JSC370 Midterm Report

2023-02-11

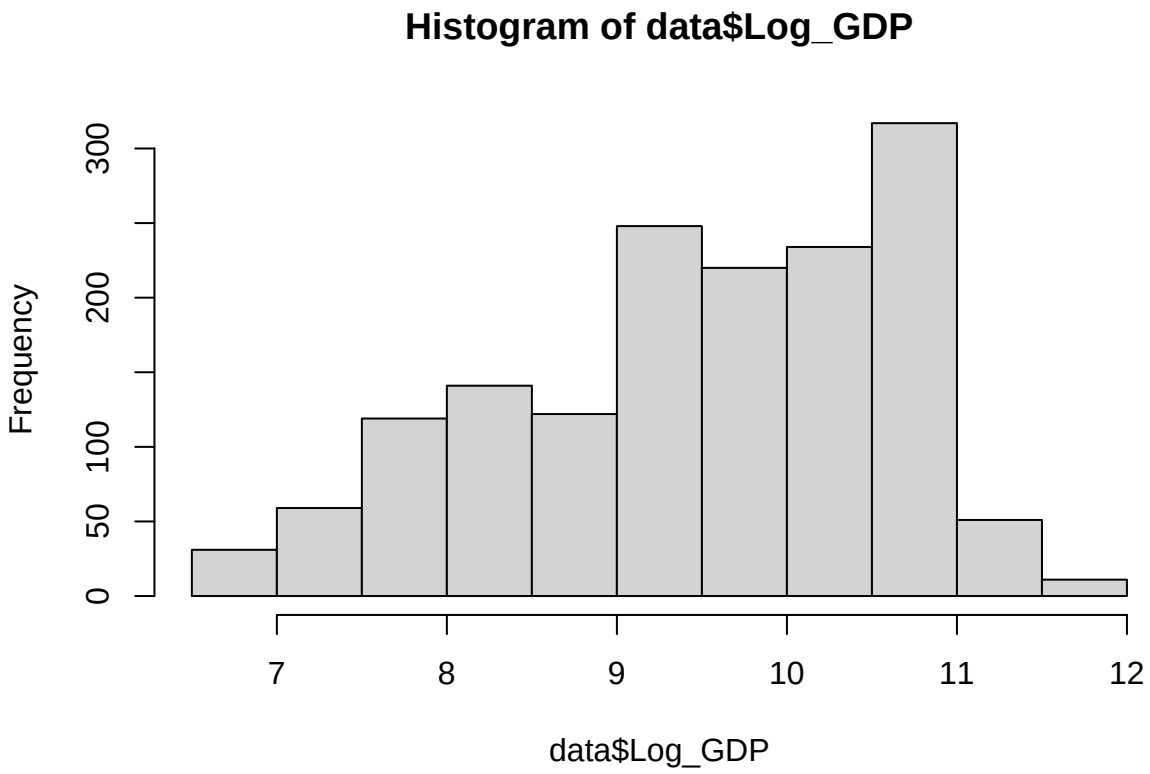
Introduction

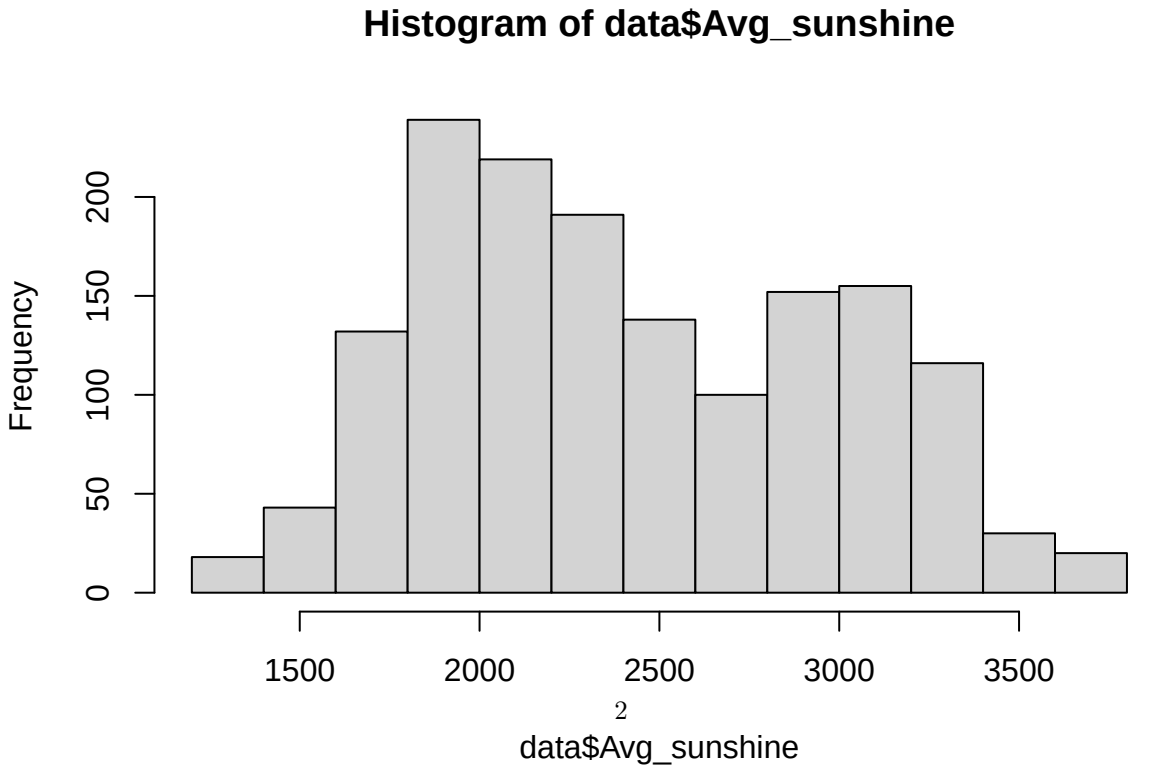
```
## # A tibble: 1 x 16
##   Country      City    Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct
##   <chr>        <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 United Stat~ Yuma   268.  271.  336.  366.  407.  415.  393.  376.  342.  320.
## # ... with 4 more variables: Nov <dbl>, Dec <dbl>, Year <dbl>, Ref. <chr>
```

