Mini Project Report On

Profiling in Stop and Frisk Program of New York Police in 2014

By - Arpit Singh **SBU ID - 110162005**

Advisor-Prof. Leman Akoglu

Department Of Computer Science

Stony Brook University

Introduction:The death of Eric Garner , a 43 old African-american on July 2014 brought the attention towards the perceived racial bias under New York Stop and Frisk Program[1]

"Stop and Frisk is practise by which a police officer initiates a stop of an individual on street allegedly based on some criminal activity"

The Center of Constitutional Right[2]

Since 2013 the practise was restricted[3], however on wake of gruesome incident, 2014 data is being analyzed to find out to what extent the profiling exist in stop and frisk practices of NYPD

Apart from analyzing racial profiling, others factors like age and build are also analyzed to find out if particular age range and body build profile are more frequent victims of stop and frisk practices. Apart from it, distribution over times of days and months of year in terms of number of person stopped and frisked is also analyzed to judge at what times of day and in what parts of year police is most active.

Exploratory data analysis is done and summary facts are presented below:

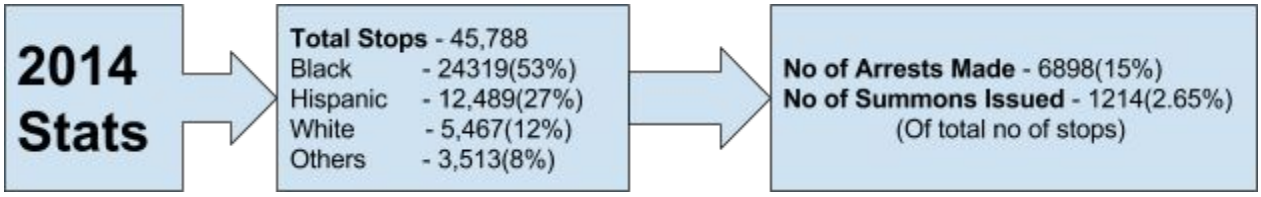

Figure 1 Summary Statistics of Stop and Frisk program by NYPD during 2014

Data Preparation and Cleaning: Data of stop and frisk is obtained from NYPD website(http://www.nyc.gov/html/nypd/html/analysis_and_planning/stop_question_and_frisk_report.shtml)

There are more than 100 fields. However , for purpose of project we needed only race,age,build,timestop,datestop and arrests made[4]. Following steps are taken for data cleaning

1. There are 8 categories of race, however only black,white,hispanic and others is taken into consideration. Black hispanic and white hispanic are combined under hispanic category while asians/pacific islander etc are joined together under other category
2. Age of as many as 201 people is mentioned incorrectly. Age is not mentioned for 54 people, for 66 people, age is mentioned to be less than 5 years and it is indicated more than 100 years for 81 people. All these are combined in unknown category.

Age not mentioned	54
Age mentioned less than 5 years	66
Age mentioned is more than 100 years	81
Total	201

Table 1: Data cleaning for discrepancies in age column

Objectives and Problem Statement: The project aimed at analyzing the program and was initiated by asking basic questions like who is being frisked? When does most frisking take place? and does people frisked share any common characterics?
Following are objectives of project:

1. Is there any biases that affects the selection of stop and frisk candidate?
2. Is particular age and build profile forms the largest share in stop and frisk program?
3. At what time of day police is most active?
4. At what time of month police is most active?
5. How many frisks resulted into arrests.?

Taking above objectives in mind, data is being analyzed.

Results and Discussion:

(1) **Race profile** of stop and frisk program is presented below. It is clear from

