

WARBY  
PARKER

# Codecademy: Learn SQL From Scratch

Usage Funnels with Warby Parker

A Capstone Project by Owen Vanicek  
07/22/18

*Codecademy*

# Table of Contents

**“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.”**

## **What is the Quiz Funnel?**

- What columns does the table have?
- What is the number of responses for each question?
- Which question(s) of the quiz have a lower completion rate(s)?
  - What do you think is the reason?

## **A/B Testing with Home Try-On Funnel.**

- What are the column names?
- Combine the funnel and select only the first ten rows.
- What are some actionable insights for Warby Parker?

## **Big Picture**

Were users who  
received more pairs to  
try on home more likely  
to make a purchase?

# Quiz Funnel - Question 1: What columns does the table have?

```
1  -- Using the * and a LIMIT statement, the query
   returns all columns from only the first ten rows.
2  -- Q Answer: question, user_id, response
3
4  SELECT *
5  FROM survey
6  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	
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

Your company utilizes a style quiz in which users progress through a series of five questions that help them narrow in on styles of glasses they may like. This table is referred to as “survey.”

Utilizing a simple SQL query of your survey table, I can show a set of ten responses and each column in the table as requested.

## Quiz Funnel - Question 2: What is the number of responses for each question?

```
1  -- Using the COUNT, DISTINCT and GROUP BY statements, the
   query returns each question and the number of unique
   responses each question received.
2  -- Q Answer: Q1 - 500, Q2 - 475, Q3 - 380, Q4 - 361, Q5 -
   270
3
4  SELECT question,
5         COUNT(DISTINCT user_id) AS 'Responses'
6  FROM survey
7  GROUP BY 1;
```

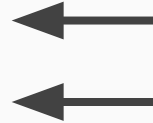
Query Results	
question	Responses
1. What are you looking for?	500
2. What's your fit?	475
3. Which shapes do you like?	380
4. Which colors do you like?	361
5. When was your last eye exam?	270

To further understand your company's style quiz, I've gone about querying the table to show us how many unique responses each question received.

This query shows us how many users move throughout the survey, how many "give up" and how many finish.

## Quiz Funnel - Question 3: Which question(s) of the quiz have lower completion rate(s)?

Warby Parker: Survey Question Completion Percentage			
Question	Responses	%	% Change
1. What are you looking for?	500	100%	0%
2. What's your fit?	475	95%	5%
3. Which shapes do you like?	380	76%	20%
4. Which colors do you like?	361	72%	5%
5. When was your last eye exam?	270	54%	25%



In copying the query results into excel, I did some further analysis to show what percentage of users completed each question. Furthermore, I showed the percent change from question to question giving some insight into which questions caused higher “give up” rates.

The data shows us that questions 3 and 5 resulted in significantly higher “give up” rates compared to questions 2 and 4.

## Quiz Funnel - Question 3 (cont'd): Which question(s) of the quiz have lower completion rate(s)?

Warby Parker: Survey Question Completion Percentage			
Question	Responses	%	% Change
1. What are you looking for?	500	100%	0%
2. What's your fit?	475	95%	5%
3. Which shapes do you like?	380	76%	20%
4. Which colors do you like?	361	72%	5%
5. When was your last eye exam?	270	54%	25%



### Why did this happen?

Looking at question 5 first, asking users when their last eye exam was in comparison to several preceding qualitative questions about style and color could catch people off guard. In fact, according to SurveyMonkey:

“Just like in conversations, when answering surveys, people don’t like to answer personal or sensitive questions. Because people feel uncomfortable or don’t want to give out much personal information, people tend to skip questions on sensitive topics or even drop out of the survey altogether.”

While not 100% positive, this could possibly explain the drop off that occurred during question 5.

Source: <https://www.surveymonkey.com/curiosity/sensitive-topics-methodology/>

## Quiz Funnel - Question 3 (cont'd): Which question(s) of the quiz have lower completion rate(s)?

Warby Parker: Survey Question Completion Percentage			
Question	Responses	%	% Change
1. What are you looking for?	500	100%	0%
2. What's your fit?	475	95%	5%
3. Which shapes do you like?	380	76%	20%
4. Which colors do you like?	361	72%	5%
5. When was your last eye exam?	270	54%	25%



### Why did this happen? (cont'd)

Now looking at question 3, this question is much harder to find fault in directly. In comparison to the other qualitative questions in the survey, it seems like it should fit right in. Given my experience working with surveys, much of my hypothesizing on why such a large drop off occurred at this question is almost purely empirical.

- Either the dip in responses was the fault of something in the UX (lack of a skip, not enough options...)
- Possibly directly related to human behavior (too lazy or impatient to complete it)
- Or it is a result of the question itself (poorly worded or doesn't generate enough enthusiasm)

Personally, I would say the lack of responses was probably both human behavior as well as the question. Perhaps some further testing on rewording that question with your marketing department is required.

# Shift into the A/B test

The Details

## Home Try-On Funnel

Take the Style Quiz → Home Try-On → Purchase

1000 users → 750 users → 495 users



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 (375 users)
- 50% of the users will get 5 pairs to try on (375 users)

## Big Picture

Were users who received more pairs to try on home more likely to make a purchase?



# Home Try-On Funnel - Question 4: What are the column names?

