TECHNICAL REPORT FOR

ACCIDENTAL DEATHS BY DRUG OVERDOSE 2012-2018

T. SOWNDARYA LAHARI

G01229147

Accidental Drug Related deaths 2012-2018

Our System,**Accidental Drug related Deaths 2012-2018** is a listing of deaths associated with drug overdose Of the Usa between the years 2012 to 2018.

The source system is an excel file which consists of the person details like person id,resident city,manner of Death,Type of Drug they used while they died,city of death,their gender race etc and the reason of their death due to different reasons of drug usage.

Data in the source system file is derived from the website data.gov .A

“Y” value under differ types of drug names indicates the substance/drug was detected in the patient.

Note:The dataset is intended for public access and use by the US government and no license information was provided.

Data Dictionary:

ID	**int**
Date	**date**
Datatype	**nvarchar**
Age	**float**
sex	**varchar**
Race	**varchar**
Residence City	**varchar**
Residence county	**varchar**
Residence State	**varchar**
Death city	**varchar**

Death county	**varchar**
Location	**varchar**
LocationIf Other	**varchar**
DescriptionOf Injury	**varchar**
Injury place	**varchar**
Injury city	**varchar**
Injury county	**varchar**
Injury state	**varchar**
cod	**varchar**
Other significant	**varchar**
heroin	**varchar**
cocaine	**varchar**
Fentanyl	**varchar**
Fentanyl Analogue	**varchar**
Oxycodone	**varchar**
Oxymorphone	**varchar**
ethanol	**varchar**
hydrocodone	**varchar**
Benzodiazepine	**varchar**

methadone	**varchar**
Amphet	**varchar**
Morphine Not heroin	**varchar**
tramadol	**varchar**
Hydromorphone	**varchar**
Other	**varchar**

Reason for selecting the data set:

The reasons for selecting the data set

-> to analyse the pattern of the drug usage according to the different categories of race ,gender and age.

-> To analyse the changing pattern of the drug usage according to the year.

Section 3:

- Basic introduction

The data is based on Accidental Drug Deaths from 2012-2018 .The columns include Id,date ,Age ,Sex ,Race,ResidenceCity,ResidenceCounty,ResidenceState etc. The data is of which drug is used by the people in the state , county,City depending on the age ,sex and Race.There are different types of drugs like Morphine,Heroin,hydromorphone,Benzodiazepine,Methadone,Amphet,Tramad,Morphine_not heroin,Other,Manner of death.

The Accidental drug deaths data consists of all the categorical data and there will be no statistics majority.The only numerical data present in the data is age .

The data consists of 41 columns and 5106 rows.

- **Statement of Purpose**

The purpose of the dataset is to show the drugs usage and the people of different age groups addicted to drugs and The way the dead .The people who belong to different cities and states and counties.

The purpose of the project is to project is to identify

->Deaths according to age group.

->Deaths according to city

-> The top 10 number of total Deaths in each city

->Accidental drug deaths by Gender

->Accidental Drug deaths by Race

-> number of deaths by age

The significance is known by plotting the graphs with the data as mentioned above.So there is an increase in the deaths in Hartford city and there are more deaths in the white people, more deaths are in the people in the age of 50 -60 years.

Section 4:

Research question(s) or hypotheses

So these are the research questions i want to do and i have done visualization

The Deaths according to age group is visualized and.Deaths according to city The top 10 number of total Deaths in each city ,Accidental drug deaths by Gender,Accidental,Drug deaths by Race, number of deaths by age

The statistical analysis is done in r

section 5: Review of Literature

Overview:

Investigation of the exacting content from death declarations can be utilized to recognize designs in the particular medications most every now and again engaged with sedate overdose passings. From 2010 through 2014, the best 10 medications included were the equivalent, however the relative positioning and age-balanced rates for passings including these medications changed. Exacting content examination likewise uncovered that many medication overdose passings included numerous medications. Discoveries ought to be deciphered considering the improvement in the nature of the information that came about because of better announcing of explicit medications on death testaments from 2010 through 2014. Relative increments in the passing rates including explicit medications and the rankings of these medications might be influenced by enhancements in revealing, genuine increments in the quantities of death, or both.Keywords: narcotic • benzodiazepine • energizer • demise declaration • National Vital Statistics SystemIntroductionFrom 1999 through 2014, the age-balanced rate for sedate overdose passings dramatically increased, from 6.1 per 100,000 populace in 1999 to 14.7 in 2014 (1). Various investigations utilizing National Vital Statistics System mortality information (NVSS–M) have educated the country about this developing general wellbeing concern (1–4). The International Classification of Diseases, Tenth Revision (ICD–10), which is the grouping framework right now used to order the basic and various reasons for death, is restricted Drugs Most Frequently Involved in Drug Overdose Deaths: United States, 2010–2014by Margaret Warner, Ph.D., National Center for Health Statistics; James P. Trinidad, M.P.H., M.S., U.S. Food

and Drug Administration; Brigham A. Bastian, B.S., Arialdi M. Miniño, M.P.H., and Holly Hedegaard, M.D., M.S.P.H., National Center for Health Statistics

Introduction:

From 1999 through 2014, the age-balanced rate for sedate overdose passings dramatically increased, from 6.1 per 100,000 populace in 1999 to 14.7 in 2014 (1). Different investigations utilizing National Vital Statistics System mortality information (NVSS–M) have educated the country about this developing general wellbeing concern (1–4). The International Classification of Diseases, Tenth Revision (ICD–10), which is the characterization framework right now used to order the fundamental and numerous reasons for death, is constrained Drugs Most Frequently Involved in Drug Overdose Deaths: United States, 2010–2014by Margaret Warner, Ph.D., National Center for Health Statistics; James P. Trinidad, M.P.H., M.S., U.S. Food and Drug Administration; Brigham A. Bastian, B.S., Arialdi M. Miniño, M.P.H., and Holly Hedegaard, M.D., M.S.P.H., National Center for Health Statistics.

Results were accounted for as numbers, rates, or rates for each medication of intrigue, alluded to as the referent drug.Age-balanced passing rates were determined utilizing the immediate technique and the 2000 standard U.S. populace (15). Patterns in death rates and number of passings were assessed utilizing the JoinPoint Regression Program (19). Any notice of a normal percent change (APC) in this report demonstrates a factually critical .

