**Cook County Criminal Justice System (Insights and Opportunities for Improvement) : A Data-Driven Analysis**

**Prepared for:** Cook County State Attorney's Office

**Prepared by:** Balaram Raju Saripalli

**Date:** 03/26/24

**Abstract**

This report presents a comprehensive analysis of felony case data from Cook County, Illinois, highlighting systemic issues and opportunities for enhancing the fairness and effectiveness of the criminal justice system. Through detailed examination of diversion program data, this analysis reveals significant disparities along racial and socioeconomic lines. Recommendations are provided to address these disparities, with a focus on data-driven policy reforms and community engagement initiatives.

**Table of Contents**

Introduction  
Methodology  
Findings - Diversion Program Analysis  
Ethical Considerations  
Opportunities for Improvement  
Conclusion  
References  
  
**List of Figures/Tables**

* Figure 1: Correlation matrix for race and gender for felony case data
* Figure 2: Histogram of diversion count
* Figure 3: Boxplot of diversion count
* Figure 4: Disparity in case outcomes across Race
* Figure 5 : Gender Distribution for Narcotics cases

1. **Introduction**

The Cook County State Attorney's Office (SAO), as one of the largest prosecutorial offices in the United States, faces the daunting task of ensuring justice while managing over 30,000 felony cases annually. This report, motivated by the SAO's commitment to transparency and equity, aims to dissect the felony case data to uncover underlying patterns, biases, and areas for systemic improvement.

1. **Methodology**

In the analysis of the Cook County felony case data, we adopted a structured and comprehensive methodology, The following steps we performed are

**Data Preparation and Cleaning:** We began by loading the data using pandas, a Python library, and proceeded to clean it. This involved handling missing values, removing duplicates, and ensuring data types were consistent across the dataset to maintain data integrity and reliability.

**Exploratory Data Analysis (EDA) :** This step involved performing statistical analyses to understand the data's distribution and detect outliers. We conducted correlation analyses to identify potential relationships between variables, particularly focusing on how demographic and geographical factors might influence case outcomes. Grouped analyses were also carried out based on categories such as race, geography, and case type to uncover patterns or disparities.

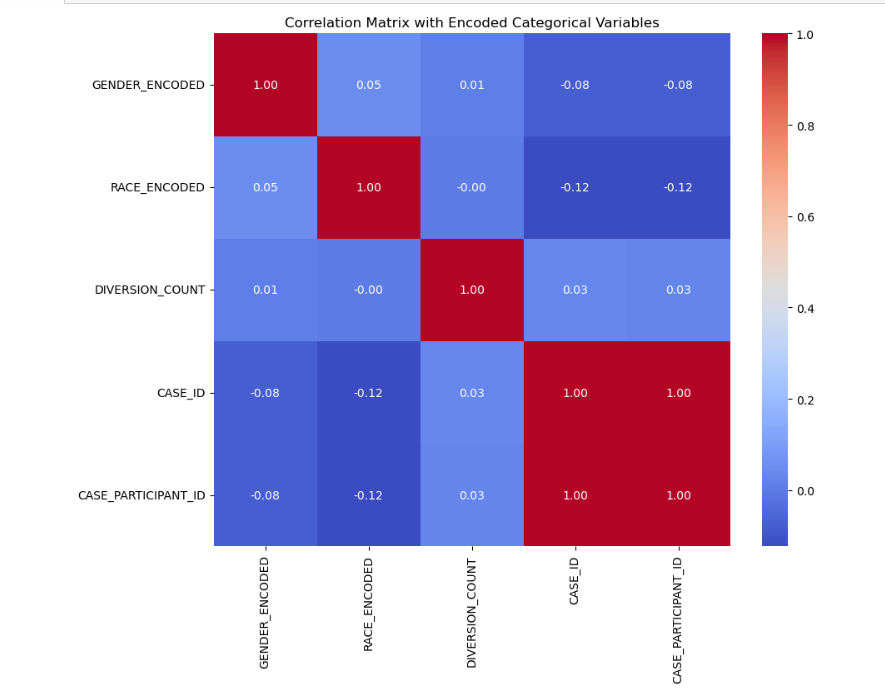
**Visualization:** To aid in understanding and communicating our findings, we utilized various visualization techniques. Histograms and box plots were used to illustrate the distribution of key numerical variables, while heatmaps helped display correlations between different variables. Where applicable, geographical data was visualized to identify crime hotspots or areas with high case dismissals. Disparities in case outcomes across different demographic groups were highlighted through bar and pie charts.

