**MIS581 Final Project Components**

Nitin Tambe

Colorado State University Global

MIS581: Capstone - Business Intelligence and Data Analytics

Dr. Justin Bateh

04/10/2022

Table of Contents

[**INTRODUCTION** 3](#_Toc99279429)

[**OBJECTIVES** 3](#_Toc99279430)

[**OVERVIEW OF STUDY** 3](#_Toc99279431)

[**RESEARCH HYPOTHESIS** 4](#_Toc99279432)

[**LITERATURE REVIEW** 5](#_Toc99279433)

[**RESEARCH DESIGN** 6](#_Toc99279434)

[**Methodology** 6](#_Toc99279435)

[**Methods** 6](#_Toc99279436)

[**Limitations** 7](#_Toc99279437)

[**Ethical considerations** 8](#_Toc99279438)

[**FINDINGS** 9](#_Toc99279439)

[**Tableau** 9](#_Toc99279440)

[**R Programming** 11](#_Toc99279441)

[**SAS Enterprise Miner** 13](#_Toc99279442)

[**CONCLUSION** 16](#_Toc99279443)

[**RECOMMENDATIONS** 17](#_Toc99279444)

[**REFERENCES** 18](#_Toc99279445)

# **INTRODUCTION**

Opioid is mainly used as a pain killer as it works on specific receptors of brain or the nervous system, to reduce the pain intensity. It is prescribed by doctors to mainly treat moderate to severe pain, but it has lot of side effects and risks of addiction. As per CDC - Centers for Disease Control and Prevention, almost 130 Americans die every day from opioid overdose, this only means the Opioid epidemic is far from over. To put this into another perspective since 1999, nearly 841, 000 people died due to drug overdose and over 70% of those drug overdose was related to Opioid. The impact of this epidemic is not restricted to adult population only, it is also impacting pediatric patients. Between 1999 and 2016, approximately 9,000 American children died due to the misuse of prescription and illicit opioids (Gaither et al. 2018). CDC also confirms children aged 12 to 17 misused opioids and that estimated number is 769,000.

After merging with pharmacy benefit manager Express-scripts, CIGNA along with Express scripts have done multiple enhancements to the Opioid management program to address various gaps and improve patient safety. In order to help prevent the progression of opioid overdose. Express Scripts and Cigna combined together continue to provide additional education and support patient to ensure proactive care and safe disposal of opioid.

# **OBJECTIVES**

This data would be used to find any data trends related to opioid which can help CIGNA and express-scripts combination to put some measures in the point of care and point of sale to help tack the opioid epidemic.

# **OVERVIEW OF STUDY**

A hypothesis is based on a previous theory which can be tested and proved correct or otherwise. It is a statement which one should be able to test with the help of scientific research. For testing a hypothesis which is proving relationship between two or more variables one need to write the hypothesis before starting the experiment or starting data collection. It is a prediction about what the research is going to discover or in other words a prediction what an experiment or research is going to prove. It is often based on a theory or a knowledge which is tested. The hypothesis proposes the relationship between the variables i.e., independent and dependent variable. A dependent variable is observed throughout the experiment. Independent variables are controlled within the experiments and changed within the experiment.

In this paper the focus is to find the relationship between the gender, race and time with the opioid related mortality. The data gathered from the WONDER website provided by CDC, has details on the gender and race. This data is for the deaths related to opioid related ICD codes from 1990 to 2020.

# **RESEARCH HYPOTHESIS**

In order to find the relationship between the mortalities and the other variables in the data there are few questions

* How mortality has changed over a period of time
* How gender affects the mortality rate
* How race affects mortality rate

These are some of the important questions which should be answered in order to get the better understanding of the data. It will also help in coming up with the measures which can help pharmacy benefit manager company to put extra check around dispensing opioid drugs.

* How mortality has changed over a period of time
  + H0 – Mortality counts are same over a period of time. The numbers are not increasing for opioid related deaths
  + Ha – Mortality counts are NOT same for opioid related deaths over a period of time
* How gender affects the mortality rate
  + H0 – Mortality count is same for both genders
  + Ha – Mortality count is NOT same for both genders
* How race affects mortality rates
  + H0 – Mortality counts are same across all races
  + Ha – Mortality counts are NOT same across all races

# **LITERATURE REVIEW**

The risk and danger of Opioid crisis is not hidden. The Food and Drug Administration (FDA) has announced a comprehensive review of the agency's opioids policies. CDC has also confirmed to fundamentally re-examine the risk-benefit paradigm for opioids use and to ensure that it has consider wider public health (Furlow, 2016). As per the (McCarthy, 2016), President Barack Obama had proposed new funding to increase access to treatment for misuse of prescribed opioids which was estimated to be around $1.1bn. In the 2022 state of the union addresses, current president Biden also emphasized on the Opioid crisis.

Historically, it was observed that men prevalent in the opioid abuse, in the recent times the trend seems to be changing. A more percentage of adolescence girls are now abusing it at a higher rate than their male counterparts. There is increase in the prevalence of fatal overdose of this drug in women than the men. Women are using it for stress management. ( Chartoff & McHugh 2016).