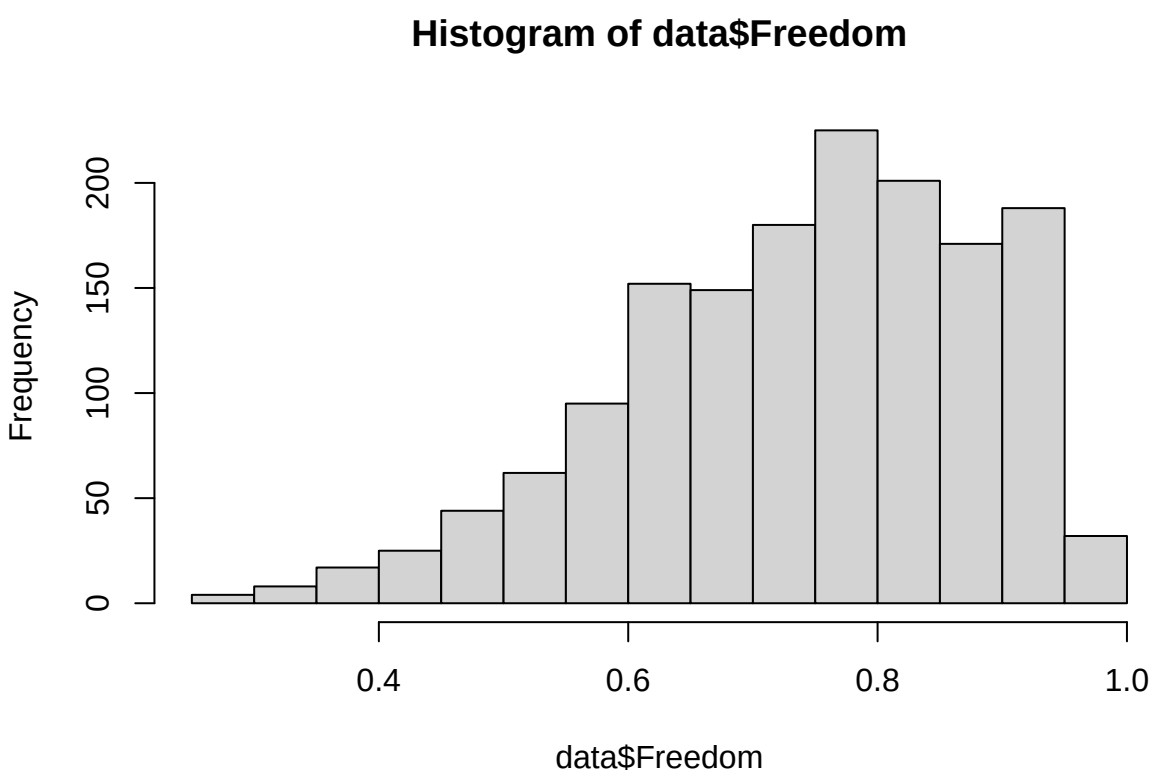
```
## # A tibble: 1 x 16
##   Country      City    Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct
##   <chr>        <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Faroe Islan~ Tórs~    14    36    71   106   124   125   111    98    80    49
## # ... with 4 more variables: Nov <dbl>, Dec <dbl>, Year <dbl>, Ref. <chr>
```

