Comparing World Quality of Life Measures: Parametric vs. Nonparametric Approaches

Katherine M. Prioli

December 20, 2018

Abstract

BACKGROUND Increasing globalization has generated interest in comparing countries by key quality of life (QoL) measures such as life expectancy, gender equality, and happiness, among others. When considering how countries compare by QoL, it is vital to understand how the measures are related, if at all, and in what ways the statistical tests chosen affect the results. METHODS A dataset containing country names and selected QoL measures for 2016 was explored through descriptive statistics and sensible univariate and bivariate visualizations. Six sets of hypotheses pertaining to relationships within the data were generated and tested with both nonparametric methods and their parametric equivalents, and results of these paired analyses were compared. RESULTS Text goes here CONCLUSION text goes here

Background

With increasing economic globalization, a natural topic of interest is how the world's nations compare with respect to quality of life (QoL). Several organizations monitor global QoL indicators and report single-dimension or aggregate values for indicators of interest. For example, the World Bank reports Gross Domestic Product (GDP), which is a single-dimension indicator often strongly predictive of QoL in a given country (The World Bank (2018)). Additionally, the World Health Organization reports infant mortality rate, life expectancy at birth, and life expectancy at 60 years of age (World Health Organization (2018b), World Health Organization (2018a)).

Other quality of life measures represent compound scores or indices based on several inputs. For example, the United Nations calculates an annual Human Development Index (HDI), representing the developmental level of each country on a scale of zero to one based on several factors, including life expectancy at birth, years of schooling, and per-capita income (The United Nations Development Programme (2018b)). The HDI also categorizes countries into four levels of development (low, medium, high, and very high). Similarly, the Social Progress Imperative publishes the Social Progress Index (SPI), ranging from 0 to 100, and comprising over 50 dimensions in three broad categories: basic human needs (e.g., nutrition, safety), foundations of wellbeing (e.g., basic knowledge, environmental quality), and opportunity (e.g., personal rights, freedoms) (Social Progress Imperative (2018b)). The World Economic Forum's Global Gender Gap Index reports a gender equality index, scaled from 0-1, based on measurements of gender-related gaps in such dimensions as economic participation, level of education, health and survival, and political offices held (World Economic Forum (2016b)). Finally, the World Happiness Report calculates a score from 0-10 by considering per-capita GDP, healthy life expectancy, social support, freedoms, and perception of corruption, among others (Helliwell, Layard, and Sachs (2018)).

The objective of this analysis is to explore the distributions of and relationships between key QoL indicators using both nonparametric and parametric methods, and to assess the appropriateness of each method used.

Methods

The dataset used in this analysis, titled `alldata`, was generated for the MAT 8790 course (Prioli (2018b)). It consists of country-level variables for calendar year 2016 as described in Table 1.

Table 1. `alldata` dataframe contents.

Source	Variable Name	Description
`countrycode` package	`country`	Country names
Social Progress Imperative (2018a)	`SPI`	Social Progress Index value (scale of 0:100)
The World Bank (2018)	`GDP_USD_2018`	2016 Gross Domestic Product (valued in $US 2018)
The United Nations Development Programme (2018a)	`HDIrank`	Human Development Index ranking
The United Nations Development Programme (2018a)	`HDIindex`	HDI index value (scale of 0:1)
The United Nations Development Programme (2018a)	`HDI_cat`	HDI index category (5 levels)

Source	Variable Name	Description
Helliwell, Layard, and Sachs (2018)	`happiness`	World Happiness Score (scale of 0:10)
World Economic Forum (2016a)	`genderequality_index`	Gender Equality Index (scale of 0:1)
World Health Organization (2018b)	`infantmort`	Infant mortality rate
World Health Organization (2018a)	`birth_MF`	Life expectancy at birth, males & females
World Health Organization (2018a)	`sixty_MF`	Life expectancy at 60 years, males & females

All variables pertain to the calendar year 2016. Missing values were omitted from the dataset to ensure that the tests of interest could be performed.

For each variable except `country`, descriptive statistics were run and a sensible visualization was generated, following which a correlation matrix was produced to examine pairwise relationships between continuous variables.

Nonparametric and Parametric Analyses

Based on the data exploration results, six sets of formal hypotheses were generated about the data (Table 2), and sensible nonparametric tests and their parametric equivalents were chosen to test these hypotheses.

