Table of Contents

Table of Tables

Table of Figures

Introduction

Qualifications

My name is Dr. J. Tyler Blackburn. I am the Policy Research Director at the Harris County Public Defender Office (Houston, Texas). This work was done as an independent consultant and in no way reflects my employer's position, nor did I use its resources. I have almost 20 years of experience in criminal legal system research with primary source data including analysis of police stop/arrest records for a variety of program evaluations for police and sheriff's offices. I also have a decade of litigation and monitoring support for federal civil rights cases.

Assignment

I began working for LatinoJustice PRLDEF in March 2025 to answer three related questions:

1. Are the data being provided, starting with January 1, 2024, adequate to produce an analysis for racial disparities in stops and post-stop outcomes?

2. Is there evidence of racial disparities in traffic stops?

3. Is there evidence of racial or other disparities in post-stop outcomes, including car searches and arrests?

The work to answer these questions was compensated at the rate of $150 per hour. The payments are independent of the analysis's results or litigation outcome(s). As of May 2025, this document reflects the results of my work. I consider this research ongoing; therefore, it is possible that as more or different information is shared, the analysis and document could change.

Executive Summary

Based on the work detailed through the course of this document, the answers the original questions are:

1. Are the data being provided, starting with January 1, 2024, adequate to produce an analysis for racial disparities in stops and post-stop outcomes?

Suffolk County Police Department has data available on traffic stops for 2024. The dataset contains 53 variables for over 158,000 stops. There are a limited number of duplicate values, which were removed. There were also a small number of records in which data was not recorded or missing; these missing observations did not negatively affect the analysis performed.

In 2022, Dr. Michael K. Smith[1] outlined the variables that SCPD needed to collect in order to evaluate traffic stop patterns. All but one of the recommended variables are now collected. The SCPD does not report information on passengers; the data reflect only driver information. There are two additional variables that SCPD does not collect and which could serve to enrich the data analysis: 1) arrest reason and 2) asked to consent to a search.

I assessed the entries for internal coherence and found a limited number of concerns for 365 stops and 13 searches (driver and/or vehicle). In these stop records, the recorded data for one variable contradicted data for another variable in the same stop. An example of violating internal coherence could be when a record indicated no summonses were issued but had a value recorded number of summonses issued. Separately, there are also missing values for some variables, which causes the system to make default entries. For example, missing values for driver searches (100 of 158,238) default to no, "driver not searched," and a missing value for vehicle exit reason (1 of 4,115 vehicle exits) defaults to "verbal request to exit vehicle." There were also 95 drivers with missing values for race/ethnicity for which no assumption could be made.

Comparing the dataset with the SCPD's online dashboard reveals differences between the records available and the public reporting. First, there are some duplicate stops in Quarter 2. These were removed and did not affect the analysis. Second, there were also discrepancies between the data and dashboard stops in 0.5% of the records. These errors appear to be evenly distributed across the full 158,238 stop dataset. For example, there is a 3.2% difference in the dataset and the dashboard report regarding driver or vehicle searches regardless of the race/ethnicity of the driver.

2. Is there evidence of racial disparities in traffic stops?

The veil of darkness analysis—a method of comparing stops at the same time of day when those times are in daylight or darkness to determine whether daylight increases racial disparity in stops—shows some disparities between Hispanic[2] and non-Hispanic as well as Hispanic and White non-Hispanic drivers. Overall, the nighttime and daylight average stops by race/ethnicity were consistent, though there was more variation after dark.

---

[1] U.S. District Court for the Eastern District of New York. (2020, July 15). Case 2:15-cv-02431-WFK-LB, Document 252-20. https://www.courtlistener.com/docket/4324349/1-21-v-the-county-of-suffolk/

[2] The SCPD uses the following categories for "Apparent race/ethnicity": "a. Hispanic, b. Black/African American, c. Asian/Pacific Islander, d. White non-Hispanic, e. Other." This report uses the term "Hispanic" for consistency with the data collected by the SCPD. The United States Census reports "Hispanic or Latino Origin" as a separate "ethnicity" category different from "race.". Because these are the categories used in the data reporting, they are the source of the data and analyzing the data requires using them. Still, readers should take caution and recognize that a) racial identity may not always be apparent and b) these categories tend to under-represent certain people, particularly Afro-Latinos.

Notably, the differences were not in a singular direction: after dark there were periods of both higher and lower proportion of expected stops for Hispanic drivers for certain time bands. The data neither confirm, nor disprove disparities in stop volume.

Moreover, literature emphasizes that the veil of darkness analysis may be less useful for Hispanic drivers whose race/ethnicity is not always apparent. The analysis is further limited because there is no way to ascertain the race or ethnicity of all drivers on the road. . The veil of darkness analysis may also be less reliable when considering a population of drivers working in daylight-dependent industries, like agriculture and construction, with seasonal variation in driving habits.

3. Is there evidence of racial or other disparities in post-stop outcomes, including car searches and arrests?

There are statistically significant differences between post-stop outcomes for Hispanic and non-Hispanic drivers across several categories. Hispanic drivers are issued summonses at higher rates, receive more summonses, are asked to exit the vehicle more often, are searched more frequently, and are arrested more often than non-Hispanic drivers. Yet they are less likely to be handcuffed once out of their vehicle, and searches of their vehicles result in lower “hit” rates. While the lack of an arrest reason variable makes further and more robust analysis of arrest impossible, the higher arrest rates for Hispanic drivers, specifically for those without a driver or vehicle search, is statistically significant and would benefit from further review.

Data Used

The Suffolk County Police Department hosts data on its website[3] in four separate files in comma separated value ("CSV") format. There is one file for each quarter of calendar year 2024 (Jan to March, April to June, July to September, and October to December). I downloaded each and combined them into one file to answer these questions.

The structure of the files is the same for each quarter, though the variable names did change slightly from Quarter 1 to Quarter 2. Quarter 3 has an additional variable called "StopCategories" that the other quarters do not have; it seems to collapse the categories in the variable "ReasonForStop," which all four quarters have.

I used the combined dataset to answer three questions: first, are the data adequate to produce an analysis for racial disparities in stops and post-stop outcomes; second, is there evidence of racial disparities in traffic stops; and third, is there evidence of racial or other disparities in post-stop outcomes, including car searches and arrests.

Analysis and Findings

Question 1: Are the data being provided, starting with January 1, 2024, adequate to produce an analysis for racial disparities in stops and post-stop outcomes?

The data are adequate to perform exploratory analysis on the racial disparities in stops and assess post-stop outcomes. Out of over 158,000 entries, there were only 130 duplicate values, incoherent entries, or external validation problems. There are between 1 and 100 missing values depending on the variable. Analysis is fully possible, and the conclusions are unlikely to change as there are so few records impacted.

The SCPD does not collect data on passengers, request for search consent, or arrest reason; including them would be informative. The missing variables do not change the analysis completed but do foreclose some additional questions (*e.g.* whether some groups are more likely to be asked for a consent search or whether part of the arrest could be disproportionality explained by open warrants). The findings and recommendations are listed after a discussion of variables necessary and the internal and external coherency checks.

Unique Data

There are 158,368 observations for 53 variables in the combined dataset.

All but two variables have internal order such that specific, distinct, and discrete answers occur. For example, the variable "PassengerInVehicle" has two possible answers: "YES" or "NO" and no other response is found. The "Gender" variable did have a quarter where MALE and FEMALE were entered as "M" and "F," which was easily remedied. The variable

[3] Suffolk County Police Department. (n.d.). Traffic Stop Dashboard. Suffolk County Transparency Hub. https://suffolkpd.org/Transparency-Hub/Traffic-Stop-Dashboard

"OtherOutcomeOfSearch" has 34 unique entries for 48 individuals and appear to be free text. To a point this makes sense, because a dropdown menu for any contraband possible would be impossible; however, it does result in non-uniform entries: "KNIFE," "knife," and "KNIVES" each appear. These aberrations do not impact the analysis and are noted for thoroughness only.

The variable "CCNumber" is a unique identifier for each stop, i.e., each stop is assigned a 10 digit number, and the numbers are not reused. I found 130 duplicate "CCNumber" observations in the dataset and checked all data input into each of the 53 variables. Of the 130 duplicate numbers, the majority (91%) appeared in two identical rows, eight CCNumbers appeared in three identical rows, and three appeared in four identical rows. All duplicates occurred in Quarter 2. The first instance of each CCNumber was retained for analysis and the remaining 130 duplicates were discarded. The 158,368 observations become 158,238 unique rows (stops).

Table 1: Duplicate CCNumbers by Number of Duplicates

	CCNumbers	%	Useless Rows
Two identical rows	105	91%	105
Three identical rows	8	7%	16
Four identical rows	3	3%	9
Total	**116**	**100%**	**130**

Presence of Variables

Dr. Michael R. Smith identified 22 variables Suffolk County Police Department was charged with collecting.[4] This section reviews the presence of the variable and the quality of recorded information.

Registration and Location

The first are vehicle registration and location of stop. I assumed "PlateState" combined with "Plate" (plate number) was registration. There were 158,238 observations recorded for "PlateState" with no blanks or NULLs. The plate number was present for 158,109 (99.9%) of observations. There were 30 NULL, where NULL includes any plate number with only four or more dashes (----) and no other character, and 99 blank for a total 129 missing. Note, only nine of these were recorded as being stopped for a registration problem, and they all included a "PlateState" entry, so the missing plate number does not suggest the car had no plate.

Regarding location, "Street Address" is some combination of "HouseNumber," "StreetPrefix," and "StreetName," which occur for all but four (0.003%) of the unique CCNumbers. Additionally, Latitude/Longitude coordinates were recorded in over 95% of

[4] U.S. District Court for the Eastern District of New York. (2020, July 15). Case 2:15-cv-02431-WFK-LB, Document 252-20. https://www.courtlistener.com/docket/4324349/1-21-v-the-county-of-suffolk/

the stops. There were 33 blank coordinates, and 6,460 coordinates called NULL, for a total of 3.1% missing. When all of these location options are combined, only one stop has no information on location.

Table 2: Useable and Unusable Entries for Registration and Location of Stop

Variable	Useable Observations		NULL	Blank	Total Missing	
	N	%	N	N	N	%
PlateState	158,238	100%	0	0	0	0.0%
Plate Number	158,109	99.9%	30	99	129	0.1%
Street Address	158,234	99.997%		4	4	0.003%
Lat/Long Coordinates	151,739	95.9%	6,460	33	6,499	3.1%

Finding: The dataset provides robust data on vehicle registration and location information when all information on location is used.

Data at Conclusion of Stop

There is a second set of nine (9) variables entered at the conclusion of a traffic stop. They include: (1) reason for stop, (2) duration of stop, (3) whether force was used during the stop, (4) total tickets issued, (5) whether there were any equipment violations, (6) whether the vehicle was searched, (7) reason for the search, (8) outcome of the search, and (9) whether a canine was called.

See table 3 to follow this discussion.

1. Reason for stop is answered using the variable "ReasonforStop," which is fully populated with 11 answer possibilities.
2. Duration of stop is fully populated with 158,238 using the variable "DurationofStop;" there are four answers possible using minute ranges. The minute range possibilities overlap. There is one 30+ and one 60+ option and it was assumed 30+ means 30 to 60 minutes and 60+ is over 60 minutes.
3. The use of force question is answered by "ForceUsedDuringStop" variable, which has two possible answers: yes or no. This is filled in for 158,237 CCNumbers. There is one blank.
4. The number of summonses issued is answered using two variables, assuming "ticket" and "summons" are synonymous. If the driver is issued a summons (75,344 dispositions), then the variable "TotalSummons" should have an entry recorded. There were 16 records of a summons issued with NULL as the number (0.02% of observations that should have a number).
5. There are two ways this dataset captures equipment violations. The first is with the variable "EquipmentViolation," which has 20,507 entries (obstructed view, other, and window tint) that are not NULL. The second is when the reason for stop is given as equipment violation; there are 20,385 of these entries. The 122 violation

difference between the two sets may mean that some individuals were stopped for another reason and then an equipment violation was found. It may also mean these data are incorrectly recorded.

6. Vehicle search is captured as a yes/no question using the variable "VehicleSearch," which is filled in for all 158,238 observations. There were 1,264 searches recorded.
7. The reason for the vehicle search is filled in for all 1,264 searches recorded as yes using three possible answers: founded suspicion, plain view, and probable cause.
8. The search's outcome is also filled in for all 1,264 searches using four categories including one to indicate nothing was found ("negative result").
9. A canine call was recorded under the variable "CanineCalled" and has two possible answers (yes or no), which were recorded for all 158,238 calls.

Table 3: Useable and Unusable Conclusion of Traffic Stop Variable Entries

Variable Name	**Useable Observations**		NULL	Blank	**Total Missing**	
	N	**%**	N	N	**N**	**%**
1 ReasonForStop (11)	158,238	100%	0	0	0	0.0%
Cell call/text (9,001 or 5.7%), Equipment Violation (20,385 or 12.9%), Moving Violation (27,496 or 17%), Other Moving Violation (20,269 or 13%), Probable Cause of a Crime (249 or 0.2%), Reasonable Suspicion of a Crime (504 or 0.3%), Red Light (2,610 or 1.6%), Registration/ Inspection/ Plate Violation (35,934 or 23%), Seat Belt (2,917 or 1.82%), Speeding (21,695 or 13.7%), Stop Sign (17,178 or 11%)						
2 DurationOfStop (4)	158,238	100%	0	0	0	0.0%
In minutes: 0-15 (131,345 or 83%), 16-30 (21,358 or 13.5%), 30+ (3,710 or 2.3%), 60+ (1,825 or 1.2%)						
3 ForceUsedDuringStop (2)	158,237	100%	0	1	1	0%
Yes (189 or 0.12%), No (158,048 or 99.88%)						
4 TotalSummonses (10) *if DriverDisposition=Summons* (75,344)	75,328	99.98%	16	0	16	0.02%
1 (40,965 or 54%), 2 (13,206 or 18%), 3 (7,172 or 10%), 4 (4,897 or 7%), 5 (3,597 or 5%), 6 (2,059 or 3%), 7 (1,305 or 2%), 8 (891 or 1%), 9 (417 or 1%), 10 (819 or 1%); Average = 2.2 and Median = 1						
5a *EquipmentViolation* (3) is not blank	20,507	100%	0	0	0	0%
Obstructed View (1,034 or 5%), Other Equipment Violation (14,839 or 72%), Window Tint (4,634 or 23%)						
5b If *ReasonForStop* = *equipment violations* & EquipmentViolation (3)	20,384	100%	1	0	1	0%
Obstructed View (1,024 or 5%), Other Equipment Violation (14,733 or 72%), Window Tint (4,627 or 23%)						
6 VehicleSearched (2)	158,238	100%	0	0	0	0.0%
Yes (1,264 or 0.8%), No (156,974 or 99.2%)						
7 If *VehicleSearch = YES* ReasonForSearch (3)	1,264	100%	0	0	0	0.0%
Founded suspicion of a crime with recorded consent (413 or 32.7%), Probable cause of a crime / automobile exception (477 or 37.7%), Plain view (374 or 29.6%)						

Variable Name	Useable Observations		NULL	Blank	Total Missing	
8 If *VehicleSearch* = *YES* OutcomeOfSearch (4)	1,264	100%	0	0	0	0.0%
Drugs (259 or 20.5%), Weapon (27 or 2.1%), Negative Result (783 or 62.0%), Other Contraband/ Evidence (195 or 15.4%)						
8 CanineCalled (2)	158,238	100%	0	0	0	0.0%
Yes (116 or 0.07%), No (158,122 or 99.93%)						

Finding: Variables collected regarding the stop's purpose and disposition are present in the dataset and contain useable entries.

