Analysis of the Covid-19 Impacts on the Unemployment Rate in the U.S.

Group 8: Yuzhu Zhang, Lu lu, Ouni Guan, and Ruofan Jiang

Columbia University

APAN 5205: Project Proposal

Dr. Vishal Lala

March 02, 2022

The Statement of Research Problem

2019 was the year a decade after the end of the Great Recession. The U.S. economy continued to expand, and the labor market remained strong by the historical standard. The unemployment of the year was 3.5%, which was the lowest rate since 1969 (Edwards & Smiths, 2020). However, the situation changed dramatically due to the Covid. Thousands of companies went bankrupt, and countless employees were dismissed. Now, the covid era has lasted for three years. The unemployment rate experienced a surge and a recovery, which was 3.9% at the end of 2021. However, what was concealed under the satisfying number was the long-time inequality in the U.S labor market. It is still unknown how the pandemic influences different sectors. Therefore, our group will analyze how Covid affects the unemployment rate in the U.S. And, more specifically, investigate the category of gender, race, and education level. As a result, we can further predict the future trend of unemployment rate of those categories and make suggestions accordingly.

Literature Review

Employment inequality, as a severe problem, persists around the world, including in the United States. There are significant and constant differences in labor market outcomes between race, gender, and ethnicities that constitute an important dimension of economic inequality (Altonji & Blank, 1999). Based on a study examining unemployment data from 2005 to 2017 by race and gender, men tended to be disproportionately affected by unemployment during the recession compared to women, but men were able to recover their jobs more quickly. Moreover, when considering unemployment from gender and race perspective, more disparities appeared. According to the study, male African Americans showed the highest unemployment rate from 2005 to 2017, and African American women tended to show the highest unemployment rate among women (Adjeiwaa-Manu,2017). Additionally, some previous research has shown associations between education and the incidence and duration of unemployment. Farber reveals that job losers with higher levels of education have higher post-displacement employment rates and are more likely to be reemployed full-time (Farber, 1999). These past studies demonstrated that inequality does exist in the U.S job market, especially among communities of color, and overall females tended to show a lower unemployment rate than males.

Nonetheless, recent research on the unemployment rate reveals dissimilar results from the previous studies. The appearance of COVID-19 pandemic from 2020 has forced many businesses to close, leading to an unprecedented disruption of commerce in all sectors of the economy all over the world, including the United States. Consequently, the unemployment rate has increased robustly. Recently, research conducted by Andrés Villarreal and Wei-hsin Yu examined the gender disparities in employment during the COVID-19 pandemic in 2020. In contrast to previous recessions in which males were more likely to experience a job loss (Alon et al. 2020a), the pandemic has had a greater effect on women's employment than men's due to both labor demand and supply factors (Villarreal & Yu, 2022). Therefore, our group speculates

that the existence of the COVID-19 pandemic would have impacted the structure of the labor market in the United States. In other words, the unemployment rate by different demographic sectors including gender, race, educational level, and ethnicity would also be affected profoundly.

Research Question

How does Covid impact the current and future unemployment rate in people of different gender, race, state and education level in the U.S?

Data Description

Our data is from the U.S Bureau of Labor Statistics, which is a federal agency that collects U.S. labor markets data. The dataset includes the monthly unemployment rate from 2010 to 2021. There are 12 variables in total. The data can be divided into four segments. The first segment demonstrates the unemployment rate related to education level, which is categorized into primary school, high school, associates, and professional degrees. A decreasing trend can be discovered for the unemployment rate with a higher education level. The second segment is race. White, Black, Asian, and Hispanic are four groups considered. The third segment is gender, including men and women. Utilizing the dataset could answer whether Covid plays a role in the inequality of the U.S. labor market. The last segment is state. Four representative states from east, west, north and south are selected from the 50 states in the United States as the survey object, which are California, New York City, Floria, and Texas.

Summary of data cleaning

The original datasets include two parts: the first part is the unemployment rate in the US before 2020 April, and the second part is the unemployment rate in the US after 2020 April. The first step is to import data into two data frames, respectively called “unemployment” and “additional data”, as the original data are in formats of csv and xlsx. After importing the data, Our group has found that the original data set put all months together. So I regrouped the data, making it in the rank of years.

The additional data table contains two rows of null values, and three rows describing the total monthly unemployment rate, the total number of people who are unemployed and who are not in the labor force. Our group evaluates these five rows as being redundant as that information can be generated from the remaining data, so we remove rows 1 to 6, 9,14, and 15. The additional data table consists of reversed rows and columns as the unemployment table, so we applied a transpose function on the additional data table to match its style with the unemployment table. To make merging processes easier and more efficient, we rename columns in the additional data table to match column names of the unemployment table. The unemployment table also has a redundant column named “date” which is a combination of columns “Month” and “Year”, so we remove it.

After reshaping the tables, we applied the rbind function to bind two tables together. But the bound table has a column of Month which contains only character type values in non-chronological order. To make data easier to read and comprehend, we first assign factor levels to the Month column using the index of month.abb, and then sort the bound table by Month and Year columns.

