<Group 1>

Effect of Gender, Age Group, and Human Development Index on Suicide Rate

2019. 12. 10

Seim An
Rutvik Bapat
Akash Chowdhury
Ngone Lo
Mounica Thanam
Yuxin Yan

A final paper submitted in conformity with the requirements

for the course of INF1344H

Faculty of Information

University of Toronto

ABSTRACT

This paper aims to discover the variation in suicide rate regarding gender, age group and Human Development Index (HDI) for 87 countries. The datasets are sourced from Kaggle and United Nations Human Development Reports.

A study conducted across 91 countries [1] stated that men were more likely to commit suicide than women. The same study also found evidence of a linear correlation between HDI and suicide rate. Another research [2] found that suicide rate varied across age groups.

To examine the above-mentioned hypothesis, we conducted various statistical tests - descriptive statistics, T-test, ANOVA and OLS Regression. Our results demonstrated that men have higher suicide rate than women; both men and women with the age group of 55+ have the highest suicide rate; in case of HDI, a linear relation was found, however, the variance in suicide rate was not properly captured by HDI.

ACKNOWLEDGEMENTS

We would like to express gratitude to Professor Jia Xue and Teaching Assistant Jiannan Li.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The World Health Organization has estimated that close to 800,000 people die by suicide every year, that is 1 person every 40 seconds. Suicide has become one of the leading causes of death globally. Studies [1][2] have shown that several factors such as hopelessness, mental disorders, alcoholism as well as stressful events, such as financial problems and interpersonal communication difficulties, are associated with suicidal behavior.

Looking at the global suicide rate for the past few decades [4], the male population had a higher suicide rate than the female population. Studies [5] suggest that men value independence and decisiveness, and they regard acknowledging a need for help as weakness and hence end up in suicide. Also, the greater suicide rate in men could partly be due to their greater rate of alcoholism [5].

In the year 2016, suicide was considered as the second leading cause of death among 15-29-year-olds globally [4]. Meanwhile, based on a study [3] across the United States in 2017, the National Institute of Mental Health (NIMH) stated that the suicide rate was highest for men among the age group 65 and older, and 45-54 among women. In our discussion, we combined the age groups to reduce the number of age groups and also to test the hypothesis mentioned by the NIMH in our dataset. The final combined age groups in our dataset were 5-24, 25-34, 35-54 and 55 + years.

Socioeconomic conditions such as unemployment, poverty, and homelessness may also increase the risk of suicide. One of the measures for socioeconomic conditions is the Human Development Index (HDI). HDI discovers the changes in the level of human development over time in each country [6]. In order to acquire HDI, three principal dimensions are needed: “mean years of schooling and expected years of schooling, life expectancy at birth, and gross national income per capita” [7]. HDI is the geometric mean of normalized indices for each of the three dimensions.

Looking at the above reasons and studies, we realized that suicide has become a public health concern for all countries, and we believe that it takes global efforts to help reduce the suicide rate. Hence, we decided to explore these factors against suicide rate.

1.1. RESEARCH QUESTION

Our research question is focused towards studying the relationship between suicide rate, gender, age group, and HDI, and they are represented as follows:

- Is there any gender biases influencing the suicide rate?
- Does the age group have an impact on the suicide rate?
- Does the HDI of a country which points at different socio-economic and human factors have any effect on the suicide rate ?

1.2. HYPOTHESIS

We seek to check the existing hypotheses in the past work done by researchers, in our dataset, hence our hypothesis for various tests is of the form,

1. T- test

   H0: Men and women have the same suicide rate

   Ha: Men have a higher suicide rate than women

2. ANOVA

   H0: All age groups have the same suicide rate

   Ha: At least one of the age groups has a significantly different suicide rate

3. Ordinary Least Square Regression

   H0: There is no linear relationship between HDI and suicide rate

   Ha: There is a linear relationship between HDI and suicide rate

LITERATURE REVIEW

Our literature review comprises of three articles. The first article “Suicide rate in relation to the Human Development Index and other health-related factors: A global ecological study from 91 countries” (Khazaei et. al., 2017) stated that, across 91 countries, with data obtained from the World Bank Report 2013, men were more likely to commit suicide than women, as it showed suicide rates of 16.3 per 100,000 for men and 4.6 per 100,000 for women.

