# MCM 2022: Problem F

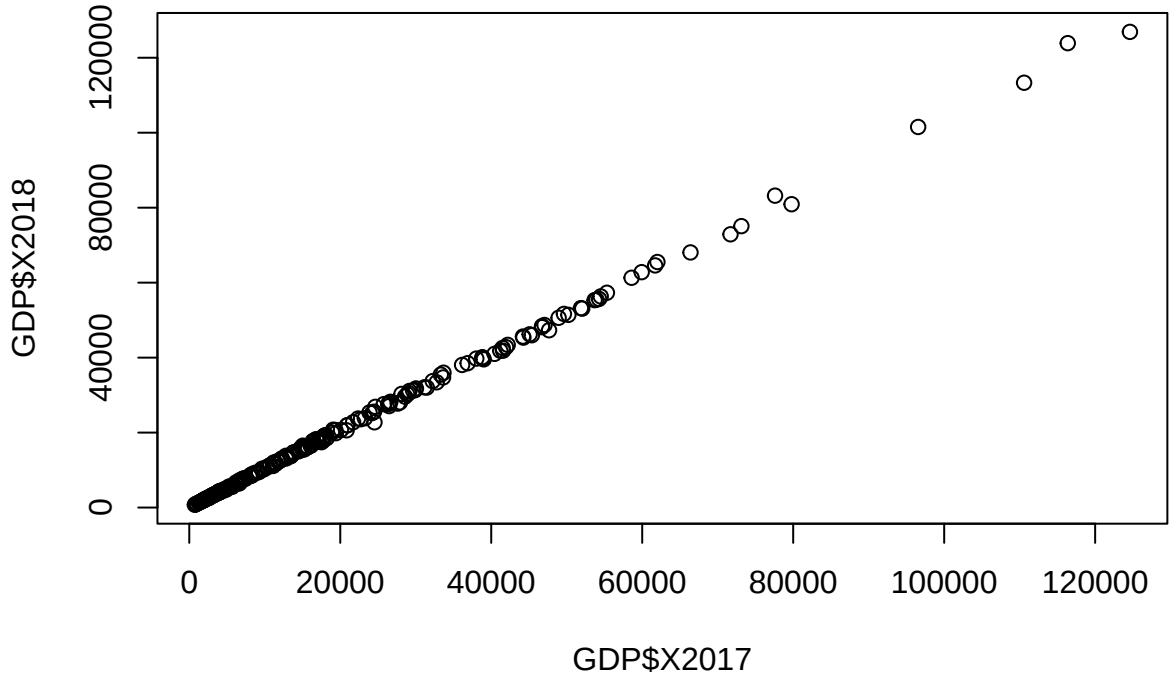
Katherine, CJ

2/18/2022

```
GDP = read.csv(file = "GDP.csv", header = TRUE, sep = ",")
summary(GDP$X2018)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.      NA's
##     744.2    4953.2   14623.7   21134.6   29900.0 126898.4        29
```

```
plot(GDP$X2017, GDP$X2018)
```




```
GDP2018 = GDP$X2018
USGDP =  subset(GDP, Country == 'United States')
UKGDP =  subset(GDP, Country == 'United Kingdom')
ChinaGDP =  subset(GDP, Country == 'China')
MexicoGDP =  subset(GDP, Country == 'Mexico')

UniRank = read.csv(file = "2020-QS-World-University-Rankings.csv", header = TRUE)
table(UniRank$Country) ## HARD
```

```
##
##                          Argentina          Australia
##                  1               13                 35
##            Austria          Azerbaijan            Bahrain
##                  8                1                  1
```

```
##           Bangladesh              Belarus              Belgium
##                    2                    2                    8
##               Brazil               Brunei             Bulgaria
##                   19                    2                    1
##               Canada                Chile                China
##                   26                   11                   42
##             Colombia           Costa Rica              Croatia
##                   12                    3                    1
##                 Cuba       Czech Republic              Denmark
##                    1                   10                    6
##              Ecuador                Egypt              Estonia
##                    4                    5                    3
##              Finland               France              Germany
##                    9                   38                   47
##               Greece            Hong Kong              Hungary
##                    6                    7                    7
##                India            Indonesia                 Iran
##                   25                    9                    6
##                 Iraq              Ireland               Israel
##                    3                    8                    7
##                Italy                Japan               Jordan
##                   35                   41                    4
##           Kazakhstan               Kuwait               Latvia
##                   10                    1                    3
##              Lebanon            Lithuania                Macau
##                    7                    4                    1
##             Malaysia               Mexico          Netherlands
##                   20                   13                   13
##          New Zealand               Norway                 Oman
##                    8                    5                    1
##             Pakistan               Panama             Paraguay
##                    7                    1                    1
##                 Peru          Philippines               Poland
##                    3                    4                   16
##             Portugal          Puerto Rico                Qatar
##                    7                    1                    1
##              Romania               Russia         Saudi Arabia
##                    2                   25                    8
##               Serbia            Singapore             Slovakia
##                    1                    3                    3
##             Slovenia         South Africa          South Korea
##                    2                    8                   30
##                Spain            Sri Lanka               Sweden
##                   27                    1                   10
##          Switzerland               Taiwan             Thailand
##                    9                   16                    8
##               Turkey              Ukraine United Arab Emirates
##                    9                    6                    8
##       United Kingdom        United States              Uruguay
##                   86                  159                    2
##            Venezuela              Vietnam
##                    4                    2
```

```
HIndex2021 = read.csv("world-happiness-report-2021.csv", header = TRUE, sep = ",")
summary(HIndex2021$Ladder.score)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.523   4.852   5.534   5.533   6.255   7.842
```

```
USHIndex =  subset(HIndex2021, Country.name == 'United States')
UKHIndex =  subset(HIndex2021, Country.name == 'United Kingdom')
ChinaHIndex =  subset(HIndex2021, Country.name == 'China')
MexicoHIndex =  subset(HIndex2021, Country.name == 'Mexico')

SpaceProgram = read.csv('MathModelingSpacePrograms[2022data].xlsx - Sheet1.csv', header = TRUE, sep = "
summary(SpaceProgram$Government.Expenditures.for.Space.Programs...10s.millions.USD.)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    10.0    28.0    72.0  1435.6   206.2 54589.0
```

```
USSpaceProgram =  subset(SpaceProgram, Country.== 'USA')
UKSpaceProgram =  subset(SpaceProgram, Country.== 'UK')
ChinaSpaceProgram =  subset(SpaceProgram, Country.== 'China')
MexicoSpaceProgram = 0

WorldPop = read.csv("WorldPopulation1960-2020.csv", header = TRUE, sep = ",")
summary(WorldPop$X2018)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.      NA's
## 1.068e+04 1.740e+06 1.046e+07 3.102e+08 6.193e+07 7.592e+09         2
```

```
USAWP = subset(WorldPop, Country.Name == 'United States')
UKWP = subset(WorldPop, Country.Name == 'United Kingdom')
ChinaWP = subset(WorldPop, Country.Name == 'China')
MexicoWP = subset(WorldPop, Country.Name == 'Mexico')

GiniC = read.csv("GiniCoefficient 2020 - Sheet1.csv", header = TRUE, sep = ",")
summary(GiniC$Latest)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2220  0.2807  0.3090  0.3266  0.3458  0.6180
```

```
USGini = subset(GiniC, Country == 'United States')
UKGini = subset(GiniC, Country == 'United Kingdom')
ChinaGini = 0
MexicoGini = subset(GiniC, Country == 'Mexico')
```

Our skeleton equation to measure global equity: $Score = aq[GDP] + br[university : populationratio] + cs[spaceprogrambudget] + dt[HI] + eu[Gini]$

Example test

```
## remove or mess around with population and space prog
USScore =  (USGDP$X2018*(10^(-3)*(4/11)) + 159/USAWP$X2018*(10^(7))*(2/11) + USSpaceProgram$Government.

UKScore =  (UKGDP$X2018*(10^(-3)*(4/11)) + 86/UKWP$X2018*(10^(7))*(2/11) + UKSpaceProgram$Government.Exp

ChinaScore =  (ChinaGDP$X2018*(10^(-3)*(4/11)) + 42/ChinaWP$X2018*(10^(7))*(2/11) + ChinaSpaceProgram$G

MexicoScore =  (MexicoGDP$X2018*(10^(-3)*(4/11)) + 13/MexicoWP$X2018*(10^(7))*(2/11) + ChinaSpaceProgra
```

```
USScore
```

```
## [1] 26.49708
```

```
UKScore
```

```
## [1] 21.45498
```

```
ChinaScore
```

```
## [1] 7.265199
```

```
MexicoScore
```

```
## [1] 9.658596
```

