Prescriptive Analysis of Unemployment

A Project Report

Submitted by

Jeet Shah
Hitansh Shah
Vansh Shah

Under the Guidance of

Rachit Garg and Amey Raut

in partial fulfillment for the Project in

Implementation of Technology

Computer Engineering
At

Mukesh Patel School of Technology Management and Engineering, Mumbai

October 2021

DECLARATION

I, **Vansh Shah**, Roll No. **C093** B.Tech Integrated (Computer Engineering), V semester understand that plagiarism is defined as anyone or combination of the following:

1. Un-credited verbatim copying of individual sentences, paragraphs or illustration (such as graphs, diagrams, etc.) from any source, published or unpublished, including the internet.
2. Un-credited improper paraphrasing of pages paragraphs (changing a few words phrases, or rearranging the original sentence order)
3. Credited verbatim copying of a major portion of a paper (or thesis chapter) without clear delineation of who did wrote what. (Source: IEEE, The institute, Dec. 2004)
4. I have made sure that all the ideas, expressions, graphs, diagrams, etc., that are not a result of my work, are properly credited. Long phrases or sentences that had to be used verbatim from published literature have been clearly identified using quotation marks.
5. I affirm that no portion of my work can be considered as plagiarism and I take full responsibility if such a complaint occurs. I understand fully well that the guide of the seminar/ project report may not be in a position to check for the possibility of such incidences of plagiarism in this body of work.

Signature of the Students: 

Names: Jeet Shah, Hitansh Shah, Vansh Shah

Roll Nos. C088, C091, C093

Place: Mumbai

Date: 24-10-2021

Annexure-III

CERTIFICATE

This is to certify that the project entitled “Prescriptive Analysis of unemployment” is the bonafide work carried out by Vansh Shah of B.Tech Integrated, MPSTME (NMIMS), Mumbai, during the V semester of the academic year 2020-2021, in partial fulfillment of the requirements for the Course Implementation of Technology.

_____

<Name of the Mentor>

Internal Mentor

_____ _____

Examiner 1 Examiner 2

Table of contents

INTRODUCTION

COVID-19, declared by the World Health Organization as a Public Health Emergency of International Concern, has claimed over 2.7 million lives worldwide. In the absence of vaccinations, social distancing and lockdowns emerged as the means to curb infection spread, with the downside of bringing the world economy to a standstill.

This pandemic has increased the unemployment rates across the world drastically due to various factors such as lockdown, migration to online environment, psychological factors such as fear of catching the disease, etc.

Our analysis demonstrates that the declining GDP captures the job loss trends in the COVID-19 period fairly accurately. The sudden surge in the COVID numbers made contagion containment an absolute imperative and greatly hurt the revenue derived from the leisure, education, construction and government sectors. This finding is corroborated by the high job loss trends in the bars and diners. Through regression studies, we observe that states that relaxed lockdown regulations tend to exhibit lower job losses, albeit at a higher contagion rate.

Literature Survey

Sr No	Title of Paper or Article	Authors	Concepts discussed in paper	Concepts use-full in project
1)	Identifying key indicators of job loss trends during COVID-19 and beyond	Satyaki Roy , Ronojoy Dutta and Preetam Ghosh	1) Multiple linear regression 2) Pearson Correlation Coefficient 3) Frequency analysis 4) Regression analysis 5) Utilizing the GDP statistics to compare the impending economic crisis with that of the great depression of 2008–09 6) Relationship between unemployment rate and GDP 7) Sectors and ethnic groups affected by job losses	1) Multiple linear regression 2) Regression analysis 3) Sectors and ethnic groups affected by job losses
2)	Projecting Unemployment Durations: A Factor-Flows Simulation Approach With Application to the COVID-19 Recession	Gabriel Chodorow-Reich, John Coglianese	1) Estimate individual transition hazards across employment, temporary layoff, permanent layoff, quitter, entrant, and out of the labor force, to forecast the duration distribution of unemployment 2) Relate the aggregate components to the overall unemployment rate using a factor model to forecast the duration distribution of unemployment. 3) Combine the individual duration dependence, factor structure, and an auxiliary forecast of the unemployment rate	1) Methodologies provided helps to plot garphs which is helpful to forcast duration distribution of unemployment 2) Considering the patterns we can distinguish the numbers of unemployment and re – employment 3) This paper also points out various state insurance benefits that an individual can claim during unemployment .

3)	Unemployment in the time of COVID-19: A research agenda	David L. Blusteina, Ryan Duffyb, Joaquim A. Ferreirac, Valerie Cohen-Scalid, Rachel Gali Cinamone, Blake A. Allanf	1) Conditional factors causing unemployment during Covid-19. 2) Psychological factors causing unemployment during Covid-19. 3) Economic factors causing unemployment during Covid-19. Strategies to deal with the unemployment caused by Covid-19.	1) The paper gives information on the factors causing Unemployment which are useful to plot a graph of unemployment rate. 2) By providing strategies, a prediction model predicting a 100% success graph can be plotted. Taking the factors into account, medians can be plotted around steep slopes of the graph indicating the factor in action.

