The Determination of Government Structure, Social Support, and Economic Development on Citizen Happiness Level

Shashwat Bansal, Allen Chen, Anni Wang, Ram Mukund Kripa, and Robert Lee

Introduction

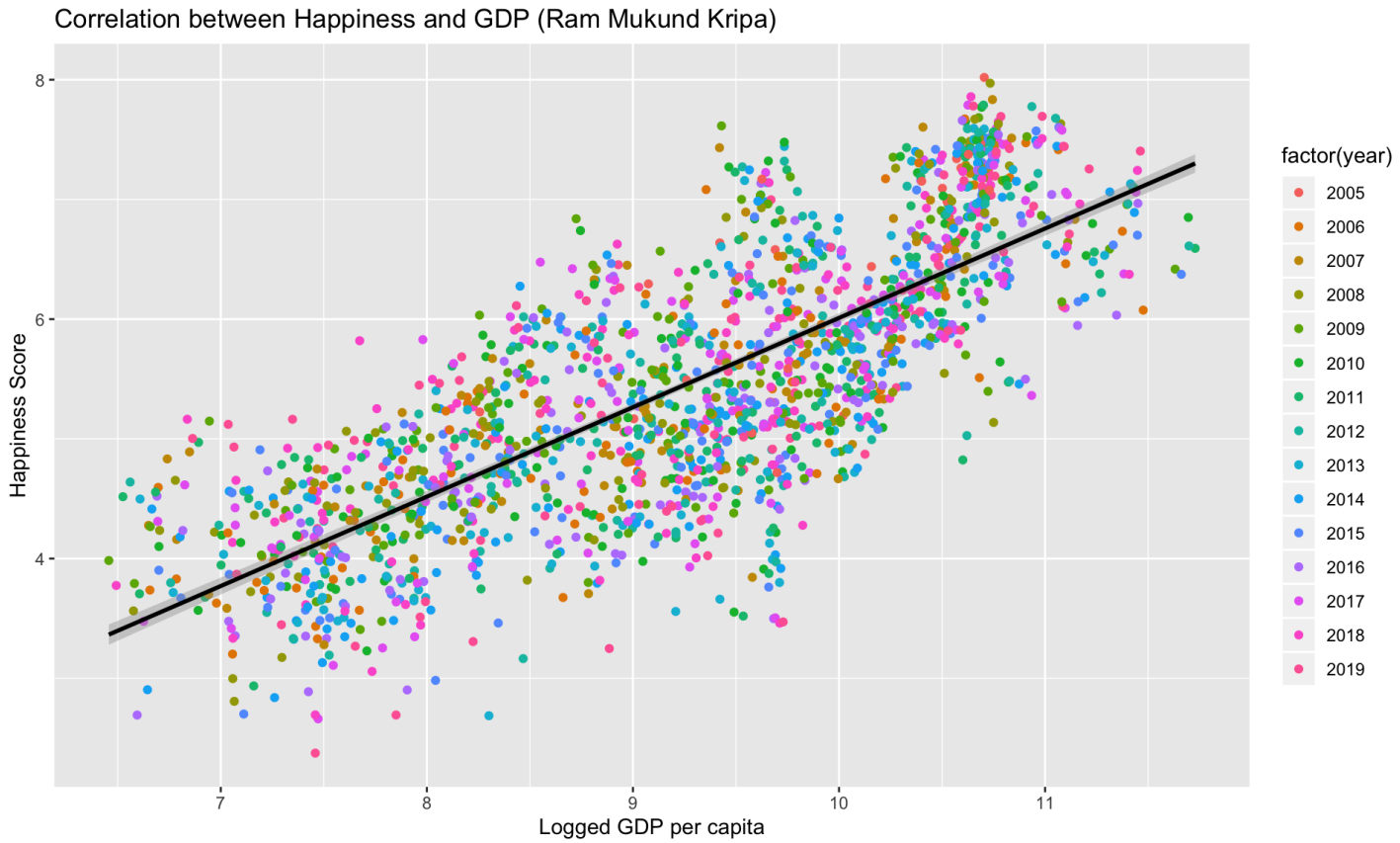
Utilizing the World Happiness Report of 2019 and the historical records of country happiness published by previous reports, the goal of the primary investigation is to determine the primary causation for “happiness” by analyzing the relationships between citizen happiness, availability of social support, regional indicators, GDP per capita, life expectancy, freedom to make choices, generosity, and perceptions of corruption. Although we recognize that a positive correlation between two variables does not necessarily determine the causation of happiness within a specific country, positive correlations between variables (such as availability of social support), etc. indicate other factors related towards the positive correlation, revealing insight into areas of expansion, future growth, and current problems within the economies of listed countries. We developed our own “segmenting” strategies with the intent to utilize hypothesis tests that included t-tests, z-tests, ANOVA, Tukey, and chi-squared tests - our strategy included separating countries based on their log GDP per capita into rich, middle, and poor countries. Through our project, we intend to address the following questions with our data analysis and supporting supplementary research:

1. What is the effect of GDP per capita and categories of wealth on happiness levels?
2. How are GDP per capita and country wealth related to perceptions of corruption - and specifically, is the level of corruption related to happiness levels?
3. Does the amount of social support in the respective group of countries affect the healthy life expectancy of the people? Also, does the healthy life expectancy affect the happiness level in each respective group of countries?
4. What is the relative affect of freedom of life choices on happiness levels, specifically the relationship between social support and freedom to make life choices?

Section A: GDP per capita and categories of wealth on happiness levels

Exploratory Data Analysis

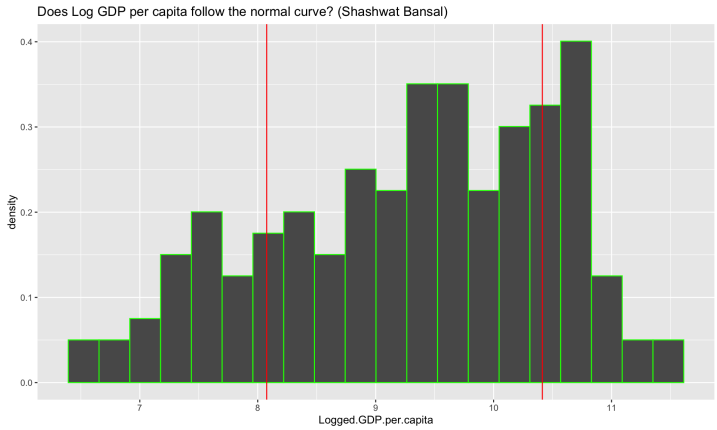
One of the first ideas that we got from this dataset was to explore the correlation between GDP and Happiness. Higher values of Log GDP per capita seemed to correspond to higher happiness ladder scores. This makes intuitive sense. While correlation does not imply causation, it would seem logical to conclude that higher GDP implies happier people.




```
##
## Call:
## lm(formula = Log.GDP.per.capita ~ Life.Ladder, data = dataprev)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.17901 -0.52264  0.02385  0.54186  2.08052
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.84350    0.08550   56.65   <2e-16 ***
## Life.Ladder  0.80806    0.01537   52.56   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7362 on 1817 degrees of freedom
##   (29 observations deleted due to missingness)
## Multiple R-squared:  0.6032, Adjusted R-squared:  0.603
## F-statistic:  2762 on 1 and 1817 DF,  p-value: < 2.2e-16
```

We also looked at the p value associated with the slope of the regression line produced by the linear modeling function. It was approx. 0, which seemed to indicate that this correlation is real, not just due to chance.

The idea of segmenting the dataset into rich, poor and middle countries was formulated soon after this plot was observed. We looked at the distribution of Log GDP per capita to decide how to segment the countries.



Since the distribution looked roughly normal, we defined rich countries as those having Log GDP per capita >= one SD above average. We also defined poor countries as those having Log GDP per capita <= one SD below average. The rest were defined as middle countries

Data Analysis

Now, we wanted to see if the likelihood of people in rich countries being happier than average is higher than that of poor and middle countries.

Our notion of happiness is defined by having an above average happiness ladder score. Since we defined these three categories and wanted to compare them, a chi squared test seemed like the logical thing to do.

Chi Squared Test

Null Hypothesis: a) The results are due to chance b) Rich, poor and middle countries are equally likely to report above average happiness ladder scores

Alternative Hypothesis: a) The results are not due to chance b) Rich, poor, and middle countries vary in their likelihood of having above average happiness ladder scores.

Test Statistic:

```
## [1] "This is table on which chi sq test is being done"
```

label <fctr>	more <int>	less <int>
rich	345	5
middle	514	592
poor	5	358

3 rows

```
## [1] "chi sq statistic is "
```

```
## [1] 676.202
```

```
## [1] "p value is"
```

```
## [1] 0
```

The p value is approximately zero. Hence we can reject the null at the 1% level, and any variations is not due to chance variation.

