Britney Zhu

AMS 315

Final Project Report

Unemployment Unplugged: Analyzing the COVID Job Market Meltdown

1. Introduction

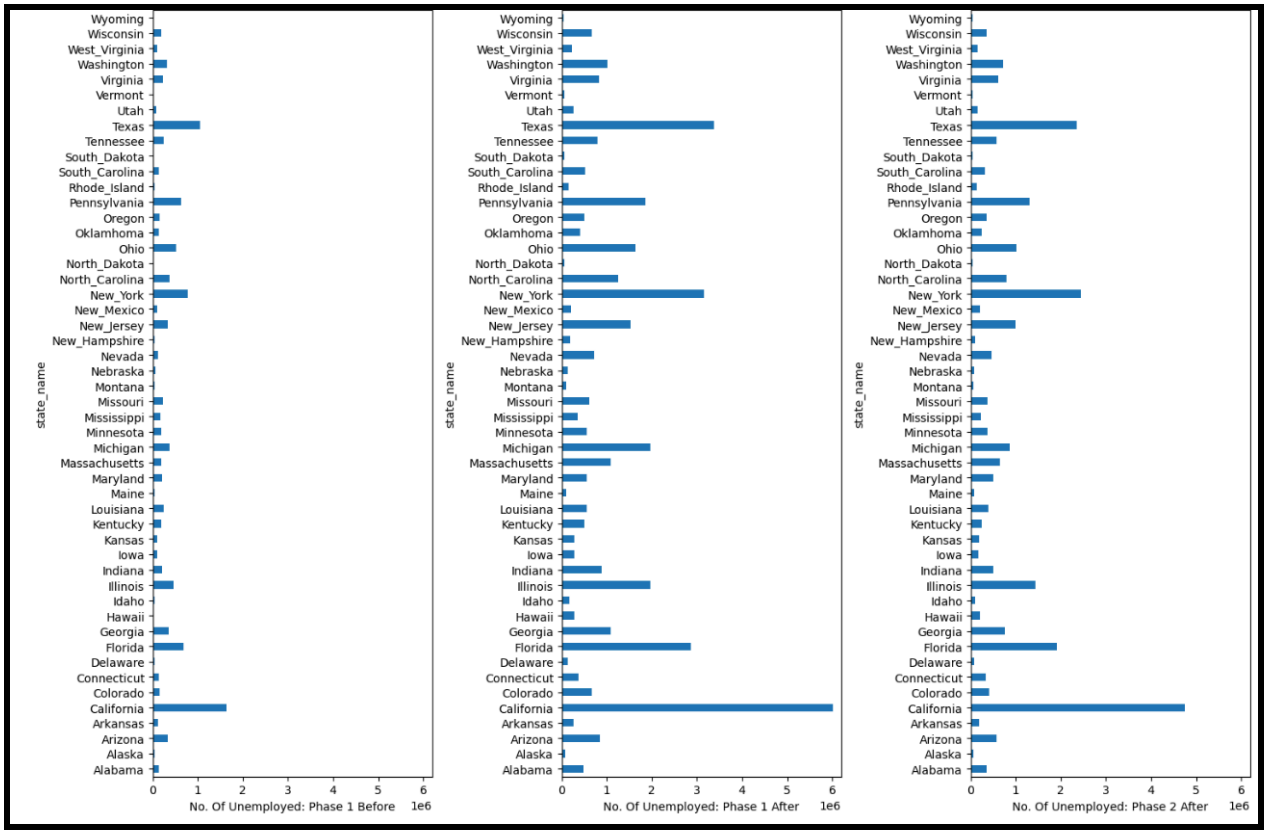
The unemployment rate has surged nationally as the pandemic took over the world. According to the U.S. Bureau of Statistics, in the second quarter of 2020, the unemployment rate was 13%. Many businesses had to get suspended due to closures and the unemployment rate tripled compared to 2019 (BLS 2021). Unfortunately, many people lost their jobs and the economy was in turmoil.

