Data Analysis: Understanding the Impact of the COVID-19 Pandemic

Abhishek Kalra
ak7468@nyu.edu
NYU Tandon School of Engineering

Rakshita
rr3653@nyu.edu
NYU Tandon School of Engineering

Fola Tubi
ft718@nyu.edu
NYU Tandon School of Engineering

Abstract—

Purpose - COVID-19 has infected and killed over 12 million and 250k people, respectively in the U.S., with New York alone accounting for ~9% of deaths, the highest in the country. Beyond the public health emergency, public officials will need to come to terms with the pandemic's impact on socioeconomic outcomes in NYC. The purpose of our study is to quantify this impact.

Design/methodology/approach – Data was acquired from a variety of sources but primarily from the NYC open data repository and evaluated using various quantitative approaches including deriving measures of excess quantities (p-scores) as well as Pearson correlation and regression analyses

Findings – Our analysis indicates that Covid-19 has resulted in an above average rise in loss of employment and a deterioration in non-COVID related health care. However, we found no evidence to claim that the pandemic has resulted in an above average deterioration in mental health while our analysis on the pandemic's impact on public safety yielded inconclusive results. Despite the observed increase in employment loss, we did not see a corresponding rise in homelessness, our chosen measure of poverty and income security, but, rather, found a negative correlation between the pandemic and poverty.

Limitations – Our analysis period coincided with the government sponsored income support programs enacted to preserve living standards during the pandemic, thereby likely distorting the data. A longer-term study would be needed to better understand how the socioeconomic effects of the pandemic evolve over an extended period of time.

Keywords—

Covid-19, Data Analysis, Pandemic, NYC, Regression, Visualization

1. Introduction and Problem Statement

As of December 6th, 2020 COVID-19 had infected over 66m people, caused over 1.5m deaths, and devastated living standards across the world, according to the COVID Tracking Project. While governments around the world have largely responded by mobilizing resources to tame the immediate public health crisis, the indirect damages in the form of economic scaring, a deterioration in mental hygiene, diminished learning outcomes, and exacerbation of social inequalities, among other legacies, stemming from the pandemic-induced shutdowns deployed by governments will arguably pose no less a challenge to public authorities.

Big cities were particularly vulnerable to a virus that thrives within societies characterized by high population density, exemplified by the course of virus in New York City (NYC), which with ~3% of the United States' population, accounts for ~9% of the deaths officially caused by the disease (Covid Tracking Project). Furthermore, given NYCs largely service driven economy and reliance on crowded public transportation, city-wide mandated shut-downs and social distancing requirements have left the city facing a formidable socioeconomic crisis just as the city's fiscal position - largely underpinned by sales, personal income, and property taxes - comes under severe strain with personal income tax revenue expected to drop by $2 billion this fiscal year, sales tax revenue down by ~11% year-to-date, and a ~4.7% rise in the rate of delinquent property tax payments that were due July 1st, culminating in a $9 billion revenue deficit and a $7 billion reduction in the city's annual budget (Reuters, 2020). Without state and federal funds, itself not a forgone conclusion with New York State facing its own deficit challenges and federal legislators at a stalemate in negotiations over further stimulus measures, the city faces the prospect of spending cuts just when the opposite is needed to help rebuild the cities vibrancy.

Complicating the situation further is the disproportionate impact the pandemic has had on communities of color which are overrepresented in several of the low-wage service sectors that were decimated by the pandemic, including bricks-and-mortar retail, leisure, and hospitality, all of which do not lend themselves to remote working, thereby exposing these communities to a higher risk of infection and/or unemployment.

Given the unprecedented scale of the challenges ahead and the need to allocate scarce fiscal funding efficiently, this study seeks to quantify how COVID-19 has impacted various socioeconomic outcomes in NYC in order to inform a targeted, high-impact fiscal policy moving forward. We will be focusing our study on investigating the hypotheses detailed in the section below.

2. Literature Review and Hypotheses

For the present study, several related works are selected as a basis for further exploration.

2.1. Poverty and Income

Using high frequency data from the Basis Monthly Current Population Survey - validated by comparing historical estimates derived from it to those based on official surveys - income distribution and poverty have been measured nationally over the course of the pandemic (Han, Meyer, & Sullivan, August 2020). Using income as a measure of poverty, the paper finds that, leading up to the pandemic, poverty had been steadily decreasing, falling by 0.9 percentage points between November 2019 and February 2020 but subsequently accelerated to 1.5 percentage points (~14%) between February and June. Within demographic groups, poverty was found to have fallen by 17.1%, 16.1%, and 11.1% for individuals aged 65+, 18-64, and 0 – 17, respectively while declines were observed across all racial and gender groups with those categorized as other (neither white nor black) experiencing a ~25% decline in poverty. The paper attributes the observed decline in poverty primarily to government assistance, including the CARES Act signed into law on March 27th which, through the Economic Impact Payments, Pandemic Unemployment Compensation, and Pandemic Unemployment Assistance programs supported the incomes of millions of American individuals and families.

Income inequality has been proposed as one of the legacies of previous pandemics/epidemics (Furceri, Lougani, Ostry, & Pizzuto, 2020). The paper investigate previous pandemics and major epidemics over the past two decades and their impact on income inequality. Focusing on the SARS (2003), H1N1 (2009), MERS(2012), Ebola(2014), and Zika (2016) pandemics/epidemics and a sample of 175 countries impacted, the authors find that five years after each event, on average, both the market and net Gini coefficients, widely used measures of inequality , were above the pre-shock trends by approximately 0.75% and 1.25%, respectively, which the paper characterizes as quantitatively significant since Gini coefficients typically vary slowly over time. The authors cite these findings to support the premise that COVID-19 could intensify income inequality both within and across countries.

Building on the above findings, we seek to investigate trends in income and poverty over the course of the pandemic in NYC, with particular emphasis on employment, homelessness, and income inequality.

Therefore, we postulate the following:

- H1: COVID-19 has likely resulted in an above average rise in employment loss in NYC
- H2: COVID-19 has likely resulted in an above average rise in homelessness in NYC

2.2. Health Safety

The pandemic has had a profound effect on general health care and has affected the way patients are using health-care facilities (Mercier, Grégoire et al., 2020). Numerous studies made since the onset of pandemic have noted a significant decline in hospital admission for elective and emergency procedures. Sharma et al. (2020) in their study noted a significant decline of 40-60% in acute myocardial infarction admissions in the United States during the onset of the pandemic in March. Mesiner et al. (2020) echoed similar findings in their study to evaluate the effect of a nationwide lockdown in France on admissions to hospitals for acute myocardial infarction.

Building on the above findings, we seek to investigate the same in our study. We undertake an assessment of the healthcare infrastructure availability for varying medical needs including adult, pediatric, ventilators and traumatic injury healthcare under the heightened strain exerted on medical infrastructure due to pandemic. Furthermore, recognizing the debilitating economic effects of COVID shutdowns, the study also assess access to healthcare facilities by analyzing enrollment in healthcare-safety net programs such as Medicaid and the Children's Health Insurance Program (CHIP).

Therefore, we postulate:

- H3: COVID-19 has likely resulted in an above average deterioration in health care security in NYC

2.3. Public Safety

To flatten the infection curve, statewide stay-at-home orders and business closures are being instituted across the United States which seem to have had a mixed impact on public safety. Recent Studies abound on both side of the spectrum wherein Jacoby et al., 2020 reported that weekly calls for service dropped "at least" 12% between February 2 and March 28, 2020 across 30 police agencies. Alternatively, Ashby (2020), in his analysis on over a dozen U.S. cities of various sizes comparing crime rates across six serious criminal offenses, concluded that data supporting reduced crime rates to be inconclusive. Current research investigates the same premise over a longer period.

Therefore, we postulate:

- H4: COVID-19 has likely resulted in an above average deterioration in public safety in NYC

2.4. Mental Health

The COVID-19 healthcare pandemic has taken a severe toll on physical wellbeing. With over 1.5m deaths and millions hospitalized, the physical impact on human life has been unprecedented. To exacerbate the situation, non-pharmaceutical interventions (NPIs), although essential to halt transmission of the virus, have led to physical isolation, closure of schools (with untold effects on the development and wellbeing of children), widespread job losses and reduced public safety. Coagulation of these factors has been detrimental to the mental wellbeing of the population as noted by WHO (2020) in their survey across 130 member counties. Similar findings were echoed in a KFF Tracking Poll conducted in mid-July wherein 53% of adults in the United States reported that their mental health has been negatively impacted due to worry and stress over the coronavirus, up from 32% in March. Owing to the multipronged nature of COVID-19's impact on physical, health, social and economic wellbeing, all of which influence the metal hygiene, our study aims to investigate the ensuing impacts in NYC.

Hence, we postulate the following:

- H5: COVID-19 induced deterioration in health care security has negatively affected mental wellness.
- H6: COVID-19 induced deterioration in economic security has negatively affected mental wellness.
- H7: COVID-19 induced deterioration in public safety has negatively influenced mental wellness.

3. Research Methods, Architecture, and Design

3.1. Data

The following datasets were used to support our analysis:

Table 1: Datasets

Variable	**Indicator**	**Dataset**
Healthcare	Medicaid Enrollees	Citywide HRA- Administered Medicaid Enrollees
	Chip Enrollees	Child Health Plus Program Enrollment: Beginning 2009
	Healthcare Infrastructure Availability	Nursing Home Weekly Bed Census: Last Submission
COVID Cases	COVID Case-Hospitalization-Deaths	COVID-19 Daily Counts of Cases, Hospitalizations, and Deaths
Public Safety	Shooting Incidents	NYPD Shooting Incident Data (Historic) NYPD Shooting Incident Data (Year To Date)
	Arrest Counts	NYPD Arrest Data (Year to Date) NYPD Arrests Data (Historic)

Table 1: Datasets (continued)

Variable	Indicator	Dataset
Mental Health	Mental Health	Anxiety and Depression Household Pulse Survey National Health Interview Survey
Homelessness	Number of adults in homeless shelters	DHS Daily Report
	Number of children in homeless shelters	
	Homeless shelter census counts by borough	Individual Census by Borough, Community District, and Facility Type
Income	Nonfarm employment	New York City Seasonally Adjusted Employment
	Private sector employment	
	Leisure and Hospitality employment	
	Education & Health Services employment	
	Financial services employment	
	Information services employment	
	Manufacturing employment	
	Public Administration employment	
	Trade, transportation, and utilities employment	
	Retail trade employment	
	Wholesale trade employment	
	Profession & Business services employment	
	% of individuals reporting loss of employment income by demographic grouping	United States Census Bureau Household Pulse Surveys
	Advances to NY state’s unemployment fund	Advances to State Unemployment Funds (Social Security Act Title XII)

Methodology

Our study adopted various quantitative approaches to test for the hypothesized relationships between the pandemic and the aforementioned socioeconomic outcomes.

Evaluation Criterion

a. Excess Quantities (P-Scores)

Borrowing from the concept of excess mortality used in epidemiology and public health to approximate the number of deaths from all causes during a crisis in excess of what would have been expected under normal circumstance, our study generalizes this model, applying it, where appropriate, to derive a measure known as the p-score to approximate the pandemics impact on enumerable socioeconomic measures. Using the transformed datasets, our study relied on the Apache Spark SQL engine to, for each feature; compute its average on the date in question for the period defined below.

We define the p-score as follows:

$$\frac{X_{\mathrm{day\ n\ 2020}} - X_{day\ n\ avg}}{X_{day\ n\ avg}} * 100$$

where X is one of our measured quantities and n is a day of the month. We have used 10 years of historical data to compute the average value of X for our analysis relating to employment and 2 years for that relating to homelessness by borough. Our remaining analyses use a 4-year period.

We believe that p-scores offer a wide-ranging measure of COVID's impact, both direct and indirect. However, we also note a limitation with respect to using average values to approximate what an expected value should be under normal circumstances which may not accurately reflect structural trends in the data.

b. Pearson Correlation Method

Within each socioeconomic segment, our study computed coefficient values for the following features:

Table 2: Features used within each socioeconomic segment

Homelessness	CVD-19 cumulative cases, CVD-19 cumulative hospitalizations, CVD-19 cumulative deaths, number of adults in homeless shelters, number of children in homeless shelters
Employment	CVD-19 cumulative cases, CVD-19 cumulative hospitalizations, CVD-19 cumulative deaths, total non-farm payroll
Health Care Security (Infrastructure Availability)	CVD-19 cumulative cases, CVD-19 cumulative hospitalizations, Nursing Home Beds Availability, Ventilator Beds Availability
Health Care Security (Healthcare Safety-Net Insurance Enrollments)	CVD-19 cumulative cases, Medicaid Enrollments, Children's Health Insurance Program (CHIP) enrollments
Public Safety	CVD-19 cumulative cases, arrest counts, shooting incidents

c. Linear Regression

To establish causation, our study employed multi-variate linear regression models with cumulative cases as the independent variable and the features as the dependent variables. Further, we also built a multiple linear regression model to study the impact of the pandemics disruption to economic, social, healthcare and public safety on the mental well-being of NYC's residents.

Dependent Variable	**Independent Variables**
Mental Wellbeing	number of people employed, arrest counts, Medicaid and CHIPs enrollments, Nursing Home Beds Availability, Ventilator Beds Availability

Architecture/design

a. Extraction

Python's data analysis and manipulation library was used to profile the raw datasets in our repository in order to extract metadata including schema semantics, data types, data statistics, and value distributions, thereby facilitating the discovery of the following data quality issues and adopted remedies:

Table 3: Data Quality Issues and their solutions

Data quality Issues	**Solution**
Duplicated rows	Deleted from dataset
Missing values for certain days in the time period of analysis	Replace by appropriate mean value. Seven day running in average in the case of daily data; 3 month running average in the case of monthly data.
Number as Strings	Converted to Float or Integer per requirement
Inconsistent date formats	Converted all dates to the following format: "YYYY-MM-DD"
Inconsistent formatting of column headers	Striped all white space/new line characters and replaced with "_"; created standardized header format with "_" separating each substring in the header
Inconsistent shape of datasets	Performed transformation on incompatible datasets

b. Integration / Aggregation

Our study primarily employed the Apache Spark SQL engine on NYU's 48-node DUMBO cluster to perform schema matching, integration, and aggregation, thereby facilitating the following data transformations:

- Joining the COVID-19 datasets with each of the datasets used for our analysis, resulting in an intermediate dataset with the following schema:

  (date, cases, hospitalization, deaths, deaths_probable, $Feature_1$*,* $Feature_2$*, …,* $Feature_n$*)*

  where $feature_i$ is one of the quantities of interest mentioned in Table 2

- Creation of additional columns to compute the cumulative counts of cases, hospitalizations, and deaths as well as a running average of each feature over an appropriate period, as defined below, used to replace null values, resulting in the following schema:

  (date, cases, cumulative case, hospitalization, cumulative hospitalization, deaths, deaths_probable, cumulative deaths, $Feature_1$, $Feature_{1avg}$,

  $Feature_2$, $Feature_{2avg}$, … , $Feature_n$, $Feature_{navg}$)

  where $Feature_{i\ avg}$ is a running average of one our measured quantities.

c. Visualization

The results of our data analysis were visualized using Tableau software and Python's Matplotlib modules.

