

Warby Parker Capstone

Learn SQL from Scratch

Erwin Kuipers - The Netherlands

May 2019

Example Table of Contents

1. Examine Quiz funnel
2. Examine Home Try-on Funnel

Quiz Funnel

Question 1

What columns does the 'survey' table have?

- SELECT statement, LIMIT by 10 rows
- See screenshot; it contains 3 columns with the question, the user_id and the response. All are 'text'.
- The table contains 1.986 rows
- user_id is the unique key which can be used for a join

```
SELECT *  
FROM   quiz  
LIMIT  10;
```

Query Results		
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

Database Schema

survey		1986 rows
question		TEXT
user_id		TEXT
response		TEXT

Question 2

Create a quiz funnel using the 'GROUP BY' command. What are the responses for each question?

- In total 500 users started the survey, and 270 of them completed it.
- Used GROUP BY to create the question bucketing

```
SELECT question,  
COUNT (user_id) AS Responses  
FROM survey  
GROUP BY 1;
```

Query Results

question	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

Question 3

Using spreadsheet to calculate the percentage of user who answer each question:

Query Results	
question	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

Completion	drop-off
100%	
95%	-5%
76%	-19%
72%	-4%
54%	-18%

Drop-off between question 2 and 3 is the biggest with 19%. Reasoning could be that customer do not know which shapes they like or that there are not enough options to choose from.

Home Try-On Funnel

Question 4

Examine the first five rows of each table (Quiz, home_try_one, purchase)

- Used SELECT statement with LIMIT to get results
- See screenshot below for the column names of each table
- User_id is in all tables and can be used when joining

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

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

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

Query Results					
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-a301-27dd93b12e2	Women's Styles	Narrow	Square	Two-Tone	
ce965c4d-7a2b-4db6-9847-601747fa7812	Women's Styles	Wide	Rectangular	Black	
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-acc6-49a7bb46c586	3 pairs		347 Madison Square N		
3bc8f97f-2336-4dab-bd86-e391609dab97	5 pairs		182 Cornelia St		
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

→ Quiz table

→ Home_try_on table

→ Purchase table

Question 5

Create predefined table:

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc	True	3	False
291f1cca	True	5	False
75122300	False	NULL	False

- Used UPDATE to delete the word 'pairs' from the number_of_pairs output to create the predefined table
- Used LEFT JOIN on user_id when combining the tables
- Used CASE for to create 'True' and 'False' output in column 1 and 3

Query Results

user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	True	3	False
291f1cca-e507-48be-b063-002b14906468	True	3	True
75122300-0736-4087-b6d8-c0c5373a1a04	False	Ø	False
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	True	5	False
ce965c4d-7a2b-4db6-9847-601747fa7812	True	3	True
28867d12-27a6-4e6a-a5fb-8bb5440117ae	True	5	True
5a7a7e13-fbcf-46e4-9093-79799649d6c5	False	Ø	False
0143cb8b-bb81-4916-9750-ce956c9f9bd9	False	Ø	False
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	True	5	False
b1dded76-cd60-4222-82cb-f6d464104298	True	3	False

```
--DELETE THE WORD PAIRS FROM home_try_on--
UPDATE home_try_on SET number_of_pairs =
REPLACE(number_of_pairs, 'pairs', '');
--FUNNELDATA--
WITH 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,
    number_of_pairs,
    CASE
        WHEN p.user_id IS NOT NULL THEN 'True'
        ELSE 'False'
    END 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 p
ON
    q.user_id = p.user_id )
--OUTPUT--
SELECT
    *
FROM
    Funnel
LIMIT
    10;
```

Question 6 - Analyse data from 'Funnel' table

What is the overall conversion rate?

- Changed the CASE into 1's and 2's in order to use SUM statement for the calculation
- Used ROUND to create a two decimal output

Query Results
Conv.Rate_Overall
0.49

```
WITH Funnel AS (  
SELECT  
  DISTINCT q.user_id,  
  CASE  
    WHEN h.user_id IS NOT NULL THEN '1'  
    ELSE '0'  
  END AS is_home_try_on,  
  number_of_pairs,  
  CASE  
    WHEN p.user_id IS NOT NULL THEN '1'  
    ELSE '0'  
  END 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 p  
ON  
  q.user_id = p.user_id )  
SELECT  
  ROUND(1.0 * SUM (is_purchase) / COUNT (user_id),2) AS  
'Conv.Rate_Overall'  
FROM  
  Funnel;
```

Question 6 - Analyse data from 'Funnel' table

What are the step by step conversion rates?

