US Unemployment Rate and GDP

DNSC 6211: Programming for Analytics

Amit Talapatra
Sonya Tahir
Wei Zheng
Wendy Zhang

December 11th, 2015

Abstract

Our project is intended to analyze the relationship between unemployment rate and GDP for all the states within the US from 1976 to 2015. We wanted to explore whether the unemployment rate in different states was affected by GDP, and if so, how it varied across the time span of four decades. We collected state-wise data for unemployment and GDP data for the country from two different sources. We plotted the data in different ways to analyze the relationship between these two variables. We would expect that as the national GDP increases, the unemployment rate should go down. This means that the two variables should have an inverse relationship. The relationship should also hold at state level with some variation due to the economic status of each state. This relationship was examined by analyzing the data using various plots. We also compared unemployment rates among all the states over the years.

Contents

1 Introduction

The 2007-2009 recession showed that losses in GDP can be paired with significant spikes in unemployment. At the beginning of the recession, the national unemployment rate was 5%, and this doubled to around 10% by the end of the recession, just two years later. In this analysis, we wanted to investigate the impact of US GDP on unemployment rates around the country. Particularly, we wanted to see if the impact varied among different states. Intuitively, there is an inverse relationship between GDP and unemployment, i.e. when GDP goes up, the unemployment rate goes down and vice versa. However, we do not expect to see a linear relationship because of the numerous other factors that could affect unemployment rate despite the changes in GDP. We intended to examine if the inverse relationship or pattern holds at the state level as well. It may be that some of the states are more economically sound and therefore tend to be affected to a greater extent due to the fluctuations in GDP. On the other hand, the states that are not major contributors to the national GDP growth would exhibit less apparent pattern than otherwise. To compare the states, we looked at the unemployment data in isolation of GDP over the span of four decades. This allowed us to observe the changes in each state's unemployment rate in comparison with all the other states.

2 Background

We started with the questions and assumptions outlined in the introduction section and searched for datasets that would serve the purpose we intended with the analysis. We chose the datasets due to a variety of factors that include an effort to make relevant comparison and making sure data is available within the same time frame, and an intent to keep the data consistent and complete. We didn't give up on any dataset as we started with a clear purpose in mind. The result of the analysis, however, did provide insight and further confirmed our initial observation/assumption of the unlikelihood of a consistent linear relationship between GDP and unemployment rate, especially at a micro level when individual states are examined. We were able to better understand the inverse relationship by graphing all the states and looking at the overall trend. Grouping the state graphs together confirmed the trend and some states' unemployment rate were clearly more sensitive to the change in national GDP level than others.

3 Method

We intended to investigate the relationship between GDP and unemployment rate. As discussed previously, we expected to see an inverse relationship, i.e., we expected the unemployment rate to go down with increase in GDP. However, we did not expect to see this inverse relationship throughout the entire time period under observation. For example, in some years, the unemployment could go up despite increase in GDP. This may be because of a variety of other factors that were not taken into consideration in our study. This relationship between GDP and unemployment was observed and verified by the data that we analyzed. We saw that GDP continuously went up throughout the years. The overall trend in unemployment was downward with a few upward movements as expected. We were able to see the peak points in unemployment rate for most states, if not all, at various data points during the time period we were observing, despite consistent increase in GDP level. We were able to reflect upon recent economic downturns to assume that they are at least partially due to the recessions US had gone through. However, our data couldn't confirm or deny such assumptions. While we were expecting such fluctuations and anomalies the extent of such are not explained by the data we

have. We thought we could provide the option to explore the changes in unemployment rate within a specific time range. We added the functionality of selecting any time period within the overall time range under observation. We then calculated the change in unemployment in these two years for the 10 worst states, i.e. the 10 states that for which the unemployment went up the most during the two years under question.

4 Organization

We worked together on all parts of the project. We discussed what questions we wanted to explore and where we could get the data accordingly. We thought we could work with economic data and see the relationship of some variable with changes in GDP. We looked at a few websites to see what possible options we have. Unemployment data seemed interesting and we thought it might have a strong correlation with GDP. We wanted to test that and so we looked for GDP and unemployment data on various websites to see which source would provide the data while satisfying the requirements of this assignment, such as API and web scraping capability. We also discussed how we wanted to plot the data to best demonstrate our findings and analyses.