```
1  -- Using nearly the same syntax as we did in Q1, we can
   return all the columns for each table in the Home_Try_On
   Funnel.
2  -- Q Answer: For Quiz table: user_id, style, fit, shape,
   color
3  -- Q Answer: For Home_Try_On table: user_id,
   number_of_pairs, address
4  -- Q Answer: For Purchase table: user_id, product_id,
   style, model_name, color, price
5
6  SELECT *
7  FROM quiz
8  LIMIT 5;
9
10 SELECT *
11 FROM home_try_on
12 LIMIT 5;
13
14 SELECT *
15 FROM purchase
16 LIMIT 5;
```

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-a301-27dd93b12e2	Women's Styles	Narrow	Square	Two-Tone	
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black	
user_id	number_of_pairs		address		
d8add87-3217-4429-9a01-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-acc-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

Shifting our attention away from the Quiz funnel to the Home Try-On funnel. I ran a very familiar query to return the first five rows from each of our tables: (quiz, home\_try\_on, purchase).

This allows us to see the names of the columns (written out in the code) along with getting a brief overview of the data in each table.

# Home Try-On Funnel - Prompt 5: Combine the funnel and select only the first ten rows.

```
1  -- Using a series of LEFT JOINS, I have gone
   about combining each table on the primary key
   'user_id.' Progressively joining down the funnel
   prevents unwanted rows as well.
2  -- Results for 'is_home_try_on' and 'is_purchase'
   are shown in binary terms. 1 = True, 0 = False.
3  -- Prompt Results: Check Query Result
4
5  SELECT DISTINCT q.user_id,
6     h.user_id IS NOT NULL AS 'is_home_try_on',
7     h.number_of_pairs,
8     p.user_id IS NOT NULL AS 'is_purchase'
9  FROM quiz AS 'q'
10 LEFT JOIN home_try_on AS 'h'
11     ON q.user_id = h.user_id
12 LEFT JOIN purchase AS 'p'
13     ON p.user_id = q.user_id
14 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	∅	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

As explained in the code, I utilized several LEFT JOINS and a couple IS NOT NULL statements which return us a 1 if any given user\_id exists in the joined tables and a 0 if it does not.

With this query, we can now see clearly who tried on the glasses at home, how many pairs they had, and whether or not they purchased.

# Insights For Warby Parker

## Home Try-On Funnel - Prompt 6: Overall Conversion Rates

```
1 -- Continuing using our LEFT JOINS, I have gone about calculating overall conversion
  rates by aggregating across all rows.I have done by using WITH and SUM to aggregate
  across our previous table.
2 -- Results: quiz_to_home - 75% conversion rate, home_to_purchase - 66% conversion rate
3
4 With funnels AS (
5 SELECT DISTINCT q.user_id,
6     h.user_id IS NOT NULL AS 'is_home_try_on',
7     h.number_of_pairs,
8     p.user_id IS NOT NULL AS 'is_purchase'
9 FROM quiz AS 'q'
10 LEFT JOIN home_try_on AS 'h'
11     ON q.user_id = h.user_id
12 LEFT JOIN purchase AS 'p'
13     ON p.user_id = q.user_id)
14 SELECT COUNT(*) AS 'quiz_total',
15     SUM(is_home_try_on) AS 'num_home_try_on',
16     SUM(is_purchase) AS 'num_purchase',
17     ROUND(1.0 * SUM(is_home_try_on) / COUNT(user_id),2) AS 'quiz_to_home',
18     ROUND(1.0 * SUM(is_purchase) / SUM(is_home_try_on),2) AS 'home_to_purchase'
19 FROM funnels;
```

Query Results				
quiz_total	num_home_try_on	num_purchase	quiz_to_home	home_to_purchase
1000	750	495	0.75	0.66

As explained in the code, I maintained our previous JOINS and added a WITH and several SUM statements to aggregate across all rows.

This provides us with some overall conversion rates for this test. Overall, maintaining the small dropoff (9%) from quiz\_to\_home to home\_to\_purchase is fantastic. Even better is that 66% of those who tried the glasses on at home purchased. A conversion rate that high is phenomenal for any product!

