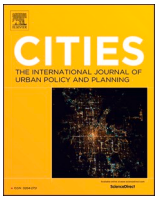
Cities 154 (2024) 105398

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities



Unveiling inequality: A deep dive into racial and gender disparities in US court case closures

Yoshiyasu Takefuji

Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan

ARTICLE INFO

Keywords:
Open data
Generative AI
Disparity trends
SEX and race

ABSTRACT

This study uses generative AI and public datasets to examine racial and gender disparities in US court case closures from 2019 to 2024. It finds significant disparities using ANOVA, Chi-Square, and Fisher's method, with an increasing trend over time. Gender disparity, though less significant in 2024, persists. Further research is needed to identify root causes and develop targeted interventions. Continuous monitoring is essential to evaluate their effectiveness. Promoting transparency, investing in research, and implementing robust monitoring systems are crucial steps towards a fairer justice system for all.

1. Introduction

This paper exposes racial and gender inequality in US court case closures from 2019 to 2024 using generative AI and public datasets. It finds significant disparities through ANOVA, Chi-Square, and Fisher's method. Targeted interventions are needed to address these disparities, but ongoing monitoring is essential to assess their effectiveness. Efforts should focus on transparency, investing in research to uncover root causes, and implementing robust monitoring systems. By taking these steps, we can work towards a more equitable justice system where fairness is not determined by race or gender.

The open data movement in the US has led to the disclosure of numerous datasets (Catone, 2023), enabling researchers to leverage generative AI for evidence-based outcomes (Paz-Pacheco, 2024). This is particularly beneficial for those without programming skills, as it allows them to visualize disparity trends by sex and race (Takefuji, 2023). As of July 22, 2024, the data.gov site hosts 301,420 datasets, making it the world's largest dataset repository. For this research, six datasets were identified using 'court' as a keyword. These datasets were utilized to visualize disparity trends by sex and race.

In statistical tests, the *p*-value is a crucial concept. It measures the strength of evidence in support of a null hypothesis, which is a general statement or default position asserting no relationship between two measured phenomena. In the context of the ANOVA and Chi-Square tests performed in this study, the null hypothesis would state that there is no difference in case closure rates across different races or genders, or no association between race/gender and case closure status. The *p*-value quantifies the probability of observing the data under the assumption that the null hypothesis is true. A smaller *p*-value indicates stronger evidence in favor of the alternative hypothesis.

A common threshold for statistical significance is a *p*-value of 0.05. If the p-value is less than 0.05, it typically indicates strong evidence against the null hypothesis, leading to its rejection and the acceptance of the alternative hypothesis. In other words, a *p*-value less than 0.05 may suggest a statistically significant difference or association in the data.

This study employs three tests: ANOVA (Chatzi & Doody, 2023), Chi-Square (Ordak, 2023), and a combination of *p*-values using Fisher's method (Park et al., 2023). ANOVA, or Analysis of Variance, compares the means of more than two groups. It assumes normal data distribution and equal variances across groups, known as homoscedasticity. ANOVA can compare more than two groups simultaneously and is robust to departures from normality. However, it may be less powerful when these assumptions are violated. While ANOVA can identify differences between groups, it does not specify which groups are statistically significantly different from each other, necessitating post-hoc tests. This study shows the distribution and the changes of racial and gender disparities in US court case closures from 2019 to 2024.

2. Methods

Our study focused on the racial and gender disparities in US court case closures specifically within the Austin city court system. We analyzed the data at the city level, examining how these disparities were distributed across different neighborhoods and districts within Austin.

The overlooked capability of generative AI lies in searching datasets over the Internet. Researchers should download six datasets manually

E-mail address: takefuji@keio.jp.

https://doi.org/10.1016/j.cities.2024.105398
Received 8 August 2024; Received in revised form 12 August 2024; Accepted 16 August 2024

from the data.gov site (Gov, n.d.-c; Gov, n.d.-b; Gov, n.d.-a; Gov, n.d.-f; Gov, n.d.-e; Gov, n.d.-d). Rename these datasets to 2019.csv, 2020.csv, …, 2024.csv respectively.

