Accidental Drug Deaths 2012- 2018

CalStateLAseal_blkgold

OCTOBER 31

California State University LA
Authored by: Ambika Sharma - 401349647

Table of Content

AIM

To Identify Factors Between Drug Use and Human Mortality

The ever-increasing practice of using several drugs in the United States from the period 2012-2017 has caused the rates of mortality from drug abuse to hit the roof.

Though along with the use of prescribed medication there is a prevalence of illegal use of opioids, heroin, cocaine, meth, and the likelihood of users becoming dependent on a drug on long-term spike after just five days of consumption.

Unfortunately, the epidemic has affected almost every age group in every U.S. population. The impact is so serious that every day more than 130 Americans die from a drug overdose[1]. In this study, we are going to examine whether there is any correlation between the prescription drugs and deaths from overdose. Further, our objective is to find associations among risk factors and identify statistically significant factors in individuals or groups who are susceptible to drug abuse. Finally, we aim to develop a risk model to predict a patient's risk of drug abuse or death from future drug use.

> ***"Nearly 841,000 people have died since 1999 from drug overdose."*** **(1)**

[1] **Wide-ranging online data for epidemiologic research (WONDER). Atlanta, GA: CDC, National Center for Health Statistics; 2016. Available at http://wonder.cdc.gov.**

INTRODUCTION

Brief

Deaths from drug overdose continue to be a communal health burden in the United States[2]. This report uses the data from the SAS to update statistics on deaths from drug overdose in the United States, including information on trends from 2012 through 2018 by sex and age group, and by specific types of drugs involved (i.e., opioids, heroin, cocaine, meth, and stimulants). I have tried to derive several graphs from the data in SAS.

Numerous data quality metrics, including the percent extensiveness in overall death reporting, percentage of deaths with cause of death pending further investigation, and the percentage of drug overdose deaths with specific drugs or drug classes reported are included to aid in understanding of provisional data as these measures are related to the precision of provisional counts.

Illicit drugs, and the global and domestic criminal organizations that traffic them, continue to represent significant threats to public health, law enforcement, and national security in the United States. The opioid risk (controlled prescription drugs, synthetic opioids, and heroin) continues at ever-increasing epidemic levels, affecting huge portions of the United States. Meanwhile, the restorative threat (methamphetamine and cocaine) is worsening and spreading as traffickers continue to sell increasing amounts outside of each drugs' traditional markets. New psychoactive substances (NPS) remain stimulating, and the inland marijuana situation is sprouting as state-level medical and recreational legalization continues.

Drug poisoning deaths are the leading cause of grievance death in the United States, which is drastic, as these are not natural cause and can be controlled and investigated.

[2] **Hedegaard H, Miniño AM, Warner M. Drug overdose deaths in the United States, 1999–2018. NCHS Data Brief, no 356. Hyattsville, MD: National Center for Health Statistics. 2020. Available from: https://www.cdc.gov/nchs/data/databriefs/db356-h.pdf**

DATA COLLECTION AND PREPARATION (DATA CLEANING)

Raw Data –

Collecting data available in SAS, with Field Name, Data Description and describing the type of data as

Data Type – 1 that tells whether the Data is Quantitative or Qualitative in nature.

Data Type – 2 is about the data type to be Number, decimal, float, date, string etc.

Field Name	Data Description	Data Type – 1	Data Type – 2
ID	Unique identifier	Quantitative – discreet	Number
Date	Date of incident	Qualitative	Date
DateType	Provides additional information about the date. The date can pertain to when the death took place or when it was reported.	Qualitative	String
Age	Age of individual that died	Quantitative – continuous	Number
Sex	Gender of individual that died	Qualitative	String
Race	Race of individual that died	Qualitative	String
ResidenceCity	City of Residence of the individual that died	Qualitative	String
ResidenceCounty	County of Residence of the individual that died	Qualitative	String
ResidenceState	State of Residence of the individual that died	Qualitative	String
DeathCity	City in which the individual died	Qualitative	String

DeathCounty	County in which the individual died	Qualitative	String
Location	Location in which the individual died i.e., hospital or residence	Qualitative	String
LocationIfOther	If the location is something other than hospital or residence, an alternate location is provided	Qualitative	String
DeathCityLat	Latitude of the city in which the individual died	Quantitative – discreet	Float/Decimal
DeathCityLong	Longitude of the city in which the individual died	Quantitative – discreet	Float/Decimal
Heroin	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Cocaine	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Fentanyl	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
FentanylAnalogue	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Oxycodone	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Oxymorphone	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Ethanol	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Hydrocodone	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Benzodizepine	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Methadone	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Amphet	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String
Tramad	A "Y" or NULL is included to indicate if this drug was cause of death	Qualitative	String

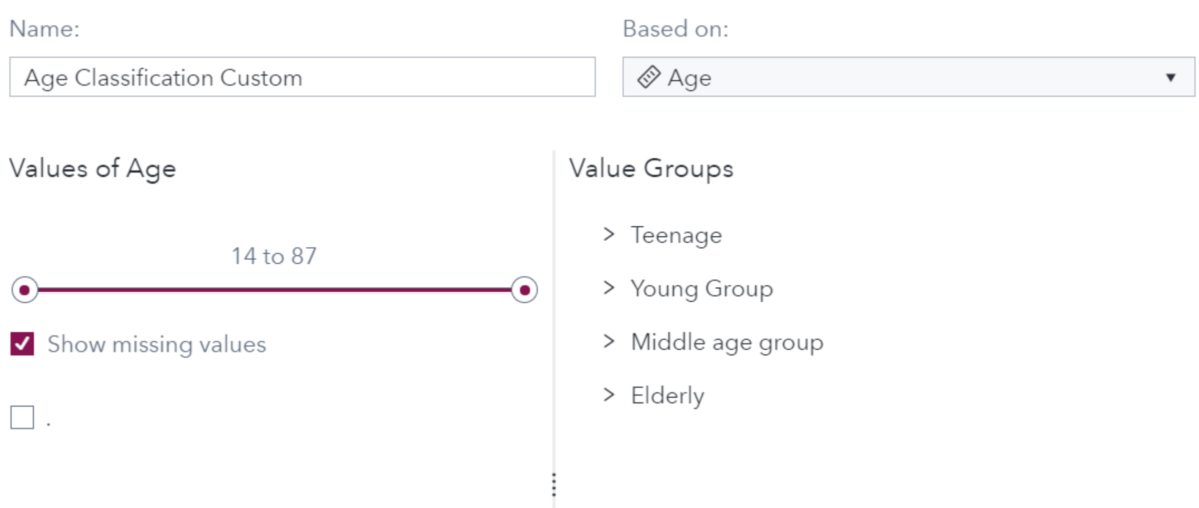
Morphine_NotHeroin	A “Y” or NULL is included to indicate if this drug was cause of death	Qualitative	String
Hydromorphone	A “Y” or NULL is included to indicate if this drug was cause of death	Qualitative	String
Other	A “Y” or NULL is included to indicate if this drug was cause of death	Qualitative	String
OpiateNOS	A “Y” or NULL is included to indicate if this drug was cause of death	Qualitative	String
AnyOpioid	A “Y” or NULL is included to indicate if this drug was cause of death	Qualitative	String

Fields Used -

Fields which glow

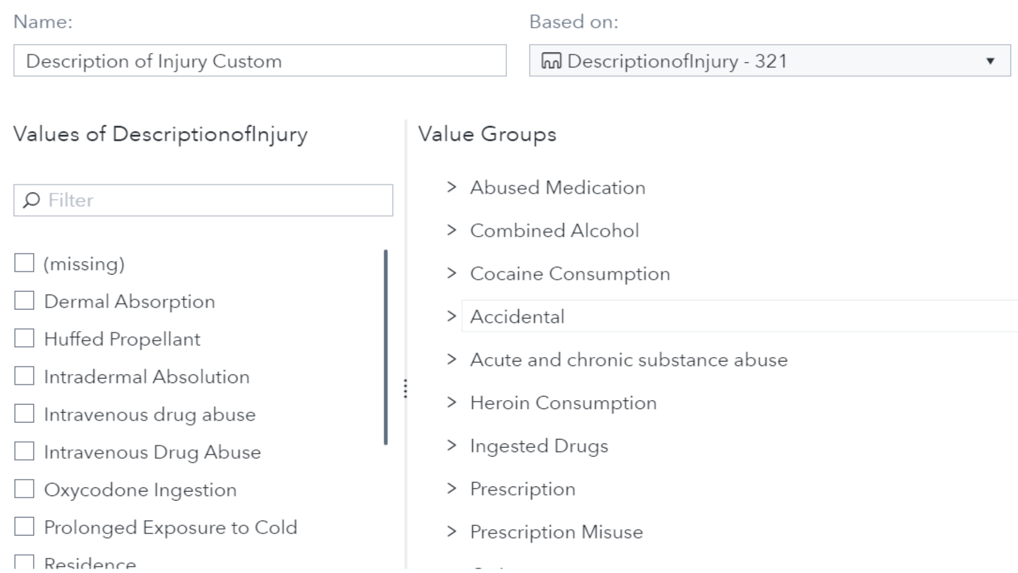
- **AGE –**

Data available in SAS, had age with distinct value, I created the age as a custom category with age range defining four groups – **Teenage (14 to 20), Young Group (21 to 30), Middle age group (36 to 55) and Elderly (56 and above).**



