

Association between Income and Life Expectancy in the United States

Statistical Project using Excel

1923114

Introduction

In this project I will study the relationship between life expectancy and income, as well as the factor gender plays into income distribution. In the first part of the project, I will investigate the income distribution of households for each gender. Then I will investigate the relationship between life expectancy and income for the female population in the data. The final part of the project will be to investigate the life expectancy distribution for females by state in USA.

1. Investigation of the income distribution based on gender

The first step for interpreting the data was to separate it into income bands instead of percentiles. After that was done the data was summarized in Table 1 where we can see the total number of households for each income band separated by gender. The data from the table was also represented in Figure 1, where the trend becomes clearer. The number of females is clearly greater than the number of males for each income band. The number of households in greater income bands is greater for both males and females, but it seems to be increasing faster for males.

To give a more qualitative measure of the distribution between genders the percentage distribution of households in each income band by gender was calculated in Table 2 and represented in Figure 2. As concluded from Table 1 and Figure 1, we can see that the percentage of females is greater for each income band. However, we can now see that the percentage of males is increasing for greater income bands, with a 0.22 % difference between the percentage of males in income band 10 and income band 1.

Sum of count	incband ▾										
gnd ▾	1	2	3	4	5	6	7	8	9	10	Grand Total
M	68.166.227	68.465.824	68.698.827	68.892.603	69.037.182	69.154.434	69.246.873	69.336.489	69.424.484	69.533.150	689.956.093
F	71.354.771	71.592.203	71.700.737	71.770.338	71.834.467	71.895.339	71.955.900	72.012.247	72.075.400	72.139.723	718.331.125
Grand Total	139.520.998	140.058.027	140.399.564	140.662.941	140.871.649	141.049.773	141.202.773	141.348.736	141.499.884	141.672.873	1.408.287.218

Table 1. Pivot Table representing the number of households in each income band by gender

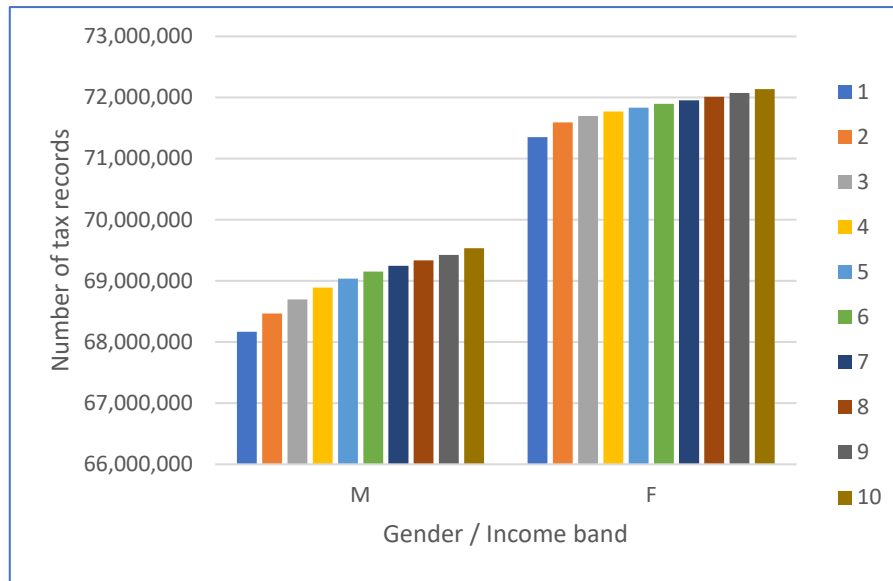


Figure 1. Chart representation of the number of households in each income band / gender

Sum of count	incband										
gnd	1	2	3	4	5	6	7	8	9	10	Grand Total
M	48,86%	48,88%	48,93%	48,98%	49,01%	49,03%	49,04%	49,05%	49,06%	49,08%	48,99%
F	51,14%	51,12%	51,07%	51,02%	50,99%	50,97%	50,96%	50,95%	50,94%	50,92%	51,01%
Grand Total	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

