

New York City Stop-Question-Frisk Analysis

Daniel Bashir, Jonathan Hayase, Mazda Moayeri, Alicia Ngo

13 December 2016

1. Introduction

Stop-question-frisk is a policy in New York that allows police officers to stop, question, and frisk pedestrians for contraband. Every time a police officer uses this policy, the officer must complete forms about the stop. The stop-question-frisk dataset for 2010 is based off of these forms. The dataset includes information about race, stop time, whether the person was frisked, etc.

From the NYCLU paper, we examined two statistical statements. First, we checked whether Black and Latino New Yorkers were more likely to be frisked than white New Yorkers and, among those frisked, whether they were less likely to be found with a weapon. Second, we checked whether 52 percent of those stopped were frisked, and of those frisked, whether a weapon was found only 2 percent of the time. We found that both statements were true.

For the first statement, we saw that Black and Latino New Yorkers are disproportionately frisked, but less likely to have a weapon, thereby demonstrating that police officers may be racially profiling through SQF. For the second statement, we see that over half of suspects stopped were frisked, though very few suspects had weapons, thereby demonstrating that police officers may be unreasonably frisking suspects.

Throughout the rest of this report, we will provide our statistical analysis of these two statements and provide our reasons for recommending the discontinuation of SQF.

2. Responses to Required Questions

Question I.

Statement

Because all of the data provided are reported by the police officers who stopped the suspects, it is possible that there is misreported data. For example, an officer may misidentify an Asian/Pacific Islander as Hispanic, or vice-versa. We would like to explore what effect such misidentifications can have on the collected data. To simplify our analysis, we will only consider the subset of the data for which suspects were identified as Asian/Pacific Island or Hispanic. Suppose that a Hispanic person has a 95% chance of correctly being identified as Hispanic, and otherwise is misidentified as Asian/Pacific Islander, and an Asian/Pacific Islander has a 95% chance of correctly being identified as Asian/Pacific Islander, and otherwise is misidentified as Hispanic. Using this subset of the 2010 data, determine the probability that a stopped suspect is Hispanic and the probability that a stopped suspect is Asian/Pacific Islander. Use this to determine the probability that someone who is identified by the officer as Hispanic is actually Hispanic and the probability that someone who is identified as Asian/Pacific Islander is actually Asian/Pacific Islander. What does this say about how we utilize the `race` column of the data?

Answer

In the first part of this problem, we have two unknowns - the actual proportions of Hispanics and Asian/Pacific Islanders. We can set up an equation for each reported amount of people identified as a certain demographic,

giving us two equations with two unknowns. The probability that a stopped suspect is Hispanic is 26.2%. The probability that a stopped suspect is Asian/Pacific Islander (API) is 2.1%. From here, we can apply Bayes' rule for conditional probabilities to determine the other desired quantities. The probability that someone who identified by the officer as Hispanic is actually Hispanic is 99.73%, leaving a very small false positive of 0.27%. The probability that someone who is identified as API is actually API is 49.6%, an extremely high false positive rate of 50.4%. Curiously, both demographics are equally recognizable in the sense that 95% of each group will have their demographic correctly reported. However, since the Hispanic group is so much larger, the 5% of misidentified Hispanics makes up for over 50% of Asian/Pacific Islander group, while the 5% of misidentified Asian/Pacific Islanders only causes a .27% false positive.

Question II.

Statement

Compare the rates at which suspects stopped in 2010 were frisked, broken down by race. Are the differences in rates between the various groups statistically significant? Which borough(s) had the largest differences? The smallest?