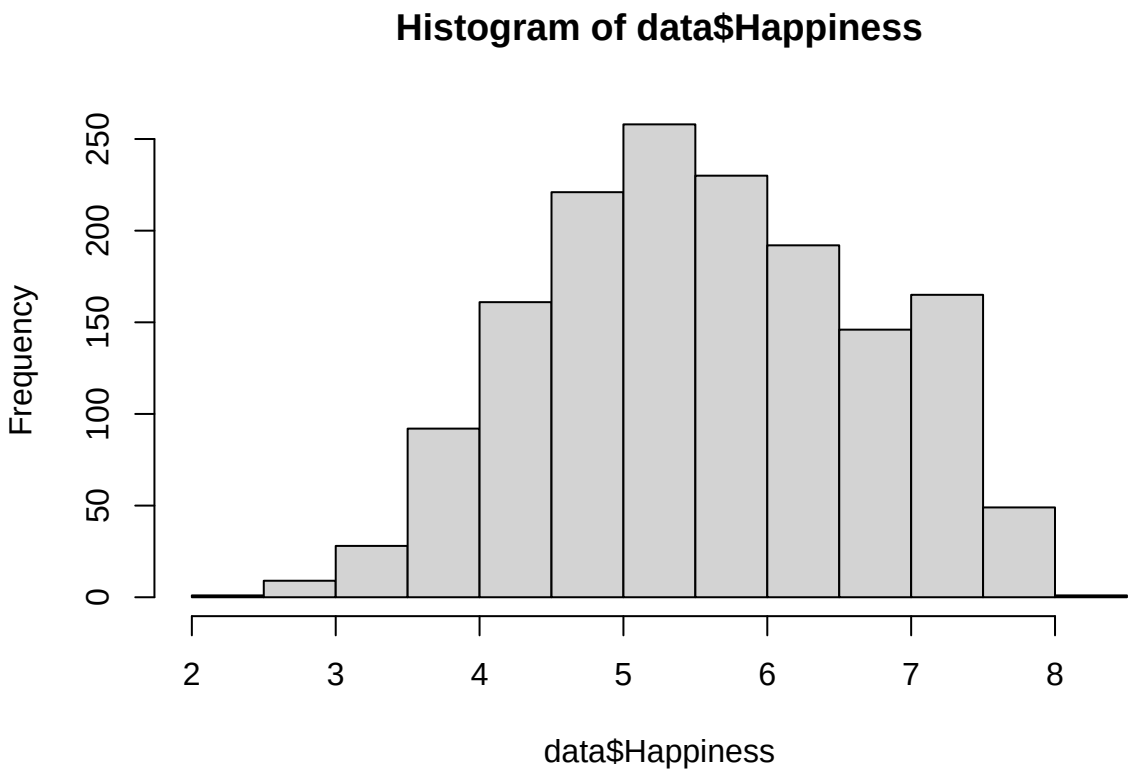
Data Wrangling/Cleaning

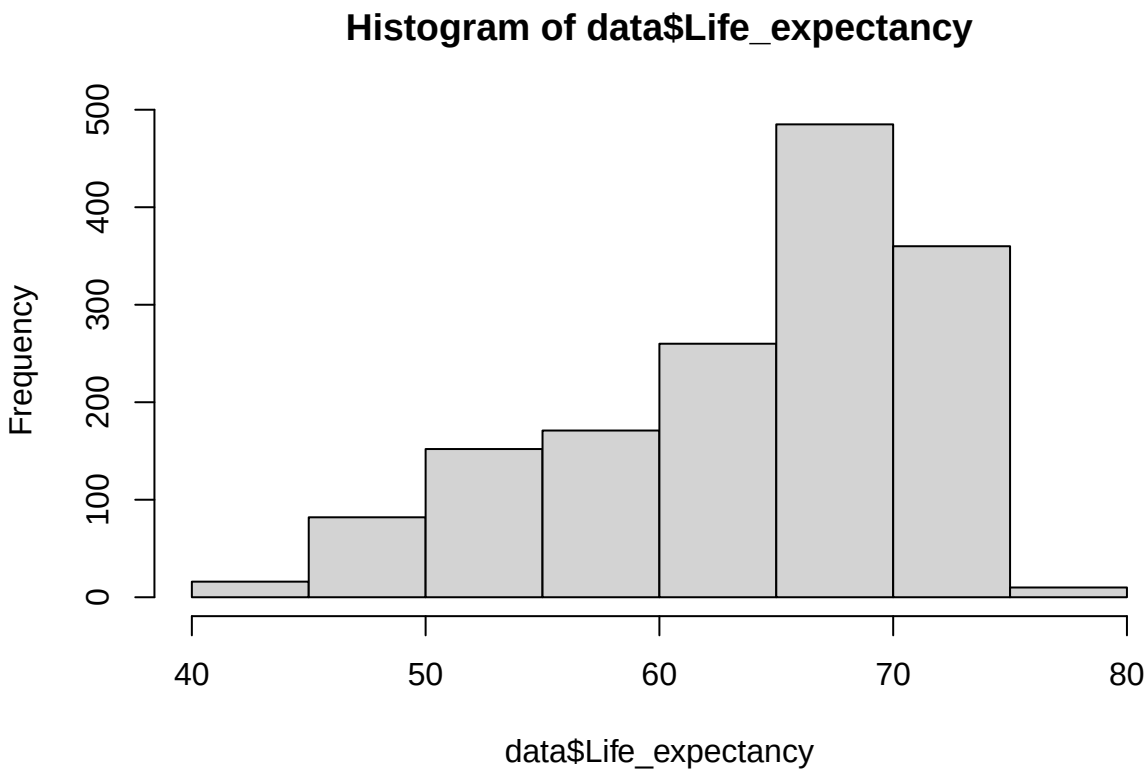
EDA histograms to look at distribution

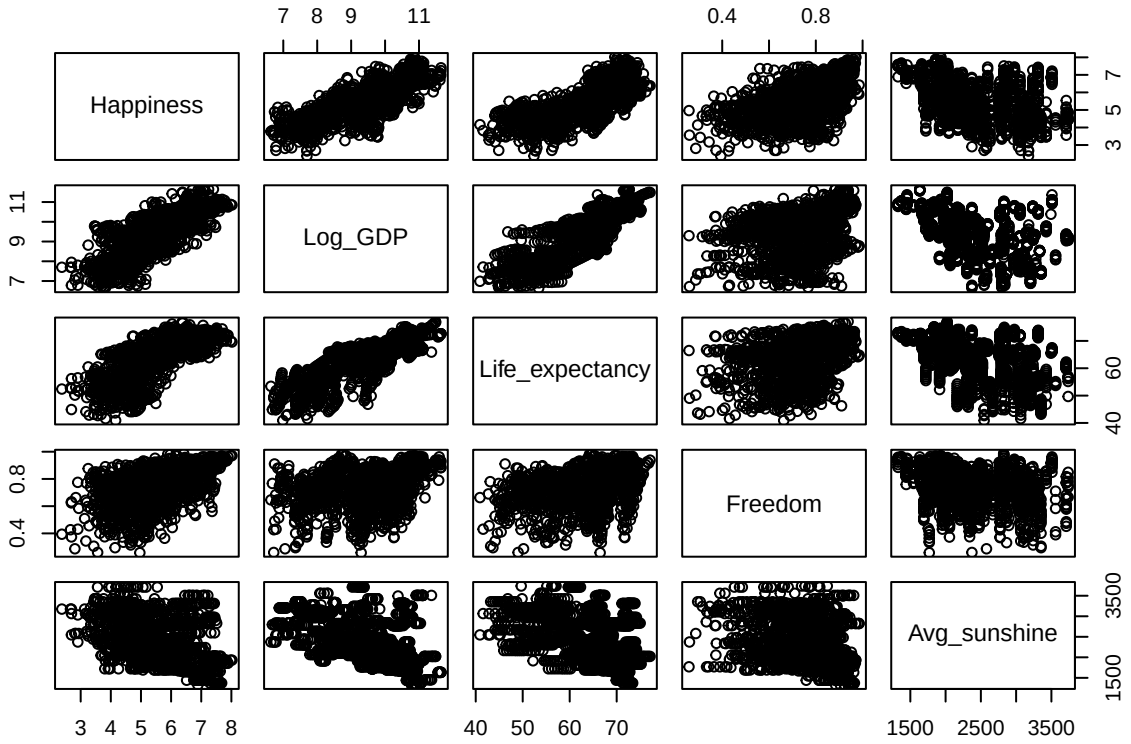





Histogram of data$Freedom
Frequency
0
50
100
150
200
0.4
0.6
0.8
1.0
data$Freedom