At last, the merged table is almost ready to be analyzed while it still contains a few repeated rows with null values. We applied na.omit function to remove rows that contain null values. The final step is to assign new indexes in numerical order to the final table and name the table "unemp".

The rationale for Analytical Techniques

In order to examine the impact of Covid-19 on the unemployment rate among people of different gender, race, state, and educational level in the U.S and further predict the impact in the future, our group would like to use two analytical techniques: linear regression model and Arima model. First of all, we used a linear regression model, which examines the relationship between the unemployment rate and the appearance of pandemic. This model helps to explore whether covid is a significant feature that has impacted the unemployment rate in the United States, and at the same time examines the impacts from other features such as gender, educational level, and race. Additionally, we decided to use the Arima model, which applies time series data to predict future trends. By using the Arima model, we were able to predict the future unemployment rate under the impact of Covid-19, different educational levels, gender and race, and then based on the prediction to provide practical and applicable recommendations to better relieve the problem of the rising unemployment rate.

Analysis and Graphs

-Linear Regression Model (Covid case and Unemployment rate)

We would like to examine whether covid is a significant factor influencing the unemployment rate. The number of confirmed covid cases in the US is downloaded from GitHub and integrated into a monthly base dataset through R studio.

First of all, we test the effect throughout the whole nation. As covid started to spread in the US from 2020 March, data from 2020 and 2021 are used. There are three categories: education level, race, and gender. Those categories are tested along with the national case number to determine their relationship. We build a linear regression model for each factor.

According to the result of the models, all the p values are less than 0.05, which indicates covid does influence the unemployment rate in the US, either positively or negatively. Specifically, people with a professional degree are severely affected. The p-value for the model is 0.00472, which is the only one less than 0.01 in the category. In other words, the unemployment rate of people with higher degrees fluctuates more seriously. For race models, basically all race groups

are significantly affected, except Black people. The 0.01356 p-value is the only one higher than 0.01. It seems Black people's unemployment rate will maintain at a high level even without the impact of covid. For the gender category, the influence of covid is similar for both men and women, which hold a p-value of 0.01151 and 0.00913 respectively. However, the impact on women is a little bit more serious.

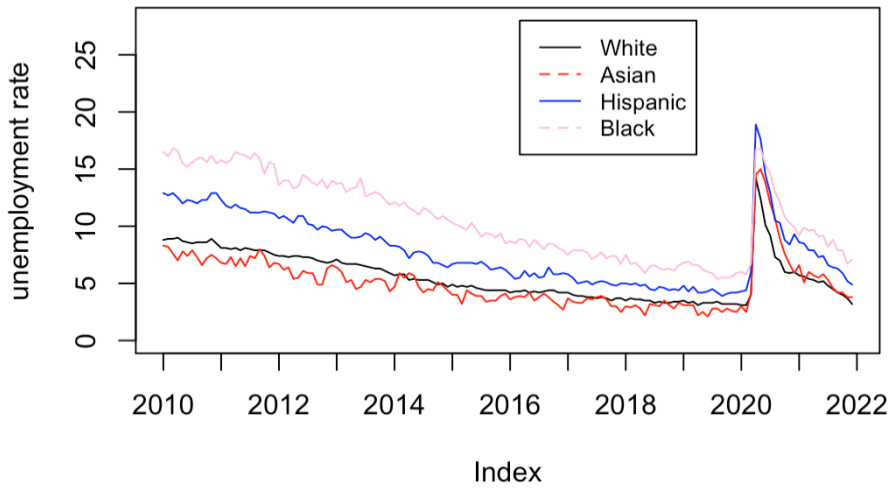
Besides, we also test the linear relationship in a smaller range – state. The top four populous states' covid case numbers and unemployment rates are collected. We build a linear regression model for each state. As predicted, all states have a p-value smaller than 0.05. In particular, Florida has a significant low p-value of 0.00283, which means the covid heavily impacts Florida's unemployment rate.

-Graphs of Unemployment Rate from 2010-2022

The four graphs below show the different unemployment rates in different categories. All four graphs have shown that the unemployment rates have been decreasing throughout the years, but with a significant raise in 2020 due to the pandemic.

