Name: Shwetha Srinivasan Natesan

G number: G01393194

Project Assignment 4 – Data Analytics Research Project

Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program

Abstract

This research aims to investigate the funds provided for the people in hospitals across the U.S affected by the corona virus during pandemic in the year 2020 based on three acts passed namely the bipartisan CARES Act, the Paycheck Protection Program and Health Care Enhancement Act (PPPHCEA), and the Coronavirus Response and Relief Supplemental Appropriations (CRRSA) Act. Totally the amount collected was $178 billion dollars and this was given to the hospitals and other health care providers. From this around $2 billion dollar was allocated by the Department of Health and Human Services to nursing home facilities that helped to reduce both Covid 19 infection rates relative their county and the mortality rates against a national benchmark through the Health Resources and Service Administration. This Fund program was named the Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program. The funds collected in each of the facility homes were contributed to reduce the infection rate in each sector. Therefore, the mortality rate due to the COVID -19 was reduced due to the proper treatment provided and allocation of beds in each of the residents. Further in this research paper, the analysis, interpretation, visualization of the data is done, and the required findings are explained.

I. Introduction

The Nursing Home Infection Control (NHIC) distribution is a component of a Provider Relief Fund (PRF) which is provided to the skilled nursing facilities and nursing homes nationwide to help fight the devastating effects of the Covid - 19 pandemic. There is an incentive payment plan called the Quality Incentive Payment (QIP) Program which was introduced to provide more funding to the nursing facilities in addition to the direct funding. These incentive payments were made based on certain performance measures and the payment can be used for the infection control expenses and the mortality rate reduction based on the Terms and Conditions applied. The QIP funding was done based on the outcome of two factors namely the COVID-19 Infection Performance Score and the COVID-19 Mortality Adjustment Rate. The Infection Performance Rate was measured by comparing the healthcare facility infection rate with the COVID – 19 infection rates for the entire county. This was done by the HRSA (Human Resources and Services Administration), and it was observed that the infection rate was greater than the county infection rate. Next the Mortality Rate was measured based on the number of admittances due to COVID-19 and the total number of infections based on the demographic characteristics of the residents. The mortality rates were reduced among the COVID facility, or the admissions based on the facility due to COVID.

There were four performance periods, and it was during September, October, November, and December 2020, for which the funds were distributed between November 2020 and February 2021. The performance period ranges for about 4 to 5 weeks according to the month. For the October month the performance period was about 5 weeks. The NHIC reporting is done based on the terms and conditions provided by the QIP which includes the cost associated with the COVID-19 testing, hiring staffs and employees for administrative support and patient care, to report the COVID-19 test results to the local, state and the federal government, Providing IT related support to the patients during their hospital days by helping them connect with their family and other expenses to improve the infection control and provide basic needs to the patients. The revenues which are lost are not reported and the interest earned via the NHIC payments should be reported separately. These payments can be used to prevent and prepare for the COVID – 19 infections. For the funding amount more than $500,000, the report must be submitted based on the quarter period, which is divided as Q1, Q2, Q3, Q4 from the month January to December and the payments is divided into two categories namely General and Administrative Expenses and Health Care Expenses. The General and Administrative Includes Rent, Insurance, Personnel, Benefits, Lease, and Utilities. The Health Care Expenses include Supplies, Equipment, Information Technology, and Facilities.

The Provider Fund data contains Nursing facility along with the Certification Number provided by the Medicare center for verification. The Total count of beds occupied by the patients during the performance period is denoted by the Total resident weeks. The Total Covid infections confirmed during the performance period, The Total Covid infections per Resident multiplied by 1000 is collected. The Average weekly number of county infections reported during performance period multiplied by 1000 and the infection performance score obtained from the different facility according to the county covid infection rate. The capped infection score compares the facility's performance score and incentive payment which are according to the terms and conditions. The most important factor, which is the Mortality Adjustment Rate is measured by the facility's infection score obtained from the mortality rate due to COVID-19. Lastly, the performance month which are September, October, November and December and the Final payment to each facility. From this information, this paper focuses on the research questions namely,

a. What are the characteristics of these acts which provided funds?
b. How was the infection rate measured and what are the measures taken to reduce it?
c. What are roles performed by the health care providers and how was the fund utilized to reduce the mortality rate?
d. What was the solution obtained from the Infection Rate and the Mortality rate based on the QIP funding?

Based on these questions the funding amount can be compared and the performance period can also be analyzed.

II. Literature Survey

The COVID -19 funding became popular during the pandemic period especially in the United States. There is a difference in reach approach of how the funds are divided and how it is being reached to the people. The purpose of this review is to compare and analyze the main differences and findings in each dataset and to bring out the common approach followed in each report.