Table 2. Analyses Performed

Analysis	Variable(s)	Null Hypothesis	Alternative Hypothesis	Nonparametric Test	Parametric Test
1	HDIindex	The sample median is equal to its mean	The sample median differs from its mean	One-Sample Sign Test	One-Sample t-Test
2	HDIindex, SPI	Human development and social progress are not associated	Human development and social progress are correlated	Kendall's Tau	Pearson's Correlation Test
3	logGDP, infantmort	There is no relationship between log(GDP) and infant mortality	There is a relationship between log(GDP) and infant mortality	Hoeffding's Test	Pearson's Correlation Test
4	happiness	Happiness is normally distributed	Happiness is not normally distributed	One-Sample Kolmogorov-Smirnov Test	Shapiro-Wilk Test
5	genderequality_index, SPI	Gender equality and adjusted social progress index have the same median	Gender equality index and adjusted social progress index differ in median	Wilcoxon Rank-Sum Test	Two-Sample t-Test
6	HDI_cat, infantmort	Infant mortality rate is the same across levels of human development	Infant mortality rate differs by level of human development	Permutation F-Test	ANOVA

A two-sided one-sample test was chosen for Analysis #1 because the `HDIindex` distribution is very non-normal, yet its median and mean appear quite similar and the standard deviation is small. The one-sample sign test assumes that the sample is random with independent draws, and the data are continuous. Its parametric equivalent, the one-sample t-test, shares these assumptions and also requires normality of the sampling distribution.

Analysis #2 was motivated by the correlation matrix analysis, which indicates a linear relationship is likely between `SPI` and `HDIindx`. Kendall's Tau was chosen as the nonparametric test, with the usual Pearson's test as the parametric alternative. Both tests assume continuous data and are equipped to detect linear dependence, and employed two-sided alternatives. Additionally, simulation is the preferred method when ties are present for Kendall's Tau, the variables were first assessed for ties before selecting a method for Kendall's Tau.

Correlation tests were chosen for Analysis #3 to determine whether increasing GDP is correlated with decreasing infant mortality. Hoeffding's Test was selected because it is sensitive to any departure from independence. Because the relationship appears to be roughly nonlinear, Pearson's test was chosen as the parametric comparator. Both tests employed two-sided alternatives and assume continuous data, and Pearson's test additionally assumes a linear relationship. These tests were carried out using `testforDEP::testforDEP()` with the options `test = "HOEFFD"` and `test = "PEARSON"` respectively.

Analysis #4 was chosen because `happiness` appears to have a symmetric, possibly normal distribution on univariate analysis. Both the one-sample Kolmogorov-Smirnov test and the Shapiro-Wilk test assume a continuous sample distribution.

Analysis #5 sought to determine whether there is a difference in median between $\frac{1}{100}$`SPI` and `genderequality_index`. Dividing `SPI` by 100 ensures it shares the same 0:1 scale as `genderequality_index`. After this rescaling, the two variables appear to have similar medians on descriptive analysis. The Wilcoxon Rank-Sum test assumes all observations are independent and come from the same distribution under H_0, whereas the two-sample t-test assumes independent samples and normality of the sampling distribution. A two-sided alternative was used for each test.

To test for a difference in infant mortality rate by human development category, Analysis #6 employed the permutation F-test and ANOVA as the nonparametric and parametric tests respectively. Both tests assume samples are independent and identically distributed, and ANOVA additionally assumes a normal distribution.

