Global Life Satisfaction

By Ye-in Jeon, Sam Kupfer, Amy Mykityshyn, Krystin Sinclair, Vishnupriya Venkateswaran
February 28, 2018

Overview

- Background of Dataset
- Data Exploration
- Data Visualization
- Background Research
- SMART Question Development
- Epicycles of Analysis (before & after research)
- Correlation
- Conclusion
- Tools and References

Background of Dataset

Organisation for Economic Cooperation and Development (OECD):
- International forum, 35 member countries
- Mission: "To promote policies that will improve the economic and social well-being of people around the world"

OECD Better Life Index:
- Data on social protection and well-being is collected annually for member countries plus others. We used data from 2015-2017.
- Data includes variables related to housing, income, jobs, community support, education, environment (air and water quality), civic engagement, health, life satisfaction, safety, and work-life balance.

Structure of Dataset

```
'data.frame':    112 obs. of  29 variables:
 $ Country                               : Factor w/ 38 levels "Australia","Austria",..: 1 2 3 5 6 7 8 9 10
11 ...
 $ Year                                  : int  2015 2015 2015 2015 2015 2015 2015 2015 2015 2015 ...
 $ Dwellings.without.basic.facilities    : num  1.1 1 2 0.2 9.4 0.9 0.9 8.1 0.6 0.5 ...
 $ Housing.expenditure                   : int  20 21 21 21 18 26 24 19 22 21 ...
 $ Rooms.per.person                      : num  2.3 1.6 2.2 2.5 1.2 1.4 1.9 1.5 1.9 1.8 ...
 $ Household.net.adjusted.disposable.income : int  31588 31173 28307 29365 14533 18404 26491 15167 27927 28799
...
 $ Household.net.financial.wealth        : int  47657 49887 83876 67913 17733 17299 44488 7680 18761 48741
...
 $ Employment.rate                       : int  72 72 62 72 62 68 73 68 69 64 ...
 $ Job.security                          : num  4.8 3.9 5 6.4 4.4 4.1 5.6 5.2 6.9 6.5 ...
 $ Long.term.unemployment.rate           : num  1.08 1.19 3.88 0.9 1.59 3.12 1.78 3.82 1.73 3.99 ...
 $ Personal.earnings                     : int  50449 45199 48082 46911 22101 20338 48347 18944 40060 40242
...
 $ Quality.of.support.network            : int  92 89 94 92 86 85 95 89 95 87 ...
 $ Educational.attainment                : int  76 83 72 89 57 92 78 90 85 73 ...
 $ Student.skills                        : int  512 500 509 522 436 500 498 526 529 500 ...
 $ Years.in.education                    : num  19.4 17 18.9 17.2 16.5 18.1 19.4 17.5 19.7 16.4 ...
 $ Air.pollution                         : int  13 27 21 15 46 16 15 9 15 12 ...
 $ Water.quality                         : int  91 94 87 91 73 85 94 79 94 82 ...
 $ Consultation.on.rule.making           : num  10.5 7.1 4.5 10.5 2 6.8 7 3.3 9 3.5 ...
 $ Voter.turnout                         : int  93 75 89 61 49 59 88 64 69 80 ...
 $ Life.expectancy                       : num  82.1 81 80.5 81.5 78.9 78.2 80.1 76.5 80.7 82.1 ...
 $ Self.reported.health                  : int  85 69 74 89 59 60 72 54 65 67 ...
 $ Life.satisfaction                     : num  7.3 6.9 6.9 7.3 6.7 6.5 7.5 5.6 7.4 6.5 ...
 $ Assault.rate                          : num  2.1 3.4 6.6 1.3 6.9 2.8 3.9 5.5 2.4 5 ...
 $ Homicide.rate                         : num  0.8 0.4 1.1 1.5 4.4 0.8 0.3 4.8 1.4 0.6 ...
 $ Employees.working.very.long.hours     : num  14.02 7.61 4.57 3.94 15.42 ...
 $ Time.devoted.to.leisure.and.personal.care : num  14.4 14.5 15.7 14.2 14.4 ...
```

