# Analysis of juvenile crime-trends during COVID-19：Based on different regions and Ethnicity

**by Lin Huang Student ID 23074062**

**2022-09-27**

# Introduction

The COVID-19 pandemic of 2020 heralds dramatic changes in societies around the world. Public health containment measures aimed at reducing the spread of the virus have greatly reduced social contact and movement of people. The initial phase of the COVID-19 pandemic saw a decline in recorded crime rates, largely due to government shutdowns and changes in the population's daily activities. By closing schools and restricting movement and gatherings, containment measures have dramatically altered the critical social and developmental environment in which teens live. Recently, there have been increasing news reports of an increase in juvenile crime in New Zealand due to factors such as higher living costs and an increase in the number of gangs as borders open.

So far, there has been no targeted analysis of the changes in juvenile delinquency base on different regions in the past five years.

This report analysis the changes in New Zealand's youth crime record before and after the introduction of containment measures over a five-year period between Aug 2017 and July 2022.

Recent research has found that the COVID-19 pandemic affects crime volumes differently by crime type and location. While "Unlawful Entry With Intent/Burglary, Break and Enter", "Robbery, Extortion and Related Offences", "Theft and Related Offences"decreased,"Sexual Assault and Related Offences" increased. As the measures change, different trends are shown. Therefore, we selected these four types of crime data for analysis.

In addition, differences in the effects of different Ethnicity will be explored.

```
library(data.table)
library(dplyr)
library(lmtest)
library(ggplot2)
library(kableExtra)
library(stringr)
library(rmarkdown)
library(vars)
library(forecast)
library(ggfortify)
library(gridExtra)
library(changepoint)
library(seasonal)
```

# Methodology

# *Design*

The unit of time analysis is Monthly.Time range between Aug 2017 and July 2022. Juvenile offender ages is from 15 and 24 years. We will remove other ages records. Crime Type present the ANZSOC Division here, and There were in total 4 crime types used in this study out of a total of 16 ANZSOC Division, with 12 ANZSOC Division being removed due to these type comprising no or very small related with Covid-19. Crime type in this report includes: “Unlawful Entry With Intent/Burglary, Break and Enter”“Sexual Assault and Related Offences” “Robbery, Extortion and Related Offences” “Theft and Related Offences”. We remove some data in Ethnicity: “Not Elsewhere Classified”“Organisation”“Not Stated” because these are missing data.

# *Data*

There are one key data sources for this study: New Zealand government Police website policedata.nz,Proceedings (offender demographics). Data sets presents descriptive information about offenders in New Zealand. Data sets includes the number of proceedings against offenders during a relevant period, and provides demographic information about offenders.

# *Measures*

Dependent variable: Crime Number(Youth offending)

Independent variables: Ethnicity,District,Time

# *Analytic approach*

1.Data on monthly juvenile crime were obtained from the official statistics of the New Zealand Police(policedata.nz) for the period between Aug 2017 and July 2022.

Funtion:

- **fread**
- **colnames**
- **sum**

2.Monthly proceedings (offender demographics) were measured based on several offense categories, I choose 4 of them to anylisis: “Unlawful Entry With Intent/Burglary, Break and Enter”, “Robbery, Extortion and Related Offences”, “Theft and Related Offences”,“Sexual Assault and Related Offences”.

Funtion:

- **aggregate**
- **summary**
- **filter**
- **gsub**
- **as.yearmon**

3.Perpetrator ages ranged between 15 and 24 years old as juvenile offender. Clean data, sort out those persons between the ages of 15 and 24 years Everyone is equal before the law, but juvenile crimes do get lighter sentences. Minor rights protection organizations will absolve minors who violate the law; juvenile courts often do not impose severe sentences, and the ultimate punishment for assault, robbery, and wounding is community labor.

Funtion:

- **levels**
- **ggplot**

- **geom_point**
- **geom_line**
- **ggtitle**
- **facet_grid**
- **mean**
- **var**
- **median**

4.Filter by Ethnicity and district

Funtion:

- **geom_boxplot**
- **geom_bar**
- **geom_line**

5.Time-Series Analysis

Time series can reflect the development and change status of a phenomenon. Through the analysis of the time series, the trend and law of the development and change of the phenomenon can be reflected, and then through the determination of various factors affecting the time series, the internal reasons of the phenomenon change can be further explained, and reliable data support for prediction and decision-making can be provided. The trend of juvenile delinquency development in New Zealand can be known through time series. Under normal circumstances, there are four kinds of numerical change laws of time series: long-term change trend, seasonal change law, periodic change law and irregular change. Different numerical variation laws are determined by different influencing factors. These influencing factors include long-term and short-term factors;

some predictable and controllable factors, and some unknown and uncontrollable factors; these factors interact and influence, so that the change trend of the time series presents different characteristics. According to the different influences of influencing factors on the numerical change trend of time series, it can be divided into four influencing factors: long-term trend influencing factors, seasonal variation influencing factors, cyclic variation influencing factors and irregular variation influencing factors. The fluctuation of the juvenile delinquency rate belongs to irregular changes, that is, these irregular changes are numerical changes caused by the announcement and release of the lockdown, which are caused by these random factors. The effects of these factors are unpredictable and irregular. Influence deformation is irregular change.

One of the main threats to internal validity in single-group interrupted time series analysis is the possibility that other factors may influence the outcome variable and are not controlled for in the model. This may lead to bias in the results. Accordingly, it is feasible that a change in lockdown policy could affect crime rates.

Funtion:

- **autoplot**
- **ts**
- **grid.arrange**
- **forecast**

# Results

## *Data processing*

```
ANZSOC_AEG_Data<-fread ("ANZSOC_AEG_Full Data_data.csv",stringsAsFactors=TRUE)
colnames(ANZSOC_AEG_Data)
```

```
## [1] "ANZSOC Division"     "Police District"     "Year Month"
## [4] "Proceedings"         "Age Group"           "Ethnicity"
## [7] "Person/Organisation"
```

```
sum(ANZSOC_AEG_Data$Proceedings)
```

```
## [1] 694800
```