The data provided is designed to facilitate the analysis of various charges brought forth in the Downtown Community Court of Austin City. The abbreviations used for "Race" and "Case Status" are as follows: "A" stands for Asian, "B" for Black, "BA" for Black or African American, "CD" when the client does not know, "CR" when the client refused, "DNC" when data is not calculated, "H" for Native Hawaiian or Other Pacific Islander, "L" for Hispanic or Latino, "ME" for Middle Eastern, "MR" when multiple races are identified, "N" for Native American or Alaskan, "O" for Other, "U" for Unknown, and "W" for White. In terms of "Case Status", "Y" signifies a closed case, "N" indicates the case is not closed, "TERM" and "TERMA" both denote a terminated case which is equivalent to a closed case, "ACT" represents an active case, and "IN" stands for an inactive case which does not imply the case is closed. Only the statuses "TERM" and "TERMA" confirm a closed case. This information is crucial for a comprehensive understanding of the data.

A query plays a crucial role in generating correct Python code. The following is the initial query fed to generative AI using Copilot. Before crafting the query, researchers should be familiar with variables in the datasets.

Initial Query: use 6 csv files such as 2019.csv up to 2024.csv for 6 years and compute bias based on *p*-value (ANOVA test,Chi-Square test, and combined *p*-values using Fisher's method) using variables such as 'Case Closed', 'RACE', and 'Defendant Gender'. 'Case Closed', 'Defendant Gender', and 'RACE' contain strings while some cells are empty. Draw a graph of trend black lines of three bias tests between 2019 and 2024. Black lines should have 4 different line styles and 2 different line width(1,2). show Python code.

3. Results

The final version of the code, which has been meticulously developed and tested, is conveniently attached in the Appendix A for reference. In addition to this, a comprehensive collection of six distinct datasets, along with the corresponding code, are readily accessible at the specified GitHub site (GitHub, n.d.). The program, q1.py for generating the distribution of racial and gender disparities in US court case closures from 2019 to 2024 and q2.py are publicly available for the changes in these disparities over the same period (GitHub, n.d.).

The graphical representation of our findings, as depicted in Fig. 1, provides a clear illustration of the disparity trends by gender and race over a six-year period from 2019 to 2024. The evidence, substantiated by the calculated *p*-value, indicates a gradual enhancement in these disparity trends by sex and race. Red line in Fig. 1 indicates p-value = 0.05 for reference.

It's of paramount importance to underscore that the only ANOVA test, which was specifically targeted at gender, did not reveal any significant correlation with disparity in 2024. This observation implies that gender might not be a major determinant in the observed disparities. In contrast, other tests have demonstrated strong associations, suggesting that factors other than gender could be contributing to these trends from 2019 to 2024. Interestingly, as depicted in Fig. 1, the disparity trends for the years 2019 and 2020 appear to be diminished.

A more in-depth investigation is required to identify these contributing factors and to fully understand their impact on the observed disparities. Gaining such insights will be crucial in devising more effective strategies aimed at addressing and potentially mitigating these disparities.

To summarize, the results emphasize the distinct disparity trends as illustrated in Fig. 1. It is imperative for future research to investigate the underlying causes and reasons for these trends. Such studies will offer a more holistic understanding of the situation, thereby informing the development of interventions that are more targeted and effective. Furthermore, continuous monitoring and evaluation of these trends will be essential to assess the impact of these interventions and make necessary adjustments.

Fig. 2 illustrates the distribution of racial and gender disparities in US court case closures from 2019 to 2024, while Fig. 3 depicts the changes in these disparities over the same period. In Fig. 2, starting from 2022, the proportions of both White and Black individuals increase; however, the proportion of Black individuals stabilizes, whereas the proportion of White individuals continues to rise. In Fig. 3, the changes in the RACE distribution remain nearly zero until 2022. The changes stabilize in 2023, with only the White population showing a positive change in proportion. In Fig. 3, the changes in the gender distribution intersect between 2020 and 2021. While the proportion of Female individuals increases until 2022, there is a small drop in 2023, followed by an increase in 2024. Conversely, Male individuals exhibit opposite behaviors.