Recommendation: Put call times on/off instead of duration for maximum accuracy or, at minimum, change the duration to distinct non-overlapping categories.

Recommendation: Consider creating a category for each stop reason and then give details. The data collection and set may be more complicated, but the data will be more complete.

Individuals Involved in Stop

Third, for each individual involved in the stop, the officer must record following 11 points of information: (1) whether the individual is a driver or passenger, (2) approximate age, (3) gender, (4) race/ethnicity, (5) whether a search of their person was conducted, (6) reason for the search, (7) outcome of the search, (8) whether the individual was asked to exit the vehicle, (9) where the individual was placed, (10) whether the individual was restrained, and (11) disposition of the stop.

It is imperative to highlight there is no variable indicating passenger or driver for the recorded observations. The structure of the dataset implies these data all reflect the driver, because of variable names, (e.g., "Was**Driver**Handcuffed," "**Driver**Searched," "**Driver**Disposition," etc.) Although there is a variable indicating the presence of passengers in the vehicle ("PassengerinVehicle" with "Yes" or "No" as options), there is no way to indicate whether passengers were handcuffed, searched, or arrested. Though the name of some variables indicate outcomes were for the driver, others are ambiguous (e.g. "ExitedVehicle," "ForceUsedDuringStop," and "PoliceActionTaken"). Given this framework, for the purposes of this analysis, all variables are assumed to be driver specific.

See Table 4 to follow this discussion.

1. There is not a variable to distinguish passenger from driver, as stated above.
2. There is a variable called "ApproximateAge" that reports the driver's age in seven different ranges. There are 97 missing "NULL" entries or 0.06%.
3. Gender is recorded as male or female for all but 96 of the unique CCNumbers.
4. Race/Ethnicity is recorded by the officers for all but 95 unique CCNumbers.

5. The search of the driver is recorded as a binary variable (yes or no) for 158,138 drivers. There were 44 NULLS and 56 blanks for 100 missing records.
6. There are three possible reasons given for the driver search and there is a reason for all 1,594 drivers coded as "yes" for driver searched.
7. The outcome of the search is recorded for all 1,594 drivers coded as "yes" for driver searched. One outcome is negative result, which indicates an unproductive search.
8. There is a variable called "ReasonforVehicleExit" that has three answer options. One option is "Verbal Request / Verbal Command," which is assumed to be synonymous with "asked to exit." The other two options are "Forced Used" and "Spontaneous" – the latter is not a request and the former assumed too ambiguous to be a request. There is one stop in which "ExitedVehicle" is recorded as "YES," but the reason is not given. It is unclear if this person received a verbal request, was removed by force, or spontaneously exited the vehicle.
9. It is unclear if "where the person was placed" refers to post vehicle exit or following being handcuffed by the officer. In the data, the variable seems to follow being handcuffed both logically (almost all "was driver handcuffed" = yes have a response on in or out of the police vehicle) and in the file's order.
10. The drivers who are coded as "YES" exited vehicle (6,133) have entries for "WasDriverHandcuffed" ("YES" or "NO") in 6,130 rows and the data are missing or coded NULL for three. Assuming this is the only restraint, the data are available to answer this question for drivers only.
11. There were 158,137 drivers of 158,238 or 99.98% with one of seven possible dispositions recorded.

Table 4: Useable and Unusable Person Information Recorded at a Stop

Variable	Useable Observations		NULL	Blank	Total Missing	
	N	%	N	N	N	%
1 NA: Driver or Passenger					158,238	100%
Not included						
2 ApproxAge (7)	158,141	99.94%	97	0	97	0.06%
LT 16 (132 or 0%); 16-25 (33,203 or 21%); 26-35 (47,665 or 30%); 36-45 (34,927 or 22%); 46-55 (23,787 or 15%); 56-65 (13,802 or 9%); 66+ (4,625 or 3%)						
3 Gender (2)	158,142	99.94%	96	0	96	0.06%
Male (112,024 or 71%); Female (46,118 or 29%)						
4 RaceEthnicity (5)	158,143	99.94%	95	0	95	0.06%
Asian/Pacific Islander (4,930 or 3%); Black/African American (29,643 or 19%); Hispanic (45,083 or 29%); Other (8,431 or 5%); White non-Hispanic (70,056 or 44%)						
5 DriverSearched (2)	158,138	99.94%	44	56	100	0.06%
Yes (1,594 or 1%); No (156,544 or 99%)						
6 *If DriverSearch = YES*, ReasonForSearch (3)	1,594	100%	0	0	0	0%
Reasonable suspicion of a crime / frisk for weapon (658 or 41%); Probable cause of a crime / vehicle exception (788 or 49%); Founded suspicion of a crime with recorded consent (148 or 9%)						
7 *If DriverSearch = YES*, OutcomeOfSearch (4)	1,594	100%	0	0	0	0%

Variable	Useable Observations		NULL	Blank	Total Missing	
	N	%	N	N	N	%
Drugs (140 or 8.8%), Weapon (11 or 0.7%), Negative Result (1,212 or 76.0%), Non Criminal Object (184 or 11.5%), Other Contraband/ Evidence (47 or 3.0%)						
8 ReasonForVehicleExit = VERBAL REQUEST / VERBAL COMMAND	4,114	100%	1	0	1	0%
ReasonForVehicleExit has three options, this is the whole of the subgroup assumed to be "request"						
9 *If WasDriverHandcuffed = Yes (1,260)* then WhereWasDriverPlaced (2)	1,255	99.6%	5	0	5	0.4%
In Police Vehicle (1,010 or 80%); Outside Police Vehicle (245 or 20%)						
10 *If ExitedVehicle = Yes (6,133)* WasDriverHandcuffed (2)	6,130	99.95%	1	2	3	0.05%
Yes (1,245 or 20%), No (4,885 or 80%)						
11 DriverDisposition (7)	158,137	99.94%	101	0	101	0.06%
Aided (4 or 0%); Arrest (1,223 or 1%); Field Appearance Ticket (278 or 0%); Lights on Voucher (300 or 0%); No Police Action (2,451 or 2%); Summons (75,344 or 48%); Verbal Warning (78,537 or 50%)						

Finding: There are no passenger indicators.

Recommendation: If the passenger outcome is of interest, record the passenger outcomes using the same CCNumber but with a different person identifier and an additional variable called Passenger/Driver. The unduplicated CCNumber with the row Passenger/Driver = Driver will be used to assess the stop and driver outcomes. Each unique passenger person identifier can be used to assess passenger outcomes.

Finding: A search reason *founded suspicion of a crime with recorded consent* only captures the drivers who gave consent, but there is not information to calculate how many were asked to consent.

Recommendation: The set should include a variable *consent requested* with "Yes" and "No" options for searches.

Finding: The dataset does not indicate why or for what offense a driver was arrested.

Recommendation: Create a variable for why an arrest occurred. The entries could be outstanding warrant, result of stop, result of search, or a list of offenses with some demarcation of warrants (not arising from this stop, simply found at this stop).

Internal Observation Coherence

Internal coherence was checked for the dataset in two ways. The first was a check to assess the responses as they exist in conversation with each other. As noted in Tables 3 to

4 above, some of these variables are nested within each other. For example, all 1,260 drivers handcuffed should have a response for "WhereWasDriverPlaced." The inverse is also true: if the driver was not handcuffed, then "WhereWasDriverPlaced" should be NULL or blank but never inside/outside the police vehicle. The second is a check on missing data, which was recorded above but assessed for impact here.

The internal group coherence concerns were calculated first, then the missing data were assessed. If a person had a grouping that was irrational, they were not assessed in the missing data group. If a person had missing data for one variable, they may also have missing data for another variable. For example, 95 drivers were missing entries on all socio-demographic variables.

There are 247 stops with incoherent data groupings. Since CCNumbers may be in multiple categories, the number of concerns do not add up to the unique stops with concerns. They include:

Table 5: Incoherent Data Groups

Number	Explanation
67	Coded NO on "ExitedVehicle," but have a reason for exit listed
57	Coded NO on "VehicleSearched," but have a reason for search listed
83	Coded NO on "PassengerInVehicle," but have a number listed in "PassengerCount"
5	Coded YES on "WasDriverHandcuffed," but have NULL for "WhenWasDriverHandcuffed"
16	Coded NO on "WasDriverHandcuffed," but have information for "WhenWasDriverHandcuffed"
2	Coded NULL on "WasDriverHandcuffed," but have information for "WhenWasDriverHandcuffed"
6	Coded NO on "WasDriverHandcuffed," but have information for "WhereWasDriverPlaced"
33	Coded NO on "DriverSearched" but reason for search given
12	A subset of the 33 coded NO on "DriverSearched" but reason for search given also have a listed "OutcomeOfSearch"
6	"DriverDisposition" is NULL and "TotalSummonses" is 1 or more
247	**Unique stops with one or more incoherent entries**

Data are missing from 118 stops that do not have a coherence concern listed in Table 5 above. Some of these duplicate information from the review of variables necessary in Tables 3 and 4. They include:

Table 6: Missing Data Creating Incoherent Groups

Number	Explanation
1	"EquipmentViolation" missing when "ReasonforStop" is "Equipment Violation"
1	"ForceUsedDuringStop" missing
97	Missing socio-demographic variables – 97 without age, 96 without gender, and 95 without race. The 95 without race also have no entries for age and gender.
100	NULL or blank for "DriverSearched;" none have a reason for search or search outcome listed, so perhaps they can be assumed to be NO. Of this group 97 have a different variable with missing information (socio-demographic).
96	NULL or blank "VehicleExit;" none have reasons for vehicle exit, so perhaps they can be assumed to be NO. There were 95 were already coded as problematic due to missing information.
16	"DriverDisposition" is Summons but NULL summonses are recorded
118	**Unique stops with one or more missing entries**

In total, there were 365 stops and 13 searches (driver and/or vehicle) that lacked information or had observations recorded that did not make sense within the full narrative of the observation. For example, there was one stop that had "EquipmentViolation" missing an entry when "ReasonforStop" is Equipment Violation. Of that group, the 95 without race/ethnicity and 1 without use of force (total 96 observations) were removed from the 158,238 unduplicated stops leaving 158,142 stops. This is a reduction of 0.2% in total records, which is at most a rounding error in aggregate, and 0.6% of searches (vehicle and or driver).

If the missing information corresponded with one specific type of record, , for example if all 96 observations were Hispanic drivers searched during the stop, then the missing data would have an impact. There is no reason to conclude that these errors are anything but random.

External Observation Coherence

There is a public facing dashboard of the stops from the Suffolk County Police Department.[5] The data collected from the Traffic Stop Report and the dashboard analysis should match to the overall number of stops and the data within each category should correspond. For example, the total number of stops in the Excel sheets and subtotals for stops by race/ethnicity and reason for stop should be the same as what appears on the dashboard. This is not the case. It also means the raw data should be able to replicate the dashboard's analysis. It cannot.

[5] Suffolk County Police Department. (n.d.). Traffic Stop Dashboard. Suffolk County Transparency Hub. https://suffolkpd.org/Transparency-Hub/Traffic-Stop-Dashboard

The overall total and by-quarter total stops do not match. The unduplicated raw data have 850 fewer observations than reported on the public dashboard (158,238 versus 159,088), which means the raw data are missing at least 0.5% of the dashboard's data. It is not a problem of duplicates or unduplicated by CCNumber as the raw data have fewer observations than reports on the public dashboard. Quarters 1 through 3 dashboard numbers on March 29, 2025 and June 10, 2025 did not match the raw data numbers (numbers on the dashboard were different on each date), but Quarter 4 raw data and dashboard values do match. Although 0.5% is not likely to impact the overall analysis when considering the full dataset, the lack of consistency between the raw data and the dashboard is concerning.

Table 7: External Observation Coherence (Dashboard Comparison) Quarterly Records

	Total Stops Excel with Duplicates	Total Stops Excel Unduplicated	Total Stops Dashboard	Numerical Difference from Unduplicated	Percent Missing in Raw Data
Quarter 1	NA	41,315	41,472	157	0.4%
Quarter 2	41,410	41,280	41,675	395	0.9%
Quarter 3	NA	39,063	39,361	298	0.8%
Quarter 4	NA	36,580	36,580	0	0.0%
Total	158,368	158,238	159,088	850 More on DB	0.5%

As set forth in Table 8 below, however, subtotals, however, are much smaller for certain groups or outcomes. If all 850 missing stops share certain characteristic, observable patterns will be lost. A non-exhaustive list of subcategories illustrates the point.

- As of March 29, 2025, the dashboard reported 45,402 stops for Hispanic drivers in 2024, 319 fewer than the raw data with duplicate entries removed. This mean 0.7% of the records are missing.
- The dashboard shows four subgroups that make up Equipment Violation as of June 10, 2025: Registration, Other Equipment Violation, Window Tint, and Obstructed View.[6] The raw data report 56,289 stops in 2024 for equipment violations while the online dashboard reports 56,743, nearly five hundred more. . T.
- When the raw data's "ReasonforStop" = EQUIPMENT VIOLATION is added to REGISTRATION, the total is 777 stops fewer than the dashboard (1.2% of records different from the dashboard).

