code|cademy

WARBY PARKER
eyewear

Optimizing Your Funnels with SQL

Any fool can complicate things. It takes genius to simplify them.

Albert Einstein

CONTENTS

—

—

A random walk thru
The Data

The Survey Table

Three fields with text data:
- Question
- User_id
- Response

The survey table contains 5 survey questions along with 1986 responses from 500 participants. The table below shows that although 500 participates started the survey, only 270 finished it. 230 participants quit during the 5 question quiz.

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

Q 1: What columns does the [survey] table have?

```
SELECT *
FROM survey
LIMIT 10
;
```

Q 2: What is the number of responses for each question in the quiz?

```
SELECT question,
    COUNT (user_id) AS "User IDs",
    COUNT (response) AS "Responses"

FROM survey

GROUP BY question
;

SELECT COUNT(DISTINCT user_id)
FROM survey
;
```

Drop Offs: Look at the Completion Rates.

When you look at completion rates for each question the differences are not as pronounced as when you convert the same metric to a drop off rate. Question 5 has the lowest completion rate, which is sad for the 91 people who came so close to finishing but didn't!

The second significant drop off occurs for question 3; 98 participants quit on this question. Almost 40% (37.8%) of participants end the survey on these two questions. Why is that?

Q 3.a: Which questions of the quiz have a lower completion rate?

```
-- No SQL for this question
```

Questions	Unique Users	% of Users	Completion Rate	Drop Off
1. What are you looking for?	500	100%	n/a	n/a
2. What's your fit?	475	95%	95%	5%
3. Which shapes do you like?	380	76%	80%	20%
4. Which colors do you like?	361	72%	95%	5%
5. When was your last eye exam?	270	54%	75%	25%

The Style Quiz Funnel

What are you looking for?

The most common results from the Style Quiz.

Men's Styles

Women's Styles

I'm not sure. Let's skip it.

What are you looking for?

```
SELECT question AS Question,
  response AS Response,
  COUNT (response) AS "Response
Count"

FROM survey

WHERE Question like "1.%"
GROUP BY response
ORDER BY "Response Count" DESC
;
```

242

Men's Styles

209

Women's Styles

I'm not sure. Let's skip it. 49 (9.8%)

What's your fit?

Please choose one

Narrow

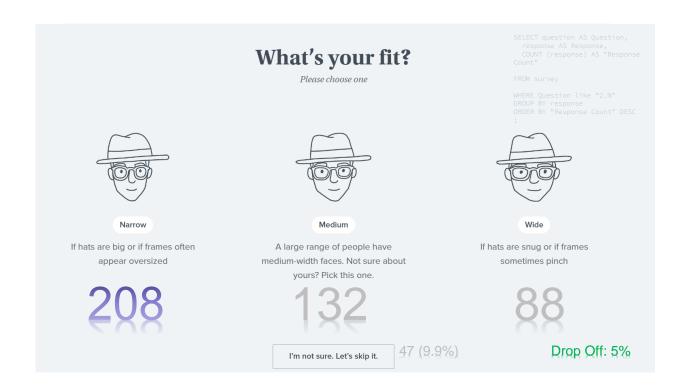If hats are big or if frames often appear oversized

Medium

A large range of people have medium-width faces. Not sure about yours? Pick this one.

Wide

If hats are snug or if frames sometimes pinch

I'm not sure. Let's skip it.

What's your fit?

Please choose one

SELECT question AS Question,
  response AS Response,
  COUNT (response) AS "Response
Count"

FROM survey

WHERE Question like "2.%"
GROUP BY response
ORDER BY "Response Count" DESC
;

