MATH 4334: Mathematical Modeling (HW #1a)

Reece Iriye

August 29, 2023

Exercise 1

It is commonly claimed in the United States that "immigrants depress wages and take jobs away from native-born Americans." Humans are innately story-telling machines, so such a narrative can easily be compelling when used by politicians to rally support. I want to challenge this assertion with factual information and test whether or not this claim can be disproven. Observing reliable data—particularly data collected by the federal government regarding immigration populations and median incomes within each U.S. state—can potentially provide reason to disprove such a hypothesis.

Data Collection

The data being analyzed is derived from the American Community Survey (ACS) which provides estimates of foreign-born population proportions across various states from 2012 to 2016. Additionally, information regarding the median wages in each state for 2022 is sourced from the U.S. Bureau of Labor Statistics. Despite the temporal differences between the datasets, comparing them can still provide valuable insights, as economic trends, especially median wages and foreign-born population proportions, don't drastically change overnight. While comparing these two datasets may not provide an exact correlation, observing the patterns and intersections can offer some insights into the dynamics between wage levels and immigrant in different states.

I am particularly concerned about data quality in the ACS data, because that data is an estimate within each state extrapolated from a small portion of presumably randomly selected people in the United States. This survey, however, is not random in nature, because the immigrant population estimations in each state are estimated based off of a survey where participants are selected via. a random sample of addresses in every state, as seen here. While the data is likely at least somewhat accurate, it only takes home-owners into account when modeling a predicted value for each state. By taking only home-owners into account, the nature of these estimated foreign-born citizen calculations may be incorrect. According to the Pew Research Center, home-ownership amongst all Americans in 2008 was approximately 67.8% in 2008, whereas the immigrant home-ownership rate at the time was around 52.9%. Assuming these calculations are accurate, we can conclude that a survey of home-owners is likely not representative for estimating the number of foreign-born citizens in each state.

Data relevance is also concerning in the ACS data, because it specifically estimates foreign-born *citizens* and not individuals applying for citizenship in the United States or undocumented immigrants. The rhetoric behind the claim that "immigrants depress wages" in the United

States often focuses specifically on undocumented immigrants, but these numbers are not reflected in the ACS data whatsoever, causing some complications in our analysis.

I am not nearly as concerned about the median wage data in each state, as median wage calculations are not as complicated to retrieve. The federal government regularly requires citizens to log their income tax return with the IRS every single year, making this information more accessible by the government. By not reporting income to the federal government without a legal exception, U.S. citizens can be imprisoned for tax evasion. Meanwhile, immigration data is difficult to collect, and because of this phenomenon, data quality is more of a concern in that space.

Plotting the Data

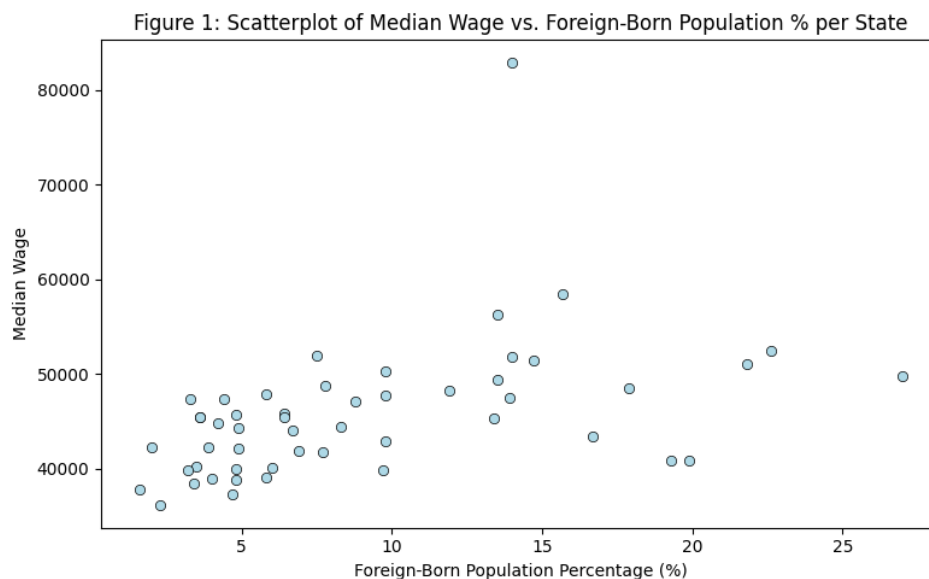Figure 1: Scatterplot of Median Wage vs. Foreign-Born Population % per State

