Factors Affecting Covid Vaccination

Dominic Hoar-Weiler

dhoarwei@andrew.cmu.edu

Introduction

The COVID19 pandemic forced the world to seemingly come to a halt for an entire year. People no longer interacted in the same way they usually would. Upon return to normal activities, however, people had to wear masks. Throughout the entirety of the pandemic, efforts to vaccinate the population were extremely important. Vaccinating the population meant returning to the old way of life, free of the isolation that the pandemic forced us into. However, there were certain people that made this process a lot harder by refusing to become vaccinated. In this study, we will explore factors that may predict the vaccine acceptance rate.

Data

The data that we will be looking at in this study was compiled by a combination of Carnegie Mellon University's COVIDcast project by the Delphi research group and the Kaiser Family Foundation. The Delphi research group obtained some of their findings using a very large database of medical insurance claims and large random samples of Facebook users age 18 and older. We have data on all fifty states in the United States of America. Each row in the data provides information and data on one of those states. For each state, the data records the **vaccine acceptance rate** among respondents to the COVIDcast survey (in August 2021), the percentage of the state population that is **uninsured**, the total **private health insurance spending** (in 2014 in millions of dollars), the **number of births** (in 2019), the **infant mortality rate** (in 2018), the **fire arm death rate** per 100,000 residents (in 2018), the **median annual household income** of residents in that state (in dollars), the **political affiliation of the governor** in that state (either Republican or Democrat), the **state senate majority** political affiliation, the **state house majority** political affiliation, the **total gross state product** (in millions of dollars), the **state unemployment claims** (in the week 8/28/2021), the **average monthly Supplemental Nutrition Assistant Program participants** (in 2019), the percentage of adults who **smoke**, the **drug overdose death rate** (in overdoses per 100,000 residents), the hospital **adjusted expenses per impatient day** (in 2019), and the **total population** (in 2019).

The primary focus of this data analysis will be the **vaccination acceptance rate**, a quantitative variable measuring the percentage of the total survey respondents that accepted the covid vaccine in a given state. These results were obtained through large random samples of people 18 years old and above on Facebook. Because not all adults are on Facebook, it is important to recognize that the sample could be slightly biased. This statistic, of course, only estimates the total percentage of the state population that accepted the covid vaccine in that state. The mean vaccination acceptance rate across all states was 86.461% and the median was 86.039%, with a standard deviation of 4.370%. In figure 1, we see that the distribution of the vaccination acceptance rate of the states is unimodal and fairly symmetrical. There is also a clear outlier: Wyoming with a vaccination acceptance rate of 74.272%. The research questions will also include the states' **governor political affiliation**, **total private health insurance spending** and **drug overdose death rate**.

The variable **governor political affiliation** is a categorical variable that indicates the political affiliation of the governor in a given state, either Republican or Democrat. The majority of governors at this time were Republican, numbering 27 (54.0%). On the other hand, 23 (46.0%) of the states had governors who were affiliated with the Democratic party. The **total private health insurance spending** is a quantitative variable measuring each state's total money spent on private health insurance in millions of dollars. The mean money spent on private health insurance was $17,429.300 million, and the median was $11,454.000 million. The standard deviation of the private health insurance spending was $19,415.867 million. In figure 2, we see that the distribution of the states' private health insurance spending is unimodal and right skewed, with three outliers. California, Texas and New York are outliers in private health insurance spending with $104,018 million, $67,494 million, and $61,730 million spent respectively. The **drug overdose death rate** is a quantitative variable that measures each state's rate of drug overdose deaths per 100,000 residents. The mean drug overdose death rate was 22.738 deaths per 100,000 residents, and the median was 20.600 deaths per 100,000 residents. The standard deviation of the drug overdose death rate was 9.872 deaths per 100,000 residents. In figure 3, we see that the distribution of the states' drug overdose death rates is unimodal and right skewed, with two outliers. The two outliers are West Virginia and Delaware with 52.8 and 48 deaths per 100,000 residents respectively.