The first report is the HHS Provider Relief Fund which has similar content as the Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program. The author for this report is HHS ASPA (Assistant Secretary for Public Affairs by US Department of Human and Health Services). Basically, the PRF money is provided by the HHS based on the acts passed during the pandemic period namely the bipartisan CARES Act, PPPHCEA and CRRSA acts. The total amount collected was $178 billion dollars from that the department allocated $50 million dollars to Medicare facilities and providers impacted by the corona virus in the PRF mode. In this the dataset consists of the US states and the health care providers in each state and city and how much payments they have accepted. There is a visualization shown which clearly shows each state and the amount collected by each provider. When we compare both the Provider Relief Fund, we can see that the common attributes are the state, city, and the amount collected. The difference would be is that the Nursing Home Incentive Care Program consists of extra attributes like zip code and detailed data about the Total Resident Weeks, Total Covid Infections, Facility Infection Rate Per 1000 Resident Weeks, County Infection Rate Per 1000 Resident Weeks, Infection Performance Score, Mortality Adjustment, and Performance Month. The main goal of the Nursing Fund is to reduce the infection caused by the covid 19 and to reduce the mortality rate. So, the other attributes like the Infection Rate per 1000 Resident Weeks and the Mortality Adjustment plays an important role in this report. Whereas in the HHS Provider Fund, the main columns are the payment collected and the Medi care provider. The number of times the provider has accepted the payment, and it is depicted as, if the provider has accepted the payment twice then both the payments are entered. If the provider has received two payments but has accepted only one, then that payment is entered. If the provider receives only one payment and accepted, it then it is entered. So, this would be the major difference between these two reports and the analysis would be that both are based on the fund acceptance, but one is for the nursing fund to work on the mortality rates and other depicts the payment acceptance by each healthcare providers and their acceptance count.

The second report is the Coronavirus Relief Fund (CRF). The CRF is a fund program which gives information about the fund amount collected by each sector, totally $150 billion dollars were collected by the CRF and about 964 recipients have accepted it. In this, 864 recipients are supposed to report under the CARES Act and about 817 prime recipients reported that they have received about $149.4 billion and in this about $148.9 billion dollars were sent to 88,864 sub recipients. The report consists of a visualization of the US map consisting of each state which shows the prime recipient amount collected and the number of prime recipients in that state. The dataset consists of each state, its prime recipient, which is the health care sector, total funds received, number of sub recipients and the sub – award amount. Further there is information

about the sub recipient details such as the name, sub – award amount, money spent, and in which category was the amount donated to. When two reports are compared the CRF talks about the entire funding resources in each sector like Nursing assistance, public healthcare support, food assistance, economic support, etc., needed during the pandemic period. In the Nursing Home Quality Incentive Program, it only talks about one sector which is nursing and infection prevention. So, from the analysis, the CRF is done for all the required sectors needed during the covid period and the fund collection and spendings by each recipient.

The third report is the Education Stabilization Fund (ESF). ESF was introduced by the US Department of Education to intercept, protect and to answer the impacts caused by the corona virus on the education of students. The ESF was established by the Coronavirus Aid, Relief, and Economic Security Act (CARES Act) and further it was signed into other pandemic acts. The amount provided by ESF was $263 billion dollars and this was implemented in the state recovery and reinstalment, and this was managed by the US Department of Education. The ESSR fund provided $189.5 billion dollars. This report gives details about each act which provided the fund to the education agencies, and it also shows each the fund provided to each category such as student property, universities, fund for improvement for education, etc. When compared the two datasets are totally different because this one is to improve the damage caused by the Covid 19 in the education sectors and to improve them, but the Nursing fund comes under the health care sector. The one common factor is Covid 19, and the funding introduced based on it. The con in this report will be that even though the states are mentioned, and which funding act provides it, the sector which received these funding payments are not given. So, this would be a drawback in this report because to which sector the money has been reached should be known so that it will make sense when the dataset is analyzed.

The significance of these reports is that the information given is clear, which is about the funding acts made in pandemic period and how the funding is done to different sectors. All the three reports consist of common thing that is the Covid 19 CARES Act from that the funding is done, in HHS the provides funds to the Medi care sectors in each state, the CRS is the overall funding in different sectors and the ESF is the mainly for education affected by the corona virus. So, the Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program gives clear and more information about the object mentioned when compared with the other three datasets. The CRS also gives detailed information about the subrecipient details which is more efficient to understand. The HHS also shows a detailed report but could have included more information.