Paper 1: Identifying key indicators of job loss trends during COVID-19 and beyond

Description:

COVID-19, declared by the World Health Organization as a Public Health Emergency of International Concern, has claimed over 2.7 million lives worldwide. In the absence of vaccinations, social distancing and lockdowns emerged as the means to curb infection spread, with the downside of bringing the world economy to a standstill.

In this work, the authors explore the epidemiological, socioeconomic and demographic factors affecting the unemployment rates of United States that may contribute towards policymaking to contain contagion and mortality while balancing the economy in the future. They identify the ethnic groups and job sectors that are affected by the pandemic and demonstrate that Gross Domestic Product (GDP), race, age group, lockdown severity and infected count are the key indicators of post-COVID job loss trends.

In this work, the authors have compiled comprehensive datasets on the businesses in US affected by COVID-19, ethnic groups employed by a sector and population data from the US Census Bureau Estimates and diverse features from the 50 states of USA – this dataset includes features such as airport traffic, homeless numbers and variations in lockdown dates. They collate these datasets to identify the specific ethnic groups and job sectors that are worse affected by the pandemic. They utilize the GDP statistics to compare the impending economic crisis with that of the great depression of 2008–09. Following this, they apply multiple linear regression to show that race, age group, lockdown severity and infected count play a significant role in determining the job loss trends in each state.

- **Summary of the methods:** They estimate the Pearson Correlation Coefficient to estimate the relationship between GDP and job loss trends. They also perform frequency analysis to connect the reports of the specific employment sectors in the COVID era with those from prior recession trends. They utilize regression analysis to identify the various socioeconomic and demographic factors contributing towards the unemployment in the US states measured in terms of the job loss frequencies across each US state.

- **Observations:** Our analysis demonstrates that the declining GDP captures the job loss trends in the COVID-19 period fairly accurately. The sudden surge in the COVID numbers made contagion containment an absolute imperative and greatly hurt the revenue derived from the leisure, education, construction and government sectors. This finding is corroborated by the high job loss trends in the bars and diners. Through regression studies, we observe that states that relaxed lockdown regulations tend to exhibit lower job losses, albeit at a higher contagion rate.

Paper 2: Projecting Unemployment Durations: A Factor-Flows Simulation Approach with Application to the COVID-19 Recession

Description:

Projecting Unemployment Durations: A Factor-Flows Simulation Approach With Application to the COVID-19 Recession by Gabriel Chodorow-Reich & John Coglianese is a research paper based on long term unemployment poses unique challenges to society . The duration distribution is divided into 3 factors : estimate indivisual hazard across employment ,relate the aggregate components to the overall unemployment rate, combine the individual duration dependence

From the above paper we can analyse the data and plot the graph of unemployment and reemployment rate using various algorithm and have a predictive and statistical study of the duration and distribution of unemployment rate during and after the COVID-19 pandemic

Hence in this paper we can see rise and fall of unemployment , A larger number of individuals remain unemployed for very long durations in the pessimistic scenario, providing a possible rationale for economic-based triggers to govern additional extensions. with the study and government aid this can benefit a large amount of unemployed individualds

Paper 3: Unemployment in the time of COVID-19: A research agenda

Description:

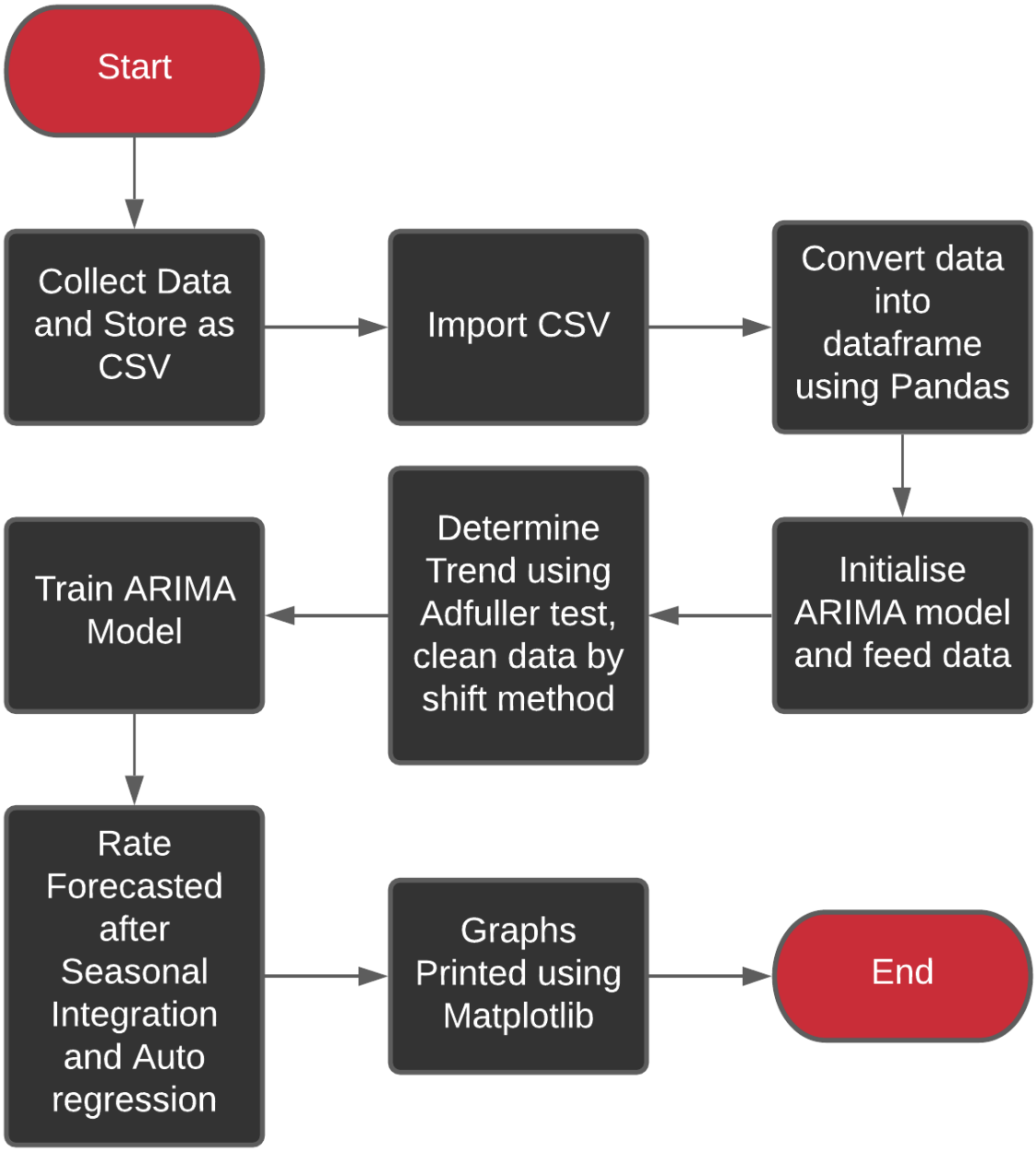
“Unemployment in the time of COVID-19: A research agenda” is a research paper published by Elsevier Publishing which explains the main factors causing mass unemployment during Covid-19. The factors have been divided into three categories namely : Conditional, Psychological and One’s Financial Well being. The final part of this research paper focusses on possible strategies that can be implemented by the Government to recover from such mass job less once normal or near normal conditions are restored.

From this research paper, one can create a predictive model by analyzing the factors given for the rise of unemployment rate, and the frequency of such events occurring, hence slopes on the graph

can be made. With the statistical data and the past schemes launched by Government, one can also predict the success of each scheme and also suggest the changes that can be made to that scheme to make it more successful. This paper has given strategies to combat the rise of the unemployment rate and also explained the process of job creation during such scenario.

This paper also tackles the psychological reason for unemployed people to not look for a job during a pandemic situation. Factors like fear of exposure to virus and fear of transmitting the virus to one's family are the main factors why people would not look for a hands on job during such a pandemic scenario. This data helps in predicting the rise in unemployment rate in the industrial sector where presence of the employee is required. Thus using the data provided by the research paper, a predictive model can be made.

Methodology & Design

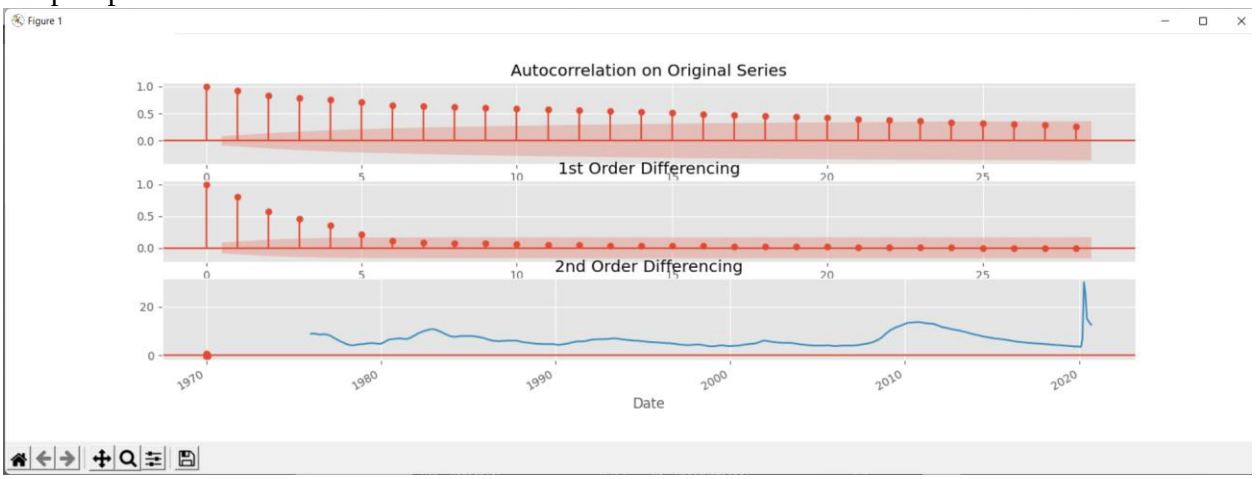


Result & Analysis

- P value : obtained using adfuler test to check if the data is stationary or not

```
p-value :  0.21210605035299013
p-value :  0.016362993009401496
p-value :  3.845014724056l3e-08
```