4. Analysis and Findings

4.1. Impact of Covid-19 on poverty and income insecurity

In figure 1, we report monthly trends in employment in key economic sectors against cumulative COVID-19 cases. Then in figure 2, we show the correlation coefficients between our measured variables.

The results indicate a general deterioration in labor market conditions in NYC with total non-farm employment, during the peak of the initial wave of the pandemic (March – May), as much as 11% below what would have been expected under normal circumstances, down from ~12% above normal before the pandemic. Sector level analysis indicates a material contraction within the leisure and hospitality sector which was as much as 63% below what would have been expected under normal circumstances down from 13% above normal before the pandemic. Manufacturing, retail, and wholesale trade also saw significant contractions, down by 36%, 27%, and 18%, respectively relative to what would have normally been expected during the peak of the pandemic. Despite the sector wide contraction in the

labor market, our data indicates that some sectors saw relative stability versus their 10-year averages. At the peak of the initial wave of the pandemic, relative to expected values under normal circumstances: Education and health services was 5% above (down from ~21% above pre-pandemic); professional and business services was 1% above (down from 18% above pre-pandemic); information was 6% above (down from ~12% above pre-pandemic); financial sector was 1% above (down from 6% above pre-pandemic); and public administration was relatively flat at 1% above, although we observed a sharp contraction in the data in July (12% below normal) before a subsequent bounce back up to normal trend.
These results were largely expected, particularly with respect to the sectors that rely on social interaction - namely leisure, hospitality, and retail – whose operations were severely curtailed by the government mandated shutdown measures. Also expected was the relative outperformance of the industries that have quickly adapted to the new remote working culture necessitated by the pandemic, namely financial activities, education, professional, and business services. Given the essential nature of health care, we also expected to observe some outperformance there as well. The relationship between the measured variables are summarized by the correlation matrix in figure 2 which show negative correlations between cases and the service oriented sectors and more neutral correlations between cases and the sectors able to sustain remote working.

In figures 10 - 12, we report the share of adults who have reported experiencing a loss of employment income since March 13, 2020 by age, gender, and race/ethnicity, respectively. The data indicate that those in the 18 – 24-year-old age bracket have fared the worst with 63% of respondents to the Census Bureau's household pulse survey reporting that they have experienced a loss of employment income since the onset of the pandemic. Those characterized as 65 and above fared the best, with an average of 35% of respondents within that age category reporting a loss of employment income during the year. Averages for the remaining categories were clustered around ~59%. Looking at the data by gender, we observe similar percentage of individuals of both sexes reporting they have lost employment income, with an average of 55% of male and female respondents reporting in the affirmative.

By race, the data suggests that those in the Hispanic and Latino demographic have fared the worst, with an average of 66% of respondents saying they had lost employment income over the course of the pandemic. Second worst hit were those in the African American demographic with an average of 58% reporting a loss of income, followed by those in the Asian demographic with an average of 55% responding in the affirmative. The White demographic group fared the best, with an average of 49% reporting a loss of income. This data seems to confirm the argument that the pandemic has disproportionately impaired the economic prospects of the young as well as communities of color, thereby further exacerbating intergenerational and racial inequalities, a finding that concurs with (Furceri, Loungani, Ostry, & Pizzuto, 2020).

Given the significant deterioration in labor market conditions during the year, we were intuitively expecting to see a rise in indicators of poverty, particularly homelessness. In figure 4, we report the trend in excess homelessness over the course of the pandemic to November 2020. While the number of adults in homeless shelters across the city was, on average, ~3.9% above expected value under normal conditions, that of children was ~15.2% below expected value, yielding a consolidated average of ~3.5% below expected value. Furthermore, although the p-score for adults was above the origin for most of the year, the data indicate a trend of improvement throughout the year with a p-score of ~(1.15)% at the end of November, down from 8.3% in February. Figure 4 shows the trend in cases by borough while figure 5 shows excess homelessness also by borough. The data indicates excess levels of homelessness in Queens, as much as 13% during the year relative to expected value, which was consistent with the course of the pandemic in that borough, as shown in Figure 4. However, despite the boroughs of Brooklyn and the Bronx recording the 2nd and 3rd greatest number of cases, they each experienced relative outperformance with the number of individuals housed in homeless shelters in each borough down as much as 20% and 18%, respectively relative to expected values. Staten Island with the fewest cases also outperformed, with the number of individuals in homeless shelters down 29% relative to expected value. However, we were surprised to discover excess homelessness in Manhattan, up as much 12% relative to expected value, considering it was the borough with the 2nd fewest number of cases.

The trends observed for homelessness might seem counter intuitive at first glance. However, in figure 9 we report the trend of the Federal Government's outstanding advance balance to New York State's unemployment fund, pursuant to the panoply of income support programs ushered into law by Congress during the pandemic. We observed a steep rise in borrowing from a zero balance at the beginning of the pandemic to over $8 billion in a matter of 4 months. Such a substantial increase in fiscal support is likely to have shielded the incomes of those who lost their jobs or faced reduced hours as a result of the pandemic, thereby preserving living standards for the time being, which would largely explain why we saw a gradual improvement in our homelessness measure, contrary to our initial hypothesis.

4.2. Impact of Covid-19 on healthcare security

Healthcare Infrastructure Availability

A time series graph of healthcare infrastructure availability over the period extending from September '19 to October '20 was plotted to understand the impact of COVID-19 on healthcare provision (Figure 13). Infrastructure availability was assessed by analyzing availability of nursing home infrastructure (being repurposed during pandemic) to treat COVID patients. The availability has been measured using the p-scores adjusted for number of cases to give a percentage of increase/decrease in the availability. As seen in the graph, there was an overall increase in the medical bed availability as nursing homes were incentivized by federal and state agencies to admit COVID-19 patients (NPR,2020). However, the repurposing was not found to be sufficient to keep up with the increased numbers of hospitalizations as cases surged. The number of available ventilator and general care beds plummeted significantly as COVID-19 hospitalizations increased and remained at a low level due to the lingering impacts of the first wave.

To quantify the visualized behavior, we calculated the correlation between cases and healthcare infrastructure availability and found the relationship to be negatively correlated. Furthermore, to establish a possible causation between COVID cases and healthcare infrastructure unavailability, we ran a multivariate linear regression mode. Both, pediatric bed and nursing home bed availability, were found to have a statistically significant inverse relationship with COVID cases (Figure 14).

Healthcare Insurance Enrollments

A timeseries graph of the enrollment in health safety-net insurance schemes, namely Medicaid & CHIPS over the period extending from September ’19 to October ’20 is show in Figure 15. Access to healthcare services, a key indicator of Healthcare safety, was assessed by analyzing the enrollments in Medicaid and CHIPS. Enrollments have been measured using the p-scores to give a percentage of increase/decrease in the enrollments for the corresponding month. As seen in the graph, there was an overall increase in Medicaid enrollments as shutdowns led to job losses. CHIP enrollments also went down as schools were shut down leading to increased dropouts of low-income students who would have otherwise qualified. Increased enrollment in the safety-net programs is contentious as it is often associated with substandard levels of care compared to pricier employment based and private insurance schemes.

To quantify the visualized behavior, we computed the correlation between COVID cases and healthcare insurance enrollments and found the relationship to be negatively correlated with Medicaid enrollments (Figure 16). After lagging initially during March-April due to the requirement to demonstrate below-threshold income for at least a month, Medicaid enrollments picked up gradually thereafter. During the deconfinement period, as employees began returning to work, there was a corresponding fall in Medicaid enrollments in October ’20. Conversely, CHIP enrollments was found to have a positive correlation that can be attributed to the underlying data timeframe. CHIP enrollments were stable prior to pandemic and began declining in May along with COVID cases, as low income students were not able to keep up with remote learning requirements. Furthermore, to establish a possible causation we assessed the impact of COVID cases on safety-net insurance enrollments by using multi-variate regression. We found Medicaid and CHIPs enrollment to have a statistically significant inverse and direct relationship with COVID case, respectively.

4.3. Impact of Covid-19 on public safety

A Timeseries graph of arrests and shooting incidents over the period extending from September'19 to October'20 was plotted to understand the impact of COVID-19 on public safety (Figure 17). Public safety was assessed by analyzing both the count of arrests made and shooting incidents. Both metrics have been measured using p-scores to give a percentage of increase/decrease in the occurrence. As seen in the graph, there was an overall increase in the shooting incidents during the period May-August and have been gradually coming down since July. The period corresponds with a high rate of job losses previously mentioned. However, the same period also saw Black Live Matter (BLM) protests in wake of racial injustice leading to many violent protests. Arrests counts showed a fluctuating pattern through this period, with reduced measures of organized crime but an increase in arrests related to the protests and property crime.

To quantify the visualized behavior, we computed the correlation between COVID cases and measures of public safety and found the relationship to be negatively correlated with arrest counts. Furthermore, to establish a possible causation we assessed the impact of COVID cases on both arrest and shooting incidents and found only arrest counts to have a statistically significant inverse relationship with COVID cases (Figure 18). Our observed mixed results echo similar findings made by Ashby (2020) in his analysis and require future investigation with a broader indictor set to elucidate the impact.

Table 4: Hypotheses Testing

Factors	Hypothesis	coef	std err	t	P>|t|	[0.025	0.975]	Supported?
Healthcare Security	Cases->Total Medicaid Enrollees	-0.8364	0.298	-2.802	0.011	-1.457	-0.216	Yes
	Cases->CHIPs Enrollees	0.1662	0.046	3.642	0.002	0.071	0.261	Yes
	Cases->Ventilator Bed Availability	-0.0000026	0.0000009	-2.9990000	0.0070000	-0.0000044	-0.0000008	Yes
	Cases->Nursing Home Bed Availability	-0.0001	0.0000465	-3.022	0.006	0	-0.0000438	Yes
Income Security/Poverty	Cases->Employed	-0.0063	0.003	-2.513	0.02	-0.012	-0.001	Yes
	Cases->People in Homeless Shelters	-0.0128	0.019	-0.683	0.502	-0.052	0.026	No
Public Safety	Cases->Shooting Incidents	-0.0001	0.001	-0.228	0.822	-0.001	0.001	No
	Case->Arrest Count	-0.0562	0.024	-2.387	0.026	-0.105	-0.007	Yes

4.4. Impact of Covid-19's effects on mental health affected by of public safety, healthcare security, poverty, and income insecurity

To understand the impact of COVID and the ensuing non-pharmaceutical interventions on the mental wellbeing of the populace, we modeled mental health as a function of economic, healthcare and public/physical wellbeing. As shown in the previous sections, COVID has had a significant impact on all of these aspects. We therefore ran a linear regression model to test the proposed causation. Prior to regression, we checked for any possible multi-co-linearity among the indicators by calculating variance inflation factors (VIFs). Nursing home bed availability and Ventilator bed availability were found to have higher values of VIF due to colocation with the same dataset. Therefore, ventilator beds availability was removed from the model.

The model was subsequently run with 5 independent variables. However, all the relationships were found to be statistically insignificant with respect to mental hygiene. We encourage caution in interpreting these results as mental health is known to be a taboo topic, despite recent progress in promoting awareness. There is an inherent tendency of participants to misrepresent their mental health (Panchal et al., 2020) and large sections of society still maintain a stigma on mental disorders. As a future work stream, an analysis of the "NYC Cause of Death" dataset should be undertaken to understand the trend of suicide cases during and post COVID. As of the time of this study, the data was not available for the time period of analysis.

Table 5: Mental Health Regression Results

OLS Regression Results

Dep. Variable:	Mental_Anxiety_Score	R-squared:	0.445
Model:	OLS	Adj. R-squared:	0.237
Method:	Least Squares	F-statistic:	2.139
Date:	Tue, 08 Dec 2020	Prob (F-statistic):	0.105
Time:	19:38:51	Log-Likelihood:	7.2343
No. Observations:	23	AIC:	-0.4687
Df Residuals:	16	BIC:	7.480
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	1.1708	0.405	2.893	0.011	0.313	2.029
nursing_home_beds_available_prop_pandemic	0.7366	0.760	0.969	0.347	-0.874	2.348
Ventilator_Beds_Available_prop_pandemic	-0.9066	0.845	-1.073	0.299	-2.697	0.884
arrest_count	-0.1307	0.289	-0.452	0.657	-0.743	0.482
Employed	-0.1562	0.272	-0.574	0.574	-0.733	0.420
CHIP_Enrolless	-0.1742	0.401	-0.434	0.670	-1.024	0.676
Total_Medicaid_Enrollees	-0.3298	0.422	-0.782	0.446	-1.224	0.564

5. Limitations and Future Directions

The research study makes an effort to analyze a diverse set of socio-economic indicators maintained publically. A few of the datasets lacked more recent data points due to either being reported at the end of the year (e.g., mental health) or being delayed because of the pandemic (e.g., cause of death, school enrollments). Further, given the nature of some our datasets featuring self-reported information, it is possible that our analysis was unable to capture the real impact of the pandemic, thereby distorting our findings. It is also possible that the number of arrests was underreported as increased social isolation has likely exacerbated the number of unreported cases of crimes, such as domestic violence, the majority of which are typically not reported. Therefore, the inclusion of a diverse set of parameters on public safety, may help capture the real trend. Future research can broaden the mental health indicator sets to include cause of death ("suicide") for a better representation.

Finally, our analysis period coincided with the government sponsored income support programs enacted to preserve living standards during the pandemic, thereby likely distorting the data. A longer-term study would be needed to better understand how the socioeconomic effects of the pandemic evolve over an extended period of time.

References

Covid Tracking Project. (2020, December 6). Retrieved from Covid Tracking Project: https://covidtracking.com

Department of Health and Mental Hygiene. (2020, November 30). *COVID-19 Daily Counts of Cases, Hospitalizations, and Deaths*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/Health/COVID-19-Daily-Counts-of-Cases-Hospitalizations-an/rc75-m7u3

Department of Homeless Services (DHS). (2020, November 30). *DHS Daily Report*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/Social-Services/DHS-Daily-Report/k46n-sa2m

Furceri, D., Loungani, P., Ostry, J. D., & Pizzuto, P. (2020, May 1). Will Covid-19 affect inequality? Evidence from past pandemics. *Covid Economics, Vetted and Real-Time Papers*, pp. 138-157.

Han, J., Meyer, B. D., & Sullivan, J. X. (August 2020). Income and Poverty in the COVID-19 Pandemic. *NBER Working Paper No. 27729*, JEL No. H53,I32,J65.