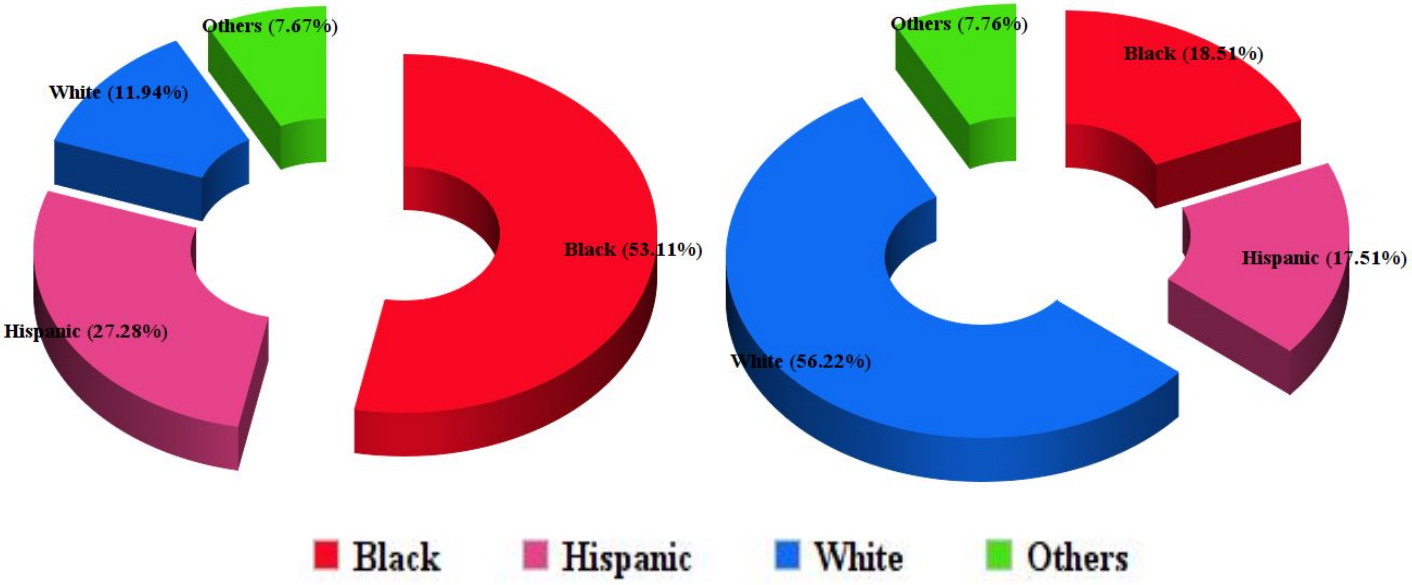

Figure 2(a)Racial Profile of Stop and Frisk 2(b)Demography of New York state

From the pie chart shown above, black community form 53.31 % of all stop and frisk practice despite forming only 16.6% of population. Hispanic community forms 27.28% of this practise while their actual population is just 18.6% of total New York population.Comparative bar graph is drawn to show the extent of racial profiling in stop and frisk activity,It looks very clear that black community is most vulnerable target of stop and frisk practices followed by hispanic population. Frisk rate of white community is low 11.94% despite forming 56.5% of total population. The statistics show the inherent bias in stop and frisk practices of NYPD. Racial profiling is based on assumption that a member of certain community are more likely to perform certain kind of crimes.