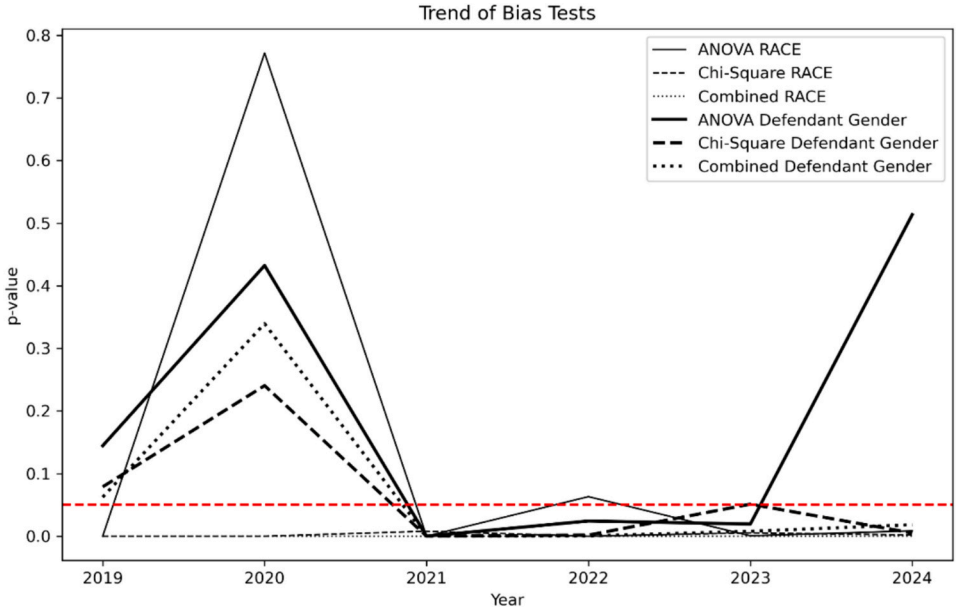


Fig. 1. Disparity trends by sex and race from 2019 to 2024.

