORDER BY 3 DESC;
```
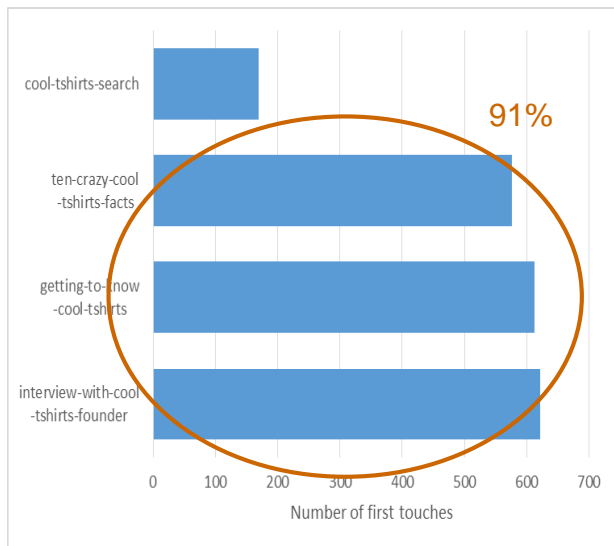
2.4 What is the _typical_ user journey?

First touches are driven by story-based campaigns…

- Four campaigns are responsible for first touches
- Of these, 91% of first touches are generated by three story-based campaigns

… but visitors require further prompting to stay on the site

- Two campaigns drive 45% of last touches on the site
- The campaigns responsible for first touches drive only 34% of last touches
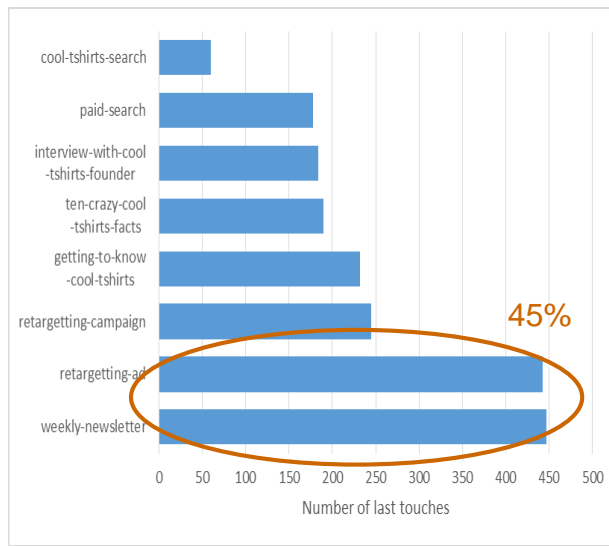
… and those further prompts deliver most of the sales

- The top three campaigns that deliver most of the last touches also drive most of the final purchases (78%)
- Overall, 18% of first touches to the site translate into sales
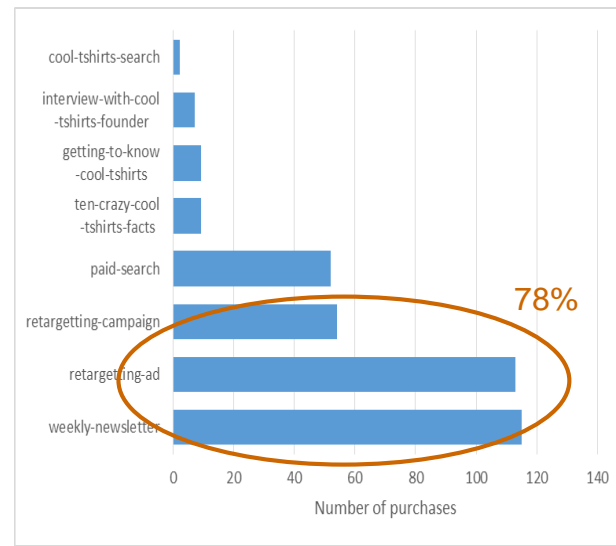
3. Optimizing the campaign budget

3.1 Which five campaigns should CoolTShirts re-invest in?

- If the limit is five campaigns, it is recommended that:
  - The three story-based campaigns highlighted in green are continued, since they are all good at delivering first touches
  - The top two campaigns that drive purchases (highlighted in blue) should be continued.
- As a portfolio, these five campaigns cover all the bases: they generate first and final touch interest and final purchases.
- If CoolTShirts has additional funds at some stage it should consider also re-investing in the retargeting-campaign campaign since it delivers a high proportion of last touches (12%) as well as a reasonable percentage of purchases (15%).

Campaign	First touches	Last touches	Purchases
interview-with-cool-tshirts-founder	622	184	7
getting-to-know-cool-tshirts	612	232	9
ten-crazy-cool-tshirts-facts	576	190	9
cool-tshirts-search	169	60	2
weekly-newsletter	0	447	115
retargetting-ad	0	443	113
retargetting-campaign	0	245	54
paid-search	0	178	52