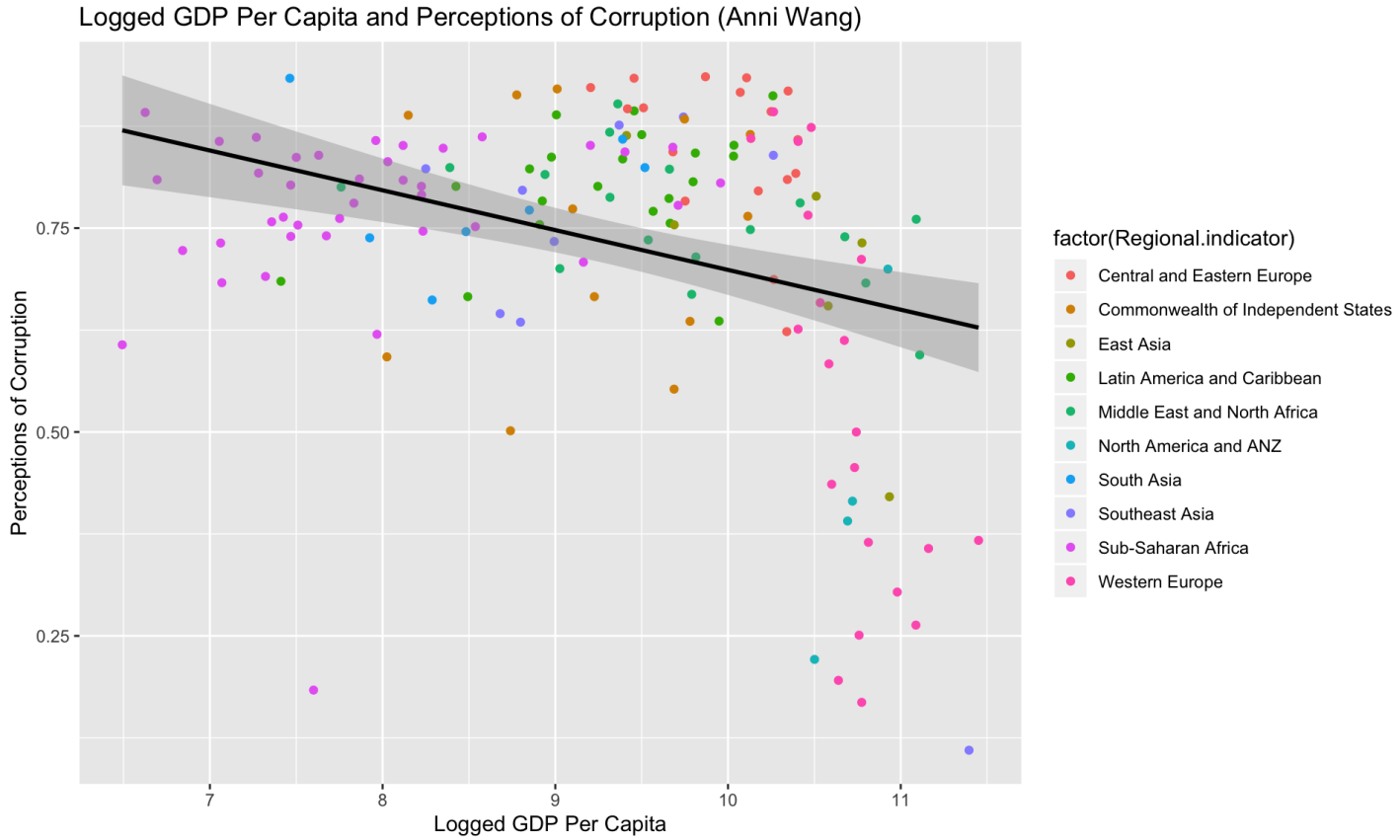
This matched our intuition that richer countries were more likely to have happier people.

Section B: GDP per capita and country wealth related to perceptions of corruption

Exploratory Data Analysis/ Description of Data

The primary goal of this section is to evaluate whether GDP per capita and country wealth are related to perceptions of corruption - and specifically, whether levels of corruption and country wealth are related to happiness levels. Utilizing the country classification system defined previously within our code, countries were segmented into rich, poor, and middle groups depending on their logged GDP per capita - countries above one standard deviation of the mean log GDP per capita were classified as “rich,” countries between one standard deviation above and below the mean were classified as in the “middle,” and countries below one standard deviation below the mean were classified as “poor.” Similarly, a rating system of high corruption, some corruption, and low corruption was created modelled on the wealth classification system as mentioned.