Human Resource Administration (HRA). (2020, November 27). *Citywide HRA- Administered Medicaid Enrollees*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/Social-Services/Citywide-HRA-Administered-Medicaid-Enrollees/33db-aeds

Human Resources Administration (HRA). (2020, October 19). *Emergency Food Assistance Program (Quarterly Report)*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/Social-Services/Emergency-Food-Assistance-Program-Quarterly-Report/mpqk-skis

Mayor's Office of Management & Budgets (OMB). (2020, November 20). *New York City Seasonally Adjusted Employment*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/City-Government/New-York-City-Seasonally-Adjusted-Employment/5hjv-bjbv

NYC Office of the Comptroller. (2020). Retrieved from New York City Comptroller: https://comptroller.nyc.gov/

Police Department (NYPD). (2020, November 5). *NYPD Arrest Data*. Retrieved from NYC Open Data: https://data.cityofnewyork.us/Public-Safety/NYPD-Arrest-Data-Year-to-Date-/uip8-fykc

Reuters. (2020, September 23). *New York City mayor announced more furloughs to counter budget shortfall*. Retrieved from Reuters: https://www.reuters.com/article/us-new-york-budget/new-york-city-mayor-announces-more-furloughs-to-counter-budget-shortfall-idUSKCN26E2JM

Sharma M, Lioutas V, Madsen T, et al. Decline in stroke alerts and hospitalisations during the COVID-19 pandemic Stroke and Vascular Neurology 2020;svn-2020-000441. doi: 10.1136/svn-2020-000441

Hospital admissions for acute myocardial infarction before and after lockdown according to regional prevalence of COVID-19 and patient profile in France: a registry study

Mesnier, Jules et al. The Lancet Public Health, Volume 5, Issue 10, e536 - e542

Understanding the effects of COVID-19 on health care and systems *Mercier, Grégoire et al. The Lancet Public Health, Volume 5, Issue 10, e524*

Ashby MPJ. Initial evidence on the relationship between the coronavirus pandemic and crime in the United States. Crime Science. 2020;9(6):1–16.

Jacoby, K., Stucka, M., & Phillips, K. (2020, April 16). Crime rates plummet amid the coronavirus pandemic, but not everyone is safer in their home. Retrieved May 28, 2020, from https://www.usatoday.com/story/news/investigations/2020/04/04/coronavirus-crime-rates-drop-and-domestic-violence-spikes/2939120001/

The intersection of COVID-19 and mental health. The Lancet Infectious Diseases, Volume 20, Issue 11, 1217 WHO 2020. The impact of COVID-19 on mental, neurological and substance use services. Accessed on https://www.who.int/publications/i/item/978924012455

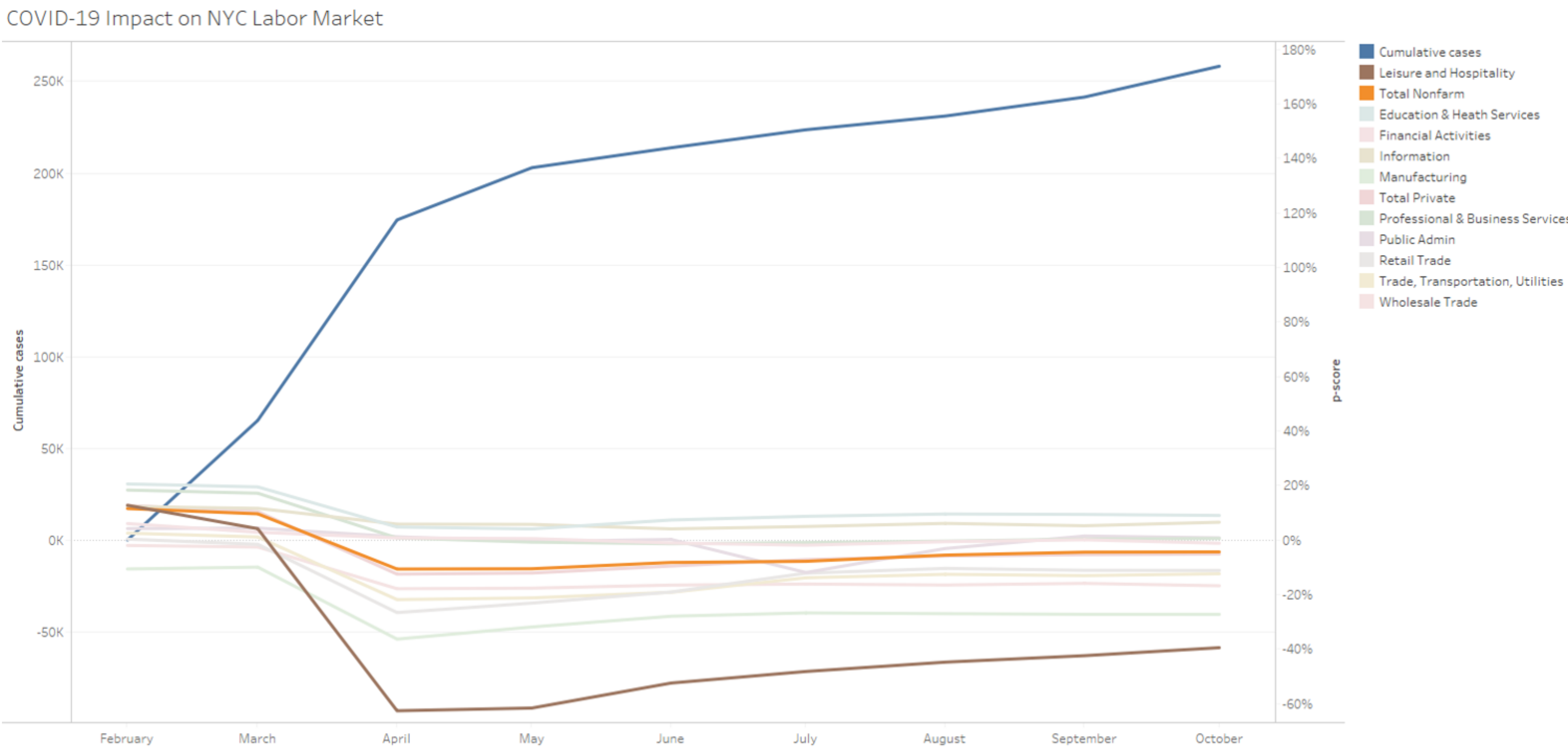
Kaiser Family Foundation, Panchal, N., et al. (2020). The Implications of COVID-19 for Mental Health and Substance Use.

Nursing Home Residents Moved Out To Make Way For COVID-19 Patients Accessed on 12/08/20 using https://www.npr.org/sections/coronavirus-live-updates/2020/08/04/897239846/nursing-home-residents-moved-out-to-make-way-for-covid-19-patients

Blumenthal D, Fowler EJ, Abrams M, Collins SR. Covid-19 - Implications for the Health Care System. N Engl J Med. 2020 Oct 8;383(15):1483-1488. doi: 10.1056/NEJMsb2021088. Epub 2020 Jul 22. Erratum in: N Engl J Med. 2020 Jul 23;: PMID: 32706956.

Appendix

Figure 1: Monthly Trends in Employment



Excess Employment, as measured by p-scores, for select industries.
Source: Mayor's Office of Management & Budgets (OMB)

Figure 2 – COVID and Employment Census Correlation Matrix

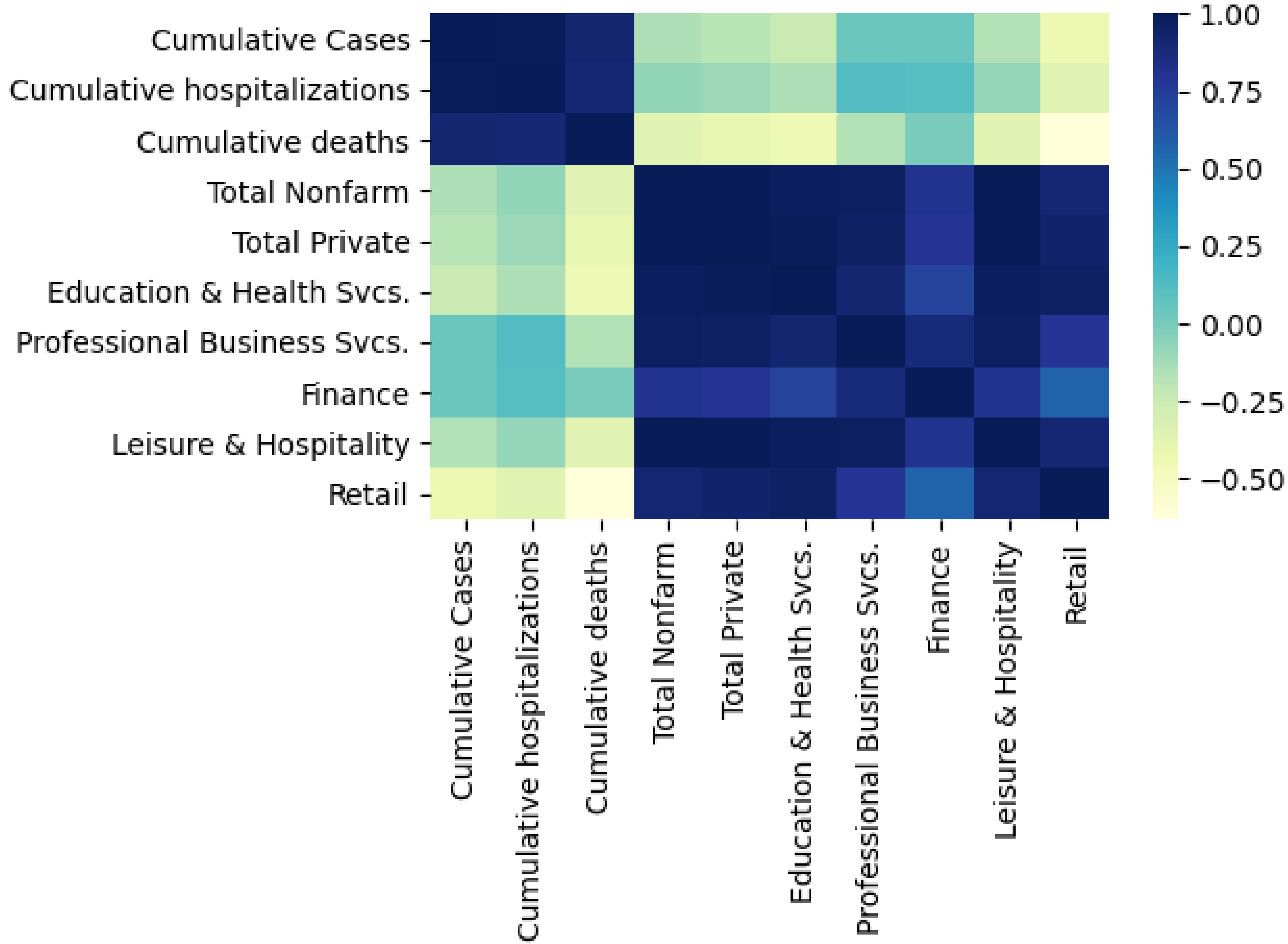
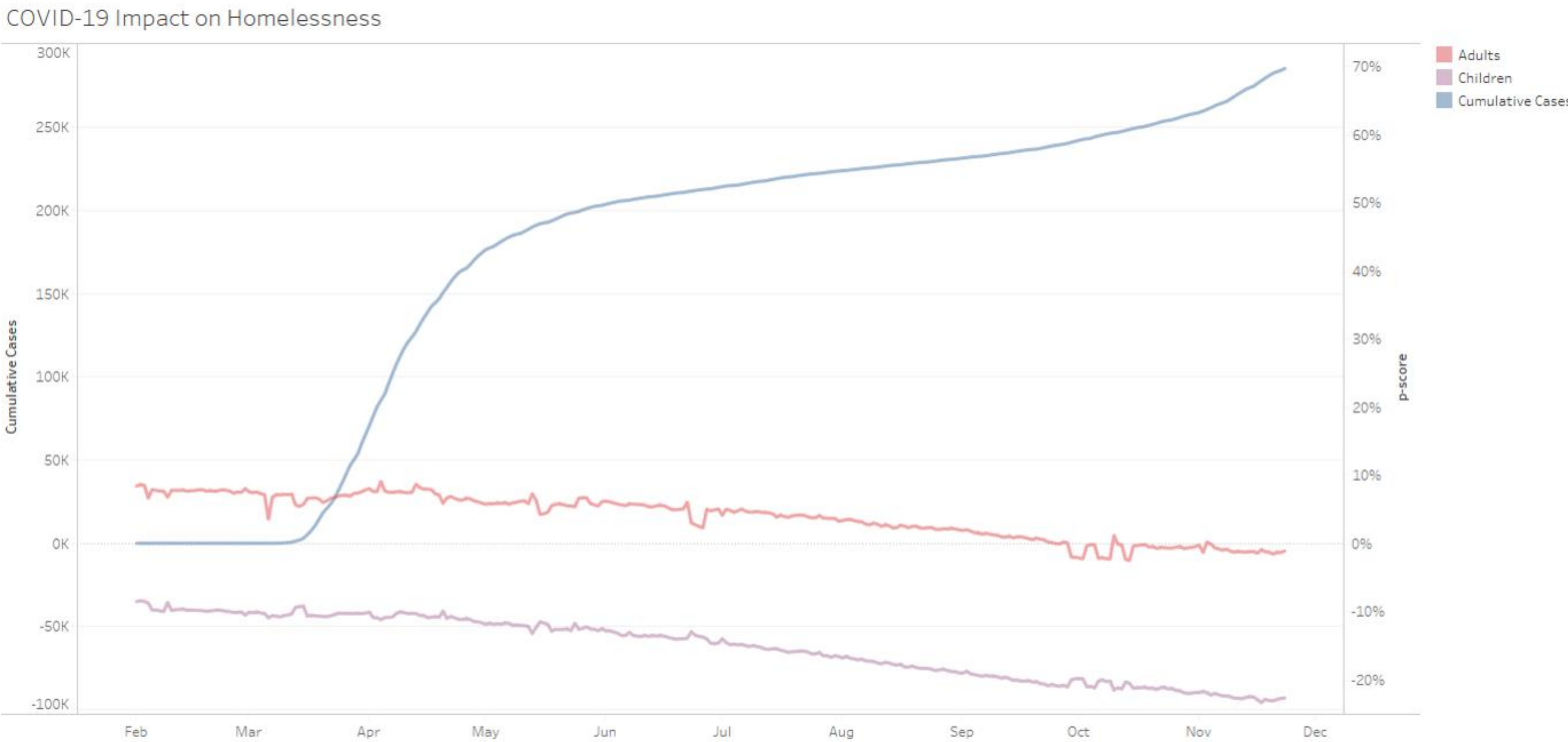
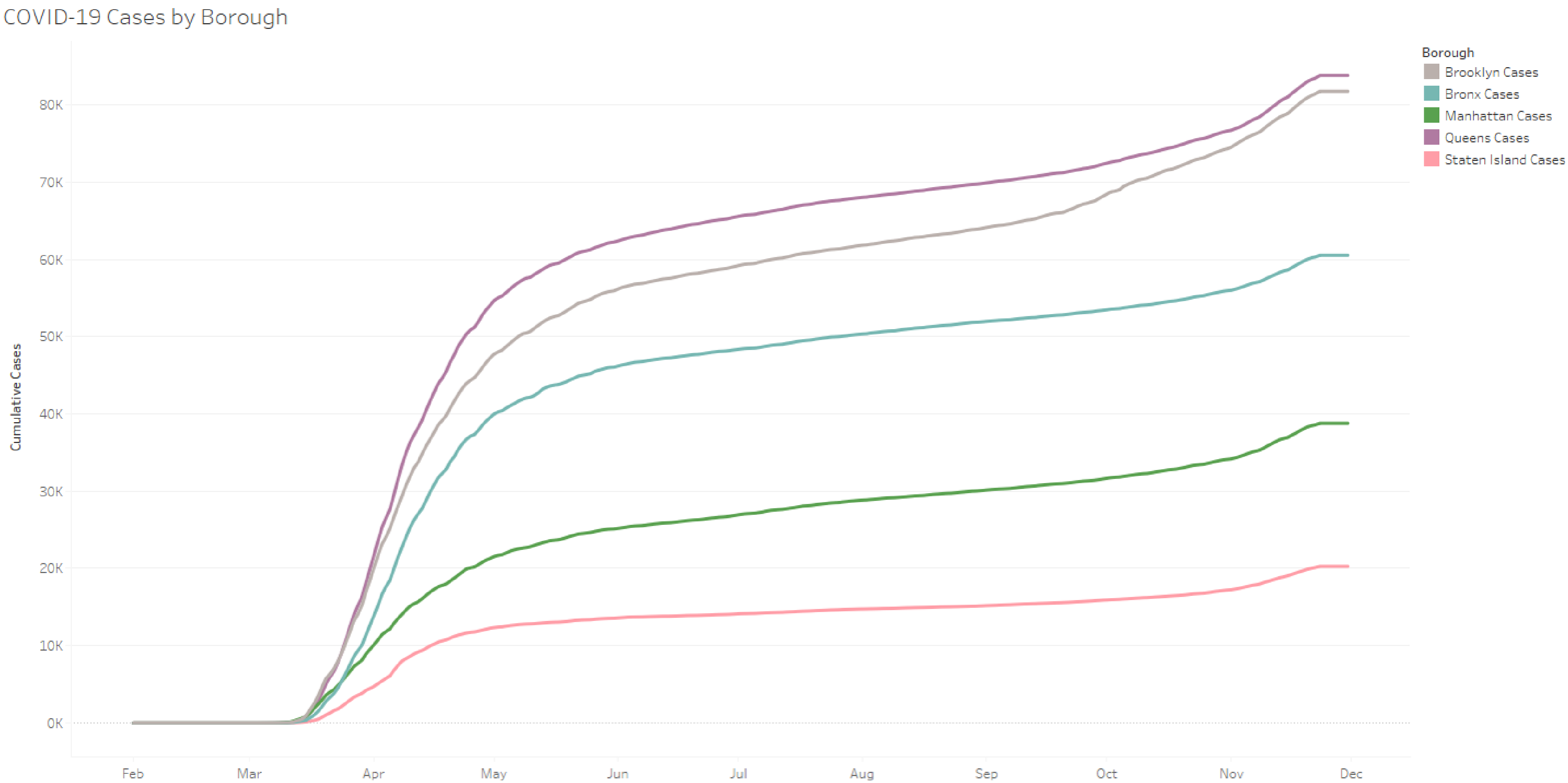


