What Relates to Life Expectancy

A PROPOSAL TO IDENTIFY ELEMENTS RELATED TO LIFE EXPECTANCY FROM THE GLOBAL PERSPECTIVE

H Karen Tam
February 10, 2018

Content

- Problems
  - Inequality on life expectancy
- Dataset
  - Gapminder
- Data Wrangling
- Exploratory Analysis
  - Density plots
  - Descriptive statistics
- In-depth Analysis
  - Pearson Correlation Coefficient
  - Multiple Linear Regression
- Recommendations
- Potential Improvements

Problems

Do poor people have shorter life spans?

All people have the basic right to survive.

Problems

- Who is interested?
  - Government agencies
  - Public health research organizations
  - Community nonprofits
  - Healthcare facilities
- Motivation
  - Improve equality in life expectancy.
  - Seek answers to close the disparity gap in life spans of the rich and the poor.

Dataset

- Data are from Gapminder.org
- A final dataset was created by combining datasets for the following variables:
  - Life Expectancy
  - Income
  - Total Health Spending
  - Government Health Spending
  - Improved Drinking Water Source
  - Food Supply
  - Various Cancer Deaths

Dataset

- This project evaluates the following variables:

Variables	Years	Number of Countries/ Territories
Life Expectancy	1960 - 2016	209
Income	1960 - 2011	200
Total Health Spending	1995 - 2010	189
Government Health Spending	1995 - 2010	191
Improved Drinking Water Source	1990 - 2010	201
Food Supply	1961 - 2007	176
Cancer Death	1995 - 2002	149

Data Wrangling

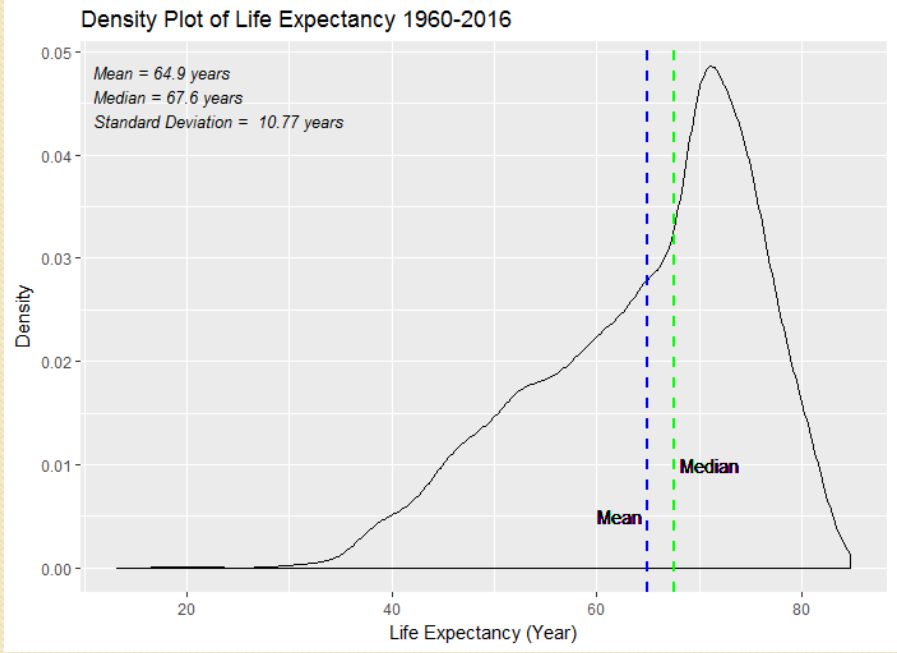
- Data Wrangling
  - Reshape the data for plots and statistical analysis
  - Remove missing value and join the data from different variable datasets
  - Examine outliers

