codecademy

Usage Funnels with Warby Parker

Learn SQL from Scratch

Timothee Busch

9 March 2019

Table of Contents

B) A/B Testing with Home Try-On Funnel

1. Getting familiar with the tables:
    - quiz
    - home_try_on
    - purchases
2. Combining the three tables
3. Conversion rates
4. Purchase rates between customers who tried on 3 or 5 pairs of glasses
5. Purchase rates: 3 vs 5 pairs and different styles
6. Most common result in terms of style
7. Most common type of purchase

A) Quiz funnel

A)1. Getting familiar with the table survey

In order to be able to make a funnel out of the 5 questions of our survey, I need first to have a feel of the table survey:

- I selected a few rows to determine how many columns the table has.
- The different column names of this table are:
    - Question (list the questions to the user),
    - User_id (unique identifier of users)
    - Responses (list of answers the user gave)

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

On the next page, you can find the table with the first ten rows

Question	User_id	Responses
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

A)2. Number of responses for each questions

After getting to know the table a bit better, it’s time to make a funnel: we want to know how many users answered the 5 different questions. With the code I typed, I hereby determined the count of distinct users who answered the different questions, and ordered it from most answered to least answered.

```
SELECT question,
COUNT(DISTINCT user_id) AS 'answers'
FROM survey
GROUP BY 1
ORDER BY 2 DESC;
```

Questions	Answers
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

A)3 Percentage of questions answered

1. What are you looking for?	100.00%
2. What's your fit?	95.00%
3. Which shapes do you like?	76.00%
4. Which colors do you like?	72.20%
5. When was your last eye exam?	54.00%

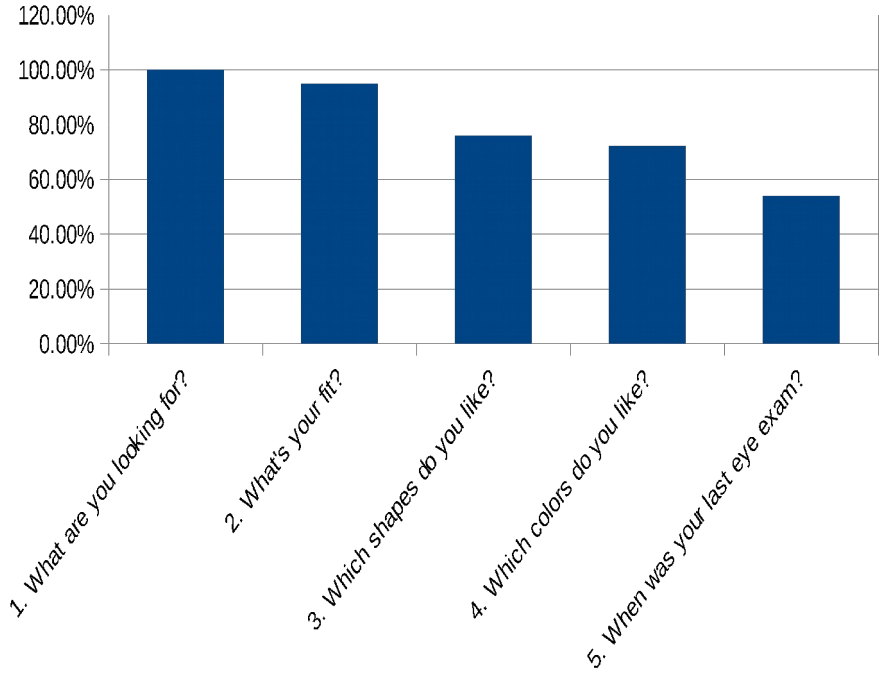
Then we want to have the percentage of questions answered by our users:

- I first queried the count of the total DISTINCT users who answered the questions.
- On the second query, I selected the questions, the count of distinct users who answered each questions divided by the total number of distinct users_id multiplied by 1.0 (to convert the result to a float) to have the percentage

```
SELECT
     COUNT(DISTINCT user_id)
FROM survey;
SELECT question,
1.0 *  COUNT(DISTINCT user_id)/500
       AS 'percentage answers'
FROM survey
GROUP BY 1;
```