The same study also found “a significant direct correlation between HDI levels of countries and the incidence of suicide”. They discussed that results from similar studies were contradictory - some showed an inverse relationship between income levels (as a component of HDI) and suicide rate, other studies showed otherwise.

The second article “Suicide” (Ritchie et. al., 2019) indicated that even though suicide was a leading cause of death among young people, it did not necessarily mean that suicide was more likely to occur on young people than old people. In fact, by comparing the suicide rate between age groups of 5-14, 15-49, 50-69 and 70+ through the year 1990 to 2017, “suicides globally follow a standard pattern of the older the age group, the higher the death rate”.

The third study “Suicide” by the National Institute of Mental Health (NIMH) of the United States discussed both gender and age groups. It pointed out that “among females, the suicide rate was highest for those aged 45-54 (10.0 per 100,000). Among males, the suicide rate was highest for those aged 65 and older (31.0 per 100,000).”

METHODS

3.1. DATASET

The main dataset considered to test the hypothesis, called “Suicide Rates Overview 1985 to 2016”, was sourced from Kaggle [8]. It compares socio-economic information with suicide rates by year and country. This compiled dataset pulled from four other datasets linked by time and place and was built to find signals correlated to increased suicide rates among different cohorts globally, across the socio-economic spectrum.

It originally consisted of 12 variables, 27820 samples for 102 countries from 1985 to 2016. While checking the database as part of data pre-processing, we sourced the missing values for HDI from the UN Development Programme Human Development Reports [7].

For the purpose of our study, the dataset was restricted to the timeline 2011-2015, 87 countries, and the variables we focused on included country, gender, age-group, number of suicides, population and HDI. We also calculated variables such as the average suicide rate and average HDI.

3.2. MEASUREMENT

Variable	Categories / Scale	Level of Measurement
Independent Variable		
Gender	Male, Female	Nominal
Age-Group	5-24 years 25-34 years 35-54 years 55+ years	Ordinal
HDI	0.63 - 0.95	Ratio
Dependent Variable		
Suicide Rate per 100K	0 - 81	Ratio

3.2.1. DEPENDENT VARIABLE

The dependant variable is the suicide rate per 100K which has a scale of 0 - 81. Its level of measurement is ratio.

3.2.2. INDEPENDENT VARIABLES

Following are descriptions of the independent variables considered for the project:

- Gender: Male and Female. Its level of measurement is nominal.
- Age-group: 5-24 years, 25-34 years, 35-54 years and 55+ years. Its level of measurement is ordinal.
- HDI: Measured on a continuum of 0.63-0.95. Its level of measurement is ratio.

3.3. DATA ANALYSIS

For the first part of our data analysis, we did some data exploration. Regarding gender, box plot diagrams by gender as well as descriptive statistics with regards to mean, standard deviation and sample size were used to get a better understanding of the distribution of suicide rate for males and females. We also ran a t-test to test for a significant difference in the mean suicide rate for men and women. Regarding age group, focusing on the general population first, box plot diagrams by age group as well as descriptive statistics with regards to mean, standard deviation and sample size were used to get a better understanding of the distribution of suicide rate for each age group. Additionally, to determine if there is a significant difference in the mean suicide rate by age group, we ran an ANOVA test. To get a more specific comparison of the suicide rate between the age groups, we also ran the Tukey's Honest Significant Difference test, which compares between every two groups. The same procedure was used for the male population and for the female population. Finally, for HDI, we also did some descriptive analysis and used box plot diagrams to get a glimpse of its distribution.