1. <u>Race</u>

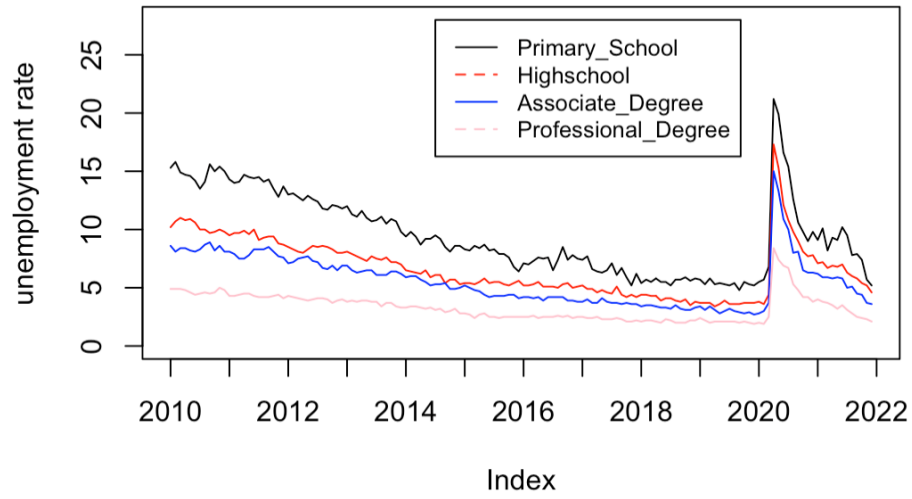
The first graph is grouped by race, White, Asian, Black, and Hispanic. The Black people has the highest unemployment in the past 12 years, followed by Hispanic, White, then Asian. This may be due to the history of society and the immigration problem. When a large number of immigrants inflow, they are limited to low-level jobs which don't require a lot of skills. These jobs are easily replaced, so when Covid-19 hits, the unemployment rate for these jobs is high.



2. <u>Academic Background</u>

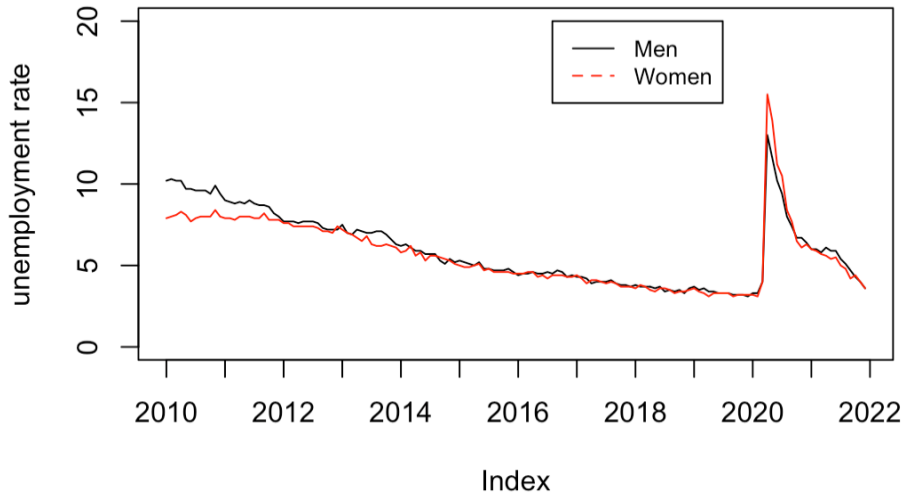
The second graph is grouped by academic background, Primary school, High school, Associate Degree, and Professional Degree. The Professional degree has the lowest unemployment rate. The higher education level they have, the harder they can lose their job. Jobs that require specific

techniques are hard to be replaced, therefore even though Covid-19, those people are still able to hold their positions.



3. Gender

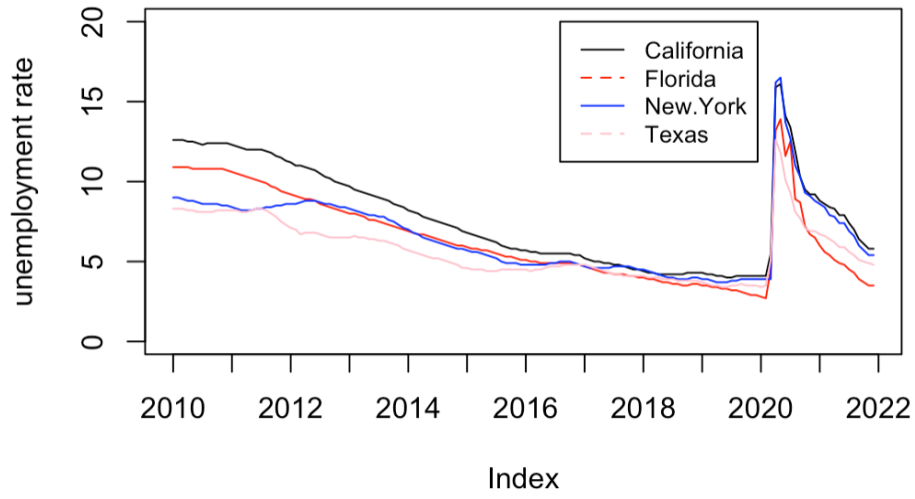
The third graph is grouped by genders, Men and Women. There is not a big difference between those two before 2020. However, after 2020, women's unemployment rate is greatly higher than men's. This may be due to gender discrimination. Women have always been identified as a more vulnerable group than men. So, during the pandemic, companies are more inclined to lay-off female employees.