I and my partner Myra Arif started our AMS 325 Project using the programming language Python in the application Jupyter Notebook. Originally our project plan was that we hope to find the correlation between anxiety and depression regarding the spread of COVID-19. We are hoping to prove the hypothesis that the symptoms of anxiety and depression will increase after the first COVID case occurs in each state. However, we could not due to limitations of the available data in the CDC data set. This is because some states have early COVID dates dating back to February 2020. We need data from 3 months before the date, and CDC's earliest date was April 2020. So we had to change the whole project. The plan will still be very similar but what we are trying to find changed. We scraped the CDC data set, so the new project hypothesis is that the unemployment rate for each state will increase with each phase.

2. Methodology

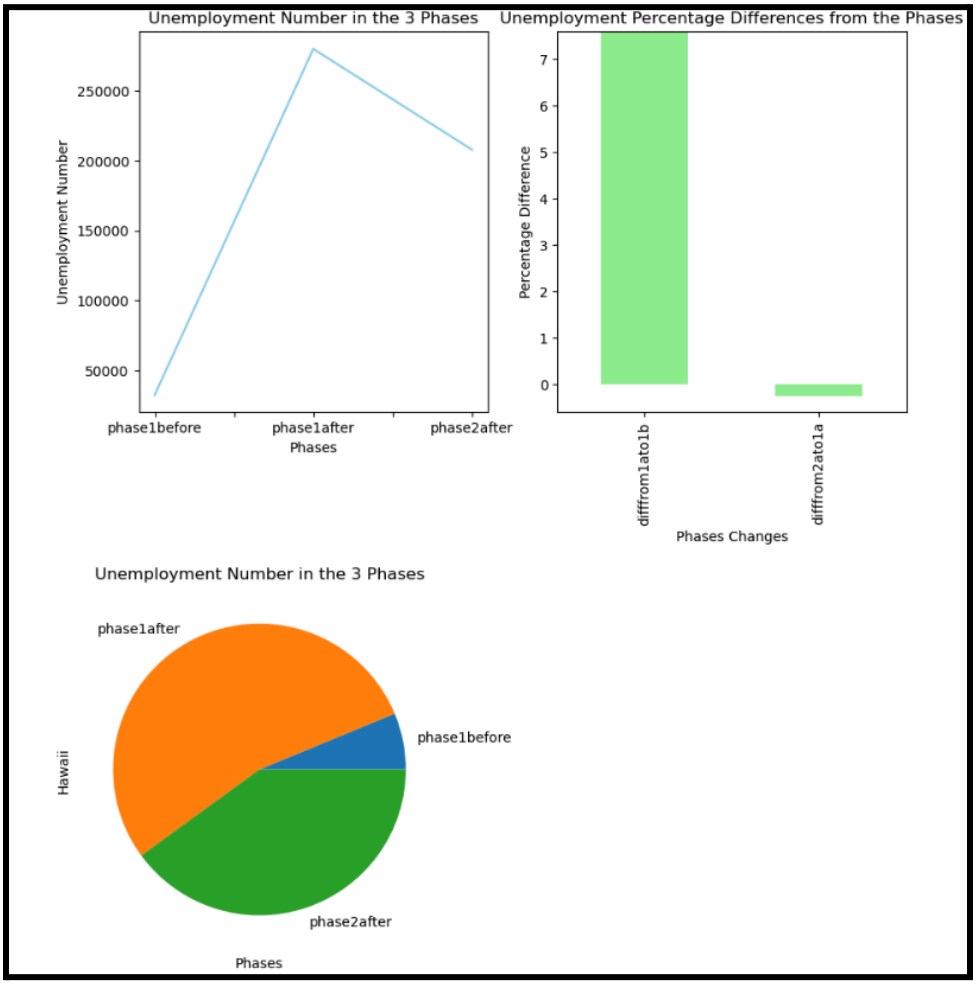
The datasets we used are the dates of the first case of COVID-19 in each state, the unemployment rate, and the census of each state. First, we put the first COVID date of each state with the available unemployment rate data we found online. After this, we changed all the unemployment rate into actual number of unemployed people using the census of each state in different months. Then, we made 3 phases. “phase1before” represents the 3 months before the first Covid date. “phase1after” represents the 3 months after the Covid date. “phase2after” represents the 4-6 months after the Covid date. We hope to find the averages of the unemployed number for each state. Since we were using dates and 3-month intervals, a lot of our initial code was converting dates into numbers we could use to find each 3-month interval. We created functions to find each of the 3 phases. It was difficult at first because for example if an initial date was February 2020, then the 3 months prior are January 2020, December 2019, and November 2019. We then created visuals to represent the data we found, mainly histograms, bar graphs, line graphs, and pie charts. Myra was in charge of initializing and finding the averages of the number of unemployed people. I was in charge of the visualization and the finding component of the project. Myra was stuck on the coding for average so I helped assist her in that too. The techniques we used were python numpy, matplotlib, pandas, datetime function, etc, and the lessons of data analysis.

