



LEARN SQL FROM SCRATCH

ATTRIBUTION QUERIES

DEPENDER BEDI

2019-06-10

codecademy

TABLE OF CONTENTS

1. GET FAMILIAR WITH THE COMPANY.

2. WHAT IS THE USER JOURNEY?

3. OPTIMIZE THE CAMPAIGN BUDGET.

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.
WHAT PAGES ARE ON THEIR WEBSITE?

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 CAN RE-INVEST IN 5 CAMPAIGNS. WHICH SHOULD THEY PICK AND WHY?

I. GET FAMILIAR WITH COOLTSHIRTS

- HOW MANY CAMPAIGNS AND SOURCES DOES COOLTSHIRTS USE?
- WHICH SOURCE IS USED FOR EACH CAMPAIGN?

- `utm_campaign` identifies the specific ad or email a potential customer will view. CoolTShirts has 8 different campaigns.
- `utm_source` identifies the medium used to promote the campaign to a specific target. CoolTShirts has 6 different sources.
- In this dataset, there is a one to many relationship between sources and campaigns. One source can have multiple campaigns.
- A `SELECT COUNT(DISTINCT...)` query allows us to determine the number of unique campaigns and sources by counting the number of rows containing the specific string.

```
1 --1. How many campaigns and sources does CoolTShirts use?
2 SELECT COUNT(DISTINCT utm_campaign) AS 'Number of Campaigns',
3 COUNT (DISTINCT utm_source) AS 'Number of Sources'
4 FROM page_visits;
5
6 --Which source is used for each campaign?
7 SELECT DISTINCT utm_campaign AS 'Campaign',
8 utm_source AS 'Source'
9 FROM page_visits;
```

Query Results	
Number of Campaigns	Number of Sources
8	6
Campaign	Source
getting-to-know-cool-tshirts	nytimes
weekly-newsletter	email
ten-crazy-cool-tshirts-facts	buzzfeed
retargeting-campaign	email
retargeting-ad	facebook
interview-with-cool-tshirts-founder	medium
paid-search	google
cool-tshirts-search	google

I. GET FAMILIAR WITH COOL TSHIRTS

■ WHAT PAGES ARE ON THE COOLTSHIRTS WEBSITE?

- CoolTShirts has 4 different pages on its website.
- A SELECT DISTINCT query returns the unique values of the page_name column.

```
1  --2. What pages are on the CoolTshirts website?
2  SELECT DISTINCT page_name AS 'Page'
3  FROM page_visits;
```

Query Results

Page

1 – landing_page

2 – shopping_cart

3 – checkout

4 - purchase

2. WHAT IS THE USER JOURNEY?

■ HOW MANY FIRST TOUCHES IS EACH CAMPAIGN RESPONSIBLE FOR?

- There were 1979 unique user IDs in this data set.
- Of the 8 available campaigns, this query shows all first touch traffic was generated from 4 campaigns.
- First touch is the first source for each unique user_id. This is found by linking the user_id and MIN timestamp for each user in a temporary table.
- Source and Campaign are linked to this and a COUNT query results in number of rows with the specific campaign

```
1  --3. How many first touches is each campaign responsible for?
2  WITH first_touch AS
3  (SELECT user_id,
4  MIN(timestamp) as first_touch_at
5  FROM page_visits
6  GROUP BY user_id),
7  ft_attr AS
8  (SELECT ft.user_id,
9  ft.first_touch_at,
10 pv.utm_source,
11 pv.utm_campaign
12 FROM first_touch ft
13 JOIN page_visits pv
14 ON ft.user_id = pv.user_id
15 AND ft.first_touch_at = pv.timestamp)
16 SELECT ft_attr.utm_source AS 'Source',
17 ft_attr.utm_campaign AS 'Campaign',
18 COUNT(*) AS 'Number'
19 FROM ft_attr
20 GROUP BY 1, 2
21 ORDER BY 3 DESC;
```

Query Results

Source	Campaign	Number
medium	interview-with-cool-tshirts-founder	622
nytimes	getting-to-know-cool-tshirts	612
buzzfeed	ten-crazy-cool-tshirt-facts	576
google	cool-tshirts-search	169

2. WHAT IS THE USER JOURNEY?

■ HOW MANY LAST TOUCHES IS EACH CAMPAIGN RESPONSIBLE FOR?

