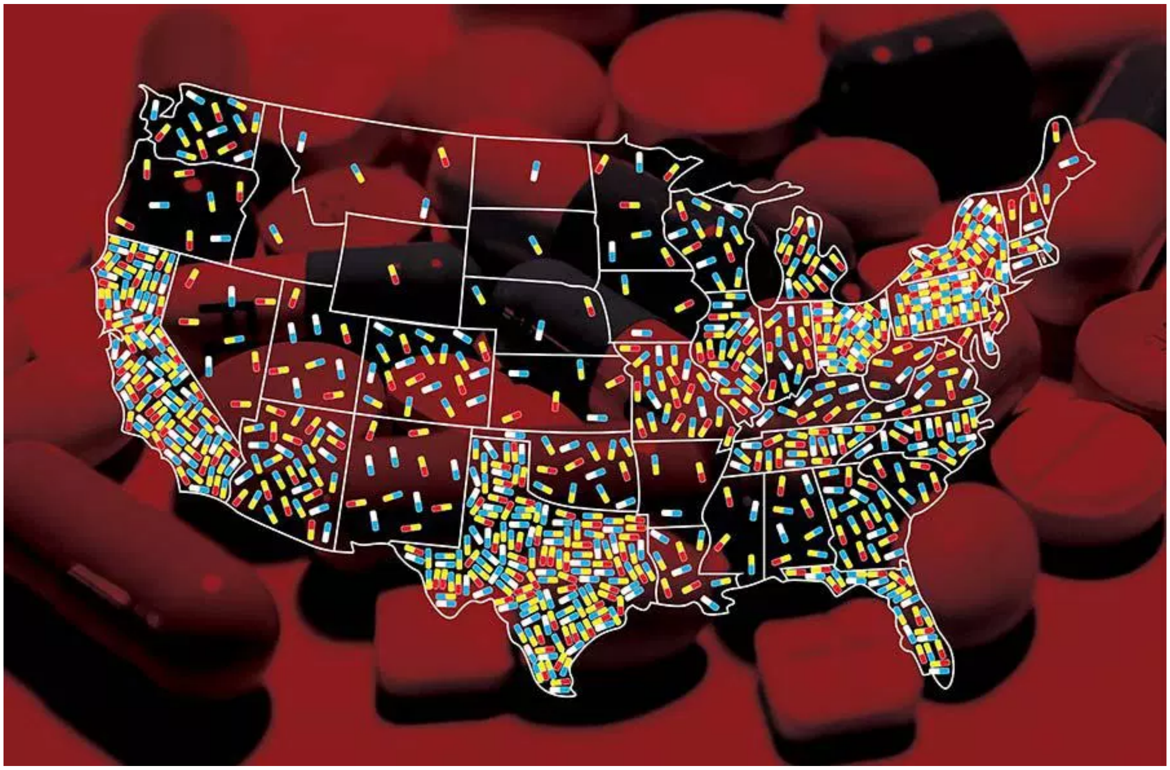
Data Analytics Project on Opioid-Related Deaths



More Frightening than Terrorism: The Opioid Epidemic

Submitted By:

Mario Espinosa
Jamie Quolas
Neha Singh
Carolyn Super
Zhenglong Wu

Table of Contents

Executive Summary:

Addiction and overdoses associated with prescription and illicit opioids have been characterized by the U.S. Department of Health and Human Services as a national crisis. Since 2000, the rate of overdose deaths involving opioids has increased fourfold. In last three years **2015, 2016, 2017** opioid overdose deaths increased to **3,091, 42,249,** and **49,068** respectively, this number is greater than any other cause of deaths number including, deaths from suicide, road accidents, terrorism etc. Every day, more than **115** people in the United States die after overdosing on opioids. The misuse and addiction to opioids—including prescription pain relievers, heroin, and synthetic opioids such as fentanyl— has become a serious national crisis that affects public health as well as social and economic welfare. The Centers for Disease Control and Prevention estimates that the total "economic burden" of prescription opioid misuse alone in the United States is **$78.5 billion** a year, including the costs of healthcare, lost productivity, addiction treatment, and criminal justice involvement. This facts and findings motivated us to do our data analytics project on opioid overdose deaths. To pursue this analysis, the data has been gathered from CDC- Centers for Disease Control and Prevention. The data included the death statistics along with many other factors like, population, state, crude.rate, prescriptions along with the years from 2006 to 2014. Standard approach has been followed by performing the exploratory analysis, forming hypothesis and conducting the regression analysis on the data. The data has been modelled using 4 models and finding the one which predicts the data best and also is practically suitable. The key findings from our analysis suggest as:

1. The population increases opioid overdose death rate increases irrespective of other factors and it comes out to be a significant factor in our study.
2. Prescription rate also has impact on number of opioids overdose deaths but it's not as significant as increase in population.
3. Analysis have identified both Michigan and Maine have good policies and treatment of opioids prescriptions as their opioid overdose death rates have decreased over the last 8 year while the population has increased over the year. This suggests that government official or health care policy maker can study what measure they are taking to combat with opioid problem and can implement those policy or measures to state with higher number of deaths like Florida.
4. The following states have the most opioid overdose deaths without any influence of population and prescriptions: West Virginia, Nevada, Utah, Maryland and Massachusetts.

Problem Significance:

In this project we want to study the relation between opioid overdose deaths with respect of number of prescription statewise in United States. We want to see some patterns or trends in data and want to give some inputs to government officials, health care professionals who are responsible for making health care related policies and regulation and responsible for its proper implementation.

In the late 1990s, pharmaceutical companies reassured the medical community that patients would not become addicted to prescription opioid pain relievers, and healthcare providers began to prescribe them at greater rates. This subsequently led to widespread diversion and misuse of these medications before it became clear that these medications could indeed be highly addictive. Opioid overdose rates began to increase. In 2015, more than 33,000 Americans died as a result of an opioid overdose, including prescription opioids, heroin, and illicitly manufactured fentanyl, a powerful synthetic opioid. That same year, an estimated 2 million people in the United States suffered from substance use disorders related to prescription opioid pain relievers, and 591,000 suffered from a heroin use disorder (not mutually exclusive).

What do we know about the opioid crisis?

- Roughly 21 to 29 percent of patients prescribed opioids for chronic pain misuse them.
- Between 8 and 12 percent develop an opioid use disorder.
- An estimated 4 to 6 percent who misuse prescription opioids transition to heroin.
- About 80 percent of people who use heroin first misused prescription opioids.
- Opioid overdoses increased 30 percent from July 2016 through September 2017 in 52 areas throughout 45 states.
- The Midwestern region saw opioid overdoses increase 70 percent from July 2016 through September 2017.
- Opioid overdoses in large cities increase by 54 percent in 16 states.

This issue has become a public health crisis with devastating consequences including increases in opioid misuse and related overdoses, as well as the rising incidence of neonatal abstinence syndrome due to opioid use and misuse during pregnancy.

The practical application of our data analysis can be used by medical professionals in determining pain treatment options for their patients. In observing any correlation between an increase in prescription rates and an increase in deaths, we can reasonably conclude that a decrease in prescription rates would lead to a corresponding decrease in deaths. It can also be used by the federal government, providing the necessary knowledge to effectively distribute funds from the federal drug-abuse program to the various state-run programs. Assuming a key mission of the federal program is to reduce drug-related deaths, allocating funding to states that have a higher number of deaths in proportion to their population will provide the resources needed to combat the issue.

