Arson Data Report

Kirstie Turnbull

May 4, 2017

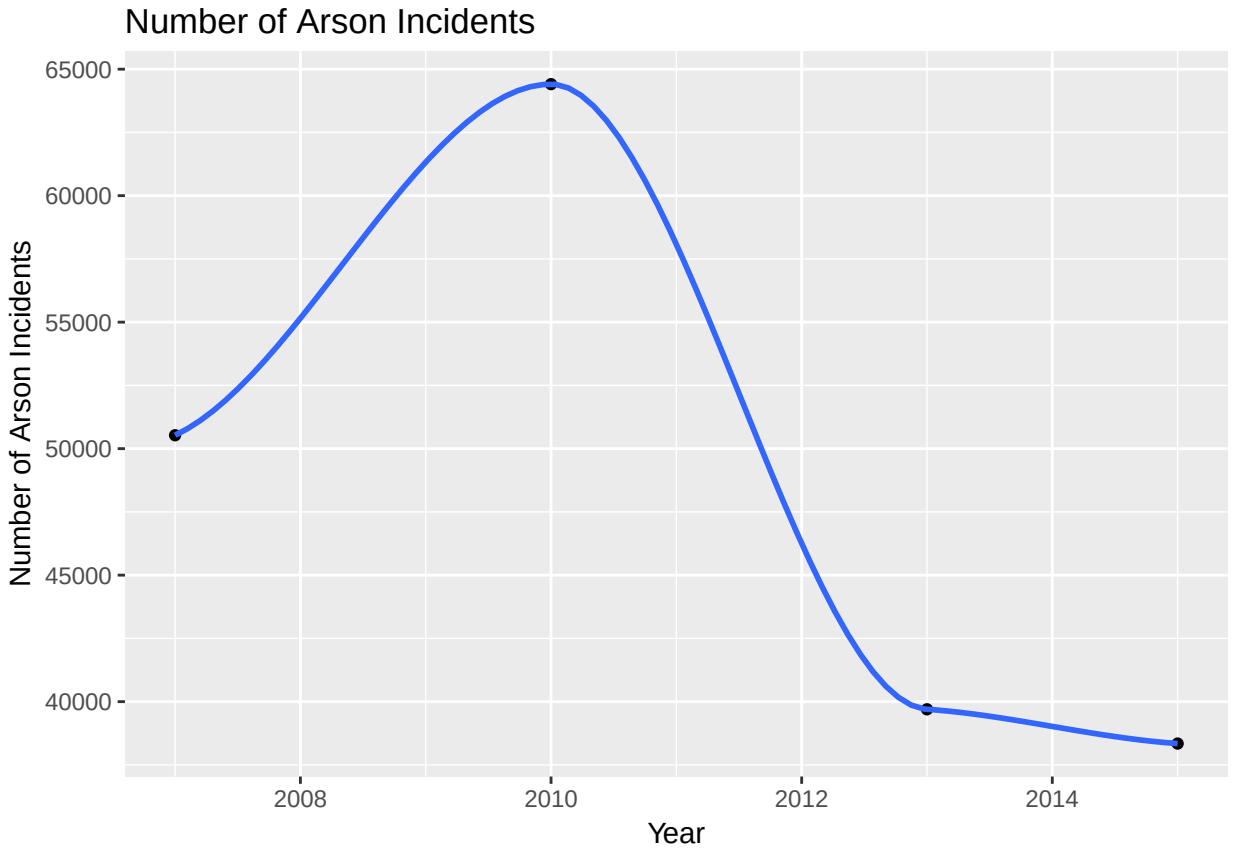
MA 415 Final Project: Arson Data

For my final project, I have chosen to focus on the arson data within the National Fire Incident Reporting System data that we were given by Diane. Diane mentioned that she had noticed a change in the data from the state of Massachusetts in the year 2011, and she asked that we look at the national data to see if we could find anything interesting. I have chosen to look at the national arson data from the years 2007, 2010, 2013, and 2015 (two years before 2011, two years after 2011).

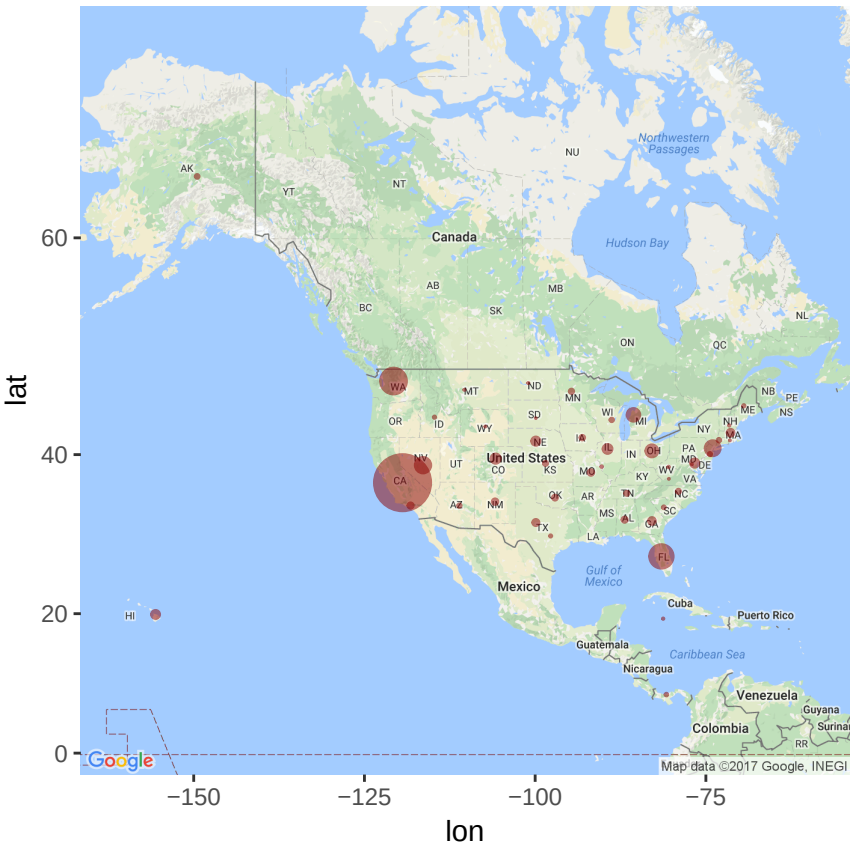
Figure 1: Arson Investigation

```
##    Year Number of Arson Incidents
## 1: 2007                     50530
## 2: 2010                     64408
## 3: 2013                     39697
## 4: 2015                     38343
```

```
## `geom_smooth()` using method = 'loess'
```



I am also interested in what it would look like to map this data set and figure out what states had the largest number of incidents of arson, so I am going to map the data from 2007.