Figure 1 showcases a scatterplot comparing median wages and foreign-born population percentages per state. A notable outlier in this plot is Washington, D.C. Its deviation from the general trend invites a deeper examination into the distribution of median wages among states to conclude if excluding this outlier would enhance the dataset's clarity.

Figure 2: Histogram of Median Annual Wages per U.S. State
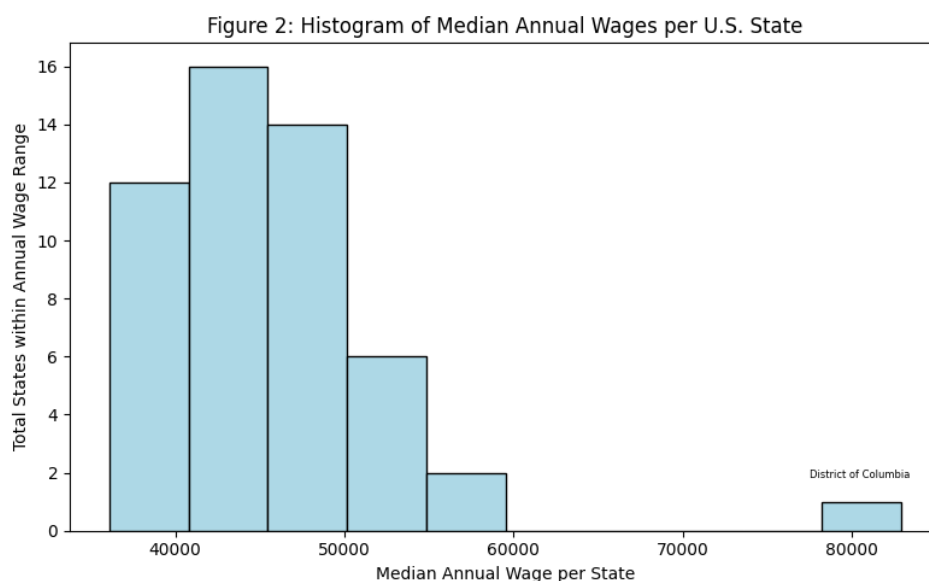
In the histogram pictured in *Figure 2,* by observing the cluster of median annual wage data points less than $60,000, we can see that Washington D.C. deviates from the rest of the dataset. Because of its abnormally high median annual wage, and because Washington D.C. is not even a state to begin with, I believe removing it will allow us to better interpret the relationship that the scatterplot in *Figure 1* tried to display.
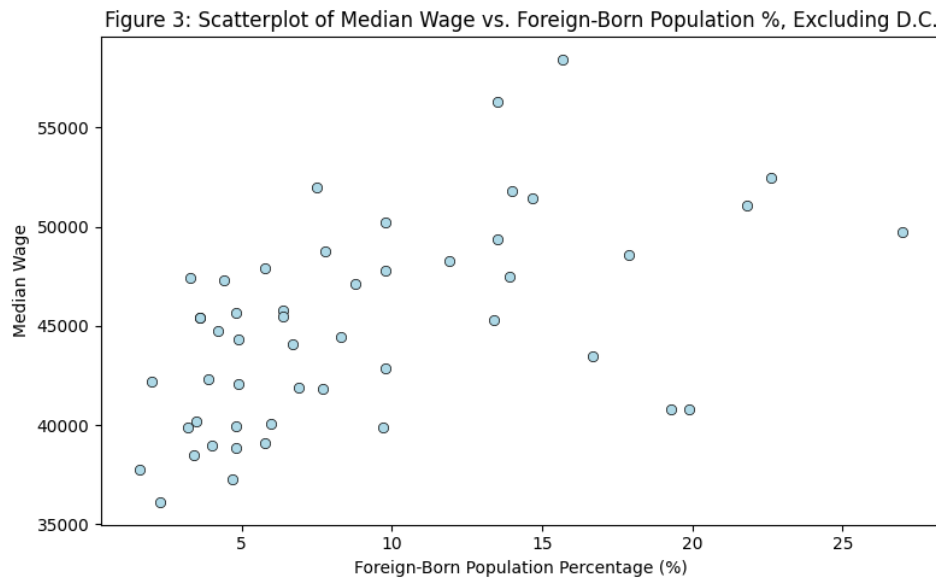

