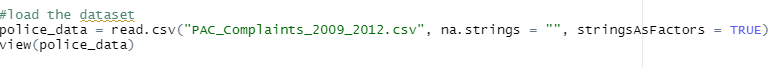
**Visualization with Police Complaints Data**

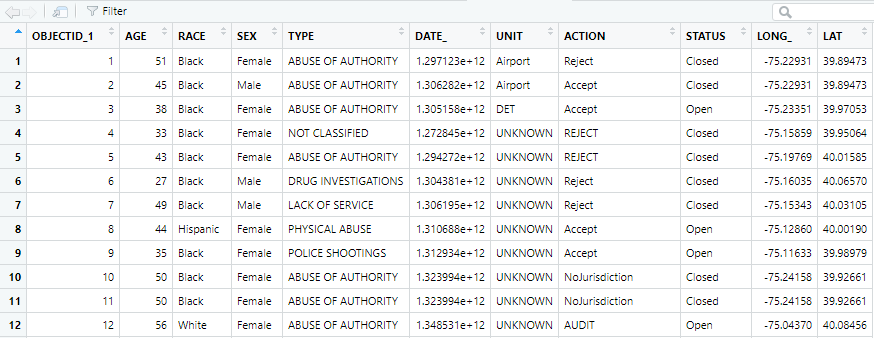
The dataset has been taken from [Police\_data\_website](https://www.opendataphilly.org/dataset/philadelphia-police-advisory-commission-complaints) shows information about Complaints filed with the Police Advisory Commission against Philadelphia Police officers. The information comes directly from Police Advisory Commission Complaint Database.

DATA PREPARATION:

* Data Import: following code uploads the data into R studio.



Set the working directory to read the file from the location and load the data into R. view the dataset police\_data.

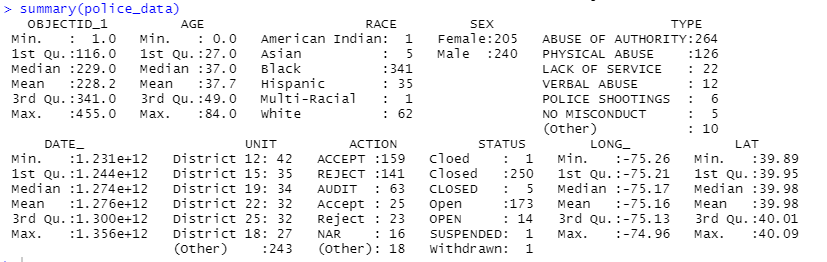


* Data description: The dataset has 445 rows and 11 columns including Complaint ID, Age, Race, Sex of the people, Type of Complaint, Date , Unit, Action on the complaint, Status of the complaint, Longitude, Latitude.

Text, letter

Description automatically generated

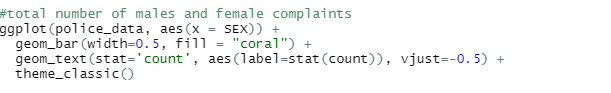
Since the dataset has categorical values, check the summary of the dataset to see the factor levels of Sex, Type, Race, Unit, Action, and status variables.



DATA CLEANING: When categorical variables turned into factors empty string is created as an level for null values. Therefore, we will clean the data by removing the empty level data from the factors.

DATA VISUALIZATION:

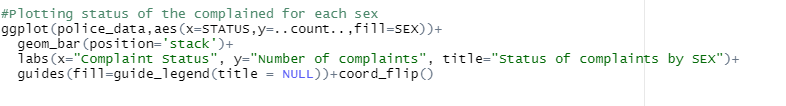
Now the dataset is ready for visualization the graphs using ggplot2 library in R. The following code will display the number of complaints raised by gender.



Chart, bar chart

Description automatically generated

From the graph we can say that there are 205 complaints raised by females and 240 complaints raised by the males.



The above code displays the Status of the Complaints raised by male and females.

Chart

Description automatically generated

From the graph, we can say that most the complaints are closed but nearly 175 complaints are open.

A picture containing text

Description automatically generated

Above code will display the Actions taken on the complaints with their status to find the working flow patterns of the complaints.

Chart, bar chart

Description automatically generated

We can observe that, the rejected cases are closed, and audit cases are either open or closed but hardly withdrawn or suspended. Many of the cases are accepted and opened for audit.



Chart, histogram

Description automatically generated

The histogram describes the age range of the complaints. The most complaints are received by the age group between 20 to 60. Let’s plot the age distribution in details with other variables.

Text

Description automatically generated

Chart

Description automatically generated

Most complaints raised by age group 20 to 60 where most of them are closed or open with a few suspended complaints.

Text

Description automatically generated

Plotting the Action items with Sex division and Race distribution. From the below graph we can say that most of the cases are either accepted, rejected or in audit. Black, White and Hispanic people seem to have more complains slightly males over females.

Calendar

Description automatically generated