# **RESEARCH DESIGN**

## **Methodology**

The tools like Tableau, R, PostgreSQL and SAS are being used in this exercise to run various analytical techniques. PostgreSQL is used to load the data and use it load in to the Tableau for various data dashboards

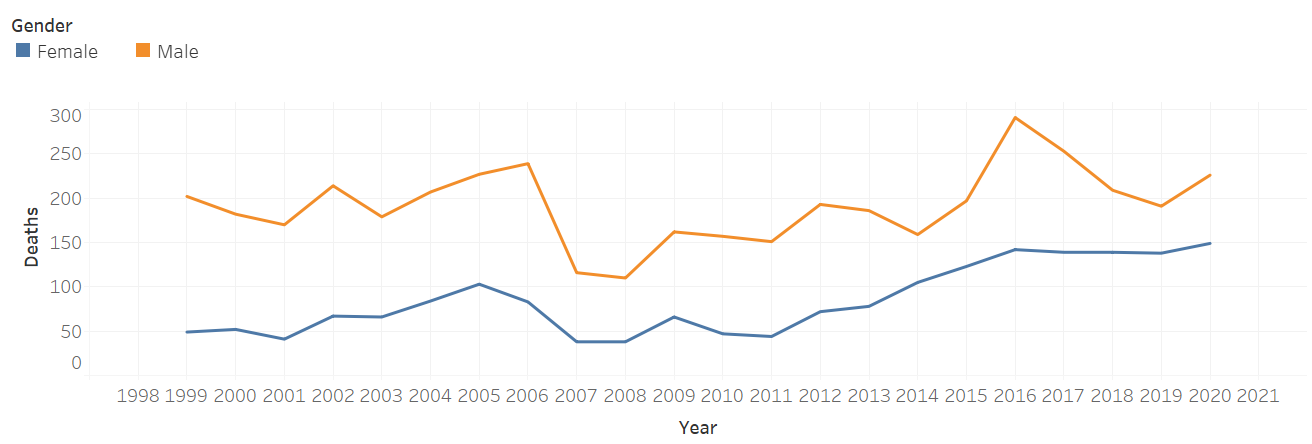
R has in built functions which can help create the correlation matrix for the given dataset. Regression can be developed in R which can be used to prove or not prove the hypothesis. R also has lot of plot functions which will be used to plot the various data presentations.

## **Methods**

Tableau provides feature which can be easily used for the data visualization. Data visualization can help in identifying the trends in the data which can be further investigated. Figure 1 and 2 shows the Tableau dashboard where the data is shown over a period of timeframe comparing the death counts by race and gender. This clearly shows the trends in the data.

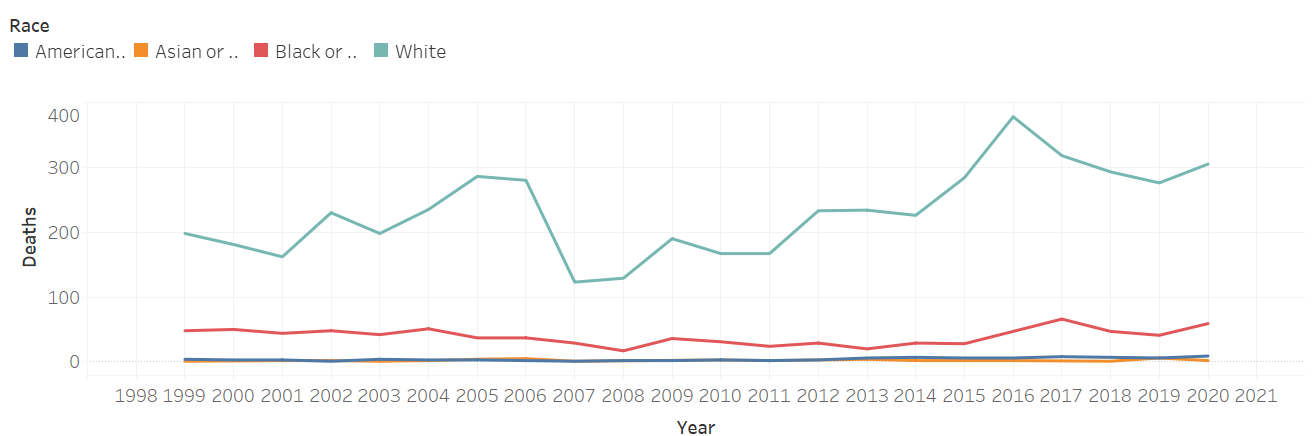
**Figure 1**

*Mortality by Gender*



**Figure 2**

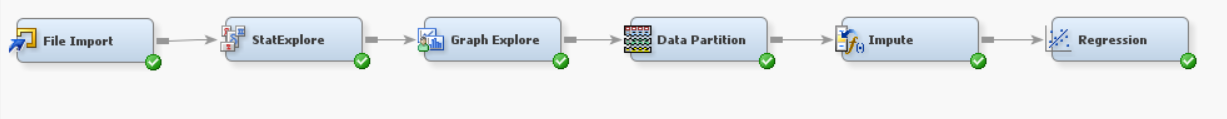
*Mortality by Race*



SAS studio has in built wizards which can be used to create the data models and visualizations for the data. Regression analysis can help identify the p value which can help prove or not prove the hypothesis. Filters are used to remove null values, SatExplore helps in understanding more about the data.

