SQL Capstone - Churn Rates

Learn SQL from Scratch – Jinwen Chen

You can use these templates in the capstone presentation for Learn SQL from Scratch!

Copy the ones you'd like to use and write in your own content.

codecademy

Churn Rates

Learn SQL from Scratch

By Jinwen Chen

Questions answered

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

1. Get familiar with Codeflix

1.1 Get familiar with the table structure of subscriptions

By selecting everything column from the subscriptions table with the limit of 100 rows, I can visualize the structure of the table, column headers and data format within each column.

Query:
select * from subscriptions limit 100;

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

1.1 Get familiar with the table structure of subscriptions

-- How many months has the company been operating?

Answer: the company has been operating since December 2016 and our data includes its operation results for Dec 2016, Jan 2017, Feb 2017 and Mar 2017.

We would either select the unique subscription start date to visualize the data. Or we could select the earliest and latest subscription start date to get a sense of the range.

min(subscription_start)	max(subscription_start)
2016-12-01	2017-03-30

```
--Question 2
select distinct subscription_start
from subscriptions
order by subscription_start asc;

--Or
select min(subscription_start),
max(subscription_start) from subscriptions;
```

1.1 Get familiar with the table structure of subscriptions

-- Which months do you have enough information to calculate a churn rate?

Based on the information provided for the prior question, although we have four months of operation data, we could only get the churn rate for Jan, Feb and Mar in 2018, because the user can not cancel their service within the same month of sign-up.

1.1 Get familiar with the table structure of subscriptions

-- What segments of users exist?

Answer: two segments exists,

We would select the unique segment from subscription table.

segment
87
30

```
select distinct segment from subscriptions;
```

2. What is the overall churn rate?

2. Overall churn rate by segment

- Step1: Create temp table ‘months’ to list out first day and last day for Jan, Feb and Mar 2018.
- Step 2: Create temp table ‘seg’ to list out unique segments from subscriptions.
- Step 3: Create temp table ‘cross_join’ subscriptions table, months table and seg table. So each record in subscription table is displayed 6 times with different months and seg combinations
- Step 4: Create temp table ‘status’, with id, first day, seg, subscription_start, subscription_end, segment, ‘is_active’, ‘is_canceled’
- Step 5: Sum up is_active (sum_active) and is_canceled (sum_canceled)
- Step 6: calculate churn rate by using sum_canceled / sum_active

churn_rates
0.221745350500715

3. Compare the churn rates between segments

2.1 Overall churn rate by segment

- Conclusion: segment 30 outperform segment 87 by having lower churn rate.
- Step1: Create temp table ‘months’ to list out first day and last day for Jan, Feb and Mar 2018.
- Step 2: Create temp table ‘seg’ to list out unique segments from subscriptions.
- Step 3: Create temp table ‘cross_join’ subscriptions table, months table and seg table. So each record in subscription table is displayed 6 times with different months and seg combinations
- Step 4: Create temp table ‘status’, with id, first day, seg, subscription_start, subscription_end, segment, ‘is_active’, ‘is_canceled’
- Step 5: Sum up is_active (sum_active) and is_canceled (sum_canceled) for each segment
- Step 6: calculate churn rate by using sum_canceled / sum_active

segment	churn_rates
30	0.0944262295081967
87	0.374508261211644

2.1 Overall churn rate by segment by month

- Conclusion: segment 30 outperform segment 87 by having lower churn rate. For both segments, March outperform January.
- Step1: Create temp table ‘months’ to list out first day and last day for Jan, Feb and Mar 2018.
- Step 2: Create temp table ‘seg’ to list out unique segments from subscriptions.
- Step 3: Create temp table ‘cross_join’ subscriptions table, months table and seg table. So each record in subscription table is displayed 6 times with different months and seg combinations
- Step 4: Create temp table ‘status’, with id, first day, seg, subscription_start, subscription_end, segment, ‘is_active’, ‘is_canceled’
- Step 5: Sum up is_active (sum_active) and is_canceled (sum_canceled) for each segment and each month
- Step 6: calculate churn rate by using sum_canceled / sum_active

month	segment	churn_rates
2017-01-01	30	0.0756013745704467
2017-02-01	30	0.0733590733590734
2017-03-01	30	0.11731843575419
2017-01-01	87	0.251798561151079
2017-02-01	87	0.32034632034632
2017-03-01	87	0.485875706214689