Technical Appendix:
Investigation of Aurora Police Department and Aurora Fire Rescue

Table of Contents

Technical Appendix

Scope of Assignment

The assignment involves conducting systematic empirical analysis in connection with the pattern and practice investigation of the Aurora Police Department ("APD") and related entities being conducted by the Office of the Attorney General for the State of Colorado.

The objective of the analysis is to conduct empirical tests for results of statistical significance indicative of disparity or disproportionality across race/ethnicity and other demographics (e.g., age, gender).

Technical oversight and support for the scope of this assignment necessitated access to capabilities and extensive experience with relational databases, dataset construction and statistical analysis.

The scope of assignment entailed the following tasks:

A. To query from a *relational database* maintained by the city of Aurora, records applicable to preparing a dataset for empirical analysis of interactions, arrests and use-of-force incidents involving Aurora police and subjects categorized by certain demographic characteristics (e.g., race/ethnicity, age, gender). This also involved a range of data processing tasks to prepare the dataset for empirical analysis.

B. To conduct an empirical analysis of the aforementioned dataset, in accordance with fundamental methodological principles and practices of statistical data analysis, and to motivate supplemental analysis relevant to further inquiry and analysis, based upon the indicative results of the empirical analytical results. Upon pre-processing and then examining the available data and conferring regarding properties of the data within the context of the assignment, the consensus judgement was to employ chi-squared tests for homogeneity across interactions with APD officers across categories (e.g., race/ethnicity and other related demographics) of subjects interacting with APD.

C. To review relevant literature, in order to highlight applicable fundamental empirical principles and illustrate foundational methodological practices related to procedural governance of data collection, acquisition, hygiene, curation, lineage and provenance, for prospective dashboard analytics with an extensive range of functionality for relevant use-case applications, motivated by both the indicative and exploratory empirical results.

This technical appendix describes an independent and objective statistical analysis for which the participating experts reserve the right to update and/or supplement the opinions and methodology contained in this memorandum should new information become available. The analysis is limited to the scope of assignment within the context of statistical analysis and related disciplines as described herein, and to which the results may be subject to further investigation. The statistical analysis as expressed is specific to the scope of assignment as described herein, and should not be deemed valid if taken out of context or misapplied.

This appendix represents the totality of the analysis conducted and the opinions presented by the Data Analytics team, as a self-contained independent complement to the exposition in the primary report (Investigation of Aurora Police Department and Aurora Fire Rescue), to provide substantive background and context for the exploratory empirical and statistical analysis conducted.

This analysis and corresponding technical appendix are not legal analysis – and should not be construed as such – nor is the terminology adopted herein being employed in a legal context, and does not constitute any legal conclusion. Therefore, it should be further noted more specifically that within this technical appendix the use of terminology related to notions of disproportionality and disparity are being employed solely for analytical purposes and not as legal terminology. It is understood that the findings and legal conclusions in the Investigation of Aurora Police Department and Aurora Fire Rescue Office conducted by the Attorney General for the State of Colorado rely in part on the empirical analysis as described in this Technical Appendix.

Summary of Empirical Results

As further described in Section IV, the empirical results as presented exhibit a persistent pattern across interactions, arrests and use of force ("UoF") occurrences by APD which are indicative of both disproportionality and disparity conditioned on the race/ethnicity of subjects. These results precipitate further analysis as described in Section V.

Backgrounds/Qualifications of the Data Analytics Team

Analysis roles and tasks were allocated and apportioned across the Data Analytics Team, comprised of participating experts from Risk Economics®, Inc.[1] and Compass Lexecon,[2] based upon respective domain expertise and experience, specialist skills (and access to relevant

---

[1] On November 18, 2020 Risk Economics® was formally commissioned by the Office of the Attorney General for the State of Colorado to assist as a technical advisor *pro bono* in the pattern and practice investigation of the Aurora Police Department (APD) and related entities. As co-founding principals of Risk Economics® Inc. (https://riskeconomicsinc.com), Dr. Mordecai and Ms. Kappagoda each participated in the methodological aspects of this analysis as research in the public interest being conducted at RiskEcon® Lab for Decision Metrics @ Courant Institute of Mathematical Sciences NYU (https://wp.nyu.edu/riskeconlab/), and in their respective capacities as Visiting Scholars at Courant since 2011 (https://math.nyu.edu/dynamic/people/visitors/). RiskEcon® Lab Industry Research Associate Nicholas F. Joseph is acknowledged for review of this exposition and helpful comments.

[2] On January 15, 2021 Compass Lexecon was formally commissioned by the Office of the Attorney General for the State of Colorado to provide supplemental technical and data analysis support *pro bono* for the pattern and practice investigation of the APD and related entities, in close coordination with Dr. Mordecai and Ms. Kappagoda. Compass Lexecon (https://www.compasslexecon.com/) is a leading consulting firm providing analysis of complex issues related to economics, finance for legal and regulatory proceedings, strategic decisions, and public policy.

resources), primarily related to the following: (i) data querying, retrieval and dataset construction; (ii) hypothesis/model specification and implementation; (iii) experimental design, empirical analysis and interpretation of results; (iv) technical oversight and peer review of relevant literature on data curation and provenance and empirical methodology.

The backgrounds and qualifications of the Data Analytics Team comprise a diverse and extensive range of experience and expertise across the following relevant quantitative and analytical fields, disciplines and practice domains: econometrics, statistics, economics, quantitative social science (e.g., demographics).

The Data Analytics Team working on this Technical Appendix was comprised of the following participants:

Dr. David K. A. Mordecai: Dr. Mordecai is President of Risk Economics®, an advisory firm which specializes in risk and liability management (as well as forensic analytics), at the intersection of commercial and industrial process engineering. He is also Visiting Scholar at Courant Institute of Mathematical Sciences New York University (NYU), leading research activities at RiskEcon® Lab for Decision Metrics @ Courant Institute. His research focuses on applying a range of computational statistics, economics and related methods to forensic analysis and damage estimation across industrial, commercial, and societal domains. He holds a PhD with concentrations in Econometrics/Statistics and Economics/Industrial Organization from the University of Chicago Graduate School of Business, as well as an MBA in Finance from NYU.

Ms. Samantha Kappagoda: Ms. Kappagoda, Chief Economist at Risk Economics®, is also Visiting Scholar at Courant Institute of Mathematical Sciences New York University (NYU), co-leading research activities at RiskEcon® Lab for Decision Metrics @ Courant Institute. Her primary focus is the development of analytics, for modeling trends and patterns across global asset classes, markets, regions, sectors and industries. Ms. Kappagoda holds an MBA in Analytic Finance, Statistics, and International Business from the University of Chicago Graduate School of Business, a Master’s degree in Economics from the University of Toronto, as well as a Bachelor’s degree in Mathematics from Imperial College London.

Mr. Michael Kwak: Mr. Kwak is an Executive Vice President at Compass Lexecon, a leading economic consulting firm that provides support and analysis to law firms, corporations, and government clients. Mr. Kwak specializes in the development and implementation of econometric analyses and statistical analyses in the fields of finance and economics. Specifically he has extensive experience employing statistical analysis in assessing allegations of discriminatory practices related to age, gender, and race in various ERISA and EEOC discrimination matters. He holds a Bachelor’s degree in Economics from Columbia University and a Master’s degree in Economics from NYU.

Mr. Mihir Gokhale: Mr. Gokhale is a Vice President at Compass Lexecon, and regularly consults on the development and implementation of statistical and empirical analyses. He holds an MBA and a Bachelor's degree in Economics and Politics from New York University.

Mr. Noah Mathews: Mr. Mathews is a Senior Analyst at Compass Lexecon. He holds a Bachelor's degree in Economics from Grinnell College, and previously served as a Senior Research Assistant at the Federal Reserve.

Mr. Peter Horvath: Mr. Horvath is a Senior Analyst at Compass Lexecon. He holds an MA in Political Economy from NYU and a Bachelor's degree in Politics and German from the University of Bath.

Introduction

As detailed below, these empirical results are indicative of disproportionality and disparity of interactions, arrests, and UoF incidents by APD predominantly involving Black/African American and Hispanic subjects across districts comprising Aurora.[3] [4] [5] As further articulated in the discussion of empirical results, disproportionality and disparity are generally exhibited by (i) the *representative ratio* of observed APD officer interactions involving racial/ethnic subgroups (in comparison to the white subgroup), each relative to their respective populations; (ii) the representative ratio of the observed number of arrests of racial minorities versus whites relative to their respective populations and corresponding number of APD officer interactions; and (iii) the representative ratio of the observed number of UoF incidents involving racial minorities versus

[3] The empirical analysis was performed on data over the period January 2018 – February 2021. It should be further noted that this Technical Appendix employs the naming conventions "Black/African American" and "White/Non-Hispanic" to denote corresponding subgroups referred to as "Black" and "white" in Investigation of Aurora Police Department and Aurora Fire Rescue published by the Office of the Attorney General for the State of Colorado.

[4] As stated previously, for the purposes of this statistical analysis, terms in this Technical Appendix regarding disproportional and disparate occurrence are neither adopted nor employed in a legal context and nothing in this appendix constitutes a legal analysis or asserts a legal conclusion. In contrast to the definitions adopted by the labor economics and employment discrimination literature (*see* 31), this appendix adopts the terms *disproportionality* and *disparity* as commonly employed by the criminal justice and social work literature. *See*, *e.g.*, https://oxfordre.com/socialwork/view/10.1093/acrefore/9780199975839.001.0001/acrefore-9780199975839-e-899 "The term 'disproportionality' refers to the ratio between the percentage of persons in a particular racial or ethnic group at a particular decision point or experiencing an event (such as maltreatment, incarceration, school dropouts) compared to the percentage of the same racial or ethnic group in the overall population… Whereas disproportionality refers to the state of being out of proportion, 'disparity' refers to a state of being unequal. In health and social service systems, disparity is typically used to describe unequal outcomes experienced by one racial or ethnic group when compared to another racial or ethnic group (in contrast, disproportionality compares the proportion of one racial or ethnic group to the same racial or ethnic group in the population)."

[5] Knox, Lowe and Mummolo (2020) discuss post-treatment selection decisions, i.e., detainment (e.g., stops, arrests) as a source of sample selection (i.e., *survivor bias*) and as a type of specification error. This Knox et al. discussion is fundamentally motivated by extensive research pioneered by Heckman (1979) and Heckman et al. (e.g., see Bibliography) on data contamination due to non-random selection, prominently and extensively citing relevant research by Heckman among others, and specifically the critique by Durlauf and Heckman (2020) of the police use of force analysis conducted by Fryer (2019), e.g., see intuitive discussion at https://fivethirtyeight.com/features/why-statistics-dont-capture-the-full-extent-of-systemic-bias-in-policing/.

whites relative to their respective populations and corresponding number of APD officer interactions.

The remainder of this appendix describes the statistical analyses conducted by the aforementioned Data Analytics Team, and the results of these analyses. Section I describes the underlying statistical principles of the *chi-squared test of homogeneity* as the primary analytical tool employed for this analysis. Section II describes the sources of data accessed within databases maintained by Aurora IT as repositories for records associated with the APD, and the construction of a dataset for analysis from those sources of data. Section III describes the cleaning and filtering process implemented on the dataset described in Section II. Section IV describes the empirical analysis and results. Section V discusses prospective further analysis based on these preliminary empirical results. Section VI discusses best practices and generally accepted practices for data curation, provenance, and analysis as motivated by Sections IV and V.

I. Chi-Squared Tests of Homogeneity

The objective of this section is to present a general yet technically accurate exposition of the *chi-squared statistic* and its use in the *chi-squared test of homogeneity* within this context for each aforementioned empirical result. In probability theory and statistics, the *chi-squared distribution* (sometimes referred to as the χ^2-distribution) with k degrees of freedom is the distribution of a sum of the squares of k independent standard normal random variables.[6] The chi-squared distribution is a special case of the *gamma distribution*, and among the most widely employed probability distributions in inferential statistics, most notably for hypothesis testing and computation of confidence intervals. A special case of the more general *noncentral chi-squared distribution* is sometimes referred to as the *central chi-squared distribution*.[7]

The primary reason the chi-squared distribution is used extensively in hypothesis testing is its relationship to the *normal distribution*. The simplest chi-squared distribution is the square of a *standard normal distribution*.[8]

[6] Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures".
[7] Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures".
[8] The chi-squared test is the most commonly employed of the applied statistical contingency tests, which compares discrete distributions, i.e., data allocated to two or more categories. The primary rationale most commonly cited for the extensive application of the chi-squared statistic for hypothesis testing – and likelihood ratio testing (e.g., see the *Neyman-Pearson* lemma) – is its corresponding fundamental (asymptotic) relationship as an approximation to the standard normal distribution for large samples of data, as well as its relationship to the *Student t-distribution* commonly applied to tests of statistical significance, among other interrelated statistical distributions applicable to inference and estimation of variance (e.g. the *binomial*, *F-distribution*, *Beta*, *LaPlace*, *Maxwell*, *Pareto*, *Rayleigh*, *Uniform*, *Gamma*, as well as the *exponential* and *Erlang* distributions, the latter two being special cases of the Gamma distribution). It can be shown that a chi-squared distribution computed by squaring a standard normal distribution exhibits one degree of freedom, and as the sample size increases the chi-squared converges (asymptotically) to the normal distribution, and that the chi-squared variable with k degrees of freedom is defined in terms of *sums of the squares* of *k independent standard normal random variables*. Furthermore, a k-dimensional *Gaussian* (i.e., normally distributed) random variable and covariance matrix with rank k is distributed chi-squared with k degrees of freedom which generalizes the chi-squared distribution to other linear combinations of standard

As is customary, the chi-squared tests employed herein are fundamentally tests of observed relative proportions of numerical counts data. By way of illustration, given some number of observations *n*, each such observation is assigned to two (or more) categories *k* (e.g., gender, race/ethnicity, age, income, calls/charges, infraction types/classes, participant roles, etc.), which relate to a collection of representative features for either a type of incident, or a category for those test subjects involved, each related to both the test sample and the populations (and subpopulations) being analyzed.[9] For each of these *n* observations comprising each category, some specified event either occurs or does not occur (i.e., mutually exclusive occurrence), e.g., interactions, arrests, UoF incidents, etc. In this instance, descriptive empirical features for classifying test subjects reasonably include both demographic (e.g., age, gender, race/ethnicity) and socioeconomic (e.g., income) statistics. Descriptive empirical features for classifying incidents for the purposes of this analysis might reasonably include geographic and/or spatial location – in conjunction with (region-specific) distributions of demographic and socioeconomic characteristics – as well as call types and infraction types, among others.

Among its other analytical implementations and applications (e.g., *independence*, *goodness-of-fit*), chi-squared tests can be employed to evaluate whether the observed relative proportion(s) of group A to which a specified incident occurred are comparably equivalent to the observed relative proportions of group B to which the specified incident also occurred, i.e., whether the relative proportion of a specified incident is *homogenous* across subgroups.[10] This is an application of the *chi-squared test of homogeneity*.

normal distributions (i.e., weighted sums of independent random variables each with zero mean and unit variance). Contingency testing, comparing classes of categorical data as arrayed in 2x2 matrices (and by extension to general comparisons of variation across more than 2x2 classes of categories by implementing the *Pearson* chi-squared methodology) is closely related both to estimating (binomial) confidence intervals, and to the simplest 2x1 case *population z-test* for two independent proportions from the same population employing the binomial distribution test for *goodness of fit*. In the simplest case, the chi-squared test is mathematically equivalent to the z-test, since the critical values of the chi-squared distribution for one degree of freedom are equivalent to the square of the corresponding critical values of *z*. By way of further background see e.g., Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures".

[9] As explained above, the Pearson chi-squared methodology, which extends the 2x2 contingency analysis to arrays with categories>2, is derived from the *z-statistic* (which is based upon the normal distribution). The critical values of the chi-squared with one degree of freedom are the squared corresponding critical values of the z-statistic. Hence, the standard 2x2 chi-squared test is equivalent to the z-test of independent proportions drawn from the same population, which is further based upon the 2x1 goodness of fit chi-squared as an implementation of the Binomial test (i.e., *population z-test*), for which each sample observation is compared to a binomially-distributed predicted value. Analogous to the more general Pearson test extended to arbitrary numbers of rows *r* and columns *c*, the aforementioned *z-test* can be extended to *r*x1 chi-squared tests (thereby evaluating an arbitrary number of rows), in order to identify statistical significance in differences across multiple categorical values, although additional analysis will be necessary to ascertain which categorical values exhibit significant variation. Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures". By way of comparison, see subsequent section *Implementation of the Generalized Pearson chi-squared test* (see Footnote 40), related to the Fisher exact test and its chi-squared approximations, as well as Footnote 36 discussing the effect of generally-accepted continuous to discrete adjustments to chi-squared calculations.

[10] Although the chi-squared test is conceptually simple in principle, in practice the underlying operations tend to become increasingly combinatorially complex as permutations and/or combinations scale as the number of categories (columns) and observations (rows) extend beyond the simple 2x2 case. Chi-squared tests, solely relevant for count(s) data, i.e., occurrence(s), are *omnibus* tests which mean that any pattern of departures weighs against H_0 (especially for contingency tables >2x2), for which observations are independent, countable, and classifiable (i.e.,

==

An illustrative hypothetical example of testing for homogeneity. By way of illustration, within a hypothetical city during any given year, police officers might interact with 10% of the population of that city. If the population of this hypothetical city is 100,000 residents, of which 20% of are Black/African American, and the remaining 80% are White/Non-Hispanic, the population consists of 20,000 Black/African American residents and 80,000 White/Non-Hispanic residents. If occurrences of police officer interaction are *homogeneous* across subgroups, of the 10,000 residents of the city expected to have interacted with police (i.e., 100,000 residents * 10%), one would expect that 2,000 of those residents would be Black/African American (i.e., 100,000 * 20% * 10%) and 8,000 of those residents would be White/Non-Hispanic (i.e., 100,000 * 80% * 10%). However, an observed 8,000 of the aforementioned 10,000 police interactions with residents to involve Black/African American residents, versus 2,000 of the 10,000 police interactions to involve White/Non-Hispanic residents, would be indicative evidence of *heterogeneity*, or *non-homogeneity*.

In other words, as described, this illustration implies that while 10% of the population of this hypothetical city has been involved with a police interaction, the ratios of incidents differ across racial/ethnic groups. In this illustration, the Black/African American ratio of incidents would be 40% (i.e., 8,000 / 20,000), whereas the White/Non-Hispanic incidence

---

mutually exclusive and exhaustive categories in accordance with the underlying notions of *separability* and *partitioning*). Additionally, depending on a specific test, for purposes of interpretation and exposition of each chi-squared statistic (versus its corresponding p-value), relative observed proportions for each respective test may need to be weighted in accordance with the appropriate relative Aurora population(s) and in some instances, the respective corresponding relative subpopulation(s), e.g., in a specified zip code or district. It should be noted that, as further explained in a subsequent section (see Footnote 32), these respective reference populations (or subpopulations) are not intended to be discussed in terms of *baselines* – a technical term of art from *structural equation modeling* for causal inference and attribution of unobserved *latent variables* and *path analysis* – highly contested fields of analysis outside the scope of these tests. Apart from causal attribution analysis, the more foundational methodological application of reference class feature selection is ubiquitous to the context of sampling classification in general, and fundamental to statistical analysis in law and economics (particularly relevant to issues of reliability, robustness with respect to the representativeness of reference class samples and selection bias). As applied in the practice of statistical legal analysis, model specification, variable selection and feature validation methods are employed as tools for empirically testing hypotheses related to reference classes for evidential support. Both theoretically and in practice, it is generally acknowledged in such empirical analysis that (a) the *reference class problem* in the context of broader evidential value depends on more than the statistical distribution (or likelihood ratio) exclusively, (b) methodologically different reference class associations might yield different inferences, and (c) by modifying a particular reference class, different comparative results might proceed from the same item of evidence. Such scholarship further acknowledges that the probative value of evidence depends on much more than the likelihood ratio corresponding to a particular reference class sample, and that *a fortiori*, a particular reference class in isolation cannot fully encompass the probative value of empirical evidence. However, empirical evidence may be particularly probative for comparing both quantitative and qualitative differences that may be identifiable between different reference class populations, each being comprised of more or less heterogeneous composites of distinguishing features. See e.g., Cheng, E.K. (2009) “A Practical Solution to the Reference Class Problem,” 109 COLUM. L. REV. 2081, pp. 2095-97; Franklin, J. (2010) “Feature Selection Methods For Solving The Reference Class Problem: Comment on Edward K. Cheng, “A Practical Solution to the Reference Class Problem”” 110 COLUM. L. REV. SIDEBAR, pp. 12–23; Nance, D.A. (2007) “The Reference Class Problem and Mathematical Models of Inference” The International Journal of Evidence & Proof Vol. 11 Issue: 4, pp. 259-273.

would be 2.5% (i.e., 2,000 / 80,000) even though the *average* ratio of incidents equals 10% (i.e., 10% = 40% * 20,000 + 2.5% * 80,000)/(100,000).
===

It should be noted that a chi-squared test statistic is not intended to compute probabilities of occurrence corresponding to a respective category or subgroup, and is neither appropriate to attribute a notion of *frequencies*, nor to extrapolate *rates* of incidence corresponding to categories or subgroups. Instead, the chi-squared statistic specified for this test describes the distribution of relative proportions of the incidents (as counts) corresponding to each of the respective categories specified, to evaluate whether those proportions are comparably equivalent. In this sense, these chi-squared tests compare the ratios of observed samples of occurrences (as counts) to another ratio of counts, being that of the respective subgroup population relative to the overall population. As also indicated in Footnote 9, and discussed further in Footnote 36, given the generally-adopted assumption that discrete (e.g., binomial) probabilities can be (asymptotically) approximated by a continuous distribution, adjustments for the discrete to continuous approximation are customarily applied.

Fundamentally, chi-squared tests – whether as tests of independence, homogeneity or goodness-of-fit – compare proportional differences between sample or population variances relative to observations recorded. The statistical significance for each of these observed proportional differences is commonly exhibited using the computed p-value relative to a *null hypothesis* for the specified test statistic (as either a one-tailed test or a two-tailed test). In the context of these tests, which assume homogeneity as the null hypothesis, the p-value quantifies the rarity of the result relative to random chance as characterized by the corresponding *cumulative distribution function* (CDF). Although computing the test statistic and corresponding p-value for a dataset is typically a simple calculation, computing the sampling distribution under the null hypothesis, and then computing its corresponding CDF can often be more difficult.

Implementation of Chi-squared Tests for Homogeneity

As described above, chi-squared tests of homogeneity compare the relative proportionality of counts for occurrences across categorical variables between two or more populations. In this analysis, the occurrences include interactions, arrests, and UoF incidents, while the categorical variables include, e.g., racial/ethnic groups, gender, district, etc. The objective is to test whether the observed proportions of event occurrences differ across categorical variables (as proxies for subpopulations) by comparing whether or not two (or more) observed proportions are statistically different from each other (based on the p-value corresponding to its respective chi-squared statistic), i.e., whether or not the observed differences are inconsistent with random error based upon the respective degrees of freedom (as described below).[11] For each demographic subgroup

[11] For the statistical significance exhibited at the specified *critical threshold value* $\alpha = .05$ (i.e., the 5% significance level), the corresponding p-value for the chi-squared statistic associated with a respective category is a function of

(e.g., race/ethnicity, age, gender), the test statistic is a ratio between the *estimated (squared) difference* in the observed counts relative to the expected counts of that respective subgroup (under the null hypothesis that the respective proportion for each subgroup is equivalent to that of the control group, e.g., White/Non-Hispanic) divided by its expected counts. The equation below demonstrates the formulation of the chi-squared statistic:

$$\chi^2_{df} = \Sigma_i^n \frac{(o_i - e_i)^2}{e_i}$$

The chi-squared statistic is calculated as the sum across the number of groups within a categorical variable (i=1, 2, …, *n*) of the squared difference between the observed proportion (o_i) and the expected proportion (e_i) divided by the expected proportion. The *observed* proportion is the empirically observed proportion of each subgroup, while the *expected* proportion is the proportion that would be expected in the case of homogeneity across subgroups. In the case of homogeneity, proportions calculated within subgroups should be equal to the proportion across subgroups. Intuitively, the chi-squared statistic represents a measure of how much observed proportions differ from expected homogenous proportions. The larger that deviation is, the less likely the observed proportions are indicative of homogeneity (subject to the degrees of freedom corresponding to the number of categories specified).[12]

In order to assess this likelihood, a p-value is calculated for a given chi-squared statistic. The p-value is the probability of observing a chi-squared statistic from the cumulative probability density function of the chi-squared distribution assuming that there is no difference in proportions between the subgroups (subject to the degrees of freedom constraint) for a chosen *evidential threshold*, often referred to as *alpha* (i.e., α). For a specified level of degrees of freedom, the larger the chi-squared statistic, the smaller the p-value. Using a commonly accepted *critical threshold value* α = 5%, observed p-values that are less than 5% are considered to be statistically significant and are considered to be sufficient to reject the null hypothesis that there is no difference between the cohorts.

