Jack Rellamas
Dr. Gao
Math 167R
May 13, 2024

Factors of Happiness

Introduction:

The main, overarching, question that this report is tackling is "What are the largest influences on happiness in a population?" If we can figure out what factors most positively influence happiness, at least on a country wide level, then having that knowledge can inform the decisions of state policy makers or wealthy and influential philanthropists around the world who want to make people live happier lives. This knowledge can also inform people on an individual level so people can prioritize what we know makes the average person happier.

The World Happiness Report is an annual report that started in 2012 to track happiness around the world and to bring attention to happiness as an important indicator of the success of a country beyond the existing economic and health indicators. In the 2022 dataset, there are six factors that were tracked alongside the happiness index for each country. Happiness was scored on the Cantril Ladder. Participants were asked how happy they feel on a scale of 1 to 10, and how happy they think they will be in 5 years. The factors being GDP per capita, healthy life expectancy, social support (scored by friends or family members one could reach out to during hard times), freedom to produce life decisions, generosity (scored by how many people gave to charity in the last month), and perceptions of corruption (of government and business).

One problem or bias that might arise with the World Happiness Report dataset is that the definition of happiness differs greatly from person to person, let alone from culture to culture. The lead editors of this project are primarily economists that come from prestigious universities in the United States, Canada, and the United Kingdom including Colombia, the London School of Economics, and the University of British Columbia. It is safe to say that this report will have happiness defined through a western perspective. Happiness in the west is defined, in a very general sense, through personal freedoms, self esteem, and generally more "peppy" emotions like excitement. In contrast, people from eastern countries are more likely to define happiness as having peace, serenity, and place less importance on the individual and more importance on the good of the society and their community. There are an infinite amount of ways that happiness can be defined, but these are two general divergent ways that people can define happiness.

The other dataset that was used during this report was the dataset for the 2022 Meta-Gallup Global State of Social Connections Study. This Meta and Gallup collaboration seeked to measure how connected and how lonely people felt on a country by country basis. How connected, or not connected, people feel to others in their community has effects on many different aspects of a person's life, including one's own mental and physical health. Both connectedness and loneliness were rated on a scale of very or fairly, a little, to not at all connected and lonely.

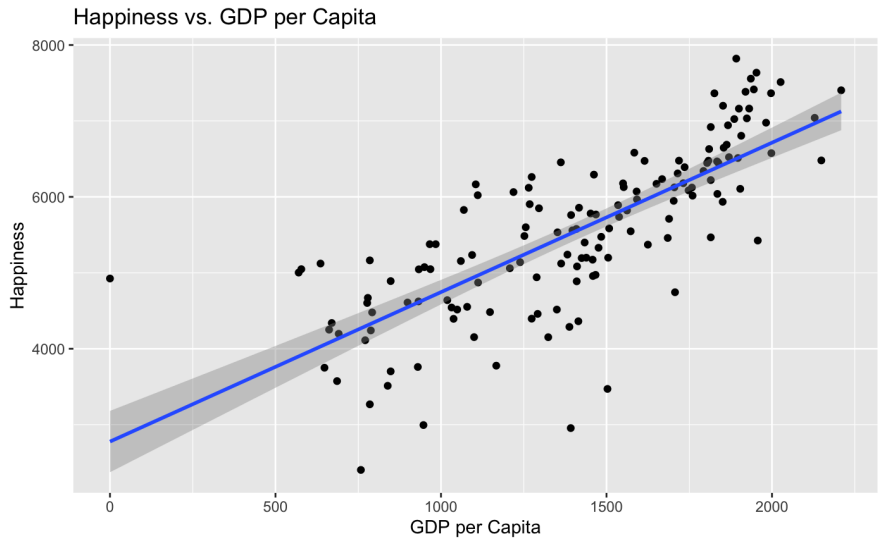
With all of that in mind, this report sought to find out the answer to three questions with all of this data. First, what are the 3 strongest predictors of happiness out of the six explanatory variables from the World Happiness Report dataset? The top three factors that we suspect will have the most impact on the happiness score are GDP per capita, healthy life expectancy, and generosity. Freedom to make life decisions and corruption seem like less important factors since there are many more restrictive societies, at least from a western perspective, like those in East Asia that are economically developed, and their citizens are still generally happy since many of the restrictions are put in place with the intent of helping the society as a whole. Corruption of government and of business does not seem like that big of an influence on the day to day life of the average citizen. Only when corruption affects the material and financial reality of the people, such as a corrupt government causing runaway inflation will it then affect the happiness of the people in a country.

