Project1 DAT 301

2022-12-4

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF,

```
library("reshape2")
library(ggplot2)
library(tidyr)
```

```
##
## Attaching package: 'tidyr'
```

```
## The following object is masked from 'package:reshape2':
##
##     smiths
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(readr)
library(caret)
```

```
## Loading required package: lattice
```

Question: I will be analyzing the data that was collected from Gallup to calculate the world happiness score. I will be coding a linear regression model in order to determine which measured factor has the greatest impact on the world happiness score.

Overview of the data: The data I am using is from Kaggle which is a csv file that combines 5 years of cleaned data from Gallup's world happiness center's research, of which the subscores total to the happiness score.They host "the most comprehensive and farthest-reaching survey of the world." it reaches over 99% of the worlds population through the collaboration with independent researchers and other private orgs. Gallup polls are widely used in policy-making decisions and look at how the human population is experiencing their quality of life, as well as other public opinion polls. The file is the cleaned data as well as several categorical variables. The original data was collected from transparency.org. The data is already very clean and so I didnt't need to wrangle it that much. I created new data frames from it though that show a few different pieces of information about the data as well as ones for each year. AS well as the one that I use which combines the

The variables that go into calculating the happiness score are family, health, freedom, generosity, government trust,and gdp per capita. there are also categorical variables for the continent, country, and the year the data was from. There's 3 other numerical variables that are in the data that we won't specifically be looking

at in this assignment, they are the dystopia residual, social support, and cpi score. I will rbe doing more exploratory data analysis on those in the python assignment.

the happiness score is what is calculated by summing the measured factors included. The happiness score is based on a 1-10 scale and each subscore calculates to a portion of that.

```
WorldHappiness_df <- read.csv("~/Downloads/WorldHappiness_Corruption_2015_2020.csv")
summary(WorldHappiness_df)
```

```
##    Country          happiness_score gdp_per_capita       family
##  Length:792         Min.   :2.567   Min.   :0.0000   Min.   :0.000
##  Class :character   1st Qu.:4.591   1st Qu.:0.6442   1st Qu.:0.000
##  Mode  :character   Median :5.486   Median :0.9945   Median :0.000
##                     Mean   :5.473   Mean   :0.9292   Mean   :0.505
##                     3rd Qu.:6.301   3rd Qu.:1.2287   3rd Qu.:1.040
##                     Max.   :7.809   Max.   :2.0960   Max.   :1.611
##      health          freedom         generosity     government_trust
##  Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.00000
##  1st Qu.:0.5101   1st Qu.:0.3254   1st Qu.:0.1258   1st Qu.:0.05286
##  Median :0.6854   Median :0.4396   Median :0.1970   Median :0.08900
##  Mean   :0.6487   Mean   :0.4270   Mean   :0.2124   Mean   :0.12572
##  3rd Qu.:0.8156   3rd Qu.:0.5463   3rd Qu.:0.2732   3rd Qu.:0.15425
##  Max.   :1.1410   Max.   :0.7240   Max.   :0.8381   Max.   :0.55191
##  dystopia_residual  continent             Year      social_support
##  Min.   :0.000     Length:792         Min.   :2015   Min.   :0.0000
##  1st Qu.:0.000     Class :character   1st Qu.:2016   1st Qu.:0.0000
##  Median :1.732     Mode  :character   Median :2018   Median :0.1762
##  Mean   :1.379                        Mean   :2018   Mean   :0.6093
##  3rd Qu.:2.237                        3rd Qu.:2019   3rd Qu.:1.2683
##  Max.   :3.602                        Max.   :2020   Max.   :1.6440
##    cpi_score
##  Min.   :11.00
##  1st Qu.:30.00
##  Median :38.00
##  Mean   :44.33
##  3rd Qu.:57.00
##  Max.   :91.00
```

```
head(WorldHappiness_df)
```

```
##       Country happiness_score gdp_per_capita   family    health   freedom
## 1      Norway           7.537       1.616463 1.533524 0.7966665 0.6354226
## 2     Denmark           7.522       1.482383 1.551122 0.7925655 0.6260067
## 3     Iceland           7.504       1.480633 1.610574 0.8335521 0.6271626
## 4 Switzerland           7.494       1.564980 1.516912 0.8581313 0.6200706
## 5     Finland           7.469       1.443572 1.540247 0.8091577 0.6179509
## 6 Netherlands           7.377       1.503945 1.428939 0.8106961 0.5853845
##   generosity government_trust dystopia_residual continent Year social_support
## 1  0.3620122        0.3159638          2.277027    Europe 2015              0
## 2  0.3552805        0.4007701          2.313707    Europe 2015              0
## 3  0.4755402        0.1535266          2.322715    Europe 2015              0
## 4  0.2905493        0.3670073          2.276716    Europe 2015              0
## 5  0.2454828        0.3826115          2.430182    Europe 2015              0
## 6  0.4704898        0.2826618          2.294804    Europe 2015              0
##   cpi_score
## 1        88
```

```
## 2         91
## 3         79
## 4         86
## 5         90
## 6         84
```

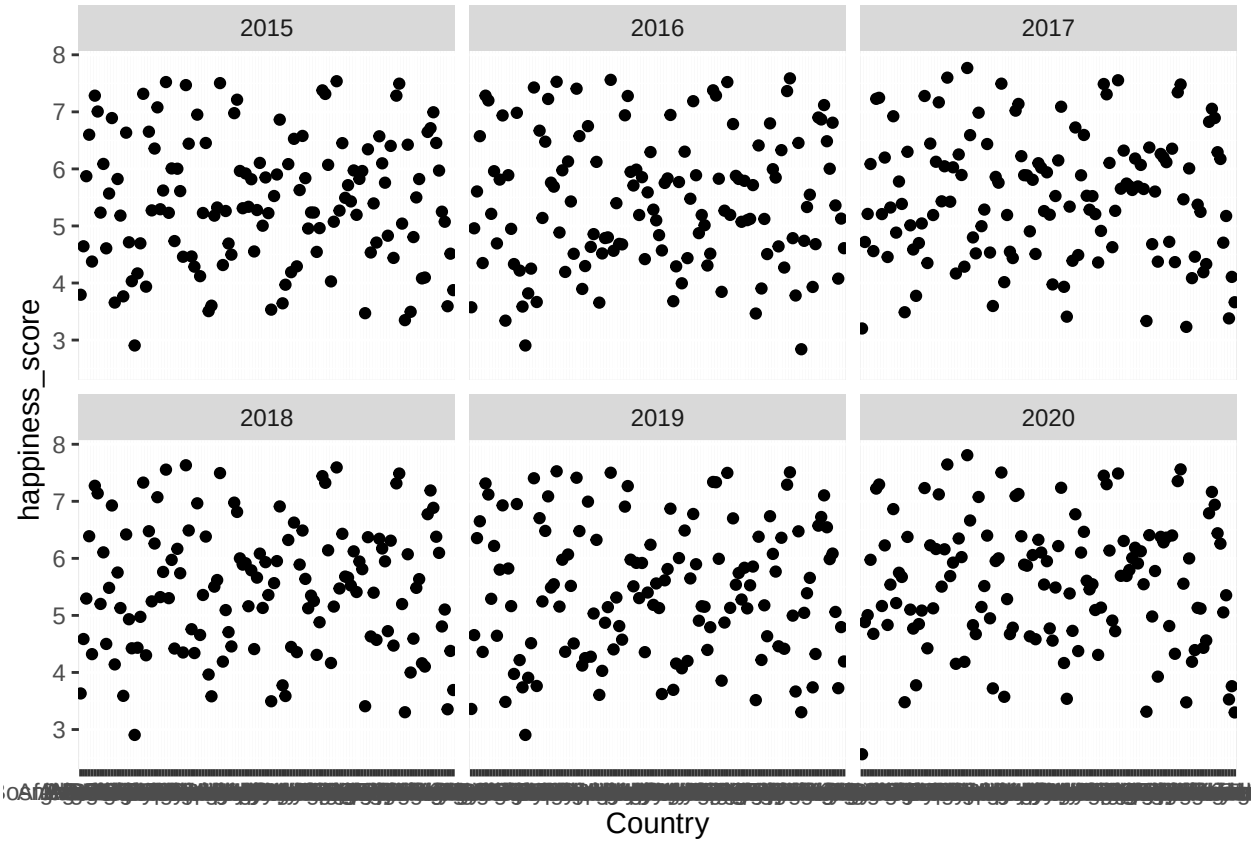
```
WorldHappiness_df <- transform(WorldHappiness_df, Country = factor(Country))
```

```
continent_mean <- WorldHappiness_df %>%
  group_by(continent) %>%
  summarize(meanHappiness = mean(happiness_score))
```

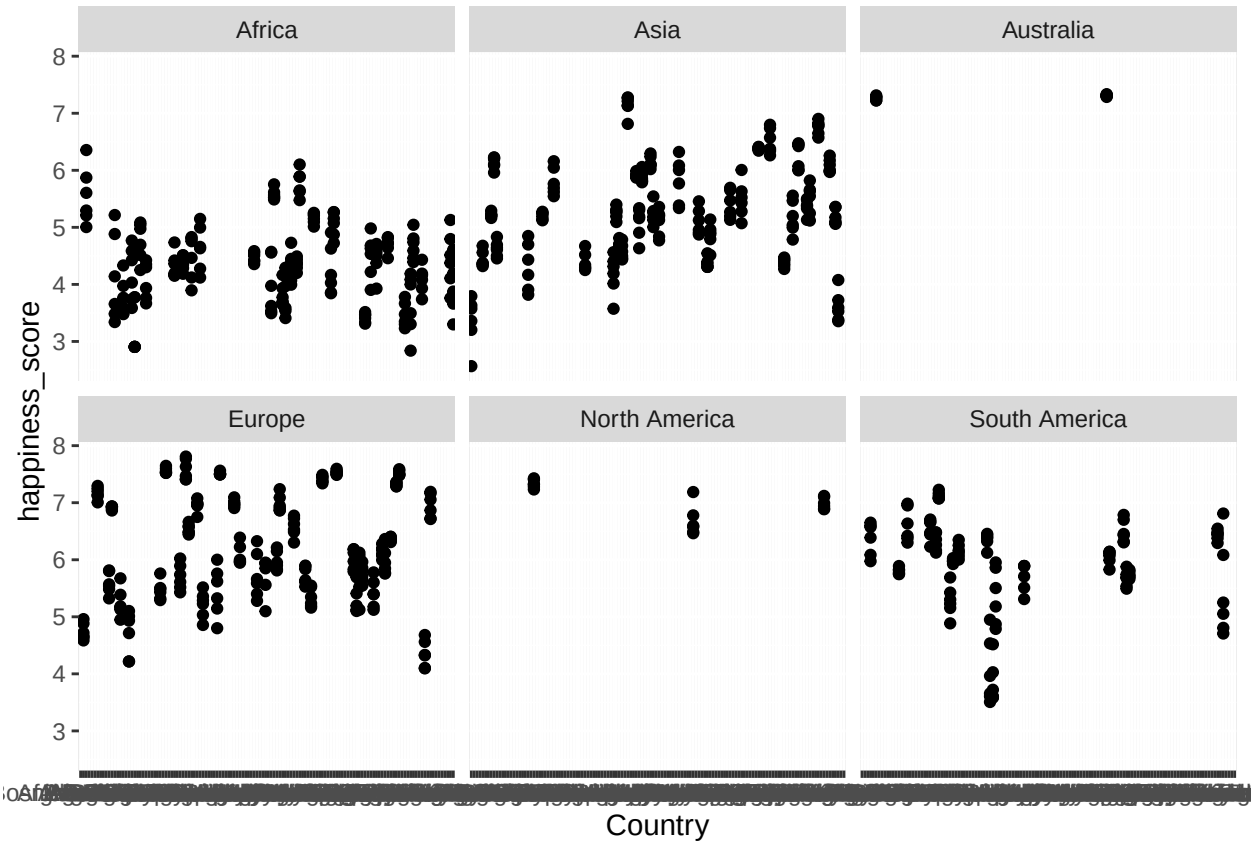
```
mean_happiness <- mean(WorldHappiness_df$happiness_score) #5.473
sd_happiness <- sd(WorldHappiness_df$happiness_score) #1.124
max_happiness <- max(WorldHappiness_df$happiness_score) #7.809
min_happiness <- min(WorldHappiness_df$happiness_score) #2.567
```

Data Visualization: This graph shows the different distribution of the happiness scales by year, I was hoping to see some sort of noticeable difference between the years but that doesn't seem to be the case. Even in 2020, pandemic year there was no noticeable difference.

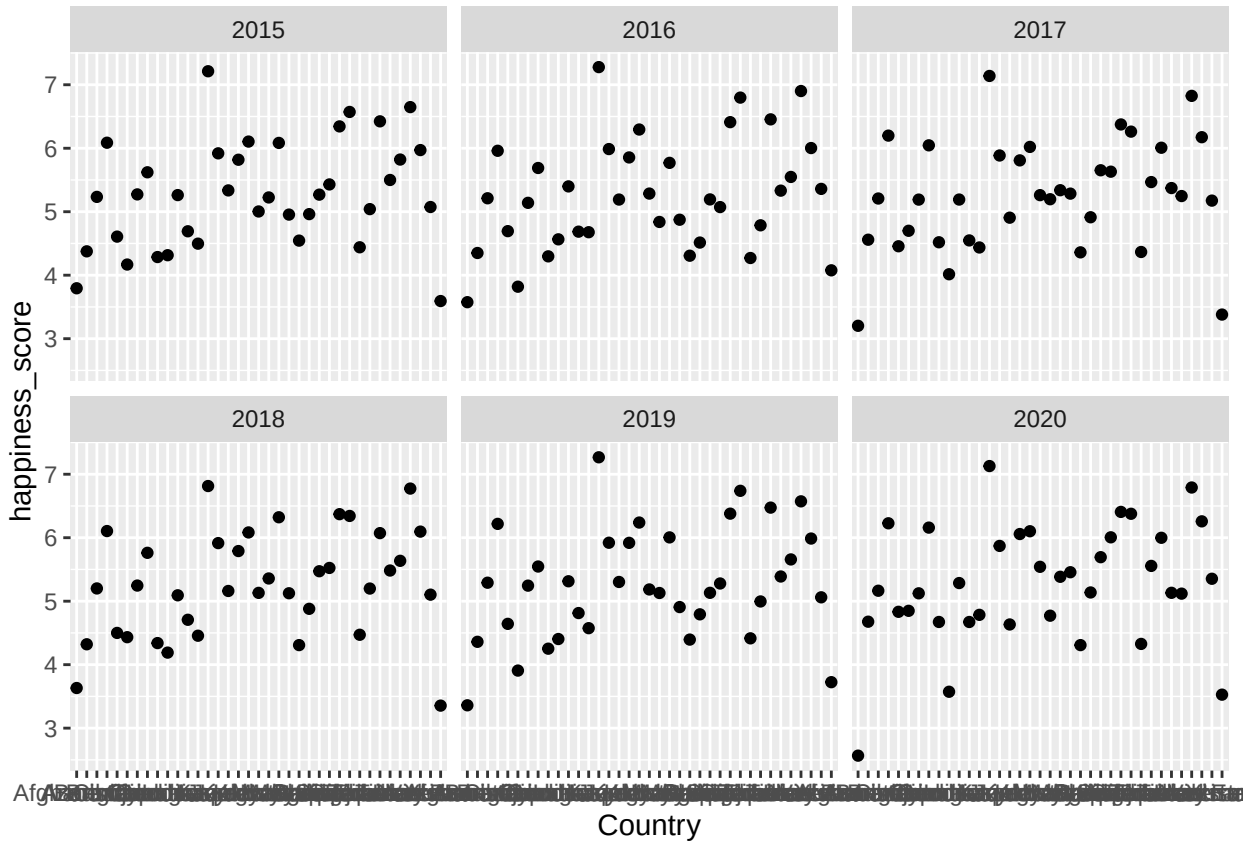
```
ggplot(WorldHappiness_df, aes(x = Country, y = happiness_score)) + geom_point() +facet_wrap(~ Year)
```