From 2010 through 2014, the quantity of medication overdose passings every year expanded 23%, from 38,329 of every 2010 to 47,055 of every 2014 . During this equivalent timeframe, the particularity of medication data in the exacting content fields of the demise authentication improved. There was a consistent increment in the level of medication overdose passings including in any event one explicit medication, from 67% in 2010 to 78% in 2014. There was a decrease in the level of passings including just a medication class (e.g., "Narcotic") yet not a particular medication (from 4% of medication overdose passings in 2010 to 3% in 2014). An audit of the strict content for these passings demonstrated that the referenced class was either a narcotic or sedative (going from 69% in 2014 to 74% in 2011). There were additionally

consistent decreases in the level of passings without notice of a particular medication or medication class, from 29% of medication overdose passings in 2010 to 19% in 2014. These passes every now and again included numerous drugs(e.g., notices of "POLYPHARMACY," "MULTIDRUG," and "Medications").

The Overdose:

the quantity of medication overdose passings in 2014 by the quantity of explicit medications associated with the demise . Of the 36,667 medication overdose passings with in any event one notice of a particular medication, 52% referenced just a single explicit medication (18,931 passings), 26% referenced two (9,351 passings), 12% referenced three (4,521 passings), 6% referenced four (2,041 passings), and 5% referenced at least five (1,823 passings). Among sedate overdose passings with at any rate one notice of a particular medication, the normal number of explicit medications referenced was 1.9. Table C shows the level of medication overdose passings with corresponding medications for tranquilize overdose passings including the main 10 medications in 2014. The level of passings including accompanying medications differed by referent medication. For instance, most of the medication overdose passings including methamphetamine didn't include different medications. Conversely, among passings including alprazolam and diazepam, over 95% included different medications. The normal number of attendant medications included (barring the referent medication) likewise changed among the main 10 medications associated with tranquilize overdose passings. For instance, tranquilize overdose passings including diazepam or alprazolam had on normal in excess of two extra medications associated with death. Medication overdose passings including fentanyl, heroin, cocaine, or methamphetamine had on normal less than two extra medications engaged with death. The percent dispersion of the quantity of accompanying medications for overdose passings including the main 10 medications in 2014 . For instance, for tranquilize overdose passings including methamphetamine, 55% had no attending makes reference to, 25% referenced one other medication, 18% referenced two to four different medications, and 1% referenced at least five medications. Interestingly, for sedate overdose passings including diazepam, 3% .

CHEMICALS CATEGORIES:

Drugs, substances, and certain chemicals used to make drugs are classified into five (5) distinct categories or schedules depending upon the drug's acceptable medical use and the drug's abuse or dependence potential. The abuse rate is a determinant factor in the scheduling of the drug; for example, Schedule I drugs have a high potential for abuse and the potential to create severe psychological and/or physical dependence. As the drug schedule changes .Schedule II, Schedule III, etc., so does the abuse potential-- Schedule V drugs represent the least potential for abuse. A Listing of drugs and their schedule are located at Controlled Substance Act (CSA) Scheduling or CSA Scheduling by Alphabetical Order. These lists describe the basic or parent chemical and do not necessarily describe the salts, isomers and salts of isomers, esters, ethers and derivatives which may also be classified as controlled substances. These lists are intended as general references and are not comprehensive listings of all controlled substances.

Discussion:

In the United States, medicated overdoses brought about 702,568 passings during 1999–2017, with 399,230 (56.8%) including narcotics. From 2016 to 2017, demise rates from all narcotics expanded, with increments driven by manufactured narcotics. Passings including IMF have been seen fundamentally east of the Mississippi River; in any case, ongoing increments happened in eight states west of the Mississippi River, including Arizona, California, Colorado, Minnesota, Missouri, Oregon, Texas, and Washington.

Medication overdose demise rates from 2013 to 2017 expanded in many states; the impact of engineered narcotics on these rate increments was seen in roughly one fourth of all states during this equivalent 5-year time frame. Overdose passings including cocaine and psychostimulants additionally have expanded as of late (1,6). Generally speaking, the overdose scourge keeps on exacerbating, and it has become progressively mind boggling by co-contribution of remedy and illegal medications (7,8). For instance, in 2016,

manufactured narcotics (fundamentally IMF) were engaged with 23.7% of passings including solution narcotics, 37.4% including heroin, and 40.3% including cocaine (9). Likewise, demise rates are expanding over numerous segment gatherings. For instance, in spite of the fact that passing rates including narcotics stayed most noteworthy among whites, moderately enormous increments over a few medication classes were seen among blacks.

The discoveries in this report are liable to in any event five restrictions. To start with, at dissection, substances tried for fluctuate by time and locale, and enhancements in toxicologic testing may represent some revealed increments. Second, the particular kinds of medications included were excluded from 15% of medication overdose demise endorsements in 2016 and 12% in 2017, and the level of death testaments with at any rate one medication indicated ran among states from 54.7%–99.3% in 2017, restricting rate correlations between states. Third, since heroin and morphine are processed also (10), some heroin passings may have been misclassified as morphine passings, bringing about underreporting of heroin passings. Fourth, potential race misclassification may have prompted belittles for specific classifications, principally for American Indian/Alaska Natives and Asian/Pacific Islanders.Finally, most state-explicit examinations were confined to DC and a subset of states with satisfactory medication explicitness, restricting generalizability.

Through 2017, the medication overdose pestilence keeps on intensifying and develop, and the inclusion of numerous sorts of medications (e.g., narcotics, cocaine, and methamphetamine) underscores the desperation to acquire all the more opportune and neighborhood information to advise general wellbeing and open security activity. Albeit solution narcotic and heroin-included passing rates were steady from 2016 to 2017, they

stayed high. Some primer markers in 2018 point to potential upgrades dependent on temporary data; be that as it may, affirmation will rely upon aftereffects of pending clinical examinations and investigation of conclusive information. By and large, passings including engineered narcotics keep on driving increments in overdose passings. CDC finances 32 states and DC to gather all the more convenient and thorough medication overdose information, remembering improved toxicologic testing for narcotic included lethal overdoses. CDC is subsidizing counteraction exercises in 42 states and DC. CDC likewise is utilizing crisis subsidizing to help 49 states, DC, and four domains to expand their observation and reaction abilities and empower thorough network level reactions with execution of novel, proof based intercessions. Proceeded with endeavors to guarantee safe endorsing rehearses by following the CDC Guideline for Prescribing Opioids for Chronic Pain are upgraded by access to nonopioid and nonpharmacologic medicines for torment. Other significant exercises incorporate expanding naloxone accessibility, extending access to medicine, helping treatment, upgrading general wellbeing and open security organizations, and amplifying the capacity of wellbeing frameworks to connect people to treatment and mischief decrease administrations.