4.1 Workflow

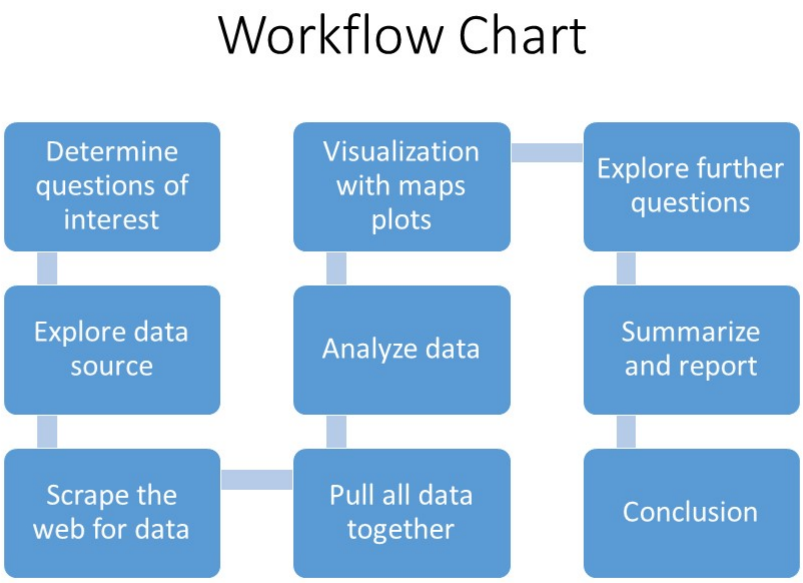


Figure 1: The project workflow

We started with deciding the question that we wanted to answer. We explored the possible data sources and decided on the final data sources that we will use for the analysis. We scraped

the web for the data. We combined all the data we got from the web and put it into a database. We also converted to csv files so that we could use the data from R. We analyzed the data from various angles by making various visualizations and graphs. We also made a heat map to compare various states. We looked at all the maps and graphs to make sure that we had answered the questions that we set out to explore. We thought we could explore the differences between two years, in particular for the years of recession. We added that functionality and showed the results in a table. We compiled everything in the form of a report describing in detail the data sources, the analyses, the results we got and the conclusion we drew from the whole analysis.

4.2 Project structure

We obtained the GDP Data from the Bureau of Economic Analysis website using web scraping.

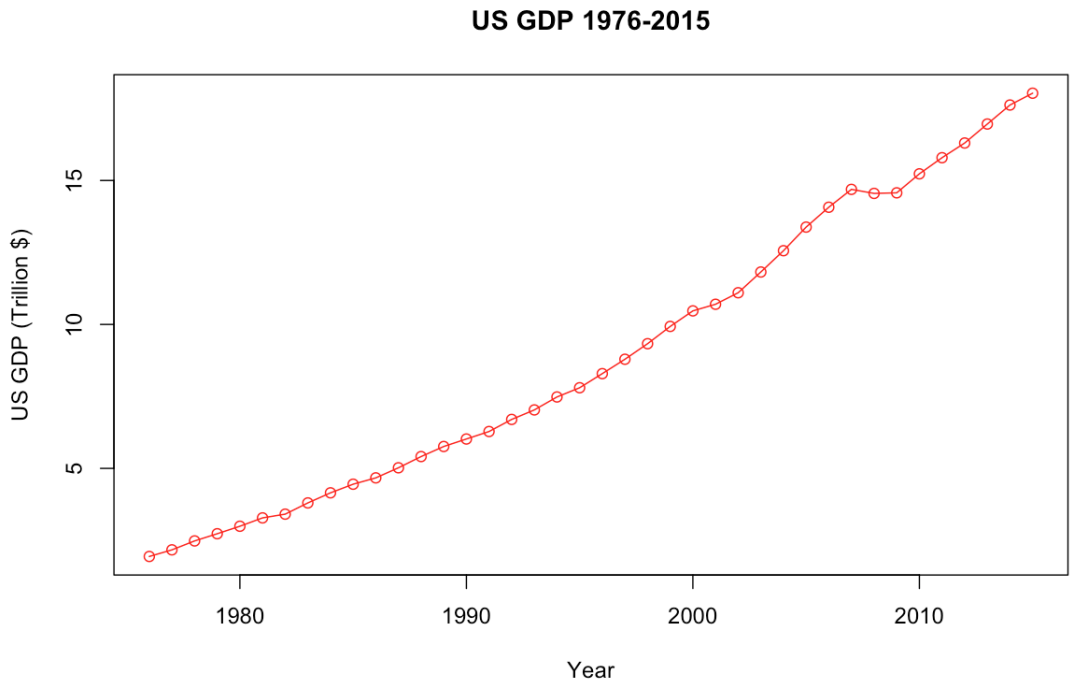
`http://www.multpl.com/us-gdp/table/by-year`

