codecademy

Codeflix
User Churn Analysis

Data analysis with SQL

Catherine Kelly

Introduction

- Codeflix, a startup video streaming company, wants to analyse how well the business is doing. In particular, they are interested in the subscription churn rates of users acquired through two different channels.
- The owners would like to answer the following questions

  1) What is the overall churn trend since the company started?

  2) Comparison of the churn rates between user segments

  3) Which segment of users should the company focus on expanding?

Overall subscription churn rate

- The churn rate over the first two months was low at 16.2 and 19.0 % for January and February, respectively.
- The churn rate was found to increase each month with the largest change experienced in March 2017.
- This suggests that something is going wrong with the users experience

Month	Churn rate (%)
January 2017	16.2
February 2017	19.0
March 2017	27.4

Overall subscription churn rate – SQL 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
  UNION
  SELECT
  '2017-03-01' AS first_day,
  '2017-03-31' AS last_day),
cross_join AS (
  SELECT *
  FROM subscriptions
  CROSS JOIN months),
status AS (
SELECT id, first_day AS month,
CASE
  WHEN (subscription_start < first_day) AND
(subscription_end > first_day OR
subscription_end IS NULL) THEN 1
  ELSE 0
  END AS is_active,
```

```
CASE
  WHEN (subscription_end BETWEEN first_day AND
last_day) THEN 1
  ELSE 0
  END AS is_canceled
FROM cross_join),
status_aggregate AS (
SELECT month, SUM(is_active) AS sum_active,
SUM(is_canceled) AS sum_canceled
FROM status
GROUP BY month),
churn_rate AS(
  SELECT month, 100.0 * sum_canceled /
sum_active AS churn_rate
  FROM status_aggregate
  GROUP BY month)
SELECT *
FROM churn_rate;
```

Subscription churn rate by segment

- The churn rate was low for users in segment 30 and remained consistent over the first two months
- The churn rate was noticeably higher for users in segment 87 and a large increase in the churn rate was experienced each month.
- Although users in segment 30 had a low churn rate, a noticeable increase again occurred in March 2017

Month	Churn rate Segment 87 (%)	Churn rate Segment 30 (%)
January 2017	25.2	7.6
February 2017	32.0	7.3
March 2017	48.6	11.7

Subscription churn rate by segment – SQL 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
  UNION
  SELECT
  '2017-03-01' AS first_day,
  '2017-03-31' AS last_day),
cross_join AS (
  SELECT *
  FROM subscriptions
  CROSS JOIN months),
status AS (
SELECT id, first_day AS month,
CASE
  WHEN (segment = 87) AND (subscription_start <
first_day) AND (subscription_end > first_day OR
subscription_end IS NULL) THEN 1
  ELSE 0
  END AS is_active_87,
CASE
  WHEN (segment = 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 = 87) AND (subscription_end BETWEEN
first_day AND last_day) THEN 1
  ELSE 0
  END AS is_canceled_87,
CASE
  WHEN (segment = 30) AND (subscription_end BETWEEN
first_day AND last_day) THEN 1
  ELSE 0
  END AS is_canceled_30
FROM cross_join),
status_aggregate AS (
SELECT month, SUM(is_active_87) AS sum_active_87,
SUM(is_active_30) AS sum_active_30,
SUM(is_canceled_87) AS sum_canceled_87,
SUM(is_canceled_30) AS sum_canceled_30
FROM status
GROUP BY month),
churn_rate AS(
  SELECT month, 100.0 * sum_canceled_87 /
sum_active_87 AS churn_rate_87, 100.0 *
sum_canceled_30 / sum_active_30 AS churn_rate_30
  FROM status_aggregate
  GROUP BY month)
SELECT *
FROM churn_rate;
```

Conclusions

- The overall churn rate over the first two months was low, however a large increase was observed in March 2017
- The churn rate was noticeably lower for the users in segment 30 compared to segment 87
- Codeflix need to look at what happened in March 2017 that may have increase the churn rate
- The company also needs to examine why they are seeing a higher churn rate with users in segment 87
- Users in segment 30 appear to like Codeflix and therefore the company should target marketing to this demographic