**Figure 3**

*SAS Enterprise Miner - Model*



## **Limitations**

The data used for the analysis is available on the CDC website which restricts the information at specific level. There is no information on the individual’s health and other medications these individuals had. It can be useful to understand how these other factors impacted or contributed for the mortality rates. Certainly, the data available can help us reach some conclusions but those would need to be studied thoroughly to ensure the steps taken for the eradication of this epidemic are concrete.

## **Ethical considerations**

The data used for this exercise is available on the CMS – WONDER. The data can be unloaded and used for specific purpose only. The data used does not include any individual’s private information which can identify specific person. While creating visualizations designer should ensure the visualizations are easy to understand and simple. In visualizations it is easy to manipulate the viewers by incorrect representation of data. It is very important to follow the ethical principles while representing the data. The designers should avoid viewers with misleading graphs or charts.

The opioid epidemic has been in focus for a while but even after decades of spending money and awareness on the topic. CDC has not been able to contained it. Using the data analytics models a trigger can be send to dispensing pharmacy to double check the prescriptions and with the patients about the need of the drug. Also, the PBM’s like express scripts can follow up with the patients who have taken smaller dosage of the drug, to confirm that have stopped the usage whenever they do not have any need of the drug. Taking these medications when they are needed can cause addiction which is one of the primary reasons for the overdose and abuse.

# **FINDINGS**

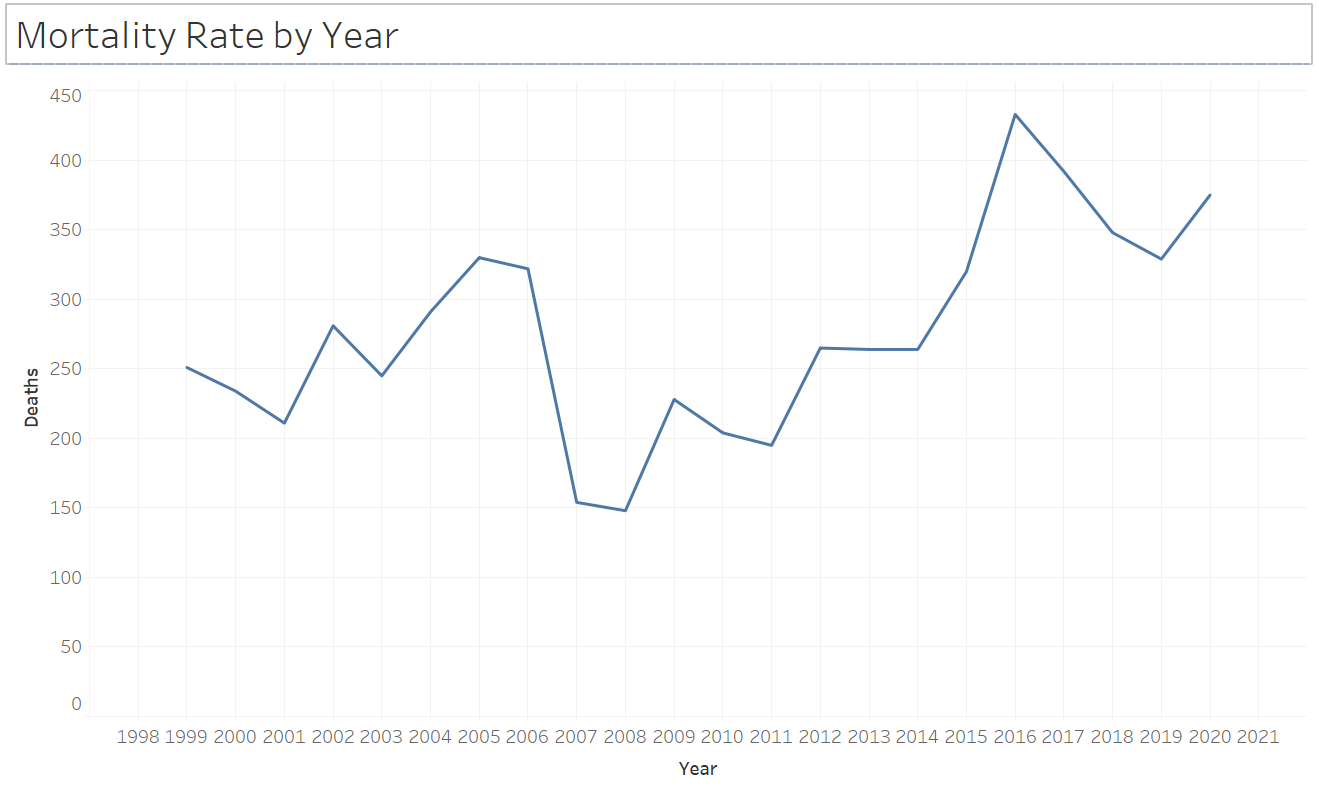
For this paper various tools were used to analyses the dataset. Tableau was used to create data visualization charts, R and SAS Enterprise Miner were used to create the regression models to verify the hypothesis.

