code|cademy

First- and Last-Touch Attribution with CoolTShirts.com

Learn SQL from Scratch
John L. Clark
July 28, 2018

Project Objective

This project's objective is to:

- Introduce and familiarize the company

- Explore the user journey

- Optimize the marketing campaign budget

Introducing CoolTShirts.com

CoolTShirts.com

CoolTShirts.com is an innovative apparel shop, is running a number of marketing campaigns. There is a need to optimize the economics of marketing efforts and to bring understanding to how existing campaigns are working to provide focus for future marketing endeavors.

CoolTShirts.com Campaigns and Sources

CoolTShirts.com employs a database to capture campaign and source metrics. This database includes a file (page_visits) of user visits with these attributes:

o   Unique identifier for each visitor to a webpage

o   The time the visitor came to the page

o   The title of the section of the page visited

o   Source identifying which site sent the traffic

o   Identification of the type of link used

o   Identification of the specific ad or email blast

CoolTShirts.com Campaigns

This query of the user visits database lists the distinct CoolTShirts.com campaigns.

```
SELECT COUNT(DISTINCT utm_source) AS 'Count of distinct campaigns'
FROM page_visits;
```

Query Results
Count of distinct campaigns
6

This query of the user visits database lists the distinct CoolTShirts.com sources.

```
SELECT COUNT(DISTINCT utm_campaign) AS 'Count of distinct sources'
FROM page_visits;
```

Query Results
Count of distinct sources
8

CoolTShirts.com Campaigns (continued)

This query relates CoolTShirts.com campaigns and sources.

```sql
SELECT DISTINCT utm_campaign AS 'Campaign',
    utm_source AS 'Source'
FROM page_visits;
```

Query Results	
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

CoolTShirts.com Campaigns (continued)

This query shows the pages on the CoolTShirts.com website.

```
select page_name as 'Web Page Name' from page_visits group
by 1 order by 1;
```

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

Exploring the User Journey

The User Journey

Understanding the user journey sheds light on the effectiveness of the marketing past campaigns. It is important to fully understand:

- 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?

CoolTShirts.com First Touch by Campaign

This query shows the first touch counts by campaign

```sql
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) AS 'first_touch_at'
    FROM page_visits
    GROUP BY user_id)
SELECT pv.utm_campaign AS 'Campaign', pv.utm_source AS 'Source',
    COUNT(*) AS 'Count of First Touches'
FROM first_touch AS 'ft'
JOIN page_visits AS 'pv'
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
GROUP BY 1, 2 ORDER BY 3 DESC;
```

Query Results		
Campaign	Source	Count of First Touches
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

CoolTShirts.com Last Touch by Campaign

This query shows the last touch counts by campaign

```sql
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) AS 'last_touch_at'
    FROM page_visits
    GROUP BY user_id)
SELECT pv.utm_campaign AS 'Campaign',
    pv.utm_source AS 'Source',
    COUNT(*) AS 'Count of Last Touches'
FROM last_touch AS 'lt'
JOIN page_visits AS 'pv'
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
GROUP BY 1, 2 ORDER BY 3 DESC;
```

Query Results		
Campaign	Source	Count of Last Touches
weekly-newsletter	email	447
retargetting-ad	facebook	443
retargetting-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	178
cool-tshirts-search	google	60

CoolTShirts.com Visitors Making a Purchase

This query shows the number of users who made a purchase – by counting the distinct users who visited the purchase web page

```sql
SELECT COUNT(DISTINCT user_id) AS 'Number of Visitors Making a Purchase'
FROM page_visits
WHERE page_name = '4 - purchase';
```

Query Results
Number of Vistors Making a Purchase
361

CoolTShirts.com Purchases by Campaign

This query shows the number purchases by campaign by counting last touches on the purchase page – grouped by campaign.

```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)
SELECT pv.utm_campaign AS 'Campaign',
    pv.utm_source AS 'Source',
    COUNT(*) AS 'Purchases'
FROM last_touch AS 'lt'
JOIN page_visits AS 'pv'
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
GROUP BY 1, 2 ORDER BY 3 DESC;
```

Query Results		
Campaign	Source	Purchases
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

Optimizing the CoolTShirts.com Campaign Budget

Optimizing the CoolTShirts.com Marketing Campaign Budget

Based on the analysis of the traffic on the CooTShirts.com website, we can draw these observations:

o   The cool-tshirts-search campaign drew the fewest first touches, last touches, and purchases

o   Of the four campaigns that drove first touches, the top three drew solid response

o   Together, weekly-newsletter and retargeting-ad campaigns were responsible for 63% of purchases

o   The retargeting-campaign and paid-search campaigns were responsible for 29% of purchases

Conclusions and recommendations for marketing spend:

o   Continue to conduct the three top campaigns drawing first touch traffic (interview-with-cool-tshirts-founder, getting-to-know-cool-tshirts, ten-crazy-cool-tshirts-facts)

o   Continue to conduct the two top campaigns that drive purchases, weekly-newsletter and retargeting-ad

o   Discontinue the cool-shirts-search campaign

o   If economically possible, continue to conduct the #3 and #4 campaigns that drive purchases, retargeting-campaign and paid-search

Appendix

Discrepancy found in page_visits file

I found a discrepancy between the instructions definition for the page_visits file. The utm_medium attribute did not appear in the actual file.

Database Schema	
page_visits	569
page_name	TEXT
timestamp	TEXT
user_id	INTEGER
utm_campaign	TEXT
utm_source	TEXT

page_visits	
A table describing each time a user visits the CoolTShirts website	
Column	**Description**
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 (i.e., google, newsletter, or facebook_ad)
utm_medium	Identifies what type of link was used (i.e., cost-per-click or email)
utm_campaign	Identifies the specific ad or email blast (i.e., june-21-newsletter or memorial-day-sale)