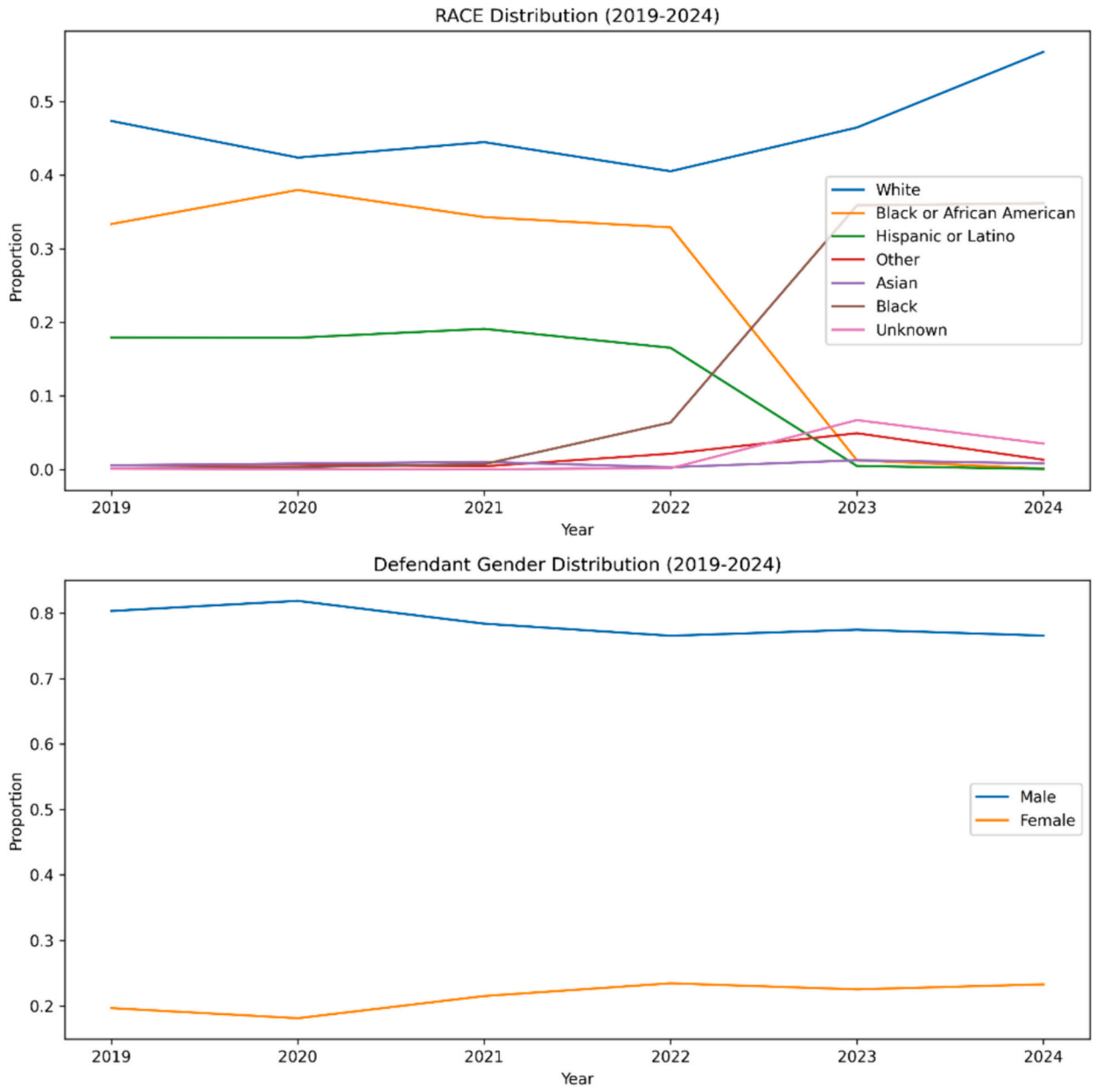


Fig. 2. distribution of racial and gender disparities in US court case closures from 2019 to 2024.

4. Discussion

A concise literature review was conducted on court disparity using recent peer-reviewed publications in 2024 from the National Library of Medicine. Singh et al. investigated the impact of economic downturns on involuntary psychiatric hospitalizations among African American males (Singh et al., 2024). Using data from four US states, they found that a decline in employment precedes a 6 % increase in hospitalizations requested by law enforcement or court orders among this group. They suggested that economic stressors may exacerbate mental health issues, leading to an increase in involuntary commitments, highlighting the unique vulnerability of racial/ethnic minorities during economic contractions. However, they did not clarify court disparity.

Wang et al. explored the impact of school racial segregation on children's health in the U.S (Wang et al., 2024). They found that the end of court-ordered desegregation, facilitated by Supreme Court decisions since 1991, has led to increased school segregation. This has resulted in improved mental health among Black children and better self-reported health among White children. However, long-term effects include worse self-reported health and higher risk of asthma among Black children. Interventions to mitigate school segregation are crucial for reducing racial health disparities (Wang et al., 2024).

Zare reported that recent high-profile police encounters have intensified scrutiny of law enforcement and deepened community mistrust (Zare, 2024). Disparities in policing, reflecting systemic oppression, required exploration of historical discrimination, socioeconomic inequalities, and power dynamics. Theories explaining these disparities included majority-minority communities, conflict theory of law, and minority threat hypothesis. The study by Ward et al. fills a research gap by focusing on nonfatal shootings, providing valuable insights into policing disparities. Understanding these factors is crucial for addressing disparities and reducing police violence (Zare, 2024).

Oh et al. addressed that the U.S. Supreme Court's 2023 decision, which dismantled race-conscious college admissions, has intensified concerns about increasing racial inequality in higher education (Oh et al., 2024). Despite an overall increase in college entrance rates, Black

