codecademy

Warby Parker

Analyze Data with SQL

Rodolfo Mares

June 2020

Table of Contents

1. Quiz Funnel

1.1 Table Structure

Style Quiz uses a table called ***survey*** which contains all user's responses to the different questions they asked in their Style Quiz:

- What are you looking for?
- What is your fit?
- Which shapes do you like?
- Which colors do you like?
- When was your last eye exam?

```
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

1.2 Number of Responses for Each Question

Users will “give up” at different points in the survey.

- How many users move from Question 1 to Question 2, etc.?
- What is the number of responses for each question?

```
SELECT question,
        COUNT(DISTINCT user_id) AS users
 FROM survey
 GROUP BY question;
```

Query Results

question	users
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

1.3 Percentage of users who answers each question

Calculate the percentage of users who answer each question.

- Which question(s) of the quiz have a lower completion rates?

Question #3 and Question #5 are the ones with the lower completion rates.

- What do you think is the reason?

The 3 options available in Question #3 (Round, Rectangular & Square) are generic cartoonish drawings and look quite similar. As a frame user, until you see the product you can decide which one, you like to try.

With regards to Question #5, maybe most of the users don't know or haven't had an eye exam and don't want to share that information.

Question	Percentage Completed
1. What are you looking for?	100%
2. What's your fit?	95%
3. Which shapes do you like?	80%
4. Which colors do you like?	95%
5. When was your last eye exam?	75%

2. Home Try-On Funnel

2.1 Tables Structure

The data is distributed in 3 tables:

- quiz
- home_try_on
- purchase

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

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

2.2 LEFT JOIN all tables

Create a new table with the following layout:

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc	True	3	False
291f1cca	True	5	False
75122300	False	NULL	False

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

```
SELECT 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 AS 'q'
LEFT JOIN home_try_on AS 'h'
  ON q.user_id = h.user_id
LEFT JOIN purchase AS 'p'
  ON h.user_id = p.user_id
LIMIT 10;
```

3. Overall Conversion Rates

3.1 From Quiz to Home_Try_On

Analyze from the total number of users that took the quiz.

- How many move to Home_Try_On?
- How many refused?

Query Results

Quiz_Customers	is_home_try_on
250	False
750	True

75 % take a chance

```
WITH data_analysis AS (
  SELECT 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 AS 'q'
  LEFT JOIN home_try_on AS 'h'
    ON q.user_id = h.user_id
  LEFT JOIN purchase AS 'p'
    ON h.user_id = p.user_id)

SELECT COUNT(*) AS Quiz_Customers,
  is_home_try_on
FROM data_analysis
GROUP BY is_home_try_on;
```

3.2 Home_Try_On to Purchase

Analyze from the total number of Home_Try_On.

- How many made a purchase?
- How many did not make a purchase?

Query Results	
Quiz_Customers	is_purchase
255	False
495	True

66 % success rate!

```
WITH data_analysis AS (
  SELECT 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 AS 'q'
  LEFT JOIN home_try_on AS 'h'
    ON q.user_id = h.user_id
  LEFT JOIN purchase AS 'p'
    ON h.user_id = p.user_id)

SELECT COUNT(*) AS Quiz_Customers,
  is_purchase
FROM data_analysis
WHERE is_home_try_on = 'True'
GROUP BY is_purchase;
```

3.3 Which is better, 3 or 5 pairs?

Analyze from the total number of Purchases.

- How many where made with 3 pairs?
- How many where made with 5 pairs?

Query Results

Customers	number_of_pairs
201	3 pairs
294	5 pairs

5 pairs by 18.8%

```
WITH data_analysis AS (
  SELECT 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 AS 'q'
  LEFT JOIN home_try_on AS 'h'
    ON q.user_id = h.user_id
  LEFT JOIN purchase AS 'p'
    ON h.user_id = p.user_id)

SELECT COUNT(*) AS Customers,
  number_of_pairs
FROM data_analysis
WHERE is_home_try_on = 'True' AND
      is_purchase = 'True'
GROUP BY number_of_pairs;
```

4. Trends

4.1 Women or Men?

- Which is the most common Style Pruchase?

Women

- Which is the most common Style Browse?

Women

```
SELECT COUNT(*) AS Total,
       style AS Style_Purchased
FROM purchase
GROUP BY style
ORDER BY style DESC;

SELECT COUNT(*) AS Total,
       style AS Style_Browsed
FROM quiz
GROUP BY style
ORDER BY style DESC;
```

Query Results

Total	Style_Purchased
252	Women's Styles
243	Men's Styles

Query Results

Total	Style_Browsed
469	Women's Styles
432	Men's Styles
99	I'm not sure. Let's skip it.

4.2 Top Selling Models

- Which is the Top Selling Model?

Eugene Narrow

```
SELECT model_name,
       COUNT(*) AS Total
FROM purchase
GROUP BY model_name
ORDER BY Total DESC;
```

Query Results	
model_name	**Total**
Eugene Narrow	116
Dawes	107
Brady	95
Lucy	86
Olive	50
Monocle	41