SQL Capstone Project

Course: Learn SQL from scratch

Project option: Usage Funnels with Warby Parker

Quiz Funnel: Question 1

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

- I have selected all columns with "*"
- From the table called "survey"
- And limited to 10 rows as I don't need more information and it reduces the loading time.

Answer:
1. question
2. user_id
3. response

```
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. 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

Quiz Funnel: Question 2

Question:
What is the number of responses for each question?

- I have selected 2 columns with "question" and "count(distinct user_id)" which will count the user_id for each question and will rename the column to "Total".
- From the table called "survey"
- And grouped the questions to see the value for each question.

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

```
SELECT question,
 COUNT(DISTINCT user_id) as 'Total'
 FROM survey
 GROUP BY question;
```

Question	Total
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

Quiz Funnel: Question 3

Question 1:
Which question(s) of the quiz have a lower completion rates?

Answer 1:
Question 5 and Question 3 has the 2 lowest completion rates which you can see in the excel table below.
Question 3 with a 80% completion rate.
Question 5 with a 75,06% completion rate.

Question 2:
What do you think is the reason?

Answer 2:
- People lose interest the more questions comes up.
- People don't expect multiple questions, they might answer the 1st question as they think they will get instant results.
- Question 5 is quite personal and people do not want to feel embarrassed by answering a quiz.

Question	Total	Calculation	Answer in %
1. What are you looking for?	500	500/500 * 100	100
2. What's your fit?	475	475/500 * 100	95
3. Which shapes do you like?	380	380/475 * 100	80
4. Which colors do you like?	361	361/380 * 100	95
5. When was your last eye exam?	271	271/361 * 100	75,06925208

Home Try-On Funnel: Question 4 Part 1

Question:
What are the column names?

- I have selected all columns with "*"
- From the table called "quiz"
- And limited to 5 rows as I don't need more information and it reduces the loading time.

Answer:
1. user_id
2. Style
3. fit
4. shape
5. color

```
SELECT *
 FROM quiz
 LIMIT 5;
```

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

Home Try-On Funnel: Question 4 Part 2

Question:
What are the column names?

- I have selected all columns with "*"
- From the table called "home_try_on"
- And limited to 5 rows as I don't need more information and it reduces the loading time.

Answer:
1. user_id
2. Number_of_pairs
3. address

```
SELECT *
 FROM home_try_on
 LIMIT 5;
```

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

Home Try-On Funnel: Question 4 Part 3

Question:
What are the column names?

- I have selected all columns with "*"
- From the table called "purchase"
- And limited to 5 rows as I don't need more information and it reduces the loading time.

Answer:
1. user_id
2. Product_id
3. Style
4. Model_name
5. Color
6. price

```
SELECT *
 FROM purchase
 LIMIT 5;
```

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

Home Try-On Funnel: Question 5

As indicated, I need 4 columns. User_id, is_home_try_on, number_of_pairs and is_purchase.

1. Selected Column "User" as distinct to give me unique values in the "User" Column.
2. Created 3 case statements to get the specific information requested
3. Is_home_try_on will be true if there is NULL value else fals
4. Number_of_pairs will take away the word pairs and only display the value
5. Is_purchase will be true if there is NULL value else false
6. Ended the case statements
7. Selected from the "quiz" table
8. User a left join to join the tables quiz and home_try_on on the user_id column
9. User a left join to join the tables purchase and quiz on the user_id column
10. Used the limit function as I only want 10 rows of information and I want to reduce loading time.

```sql
SELECT
  DISTINCT quiz.user_id as 'User',
    CASE
      WHEN home_try_on.user_id IS NOT NULL THEN 'True'
      ELSE 'False'
    END  as 'is_home_try_on',
    CASE
      WHEN home_try_on.number_of_pairs = '3 pairs' THEN '3'
      WHEN home_try_on.number_of_pairs = '5 pairs' THEN '5'
      ELSE 'NULL'
    END  as 'number_of_pairs',
    CASE
      WHEN purchase.user_id IS NOT NULL THEN 'True'
      ELSE 'False'
    END  as 'is_purchase'
FROM quiz
LEFT JOIN home_try_on
  ON quiz.user_id = home_try_on.user_id
LEFT JOIN purchase
  ON purchase.user_id = quiz.user_id
LIMIT 10;
```

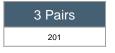| User	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	True	3	False
291f1cca-e507-48be-b063-002b14906468	True	3	True
75122300-0736-4087-b6d8-c0c5373a1a04	False	NULL	False
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	True	5	False
ce965c4d-7a2b-4db6-9847-601747fa7812	True	3	True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

Home Try-On Funnel: Question 6
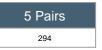
Question:

What are some actionable insights for Warby Parker?

Answer:

1. 20.1% of users that tried 3 pairs, bought a pair.
2. 29.4% of users that tried 3 pairs, bought a pair.
3. 49.5% of users that tried glasses on at home bought the product.
4. 1000 user

3 Pairs
201

5 Pairs
294

Home try on
495

Total
1000

```
1. SELECT
     count(DISTINCT quiz.user_id)
   FROM quiz
   LEFT JOIN home_try_on
     ON quiz.user_id = home_try_on.user_id
   LEFT JOIN purchase
     ON purchase.user_id = quiz.user_id
   WHERE purchase.user_id is not null
   AND number_of_pairs = '3 pairs';

2. SELECT
     count(DISTINCT quiz.user_id)
   FROM quiz
   LEFT JOIN home_try_on
     ON quiz.user_id = home_try_on.user_id
   LEFT JOIN purchase
     ON purchase.user_id = quiz.user_id
   WHERE purchase.user_id is not null
   AND number_of_pairs = '5 pairs';

3. SELECT
     count(DISTINCT quiz.user_id)
   FROM quiz
   LEFT JOIN home_try_on
     ON quiz.user_id = home_try_on.user_id
   LEFT JOIN purchase
     ON purchase.user_id = quiz.user_id
   WHERE home_try_on.user_id is not null
   AND purchase.user_id is not NULL;

4. SELECT
     count(DISTINCT quiz.user_id)
   FROM quiz;
```