Narrow

If hats are big or if frames often appear oversized

208

Medium

A large range of people have medium-width faces. Not sure about yours? Pick this one.
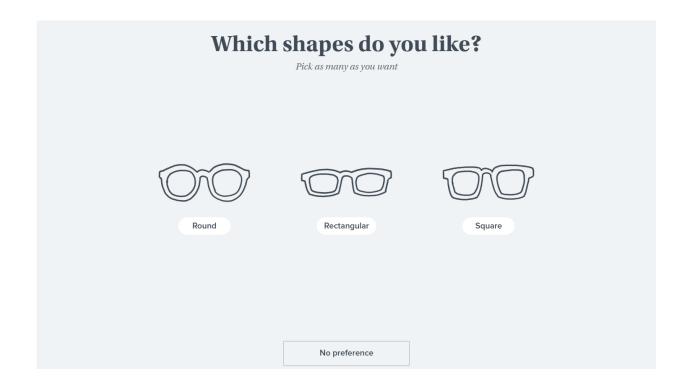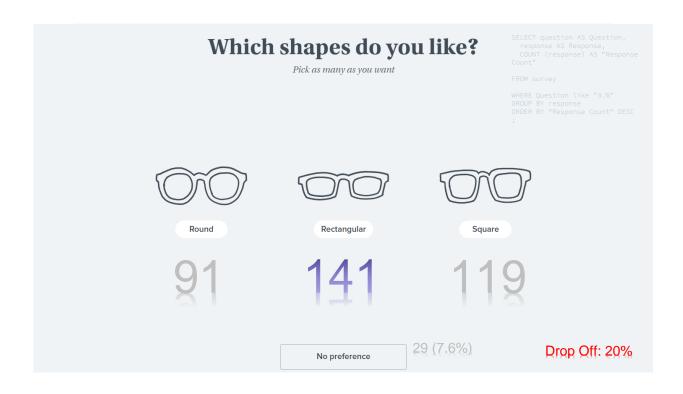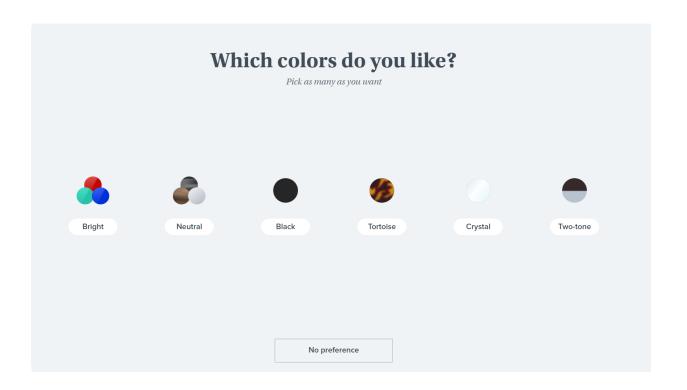
132

Wide

If hats are snug or if frames sometimes pinch

88

I'm not sure. Let's skip it.

47 (9.9%)

Drop Off: 5%

Which shapes do you like?

Pick as many as you want

Round

Rectangular

Square

No preference

Which shapes do you like?

Pick as many as you want

SELECT question AS Question,
  response AS Response,
  COUNT (response) AS "Response
Count"

FROM survey

WHERE Question like "3.%"
GROUP BY response
ORDER BY "Response Count" DESC
;

Round

Rectangular

Square

91

141

119

No preference

29 (7.6%)

Drop Off: 20%

Which colors do you like?

Pick as many as you want

Bright

Neutral

Black

Tortoise

Crystal

Two-tone

No preference

Which colors do you like?

Pick as many as you want

Despite this direction, no participants choose more than one color. I took the quiz myself, and missed this direction as well.

```
SELECT question AS Question,
    response AS Response,
    COUNT (response) AS "Response Count",
    COUNT(user_id)

FROM survey

WHERE Question like "4.%"
GROUP BY response
ORDER BY "Response Count" DESC
;

SELECT user_id,
    COUNT(response) AS "RES"

FROM survey

WHERE question like "4.%"

GROUP BY user_id
ORDER BY 2 DESC
;
```

Bright	Neutral	Black	Tortoise	Crystal	Two-tone
0	36	112	117	69	27

Why? Is it because they are cheap? I don't think so!

No preference   0 (0%)

Drop Off: 5%

When was your last eye exam?

Please choose one

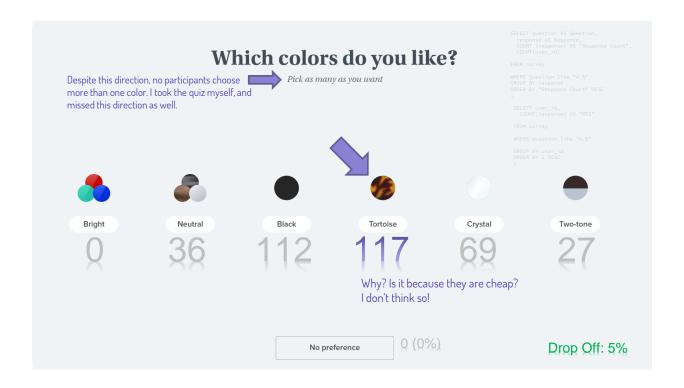< 1 Years   1 - 3 Years   3+ Years
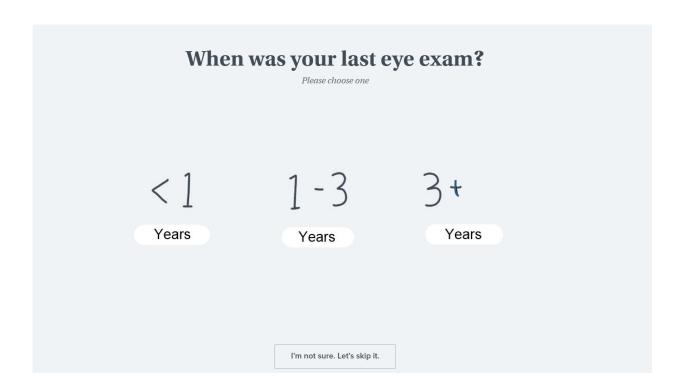
I'm not sure. Let's skip it.

