codecademy

Warby Parker

Learn SQL from Scratch

Elzbieta Quintana

2/13/2019

Table of Contents

1. Warby Parker Overview

1.1 Warby Parker Overview

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.

1.2 Warby Parker Marketing Funnels

In this Capstone Project, the analysis is performed on different Warby Parker's marketing funnels in order to calculate conversion rates.

Warby Parker's purchase funnel is:
Take the Style Quiz → Home Try-On → Purchase the Perfect Pair of Glasses

The analysis will be performed on the following tables:

- Survey
- Quiz
- Home_try_on
- Purchase

2. Survey Funnel

2.1 Style Quiz – “Survey” table

Warby Parker has a Style Quiz that asks questions to determine the glasses frame that will fit perfectly the customer:

- These questions are stored in the table called “survey”
- Survey has the following columns: question, user_id and response (see the example of the results)

```
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

2.2 Number of responses to each question in a Style Quiz

Here we calculated the number of distinct users answering each question. The number of responses for each question is as follows:

- Question 1 – 500 responses
- Question 2 – 475 responses
- Question 3 – 380 responses
- Question 4 – 361 responses
- Question 5 – 270 responses

```
SELECT question, COUNT(DISTINCT user_id)
FROM survey
GROUP BY 1
ORDER BY 1;
```

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

2.3 Question Completion Rate

Using prior results in Excel, you can calculate the percentage of users who answer each question.

- The third column was calculated in the following way: the number of people completing each step by the number of people completing the previous step
  - Question number 5 has the lowest completion rate (75%) while question 2 and 5 have the highest (95%)
  - This suggest that doctor's visits are more sensitive subject and people might be reluctant to answer it
- The fourth column indicates how many people completed each step as compared to the number at the start of the survey
  - We can see that 54% of people who started the survey actually completes it

question	COUNT(DISTINCT user_id)	% Completing this Question to Prior Question	% Completing this Question to First Question
1. What are you looking for?	500	100%	100%
2. What's your fit?	475	95%	95%
3. Which shapes do you like?	380	80%	76%
4. Which colors do you like?	361	95%	72%
5. When was your last eye exam?	270	75%	54%

3. Conversion Rates

3.1 Data Sets

Warby Parker's purchase funnel is:
Take the Style Quiz → Home Try-On → Purchase the Perfect Pair of Glasses

The data is distributed across three tables:
- quiz
- home_try_on
- Purchase

The quiz table contains the following columns:
- User_id
- Style
- Fit
- Shape
- Color

The home_try_on table contains the following columns:
- User_id
- Number_of_pairs
- Address

The purchase table contains the following columns:
- 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;
```

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

3.2 Combine Three Tables

First, we want to combine the information from the three tables (quiz, home_try_on and purchase) into one table with the following layout:

- User_id
- Is_home_try_on
- Number_of_pairs
- Is_purchase

I selected first 10 rows from this table to have a glance how the table would look like.

user_id	is_home_try_on	number_of_pairs	is_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	Ø	False
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	True	5 pairs	False
ce965c4d-7a2b-4db6-9847-601747fa7812	True	3 pairs	True
28867d12-27a6-4e6a-a5fb-8bb5440117ae	True	5 pairs	True
5a7a7e13-fbcf-46e4-9093-79799649d6c5	False	Ø	False
0143cb8b-bb81-4916-9750-ce956c9f9bd9	False	Ø	False
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	True	5 pairs	False
b1dded76-cd60-4222-82cb-f6d464104298	True	3 pairs	False

```
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 AS 'q'
LEFT JOIN home_try_on AS 'h'
ON q.user_id = h.user_id
LEFT JOIN purchase AS 'p'
ON h.user_id = p.user_id
LIMIT 10;
```

3.3 Aggregated Conversion Rate

The columns of the table below represents the following information:

- Num_quiz – number of people who took quiz
- Num_tried – number of people who tried the glasses on
- Num_purchased – number of people who purchased the glasses
- Tried_to_quiz – percentage of users trying the glasses at home to the total number of users who took quiz
- Purchased_to_tried – Percentage of users who purchased the glasses to the total number of users who tried the glasses at home
- Purchased_to_quiz – Percentage of users who purchased the glasses to the total number of users who took quiz

75% of people who take the quiz take the glasses home to try on and 66% of people who try the glasses at home, purchase them. Almost 50% of users who take the quiz, actually buys the glasses.

The management should look into the ways to improve the conversion rate for the customers who take the glasses to try on to the final stage of purchasing them.

num_quiz	num_tried	num_purchased	tried_to_quiz	purchased_to_tried	purchased_to_quiz
1000	750	495	0.75	0.66	0.495

```
WITH funnels AS
(SELECT DISTINCT q.user_id, 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 q.user_id = h.user_id
LEFT JOIN purchase AS 'p'
ON h.user_id = p.user_id
)
SELECT COUNT(*) AS 'num_quiz',
SUM(is_home_try_on) AS 'num_tried',
SUM(is_purchase) AS 'num_purchased',
1.0 * SUM(is_home_try_on) / COUNT(*) AS
'tried_to_quiz',
1.0 * SUM(is_purchase) / SUM(is_home_try_on) AS
'purchased_to_tried',
1.0 * SUM(is_purchase) / COUNT(*) AS
'purchased_to_quiz'
FROM funnels;
```

3.4 Home Try-On with A/B Test – Steps

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

To determine whether or not users who get more pairs to try on at home will be more likely to make a purchase, we will perform the following steps:

- Step 1: Combine the information from the two tables (home_try_on and purchase) into one table
- Step 2: Calculate the difference in purchase rates between customers who had 3 number of pairs with ones who had 5.

