codecademy

Warby Parker

Funnels Usage – SQL and Excel Project

Fırat Olçum

28-07-2022

Table of Contents

1. What is Warby Parker

Warby Parker is a socially conscious company that operates with the objective to offer designer eyewear at a revolutionary price.

When costumers interact with Warby's website, they are first encouraged to fill out a quiz to help them narrow down the options.After completing the survey, they can select a set number of glasses to try at home before deciding whether they want to make a purchase.

We will explore the data collected along the customer journey to retrieve insight that addresses the company's objective.

2. Quiz Funnel

2.1 Columns on the Survey table

- The Survey table consist of the three columns: question, user_id and response.

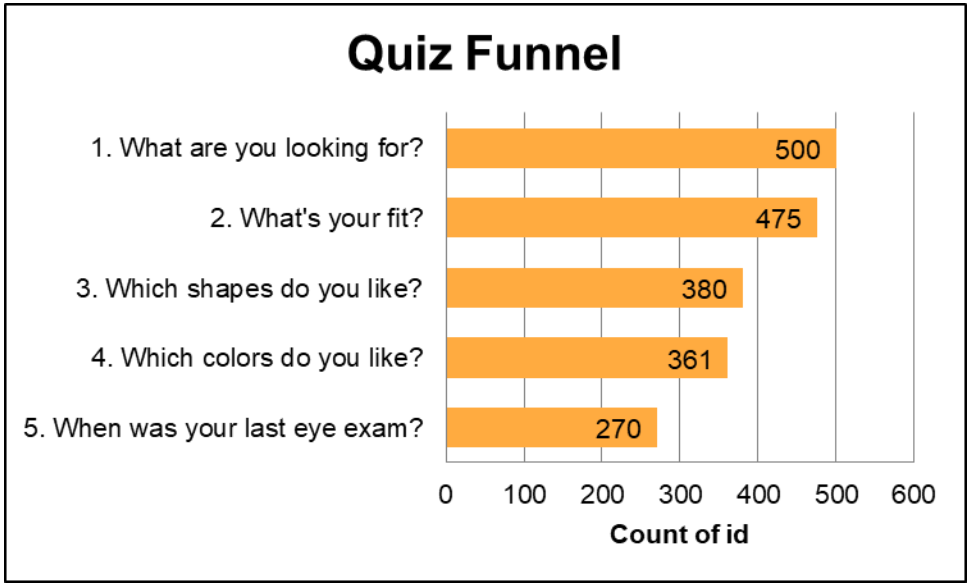
- The quiz column consists of the following questions:

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?

- As you can see from the table on the right, for every question answered we know the user id and their response.

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 for each question




```
--SQL Code

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

Using GROUP BY statement we organize the data to COUNT the number of users that answered each question. The code on the right produces the numbers in the chart on the left. As you can see from the chart, the number of respondents is decreasing for each question.

2.3 Question of the quiz with a lower completion rate

question	NumberOf Responses	PriorNumber OfResponses	PercentComplete ThisQuestion
1. What are you looking for?	500		
2. What's your fit?	475	500	95.0
3. Which shapes do you like?	380	475	80.0
4. Which colors do you like?	361	380	95.0
5. When was your last eye exam?	270	361	75.0

Looking at the table on left, we can see that questions 1, 2 and 4 had the highest answer rate, followed by question 3 and 5 with the lowest ones.

