Intro to Stats - Final Project

Todd Estep

11/17/2024

Minimum Wage & Cost of Living in the USA

Introduction

The minimum wage and its ability to support someone who works full-time to live and support themself with dignity is important to me. I beleive all people should be able to work and support themselves, and that they should benefit from the fruit of their labor. Oftentimes, the benefit of their labor is lost to exploitation by others in privileged positions. Of course employers should gain from their investment in their employees, but if that investment does not allow the employee to live with dignity, then the employer should not be able to benefit from their labor.

The minimum wage affects nearly everyone, even those with salaries well above it, because it affects the cost of basic goods, and higher salaries are a reflection of the minimum wage. I became interested in this topic when I heard a story of someone who had saved money by working a minimum wage job in the 1990's, and I thought how impossible that would be today. I also reflected back on my own past, including times when I worked jobs for less than the minimum wage as a teenager with little options. I also knew that the federal minimum wage hasn't increased since 2009, with heavy inflation in recent years, especially in rent and housing prices.

I think data science is perfect for analyzing this because it there is lots of public data on this topic, including what the minimum wage is in different areas, the cost of living in those areas, how many people are living on a minimum wage, public opinion on minimum wage and more. By analyzing the large datasets, we can make sense about the severity of the issue and make useful graphs that are easy to understand and improve our knowledge in the topic.

Research Questions

1. What is the minimum wage?

The federal minimum wage is $7.25 per hour, but many places in the US have set their own minimum wage by state or county. I would like to try and zoom in to the county level and see what the minimum wage is in each of the 3,144 counties in the US.

2. What is the cost of living?

The cost of living varies a lot by local econonimies. For this project I would like to zoom in to the county level, and for that I will need the cost of living in of the 3,144 counties in the US.

3. Which counties are most affected by the gap of minimum wage and cost of living?

Are there places in the US that are easier to live while earning minimum wage? Would they be traditionally rural or urban areas? Are there nearby counties with drastically different cost of living and wage ratios?

4. Does the minimum wage affect the cost of living?

I am curious if a higher minimum wage has a direct affect on the cost of living. Does an increase in labor costs make living in the area more difficult?

5. Does the minimum wage affect the average salary?

I would like to know how much the minimum wage affects other wages. Are higher average salaries a reflection of a higher minimum wage?

6. What is the best way to communicate this information?

I am interested in using maps to answer these questions, but I'm also curious if there are other graphs that will be useful in communicating this information. When I combine the datasets, I am interested about what interactions I can find.

Approach

I plan on using a variety of approaches in answering these questions, but primarily I would like to zoom in on the county level in the US. With minimum wage being set at a local level (and even some individual cities), and the cost of living varying widely too, I think analyzing the 3,144 counties in the US provides a good depth in the data where I can gain a variety of insights.

Because the counties refer to geographic locations, I think using maps will be essential in analyzing this data. It will also be useful in seeing if there are larger groups where the problem is more severe, or if there are areas that are quite livable on the minimum wage.

In addition to geography, I will also look at how the minimum wage affects people by demographics, if certain groups are more affected by others. I am also interested in looking at minimum wage historically and how it compares to inflation. Finally, I'd also like to see how widespread minimum wage is, how many people it affects, and how many people earn below minimum wage.

Data

The Econonimc Policy Institute - Minimum Wage Tracker
URL: www.epi.org/minimum-wage-tracker/#/min_wage
Published Oct 1, 2024. Accessed Noc 17, 2024.

This site has a list of all the minimum wages, including highlighting differences. I will have to sort the data into a csv spreadsheet

The Economic Policy Institute - Cost of Living by County
URL: www.epi.org/resources/budget/budget-map/
Published Jan 2024. Accessed Nov 17, 2024.

This has a downloadable csv with data covering the cost of living by county, including comparing costs of single adults vs families.

The US Bureau of Labor Statistics - Characteristics of Minimum Wage Workers
URL: www.bls.gov/opub/reports/minimum-wage/2023/
Published May 2024. Accessed Nov 17, 2024.

This site offers lots of data on who is working minimum wage by demographics, including sex, age, race and industry. It will be good in helping me determine if the minimum wage is disproportionately exploiting and affecting a group of people.