Figure 3: Covid-19’s Impact on Homelessness



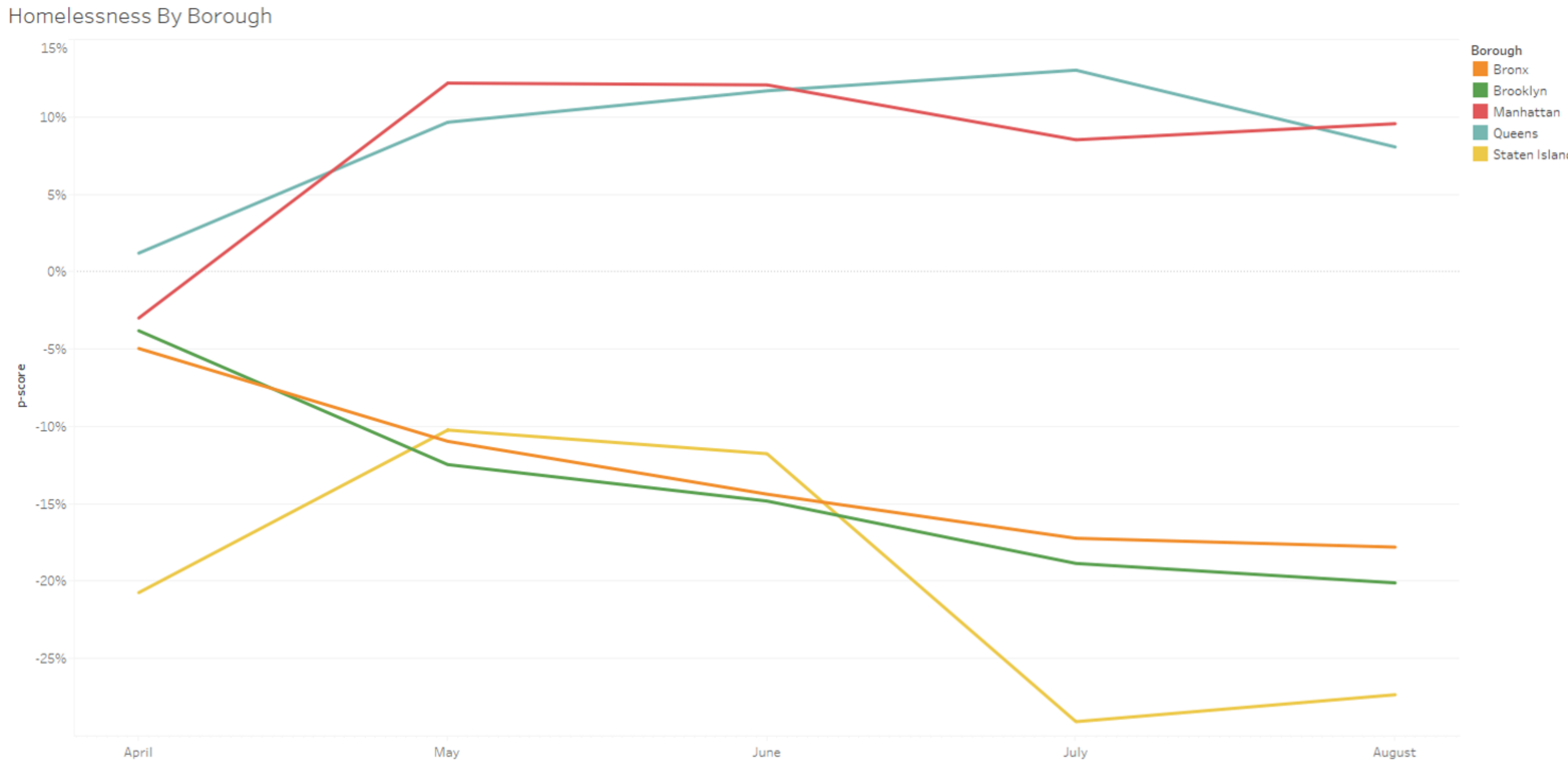
Trend of cumulative COVID-19 cases and excess homelessnes.
Source: NYC Department of Homeless Services

Figure 4 Covid-19 Cases by Borough



The trends of COVID-19 cases in Brooklyn, Bronx, Manhattan, Queens, and Staten Island. Color shows details for each borough.
Source: NYC Department of Health and Mental Hygiene

Figure 5: Homelessness by Borough



The trends of excess homelessness for Brooklyn, Bronx, Manhattan, Queens and Staten Island for Month. Color shows details about Brooklyn, Bronx, Manhattan, Queens and Staten Island.
Source: NYC Department of Homeless Services