```
--SQL Code

WITH temp_table AS
(
SELECT DISTINCT question,
                COUNT( user_id) OVER(PARTITION BY
question) NumberOfResponses
FROM survey
),
temp_table2 AS
(
SELECT *, LAG(NumberOfResponses) OVER()
PriorNumberOfResponses
FROM temp_table
)
SELECT *, ROUND((100.0 * NumberOfResponses /
PriorNumberOfResponses), 0) AS
PercentCompleteThisQuestion
FROM temp_table2
```

2.4 Possible reasons behind lower completion rates

- It is interesting to note that each question has a neutral or skip option as an answer, which means that users can always move on to the next question without making a concrete choice as opposed to not completing the quiz all together.
- Even given the option to skip, almost half the users still decided to discontinue the quiz on questions 3 and 5, which directs us to having a closer look at the questions.
- Despite also being a question about the user's preference, question 3 'Which shapes do you like' has a far larger response rate drop than question 4 'Which colors do you like?'.
- This might be because we are more likely to already have a favorite color that we would choose as our answer. Whereas the shape of glasses is more complex and usually requires previous try on sessions in order to know one's preference. A first time purchaser would not have such exposure for example.
- Like question 2 'What's your fit?', question 5 'When was your last eye exam?' is measurable and objective. However, the latter has a far larger response rate drop. Is the issue also within the question itself? Or could we attribute it to the fact that it's the final question of the quiz funnel, and therefore that a large drop would be expected?

3. Home Try-On Funnel

3.1 Looking at the data (limit to 3 rows)

quiz table

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

home_try_on table

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

purchase table

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

3.1 home_try_on_funnel table

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	N/A	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	N/A	False

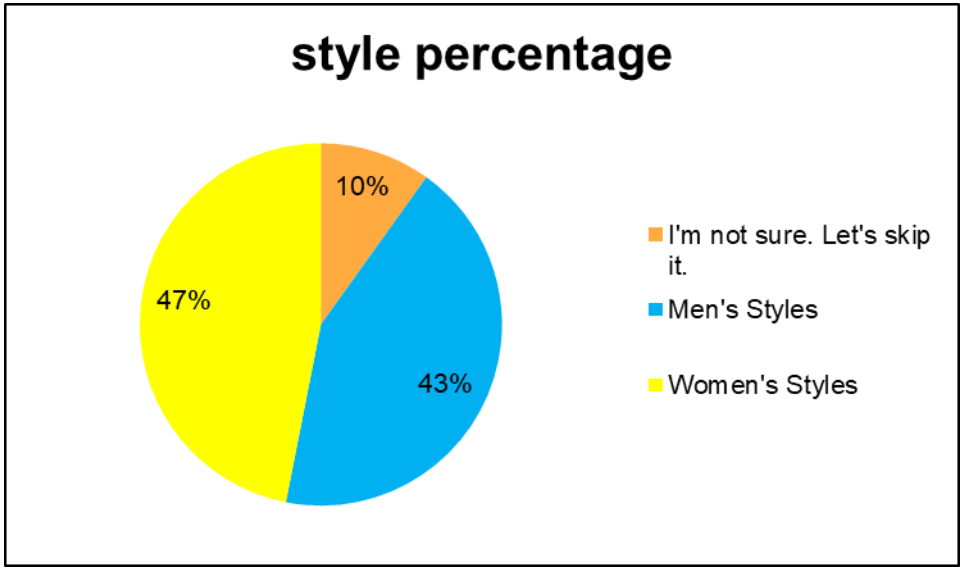
```
--SQL Code

WITH home_try_on_funnel AS
(
SELECT DISTINCT q.user_id,
       CASE WHEN h.user_id IS NOT NULL THEN 'True'
ELSE 'False' END AS is_home_try_on,
       CASE WHEN h.number_of_pairs IS NULL  THEN
'N/A' ELSE h.number_of_pairs END AS
number_of_pairs,
       CASE WHEN p.user_id IS NOT NULL THEN 'True'
ELSE 'False' END AS is_purchase
FROM quiz q
LEFT JOIN home_try_on h
ON q.user_id = h.user_id
LEFT JOIN purchase p
ON q.user_id = p.user_id
)
SELECT *
FROM home_try_on_funnel
LIMIT 7;
```

Using a WITH clause and two LEFT JOINs we can combine all three tables into one.This table includes the id of all the users who completed the quiz, as well as whether they tried-on glasses at home, received 3, 5 or no pairs, and made a purchase.

4. Actionable Insights

4.1.1 quiz table style column

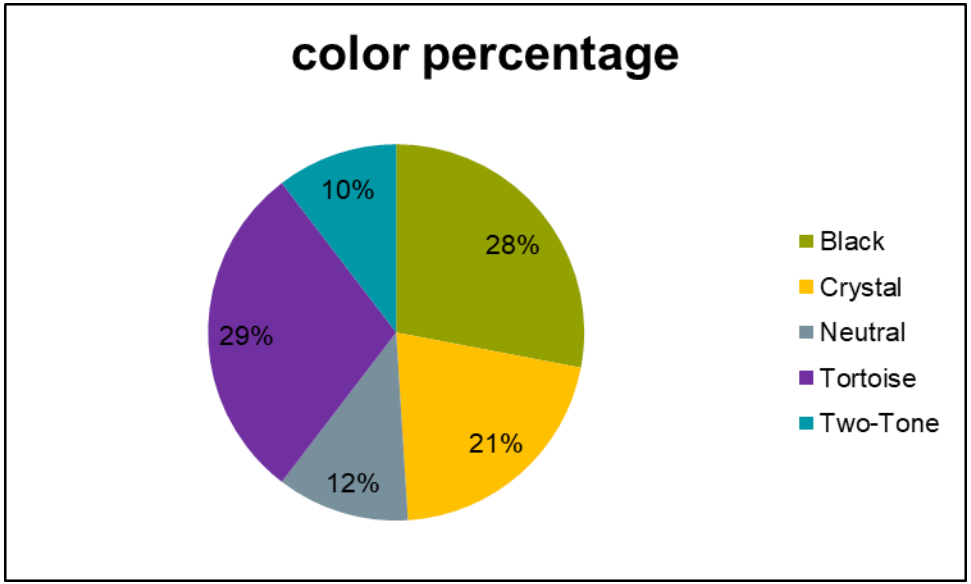