As we can observe on the diagram, the completion percentage drops as the users go through the different questions. The final question is answered only by 54% of the total number of users who had answered the first question.
Different reason could explain this drop:

- Throughout the survey, some users realised that the shapes proposed are not suited to them (drop of 19% from question 2 to 3). Eventually, the survey could propose more shapes of glasses, so that the customer could stay more interested.
- The last question, question 5, although essential if one goes to buy a pair of glasses, could be interpreted by some users as an intrusion of their medical situation. Since this survey is only there for marketing purposes, we could scrap this last question entirely.



B) A/B Testing with Home Try-On Funnel

B)1. Getting familiar with the table quiz, home_try_on and purchase

We will make a funnel using those three tables to determine the purchase rate of customers taking the quiz and trying at home 3 or 5 glasses.
The table names are:
quiz, home_try_ on and purchase
Column names from each table are:

- Table quiz: User_id, style, fit, shape and color.
- Table home_try_on: User_id, number_of_pairs and Address.
- Table purchase: User_id, product_id, style, model_name, color and price.

You will find the three tables on the next page.
The User_id column relates the different table to one other

```
SELECT * FROM quiz LIMIT 5;
SELECT * FROM home_try_on LIMIT 5;
SELECT * FROM purchase LIMIT 5;
```

Quiz

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

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

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

B)2. Combining the three tables

We then combine those three tables using a left join, we join them on their user_id column, to compare if the same matching user_id appears is those three table. If it does not match, we will get a null result. We will only have four columns from those tables:

- Distinct User_id from quiz
- user_id from home_try_on who is True or False depending if user_id has an entry in home-try_on, when no matching user_id is found in home_try_on then result is False (0) otherwise result is True (1).
- Numbers of pairs
- user_id from purchase which is true or false depending if user_id has an entry in purchase, when no matching user_id is found in home_try_on then result is False (0) otherwise result is True (1).

```
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 q.user_id = p.user_id
LIMIT 10;
```

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	null	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	null	0

B)3. Conversion rates

With the result of this previous table we can then aggregate those results and calculate the overall conversion rate.
We use the previous code and create the temporary table funnel.
We count the total number of distinct users who first took the quiz, rename this column total users, sum the total amount of users who home tried some of the glasses and rename this column total home tried and finally the sum of the total amount of users who purchased those glasses and rename this column total purchased.
So from 1000 users who took the quiz, 750 tried glasses at home (5 or 3 pairs) and 495 purchased glasses.

```
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 q.user_id = p.user_id)
SELECT COUNT(DISTINCT user_id) as 'total users',
       SUM(is_home_try_on) as 'total home tried',
       SUM(is_purchase) as 'total purchased'
FROM funnel;
```

total users	total home tried	total purchased
1000	750	495

Then I calculated the percentage of the conversion with almost the same code just multiplying all the columns by 1.0 and dividing it by the count of distinct users. I also added to this result the percentage rate of users buying glasses after receiving the glasses at home.

total users	total home tried	total purchased	Purchased when home tried
100%	75%	49%	66%

```
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 q.user_id = p.user_id)
SELECT
  1.0 * COUNT(DISTINCT user_id)/
    1000 AS 'total users',
  1.0 * SUM(is_home_try_on)/
    1000 AS 'total home tried',
  1.0 * SUM(is_purchase)/
    1000 AS 'total purchased',
  1.0*SUM(is_purchase)/
     SUM(is_home_try_on) AS 'Purchased when
     home tried'
FROM funnel;
```

B)4. Purchases rates between customers who had 3 or 5 glasses

```
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 q.user_id = p.user_id)
SELECT    COUNT(user_id)as 'number of users',
          number_of_pairs AS 'number of pairs',
          SUM(is_purchase) AS 'sum of purchases',
          1.0 *SUM(is_purchase) /
     COUNT(user_id) AS 'percent bought'
FROM funnel
WHERE number_of_pairs IS NOT NULL
GROUP BY number_of_pairs;
```

number of user	number_of_pairs	sum of purchase	percent bought
379	3 pairs	201	53%
371	5 pairs	294	79%