Variables of Interest

- **Economic**
  - Personal earnings *(US dollars)*
  - Employment rate *(percentage of working-age population)*
  - Household net adjusted disposable income *(US Dollars)*
  - Rooms per person *(total number of rooms divided by number of people in dwelling)*
  - Long term unemployment *(percentage of labor force, 1 year or more)*
- **Mental and Physical Health**
  - Homicide rate *(rate per 100,000 people)*
  - Life expectancy *(years, people born today)*
  - Self reported health *(survey, percentage reporting "good" or better health)*
  - Time devoted to leisure and personal care *(hours per day)*
  - Quality of support network *(survey, age 15 or older)*
 - Dwelling without basic facilities *(percentage of population without indoor toilets)*
- **Environment and Education**
  - Air pollution *(micrograms per cubic meter)*
  - Years of education *(years starting at age 5)*
  - Student skills *(average score in reading, math, and science on OECD Programme for International Student Assessment)*
  - Water quality *(survey, percentage satisfied with quality of water)*

Descriptive Statistics

Variable	Count	Mean	Std. Dev	Min	25%	Median	75%	Max
Life satisfaction	112	6.54	0.77	4.80	5.90	6.70	7.30	7.60
Personal earnings	112	$36,482	$13,965	$5,429	$23,593	$36,504	$48,308	$62,636
Employment rate	112	66.84%	7.97%	43.00%	62.00%	67.50%	73.00%	86.00%
Household net adjusted disposable income	112	$24,294	$7,776	$8,712	$18,252	$23,965	$29,567	$44,049
Rooms per person	112	1.64	0.45	0.7	1.2	1.7	1.9	2.5
Time devoted to leisure and personal care	112	14.84	0.70	12.24	14.48	14.90	15.17	16.36
Homicide rate	112	2.95	5.64	0.20	0.60	1.00	1.53	27.60
Life expectancy	112	79.51	4.23	56.80	78.15	81.05	81.92	83.90
Air pollution	112	15.63	7.05	3.00	11.00	15.00	18.00	46.00
Years of education	112	17.41	1.31	14.40	16.40	17.40	18.12	21.20

Life Satisfaction by OECD Country Geomap
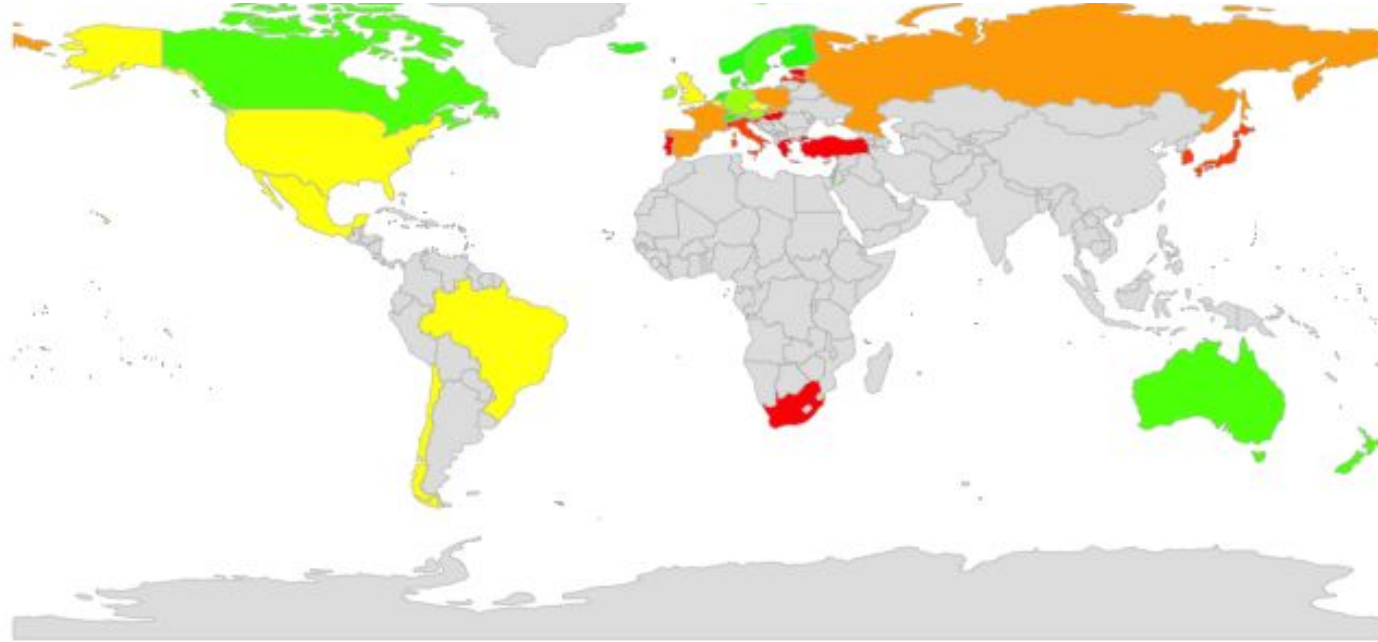
Life satisfaction was a self-reported measure with a scale of 1 (low) to 10 (high)

