

FIRST & LAST TOUCH ATTRIBUTIONS

LEARN SQL FROM SCRATCH

LING LING LU

7.16.2018

TABLE OF CONTENTS

- Get familiar with CoolTShirts
- What is the user journey?
- Optimize the campaign budget

1. Get 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`.
- What pages are on their website?

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?

3. Optimize the campaign budget

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

OVERVIEW OF PAGE_VISITS TABLE

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 COOLTSHIRTS

❖ How many campaigns and sources does CoolTShirts use and how are they related? Which source is used for each campaign?

- 8 campaigns
- 6 sources
- A single source can have multiple campaigns, like email and google do shown in the query results on the right.
- Likewise, a single campaign could be run across multiple sources, but in this dataset, different campaigns appear to have been run across different sources – none overlap in the current data.

❖ What is the difference between `utm_campaign` and `utm_source`?

- `utm_campaign` identifies the specific entity (or campaign) that traffic originated from (e.g., a specific ad, email blast, newsletter, etc.).
- `utm_source` identifies which website the traffic originated from (e.g., google, buzzfeed, facebook_ad, etc.).

```
-- COUNT DISTINCT number of campaigns.
SELECT COUNT(DISTINCT utm_campaign) AS
'Number of Campaigns'
FROM page_visits;

-- COUNT DISTINCT number of sources.
SELECT COUNT(DISTINCT utm_source) AS
'Number of Sources'
FROM page_visits;

-- List sources used per campaign to
show how they are related.
SELECT DISTINCT utm_campaign AS
'Campaign',
       utm_source AS 'Source'
FROM page_visits;
```

Query Results	
Number of Campaigns	
8	
Number of Sources	
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 FAMILIAR WITH COOLTSHIRTS (CONT.)

- ❖ **What pages are on the CoolTShirts's website?**
(Find the distinct values of the **page_name** column.)
- landing_page
 - shopping_cart
 - checkout
 - purchase
 - There are only 4 pages on CoolTShirt's website, which seems a bit low, but using a simple DISTINCT query yielded only 4 page results.

```
1  -- DISTINCT values of page_name
2  SELECT DISTINCT page_name AS 'Page
   Name'
3  FROM page_visits;
```

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

2.1 WHAT IS THE USER JOURNEY?

❖ How many first touches is each campaign responsible for?

- In order to identify first touches, `MIN(timestamp)` is used to determine which campaign from which source users originated from.
- Based on the query results shown on the right, 4 out of the 8 total campaigns generated the majority of first touch traffic.

Query Results		
Source	Campaign	First Touch 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 WITH first_touch AS (  
2     SELECT user_id,  
3           MIN(timestamp) as first_touch_at  
4     FROM page_visits  
5     GROUP BY user_id),  
6 ft_attr AS (  
7     SELECT ft.user_id,  
8           ft.first_touch_at,  
9           pv.utm_source,  
10          pv.utm_campaign  
11    FROM first_touch ft  
12    JOIN page_visits pv  
13      ON ft.user_id = pv.user_id  
14      AND ft.first_touch_at = pv.timestamp  
15 )  
16 SELECT ft_attr.utm_source AS 'Source',  
17        ft_attr.utm_campaign AS 'Campaign',  
18        COUNT(*) AS 'First Touch Count'  
19 FROM ft_attr  
20 GROUP BY 1, 2  
21 ORDER BY 3 DESC;
```

2.1 WHAT IS THE USER JOURNEY? (CONT.)

❖ How many last touches is each campaign responsible for?

- In order to identify which last touches each campaign is responsible for, `MIN(timestamp)` is swapped out for `MAX(timestamp)`.
- Interestingly enough, the results show lots of traffic through different campaigns from different sources, but campaigns that had high traffic with first touch have lower numbers with last touch. Prospective customers (i.e., leads) often became actual customers with campaigns that acted more as “reminders”, like the email ‘weekly-newsletter’ and facebook’s ‘retargetting-ad’.

Query Results		
Source	Campaign	Last Touch Count
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

```
1 WITH last_touch AS (  
2     SELECT user_id,  
3           MAX(timestamp) as last_touch_at  
4     FROM page_visits  
5     GROUP BY user_id),  
6 lt_attr AS (  
7     SELECT lt.user_id,  
8           lt.last_touch_at,  
9           pv.utm_source,  
10          pv.utm_campaign,  
11          pv.page_name  
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 )  
17 SELECT lt_attr.utm_source AS 'Source',  
18        lt_attr.utm_campaign AS 'Campaign',  
19        COUNT(*) AS 'Last Touch Count'  
20 FROM lt_attr  
21 GROUP BY 1, 2  
22 ORDER BY 3 DESC;
```


2.1 WHAT IS THE USER JOURNEY? (CONT.)

❖ How many visitors make a purchase?

- Out of 1979 visitors, 361 of those visitors became actual customers by making a purchase.
- $361/1979 = 18.2\%$ converted from visitor to customer.

