# Factors Surrounding Suicide Rates Across Countries

Ridhika Agrawal, Yuejun Chen, and Yuya Kawakami

**Abstract:**

Reducing suicide rate has been a focal point of many disciplines for decades. It is noted that individuals' suicides not only affect their families and people surrounding them, they also have societal costs. Moreover, suicide is a major preventable cause of death on a global level. In this study, we study the relationship between suicide rate and variables from different social dimensions across countries, hoping to identify valuable factors that correlate with suicide rate for future research. Our variables include happiness index, temperature, religion, and mental health resources. We used a multivariate regression model and ANOVA to evaluate the significance of each variable. Our result showed that countries in which Islam dominates tend to have lower suicide rate. Depending on the income group of the country, each variable is correlated with suicide rate differently, with temperature showing significance in some income groups.

**Background:**

The research question of this paper was: what factors correlate with suicide rates? We are interested in studying this topic because we believe that a country's suicide rate reflects the general mental well-being of its residents[6,7]. Hence, we hope to offer policy makers additional consideration when developing programs/policies aimed at improving the mental health of the nation.

From our literature reviews, we found that subjective well-being and temperature are shown to be positively correlated with suicide rate[1,2]. Therefore, we include happiness index and annual average temperature of each country in our model to study suicide on a international level. We also included the availability of mental health resources, which is quantified by the number of psychiatrists per 100,000 residents, because previous research also shows that mental disorders are among the strongest predictors of suicide[4]. Our prior research also indicated cultural influences in suicide rates[3]; we quantified this by including the dominant religion of each country to improve our model.

**Data Cleaning and Methods:**

We adopted multivariable regression model to find the correlation between different factors and suicide rate as well as ANOVA to examine the difference in suicide rate between different countries that have religions. We obtained our data from different scholarly and governmental sources (Refer to Appendix D). The suicide rate and happiness index are both collected in 2015 across countries. The number of psychiatrists per 100,000 residents was collected in 2014. The data for religion were collected in 2010. The annual average temperature was last updated in 2011. Since all the data were collected between 2010 to 2015, we are confident that they are relevant indicators for each variable in 2015 and thus can be incorporated into our analysis.

We cleaned our data by relabeling columns to make it more readable, filtering out irrelevant or unworkable observations, and making adjustments to imperfect but workable observations. We transformed the religion data set by selecting top four religions (Christianity, Buddhism, Islam, Judaism), non-religious (atheism) and other religions not included in the top-four religions and atheism (other) as our categorical variable. We then selected the religion that makes up the highest portion of the population in the country as the dominant religion. We also transformed temperature by changing Celsius to Fahrenheit as the unit to transform negative values for creating the regression models.

After we obtained the data from the aforementioned sources, we joined the explanatory variables (religion, happiness score, annual average temperature, number of psychiatrists per 100,000 residents) with suicide rates by country.

We used the religion data set, joined with the suicide rates, to create a boxplot (Refer to Appendix A.1) and conducted an ANOVA test and a tukey's HSD test (Refer to Appendix A.2, A.3).

For creating the regression models and graphs, we joined the data sets containing information about happiness scores, annual average temperature and number of psychiatrists per 100,000 residents by country. We also added an income-category column (high income, upper middle income, lower middle income and low income). We created three different graphs

for the three explanatory variables faceted by the income-category. We applied a log-transformation on the psychiatrists per capita data since we noticed that the data was heavily right skewed (histogram can be found in Appendix B.2b). We then split up the data set containing suicide rates, the explanatory variables, income-category, into four different data sets by the income-category. We used this for creating four regression models of suicide rates per 100,000 residents vs. happiness scores, annual average temperature and number of psychiatrists per 100,000 residents.

When we joined the data sets to create the bigger data set, we lost countries for which one of the data sets did not have values so for our final data set we the number of observations corresponded to the data set with the lowest number of observations. This ensured that all the data was available for all the countries that we included.

**Results and Discussion:**

From our ANOVA table (Refer to Appendix A.2), we obtained a result that is significant at the level of 0.01. Therefore, we concluded that at least one pair of means are statistically significant. We then used the Tukey's test to further our analysis and found that the difference between Islam and Buddhism and the difference between Islam and Christian are significant (Refer to Appendix A.3). We then drew the conclusion that countries in which Islam dominates have lower suicide rate then countries in which Christianity or Buddhism dominate.