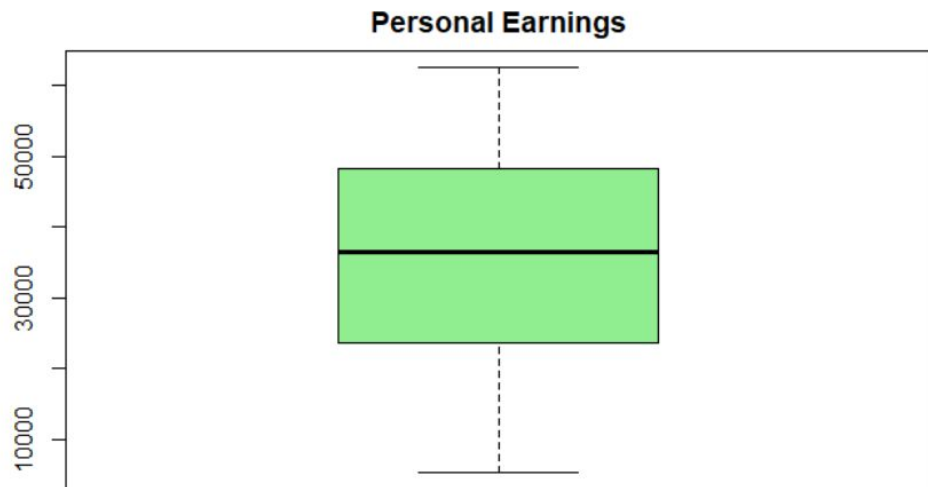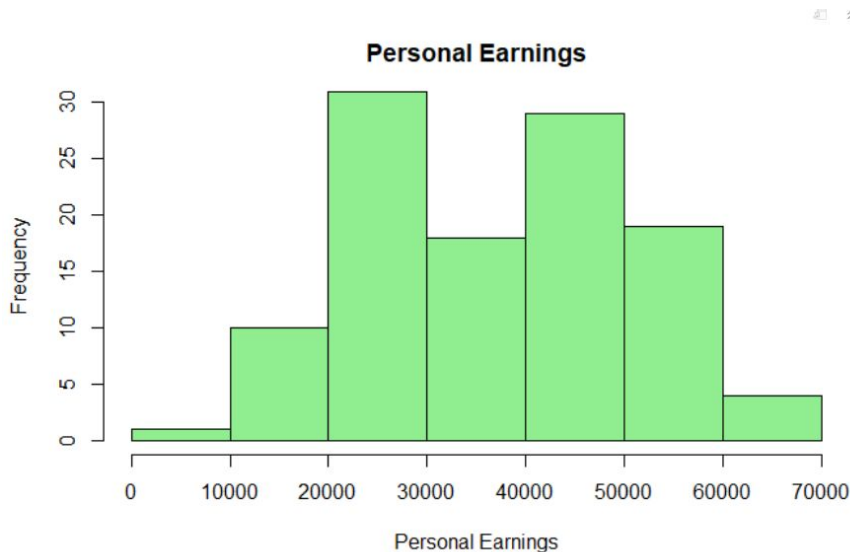4.8   7.5