Medications, substances, and certain synthetic concoctions used to make drugs are characterized into five (5) particular classes or timetables relying on the medication's satisfactory clinical use and the medication's maltreatment or reliance potential. The maltreatment rate is a determinant factor in the planning of the medication; for instance, Schedule I tranquilizes have a high potential for misuse and the possibility to make extreme mental and additionally physical reliance. As the medication plan changes- - Schedule II, Schedule III, and so forth., so does the maltreatment potential- - Schedule V drugs speaks to minimal potential for misuse. A Listing of medications and their timetable are situated at Controlled Substance Act (CSA) Scheduling or CSA Scheduling by Alphabetical Order. These rundowns portrays the essential or parent synthetic and don't really depict the salts, isomers and salts of isomers, esters, ethers and subsidiaries

which may likewise be named controlled substances. These rundowns are planned as general references and are not exhaustive postings of every single controlled substance.

If it's not too much trouble note that a substance need not be recorded as a controlled substance to be treated as a Schedule I substance for criminal arraignment. A controlled substance simply is a substance which is expected for human utilization and is fundamentally or pharmacologically considerably like or is spoken to as being like a Schedule I or Schedule II substance and isn't an affirmed prescription in the United States. (See 21 U.S.C. 802(32)(A) for the meaning of a controlled substance simple and 21 U.S.C. 813 for the timetable.)

Calendar I

Calendar I medications, substances, or synthetic substances are characterized as medications with no right now acknowledged clinical use and a high potential for misuse. A few instances of Schedule I tranquilizes are:

heroin,lysergic,corrosive,diethylamide(LSD),pot(cannabis),
3,4-methylenedioxymethamphetamine (joy), methaqualone, and peyote

Calendar II

Calendar II medications, substances, or synthetic compounds are characterized as medications with a high potential for misuse, with utilize possibly prompting serious mental or physical reliance. These medications are likewise viewed as perilous. A few instances of Schedule II drugs are:

Mix items with under 15 milligrams of hydrocodone per measurement unit (Vicodin), cocaine, methamphetamine, methadone, hydromorphone (Dilaudid), meperidine (Demerol), oxycodone (OxyContin), fentanyl, Dexedrine, Adderall, and Ritalin

Timetable III :

Timetable III medications, substances, or synthetic compounds are characterized as medications with a moderate to low potential for physical and mental reliance. Timetable III medications misuse potential is not as much as Schedule I and Schedule II tranquilizes however more than Schedule IV. A few instances of Schedule III medications are:

Items containing under 90 milligrams of codeine for every dose unit (Tylenol with codeine), ketamine, anabolic steroids, testosterone

Calendar IV :

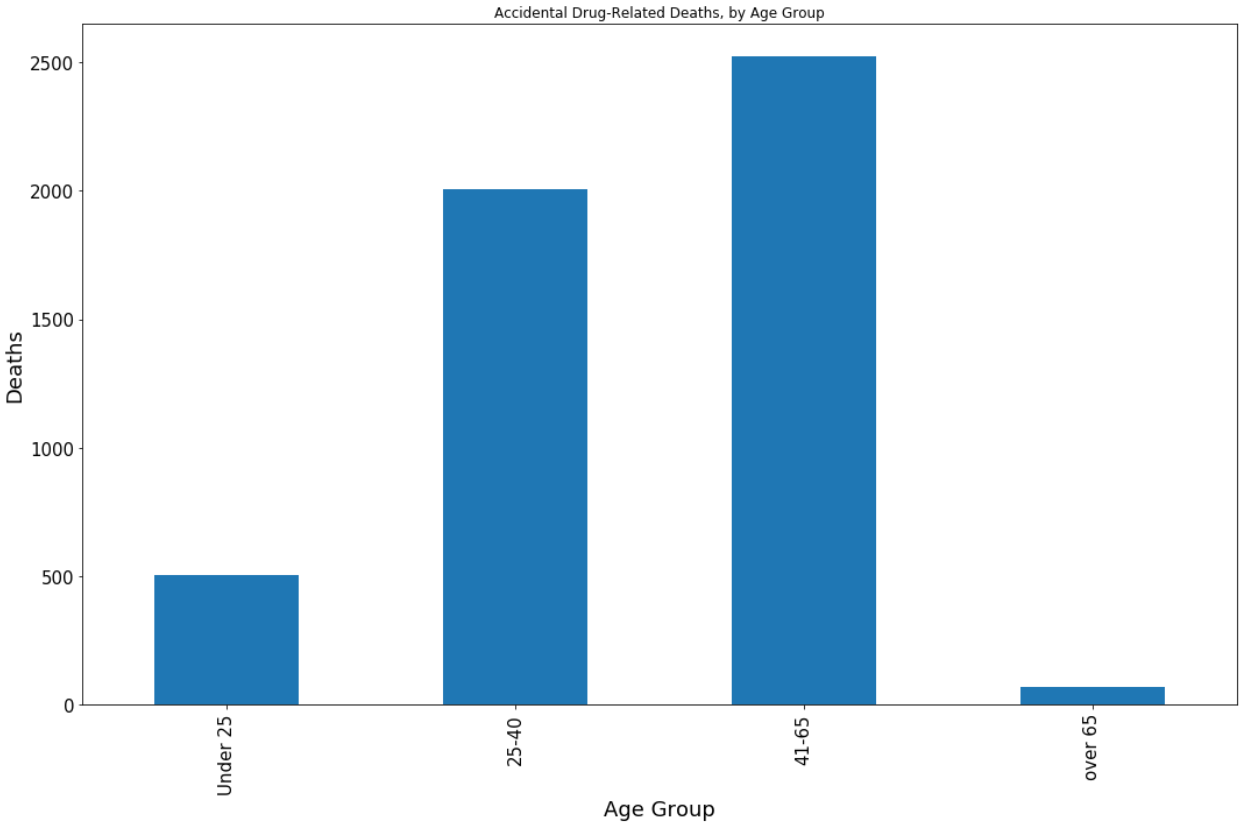
Calendar IV medications, substances, or synthetics are characterized as medications with a low potential for misuse and generally safe of reliance. A few instances of Schedule IV drugs are:

Calendar V

Calendar V medications, substances, or synthetic substances are characterized as medications with lower potential for maltreatment than Schedule IV and comprise of arrangements containing restricted amounts of specific opiates. Timetable V drugs are commonly utilized for antidiarrheal, antitussive, and pain relieving purposes. A few instances of Schedule V drugs are:

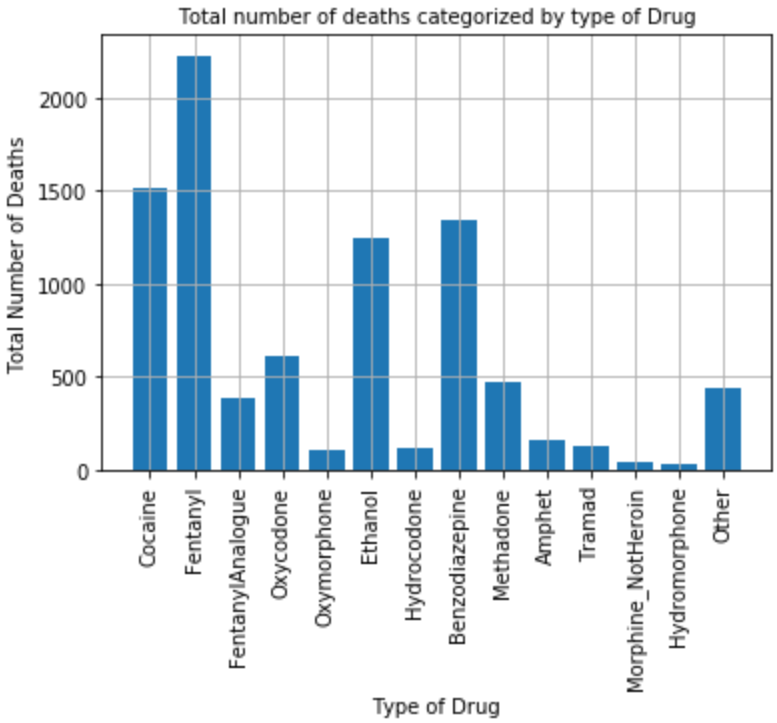
hack arrangements with under 200 milligrams of codeine or per 100 milliliters (Robitussin AC), Lomotil, Motofen, Lyrica, Parepectolin

Section 6: Methods



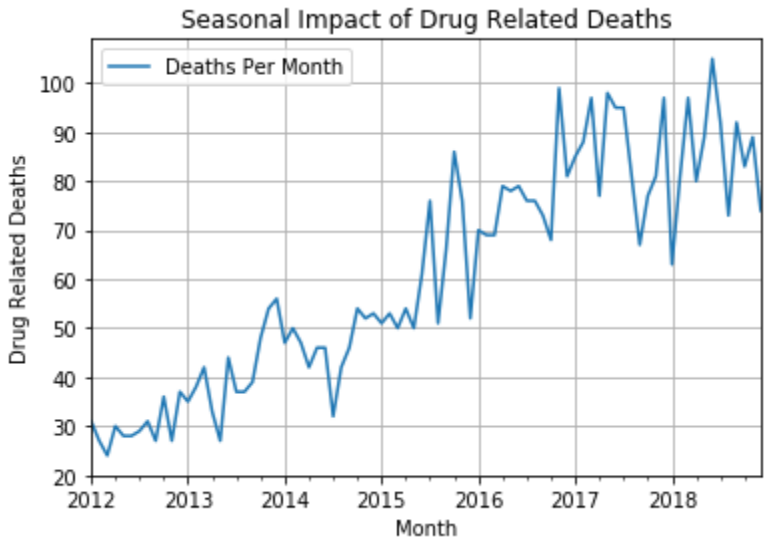
The plot describes the Accidental drug related deaths by age group . The age group of 41-65 are more in number of deaths by the age group and the medium number of deaths are recorded in the age group of 25-40 years. The x-axis is represented by the age group and y axis is represented by the deaths.

So i observe that there are less number of people over 65 who are used to drugs and deaths.

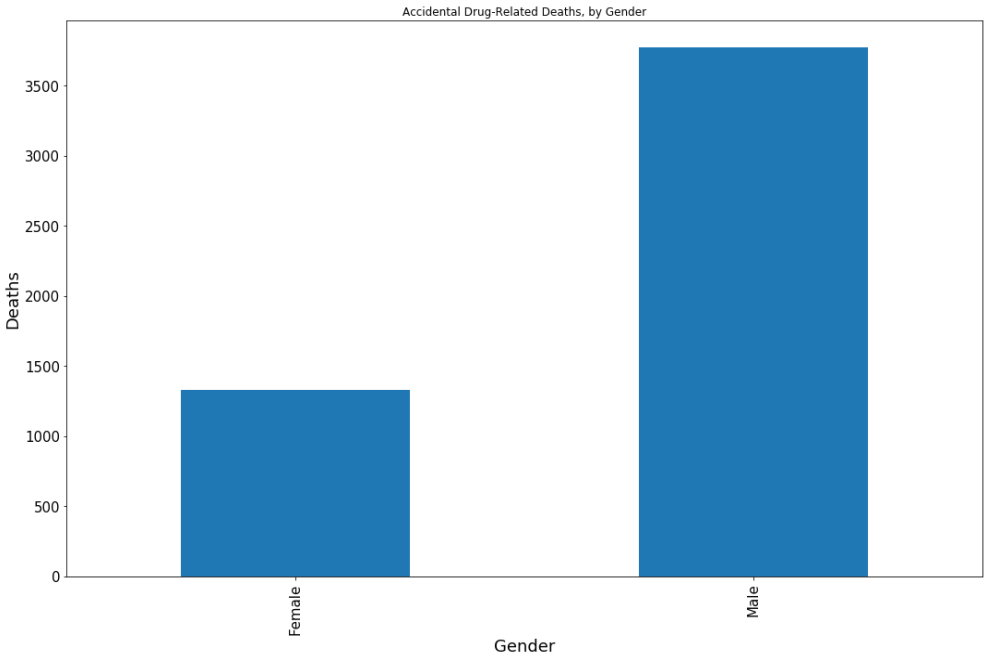


Description: This plot describes the Total number of deaths categorized by type of Drug .The X- axis is plotted on Type of Drug and Y axis is plotted on total number of deaths. The plot is done on the different type of drugs like Cocaine,Fentanyl,FentanylAnalogue,Oxycodone,Oxymorphone,Ethanol,Hydrocodone,Benzodiazepine,Methadone,Amphet,Tramad,Morphine_NotHeroin,Other.