As an extension of the first section, which discusses the relation of logged GDP per capita to happiness, the primary goal of this section is to investigate whether a correlation between GDP per capita and perceptions of corruption exist, and to determine whether subsequent perceptions of corruption are therefore related to happiness.



According to the scatterplot, there is a strong negative correlation between Logged GDP per capita and perceptions of corruption. However, Rwanda exhibits the opposite graphical behavior, in that it displays both low logged GDP per capita as well as low perceptions of corruption, whereas other countries with low logged GDP per capita tend to have extremely high perceptions of corruption. Accoriding to BBC, Rwanda mainly has low levels of corruption due to its police state with severe punishments instituted for infractions - Transparency International has difficulty in producing comparisons of Rwanda against neighboring countries in terms of corruption due to extremely low reports and accounts of government bribery. According to the Rwanda Governance Scorecard 2019 (Rwanda Governance Board 2019), citizens in Rwanda place high trust within the Rwanda Police Force at a level of 96.5%. According to Transparency International Rwanda, “trust has been maintained over the last years in response to years of public sector purges that helped to significantly improve the police and military’s accountability to civilians.” The government of Rwanda has implemented policies intended to curb corruption within its police force, including doubling of salaries, professionalism training, and creation of an ethics centre intended to develop the discipline of recruits and existing staff.

Data Analysis

In order to evaluate whether GDP per capita and country wealth are related to perceptions of corruption - and specifically, whether levels of corruption are related to happiness levels, the chi-squared, t-test, ANOVA, and Tukey hypothesis tests were utilized.

A t-test was utilized to address whether there was a significant difference between how rich and middle countries perceive corruption. The Welch’s Two Sample T-Test was selected for this question - as I intend to observe whether there are differences between the means of corruption perceptions between two samples that were possibly unequal in their size.

```
##
##  Welch Two Sample t-test
##
## data:  Perceptions.of.corruption by countrytypes
## t = -8.1339, df = 30.128, p-value = 4.296e-09
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.4033847 -0.2414955
## sample estimates:
##   mean in group richcountry mean in group middlecountry
##                   0.4799714                   0.8024115
```

Assuming that the null hypothesis is that rich and middle countries are equal in terms of their average perception of corruption (and that any difference is due to chance variation), and that the alternative hypothesis is that rich and middle countries are not equal in terms of their average perception of corruption (and that any difference is not due to chance variation), Welch's Two Sample t-test, the p-value returned is 4.296e-09, which is highly significant. Therefore, the null hypothesis is rejected. It appears that there is a significant difference in terms of how wealthy and middle countries perceive corruption. Upon analysis of the countries that were placed in the rich category and middle categories, most of these countries operate on a very democratically-centered government in which citizens are actively participating and involved in public politics. Many countries placed in the middle, such as the Ivory Coast for example, have historically very totalitarian and dictator forms of government, contributing to a high perception of corruption within the countries.

I then wanted to address: Is there a significant difference between how middle and poor countries perceive corruption?

```
##
##  Welch Two Sample t-test
##
## data:  Perceptions.of.corruption by countrytypes
## t = 1.9969, df = 37.438, p-value = 0.05315
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.000748847  0.105631616
## sample estimates:
## mean in group middlecountry   mean in group poorcountry
##                   0.8024115                   0.7499701
```

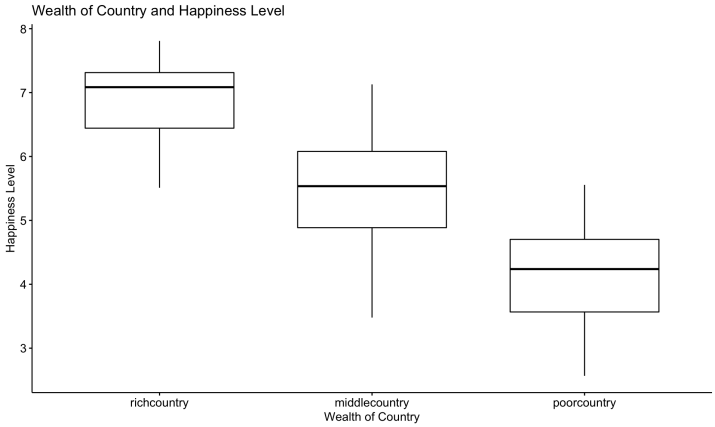
I assume the same null hypothesis and alternative hypothesis as in the test above. Interestingly, the p-value is 0.05315, which is greater than 0.05. The averages are 0.8024115 for middle countries, and 0.7499701 for poor countries. There is not a significant difference between perceptions of corruption between middle and poor countries, leading to the failure of rejecting the null hypothesis. Upon analysis of governmental structure in middle and poor countries, several countries share similarities in both the middle and poor categories in terms of historical totalitarian government structures, such as the former historical dictator of Niger Sani Abacha, and Mobutu Sese Seko of the Democratic Republic of the Congo.