# *Sort out by type*

```
crime_data<-aggregate(ANZSOC_AEG_Data$Proceedings,by=list(ANZSOC_AEG_Data$`ANZSOC Division`,ANZSOC_AEG_Data$`Year Month`,ANZSOC_AEG_Data$Ethnicity,ANZSOC_AEG_Data$`Age Group`,ANZSOC_AEG_Data$`Police District`),sum)
colnames(crime_data)<- c("Type","Year Month","Ethnicity","Age","District","Number")
levels(ANZSOC_AEG_Data$`ANZSOC Division`)
```

```
##  [1] "Abduction, Harassment and Other Related Offences Against a Person"
##  [2] "Acts Intended to Cause Injury"
##  [3] "Dangerous or Negligent Acts Endangering Persons"
##  [4] "Fraud, Deception and Related Offences"
##  [5] "Homicide and Related Offences"
##  [6] "Illicit Drug Offences"
##  [7] "Miscellaneous Offences"
##  [8] "Offences Against Justice Procedures, Govt Sec and Govt Ops"
##  [9] "Prohibited and Regulated Weapons and Explosives Offences"
## [10] "Property Damage and Environmental Pollution"
## [11] "Public Order Offences"
## [12] "Robbery, Extortion and Related Offences"
## [13] "Sexual Assault and Related Offences"
## [14] "Theft and Related Offences"
## [15] "Traffic and Vehicle Regulatory Offences"
## [16] "Unlawful Entry With Intent/Burglary, Break and Enter"
```

```
summary(crime_data)
```

```
##                                                          Type
##  Traffic and Vehicle Regulatory Offences                   :28507
##  Dangerous or Negligent Acts Endangering Persons           :26538
##  Acts Intended to Cause Injury                             :25329
##  Theft and Related Offences                                :18943
##  Public Order Offences                                     :18682
##  Offences Against Justice Procedures, Govt Sec and Govt Ops:18064
##  (Other)                                                   :88668
##    Year Month              Ethnicity          Age
##  Apr-20 :  4048   European      :75237   25-29  :28561
##  May-20 :  4023   Maori         :71774   20-24  :28341
##  Dec-19 :  3965   Pacific Island:26433   15-19  :26820
##  May-18 :  3948   Not Stated    :23804   30-34  :26039
##  May-19 :  3946   Indian        : 9922   35-39  :23045
##  Aug-17 :  3945   Asian         : 7462   40-44  :19793
##  (Other):200856   (Other)       :10099   (Other):72132
##              District         Number
##  Counties/Manukau:25930   Min.   : 1.000
##  Auckland City   :23586   1st Qu.: 1.000
##  Waitemata       :20663   Median : 2.000
##  Wellington      :20577   Mean   : 3.092
##  Canterbury      :19344   3rd Qu.: 4.000
##  Bay Of Plenty   :19212   Max.   :79.000
##  (Other)         :95419
```

```
crime_selected_type<-filter(crime_data,crime_data$Type =="Unlawful Entry With Intent/
Burglary, Break and Enter"|crime_data$Type=="Sexual Assault and Related Offences" |cr
ime_data$Type=="Robbery, Extortion and Related Offences" |crime_data$Type=="Theft and
Related Offences" )
levels(crime_selected_type$`Year Month`)
```

```
##  [1] "Apr-18" "Apr-19" "Apr-20" "Apr-21" "Apr-22" "Aug-17" "Aug-18" "Aug-19"
##  [9] "Aug-20" "Aug-21" "Dec-17" "Dec-18" "Dec-19" "Dec-20" "Dec-21" "Feb-18"
## [17] "Feb-19" "Feb-20" "Feb-21" "Feb-22" "Jan-18" "Jan-19" "Jan-20" "Jan-21"
## [25] "Jan-22" "Jul-18" "Jul-19" "Jul-20" "Jul-21" "Jul-22" "Jun-18" "Jun-19"
## [33] "Jun-20" "Jun-21" "Jun-22" "Mar-18" "Mar-19" "Mar-20" "Mar-21" "Mar-22"
## [41] "May-18" "May-19" "May-20" "May-21" "May-22" "Nov-17" "Nov-18" "Nov-19"
## [49] "Nov-20" "Nov-21" "Oct-17" "Oct-18" "Oct-19" "Oct-20" "Oct-21" "Sep-17"
## [57] "Sep-18" "Sep-19" "Sep-20" "Sep-21"
```

```
crime_selected_type$`Year Month`<-gsub('[-]','',crime_selected_type$`Year Month`)
crime_selected_type$`Year Month`<-as.yearmon(crime_selected_type$`Year Month`, "%b%y"
)
```

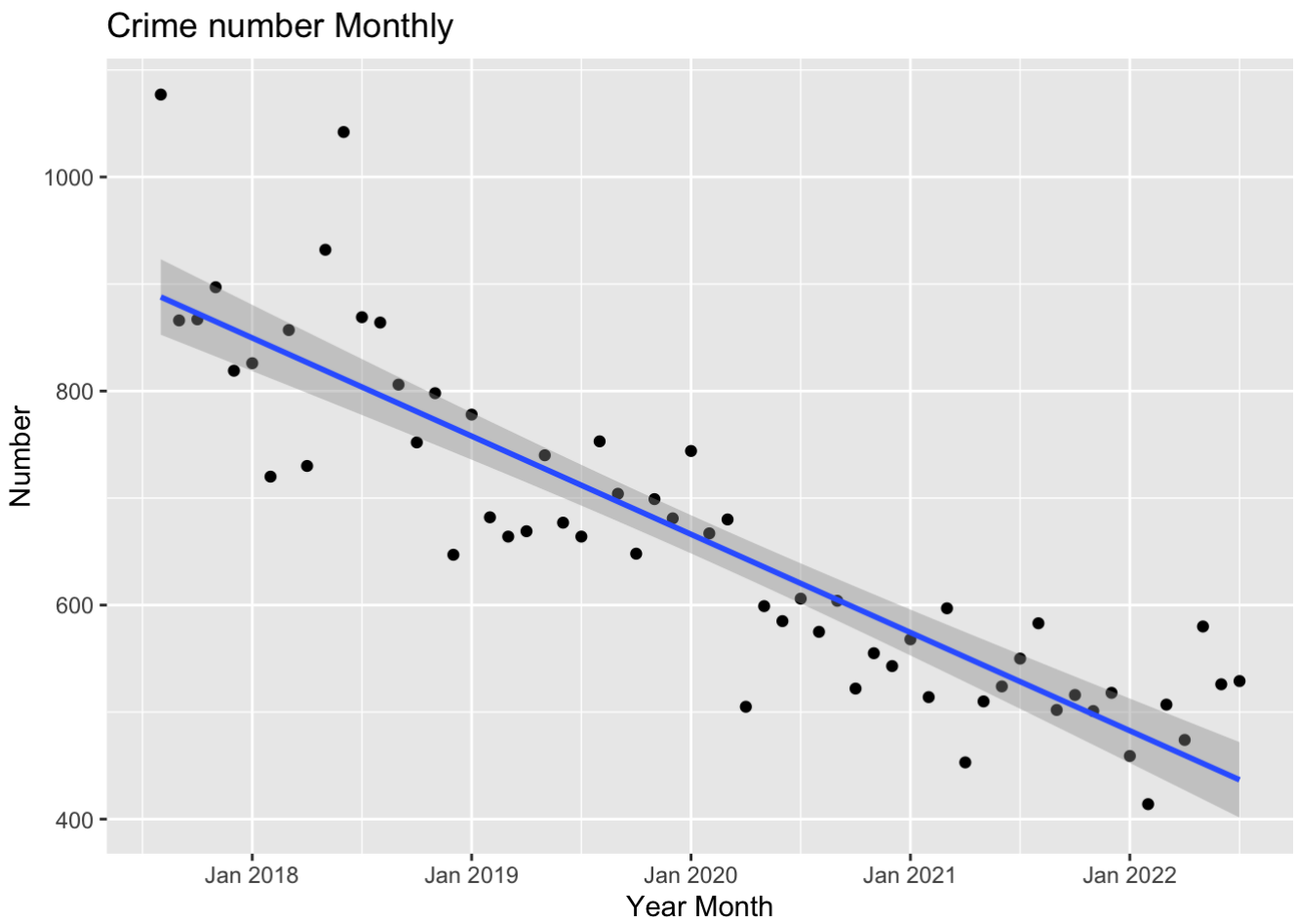
# *select youth offender by age*

```
levels(crime_selected_type$`Ethnicity`)
```

