codecademy

Capstone: Warby Parker

Learn SQL from Scratch

Brad Lokitz

January 2, 2019

Table of Contents

1. Warby Parker – A transformative lifestyle brand!

1.1 Warby Parker: A Transformative Lifestyle Brand

- Warby Parker is a transformative lifestyle brand with a lofty objective: to offer designer eyewear at a revolutionary price while leading the way for socially conscious businesses. Founded in 2010 and named after two characters in an early Jack Kerouac journal, Warby Parker believes in creative thinking, smart design, and doing good in the world. For every pair of eyeglasses and sunglasses sold, a pair is distributed to someone in need.

- Find the perfect frame – the Warby Parker website directs user to a style quiz with 5 questions to help find the perfect frame:
    1. What are you looking for?
    2. What's your fit?
    3. Which shapes do you like?
    4. Which colors do you like?
    5. When was your last eye exam?

- Once the user has found the perfect frame they are given the opportunity to try the frames at home for free. Warby Parker has designed an A/B test where 50% of the users receive 3 pairs of glasses to try on at home and 50% receive 5 pairs of glasses and wants to know the conversion rates for each.

2. Quiz Funnel – find your perfect frame!

2.1 Style quiz – find your perfect frame !

- The style quiz is used to help users find perfect frame. 5 Questions are asked 1) What are you looking for?; 2) What's your fit?; 3)Which shapes do you like?; 4) Which colors do you like?; 5)When was your last eye exam?
- The answers are stored in an sql table (see example table below) with 3 columns: question, user_id, and response. The sql code shown in the grey box is used to pull up the table. The query is limited to 10 rows.

```
-- You can put your query here

SELECT *
FROM survey
limit 10;
```

question	user_id	response
1. What are you looking for?	005e7f99-d48c-4fce-b605-10506c85aaf7	Women's Styles

2.2 Style quiz response rate

- Users will "give up" at different points in the survey. An analysis of how many users move from Question 1 to Question 2, etc.
- You can use sql to count the number of users that answer each question by using the count command and only counting distinct user_id from the survey table and then group the results by question. The query is given in the grey box.

    - 500 people took the quiz and answered question 1;
    - 95% (475 people) of the people go on to answer question 2;
    - 76% (380 people) of the people answer question 3;
    - 72% (361 people) of the people answer question 4;
    - 54% (270 people) of the people answer question 5;

- The numbers show that fewer people answer the questions about the shapes (question 3) and colors (question 4) they like but the question with the lowest response rate is question 5, When was your last eye exam?". Questions 3 (shapes) and 4 (colors) may see an initial drop off in response if respondants if are unsure of their favorite shape and/or colors or can't make an quick decision. An option might be to prompt the user by suggesting the most popular shapes and color based on the response to the fit question. Question 5 probably is the lowest because people don't want to share what they consider personal data and/or they are embarrassed about how long its been since their last eye exam. Consider partnering with local eye doctors and change the question (or add an option) to: "Suggest local eye doctors so I can get an eye exam?".

question	Count (distinct 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

```
-- You can put your query here

select question,
count(Distinct user_id)
from survey
group by question;
```

3. 3 vs 5 Pairs– A/B testing with home try-on funnel

3.1 A/B home try-on test

- Warby Parker purchase funnel is defined as: 1) Take the Style Quiz => 2) Home Try-on => 3) Purchase the perfect pair of glasses
- They are interested in determining if the number of pairs of glasses tried on during the home try-on impacts the purchase conversion rate.
- An A/B test was designed for the home try-on stage
  - A customers got 3 pairs to try-on (50% of total)
  - B customers got 5 pairs to try-on (50% of total)
- Data stored in 3 tables (quiz, home_try_on, purchase) with different columns (shown below) and query to load table shown right.

quiz table				
user_id	**style**	**fit**	**shape**	**color**

home_try_on table		
user_id	**Number_of_pairs**	**address**

purchase table					
user_id	**product_id**	**style**	**model_name**	**color**	**price**

```
-- You can put your query here

select *
from quiz
limit 5;

Select *
from home_try_on
limit 5;

Select *
from purchase
limit 5;
```

