codecademy

Warby Parker Usage Funnels

Learn SQL from Scratch: Capstone

Zach Rachuba

Dec. 20, 2018

The Quiz Funnel

1. Getting Familiar

What columns does the table have?

- The table survey has the columns question, user_id, and response.
- *Only the first four rows are shown for visibility.*

```
--QUERY:—

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

2. Survey Funnel: Response per Question

What is the number of responses for each question?

- 500 for the first, 475 for the second, 380 for the third, 361 for the fourth, and 270 for the fifth.

```
-- QUERY:--

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

COUNT(DISTINCT user_id)	question
500	1. What are you looking for?
475	2. What's your fit?
380	3. Which shapes do you like?
361	4. Which colors do you like?
270	5. When was your last eye exam?

3. Survey Funnel: Conversion

Which question(s) of the quiz has/have a lower completion rate(s)? What do you think is the reason?

The completion rate table is shown to the right, calculated using the query results from the previous slide: "% Completing" is the count of users finishing each question divided by the same for the previous question.

- The third and fifth questions have lower completion rates: 80% and 74.8%, respectively.
- The third's lower response may be due to responders' unfamiliarity with choosing an eyewear shape they like (they may have never done so/not know the different shapes).
- The fifth question may seem intrusive/an invasion of privacy, or users may be uncomfortable admitting they have never had an eye exam/not had one for a long time.

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

The Home Try-On Funnel

4. Survey Funnel: Response per Question

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

- The columns from quiz are user_id, style, fit, shape, and color; those from home_try_on are user_id, number_of_pairs, and address. The table purchase has columns user_id, product_id, style, model_name, color, and price. [Results are trimmed for space.]

```
-- QUERY:--
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

quiz

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

home_try_on

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

purchase

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

5. Home Try-On Funnel: Table Join

Create a table with the columns user_id, number_of_pairs, is_home_try_on, is_purchase. Select only the first 10 rows from this table (otherwise, the query will run really slowly).

- *Below results are trimmed for visibility.*

```
--QUERY:--
SELECT DISTINCT
             q.user_id,
             h.number_of_pairs 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;
```

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

6.1 Hit Rate by Number of Pairs

What are some actionable insights for Warby Parker?

- We can view the three-versus-five pair distinction for the campaign participants as an A/B test: two parallel campaigns with one factor (number of pairs) differing to compare which is more effective. With the results from the previous query (a join of the relevant funnel tables) as a temporary table, total purchases and total try-ons by number of pairs can be summed using COUNT, SUM, and GROUP BY. The proportions of each that made a purchase can then be calculated.
- To make the results clearer, non-purchases have been excluded with a HAVING clause, and hit_rate has been rounded to two decimal places.
- With similar sample sizes (371 to 379), five pairs has a much higher purchase rate of 79% compared to 53% for three pairs. Other things being equal, five pairs may be a better number to have users try on.

purchases	number_of_pairs	tryons	hit_rate
201	3 pairs	379	0.53
294	5 pairs	371	0.79

```
--QUERY:--
WITH funnel AS (
  SELECT DISTINCT
              q.user_id,
              h.number_of_pairs 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)
SELECT SUM(is_purchase) AS purchases, number_of_pairs,
COUNT(user_id) AS tryons, ROUND(1.0 * SUM(is_purchase)
/ COUNT(user_id), 2) AS hit_rate
FROM funnel
GROUP BY 2
HAVING SUM(is_purchase) > 0;
```

6.2 Revenue Stream by Product

What are some actionable insights for Warby Parker?

- The top three models (Dawes, Lucy, Eugene Narrow) make up a large portion of revenue and total purchases. These popular models should be prioritized for deals/specials.
- The revenue gap is large in men's styles: the Dawes made $16,060 from 107 hits compared to $9025 from the 95 Brady hits. The Brady purchasers may decide against buying the Dawes because of the higher price point, but Warby's could experiment with a price for the Dawes lower than $150 but more than $95—it's possible that enough buyers would switch from the Brady to the Dawes to more than offset the revenue loss from a lower price.
- The revenue gaps between different colors of the same style are surprisingly low, especially for women, which could suggest that demand exists for many different colors: almost the same amount bought the different colors for the Lucy, and the difference is only 8 for the Eugene Narow. Warby's should experiment with additional colors to get purchases from those who did not have their preferred color available.

```
--QUERY:--
SELECT COUNT(user_id) AS hits, model_name, 1.0 * price *
COUNT(user_id) AS revenue, price, style, color
FROM purchase
GROUP BY 2, 6
ORDER BY 3 DESC;
```

hits	model_name	revenue	price	style	color
63	Dawes	9450.0	150	Men's Styles	Driftwood Fade
44	Dawes	6600.0	150	Men's Styles	Jet Black
44	Lucy	6600.0	150	Women's Styles	Elderflower Crystal
42	Lucy	6300.0	150	Women's Styles	Jet Black
62	Eugene Narrow	5890.0	95	Women's Styles	Rosewood Tortoise
54	Eugene Narrow	5130.0	95	Women's Styles	Rose Crystal
52	Brady	4940.0	95	Men's Styles	Layered Tortoise Matte
50	Olive	4750.0	95	Women's Styles	Pearled Tortoise
43	Brady	4085.0	95	Men's Styles	Sea Glass Gray
41	Monocle	2050.0	50	Men's Styles	Endangered Tortoise

6.3 Purchases by Purchaser

What are some actionable insights for Warby Parker?

- With the purchases table, we can see how many items each purchaser bought.
- All buyers made only one purchase! Warby could experiment with deals (like buy one full price, get a discount on additional purchases) to change this, and sell more products at once.

```
--QUERY:--
SELECT user_id, COUNT(DISTINCT product_id)
FROM purchase
GROUP BY 1
ORDER BY 1 DESC
LIMIT 5;
```

user_id	COUNT(DISTINCT product_id)
ffef9d8c-7a8d-475d-a9b2-856b81c05b66	1
fee44777-bbd7-44ea-8102-291c89d0519b	1
fdf35090-5080-4d90-9204-7af788268837	1
fdd0cc7d-b9fc-404e-9d95-f69b230d2e72	1
fdb36b83-f7e5-475a-9165-336fc3a39ccc	1