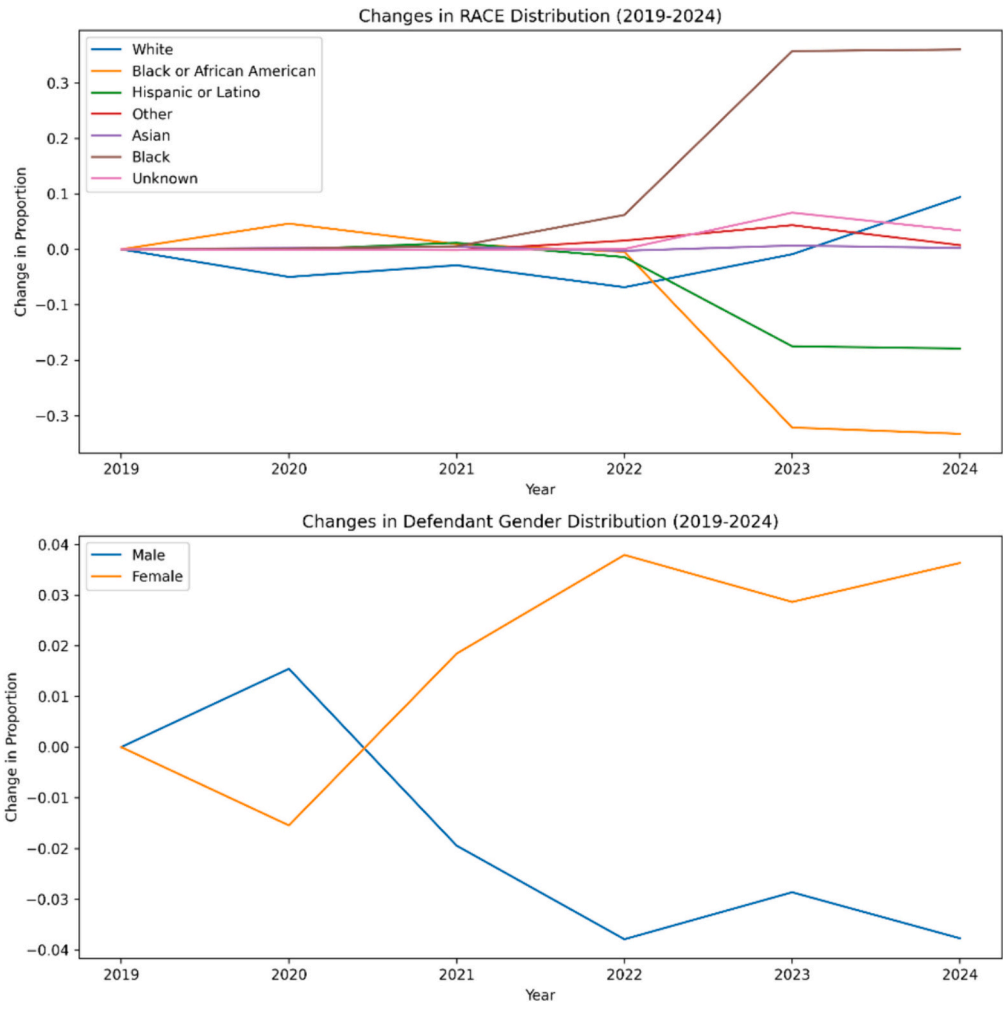


Fig. 3. changes of racial and gender disparities in US court case closures from 2019 to 2024.

and Hispanic youth were less likely than their White peers to attend four-year selective colleges. This disparity has widened over the decades, even after accounting for socioeconomic factors. Their findings highlighted the urgent need for policy interventions to address this growing racial inequality in higher education (Oh et al., 2024).

Hailu et al. investigated the association between county-level jail incarceration inequity and severe maternal morbidity (SMM) risk in California (Hailu et al., 2024). Their study found that Black and Hispanic or Latinx individuals residing in counties with high Black-White jail incarceration inequity have increased odds of SMM. The findings underscored the adverse maternal health consequences of structural racism within the criminal-legal system, highlighting the need for community-based alternatives to address these inequities (Hailu et al., 2024).

Zimmerman et al. studied racial and ethnic disparities in potential years of life lost due to violent death in the U.S (Zimmerman et al., 2024). Findings revealed that persons of color, particularly non-Hispanic African Americans and Asian or Pacific Islanders, are disproportionately impacted by violence, dying eleven or more years earlier than their non-Hispanic White counterparts. These disparities, partly explained by individual differences, incident characteristics, and contextual factors, underscored the high societal cost of homicide and suicide, particularly among persons of color (Zimmerman et al., 2024).

The current studies in 2024 revealed significant racial disparities across various societal domains, including involuntary psychiatric hospitalizations, school segregation, law enforcement, higher education, and maternal health. These disparities, often exacerbated by systemic oppression and economic downturns, highlight the urgent need for policy interventions and community-based alternatives to address these inequities and reduce the societal cost of violence. This paper provides a response to recent research findings and addresses the disparities in court case closures, substantiated by empirical evidence.

Once we identify the root causes of these disparities, it's essential to devise targeted interventions to address them. However, the creation of interventions alone is insufficient. We need to establish a system for ongoing monitoring to assess the effectiveness of these interventions. Our efforts should be centered on fostering transparency by utilizing publicly available datasets, investing in research to discover the underlying causes of these disparities, formulating targeted interventions based on the findings of this research, and setting up a robust system for continuous monitoring to gauge the impact of these interventions. By undertaking these measures, we can work towards achieving a justice system that is more equitable, where fairness is not determined by one's race or gender.

We have outlined the following recommendations: For local practice, we suggest implementing targeted policies to address identified disparities in court case closures within Austin, including training programs for legal professionals on implicit bias and equitable practices. Additionally, we recommend developing community outreach programs to support affected populations and ensure they have access to legal resources and representation. For international practice, we propose facilitating international collaboration and knowledge exchange on best practices for reducing racial and gender disparities in the legal system. Furthermore, we advocate for the adoption of global standards and

guidelines to promote equity and justice in court case closures worldwide. By implementing these recommendations, we aim to create a more equitable justice system both locally and globally.