As described further in a subsequent section of this appendix, relevant extensions of the chi-squared test for homogeneity entail implementation and interpretation of the *Pearson chi-squared test* for *goodness-of-fit* to a *uniform distribution* – i.e., the similarity to the equal likelihood of observing the relative frequencies of incidents – as well as *logistic* (*logit*) *regressions* (to estimate the odds (*log odds*) of a particular incident occurrence). Additional model specifications and tests

the degrees of freedom, i.e., the relationship between the number of variables (categories) relative to the number of observations. Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures".

[12] As previously explained, the chi-squared statistic is also characterized by the degrees of freedom for the test, i.e., a constraint which for estimating systems of equations in general represents the tradeoff between the specified number of parameters (e.g., categories) as tabular data fields arrayed in columns versus the number of observations (e.g. counts) as tabular data fields arrayed in rows. For the implementation of these tests, the degrees of freedom constraint is customarily specified to be equivalent to the number of categories minus one (i.e., $c - 1$). Sheskin, D.J. (1997) "Handbook of Parametric and Nonparametric Statistical Procedures".

are preliminarily explored subject to further investigation, which is dependent upon supplemental acquisition of suitably curated additional data.

II. Data Available from the Aurora Police Department

A series of discussions between the APD, the Office of the Colorado Attorney General ("COAG") and the Data Analytics Team identified the following two SQL mirrors of two data repositories to be relevant for the corresponding scope of the empirical analysis as previously described: (a) the Records Management System ("RMS") database, a repository of data records associated with activities that occur in the day-to-day operations of APD, e.g., arrests, traffic stops, follow-ups, etc; and (b) the Administrative Investigations Management ("AIM") database which is comprised of recorded APD administrative data, including investigations into UoF incidents.

RMS Database

The data records in the SQL mirror of the RMS database are generally structured as follows. Distinct interactions between an APD officer and a subject are identified by a "Report Number" or "Primary Key." A Report Number may have one to many observations in the RMS database, where each observation under a particular Report Number corresponds to a subject (non-APD-officer) involved in an interaction in some capacity. This capacity is identified by two variables: "Role" and "Role Expansion." Each observation provides demographic detail about a subject, including first and last name, date of birth, race, ethnicity, and sex. The difference between "race" and "ethnicity" as delineated by the data records is discussed further in Section III. The data records relating to a subject involved/Report Number include but are not limited to information related to the case type (i.e., the nature of the interaction, for example, suspicious activity, suspicious occurrence, or domestic disturbance), call type (i.e., the nature of the call that triggered the interaction in the event a call did so), and charges filed against a subject to the extent charges were filed. The RMS database is comprised of a heterogenous collection of tables. The date ranges for these tables are comprised of records extending at least as early as the year 2000.

AIM Database

The data records in the AIM database pertaining to UoF incidents include a variable "Primary Key" which, after some syntax modifications, can be used to join the UoF AIM data with the RMS data. AIM database records comprise variables that indicate event type and a numerical value indicating whether the event type corresponds to a UoF incident. The AIM data indicate whether a UoF incident occurred in relation to a particular individual subject, corresponding to a particular

Report Number, and also indicate the "tier" of UoF employed as determined by the Force Review Board.[13] The earliest UoF incident recorded in the AIM database is dated June 2012.
The analysis primarily focuses on data corresponding to interactions between APD officers and subjects within APD jurisdiction from the RMS database (including, for example, demographic information related to subjects and APD officers involved in the interactions; geolocational and chronological metadata specifying the time and geographic location of the interaction; classifications of the type of interaction; whether arrests were made; nature of the arrest; charges alleged; etc.) and whether a UoF incident occurred, as well as additional fields recording details associated with the respective UoF incidents.

In order to identify and retrieve data with the objective of preparing a dataset for the aforementioned scope of empirical analysis, relevant records comprising the RMS and AIM databases were accessed as follows:

(1) Queried data records from eleven tables in the SQL mirror of the RMS database through a virtual machine on the APD system, arranged by the APD for access by the Data Analytics Team;
(2) Merged the queries of each of the eleven together on event-specific, subject-specific, and other "key" variables identified by representatives of APD;
(3) Queried data records from four tables specifically related to Use of Force investigations identified by the employees of APD housed in a SQL mirror of the AIM database through a virtual machine hosted on APD servers;
(4) Joined the results of step (3) onto the results of step (2) using a combination of a transformation of the "primary key" variable, a "first name" variable, and a "last name" variable.

The above four steps result in a dataset corresponding to interactions between APD officers and subjects in Aurora and records of UoF incidents with some data fields dating back at least ten years (i.e., at least 2010). Notably, the *Case Type* variable, specifying the category of interaction between the APD officer(s) and subjects for incidents, was not populated prior to 2018 thereby limiting the date range feasible for performing reasonably comparative empirical analysis to the period between January 2018 – February 2021.[14]

[13] APD Policies: Directives Manual Chapter 5 *Weapons and Use of Force* discusses review by FRB of Tier II and III UoF incidents (see Section 5.4.7 Reporting and Investigating the Use of Tools Weapons and Physical Force https://public.powerdms.com/AURORAPD/tree/documents/107). It should be noted that either systematic or idiosyncratic conflation of Tiers II/III should be examined as a prospective procedural source of specification error. Also see Footnote 5 regarding specification error and data contamination.
[14] As a result of the unpopulated Case Type field prior to 2018, in order to preserve consistency and comparability, the analysis focused on data for the January 2018 – February 2021 period, over which the Case Type field is populated.

Although certain corresponding database fields were utilized to reconcile RMS and AIM databases, those database fields corresponding to identifiable personal information were neither retrieved nor extracted.[15] [16] [17]

III. Data Filtering and Supplemental Variable Generation

Filtering

For analysis of interaction data described *supra* Section I (and limited to interactions from January 2018 – February 2021 due to empty set for the *Case Type* variable prior to 2018), the implemented series of filters include the following:

(1) The analysis limits the dataset to observations where each person is eighteen years of age or older on the date of an interaction (i.e., non-juveniles).[18] This step results in the omission of 19.47% of the observations.
(2) Although, as previously mentioned, the RMS database records each subject associated with a particular interaction (as indexed by the GO number), the dataset prepared for analysis limits the data to "subjects" of an interaction, i.e., observations in the data where the "role expansion" variable exhibits a value corresponding to the following categories: "arrestee," "driver/victim," "summons recipient," "offender/suspect," "subject," "victim/arrestee," or "involved." This step results in the omission of 41.93% of the remaining observations.[19]
(3) The analysis omits instances related to GO report numbers corresponding to case types that either would not or did not initiate a "call for service" from an APD dispatcher, in order to focus exclusively on documented interactions between an APD officer and subjects in Aurora. This step results in the omission of 1.09% of the remaining observations.[20]

---

[15] To identify and track subjects across the dataset, the Data Analytics Team utilized the variable "pin" (i.e., personal identification number), a numeric variable (e.g., 94596) that is used to identify individuals throughout the RMS database.
[16] With regard to applicability of the relevant policies and procedures in accordance with the amendment, as previously indicated, the Data Analytics Team was granted access to digitally-imaged metadata and specific data queries (e.g., as needed for validation) hosted on a data repository compliant with the applicable policies, see https://www.fbi.gov/file-repository/cjis_security_policy_v5-9_20200601.pdf/view, the more relevant subject matter is articulated in Sections 5.9 and 5.10 regarding access and security (i.e., "common sense approaches to sensitive data").
[17] In order to comply with CJIS policies, queries were performed on the APD server, and data fields were depersonalized.
[18] As described previously, each observation from the RMS database is associated with a specific subject and GO number.
[19] Steps (1) and (2) have the combined effect of omitting all incidents where the "subject" of an interaction is less than 18 years old.
[20] It is understood based on conversations with APD personnel that the omission of instances associated with case types that do not initiate a "call for service" eliminates from the dataset records of police activity unrelated to police interactions with subjects (e.g., administrative or informational calls).

(4) Some portion of the GO report numbers in the January 2018 – February 2021 dataset have been assigned multiple case types. Report numbers with multiple case types are omitted from the dataset, in order to avoid double counting incidents, and to avoid subjectively selecting which of the multiple case type should be associated with a particular report number.[21] This step results in the omission of 0.81% of the remaining observations.

Supplemental Categorical Indicator Variables

The following *categorical indicator variables* are tabular data fields arranged in columns (with each column attributed to a specific category) and assigned a binary [0,1] value corresponding to its categorical association.

(1) A variable that represents both race and ethnicity: The RMS database consists of two distinct variables for race[22] and ethnicity.[23] A supplemental "race/ethnicity" variable is produced which assigns the "White" race category as "Hispanic or Latino" when the variable race is "White" and the variable ethnicity is "Hispanic or Latino," and as "White/Non-Hispanic" otherwise. The "race/ethnicity" variable is equivalent to the "race" variable otherwise.

(2) A variable that indicates whether a subject in an interaction is alleged to have resisted arrest. This variable is equal to 1 if the Uniform Crime Reporting (UCR) Code associated with any charge against the subject is "4801,"[24] and equal to 0 if no UCR Code associated with any charge against the subject associated with resisting arrest.

(3) A variable that indicates whether a subject in a given interaction is alleged to have failed to obey an order from a police officer. This variable is equal to 1 if the statute associated with any charge against the subject is either "94-110(5)" or "27-69(5),"[25] and equal to 0 if no statute associated with any charge against the subject corresponds to failing to obey an order from a police officer.

---

[21] In order to mitigate effects of double counting, interactions corresponding to GO numbers associated with multiple case types were omitted. However, subject to supplementation dependent upon the availability of suitably curated data, the prospective net effects of survivor bias and other sample selection biases requires further analysis. By way of illustration, the following empirical question remains open: what would be the net effect of including multiple case type interactions categorized by the most severe versus the least severe case types?

[22] The race variable ('race_expansion') contains the following unique entries: "WHITE," "UNKNOWN," "BLACK/AFRICAN AMERICAN," "ASIAN," "AMERICAN INDIAN/ALASKAN N," "NATIVE HAWAIIAN/PACIFIC I," and NA (empty).

[23] The ethnicity variable ('ethnicity_expansion') is comprised of the following entries: "NON-HISPANIC," "HISPANIC OR LATINO," "UNKNOWN," and NA (empty).

[24] The translation of UCR Code "4801" in the RMS database is "OBSTRUCT POLICE RESIST OFFICER." Statute 94-110(5) reads "[f]ails to obey a lawful order or command by a peace officer, firefighter, marshal, or detention officer acting under the color of official authority which causes or is likely to cause harm or a serious inconvenience." The translation of statute 27-69(5) in the RMS database is "Disorderly Conduct/Fail to Obey Order.

[25] The translation of UCR Code "4801" in the RMS database is "OBSTRUCT POLICE RESIST OFFICER." Statute 94-110(5) reads "[f]ails to obey a lawful order or command by a peace officer, firefighter, marshal, or detention officer acting under the color of official authority which causes or is likely to cause harm or a serious inconvenience." The translation of statute 27-69(5) in the RMS database is "Disorderly Conduct/Fail to Obey Order".

(4) A variable that identifies whether the sole charge against a subject in a given interaction is "*resisting arrest*." This variable is equal to 1 if the only Uniform Crime Reporting (UCR) Code associated with any charge against the subject is "4801," and equal to 0 if a charge other than resisting arrest is recorded.
(5) A variable that identifies whether the sole charge against a subject in a given interaction is "*failure to obey*" an order from a police officer. This variable is equal to 1 if the only statute(s) associated with any charge against a subject is either "94-110(5)" or "27-69(5)" and equal to 0 if a charge other than failure to obey is recorded.
(6) Using a concordance between "Case Types" and "Call Types" provided by APD, we add a variable for "Call Type" based on an event "Case Type."[26]

IV. Chi-Squared Test Results

This section discusses results of a statistical analysis of the dataset prepared from the APD databases. As described in Section III, this statistical analysis primarily focuses on differences in relative proportionality of (i) APD interactions, (ii) APD arrests, and (iii) APD UoF incidents across sociodemographic subpopulations within the city of Aurora. Specifically, the analysis employs chi-squared tests to examine empirical evidence indicative of *disproportionate* and *disparate* occurrences of interactions, arrests, and UoF incidents across different racial/ethnic groups.[27] In this context, disproportionality refers to marginal differences in the number of counts relative to the categorical composition of the population, whereas disparity refers to marginal differences in occurrence across categories between comparable types of incidents (e.g., call types).

The analysis considers occurrences of interactions, arrests and UoF incidents to be *disproportionate* if a ratio of interactions, arrests, and/or UoF incidents relative to the subpopulation for specified racial/ethnic groups is higher relative to the corresponding ratio for the White/Non-Hispanic racial/ethnic group.

As explained in detail in Section III, chi-squared tests of homogeneity are methods of analysis that can be applied to test whether ratios of incidents for specified subgroups are (statistically) significantly different from the expected ratios of incidents for those specified subgroups. Figures 1.A, 1.B, and 1.C present chi-squared tests of the ratio (relative to the respective subgroup populations) of (i) interactions, (ii) arrests, and (iii) UoF incidents across each of the different racial/ethnic groups available in the APD database versus the ratio (relative to the respective subgroup population) of (i) interactions, (ii) arrests and (iii) UoF incidents for the White/Non-

[26] Each "Case Type" can be classified as one unique "Call Type." As such, "Call Type" is a more aggregated version of "Case Type." Case type can be found in the CAD Complaint Data Table, and both variables are populated by the police dispatcher as they field a call for service.
[27] See Footnotes 4 and 31 regarding operative definitions for *disproportionate* and *disparate* as terms of art within the scope of this analysis.

Hispanic subgroup for the year 2019.[28] Figures 1.A – 1.C present indicative evidence of *disproportionate* ratios of interactions, arrests, and UoF incidents involving Black/African American subjects and Non-White subjects, versus corresponding ratios involving White/Non-Hispanic subjects in the City of Aurora in the year 2019.

Figure 1.A implements chi-squared tests of the ratio of police interactions to population for each non-white subgroup relative to the White/Non-Hispanic subgroup over the year 2019. As an example, the chi-squared statistic associated with the Black/African American race/ethnicity subgroup ratio of Interactions to Population ("IPR") (22.3%) versus the same ratio for the White/Non-Hispanic subgroup ratio (9.8%) is 5,214.8, with an associated p-value of 0.00. That is, Figure 1.A shows that the difference between the IPRs for Black/African American subjects and White/Non-Hispanic subjects is statistically significant at the 5% significance level. This is indicative evidence of *disproportionality* among the Black/African American and White/Non-Hispanic subgroup interactions within the City of Aurora.[29] Furthermore, the observed difference in the Black/African American subgroup IPR being higher than the White/Non-Hispanic subgroup IPR implies that over the year 2019, Black/African American subjects experienced a higher number of interactions with the police relative to their corresponding proportion of an aggregate subpopulation comprised solely of Black/African American subjects and White/Non-Hispanic subjects.

Figure 1.A suggests statistically significant evidence of disproportionality in the IPRs for three other racial/ethnic subgroups (relative to the respective population composition(s) across subjects): (i) American Indian/Alaska Native, (ii) Asian, and (iii) Hispanic or Latino. The IPR for the Hispanic or Latino subjects is 10.7% (compared to corresponding ratio of 9.8% for White/Non-Hispanic subjects), and the chi-squared test for Hispanic or Latino subjects relative to White/Non-Hispanic subjects results in a chi-squared statistic of 44.3 and an associated p-value of 0.00, indicating that Hispanic or Latino subjects experienced higher amounts of interactions with the police relative to their corresponding proportion of an aggregate subpopulation comprised solely of Hispanic or Latino subjects and White/Non-Hispanic subjects, though Black/African American subjects appear to have experienced even higher interactions per subgroup population relative to Hispanic or Latino subjects or White/Non-Hispanic subjects. Figure 1.A shows that the chi-squared tests associated with American Indian/Alaska Native subjects and Asian subjects are also associated with statistics and p-values that indicate disproportionality relative to White/Non-Hispanic subjects. However, the IPRs for subjects of both subgroups (2.8% and 4.0%, respectively) being *lower* than the corresponding IPR for White/Non-Hispanic subjects (9.8%), implies that American Indian/Alaska Native subjects and Asian subjects each respectively exhibited *fewer* interactions per their respective populations as compared to the White/Non-Hispanic subjects in 2019.

[28] A comparably similar analysis conducted examining the total number of incidents across specified subgroups during the period January 2018 – February 2021 yielded similar results. *See* Figures 1.D – 1.F.

[29] In other words, the chi-squared test suggests that the "ratio between the percentage of persons in a particular racial or ethnic group at a particular decision point or experiencing an event" – in this case, an interaction with the police – exhibits a statistically significantly difference relativev to the "percentage of the same racial or ethnic group in the overall population." See Footnote 4.

Finally, Figure 1.A reports the result of a chi-squared test that compares subjects from a "Non-White" subgroup (i.e., subjects from all subgroups exclusive of subjects from the White/Non-Hispanic subgroup) in comparison to subjects from the White/Non-Hispanic subgroup. This particular test provides indicative evidence of disproportionality across subjects from the Non-White subgroup versus subjects from the White/Non-Hispanic subgroup. Furthermore, the observed difference in the Non-White subgroup IPR being higher than the White/Non-Hispanic subgroup IPR implies that over the year 2019, Non-White subjects experienced higher amounts of interactions with the police relative to their corresponding proportion to an aggregate subpopulation comprised solely of Non-White subjects and White/Non-Hispanic subjects.

Figure 1.B reports the results of chi-squared tests that examine, for subjects in each subgroup, the ratio of arrests to population ("APR"). The APR results indicate statistically significant evidence that arrests occurred *more* frequently for the subjects in the Black/African American, Hispanic or Latino, and Non-White subgroups relative to the corresponding proportions of their aggregate subpopulations with White/Non-Hispanic subjects, and *less* frequently for the American Indian/Alaska Native and Asian subgroups relative to the corresponding proportions of their respective aggregate subpopulations with White/Non-Hispanic subjects, in 2019.

Figure 1.C reports the results of chi-squared tests that examine the number of UoF Incidents relative to population for each subgroup ("UPR"). The UPR results indicate statistically significant evidence that UoF Incidents occurred *more* frequently for subjects from the Black/African American subgroup relative to their corresponding proportion of an aggregate subpopulation solely comprised of White/Non-Hispanic subjects and Black/African American subjects, and *less* frequently for the Asian subgroup relative to the corresponding proportion of an aggregate subpopulation solely comprised of Asian subjects and White/Non-Hispanic subjects.

The subsequent results present indicative evidence of *disparate* ratios of arrests (relative to interactions) and UoF incidents (relative to arrests) by racial/ethnic groups relative to the White/Non-Hispanic group in the City of Aurora. This analysis of disparity considers whether the respective number of arrests relative to interactions and/or UoF incidents relative to arrests across subjects from different racial/ethnic subgroups is significantly different in instances comprised of comparatively similar occurrences.[30] [31] [32]

Figures 2.A, 2.B and 2.C present a chi-squared test comparing the ratios of (i) arrests *given* interactions ("AIR"), (ii) uses of force *given* arrests ("UAR") and (iii) UoF given interactions ("UAI") during the year 2019. For both of these figures, the statistically significant results persist for the Black/African American subjects and Non-White subjects, which are indicative evidence of disparity for (i), (ii) and (iii), aside from the indicative evidence of disproportionality presented in Figures 1.A – 1.C.

Figures 2.D, 2.E and 2.F present the same analysis presented in Figures 2.A and 2.B, but over the entire period of data, 2018 – February 2021. For both of these figures, the statistically significant results for the Black/African American subjects and Non-White subjects persist.

---

[30] For the purposes of analysis at this stage, these particular tests of disparity across occurrences are specified to treat observed interactions as being comparatively similar, and all arrests as being comparatively similar (without differentiating by call types, case types, age, gender, etc.). In subsequent analysis in this technical appendix, tests of disparity further differentiate between case types and other comparable distinctions across occurrences.
[31] As previously indicated (*see* Footnote 4), for the purposes of this statistical analysis, terms regarding disproportional and disparate occurrence(s) are neither adopted nor employed in a legal context, and nothing in this appendix constitutes a legal analysis or asserts a legal conclusion. this analysis adopts the definition of "disparate" as commonly employed by the criminal and social justice literature. This analysis neither adopts the use of, nor addresses inferences associated with the differentiated terminology (i.e., "treatment" and "impact"), as is periodically adopted by the field of labor economics and public policy for discussions regarding empirical and theoretical analysis of the legal concepts related to disparity. See Rodgers, W. M. ed. (2006) "Handbook on the Economics of Discrimination". By way of further illustration, according to one EEOC reference, *disparate treatment* occurs when an employer treats some individuals less favorably than other similarly situated individuals because of their race, color, religion, sex, or national origin. *Disparate impact* can result from neutral employment policies and practices which are applied evenhandedly to all employees and applicants, but which have the effect of disproportionately excluding women and/or minorities. *See*, *e.g.,* https://www.eeoc.gov/laws/guidance/cm-604-theories-discrimination (retrieved on or around June 10, 2021).
[32] Although these chi-squared tests are tests of the variance of relative proportions across ratios between categories normalized by corresponding sample populations or subpopulations as denominators (subject to the particular test specification), outside the scope of the current analysis are both *saturated* and *baseline* model implementations of structural equation model (SEM) specifications, which entail the simultaneous estimation of means, variances, and covariances in order to infer latent variables underlying systems of equations. SEMs are statistical methods typically applied to causal attribution analysis as a means of attributing observations to unobservables primarily in the social sciences. Entailing joint hypothesis tests of systems of simultaneous equations being highly dependent on assumptions underlying a particular structural specification, this field of inquiry although extensive remains highly contested and subject to extensive controversy — i.e., susceptibility to attribution, confirmation and reinforcement biases — generally related to intrinsic model misspecification errors (e.g., *model inconsistency*, instability) inherent to *undetermined* vs *overdetermined systems of equations* with either no solution or no unique solution. See Tarka, Piotr (2017) "An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences"; Bollen, K.A., J. Pearl (2013) "Eight myths about causality and structural equation models survey the history of the causal interpretation and sources of associated confusion and controversy"; Pearl, J. (2000) "Causality: models, reasoning and Inference discusses both parametric and nonparametric extensions of SEMs particularly in the context of causal and counter factual interpretations". Also see "Kline, R.B. (2016) Principles and practice of structural equation modeling"; Kaplan, D. (2009) Structural equation modeling: foundations and extensions"; Bollen, K.A. (1989) "Structural equations with latent variables; Duncan, O.D. (1975) Introduction to structural equation models".

These particular tests of disparity across occurrences as described above are specified to treat observed interactions as being comparatively similar, and all arrests as being comparatively similar (without differentiating by call types, case types, age, gender, etc.). The analysis in the subsequent section describes tests of disparity which further differentiate between case types and other comparable distinctions across occurrences.

Further Analysis Conducted Employing Additional Categorical Indicator Variables

As discussed *supra*, UoF incidents are classified by the APD in "tiers," where Tier 1 is the lowest level of force, Tier 2 is the middle level, and Tier 3 is the highest level.[33] APD data appear to agglomerate Tier 2 and Tier 3 UoF as one level of classification (identified as "Tier 2/3" in the data) over the period. Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across all tiers of UoF.

Figures 3.A, 3.B present chi-squared tests that examine differences in UAR and UAI by racial/ethnic group for each UoF tier. For both these figures, the statistically significant results pertaining to UAR and UAI disparity for Black/African American subjects and Non-White subjects persists across each category of UoF tier observed in the data.

The data comprise three "Districts" (District 1, District 2 and District 3) which represent geographic divisions within the City of Aurora. Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across all APD districts pooled together. Figures 4.A, 4.B and 4.C present chi-squared tests that examine differences between AIR, UAR and UAI by racial/ethnic group within each district. For both of these figures, the statistically significant results pertaining to AIR, UAR and UAI disparity for Black/African American subjects and Non-White subjects persist within each district observed in the data.

The data indicate assigned levels of case and class "severity" (i.e., "misdemeanor" or "felony") to each relevant record. Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across all levels of case/class severity pooled together. Figures 5.A, 5.B and 5.C present chi-squared tests that examine differences between AIR, UAR and UAI by racial/ethnic group for each assigned level of severity. For both of these figures, the statistically significant results pertaining to AIR, UAR and UAI disparity for Black/African American subjects and Non-White subjects persists within each assigned level of severity observed in the data.