Data Source/Preparation:

The portion of data containing information about state populations and opioid-related deaths was extracted from the National Center for Health Statistics and compiled with additional information from the CDC's WONDER database concerning the number of prescriptions. Accompanying the variables deaths, state, population, and prescriptions were various ratios that provided little additional insight to our analysis. In our data, opioid deaths are defined as deaths in which the underlying cause was drug overdose. The drugs included in the variable are heroin, opium, other opioids, methadone, other synthetic narcotics, and other unspecified narcotics.

We were unable to perform analysis on the state of North Dakota. The data for all but one year presented were suppressed. In other words, according to the data, disclosing the information could lead to the disclosure of individual identities. We do not believe that excluding North Dakota impacted our analysis of the other states. Wyoming, Arkansas, and District of Columbia each had one year that data on the number of opioid-related deaths was unavailable. The one year of data was excluded from the analysis of those states. We do not believe it significantly impacted our models.

Hypotheses:

We have structured our hypotheses around two important implications of our data to the public.

H1a: βPopulation > 0

The first set of hypotheses concern the number of opioid related deaths in a state compared to the states' population. This will inform us which states have the highest rate of death and answer the question of where the government should be allocating its resources to when assessing the funding of treatment centers, programs, and other projects related to the war on drugs.

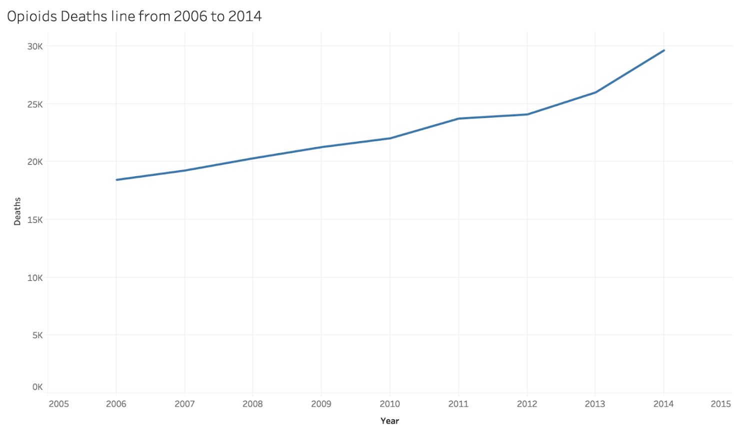
H2a: βPrescription > 0

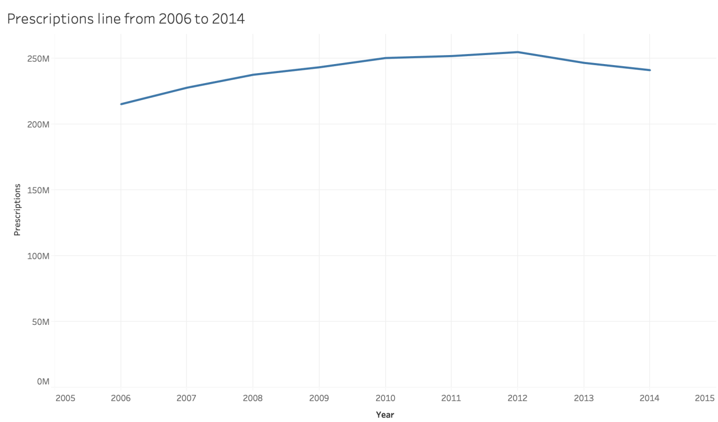
The second set of hypothesis correspond the number of prescriptions to the number of deaths and will tell us if the increase in opioid prescriptions are related to the rise in opioid related deaths, or if the rise in deaths is due to another factor not included in our data.

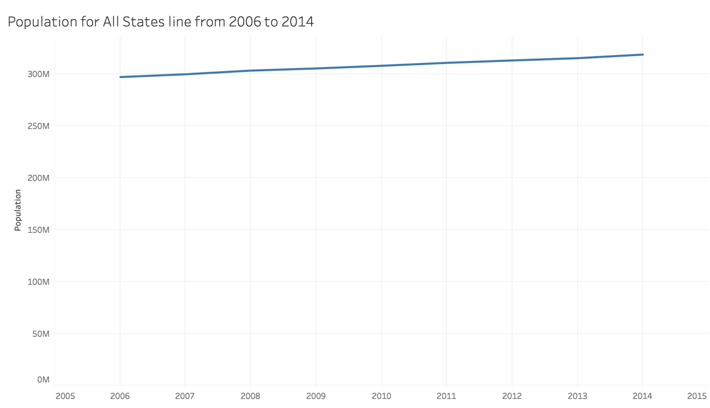
H3a: βYear > 0

The second set of hypothesis identifies trends over the 8 years. Because of the addictive nature of opioid drugs, we expect the deaths to opioid overdose to increase each year.

Descriptive Analysis





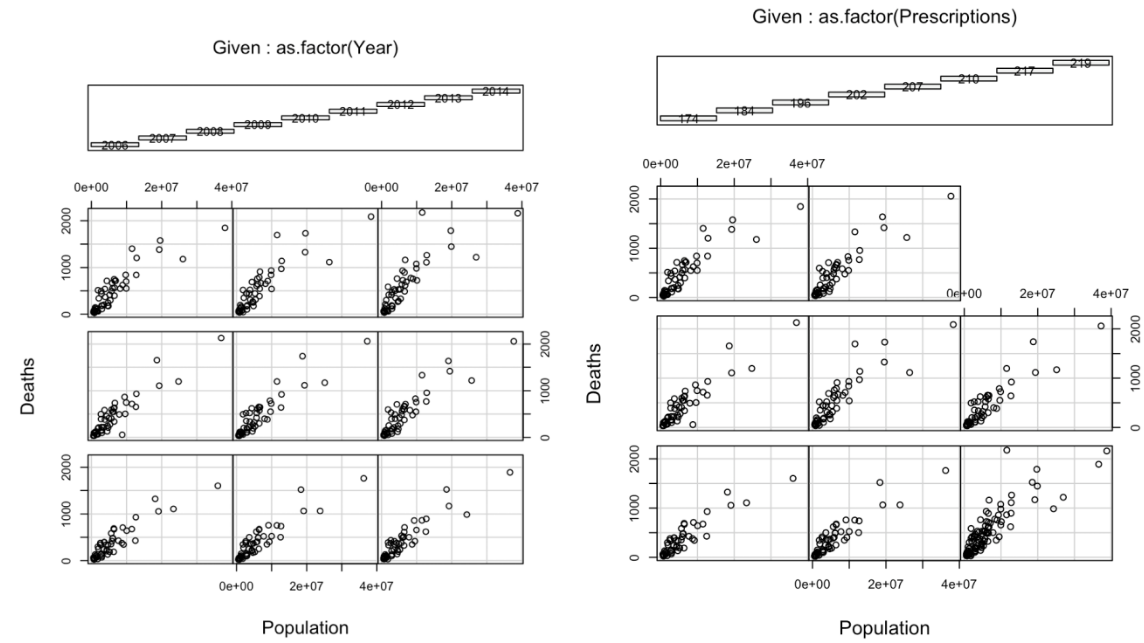


The number of opioid related deaths in the United States are steadily increasing at a relatively even rate across all years presented. The number of prescriptions increases steadily from 2006, peaking in 2011 at nearly 220 million. However, this trend reverses following 2011, with less than 200 prescriptions issued nationwide in 2014.