Generate validations

```
data = read.csv("MCM Full Data - Sheet1.csv", header = TRUE, sep = ",")

N = 39
result = numeric(N)
for (i in 1:N) {
  result[i] = data$GDP2018[i]*(10^(-3)*(4/11)) + (data$TotalSRank[i]/data$WP2020[i]*(10^(7)*(2/11)) + da
}

result
```

```
##  [1] 21.963693 23.205300 21.291338  9.823147 20.216850 12.033861  9.494698
##  [8] 17.461846 23.525291 18.573733 22.012013 18.953832 21.786676 13.139426
## [15] 13.940385 22.206022 34.608671 17.497542 17.782440 17.644332 16.586578
## [22] 15.374968 17.039641 42.712926  9.113831 23.325795 26.897511 13.489244
## [29] 14.771050 11.928714 11.742356 16.825810  7.358844 16.945156 22.538980
## [36] 28.182799 11.933425 20.697646 25.917245
```

```
## Create subsets
EURegion = subset(data, Region == 'Europe')
AMERegion = subset(data, Region == 'America')
MERegion = subset(data, Region == 'Middle East')

summary(EURegion$GE.Index)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   9.823  16.100  18.954  20.335  22.872  42.713
```

```
sd(EURegion$GE.Index)
```

```
## [1] 7.020712
```

```
summary(AMERegion$GE.Index)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   9.114  10.129  16.125  16.457  21.527  25.917
```

```
sd(AMERegion$GE.Index)
```

```
## [1] 7.154248
```

```
summary(MERegion$GE.Index)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   11.93   13.32   14.72   14.72   16.11   17.50
```

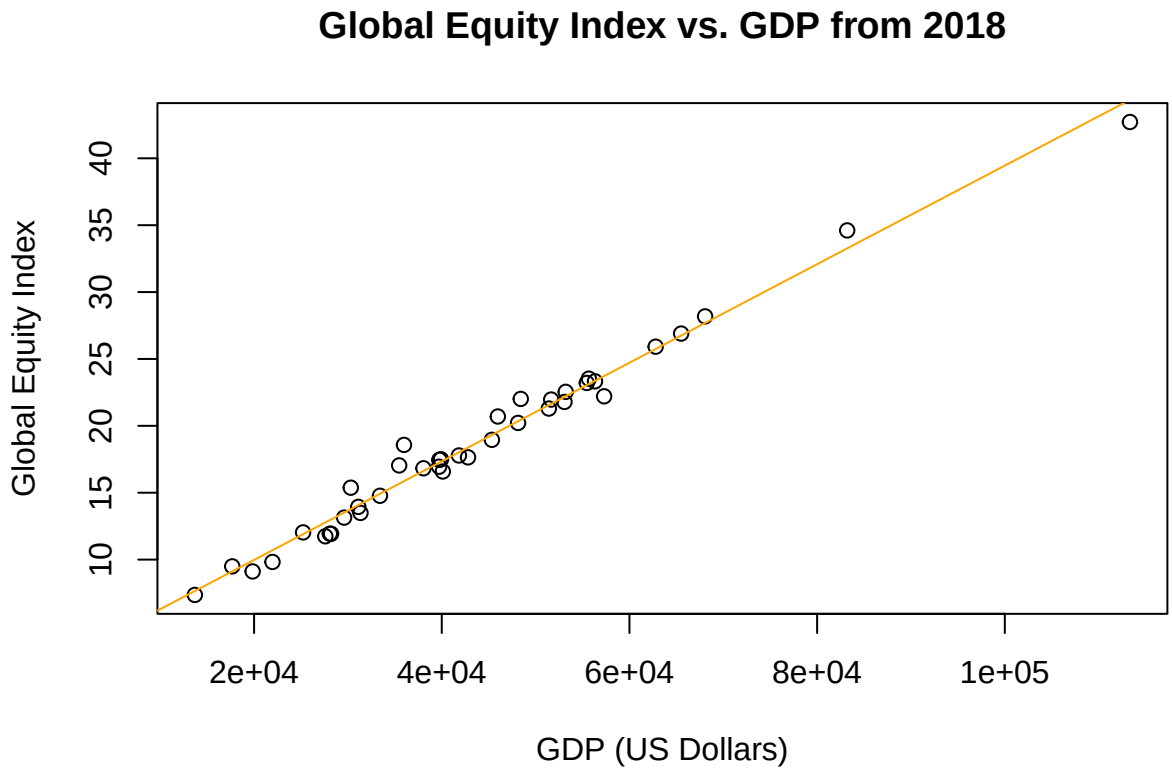
```r
sd(MERegion$GE.Index)
```

```
## [1] 3.934425
```

```r
plot(data$GDP2018, data$GE.Index, xlab = "GDP (US Dollars)", ylab = "Global Equity Index", main = "Globa
lin.reg1 = lm(data$GE.Index ~ data$GDP2018)
lin.reg1
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$GDP2018)
##
## Coefficients:
##  (Intercept)  data$GDP2018
##    2.6076441     0.0003685
```

```r
abline(lin.reg1, col="orange")
```




```r
cor(data$GDP2018, data$GE.Index)
```

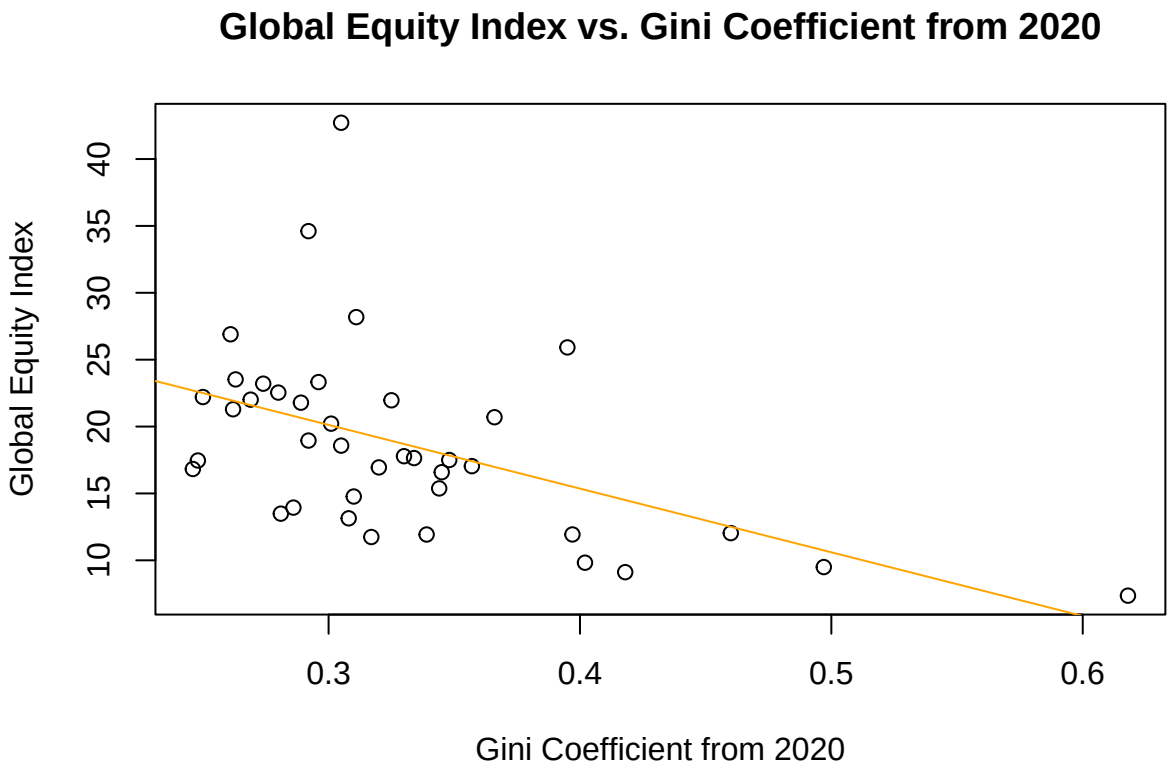
```
## [1] 0.9918337
```

```r
plot(data$Gini2020, data$GE.Index, xlab = "Gini Coefficient from 2020", ylab = "Global Equity Index", ma
lin.reg2 = lm(data$GE.Index ~ data$Gini2020)
lin.reg2
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$Gini2020)
##
## Coefficients:
##   (Intercept)  data$Gini2020
```

```
##         34.41         -47.63
```

```
abline(lin.reg2, col="orange")
```