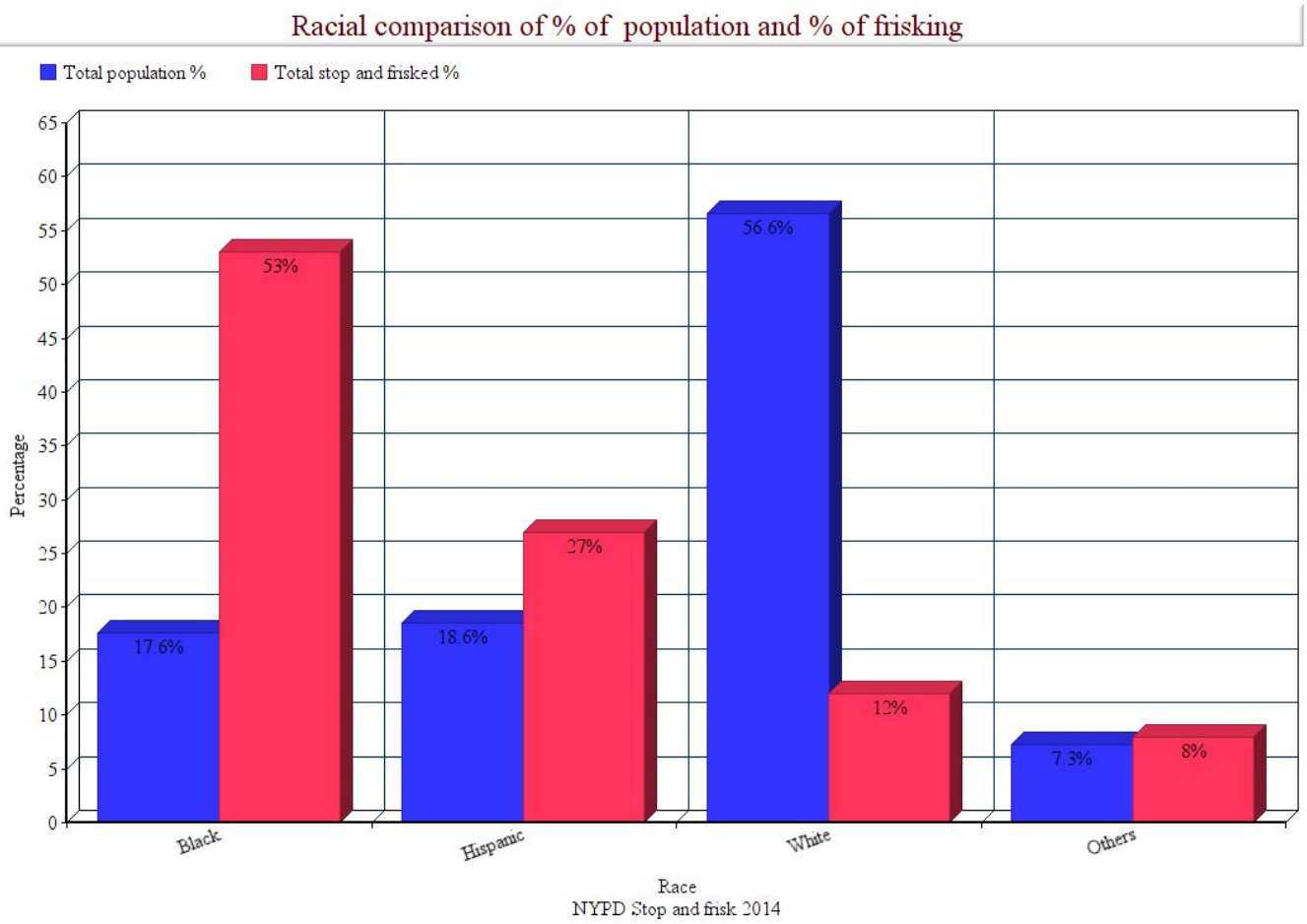

Figure 3 Comparative analysis of % of population in demography and % of people frisked per community

(2) Build Profile : Build profile is also analyzed to determine physique of persons most vulnerable to being stopped. Figure 4 shows build profile. Persons with 'Medium' builds form 53.4% followed by persons with 'thin' build which forms 35.5% of frisk.

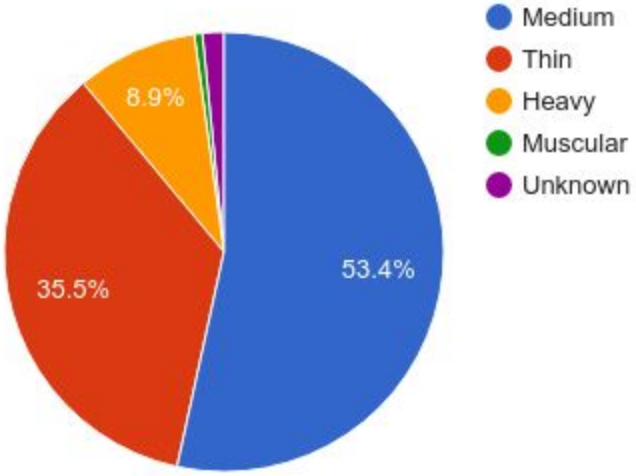

Figure 4 : Build profile of stop and frisk program

(3)Age Profile: Age profile of victims of Stop and Frisk data is presented below:

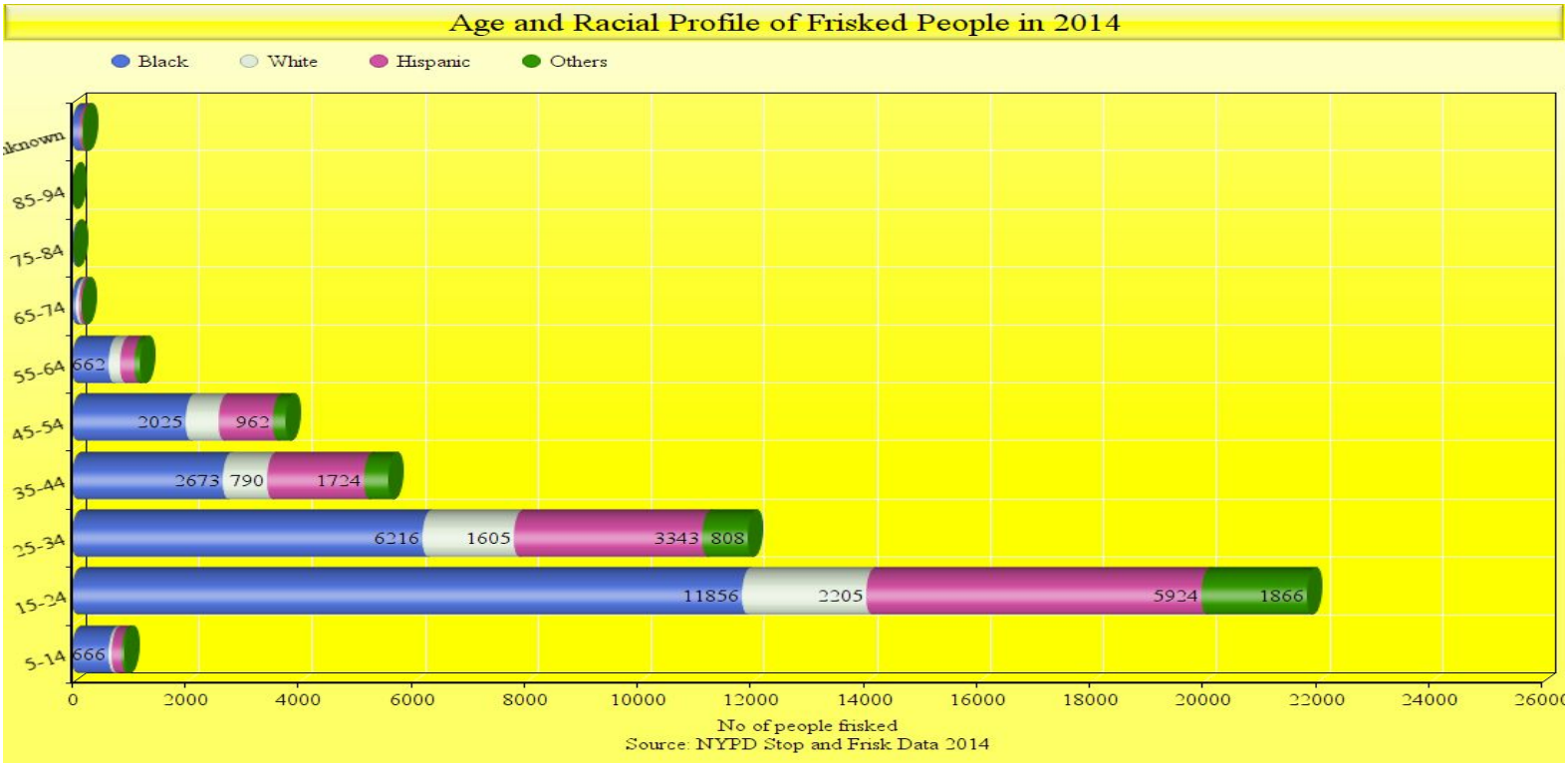

Figure 5 : Age and Racial profile in Stop and Frisk Data

Thus from above graph , it is clear that age group of 15 to 24 and 24 to 34 forms the major chunk of victims of stop and frisk data. It looks Black Youth is biggest target of stop and Frisk program as nearly 55% of all youths frisked. **Another fact is black kids and teen forms as much 70% of all kids and teen frisked.** As evident from graph, hispanic youth and middle aged also form huge chunk of persons(around 27%) frisked during 2014.
One positive found is there is considerable less frisking of people who are more than 55 years of age but black still forms 53.34% of aged population frisked between 55 to 64. Police has spared old age veterans from frisking as no of frisking is very less beyond age of 65.