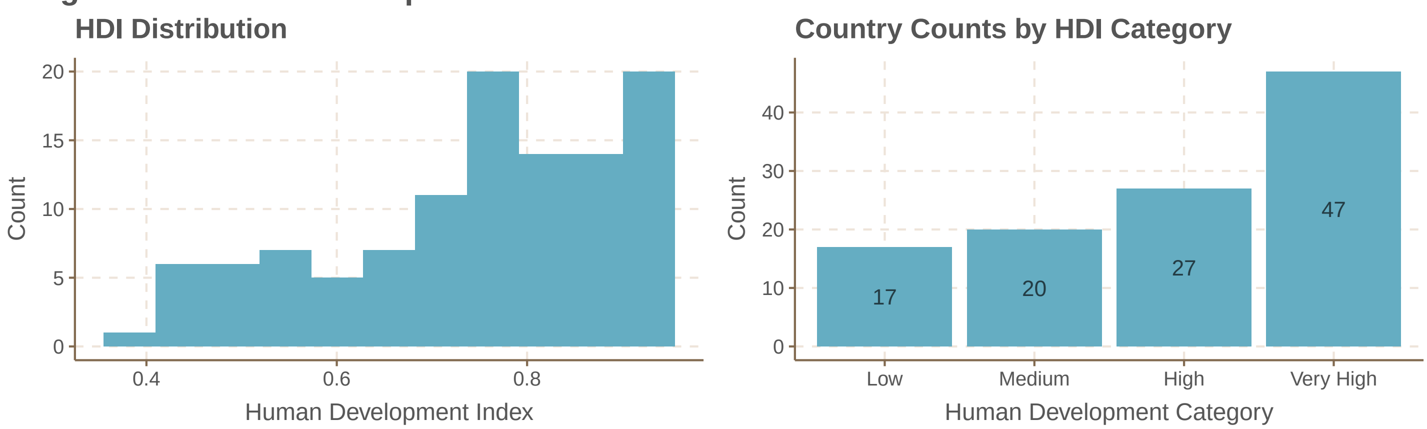
Results of these six paired analyses were compared in the context of the data and assumptions needed. All tests were performed at level $\alpha = 0.05$ in R. The dataset, full code, and this report are available in an online repository (Prioli (2018a)).

Results

Descriptive Statistics and Visualizations

Descriptive statistics and univariate visualizations for the dataset are presented in Figs. 1-8. Figure 1 depicts the distribution and descriptive statistics of the `HDIindex` variable, along with counts by country for the `HDI_cat` variable. `HDIindex` has a very nonnormal appearance, yet has similar mean and median, and a small standard deviation as compared to the mean. The highest category of human development is the most represented in the data.

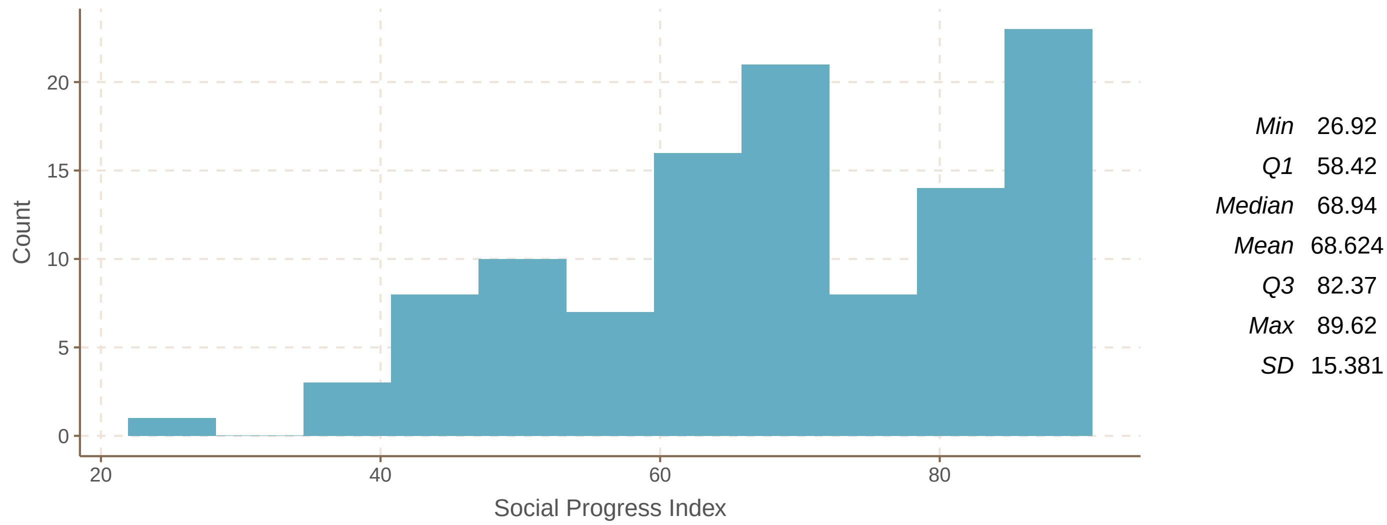
Figure 1. Human Development Index



Min	Q1	Median	Mean	Q3	Max	SD
0.405	0.653	0.772	0.745	0.867	0.951	0.15

`SPI`, described in Fig. 2, appears to have a multimodal distribution, with mean and median quite similar in value.

