Calvin Ochi, Jessie Le, Leon Zhou

STAT 316

Professor Roback

30 Oct. 2019

**Project Stage I**

#### **Research questions**

How does life expectancy differ between countries, continents, and regions? What factors affect differences in life expectancy rates between countries? Which factors are most closely associated with adult mortality and infant death per 1000 people?

Rationale for each variable: Each country will be an observational unit associated with the year the summary statistics were collected. Our response variable is life expectancy, which we will use as a general measure of a country's quality of public health. A country’s socioeconomic status can be described by the countries development variables: percentage of GDP spent on healthcare, total money spent on healthcare in US dollars, gross-domestic product (GDP), human development index score(1-10), and average years of schooling. We can account for overdispersion between countries by accounting for variables like GDP and total money spent on healthcare. Health and immunization factors include: alcohol consumption in liters, body-mass index (BMI), and prevalence of thinness among children aged 5-19 as well as immunizations of Hepatitis B, Polio, and Diphtheria among one-year olds, and the number of measles cases per 1000 people. The values of these variables may illustrate the availability of immunizations and the effectiveness of current preventative medicine practices. Variables describing mortality include adult mortality, infant deaths, and the number of deaths for children below 5 years of age per 1000 people. These statistics may be used to describe how available healthcare providers are amongst the general population, as young children are more vulnerable to disease and often require more frequent medical attention. The number of deaths from HIV/AIDS per 1000 live births is included, since HIV/AIDS prevalence has been known to decrease life expectancy for adults and children born to women with HIV/AIDS.

Reason of interest: Life expectancy is a common statistic used to describe differences between countries in quality of life. We are interested in examining the general concept that a country's level of development, wealth, and education systems are good predictors of the average citizens quality of life and prosperity. From this project, we hope to gain insights about how to increase life expectancy in a country by assessing governmental support, population’s general lifestyle habits, and a country’s healthcare system.

#### **References**

1. Cardona, C. and Bishai D. (2018). The slowing pace of life expectancy gains since 1950. *BMC Public Health,* *18(151)*.

There have been immense breakthroughs in biomedical technology and increases in healthcare expenditure since the 1950s. However, the life expectancy at birth (LEB) does not reflect these adjustments. Globally, life expectancy has continued to increase annually, but improvements in LEB are slowing down, after controlling for socioeconomic and environmental factors. The decelerating rate of LEB growth is greatest among countries in the lowest starting LEB stratum. A plausible explanation is the change in country composition in each LEB stratum over different decades. Countries that strive for better health policies have moved up to higher LEB strata. Based on these observations, it is concluded that technology improvements and health spending are not the major contributions for LEB gains. Other changes in national health policies should be investigated to determine what is slowing the rate of life expectancy at birth rates.

This reference is relevant to our research question because we are also looking into similar predictors for life expectancy change. This study can give us an idea of what trends we may expect to observe between life expectancy and the variable total expenditure of GDP.

1. Douglas Barthold D., Nandi A., Mendoza Rodríguez J. M., & Heymann J. (2014). Analyzing Whether Countries Are Equally Efficient at Improving Longevity for Men and Women.”*Am. J. Public Health, 104(1)*.

Douglas et al. models life expectancy versus total health expenditure per capita, and other controlling variables, to determine the efficiency of health spending allocation among 27 countries apart of the Organization for Economic Cooperation and Development (OECD) during the 1990s and 2000s. The effect of health expenditure varies between countries, but the majority of responses indicate that life expectancy increases with health expenditure.

Men may experience greater increases in life expectancy than women per the same amount of increase in health expenditure. To reduce this disparity, redistribution of resources may be necessary to increase women’s life expectancy at a similar rate to men.

This reference describes different effects on life expectancy across countries for men and women. Further variables which are relevant to healthcare policies should be taken into consideration.

1. Khan, A., Khan S., &Khan, M. (2016, November). Factors effecting life expectancy in developed and developing countries of the world (An approach to available literature). *International Journal of Yoga, Physiotherapy and Physical Education, 1(1),* 04-06.

This article has reviewed various published articles that are about life expectancy between developing countries and developed countries and summarizes important factors that may affect life expectancy in a specific country. These factors include: financial income, literacy, accessibility to health facilities, and one’s daily lifestyle. This article has provided us with some background knowledge about the differences in life expectancy between different countries, regions, and country’s status level (developed vs. developing).

1. Mäki, N., Martikainen, P., Eikemo, T., Menvielle, G., Lundberg, O., Östergren, O.,...the EURO-GBD-SE consortium. (2013, October). Substudy 1: Educational differences in disability-free life expectancy: a comparative study of long-standing activity limitation in eight European countries. *Soc Sci Med, 94,* 6-20.