(4) Month wise distribution of number of stop and frisks : Figure no 6 shows yearly distribution of stop and frisks. The graph also shows no of persons stopped as well as no of person frisked during a particular month.We observe that frisking occurs briskly during start of year and all time high during January.However it goes down during latter half of year and becomes all time low at the end of the year owing to festival season.The percentage of arrests made seems fairly constant throughout the year hovering around 13% to 18% of total stops during that month.However, point to be investigated is why stop and frisk went down in latter part of year suddenly after June.

Mean From January to June - 4694
Mean From July to December- 3188

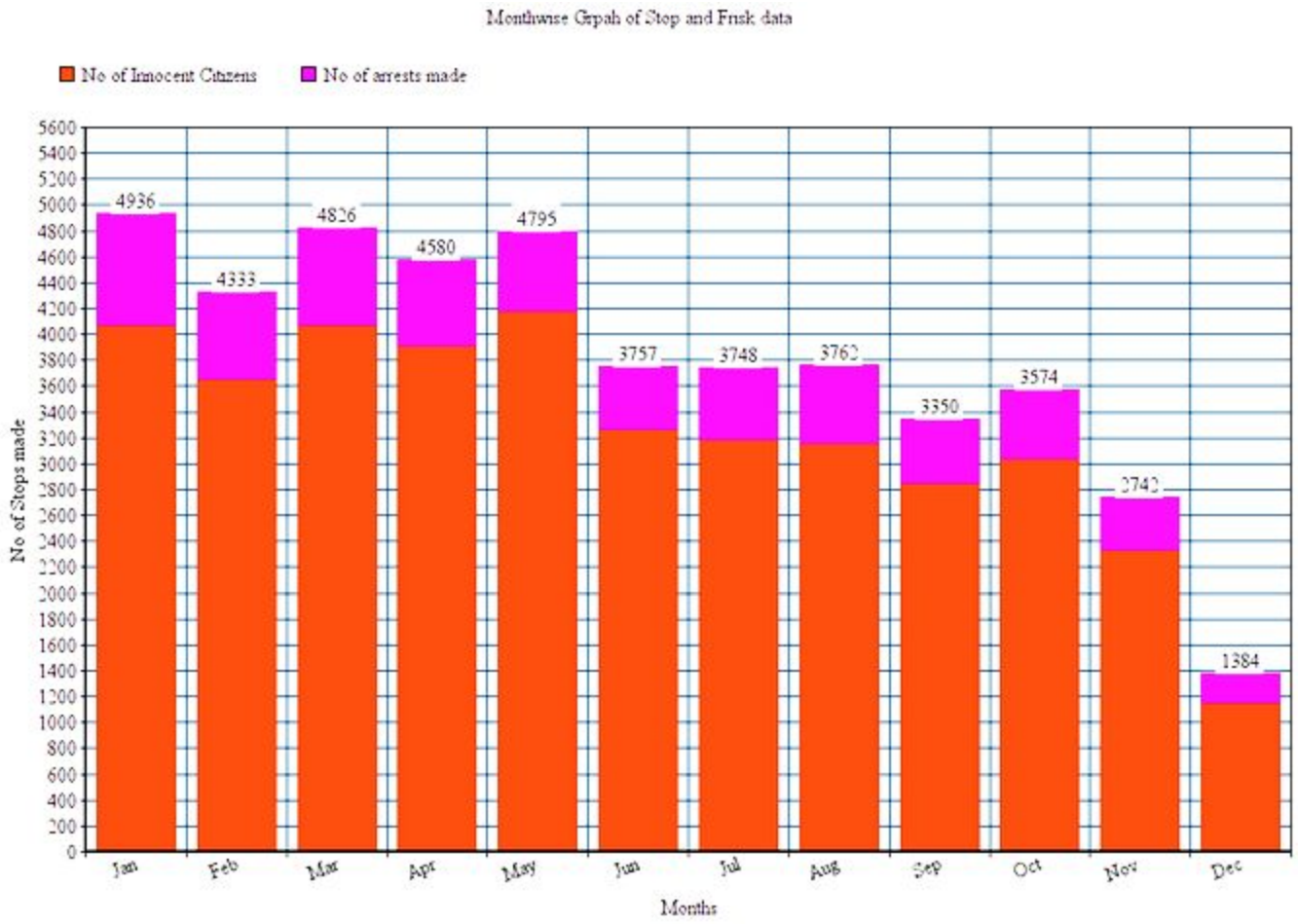

Figure 6 : Number of Persons frisked and number of arrests made per month

(5) Timewise distribution of frisking and arrest rate : Temporal distribution of data throughout the day is being analyzed to find out the time of day when police becomes most active.Corresponding arrests made is also shown to indicate the success or futility of practise.