Based on this map, in 2007, California, Florida, and Washington had the greatest number of incidents of arson.

After an intitial exploration of the data, I narrowed my investigation to just a few specific variables that I found particularly interesting: civilian deaths, race and age of person arrested for committing arson, incendiary devices, and time of year of incidents. First I am going to look at civilian deaths.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##     1.0     1.0     2.0     3.1     4.0    36.0   38654
```

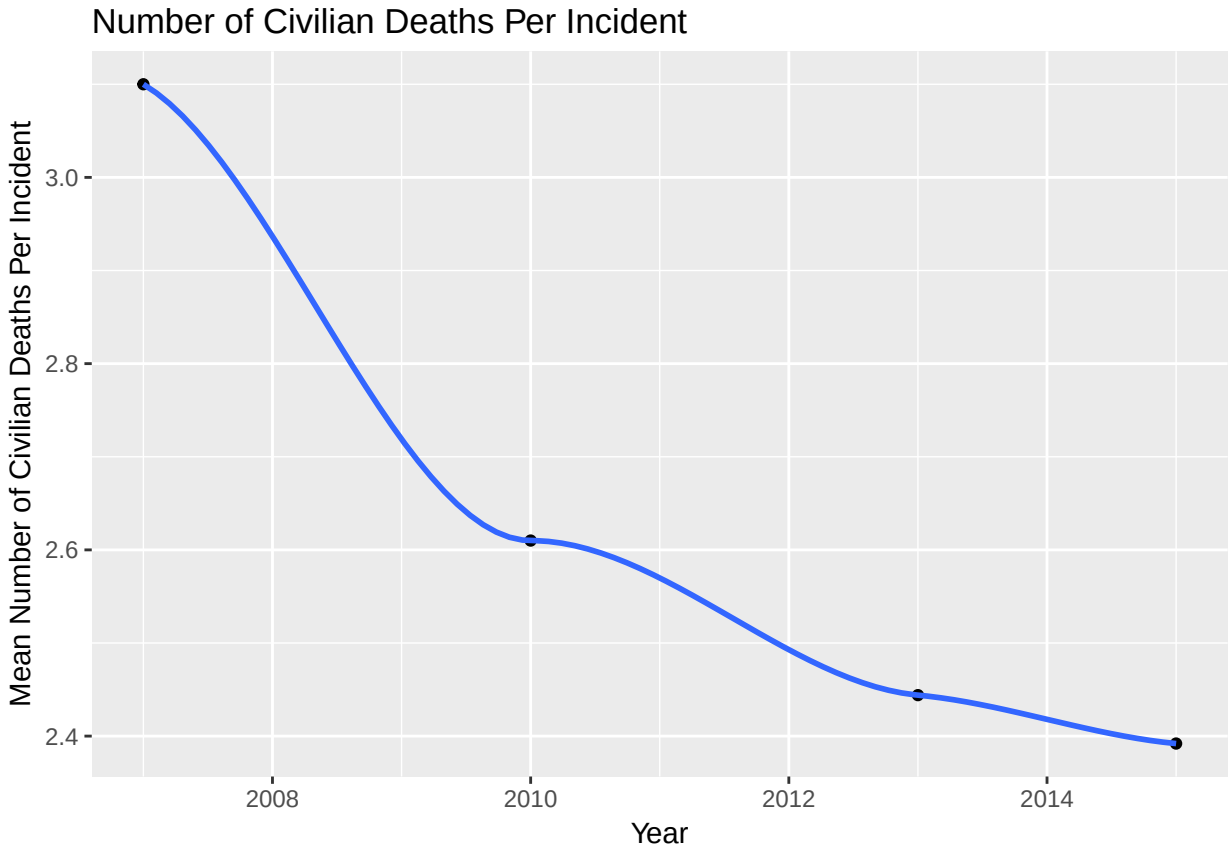
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##    1.00    1.00    2.00    2.61    3.00   26.00   50426
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   1.000   2.000   2.444   3.000  24.000   30740
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   1.000   1.000   2.000   2.392   3.000 153.000   28616
```

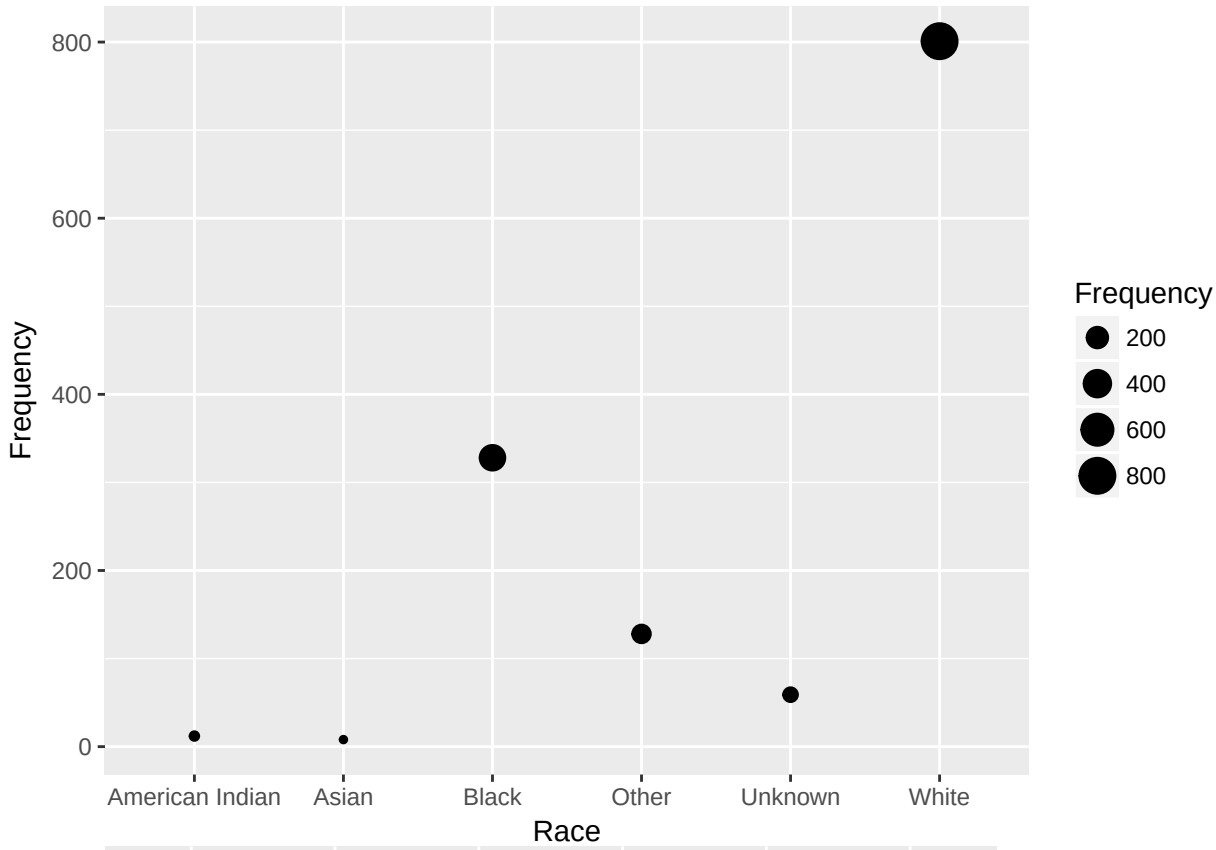
```
##    Year Mean Number of Civilian Deaths Per Incident
## 1: 2007                                       3.100
## 2: 2010                                       2.610
## 3: 2013                                       2.444
## 4: 2015                                       2.392
```

