**BIST 5225-Final Project**

**Data Visualization on** **Accidental Drug-related Deaths via R**

***Name:***

***Zhenghao Yu***

**Contact Information:**

**Zhenghao.Yu@uconn.edu**

**Contents:**

* **Part 1: Aims**
* **Part 2: Summary and Inferences**
* **Part 3:** **Process and Graphic Descriptions:**
* **Part 4: Reference**

**1. Aims:**

The primary aim is to find the patterns of accidental drug-related deaths during 2012-2018 in Connecticut associated with people’s information related to their death.

Then, people can use obtained graphic information to prove or disprove assumptions and get new inferences.

**2. Summary:**

*\*software: Rstudio*

1. **Locations (Coordinates-related):** For both Males and Females, Accidental drug-related deaths during 2012-2018 in Connecticut concentrated to four big cities: Hartford, Waterbury, new Haven and Bridgeport, which could make sense since the population density is high in these four areas. Also, numbers of deaths in other cities were also high.

However, it is known that the population of Waterbury is lower than that of Stamford, but the number of accidental drug-related deaths of Waterbury is higher than that of Stamford. We can see that the population could not be the only reason why accidental drug-related deaths happened mostly on these four areas.

2. **Locations (Types):** Accidental drug-related deaths during 2012-2018 in Connecticut mostly happened in people’s residences (52.4%) and Hospital (31.9%).

3. **Races (Numbers):** Accidental drug-related deaths during 2012-2018 in Connecticut mostly happened three races: "White"78.4%, "Hipanic,White" 11% and "Black” 8.5%. Compared with those three races, the number of related deaths of other races is close to zero.

4. **Races& Sex:** For those three races, accidental drug-related deaths during 2012-2018 in Connecticut mostly happened to males rather than females.

5. Accidental drug-related deaths during 2012-2018 in Connecticut converged to people with ages of about 40 roughly, but the highest number of deaths are close to ages of about 28 and 50(two peaks).

6. **Races& Ages:** For Black people and White& Hispanic people, the number of accidental drug-related deaths converge between age 40-50. For White people, the number of deaths has two peaks on about age 30 and 50.

After analyzing the external data of population estimates in Connecticut by July 1st 2018, it is shown that among the ratios of accidental drug-related deaths in population of each race, the ratio of white people is way higher than that of Black people and that of White& Hispanic people.

7. (From 3&6) The overall trend of number of accidental drug-related deaths vs. age is similar to that of white people since the number of related deaths mostly happened to White people in 2012-2018 in Connecticut, which is 78.4% (obtained in pie chart).

8. **Comparison involvements of Heroin and Cocaine:** For people who were younger than 10 years old, the number of accidental drug-related deaths related to heroin and those related to cocaine were close. However, for people with ages of more than 10, the number of deaths related to cocaine was way higher than that related to Heroin.

The reason could be that cocaine had the higher accessibility than Heroin during 2012-2018 in Connecticut.