B)4. Purchase rates between customers who tried 3 or 5 glasses

I calculated the different purchase rates between users who tried three pairs of glasses and the ones who tried five pairs of glasses.
I created a temporary table with the following columns: the count of user_id, the number of pairs, the sum of the users who made a purchase, and a percentage column calculating the percentage of purchases per glasses bought.
I eliminated from the result the instances where users did not try any pairs, and grouped the results by the number of pairs tried on so that we can observe the count of users who tried 3 pairs of glasses and the ones who tried 5 pairs of glasses, also we can see the purchase percentage of users who tried 3 glasses or 5.
We can clearly see in the results that the customers trying 5 pairs ended up purchasing more (294 against 201 pairs) with a percentage rate of 79% for the users who tried 5 pairs against a percentage rate of only 53% for the others.

B)5. Purchases rates 3 vs 5 glasses split by style

Reusing this temporary table funnel, we compare again the results between customers who tried three pairs, and the ones who tried 5 but this time we split the results by style. We have two different styles: Men's style and Women's style.

```
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',
    q.style
  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
     COUNT(user_id) AS 'number of user',
     number_of_pairs,
     SUM(is_purchase) AS 'sum of purchase',
     1.0 *SUM(is_purchase) / COUNT(user_id) AS 'percent bought',
     style
FROM funnel
WHERE number_of_pairs IS NOT NULL
AND (style LIKE "Men's Styles"
OR style LIKE "Women's Styles")
GROUP BY number_of_pairs, style
ORDER BY 3 DESC;
```

number of user	number_of_ pairs	sum of purchase	percent bought	style
184	5 pairs	154	83%	Women's Styles
155	5 pairs	140	90%	Men's Styles
165	3 pairs	103	62 %	Men's Styles
177	3 pairs	98	55%	Women's Styles

Users who chose Women’s style on the quiz end up buying more glasses than the one choosing men’s style.
But the percentage of purchases is higher with users wanting men’s style.
Assuming that the users wanting men’s style are men and the one wanting women’s style
are women we can conclude that regardless of the number of pairs tried, women were likely to try on a larger number of pairs of glasses than men, but men tended to buy more of the pairs of glasses that they tried on.

B)6. Most common result of style

To see the most common result of the style question in the quiz, we type type this query below

```
SELECT
     style,
     count(*)
FROM quiz
GROUP BY 1
ORDER BY 2 DESC;
```

We can see the result of this query in this table below

Style	Count(*)
Women's style	469
Men's Style	432
I'm not sure. Let's skip it.	99

As we can see, most users answered that they where looking for Women's style kind of glasses.

B)6. Most common types of purchases

We also want to know the three most common types of purchases made: we look at which types of glasses were sold the most, according to its style, model and color.
Limit this result to three.
We can see that the type which is bought the most is the Men’s Style Dawes model with the color Driftwood Fade.

```
SELECT
     style,
     model_name,
     color,
     COUNT(*) AS 'number bought'
FROM purchase
GROUP BY 1,2,3
ORDER BY 4 DESC
LIMIT 3;
```

style	model_name	color	Number bought
Men's Styles	Dawes	Driftwood Fade	63
Women's Styles	Eugene Narrow	Rosewood Tortoise	62
Women's Styles	Eugene Narrow	Rose Crystal	54

Conclusion

Looking at the result we have got from the survey and the home-try-on funnel, we can conclude by with those advices: In the survey, there should be more choices of shapes of glasses and the last question about the eye exam is not very well answered by a lot of potential customers, they could consider do not put it in the survey and ask it later when the customer made already his choice in glasses. On the home-try-on, users who tried 5 pairs of glasses end up buying much more than the one trying 3 pairs. Recommended to generally sent 5 pairs of glasses to the customers instead of 3 since the purchase rates is much higher (79% vs 53%).