As discussed earlier, starting in January 2018, the data indicate "Case Types" to each relevant record. Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across Case Types. Figures 6.A, 6.B and 6.C. present chi-squared tests that examine differences between AIR, UAR and UAI for the top three most observed case types in the data from January 2018 – February 2021. For these three figures, the statistically significant results pertaining to AIR, UAR

[33] APD Policies: Directives Manual Chapter 5 Weapons and Use of Force (Section 5.4.7 Reporting and Investigating the Use of Tools Weapons and Physical Force https://public.powerdms.com/AURORAPD/tree/documents/107).

and UAI disparity for Black/African American subjects and Non-White subjects persist for the Case Type "SUSPO" (i.e., suspicious occurrence), but do not persist for Case Types "DISTR-", "DOMES-" (corresponding to disturbance/noise complaint and domestic dispute, respectively).

The data indicate an observed "sex" variable (comprised of the values "Male" and "Female") for each subject record. Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across both these values as recorded in the data. Figures 7.A, 7.B and 7.C present chi-squared tests that examine differences between AIR, UAR and UAI by "Male" and "Female" gender categories. For these two figures, the statistically significant results pertaining to AIR, UAR and UAI disparity for Black/African American subjects and Non-White subjects persists for both "Male" and "Female" gender categories.

The data indicate the date of birth of a subject and the date of an interaction/arrest/UoF, from which the age of a subject is calculated.[34] Figures 2.A – 2.F present indicative evidence of disparity in AIR, UAR and UAI across all subject ages. Figures 8.A, 8.B and 8.C present chi-squared tests that examine the differences between AIR, UAR and UAI for aggregate subpopulations comprised solely of Black/African American subjects and White/Non-Hispanic subjects, by subject age group. For these two figures, the statistically significant results pertaining to Black/African American subjects persist for the age groups 18-21, 22-29, and 30-49, but do not persist for the age groups 50-64 or 65-98.

As highlighted throughout this section, the persistent indicative pattern of disproportionality and disparity across race/ethnicity subgroups exhibited by the chi-squared test results is compelling empirical evidence for further inquiry.[35] [36] Subsequent analysis involving estimates and inferences

---

[34] The data also contain a field entitled "age group" with four distinct values. This analysis computes the age of a subject as the difference between the date of an interaction and the subject's date of birth, and subsequently creates a different age group variable by assigning the age of the subject to one of the following categories: 18-21; 22-29; 30-49; 50-64; 65-98; and 98+.

[35] This indicative scope of analysis has been conducted by employing available data with the implicit assumption that these data are the actual population of police interaction for the relevant data period without further performing an analysis of sampling error, representativeness, or the statistical leverage of influential observations. During the course of the analysis, access was solely granted to the RMS and AIM data repositories maintained by Aurora IT for APD. No analysis of sampling error and representativeness, including sampling and resampling methods (e.g., bootstrap, jackknife, capture-recapture) has been conducted.

[36] As previously indicated (see Footnote 9), generally-accepted adjustments applied to the fundamental chi-squared statistic are customary in practice to address the approximation of the discrete binomial probabilities using a continuous distribution (which can introduce some error thereby necessitating the adjustment). The R software codebase (Version 4.1.0) applies a *Yates continuity correction* to chi-squared tests of homegeneity for 2x2 matrices as follows:

$$\Sigma_i \frac{(|o_i - e_i| - 0.5)^2}{e_i}$$

It is generally acknowledged in the statistics literature that this adjustment to the numerator of the chi-squared equation typically introduces *Type-II error* (i.e., a false negative, the failure to reject the *null hypothesis*), thereby indicating this adjustment to be conservative. Furthermore, empirical inspection exhibits *de minimis* effects on the results. See Yates, F (1934). "Contingency table involving small numbers and the χ2 test". Supplement to the Journal of the Royal Statistical Society 1(2): 217–235. Also see Agresti, Alan; Coull, Brent A. (1998). "Approximate is better than 'exact' for interval estimation of binomial proportions". The American Statistician. 52 (2): 119–126. Brown, Lawrence D.; Cai, T. Tony; DasGupta, Anirban (2001). "Interval Estimation for a Binomial regarding other types of empirically observable differences[37] in interactions, arrests or UoFs will be subject to further analysis entailing the use of substantially expanded datasets to be curated in accordance with generally-accepted methodological practices for rigorous statistical analysis, as generally highlighted by the methodological references presented in Section VI.[38]

V. Supplemental Empirical Inquiry[39]

Implementation of a Generalized Expansion of the Pearson Chi-squared Test

The empirical results discussed above in Sections III and IV have employed chi-squared tests of homogeneity using a 2x2 contingency (i.e., cross tabulation or crosstab) table structure which displays the occurrence distribution of variables,[40] where each racial/ethnic subgroup is tested with the White/Non-Hispanic subgroup to determine whether ratios of Interactions, Arrests and UoF

Proportion". Statistical Science. 16 (2): 101–133. Devore, Jay L., Probability and Statistics for Engineering and the Sciences, Fourth Edition, Duxbury Press, 1995. Feller, W., On the normal approximation to the binomial distribution, The Annals of Mathematical Statistics, Vol. 16 No. 4, Page 319-329, 1945. Thulin, Måns (2014) "The cost of using exact confidence intervals for a binomial proportion". Electronic Journal of Statistics. 8 (1): 817–840.

[37] Such observed differences being relative differences in e.g., *frequencies*, *rates*, *probabilities* (or alternatively *odds* or *log odds* of occurrences), or the *propensities* for types of interactions, arrests or UoFs, each corresponding to the respective reference population (which is subject to the specification of the particular test).

[38] The research literature on differential police interactions and testing for discriminatory behaviors is extensive and remains active. Areas actively being researched include controlling for the allocation of police resources and the distribution of crime types relative to population demographics. By way of illustration, Pierson, Simoiu, Overgoor et al. (2020) "A Large Scale Analysis of Racial Disparities in Police Stops Across the United States" and Simoiu, Corbett-Davies and Goel (2017) "The Problem of Infra-Marginality in Outcome Tests" highlight statistical testing of biased decision making and the challenges to rigorous assessment, predominantly due to well-known limitations with the two most common statistical tests for discrimination, i.e., *benchmarking* and *outcome testing*, as well as other methodological attempts to address limitations associated with benchmarking and outcome testing. More recent statistical tests for discrimination (e.g., threshold testing) attempt to mitigate infra-marginality by jointly estimating *decision thresholds* and *risk distributions*. As described by Ayres (2002), one limitation of benchmarking (as a specific instance of *omitted variable bias*), in the literature is referred to as the *qualified pool* or *denominator* problem. In order to address this bias, Becker (1957, 1993) proposed the outcome test, which is based not on the rate at which decisions are made, but on the success rate of those decisions. It should be noted that outcome tests despite being widely adopted across a diverse range of domains of analysis, in particular that of policing, e.g., see Goel, Rao and Shroff (2016, 2017), Ayres (2002), Knowles, Persico and Todd (2001), tend to be imperfect barometers of bias.

[39] The proposed supplemental analyses and any preliminary exploratory results are closely related to the motivations for the subsequent section on *Applicable Empirical Principles and Methodological Practices for Data Acquisition, Curation, and Provenance and Analysis*.

[40] The *Fisher exact test* – which relies upon the hypergeometric distribution, assuming no association between the observed occurrence(s) and category of interest (e.g., race/ethnicity) – is an exact test of equivalence between two proportions. The Fisher exact test (and its chi-squared approximation) assumes *fixed margins* of the four-fold 2x2 table, i.e., all fixed margin totals for which the number of possible outcomes that might have occurred keeping numbers of subpopulation counts and populations counts constant (i.e., the number of possible permutations of occurrences no greater than the observed number of occurrences) is divided by the total number of all possible permutations of occurrences and non-occurrences irrespective of race/ethnicity. However, if the objective is to prospectively emulate the behavior of the process it is inappropriate to assume fixed margins. By way of further background regarding cross-tabular data structure, see e.g., Gray, Bosworth, Layman and Pirahesh (1996) "Data Cube: Aggregation Operator Generalizing Group-By, Cross-Tab and Sub-Totals".

Incidents are heterogenous in terms of disproportionality and/or disparity. A more generalized expansion of the chi-squared test – often commonly referred to as the *generalized Pearson chi-squared test* – to $m > 2$ categories can be used to test whether the same ratios are homogenous across *all* racial/ethnic subgroups, including the White/Non-Hispanic subgroup.[41] Effectively, these specifications of the Pearson chi-squared test compare the aforementioned observed ratios to the expected ratios under a null hypothesis of a *uniform distribution* of incidents. If these ratios are equally distributed (in accordance with the properties of the uniform distribution), the distribution for the ratios of counts across subgroups should be closely approximated by the denominator of the ratio for each subgroup (i.e., population in the tests of disproportionality, and incidents and/or arrests in the tests of disparity).[42] [43]

As with the chi-squared tests of homogeneity, associated p-values are calculated for a Pearson chi-squared statistic. A p-value less than the commonly accepted critical value threshold level of $\alpha = 0.05$ suggests that the ratios for distributions of counts across each subgroup is *not* uniform (i.e., *heterogenous or non-homogenous*).[44] This expansion of the chi-squared specification generalizes the test to more than a 2x2 matrix of subgroups, although the results are not appropriate for direct comparisons across ratios associated with specified subgroups.[45]

Figures 1.G and 1.H present the Pearson chi-squared test of disproportionality for IPR, APR and UPR, while Figures 2.G - 2.L present the Pearson chi-squared test of disparity for AIR, UAR and

---

[41] As an alternative to Pearson chi-squared tests, loglinear models (related to logistic regression i.e., logit models) estimate co-occurrence between more than two categorical variables for both hypothesis testing and model specification with the objective to find the most parsimonious (i.e., least complex) model, in order to explain the variance in observed occurrences. The likelihood ratio, applicable to confidence and significance measures for statistical hypothesis testing generalizes to multi-way contingency tables and non-count distributions.

[42] It should be noted that being dependent upon the sample size, for instances in which chi-squared statistics are not suitable measures for the *within-table degree of association*, the *phi-squared* (i.e., *mean squared contingency*) measure of association commonly adopted for $2x2$ tables, due to certain convenient properties does not generalize to higher dimensional contingency tables ($> 2x2$ tables), e.g., does not sum to 1 even when the range of the chi-squared remains unbounded and the categorical attributes are independent. For the general rxc case, *phi-squared* ranges between zero and $r - 1$ and $c - 1$ with the upper limit attainable solely when the matrix is square i.e., the system of equations is determined ($r = c$). Another measure of association between attributes on two sizes of a $2x2$ contingency table is the *tau-B* (τ_B) *statistic*, an error reduction measure, commonly defined as the percentage decrease in the expected number of classification errors (due to knowledge of the *conditioning factor*), often standardized by dividing by the lesser of $r - 1$ and $c - 1$.

[43] Like the odds ratio, the τ_B measure exhibits the same result, both retrospectively and prospectively, although, unlike the odds ratio, solely for $2x2$ tables. It should be noted that although outside the scope for this indicative analysis at this stage, further inquiry regarding estimates of odds, log odds or probabilities may necessitate sampling, in order to compute confidence intervals or prediction intervals.

[44] Although relative proportions as ratios of observed counts are the subject of study, these are not to be confused with *ratio estimators* i.e., each a ratio of means of two random variables, nor does this analysis directly employ *ratio distributions* (each a distribution of the ratio of two random variables, e.g. Cauchy), per se.

[45] For comparisons across multiple proportions (>2x2 tables), hypotheses regarding multinomial distributions of cell frequencies employ chi-squared tests for goodness of fit in higher dimensional contingency tables (2xk tables), where the source distribution of the data is unspecified. With respect to pooled marginal proportions (the sum of marginal proportions times the corresponding sample size) for the ratio of observed relative to expected cell frequencies, the cell frequencies for each sample obtained with expected (population-level) cell frequencies estimated is proportional to the marginal frequencies computed using the method of *maximum likelihood*.

UAI, both for all racial/ethnic subgroups. Figures 1.G and 1.H present statistical evidence that the IPR, APR and UPR are disproportionate across all racial/ethnic subgroups. Figures 2.G - 2.L present statistical evidence that the AIR, UAR and UAI are disparate across all racial/ethnic subgroups.

Implementation of Logit Regression

A binomial logistic ("logit") regression is a statistical tool commonly used to model the probability of observing a binary outcome event, i.e., whether a certain event is observed to occur. Logit regressions can be specified to estimate the association between certain independent variables (e.g., variables pertaining to demographic, geolocational, and/or other specified variable) and the probability of observing a certain event. For a set of $i = 1, 2, \dots, n$ observed binary outcomes and independent variables, a logit regression is specified generally as follows:

$$P(Event\ Occurs) = \alpha + \Sigma_{j=1}^{m} \beta_j X_{i,j} + \epsilon_i$$

Where the $X_{i,j}$ represent m independent variables for each observation i, the β_j represent the coefficients on the m independent variables, the α represents a constant, and the ϵ_i represent error terms for each observation. Logit regressions assume that the error terms ϵ_i follow a logistic distribution. The estimation of a logit model allows for the assessment of statistical significance of the association between independent variables simultaneously.[46] It should be noted that the logit model coefficients do not directly correspond to probabilities or odds, but to *log odds*.[47]

Figure 9.A presents the results of a logit model specification, which regresses observed incident occurrences (i.e., arrest, UoF) as *indicator variables* with respect to categorical indicator variables (also binary [0,1] variables) for race/ethnicity of the test subject. For each test subject, the corresponding categorical binary indicator variable is equal to the value one if the data field in that category associates the respective subject with that specific race/ethnicity, and is equal to zero otherwise.

Figure 9.A, Panel A presents the results of a logit model specification which regresses observed arrest occurrences with respect to race/ethnicity categorical indicator variables. Figure 9.A, Panel B presents a comparably similar specification as in the corresponding panels in Figure 9.A, but regresses whether a UoF incident is observed on race/ethnicity indicators. In both panels, no race/ethnicity indicator for White/Non-Hispanic is included. This allows for the interpretation of the coefficient for each race/ethnicity indicator as the incremental effect of subjects belonging to

[46] In the case that all independent variables are *categorical*, i.e. exhibit values from a set of discrete non-ordered values, a *log-linear* analysis of co-occurrence counts, although also appropriate, is outside the scope for the stage of analysis as discussed in this appendix.

[47] The expected values estimated by the logit coefficients are natural logarithms of the ratio of each probability p to its complement $1 - p$.

a certain racial/ethnic subgroup relative to subjects belonging to the White/Non-Hispanic subgroup.

Both panels in Figure 9.B present comparably similar specifications as in the corresponding panels in Figure 9.A, but include an additional binary indicator variable for the race/ethnicity subgroup “Unknown” and include observations in this race/ethnicity subgroup in the regression sample. The results presented in Figure 9.A persist when including these observations and the additional binary indicator variable.

Both panels in Figure 9.C present comparably similar specifications as in the corresponding panels in Figures 9.A and 9.B, but excludes both binary indicator variables for the race/ethnicity subgroups “Unknown” and “Native Hawaiian/Pacific Islander” from the regression sample, in order to further test sensitivity of the model specification to these omitted variables. The results presented in Figure 9.A persist when excluding these binary indicator variables.

In both panels, the coefficients for the Black/African American, Hispanic or Latino, and American Indian/Alaskan Native race/ethnicity indicators are statistically significant and positive, presenting indicative evidence that subjects belonging to these racial/ethnic subgroups were observed to have relatively higher propensities to experience an arrest or UoF relative to White/Non-Hispanic subjects over the time period January 2018 – February 2021. In contrast, the coefficient for the Asian race/ethnicity indicator is statistically significant and negative in Panel B, presenting indicative evidence that Asian subjects were observed to have relatively lower propensities to experience a UoF incident relative to White/Non-Hispanic subjects over the time period January 2018 – February 2021.

Other Preliminary Exploratory Analysis

Geolocation Data Supplementation. At the request of the Colorado AG team, these preliminary zip code-level tests, conditioned on both median income and race/ethnicity, are specified to indicatively examine the extent to which income effects are confounds for race/ethnicity and whether household income differentials subsume the observed disproportionality by racial/ethnicity across the aforementioned zip codes. As previously indicated, the geospatial distribution of incidents and police resources conditioned on the distribution of the population by race/ethnicity and other demographics (e.g., age, socioeconomic factors) may be necessary for additional inference, but currently outside the scope at this preliminary phase of analysis.[48] [49] [50] Although more extensive geolocation and spatial analysis is outside the scope for the current phase

[48] *Spatial analysis* (i.e., *spatial statistics*) includes an extensive range of formal methods to analyze entities using topological, geometric, topographical or geographic properties, many such methods being still in early development, although with a diverse scope of application at various scales and resolutions. In a somewhat more limited sense, spatial analysis applied to structures at human scale, most notably entails analysis of geographic data, e.g., geolocation (i.e., position, navigation and timing), geospatial (satellite, aerial) imaging, geographic information systems (GIS) mapping, as well as cadastral, cartographic and spatial statistics/econometrics, among methods. Many complex issues in spatial analysis, remaining neither clearly defined nor completely resolved, form the basis for current research, the most fundamental of these being the problem of defining the spatial location of entities being studied. Issues to be addressed in a geospatial analysis include the following: *spatial dependence* (and heterogeneity), *spatial autocorrelation* and *spatial association*, as well as *spatial measurement scale* and *spatial sampling*. Spatial sampling entails specifying the number of locations in a geographic space for reliably measuring phenomena subject to dependency and heterogeneity. See: Berry B. J., F. Horton (1971) Geographic Perspectives on Urban Systems, John Wiley; Berry B.J., K.B. Smith eds. (1972) City Classification Handbook: Methods and Applications, John Wiley; Tucker L.R. (1964) "The extension of Factor Analysis to three-dimensional matrices", in Frederiksen N & H Gulliksen eds, Contributions to Mathematical Psychology, Holt, Rinehart and Winston. Also see: Wang, J.F., T.L. Zhang, B.J. Fu (2016) "A measure of spatial stratified heterogeneity". Ecological Indicators. 67: 250–256; Brunsdon, C., A.S., Fotheringham, M.E. Charlton (1996). "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity". Geographical Analysis. 28 (4): 281–298. Banerjee, S., B.P. Carlin, A.E. Gelfand, Alan E. (2014), Hierarchical Modeling and Analysis for Spatial Data, Second Edition, Monographs on Statistics and Applied Probability (2nd ed.), Chapman and Hall/CRC.

[49] Fundamental adjustments in spatial analysis address, e.g., mathematical properties of the spatial relationships (i.e., geometric, topological, topographical features), graphical properties for presentation of spatial data (e.g., correspondence, scale and resolution related to *tiles* or *rasters* as homogenized spatial units), etc. Administrative (e.g., census), socioeconomic and demographic data which aggregate data into location-based units, i.e., regions, presents a number of statistical issues to be addressed, related to the scale and sampling properties of spatial distributions within such regions (e.g., MAUP). The *modifiable areal unit problem* (MAUP) is an issue for the analysis of spatial data arranged in zones, typically dependent on the particular shape or size of the zones used in the analysis, since spatial analysis and modeling often involves aggregate spatial units such as census tracts or traffic analysis zones. Such areal units (typically reflecting data collection and/or modeling convenience) often tend to be somewhat arbitrary, and subject to modification, inherently heterogeneous and present artifacts related to the degree of spatial aggregation or the placement of boundaries. Since results derived from an analysis of these zones depends directly on the properties of the zones being studied, the aggregation of point data into zones of different shapes and sizes can lead to opposite conclusions. The *locational fallacy* refers to error related to a particular spatial characterization chosen for analysis, i.e., location-based spatial characterization, which can be overly simplistic or just wrong. Reducing spatial activities to a single spatial point, i.e., a residential address, often impairs analysis (e.g., epidemiological studies of infectious disease transmission occurrences at work or at school versus at the subject location of residence). Implicitly, spatial characterization might also inherently limit the scope and subject of study, e.g., spatial analysis of crime data typically can only describe specific types of crime incidents which can be described explicitly in location-specific spatial (and often both spatial and temporal) terms – maps of assault, arson or burglary, but not maps of embezzlement – with inherent implications for policing resource allocation and enforcement practices. See Banerjee, S., A.E. Gelfand, A.O. Finley, H. Sang (2008) "Gaussian predictive process models for large spatial datasets" Journal of the Royal Statistical Society, Series B. 70 (4): 825–848; Datta, A, S. Banerjee, A.O., Finley, A.E. Gelfand (2016). "Hierarchical Nearest Neighbor Gaussian Process Models for Large Geostatistical Datasets" Journal of the American Statistical Association, 111 (514): 800–812. More specifically, on a related note, with respect to risks of underlying bias, the application of spatial crime analysis to police resource allocation remains both controversial and widely disputed, see e.g., Akpinar, N., M. De-Artaga and A. Chouldechova (2021), "The Effect of Differential Victim Crime Reporting on Predictive Policing Systems"; Fitzpatrick, D.J., W.L. Gorr and D.B. Neill (2019), "Keeping Score: Predictive Analytics in Policing".

[50] Urban and regional analysis of large tables of spatial data from censuses and surveys typically apply factor analysis methods to transform correlated variables of the census into fewer independent factors (e.g., *principal components*, the eigenvectors of the data correlation matrix weighted by the inverse of their eigenvalues). The choice of spatial distance metric is critical, e.g., the Euclidean metric (principal component analysis), the chi-

of analysis, zip code level summary statistics for socioeconomic (median income) and demographics (median age) based on census data from the American Community Survey ("ACS") 5-Year estimates and indicative results are presented in Figures 11.A - 13.C.[51] [52]

Conditional on unobserved demographic characteristics, the distribution of racial/ethnic subgroup populations with respect to police interactions/arrests/uses of force may be heterogenous. For example, one might observe higher frequencies of police interactions/arrests/uses of force in populations with lower levels of income relative to populations with higher levels of income. Furthermore, certain race/ethnicity subgroup populations may be overrepresented in populations with lower levels of income relative to populations with higher levels of income. In this case, statistical tests of disproportionality and disparity for the whole population of Aurora may not be applicable to subpopulations of Aurora with differing income levels and race/ethnic compositions.[53]

As an indicative test exploring whether the statistical significance of the aforementioned results persist across income levels, quartiles for each zip code observed in the dataset were calculated for median household income reported in the 2019 ACS 5-Year Estimates. Figure 11.A exhibits the range of median household income and the zip codes within each quartile. Income Quartile 1 represents the zip codes with the lowest median incomes, while Income Quartile 4 represents the zip codes with the highest median incomes. Figure 11.B exhibits the respective populations obtained from the 2019 ACS 5-Year Estimates for each Income Quartile. As median household income increases from Income Quartile 1 to Income Quartile 4, the relative population of the Black/African American subgroup decreases from 14.6% to 4.8%, while the relative population of the White/Non-Hispanic subgroup increases from 30.6% to 76.9%.

Figures 12.A, 12.B and 12.C present tests of disproportionality for the year 2019 by Income Quartile for IPR, APR and UPR, respectively. Each of these figures exhibits persistence of statistically significant results for the Black/African American subgroup and the Non-White subgroup, with the exception of the UPR test for Income Quartile 2 and Income Quartile 4 of the

squared distance (*correspondence analysis*) or the generalized Mahalanobis distance (*discriminant analysis*) being among the more widely used. See Rummel, R.J. (1970) "Applied Factor Analysis" Evanston, ILL: Northwestern University Press.

[51] While the RMS database provided some geolocation data corresponding to the locations where certain interactions occurred, some observations exhibit an interaction address without a corresponding zip code, which requires matching zip codes with the exhibited address fields, GPS locations, or latitudes-longitudes. Although outside the current scope of the analysis conducted, a supplemental analysis is proposed to be conducted at the zip code level, employing the ArcGIS geocoding tool vis-à-vis Version 4.1.0 R Studio (an Integrated Development Environment for R, a well-established and widely-adopted programming language for statistical computing) to map zip codes to incident data for instances where the address for an incident is available.

[52] Further demographic inference would entail fuller income and age distributions (e.g., quintiles, deciles) for spatial analysis at the zip code level. See Wood (2020) which references two general data sources (FOIA and litigation versus private data agreements with police departments), as well as the functionality of cartographic data visualization exhibited by Citizens Police Data Project (CPDP), as further described in the subsequent section.

[53] For zip code level tests in this analysis, zip codes are employed to assign demographic and income distribution for the zip codes comprising the city of Aurora. It is acknowledged that certain zip codes within the Aurora city limits might also include geographic areas that are contiguous to the Aurora city borders, e.g., greater Denver.

Non-White subgroup, as well as the exception of IPR and APR for the Non-White subgroup for Income Quartile 1.

Figures 13.A, 13.B and 13.C present a comparably similar analysis over the entire sample period January 2018 – February 2021, for which the statistically significant results persist for Black/African American subgroup and the Non-White subgroup, with the exception of the Non-White subgroup for IPR and APR in Income Quartile 1.

