



CoolTShirts Capstone Project

Learn SQL from Scratch

Mike Cox

June 28, 2018

Table of Contents

1. Getting Familiar with CoolTShirts
2. The User Journey
3. Optimizing the Campaign Budget
4. Additional Thoughts – Source of Traffic

1. Getting Familiar with CoolTShirts

1.1 How many campaigns and Sources does CoolTShirts use? Which Source is used for each Campaign?

- Campaign is a specific plan to market the CoolTShirts product.
- Source is the channel through which the Campaign is marketed, such as print media, online media, social media, email, etc.
- There can be a one to many relationship between Campaign and Source. For example, “retargeting” was Sourced using email and Facebook.
- Sources can also be used for multiple Campaigns, such as “google” was a Source for “Paid” and “CoolTShirts” searches

Campaign Count	Source Count
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

```
1 --Getting distinct Campaign Count
2 SELECT COUNT(DISTINCT utm_campaign) AS 'Campaign Count'
3 FROM page_visits;
4
5 --Getting distinct Sources
6 SELECT COUNT(DISTINCT utm_source) AS 'Source Count'
7 FROM page_visits;
8
9 --Determining relationship between Campaign and Source
10 SELECT DISTINCT utm_campaign AS Campaigns,
11                 utm_source AS Sources
12 FROM page_visits;
```

1.2 What pages are on the CoolTShirts website?

- Websites can contain many pages.
- The CoolTShirts website contains four unique pages.
- Each page serves a different purpose and can be accessed by the customer from different Campaigns and Sources, which will prove important later in the project analysis.

```
13
14 --Determining pages that are contained on the CoolTShirts website
15 SELECT DISTINCT page_name AS Page
16 FROM page_visits;
```

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

2. The User Journey

2.1 How many first touches is each Campaign responsible for?

- Joining user id to the MIN timestamp allows the ability to calculate the count of unique first touches.
- Associating this with the Source allows the ability to determine which Sources led users to the CoolTShirts website for their first touch.
- Finally, by joining that with the Campaign, one can now determine by Campaign and Source the combination that is the biggest first touch draw to the CoolTShirts website.
- Of the eight Campaigns, four drew first touches from customers for a total of 1,979 unique first touches.

Source	Campaign	Count
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

```
1  --Determining the unique first touches by id by creating a
2  --temp table
3  WITH first_touch AS(
4      SELECT user_id,
5             MIN(timestamp) as first_touch_at
6      FROM page_visits
7      GROUP BY user_id),
8
9  --Determining the unique first touches
10 --by Source with Campaign joined by id and timestamp by
11 --creating a second temp table
12 fl_attr AS(
13     SELECT ft.user_id,
14            ft.first_touch_at,
15            pv.utm_source,
16            pv.utm_campaign
17     FROM first_touch ft
18     JOIN page_visits pv
19         ON ft.user_id = pv.user_id
20         AND ft.first_touch_at = pv.timestamp)
21
22 --Putting it together, determining the total number of unique
23 --first touches by id with Source and Campaign
24 SELECT fl_attr.utm_source AS Source,
25        fl_attr.utm_campaign AS Campaign,
26        COUNT (*) AS Count
27 FROM fl_attr
28 GROUP BY 1, 2
29 ORDER BY 3 DESC;
```

2.2 How many last touches is each Campaign responsible for?

- Joining user id to the MAX timestamp allows the ability to count the number of last touches by user.
- Joining this with the Source allows the ability to determine the Source that led to the user's last touch.
- Finally, by joining Source with the Campaign, one can now determine the Campaign and Source combination that led to the last touch to the CoolTShirts website.

Source	Campaign	Count
email	weekly-newsletter	447
facebook	retargeting-ad	443
email	retargeting-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

The top two Source/Campaign combinations were email/weekly-newsletter and facebook/retargeting-ad, which were responsible for 45% of the last touches.

```
1  --Determining the unique last touches by id by creating a
2  --temp table
3  WITH last_touch AS(
4      SELECT user_id,
5             MAX(timestamp) as last_touch_at
6      FROM page_visits
7      GROUP BY user_id),
8
9  --Determining the unique last touches
10 --by Source with Campaign joined by id and timestamp by
11 --creating a second temp table
12 lt_attr AS(
13     SELECT lt.user_id,
14            lt.last_touch_at,
15            pv.utm_source,
16            pv.utm_campaign
17     FROM last_touch lt
18     JOIN page_visits pv
19         ON lt.user_id = pv.user_id
20         AND lt.last_touch_at = pv.timestamp)
21
22 --Putting it together, determining the total number of
23 --last touches by id with Source and Campaign
24 SELECT lt_attr.utm_source AS Source,
25        lt_attr.utm_campaign AS Campaign,
26        COUNT (*) AS Count
27 FROM lt_attr
28 GROUP BY 1, 2
29 ORDER BY 3 DESC;
```


2.3 How many visitors make a purchase?

- By counting the distinct user ids with the last touch on the '4 – purchase page', one can determine the number of unique visitors who completed a purchase on the CoolTShirts website.
- The conversion rate to purchase for visitors is 18%. This is determined by the dividing the total number of unique visitors who purchased (361) and dividing by unique first touches (1,979 determined in step 2.1).
- $361 / 1,979 = 18\%$

