codecademy

Capstone: Warby Parker

Learn SQL from Scratch

Kevin Wheaton

1/28/19

Table of Contents

1. Quiz questions

1.1 What does this table tell us?

This table shows us the five questions that are visible to users and their responses. Using this data you can cater to each user individually by recommending them items based on their responses.

```
SELECT *
FROM survey
LIMIT 10;
```

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?	005e7f99-d48c-4fce-b605-10506c85aaf7	Round
4. Which colors do you like?	005e7f99-d48c-4fce-b605-10506c85aaf7	Two-Tone
5. When was your last eye exam?	005e7f99-d48c-4fce-b605-10506c85aaf7	<1 year

1.2 Where did users give up in the quiz funnel?

This table shows us the total amount of responses to each question and their completion rate. You can see that the third and fifth question have the lowest completion rate. This could be for a few reasons:

- The user just wants to browse the selections and only needs to give the bare minimum information which doesn't need to include shape preferences.
- If the user wants to go more in depth by answering the shape question, then its not too far of a stretch to answer what colors the user prefers.
- The last question can be seen as an attack or an insult to the user. It is also a somewhat personal question. Since you just answered four questions about product preferences, this question can imply your taste is bad. It is also a question that even friends will not just ask you out of pure interest. Unlike the other questions, this one can have a perceived wrong answer and make the user feel bad.

Question	Count(Distinct user_id)	Completion Rate
1. What are you looking for	500	100%
2. What's your fit?	475	95%
3. Which shapes do you like?	380	80%
4. Which colors do you like?	361	95%
5. When was your last eye exam?	270	74.79%

```
SELECT question,
Count(distinct user_id)
From survey
Group By question;
```

2. Home Try-On A/B Test

2.1 Setting up the A/B test

This gives us three separate tables, each with their own columns and a user_id to tie them together. Now we see each user_id in all three tables can be joined together, so that the information in all three tables can be combined into one. Below is a picture of the table that the query to the right will show.

Query Results

user_id	style	fit	shape	color
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	Women's Styles	Medium	Rectangular	Tortoise
291f1cca-e507-48be-b063-002b14906468	Women's Styles	Narrow	Round	Black
75122300-0736-4087-b6d8-c0c5373a1a04	Women's Styles	Wide	Rectangular	Two-Tone
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	Women's Styles	Narrow	Square	Two-Tone
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black

user_id	number_of_pairs	address
d8addd87-3217-4429-9a01-d56d68111da7	5 pairs	145 New York 9a
f52b07c8-abe4-4f4a-9d39-ba9fc9a184cc	5 pairs	383 Madison Ave
8ba0d2d5-1a31-403e-9fa5-79540f8477f9	5 pairs	287 Pell St
4e71850e-8bbf-4e6b-accc-49a7bb46c586	3 pairs	347 Madison Square N
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs	182 Cornelia St

user_id	product_id	style	model_name	color	price
00a9dd17-36c8-430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black	150
00e15fe0-c86f-4818-9c63-3422211baa97	7	Women's Styles	Lucy	Elderflower Crystal	150
017506f7-aba1-4b9d-8b7b-f4426e71b8ca	4	Men's Styles	Dawes	Jet Black	150
0176bfb3-9c51-4b1c-b593-87edab3c54cb	10	Women's Styles	Eugene Narrow	Rosewood Tortoise	95
01fdf106-f73c-4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black	150

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

2.2 Combining all three Tables

By using a left join we can get all the records from the first table, which is quiz, and then get the records from each subsequent table from the selected columns. We get NULL as data when there are no related records. There are several ways Warby-Parker can use this data:

- You can calculate purchase rates between customers who tried the home_try_on and those who didn't
- You can calculate if customers who had more pairs of shoes would be more likely to purchase more products.
- You can calculate if answering the quiz questions has a better chance of making a sale by comparing conversion rates from the user_id's who answered the question, and those who didn't, to user_id's who made a purchase.

Query Results

user_id	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	NULL	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	NULL	False
0143cb8b-bb81-4916-9750-ce956c9f9bd9	False	NULL	False
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	True	5 pairs	False
b1dded76-cd60-4222-82cb-f6d464104298	True	3 pairs	False

```
SELECT Distinct quiz.user_id,
   CASE
      WHEN
      home_try_on.user_id IS NOT NULL
      THEN 'True'
      ELSE 'False'
   END AS home_try_on,
   CASE
      WHEN
      number_of_pairs.user_id IS NULL
      THEN 'NULL'
      ELSE 'number_of_pairs'
   END AS number_of_pairs,
   CASE
      WHEN
      purchase.user_id IS NOT NULL
      THEN 'True'
      ELSE 'False'
   END AS is_purchase
FROM quiz
LEFT JOIN home_try_on
   ON home_try_on.user_id = quiz.user_id
LEFT JOIN purchase
   ON purchase.user_id = home_try_on_id
LIMIT 10;
```