Life Satisfaction by OECD Country Geomap

Code:

```r
```{r 2017_ls_map}
oecd17 <- subset(oecd, subset = Year == 2017)
library(rworldmap)
oecd17_joined <- joinCountryData2Map(oecd17, joinCode ="NAME", nameJoinColumn = "Country")
mapCountryData(oecd17_joined, nameColumnToPlot="Life.satisfaction", colourPalette = c("red",
"yellow", "green"), missingCountryCol = "gray89", addLegend = TRUE, mapTitle = "Life
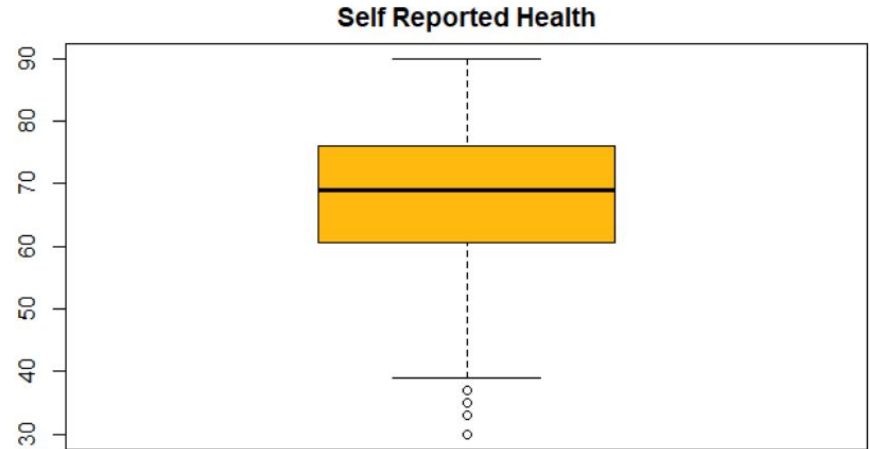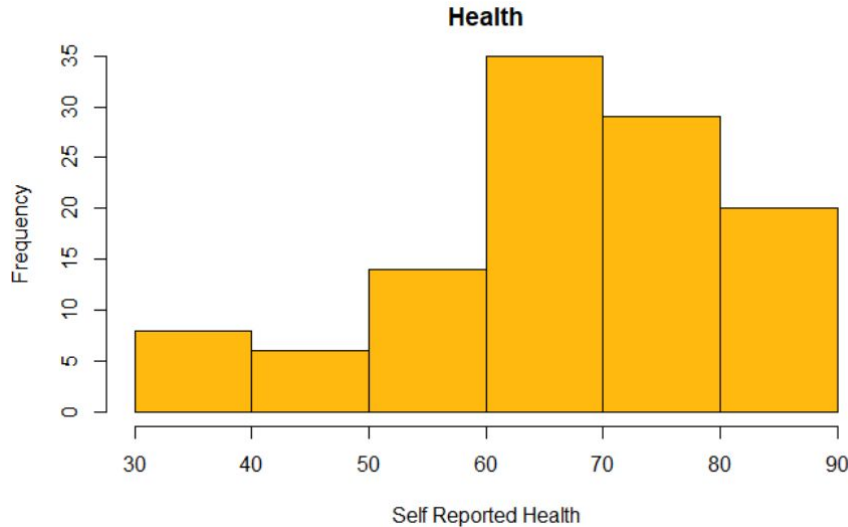Satisfaction in OECD Countries" )
```
```

Data Visualization
Economic (personal earnings)


```r
```{r data visualization}
hist(oecd$Personal.earnings, xlab="Personal Earnings", main="Personal Earnings",
col="lightgreen")
boxplot(oecd$Personal.earnings, main="Personal Earnings", col="lightgreen")
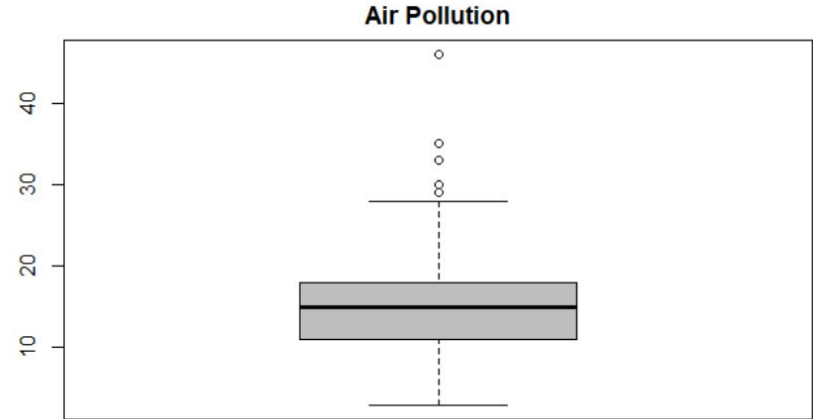```