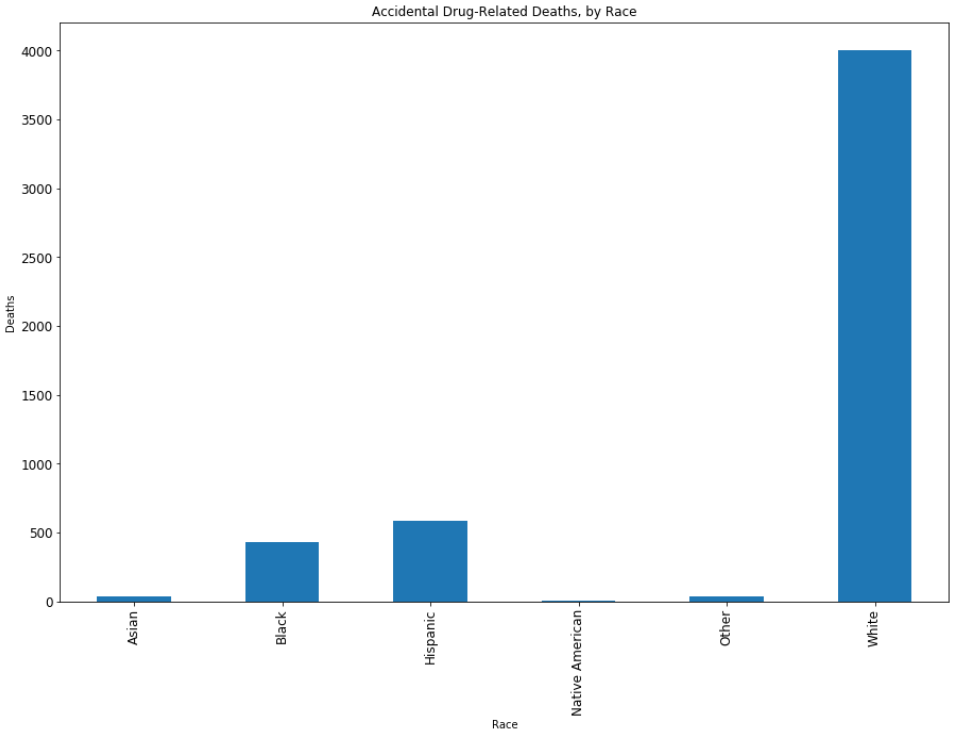
The total number of deaths by the Drug Fentanyl are more as plotted in the graph.The less count is for Morphine_NotHeroin ,Tramad as observed.



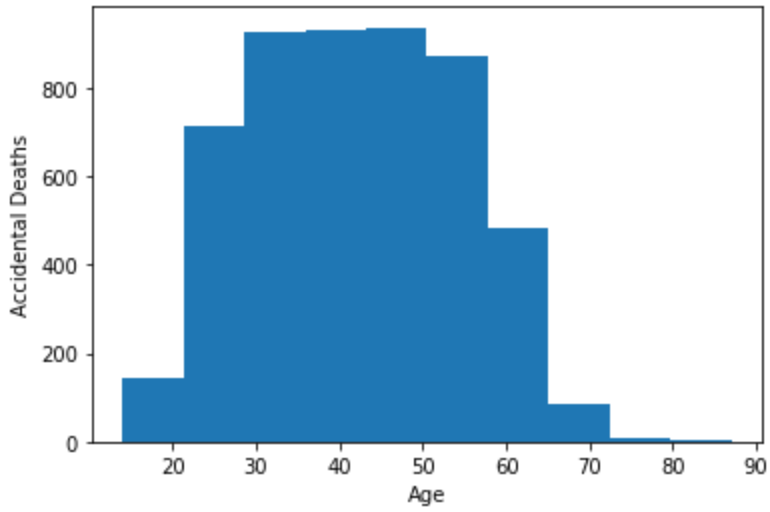
The Seasonal Impact of Drug Related Deaths based on months by deaths. I have taken start date and end dates and plotted the graph.The graph is fluctuating and it had raisin in the year 2016 and 2018.The X- axis is taken as month and Y axis Drug related Deaths.



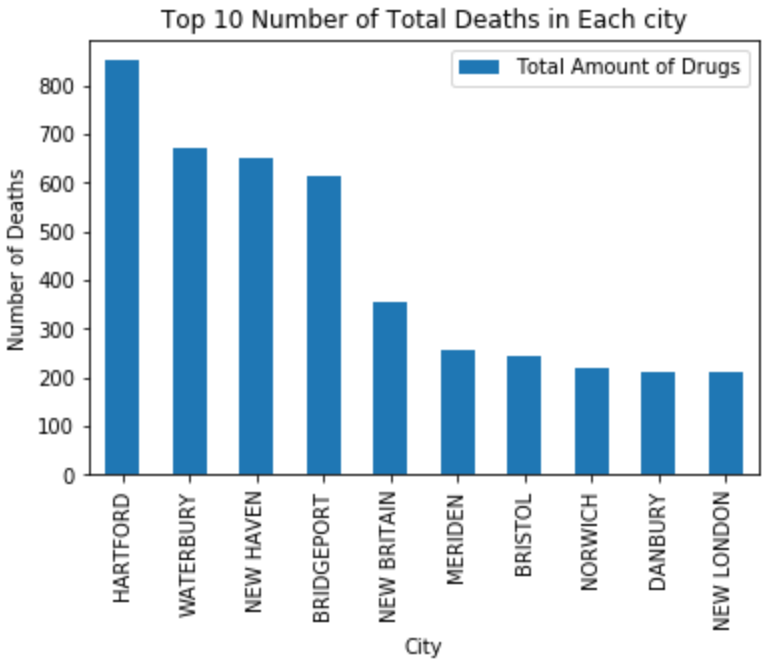
The plot done on the Gender and Deaths . I observed that the more deaths are occured in males compared with females. So I concluded that males are more addicted to the drugs than females.