It is observed that after the wee hours of morning, no of stop and frisks increase linearly with sharp rise around 8 PM to 11 PM. It remains stable during afternoon hours.It is also observed that as number of frisk increases, number of arrests also increase however not in the proportion of increases of frisks. Since 8PM to 11PM seems like most busy hours when people are on street , it results into increase in no of people being frisked,However, no of frisks drops during night as it reduces to 236 from 3246 during 12:00 AM to 7:00 AMWe observe that no of arrest made also reduces to 28 at 7AM and then increase linearly to 576 at 9PM. However in same time range number of people frisked increase from 236 to 3857. **Thus number of innocent people being frisked increase from 228 to 3249 from 7 AM to 9 PM**

Another visible observation is no of people frisked reduces steeply from midnight to morning hours and then increases steeply from morning to afternoon.Increase becomes much more steep from 8PM to late night. Thus forming a sort of cycle over day

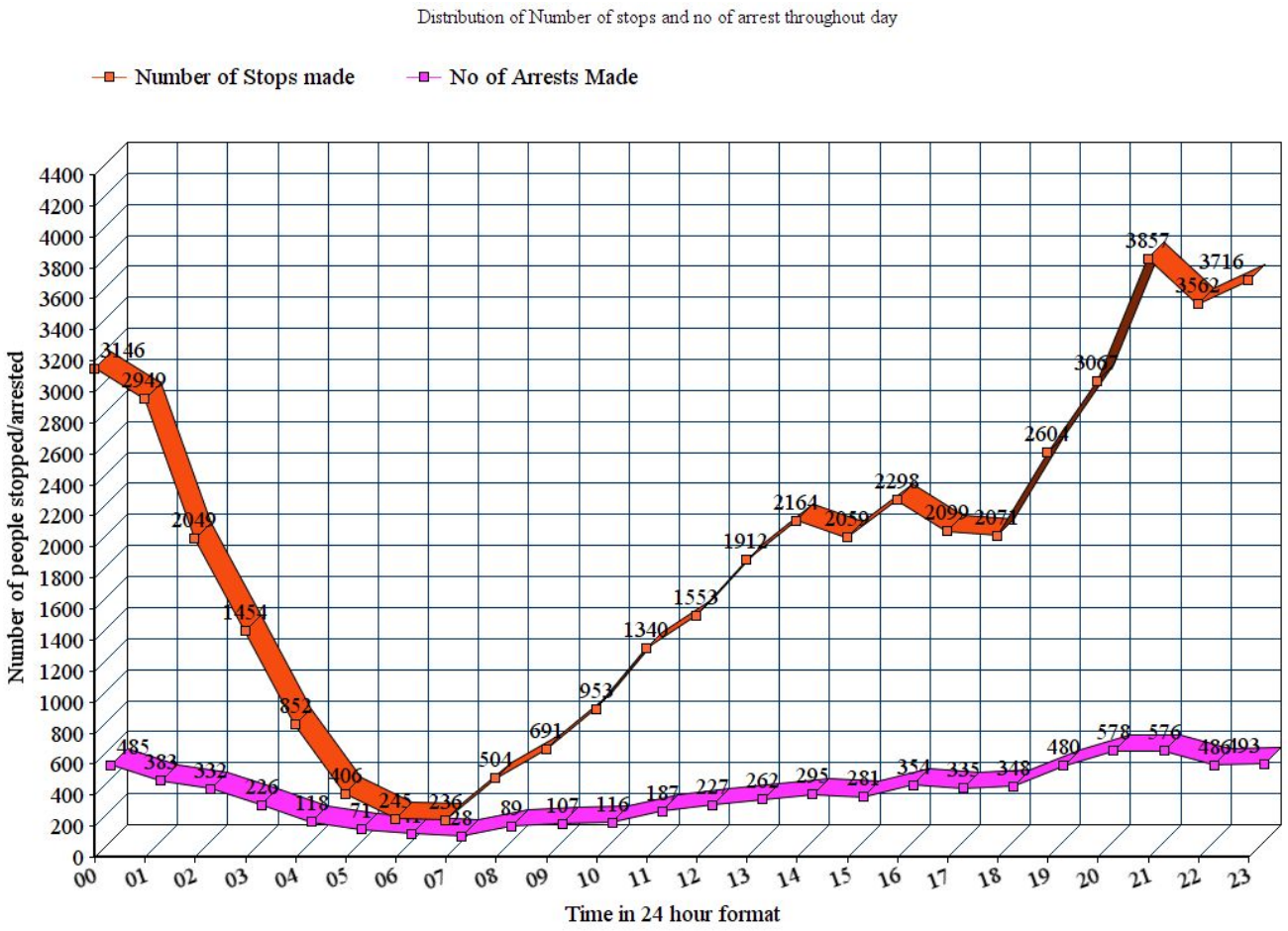

Figure 7 : Time wise distribution of Number of stops made and no of arrests made over 24 hours of a day

(6) Arrest distribution during the day: Now, arrest success rate throughout day is also analyzed. Figure 8 plots arrest % of total number of people frisked. We observe that arrest percent is fluctuating graph. It seems to indicate that police makes good number of arrest over one hour and futile suspicion over next hour. Arrest percent remains flat low during afternoon hours indicating that no of innocent people caught during afternoon are high. It is also observed that there is increase in arrest percent around 7PM to 8PM(19:00 to 20:00h). However one interesting point is **increase in arrest percent is not as steep as increase in stop and frisk around 8 to 11PM when compared from fig 7.** Thus it indicates, that police becomes hyperactive during evening hours,resulting in more number of arrests.Interestingly arrest percent start decreasing from 8PM reducing from 18.84% to 13.26% at midnight indicating innocent people becomes more victim of stop and frisk during the late night.
Interestingly, no of stop and frisk reduce to very low around 8AM, however success rate at 5AM and 8 AM increases indicating that more person are likely to be arrested and less number of innocents harassed during that time period.

Figure 8 : Distribution of Arrest made from stop and frisk over the 24 hours of day

