The Relationship between Household Income and Life Expectancy

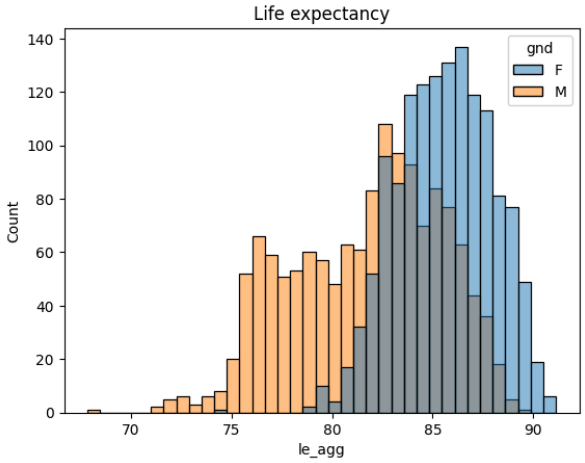
Econ 0150 Final Project
Megan Suess and Amina Taivan

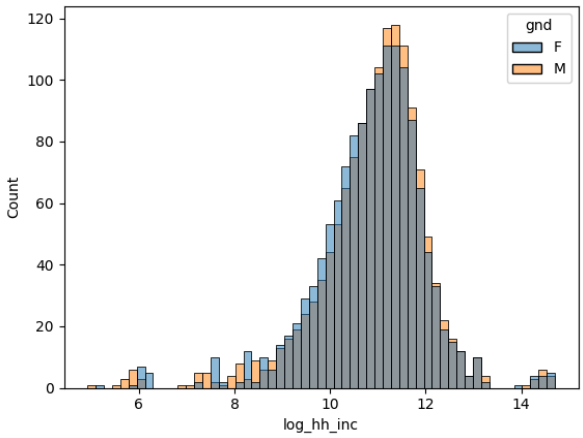
Introduction

The question being explored is whether household income influences life expectancy.

Data Description

The main variables include household income and life expectancy. The data was aggregated by income percentile and gender. Household income was converted to a logarithmic scale to allow for comparison. Household income values in the data span from 1999 to 2014; all household incomes were converted to equate to 2012 values using the consumer price index.





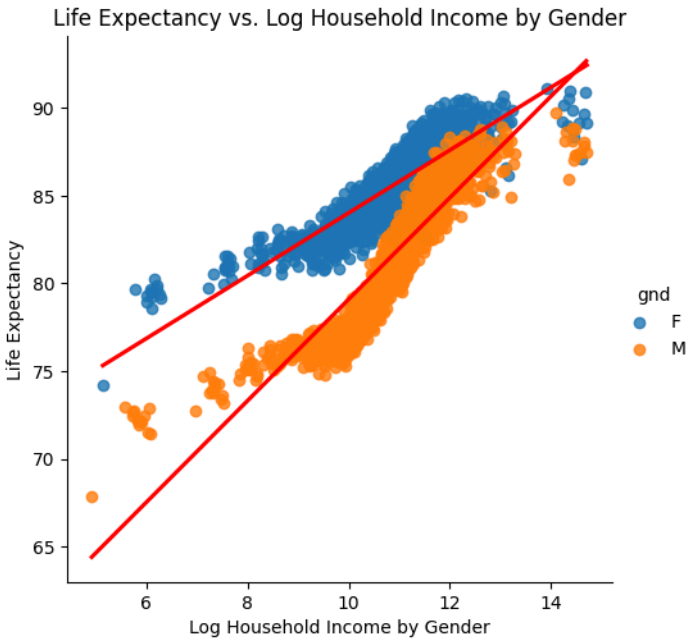
Data was found on the Health Inequality Project website. The data contains life expectancy estimates at age 40 by income and sex at various levels of geographic and temporal aggregation. The data is released to the public domain under a CC0 license and was sourced from JAMA Health Disparities, published online on April 26, 2016. There are no issues with reliability or source credibility.

Methodology

The general linear model was used to analyze the relationship between log household income and life expectancy.

