codecademy

CALCULATING CHURN RATES

DATA ANALYSIS SQL PROJECT

JINDRISKA MAZUROVA

09.10.2023

1. Introduction

What is Churn rate?

A common revenue model for SaaS (Software as a service) companies is to charge a monthly subscription fee for access to their product. Frequently, these companies aim to continually increase the number of users paying for their product. One metric that is helpful for this goal is churn rate.

Churn rate is the percent of subscribers that have canceled within a certain period, usually a month. For a user base to grow, the churn rate must be less than the new subscriber rate for the same period.

To calculate the churn rate, we only will be considering users who are subscribed at the beginning of the month. The churn rate is the number of these users who cancel during the month divided by the total number:

churn rate = cancellations / total subscribers

2. Company and dataset

Company - We analyze churn rates for two groups of Codeflix users after four months launching the service and recommend which segment represents a better long-term user base. Codeflix requires a minimum subscription length of 31 days.

Dataset -

Database Schema subscriptions	
name	**type**
id	INTEGER
subscription_start	TEXT
subscription_end	TEXT
segment	INTEGER

2.1. First we created a query to determine our **two subscription channels - 30 and 87**

```
SELECT segment, COUNT(*) FROM subscriptions
GROUP BY 1;
```

segment	count(*)
30	1000
87	1000

2.2 We determined the start and the end period for subscriptions and as next the first and the last day for churn rates with the **temporary table months**.

start	end
2016-12-01	2017-03-30

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

3. Status table

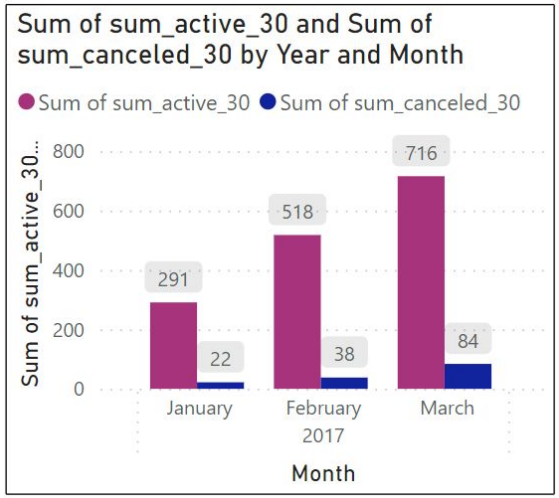
3.1. As next we created the **temporary table cross_join**, joining months and subscriptions.

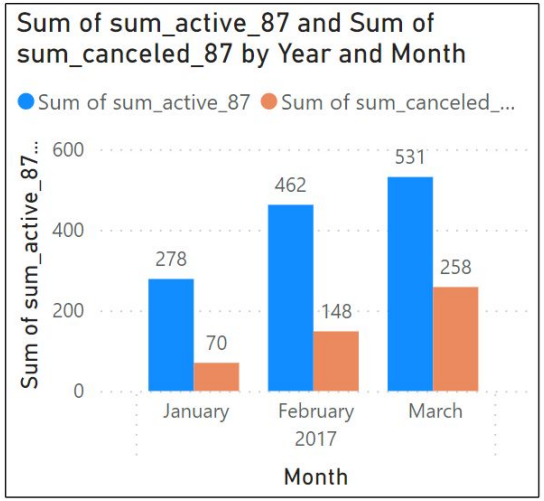
3.2. Then we created the **temporary table status**, that contains column_is_active or column_is_cancelled for showing, if the user's subscription was active or cancelled for a selected month and 0 otherwise .

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	1	0
1	2017-03-01	0	0	0	0
2	2017-01-01	1	0	1	0
2	2017-02-01	0	0	0	0
2	2017-03-01	0	0	0	0
3	2017-01-01	1	0	0	0
3	2017-02-01	1	0	0	0
3	2017-03-01	1	0	1	0
4	2017-01-01	1	0	0	0

4. Subscription status by month

4.1. Then we created a **temporary table status_aggregate**, which calculated the sum of active and canceled subscriptions for every month. You can see, that the segment 30 has much lower drop rates than the segment 87 .





5. Churn rates and conclusion

5.1. Last, we calculated **churn rates in percentage** for every segment and month.

5.2. In conclusion, our analysis of churn rates for different segments and months indicates that the segment 30 has much better churn rates. Maybe they could recommend the segment 87, what is the difference in communication with the their users. The segment 87 could also ask their users for a feedback and recommendation, what they could make better.

Year	Month	ChurnRate_30_%	ChurnRate_87_%
2017	January	8	25
2017	February	7	32
2017	March	12	49