Figure 2. Social Progress Index Distribution

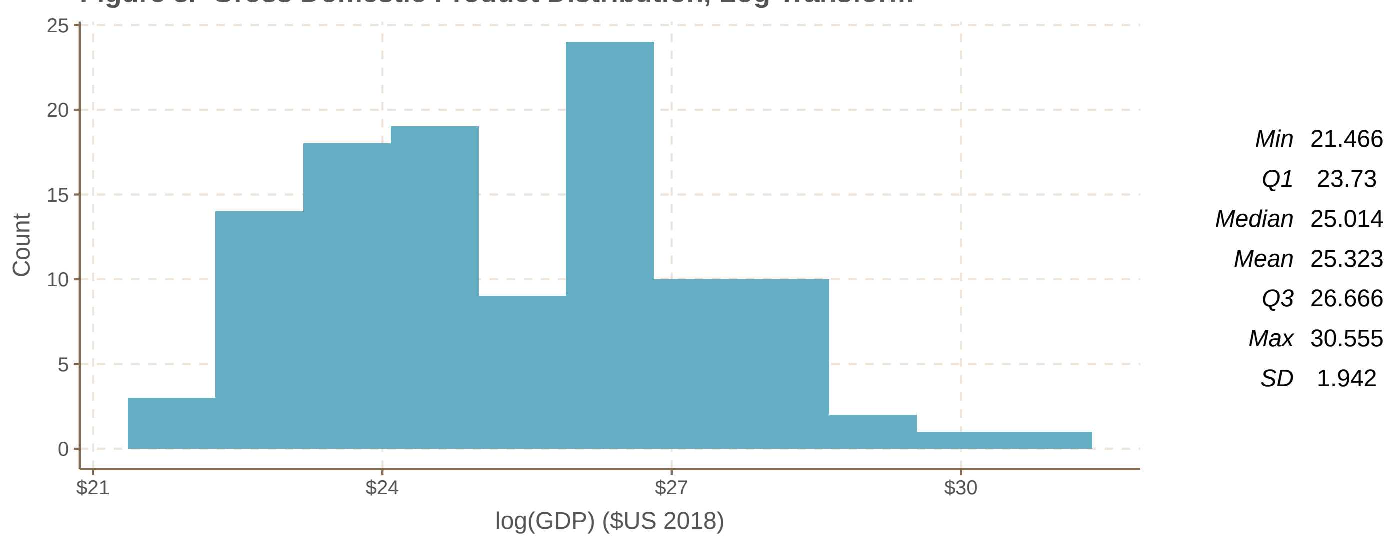


Summary statistics of the `GDP_USD_2018` variable (Table 3) show a wide range for this variable, with median an order of magnitude less than mean, and standard deviation an order of magnitude greater than mean. After log transformation, the distribution appears possibly bimodal with mean and median reasonably close in value and standard deviation smaller than both.

Table 3. GDP SUmmary Statistics

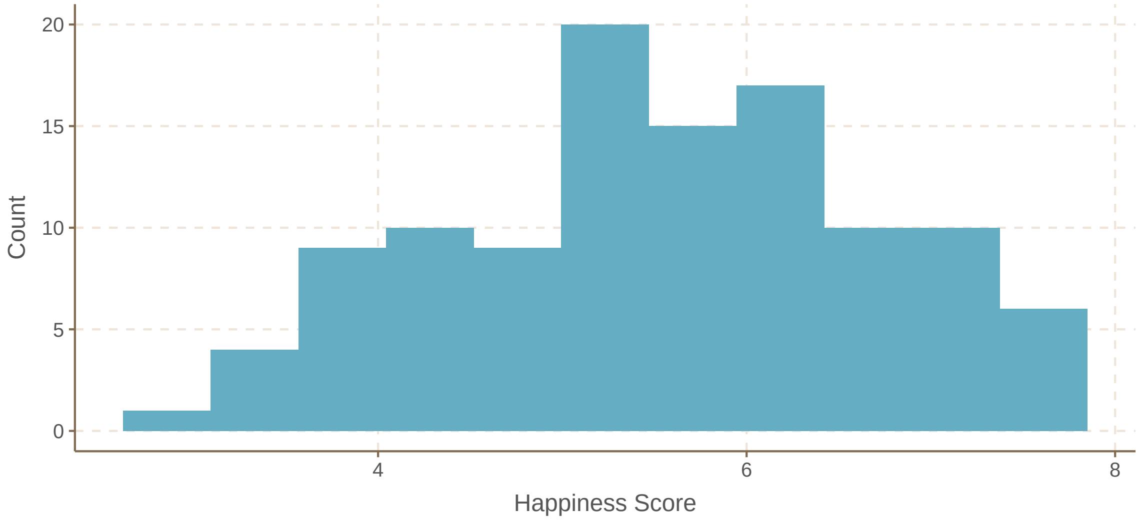
Min	Q1	Median	Mean	Q3	Max	SD
2101	20228.99	73000.98	650859.6	380937.5	18624500	2138824