Next, does connectedness have an effect on the happiness of a country? Does loneliness have a negative effect on the (un)happiness of a country? We suspected that higher senses of connectedness will have a higher positive correlation on the happiness score, and the higher senses of loneliness with a greater absolute correlation with unhappiness. That is, loneliness and happiness will be inversely correlated.

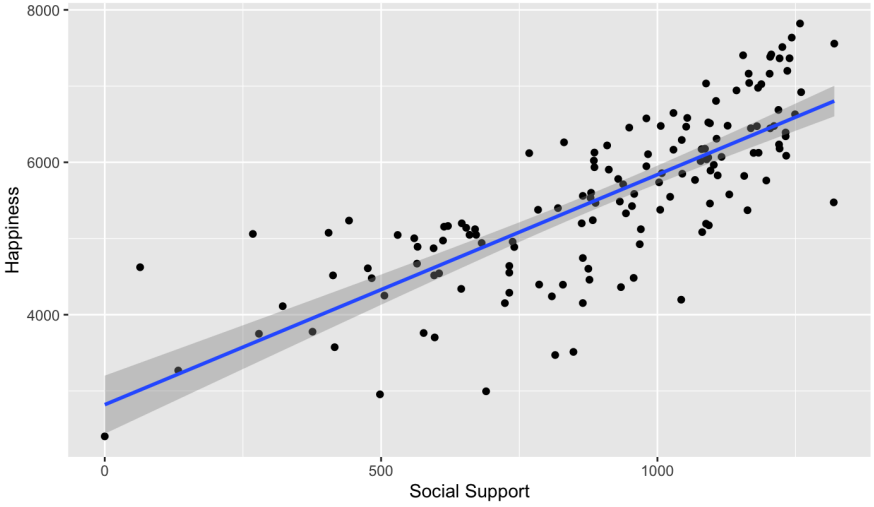
Finally, to what degree does loneliness affect the financial outcomes and the health outcomes of a country? That is to say, how does loneliness affect GDP per capita and healthy life expectancy? It is natural to suspect that loneliness will be inversely correlated with the two aforementioned factors.

Exploratory data analysis:

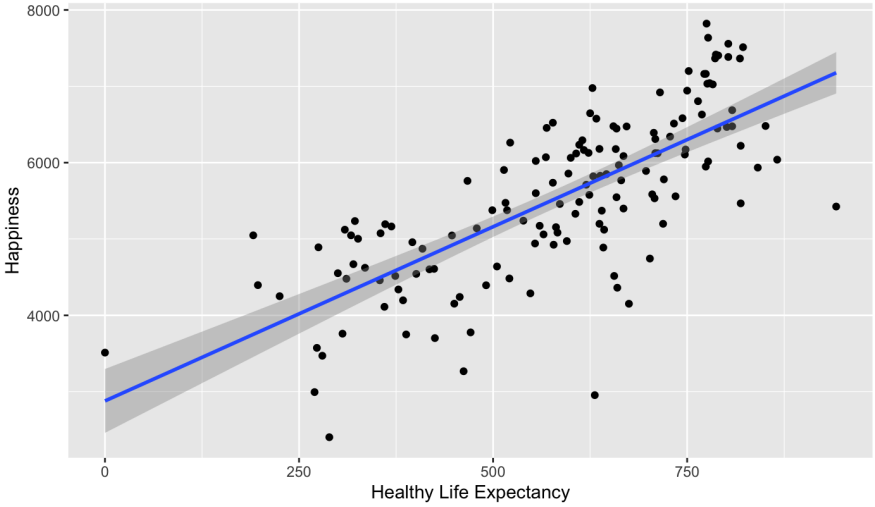
Subquestion #1:



