**DASC 5300/CSE 5300 Foundations of Computing**

**Instructor: Sharma Chakravarthy**

**Project I: Python Programming and Data Analysis**

**Group 8 (Year: 2013 and 2014)**

**Team Members**

1. Palavalli, Sai Dinesh Reddy
2. Dmello, Lolita Louis

**Grading Scheme**

|  |  |
| --- | --- |
| Histogram of Boroughs | 20 |
| Histogram of age-based histogram | 20 |
| Plot of Location-based arrests for each Borough | 20 |
| Q/A performance during demo | 30 |
| Challenges encountered and solution | 5 |
| Weight | 15% of total |
| Total Points | 100 |

**Index**

|  |  |
| --- | --- |
| **Topic** | **Page no** |
| Analysis of the achieved graphs |  |
| A. Histogram of Boroughs | **3** |
| B. Histogram of age-based histogram | **4** |
| C. plot of Location-based arrests for each Borough | **5** |
| Overall Status | **6** |
| Division of labour | **6** |
| Problem encountered and solution | **7** |
|  |  |

1. **Histogram of Boroughs**

After proper analysis of the graph and with the help of the additional information from the [internet](https://worldpopulationreview.com/boroughs/manhattan-population). Following conclusions may be drawn.

* + - Chart, bar chart

      Description automatically generatedQueens’s population being more than Manhattan’s population. Despite this fact Queen have fewer number of crime cases. Fewer number of crime cases could be due to its low unemployment rate in Queens than Manhattan.
    - From the fig 1.2 we can conclude that most of the crimes (M and V) is done by Blacks. Manhattan being the 3rd largest populated town with 15% of the Black race which might be the reason having high crime rate.
* Staten Island’s population is less than 50% of any one of the boroughs. Even the black’s and Hispanic’s population is less (these two races are responsible for the greatest number of crimes as per fig 1.2 Staten Island’s unemployment rate being the lowest. So, we can conclude that these might be the factors effecting crime rate of any given borough.

Fig 1.1 Histogram of Number of Misdemeanour and Violation cases of each Borough from year 2013-2014

Few things are evident from these figures:

Chart, pie chart

Description automatically generated

* Black race had the highest amount of arrest rate for misdemeanor and Violation. The combined crime rate of black and Hispanic people was about 83%
* The second highest arrest rate was of White Hispanic people.
* Staten Island had the least number of cases for misdemeanor and almost negligible for violation. Another observation which can be made is that there was a decrease in misdemeanor and violation cases from the year 2013 to 2014 except for Queens and Staten Island for misdemeanor and Bronx for a violation where there was an increase in cases.

Fig1.2: chart represents rate of arrest cases of Misdemeanour & Violation from year 2013 to 2014 of each race

1. **Histogram of age-based histogram**

After proper analysis of the graph Following conclusions may be drawn

**Chart, bar chart

Description automatically generated**

Fig 2. Histogram of Number of Misdemeanour and Violation cases of each age group for each Borough from the year 2013-2014

1. The age group 25-44 has the greatest number of arrest cases in all 5 boroughs.
2. The second highest arrest cases were from the age group 18 -24.
3. Age group 65+ had the least number of arrest and were negligible as compared to other age groups.

Overall, it is evident from the graph that adults are more likely to perform crimes as compared to teens and old people in New York City.

**C. Plot of Location-based arrests for each Borough**

**A picture containing text, wire, light, line

Description automatically generated**

Fig3 google map of the number of arrests in a radius of 3 miles from that location for each Borough

After plotting the number of arrests in a radius of 3 miles from the location (Randomly picked) for each Borough on the map, we could see higher cases in Manhattan.

**Overall Status**

* Assigned project was completed on time with proper division of the work between the teammates.
* The dataset consisted of 4798339 number of rows and 18 columns. Year 2013 and 2014 data was filtered out for the analysis as per the year assigned for each group (group 8). Exploratory data analysis was performed on the filtered data set. Graphs and other insights can be found in this report. Additional information was used for the analysis of the graph. So better conclusion could be drawn.

**Division of labour**

To carry out the project and continuously monitor the progress of the work we carried it on google colab ([Link for the code](https://colab.research.google.com/drive/1emnjV4_56hGg45vUfGH8PqS9I3YC90L5?usp=sharing))

|  |  |  |
| --- | --- | --- |
| **Time spent** | **Task** | **Member name (Hours dedicated)** |
| 7 hr | Plotting the graphs | Dinesh Reddy(5), Lolita (3) |
| 6 hr | Analysing the graphs | Dinesh Reddy(3), Lolita(6) |
| 10 hr | Report | Dinesh Reddy(5), Lolita (5) |

**Problem Encountered**

1. While examining we couldn't understand why Staten Island had the least crime rate as compared to other boroughs and after analysing, we understood that due to the low population rate and better income rate it was Lowest.
2. From the analysis perspective, the data attributes provided were not informational enough to understand the reasons behind the trend, hence we had to search for extra information from the internet (eg. The population rate, employment rate)
3. During the analysis of the graphs, couldn’t reach a conclusion for less crime rate in Brooklyn when compared to Manhattan even though
   1. Brooklyn’s population and unemployment was high when compared to Manhattan’s population and unemployment
   2. Brooklyn’s education and income rate lower than that of Manhattan’s education and income

there could be multiple reasons for this trend. One of possibilities could be all the criminals come to Manhattan as there are more valuable products to steal (as more educated and cost of owning a property is high when compared)

\*\*\*\*End\*\*\*\*