From our data on Happiness Score, Psychiatrists Per Capita and Annual Average Temperature of respective countries, we created three visualizations as seen in Appendix B (Note that 'Psychiatrists Per Capita' underwent a log-transformation as explained above). Moreover, since poverty increases the propensity towards suicide[5], we decided to facet our plots by the Income Level of each country as defined by the World Bank.

Appendix B.1 describes the relationship between happiness score and suicide rate per capita. Before our analysis, we hypothesized that a higher happiness score would be correlated with a lower suicide rate. Our visualization reveals that this trend is, in general, true. However, the strength of correlation varies across the different income groups.

Appendix B.2.a graphs the log of psychiatrists per capita and suicide rate per capita of each country. Contrary to our initial thought that a higher number of psychiatrists would be correlated with a lower suicide rate, we notice that, in higher income countries, a higher number of psychiatrists is correlated to a higher suicide rate. This can be explained by considering the fact that high income countries with high suicide rate have enough capital and incentive to educate psychiatrists. However, in Low Income and Lower Income Countries, a higher number of psychiatrists is correlated with a lower suicide rate.

Next, we examine the relationship between the annual average temperature and suicide rate per capita with Appendix B.3. We hypothesized there is a negative correlation between average temperature and suicide rate per capita. This was true for all income groups except for low income countries. This can be explained by the fact that extreme (high) temperatures would be more of an inconvenience than a luxury, due to the possibility of heat strokes and dehydration caused by inability to afford cooling provisions.

It should be noted that the correlation between suicide rate and the respective explanatory variables is weak in all cases.

We created four regression models for the four income groups (High income, Upper middle income, Lower middle income and Low income). Each has varying different R-squared values as can be seen in Appendix C (High income: 35.8%, Upper middle income: 7.427%, Lower middle income: 8.815% and Low income: 23.18%).

Interpreting the regression model for high income groups (Refer to Appendix C.1), we can see that there is a positive correlation between suicide rates per 100,000 people and Psychiatrists per 100,000 residents and a negative correlation between suicide rates per 100,000 people and happiness score and temperature. For every increase in psychiatrists per 100,000 residents by one unit, the suicide rates per 100,000 people goes up by 0.01340% holding everything else constant. For every increase in Happiness score by one unit, the suicide rates per 100,000 people goes down by 1.71007% holding everything else constant. Similarly, For every increase in Annual average temperature by one Fahrenheit, the suicide rates per 100,000 people goes down by 0.18574% holding everything else constant. The percentage of variation in suicide rates that is explained by our explanatory variables (our model) is 35.8% (r-squared value), which is relatively high. Further, the p-value for the model is 0.002415, which is statistically significant indicating that this is a strong model. The normal probability plot (Refer to Appendix C.1.a) has a reverse s-shape which indicates that we might have more variance than one would expect in a normal distribution.

The other regression models, the details of which can be found in Appendix C.2, C.3 and C.4 can be interpreted in a similar way.

**Conclusion and Future Work**

From our ANOVA test we concluded that since there was a significant difference in suicide rates between Islamic countries and Buddhist or Christian countries, it implies that there is a correlation between what the predominant religion is in a country and the suicide rates of the country.

Next our visualizations indicate a negative correlation between happiness score and suicide rates within each income category. However, since this relation is weak, we cannot make a stronger conclusion. Next, there is a negative correlation between log of number of psychiatrists and suicide rates for lower income countries and a positive correlation between log of number of psychiatrists and suicide rates for higher income countries. Although there can be explanations for this difference, we again, cannot conclude something of significance. By the third visualization we can see that there is a negative correlation between average temperature and suicide rates for the higher three income groups, but there is a positive correlation between average temperature and suicide rates for the lowest income group. Again, there is no strong correlation between the two variables.

Finally, from our three-variable regression model, although the R-squared value is relatively high and the p-value agrees with the relevance of our model, due to the reverse s-shaped normal probability plot, its not correct to assume normality; it might present some weakness.

This exploration into factors that affect suicide rates is a start but is not complete; since there are several other factors (alcohol usage, unemployment rates, health indicators, gender differences, government policies) that could influence suicide rates, including them into models

would be a potential extension of this exploration. However, since suicide rates depend on so many variables with varying degrees of significance and will always have anomalies across countries, coming up with a single pertinent model that predicts suicide rates is difficult, but could prove potentially useful.

**References**


1. Page, Lisa A., et al. “Relationship between Daily Suicide Counts and Temperature in England and Wales.” British Journal of Psychiatry, vol. 191, no. 02, Aug. 2007, pp. 106–112., doi:10.1192/bjp.bp.106.031948.

