

## Life Satisfaction

We are interested in predicting life satisfaction on Twitter in English-speaking US citizens. We used monthly Twitter data from the U.S. (Jan 2017 – March 2020) to estimate life satisfaction reports on social media. We used Twitter data because no polling data for monthly life satisfaction over the required time period exists. Estimates are calculated using a ridge regression model trained on *latent Dirichlet allocation* topics, selected using univariate feature selection and dimensionally reduced using randomized principal component analysis, to predict Cantril ladder life satisfaction scores. Given that Cantril ladder scores range from 0 to 10, the current estimates have an equivalent range. Such twitter-based estimates closely follow nationally representative polls (1).

## Reference:

D. Witters, J. Harter, “In U.S., Life Ratings Plummet to 12-Year Low” (2020), (available at <https://news.gallup.com/poll/308276/life-ratings-plummet-year-low.aspx>).