codecademy

Learn SQL from Scratch – Usage funnels with Warby Parker

Ethan Skeels
26/09/2018

Table of Contents

1. Quiz Funnel

1.1 Survey table column names

The table below outlines the three different columns that are used to store data in the *survey* table along with the first 10 rows of data.

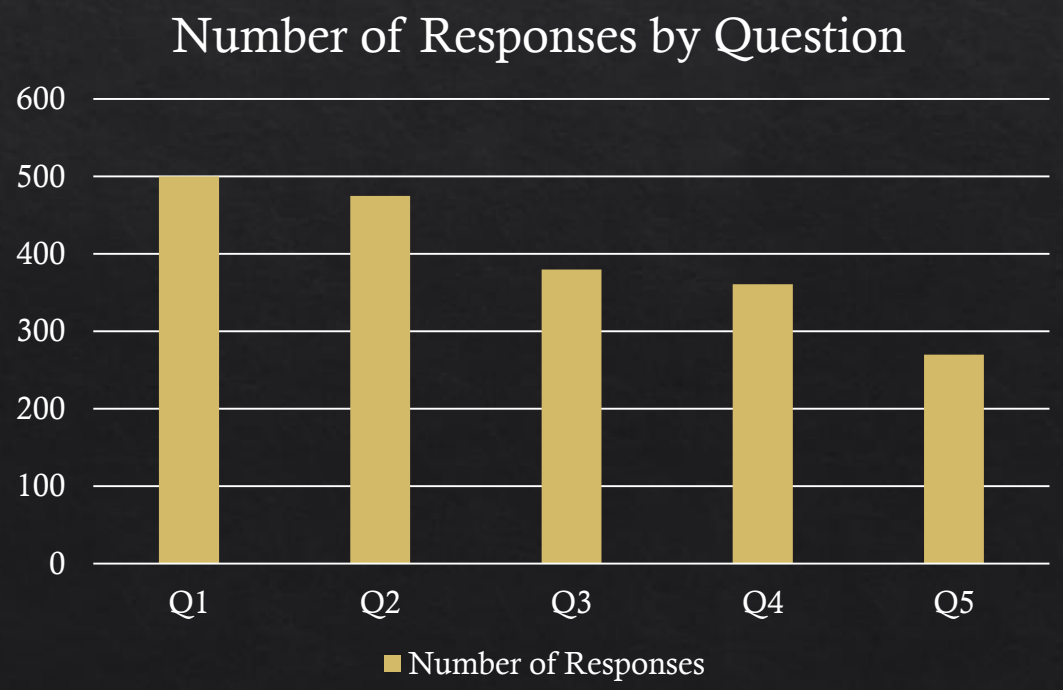
question	user_id	response
1. What are you looking for?	005e7f99-d48c-4fce-b605-10506c85aaf7	Women's Styles
2. What's your fit?	005e7f99-d48c-4fce-b605-10506c85aaf7	Medium
3. Which shapes do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	Round
4. Which colors do you like?	00a556ed-f13e-4c67-8704-27e3573684cd	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

1.2 Number of responses per question

In order to find the number of responses per question from the quiz you first need to select the five questions, secondly you need to count the distinct or unique IDs which are associated with each question. This is completed using the *COUNT* and *DISTINCT* functions. However, in order to show the results of the COUNT function per question you need to group the results by the question itself, otherwise you would not get the correct result. In order to make the columns easier to understand the AS function was used to rename them.

```
SELECT question AS "Question", COUNT(DISTINCT
user_id) AS "Number of Responses"
FROM survey
GROUP BY question;
```