4. States

The fourth graph is grouped by representative states, which are California, New York City, Floria, and Texas. Before the pandemic, California's overall unemployment rate was usually slightly above the other states' rate. However, when the pandemic started at the beginning of 2020, New York surpassed California and all of the four states reached their highest peak of unemployment rate. This might be due to the fact that New York got itself into a worse situation with comparatively more identified covid cases and deaths than California.

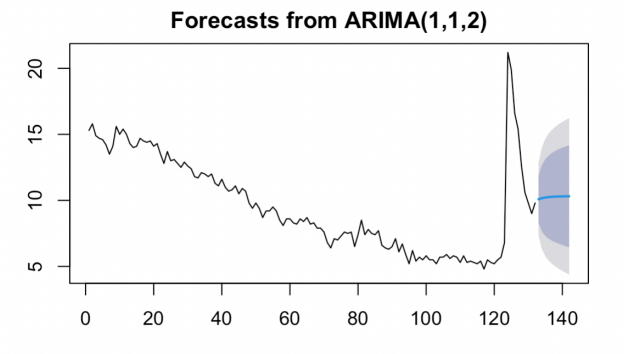
Additionally, use 15% as a dividing point, these states are divided into 2 groups. CA and NY had much higher unemployment rates than FL and TX. This result might be attributed to the fact that CA and NY are heavily reliant on leisure, entertainment, and tourism, which are the most vulnerable industries with the lockdown and social distancing policies under the Covid. States like Florida and Texas, though experiencing exponential increase in unemployment rate, has been less affected by the Covid than NY and CA.



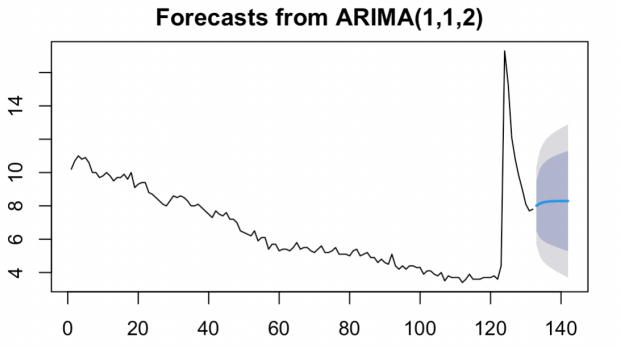
-Arima Model

Before application of the Auto ARIMA model, we have splitted the data into train and test groups. Train group includes all unemployment rates before 2021 Jan and the test group includes the unemployment rate from 2021 Jan until 2021 Dec. We later check the accuracy of the model by calculating rmse of test and train groups with forecasted results and original data.

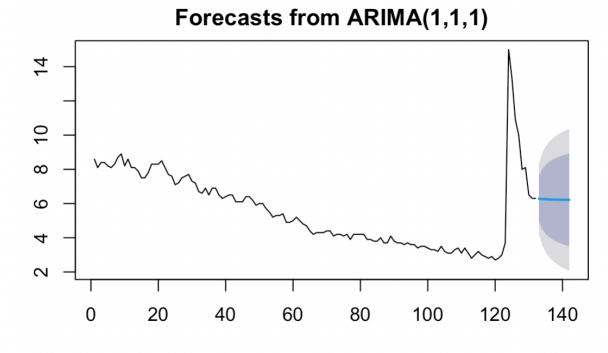
This is the Arima for Primary school. The trend of Primary school diplomas is basically keeping at the same level. In the real world, jobs like Starbucks assistants or gasoline workers are easy to replace. They have a huge chance to get fired if the business is affected by an external force. Thus when Covid came, the primary school's unemployment rate was incredibly high. Now that Covid is under control, the economy is recovering, and companies need more basic labor jobs, the unemployment rate of primary school will be relatively smooth again.



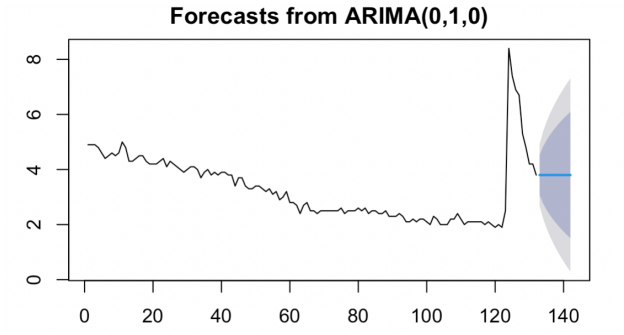
This is the Arima for High school. Unemployment for high school diplomas will slightly rise in the future. According to historical data, the unemployment rate in high schools is lower than that in primary schools. The high school unemployment rate has remained around 4.3, while the elementary school unemployment rate is around 6.4. The current situation may be due to companies needing cheaper labor forces to complete basic work in order to save money after the epidemic. The weekly salary for those with a high school diploma is about $781, and that of a junior high school diploma is $619. So the company may give priority to the admission of people with a primary school diploma.