Finally, a chi-square test for independence was conducted to verify whether perceptions of corruption are independent of happiness level. First, I defined whether countries were considered high, medium, or low in terms of corruption perception utilizing historical data. The methodology is modeled from the rich, middle, and poor countries, as in highly corrupt countries are greater than one standard deviation above the mean, somewhat corrupt countries are between one standard deviation above and below the mean, and low corruption countries remain one standard deviation below the mean level of corruption. Happiness was split into two sections - countries were considered more happy if they are above the average level of happiness in the WHR20_previous_years data.

```
##
##  Pearson's Chi-squared test
##
## data:  corruptiondatatable
## X-squared = 159.08, df = 2, p-value < 2.2e-16
```

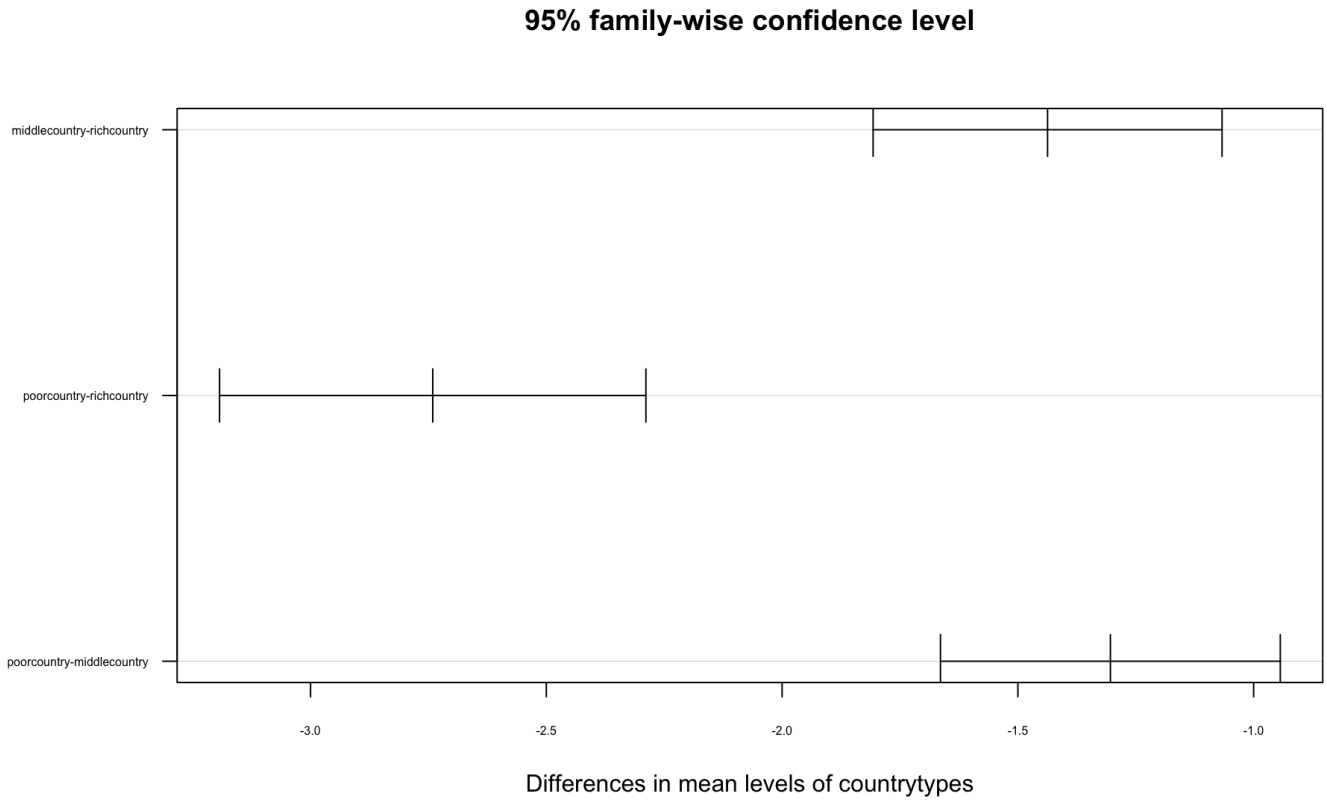
The p-value for this Chi-squared test is < 2.2e-16. There is an undeniable relationship between happiness and perceptions of corruption - which is intuitive, as high perceptions of corruption likely limits individual freedoms, and therefore leads to low happiness levels. According to Nina Gorovaia and Stavros A. Zenios - development is a process of expanding the real freedoms and activities that people enjoy, which is directly related to increasing happiness (which is also interchangeably defined as life satisfaction).