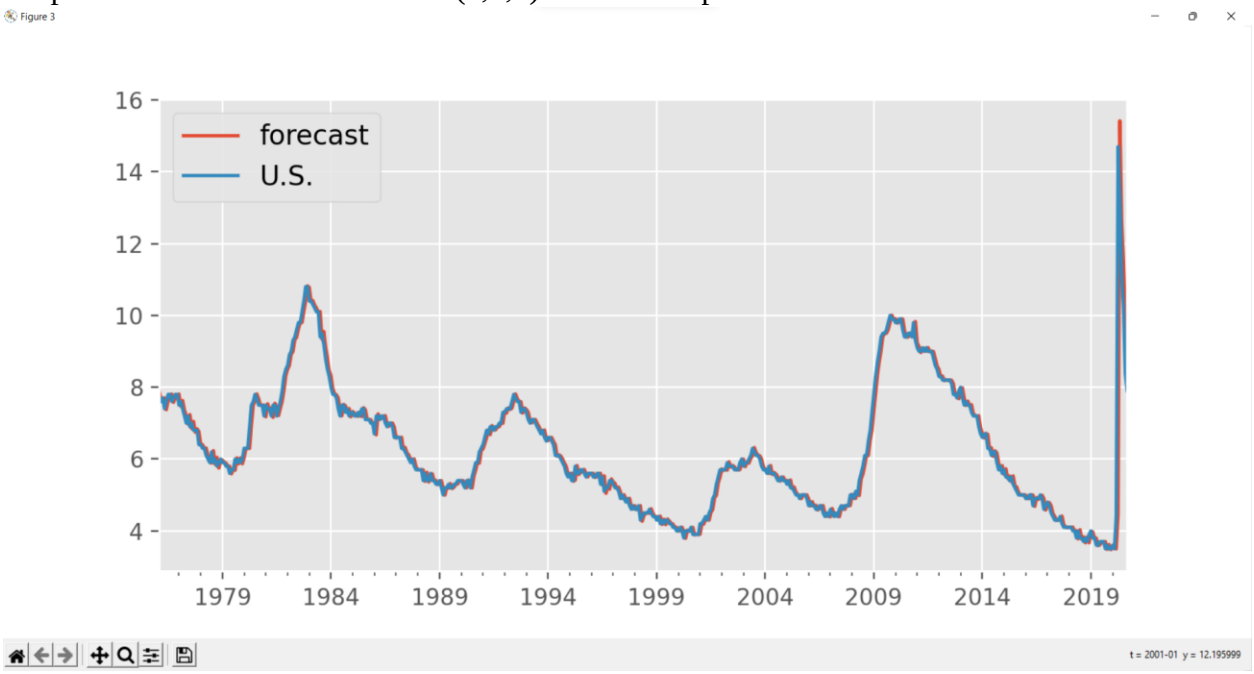
- Graph : plots the P value obtained before and after differentiation