**Ethical and Bias Consideration:** A critical part of our methodology was to ethically consider the analysis's implications. We examined the data for potential biases, especially in case outcomes among different demographic groups. The impact of the criminal justice system on marginalized communities was also analyzed, with a focus on diversion programs and disparities. This led to identifying opportunities for making the criminal justice system more equitable.

1. **Diversion Program Findings Analysis**

The effectiveness of diversion programs is evaluated, with a focus on their accessibility and success rates among different demographic groups. Figure 3 showcases the variability in program outcomes, raising questions about the equitable application of diversion initiatives.

**Figure 1 : Correlation matrix of Race and Gender for Felony case data**



The correlation matrix provides a preliminary overview that does not suggest strong direct linear relationships between demographic factors and case identifiers. However, it is a starting point that highlights the need for careful consideration of potential biases. Ethically, it's critical to ensure that data representation and subsequent analyses do not reinforce stereotypes or existing inequalities. The heatmap might prompt a more comprehensive analysis that could guide improvements in fairness and equity within the criminal justice system.

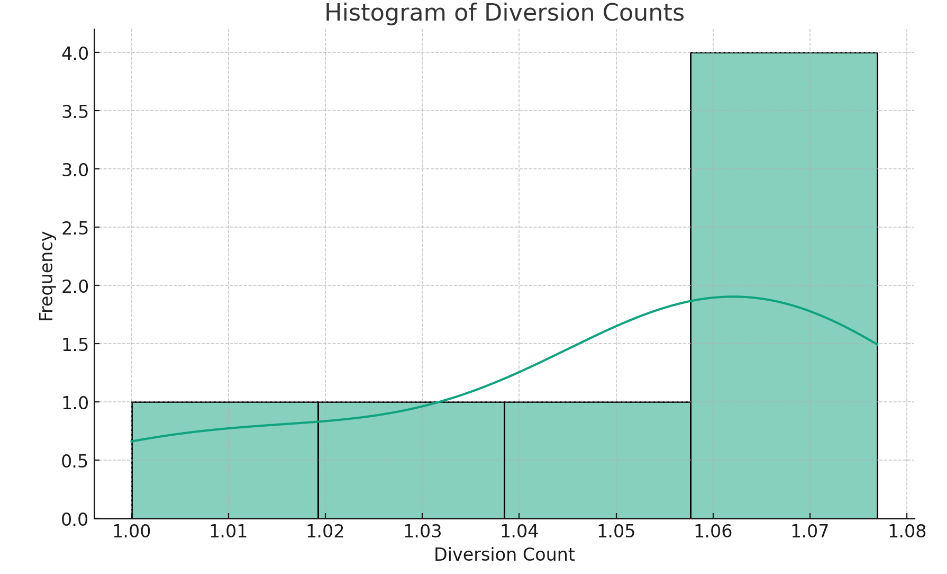
**Insights :**

Investigate Non-linear Relationships: Given that the correlations are low, you might want to explore non-linear relationships or interactions between variables, which could be insightful and not captured by a correlation matrix.