A decrease of 20 million prescriptions issued seems like a large amount, but it is less than a 10% decrease in total prescriptions. We think there may still be some correlation between prescriptions issued and opioid-related deaths, but given this preliminary analysis, it is clear there are other factors at play as well.

The population across all states is increasing by a nominal amount each year, and there is a positive linear correlation between population and number of deaths as would be expected. Our analysis will tell us what percentage of the increase in deaths is attributable to population increase and what percentage is a result of other factors.

The Corplot for correlation of all variables. [Overdose is All Drug Overdose Deaths]



The relationship between Deaths and Populations depend on different year or different prescription. [2006 - 2014]

In the Corplot, we can see that Deaths has a high correlation with Population and Prescriptions, and Population and Prescriptions also highly correlated. So, we should pay attention to multicollinearity when build models.

In the relationship between Deaths and Populations depend on different year or different prescriptions, the patterns between Population and Deaths doesn’t change too much with the Year. The prescriptions is also same.

Models:

Observing the exponential nature of the opioid death histogram, we try to normalize the data through a log function. All models use the logarithm of opioid deaths as a dependent variable and population and prescription of opioid drugs as predictor variables. In the temporal model, we add the Year as the time variable.

Basic Model

Deaths = βPopulation + βPrescriptions + ε

This model is the basic model that use linear regression to explain the opioids death based on population and prescriptions. (The output is in the Figure 1 in Appendix)

Exponential Model

Log(Deaths) = βPopulation + βPrescriptions + ε

Due to the Deaths is not normal distributed; we did some transformation and build the second model. This model uses a non-linear regression and takes the logarithm of opioid deaths as function of population and prescription. (The output is in the Figure 2 in Appendix)

Panel Model

Log(Deaths) = βPopulation + βPrescriptions + ε

This model explores the panel effects of various states and year. The pooling model confirms the exponential model; while the fixed model explores the differences between each state. Finally, the random model show the spread of differences between each state. (The output is in the Figure 4 in Appendix)

Mixed Model

Log(Deaths) = βPopulation + βPrescriptions + β(State) + ε

This model combines the fixed effects and random efforts into a single cohesive model. It has similar results as the previous model. (The output is in the Figure 5 in Appendix)

Temporal Model

Log(Deaths) = βYear + βPrescriptions + ε

In the the temporal model, we explore the trends of each state across the years, 2006-2014. We exam each state logarithm of opioids deaths as a function of year, population and prescription. We found 19 states that have significant variables in the model (see Appendix 1). For the lag analysis, we randomly picked four states with significant variables and four states with non-significant variables in the temporal analysis. We no presences of significant variables in the lagged model for all 8 states.

Quality Checks:

Model Comparison

1. Exponential Model

A log-linear model is used for the first and second hypothesis. In the multicollinearity test, the VIF for Population and Prescriptions are all less than 10. The exponential Model is Heteroscedastic and Non-linear on the chart, and also the QQ plot doesn't look normal, especially at high theoretical. In the shapiro.test, the p-value in residuals of exponential model is less than 0.05; we can reject H0 that residuals are normal. Therefore, this model is a bad model. (The output is in the Figure 2 in Appendix)

2. Panel Model

In the panel model, we built pooling, fixed and random models. In the multicollinearity test, the VIF for Population and Prescriptions are all less than 10. In the plmtest of pooling model, the p-value < 0.05, which means data, shows panel effect. In the pFtest between fixed model and pooling model, the p-value < 0.05, which means fixed model, is better than pooling model. In the phtest between fixed model and random model, the p-value is 0.08(>0.05), we cannot reject fixed model is worse than random model.With the fixed effect, the median of the residuals were just slightly below zero (-.005) and only a small drop in the adjusted R^2 from the R^2 ensued. (The output is in the Figure 4 in Appendix)

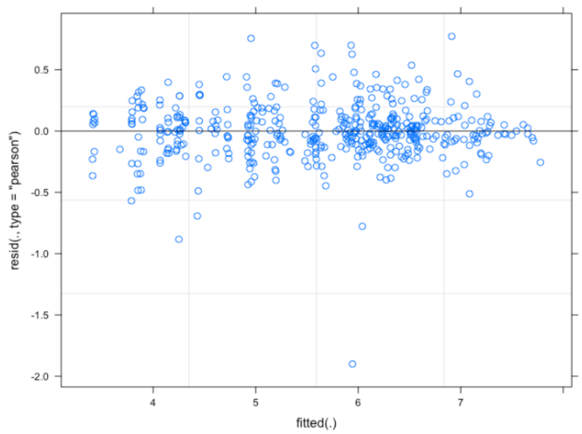
3. Mixed Model

In the multicollinearity test, the VIF for Population and Prescriptions are all less than 10. Based on a visual test, the residuals seem to be centered around zero which indicates that they are not biased. Also, the variance is pretty evenly spread with minor concentrations around 6 and 7. This indicates for the most part it is homoskedastic. This plot shows no real patterns so it shows a pretty random spread above and below zero; indicating that it is multivariate normality. (The output is in the Figure 5 in Appendix)

4. Temporal Model

In Temporal model, we build 51 OLS regression models for 51 states, and then we pick out Michigan, which are the only significant models in these 51. In the Bartlett test, the p-value < 0.05, we must reject H0 that the residual and fitted values have equal error variance, the residuals are heteroskedastic. (The output is in the Figure 6 in Appendix)

All in all, The mixed model is the best model, because it captured the exponential model and represented the panel effects. And it is homoskedastic and multivariate normal.



Outcome:

In the model we built, we can figure out that in the mixed model :

- When other variables are constant, 1 million population increases, 3.15% of Opioid death increase in a year.(not significant)
- When other variables are constant, 100 thousand prescription increases, and 1.47% of Opioid death increase in a year. (significant)
- Without any effect of Population and Prescription, we expect that 111.3 [exp(4.712)] Opioid death in a year.
- The random effect of Different States on Death is at 0.6737 standard deviation, which means different state will have almost 23% different opioid deaths without any other effects.

Recommendations:

Based on our analysis, the following states have the most opioid overdose deaths without any influence of population and prescriptions: West Virginia, Nevada, Utah, Maryland and Massachusetts. Our recommendation to their health professionals and state policymakers is to prioritize their allocation of resources for state-run drug programs for a greater proportion of funding for these states if the goal is to be effective in reducing the number of opioid-related deaths. In addition, we have identified both Michigan and Maine have good policies and treatment of opioids prescriptions as their opioid overdose death rates have decreased over the last 8 years. Other states should use their policies and treatment on opioid as a benchmark for example Maine's New Opioid Prescribing Law has this few interesting points that can be implemented in other state too:

- Doctors will be monitoring a database called the "Prescription Monitoring Program" when they write prescriptions for opioid pain relievers
- People taking opioids for chronic pain may be asked to come to their doctor's office more frequently for visits or other types of monitoring tools.
- Some people already taking opioids for chronic pain will be asked to work closely with their doctor to slowly decrease their dose of medication.