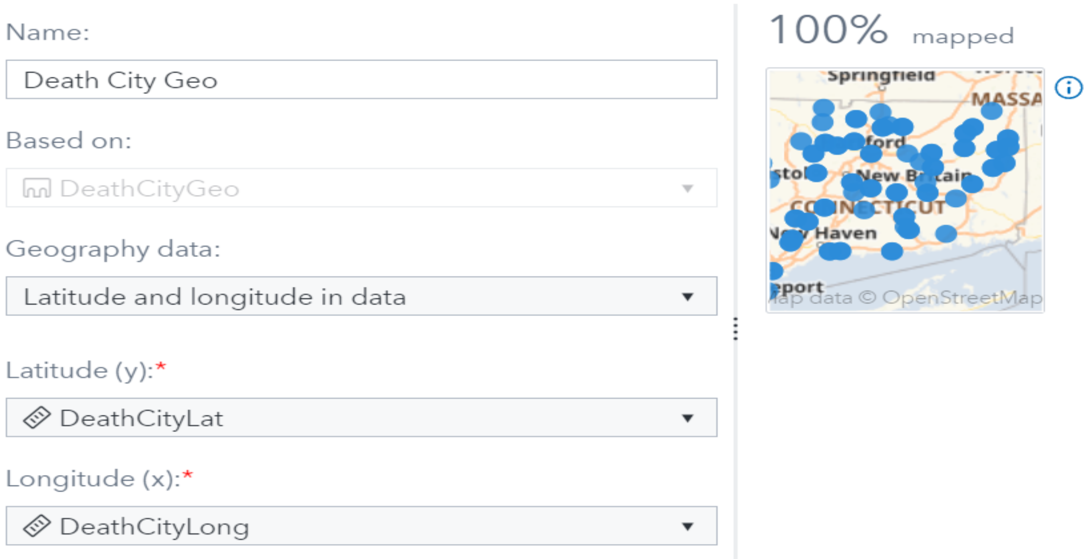
- **Description of Injury –**

Data available in SAS, had several values with repetitive information of how a patient is injured, I **grouped** those values under **19 topics** and assembled them later in the graphs. The remaining values were also grouped as Others.



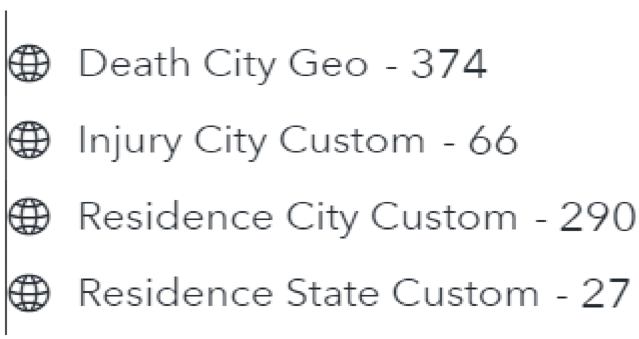
- **Injury City –**

Data available in SAS, had **latitude and longitude value** of Injury City, which were combined using geographical category and mapped in the data.



- **Death City –**

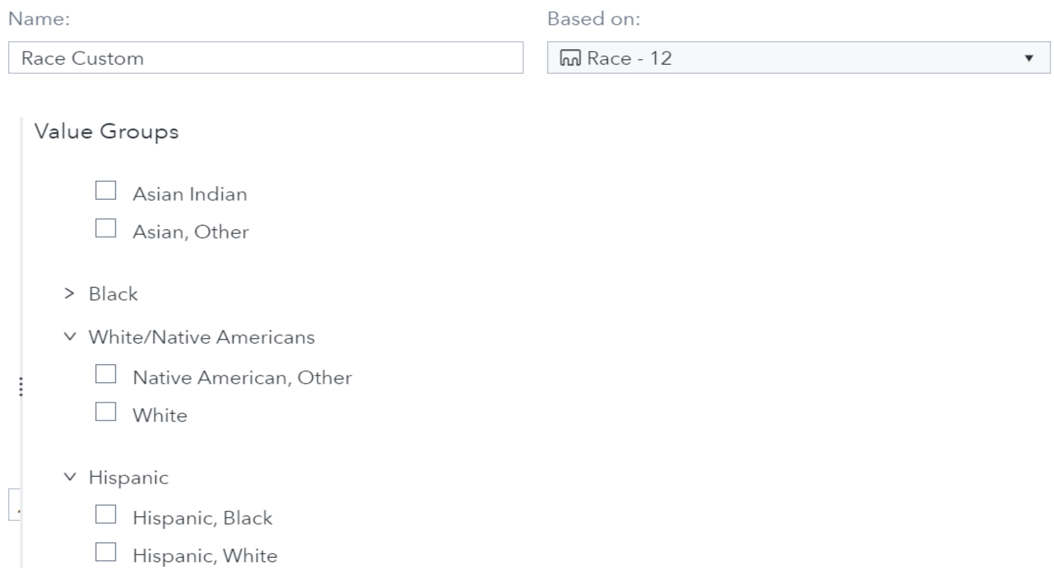
Data available in SAS, had values like death city Lat and Long values which were customized as what I did with Injury City. Similarly, Residence City and Residence State were formed using custom category under geography.



- **Race –**

Data available in SAS, with Race had redundant data, so I made a custom category for Race where the values were grouped. Under Asian – Asian Indian and Asian, other. Under

White/Native Americans – White, and Native Americans and so on.



- **Sex –**

Data available in SAS, it was used as is – with three categorization – Male, Female and Unknown.

- **Consumption of several Drugs (Heroin, Amphet, Opioid, Cocaine, Meth and Morphine) –**

Data available in SAS, I chose few out of many drugs and customized them as and made them a calculated value.

Made a measure, from Heroin, Amphet, Opioid, Cocaine, Meth and Morphine consumption and created few graphs according to it.

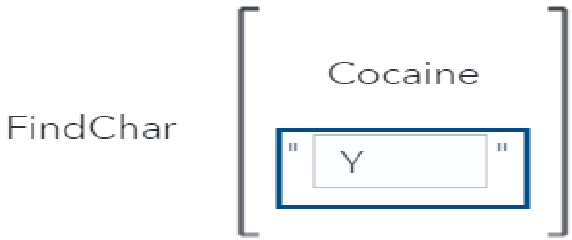
How the calculation was generated?

This was the most challenging calculation, as I was unable to figure out how to calculate number of people from just two values for each drug which was **“Y”** and **“N”.** *I took the character and found out if the string contains Y then return “1” else do not detect anything*

<u>and then utilized the sum of all the 1's which constitutes the sum of people who consumed any of the above drug and voila, our measure was ready.</u>

For example -

In drug Cocaine we find the character "Y"



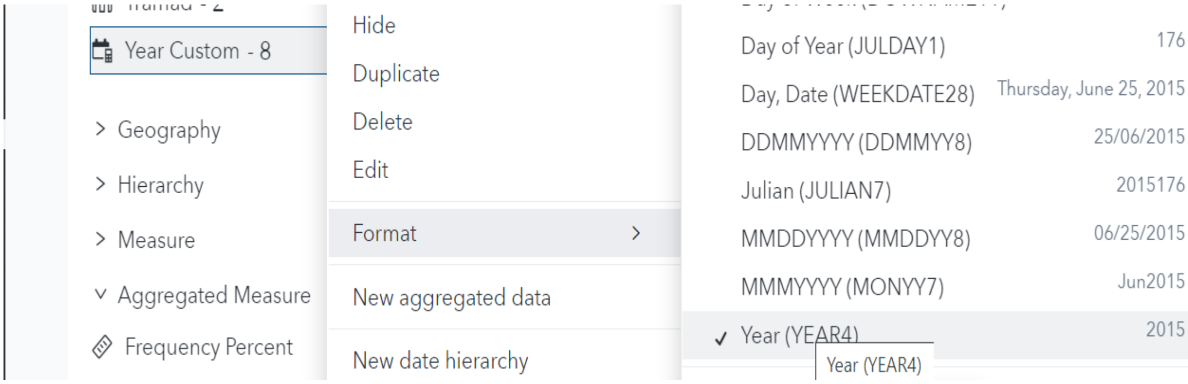
And we preview the values as 0s and 1s- the sum of 1's makes us aware of the total number of people who consumed cocaine and similarly the same measure is used for few drugs and I have further used them in graphs.