This article examined the effect different levels of education has on disability-free life expectancy across eight european countries. It concluded that for those european countries they examined, higher education has led to longer life expectancy and better health in general.

The variable schooling is included in our dataset. So, we plan to investigate if education is a

good predictor for life expectancy and whether our results would be consistent with this article.

.

1. Vaupel, J. W., Zhang, Z., &Raalte, A. A. (2011). Life expectancy and disparity: an international comparison of life table data. *BMJ Open, 1,* 1-6.

This study looked at the relationship between life disparity and life expectancy across 40 developed countries from 1840 to 2009. It examined the relationship between prevention of premature death and its effect on life expectancy and concluded that populations with high life expectancy also have lower life disparity and vice versa. In addition, it has found that countries with high life expectancy have been successfully reducing premature deaths for all ages.

This article provided insights into how premature deaths could related to life expectancy. In our dataset, we have variables related to mortality such as adult mortality (death between age 15 to 60) and infant death, which could potentially serve as predictors when modeling life expectancy.

#### **Variable Chart**

### **Variable summary Table**

| **Variable names** | **Description** | **Role** | **Type** | **Values** | **Units** |
| --- | --- | --- | --- | --- | --- |
| Country |  | Explanatory | Character | Country name | N/A |
| Year | Summary statistics taken each year | Explanatory | Numerical | 2000-2015 | years |
| Status | Country status | Explanatory | Categorical | Developing, developed | N/A |
| Lifeexpectancy | Average lifespan by year and country | Response | Numerical | 36.3-89 | Age in years |
| AdultMortality |  | Response | Numerical | 1-723 | /1000 population |
| infantdeath | Number of infant death | Response | Numerical | 0-1800 | /1000 population |
| Alcohol | In pure alcohol term | Explanatory | numerical | 0.01-17.9 | litre |
| percentageexpenditure | Expenditure on health as a percentage of Gross Domestic Product per capita(%) | Explanatory | numerical | 0-19.5k | Currency |
| HepatitisB |  | Explanatory | numerical | 1-99 | % among 1-year-old |
| Measles |  | Explanatory | Numerical | 0-212k | /1000 population |
| BMI | Average BMI | Explanatory | numerical | 1-87.3 | BMI index |
| under-fivedeaths | Children under age 5 | Explanatory | Numerical | 0-2500 | /1000 population |
| Polio | Polio (Pol3) immunization coverage among 1-year-olds (%) | Explanatory | Numerical | 3-99 | % among 1-year-old |
| Totalexpenditure |  | Explanatory | Numerical | 0.37-17.6 | % |
| Diphtheria |  | Explanatory | Numerical | 2-99 | % among 1-year-old |
| HIV/AIDS |  | Explanatory | Numerical | 0.1-50.6 | per 1000 live births HIV/AIDS (0-4 years) |
| GDP |  | Explanatory | Numerical | 1.68-119k | USD |
| Population | Population of the country | Weights | Numerical | 34-1.29b | counts |
| thinness1-19years | Prevalence of thinness among children and adolescents for Age 10 to 19 | Explanatory | Numerical | 0.1-27.7 | % |
| thinness5-9years | prevalence of thinness among children for Age 5 to 9 | Explanatory | Numerical | 0.1-28.6 | % |
| Incomecompositionofresources |  | Explanatory | Numerical | 0-0.95 | Human development index |
| Schooling |  | Explanatory | Numerical | 0-20.7 | years |

#### **Analysis plan**

One of our references above mentions that population density is an important factor predicting life expectancy. Thus, we are planning to incorporate land area of each country into our dataset and calculate population density.

We plan to group countries into seven continents to examine regional level for multilevel models.

We will model life expectancy in a multilevel regression analysis among 193 countries between 2000-2015. We will examine immunization factors, mortality factors, economic factors, social factors and other health related factors.

Potentially, we could classify life expectancy into two groups: high life expectancy and low expectancy. Then, we could use binomial regression to model the odds of of each country having high life expectancy.

There might also be some interaction effects when predicting life expectancy, i. e. status of the country (developed vs. developing) and average years of schooling. Potential interaction terms might be introduced when modeling.

In addition, we could potentially model adult mortality and infant deaths using poisson regression (ZIP, Hurdle).

Each country has multiple observations taken for each sequential year, so the independence assumption may be violated. We can use variables such as years as a level for a multilevel regression model. Then we would regroup countries into different continents to produce another level for our multilevel regression model.

#### **Data source**

We obtained this data from Kaggle which was originally sourced from the Global Health Observatory (GHO) data repository under World Health Organization (WHO). Included in this dataset are average life expectancy values by country and year with multiple other explanatory variables like BMI, alcohol consumption, population, etc. The dataset has been stored in the project folder on R.

Link to the dataset: <https://www.kaggle.com/augustus0498/life-expectancy-who>