```
--SQL Code

WITH temp_table AS
(
SELECT DISTINCT style,
       COUNT(user_id) OVER(PARTITION BY style) AS
style_counts
FROM quiz
ORDER BY style_counts DESC
)
SELECT style,
       100.0 * style_counts /SUM(style_counts)
OVER() AS style_percentage
FROM temp_Table
```

You can see the style preferences of the participants participating in the quiz as a percentage. While 47% of the participants chose woman styles, 43% chose men styles. and 10% are undecided. The 'unisex' option can be included in the quiz as a suitable style for this 10%.

4.1.2 quiz table style column

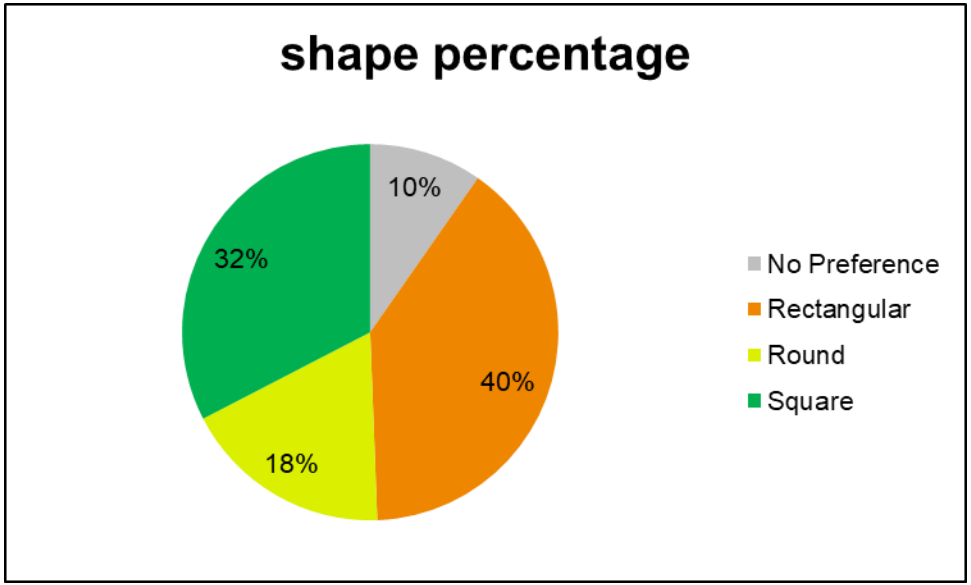