I now intend to address the second part of my question: whether country wealth is related to happiness levels. An ANOVA test was utilized to determine whether there are significant differences in the level of happiness, between countries that are rated wealthy countries, medium wealthy countries, and poor countries. My null hypothesis is that the three categories of countries have similar mean happiness levels, and that variation is due to chance. Similarly, the alternative hypothesis states that at least one category is different than the two others, and that this is not due to chance variation.



According to the box plots above, a clear difference can be seen between the averages of the rich country, middle country, and poor countries in terms of happiness levels. However, the p value returned by the ANOVA test is <2e-16 - as trends are shown in the plot, at least one of the categories of richcountry, middlecountry, and poorcountry is significantly different in terms of their mean happiness level.

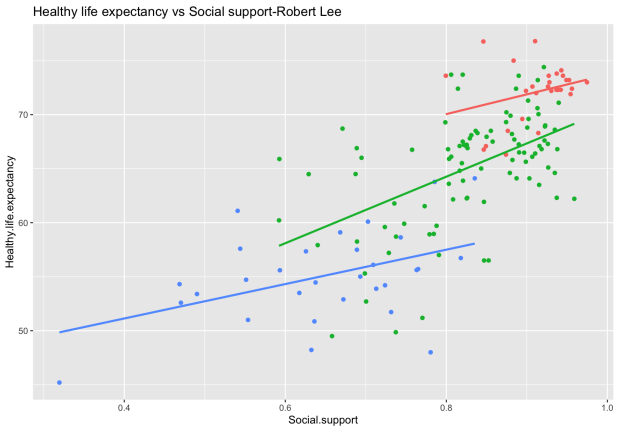
```
##               Df Sum Sq Mean Sq F value Pr(>F)
## countrytypes   2 108.81   54.41     103 <2e-16 ***
## Residuals    150  79.23    0.53
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Similarly, as trends are shown above - middle countries and rich countries are visually much more similar compared to poor countries and middle countries, but poor countries and rich countries are much more difference in terms of their mean levels.

Section C: Social Support and Healthy Life Expectancy

Exploratory Data Analysis/ Description of Data



Scatter plot of the countries with regression line to test whether there is a linear relationship between healthy life expectancy and social support

Scatter plot of the countries with regression line to test whether there is a linear relationship between healthy life expectancy and happiness levels

Data Analysis

```
##
##  Welch Two Sample t-test
##
## data:  Healthy.life.expectancy by countrytypes
## t = 9.5037, df = 89.285, p-value = 3.304e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  5.514169 8.429232
## sample estimates:
##   mean in group richcountry mean in group middlecountry
##                    72.10938                    65.13768
```

Then, a T test that compares whether a country being rich or mid affects the health life expectancy.

```
##
##  Welch Two Sample t-test
##
## data:  Healthy.life.expectancy by countrytypes
## t = 10.622, df = 58.517, p-value = 2.835e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   8.145702 11.928048
## sample estimates:
## mean in group middlecountry   mean in group poorcountry
##                    65.13768                    55.10080
```

Then, a T test that compares whether a country being mid or poor affects the health life expectancy.

