codecademy

Capstone Project

Learn SQL from Scratch

Paulo Porto

June 27, 2018

Table of Contents

1. Style Quiz

1.1 Style Quiz

A style quiz was created to collect information from users:

- Users will respond different questions to determine profile and preferences
- Quiz responses will be stored in a table with the following structure:

Question	User_id	Response
Text field	Text field	Text field
Pre-defined values	Stores the user unique identifier	Stores responses from questions
Question * … 1 User Id	Primary key	User id 1 …. * Responses

```
-- Style Quiz funnels

SELECT *
FROM survey
LIMIT 10;
```

1.2 Style Quiz - funnels

- Based on the data collected, it becomes clear that the number of users were declining as the quiz progresses
- The numbers for each question also accounts for answers like “skip it to next questions” and “no preference”.
- For that reason, the numbers represents abandonment.

```
-- Total responses per questions

SELECT question, COUNT (response) AS number_responses
FROM survey
GROUP BY question;
```

question	number_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 Style Quiz - funnels

- Transition rates:
  - Question 1 → Question 2: 95% of users continued
  - Question 2 → Question 3: 80% of users continued
  - Question 3 → Question 4: 95% of users continued
  - Question 4 → Question 5: 95% of users continued

- The overall transition rate (From Question 1 → Question 5), based on the results collected, 54% of all users completed the whole quiz.

```
-- Transition rates

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

question	number_responses	% of Respondents
1. What are you looking for?	500	100%
2. What's your fit?	475	95%
3. Which shapes do you like?	380	76%
4. Which colors do you like?	361	72%
5. When was your last eye exam?	270	54%

2. Home Try-On Funnel

2.1 Home Try-On Funnel – Data source

- To perform analysis on Home Try-On and Purchases, data will be collected from 3 different tables:
  - Quiz
  - Home_try_on
  - Purchase
  - Will use “User_id” as the common data-point to perform the cross-data analysis

```
-- Home Try-On Table

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

quiz	home_try_on	purchase
User_id	User_id	User_id
Style	Number_of_pairs	Product_id
Fit	address	Style
Shape		Model_name
Color		Color
		price

2.2 Home Try-On Funnel – Data collection

- To perform analysis on Home Try-On and Purchases, data will be collected from 3 different tables: Purchase, Quiz and Home_Try
- See a sample of “Funnels” table on the side

```
-- Home Try-On Table

SELECT DISTINCT q.user_id,
CASE WHEN
   h.user_id IS NOT NULL
   THEN 'True'
   ELSE 'False'
   END AS 'is_home_try_on',
h.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 p.user_id = q.user_id
LIMIT 10;
```

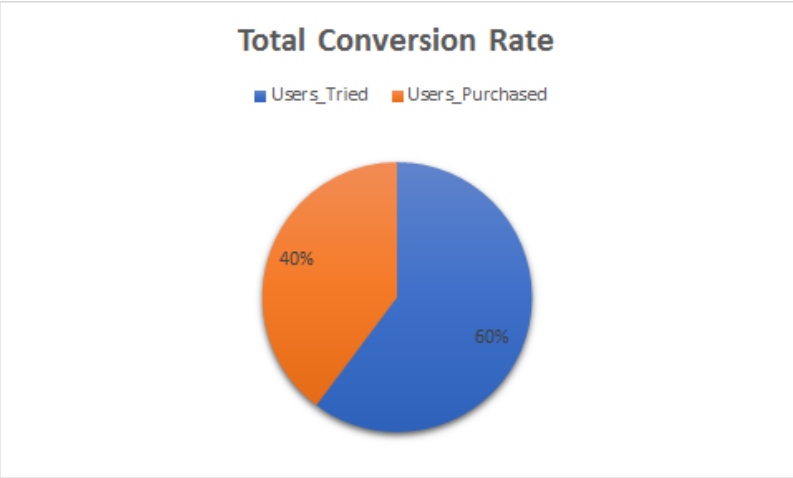
user_id	is_home_try_on	number_of_pairs	is_purchase
4e8118dc-bb3d-49bf-85fc-cca8d83232ac	TRUE	3 pairs	FALSE
291f1cca-e507-48be-b063-002b14906468	TRUE	3 pairs	TRUE
75122300-0736-4087-b6d8-c0c5373a1a04	FALSE	NULL	FALSE
75bc6ebd-40cd-4e1d-a301-27ddd93b12e2	TRUE	5 pairs	FALSE
ce965c4d-7a2b-4db6-9847-601747fa7812	TRUE	3 pairs	TRUE
28867d12-27a6-4e6a-a5fb-8bb5440117ae	TRUE	5 pairs	TRUE

3. Funnel analysis

3.1 How many users tried out glasses and ended up buying?

- Based on the “Funnels” table created, it is able to identify the Conversion Rate.
- The conversion rate from users who tried out a pair of glasses that ended up buying is represented by 40% of users.