```
--SQL Code

WITH temp_table AS
(
SELECT DISTINCT color,
       COUNT(user_id) OVER(PARTITION BY color) AS
color_counts
FROM quiz
ORDER BY color_counts DESC
)
SELECT color,
       100.0 * color_counts /SUM(color_counts)
OVER() color_percentage
FROM temp_table
```

You can see the color preferences of the participants participating in the quiz as a percentage. We see that the most preferred colors are tortoise and black.

4.1.3 quiz table shape column

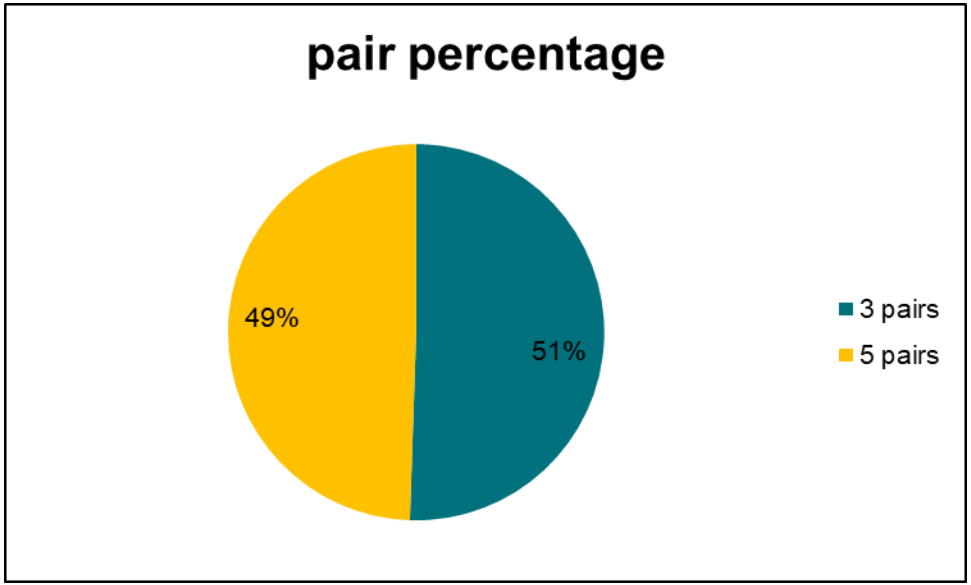



```
--SQL Code

WITH temp_table AS
(
SELECT DISTINCT shape,
       COUNT(user_id) OVER(PARTITION BY shape) AS
shape_counts
FROM quiz
ORDER BY shape_counts DESC
)
SELECT shape,
       100.0 * shape_counts /SUM(shape_counts)
OVER() shape_percentage
FROM temp_table
```

You can see the shape preferences of the participants participating in the quiz as a percentage. We see that the most preferred shapes are rectangular and square. 10% of the participants did not make any choice. Other option can be put here.

4.2.1 home_try_on table number_of_pairs column

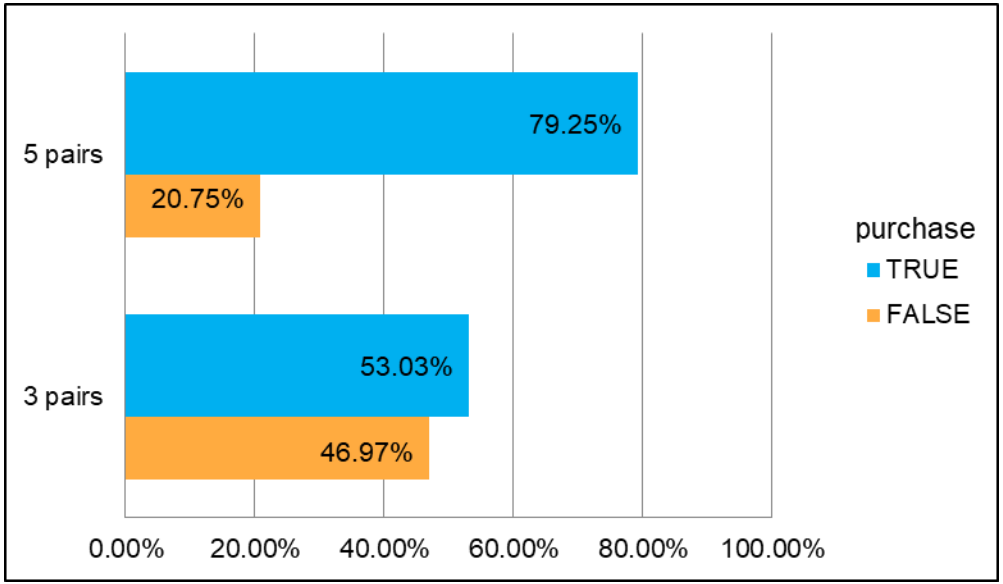