Figure 3. Gross Domestic Product Distribution, Log Transform



Per Fig. 4, the World Happiness Score appears roughly normally distributed.

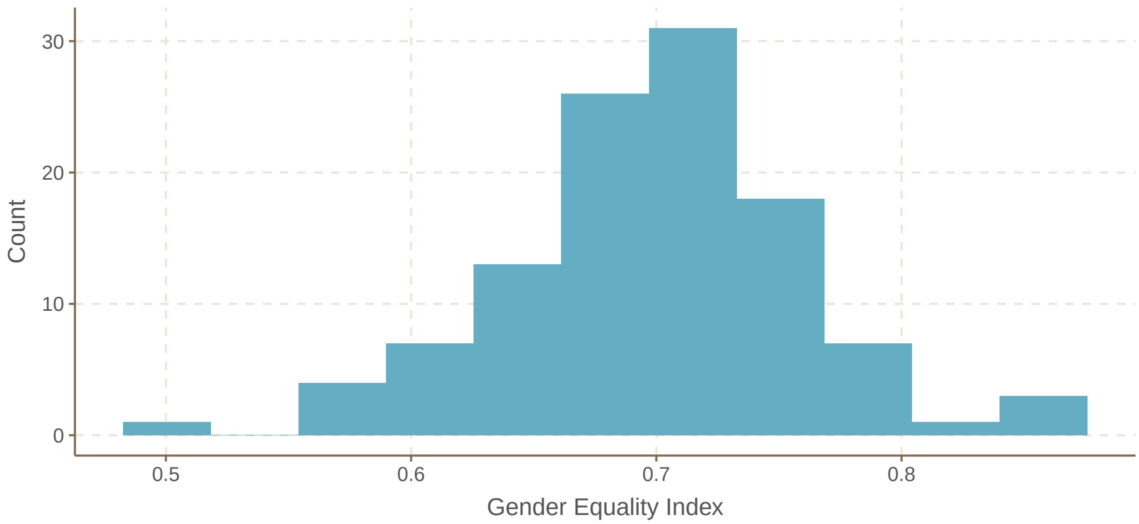
Figure 4. Happiness Score Distribution



Min	2.903
Q1	4.624
Median	5.578
Mean	5.553
Q3	6.339
Max	7.66
SD	1.139

Figure 5 shows a roughly symmetric, possibly trimodal distribution for `genderequality_index`, with mean and median very close in value, and standard deviation approximately a tenth of either measure of location.

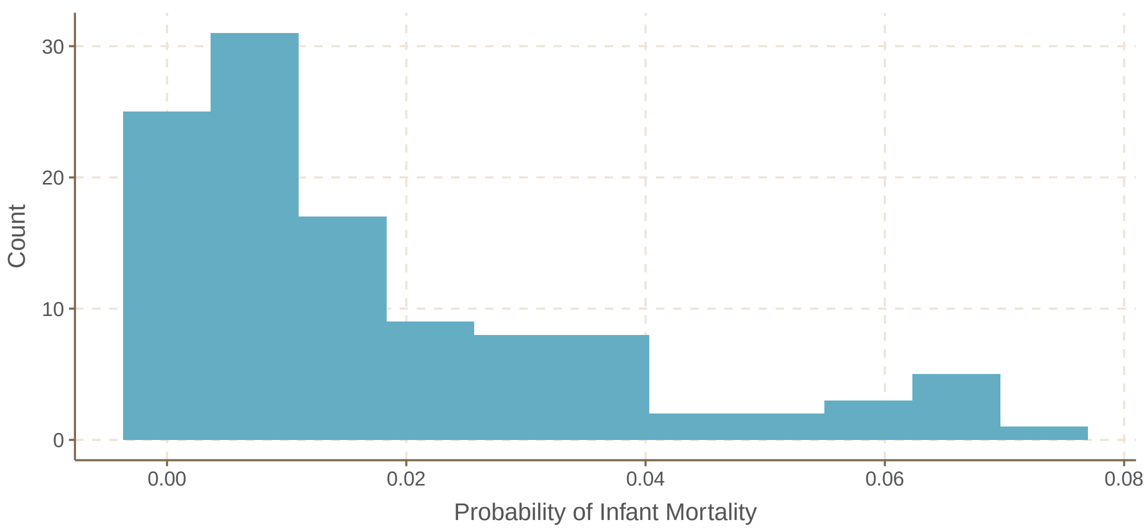
Figure 5. Gender Equality Index Distribution



