NYDP Shooting Incident Data

2024-03-04

Dataset Description

This is a breakdown of every shooting incident that occurred in NYC going back to 2006 through the end of the previous calendar year.

This data is manually extracted every quarter and reviewed by the Office of Management Analysis and Planning before being posted on the NYPD website. Each record represents a shooting incident in NYC and includes information about the event, the location and time of occurrence. In addition, information related to suspect and victim demographics is also included. This data can be used by the public to explore the nature of shooting/criminal activity. Please refer to the attached data footnotes for additional information about this dataset.

Source: https://catalog.data.gov/dataset/nypd-shooting-incident-data-historic

Step 1: Install And/Or Import Libraries

(Optional): In case you didn't install any of these packages below, please feel free to install it as it required, using these command below.

```
# install.packages("tidyverse")
# install.packages("lubridate")
# install.packages("ggplot2")
library(tidyverse)
library(lubridate)
library(ggplot2)
```

Step 2: Load datatable

```
dataUrl <- "https://data.cityofnewyork.us/api/views/833y-fsy8/rows.csv?accessType=DOWNLOAD"

shootingData = read_csv(dataUrl)
```

```
## Rows: 27312 Columns: 21
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (12): OCCUR_DATE, BORO, LOC_OF_OCCUR_DESC, LOC_CLASSFCTN_DESC, LOCATION...
## dbl   (7): INCIDENT_KEY, PRECINCT, JURISDICTION_CODE, X_COORD_CD, Y_COORD_CD...
## lgl   (1): STATISTICAL_MURDER_FLAG
## time  (1): OCCUR_TIME
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
glimpse(shootingData)
```

```
## Rows: 27,312
## Columns: 21
## $ INCIDENT_KEY            <dbl> 228798151, 137471050, 147998800, 146837977, 58~
## $ OCCUR_DATE              <chr> "05/27/2021", "06/27/2014", "11/21/2015", "10/~
## $ OCCUR_TIME              <time> 21:30:00, 17:40:00, 03:56:00, 18:30:00, 22:58~
## $ BORO                    <chr> "QUEENS", "BRONX", "QUEENS", "BRONX", "BRONX",~
## $ LOC_OF_OCCUR_DESC       <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ PRECINCT                <dbl> 105, 40, 108, 44, 47, 81, 114, 81, 105, 101, 2~
## $ JURISDICTION_CODE       <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 2, 2~
## $ LOC_CLASSFCTN_DESC      <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ LOCATION_DESC           <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, "MULTI DWE~
## $ STATISTICAL_MURDER_FLAG <lgl> FALSE, FALSE, TRUE, FALSE, TRUE, TRUE, FALSE, ~
## $ PERP_AGE_GROUP          <chr> NA, NA, NA, NA, "25-44", NA, NA, NA, NA, "25-4~
## $ PERP_SEX                <chr> NA, NA, NA, NA, "M", NA, NA, NA, NA, "M", NA, ~
## $ PERP_RACE               <chr> NA, NA, NA, NA, "BLACK", NA, NA, NA, NA, "BLAC~
## $ VIC_AGE_GROUP           <chr> "18-24", "18-24", "25-44", "<18", "45-64", "25~
## $ VIC_SEX                 <chr> "M", "M", "M", "M", "M", "M", "M", "M", "M", "~
## $ VIC_RACE                <chr> "BLACK", "BLACK", "WHITE", "WHITE HISPANIC", "~
## $ X_COORD_CD              <dbl> 1058925.0, 1005028.0, 1007667.9, 1006537.4, 10~
## $ Y_COORD_CD              <dbl> 180924.0, 234516.0, 209836.5, 244511.1, 262189~
## $ Latitude                <dbl> 40.66296, 40.81035, 40.74261, 40.83778, 40.886~
## $ Longitude               <dbl> -73.73084, -73.92494, -73.91549, -73.91946, -7~
## $ Lon_Lat                 <chr> "POINT (-73.73083868899994 40.662964620000025)~
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

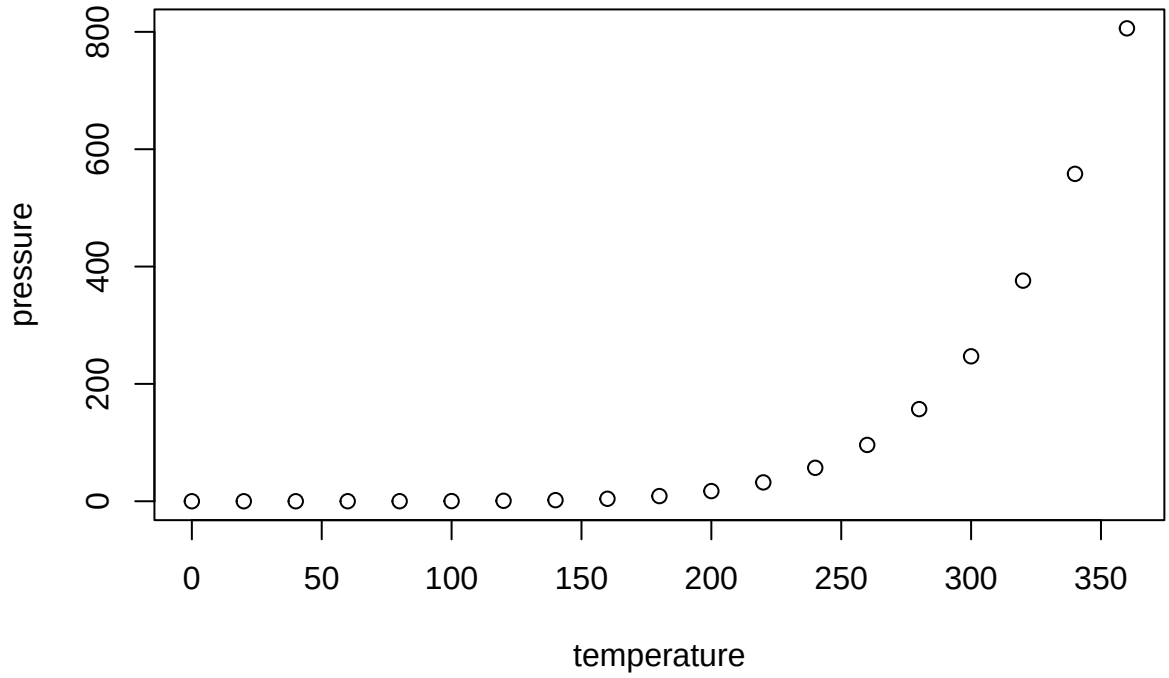
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.