2. Daly, Mary C., and Daniel J. Wilson. “Happiness, Unhappiness, and Suicide: An Empirical Assessment.” Journal of the European Economic Association, vol. 7, no. 2-3, 1 May 2009, pp. 539–549., doi:10.1162/jeea.2009.7.2-3.539.

3. Jarvis, George K., and Herbert C. Northcott. “Religion and Differences in Morbidity and Mortality.” Social Science & Medicine, vol. 25, no. 7, 1987, pp. 813–824., doi:10.1016/0277-9536(87)90039-6.

4. Stein, Dan J., et al. “The Cross-National Epidemiology of Mental Disorders.” PLOS Medicine, 11 Aug. 2009, pp. 3–8., doi:10.1017/9781316336168.002.

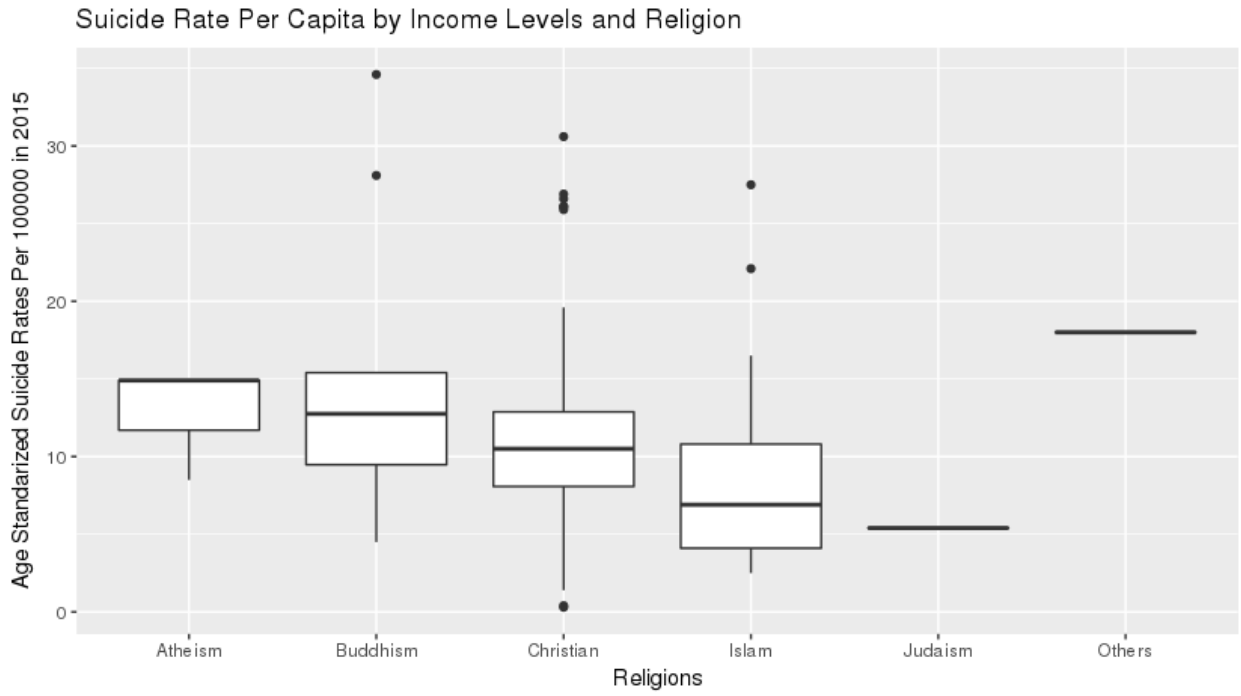
5. Stack, Steven. “Suicide: A 15‐Year Review of the Sociological Literature Part I: Cultural and Economic Factors.” Suicide and Life-Threatening Behavior, 30 Dec. 2010, doi:10.1107/s0108270109054134/gg32073asup4.hkl.

6. “The Importance of Research in Understanding Suicide.” Austen Riggs Center, www.austenriggs.org/blog-post/importance-research-understanding-suicide.

7. “The Relationship Between Suicide and Mental Health.” Canadian Association for Suicide Prevention, 25 May 2016, suicideprevention.ca/the-relationship-between-suicide-and-mental-health/.