The plot describes the Race and death .I observed that the more deaths are occured in the white race and less deaths are occured in Native american and asian and other races.The X axis is plotted for the Race and Y axis is plotted for the death count.

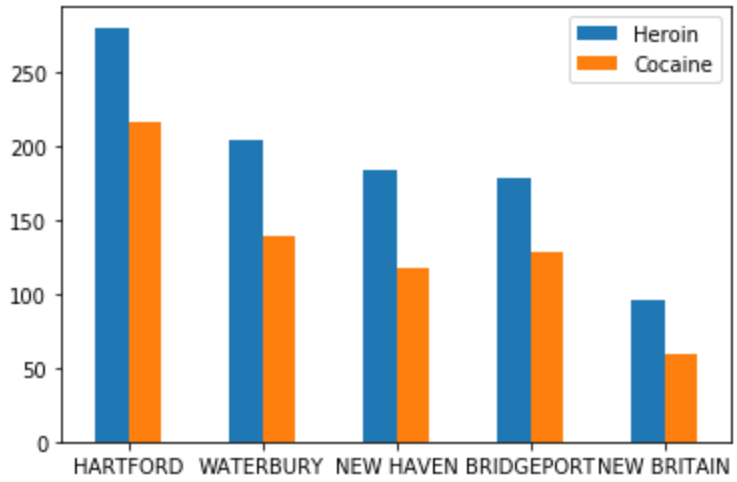


The plot describes the age and the accidental deaths .The more deaths occured in the age of 40 ,50 to 60 years people . So I observed that the age group of 40 -50and 50-60 year people are getting addicted to drugs. The least are the 20 year age group.



The graph describes the city and number of deaths. The highest number of deaths are recorded HartFord and the lowest is recorded in New london.

The total amount of drugs used by the people in the waterbury and New heaven are in the amount .So they are on the same level of the plot . The x axis represents the city and the Y axis is represented by the number of deaths.



The bar graph is plotted by taking the Heroin and cocaine values in the different places.The Hartford place had the Highest count in the usage and similarly the cocaine

usage is more in that place . The Waterbury is the second place .The least is recorded in Britain with 100 and 50 differences between the two drugs.

The R code:

I have used Library of Mass

```
library(MASS)



library(class)
#Reading the csv
data1<- read.csv("accident.csv",header= TRUE)
```

The data1 is assigned to the BD variable

```
BD<- data1
#omitting na values
na.omit(data1)
data1
modell <- glm(data=data1,Age ~ Race+Sex+DeathCity+Location )
modell
summary(modell)

modelll <- glm(data=data1,Age ~ Race+Sex+DeathCity )
modelll
summary(modelll)

confint.lm(modell)



exp(coef(modell))
```

```
exp(cbind(OR = coef(modell), confint.lm(modell)))
```

```
dim(data1)
```

```
summary(data1)
```

```
(Dispersion parameter for gaussian family taken to be 147.6656)

    Null deviance: 776335  on 5101  degrees of freedom
Residual deviance: 718393  on 4865  degrees of freedom
  (3 observations deleted due to missingness)
AIC: 40196

Number of Fisher Scoring iterations: 2
```

```
> confint.lm(modell)
                                2.5 %     97.5 %
(Intercept)                 13.8343484 59.9220757
RaceAsian Indian           -21.9091566 -2.1712156
RaceAsian, Other           -21.6377897 -2.8022258
RaceBlack                   -8.7009680  6.4606387
RaceChinese                -44.9195898 -6.0878090
RaceHawaiian               -14.0524204 35.5427161
RaceHispanic, Black        -13.7328582  4.1410074
RaceHispanic, White        -12.7372471  2.3836577
RaceNative American, Other -26.4466761 23.2825826
RaceOther                  -19.8013474  0.9434369
RaceUnknown                -16.8332680  1.1669567
RaceWhite                  -13.3696847  1.6431187
SexFemale                  -15.4291999 12.0291843
SexMale                    -16.0449766 11.3938261
SexUnknown                 -10.1290552 45.3096218
DeathCity06340              -5.7332352 52.4111941
DeathCityAMSTON            -28.6260211 29.4909654
DeathCityANDOVER           -11.8490419 31.6302745
DeathCityANSONIA            -8.7629043 26.4957075
DeathCityASHFORD           -11.1449013 27.8326412
DeathCityAVON              -21.2836215 18.6139214
DeathCityBAKERSVILLE       -25.7332352 32.4111941
DeathCityBALTIC            -17.2146715 26.2669090
DeathCityBANTAM            -24.6260211 33.4909654
DeathCityBARKHAMSTED       -28.7332352 29.4111941
DeathCityBEACON FALLS       -6.5620695 34.7242395
DeathCityBERLIN             -1.7007644 33.9472690
DeathCityBETHANY           -17.0193490 22.9798064
DeathCityBETHEL            -11.2574992 24.7350985
DeathCityBETHLEHEM          -6.6260211 51.4909654
DeathCityBLOOMFIELD         -4.1342203 32.8757081
DeathCityBOLTON             -4.2137802 43.3596643
```

Section 7: Results

The Data is cleaned in python and plotted for different data columns .The data is visualized according to the data by age and gender and several factors are taken into consideration.

By visualization i came to know that the people in the age 40-50 ,50-60 years are habituated to drugs . Compared to females the males are the most in number to be addicted to different types of Drugs.

By the Gender the male are the highest in number to die by the addiction of Drugs. The whites are the mainly addicted to drugs so they are more in number in the deaths by the overdose of the drug.

