Data Science Fundamentals: Project Proposal
What Makes People in a Country Happy?

Objective: To use additional data sources and measurements to produce new indicators of happiness around the world. Secondarily, to validate the self-reported indicators of happiness and well-being around the world as measured by the World Happiness Report.

Background:

The World Happiness Report (WHR) (Helliwell et al, 2020) is an annual report compiled by a global research group in collaboration with the United Nations, that attempts to measure the happiness of citizens in each nation. The report is a product of the Sustainable Development Solutions Network (a program of the United Nations), in conjunction with Columbia University, the London School of Economics, the University of British Columbia, and recently the University of Oxford. The report uses a variety of indicators to measure a country's happiness, including objective measures such as GDP per capita and life expectancy at birth. Data for these metrics are taken from the World Bank's World Development Indicators (for GDP) and the World Health Organization (for life expectancy). The self-reported measurements for the Report come from Gallup World Poll data.

The following metrics are used to generate a country's happiness index according to the World Happiness Report (WHR 2020, Figure 2.1, Table 2.1):

1. GDP Per Capita (Measured in terms of Purchasing Power Parity and logged)
2. Life Expectancy at Birth
3. Self-Reported Responses to Gallup World Poll questions on the following variables:
    a. Social Support
    b. Freedom to Make Life Choices
    c. Generosity
    d. Perceptions of Corruption
    e. Positive Affect
    f. Negative Affect

The WHR and its indicators of national happiness is an authoritative voice on national happiness measurements in cross-national research. However, it has also been subject to a wide range of critiques, both on methodological grounds and theoretical grounds. Some argue that the way the report formats its indices (by using a simple additive method) is flawed, and other models produce better measurements (Carlsen, 2019; Carlsen 2017), while others have attempted to utilize more computationally sophisticated measures to study national happiness (Dixit, Chaudhary, and Sahni, 2020). Other scholars have attempted to measure alternative indices of happiness that incorporate features specific to marginalized communities. The Gay Happiness Index (Berggren, Bjørnskov and Nilsson, 2016) was created to study factors that contribute specifically to gay men's happiness, while others have studied the relationship of happiness indicators to gender equality measurements (Kabene et al, 2017).

Approach:

1. Incorporate additional features to create new happiness rankings

Our main goal is to attempt to create better happiness ratings by incorporating additional indicators that are not factored into the World Happiness Report's main country indices. The indicators used to create the country indices on the world happiness report are fairly limited, and predominantly rely on self-report data collected by Gallup World Poll, and we are interested in complementing these self-report measures with additional features. We attempt to integrate a more expansive range of objective metrics to complement the self-reported well-being of country citizens.

Some potential indicators that we plan to implement into our rankings are as follows: (Note: This is a preliminary list, which may change as our approach is finalized)

- Maternal Mortality Rate
- Adolescent Birth Rate
- % Female/% Male Enrolled in Secondary Education
- % of Parliamentary Seats Held by Women
- Adult Literacy Rate
- % Natural Resource Depletion
- Life Expectancy at Birth
- Income Inequality (GINI Coefficient)
- Female and Male Suicide Rates