Answer

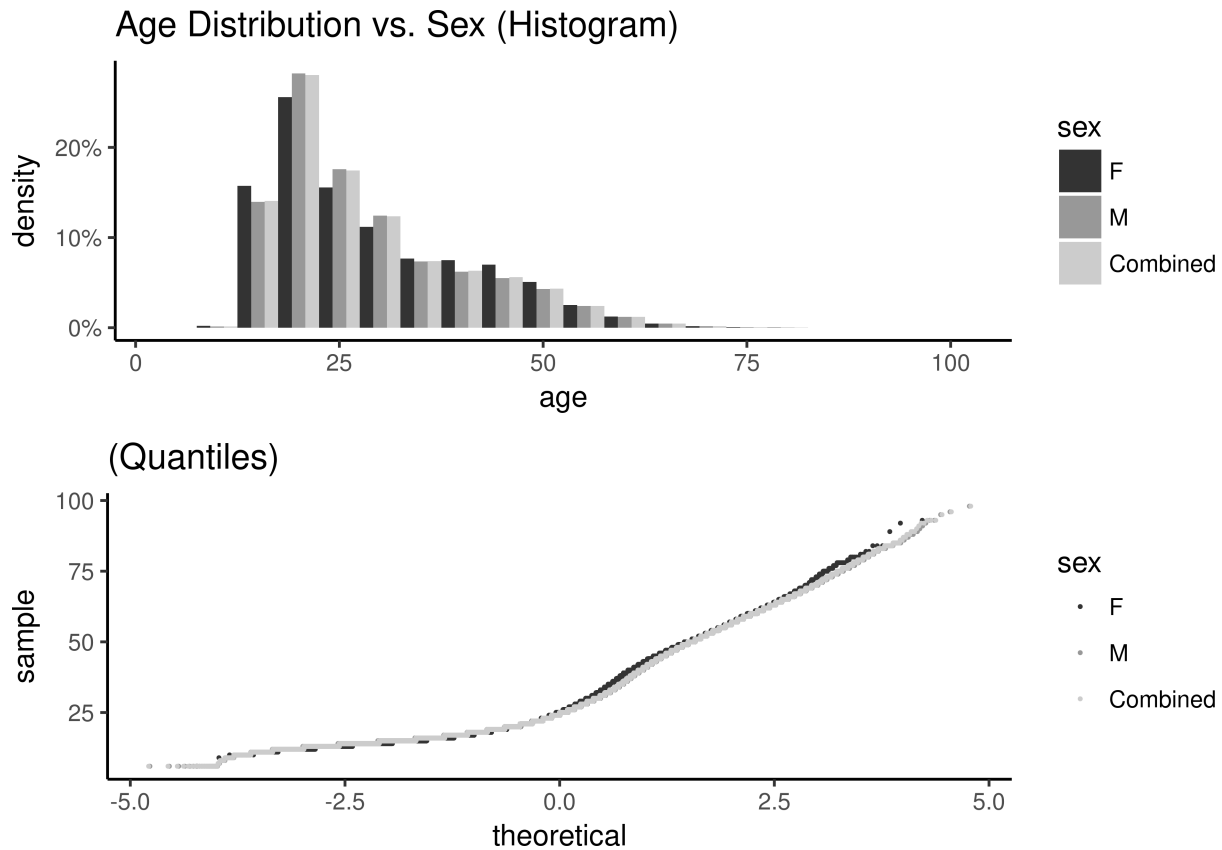
The following t-tests were significant in comparing the frisking rate by race: Whites had a significantly smaller probability of being frisked when compared to every other demographic group. Most groups had significant differences in their respective probabilities of being frisked, with black hispanics having the highest probability, followed by White Hispanics, Blacks, Natives, and finally, Asian/Pacific Islanders (in that order). Natives and Asian/Pacific Islanders did not have a significant difference in probabilities, however. When broken down by city, we found that there were significant differences in rates between every borough except Brooklyn and Manhattan. The order (from highest rate to lowest) was Bronx, Queens, Brooklyn, Manhattan, Staten Island. When looking at demographics in each borough, we found that the Bronx had 28% greater proportion of minorities (blacks and hispanics) than Staten Island, which was consistent with the conclusions from the earlier part of the problem.

Question III.