```
1 SELECT COUNT(DISTINCT user_id) AS  
   'Customers with Purchases', page_name  
   AS 'Page Name'  
2 FROM page_visits  
3 WHERE page_name = '4 - purchase';
```

Query Results	
Customers with Purchases	Page Name
361	4 - purchase

- ❖ Below is an extra query to determine the total number of customers who did not make a purchase.

```
1 SELECT COUNT(DISTINCT user_id) AS  
   'Customers with Purchases', page_name  
   AS 'Page Name'  
2 FROM page_visits  
3 WHERE page_name IS NOT '4 - purchase';
```

Query Results	
Customers with Purchases	Page Name
1979	3 - checkout

2.1 WHAT IS THE USER JOURNEY? (CONT.)

❖ How many last touches on the purchase page is each campaign responsible for?

- By adding a **WHERE** clause to the previous last touch query, the results provide valuable information, showing which campaigns correlated with customers making a purchase.
- Again, it appears that the campaigns that act more as “reminders” retargeting visitors is more effective in turning a visitor into a customer (i.e., the first 3 campaigns listed in the query results on the right).

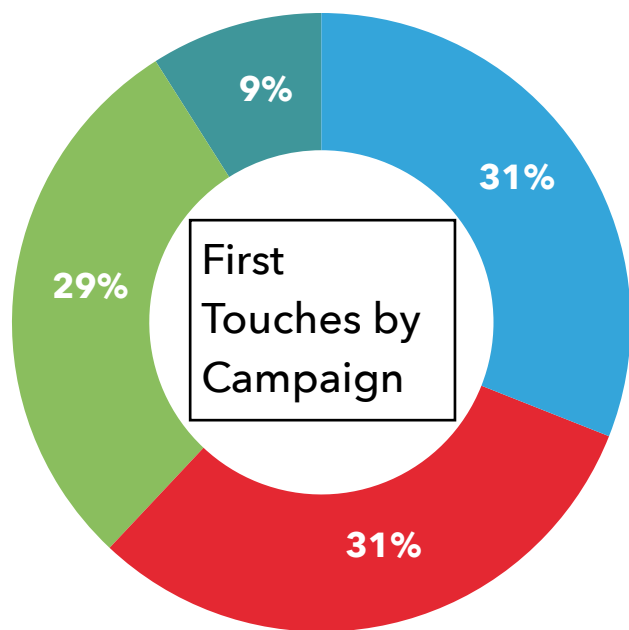
Query Results		
Source	Campaign	Last Touch Count
email	weekly-newsletter	115
facebook	retargeting-ad	113
email	retargeting-campaign	54
google	paid-search	52
buzzfeed	ten-crazy-cool-tshirts-facts	9
nytimes	getting-to-know-cool-tshirts	9
medium	interview-with-cool-tshirts-founder	7
google	cool-tshirts-search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