## Home Try-On Funnel - Prompt 6: Conversion Rates Comparison

```
1 -- By adding a GROUP BY statement to our previous query, we can segment and get conversion
  rates for TEST A and TEST B to compare directly.
2 -- Results: reference the attached table.
3
4 With funnels AS (
5 SELECT DISTINCT q.user_id,
6     h.user_id IS NOT NULL AS 'is_home_try_on',
7     h.number_of_pairs,
8     p.user_id IS NOT NULL AS 'is_purchase'
9 FROM quiz AS 'q'
10 LEFT JOIN home_try_on AS 'h'
11     ON q.user_id = h.user_id
12 LEFT JOIN purchase AS 'p'
13     ON p.user_id = q.user_id)
14 SELECT COUNT(*) AS 'quiz_total',
15     SUM(is_home_try_on) AS 'num_home_try_on',
16     SUM(is_purchase) AS 'num_purchase',
17     ROUND(1.0 * SUM(is_home_try_on) / COUNT(user_id),2) AS 'quiz_to_home',
18     ROUND(1.0 * SUM(is_purchase) / SUM(is_home_try_on),2) AS 'home_to_purchase',
19     number_of_pairs
20 FROM funnels
21 GROUP BY number_of_pairs;
```

Query Results					
quiz_total	num_home_try_on	num_purchase	quiz_to_home	home_to_purchase	number_of_pairs
250	0	0	0.0	Ø	Ø
379	379	201	1.0	0.53	3 pairs
371	371	294	1.0	0.79	5 pairs

Here is where we can really begin to derive some insight for Warby Parker. By adding a GROUP BY statement to our previous query we can see directly which test, A or B, performed better.

From the query, we can see that we have a 53% conversion rate for home\_to\_purchase for 3 pairs (Test A) and a 79% conversion rate for 5 pairs (Test B).

Based on this test alone, we can see that 5 pairs yields higher conversion rates compared to three. But let's see if we can provide any further detail into why this was the case.

# Testing the Validity of the A/B Test

One of the biggest ways that A/B tests fail is that they end up testing multiple variables (even if not on purpose). Confounding variables can end up yielding a false positive result, often times tricking companies into believing their test was a success.

For Warby Parker, with many other variables like Gender, Style, Color, and Price, it would be a good business practice for any analyst to test those other variables for a potential impact on the A/B test.

The next slides contain many of the SQL syntax and their results I used to establish some numbers for my testing. My excel work containing my analysis showing their impact (or lack thereof) will follow those.

# Home Try-On Funnel - Prompt 6: A/B Test Validity

```
1  -- Here I am utilizing several queries to return some numbers
   based on the quiz.
2  -- My goal here was to begin testing to see if there were any
   other factors that may have influenced the A/B test.
3
4  SELECT style, COUNT(*)
5  FROM quiz
6  GROUP BY style;
7
8  SELECT fit, COUNT(*)
9  FROM quiz
10 GROUP BY fit;
11
12 SELECT shape, COUNT(*)
13 FROM quiz
14 GROUP BY shape;
15
16 SELECT color, COUNT(*)
17 FROM quiz
18 GROUP BY color;
```

Query Results	
style	COUNT(*)
I'm not sure. Let's skip it.	99
Men's Styles	432
Women's Styles	469
fit	COUNT(*)
I'm not sure. Let's skip it.	89
Medium	305
Narrow	408
Wide	198
shape	COUNT(*)
No Preference	97
Rectangular	397
Round	180
Square	326
color	COUNT(*)
Black	280
Crystal	210
Neutral	114
Tortoise	292
Two-Tone	104

# Home Try-On Funnel - Prompt 6: A/B Test Validity

```
1 -- Here I am utilizing several queries to return some numbers
  based on the purchase table.
2 -- My goal here was to begin testing to see if there were any
  other factors that may have influenced the A/B test.
3
4 SELECT product_id, COUNT(*)
5 FROM purchase
6 GROUP BY product_id;
7
8 SELECT style, COUNT(*)
9 FROM purchase
10 GROUP BY style;
11
12 SELECT model_name, COUNT(*)
13 FROM purchase
14 GROUP BY model_name;
15
16 SELECT color, COUNT(*)
17 FROM purchase
18 GROUP BY color;
```

Query Results	
product_id	COUNT(*)
1	52
2	43
3	63
4	44
5	41
6	50
7	44
8	42
9	54
10	62
style	COUNT(*)
Men's Styles	243
Women's Styles	252
model_name	COUNT(*)
Brady	95
Dawes	107
Eugene Narrow	116
Lucy	86
Monocle	41
Olive	50

color	COUNT(*)
Driftwood Fade	63
Elderflower Crystal	44
Endangered Tortoise	41
Jet Black	86
Layered Tortoise Matte	52
Pearled Tortoise	50
Rose Crystal	54
Rosewood Tortoise	62
Sea Glass Gray	43
price	COUNT(*)
50	41
95	261
150	193



## Home Try-On Funnel - Prompt 6: A/B Test Validity

```
1 -- Here I am checking the average price between both
  tests.
2 -- This confirms that there was a negligible price
  difference between test A and B showing that price did
  not impact the test.
3
4 SELECT AVG(price), number_of_pairs
5 FROM home_try_on
6 INNER JOIN purchase ON home_try_on.user_id =
  purchase.user_id
7 GROUP BY number_of_pairs;
```

Query Results	
AVG(price)	number_of_pairs
113.258706467662	3 pairs
112.34693877551	5 pairs

This is a test that I ran based on a simple SQL query to show the avg price between 3 and 5 pairs.