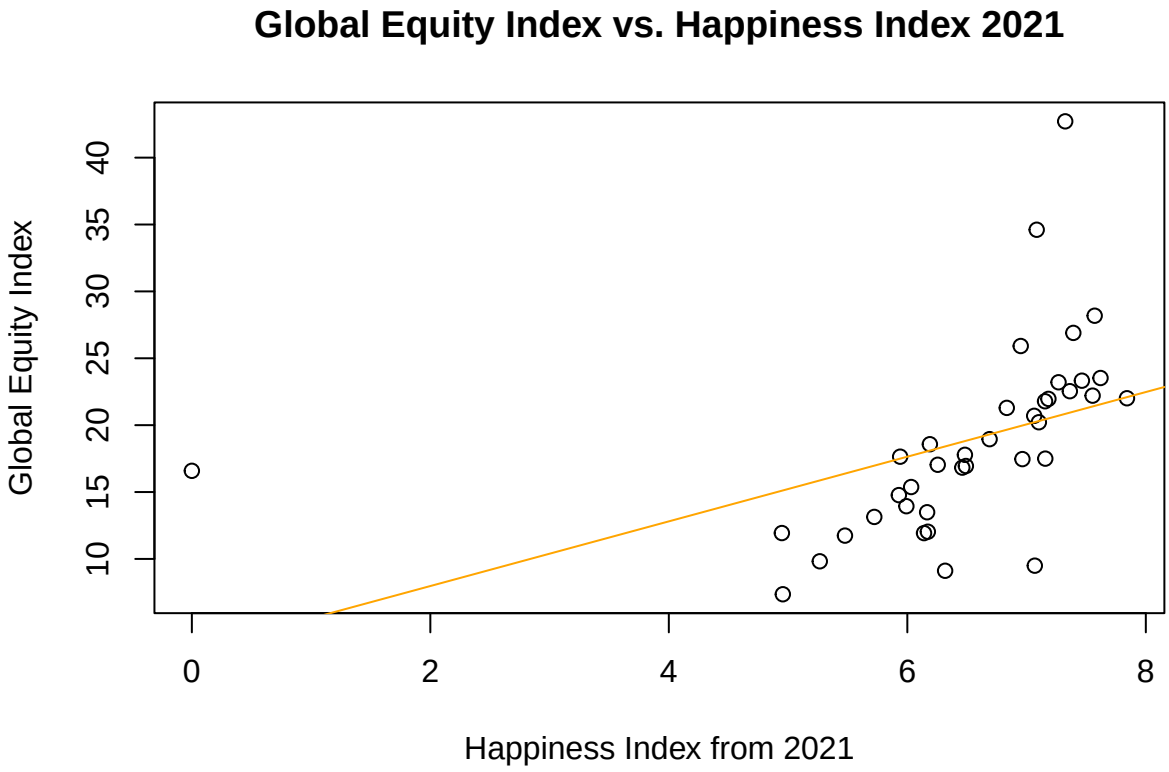

```
cor(data$Gini2020, data$GE.Index)
```

```
## [1] -0.5039276
```

```
plot(data$HI12021, data$GE.Index, xlab = "Happiness Index from 2021", ylab = "Global Equity Index", mai
lin.reg3 = lm(data$GE.Index ~ data$HI12021)
lin.reg3
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$HI12021)
##
## Coefficients:
##  (Intercept)  data$HI12021
##        3.139         2.417
```

```
abline(lin.reg3, col="orange")
```



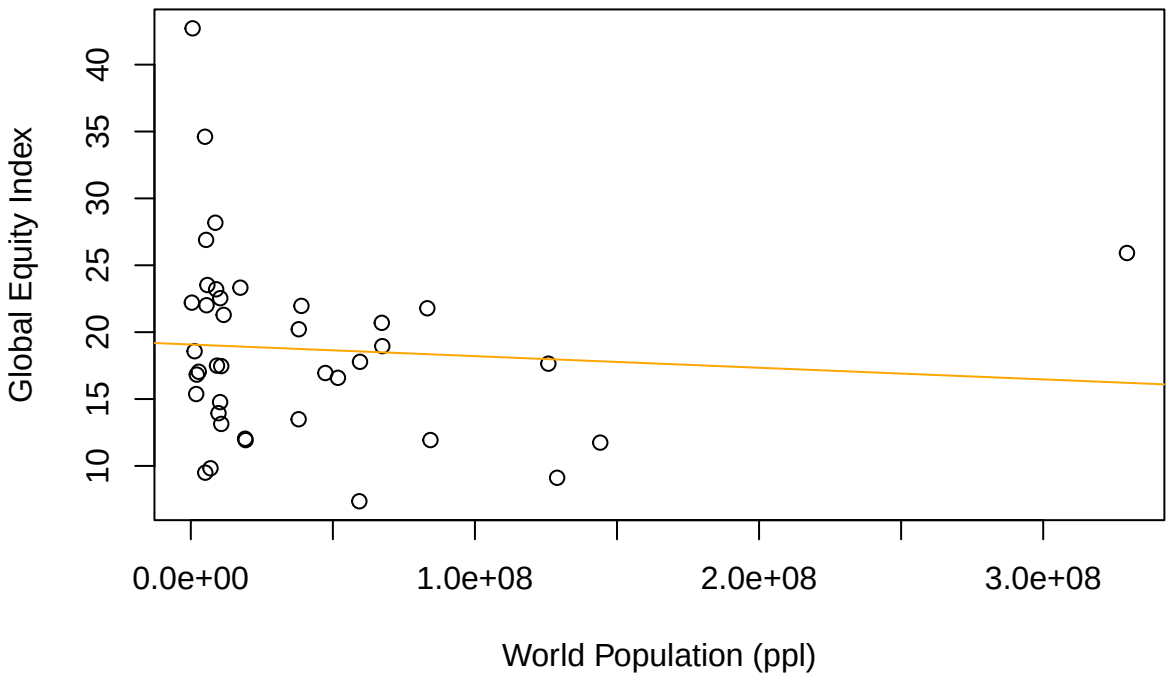

```
cor(data$HI12021, data$GE.Index)
```

```
## [1] 0.4474737
```

```
plot(data$WP2020, data$GE.Index, xlab = "World Population (ppl)", ylab = "Global Equity Index", main =
lin.reg4 = lm(data$GE.Index ~ data$WP2020)
lin.reg4
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$WP2020)
##
## Coefficients:
## (Intercept)  data$WP2020
##   1.908e+01   -8.694e-09
```

```
abline(lin.reg4, col="orange")
```



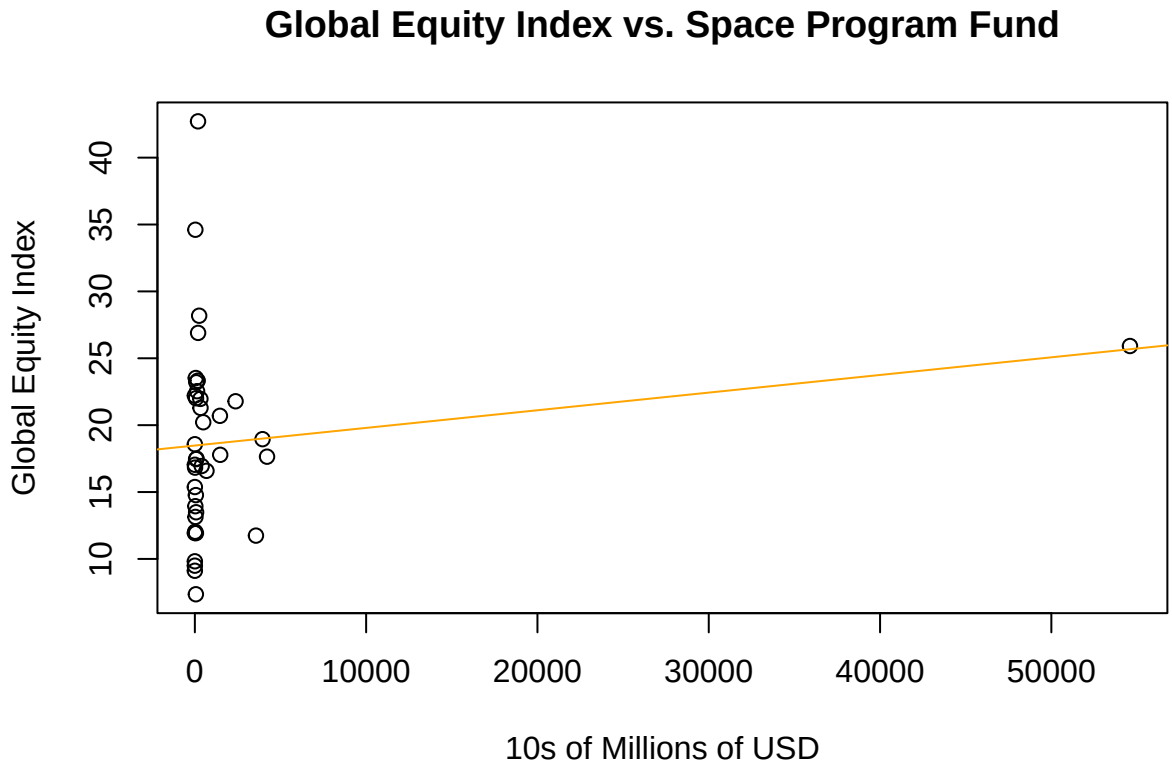

```
cor(data$WP2020, data$GE.Index)
```

```
## [1] -0.07567746
```

```
plot(data$SPFund, data$GE.Index, xlab = "10s of Millions of USD", ylab = "Global Equity Index", main =
lin.reg5 = lm(data$GE.Index ~ data$SPFund)
lin.reg5
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$SPFund)
##
## Coefficients:
## (Intercept)  data$SPFund
##   1.848e+01    1.319e-04
```

```
abline(lin.reg5, col="orange")
```




```
cor(data$SPFund, data$GE.Index)
```

