codecademy/PRO

Codecademy Intensive: Learn SQL from Scratch

Terence Bennett
July 17th, 2018

Purpose: The following slides show my answers for Codecademy's 'Funnel with Warby Parker' problem set.

1. Here is a breakdown of the Survey table.

```
# Returns all columns first 10 rows of Survey table
SELECT *
FROM survey
LIMIT 10;
```

Query Results		
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

2.   Here is a breakdown of the quiz funnel and the number of users who answered each question.

Three fields are returned by the query: Question, # of User Ids, and # of Responses

500 participants started the survey and 270 participants ended the survey. 230 participants did not complete the survey.

```
project.sqlite                                    ⤢

1    /*Returns Funnels metrics for quiz*/
2    SELECT question,
3    COUNT(user_id) AS 'User Ids',
4    COUNT(response) AS 'Responses'
5    FROM survey
6    WHERE response IS NOT NULL
7    GROUP BY question;
```

Query Results		
question	User Ids	Responses
1. What are you looking for?	500	500
2. What's your fit?	475	475
3. Which shapes do you like?	380	380
4. Which colors do you like?	361	361
5. When was your last eye exam?	270	270
Database Schema		
home_try_on		750 rows

3.   This excel table was created to show the drop off rate throughout the quiz funnel. Question 5 is highlighted for having the highest drop off rate. This high drop off rate is likely due to the fact that many people do not know the date of their last exam. Removing or changing this question would likely lead to higher completion rates.

Questions	Number of participants that completed each questions	Completion rate	Drop off rate
1. What are you looking for?	500	100.00%	0.00%
2. What's your fit?	475	95.00%	5.00%
3. Which shapes do you like?	380	80.00%	20.00%
4. Which colors do you like?	361	95.00%	5.00%
5. When was your last eye exam?	270	74.79%	25.21%

4. Below are screenshots of SQL code and tables.

```
project.sqlite

1  SELECT *
2  FROM purchase
3  LIMIT 5;
4
```

Query Results

user_id	product_id	style	model_name	color
430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black
818-9c63-3422211baa97	7	Women's Styles	Lucy	Elderflower Crystal
4b9d-8b7b-f4426e71b8ca	4	Men's Styles	Dawes	Jet Black
4b1c-b593-87edab3c54cb	10	Women's Styles	Eugene Narrow	Rosewood Tortoise
4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black

```
project.sqlite

1  SELECT *
2  FROM home_try_on
3  LIMIT 5;
4
```

Query Results

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

```
project.sqlite

1  SELECT *
2  FROM quiz
3  LIMIT 5;
4
```

Query Results

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

5. This query is a LEFT JOIN of the three tables on user_id. Started from the left, the fields show the user_id, a binary return of whether they tried them on at home, the number of pairs they tried on, and whether or not they purchased glasses.

This query shows 70% of users tried on glasses at home and 43% of those that tried them on at home decided to buy a pair of glasses.

```
 1   SELECT DISTINCT
 2     q.user_id,
 3     h.user_id IS NOT NULL AS 'is_home_try_on',
 4     number_of_pairs,
 5     p.user_id IS NOT NULL AS 'is_purchase'
 6   FROM quiz AS 'q'
 7   LEFT JOIN home_try_on AS 'h'
 8   ON q.user_id = h.user_id
 9   LEFT JOIN purchase AS "p"
10   ON p.user_id = h.user_id
11   LIMIT 10;
```

project.sqlite

Query Results			
user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	1	3 pairs	0
291f1cca-e507-48be-b063-002b14906468	1	3 pairs	1
75122300-0736-4087-b6d8-c0c5373a1a04	0	Ø	0
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	1	5 pairs	0
ce965c4d-7a2b-4db6-9847-601747fa7812	1	3 pairs	1
28867d12-27a6-4e6a-a5fb-8bb5440117ae	1	5 pairs	1
5a7a7e13-fbcf-46e4-9093-79799649d6c5	0	Ø	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9	0	Ø	0
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	1	5 pairs	0
b1dded76-cd60-4222-82cb-f6d464104298	1	3 pairs	0
Database Schema			

6. This query borrows from the previous LEFT JOIN. Starting from the left, it lists the number of people who browsered, who decided to try on, who decided to purchase, and then the funnels between each of the three categories. Based on this data, the funnels appear to be very effective. Questions 5 on the quiz should be updated to increase completion rated and Warby Parker should spend money on advertising to increased throughput on their funnels.

```
1   WITH funnel AS (SELECT DISTINCT
2     q.user_id,
3     h.user_id IS NOT NULL AS 'is_home_try_on',
4     number_of_pairs,
5     p.user_id IS NOT NULL AS 'is_purchase'
6   FROM quiz AS 'q'
7   LEFT JOIN home_try_on AS 'h'
8   ON q.user_id = h.user_id
9   LEFT JOIN purchase AS "p"
10  ON p.user_id = h.user_id)
11
12  SELECT COUNT(user_id) AS 'num_browse',
13  SUM(is_home_try_on) AS 'num_try_on',
14  SUM(is_purchase) AS 'num_purchase',
15  1.0 * SUM(is_home_try_on) / COUNT(user_id) AS
    'Funnel 1',
16  1.0 * SUM(is_purchase) / SUM(is_home_try_on) AS
    'Funnel 2'
17  FROM funnel;
```

project.sqlite

Query Results

num_browse	num_try_on	num_purchase	Funnel 1	Funnel 2
1000	750	495	0.75	0.66

Database Schema

home_try_on	750 rows
user_id	TEXT
number_of_pairs	TEXT
address	TEXT

purchase	495 rows
user_id	TEXT
product_id	INTEGER
style	TEXT
model_name	TEXT
color	TEXT
price	INTEGER

survey	1986 rows
question	TEXT
user_id	TEXT
response	TEXT

quiz	1000 rows