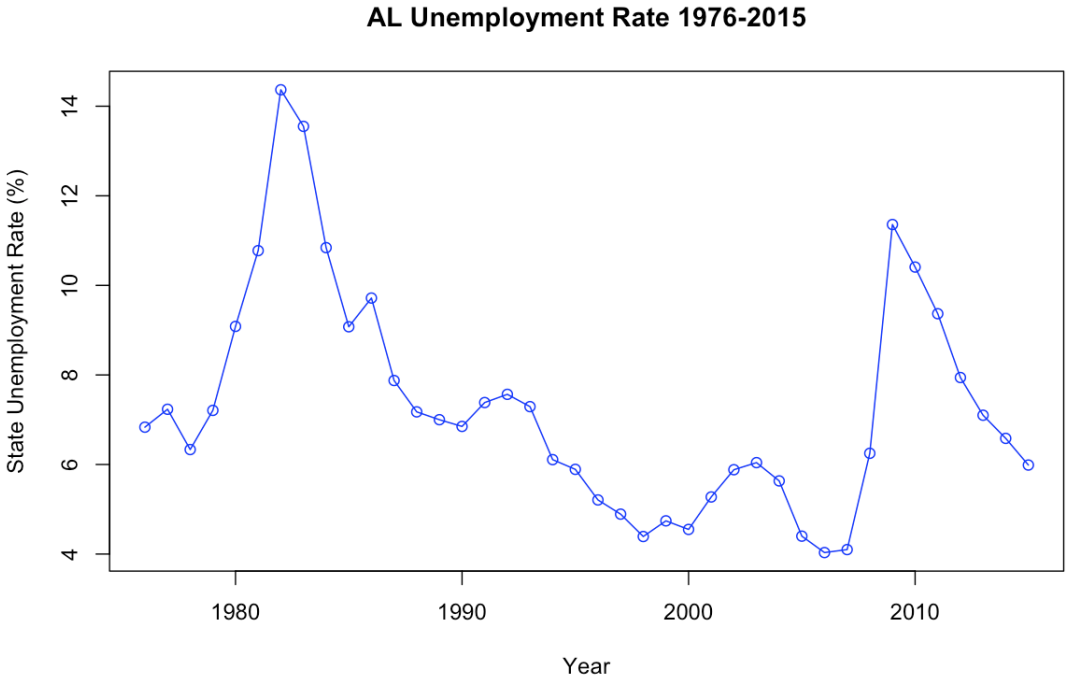
We obtained the Unemployment Data by State from the website of the Federal Reserve Economic Database using the Federal Reserve Economic Database API.