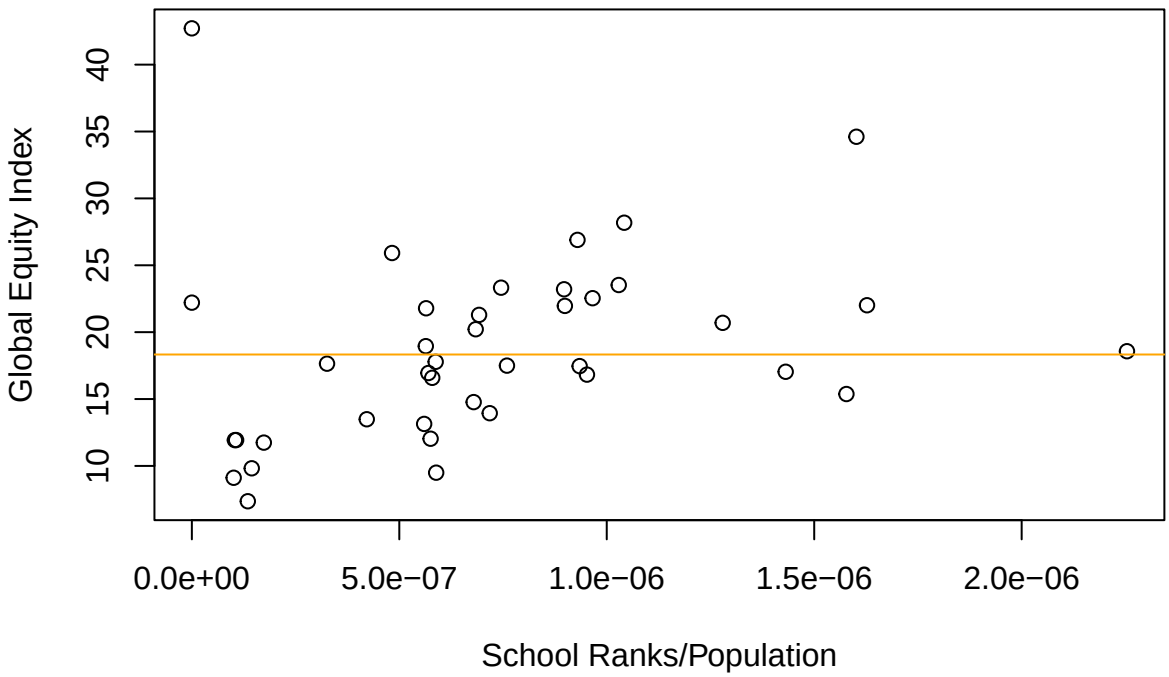
```
## [1] 0.1640307
```

```
plot(data$TotalSRank/data$WP2020, data$GE.Index, xlab = "School Ranks/Population", ylab = "Global Equity
lin.reg6 = lm(data$GE.Index ~ data$TotalSRank/data$WP2020)
lin.reg6
```

```
##
## Call:
## lm(formula = data$GE.Index ~ data$TotalSRank/data$WP2020)
##
## Coefficients:
##                   (Intercept)               data$TotalSRank
##                     1.833e+01                     1.003e-02
## data$TotalSRank:data$WP2020
##                     9.663e-11
```

```
abline(lin.reg6, col="orange")
```

```
## Warning in abline(lin.reg6, col = "orange"): only using the first two of 3
## regression coefficients
```




```
cor(data$TotalSRank/data$WP2020, data$GE.Index)
```

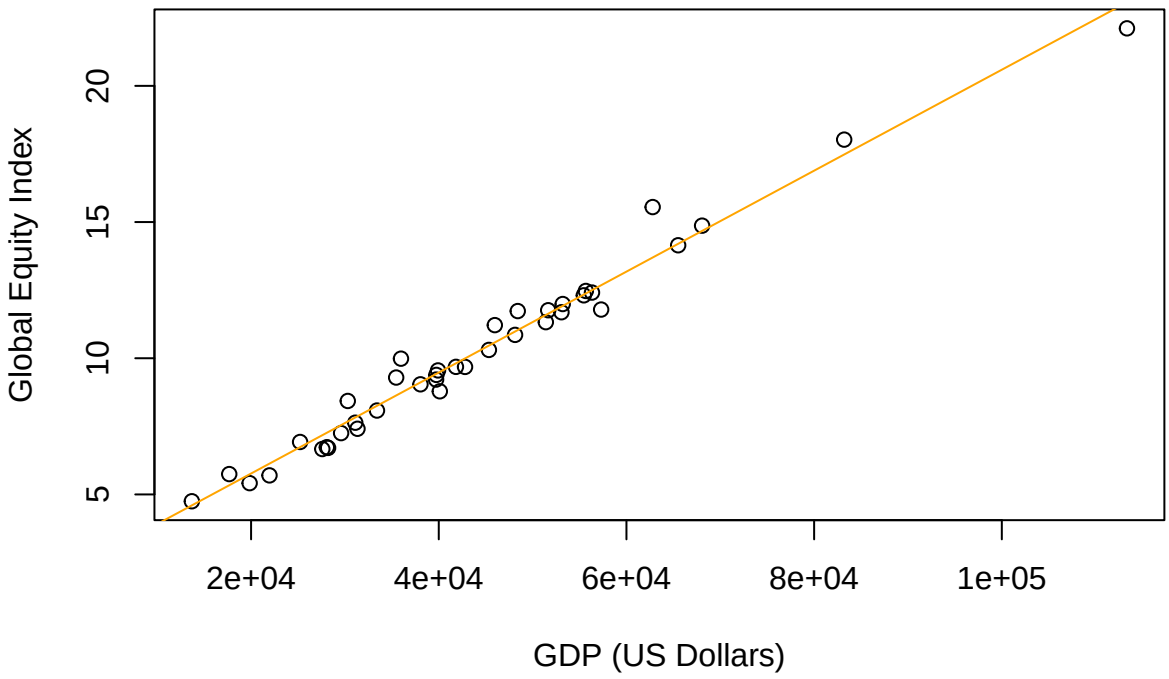
```
## [1] 0.2641452
```

Weight on Space Program Fund

```
plot(data$GDP2018, data$GE.IndexSpace, xlab = "GDP (US Dollars)", ylab = "Global Equity Index", main =
lin.reg1 = lm(data$GE.IndexSpace ~ data$GDP2018)
lin.reg1
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$GDP2018)
##
## Coefficients:
##  (Intercept)  data$GDP2018
##    2.0615898     0.0001853
```

```
abline(lin.reg1, col="orange")
```



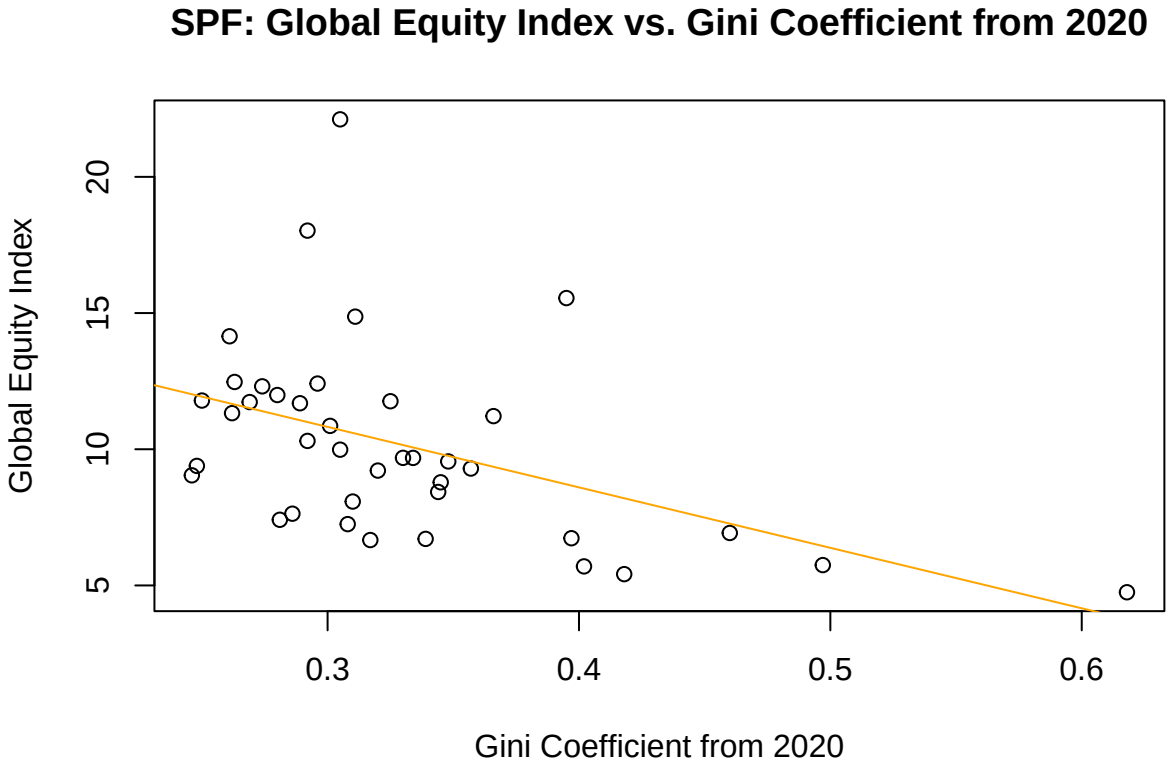

```
cor(data$GDP2018, data$GE.IndexSpace)
```

```
## [1] 0.9877922
```

```
plot(data$Gini2020, data$GE.IndexSpace, xlab = "Gini Coefficient from 2020", ylab = "Global Equity Inde
lin.reg2 = lm(data$GE.IndexSpace ~ data$Gini2020)
lin.reg2
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$Gini2020)
##
## Coefficients:
##   (Intercept)  data$Gini2020
##         17.48         -22.20
```

```
abline(lin.reg2, col="orange")
```