Statement

Compare the distribution of ages for male suspects in 2010 with the distribution of ages for female suspects in 2010. Use `qqnorm` to determine if they are normally distributed, and compare them with each other by using `qqplot`. Are the age distributions the same? Compare the age distribution for male suspects to that of the entire population of suspects that were stopped in 2010. Type `?qqplot` for help on how to use the command. Note that some ages are reported as 0 or 999 if the officer did not know the age, so you may want to throw out the extraneous data first.

Answer



The distributions appear similar, with subtle differences. The shapes of the distribution appear approximately equivalent, and the distributions do not appear normal. The distribution for male pedestrians is much more similar to the overall distribution, because more than ten times as many males were stopped as females. However despite this, the distribution of ages for female pedestrians is also very close to the combined and male distributions.

Question IV.

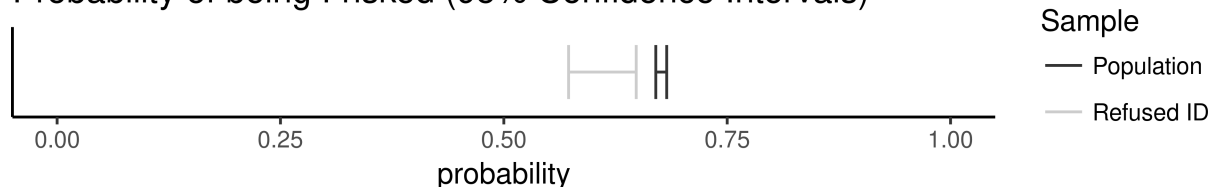
Statement

Find the probability, with a 95% confidence interval, that a suspect was frisked (a) for the entire population in 2015, and (b) for suspects in 2015 who refused to provide identification, and determine whether suspects who refused to provide identification had a different probability of being frisked than the population at large.

Answer

The 95% confidence interval for the probability that a person is frisked is (0.670, 0.682). However, when we restrict the sample to only pedestrians who refused to provide and ID, the probability that the 95% confidence interval for the probability the person is frisked becomes (0.572, 0.648).

Probability of being Frisked (95% Confidence Intervals)



As the two probability regions are disjoint, we may conclude with at least 95% confidence that the probability of being frisked after refusing identification is not the same as the probability of being frisked overall.

Question V.

Statement

For the 2010 data, decide which of the following binary factors: `arstmade`, `searched`, `inside`, `sumissue`, `frisked`, `weap`, `contrabn`, `radio`, `pf` had a significant effect on the length of the stop (`perstop`) by using linear regression. Make sure to check your residuals for normality, and apply an appropriate transformation to `perstop` or remove outlier points if it does not look normal (see your notes from Lecture 11 to review how to do this). Note that `perstop` is a discrete variable, so you are looking for an approximately normal distribution for the residuals. Consider the p-values for the coefficients and the R^2 value for your regression model. What do they indicate about how the factors affect the length of the stop? Recall that $R^2 = \frac{SSR}{SST}$. How much of the variability in `perstop` is due to the explanatory variables you have selected? Why does this make sense?

