Capstone: Funnels with Warby Parker

Learn SQL from Scratch
Sam Gross
16th June 2018

Table of Contents

Quiz Funnel

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

We can find out the columns in the survey table by using a simple `SELECT *`

(Whilst the code uses `LIMIT 10`, only five rows are included in the example table below for the sake of space)

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.

```
SELECT *
FROM survey
LIMIT 10;
```

2. What is the number of responses for each question?

We can summarise the number of responses to each `question` by counting the total amount of unique `user_id` values that appear beside them.

question	COUNT (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

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

3. Which question(s) of the quiz have a lower completion rates? What do you think is the reason?

If 500 users account for 100% of participants, then we see the completion rate drop to 95%, 76%, 72.2%, and finally 54% of this total. This contraction of user numbers is typical for a usage funnel, though we might infer that users are confident of the broad style (Q1) and fit (Q2) of their eyewear, but are less able to decide on specific traits (Q3, Q4) and the date of their last eye exam (Q5).

question	COUNT (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

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

Home-Try-On Funnel

4. Examine the first five rows of each table. What are the column names?

The columns of each table are as follows, with one example row to show the contents. Each query is kept separate using ;

user_id	style	fit	shape	color
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	Women's Styles	Medium	Rectangular	Tortoise

user_id	number_of_pairs	address
d8addd87-3217-4429-9a01-d56d68111da7	5 pairs	145 New York 9a

user_id	product_id	style	model_name	color	price
00a9dd17-36c8-430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black	150

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

5. Use a LEFT JOIN to combine the three tables… Select only the first 10 rows from this table

funnel is created using the `quiz`, `home_try_on` and `purchase` tables. We ascertain if users from `quiz` have tried or purchased items by checking whether their `user_id` appears in the subsequent tables.

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

```
WITH funnel AS (
  SELECT quiz.user_id,
 CASE
   WHEN home_try_on.user_id IS NOT NULL
      THEN 'True'
      ELSE 'False'
   END AS 'is_home_try_on',
 home_try_on.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 quiz.user_id = purchase.user_id)

SELECT *
FROM funnel
LIMIT 10;
```

Actionable Insights

Conversion Rates: Quiz, Try-On, and Purchase

We can quickly count the instances of `True` and `False` results for the `is_home_try_on` and `is_purchase` stages of the funnel. From a total of 1000 users, 750 tried glasses at home, and 495 made an actual purchase. The drop off between quiz takers and users who go to the trouble of trying glasses is understandable, but these figures indicate that 34% of customers are receiving multiple pairs of glasses, trying them, and sending them back without buying anything.
Perhaps something needs to be done to improve the site's recommendations?

is_home_try_on	COUNT (*)
False	250
True	750

```
SELECT is_home_try_on, COUNT (*)
FROM funnel
GROUP BY is_home_try_on;
```

is_purchase	COUNT (*)
False	505
True	495

```
SELECT is_purchase, COUNT (*)
FROM funnel
GROUP BY is_purchase;
```

Split Test: More Pairs, More Purchases

With the following query applied to our funnel table, we can count the total number of users that ultimately purchased glasses, and separate them into two groups based on which half of the split test they were in.

Of the 379 customers sent **3 pairs**, 201 (**50%**) went on to buy.
Of the 371 customers sent **5 pairs**, 294 (**79.25%**) went on to buy.

From these two figures alone, we can see that customers who are sent **5 pairs** of glasses are almost **30% more likely** to make a purchase.
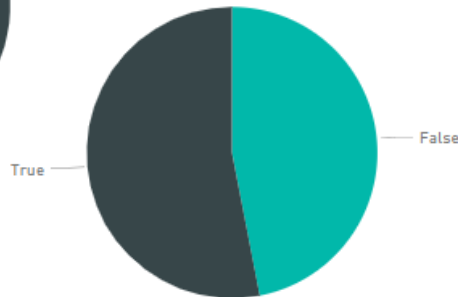
```
SELECT number_of_pairs, COUNT(number_of_pairs),
COUNT(is_purchase)
FROM funnel
WHERE is_purchase = 'True'
GROUP BY number_of_pairs;
```
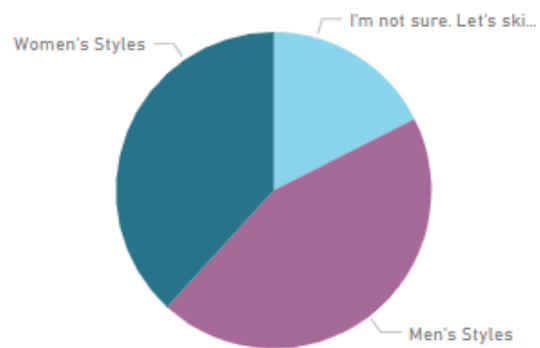

