codecademy

Capstone: Funnels with Warby Parker Project

Learn SQL from Scratch

Issac Hernandez

July 2018

Table of Contents

1. Get Familiar With Warby Parker

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.

In this Capstone Project, we will analyze Warby Parker's different marketing funnels in order to calculate conversion rates. Here are the funnels and the tables that you are given: Survey, Quiz, Home Try-On, and Purchase

Survey Funnel

1. To help users find their perfect frame, Warby Parker has a Style Quiz that has the following questions:

   "What are you looking for?"
   "What's your fit?"
   "Which shapes do you like?"
   "Which colors do you like?"
   "When was your last eye exam?"

The users' responses are stored in a table called *survey*.

```
--Columns in the survey

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
5. When was your last eye exam?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	<1 Year

Survey Funnel

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

```
--Number of people who answered each
question

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

- 54% of people who started the quiz completed it

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

Quiz Funnel

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

   Which question(s) of the quiz have a lower completion rates? What do you think is the reason?
   Add this finding to your presentation slides!

```
--Number of people who answered each
question

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

- Question 5 had the lowest answer rate (270), perhaps it is because users had tired of answering questions, or maybe it had been longer than 1 year and users did not want to answer

Question	COUNT(DISTINCT user_id)	Percentage
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%

Home Try-On Funnel

4. 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?

```
-- 3 Tables For The Home Try-On Funnel

SELECT *
FROM quiz
LIMIT 5;

SELECT *
FROM home_try_on
LIMIT 5;

SELECT *
FROM purchase
LIMIT 5;
```

- FOR QUIZ-there are 5 columns: user_id, style, fit, shape, and color
- FOR HOME_TRY_ON there are 3 columns: user_id, number_of_pairs, and address
- FOR PUCHASE there are 6 columns: user_id, product_id, style, model_name, color, and price

* All 3 tables have the *user_id* column

Home Try-On Funnel

A/B Testing From Users Who Elected To Try Frames At Home

- 750 people elected to try frames at home
- Nearly half were sent either 3 pairs or 5 pairs to try on

```
--Number of people who tried frames at home

SELECT COUNT(*)
FROM home_try_on;

-- People received either 3 or 5 frames to try

SELECT number_of_pairs, COUNT(number_of_pairs)
FROM home_try_on
GROUP BY number_of_pairs;
```

number_of_pairs	COUNT(number_of_pairs)
3 pairs	379
5 pairs	371

Home Try-On Funnel

5. 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

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).

Issac Hernandez_Learn SQL from Scratch

```
--Funnel from home try-on to purchase

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;
```

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	⌀	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	⌀	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9	0	⌀	0
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	1	5 pairs	0
b1dded76-cd60-4222-82cb-f6d464104298	1	3 pairs	0

Conversion Rates: Quiz → Home-Try-On → Purchases

- 75% of users who took the *quiz* went on to try frames at home
- 66% of users who tried frames at home went on to purchase a frame

COUNT QUIZ
1000
COUNT HOME TRY ON
750
COUNT PURCHASE
495

```
-- Number of users who took the quiz, vs. who
tried frames at home vs. who purchased frames

SELECT COUNT(DISTINCT user_id)
FROM quiz;

SELECT COUNT(*)
FROM home_try_on;

SELECT COUNT(*)
FROM purchase;

--Conversion Rate from Quiz to Home Try-On

SELECT 750./1000;

--Conversion Rate from Home Try-on to Purchase

SELECT 495./750;
```

Issac Hernandez_Learn SQL from Scratch

Sales Analysis

The top 3 purchased frames are:

1. Eugene Narrow
2. Dawes
3. Brady

* Note that the 4th frame brought in the second highest sales ($12,900)

The 3 highest grossing frames are:

1. Dawes - $16,050
2. Lucy - $12,900
3. Eugene Narrow - $11,020

style	model_name	COUNT(style)	SUM(price)	price
Women's Styles	Eugene Narrow	116	11020	95
Men's Styles	Dawes	107	16050	150
Men's Styles	Brady	95	9025	95
Women's Styles	Lucy	86	12900	150
Women's Styles	Olive	50	4750	95
Men's Styles	Monocle	41	2050	50

```
-- Sales Analysis

SELECT style, model_name,
COUNT(style), SUM(price), price
FROM purchase
GROUP BY model_name
ORDER BY 3 DESC;
```


codecademy