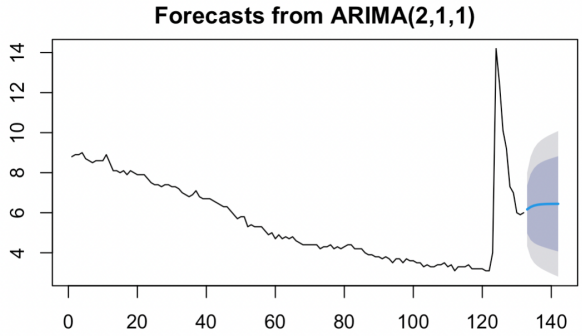
This is the Arima for the Associate degree. It is inevitable that companies will lay off staff due to the pandemic. But for those with a college degree, who have certain expertise, their unemployment rate will level off after the pandemic is over. Their future employment will not be affected by Covid.



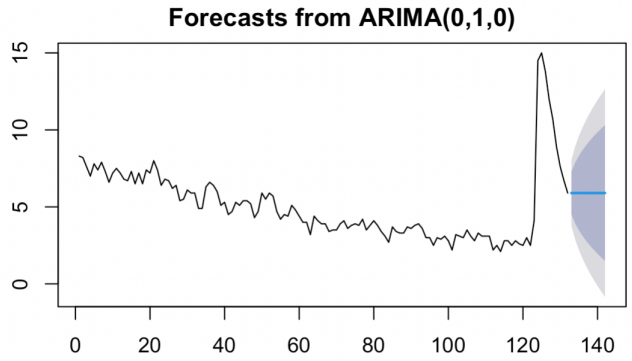
This is the Arima for Professional Degree. The projected future unemployment rate is smoothing out because this group is people with the highest intellectual background. They are harder to replace. They can still stand their ground regardless of external problems. Compared to the other three qualifications, the professional degree has the lowest unemployment rate, but the pandemic did raise it to its peak in recent years. It also proves that the pandemic has indeed affected the entire economic market.



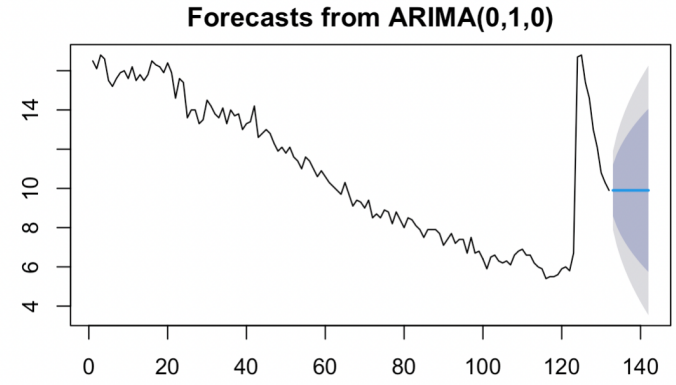
This is the Arima for White people. The future trend of White people will be a slight rise again. This shows that white people are affected by COVID-19. This may be due to more immigrants entering the United States. More people competing for limited jobs has contributed to a higher unemployment rate.



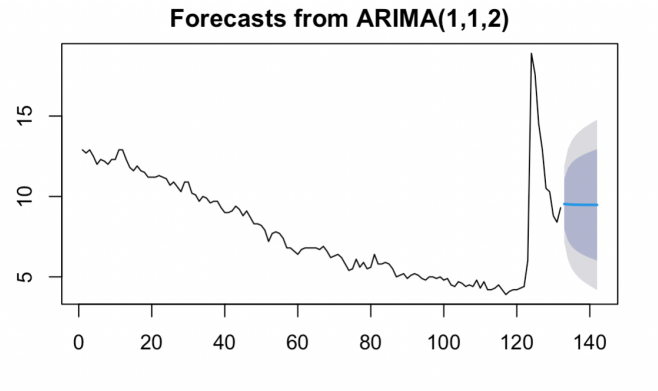
This is the Arima for Asian people. Future unemployment forecasts for Asian people are also smooth. Clearly, the market likes Asian people since they are hard-working. Asians had the lowest unemployment rate in the past. But when the pandemic hit, it climbed to unprecedented heights. This may be because the virus was discovered in China, and some local U.S. companies are laying off Asians for the fear of COVID-19. But after the epidemic passes, the Asian curve will return to normal.



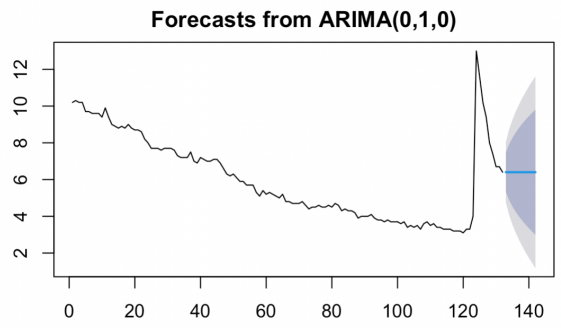
This is Arima for Black people. The future trend of Black people is smooth. Although the unemployment rate of Black people was significantly increased at the time of the Covid-19 outbreak, it has been declining for the past 10 years. And the future trends of people and Asians are relatively similar. This proves that society has achieved remarkable results in eliminating prejudice against race.