```
1 WITH last_touch AS (  
2     SELECT user_id,  
3           MAX(timestamp) as last_touch_at  
4     FROM page_visits  
5     WHERE page_name = '4 - purchase'  
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          pv.page_name  
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 )  
18 SELECT lt_attr.utm_source AS 'Source',  
19        lt_attr.utm_campaign AS 'Campaign',  
20        COUNT(*) AS 'Last Touch Count'  
21 FROM lt_attr  
22 GROUP BY 1, 2  
23 ORDER BY 3 DESC;
```

2.2 TYPICAL USER JOURNEY – OVERVIEW

First Touches by Campaign

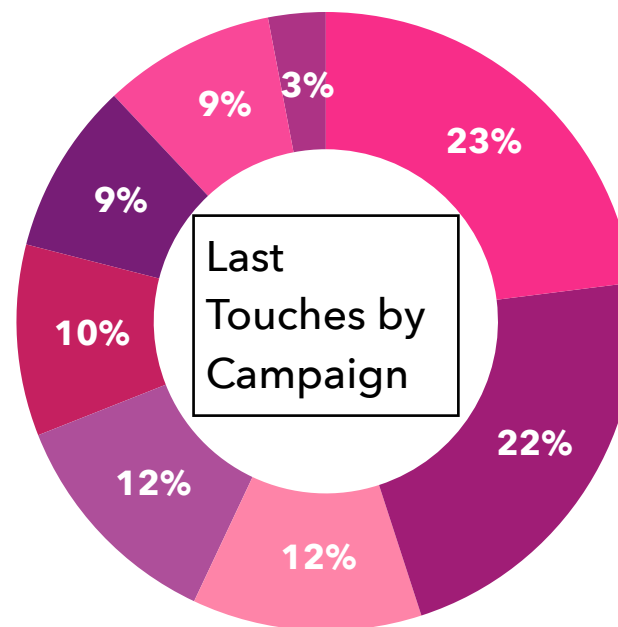
- ❖ 91% of the "About" types of campaigns peaks users' interests to venture to CoolTShirts's website to learn more.



- interview-with-cool-tshirts-founder
- getting-to-know-cool-tshirts
- ten-crazy-cool-tshirts-facts
- cool-tshirts-search
- weekly-newsletter

Last Touches by Campaign

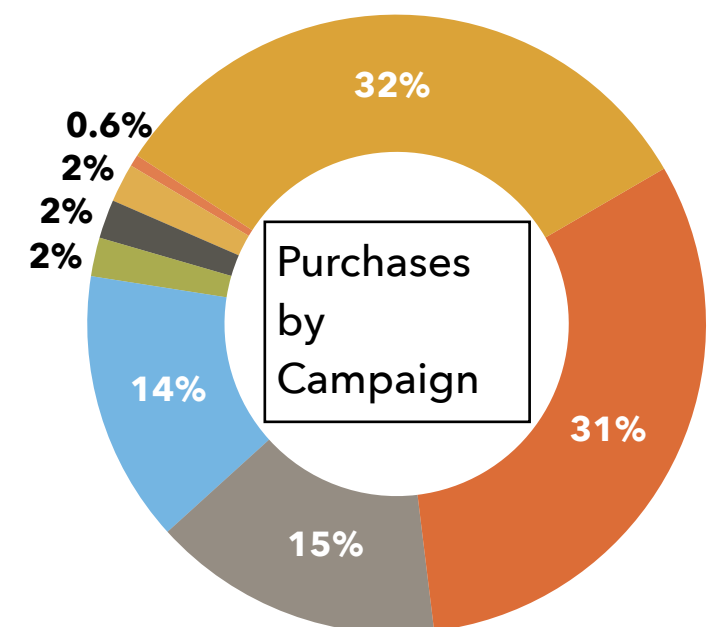
- ❖ 34% of users attained from "About" types of campaigns lose interest after their initial visit.
- ❖ 45% stay engaged through "Reminder" types of campaigns.



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

Purchases by Campaign

- ❖ 78% of "Reminder" types of campaigns correlates with purchases made by users.



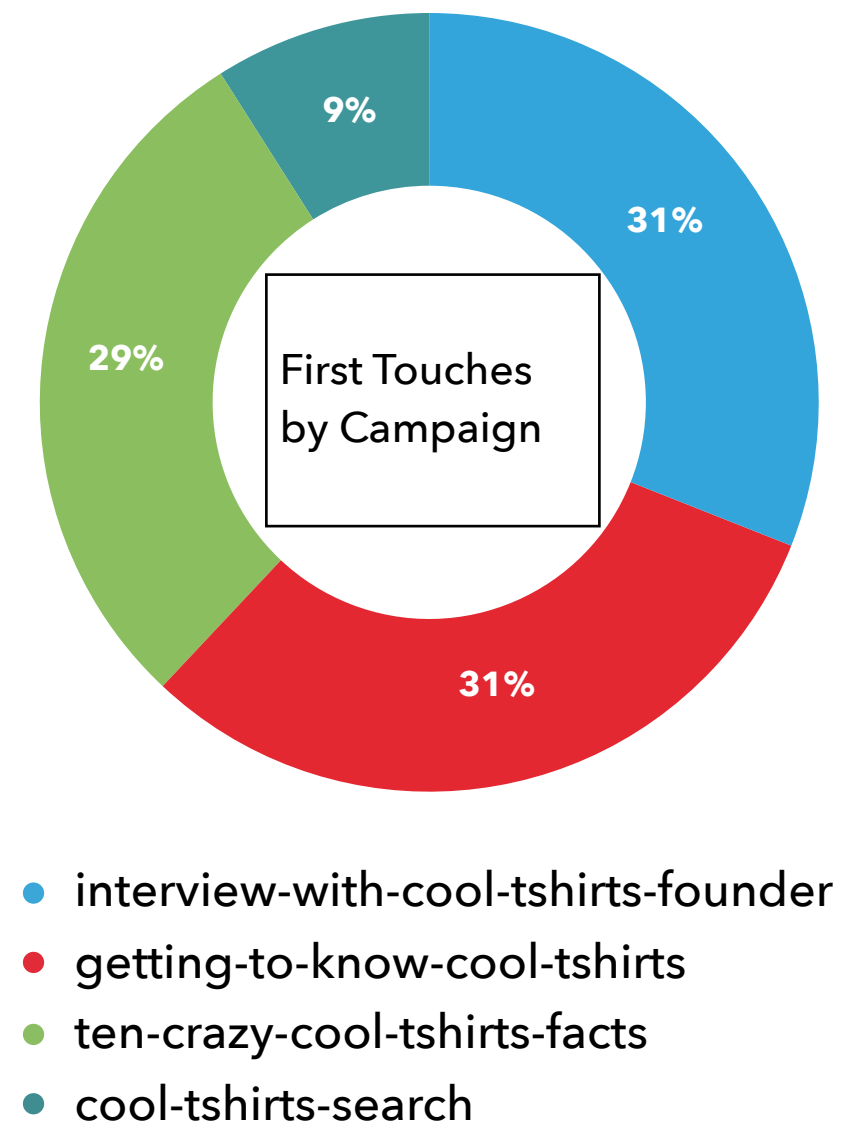
- weekly-newsletter
- retargeting-ad
- retargeting-campaign