Michigan providers will now be required to check the state's prescription database before they prescribe painkillers and powerful medications, under legislation by Michigan Lt. Governor Brian Calley. The legislation, signed into law, will also put a limit on the number of opioids prescribed to patients for acute pain and establish a "bona fide physician-patient relationship before prescribing controlled substances." Further, the law specifies specific penalties for providers who fail to meet these requirements.

https://www.affirmhealth.com/blog/opioid-prescribing-guidelines-a-state-by-state-overview this is the link where people can go and compare state wise diffrence in laws regarding opioid and what different this state are doing.

Appendix A

Basic Model

```
> summary(basic)

Call:
lm(formula = Deaths ~ Population + Prescriptions, data = opioids)

Residuals:
    Min      1Q  Median      3Q     Max
-538.93  -97.57  -40.17  120.75 1331.82

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.130e+01  1.423e+01   5.010 7.87e-07 ***
Population    3.325e-05  4.082e-06   8.145 3.87e-15 ***
Prescriptions 3.724e-05  6.330e-06   5.883 7.95e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 195.7 on 445 degrees of freedom
Multiple R-squared:  0.795,     Adjusted R-squared:  0.7941
F-statistic: 862.8 on 2 and 445 DF,  p-value: < 2.2e-16
```

```
> shapiro.test(basic$residuals)

        Shapiro-Wilk normality test

data:  basic$residuals
W = 0.92782, p-value = 7.176e-14
```

```
> vif(basic)
   Population Prescriptions
     9.020768      9.020768
```

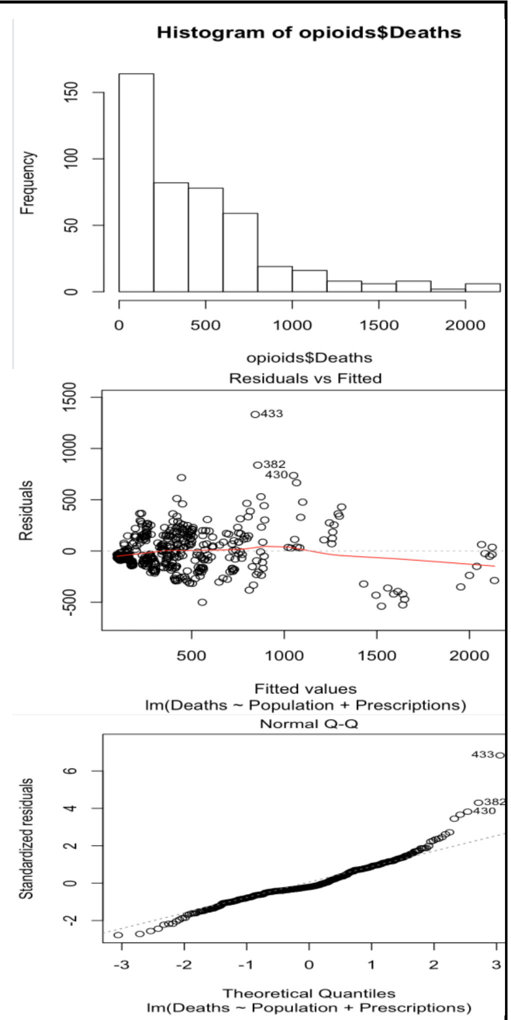


Figure 1. Basic Model

Exponential Model

```
> summary(exponential)

Call:
lm(formula = log(Deaths) ~ Population + Prescriptions, data = opioids

Residuals:
     Min       1Q   Median       3Q      Max 
-1.73191 -0.54206 -0.00285  0.57960  1.61226 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)    4.707e+00  4.950e-02  95.084   <2e-16 ***
Population    -2.415e-08  1.420e-08  -1.701   0.0897 .  
Prescriptions  2.240e-07  2.202e-08  10.174   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6808 on 445 degrees of freedom
Multiple R-squared:  0.5993,	Adjusted R-squared:  0.5975 
F-statistic: 332.7 on 2 and 445 DF,  p-value: < 2.2e-16

> shapiro.test(exponential$residuals)

	Shapiro-Wilk normality test

data:  exponential$residuals
W = 0.97724, p-value = 1.807e-06

> vif(exponential)
   Population Prescriptions 
     9.020768      9.020768 
```

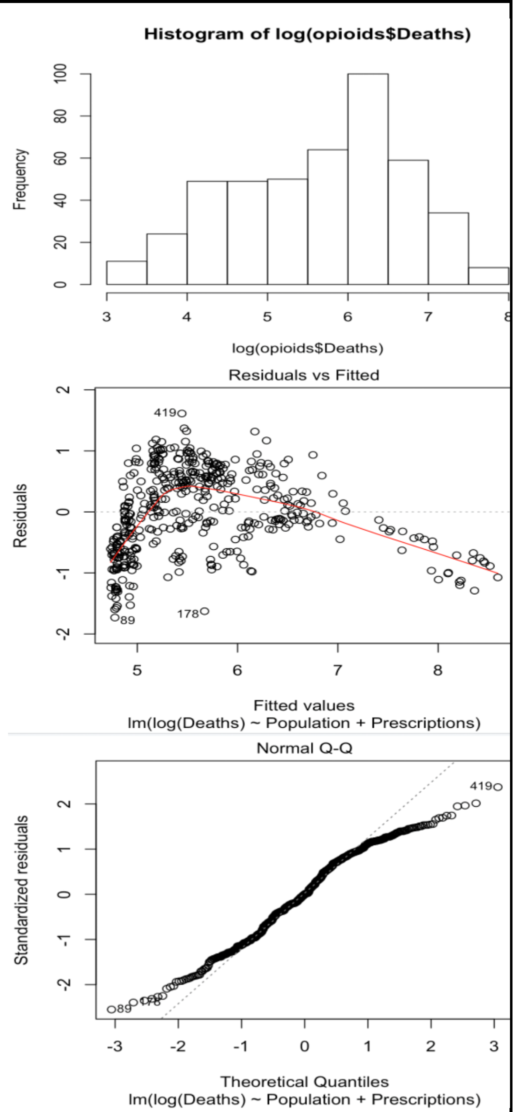


Figure 2. Exponential Model

Evaluate:

- From the chart, we can see the model is Heteroscedastic and Non-linear.
- QQ plot doesn’t look normal, especially at high theoretical.
- In the shapiro.test, since p-value <0.05, we can reject H0 that residuals are normal.

Interpretation:

- Population is a significant variable increasing Opioid Death.
- When other variables are constant, 1 million population increases, 2.42% of Opioid death decrease in a year.
- When other variables are constant, 100 thousand prescription increases, 2.24% of Opioid death increase in a year.
- Without any effect of Population and Prescription, we expect that 110.7 [exp(4.707)] Opioid death in a year.