III. Methods and Tools

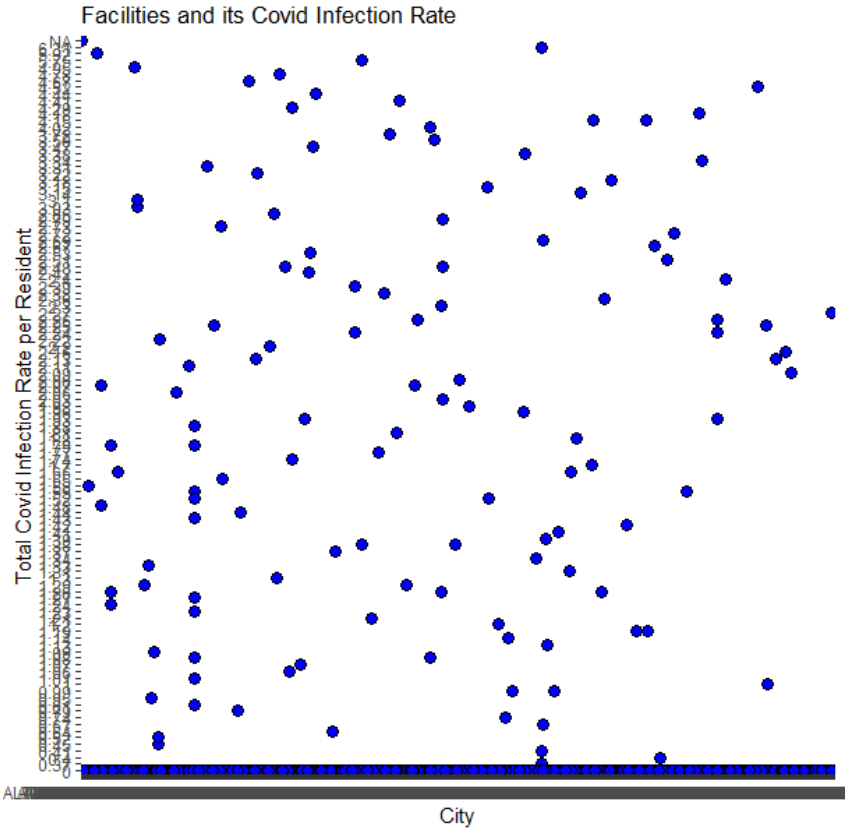
The Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program was available online to the public and it was provided by the Health Resources and Services Administration. The dataset was already in the csv format, so it was easier to load and interpret. The dataset

contains Geographical information such as the city, State and Zip code and the coverage was based on the US country. The dataset consists of 33,440 data values and 15 columns. Each row represents the Nursing Home Facility and the funding that occurred between September 2020 to December 2020. The size of the dataset is around 3930 KB which is equal to 3.93 MB, which is not quite large, but it was easier to load.

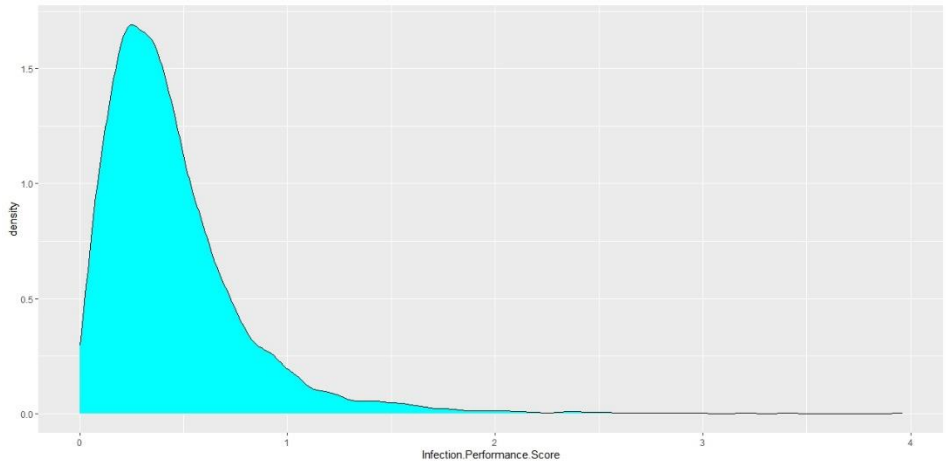
The tools used to analyze and visualize this dataset are R, Python and SQL. R is primarily used to do visualization because it was more efficient with the help of the ggplot function. In Python few visualizations and more statistical summaries were done. There wasn't any need to clean the data as the dataset was already proper with all the NOIR data types. The Facility Name, City, State, Performance Month and Zip code are Nominal, the Infection Performance Score and Total Resident Weeks are Ratio, the Certification Number is Ordinal, and the Mortality Adjustment is Interval. SQL was implemented in the MY SQL Workbench version 8.0 to load the data and perform few queries.