Number of rows to show: 50

Cocaine Consumption	Cocaine
0	
1	Y
1	Y
0	
0	
0	
0	
1	Y
0	
0	
1	Y

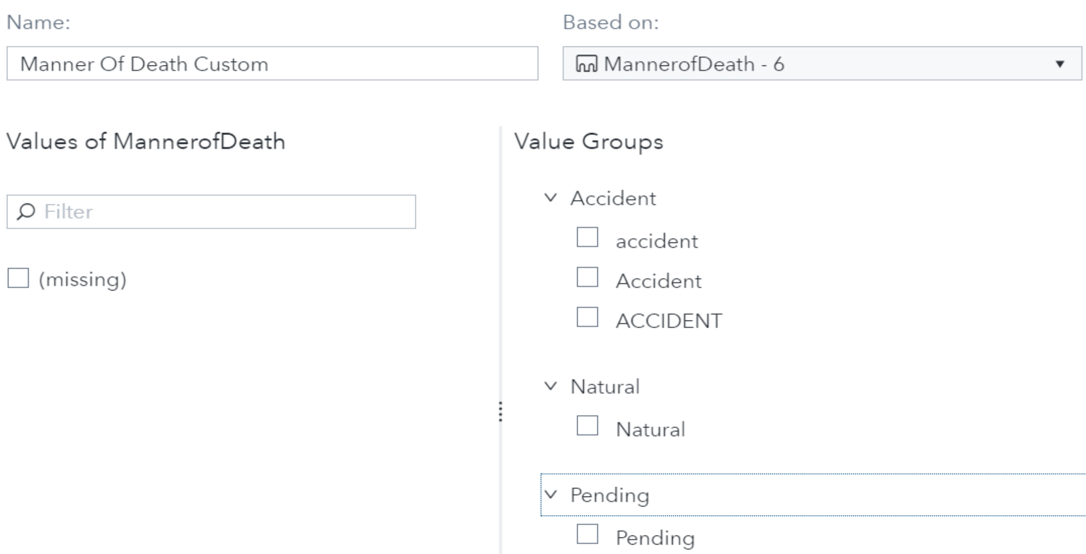
- **Date –**

Data available in SAS, I just changed the format of the date and got the **years,** and then I have used the year recording the number of deaths by consumption of drugs every year.



- **Manner of Death –**

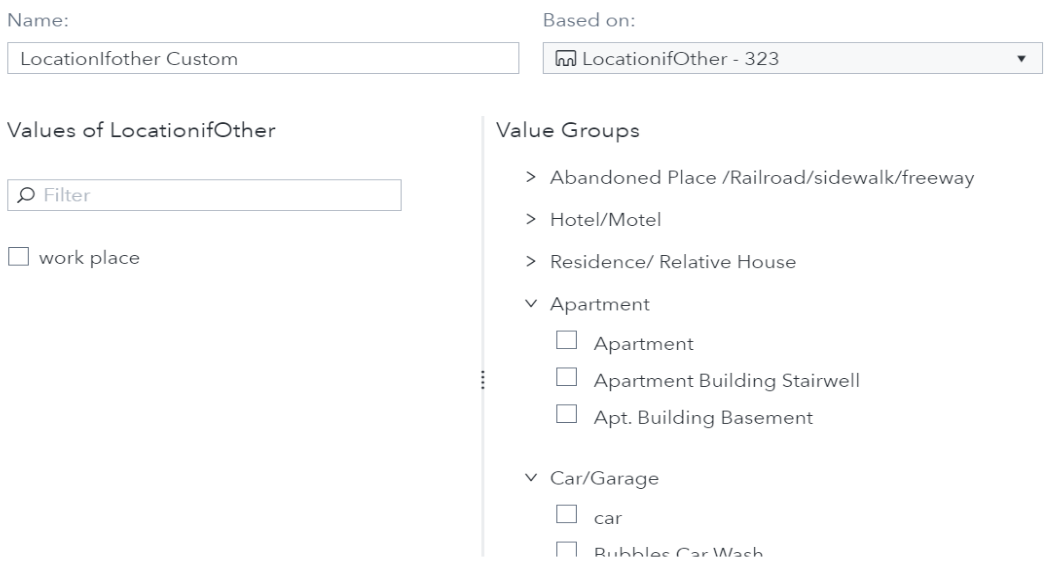
Data available in SAS, created another custom field – with three groups and assigning the repetitive value under a group which was Accident and other 2 groups accordingly.



- **Location if other –**

Data available in SAS, created another custom category of where the death occurred –

created 6 groups of different location as below:



METHODOLOGY

Three Phases of Methodology –

1. **Descriptive Analysis:** It was performed to state the key metrics that would be the decisive points of the learning. This included exploratory investigation by examining the trends of recorded deaths due to several types of drugs consumption and by analyzing the data of various age-groups that were impacted by this epidemic. The state-wise mortality data owing to drug abuse was examined too.

2. **Correlation Analysis:** Various studies stated that physicians, insurance companies, pharma companies and even the individuals misusing the prescription, are all the reason for fueling up the crisis. The study in JAMA Network Open Drug states that drug companies have been spending billions of dollars to market their products to doctors, and other prescribers with speaking fees, free dinners, paid trips, and more [3]. My report whereas focused to find the correlation between:

i. Prescription rate vs drug overdose deaths.

ii. Drug consumption vs Race.

iii. Drug Consumption vs Sex.

iv. Drug Consumption vs Year

3. **Predictive Analysis:** There are several factors at an individual level that may have been contributing to the rise of the epidemic. For instance, lack of awareness on the effective use of these drugs, high consumer demand for drugs, broken health monitoring program and unavailability of predictive tools to predict the abuser of the prescribed drug. Therefore, a predictive analysis was performed to construct two different models.

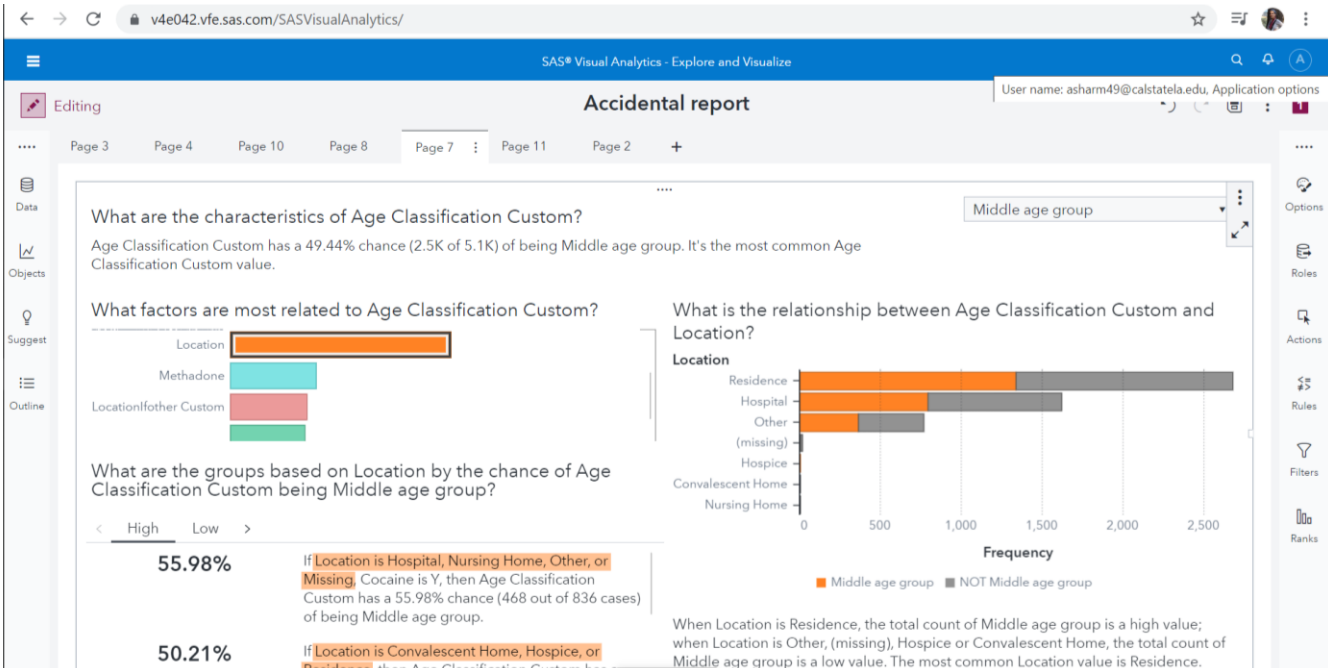
i. Prediction for Race Custom?

ii. What values for the most important factors should be used to predict?