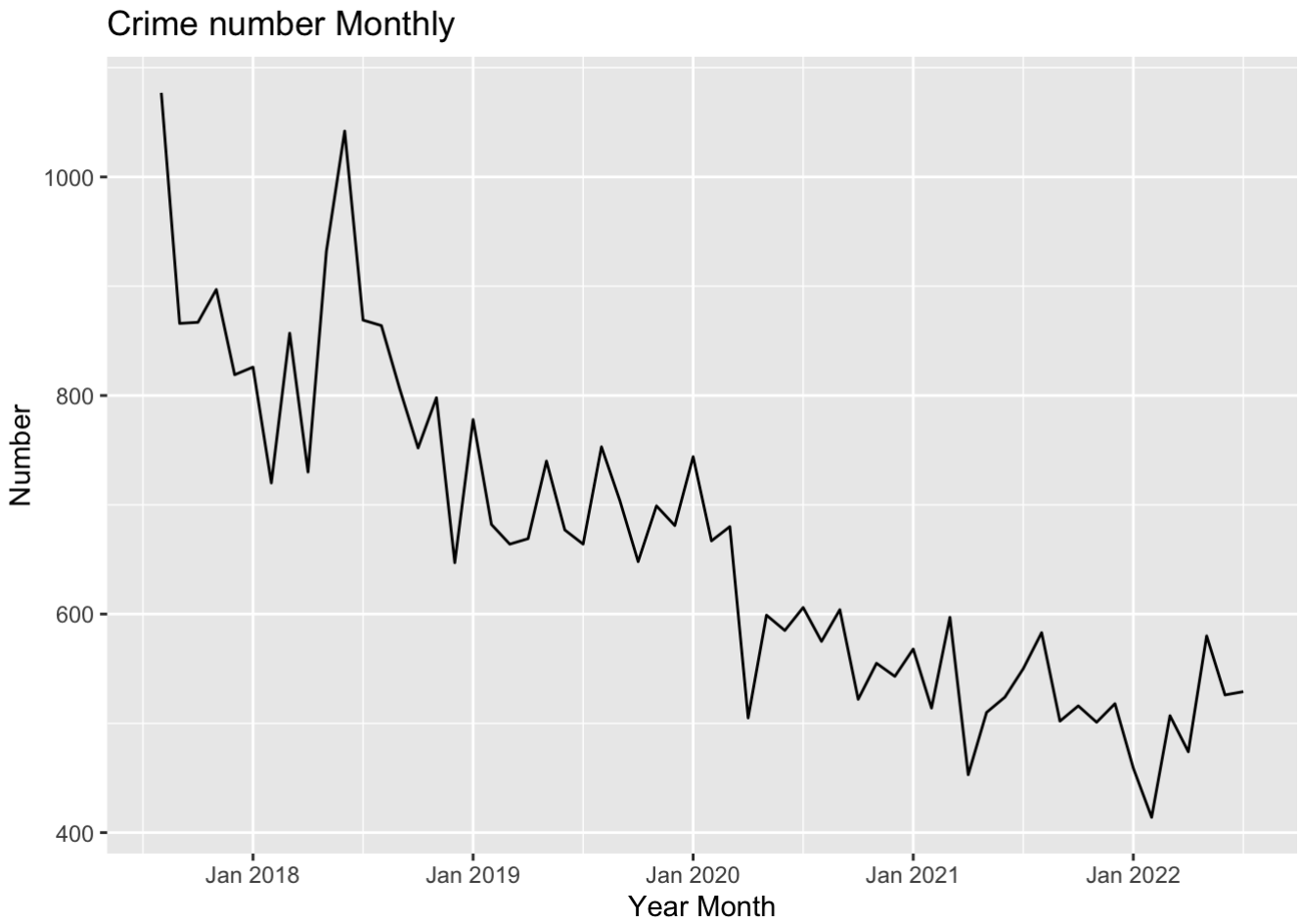
```
##  [1] "African"                   "Asian"
##  [3] "European"                  "Indian"
##  [5] "Latin American/Hispanic"   "Maori"
##  [7] "Middle Eastern"            "Not Elsewhere Classified"
##  [9] "Not Stated"                "Organisation"
## [11] "Pacific Island"
```

```
#crime_selected_type$Age<-gsub('Oct','12',crime_selected_type$Age)
crime_selected_type_age<-filter(crime_selected_type,crime_selected_type$Age =="15-19"
|crime_selected_type$Age =="20-24")

crime_total_data<-aggregate(crime_selected_type_age$Number,by=list(crime_selected_typ
e_age$`Year Month`),sum)
colnames(crime_total_data)<- c("Year Month","Number")

ggplot(crime_total_data, aes(`Year Month`, Number))+ geom_point()+geom_smooth(method=
"lm")+ggtitle("Crime number Monthly")
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
ggplot(data=crime_total_data,aes(x=`Year Month`,y=Number))+geom_line()+ggtitle("Crime
number Monthly")
```




```
mean(crime_total_data$Number)
```

```
## [1] 662.3
```

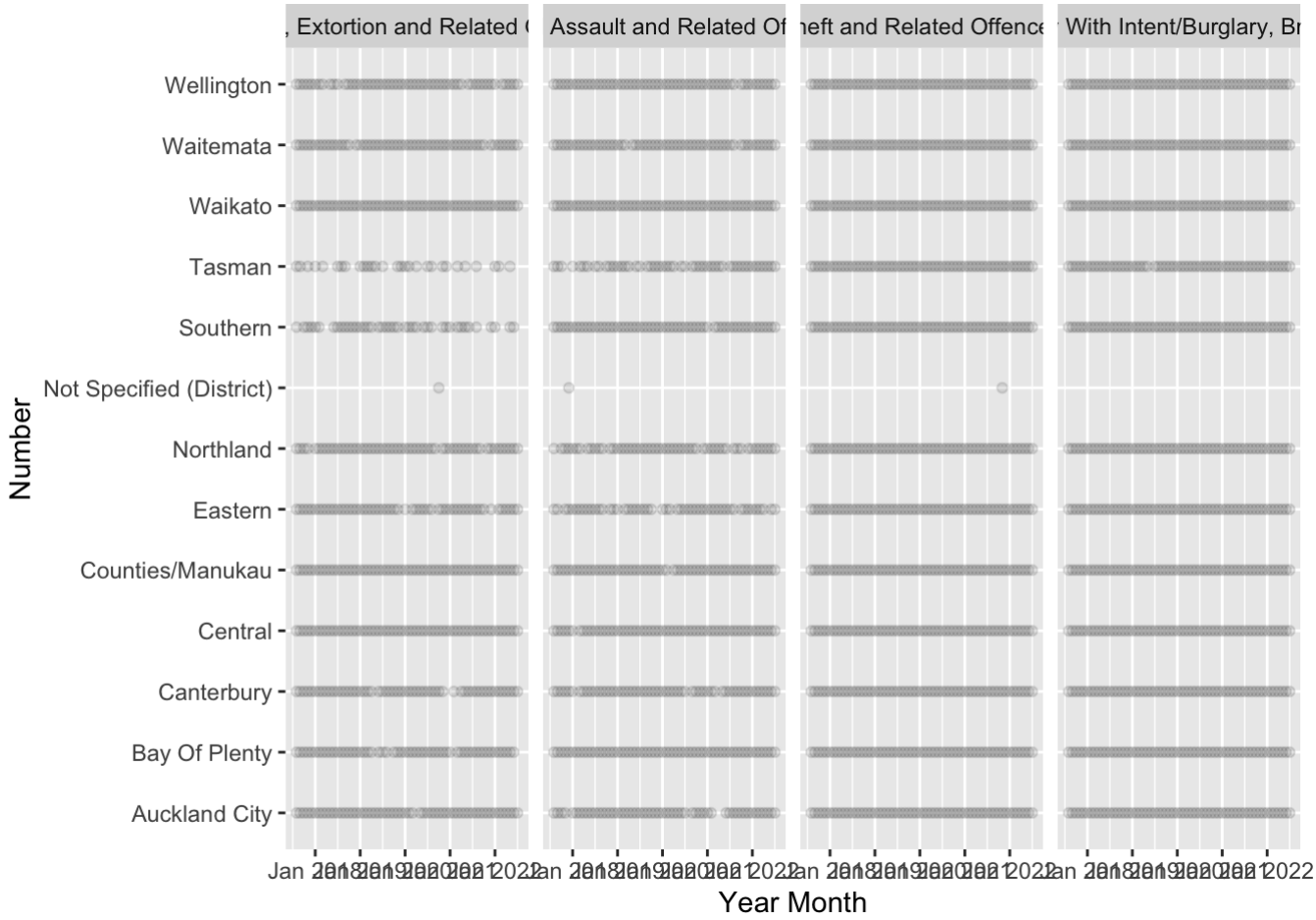
```
var(crime_total_data$Number)
```

```
## [1] 22515.06
```

```
median(crime_total_data$Number)
```

```
## [1] 656
```

```
crime_type_data<-aggregate(crime_selected_type_age$Number,by=list(crime_selected_type
_age$`Year Month`,crime_selected_type_age$Type,crime_selected_type_age$District),sum)
colnames(crime_type_data)<- c("Year Month","Type","Number","District")

ggplot(data=crime_type_data, aes(x=`Year Month`, y=Number)) +geom_point(alpha = 0.1)
 +facet_grid(~Type)
```