3. Results



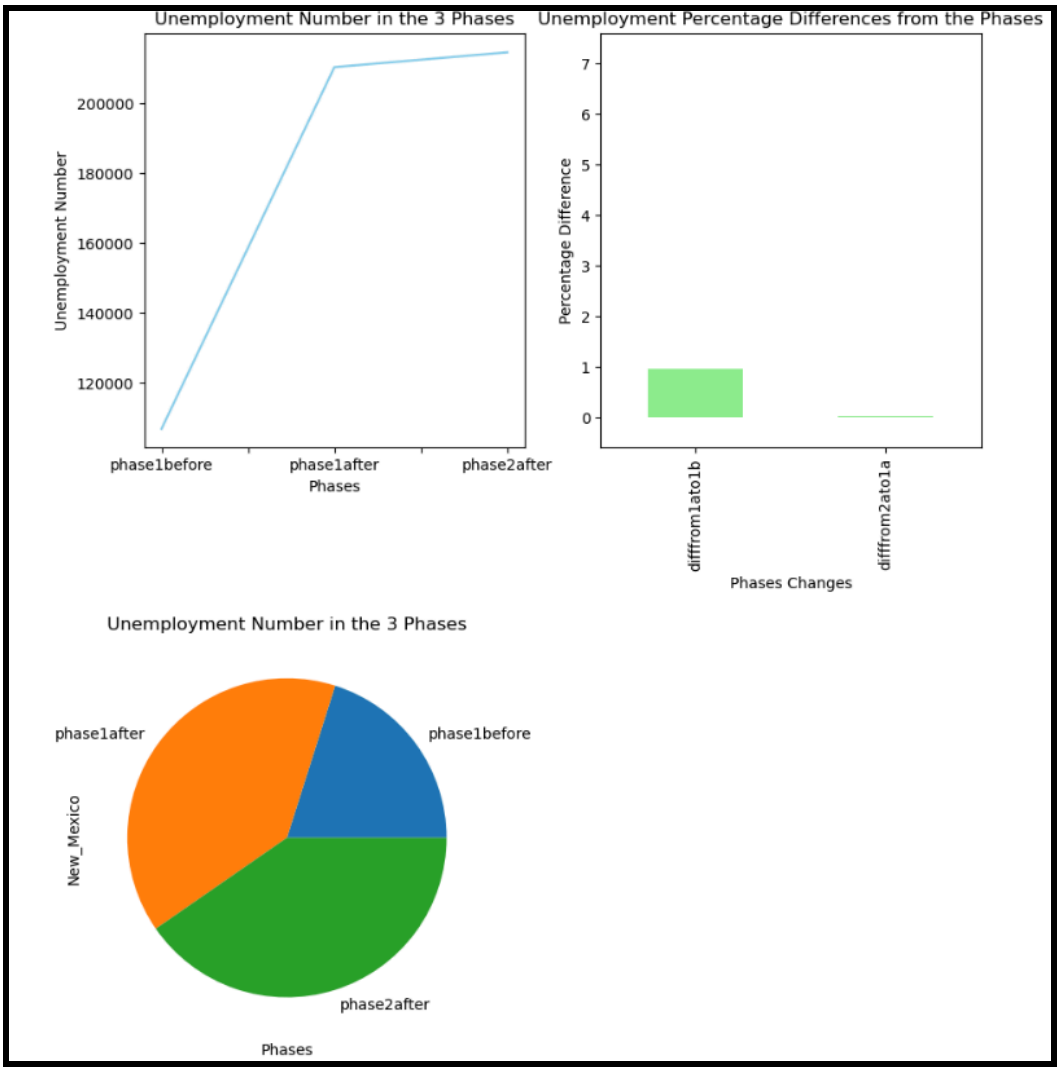
We created a subplot of 3 histograms of the 3 phases. We can see there is an increase when it comes to the unemployment number in each state. Originally I thought it was a bad idea to include the unemployment number because every state has a different population. But I realize that instead of comparing states, we just need to compare the 3 phases in that state. So it shouldn't matter really if we used unemployed numbers instead. Looking at the numbers, it is easier to grasp how big of an impact covid date had. While only looking at the percentages, the impact wouldn't seem as large. But seeing California going from 1.8 million unemployed to 6 million 3 months after the first COVID-19 date just shows how impactful it was. But since this plot includes all the states, it can be overwhelming.