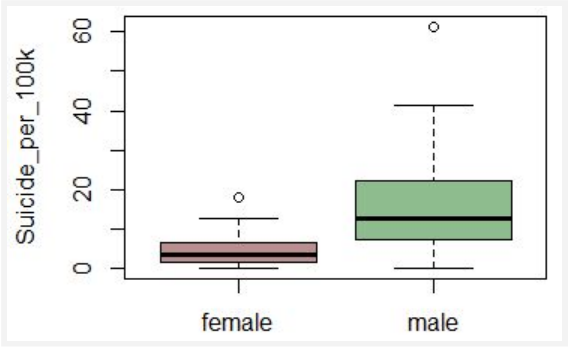
For the second part of our data analysis, we constructed two linear regression models. The first model looks at the suicide rate with relation to HDI on its own. For the second model, we constructed a linear regression model with men and 55+ years as reference groups.

RESULTS & DISCUSSION

4.1. DESCRIPTIVE & COMPARATIVE ANALYSIS RESULTS

Figure 1 shows the box plots of the suicide rate for males and females. *Table 1* shows descriptive statistics of the suicide rate for males and females and the result of the t-test looking for a significant difference in the mean suicide rate for men and women.

<Figure 1: Box Plots of Suicide Rate (SR) by Gender>



<Table 1: Suicide Rate (SR) Statistics by Gender>

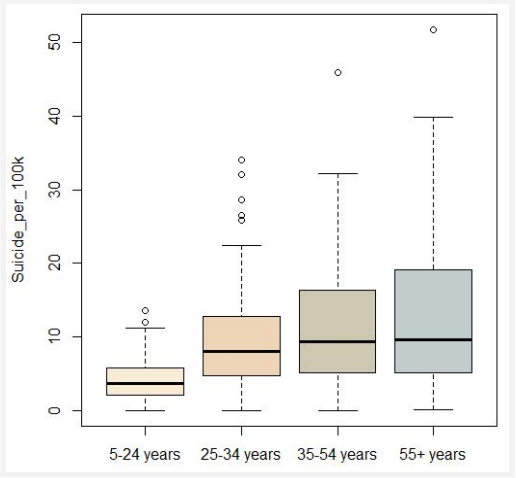
	N	Mean Suicide Rate per 100k (SD)
Women	87	4.236 (3.385)
Men	87	15.439 (11.582)
t(101)= 8.659***		

***p<.001;**p<.01;*p<.05.

For both men and women, the distribution of the suicide rate is slightly right-skewed. We noted the presence of at least one outlier for both men and women. Since we have 87 countries in our final dataset, both genders have a sample size of 87. The mean and standard deviation of men's suicide rate is much higher than women's. For the t-test to test for a significant difference in the mean suicide rate of men and women, since $p < 0.05$, we reject our null-hypothesis. Hence, men have a significantly higher suicide rate than women.

Figure 2 shows the box plots of suicide rate by age group while *Table 2* shows descriptive statistics of suicide rate for each age group as well as the results of the ANOVA and Tukey tests to determine a significant difference in the mean suicide rate by age groups and test for significant difference between groups by pair.

<Figure 2: Box Plots of SR by Age Group>



<Table 2: SR Statistics by Age Group>

	N	Mean Suicide Rate per 100k (SD)
5-24 years	87	4.109 (2.846)
25-34 years	87	9.499 (7.289)
35-54 years	87	11.591 (8.715)
55+ years	87	12.668 (10.229)
F(3, 344)=20.85***		
Difference		
55+ years - 5-24 years		8.559***
55+ years - 25-34 years		3.169*
55+ years - 35-54 years		1.077
35-54 years - 25-34 years		2.093
35-54 years - 5-24 years		7.482***
25-34 years - 5-24 years		5.389***