Panel Model

```
> summary(pooled)
Pooling Model

Call:
plm(formula = log(Deaths) ~ Population + Prescriptions, data = panel
    model = "pooling")

Unbalanced Panel: n = 51, T = 1-9, N = 448

Residuals:
      Min.    1st Qu.     Median    3rd Qu.       Max.
-1.7319071 -0.5420639 -0.0028485  0.5795981  1.6122568

Coefficients:
                 Estimate  Std. Error t-value Pr(>|t|)
(Intercept)    4.7071e+00  4.9504e-02 95.0836  < 2e-16 ***
Population    -2.4152e-08  1.4199e-08 -1.7009  0.08965 .
Prescriptions  2.2402e-07  2.2019e-08 10.1743  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    514.72
Residual Sum of Squares: 206.26
R-Squared:      0.59928
Adj. R-Squared: 0.59748
F-statistic: 332.746 on 2 and 445 DF, p-value: < 2.22e-16
```

Pooling Model Interpretation:

- Prescription is a significant variable increasing Opioid Death.
- When other variables are constant, 1 million population increases, 2.42% of Opioid death decrease in a year. (not signicant) XXX
- When other variables are constant, 100 thousand prescription increases, 2.24% of Opioid death increase in a year.
- Without any effect of Population and Prescription, we expect that 110.7 [exp(4.707)] Opioid death in a year.

```
> summary(fixed)
Oneway (individual) effect Within Model

Call:
plm(formula = log(Deaths) ~ Population + Prescriptions, data = pane
    model = "within")

Unbalanced Panel: n = 51, T = 1-9, N = 448

Residuals:
       Min.      1st Qu.       Median      3rd Qu.        Max.
-1.89779420 -0.11772007 -0.00013198  0.11311760  0.75451648

Coefficients:
                Estimate Std. Error t-value Pr(>|t|)
Population    1.1717e-07 5.4115e-08  2.1652  0.03097 *
Prescriptions 1.1345e-07 3.6322e-08  3.1233  0.00192 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    25.461
Residual Sum of Squares: 23.762
R-Squared:      0.066742
Adj. R-Squared: -0.056117
F-statistic: 14.1243 on 2 and 395 DF, p-value: 1.1894e-06
```

Within Model Interpretation:

- Median is very close to 0 as we would expect
- When other variables are constant, 100 thousand population increases, 1.17% of Opioid death increase in a year.
- When other variables are constant, 100 thousand prescription increases, 1.13% of Opioid death increase in a year.
- Without any effect of Population and Prescription, we expect that (25.804, 250.153) Opioid death in a year. [(exp(SouthDakota),exp(WestVirginia))]. This would be a good table to show Top 5 or Top 10 states that have the highest Opioids Death without effect of Population and Prescriptions.

```
> summary(random)
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = log(Deaths) ~ Population + Prescriptions, data = p
    model = "random")

Unbalanced Panel: n = 51, T = 1-9, N = 448

Effects:
                  var std.dev share
idiosyncratic 0.06016 0.24527 0.124
individual    0.42486 0.65181 0.876
theta:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.6478  0.8755  0.8755  0.8746  0.8755  0.8755

Residuals:
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-1.87944 -0.10787 -0.00586  0.00218  0.12605  0.80843

Coefficients:
                Estimate Std. Error t-value  Pr(>|t|)
(Intercept)   4.7131e+00 1.2602e-01 37.4006 < 2.2e-16 ***
Population    3.0309e-08 2.3692e-08  1.2793    0.2015
Prescriptions 1.4786e-07 3.1753e-08  4.6564 4.254e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    32.983
Residual Sum of Squares: 27.064
R-Squared:      0.18149
Adj. R-Squared: 0.17781
F-statistic: 48.662 on 2 and 445 DF, p-value: < 2.22e-16
```

```
vif(pooled)
 Population Prescriptions
   9.020768      9.020768
vif(random)
 Population Prescriptions
    3.25512       3.25512
```

```
> plmtest(pooled)

	Lagrange Multiplier Test - (Honda) for unbalanced panels

data:  log(Deaths) ~ Population + Prescriptions
normal = 36.087, p-value < 2.2e-16
alternative hypothesis: significant effects

> pFtest(fixed, pooled)

	F test for individual effects

data:  log(Deaths) ~ Population + Prescriptions
F = 60.674, df1 = 50, df2 = 395, p-value < 2.2e-16
alternative hypothesis: significant effects

> phtest(fixed, random)

	Hausman Test

data:  log(Deaths) ~ Population + Prescriptions
chisq = 4.9587, df = 2, p-value = 0.0838
alternative hypothesis: one model is inconsistent
```

Figure 4. Panel Model

	Estimate	Std.	Error	t-value	Pr(>|t|)	
WestVirginia	250.153568	5.52	0.126235	43.7445	2.20E-16	***
Nevada	249.17962	5.52	0.1518	36.3517	2.20E-16	***
Utah	219.818235	5.39	0.153246	35.1905	2.20E-16	***
Maryland	205.371801	5.32	0.283722	18.7678	2.20E-16	***
Massachusetts	203.330554	5.31	0.318884	16.667	2.20E-16	***
Oklahoma	193.537056	5.27	0.202152	26.0471	2.20E-16	***
NewMexico	188.596158	5.24	0.125968	41.5947	2.20E-16	***
Kentucky	171.397942	5.14	0.234291	21.9556	2.20E-16	***
Washington	169.610223	5.13	0.326236	15.7355	2.20E-16	***
Oregon	151.024489	5.02	0.199754	25.1181	2.20E-16	***
Missouri	148.76353	5.00	0.293391	17.0501	2.20E-16	***
……						
Montana	50.5492951	3.92	0.094072	41.7014	2.20E-16	***
Mississippi	47.0538175	3.85	0.166985	23.0637	2.20E-16	***
NewYork	46.7861865	3.85	0.931638	4.1278	4.47E-05	***
District of Columbia	42.0373286	3.74	0.091664	40.7854	2.20E-16	***
Wyoming	38.7698545	3.66	0.090672	40.3392	2.20E-16	***
Nebraska	33.8087958	3.52	0.118463	29.7201	2.20E-16	***
NorthDakota	29.694861	3.39	0.247746	13.6873	2.20E-16	***
Florida	29.056029	3.37	0.891962	3.7773	0.0001829	***
SouthDakota	25.8042705	3.25	0.090458	35.9344	2.20E-16	***
Texas	7.89421332	2.07	1.183572	1.7457	0.0816452	.
California	2.43817059	0.89	1.772557	0.5028	0.6153827	

