codecademy

Churn Rates with Codeflix

Learn SQL from Scratch

Sarah Chabolla

Table of Contents

1. Get familiar with Codeflix

Get familiar with Codeflix

1. How many different segments do you see?
   - 2 segments: 87 and 30.
   - **Table 1**: First 2 rows of subscription table.

```
SELECT *
FROM subscriptions
LIMIT 100;
```

Table 1

id	subscription_start	subscription_end	segment
1	2016-12-01	2017-02-01	87
2	2016-12-01	2017-01-24	87
...			

Get familiar with Codeflix - cont'd

2. Which months will you be able to calculate churn for?
    - **Table 2**: January 2017-March 2017
    - Although Codeflix opened in December 2016, users cannot cancel their subscriptions in the same month they started their subscription, so the first month we can calculate churn for is January 2017.
3. To get started, create a temporary table of months. (**Table 3**)
4. Create a temporary table, cross_join, from subscriptions and your months. Be sure to SELECT every column.
    - **Table 4**: First 2 rows of cross_join table.

Table 2

MIN(subscription_start)	MAX(subscription_start)
2016-12-01	2017-03-30

Table 3

first_day	last_day
2017-01-01	2017-01-31
2017-02-01	2017-02-28
2017-03-01	2017-03-31

Table 4

id	subscription_start	subscription_end	segment	first_day	last_day
1	2016-12-01	2017-02-01	87	2017-01-01	2017-01-31
1	2016-12-01	2017-02-01	87	2017-02-01	2017-02-28
...					

2. What is the overall churn trend since the company started?

What is the overall churn trend since the company started?

5. Create a temporary table, status, from the cross_join table you created.
    - This table shows whether or not each subscription was active during each month.
    - **Table 5**: First 4 rows of status table.

Table 5

id	month	is_active_87	is_active_30
1	2017-01-01	1	0
1	2017-02-01	0	0
1	2017-03-31	0	0
2	2017-01-01	1	0
...			

What is the overall churn trend since the company started? - cont'd

6. Add an is_canceled_87 and an is_canceled_30 column to the status temporary table. This should be 1 if the subscription is canceled during the month and 0 otherwise.
   - **Table 6**: First 2 rows of status table.

Table 6

id	month	is_active_87	is_active_30	is_canceled_87	is_canceled_30
1	2017-01-01	1	0	0	0
1	2017-02-01	0	0	0	0
...					

What is the overall churn trend since the company started? - cont'd

7. Create a status_aggregate temporary table (**Table 7**) that is a SUM of the active and canceled subscriptions for each segment, for each month.
   - Codeflix has had growth in active subscriptions month over month with the biggest leap between January and February and a much smaller increase from February to March.
   - At a glance, segment 30 has much fewer canceled subscriptions than segment 87, but in both segments, the number of canceled subscriptions increases each month.

Table 7

month	sum_active_87	sum_active_30	is_canceled_87	sum_canceled_30
2017-01-01	278	291	64	21
2017-02-01	462	518	138	37
2017-03-01	531	716	237	79

3. Compare the churn rates between user segments.

Compare the churn rates between user segments.

8. Calculate the churn rates for the two segments over the three month period. Which segment has a lower churn rate?

- Month by month churn in **Table 8**, total in **Table 9.**
- The churn rate of segment 30 is much lower than that of segment 87.
- Not only does the churn rate of segment 30 increase at a much more gradual rate, it stayed virtually the same from January to February, where segment 87 make a leap of 7 percentage points during the same time period.

Conclusion: Codeflix should focus on increasing the 30 user segment in order to build a more sustaining user base and increase their customer retention.

Table 8

month	churn_rate_87	churn_rate_30
2017-01-01	.2302	.0721
2017-02-01	.2987	.0714
2017-03-01	.4463	.1103

Table 9

churn_rate_87	churn_rate_30
0.3453	0.0893