code|cademy

Capstone Project – Usage Funnels with Warby Parker

Learn SQL from Scratch
Alexander Campo
3/15/2019

Table of Contents

1. Getting to Know the Table Schema

1. Getting to Know the Table Schema

To help users find their perfect frame, Warby Parker has a Style Quiz that has the following questions:
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 users' responses are stored in a table called survey.
Select all columns from the first 10 rows. What columns does the table have?

```
-- SQL 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

2. How Many Users Answer Each Survey Question?

2. How Many Users Answer Each Survey Question?

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
-- SQL Query

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

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

3. Analyzing Survey Question Completion Rates

3. Analyzing Survey Question Completion Rates

Using a spreadsheet program like Excel or Google Sheets, calculate the percentage of users who answer each question.:

Click on Spreadsheet to the right to see Calculations:

question	COUNT(DISTINCT user_id)	% Who Complete Question		
1. What are you looking for?	500	100%		
2. What's your fit?	475	95%		
3. Which shapes do you like?	380	80%		
4. Which colors do you like?	361	95%		
5. When was your last eye exam?	270	75%		

Findings:
Which question(s) of the quiz have a lower completion rates?
```
Questions 3 & 5 had the lowest completion rates of the 5 question survey.
3. Which shapes do you like?
5. When was your last eye exam?
```

What do you think is the reason?
```
A possible reason for the low completion rates on questions 3 and 5 may be a result of the customer simply not knowing the
answer.  A casual visitor to the survey may not purchase glasses often and when faced with a question like 'what shapes do
you like?' may not know the shapes of glasses frames.  This could cause them to give up in frustration or possibly navigate
away from Warby Parker to find more information.  Question 5 is more than just a preference question but a real fact of the
visitor, who may not know the date of their last eye exam.  If they don't know they may fear follow up questions about their
prescription, which they may also not know, and would give up before completing the survey.
```

4. Getting to Know the Tables of Warby Parker's Purchase Funnel

4. Getting to Know the Tables of Warby Parker's Purchase Funnel

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
What are the column names?

quiz	home_try_on	purchase
user_id	user_id	user_id
style	number_of_pairs	product_id
fit	address	style
shape		model_name
color		color
		price

```sql
-- SQL Query

SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

5. Joining Tables for Analysis

5. Joining Tables for Analysis

We'd like to create a new table with the following layout:

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc	True	3	False
291f1cca	True	5	False
75122300	False	NULL	False

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_on will 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).

```sql
-- SQL Query

SELECT DISTINCT quiz.user_id,
  CASE
    WHEN home_try_on.user_id IS NOT NULL THEN 'TRUE'
    WHEN home_try_on.user_id IS NULL THEN 'FALSE'
  END AS 'is_home_try_on',
  CASE
    WHEN home_try_on.number_of_pairs = '3 pairs' THEN '3'
    WHEN home_try_on.number_of_pairs = '5 pairs' THEN '5'
    ELSE 'NULL'
  END AS 'number_of_pairs',
  CASE
    WHEN purchase.user_id IS NOT NULL THEN 'TRUE'
    WHEN purchase.user_id IS NULL THEN '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 purchase.user_id = quiz.user_id
LIMIT 10;
```

5.1 Joining Tables for Analysis – Resulting Table

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	TRUE	3	FALSE
291f1cca-e507-48be-b063-002b14906468	TRUE	3	TRUE
75122300-0736-4087-b6d8-c0c5373a1a04	FALSE	NULL	FALSE
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	TRUE	5	FALSE
ce965c4d-7a2b-4db6-9847-601747fa7812	TRUE	3	TRUE
28867d12-27a6-4e6a-a5fb-8bb5440117ae	TRUE	5	TRUE
5a7a7e13-fbcf-46e4-9093-79799649d6c5	FALSE	NULL	FALSE
0143cb8b-bb81-4916-9750-ce956c9f9bd9	FALSE	NULL	FALSE
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	TRUE	5	FALSE
b1dded76-cd60-4222-82cb-f6d464104298	TRUE	3	FALSE

6. Actionable Insights for Warby Parker

6.1 Conversion Rates

Using the Query to the right we are able to determine the overall conversion rates from Quiz to Purchase.
Based on the Query it appears a total of 1000 visitors took the Quiz to some extent.
Following along the funnel we find that only 495 of those users actually completed a purchase from Warby Parker. That's an overall conversion Rate of 49.5%

We can also see that the Quiz to Home Try On Conversion Rate is 75%

And that the Home Try on to Purchase is 66%

Based on those we can determine that the Home Try On does have a positive effect on the overall purchase rate.

```
-- SQL Query

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(user_id),
SUM(is_home_try_on),
SUM(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_purchase'
FROM funnels;
```

COUNT(user_id)	SUM(is_home_try_on)	SUM(is_purchase)	quiz_to_home_try_on	home_try_on_to_purchase
1000	750	495	0.75	0.66

6.2 Conversion Rates Based on Home Try on Numbers

Using the Query to the right we are able to determine the overall conversion rates from Home Try On to Purchase based on the number of pairs the user received.

