LEARN SQL FROM SCRATCH

codecademy

ALLIE BAYNES

CAPSTONE: ATTRIBUTION QUERIES

2019-03-11

Table of contents

- ABOUT COOLTSHIRTS
- THE USER JOURNEY
- OPTIMIZE CAMPAIGN BUDGET

code|cademy

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 reinvest in 5 campaigns. Which should they pick and why?

ABOUT COOLTSHIRTS

codecademy

About CoolTShirts

How many campaigns and sources does CoolTShirts use? How are they related?

- Campaigns drive awareness for a product in an effort to increase sales.

- The campaign messages are communicated through different channels, or sources, such as Facebook, websites, and email.

- In the case of CoolTShirts, each campaign is only being run through a single source, though campaigns in general may be ran through multiple sources.

```
1   --count campaigns, rename aggregate to alias
2   SELECT
3     COUNT(DISTINCT utm_campaign) as 'Campaign Count'
4   FROM page_visits;
5   --count sources
6   SELECT
7     COUNT(DISTINCT utm_source) as 'Source Count'
8   FROM page_visits;
9   --match source with each campaign to show relationship,
    columns are separated by a comma
10  SELECT DISTINCT utm_campaign as 'Campaign',
11    utm_source AS 'Source'
12  FROM page_visits;
```

Campaign Count	Source Count
8	6

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

ABOUT COOLTSHIRTS - CONT.

What pages are on the CoolTShirts website?

- From the **page_visits** database schema, we see that the data includes a column for **page_names.**

- Find the unique values in this column by running a simple SELECT DISTINCT query.

- The results show that there are 4 pages, which are only a subset of pages on the entire site. These are likely representative of the pages in which the campaign linked the user to.

- It is important to note that users will only reach the purchase page <u>after</u> they have completed a purchase.

- For later use, run COUNT query to determine number of unique site visitors using DISTINCT command.

```
1   --find unique website page names
2   SELECT DISTINCT page_name as 'Page Names'
3   FROM page_visits;
```

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

```
1   --count unique site visitors
2   SELECT COUNT(DISTINCT user_id) AS 'Unique Users'
3   FROM page_visits; |
```

Unique Users
1979

THE USER JOURNEY

codecademy

THE USER JOURNEY

How many first touches is each campaign responsible for?

- Use the first touch query and MIN(timestamp) to give the first timestamp for a specific user. This will help us understand which source & campaign led the potential customer to initially discover and click through to the site.
- Although there are 8 active campaigns running, only 4 of them were responsible for user first touches.

*Source column included per recommendation in project's Hint.

Source	Campaign	First 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

```sql
1   --create temp table first_touch that is the set of all
    1st touches by user id, join to page_visits table, & run
    query to count 1st touches associated with sources &
    campaigns, group by source & campaign, sort by count
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   SELECT
8     pv.utm_source AS 'Source',
9     pv.utm_campaign AS 'Campaign',
10    COUNT(utm_campaign) AS 'First Touches'
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  GROUP BY 1, 2
16  ORDER BY 3 DESC;
```

THE USER JOURNEY - CONT.

```
1    --create temp table last_touch that is the set of
     all 1st touches by user id, join to page_visits
     table, & run query to count 1st touches associated
     with sources & campaigns, group by source &
     campaign, sort by count
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    SELECT
8      pv.utm_source AS 'Source',
9      pv.utm_campaign AS 'Campaign',
10     COUNT(utm_campaign) AS 'Last Touches'
11   FROM last_touch lt
12   JOIN page_visits pv
13     ON lt.user_id = pv.user_id
14     AND lt.last_touch_at = pv.timestamp
15   GROUP BY 1, 2
16   ORDER BY 3 DESC;
```

Source	Campaign	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

How many last touches is each campaign responsible for?

- Adjust the first touch query and use MAX(timestamp) to give the last timestamp for a specific user. This helps us determine how the customer v drawn back to the site, especially when making a final purchase.
- All 8 campaigns were responsible for a last touch.
- While article-based campaigns accounted for 90% of first touches, retargeting campaigns account over half of last touches.

THE USER JOURNEY - CONT.