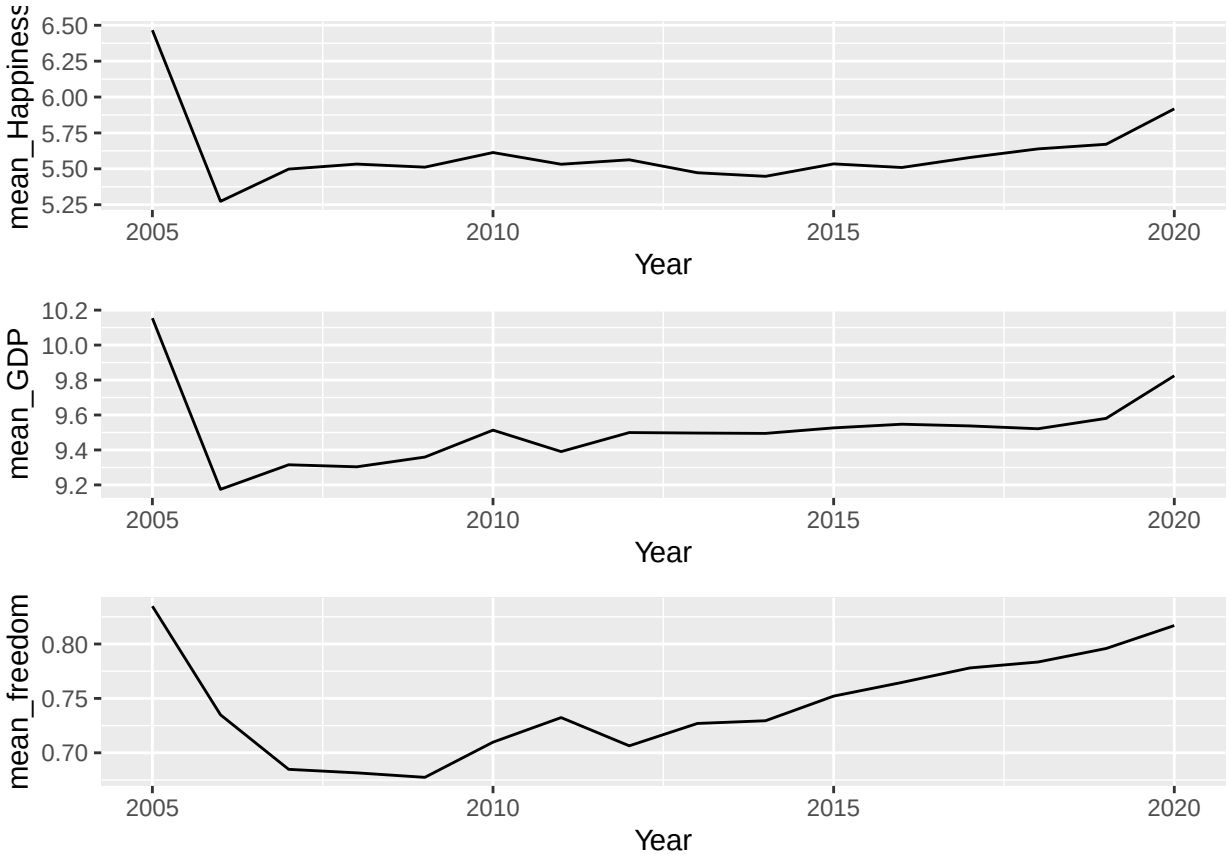

Histogram of data$Happiness
Frequency
0
50
100
150
200
250
2
3
4
5
6
7
8
data$Happiness