Let's call Phase 1: 3 months before the first covid date, Phase 2: 3 months after the first covid date, and Phase 3: 4-6 months after the first covid date for simplicity. We can already infer that Phase 2 was impacted by the first COVID-19 date as California's unemployed number went from 1.8 million to 6 million people. But then in Phase 3, it lowered but still overall a drastic increase in unemployed people. But is it for all the states? We needed to find the range of maximum and minimum to find out. Let's try to find which states have the maximum percentage difference between phases and which one has the minimum percentage difference.



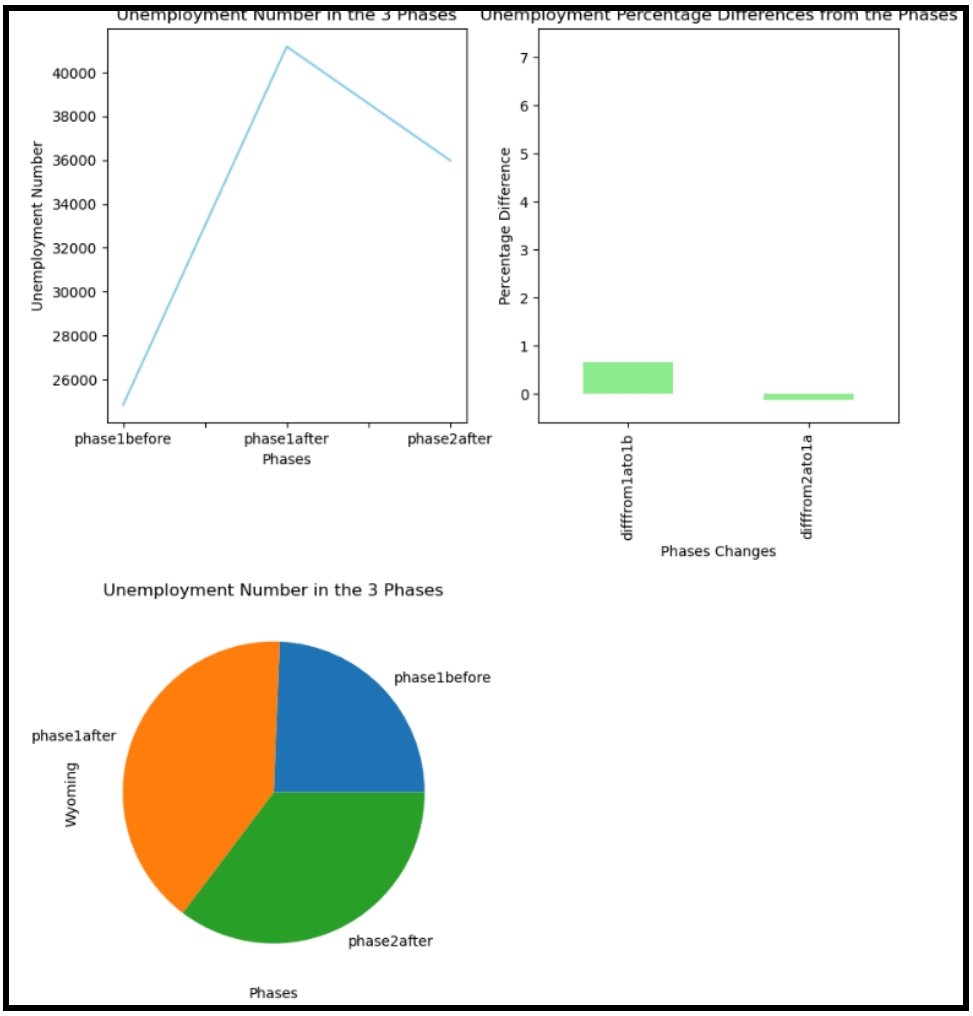
```
first_covid_case          NaN
phase1before         32585.63
phase1after         280218.82
phase2after         208107.24
difffrom1ato1b           7.60
difffrom2ato1a          -0.26
Name: Hawaii, dtype: float64
```