Figure 3: Scatterplot of Median Wage vs. Foreign-Born Population %, Excluding D.C.

Figure 3 provides a closer look into the relationship between our variables of interest without the Washington D.C. outlier disrupting us from evaluating the overall trend between these variables.

Relevant Insights

A roughly positive correlation between foreign-born population percentage and median wage per state can be inferred from *Figure 3*, which counteracts the claim that "immigrants depress wages" entirely. This positive trend can be seen particularly between the 0-16% range for foreign-born population percentages. The trend shown in *Figure 3* does not include any content that proves or disproves the premise that immigrants "take jobs away from native-born Americans"; instead, it only allows us to make an alternative hypothesis about the positive correlation between foreign-born population and median wage.

Given the nature of the data collection process in the ACS foreign-born population dataset for each state, I believe an alternative hypothesis can be made: "As the proportion of home-owning foreign-born citizens increases in each U.S. state, wages increase as well". This alternative hypothesis resolves my data relevance concerns from earlier and mitigates some of my concerns about data quality as well. Moreover, this hypothesis can likely remain true based on the scatterplot even if the predictions are not entirely accurate. My only major concern that remains is that the ACS and U.S. Bureau of Labor Statistics datasets both evaluate different time intervals. Nevertheless, assuming these values did not drastically change over the years, the alternative hypothesis reflects the realities seen in both these datasets and is entirely data-driven.

Exercise 2

A common argument in politics in regards to the economy is that "high taxes strangle the economy. Low taxes increase economic growth, which benefit everyone." While this argument can easily be attached to a narrative and justified by claiming that lower taxes enable spending—allowing the market to naturally flourish and grow as a result—I will observe the data behind taxation and economic growth and find evidence to either reject or fail to reject this claim.

Data Collection

The data being analyzed is derived from the Tax Policy Center, which provides total tax burden percentage values in each U.S. state in 2023. WalletHub argues that tax burden is a better metric for observing how much each state issues taxes in a given year, because "tax burden measures the proportion of total personal income that residents pay toward state and local taxes," meanwhile tax rates "vary widely based on an individual's circumstances". Additionally, information regarding per-capita GDP in each state in 2022 is sourced from by the Statistica Research Department. These datasets capture a timeframe that is lagged by 1 year, which may be somewhat significant because some states had elections and shifts in power, these people had only taken their oath of office just months before the tax burden data had been collected. Hence, there likely was not much of a change in tax policy in the meantime in each state, meaning that tax burden data is likely similar to the same data in 2022.

Data relevance is somewhat concerning for me in this situation, because tax burden is a metric that encompasses property tax burden, individual income tax burden, and total sales & excise tax burden. It only focuses on taxes issued on the state and local level and does not incorporate important taxes like the federal income tax at all. Hence, using this state-level tax data as opposed to overall tax burden and national per-capita GDP over the course of x years could lead to this issue. Nevertheless, I will plot the data and evaluate the underlying relationship to understand if the total tax burden in each state leads to a certain value in real per-capita GDP in that respective state.

