SQL FROM SCRATCH

CAPSTONE PROJECT- LARA CHAVES

QUIZ FUNNEL

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

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

```
SELECT *
FROM survey
LIMIT 10
;

/*
1. The table has the following
columns:
- question
- user_id
- response
*/
```

QUIZ FUNNEL

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?

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

/*
2. The number of responses for
each question is:
- Question 1: 500
- Question 2: 475
- Question 3: 380
- Question 4: 361
- Question 5: 270
*/
```

QUIZ FUNNEL

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!

```
/*
3. Completion Rates:
- Question 1: 100%
- Question 2: 95%
- Question 3: 82%
- Question 4: 95%
- Question 5: 74%
The questions with lower completion
rates are 'Which shapes do you like?'
and 'When was your last eye exam?'
For the first one I think that people
are not sure about what they want
and for the second one is possible
that they are not sure about the
time without checking.
*/
```

HOME TRY-ON FUNNEL

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.

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

HOME TRY-ON FUNNEL

4. Examine the first five rows of each table.

What are the column names?

```
/*
4. The quiz table has the following
columns:
- user_id
- style
- fit
- shape
- color

The home_try_on table has the
following columns:
- user_id
- number_of_pairs
- address

The purchase table has the
following columns:
- user_id
- product_id
- style
- model_name
- color
- price
*/
```

HOME TRY-ON FUNNEL

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_onwill 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).

```
SELECT DISTINCT quiz.user_id,
  home_try_on.user_id IS NOT NULL AS
  'is_home_try_on',
  home_try_on.number_of_pairs,
  purchase.user_id IS NOT NULL 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 =
  home_try_on.user_id
LIMIT 10
;

/*
5. Seven of the ten Users tried items
at home and three of the ten Users
made a purchase
*/
```

HOME TRY-ON FUNNEL

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

What are some actionable insights for Warby Parker?

```
WITH funnels AS
(
  SELECT DISTINCT quiz.user_id,
    home_try_on.user_id IS NOT NULL AS
    'is_home_try_on',
    home_try_on.number_of_pairs,
    purchase.user_id IS NOT NULL 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 =
    home_try_on.user_id
)
SELECT COUNT (*) AS 'num_browse',
  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
  'browse_to_home_try_on',
  1.0 * SUM(is_purchase) /
  SUM(is_home_try_on) AS
  'home_try_on_to_purchase'
FROM funnels
;
```