codecademy

Funnels Warby Parker

Learn SQL from Scratch

Ana Carolina Cronemberger C Mendonca

July 24, 2018

https://gist.github.com/a166c244662b5ff2c349ce7f492d689a

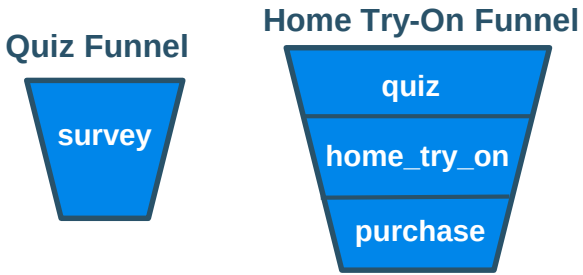
Table of Contents

1. Get familiar with Warby Parker

1.1 Presenting Tables and Funnels

What are Warby Parker´s Tabels and Funnels?

- Warby Parker uses four tables to determine its marketing funnels in order to calculate conversion rates



Survey Table	Quiz Table	Home_try_on Table	Purchase Table
question	user_id	user_id	user_id
user_ID	style	number_of_pairs	product_id
response	fit	address	style
	shape		model_name
	color		color
			price

```sql
-- Get to know survey table
SELECT *
FROM survey
LIMIT 10;
```

```sql
-- Get to know quiz table
SELECT *
FROM quiz
LIMIT 5;

-- Get to know home_try_on table
SELECT *
FROM home_try_on
LIMIT 5;

-- Get to know purchase table
SELECT *
FROM purchase
LIMIT 5;
```

2. Quiz Funnel

2.1 The Quiz Funnel

What is the Quiz Funnel?

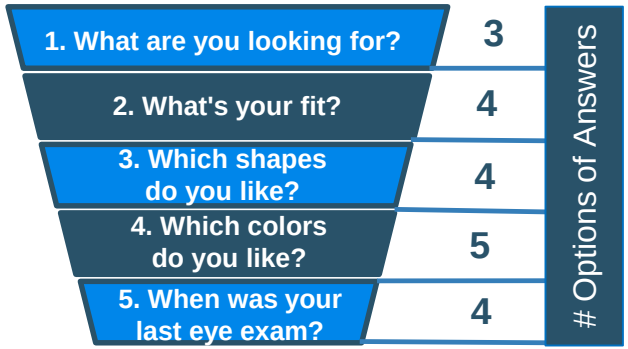
- The first step the customer come through. It helps users find their perfect frame.

What are the questions form the Quiz Funnel?

- Find the distinct values of the question column. A simple SELECT DISTINCT query answers this question quickly.

How many possible answers per question?

- Count the distinct responses for each question by using COUNT(DISTINCT), and GROUP BY and ORDER BY DISTINCT question.




```sql
-- List questions
SELECT DISTINCT(question)
FROM survey;

-- Questions and correspondet answers
SELECT DISTINCT(question),
COUNT(DISTINCT(response))
FROM survey
GROUP BY 1
ORDER BY 1;
```

2.2 The Quiz Funnel

Users can "give up" at different points in the survey. Build a funnel to identify in which moment the customers are leaving.

What is the number of responses for each question?

- Count the distinct user_id for each question by using COUNT(DISTINCT), and GROUP BY and ORDER BY DISTINCT question

Which question(s) of the quiz have a lower completion rates?

- Using Excel Sheet: # of Responses of an specific question divide by the # of Responses of the previous question, the result * 100%

What do you think is the reason?

- The fifth question is not referring to the glasses, but it is related to the buyer. It is a personal question, and it is related to the buyer's personal health. They may do not want to answer it.
- Considering that each added question is going to decrease the number of customers in the journey, the company should delete the question and leave the quiz with only four question, which is enough to establish which glasses to send in the Home Try-On Stage.

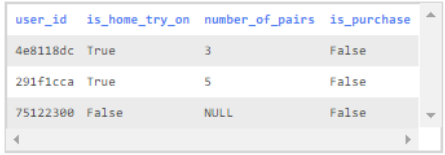
```
18  -- number of responses for each question
19  SELECT question,
20    COUNT(DISTINCT user_id)
21  FROM survey
22  GROUP BY 1
23  ORDER BY 1 ASC;
24
```

question	# of Responsers	% of users who answer each question.:
1. What are you looking for?	500	100%
2. What's your fit?	475	95%
3. Which shapes do you like?	380	80%
4. Which colors do you like?	361	95%
5. When was your last eye exam?	270	75%