Read per-capita GDP is a strong measure of economic performance in each state, as it filters out the noise caused by population size in standard GDP calculations and, instead, represents overall economic performance in a region in comparison to the population being directly affected by how that economy is performing.

Data quality should not be a concern, as broad financial data related to the state of the economy and taxation in the United States regularly collected. The Tax Policy Center is a highly respected non-partisan thinktank based in Washington D.C., so the data they provided related to tax burden per state is likely legitimate. Additonally, the Statistica Research Department regularly pairs with universities and provides data to support research across multiple disciplines.

Plotting the Data

Figure 4: Scatterplot of Real Per Capita GDP vs. Total Tax Burden
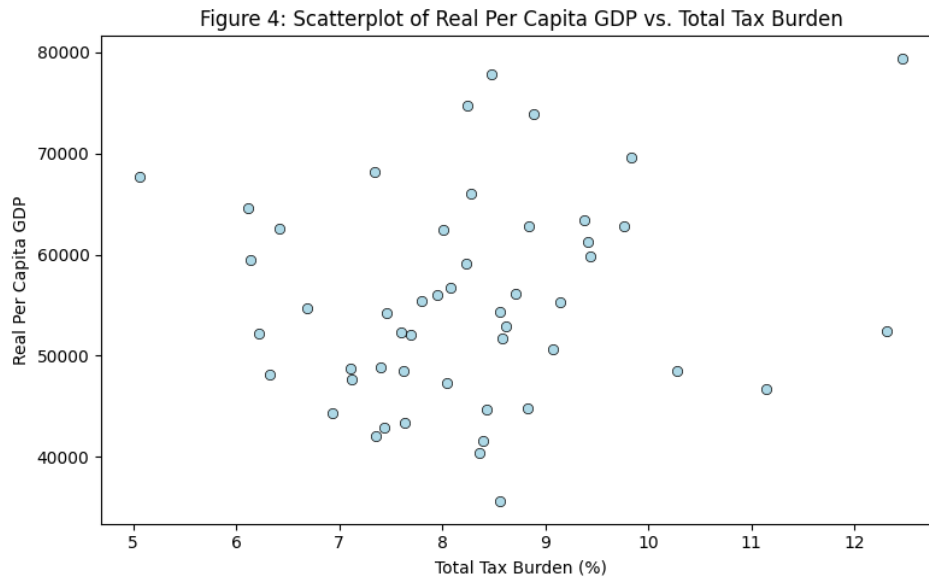
Figure 4 displays a scatterplot representing the relationship between 2023 total tax burden percentages versus 2022 real per-capita GDP values by state. Washington D.C. was not included in this graph, as the Tax Policy Center did not provide a tax burden value for Washington D.C.

Relevant Insights

No obvious relationship between tax burden and real per-capita GDP can be ascertained from *Figure 4*. A slightly positive correlation between tax burden and real per-capita GDP per state could potentially be identified, but no significantly obvious trend can be seen here.

Given the observed trend (or lack thereof) from *Figure 4*, I would assert an alternative hypothesis: "There is not 1 straight-forward negative relationship between the amount a state taxes its citizens and economic performance within the state". Instead, other factors like infrastructure quality, presence of well-performing universities, and educational strength may play a much stronger role in determining the performance of a state's economy, rather than just tax rates alone.

Exercise 3

Data Collection

The data being analyzed is derived from YouGov's 2013 research, which provides gun ownership rate estimations across all states based on 4,000 US adults aged 18 or older. Assuming the people surveyed are representative of the entire population, this data is somewhat reliable, but I am somewhat concerned about the sample size not being big enough to represent each individual state's population. If 4,000 adults in total are surveyed, then only around 80 people in each state are evaluated and used to create a metric of gun ownership in each state. Hence, I believe the estimates might be biased towards an unrepresentative sample.