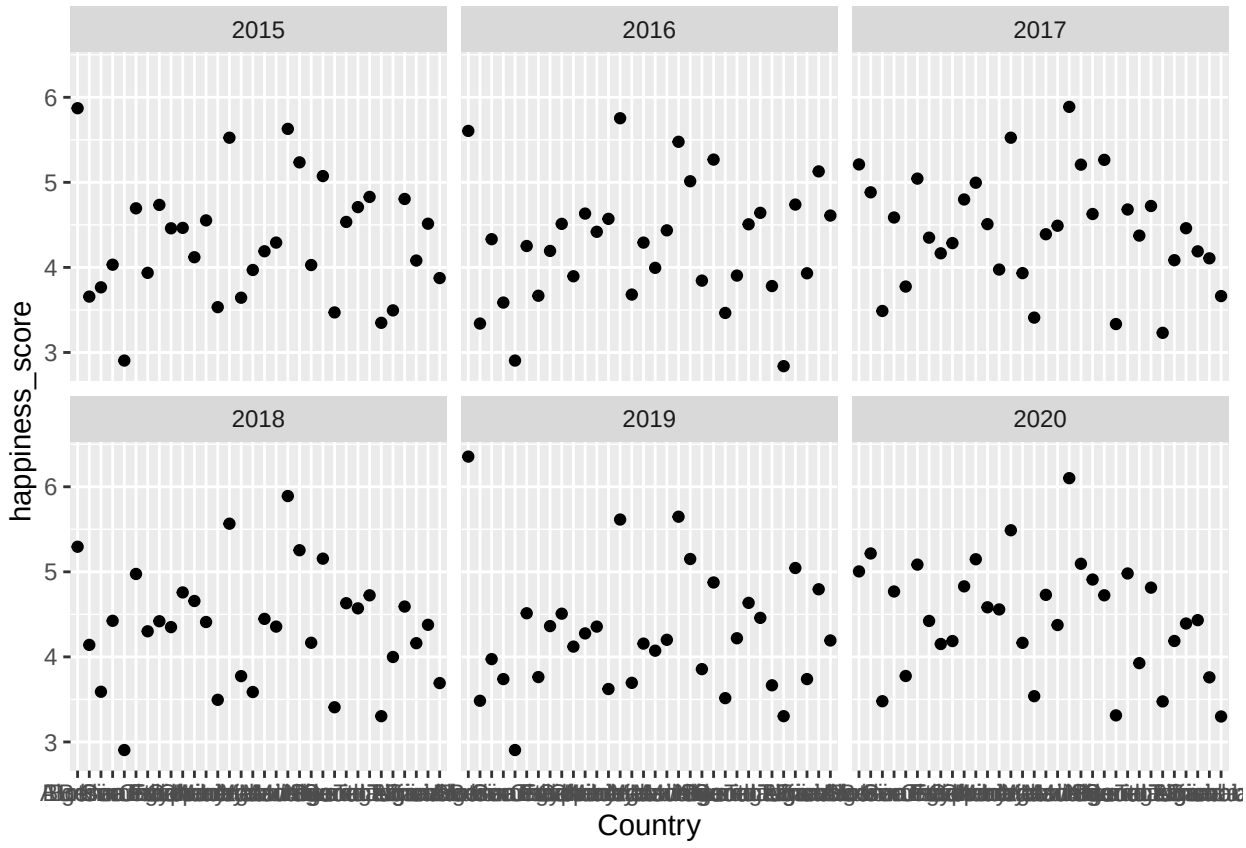

```
WorldHappiness_df %>%
  group_by(continent) %>%
  ggplot(aes(x = Country, y = happiness_score)) + geom_point() + facet_wrap(~ continent)
```



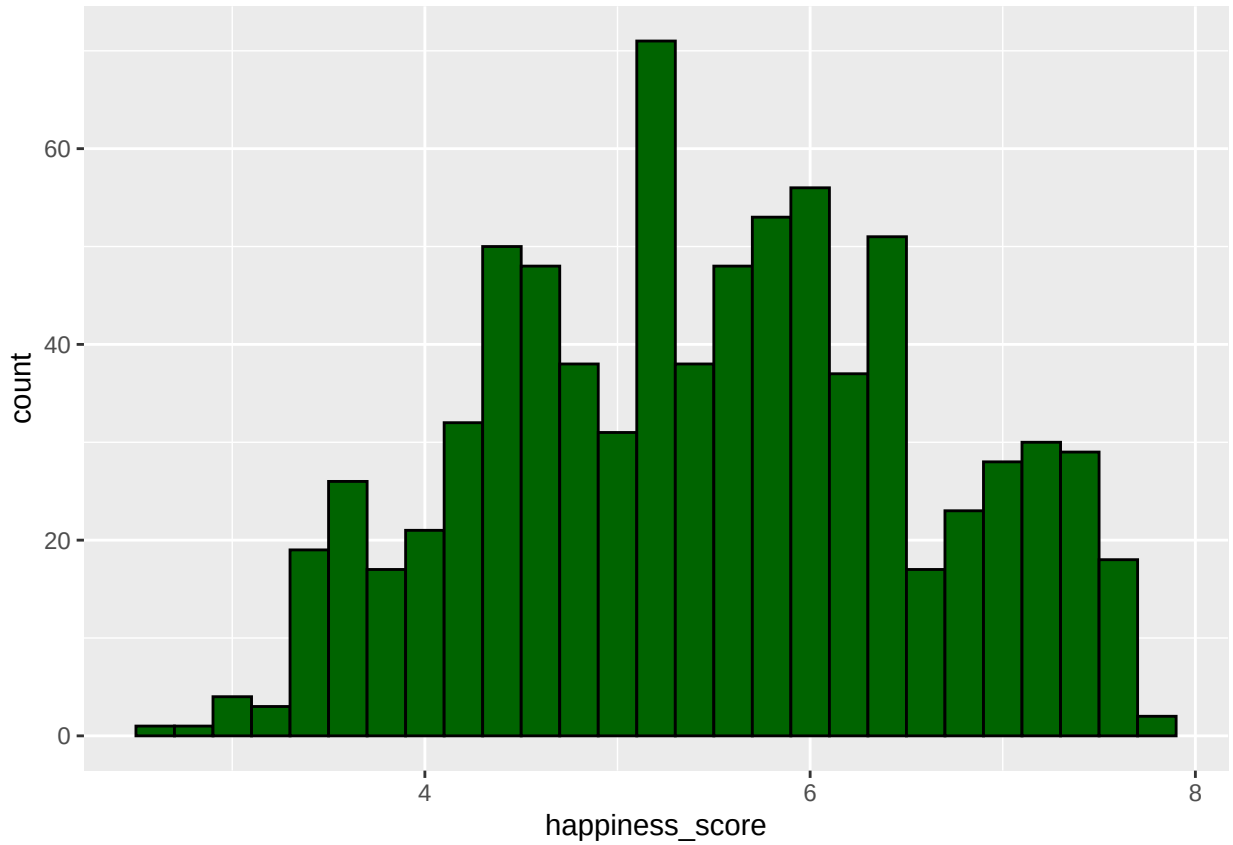

```
WorldHappiness_df %>%
  filter(continent == "Asia") %>%
  ggplot(aes(x = Country, y = happiness_score)) + geom_point() +facet_wrap(~ Year)
```



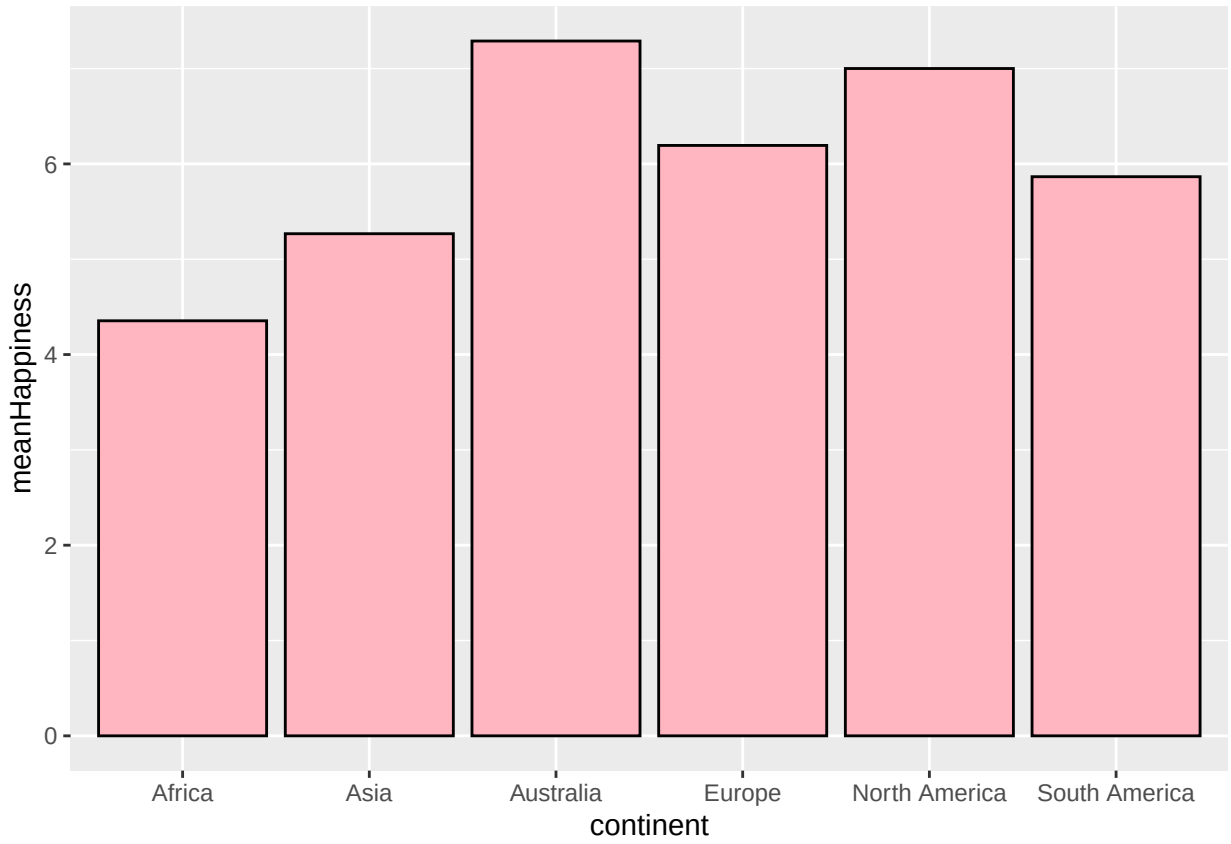

```
WorldHappiness_df %>%
  filter(continent == "Africa") %>%
  ggplot(aes(x = Country, y = happiness_score)) + geom_point() +facet_wrap(~ Year)
```



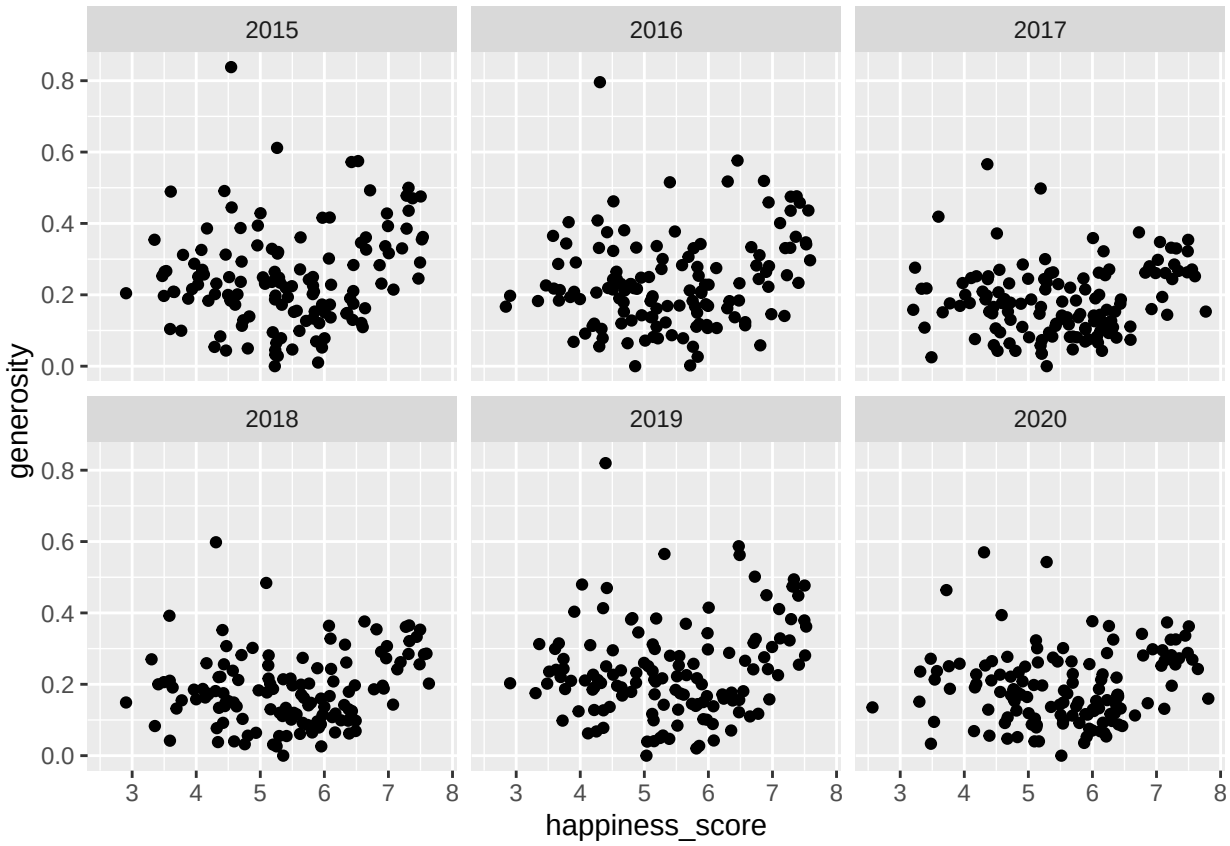

```
ggplot(WorldHappiness_df, aes(happiness_score)) +
geom_histogram(binwidth = .2, fill = 'darkgreen', color = "black")
```