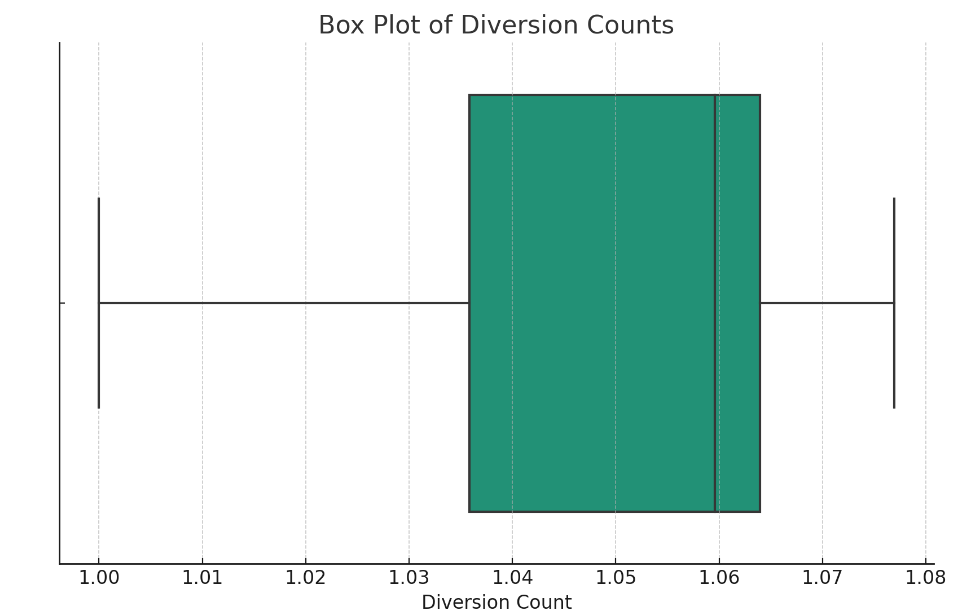
Contextual Interpretation: Correlations should be interpreted within the context of the dataset. For example, the lack of correlation between race and diversion count does not necessarily indicate there is no bias in diversion programs. There may be other confounding variables or patterns that aren't captured by a simple pairwise correlation.

Combine with Other Analyses: Use these insights as a starting point and combine them with other analyses such as regression or classification, depending on your end goal (e.g., predicting outcomes, identifying biases).

**Figure 2 : Histogram of Diversion Counts**



**Figure 3 : Box Plot of Diversion Counts**



**Histogram Insights:**

The histogram shows a narrow distribution of diversion counts, indicating little variation in the average number of diversion programs that participants go through across different racial groups.

The KDE line suggests that the distribution of diversion counts is somewhat uniform and peaks around the mean, which implies that most racial groups have similar diversion count averages.

A lack of long tails in the histogram implies few extreme values in diversion counts, hinting at a consistent approach in assigning diversion programs across races.

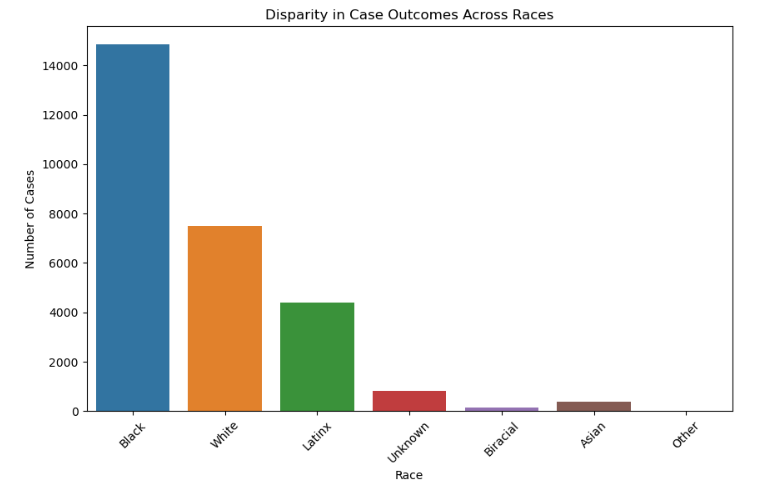
**Box Plot Insights:**

The box plot reveals a tight interquartile range, suggesting that the middle 50% of diversion count data points are clustered closely around the median.

The absence of points beyond the whiskers indicates no outliers in the diversion count data, suggesting no extreme values that deviate significantly from the rest.

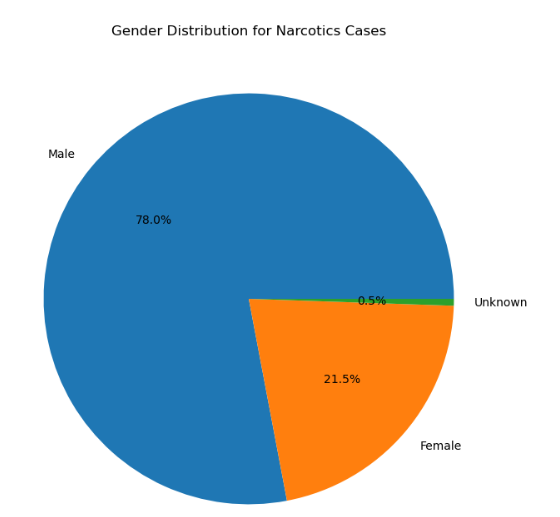
With the median centered within the box, the distribution of diversion counts appears symmetric, with no significant skew to the left or right