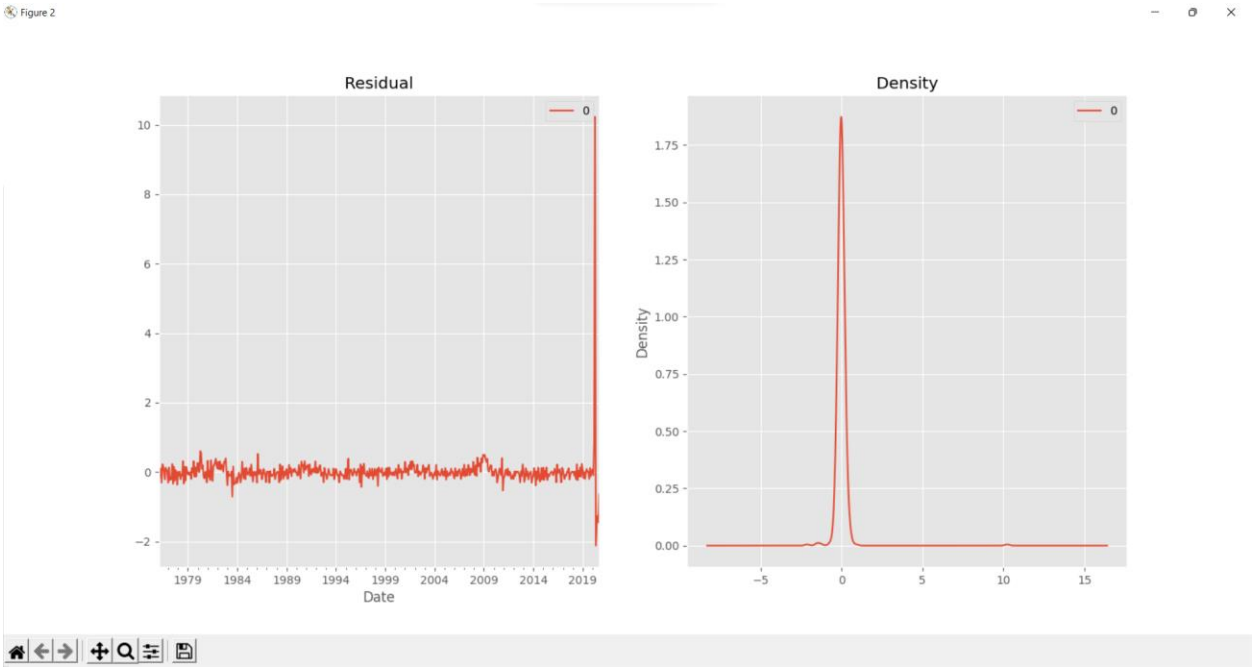
- ARIMA SUMMARY1: Summary or Result of Auto Regressive Model with (p,d,q) as (1,1,1)

```
                             ARIMA Model Results
==============================================================================
Dep. Variable:                 D.U.S.   No. Observations:                  536
Model:                 ARIMA(1, 1, 1)   Log Likelihood                -384.006
Method:                       css-mle   S.D. of innovations              0.495
Date:                Sun, 24 Oct 2021   AIC                            776.012
Time:                        20:26:48   BIC                            793.148
Sample:                    02-01-1976   HQIC                           782.716
                         - 09-01-2020
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.0001      0.022     -0.007      0.995      -0.044       0.044
ar.L1.D.U.S.  -0.7118      0.185     -3.852      0.000      -1.074      -0.350
ma.L1.D.U.S.   0.7870      0.166      4.751      0.000       0.462       1.112
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1           -1.4049           +0.0000j            1.4049            0.5000
MA.1           -1.2707           +0.0000j            1.2707            0.5000
-----------------------------------------------------------------------------
```

- Graph : Plots the result of ARIMA (1,1,1) model with prediction



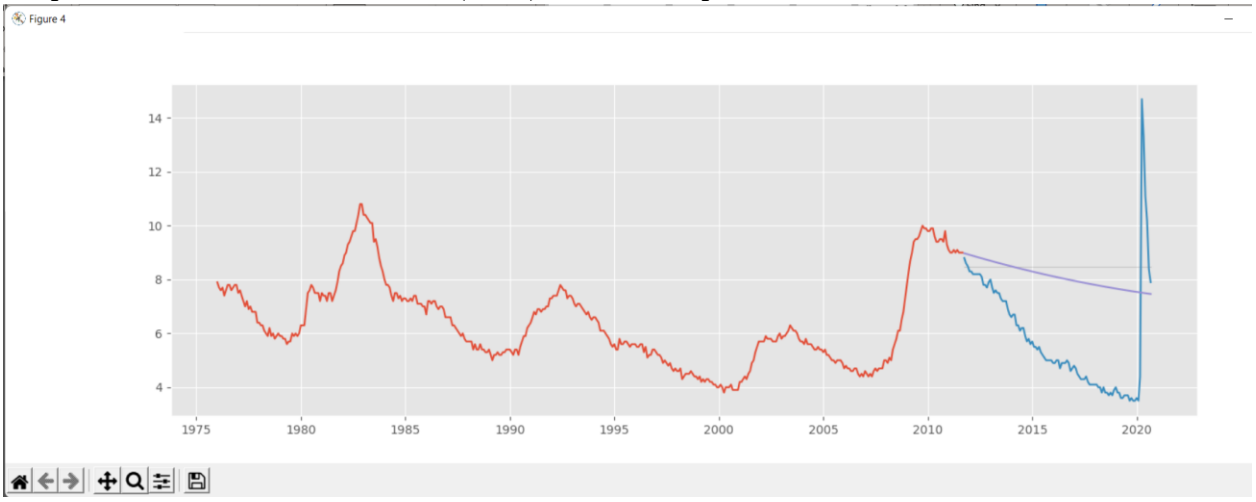
- Graph : plots residual and Density of ARIMA (1,1,1) model