***p<.001;**p<.01;*p<.05.

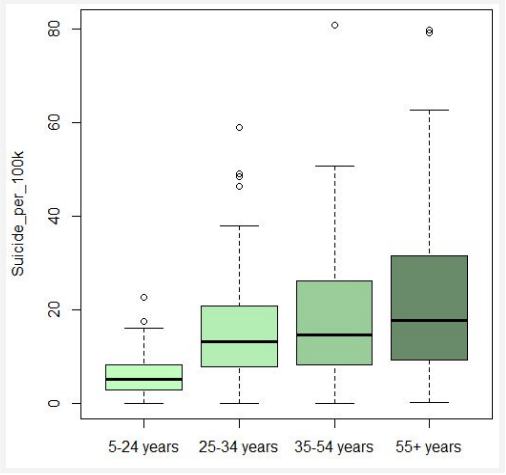
The distribution of the suicide rate is slightly right-skewed for all age groups. We also noted the presence of at least one outlier for all age groups. Since we have 87 countries in our final dataset, all age groups have a sample size of 87. The age group of 55+ years has the highest mean and standard deviation. For the ANOVA test, since $p < 0.05$, we reject our null-hypothesis. Hence, there is at least one age group with a significantly different mean suicide rate than the rest. Furthermore, from the results of the Tukey test, we can conclude that the differences in the mean suicide rate between 55+ years and 5-24 years as well as 55+ years and 25-34 years are statistically significant. However, the difference between the

55+ years and 35-54 years is not significant. Other age groups with a significant mean suicide rate between them are 35-54 years and 5-24 years, and 25-34 years and 5-24 years.

When we look at men only, or women only, the trend is the same compared to the general population: the box plots show a slightly right-skewed distribution across all age groups. The same trend also applies to the descriptive statistics of suicide rate by age group for both the male and female population: the age group 55+ years has the highest mean and standard deviation. The ANOVA and Tukey tests also show similar outcomes for both genders: there is at least one age group with a significantly different mean suicide rate than the rest; the differences between 55+ years and 5-24 years, 55+ years and 25-34 years, 35-54 years and 5-24 years, and 25-34 years and 5-24 years are statistically significant while the difference between 55+ years and 35-54 years is not significant.

Figures 3 and *4* and *Tables 3* and *4* show the box plots of suicide rate by age group, descriptive statistics of the suicide rate for each age group as well as the results of the ANOVA and Tukey tests for the male and female population respectively.

<Figure 3: Box Plots of SR by Age Group for Men>

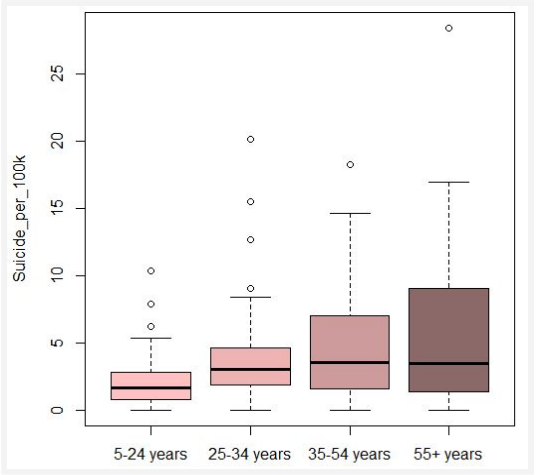


<Table 3: SR Statistics by Age Group for Men>

	N	Mean Suicide Rate per 100k (SD)
5-24 years	87	6.021 (4.266)
25-34 years	87	15.240 (11.943)
35-54 years	87	18.732 (14.336)
55+ years	87	21.905 (17.662)
F(3, 344)=24.18***		
Difference		
55+ years - 5-24 years		15.884***
55+ years - 25-34 years		6.665**
55+ years - 35-54 years		3.172
35-54 years - 25-34 years		3.493
35-54 years - 5-24 years		12.712***
25-34 years - 5-24 years		9.219***