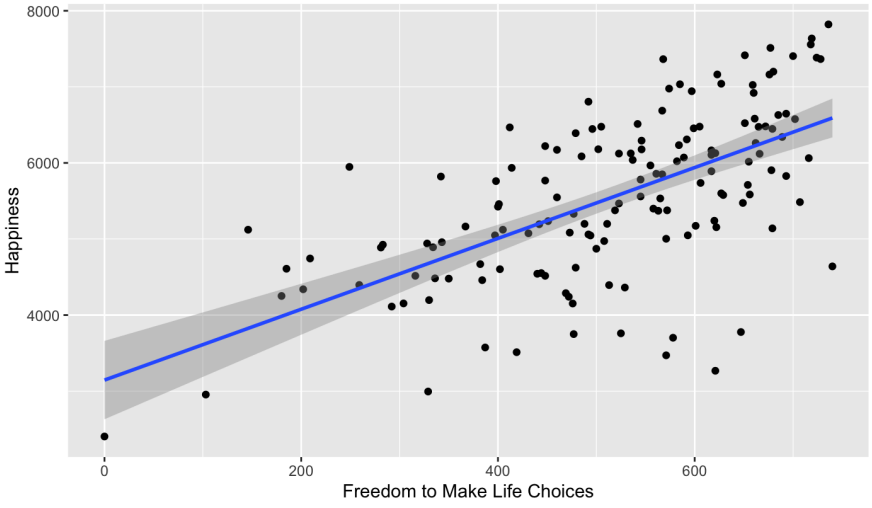
Happiness vs. Social Support

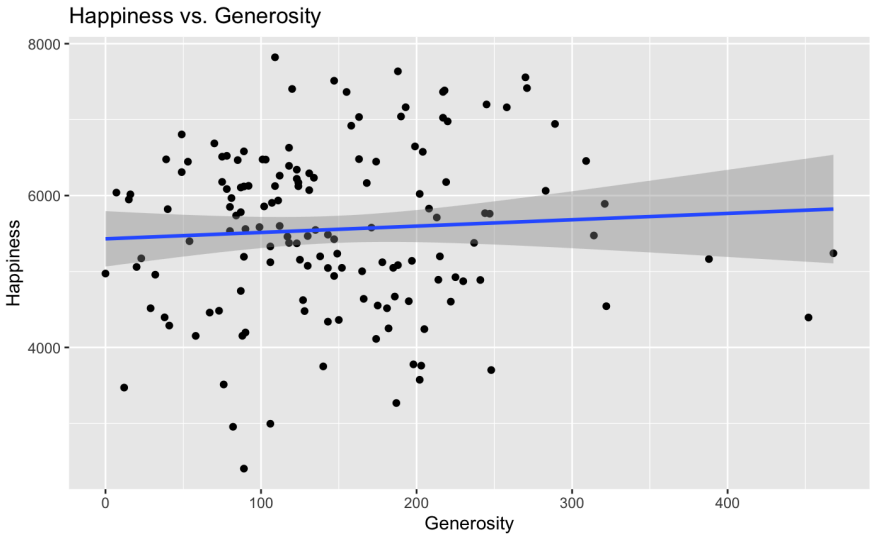


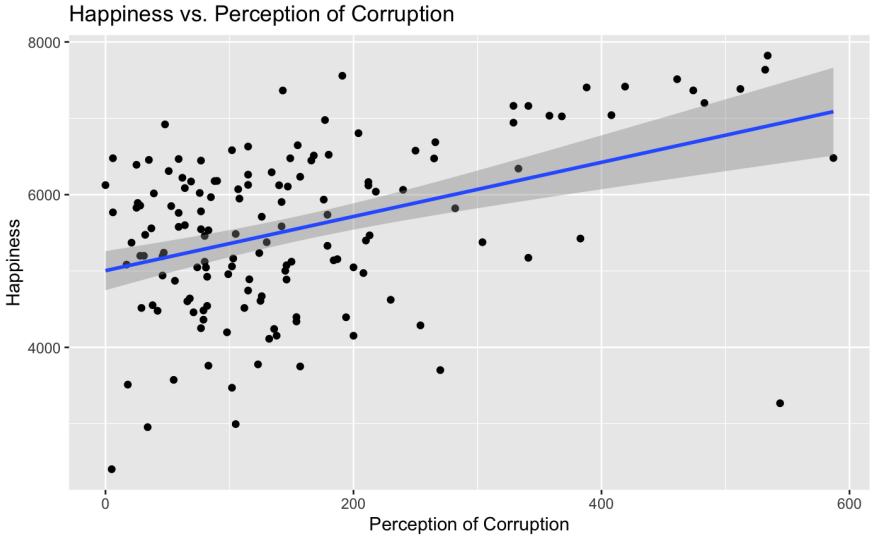
Happiness vs. Healthy Life Expectancy



Happiness vs. Freedom to Make Life Choices







	p-value	β_1	Adjusted R-squared
GDP per capita	< 2.2e-16	0.29628	0.5803
Social Support	< 2.2e-16	0.2005	0.6024
Healthy Life Expectancy	< 2.2e-16	0.12011	0.5448
Freedom to Make Life Choices	< 2.2e-16	0.083854	0.3862