Figure 6 – Correlation matrix for COVID and DHS census

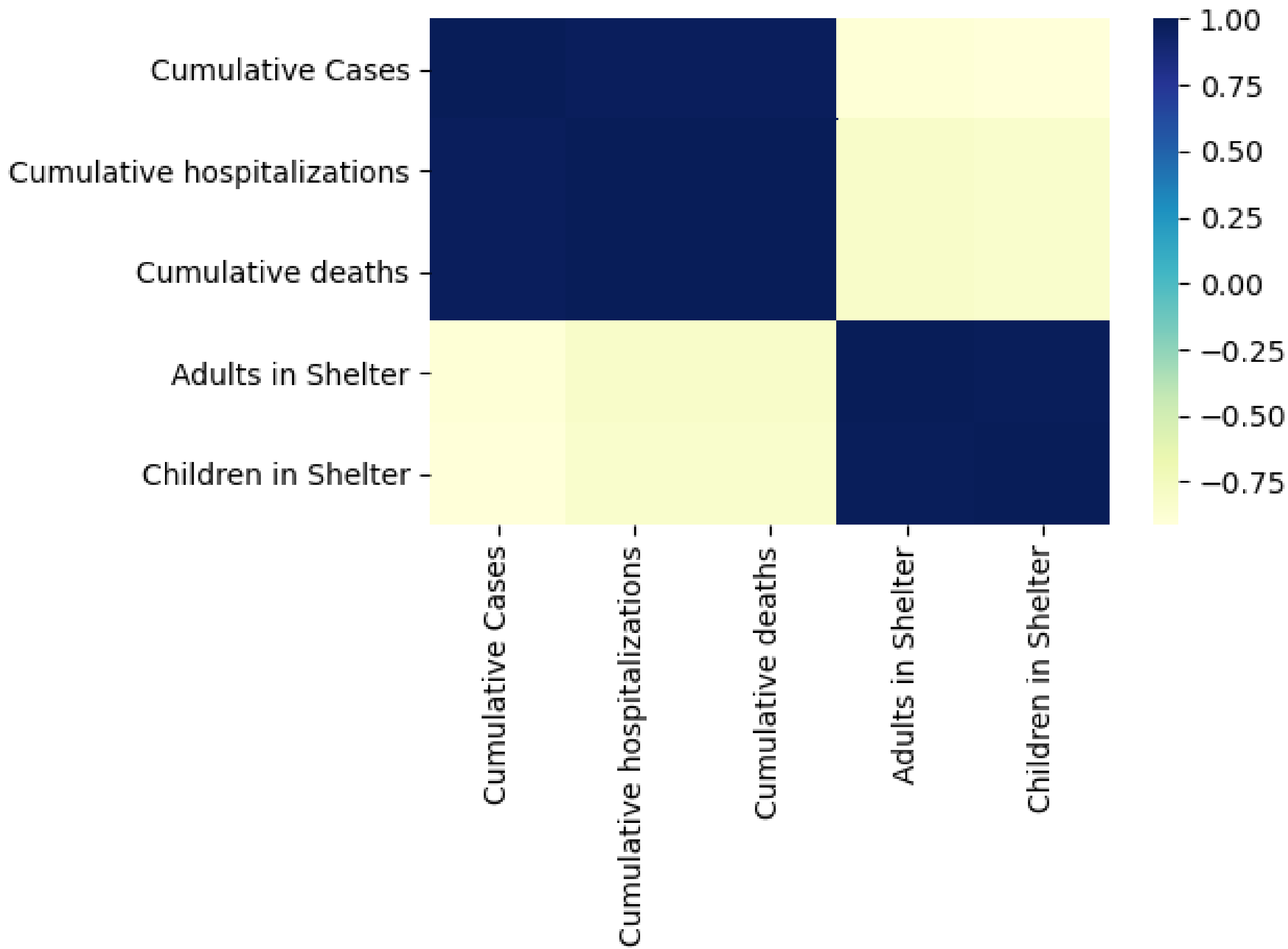


Figure 7 – Adults in Homeless Shelters Regression Plot

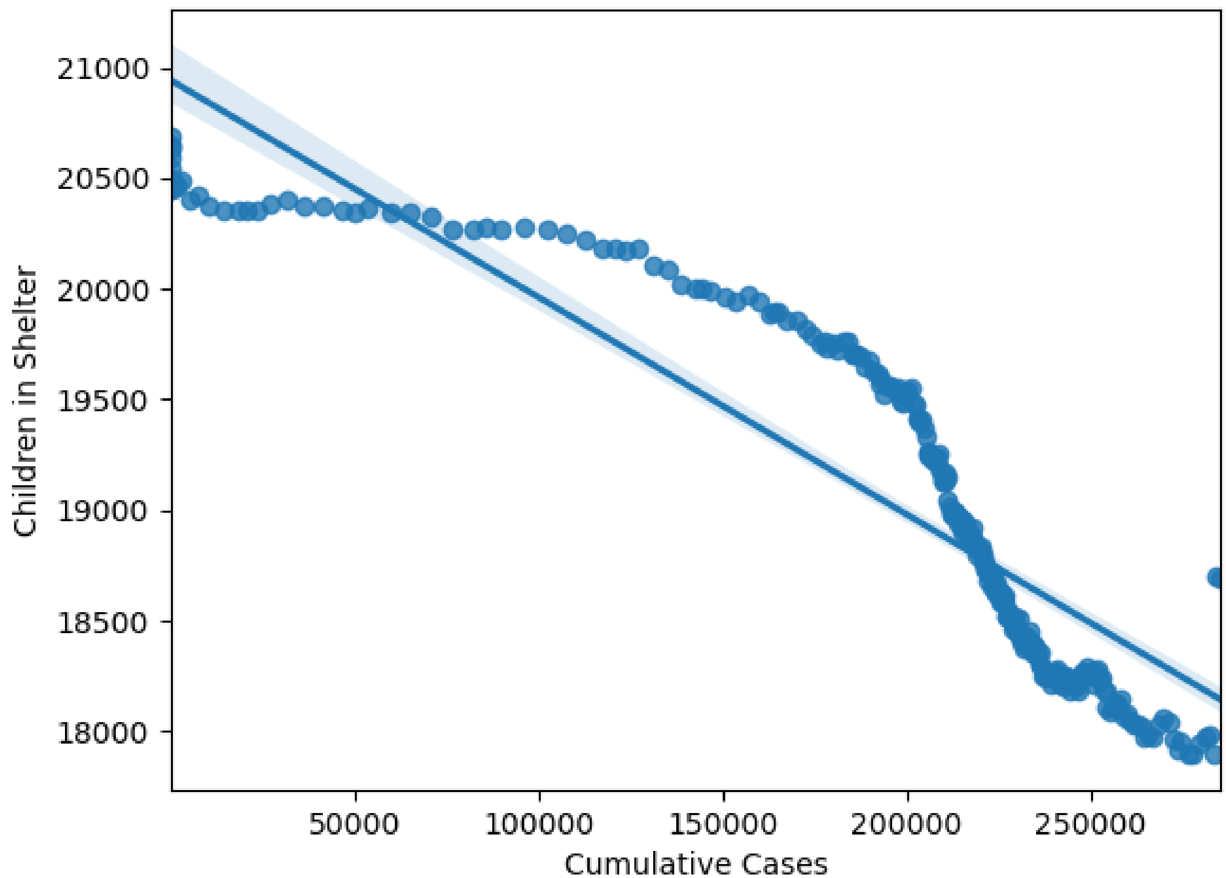


Figure 8 – Children in Homeless Shelters Regression Plot

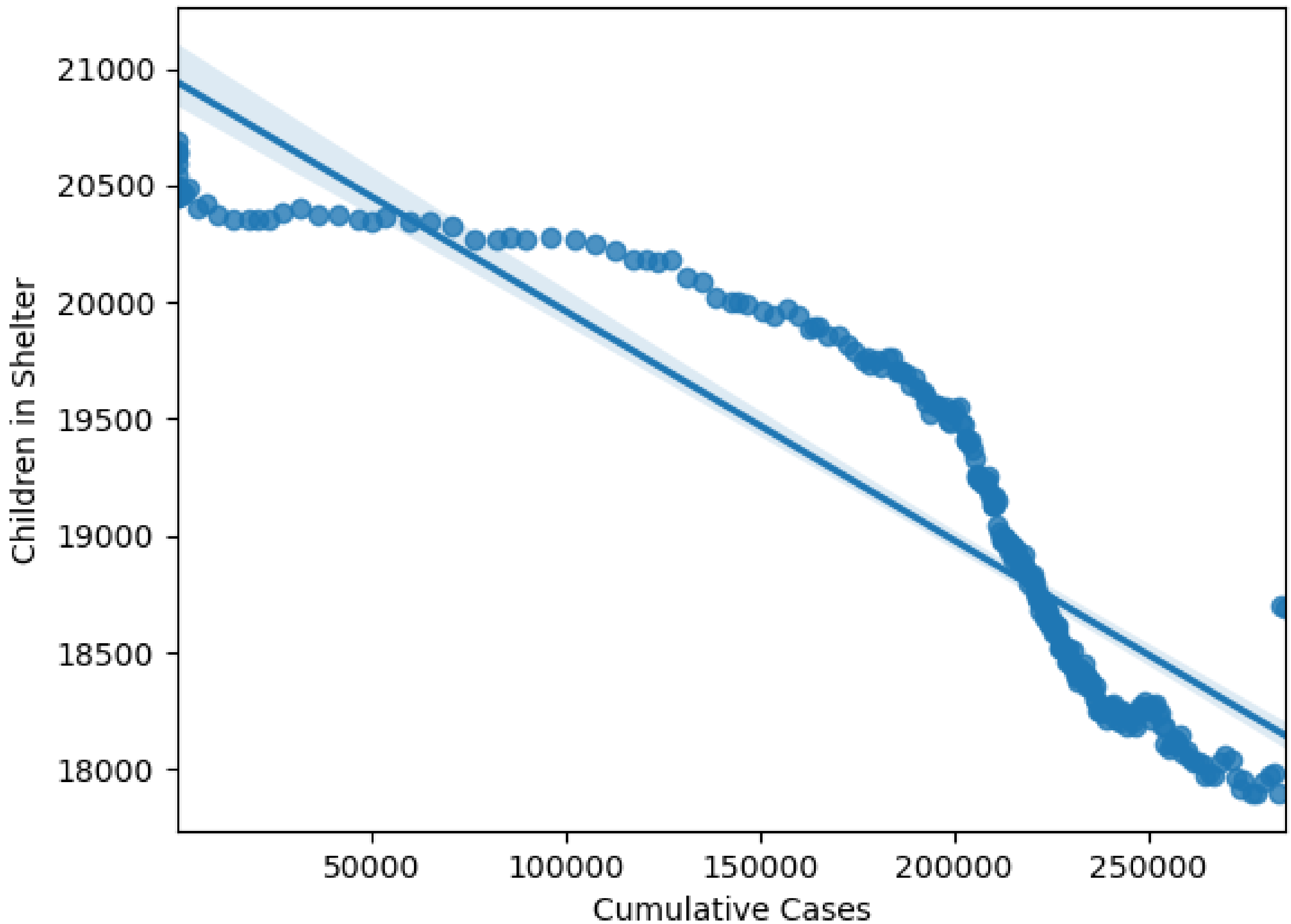


Figure 9: Advances in New York state Unemployment Fund

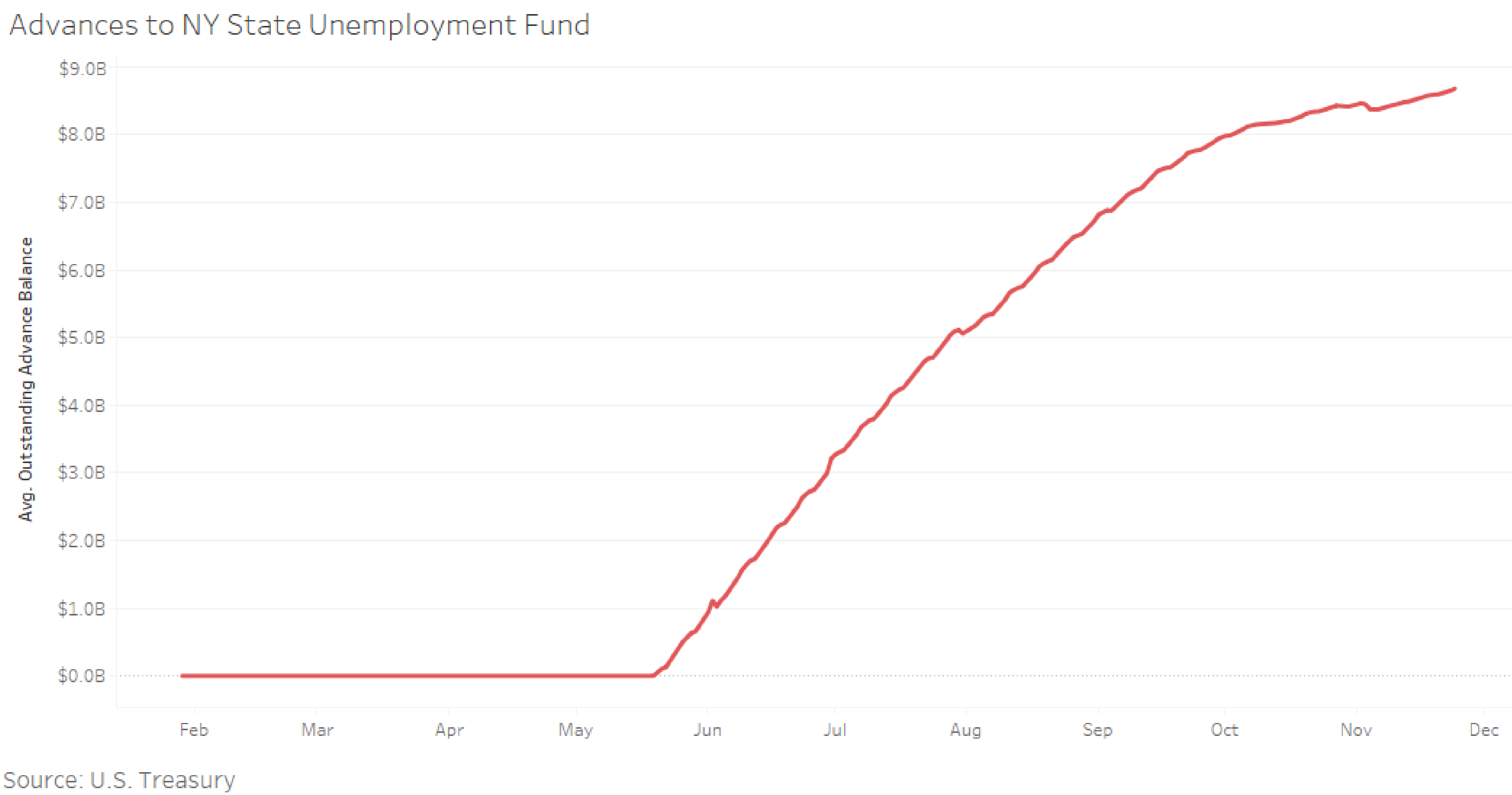
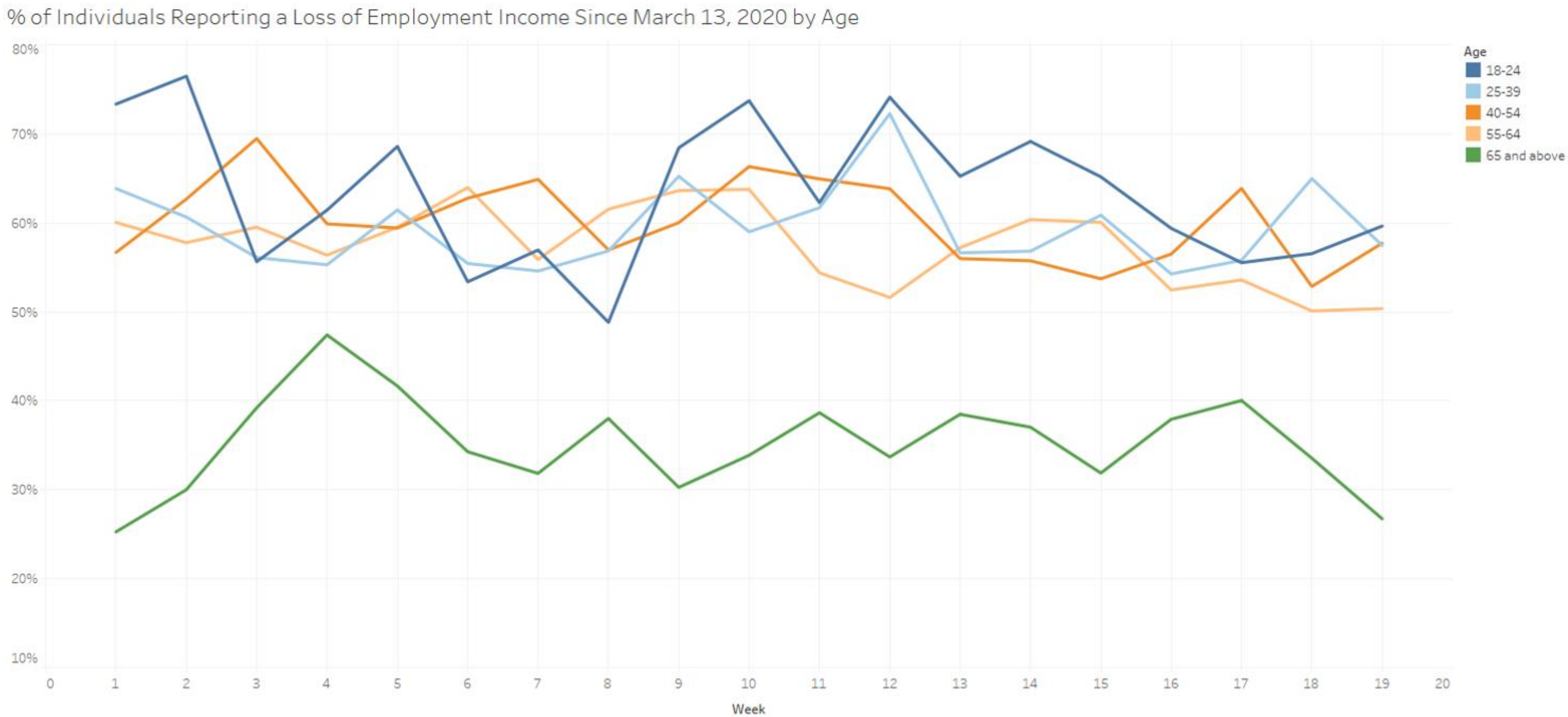


Figure 10: Loss of Employment Income (by age)



Source: United States Census Bureau Household Pulse Survey.
Week 1: Apr 23 - May 5. Week 2: May 7 - May 12. Week 3: May 14 - May 19. Week 4: May 21 - 26. Week 5: May 28 - June 2. Week 6: Jun 4 - Jun 9.
Week 7: Jun 11 - Jun 16. Week 8: Jun 18 - June 23. Week 9: Jun 25 - Jun30. Week 10: Jul 2 - Jul 7. Week 11: Jul 9 - Jul 14. Week 12: Jul 16 - Jul 21.
Week 13: Aug 19 - Aug 31. Week 14: Sep 2 - Sep 14. Week 15: Sep 16 - Sep 28. Week 16: Sep 30 - Oct 12. Week 17: Oct 14 - Oct 26.
Week 18: Oct 28 - Nov 9. Week 19: Nov 11 - Nov 23

Figure 11: Loss of Employment Income (by gender)

Source: United States Census Bureau Household Pulse Survey.
Week 1: Apr 23 - May 5. Week 2: May 7 - May 12. Week 3: May 14 - May 19. Week 4: May 21 - 26. Week 5: May 28 - June 2. Week 6: Jun 4 - Jun 9.
Week 7: Jun 11 - Jun 16. Week 8: Jun 18 - June 23. Week 9: Jun 25 - Jun30. Week 10: Jul 2 - Jul 7. Week 11: Jul 9 - Jul 14. Week 12: Jul 16 - Jul 21.
Week 13: Aug 19 - Aug 31. Week 14: Sep 2 - Sep 14. Week 15: Sep 16 - Sep 28. Week 16: Sep 30 - Oct 12. Week 17: Oct 14 - Oct 26.
Week 18: Oct 28 - Nov 9. Week 19: Nov 11 - Nov 23

Figure 12: Loss of Employment Income (by race/ethnicity)