We found that Hawaii has the highest percentage difference from Phase 1 to Phase 2. This means that Hawaii was the state that was most impacted by the first COVID-19 date from 3 months before to 3 months after. A whopping percentage difference of 760%. The line graph is used to see the skyrocket of the amount of people unemployed from Phase 1 to Phase 2. Followed by some decrease. The pie chart is used to see the huge amount of people who are unemployed in Phase 2 and Phase 3. The bar graph shows a percentage difference of 760% then a -26% in Phase 3.



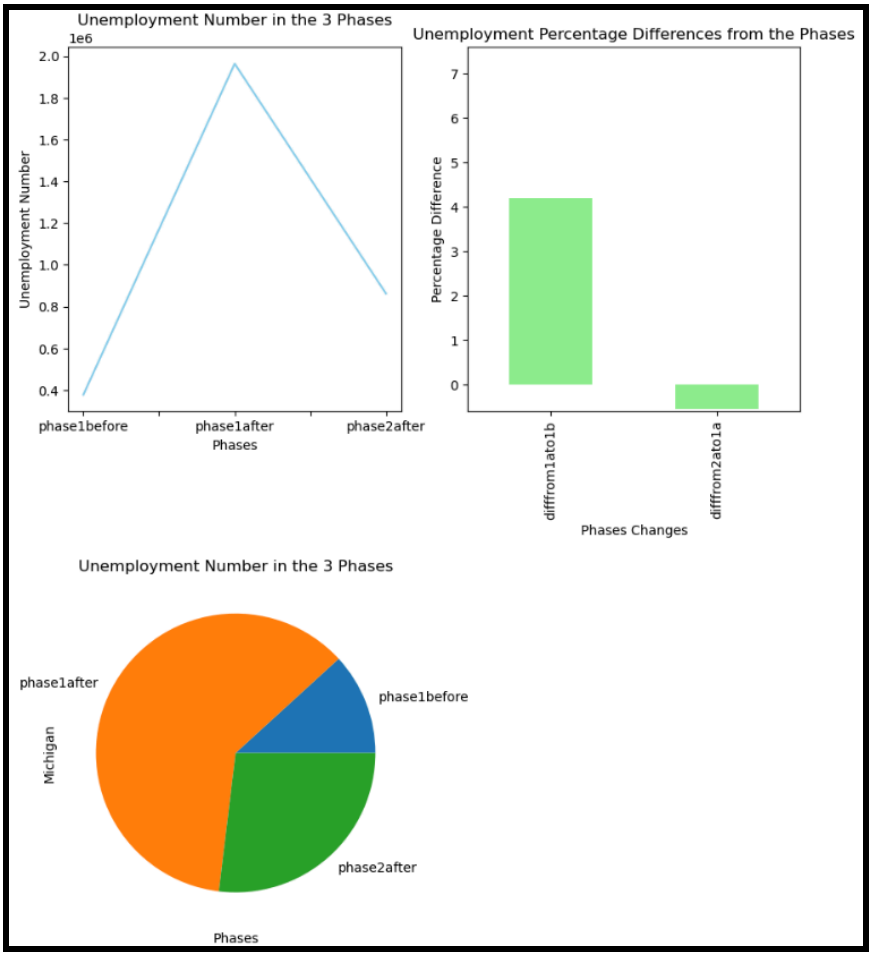

```
first_covid_case         NaN
phase1before        106958.22
phase1after         210344.89
phase2after         214580.02
difffrom1ato1b           0.97
difffrom2ato1a           0.02
Name: New_Mexico, dtype: float64
```

