Learn SQL From Scratch Capstone Project: Usage Funnels with Warby Parker

Mohammed Mahfuz 7/24/2018

Warby Parker

Warby Parker is a transformative lifestyle brand with a lofty objective: to offer designer eyewear at a revolutionary price while leading the way for socially conscious businesses. Founded in 2010 and named after two characters in an early Jack Kerouac journal, Warby Parker believes in creative thinking, smart design, and doing good in the world. For every pair of eyeglasses and sunglasses sold, a pair is distributed to someone in need.

Question 1

To help users find their perfect frame, Warby Parker has a Style Quiz that has the following questions:

1. "What are you looking for?"
2. "What's your fit?"
3. "Which shapes do you like?"
4. "Which colors do you like?"
5. "When was your last eye exam?"

The users' responses are stored in a table called `survey`.

Select all columns from the first 10 rows. What columns does the table have?

```
Select *
From survey
Limit 10;
```

ANSWER:

The columns are ‘question’, ‘user_id’ and ‘response’

Question 1 Query Results

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

Question 2

Users will "give up" at different points in the survey. Let's analyze how many users move from Question 1 to Question 2, etc.

Create a quiz funnel using the `GROUP BY` command.

What is the number of responses for each question?

```
Select question, COUNT(DISTINCT user_id) AS 'Distinct Users'
From survey
Group by 1;
```

ANSWER:

Q1: 500

Q2: 475

Q3: 380

Q4: 361

Q5: 270

Question 2 Query Results

Query Results	
question	Distinct 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

Question 3

Using a spreadsheet program like Excel or Google Sheets, calculate the percentage of users who answer each question.:

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

What do you think is the reason?

Add this finding to your presentation slides!

question	Distinct Users	% completion
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%

ANSWER:

Q1: 100%

Q2: 475/500 = 95%

Q3: 380/475 = 80%

Q4: 361/380 = 95%

Q5: 270/361 = 74.79%

Q5 and Q3 have a lower completion rate. I believe Q5 has the lowest completion rate because people may not remember when they had their last eye exam and don't really want to do any work to figure out when it was. Also, it doesn't necessarily have anything to do with the style questions so it may throw them off guard.

Question 4

Warby Parker's purchase funnel is:

Take the Style Quiz → Home Try-On → Purchase the Perfect Pair of Glasses

During the Home Try-On stage, we will be conducting an A/B Test:

- 50% of the users will get **3**pairs to try on
- 50% of the users will get **5**pairs to try on

Let's find out whether or not users who get more pairs to try on at home will be more likely to make a purchase.

The data will be distributed across three tables:

- `quiz`
- `home_try_on`
- `purchase`

Examine the first five rows of each table

What are the column names?

ANSWER:

```
Select *
From quiz
Limit 5;

Select *
From home_try_on
Limit 5;

Select *
From purchase
Limit 5;
```

ANSWER:

Quiz: user_id, style, fit, shape, color

Home_try_on: user_id, number_of_pairs, address

Purchase: user_id, product_id, style, model_name, color, price

Question 4 Query Results

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
0a9dd17-36c8-430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black	150
0e15fe0-c86f-4818-9c63-3422211baa97	7	Women's Styles	Lucy	Elderflower Crystal	150
17506f7-aba1-4b9d-8b7b-f4426e71b8ca	4	Men's Styles	Dawes	Jet Black	150
176bfb3-9c51-4b1c-b593-87edab3c54cb	10	Women's Styles	Eugene Narrow	Rosewood Tortoise	95
01fdf106-f73c-4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black	150

Question 5

We'd like to create a new table with the following layout:

Each row will represent a single user from the browse table:

- If the user has any entries in `home_try_on`, then `is_home_try_on` will be 'True'.
- `number_of_pairs` comes from `home_try_on` table
- If the user has any entries in `is_purchase`, then `is_purchase` will be 'True'.

Use a `LEFT JOIN` to combine the three tables, starting with the top of the funnel (browse) and ending with the bottom of the funnel (purchase).

Select only the first 10 rows from this table (otherwise, the query will run really slowly).

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc	True	3	False
291f1ca	True	5	False
75122300	False	NULL	False

```
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
Limit 10;
```

Question 6

Once we have the data in this format, we can analyze it in several ways:

- We can calculate overall conversion rates by aggregating across all rows.
- We can compare conversion from `quiz`→`home_try_on` and `home_try_on`→`purchase`.
- We can calculate the difference in purchase rates between customers who had 3 `number_of_pairs` with ones who had 5.
- And more!

We can also use the original tables to calculate things like:

- The most common results of the style `quiz`.
- The most common types of `purchase` made.
- And more!

What are some actionable insights for Warby Parker?

ANSWER:

Warby Parker can change Question 5 in their Style Quiz since it is near the end of the survey and they only have ~75% completion rate which means they lose 25% of customers who were a single question away from getting an answer that could have led them to purchase a pair of glasses.

Thank you for a wonderful course!