This shows that the raw data are not missing the records, just coding them more granularly. A similar problem with granularity occurs with the variable for search. The dashboard does not clarify if the data reflect driver or vehicle searches. The 2,166 observations suggest that it reflects that either the driver, the vehicle, or both were

[6] Because "registration" is a separate entry in the raw dataset, the "Equipment Violation" and "Registration" entries in the raw dataset must be combined in order to compare them to the dashboard.

searched, as 2,097 unique stops fall into this category.[7] That is a difference of 69 stops between the dashboard and the raw data, meaning 3.2% of records are different or missing. This increases to 3.4% for searches when the driver is Hispanic. The dashboard reports 707 searches while the dataset reports only 683.

Table 8: External Observation Coherence (Dashboard Comparison) on Searches

	Total Excel Duplicates	Total Excel Unduplicated	Total Dashboard	Numerical Difference from Unduplicated	Percent Missing in Raw Data
Stop Hispanic	45,167	45,083	45,402	319	0.7%
Equip Violation	20,410	20,385[8]	57,066	36,681	
Registration	35,934	35,904	36,407	503	
EV & Registration	56,344	56,289	57,066	777	1.2%
Search (Driver)	1,616	1,594	NA – Not Separate		
Search (Vehicle)	1,295	1,264	NA – Not Separate		
Search Either	2,132	2,097	2,166	69	3.2%
His Search Either	696	683	707	24	3.4%

Finding: Public information and raw data do not produce the same results.

Recommendation: If data are recoded, cleaned, or updated in any way, the new set needs to be uploaded to the Raw Data repository on the website.

Finding: Public information about search seems to produce a binary yes/no variable that reports if the driver and/or the vehicle is searched.

Recommendation: Clarification on the website or inclusion of that variable on the raw sets would improve utility.

Finding: Equipment Violation includes other observations on the dashboard that in a way that is not clear in the dataset.

Recommendation: Clarification on the website or inclusion of that variable on the raw sets would improve utility.

The missing race/ethnicity observations show White non-Hispanic, Black/African American, and Hispanic are all missing about the same proportion of stop records. The duplicated records are lower than the dashboard numbers, meaning the change between

[7] Further evidence is that the quarter 4 number, the only quarter than matched the dashboard, reports the same number when both search types are combined and the Asian/Pacific Islander and Other "search either" is also the same number.
[8] ReasonforStop = EQUIPMENT VIOLATION

the raw data and the dashboard is not occurring because of the duplicate exclusions. Only one record with Race/Ethnicity was coded NULL, so it is not a problem of subsequently cleaned data.

Table 9: External Observation Coherence (Dashboard Comparison) Search by Race/Ethnicity

	Total Excel Duplicates	Total Excel Unduplicated	Total Dashboard	Numerical Difference from Unduplicated	Percent Missing in Raw Data
Asian/PI	21	21	21	0	0%
Black/AA	706	698	720	22	3.1%
Hispanic	696	683	707	24	3.4%
White non-His	668	664	688	24	3.5%
Search Either	**2,132**	**2,097**	**2,166**	**69**	**3.2%**

Finding: Public information and raw data do not produce the same results.

Recommendation: If data are recoded, cleaned, or updated in any way, the new set needs to be uploaded to the Raw Data repository on the website.

Data Adequacy

The data are adequate to do exploratory analysis on the racial disparities in stops and assess post-stop outcomes. There were a limited number of duplicate values, which were easily addressed, and almost all variables were present. There is a notable absence of information on passengers. All data reflect drivers only. The set would benefit from information on both (1) arrest reason and (2) if the driver was asked to consent to a search.

The dataset seemed to have a limited number of incoherent data entries: there were 365 stops and 13 searches where the variables produced inconsistent results that suggest some type of error. For example, a stop with the variable "ReasonforStop" entered as "Equipment Violation" but the variable "EquipmentViolation" is internally incoherent. The rate of internal incoherence is only 0.2% of total stops and 0.6% of searches and therefore unlikely to affect the analysis. The external validation showed that the raw data contain about 850 fewer records (about 0.5%) than the public-facing dashboard. None of these proportions are concerning unless there is reason to believe there are systematic problems for certain stop types or subgroups, which was not indicated in the external coherence analysis for searches or race/ethnicity. It is, however, a concern for stop reasons. The example of Equipment Violations and searches showed the raw data has more granular information.

Question 2: Is there evidence of racial disparities in traffic stops?

A "veil of darkness" analysis (described below) was performed on the 2024 traffic stop data matched to the recorded sunrise and sunset times for Suffolk County to create the light and dark periods. The analysis suggested some difference in stop rates for Hispanic versus non-Hispanic drivers and Hispanic versus White non-Hispanic drivers. However, Hispanic drivers were both over and underrepresented for certain time bands in the dark period, making it difficult to calculate the difference and what it signifies.

These discrepancies emphasize why the literature does not recommend using the veil of darkness analysis for Hispanic drivers. Because the ethnicity of Hispanic drivers is not always readily apparent, error rates for correctly identifying Hispanic drivers can be high, particularly when race or ethnicity is reported by the officer based on their perception, as is the case in Suffolk County. A chi-square test (to assess whether the distribution between categorical variables – race/ethnicity and stop – is the same) did not produce a statistically significant p-value for the differences between light/dark or time bands.

Data

As detailed above, there were 130 observations removed because they were duplicate records, 95 removed because of NULL entries for race/ethnicity, and 1 missing a "yes or no" code for use of force. Racial disparity was evaluated using a universe of 158,142 stops for which the record is assumed to be kept for the driver.

There is one variable for race/ethnicity in the dataset. For each stop, the driver is categorized as: Asian/Pacific Islander (4,930 or 3%); Black/African American (29,643 or 19%); Hispanic (45,083 or 29%);[9] Other (8,431 or 5%); or White non-Hispanic (70,056 or 44%) only. These parameters necessarily exclude certain intersectional identities such as an individual who is both Hispanic and Black/African American because only one response is possible. The response is the officer's opinion and not self-reported identification from the driver.

The stop rates show significant disparities from the population of Suffolk County, as reported by the United States Census, but caution should be used when making such comparisons. The demographic makeup of drivers on the road differs from the

[9] The category "Hispanic" is definitionally different than "Latino". Historically, the first term referred to individuals from or descending from Spanish-speaking countries (Spain, the Philippines, and Spanish-speaking Latin America) and the latter referring to persons from or descending from a Latin-American country regardless of use or connection to the Spanish language. Currently, the United States Census identifies "Hispanic or Latino Origin" as an "ethnicity" that may be included in a person's demographic information regardless of their "race." As noted above, these terms and categories all come with limitations. But these are the terms that are used in the data that was collected and analyzing these data requires adopting those terms.

demographic makeup of a county's entire population and may change over time based on the school year or other seasonal concerns. Not everyone counted by the Census has a drivers' license, and not everyone with a drivers' license drives. Still, as a baseline, it is useful to know that the population of Suffolk County is 4.4% Asian or Pacific Islander, 6.8% Black, 22% Hispanic or Latino, and 63% White. The significant disparities between the racial composition of the county population and the people stopped by the Suffolk County Police Department, while not alone indicative of racial profiling, should be analyzed.

Veil of Darkness

"Veil of Darkness" analysis is an attempt at a natural experiment to determine if certain racial/ethnic groups are disproportionately stopped by comparing the volume and proportion of stops during the "day" versus the "night". During the day, the racial/ethnic identity of the driver is more likely to be known, as it is visible to the officer. During the night (that is, under the veil of darkness) the identity is far more difficult to discern. The analysis considers evening stops that take place before sunset (day/light) and after sunset (night/dark). This allows the location to provide its own benchmark. Figure 2 below for the operationalization of this method.

The approach overcomes one issue with measuring demographic data in police stops against the demographic data of the population as a whole. General population demographic data do not capture the racial/ethnic makeup of who uses the road and when they use it.[10] The veil of darkness analysis is typically applied to testing disparities for Black/African American stops. Replication literature notes the ethnicity of Hispanic drivers is not always apparent, even during daylight hours.[11]

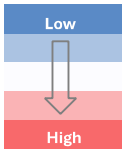
Figure 1, below, shows a heatmap of stops across 2024 by day of week and time of day of stop. It runs from bluest (lowest volume) to reddest (highest volume). The two high-volume periods are the morning and evening commutes, from 8:00 a.m. to noon and from 4:00 pm to 6:00 pm. On the weekends, the heaviest stop times are from 9:00 a.m. to noon, followed by a longer break from 1:00 p.m. until 4:00 p.m., and then a resurgence through 6:00 p.m.. The 11:00 p.m. hour also has noticeable increase for an hour every day of the week and on Friday for two hours until 1:00 a.m. Mornings from 2:00 a.m. to 6:00 a.m. have low numbers of stops every day regardless of weekend/work week division.

[10] Grogger, J. and Ridgeway, G. (2006) Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness, Journal of the American Statistical Association, 101:475, 878-887, DOI: 10.1198/016214506000000168

[11] Pierson, E., Simoiu, C., Overgoor, J. et al. A large-scale analysis of racial disparities in police stops across the United States. Nat Hum Behav 4, 736–745 (2020). https://doi.org/10.1038/s41562-020-0858-1

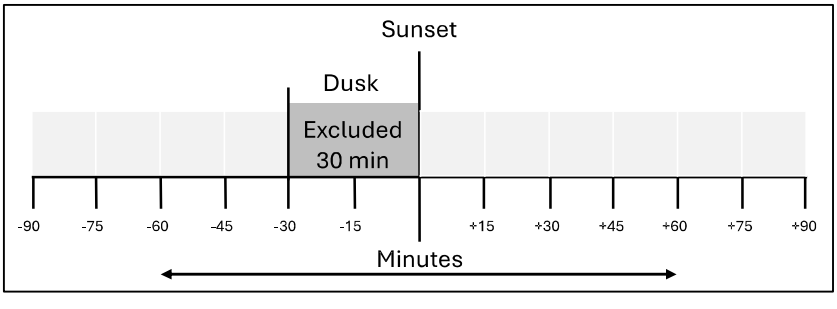
Figure 1: Distribution of Stop Volume by Time and Day of Week Over a Year

	Sun	Mon	Tues	Wed	Thur	Fri	Sat
0	741	662	747	808	777	876	955
1	574	409	430	454	469	550	676
2	328	199	223	244	229	303	352
3	190	127	132	132	144	161	222
4	106	99	116	111	101	131	103
5	52	79	100	102	99	104	90
6	39	166	193	219	200	223	107
7	327	942	1,014	990	1,025	959	453
8	689	1,648	1,701	1,783	1,604	1,614	942
9	1,074	1,860	1,956	2,056	1,853	1,884	1,306
10	1,283	1,734	1,979	1,955	1,894	1,956	1,473
11	1,150	1,517	1,923	1,925	1,647	1,777	1,392
12	825	1,141	1,471	1,357	1,356	1,285	1,025
13	674	1,008	1,351	1,209	1,159	1,123	896
14	312	514	833	761	799	606	460
15	686	1,490	1,641	1,695	1,696	1,300	848
16	1,190	2,133	2,440	2,278	2,472	2,082	1,328
17	1,013	1,782	2,015	1,974	2,082	1,935	1,215
18	750	1,187	1,393	1,278	1,410	1,287	831
19	625	885	1,039	916	958	1,019	631
20	476	712	944	867	814	957	676
21	490	567	740	738	730	761	563
22	632	712	884	853	904	859	640
23	855	1,003	1,096	1,033	1,100	1,090	837



Sunset times for 2024 were retrieved for Suffolk County from the website "Date and Time" for each date in the year.[12] The period coded "light" ending 30 minutes before sunset and beginning 60 minutes before that for the specific date of the stop. The period coded "dark" starts at sunset and continues for 90 minutes. The cohorts replicate Grogger and Ridgeway who looked at the 2.5 hour period of evening stops. The entire period reviewed is "evening" – 90 minutes before and after sunset, excluding the half hour dusk period because it is neither dark, nor light during this period when the sun is six degrees below the horizon. [13] The process involves comparing the racial proportions of drivers who were stopped during daylight with the demographics of those stopped after dark. Figure 2 illustrates the periods.

Figure 2: Veil of Darkness Analysis



[12] timeanddate.com. (n.d.). Sunrise and sunset times in Suffolk County. https://www.timeanddate.com/sun/@5140076

[13] Grogger, J. and Ridgeway, G. (2006) Testing for Racial Profiling in Traffic Stops from Behind a Veil of Darkness, Journal of the American Statistical Association, 101:475, 878-887, DOI: 10.1198/016214506000000168

Because the time the sun sets changes over the seasons, the same time of day is coded as "day" or "night" depending on the time of year. Therefore, the process proposes to correct for patterns of driving based on time of day. If raced-based discrepancies are present at 5:30 p.m. in July when officers can see a driver's race, but not at 5:30 p.m. in February, when officers cannot see a driver's race, that provides evidence that the racial disparities in stops are based on officers' perceptions of race.