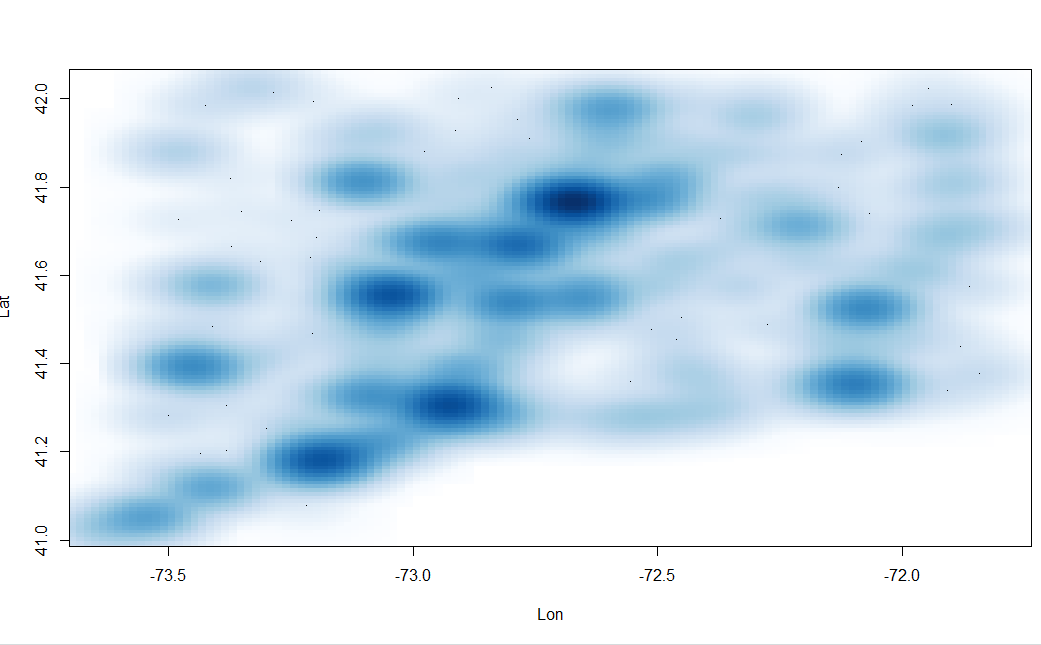
**3.** **Analysis and Graphic Descriptions:**

**Section1: Locations (Coordinates and types)**

**## Scatterplot for coordinates of the accidental drug-related deaths:**

**#scatterplot colored by smooth density**

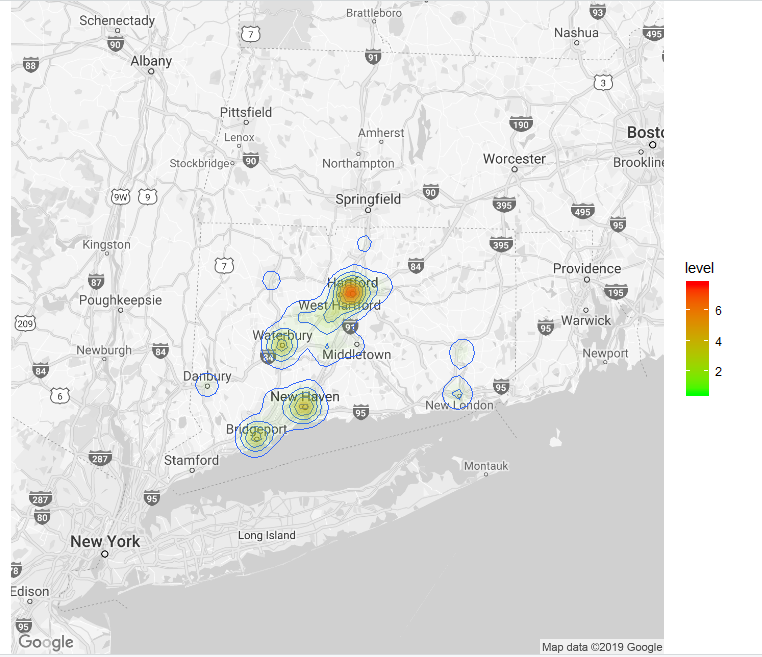
In order to find where the accidental drug-related deaths happened mostly during 2012-2018, I decide to first see the **scatterplot colored by smooth density** to help us find the possible coordinates roughly.