Life Expectancy = B_0 + B_1*gender + B_2*log_income + B_3 *(gender x log_income) + e

The predictor variable is log household income, and the outcome variable is life expectancy. Limitations of the data might include the nonlinearity in the results because life expectancy rates are nonlinear.

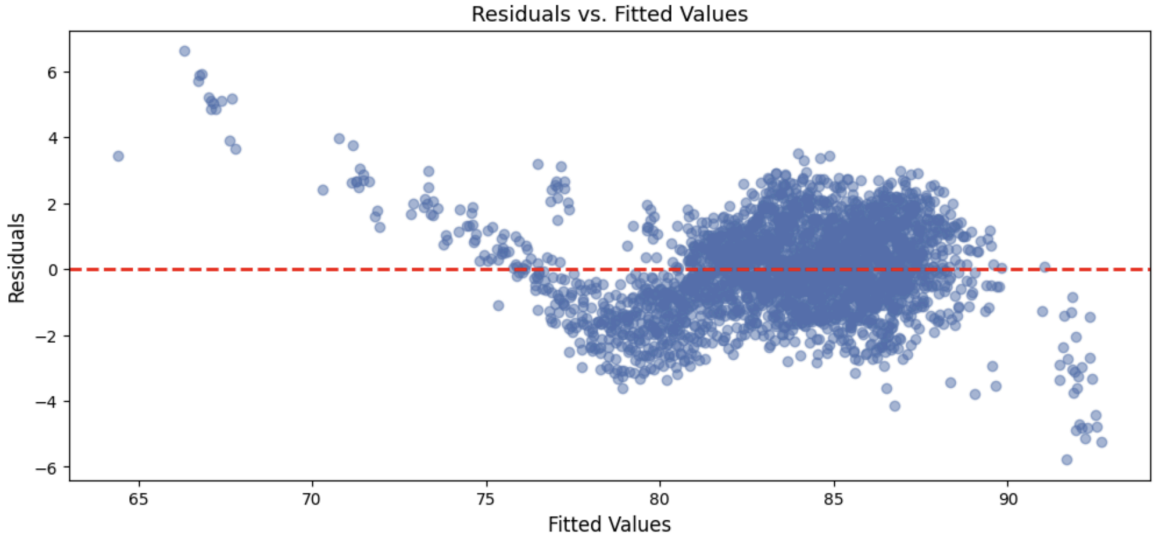


Analysis and Results

The intercept of 66.12 in simple terms means that if the log-income is 0(income = $1), the average female life expectancy would be about 66 years. On average, males have a lower life expectancy than females by 15 years, indicated by the negative coefficient on the male indicator. A 1% increase in household income equates to a life expectancy increase of about 0.018 years(6.6 days) for females. For males, the increase is about 0.029 years(10.6 days). The model reports a p-value of 0.000 for log household income, meaning it is statistically significant. This means we can reject the null hypothesis of no association between the two variables. Changes to household income are meaningfully related to changes in life expectancy.

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	66.1240	0.355	186.057	0.000	65.427	66.821
gnd[T.M]	−15.9532	0.496	−32.154	0.000	−16.926	−14.980
log_hh_inc	1.7893	0.033	54.966	0.000	1.726	1.853
log_hh_inc:gnd[T.M]	1.1034	0.045	24.329	0.000	1.015	1.192

Residuals are mostly randomly scattered and have fairly constant variance at each level of the fitted value. Some limitations of the model are that it fits lower and middle-income households better than higher-income households. There is more variability in high-income levels, and the model is less accurate for high-income households. It exhibits a curved pattern at higher income levels (heteroskedasticity).



Conclusion

Our model examines the relationship between household income and life expectancy, accounting for gender. The scatterplot we created showed a clear positive association between household income and life expectancy. The analysis illustrates that individuals with higher household income tend to live longer. This can be seen for both males and females. The model concludes that a key factor in higher life expectancy is income; each increase in log income is associated with an increase in longevity. There were a few limitations. The residual plot was mostly centered around zero, but points were spread out at higher income levels, meaning the model didn't fit all income groups. Additionally, life expectancy varied more at higher incomes, so the model couldn't capture the full pattern.

References

https://healthinequality.org/data/