codecademy

WARBY PARKER

Warby Parker Survey Data Analysis with SQL

Jose Luis Garcia Tucci

26/04/2022

codecademy

WARBY PARKER

Table of Contents

1. Context

codecademy

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 project, different Warby Parker's marketing funnels are to be analyze in order to calculate conversion rates.

Note: data analyzed in this case study is fictional and is only for Codecademy Analyze Data with SQL Skill Path course fulfillment

codecademy

WARBY PARKER

2. Quiz Funnel

Question 1

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 let's query first 10 lines

Database Schema	
survey	
name	type
question	TEXT
user_id	TEXT
response	TEXT
Rows: 1986	

```
-- SQL Query
SELECT *
FROM survey
LIMIT 10;
```

Query results		
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

Questions 2 & 3

What is the number of responses for each question?
Which question(s) of the quiz have a lower completion rates?
What could be the reason?

```
-- SQL Query

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

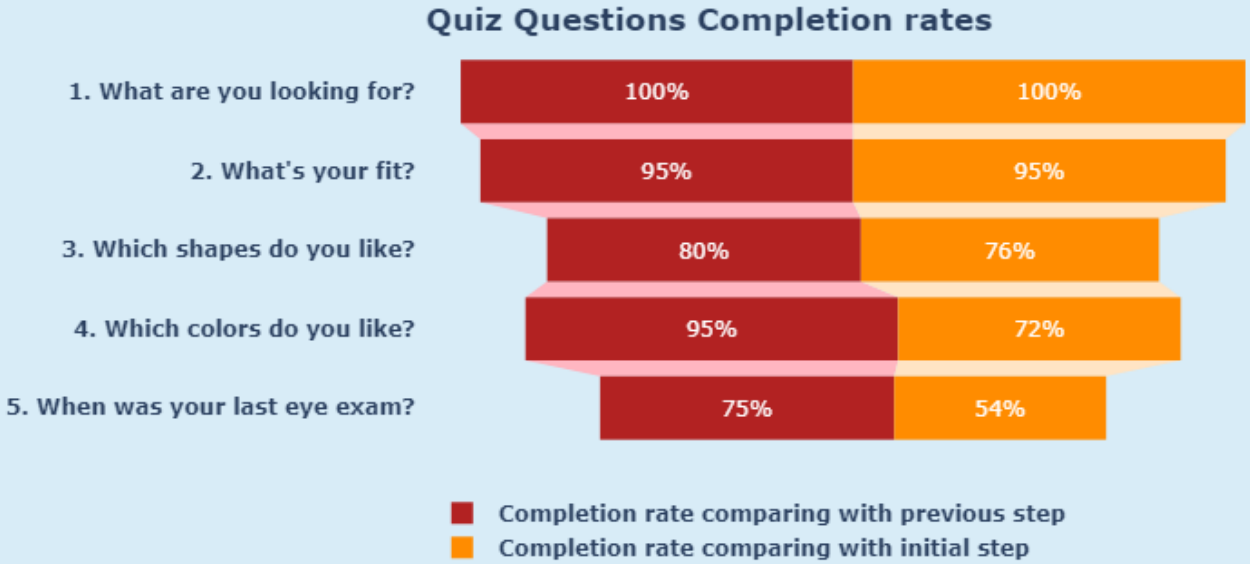
Query Results	
question	user_count
1. What are you looking for?	500
2. What's your fit?	475
3. Which shapes do you like?	380
4. Which colors do you like?	361
5. When was your last eye exam?	270

Question Number	User Count Completing the Question	Completion rate compare with previous step %
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%

Only 270 users responded question 5 this being the less answered question.

Questions 2 & 3

What could be the reason for this completion rates?



Only 54% of the initial users that started the survey answered up to question number 5, it also can be seen 15% of users who answered question 2. choose to skip question 3.

Reason for people skipping question number 2 might be that making a decision o the shape without a fit test might be difficult and user might skip this to get all possible options at the end, but when it comes to color choice people somehow have this preferences set from a long time ago.

On question number 5 where nearly half of the users that started the survey reject answering this can be due to personal decision on not sharing medical information which might be sensitive to many people.

3. Purchase Funnel

codecademy

WARBY PARKER

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 is distributed across three tables:
- quiz
- home_try_on
- purchase

Note: data analyzed in this case study is fictional and is only for Codecademy Analyze Data with SQL Skill Path course fulfillment

Question 4

Let’s query first 5 lines for tables quiz and home_try_on (purchase table is on next slide)