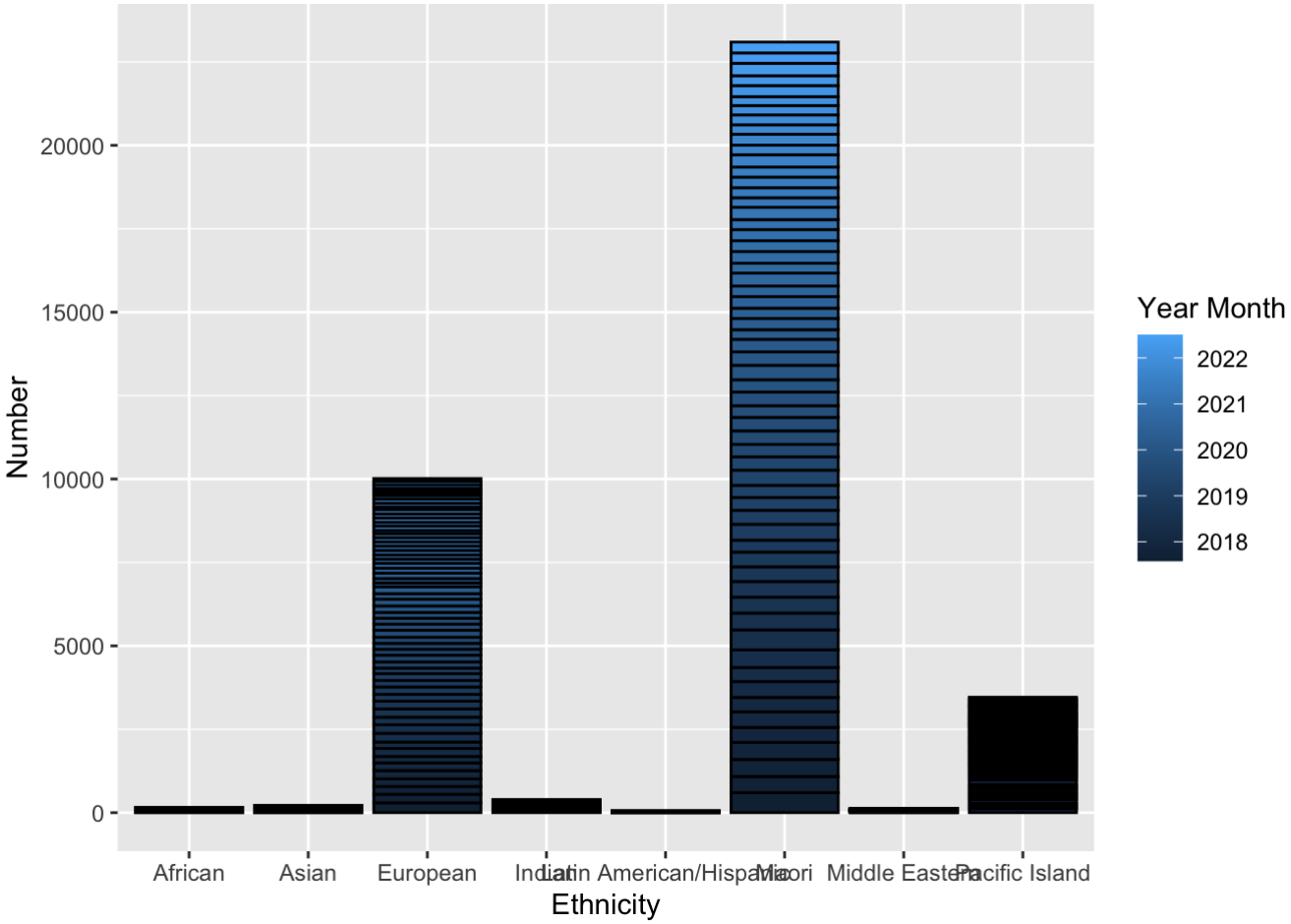



```
#ggplot(data=crime_type_data, aes(x=`Year Month`, fill = Type)) +geom_histogram(bins
 = 50)
```

# *Filter by Ethnicity and district*

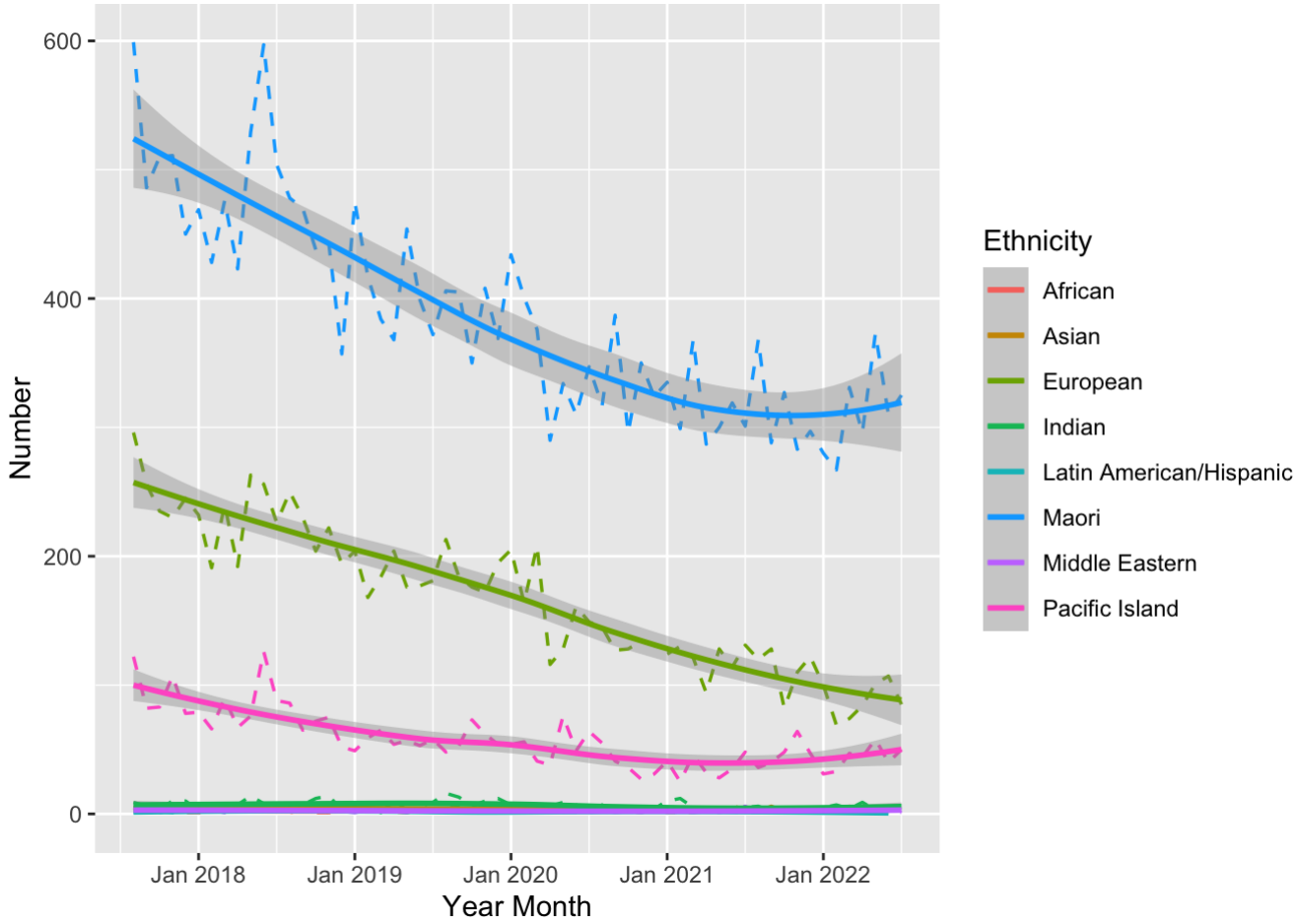
We can see that Māori youth are the most criminal, followed by Europeans. Maybe it has something to do with their education level, family income and population proportion, we need to do further data extraction and correlation analysis.

```
crime_Ethnicity_data<-aggregate(crime_selected_type_age$Number,by=list(crime_selected
_type_age$`Year Month`,crime_selected_type_age$Ethnicity),sum)
colnames(crime_Ethnicity_data)<-c("Year Month","Ethnicity","Number")

crime_Ethnicity_data$Ethnicity[crime_Ethnicity_data$Ethnicity =="Not Elsewhere Classi
fied"|crime_Ethnicity_data$Ethnicity =="Organisation"|crime_Ethnicity_data$Ethnicity
 =="Not Stated"]<-NA
crime_Ethnicity_data<-na.omit(crime_Ethnicity_data)

#barplot
ggplot(data=crime_Ethnicity_data, aes(x=Ethnicity, y=Number,fill=`Year Month`)) +geom
_bar(stat="identity",color="black")
```




```
# line types
ggplot(data=crime_Ethnicity_data, aes(x=`Year Month`, y=Number, group=Ethnicity,color
=Ethnicity)) +
  geom_line(linetype="dashed", size=0.6)+geom_smooth(method="loess")
```