- SELECT both the absolute numbers per step and calculated the conversion rates per step

Query Results		
Step1_Users_Quiz	Step2_Users_HomeTry	Step3_Users_Purchase
1000	750	495

Conv.Rate_Step1_to_2	Conv.Rate_Step2_to_3
0.75	0.66

```
WITH Funnel AS (  
  SELECT  
    DISTINCT q.user_id,  
    CASE  
      WHEN h.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END AS is_home_try_on,  
    number_of_pairs,  
    CASE  
      WHEN p.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END 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 p  
  ON  
    q.user_id = p.user_id )  
SELECT  
  COUNT (user_id) AS 'Step1_Users_Quiz',  
  SUM (is_home_try_on) AS 'Step2_Users_HomeTry',  
  SUM (is_purchase) AS 'Step3_Users_Purchase',  
  1.0 * SUM (is_home_try_on) / COUNT (user_id) AS  
  'Conv.Rate_Step1_to_2',  
  1.0 * SUM (is_purchase) / SUM (is_home_try_on) AS  
  'Conv.Rate_Step2_to_3'  
FROM  
  Funnel;
```

Question 6 - Analyse data from 'Funnel' table

What is the difference in purchases between customer who had 3 pairs or 5 pairs to test at home?

- Used ROUND to have a two decimal output in column 3
- Used GROUP BY to create the number_of_pairs bucketing
- ORDER BY column 2

Query Results		
number_of_pairs	Total Users	% User Reached Purchase
3 pairs	379	0.53
5 pairs	371	0.79
Ø	250	0.0



Interesting; users who did not apply for pairs to test did not reached the final step of the funnel. Marketing and sales need to be focussed on pushing test pairs the market.

```
WITH Funnel AS (  
  SELECT  
    DISTINCT q.user_id,  
    CASE  
      WHEN h.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END AS is_home_try_on,  
    number_of_pairs,  
    CASE  
      WHEN p.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END 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 p  
  ON  
    q.user_id = p.user_id )  
SELECT  
  number_of_pairs,  
  COUNT (user_id) AS 'Total Users',  
  ROUND(1.0 * SUM (is_purchase) / COUNT (user_id),2) AS '%  
User Reached Purchase'  
FROM  
  Funnel  
GROUP BY  
  1  
ORDER BY  
  2 DESC;
```

Question 6 - Analyse data from 'Funnel' table

What is the impact of the number of pairs tried at home on the average order value?

- Adding a calculation to the existing SQL code (AvgOrderValue)

Query Results			
number_of_pairs	Total Users	% User Reached Purchase	AvgOrderValue
3 pairs	379	0.53	60
5 pairs	371	0.79	89
Ø	250	0.0	Ø



Adding a datapoint to the previous insight: customers who tried on 5 pairs have a higher order value!

```
WITH Funnel AS (  
  SELECT  
    DISTINCT q.user_id,  
    CASE  
      WHEN h.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END AS is_home_try_on,  
    number_of_pairs,  
    CASE  
      WHEN p.user_id IS NOT NULL THEN '1'  
      ELSE '0'  
    END AS is_purchase,  
    price  
  FROM  
    quiz AS q  
  LEFT JOIN  
    home_try_on AS h  
  ON  
    q.user_id = h.user_id  
  LEFT JOIN  
    purchase AS p  
  ON  
    q.user_id = p.user_id )  
SELECT  
  number_of_pairs,  
  COUNT (user_id) AS 'Total Users',  
  ROUND(1.0 * SUM (is_purchase) / COUNT (user_id),2) AS '%  
User Reached Purchase',  
  SUM (price) / COUNT (user_id) AS AvgOrderValue  
FROM  
  Funnel  
GROUP BY  
  1  
ORDER BY  
  2 DESC;
```

Question 6 - Analyse data from 'Funnel' table

Survey: most popular answer given to the survey question

- Used MAX statement and GROUP by question

```
SELECT
  question,
  MAX(response) AS MaxResp
FROM
  Survey
GROUP BY
  1;
```

Query Results	
question	MaxResp
1. What are you looking for?	Women's Styles
2. What's your fit?	Wide
3. Which shapes do you like?	Square
4. Which colors do you like?	Two-Tone
5. When was your last eye exam?	Not Sure. Let's Skip It



Interesting; models that fulfill the above needs would probably have a high demand. Also, an eye exam could be a good additional proposition to go to market with, next to the actual glasses.