USA Facts: Minimum Wage in America - How Many people are earning $7.25 an Hour
URL: usafacts.org/articles/minimum-wage-america-how-many-people-are-earning-725-hour/
Published Feb 27, 2024. Acccessed Nov 17, 2024.

This site offers a good historical comparison of the federal minimum wage when adjusted for inflation, with a downloadable csv.

Required Packages

- **readr:** To import the data.
- **dplyr:** To edit the data, including the filter feature to focus on single adults, mutate to add columns like the minimum wage annual salary and the annual salary to cost of living ratio.
- **usmap:** To make county level maps of the USA.
- **ggplot2:** To edit the maps of the USA with ggplot2 elements, like legend text size and colors.
- **gridExtra:** To fit the graphs nicely on a page.

Table and Plot Needs

- **Tables:** I will need a table with the minimum wage for each county, along with the cost of living for each county. I will also need the fips number for each county, as that is how the data is graphed. The table will also need to be a complete list of all the counties in the USA, this includes 3144 counties in the USA, so it will be a large table.
- **Plots:** I will neep geographic map plots of the USA. I will need them to show each county, and then I will color each county according the data.

Questions for Future Steps

Right now my biggest concern is working with map graphs. I have not done that before, but my initial research suggests that I should be able to do it with practice. I am confident in my ability to organize the data into something manageable, like focusing on single adults or finding which demographics to focus on.

I am a little concerned that my idea to focus on the county level may be too ambitious, but I think I will be able to do it, because many states have statewide minimum wages, I will just need to make adjustments for the differences, which isn't too many counties. My cost of living dataset already includes the cost of living by county, so I will just need to make a subset of that for single adults.

Milestone 2

How did you import and clean your data?

To simplify the data, and increase focus and clarity, I decided to focus on single adults with no children. My original data had a family column with different costs for different family sizes. I used the dplyr filter method to filer the family to "1p0c", denoting 1 parent no children. Since it originally included 4 different family sizes, this reduced the data by 75%. Once the family sizes were set to all be single adults, the column was redundant so I deleted it.

In the cost of living data, my original data included monthly and annual costs. Again, to increase focus and clarity, I decided to work with the annual costs. I used the dplyr select function to delete the monthly costs. The original dataset also included categories for annual costs, like food, transportation, taxes and more. I really only wanted to focus on the total cost, so I used the dplyr select to remove the individual categories and focused on the total annual cost.

```
min_wage_data <- read_csv("min_wage_data.csv")
min_wage_data <- min_wage_data %>% mutate(median_ratio =
                                            median_family_income/a_income)
summary(min_wage_data$min_wage)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   7.250   7.250   7.250   9.711  12.000  17.500
```

A seperate dataset had the minimum wages by county, but it wasn't in a csv file, so I added that data manually in Excel. With many states having the same minimum wage with only a few adjustments for individual counties, it wasn't too difficult to add a minimum wage for all 3,144 counties.