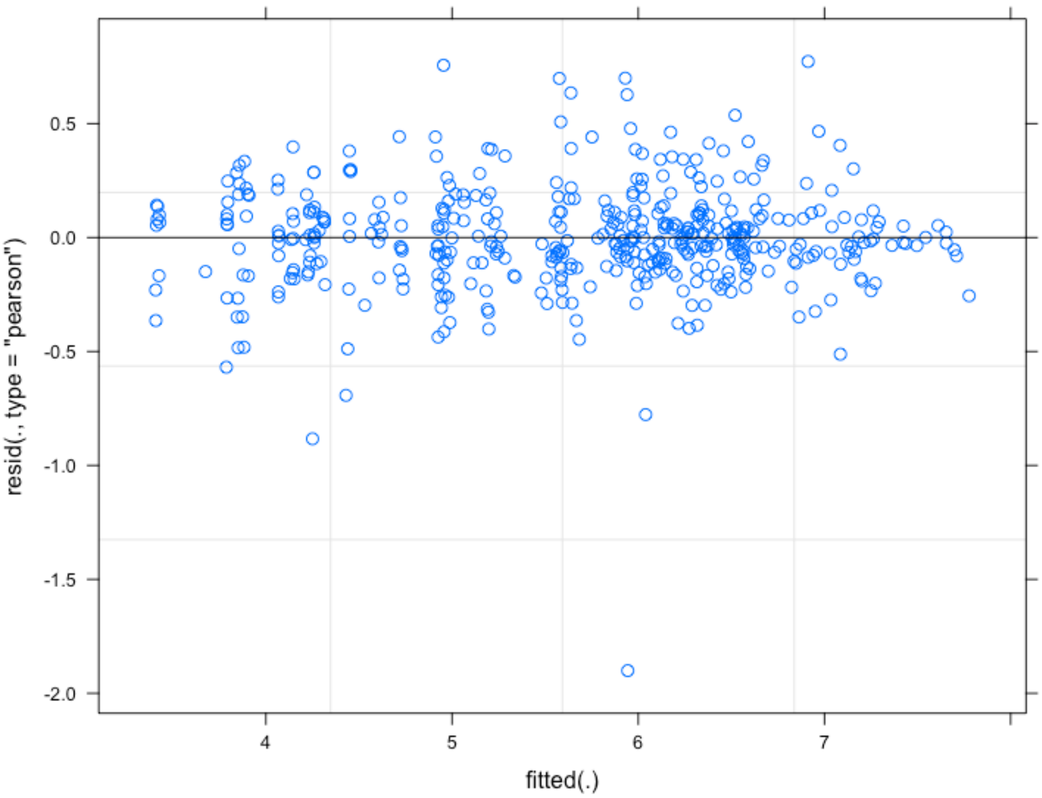
Mixed Model

```
> summary(mixed)
Linear mixed model fit by REML ['lmerMod']
Formula: log(Deaths) ~ Population + Prescriptions + (1 | State)
   Data: opioids

REML criterion at convergence: 294.1

Scaled residuals:
    Min      1Q  Median      3Q     Max
-7.7328 -0.4572 -0.0157  0.4363  3.1438

Random effects:
 Groups   Name        Variance Std.Dev.
 State    (Intercept) 0.4539   0.6737
 Residual             0.0604   0.2458
Number of obs: 448, groups:  State, 51

Fixed effects:
               Estimate Std. Error t value
(Intercept)   4.712e+00  1.291e-01  36.492
Population    3.146e-08  2.395e-08   1.313
Prescriptions 1.465e-07  3.187e-08   4.597

Correlation of Fixed Effects:
            (Intr) Popltn
Population  -0.158
Prescriptns -0.239 -0.827
fit warnings:
Some predictor variables are on very different scales: consider rescaling
```

Figure 5. Mixed Model

Fixed effect: Population, Prescription
Random effect: State to Intercept
Interpretation:

- Median is not exactly zero.
- When other variables are constant, 1 million population increases, 3.15% of Opioid death increase in a year.(not signicant)
- When other variables are constant, 100 thousand prescription increases, 1.47% of Opioid death increase in a year. (significant)
- Without any effect of Population and Prescription, we expect that 111.3 [exp(4.712)] Opioid death in a year.
- The random effect of Different States on Death is at 0.6737 standard deviation, which means different state will have almost 23% different opioid deaths without any other effects.

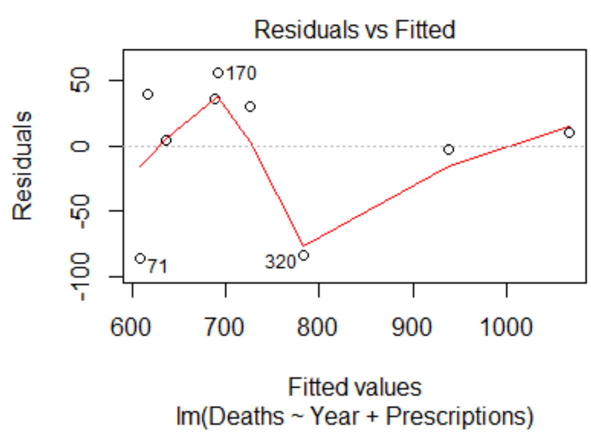


Based on a visual test, the residuals seem to be centered around zero which indicates that they are not biased. Also, the variance is pretty evenly spread with minor concentrations around 6 and 7. This indicates for the most part it is homoskedastic. This plot shows no real patterns so it shows a pretty random spread above and below zero; indicating that it is multivariate normality.

Temporal Model

```
[1] Michigan
51 Levels: Alabama Alaska Arizona Arkansas California ... Wyoming
> summary(tm1)

Call:
lm(formula = Deaths ~ Year + Prescriptions, data = temp)

Residuals:
    Min      1Q  Median      3Q     Max
-86.346  -3.174   9.558  35.717  55.145

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.163e+05  3.854e+04  -5.612  0.00137 **
Year           1.092e+02  1.952e+01   5.594  0.00139 **
Prescriptions -2.671e-04  8.258e-05  -3.235  0.01780 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 59.59 on 6 degrees of freedom
Multiple R-squared:  0.9012,    Adjusted R-squared:  0.8682
F-statistic: 27.35 on 2 and 6 DF,  p-value: 0.0009655
```

```
        Bartlett test of homogeneity of
variances

data:  list(tm1$residuals, tm1$fitted.values)
Bartlett's K-squared = 7.7709, df = 1, p-value
= 0.005309

> vif(tm1)
         Year Prescriptions
     6.440052      6.440052
```



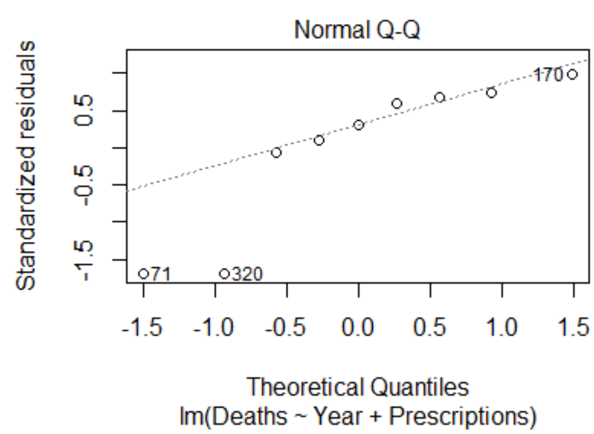


Figure 6. Temporal Model

Temporal Analysis by each state

Interpretation: (Blue Text) States indicates strong significant results for the independent variables (Florida, Maine, Michigan). These are the states that need some attention when it comes to trying to reduce opioid impact. There could possibly be programs, funding, or government intervention implemented to help identify

<table>
<tr>
<td>
<pre>
> ST[10]
[1] Florida
51 Levels: Alabama Alaska Arizona Arkansas California ... Wyoming
> summary(tm1)

Call:
lm(formula = Deaths ~ Year + Prescriptions, data = temp)

Residuals:
    Min   1Q Median   3Q   Max
-58.949 -35.228 -1.301 12.158 80.102

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.011e+04 1.470e+04 -1.367 0.220486
Year   9.421e+00 7.260e+00   1.298 0.242076
Prescriptions 1.759e-04 2.377e-05   7.401 0.000313 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 53.76 on 6 degrees of freedom
Multiple R-squared: 0.9026,	Adjusted R-squared: 0.8702
F-statistic: 27.81 on 2 and 6 DF, p-value: 0.0009235
</pre>
</td>
<td>
Interpretation:

When other variables are constant, 1 Year increase, 942% of Opioid death increase in a year
When other variables are constant, 100 prescription increases, 1.76% of Opioid death increase in a year. (significant)
Without any effect of Year and Prescription, we expect that 0 [exp(-20110)] Opioid death in a year.