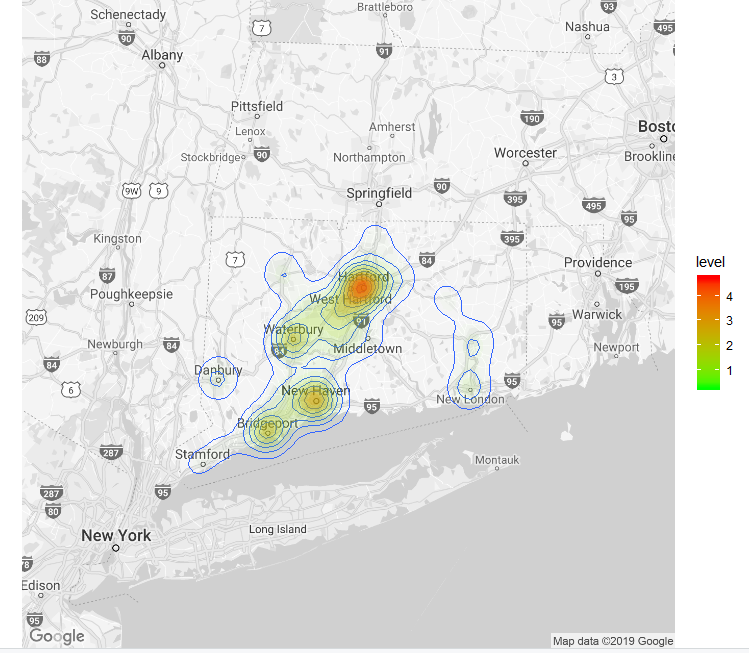
Since from the above scatterplot, we can see the accidental deaths related to drugs in 2012-2018 concentrate on four points within latitude 41.2~41.9, longitude -73.3~--72.7, I suspect that these four spots are big cities or really close to big cities, since the population density is high in these areas. Then I can use geographic heatmap on google map to test assumptions.

**#Geographic Heatmap on GoogleMaps**

Male:



Female:



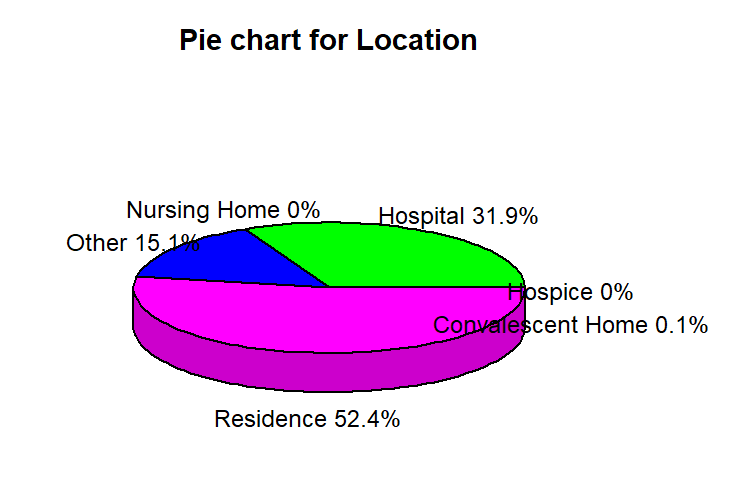
Also, in order to find if sex has high influence on accidental e divided our data by sex and use geographic heatmap on google map to test our assumption.

From above graphs, both parts of accidental drug-related deaths numbers concentrate to four big cities: Hartford, Waterbury, new Haven and Bridgeport, which could make sense since the population density is high in these four areas.

However, it is known that the population of Waterbury, 109,250, is lower than that of Stamford, 128,851, but by r, we get that the number of accidental drug-related deaths of Waterbury, 373, is higher than that of Stamford, 81, during 2012- 2018. Therefore, the population could not be the only reason why accidental drug-related deaths happened mostly on these four areas.

Reference: <https://www.molecularecologist.com/2016/03/geographical-heat-maps-in-r/>