  - Total users who tried = 750
  - Total users who bought = 495



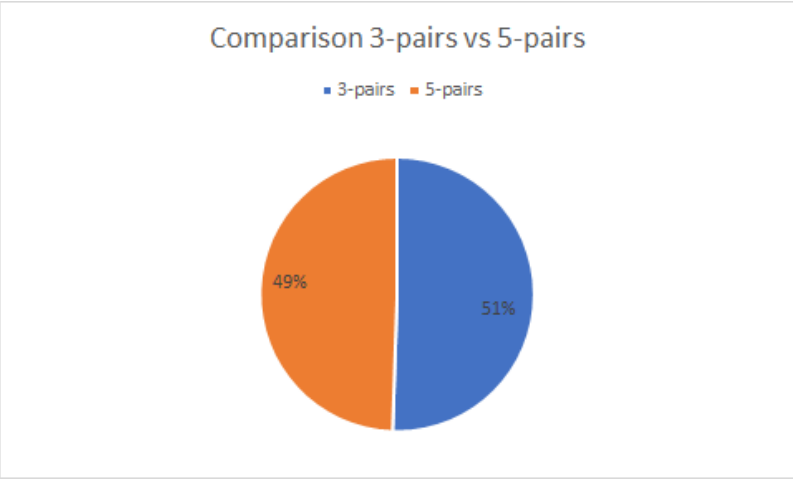

```
-- Conversion rate

WITH funnels 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 AS 'q'
LEFT JOIN home_try_on AS 'h'
   ON q.user_id = h.user_id
LEFT JOIN purchase AS 'p'
   ON p.user_id = q.user_id)
SELECT
CASE WHEN
is_home_try_on = 1
THEN 'True'
ELSE 'False'
END AS "Users_Tried",
number_of_pairs,
CASE WHEN
is_purchase = 1
THEN 'True'
ELSE 'False'
END AS "Users_Purchased"
FROM funnels;
```

3.2 How many users tried 3-pairs vs. 5-pairs of glasses?

- From all users who ended trying a pair of glasses, 51% picked “5-pairs” of glasses and 49% of users chose trying “3-pairs”

  - Total users who tried = 379
  - Total users who bought = 371
  - From 750 users who tried a pair of glasses



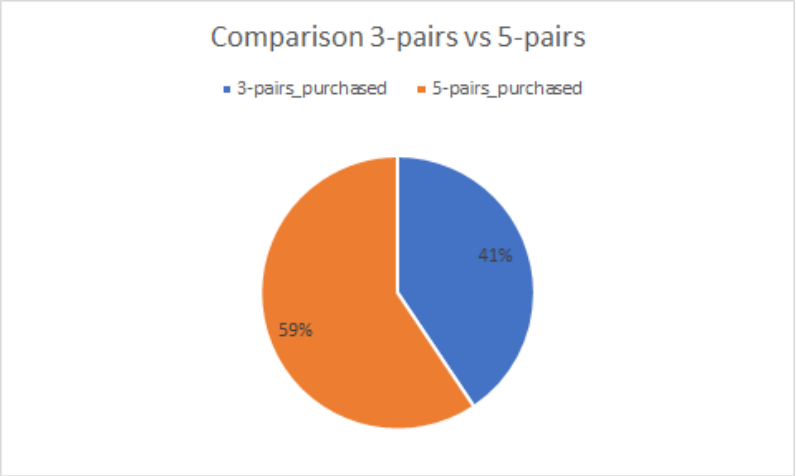

```
-- 3-pairs vs 5-pairs

WITH funnels 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 AS 'q'
LEFT JOIN home_try_on AS 'h'
   ON q.user_id = h.user_id
LEFT JOIN purchase AS 'p'
   ON p.user_id = q.user_id)
SELECT DISTINCT user_id,
CASE WHEN
is_home_try_on = 1
THEN 'True'
ELSE 'False'
END AS "Users_Tried",
number_of_pairs,
CASE WHEN
is_purchase = 1
THEN 'True'
ELSE 'False'
END AS "Users_Purchased"
FROM funnels;
```

3.3 How many users ended up buying when trying 3 pairs vs. 5 pairs?

- From users who ended up buying after trying out glasses, 41% ended up buying after trying 3 pairs of glasses and 59% finished purchasing after looking at 5 pairs of glasses

  - Number of purchases for 3-pairs = 201
  - Number of purchases for 5-pairs = 294
  - From 750 users who tried a pair of glasses.