```
--SQL Code

WITH temp_table AS
(
SELECT number_of_pairs,
       COUNT(DISTINCT user_id) pair_counts
FROM home_try_on
GROUP BY number_of_pairs
)
SELECT number_of_pairs,
       100.0 * pair_counts / SUM(pair_counts)
OVER() AS pair_percentage
FROM temp_table
```

You can see the shape preferences of the participants participating in the home_try_on as a percentage. There is no distinguishing difference between the participants who tried 3 pairs and 5 pairs of glasses at home. It is not possible to draw an insight from this chart.

4.2.2 home_try_on table purchase insight

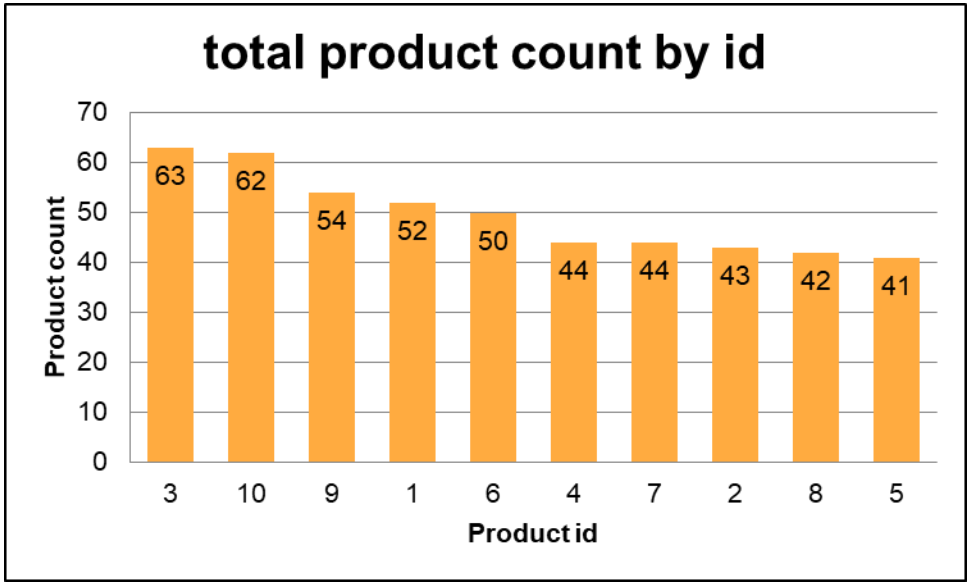



```
--SQL Code

WITH home_try_on_funnel AS
(
SELECT DISTINCT q.user_id, h.number_of_pairs,
       CASE  WHEN p.user_id IS NOT NULL THEN 'True'
             ELSE 'False' END AS is_purchase
FROM quiz q
LEFT JOIN home_try_on h
ON q.user_id = h.user_id
LEFT JOIN purchase p
ON q.user_id = p.user_id
)
SELECT number_of_pairs, is_purchase,
COUNT(is_purchase) AS TotalCountOfPurchase
FROM home_try_on_funnel
GROUP BY  number_of_pairs, is_purchase
HAVING number_of_pairs IS NOT NULL;
```

You can see the purchase rates of the products tried as 3 pairs and 5 pairs. 26% more of the users that received 5 pairs of glasses completed a purchase compared to the users that only received 3 pairs.

4.3.1 purchase table product_id column

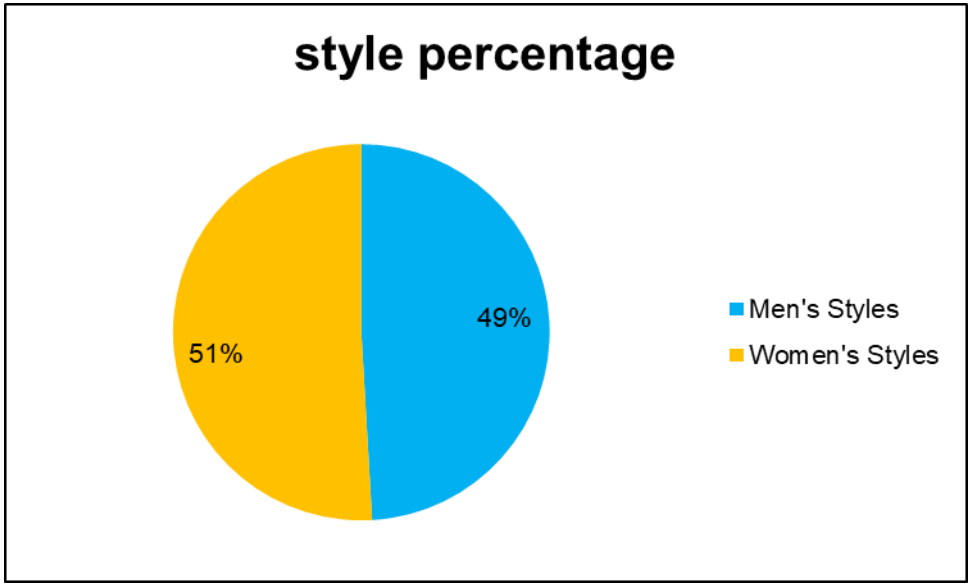