***p<.001;**p<.01;*p<.05.

<Figure 4: Box Plots of SR by Age Group for Women>



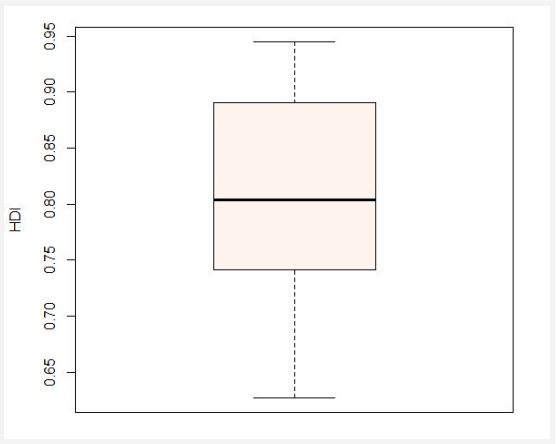
<Table 4: SR Statistics by Age Group for Women>

	N	Mean Suicide Rate per 100k (SD)
5-24 years	87	2.099 (1.789)
25-34 years	87	3.662 (3.217)
35-54 years	87	4.605 (3.787)
55+ years	87	5.462 (5.244)
F(3, 344)=12.95***		
Difference		
55+ years - 5-24 years		3.363***
55+ years - 25-34 years		1.800**
55+ years - 35-54 years		0.857
35-54 years - 25-34 years		0.943
35-54 years - 5-24 years		2.506***
25-34 years - 5-24 years		1.563*

***p<.001;**p<.01;*p<.05.

Finally, *Figure 5* shows a box plot of HDI. HDI is relatively evenly spread out. The mean is 0.808, the standard deviation is 0.0862, and the sample size is 87.

<Figure 5: Box Plot of HDI>



4.2. OLS REGRESSION RESULTS

We build two linear regression models involving gender, age group and HDI as independent variables and suicide rate as the dependent variable. Model 1 looks at the suicide rate with relation to HDI on its own, and Model 2 at suicide rate with relation to gender age group, with men and 55+ years as reference groups. *Table 5* shows the results of the two models.

<Table 5: Linear Regression Models of Suicide Rate>

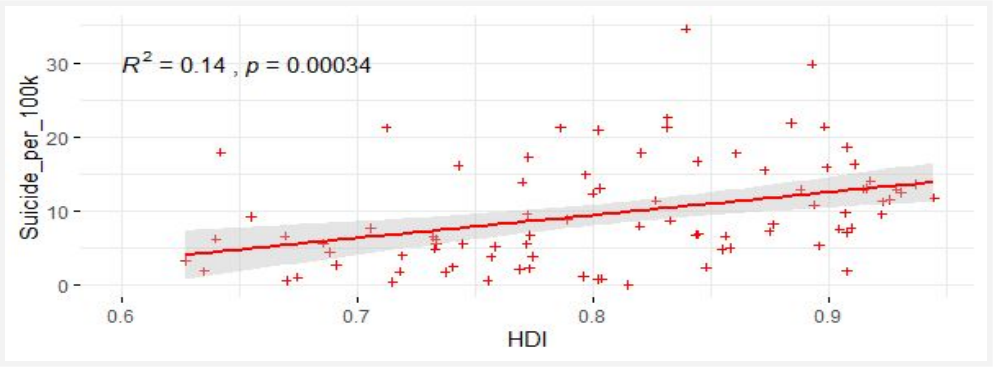
Variable	Model 1		Model 2	
	b	se	b	se
HDI	30.858***	8.263	-	-
Female	-	-	-11.518***	0.746
Age Group	-	-	-	-
5-24 years	-	-	-9.624***	1.056
24-34 years	-	-	-4.233***	1.056
35-54 years	-	-	-2.015	1.056
Intercept	-15.265*	6.715	19.442***	0.835
R-square	0.141***	-	0.3227***	-
N	87	-	696	-

***p<.001;**p<.01;*p<.05.

For Model 1, we can see that HDI explains about 14% of the variance found in the suicide rate. The intercept -15 translates into a negative suicide rate when HDI is zero. However, in real life, HDI can never be zero. Hence, the intercept does not make sense here. Suicide rate increases by 30 for every one unit increase in HDI. For the second Model 2, looking at the intercept, we would expect a suicide rate of 19 per 100k for men aged 55+ years. The beta coefficients are negative because our reference groups have the highest suicide rate. Compared to men, we would expect women to have a suicide rate 11.518 lower as long as average the age group is the same. Regarding age group, as long as the gender is the same, compared to the 55+ years group, the 5-24 years group has a mean suicide rate

9.624 lower, the 35-54 years group has a mean suicide rate 2.015 lower, and the 24-34 years group has a mean suicide rate 4.233 lower. All beta coefficients are significant except for the 35-54 years groups. Model 2 explains the variance found in the suicide rate better at 32 %. Both models are statistically significant. Figure 6 shows a visualization of Model 1.