Min	0.516
Q1	0.669
Median	0.7
Mean	0.699
Q3	0.735
Max	0.874
SD	0.06

The probability of infant mortality is heavily right-skewed per Fig. 6.

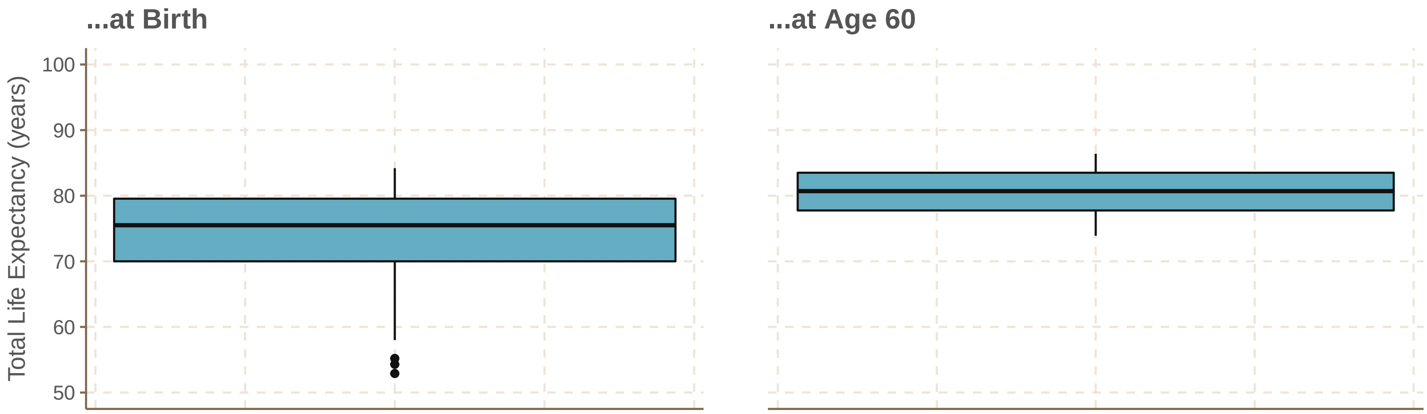
Figure 6. Infant Mortality Distribution



Min	0.002
Q1	0.004
Median	0.011
Mean	0.018
Q3	0.027
Max	0.075
SD	0.019

Figure 7 explores total life expectancy at birth vs. age 60. Life expectancy at birth appears left-skewed, but total life expectancy at sixty looks reasonably symmetric. These data indicate that, across all countries, a person who survives until at least 60 years has a longer total life expectancy than does the general world population at birth.

Figure 7. Total Life Expectancy...