3.4 Home Try-On with A/B Test – Step 1

Step 1 Combine the information from the three tables (quiz, home_try_on and purchase) into one table

- The table below provides information about the which type of variant of “Try On” (3 pairs of glasses or 5 pair of glasses) for each unique user ID and whether the purchase was made
- The results were limited to the first 100 customers for the ease of query calculation

number_of_pairs	user_id	is_purchase
5 pairs	d8addd87-3217-4429-9a01-d56d68111da7	1
5 pairs	f52b07c8-abe4-4f4a-9d39-ba9fc9a184cc	0
5 pairs	8ba0d2d5-1a31-403e-9fa5-79540f8477f9	0
3 pairs	4e71850e-8bbf-4e6b-accc-49a7bb46c586	0
5 pairs	3bc8f97f-2336-4dab-bd86-e391609dab97	1
5 pairs	4c10e298-53c8-4009-adda-bbcaecb7e8b6	0
5 pairs	5a3ee321-517d-4a21-a351-d6815ab2edd5	1
3 pairs	4d895ccf-4877-4f13-8183-13d7d0a20a47	1
3 pairs	39e8a811-75b9-4dc3-bdff-c92b6db0431d	1
5 pairs	9d2656a0-d066-4b42-bce1-77825f34ded9	0
3 pairs	a0ca93b5-a5ed-4660-9a1c-e91c917da388	0

```
WITH funnels AS
(SELECT DISTINCT h.number_of_pairs,
h.user_id,p.user_id IS NOT NULL AS 'is_purchase'
FROM home_try_on AS 'h'
LEFT JOIN purchase AS 'p'
ON h.user_id = p.user_id
)
SELECT *
FROM funnels
LIMIT 100;
```

3.4 Home Try-On with A/B Test – Step 2

Step 2 Calculate the difference in purchase rates between customers who had 3 number of pairs with ones who had 5.

- Based on the table below 379 people tried 3 pairs of glasses, while 371 people tried 5 pairs of glasses, so the test was not split equally to each group. Since the difference is not significant as compared to the whole population, this variance doesn't impact the interpretation of the results in the next bullet point
- The results show that the purchase rate of customers who tried 5 pair of glasses was greater (at 79% purchase rate) than ones who tried 3 pair of glasses (at 53%)

By increasing the number of pair of glasses that customers can take home to try on, the management will be able to increase the sales of the glasses.

number_of_pairs	num_tried	num_purchased	purchased_to_tried
3 pairs	379	201	0.530343007915567
5 pairs	371	294	0.792452830188679

```
WITH funnels AS
(SELECT DISTINCT h.number_of_pairs,
h.user_id,p.user_id IS NOT NULL AS 'is_purchase'
FROM home_try_on AS 'h'
LEFT JOIN purchase AS 'p'
ON h.user_id = p.user_id
)
SELECT number_of_pairs, COUNT(*) AS 'num_tried',
SUM(is_purchase) AS 'num_purchased',
1.0 * SUM(is_purchase) / COUNT(*) AS
'purchased_to_tried'
FROM funnels
GROUP BY 1;
```

4. Other Analysis

4.1 Popular Fits, Shapes and Colors from Quiz

From the “quiz” table, we analyzed the types of Fits, Shapes and Colors that were the most popular based on the survey on the potential customers.

Based on the answers provided by the users, the glasses with the narrow fit, rectangular and square shapes and tortoise and black colors were the most desirable by customers.

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

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

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

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

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

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

4.2 Popular Selling Models

From the “purchase” table, we analyzed the models and colors that were the most popular based on the sales data.

Based on the data, the model “Eugene Narrow” and the color “Jet Black” were the most selling types.

```
SELECT model_name, COUNT(*)
FROM purchase
GROUP BY model_name
ORDER BY 2 DESC;

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

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

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

4.3 Women’s vs Men’s Styles Sold

The analysis at the table below that there is not much difference in the number of pair of glasses purchased by men or women.

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

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

4.4 Most Sales at Price Level

In the table below, we analyzed the following:

- Price – amount per one pair of glasses
- Num_of_glasses_sold – the number of glasses sold at the price level
- Revenue_at_price_level – price multiplied by the number of glasses sold

Based on the results, the highest price ($150) per glasses brings the most revenue, while the average price (about $100) has the highest number of glasses sold. Since the variance between the total revenue at the highest price and the revenue at the average price is not significant, the management should keep the price level between $100 and $150 to generate the most of revenue.

price	num_of_glasses_sold	revenue_at_price_level
95	261	24795
150	193	28950
50	41	2050

```
SELECT price, COUNT(*) AS 'num_of_glasses_sold',
price * COUNT(*) AS 'revenue_at_price_level'
FROM purchase
GROUP BY 1
ORDER BY 2 DESC;
```

5. Summary

5.1 Summary

Based on the analysis, the management could try the following to increase the sales:

- Drop or change the last question on the survey, so that more people don’t give up on going to the next step of trying the glasses at home
- The management should look into the ways to improve the conversion rate for the customers who take the glasses to try on to the final stage of purchasing them. For example:
  - Give more pairs of glasses to try at home
  - Create more models with the most popular fits and shapes, like narrow and medium fit and rectangular or square shape
  - The management should keep the price level between $100 and $150 to generate the most of revenue.