The persistence of statistical significance exhibited by the net effect for race/ethnicity, particularly for the Black/African American category across zip codes sorted into median household income quartiles implies that despite any underlying correlation between race/ethnicity and income, the disproportionality by race/ethnicity is not subsumed by differences in income.

Discussion Regarding Prospective Analysis of Repeat Interactions/Arrests/UoF Incidents. Notably, the chi-squared analyses above do not assume *unique interactions* (e.g., per Figures 1.A, 1.B and 1.C that 22.3%, 10.0%, and 0.3% of the Black/African American population of Aurora each had interactions, were arrested, and/or were subjected to UoF incidents). For example, the 13,570 interactions between APD officers and Black/African American subjects, or the 16,092 interactions between APD officers and White/Non-Hispanic subjects, noted in Figure 1.A may include repeated interactions with a subset of Black/African American or White/Non-Hispanic subjects (i.e., the total number of interactions with each subgroup occurs with fewer than 13,570 Black/African American subjects or 16,092 White/Non-Hispanic subjects).[54]

Chart 1 shows the cumulative proportion of interactions (along the y-axis) indexed by the number of subjects with zero, one, two… through up to twenty-eight ***prior recorded*** interactions as of 2019 (along the x-axis), for White/Non-Hispanic and for Black/African American racial/ethnic groups, respectively. Notwithstanding yet to be explored or specified systematic conditions (e.g., record keeping conventions, regime changes related to modified practices and procedures) which will require more comprehensively rigorous data acquisition, curation and provenance, however limited (e.g., censored and truncated), these preliminary exploratory results apparently indicate that during the period January 2018 to February 2021, the data corresponding to Black/African American subjects in the City in Aurora did exhibit higher occurrences of "recurring interactions" with APD officers than White/Non-Hispanic subjects.[55]

Figures 10.A and 10.B present comparably similar analyses to Figures 1.A and 1.D, respectively, but examine the ratios of *unique* Interactions to Population ("uIPR") in 2019 for each racial/ethnic

[54] See discussion in Section VI, and in particular, Footnotes 68 and 77 regarding causal inference modeling applied to addressing racial factors in multi-stage police-civilian interactions (e.g., *differential policing* or *selective enforcement*) and the interrelated role of procedural decisions with respect to sample selection issues (e.g., censoring, truncation, survivor bias and omission bias) in administrative records and the risk of statistical bias.
[55] *See* Chart 1 as well as footnotes regarding the procedural discretion and data contamination related to *differential policing* or *selective enforcement*.

subgroup.[56] For these two figures, the statistically significant results pertaining to Black/African American subjects and Non-White subjects persist. Similarly, Figures 10.C and 10.D present comparably similar analyses to 1.B and 1.E, but examine the ratios of unique Arrests to Population ("uAPR") in 2019 and January 2018 – February 2021; and Figures 10.E and 10.F present comparably similar analyses to 1.C and 1.F, but examine the ratios of unique Uses of Force to Population ("uUPR") in 2019 and January 2018 – February 2021. For these figures, the statistically significant results pertaining to Black/African American subjects and Non-White subjects persist.

These indicative empirical test results yield inferences generally consistent with inferences from analyses conducted by other researchers across numerous police jurisdictions, as cited throughout this appendix (particularly in this and the subsequent section).[57]

In contrast to the more extensive empirical analysis conducted, which implemented chi-squared tests of homogeneity presented in the previous section, in this section the exploratory supplemental results which have been presented motivate and illustrate supplemental indicative empirical results of relative differences in estimated *frequencies*, *rates*, *probabilities* – or alternatively *odds* or *log odds* – of occurrences, or *propensities* of interactions, arrests or UoFs, as highlighted by the references cited in the subsequent section. These indicative results yield inferences generally consistent with results subsequently highlighted in Section VI.

VI. Applicable Empirical Principles and Methodological Practices for Data Collection, Acquisition, Curation, Lineage, Provenance and Analytics

The applicable types of model specifications and corresponding empirical results related to disparity and disproportionality exhibited by police activity and interactions are primarily dependent upon the granularity, comparability, consistency and completeness of available data. What is readily apparent from reviewing a cross-section of the applicable databases maintained, datasets employed and empirical analyses conducted is the emphasis on reasonably sufficient volumes and adequately fine-grained granularity of data necessary for reliably robust monitoring, supervision and analytics, as well as the necessity for consistent and disciplined governance of

[56] "Unique" interactions are determined by the variable "*pin*" (i.e., personal identification number). For a pin associated with more than one observed interaction over 2019 (Figure 10.A) or January 2018 – February 2021 (Figure 10.B), any observed interactions after the first observed interactions are excised. For a pin with only one observed interaction over 2019 (Figure 10.A) or January 2018 – February 2021 (Figure 10.B), no observed interactions are removed.

[57] See e.g., Barsamian Kahn, K., Goff. P.A., Lee, J.K. and Motamed, D. (2016) "Protecting Whiteness: White Phenotypic Racial Stereotypicality Reduces Police Use of Force" Social Psychological and Personality Science Journal (Sage); Chicago Police Accountability Task Force (2016) "Recommendations for Reform"; Ross, C.T. (2015) "A Multi-Level Bayesian Analysis of Racial Bias in Police Shootings at the County-Level in the United States," PloS One pp. 2011–2014; US Attorney's Office Western District of Washington and US Department of Justice Civil Rights Division (2011) "Investigation of the Seattle Police Department"; US Department of Justice Civil Rights Division (2015) "Investigation of the Ferguson Police Department"; US Department of Justice Civil Rights Division (2011) "United States' Investigation of the Maricopa County Sheriffs Office".

data collection, acquisition, hygiene, curation, provenance, and other stewardship practices for the relevant demographic, procedural, administrative and supervisory data.

The reliable implementation of appropriate model specifications employed in the literature for estimating differences in relative frequencies or rates of occurrence (e.g., *Poisson regressions*,[58] *negative binomial regressions*,[59] *multinomial regressions*[60]), *associations* or interrelationships between co-occurrences (e.g. *loglinear regressions*[61]), or for addressing data imbalances, sample selection biases,[62] censored[63] or truncated data (e.g., *recurring events*,[64] *proportional hazards*[65] and *competing risk* models[66]), or other omitted variable biases (e.g. *propensity score weighted*

[58] In statistics, Poisson regression is a generalized linear model for regression analysis of count data and contingency tables. Poisson regression assumes the response variable has a Poisson distribution, and assumes the logarithm of its expected value can be modeled by a linear combination of unknown parameters. In more concise terms, Poisson regression models are generalized linear models with the logarithm as the (canonical) link function, and the Poisson distribution function as the assumed probability distribution of the response.

[59] Negative binomial regression is a widely-adopted generalization of Poisson regression which relaxes the restrictive assumption that the variance is equal to the mean as specified in the Poisson regression model. The most widely implemented negative binomial regression model, commonly known as *NB2*, is based on the Poisson-gamma mixture distribution, which models the Poisson heterogeneity with a gamma distribution. Since a characteristic of the Poisson distribution is that its mean is equal to its variance, in certain circumstances, the empirically observed variance might greater than the observed mean, referred as *overdispersion*. Two common reasons for overdispersion might be either the omission of relevant explanatory variables or dependent observations. Under certain circumstances, the problem of overdispersion can be solved by using *quasi-likelihood* estimation or a negative binomial distribution.

[60] Multinomial logistic regression, a classification method that generalizes logistic regression to multiclass problems, (i.e., with more than two possible discrete outcomes) to estimate probabilities for different outcomes of a categorical dependent variable. It is also possible to formulate multinomial logistic regression as a latent variable model, following the two-way latent variable model described for binary logistic regression. This formulation is common in the theory of discrete choice models, and makes it easier to compare multinomial logistic regression to the related multinomial *probit* model, as well as to extension to more complex model specifications.

[61] A Poisson regression model is sometimes also referred to in terms of a log-linear model, especially when used to model co-occurrences across contingency tables.

[62] The statistical term *sample selection bias* generally refers to substantive inconsistency or incongruity between a test sample and the subject population (or subpopulation) being analyzed. *Sampling bias* is commonly described as a limit in the generalizability of results related to the inherent sensitivity of empirical results related to systematic discrepancies in a test sample relative to the subject population (or subpopulation), which occurs when a sample is imbalanced (i.e., some members of the subject population are systematically more likely to be selected in a test sample than others). Sampling bias tends to occur within samples when certain underlying variables systematically under-represented or over-represented with respect to the actual distribution of those variables. See Cochran, W.G. (1977) "Sampling Techniques" Wiley. On a separate but related note, further analysis entails supplementing fundamental analysis of sampling error and representativeness, with sampling and resampling methods (e.g., bootstrap, jackknife, capture-recapture), established statistical methodologies also widely accepted in administrative and judicial proceedings.

[63] *Censoring* is a type of missing data problem in which time to event is not observed for reasons such as termination of study before a subset of those subjects to have exhibited the occurrence of interest is omitted from the dataset prior to exhibiting an occurrence.

[64] Recurrent event analysis (i.e., the modelling of *time-to-event* data) refers to *recurring event* or *repeated event* models which relax the assumption of a singular event occurrence per subject.

[65] *Proportional hazards* regression (i.e., *Cox regression*) is a methodology for analyzing the effect of several variables upon the time to the occurrence of a specified event.

[66] *Competing risks* regressions focus on the cumulative incidence function, which indicates the probability of a specified event occurring before a specified time horizon.

regressions[67]) impose requirements on the dataset being employed. In addition to sound practices and principles of data acquisition, curation and provenance exercised consistently with suitable rigor,[68] the sequence of procedures applied to pre- and post- processing of data are critical to reliability of results.[69]

Illustrative Highlights of Applicable Fundamental Empirical Principles and Foundational Methodological Practices for Analytical Robustness

The following references are presented to highlight applicable fundamental empirical principles and to illustrate foundational methodological practices related to prospective analytics for further inquiry (as described in the previous section), subject to established guidelines for *interpretability*, *repeatability*, *reproducibility* and *replicability* of results.[70]

Analysis of Complaint and Roster Data. **Wood et al. (2019)** implements a *Bayesian negative binomial model* (Bürkner 2017) in order to estimate the marginal effects of gender, race and ethnicity, age, and officer tenure – the latter measured as the year of appointment to the Chicago Police Department (CPD) – on the *rate* (i.e., *frequency of occurrence*) for civilian-facing complaints received per year, in which the outcome variable of the negative binomial model is a count of civilian complaints received for each officer in the period 2010 to 2016. The dataset includes any CPD officer active for at least part of the period 2010 to 2016, including those who did not receive a complaint within this period, i.e., the length of time that each officer was at risk of receiving a complaint as an offset. This enables the estimation of a rate of complaints per year and then to estimate the same model for department-facing complaints. The model gauges how the frequency of complaints differs across officers according to specified attributes (e.g., gender, age, race/ethnicity). With regard to associated social network patterns of officer misconduct, this analysis further infers social network structure employing data on 16,503 complaints and 15,811 police officers over the six-year period in Chicago, and conducts *co-complaint analysis*, which estimates a *Bayesian exponential random graph model* for each *district-level civilian and department co-complaint network*. The analysis by Wood et al. employs two primary sources of data: (i) complaints filed against officers in the CPD from January 2010 to June 2016 and (ii) roster data on all officers who were active in the CPD during this time period. The complaints and roster

[67] See e.g., Freedman and Berk (2008) "Weighting Regressions by Propensity Scores".
[68] See Knox and Mummolo (2020a, 2020b) discussion of research design and causal inference modeling applied to addressing racial factors in multi-stage police-civilian interactions (e.g., discretionary practices involving *differential policing*, *selective enforcement*), Knox, Lowe and Mummolo (2020a) discussion of the interrelated role of procedural decisions with respect to sample selection issues (e.g. censoring, truncation, survivor bias and omission bias) in administrative records resulting in and the risk of *specification error* and statistical bias, as well as Knox, Lowe and Mummolo (2020b) regarding *data contamination* due to post-treatment selection. Also see Georger, Mummolo and Westwood (2020) on data-related impediments to evidence-based policy related to policing.
[69] See e.g., Irvine, J. "Transforming Data into Information: Enabling Detection and Discovery" MITRE #14-2487
[70] See the National Institute of Standards and Technology (NIST) report by Plant and Hanisch (2018) on *reproducibility* and *replicability* in science to the National Academies of Sciences, Engineering and Medicine Committee on Reproducibility and Replicability in Science.

data are part of a larger dataset obtained by the NGO *Invisible Institute* through a series of *Freedom of Information Act* (FOIA) and litigation requests and subsequently made available to the public.

Cartographical Mapping and Spatiotemporal Analysis of Geolocation Data. By way of *cartographical visualization* for *spatiotemporal analysis* of geolocated police activity and resource allocation, subsequent related work in progress as presented by **Wood (2020)** – which references two general data sources (FOIA and litigation vs private data agreements with police departments), as well as the functionality of cartographic data visualization exhibited by *Citizens Police Data Project* (CPDP)[71] – highlights a range of numerous unanswered queries regarding relevant statistics (e.g., police use of firearms, distribution of active police officers deployed across law enforcement agencies, complaints filed, sustained or procedurally interrupted, false arrests, etc.) which carefully curated cartographic analysis and visualization can be employed to address. For illustrative purposes, this ongoing work by Wood overlays population (census) demographics and police force demographics and rapid response times by districts, police beats, historical redlining zones, UoF from tactical response reports (2004-2020) by UoF type and race/ethnicity with varying degrees of granularity. By way of illustration, the ongoing Wood visualization analysis plots relative police deployment (officers per thousand residents) for both rapid response and traffic enforcement corresponding to population proportions of race/ethnicity (percentages).[72]

Analysis of Extensive and Detailed Traffic Stop Data. Focusing on the volume of police activity comprised of more than 20 million traffic stops annually,[73] through an iterative series of public data requests across all 50 states (i.e., requisitions submitted to 50 state patrol agencies and 100 municipal police departments) resulting in aggregated data comprised of 255 million records,[74]

[71] The CPDP dataset includes arrests, assignment, attendance, calls for service, complaints, demographics, hiring, overtime, ranks and promotions, and UoF incidents (see https://cpdp.co).

[72] It is further noted that the CPD maintains a public dashboard that displays information related to CPD use of force incident data. The Use of Force Dashboard is updated on the first of each month, and displays the most frequent type of Use of Force utilized by CPD. Users can sort and filter by year, location, subject demographics and use of force options, and can also view multiple time periods, dating back to 2015. Described as a central component of ongoing efforts by CPD toward police reform and transparency, as facilitated by the consent decree which requires CPD to collect and maintain the data and records necessary to accurately evaluate its use of force practices, and to facilitate transparency and accountability regarding those practices. See https://home.chicagopolice.org/statistics-data/data-dashboards/use-of-force-dashboard/.

[73] In addition to publications by the US Department of Justice (DoJ) Bureau of Justice Statistics (BJS) Davis, Whyde and Langton (2018) URL: http://www.bjs.gov/index.cfm?ty=pbdetail&iid=6406 and Langton and Durose (2013) URL: https://www.bjs.gov/content/pub/pdf/pbtss11.pdf, the analysis also cites Baumgartner, Epp and Shoub (2018) and Epp, Maynard-Woody and Hayder-Markel (2014) as supporting references regarding traffic stop figures. Also cited are statistical analyses conducted on stops, searches, restraints and arrests reported in diverse municipal jurisdictions, Boston, Cincinnati, Nashville, New York City and Oakland.

[74] As explained in detail in the Methods section of Pierson, Simoiu, Overgoor et al. (2020), an aggregate total of 225 million stops, comprised of 221 million stops conducted by 33 state patrol agencies and 34 million stops conducted by 56 municipal police departments "provided in idiosyncratic formats with varying levels of specificity", which required the application of a range of automated and manual procedures to produce the primary dataset for each of the 94,778,505 recorded stops, as follows: (i) extract and normalize the date and time; (ii) the county (for state patrols) or municipal police subdivision (e.g., beat, precinct, zone); (iii) race, gender and age of driver; (iv) reason for stop (e.g. speeding); (v) whether or not a search was conducted; (vi) legal justification for the search (e.g. probable cause, consent); (vii) whether the search discovered contraband; (viii) stop outcome (e.g., citation, arrest). The preparation of the primary dataset was extensive, and for purposes of transparency and reproducibility the raw

Pierson, Simoiu, Overgoor et al. (2020) compiled for analysis a primary dataset detailing approximately 95 million traffic stops conducted by 21 state patrol agencies and 35 municipal police departments over a period of approximately ten years (i.e., 2011-2018). This primary dataset gleaned from the aggregated data to consist of those stops involving drivers classified as "white, black or Hispanic" and the ensuing analyses were restricted to the subset of jurisdictions for which the required corresponding fields were available. **Rivera and Rosenbaum (2020)** employ a data sample of police traffic stops for San Diego (242,000 stops January 1, 2015 to March 31, 2017) and San Francisco (125,000 stops from January 1, 2015 to June 30, 2016) obtained by the Stanford Open Policing Project database to perform outcome and threshold testing based upon race, age and gender of driver, whether or not a search was conducted and whether the search found contraband.[75]

Simoiu, Corbett-Davies and Goel (2017) assembled a comprehensive dataset of 9.5 million traffic stops conducted by the 100 largest police departments in North Carolina between January 2009 and December 2014 that was obtained via a public records request filed with the state. Several variables are recorded for each stop, including the race of the driver (White, Black, Hispanic, Asian, Native American, or "other"), officer department, the reason for the stop, whether a search was conducted, the type of search, the legal basis for that search, and whether contraband (e.g., drugs, alcohol, or weapons) was discovered during the search. In this analysis, "Hispanic" includes any subject whose ethnicity was recorded as Hispanic, irrespective of their recorded race (e.g., it includes both white and black Hispanics). Due to lack of data, the analysis omitted Native American subjects, who comprised fewer than 1% of all recorded stops and also omitted the 1.2% of stops where the driver's race was not recorded or was listed as "other". This reference also presents an extensive survey of other research in order to highlight statistical limitations and mitigants for the most common methods employed (e.g., benchmarking and outcome tests). Among the somewhat more recent research surveyed by this reference are as follows:

- The implementation by **Gelman, Fagan and Kiss (2007)** of a hierarchical Bayesian model to specify a benchmark using neighborhood- and race- specific crime rates (which, as previously explained, should be conditioned upon police resource allocation),
- The adoption by **Ridgeway (2006)** of propensity score-based benchmarking of post-stop police actions as an attempt to match non-white and white drivers using demographics, time, location and purpose of stop, and
- Benchmarks constructed by **Grogger and Ridgeway (2006)** for "veil of darkness" tests of stops at night.[76]

data, standardized data and code employed to process and analyze the records are hosted at (https://openpolicing.stanford.edu).

[75] *Significance* August 2020 Royal Statistical Society (also see https://openpolicing.stanford.edu/data/).

[76] By way of further background, also see the analysis by Dr. Joseph B. Kadane cited in the New Jersey v. Soto opinion of the late Judge Robert E. Francis J.S.C. – 324 NJ Superior 66 734 A.2d 350 (Decided March 5, 1996; Approved for Publication July 15, 1999) – who subsequently authored a seminal publication on the topic: Kadane and Terrin (1997) "Missing Data in the Forensic Context" Royal Statistical Society Series A 160 Part 2, and Terrin and Kadane (1998) "Counting Cars in a Legal Case Involving *Differential Enforcement*" *Chance* American Statistical Association.

Applicability of Corresponding Time-stamped and Geolocated Incident Data with Extensive Officer Deployment, Demographics and Employment Records. **Ba, Knox, Mummolo and Rivera (2021)** highlights general limitations typically common to datasets employed to conduct statistical analysis of policing activity with a particular emphasis on the absence of adequately fine-grained data on officer deployment, and describes their multi-year data acquisition process via open records requests which included the following data fields: officer demographics, language skills, daily shift assignments, career transitions which were associated with time-stamped, geolocated records of corresponding officer stops, arrests and UoF incidents. The dataset is aggressively *pruned* "to maximize analytic validity" of fine-grained data regarding "daily patrol assignments which vary *exogenously* on the basis of fixed rules and pre-assigned rotations." The objective of such data pruning is a resulting dataset comprised of a panel of 2.9 million officer shifts and 1.6 million enforcement events, involving approximately 7,000 officers during the 2012-2015 period for the analysis of cross-sectional differences across comparably equivalent instances related to demographic heterogeneity (e.g., race/ethnicity).[77]

Omitted (Included) Variable Bias in Disparate Impact Testing. **Jung, Corbett-Davies, Shroff and Goel (2019)** implement a *risk-adjusted regression* to address misspecification resulting from either missing variable or confounds irrelevant variables, and test for disparate impact across 2.2 million police stops of pedestrians in New York City.[78]

Systematic Computational Linguistic Analysis of Body Camera Video and Audio Content as Administrative, Supervisory and Procedural Data. Systematic analysis of officer body-worn camera content demonstrates the applicability of language-based analytics to police–civilian interactions and the use of computational linguistic techniques to automatically measure verbal indicators of respect displayed by officers during civilian interactions as well as demonstrating the utility of body camera video and audio content as a documented source of administrative, supervisory and procedural data (versus solely archival evidence). Methodological research by **Voigt, Camp et al. (2017)** notes that despite the rapid proliferation of body-worn cameras, law enforcement has yet to adopt any systematic analysis of the volumes of highly detailed audio and video content generated. This research further demonstrates as a proof-of-principle, the practical application of computational linguistics methods to systematically identify verbal signals of interaction patterns from transcripts, informed by a *thin-slicing study* of participant ratings of

[77] Related research by Knox and Mummolo (2020b) and others – e.g., see Schimmack and Carlsson (2020) – critiqued fundamental methodological deficiencies in previously published research by Johnson, Tress, Burkel, Taylor and Cesario (2019) related to population vs subgroup inferences, which resulted in the correction and retraction of the Johnson et al. (2019) article. Certain of these critiques are particularly illustrative of the importance of proportionality of subgroup population relative to the aggregate population (see PNAS 117). By way of further background, see Princeton University (2020) "A Cautionary Tale About Measuring Racial Bias in Policing". Further research by Knox et al. e.g., employs *probabilitistic causal graph inference models* of the *data-generating process* to rigorously specify error bounds and thereby identify numerous implicit assumptions, inconsistencies and discrepancies in Fryer (2019), Gaebler, Cai, Basse, Shroff, Goel and Hill (2020), and Johnson, Tress, Burkel, Taylor and Cesario (2019) – see Duarte, Knox and Mummolo (2021) "Ɛ-sharp Bounds for Partially Observed Causal Processes: Testing for Racial Bias in Policing by Fusing Incomplete Records".

[78] The code and data available for replicating the analysis is hosted at https://github.com/stanford-policylab/risk-adjusted-regression.

officer utterances. The methodological approach applied computational linguistic tools and techniques to differentiate systematic disparities in officer speech across interactions based upon the race of the civilian subject, after controlling for the race of the officer, the severity of the infraction, the location of the stop, and the outcome of the stop. This dataset consisted of transcribed body camera audio and video for vehicle stops of white and black community members conducted by the Oakland Police Department during the month of April 2014 examined 981 stops of black (N = 682) and white (N = 299) drivers during this period (i.e., 68.1% of the 1,440 stops of white and black drivers in this period). The 981 stops were conducted by 245 different officers.[79]

Highlights of Foundational Governance Principles and Practices and Other Observations Applicable to Administrative and Supervisory Data Curation and Provenance

Fundamentally, in the context of principles of experimental design, statistical analyses of policing practices can be framed as *natural experiments*[80] implemented with data generated by police-civilian interactions (being de facto *ad hoc field experiments*[81]), conducted by officers as dictated by procedural and administrative practices instituted. As such, in order to implement and maintain a robust data repository and data dashboard, addressing procedural measures, administrative data records and statistical implications is absolutely necessary, being both intrinsically interrelated and fundamentally critical to foundational principles and practices for reasonably sound governance of data collection, acquisition, lineage, provenance and curation.

---

[79] Rob Voigt, Nicholas P. Camp, Vinodkumar Prabhakaran, William L. Hamilton, Rebecca C. Hetey, Camilla M. Griffiths, David Jurgens, Dan Jurafsky, and Jennifer L. Eberhardt (2017) PNAS 114 (25) 6521-6526 (https://doi.org/10.1073/pnas.1702413114). The reference cites Oakland Police Department policy requiring officers to activate body cameras prior to contact with the driver and to record for the duration of the stop, and further note that the resulting 183 hours of content in these interactions, yielded 36,738 usable officer utterances for language analysis. The reference includes an appendix describing the data sampling process for inclusion criteria (see https://www.pnas.org/content/pnas/suppl/2017/05/30/1702413114.DCSupplemental/pnas.1702413114.sapp.pdf), and cites several relevant sources, e.g., President's Task Force on 21st Century Policing (2015) Final Report of the President's Task Force on 21st Century Policing; The White House (2014) Fact sheet: Strengthening community policing. (Press release: https://obamawhitehouse.archives.gov/the-press-office/2014/12/01/fact-sheet-strengthening-community-policing); Reaves, B. (2015) Local Police Departments, 2013: Personnel, Policies, and Practices (US Dep Justice, Washington, DC); Eith C, Durose M (2011) Contacts Between Police and the Public (Bur Justice Stat, Washington, DC); Langton L, Durose M (2013) Special Report: Police Behavior During Traffic and Street Stops, 2011 (Bur Justice Stat, Washington, DC).