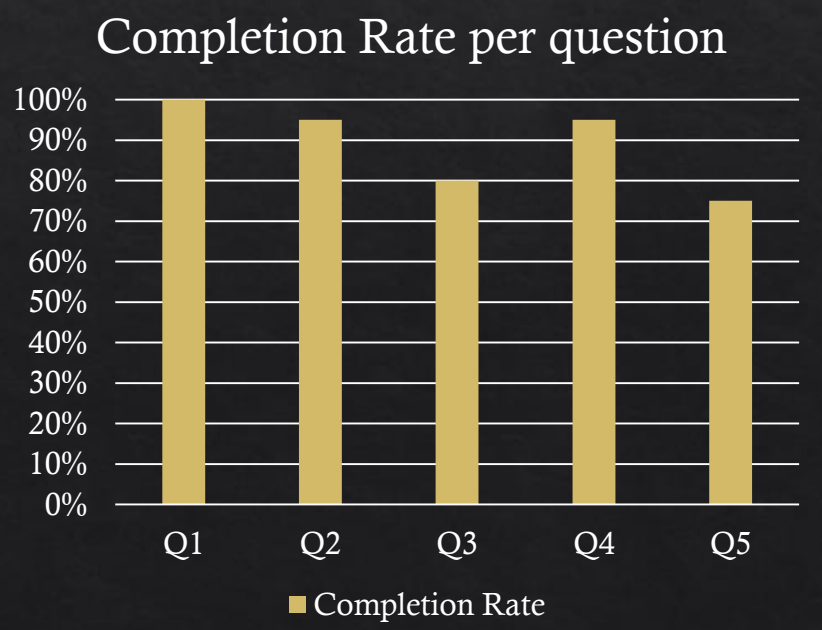
Question	Number of Responses
1. What are you looking for?	500
2. What's your fit?	475
3. Which shapes do you like?	380
4. Which colors do you like?	361
5. When was your last eye exam?	270



1.3 Survey Questions Completion Rates

The question with the lowest completion rate is question five ('When was your last eye exam?') followed by question three ('What shapes do you like?'). The likely reason that question five has the lowest completion rate is that it is the hardest question to answer. It is quite likely that a lot of people do not know when their last eye test was and so are not able to answer the question. Given that it is the last question on the quiz is not a likely reason for the low completion rate because this quiz is designed to help "users find the perfect frame". This means they want to do this quiz to aid their glasses decision.

Questions	Number of Responses	Completion Rate
1. What are you looking for?	500	100%
2. What's your fit?	475	95%
3. What shapes do you like?	380	80%
4. What colors do you like?	361	95%
5. When was your last eye exam?	270	75%



2. Home Try-on Funnel

2.1 Purchase funnel tables

The tables below outline the different columns spread across the three different tables that are used to store data.

‘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
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	Women's Styles	Narrow	Square	Two-Tone
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black

‘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
4e71850e-8bbf-4e6b-accc-49a7bb46c586	3 pairs	347 Madison Square N
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs	182 Cornelia 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
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

2.2 Purchase results from A/B test

It is possible to perform some basic analysis on the A/B test to see if there is any difference in purchasing patterns depending on whether the user was given three pairs or five pairs when they were part of the A/B test.
The results show that more purchases are made when the users are given the extra two pairs of glasses to try on. SQL has also been used to calculate the percentage of people who buy glasses after trying on the respective quantity of glasses from the A/B test. Only 53% of users buy glasses when receiving three different pairs where as 79% of users buy glasses when they get the five pairs.

Unique Pairs	Number of pairs	# of purchases per number of pairs	% people who buy glasses
3 pairs	379	201	53
5 pairs	371	294	79

```
SELECT number_of_pairs AS "Unique Pairs",
COUNT(number_of_pairs) AS "Number of pairs",
COUNT(DISTINCT purchase.user_id) AS "# purchases
per number of pairs", ROUND(100 *  COUNT(DISTINCT
purchase.user_id) / COUNT(number_of_pairs),0) AS
"% people who buy glasses"
FROM home_try_on
LEFT JOIN purchase ON home_try_on.user_id =
purchase.user_id
GROUP BY number_of_pairs;
```