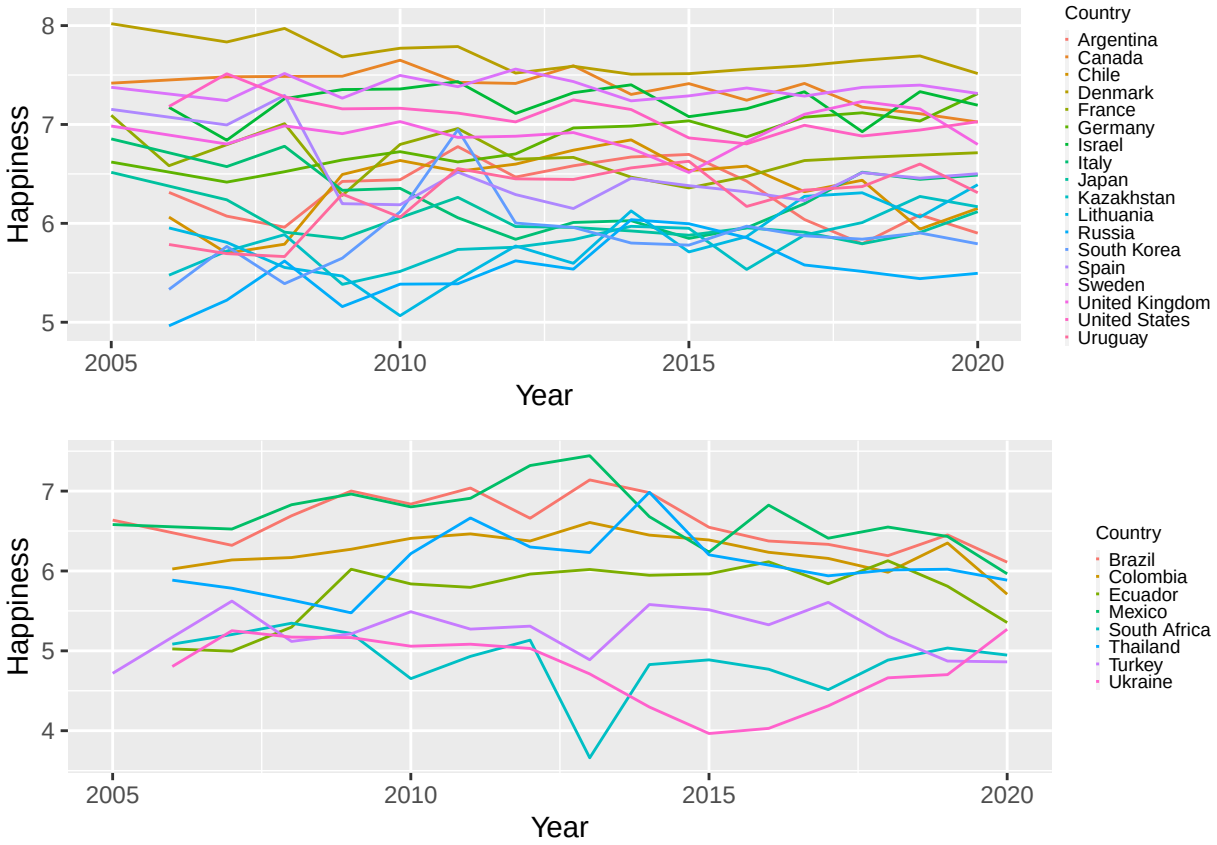

Histogram of data$Life_expectancy
Frequency
0
100
200
300
400
500
40
50
60
70
80
data$Life_expectancy



Preliminary Analysis

0. Summary table

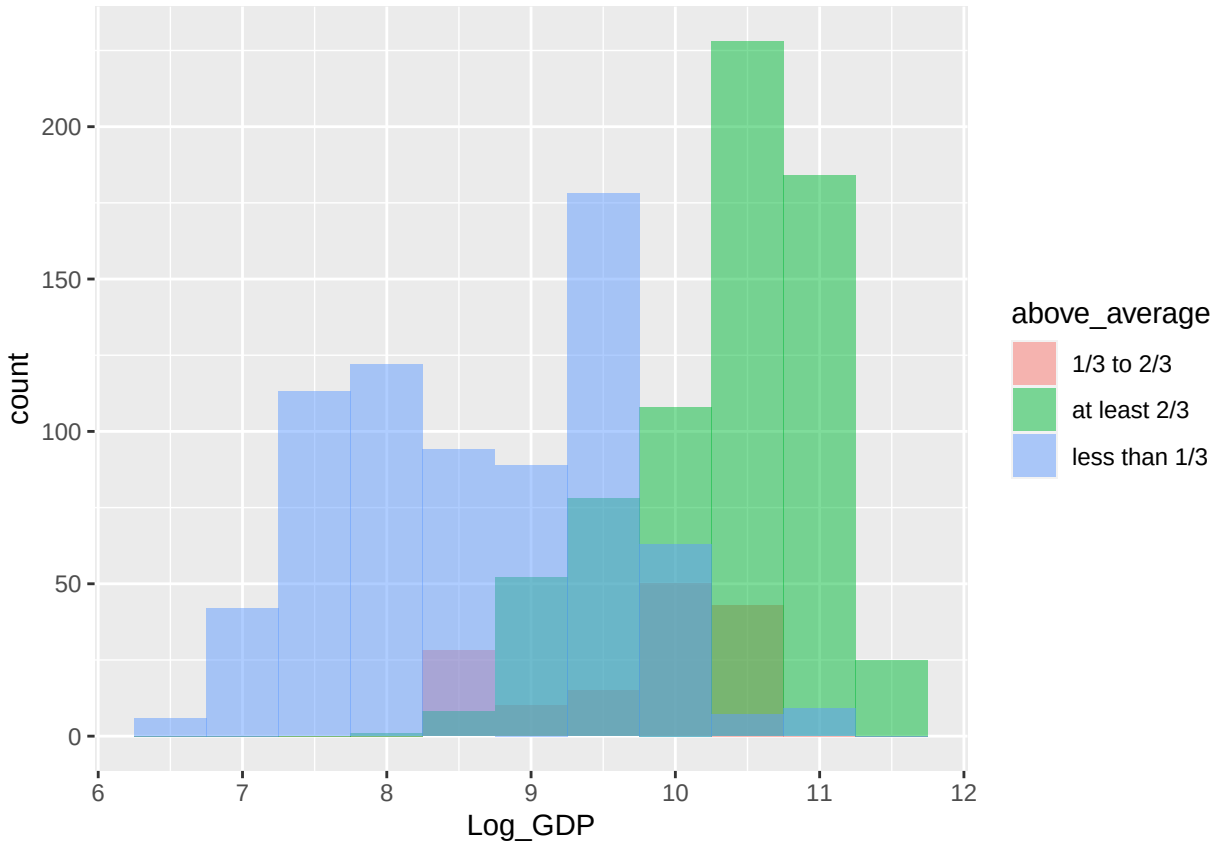

mean_Happiness
6.50
6.25
6.00
5.75
5.50
5.25
2005
2010
2015
2020
Year
mean_GDP
10.2
10.0
9.8
9.6
9.4
9.2
2005
2010
2015
2020
Year
mean_freedom
0.80
0.75
0.70
2005
2010
2015
2020
Year