[80] A natural experiment is an empirical study in which individuals (or clusters of individuals) are exposed to the experimental and control conditions that are determined by nature or by other factors outside the control of the investigators. See DiNardo, J. (2008). "Natural experiments and quasi-natural experiments". In Durlauf, Steven N.; Blume, Lawrence E (eds.). The New Palgrave Dictionary of Economics; Dunning, T. (2012) Natural Experiments in the Social Sciences: A Design-Based Approach; Rosenzweig, M. R., K.I. Wolpin (2000). "Natural 'Natural Experiments' in Economics". Journal of Economic Literature. 38 (4).

[81] Although under the assumptions of random assignment, excludability and non-interference, outcomes of field experiments are considered to be unbiased, these assumptions are violated by asymmetries in assignment, administration or measurement. See e.g., Harrison, G. W., J. A. List (2004). "Field experiments," Journal of Economic Literature. 42 (4) and Rubin, Donald B. (2005) "Causal Inference Using Potential Outcomes" Journal of the American Statistical Association 100 (469).

By way of general background and context regarding procedural "mechanics" and operative relevance to forensic reliability – particularly in the context of reasonably suitable and appropriate convention, as well as established principles and practice standards for data integrity – *data curation*, i.e., the organization and integration of data collected from various sources, involves annotation, publication and presentation of the data such that the value of the data is maintained over time, and the data remains available for preservation and reuse.[82]

Data curation entails requisite processes not limited solely to production, processing, maintenance or management of data repositories, but extends to processes of digital forensics, information retrieval, search and visualization of scientific and technical text as well as other research content and modalities. In general, data curation encompasses a range of activities and processes implemented for data production, processing, storage, management, maintenance, verification and validation involving both data and corresponding *metadata* (e.g., time stamps).[83] Application in the context of legal and policy domains increasingly requires expertise in analytical practices of data curation as emphasis on disciplined curation of data has become more predominant, particularly for software processing in high volume on more complex data systems. *Data governance* entails capabilities that enable high data quality, integrity and fidelity across the complete lifecycle of a *data corpus*, and the corresponding data controls implemented to support use-case application objectives.

The **New Jersey State Police Review**, commissioned as the result of a consent decree subsequent to the State of New Jersey v Soto case, specifically focused on activities of state troopers assigned to patrol the NJ Turnpike, documents numerous observed interrelationships between procedural practices of *differential selective enforcement*, as well as prescriptive remedies, reforms and interventions (i.e., *Standard Operating Procedures*), and corresponding statistical implications (i.e., biased outcomes). The primary emphasis of the review was on the reliance by the officer on

[82] Data governance plays a central role in in metadata management. Related operative principles relevant to data governance include the following: *data lineage* and *data provenance*. Data lineage provides an “audit trail” comprised of technical metadata documenting data transformations, which might include e.g., data quality test results, reference data values, data models, vocabulary, data stewards. The objective of data provenance is documentation of data transformations with sufficient detail to enable reproducibility, by tracking data across transformations, analyses and interpretations. Datasets are deemed to be reliable when the underlying data process is verifiably reproducible. Ikeda, R., Park H., Widom, J.. Provenance for generalized map and reduce workflows. In Proc. of CIDR, January 2011; Cui, Y. and Widom, J. (2003) Lineage tracing for general data warehouse transformations. VLDB Journal, 12(1).

[83] Metadata is generally defined as types of standardized contextual data (e.g., labels, tags) specifying one or more characteristics or properties of corresponding data elements (as content), in order to facilitate acquisition, collection, indexation, storage, transmission, transformation, as well as other administration and usage of specific data, by describing the contents and context of data or data files. Reference metadata describes content and quality of statistical data. Statistical metadata (also called process data) describe processes that collect, process, or produce statistical data. Dippo, C., Sundgren, B. (2000) “The Role of Metadata in Statistics” Bureau of Labor Statistics; Directorate, OECD Statistics. “OECD Glossary of Statistical Terms - Reference metadata Definition” (stats.oecd.org); Zeng, M. (2004) “Metadata Types and Functions” (https://marciazeng.slis.kent.edu/metadatabasics/types.htm) National Information Standards Organization (NISO); NISO (2001) “Understanding Metadata”.

"race, ethnicity or national origin in conjunction with other factors in selecting vehicles to be stopped" and other discretionary decisions during a traffic stop.[84]

The **Fort Worth Police Department Expert Review Panel Preliminary Observations and Recommendations** discusses implementation of an *early intervention system* (EIS) to risky and problematic trends before a serious incident occurs, including uses of force, external community member complaints, stops, and arrests, domestic violence allegations, missed court appearances and other conduct as indicators. The discussion regarding EIS implementation references a document which further cites numerous sources, including relevant US Department of Justice references, as well as technical assistance guides, related consent decrees and case studies on implementation, maintenance and application of such monitoring systems.[85] This reference describes key components of EIS (i.e., identification, evaluation, intervention and monitoring), as well as key performance indicators as flags (e.g., yellow, red) for risk mitigation. It further summarizes and outlines functional, policy, practice and operational considerations.

A collaboration involving US Department of Justice Community-Oriented Policing Services (COPS) and the California Attorney General has published a guide for the collection and use of stop data. Key observations include the following: (i) of the approximately 20 states requiring data collection on vehicle stops, requirements vary widely with a lack of cohesive curation practices to enable standardization and analysis, (ii) stop data can be applicable to examining law enforcement policies and practices, and with resource allocation and more systematic observation of disproportionality and disparity, (iii) all law enforcement agencies conducting stops should collect stop data (including specialized units). Irrespective of data collection methods (paper form, handheld mobile device, mobile data computer), it is critical that data be practicably complete, accurate, subject to robust analysis, and be publicly available "in a way that is contextualized and easy to understand."[86] [87]

[84] Among concerns the report highlights are the following: (i) the extent of missing information regarding racial characteristics of detained motorists; (ii) officer discretion; (iii) the tautological misuse of racial statistics to validate pre-existing stereotypes (i.e., *self-confirming* and *reinforcement bias*); (iv) the limitations of crime analysis statistics for police resource allocation; (v) the critical importance of reliable study of racial and ethnic characteristics of the population of drivers using the NJ Turnpike (and proposed a population survey in consultation with Civil Rights Division of the US Department of Justice).
See https://www.state.nj.us/lps/intm_exe.pdf; https://www.state.nj.us/lps/intm_419.pdf.

[85] See Best Practices in Early Intervention Systems (https://www.policefoundation.org/publication/best-practices-in-early-intervention-system-implementation-and-use-in-law-enforcement-agencies/) also referred to as early warning systems (EWS) and early warning and intervention systems (EWIS). Other use-case applications for monitoring dashboards being explored have been police de-escalation training (e.g. Camden NJ Police Department).

[86] "Collecting, Analyzing, and Responding to Stop Data: A Guidebook for Law Enforcement Agencies, Government, and Communities"
https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/5f7335d7294be10059d32d1c/1601385959666/COPS-Guidebook+Final+Release+Version.pdf.

[87] In discussing the absence of adequately curated policing data, an article discussed undue reliance by most police departments on legacy IT for record keeping and record management, in contrast to investments in surveillance and other policing equipment. See "A Major Challenge to Policing Reform: the Absence of Good Data" (https://slate.com/news-and-politics/2020/07/policing-project-barry-friedman-pesca-gist-interview-transcript.html).

Standardization of Data from State-Maintained Crime Statistics Data Portals. The **Uniform Crime Reporting (UCR) and Colorado Crime Statistics** data portal maintained by the Colorado Bureau of Investigation (CBI) exhibits crime statistics reported by Colorado law enforcement agencies.[88]

> "… January 1, 2021, marks a new era in the Uniform Crime Reporting (UCR) Program's partnership with law enforcement to provide more meaningful data to help understand crime in our communities and in our nation as a whole. Not only is the UCR Program completing its transition to the National Incident-Based Reporting System (NIBRS), but it is also completing its migration from traditional electronic publications to dynamic data presentations through the Crime Data Explorer (CDE). With each of these changes endorsed by law enforcement, the UCR Program remains committed to making available the types of data that aid in combatting crime and promoting transparency.
>
> Since the CJIS Advisory Policy Board recommended the FBI transition to the NIBRS-only data collection five years ago, thousands of agencies have made the move from the Summary Reporting System (SRS) to the more detailed data of NIBRS, and thousands more are committed to making the switch. As of October 31, 2020, 43 states were NIBRS-certified, i.e., the states have records management systems that meet the FBI's requirements for collecting crime data according to established technical specifications. At that time, 8,742 law enforcement agencies representing 48.9 percent of the population were reporting NIBRS data to the UCR Program. The FBI also collaborated with federal and tribal agencies to develop the NIBRS Collection Application as a solution for these agencies to submit their data. The UCR Program conducted virtual training for approximately 45 federal and tribal agencies since from June through November 2020. The FBI continues to assist all law enforcement agencies with their transition through training, data integration, and technical assistance with NIBRS data specifications and reporting requirements. …"[89]

A key difference between the UCR summary versus the NIBRS is that NIBRS statistics, being more granular, "counts every crime recorded for each incident", versus UCR which reports "the single most serious offense", with the result that NIBRS counts are higher than UCR summary counts.[90]

---

[88] The UCR and Colorado Crime Statistics data portal indicates that it is successor to the Crime in Colorado report due to the transition from the FBI Uniform Crime Reporting (UCR) summary measurement to the more detailed National Incident-Based Reporting System (NIBRS) advocated and "soon to be required by the FBI." See https://cbi.colorado.gov/sections/crime-information-management-unit/uniform-crime-reporting-ucr-and-colorado-crime.

[89] See https://ucr.fbi.gov/nibrs/2019. Also see https://www.waspc.org CJISPDF on incident-Based Reporting and RMS.

[90] Edwards, Lee and Esposito (2019) indicate inadequate tracking of officer-involved deaths and further highlight fundamental limitations acknowledged by the BJS of the Arrest-Related Deaths data, as well as corresponding limitations to the National Vital Statistics system (NVSS) which undercounts law enforcement related deaths, and the National Violent Death Reporting System (NVDRS), which although exhibits better coverage of police-involved deaths than NVSS, currently lacks geographic and temporal coverage.

Interactive dashboard of enforcement patterns by charges and demographics. The ***John Jay College Data Collaborative Research Network on Misdemeanor Justice*** maintains an interactive online dashboard to monitor cross-jurisdictional enforcement patterns (e.g., general trends in misdemeanor arrest rates) by charge and by demographics (i.e., race, age , gender), across eight geographically diverse jurisdictions (Durham NC, Los Angeles CA, Louisville KY, New York City, Prince George's County MD, Seattle WA and St. Louis MO), based upon reports from local researchers and criminal justice agencies.[91]

As illustrated by these references, interpretation and inferences of these analyses as natural experiments are subject to inherent sample selection bias which must be judiciously mitigated. These highlights also illustrate that in designing a data-driven monitoring dashboard, it will be critical to consider: (i) what is "under the hood", i.e., the properties of the underlying data and computations might constitute the visualization: the underlying data structure (e.g., a suitably configured dataset of binary [0,1] variables), (ii) vendor capabilities and capacity for scalable adaptation and integration, [92] (iii) and the time and resources a prospective dataset might require and what analytical objectives might be entailed (e.g. *OLAP data*,[93] *clustering graphs*,[94] *capture-recapture methods*,[95] and *factor analysis*,[96] among other relevant methods), (iii) and model specification and implementation considerations for e.g., Poisson, negative binomial, logit/probit, multinomial and loglinear regression, propensity analysis, as well as proportional hazard and competing risk models with fixed, mixed and/or random effects.

---

[91] Appendix B of the *Misdemeanor Enforcement Trends Across Seven U.S. Jurisdictions October 2020* describes its data limitations.

[92] The *Capability Maturity Model* (CMM) was originally developed as a tool for objectively assessing the ability of government contractors' processes to implement a contracted software project. The approach has been adopted beyond software development, and is widely applied as a general model of the maturity of process (e.g., IT service management) across IT. The *Capability Maturity Model Integration* (CMMI) project was formed to address the problem of implementing multiple models for software development processes, and has superseded the CMM model, though the CMM model continues to be a general theoretical process capability model used in the public domain. See Humphrey, W. S. (1988) "Characterizing the software process: A maturity framework" IEEE Software. 5 (2); https://resources.sei.cmu.edu/library/asset-view.cfm?assetid=11955; https://ieeexplore.ieee.org/document/2014.

[93] Gray, Bosworth, Layman and Pirahesh (1996) "Data Cube: Aggregation Operator Generalizing Group-By, Cross-Tab and Sub-Totals".

[94] See Nisbet, R., Miner, G. et al. (2018) "Advanced Data Mining Algorithms" in *Handbook of Statistical Analysis and Data Mining Applications* (2nd Edition). Also see Brandao and Moro (2017) "Social professional networks: A survey and taxonomy" *Computer Communications* Vol. 100.

[95] See e.g. Rivera and Rosenbaum (2020).

[96] Factor analysis refers to a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. See e.g., Bandalos, Deborah L. (2017) "Measurement Theory and Applications for the Social Sciences"; Bartholomew, D.J.; Steele, F.; Galbraith, J.; Moustaki, I. (2008) "Analysis of Multivariate Social Science Data"; Cattell, R. B. (1952) "Factor analysis"; Cattell, R. B. (1978) "Use of Factor Analysis in Behavioral and Life Sciences"; Child, D. (2006) "The Essentials of Factor Analysis"; Fruchter, B. (1954) "Introduction to Factor Analysis"; Gorsuch, R. L. (1983) "Factor Analysis"; Harman, Harry H. (1976) "Modern Factor Analysis"; Jolliffe I.T. (2002) "Principal Component Analysis"; McDonald, R. P. (1985) Factor Analysis and Related Methods; Velicer, W.F. (1976) "Determining the number of components from the matrix of partial correlations".

Although some contiguous domain considerations which might be outside of the current scope of this assignment, inherent dashboard and database-related forensic implications for robust and reliable statistical analytics are readily apparent. As previously discussed and referenced, such methods when correctly specified and implemented represent customary generally-accepted methodologies for objective, forensically reliable and robust statistical analysis across a suitably representative dataset.[97]

[97] Cloud-based architectures are increasingly enabling adoption of established techniques from operations research (logistics) and statistical process control to police resource allocation, and the application of search and information retrieval techniques, as well as widely-adopted statistics-based methodological frameworks employing graphical analysis beyond SQL data base, e.g. NoSQL and graphical search and dynamic data models e.g., *Dynamic Distributed Dimensional Data Model* (D4M). A D4M query returns a sparse matrix or graph for statistical signal processing or graph analysis of a database regardless of whether is structured or unstructured. Whether graphical, numeric or string data, the primary objective of D4M is to process heterogeneous data (https://d4m.mit.edu/), with interoperability between diverse databases, by combining advantages of five processing technologies (e.g., sparse linear algebra, associative arrays, fuzzy algebra, distributed arrays, and triple-store/NoSQL databases such as *Hadoop HBase* and *Apache Accumulo*) to provide a database and computation system that addresses the problems associated with expanding volumes of high-dimensional data (and corresponding metadata) analyze on demand, in addition to the adoption of adaptive algoriths to analyze data streams. See https://arxiv.org/ftp/arxiv/papers/1702/1702.03253.pdf; https://arxiv.org/pdf/1708.02934.pdf.

VII. Bibliography

Agresti, A. and B.A. Coull (1998), "Approximate is Better than 'Exact' for Interval Estimation of Binomial Proportions," The American Statistician, Vol. 52: No. 2, pp. 119–126.

Akpinar, N., M. De-Artaga and A. Chouldechova (2021), "The Effect of Differential Victim Crime Reporting on Predictive Policing Systems," ACM Virtual Event, March 2021.

Aurora Police Department (1998), "Section 5.4.7. Reporting, and Investigating the Use of Tools, Weapons and Physical Force" in APD Policies: Directives Manual Chapter 5 Weapons and Use of Force https://public.powerdms.com/AURORAPD/tree/documents/107).

Ayres, I. (2002), "Outcome Tests of Racial Disparities in Police Practices," Justice Research and Policy, Vol. 4: pp. 131-142, NCJ 199116.

Ba, B., D. Knox, J. Mummolo and R. Rivera (2021), "The Role of Officer Race and Gender in Police-Civilian Interactions in Chicago," Science, Vol. 371: Issue 6530, pp. 696-702.

Bandalos, D. L. (2017), "Measurement Theory and Applications for the Social Sciences," Guilford Press.

Banerjee, S., B.P. Carlin and A.E. Gelfand (2014), "Hierarchical Modeling and Analysis for Spatial Data," Second Edition, Chapman and Hall/CRC.

Banerjee, S., A.E. Gelfand, A.O. Finley and H. Sang (2008), "Gaussian Predictive Process Models for Large Spatial Datasets," Journal of the Royal Statistical Society, Series B: 70 (4), pp. 825–848.

Bartholomew, D.J., F. Steele, I. Moustaki, and J. Galbraith (2008), "Analysis of Multivariate Social Science Data," Chapman & Hall.

Baumgartner, F.R., D.A. Epp, and K. Shoub (2018), "Suspect Citizens: What 20 Million Traffic Stops Tell Us About Policing and Race," Cambridge University Press.

Becker, G.S. (1957), "The Economics of Discrimination," University of Chicago Press.

Becker, G.S. (1993), "Nobel Lecture: The Economic Way of Looking at Behavior," Journal of Political Economy, Vol. 101: No. 3, pp. 385-409.

Berry B.J. and K.B. Smith (Eds.) (1972), "City Classification Handbook: Methods and Applications," John Wiley.

Blalock, H.M. (1960), "Social Statistics," McGraw-Hill.

Bollen, K.A. (1989), "Structural Equations with Latent Variables," Wiley.

Bollen, K.A. and J. Pearl (2013), "Eight Myths About Causality and Structural Equation Models," in S.L. Morgan (Ed.), Handbook of Causal Analysis for Social Research, Chapter 15, pp. 301-328, Springer.

Brown, L.D., T. Cai and A. DasGupta (2001), "Interval Estimation for a Binomial Proportion," Statistical Science, Vol. 16: No. 2, pp. 101–133.

Brunsdon, C., A.S. Fotheringham and M.E. Charlton (1996), "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity," Geographical Analysis, Vol. 28: Issue 4, pp. 281–298.

Cattell, R.B. (1952), "Factor Analysis: An Introduction and Manual for the Psychologist and Social Scientist," Harper.

Cattell, R.B. (1978), "The Scientific Use of Factor Analysis in Behavioral and Life Sciences," Springer.

Cheng, E.K. (2009), "A Practical Solution to the Reference Class Problem," 109 Colum. L. Rev. 2081, pp. 2095-97.

Chicago Police Accountability Task Force (2016), "Recommendations for Reform."

Child, D. (2006), "The Essentials of Factor Analysis," Bloomsbury Press.

Cochran, W.G. (1977), "Sampling Techniques," Wiley.

Cramer, H. (1946), "Mathematical Methods of Statistics," Princeton University Press.

Cui, Y. and J. Widom (2003), "Lineage Tracing for General Data Warehouse Transformations," VLDB Journal, 12(1).

Datta, A, S. Banerjee, A,O, Finley, A.E. Gelfand (2016), "Hierarchical Nearest Neighbor Gaussian Process Models for Large Geostatistical Datasets," Journal of the American Statistical Association. Vol. 111: No. 514: pp. 800–812.

Davis, E., A. Whyde and L. Langton (2018), "Contacts Between Police and the Public, 2015," U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, NCJ 251145.

Devore, J.L. (1995), "Probability and Statistics for Engineering and the Sciences," Fourth Edition, Duxbury Press.

DiNardo, J. (2008), "Natural Experiments and Quasi-Natural Experiments," in Durlauf, S.N. and L.E. Blume (Eds.), The New Palgrave Dictionary of Economics.

Dippo, C. and B. Sundgren (2000), "The Role of Metadata in Statistics," Bureau of Labor Statistics.

Duarte, G., D. Knox and J. Mummolo (2021), "Ɛ-sharp Bounds for Partially Observed Causal Processes: Testing for Racial Bias in Policing by Fusing Incomplete Records," Presentation at NYU Center for Data Science on February 17, 2021. Also see Presentation at Yale University Quantitative Research Methods Workshop on February 25, 2021.

Duncan, O.D. (1975), "Introduction to Structural Equation Models," Elsevier.

Dunning, T. (2012), "Natural Experiments in the Social Sciences: A Design-Based Approach," Cambridge University Press.

Durlauf, S.N. and J.J. Heckman (2020), "An Empirical Analysis of Racial Differences in Police Use of Force: A Comment," Journal of Political Economy, Vol. 128: No. 10.

Edwards, F., H. Lee and M. Esposito (2019), "Risk of Being Killed by Police Use of Force in the United States by Age, Race–Ethnicity, and Sex," Proceedings of the National Academic of Sciences, Vol. 116: No. 34, pp. 16793-16798.

Eith C. and M. Durose (2011), "Contacts Between Police and the Public," US Department of Justice, Bureau of Justice Statistics, NCJ 234599.

Epp, C., S. Maynard-Woody and D.P. Hayder-Markel (2014), "Pulled Over: How Police Stops Define Race and Citizenship," University of Chicago Press.

Feller, W. (1945), "On the Normal Approximation to the Binomial Distribution," The Annals of Mathematical Statistics, Vol. 16: No. 4, pp. 319-329.

Fisher, R.A. (1973), "Statistical Methods for Research Workers," in Yates, (Ed.) Statistical Methods for Research Workers, Statistical Methods and Scientific Inference, and Design of Experiments.

Fitzpatrick, D.J., W.L. Gorr and D.B. Neill (2019), "Keeping Score: Predictive Analytics in Policing," Annual Review of Criminology, Vol. 2: 7.1-7.19.

Fleckenstein, M. and L. Fellows (2018), "Modern Data Strategy," Springer.

Fleiss, J.L. (1981), "Statistical Methods for Rates and Proportions," John Wiley & Sons.

Franklin, J. (2010), "Feature Selection Methods For Solving The Reference Class Problem: Comment on Edward K. Cheng, "A Practical Solution to the Reference Class Problem"," 110 Colum. L. Rev. Sidebar, pp. 12–23.

Freedman, D.A. and R.A. Berk (2008), "Weighting Regressions by Propensity Scores," Evaluation Review, Vol. 32: Issue 4, pp. 392-409.

Fruchter, B. (1954) "Introduction to Factor Analysis," Van Nostrand.

Fryer, R. (2018), "Reconciling Results on Racial Differences in Police Shootings," The American Economic Review, Vol. 108, pp. 228–33.

Fryer, R. (2019), "An Empirical Analysis of Racial Differences in Police Use of Force, Journal of Political Economy, Vol. 127: No. 3.

Fürber, C. (2015), "Data Quality Management with Semantic Technologies," Springer.

Gabriel, K.R. (1966), "Simultaneous Test Procedures for Multiple Comparisons on Categorical Data," Journal of the American Statistical Association, Vol. 61: No. 316, pp. 1081-1096.

Gaebler, J., W. Cai, G. Basse, R. Shroff, S. Goel and J. Hill (2020), "A Causal Framework for Observational Studies of Discrimination," ArXiv Preprint.

Gelman, A., J. Fagan and A. Kiss (2007), "An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias," Journal of the American Statistical Association, Vol. 102: No. 479, pp. 813-823.

Goel, S., J.M. Rao and R. Shroff (2016), "Precinct or Prejudice? Understanding Racial Disparities in New York City's Stop-and-Frisk Policy," Annals of Applied Statistics, Vol. 10: No. 1, pp. 365-394.

Goel, S., J.M. Rao and R. Shroff (2017), "Personalized Risk Assessments in the Criminal Justice System", American Economic Review, Vol. 106: No. 5, pp. 119-123.

Goerger, S., J. Mummolo and W. Westwood (2020), "Which Police Departments Want Reform? Barriers to Evidence-Based Policymaking," Working paper.

Goodman, L.A. and W.H. Kruskal (1979), "Measures of Association for Cross-Classification," Journal of the American Statistical Association, Vol. 49: No. 268, pp. 732-764.

Gorsuch, R.L. (1983), "Factor Analysis," Lawrence Erlbaum and Associates.