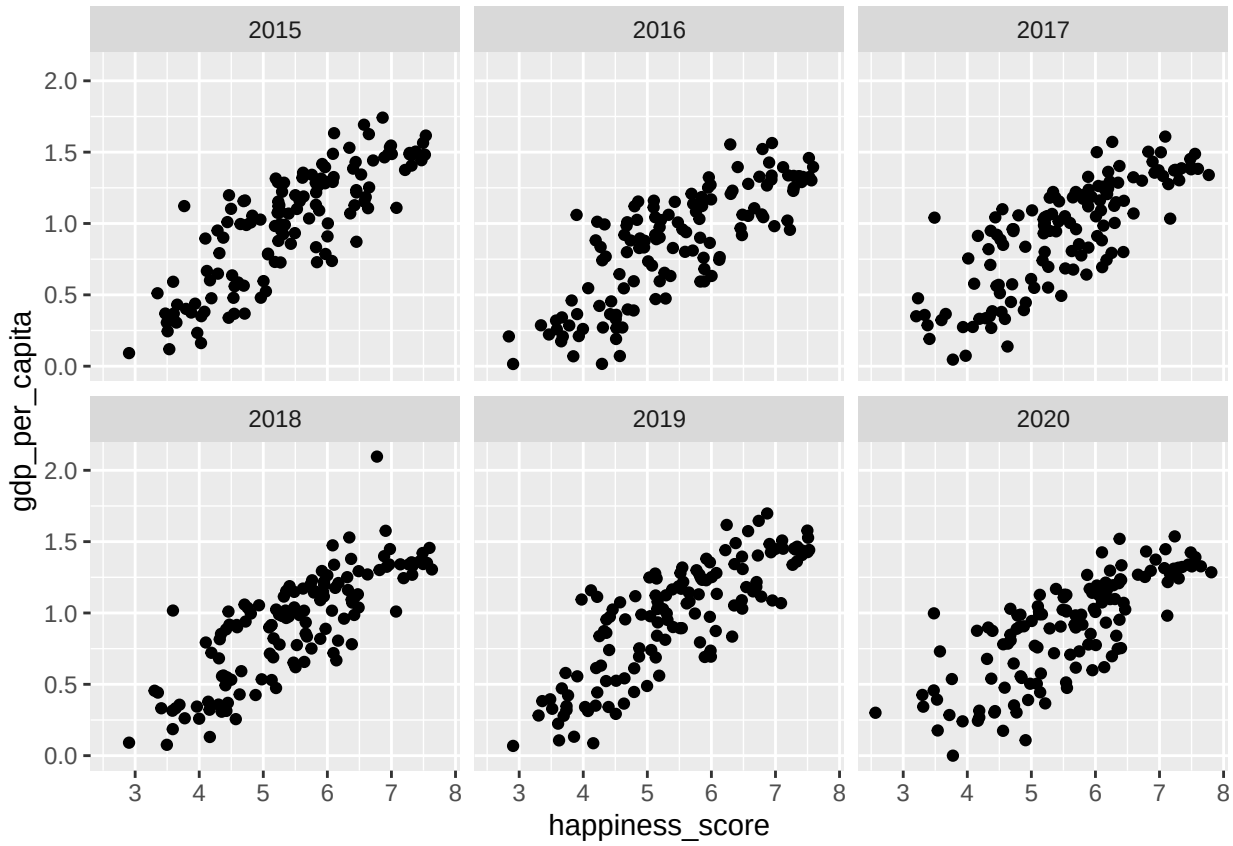

```
ggplot(continent_mean, aes(x = continent, y = meanHappiness)) + geom_bar(stat = "identity", fill = 'ligh
```



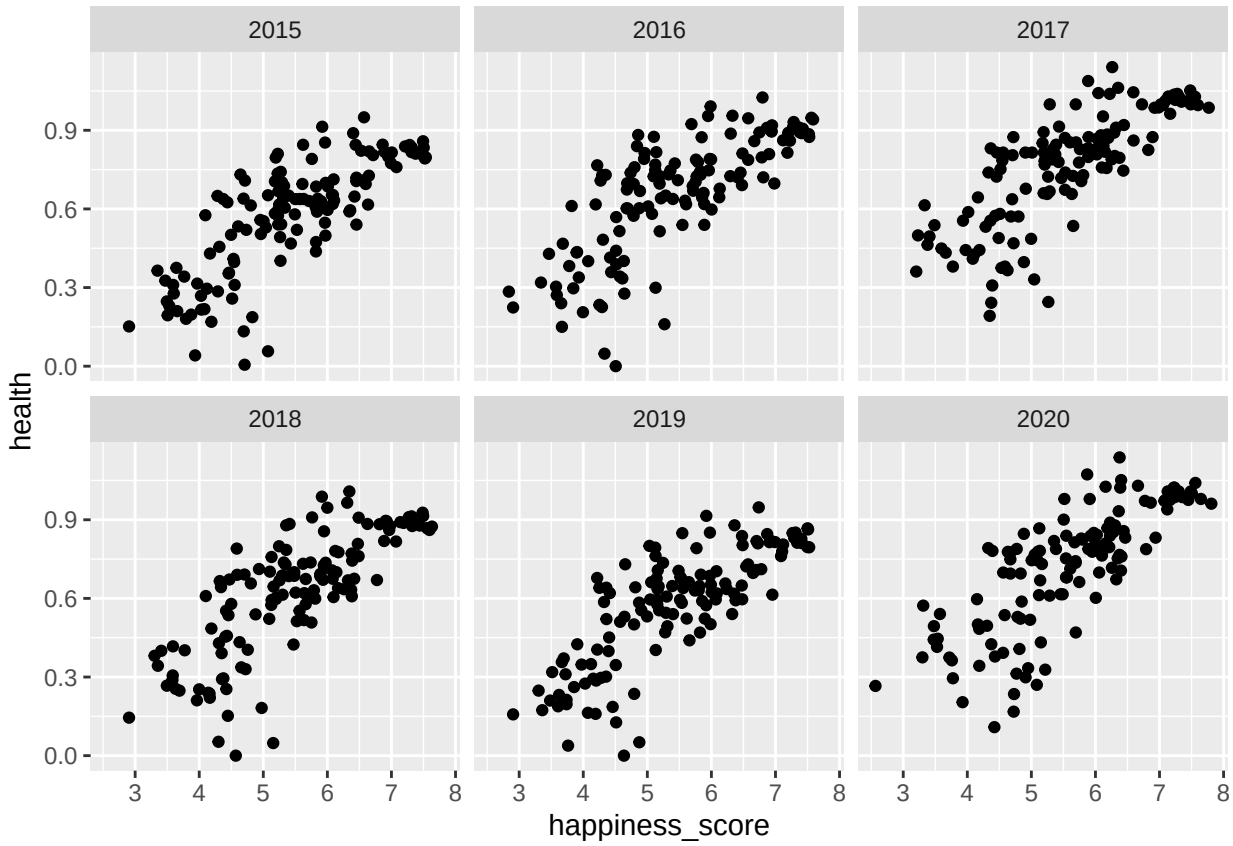

```
ggplot(WorldHappiness_df, aes(x = happiness_score, y = generosity)) + geom_point() +facet_wrap(~ Year)
```




```
ggplot(WorldHappiness_df, aes(x = happiness_score, y = gdp_per_capita)) + geom_point() + facet_wrap(~ Y
```

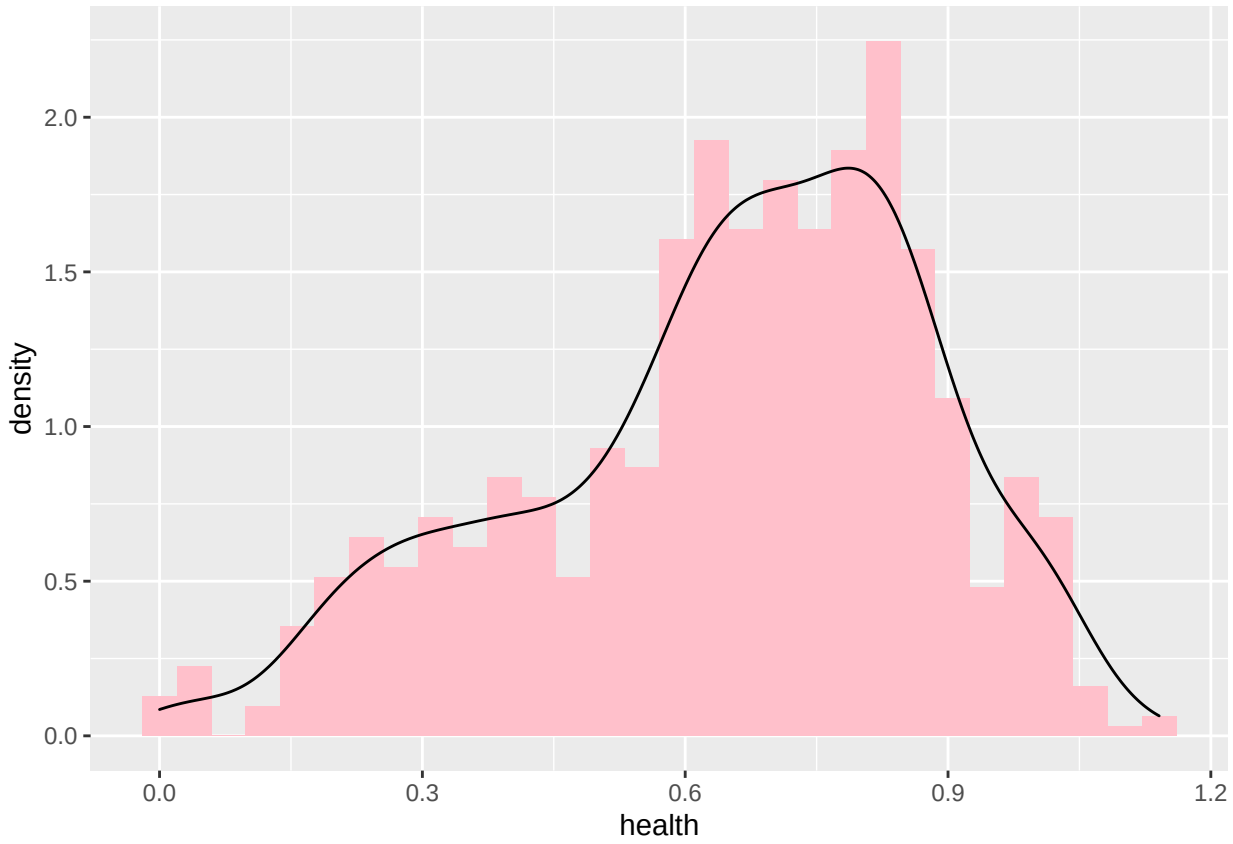



```
WorldHappiness_df %>%
  ggplot(aes(x = happiness_score, y = health )) + geom_point() +facet_wrap(~ Year)
```



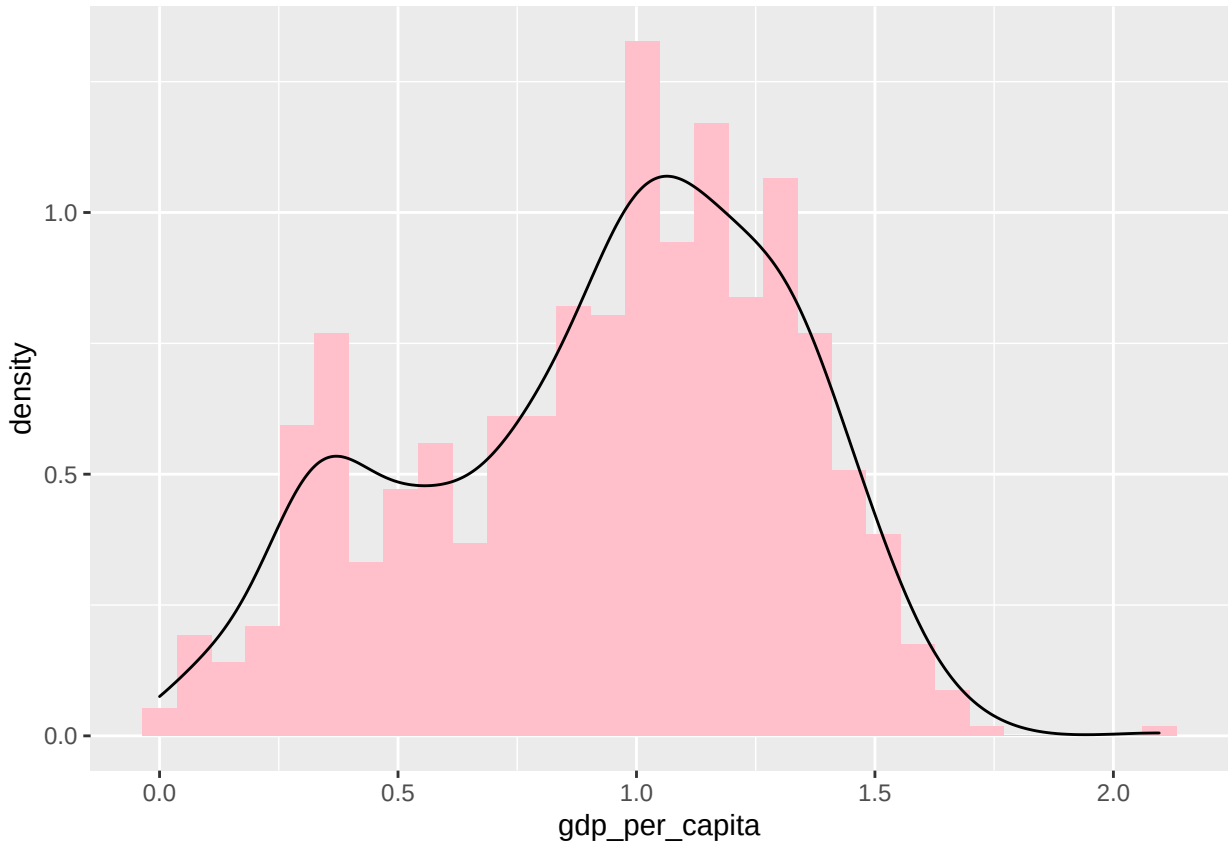

```
ggplot(WorldHappiness_df, aes(health)) +geom_histogram(aes(y = ..density..), fill = "pink") + geom_densi
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



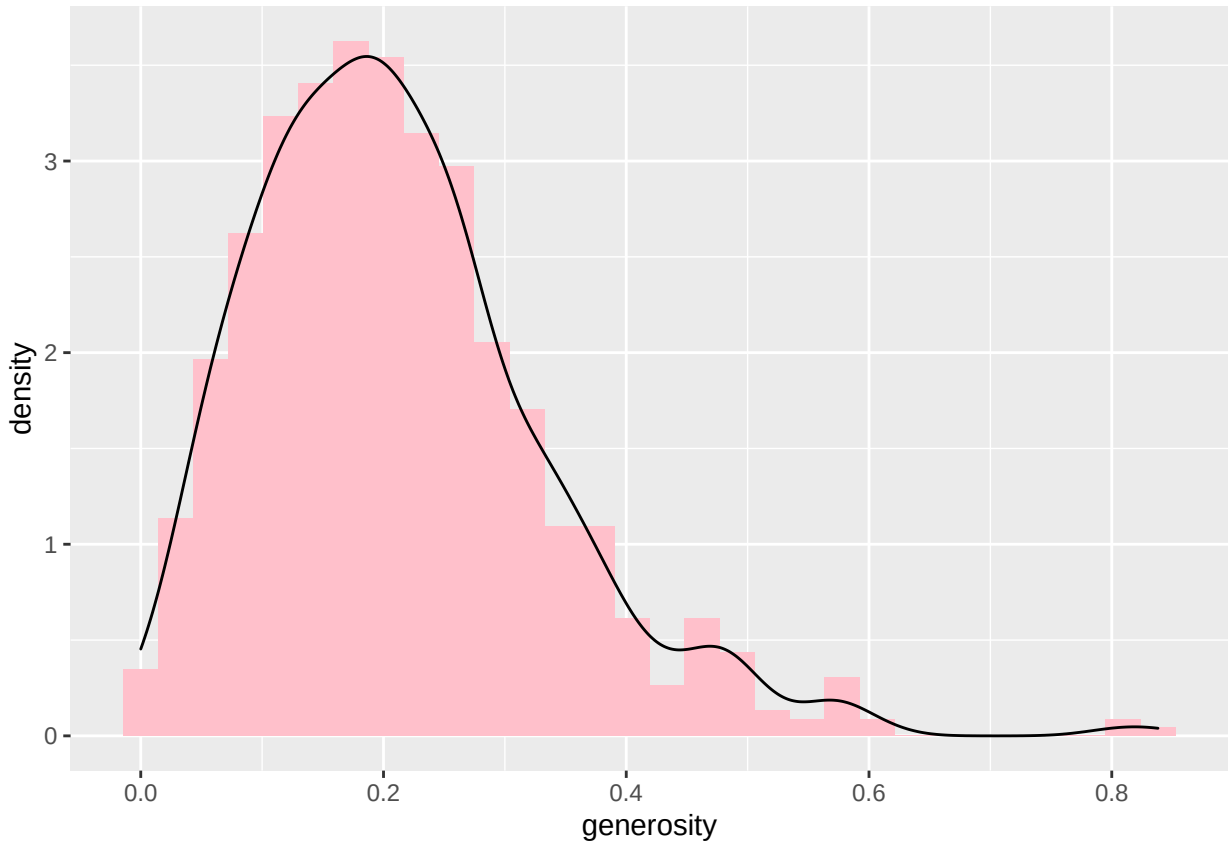