Figure 1

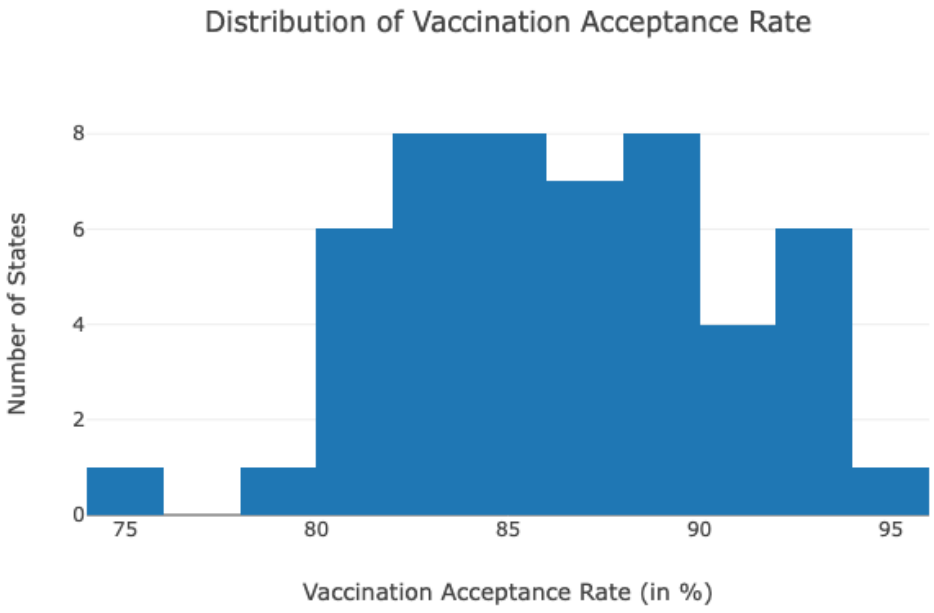


Figure 2

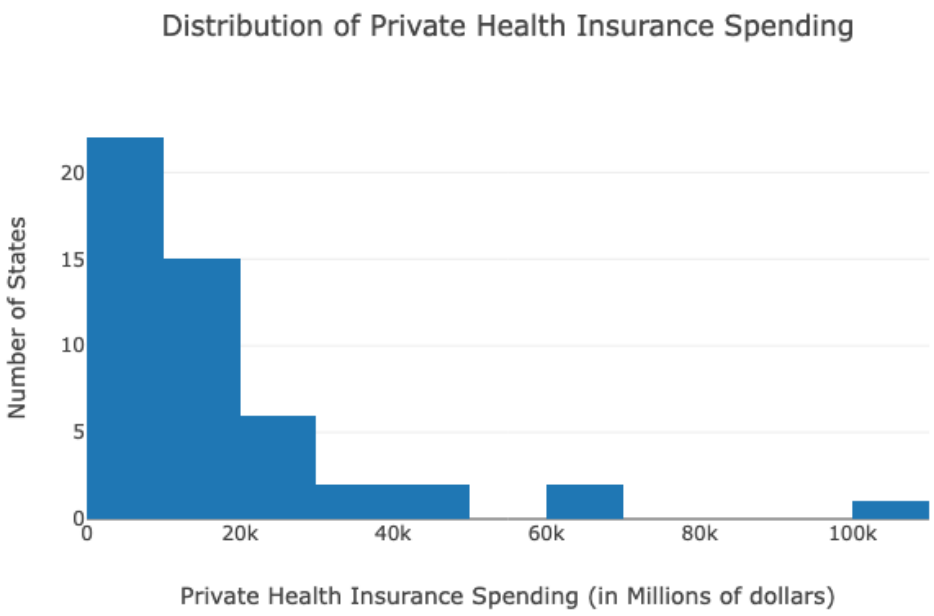
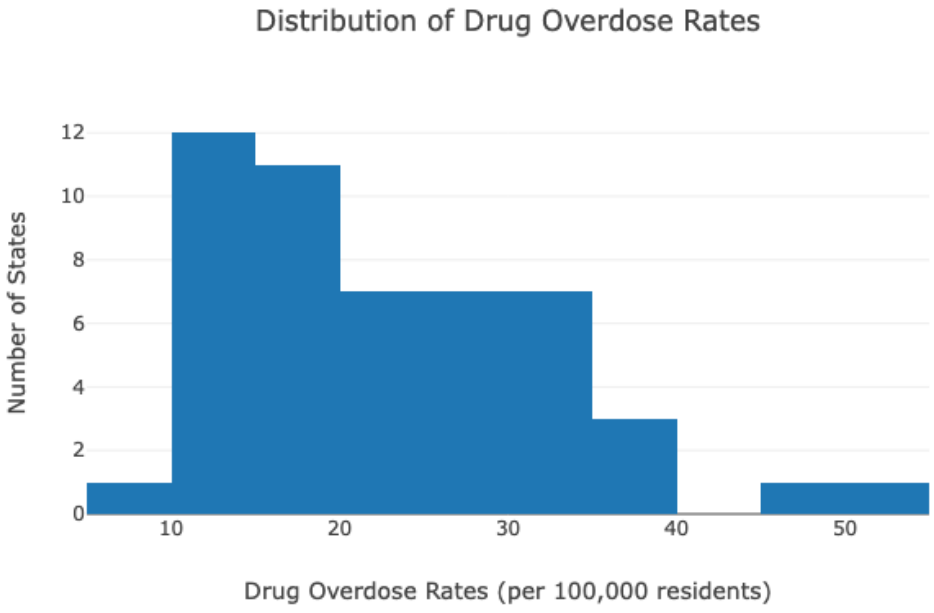


