Warby Parker

David Petrushka – Learn SQL from Scratch Capstone

Question 1

- This SQL query brings up every column in the Survey Table
- Limiting the query to the first 10 lines

Question 2

- Here we create the first funnel in our data – Segregating the number of individual Distinct User IDs who answer each question

- By doing this, we can establish user behavior by seeing where users stop responding to our survey

- Answer:
  - 1. 500
  - 2. 475
  - 3. 380
  - 4. 361
  - 5. 270

```
project.sqlite

1  SELECT question,
2      COUNT(DISTINCT user_id)
3  FROM survey
4  GROUP BY question;
```

Query Results

question	COUNT(DISTINCT user_id)
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

Question	#responded	% Completion Rate
1. What are you looking for?	500	100.0%
2. What's your fit?	475	95.0%
3. Which shapes do you like?	380	80.0%
4. Which colors do you like?	361	95.0%
5. When was your last eye exam?	270	74.8%

The question with the lowest Completion Rate is **Question 5**

Rationale:

1. This asks for a concrete answer (What is the exact date) – people may not know the exact date of their last eye exam and may not be willing to go find that answer

2. This was the last question on the quiz – after 5 Questions people have a lower willingness to comply with quiz directions

Question 4

- This question asks us to define the column names for 3 tables
  - **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

```
project.sqlite

1   SELECT *
2   FROM quiz
3   LIMIT 5;
4
5   SELECT *
6   FROM home_try_on
7   LIMIT 5;
8
9   SELECT *
10  FROM purchase
11  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

Question 5

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

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	0	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	0	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9	0	0	0
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	1	5 pairs	0
b1dded76-cd60-4222-82cb-f6d464104298	1	3 pairs	0

Database Schema

project.sqlite

Questions 6

The data output from the SQL Query provides a funnel that allows us to see the point where people either opt out of the purchase process or the factors that led them to make a purchase

A few Actionable Insights that could be derived from this data:

A/B Testing on purchase rate from Number of Pairs

Likelihood of customers making a second purchase

Purchase rates based on Models used