```
ggplot(WorldHappiness_df, aes(gdp_per_capita)) +geom_histogram(aes(y = ..density..), fill = "pink") + g
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(WorldHappiness_df, aes(generosity)) +geom_histogram(aes(y = ..density..), fill = "pink") + geom_
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



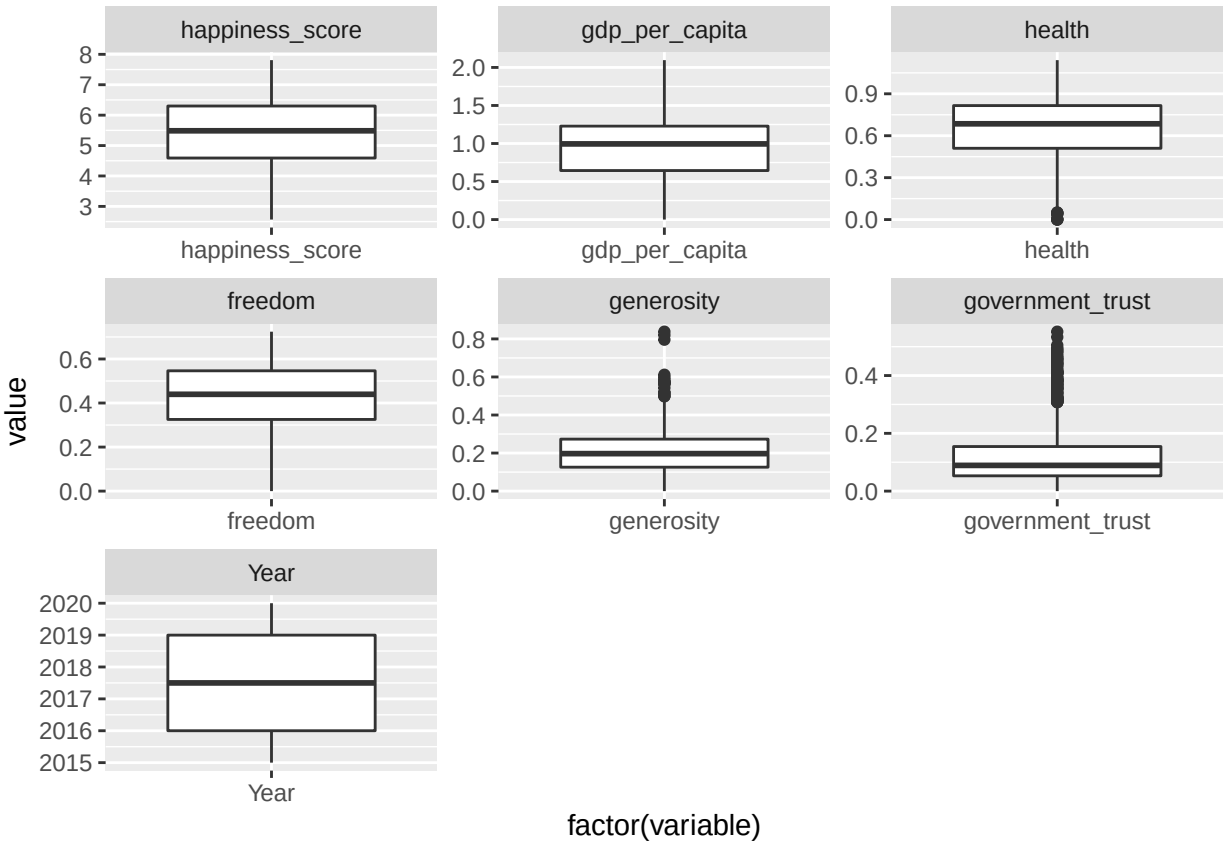
Box Plots: in order to better observe the outliers in the data I created a box plot to look at the outliers of the variables. before I do this though I created a data frame that only has the factors that sum to the happiness score along with the country and year. this is so that we aren't concerned with the other variables that will not be calculated in the coming regression models.

```
happiness_scores <- WorldHappiness_df %>%
  select(Country, happiness_score, gdp_per_capita, health, freedom, generosity, government_trust, Year)
```

```
library(reshape2)
meltData <- melt(happiness_scores)
```

```
## Using Country as id variables
```

```
p <- ggplot(meltData, aes(factor(variable), value))
p + geom_boxplot() + facet_wrap(~variable, scale="free")
```



linear regression: in this section I looked at how different scores effected the happiness scores the most. I ran linear model regressions on the factors of health and gdp per capita in regard to the happiness score. looking at the results of these we can see that the model worked much better for the gdp and they are very correlated the r value is .97 which means that it is very close to 1 which would be an exact fit. it was much more accurate than the preceding health factor which came up with an r value of .8877, much less accurate than that of the gdp model. I also ran a regression of one of the least correlated factors to the happiness score, generosity. this calculated an r value of 0.0352 making it not correlated or accurate. Due to these models come to the conclusion that the gdp has the most significant impact on the happiness score of a specific country. those that had a significant relationship over the course of all 6 years that data was collected from are health, freedom and family. the only conditions that didn't have a significant impact were government trust and generosity.

```
library(caret)
index <- createDataPartition(happiness_scores$health, p = .70, list = FALSE)
train <- happiness_scores[index, ]
test <- happiness_scores[-index, ]
```

```
health_model <- lm(health ~ ., dat = train)
summary(health_model)
```

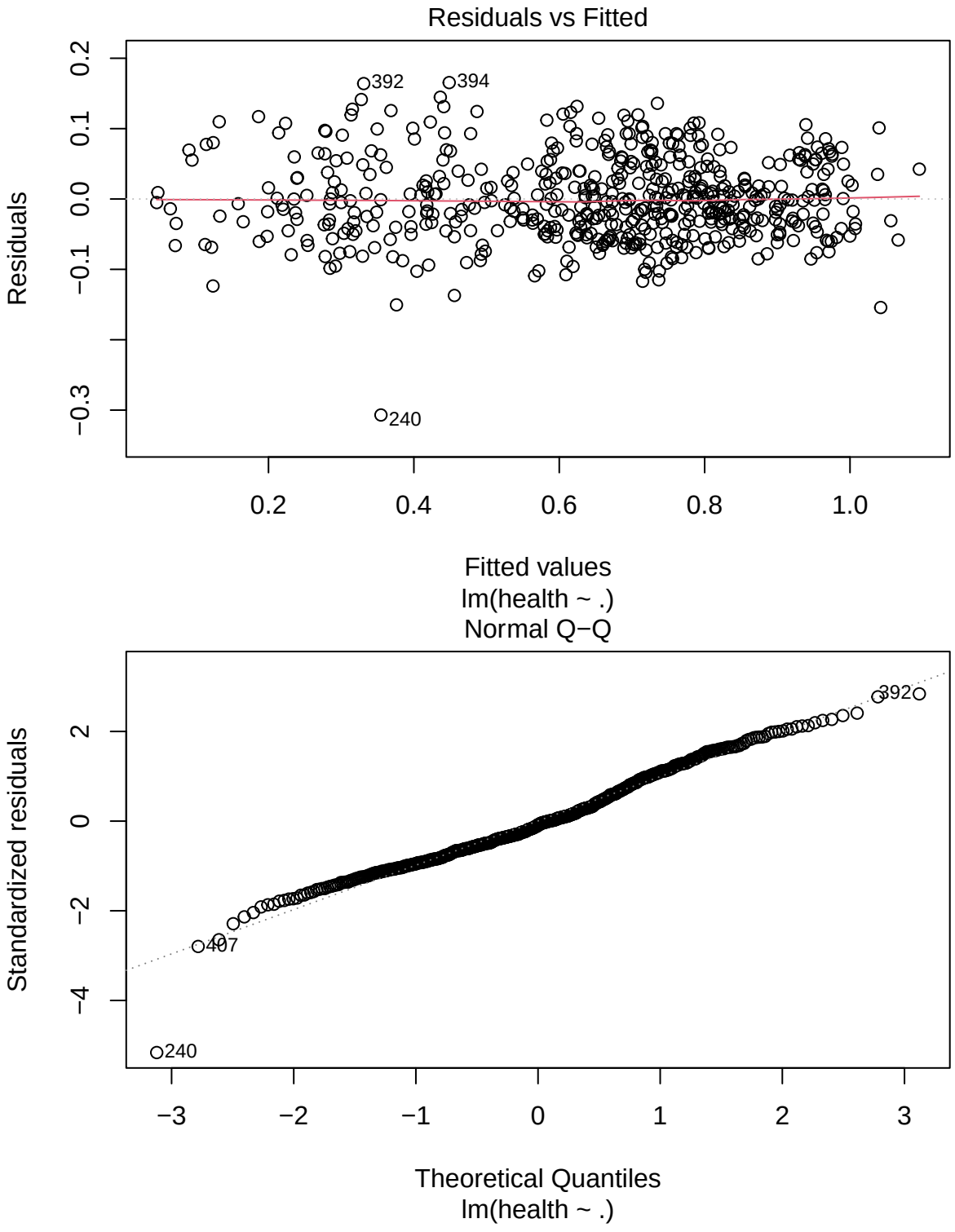
```
##
## Call:
## lm(formula = health ~ ., data = train)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.306944 -0.039617 -0.004991  0.039599  0.165421
##
```