```
--SQL Code

SELECT product_id,
                COUNT(DISTINCT user_id ) AS
product_count
FROM purchase
GROUP BY product_id
ORDER BY product_id
```

You can see the product preferences of the participants participating in the purchase as a count. Products with 3 and 10 prdouct ids were sold more than others.

4.3.2 purchase table style column

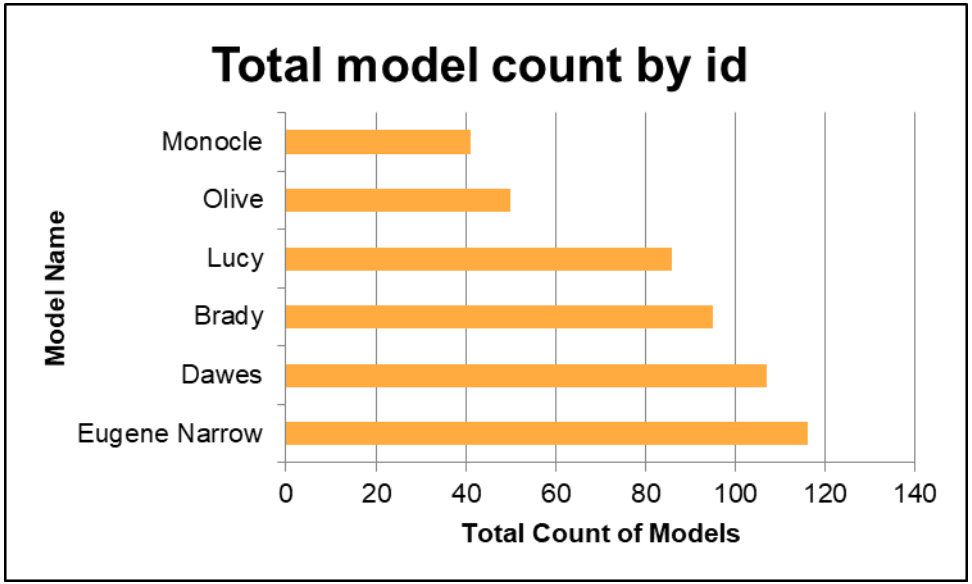