3.2 Combing tables using left join

- In order to analyze the data and get actionable results it was necessary to join 3 tables (quiz, home_try_on, and purchase) into one table using the Left Join command on the user_id key.
- The resulting table (shown below) has 4 columns and makes it easy to see which customers received 3 pairs vs 5 pairs for home try-on and if they made a purchase. The query is shown on the right.

user_id	is_home_try_on	number_of_pairs	ls_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

```
-- You can put your query here

SELECT DISTINCT q.user_id,
Case When h.user_id IS NOT NULL then 'True' else 'False' End As
'is_home_try_on',
  h.number_of_pairs,
  Case When p.user_id IS NOT NULL then 'True' else 'False' End 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;
```

3.3 Conversion rates

- Data Analysis/insights:
- 75% of customers (750/1000) who took quiz went on to the home try-on
- 49.5% of the total customers (495/1000) who took the quiz purchased glasses
- 66% (495/750) customers who did home try-on purchased glasses
- 53.6% (201/495) of the A group (3 pairs) purchased glasses
- 78.4% (294/495) of the B (5 pairs) purchased glasses

- The conversion rates indicate that customers who took the quiz were likely (75%) to try-on glasses at home. The B group– customers who received 5 pairs of glasses to try-on - were roughly 25% more likely to purchase glasses over the A group. My recommendation to Warby Parker would be to send all customers 5 pairs of glasses to try-on and should lead to an overall increase in the conversion rate of customers who do home try-on and purchase glasses. Warby Parker might consider a second A/B test were customers are given either 5 or 7 pairs of glasses to see if 5 is the right number or the number should be increased to 7.

```
-- You can put your query here

With  funnels 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 (*) as 'num_quiz',
    sum(is_home_try_on) as 'num_home_try_on',
    sum(is_purchase) as 'num_is_purchase',
    1.0 * sum(is_home_try_on) / count(user_id) as 'quiz_to_home_try_on',
    1.0 * sum(is_purchase) / sum(is_home_try_on) as
'home_try_on_to_is_purchase',
    1.0 * Count(case when is_purchase = 1 and number_of_pairs = '3 pairs' then 1
else Null End)/(sum(is_home_try_on)/2) as '3_pairs_to_purchase_rate',
    1.0 *count(case when is_purchase = 1 and number_of_pairs = '5 pairs' then 1
else Null end)/ (sum(is_home_try_on)/2) as '5_pairs_to_purchase_rate'
from Funnels;
```

3.4 Further insights

- The individual tables can also be analyzed to get information on the most popular products and best selling items.
    - Women' styles are slightly more popular than Men's (469 to 432)
    - Narrow fit is the most popular and wide fit the least
    - Straight shapes (rectangular and square) are more popular than round
    - Tortoise (292) and black (280) are most popular colors
    - Most popular Product_ids and times ordered (#) are: 3 (63), 10 (62), 9 (54), 1 (52), 6 (50)
    - Women's styles have been purchases more than mens (252 to 243) which matches quiz responses
    - Eugene Narrow is the most purchased model (116) followed ty Dawes (107) and Brady (95)
    - Jet black is most purchased single color (86) followed by driftwood fade (63) and rosewood tortoise (62). There are multiple options for tortoise so if they were combined they would be most popular.
    - $95 is the most popular price point by 1.4x over $150.

- Warby Parker can use the above information to guide production and inventory decisions. It can also be used to determine if products should be discontinued and/or new products developed. It can also suggest the most popular/best selling glasses to customers who do not answer certain questions on the quiz.

```
-- You can put your query here

select style,
count(*) as 'times_selected'
from quiz
group by 1
order by 2 desc;

select fit,
count(*) as 'times_selected'
from quiz
group by 1
order by 2 desc;

select shape,
count(*) as 'times_selected'
from quiz
group by 1
order by 2 desc;

select color,
count(*) as 'times_selected'
from quiz
group by 1
order by 2 desc;

select product_id,
count(*) as 'times_purchased'
from purchase
group by 1
order by 2 desc;
```

```
-- query continued


select style,
count(*) as 'times_purchased'
from purchase
group by 1
order by 2 desc;

select model_name,
count(*) as 'times_purchased'
from purchase
group by 1
order by 2 desc;

select color,
count(*) as 'times_purchased'
from purchase
group by 1
order by 2 desc;

select price,
count(*) as 'times_purchased'
from purchase
group by 1
order by 2 desc;
```

4. Thank You!