Stops, racial breakdown

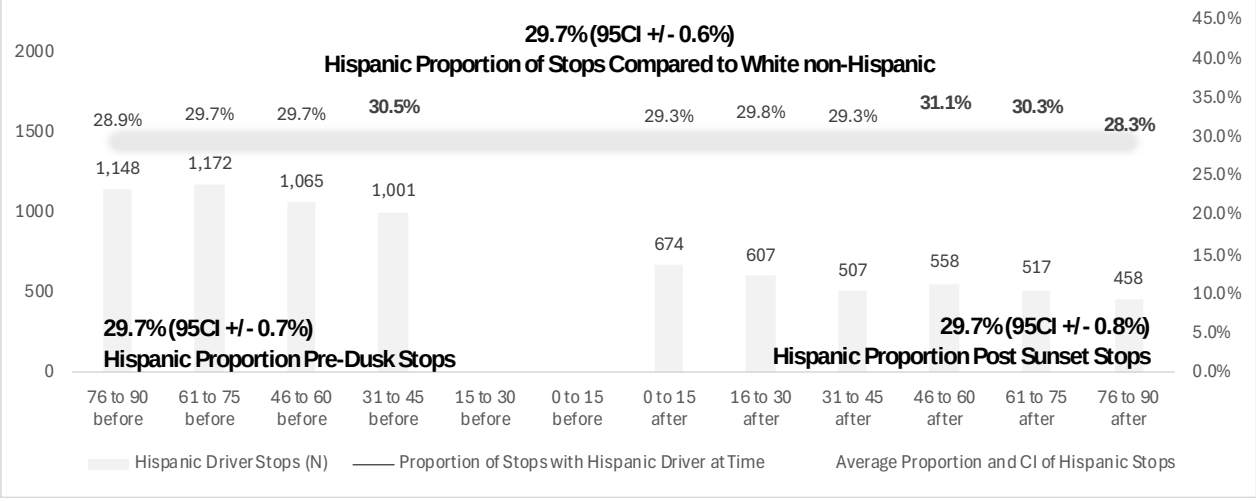
There were 14,790 stops that occurred in the period from 90 minutes before sunset to 30 minutes before sunset (Light) and 11,180 during the 90 minute period after sunset. This is a total of 25,970 stops or 16% of the total stops in the datasets. Hispanic drivers made up 28.5% of all drivers stopped and 29.7% of those stopped in the "evening" subset. In fact, Hispanic drivers made up exactly 29.7% of those stopped in both the "light" and "dark" portions of the "evening"—showing no proportional change at all. White non-Hispanic drivers made up 44.3% of all drivers stopped, and 44.0% of the drivers stopped during the "evening" subset. This rate did not change significantly between the "light" and "dark" periods.

Table 10: Veil of Darkness Stops by Race/Ethnicity

	All Stops		Evening Exclude 30m of Dusk		Evening Light		Evening Dark	
	N	%	N	%	N	%	N	%
Hispanic	45,083	28.5%	7,707	29.7%	4,386	29.7%	3,321	29.7%
Asian/PI	4,930	3.1%	825	3.2%	445	3.0%	380	3.4%
Black/AA	29,643	18.7%	4,573	17.6%	2,620	17.7%	1,953	17.5%
Other	8,430	5.3%	1,427	5.5%	819	5.5%	608	5.4%
White NH	70,056	44.3%	11,438	44.0%	6,520	44.1%	4,918	44.0%
Total NH	113,059	71.4%	18,263	70.3%	10,404	70.3%	7,859	70.3%
Total	**158,142**		**25,970**		**14,790**		**11,180**	

Figure 3 breaks down this data over shorter time increments. While 29.7% of those stopped during the entire "evening" period were Hispanic, this figure changed over time. Pre-dusk, the rate gradually increases and is outside the regular variation for the last fifteen minutes before dusk (at 30.5%). After dark, the discrepancy increases and then drops sharply. From 46 to 75 minutes after sunset, Hispanic drivers were overrepresented among those stopped, but from 76 to 90 minutes after sunset, they made up only 28.3% of the drivers stopped. A chi-square test (a test for association between categorical variables to assess if the distribution is between the two is the same) did not produce a statistically significant p-value for the differences between stop rates during the light and dark periods.

Figure 3: Number of Stops of Hispanic Drivers Before and After Dusk as a Proportion of Total Stops



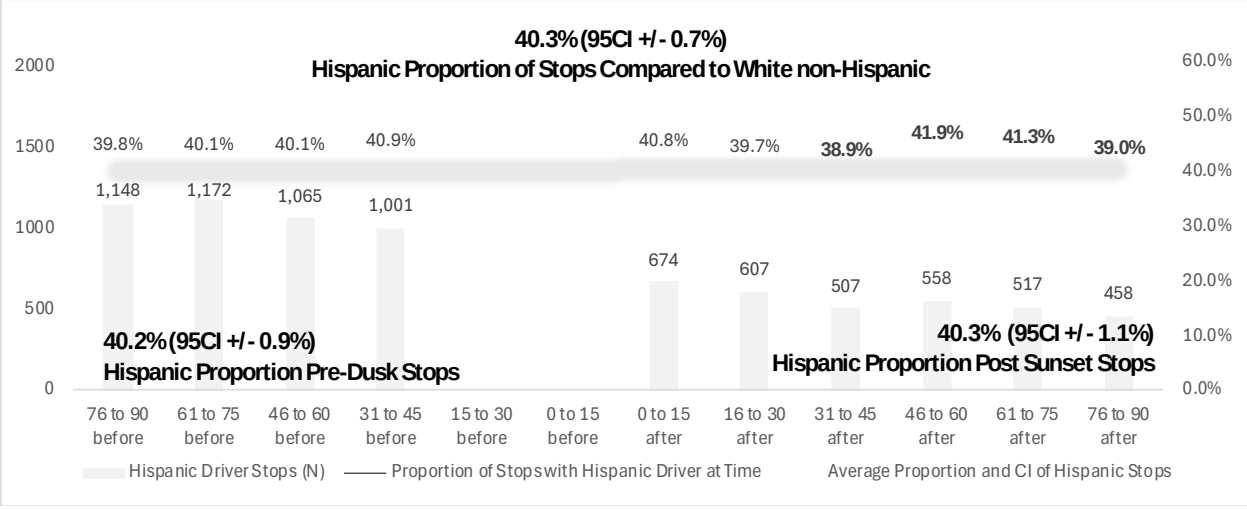
Conducting the analysis on the universe of only Hispanic and White non-Hispanic drivers (that is, excluding drivers categorized as Black, Asian, and Other from the analysis) yields similar results. The variation in proportion of stops in the evening do not differ by before/after sunset (dusk excluded). Hispanic drivers made up 40.3% of the evening stops in total and 40.2% of stops when it was light in the evening. White non-Hispanic drivers were 59.7% of the total evening stops and 59.8% of those in the light part of the evening (excluding dusk).

Table 11: Veil of Darkness Stops for Hispanic and White non-Hispanic Drivers

	All Stops		Evening Exclude 30m of Dusk		Evening Light		Evening Dark	
	N	%	N	%	N	%	N	%
Hispanic	45,083	39.2%	7,707	40.3%	4,386	40.2%	3,321	40.3%
White NH	70,056	60.8%	11,438	59.7%	6,520	59.8%	4,918	59.7%
Total	**115,139**		**19,145**		**10,906**		**8,239**	

The difference remains relatively constant comparing Hispanic and White non-Hispanic drivers when separated by 15 minute increment time bands. The 95% confidence interval was wider in the dark part of the evening for Hispanic drivers. During the period, natural variation would suggest that between 39.6% to 41.0% of the drivers stopped would be Hispanic. The were two periods with stops below the confidence interval and two periods with stops above the confidence interval. During the light part of the evening, all stops fell within the confidence interval. A chi-square test did not produce a statistically significant p-value for the differences between light/dark or time bands.

Figure 4: Number of Stops of Hispanic Drivers Before and After Dusk as a Proportion of Hispanic and White non-Hispanic Drivers Stopped



In both comparisons, Hispanic drivers as a percentage of all drivers stopped during the evening light remained in the confidence interval but had wider variation after dark. But since the variation included periods in which Hispanic drivers were both over- and under-represented, they do not indicate SCPD targeting people based on race.

One further open question is whether the driver populations are indeed constant at time of day. It is worth questioning whether the proportion of all drivers who are Hispanic may change seasonally. i.e., disproportionate participation in vocations predicated on presence of sunlight. In order to isolate "commutes," I conducted a veil of darkness analysis for the 6:00P hour in 15 minute increments. This too did not produce any additional information differences between Hispanic and non-Hispanic drivers.

Finding: The veil of darkness analysis on Hispanic drivers compared to all other drivers shows Hispanic drivers are stopped both more and less frequently after dark depending on the time band. When Hispanic drivers are compared to only White non-Hispanic drivers, this variation was more pronounced in the dark time increments as there were four increments outside the confidence interval, which was one additional from the comparison of Hispanic drivers to all other non-Hispanic drivers (including other minority groups).

The veil of darkness analysis here does not show disproportionality in stops. Still, the literature on veil of darkness analysis cautions to not use the analysis for Hispanic drivers because the ethnicity of Hispanic drivers is not always readily apparent.

Question 3: Is there evidence of racial or other disparities in post-stop outcomes, including car searches and arrests?

The majority of stops resulted in a summons or verbal warning regardless of the initial reason for a stop except for stops initiated because of "Probable Cause of a Crime." For these stops, arrest (45%) was more likely than summons and verbal warning combined (38%). When these two results were assessed, Hispanic drivers were more likely to receive a summons than non-Hispanic drivers and received more summonses when they were issued than non-Hispanic drivers did.

Hispanic drivers were also more likely to be asked to exit their vehicles during a stop (though this was an infrequent outcome) but less likely to be handcuffed once removed. Most significantly, SCPD officers searched Hispanic drivers more frequently than they searched non-Hispanic drivers. And when officers reported that the reason for the search was "Reasonable Suspicion," they were less likely to find any unlawful material than when they searched non-Hispanic drivers. This finding suggests that the searches of Hispanic drivers were conducted, on average, with a lower level of objectively credible evidence than searches of non-Hispanic drivers. Additionally, Hispanic drivers were arrested more frequently, though arrests were uncommon during stops. Because there is no information captured regarding the reason for an arrest, it is difficult to control for factors leading to arrest. It is also difficult to determine the timeline for search and arrest, *i.e.*, whether a search led to an arrest or the driver had a warrant and is then searched before being placed in the patrol car.

Stop Outcomes

The measurable post-stop outcomes include 1) stop disposition, 2) number of summonses issued if any, 3) exit vehicle (y/n), and (4) handcuffed (y/n). In addition, car search (y/n) and driver search (y/n), along with the "hit rates" (indication of productive search), were also assessed. These were all cross-referenced with the data categories used by the SCPD for drivers entered as Hispanic versus those entered as non-Hispanic (Asian, Black/African American, Other, and White non-Hispanic). In some instances, in order to account for racially disparate treatment of other groups, post-stop outcomes for Hispanic drivers were compared only to White non-Hispanic drivers. As previously noted, these categories and terms come with disadvantages; however, the terms are used in SCPD's data and analyzing these records requires adopting their use.

For example, of the Black/African American drivers who were stopped by the SCPD during the period, nearly 1.2% were arrested, which is double the rate of 0.6% of White non-Hispanic drivers who are arrested after a stop. But combining these figures when comparing the groups to Hispanic Drivers will suggest 0.9% of non-Hispanic drivers are arrested after being stopped and hides the discrepancy between Hispanic and non-Hispanic White drivers. Therefore, throughout this section, rates for Hispanic drivers are compared both to all non-Hispanic drivers and to White non-Hispanic drivers.

Statistical Models

There are two statistical models used throughout this section to test if the mean outcome for Hispanic drivers is different from the mean outcome for non-Hispanic and White non-Hispanic drivers. The first statistical model is a "t-test" and the second is a "binary logistic regression."

A t-test compares means between two race/ethnicity groups by testing the null hypothesis "these population means are equal" and the alternative hypothesis "these population means are not equal." The test generates a "p-value", which is a number between 0 and 1 used to measure the likelihood that a given result is the product of chance. Here, the p-value measures likelihood of seeing a difference in the measured averages if the population means for Hispanic and non-Hispanic (or White non-Hispanic) were equal. So, if the p-value is greater than 0.05, the data supports the null hypotheses and suggests no difference in outcomes for Hispanic drivers compared to non-Hispanic drivers. If the p-value is less than or equal to 0.05, it represents the likelihood that the difference in average outcomes for Hispanic and non-Hispanic drivers would occur if the SCPD treated the two populations identically. The likelihood that the discrepancy is random noise is less than 5% when $p<.05$, less than 1% when $p<.01$, or less than 0.1% when $p<.001$.

A "binary logistic regression" models the relationship between a binary dependent variable, (a variable with only two possible outcomes, such as (yes/no) and one or more independent variables. The null hypothesis is that "there is no relationship between the dependent and independent variables" and the alternative hypothesis is "one or more of the independent variables has a significant impact on the dependent variable." A p-value is generated to indicate statistical significance for each independent variable. When the p-value is less than 0.05, the null hypothesis is rejected. In these instances, as above, the p-value represents the probability that the outcome was not related to the independent variable, and therefore the likelihood that the independent variable did not impact the outcome.

The output includes a β value for each independent variable. A positive number means the variable moves in the same direction as the independent variable. For example, if the dependent variable is coded as 0 for "no arrest" and 1 for "arrest" and the independent variable is coded as 0 for "non-Hispanic" and 1 for "Hispanic" then a positive β value means that "Hispanic" has a higher likelihood of arrest. The variables move in the same direction.

The Odd's Ratio shows the impact of that relationship. When the β value is positive, the Odd's Ratio *minus* 1 represents the percentage increase in odds for the outcome (in this example, arrest for Hispanic drivers) if the p-value is less than 0.05. When the β value is negative, then there is an inverse relationship between the independent and dependent variables. In our example, that would be the case if Hispanic drivers were arrested at a lower rate than non-Hispanic drivers. The impact would then be 1 *minus* the Odd's Ratio, which indicates the percentage decrease in the likelihood of arrest if the p-value is less than 0.05. This description is admittedly challenging for persons unfamiliar with this form

of statistical analysis; each binary logistic regression will include a table with the β value, p-value, and Odd's Ratio and the text will explain when there is statistical significance and what that means in terms of more/less likely as an outcome for Hispanic drivers.

The logistic regression uses three additional control variables: (i) Highway Patrol versus precinct officers, (ii) out-of-state plates, and (iii) drivers aged 25 or younger.

The Stonewall Analytics report on 2023 traffic stops found a difference in results when the stop was performed by Highway Patrol versus precinct officers; that analysis was repeated here.[14] The out-of-state plates control is a proxy for state residents. Historical research has been mixed on residency: some studies suggest that out-of-state drivers are more likely to receive a summons because the drivers will not return to the state to challenge the ticket, and others show that out-of-state drivers are less likely to get a summons because local public safety is more sensitive to enforcement.[15] Young drivers have been shown to be more likely to receive summonses[16] and be searched.[17]

Stop Disposition

The interaction between race/ethnicity (Hispanic vs non-Hispanic), stop reason, and outcome showed a few early patterns. The reason for stops had variation by race/ethnicity. The majority (about 98%) of stops were disposed with either a summons or a verbal warning, and the disposition of stops varied by both race/ethnicity and type of stop showing an interesting aberration for the stop reasons *Probable Cause of a Crime* and *Reasonable Suspicion of a Crime*.