Additionally, information from 2014 regarding firearm death rate per 100,000 is sourced from the Center for Disease Control and Prevention. This dataset is reliable, as it is difficult to for deaths not caused by gun violence to be recorded, and the CDC is a reliable source of information when it comes to logging causes of deaths. This dataset does not include Washington, D.C., unlike the previous dataset, so Washington, D.C. will not be incorporated into my analysis.

Both these datasets span a temporal range that only lags by 1 year, which is reliable for analysis in my opinion, because gun ownership and death rates likely did not drastically change in many U.S. states over the course of 1 year.

Plotting the Data

Figure 5: Scatterplot of Gun Death Rate vs. Gun Ownership Rate
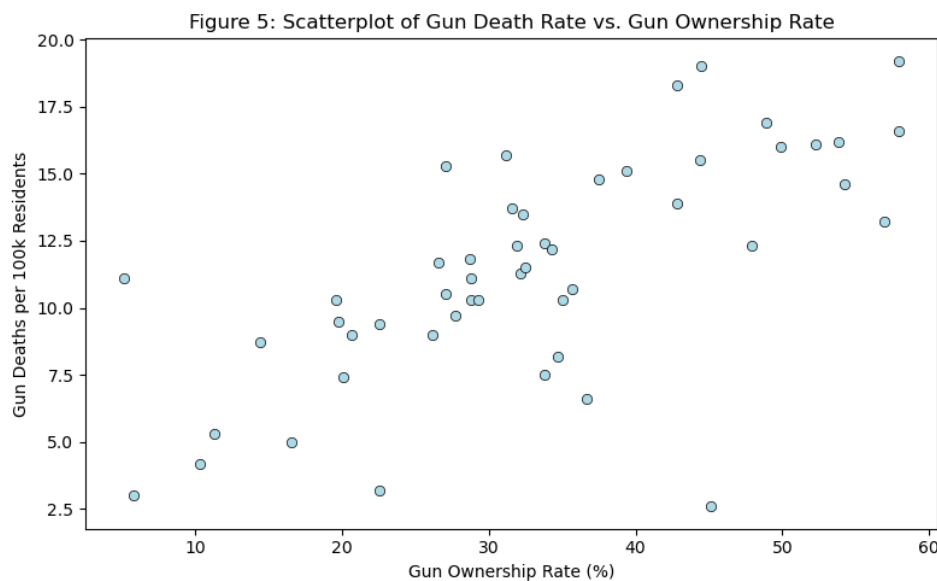
This scatterplot represents the relationship between gun deaths from 2014 versus gun ownership rates in 2013 in every state.

Relevant Insights

We can see in this graph that there exists a weak but positive correlation in the data. This positive correlation between gun ownership rate and gun death rate in each U.S. state does not support the claim that "if guns are outlawed, then only outlaws will have guns. An armed society is a polite society". In fact, it contradicts it and suggests the opposite trend. While the data integrity of gun ownership rate is questionable based on the sampling methods used to extrapolate these values, as long as the true values do not deviate much from these predicted

values based from our sample, this trend in the plot can allow us to form another hypothesis that an increase in gun ownership in a state actually can directly lead to higher gun deaths per capita in that same state. Other variables that have not been studied, such as education levels, may play a role in this equation as well, but we cannot make a well-informed hypothesis about that data unless we specifically look at that data in particular.

Exercise 4

It is a common talking point amongst liberals in the United States that "higher minimum wages do not lead to a decrease in employment. They lead to a more equally distributed pool of wealth by allowing for workers with these minimum-wage jobs to have a bigger share of the overall distributed wealth." I want to challenge this assertion by looking at the relationship between minimum wages in states and unemployment rates.