- Last touch provides the last source for each user_id.
- This is similar to the first touch query, switching the MIN timestamp for MAX.

```
1  --4. How many last touches is each campaign responsible for?
2  WITH last_touch AS
3  (SELECT user_id,
4  MAX(timestamp) AS last_touch_at
5  FROM page_visits
6  GROUP BY user_id),
7  lt_attr AS
8  (SELECT lt.user_id,
9  lt.last_touch_at,
10 pv.utm_source,
11 pv.utm_campaign
12 FROM last_touch lt
13 JOIN page_visits pv
14 ON lt.user_id = pv.user_id
15 AND lt.last_touch_at = pv.timestamp)
16 SELECT lt_attr.utm_source AS 'Source',
17 lt_attr.utm_campaign AS 'Campaign',
18 COUNT(*) AS 'Number'
19 FROM lt_attr
20 GROUP BY 1, 2
21 ORDER BY 3 DESC;
```

Query Results		
Source	Campaign	Number
email	weekly-newsletter	447
facebook	retargeting-ad	443
email	retargeting-campaign	245
nytimes	getting-to-know-cool-tshirts	232
buzzfeed	ten-crazy-cool-tshirt-facts	190
medium	interview-with-cool-tshirts-founder	184
google	paid-search	178
google	cool-tshirts-search	60

2. WHAT IS THE USER JOURNEY?

- HOW MANY VISITORS MAKE A PURCHASE?
- HOW MANY LAST TOUCHES ON THE PURCHASE PAGE IS EACH CAMPAIGN RESPONSIBLE FOR?

- Of the 1979 unique user IDs, only 361 make a purchase, representing approximately an 18% conversion rate.
- The WHERE statement filters the purchase page, providing the data on which campaign results in the most purchases.

```
1 --5. How many visitors make a purchase?
2 SELECT COUNT(DISTINCT user_id) AS 'Visitors that purchase'
3 FROM page_visits
4 WHERE page_name = '4 - purchase';
```

```
1 --6. How many last touches on the purchase page is each campaign responsible for?
2 WITH last_touch AS
3 (SELECT user_id,
4  MAX(timestamp) AS last_touch_at
5  FROM page_visits
6  WHERE page_name = '4 - purchase'
7  GROUP BY user_id),
8 lt_attr AS
9 (SELECT lt.user_id,
10  lt.last_touch_at,
11  pv.utm_source,
12  pv.utm_campaign
13  FROM last_touch lt
14  JOIN page_visits pv
15  ON lt.user_id = pv.user_id
16  AND lt.last_touch_at = pv.timestamp)
17 SELECT lt_attr.utm_source AS 'Source',
18  lt_attr.utm_campaign AS 'Campaign',
19  COUNT(*) AS 'Number'
20 FROM lt_attr
21 GROUP BY 1, 2
22 ORDER BY 3 DESC;
```

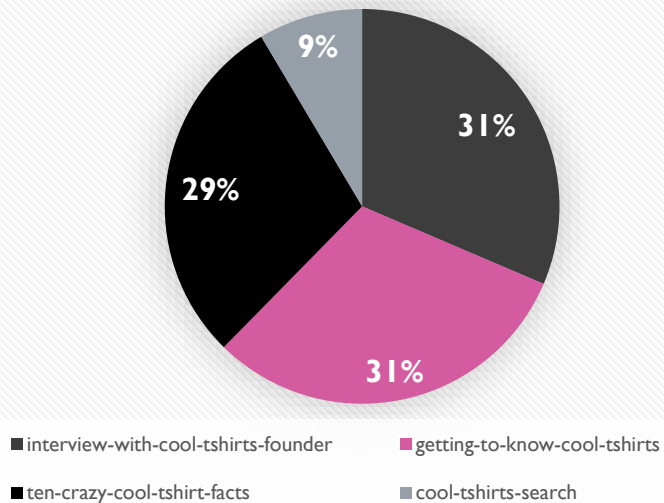
Query Results		
Visitors that purchase		
361		

Query Results		
Source	Campaign	Number
email	weekly-newsletter	115
facebook	retargeting-ad	113
email	retargeting-campaign	54
google	paid-search	52
buzzfeed	ten-crazy-cool-tshirt-facts	9
nytimes	getting-to-know-cool-tshirts	9
medium	interview-with-cool-tshirts-founder	7
google	cool-tshirts-search	2