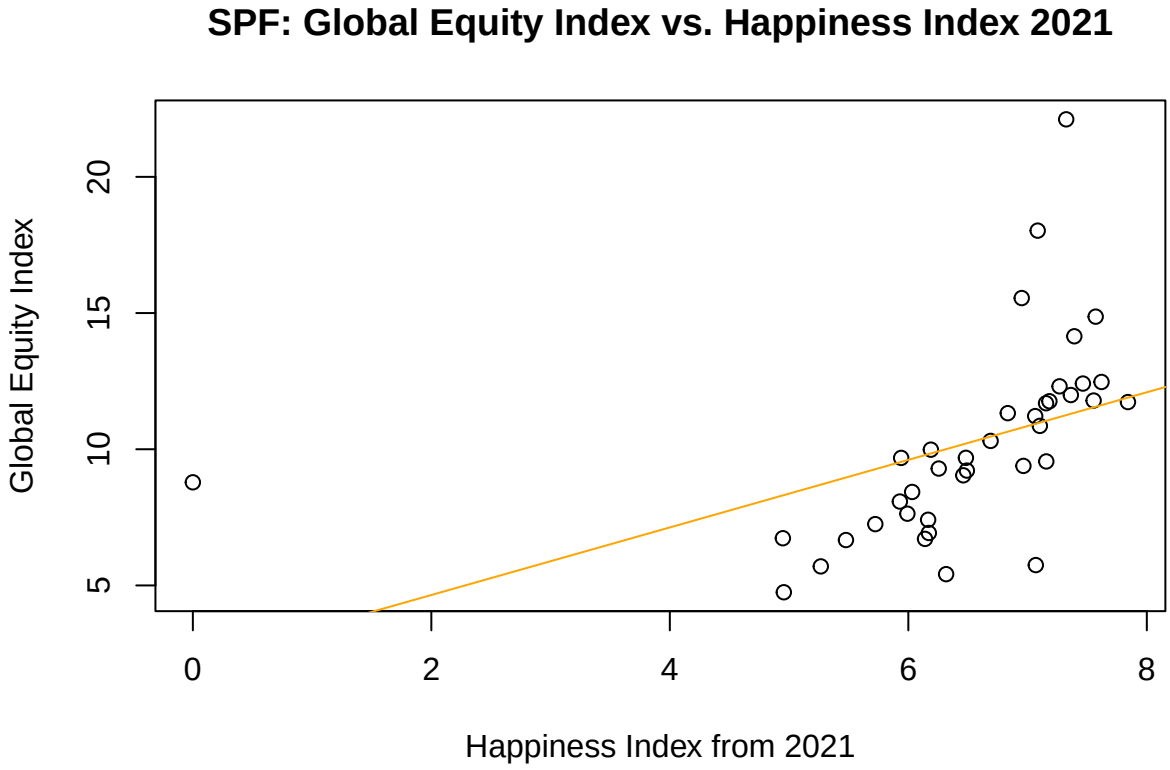

```
cor(data$Gini2020, data$GE.IndexSpace)
```

```
## [1] -0.4652172
```

```
plot(data$HI12021, data$GE.IndexSpace, xlab = "Happiness Index from 2021", ylab = "Global Equity Index"
lin.reg3 = lm(data$GE.IndexSpace ~ data$HI12021)
lin.reg3
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$HI12021)
##
## Coefficients:
##  (Intercept)  data$HI12021
##        2.168         1.240
```

```
abline(lin.reg3, col="orange")
```



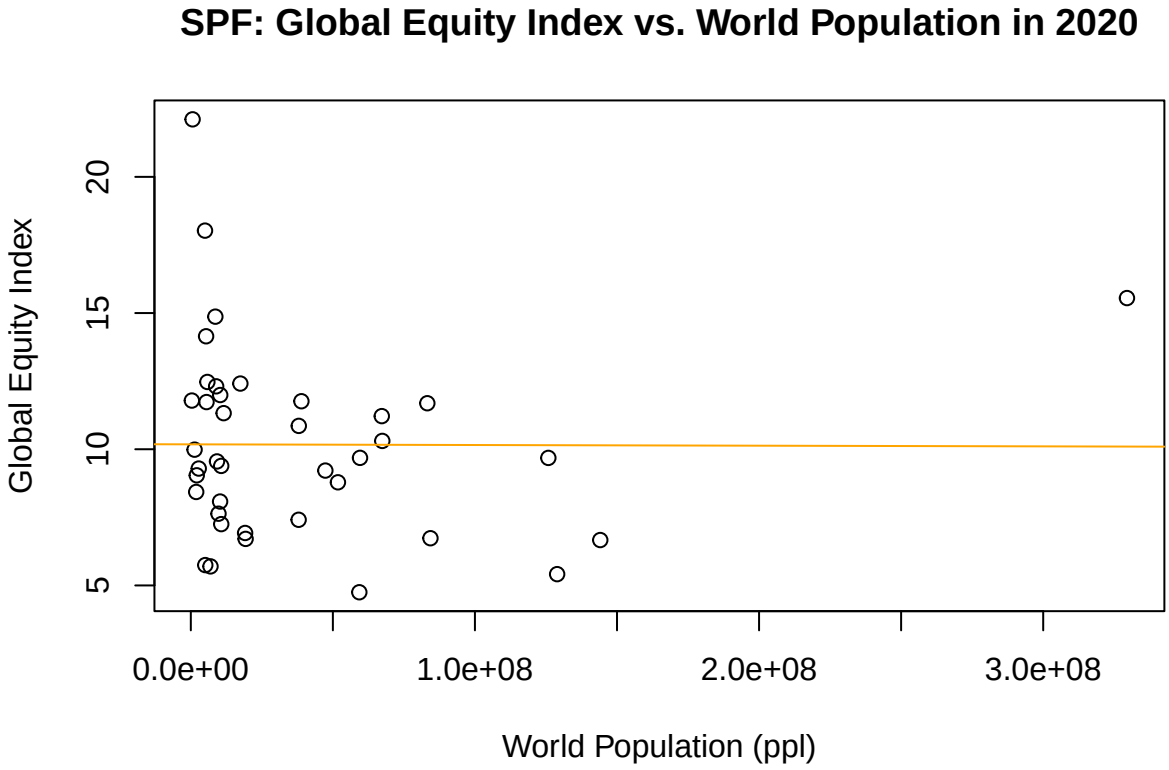

```
cor(data$HI12021, data$GE.IndexSpace)
```

```
## [1] 0.4547977
```

```
plot(data$WP2020, data$GE.IndexSpace, xlab = "World Population (ppl)", ylab = "Global Equity Index", ma
lin.reg4 = lm(data$GE.IndexSpace ~ data$WP2020)
lin.reg4
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$WP2020)
##
## Coefficients:
## (Intercept)  data$WP2020
##   1.018e+01   -2.490e-10
```

```
abline(lin.reg4, col="orange")
```