Generosity	0.4444	4.859e-03	-0.002848
Perception of Corruption	1.74e-07	0.04883	0.1675

For these plots, we let the null hypothesis, H_0, be that each of the factors have no effect on the happiness score of a country. The alternative hypothesis being that there is a correlation between the independent variable and happiness.

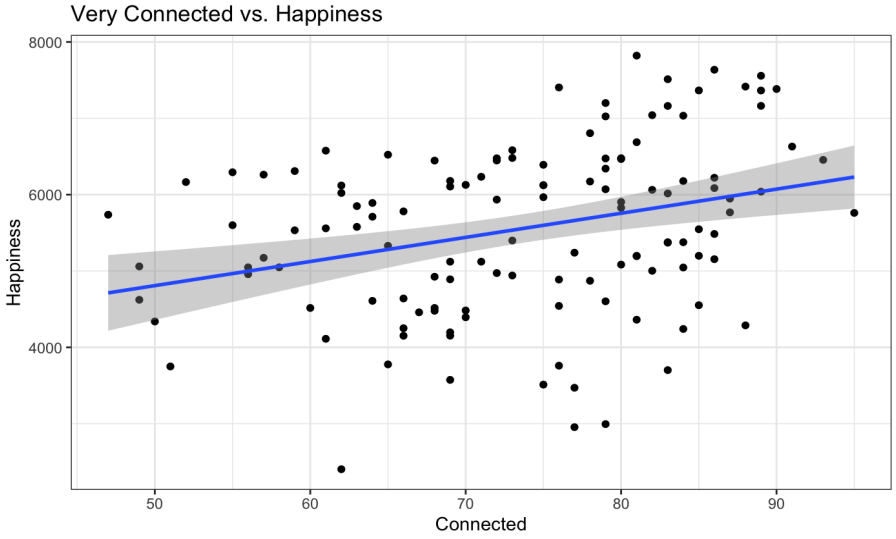
Through plotting happiness against the six factors of happiness from the World Happiness Report dataset, it can be seen that the three most influential factors of happiness are GDP per capita, social support and healthy life expectancy, in that order. Their adjusted R-squared values are greater than 0.5, which means that the relationship between the independent variables and happiness explains more than 50 percent of the variation in the data. Additionally, they have the highest β_1 values out of the 6 explanatory variables.

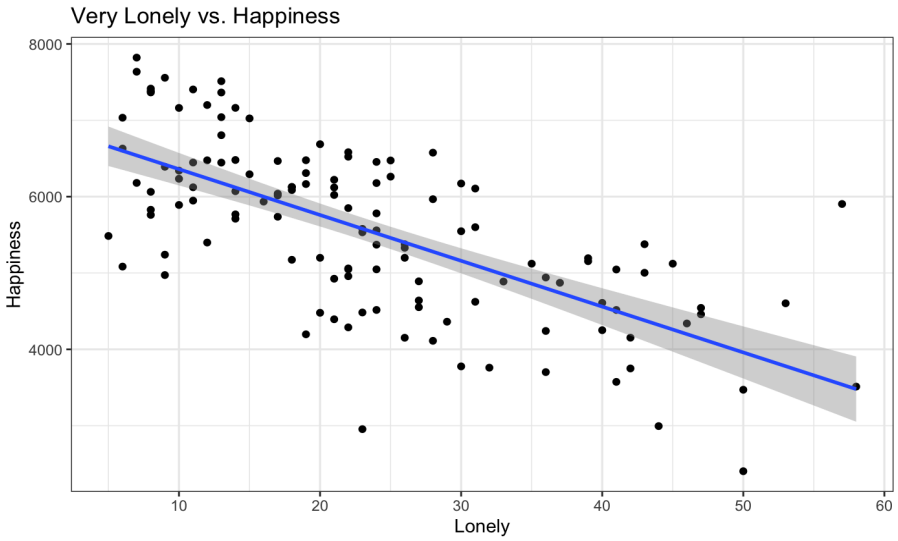
All three of these factors have a clear positively correlated linear relationship between happiness and these three variables. These results make intuitive sense. Having more money means more access to a greater quantity of life's wants and needs like food, technology, or entertainment. It also means greater access to higher quality things like healthcare and food. Social support should correlate with happiness too. People are social creatures, and our high level of collaboration is a large part of what made humans become the apex predator of the animal kingdom. Additionally, having a better social support structure and knowing there is someone to lean on during tough times takes some stress and anxiety off of the individual, even if it is only psychological. As mentioned before, being more connected to a community has positive effects on a person's mental and physical health, which would make sense that healthier people are more happy. Healthy life expectancy being positively correlated with happiness comes as no surprise either. Anyone who's ever been sick or injured, or had to deal with an underlying health condition knows how stressful it can be. A healthy life allows for more focus and attention on other things besides one's own health. Freedom to make life choices was also a statistically significant predictor of greater levels of happiness, but it just was not as strong as the previous three factors. This too makes sense, since having more agency gives people a greater sense of control over their life. A person is much happier when they quit a job versus when they are fired from their job, even though the outcome for both is the person not having the job anymore.

