**Group 7: •** Aggarwal, Brahm, b3aggarw **•** Chellamuthu, Shanmuga, spchella **•** Mammadov, Rashad, r2mammad **•** Mazloomi, Rod, rmazloom **•** Sanchez, Monica, m7sanche **•** Sheikh, Hassan, h25sheik

**Group Assignment: Statistical Data Analysis**

**Topic Proposal**

**Research question:** What factors are more important for national happiness? Are these factors constant over time?

**Data source:** For this project, we will use data from the World Happiness Report found in [kaggle](https://www.kaggle.com/unsdsn/world-happiness). The Report is a publication of the Sustainable Development Solutions Network, that was launched in 2012 and has been published annually till 2020, excluding 2014. The Report uses data from the Gallup World Poll and ranks countries by how happy their citizens perceive themselves. The rankings are based on responses to the main life evaluation questions, in which participants are asked to rate their lives on a 0 (worst possible life) to 10 (best possible life) scale. These rankings are stablished in comparison to Dystopia, an imaginary county with the world’s least-happy people among the main six key variables: levels of GDP, life expectancy, generosity, social support, freedom, and corruption.

**Database description:** [kaggle's](https://www.kaggle.com/unsdsn/world-happiness) dataset includes data from the 2015-19 waves of the World Happiness Report. Each wave contains the national happiness score and the country ranking. Additionally, the dataset contains the contribution to happiness evaluation from six main factors: levels of GDP, life expectancy, generosity, social support, freedom, and corruption. The dataset also includes a variable called ‘Dystopia residual’, that reflects the extent to which the six main factors over- or under-explain the average 2014-2016 life evaluations.