When was your last eye exam?

Please choose one

```
SELECT question AS Question,
  response AS Response,
  COUNT (response) AS "Response
Count",
  COUNT (user_id)

FROM survey

WHERE Question like "5.%"
GROUP BY response
ORDER BY "Response Count" DESC
;
```

< 1
Years
141

1 - 3
Years
56

3+
Years
37

I'm not sure. Let's skip it.   36 (10%)

Drop Off: 25%

The Survey Table

On question 3, about shape, we see a significant drop off with only 7% selecting no preference (avg. no preference selection is around 10%). I thought selecting "No preference" might shed light on survey fatigue, No luck.
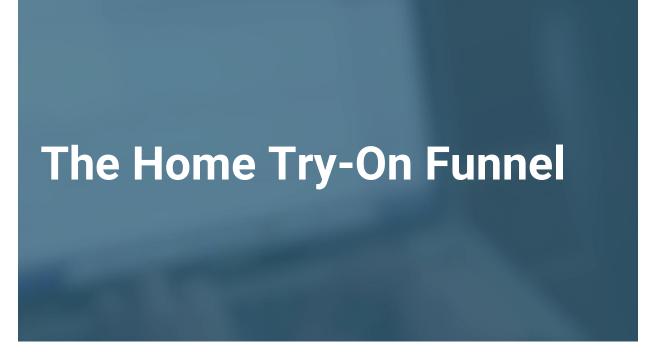
On question 5, about the last eye exam, there is significant drop off as well. Most respondents answered less than a year. That isn't hard to remember. A case can be made for fatigue.

I don't see a clear reason for why. I would test the order of the questions. Possibly, these question are more difficult to answer. Maybe users start to fatigue on the third question. The colors question showed strong participation. Try using that as the third question. I strongly recommend testing the order of questions.

Questions	Unique Users	% of Users	Completion Rate	Drop Off
1. What are you looking for?	500	100%	n/a	n/a
2. What's your fit?	475	95%	95%	5%
3. Which shapes do you like?	380	76%	80%	20%
4. Which colors do you like?	361	72%	95%	5%
5. When was your last eye exam?	270	54%	75%	25%

Q 3.b: Why?

-- No SQL for this question

The Home Try-On Funnel

The Home Try-On Tables

This data funnel pulls from data organized into three tables.

The quiz table has five fields with text data: user_id, style, fit, shape, & color.

The home_try_on table has three fields with text data: user_id, number_of_pairs, & address.

The purchase table has six fields with text and integer data: user_id, product_id, style, model_name, color, price.

Note that below is the schema, not the results of the query. The results were not material to this question.

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

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

Q 4: What are the column names [of each table]?

```
SELECT *
FROM quiz
LIMIT 5
;

SELECT *
FROM home_try_on
LIMIT 5
;

SELECT *
FROM purchase
LIMIT 5
;
```

Overall Conversion Rates

Of those users that took the quiz, 75% moved on to the home trial.
Of those users that completed the home trial, 66% made a purchase. This 66% is misleading. We A/B tested sending out 5 trail pairs of glasses VS 3 pairs. The results are unmistakable.

	Quiz	Home Trial	Purchase
Unique Users	1000	750	495
Conversion Rate	n/a	75%	66%

Q 6.a: calculate overall conversion rates by aggregating across all rows and compare conversions.

```
WITH base_table AS(
SELECT q.user_id,
  h.user_id IS NOT NULL AS
'is_home_try_on',
  h.number_of_pairs AS 'AB_TEST',
  b.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 b -- 'b' for buy. p
is too close to q
  ON q.user_id = b.user_id
)
SELECT
  --AB_TEST,
  COUNT(user_id) as "Quiz",
  SUM(CASE WHEN is_home_try_on = 1 THEN
1 ELSE 0 END) as "Home Trial",
  SUM(CASE WHEN is_purchase = 1 THEN 1
ELSE 0 END) as     "Purchase"

 FROM base_table
;
```

What are some
actionable insights
For Warby Parker ?

Five Pairs Drive More Sales!

At almost 80% sales conversion, offering 5 trail pairs clearly drives more sales than the standard offering of 3 trial pairs. The results of the A/B test demonstrate this outcome with a 1% margin of error.

Offer 5 trail pairs.

A/B Test	Home Trial	Purchase	Conversion
3 pairs	379	201	53.0%
5 pairs	371	294	79.2%
All	750	495	66%

This is significant at a 99% confidence interval.