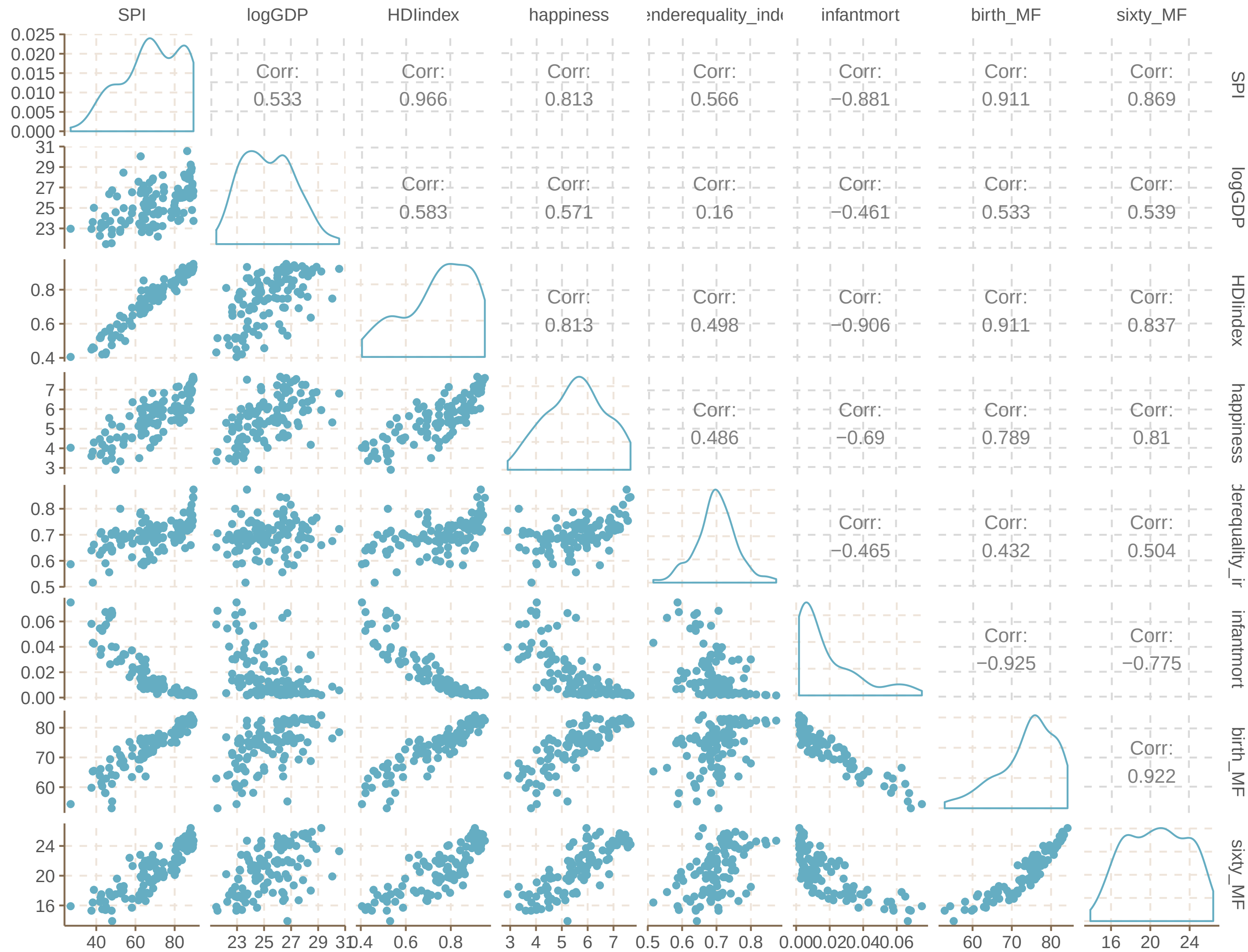
Min	Q1	Median	Mean	Q3	Max	SD
52.9	70	75.5	73.959	79.55	84.2	7.27

Min	Q1	Median	Mean	Q3	Max	SD
73.9	77.75	80.7	80.662	83.5	86.4	3.107

Figure 8 depcits pairwise relationships between continuous variables. Strong positive linear relationships are seen between `HDIindex` and `SPI`, `HDIindex` and `birth_MF`, and `SPI` and `birth_MF`, indicating that higher levels of social progress and human development generally correlate with a longer life expectancy at birth. Additionally, positive linear relationships are seen between `happiness` and `birth_MF`, between `happiness` and `HDIindex`, and between `happiness` and `SPI`, indicating that increasing levels of human development and social progress generally lead to increased happiness (though there is considerable spread in the data), and that happiness and longevity may also be correlated.

Strong negative relationships are seen between `infantmort` and `birth_MF`, between `HDIindex` and `infantmort`, and between `SPI` and `infantmort`, though the latter two of these may not necessarily be linear. These indicate that increasing human development and social progress levels correlate with lower levels of infant mortality, and that as infant mortality decreases, life expectancy at birth increases.