The Usage of Sql:

The Sql is used for the data to write some queries:

The SQL DEVELOPER is used to write the sql part

First I have created the table named drugs12

```
CREATE TABLE DRUGS12(
Age     INT,
Sex       varchar(255),
Race     varchar(255),
DeathCity        varchar(255),
COD     varchar(255),
Heroin  varchar(255),
Cocaine          varchar(255),
Fentanyl varchar(255),
FentanylAnalogue varchar(255),
Oxycodone       varchar(255),
```

Oxymorphone varchar(255),

Ethanol varchar(255),

Hydrocodone varchar(255),

Benzodiazepine varchar(255),

Methadone varchar(255),

Amphet varchar(255),

Tramadol varchar(255),

Morphine_NotHeroin varchar(255),

Hydromorphone varchar(255),

Other varchar(255),

OpiateNOS varchar(255),

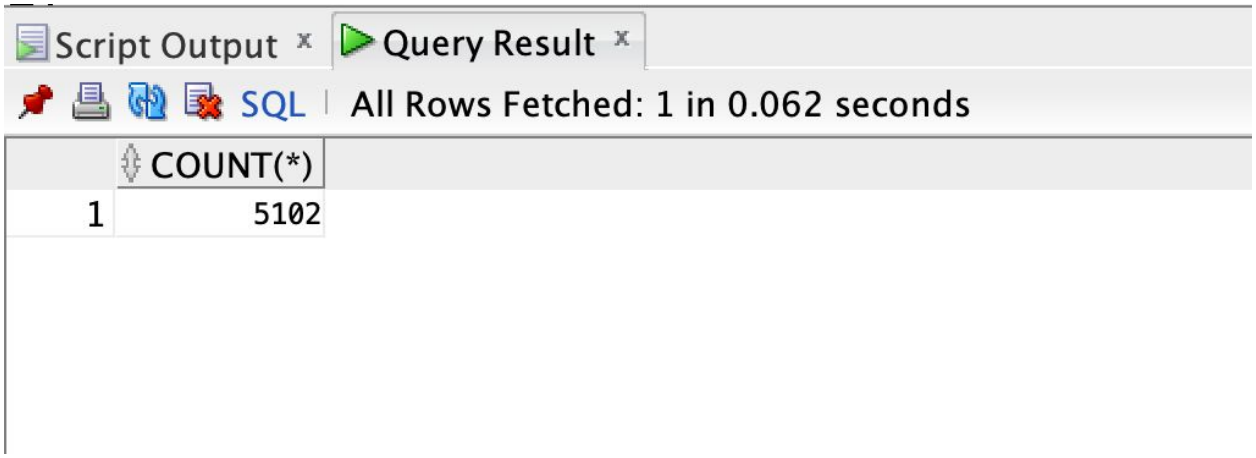
AnyOpioid varchar(255));

The table is created .

The second step is to import the dataset into the tabel. After successfully importing the csv .

Query 1:

select count(*) from drugs12;



Query 2:

The query displays the count of age where the age greater than 50.

select count(age)

from drugs12

where age > '50';

```
select count(*) from drugs12;

select count(age)
from drugs12
where age > '50';
```

ipt Output | Query Result

SQL | All Rows Fetched: 1 in 0.022 seconds

COUNT(AGE)
1451

Query 3:

The distinct function is used in this query.

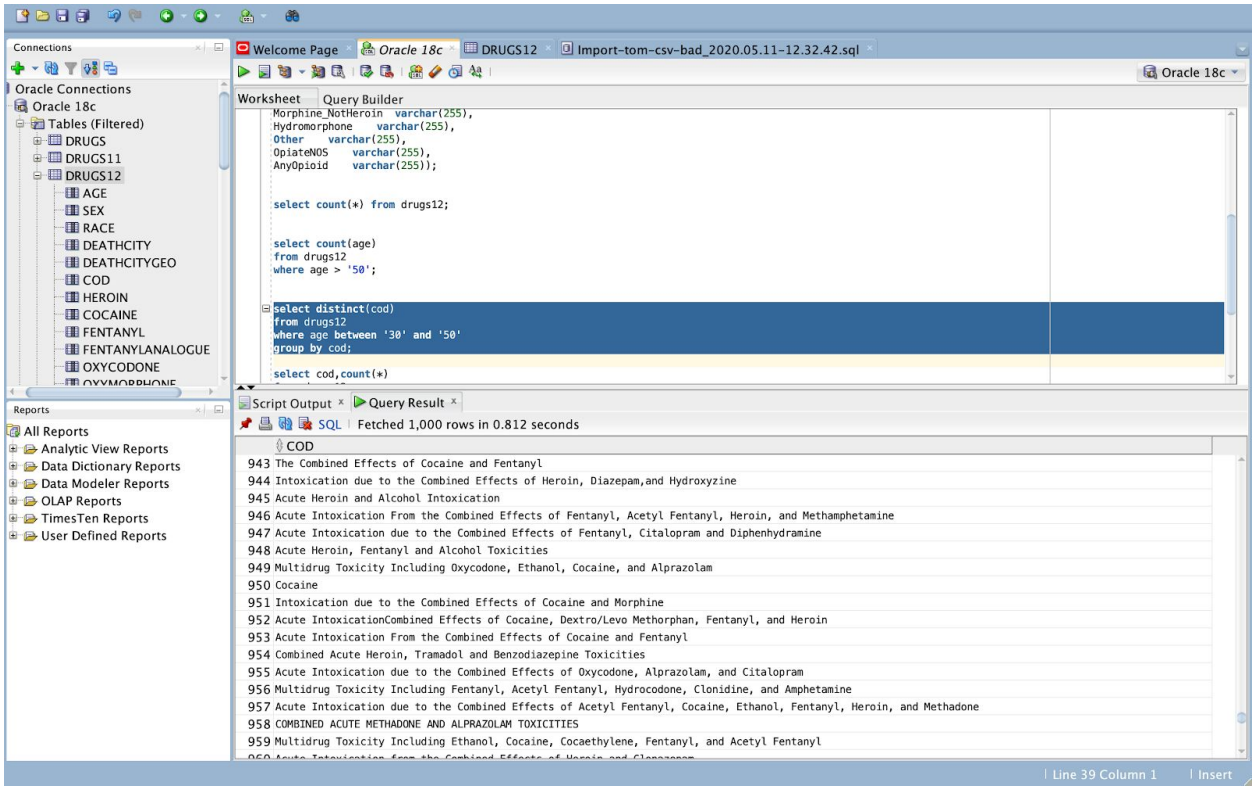
The query represents age between 30 and 50 and group by cod.

