# Analyze Data with SQL

## Warby Parker marketing funnels

*Shaheen Nazar*
*MSIS, SCU*

# About Warby Parker

Warby Parker is a transformative lifestyle brand

- 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.

- Objective: To offer designer eyewear at a revolutionary price while leading the way for socially conscious businesses.

- Mission: For every pair of eyeglasses and sunglasses sold, a pair is distributed to someone in need.

# What is a Funnel?

A **funnel** is a marketing model which illustrates the theoretical customer journey towards the purchase of a product or service.

Oftentimes, we want to track how many users complete a series of steps and know which steps have the greatest number of users giving up.

Some examples include:

➢ Answering each part of a 5 Question survey on customer satisfaction

➢ Browsing a selection of products → Viewing a shopping cart → Making a purchase

Generally, we want to know the total number of users in each step of the funnel, as well as the percent of users who complete each step.

Using SQL, we can dive into complex funnels and event flow analysis to gain insights into the users' behavior.

# Quiz Funnel : Task 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.

▶ Select all columns from the first 10 rows.

▶ What columns does the table have?

# Solution with Query

The survey table has the following columns:

- question
- user_id
- response

```
-- query here

SELECT *
FROM survey
LIMIT 10;
```

| question | user_id | response |
|---|---|---|
| **Query Results** | | |
| 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 |

# Quiz Funnel : Task 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 BY command.

▶ *What is the number of responses for each question?*

# Solution with Query

The survey table has the following columns:

- question
- user_id
- response

```
-- query here

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 |

# Quiz Funnel : Task 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?

# Solution

- We got the percentage by dividing the number of people completing each step by the number of people completing the previous step.

- The percentage drop in 3rd question shows that customers are quite unsure of the shapes they like.

- Last question shows the maximum drop, the reason might be "customers are quite hesitant to share their last date of eye exam" or they don't even remember the date.

| Question Number | Percent Completing this question |
| --- | --- |
| 1 | 100% |
| 2 | 95% |
| 3 | 80% |
| 4 | 95% |
| 5 | 75% |

# Home Try-On Funnel: Task 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 if 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?

# 4.1 Solution with Query

The quiz table has the following columns:

- user_id
- style
- fit
- shape
- color

```
-- query here

SELECT *
FROM quiz
LIMIT 10;
```

| 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 |
| 28867d12-27a6-4e6a-a5fb-8bb5440117ae | Women's Styles | Narrow | Rectangular | Black |
| 5a7a7e13-fbcf-46e4-9093-79799649d6c5 | Women's Styles | Wide | Rectangular | Tortoise |
| 0143cb8b-bb81-4916-9750-ce956c9f9bd9 | Women's Styles | Wide | Rectangular | Two-Tone |
| a4ccc1b3-cbb6-449c-b7a5-03af42c97433 | Women's Styles | I'm not sure. Let's skip it. | Square | Tortoise |
| b1dded76-cd60-4222-82cb-f6d464104298 | Women's Styles | Narrow | Rectangular | Crystal |

# 4.2 Solution with Query

The home_try_on table has the following columns:

- user_id
- number_of_pairs
- address

```
-- query here

SELECT *
FROM home_try_on
LIMIT 10;
```

| 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 |
| 4c10e298–53c8–4009–adda–bbcaecb7e8b6 | 5 pairs | 312 Frawley Cir |
| 5a3ee321–517d–4a21–a351–d6815ab2edd5 | 5 pairs | 301 Fred Douglass Cir |
| 4d895ccf–4877–4f13–8183–13d7d0a20a47 | 3 pairs | 77 Margaret Corbin Dr |
| 39e8a811–75b9–4dc3–bdff–c92b6db0431d | 3 pairs | 482 Coenties Slip |
| 9d2656a0–d066–4b42–bce1–77825f34ded9 | 5 pairs | 364 De Peyster St |

# 4.3 Solution with Query

The purchase table has the following columns:
- user_id
- product_id
- style
- model_name
- Color
- price

