Table of Contents

- ## Get familiar with CoolTShirts
  - 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?

- ## 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?

- ## Optimize the campaign budget
  - CoolTShirts can re-invest in 5 campaigns. Which should they pick and why?

Get familiar with CoolTShirts

How many campaigns and sources does CoolTShirts use? Which source is used for each campaign?

CoolTShirts has 6 different sources and 8 different campaigns. The difference between the two is:
- A source is the type of contact made, such as an email or site which the campaign is on such as Buzzfeed.
- A campaign is a message designed to draw the attention of group of prospective customers towards your website.
- Sources can run multiple campaigns and campaigns can be run across multiple sources.

What pages are on their website?

There are four pages on their website: Landing page, Shopping cart, Checkout and Purchase.

These can be found by simply finding the distinct values of the page_name column using a SELECT DISTINCT query.

```
1   SELECT DISTINCT page_name AS 'Page Names'
2   FROM page_visits;
```

The page names shown appear to depict the process the user goes through and are a subset of actual pages on the site.

Campaigns	Source
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

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

What is the user journey?

How many first touches is each campaign responsible for?

Using MIN timestamp for each user allows us to identify each individual users first touch, which we can then join with page_visits to pull the source and campaign to group by and figure out the total number of first touches.

Source	Campaign	Total 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

```
WITH first_touch AS (
    SELECT user_id,
        MIN(timestamp) as first_touch_at
    FROM page_visits
    GROUP BY user_id),
ft_attr AS (
SELECT ft.user_id,
    ft.first_touch_at,
    pv.utm_source,
            pv.utm_campaign
FROM first_touch ft
JOIN page_visits pv
    ON ft.user_id = pv.user_id
    AND ft.first_touch_at = pv.timestamp
)
SELECT ft_attr.utm_source AS 'Source',
                ft_attr.utm_campaign AS
'Campaign',
      COUNT(*) AS 'Total touches'
FROM ft_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

What is the user journey? Pt 2.

How many last touches is each campaign responsible for?

Using MAX timestamp for each user allows us to identify each individual users last touch, which we can then join with page_visits to pull the source and campaign to group by and figure out the total number of last touches.

Source	Campaign	Total 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

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) as last_touch_at
    FROM page_visits
    GROUP BY user_id),
lt_attr AS (
SELECT lt.user_id,
    lt.last_touch_at,
    pv.utm_source,
            pv.utm_campaign
FROM last_touch AS 'lt'
JOIN page_visits AS 'pv'
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source AS 'Source',
                lt_attr.utm_campaign AS
'Campaign',
        COUNT(*) AS 'Total touches'
FROM lt_attr
GROUP BY 1, 2
ORDER BY 3 DESC;
```

What is the user journey? Pt 3.

How many visitors make a purchase?
361 Customers made a purchase, which equates to 18.24% (361/1979).

How many last touches on the purchase page is each campaign responsible for?
Breakdown of last touches shown below, obtained with the standard last touch queries but filtered by '4 - purchase' using a WHERE query.

Source	Campaign	Total Touches
email	weekly-newsletter	114
facebook	retargetting-ad	112
email	retargetting-campaign	53
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

```
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) AS 'last_touch_at'
    FROM page_visits
    GROUP BY user_id),
lt_attr AS (
  SELECT lt.user_id,
         lt.last_touch_at,
         pv.utm_source,
         pv.utm_campaign,
         pv.page_name
  FROM last_touch AS lt
  JOIN page_visits AS pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
       lt_attr.utm_campaign,
       page_name,
       COUNT(*)
FROM lt_attr
WHERE page_name = '4 - purchase'
GROUP BY 2, 3
ORDER BY 4 DESC;
```

What is the typical user journey?

What is the typical user journey?
On the CoolTShirts website, a typical journey follows this format:
Landing Page -> Shopping Cart -> Checkout -> Purchase

Each page has a substantial drop off, with the move from Checkout to Purchase being the biggest, below depicts this:

Page Names	Total Touches
1 - landing_page	2000
2 - shopping_cart	1900
3 - checkout	1431
4 - purchase	361

```
1  SELECT page_name, MIN(timestamp), COUNT(*)
2  FROM page_visits
3  GROUP BY page_name;
```

```sql
WITH last_touch AS (
    SELECT user_id,
        MAX(timestamp) AS 'last_touch_at'
    FROM page_visits
    GROUP BY user_id),
lt_attr AS (
  SELECT lt.user_id,
         lt.last_touch_at,
         pv.utm_source,
         pv.utm_campaign,
         pv.page_name
  FROM last_touch AS lt
  JOIN page_visits AS pv
    ON lt.user_id = pv.user_id
    AND lt.last_touch_at = pv.timestamp
)
SELECT lt_attr.utm_source,
       lt_attr.utm_campaign,
       page_name,
       COUNT(*)
FROM lt_attr
WHERE page_name = '4 - purchase'
GROUP BY 2, 3
ORDER BY 4 DESC;
```

Optimize the campaign budget

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

The short answer to this is that they should re-invest in the following campaigns:
Interview-with-cool-tshirts-founder
Getting-to-know-cool-tshirts
Ten-crazy-cool-tshirts-facts
Weekly-newsletter
Retargetting-ad

These yield the highest results, with the first three providing the highest first touch percentage and the last two have the highest return rate, with purchase % of just over 30% for both.

Between the highest first touch and total purchase count it would be optimal to re-invest in these 5 campaigns as they yield the most success.

Campaign	Total Touches
interview-with-cool-tshirts-founder	622
getting-to-know-cool-tshirts	612
ten-crazy-cool-tshirts-facts	576

Campaign	Purchases
weekly-newsletter	114
retargetting-ad	112