select distinct(cod)

from drugs12

where age between '30' and '50'

group by cod;



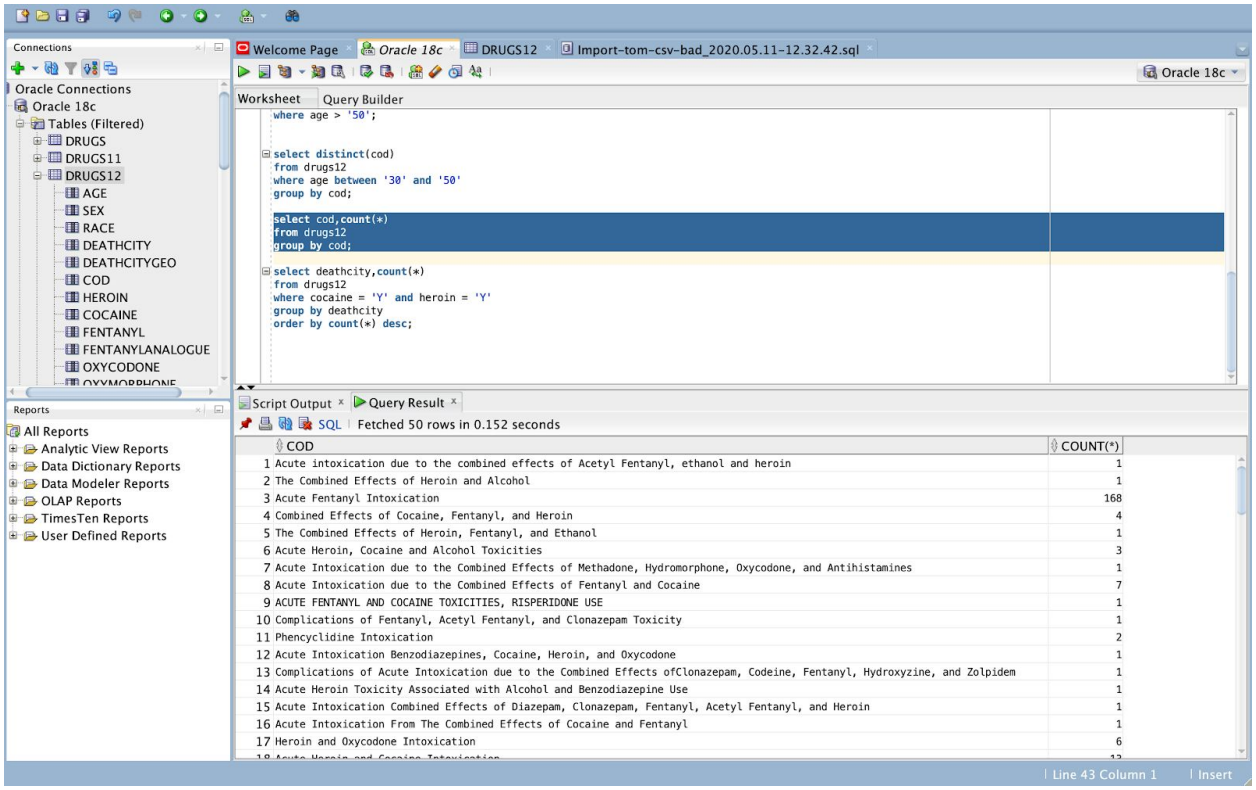
Query 4:

Here I used groupby function to get all the cod in group.

select cod,count(*)

from drugs12

group by cod;



Query 5:

In this query i have selected death city from the table where cocaine and heroin drug is used.

The count of the drugs where in different places are displayed.

select deathcity,count(*)

from drugs12

where cocaine = 'Y' and heroin = 'Y'

group by deathcity

order by count(*) desc;

```
select deathcity,count(*)
from drugs12
where cocaine = 'Y' and heroin = 'Y'
group by deathcity
order by count(*) desc;
```

ript Output | Query Result

SQL | Fetched 50 rows in 0.051 seconds

	DEATHCITY	COUNT(*)
1	HARTFORD	102
2	WATERBURY	79
3	BRIDGEPORT	57
4	NEW HAVEN	52
5	NEW BRITAIN	30
6	MERIDEN	24
7	NEW LONDON	19
8	NORWICH	19
9	DANBURY	18
0	TORRINGTON	18
1	ENFIELD	16
2	BRISTOL	16
3	MANCHESTER	15
4	MILFORD	13
5	MIDDLETOWN	13
6	STAMFORD	13
7	STRATFORD	12
8	WEST HAVEN	12

Section 8: Discussion

- **Restate Overall Research Question**

So, my data is about the accidental deaths of the people by the overdose of the drug .

I have taken the side to research things like The drug usage of people by gender. Drug usage by the county ,state. And, then the which age groups are mainly addicted by the drugs and deaths by the drug overdose.

The results i have observed is that The males are in majority in number to die due to drug overdose. The White race are the people who ate more effected by the death. The age group of 40-60 are more addicted to drugs.

Section 9: Recommendations (if applicable, generally related to practice)

I recommend that there should be a lot of data to predict the outcome and there are totally missing values and completely missing values in the data which is very difficult to analyze.

Section 10: Limitations

The limitations of the dataset is if there are many missing columns which is very difficult to interpret the result and The incomplete data takes lots of time to clean while interpreting the result and gives incomplete results where I have imputed the values in many columns.

So if there is complete data we can predict the results in an accurate way and best visualization.

Another backdrop is that the data set completely consists of categorical values where it made much difficult task to interpret the value in the column by the techniques

Reference:

Scholl, L., Seth, P., Kariisa, M., Wilson, N., & Baldwin, G. (2018, January 4). Drug and Opioid-Involved Overdose Deaths - United States, 2013-2017. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6334822/

Drug Scheduling. (n.d.). Retrieved from https://www.dea.gov/drug-scheduling

“Accidental Drug Related Deaths 2012-2018.” Data.gov. Publisher data.ct.gov, May 8, 2019. https://catalog.data.gov/dataset/accidental-drug-related-deaths-january-2012-sept-2015.