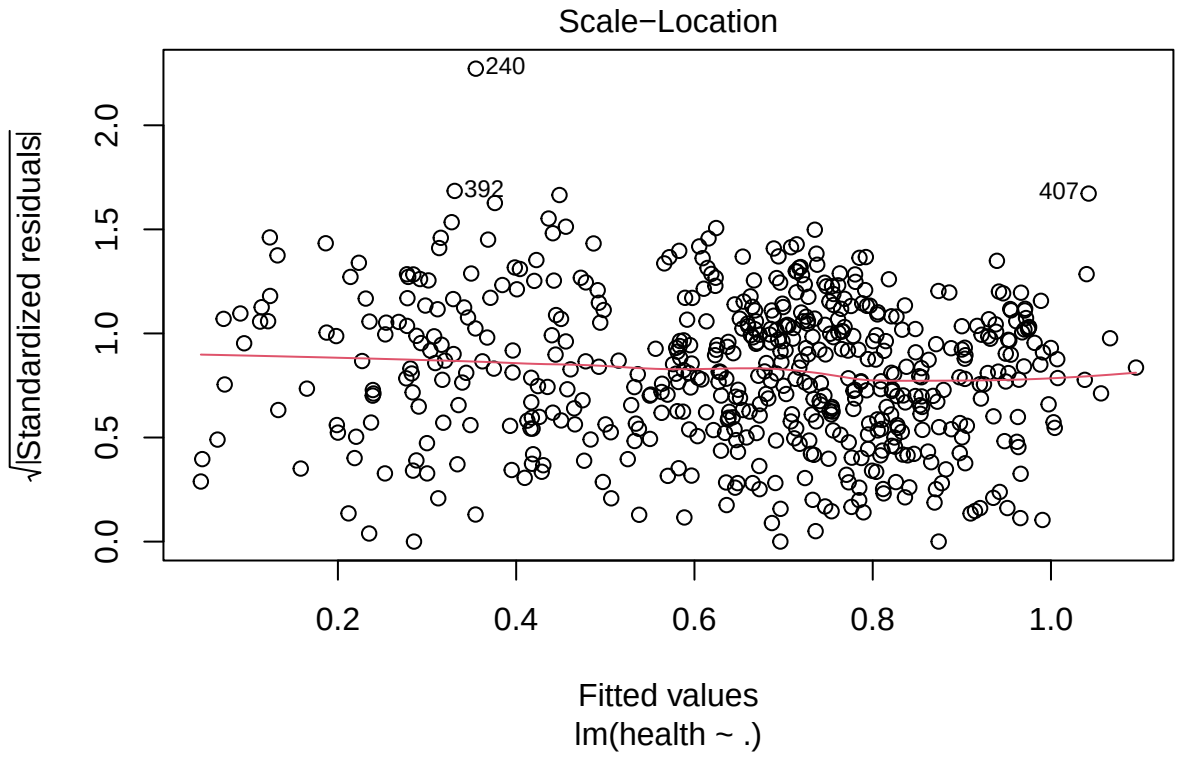
```
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  -3.3617543  3.7928012  -0.886 0.375938
## CountryAlbania                0.7220328  0.0601690  12.000  < 2e-16 ***
## CountryAlgeria                0.5497089  0.0636933   8.631  < 2e-16 ***
## CountryArgentina              0.6487476  0.0718951   9.024  < 2e-16 ***
## CountryArmenia                0.5513094  0.0550566  10.014  < 2e-16 ***
## CountryAustralia              1.0996041  0.0925725  11.878  < 2e-16 ***
## CountryAustria                0.9855055  0.0895274  11.008  < 2e-16 ***
## CountryAzerbaijan             0.5239070  0.0659615   7.943 1.85e-14 ***
## CountryBahrain                0.8380525  0.0807644  10.377  < 2e-16 ***
## CountryBangladesh             0.3478125  0.0552408   6.296 7.72e-10 ***
## CountryBelarus                0.6246656  0.0661851   9.438  < 2e-16 ***
## CountryBelgium                0.9463353  0.0854854  11.070  < 2e-16 ***
## CountryBenin                 -0.0095517  0.0552718  -0.173 0.862882
## CountryBolivia                0.4131483  0.0671124   6.156 1.75e-09 ***
## CountryBosnia and Herzegovina 0.6886897  0.0695912   9.896  < 2e-16 ***
## CountryBotswana               0.2863053  0.0667882   4.287 2.25e-05 ***
## CountryBrazil                 0.5544912  0.0699529   7.927 2.07e-14 ***
## CountryBulgaria               0.6777364  0.0652406  10.388  < 2e-16 ***
## CountryBurkina Faso          -0.0537734  0.0499913  -1.076 0.282702
## CountryBurundi               -0.2033075  0.0538448  -3.776 0.000183 ***
## CountryCambodia               0.3926035  0.0572323   6.860 2.49e-11 ***
## CountryCameroon              -0.0322842  0.0526139  -0.614 0.539810
## CountryCanada                 1.0689552  0.0903066  11.837  < 2e-16 ***
## CountryChad                  -0.1985653  0.0444447  -4.468 1.02e-05 ***
## CountryChile                  0.8312802  0.0776938  10.699  < 2e-16 ***
## CountryChina                  0.6652234  0.0681996   9.754  < 2e-16 ***
## CountryColombia               0.5289246  0.0741911   7.129 4.47e-12 ***
## CountryCosta Rica             0.7097093  0.0771764   9.196  < 2e-16 ***
## CountryCroatia                0.7094018  0.0682495  10.394  < 2e-16 ***
## CountryCyprus                 0.9866752  0.0739297  13.346  < 2e-16 ***
## CountryDenmark                0.9884729  0.0972498  10.164  < 2e-16 ***
## CountryDominican Republic     0.5668658  0.0671117   8.447 5.01e-16 ***
## CountryEcuador                0.6263622  0.0676664   9.257  < 2e-16 ***
## CountryEgypt                  0.4564310  0.0569186   8.019 1.08e-14 ***
## CountryEl Salvador            0.4440234  0.0637251   6.968 1.26e-11 ***
## CountryEstonia                0.8454318  0.0965665   8.755  < 2e-16 ***
## CountryEthiopia               0.1223258  0.0494705   2.473 0.013806 *
## CountryFinland                0.9187222  0.0951780   9.653  < 2e-16 ***
## CountryFrance                 0.9228568  0.0822715  11.217  < 2e-16 ***
## CountryGabon                  0.3642020  0.0636015   5.726 1.96e-08 ***
## CountryGeorgia                0.5324766  0.0632016   8.425 5.86e-16 ***
## CountryGermany                1.0195266  0.0878410  11.606  < 2e-16 ***
## CountryGhana                  0.2093112  0.0551873   3.793 0.000171 ***
## CountryGreece                 0.7663224  0.0678283  11.298  < 2e-16 ***
## CountryGuatemala              0.4603751  0.0665885   6.914 1.77e-11 ***
## CountryGuinea                -0.0155681  0.0475336  -0.328 0.743441
## CountryHaiti                  0.1857828  0.0478869   3.880 0.000122 ***
## CountryHonduras               0.5030346  0.0569410   8.834  < 2e-16 ***
## CountryHungary                0.6900304  0.0668674  10.319  < 2e-16 ***
## CountryIceland                1.0801700  0.0926577  11.658  < 2e-16 ***
## CountryIndia                  0.3597059  0.0650906   5.526 5.77e-08 ***
## CountryIndonesia              0.6885825  0.0677378  10.165  < 2e-16 ***
```

```
## CountryIran                0.7907394  0.0666680  11.861  < 2e-16 ***
## CountryIraq                0.4952087  0.0622326   7.957 1.67e-14 ***
## CountryIreland             1.1059998  0.0910201  12.151  < 2e-16 ***
## CountryIsrael              0.9750701  0.0812627  11.999  < 2e-16 ***
## CountryItaly               0.9658645  0.0751084  12.860  < 2e-16 ***
## CountryJamaica             0.5386991  0.0633696   8.501 3.36e-16 ***
## CountryJapan               0.9784374  0.0805031  12.154  < 2e-16 ***
## CountryJordan              0.5525912  0.0860334   6.423 3.64e-10 ***
## CountryKazakhstan          0.6096021  0.0739665   8.242 2.21e-15 ***
## CountryKenya               0.3024315  0.0530994   5.696 2.32e-08 ***
## CountryKosovo              0.5594152  0.0655485   8.534 2.63e-16 ***
## CountryKuwait              0.8272293  0.0867172   9.539  < 2e-16 ***
## CountryKyrgyzstan          0.4541805  0.0568659   7.987 1.36e-14 ***
## CountryLatvia              0.6653602  0.0688268   9.667  < 2e-16 ***
## CountryLebanon             0.7446767  0.0613477  12.139  < 2e-16 ***
## CountryLiberia            -0.0762342  0.0503113  -1.515 0.130465
## CountryLibya               0.4880856  0.0714334   6.833 2.96e-11 ***
## CountryLithuania           0.6215193  0.0710058   8.753  < 2e-16 ***
## CountryLuxembourg          1.1202531  0.0974306  11.498  < 2e-16 ***
## CountryMadagascar          0.1008964  0.0482561   2.091 0.037144 *
## CountryMalawi              0.0137439  0.0519445   0.265 0.791458
## CountryMalaysia            0.7989272  0.0753684  10.600  < 2e-16 ***
## CountryMali               -0.1074422  0.0508833  -2.112 0.035319 *
## CountryMalta               1.0946244  0.0843526  12.977  < 2e-16 ***
## CountryMauritania          0.1115676  0.0779952   1.430 0.153337
## CountryMauritius           0.7341277  0.0728832  10.073  < 2e-16 ***
## CountryMexico              0.6463416  0.0748892   8.631  < 2e-16 ***
## CountryMoldova             0.3861664  0.0588121   6.566 1.54e-10 ***
## CountryMongolia            0.5612134  0.0602770   9.311  < 2e-16 ***
## CountryMontenegro          0.6906650  0.0642995  10.741  < 2e-16 ***
## CountryMorocco             0.4316964  0.0687753   6.277 8.65e-10 ***
## CountryMyanmar             0.5183388  0.0662599   7.823 4.26e-14 ***
## CountryNepal               0.3585106  0.0552445   6.490 2.44e-10 ***
## CountryNetherlands         1.0678157  0.0909013  11.747  < 2e-16 ***
## CountryNew Zealand         1.0589240  0.1009560  10.489  < 2e-16 ***
## CountryNicaragua           0.5354742  0.0601476   8.903  < 2e-16 ***
## CountryNiger              -0.1068090  0.0494478  -2.160 0.031338 *
## CountryNigeria            -0.0768247  0.0574240  -1.338 0.181673
## CountryNorway              1.0718048  0.0967382  11.079  < 2e-16 ***
## CountryPakistan            0.2882131  0.0545031   5.288 2.00e-07 ***
## CountryPanama              0.7299630  0.0766850   9.519  < 2e-16 ***
## CountryParaguay            0.5540935  0.0638392   8.680  < 2e-16 ***
## CountryPeru                0.5984639  0.0699339   8.558  < 2e-16 ***
## CountryPhilippines         0.3601972  0.0685829   5.252 2.40e-07 ***
## CountryPoland              0.8001917  0.0838641   9.542  < 2e-16 ***
## CountryPortugal            0.8392020  0.0825422  10.167  < 2e-16 ***
## CountryRomania             0.6800382  0.0778535   8.735  < 2e-16 ***
## CountryRussia              0.5340552  0.0714636   7.473 4.63e-13 ***
## CountryRwanda              0.2357041  0.0659142   3.576 0.000390 ***
## CountrySaudi Arabia        0.7301114  0.0825653   8.843  < 2e-16 ***
## CountrySenegal             0.1602816  0.0493659   3.247 0.001261 **
## CountrySerbia              0.6453158  0.0600321  10.750  < 2e-16 ***
## CountrySierra Leone       -0.1810338  0.0505561  -3.581 0.000383 ***
## CountrySingapore           1.2673609  0.0970903  13.053  < 2e-16 ***
```

```
## CountrySlovakia                0.7467353  0.0716840  10.417  < 2e-16 ***
## CountrySlovenia                0.8954778  0.0772436  11.593  < 2e-16 ***
## CountrySouth Africa            0.2316775  0.0612176   3.784 0.000177 ***
## CountrySpain                   0.9801774  0.0808135  12.129  < 2e-16 ***
## CountrySri Lanka               0.7645325  0.0694730  11.005  < 2e-16 ***
## CountrySweden                  1.0469215  0.0936406  11.180  < 2e-16 ***
## CountrySwitzerland             1.0778952  0.0953193  11.308  < 2e-16 ***
## CountryTajikistan              0.3507794  0.0539892   6.497 2.33e-10 ***
## CountryTanzania                0.2510172  0.0524399   4.787 2.36e-06 ***
## CountryThailand                0.7961851  0.0796175  10.000  < 2e-16 ***
## CountryTogo                    0.0003929  0.0534940   0.007 0.994143
## CountryTunisia                 0.6318151  0.0673040   9.387  < 2e-16 ***
## CountryTurkey                  0.6290304  0.0668097   9.415  < 2e-16 ***
## CountryTurkmenistan            0.4548073  0.0716984   6.343 5.85e-10 ***
## CountryUganda                  0.0411034  0.0490476   0.838 0.402493
## CountryUkraine                 0.5460438  0.0552455   9.884  < 2e-16 ***
## CountryUnited Arab Emirates    0.9554428  0.0962126   9.931  < 2e-16 ***
## CountryUnited Kingdom          1.0489822  0.0850797  12.329  < 2e-16 ***
## CountryUnited States           0.9828940  0.0886866  11.083  < 2e-16 ***
## CountryUruguay                 0.7412177  0.0754998   9.817  < 2e-16 ***
## CountryUzbekistan              0.5140904  0.0694828   7.399 7.60e-13 ***
## CountryVenezuela               0.5476405  0.0631295   8.675  < 2e-16 ***
## CountryVietnam                 0.5408864  0.0711024   7.607 1.87e-13 ***
## CountryYemen                   0.0857266  0.0496706   1.726 0.085104 .
## CountryZambia                  0.1057331  0.0653010   1.619 0.106166
## CountryZimbabwe                0.0175965  0.0477188   0.369 0.712497
## happiness_score                0.0422761  0.0129888   3.255 0.001227 **
## gdp_per_capita                -0.4355235  0.0506090  -8.606  < 2e-16 ***
## freedom                       -0.0460220  0.0615719  -0.747 0.455211
## generosity                    -0.6047343  0.0727337  -8.314 1.31e-15 ***
## government_trust              -0.0727879  0.1006932  -0.723 0.470165
## Year                           0.0018724  0.0018792   0.996 0.319623
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06719 on 418 degrees of freedom
## Multiple R-squared:  0.938,  Adjusted R-squared:  0.9177
## F-statistic: 46.17 on 137 and 418 DF,  p-value: < 2.2e-16
```

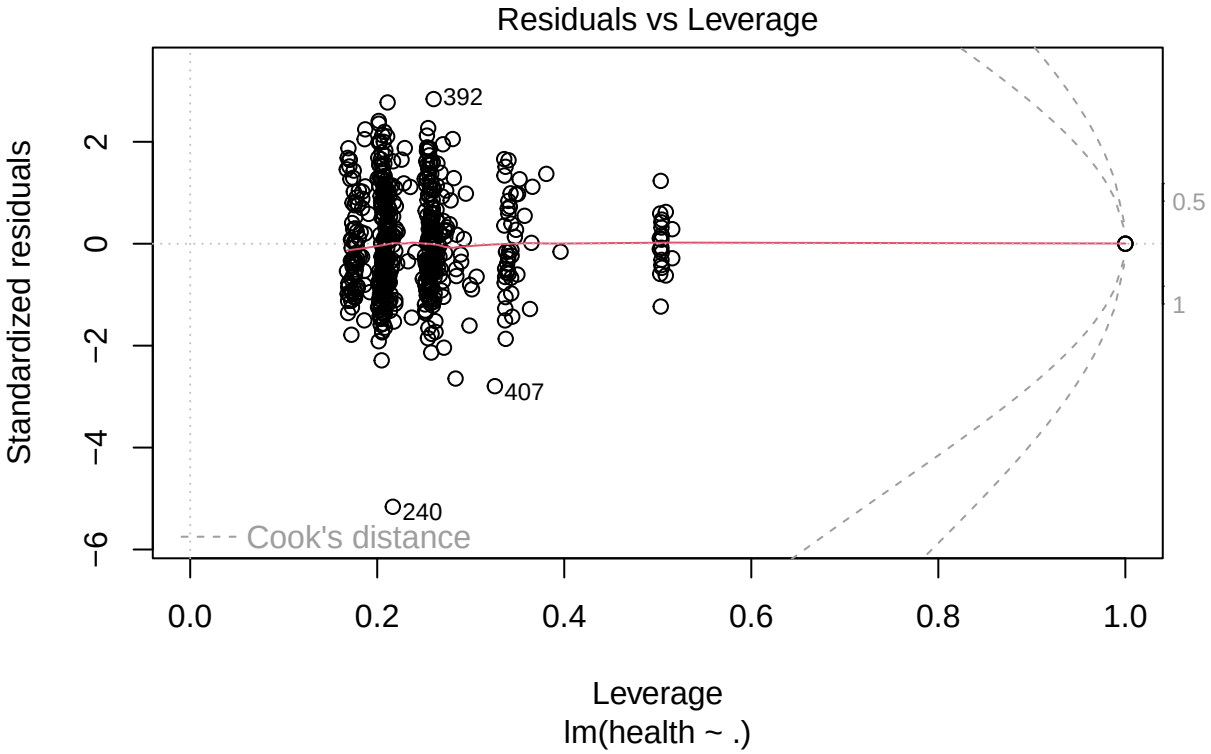
```
plot(health_model)
```


Residuals vs Fitted
Residuals
Fitted values
lm(health ~ .)
392
394
240
0.2
0.1
0.0
−0.1
−0.3
0.2
0.4
0.6
0.8
1.0
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(health ~ .)
392
407
240
2
0
−2
−4
−3
−2
−1
0
1
2
3




```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```



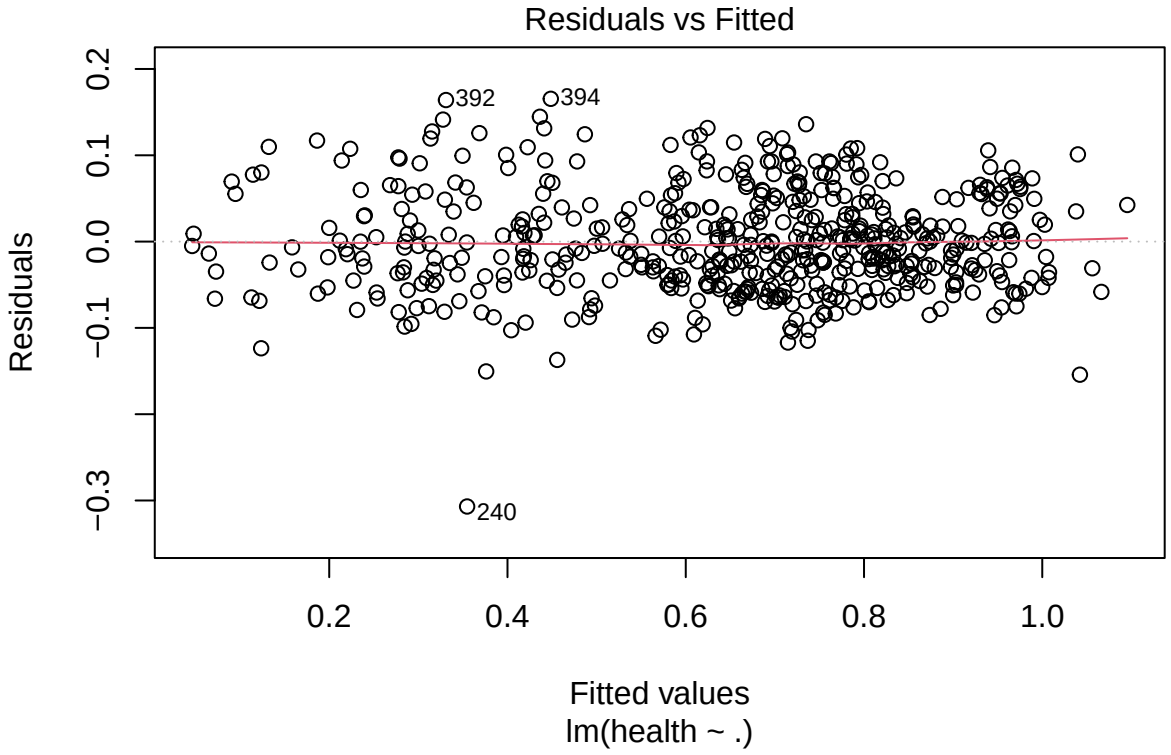