5. Conclusion

In light of the identified disparities in gender and race within US court case closures, our attention should be directed towards several key actions:

Transparency and Accountability: Promote transparency and accountability in the justice system by making use of publicly available datasets and ensuring open communication about the progress and results of these interventions.

Invest in Research: Allocate resources to research aimed at identifying the root causes of these disparities.

Develop Targeted Interventions: Based on the findings of the research, develop interventions specifically designed to address the identified causes of disparities.

Implement Continuous Monitoring: Establish a robust system for continuous monitoring to assess the effectiveness of the implemented interventions over time.

By focusing on these actions, we can strive to create a more equitable justice system where everyone, regardless of their race or gender, is treated with fairness and dignity. This is not just a legal imperative, but a moral one, as it goes to the heart of what justice truly means.

Funding

This research has no fund.

CRediT authorship contribution statement

Yoshiyasu Takefuji: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The author has no conflict of interest.

Data availability

Data are publicly available at Gov sites (Gov, n.d.-c; Gov, n.d.-b; Gov, n.d.-a; Gov, n.d.-f; Gov, n.d.-e; Gov, n.d.-d).

Acknowledgements

Not applicable.

Appendix A

```
disparity.py
import pandas as pd
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt
# Initialize dictionaries to store p-values for each year
pvals_anova = {}
pvals_chi = {}
pvals_combined = {}
# Ask user for input
print("Please select one of the following options:")
print("1. Evaluate 'RACE'")
print("2. Evaluate 'Defendant Gender'")
print("3. Evaluate both 'RACE' and 'Defendant Gender'")
choice = int(input("Enter your choice (1, 2, or 3): "))
# Determine variables to evaluate based on user input
if choice == 1:
variables = ['RACE']
elif choice == 2:
variables = ['Defendant Gender']
elif choice == 3:
variables = ['RACE', 'Defendant Gender']
else:
print("Invalid choice. Please enter 1, 2, or 3.")
exit()
# Define line styles and widths for each variable and test
styles = {'RACE': {'linestyles': ['-', '–', ':', '-.'], 'linewidths': [1]},
'Defendant Gender': {'linestyles': ['-', '–', ':', '-.'], 'linewidths': [2]}}
# Loop over each year
for year in range(2019, 2025):
# Load the data
df = pd.read_csv(f'{year}.csv')
for variable in variables:
# Handle empty cells by filling with a default value or removing
df = df.dropna(subset=['Case Closed', variable])
# Convert string values to categories for analysis
df['Case Closed'] = df['Case Closed'].astype('category').cat.codes
df[variable] = df[variable].astype('category').cat.codes
# Perform ANOVA test
fval_anova, pval_anova = stats.f_oneway(*[df.loc[df[variable] ==
val, 'Case Closed'] for val in df[variable].unique()])
pvals_anova.setdefault(variable, []).append(pval_anova)
# Perform Chi-Square test
contingency_table = pd.crosstab(df[variable], df['Case Closed'])
chi2, pval_chi, dof, expected = stats.chi2_contingency
(contingency_table)
pvals_chi.setdefault(variable, []).append(pval_chi)
# Combine the p-values using Fisher's method
combined_pval = stats.combine_pvalues([pval_anova, pval_chi],
method='fisher')[1]
pvals_combined.setdefault(variable, []).append(combined_pval)
# Plot the trends
plt.figure(figsize=(10, 6))
years = list(range(2019, 2025))
for variable in variables:
plt.plot(years, pvals_anova[variable], label=f'ANOVA {variable}',
linestyle=styles[variable]['linestyles'][0], linewidth=styles[variable]
['linewidths'][0], color='black')
plt.plot(years, pvals_chi[variable], label=f'Chi-Square {variable}',
linestyle=styles[variable]['linestyles'][1], linewidth=styles[variable]
['linewidths'][1], color='black')
plt.plot(years, pvals_combined[variable], label=f'Combined {vari-
able}', linestyle=styles[variable]['linestyles'][2], linewidth=styles[var-
iable]['linewidths'][2], color='black')
# Add a horizontal line at p-value=0.05
plt.axhline(y=0.05, color='r', linestyle='–')
plt.xlabel('Year')
plt.ylabel('p-value')
plt.title('Trend of Bias Tests')
plt.legend()
plt.savefig('result.png',dpi=300)
plt.show()
```

References

Catone, M. C. (2023). The role of open data in digital society: The analysis of scientific trending topics through a bibliometric approach. *Frontiers in Sociology, 8*, Article 1134518. Published 2023 Feb 1 https://doi.org/10.3389/fsoc.2023.1134518.