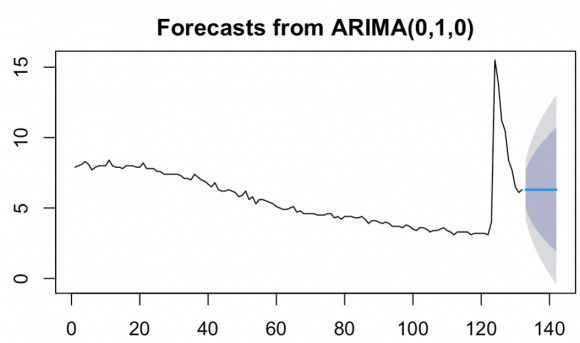


This is Arima for Hispanics. The future trend of the Hispanic population will have a slight downward trend. It could be because most Hispanic workers are more likely to have low-paid service positions and now most companies are in desperate need of cheap labor. Thus, the Hispanic population's unemployment rate goes down.



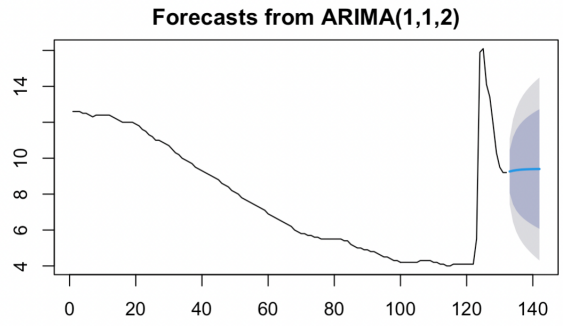
This is Arima for Men and Women. Future unemployment projections for both men and women trend horizontally. Comparing the unemployment rates of men and women, the unemployment rate of women is higher than that of men. Women are more affected by the epidemic than men. This may be due to gender inequality issues. Because women have long been in a weaker position in society, when the pandemic hit, women were still more unemployed than men, even when both genders lost their place in the market. But the pandemic will end one day, which is why the chart predicts that the unemployment rate will remain flat in the future.



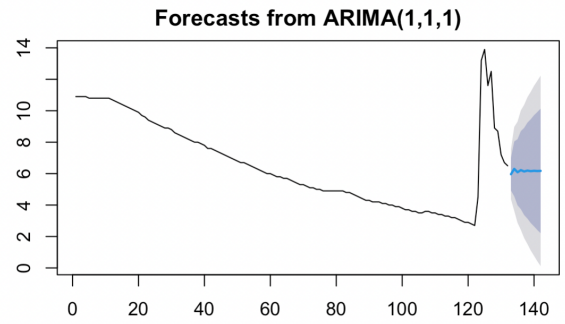


This is Arima for CA. Future projection trends first uprising and then horizontally. The result indicates that Covid possibly have long-lasting effects on CA's unemployment rate and contribute to CA's lagging employment recovery in the future.

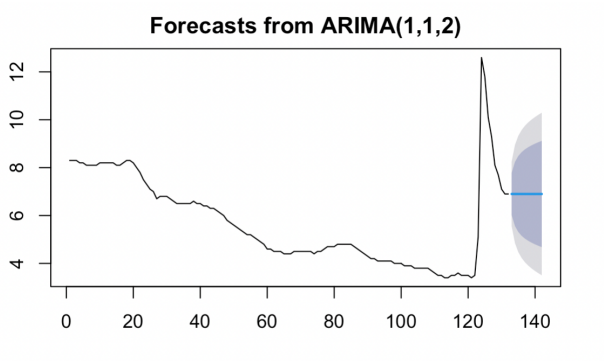
In fact, even pre-pandemic, California's overall unemployment rate was usually slightly above the national rate. During the pandemic, due to the fact that so many residents in CA work in the leisure and hospitality industries, there were massive layoffs at the beginning of the pandemic. California is also heavily reliant on tourism and entertainment, which highly depends on traveling and the interaction between people. Under Covid, the demise of tourism and entertainment leads to higher unemployment rate in California. It is reasonable to expect that a similar situation such as massive lockdown would lead to possible raising of unemployment.