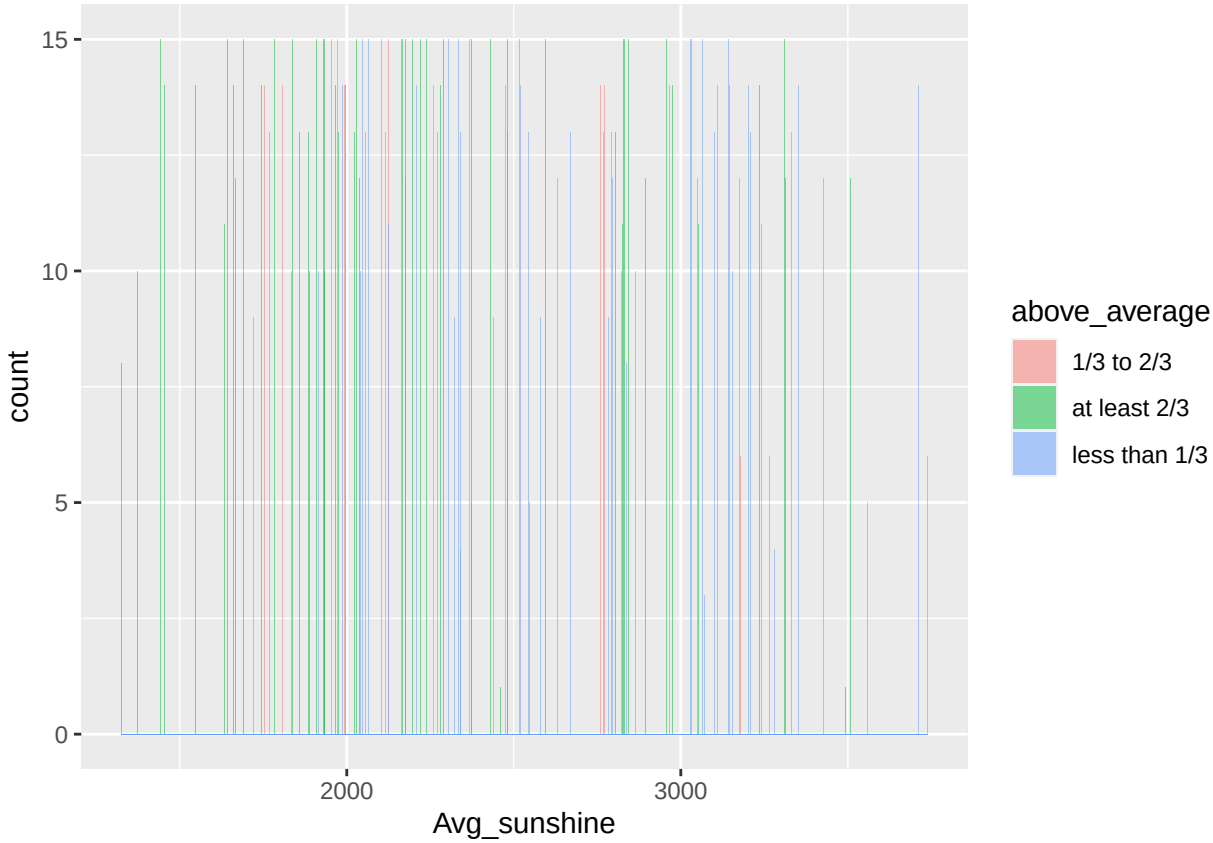


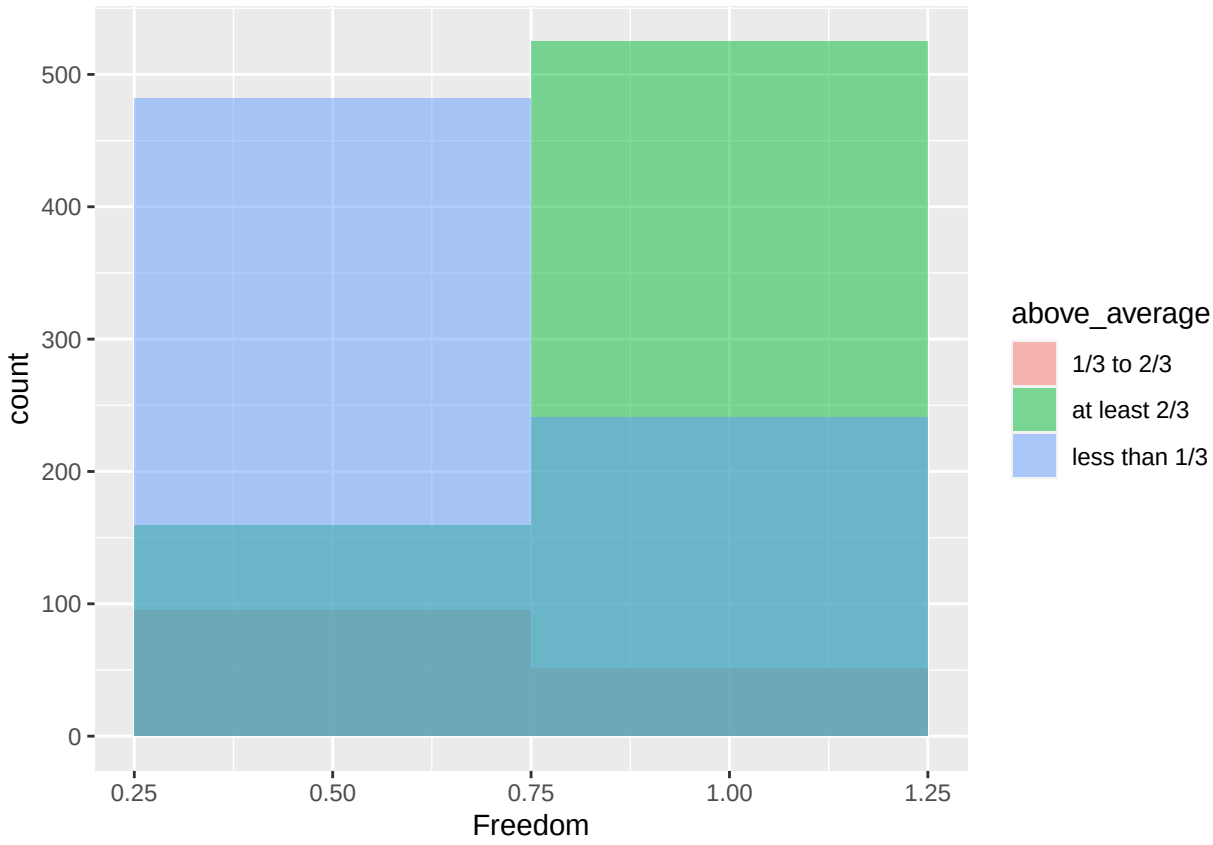
Response

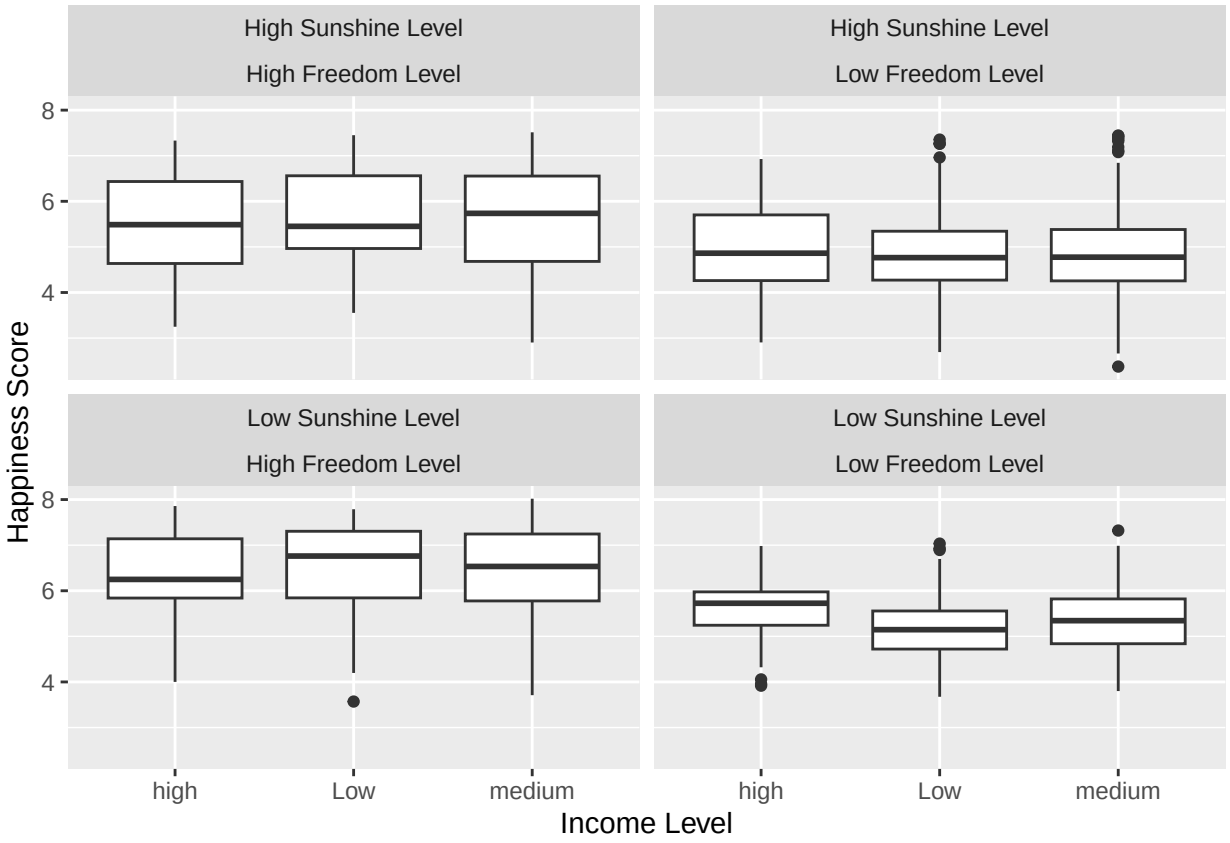
No specific trend is observed.

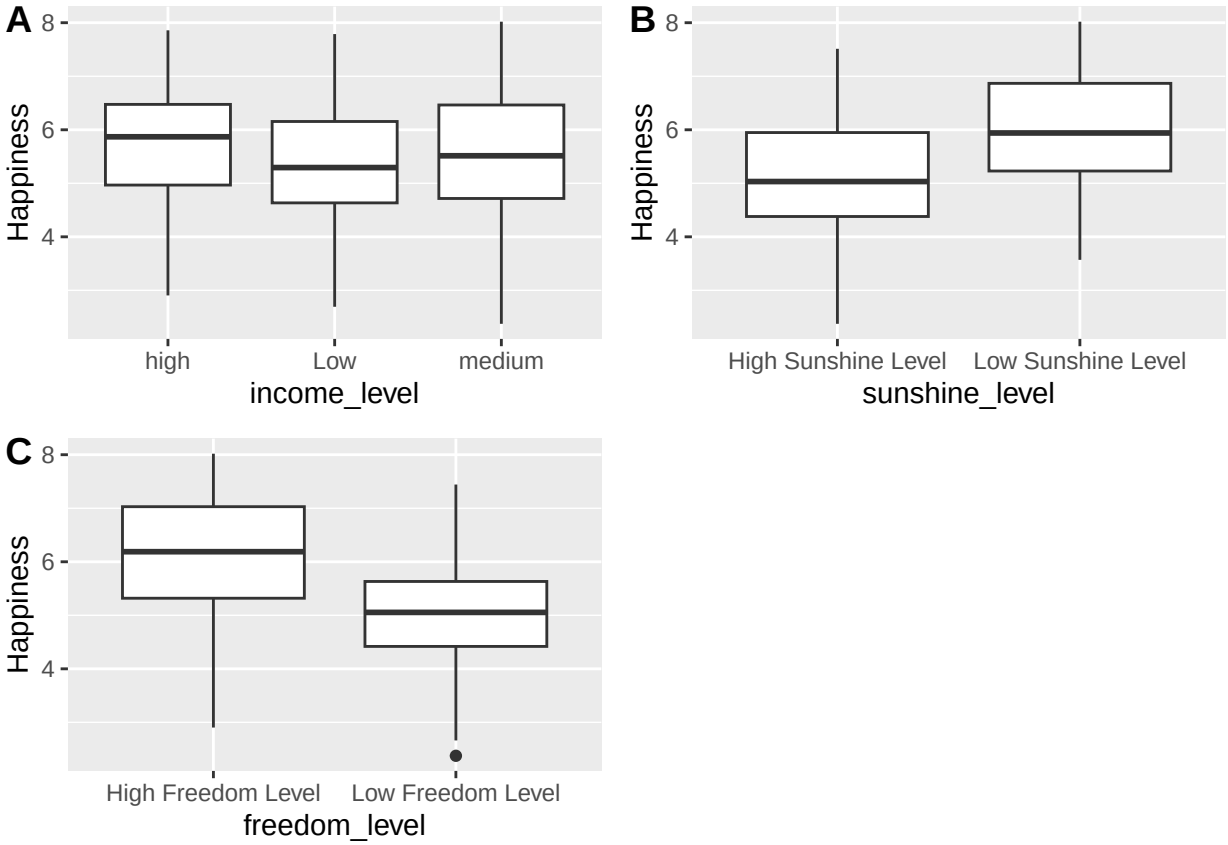
1. For countries with higher mean average happiness at least $\frac{2}{3}$, $\frac{1}{3} - \frac{2}{3}$, less than $\frac{1}{3}$, what are the distribution for the three variables?