IV. Results

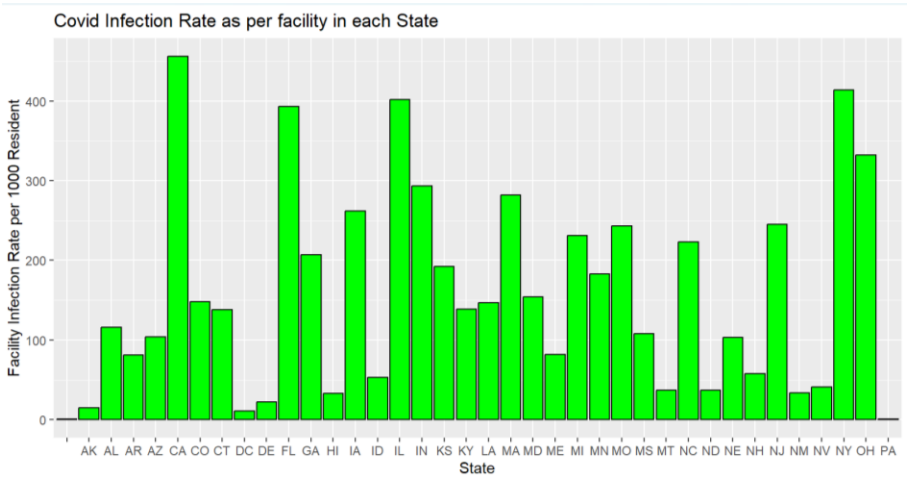
The Facility Infection Rate per 1000 resident and its corresponding city is plotted using scatter plot.



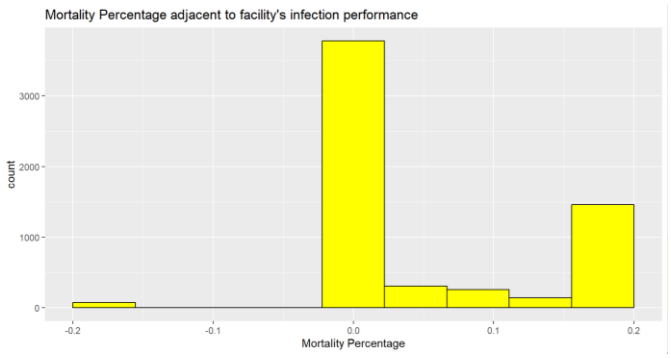
The density plot for the Infection Performance Score is done and the highest value would be 5.6.



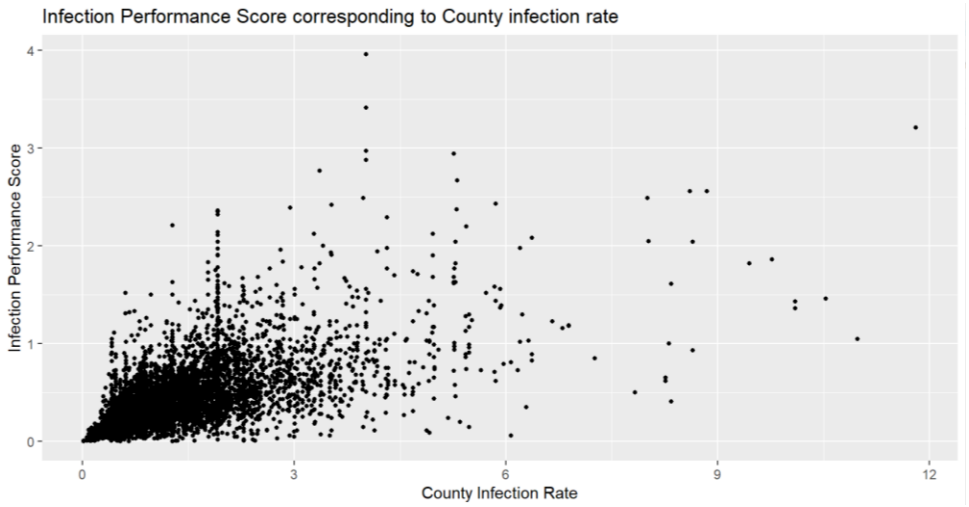
The bar graph for the Covid infection rate per resident in each state is plotted and California has the highest infection rate.



The mortality adjustment rate to facility's infection rate is measured using the histogram at one point both the values are at peak.



The scatter plot for the Infection performance score corresponding to the county infection rate is done.

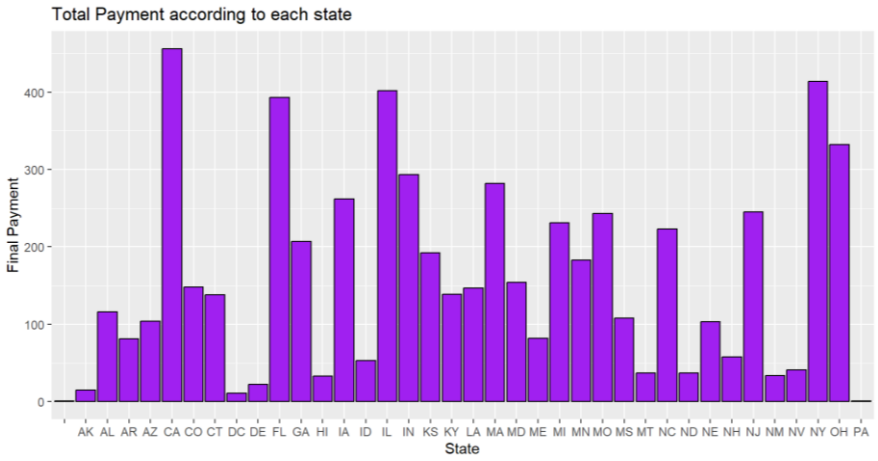


The summary of each values is done using the summary function and each of the statistical values are obtained.