Drivers coded Hispanic were disproportionately stopped for Equipment Violation, Red Light, and Registration/Inspection stops than those not coded Hispanic. Hispanic drivers had the highest rate of Equipment Violation stops compared to all other race/ethnicity categories at 18% compared to 11% non-Hispanic drivers. Black/African American drivers were the second highest at 16%, and White non-Hispanic drivers were the lowest at 9%. Probable Cause of a Crime, Reasonable Suspicion of a Crime, Seat Belt, and Stop Sign made up comparable proportions of total stops for Hispanic, non-Hispanic, and White non-Hispanic drivers. While 15% of the non-Hispanic drivers who were stopped were

---

[14] Stonewall Analytics. February 2025. 2023 Traffic Stop Data Analysis: Suffolk County Police Department.

[15] Roach, M. 2015. Is the Highway Patrol Really Tougher on Out-of-State Drivers? An Empirical Analysis. BE J. Econ. Anal. Policy: 15(2): 769–796

[16] Chenhui Liu and Anuj Sharma. 2019. Are you going to get a ticket or a warning for speeding? An autologistic regression analysis in Burlington, VT, Transportation Research Interdisciplinary Perspectives, Volume 1, 2019, https://doi.org/10.1016/j.trip.2019.100001.

[17] Rosenfeld, R., Rojek, J., & Decker, S. (2012). Age matters: Race differences in police searches of young and older male drivers. Journal of Research in Crime and Delinquency, 49(1), 31–55.

stopped for speeding, only 11% of the Hispanic drivers who were stopped were stopped for speeding.

Table 12: Stop Reason by Race/Ethnicity

	Hispanic		Non- Hispanic		Asian / PI		Black / AA		Other		White Non-His		Total	
Cell Call/ Text	2,204	5%	6,793	6%	232	5%	1,115	4%	531	6%	4,915	7%	8,997	6%
Equip Vio	7,903	18%	12,464	11%	536	11%	4,689	16%	889	11%	6,350	9%	20,367	13%
Moving Vio	6,829	15%	20,641	18%	944	19%	4,210	14%	1,696	20%	13,791	20%	27,470	17%
Other Moving Vio	5,457	12%	14,808	13%	729	15%	3,582	12%	1,678	20%	8,819	13%	20,265	13%
Prob Cause Crime	78	0.2%	171	0.2%	3	0%	68	0%	5	0%	95	.1%	249	.2%
Reas Susp of a Crime	141	0.3%	361	0.3%	9	0%	141	0%	10	0%	201	.3%	502	.3%
Red Light	805	1.8%	1,805	1.6%	95	2%	407	1%	110	1%	1,193	2%	2,610	2%
Reg / Insp / Plt Vio	10,872	24%	25,032	22%	827	17%	8,743	29%	1,203	14%	14,259	20%	35,904	23%
Seat Belt	930	2%	1,987	2%	42	1%	448	2%	126	1%	1,371	2%	2,917	2%
Speed	5,120	11%	16,568	15%	684	14%	3,658	12%	1,572	19%	10,654	15%	21,688	14%
Stop Sign	4,744	11%	12,429	11%	829	17%	2,582	9%	610	7%	8,408	12%	17,173	11%
Total	45,083		113,059		4,930		29,643		8,430		70,056		158,142	

As these stop types were distributed differently across race/ethnicity, dispositions cannot be explored by race/ethnicity alone. Hence, stop type was also assessed. Stop types had seven possible outcomes: Aided, Arrested, First Appearance Ticket, Lights-On Voucher, No Police Action Taken, Summons Issued, and Verbal Warning (Education). There were four “aided” and seven “NULL” excluded from the table below.

Half of all stops ended with a Verbal Warning/Education (50%) but that varied based on the reason for the stop. Only 18% of stops ended in a verbal warning where the officer observed Probable Cause of a Crime, while 64% of stops ended in verbal warning due to Equipment Violation. Just under half of all stops ended with a summons (48%). This too varied based on the reason for the stop; only 20% of all stops made after an officer observed Probable Cause or Reasonable Suspicion of a Crime resulted in a summons while two-thirds of all stops initiated for a Seat Belt violation did. Only 1% of stops resulted in an arrest, though unsurprisingly a high of 45% of Probable Cause of a Crime stops resulted in an arrest.

Table 13: Driver Disposition by Stop Reason

	Arrest		Field Appearance		Lights-On Voucher		No Police Action		Summons		Verbal Warning / Education		Total
Cell Call/Text	10	0%	0	0%	5	0%	73	1%	5,068	56%	3,841	43%	8,997
Equipment Violation	174	1%	43	0%	270	1%	381	2%	6,436	32%	13,061	64%	20,367
Moving Violation	153	1%	16	0%	2	0%	259	1%	14,204	52%	12,831	47%	27,470
Other Moving Violation	254	1%	35	0%	5	0%	409	2%	9702	48%	9,860	49%	20,265
Prob Cause of a Crime	113	45%	10	4%	0	0%	31	12%	49	20%	46	18%	249
Reasonable Susp of a Crime	57	11%	21	4%	0	0%	119	24%	102	20%	203	40%	502
Red Light	29	1%	4	0%	3	0%	5	0%	1,166	45%	1,403	54%	2,610
Reg / Insp / Plate Vio	230	1%	133	0%	4	0%	646	2%	16,583	46%	18,304	51%	35,904
Seat Belt	16	1%	1	0%	2	0%	7	0%	1,953	67%	938	32%	2,917
Speeding	103	0%	5	0%	4	0%	452	2%	12,802	59%	8,322	38%	21,688
Stop Sign	84	0%	10	0%	5	0%	69	0%	7,278	42%	9,727	57%	17,173
Total	**1,223**	**1%**	**278**	**0%**	**300**	**0%**	**2,451**	**2%**	**75,343**	**48%**	**78,536**	**50%**	**158,142**

Summons and verbal warning/education as a subset was used to create a dichotomous variable (a variable with only two possible values – in this case summons or verbal warning). There are limited observations in the other disposition categories, and they get even smaller when separating by race/ethnicity. When considering only stops that ended in a summons or a warning, Hispanic drivers were issued a summons 55% of the time they were stopped and non-Hispanic drivers were issued a summons only 46% of the time they were stopped. The summons rate was higher for every stop type except Probable Cause of a Crime (which had a very small sample size of 31 Hispanic and 66 non-Hispanic drivers).

Table 14: Number and Proportion of Hispanic and Non-Hispanic Drivers Given Summons or Verbal Warning by Stop Reason

	Hispanic				Non-Hispanic			
	Summons		Verbal Warning / Education		Summons		Verbal Warning / Education	
	N	%	N	%	N	%	N	%
Cell Call/Text	1,478	68%	705	32%	3,590	53%	3,136	47%
Equipment Violation	2,901	38%	4,731	62%	3,535	30%	8,330	70%
Moving Violation	4,175	62%	2,550	38%	10,029	49%	10,281	51%
Other Moving Violation	3,066	58%	2,209	42%	6,636	46%	7,651	54%
Prob Cause of a Crime	13	42%	18	58%	36	56%	28	44%
Reasonable Susp of a Crime	28	33%	57	67%	74	34%	146	66%
Red Light	466	59%	321	41%	700	39%	1,082	61%
Reg / Insp / Plate Vio	5,514	52%	5,053	48%	11,069	46%	13,251	54%

	Hispanic				Non-Hispanic			
	Summons		Verbal Warning / Education		Summons		Verbal Warning / Education	
Seat Belt	684	74%	237	26%	1,269	64%	701	36%
Speeding	3,552	71%	1,474	29%	9,250	57%	6,848	43%
Stop Sign	2,437	52%	2,264	48%	4,841	39%	7,463	61%
Total	**24,314**	**55%**	**19,619**	**45%**	**51,029**	**46%**	**58,917**	**54%**

Finding: Proportionately more Hispanic drivers received a summons as opposed to a warning compared to non-Hispanic drivers overall whenever the initial stop was made based on a suspected moving violation. The only times that Hispanic drivers were more likely than non-Hispanic drivers to receive a warning was when the officer reported that the initial reason for the stop was probable cause or reasonable suspicion that the driver had committed a criminal offense.

A binary logistic regression was run for all stop types for summons versus verbal warning for Hispanic versus non-Hispanic drivers. Hispanic drivers were issued proportionally more summonses compared to warnings. Because the dataset does not include natural control measures, one was constructed for Highway Patrol versus precinct officers, another for out-of-state plates, and the age cohort was narrowed to drivers aged 25 or younger. The dependent variable is summons (1) compared to warning (0). Probable Cause and Reasonable Suspicion stops (which are more likely to result in an arrest) are excluded in this discussion for three reasons: first, there was no statistically significant outcome; second, over 98% of the stops outside these categories resulted in a summons or warning; and third, these two stop types are addressed in depth later during a discussion of arrests.

The binary Hispanic variable was related to receiving a summons versus a warning for all stop types. It was at the $p<.001$ level for every stop type except Seat Belt, which was at the $p<.05$ level. Being a Hispanic driver increased the odds of getting a summons in lieu of a warning by 80% for Cell/Text stops, 44% for Equipment Failures, 70% on Moving Violations, 63% on Other Moving Violations, over two-fold on Red Light stops, 32% on Registration stops, 59% on Seat Belt stops, 72% on Speeding, and 61% on Stop Light stops.

Table 15: Binary Logistic Regression by Stop Reason on Likelihood of Receiving a Summons or a Verbal Warning

	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Cell Call/Text					
Hispanic (0 Non, 1 Hispanic)	.588	.052	125.992	<.001***	1.800
NY Plates (0 No, 1 Yes)	-.456	.087	27.548	<.001***	.634
25 or Younger (0 No 26 and older, 1 Yes)	.334	.063	28.144	<.001***	1.397
Patrol (0 Highway, 1 Precinct)	-.419	.046	81.976	<.001***	.657
Constant	.791	.089	79.143	<.001	2.206
Equipment Failure					

	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic (0 Non, 1 Hispanic)	.366	.031	137.384	<.001***	1.442
NY Plates (0 No, 1 Yes)	.079	.049	2.650	.104	1.083
25 or Younger (0 No 26 and older, 1 Yes)	.092	.035	6.992	.008	1.096
Patrol (0 Highway, 1 Precinct)	-.426	.049	74.234	<.001***	.653
Constant	-.569	.064	80.022	<.001	.566
Moving Violation					
Hispanic (0 Non, 1 Hispanic)	.528	.029	327.448	<.001***	1.695
NY Plates (0 No, 1 Yes)	-.313	.050	39.206	<.001***	.731
25 or Younger (0 No 26 and older, 1 Yes)	.360	.030	141.280	<.001***	1.433
Patrol (0 Highway, 1 Precinct)	-.513	.027	357.092	<.001***	.599
Constant	.543	.052	111.045	<.001	1.722
Other Moving Violation					
Hispanic (0 Non, 1 Hispanic)	.488	.033	214.737	<.001***	1.630
NY Plates (0 No, 1 Yes)	-.397	.049	66.481	<.001***	.672
25 or Younger (0 No 26 and older, 1 Yes)	.387	.039	97.707	<.001***	1.472
Patrol (0 Highway, 1 Precinct)	-.589	.029	401.126	<.001***	.555
Constant	.456	.049	85.679	<.001	1.578
Red Light					
Hispanic (0 Non, 1 Hispanic)	.785	.088	80.230	<.001***	2.193
NY Plates (0 No, 1 Yes)	-.366	.162	5.126	.822	.693
25 or Younger (0 No 26 and older, 1 Yes)	.246	.103	5.699	.829	1.279
Patrol (0 Highway, 1 Precinct)	-.263	.266	.976	.303	.769
Constant	.124	.305	.165	.412	1.131
Registration/Inspection					
Hispanic (0 Non, 1 Hispanic)	.277	.024	138.679	<.001***	1.319
NY Plates (0 No, 1 Yes)	.294	.041	52.607	<.001***	1.342
25 or Younger (0 No 26 and older, 1 Yes)	.021	.028	.525	.469	1.021
Patrol (0 Highway, 1 Precinct)	-.443	.049	82.552	<.001***	.642
Constant	-.038	.061	.398	.528	.962
Seat Belt					
Hispanic (0 Non, 1 Hispanic)	.467	.089	27.450	<.001***	1.594
NY Plates (0 No, 1 Yes)	-.114	.151	.572	.450	.892
25 or Younger (0 No 26 and older, 1 Yes)	.122	.107	1.295	.255	1.130
Patrol (0 Highway, 1 Precinct)	-.210	.105	3.989	.046*	.811
Constant	.849	.173	24.088	<.001	2.338

	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Speeding					
Hispanic (0 Non, 1 Hispanic)	.545	.035	237.000	<.001***	1.724
NY Plates (0 No, 1 Yes)	-.341	.060	32.167	<.001***	.711
25 or Younger (0 No 26 and older, 1 Yes)	.546	.033	268.144	<.001***	1.726
Patrol (0 Highway, 1 Precinct)	-.362	.029	158.400	<.001***	.696
Constant	.667	.060	122.116	<.001	1.949
Stop Light					
Hispanic (0 Non, 1 Hispanic)	.477	.035	188.138	<.001***	1.611
NY Plates (0 No, 1 Yes)	-.274	.069	15.810	<.001***	.760
25 or Younger (0 No 26 and older, 1 Yes)	.287	.038	58.067	<.001***	1.332
Patrol (0 Highway, 1 Precinct)	.048	.298	.026	.872	1.049
Constant	-.277	.304	.826	.363	.758

* p<.05; ** p<.01; *** p<.001

Finding: In 2024, Hispanic drivers received summonses versus warnings at consistently higher rates than non-Hispanic drivers, regardless of the original reason for the stop. The increased rate of summonses for Hispanic drivers compared to non-Hispanic drivers is less than 1% likely to have occurred due to natural variation.

On average, SCPD officers issued 2.2 summonses for each stop when they issued any summonses. For Hispanic drivers, the average number was 2.4; non-Hispanic drivers received an average of 2.1. Therefore, Hispanic drivers were not only more likely to receive a summons but also received 12% more summonses at a statistically significant rate when they received them.

This pattern applied to every stop type except for Probable Cause of a Crime and Reasonable Suspicion of a Crime, each of which result in summonses or warnings less frequently.