```
--SQL Code

SELECT style,
       COUNT(DISTINCT user_id) AS style_count
FROM purchase
GROUP BY style
```

You can see the style preferences of the participants participating in the purchase as a percentage. While 51% of the customers preferred women's styles, 49% preferred men's styles. We can say that there is an equal distribution in terms of gender.

4.3.3 purchase table model_name column

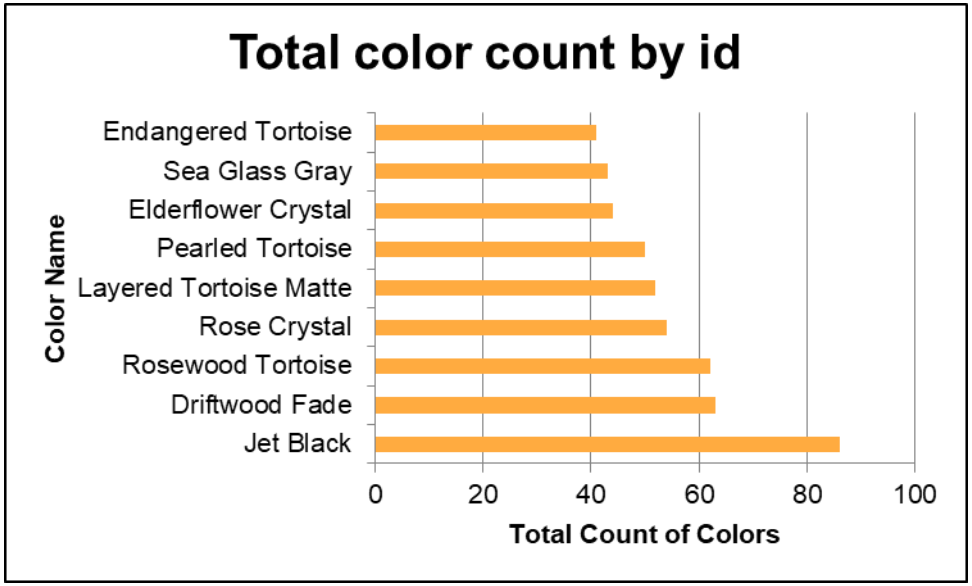



```
--SQL Code

SELECT model_name,
       COUNT(DISTINCT user_id) AS model_count
FROM purchase
GROUP BY model_name
```

You can see the model preferences of the participants participating in the purchase as a count. We see that the most preferred models by customers are ‘Eugine Narrow’ and ‘Dawes’ models. The least preferred are ‘Monocle’ and ‘Olive’. Discounts on these two models can increase their sales.

4.3.4 purchase table color column

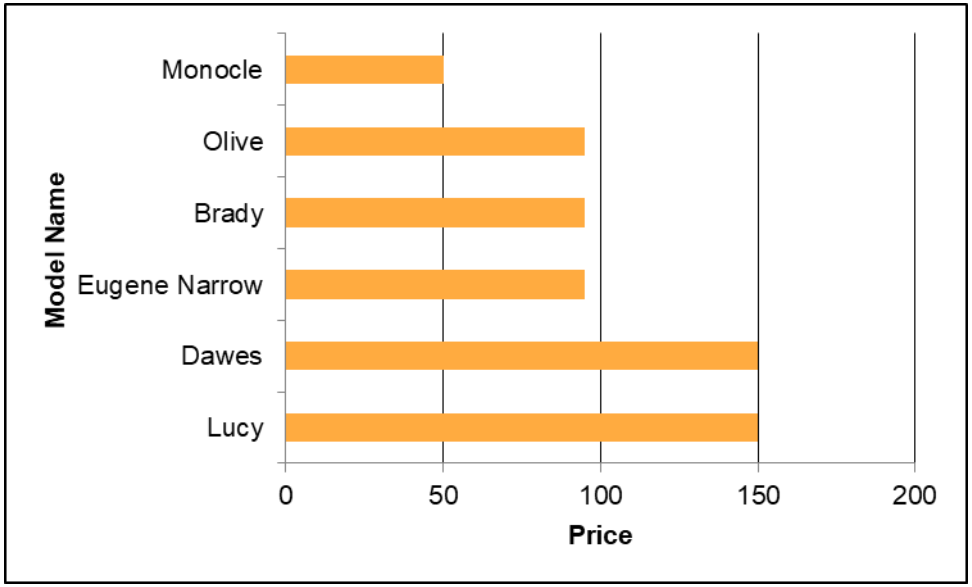