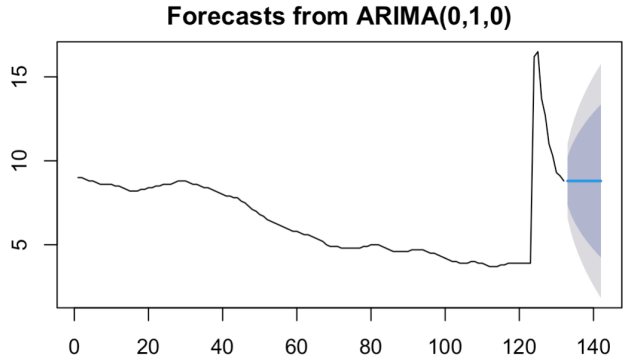
This is arima for Florida. The forecast shows an almost horizontal trend of unemployment rate for Florida. Compared to other states, Florida did not receive as much of the covid impact as other states, but still reached its highest unemployment rate in the last 30 years. Fortunately, the economy bounced back due to several rounds of massive federal stimulus money since May. Additionally, federal-driven revisions and annual U.S. Department of Labor revisions showed the state has recovered quicker than previously reported, and there was continued growth across all industries. As more people in Florida are encouraged to seek employment, the unemployment rate might slightly increase due to the increasing number of people entering the workforce and seeking job opportunities.



This is arima for Texas. The forecast shows a horizontal trend in Texas' unemployment rate. It indicates that after the covid, Texas' unemployment rate will continue to be stable in the following months. This is attributed to the fact that although many of the people living in Texas are facing obstacles when reentering the workforce, the Texas' government will address these problems via promoting new workforce policies. Especially for people living in south Texas where many of them work in the oil or mining fields. The stable predicted trend implies further capabilities for Texas to recover from the damage brought by Covid and to manage more opening positions in the job market.



This is arima for New York. The forecast shows a horizontal trend of New York's unemployment rate. The prediction implies possible economic recovery and a more stable workforce in NY. Due to its position in the world, the NY government provides more opening positions and has more opportunities. The predicted horizontal trend indicates that NY will recover well from the impact of Covid.



Conclusions & Recommendations

After analyzing the impact of pandemic on the unemployment data of the United States under different categories of people including educational level, race, gender, region by using linear regression model and Arima model, we are able to confirm the huge impact caused by pandemic to different categories of people and predict the future trend of unemployment as well.

Firstly, we used a linear regression model to examine whether covid-19 significantly impacted the unemployment rate in the U.S. and how the pandemic impacted different categories of people in gender, race and educational level. Based on the result, we concluded that covid-19 did significantly influence the unemployment rate in the U.S. The unemployment rate of people with different educational levels were all significantly impacted by the pandemic. Regarding race as a factor, the unemployment rate of all races were impacted significantly, except for Black people who maintained a high unemployment rate even before the appearance of pandemic. Moreover, in terms of gender, the impact on the unemployment rate toward females and males are similarly significant, with a little higher on females.

According to the Arima model, which is used to predict the impact of Covid-19 on the unemployment rate of different categories of people in the future, discrepancies in each category have appeared. People with different academic backgrounds have suffered losing their positions as the pandemic started and it is reasonable because the economy was once in a bad position, companies start to fire people maintaining a low cost. However, the higher degrees people have, the more stable their jobs tend to be. People with primary school and high school degrees tend to have higher unemployment rates, since they are easy to be replaced, so the market easily cuts them off if there is a pandemic or other big issue in society. Moreover, the predictions of the Arima model indicate that the unemployment rate of people with different academic backgrounds will become smooth in the future, except for people with high school degrees, whose unemployment rate might rise slightly. As the pandemic gradually starts to be controlled and the economy recovers, the employment market tends to return to normal.

Regarding the factor of race, the unemployment rate of all races including White, Asian, Black and Hispanic people were all affected by the Covid-19. Asian people's employment problems were impacted a lot as well during the pandemic, which might be hypothesized as "asian hate", since the virus was first discovered in Asia. Predictions show that the future trends of the unemployment of Black and Asian people are similar, which tend to be smooth, but the future trend of unemployment of White people seems to rise. On the contrary, the future trend of unemployment of Hispanic people tends to decrease according to prediction. We conjecture that Hispanic workers are more likely to have low-paid jobs which are highly needed post pandemic. In terms of gender as a factor, the predictions of the Arima model shows that the unemployment rate of both men and women tends to go flat in the future. While comparing the unemployment rates of men and women during the pandemic, women tend to experience higher impacts on career from the Covid-19.

To compare the impacts brought by covid on different regions, we examine covid case data of four representative states in the U.S, including New York, Texas, Florida and California. Through analysis, we discovered that the future trend of unemployment rate in CA will rise, which suggests that the pandemic might have a long-lasting impact on the employment market in CA. Moreover, Texas might face a similar problem that the future trend of unemployment might rise and it might still need time to recover. As for the unemployment in Florida, the forecast shows a smooth trend in general, but a little fluctuation. Compared to other states, Florida did not experience as much of the covid impact as other states. The future trend of unemployment in New York and Texas are similar, nearly horizontal, which implies possible economic recovery and a more stable workforce in these two states.

