short shirt - learn sql from scratch capstone

cool tshirts - j puryear

toc

1. sourcing - familiarity prints contempt
2. journey - users travails
3. the way ahead - optimizing the campaign

1. sourcing - familiarity prints contempt

sourcing 1.1

sources numbers

- utm_source - Identifies which touchpoint sent the traffic (e.g. google, email, or facebook)

```
SELECT count(DISTINCT
utm_source) AS sources
FROM page_visits;
```

sources
6

sourcing 1.2

campaigns numbers

- utm_campaign - Identifies the specific ad or email blast (e.g. retargetting-ad or weekly-newsletter)

```
SELECT count(DISTINCT
utm_campaign) AS campaigns
FROM page_visits;
```

campaigns
8

sourcing 1.3

sources and campains relations

campaigns	sources	cnt
ten-crazy-cool-tshirts-facts	buzzfeed	1198
retargetting-campaign	email	300
weekly-newsletter	email	565
retargetting-ad	facebook	558
cool-tshirts-search	google	313
paid-search	google	231
interview-with-cool-tshirts-founder	medium	1178
getting-to-know-cool-tshirts	nytimes	1349

```
SELECT utm_campaign AS
campaigns,
        utm_source AS
sources,
        count(*) as cnt
FROM page_visits
GROUP BY 2, 1;
```

sourcing 1.4

pages on cool tshirts

```
SELECT DISTINCT page_name
FROM page_visits;
```

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

2. journey - users travails

journey 2.1

first touch by campaigns

campaign	cnt
cool-tshirts-search	169
getting-to-know-cool-tshirts	612
interview-with-cool-tshirts-founder	622
ten-crazy-cool-tshirts-facts	576

```
WITH firsts AS (
     SELECT user_id,
            MIN(timestamp) AS
touch
     FROM page_visits
     GROUP BY user_id)
SELECT utm_campaign AS
campaign,
       COUNT(DISTINCT
v.user_id) AS cnt
FROM page_visits v
     JOIN firsts f
         ON f.user_id =
v.user_id
         AND touch = timestamp
GROUP BY 1;
```

journey 2.2

last touch by campaign

campaign	cnt
cool-tshirts-search	60
getting-to-know-cool-tshirts	232
interview-with-cool-tshirts-founder	184
paid-search	178
retargetting-ad	443
retargetting-campaign	245
ten-crazy-cool-tshirts-facts	190
weekly-newsletter	447

```
WITH firsts AS (
    SELECT user_id,
        MIN(timestamp) AS
touch
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign AS campaign,
    COUNT(DISTINCT v.user_id)
AS cnt
FROM page_visits v
    JOIN firsts f
        ON f.user_id =
v.user_id
        AND touch = timestamp
GROUP BY 1;
```

journey 2.3

customers who purchased

```
SELECT COUNT(DISTINCT user_id)
as purchasers
FROM page_visits
WHERE page_name LIKE '4%';
```

purchasers
361

journey 2.4

campaigns that resulted in final purchases

campaign	cnt
cool-tshirts-search	2
getting-to-know-cool-tshirts	9
interview-with-cool-tshirts-founder	7
paid-search	52
retargetting-ad	113
retargetting-campaign	54
ten-crazy-cool-tshirts-facts	9
weekly-newsletter	115

```
WITH lasts AS (
    SELECT user_id,
        MAX(timestamp) AS
touch
    FROM page_visits
    GROUP BY user_id)
SELECT utm_campaign AS
campaign,
    COUNT(DISTINCT
v.user_id) AS cnt
FROM page_visits v
    JOIN lasts l
        ON l.user_id =
v.user_id
        and touch = timestamp
GROUP BY 1;
```

journey 2.5

typical user journey

- the most common user runs away by after visiting the landing page
- the second most common user shops a bit and goes the the shopping cart
  - by this point some have signed up for the weekly newsletter
- amazingly a large number go directly from the newsletter to checkout

3. the way ahead - optimizing the campaign

the way ahead 1.1

Assuming this data is correct and our premises on it are as well it would appear that no one shows up except for unique content: local search, getting to know, the interview, and 10 crazy facts. Without these there are no customers. The best way to get them to come back is the newsletter and retargeting ad. I can't imagine that unique content like that can be rewritten, but the founder seems to draw attention. Focusing on that seems wise. So adding new content to get first time viewers and continuing the newsletter and retargeting seem like the top picks. The retargeting campaign and paid search are the toss up more time may tell but currently the campaign is on top.