## Appendix:

## Appendix A.1




## Appendix A.2

Anova Table

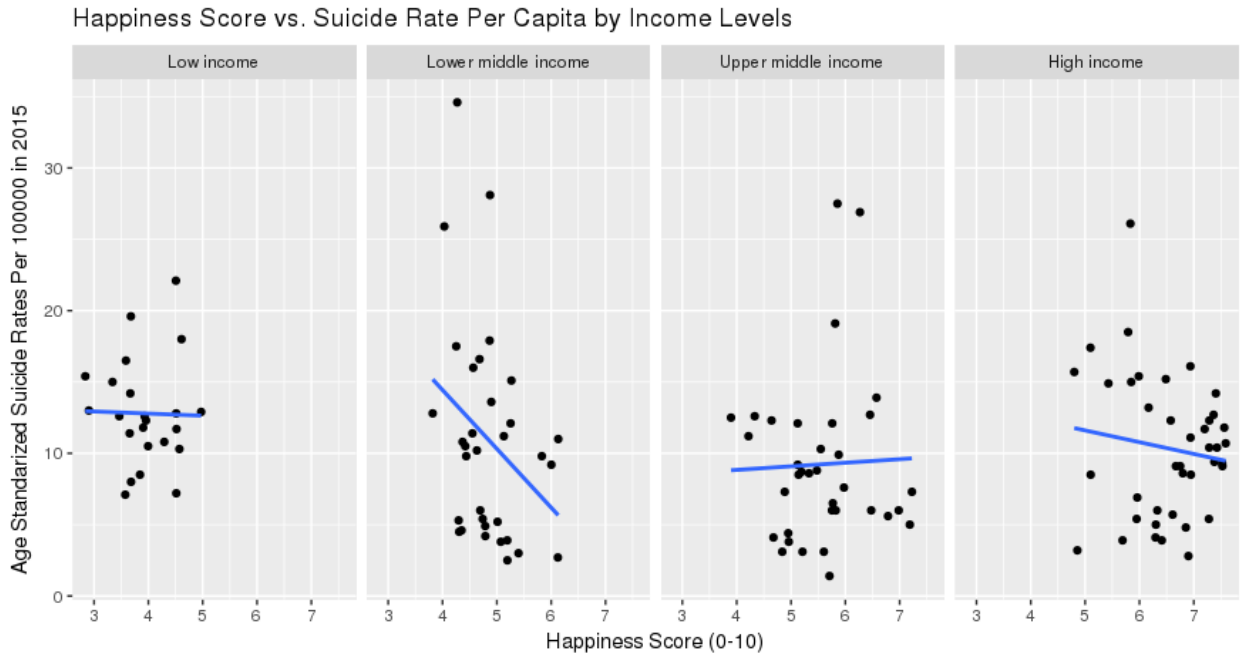
```
            Df Sum Sq Mean Sq F value  Pr(>F)
Religion     5    609   121.7   3.734 0.00308 **
Residuals  176   5738    32.6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
37 observations deleted due to missingness
```