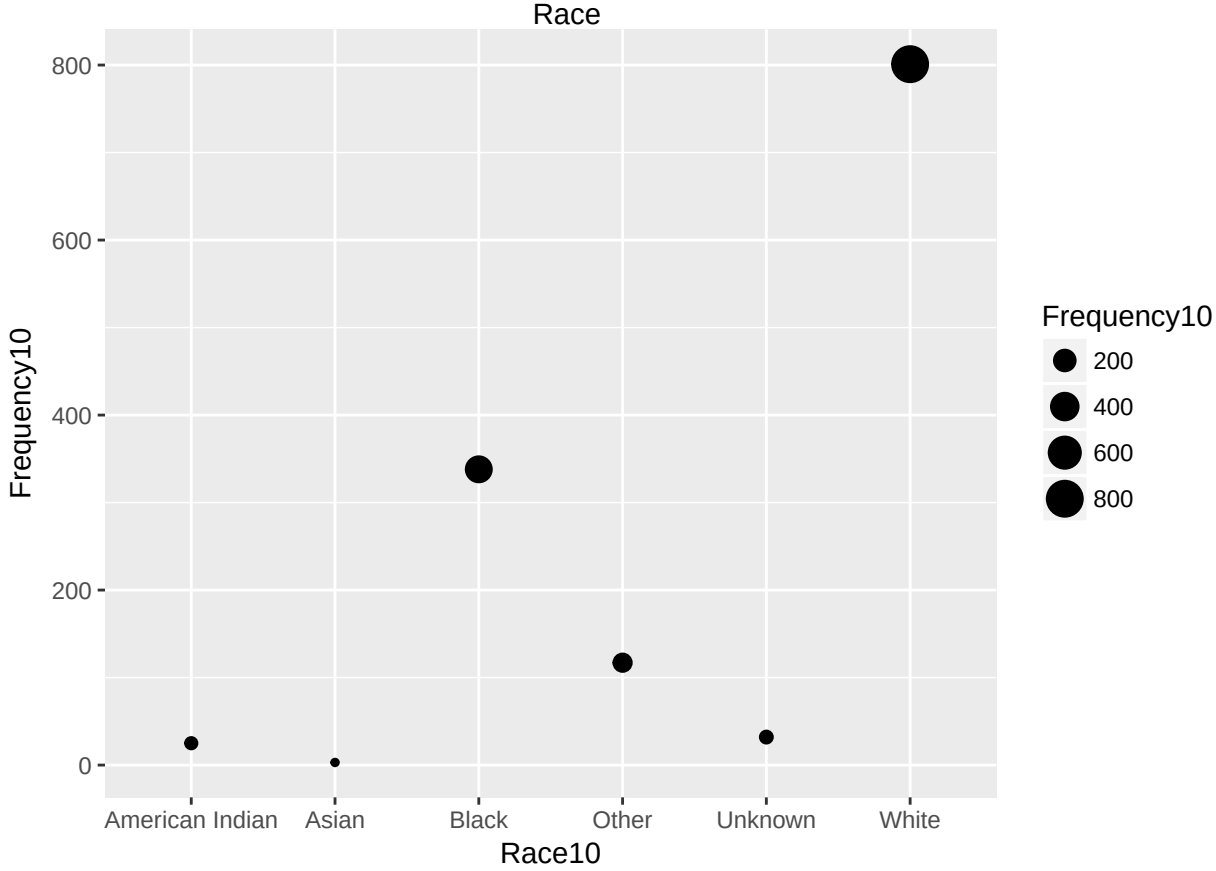
```
## `geom_smooth()` using method = 'loess'
```