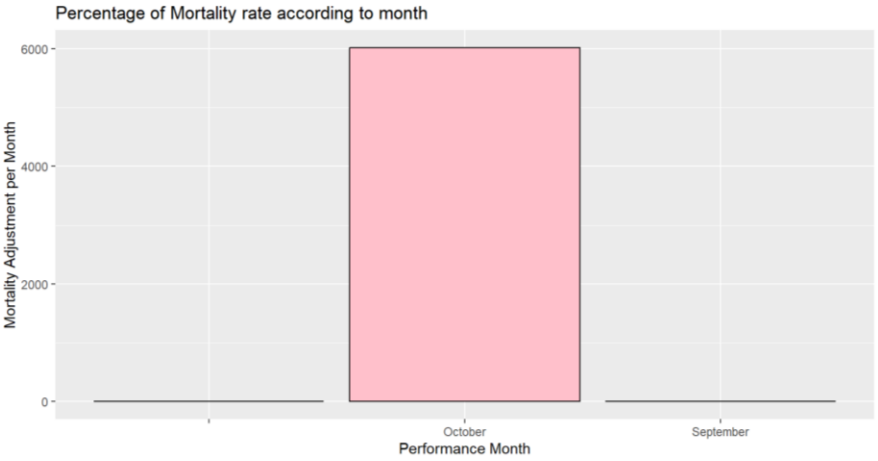
```
Console  Terminal  Background Jobs
R 4.2.2 · ~/AIT R/
   Record.No         CCN             Facility.Name         Address
 Min.   :   1   Length:6021        Length:6021        Length:6021
 1st Qu.:1506   Class :character   Class :character   Class :character
 Median :3011   Mode  :character   Mode  :character   Mode  :character
 Mean   :3011
 3rd Qu.:4516
 Max.   :6021

     City              State              Zip.Code     Total.Resident.Weeks
 Length:6021        Length:6021        Min.   : 1001   Min.   :   2.0
 Class :character   Class :character   1st Qu.:28025   1st Qu.: 224.0
 Mode  :character   Mode  :character   Median :46306   Median : 338.0
                                       Mean   :45831   Mean   : 378.7
                                       3rd Qu.:64468   3rd Qu.: 470.0
                                       Max.   :99901   Max.   :2705.0
                                       NA's   :1       NA's   :1
 Total.Covid.Infections Facility.Infection.Rate.Per.1000.Resident.Weeks
 Min.   :0.00000        Min.   :0.00000
 1st Qu.:0.00000        1st Qu.:0.00000
 Median :0.00000        Median :0.00000
 Mean   :0.02243        Mean   :0.04553
 3rd Qu.:0.00000        3rd Qu.:0.00000
```

```
Console  Terminal  Background Jobs
R 4.2.2 · ~/AIT R/
 County.Infection.Rate.Per.1000.Resident.Weeks Infection.Performance.Score
 Min.   : 0.020                                Min.   :0.0000
 1st Qu.: 0.710                                1st Qu.:0.2200
 Median : 1.090                                Median :0.3700
 Mean   : 1.392                                Mean   :0.4457
 3rd Qu.: 1.840                                3rd Qu.:0.5700
 Max.   :11.810                                Max.   :3.9600
 NA's   :1                                     NA's   :1
 Infection.Performance.Score.Capped Mortality.Adjustment Performance.Month
 Mode:logical                       Min.   :-0.20000     Length:6021
 NA's:6021                          1st Qu.: 0.00000     Class :character
                                    Median : 0.00000     Mode  :character
                                    Mean   : 0.05497
                                    3rd Qu.: 0.14000
                                    Max.   : 0.20000
                                    NA's   :1

 Final.Payment
 Min.   :   170
 1st Qu.: 29859
 Median : 49889
 Mean   : 60250
 3rd Qu.: 76979
 Max.   :507533
```

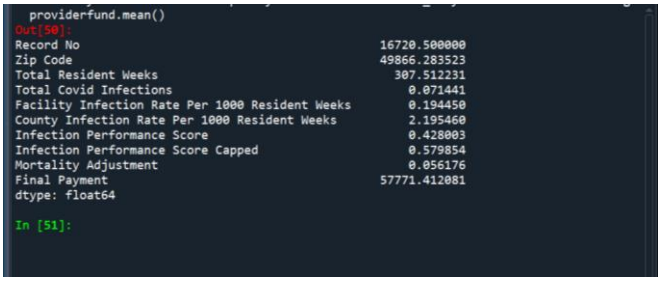
The Total Funding Payment provided by the QIP is calculated and the highest state is California having $718,593.