```
##
##  Welch Two Sample t-test
##
## data:  Healthy.life.expectancy by countrytypes
## t = 18.319, df = 49.06, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  15.14283 18.87432
## sample estimates:
## mean in group richcountry mean in group poorcountry
##                  72.10938                  55.10080
```

A T test that compares whether a country being rich or poor affects the health life expectancy For both the data for the low-gdp countries and mid-gdp countries we use the linear model, which had an r^2 value that indicated a linear relationship and a low p-value under 0.05 respectively based on a t-test that indicated significant evidence whether there is a relationship between the amount of social services and the life expectancy. However, the high-gdp countries had an r^2 value and had a p-value of 0.1446 that did not indicate a strong linear relationship between the amount of social services and the life expectancy or provide evidence that there is a relationship between the amount of social services and life expectancy.

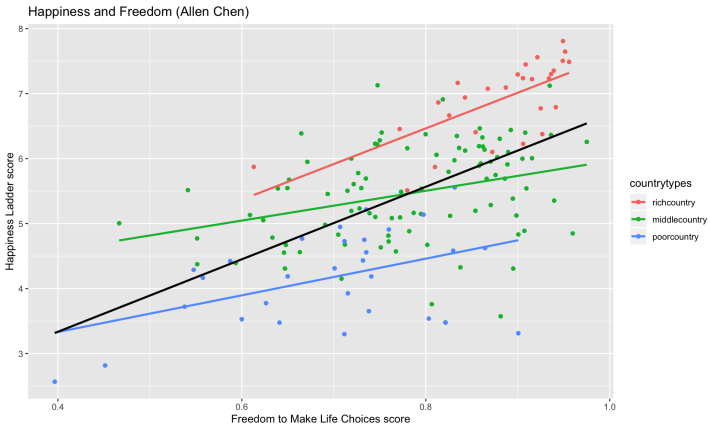
Through t-test, we observed a low p-value in the t-test comparing the healthy life expectancy between the rich and poor countries, rich and middle, and poor and middle, therefore we have significant evidence that there is a difference in life-expectancy between each of these two groups of countries.

Section D: Freedom of Life Choices on Happiness Levels

Exploratory Data Analysis

The debate over freedom has seen new battlegrounds in light of the Covid-19 pandemic, as strong totalitarian lockdowns have proved to be effective in restricting the spread of the virus, and democratic countries find themselves struggling to balance their citizens' freedoms with the necessity of a national lockdown. Are the benefits of freedom real? Are they the same across different societies? What factors contribute to freedom in a society? With a little bit of exploratory data analysis, perhaps one can glean new insights into these questions.

Let's start by looking at the correlation between Freedom to Make Life Choices, a score between 0 and 1 given by citizens of the country being evaluated, and the overall happiness Ladder Score of a country, a value between 0 and 10 that measures the average response of a country's citizen's when asked how happy they were.




```
## [1] "Correlation between happiness ladder score and freedom"
```

```
## [1] 0.5905968
```

```
## [1] "Correlation between happiness and freedom in rich countries"
```

```
## [1] 0.6871969
```

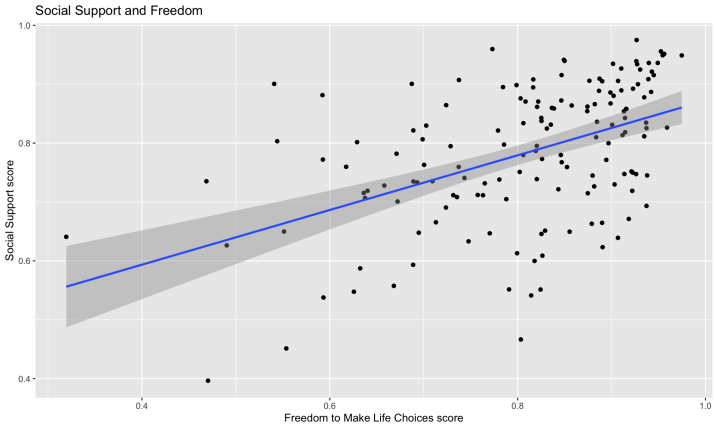
```
## [1] "Correlation between happiness and freedom in poor countries"
```

```
## [1] 0.4617554
```

Using our predefined definitions of rich, middle and poor countries (the upper and lower 15% of countries ranked by log GDP), we can look at this same correlation between freedom and happiness within rich and poor countries. On the plot, each country type has been color-coded and their individual regression lines have been drawn.

It would appear that the correlation is stronger in rich countries than it is in poor countries.

What factors are correlated with a country's freedom, then? Looking at both the overall data and then the individual rich and poor segments, it seems that there is in fact a strong correlation between a nation's level of social support and the self-percieved freedom of their citizens.



Data Analysis

Within this exploration of the data, a recurring theme in the analysis of the correlations between freedom, happiness, and social support was that rich countries seem to have a greater correlation than poorer ones. We can see whether these differences are due to random chance or not by using a hypothesis test.

To start, let's see whether the difference in average freedom between rich and poor countries is real. Our null hypothesis is that the observed difference is due to chance, giving us an expected difference of zero. Our alternative states that the observed difference is real and not chance, and that there are factors between rich and poor countries that cause their freedom scores to differ.