Figure 8. Correlation Matrix, Continuous Variables



Nonparametric and Parametric Analyses

For Analysis 1, the one-sample sign test for $H_0 : m = \texttt{mean(alldata\$HDIindex)}$ versus $H_A : m \neq \texttt{mean(alldata\$HDIindex)}$ yielded a p-value of $p_{sign} = 0.036$. Since $p_{sign} < \alpha$, the null hypothesis was rejected, leading to the conclusion that the median and mean HDI index values are not equal. However, the parametric test assessed a complementary set of hypotheses − that is, $H_0 : \mu = \texttt{median(alldata\$HDIindex)}$ versus $H_A : \mu \neq \texttt{median(alldata\$HDIindex)}$ − and resulted in $p_t = 0.055$. Since $p_t > \alpha$, the null hypothesis was retained with the conclusion that, at the $\alpha = 0.05$ level, there is insufficient evidence to assert that the mean and median are different in the parametric case. The sign test is appropriate for this data; however, due to the clear nonnormality of the `HDIindex` data, the t-test is inappropriate because the assumption of normality is violated. Thus the sign test results were retained.

For Analysis 2, ties were not present in either `SPI` or `HDIindex`; thus, Kendall's Tau was tested using the usual `cor.test(x, y, method = "kendall")` approach. When testing for independence in both the nonparametric and parametrics cases, the p-values are less than 0.0001, with correlation coefficients $r_\tau = 0.843$ and $r = 0.966$ respectively. Both are statistically significant and represent strong evidence of a linear relationship between social progress and human developmental level, thus H_0 was rejected with the conclusion that, at the $\alpha = 0.05$ level, social progress and human development are correlated. Both tests assume continuous data, which is satisfied, and both yield the same p-value. The nonparametric method is more conservative, with a slightly smaller correlation coefficient for Kendall's Tau as compared to Pearson's, so the nonparametric method results were retained.

When testing for independence in Analysis 3 for both the nonparametric and parametrics cases, the p-values are less than 0.0001, leading to the rejection of H_0 (i.e., independence) and the conclusion that, at the $\alpha = 0.05$ level, there is sufficient evidence of an association between log(GDP) and infant mortality. The correlation matrix shows that this is an inverse relationship - i.e., as log(GDP) increases, infant mortality decreases. This relationship was sufficiently linear to yield a significant p-value on Pearson's test, so although the scatterplot in Fig. 8 appears nonlinear, it may in fact be linear.

In Analysis 4, the one-sample Kolmogorov-Smirnov test for H_0 : `happiness` is normally distributed vs. for H_A : `happiness` is not normally distributed yielded p-value $p_{KS} = 0.915$, whereas the p-value for the Shapiro-Wilk test is $p_{SW} = 0.129$. Both exceed alpha; thus, H_0 is retained for both tests and I conclude that there is insufficient evidence to assert that `happiness` is not normally distributed. Both tests assume a continuous distribution, which is satisfied by the data, and the tests agree, so either the nonparametric or parametric approach is acceptable in this case. Notably, though neither p-value is significant, p_{SW} is much smaller than p_{KS}.

Discussion

Limitations

Conclusion

References

Helliwell, John F., Richard Layard, and Jeffrey D. Sachs. 2018. "World Happiness Report." http://worldhappiness.report/ed/2018/.

Prioli, Katherine M. 2018a. "MAT_8452_Final_Project." https://github.com/kmprioliPROF/MAT_8452_Final_Project.

———. 2018b. "MAT_8790_Final_Project." https://github.com/kmprioliPROF/MAT_8790_Final_Project.

Social Progress Imperative. 2018a. "Social Progress Index." https://www.socialprogress.org/?tab=4.

———. 2018b. "Social Progress Index - Overview." https://www.socialprogress.org/index.

The United Nations Development Programme. 2018a. "Human Development Index." http://hdr.undp.org/en/data.

———. 2018b. "Human Development Reports - Human Development Index (HDI)." http://hdr.undp.org/en/content/human-development-index-hdi.

The World Bank. 2018. "Gross Domestic Product." https://data.worldbank.org/indicator/ny.gdp.mktp.cd?view=map&year_high_desc=true.

World Economic Forum. 2016a. "Gender Equality." http://reports.weforum.org/global-gender-gap-report-2016/rankings/.

———. 2016b. "Measuring the Global Gender Gap." http://reports.weforum.org/global-gender-gap-report-2016/measuring-the-global-gender-gap/.

World Health Organization. 2018a. "Life Expectancy." http://apps.who.int/gho/data/view.main.SDG2016LEXv?lang=en.

———. 2018b. "Probability of Dying Per 1000 Live Births." http://apps.who.int/gho/data/view.main.182?lang=en.