<Figure 6: Visualization of Model 1: SR in relation to HDI>



Although weak, we can see a positive correlation.

We did not run a model with HDI, age group and gender together as independent variables, because for any country, all age groups and gender classes share the same HDI, making these three variables not isolated from each other.

CONCLUSION

In conclusion, our results suggest that male gender, higher age group and HDI components are associated with increased risk of suicide. Additionally, gender and age group explain about 32% of the variance found in suicide rate while HDI explains about 14% of that variance. Hence, detecting stresses with a high suicide risk and reducing social factors influencing these is necessary to prevent an epidemic in the world. Community centric programs and availability of economic resources may reduce suicide rates.

We framed our analysis in the context of the gendered socialisation of men and women which might influence different psychological factors related to suicide, and the probability of seeking help. We can assume that these in turn contribute to the variation in suicide rates with respect to gender. We also, however, note that suicide rates of men and women is not enough by itself to validate this hypothesis, as we exist in a complex web of societal forces, but the data gathered has been restricted to only two genders. The studies taken into account have only been focussed on suicide rate, which could overlook the possibility of men using more lethal means to commit suicide when compared to women, and the number of non-binary identifying people committing suicide or their suicide rate is currently unknown in this study. Similarly, there might be varying factors that have not been taken into account since the data is itself has not captured it. It is against this backdrop that we should interpret the result of men having higher suicide rate with a deeper lens.

REFERENCES

[1] Khazaei, S., Armanmehr, V., Nematollahi, S., Rezaeian, S., & Khazaei, S. (2017). Suicide rate in relation to the Human Development Index and other health related factors: A global ecological study from 91 countries. *Journal of Epidemiology and Global Health*, *7*(2), 131–134. doi: 10.1016/j.jegh.2016.12.002

[2] Ritchie, H., Roser, M., & Ortiz-Ospina, E. (2015, June 15). Suicide. Retrieved December 8, 2019, from https://ourworldindata.org/suicide.

[3] National Institute of Mental Health. (2019, April). Suicide. Retrieved December 10, 2019, from https://www.nimh.nih.gov/health/statistics/suicide.shtml.

[4] World Health Organization. (2019). Evolution 1950-2000 of global suicide rates (per 100'000). Retrieved December 6, 2019, from https://www.who.int/mental_health/prevention/suicide/evolution/en/.

[5] Murphy, G. E. (1998). Why women are less likely than men to commit suicide. *Comprehensive Psychiatry*, *39*(4), 165–175. doi: 10.1016/s0010-440x(98)90057-8

[6] Chappelow, J. (2019, October 20). Human Development Index (HDI) Definition. Retrieved December 3, 2019, from https://www.investopedia.com/terms/h/human-development-index-hdi.asp.

[7] United Nations Development Programme. (n.d.). Human Development Index (HDI). Retrieved December 10, 2019, from http://hdr.undp.org/en/content/human-development-index-hdi.

[8] Suicide Rates Overview 1985 to 2016: Compares socio-economic info with suicide rates by year and country. Retrieved Oct 22, 2019, from https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016

LIST OF APPENDICES

Descriptive statistics R output for suicide rate by gender

```
> tapply(SRGENDER$Suicide_per_100k, SRGENDER$Gender, mean)
   female      male
 4.235511 15.438710
> tapply(SRGENDER$Suicide_per_100k, SRGENDER$Gender, sd)
   female      male
 3.385451 11.582511
> tapply(SRGENDER$Suicide_per_100k, SRGENDER$Gender, length)
female   male
    87     87
```