## **Tableau**

Tableau is very useful in creating meaningful data visualization with minimum efforts. It takes various data types as input such as excel or can easily connect to databases. Below charts show various visualizations which helps in identifying the trends in the data.

**Figure 4**

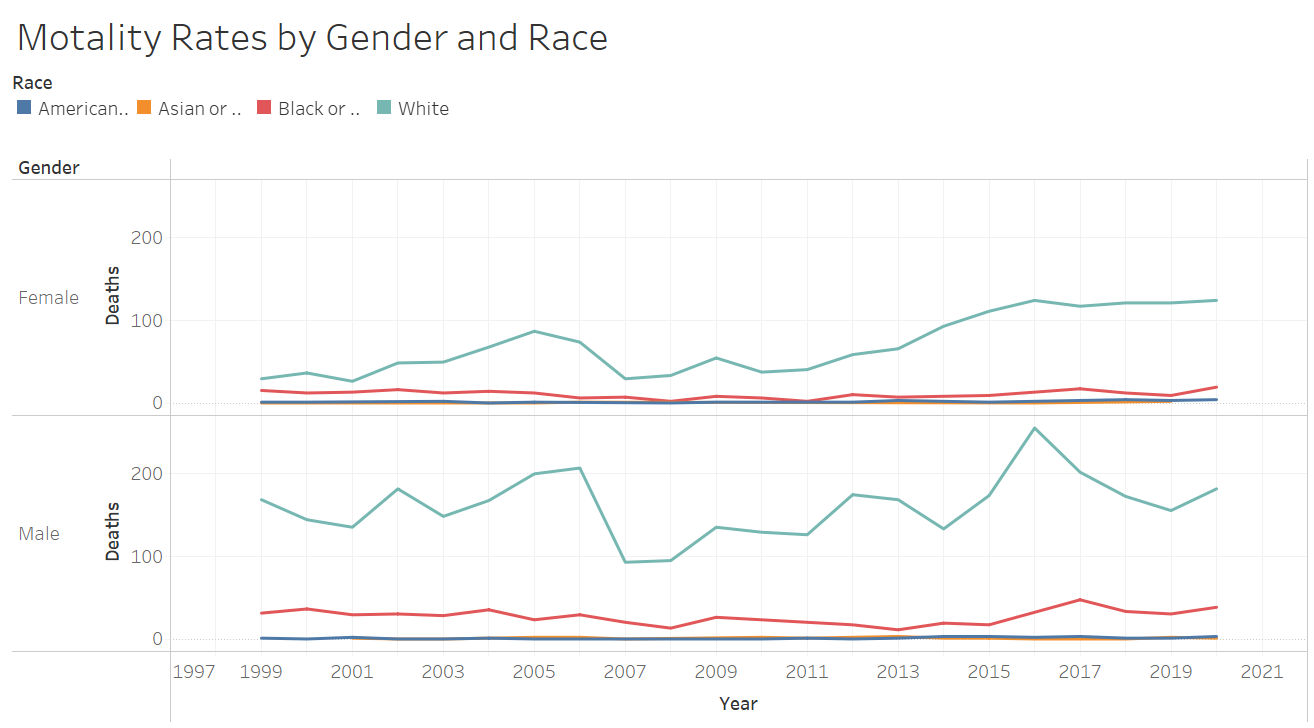
*Mortality rates by timeframe - year*



Notes – The chart clearly shows the increase in the mortality rate caused by Opioid overdose.

**Figure 5**

*Mortality rates comparison chart – Race & Gender*



Notes – The chart shows the difference in the mortality rates by Race and Gender. Irrespective of the gender Whites are showing consistent increase in mortality rates due to opioid overdose. In the past 10 years the female mortality rate has been increasing.