</td>
</tr>
<tr>
<td>
<pre>
[1] Maine
51 Levels: Alabama Alaska Arizona Arkansas California ... Wyoming
> summary(tm1)

Call:
lm(formula = Deaths ~ Year + Prescriptions, data = temp)

Residuals:
     Min   1Q   Median   3Q   Max
-12.4344 -8.4432 -0.3977   9.5347 11.8427

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.239e+04 2.833e+03 -4.373 0.00470 **
Year   6.422e+00 1.412e+00   4.549 0.00390 **
Prescriptions -3.465e-04 7.646e-05 -4.532 0.00397 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.91 on 6 degrees of freedom
Multiple R-squared: 0.8657,	Adjusted R-squared: 0.821
F-statistic: 19.34 on 2 and 6 DF, p-value: 0.002421
</pre>
</td>
<td>
Interpretation:

When other variables are constant, 1 Year increase, 642.2% of Opioid death increase in a year
When other variables are constant, 100 prescription increases, 3.47% of Opioid death decrease in a year.
Without any effect of Year and Prescription, we expect that 0 [exp(-12390)] Opioid death in a year.

</td>
</tr>
<tr>
<td>
<pre>
[1] Michigan
51 Levels: Alabama Alaska Arizona Arkansas California ... Wyoming
> summary(tm1)

Call:
lm(formula = Deaths ~ Year + Prescriptions, data = temp)

Residuals:
    Min   1Q Median   3Q   Max
-86.346 -3.174   9.558 35.717 55.145

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.163e+05 3.854e+04 -5.612 0.00137 **
Year   1.092e+02 1.952e+01   5.594 0.00139 **
Prescriptions -2.671e-04 8.258e-05 -3.235 0.01780 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 59.59 on 6 degrees of freedom
Multiple R-squared: 0.9012,	Adjusted R-squared: 0.8682
F-statistic: 27.35 on 2 and 6 DF, p-value: 0.0009655
</pre>
</td>
<td>
Interpretation:

When other variables are constant, 1 Year increase, 10900% of Opioid death increase in a year
When other variables are constant, 100 prescription increases, 2.67% of Opioid death decrease in a year. (significant)
Without any effect of Year and Prescription, we expect that 0 [exp(-216300)] Opioid death in a year.

</td>
</tr>
</table>

```
[1] West Virginia
51 Levels: Alabama Alaska Arizona Arkansas California ... Wyoming
> summary(tm1)

Call:
lm(formula = Deaths ~ Year + Prescriptions, data = temp)

Residuals:
     Min       1Q   Median       3Q      Max 
-184.802  -22.478    3.504   41.735  129.434 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)  
(Intercept)   -7.350e+04  2.726e+04  -2.697   0.0357 *
Year           3.694e+01  1.347e+01   2.742   0.0336 *
Prescriptions -1.198e-04  2.635e-04  -0.455   0.6654  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 100.7 on 6 degrees of freedom
Multiple R-squared:  0.5996,	Adjusted R-squared:  0.4662 
F-statistic: 4.493 on 2 and 6 DF,  p-value: 0.06417
```

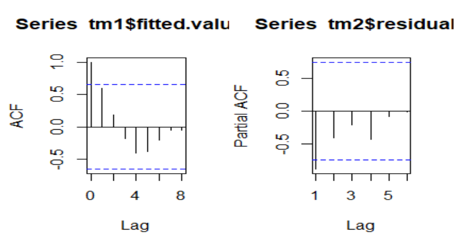
States with not significant variables...

Arizona	Arkansas	California	Connecticut
Georgia	Hawaii	Idaho	Indiana
Iowa	Kansas	Kentucky	Massachusetts
Minnesota	Mississippi	Montana	Nebraska
New Jersey	New Mexico	Oklahoma	Oregon
Rhode Island	South Carolina	South Dakota[	Texas
Utah	Vermont	Virginia	West Virginia
Wisconsin	Wyoming	North Dakota	

Lag Analysis

Chose 4 states with not significant variables...lag analysis still did not reflect any significant variables.

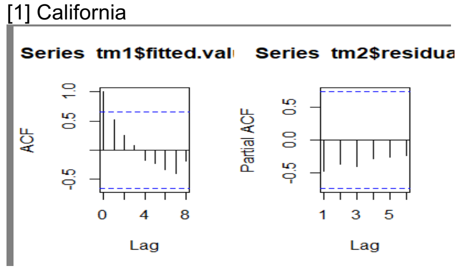
[1] Arizona



Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + lag1Deaths,
    data = temp)

Residuals:
   100   150   199   249   299   349   399
-1.662 6.594 -6.318 -2.362 8.617 -7.875 3.006

[1] California



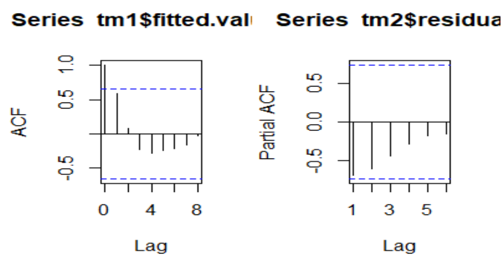
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + lag1Deaths +
    lag2Deaths, data = temp)

Residuals:
   103   153   202   252   302   352   402
-93.804 147.091 -15.927 -22.371 -25.270   8.730   1.551

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.110e+04  4.548e+04   0.244    0.830
Year          -5.452e+00  2.332e+01  -0.234    0.837
Population     2.507e-04  2.609e-03   0.096    0.932
Prescriptions -4.842e-04  9.934e-04  -0.487    0.674
lag1Deaths    -9.137e-02  2.281e-01  -0.401    0.727

Residual standard error: 10.89 on 2 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.6408,    Adjusted R-squared:  -0.07753
F-statistic: 0.8921 on 4 and 2 DF,  p-value: 0.5893
```

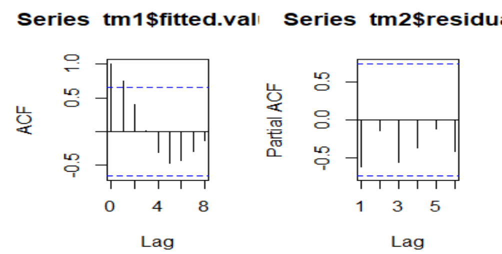
```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.021e+05  1.363e+06  -0.148    0.906
Year           1.105e+02  7.137e+02   0.155    0.902
Population    -2.973e-04  1.942e-03  -0.153    0.903
Prescriptions -4.263e-04  3.977e-04  -1.072    0.478
lag1Deaths     3.446e-01  6.489e-01   0.531    0.689
lag2Deaths     7.039e-01  1.044e+00   0.674    0.622