3. Home Try-On Funnel

3.1 Home Try-On Funnel

To build the Home Try-On Funnel is necessary to create a new Table that JOINs the three tables (quiz, home_try_on, and purchase) following the layout:



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

- Replace the natural result for the aggregates IS NULL (0) and IS NOT NULL (1) to False and True, respectively, by creating a SELECT CASE WHEN _____ THEN_____ ELSE _____ END
- Finally, LEFT JOIN the tables ON the column they have in common

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	NULL	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	NULL	FALSE
0143cb8b-bb81-4916-9750-ce956c9f9bd9	FALSE	NULL	FALSE
a4ccc1b3-cbb6-449c-b7a5-03af42c97433	TRUE	5	FALSE
b1dded76-cd60-4222-82cb-f6d464104298	TRUE	3	FALSE

```sql
-- Create New Table
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 'NULL'
  WHEN h.number_of_pairs = '3 pairs' THEN '3'
  WHEN h.number_of_pairs = '5 pairs' THEN '5'
  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 p.user_id = q.user_id
LIMIT 10;
```

3.2 Home Try-On Funnel

Calculate overall conversion rate

- Make the LEFT JOIN of the three tables as a **new table named "funnel"** using WITH AS.
- SELECT COUNT to find the total number of users
- COUNT DISCTINCT using CASE WHEN _____ THEN____ ELSE _____ END to find the number of users who made a purchase
- Create a new table from "funnel" using WITH AS named "new"
- SELECT Dividing num_purchase by num_quiz and multiply by 1.0 to get the Conversion Rate
- The conversion rate shows that less than 50% of the customers who start the quiz makes a purchase. The company is loosing 50% of chance to make a sale

num_quiz	num_purchase	overall conversion rate
1000	495	49.5%

```sql
-- overall conversion rate
WITH new AS (
  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',
CASE
  WHEN h.number_of_pairs IS NULL THEN 'NULL'
  WHEN h.number_of_pairs = '3 pairs' THEN '3'
  WHEN h.number_of_pairs = '5 pairs' THEN '5'
  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 p.user_id = q.user_id)
SELECT COUNT(*) AS 'num_quiz',
COUNT(DISTINCT CASE
      WHEN is_purchase = 'True' THEN user_id
      END) AS 'num_purchase'
FROM funnel)
SELECT num_quiz, num_purchase,
1.0 * num_purchase / num_quiz AS 'overall conversion rate'
FROM new;
```

3.3 Home Try-On Funnel

Compare conversion from quiz→home_try_on and home_try_on→purchase

- Used to identify in which phase of the funnel the customer “gives up”
- The key solution for that query is how to count the number of distinct True events in the is_home_try_on and is_purchase columns. Use COUNT DISTINCT with a CASE _____ THEN_____ END
- Create a “new” table from “funnel” using WITH AS, and then find the conversion rates using SELECT dividing funnel level by prior funnel level and multiplying by 1.0
- The result shows that 25% of the customers left from quiz to home_try_on, while 44% left from home_try_on to purchase. What means that the customers are not purchasing because the glasses did not fit them

um_quiz	num_try_on	num_purchase	Percentage of users from quiz to home try on	Percentage of users from home try on to purchase
1000	750	495	75.0%	66.0%

```sql
-- conversion from quizhome_try_on and home_try_onpurchase.
WITH new AS (
  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',
CASE
  WHEN h.number_of_pairs IS NULL THEN 'NULL'
  WHEN h.number_of_pairs = '3 pairs' THEN '3'
  WHEN h.number_of_pairs = '5 pairs' THEN '5'
  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 p.user_id = q.user_id)
SELECT COUNT(*) AS 'num_quiz',
COUNT(DISTINCT CASE
      WHEN is_home_try_on = 'True' THEN user_id
      END) AS 'num_try_on',
COUNT(DISTINCT CASE
      WHEN is_purchase = 'True' THEN user_id
      END) AS 'num_purchase'
FROM funnel)
SELECT num_quiz, num_try_on, num_purchase,
1.0 * num_try_on / num_quiz AS 'Percentage of users from quiz to home try on',
1.0 * num_purchase / num_try_on AS 'Percentage of users from home try on to purchase'
FROM new;
```

3.4 Home Try-On Funnel

A/B Test

To conduct an A/B test, during the Home Try-On stage:

- 50% of the users will get 3 pairs to try on
- 50% of the users will get 5 pairs to try on

Find out whether or not users who get more pairs to try on at home will be more likely to make a purchase.