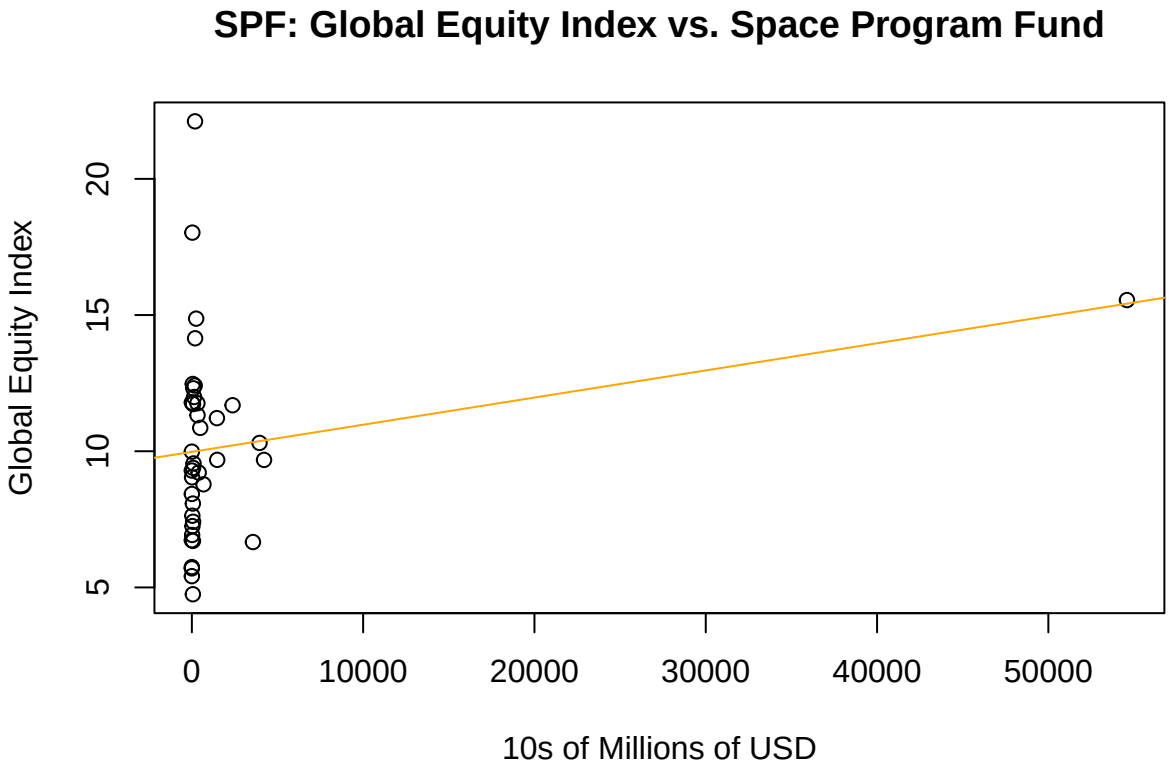

```
cor(data$WP2020, data$GE.IndexSpace)
```

```
## [1] -0.004292759
```

```
plot(data$SPFund, data$GE.IndexSpace, xlab = "10s of Millions of USD", ylab = "Global Equity Index", ma
lin.reg3 = lm(data$GE.IndexSpace ~ data$SPFund)
lin.reg3
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$SPFund)
##
## Coefficients:
## (Intercept)  data$SPFund
##   9.977e+00    9.969e-05
```

```
abline(lin.reg3, col="orange")
```




```
cor(data$SPFund, data$GE.IndexSpace)
```

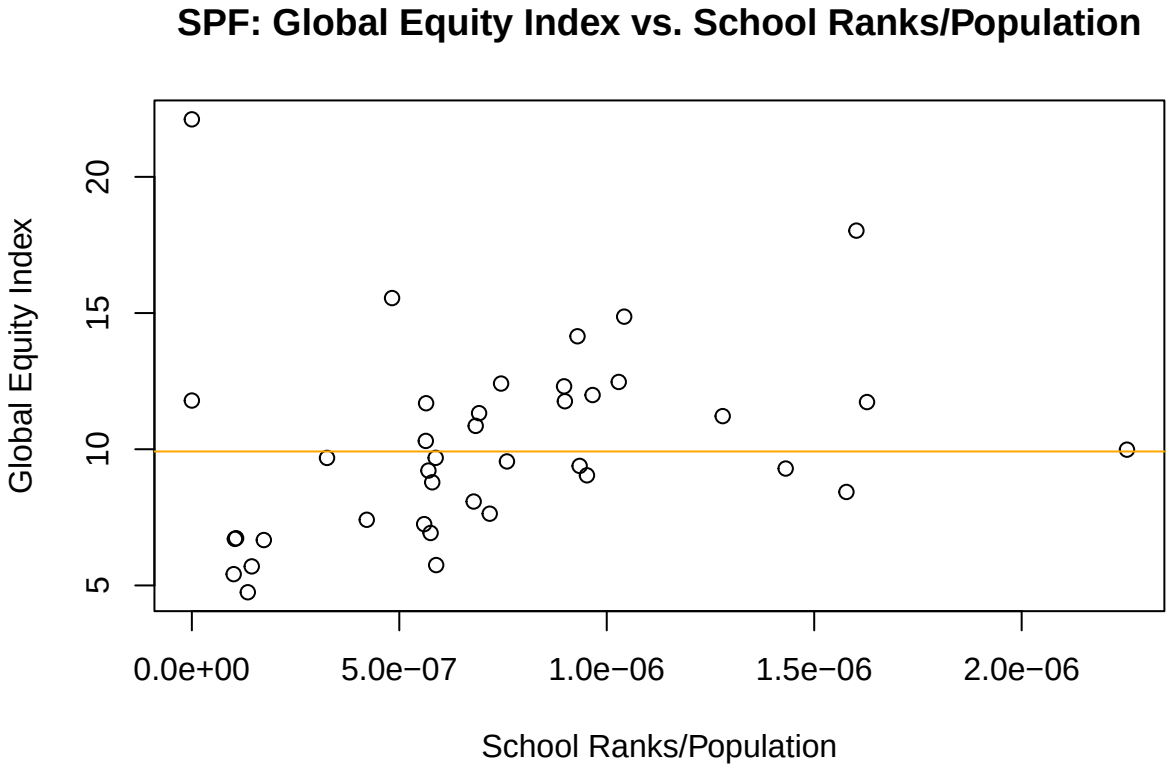
```
## [1] 0.2455124
```

```
plot(data$TotalSRank/data$WP2020, data$GE.IndexSpace, xlab = "School Ranks/Population", ylab = "Global I
lin.reg6 = lm(data$GE.IndexSpace ~ data$TotalSRank/data$WP2020)
lin.reg6
```

```
##
## Call:
## lm(formula = data$GE.IndexSpace ~ data$TotalSRank/data$WP2020)
##
## Coefficients:
##                   (Intercept)              data$TotalSRank
##                     9.918e+00                    2.918e-03
## data$TotalSRank:data$WP2020
##                     8.979e-11
```

```
abline(lin.reg6, col="orange")
```

```
## Warning in abline(lin.reg6, col = "orange"): only using the first two of 3
## regression coefficients
```




```
cor(data$TotalSRank/data$WP2020, data$GE.IndexSpace)
```

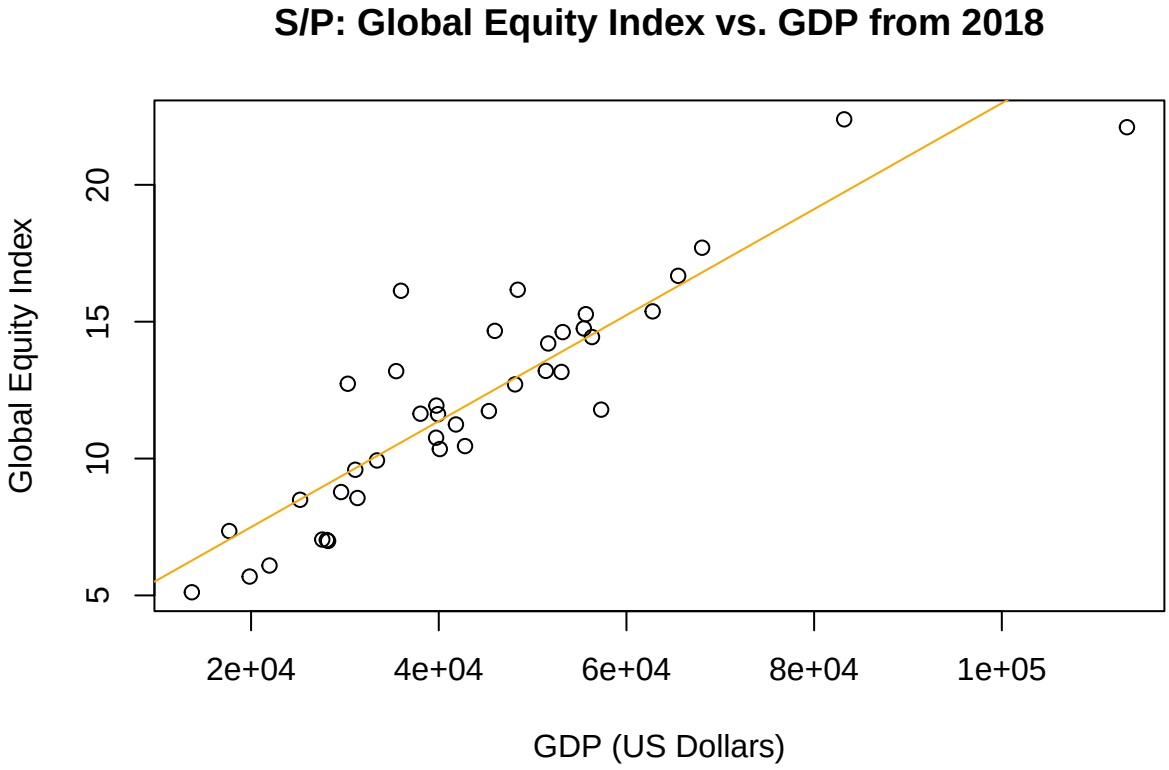
```
## [1] 0.2485563
```

Weight on Schools / Population

```
plot(data$GDP2018, data$GE.IndexSchool, xlab = "GDP (US Dollars)", ylab = "Global Equity Index", main =
lin.reg1 = lm(data$GE.IndexSchool ~ data$GDP2018)
lin.reg1
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$GDP2018)
##
## Coefficients:
##  (Intercept)  data$GDP2018
##    3.6233998     0.0001936
```

```
abline(lin.reg1, col="orange")
```



