Why Data Science?

Exploring Data with SQL - Continued

3 min

Next, Catherine wants to take a look at the churn rate.

Churn rate is the percent of subscribers to a monthly service who have canceled. For example, in January, let's say Codecademy had 1,000 learners. In February, 200 learners sign up, and 250 cancel.

The churn rate for February would be:

$$\frac{cancellations}{total\ subscribers} = \frac{250}{1000+200} = 20.8\%$$

$$cancellations = 1000+200$$

$$250 = 20.8\%$$

Catherine wants to analyze the churn rates for Codecademy for the past few months so she writes another SQL query.

Instructions

1. Checkpoint 1 Passed

1.

Click Run, to see Catherine's analysis for the churn rate in March 2017.

What recommendations would you make to Codecademy based on Catherine's analysis?

(This query might take some time to load because the pro_users table has 118,135 rows!)

In the result, there should be three columns:

- enrollments

- march_cancellations

- churn_rate

test.sqlite
```
SELECT COUNT(DISTINCT user_id) AS 'enrollments',
  COUNT(CASE
      WHEN strftime("%m", cancel_date) = '03'
      THEN user_id
  END) AS 'march_cancellations',
  ROUND(100.0 * COUNT(CASE
      WHEN strftime("%m", cancel_date) = '03'
      THEN user_id
  END) / COUNT(DISTINCT user_id)) AS 'churn_rate'
```

```sql
FROM pro_users

WHERE signup_date < '2017-04-01'

 AND (

  (cancel_date IS NULL) OR

  (cancel_date > '2017-03-01')

 );
```

Query Results		
enrollments	march_cancellations	churn_rate
16435	4165	25.0

Database Schema	
pro_users	
name	type
user_id	TEXT
signup_date	TEXT
cancel_date	TEXT

Rows: 118135