Q 6.a: calculate the difference in purchase rates between customers who had 3 "number_of_pairs" with ones who had 5.

```sql
WITH base_table AS(
SELECT q.user_id,
  h.user_id IS NOT NULL AS 'is_home_try_on',
  h.number_of_pairs AS 'AB_TEST',
  b.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 b -- 'b' for buy. p is
too close to q
          ON q.user_id = b.user_id
)

SELECT
    AB_TEST,
    COUNT(user_id) as "Quiz",
    SUM(CASE WHEN is_home_try_on = 1 THEN 1
ELSE 0 END) as "Home Trial",
    SUM(CASE WHEN is_purchase = 1 THEN 1 ELSE
0 END) as    "Purchase"

 FROM base_table

 GROUP BY AB_TEST
 HAVING "Home Trial" > 0
 ORDER BY AB_TEST
 ;
```

Tortoise IS the new black.

For women, there is a clear preference for the Eugene Narrow Rosewood Tortoise over the control pair (Lucy Jet Black) with a 5% margin of error. The same frames in the rosewood crystal color are the second most popular, besting the control with a 10% margin of error.

Tortoise and black colors are the most popular for both men and women. The survey shows a measurable statistically significant preference for these colors compared to the fourth place "Neutral" color. Women preferred tortoise to black with a 10% margin of error.

```sql
SELECT model_name AS "Model Name",
    COUNT (user_id) AS Total,
    SUM(CASE WHEN color = "Elderflower
Crystal" then 1 else 0 END) AS "Elderflower
Crystal",
    SUM (CASE WHEN color = "Jet Black" AND
style = "Women's Styles" THEN 1 ELSE 0 END)
AS "Jet Black (W)",
    SUM (CASE WHEN color = "Pearled
Tortoise" THEN 1 ELSE 0 END) AS "Perled
Tortoise",
    SUM (CASE WHEN color = "Rose Crystal"
THEN 1 ELSE 0 END) AS "Rose Crystal",
    SUM (CASE WHEN color = "Rosewood
Tortoise" THEN 1 ELSE 0 END) AS "Rosewood
Tortoise"

FROM purchase

GROUP BY model_name
ORDER BY 2 DESC

 ;
```

Model Name	Total	Elderflower Crystal	Jet Black (W)	Pearled Tortoise	Rose Crystal	Rosewood Tortoise
Eugene Narrow	116	0	0	0	54	62
Dawes	107	0	0	0	0	0
Brady	95	0	0	0	0	0
Lucy	86	44	42	0	0	0
Olive	50	0	0	50	0	0
Monocle	41	0	0	0	0	0

Men are Indiferent . . . for now.

For the men, the most purchased item is the Dawes Driftwood Fade glasses. There is a measurable preference for these glasses compared to a control group of glasses. Brady Layered Tortoise placed second and was statistically indistinguishable from the control group or the Dawes. This matches our expectations. Men's preference for tortoise over black was slightly higher but indistinguishable with any meaningful measure of confidence.

Pay attention to this trend, we might still be in the early stages.

```
SELECT model_name AS "Model Name",
    COUNT (user_id) AS Total,
    SUM(CASE WHEN color = "Driftwood Fade"
THEN 1 ELSE 0 END) AS Driftwood,
    SUM(CASE WHEN color = "Endangered
Tortoise" THEN 1 ELSE 0 END) AS "Endangered
Tortoise",
    SUM(CASE WHEN color = "Jet Black" AND
style = "Men's Styles" THEN 1 ELSE 0 END)
AS "Jet Black (M)",
    SUM(CASE WHEN color = "Layered Tortoise
Matte" THEN 1 ELSE 0 END) AS "Layered
Tortoise Matte",
    SUM(CASE WHEN color = "Sea Glass Gray"
THEN 1 ELSE 0 END) AS "Sea Glass Gray"

FROM purchase

GROUP BY 1
ORDER BY 2 DESC

;
```

Model Name	Total	Driftwood	Endangered Tortoise	Jet Black (M)	Layered Tortoise Matte	Sea Glass Gray
Eugene Narrow	116	0	0	0	0	0
Dawes	107	63	0	44	0	0
Brady	95	0	0	0	52	43
Lucy	86	0	0	0	0	0
Olive	50	0	0	0	0	0
Monocle	41	0	41	0	0	0

Take-aways

1. Offer 5 pairs of glasses for the home trial (Slide 23).

2. Raise prices for Tortoise products in the women's style lines. Early adopters will not be price sensitive.

3. Monitor Tortoise products in the men's lines. Test prices. Early adopters will not be price sensitive.

4. Do some testing with the order of questions in the survey quiz.

5. The survey identifies that participants can identify more than one color, but none did. This is an issue with the design of the quiz.