```
-- SQL Query
SELECT *
FROM purchase
LIMIT 5;
```

Database Schema	
purchase	
name	type
user_id	TEXT
style	TEXT
fit	TEXT
shape	TEXT
color	TEXT
Rows: 1000	

Query results				
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

```
-- SQL Query
SELECT *
FROM home_try_on
LIMIT 5;
```

Database Schema	
home_try_on	
name	type
user_id	TEXT
number_of_pairs	TEXT
address	TEXT
Rows: 750	

Query results		
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

Question 4

Let’s query first 5 lines for purchase table

```
-- SQL Query
SELECT *
FROM purchase
LIMIT 5;
```

Database Schema	
purchase	
name	type
user_id	TEXT
product_id	TEXT
style	TEXT
model_name	TEXT
color	TEXT
price	INTEGER
Rows: 495	

Query results

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

Question 5

Create a 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

```
-- SQL Query
WITH purchase_funnel
AS (SELECT DISTINCT quiz.user_id,
    CASE
    WHEN home_try_on.user_id IS NOT NULL THEN 'True'
    ELSE 'False' END AS 'is_home_try_on',
    CASE
    WHEN home_try_on.number_of_pairs IS NULL THEN 'NULL'
    ELSE home_try_on.number_of_pairs END AS 'number_of_pairs',
    CASE
    WHEN purchase.user_id IS NOT NULL THEN 'True'
    ELSE 'False' END AS 'is_purchase',
    CASE
    WHEN home_try_on.number_of_pairs IS '3 pairs' THEN 'A'
    WHEN home_try_on.number_of_pairs IS '5 pairs' THEN 'B'
    ELSE 'N/A'END AS 'test'
    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 *
FROM purchase_funnel
LIMIT 10;
```

Query results

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

Test column was added for ease of analysis

Question 6

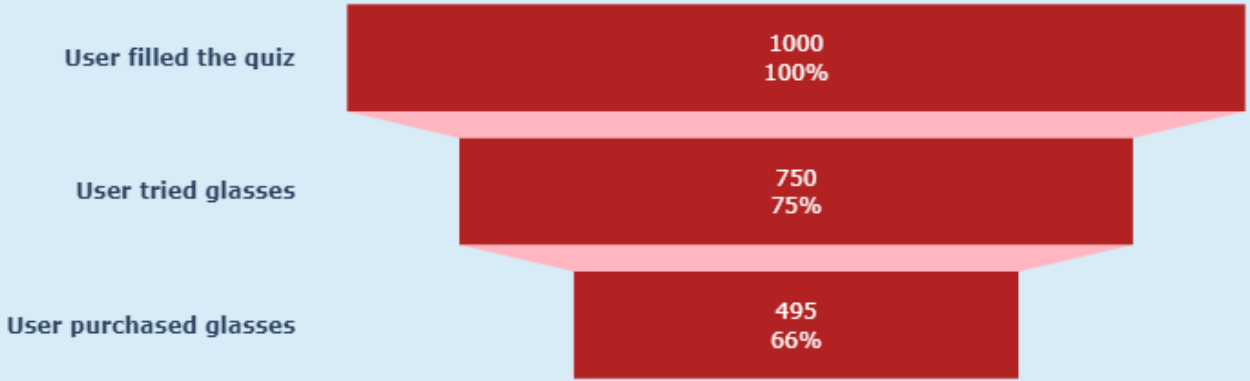
Conversion Rates and different analysis using purchase_funnel table created in previous slide

Let's make a query to get relationship how many users filled the quiz, how many tried glasses and how many purchased glasses at the end

