codecademy

Capstone project: Funnels with Warby Parker

Learn SQL from Scratch submission

Neil Ballantyne

DATE

1. Exercise submissions

Exercise 1

Viewing columns of the *survey* table:

- question
- user_id
- response

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

Exercise 2

Showing the number of responses for each questions in a survey.

```
SELECT question,
 COUNT(DISTINCT user_id) AS 'number_answers'
 FROM survey
 GROUP BY question
;
```

question	number_answers
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

Exercise 3

Showing the steps as a funnel with percentage of users who'd completed previous step.

Final step has highest drop off. Hypothesis for this being that this step requires recall of a specific piece of information not likely to be immediately available.

question	Percentage complete
1. What are you looking for?	100.0%
2. What's your fit?	95.0%
3. Which shapes do you like?	80.0%
4. Which colors do you like?	95.0%
5. When was your last eye exam?	74.8%

Spreadsheet at https://docs.google.com/spreadsheets/d/1OqTc_VAjef8NtrsGmcE9RBaUqw2iTKrctfhBeaD44z4/edit?usp=sharing

Exercise 4

Viewing columns of the *quiz, home_try_on* and *purchase* 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

```
SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

Exercise 5

Joined table showing results from three combined tables (showing only 8 here for brevity)

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	NULL	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	NULL	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9	0	NULL	0

```
SELECT
DISTINCT q.user_id,
h.user_id 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
;
```

Exercise 6 – Overall conversion rate

Overall conversion rate from quiz users to conversion, regardless of experiment, is 49.5%

num_quiz	num_purchase	conversion_%
1000	495	49.5

```
 WITH funnel AS
 (SELECT DISTINCT q.user_id,
   h.user_id 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
COUNT(user_id) As 'num_quiz',
SUM(is_purchase) AS 'num_purchase',
100 * (1.0 * SUM(is_purchase) /
COUNT(user_id)) AS 'conversion_%'
FROM funnel
;
```

Exercise 6 – Conversion rate at each step of funnel

There is a 75% conversion from taking the quiz to trying some glasses on and then 66% from trying on to purchase.

As seen, from completing the quiz there is a 49.5% total conversion rate

```
…

SELECT
COUNT(user_id) AS 'num_quiz',
SUM(is_home_try_on) AS 'num_home_try_on',
SUM(is_purchase) AS 'num_purchase',
ROUND (100 * (1.0 * SUM(is_home_try_on) /
COUNT(user_id)),1) AS 'quiz_to_try_on',
ROUND (100 * (1.0 * SUM(is_purchase) /
SUM(is_home_try_on)),1) AS 'try_on_to_purchase',
ROUND (100 * (1.0 * SUM(is_purchase) /
COUNT(user_id)),1) AS 'overall_conversion'
FROM funnel
;
```

num_quiz	num_home_try_on	num_purchase	quiz_to_try_on	try_on_to_purchase	Overall_conversion
1000	750	495	75.0	66.0	49.5

Exercise 6 – Conversion rate by variant

Despite having a small decrease in the number of home-try-on orders, there is a significantly better conversion in the ‘5 pairs’ variant with a 49.4% increase (or 26.2 percentage points).

The company should consider rolling out this variant. The hypothesis would be that having more choice increase the chances of the customer finding something they like.

Variant	num_home_try_on	num_purchase	conversion_%
3 pairs	379	201	53.0
5 pairs	371	294	79.2

```
…
SELECT
number_of_pairs AS 'Variant',
SUM(is_home_try_on) AS 'num_home_try_on',
SUM(is_purchase) AS 'num_purchase',
ROUND (100 * (1.0 * SUM(is_purchase) /
SUM(is_home_try_on)),1) AS 'conversion_%'
FROM funnel
    WHERE number_of_pairs IS NOT NULL
    GROUP BY 1
;
```

Exercise 6 – Further insights from funnel

Comparing conversion of men's versus women's styles.

Men's styles convert better than women's so the company may want to experiment with targeting male customers more.

style	num_home_try_on	num_purchase	conversion_%
Men's Styles	320	243	75.9
Women's Styles	361	252	69.8
I'm not sure. Let's skip it.	69	0	0.0

```
WITH funnel AS
 (SELECT DISTINCT q.user_id,
   h.user_id 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
style,
SUM(is_home_try_on) AS 'num_home_try_on',
SUM(is_purchase) AS 'num_purchase',
ROUND (100 * (1.0 * SUM(is_purchase) /
SUM(is_home_try_on)),1) AS 'conversion_%'
FROM quiz q
LEFT JOIN funnel f
on q.user_id = f.user_id
GROUP BY 1
ORDER BY 4 DESC
 ;
```

Exercise 6 – Further insights from funnel

The same query can be used to look at conversions for the other aspects – fit, colour, shape

Insights from this:
People who selected the '*I'm not sure*' option for fit and shape ends up converting best. (70.3% and 74.6% respectively)

The company may want to consider whether this is a necessary question – it may just add friction to the process by getting people to state a preference in a decision that may be easily changed.

Black is the best converting colour with 68.2% conversion. Neutral converts the least (60.8%).

```
WITH funnel AS
 (SELECT DISTINCT q.user_id,
   h.user_id 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
--switch out fit/shape/style/color
fit,
--color,
--style,
--shape,
SUM(is_home_try_on) AS 'num_home_try_on',
SUM(is_purchase) AS 'num_purchase',
ROUND (100 * (1.0 * SUM(is_purchase) /
SUM(is_home_try_on)),1) AS 'conversion_%'
FROM quiz q
LEFT JOIN funnel f
on q.user_id = f.user_id
GROUP BY 1
ORDER BY 4 DESC
 ;
```

Exercise 6 – Purchase insights

Eugene Narrow is the best selling model. The best selling men's style is Dawes.

style	model_name	purchases
Women's Styles	Eugene Narrow	116
Men's Styles	Dawes	107
Men's Styles	Brady	95
Women's Styles	Lucy	86
Women's Styles	Olive	50
Men's Styles	Monocle	41

```
SELECT
q.style,
p.model_name,
COUNT(p.user_id) AS 'purchases'
FROM
quiz q
LEFT JOIN purchase p
ON p.user_id = q.user_id
GROUP BY 2
HAVING p.model_name IS NOT NULL
ORDER BY 3 DESC
;
```