# Data Visualization
## Mental and Physical Health (self reported health)



```
hist(oecd$Self.reported.health, xlab="Self Reported Health", main="Health",
col="darkgoldenrod1")
boxplot(oecd$Self.reported.health, main="Self Reported Health", col="darkgoldenrod1")
```

Data Visualization
Environment / Education (air pollution)
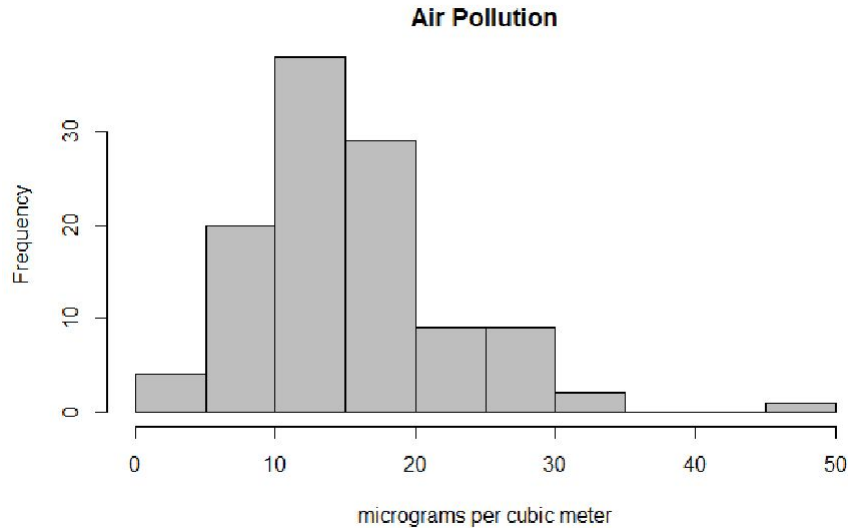

```
hist(oecd$Air.pollution, xlab="micrograms per cubic meter", main="Air Pollution", col="gray")
boxplot(oecd$Air.pollution, main="Air Pollution", col="gray")
```

Background Research

There have been many attempts to study life satisfaction quantitatively.
"Variance in satisfaction between nations has been studied; it has been shown that living conditions are a major determinant of life satisfaction."

Survey results show that health, age, and financial stability have significant relationships with life satisfaction.
"Economically prosperous countries tend to experience [life satisfaction] more when compared to poorer nations."
"more highly educated countries generally experience higher levels of satisfaction, but with this education comes opportunity for aversive consequences: ...job competition, or even lack of jobs"
"The present analysis shows not only that health is the strongest variable, but that it alone accounts for the large majority of the explained variance in life satisfaction"

Different studies have come to different conclusions about what variables are most important to life satisfaction

SMART Question Development

Keeping in mind the OECD mission and dataset, we are interested in what affects life satisfaction. This is a very general question that we tried to make more specific for quantitative study.

Most general: What characteristics contribute most to life satisfaction?

Groupings: Do economic, mental and physical health, and environment and education contribute to life satisfaction?

More specific: How much do economic, mental and physical health, and environment and education contribute to life satisfaction?

Most specific: Which of the variables correlate strongly ($|r| > .6$) with life satisfaction in countries considered by the OECD from 2015-2017?

SMART Question

Most specific: Which of the variables correlate strongly (absolute value of Pearson correlation > .6) with life satisfaction in countries considered by the OECD from 2015-2017?

<u>S</u>pecific: About a specific set of countries, time frame, and focused on 10 variables

<u>M</u>easurable: Can be measured with a correlation coefficient.
Yes if (|r|> .6); no if (|r|< .6)

<u>A</u>chievable: Dataset is small enough to be analyzed in R

<u>R</u>elevant: Useful for researchers to understand what factors drive life satisfaction and what can be done to achieve progress

<u>T</u>ime Oriented: We believe this question is straightforward enough to answer in the given time. The data is already collected