How many visitors make a purchase?
- 361 of the 1,979 visitors make a purchase, an 18% conversion rate.

```sql
1  --count unique visitors that make a purchase
2  SELECT COUNT(DISTINCT user_id) AS 'Visitors that Purchase'
3  FROM page_visits
4  WHERE page_name = '4 - purchase';
```

Visitors that Purchase
361

```sql
1   --create temp table to find last touches by user_id
2   WITH last_touch AS (
3       SELECT user_id,
4              MAX(timestamp) as last_touch_at
5       FROM page_visits
6   --use WHERE clause
7       WHERE page_name = '4 - purchase'
8       GROUP BY user_id)
9   --join pv table to temp table on user_id and
    timestamp, count last touches associated with
    sources & campaigns
10  SELECT
11    pv.utm_source AS 'Source',
12    pv.utm_campaign AS 'Campaign',
13    COUNT(utm_campaign) AS 'Count'
14  FROM last_touch lt
15  JOIN page_visits pv
16    ON lt.user_id = pv.user_id
17    AND lt.last_touch_at = pv.timestamp
18  GROUP BY 1, 2
19  ORDER BY 3 DESC;
```

Source	Campaign	Count
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

How many last touches on the purchase page is each campaign responsible for?
- Identify users with last touches when visiting the purchase page and associate them with the source and campaign responsible for the purchase.
- Results further confirm how article-based campaigns decrease in conversion rate.
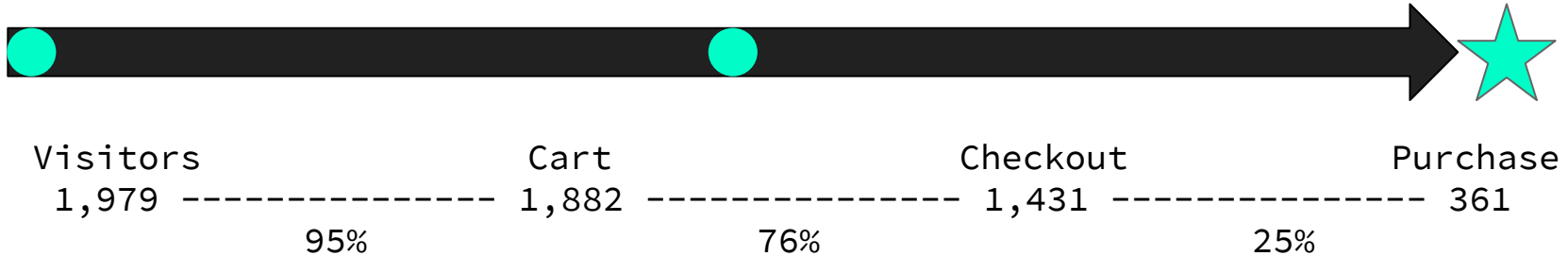
THE TYPICAL USER JOURNEY

When thinking about the typical user journey, consider the steps throughout the process.

- First touch: a user's first-ever interaction (attributed to a specific campaign and source)

- Opportunity creation: user intentionally moves toward a purchase decision (e.g. submits contact info, adds items to cart, or goes to checkout page)

- Conversion: user makes a purchase

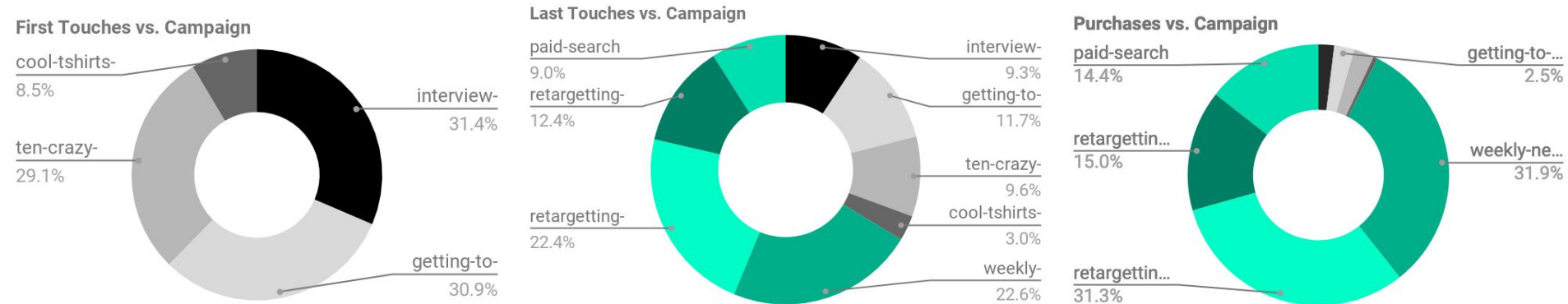**First Touch**　　　　　　　　　　**Opportunity Creation**　　　　　　　　　　**Conversion**

```
   Visitors          Cart            Checkout          Purchase
   1,979 -------------- 1,882 -------------- 1,431 -------------- 361
              95%                 76%                 25%
```

THE TYPICAL USER JOURNEY

- Article-based campaigns are responsible for sparking initial interest.
  - a. 91% of website first touches come from these articles
- Users need friendly reminders to revisit the site.
  - a. 95% of visitors return back to the site, with a majority touching the site 2-4 total times.
- Most purchases are the result of these campaigns that draw customers back to the site.
  - a. 78% of purchases come from retargeting campaigns