```r
test$predicted_health <- predict(health_model, test)
head(test[, c("health", "predicted_health")])
```

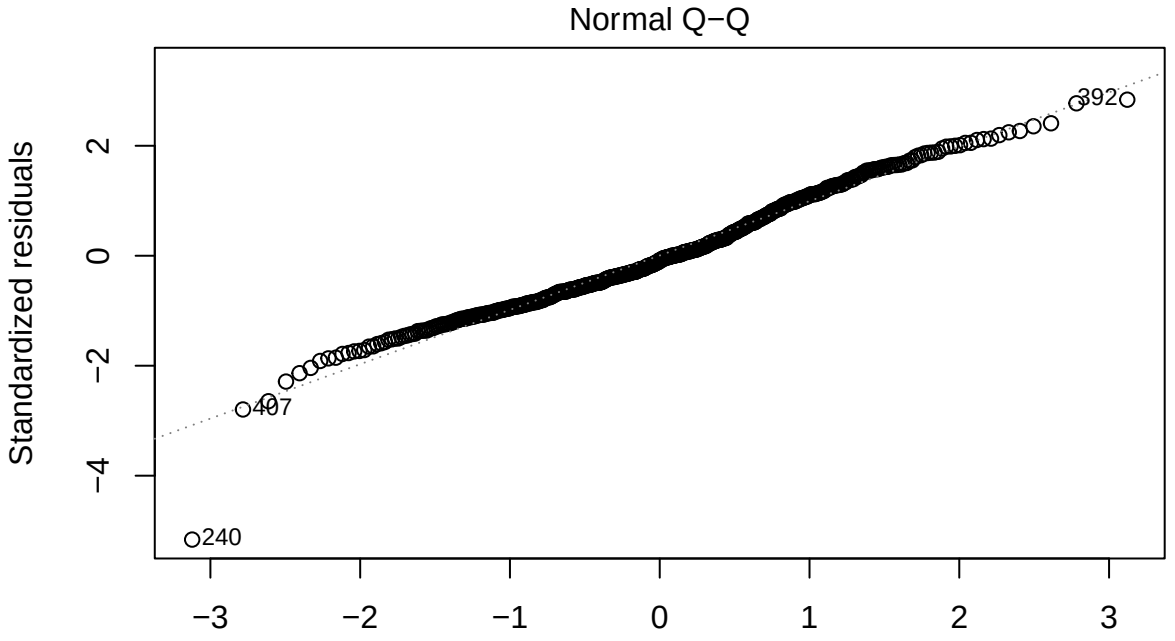
```
##       health predicted_health
## 1  0.7966665        0.8264892
## 4  0.8581313        0.8933914
## 7  0.8345577        0.8328120
## 17 0.8180919        0.8282351
## 19 0.8053359        0.7754734
## 20 0.8194797        0.7572249
```

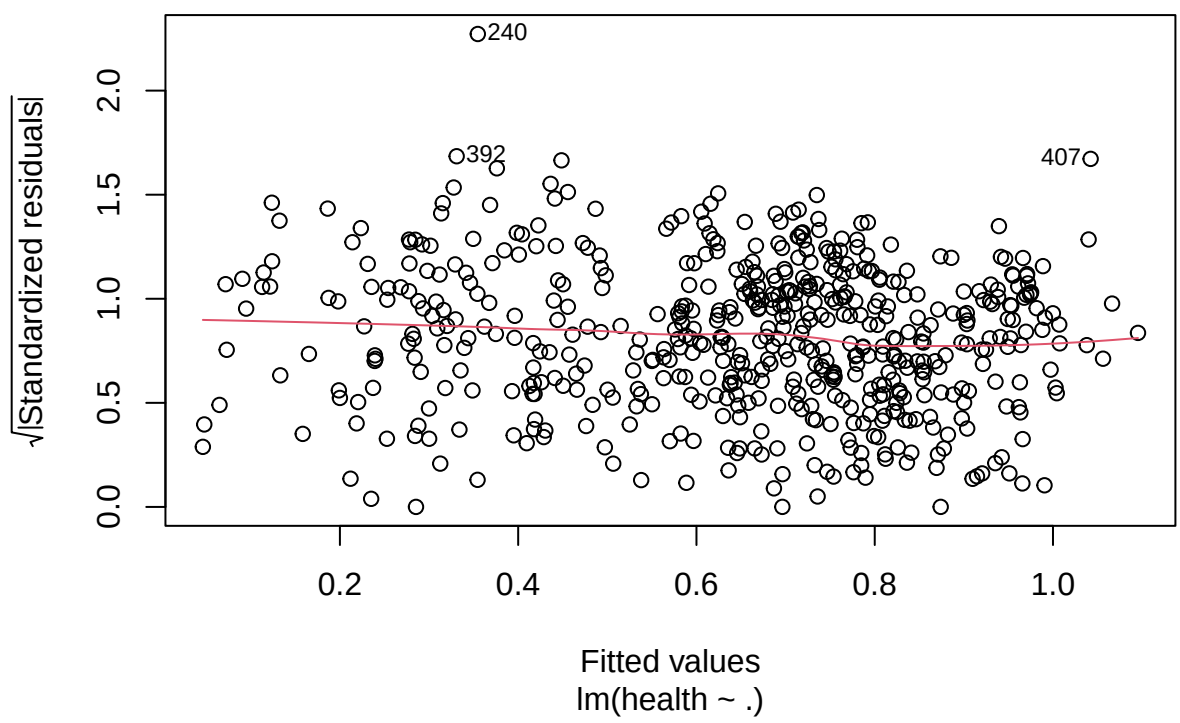
```
(mfrow=c(2,2))
```

```
## [1] 2 2
```

```
plot(health_model)
```

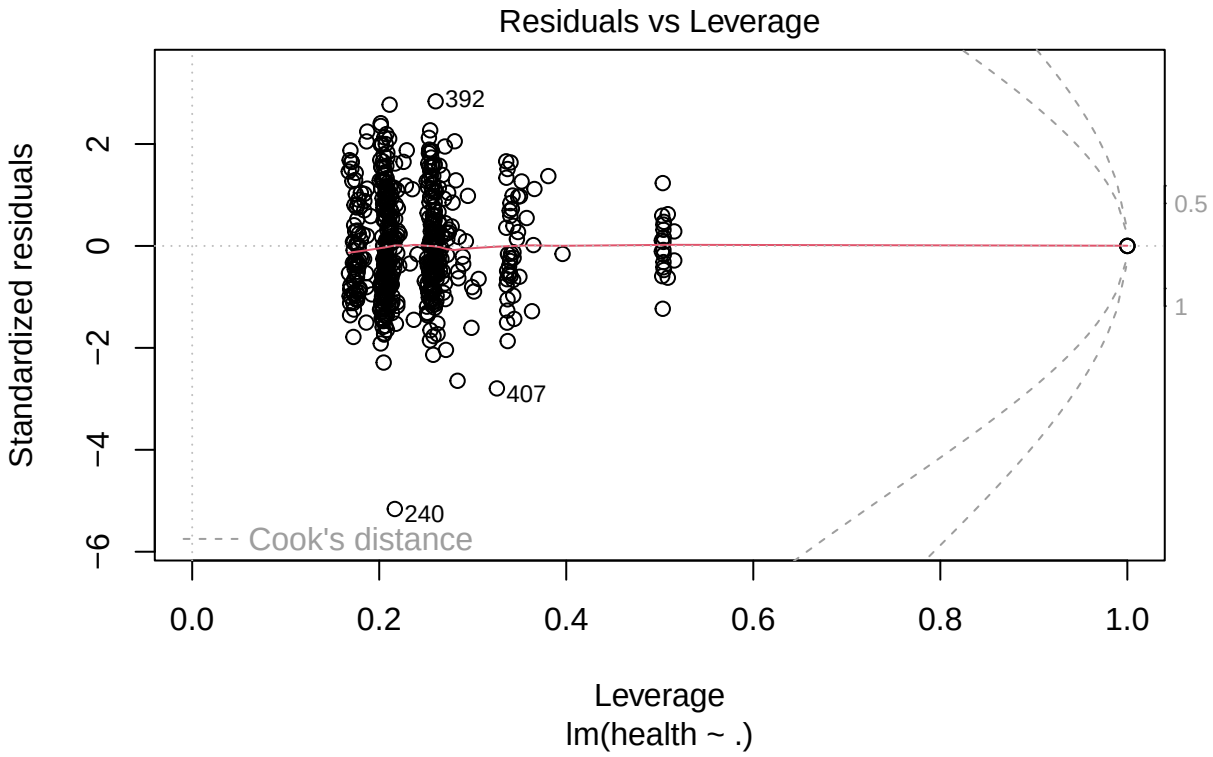







```
## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced

## Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```




```
par(mfrow=c(1,1))
```

```
test$predicted_health <- predict(health_model, test)
head(test[,c("health", "predicted_health")])
```

```
##       health predicted_health
## 1  0.7966665        0.8264892
## 4  0.8581313        0.8933914
## 7  0.8345577        0.8328120
## 17 0.8180919        0.8282351
## 19 0.8053359        0.7754734
## 20 0.8194797        0.7572249
```

```
real <- test$health
prediction <- test$predicted_health
rss <- sum((prediction - real) ^ 2)
tss <- sum((real - mean(real)) ^ 2)
rsq <- 1 - rss/tss
rsq
```

```
## [1] 0.895839
```

```
library(caret)
index <- createDataPartition(happiness_scores$gdp_per_capita, p = .70, list = FALSE)
train <- happiness_scores[index, ]
test <- happiness_scores[-index, ]
```

```
gdp_model <- lm(gdp_per_capita ~ ., dat = train)
summary(gdp_model)
```

