codecademy

Capstone : Warby Parker

Learn SQL from Scratch

Kwun Lai

July, 2018

Table of Contents

1.1 Quiz Funnel

question	user_id	response
1. What are you looking for?	005e7f99-d48c-4fce-b605-10506c85aaf7	Women's Styles
2. What's your fit?	005e7f99-d48c-4fce-b605-10506c85aaf7	Medium
3. Which shapes do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Round
4. Which colors do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Two-Tone
1. What are you looking for?	00a556ed-f13e-4c67-8704-27e3573684cd	I'm not sure. Let's skip it.
2. What's your fit?	00a556ed-f13e-4c67-8704-27e3573684cd	Narrow
5. When was your last eye exam?	00a556ed-f13e-4c67-8704-27e3573684cd	<1 Year
3. Which shapes do you like?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Square
5. When was your last eye exam?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	<1 Year
2. What's your fit?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Medium

```
SELECT * FROM survey LIMIT 10;
/*
Query selects first 10 entries of the 'survey'
database and displays it
*/
```

The column names for the 'survey' table are:

- question: string holding the question asked
- user_id: unique user identifier
- response: string holding user's response

1.2 Quiz Funnel

500 users answered Question 1
475 users answered Question 2
380 users answered Question 3
361 users answered Question 4
270 users answered Question 5

Question	# Answered
1. What are you looking for?	500
2. What's your fit?	475
3. Which shapes do you like?	380
4. Which colors do you like?	361
5. When was your last eye exam?	270

```
SELECT
question AS 'Question',
count(user_id) AS '# Answered'
FROM survey
GROUP BY 1;

/*
Query selects column question, capitalizes it for
style, and counts distinct user_ids and renames the
column as "# Answered" for readability. Then it sorts
the columns by 'question'.
*/
```

1.3 Quiz Funnel

The last question (When was your last eye exam?) had the lowest completion rate. 74.79% of users answered the question. Why? The first 4 questions are easy to answer. They ask the user likes, what their personal style is and personal preferences. These answers come instantly and are easily answerable after thinking for a few seconds. Whereas the last question asks for information that takes more than a few seconds to retrieve. The user might have to navigate away from the page to look up past calendars, pull up their phone and scroll through appointments or even physically move to ask a spouse. Also, with this last question, asking about when a user's last eye exam is quite the jump from asking what color they like their frames. It is inconsistent with eyewear style questions. Maybe this question is better asked after the user is done selecting what frames they like the most; near the end of the checkout process.

Question	# Answered	% Answered
1. What are you looking for?	500	100.00%
2. What's your fit?	475	95.00%
3. Which shapes do you like?	380	80.00%
4. Which colors do you like?	361	95.00%
5. When was your last eye exam?	270	74.79%

2. Home-Try-On Funnel

2.1 Home Try-On Funnel

The column names for each table are:
quiz.user_id: unique user identifier
quiz.style: a list of gender eyewear styles (optional)
quiz.fit: a list of facial fits
quiz.shape: a list of possible eyewear shapes
quiz.color: a list of possible color options

home_try_on.user_id: unique user identifier
home_try_on.number_of_pairs: whether the user has 3 or 5 pairs (A/B)
home_try_on.address: address of the user

purchase.user_id: unique user identifier
purchase.product_id: unique product identifier
purchase.style: gender of eyewear style
purchase.model_name: model name of eyewear
purchase.color: color of eyewear
purchase.price: price of the eyewear

```
SELECT * FROM quiz LIMIT 5;
SELECT * FROM home_try_on LIMIT 5;
SELECT * FROM purchase LIMIT 5;
/*
This query selects first 5 rows of the quiz,
home_try_on and purchase tables.
*/
```

2.2 Home Try-On Funnel - Query

This query combines the quiz, purchase and home_try_on tables and does some formatting so it easier to read by humans. We use case statements for h.user_id and p.user_id fields because we what true and false instead of 1s and 0s. We joined these via the user_id primary key from each table, as they are unique to each user. The output of this query (the first 10 entries) is on the next slide.

```
SELECT DISTINCT q.user_id,
   CASE
      WHEN h.user_id IS NOT NULL THEN 'True'
      ELSE 'False'
   END AS 'is_home_try_on',
   h.number_of_pairs,
   CASE
      WHEN p.user_id IS NOT NULL THEN 'True'
      ELSE 'False'
   END AS 'is_purchase'
FROM quiz q
LEFT JOIN home_try_on h
   ON q.user_id = h.user_id
LEFT JOIN purchase p
   ON p.user_id = q.user_id
LIMIT 10;
/*
Case statement used to enhance 'readability'. Instead
of 1 and 0, we have true or false
*/
```