Data Collection

The data being analyzed is derived from the U.S. Bureau of Laor Statistics, which provides unemployment rates across all states from July 2023. Unemployment rate is a somewhat misleading metric in its definiton because it equals $\frac{unemployed}{employed}$. The denominator does not include unemployed people, which amplifies the metric a little bit as opposed to it being the sum of employed and unemployed people, but nevertheless, it grows as the number of unemployed people grows. I have no concerns with the quality of this data.

Additionally, information about minimum wages in each state is sourced from Paycom. The data is from 2023, and I believe it is reliable, as these are standard rates that can easily be found on the internet for each state. The dataset does not take into consideration the possibility of local rules being different, but nevertheless, most of these values are entirely accurate, and some values are exact with exceptions based on health, type of company, etc.

Plotting the Data

Figure 6: Scatterplot of Unemployment Rate vs. Minimum Wage
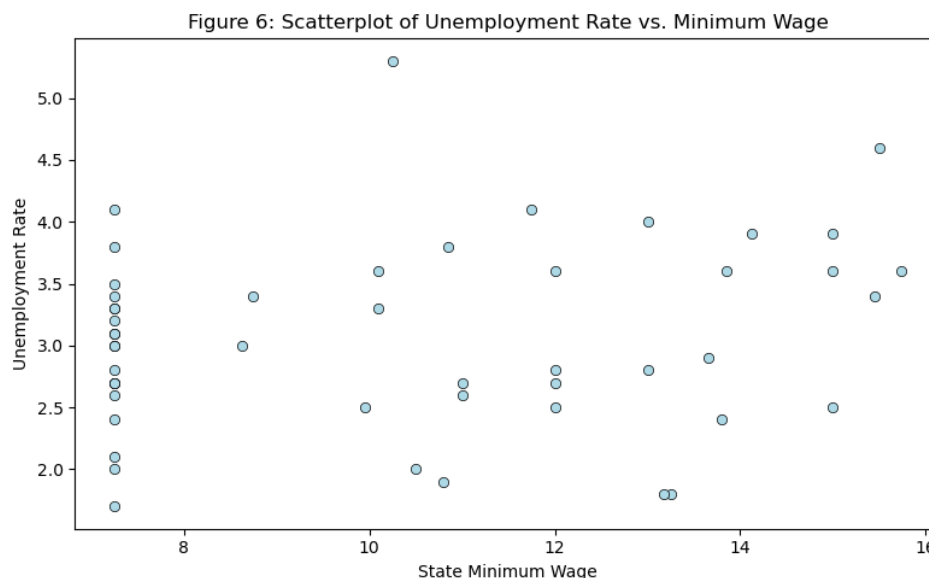
Relevant Insights

We can see in this graph that there is not an obvious relationship between state minimum wage and unemployment rate. It's clear that the relationship between minimum wage and unemployment is not as straight-forward as some liberals and conservatives make it out to be. Some states have hiked up their minimum wage and still have low unemployment, while others have kept wages low but are not exactly booming with jobs.

Given the data quality for both these values is high, this lack of a trend allows me to sympathize with the claim often made by liberals that " "higher minimum wages do not lead to a decrease

in employment", and it does not provide me with enough information to outright reject it. The relationship found in *Figure 6* leads me to believe that other factors, such as population density and education level, may play a higher role in actually determining the unemployment rate in a state.

Exercise 5

I am interested in evaluating some sort of standardized index that measures freedom across multiple countries. I want to investigate whether or not a high freedom index is associated with a low cost in Big Macs from McDonald's, and vice versa. This relationship can actually be important in assessing the relationship between how authoritarian a country is and, in turn, how that compares to prices of goods in these same regions.

Data Collection

The metric I am using to quantify freedom in each country is Freedom House's 2023 Freedom in the World (FIW) Report. Freedom House calculates the FIW of each country and territory of interest by "using a combination of on-the-ground research, consultations with local contacts, and information from news articles, nongovernmental organizations, governments, and a variety of other sources". All of the information taken into consideration to calculate FIW for each country "analyzes the electoral process, political pluralism and participation, the functioning of the government, freedom of expression and of belief, associational and organizational rights, the rule of law, and personal autonomy and individual rights".