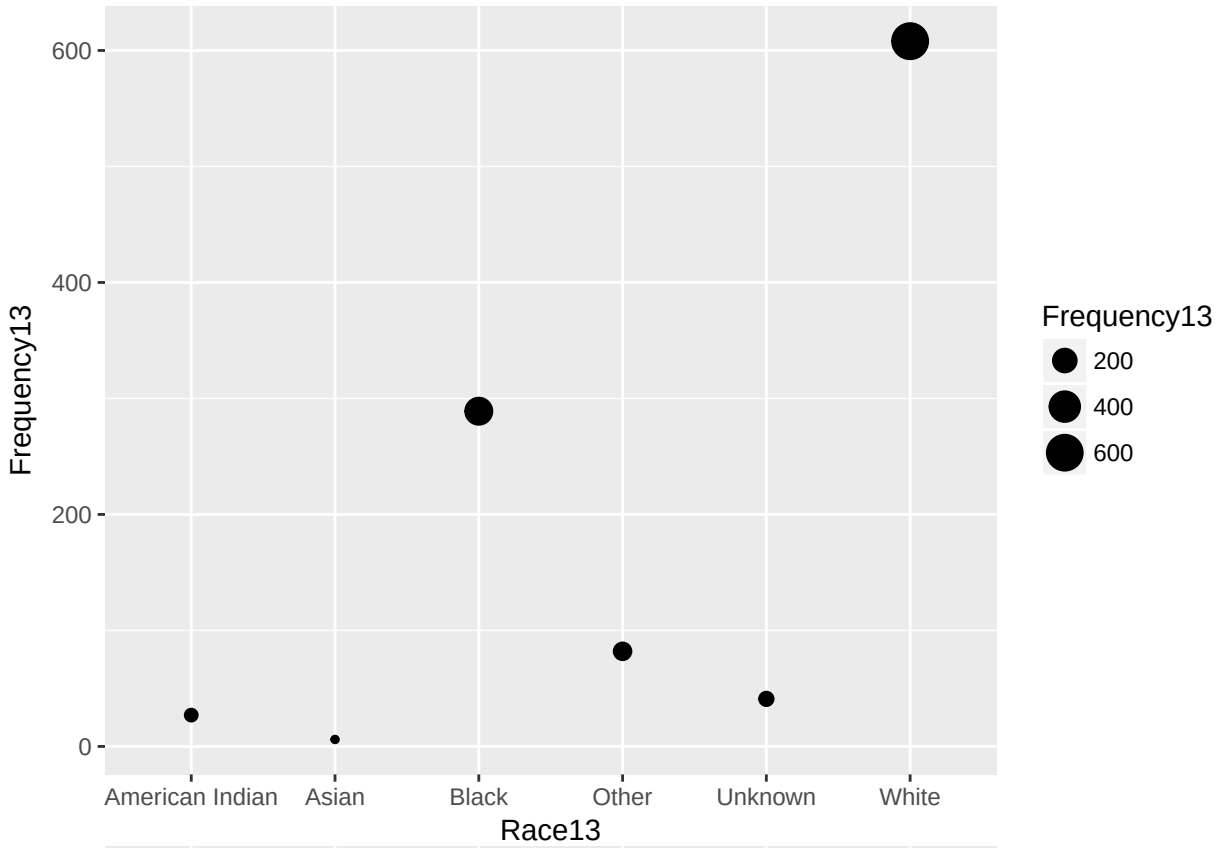


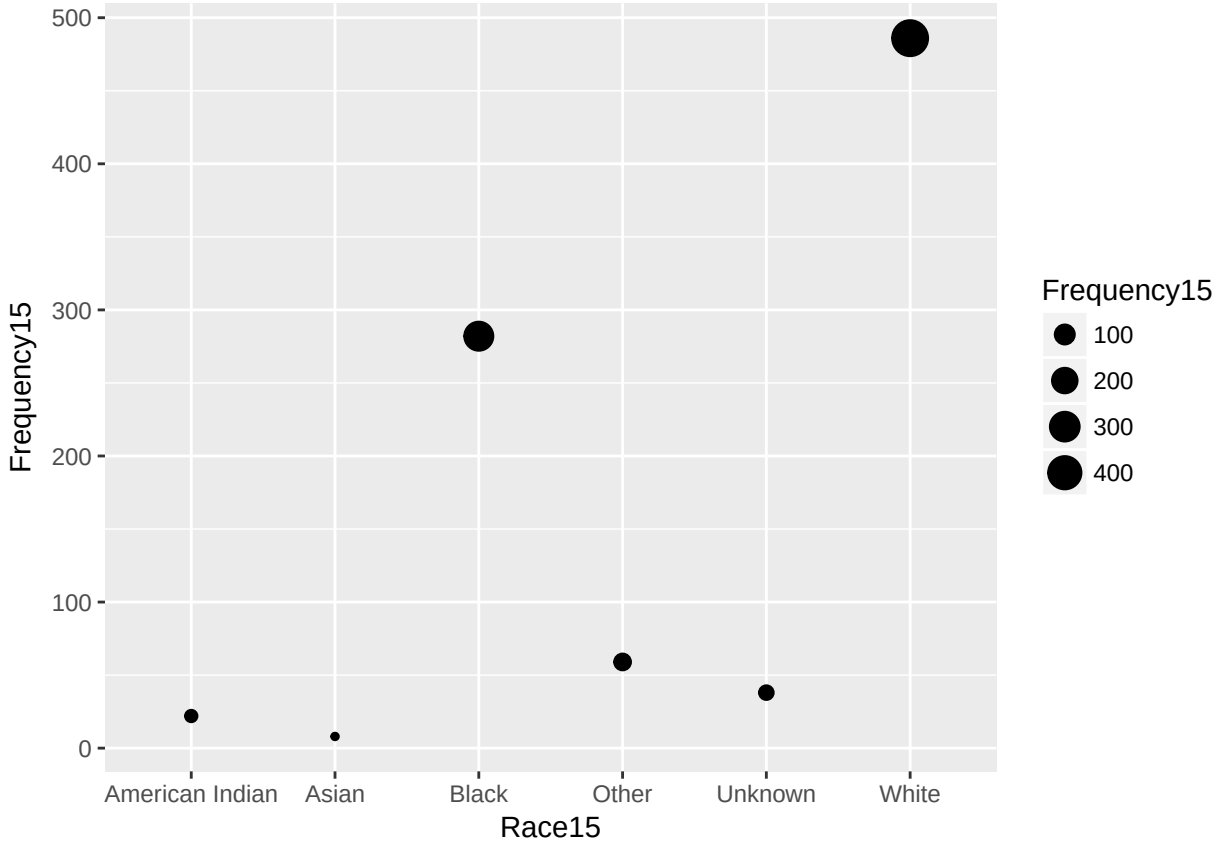

```
##
##  Welch Two Sample t-test
##
## data:  data_2007$`Civilian Casualties` and data_2015$`Civilian Casualties`
## t = 15.76, df = 21554, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.6182929 0.7939350
## sample estimates:
## mean of x mean of y
##  3.098013  2.391899
```

Clearly, as seen in the "Number of Civilian Deaths" graph and as proven in the two-sample hypothesis test, the number of civilian deaths due to arson incidents has significantly decreased from 2007 to 2015. Now I want to look more into the race and age of the person who committed arson. I am going to look at race in this PDF document, and age in the Shiny app.