Results of EDA:
Correlation

```{r corrplots}
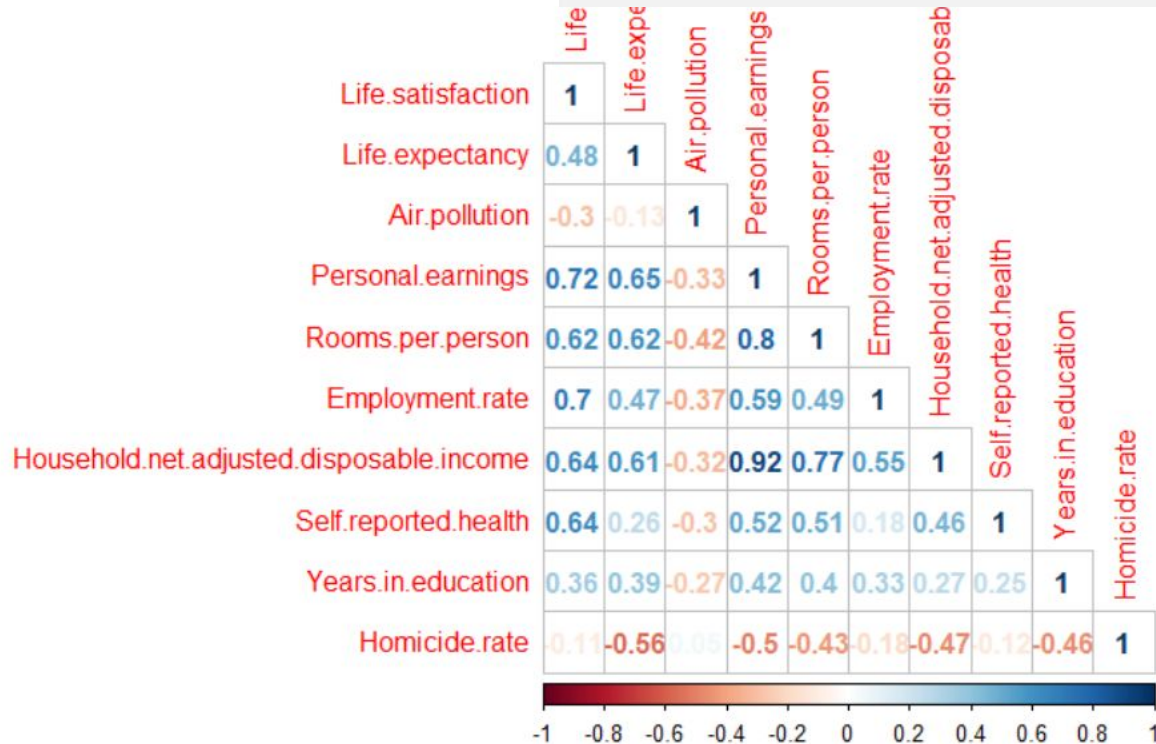oecdused <- c("Life.satisfaction", "Life.expectancy", "Air.pollution", "Personal.earnings",
"Rooms.per.person", "Employment.rate", "Household.net.adjusted.disposable.income",
"Self.reported.health", "Years.in.education", "Homicide.rate")
columnsubset <- oecd[oecdused]
oecd_cor <- cor(columnsubset, method = "pearson")
#install.packages("corrplot")
 library(corrplot)
##loadPkg(corrplot)
corrplot(oecd_cor, method = "number", type = "lower")
```

	Life	Life.expe	Air.pollution	Personal.earnings	Rooms.per.person	Employment.rate	Household.net.adjusted.disposab	Self.reported.health	Years.in.education	Homicide.rate
Life.satisfaction	1									
Life.expectancy	0.48	1								
Air.pollution	-0.3	-0.13	1							
Personal.earnings	0.72	0.65	-0.33	1						
Rooms.per.person	0.62	0.62	-0.42	0.8	1					
Employment.rate	0.7	0.47	-0.37	0.59	0.49	1				
Household.net.adjusted.disposable.income	0.64	0.61	-0.32	0.92	0.77	0.55	1			
Self.reported.health	0.64	0.26	-0.3	0.52	0.51	0.18	0.46	1		
Years.in.education	0.36	0.39	-0.27	0.42	0.4	0.33	0.27	0.25	1	
Homicide.rate	-0.11	-0.56	0.05	-0.5	-0.43	-0.18	-0.47	-0.12	-0.46	1

-1 -0.8 -0.6 -0.4 -0.2 0 0.2 0.4 0.6 0.8 1

Conclusions

Which of the variables correlate strongly(|r| > .6) with life satisfaction in countries considered by the OECD from 2015-2017?