Gray, J., A. Bosworth, A. Layman and H. Pirahesh (1996), "Data Cube: Aggregation Operator Generalizing Group-By, Cross-Tab and Sub-Totals," Institute of Electrical and Electronics Engineers (IEEE): Proceedings of the Twelfth International Conference on Data Engineering, pp. 152-159.

Grogger J. and G. Ridgeway (2006), "Testing for Racial Profiling in Traffic Stops From Behind a Veil of Darkness," Journal of the American Statistical Association, Vol. 101, pp. 878-887.

Harman, H.H. (1976), "Modern Factor Analysis," University of Chicago Press.

Harrison, G.W. and J.A. List (2004), "Field Experiments," Journal of Economic Literature, Vol. 42: No. 4, pp. 1009-1055.

Heckman, J. J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," Annals of Economic Social Measurement, Vol.5: No. 4, pp.475–492.

Heckman, J. J. (1977), "Sample Selection Bias as a Specification Error (with an Application to the Estimation of Labor Supply Functions)," NBER Working Paper No. 0172.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error," Econometrica, Vol. 47: No. 1, pp.153–161.

Heckman, J., N. Hohmann and J. Smith with M. Khoo (1998a), "Substitution and Dropout Bias in Social Experiments: Evidence From an Influential Social Experiment," Quarterly Journal of Economics.

Heckman, J., H. Ichimura, J. Smith and P. Todd (1998b), "Characterizing Selection Bias Using Experimental Data," Econometrica Vol. 66, pp. 1017-1098.

Heckman, J., H. Ichimura and P. Todd (1998c), "Matching as an Econometric Evaluation Estimator, Review of Economic Studies, Vol. 65: No. 2, pp. 261-294.

Herzog, T.N., F.J. Scheuren and W.E. Winkler (2007), "Data Quality and Record Linkage Techniques," Springer.

Holland, P. (1986), "Statistics and Causal Inference," Journal of the American Statistical Association, Vol. 81: No. 396, pp. 945-960.

Holland, P. (1988), "Causal Inference, Path Analysis and Recursive Structural Equation Models," in C. Clogg, (Ed.), Sociological Methodology (American Sociological Association, Washington, DC) pp. 449-484.

Horrace, W.C., and S.M. Rohlin (2016), "How Dark is Dark? Bright Lights, Big City, Racial Profiling," Review of Economics and Statistics, Vol. 98: No. 2, pp. 226–232.

Humphrey, W.S. (1988), "Characterizing the Software Process: A Maturity Framework," Institue of Electrical and Electronic Engineers (IEEE), Software, Vol. 5: Issue 2.

Ikeda, R., H. Park, and J. Widom (2011), "Provenance for Generalized Map and Reduce Workflows," In Proceedings for Conference on Innovative Data Systems Research.

Irvine, J. (2014), "Transforming Data into Information: Enabling Detection and Discovery," MITRE #14-2487.

Johnson, D., T. Tress, N. Burkel, C. Taylor and J. Cesario (2019), "Officer Characteristics and Racial Disparities in Fatal Officer-Involved Shootings," Proceedings of the National Academy of Sciences, Vol. 116: No. 32, pp. 15877-15882.

Jolliffe I.T. (2002), "Principal Component Analysis," Springer-Verlag.

Jung, J., S. Corbett-Davies, R. Shroff and S. Goel (2019), "Omitted and Included Variable Bias in Tests for Disparate Impact," Working Paper.

Jurafsky, and J.L. Eberhardt (2017), "Language from Police Body Camera Footage Shows Racial Disparities in Officer Respect," Proceedings of the National Academy of Sciences, Vol. 114: No. 25, pp. 6521-6526.

Kadane, J.B. and N. Terrin (1997), "Missing Data in the Forensic Context," Journal of the Royal Statistical Society, Vol. 160: No. 2, pp. 351-357.

Kahn, K.B., P.A. Goff, J.K. Lee and D. Motamed (2016), "Protecting Whiteness: White Phenotypic Racial Stereotypicality Reduces Police Use of Force," Social Psychological and Personality Science Journal (Sage).

Kaplan, D. (2009), Structural Equation Modeling: Foundations and Extensions," Sage.

Kline, R.B. (2016), "Principles and Practice of Structural Equation Modeling," Guilford Press.

Knowles, J., N. Persico and P. Todd (2001), "Racial Bias in Motor Vehicle Searches: Theory and Evidence," Journal of Political Economy, Vol. 109: No. 1.

Knox D., W. Lowe and J. Mummolo (2020a), Administrative Records Mask Racially Biased Policing, American Political Science Review, Vol. 114: Issue 3.

Knox D., W. Lowe and J. Mummolo (2020b), "Can Racial Bias in Policing Be Credibly Estimated Using Data Contaminated by Post-Treatment Selection?" ArXiv Preprint.

Knox D., and J. Mummolo (2020a), "Toward a General Causal Framework for the Study of Racial Bias in Policing," Journal of Political Institutions and Political Economy, Vol. 1: No. 3, pp. 341-378.

Knox, D., and J. Mummolo (2020b), "Making Inferences about Racial Disparities in Police Violence," Proceedings of the National Academy of Sciences, Vol. 117: No 3, pp. 1261–1262.

Langton L. and M. Durose (2013), "Special Report: Police Behavior During Traffic and Street Stops, 2011," U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics, September 2013.

Mahanti, R. (2019), "Data Quality: Dimensions, Measurement, Strategy, Management and Governance," Quality Press.

McDonald, R.P. (1985), "Factor Analysis and Related Methods," Erlbaum.

Mebius, C., L. Dobrowolski, T. Keesee, R. N. Johnson-Ahorlu, J. Tindel, D. Johnson, B. Chen, C. Eidson, A. LeGrand, M. Ponomarenko, C. Socci, R. Tholin, R. Chance, J. Reich, K. Walker, T. Simmons, and A. Burke, "Collecting, Analyzing, and Responding to Stop Data: A Guidebook for Law Enforcement Agencies, Government, and Communities," Office of Community Oriented Policing Services, U.S. Department of Justice cooperative agreement number 2016-CK-WX-KOJ6.

Miller, R.G. (1981), "Simultaneous Statistical Inference," Springer.

National Information Standards Organization (2001), "Understanding Metadata," NIST.

Nance, D.A. (2007), "The Reference Class Problem and Mathematical Models of Inference," The International Journal of Evidence & Proof, Vol. 11: Issue 4, pp. 259-273.

National Information Standards Organization (2019), "NIST Big Data Interoperability Framework: Vol. 4, Security and Privacy," NIST.

OECD Statistics Directorate, "OECD Glossary of Statistical Terms - Reference Metadata Definition," (stats.oecd.org).

Pearl, J. (2000), "Causality: Models, Reasoning and Inference," Cambridge University Press.

Peeples, L. (2019), "What the Data Say About Police Shootings," Nature, Vol. 573: Issue 7772, pp. 24-26.

Pesca, M. (2020), “A Major Challenge to Policing Reform: the Absence of Good Data,” Slate.

Pierson, E., C. Simoiu, J. Overgoor, S. Corbett-Davies, D. Jenson, A. Shoemaker, V. Ramachandran, P. Barghouty, C. Phillips, R. Shroff and S. Goel (2020), “A Large-Scale Analysis of Racial Disparities in Police Stops Across the United States,” Nature Human Behavior, Vol. 4, pp. 736–745.

Plant, A.L. and R.J. Hanisch (2018), “Reproducibility and Replicability in Science, A Metrology Perspective,” A Report to the National Academies of Sciences, Engineering and Medicine Committee on Reproducibility and Replicability in Science.

President’s Task Force on 21st Century Policing (2015), “Final Report of the President’s Task Force on 21st Century Policing,” Office of Community Oriented Policing Services.

Princeton University (2020), “A Cautionary Tale About Measuring Racial Bias in Policing.”

Reaves, B.A. (2015), “Local Police Departments, 2013: Personnel, Policies, and Practices,” US Department of Justice, Bureau of Justice Statistics.

Rummel, R.J. (1970), “Applied Factor Analysis,” Northwestern University Press.

Russom, P. (2006), “Liability and Leverage - A Case for Data Quality,” Vol. 16: Issue 8.

Ridgeway, G. (2006), “Assessing the Effect of Race Bias in Post-Traffic Stop Outcomes Using Propensity Scores,” Journal of Quantitative Criminology, Vol. 22: No. 1, pp. 1-29.

Rivera, R. and J. Rosenbaum (2020), “Racial Disparities in Police Stops in US Cities,” Significance, Vol. 17: Issue 4, pp. 4-5.

Rodgers, W.M. (Ed.) (2006) “Handbook on the Economics of Discrimination,” Edward Elgar.

Rosenzweig, M.R. and K.I. Wolpin (2000). “Natural 'Natural Experiments' in Economics,” Journal of Economic Literature, Vol. 38: No. 4, pp. 827-874.

Ross, C.T. (2015), “A Multi-Level Bayesian Analysis of Racial Bias in Police Shootings at the County-Level in the United States,” PloS One, pp. 2011–2014.

Rubin, Donald B. (2005), “Causal Inference Using Potential Outcomes,” Journal of the American Statistical Association, Vol. 100: Issue 469, pp. 322-331.

Schimmack, U. and R. Carlsson (2020), "Young Unarmed Nonsuicidal Male Victims of Fatal Use of Force are 13 Times More Likely to be Black than White," Proceedings of the National Academy of Sciences, Vol. 117: No. 3, pp. 1263.

Sheskin, D.J. (1997), "Handbook of Parametric and Nonparametric Statistical Procedures," Chapman & Hall.

Simoiu, C., A. Corbett-Davies and S. Goel (2017), "The Problem of Infra-Marginality in Outcome Tests," The Annals of Applied Statistics, Vol. 11: No. 3, pp. 1193–1216.

Smallwood, R.F. (2014), "Information Governance: Concepts, Strategies, and Best Practices," John Wiley and Sons.

Snedecor, G.W. and W.G. Cochran. (1980), "Statistical Methods," Wiley-Blackwell.

Tarka, P. (2017), "An Overview of Structural Equation Modeling: Its Beginnings, Historical Development, Usefulness and Controversies in the Social Sciences," Quality & Quantity, Vol. 52: No. 1, pp. 313–354.

Terrin, N. and J.B. Kadane (1998), "Counting Cars in a Legal Case Involving Differential Enforcement," American Statistical Association, Chance, Vol. 11: No. 3, pp. 25-27.

Thulin, M. (2014), "The Cost of Using Exact Confidence Intervals for a Binomial Proportion," Electronic Journal of Statistics, Vol. 8: No. 1, pp. 817–840.

Tucker L.R. (1964), "The Extension of Factor Analysis to Three-Dimensional Matrices," in Frederiksen, N. and H. Gulliksen (Eds.), Contributions to Mathematical Psychology, Holt, Rinehart and Winston.

Voigt, R., N.P. Camp, V. Prabhakaran, W.L. Hamilton, R.C. Hetey, C.M. Griffiths, and D. Jurgens (2017), "Language from Police Body Camera Footage Shows Racial Disparities in Officer Respect," Proceedings of the National Academy of Sciences of the United States of America, Vol. 114: No. 25, pp. 6521–6526.

The White House (2014), "Fact Sheet: Strengthening Community Policing."

US Attorney's Office Western District of Washington and US Department of Justice Civil Rights Division (2011), "Investigation of the Seattle Police Department."

US Department of Justice Civil Rights Division (2015), "Investigation of the Ferguson Police Department."

US Department of Justice Civil Rights Division (2011), "United States' Investigation of the Maricopa County Sheriffs Office."

Velicer, W.F. (1976), “Determining the Number of Components from the Matrix of Partial Correlations,” Springer.

Wang, J.F., T.L. Zhang and B.J. Fu (2016), "A Measure of Spatial Stratified Heterogeneity," Ecological Indicators, Vol. 67, pp. 250–256.

Wood, G., D. Roithmayr and A.V. Papachristos (2019), “The Network Structure of Police Misconduct,” Socius, Vol. 5, pp. 1–8.

Wood, G. (2020), “A Cartography of Policing and the Use of Force in Chicago,” Presentation at NYU Center for Data Science on November 4, 2020.

Yates, F. (1934), “Contingency Table Involving Small Numbers and the $\chi 2$ Test,” Supplement to the Journal of the Royal Statistical Society, Vol. 1: No. 2, pp. 217–235.

Zeng, M. (2004), “Metadata Types and Functions,” National Information Standards Organization (NISO).

Figure 1.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Interactions to Population Counts
2019

Race/Ethnicity[1]	Population	Interactions	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	90	2.8%	159.6	0.000
Asian	23,917	958	4.0%	778.3	0.000
Black/African American	60,909	13,570	22.3%	5,214.8	0.000
Hispanic or Latino	101,562	10,840	10.7%	44.3	0.000
Native Hawaiian/Pacific Islander	1,296	117	9.0%	0.8	0.361
Non-White (All)	190,887	25,575	13.4%	957.1	0.000
White/Non-Hispanic	163,765	16,092	9.8%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Population Counts
2019

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	51	1.6%	43.5	**0.000**
Asian	23,917	377	1.6%	313.5	**0.000**
Black/African American	60,909	6,094	10.0%	2,972.8	**0.000**
Hispanic or Latino	101,562	4,355	4.3%	22.8	**0.000**
Native Hawaiian/Pacific Islander	1,296	37	2.9%	3.6	0.057
Non-White (All)	190,887	10,914	5.7%	593.8	**0.000**
White/Non-Hispanic	163,765	6,394	3.9%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Population Counts
2019

Race/Ethnicity[1]	Population	UoF Incidents	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	3	0.1%	0.0	0.846
Asian	23,917	6	0.0%	9.4	**0.002**
Black/African American	60,909	203	0.3%	182.8	**0.000**
Hispanic or Latino	101,562	102	0.1%	2.0	0.162
Native Hawaiian/Pacific Islander	1,296	0	0.0%	1.1	0.298
Non-White (All)	190,887	314	0.2%	45.3	**0.000**
White/Non-Hispanic	163,765	137	0.1%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.D
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Interactions to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Interactions	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	259	8.1%	485.0	0.000
Asian	23,917	2,793	11.7%	2,367.2	0.000
Black/African American	60,909	39,873	65.5%	15,008.2	0.000
Hispanic or Latino	101,562	31,629	31.1%	82.8	0.000
Native Hawaiian/Pacific Islander	1,296	272	21.0%	29.5	0.000
Non-White (All)	190,887	74,826	39.2%	2,573.2	0.000
White/Non-Hispanic	163,765	47,745	29.2%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.E
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	137	4.3%	135.9	**0.000**
Asian	23,917	1,075	4.5%	906.0	**0.000**
Black/African American	60,909	16,904	27.8%	7,760.4	**0.000**
Hispanic or Latino	101,562	11,846	11.7%	12.1	**0.001**
Native Hawaiian/Pacific Islander	1,296	106	8.2%	10.5	**0.001**
Non-White (All)	190,887	30,068	15.8%	1,340.9	**0.000**
White/Non-Hispanic	163,765	18,334	11.2%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.F
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	UoF Incidents	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	5	0.2%	0.7	0.395
Asian	23,917	12	0.1%	32.2	**0.000**
Black/African American	60,909	605	1.0%	596.1	**0.000**
Hispanic or Latino	101,562	310	0.3%	14.4	**0.000**
Native Hawaiian/Pacific Islander	1,296	2	0.2%	0.3	0.578
Non-White (All)	190,887	934	0.5%	162.7	**0.000**
White/Non-Hispanic	163,765	374	0.2%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.G
Pearson Chi-Squared Tests
Ratios of Interactions, Arrests, and UoF Incidents to Population Counts
2019

Race/Ethnicity[1]	Population	Interactions	(as % of Population)	Arrests	(as % of Population)	UoF Incidents	(as % of Population)
American Indian/Alaska Native	3,203	90	2.8%	51	1.6%	3	0.1%
Asian	23,917	958	4.0%	377	1.6%	6	0.0%
Black/African American	60,909	13,570	22.3%	6,094	10.0%	203	0.3%
Hispanic or Latino	101,562	10,840	10.7%	4,355	4.3%	102	0.1%
Native Hawaiian/Pacific Islander	1,296	117	9.0%	37	2.9%	0	0.0%
White/Non-Hispanic	163,765	16,092	9.8%	6,394	3.9%	137	0.1%
		Chi-Squared Statistic:	7810.6		4287.3		255.1
		P-Value:	**0.000**		**0.000**		**0.000**

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 1.H
Pearson Chi-Squared Tests
Ratios of Interactions, Arrests, and UoF Incidents to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Interactions	(as % of Population)	Arrests	(as % of Population)	UoF Incidents	(as % of Population)
American Indian/Alaska Native	3,203	259	8.1%	137	4.3%	5	0.2%
Asian	23,917	2,793	11.7%	1,075	4.5%	12	0.1%
Black/African American	60,909	39,873	65.5%	16,904	27.8%	605	1.0%
Hispanic or Latino	101,562	31,629	31.1%	11,846	11.7%	310	0.3%
Native Hawaiian/Pacific Islander	1,296	272	21.0%	106	8.2%	2	0.2%
White/Non-Hispanic	163,765	47,745	29.2%	18,334	11.2%	374	0.2%
		Chi-Squared Statistic:	22900.5		11596.5		814.1
		P-Value:	**0.000**		**0.000**		**0.000**

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.

[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.

[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 2.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
2019

Race/Ethnicity[1]	Interactions	Arrests	(as % of Interactions During The Period)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	90	51	56.7%	10.0	**0.002**
Asian	958	377	39.4%	0.0	0.841
Black/African American	13,570	6,094	44.9%	80.6	**0.000**
Hispanic or Latino	10,840	4,355	40.2%	0.5	0.476
Native Hawaiian/Pacific Islander	117	37	31.6%	2.9	0.091
Non-White (All)	25,575	10,914	42.7%	35.0	**0.000**
White/Non-Hispanic	16,092	6,394	39.7%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and Arrests show the total number of interactions and arrests, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 41.9 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 591.1 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
2019

Race/Ethnicity[1]	Arrests	UoF Incidents	(as % of Arrests During The Period)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	51	3	5.9%	1.8	0.179
Asian	377	6	1.6%	0.3	0.590
Black/African American	6,094	203	3.3%	16.2	**0.000**
Hispanic or Latino	4,355	102	2.3%	0.4	0.534
Native Hawaiian/Pacific Islander	37	0	0.0%	0.1	0.742
Non-White (All)	10,914	314	2.9%	8.3	**0.004**
White/Non-Hispanic	6,394	137	2.1%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and UoF Incidents show the total number of interactions and UoF incidents, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 7.2 and a p-value of 0.01 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 10.4 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
2019

Race/Ethnicity[1]	Interactions	UoF Incidents	(as % of Arrests During The Period)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	90	3	3.3%	3.9	**0.049**
Asian	958	6	0.6%	0.3	0.576
Black/African American	13,570	203	1.5%	26.4	**0.000**
Hispanic or Latino	10,840	102	0.9%	0.5	0.482
Native Hawaiian/Pacific Islander	117	0	0.0%	0.2	0.620
Non-White (All)	25,575	314	1.2%	12.7	**0.000**
White/Non-Hispanic	16,092	137	0.9%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and UoF Incidents show the total number of interactions and UoF incidents, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 3.5 and a p-value of 0.06 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 38.6 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.D
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
January 2018 - February 2021

Race/Ethnicity[1]	**Interactions**	**Arrests**	**(as % of Interactions During The Period)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	259	137	52.9%	22.3	**0.000**
Asian	2,793	1,075	38.5%	0.0	0.941
Black/African American	39,873	16,904	42.4%	144.0	**0.000**
Hispanic or Latino	31,629	11,846	37.5%	7.2	**0.007**
Native Hawaiian/Pacific Islander	272	106	39.0%	0.0	0.896
Non-White (All)	74,826	30,068	40.2%	38.7	**0.000**
White/Non-Hispanic	47,745	18,334	38.4%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.Interactions and Arrests show the total number of interactions and arrests, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 211.3 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 1357.9 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.E
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
January 2018 - February 2021

Race/Ethnicity[1]	Arrests	UoF Incidents	(as % of Arrests During The Period)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	137	5	3.6%	1.0	0.307
Asian	1,075	12	1.1%	4.0	**0.046**
Black/African American	16,904	605	3.6%	76.6	**0.000**
Hispanic or Latino	11,846	310	2.6%	10.6	**0.001**
Native Hawaiian/Pacific Islander	106	2	1.9%	0.1	0.815
Non-White (All)	30,068	934	3.1%	48.9	**0.000**
White/Non-Hispanic	18,334	374	2.0%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Arrests and UoF Incidents show the total number of arrests and UoF incidents, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 44.6 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 55.3 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.F
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
January 2018 - February 2021

Race/Ethnicity[1]	Interactions	UoF Incidents	(as % of Interactions During The Period)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	259	5	1.9%	3.0	0.084
Asian	2,793	12	0.4%	3.9	**0.048**
Black/African American	39,873	605	1.5%	105.3	**0.000**
Hispanic or Latino	31,629	310	1.0%	8.4	**0.004**
Native Hawaiian/Pacific Islander	272	2	0.7%	0.1	0.798
Non-White (All)	74,826	934	1.2%	59.2	**0.000**
White/Non-Hispanic	47,745	374	0.8%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Arrests and UoF Incidents show the total number of arrests and UoF incidents, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 23.1 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 145.6 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.G
Pearson Chi-Squared Test
Ratios of Arrests to Interactions
2019

Race/Ethnicity[1]	Interactions	Arrests	(as % of Interactions During The Period)
American Indian/Alaska Native	90	51	56.7%
Asian	958	377	39.4%
Black/African American	13,570	6,094	44.9%
Hispanic or Latino	10,840	4,355	40.2%
Native Hawaiian/Pacific Islander	117	37	31.6%
White/Non-Hispanic	16,092	6,394	39.7%
		Chi-Squared Statistic:	108.4
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and Arrests show the total number of interactions and arrests, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 2600.5 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 667.3 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.H
Pearson Chi-Squared Test
Ratios of UoF Incidents to Arrests
2019

Race/Ethnicity[1]	Arrests	UoF Incidents	(as % of Arrests During The Period)
American Indian/Alaska Native	51	3	5.9%
Asian	377	6	1.6%
Black/African American	6,094	203	3.3%
Hispanic or Latino	4,355	102	2.3%
Native Hawaiian/Pacific Islander	37	0	0.0%
White/Non-Hispanic	6,394	137	2.1%
		Chi-Squared Statistic:	23.9
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and UoF Incidents show the total number of interactions and UoF incidents, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 30.2 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 26.2 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.I
Pearson Chi-Squared Test
Ratios of UoF Incidents to Interactions
2019

Race/Ethnicity[1]	Interactions	UoF Incidents	(as % of Interactions During The Period)
American Indian/Alaska Native	90	3	3.3%
Asian	958	6	0.6%
Black/African American	13,570	203	1.5%
Hispanic or Latino	10,840	102	0.9%
Native Hawaiian/Pacific Islander	117	0	0.0%
White/Non-Hispanic	16,092	137	0.9%
		Chi-Squared Statistic:	39.1
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Interactions and Arrests show the total number of interactions and arrests, respectively, during 2019.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 96.1 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 68.2 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.J
Pearson Chi-Squared Test
Ratios of Arrests to Interactions
January 2018 - February 2021

Race/Ethnicity[1]	Interactions	Arrests	(as % of Interactions During The Period)
American Indian/Alaska Native	259	137	52.9%
Asian	2,793	1,075	38.5%
Black/African American	39,873	16,904	42.4%
Hispanic or Latino	31,629	11,846	37.5%
Native Hawaiian/Pacific Islander	272	106	39.0%
White/Non-Hispanic	47,745	18,334	38.4%
		Chi-Squared Statistic:	240.1
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.Interactions and Arrests show the total number of interactions and arrests, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 6835.4 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 1567.6 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.K
Pearson Chi-Squared Test
Ratios of UoF Incidents to Arrests
January 2018 - February 2021

Race/Ethnicity[1]	Arrests	UoF Incidents	(as % of Arrests During The Period)
American Indian/Alaska Native	137	5	3.6%
Asian	1,075	12	1.1%
Black/African American	16,904	605	3.6%
Hispanic or Latino	11,846	310	2.6%
Native Hawaiian/Pacific Islander	106	2	1.9%
White/Non-Hispanic	18,334	374	2.0%
		Chi-Squared Statistic:	91.4
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Arrests and UoF Incidents show the total number of arrests and UoF incidents, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 101.5 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 98.2 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 2.L
Pearson Chi-Squared Test
Ratios of UoF Incidents to Interactions
January 2018 - February 2021

Race/Ethnicity[1]	Interactions	UoF Incidents	(as % of Interactions During The Period)
American Indian/Alaska Native	259	5	1.9%
Asian	2,793	12	0.4%
Black/African American	39,873	605	1.5%
Hispanic or Latino	31,629	310	1.0%
Native Hawaiian/Pacific Islander	272	2	0.7%
White/Non-Hispanic	47,745	374	0.8%
		Chi-Squared Statistic:	128.1
		P-Value:	0.000

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.Interactions and Arrests show the total number of interactions and arrests, respectively, between January 2018 and February 2021.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Orange P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 285.4 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 222.5 and a p-value of 0.00 for a test of Non-White versus White/Non-Hispanic.