- paid-search
- ten-crazy-cool-tshirts-facts
- getting-to-know-cool-tshirts
- interview-with-cool-tshirts-founder
- cool-tshirts-search

2.2 WHAT IS THE TYPICAL USER JOURNEY?

❖ First Touches by Campaign

- The top 3 campaigns are more **"About"** types of campaigns, where users can learn a bit more about the company, the company's philosophy or values, and about the product.
- Based on the current dataset, **91%** of the **"About"** types of campaigns peaks users' interests to venture to CoolTShirts's website to learn more.
- Only **9%** proactively search either for CoolTShirts's website or via a keyword search of 'cool tshirts' (best guess based on the data available).

Query Results		
Source	Campaign	First Touch 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

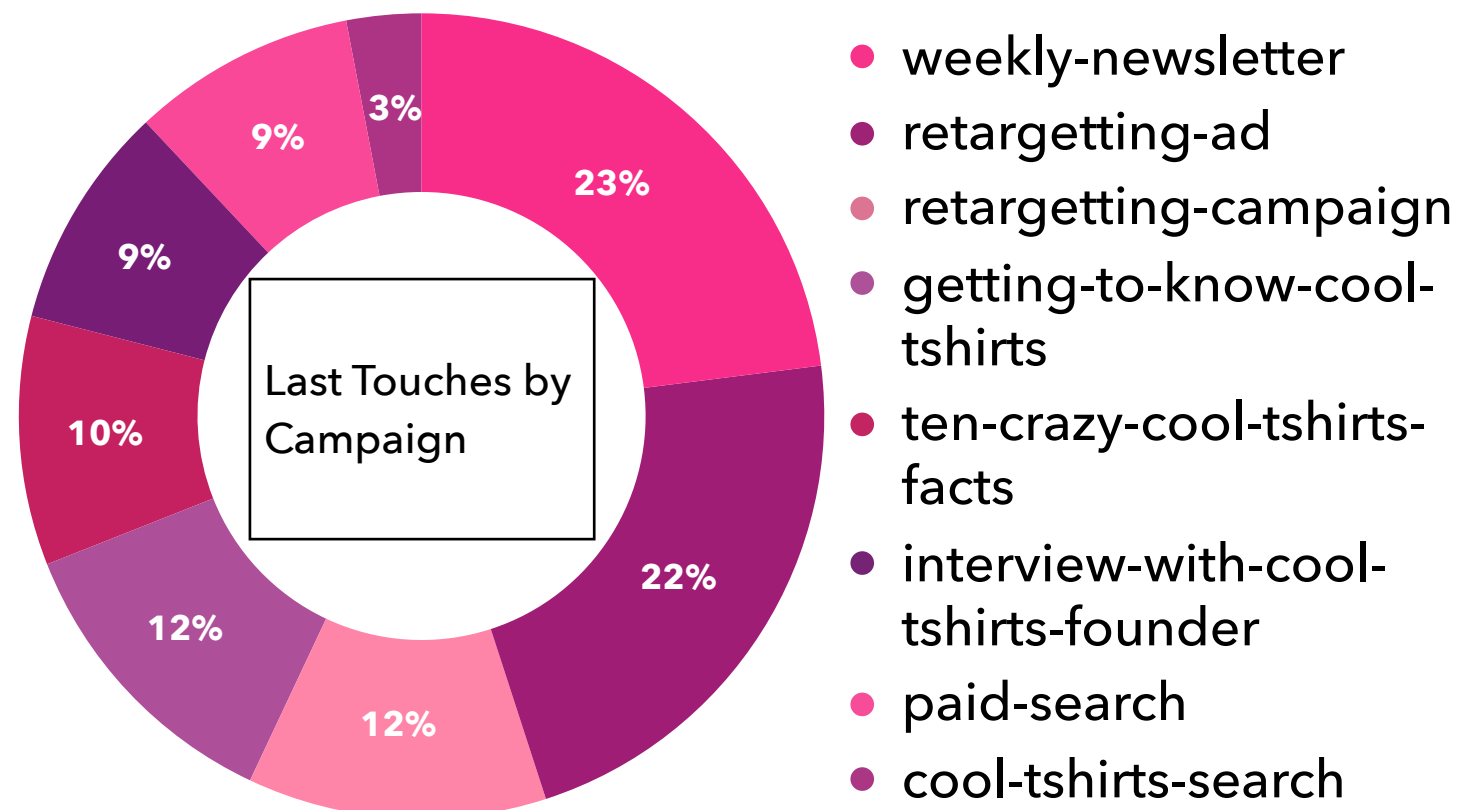


2.2 WHAT IS THE TYPICAL USER JOURNEY? (CONT.)

❖ Last Touches by Campaign

- While the “About” types of campaigns generated a lot of *initial* traffic to the site, it did not interest enough users to continue exploring the website as **34%** of the first touch users **do not return after their first visit**.
- However, only **7%** of those users acquired from the “About” types of campaigns actually make a purchase.
- **45%** of users **stay engaged** from “Reminder” types of campaigns – the top 2 being ‘weekly-newsletter’ and ‘retargetting-ad’.

