Cierra Wickliff

Data Science for Political Science

**Introduction**

My research question: does the income of a country correlate to health? To measure health, I look at the life expectancy of a country. I hypothesize that a higher income group correlates to a longer life expectancy. This is a descriptive claim. Examining the relationship between income and life expectancy is important for a variety of reasons, such as helping determine a country's population trends over time. Additionally, life expectancy is a good indicator of health within a country. If someone is trying to improve the health of a country, looking at the income group can be helpful as there could be a confounding variable between the two.

Because life expectancy has increased over time, I only looked at entries from 2019. There were four income classifications: Low income, Lower middle income, Upper middle income, and High income. I compared the mean life expectancy for each income group. Each income group varies between 5-8 years between each other, showing a correlation between life expectancy and income. This means that the income group could be either a causal or confounding variable for life expectancy.

**Background**

It is important to find if there is a relationship between life expectancy and income for health, economic, and socio-political reasons. Finding if there is a correlation can lead to further research questions. For example, what is the confounding variable between income and life expectancy, or is there even a causal relationship? It’s possible that a better income can enable a country’s people to have better access to healthcare resources, less dangerous jobs, or better nutrition, all of which could affect life expectancy.

Surprisingly, a study found that “socio-economic factors like per capita income, education, health expenditure, access to safe water, and urbanization cannot always be considered to be influential in determining life expectancy in developing countries” (Kabir 2008). This means that for many countries, a socio-economic indicator like income group would not be a correlating factor for life expectancy. Another study found “that doubling annual pharmaceutical expenditures adds about one year of life expectancy for males at age 40 and slightly less than a year of life expectancy for females at age 65” (Shaw et al., 2005). Although not a large increase in life expectancy, a year added is a difference. Doubling pharmaceutical expenditures would imply a country has a larger disposable income for pharmaceuticals and thus a higher income group.

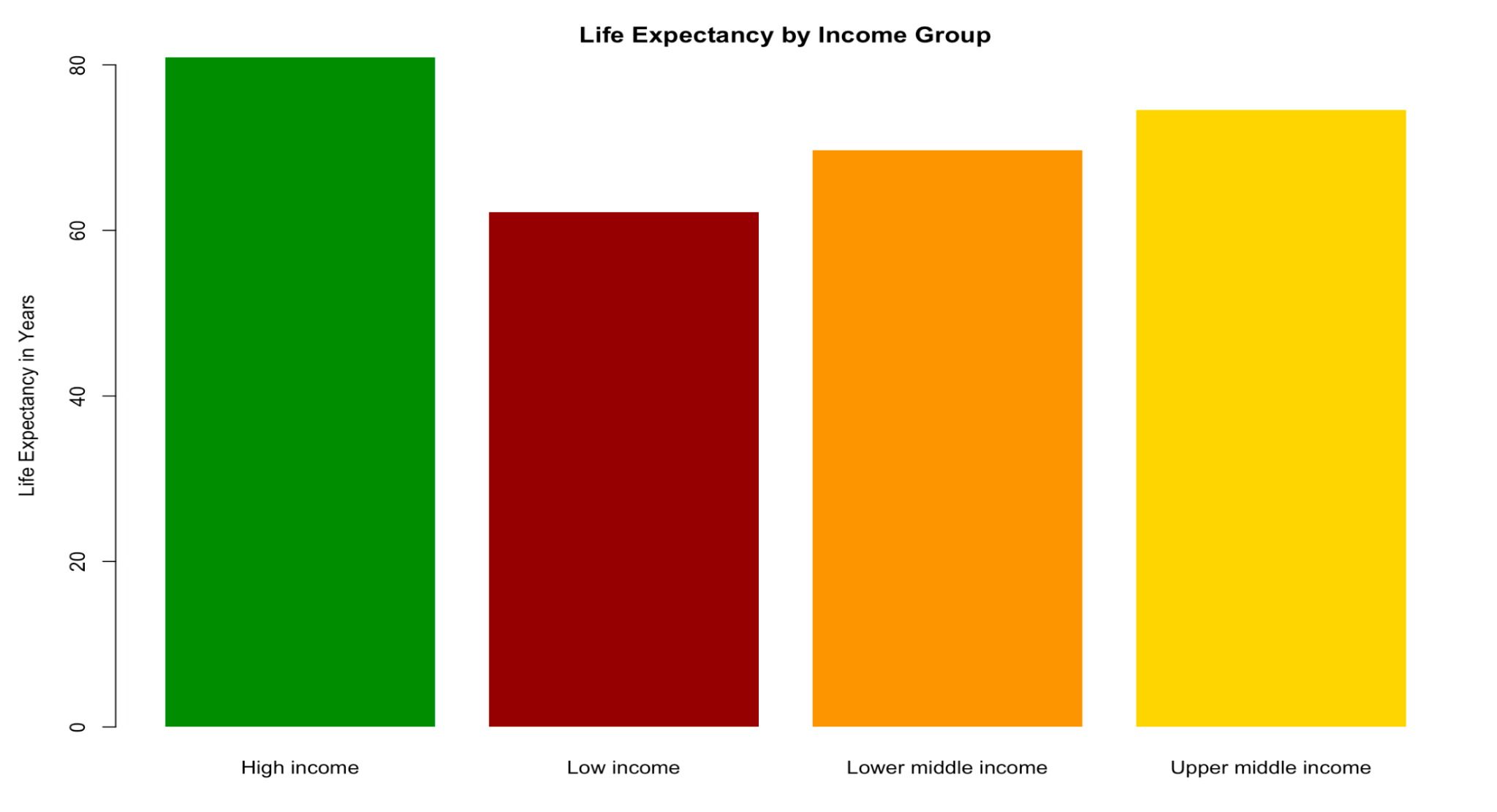
**Data and Approach**

I used data from the World Bank 1980-2020 from 216 countries on dozens of world development indicators. The entire data set has 8856 rows. Each row represents a country in a given year. The income statistic is assigned based on the GNI per capita from the previous year. Because life expectancy has increased over time, I decided to only use data that had the income group classified from the year 2019. I subsetted the data, which then included 197 countries. My key variables were income group, coded as “income,” and life expectancy, coded as “SP.DYN.LE00.IN.” I wanted to look at life expectancy by income group. To visualize the data, I did a barplot of each income group’s median life expectancy. I specifically chose median instead of mean so as not to skew the data due to outliers.

**Results**

Low income countries have a life expectancy of 62.82 years. Lower middle income countries have a life expectancy of 68.30 years. Upper middle income countries have a life expectancy of 73.91 years. High income countries have a life expectancy of 80.12 years. Each income group varies between 5-8 years between each other, showing a correlation between life expectancy and income. This means that the income group could be either a causal or confounding variable for life expectancy, supporting my hypothesis.

Overall, the mean life expectancy is 72.93 years. The median is 74.24 years. The lowest income group is more than a decade lower than the mean life expectancy, and the highest income group is more than 7 years higher than the mean.



The figure shows that the bar for low income is lowest and the bars increase as the income group increases.

**Conclusion**

My research question was does the income of a country correlate to health? My results align with my expectations that a higher income for a country correlates with a higher life expectancy. These findings are important because it implies that income is a cause or at the very least correlated with life expectancy and thus health.

Works Cited

Kabir, Mahfuz. The Journal of Developing Areas, Vol. 41, No. 2 (Spring, 2008), pp. 185-204.

Shaw, et al. Southern Economic Journal, Vol. 71, No. 4 (Apr., 2005), pp. 768-783.

World Bank 1980-2020.