Through the analysis and prediction of unemployment rate of different demographic groups under the impacts of pandemic, we are able to aware that although the impacts brought by the

pandemic to the employment of each demographic groups are huge, it generally will return back to its normal status except for several demographic groups such as people with high school degree, White people, and people in CA. The result of the Arima model indicates an uprising trend in the unemployment rate in the future. Thus, to better deal with these inequalities , some public actions can be an effective way to accelerate legal measures and policies. Leaders from reputable companies can put out public statements to accelerate social change which might require governmental and society support. For example, they can advocate for the expansion of unemployment benefits and building economic relief programs for particular regions or colored people whose employment were significantly impacted. At the organizational level, it is recommended to provide tailored support for employees who were most impacted by the pandemic by building stronger company camaraderie and at the same time challenge discrimnation behavior through effective training programs or strict intervention and punishment. In addition, nowadays the pandemic is accelerating remote-work and independent-work platforms, which make digital literacy much more important. Business leaders and governments are expected to address race and gender stereotypes that might inhibit some groups of people's access to digital techs and improve their digital literacy. Also, measures to promote gender and race diversity in funding for entrepreneurs of minorities are encouraged to further eliminate biases in recruitment and selection processes. Lastly, people with low educational levels such as people with primary and high school degrees are encouraged to get exposed to educational opportunities to acquire more skills and expertise.

References

Bureau Labor Statistic. (2010-2021). *Topic: Effects of COVID-19 Pandemic on the Employment Situation News Release and Data.* Bureau Labor Statistic. Retrieved December 2021, from https://www.bls.gov/covid19/effects-of-covid-19-pandemic-and-response-on-the-employment-situation-news-release.htm

Edwards, R., & Smith, S. M. (2020, April). Job market remains tight in 2019, as the unemployment rate falls to its lowest level since 1969 : Monthly Labor Review. U.S. Bureau of Labor Statistics. Retrieved March 3, 2022, from https://www.bls.gov/opub/mlr/2020/article/job-market-remains-tight-in-2019-as-the-unemployment-rate-falls-to-its-lowest-level-since-1969.htm

Altonji, J., & Blank, R. (1999). *Race and gender in the labor markets.* Amsterdam and Boston: Elsevier, North-Holland. Retrieved March 3, 2022, from https://www.federalreserve.gov/econres/feds/files/2021080pap.pdf

Adjeiwaa-Manu, N. (2017, August). *Unemployment Data by Race and Ethnicity.* Center for Global Policy Solutions.Retrieved March 3, 2022, from https://globalpolicysolutions.org/wp-content/uploads/2017/07/Unemployment-Data-by-Race.pdf

Farber, Henry S. 1999. *Mobility and Stability: The Dynamics of Job Change in Labor Markets.* Handbook of Labor Economics vol. 3. New York: Elsevier.

Alon, T., Doepke, M., Olmstead-Rumsey, J., & Tertilt, M. (2020a). *This time it's different: The role of women's employment in a pandemic recession* (NBER Working Paper 27660). Cambridge, MA: National Bureau of Economic Research.

Villarreal, A., & Yu, W.-hsin. (2022, February 1). *Research note: Gender differences in employment during the covid-19 epidemic*. Demography. Retrieved March 3, 2022, from https://read.dukeupress.edu/demography/article/59/1/13/293749/Research-Note-Gender-Differences-in-Employment

Gedye, G. (2021, December 10). Why does California have the highest jobless rate in the country? CalMatters. Retrieved April 25, 2022, from https://calmatters.org/economy/2021/12/california-unemployment-rate/

Turner , J.. *Florida unemployment rate at 3.5%*. https://www.wctv.tv. Retrieved April 25, 2022, from https://www.wctv.tv/2022/03/14/florida-unemployment-rate-35/

Novak, A., Ferman, M., & Cai, M. (2020, April 17). *Texas unemployment rate: How coronavirus impacted the economy*. The Texas Tribune. Retrieved April 26, 2022, from https://apps.texastribune.org/features/2020/texas-unemployment/#:~:text=Texas%20lost%201.4%20million%20jobs,billion%20in%20revenue%20were%20regained.

Florida, U. of S. (n.d.). *2021 E-Insights Report*. Key Insights from COVID-19 Analysis | 2021 E-Insights Report | State of the Region | USF Muma College of Business. Retrieved April 26, 2022, from https://www.usf.edu/business/state-of-the-region/e-insights-2021/section-2-06-key-insights-from-covid-19-analysis.aspx

Data Sources:

U.S. Bureau of Labor Statistics. (n.d.). Monthly employment situation news releases and covid-19 impact summaries. U.S. Bureau of Labor Statistics. Retrieved April 28, 2022, from https://www.bls.gov/covid19/effects-of-covid-19-pandemic-and-response-on-the-employment-situation-news-release.htm

Department of Consumer and Worker Protection (DCWP). (2022, April 22). Consumer Services Mediated Complaints: NYC open data. Consumer Services Mediated Complaints | NYC Open Data. Retrieved April 28, 2022, from https://data.cityofnewyork.us/Business/Consumer-Services-Mediated-Complaints/nre2-6m2s

CSSEGISandData. (n.d.). Covid-19/csse_covid_19_data/csse_covid_19_time_series at master · CSSEGISANDDATA/covid-19. GitHub. Retrieved April 28, 2022, from https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series