Table 2. Pivot Table representing the percentage distribution households in each income band by gender

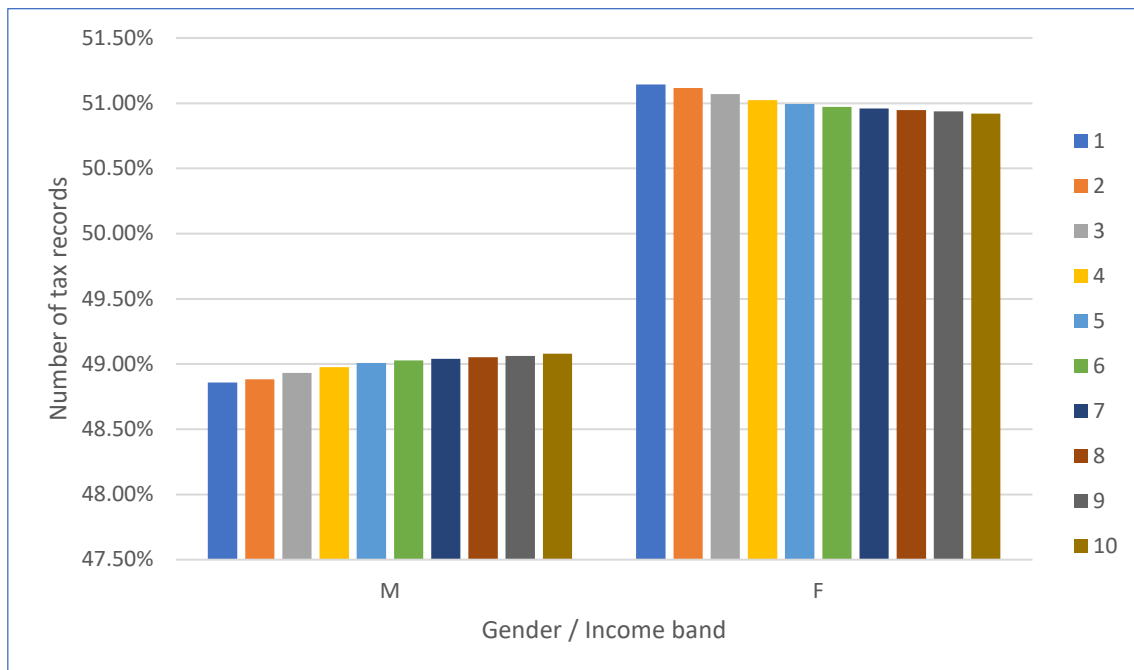


Figure 2. Chart representation of the percentage distribution in each income band by gender

2. Investigation of the life expectancy based on income for females

The first step in data processing for this investigation was to calculate the average decadic logarithm of the life expectancy for each income band. The results of this calculation can be seen in Table 3. From the table we can see that the life expectancy is increasing as income is increasing. The data was also represented in a graph in Figure 3 where we can see that the decadic logarithm of the life expectancy follows an approximately linear trend. The nonlinearity seems to indicate that the relationship between life expectancy and income is stronger first and last three income bands, and weaker for income bands closer to the center.

incband	Average of log le
1	1,909124727
2	1,916394554
3	1,92075634
4	1,924628981
5	1,928114929
6	1,931125271
7	1,934105639
8	1,937158509
9	1,941189154
10	1,946439785
Grand Total	1,928903789