```
## [1] 0.8798518 0.6943083
```

```
## [1] "These are the mean freedom scores of the rich and poor countries, respectively"
```

```
## [1] 0.0258756
```

```
## [1] "Using R and the law of averages, we find the SE of the difference of means"
```

```
## [1] 0.1855435
```

```
## [1] "This is the observed difference of the means"
```

```
## [1] 3.73368e-13
```

```
## [1] "And this is the p-value"
```

The p-value in this case is highly significant, which tells us that the difference in the means is most likely not due to chance. There is a real, significant difference between the freedom of rich countries and the freedom of poor countries.

Conclusion

In general, wealth is an indicator for happiness ladder scores - rich, poor, and middle countries have different likelihoods of reporting above average happiness ladder scores. Combining these findings with the findings from the scatterplot one in Section A, we can infer that richer countries are more likely to report above average happiness scores than poor and middle countries.In terms of likelihood of reporting above average happiness ladder scores, generally countries follow the trend that wealthier countries will have the greatest likelihood of reporting a higher happiness ladder score, followed by middle countries, and then poor countries. In terms of the relationship between GDP per capita and perceptions of corruption, there appears to be a strong negative correlation between Logged GDP per capita and perceptions of corruption. Analysis in section B revealed that wealthier countries appeared to follow a more democratically structured government, in which citizens had high levels of civic and political participations and freedoms, leading to a lower perception of corruption. Many countries placed in the middle and lower sectors of wealth, such as the Ivory Coast for example, have historically very totalitarian and dictator forms of government, contributing to a high perception of corruption within the countries. While keeping in mind the strong dependent relationship between happiness ladder scores and perception of corruption - countries with higher perceptions of corruption (which were usually in the lower economic sector of middle and lower wealth) had much lower ladder happiness scores than wealthier countries (with very democratically centered governments and traditionally high personal freedom levels), which had much higher ladder happiness scores. In fact, according to Section D, wealthier governments are more able to provide higher levels of social support, leading to higher happiness ladder scores. In low-gdp countries and mid-gdp countries have more of a correlation between social services and life expectancy than the high-gdp countries because with the LDC countries because the lack social support from the government would most likely affect their living conditions and life expectancy, which in turn affects the happiness level of the people. In high-gdp countries there is a low correlation between social services and life expectancy because with high-gdp countries, because of the prevalence of social support provided, there is a less of an effect on the life expectancy, which in turn has less of an effect on the over happiness of the people in high-gdp countries.We can see an example of the effect of lack of social services in poor countries that effect life expectancy and happiness. For example, according to an nih study, fanthun and colleagues showed that low social support for mother's and care takers increased the mortality rate, which would decrease life expectancy. Also, in the nih study, nilsson and colleagues analyzed the relationship between social capital and self expressed quality of life in rural bangladesh. They analyzed that these lack of social capital provided by the government decreased the self expressed quality of life, which provides some evidence that the lack of social support from government can lead to lack of happiness.

References

https://www.bbc.com/news/world-africa-10726324 (https://www.bbc.com/news/world-africa-10726324)

https://knowledgehub.transparency.org/helpdesk/curbing-corruption-in-the-police-in-rwanda (https://knowledgehub.transparency.org/helpdesk/curbing-corruption-in-the-police-in-rwanda)

https://www.newtimes.co.rw/section/read/30573 (https://www.newtimes.co.rw/section/read/30573)

https://www.britannica.com/biography/Mobutu-Sese-Seko (https://www.britannica.com/biography/Mobutu-Sese-Seko)

https://www.cnn.com/2019/06/05/africa/nigeria-abacha-stolen-loot-jersey-intl/index.html (https://www.cnn.com/2019/06/05/africa/nigeria-abacha-stolen-loot-jersey-intl/index.html)

https://www.researchgate.net/publication/233406564_Does_freedom_lead_to_happiness_Economic_growth_and_quality_of_life (https://www.researchgate.net/publication/233406564_Does_freedom_lead_to_happiness_Economic_growth_and_quality_of_life)

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4318243/ (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4318243/)