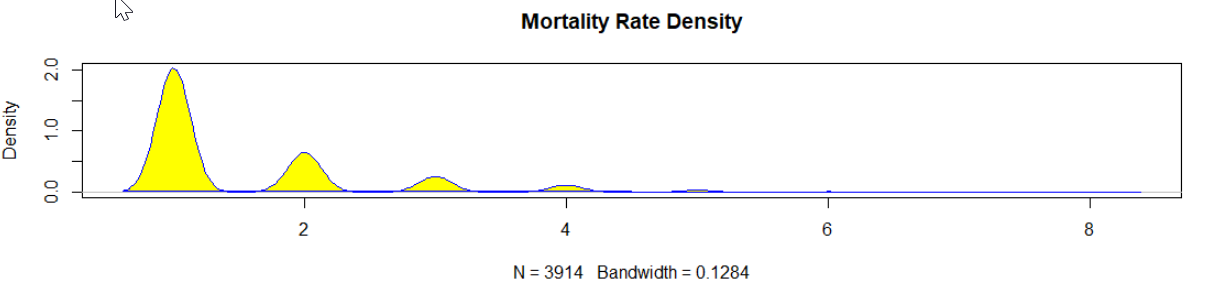
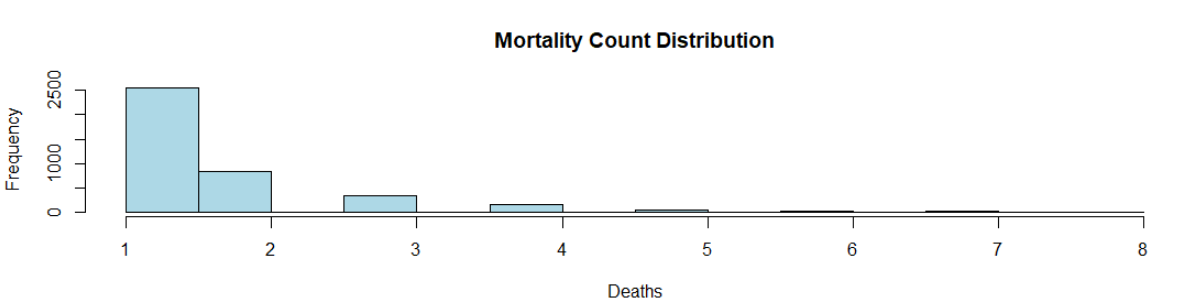
The tableau charts clearly show the mortality rate is increasing, while the numbers for other population seems to be same or have not increased by significance amount, the white population seems to be affected most by opioid overdose. The female population has severe affect in past decade. The numbers for the while female population are increasing consistently.

## **R Programming**

R is very quick to learn and create various data analytics models. In this paper, R was used to create the regression and to create few visualizations.

**Figure 6**

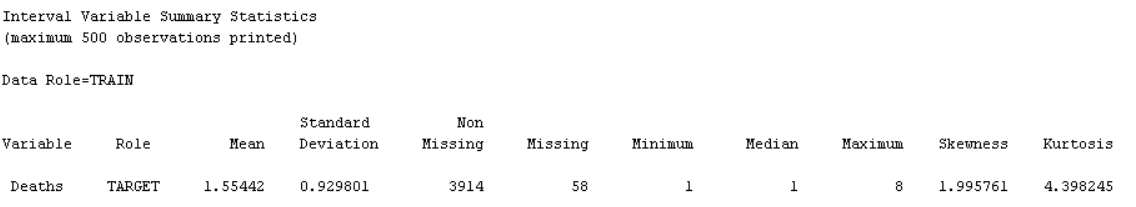
*Distribution and Density for Mortality by year*



Notes – Above charts shows the distribution of the mortality counts.

**Figure 7**

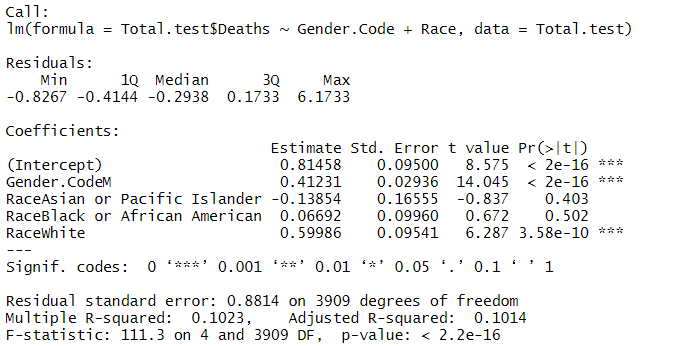
*Summary Statistics*



Note - The skewness is positive which means the data is right skewed. Kurtosis shows the distribution of the data. The value of 4.4 shows the heavily tailed data.

**Figure 8**

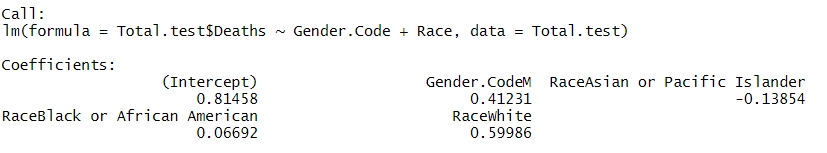
*Linear regression model - I*



Notes – The dataset created for this analysis had only 2 columns for Race and Gender each. As it is evident from the P value the Male gender and whites have a significant impact on the mortality rates caused by opioid overdose.