2.2 Home Try-On Funnel - Query Results

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	True	3 pairs	False
291f1cca-e507-48be-b063-002b14906468	True	3 pairs	True
75122300-0736-4087-b6d8-c0c5373a1a04	False		False
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	True	5 pairs	False
ce965c4d-7a2b-4db6-9847-601747fa7812	True	3 pairs	True
28867d12-27a6-4e6a-a5fb-8bb5440117ae	True	5 pairs	True
5a7a7e13-fbcf-46e4-9093-79799649d6c5	False		False
0143cb8b-bb81-4916-9750-ce956c9f9bd9	False		False
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	True	5 pairs	False
b1dded76-cd60-4222-82cb-f6d464104298	True	3 pairs	False

2.3 Home Try-On Funnel - Analysis 1 - Conversion Rates

```
WITH warby_funnel_t AS (
SELECT DISTINCT q.user_id,
   h.user_id IS NOT NULL AS
'is_home_try_on',
   h.number_of_pairs,
   p.user_id IS NOT NULL AS
'is_purchase'
FROM quiz q
LEFT JOIN home_try_on h
   ON q.user_id = h.user_id
LEFT JOIN purchase p
   ON p.user_id = q.user_id
)
SELECT
COUNT(*) AS 'Users',
SUM(is_home_try_on) AS 'Total Home Try
On',
SUM(is_purchase) AS 'Purchased'
FROM warby_funnel_t;
```

```
WITH warby_funnel_3 AS (
SELECT DISTINCT q.user_id,
   h.user_id IS NOT NULL AS
'is_home_try_on',
   h.number_of_pairs,
   p.user_id IS NOT NULL AS
'is_purchase'
FROM quiz q
LEFT JOIN home_try_on h
   ON q.user_id = h.user_id
LEFT JOIN purchase p
   ON p.user_id = q.user_id
WHERE h.number_of_pairs = '3 pairs'
)
SELECT
COUNT(*) AS 'Users',
SUM(is_home_try_on) AS '3 Pairs Home
Try On',
SUM(is_purchase) AS 'Purchased'
FROM warby_funnel_3;
```

```
WITH warby_funnel_5 AS (
SELECT DISTINCT q.user_id,
   h.user_id IS NOT NULL AS
'is_home_try_on',
   h.number_of_pairs,
   p.user_id IS NOT NULL AS
'is_purchase'
FROM quiz q
LEFT JOIN home_try_on h
   ON q.user_id = h.user_id
LEFT JOIN purchase p
   ON p.user_id = q.user_id
WHERE h.number_of_pairs = '5 pairs'
)
SELECT
COUNT(*) AS 'Users',
SUM(is_home_try_on) AS '5 Pairs Home
Try On',
SUM(is_purchase) AS 'Purchased'
FROM warby_funnel_5;
```

2.3 Home Try-On Funnel - Analysis 1 - Conversion Rates

Here is the output of the previous 3 queries. We used 3 separate queries to independently test the outcome of each funnel; warby_funnel_3 being 3 pairs and warby_funnel_5 being 5 pairs. What does this tell us? 75% of users used Home Try-On. 25% did not. Of the 750 users who tried on, 66% of them eventually went on to purchase glasses. User percentage who opted for Home Try-On between 3 and 6 pairs is exactly the same; meaning offering a different number of pairs did not influence their selection. What drove them to participate was the fact that a free trial was offered. What is interesting is the percentage of users to ended up buying after the Home Try-On. With 3 pairs, 53% of users ended up buying. With 5 pairs, it is closer to 80% of users. That is a huge difference. We can conclude that giving more options equals a higher chance that a purchase will be made. It gives the user greater room for experimentation and fine tuning their personal style.

Users	Total Home Try On	Purchased
1000	750	495
Users	**3 Pairs Home Try On**	**Purchased**
379	379	201
Users	**5 Pairs Home Try On**	**Purchased**
371	371	294

2.3 Home-Try-On Funnel - Analysis 2 - Style Popularity

We can query the style table with the queries on the right to see which is the most popular in whatever category. The last query allows you to find how many of each given column was chosen. From the queries of the data, we see that women's styles are slightly more popular than men's style, 46% to 43% respectively. Rectangular glasses (39%) with narrow fit (40%) are the most popular styles chosen from the table. The most popular color being tortoise shell at 29% slightly edging out plain black at 28%. What this tells us women's tortoise color, narrow fit, rectangular glasses are the most popular. We either need to boost our styles in that shape or more aggressively market the other styles.

```
SELECT
   style,
   fit,
   shape,
   color,
   COUNT(user_id)
FROM quiz
GROUP BY 1,2,3,4;
```

2.3 Home-Try-On Funnel - Analysis 3 - Purchase Trends

We can query the purchase table with the query on the right to see which is the most bought. The most bought men's style was the most expensive one, Dawes. The most bought women's style was the mid range Eugene Narrow. Those two styles represent 44% of the men's sales and 46% of the women's.

```
SELECT
   product_id,
   style,
   model_name,
   color,
   price,
   COUNT(user_id)
FROM purchase
GROUP BY 1,2,3,4,5;
```