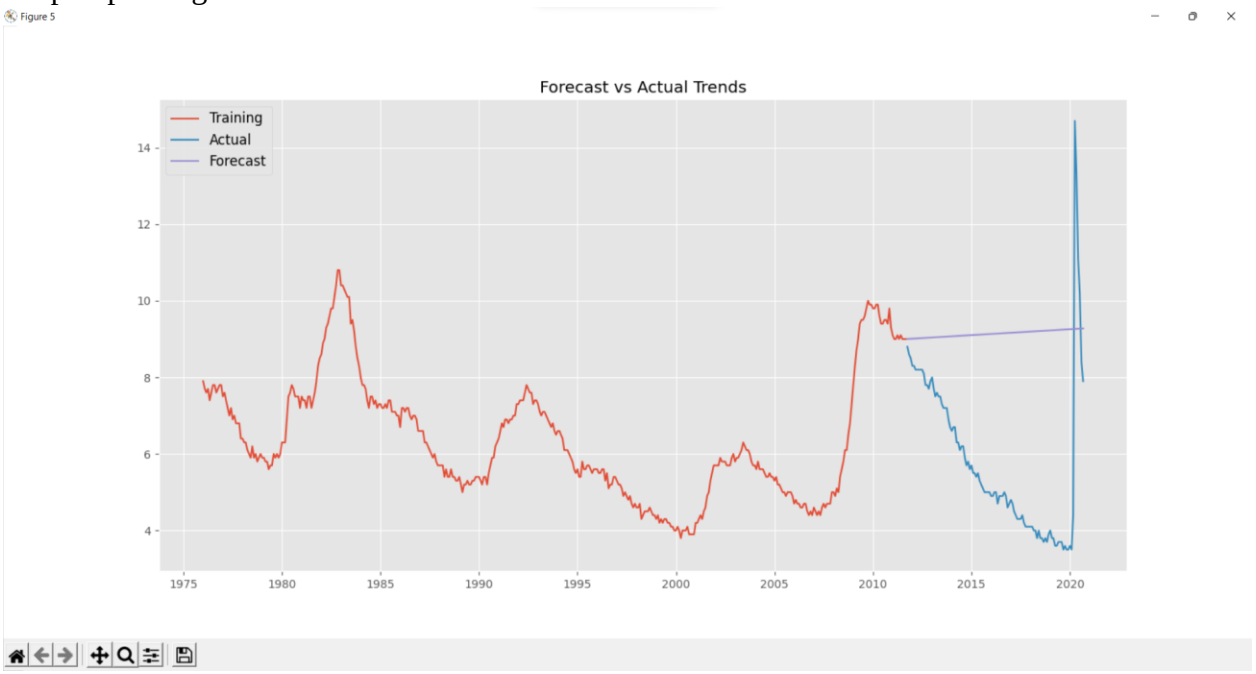
- ARIMA SUMMARY1: Summary or Result of Auto Regressive Model with (p,d,q) as (3,2,1)

```
                             ARIMA Model Results
==============================================================================
Dep. Variable:                D2.U.S.   No. Observations:                  427
Model:                 ARIMA(3, 2, 1)   Log Likelihood                 164.471
Method:                       css-mle   S.D. of innovations              0.164
Date:                Sun, 24 Oct 2021   AIC                           -316.941
Time:                        20:26:49   BIC                           -292.601
Sample:                  03-01-1976   HQIC                          -307.327
                       - 09-01-2011
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          9.465e-05      0.002      0.055      0.956      -0.003       0.003
ar.L1.D2.U.S.    -0.2503      0.109     -2.302      0.021      -0.463      -0.037
ar.L2.D2.U.S.    -0.0799      0.103     -0.772      0.440      -0.283       0.123
ar.L3.D2.U.S.     0.0223      0.079      0.283      0.777      -0.132       0.177
ma.L1.D2.U.S.    -0.7192      0.097     -7.446      0.000      -0.908      -0.530
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1           -1.4187           -2.2305j            2.6435           -0.3402
AR.2           -1.4187           +2.2305j            2.6435            0.3402
AR.3            6.4199           -0.0000j            6.4199           -0.0000
MA.1            1.3904           +0.0000j            1.3904            0.0000
-----------------------------------------------------------------------------
```