The Percentage of the Mortality rate corresponding to the performance month is shown where October has the highest mortality rate.



The mean values of each variable is calculated using Python.

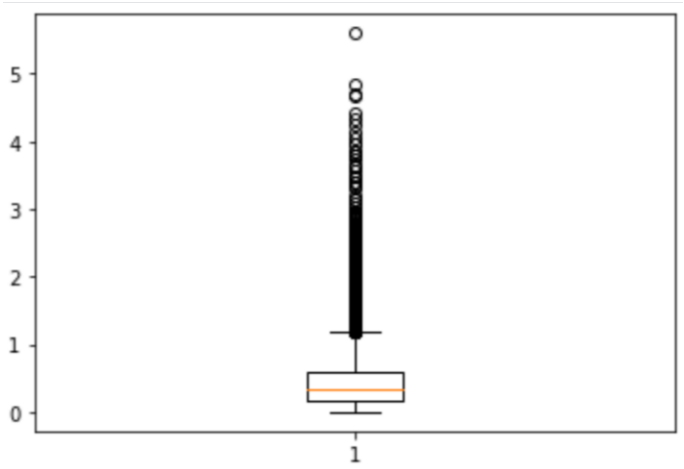


```
providerfund.mean()
Out[50]:
Record No                                          16720.500000
Zip Code                                           49866.283523
Total Resident Weeks                                 307.512231
Total Covid Infections                                 0.071441
Facility Infection Rate Per 1000 Resident Weeks        0.194450
County Infection Rate Per 1000 Resident Weeks          2.195460
Infection Performance Score                            0.428003
Infection Performance Score Capped                     0.579854
Mortality Adjustment                                   0.056176
Final Payment                                      57771.412081
dtype: float64

In [51]:
```

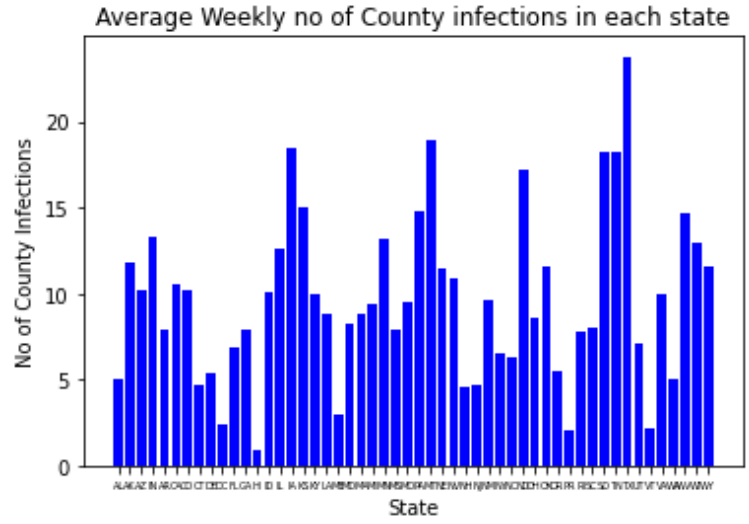
The describe function is implemented to find the percentage values for each of the variables. The values at 25% which is the lowest, 50% which is the median and 75% which is the maximum is obtained.

```
In [51]: providerfund.describe()
Out[51]:
          Record No      Zip Code  ...  Mortality Adjustment  Final Payment
count  33440.000000  33440.000000  ...          33440.000000   33440.000000
mean   16720.500000  49866.283523  ...              0.056176   57771.412081
std     9653.440837  28039.649188  ...              0.092040   49332.373178
min        1.000000    660.000000  ...             -0.200000     101.000000
25%     8360.750000  28304.000000  ...              0.000000   23423.250000
50%    16720.500000  48380.000000  ...              0.000000   45627.500000
75%    25080.250000  74340.000000  ...              0.160000   77981.500000
max    33440.000000  99929.000000  ...              0.200000  718593.000000

[8 rows x 10 columns]

In [52]:
```

Using Python, the boxplot for the Infection Performance Score is plotted and the values above 1.5 are larger. The median value is around 0.5 and the largest value is 5.6.



The County Infection Rate for each state is plotted using Python and it is observed that Texas has the highest County Infection Rate.



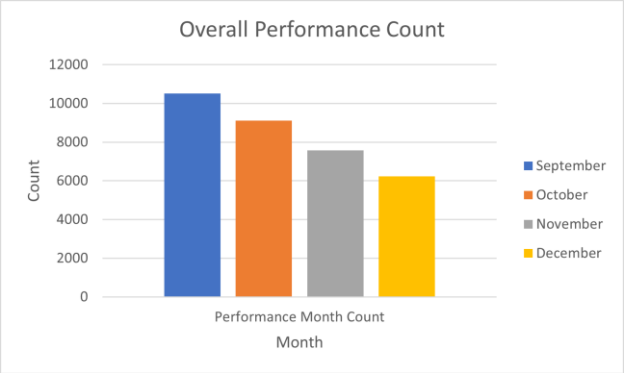
The SQL is used to load the dataset and it was loaded successfully.

