

Capstone: Attribution

Learn SQL from Scratch

Robert Warmbier

4 July 2018

SQL notes

- Please note: SQL coding for this is located on code.sql file
- There are 6 different 'segments' on the file
 - - 1. distinct campaigns; distinct sources, and distinct number of campaigns and sources
 - - 2. distinct page names from the page visits table
 - - 3. first_touches for each campaign
 - - 4. last touches for each campaign
 - - 5. number of visitors who made a purchase
 - - 6. number of last touches on the purchase page for each campaign

CoolTShirts' user data

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)

1. Getting 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`.

- They're related as far as news/website sources.

The `utm_source` and `utm_campaign` are defined as follows:

- `utm_source` - Identifies which site sent the traffic (i.e., google, newsletter, or facebook_ad)
- `utm_campaign` - Identifies the specific ad or email blast (i.e., june-21-newsletter or memorial-day-sale)

Number of campaigns
and sources

source	campaign
6	8

sources

- nytimes
- email
- buzzfeed
- facebook
- medium
- google

Number of campaigns
and sources

source	campaign
6	8

campaigns

- getting-to-know-cool-tshirts
- weekly-newsletter
- ten-crazy-cool-tshirts-facts
- retargetting-campaign
- retargetting-ad
- interview-with-cool-tshirts-founder
- paid-search
- cool-tshirts-search

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?

- A typical user journey:

User Journey
1 - landing_page
2 - shopping_cart
3 - checkout
4 - purchase

- First touches each campaign is responsible for:

source	campaign	first_touches
buzzfeed	ten-crazy-cool-tshirts-facts	576
google	cool-tshirts-search	169
medium	interview-with-cool-tshirts-founder	622
nytimes	getting-to-know-cool-tshirts	612

- Last touches is each campaign responsible for:

source	campaign	last_touch
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

- Number of visitors making a purchase:

visitors_purchase

361

- Last touches on the purchase page each campaign responsible for:

source	campaign	last_touches
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

3. Optimize the campaign budget

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

■ Re-invest campaigns

- Based upon the data presented on a previous slide, where the customers last-touched the item and purchased on the item, the company should re-invest in the 5 best campaigns as presented on the following table

number	source	campaign	last_touches
1	email	weekly-newsletter	115
2	facebook	retargetting-ad	113
3	email	retargetting-campaign	54
4	google	paid-search	52
5	buzzfeed	ten-crazy-cool-tshirts-facts	9