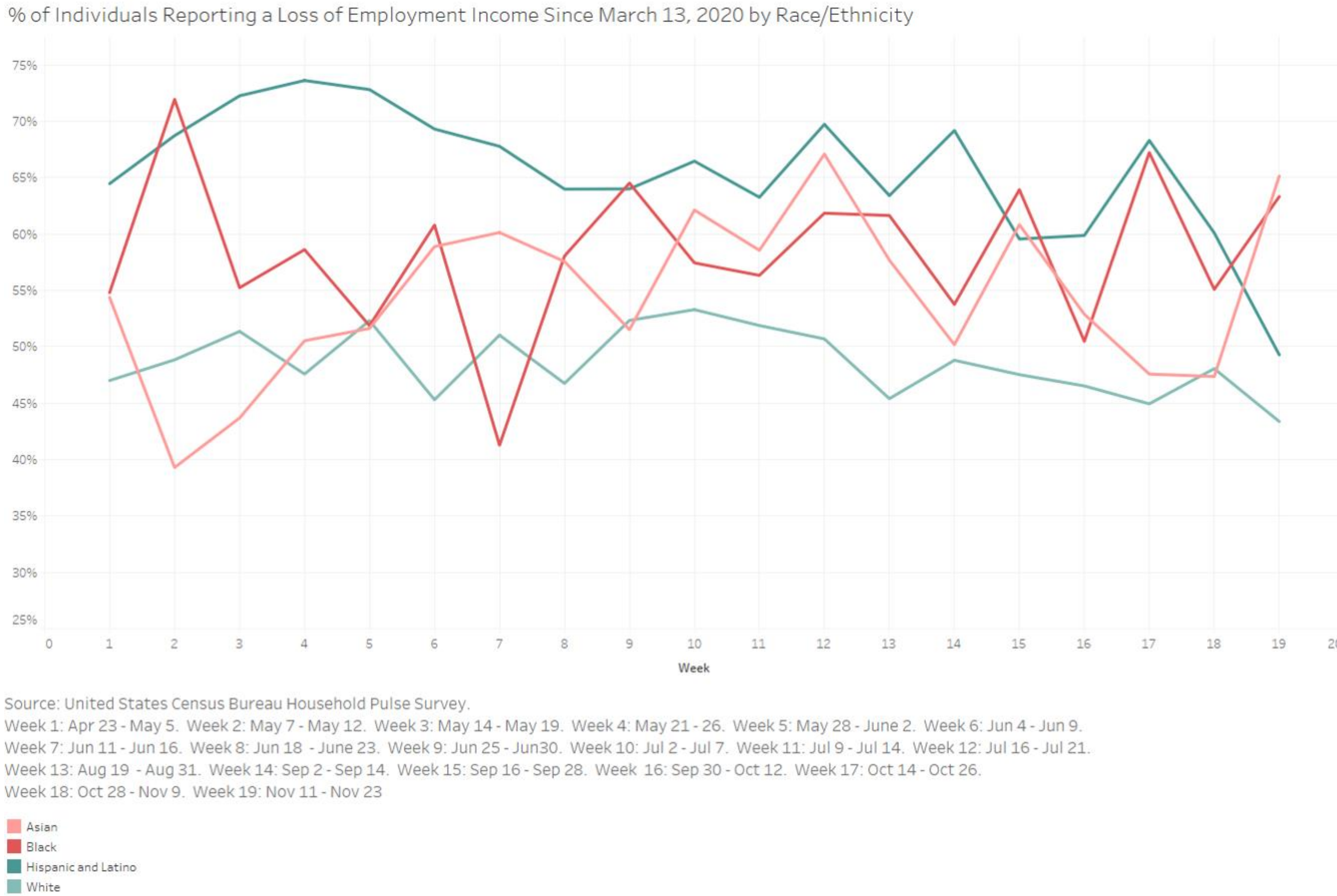


Figure 13: Availability of Nursing Home Beds and Ventilators

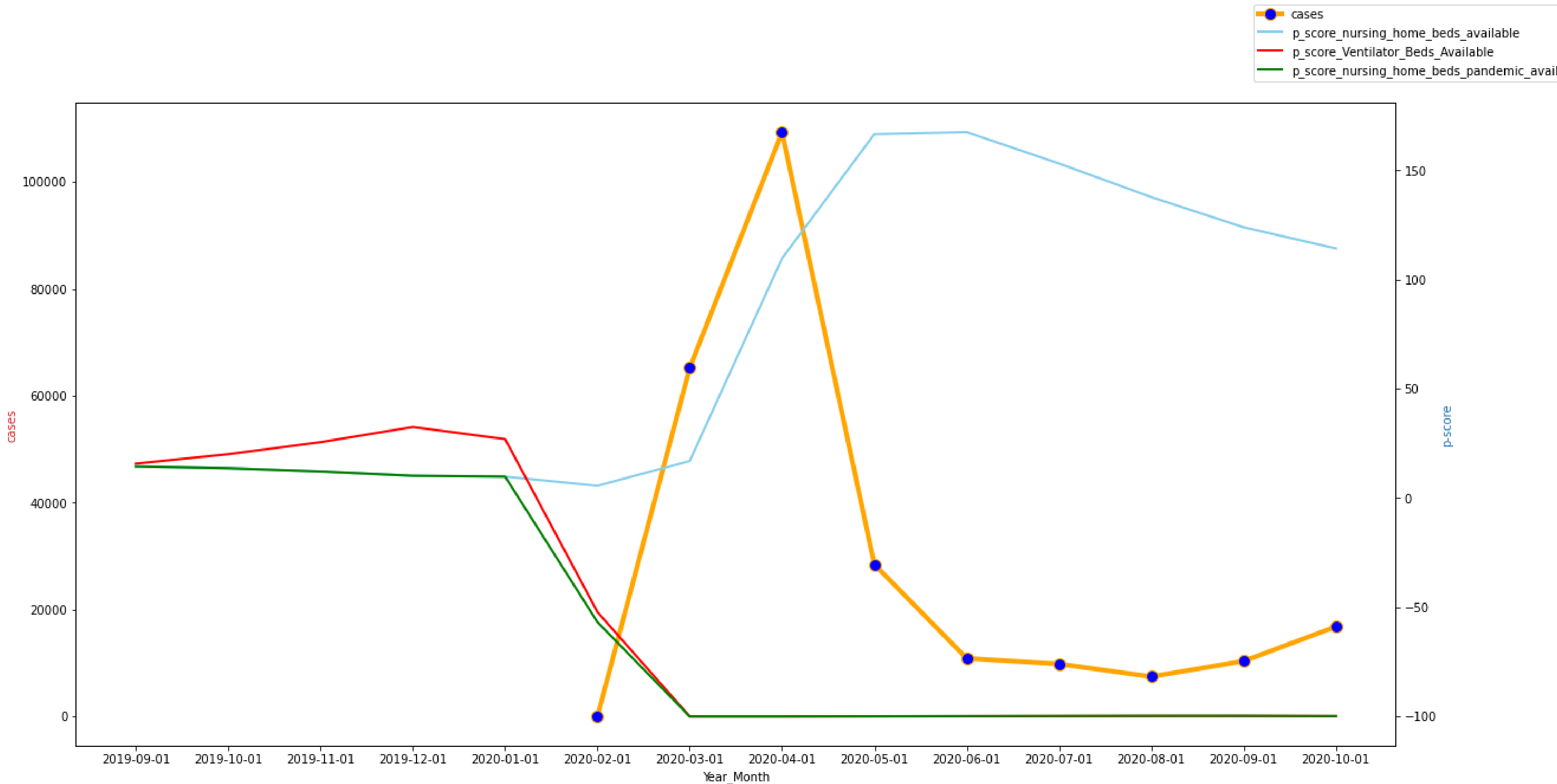
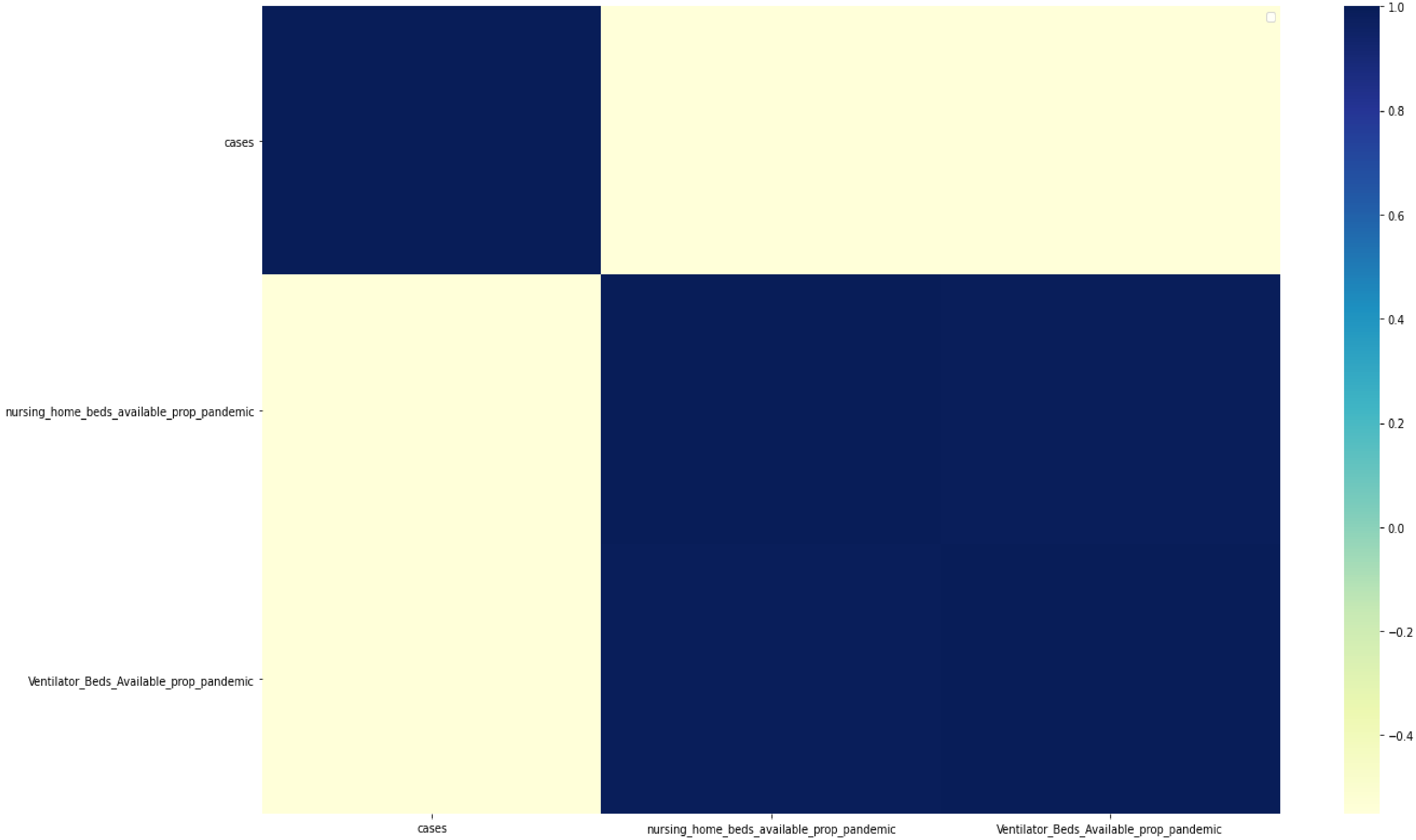
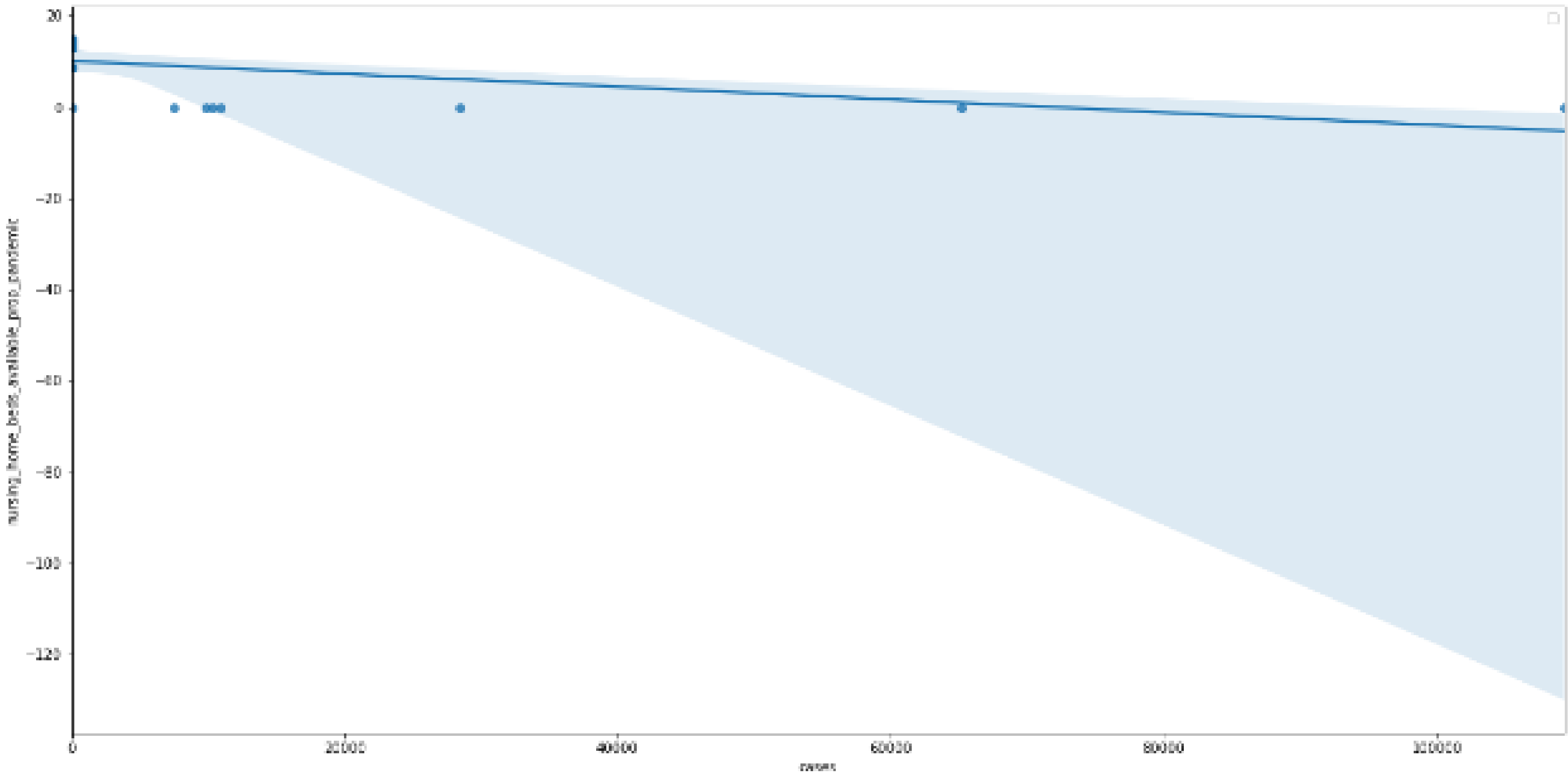


Figure 14: Correlation Matrix and Regression

Plots for Infrastructure Avilability in Healthcare Settings





20
0
−20
−40
−60
−80
−100
−120
nursing_home_beds_available_prop_pandemic
0
20000
40000
60000
80000
100000
cases