With some initial analysis, it can be said that most of these variables are correlated to the current ranks of the countries in some way. Considering the Life expectancy at birth for all countries, it is seen that on average, a person has a life expectancy of around 73 years at birth around the world. This has been plotted to a graph on the right. The x-axis denotes the rank of the country and the y-axis represents the <u>life expectancy at birth</u>. This shows a good correlation for the use of this attribute for our analysis. A similar trend was observed with some other variables like Mortality Rate and GDP per Capita.
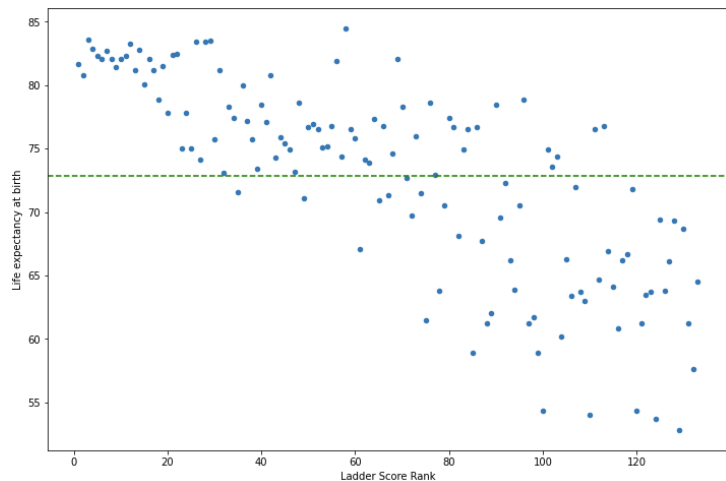
On the contrary, while understanding how suicide numbers in a country relate to its happiness, we saw no correlation in them. This is rather surprising because when checked for only female suicides, it seems that better-ranked countries have them below the average, but
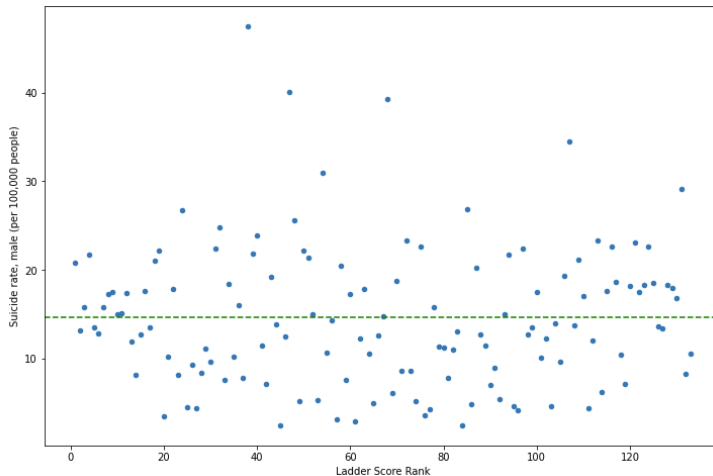
male suicides seem to not affect the current ranking in any way. This does not make sense as there have been studies conducted showing how suicides do affect a country's happiness (Daly and Oswald; 2011). This can be seen in the figure plotted on the left, where the x-axis is the rank of the countries, the y-axis is the <u>number of male suicides per 100k people</u>, while the green line represents the average. As part of the project, we intend to research if such a variable should matter when scoring a country's happiness and try to accommodate them into our scoring functions.

2. *Secondarily, validate self-report measurements on a subset of countries using new data sources*

The World Happiness Report's Indices are calculated in large part from self-reported data on citizens' well-being collected by Gallup World Poll data. An additional goal of this project is to validate those self-report measurements since previous research on self-report survey instruments has shown self-report data to be unreliable in some cases. Our goal is to use alternative sources of data to ascertain whether these happiness indices are robust measurements when different data sources are used. Additionally, we feel that, since a large majority of the self-report measurements were calculated simply by averaging responses to a binary question, alternative sources of data that utilize different metrics to study happiness may provide better indicators of subjective well-being.

Since Gallup World Poll does not provide access to its data except to large organizations with deep pockets, we turn to alternative public sources of self-report well-being measurements. Our attempts to validate existing self-report data utilize two sources of data:

● *The World Values Survey:* Extensively used in cross-national research and studies on happiness (Bruni and Stanca, 2006; 2008; Bixter, 2015; Leigh and Wolfers, 2006), the World Values Survey collects data on a wide variety of social, political, economic and well-being indicators by sampling individuals from a wide range of countries. The most recent wave of the World Values Survey began in 2017 and is ongoing, however, we are currently able to get data from individuals in 48 different countries. While Gallup world Poll data utilizes binary metrics to study outcomes of interest, many variables on the World Values Survey are measured using Likert Scale style questions, which will likely provide a better measure of how happy a nation is than a binary metric.
● *The European Social Survey:* Another cross-national survey specific to Europe, this study is conducted every two years and has also been widely used in previous research (Bilsky, Janik and Schwartz, 2011; Davidov, 2010; Georgellis, Tsitsanis, and Yin, 2009). This

survey includes Likert-scale measurements of happiness similar to the World Values Survey, and by using this dataset we can get a complete picture of happiness across European Union member states (27 nations).

Top10 Ladder - Seats held in parliament by Women

Top10 Support - Life Expectancy at Birth
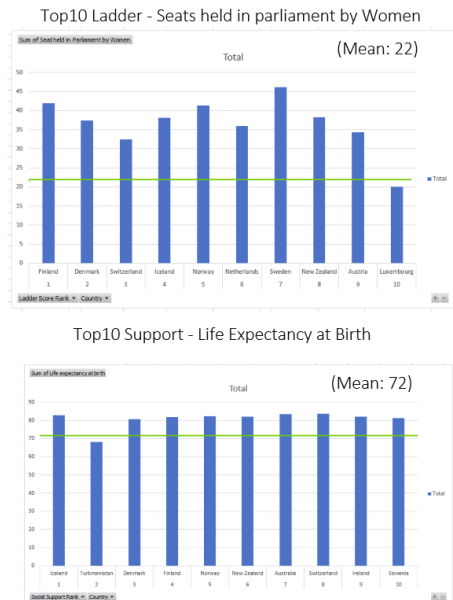
Comparing different data sources to the existing self-report data on national happiness allows us to understand to what extent these measurements are robust. While comparing the currently ranked countries with a few well-correlated variables, we come across some countries that make the Top 10 but have a negatively related value. This helps to understand the importance of understanding the scoring functions used by the report, as there are various variables affecting the overall rank. Producing comparable scoring functions in our analysis shall also help with the validation. Global happiness rankings should not change very much depending on the data sources we use to measure self-reported happiness, and if they do, there is an argument to be made for adjusting the methodology of the World Happiness Report to incorporate additional metrics or switch from binary measurements of happiness to more expansive ones.