Table 16: Comparison of Mean Number of Summonses Issued for Hispanic/Non-Hispanic Drivers (t-test)

	Hispanic		Non-Hispanic		*t*			
	M	SD	M	SD	Total Issued	Score	*p*	Cohen's d
Cell Call/Text	1.63	1.16	1.42	1.03	5,067	-6.07	p<.001***	1.067
Equipment Violation	2.99	2.06	2.81	2.12	6,436	-3.37	p<.001***	2.096
Moving Violation	1.83	1.48	1.61	1.33	14,201	-8.55	p<.001***	1.374
Other Moving Violation	2.28	1.76	1.96	1.67	9,698	-8.73	p<.001***	1.699
Prob Cause of a Crime	1.46	.66	2.33	1.97	49	2.32	p<.05*	1.733

	Hispanic		Non-Hispanic		*t*			
	M	**SD**	**M**	**SD**	**Total Issued**	**Score**	***p***	**Cohen's d**
Reasonable Suspicion of a Crime	2.86	1.82	2.88	1.93	102	0.05	-	1.901
Red Light	2.1	1.54	1.9	1.46	1,166	-2.35	p<.05*	1.495
Reg / Insp / Plate Vio	3.6	2.23	3.4	2.35	16,579	-4.04	p<.001***	2.313
Seat Belt	2.1	1.68	1.8	1.73	1,953	-3.17	p<.01**	1.714
Speeding	1.8	1.49	1.6	1.33	12,800	-7.88	p<.001***	1.378
Stop Sign	1.9	1.43	1.6	1.29	7,276	-7.49	p<.001***	1.338
Total	**2.4**	**1.92**	**2.1**	**1.86**	**75,327**	**-20.213**	**p<.001*** **	**1.878**
								Two sided t-test
								* p<.05; ** p<.01; *** p<.001

Finding: On average, Hispanic drivers received more summons overall and by stop type.

Exit Vehicle

There were 6,132 drivers who exited the vehicle for the following reasons: Force (18 or 0.3%), Spontaneous (2,017 or 32.9%), and Verbal Request (4,097 or 66.8%). Force represents an extremely small sample size and spontaneous exits are not necessarily the result of officer action. A request to exit the vehicle, however, occurs based on the officer's discretion due to circumstances of the driver and stop. Requested exits, per Dr. Smith's original data lists, were isolated to explore disproportionality.[18] Hispanic drivers were requested to exit the vehicle in 3.4% of stops compared to 2.3% of non-Hispanic drivers.

Table 17: Number and Proportion of Stops with Requests to Exit the Vehicle for Hispanic and Non-Hispanic Drivers

	Verbal Request to Exit	Total Stops	Percent of Stops with Request to Exit
Hispanic	1,522	29,643	3.4%
Non-Hispanic	2,575	113,059	2.3%
Total	**4,097**	**158,142**	**2.6%**

Hispanic drivers were asked to exit the vehicle at a rate 1.4 times higher than non-Hispanic drivers at a p<.001 level of statistical significance. That is, the likelihood that this discrepancy is due to natural variation is less than one-tenth of one percent. The other control variables, less age, were significant as well. Drivers with NY plates were less likely to be asked to exit the vehicle. Those stopped by a precinct patrol officer were more likely to be asked to exit the vehicle than those pulled over by a Highway Patrol officer. These

[18] U.S. District Court for the Eastern District of New York. (2020, July 15). Case 2:15-cv-02431-WFK-LB, Document 252-20. https://www.courtlistener.com/docket/4324349/1-21-v-the-county-of-suffolk/

results may have to do with stop reason or deployment (it is far less safe to exit a vehicle on the highway).

Table 18: Binary Logistic Regression of Requests to Exit the Vehicle for Hispanic and Non-Hispanic Drivers

Requested to Exit Vehicle	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic (0 Non, 1 Hispanic)	.368	.033	123.705	<.001***	1.444
NY Plates (0 No, 1 Yes)	-.707	.045	250.305	<.001***	.493
25 or Younger (0 No 26 and older, 1 Yes)	.061	.038	2.597	.107	1.063
Patrol (0 Highway, 1 Precinct)	.152	.039	14.988	<.001***	1.164
Constant	-3.286	.082	1603.528	<.001	.037

* p<.05; ** p<.01; *** p<.001

Finding: Hispanic drivers were more likely to be asked to exit the vehicle than non-Hispanic drivers.

Handcuffed Post Vehicle Exit

An officer decides whether to handcuff the driver after a request to exit the vehicle. There was limited difference in the overall handcuff rate by vehicle exit reason between Hispanic (19.5%) and non-Hispanic (20.7%) drivers. Hispanic drivers compared to non-Hispanic drivers with a vehicle exit were handcuffed less frequently after they were asked to exit the vehicle. (Note that only eighteen drivers were removed from their cars by force, making the sample size too small to be valuable).

Table 19: Hispanic and Non-Hispanic Drivers Handcuffed by Vehicle Exit Reason

Why Vehicle Exit	Hispanic Driver			Non-Hispanic Drivers		
	N HCuffed	N Exits	% Cuffed	N HCuffed	N Exits	% Cuffed
Force	5	7	71.4%	9	11	81.8%
Spontaneous	26	602	4.3%	49	1,415	3.5%
Request	385	1,522	25.3%	771	2,575	29.9%
Total	**416**	**2,131**	**19.5%**	**829**	**4,001**	**20.7%**

A binary logistic regression was done on handcuffed drivers (yes or no by exit reason). Hispanic drivers are disproportionately asked to exit the vehicle, see Table 18, but their odds of being handcuffed were 21% lower than the comparison group (p<.01). This finding is consistent with the hypothesis that Hispanic drivers are disproportionately asked to exit their vehicles compared to non-Hispanic drivers. Once asked to exit, they are handcuffed less frequently, suggesting, on average, more compliant behavior after exiting the vehicle.

Table 20: Binary Logistic Regression on Drivers Handcuffed by Vehicle Exit Reason

Vehicle Exit Reason		β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Force	Hispanic (0 Non, 1 Hispanic)	-.588	1.145	.264	.608	.556
	Constant	1.504	.782	3.702	.054	4.5
Spontaneous	Hispanic (0 Non, 1 Hispanic)	.231	.248	.869	.351	1.260
	Constant	-3.327	.145	523.610	<.001	.036
Request	Hispanic (0 Non, 1 Hispanic)	-.233	.073	10.174	.001**	.792
	Constant	-.850	.043	390.325	.001	.427

* p<.05; ** p<.01; *** p<.001

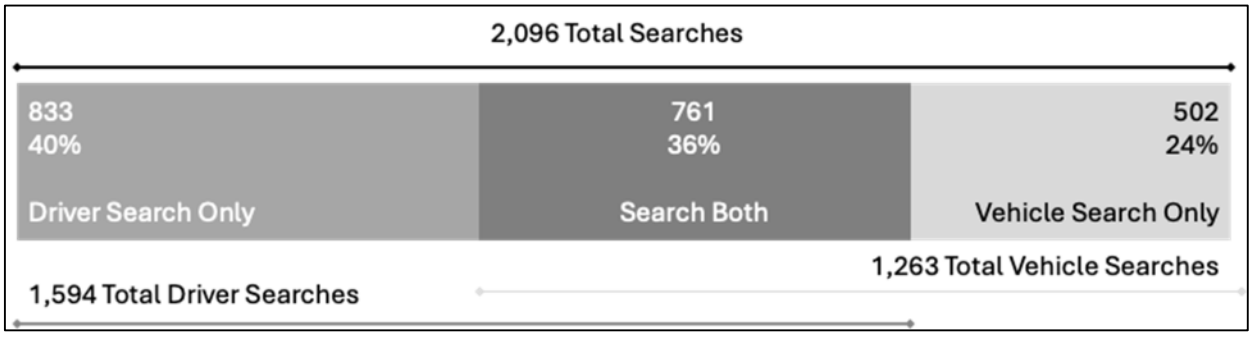
Finding: Though more Hispanic drivers are asked to exit the vehicle, they are handcuffed at a lower rate.

Search Rate

The two variables (driver search and vehicle search) can be combined into one flag as discussed in Table 9. This is what SCPD did for the dashboard and makes sense as a reflection of what happens on a stop. There were 1,263 vehicle searches, 1,594 driver searches, and a combined 2,096 searches (in 761 stops, officers searched both the driver and the vehicle). As a result of the overlap, analyzing driver stops and vehicle stops separately may over- or understate a disproportionality.

This becomes even more important when discussing search results. If an officer finds contraband when searching a driver, they are more likely to search the vehicle and vice-versa. Considering the search types independently can underreport the "hit" rate. For example, if an officer searches a vehicle, finds contraband, and arrests the driver, the officer will search the driver as part of the arrest process. As discussed below, if the driver and vehicle search are combined, the stop looks like two unconnected searches with a 50 percent hit when it is really one search event.

Figure 5: Search of Vehicle, Driver, or Both Vehicle and Driver



The likelihood that SCPD would conduct a search of either the driver or the vehicle after stopping Hispanic drivers was 2% compared to 1% of non-Hispanic drivers. A preliminary t-test indicates that this difference is statistically significant at the p<.001 level for Hispanic drivers. This test rejects the null hypothesis that the search rate for Hispanic and non-Hispanic drivers is the same.

Table 21: t-test on Average Proportion of Searches (Vehicle and/or Driver)

	Hispanic 45,083		**Non-Hispanic 113,059**		***t***		
	M	SD	M	SD	Score	*p*	Cohen's d
Vehicle, Driver, or Vehicle & Driver Search	.02	.122	.01	.111	-3.997	<.001***	.114
							Two sided t-test * p<.05; ** p<.01; *** p<.001

A binary logistic regression confirms this finding. When the Hispanic driver binary flag is the only independent variable, the binary logistic regression finds Hispanic drivers are 1.2 times more likely to have their vehicle and/or person searched at a stop. This difference is likewise statistically significant at p<.001. When the additional control variables are included, the Hispanic binary flag is significant at the p<.05 level and shows Hispanics were 1.1 times more likely to be searched than their non-Hispanic counterparts. Stops by precinct officers increased search rate by almost 7.5 times that of stops by Highway Patrol and was significant at the p<.001 level.

Table 22: Binary Logistic Regression on Search (Vehicle and/or Driver) Occurrence

Search = 1	Variable	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic Only	Hispanic (0 Non, 1 Hispanic)	.195	.047	17.279	<.001***	1.215
	Constant					
Hispanic & Controls	Hispanic (0 Non, 1 Hispanic)	.101	.047	4.562	.033*	1.106
	NY Plates (0 No, 1 Yes)	-.721	.063	130.725	<.001***	.486
	25 or Younger (0 No 26 and older, 1 Yes)	.130	.052	6.161	.013*	1.138
	Patrol (0 Highway, 1 Precinct)	2.01	.112	324.138	<.001***	7.460
	Constant					

* p<.05; ** p<.01; *** p<.001

Finding: Hispanic drivers are more likely to be searched using the combined search flag (driver, vehicle, or driver and vehicle).

Search Reason

Of the 1,594 driver searches, 148 or 9% were recorded as based on founded suspicion with recorded consent, 658 or 41% were identified as based on reasonable suspicion, and 788 were stated as based on probable cause. Of the 1,263 vehicle searches, 412 or 32.6% were identified as based on founded suspicion with recorded consent, 374 or 29.6% were based on plain view, and 477 or 37.8% were based on probable cause. The entries for the 761 stops in which the SCPD searched both the driver and the vehicle recorded independent reasons for each search, e.g., consent on one and probable cause on the other. It is unclear if the driver or vehicle search occurred first, so to create a combined search reason, the lowest standard from either was used. This means, if either vehicle or driver search is justified as *founded suspicion*, the reason is founded suspicion; if either is *reasonable suspicion* or *plain view* and neither was founded, it is coded reasonable suspicion; and finally, if both are coded *probable cause* or one is and the other is NULL, then the reason was coded probable cause.

This calculation results in the following breakdown for reasons for the 2,096 combined searches: 491 or 23% founded suspicion, 833 or 40% reasonable suspicion,[19] and 772 or 37% probable cause. The distribution of search types for Hispanic and non-Hispanic drivers was comparable; however, the measure of disproportionality is not defined because there is not a record of how many drivers were asked to consent to a search.

[19] A plain view vehicle search was treated as the intermediate reasonable suspicion for the created combination variable.

Table 23: Number and Proportion of Combined Search Reasons for Driver/Vehicle Searches for Hispanic and Non-Hispanic Drivers

	Founded Suspicion of a Crime with Recorded Consent		Reasonable Suspicion/Frisk for Weapon		Probable Cause of a Crime / Automobile Exception		Total
	N	%	N	%	N	%	
Hispanic	148	21.7%	270	39.5%	265	38.8%	**683**
Non-Hispanic	343	24%	563	40%	507	36%	**1,413**
Total	**491**	**23%**	**833**	**40%**	**722**	**37%**	**2,096**

Finding: There is no variable recording who was asked to consent for a search, only the implication of those who said yes (recorded consent).

Recommendation: Record who was asked to consent to a search.

Finding: Distribution of search types using a combined search variable (driver and vehicle) show Hispanic and non-Hispanic search rates are comparable.

Hit Rate

The "hit rate" represents the rate at which officers recover what they are looking for (generally evidence of a crime) when they search a person or vehicle. If a certain group of people being searched has a uniformly lower hit rate, that suggests that officers are searching that group with a lower level of objective evidence than a group for which there is a higher hit rate.

Both types of searches can be combined into a "hit" binary that captures the full experience of the search without creating false negatives when reviewed separately. For example, if the SCPD stops a driver, searches him, and finds contraband in his pocket, then "Driver Search" will be a hit for contraband. If officers then search the vehicle and find no contraband, that will register as a non-hit even though the overall encounter uncovered contraband. There are 11 such results (driver search = drugs and vehicle search = negative) in Table 24, which could look like an unnecessary search without reviewing the context of the entire encounter.

Table 24: Explanation of Creating a Combined Hit Rates for Driver/Vehicle Searches

		Driver Search						
		Drugs	Weapon	Negative	Non-Criminal Object	Not Searched (NULL)	Other Contraband / Evidence	**Total Driver**
Vehicle Search	Drugs	92	0	117	12	32	7	**260**
	Weapons	2	6	13	0	5	1	**27**
	Negative	11	2	338	37	388	7	**783**
	Not Searched (NULL)	31	3	670	112	0	15	**831**
	Other Contraband/ Evidence	4	0	74	23	77	17	**195**
	Total Vehicle	**140**	**11**	**1,212**	**184**	**502**	**47**	**2,096**

The codes are as follows:

- Those who were only subject to a driver search and had a hit are considered *yes: hit*.
- Those who were only subject to a vehicle search and had a hit are also considered *yes: hit*.
- Drivers subject to both need only a hit on one to be coded *yes: hit*.
- Those who were subject to either or both with nothing found on any search are coded *no,* they are the darker great box in the middle of Table 24.

