**Project Team: Safaa, Christa, Raphael, Joe, and Kristen**

**Github Link:** [**https://github.com/marinellijoseph/Project3\_World\_Happiness**](https://github.com/marinellijoseph/Project3_World_Happiness)

**Project Theme**: *World Happiness – Survey, Social Media sentiment analysis, News Articles and Corruption Perception index*

**Potential Data Sources:**

* World Happiness survey data
  + <https://www.kaggle.com/unsdsn/world-happiness>
* Corruption Perceptions Index
  + <https://www.transparency.org/research/cpi/overview>
* Twitter API
* Recent News Articles from New York Times for top two countries and bottom two countries from 2017 ranking
  + Norway: <https://www.nytimes.com/topic/destination/norway>
  + Denmark: <https://www.nytimes.com/topic/destination/denmark>
  + Burundi: <https://www.nytimes.com/topic/destination/burundi>
  + Central African Republic: <https://www.nytimes.com/topic/destination/central-african-republic>

Note: Due to the diversity of data sources, we expect that some data management steps would be needed in order to standardize and merge data

**Research topic and rational:**

Analysis of the world happiness survey data and contrasting it with corruption, news, and social media. Our goal is to create a dashboard using Tableau with the results of the analysis including maps and graphical displays. Our data focus will be on 2017 survey data but we might consider looking at trend analysis and try predicting the drivers for rank changes and 2018 best and worst.

Some of the research questions might cover the following:

* Which countries do score higher in happiness according to different data sources
* Clustering countries by Happiness indicators using K-means
* Text analytics on news articles to explore the content of the first page of news in top 2 and bottom two countries
* Explore correlation between happiness and corruption