The final outcome from the project would contain a .ipynb notebook and all the used data sources. The .ipynb files would showcase all the steps needed to clean the data, perform any needed analysis, feature generation, and visualizations, to validate our process and describe how new features could affect the overall ranking. Along with this, we plan to make a small web application that would contain a world SVG with colored countries according to their ranks. This can help to compare our results with the previous report's ranks and having an option to select some of the new features would

help make this visualization better. Future work can build on our attempts to validate the World Happiness Report's rankings in several ways, but most notably by looking at additional methods of calculating happiness indices. Previous work has suggested that using an additive index as the WHR does may not be the best way to calculate these measurements, and other methods might prove fruitful.

References:

[1] Berggren, N., Bjørnskov, C., & Nilsson, T. (2017). What Aspects of Society Matter for the Quality of Life of a Minority? Global Evidence from the New Gay Happiness Index. *Social Indicators Research*, *132*(3), 1163–1192. https://doi.org/10.1007/s11205-016-1340-3

[2] Bilsky, W., Janik, M., & Schwartz, S. H. (2011). The Structural Organization of Human Values-Evidence from Three Rounds of the European Social Survey (ESS). Journal of Cross-Cultural Psychology, 42(5), 759–776. https://doi.org/10.1177/0022022110362757

[3] Bixter, M. T. (2015). Happiness, political orientation, and religiosity. Personality and Individual Differences, 72, 7–11. https://doi.org/10.1016/j.paid.2014.08.010

[4] Bruni, L., & Stanca, L. (2006). Income aspirations, television and happiness: Evidence from the world values survey. Kyklos, 59(2), 209–225.

[5] Bruni, L., & Stanca, L. (2008). Watching alone: Relational goods, television and happiness. Journal of Economic Behavior & Organization, 65(3–4), 506–528. https://doi.org/10.1016/j.jebo.2005.12.005

[6] Carlsen, L. (2018). Happiness as a sustainability factor. The world happiness index: A posetic-based data analysis. *Sustainability Science*, *13*(2), 549–571. https://doi.org/10.1007/s11625-017-0482-9

[7] Carlsen, L. (2020). How Happy Are we Actually? A Posetic Analysis of the World Happiness Index 2016–2019 Denmark as an Exemplary Case. *International Journal of Community Well-Being*, *3*(3), 311–322. https://doi.org/10.1007/s42413-019-00044-8

[8] Davidov, E. (2010). Testing for comparability of human values across countries and time with the third round of the European Social Survey. International Journal of Comparative Sociology, 51(3), 171–191. https://doi.org/10.1177/0020715210363534

[9] Mary C. Daly, Andrew J. Oswald, Daniel Wilson & Stephen Wu (2011). University of Warwick. "Happiest places have highest suicide rates, new research finds." ScienceDaily. ScienceDaily, 21 April 2011. www.sciencedaily.com/releases/2011/04/110421082641.htm

[10] Dixit, S., Chaudhary, M., & Sahni, N. (2020). Network Learning Approaches to study World Happiness. *ArXiv:2007.09181 [Cs, Stat]*. http://arxiv.org/abs/2007.09181

[11] Georgellis, Y., Tsitsianis, N., & Yin, Y. P. (2009). Personal Values as Mitigating Factors in the Link Between Income and Life Satisfaction: Evidence from the European Social Survey. Social Indicators Research, 91(3), 329–344. https://doi.org/10.1007/s11205-008-9344-2

[12] Helliwell, John F., Richard Layard, Jeffrey Sachs, and Jan-Emmanuel De Neve, eds. 2020. World Happiness Report 2020. New York: Sustainable Development Solutions Network https://worldhappiness.report/ed/2020/

[13] Kabene, S., Baadel, S., Jiwani, Z., & Lobo, V. (2017). *Women in Political Positions and Countries' Level of Happiness*. *18*(4), 10. https://vc.bridgew.edu/jiws/vol18/iss4/15/

[14] Leigh, A., & Wolfers, J. (2006). Happiness and the human development index: Australia is not a paradox. Australian Economic Review, 39(2), 176–184.