Query Results		
Source	Campaign	Last Touch Count
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

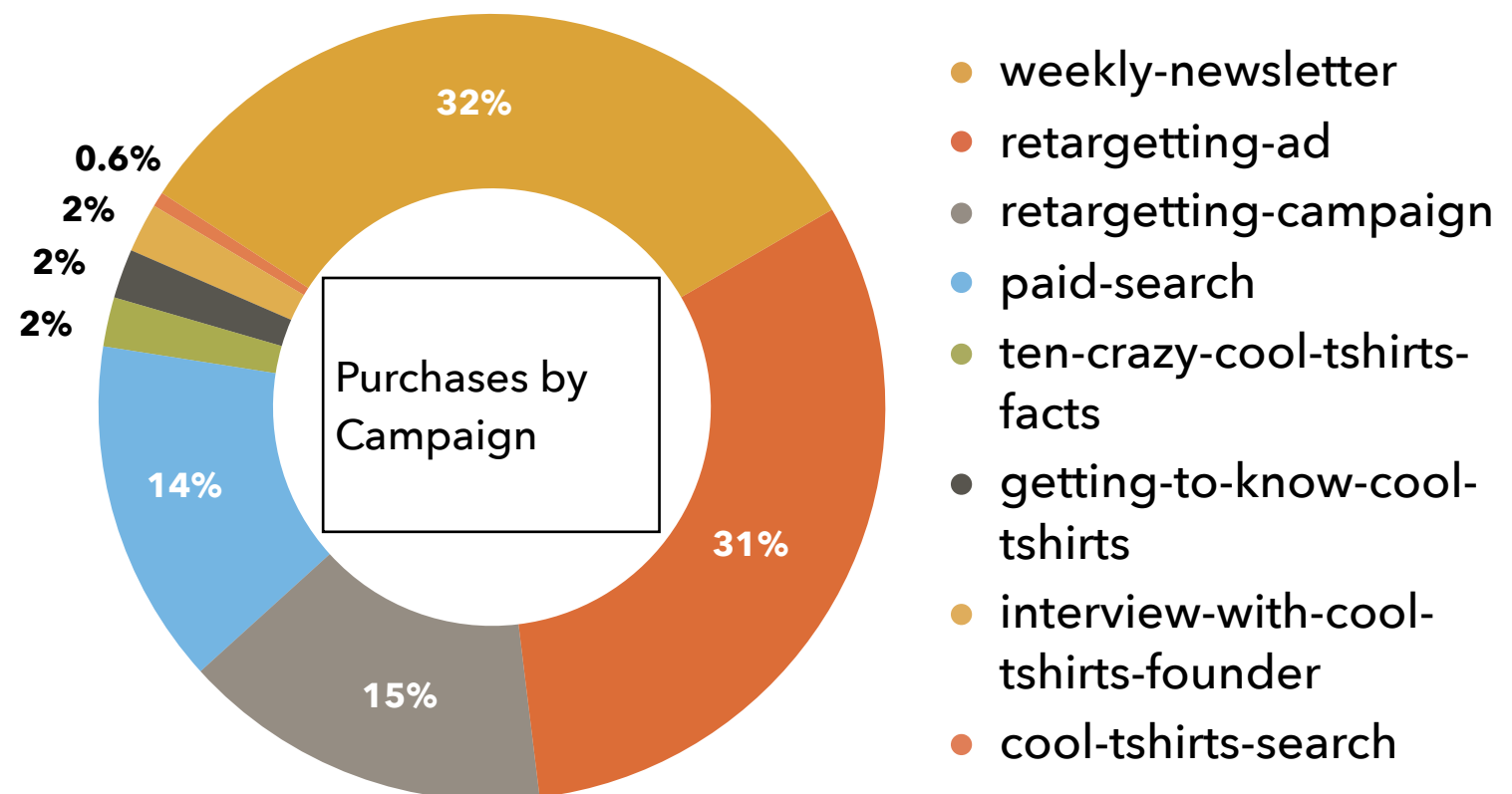


2.2 WHAT IS THE TYPICAL USER JOURNEY? (CONT.)

❖ Purchases by Campaign

- **“Reminder” types of campaigns correlates with 78% of purchases made by users:**
 - ✦ weekly-newsletter = 32%
 - ✦ retargeting-ad = 31%
 - ✦ retargeting-campaign = 15%
- **The top 2 last touch “Reminder” type of campaigns accounts for 63% of total purchases** (i.e., campaigns via email and via Facebook).

Query Results		
Source	Campaign	Last Touch 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



3. OPTIMIZE THE CAMPAIGN BUDGET

✦ Why CoolTShirts should reinvest in these 5 campaigns:

- CoolTShirts should reinvest in all 3 first touch "About" type of campaigns highlighted in peach below because these generate large amounts of initial activity from widely viewed sources, with all types of audiences. These sources help network CoolTShirts passively.
 - ✦ However, it may also help increase interest if they also sent these campaigns by email later on if users navigated to the site through other means (e.g., if the user discovered CoolTShirts from a Google search and signed up for their weekly newsletter, then maybe the first email after a week of inactivity could be one of the campaigns highlighted in peach below, which would convert that campaign from an "About" type to an "Reminder" type of campaign).
- CoolTShirts should also reinvest in the first 2 "Reminder" style of campaigns highlighted in pink below because these two campaigns accounts for 63% of purchases made by users. Something about these campaigns converts users from being a "visitor" to a "customer".
 - However, if additional funds are available, then it may be worth reinvesting in the 'retargetting-campaign' option as well since this campaign generates 15% of revenue. Though, experimenting with this campaign on different sources may help develop better strategies in increasing its value to contribute to purchases, such as sending a 'retargetting-campaign' every few months on each of the 3 sources that generate a higher initial traffic of users (medium, nytimes, buzzfeed).

Source	Campaign	First Touch	%	Last Touch	%	Purchases	%
medium	interview-with-cool-tshirts-founder	622	31%	184	9%	7	2%
nytimes	getting-to-know-cool-tshirts	612	31%	232	12%	9	2%
buzzfeed	ten-crazy-cool-tshirts-facts	576	29%	190	10%	9	2%
google	cool-tshirts-search	169	9%	60	3%	2	0.6%
email	weekly-newsletter	---	---	447	23%	115	32%
facebook	retargetting-ad	---	---	443	22%	113	31%
email	retargetting-campaign	---	---	245	12%	54	15%
google	paid-search	---	---	178	9%	52	14%