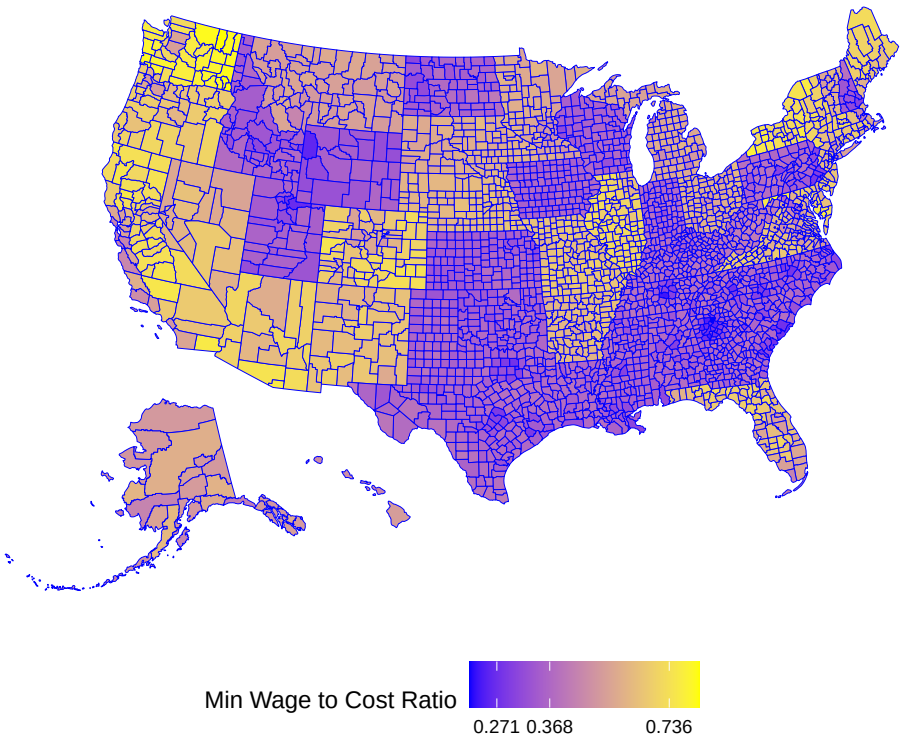
Once the minimum wage data was in, I could add some columns with some simple calculations, like the annual salary on a minimum wage. I used a standard 8 hour work day and 220 work days in a year. The dplyr mutate function added the column easily using $annualsalary = minwage * 220 * 8$.

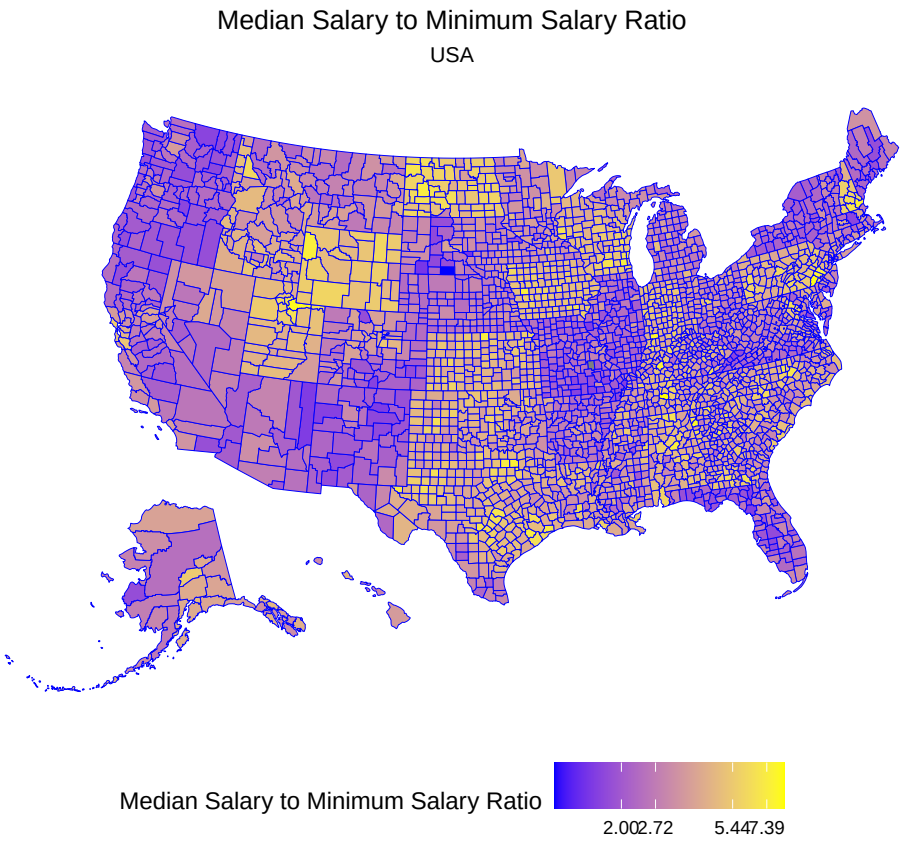
I also wanted the ratio of annual salary to annual cost, so I used dplyr mutate to add a column, annual wage to annual cost of living $wageratio = \frac{annualsalary}{annualwage}$.