## Appendix A.3

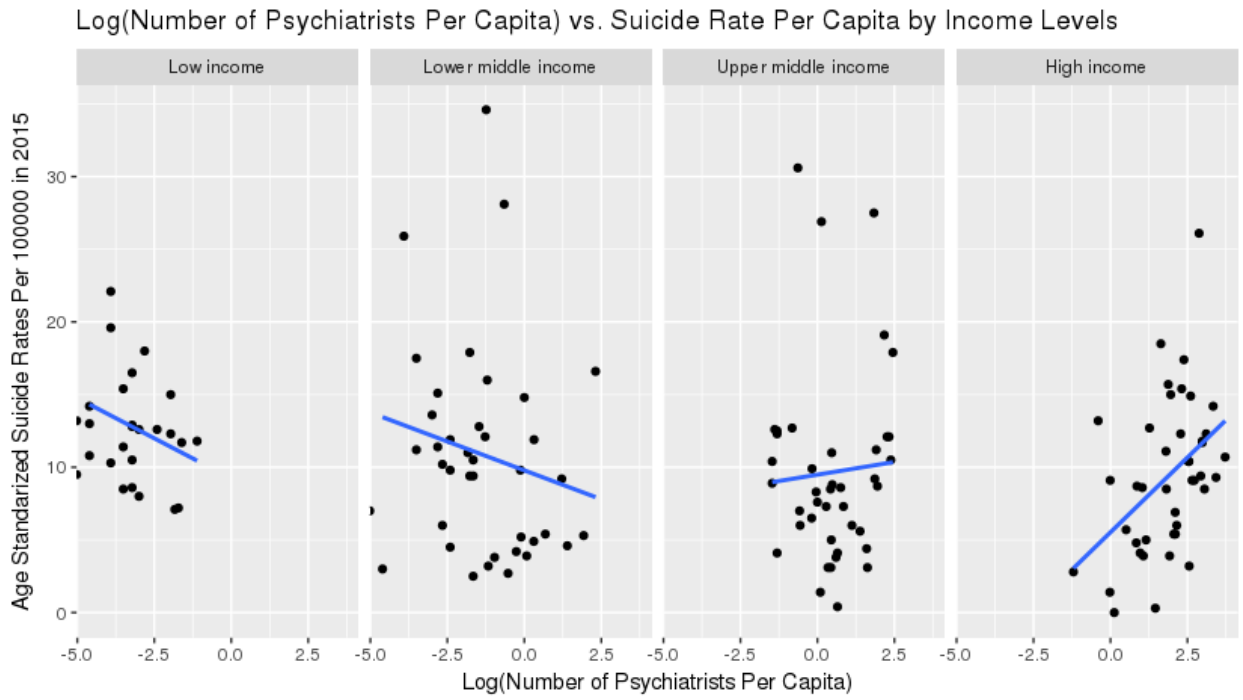
Result from Tukey Method

```
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = AgeStandardizedSuicideRatesPer100000Population2015 ~ Religion, data =
religionsuicidetemp)


$Religion
                          diff        lwr         upr     p adj
Buddhism-Atheism     2.373333  -8.456662 13.20332880 0.9885278
Christian-Atheism   -1.746328 -11.364848  7.87219293 0.9951994
Islam-Atheism       -4.587075 -14.372061  5.19791166 0.7561116
Judaism-Atheism     -7.366667 -26.363735 11.63040208 0.8736815
Others-Atheism       5.233333 -13.763735 24.23040208 0.9682793
Christian-Buddhism  -4.119661  -9.538188  1.29886555 0.2473494
Islam-Buddhism      -6.960408 -12.669213 -1.25160338 0.0073264
Judaism-Buddhism    -9.740000 -26.994945  7.51494458 0.5822026
Others-Buddhism      2.860000 -14.394945 20.11494458 0.9968771
Islam-Christian     -2.840747  -5.636742 -0.04475197 0.0440664
Judaism-Christian   -5.620339 -22.141848 10.90116971 0.9237165
Others-Christian     6.979661  -9.541848 23.50116971 0.8279641
Judaism-Islam       -2.779592 -19.398565 13.83938139 0.9967404
Others-Islam         9.820408  -6.798565 26.43938139 0.5318866
Others-Judaism      12.600000 -10.666563 35.86656252 0.6256402
```