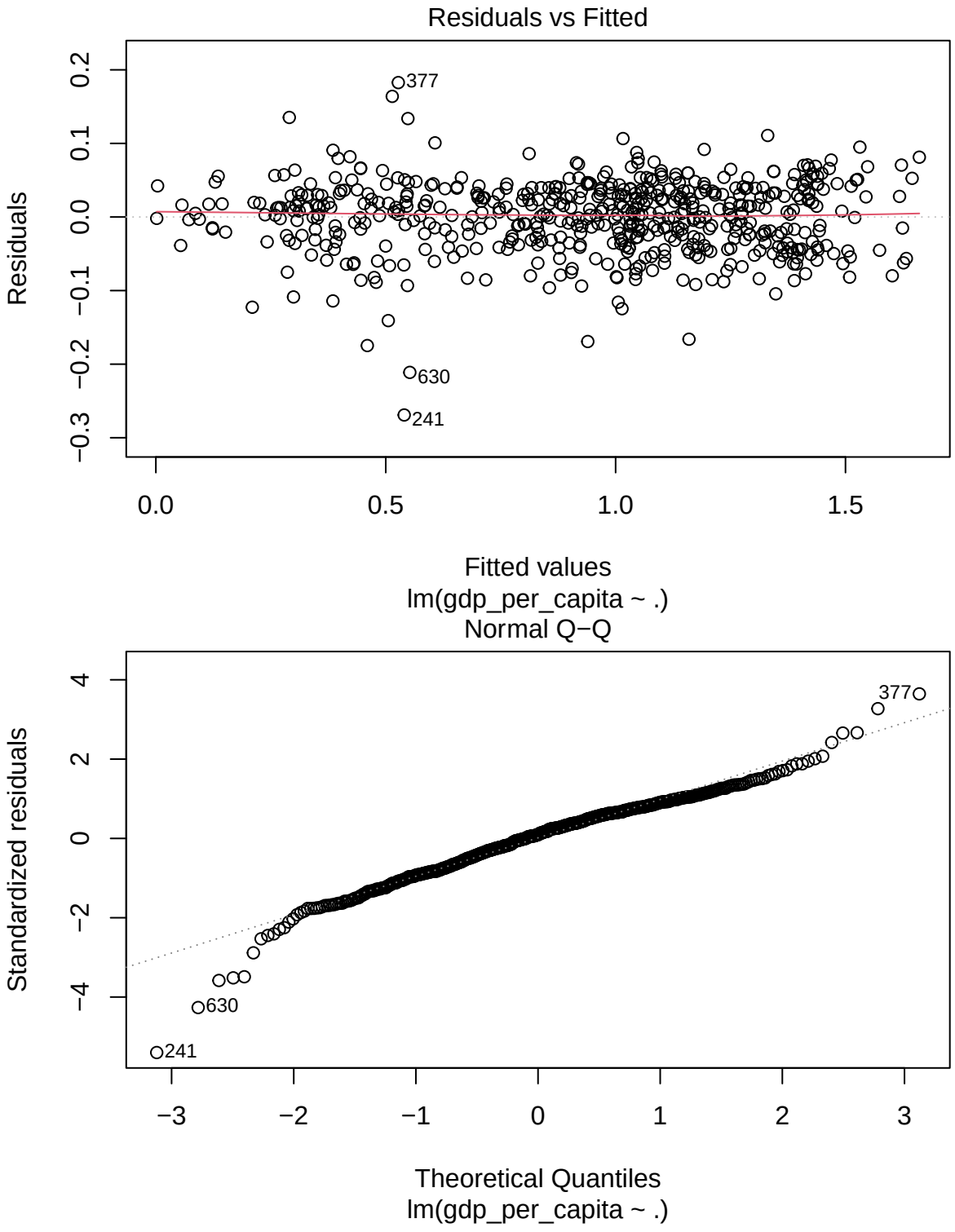
```
##
## Call:
```

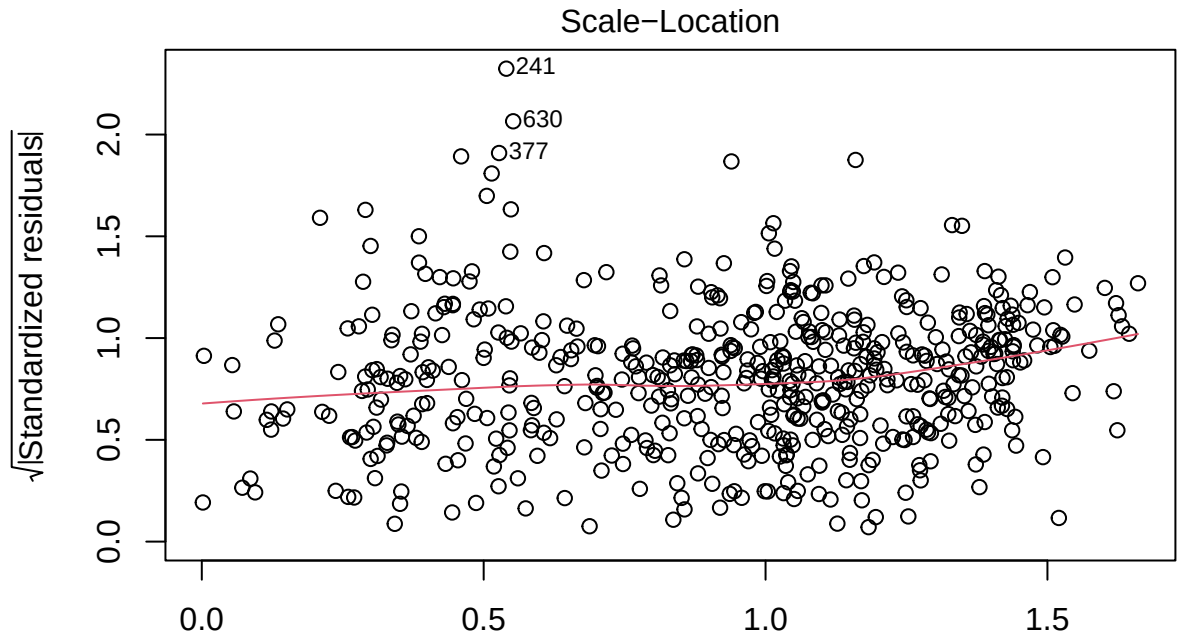
```
## lm(formula = gdp_per_capita ~ ., data = train)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.269092 -0.031185  0.004792  0.032514  0.182653
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   12.151985   3.178981   3.823 0.000152 ***
## CountryAlbania                 0.793053   0.057230  13.857  < 2e-16 ***
## CountryAlgeria                 0.734890   0.052395  14.026  < 2e-16 ***
## CountryArgentina               0.867534   0.054977  15.780  < 2e-16 ***
## CountryArmenia                 0.591888   0.047468  12.469  < 2e-16 ***
## CountryAustralia               1.277000   0.067461  18.929  < 2e-16 ***
## CountryAustria                 1.225873   0.062277  19.684  < 2e-16 ***
## CountryAzerbaijan              0.817802   0.049193  16.624  < 2e-16 ***
## CountryBahrain                 1.175156   0.055386  21.218  < 2e-16 ***
## CountryBangladesh              0.341983   0.048297   7.081 6.10e-12 ***
## CountryBelarus                 0.811614   0.051697  15.700  < 2e-16 ***
## CountryBelgium                 1.211575   0.061140  19.816  < 2e-16 ***
## CountryBenin                   0.058967   0.043621   1.352 0.177171
## CountryBolivia                 0.522356   0.052293   9.989  < 2e-16 ***
## CountryBosnia and Herzegovina  0.689719   0.049374  13.969  < 2e-16 ***
## CountryBotswana                0.751132   0.047816  15.709  < 2e-16 ***
## CountryBrazil                  0.754541   0.054856  13.755  < 2e-16 ***
## CountryBulgaria                0.887349   0.048850  18.165  < 2e-16 ***
## CountryBurkina Faso           -0.022175   0.044610  -0.497 0.619393
## CountryBurundi                -0.305092   0.041319  -7.384 8.40e-13 ***
## CountryCambodia                0.363532   0.050366   7.218 2.51e-12 ***
## CountryCameroon                0.148428   0.045405   3.269 0.001168 **
## CountryCanada                  1.237571   0.070012  17.677  < 2e-16 ***
## CountryChad                   -0.049970   0.044726  -1.117 0.264526
## CountryChile                   0.956816   0.057076  16.764  < 2e-16 ***
## CountryChina                   0.826883   0.050532  16.363  < 2e-16 ***
## CountryColombia                0.736184   0.053987  13.636  < 2e-16 ***
## CountryCosta Rica              0.836284   0.060669  13.784  < 2e-16 ***
## CountryCroatia                 0.916572   0.050235  18.246  < 2e-16 ***
## CountryCyprus                  1.101656   0.055219  19.951  < 2e-16 ***
## CountryDenmark                 1.244636   0.069344  17.949  < 2e-16 ***
## CountryDominican Republic      0.790272   0.053484  14.776  < 2e-16 ***
## CountryEcuador                 0.692201   0.054671  12.661  < 2e-16 ***
## CountryEgypt                   0.659300   0.048984  13.459  < 2e-16 ***
## CountryEl Salvador             0.569874   0.053704  10.611  < 2e-16 ***
## CountryEstonia                 1.045677   0.055087  18.982  < 2e-16 ***
## CountryEthiopia                0.026587   0.042737   0.622 0.534210
## CountryFinland                 1.161280   0.069761  16.647  < 2e-16 ***
## CountryFrance                  1.174161   0.058467  20.082  < 2e-16 ***
## CountryGabon                   0.756483   0.044966  16.824  < 2e-16 ***
## CountryGeorgia                 0.659612   0.051891  12.711  < 2e-16 ***
## CountryGermany                 1.235531   0.062937  19.631  < 2e-16 ***
## CountryGhana                   0.306514   0.043600   7.030 8.46e-12 ***
## CountryGreece                  0.961846   0.050864  18.910  < 2e-16 ***
## CountryGuatemala               0.546213   0.057507   9.498  < 2e-16 ***
## CountryGuinea                  0.048928   0.048062   1.018 0.309258
```

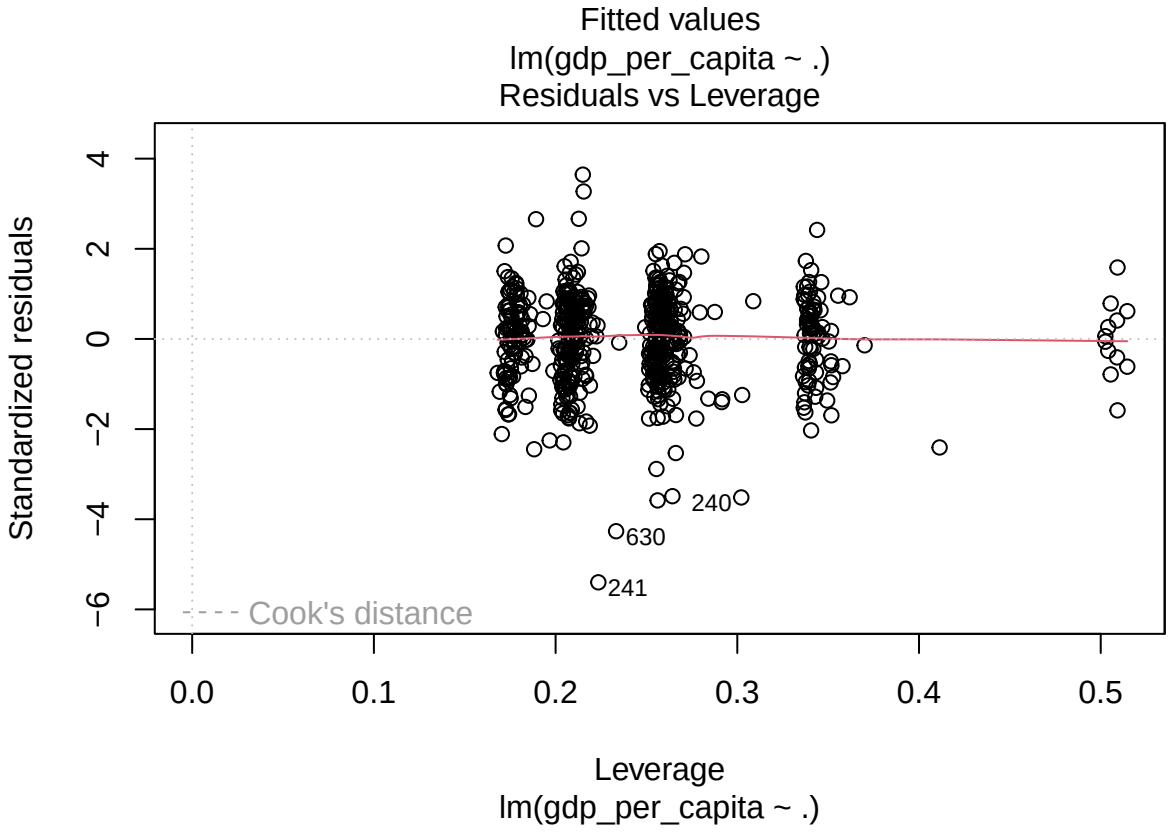
```
## CountryHaiti                0.007948   0.042310   0.188 0.851089
## CountryHonduras             0.428784   0.049572   8.650  < 2e-16 ***
## CountryHungary              0.929059   0.058548  15.868  < 2e-16 ***
## CountryIceland              1.248660   0.068360  18.266  < 2e-16 ***
## CountryIndia                0.535183   0.047153  11.350  < 2e-16 ***
## CountryIndonesia            0.657584   0.061459  10.700  < 2e-16 ***
## CountryIran                 0.861949   0.051362  16.782  < 2e-16 ***
## CountryIraq                 0.743116   0.046138  16.106  < 2e-16 ***
## CountryIreland              1.284167   0.067344  19.069  < 2e-16 ***
## CountryIsrael               1.114822   0.062572  17.817  < 2e-16 ***
## CountryItaly                1.077549   0.064165  16.793  < 2e-16 ***
## CountryJamaica              0.606468   0.056369  10.759  < 2e-16 ***
## CountryJapan                1.197232   0.056821  21.070  < 2e-16 ***
## CountryJordan               0.655880   0.047947  13.679  < 2e-16 ***
## CountryKazakhstan           0.921628   0.052421  17.581  < 2e-16 ***
## CountryKenya                0.236652   0.048064   4.924 1.23e-06 ***
## CountryKosovo               0.604077   0.049146  12.291  < 2e-16 ***
## CountryKuwait               1.295269   0.057641  22.471  < 2e-16 ***
## CountryKyrgyzstan           0.314410   0.050324   6.248 1.03e-09 ***
## CountryLatvia               0.924151   0.051295  18.016  < 2e-16 ***
## CountryLebanon              0.811741   0.047225  17.189  < 2e-16 ***
## CountryLiberia             -0.211420   0.052330  -4.040 6.36e-05 ***
## CountryLibya                0.801781   0.053028  15.120  < 2e-16 ***
## CountryLithuania            0.943496   0.052394  18.008  < 2e-16 ***
## CountryLuxembourg           1.486629   0.062998  23.598  < 2e-16 ***
## CountryMadagascar          -0.045669   0.043901  -1.040 0.298817
## CountryMalawi              -0.082971   0.045514  -1.823 0.069021 .
## CountryMalaysia             1.018464   0.052165  19.524  < 2e-16 ***
## CountryMali                 0.003662   0.048168   0.076 0.939427
## CountryMalta                1.160428   0.060923  19.047  < 2e-16 ***
## CountryMauritania           0.213494   0.048038   4.444 1.13e-05 ***
## CountryMauritius            0.902918   0.053889  16.755  < 2e-16 ***
## CountryMexico               0.814403   0.060861  13.381  < 2e-16 ***
## CountryMoldova              0.390386   0.049673   7.859 3.31e-14 ***
## CountryMongolia             0.678968   0.047217  14.380  < 2e-16 ***
## CountryMontenegro           0.790841   0.051960  15.220  < 2e-16 ***
## CountryMorocco              0.554114   0.048719  11.374  < 2e-16 ***
## CountryMyanmar              0.362843   0.054059   6.712 6.27e-11 ***
## CountryNepal                0.212578   0.045711   4.650 4.45e-06 ***
## CountryNetherlands          1.253160   0.065761  19.056  < 2e-16 ***
## CountryNew Zealand          1.180191   0.068106  17.329  < 2e-16 ***
## CountryNicaragua            0.463387   0.054419   8.515 3.03e-16 ***
## CountryNiger               -0.192265   0.045412  -4.234 2.83e-05 ***
## CountryNigeria              0.286924   0.050386   5.695 2.33e-08 ***
## CountryNorway               1.353151   0.067550  20.032  < 2e-16 ***
## CountryPakistan             0.357448   0.050530   7.074 6.38e-12 ***
## CountryPanama               0.953786   0.055735  17.113  < 2e-16 ***
## CountryParaguay             0.641531   0.050999  12.579  < 2e-16 ***
## CountryPeru                 0.719414   0.052794  13.627  < 2e-16 ***
## CountryPhilippines          0.536311   0.052060  10.302  < 2e-16 ***
## CountryPoland               0.982645   0.054348  18.081  < 2e-16 ***
## CountryPortugal             1.054621   0.054196  19.459  < 2e-16 ***
## CountryRomania              0.888353   0.060069  14.789  < 2e-16 ***
## CountryRussia               0.895243   0.050247  17.817  < 2e-16 ***
```

```
## CountryRwanda                  0.191177   0.060119   3.180 0.001582 **
## CountrySaudi Arabia            1.175197   0.056234  20.898  < 2e-16 ***
## CountrySenegal                 0.155516   0.045410   3.425 0.000676 ***
## CountrySerbia                  0.758700   0.047765  15.884  < 2e-16 ***
## CountrySierra Leone           -0.086255   0.045470  -1.897 0.058521 .
## CountrySingapore               1.500183   0.066601  22.525  < 2e-16 ***
## CountrySlovakia                1.003029   0.053196  18.855  < 2e-16 ***
## CountrySlovenia                1.106213   0.059050  18.733  < 2e-16 ***
## CountrySouth Africa            0.633165   0.045192  14.010  < 2e-16 ***
## CountrySpain                   1.125476   0.058044  19.390  < 2e-16 ***
## CountrySri Lanka               0.766957   0.050710  15.124  < 2e-16 ***
## CountrySweden                  1.266495   0.068518  18.484  < 2e-16 ***
## CountrySwitzerland             1.319629   0.069184  19.074  < 2e-16 ***
## CountryTajikistan              0.247590   0.047695   5.191 3.27e-07 ***
## CountryTanzania                0.179967   0.044871   4.011 7.17e-05 ***
## CountryThailand                0.867865   0.057929  14.981  < 2e-16 ***
## CountryTogo                   -0.037841   0.047815  -0.791 0.429152
## CountryTunisia                 0.684407   0.048295  14.171  < 2e-16 ***
## CountryTurkey                  0.894486   0.047765  18.727  < 2e-16 ***
## CountryTurkmenistan            0.777223   0.050835  15.289  < 2e-16 ***
## CountryUganda                  0.003208   0.048404   0.066 0.947197
## CountryUkraine                 0.576763   0.044902  12.845  < 2e-16 ***
## CountryUnited Arab Emirates    1.336685   0.061710  21.661  < 2e-16 ***
## CountryUnited Kingdom          1.156713   0.065653  17.619  < 2e-16 ***
## CountryUnited States           1.254753   0.059284  21.165  < 2e-16 ***
## CountryUruguay                 0.929312   0.056342  16.494  < 2e-16 ***
## CountryUzbekistan              0.539039   0.059276   9.094  < 2e-16 ***
## CountryVenezuela               0.729261   0.050156  14.540  < 2e-16 ***
## CountryVietnam                 0.537169   0.052025  10.325  < 2e-16 ***
## CountryYemen                   0.191891   0.043132   4.449 1.11e-05 ***
## CountryZambia                  0.284066   0.045845   6.196 1.39e-09 ***
## CountryZimbabwe                0.031579   0.041519   0.761 0.447340
## happiness_score                0.031976   0.010868   2.942 0.003439 **
## health                        -0.337953   0.038468  -8.785  < 2e-16 ***
## freedom                       -0.176411   0.049919  -3.534 0.000455 ***
## generosity                    -0.129540   0.066094  -1.960 0.050668 .
## government_trust              -0.055303   0.088440  -0.625 0.532109
## Year                          -0.005832   0.001575  -3.702 0.000242 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05656 on 418 degrees of freedom
## Multiple R-squared:  0.9836, Adjusted R-squared:  0.9782
## F-statistic: 182.8 on 137 and 418 DF,  p-value: < 2.2e-16
```

```
plot(gdp_model)
```