[3] **https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2720914**

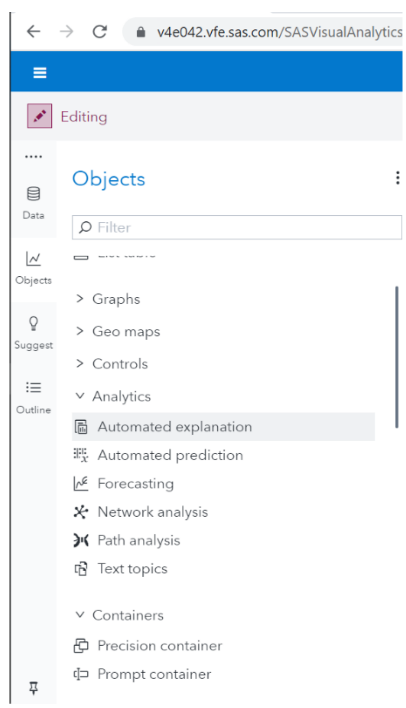
ANALYSIS AND IMPLEMENTATION

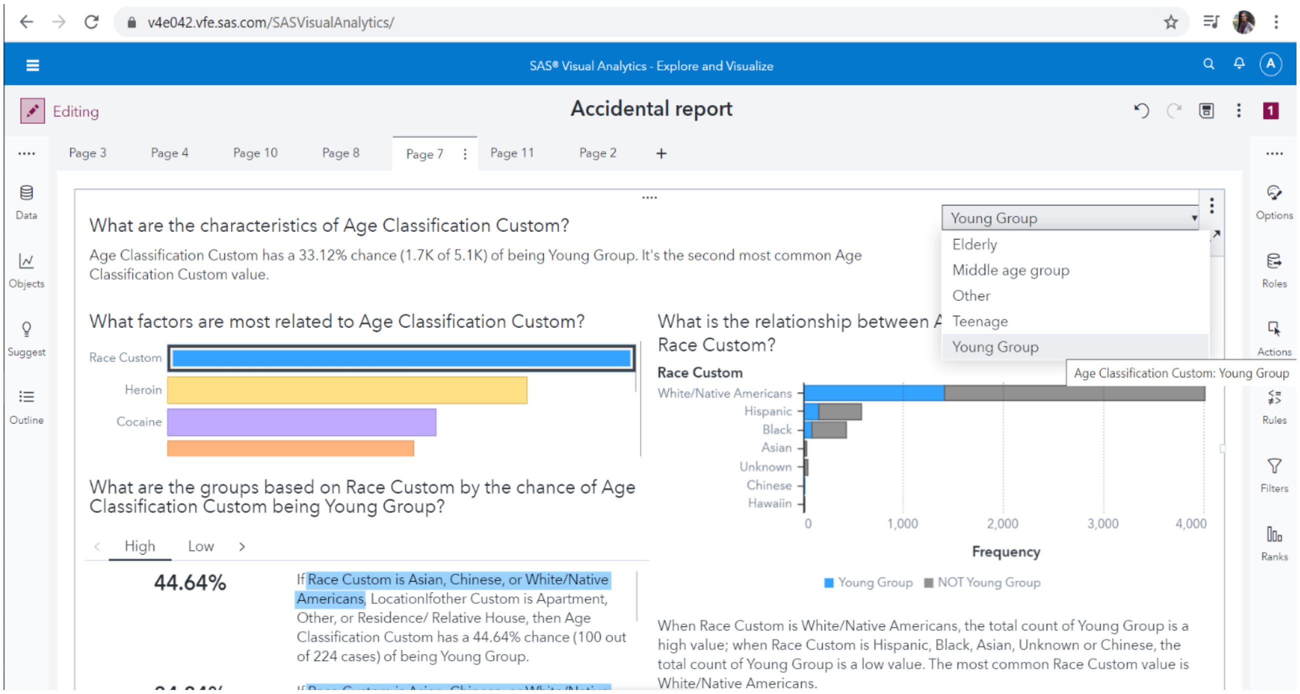
Analysis –



The above is the SAS analysis which explain the raw data. This analytic object is appropriate for anyone who wants to skim their data in the first go and provide tn overview, of what data is and what it is about.

Playing with data was a fun activity, the drop down in the above right corner allowed me to select a specific age group – **Young, Elderly, Middle age, Teenage group** etc. according to the custom field which I created initially.





Key points to grab from this analysis:

What are the characteristics of age classification custom?

Age classification custom has a ***33.12% chance (1.7k of 5.1k) of being young group***. It's the second most common age classification custom value.

What factors are most related to age classification custom?

This is based on factors like race, consumption of drugs, location of death, sex etc.

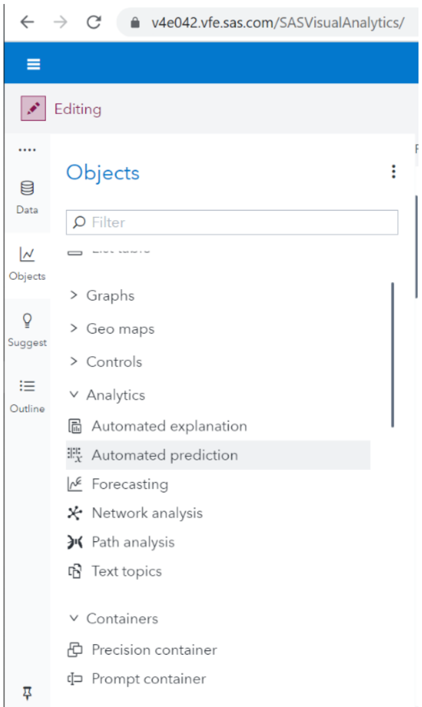
What is the relationship between age classification custom and race custom?

When race custom is white/native Americans, the total count of young group is High. When race custom is Hispanic, black, Asian, unknown, or Chinese, the total count of young group is a low value. The most common race custom value is ***white/native Americans.***

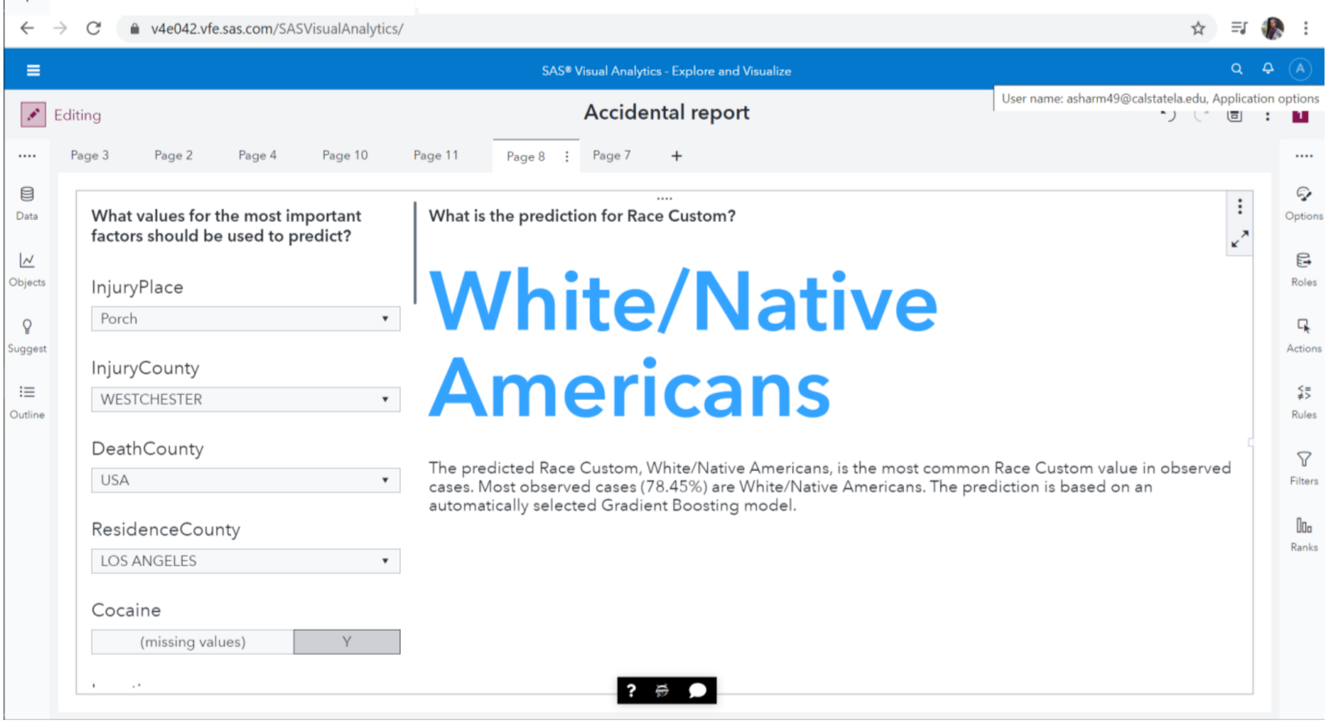
What are the groups based on race custom by the chance of age classification custom being young group?

If race custom is **Asian, Chinese, or white/native Americans**, location if other custom is apartment, other, or residence/ relative house, then age classification custom has a ***44.64%*** chance (100 out of 224 cases) of being young group.

Analysis –



The second analysis was generated through another analytical option, which allows to predict data. With a given response, the predictive analysis provides an automated prediction, for example: I took the response data “**race**”. I made a custom field – **Race Custom** – where I grouped several users and categorize them according to their race – Hispanic (inclusive of Hispanic black and Hispanic), White/Native Americans (inclusive of whites and native Americans), Asians (Inclusive of Chinese, Indians, and other Asians), Black etc.



What is the prediction for Race Custom?

White/Native Americans - The predicted Race Custom, White/Native Americans, is the most common Race Custom value in observed cases. Most observed cases (78.45%) are White/Native Americans. The prediction is based on an automatically selected Gradient Boosting model.