5 Pairs

3 Pairs

number_of_pairs	COUNT (is_purchase)
3 pairs	201
5 pairs	294
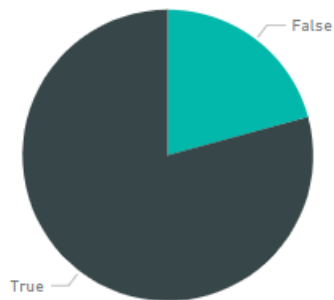
The Survey Table

The original `survey` table can be separated by `question` and `response`, and then a `COUNT` can be taken of how many users selected which options. The results (visualisations overleaf) show a variety of thing; an even gender split, demand for a narrow fit, colour preferences, and lack of recent eye examinations in customers. (Q2 and Q3 including below as examples)
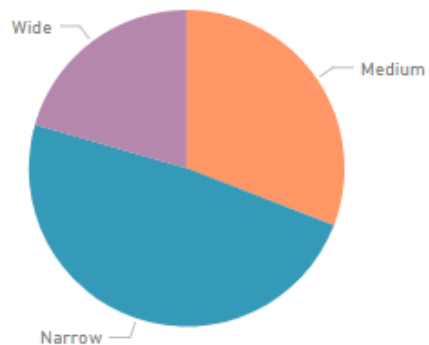
question	response	COUNT (user_id)
2. What's your fit?	Medium	132
2. What's your fit?	Narrow	208
2. What's your fit?	Wide	88
3. Which shapes do you like?	No Preference	29
3. Which shapes do you like?	Rectangular	141
3. Which shapes do you like?	Round	91
3. Which shapes do you like?	Square	119

```
SELECT question, response, COUNT (user_id)
FROM survey
GROUP BY response
ORDER BY question;
```
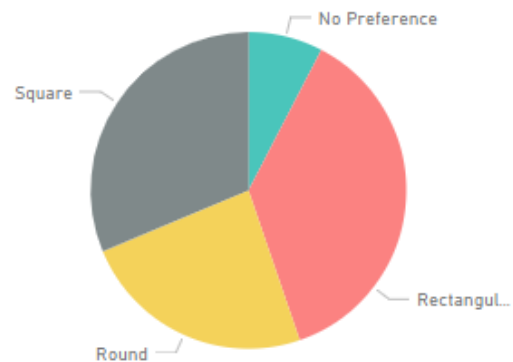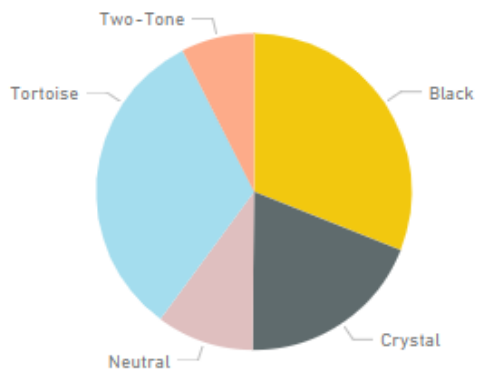
Looking For?

- I'm not sure. Let's ski...
- Women's Styles
- Men's Styles

FIT

- Wide
- Medium
- Narrow

SHAPE

- No Preference
- Square
- Rectangul...
- Round

COLOR

- Two-Tone
- Tortoise
- Black
- Crystal
- Neutral

EYE EXAM

- Not Sure. Let's Skip It
- 3+ Years
- 1-3 Years
- <1 Year
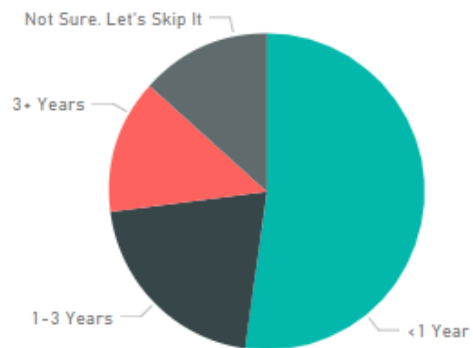
Purchases

Every single item bought is recorded in the `purchase` table. We can see every product's performance; glasses priced at $95 are more likely to sell. Both Eugene Narrow models are in the top three most purchased. Monocles sell worst, despite being the cheapest item by a large margin, and the only 'Tortoise' color not to make the top five.

```
SELECT style, model_name, color, price, COUNT
(user_id) AS bought
FROM purchase
GROUP BY product_id
ORDER BY bought DESC;
```

style	model_name	color	price	bought
Men's Styles	Dawes	Driftwood Fade	150	63
Women's Styles	Eugene Narrow	Rosewood Tortoise	95	62
Women's Styles	Eugene Narrow	Rose Crystal	95	54
Men's Styles	Brady	Layered Tortoise Matte	95	52
Women's Styles	Olive	Pearled Tortoise	95	50
Men's Styles	Dawes	Jet Black	150	44
Women's Styles	Lucy	Elderflower Crystal	150	44
Men's Styles	Brady	Sea Glass Gray	95	43
Women's Styles	Lucy	Jet Black	150	42
Men's Styles	Monocle	Endangered Tortoise	50	41