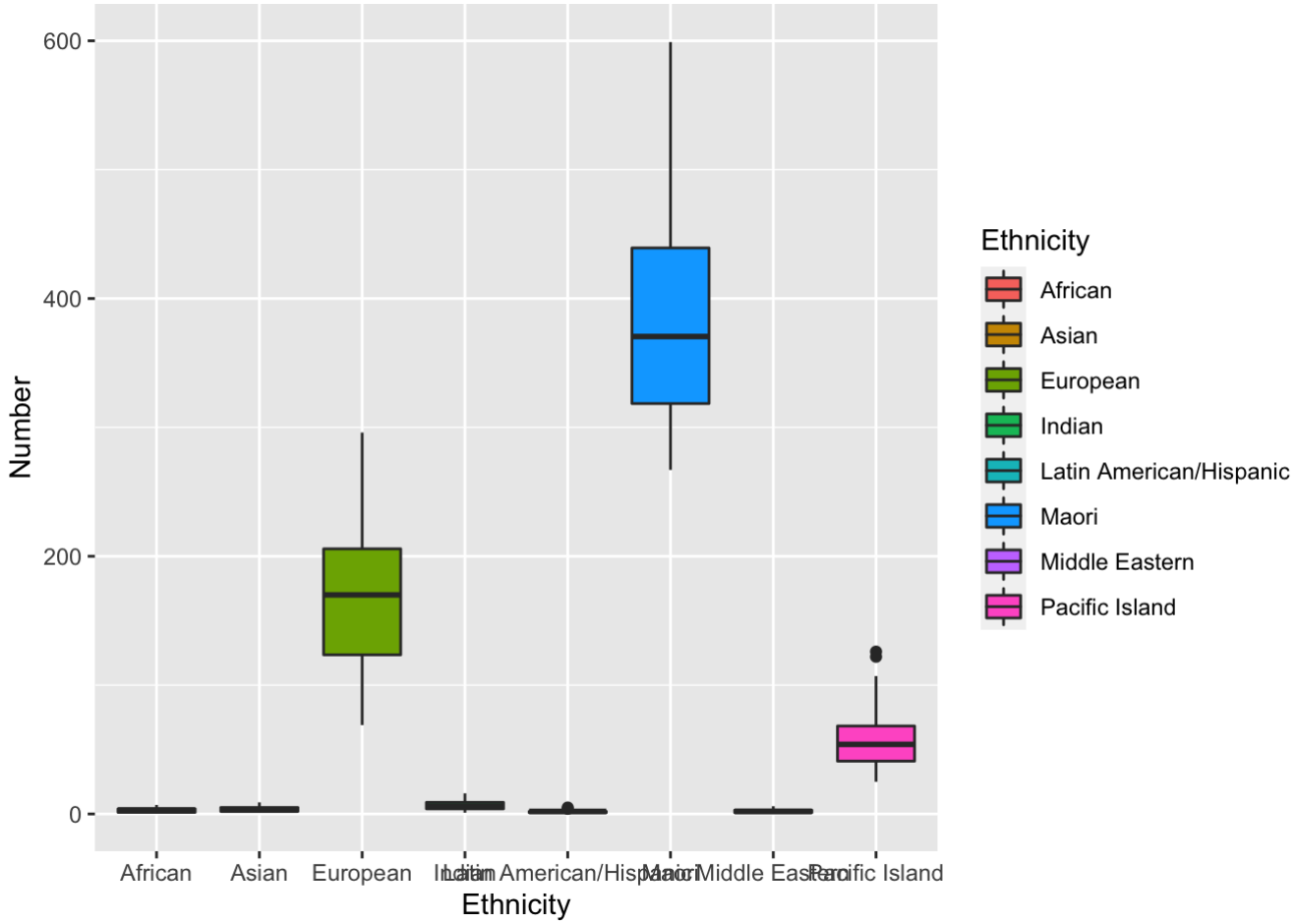
```
## `geom_smooth()` using formula 'y ~ x'
```




```
  geom_point(color="red", size=1)
```

```
## geom_point: na.rm = FALSE
## stat_identity: na.rm = FALSE
## position_identity
```

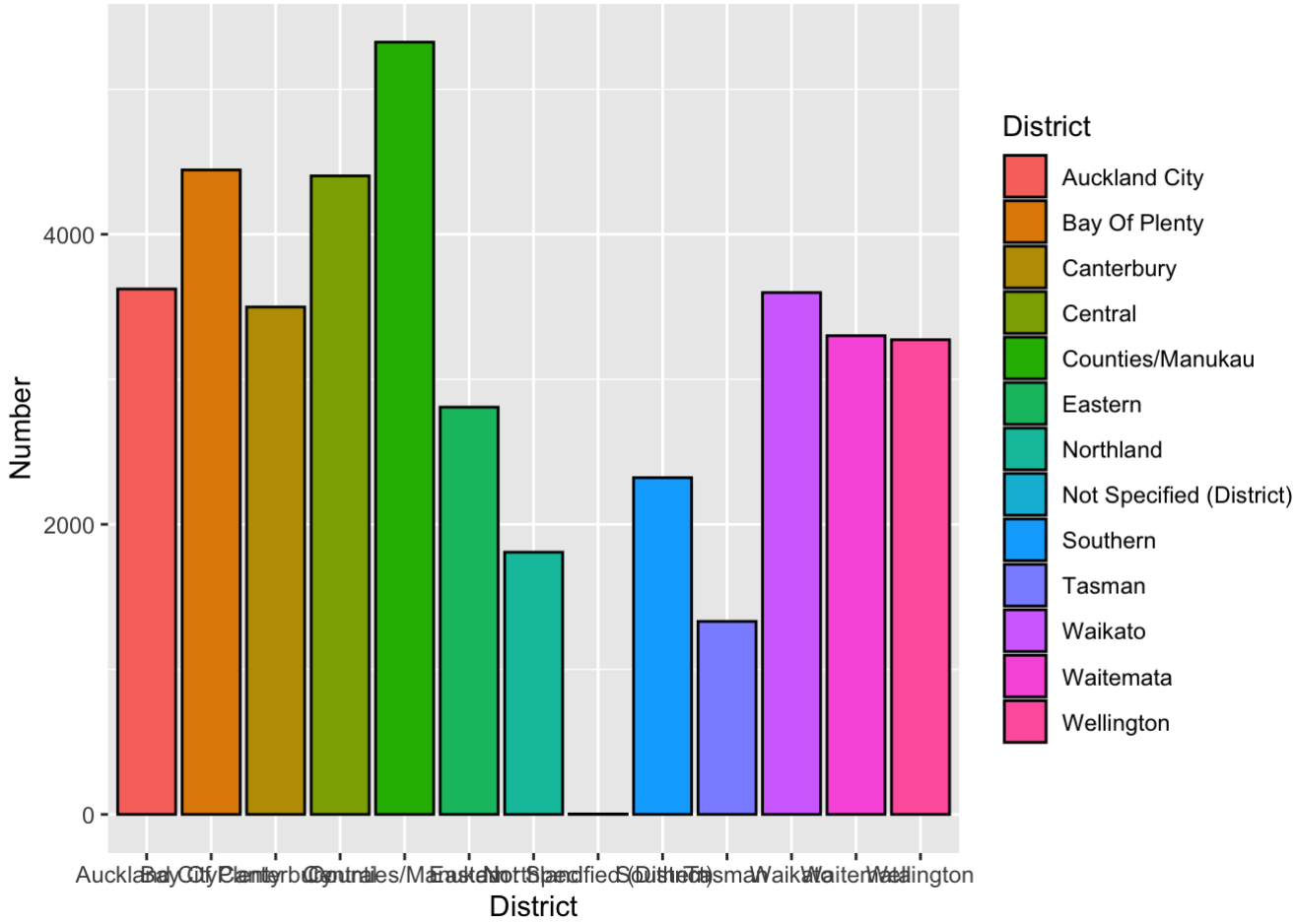
```
ggplot(data=crime_Ethnicity_data, aes(x=Number, y=Ethnicity,fill =Ethnicity)) +geom_b
oxplot()+coord_flip()
```




```
crime_district_data<-aggregate(crime_selected_type_age$Number,by=list(crime_selected_
type_age$District),sum)
summary(crime_district_data)
```

```
##              Group.1        x
##  Auckland City    :1   Min.   :   3
##  Bay Of Plenty    :1   1st Qu.:2322
##  Canterbury       :1   Median :3301
##  Central          :1   Mean   :3057
##  Counties/Manukau:1   3rd Qu.:3623
##  Eastern          :1   Max.   :5325
##  (Other)          :7
```

```
colnames(crime_district_data)<-c("District","Number")

#barplot
ggplot(data=crime_district_data, aes(x=District, y=Number,fill=District)) +geom_bar(s
tat="identity",color="black")
```