Figure 3



Relationship between Vaccination Acceptance Rate and Governor Political Affiliation

First, we are interested in exploring whether there is a relationship between vaccination acceptance rate and the governor's political affiliation. Our question is: is there a relationship between vaccination acceptance rate and the governor's political affiliation? We hypothesize that there will be relationship between vaccination acceptance rate and the governor's political affiliation, as Republican ideology has traditionally focused more on personal freedoms, including one's ability to

refuse vaccinations. We think the governor's political affiliation is a good indicator of whether the state espouses more of a Republican or Democratic ideology in general. As such, we think the vaccination rate in states with Democratic governors will be higher than in states with Republican governors. To formalize this hypothesis, we will focus on the mean vaccination acceptance rate in states with Republican governors and the mean vaccination acceptance rate in states with Democratic governors. Let muR and muD denote the true mean vaccination acceptance rate (in %) of residents in states with Republican affiliated governors and states with Democrat affiliated governors respectively. Our null and alternative hypotheses are

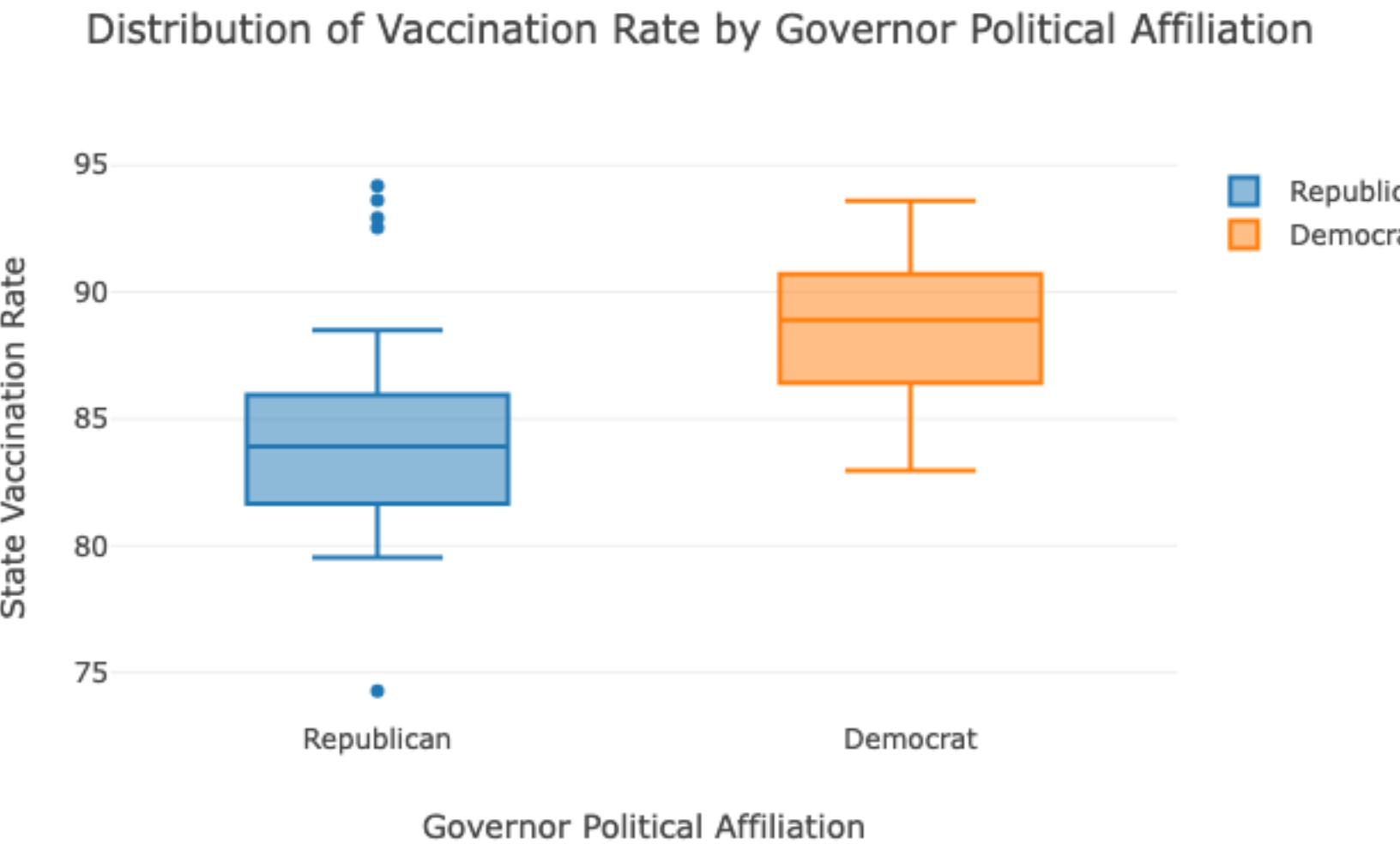
H0: muD - muR = 0

Ha: muD - muR > 0

Our estimators of the true means are X-barD = 88.655% and X-barR = 84.592%.

Before formally testing the hypotheses, we visually compare the distribution of vaccination acceptance rates for the two governor political affiliations. Figure 4 displays a boxplot of vaccination acceptance rates for each group. The figure shows that the majority of vaccination acceptance rates in states with Republican governors are less than 85.845% while the majority of vaccination acceptance rates in states with Democrat governors are greater than 86.423%. This does suggest that there is a relationship between governor political affiliation and the state's vaccination acceptance rate.

Figure 4



In order to perform a test, we must first check the conditions. No state governor can be both affiliated with the Republican party and the Democratic party so the two groups are independent from each other. The study states that the polls' participants were obtained through random sampling. As such, we satisfy this condition for the vaccination acceptance rates. The study doesn't specify the sample sizes of the polls in each state, but for the sake of the test, we can assume that the total population of each state is greater than 20 times the polls' sample sizes. The study states that the sample size of people 18 years old and above was large, so by this, we can assume it is large enough to justify a t-distribution for the t-test. Since the true standard deviations for the vaccination acceptance rates for each group is unknown, we use a t-test.