Exploratory Analysis

Income

Mean: US$6,616
Median: US$1,896
Standard Deviation: US$10,671

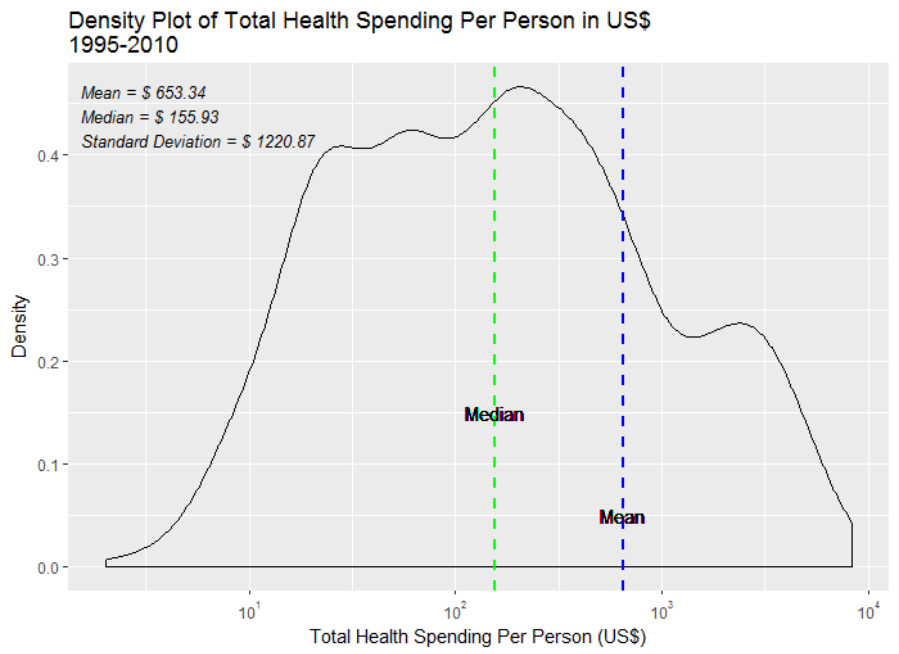
Life Expectancy



Mean: 65 years
Median: 68 years
Standard Deviation: 11 years

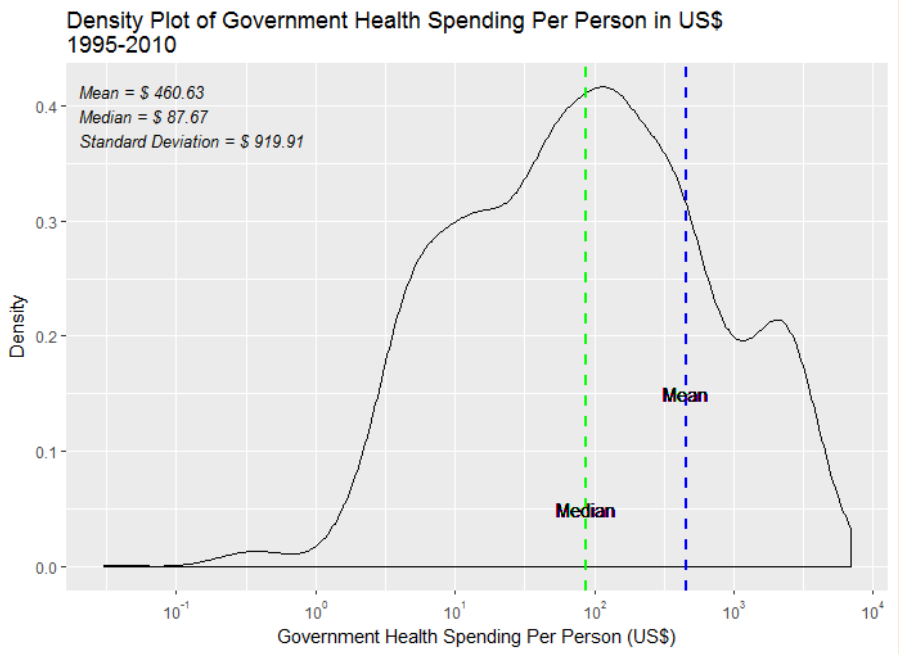
Exploratory Analysis

Total Health Spending



Mean: US$653
Median: US$156