```
-- SQL Query
SELECT COUNT(*) AS 'User filled quiz',
       SUM(CASE
           WHEN is_home_try_on ='True'
           THEN 1 ELSE 0 END) AS 'User tried glasses',
        SUM(CASE
           WHEN is_purchase ='True'
           THEN 1 ELSE 0 END) AS 'User purchased glasses'
FROM purchase_funnel;
```

Query results		
User filled quiz	User tried glasses	User purchased glasses
1000	750	495

Completion rate comparing with previous step



66% of users that tried glasses at home ended up purchasing glasses.

As hidden insight in the data is that 250 user that filled the quiz did not tried glasses that is 25% of the population that filled the quiz,

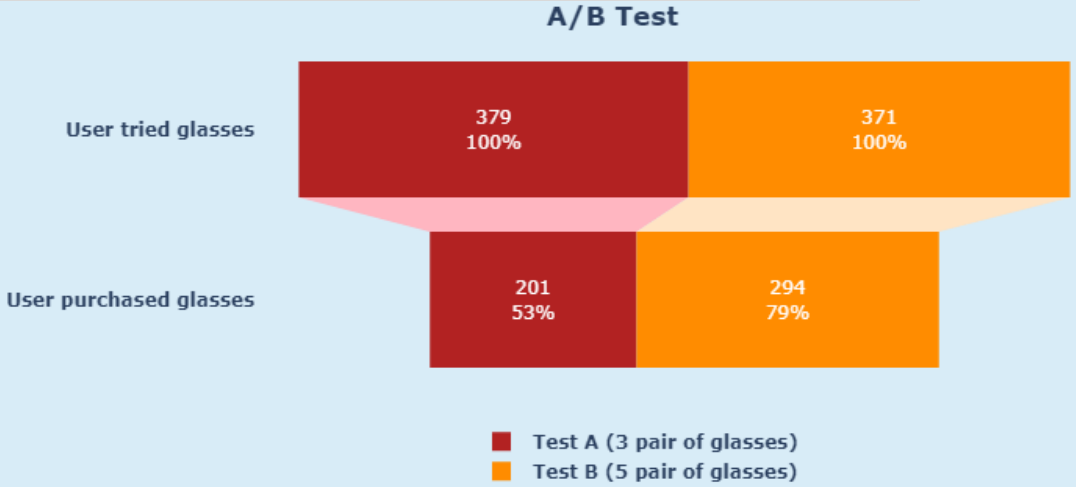
Question 6

Conversion Rates and different analysis using purchase_funnel table created in previous slide

Taking the analysis further let's query test group (A=User tried 3 pairs, B= User tried 5 pairs)

```
-- SQL Query
SELECT test,
       COUNT(*) AS 'User filled quiz',
       SUM(CASE
           WHEN is_home_try_on ='True'
           THEN 1 ELSE 0 END) AS 'User tried glasses',
        SUM(CASE
           WHEN is_purchase ='True'
           THEN 1 ELSE 0 END) AS 'User purchased glasses'
FROM purchase_funnel
WHERE TEST IS NOT 'N/A'
GROUP BY test;
```

Query results		
test	User tried glasses	User purchased glasses
A	379	201
B	371	294



From the A/B test results 79% of users that tried 5 pair of glasses ended up purchasing against 53% in the group that tested 3 pairs.

Question 6

Conversion Rates and different analysis using purchase_funnel table created in previous slide

Let's look which product and styles are more popular in sales