- Graph : Plots the result of ARIMA (3,2,1) model with prediction



- Graph : plotting Forecast Trend and Actual Trend

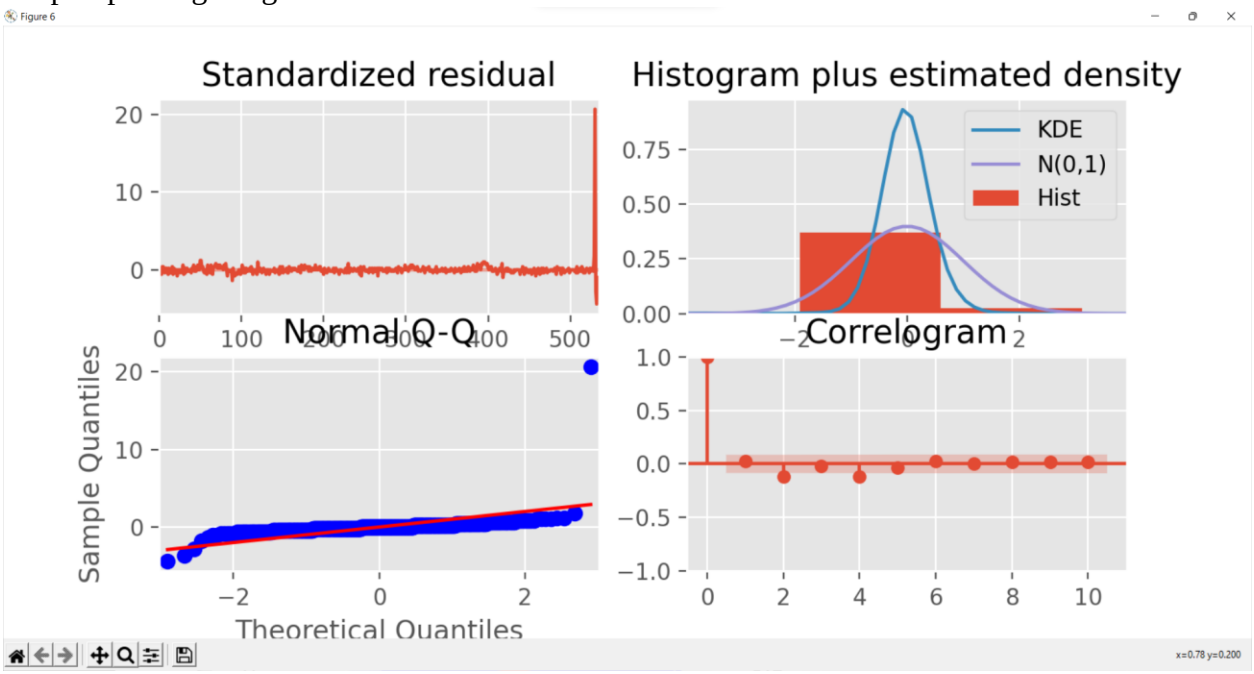


- ARIMA: Result of Auto ARIMA

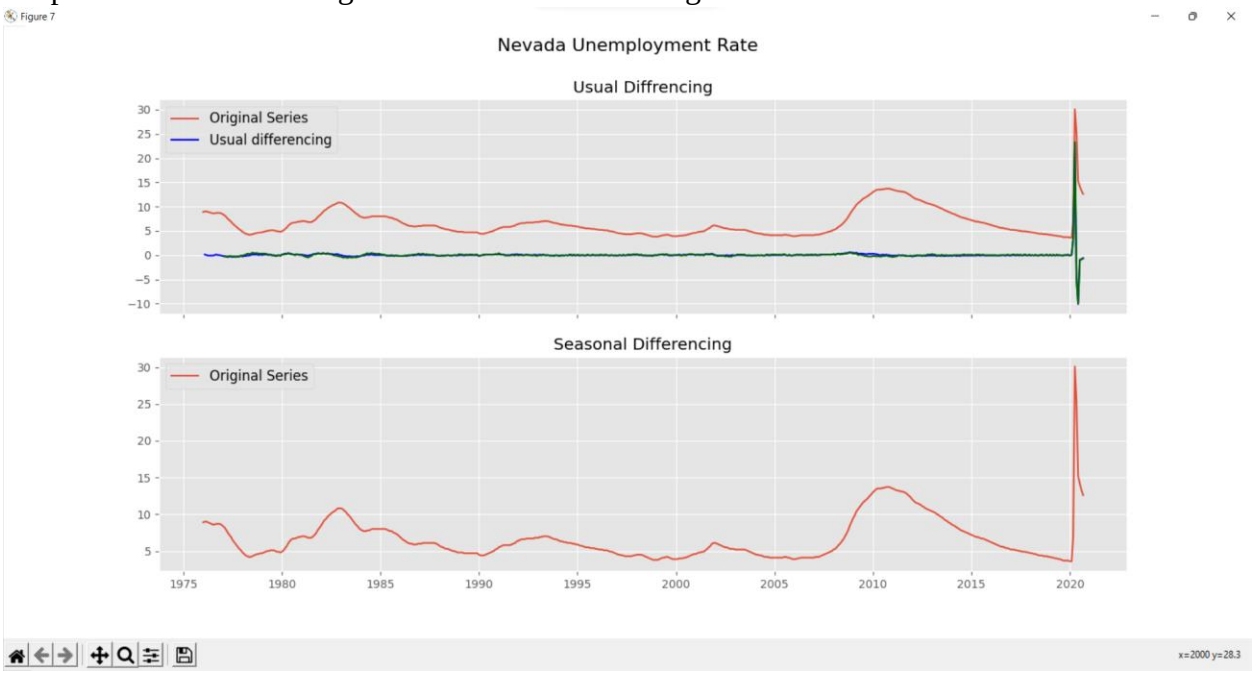
```
Performing stepwise search to minimize aic
 ARIMA(1,1,1)(0,0,0)[0] intercept   : AIC=776.012, Time=0.36 sec
 ARIMA(0,1,0)(0,0,0)[0] intercept   : AIC=777.953, Time=0.08 sec
 ARIMA(1,1,0)(0,0,0)[0] intercept   : AIC=779.484, Time=0.07 sec
 ARIMA(0,1,1)(0,0,0)[0] intercept   : AIC=779.337, Time=0.13 sec
 ARIMA(0,1,0)(0,0,0)[0]             : AIC=775.953, Time=0.03 sec

Best model:  ARIMA(0,1,0)(0,0,0)[0]
Total fit time: 0.696 seconds
```