800
600
400
200
0
Frequency
American Indian
Asian
Black
Other
Unknown
White
Race
Frequency
200
400
600
800


800
600
400
200
0
Frequency10
American Indian
Asian
Black
Other
Unknown
White
Race10
Frequency10
200
400
600
800

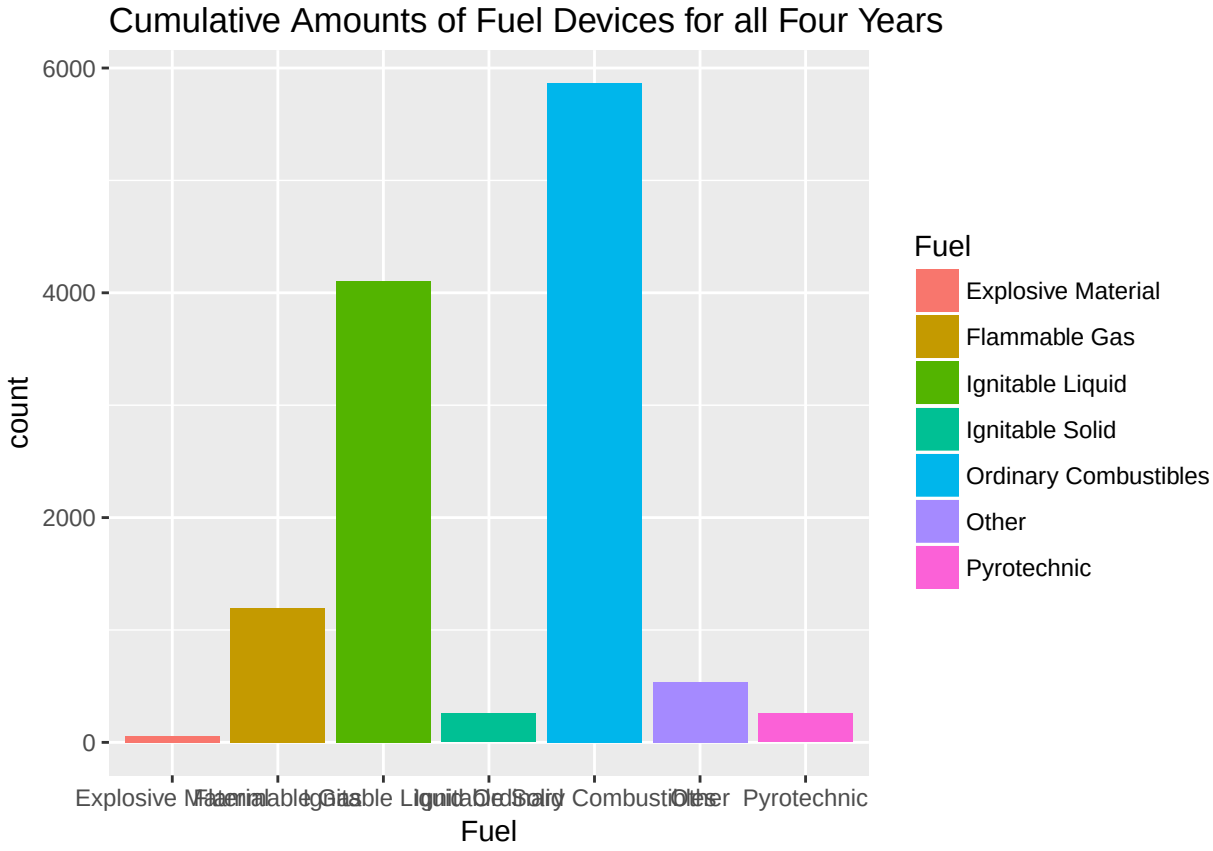




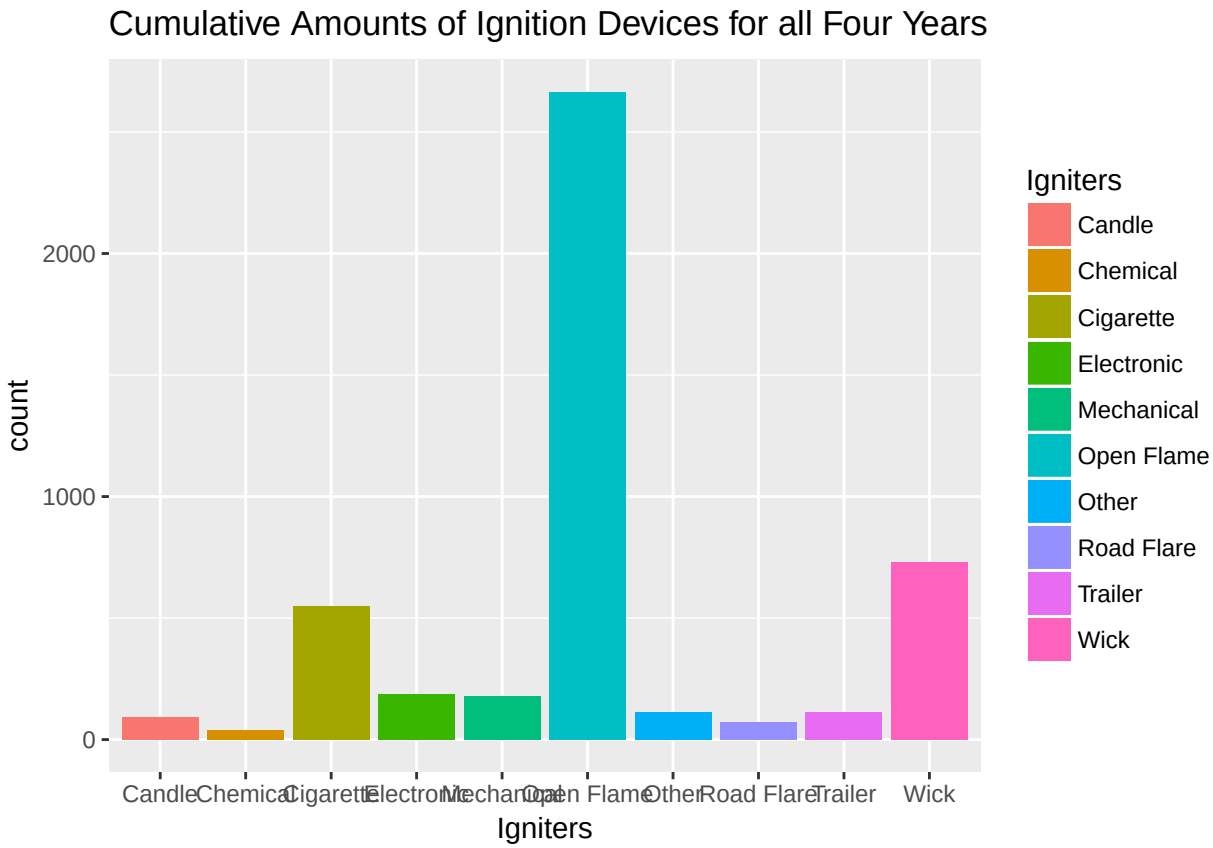
From all of this information, it is clear that white people were much more likely to commit arson. In all four

years that I investigated, many more white people committed arson than any other demographic, so this is likely a trend that continues today. Next I want to look at all of the incendiary devices that were used over the course of all four years.