The combined 2,096 driver and vehicle searches above are analyzed together for Hispanic versus non-Hispanic drivers in Table 25. The "hit rate" flag identifies when a search recovers contraband. Of the 2,096 searches total, 1,545 (74%) are coded as "not productive," which means the search did not recover any contraband, regardless of whether the driver, vehicle, or both were searched. Hispanic drivers have a slightly lower hit rate of 24% compared to 27% for non-Hispanic drivers searched and a hit rate seven percentage points lower than non-Hispanic White drivers. Note, non-Hispanic White drivers are a subset of non-Hispanic drivers. These drivers had a 31% hit rate.

Table 25: Combined Hit Rate for Driver and/or Vehicle Searches for Hispanic, Non-Hispanic Drivers, Non-Hispanic White Drivers

	Hispanic		**Non-Hispanic**		**Non-Hispanic White**		**Total**	
	N	%	N	%	N	%	N	%
Not Productive	520	76%	1,025	73%	457	69%	1,545	74%
Hit	163	24%	388	27%	207	31%	551	26%
Total	**683**		**1,413**		**664**		**2,096**	

The hit rates for searches conducted based on Founded Suspicion with Recorded Consent were effectively equivalent between Hispanic drivers (18%) and non-Hispanic drivers (17%), but nine percentage points lower than non-Hispanic White drivers (27%). The hit

rate for Reasonable Suspicion searches is six percentage points lower for Hispanic drivers searched than for non-Hispanic drivers and 10 percentage points lower than non-Hispanic White drivers. Reasonable suspicion had the highest hit rate for all cohorts. Probable Cause searches retain the three percentage point difference for Hispanic (22%) and non-Hispanic (25%) searches seen for overall searches in Table 25 (24% Hispanic drivers and 27% non-Hispanic drivers). Non-Hispanic White drivers had a hit rate of 26% for probable cause searches, which was the most comparable hit rate across race/ethnicity by search reason.

Table 26: Combined Hit Rate by Search Reason for Hispanic, Non-Hispanic Drivers, Non-Hispanic White Drivers

	Founded Suspicion of a Crime with Recorded Consent		Reasonable Suspicion/Frisk for Weapon		Probable Cause of a Crime / Automobile Exception	
	N	%	N	%	N	%
Hispanic Not Productive	121	82%	193	71%	206	78%
Hispanic Hit Rate	27	18%	77	29%	59	22%
Total Hispanic	**148**		**270**		**265**	
Non-Hispanic Not Productive	283	83%	364	65%	378	75%
Non-Hispanic Hit Rate	60	17%	199	35%	129	25%
Total Non-Hispanic	**343**		**563**		**507**	
Non-Hispanic White Not Productive	113	73%	161	61%	183	74%
Non-Hispanic White Hit Rate	41	27%	101	39%	65	26%
Total Non-Hispanic White	**154**		**262**		**248**	

The differences in hit rates for Founded Suspicion and Probable Cause Searches were not statistically significant. The hit rate difference for Reasonable Suspicion was significant at the $p<.05$ level for Hispanic compared to non-Hispanic drivers. That is, the finding that SCPD officers who search Hispanic drivers identified contraband at a rate 27% lower than non-Hispanics is less than 5% due to natural variation. The hit rate difference for Reasonable Suspicion was also significant at the $p<.05$ level for Hispanic compared to non-Hispanic White drivers and Hispanic drivers were 36% less likely to have contraband found in a Reasonable Suspicion search.

Table 27: Binary Logistic Regression of Combined Hit Rate by Combined Search Reason (Hispanic and Non-Hispanic & Hispanic and Non-Hispanic White Drivers)

Hit Rate = 1	Variable	β	$SE\,\beta$	Wald's χ^2	p	Odd's Ratio (e^{β})
Founded Suspicion	Hispanic (0 Non, 1 Hispanic)	.051	.256	.040	.842	1.052
	Constant	-1.551	.142	119.103	<.001	.212
	Hispanic (0 NH White, 1 Hispanic)	-.486	.280	3.009	.083	.615

Hit Rate = 1	Variable	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
	Constant	-1.014	.182	30.921	<.001	.363
Reasonable Suspicion	Hispanic (0 Non, 1 Hispanic)	-.315	.161	3.826	.050*	.730
	Constant	-.604	.088	46.914	<.001	.547
	Hispanic (0 NH White, 1 Hispanic)	-.453	.185	5.976	.015*	.636
	Constant	-.466	.127	13.494	<.001	.627
Probable Cause	Hispanic (0 Non, 1 Hispanic)	-.175	.179	.954	.329	.839
	Constant	-1.075	.102	111.162	<.001	.341
	Hispanic (0 NH White, 1 Hispanic)	-.215	.207	1.086	.297	.806
	Constant	-1.035	.144	51.390	<.001	.355

* p<.05; ** p<.01; *** p<.001

Finding: Hispanic drivers searched (person and/or vehicle) because the officer stated there was Reasonable Suspicion had lower hit rates. This finding suggests that, overall, Hispanic drivers who were searched were searched at a lower objective evidentiary threshold than non-Hispanic drivers and White non-Hispanic drivers.

Arrest

There is no variable indicating for what or why an individual was arrested. An arrest may be related to the alleged violation precipitating the stop, the search (if applicable) conducted during the stop, or an outstanding warrant discovered during the stop. Therefore, relationships shown between "hit rate" and arrest or stop type and arrest may be driven by factors unknown from this dataset.

Arrests were an uncommon outcome of driver stops. Arrests occurred following 1,223 (0.77%) of 158,142 stops. While Hispanic drivers made up 29% of all drivers stopped, 35% of the drivers who were arrested were Hispanic. In 2024, 0.94% of Hispanic drivers (425 of 45,083) who had been stopped by the SCPD were arrested. Only 0.71% of the non-Hispanic drivers who were stopped ended up being arrested (798 of 113,059).

Table 28: Arrest Number and Proportion of Stops

	Arrest	% of Stops Ending in Arrest	Total Stops
Hispanic Drivers	425	0.94%	45,083
Non-Hispanic Drivers	798	0.71%	113,059
Total Drivers	1,223	0.77%	158,142

A t-test shows these averages differ in a statistically significant way but do not point to a direct cause. More exploration is necessary to determine if additional factors are driving this disparity.

Table 29: t-test Comparing Average Arrest Rates for Hispanic and Non-Hispanic Drivers

	Hispanic 45,083		**Non-Hispanic 113,059**		***t***		
	M	SD	M	SD	Score	*p*	Cohen's d
Arrest	.0094	.0094	.0071	.08372	-4.566	<.001***	.08759
							Two sided t-test
							* p<.05; ** p<.01; *** p<.001

The dataset lacks a variable noting the reason for arrest. The binary logistic regression shows Hispanic drivers stopped had an arrest rate 26% higher than non-Hispanic drivers who had been stopped. This finding was statistically significant at the p<.001 level. Drivers with NY plates or stopped by Highway Patrol officers were less likely to be arrested.

Table 30: Binary Logistic Regression on Arrest Rate

Arrest = 1	**β**	***SE* β**	**Wald's χ^2**	***p***	**Odd's Ratio (e^{β})**
Hispanic (0 Non, 1 Hispanic)	.234	.061	14.844	<.001***	1.264
NY Plates (0 No, 1 Yes)	-.210	.075	7.936	.005**	.810
25 or Younger (0 No 26 and older, 1 Yes)	-.721	.081	78.536	.107	.486
Patrol (0 Highway, 1 Precinct)	1.464	.115	161.296	<.001***	4.324
Constant	-5.512	.132	1752.334	<.001	.004
					* p<.05; ** p<.01; *** p<.001

A binary logistic regression was performed for each stop reason to investigate whether Hispanic overrepresentation was concentrated for certain stop types. This showed that while Hispanic drivers were not more likely to be arrested that non-Hispanic drivers when stopped for most violations, they were 87% more likely to be arrested when stopped for a Moving Violation (p<.001), 39% more likely to be arrested when stopped for Other Moving Violation Stop (p<.05), and 52% more likely to be arrested when stopped for Speeding (p<.05).

Table 31: Binary Logistic Regression on Arrest Rate by Stop Type when Hispanic Variable Had Statistical Significance

Arrests = 1	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Moving Violation					
Hispanic (0 Non, 1 Hispanic)	.624	.167	13.963	<.001***	1.866
NY Plates (0 No, 1 Yes)	-.467	.274	2.911	.088	.627
25 or Younger (0 No 26 and older, 1 Yes)	-.070	.198	.126	.722	.932
Patrol (0 Highway, 1 Precinct)	1.333	.265	25.283	<.001***	3.791
Constant	-6.014	.351	293.602	<.001	.002
Other Moving Violation					
Hispanic (0 Non, 1 Hispanic)	.331	.133	6.215	.013*	1.392
NY Plates (0 No, 1 Yes)	-.393	.187	4.409	.036*	.675
25 or Younger (0 No 26 and older, 1 Yes)	-.316	.182	3.015	.083	.729
Patrol (0 Highway, 1 Precinct)	2.230	.224	99.261	<.001***	9.304
Constant	-5.762	.269	458.409	<.001	.003
Speeding					
Hispanic (0 Non, 1 Hispanic)	.421	.212	3.933	.047*	1.523
NY Plates (0 No, 1 Yes)	-.845	.290	8.496	.004**	.429
25 or Younger (0 No 26 and older, 1 Yes)	-.051	.224	.052	.820	.950
Patrol (0 Highway, 1 Precinct)	.798	.212	14.095	<.001***	2.220
Constant	-5.156	.312	272.833	<.001	.006

* p<.05; ** p<.01; *** p<.001

Finding: For all stops combined, the likelihood of arrest increased when the driver was Hispanic.

Finding: The Hispanic binary flag was statistically significant for arrests during Moving Violation, Other Moving Violation, and Speeding stop types.

Prior research about disproportionate stops by race or ethnicity, including research focused on Suffolk County, have found a relationship between arrest and Hispanic ethnicity.[20] That research has not typically compared Hispanic to White non-Hispanic groups. The two calculations above are repeated below comparing Hispanic drivers to White non-Hispanic drivers. Conflating White non-Hispanic drivers and other non-Hispanic drivers can hide discrepancies, particularly involving Black/African American drivers, who are arrested at dramatically higher rates than any other race or ethnicity category. Black drivers made up only 20% of the drivers stopped (29,643 of 144,782) but 30% the people arrested were Black (354 of 1,182). These drivers were arrested at double the proportion of White non-Hispanic drivers (1.2% versus 0.6%). As they were about one-third of the arrest population, the differences artificially converged to an average 0.9% when combined with White non-Hispanic drivers.

[20] Stonewall Analytics. February 2025. 2023 Traffic Stop Data Analysis: Suffolk County Police Department.

Table 32: Number and Proportion of Arrests by Race/Ethnicity

	Arrest	% of Arrests	% of Drivers Arrested	Total Drivers	% of Drivers
Black/African American	354	30%	1.2%	29,643	20%
Hispanic	425	36%	0.9%	45,083	31%
White Non-Hispanic	403	34%	0.6%	70,056	48%
Total	**1,182**	**100%**	**0.8%**	**144,782**	**100%**

The t-test shows the mean arrest rate between Hispanic and White non-Hispanic drivers should not be considered equal. This is the same finding as Table 29 and at the same p<.001 level.

Table 33: t-test Comparing Average Arrest Rates for Hispanic and White Non-Hispanic Drivers

	Hispanic 45,083		White Non-Hispanic 70,056		*t*		
	M	SD	M	SD	Score	*p*	Cohen's d
Arrest	.0094	.0966	.0058	.0756	6.838	<.001***	.08448
							Two sided t-test
							* p<.05; ** p<.01; *** p<.001

The binary logistic analysis shows the odds Hispanic drivers were arrested were about 55% higher than White non-Hispanic drivers after being stopped by the SCPD. The odds ratio increased from 1.264 when Hispanic drivers were compared to all other non-Hispanic drivers (see Table 30) to 1.549 when Hispanic drivers were compared only to White non-Hispanic drivers. Both flags were statistically significant (p<.001).

Table 34: Binary Logistic Regression Comparing Average Arrest Rates for Hispanic and White Non-Hispanic Drivers

Arrest = 1	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic (0 White non-Hispanic, 1 Hispanic)	.438	.071	38.26	<.001***	1.549
NY Plates (0 No, 1 Yes)	-.624	.106	34.71	<.001***	.536
25 or Younger (0 No 26 and older, 1 Yes)	-.280	.092	9.36	.002**	.756
Patrol (0 Highway, 1 Precinct)	1.463	.144	103.78	<.001***	4.319
Constant	-5.779	.170	1,155.11	<.001	.003

* p<.05; ** p<.01; *** p<.001

The difference by stop type was statistically significant for Moving Violation, Other Moving Violation, and Speeding (just like the Hispanic/non-Hispanic difference). In this model, the difference for Stop Sign was also statistically significant. The odd's ratios increased when

using a Hispanic/ White non-Hispanic flag. Hispanic drivers stopped for Moving Violations were 2.2 times as likely to be arrested than their White non-Hispanic counterparts, 1.5 times as likely on Other Moving Violations, about twice as likely for Speeding, and 1.9 times as likely for Stop Sign stops. The Hispanic flag was not significant for the remaining seven stop types.