- Make the LEFT JOIN of the three tables as a **new table named "funnel"** using WITH AS
- SELECT DISTINCT number_of_pairs and COUNT how many responses they got AS number of try-on
- To COUNT the number of purchases per pair, create a DISTINCT CASE _____ THEN_____ END

pairs	num_try_on	num_purchase
3	379	201
5	371	294
NULL	250	0

```sql
-- A/B Test
WITH new AS (
  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',
CASE
  WHEN h.number_of_pairs IS NULL THEN 'NULL'
  WHEN h.number_of_pairs = '3 pairs' THEN '3'
  WHEN h.number_of_pairs = '5 pairs' THEN '5'
  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 p.user_id = q.user_id)
SELECT DISTINCT(number_of_pairs) AS 'pairs', COUNT(*) AS 'num_try_on',
  COUNT(DISTINCT CASE
     WHEN is_purchase = 'True' THEN user_id
     END) AS 'num_purchase'
FROM funnel
GROUP BY 1)
SELECT pairs, num_try_on, num_purchase,
  round(1.0 * num_purchase / num_try_on,2) AS 'conversion rate'
FROM new;
```

3.5 Home Try-On Funnel

A/B Test -- continuing

- Calculate the difference in purchase rates between customers who had 3 number_of_pairs with ones who had 5 by creating a new table named “new” from the previous table “funnel” using WITH AS
- SELECT dividing num_purchase by num_try_on and multiply by 1.0 to get the Conversion Rate
- The customers who get 5 pairs are more likely to make a purchase than the customers that get 3 pairs. Conversion rate for 5 pairs is 79%, while the 3 pairs is 53%
- If the company sends more option for the customer to try-on it increases the chances of selling

pairs	num_try_on	num_purchase	conversion rate
3	379	201	53.0%
5	371	294	79.0%
NULL	250	0	0.0%

```
-- A/B Test
WITH new AS (
  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',
CASE
  WHEN h.number_of_pairs IS NULL THEN 'NULL'
  WHEN h.number_of_pairs = '3 pairs' THEN '3'
  WHEN h.number_of_pairs = '5 pairs' THEN '5'
  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 p.user_id = q.user_id)
SELECT DISTINCT(number_of_pairs) AS 'pairs', COUNT(*) AS 'num_try_on',
  COUNT(DISTINCT CASE
     WHEN is_purchase = 'True' THEN user_id
     END) AS 'num_purchase'
FROM funnel
GROUP BY 1)
SELECT pairs, num_try_on, num_purchase,
  round(1.0 * num_purchase / num_try_on,2) AS 'conversion rate'
FROM new;

```

3.6 Home Try-On Funnel

The most common results of the style quiz

- Women's Style. Women are more likely to finish the purchase, although the men are just a little behind

The most common model of purchase made

- Eugene Narrow is the most purchased model; it represents 23% of total sales.

```sql
-- The most common results of the style quiz
SELECT DISTINCT(style), COUNT(*)
FROM quiz
GROUP BY style
ORDER BY style DESC;

-- The most common types of purchase made
SELECT DISTINCT(model_name), COUNT(*)
FROM purchase
GROUP BY 1
ORDER BY 2 DESC;
```

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

model_name	COUNT(*)
Eugene Narrow	116
Dawes	107
Brady	95
Lucy	86
Olive	50
Monocle	41

3.7 Home Try-On Funnel

What are some actionable insights for Warby Parker?

Matching Women´s style and Shape

- Very few women do not know their preferred shape, only 46 of them. Round glasses is a small portion of the business
- The most popular shape is Rectangular (184 women), followed by Square (158 Women), they represent 73% of sales. Warby Parker should invest in more assortment/options of rectangular and square shapes.

Price

- Counting the number of sales grouped by the Selling Prices, just a few customers purchased the cheapest glasses, the great majority are not concerned about the price
- More than 50% of sales are in the $ 95 selling price, the intermediate price level. However, the more expensive glasses have good acceptance, representing 40% of sales
- Warby Parker does not need to care about cheap products; it can offer quality glasses to its clients.

```
178  -- prices from purchase
179  SELECT DISTINCT(price), COUNT(*)
180  FROM purchase
181  GROUP BY 1
182  ORDER BY 1;
183
184  -- Match style and Shape from quiz
185  SELECT DISTINCT(shape), COUNT(*)
186  FROM quiz
187  WHERE style LIKE 'Women%'
188  GROUP BY 1;
189
```

shape	# of Women
Rectangular	184
Square	158
Round	81
No Preference	46

price	COUNT(*)
50	41
95	261
150	193