Answer

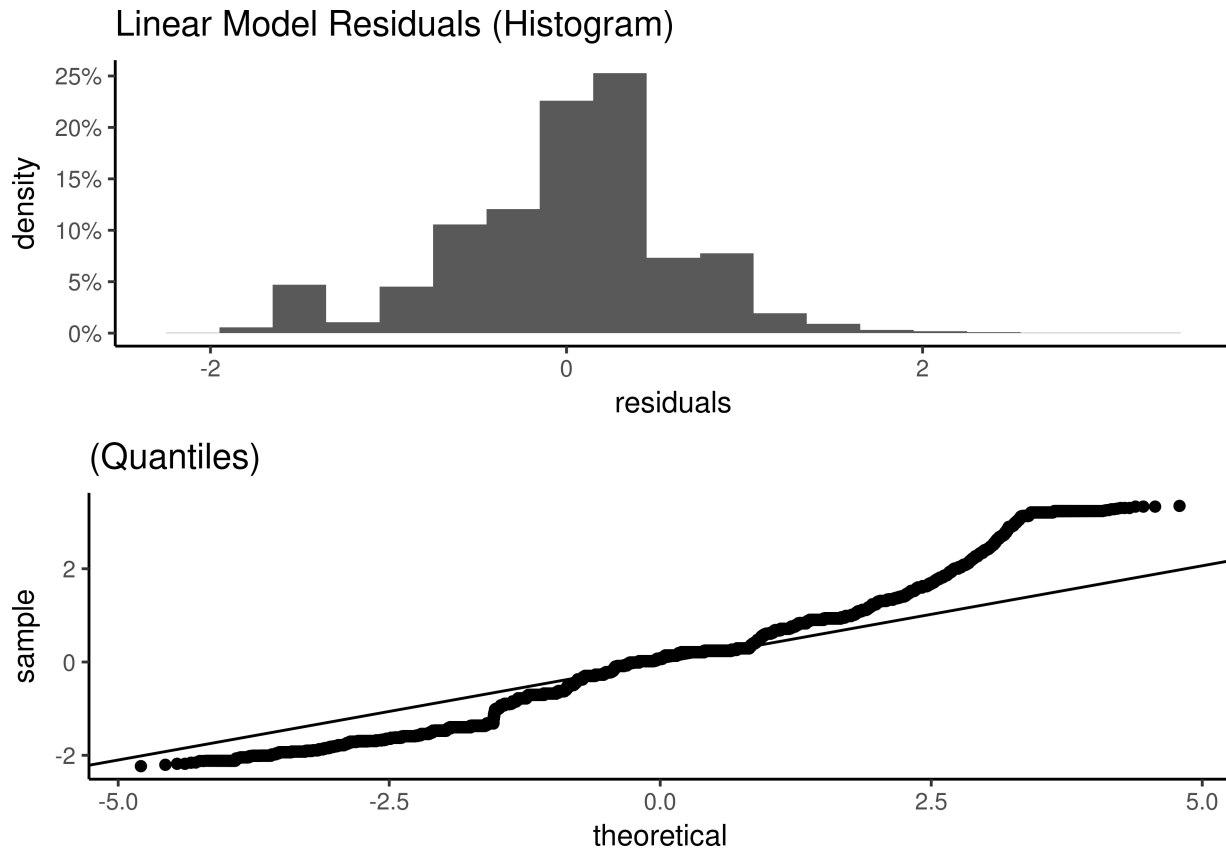
```
## Start: AIC=-532647.2
## perstop ~ arstmade + searched + inside + sumissue + frisked +
##      weap + contrabn + radio + pf
##
##           Df Sum of Sq    RSS    AIC
## - weap      1      0.3 247654 -532649
## <none>                247653 -532647
## - contrabn   1      6.9 247660 -532633
## - frisked    1     100.6 247754 -532405
## - inside     1     260.6 247914 -532017
## - arstmade    1     320.9 247974 -531871
## - searched    1     425.7 248079 -531617
## - pf          1     479.0 248132 -531488
## - sumissue    1    3815.2 251468 -523462
## - radio       1    5553.7 253207 -519322
##
## Step: AIC=-532648.6
## perstop ~ arstmade + searched + inside + sumissue + frisked +
##      contrabn + radio + pf
##
##           Df Sum of Sq    RSS    AIC
## <none>                247654 -532649
## - contrabn   1      6.8 247660 -532634
## - frisked    1     100.4 247754 -532407
## - inside     1     261.0 247915 -532018
## - arstmade    1     324.2 247978 -531865
```

```

## - searched 1      432.4 248086 -531602
## - pf      1      478.8 248132 -531490
## - sumissue 1      3816.7 251470 -523460
## - radio   1      5555.1 253209 -519320

##
## Call:
## lm(formula = perstop ~ arstmade + searched + inside + sumissue +
##      frisked + contrabn + radio + pf, data = lmdata)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.2332 -0.2993  0.0686  0.2625  3.3470
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.368522   0.001521  899.715 < 2e-16 ***
## arstmade      0.115168   0.004106   28.046 < 2e-16 ***
## searched     0.110847   0.003422   32.389 < 2e-16 ***
## inside      -0.050972   0.002025  -25.165 < 2e-16 ***
## sumissue     0.312103   0.003243   96.233 < 2e-16 ***
## frisked      0.029410   0.001884   15.610 < 2e-16 ***
## contrabn    -0.027080   0.006685   -4.051 5.1e-05 ***
## radio       0.223241   0.001923  116.099 < 2e-16 ***
## pf          0.073877   0.002167   34.086 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.642 on 600907 degrees of freedom
## Multiple R-squared:  0.04838,    Adjusted R-squared:  0.04837
## F-statistic: 3819 on 8 and 600907 DF,  p-value: < 2.2e-16