Chatzi, A., & Doody, O. (2023). The one-way ANOVA test explained. *Nurse Researcher, 31* (3), 8–14. https://doi.org/10.7748/nr.2023.e1885

GitHub. Disparity trends by sex and race Accessed on July 28 https://github.com/y-takefuji/disparity.

Gov. Downtown Austin Community Court DACC FY2019 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/9e8e-ghwx/rows.csv?accessType=DOWNLOAD.
Gov. Downtown Austin Community Court DACC FY2020 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/vgqi-hjzx/rows.csv?accessType=DOWNLOAD.
Gov. Downtown Austin Community Court DACC FY2021 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/448k-h8t4/rows.csv?accessType=DOWNLOAD.
Gov. Downtown Austin Community Court DACC FY2022 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/56au-zawf/rows.csv?accessType=DOWNLOAD.
Gov. Downtown Austin Community Court DACC FY2023 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/4tje-ntwx/rows.csv?accessType=DOWNLOAD.
Gov. Downtown Austin Community Court DACC FY2024 caseload information Accessed on July 28 https://data.austintexas.gov/api/views/mv2b-q2wb/rows.csv?accessType=DOWNLOAD.
Hailu, E. M., Riddell, C. A., Bradshaw, P. T., Ahern, J., Carmichael, S. L., & Mujahid, M. S. (2024). Structural racism, mass incarceration, and racial and ethnic disparities in severe maternal morbidity. *JAMA Network Open, 7*(1), Article e2353626. Published 2024 Jan 2 https://doi.org/10.1001/jamanetworkopen.2023.53626.
Oh, B., Tilbrook, N., & Shifrer, D. (2024). Shifting tides: The evolution of racial inequality in higher education from the 1980s through the 2010s. *Socius, 10*. https://doi.org/10.1177/23780231231225578 (10.1177/23780231231225578).
Ordak, M. (2023). ChatGPT's skills in statistical analysis using the example of allergology: Do we have reason for concern? *Healthcare (Basel), 11*(18), 2554 (Published 2023 Sep 15) https://doi.org/10.3390/healthcare11182554.
Park, S. H., Ha, S., & Kim, J. K. (2023). A general model-based causal inference method overcomes the curse of synchrony and indirect effect. *Nature Communications, 14*(1), 4287. Published 2023 Jul 24 https://doi.org/10.1038/s41467-023-39983-4.
Paz-Pacheco, E. (2024). Generative Artificial Intelligence (AI) in scientific publications. *Journal of the ASEAN Federation of Endocrine Societies, 39*(1), 4–5. https://doi.org/10.15605/jafes.039.01.01
Singh, P., Catalano, R., & Bruckner, T. A. (2024). Racial disparities in law enforcement/court-ordered psychiatric inpatient admissions after the 2008 recession: A test of the frustration-aggression-displacement hypothesis. *Social Psychiatry and Psychiatric Epidemiology, 20*. https://doi.org/10.1007/s00127-024-02627-z
Takefuji, Y. (2023). Generative AI for analysis and identification of Medicare improper payments by provider type and HCPC code. *Exploratory Research in Clinical and Social Pharmacy, 12*, Article 100387. Published 2023 Dec 1 https://doi.org/10.1016/j.rcsop.2023.100387.
Wang, G., White, J. S., & Hamad, R. (2024). The end of court-ordered desegregation and U.S. children's health: Quasi-experimental evidence. *American Journal of Epidemiology*. https://doi.org/10.1093/aje/kwae082. Published online May 22.
Zare, H. (2024). Disparities in policing from theory to practice. *American Journal of Public Health, 114*(4), 384–386. https://doi.org/10.2105/AJPH.2024.307588
Zimmerman, G. M., Fridel, E. E., & Trovato, D. (2024). Disproportionate burden of violence: Explaining racial and ethnic disparities in potential years of life lost among homicide victims, suicide decedents, and homicide-suicide perpetrators. *PLoS One, 19*(2), Article e0297346. Published 2024 Feb 7 https://doi.org/10.1371/journal.pone.0297346.