codecademy

Capstone - Funnels

Learn SQL from Scratch

Deidania Chobot]

November 21, 2018

Table of Contents

Usage Funnels with Warby Parker

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 different Warby Parker's marketing funnels in order to calculate conversion rates. Here are the funnels and the tables:

Quiz Funnel:
survey

Home Try-On Funnel:
quiz
home_try_on
purchase

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.

What columns does the table have?

The survey table has three columns:
Question
User_id
Response

survey 1986 rows	
question	TEXT
user_id	TEXT
response	TEXT

```
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
1. What are you looking for?	00a556ed-f13e-4c67-8704-27e3573684cd	I'm not sure. Let's skip it.
2. What's your fit?	00a556ed-f13e-4c67-8704-27e3573684cd	Narrow
5. When was your last eye exam?	00a556ed-f13e-4c67-8704-27e3573684cd	<1 Year
3. Which shapes do you like?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Square
5. When was your last eye exam?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	<1 Year
2. What's your fit?	00bf9d63-0999-43a3-9e5b-9c372e6890d2	Medium

2. What is the quiz funnel:

Users will "give up" at different points in the survey. Let's analyze how many users move from Question 1 to Question 2, etc.
Here we create a quiz funnel using the GROUP BY command.

What is the number of responses for each question?

Here's the number of responses for each question asked:

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

```
SELECT question, COUNT(DISTINCT user_id)
FROM survey
WHERE question IS NOT NULL
GROUP BY 1;
```

Using Excel, we calculate the percentage of users who answer each question. We divide the number of people completing each step by the number of people completing the *previous* step:

Which question(s) of the survey have a lower completion rate?

At 75%, question 5 had the lowest completion rate.

What do you think is the reason?

Question 5 has a low completion rate probably because most people don't remember when was their last eye exam or maybe they did not want to answer this question.

question	COUNT(DISTINCT user_id)
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%

500	
475	0.95
380	0.8
361	0.95
270	0.747922

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:

Home_try_on table	Purchase table	Quiz Table
User_id	User_id	User_id
Number_of_pairs	Product_id	Style
address	style	fit
	Model_name	shape
	color	color
	price	

home_try_on 750 rows	
user_id	TEXT
number_of_pairs	TEXT
address	TEXT

purchase 495 rows	
user_id	TEXT
product_id	INTEGER
style	TEXT
model_name	TEXT
color	TEXT
price	INTEGER

quiz 1000 rows	
user_id	TEXT
style	TEXT
fit	TEXT
shape	TEXT
color	TEXT

Here are the first five rows for each of the three tables:

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
75122300-0736-4087-b6d8-c0c5373a1a04	Women's Styles	Wide	Rectangular	Two-Tone
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	Women's Styles	Narrow	Square	Two-Tone
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black

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
8ba0d2d5-1a31-403e-9fa5-79540f8477f9	5 pairs	287 Pell St
4e71850e-8bbf-4e6b-accc-49a7bb46c586	3 pairs	347 Madison Square N
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs	182 Cornelia St

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
017506f7-aba1-4b9d-8b7b-f4426e71b8ca	4	Men's Styles	Dawes	Jet Black	150
0176bfb3-9c51-4b1c-b593-87edab3c54cb	10	Women's Styles	Eugene Narrow	Rosewood Tortoise	95
01fdf106-f73c-4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black	150

```
--select all columns from quiz table
SELECT *
FROM quiz
LIMIT 5;

--select all columns from home_try_on table
SELECT *
FROM home_try_on
LIMIT 5;

--select all columns from purchase table
SELECT *
FROM purchase
LIMIT 5;
```

3. A/B testing with home try-on funnels:

Here's a new table where each row represents a single user;
Where if the user tried a pair of glasses at home, is_home_try_on is true/1 and if the user made a purchase, is_purchase is true/1.
We also know how many pairs of glasses if any did each user receive to try on.

A left join is used to combine the three tables home_try_on, quiz and purchase.

In conclusion, of the ten users shown here, users that received 3 pairs made more purchases than those that received 5 or none. We need more data to know for sure.

user_id	number_of_pairs	is_home_try_on	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	3 pairs	1	0
291f1cca-e507-48be-b063-002b14906468	3 pairs	1	1
75122300-0736-4087-b6d8-c0c5373a1a04		0	0
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	5 pairs	1	0
ce965c4d-7a2b-4db6-9847-601747fa7812	3 pairs	1	1
28867d12-27a6-4e6a-a5fb-8bb5440117ae	5 pairs	1	1
5a7a7e13-fbcf-46e4-9093-79799649d6c5		0	0
0143cb8b-bb81-4916-9750-ce956c9f9bd9		0	0
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	5 pairs	1	0
b1dded76-cd60-4222-82cb-f6d464104298	3 pairs	1	0

```
SELECT DISTINCT q.user_id,
h.number_of_pairs,
h.user_id IS NOT NULL AS 'is_home_try_on',
p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS 'q'

LEFT JOIN home_try_on AS 'h'
ON h.user_id = q.user_id

LEFT JOIN purchase AS 'P'
ON p.user_id = q.user_id

LIMIT 10;
```

Once we have the data in this format, we can analyze it in several ways:
We can calculate overall conversion rates by aggregating across all rows.
We can compare conversion
from quiz→home_try_on and home_try_on→purchase.

Here we see that 75% of users go from taking the quiz to trying at home and 66% of users go from trying at home to purchasing.

number_of_pairs	num_quizes	num_tries	num_purchase	percentUserQtoT	percentUserTtoP
	1000	750	495	0.75	0.66