T-test R output for suicide rate by gender

```
> t.test(SRGENDER$Suicide_per_100k ~ SRGENDER$Gender, data=SRGENDER)

        Welch Two Sample t-test

data:  SRGENDER$Suicide_per_100k by SRGENDER$Gender
t = -8.6596, df = 100.59, p-value = 8.16e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -13.769745  -8.636653
sample estimates:
mean in group female   mean in group male
            4.235511            15.438710
```

Descriptive statistics R output for suicide rate by age group

```
> tapply(SRAGE$Suicide_per_100k, SRAGE$Age_Group, mean)
25-34 years 35-54 years  5-24 years   55+ years
   9.498539   11.591324    4.109088   12.668038
> tapply(SRAGE$Suicide_per_100k, SRAGE$Age_Group, sd)
25-34 years 35-54 years  5-24 years   55+ years
   7.289691    8.715347    2.846418   10.229692
> tapply(SRAGE$Suicide_per_100k, SRAGE$Age_Group, length)
25-34 years 35-54 years  5-24 years   55+ years
         87          87          87          87
```

ANOVA R output for suicide rate by age group

```
> results = aov(Suicide_per_100k ~ Age_Group, data=SRAGE)
> summary(results)
             Df Sum Sq Mean Sq F value   Pr(>F)
Age_Group     3   3782  1260.6   20.85 1.96e-12 ***
Residuals   344  20799    60.5
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Tukey test R output for suicide rate by age group

```
> TukeyHSD(results)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Suicide_per_100k ~ Age_Group, data = SRAGE)

$Age_Group
                             diff         lwr       upr     p adj
35-54 years-25-34 years  2.092785  -0.9508240  5.136394 0.2871003
5-24 years-25-34 years  -5.389451  -8.4330596 -2.345841 0.0000400
55+ years-25-34 years    3.169499   0.1258899  6.213108 0.0375950
5-24 years-35-54 years  -7.482236 -10.5258446 -4.438626 0.0000000
55+ years-35-54 years    1.076714  -1.9668951  4.120323 0.7977816
55+ years-5-24 years     8.558950   5.5153404 11.602559 0.0000000
```

Descriptive statistics R output for suicide rate by gender for men only

```
> tapply(SRAGEMALE$Suicide_per_100k, SRAGEMALE$Age_Group, mean)
25-34 years 35-54 years  5-24 years   55+ years
  15.239693   18.732209    6.020566   21.904903
> tapply(SRAGEMALE$Suicide_per_100k, SRAGEMALE$Age_Group, sd)
25-34 years 35-54 years  5-24 years   55+ years
  11.943241   14.335623    4.266471   17.661934
> tapply(SRAGEMALE$Suicide_per_100k, SRAGEMALE$Age_Group, length)
25-34 years 35-54 years  5-24 years   55+ years
         87          87          87          87
```

ANOVA R output for suicide rate by age group for men only

```
> results = aov(Suicide_per_100k ~ Age_Group, data=SRAGEMALE)
> summary(results)
             Df Sum Sq Mean Sq F value   Pr(>F)
Age_Group     3  12301    4100   24.18 3.19e-14 ***
Residuals   344  58334     170
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Tukey test R output for suicide rate by age group for men only

```
> TukeyHSD(results)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Suicide_per_100k ~ Age_Group, data = SRAGEMALE)

$Age_Group
                             diff        lwr       upr     p adj
35-54 years-25-34 years   3.492516  -1.604664  8.589697 0.2901593
5-24 years-25-34 years   -9.219127 -14.316307 -4.121947 0.0000257
55+ years-25-34 years     6.665210   1.568030 11.762390 0.0045364
5-24 years-35-54 years  -12.711643 -17.808823 -7.614463 0.0000000
55+ years-35-54 years     3.172694  -1.924486  8.269874 0.3760432
55+ years-5-24 years     15.884337  10.787157 20.981517 0.0000000
```