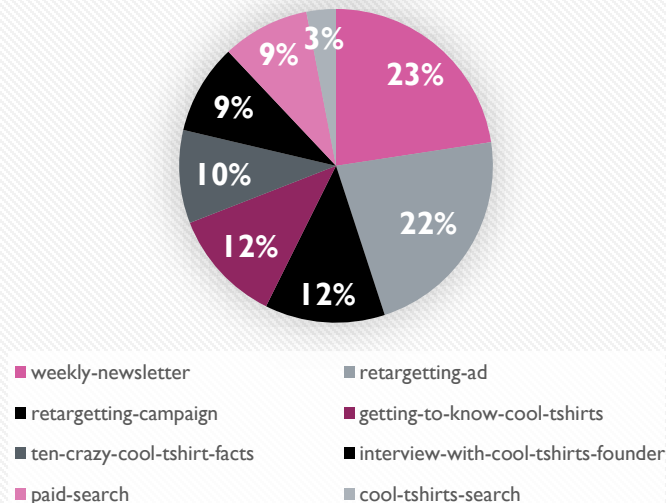
2. WHAT IS THE USER JOURNEY?

■ WHAT IS THE TYPICAL USER JOURNEY?

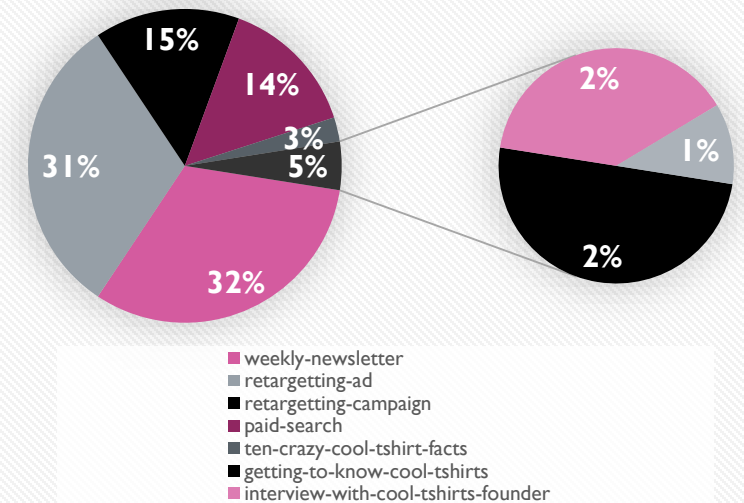
First Touches by Campaign



Last Touches by Campaign



Purchases by Campaign



- 91% of users are introduced to CoolTShirts via campaigns on third parties Medium, NYTimes and BuzzFeed.

- Data suggests while the best first touch campaigns are great at generating traffic to the site, they have poor conversion rates which is why a follow up campaign is required.

- The follow up campaigns accounted for 78% of the purchases.

3. OPTIMIZING THE CAMPAIGN BUDGET

- COOLTSHIRTS CAN RE-INVEST IN 5 CAMPAIGNS.WHICH SHOULD THEY PICK AND WHY?

- Making minor amendments to the 'First Touch', 'Last Touch' and 'Last Touch on purchase page' queries provides the distribution by campaign.
- Depending on the price per click, data suggests re-investment in the 3 campaigns providing the greatest proportion of first touches. These build the most brand awareness.
- The weekly-newsletter and retargeting-ad account for 63% of the purchases.As such, both of these campaigns should be re-invested in.

```
15 SELECT ft_attr.utm_campaign AS 'Campaign',
16 100.0 * COUNT(*) / 1979 AS '% First Touch'
17 FROM ft_attr
18 GROUP BY 1
19 ORDER BY 2 DESC;
```

```
15 SELECT lt_attr.utm_campaign AS 'Campaign',
16 100.0 * COUNT(*) / 1979 AS '% Last Touch'
17 FROM lt_attr
18 GROUP BY 1
19 ORDER BY 2 DESC;
```

```
16 SELECT lt_attr.utm_campaign AS 'Campaign',
17 100.0 * COUNT(*) / 361 AS '% Purchases'
18 FROM lt_attr
19 GROUP BY 1
20 ORDER BY 2 DESC;
```

Campaign	% First Touch	% Last Touch	% Purchase
weekly-newsletter	0%	23%	32%
retargeting-ad	0%	22%	31%
retargeting-campaign	0%	12%	15%
paid-search	0%	9%	14%
ten-crazy-cool-tshirt-facts	29%	10%	2%
getting-to-know-cool-tshirts	31%	12%	2%
interview-with-cool-tshirts-founder	31%	9%	2%
cool-tshirts-search	9%	3%	1%