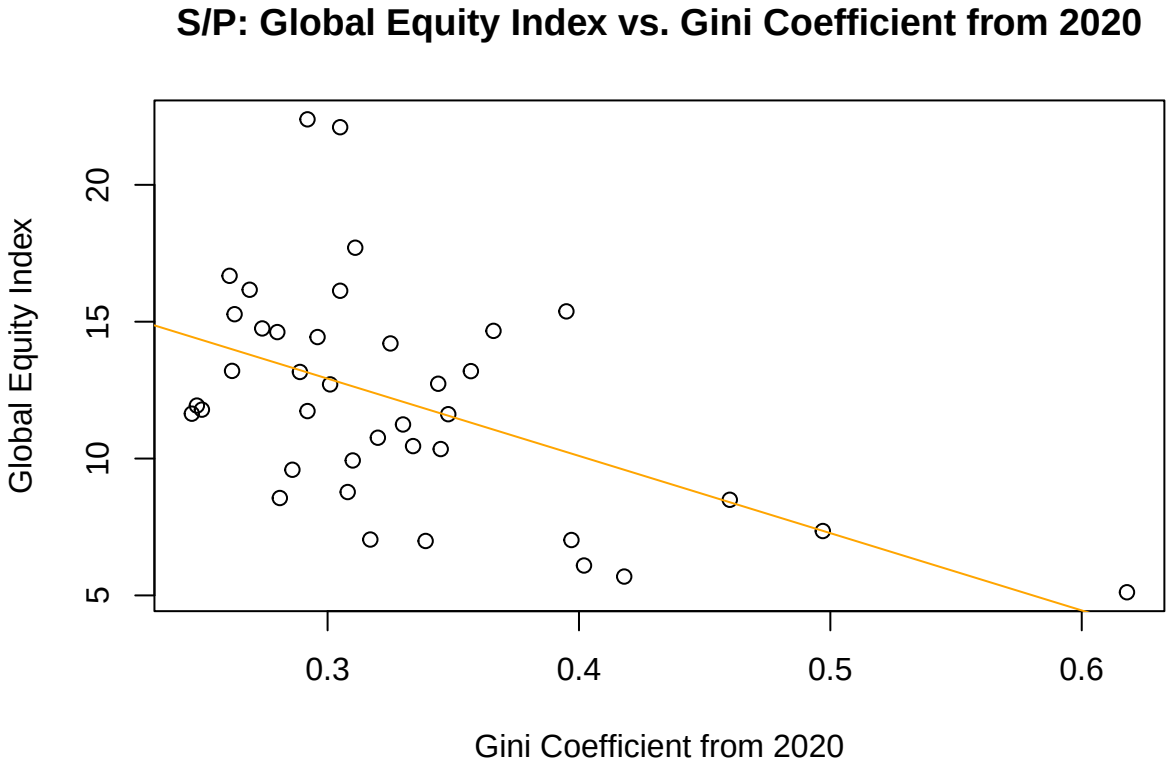

```
cor(data$GDP2018, data$GE.IndexSchool)
```

```
## [1] 0.898549
```

```
plot(data$Gini2020, data$GE.IndexSchool, xlab = "Gini Coefficient from 2020", ylab = "Global Equity Inde
lin.reg2 = lm(data$GE.IndexSchool ~ data$Gini2020)
lin.reg2
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$Gini2020)
##
## Coefficients:
##   (Intercept)  data$Gini2020
##         21.39         -28.23
```

```
abline(lin.reg2, col="orange")
```



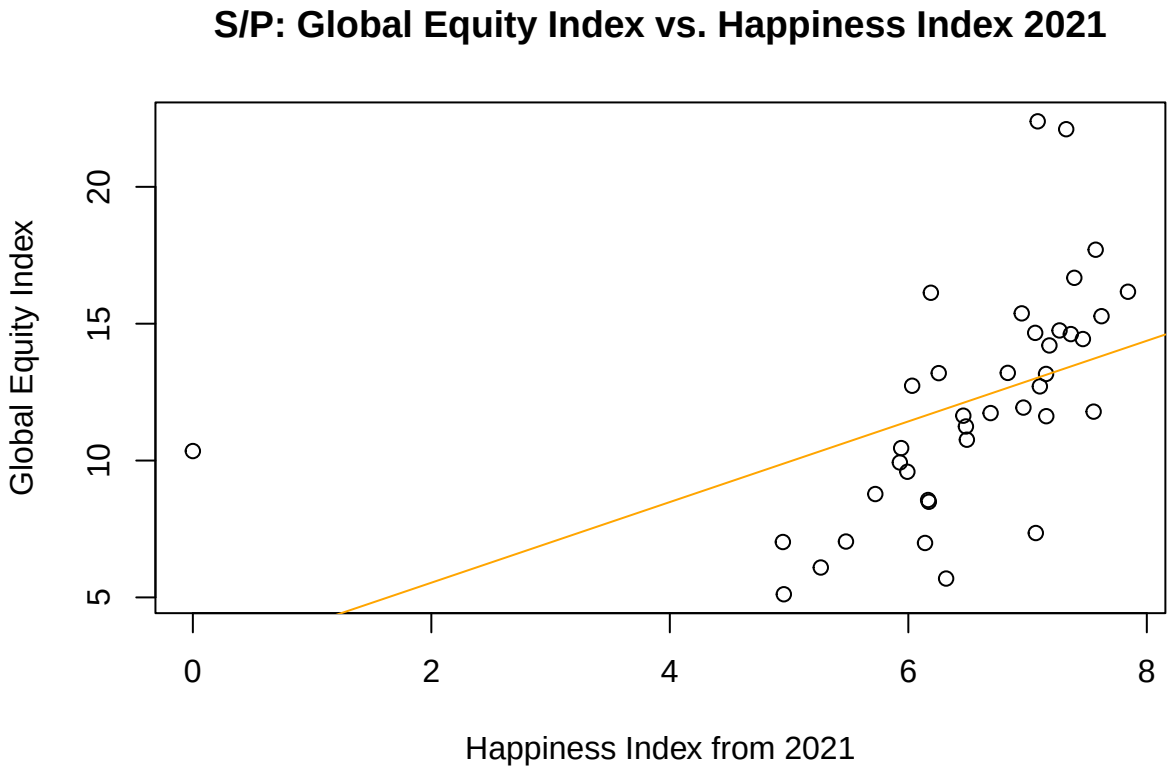

```
cor(data$Gini2020, data$GE.IndexSchool)
```

```
## [1] -0.5150653
```

```
plot(data$HI12021, data$GE.IndexSchool, xlab = "Happiness Index from 2021", ylab = "Global Equity Index
lin.reg3 = lm(data$GE.IndexSchool ~ data$HI12021)
lin.reg3
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$HI12021)
##
## Coefficients:
##  (Intercept)  data$HI12021
##        2.591         1.473
```

```
abline(lin.reg3, col="orange")
```



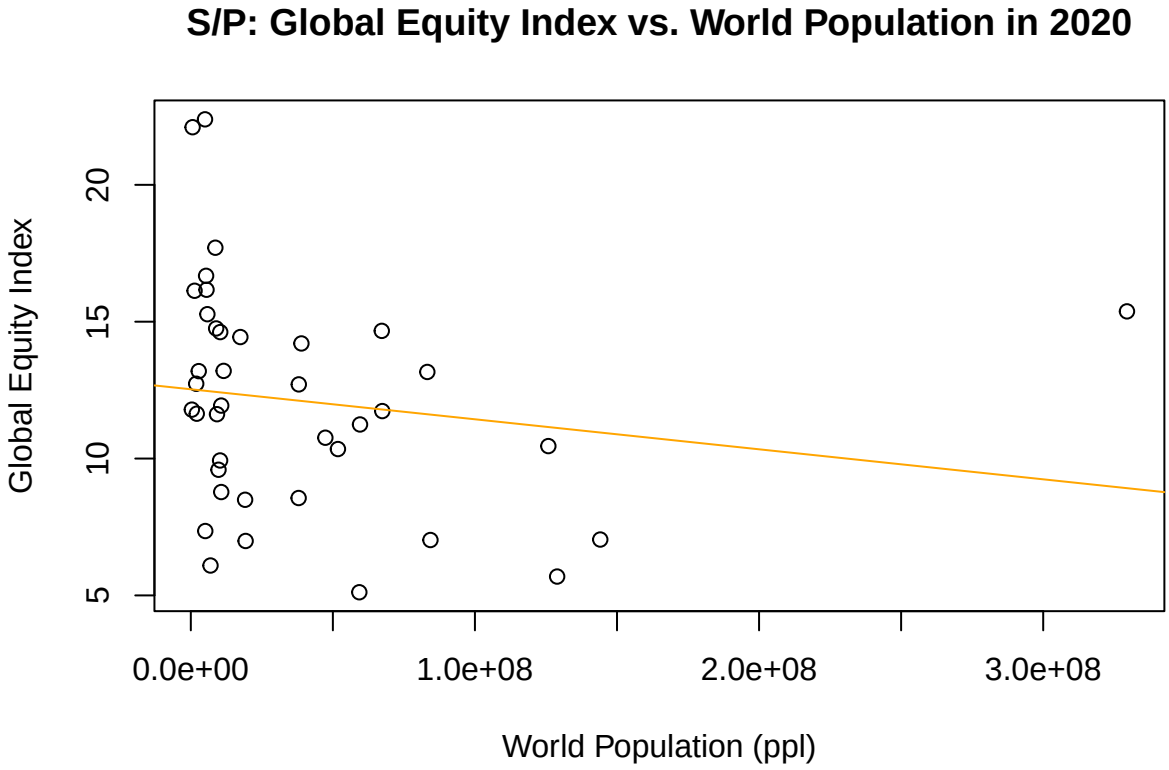

```
cor(data$HI12021, data$GE.IndexSchool)
```

```
## [1] 0.4702592
```

```
plot(data$WP2020, data$GE.IndexSchool, xlab = "World Population (ppl)", ylab = "Global Equity Index", ma
lin.reg4 = lm(data$GE.IndexSchool ~ data$WP2020)
lin.reg4
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$WP2020)
##
## Coefficients:
## (Intercept)  data$WP2020
##   1.253e+01   -1.097e-08
```

```
abline(lin.reg4, col="orange")
```