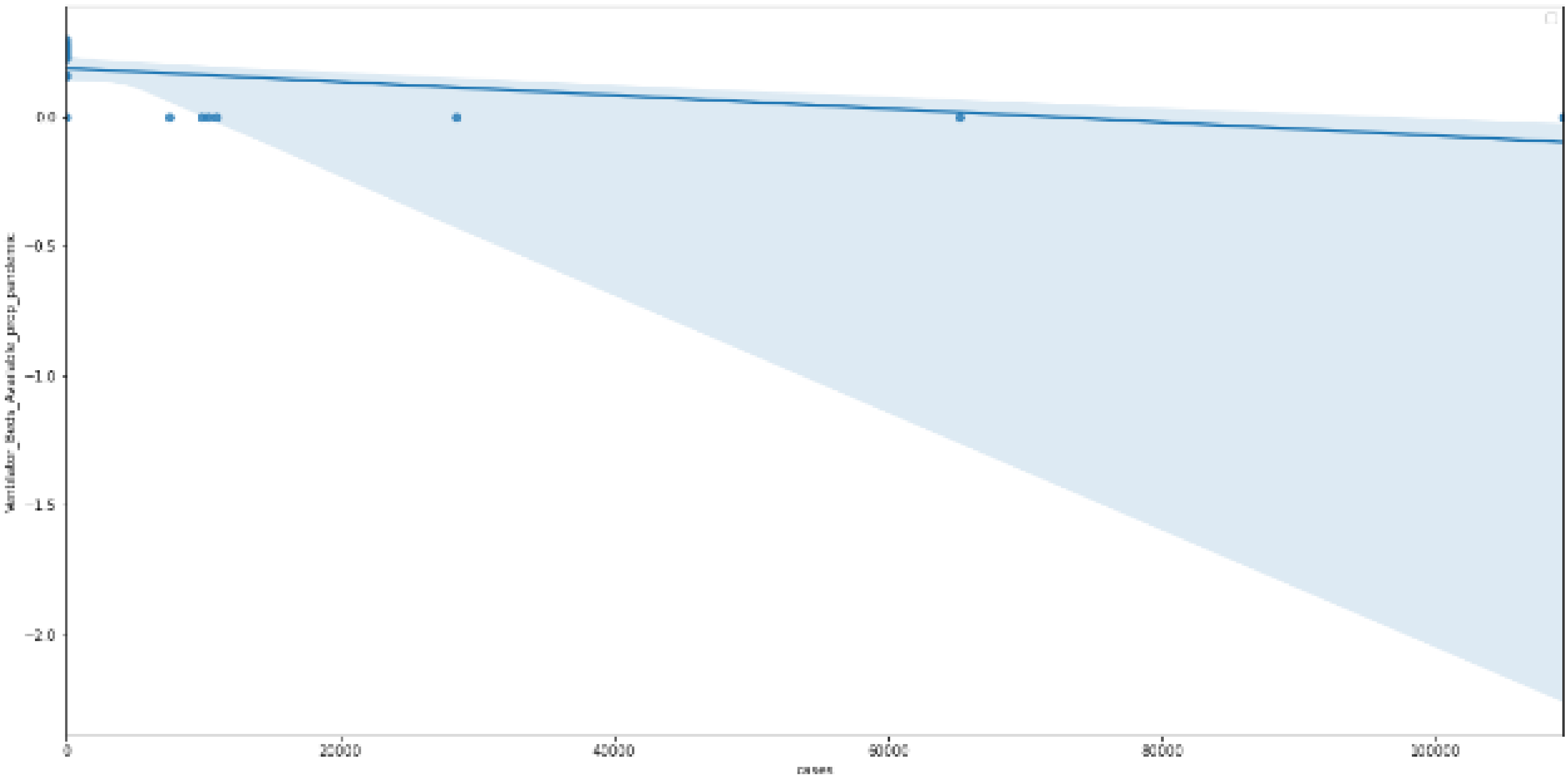

0.0
−0.5
−1.0
−1.5
−2.0
0
20000
40000
60000
80000
100000
cases

Figure 15: Covid Cases vs Safetynet Healthcare Enrollments

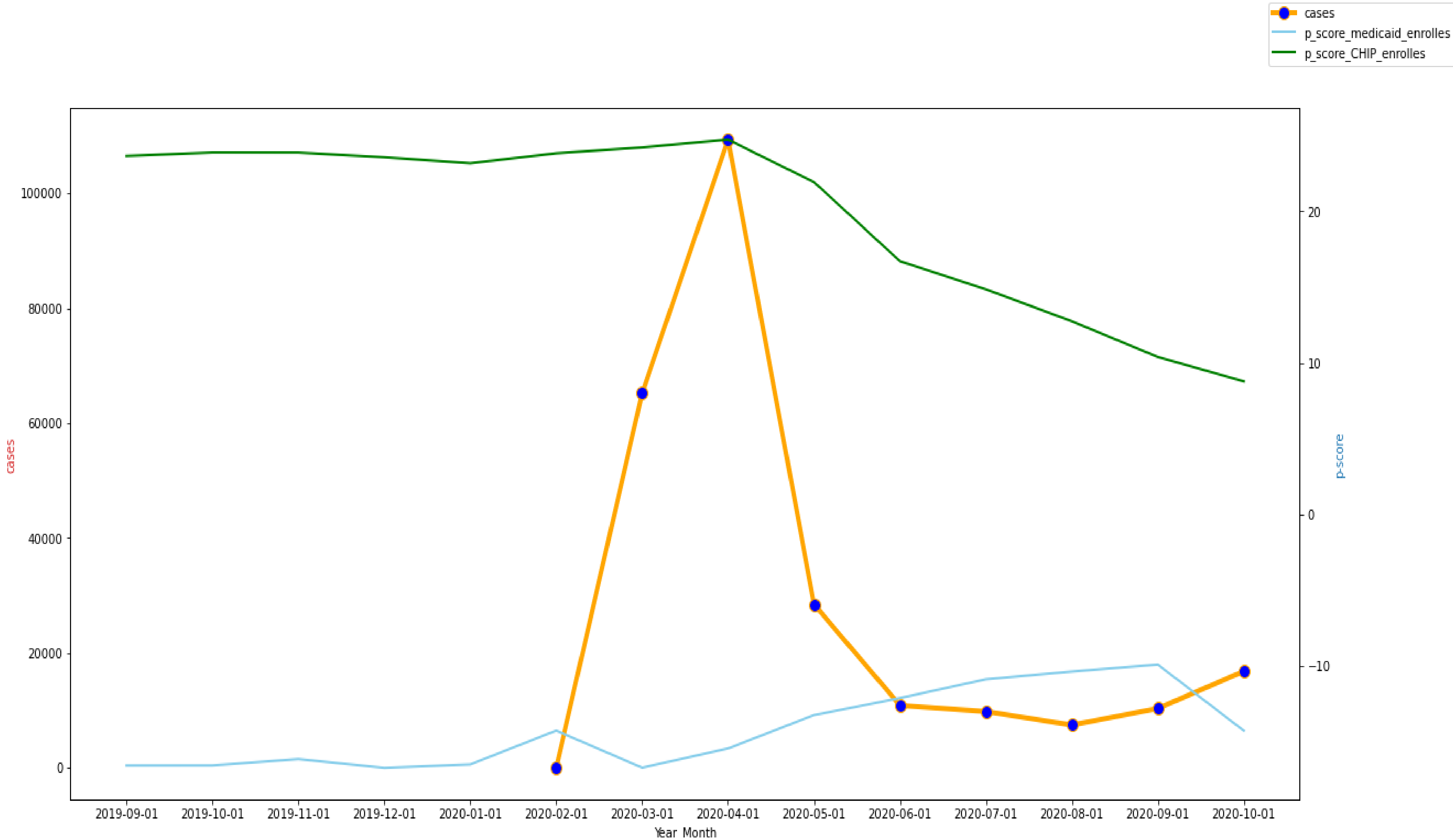
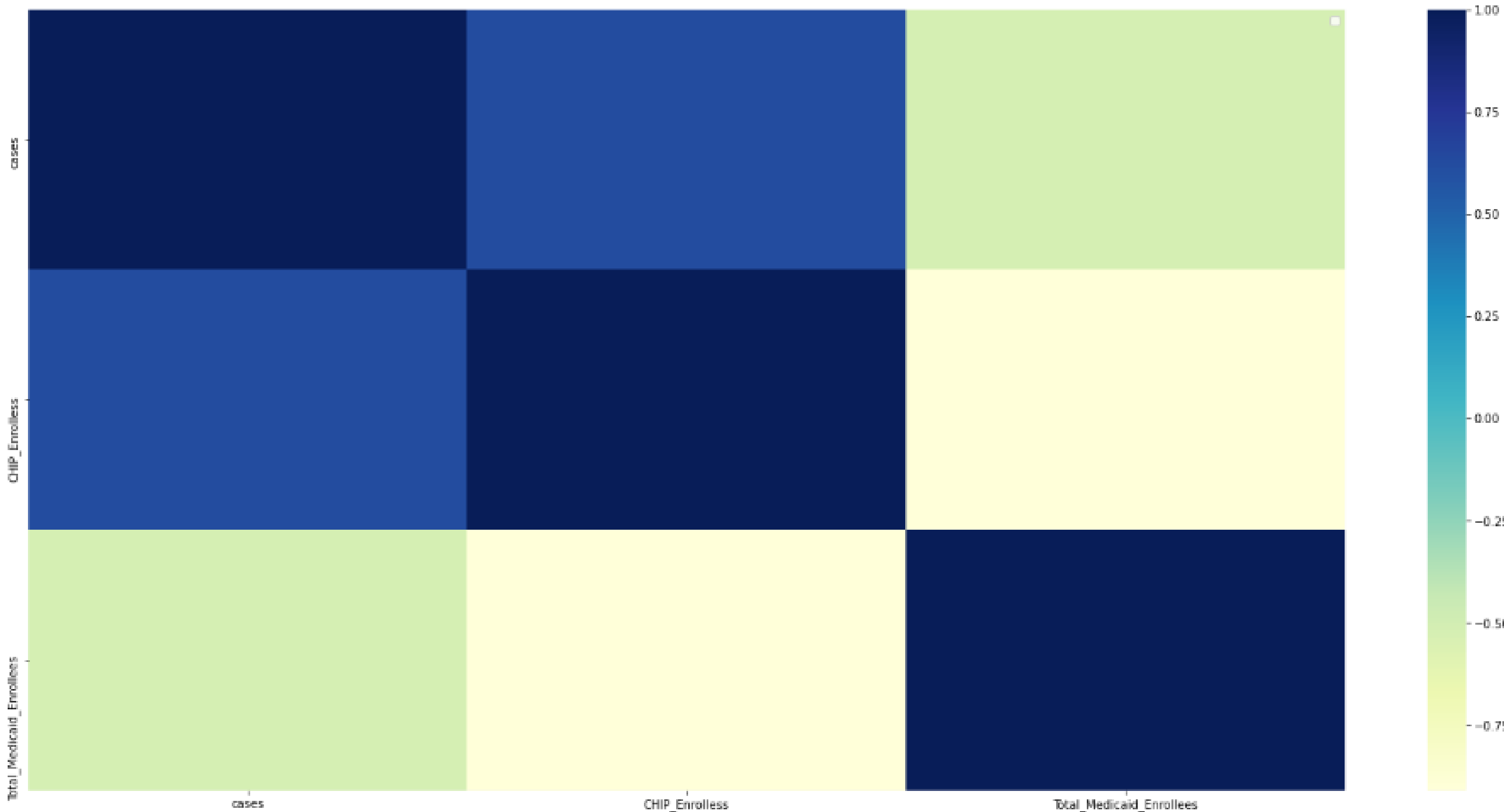
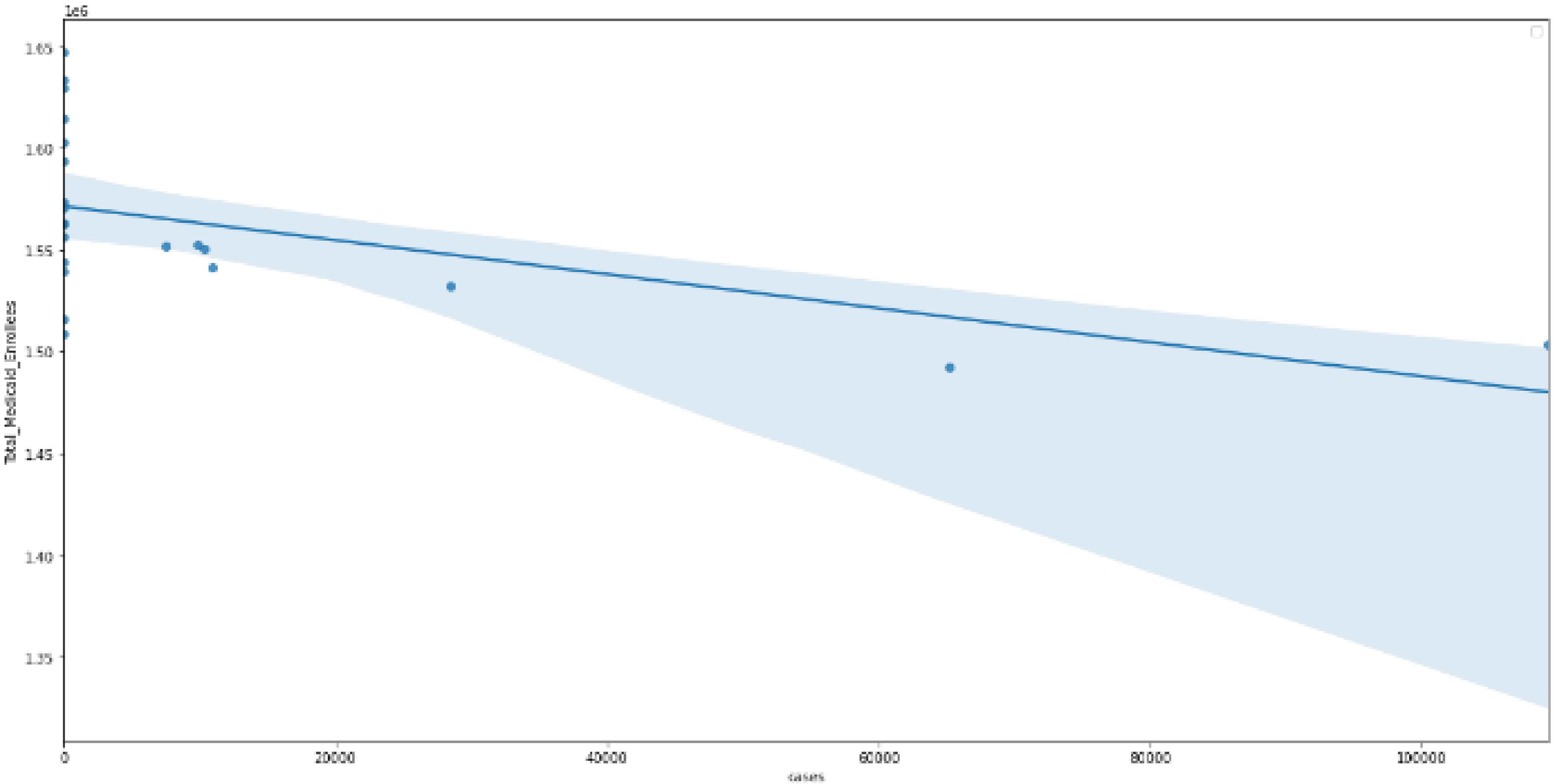
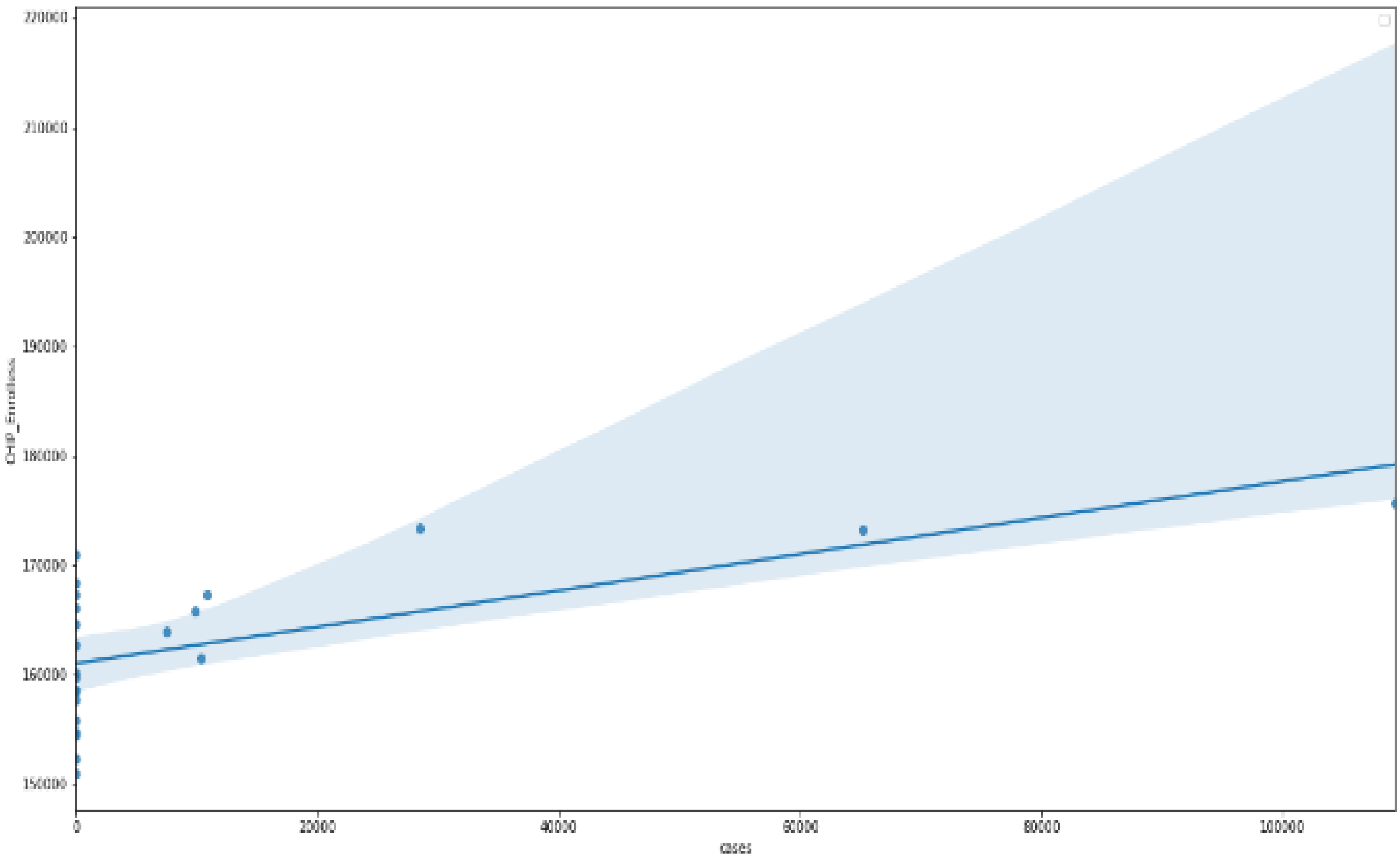