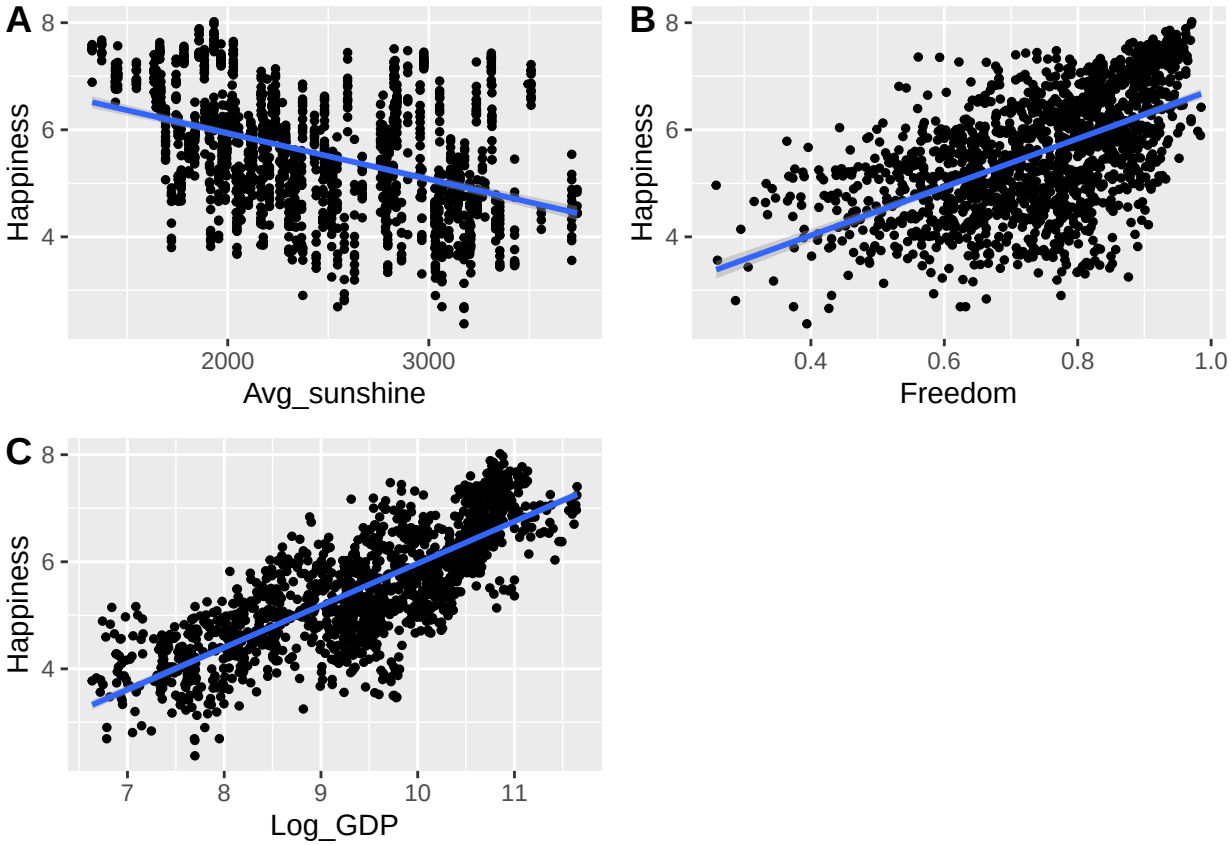

count
200
150
100
50
0
6
7
8
9
10
11
12
Log_GDP
above_average
1/3 to 2/3
at least 2/3
less than 1/3


15
10
5
0
count
2000
3000
Avg_sunshine
above_average
1/3 to 2/3
at least 2/3
less than 1/3


500
400
300
200
100
0
count
0.25
0.50
0.75
1.00
1.25
Freedom
above_average
1/3 to 2/3
at least 2/3
less than 1/3


High Sunshine Level
High Freedom Level
High Sunshine Level
Low Freedom Level
Low Sunshine Level
High Freedom Level
Low Sunshine Level
Low Freedom Level
Happiness Score
8
6
4
high
Low
medium
Income Level


A
8
6
4
Happiness
high
Low
medium
income_level
B
8
6
4
Happiness
High Sunshine Level
Low Sunshine Level
sunshine_level
C
8
6
4
Happiness
High Freedom Level
Low Freedom Level
freedom_level


A
B
C
Happiness
Avg_sunshine
Freedom
Log_GDP
8
6
4
2000
3000
0.4
0.6
0.8
1.0
7
8
9
10
11