code_cademy

Analysis of First- and Last-Touch Attribution for CoolTShirts Website

Learn SQL from Scratch

Matt McNearney

August 31, 2018

Table of Contents

1. Background

1-1 Background: Database Schema

The CoolTShirts company has been implemented several marketing campaigns at various sources with a goal to bring customers to its site and ultimately to buy products. The intent of this report is to show an analysis of which campaigns and which sources have yielded the best results. This will help CoolTShirts focus it future marketing efforts in the most effective manner.

To track the results of these various campaigns and sources, the company captured data in a page_visits table, using the following schema.

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

1-2 Background: Sample Data

A sample of the data collected is shown below.

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

```
-- Query used to show a sample of the data

SELECT *
FROM page_visits
LIMIT 10;
```

1-3 Background: Campaigns and Sources

CoolTShirts conducted a total of 8 marketing campaigns in 6 different sources. Campaigns were tracked using utm_campaign, which identifies the marketing tool used to draw the customer to the CoolTShirt website. The location from which the customer came to the website was tracked using utm_source.

The campaigns and the sources where each of those campaigns was conducted are as follows:

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

```
-- Query used to count utm_campaign

SELECT COUNT(DISTINCT utm_campaign)
FROM page_visits;

-- Query used to count utm_source

SELECT COUNT(DISTINCT utm_source)
FROM page_visits;

-- Query used to show relationship between campaigns
and sources

SELECT DISTINCT utm_campaign, utm_source
FROM page_visits
ORDER BY 2 ASC;
```

1-4 Background: Website Pages

The data used to analyze these campaigns came from visits and transactions on the CoolTShirt website, which consists of four pages:

Query Results
page_name
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

```
-- Query used to identify distinct pages within the
website

SELECT DISTINCT page_name
FROM page_visits;
```

2. Analysis

2-1 Analysis: First Touch

Four campaigns were used to bring visitors to the CoolTShirts websites. They are listed below along with the sources where the campaigns were placed and the number of visitors generated by each campaign.

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

```
-- Query used to identify first-touches and counts for
campaigns that brought initial visits to the website

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

2-2 Analysis: Last Touch

The overall result of each campaign is shown below, with the total counts of users whose last touch was the result of the respective campaign.

	Query Results	
lt_attr.utm_campaign	**lt_attr.utm_source**	**COUNT(*)**
weekly-newsletter	email	447
retargeting ad	facebook	443
retargeting-campaign	email	245
getting-to-know-cool-tshirts	nytimes	232
ten-crazy-cool-tshirts-facts	buzzfeed	190
interview-with-cool-tshirts-founder	medium	184
paid-search	google	170
cool-tshirts-search	google	60

```
-- Query used to identify last-touches and counts for
each campaign

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,
        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)

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

2-3 Analysis: Users Making Purchases

There were a total of 361 unique users who made purchases on the site.

Query Results
COUNT(DISTINCT user_id)
361

```
-- Query used to identify users making purchases

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

2-4 Analysis: Last Touches on Purchase Page

Below is a list that shows the total number of purchases (last touch on purchase page) associated with each campaign.

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

```sql
-- Query used to identify users making purchases

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,
        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)

SELECT lt_attr.utm_campaign,
    lt_attr.utm_source,
    COUNT(*)
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```
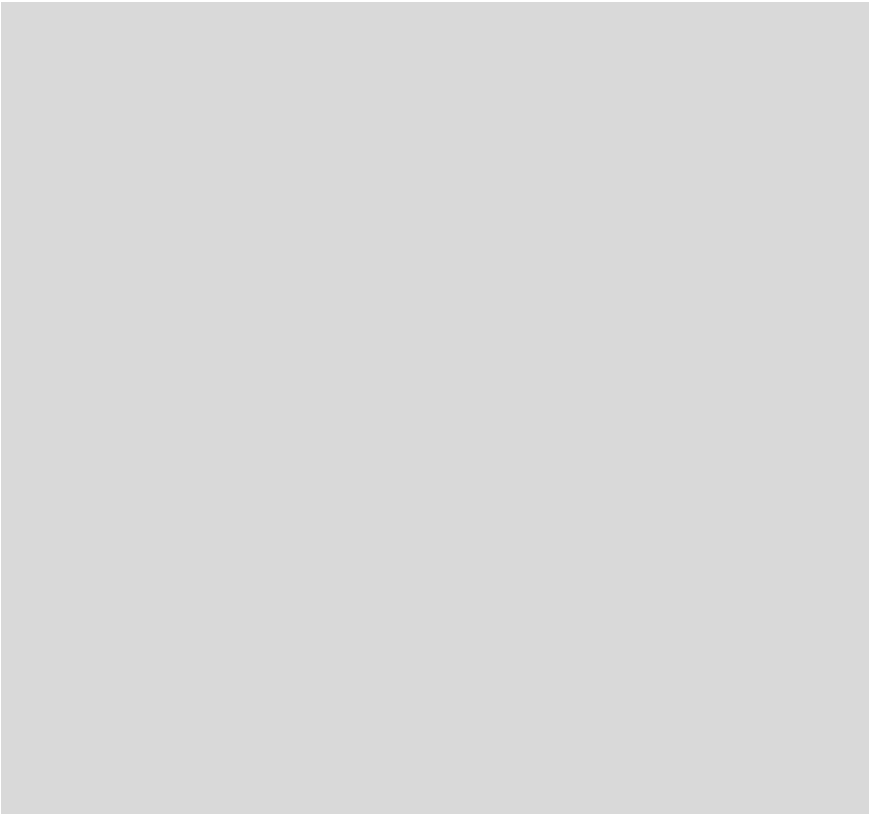
3. Conclusions