My data also included the median income by county, so I decided to compare that to the minimum salary to see if there were some states that had a large ratio of median salary to minimum salary. I used dplyr mutate to add another column $medianratio = \frac{medianannualsalary}{minimumannualsalary}$. This may be useful to see how much median salaries depend on the minimum as a baseline, and to see if there are some states or regions where the difference is much stronger.

Annual Min Wage to Annual Cost of Living

USA





What does the final data set look like?

The final data set has 12 columns: State, Fips, County, Annual Minimum Wage, Median Income, Number of Counties in State, Cost of Living Rank by State, Median Income Rank by State, Affordability Ranking by State, Median Income Ranking by State, Minimum Wage, Annual Minimum Wage, Annual Ratio of Minimum Wage to Cost of Living, Annual Ratio of Median Salary to Minimum Wage Salary.

What information is not self-evident?

When I initially made the graph, it was that the biggest differences were between states and not between counties. It is hard to see what the metropolitan areas are on the map, but it is easy to see the state lines on the map.

Another important piece of information is that even at the highest values the ratio between minimum salary and cost of living is 0.74 on the legend. My data shows a range between a minimum of 0.230 and a maximum of 0.874. This would mean that nowhere in America does the minimum wage cover the cost of living. It is also interesting to note that 7 out of the 10 lowest ratios for cost of living are in the state of Georgia, and 9 out of the 10 highest ratios for cost of living are in Washington state,

How do you plan to slice and dice the data?

I cut the tables to 6 columns: State, County, Annual Cost of Living, Minimum Wage, Annual Minimum Wage Salary and Annual Ratio of Minimum Salary to Cost of Living. I then wanted to take a look at the top and bottom 10 counties by ratio of minimum wage salary to cost of living.

state	county	a_total	min_wage	a_income	a_ratio
GA	Forsyth County	65340	7.25	15080	0.2307928
TN	Williamson County	61980	7.25	15080	0.2433043
GA	Fayette County	61056	7.25	15080	0.2469864
GA	Cherokee County	59664	7.25	15080	0.2527487
GA	Gwinnett County	58152	7.25	15080	0.2593204
WY	Teton County	57876	7.25	15080	0.2605571
GA	Paulding County	57144	7.25	15080	0.2638947
GA	Cobb County	56652	7.25	15080	0.2661865
UT	Summit County	56484	7.25	15080	0.2669783
GA	Fulton County	55800	7.25	15080	0.2702509
NC	Buncombe County	55596	7.25	15080	0.2712425
GA	Henry County	55308	7.25	15080	0.2726550

state	county	a_total	min_wage	a_income	a_ratio
WA	Ferry County	38724	16.28	33862.4	0.8744551
WA	Garfield County	39168	16.28	33862.4	0.8645425
WA	Asotin County	39252	16.28	33862.4	0.8626923
WA	Whitman County	39576	16.28	33862.4	0.8556297
WA	Wahkiakum County	39648	16.28	33862.4	0.8540759
WA	Okanogan County	39732	16.28	33862.4	0.8522702
WA	Stevens County	39936	16.28	33862.4	0.8479167
WA	Pacific County	39984	16.28	33862.4	0.8468988
MD	Allegany County	36948	15.00	31200.0	0.8444300
WA	Spokane County	40296	16.28	33862.4	0.8403415
WA	Pend Oreille County	40596	16.28	33862.4	0.8341314
WA	Lincoln County	40860	16.28	33862.4	0.8287420

How could you summarize your data to answer key questions?

I think the most important data I've found is that no county in the USA has a minimum wage salary that covers the cost of living. Even the counties with the highest ratio of minimum salary to cost of living only cover around 85% of the cost of living. This would mean that nowhere in America is a minimum wage job survivable. Some counties with the lowest ratios only cover about 25% of the cost of living.

Another key takeaway is that 9 out of the top 10 counties for living on minimum wage are in Washington state, which has the highest minimum wage in the country at $16.28 per hour. There is also one state responsible for a large portion of the bottom counties, Georgia, with 7 out of the bottom 10 counties. Georgia state technically has a minimum wage of $5.15 per hour by state law, but since federal minimum wage is $7.25 per hour, that applies and Georgia's state law is ignored.