**Figure 9**

*Linear Regression – output-II*



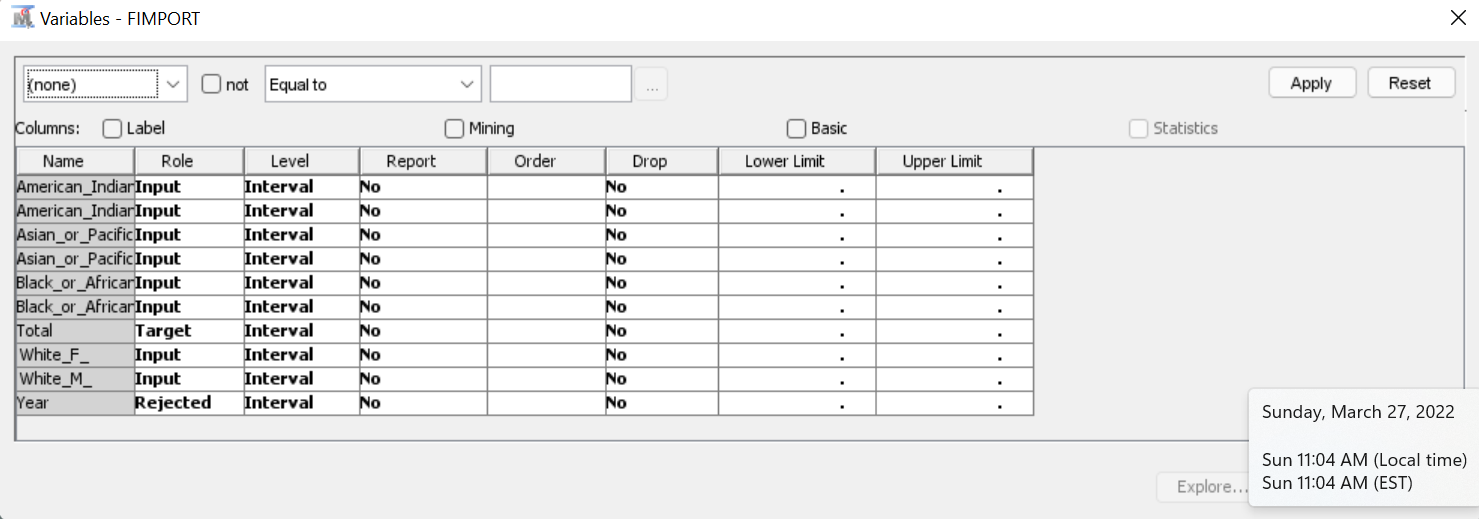
Notes – Similar observations as shown in the figure 8. White and Gender code M has a high coefficient.

## **SAS Enterprise Miner**

SAS Enterprise miner offers various tools with the user-friendly wizard to run the data analytics models.

**Figure 10**

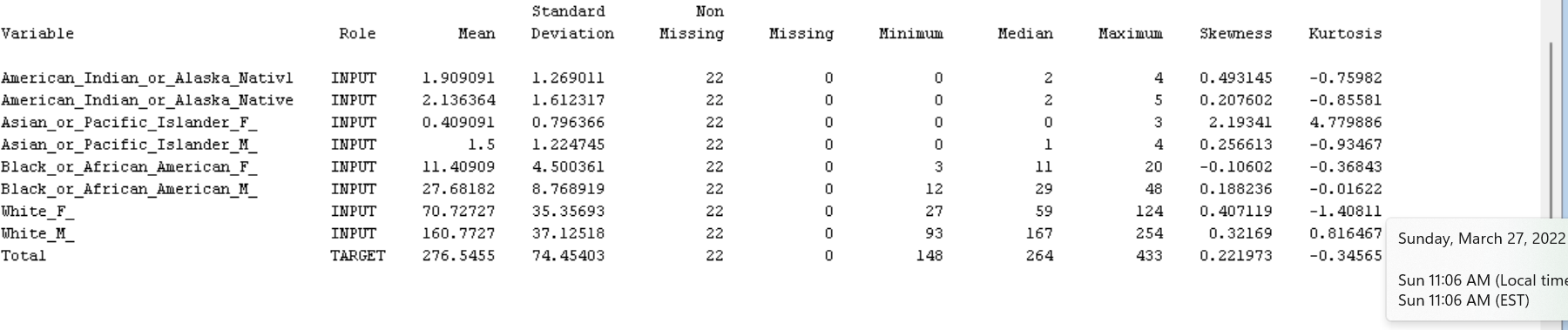
*Data Import - Properties*



Notes – Data imported from the excel CSV format. Columns were created for each variable in the file. Mortality count was defined as a Target variable whereas the race and gender were input variables.

**Figure 11**

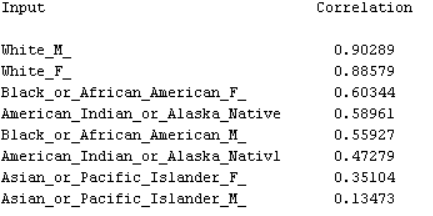
*SAS Enterprise Miner - Summary Stats*



Notes – Summary stats from the SAS enterprise miner shows same observations as the R output.

**Figure 12**

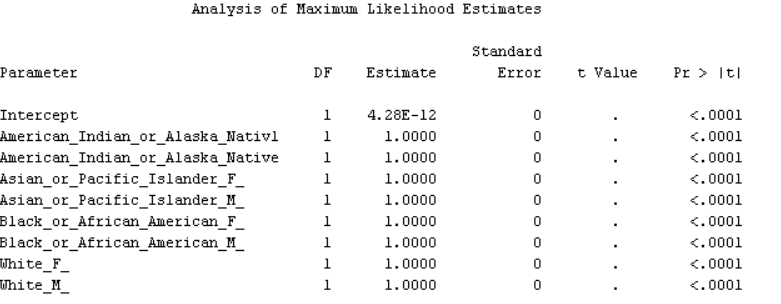
*SAS Enterprise Miner – Correlation Analysis*



Notes – Race and gender combination for the white show strong relationship with the mortality counts due to opioid.