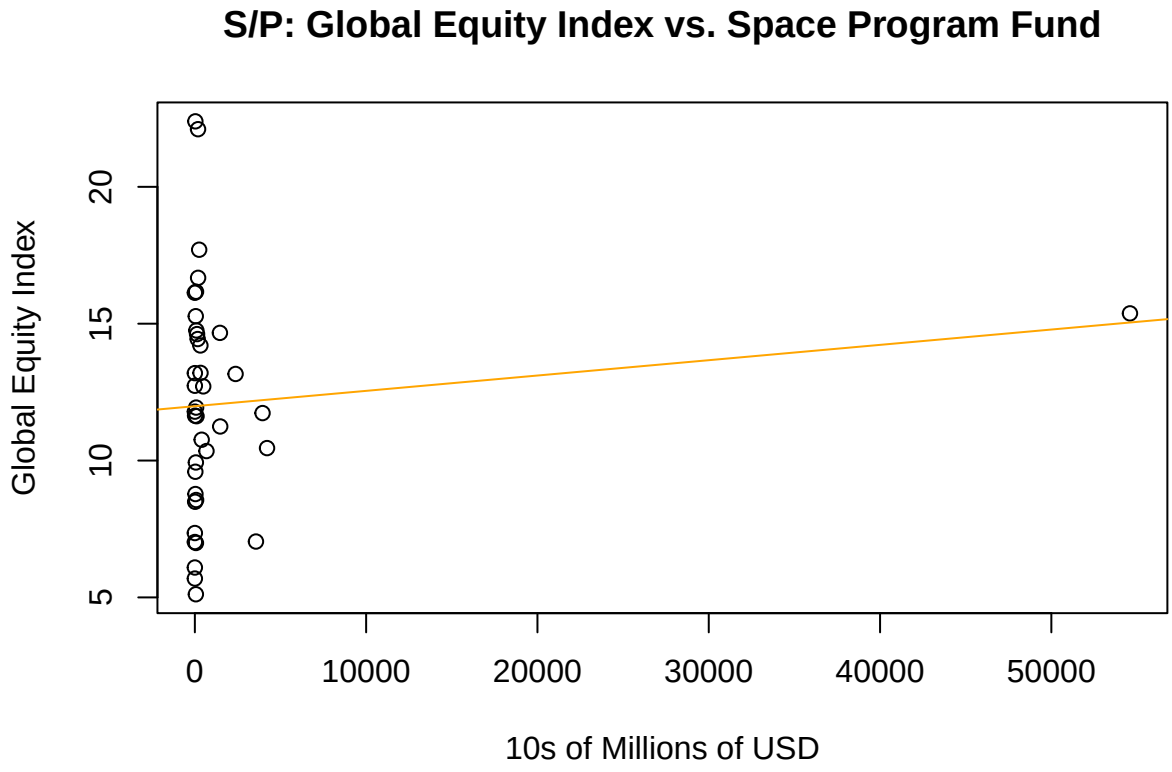

```
cor(data$WP2020, data$GE.IndexSchool)
```

```
## [1] -0.1645773
```

```
plot(data$SPFund, data$GE.IndexSchool, xlab = "10s of Millions of USD", ylab = "Global Equity Index", ma
lin.reg5 = lm(data$GE.IndexSchool ~ data$SPFund)
lin.reg5
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$SPFund)
##
## Coefficients:
## (Intercept)  data$SPFund
##   1.199e+01    5.601e-05
```

```
abline(lin.reg5, col="orange")
```




```
cor(data$SPFund, data$GE.IndexSchool)
```

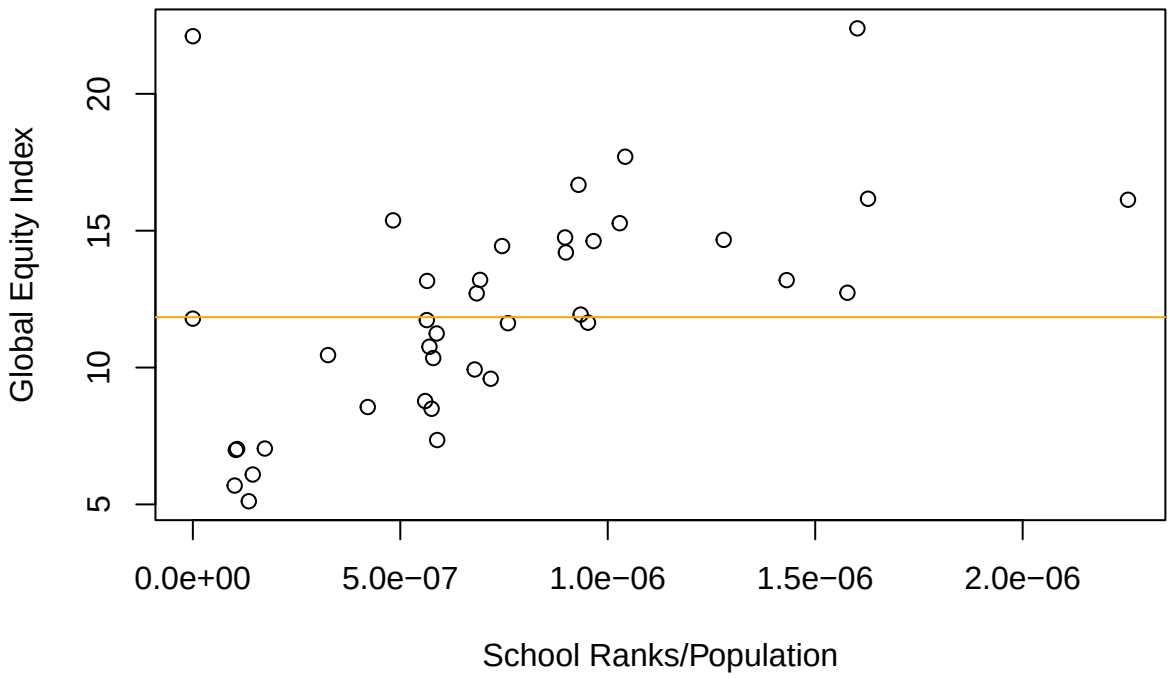
```
## [1] 0.1200979
```

```
plot(data$TotalSRank/data$WP2020, data$GE.IndexSchool, xlab = "School Ranks/Population", ylab = "Global
lin.reg6 = lm(data$GE.IndexSchool ~ data$TotalSRank/data$WP2020)
lin.reg6
```

```
##
## Call:
## lm(formula = data$GE.IndexSchool ~ data$TotalSRank/data$WP2020)
##
## Coefficients:
##                   (Intercept)               data$TotalSRank
##                     1.184e+01                     1.122e-02
## data$TotalSRank:data$WP2020
##                     1.904e-11
```

```
abline(lin.reg6, col="orange")
```

```
## Warning in abline(lin.reg6, col = "orange"): only using the first two of 3
## regression coefficients
```




```
cor(data$TotalSRank/data$WP2020, data$GE.IndexSchool)
```

```
## [1] 0.5595123
```

```
sd(data$GE.Index)
```

```
## [1] 7.014322
```

```
mean(data$GE.Index)
```

```
## [1] 18.73189
```

```
summary(data$GE.Index)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   7.359  13.715  17.644  18.732  22.109  42.713
```

```
(21.1346*(4/11)) + (18.897/31.02)*(2/11)+ .14356*(1/11) + 5.533*(1/11) + (1-3.266)*10*(3/11)
```

```
## [1] 2.132121
```