Figure 3.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By Use of Force Tier
January 2018 - February 2021

	Panel A: Tier 1					*Panel B: Tier 2 & 3*				
Race/Ethnicity[1]	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	259	2	0.8%	0.0	1.000	259	3	1.2%	7.1	**0.008**
Asian	2,793	8	0.3%	3.5	0.061	2,793	4	0.1%	0.2	0.622
Black/African American	39,873	419	1.1%	61.1	**0.000**	39,873	186	0.5%	45.1	**0.000**
Hispanic or Latino	31,629	207	0.7%	1.7	0.191	31,629	103	0.3%	10.4	**0.001**
Native Hawaiian/Pacific Islander	272	2	0.7%	0.0	0.952	272	0	0.0%	0.0	0.941
Non-White (All)	74,826	638	0.9%	29.3	**0.000**	74,826	296	0.4%	32.4	**0.000**
White/Non-Hispanic	47,745	276	0.6%	N/A	N/A	47,745	98	0.2%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 9.1, 16.7 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 81.3, 66.5 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 3.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By Use of Force Tier
January 2018 - February 2021

	Panel A: Tier 1					*Panel B: Tier 2 & 3*				
Race/Ethnicity[1]	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	137	2	1.5%	0.1	0.758	137	3	2.2%	4.1	**0.042**
Asian	1,075	8	0.7%	3.6	0.059	1,075	4	0.4%	0.2	0.618
Black/African American	16,904	419	2.5%	42.6	**0.000**	16,904	186	1.1%	34.5	**0.000**
Hispanic or Latino	11,846	207	1.7%	2.5	0.112	11,846	103	0.9%	11.7	**0.001**
Native Hawaiian/Pacific Islander	106	2	1.9%	0.0	0.938	106	0	0.0%	0.0	0.932
Non-White (All)	30,068	638	2.1%	23.0	**0.000**	30,068	296	1.0%	28.0	**0.000**
White/Non-Hispanic	18,334	276	1.5%	N/A	N/A	18,334	98	0.5%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 21.0, 25.7 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 26.0, 31.8 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 4.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By District
January 2018 - February 2021

	Panel A: District 1					*Panel B: District 2*					*Panel C: District 3*				
Race/Ethnicity[1]	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	131	0	0.0%	0.4	0.547	104	5	4.8%	12.4	**0.000**	22	0	0.0%	1.2	0.275
Asian	1,271	9	0.7%	0.2	0.642	778	1	0.1%	4.6	**0.033**	720	2	0.3%	0.5	0.472
Black/African American	20,006	327	1.6%	42.2	**0.000**	12,755	194	1.5%	18.8	**0.000**	6,806	76	1.1%	20.2	**0.000**
Hispanic or Latino	15,017	155	1.0%	2.0	0.161	11,352	110	1.0%	0.0	0.834	5,027	41	0.8%	3.9	**0.048**
Native Hawaiian/Pacific Islander	140	1	0.7%	0.1	0.801	98	1	1.0%	0.2	0.660	33	0	0.0%	0.6	0.453
Non-White (All)	36,565	492	1.3%	21.9	**0.000**	25,087	311	1.2%	7.3	**0.007**	12,608	119	0.9%	14.2	**0.000**
White/Non-Hispanic	17,598	154	0.9%	N/A	N/A	14,323	134	0.9%	N/A	N/A	15,416	85	0.6%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] Districts "OD" and "PCW" are not shown due to small sample sizes.

[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[6] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 10.1, 1.1, 3.7 and a p-value of 0.00, 0.30, 0.05 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 58.2, 30.1, 28.4 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 4.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By District
January 2018 - February 2021

	Panel A: District 1					*Panel B: District 2*					*Panel C: District 3*				
Race/Ethnicity[1]	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	62	0	0.0%	0.6	0.451	65	5	7.7%	6.7	**0.009**	8	0	0.0%	1.0	0.318
Asian	523	9	1.7%	0.4	0.543	310	1	0.3%	4.0	**0.045**	232	2	0.9%	0.5	0.477
Black/African American	8,997	327	3.6%	25.9	**0.000**	5,504	194	3.5%	19.7	**0.000**	2,270	76	3.3%	18.9	**0.000**
Hispanic or Latino	5,687	155	2.7%	3.0	0.082	4,399	110	2.5%	1.2	0.269	1,653	41	2.5%	3.7	0.055
Native Hawaiian/Pacific Islander	48	1	2.1%	0.2	0.672	42	1	2.4%	0.2	0.665	16	0	0.0%	0.2	0.657
Non-White (All)	15,317	492	3.2%	16.0	**0.000**	10,320	311	3.0%	10.6	**0.001**	4,179	119	2.8%	13.3	**0.000**
White/Non-Hispanic	6,914	154	2.2%	N/A	N/A	6,219	134	2.2%	N/A	N/A	5,004	85	1.7%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Districts "OD" and "PCW" are not shown due to small sample sizes.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 14.9, 9.4, 11.8 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 18.3, 12.7, 14.3 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 4.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Arrests to Interactions
By District
January 2018 - February 2021

	Panel A: District 1					*Panel B: District 2*					*Panel C: District 3*				
Race/Ethnicity[1]	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	131	62	47.3%	3.2	0.074	104	65	62.5%	14.5	**0.000**	22	8	36.4%	0.0	0.871
Asian	1,271	523	41.1%	1.6	0.200	778	310	39.8%	3.7	0.055	720	232	32.2%	0.0	0.926
Black/African American	20,006	8,997	45.0%	123.6	**0.000**	12,755	5,504	43.2%	0.2	0.666	6,806	2,270	33.4%	1.7	0.196
Hispanic or Latino	15,017	5,687	37.9%	6.8	**0.009**	11,352	4,399	38.8%	56.7	**0.000**	5,027	1,653	32.9%	0.3	0.591
Native Hawaiian/Pacific Islander	140	48	34.3%	1.3	0.262	98	42	42.9%	0.0	0.992	33	16	48.5%	3.2	0.076
Non-White (All)	36,565	15,317	41.9%	33.1	**0.000**	25,087	10,320	41.1%	19.4	**0.000**	12,608	4,179	33.1%	1.4	0.229
White/Non-Hispanic	17,598	6,914	39.3%	N/A	N/A	14,323	6,219	43.4%	N/A	N/A	15,416	5,004	32.5%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Districts "OD" and "PCW" are not shown due to small sample sizes.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 26.6, 264.8, 142.0 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 701.6, 243.3, 147.4 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 5.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By Misdemeanor and Felony
January 2018 - February 2021

	Panel A: Misdemeanor					*Panel B: Felony*				
Race/Ethnicity[1]	Interactions	UoF Incidents	(as % of Interactions)	Chi-Squared Statistic	P-Value	Interactions	UoF Incidents	(as % of Interactions)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	179	2	1.1%	0.0	0.862	70	3	4.3%	4.1	**0.042**
Asian	2,028	8	0.4%	2.5	0.111	604	4	0.7%	0.5	0.470
Black/African American	27,751	395	1.4%	72.4	**0.000**	10,158	193	1.9%	24.9	**0.000**
Hispanic or Latino	21,766	184	0.8%	2.3	0.126	7,785	114	1.5%	5.8	**0.016**
Native Hawaiian/Pacific Islander	200	2	1.0%	0.0	0.970	62	0	0.0%	0.0	0.851
Non-White (All)	51,924	591	1.1%	36.2	**0.000**	18,679	314	1.7%	17.9	**0.000**
White/Non-Hispanic	34,591	251	0.7%	N/A	N/A	10,517	111	1.1%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 10.3, 15.2 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 102.3, 23.2 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 5.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By Misdemeanor and Felony
January 2018 - February 2021

	Panel A: Misdemeanor					*Panel B: Felony*				
Race/Ethnicity[1]	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	100	2	2.0%	0.0	0.874	33	3	9.1%	1.3	0.261
Asian	897	8	0.9%	2.9	0.087	143	4	2.8%	0.1	0.704
Black/African American	12,607	395	3.1%	59.7	**0.000**	3,685	193	5.2%	7.6	**0.006**
Hispanic or Latino	8,969	184	2.1%	3.6	0.058	2,281	114	5.0%	4.3	**0.038**
Native Hawaiian/Pacific Islander	86	2	2.3%	0.0	0.978	16	0	0.0%	0.0	0.893
Non-White (All)	22,659	591	2.6%	32.9	**0.000**	6,158	314	5.1%	7.5	**0.006**
White/Non-Hispanic	14,746	251	1.7%	N/A	N/A	2,936	111	3.8%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[4]It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 29.6, 7.5 and a p-value of 0.00, 0.01 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 37.9, 7.4 and a p-value of 0.00, 0.01 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 5.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
By Misdemeanor and Felony
January 2018 - February 2021

	Panel A: Misdemeanor					*Panel B: Felony*				
Race/Ethnicity[1]	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	179	100	55.9%	12.2	**0.000**	70	33	47.1%	11.8	**0.001**
Asian	2,028	897	44.2%	1.9	0.163	604	143	23.7%	4.9	**0.027**
Black/African American	27,751	12,607	45.4%	48.9	**0.000**	10,158	3,685	36.3%	165.5	**0.000**
Hispanic or Latino	21,766	8,969	41.2%	11.0	**0.001**	7,785	2,281	29.3%	4.1	**0.042**
Native Hawaiian/Pacific Islander	200	86	43.0%	0.0	0.973	62	16	25.8%	0.1	0.820
Non-White (All)	51,924	22,659	43.6%	8.6	**0.003**	18,679	6,158	33.0%	79.8	**0.000**
White/Non-Hispanic	34,591	14,746	42.6%	N/A	N/A	10,517	2,936	27.9%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 384.7, 48.0 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 1123.5, 151.1 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 6.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By Case Types "DISTR-," "DOMES-," and "SUSPO"
January 2018 - February 2021

	Panel A: DISTR-					*Panel B: DOMES-*					*Panel C: SUSPO*				
Race/Ethnicity[1]	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	6	1	16.7%	0.8	0.379	11	1	9.1%	0.7	0.403	18	0	0.0%	0.1	0.720
Asian	61	1	1.6%	0.0	0.982	70	1	1.4%	0.3	0.614	69	1	1.4%	0.1	0.756
Black/African American	1,498	51	3.4%	1.4	0.230	1,692	29	1.7%	0.2	0.634	1,529	65	4.3%	17.0	**0.000**
Hispanic or Latino	744	14	1.9%	0.6	0.422	1,191	14	1.2%	0.2	0.665	1,014	27	2.7%	2.3	0.126
Native Hawaiian/Pacific Islander	10	0	0.0%	0.3	0.617	7	0	0.0%	1.6	0.203	21	1	4.8%	0.0	0.824
Non-White (All)	2,319	67	2.9%	0.2	0.623	2,971	45	1.5%	0.0	0.951	2,651	94	3.5%	11.7	**0.001**
White/Non-Hispanic	1,257	32	2.5%	N/A	N/A	1,526	22	1.4%	N/A	N/A	1,636	28	1.7%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] "DISTR-" is a call type associated with a disturbance/noise complaint. "DOMES-" is a call type associated with a domestic dispute. "SUSPO" is a call type associated with suspicious activity.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 0.1, 0.0, 10.6 and a p-value of 0.72, 1.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 0.5, 0.4, 13.4 and a p-value of 0.46, 0.85, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 6.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By Case Types "DISTR-," "DOMES-," and "SUSPO"
January 2018 - February 2021

	Panel A: DISTR-					*Panel B: DOMES-*					*Panel C: SUSPO*				
Race/Ethnicity[1]	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	5	1	20.0%	0.0	0.863	6	1	16.7%	0.2	0.621	12	0	0.0%	0.0	0.953
Asian	10	1	10.0%	0.1	0.735	27	1	3.7%	0.1	0.699	28	1	3.6%	0.2	0.683
Black/African American	589	51	8.7%	0.1	0.707	658	29	4.4%	0.0	0.931	836	65	7.8%	9.6	**0.002**
Hispanic or Latino	232	14	6.0%	0.4	0.504	461	14	3.0%	0.6	0.451	454	27	5.9%	2.2	0.142
Native Hawaiian/Pacific Islander	4	0	0.0%	0.1	0.718	3	0	0.0%	1.2	0.272	5	1	20.0%	0.5	0.494
Non-White (All)	840	67	8.0%	0.0	0.996	1,161	45	3.9%	0.0	0.905	1,335	94	7.0%	7.7	**0.006**
White/Non-Hispanic	411	32	7.8%	N/A	N/A	532	22	4.1%	N/A	N/A	717	28	3.9%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] "DISTR-" is a call type associated with a disturbance/noise complaint. "DOMES-" is a call type associated with a domestic dispute. "SUSPO" is a call type associated with suspicious activity.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 0.0, 0.1, 7.4 and a p-value of 1.00, 0.90, 0.01 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 0.0, 0.0, 8.0 and a p-value of 0.99, 0.96, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 6.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
By Case Types "DISTR-," "DOMES-," and "SUSPO"
January 2018 - February 2021

	Panel A: DISTR-					*Panel B: DOMES-*					*Panel C: SUSPO*				
Race/Ethnicity[1]	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	6	5	83.3%	4.8	**0.028**	11	6	54.5%	1.1	0.295	18	12	66.7%	2.9	0.089
Asian	61	10	16.4%	6.4	**0.012**	70	27	38.6%	0.3	0.611	69	28	40.6%	0.2	0.683
Black/African American	1,498	589	39.3%	12.7	**0.000**	1,692	658	38.9%	5.4	**0.020**	1,529	836	54.7%	36.8	**0.000**
Hispanic or Latino	744	232	31.2%	0.4	0.515	1,191	461	38.7%	4.1	**0.043**	1,014	454	44.8%	0.2	0.662
Native Hawaiian/Pacific Islander	10	4	40.0%	0.0	0.879	7	3	42.9%	0.0	0.964	21	5	23.8%	2.6	0.106
Non-White (All)	2,319	840	36.2%	4.3	**0.038**	2,971	1,155	38.9%	6.8	**0.009**	2,651	1,335	50.4%	17.0	**0.000**
White/Non-Hispanic	1,257	411	32.7%	N/A	N/A	1,526	532	34.9%	N/A	N/A	1,636	717	43.8%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] "DISTR-" is a call type associated with a disturbance/noise complaint. "DOMES-" is a call type associated with a domestic dispute. "SUSPO" is a call type associated with suspicious activity.

[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 2.3, 5.3, 11.8 and a p-value of 0.13, 0.02, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 9.5, 10.0, 26.6 and a p-value of 0.00, 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 7.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By Male/Female
January 2018 - February 2021

	Panel A: Female					*Panel B: Male*				
Race/Ethnicity[1]	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**UoF Incidents**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	107	2	1.9%	2.4	0.125	152	3	2.0%	0.5	0.474
Asian	1,156	2	0.2%	1.1	0.288	1,637	10	0.6%	2.5	0.114
Black/African American	15,378	105	0.7%	10.5	**0.001**	24,490	500	2.0%	84.9	**0.000**
Hispanic or Latino	12,783	45	0.4%	0.9	0.352	18,832	265	1.4%	11.9	**0.001**
Native Hawaiian/Pacific Islander	116	0	0.0%	0.0	0.988	156	2	1.3%	0.0	0.914
Non-White (All)	29,540	154	0.5%	2.2	0.142	45,267	780	1.7%	53.4	**0.000**
White/Non-Hispanic	20,262	86	0.4%	N/A	N/A	27,477	288	1.0%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the Non-White category results in a chi-squared statistic of 0.1, 24.0 and a p-value of 0.91, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 11.3, 123.6 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 7.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By Male/Female
January 2018 - February 2021

	Panel A: Female					*Panel B: Male*				
Race/Ethnicity[1]	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**	**Arrests**	**UoF Incidents**	**(as % of Arrests)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	43	2	4.7%	1.5	0.216	94	3	3.2%	0.0	0.897
Asian	362	2	0.6%	1.0	0.311	713	10	1.4%	2.7	0.099
Black/African American	4,992	105	2.1%	10.5	**0.001**	11,910	500	4.2%	55.9	**0.000**
Hispanic or Latino	3,465	45	1.3%	0.0	0.966	8,380	265	3.2%	9.1	**0.003**
Native Hawaiian/Pacific Islander	26	0	0.0%	0.1	0.781	80	2	2.5%	0.1	0.740
Non-White (All)	8,888	154	1.7%	4.1	**0.042**	21,177	780	3.7%	36.5	**0.000**
White/Non-Hispanic	6,562	86	1.3%	N/A	N/A	11,769	288	2.4%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 3.7, 33.3 and a p-value of 0.05, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 4.6, 41.9 and a p-value of 0.03, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 7.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
By Male/Female
January 2018 - February 2021

	Panel A: Female					*Panel B: Male*				
Race/Ethnicity[1]	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**	**Interactions**	**Arrests**	**(as % of Interactions)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	107	43	40.2%	2.6	0.106	152	94	61.8%	21.5	**0.000**
Asian	1,156	362	31.3%	0.5	0.469	1,637	713	43.6%	0.3	0.583
Black/African American	15,378	4,992	32.5%	0.0	0.888	24,490	11,910	48.6%	175.4	**0.000**
Hispanic or Latino	12,783	3,465	27.1%	103.1	**0.000**	18,832	8,380	44.5%	12.6	**0.000**
Native Hawaiian/Pacific Islander	116	26	22.4%	4.8	**0.029**	156	80	51.3%	4.2	**0.041**
Non-White (All)	29,540	8,888	30.1%	29.5	**0.000**	45,267	21,177	46.8%	107.5	**0.000**
White/Non-Hispanic	20,262	6,562	32.4%	N/A	N/A	27,477	11,769	42.8%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.

[4] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[5] The above analysis excludes observations that belong to the "Unknown" race/ethnicity subgroup. Including the "Unknown" observations in the Non-White category results in a chi-squared statistic of 308.4, 38.6 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic. Including the "Unknown" observations in the White/Non-Hispanic category results in a chi-squared statistic of 140.4, 1256.4 and a p-value of 0.00, 0.00 (respectively) for a test of Non-White versus White/Non-Hispanic.

Figure 8.A
Chi-Squared Tests
"Black/African American" Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Interactions
By Age Range
January 2018 - February 2021

Age Range	Interactions	UoF Incidents	(as % of Interactions)	Chi-Squared Statistic	P-Value
Black/African American:					
18-21	4,756	76	1.6%	12.1	**0.000**
22-29	10,024	208	2.1%	24.1	**0.000**
30-49	17,923	287	1.6%	49.7	**0.000**
50-64	5,803	31	0.5%	0.6	0.433
65-98	1,125	2	0.2%	0.1	0.732
White/Non-Hispanic:					
18-21	3,482	25	0.7%	N/A	N/A
22-29	9,745	115	1.2%	N/A	N/A
30-49	22,786	191	0.8%	N/A	N/A
50-64	8,192	35	0.4%	N/A	N/A
65-98	3,210	6	0.2%	N/A	N/A

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] Age range group "Over 98" is not shown because of the small number of interactions and UoF incidents for that group.

[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate.

[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).

[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 23.4, 52.8, 93.4, 3.7, 0.0 and a p-value of 0.00, 0.00, 0.00, 0.05, 1.00 (respectively).

Figure 8.B
Chi-Squared Tests
"Black/African American" Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Arrests
By Age Range
January 2018 - February 2021

Age Range	Arrests	UoF Incidents	(as % of Arrests)	Chi-Squared Statistic	P-Value
Black/African American:					
18-21	2,206	76	3.4%	16.0	**0.000**
22-29	4,866	208	4.3%	21.9	**0.000**
30-49	7,217	287	4.0%	34.9	**0.000**
50-64	2,241	31	1.4%	0.0	0.983
65-98	329	2	0.6%	0.1	0.715
White/Non-Hispanic:					
18-21	1,788	25	1.4%	N/A	N/A
22-29	4,589	115	2.5%	N/A	N/A
30-49	8,251	191	2.3%	N/A	N/A
50-64	2,622	35	1.3%	N/A	N/A
65-98	1,027	6	0.6%	N/A	N/A

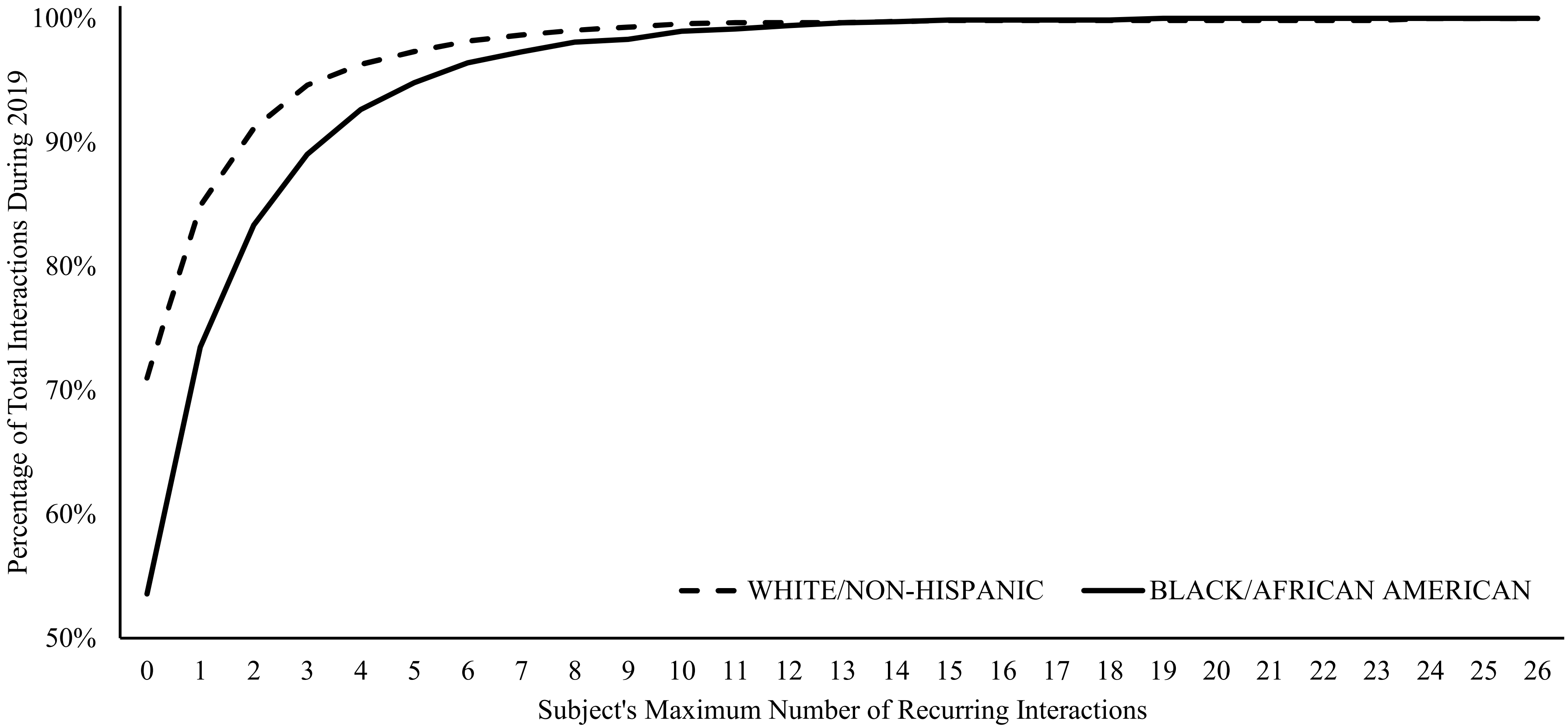
Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Age range group "Over 98" is not shown because of the small number of arrests and UoF incidents for that group.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 18.3, 24.5, 37.1, 0.1, 0.0 and a p-value of 0.00, 0.00, 0.00, 0.91, 1.00 (respectively).

Figure 8.C
Chi-Squared Tests
"Black/African American" Against "White/Non-Hispanic" Group
Ratios of Arrests to Interactions
By Age Range
January 2018 - February 2021

Age Range	Interactions	Arrests	(as % of Interactions)	Chi-Squared Statistic	P-Value
Black/African American:					
18-21	4,756	2,206	46.4%	19.7	0.000
22-29	10,024	4,866	48.5%	4.1	0.042
30-49	17,923	7,217	40.3%	69.9	0.000
50-64	5,803	2,241	38.6%	65.2	0.000
65-98	1,125	329	29.2%	2.8	0.094
White/Non-Hispanic:					
18-21	3,482	1,788	51.3%	N/A	N/A
22-29	9,745	4,589	47.1%	N/A	N/A
30-49	22,786	8,251	36.2%	N/A	N/A
50-64	8,192	2,622	32.0%	N/A	N/A
65-98	3,210	1,027	32.0%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[3] Age range group "Over 98" is not shown because of the small number of interactions and UoF incidents for that group.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
[6] The above analysis excludes observations that belong to the “Unknown” race/ethnicity subgroup. Including the “Unknown” observations in the White/Non-Hispanic category results in a chi-squared statistic of 25.9, 274.6, 588.3, 366.7, 11.2 and a p-value of 0.00, 0.00, 0.00, 0.00, 0.00 (respectively).