Table 3. Decadic logarithm of the life expectancy by income band

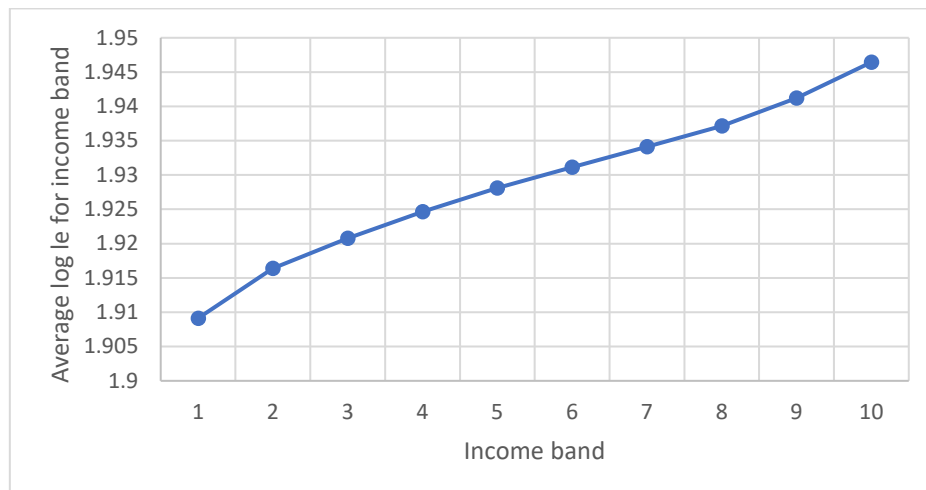


Figure 3. Graphical representation of the relationship between the decadic logarithm and income band

3. Investigation of the distribution of life expectancy of females by state in USA

The data for this investigation was given as life expectancy in each state of the USA and it was separated in quartiles. To obtain the descriptive statistics information needed for this investigation I used the Data Analysis ToolPak. The resulting data is represented in Table 4 below.

The first quartile has a mean life expectancy of 81.90 years with a range of 4.42 years. From the values of the kurtosis and skewness I can deduce that the distribution is flatter than a normal distribution and slightly positively skewed.

For the second quartile I suspect there was an error in the data, since all the life expectancy values are equal to 83.87, leading to a range of 0 and a lack of a distribution to be described.

The third quartile has a mean of 85.87 years and the range is 2.79 years. This time the distribution shows a slightly stronger and negative skewness, while the kurtosis is close to 0 which suggests that the distribution is very close to a normal distribution.

The third quartile has a mean life expectancy of 87.65 years and a range of 2.8 years. The distribution is slightly positively skewed and slightly narrower than the normal distribution.

<i>le-q1-F</i>		<i>le-q2-F</i>		<i>le-q3-F</i>		<i>le-q4-F</i>	
Mean	81,8950722	Mean	83,8563385	Mean	85,86858518	Mean	87,64737701
Standard Error	0,14600899	Standard Error	0	Standard Error	0,08588964	Standard Error	0,07986669
Median	81,6261139	Median	83,8563385	Median	85,86901855	Median	87,67614746
Mode	#N/A	Mode	83,8563385	Mode	#N/A	Mode	#N/A
Standard Deviation	1,04271278	Standard Deviation	0	Standard Deviation	0,61337472	Standard Deviation	0,57036225
Sample Variance	1,08724994	Sample Variance	0	Sample Variance	0,376228547	Sample Variance	0,325313096
Kurtosis	-0,5139063	Kurtosis	#DIV/0!	Kurtosis	0,082871978	Kurtosis	0,357296093
Skewness	0,29562027	Skewness	#DIV/0!	Skewness	-0,33606818	Skewness	0,213574965
Range	4,424263	Range	0	Range	2,792373657	Range	2,800201416
Minimum	80,0119934	Minimum	83,8563385	Minimum	84,0636673	Minimum	86,09049988
Maximum	84,4362564	Maximum	83,8563385	Maximum	86,85604095	Maximum	88,89070129
Sum	4176,64868	Sum	4276,67326	Sum	4379,297844	Sum	4470,016228
Count	51	Count	51	Count	51	Count	51

Table 4. Descriptive statistics for the life expectancy by quartiles across all states

Conclusion

In this project I found that the number of females is greater than the number of males for every income band, showing just a slight increase in the percentage of males in higher income bands.

Regarding the relationship between life expectancy and income the I found that the life expectancy is clearly increasing with income. The relationship is approximately linear, having a slightly higher gradient for the lower most and higher most income bands.

Lastly, the investigation into the distribution of life expectancy of females by state in USA found that the distribution is close to a normal distribution.