The remaining two of the six indicators were not correlated with happiness. Them being generosity and perception of corruption. One explanation for generosity is that people that give to charity are not necessarily giving to local charities. They may not be able to see the impact of their donation in their immediate surroundings, thus, they do not get the reward of feeling good about their altruism. The perception of corruption is perplexing. As seen in its plot, there are plenty of highly corrupt, yet still happy countries. In fact, there are more highly corrupt countries

that are happy than corrupt and unhappy countries. Corruption may not lead to significant impacts on a citizen's everyday life.

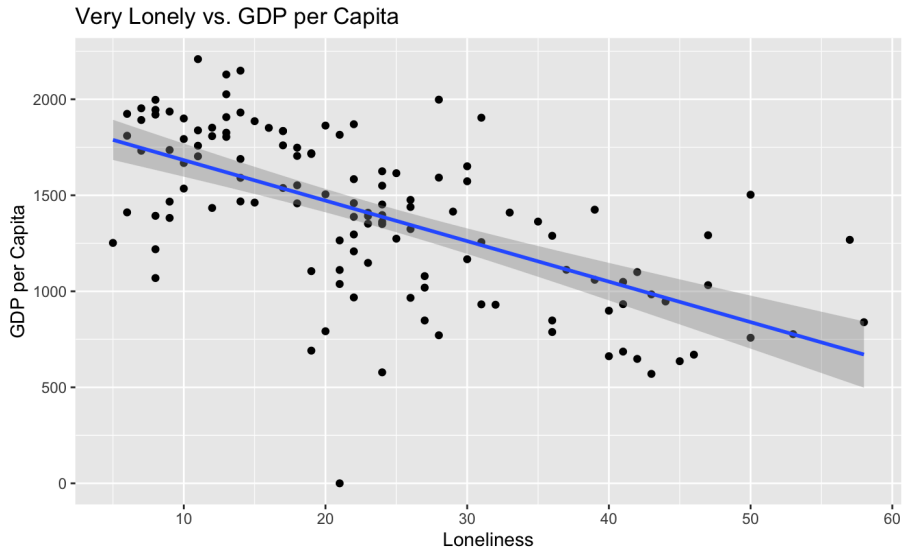
Subquestion #2:





It can be seen that there is a positive correlation with connectedness and happiness and there is a negative relationship between loneliness and happiness. An interesting fact that can be seen from these plots is that people of a country can feel fairly connected to others, yet still be unhappy. However, a country could not be very happy if it was very lonely. In the happiest countries, there is a very small proportion of the population that was very lonely. But there is much more variance of above average "connected" countries in terms of their happiness. People can be around others, and even feel connected to a community or to their family, but that there can still be conflict and stress from those relationships. However, if you are lonely, there are only negative factors that come with loneliness.

