



Warby Parker Case

Learn SQL from Scratch | Funnels

Frederico Meneses

23/06/2018

WP Case - Exercise 1

This query gives us the following:

- A table containing WP Quiz's results
- The * means that we have selected all columns available
- Columns: Question, user_id and response (see below)
- All this information was obtained from the quiz table called survey
- We limited our results to the first 10 rows/results

```
1  
2  select *  
3  from survey  
4  limit 10;
```

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.

WP Case - Exercise 2

This query gives us the following:

- The total amount of responses per question
- Columns selected: Question and 'Total_responses'
- Count (distinct user_id) – we have used this function in order to count all unique users that have replied to the quiz
- All this information was obtained from the quiz table called survey
- When using the function group by 1, we are grouping all unique users by question in order to understand how many unique users have replied to each question

```
6  select
7  question,
8  count (distinct user_id) as
   'total_responses'
9  from survey
10 group by 1;
```

question	total_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

WP Case - Exercise 3

- % = number of total answers of the current question/number of total answers of the previous question. Example: Only 75% of the total respondents have replied to the 5th question
- The 2 questions with lower completion rate are the questions number 3 (80%) and 5 (75%)
- The proposed causes, in the grey box

Questions	Total answers	%
1	500	100%
2	475	95%
3	380	80%
4	361	95%
5	270	75%

Q3 - "Which shapes do you like"?

Why:

- It requires more information next to the 3 options, similarly, to what we have in the question 2 for example, giving more examples to the customer helping me in the decision process
- it might be quite difficult to know which shapes could fit better without trying first

Q5 - "When was your last exam"?

Why:

- Think most people usually do not remember when was their last eye exam
- We might have a high level of friction as the customers needs to do an extra effort to remember the above + the Q5 is the 6th question whereas customer's patience/attention span is lower. (could be interesting to a/b test this question as if it was a Q2 or Q3 question and see if there is any significant change

WP Case - Exercise 4 – Table Quiz

This query gives us the following:

- A table containing WP Quiz's results
- The * means that we have selected all columns
- Columns: user_id, style, fit, shape and color (see below)
- All this information was obtained from the quiz table
- We limited our results to the first 5 rows/results

```
17  select *
18  from quiz
19  limit 5;
```

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-c0c5373ala04	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

WP Case - Exercise 4 – Table Home_try_on

This query gives us the following:

- A table containing WP Quiz's results
- The * means that we have selected all columns
- Columns: user_id, number_of_pairs and address(see below)
- All this information was obtained from the home_try_on table
- We limited our results to the first 5 rows/results

```
21  select *
22  from home_try_on
23  limit 5;
```

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-acc4-49a7bb46c586	3 pairs	347 Madison Square N
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs	182 Cornelia St

WP Case - Exercise 4 – Table Purchase

This query gives us the following:

- A table containing WP Quiz's results
- The * means that we have selected all columns
- Columns: user_id, product_id, style, model_name, color and price (see below)
- All this information was obtained from the purchase table
- We limited our results to the first 5 rows/results

```
25  select *
26  from purchase
27  limit 5;
```

user_id	product_id	style	model_name	color	price
00a9dd17-36c8-430c-9d76-df49d4197dcf	8	Women's Styles	Lucy	Jet Black	150
00e15fe0-c86f-4818-9c63-342221baa97	7	Women's Styles	Lucy	Elderflower Crystal	150
017506f7-abaf-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
01fd106-f73c-4d3f-a036-2f3e2ab1ce06	8	Women's Styles	Lucy	Jet Black	150

WP Case - Exercise 5

This query gives us the following:

- A table containing WP Quiz's results
- Columns selected: q.user_id, h.user_id, h.number_of_pairs and p.user_id
- "is not null" means that we only consider the variable's result if the variable has results or have entries
- All this information was obtained due to the function left join combining 3 different tables: quiz , home_try_on and purchase
- The common element among all of the tables is user_id
- We limited our results to the first 10 rows/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

```
30  select distinct q.user_id,
31  h.user_id is not null as
    'is_home_try_on',
32  h.number_of_pairs,
33  p.user_id is not null as
    'is_purchase'
34  from quiz as 'q'
35  left join home_try_on as 'h'
36  on q.user_id = h.user_id
37  left join purchase as 'p'
38  on h.user_id = p.user_id
39  limit = 10;
```


WP Case - Exercise 6 | Part 1

This query gives us the following:

- Overall Conversion % : 0.495
- Cvr% from Quiz – Home_try_on: 0.75
- Cvr% from Home_try_on – Purchase: 0.66
- 1000 people have done the Quiz
- 750 went from the Quiz to the Home_try_on
- 495 went from the Home_try_on to the purchase stage

n_quiz	n_home_try_on	n_purchase	quiz_to_home_%	home_to_purchase_%	overall_conversion%
1000	750	495	0.75	0.66	0.495

```
29 WITH t AS (  
30   select distinct q.user_id,  
31   h.user_id is not null as 'is_home_try_on',  
32   h.number_of_pairs,  
33   p.user_id is not null as 'is_purchase'  
34   from quiz as 'q'  
35   left join home_try_on as 'h'  
36   on q.user_id = h.user_id  
37   left join purchase as 'p'  
38   on h.user_id = p.user_id)  
39 SELECT  
40   count (*) as 'n_quiz',  
41   sum(is_home_try_on) as 'n_home_try_on',  
42   sum(is_purchase) as 'n_purchase',  
43   1.0 * sum(is_home_try_on)/count (user_id) as  
44   'quiz_to_home_%',  
45   1.0 * sum(is_purchase)/sum(is_home_try_on) as  
46   'home_to_purchase_%',  
47   1.0 * sum(is_purchase)/count(user_id) as  
48   'overall_conversion_%'  
49 from t;
```

WP Case - Exercise 6 | Part 2

This query gives us the following cvr% per funnel accordingly to the number of pairs that the customer has chosen .

Interesting facts:

- When the customer gets 5 pairs, the cvr% increases +61% $((0.79-0.49)/(0.49))$
- When the customer gets 3 pairs , the cvr% increases +8.16% $((0.53-0.49)/(0.49))$

*In order to get the same results regarding 5 pairs, swipe the number 3 for the number 5 in these 2 lines: **number_of_pairs = 3** and where **number_of_pairs = '3 pairs'**

Number_of_pairs	3	5	Total
Number of people	379	371	750
Home to Purchase %	0.53	0.79	0.66
Overall CVR%	0.53	0.79	0.495

n_quiz	n_home_try_on	n_purchase	quiz_to_home_%	home_to_purchase_%	overall_conversion_%
371	371	294	1.0	0.792452830188679	0.792452830188679
n_quiz	n_home_try_on	n_purchase	quiz_to_home_%	home_to_purchase_%	overall_conversion_%
379	379	201	1.0	0.530343007915567	0.530343007915567

```
93 -- when number_of_pairs = 3
94 WITH t AS (
95   select distinct q.user_id,
96   h.user_id is not null as 'is_home_try_on',
97   h.number_of_pairs,
98   p.user_id is not null as 'is_purchase'
99   from quiz as 'q'
100   left join home_try_on as 'h'
101   on q.user_id = h.user_id
102   left join purchase as 'p'
103   on h.user_id = p.user_id)
104 SELECT
105   count (*) as 'n_quiz',
106   sum(is_home_try_on) as 'n_home_try_on',
107   sum(is_purchase) as 'n_purchase',
108   1.0 * sum(is_home_try_on)/count (user_id) as
109   'quiz_to_home_%',
110   1.0 * sum(is_purchase)/sum(is_home_try_on) as
111   'home_to_purchase_%',
112   1.0 * sum(is_purchase)/count(user_id) as
113   'overall_conversion_%'
114 from t
115 where number_of_pairs = '3 pairs';
```

WP Case - Exercise 6 | Part 3

This query gives us the following:

- Gender %:
F: $252/500 = 50.4\%$
M: $248/500 = 49.6\%$
- Fit %:
Medium: $305/1000 = 30.5\%$
Narrow: $305/1000 = 40.8\%$
Wide: $305/1000 = 19.8\%$
- AOV € : 112.7€
- Favourite colors:
Black: $280/1000 = 28\%$
Crystal: $210/1000 = 21\%$
Neutral: $114/1000 = 11.4\%$
Tortoise: $292/1000 = 29.2\%$
TwoTone: $104/1000 = 10.4\%$

style	total_styles
Men's Styles	243
Women's Styles	252
fit	total_fits
I'm not sure. Let's skip it.	89
Medium	305
Narrow	408
Wide	198

color	total_colors
Black	280
Crystal	210
Neutral	114
Tortoise	292
Two-Tone	104

avg(price)
112.717171717172

```
48 select
49 style,
50 count (distinct user_id) as 'total_styles'
51 from purchase
52 group by 1;
53
54 select
55 fit,
56 count (distinct user_id) as 'total_fits'
57 from quiz
58 group by 1;
59
60 select
61 avg(price)
62 from purchase;
63
64 select
65 color,
66 count (distinct user_id) as 'total_colors'
67 from quiz
68 group by 1;
```

WP Case - Exercise 6 | Conclusions

Conclusions:

- Typical Customer:
 - Gender: Female/Male
 - Style: Narrow
 - AOV: 112.7€
 - Favourite Color: Black and Tortoise
- In order to increase the number of people answering the question 3 and 5 more often you should a/b test the position of the questions and see if the order has impact. For example, maybe put the questions that require more effort(for example the question number 5), in the middle rather than in the end of the quiz in order to reduce friction
- In order to increase overall cvr%:
All your communication strategy should reinforce and/or give exposure to the features of the typical customer

Add an incentive to Increase the number of customers trying 5 pairs instead of only 3

Find the marginal add of number of pairs in order to maximize cvr%

Use social proof regarding the style and favourite colours : "Everyone has been picking the style Narrow and in black and Tortoise in order to increase cvr%

Free delivery and/or any other bonus if you spend above 112.99€ resulting in AOV increase