2.3 New table creation

The table below shows the important information for each user. In order to uniquely identify each user their user_id is shown, followed by the column that lets us know if they tried any of the glasses on at home. The result of 'IS NOT NULL' is either '1' or '0'. This makes it hard to analyse the data. Therefore, a CASE WHEN statement is used to convert the ones and zeros to 'Yes' and 'No'. Therefore, if the user has a 'Yes' in their row then they tried the glasses on at home. The same true for the 'User purchase?' column. If they bought some glasses they get a 'Yes' recorded in their row. The A/B test is only for users who tried their glasses at home, therefore those with a NULL recorded against them are not part of the test.

user_id	User try at home?	number_of_pairs	User purchase?
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	Yes	3 pairs	No
291f1cca-e507-48be-b063-002b14906468	Yes	3 pairs	Yes
75122300-0736-4087-b6d8-c0c5373a1a04	No	NULL	No
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	Yes	5 pairs	No
ce965c4d-7a2b-4db6-9847-601747fa7812	Yes	3 pairs	Yes
28867d12-27a6-4e6a-a5fb-8bb5440117ae	Yes	5 pairs	Yes
5a7a7e13-fbcf-46e4-9093-79799649d6c5	No	NULL	No
0143cb8b-bb81-4916-9750-ce956c9f9bd9	No	NULL	No
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	Yes	5 pairs	No
b1dded76-cd60-4222-82cb-f6d464104298	Yes	3 pairs	No

```
SELECT quiz.user_id,
    CASE WHEN home_try_on.user_id IS NOT NULL = 1
THEN "Yes" ELSE "No" END AS "User try at
home?",home_try_on.number_of_pairs AS "Number of
pairs",
  CASE WHEN purchase.user_id IS NOT NULL = 1 THEN
"Yes" ELSE "No" END AS "User purchase?"
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
LIMIT 10;
```

2.4 Overall Conversion Rates

It is possible to work out conversion rates for the marketing funnel run by Warby Paker. First, a temporary table is created so that analysis can be completed on this new table. Then some simple functions are conducted (COUNT and SUM) to work out the total number of users at each level of the funnel. Then the percentage of users that move from one level to the other is calculated and displayed as a percentage.

From this analysis you can see that 66% of users who try their glasses on at home make a purchase.

Total Users	Total at Home	Total Purchases	% of users Quiz -> Home_Try_On	% of users Home_Try_On -> Purchase
1000	750	495	75	66

```
WITH funnels AS(
SELECT quiz.user_id,home_try_on.user_id IS NOT
NULL AS "User try at
home?",home_try_on.number_of_pairs AS "Number of
pairs", purchase.user_id IS NOT NULL AS "User
purchase?"
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 COUNT(*) AS "Total Users", SUM("User try at
home?") AS "Total at home", SUM("User Purchase?")
AS "Total purchases", 100 * SUM("User try at
home?") / COUNT(user_id) AS "% of users Quiz ->
Home_Try_On", 100 * SUM("User Purchase?") /
SUM("User try at home?") AS "% of users
Home_Try_On -> Purchase"
FROM funnels;
```

2.5 Insights from quiz table

It is possible to work out the most poplar results from each of the attributes in the quiz table. This is completed using the COUNT function. In summary, the most popular style is “Women’s Styles”, the most popular colour is “Tortoise”, the most popular fit is “Narrow” and the most popular shape is “Rectangular”. The full results are below:

style	COUNT(style)
Women's Styles	469
Men's Styles	432
I'm not sure. Let's skip it.	99

color	COUNT(color)
Tortoise	292
Black	280
Crystal	210
Neutral	114
Two-Tone	104

fit	COUNT(fit)
Narrow	408
Medium	305
Wide	198
I'm not sure. Let's skip it.	89

shape	COUNT(shape)
Rectangular	397
Square	326
Round	180
No Preference	97

2.6 Insights from purchase table

Below are the results from the analysis of the purchase table. They show how different styles, colours and products are selling so that the company could make decisions on what to sell and what not to sell. The product that sells the most is the product with ID 3, Women's styles are the most popular products and the products that are Jet-Black are sold the most.

product_id	COUNT(product_id)
3	63
10	62
9	54
1	52
6	50
4	44
7	44
2	43
8	42
5	41

style	COUNT(style)
Women's Styles	252
Men's Styles	243

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