# Assignment2

The relationship between different demographic factors and different types of crimes

Jiajun Li, ID:38339315

2025-09-07

## 1 Introduction

### 1.1 Background

Crime is not only a legal issue but also a serious social and public safety concern. Different demographic factors, such as age, gender, and ethnicity, may have varying degrees of severity in terms of crime types. Understanding how these variables interact can provide deeper insights into crime patterns, which can support more effective policy design and policing strategies.

### 1.2 Purpose

This study aims to examine the interactions between demographic variables (age, sex, ethnicity) and crime types. Specifically, the research questions are:
How do the demographic variables (age, sex, ethnicity) interact to shape the distribution and trends of different crime categories in New Zealand?

## 2 Methods

### 2.1 Dataset

The data used in this study were sourced from the offender demographics section of the NZ Police website. It contains monthly records of offender proceedings with attributes including:
Ethnic Group: offender ethnicity
Age Group: age bracket of the offender
Sex: male, female, or not stated
ANZSOC Division: classification of crime type
Proceedings: number of recorded offender proceedings

### 2.2 Data preprocessing

Standardized column names and replaced the spaces in column names with underscores.

Converted monthly data to annual totals for trend analysis.

Removed unspecified or invalid categories.

Crime type: grouped into four broader categories for comparability: Violent Offences (Assault, Homicide, Robbery, Sexual Offences, Harm or endanger persons, Weapons and explosives offences)
Property Offences (Burglary, Theft, Property damage, Fraud , and related offences)

Drug & Public order Offences (Drug Offences, Public order/health, Justice procedures, Offences against government)
Traffic & Other Offences (Traffic Offences and others)

Re-grouped age into five broader categories for interpretability:
Teenagers (0-19), Adults (20-59), Elderly (60+)

Aggregated proceedings by combinations of variables (Year × Age × Sex × Ethnic × Crime_type).