New Mexico has the highest percentage difference from Phase 2 and Phase 3. This shows that there are states with a slight increase in unemployment after 6 months.




```
first_covid_case        NaN
phase1before       24840.88
phase1after        41178.38
phase2after        35982.98
difffrom1ato1b         0.66
difffrom2ato1a        -0.13
Name: Wyoming, dtype: float64
```

Wyoming had the lowest percentage difference from Phase 1 to Phase 2. With a percentage increase of only 66%. Still a high increase but is not as high as Hawaii was. This still shows that there is an increase after the first covid date as the percentage difference is positive. And the line graph is upward sloping from Phase 1 to 2.




```
first_covid_case              NaN
phase1before           378164.85
phase1after           1963194.48
phase2after            862463.22
diffrom1ato1b               4.19
diffrom2ato1a              -0.56
Name: Michigan, dtype: float64
```

Lastly Michigan with the lowest percentage difference from Phase 2 to Phase 3. This shows that there is a 56% decrease in the unemployment rate from Phase 2 to Phase 3. Still an overall increase but was a lot better than it was in Phase 2.

We found out that Wyoming has the lowest percentage difference of 0.66 from phase 1 before to phase 1 after. This proves that the percentage difference is positive meaning there is a

correlation. A positive correlation was the first COVID-19 date occurred, and the following 3 months had an increase in the unemployment rate in all the states. We can infer that the lowest percentage difference is still positive. However, the 3rd phase (1st after to 2nd after) did experience a decrease when looking at Michigan (-0.56) but Michigan still experienced an increase in unemployment overall.

4. Conclusion

Remember that our project's hypothesis was the unemployment rate for each state would increase with each phase. Well, partially. There is indeed an increase from Phase 1 to Phase 2 as our minimum percentage difference was 66%. Not all states continued to still have an increase as we see the minimum percentage difference was -56%. This means that some states began to recover after Phase 2 occurred. We also had limitations in our data. There was limited data for 2019 so some phase1before calculations only consisted of two months. We also used the census population estimate for each of the states instead of the actual census population. We only looked at the 50 states (did not look at Washington DC.) However, we were still able to find patterns that helped solidify our conclusions. We were able to learn to create data time functions and apply real data to our coding.

References

- U.S. Bureau of Labor Statistics. (2021). Unemployment rises in 2020, As the Country Battles the Covid-19 Pandemic. U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2021/article/unemployment-rises-in-2020-as-the-country-battles-the-covid-19-pandemic.htm
- Unemployment rate (seasonally adjusted). KFF. (2022, May 4). https://www.kff.org/other/state-indicator/unemployment-rate/?currentTimeframe=17&sortModel=%7B%22colId%22%3A%22Location%22%2C%22sort%22%3A%22asc%22%7D
- David García-García, Enrique Morales, Cesar de la Fuente-Nunez, Isabel Vigo, Eva S. Fonfría, Cesar Bordehore,
- Identification of the first COVID-19 infections in the US using a retrospective analysis (REMEDID), https://doi.org/10.1016/j.sste.2022.100517.