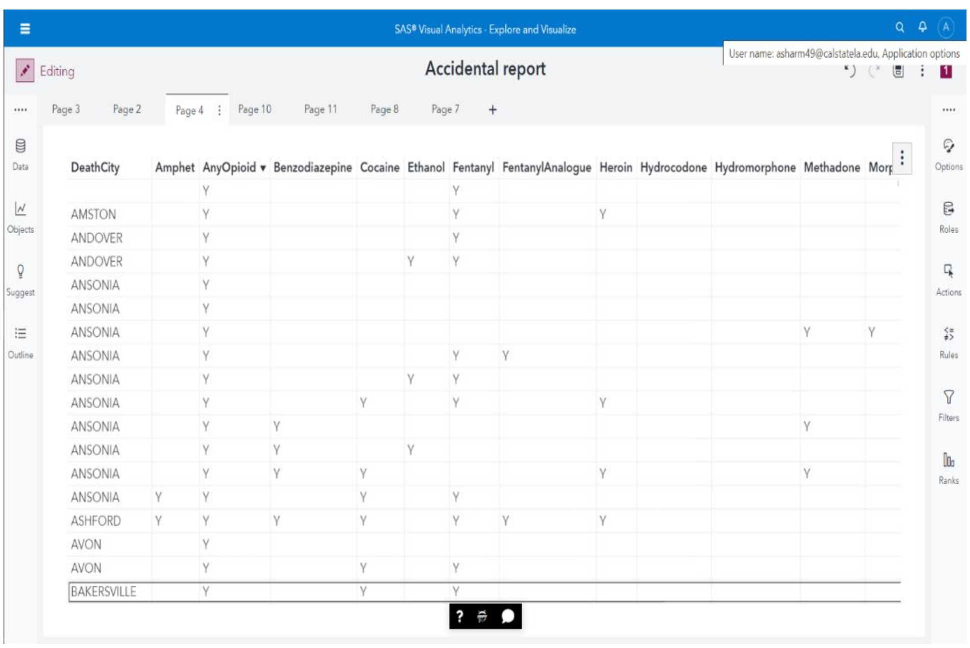
What values for the most important factors should be used to predict?

Injury Place, Injury and Death County, use of the drug be it – cocaine, heroin, ethanol were the factors used to determine the custom race.

The above two graphs gave me an idea of how prediction analysis work in SAS and were extremely helpful to provide holistic view of the data.

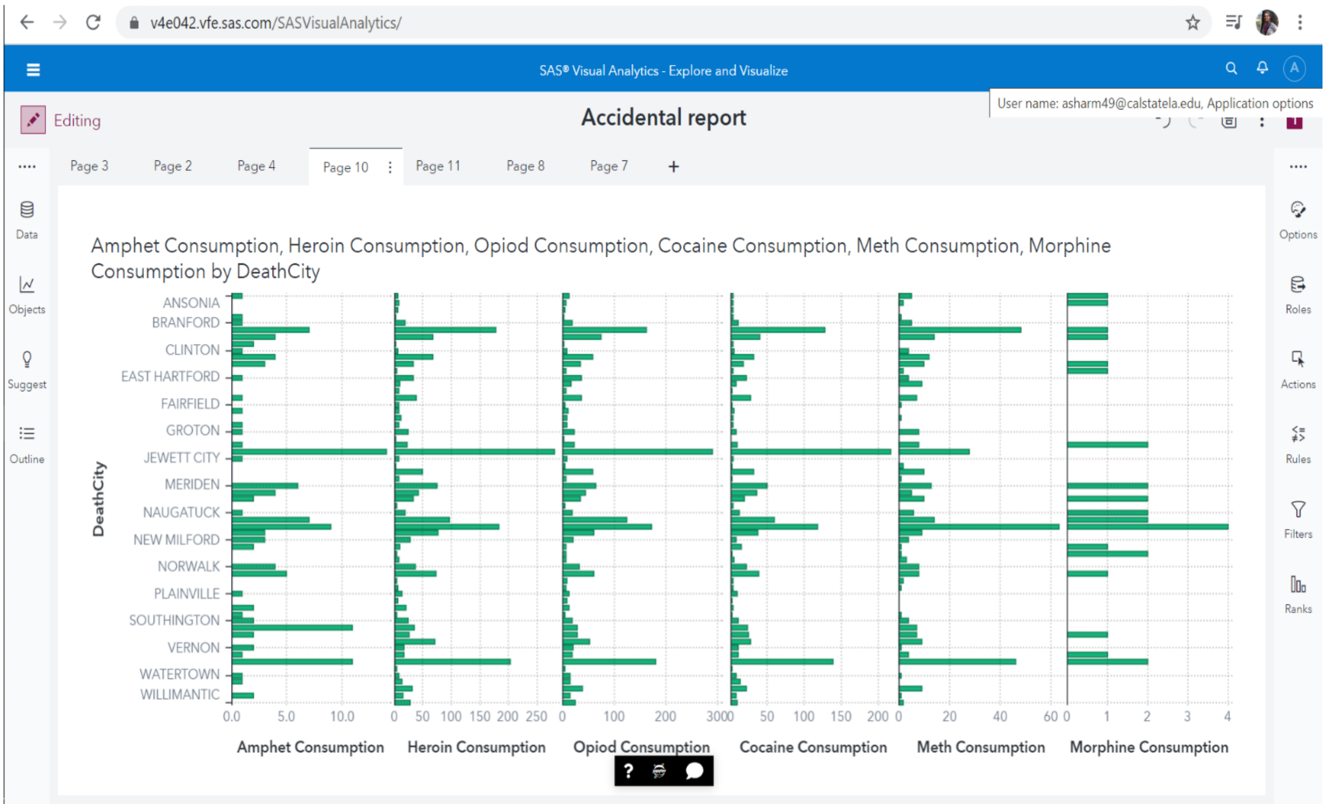
Implementation –

- **Number of deaths based on drugs consumption (Heroin, Amphet, Opioid, Cocaine, Meth, Morphine) in different states, from 2012 to 2018.**



DeathCity	Amphet	AnyOpioid	Benzodiazepine	Cocaine	Ethanol	Fentanyl	FentanylAnalogue	Heroin	Hydrocodone	Hydromorphone	Methadone	Morp
		Y				Y						
AMSTON		Y				Y		Y				
ANDOVER		Y				Y						
ANDOVER		Y			Y	Y						
ANSONIA		Y										
ANSONIA		Y										
ANSONIA		Y									Y	Y
ANSONIA		Y				Y	Y					
ANSONIA		Y			Y	Y						
ANSONIA		Y		Y		Y		Y				
ANSONIA		Y	Y								Y	
ANSONIA		Y	Y		Y							
ANSONIA		Y	Y	Y				Y			Y	
ANSONIA	Y	Y		Y		Y						
ASHFORD	Y	Y	Y	Y		Y	Y	Y				
AVON		Y										
AVON		Y		Y		Y						
BAKERSVILLE		Y		Y		Y						

This is the raw data pulled from Data Source which represents the usage based on various drugs. I have created custom calculation for few drugs using the **count of 'Y'** to identify the impacted population by City. This was especially useful to represent the report and visualization. The below graph from the raw data received and converted.

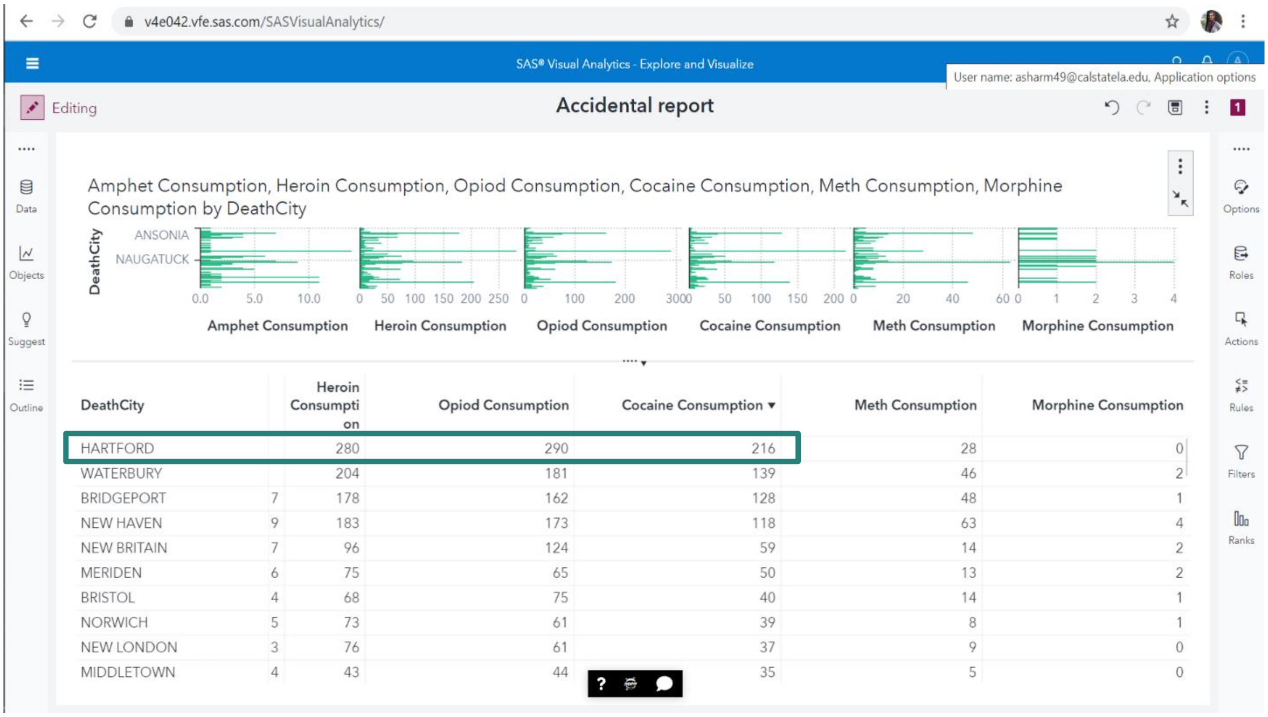


drug-consumption-in-years.mp4

This is an interactive report which represents the consumption of **various drugs vs the City in the years.** The video link attached above will give an idea of how the rate of consumption of drugs increases and decreases between 2012 to 2018. By selection of a specific drug, we can observe

- **Opioids consumption was leading factor involved in 290 deaths.**
- **Followed by Heroin, was involved in 280 deaths.**
- **Cocaine was involved in 216 deaths.**

All the above is for HARTFORD city.



