codecademy

Capstone: Calculating Churn Rates

Learn SQL from Scratch

Example Table of Contents

1. Get familiar with Codeflix

1.1 How many months has the company been operating? Which months do you have enough information to calculate a churn rate? What segments of users exist?

The company has been operating for 4 months. This can be calculated with the MIN and MAX functions (see code on the side and table below) by comparing the first subscription date (beginning_month alias) and the last subscription date (ending_month alias).

We can only calculate churn rate for January until March (3 months). We cannot calculate it for December because we do not have the active members at the beginning of the month (first subscription on 01.12.2016) and people cannot unsubscribe in the month they signed up in.

By selecting all columns from the dataset, it becomes clear that there are two segments, 30 and 87 (see code on the side).

beginning_month	ending_month
2016-12-01	2017-03-30

```
SELECT MIN(subscription_start) AS beginning_month,
MAX(subscription_start) AS ending_month
FROM subscriptions;

SELECT *
FROM subscriptions
LIMIT 100;
```

2. What is the overall churn rate by month?

1.2 What is the overall churn rate by month? What is the overall churn trend since the company started?

The overall churn rates can be seen in the table below. They were calculated by defining the months in SQL and after cross joining them with the subscriptions, by creating further temporary tables that enabled the definition of active users and cancellations for the two segments. Finally, the churn rates were calculated.

The overall churn trend since the start of operations has been an increase in the churn rates.

month	churn_rate_30	churn_rate_87
2017-01-01	0.08	0.25
2017-02-01	0.07	0.32
2017-03-01	0.12	0.49

```
Elements of the code:
WITH months AS
                  (SELECT
                                    '2017-01-01' as first_day,
                                    '2017-01-31' as last_day
                  UNION
                  SELECT
                                    '2017-02-01' as first_day,
                                    '2017-02-28' as last_day)

cross_join AS
                  (SELECT *
                   FROM subscriptions
   CROSS JOIN months
)

 CASE
    WHEN (subscription_start < first_day)
                   AND (subscription_end > first_day OR
subscription_end IS NULL) AND (segment = 30)
                                    THEN 1
                                    ELSE 0
    END AS is_active_30,
  CASE
                   WHEN subscription_end BETWEEN first_day AND
last_day AND (segment = 87)
                   THEN 1
                   ELSE 0
                   END AS is_canceled_87

 SUM(is_active_87) AS sum_active_87,

ROUND((1.0*sum_canceled_30/sum_active_30),2) AS churn_rate_30,
```

3. Compare the churn rates between segments

1.3 Which segment of users should the company focus on expanding?

The company should focus on expanding the segment 30 if it has comparable profitability as segment 87 because it has a lower churn rate. This implies that more users stay with the product for a longer time period. This in turn implies that it is the more lucrative product to sell.

Moreover, the company should assess whether there is a way to reduce the churn rate of the other product (segment 87).

Additionally, the company must consider the reasons for the rising churn rates because in both segments, the churn rates have been increasing since the introduction of the product.