Based on the results below it is clear that a Home Try on consisting of 5 pairs leads to a much higher purchase conversion rate of 79%

Those users who only received 3 pairs had a purchase conversion rate of 53%

And those who did not have a Home Try On had a purchase rate of 0%

For the Warby Parker A/B Test it is clear that providing the user with 5 pairs leads to a higher overall conversion rate for purchase

number_of_pairs	COUNT(DISTINCT(user_id))	SUM(is_purchase)	Purchase Rate
3	379	201	0.53
5	371	294	0.79
NULL	250	0	0.0

```sql
-- SQL Query

WITH funnels AS(
  SELECT DISTINCT quiz.user_id,
          CASE
  WHEN home_try_on.user_id IS NOT NULL THEN 'TRUE'
  WHEN home_try_on.user_id IS NULL THEN 'FALSE'
  END AS 'is_home_try_on',
  CASE
  WHEN home_try_on.number_of_pairs = '3 pairs' THEN '3'
  WHEN home_try_on.number_of_pairs = '5 pairs' THEN '5'
  ELSE 'NULL'
  END AS '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 = quiz.user_id)
SELECT number_of_pairs,
COUNT(DISTINCT(user_id)),
SUM(is_purchase),
  ROUND(1.0 * SUM(is_purchase) /
COUNT(DISTINCT(user_id)),2) AS 'Purchase Rate'
FROM funnels
WHERE number_of_pairs IS NOT NULL
GROUP BY number_of_pairs;
```

6.3 Conversion Rates Based on Home Try on Numbers

Using the Query to the right we are able to determine the overall conversion rates from Home Try On to Purchase based on the number of pairs the user received.

Based on the results below it is clear that a Home Try on consisting of 5 pairs leads to a much higher purchase conversion rate of 79%

Those users who only received 3 pairs had a purchase conversion rate of 53%

And those who did not have a Home Try On had a purchase rate of 0%

For the Warby Parker A/B Test it is clear that providing the user with 5 pairs leads to a higher overall conversion rate for purchase

number_of_pairs	COUNT(DISTINCT(user_id))	SUM(is_purchase)	Purchase Rate
3	379	201	0.53
5	371	294	0.79
NULL	250	0	0.0

```
-- SQL Query

WITH funnels AS(
  SELECT DISTINCT quiz.user_id,
    CASE
      WHEN home_try_on.user_id IS NOT NULL THEN 'TRUE'
      WHEN home_try_on.user_id IS NULL THEN 'FALSE'
    END AS 'is_home_try_on',
    CASE
      WHEN home_try_on.number_of_pairs = '3 pairs' THEN '3'
      WHEN home_try_on.number_of_pairs = '5 pairs' THEN '5'
      ELSE 'NULL'
    END AS '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 = quiz.user_id)
SELECT number_of_pairs,
  COUNT(DISTINCT(user_id)),
  SUM(is_purchase),
  ROUND(1.0 * SUM(is_purchase) / COUNT(DISTINCT(user_id)),2)
AS 'Purchase Rate'
FROM funnels
WHERE number_of_pairs IS NOT NULL
GROUP BY number_of_pairs;
```

6.4 Most Popular Styles & Products

Using Some simpler queries to look at the singular tables we can get an idea of the most popular styles and products

According to the Quiz the majority of users answered the questions knowing if they were looking for Men's or Women's Styles

With Women's style having a slight edge for more interest

style	COUNT(user_id)
I'm not sure. Let's skip it.	99
Men's Styles	432
Women's Styles	469

```sql
-- SQL Query

SELECT DISTINCT(style),
          COUNT(user_id)
FROM quiz
GROUP BY style;
```

6.5 Most Popular Styles & Products

Using Some simpler queries to look at the singular tables we can get an idea of the most popular styles and products

Unfortunately Not all those who answer the quiz end up purchasing with a conversion rate of
56.25% for Men's Styles
53.73% for Women's Styles

While it appears there is more interest in Women's Styles.
Those interested in the Men's styles are slightly more likely to purchase

```
-- SQL Query

SELECT DISTINCT(style),
            COUNT(user_id)
FROM purchase
GROUP BY style;
```

style	COUNT(user_id)
Men's Styles	243
Women's Styles	252
style	COUNT(user_id)

6.6 Most Popular Styles & Products

Using Some simpler queries to look at the singular tables we can get an idea of the most popular styles and products

Finally, from this last query we can see that the most popular products for purchase are Products 3 & 10

product_id	style	total
3	Men's Styles	63
10	Women's Styles	62
9	Women's Styles	54
1	Men's Styles	52
6	Women's Styles	50
4	Men's Styles	44
7	Women's Styles	44
2	Men's Styles	43
8	Women's Styles	42
5	Men's Styles	41

```sql
-- SQL Query

SELECT DISTINCT(product_id),
          style,
          COUNT(user_id) AS 'total'
FROM purchase
GROUP BY product_id
ORDER BY total DESC;
```