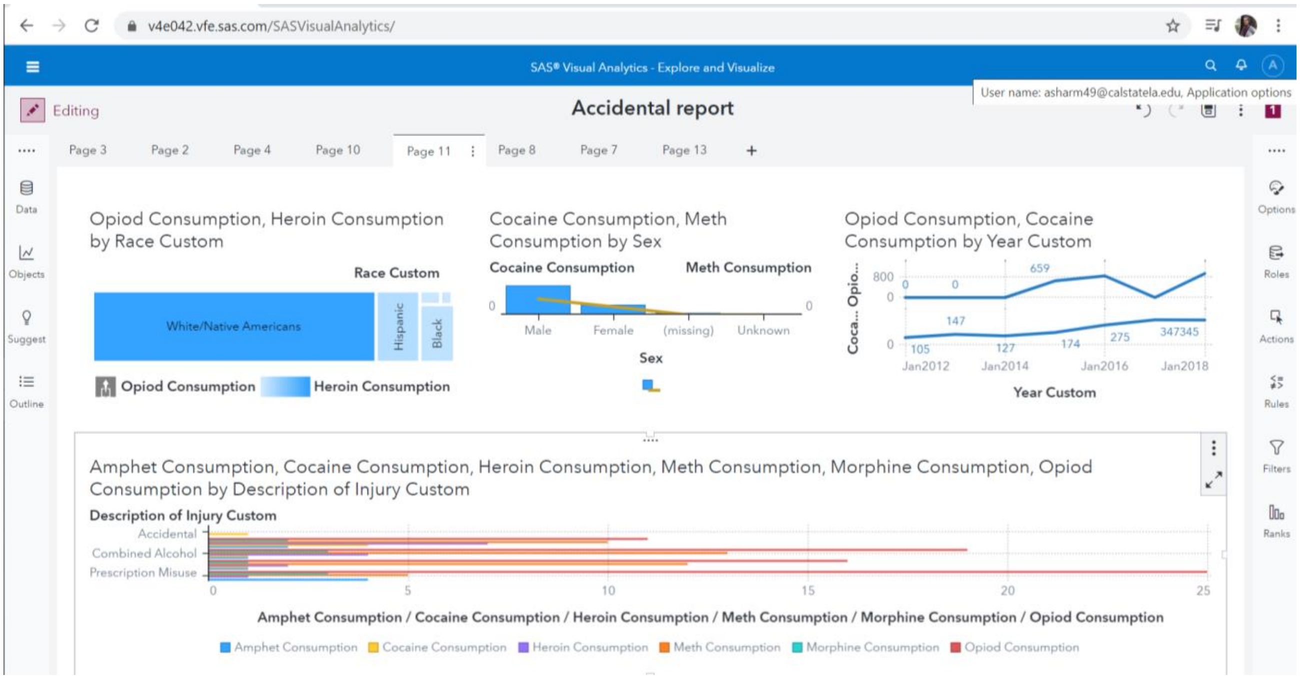
DeathCity		Heroin Consumption	Opiod Consumption	Cocaine Consumption ▾	Meth Consumption	Morphine Consumption
HARTFORD		280	290	216	28	0
WATERBURY		204	181	139	46	2
BRIDGEPORT	7	178	162	128	48	1
NEW HAVEN	9	183	173	118	63	4
NEW BRITAIN	7	96	124	59	14	2
MERIDEN	6	75	65	50	13	2
BRISTOL	4	68	75	40	14	1
NORWICH	5	73	61	39	8	1
NEW LONDON	3	76	61	37	9	0
MIDDLETOWN	4	43	44	35	5	0

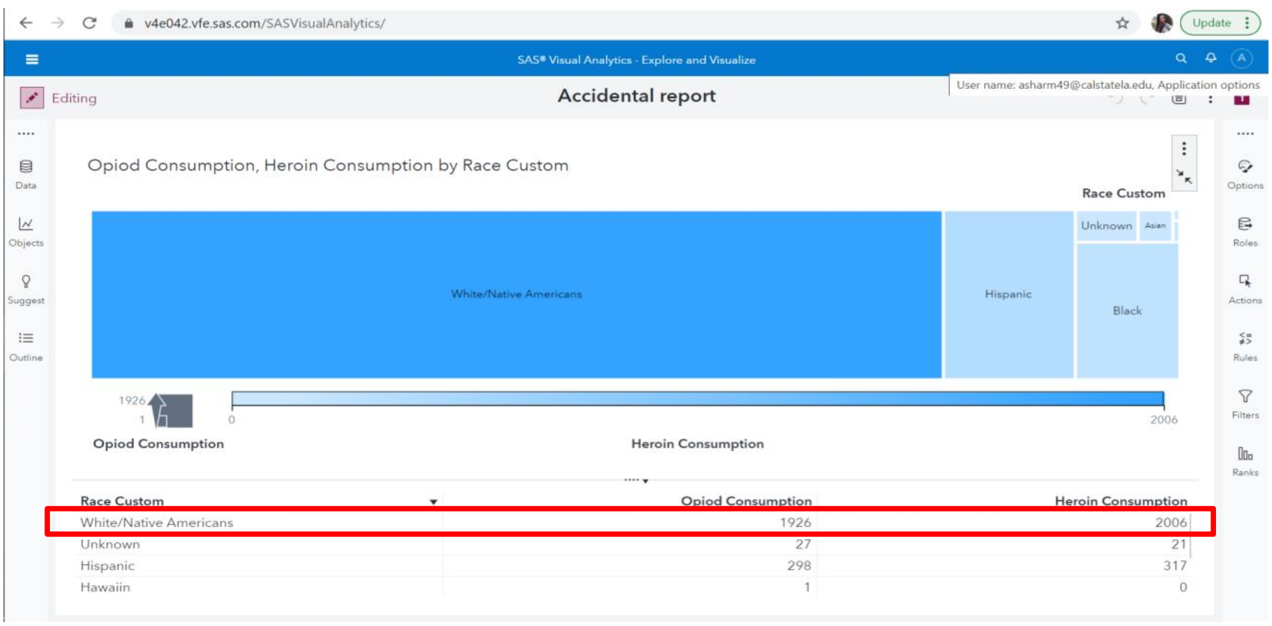
The above graph gave me an idea to further simplify it and check the number of deaths according to ***year, sex, race, and injury type.***

So, I created a dashboard with the below points –

- **Consumption of drug based on race**
- **Consumption of drug based on gender**
- **Consumption of drug based on year**
- **Description of injury based on drug consumption.**

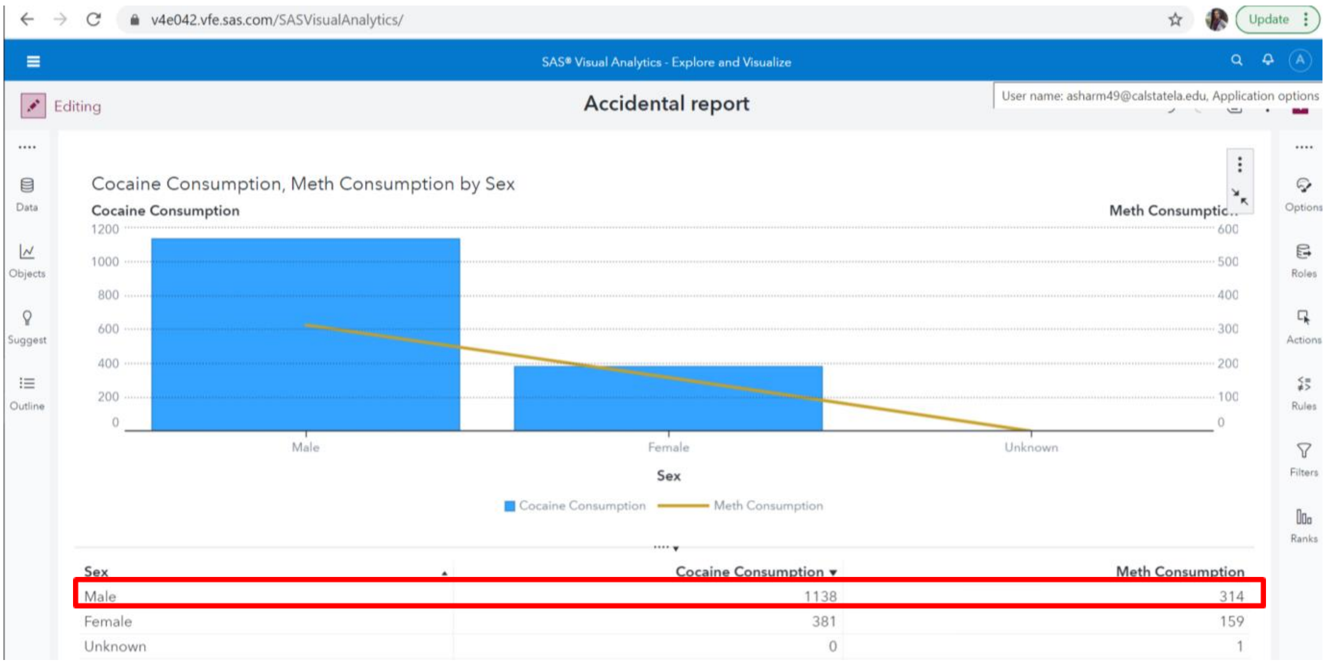
I have depicted the answers to these problems through the dashboard below and I will explain each graph according to the condition described above.