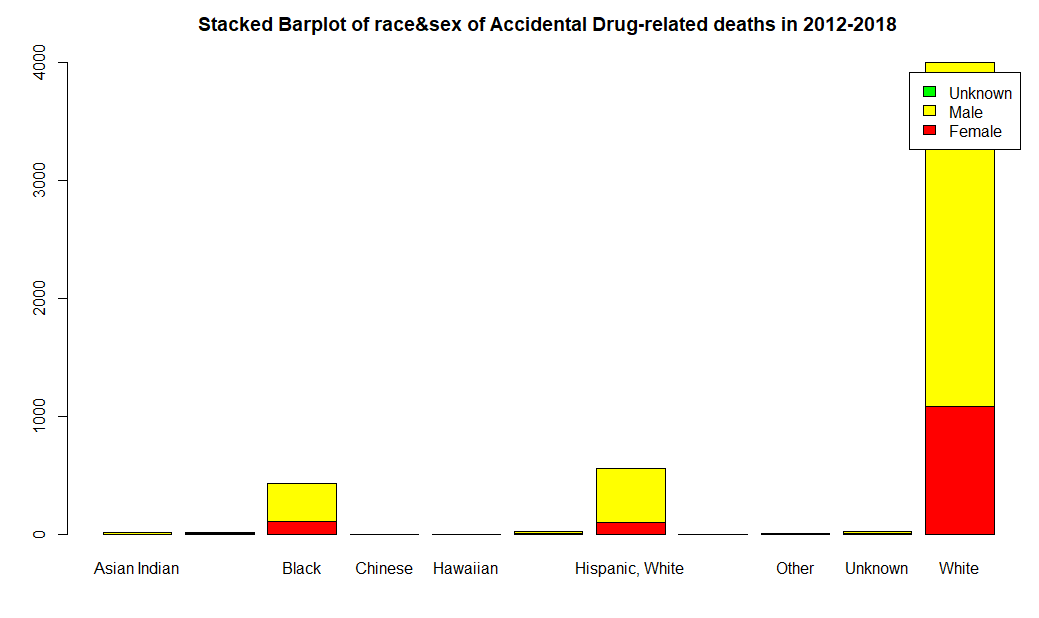
**##3D Pie Chart: Types of Locations of accidental drug-related deaths**

****

**From the above 3D-pie-chart, it is evident that the most accidental drug-related deaths in Connecticut during 2012-2018** **happened in people’s residential area(52.4%).**

