Codeflix

Churn Analysis
Overall and by Segment

Table of Contents

1. Get familiar with Codeflix
2. Compare the churn rates between segments
3. What is the overall churn rate by month?

Get Familiar with Codeflix

1. User Segments

What segments of users exist?

There are 2 segments that can be seen when querying the first 100 rows in the table:

Segment 87 and Segment 30.

Query Results			
id	subscription_start	subscription_end	segment
1	2016-12-01	2017-02-01	87
2	2016-12-01	2017-01-24	87
3	2016-12-01	2017-03-07	87
4	2016-12-01	2017-02-12	87
5	2016-12-01	2017-03-09	87
6	2016-12-01	2017-01-19	87
7	2016-12-01	2017-02-03	87
8	2016-12-01	2017-03-02	87
9	2016-12-01	2017-02-17	87
10	2016-12-01	2017-01-01	87
11	2016-12-01	2017-01-17	87
12	2016-12-01	2017-02-07	87
13	2016-12-01	Ø	30
14	2016-12-01	2017-03-07	30
15	2016-12-01	2017-02-22	30
16	2016-12-01	Ø	30
17	2016-12-01	Ø	30
18	2016-12-02	2017-01-29	87
19	2016-12-02	2017-01-13	87
20	2016-12-02	2017-01-15	87

2. Data Available

How many months has the company been operating? Which months do you have enough information to calculate a churn rate?

According to the data, the earliest start date is 12/01/2016 and the latest start date is 03/30/2017. So, the company has been operating for 4 months.

Because a user cannot start and end their subscription in the same month, we cannot calculate churn for December. We can, however, calculate churn for January, February and March of 2017.

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

Compare churn rates between segments

To calculate the churn rate by segment:
- Determine the total number of subscribers at the beginning of the month for each segment;
- Determine the total number of those subscribers who cancelled during the month;
- Calculate the percentage as total cancellations for the month divided by the total subscribers at the beginning of the month for each segment.

3. Create a table of Months

To calculate the churn trend by segment, the first step is to create a table of months of available data.

Query Results	
first_day	**last_day**
2017-01-01	2017-01-31
2017-02-01	2017-02-28
2017-03-01	2017-03-31

4. Create a table of months joined to subscriptions

The second step to calculate the churn trend is to create a new temporary table by joining the months table to the subscriptions table as illustrated in the query results shown.

first_day	last_day	id	subscription_start	subscription_end	segment
2017-01-01	2017-01-31	1	2016-12-01	2017-02-01	87
2017-02-01	2017-02-28	1	2016-12-01	2017-02-01	87
2017-03-01	2017-03-31	1	2016-12-01	2017-02-01	87
2017-01-01	2017-01-31	2	2016-12-01	2017-01-24	87
2017-02-01	2017-02-28	2	2016-12-01	2017-01-24	87
2017-03-01	2017-03-31	2	2016-12-01	2017-01-24	87
2017-01-01	2017-01-31	3	2016-12-01	2017-03-07	87
2017-02-01	2017-02-28	3	2016-12-01	2017-03-07	87
2017-03-01	2017-03-31	3	2016-12-01	2017-03-07	87
2017-01-01	2017-01-31	4	2016-12-01	2017-02-12	87
2017-02-01	2017-02-28	4	2016-12-01	2017-02-12	87
2017-03-01	2017-03-31	4	2016-12-01	2017-02-12	87
2017-01-01	2017-01-31	5	2016-12-01	2017-03-09	87
2017-02-01	2017-02-28	5	2016-12-01	2017-03-09	87
2017-03-01	2017-03-31	5	2016-12-01	2017-03-09	87

Query Results

5. Active status of each customer at the beginning of the month by segment

A determination must be made for each subscriber as to whether or not the subscriber was active at the beginning of the month and to which segment the subscriber is assigned.

id	month	is_active_87	is_active_30
1	2017-01-01	1	0
1	2017-02-01	0	0
1	2017-03-01	0	0
2	2017-01-01	1	0
2	2017-02-01	0	0
2	2017-03-01	0	0
3	2017-01-01	1	0
3	2017-02-01	1	0
3	2017-03-01	1	0
4	2017-01-01	1	0
4	2017-02-01	1	0
4	2017-03-01	0	0
5	2017-01-01	1	0
5	2017-02-01	1	0
5	2017-03-01	1	0
6	2017-01-01	1	0
6	2017-02-01	0	0
6	2017-03-01	0	0

Query Results

6. Cancelled status of each customer by segment

A determination must be made for each subscriber as to whether or not the subscriber cancelled during the month and to which segment the subscriber is assigned.

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	2017-02-01	1	0	1	0
4	2017-03-01	0	0	0	0
5	2017-01-01	1	0	0	0
5	2017-02-01	1	0	0	0
5	2017-03-01	1	0	1	0
6	2017-01-01	1	0	1	0
6	2017-02-01	0	0	0	0
6	2017-03-01	0	0	0	0

Query Results

7. Aggregate status by segment by month

Total the active and cancelled subscribers by segment by month.

Query Results				
month	sum_active_87	sum_active_30	sum_canceled_87	sum_canceled_30
2017-01-01	278	291	70	22
2017-02-01	462	518	148	38
2017-03-01	531	716	258	84

8. Calculate monthly churn by segment

Calculate the churn rates for the two segments over the three month period.

Which segment has a lower churn rate?

As you can see from the query results, Segment 30 has the lower churn. Codeflix should focus their efforts on expanding this segment.

Query Results		
month	churn_87	churn_30
2017-01-01	0.251798561151079	0.07560137457044
2017-02-01	0.32034632034632	0.07335907335907
2017-03-01	0.485875706214689	0.1173184357541

Overall churn trend

To calculate the churn rate:
- Determine the total number of subscribers at the beginning of the month;
- Determine the total number of those subscribers who cancelled during the month;
- Calculate the percentage as total cancellations for the month and dividing it by the total subscribers at the beginning of the month.

Calculate overall churn trend by month

To calculate the overall churn trend by month, the active and cancelled subscribers must be determined by month without regard to the segments.

The results show that March had the highest churn.

month	churn
Query Results	
2017-01-01	0.161687170474517
2017-02-01	0.189795918367347
2017-03-01	0.274258219727346