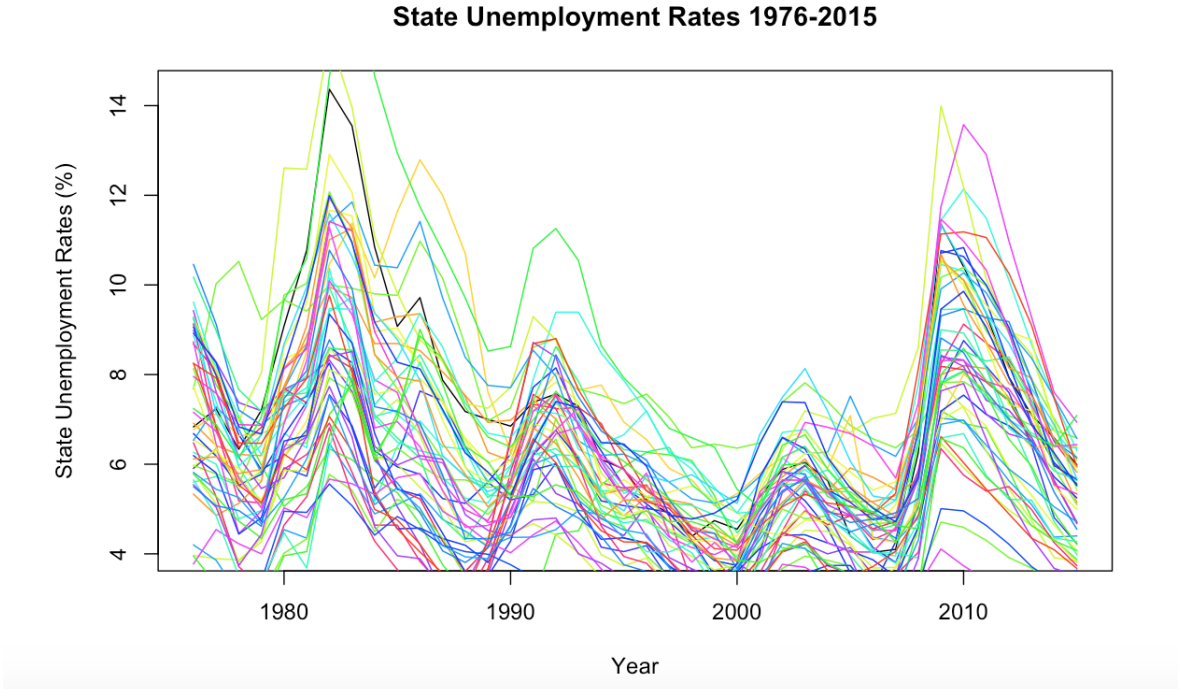
`https://research.stlouisfed.org/`

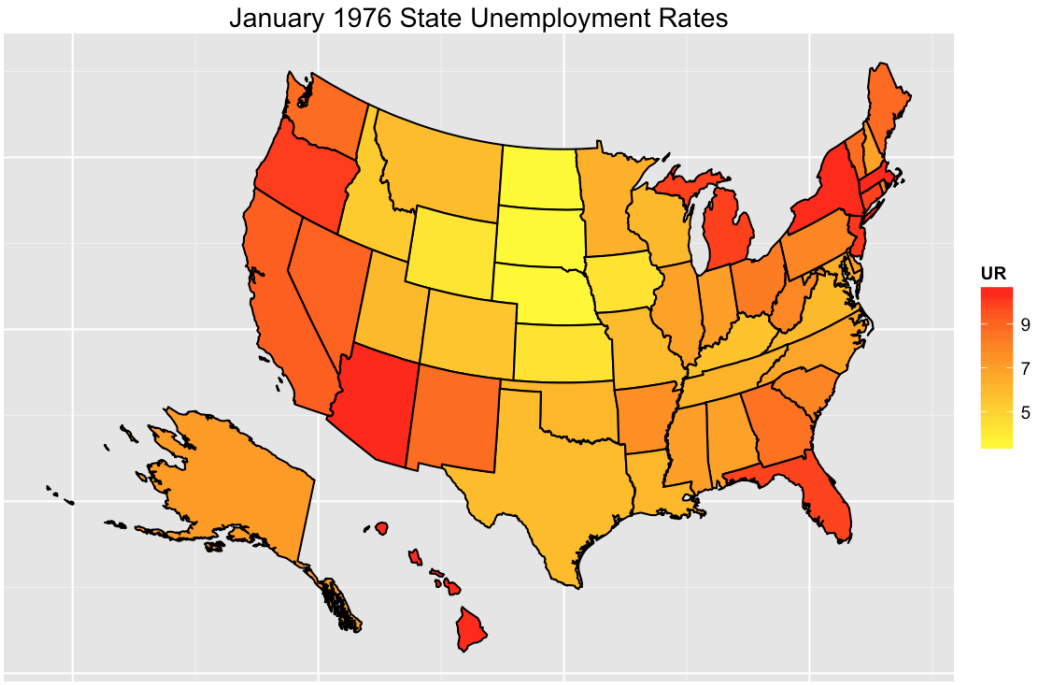
These are the two components of our research question, i.e. we wanted to analyze the impact of GDP on unemployment. We obtained data from the above mentioned sources from 1976 to 2015 for this analysis.