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Record No	CCN	Facility Name	Address	City	State	Zip Code
1	15012	HIGHLANDS HEALTH AND REHAB	380 WOODS COVE ROAD	SCOTTSBORO	AL	35768
2	15015	PLANTATION MANOR NURSING HOME	6450 OLD TUSCALOOSA HIGHWAY P O BOX 97	MC CALLA	AL	35111
3	15024	SENIOR REHAB & RECOVERY AT LIMESTONE H...	1600 WEST HOBBS STREET	ATHENS	AL	35611
4	15031	MITCHELL-HOLLINGSWORTH NURSING & REHA...	805 FLAGG CIRCLE	FLORENCE	AL	35631
5	15032	DIVERSICARE OF FOLEY	1701 NORTH ALSTON STREET	FOLEY	AL	36535
6	15035	WEST GATE VILLAGE	100 PINEVIEW AND THIRD P.O. BOX 49	BREWTON	AL	36427
7	15037	CRIMSON HEALTH AND REHAB, LLC	3312 WOODLEY ROAD	MONTGOMERY	AL	36116
8	15044	CAREGIVERS OF PLEASANT GROVE, INC	700 FIRST AVENUE	PLEASANT GROVE	AL	35127
9	15045	EAMC LANIER NURSING HOME	4800 48TH STREET	VALLEY	AL	36854
10	15048	CULLMAN HEALTH CARE CENTER	1607 MAIN AVE NE	CULLMAN	AL	35055
11	15050	OAK TRACE CARE & REHABILITATION CENTER	325 SELMA ROAD	BESSEMER	AL	35020
12	15060	TERRACE OAKS CARE & REHABILITATION CEN...	4201 BESSEMER SUPER HIGHWAY	BESSEMER	AL	35020
13	15063	DIVERSICARE OF BOAZ	600 CORLEY AVENUE	BOAZ	AL	35957
14	15065	PRATTVILLE HEALTH AND REHABILITATION, LLC	601 JASMINE TRAIL	PRATTVILLE	AL	36066

The Performance Month count is done, and September has the highest number of performance period.

Performance_Month	count(*)
September	10273
October	8912
November	7413
December	6100



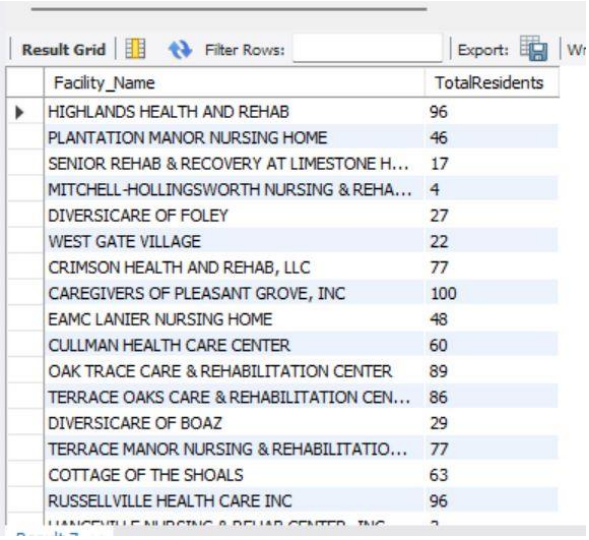
The Mortality rate and Infection Performance Score is compared here, and the negative mortality values indicate the reduced mortality rate. The zero values indicate that there is no mortality in that facility. Positive values denotes the increase in mortality rate.

Result Grid | Filter Rows: | Exp

Mortality_Adjustment	InfectionPerformanceScore
-0.2	171
-0.2	161
-0.2	138
-0.2	235
-0.2	435
-0.2	67
-0.2	22
-0.2	2
-0.2	2
-0.2	4
-0.15	4
-0.09	4
0	317
0	268
0	87
0	471

Result 6

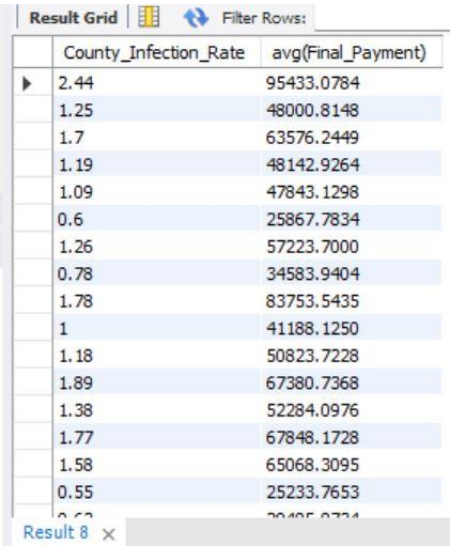
The Facility Name and its corresponding total residents is shown. The Total Resident count is the count of the beds occupied in that facility.