- Graph : plotting Diagnostics



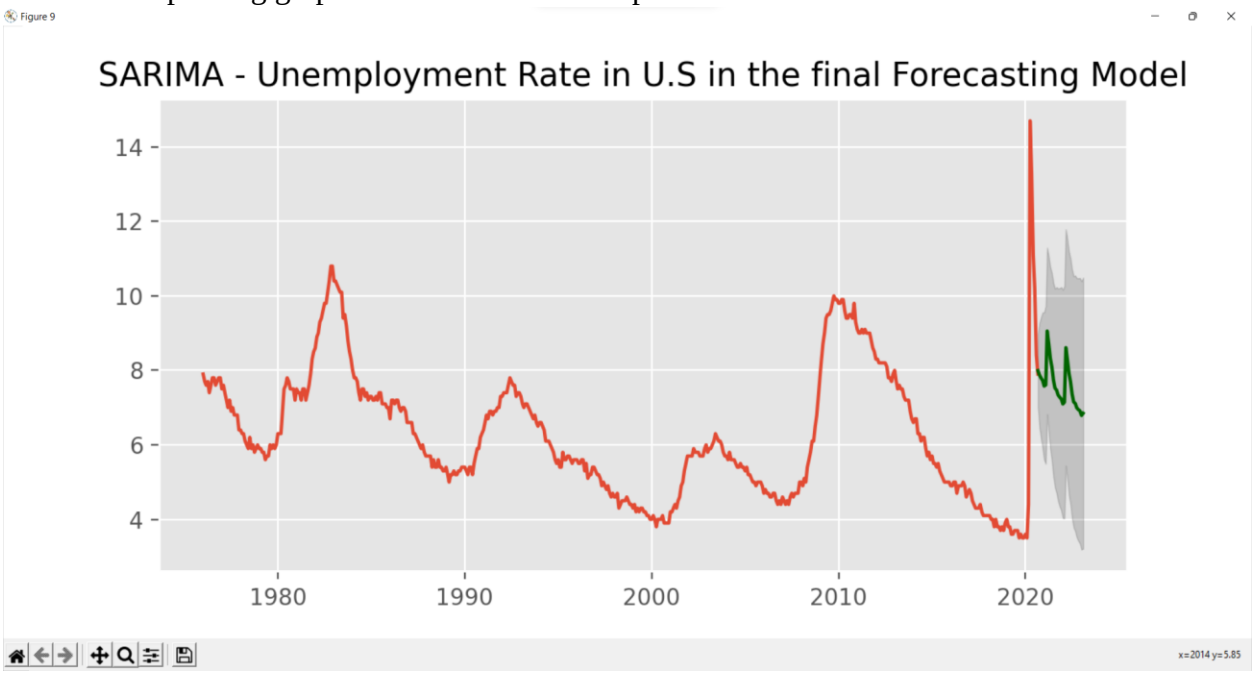
- Graph : Usual Differencing and Seasonal Differencing



- SARIMAX: Summary or Result of SARIMAX (p,d,q) as (0,1,0)

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  537
Model:               SARIMAX(0, 1, 0)   Log Likelihood                -386.976
Date:                Sun, 24 Oct 2021   AIC                            775.953
Time:                        20:26:50   BIC                            780.237
Sample:                             0   HQIC                           777.629
                                  - 537
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
sigma2         0.2481      0.001    213.883      0.000       0.246       0.250
===================================================================================
Ljung-Box (L1) (Q):                   0.47   Jarque-Bera (JB):           2595879.22
Prob(Q):                              0.49   Prob(JB):                         0.00
Heteroskedasticity (H):              17.82   Skew:                            16.21
Prob(H) (two-sided):                  0.00   Kurtosis:                       342.39
===================================================================================
```

- SARIMAX: plotting graph based on SARIMAX prediction



Conclusion

Our project of “Prescriptive Analysis of Unemployment Rate” goals to analyse various factors that caused this using technologies like data analysis, algorithms and regressive models. The analysis would then lead to prediction of future unemployment rates based on the schemes and factors and suggestion of what should be done on basis of the report.

Future Scope

- Furthermore, the trained model can be used as a reference for any similar events occurring in future and more such models can be created for calamities like natural disasters etc.
- Having prior knowledge of future unemployment rate can lead to creation and launching of effective relief schemes in case of a calamity.

References

- GabrielChodorow-Reich : May 2021
  https://ezproxy.svkm.ac.in:2054/science/article/abs/pii/S0047272721000347
- JohnCoglianeseb : May 2021
  https://www.sciencedirect.com/science/article/pii/S2590291121000590
- https://ezproxy.svkm.ac.in:2054/science/article/abs/pii/S0001879120300610
- https://becominghuman.ai/applying-machine-learning-to-forecast-the-unemployment-rate-in-australia-as-a-result-of-covid-19-b344f82ced4
- https://link.springer.com/epdf/10.1007/s10614-019-09908-9?author_access_token=G4jMN0uiQL-AnXWTrq9O-_e4RwlQNchNByi7wbcMAY5PNKGmpQuN7WXs4cV24Wkw6IhZemSli9OICzN1oY9lmuhy5bRR5mcUPMvTIruyfl_jd2nrVKzzRAS2DzMMyX-RLTIPx0xt5X3moenawNsVTQ%3D%3D