To formally test our hypotheses, we use a two sample t-test with a level alpha = 0.05. The test statistic is t = 3.8062 and the resulting p-value is approximately 0.0002. Since this p-value is less than our alpha value of 0.05, we reject our null hypothesis. In context, this means that there is sufficient evidence to suggest that the true mean vaccination acceptance rate is higher in states with governors who are affiliated with the Democratic party than in states with governors who are affiliated with the Republican party.

Relationship between Vaccination Acceptance Rate and Private Health Insurance Spending

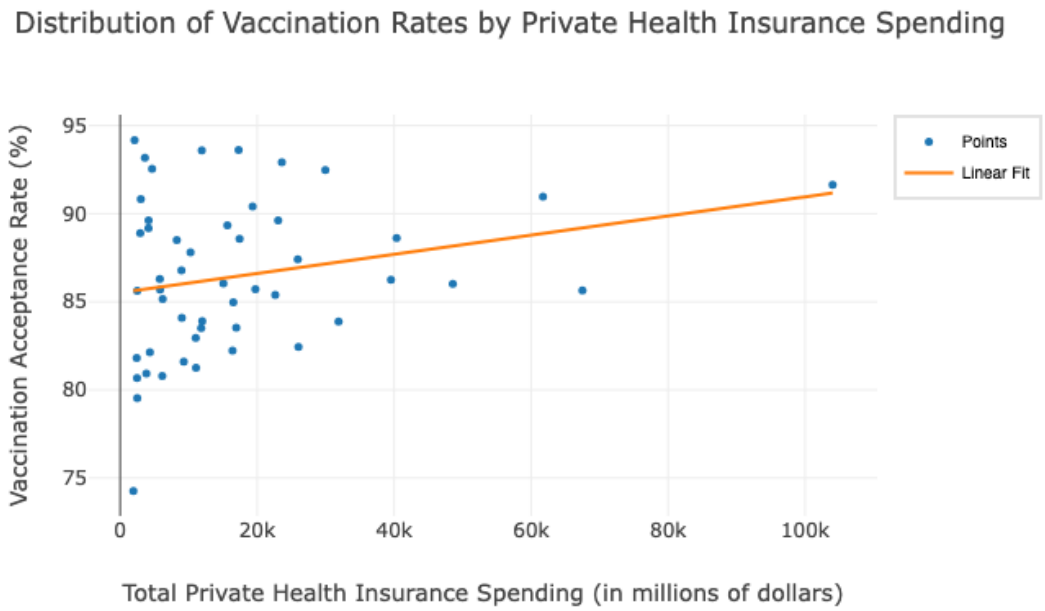
Now, we are interested in exploring whether there is a relationship between vaccination acceptance rate and private health insurance spending in each state. Our question is: is there a relationship between vaccination acceptance rate and private health insurance spending in each state? We hypothesize that states with higher private health insurance spending will have higher vaccination acceptance rates because there will likely be more accessibility to the vaccine as well as more trust in health providers. Let slope1 be the slope of the regression line crafted between vaccination acceptance rate and private health insurance spending. As such, we have our null and alternative hypotheses:

H0: slope1 = 0

Ha: slope1 > 0

Before formally testing the hypotheses, we visually compare the scatterplot of the vaccination acceptance rates and the private health insurance spending of each state. Figure 5 displays a scatterplot of these variables, with a regression model on top of the data. From this, we can see that there are some outliers in total private health insurance spending, for example at more than $100,000 million and a state with a vaccination acceptance rate of below 75%. But, they don't seem to be affecting the regression line too much, so we can proceed. We can also see the right skewness of the private health insurance spending in the figure. The regression illustrates a slight relationship between the two variables.

Figure 5



In order to justify the test for a quantitative linear relationship, we must check certain requirements.