```
--SQL Code

SELECT color,
       COUNT(DISTINCT user_id) AS color_count
FROM purchase
GROUP BY color
```

You can see the color preferences of the participants participating in the purchase as a count. We see that the most preferred color by customers is ‘Jet Black’. If you produce other models similar to this model, we can assume that they will sell well.

4.3.5 purchase table price column

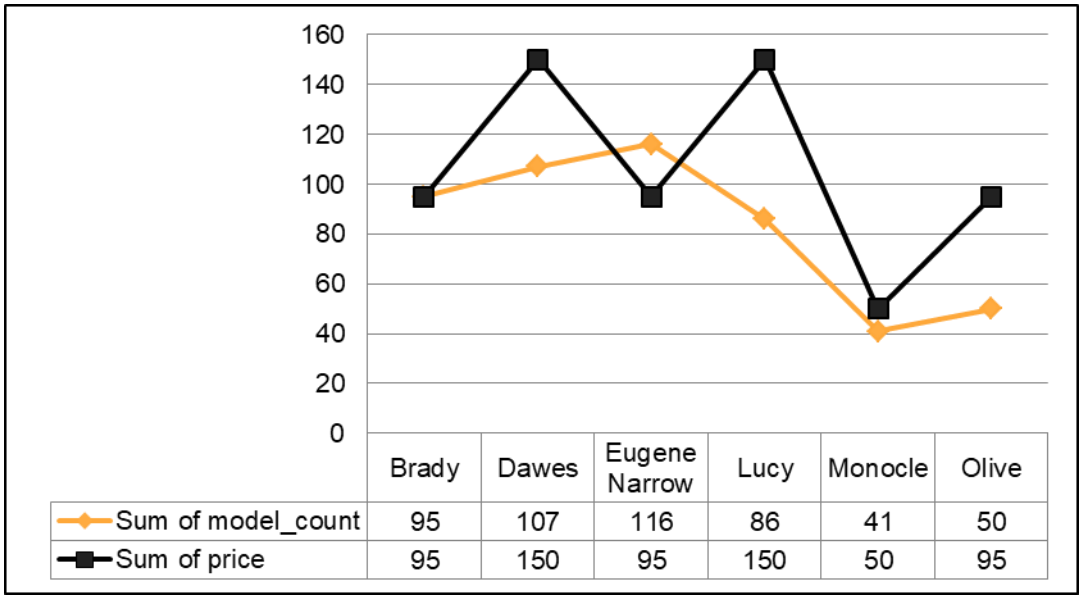



```
--SQL Code

SELECT DISTINCT(model_name), price
FROM purchase
ORDER BY price DESC
```

You can see the prices according to the names of the models. The two most expensive models are the ‘Dawes’ and ‘Lucy’.

4.3.6 purchase table price column



	Brady	Dawes	Eugene Narrow	Lucy	Monocle	Olive
Sum of model_count	95	107	116	86	41	50
Sum of price	95	150	95	150	50	95

```
--SQL Code

SELECT DISTINCT(model_name),
       COUNT(user_id) AS model_count, price
FROM purchase
GROUP BY model_name
ORDER BY price DESC, model_count DESC
```

You can see the Dawes model being the best selling model despite being the most expensive model. Customers who bought this model did not pay much attention to the price. The price of this product can be increased. In addition, although the monocle model is the cheapest model, it is the least sold model. The design of this model can be changed.

4.3.7 funnel insights

quiz_to_hometry	hometry_to_purchase	quiz_to_purchase
75%	66%	49.50%
quiz_id_count	home_try_id_count	purchase_id_count
1000	750	495

```
--SQL Code

WITH T1 AS(
SELECT COUNT(DISTINCT q.user_id) AS quiz_id_count,
       COUNT(DISTINCT h.user_id) AS home_try_id_count              ,
       COUNT(DISTINCT p.user_id) AS purchase_id_count
FROM quiz q
LEFT JOIN home_try_on h
ON q.user_id = h.user_id
LEFT JOIN purchase p
ON q.user_id = p.user_id
)
SELECT
quiz_id_count, home_try_id_count, purchase_id_count,
(100.0 * home_try_id_count / quiz_id_count) AS quiz_to_hometry,
(100.0 * purchase_id_count / home_try_id_count) AS hometry_to_purchase,
(100.0 * purchase_id_count / quiz_id_count) AS quiz_to_purchase
FROM T1
```

Almost half the users that completed the quiz made a purchase. We see that we loose 25% of users on the second stage and then even more on the third stage with 34%. We know that to retain more people from hometry to purchase, WarbyParker should offer users 5 pairs to try-on at home as it has proven to be more successful. Creating an A/B test during the quiz stage, could help determine how to drive more costumers to the try-on stage.

Thank you for your time and attention.