**2.3 Statistical method**

For descriptive analysis, I used different charts with ggplot2 to visualize the interactions rather than hypothesis testing, as the aim was to uncover interaction patterns not easily visible in dashboards.
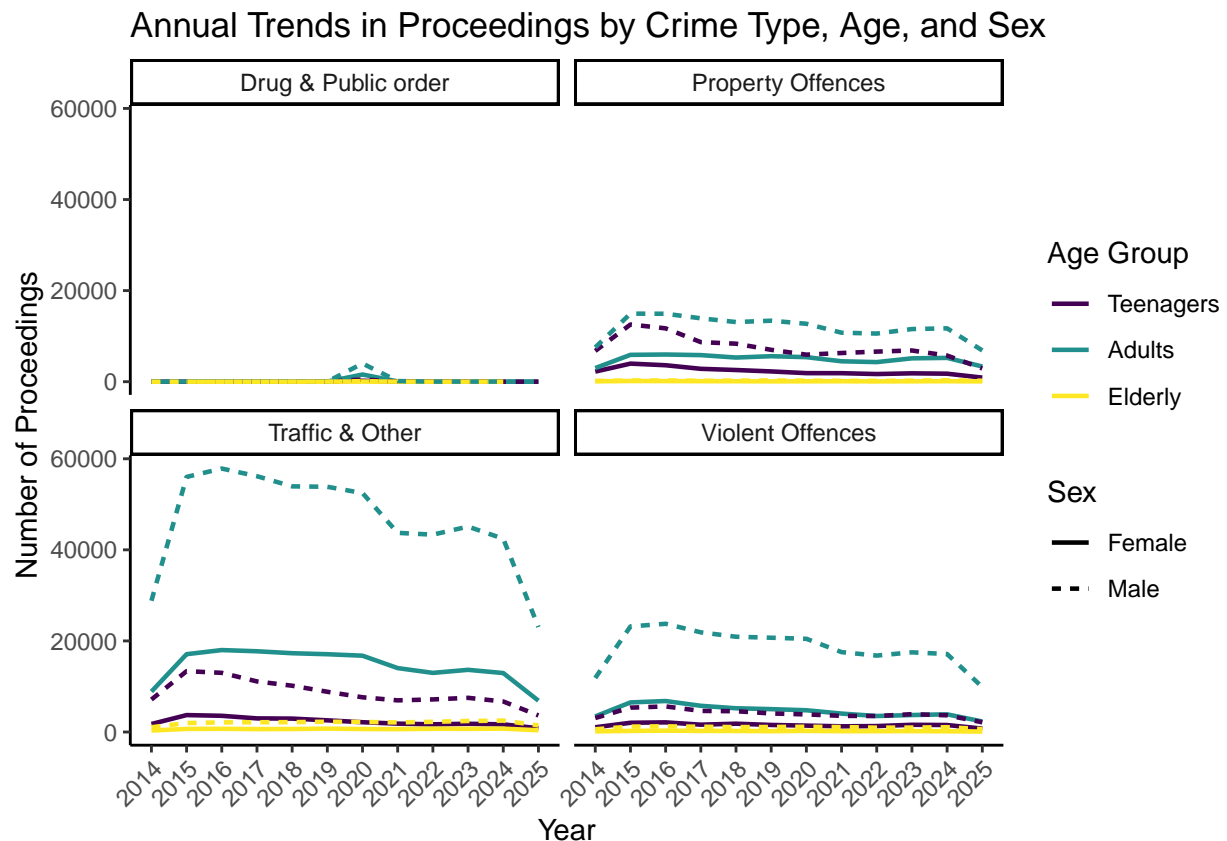Visualization:
Line charts to show long-term annual trends by sex and age.
Bar charts to highlight demographic breakdowns by sex, age, and ethnicity.
100% stacked bar charts to compare ethnic group proportions across age groups and crime. categories.
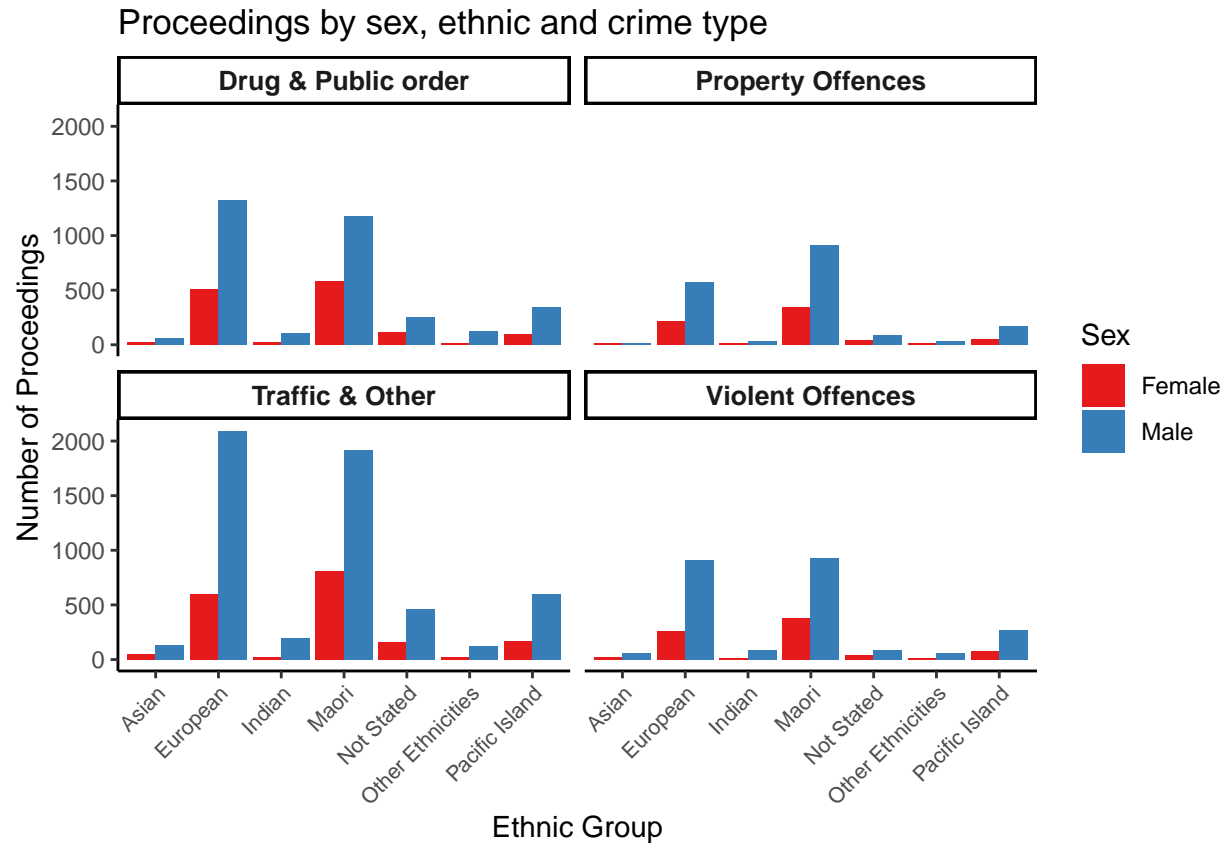
# 3 Result

