



SQL Capstone

Learn SQL from Scratch
Michael Mitchell
6/16/19

CONTENTS

- ❖ Get Familiar with the Company!
- ❖ What is the user Journey!
- ❖ Optimize the company Campaign budget

Get Familiar With CoolTShirts.com

“CoolTShirts sells shirts of all kinds, as long as they are T-shaped and cool. Recently, CTS started a few marketing campaigns to increase website visits and purchases. They would like to map their customers’ journey: from initial visit to purchase. They can use that information to optimize their marketing campaigns”.

- ❖ How many campaigns and sources does CoolTShirts use and how are they related?
- ❖ Explain the difference between **utm_campaign** and **utm_source**
- ❖ What pages are on their website?

THE SETUP

Let's find out what our queries tell us!

Get Familiar With CoolTShirts.com

For their marketing strategy, CTS decided to use **8** different campaigns and **6** different sources to improve website visits and purchases. By analyzing what traffic is attributed through which sources they can determine the best route the company can take to maximize sales.

We delve further into that analysis as we go on.

campaign	sources
getting-to-know-cool-tshirts	nytimes
weekly-newsletter	email
ten-crazy-cool-tshirts-facts	buzzfeed
retargeting-campaign	facebook
retargeting-ad	medium
interview-with-cool-tshirts-founder	google
paid-search	
cool-tshirts-search	

```
/* Query 1a - pulls data for all distinct campaigns being ran*/  
SELECT DISTINCT utm_campaign  
FROM page_visits;  
/* Counts number of distinct campaigns from page_visits table*/  
SELECT COUNT(DISTINCT utm_campaign) AS '#ofcampaigns'  
FROM page_visits;  
/* Query 1b - pulls data for all distinct sources campaigns use */  
SELECT DISTINCT utm_source  
FROM page_visits;  
/* Counts the number of sources pulled from the page_visits table*/  
SELECT COUNT(DISTINCT utm_source) AS '#ofsources'  
FROM page_visits;
```

Get Familiar With CoolTShirts.com

A **utm_source** identifies what site, media, or app sent the initial traffic to a web page. A **utm_campaign** represents the sales script or ad campaign that promotes the product.

Ultimately these are part of the parameters that shape the information used to determine how well traffic is sending people to specific pages on CTS's site. Among the next graphics we will see which campaigns utilize which sources.

Traffic needs to be driven to CTS's site to increase visibility and the likelihood of a purchase. There are 4 pages making up the site and they are the **landing page**, the **shopping cart**, the **checkout** and the **purchase page**.

Get Familiar With CoolTShirts.com

```
/* Query 1c - pulls data from page_visits table displaying which campaigns use which sources*/  
SELECT DISTINCT utm_source,  
                utm_campaign  
FROM page_visits;  
  
/* Query 2 - pulls DISTINCT page names from the page_visits table*/  
SELECT DISTINCT page_name AS 'Page Names'  
FROM page_visits;
```

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

pages

- 1 – landing_page
- 2 – shopping_cart
- 3 – checkout
- 4 - purchase

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?

The “**first_touch**” or the first time a visitor visits a site can be attributed to a specific source. Gathering data from first_touch metrics can give insight to a website’s visibility and how successful different marketing campaigns are with getting visitors. The “**last_touch**” is a touchpoint that conveys the last time a user visited a site and can be from exposure from the same source or a different source

What is the user JOURNEY?

Out of the 8 campaigns
only 4 are responsible
for all of the site's
first_touches and out
of the 4, 3 are
responsible for a
significant amount of
the visitors

```
/* Query 3 - creates a temporary table that aids in counting first touches per campaign */  
WITH first_touch AS (  
  SELECT user_id,  
         MIN(timestamp) as first_touch_at  
  FROM page_visits  
  GROUP BY user_id)  
SELECT ft.user_id,  
       ft.first_touch_at,  
       pv.utm_source,  
       pv.utm_campaign,  
       COUNT(ft.first_touch_at) AS '#of 1st 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 utm_campaign  
ORDER BY 5 DESC;
```

utm_source	utm-campaign	#of 1 st Touches
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

What is the user JOURNEY?

While not all campaigns contributed to the initial exposure to the site all campaigns and sources contributed to buyers/sales

```
/* Query 4 - creates a temporary table that aids in counting last touches per campaign */  
WITH last_touch AS (  
  SELECT user_id,  
         MAX(timestamp) as last_touch_at  
  FROM page_visits  
  GROUP BY user_id)  
SELECT lt.user_id,  
       lt.last_touch_at,  
       pv.utm_source,  
       pv.utm_campaign,  
       COUNT(lt.last_touch_at) AS '#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 utm_campaign  
ORDER BY 5 DESC;
```

utm_source	utm-campaign	#of Last Touches
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

What is the user JOURNEY?