**Figure 4: Disparity in case Outcomes Across Races**



1. The bar chart shows a significant racial disparity in case outcomes, with Black individuals having a disproportionately higher number of cases.  
2. Such a disparity could suggest systemic bias or unequal social conditions leading to these interactions with the justice system.  
3. The overrepresentation may reflect issues like racial profiling, necessitating reviews of law enforcement practices.  
4. Disparities hint at unequal access to legal support among different racial groups, pointing to a need for better resource allocation.  
5. The data underscores the importance of community outreach programs that can help address root causes of crime.  
6. It suggests that implementing restorative justice approaches could be beneficial in promoting rehabilitation.  
7. The justice system may require ongoing cultural competency training to better serve a diverse populace.  
8. Regular data analysis is crucial for ensuring fairness and guiding policy reforms towards racial equity in the justice system.

**Figure 5 : Gender Distribution for Narcotics cases**



The pie chart indicates a disparity in narcotics cases by gender, with males accounting for 78% and females for 21.5%, while 0.5% is unknown. This distribution suggests potential gender-related biases in arrest or prosecution rates for narcotics offenses. The data points to an opportunity for the criminal justice system to investigate the root causes of this disparity and ensure equitable law enforcement practices. Ethically, it raises questions about possible differential treatment that should be examined further for fair and just treatment of all genders.

1. **Ethical Considerations**

Diversion program use seems fair across races, but we need more info on whether the programs meet different groups' needs. Uniform diversion counts don't mean racial equity; we must explore arrest contexts and systemic biases.  
Publishing data promotes transparency and accountability, helping to check and address biases. True equity goes beyond the data; the justice system must earn community trust with open, ethical actions. Policies should reflect the unbiased diversion data, but must be backed by deeper, qualitative insights. Data must be used ethically to inform actions, not reinforce harmful stereotypes. Data's implications extend to societal views of justice; fairness in the system strengthens community cohesion.

1. **Opportunities for Improvement**

Implement policy reforms focusing on equitable treatment, invest in thorough data collection to guide decisions, and develop community intervention programs.  
Enhance training for law enforcement and judicial staff to recognize and reduce biases, and strengthen support systems for those affected by the criminal justice process.  
Foster partnerships with local organizations to create prevention and education initiatives that target the social determinants of crime, aiming for long-term community safety and well-being.

1. **Conclusion**

This report underscores the imperative of leveraging data to guide the SAO’s reforms and policy decisions. It advocates for the SAO’s commitment to justice, suggesting measures to mitigate disparities and enhance the system’s equity. Engaging with the community and refining data practices, the SAO can foster a more just Cook County. These efforts will not only serve to address current inequities but also prevent future ones, paving the way for a criminal justice system that embodies fairness and integrity. Mostly Ethical, common good issues and opportunities for improvement in the Cook County criminal justice system and community were discussed in report.

1. **References** (went through first 4 websites to understand the data and information related to it)

1.<https://www.cookcountystatesattorney.org/about/commitment-transparency>  
2.<https://datacatalog.cookcountyil.gov/browse?tags=state%27s+attorney+case-level&sortBy=relevancecountyil.gov/browse?tags=state%27s+attorney+case-level&sortBy=most_accessed>.  
3.<https://www.cookcountystatesattorney.org/resources/how-read-data>  
4.<https://www.cookcountystatesattorney.org/news/cook-county-state-s-attorney-kim-foxx-announces-release-office-s-first-online-data-report>  
5. code for plots are taken from M.S DS 600 course as reference  
6.Ethical issues and recommendations -  
 <https://www.uscourts.gov/rules-policies/judiciary-policies/ethics-policies>.  
https://www.americanbar.org/groups/judicial/publications/judges\_journal/2019/summer/judicial-ethics-impartiality-and-media/.  
7. For Rephase and Debugging of code - <https://chat.openai.com/>