```
WITH funnels AS
(
SELECT DISTINCT q.user_id,
h.number_of_pairs,
h.user_id IS NOT NULL AS 'is_home_try_on',
p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS '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
number_of_pairs,
COUNT(*) AS 'num_quizes',
SUM(is_home_try_on) AS 'num_tries',
SUM(is_purchase) AS 'num_purchase',

1.0 * SUM(is_home_try_on) / COUNT(user_id) AS
'percentUserQtoT',

1.0 * SUM(is_purchase) / SUM(is_home_try_on) AS
'percentUserTtoP'

FROM funnels;
```

We can calculate the difference in purchase rates between customers who had 3 number_of_pairs with ones who had 5.

We know that 250 users did the quiz but did not receive any pairs to try on. 379 users received 3 pairs to try on and 371 users received 5 pairs to try on.
We know that users that did not receive any pairs did not make a purchase.
Users that received 3 pairs made 201 purchases, whereas users who received 5 pairs made 294 purchases.
In comparing conversion rates:
53% of users that received 3 pairs made a purchase.
79% of users that received 5 pairs made a purchase.

umber_of_pairs	num_quizes	num_tries	num_purchase	percentUserQtoT	percentUserTtoP
	250	0	0	0.0	
3 pairs	379	379	201	1.0	0.530343007915567
5 pairs	371	371	294	1.0	0.792452830188679

```
WITH funnels AS
(
SELECT DISTINCT q.user_id,
h.number_of_pairs,
h.user_id IS NOT NULL AS 'is_home_try_on',
p.user_id IS NOT NULL AS 'is_purchase'
FROM quiz AS '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
number_of_pairs,
COUNT(*) AS 'num_quizes',
SUM(is_home_try_on) AS 'num_tries',
SUM(is_purchase) AS 'num_purchase',

1.0 * SUM(is_home_try_on) / COUNT(user_id) AS
'percentUserQtoT',

1.0 * SUM(is_purchase) / SUM(is_home_try_on) AS
'percentUserTtoP'

FROM funnels
GROUP BY 1
ORDER BY 1;
```

4. Insights for Warby Parker:

Style Quiz

The most common result from the style quiz was Women's Styles with 469 counts, the least common was Men's styles with 432 results and 99 where not sure of the style.

Warby Parker should target more women in their marketing campaigns!

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

```
SELECT style, COUNT(style)
FROM quiz
GROUP BY 1;
```

Fit

The most popular fit from the quiz was the Narrow with 408 counts the least popular was wide with 198 counts and a small number of people 89 was not sure of the fit.

More people are interested in Narrow fit, maybe Warkby parker should make more Narrow fitting styles.

fit	COUNT(fit)
I'm not sure. Let's skip it.	89
Wide	198
Medium	305
Narrow	408

```
SELECT fit, COUNT(fit)
FROM quiz
GROUP BY 1
ORDER BY 2;
```

Shape

The most popular shape from the quiz was Rectangular with 397 counts, the least popular was Round with 180 results. A small number 97, had no preference.

Warby Parker should make more Rectangular shaped glasses.

```
SELECT shape, COUNT(shape)
FROM quiz
GROUP BY 1
ORDER BY 2;
```

shape	COUNT(shape)
No Preference	97
Round	180
Square	326
Rectangular	397

Color

The most popular color from the quiz was Tortoise with 292 counts, the least popular was Two-Tone with 104 results.

Tortoise is a classic color that matches many outfits. Warby Parker should make more styles in Tortoise.

```
SELECT color, COUNT(color)
FROM quiz
GROUP BY 1
ORDER BY 2;
```

color	COUNT(color)
Two-Tone	104
Neutral	114
Crystal	210
Black	280
Tortoise	292

Price

Pairs that were priced at $95 sold the most with 261 purchases, pairs that were priced at $50 sold the least with 41 purchases. Warby Parker needs to have more pairs available at $95!

price	COUNT(price)
50	41
95	261
150	193

```
SELECT price, COUNT(price)
FROM purchase
GROUP BY 1
ORDER BY 1 ASC;
```

Model

The least popular model was Monocle with 41 purchases and the most popular model was Eugene Narrow with 116 purchases.

model_name	COUNT(model_name)
Monocle	41
Olive	50
Lucy	86
Brady	95
Dawes	107
Eugene Narrow	116

```
SELECT model_name, COUNT(model_name)
FROM purchase
GROUP BY 1
ORDER BY 2 ASC;
```

Color Purchased

The most popular color purchased was the Jet Black with 86 purchases and the least popular was the Endangered Tortoise with 41 purchases.

Might be a good idea to change this color's name!

color	COUNT(color)
Endangered Tortoise	41
Sea Glass Gray	43
Elderflower Crystal	44
Pearled Tortoise	50
Layered Tortoise Matte	52
Rose Crystal	54
Rosewood Tortoise	62
Driftwood Fade	63
Jet Black	86

```
SELECT color, COUNT(color)
FROM purchase
GROUP BY 1
ORDER BY 2;
```

Style Purchased

The most popular style purchased was the women’s styles with 252 purchases, but not by much since 243 men’s styles were also purchased.

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

```
SELECT style, COUNT(style)
FROM purchase
GROUP BY 1
ORDER BY 2;
```