Subquestion #3:



We can confirm from the data that loneliness, as expected, has a negative correlation with both GDP per capita and the healthy life expectancy for people. People who are lonely will be, not only less happy, but also less productive citizens in terms of their economic output. It can be argued that if you want your country to succeed, having a better sense of community is an essential part of being more competitive in the global economy. Policy makers can put community building as an economic issue. For individuals, they should push themselves to become more engaged in their community if they want to make more money, and if they want to live longer, healthier lives.

Conclusion and Future Questions:

There is clear evidence that happiness and a combination of having a good economic situation and feeling like you are connected to a community, whether it be with friends or family, are both important factors of a happier society. What that means for the state policy makers and

the philanthropists interested in building a greater society is that they should be focusing on programs and policies that raise the level of economic prosperity for all citizens of a country. There are many opposing views on the best way to accomplish this goal, and we will leave that to the respective governments to decide their own best course of action.

In regards to building better communities, there are an equally vast amount of ways people can help facilitate greater connection within communities. Building spaces like community centers, parks, libraries, etc. are ways to provide a way in which people get to meet others. Things like after school sports programs and adult sports leagues are another more natural way people can get to meet others while doing something enjoyable. Another way governments can also build closer communities is to work on reducing crime. People will feel safer and will be more willing to engage with others of their community if trust is high between citizens.

Another question that could be asked still is what other factors of a high corruption society, yet still a happy country, lead to happiness.

There are still further questions for people wanting to build a community. Firstly, other interested parties would need to see if the research confirms that countries with lower crime rates have higher rates of connection. Then they would have to determine what crime is most prevalent and how to deal with it. We know that fair and effective policing of crime can be a difficult undertaking, as even the United States, one of the most economically prosperous countries in the history of the world, still has struggles in this area of policy.

Code Appendix

```{r}
library(tidyverse)
happiness_data <- read.csv("/Users/rellamas/math_and_algos/R/final/2022.csv")
```
```{r}
# remove empty "xx" country remove dataframe
happiness_data <- happiness_data[-c(147), ]
```

```{r}
install.packages(c("cowplot", "googleway", "ggplot2", "ggrepel",
"ggspatial", "libwgeom", "sf", "rnaturalearth", "rnaturalearthdata"))
```
```{r}
library("ggplot2")
theme_set(theme_bw())
library("sf")
library("rnaturalearth")
library("rnaturalearthdata")

world <- ne_countries(scale = "medium", returnclass = "sf")
class(world)
```

```{r}
# cleaning

# remove 3 unused columns from happiness df
drops <- c("Whisker.high", "Whisker.low", "Dystopia..1.83....residual")
happiness_data <- happiness_data[ , !(names(happiness_data) %in% drops)]

# remove all instances of '*' from happiness df 'happiness score' column
happiness_data$Country <- as.character(gsub("\\*","", happiness_data$Country))  #
regex
```

```
# convert chr to doubles
# 1
happiness_data$Happiness.score <- as.numeric(gsub(",","",
happiness_data$Happiness.score))
# 2
happiness_data$Explained.by..GDP.per.capita <- as.numeric(gsub(",","",
happiness_data$Explained.by..GDP.per.capita))
# 3
happiness_data$Explained.by..Social.support <- as.numeric(gsub(",","",
happiness_data$Explained.by..Social.support))
# 4
happiness_data$Explained.by..Healthy.life.expectancy <- as.numeric(gsub(",","",
happiness_data$Explained.by..Healthy.life.expectancy))
# 5
happiness_data$Explained.by..Freedom.to.make.life.choices <- as.numeric(gsub(",","",
happiness_data$Explained.by..Freedom.to.make.life.choices))
# 6
happiness_data$Explained.by..Generosity <- as.numeric(gsub(",","",
happiness_data$Explained.by..Generosity))
# 7
happiness_data$Explained.by..Perceptions.of.corruption <- as.numeric(gsub(",","",
happiness_data$Explained.by..Perceptions.of.corruption))
```

```{r}
# gdp per capita
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..GDP.per.capita, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("GDP per Capita") +
  ylab("Happiness") +
  ggtitle("Happiness vs. GDP per Capita")

gdp_res <- lm(Explained.by..GDP.per.capita ~ Happiness.score, data = happiness_data)
summary(gdp_res)
# p-value: < 2.2e-16
```

```
# beta 1: 0.29628
# Adjusted R-squared:  0.5803
```

```{r}
# social support
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..Social.support, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Social Support") +
  ylab("Happiness") +
  ggtitle("Happiness vs. Social Support")

social_support_res <- lm(Explained.by..Social.support ~ Happiness.score, data =
happiness_data)
summary(social_support_res)
# p-value: < 2.2e-16
# beta 1: 0.2005
# Adjusted R-squared:  0.6024
```

```{r}
# healthy life expectancy
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..Healthy.life.expectancy, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm)  +
  xlab("Healthy Life Expectancy") +
  ylab("Happiness") +
  ggtitle("Happiness vs. Healthy Life Expectancy")

life_expectancy_res <- lm(Explained.by..Healthy.life.expectancy ~ Happiness.score, data
= happiness_data)
summary(life_expectancy_res)
# p-value: < 2e-16
# beta 1: 0.12011
# Adjusted R-squared:  0.5448
```

```