**Figure 13**

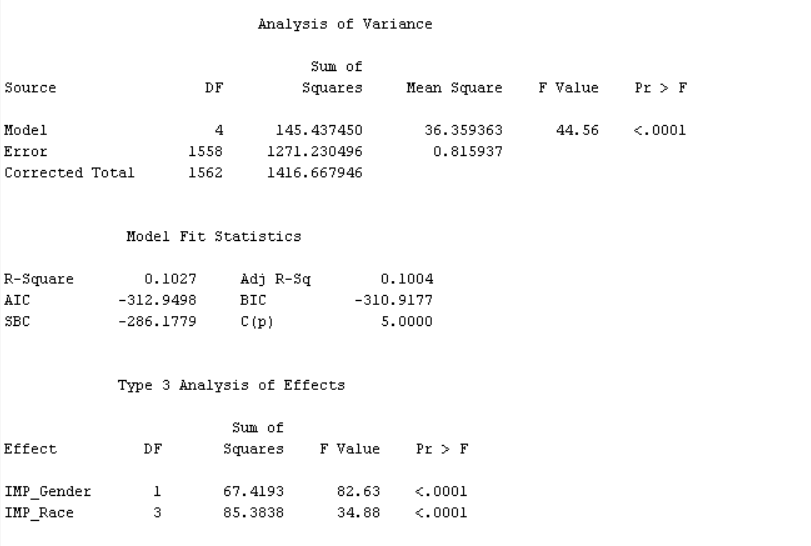
*SAS Enterprise Miner – Significance – Data Type - I*



Notes – The SAS enterprise miner shows strong relationship on the mortality counts for all the gender and race combinations.

**Figure 14**

*SAS Enterprise Miner – Model Output – Data Type - II*



Notes – Using different dataset, which was used in the previous analysis where there were only 2 columns for Race and Gender, SAS enterprise miner shows both have a strong significance.

Using the results from the various tools used for the data analysis. Below are the results on the hypothesis.

* How mortality has changed over a period of time
  + H0 – Mortality counts are same over a period of time. The numbers are not increasing for opioid related deaths
  + Ha – Mortality counts are NOT same for opioid related deaths over a period of time

From the various charts mortality counts by the opioid overdose are increasing for some gender and race these are increasing above the average. So, the null hypothesis is true in this case. Alternate hypothesis can be rejected.

* How gender affects the mortality rate
  + H0 – Mortality count is same for both genders
  + Ha – Mortality count is NOT same for both genders

Gender has shown the significant in both the models run in R as well as SAS Enterprise miner.

* How race affects mortality rates
  + H0 – Mortality counts are same across all races
  + Ha – Mortality counts are NOT same across all races

Race is also a significant factor in increase in mortality rates due to opioid overdose. Whites have shown Significant increase in the number of deaths due to opioid overdose. This is across both the genders.

# **CONCLUSION**

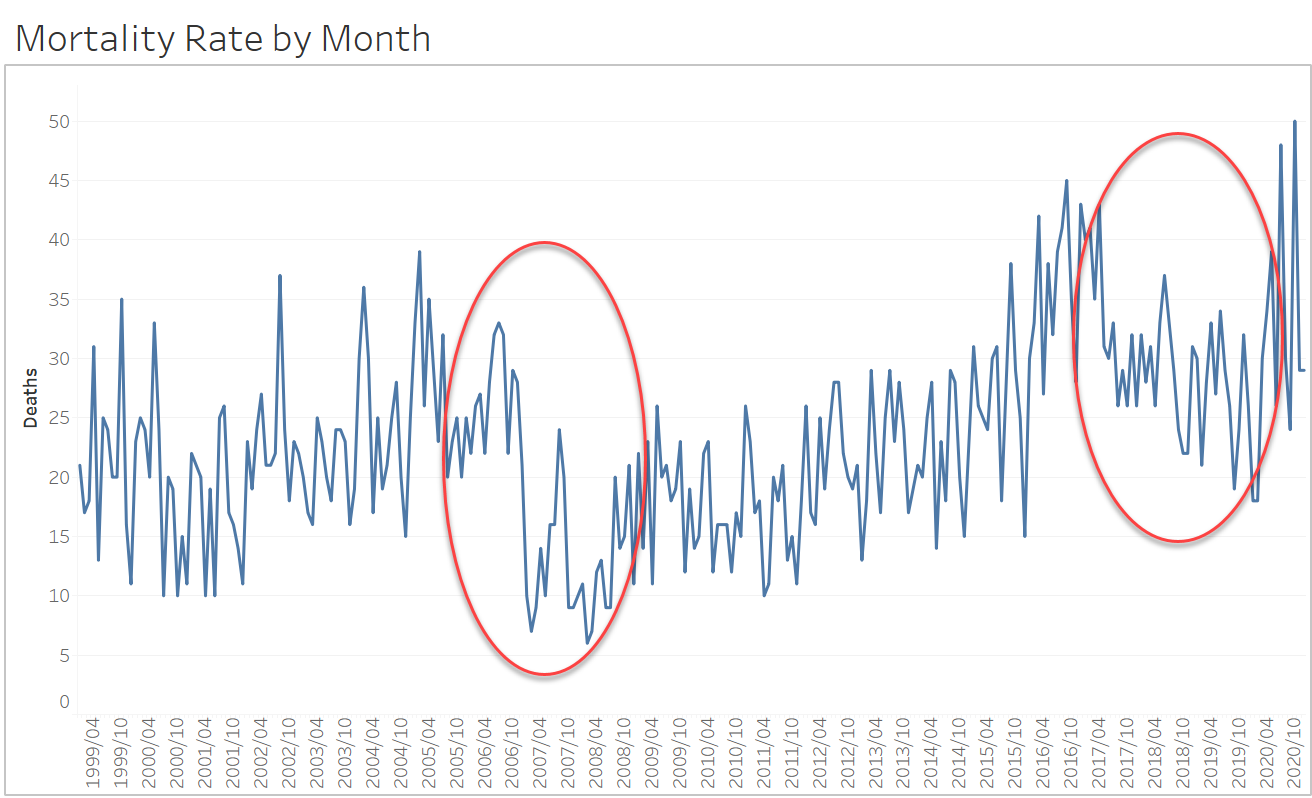
The analysis clearly shows that there is significant increase in the mortality counts by opioid overdose in America. The study was done across multiple race such as American Indian or Alaska Native, Asian or Pacific Islander, Black or African American and White. The data was also divided by the gender. The study clearly shows there is significant increase in the mortality counts in whites. It also shows the increase in the female white counts for past decade. This is in-line with the study conducted by Chartoff & McHugh in 2016.