Suggestions and Recommendations: Based on Data analyzed above, we realize average arrest rate hovers between 13% to 18% indicating a substantial at least 82% of people have to through the Stop and Frisk Program. We observed summon rate is as low as 2.65%. Thus innocent New Yorkers becomes more victim of stop and frisk program. Also , despite the court ruling of 2013, there still seem to be evidence of racial bias as blacks and hispanic people are stopped more than other communities and their stops is highly disproportioned to their ratio in population.
Following changes in Stop and Frisk policy are recommended:

1. Make Stop and frisk program to be more accountable and monitored. On evidence of racial bias, strict action should be taken
2. **A machine learning model should be developed to identify the innocent and suspicious person as well based on data from history.It wil result in less number of innocents being harressed.**
3. Proper training to police in order to remove any sort of racial bias while active on duty.
4. Use advanced technology in tracking movements of people and suspicion activity

Conclusion: Thus based on data analyzed above, we can conclude that a black or is more likely to be frisked than white and others. It is also observed that youth is main target of this program. Black teen as well as youth are likely to frisked more than any other community followed by hispanic community. There is evidence that people with medium build and thin build are more likely to be frisked. It may be because majority of new york population is either of thin build or medium build. Other factor could be inherent bias that medium and thin build allows a person to commit crime more easily.
Another interesting fact is police is more active during early part of year and becomes less active during second half of year reducing to lowest during december both in terms of number of stop and frisk as well as in terms of number of arrest made. From time distribution, we observed that police becomes hyperactive during evening and steep increase in number of people being frisked during that time period. Success rate of arrests fluctuates from 13% to 18% throughout the day indicating that still a huge percentage of at least 82% is frisked daily despite being innocent.

References:

[1] http://www.nytimes.com/2015/06/14/nyregion/eric-garner-police-chokehold-staten-island.html?_r=0
[2] https://ccrjustice.org/stop-and-frisk-human-impact
[3]https://ccrjustice.org/home/what-we-do/our-cases/floyd-et-al-v-city-new-york-et-al
[4]http://www.nyc.gov/html/nypd/downloads/pdf/analysis_and_planning/2012_sqf_file_spec.pdf