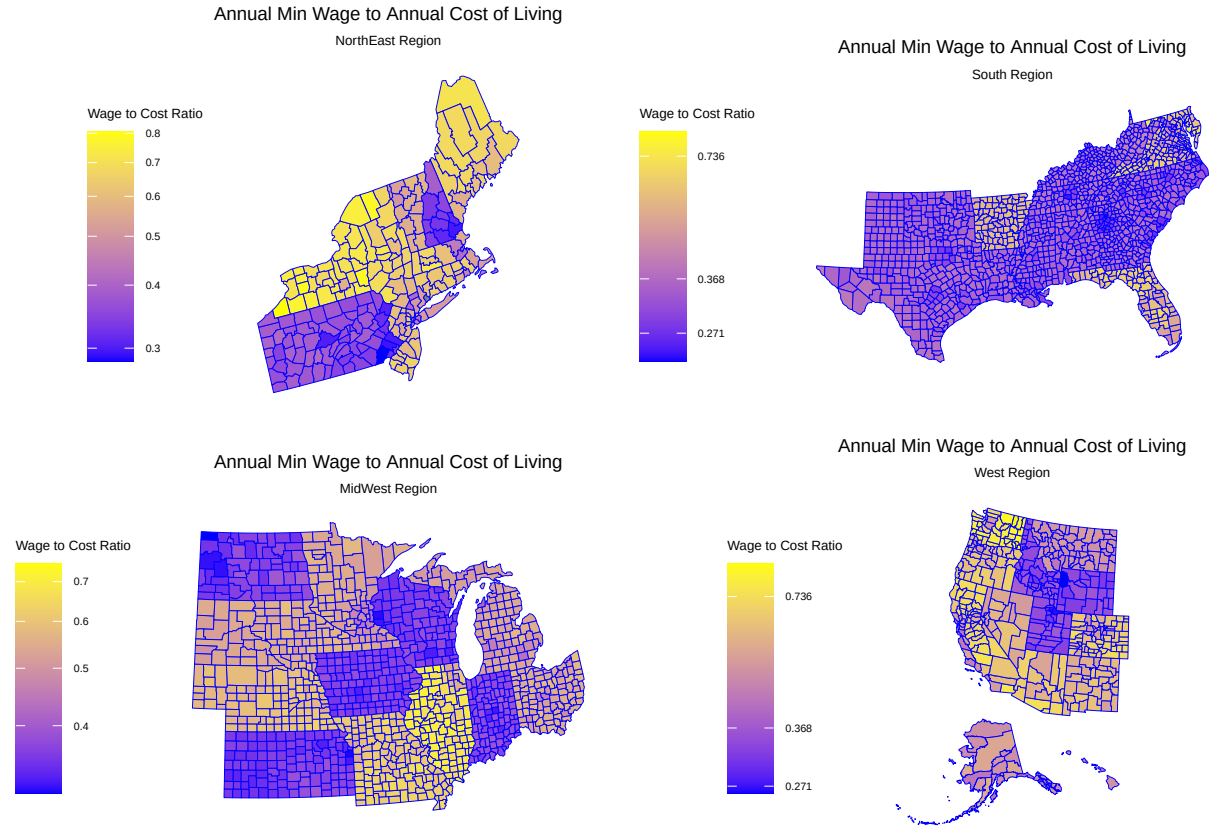
From this we can start to answer one of the most important starting questions, that is "Which counties are most affected by the minimum wage?"

Table 3: Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
a_total	3143	42915	5094	33420	39954	43920	84264
min_wage	3143	9.7	2.9	7.2	7.2	12	18
a_income	3143	20200	6054	15080	15080	24960	36400
a_ratio	3143	0.47	0.13	0.23	0.37	0.58	0.87

What types of plots and tables will help you to illustrate the findings to your questions?

I think it might be easier to see the data if I zoomed in on the states a bit more.



From this it is clear to see that the biggest divide is between states, not counties and not indivdiaul metropolitan areas. Some states are clearly "purple" with a low ratio, and most of them are still on the federal minimum wage of $7.25 per hour.

