codecademy

Codeflix: Churn Analysis

Learn SQL from Scratch

Sean Pyne

08/28/18

Table of Contents

1. Company Background

1.1 Company Background

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

- The company has been operating from December of 2016 to March of 2017. So 4 months.

- Subs are a min of a month so you can't calculate churn in Dec, all other months are possible.

#2 - What segments of users exist?
- There are two user segments – 87 and 30

```
#1
SELECT Min(subscription_start),
Max(subscription_start),Min(subscription_end),
Max(subscription_end)
FROM subscriptions;


#2
SELECT DISTINCT segment
FROM subscriptions;
```

2. Churn Analysis

2.1 Overall Churn Rates

#1 – What is the overall churn trend since the company started?

- Since the beginning of the company churn rates have steadily increased.

Month	Active	Canceled	Churn Rate
January	569	92	0.16
February	980	186	0.19
March	1247	342	0.27

```
#create month table and cross join months with
subscriptions to see the status of a sub in a given
month - is active or is cancelled then aggregate to
see monthly stats + Churn rate

--after data cleaning, aggregate and calculate churn

. . .
status_aggregate AS
(SELECT  month,  SUM(is_active) as active,
   SUM(is_canceled) as canceled
FROM status
GROUP BY month)
SELECT  month, active, canceled,
   1.0 *  canceled/active AS churn_rate
FROM status_aggregate;
```

2. Churn Analysis by Segment

3.1 Churn Rates By Segment

#1 – Compare the Churn rates by user segment.
- Churn rates tell a very different story broken out by segment as seen in the table below

#2 – Which segment should the company focus on expanding?
- The company should clearly target the 30 segment as they've seen considerably more retention in that segment.

Month	Churn Rate - 30	Churn Rate - 87
January	0.076	0.25
February	0.073	0.32
March	0.12	0.49

```
#Calculating churn by month has already been seen in
the previous slide. So this one I will break out the
addition of segments.

. . .
CASE
    WHEN (segment IS 30)
    AND (subscription_start < first_day)
    AND (subscription_end > first_day
        OR subscription_end IS NULL)THEN 1
    ELSE 0
END as is_active_30,
CASE
    WHEN segment IS 30
    AND subscription_end BETWEEN first_day
    AND last_day THEN 1
    ELSE 0
END as is_canceled_30
. . .

That is just one segment, repeat for the 87 segment--
after data cleaning, aggregate and calculate churn by
segment
```

code|cademy

Thanks!

Learn SQL from Scratch
Sean Pyne
08/28/18