Residual standard error: 178.6 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.6224,    Adjusted R-squared:  -1.265
F-statistic: 0.3297 on 5 and 1 DF,  p-value: 0.8579
```

[1] New Mexico




```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths + 
    lag2Deaths, data = temp)

Residuals:
    130     179     229     279     329     379     429 
-0.6450  0.2588  2.0503 -5.4263  6.1288 -1.6036 -0.7629 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -3.128e+05  5.466e+04  -5.723    0.110
Year           1.666e+02  2.940e+01   5.666    0.111
Population    -1.076e-02  2.395e-03  -4.492    0.139
Prescriptions  4.599e-04  3.271e-04   1.406    0.394
lag1Deaths    -4.695e-01  1.811e-01  -2.592    0.234
lag2Deaths    -1.109e+00  2.483e-01  -4.465    0.140

Residual standard error: 8.651 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.998,     Adjusted R-squared:  0.9882
F-statistic: 101.4 on 5 and 1 DF,  p-value: 0.07523
```

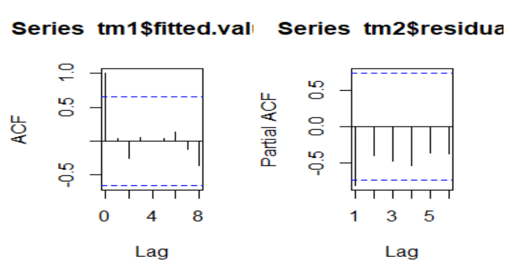
[1] West Virginia



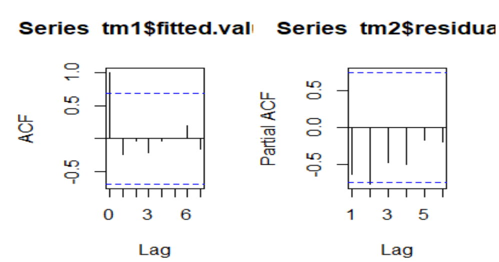

```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths + 
    lag2Deaths, data = temp)

Residuals:
    146      195      245      295      345      395      446 
 50.300 -104.815   37.722  -13.397   83.861  -56.141    2.471 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.114e+05  3.975e+05   0.280    0.826
Year          -6.962e+01  2.176e+02  -0.320    0.803
Population     1.850e-02  2.881e-02   0.642    0.637
Prescriptions -1.901e-03  2.665e-03  -0.713    0.606
lag1Deaths    -1.959e-01  6.535e-01  -0.300    0.815
lag2Deaths    -6.188e-01  6.896e-01  -0.897    0.534

Residual standard error: 159.1 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.7846,    Adjusted R-squared:  -0.2924
F-statistic: 0.7285 on 5 and 1 DF,  p-value: 0.7059
```

Chose 4 states with significant variables...lag analysis changed to model to have no significant variables.

[1] Alabama



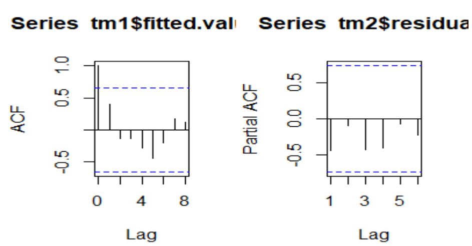
[1] Alaska



```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths + 
    lag2Deaths, data = temp)

Residuals:
     99     149     198     248     298     348     398 
  2.302  -5.519   4.248  -4.606  11.223 -11.835   4.187 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.405e+05  1.214e+05   1.158    0.454
Year          -8.331e+01  6.884e+01  -1.210    0.440
Population     6.029e-03  3.689e-03   1.634    0.350
Prescriptions -2.165e-04  7.190e-05  -3.011    0.204
lag1Deaths     1.557e-01  1.315e+00   0.118    0.925
lag2Deaths    -1.829e+00  1.207e+00  -1.515    0.371

Residual standard error: 18.94 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.9579,    Adjusted R-squared:  0.7474 
F-statistic: 4.551 on 5 and 1 DF,  p-value: 0.341
```

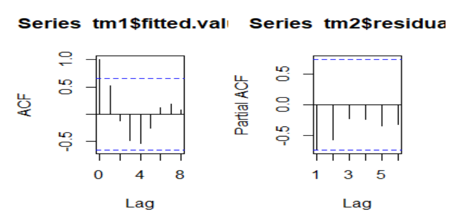
```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths, 
    data = temp)

Residuals:
   100    150    199    249    299    349    399 
-1.662  6.594 -6.318 -2.362  8.617 -7.875  3.006 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.110e+04  4.548e+04   0.244    0.830
Year          -5.452e+00  2.332e+01  -0.234    0.837
Population     2.507e-04  2.609e-03   0.096    0.932
Prescriptions -4.842e-04  9.934e-04  -0.487    0.674
lag1Deaths    -9.137e-02  2.281e-01  -0.401    0.727

Residual standard error: 10.89 on 2 degrees of freedom
  (3 observations deleted due to missingness)
Multiple R-squared:  0.6408,    Adjusted R-squared:  -0.07753 
F-statistic: 0.8921 on 4 and 2 DF,  p-value: 0.5893
```

[1] Nevada




```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths + 
    lag2Deaths, data = temp)

Residuals:
    127     176     226     276     326     376     426 
-21.037  21.763  -2.747  27.129 -21.124 -10.745   6.761 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.146e+05  4.407e+05   0.260    0.838
Year          -5.910e+01  2.287e+02  -0.258    0.839
Population     1.599e-03  6.806e-03   0.235    0.853
Prescriptions -8.011e-05  4.105e-04  -0.195    0.877
lag1Deaths     1.586e+00  1.776e+00   0.893    0.536
lag2Deaths    -3.567e-01  1.619e+00  -0.220    0.862

Residual standard error: 47.61 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.8175,    Adjusted R-squared:  -0.09505 
F-statistic: 0.8958 on 5 and 1 DF,  p-value: 0.6609
```

[1] Florida




```
Call:
lm(formula = Deaths ~ Year + Population + Prescriptions + 
lag1Deaths + 
    lag2Deaths, data = temp)

Residuals:
    108     157     207     257     307     357     407 
 -6.254   7.697   8.231 -17.080  15.483 -13.613   5.536 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.473e+05  3.152e+05   1.102    0.469
Year          -1.844e+02  1.648e+02  -1.119    0.464
Population     1.058e-03  7.749e-04   1.365    0.402
Prescriptions  3.115e-04  7.273e-05   4.283    0.146
lag1Deaths    -3.738e-02  1.913e-01  -0.195    0.877
lag2Deaths     1.015e-01  3.249e-01   0.312    0.807

Residual standard error: 30.22 on 1 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.9935,    Adjusted R-squared:  0.961 
F-statistic: 30.54 on 5 and 1 DF,  p-value: 0.1365
```