Do you plan on incorporating any machine learning techniques to answer your research questions? Explain.

I don't think machine learning is appropriate in this topic. I am not trying to predict anything and I am not proposing a model. I am simply looking at the data that is available and assessing it.

What questions do you have now, that will lead to further analysis or additional steps?

Two questions come to mind now:

1. Is the cost of living correlated with the minimum wage salary?

I have the right data to take a look at this question, and I will try to take a look, but a more serious deep dive into this is out of the scope of this project.

2. How many people depend on the minimum wage in each county?

Of course comparing the cost of living to the minimum wage seems very dire, but it is only important if a significant percentage of people actually depend on it. I do have some data on this on a federal level, and I will try to address it.

Milestone 3

Introduction

The minimum wage hasn't increased in 15 years in America. This is the longest stretch of time without an increase ever, with the 2nd longest stretch being 10 years between 1997 and 2007. With inflation being on many American's minds recently, I thought I would take a look at how the minimum wage compares to the cost of living in 2024.

Problem statement

One of the biggest arguments against raising the minimum wage is that it depends on where someone lives. Living in a countryside area has a much lower cost of living than in a metropolitan area. So I decided to examine each county in the USA and compare it's cost of living to the local minimum wage. I wanted to see if it was plausible that someone could live on the minimum wage anywhere in America, and how disparate the wage and cost of living were.

I also wanted to see if there were certain areas where the minimum wage more accurately reflected living costs, if there was a significant differnce between rural and metropolitan areas, and if there was a significant difference between states.

How the Problem was Addressed

I gathered data on the minimum wage for each county, along with the cost of living for each county. I focused on single adults with no children, to give the lowest possible cost of living. I then calculated the ratio of annual minimum wage salary to the annual cost of living. I then created a map of the USA using that ratio to see if there was anywhere in America where it would be practical to live on the minimum wage.

Analysis

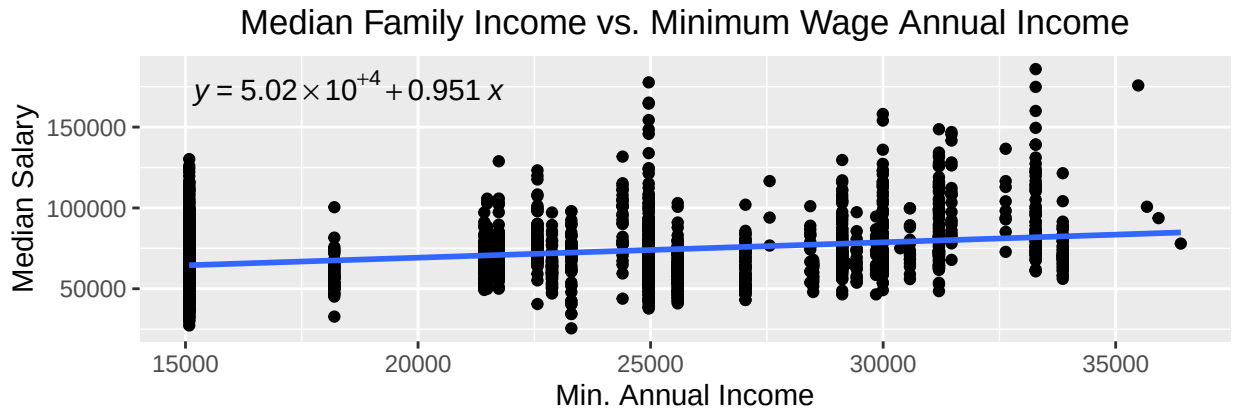
The data revealed some interesting insights:

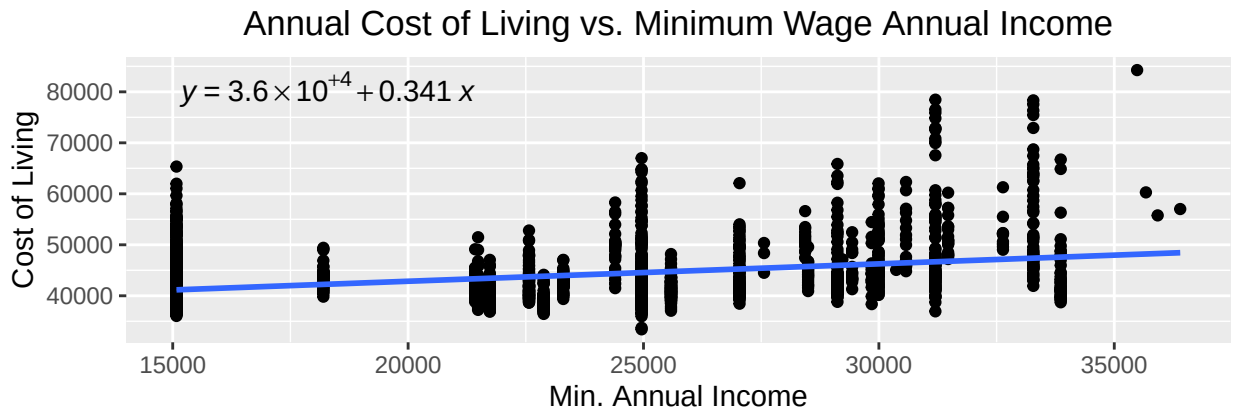
Minimum Salary Ratio: Nowhere in America did the minimum salary exceed 87% of the cost of living, with the mean ratio being only 47% of the cost of living.

City Differences and State Differences: There was not a large differnce in salary and cost ratios between metropolitan and rural areas. States with big cities and significant rural areas, like New York, Illinois, Texas and California seemed mostly consistent in their coloration, with only slight differences between counties with big cities and rural counties.

There was a big differnce between states that had the federal minimum wage or near the federal minimum wage, in that they were largely colored purple, meaning that they had a low ratio of salary to cost of living. States with a higher minimum wage of course fared better, but again it is important to keep in mind that no state had a county where the minimum wage was livable.

Minimum Wage Affect on Median Salary and on Cost of Living:






```
##
##  Pearson's product-moment correlation
##
## data:  min_wage_data$median_family_income and min_wage_data$a_income
## t = 19.346, df = 3140, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.2947422 0.3572386
## sample estimates:
##      cor
## 0.326347
```

```
##
##  Pearson's product-moment correlation
##
## data:  min_wage_data$a_total and min_wage_data$a_income
## t = 24.876, df = 3141, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3760575 0.4344864
## sample estimates:
##       cor
## 0.4056863
```

A quick correlation test between the affect of median salary on median wage has a correlation coefficient of 0.326, a weak positive correlation. I think that would be fair to say that the minimum wage has a weak, positive affect on general wages.

The correlation test on minimum wage on cost of living has a correlation coefficient of .406, a weak-medium positive correlation.

Implications

According to the US Bureau of Labor Statistics, there are just over 1 million working-age workers who earn at or below minimum wage, and this is about 1.3% of workers (Source: https://www.bls.gov/opub/reports/minimum-wage/2022/). This means there are over a million people who are working and earning below the cost of living.

Generally, the biggest argument against raising the minimum wage at a federal level is that the cost of living is so different between urban and rural areas, and that raising the minimum wage should be done at a local level. However, this data shows that the median minimum wage to cost of living ratio is 0.47, so a minimum wage worker is earning less than half of their cost of living.

I also think that this shows that raising the minimum wage doesn't have a strong negative affect on the larger economy. Raising the minimum wage affects both the cost of living and median income at about the same rate.

Limitations

I think there is a lot more work to do if I really wanted to find the affect of minimum wage on median salary, cost of living and the overall economy. Economists generate complex analyses of these problems, and I just did a quick analysis with basic information. It is obviously a much more complicated topic than I am able to analyze.

Concluding Remarks

Overall I very much enjoyed working on this project. I was able to learn new skills, reinforce old skills and really push myself in my R-programming and statistical analysis skills. I put a lot of effort into data cleaning and making the maps readable and attractive, and overall I am happy with the end product.

I still strongly believe that raising the federal minimum wage is long overdue, and that it would benefit all Americans, even those earning higher salaries. I think working poverty is completely unnecessary in America, and that it benefits no one.