```



We conclude that all of the binary factors except the presence of a weapon had a significant effect on the stop period. We concluded that the distribution of residuals for our best-fit model may be roughly approximated by a normal distribution. This indicates that the stop period was influenced largely by factors that did not directly have an effect on the safety of the cop. The R^2 value is very small, at approximately 0.049, which means that these variables explain about 5% of the variation in stop period. This makes sense because we wouldn't expect a handful of binary variables to contain enough information to determine the period of the stop.

Chosen Questions

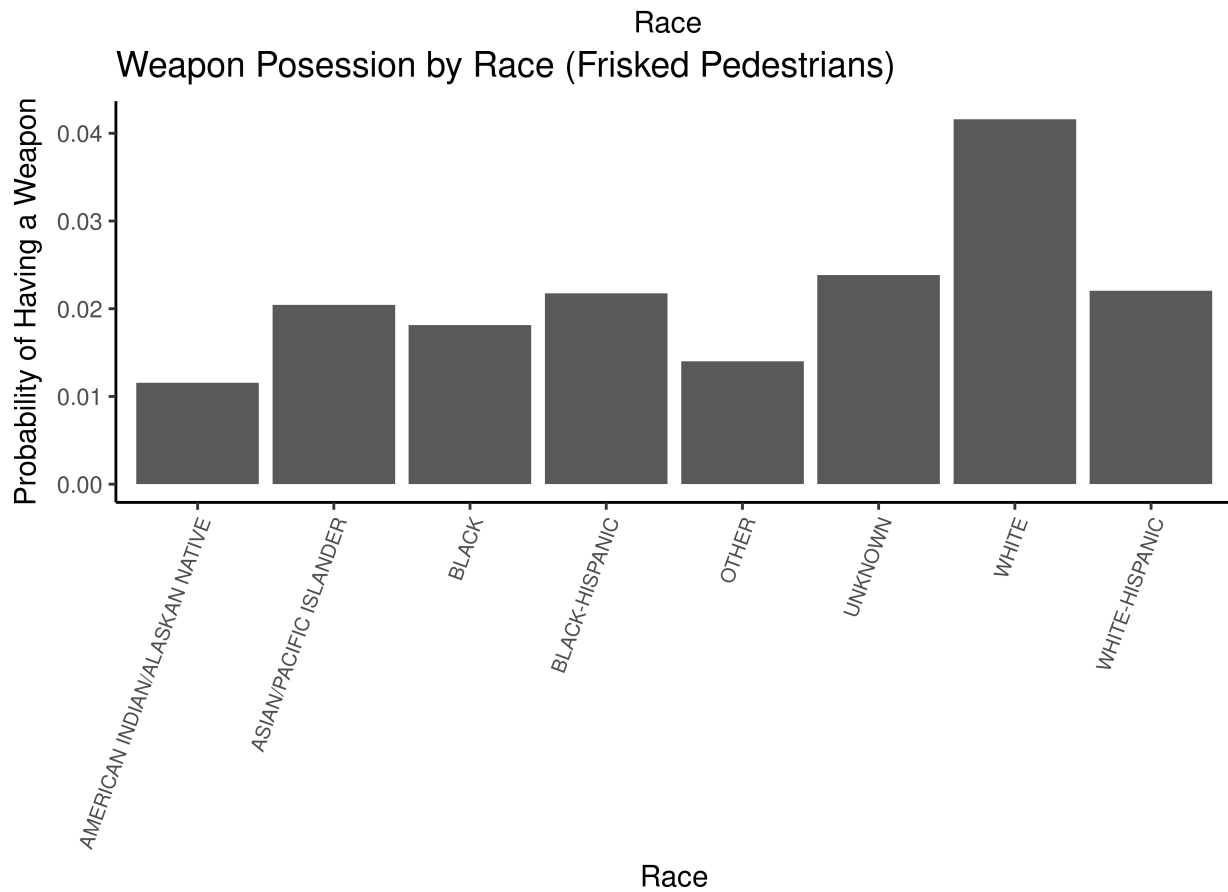
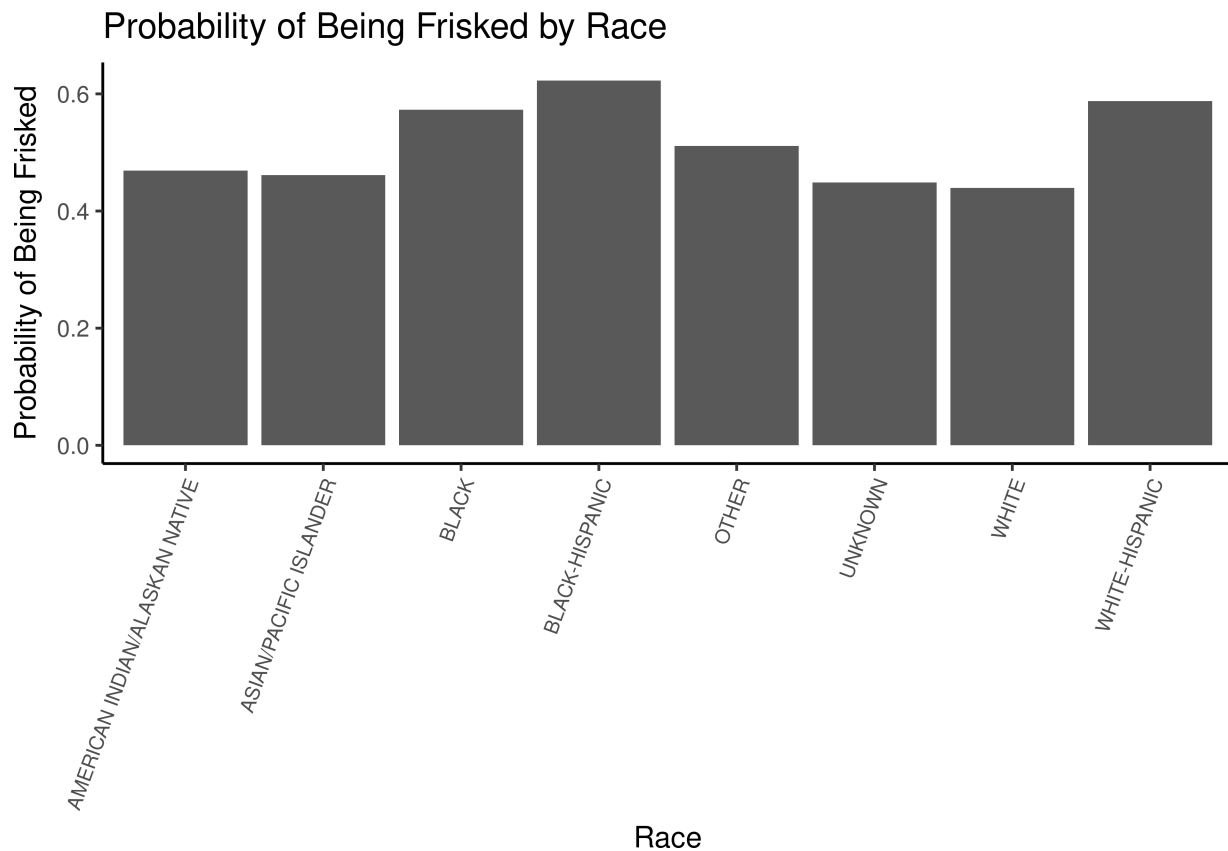
Question I.