Standard Deviation: US$1,221

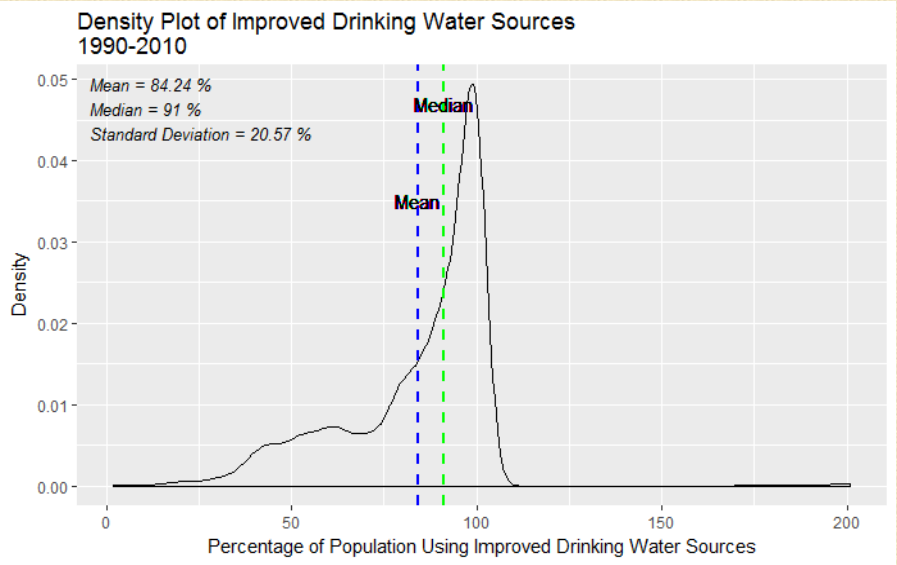
Government Health Spending



Mean: US$461
Median: US$88
Standard Deviation: US$920

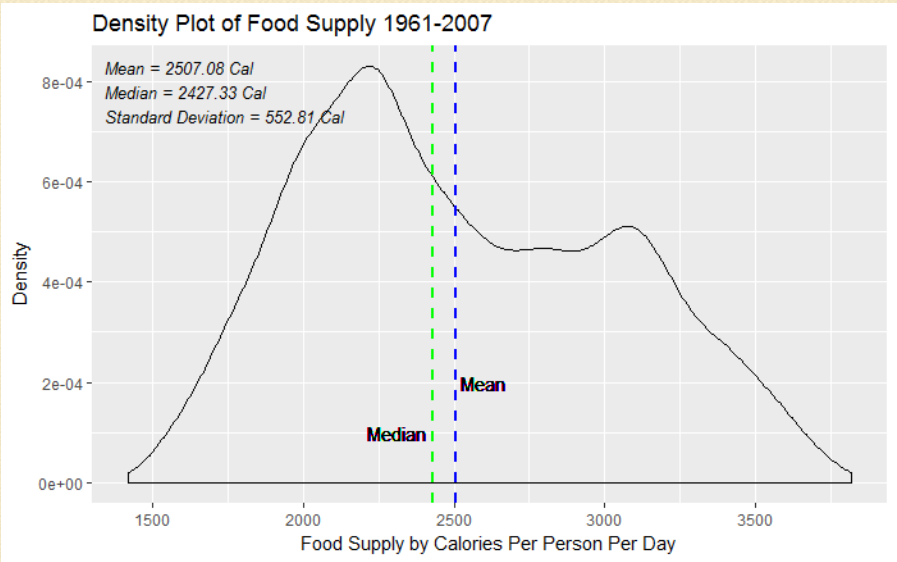
Exploratory Analysis

Drinking Water



Mean: 84 %
Median: 91%
Standard Deviation: 21%

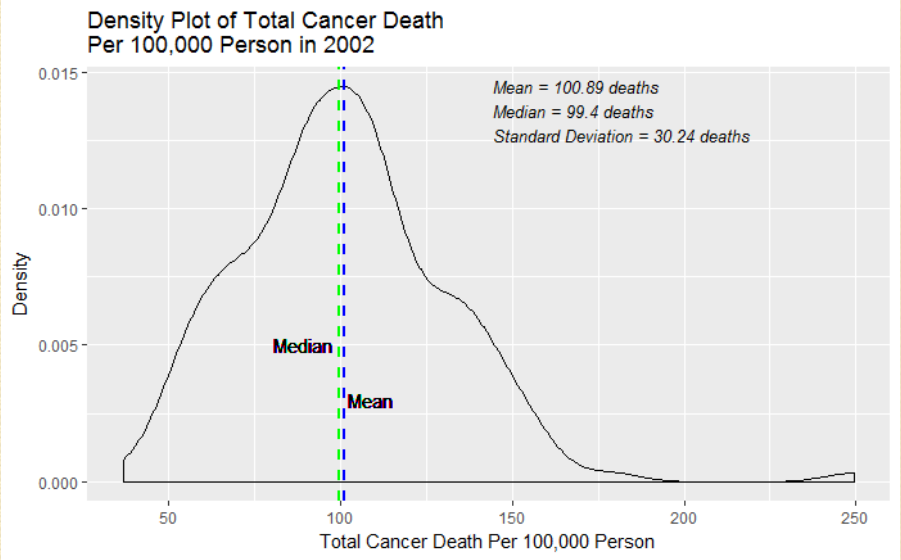
Food Supply



Mean: 2,507 Cal
Median: 2,427 Cal
Standard Deviation: 553 Cal

Exploratory Analysis

Cancer Death



Mean: 101 deaths
Median: 99 deaths
Standard Deviation: 30 deaths

In-depth Analysis

Pearson Correlation Coefficient

- Strength and direction of linear relationships between life expectancy and 6 independent variables