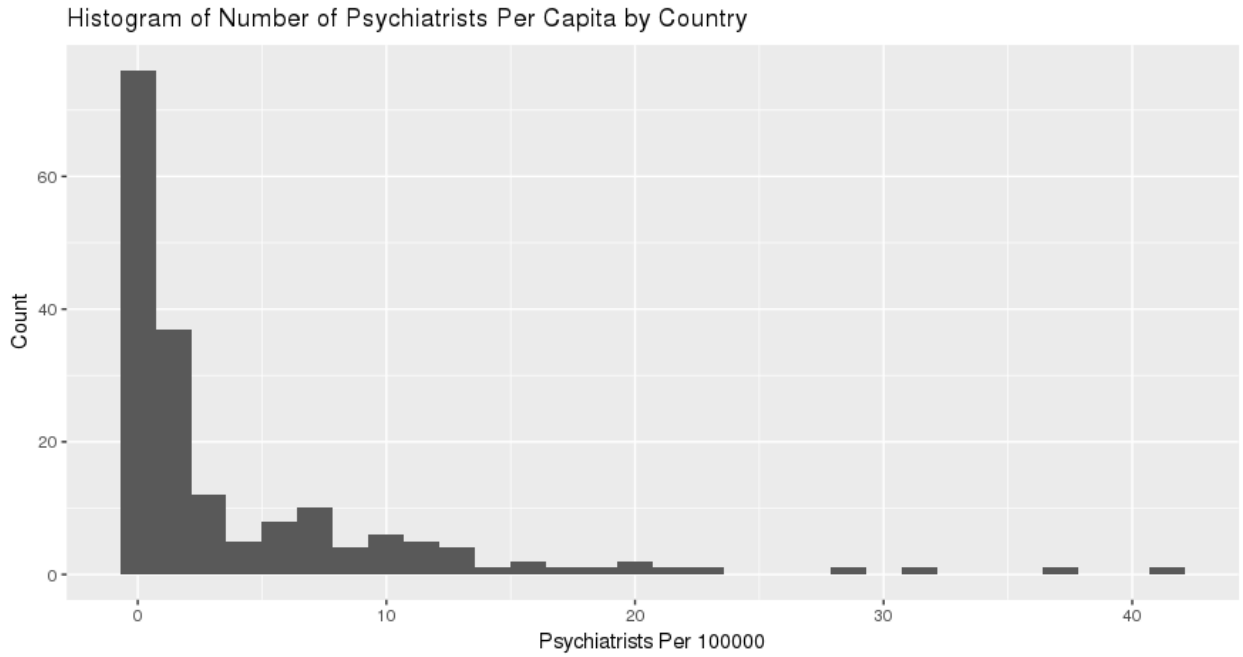
## Appendix B.1



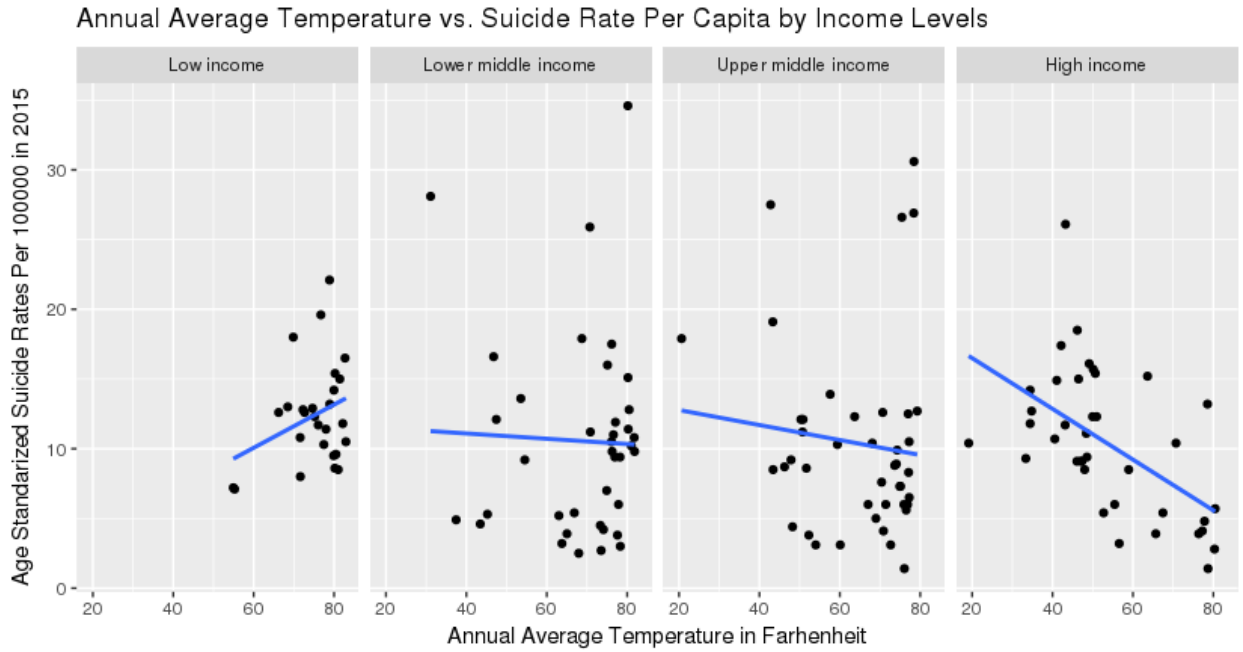

## Appendix B.2.a



## Appendix B.2.b




## Appendix B.3



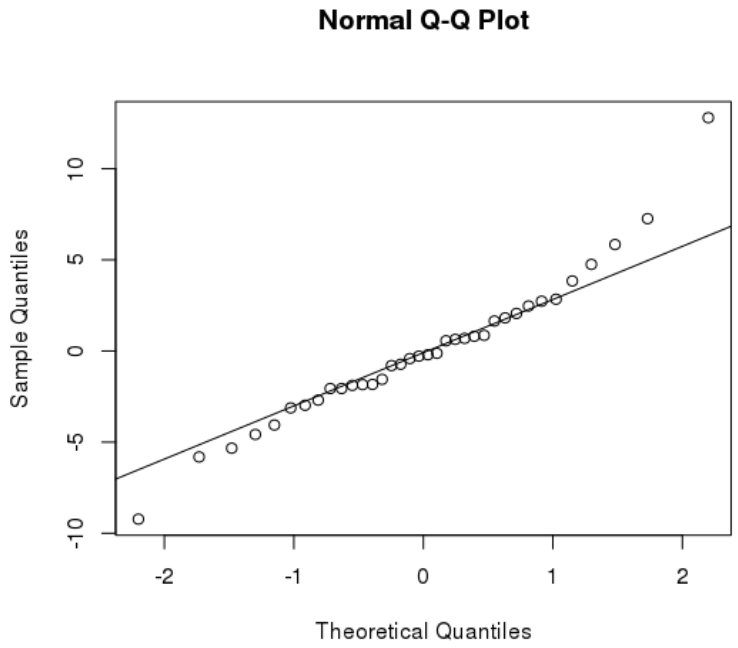
## Appendix C.1

Summary from Regression model for High Income

```
Call:
lm(formula = AgeStandardizedSuicideRatesPer100000Population2015 ~
    PsychaiatristsPer100000 + Happiness.Score + Annual.Temperature,
    data = HighIncomeVariables)

Residuals:
    Min      1Q  Median      3Q     Max
-9.2243 -2.0595 -0.2456  1.8767 12.7986

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            31.06835    6.59798   4.709 4.62e-05 ***
PsychaiatristsPer100000  0.01340    0.10000   0.134  0.89428
Happiness.Score        -1.71007    0.91578  -1.867  0.07104 .
Annual.Temperature     -0.18574    0.05656  -3.284  0.00248 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.168 on 32 degrees of freedom
  (11 observations deleted due to missingness)
Multiple R-squared:  0.358,  Adjusted R-squared:  0.2979
F-statistic: 5.949 on 3 and 32 DF,  p-value: 0.002415
```