Chart 1
Percentage of Total Interactions
By Race/Ethnicity and Number of Recurring Interactions Per Unique Subject
2019



Notes:
[1] Lines show the respective number of interactions by race/ethnicity involving unique individuals ("Subjects") with a given number of recurring interactions or less during 2019 as a percentage of total interactions during 2019. Unique individuals are identified using the variable 'pin' (person id number).
[2] It should be noted that due to the incompleteness of the available data prior to 2019, the data underlying this graph has been *truncated* and therefore does not capture preceding police interactions per unique subject prior to 2019. In addition, it should be further noted that due to a lack of data related to the allocation of police resources, this graph does not adjust for the relative population composition overall between White subjects versus Black/African American subjects, which might result in substantive modifications to this graph.

Figure 9.A
Likelihood of Arrests and UoF Incidents Conditional on Interactions
Estimation Period: January 2018 - February 2021

Panel A: Arrest Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	-0.33	0.00	-80.69	**0.000**
race/ethnicity binary indicator variable:				
AMERICAN INDIAN/ALASKAN N	0.83	0.05	17.52	**0.000**
ASIAN	-0.10	0.02	-5.38	**0.000**
BLACK/AFRICAN AMERICAN	0.45	0.01	76.54	**0.000**
HISPANIC OR LATINO	0.38	0.01	61.33	**0.000**
NATIVE HAWAIIAN/PACIFIC ISLANDER	0.82	0.07	11.37	**0.000**

Panel B: UoF Incident Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	-4.55	0.02	-229.97	**0.000**
race/ethnicity binary indicator variable:				
AMERICAN INDIAN/ALASKAN N	0.38	0.19	2.04	**0.041**
ASIAN	-0.45	0.11	-4.11	**0.000**
BLACK/AFRICAN AMERICAN	0.60	0.02	23.93	**0.000**
HISPANIC OR LATINO	0.21	0.03	7.42	**0.000**
NATIVE HAWAIIAN/PACIFIC ISLANDER	-0.36	0.41	-0.87	0.383

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] The ethnicity category "Unknown" has been excluded from the analysis presented above.

[4] Orange P-Values indicate statistical significance at the 5% level.

Figure 9.B
Likelihood of Arrests and UoF Incidents Conditional on Interactions
Estimation Period: January 2018 - February 2021

Panel A: Arrest Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	0.05	0.02	5.93	**0.015**
race/ethnicity dummies:				
AMERICAN INDIAN/ALASKAN N	0.68	0.09	52.37	**0.000**
ASIAN	0.00	0.04	0.01	0.908
BLACK/AFRICAN AMERICAN	0.33	0.02	204.03	**0.000**
HISPANIC OR LATINO	0.26	0.02	126.33	**0.000**
NATIVE HAWAIIAN/PACIFIC ISLANDER	0.47	0.08	30.98	**0.000**
UNKNOWN	-1.77	0.03	3962.94	**0.000**

Panel B: UoF Incident Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	-4.10	0.08	2609.55	**0.000**
race/ethnicity dummies:				
AMERICAN INDIAN/ALASKAN N	0.96	0.23	17.44	**0.000**
ASIAN	-0.62	0.17	13.14	**0.000**
BLACK/AFRICAN AMERICAN	0.84	0.08	105.00	**0.000**
HISPANIC OR LATINO	0.39	0.08	21.25	**0.000**
NATIVE HAWAIIAN/PACIFIC ISLANDER	-0.40	0.39	1.07	0.301
UNKNOWN	-1.43	0.13	119.70	**0.000**

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[4] Orange P-Values indicate statistical significance at the 5% level.

Figure 9.C
Likelihood of Arrests and UoF Incidents Conditional on Interactions
Estimation Period: January 2018 - February 2021

Panel A: Arrest Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	-0.33	0.00	-80.17	0.000
race/ethnicity binary indicator variable:				
AMERICAN INDIAN/ALASKAN N	0.83	0.05	17.46	0.000
ASIAN	-0.10	0.02	-5.53	0.000
BLACK/AFRICAN AMERICAN	0.44	0.01	76.14	0.000
HISPANIC OR LATINO	0.37	0.01	60.93	0.000

Panel B: UoF Incident Likelihood

Variable	Coefficient	Standard Error	Wald Test Statistic	P-Value
Intercept	-4.55	0.02	-230.28	0.000
race/ethnicity binary indicator variable:				
AMERICAN INDIAN/ALASKAN N	0.38	0.19	2.05	0.041
ASIAN	-0.45	0.11	-4.10	0.000
BLACK/AFRICAN AMERICAN	0.60	0.02	23.99	0.000
HISPANIC OR LATINO	0.21	0.03	7.45	0.000

Notes:

[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'

[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.

[3] Orange P-Values indicate statistical significance at the 5% level.

Figure 10.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique Interactions to Population Counts
2019

Race/Ethnicity[1]	Population	Unique Interactions	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	75	2.3%	128.4	0.000
Asian	23,917	838	3.5%	573.7	0.000
Black/African American	60,909	9,485	15.6%	2,506.5	0.000
Hispanic or Latino	101,562	8,751	8.6%	26.1	0.000
Native Hawaiian/Pacific Islander	1,296	89	6.9%	2.2	0.141
Non-White (All)	190,887	19,238	10.1%	404.7	0.000
White/Non-Hispanic	163,765	13,151	8.0%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 10.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique Interactions to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Unique Interactions	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	216	6.7%	392.1	0.000
Asian	23,917	2,443	10.2%	1,763.1	0.000
Black/African American	60,909	28,084	46.1%	7,322.7	0.000
Hispanic or Latino	101,562	25,610	25.2%	44.9	0.000
Native Hawaiian/Pacific Islander	1,296	218	16.8%	27.0	0.000
Non-White (All)	190,887	56,571	29.6%	1,076.7	0.000
White/Non-Hispanic	163,765	39,131	23.9%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 10.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique Arrests to Population Counts
2019

Race/Ethnicity[1]	Population	Unique Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	46	1.4%	34.8	0.000
Asian	23,917	343	1.4%	247.3	0.000
Black/African American	60,909	4,499	7.4%	1,623.7	0.000
Hispanic or Latino	101,562	3,691	3.6%	14.3	0.000
Native Hawaiian/Pacific Islander	1,296	34	2.6%	2.0	0.152
Non-White (All)	190,887	8,613	4.5%	297.5	0.000
White/Non-Hispanic	163,765	5,492	3.4%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 10.D
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique Arrests to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Unique Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	117	3.7%	118.8	0.000
Asian	23,917	974	4.1%	728.2	0.000
Black/African American	60,909	12,570	20.6%	4,233.7	0.000
Hispanic or Latino	101,562	10,055	9.9%	3.7	0.054
Native Hawaiian/Pacific Islander	1,296	95	7.3%	7.2	0.007
Non-White (All)	190,887	23,811	12.5%	624.5	0.000
White/Non-Hispanic	163,765	15,820	9.7%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 10.E
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique UoF Incidents to Population Counts
2019

Race/Ethnicity[1]	Population	Unique UoF Incidents	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	2	0.1%	0.1	0.727
Asian	23,917	6	0.0%	8.6	**0.003**
Black/African American	60,909	192	0.3%	170.9	**0.000**
Hispanic or Latino	101,562	97	0.1%	1.8	0.185
Native Hawaiian/Pacific Islander	1,296	0	0.0%	1.0	0.309
Non-White (All)	190,887	297	0.2%	41.7	**0.000**
White/Non-Hispanic	163,765	131	0.1%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 10.F
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Unique UoF Incidents to Population Counts
January 2018 - February 2021

Race/Ethnicity[1]	Population	Unique UoF Incidents	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,203	4	0.1%	1.3	0.257
Asian	23,917	12	0.1%	30.2	**0.000**
Black/African American	60,909	579	1.0%	568.8	**0.000**
Hispanic or Latino	101,562	296	0.3%	13.2	**0.000**
Native Hawaiian/Pacific Islander	1,296	2	0.2%	0.2	0.619
Non-White (All)	190,887	893	0.5%	154.3	**0.000**
White/Non-Hispanic	163,765	359	0.2%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence. Unique Subjects with At Least One Interaction, Unique Subjects with At Least One Arrest, and Unique Subjects with At Least One UoF Incident show the number of unique individuals ("Subjects") with one or more interactions, arrests, and UoF incidents, respectively, during 2019. Unique subjects are identified using the variable 'pin' (person id number).
[3] Unknown race category is not shown because population counts and frequencies of interaction/arrest/UoF incident are not comparable for that category.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for the City of Aurora (Census Place 08:04000).

Figure 11.A
Zip Code Summary by Zip Code Median Income Quartiles

	Income Quartile 1		*Income Quartile 2*		*Income Quartile 3*		*Income Quartile 4*	
	Lower Bound	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**	**Lower Bound**	**Upper Bound**
Median Household Income	$45,431	$58,992	$60,725	$70,778	$79,711	$83,186	$105,658	$136,144
Total Population	270,605		275,054		190,969		318,988	
Zip Codes	80010		80231		80019		80015	
	80011		80014		80249		80016	
	80012		80017		80013		80111	
	80045		80022		80018		80112	
	80216		80215		80102		80113	
	80219		80222		80137		80124	
	80223		80224		80207		80138	
	80247		80239		80220		80210	
	80645		80246		80230		80238	

Notes:
[1] Income Quartiles 1 - 4 represent Income Quartiles with the lowest (1) to highest (4) median incomes.
[2] Income is the median household income per the ACS.
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 11.B
Demographic Summary by Zip Code Median Income Quartiles

Race/Ethnicity[1]	*Income Quartile 1* Population	*Income Quartile 1* (as % of Total Population)	*Income Quartile 2* Population	*Income Quartile 2* (as % of Total Population)	*Income Quartile 3* Population	*Income Quartile 3* (as % of Total Population)	*Income Quartile 4* Population	*Income Quartile 4* (as % of Total Population)
American Indian/Alaska Native	3,736	1.4%	2,425	0.9%	1,211	0.6%	1,365	0.4%
Asian	14,081	5.2%	11,592	4.2%	10,369	5.4%	24,165	7.6%
Black/African American	39,480	14.6%	38,534	14.0%	31,534	16.5%	15,201	4.8%
Hispanic or Latino	130,076	48.1%	86,893	31.6%	43,889	23.0%	32,467	10.2%
Native Hawaiian/Pacific Islander	415	0.2%	785	0.3%	301	0.2%	466	0.1%
Non-White (All)	187,788	69.4%	140,230	51.0%	87,304	45.7%	73,664	23.1%
White/Non-Hispanic	82,817	30.6%	134,824	49.0%	103,665	54.3%	245,324	76.9%
Total Population	**270,605**		**275,054**		**190,969**		**318,988**	

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 12.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Interactions to Population Counts
By Zip Code Median Income Quartiles
2019

	Panel A: Income Quartile 1					*Panel B: Income Quartile 2*				
Race/Ethnicity[1]	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	3,736	72	1.9%	247.1	0.000	2,425	12	0.5%	36.2	0.000
Asian	14,081	539	3.8%	522.3	0.000	11,592	187	1.6%	27.3	0.000
Black/African American	39,480	9,421	23.9%	3,441.6	0.000	38,534	2,163	5.6%	1,007.6	0.000
Hispanic or Latino	130,076	7,838	6.0%	1,133.4	0.000	86,893	1,407	1.6%	148.2	0.000
Native Hawaiian/Pacific Islander	415	91	21.9%	55.9	0.000	785	11	1.4%	3.2	0.075
Non-White (All)	187,788	17,961	9.6%	21.0	0.000	140,230	3,780	2.7%	26.3	0.000
White/Non-Hispanic	82,817	8,415	10.2%	N/A	N/A	134,824	3,214	2.4%	N/A	N/A

	Panel C: Income Quartile 3					*Panel D: Income Quartile 4*				
Race/Ethnicity[1]	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	1,211	4	0.3%	19.3	0.000	1,365	2	0.1%	8.4	0.004
Asian	10,369	122	1.2%	47.1	0.000	24,165	110	0.5%	47.4	0.000
Black/African American	31,534	1,294	4.1%	329.9	0.000	15,201	687	4.5%	1,731.6	0.000
Hispanic or Latino	43,889	1,039	2.4%	3.7	0.055	32,467	550	1.7%	193.7	0.000
Native Hawaiian/Pacific Islander	301	10	3.3%	1.7	0.193	466	5	1.1%	0.2	0.660
Non-White (All)	87,304	2,469	2.8%	74.3	0.000	73,664	1,354	1.8%	470.0	0.000
White/Non-Hispanic	103,665	2,284	2.2%	N/A	N/A	245,324	2,163	0.9%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 12.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Population Counts
By Zip Code Median Income Quartiles
2019

	Panel A: Income Quartile 1					*Panel B: Income Quartile 2*				
Race/Ethnicity[1]	**Population**	**Arrests**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Arrests**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	3,736	43	1.2%	94.6	0.000	2,425	4	0.2%	14.7	0.000
Asian	14,081	236	1.7%	243.4	0.000	11,592	61	0.5%	17.6	0.000
Black/African American	39,480	4,643	11.8%	2,002.9	0.000	38,534	809	2.1%	364.1	0.000
Hispanic or Latino	130,076	3,296	2.5%	632.9	0.000	86,893	531	0.6%	58.3	0.000
Native Hawaiian/Pacific Islander	415	25	6.0%	1.9	0.169	785	5	0.6%	0.6	0.429
Non-White (All)	187,788	8,243	4.4%	4.4	0.036	140,230	1,410	1.0%	7.1	0.008
White/Non-Hispanic	82,817	3,788	4.6%	N/A	N/A	134,824	1,222	0.9%	N/A	N/A

	Panel C: Income Quartile 3					*Panel D: Income Quartile 4*				
Race/Ethnicity[1]	**Population**	**Arrests**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Arrests**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	1,211	3	0.2%	3.0	0.082	1,365	1	0.1%	2.1	0.144
Asian	10,369	43	0.4%	8.5	0.004	24,165	37	0.2%	13.9	0.000
Black/African American	31,534	425	1.3%	143.2	0.000	15,201	214	1.4%	516.9	0.000
Hispanic or Latino	43,889	342	0.8%	7.1	0.008	32,467	181	0.6%	67.8	0.000
Native Hawaiian/Pacific Islander	301	3	1.0%	0.5	0.462	466	4	0.9%	5.4	0.020
Non-White (All)	87,304	816	0.9%	48.1	0.000	73,664	437	0.6%	152.3	0.000
White/Non-Hispanic	103,665	677	0.7%	N/A	N/A	245,324	697	0.3%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities* . Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 12.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Population Counts
By Zip Code Median Income Quartiles
2019

	Panel A: Income Quartile 1					*Panel B: Income Quartile 2*				
Race/Ethnicity[1]	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	3,736	3	0.1%	0.1	0.737	2,425	0	0.0%	0.5	0.486
Asian	14,081	4	0.0%	6.6	**0.010**	11,592	0	0.0%	2.3	0.128
Black/African American	39,480	159	0.4%	126.7	**0.000**	38,534	23	0.1%	16.3	**0.000**
Hispanic or Latino	130,076	76	0.1%	10.6	**0.001**	86,893	15	0.0%	0.2	0.644
Native Hawaiian/Pacific Islander	415	0	0.0%	0.4	0.524	785	0	0.0%	0.2	0.692
Non-White (All)	187,788	242	0.1%	4.6	**0.031**	140,230	38	0.0%	1.5	0.228
White/Non-Hispanic	82,817	81	0.1%	N/A	N/A	134,824	27	0.0%	N/A	N/A

	Panel C: Income Quartile 3					*Panel D: Income Quartile 4*				
Race/Ethnicity[1]	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	1,211	0	0.0%	0.2	0.675	1,365	0	0.0%	0.1	0.780
Asian	10,369	2	0.0%	0.1	0.702	24,165	0	0.0%	1.4	0.240
Black/African American	31,534	16	0.1%	13.9	**0.000**	15,201	5	0.0%	14.5	**0.000**
Hispanic or Latino	43,889	8	0.0%	0.3	0.597	32,467	3	0.0%	0.6	0.444
Native Hawaiian/Pacific Islander	301	0	0.0%	0.0	0.835	466	0	0.0%	0.0	0.870
Non-White (All)	87,304	26	0.0%	5.2	**0.023**	73,664	8	0.0%	2.2	0.140
White/Non-Hispanic	103,665	15	0.0%	N/A	N/A	245,324	14	0.0%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 13.A
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Interactions to Population Counts
By Zip Code Median Income Quartiles
January 2018 - February 2021

	Panel A: Income Quartile 1					*Panel B: Income Quartile 2*				
Race/Ethnicity[1]	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	3,736	196	5.2%	759.8	0.000	2,425	42	1.7%	98.9	0.000
Asian	14,081	1,620	11.5%	1,516.3	0.000	11,592	501	4.3%	122.5	0.000
Black/African American	39,480	27,485	69.6%	9,756.0	0.000	38,534	6,468	16.8%	3,003.9	0.000
Hispanic or Latino	130,076	22,712	17.5%	3,601.4	0.000	86,893	4,251	4.9%	426.2	0.000
Native Hawaiian/Pacific Islander	415	209	50.3%	56.2	0.000	785	26	3.3%	16.0	0.000
Non-White (All)	187,788	52,222	27.8%	103.8	0.000	140,230	11,288	8.0%	75.0	0.000
White/Non-Hispanic	82,817	24,910	30.1%	N/A	N/A	134,824	9,625	7.1%	N/A	N/A

	Panel C: Income Quartile 3					*Panel D: Income Quartile 4*				
Race/Ethnicity[1]	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**Interactions**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	1,211	16	1.3%	51.8	0.000	1,365	5	0.4%	25.4	0.000
Asian	10,369	336	3.2%	174.3	0.000	24,165	333	1.4%	121.9	0.000
Black/African American	31,534	3,880	12.3%	961.5	0.000	15,201	2,009	13.2%	5,159.7	0.000
Hispanic or Latino	43,889	3,051	7.0%	3.6	0.057	32,467	1,584	4.9%	557.8	0.000
Native Hawaiian/Pacific Islander	301	20	6.6%	0.0	0.984	466	17	3.6%	2.3	0.133
Non-White (All)	87,304	7,303	8.4%	183.0	0.000	73,664	3,948	5.4%	1,412.0	0.000
White/Non-Hispanic	103,665	6,914	6.7%	N/A	N/A	245,324	6,230	2.5%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 13.B
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of Arrests to Population Counts
By Zip Code Median Income Quartiles
January 2018 - February 2021

Panel A: Income Quartile 1

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	3,736	109	2.9%	282.4	0.000
Asian	14,081	669	4.8%	671.3	0.000
Black/African American	39,480	12,605	31.9%	5,149.4	0.000
Hispanic or Latino	130,076	8,848	6.8%	1,997.1	0.000
Native Hawaiian/Pacific Islander	415	78	18.8%	11.5	0.001
Non-White (All)	187,788	22,309	11.9%	40.7	0.000
White/Non-Hispanic	82,817	10,607	12.8%	N/A	N/A

Panel B: Income Quartile 2

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	2,425	22	0.9%	27.9	0.000
Asian	11,592	187	1.6%	45.5	0.000
Black/African American	38,534	2,428	6.3%	1,144.0	0.000
Hispanic or Latino	86,893	1,557	1.8%	172.1	0.000
Native Hawaiian/Pacific Islander	785	12	1.5%	3.8	0.052
Non-White (All)	140,230	4,206	3.0%	27.7	0.000
White/Non-Hispanic	134,824	3,588	2.7%	N/A	N/A

Panel C: Income Quartile 3

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	1,211	4	0.3%	16.2	0.000
Asian	10,369	106	1.0%	43.2	0.000
Black/African American	31,534	1,226	3.9%	379.7	0.000
Hispanic or Latino	43,889	917	2.1%	3.2	0.076
Native Hawaiian/Pacific Islander	301	7	2.3%	0.2	0.639
Non-White (All)	87,304	2,260	2.6%	87.2	0.000
White/Non-Hispanic	103,665	2,018	1.9%	N/A	N/A

Panel D: Income Quartile 4

Race/Ethnicity[1]	Population	Arrests	(as % of Population)	Chi-Squared Statistic	P-Value
American Indian/Alaska Native	1,365	2	0.1%	7.9	0.005
Asian	24,165	112	0.5%	40.0	0.000
Black/African American	15,201	626	4.1%	1,473.9	0.000
Hispanic or Latino	32,467	508	1.6%	158.1	0.000
Native Hawaiian/Pacific Islander	466	9	1.9%	6.4	0.011
Non-White (All)	73,664	1,257	1.7%	399.2	0.000
White/Non-Hispanic	245,324	2,080	0.8%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.

Figure 13.C
Chi-Squared Tests Against "White/Non-Hispanic" Group
Ratios of UoF Incidents to Population Counts
By Zip Code Median Income Quartiles
January 2018 - February 2021

	Panel A: Income Quartile 1					*Panel B: Income Quartile 2*				
Race/Ethnicity[1]	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	3,736	5	0.1%	3.2	0.073	2,425	0	0.0%	1.1	0.291
Asian	14,081	10	0.1%	23.0	**0.000**	11,592	0	0.0%	5.3	**0.021**
Black/African American	39,480	460	1.2%	352.0	**0.000**	38,534	80	0.2%	95.6	**0.000**
Hispanic or Latino	130,076	235	0.2%	29.2	**0.000**	86,893	38	0.0%	0.1	0.807
Native Hawaiian/Pacific Islander	415	2	0.5%	0.5	0.484	785	0	0.0%	0.4	0.548
Non-White (All)	187,788	712	0.4%	11.6	**0.001**	140,230	118	0.1%	15.3	**0.000**
White/Non-Hispanic	82,817	244	0.3%	N/A	N/A	134,824	62	0.0%	N/A	N/A

	Panel C: Income Quartile 3					*Panel D: Income Quartile 4*				
Race/Ethnicity[1]	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**	**Population**	**UoF Incidents**	**(as % of Population)**	**Chi-Squared Statistic**	**P-Value**
American Indian/Alaska Native	1,211	0	0.0%	0.4	0.511	1,365	0	0.0%	0.2	0.678
Asian	10,369	2	0.0%	0.7	0.389	24,165	0	0.0%	3.1	0.081
Black/African American	31,534	47	0.1%	50.0	**0.000**	15,201	17	0.1%	76.5	**0.000**
Hispanic or Latino	43,889	26	0.1%	4.0	**0.045**	32,467	11	0.0%	8.6	**0.003**
Native Hawaiian/Pacific Islander	301	0	0.0%	0.1	0.743	466	0	0.0%	0.1	0.808
Non-White (All)	87,304	75	0.1%	20.4	**0.000**	73,664	28	0.0%	19.7	**0.000**
White/Non-Hispanic	103,665	37	0.0%	N/A	N/A	245,324	31	0.0%	N/A	N/A

Notes:
[1] Race and ethnicity variable is equal to "Hispanic or Latino" if 'ethnicity expansion' is equal to "Hispanic or Latino." Otherwise, the variable is equal to 'race expansion.'
[2] Income Quartiles are computed from median household income from the ACS 5-Year Community Survey for each zip code that appears in the data.
[3] All data is limited to observations where person's 'role expansion' is equal to "ARRESTEE," "DRIVER/VICT," "SUMMONS," "OFFEND/SUSP," "SUBJECT," "VICT/ARREST," or "INVOLVED" and age is greater than or equal to 18 at date of occurrence.
[4] The Chi-Squared distribution, as a one-tailed distribution, is comprised of non-negative values (i.e., frequencies or counts). Red P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is greater than the white group's incidence rate. Blue P-Values indicate statistically significant differences in observed relative proportions at the 5% significance level where the specified subject group's incidence rate is less than the white group's incidence rate.
[5] It should be noted that the Chi-Squared Statistic, although indicative of proportionality (or disproportionality) to some degree, cannot be interpreted in statistical terms as *odds* or *probabilities*. Furthermore, a reasonably suitable interpretation of the test statistic as specified is conditionally dependent on the aggregate population composition and/or a specified local subpopulation composition (e.g., overall age demographics within a particular zone versus across the metropolitan area in general).
Sources: Census Data shown in Population column comes from the 2019 American Community Survey (Five-Year Estimates) and are estimates for each zip code that appears in the dataset.