Customers Who Purchase

361

```
1 --Determining the total number of visitors who made a purchase
2 SELECT COUNT(DISTINCT user_id)
3 FROM page_visits
4 WHERE page_name = '4 - purchase';
5
```

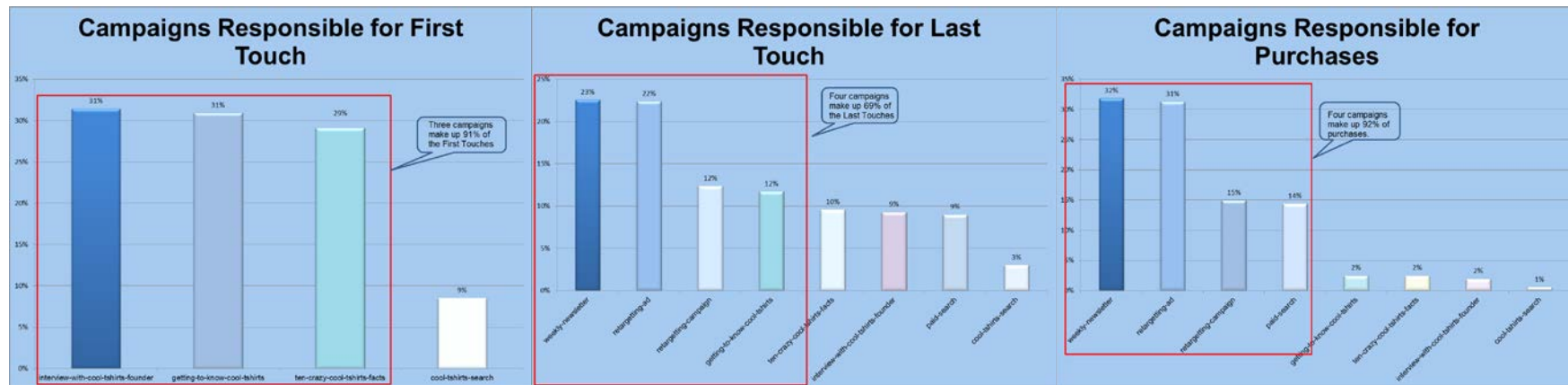
2.4 How many last touches *on the purchase page* is each Campaign responsible for?

- Joining user id to the MAX timestamp on the '4 – Purchase' page allows the ability to calculate the last touches by Campaign.
- The top two Source/Campaign combinations were email/weekly-newsletter and facebook/retargeting-ad, which were responsible for 228 purchases, or 63% of total purchases.

Source	Campaign	Count
email	weekly-newsletter	115
facebook	retargeting-ad	113
email	retargeting-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

```
1  --Last touches on the purchase page by campaign
2
3  WITH last_touch AS (
4    SELECT user_id,
5           MAX(timestamp) as last_touch_at
6    FROM page_visits
7    WHERE page_name = '4 - purchase'
8    GROUP BY user_id),
9
10  lt_attr AS (
11    SELECT lt.user_id,
12           lt.last_touch_at,
13           pv.utm_source,
14           pv.utm_campaign
15    FROM last_touch lt
16    JOIN page_visits pv
17      ON lt.user_id = pv.user_id
18      AND lt.last_touch_at = pv.timestamp)
19
20  SELECT lt_attr.utm_source AS Source,
21         lt_attr.utm_campaign AS Campaign,
22         COUNT (*) AS COUNT
23  FROM lt_attr
24  GROUP BY 1, 2
25  ORDER BY 3 DESC;
26
```

2.5 The *Typical* User Journey



- Three Campaigns drive 91% the First Touches.
- These are all story based.
- Four Campaigns drive 69% of the Last Touches.*
- Two are retargeting ads and one each story based and newsletter.
- Four Campaigns lead to 92% of the Purchases.
- Two are retargeting ads and one each news letter and paid-search.

**This not to be interpreted as a purchase*

Visitor reads about the CoolITShirts and checks out the website.