Result Grid | Filter Rows: | Export:

Facility_Name	TotalResidents
HIGHLANDS HEALTH AND REHAB	96
PLANTATION MANOR NURSING HOME	46
SENIOR REHAB & RECOVERY AT LIMESTONE H...	17
MITCHELL-HOLLINGSWORTH NURSING & REHA...	4
DIVERSICARE OF FOLEY	27
WEST GATE VILLAGE	22
CRIMSON HEALTH AND REHAB, LLC	77
CAREGIVERS OF PLEASANT GROVE, INC	100
EAMC LANIER NURSING HOME	48
CULLMAN HEALTH CARE CENTER	60
OAK TRACE CARE & REHABILITATION CENTER	89
TERRACE OAKS CARE & REHABILITATION CEN...	86
DIVERSICARE OF BOAZ	29
TERRACE MANOR NURSING & REHABILITATIO...	77
COTTAGE OF THE SHOALS	63
RUSSELLVILLE HEALTH CARE INC	96

The average payment and the county infection rate are selected to compare the payment increase and decrease based on the infection rate in that county. This is then compared with the facility infection rate and then the funding payment will be decided.



Result Grid | Filter Rows:

County_Infection_Rate	avg(Final_Payment)
2.44	95433.0784
1.25	48000.8148
1.7	63576.2449
1.19	48142.9264
1.09	47843.1298
0.6	25867.7834
1.26	57223.7000
0.78	34583.9404
1.78	83753.5435
1	41188.1250
1.18	50823.7228
1.89	67380.7368
1.38	52284.0976
1.77	67848.1728
1.58	65068.3095
0.55	25233.7653

Result 8

V. Limitations

There are few limitations from this study which I faced during interpretation. First the dataset needs more clarity because the infection rate is not precise and there are lots of duplicate values. Because of this visualization is a bit difficult to perform especially for visualization based on map. Few more analyzation would be more efficient for this dataset like what was the actual

measurement done for the Total Covid Infections. This was collected in the year 2020 and after the pandemic there is no recent study or development in this area. This makes it difficult for the researchers to lay a foundation in this domain.

VI. Conclusion

The overall analysis of this dataset gives an idea about the Funding Acts introduced during the COVID-19 pandemic and the performance of each nursing facility based on the incentive pay is obtained.

a. What are the characteristics of these acts which provided funds?

   The Provider Fund consists of the funding allocation and the HRSA incentive payment to the nursing facilities across the US. The Infection rate and mortality rate are the main decision variables which determine the payment of the incentive to the nursing facility based on the previous month performance.

b. How was the infection rate measured and what are the measures taken to reduce it?

   The infection rate in each facility and that county is measured and both are compared based on which the result is obtained, and the performance of the facility home is also obtained. The infection rate within the facility is lesser when compared which shows that the performance of the nursing home facility is good.

c. What are roles performed by the health care providers and how was the fund utilized to reduce the mortality rate?

   The health care providers are performed based on the incentive payment given to their facility and the reduction in the mortality rate in that facility gives the result of good performance. The highest payment received was California which means that the nursing facility in California has lowest mortality rate and good performance.

d. What was the solution obtained from the Infection Rate and the Mortality rate based on the QIP funding?

   The comparison of the Infection Rate and Mortality Rate was done, the lowest infection rate is 0.0 and mortality rate is -2.0 which means there is no COVID infection, and the mortality rate is reduced. The zero-mortality rate indicates that there are no deaths due to COVID. So, for these facilities it is observed that the funding is higher.

References

[1] data.cdc.gov, ‘ HHS Provider Relief Fund’ 2022[Online] Available: https://data.cdc.gov/Administrative/HHS-Provider-Relief-Fund/kh8y-3es6 [Accessed on 29 - Oct -2022].

[2] pandemicoversight.gov, ‘Coronavirus Relief Fund (CRF)’ 2022[Online] Available: https://www.pandemicoversight.gov/data-interactive-tools/interactive-dashboards/coronavirus-relief-fund [Accessed on 29 - Oct - 2022].

[3] covid-relief-data.ed.gov, ‘Education Stabilization Fund’ 2020[Online] Available: https://covid-relief-data.ed.gov/ [Accessed on 30 - Oct - 2022].

[4] catalog.data.gov/dataset, ‘Provider Relief Fund COVID-19 Nursing Home Quality Incentive Program’ 2021[Online] Available: https://catalog.data.gov/dataset/provider-relief-fund-covid-19-nursing-home-quality-incentive-program-3766a [Accessed on 30 - Oct - 2022].

[5] hrsa.gov, ‘Nursing Home Infection Control Distribution’ 2022[Online] Available: https://www.hrsa.gov/provider-relief/reporting-auditing/nursing-home-infection-control-distribution [Accessed on 5 - Dec – 2022].