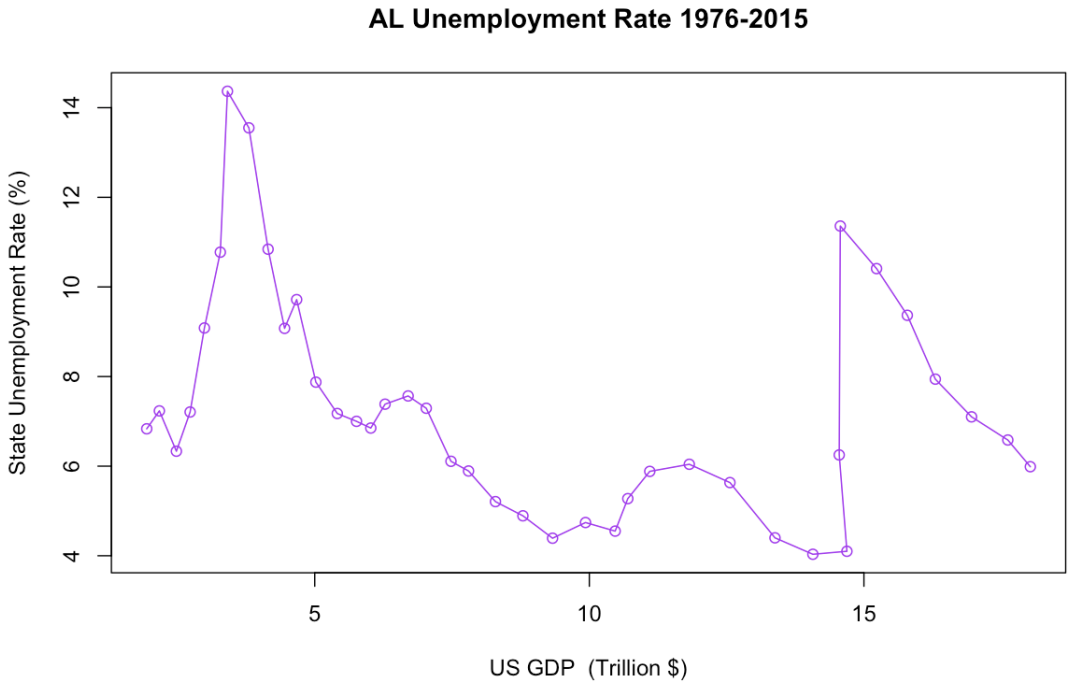
4.3 Figures and Tables




AL Unemployment Rate 1976-2015
State Unemployment Rate (%)
4
6
8
10
12
14
1980
1990
2000
2010
Year


State Unemployment Rates 1976-2015
State Unemployment Rates (%)
4
6
8
10
12
14
1980
1990
2000
2010
Year





The figures above show the analyses performed and their results. The first graph shows the

GDP trend. The second graph shows the unemployment rate for a particular state. The third graph shows the unemployment rates for all the states. The heatmap shows the comparison of unemployment rate of all the states as well. The last graph shows unemployment vs GDP. These graphs confirm our predictions and answer our research question.

5 Discussion

We believe that the result of our analysis provides valuable insight into the relationship we are trying to explore. We have substantial data over a few decades that thoroughly summarize the trends and the close relationship we examined. The heat map provides a clear visual of how unemployment rate has changes over the last 40 years. And the all states unemployment and GDP graphs provide a clear and straightforward presentation that confirms our assumptions.

5.1 Learnings

One of the better moments was being able to utilize R shiny to create an interactive heat map that allows users to observe the unemployment rate at a particular year and month. We also worked a lot more in R than we had before and learned a lot in the process.

5.2 Challenges

One of the challenges was utilizing slider input feature of R shiny to provide a slider using year and month instead of using just years. We had 477 months of data and each row corresponds to a particular year and month. We wanted to show the start and end date and increment it by month. In the end we decided to provide a data range input instead.

6 Conclusion

The result of our analysis is very insightful and outlines the relationship of the two variables under study. The results show that during 1976 - 2015, unemployment rate went down as GDP went up. This is confirmation of the intuitive prediction. We also saw that the relationship is not consistently linear which shows the affect of other variables on this relationship. This was not part of our research question. We were also able to visualize the comparison of unemployment between the states over the years. As the reason for our interest in this topic was the 2007 recession, and its affect on unemployment rates, we added an element to our Shiny app to report the top 10 states with the largest increase in unemployment rate over a selected period of time. Using the 2007 recession as an example, we saw that Alabama was the state most affected by the recession. This type of tool could be useful to pinpoint the regional effects of changes in an economy, and identifying the areas most affected would be useful in developing policies to reduce this impact.