```
-- query here

SELECT *
FROM purchase
LIMIT 10
```

### 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 |
| 021901a5-74ee-482d-9f03-0089d17ff7d6 | 3 | Men's Styles | Dawes | Driftwood Fade | 150 |
| 026e11d9-926c-4206-af7c-273e3ee6ad7f | 2 | Men's Styles | Brady | Sea Glass Gray | 95 |
| 028b47b0-e075-4b2c-bcf3-f963d8089449 | 10 | Women's Styles | Eugene Narrow | Rosewood Tortoise | 95 |
| 029e417e-a0b5-4ec7-9190-a852016619d9 | 5 | Men's Styles | Monocle | Endangered Tortoise | 50 |
| 02bd0f5a-3f04-4902-b005-82188b07785f | 7 | Women's Styles | Lucy | Elderflower Crystal | 150 |

# Home Try-On Funnel: Task 5

Create a new table with the desirable layout on the right side

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 purchase,
   then is_purchase will be True.

▶ Select only the first 10 rows from this table
   (otherwise, the query will run slowly)

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

# Solution with Query

In order to get the desired table, I used a LEFT JOIN to combine the three tables, starting at the top of the funnel (quiz) and ending with the bottom of the funnel (purchase).

```
-- query here

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;
```

**Query Results**

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

# Home Try-On Funnel: Task 6

**Now the data is in desired format, we can analyze it in several ways:**

✓ Calculate overall conversion rates by aggregating across all rows

✓ Compare conversion from

   quiz → home_try_on  and  home_try_on → purchase.

✓ Calculate the difference in purchase rates between customers who had 3 number_of_pairs with ones who had 5

**Original tables can be used to calculate things like:**

✓ The most common results of the style quiz

✓ The most common types of purchase made.

# 6.1 Building a funnel for analyzing the data

Creating a temporary table with the name "Funnel"

Querying the Funnel table for calculating overall conversion rates.

| Query Results | | |
|---|---|---|
| num_quiz | num_home_try_on | num_purchase |
| 1000 | 750 | 495 |

```
-- query here

WITH Funnel 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(*) AS 'num_quiz',
SUM(is_home_try_on) AS 'num_home_try_on',
SUM(is_purchase) AS 'num_purchase'
FROM Funnel;
```

# 6.2 Analyzing the data continued..

Querying the temporarily created Funnel table for

- Percentage of users from "quiz" to "home try on"
- Percentage of users from "home try on" to "purchase"

| Query Results | |
|---|---|
| Quiz_to_Hometrial | Hometrial_to_Purchase |
| 0.75 | 0.66 |

```
-- query here

WITH Funnel 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
SUM(is_home_try_on) * 1.0 / COUNT(user_id) AS
'Quiz_to_Hometrial',
SUM(is_purchase) * 1.0 / SUM(is_home_try_on) AS
'Hometrial_to_Purchase'
FROM Funnel;
```

# 6.3 Analyzing the data continued..

Querying the temporarily created Funnel table for :

Calculating the difference in purchase rates between customers who had 3 number_of_pairs with ones who had 5

When customers were provided with 5 pairs for home trial, the conversion rate to purchase is 79.24%, however when provided with only 3 pairs, the conversion rate to purchase got restricted to 53.03%

Hence proved that if users get more pairs to try at home, they are more likely to make a purchase.

| Query Results | | |
|---|---|---|
| AB_variant | home_trial | purchase |
| Ø | 0 | 0 |
| 3 pairs | 379 | 201 |
| 5 pairs | 371 | 294 |