# **RECOMMENDATIONS**

There are few other observations from the data visualization. There is no doubt in confirming that the mortality rates due to opioid overdose is on the rise. There are couple of observations related to data. There is increase in counts for the white females. This should be studied further for the identification of the root cause. The increase is very evident in the past decade. Another observation from the data is the chart shown in figure 15. It shows there is a decline in the rates before it starts rising again. This should be studies further to identify what causes the reduction in these numbers. This can help in controlling this epidemic as it is increasing at an alarming pace.

**Figure 15**

*Mortality counts by Months*



Notes –Mortality counts by the month shows the decline before it starts going up

# **REFERENCES**

Biden Administration Set To Release Plan For Treating Opioid Addiction. (2021). *Morning Edition*.

Blair, K. (2020, Jan 28). Measuring the Impact of Opioid Crisis Management. Retrieved from Express Scripts: <https://express-scripts.com/corporate/index.php/articles/measuring-impact-opioid-crisis-management>

Chartoff, E. H., & McHugh, R. K. (2016). Translational Studies of Sex Differences in Sensitivity to Opioid Addiction. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *41*(1), 383–384. https://doi.org/10.1038/npp.2015.272

Furlow, B. (2016). FDA confronts opioid addiction and overdose deaths. *The Lancet. Oncology*, *17*(3), e95. https://doi.org/10.1016/S1470-2045(16)00090-5

Gaither, J. R., Shabanova, V., & Leventhal, J. M. (2018). US National Trends in Pediatric Deaths from Prescription and Illicit Opioids, 1999-2016. JAMA Network Open, 1(8), e186558. https://doi.org/10.1001/jamanetworkopen.2018.6558

intellipaat. (2022, Jan 08). What is Tableau? Retrieved from intellipaat: https://intellipaat.com/blog/what-is-tableau/

Kirvan, P. (2021, Jan 6). 20 keys to a successful enterprise data protection strategy. Retrieved from techtarget: https://www.techtarget.com/searchdatabackup/tip/20-keys-to-a-successful-enterprise-data-protection-strategy

Lee, W. W., Zankl, W., & Chang, H. (2016, Dec 24). An Ethical Approach to Data Privacy Protection. Retrieved from ISACA JOURNAL: <https://www.isaca.org/resources/isaca-journal/issues/2016/volume-6/an-ethical-approach-to-data-privacy-protection>

McCarthy, M. (2016). US needs $1.1bn to treat opioid addiction. *BMJ (Clinical Research Ed.)*, *352*, i694. https://doi.org/10.1136/bmj.i694

mergr. (n.d.). Cigna Acquires Express Scripts Holding. Retrieved from mergr: https://mergr.com/cigna-acquires-express-scripts-holding#:~:text=On%20March%208%2C%202018%2C%20Cigna%20Corp.%20acquired%20healthcare,sector.%20This%20is%20Cigna%E2%80%99s%201st%20largest%20%28disclosed%29%20transaction.

NIDA. (2021, March 11). Opioid Overdose Crisis. Retrieved from National Institute of Drug Abuse: https://nida.nih.gov/drug-topics/opioids/opioid-overdose-crisis

PCMA. (2022). Pharmacy Benefit Managers (PBMs) Provide Solutions to Address the Opioid Crisis. Retrieved from Drug Benefit Solutions: https://drugbenefitsolutions.com/drug-policy/pbm-solutions-to-address-opioid-crisis/

SAS. (2022). SAS. Retrieved from SAS Enterprise Miner™ Features: https://www.sas.com/en\_us/software/enterprise-miner/features-list.html