# *Time-Series Analysis*

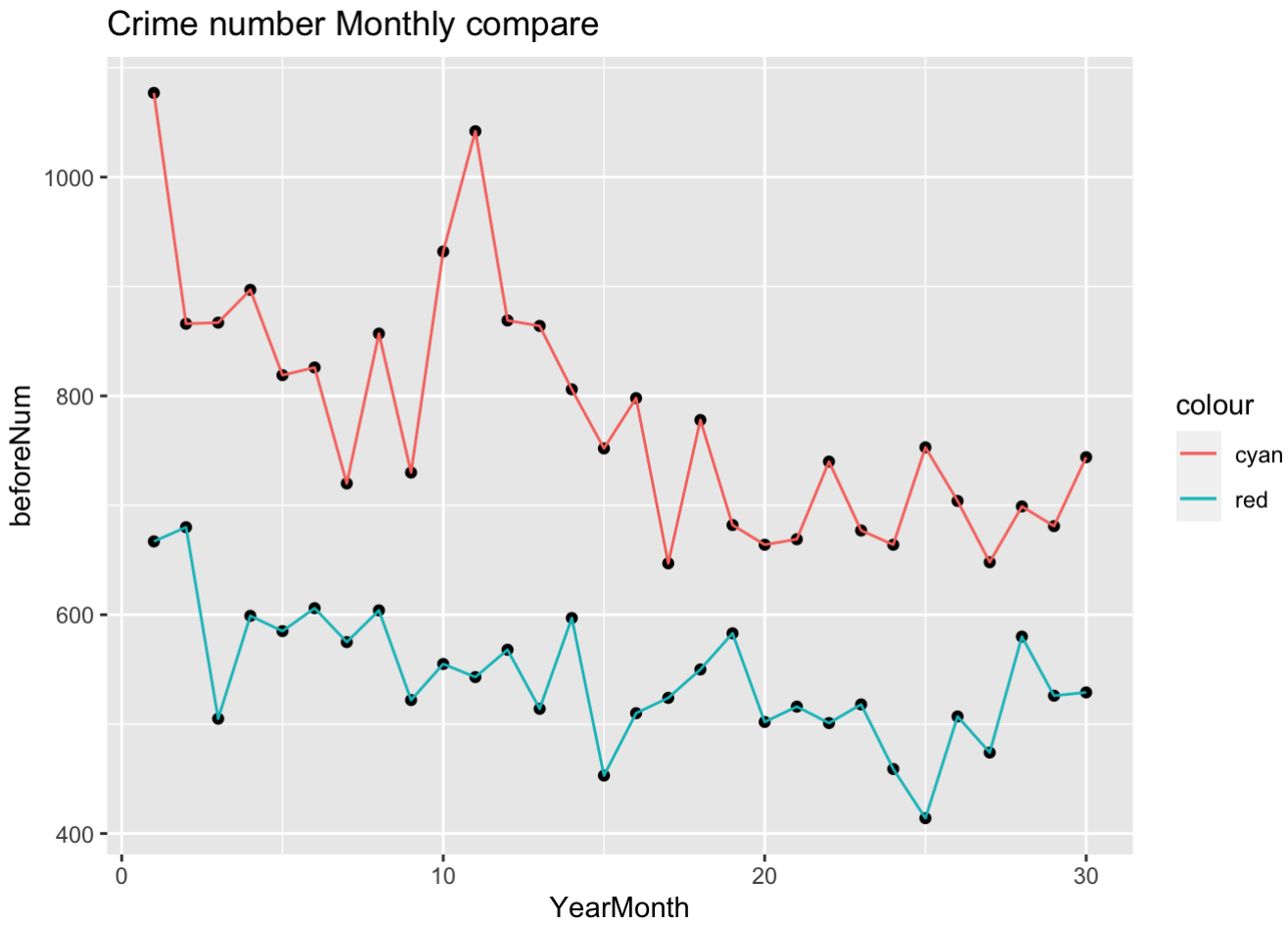
In the preliminary analysis, we focused on the temporal patterns of monthly juvenile crime rates between 8/2017-1/2020 and 2/2020-7/2022.

Before Covid-19:8/2017-1/2020; After break out： 2/2020-7/2022.

Figures 1, 2, 3, 4, 5, and 6 present crime rate trends for each of the six categories during the abovementioned period, i.e., before the outbreak of COVID-19 and during the lockdown periods, the first of which was between March 25 and May 4, 2020 and the second between September 25 and October 17, 2020.

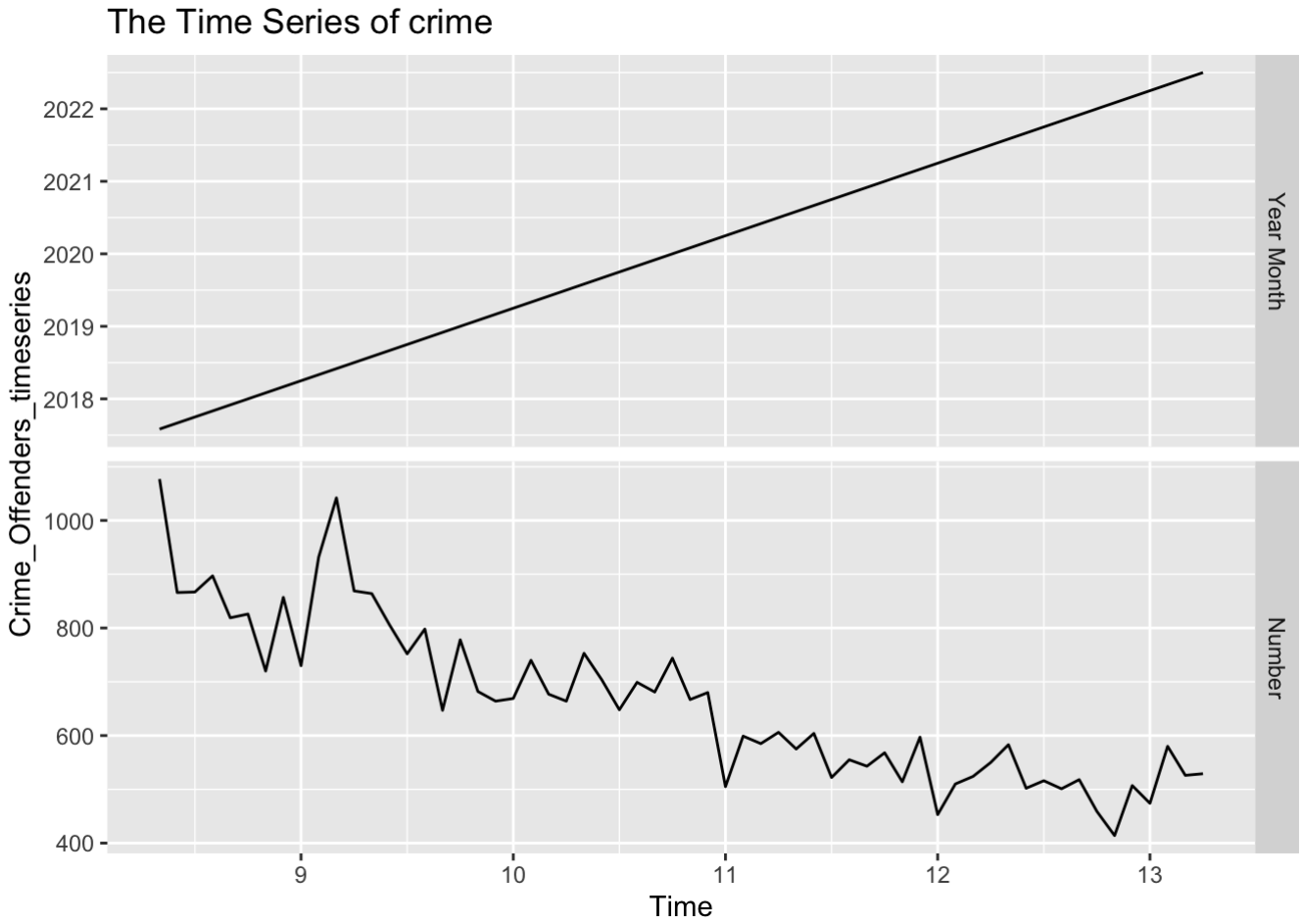
We can see the crime numbers are both decreased during these period.

```
crime_before_data<-crime_total_data[crime_total_data$`Year Month`<'Feb 2020',]
crime_after_data<-crime_total_data[crime_total_data$`Year Month`>='Feb 2020',]
crime_before_after_compare<-crime_before_data
crime_before_after_compare$YearMonth<-c(1:30)
crime_before_after_compare$beforeNum<-as.numeric(crime_before_data$Number)
crime_before_after_compare$afterNum<-as.numeric(crime_after_data$Number)
crime_before_after_compare$`Year Month`<-NULL
crime_before_after_compare$Number<-NULL
ggplot(crime_before_after_compare, aes(x=YearMonth)) +
            geom_point(aes(y=beforeNum) , ) + geom_line(aes(y=beforeNum,color="cyan"
)) +
            geom_point(aes(y=afterNum), ) + geom_line(aes(y=afterNum,color="red"))+ g
gtitle("Crime number Monthly compare")
```