Statement

Black and Latino New Yorkers were more likely to be frisked than white New Yorkers and, among those frisked, were less likely to be found with a weapon.

Analysis

This question is important because it illustrates two things: First, Black and Latino New Yorkers are disproportionately frisked in general. Second, this disproportionate frisking does not correlate with how many of them have weapons. As a result, these frisks may not be based on accurate evaluations.



In our analysis, we found that 58% of Blacks and Hispanics were frisked, while 44% of Whites were frisked. We also confirmed the statement that Black and Latino New Yorkers who were frisked were less likely to be found with a weapon than whites. Of those frisked, 2.0% of Blacks/Hispanics had a weapon, while 4.0% of Whites had a weapon. A bar graph shows that overall, the probability that someone who was frisked had a weapon was higher for Whites than for any other race.

Question II.

Statement

52 percent of those stopped were frisked. Of those frisked, a weapon was found only 2 percent of the time. And police officers conducted frisks in one-half of the 4.4 million stops of innocent people.

Analysis

This question indicates that the NY police officers may have been quite liberal with frisking, and that the decision to frisk could be based on assumptions and inaccurate evaluations in general. Although the report claimed 52% of those stopped were frisked, we found that 56% of those stopped were frisked. We found that of those frisked, 2.1% had weapons, which is very similar to the 2% that the report claimed.

Question III.

Statement

A statistical study of nearly 4.5 million stops produced at trial showed that only 6 percent of stops resulted in arrests and 6 percent resulted in summonses

Analysis

Our analysis of the third statement shows that in 2010, about 7% of stops resulted in arrests, and about the same percentage resulted in a summons. This confirms the following statement in the article that most of the people who were stopped “had been doing nothing wrong.”

Conclusions and Recommendations

We found that all statements seem to be fairly accurate. For reference, the three statements are provided below again.

1. Black and Latino New Yorkers were more likely to be frisked than white New Yorkers and, among those frisked, were less likely to be found with a weapon.
2. 52 percent of those stopped were frisked. Of those frisked, a weapon was found only 2 percent of the time. And police officers conducted frisks in one-half of the 4.4 million stops of innocent people.
3. A statistical study of nearly 4.5 million stops produced at trial showed that only 6 percent of stops resulted in arrests and 6 percent resulted in summonses

It appears that the SQF program is not very effective in finding people with weapons, and as a result may not be able to significantly reduce crime rates. Also, only a small number of stops resulted in arrests, suggesting that many stops were unnecessary. In addition, it appears that Blacks and Hispanics are disproportionately stopped and frisked, thereby suggesting that racial profiling has occurred. Thus, we recommend that the SQF program be discontinued.

Bibliography

Christopher, Dunn. "Stop and Frisk During the Bloomberg Administration." New York Civil Liberties Union (2014): n. pag. New York Civil Liberties Union. New York Civil Liberties Union, Aug. 2014. Web. 8 Dec. 2016. http://www.nyclu.org/files/publications/stopandfrisk_briefer_2002-2013_final.pdf.

The Editorial Board. "Policing the Police on Stop-and-Frisk." The New York Times. The New York Times, 22 June 2016. Web. 12 Dec. 2016. http://www.nytimes.com/2016/06/23/opinion/policing-the-police-on-stop-and-frisk.html?_r=0.