This report provide visualization grouped by consumption of Opioid and Heroin, vs the Race sample set. Contrary to popular belief the **White/ Native Americans** lead the sample set with the highest consumption of **Opioid (1926 deaths)** and **Heroin (2006 deaths)**.

There are several drugs to be personalized, I took few known drugs and customized them according to different factors, like for the graph below it provides us the usage among different Sex grouped by Cocaine and Meth. **Male** lead this trend with higher consumption of **Cocaine (1138 deaths).**



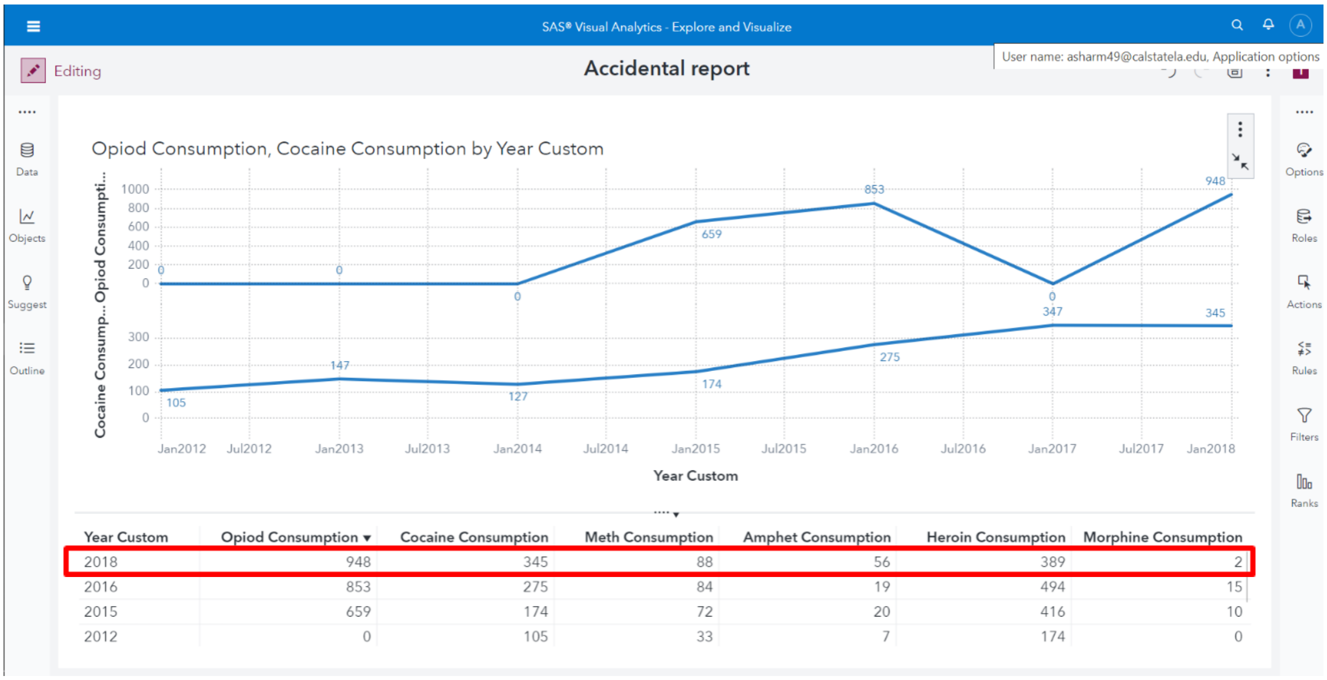
For the Opioid and Morphine consumption with respect to years:

We can observe based on the trend report the consumption of **Opioid** has gone up exponentially in the year 2015, 2016 and 2018. On the other hand, the consumption of **cocaine** has been consistent throughout the years. It is curious to see that in **year 2018 the consumption for both Opioid and Cocaine were at their peak**.

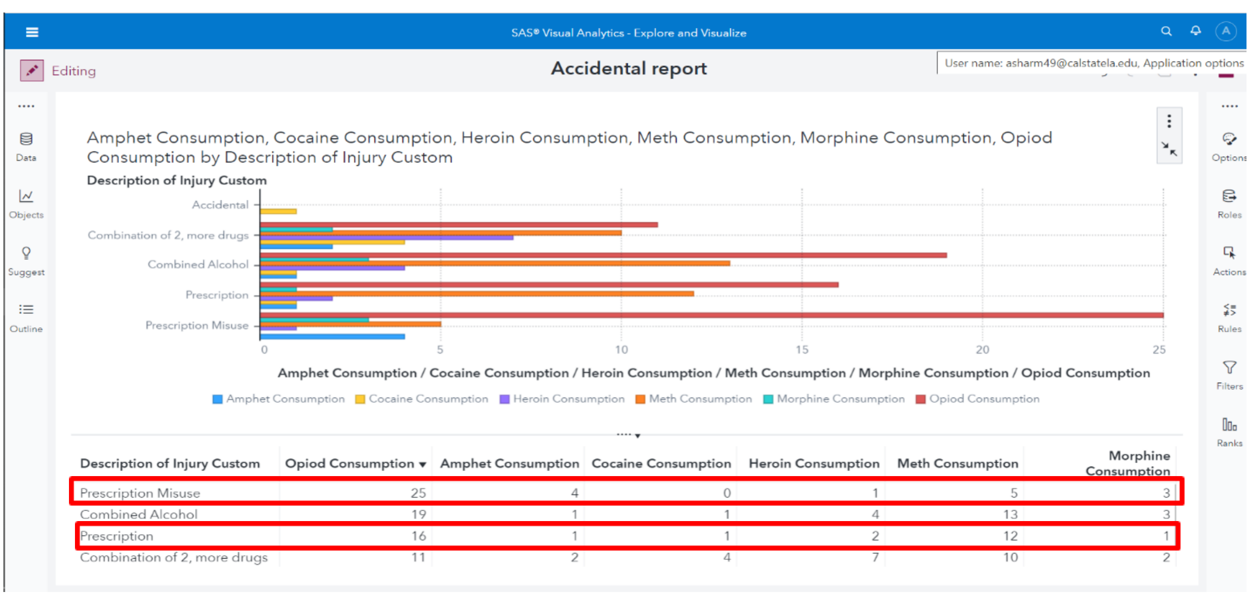
Though there are varied results for drugs every year, we see that the year **2018** has comparatively high consumption of drugs overall.

The graph below shows two sets of drugs, but in the data below that graphs depicts data labels for **count of death from the drug consumption** “Opioid, Methane, Cocaine, Heroin, Amphet, and

Morphine. **So, deaths in the year 2018 from Opioid – 948, Cocaine – 345, Meth – 88, Amphet – 56, Heroin – 389, and morphine – 2.**

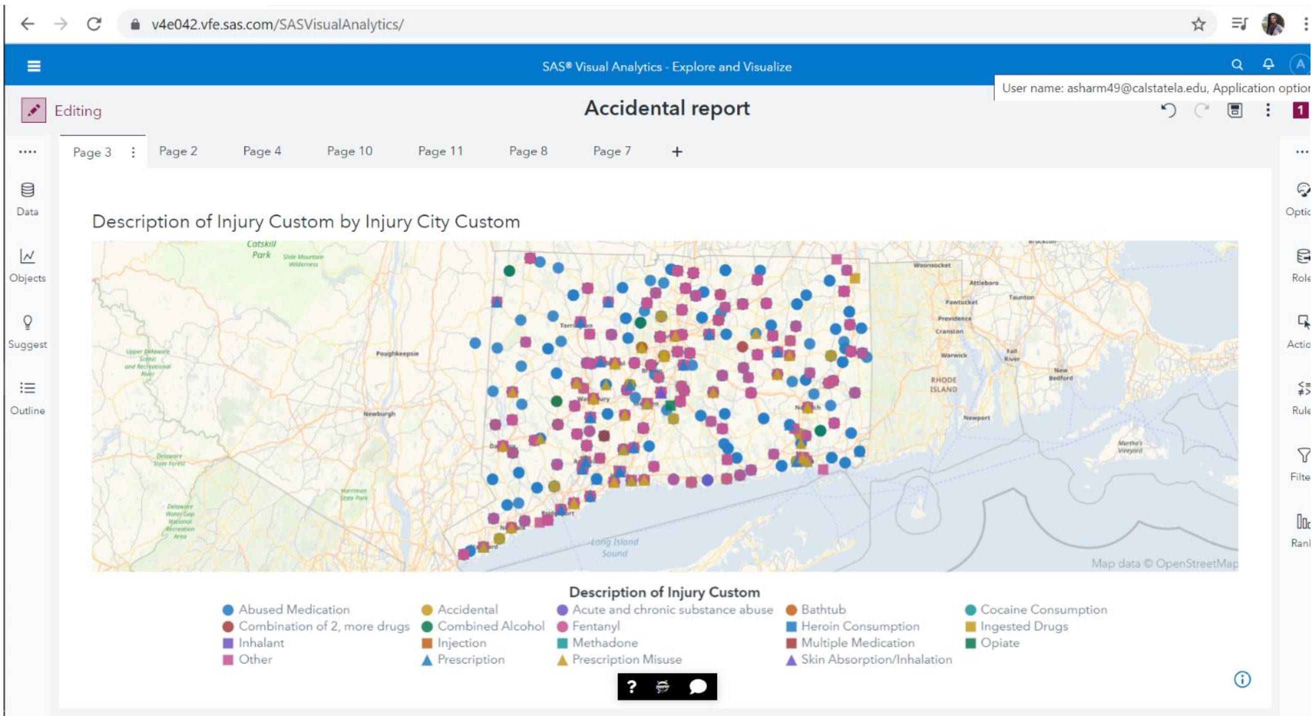


The below graph provides the **description of injury from the doctor's prescription** – how the prescription was misused and how much harm did the prescription did overall if it wasn't mixed with other drugs. So overall a pattern is seen where **opioid prescription** has been a **major contributor of deaths.**