Figure 16: Correlation Matrix and Regression Plots for Health Insurance Enrollments




1e6
1.65
1.60
1.55
1.50
1.45
1.40
1.35
Total_Medicaid_Enrollees
0
20000
40000
60000
80000
100000
cases



220000
210000
200000
190000
180000
170000
160000
150000
CHIP_Enrollees
0
20000
40000
60000
80000
100000
cases

Figure 16: Time series Plots between Shooting Incidents, Arrests and Covid-19 Cases

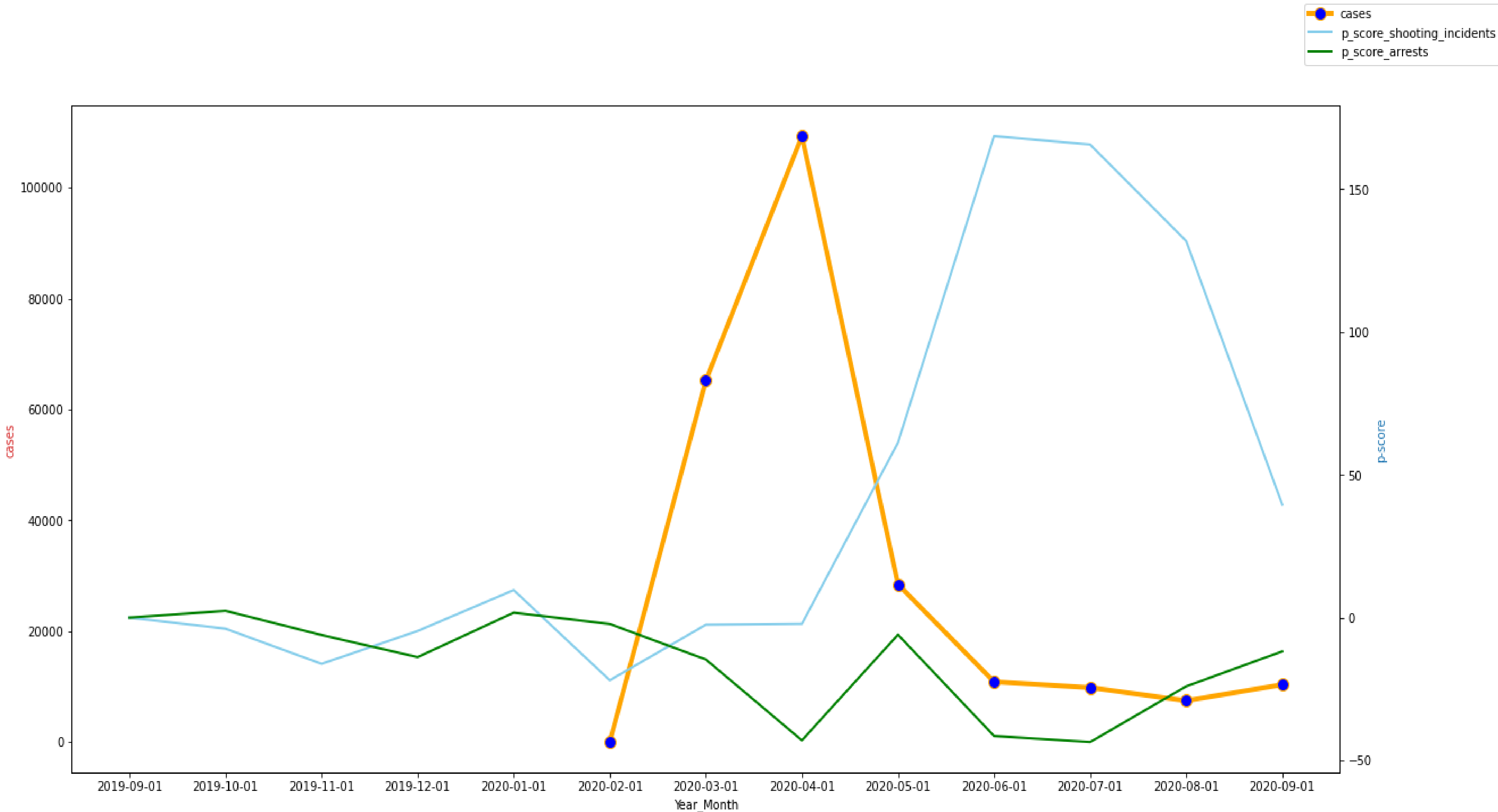
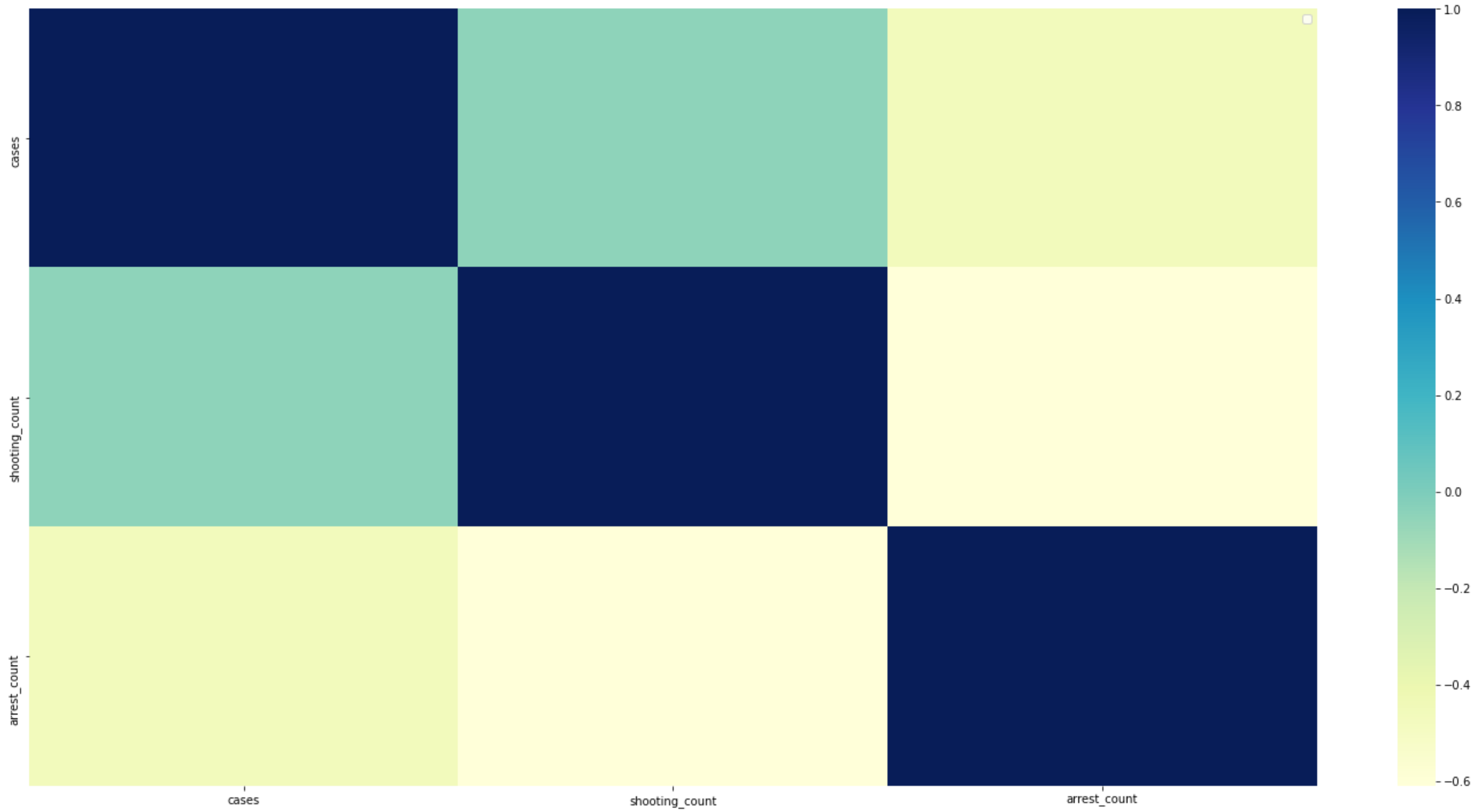
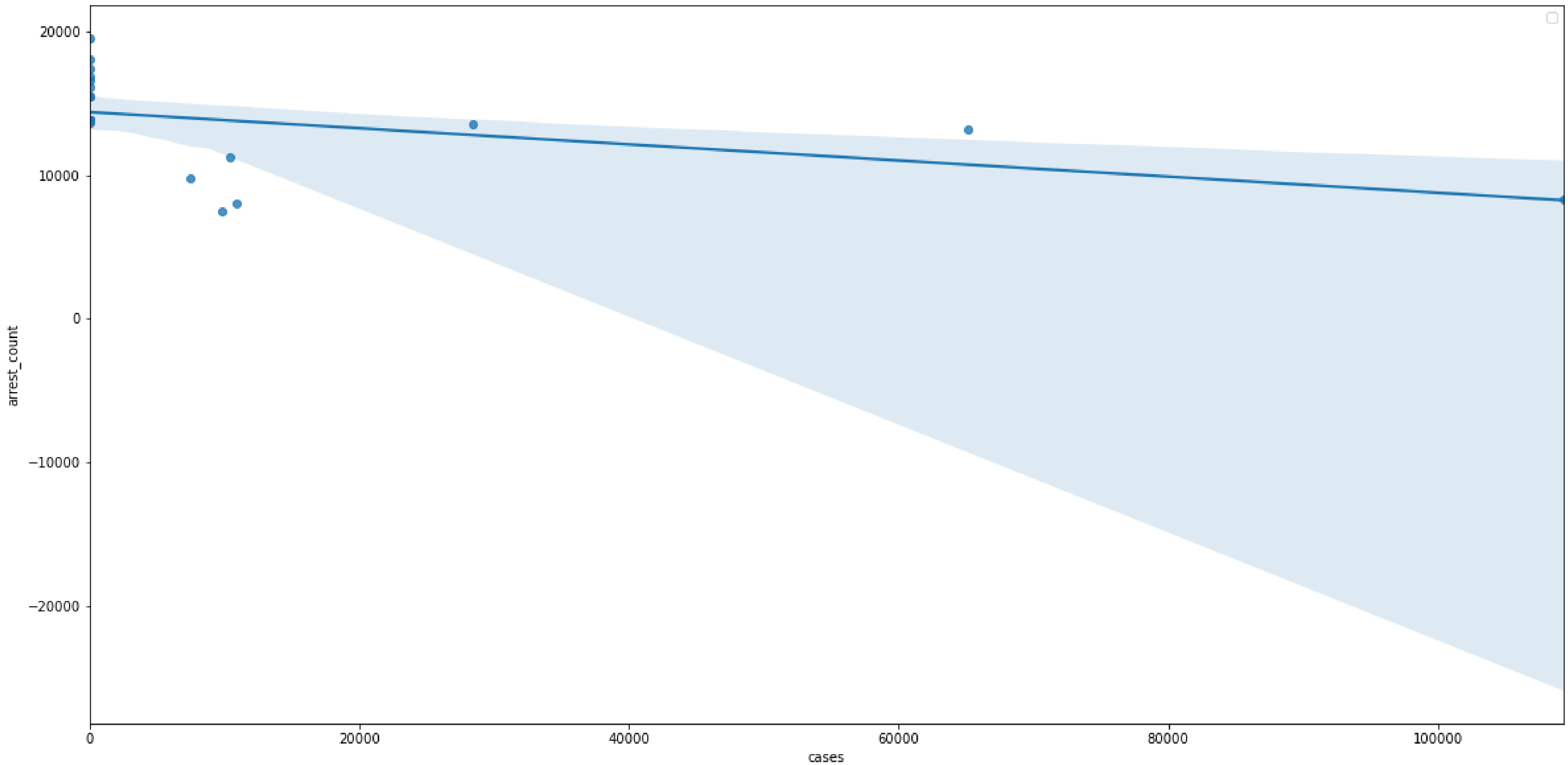


Figure 17: Correlation Matrix and Regression Plots for Public Safety




20000
10000
0
−10000
−20000
arrest_count
0
20000
40000
60000
80000
100000
cases