**3.1 Proceeding by age group, sex and crime type**



Males consistently have much higher numbers than females across all crime types. Adults dominate across categories, especially in "Traffic & Other" and "Violent Offences". In drug and public order, the number of proceedings reached its peak in 2020. Male teenagers show a noticeable decline in property Offences after
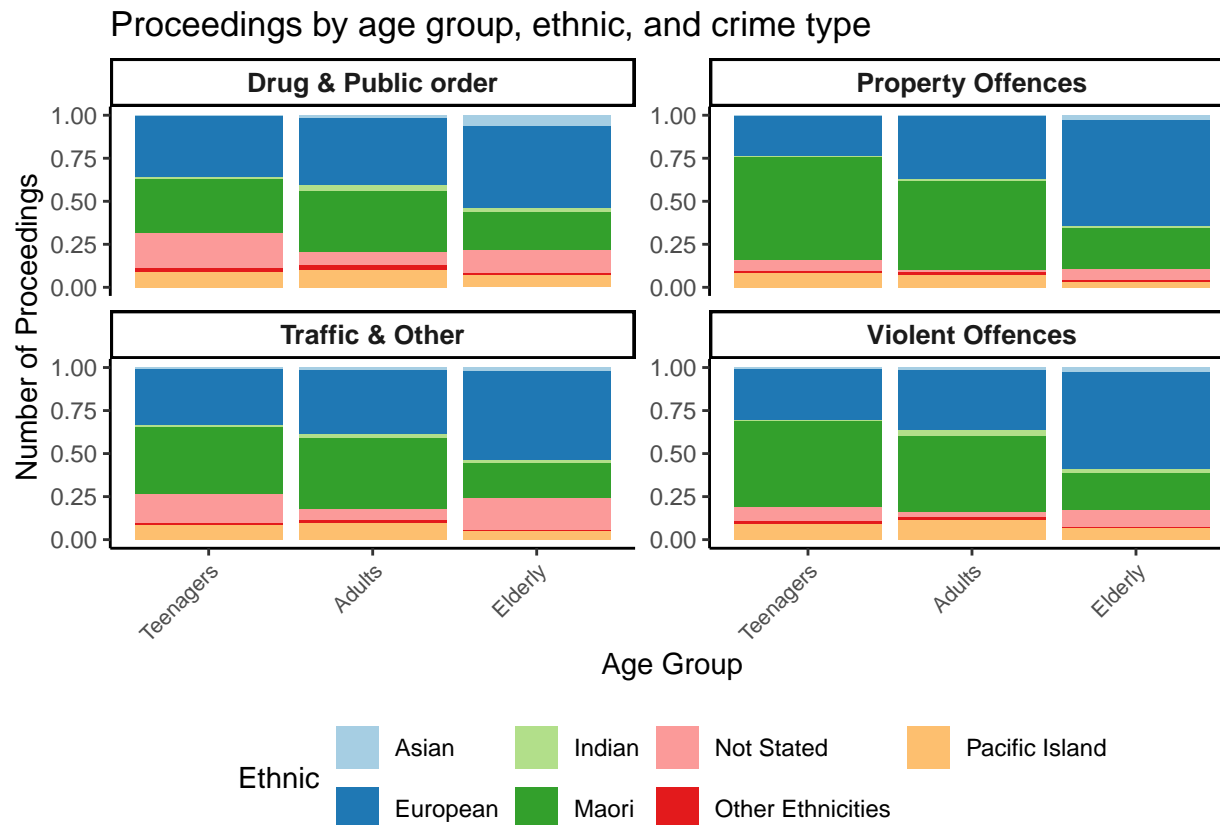
2016, even surpassing that of adult females in 2020. The elderly consistently contribute very little, indicating crime is highly concentrated among younger groups.

**3.2 Proceeding by sex, ethnic and crime type**

## Proceedings by sex, ethnic and crime type



Males within Māori and European groups drive the largest shares of proceedings, especially for "Violent Offences" and "Traffic & Other". It highlights the superimposed effect of ethnicity and gender. The ethnic groups with a relatively higher proportion of females are mainly concentrated in the categories of "Drug & Public Order" and "Property Offences" for Māori and European populations. Other ethnicities (Asian, Indian, Pacific Island) contribute smaller shares.

**3.3 Proceeding by age group, ethnic and crime type**



Proceedings by age group, ethnic, and crime type

Māori and European consistently account for the largest proportion within all age groups. Among Teenagers, Māori make up a disproportionately high share, especially in "Violent Offences" and "Property Offences". Among adults, Europeans have a stronger representation, particularly in "Traffic & Other". Asian and Indian groups remain relatively small across all age groups and crime types.

# 4 Discussion

## 4.1 Conclusion

During the trend from 2014 to 2025, a significant peak was observed in 2020, particularly in the categories of "Drug & Public Order" and "Traffic & Other". This might be related to the COVID-19 pandemic and the lock down policies. During the pandemic, the police's law enforcement priorities were adjusted, with an increase in the recording of cases related to public order and violations of lock down regulations. At the same time, increased social and economic pressures (unemployment, decreased income) might lead to an increase in crimes related to property and traffic.

Among all the crime categories, the number of males is significantly higher than that of females, especially in Property Offences and Violent Offences.

At the ethnic level, Māori and European make up the vast majority, but the distribution patterns are slightly different:

1.The proportion of Europeans is higher in property-related crimes.

2.Māori have a higher proportion in "Drug & Public" order and "Traffic & Other". This may be related to long-term social inequality, insufficient educational and employment opportunities, and the historical background of colonization.

The ethnic proportion structure within each age group remained relatively stable and similar overall. Among the youth population, ethnic differences are relatively small. This might be because at this stage, most individuals are still within the educational system and are subject to certain institutional constraints.

**4.2 Recommendations**

Preventive programs could focus on teenagers, as early intervention may prevent progression into more serious Offences in adulthood.
Increase targeted crime prevention programs for European and Māori communities.
Regarding the high proportion of the "Not Stated" category, it is suggested that the data recording be further improved to prevent the influence on the research conclusions due to data omission.