```
-- query here

WITH Funnel AS (
SELECT DISTINCT q.user_id,
h.user_id IS NOT NULL AS 'is_home_try_on',
h.number_of_pairs AS 'AB_variant' ,
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 AB_variant ,
SUM(is_home_try_on) AS "home_trial",
SUM(is_purchase) AS "purchase"
FROM Funnel
GROUP BY AB_variant;
```

# 6.4 Analyzing the data continued..

1. The most common results of the style 'quiz'
2. The most common types of 'purchase' made

The most common style in quiz table is Women's style

The model called "Lucy" is the most purchased model in Women's style and model called "Dawes" is the most purchased model in Men's style.

**Query Results**

| style | common_style |
|---|---|
| Women's Styles | 469 |
| Men's Styles | 432 |
| I'm not sure. Let's skip it. | 99 |
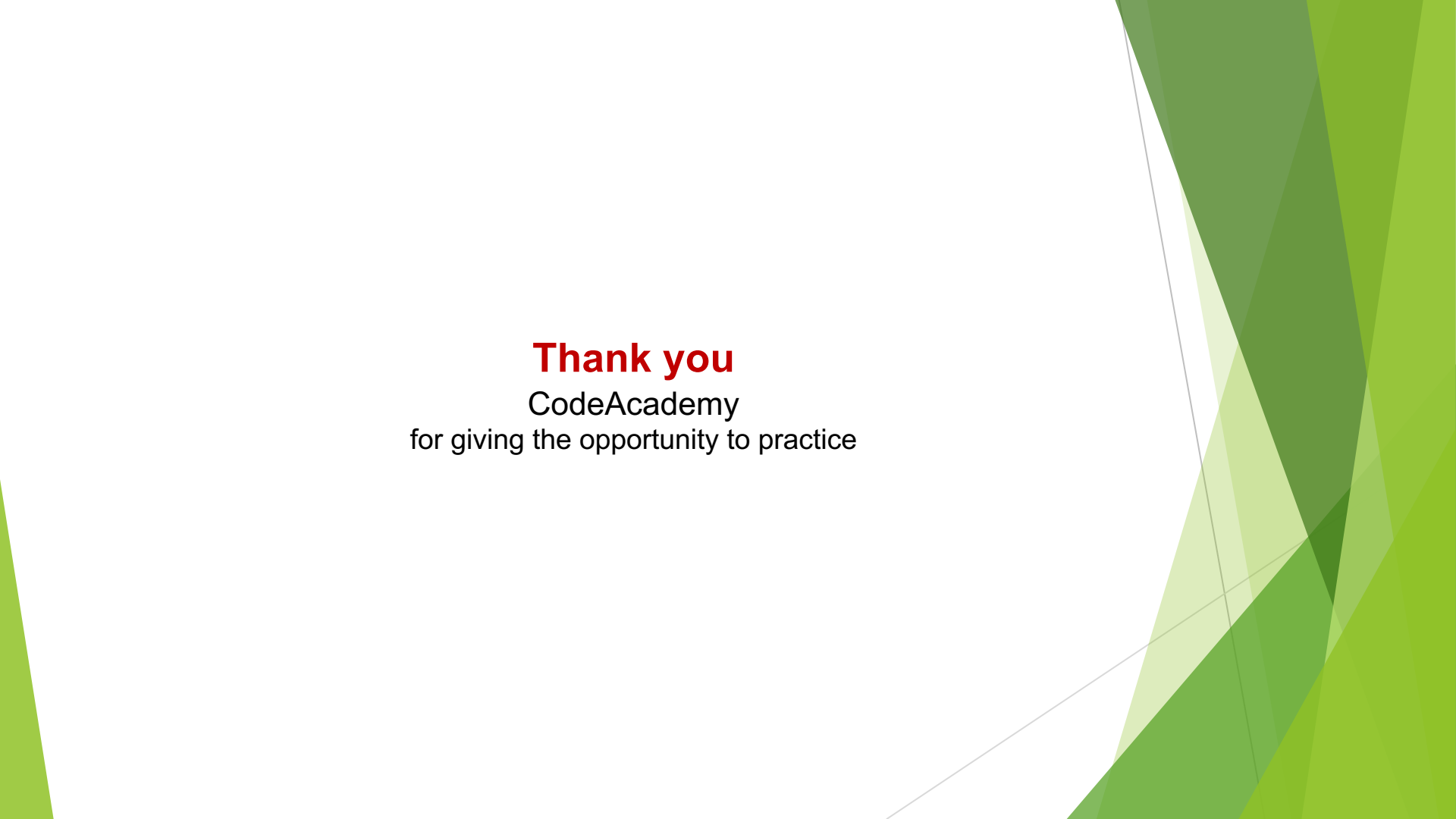
**Query Results**

| style | model_name | total_count |
|---|---|---|
| Women's Styles | Lucy | 252 |
| Men's Styles | Dawes | 243 |

```
-- query here

SELECT style, COUNT(*) AS
'common_style'
FROM quiz
GROUP BY style
ORDER BY 2 DESC;


-------------------------------------

SELECT style, model_name,
COUNT(model_name) AS 'total_count'
FROM purchase
GROUP BY style
ORDER BY 3 DESC;
```

**Thank you**
CodeAcademy
for giving the opportunity to practice