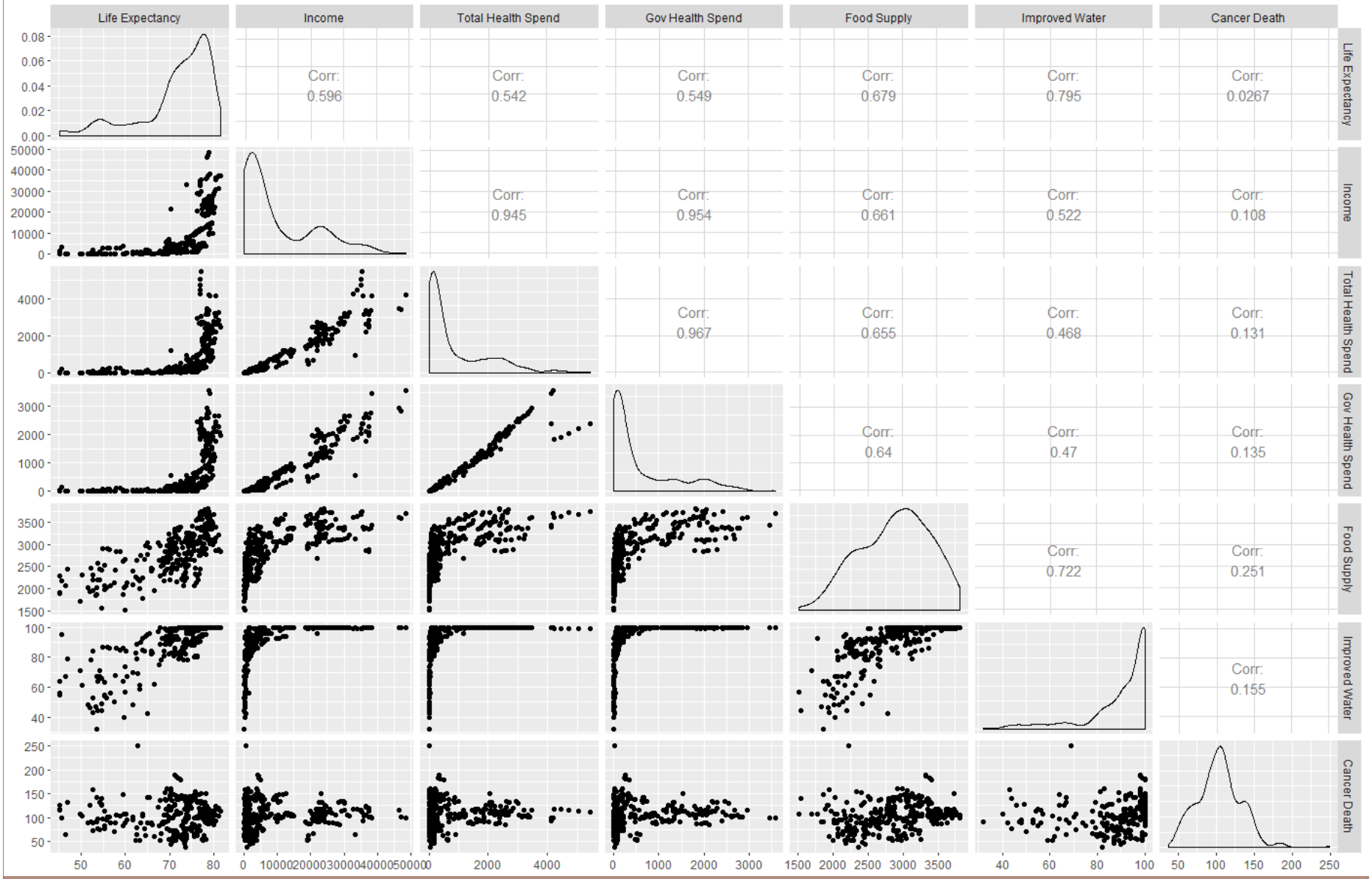
Multiple Linear Regression

- Model the relationship between life expectancy and 6 independent variables by fitting a linear equation to observed data
- Variable Selection
  - Best Subsets Regression
  - Stepwise Forward Regression
  - Random Forest
- 10-fold Cross-Validation

Pearson Correlation Coefficient
Correlation with Life Expectancy

Independent Variable	r	Significantly Correlated	Strength	Direction
Income	0.60	Yes	Moderate	Positive
Total Health Spending	0.54	Yes	Moderate	Positive
Government Health Spending	0.55	Yes	Moderate	Positive
Improved Drinking Water Source	0.80	Yes	Strong	Positive
Food Supply	0.68	Yes	Strong	Positive
Cancer Death	-0.03	No	Weak	Negative

Pairwise Comparison



Multiple Linear Regression

Variable	Amount Needed to Increase 1 Year of Life
Income	An Increase of US$7,003 in GDP Per Capita
Improved Drinking Water Source	An increase in 3% of country's population with an access to clean drinking water sources
Food Supply	An increase in 453 calories intake per person per day*
Cancer Death	A decrease in 29 cases of cancer death per 100,000 person per year

* If people faces food shortage due to extreme poverty or wars

Recommendations

- To close the inequality gap in life expectancy

Stakeholders	Recommended Actions
All Stakeholders	1. Monitor income-level, drinking water condition, and cancer death in communities. 2. Ensure that there is enough food supply for people.
Governments	1. Develop and implement policies to improve the average income, drinking water condition, and cancer death rate. 2. Ensure that citizens have 2,000-2,500 calories intake in average each day.

Recommendations

- To close the inequality gap in life expectancy

Stakeholders	Recommended Actions
Community Organizations and Health Facilities	1. Provide services to reduce low-income, drinking water condition, and high cancer death rate problems in underprivileged communities. 2. Examples: Job training, career center, environment and water source protection campaigns, and cancer early detection programs.

Potential Improvements

- More independent variables that are associated with life expectancy should be identified.
- Updated data should be used when available.