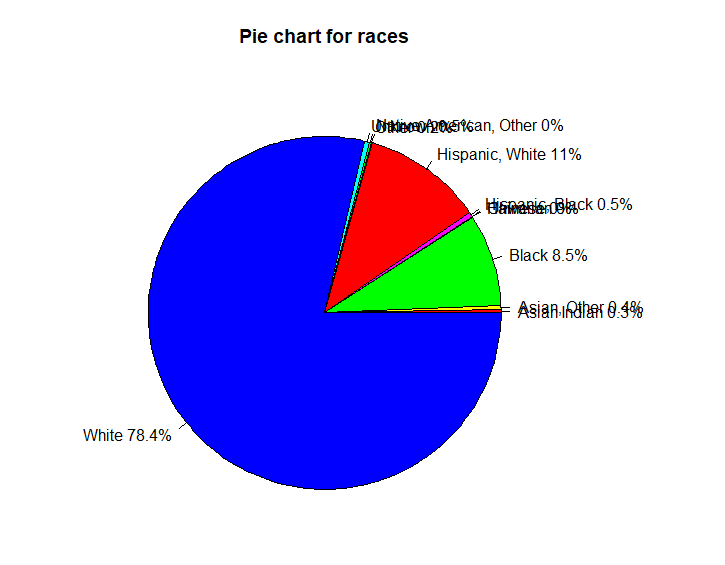
**Section 2: Race and Sex**

**#Stacked Barplot of Race& Sex of Accidental Drug-related deaths in 2012-2018**

****

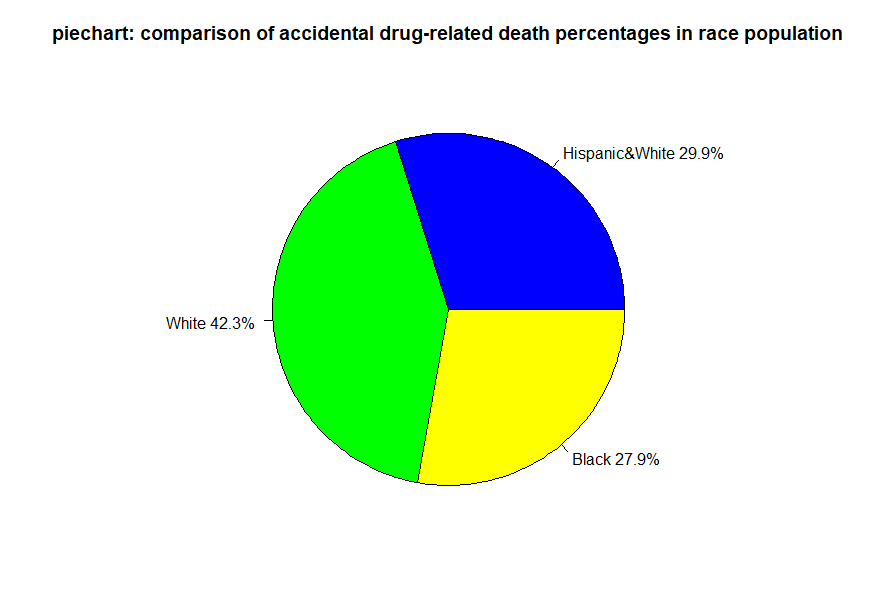
From the above stacked Barplot associated with races and sex, we can see that most accidental drug-related deaths happened to Male and three races: "White", "Hipanic,White" and "Black”.

#Compare the percentage of all races in accidental drug-related deaths:



From the above Pie chart, we can easily see that the percentage of White people in accidental drug-related deaths in 2012-2018 is 78.4% and the difference of percentage of Hispanic, White and White in related deaths is 2.5%, which is not much. Also, the related deaths of other races are way fewer than those of above three races, which is why we can barely see those races in the pie chart.

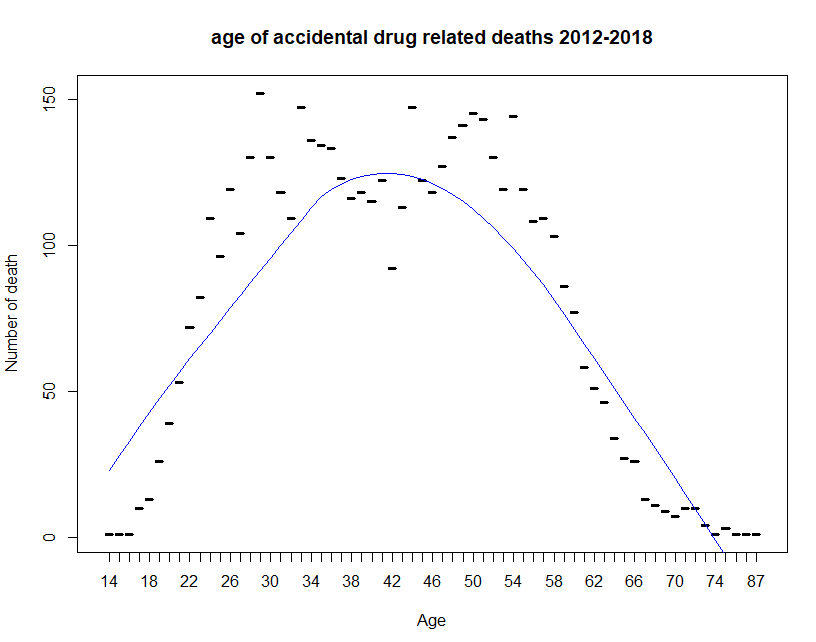
**# Comparison of ratios of accidental drug-related deaths in population of each race in Connecticut by 2018**



From the above pie-chart, we can clearly see that among three ratios of accidental drug-related deaths in population of each race, the ratio of accidental drug-related death to population of white people is way higher than that of Black people and that of White& Hispanic people.

In this way, it is evident that besides the dominating number of white people compared with that of other races, there could exist other reasons why accidental drug-related deaths happen mostly on white people in Connecticut.