Total Site Touches	Visitors
3	1057
2	448
4	359
1	98
5	11
6	6

First Touches vs. Campaign

cool-tshirts- 8.5%
interview- 31.4%
ten-crazy- 29.1%
getting-to- 30.9%

Last Touches vs. Campaign

paid-search 9.0%
retargetting- 12.4%
retargetting- 22.4%
interview- 9.3%
getting-to- 11.7%
ten-crazy- 9.6%
cool-tshirts- 3.0%
weekly- 22.6%

Purchases vs. Campaign

paid-search 14.4%
retargettin... 15.0%
retargettin... 31.3%
getting-to-... 2.5%
weekly-ne... 31.9%

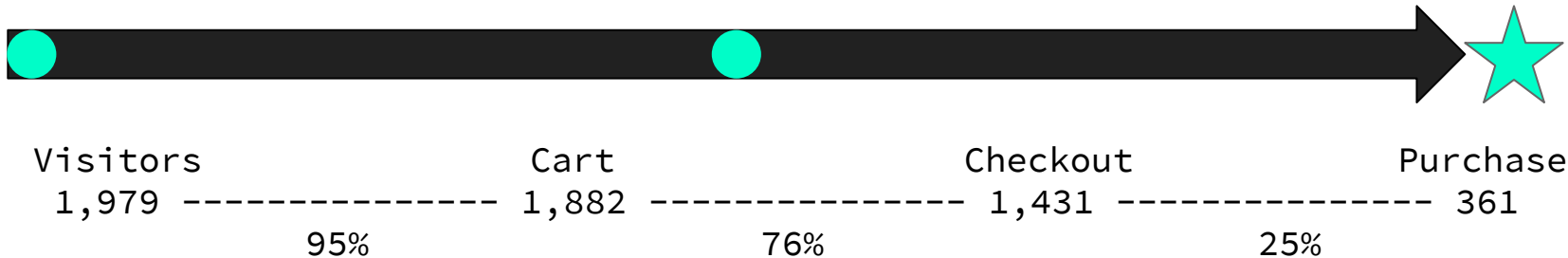# OPTIMIZE CAMPAIGN BUDGET

codecademy

OPTIMIZE CAMPAIGN BUDGET

- In the case of CoolTShirts, the last touch attribution is shown below as each user navigates their way through the site. There must be a strong interest in the product, as 95% of visitors put an item into their carts, so the current campaigns responsible for first touches (article-based) are performing well.
- The company should focus on refining retargeting campaigns. It appears that user information required for retargeting campaigns is not collected until they reach the checkout page.

First Touch   **Opportunity Creation**   **Conversion**

```
   Visitors              Cart            Checkout            Purchase
   1,979 -------------- 1,882 -------------- 1,431 -------------- 361
              95%                  76%                  25%
```

OPTIMIZE THE CAMPAIGN BUDGET – CONT.

- Arriving at the checkout page means the user was required to submit an email address.

- Consider requesting user's contact info earlier in the journey via pop-up, discount codes, etc. since retargeting campaigns result in more conversions.

- Would be interesting to know what type of contact info (email, facebook, etc.) was used for each visitor in the retargeting campaigns and how many were retargeted.

Campaign	Last Touches at Purchase	Last Touches at Checkout	Last Touches at Cart	Last Touches at Landing Page
interview-with-cool-tshirts-founder	7	31	513	623
getting-to-know-cool-tshirts	9	41	678	612
ten-crazy-cool-tshirts-facts	9	32	560	576
cool-tshirts-search	2	7	131	168
weekly-newsletter	115	450	0	0
retargetting-ad	113	445	0	0
retargetting-campaign	54	246	0	0
paid-search	52	179	0	0
TOTAL	**361**	**1,431**	**1,882**	**1,979**

OPTIMIZE THE CAMPAIGN BUDGET - CONT.

Which 5 campaigns should CoolTShirts re-invest in?

- Re-invest in 3 article-based campaigns: **interview-with-cool-tshirts-founder, getting-to-know-cool-tshirts, & ten-crazy-cool-tshirts-facts**
    a. Potential consumers are learning about the site through articles, with 91% of first touches attributed to these campaigns.
    b. Refine process for gathering consumer contact info so more initial visitors can be re-targeted. Optimize website to try and acquire emails sooner. Try a pop-up on the landing page to submit email address for a small discount.

- Re-invest in **weekly-newsletter** and **retargeting-ad**
    a. These campaigns alone contribute to 63% of purchases, so re-invest and refine campaigns.
    b. Evaluate content in newsletter and look for opportunities. Analyze why retargeting ad is performing better than the retargeting campaign and focus extra budget resources into further improving this campaign

Campaign	1st Touches	% 1st 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%
retargetting-ad	0	0%	443	22%	113	31%
retargetting-campaign	0	0%	245	12%	54	15%
paid-search	0	0%	178	9%	52	14%

THANKS!

codecademy