codecademy

Warby Parker Usage Funnels

Learn SQL from Scratch

Kameron Canbaz

January 27, 2019

Table of Contents

1. Getting to Know the Data

1.1 The `quiz` Table

Using a simple `SELECT` statement reveals the structure of the `quiz` table.

- This enables me to get familiar with the table structure and to know the types of data available.

```
SELECT *
FROM quiz
LIMIT 10;
```

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

2. Exploration and Insights

1.1 Quiz Funnel, contents

Similarly, another simple `SELECT` statement reveals the structure of the `survey` table.
- There are only three columns within this table.

```
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?	00a556ed-f13e-4c67-8704-27e3573684cd	Round
4. Which colors do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Two-Tone

1.2 Quiz Funnel, question counts

Understanding how many users answer each question is useful, as it can reveal where users stop the survey.

- Looking over the query results, it's apparent that there's a drop off in answers as soon as the third question.

```
SELECT question, COUNT(user_id) AS 'num_users'
FROM survey
GROUP BY 1;
```

question	num_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 Quiz Funnel, question counts (continued)

The table to the right shows the completion rate for each question, and below are my findings.

Which question(s) of the quiz have lower completion rates?
- Completion rates degrade after the first question, but clearly the final three questions have the lowest rates, all falling under 20%.

What do you think is the reason?
- The first two questions are rather simple and don't require much thought from the user. At question three, s/he is required to make decisions based on several options, thus making the question more involved.
- Finally, the last question requires a user to remember when her/his last exam was (if s/he had one). This information isn't likely to be readily available, and such will require the user to make an effort to answer it.

question	num_users	% of users
1. What are you looking for?	500	25.2%
2. What's your fit?	475	23.9%
3. Which shapes do you like?	380	19.1%
4. Which colors do you like?	361	18.2%
5. When was your last eye exam?	270	13.6%

1.4 Home Try-On Funnel, contents

Below are snippets from each of the tables I'll use to uncover insights about the Home Try-On funnel.

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

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

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

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

1.5 Home Try-On Funnel, contents

Below is a new table to determine if a user received a home try-on, and whether s/he purchased.

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

1.6 Home Try-On Funnel, additional analysis

75% of quiz-takers ordered a home try-on kit. Of the users who got a home try on kit, only 66% made a purchase.

On its own, this data doesn't reveal too much in terms of which type of home try-on kit performed best (3 pairs vs 5 pairs).

um_users	num_home_try_on	num_purchase	prct_home_try_on	prct_purchase
1000	750	495	0.75	0.66

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

SELECT COUNT(*) AS 'num_users',
  SUM(is_home_try_on) AS 'num_home_try_on',
  SUM(is_purchase) AS 'num_purchase',
  (1.0 * SUM(is_home_try_on) / COUNT(user_id)) AS
'prct_home_try_on',
  (1.0 * SUM(is_purchase) / SUM(is_home_try_on)) AS
'prct_purchase'
FROM funnel;
```

1.6 Home Try-On Funnel, additional analysis

To get a better idea of which home try-on kit performs better, I've split out the data.

When comparing users who received 3 pairs of glasses vs users who received 5 pairs of glasses, the purchase rates are significantly different.

As it turns out, users who received 5 pairs of glasses completed more purchases (79% purchase rate).

number_of_pairs	num_users	num_home_try_on	num_purchase	prct_home_try_on	prct_purchase
NULL	250	0	0	0.0	NULL
3 pairs	379	379	201	1.0	0.530343007915567
5 pairs	371	371	294	1.0	0.792452830188679

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

SELECT COUNT(*) AS 'num_users',
  SUM(is_home_try_on) AS 'num_home_try_on',
  SUM(is_purchase) AS 'num_purchase',
  (1.0 * SUM(is_home_try_on) / COUNT(user_id)) AS
'prct_home_try_on',
  (1.0 * SUM(is_purchase) / SUM(is_home_try_on)) AS
'prct_purchase'
FROM funnel
GROUP BY number_of_pairs;
```

1.7 Final Conclusions

Based on the data uncovered during my analysis, I have two recommendations for Warby Parker.

Survey Structure
- Consider simplifying at least the final question on the survey.
- For example, "Have you had an eye exam within the past year?" could elicit more responses as it would be easier for a user to recall.

Home Try-On Kits
- Discontinue the home try-on kits with 3 pairs of glasses, as those kits resulted in a significantly lower conversion rate.

question	num_users	% of users
1. What are you looking for?	500	25.2%
2. What's your fit?	475	23.9%
3. Which shapes do you like?	380	19.1%
4. Which colors do you like?	361	18.2%
5. When was your last eye exam?	270	13.6%

number_of_pairs	num_users	num_home_try_on	num_purchase	prct_home_try_on	prct_purchase
NULL	250	0	0	0.0	NULL
3 pairs	379	379	201	1.0	0.530343007915567
5 pairs	371	371	294	1.0	0.792452830188679