**##the number of accidental drug related deaths vs age**

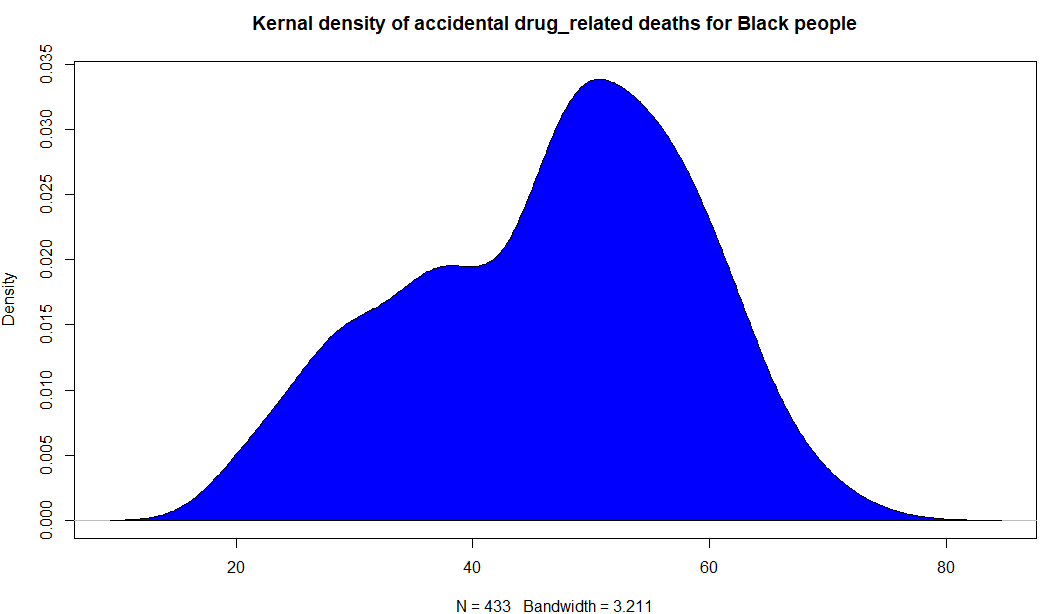


the number of accidental drug related deaths converge to people with ages of about 40 roughly, but we can also see that the highest number of deaths converge to ages of about 28 and 50.

we can see that maybe the reasons why death converge to middle ages are stress levels and economical ability.

**##Density plot for number of deaths vs. Age for black people:**

Since we have the overall plot of the number of accidental drug related deaths vs. age, we can see the density plot for a specific race. Among three races with most data, I can specifically study the **number of deaths vs. Age for black people** and figure out if there exists difference between it and overall plot.

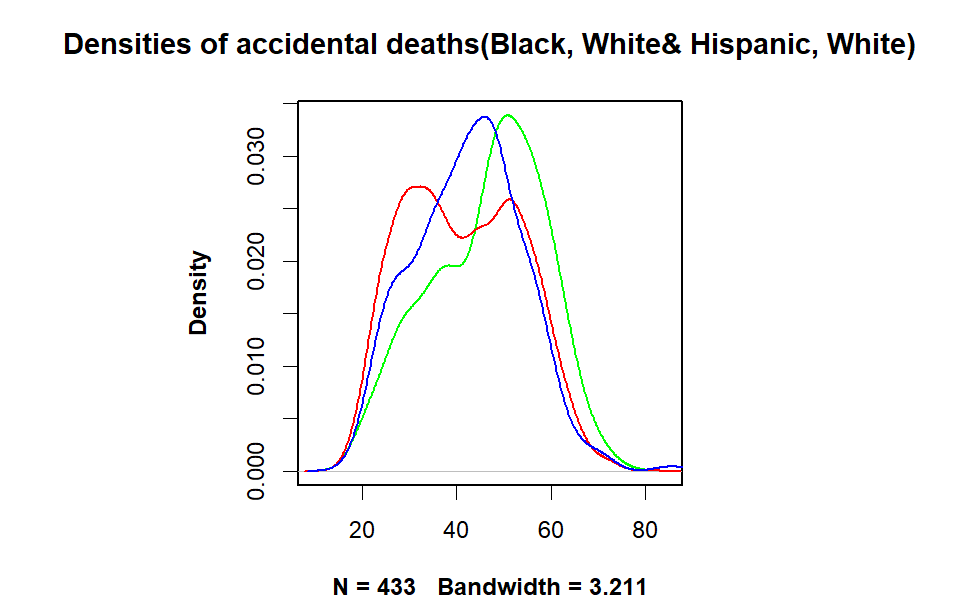


From the above density plot, the conclusion is that the number of accidental drug related deaths converges to black people with ages of about 50.

we can guess that the reason why death converge to middle ages is health condition.