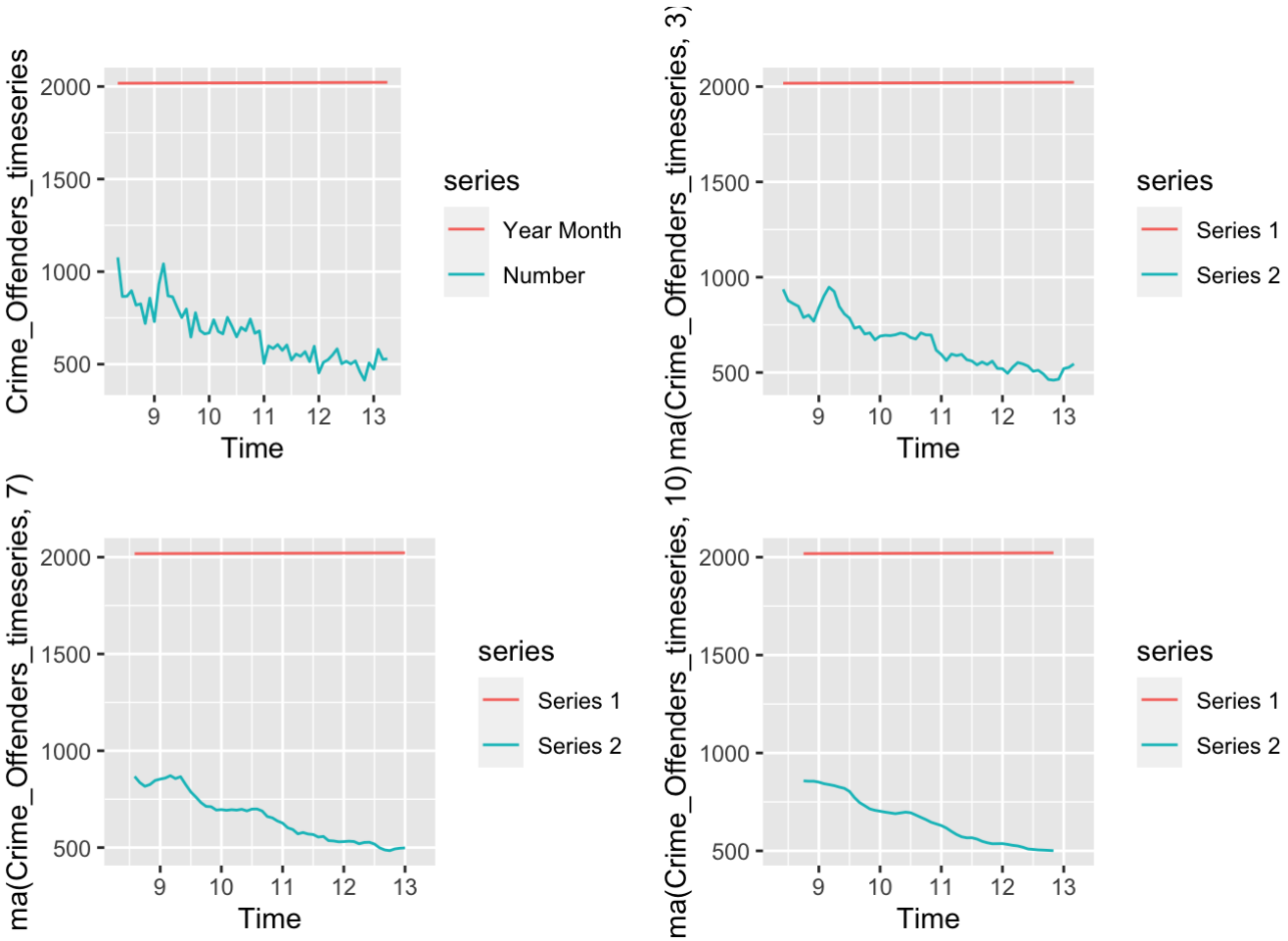



*## Forecast: It is predicted that juvenile delinquency rates will continue to decline over time.*

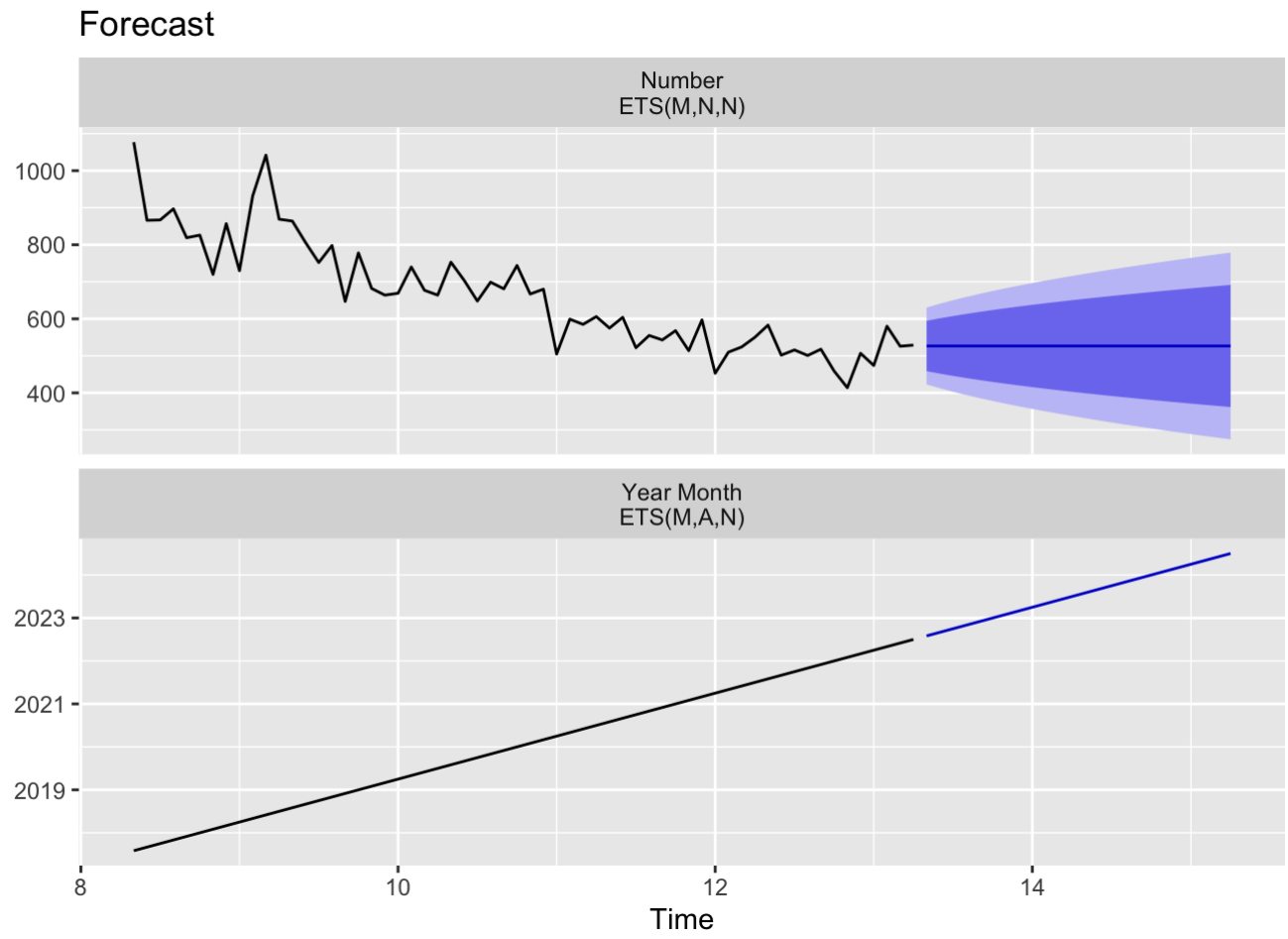
```
##Time Series
Crime_Offenders_timeseries<-ts(crime_total_data,start(2017,8),frequency=12)
autoplot(Crime_Offenders_timeseries,facets = TRUE) + ggtitle("The Time Series of crim
e")
```




```
a <- autoplot(ma(Crime_Offenders_timeseries,3))
b <- autoplot(ma(Crime_Offenders_timeseries,7))
c <- autoplot(ma(Crime_Offenders_timeseries,10))
d <- autoplot(Crime_Offenders_timeseries)
grid.arrange(d,a,b,c,ncol=2)
```