3-1 Conclusion: Typical Journey-First Touch

The journey of a typical user involves learning about CoolTShirts through an initial campaign place in an outside source, such as Medium, the New York Times, Buzzfeed, and Google. The result of these campaigns is to bring new visitors to the CoolTShirt website. This is shown in the first-touch table.
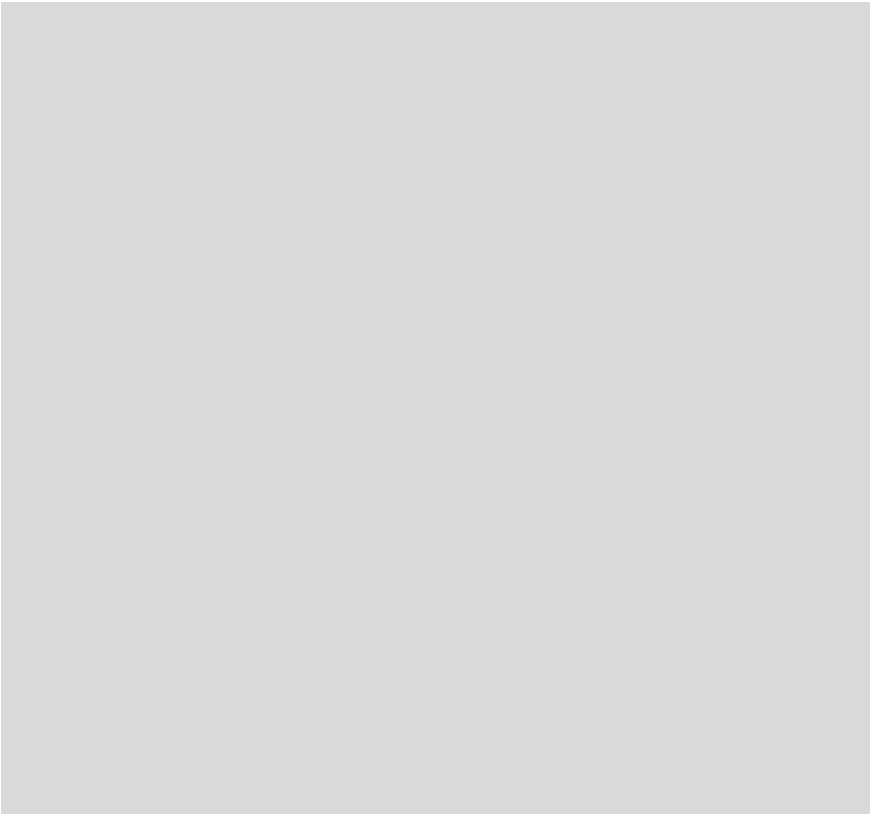
Query Results		
ft_attr.utm_campaign	**ft_attr.utm_source**	**COUNT(*)**
interview-with-cool-tshirts-founder	medium	622
getting-to-know-cool-tshirts	nytimes	612
ten-crazy-cool-tshirts-facts	buzzfeed	576
cool-tshirts-search	google	169

3-2 Conclusion: Typical Journey-Last Touch

After the initial visit, the user remains engaged with CoolTShirts through targeted follow-up campaigns, the result of which often brings the user back to make a purchase.

Query Results

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

3-3 Conclusion: Campaigns in which to Re-Invest

If CoolTShirts can re-invest in only five campaigns, it might consider going with the following:

	Campaign	Source	Objective
1	interview-with-cool-tshirts-founder	Medium	Bring visitors to website
2	getting-to-know-cool-tshirts	New York Times	Bring visitors to website
3	ten-crazy-cool-tshirts-facts	Buzzfeed	Bring visitors to website
4	weekly-newsletter	Email	Follow-up to initial website visit
5	retargeting-ad	Facebook	Follow-up to initial website visit