Since from this plot, we can see for different races, the trends of number of deaths vs age could be different. Then, I decided to compare **densities of accidental deaths for different races.**

**##Compare densities of accidental deaths(Black, White& Hispanic, White)**



In the above density plot, we use green, red and blue to indicate number of accidental drug-related deaths of Black, White and White& Hispanic people respectively.

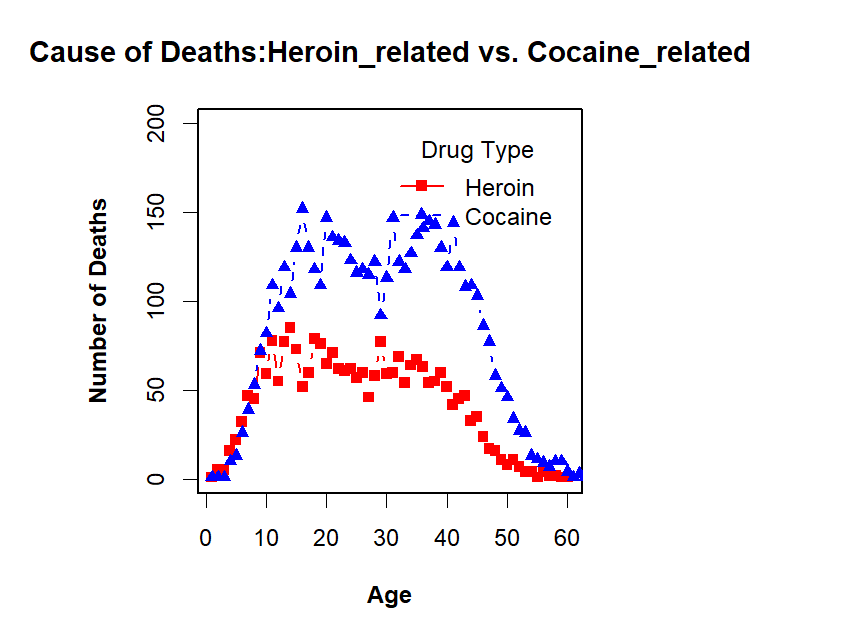
We can see that for Black people and White& Hispanic people, the number of accidental drug-related deaths converge between age 40-50. For White people, the number of deaths has two peaks on about age 30 and 50.

The reason why the overall trend of number of accidental drug-related deaths vs. age is similar to that of white people is probably that the number of deaths mostly happened to White people in 2012-2018, which is 78.4% (obtained in pie chart).

**Section 3: Drug Type associated with Age**

**##Cause of Death: Heroin vs Cocaine**

For two well-known drugs, Heroin and Cocaine, we are interested in their trends versus age which one caused more accidental deaths during 2012-2018 in Connecticut.



First, by the column COD (Cause of Death), we can extract all data with COD related to Heroin and Cocaine respectively. Then we can use legends code to compare two trends of number of deaths vs. age in one graph. As in this graph, we can tell that for people who is younger than 10 years old, the number of deaths because of heroin and cocaine are close. However, for people with ages of more than 10, the number of deaths related to cocaine is way higher than that related to Heroin. We guess the reason is that cocaine has the higher accessibility than Heroin.

**4. Reference:**

1.Lecture Notes: Plots.pdf

2.Geographic Heat Map: <https://www.molecularecologist.com/2016/03/geographical-heat-maps-in-r/>

3.Google Maps

4. Population and percentage of Connecticut shown on “U.S. Census Bureau QuickFacts: Connecticut”

[https://www.census.gov/quickfacts/fact/table/ct#](https://www.census.gov/quickfacts/fact/table/ct)

5. population estimates of all cities in Connecticut by 2017:

<https://www.connecticut-demographics.com/cities_by_population>