Table 35: Binary Logistic Regression Comparing Average Arrest Rates by Stop Type When Hispanic Driver Variable is Significant

Arrests = 1	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Moving Violation					
Hispanic (0 Non, 1 Hispanic)	.801	.190	17.736	<.001***	2.229
NY Plates (0 No, 1 Yes)	-.334	.350	.908	.341	.716
25 or Younger (0 No 26 and older, 1 Yes)	-.072	.228	.101	.751	.930
Patrol (0 Highway, 1 Precinct)	1.227	.306	16.054	<.001***	3.411
Constant	-6.215	.439	200.043	<.001	.002
Other Moving Violation					
Hispanic (0 White NH, 1 Hispanic)	.405	.151	7.143	.008**	1.499
NY Plates (0 No, 1 Yes)	-.371	.232	2.564	.109	.690
25 or Younger (0 No 26 and older, 1 Yes)	-.406	.222	3.341	.068	.666
Patrol (0 Highway, 1 Precinct)	2.314	.279	68.701	<.001***	10.115
Constant	-5.917	.342	299.752	<.001	.003
Speeding					
Hispanic (0 Non, 1 Hispanic)	.712	.246	8.362	.004**	2.038
NY Plates (0 No, 1 Yes)	-.440	.431	1.044	.307	.644
25 or Younger (0 No 26 and older, 1 Yes)	-.080	.274	.085	.770	.923
Patrol (0 Highway, 1 Precinct)	.890	.271	10.772	.001***	2.436
Constant	-5.855	.476	151.093	<.001	.003
Stop Sign					
Hispanic (0 Non, 1 Hispanic)	.628	.290	4.708	.030*	1.874
NY Plates (0 No, 1 Yes)	-1.136	.398	8.167	.004**	.321
25 or Younger (0 No 26 and older, 1 Yes)	.413	.309	1.788	.181	1.511
Patrol (0 Highway, 1 Precinct)	-3.064	.639	23.030	<.001***	.047
Constant	-1.918	.697	7.569	.006	.147

* p<.05; ** p<.01; *** p<.001

Finding: A larger proportion of Hispanic drivers were arrested than non-Hispanic drivers overall and for specific stop types. The rate of difference increased when the binary flag was Hispanic compared to White non-Hispanic and one additional stop type showed an increase in proportion arrested that was statistically significant.

Search and Arrest Relationship

Hispanic drivers who were *not* searched were arrested at much higher rates than non-Hispanic drivers who were stopped but not searched. A binary logistic regression was

performed comparing arrest rate by search type (none, unproductive, and productive). The Hispanic driver flag was statistically significant for arrests occurring without a search at the p<.001 level. Being Hispanic increased the likelihood of arrest by 92% compared to White non-Hispanic drivers (the variable was significant when assessing Hispanic versus total non-Hispanic drivers at the p<.01 level). The flag is not significant for those searched regardless of outcome.

Table 36: Binary Logistic Regression by Search and Outcome Comparing Hispanic and White non-Hispanic Drivers

Arrests = 1	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
No Search					
Hispanic (0 White NH, 1 Hispanic)	0.65	0.156	17.384	<.001***	1.915
NY Plates (0 No, 1 Yes)	-0.709	0.222	10.187	0.001**	0.492
25 or Younger (0 No 26 and older, 1 Yes)	-0.419	0.208	4.074	0.044*	0.657
Patrol (0 Highway, 1 Precinct)	1.44	0.312	21.294	<.001***	4.222
Constant	-7.339	0.366	401.827	<.001	0.001
Unproductive Search					
Hispanic (0 White NH, 1 Hispanic)	0.11	0.133	0.687	0.407	1.117
NY Plates (0 No, 1 Yes)	0.361	0.197	3.355	0.067	1.434
25 or Younger (0 No 26 and older, 1 Yes)	-0.64	0.16	16.051	<.001***	0.527
Patrol (0 Highway, 1 Precinct)	-1.06	0.356	8.875	0.003**	0.346
Constant	0.688	0.402	2.933	0.087	1.99
Search Hit					
Hispanic (0 White NH, 1 Hispanic)	0.05	0.22	0.052	0.82	1.052
NY Plates (0 No, 1 Yes)	-0.165	0.308	0.288	0.591	0.848
25 or Younger (0 No 26 and older, 1 Yes)	-0.66	0.265	6.208	0.013*	0.517
Patrol (0 Highway, 1 Precinct)	-21.327	11011	0	0.998	0
Constant	21.611	11011	0	0.998	2.429

* p<.05; ** p<.01; *** p<.001

Probable Cause of a Crime and Reasonable Suspicion of a Crime

When officers stopped a car based on Probable Cause of a Crime or Reasonable Suspicion of a Crime, the outcomes (and the racial disparities) were different. These stops, unsurprisingly, are more likely to result in arrests than other stops, and a lower proportion of summonses and warnings were issued than for other stops. They are the only two types of stops in which Hispanic drivers do not receive a disproportionately higher number of summonses compared to warnings than non-Hispanic drivers (see Table 14, above). But isolating these two types of stops is particularly useful in analyzing arrests, since people stopped for these reasons are arrested at higher rates overall.

When stopped based on Probable Cause of a Crime, Hispanic and White non-Hispanic drivers are each arrested 47.4% of the time and issued Field Appearance tickets at similar rates (3.8% versus 3.2%). The total non-Hispanic population had a lower proportion of

arrests (44.4%), but higher field tickets (4.1%), so a similar overall proportion with continued criminal legal system involvement (48.5%). Hispanic drivers were issued summonses in 17% of these stops, which was less frequently than White non-Hispanic drivers (23%) or non-Hispanic (21%) drivers. They received verbal warnings more frequently (23%) than the other two groups (about 16% for each). Although Hispanic driver stops ended with no police action 9% less frequently than the other two cohorts (10.5% White non-Hispanic and 14% non-Hispanic), the proportion of verbal warnings was so much higher that the subtotal of concluded interaction was also higher. One-third of probable cause stops for Hispanic drivers concluded with no further activity necessary for the driver (32%) versus about a quarter for White non-Hispanic drivers (26%).

Table 37: Probable Cause of a Crime Stop Dispositions for Hispanic, White non-Hispanic, and Non-Hispanic Drivers

	Hispanic		White non-Hispanic		Non-Hispanic	
	N	%	N	%	N	%
Arrest	37	47.4%	45	47.4%	76	44.4%
Field Appearance Tickets	3	3.8%	3	3.2%	7	4.1%
Subtotal Criminal Legal System Involvement	**40**	**51%**	**48**	**51%**	**83**	**49%**
Summons	**13**	**17%**	**22**	**23%**	**36**	**21%**
Verbal Warning	18	23.1%	15	15.8%	28	16.4%
No Police Action	7	9.0%	10	10.5%	24	14.0%
Sub Concluded Interaction	**25**	**32%**	**25**	**26%**	**52**	**30%**
Total	**78**	**100%**	**95**	**100%**	**171**	100%

A binary logistic regression was performed with each outcome as the dependent variable. Arrest, Field Appearance, and the subtotal Criminal Legal System Involvement did not have the Hispanic variable as statistically significant (as Hispanic vs All or Hispanic vs White non-Hispanic). This was not a surprise since the proportions are equivalent. The Hispanic driver flag was not significant for the following outcomes: Summons Issued, Verbal Warning, No Police Action, and the combined Concluded Interaction.

Notably, when stopped based on Probable Cause of a Crime, Hispanic drivers were more likely to receive a verbal warning than a summons, and more likely than White non-Hispanic drivers to receive a verbal warning as opposed to a summons. Hispanic drivers received a verbal warning in 23.1% of these stops while White non-Hispanic drivers were warned only 15.8% of the time.

In the table below, the same model was run with receiving a verbal warning as the binary dependent variable. Probable cause stop was an independent dummy variable. For Hispanic drivers, it had no statistically significant impact. But for White non-Hispanic and total non-Hispanic drivers, a probable cause stop reduced the odds of receiving a verbal warning compared to a summons by 51% and 40% respectively. There are a limited number of observations, so a strong conclusion is not possible, but the lack of statistical significance suggests either a different standard for probable cause is used for Hispanic

drivers or the same standard is not appropriate because it does not yield summons worthy behavior for Hispanic drivers.

Table 38: Binary Logistic Regression Impact of Probable Cause of a Crime Stop on Likelihood of Verbal Warning Instead of Summons for Hispanic, White Non-Hispanic, and Non-Hispanic Drivers

Verbal Warning v Summons	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic					
Probable Cause (0 No, 1 Yes)	0.471	0.367	1.65	0.199	1.602
NY Plates (0 No, 1 Yes)	0.026	0.033	0.627	0.429	1.027
25 or Younger (0 No 26 and older, 1 Yes)	-0.086	0.022	15.449	<.001***	0.917
Patrol (0 Highway, 1 Precinct)	0.834	0.026	1019.722	<.001***	2.302
Constant	-0.898	0.038	552.69	<.001	0.407
White non-Hispanic					
Probable Cause (0 No, 1 Yes)	-0.712	0.336	4.504	0.034*	0.49
NY Plates (0 No, 1 Yes)	0.1	0.034	8.461	0.004**	1.105
25 or Younger (0 No 26 and older, 1 Yes)	-0.351	0.02	300.1	<.001***	0.704
Patrol (0 Highway, 1 Precinct)	0.52	0.018	828.839	<.001***	1.682
Constant	-0.228	0.035	41.478	<.001	0.796
Non-Hispanic (Total)					
Probable Cause (0 No, 1 Yes)	-0.505	0.253	3.983	0.046*	0.603
NY Plates (0 No, 1 Yes)	0.079	0.023	11.309	<.001***	1.082
25 or Younger (0 No 26 and older, 1 Yes)	-0.328	0.016	443.453	<.001***	0.72
Patrol (0 Highway, 1 Precinct)	0.547	0.014	1483.56	<.001***	1.727
Constant	-0.279	0.025	127.442	<.001	0.757

* p<.05; ** p<.01; *** p<.001

On stops which officers reported were conducted because they had a reasonable suspicion that criminal activity had or was occurring, Hispanic, White non-Hispanic, and total non-Hispanic drivers had comparable arrest rates around 11%. Hispanic drivers received a slightly higher percentage of Field Appearance tickets, so Hispanic drivers had the highest subtotal (16%) of stops resulting in continued involvement in the criminal legal system. As with Probable Cause stops, Hispanic drivers received a higher proportion of verbal warnings compared to summonses. The proportion (40%), however, was almost equivalent with the comparison cohorts (38% for White non-Hispanic and 40% non-Hispanic) as was no police action. Resultantly, the concluded interaction subtotal was the same for Hispanic and non-Hispanic drivers at 64%, which was slightly higher than 62% for White non-Hispanic drivers.

Table 39: Reasonable Suspicion of a Crime Stop Dispositions for Hispanic, White non-Hispanic, and Non-Hispanic Drivers

Reasonable Suspicion	Hispanic		White non-Hispanic		Non-Hispanic	
	N	%	N	%	N	%
Arrest	16	11%	20	10%	41	11%
Field Appearance	7	5%	7	3%	14	4%
Subtotal Criminal Legal System Involvement	**23**	**16%**	**27**	**13%**	**55**	**15%**
Summons	**28**	**20%**	**49**	**24%**	**74**	**20%**
Verbal Warning	57	40%	77	38%	146	40%
No Police Action	33	23%	48	24%	86	24%
Sub Concluded Interaction	**90**	**64%**	**125**	**62%**	**232**	**64%**
Total	**141**		**201**		**361**	

None of these outcomes have a statistically significant Hispanic driver flag in a t-test or a binary logistic regression. Although these stops exclusively involve instances where the officer stated there was reasonable suspicion that criminal activity was underway, only one-third of those stopped were arrested, given a field appearance ticket, or issued a summons. While nearly half of those stopped because the officer stated there was probable cause that a crime had taken place, the high rate at which Reasonable Suspicion stops do not end with even a summons suggests officers are casting a wide net without much return.

Over six in ten Hispanic and non-Hispanic drivers received a Warning or No Further Police Action; four in ten received a warning. The same model was run with receiving a verbal warning (0) versus a summons (1) as the binary dependent variable. Reasonable Suspicion was an independent dummy variable. For Hispanic drivers, a Reasonable Suspicion stop type made the odds of a verbal warning 219% higher than for other stop types and was statistically significant at the p<.001 level. The Reasonable Suspicion stop type was not significant for White non-Hispanic drivers but was for total non-Hispanic drivers for whom a Reasonable Suspicion increased Verbal Warnings by 52% and this difference is statistically significant (p < .01).

Table 40: Binary Logistic Regression on Impact of Reasonable Suspicion of a Crime Stop on Likelihood of Verbal Warning Instead of Summons for Hispanic, White non-Hispanic, and Non-Hispanic Drivers

Verbal Warning v Summons	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
Hispanic					
Reasonable Suspicion (0 No, 1 Yes)	0.786	0.231	11.54	<.001***	2.194
NY Plates (0 No, 1 Yes)	-0.086	0.022	15.478	<.001***	0.917
25 or Younger (0 No 26 and older, 1 Yes)	0.027	0.033	0.645	0.422	1.027
Patrol (0 Highway, 1 Precinct)	0.834	0.026	1019.842	<.001***	2.303
Constant	-0.899	0.038	553.269	<.001	0.407

Verbal Warning v Summons	β	*SE* β	Wald's χ^2	*p*	Odd's Ratio (e^{β})
White non-Hispanic					
Reasonable Suspicion (0 No, 1 Yes)	0.15	0.184	0.67	0.413	1.162
NY Plates (0 No, 1 Yes)	-0.349	0.02	297.414	<.001***	0.705
25 or Younger (0 No 26 and older, 1 Yes)	0.099	0.034	8.347	0.004**	1.104
Patrol (0 Highway, 1 Precinct)	0.52	0.018	829.189	<.001***	1.682
Constant	-0.228	0.035	41.38	<.001	0.796
Non-Hispanic (Total)					
Reasonable Suspicion (0 No, 1 Yes)	0.419	0.143	8.533	0.003**	1.52
NY Plates (0 No, 1 Yes)	-0.327	0.016	442.573	<.001***	0.721
25 or Younger (0 No 26 and older, 1 Yes)	0.08	0.023	11.609	<.001***	1.083
Patrol (0 Highway, 1 Precinct)	0.546	0.014	1481.775	<.001***	1.727
Constant	-0.28	0.025	128.182	<.001	0.756

* p<.05; ** p<.01; *** p<.001

Finding: Hispanic drivers were over two-fold more likely to receive a verbal warning versus a summons when the stop reason was Reasonable Suspicion compared to other stop types.

Reasonable suspicion as a stop reason and/or a search reason produces lower rates of criminal legal involvement or hit rates for Hispanic drivers. Only 37% of reasonable suspicion stops had any continued system involvement, another 20% resulted in a summons, and 33% of reasonable suspicion searches produced hits. The proportions for Hispanic drivers were lower, with 17% receiving a summons and 29% search hit rates. These differences suggest variations in the applicability of the reasonable suspicion standard with Hispanic drivers.