The metric I am using to measure the cost of a Big Mac is the Big Mac Index, which is a metric defined by the Economist to measure "whether the market exchange rates for different countries' currencies are overvalued or undervalued". The Big Mac Index is not necessarily a measure of purchasing power in each country, because of other variables like tax rates and wage regulations; however, it does allow us to measure the power of the U.S. dollar in some countries as opposed to others. The only 2 issues that I see using this metric—both of which are significant—is that this analysis will only take into account countries that actually have a McDonald's. Additionally, the Big Mac Index data classifies the European Union all under one umbrella, which does not allow for me to perform an inner join on both the datasets to merge them together.

Both of these datasets are from 2023, so they span the same time range.

Plotting the Data

Figure 7: Scatterplot of Country Freedom vs. Cost of a Big Mac
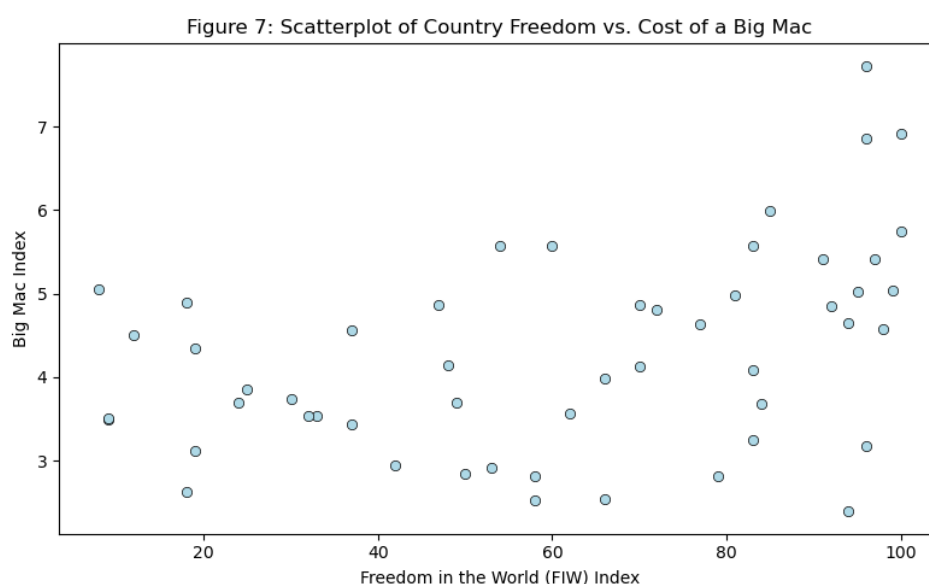
Relevant Insights

Figure 7 does not showcase an obvious trend between the Freedom in the World Index in each country and the respective Big Mac Index value for each country. The existence of only countries in this plot that have a Big Mac and the lack of E.U. countries that use the Euro as their standard of currency could pose bias that makes this trend not mimic reality, but it is the trend that we can see here.

Based on the lack of an obvious trend in *Figure 7*, I believe a hypothesis can be made: "The fairness of electoral processes, the presence of political pluralism and participation, the allowance of freedom of expression and belief, and accordance to the rule of law do not necessarily lead to an increase or decrease in prices for standardized global goods, such as the Big Mac". Just because a country may be more "free" or "democratic" according to Freedom House does not mean that standardized global goods like a Big Mac are going to be cheaper or more expensive. There are likely other economic, cultural, and logistical factors at play that affect the price of goods like the Big Mac. The phenomenon observed in *Figure 7* also raises an interesting point: While freedom and democracy are invaluable, they don't always directly correlate with a decrease in challenges that consumers might feel in their day-to-day life in a country.