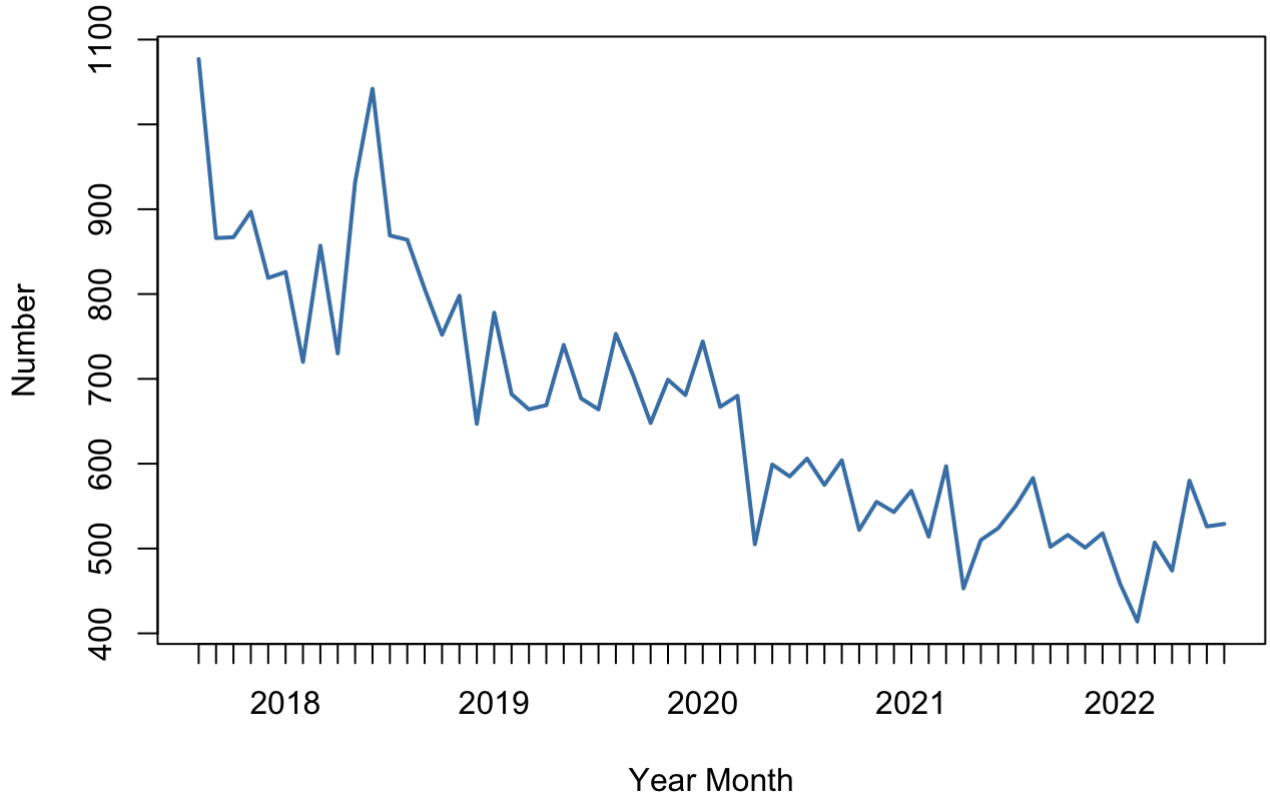

```
##forecast
forecast_Crime_Offenders_timeseries<-forecast(Crime_Offenders_timeseries)
autoplot(forecast_Crime_Offenders_timeseries)+ ggtitle("Forecast")
```




```
##change points
shapiro.test(crime_total_data$Number)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  crime_total_data$Number
## W = 0.95074, p-value = 0.01688
```

```
#cpm.res = processStream(crime_total_data$Number, cpmType = "Kolmogorov-Smirnov")
plot(crime_total_data, type = "l", col = "steelblue", lwd = 2)
```




```
#+abline(v=cpm.res$changePoints)
```

cpm.res = processStream(crime_total_data$Number, cpmType = “Kolmogorov-Smirnov”) cpm.res $changePoints [1] 13 32

$detectionTimes [1] 23 38

*change point： To analyze the impact of the closure measures on the epidemic, each time point has a turning point. Based on the previous time series, change point analyzes whether the turning point is related to the announcement and lifting of the lockdown.*

1.Shapiro-Wilk normality test p-value = 0.01688 which is less than 0.05,Check whether the data obeys the Gaussian distribution and find that it does not obey. So choose a non-Gaussian distribution method.

data: crime_total_data$Number W = 0.95074, p-value = 0.01688

2.regression analysis

regmonthcrime<-lm( `Year Month` ~ Number, data = crime_total_data)

Call: lm(formula = `Year Month` ~ Number, data = crime_total_data)

Coefficients: (Intercept) Number
2.026e+03 -8.632e-03

# Discussion

Crime numbers are decreased not only before covid-19, but also covid-19 break out.

Figures 1, 2, 3, 4, 5, and 6 present crime rate trends for each of the six categories during the abovementioned period, i.e., before the outbreak of COVID-19 and during the lockdown periods, the first of which was between March 25 and May 4, 2020 and the second between September 25 and October 17, 2020.We can see the crime numbers siginifican decrease.

We can see that Māori youth are the most criminal, followed by Europeans. While this may be due to a variety of factors, it is clear that there is a problem that needs to be addressed. Māori youth need to be given the opportunity to succeed, and to feel like they are a part of society. Europeans need to be taught to respect Māori culture and to understand the issues that Māori face. Maybe it has something to do with their education level, family income and population proportion, we need to do further data extraction and correlation analysis.

This research found out some changes in crime numbers during this period according to crime type and location. Break points could identify if the changes cause by the lockdown policy or not, which indicate that further research is required. Expanding the analysis to other countries and to online juvenile crime activity would provide further insight in this area.

# References

Sela-Shayovitz Revital & Noam Haviv. Juvenile delinquency and COVID-19: the effect of social distancing restrictions on juvenile crime rates in Israel

Molly McCarthy molly.mccarthy, Jacqueline Homel, James Ogilvie and Troy Allard.Initial impacts of COVID-19 on youth offending: An exploration of differences across communities

policedata.nz Proceedings (offender demographics) (June2018)