I wanted to see if there was potentially a significant difference in price between the two that caused more people to purchase. There was not, thus price was not a factor.

## Home Try-On Funnel - Prompt 6: A/B Test Validity

```
1  -- Here I am checking both average price between both
   tests as well as the revenues and genders.
2  -- The goal of this test was to compare several factors.
   We can see a lot of useful information from this query.
3
4  SELECT style, COUNT(*),
5         ROUND(SUM(price),2) AS 'Revenue',
6         ROUND(AVG(price),2) AS 'AVG_Price',
7         number_of_pairs
8  FROM purchase
9     INNER JOIN home_try_on
10        ON home_try_on.user_id = purchase.user_id
11  WHERE style LIKE 'Women%'
12  GROUP BY number_of_pairs;
13
14 SELECT style, COUNT(*),
15        ROUND(SUM(price),2) AS 'Revenue',
16        ROUND(AVG(price),2) AS 'AVG_Price',
17        number_of_pairs
18  FROM purchase
19     INNER JOIN home_try_on
20        ON home_try_on.user_id = purchase.user_id
21  WHERE style LIKE 'Men%'
22  GROUP BY number_of_pairs;
```

Query Results				
style	COUNT(*)	Revenue	AVG_Price	number_of_pairs
Women's Styles	98	10960.0	111.84	3 pairs
Women's Styles	154	17710.0	115.0	5 pairs
style	COUNT(*)	Revenue	AVG_Price	number_of_pairs
Men's Styles	103	11805.0	114.61	3 pairs
Men's Styles	140	15320.0	109.43	5 pairs

This is probably the most significant table in my search. It both shows the revenues between the two for both men and women. It shows the total products sold for both tests as well as compares Men and Women.

I use this table a lot in my excel analysis which follows. Check the next slide for that.

# Home Try-On Funnel - Prompt 6: A/B Test Validity

	product_id	COUNT(*)	
Layered Tortoise Matte	1	52 M	
Sea Glass Gray	2	43 M	
Driftwood Fade	3	63 M	
Elderflower Crystal	4	44 w	
Endangered Tortoise	5	41 M - Discont.	
Pearled Tortoise	6	50 w	
Jet Black	7	44 M	
Jet Black	8	42 w	
Rose Crystal	9	54 w	
Rosewood Tortoise	10	62 w	
	style	COUNT(*)	
	Men's Styles	243	
	Women's Styles	252	
	model_name	COUNT(*)	
	Brady	95 1,2	
	Dawes	107 3,7	
	Eugene Narrow	116 9,10	
	Lucy	86 4,8	
	Monocle	41 5	
	Olive	50 6	
	color	COUNT(*)	
	Driftwood Fade	63	
	Elderflower Crystal	44	
	Endangered Tortoise	41	
	Jet Black	86	
	Layered Tortoise Matte	52	
	Pearled Tortoise	50	
	Rose Crystal	54	
	Rosewood Tortoise	62	
	Sea Glass Gray	43	
	price	COUNT(*)	
	50	41 Monocle	Men
	95	261	
	150	193	

So let me breakdown this table to the left. I went through using my previous queries and determined which style equated to which pair as well as what colors they were and to which gender they were ascribed to.

What I was able to determine was that there was no significant impact on the A/B test for Gender, Price, Color, or Style, thus the A/B test was a success and is valid.

I was able to reveal something interesting however. The average price for men's glasses for 5 pairs was lower than that for women. Why? The Monocle. The monocle was only \$50 and was only sold specifically to men. Which given its novelty made it a great addition for those who were able to try on 5 pairs. Given the low sales of the monocle and some research of the product itself, it makes sense why Warby Parker discontinued the item.

# Final Recommendation

According to the results of the test, sending 5 pairs for potential customers to try on yielded more sales.

Gender, Price, Color, and Style had an insignificant impact on the test.

Low priced novelty items caused the lower average sales for men who received 5 pairs.

Overall, my recommendation would be to remove those low priced items and instead include some higher priced specialty or novelty glasses. If people are searching for novelty (which 41 customers were) why not put a high price tag on it and squeeze out extra margins?