Residuals vs Fitted
Residuals
Fitted values
lm(gdp_per_capita ~ .)
377
630
241
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(gdp_per_capita ~ .)






```
test$predicted_gdp <- predict(gdp_model, test)
head(test[, c("gdp_per_capita", "predicted_gdp")])
```

```
##    gdp_per_capita predicted_gdp
## 4       1.5649796     1.5016910
## 5       1.4435719     1.3644419
## 15      1.5357066     1.4603295
```

```
## 20      1.2527846     1.1789835
## 24      1.1531838     1.0887390
## 27      0.8720019     0.8350511
```

```
real <- test$gdp_per_capita
prediction <- test$predicted_gdp
rss <- sum((prediction - real) ^ 2)
tss <- sum((real - mean(real)) ^ 2)
rsq <- 1 - rss/tss
rsq
```

```
## [1] 0.9691698
```

```
index <- createDataPartition(happiness_scores$generosity, p = .70, list = FALSE)
train <- happiness_scores[index, ]
test <- happiness_scores[-index, ]
```

```
gen_model <- lm(generosity ~ ., dat = train)
summary(gen_model)
```

```
##
## Call:
## lm(formula = generosity ~ ., data = train)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.129165 -0.022314 -0.001849  0.022180  0.121105
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   15.637485   2.224250   7.030 8.44e-12 ***
## CountryAlbania                 0.108678   0.043139   2.519 0.012132 *
## CountryAlgeria                -0.047932   0.041096  -1.166 0.244149
## CountryArgentina               0.032558   0.048480   0.672 0.502226
## CountryArmenia                -0.012560   0.035135  -0.357 0.720913
## CountryAustralia               0.388556   0.060613   6.410 3.92e-10 ***
## CountryAustria                 0.274667   0.060776   4.519 8.09e-06 ***
## CountryAzerbaijan             -0.049407   0.042423  -1.165 0.244832
## CountryBahrain                 0.164963   0.055220   2.987 0.002980 **
## CountryBangladesh              0.020461   0.034101   0.600 0.548816
## CountryBelarus                 0.021064   0.044322   0.475 0.634863
## CountryBelgium                 0.178978   0.058397   3.065 0.002319 **
## CountryBenin                  -0.043454   0.032105  -1.354 0.176624
## CountryBolivia                 0.057124   0.041234   1.385 0.166683
## CountryBosnia and Herzegovina  0.186228   0.040312   4.620 5.12e-06 ***
## CountryBotswana               -0.102687   0.041692  -2.463 0.014181 *
## CountryBrazil                  0.045204   0.045072   1.003 0.316470
## CountryBulgaria                0.042933   0.042640   1.007 0.314571
## CountryBurkina Faso           -0.047008   0.027522  -1.708 0.088368 .
## CountryBurundi                -0.146269   0.035878  -4.077 5.47e-05 ***
## CountryCambodia                0.183873   0.035177   5.227 2.73e-07 ***
## CountryCameroon               -0.045532   0.037572  -1.212 0.226251
## CountryCanada                  0.367104   0.061187   6.000 4.29e-09 ***
## CountryChad                   -0.099923   0.026836  -3.724 0.000223 ***
## CountryChile                   0.218343   0.048762   4.478 9.74e-06 ***
## CountryChina                   0.008518   0.044454   0.192 0.848145
```

```
## CountryColombia                0.052172   0.048944   1.066 0.287058
## CountryCosta Rica              0.144750   0.050859   2.846 0.004644 **
## CountryCroatia                 0.118879   0.046029   2.583 0.010142 *
## CountryCyprus                  0.243134   0.052206   4.657 4.31e-06 ***
## CountryDenmark                 0.259126   0.064610   4.011 7.17e-05 ***
## CountryDominican Republic      0.067288   0.045125   1.491 0.136680
## CountryEcuador                 0.048337   0.042994   1.124 0.261534
## CountryEgypt                  -0.028851   0.038895  -0.742 0.458638
## CountryEl Salvador            -0.012064   0.041457  -0.291 0.771196
## CountryEstonia                 0.047389   0.048966   0.968 0.333706
## CountryEthiopia                0.028432   0.028838   0.986 0.324747
## CountryFinland                 0.192059   0.063672   3.016 0.002714 **
## CountryFrance                  0.104879   0.056423   1.859 0.063756 .
## CountryGabon                  -0.078289   0.039388  -1.988 0.047503 *
## CountryGeorgia                -0.061128   0.040250  -1.519 0.129592
## CountryGermany                 0.265731   0.058201   4.566 6.55e-06 ***
## CountryGhana                   0.052590   0.034362   1.530 0.126657
## CountryGreece                 -0.030912   0.046774  -0.661 0.509049
## CountryGuatemala               0.114222   0.040950   2.789 0.005523 **
## CountryGuinea                 -0.015427   0.027612  -0.559 0.576662
## CountryHaiti                   0.208832   0.035100   5.950 5.69e-09 ***
## CountryHonduras                0.096060   0.034668   2.771 0.005840 **
## CountryHungary                 0.030840   0.047184   0.654 0.513725
## CountryIceland                 0.393273   0.061382   6.407 4.01e-10 ***
## CountryIndia                   0.091271   0.033909   2.692 0.007395 **
## CountryIndonesia               0.389823   0.044541   8.752  < 2e-16 ***
## CountryIran                    0.243930   0.042113   5.792 1.37e-08 ***
## CountryIraq                    0.036655   0.038038   0.964 0.335790
## CountryIreland                 0.388326   0.061681   6.296 7.75e-10 ***
## CountryIsrael                  0.300307   0.055480   5.413 1.05e-07 ***
## CountryItaly                   0.171642   0.051343   3.343 0.000903 ***
## CountryJamaica                 0.083930   0.040061   2.095 0.036766 *
## CountryJapan                   0.098626   0.054685   1.804 0.072026 .
## CountryJordan                  0.031350   0.038961   0.805 0.421468
## CountryKazakhstan              0.085142   0.047372   1.797 0.073004 .
## CountryKenya                   0.206970   0.029178   7.093 5.63e-12 ***
## CountryKosovo                  0.164226   0.037918   4.331 1.86e-05 ***
## CountryKuwait                  0.132458   0.056700   2.336 0.019956 *
## CountryKyrgyzstan              0.199634   0.033386   5.980 4.80e-09 ***
## CountryLatvia                  0.070432   0.045842   1.536 0.125195
## CountryLebanon                 0.138956   0.042312   3.284 0.001109 **
## CountryLiberia                -0.017567   0.026943  -0.652 0.514763
## CountryLibya                   0.059714   0.043819   1.363 0.173701
## CountryLithuania              -0.022643   0.047545  -0.476 0.634147
## CountryLuxembourg              0.248852   0.068193   3.649 0.000296 ***
## CountryMadagascar             -0.021184   0.026747  -0.792 0.428812
## CountryMalawi                  0.031657   0.028479   1.112 0.266952
## CountryMalaysia                0.283848   0.047642   5.958 5.43e-09 ***
## CountryMali                   -0.093164   0.027545  -3.382 0.000787 ***
## CountryMalta                   0.417603   0.056616   7.376 8.84e-13 ***
## CountryMauritania             -0.070090   0.029825  -2.350 0.019234 *
## CountryMauritius               0.238890   0.046389   5.150 4.03e-07 ***
## CountryMexico                  0.039686   0.047539   0.835 0.404302
## CountryMoldova                 0.059427   0.035335   1.682 0.093351 .
```

```
## CountryMongolia                0.182955   0.038657   4.733 3.04e-06 ***
## CountryMontenegro              0.083449   0.043271   1.929 0.054466 .
## CountryMorocco                -0.064833   0.037122  -1.747 0.081458 .
## CountryMyanmar                 0.496606   0.032680  15.196  < 2e-16 ***
## CountryNepal                   0.158952   0.031975   4.971 9.73e-07 ***
## CountryNetherlands             0.370976   0.061360   6.046 3.30e-09 ***
## CountryNew Zealand             0.391103   0.061066   6.405 4.06e-10 ***
## CountryNicaragua               0.146070   0.039498   3.698 0.000246 ***
## CountryNiger                  -0.048158   0.036791  -1.309 0.191271
## CountryNigeria                -0.014943   0.032284  -0.463 0.643696
## CountryNorway                  0.292223   0.066762   4.377 1.52e-05 ***
## CountryPakistan                0.095644   0.035836   2.669 0.007906 **
## CountryPanama                  0.130443   0.050369   2.590 0.009940 **
## CountryParaguay                0.141190   0.045252   3.120 0.001933 **
## CountryPeru                    0.031770   0.044312   0.717 0.473809
## CountryPhilippines             0.052961   0.038504   1.375 0.169728
## CountryPoland                  0.089505   0.048966   1.828 0.068275 .
## CountryPortugal                0.088900   0.050113   1.774 0.076791 .
## CountryRomania                 0.058311   0.050254   1.160 0.246570
## CountryRussia                 -0.026041   0.048430  -0.538 0.591071
## CountryRwanda                 -0.014549   0.041772  -0.348 0.727793
## CountrySaudi Arabia            0.056502   0.057604   0.981 0.327222
## CountrySenegal                -0.028824   0.030036  -0.960 0.337787
## CountrySerbia                  0.097129   0.041991   2.313 0.021204 *
## CountrySierra Leone           -0.042427   0.028520  -1.488 0.137610
## CountrySingapore               0.304743   0.069847   4.363 1.62e-05 ***
## CountrySlovakia                0.093780   0.048408   1.937 0.053385 .
## CountrySlovenia                0.195311   0.052094   3.749 0.000202 ***
## CountrySouth Africa           -0.062731   0.037894  -1.655 0.098585 .
## CountrySpain                   0.182517   0.056261   3.244 0.001273 **
## CountrySri Lanka               0.268178   0.043616   6.149 1.83e-09 ***
## CountrySweden                  0.307598   0.062944   4.887 1.46e-06 ***
## CountrySwitzerland             0.273494   0.066845   4.091 5.15e-05 ***
## CountryTajikistan              0.055897   0.032685   1.710 0.087978 .
## CountryTanzania                0.102120   0.029384   3.475 0.000563 ***
## CountryThailand                0.389727   0.048129   8.097 6.20e-15 ***
## CountryTogo                   -0.072436   0.031084  -2.330 0.020263 *
## CountryTunisia                -0.058821   0.039003  -1.508 0.132278
## CountryTurkey                  0.006622   0.045162   0.147 0.883489
## CountryTurkmenistan            0.082581   0.042210   1.956 0.051079 .
## CountryUganda                  0.045096   0.029513   1.528 0.127268
## CountryUkraine                 0.095865   0.037459   2.559 0.010843 *
## CountryUnited Arab Emirates    0.216250   0.066146   3.269 0.001167 **
## CountryUnited Kingdom          0.359976   0.057352   6.277 8.67e-10 ***
## CountryUnited States           0.342936   0.057725   5.941 5.98e-09 ***
## CountryUruguay                 0.118073   0.050449   2.340 0.019731 *
## CountryUzbekistan              0.190683   0.041741   4.568 6.48e-06 ***
## CountryVenezuela              -0.020309   0.042024  -0.483 0.629161
## CountryVietnam                 0.087195   0.042128   2.070 0.039088 *
## CountryYemen                  -0.113175   0.027274  -4.150 4.04e-05 ***
## CountryZambia                  0.014807   0.029642   0.500 0.617666
## CountryZimbabwe               -0.067558   0.028631  -2.360 0.018753 *
## happiness_score                0.003367   0.007706   0.437 0.662414
## gdp_per_capita                -0.098764   0.033525  -2.946 0.003400 **
```

```
## health                         -0.233851   0.027612  -8.469 4.25e-16 ***
## freedom                        -0.069877   0.034879  -2.003 0.045778 *
## government_trust                0.130658   0.062559   2.089 0.037350 *
## Year                           -0.007581   0.001103  -6.873 2.30e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04103 on 418 degrees of freedom
## Multiple R-squared:   0.91,  Adjusted R-squared:  0.8805
## F-statistic: 30.85 on 137 and 418 DF,  p-value: < 2.2e-16
```

```
test$predicted_generosity <- predict(gen_model, test)
head(test[, c("generosity", "predicted_generosity")])
```

```
##    generosity predicted_generosity
## 2   0.3552805            0.3235025
## 3   0.4755402            0.4157911
## 4   0.2905493            0.3102894
## 13  0.3164723            0.3121059
## 19  0.4927742            0.4132195
## 21  0.3609419            0.2701128
```

```
real <- test$generosity
prediction <- test$predicted_generosity
rss <- sum((prediction - real) ^ 2)
tss <- sum((real - mean(real)) ^ 2)
rsq <- 1 - rss/tss
rsq
```

```
## [1] 0.8707545
```

Here I did linear models on all of the factors that effect the happiness scores. we can also see here as wel as previously visualized that gdp and health have the greatest impact on the happiness score for a specific country

```
family_model <- lm(family ~ happiness_score, dat = WorldHappiness_df) #.076
family_model
```

```
##
## Call:
## lm(formula = family ~ happiness_score, data = WorldHappiness_df)
##
## Coefficients:
##     (Intercept)  happiness_score
##         0.08822          0.07615
```

```
govTrust_md <-lm(government_trust ~ happiness_score, dat = WorldHappiness_df) #0.044
govTrust_md
```

```
##
## Call:
## lm(formula = government_trust ~ happiness_score, data = WorldHappiness_df)
##
## Coefficients:
##     (Intercept)  happiness_score
##        -0.11595          0.04415
```

```r
health_model <- lm(health ~ happiness_score, dat = WorldHappiness_df) #0.1556
health_model
```

```
##
## Call:
## lm(formula = health ~ happiness_score, data = WorldHappiness_df)
##
## Coefficients:
##     (Intercept)  happiness_score
##         -0.2030           0.1556
```

```r
freedom_model <- lm(freedom ~ happiness_score, dat = WorldHappiness_df) #0.07162
freedom_model
```

```
##
## Call:
## lm(formula = freedom ~ happiness_score, data = WorldHappiness_df)
##
## Coefficients:
##     (Intercept)  happiness_score
##         0.03500          0.07162
```

```r
generosity_model <- lm(generosity ~ happiness_score, dat = WorldHappiness_df) #0.01681
generosity_model
```

```
##
## Call:
## lm(formula = generosity ~ happiness_score, data = WorldHappiness_df)
##
## Coefficients:
##     (Intercept)  happiness_score
##         0.12043          0.01681
```

```r
corr_health = cor.test(WorldHappiness_df$health,WorldHappiness_df$happiness_score, method = "pearson")
corr_health
```

```
##
##  Pearson's product-moment correlation
##
## data:  WorldHappiness_df$health and WorldHappiness_df$happiness_score
## t = 32.217, df = 790, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7217587 0.7821403
## sample estimates:
##       cor
## 0.7535343
```

```r
corr_gdp = cor.test(WorldHappiness_df$gdp_per_capita,WorldHappiness_df$happiness_score, method = "pearso
corr_health
```

```
##
##  Pearson's product-moment correlation
##
## data:  WorldHappiness_df$health and WorldHappiness_df$happiness_score
## t = 32.217, df = 790, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
```

```
## 95 percent confidence interval:
##  0.7217587 0.7821403
## sample estimates:
##       cor
## 0.7535343
```

```
corr_gen = cor.test(WorldHappiness_df$generosity, WorldHappiness_df$happiness_score, method = "pearson")
corr_gen
```

```
##
##  Pearson's product-moment correlation
##
## data:  WorldHappiness_df$generosity and WorldHappiness_df$happiness_score
## t = 4.4221, df = 790, p-value = 1.114e-05
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.08669416 0.22267173
## sample estimates:
##       cor
## 0.1554191
```

Citations: https://www.r-bloggers.com/2020/05/step-by-step-guide-on-how-to-build-linear-regression-in-r-with-code/ https://www.scribbr.com/statistics/linear-regression-in-r/ https://www.kaggle.com/datasets/eliasturk/world-happiness-based-on-cpi-20152020/code?resource=download “‘

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.