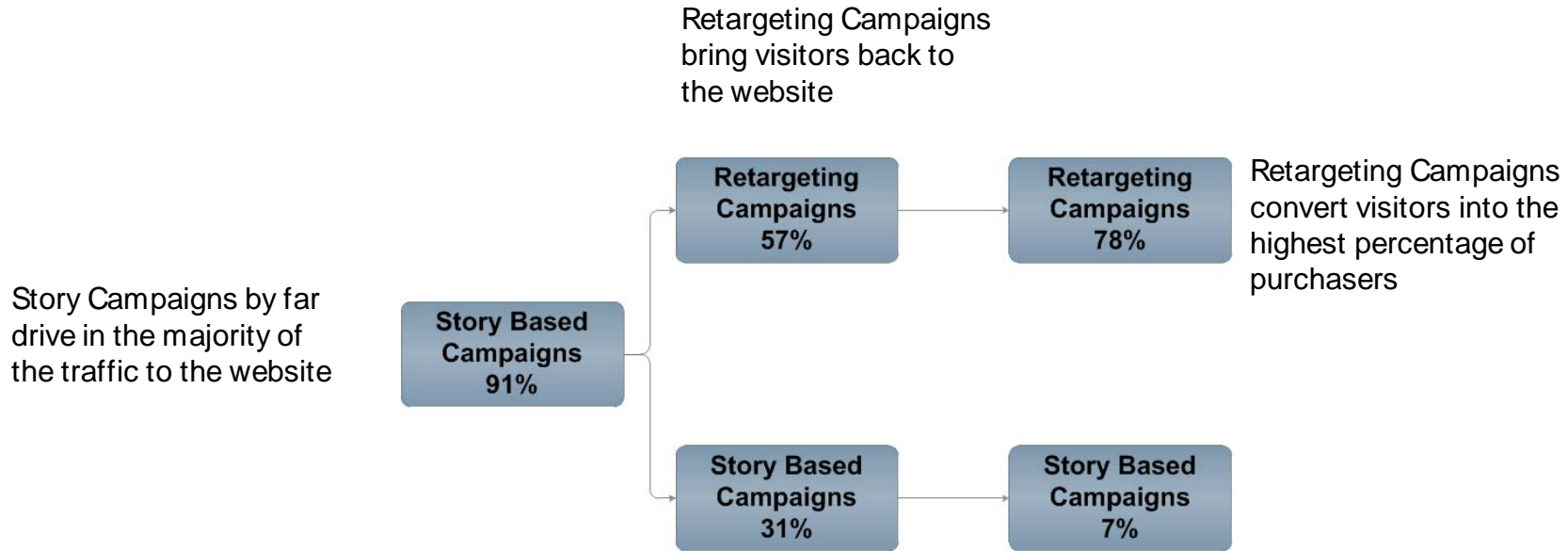
Visitor continues to last touch from Story Based Campaigns.



Visitor returns and makes purchase after receiving Retargeting Campaign.



2.5 The *Typical* User Journey *cont'd*



Visitor reads about the CoolTShirts and checks the website out.



Visitor receives a reminder from a Retargeting Campaign and returns to site.



Visitor returns and makes purchase after receiving Retargeting Campaign.



3. Optimizing the Campaign Budget

3. Optimizing the Campaign Budget

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

There are two components to determining which Campaigns that CoolTShirts should consider re-investing in:

1. What Campaigns are bringing the highest number of visitors to the website?
2. What Campaigns are leading visitors to make purchases on the website?

Campaign	First Touches	%First Touches	Last Touches	%Last Touches	Purchases	%Purchases
interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%
getting-to-know-cool-tshirts	612	31%	232	12%	9	2%
ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	2%
cool-tshirts-search	169	9%	60	3%	2	1%
weekly-newsletter	0	0%	447	23%	115	32%
retargeting-ad	0	0%	443	22%	113	31%
retargeting-campaign	0	0%	245	12%	54	15%
paid-search	0	0%	178	9%	52	14%

Recommendation

1. Story Based Campaigns are by far the best way to bring visitors to the website. However, they have the lowest rate of purchases.
2. Retargeting Campaigns lead to the highest number of purchases on the website. Due to the nature of the Campaign, they do not bring in any first touches.

The best five Campaigns in which to reinvest are a combination of the two:

The top three Story Based Campaigns (*as they bring in 91% of first touches*) and top two Retargeting Campaigns (*as they lead to 63% of overall purchases*).

The top five Campaigns are highlighted in the table.

4. Additional Thoughts – Source of Traffic

Another thing to consider is the best combination of Source and Campaign that visitors are using to access the CoolTShirts website.

First Touch by Source

Source	First Touch	%First Touch
nytimes	747	21%
email	696	20%
buzzfeed	648	19%
medium	625	18%
facebook	445	13%
google	339	10%

```
1 --Another thing to consider, Source for
2 --First Touch
3
4 SELECT utm_source AS Source,
5        COUNT (DISTINCT user_id) AS 'First Touch'
6 FROM page_visits
7 GROUP BY 1
8 ORDER BY 2 DESC;
```

The Source differs between first touches, or what drives visitors to the CoolTShirts website and the Source that visitors use when making a purchase.

Facebook seems to be one of the lowest Sources at driving a first touches. However, it is the second highest utilized Source when visitors make purchases via Retargeting Campaigns.

For driving purchases, Retargeting Campaigns (the highest driver for visitors making a purchase) combined with either email or facebook as Source will optimize budget use as these combinations account for 78% of all purchases.

Purchase by Source

Source	Purchase	%Purchase
email	169	47%
facebook	113	31%
google	54	15%
buzzfeed	9	2%
nytimes	9	2%
medium	7	2%

```
1 --Also, Source for Purchases
2
3 SELECT utm_source AS Source,
4        COUNT (DISTINCT user_id) AS 'Purchase'
5 FROM page_visits
6 WHERE page_name = '4 - purchase'
7 GROUP BY 1
8 ORDER BY 2 DESC;
```