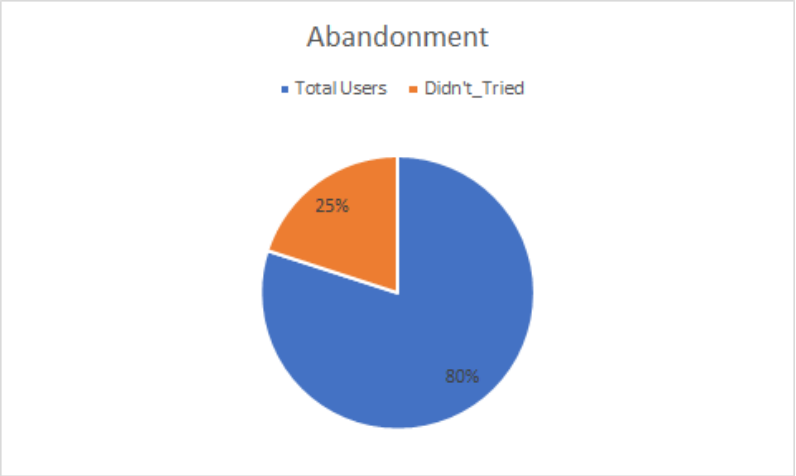


Based on this sample, users who tried 5 pairs were more susceptible to purchasing glasses.

3.4 How many users ended up buying when trying 3 pairs vs. 5 pairs?

- From total users who visited website, the abandonment rate is 20%
- Abandonment rate accounts for Users who neither tried glasses, nor ended up purchasing.

  - Total website's visitors = 1000
  - Number of abandonment = 250



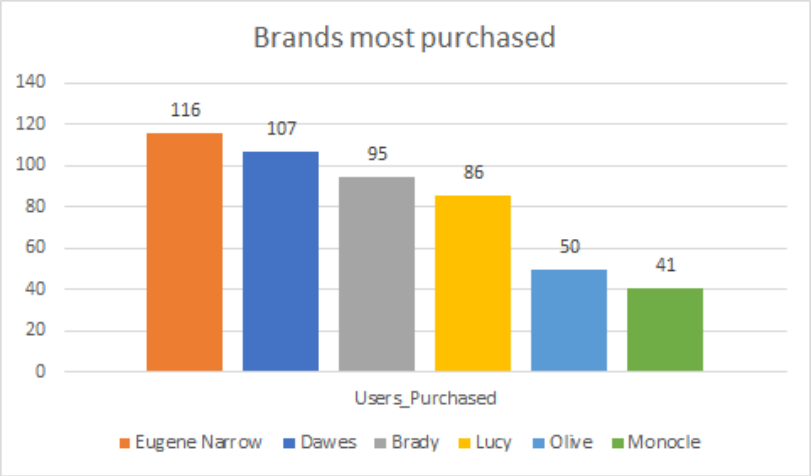
Low rate of abandonment of website indicates that it is very attractive.

4. Purchase analysis

4.1 What brands were purchased?

- From total users who purchased a pair of glasses, 23% picked the brand “Eugene Narrow”, 22% selected “Dawes” and 19% chose “Brady”.

  - Eugene Narrow - 23%
  - Dawes - 22%
  - Brady - 19%
  - Lucy - 17%
  - Olive - 10%
  - Monocle - 8%
  - From total of 495 purchases made




```
-- Brands sold

SELECT DISTINCT model_name, SUM (price) AS price
FROM purchase
GROUP BY model_name;
```

4.2 Total sales per brand?

- Total sales in this period resulted in $55,795.00
- See table for breakdown per Brand.
  - From total of 495 purchases made

```
-- Sales per brand

SELECT DISTINCT style, COUNT (user_id)
FROM purchase
GROUP BY style;
```

model_name	price($)
Brady	9,025
Dawes	16,050
Eugene Narrow	11,020
Lucy	12,900
Monocle	2,050
Olive	4,750

4.3 Total sales after trying 3-pair vs. 5-pairs?

- Total of effective purchases from users who tried 5-pairs were 19% higher then from those who tried 3 pairs of glasses.
- Preferred Brand is “Eugene Narrow”
  - 3-pair users resulted in $22,765 in purchases
  - 5-pair users totalled $33,030 in sales
  - From total of 495 purchases made

```
-- 3-pair vs 5 pair vs Brands

WITH funnels 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',
   p.style AS 'purchased_style',
   p.model_name AS 'purchased_model',
   p.price AS 'purchased_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 p.user_id = q.user_id)
SELECT DISTINCT user_id,
purchased_price,purchased_model,
```

purchased_models	3-pairs_purchased ($)	5-pairs_purchased ($)
Eugene Narrow	5,130	5,890
Olive	1,330	3,420
Dawes	7,350	8,700
Brady	3,705	5,320
Lucy	4,500	8,400
Monocle	750	1,300
TOTAL	**22,765**	**33,030**