Descriptive statistics R output for suicide rate by gender for women only

```
> tapply(SRAGEFEMALE$Suicide_per_100k, SRAGEFEMALE$Age_Group, mean)
25-34 years 35-54 years  5-24 years   55+ years
   3.662016    4.604569    2.098636    5.461963
> tapply(SRAGEFEMALE$Suicide_per_100k, SRAGEFEMALE$Age_Group, sd)
25-34 years 35-54 years  5-24 years   55+ years
   3.270879    3.787392    1.788569    5.244074
> tapply(SRAGEFEMALE$Suicide_per_100k, SRAGEFEMALE$Age_Group, length)
25-34 years 35-54 years  5-24 years   55+ years
         87          87          87          87
```

ANOVA R output for suicide rate by age group for women only

```
> results = aov(Suicide_per_100k ~ Age_Group, data=SRAGEFEMALE)
> summary(results)
             Df Sum Sq Mean Sq F value  Pr(>F)
Age_Group     3    542  180.52   12.95 4.9e-08 ***
Residuals   344   4794   13.94
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Tukey test R output for suicide rate by age group for women only

```
> TukeyHSD(results)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Suicide_per_100k ~ Age_Group, data = SRAGEFEMALE)

$Age_Group
                               diff        lwr       upr     p adj
35-54 years-25-34 years  0.9425531 -0.5186553  2.403761 0.3437421
5-24 years-25-34 years  -1.5633803 -3.0245886 -0.102172 0.0306603
55+ years-25-34 years    1.7999466  0.3387383  3.261155 0.0086888
5-24 years-35-54 years  -2.5059334 -3.9671417 -1.044725 0.0000755
55+ years-35-54 years    0.8573936 -0.6038148  2.318602 0.4296905
55+ years-5-24 years     3.3633269  1.9021186  4.824535 0.0000000
```

Descriptive statistics R output for HDI

```
> summary(SRHDI$HDI)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.6272  0.7418  0.8037  0.8081  0.8906  0.9450
> sd(SRHDI$HDI)
[1] 0.08621155
```

Model 1 R output for suicide rate and HDI

```
> model_hdi <- lm(Suicide_per_100k ~ HDI , data= SRHDI)
> summary(model_hdi)

Call:
lm(formula = Suicide_per_100k ~ HDI, data = SRHDI)

Residuals:
    Min      1Q  Median      3Q     Max
-10.869  -4.537  -1.703   3.426  23.895

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -15.265      6.715  -2.273  0.02553 *
HDI           30.858      8.263   3.734  0.00034 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6.607 on 85 degrees of freedom
Multiple R-squared:  0.1409,    Adjusted R-squared:  0.1308
F-statistic: 13.94 on 1 and 85 DF,  p-value: 0.0003398
```

Model 2 R output for suicide rate and age, group and gender

```
> Gender_releveled <- relevel(Gender_factor, ref="male")
> Age_Group_factor <- as.factor(SRMAIN$Age_Group)
> Age_releveled <- relevel(Age_Group_factor, ref="55+ years")
> model_main_2 <- lm(Suicide_per_100k ~ Gender_releveled + Age_releveled, dat
a= SRMAIN)
> summary(model_main_2)

Call:
lm(formula = Suicide_per_100k ~ Gender_releveled + Age_releveled,
    data = SRMAIN)

Residuals:
    Min      1Q  Median      3Q     Max
-19.309  -5.615  -1.284   3.328  63.455

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               19.4422     0.8345  23.297  < 2e-16 ***
Gender_releveledfemale   -11.5175     0.7464 -15.430  < 2e-16 ***
Age_releveled25-34 years  -4.2326     1.0556  -4.010 6.74e-05 ***
Age_releveled35-54 years  -2.0150     1.0556  -1.909   0.0567 .
Age_releveled5-24 years   -9.6238     1.0556  -9.117  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 9.846 on 691 degrees of freedom
Multiple R-squared:  0.3237,    Adjusted R-squared:  0.3198
F-statistic: 82.68 on 4 and 691 DF,  p-value: < 2.2e-16
```