Calculating Churn Rates

Learn SQL From Scratch – Capstone project
Mar 12, 2017 – May 06, 2017

By: Christopher W. Svensson

For: codecademy

Project Goals:

1. Examine SQL table data by months and segments.
2. Calculate overall churn rates.
3. Compare churn rates between segments.
4. Expand query to allow for multiple segments.

Examine SQL Table Data by Months and Segments.

Data Overview

We can start with a reduced, quick look at the first 100 rows of data, and other queries to see what there is to work with. Here, four columns (whose titles are shown below) have been selected. Some things that stand out are:

- Various subscription start dates from 2016-12-01 through 2017-03-30. This represents 4 months of data.
- Subscription end dates from 2017-01-01 through 2017-03-31, therefore churn rate is for **3 months only**.
- Null values in the subscription end data, indicating segments that are still active after the data sample.
- Only two segments; **30 and 87**.
- With no subscriptions ending before January of 2017, that is the start month for data analysis.

Random rows are shown here:

id	subscription_start	subscription_end	segment
34	2016-12-02	2017-02-06	30
220	2016-12-12	NULL	30
710	2017-01-10	2017-03-07	87
1938	2017-03-26	NULL	87

Calculate Overall Churn Rates.

Churn in Aggregate Form

- The churn, shown as a percent for each of the three months covered by the data. For decimal representation, each churn rate may by divided by 100.
- In general, there is a progression from month to month, with the trend toward higher churn rates for each subsequent month.
- Without segment breakdown, it is impossible to see the contribution, or effectiveness of each segment.

Month	Percent Churn
2017-01-00	16.14
2017-02-00	18.88
2017-03-00	27.16

Compare Churn Rates Between Segments.

Churn by segment

The churn rate, expressed as a percent (not a decimal) for each of the two segments, by month.

- For each segment, the churn rate for 30 is significantly less than that for segment 87.
- While the trend is up for each month, segment 30's rate actually drops a bit for February.
- Clearly, segment 30 is by far the better of the two.

Month	Percent Churn Segment 87	Percent Churn Segment 30
2017-01-00	25.09	7.56
2017-02-00	31.69	7.34
2017-03-00	47.69	11.7

Expand Query to Allow For Multiple Segments.

Churn for Multiple Segments.

By simply removing the hard coding of the segment numbers, and adjusting the queries to hold each segment data item, more segments will automatically show up in the results. A hard-coded query will still show segments 30 and 87, but all other segments would just get ignored. There are two ways of sorting the rows:

- By segment first, then by months, or
- By month first, then by segment (shown here).

To best compare segment by segment, sorting by month yields results that are easier to understand.

Month	Segment	Percent Churn
2017-01-00	30	7.56
2017-01-00	87	25.09
2017-02-00	30	7.34
2017-02-00	87	31.69
2017-03-00	30	11.7
2017-03-00	87	47.69