361 of 1979 visitors made it through to purchase something from CTS! About **18%** of visitors make a purchase!!

```
/* Query 5 - counts all distinct users that made it through to purchase page*/  
SELECT COUNT(DISTINCT user_id) AS 'Buyers'  
FROM page_visits  
WHERE page_name = '4 - purchase';  
  
/* Query 5b Extra - counts all distinct users that represent all visitors to the site*/  
SELECT COUNT(DISTINCT user_id) AS 'Visitors'  
FROM page_visits;
```

Buyers	Visitors
361	1979

What is the user JOURNEY?

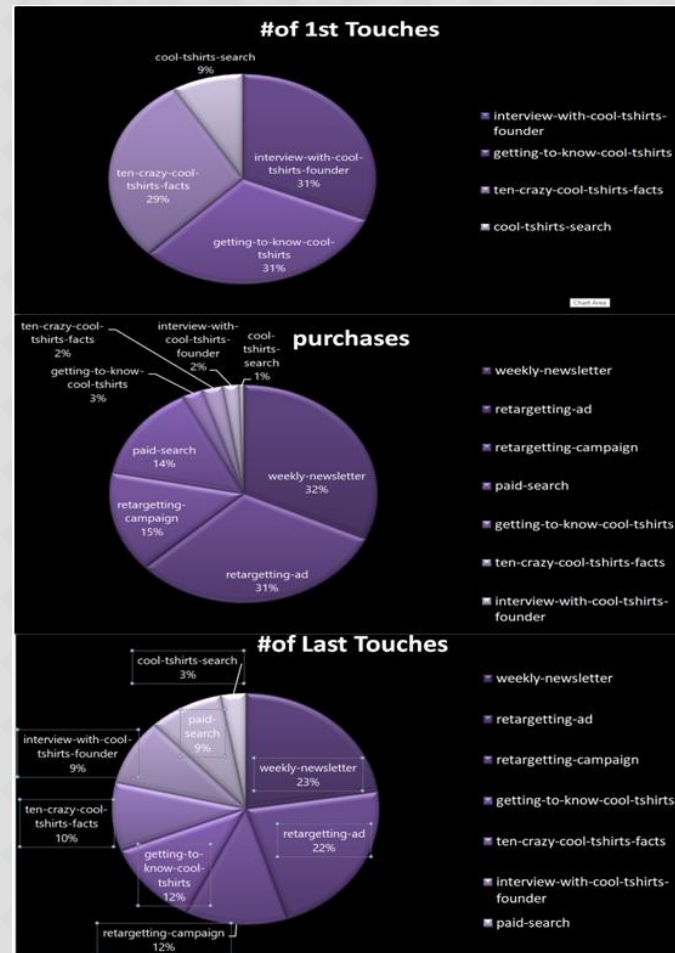
Taking a look at the last_touch data from the purchase page, we get a better idea of which campaigns and from which sources contributed to the companies sales.

```
/* Query 6 - amended last touch query that shows which campaigns led to a 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)
SELECT lt.user_id,
       lt.last_touch_at,
       pv.utm_source,
       pv.utm_campaign,
       COUNT(lt.last_touch_at) AS '#of Buyers 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 utm_campaign
ORDER BY 5 DESC;
```

utm_source	utm-campaign	#of Last Touches
email	weekly-newsletter	115
facebook	retargetting-ad	113
email	retargetting-campaign	54
google	paid-search	52
nytimes	getting-to-know-cool-tshirts	9
buzzfeed	ten-crazy-cool-tshirts-facts	9
medium	interview-with-cool-tshirts-founder	7
google	cool-tshirts-search	2

What is the user JOURNEY?

Typically it looks like most users aren't necessarily searching for CoolTShirts.com but gain interest from articles posted on news aggregate sites. The bulk of the first_touches can be attributed to buzzfeed and nytimes. After the initial interest is sparked, facebook and email blasts seem to make up the bulk of last_touches. Continuing with the trend most purchases are a result of the same type of methods being either facebook or email.



Optimize CTS's campaign budget!

To optimize the campaign budget, the company should definitely cut or drastically reduce whatever is allocated towards Google ad campaigns as the search campaigns yielded the least significant results. Maybe focusing on SEO for articles on aggregate sites with facebook and email campaigns would be the best way going forward to increase revenue.