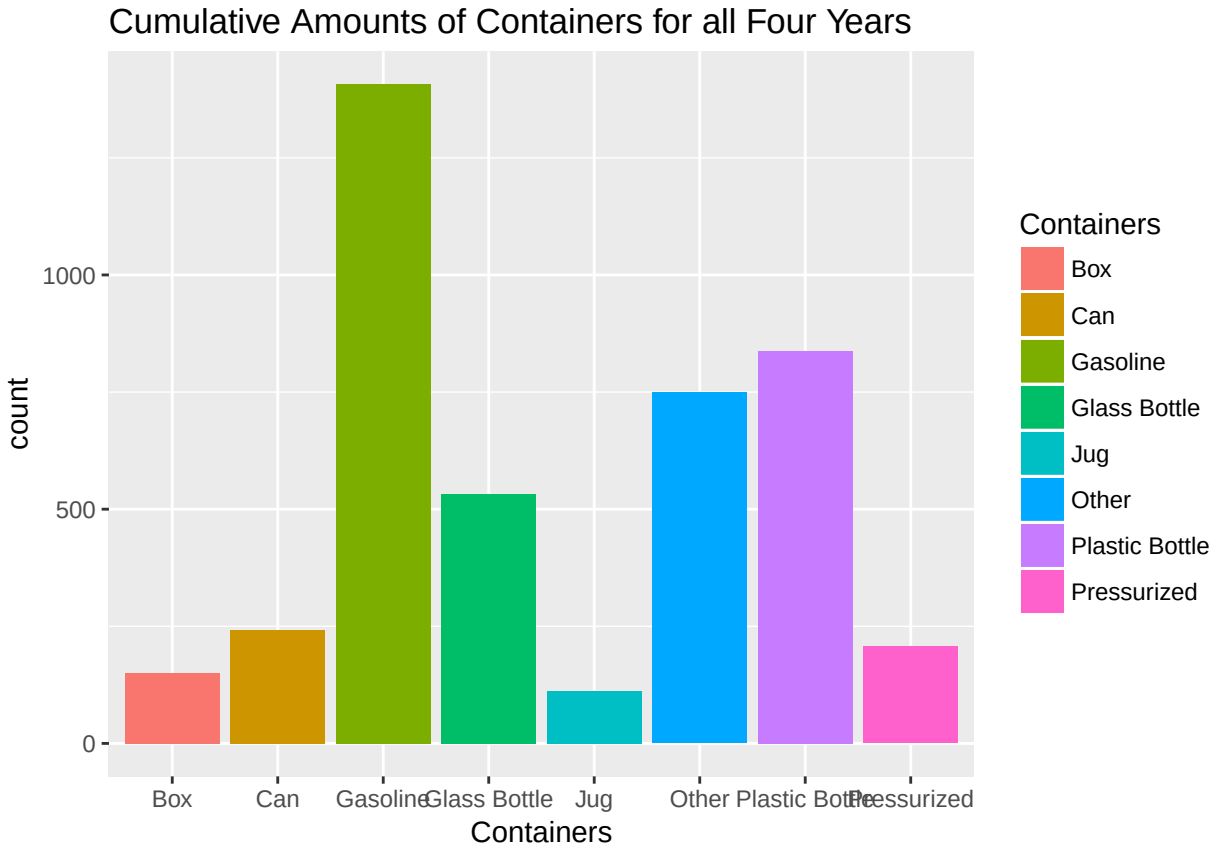
```
##                    Fuel NumberOfIncidents
## 1: Ordinary Combustibles              5866
## 2:      Ignitable Liquid              4108
## 3:         Flammable Gas              1196
## 4:                 Other               540
## 5:           Pyrotechnic               256
## 6:       Ignitable Solid               257
## 7:    Explosive Material                54
```




```
##       Igniters NmbrofIncidents
##  1: Open Flame            2664
##  2:       Wick             732
##  3:  Cigarette             551
##  4:      Other             115
##  5:    Trailer             115
##  6:     Candle              92
##  7: Road Flare              71
##  8: Mechanical             180
##  9: Electronic             186
## 10:   Chemical              39
```