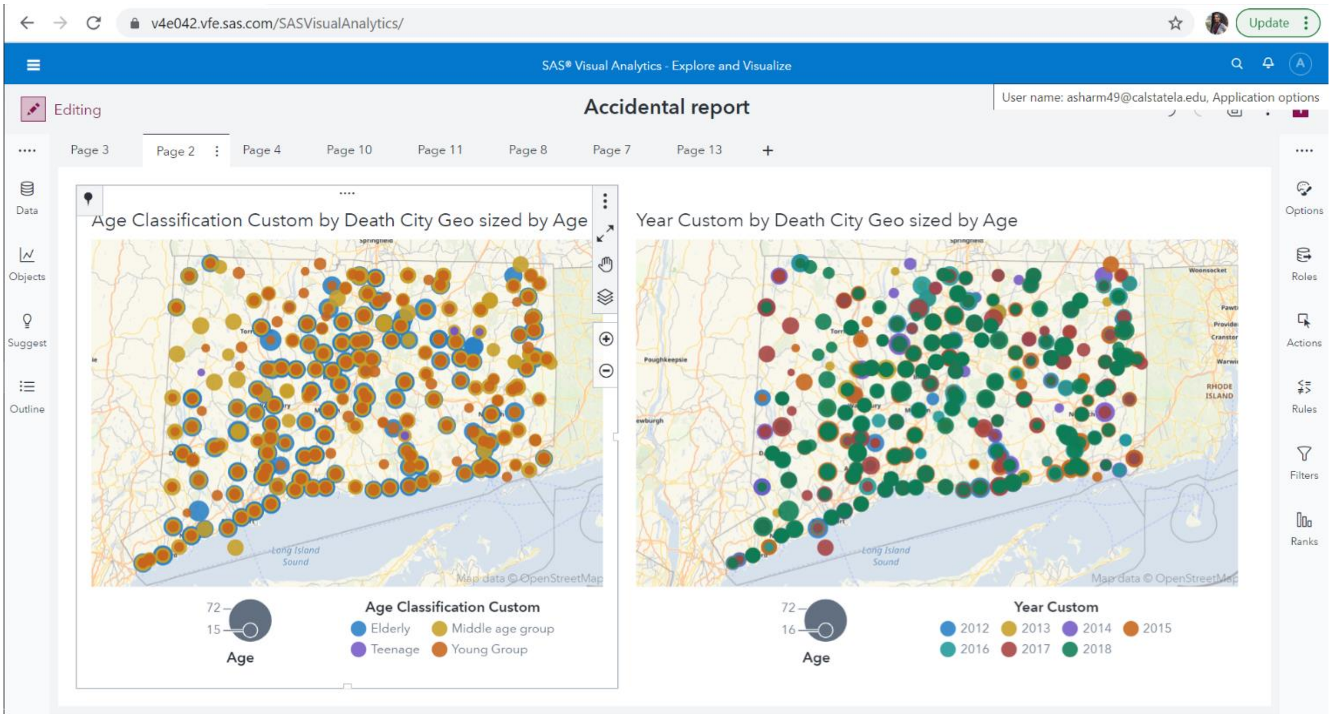
- **What are the factors behind the injury's vs the injury city?**

This Visualization provides us the details with regards to number Injuries caused versus the contributing factors. We have observed the top 10 factors which has led to injuries, with **Abused Medication being the highest contributor**. These factors can be further utilized for future studies to understand the reasons for such behavior.



- **What age group has experienced a higher death rate due to overdose?**

This Visualization provides us the description of number of deaths vs the Age groups impacted. The **Young group (21 to 30)** leads this as observed in the report, followed by **Middle aged group (36 to 55)**. This gives us indication that Age is one of the key factors and certain groups can be targeted for sensitized regarding the harmful effects of drugs.



- **Which year has seen a higher death rate?**

  This visualization provides the Number deaths in the year versus the Age Group in various cities. The year **2018** leads the sample set with mostly Middle-Aged people between the age of **37.3 & 52 years from Hartford, New Britain and Waterbury**.

The above models can be used by doctors or prescribers for individual risk management measures of drugs before they administer them for treatment.

The drug overuse disorder model can be helpful in predicting patients age group who are vulnerable to drug overdose and therefore, the model's results could be administered to vulnerable patients upon an initial visit prior to beginning for any kind of prescription. This approach will be helpful to reduce morbidity and mortality associated with several drug abuse. The risk stratification model can be integrated into a clinical method to easily identify high risk patients and provide physicians with real-time pre or post prescription decision-making alerts. These models can enable clinicians to take the required action and measures as per the risk level of the patient. For instance, recognize high-risk patients could warn the physician that the patient with what age group, race or sex would need what kind of care[4].

Key Findings:

- *In 2018, there were 67,367 drug overdose deaths in the United States, a 4.1% decline from 2017 (70,237 deaths).*
- *The age-adjusted rate of drug overdose deaths in 2018 (20.7 per 100,000) was 4.6% lower than in 2017 (21.7).*
- *For 14 states and the District of Columbia, the drug overdose death rate was lower in 2018 than in 2017.*
- *The rate of drug overdose deaths involving synthetic opioids other than methadone (drugs such as fentanyl, fentanyl analogs, and tramadol) increased by 10%, from 9.0 in 2017 to 9.9 in 2018.*
- *From 2012 through 2018, the rate of drug overdose deaths involving cocaine more than tripled (from 1.4 to 4.5) and the rate for deaths involving psychostimulants with abuse potential (drugs such as methamphetamine) increased nearly 5-fold (from 0.8 to 3.9).*

[4] Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G. Drug and Opioid-Involved Overdose Deaths – United States, 2013-2017. WR Morb Mortal Wkly Rep. ePub: 21 December 2018

MOTIVATION

Drug Interactions Can Be Deadly

I was motivated to create this report since the usage of Drugs and Opioid is an epidemic among many people from various walks of life, age, and places. Many have lost their loved ones over course of years due to this plethora of problems. Even celebrated personalities are not spared, whether it was Michael Jackson, Whitney Huston, Amy Winehouse to Lance Armstrong who were heroes and inspiration to communities, a beacon of light to achieve what many couldn't only to falter and become dependent on usage of these chemicals.

Oscar-winning actor Heath Ledger, 28.

Michael Jackson, the king of pop, 50.

Acclaimed musician and artistic innovator Prince, 57[5].

> ***"These cultural icons share a tragic connection: They all died too young from accidental overdoses of prescription painkillers – mostly synthetic opioids – or a cocktail of prescription drugs."***

The tragic loss of a renowned celebrity or athlete to a drug or alcohol-related death is a happening that often leaves many unanswered questions. While many deceases may be due to misuse of illegal constituents, these drugs are not always illicit - and an overdose may not always be intentional.

[5] **https://www.sas.com/hu_hu/insights/articles/risk-fraud/prescription-drug-monitoring.html**

CONCLUSION

Bringing the data together helps the numerous communities to work together. It helps them see and understand a bigger picture, as well as illuminate and emphasis on the outcomes we want. Everyone is aligned around **the conclusion of saving lives**, but there are different approaches and combinations. Reducing the accessibility of illegal street drugs would take a massive implementation strategy. Any approach should start at the root of the problem, for example patients getting prescriptions for opioids. We need to identify and study the patterns and spot the gaps in the system, train physicians on better ways to evaluate pain and prescribing the most efficient solution. “What’s the best outcome for patient with this type of agony? How has that consequence been achieved for other similar patients? If this is an issue of addiction or other types of difficulties, how can we get the patient plugged into the correct treatment options?” As organizations get better at collecting and managing the data, they should invest in automating that process.

Educate physicians and institutions across the health care system. Start with presenting the data, because when they see that they are adverse outliers in terms of methods and results, they change behavior. Use analytics to prepare the data. Automate the generation of reports this allows the analysts and medical informatics staff to spend less time working on the data itself and more time enabling and encouraging its use with predictive modeling and what-if scenario capabilities.