```
-- SQL Query
SELECT product_id,
       style,
       model_name,
       color,
       SUM(price) AS total_sales,
       COUNT(*) AS 'Sold Units'
FROM purchase
GROUP BY product_id
ORDER BY 6 DESC
```

Query results

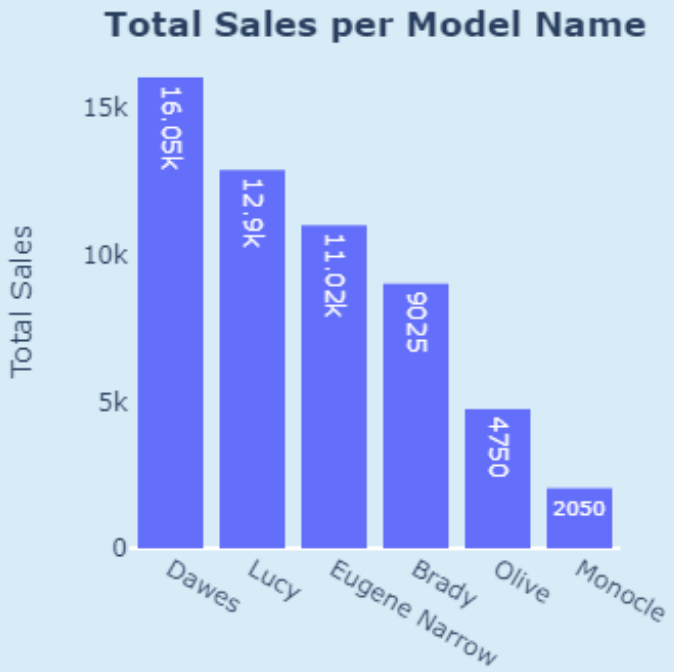
product_id	style	model_name	color	total_sales	Sold Units
3	Men's Styles	Dawes	Driftwood Fade	9450	63
10	Women's Styles	Eugene Narrow	Rosewood Tortoise	5890	62
9	Women's Styles	Eugene Narrow	Rose Crystal	5130	54
1	Men's Styles	Brady	Layered Tortoise Matte	4940	52
6	Women's Styles	Olive	Pearled Tortoise	4750	50
7	Women's Styles	Lucy	Elderflower Crystal	6600	44
4	Men's Styles	Dawes	Jet Black	6600	44
2	Men's Styles	Brady	Sea Glass Gray	4085	43
8	Women's Styles	Lucy	Jet Black	6300	42
5	Men's Styles	Monocle	Endangered Tortoise	2050	41

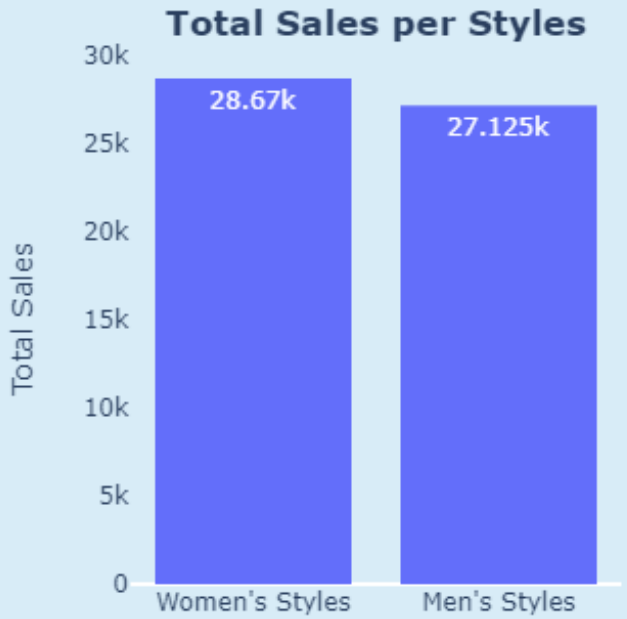
The most popular product is product_id 3 model name Dawes in Men's Style which also reported the most sales and sold units

Question 6

Conversion Rates and different analysis using purchase_funnel table created in previous slide

Let’s look which product and styles are more popular in sales





Question 6

Conversion Rates and different analysis using purchase_funnel table created in previous slide

Let’s look which product and styles are more popular in sales