```{r}
# freedom to make life choices
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..Freedom.to.make.life.choices, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Freedom to Make Life Choices") +
  ylab("Happiness") +
  ggtitle("Happiness vs. Freedom to Make Life Choices")

freedom_res <- lm(Explained.by..Freedom.to.make.life.choices ~ Happiness.score, data = happiness_data)
summary(freedom_res)
# p-value: < 2e-16
# beta 1: 0.083854
# Adjusted R-squared:  0.3862
```

```{r}
# generosity
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..Generosity, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Generosity") +
  ylab("Happiness") +
  ggtitle("Happiness vs. Generosity")

generosity_res <- lm(Explained.by..Generosity ~ Happiness.score, data = happiness_data)
summary(generosity_res)
# p-value: 0.4444
# beta 1: 4.859e-03
# Adjusted R-squared:  -0.002848
```

```{r}

```
# perception of corruption
attach(happiness_data)
ggplot(data = happiness_data,
       aes(Explained.by..Perceptions.of.corruption, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Perception of Corruption") +
  ylab("Happiness") +
  ggtitle("Happiness vs. Perception of Corruption")

corruption_res <- lm(Explained.by..Perceptions.of.corruption ~ Happiness.score, data =
happiness_data)
summary(corruption_res)
# p-value: 1.74e-07
# beta 1: 0.04883
# Adjusted R-squared: 0.1675
# qqnorm(resid(corruption_res))
# qqline(resid(corruption_res))
```

```{r}
# low corruption, low happiness
#TODO
# maybe do some analysis on these countries
# like see what other factors of happiness they are lacking in
lo_corruption_lo_happy <- happiness_data %>%
  filter(Happiness.score < median(Happiness.score)) %>%
  filter(Explained.by..Perceptions.of.corruption < 200)
print(median(happiness_data$Happiness.score))
print(median(happiness_data$Explained.by..Perceptions.of.corruption))

ggplot(data = lo_corruption_lo_happy,
       aes(Explained.by..Perceptions.of.corruption, Happiness.score)) +
  geom_point()
```

```{r}
# combining happiness, connectedness, and loneliness into one dataframe
#       will have to remove 4 countries
```

```
df2 <- merge(connected, lonely, by="Country")
df <- merge(happiness_data, df2)
# remove all instances of '*' from happiness df 'happiness score' column
df$Very.Fairly.x <- as.numeric(gsub("\\%","", df$Very.Fairly.x))  # regex
df$Very.Fairly.y <- as.numeric(gsub("\\%","", df$Very.Fairly.y))  # regex

ggplot(data = df,
       aes(Very.Fairly.x, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Connected") +
  ylab("Happiness") +
  ggtitle("Very Connected vs. Happiness")

ggplot(data = df,
       aes(Very.Fairly.y, Happiness.score)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Lonely") +
  ylab("Happiness") +
  ggtitle("Very Lonely vs. Happiness")
```

```{r}
# 3
ggplot(data = df,
       aes(Very.Fairly.y, Explained.by..GDP.per.capita)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Loneliness") +
  ylab("GDP per Capita") +
  ggtitle("Very Lonely vs. GDP per Capita")

ggplot(data = df,
       aes(Very.Fairly.y, Explained.by..Healthy.life.expectancy)) +
  geom_point() +
  geom_smooth(method = lm) +
  xlab("Loneliness") +
  ylab("Healthy Life Expectancy") +
  ggtitle("Very Lonely vs. Healthy Life Expectancy")
```

```
```