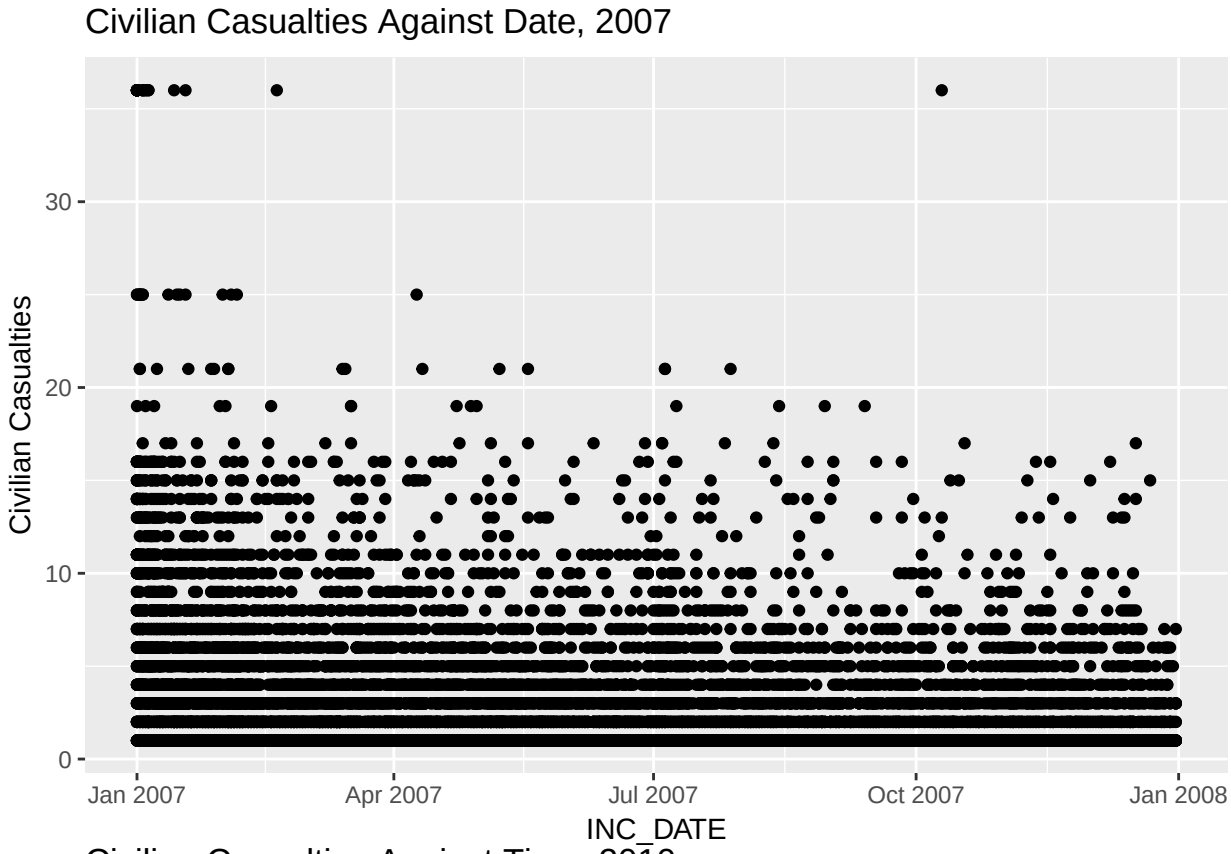


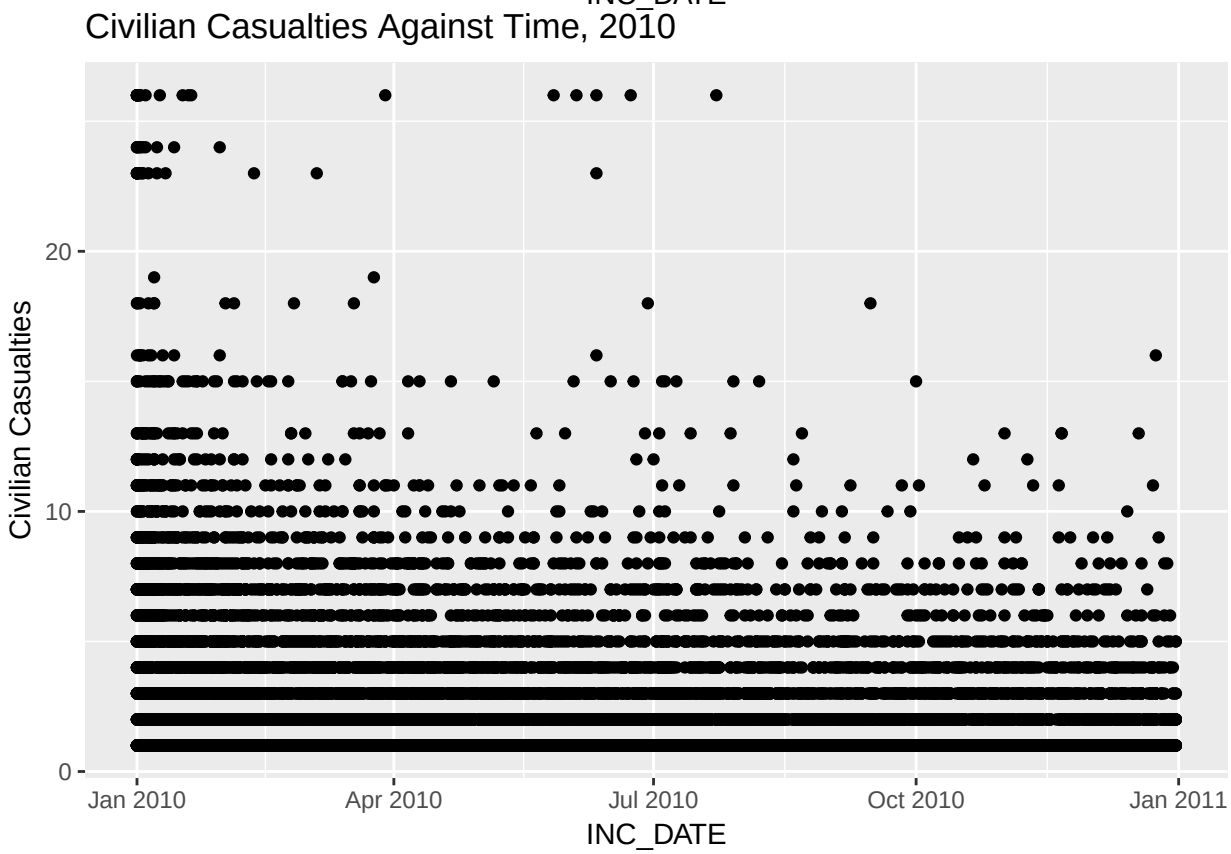

```
##        Containers NumIncidents
## 1:       Gasoline         1407
## 2: Plastic Bottle          838
## 3:          Other          749
## 4:   Glass Bottle          533
## 5:            Can          242
## 6:    Pressurized          207
## 7:            Box          151
## 8:            Jug          112
```