- **Economic**
    - Personal earnings *(US dollars):* **r=.72**
    - Employment rate *(percentage of working-age population):* **r=.70**
    - Household net adjusted disposable income *(US Dollars):* **r=.64**
    - Rooms per person *(total number of rooms divided by number of people in dwelling):* **r=.62**
- **Mental and Physical Health**
    - Self reported health *(survey, percentage reporting "good" or better health):* **r=.64**

Economic variables had the strongest correlations to life satisfaction. Health also has a strong correlation. These results agree with the results of some of our background research

Epicycles of Analysis
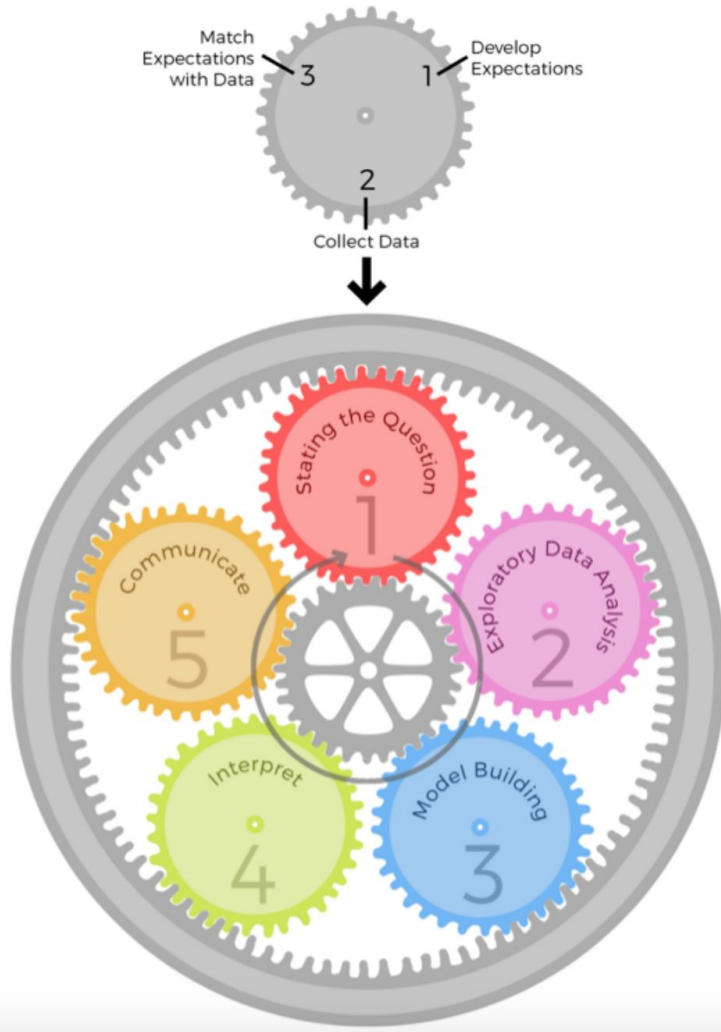
1. **Question:** Which of the variables correlate strongly (absolute value of Pearson correlation > .6) with life satisfaction in countries considered by the OECD from 2015-2017?

2. **EDA:** Descriptive Statistics and Visualization

3. **Build model:** Correlation

4. **Interpret:** strong relationship is $|r|>.6$

5. **Communicate:** Present in Class

Tools, Packages & Charts

- **R**
  - **Packages**
    - rworldmap
    - corrplot
  - **Charts**
    - Histogram
    - Correlation Matrix
    - Boxplot
    - Geomap

References

Rogers, S. (2012). Better life: relaunching the happiness index. Retrieved February 24th 2018, from https://www.theguardian.com/news/datablog/2012/may/22/better-life-index-oecd

Chompoo (2015). Life Satisfaction and its seven contributing factors. Retrieved February 24th 2018, from https://positivepsychologyprogram.com/life-satisfaction/

Palmore, E. & Luikart, C. (2018). Health and Social Factors Related to Life Satisfaction . Retrieved February 24th 2018 from https://www.jstor.org/stable/pdf/2136974.pdf

(2017)Oecd. Better Life Index -Edition 2017. Retrieved February 10th 2018 from http://stats.oecd.org/Index.aspx?DataSetCode=BLI

Questions