First, the graph above satisfies the linear relationship check, as the distribution looks to be pretty linear. Since random sampling was used to obtain the sample statistics for the vaccination acceptance rates, we satisfy the randomization requirement. Next, the data satisfies the independence condition as, in figure 5, we can see that the points deviations from each other appear reasonably random and without any pattern. Because the study states that the sample sizes for each state's vaccination acceptance polls were large, we can say that the data satisfies a t-distribution. Lastly, the vertical spread around the line is reasonably consistent, so we fulfill the spread requirement. As such, we have justified a test for quantitative linear relationship. In figure 6 below, we see the result of such a test.

Figure 6: Regression of Vaccination Acceptance on Total Private Health Insurance Spending (model id: slm2)

Variable	Coefficient	Std. Error	t	p-value
Intercept	85.5132	0.8180	104.5456	0
total_private_health_insurance_spending	0.0001	0.0000	1.7244	0.091

To formally test our hypotheses, we used a simple linear regression test with a level alpha = 0.05. The statistic is slope1 = 0.0001, meaning for every million dollar increase in private health insurance spending, the vaccination acceptance rate increases by 0.0001. The resulting p-value is approximately 0.091. Since this p-value is greater than our alpha value of 0.05, we do not reject our null hypothesis. In context, this means that there is not sufficient evidence to suggest that there is a relationship between private health insurance spending and vaccination acceptance rate.

Relationship between Vaccination Acceptance Rate and Drug Overdose Death Rates

Lastly, we want to investigate whether there is a relationship between vaccination acceptance rate and drug overdose death rates. Our question is: is there a relationship between vaccination acceptance rate and drug overdose death rates in each state? We hypothesize that states with higher drug overdose death rates will have lower vaccination acceptance rates as residents would be more skeptical of new drugs, such as the vaccine. Let slope represent the slope of the regression line crafted between vaccination acceptance rate and drug overdose death rate. As such, we have that our null and alternative hypotheses are:

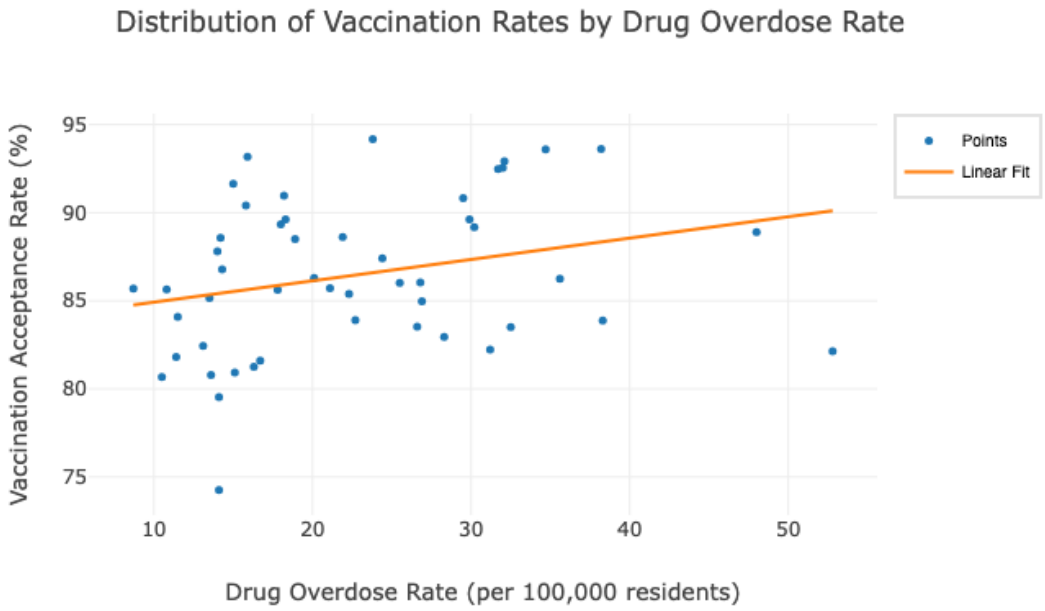
H0: slope = 0

Ha: slope < 0

Before formally testing the hypotheses, we visually analyze the scatterplot of the vaccination acceptance rate and the drug overdose death rate of each state. Figure 7 displays a scatterplot of these variables, with a regression model on top of the data. From this, we can see that there are some outliers; there is a state with a drug overdose death rate greater than 50 deaths per 100,000 residents and a state with a vaccination acceptance rate of below 75%. Since there are only two

clear outliers, we will proceed without removing them. The regression illustrates a slight positive relationship between the two variables, contrary to our prediction of a slightly negative one.

Figure 7



In order to justify the test for a quantitative linear relationship, we must once again check certain requirements. First, the graph above satisfies the linear relationship check, as the distribution appears to be pretty linear. Similar to the last example, since random sampling was used to obtain the sample statistics for the vaccination acceptance rates, we satisfy the randomization requirement. Next, the data satisfies the independence condition as, in figure 7, the deviations between points seem to be patternless. Because the study states that the sample sizes for each state's vaccination acceptance polls were large, we can say that the data satisfies a t-distribution. Lastly, the vertical spread around the line is reasonably consistent, so we fulfill the spread requirement for a test. As such, we have justified a test for quantitative linear relationship. In figure 8 below, we see the result of such a test.

Figure 8: Regression of Vaccination Acceptance on Drug Overdose Death Rate (model id: slm3)

Variable	Coefficient	Std. Error	t	p-value
Intercept	83.7028	1.5208	55.0384	0
drug_overdoses	0.1213	0.0614	1.9739	0.0542

To formally test our hypotheses, we used a simple linear regression test with a level alpha = 0.05. The statistic is slope = 0.1213, meaning for every increase in the drug overdose death rate, the vaccination acceptance rate increases by 0.1213. The resulting p-value is approximately 0.0542.

Since this p-value is greater than our alpha value of 0.05, we do not reject our null hypothesis. In context, this means that there is not sufficient evidence to suggest that there is a relationship between drug overdose death rate and vaccination acceptance rate. However, since our p-value is very close to our alpha level, it is important to recognize that our outliers may have affected our outcomes. If we were to preform the test without these outliers, we may get a different result. However, for the sake of this test, we did not find sufficient evidence of an association.

Discussion

All 50 states in the United States had data compiled on a myriad of different topics. This data included the results of a survey on vaccination acceptance rate from a sample group in each state. The survey's results and the other data compiled on the 50 states were then examined for potential relationships between the vaccination acceptance rates, the political affiliation of the governors, the total private health care spending, and the drug overdose death rates. We were able to find a significant correlation between the governor's political affiliation and the vaccination acceptance rate in a state. We found sufficient evidence that Democrat Governor's correlate to higher vaccine acceptance rates in states compared to Republican Governors. That being said, we weren't able to find any evidence that either the total private health care spending or the drug overdose death rate had an association with the vaccination acceptance rate of a state.

These findings reveal how political vaccinations are. Democrat states correlate to higher vaccination acceptance rates. In future analyses, it would be beneficial to examine some other factors that might help reveal more of the factors that influence vaccination acceptance rates. For example, while we initially found no evidence of an association between drug overdose death rate and vaccination acceptance rate, we might try removing some of the outliers to see if a trend develops. We may also want to explore the relationship between the median annual household income and the vaccination acceptance rate in each state. This is because it wouldn't be too surprising that wealthier states had more accessibility to vaccines and thus had higher vaccination acceptance rates.

In summary, this analysis can help us be aware of some of the factors that contribute to lower vaccination rates. In the future, we can use these findings to help motivate more targeted marketing campaigns to increase the amount of people who become vaccinated. In other words, these findings can be used for the advancement of the public good.