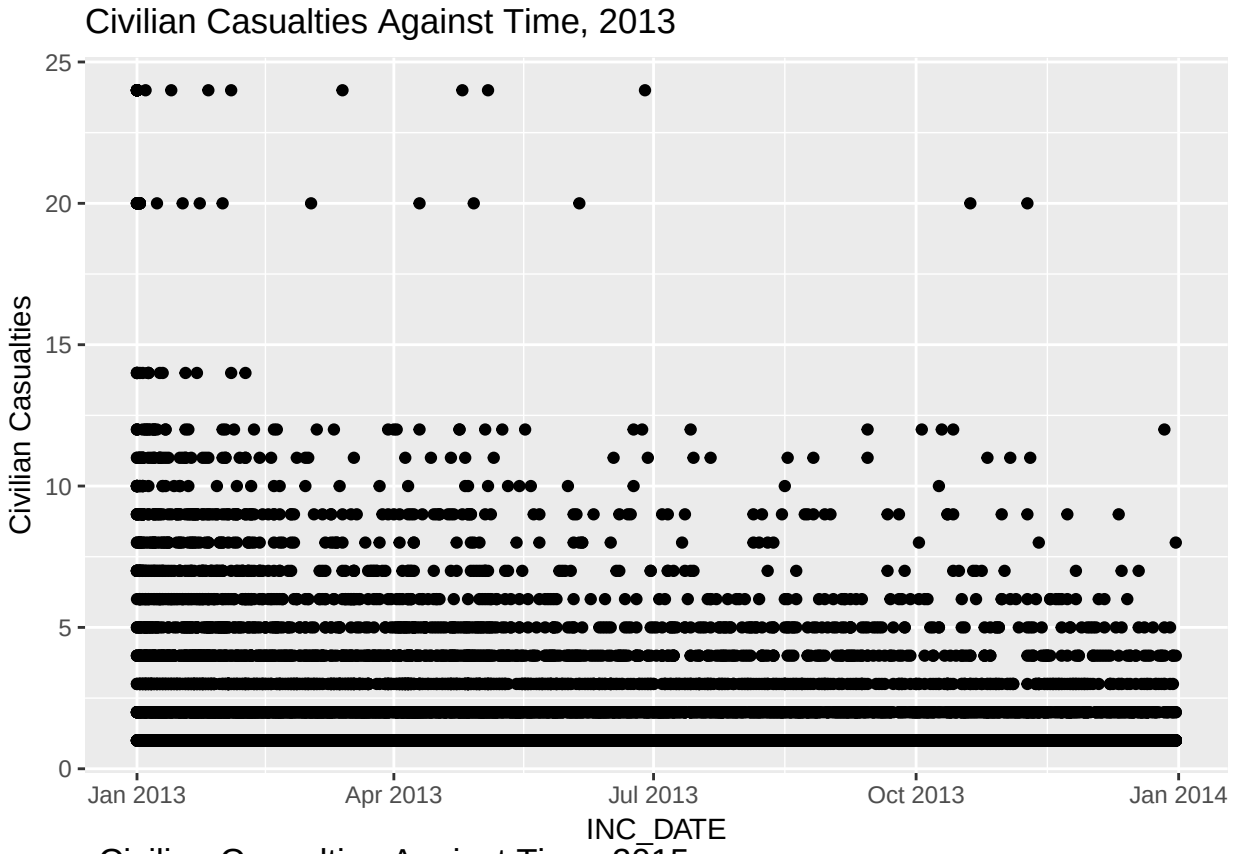


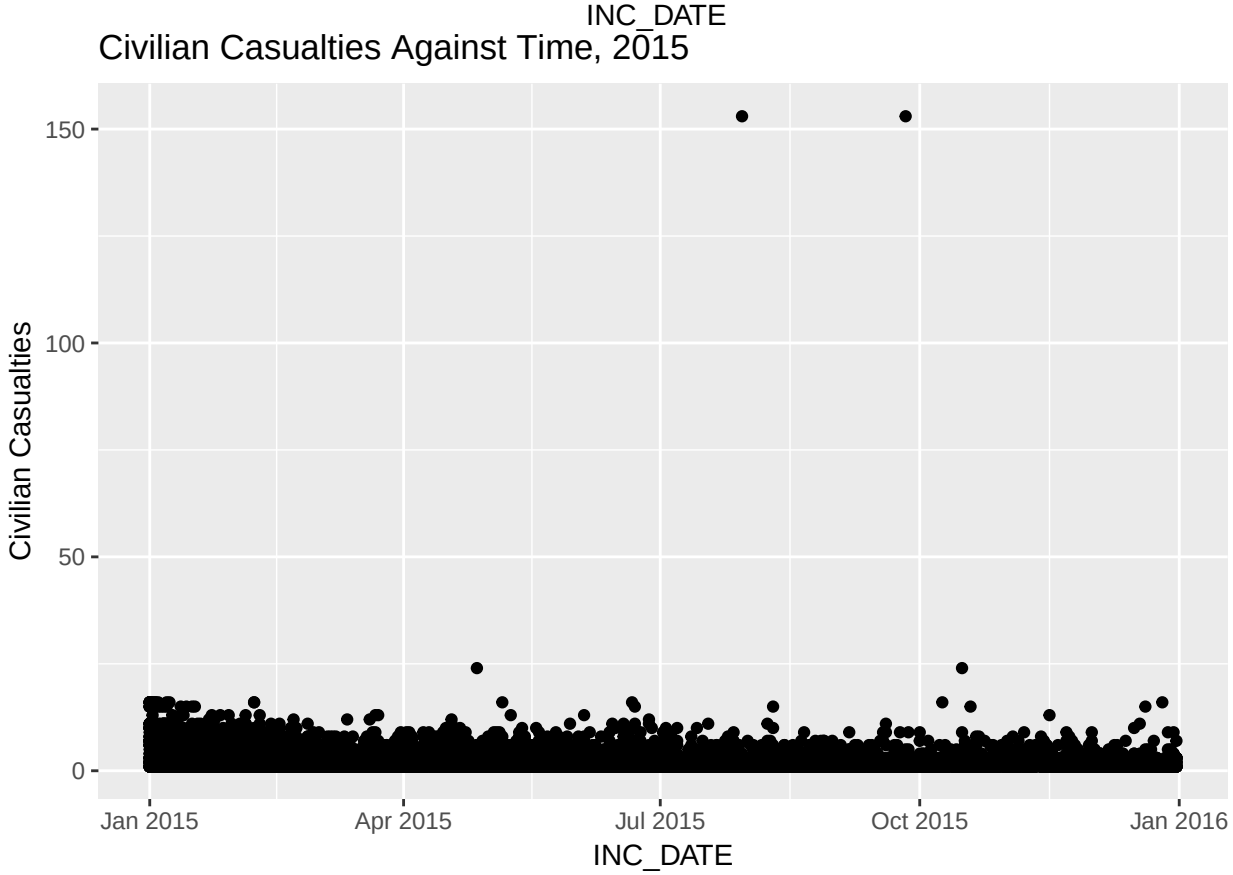
Clearly, the most popular forms of fuel are ordinary combustibles (which likely includes fuel that is meant to start controlled fires) and ignitable liquid. The most popular ignition devices are open flames, wicks and cigarettes and the most common containers that arsonists use are gasoline containers and plastic bottles. I had hoped that by further investigating this variable, I would be able to give suggestions for what products store owners could watch out for people buying in large quantities, because it could indicate that they were going to commit arson. Unfortunately, these are all pretty normal products, so they would be very difficult to regulate.

Now I am going to look into whether or not there are certain peak times of year that more civilians die due to incidents of arson.


Civilian Casualties Against Date, 2007
Civilian Casualties
30
20
10
0
Jan 2007
Apr 2007
Jul 2007
Oct 2007
Jan 2008
INC_DATE


Civilian Casualties Against Time, 2010
Civilian Casualties
20
10
0
Jan 2010
Apr 2010
Jul 2010
Oct 2010
Jan 2011
INC_DATE





While the number of civilian casualties due to incidents of arson do seem to remain mostly constant throughout

the year in all four years, there appears to be a slight decreasing trend throughout the year- there are generally more civilian casualties in January/February than in November/December. What I find really interesting about these plots is that in the 2015 plot, there are two outstanding observations in which over 150 civilians died. I want to look into this.

After further attemptng to research these two incidents of arson, I could not find any news reporting of these on the internet. I find this very strange, because I could find other reports of incidents of arson that resulted in much fewer deaths at other times of the year. This leads me to believe that the number of civilian casualties has been misrecorded here, which is kind of concerning.