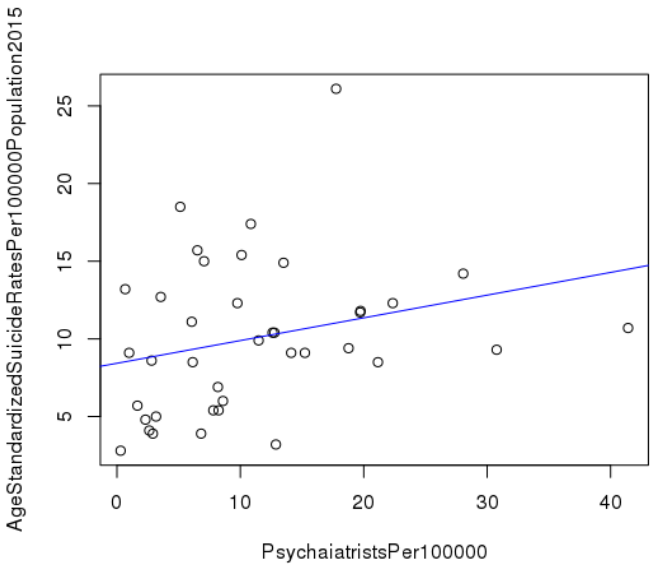
**Appendix C.1.a**

Normal Probability Plot



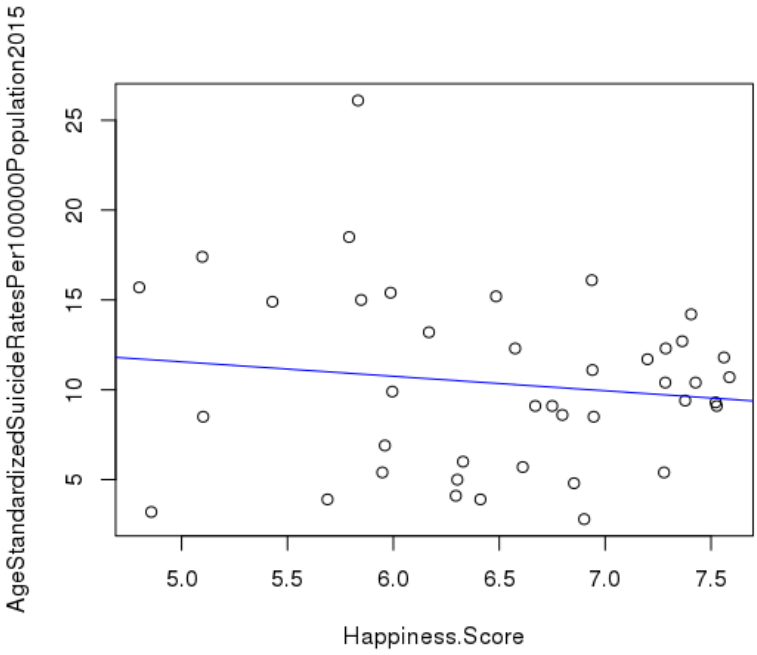

**Appendix C.1.b**

Plot For Suicide Rate vs. Psychiatrists Per 100000 for High Income



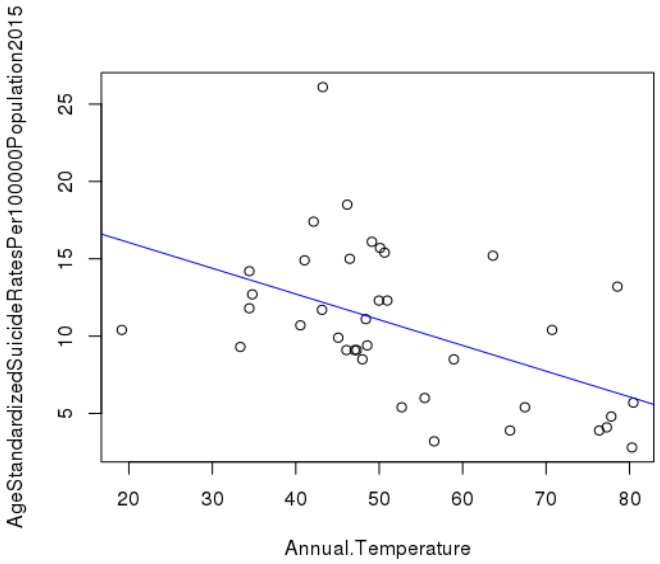
**Appendix C.1.c**

Plot For Suicide Rate vs. Happiness Score for High Income




**Appendix C.1.d**

Plot For Suicide Rate vs. Psychiatrists Per 100000 for High Income



## Appendix C.2

Summary from Regression model for Upper Middle Income

```
Call:
lm(formula = AgeStandardizedSuicideRatesPer100000Population2015 ~
    PsychaiatristsPer100000 + Happiness.Score + Annual.Temperature,
    data = UpperIncomeVariables)

Residuals:
   Min     1Q Median     3Q    Max
-7.230 -3.743 -1.095  1.115 18.587

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)             6.82048    9.91015   0.688    0.497
PsychaiatristsPer100000 0.51354    0.56650   0.907    0.373
Happiness.Score         0.30895    1.52802   0.202    0.841
Annual.Temperature     -0.01314    0.13994  -0.094    0.926

Residual standard error: 6.229 on 27 degrees of freedom
  (9 observations deleted due to missingness)
Multiple R-squared:  0.07427,Adjusted R-squared:  -0.02859
F-statistic: 0.722 on 3 and 27 DF,  p-value: 0.5476
```

## Appendix C.3

Summary from Regression model for Lower Middle Income

```
Call:
lm(formula = AgeStandardizedSuicideRatesPer100000Population2015 ~
    PsychaiatristsPer100000 + Happiness.Score + Annual.Temperature,
    data = LowerIncomeVariables)

Residuals:
   Min     1Q Median     3Q    Max
-9.735 -5.441 -1.427  3.936 22.490

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)            33.14844   13.95229   2.376   0.0239 *
PsychaiatristsPer100000 -0.56082    0.73565  -0.762   0.4516
Happiness.Score         -3.55992    2.21209  -1.609   0.1177
Annual.Temperature      -0.07075    0.10608  -0.667   0.5098
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.507 on 31 degrees of freedom
  (8 observations deleted due to missingness)
Multiple R-squared:  0.08815,Adjusted R-squared:  -9.891e-05
F-statistic: 0.9989 on 3 and 31 DF,  p-value: 0.4064
```

## Appendix C.4

Summary from Regression model for Low Income

```
Call:
lm(formula = AgeStandardizedSuicideRatesPer100000Population2015 ~
    PsychaiatristsPer100000 + Happiness.Score + Annual.Temperature,
    data = LowIncomeVariables)

Residuals:
    Min      1Q  Median      3Q     Max
-6.0399 -1.6541 -0.1523  1.2592  7.6708

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)             -1.7094     9.7348  -0.176   0.8625
PsychaiatristsPer100000 -10.9025     9.6016  -1.135   0.2703
Happiness.Score           0.2906     1.3778   0.211   0.8352
Annual.Temperature        0.1908     0.1003   1.903   0.0723 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.554 on 19 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.2318, Adjusted R-squared:  0.1106
F-statistic: 1.911 on 3 and 19 DF,  p-value: 0.162
```

## Appendix D

| Variable | Source |
|---|---|
| Suicide Rate | World Health Organization (WHO) |
| Happiness Index | World Happiness Report |
| Number of Psychiatrists | World Health Organization (WHO) |
| Religion | World Religion Project |
| Temperature | The World Bank |
| Income Group | The World Bank |