

# Exploring the Crime Landscape in British Colombia, Canada

Casper Kacper Ruta

2024-02-25

```
# Loading and Installing the Necessary Documents.
packages <- c("here",
              "readr",
              "stringr",
              "clock",
              "dplyr",
              "ggplot2",
              "rmarkdown",
              "knitr",
              "magrittr",
              "glue",
              "xfun",
              "fs",
              "tidyr")

installed_packages <- packages %in% rownames(installed.packages())

if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])
}

library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

# Basic Summary of the Data
crime_data <- read.csv("bc_crime.csv")
summary(crime_data)
```

```
##      TYPE      YEAR      MONTH      DAY
```

```
## Length:881242      Min.   :2003      Min.   : 1.000      Min.   : 1.00
## Class :character    1st Qu.:2007      1st Qu.: 4.000      1st Qu.: 8.00
## Mode  :character    Median :2012      Median : 7.000      Median :15.00
##                      Mean   :2012      Mean   : 6.523      Mean   :15.38
##                      3rd Qu.:2018      3rd Qu.: 9.000      3rd Qu.:23.00
##                      Max.   :2023      Max.   :12.000      Max.   :31.00
##
##      HOUR           MINUTE      HUNDRED_BLOCK      NEIGHBOURHOOD
## Min.   : 0.00      Min.   : 0.00      Length:881242      Length:881242
## 1st Qu.: 7.00      1st Qu.: 0.00      Class :character    Class :character
## Median :14.00      Median : 5.00      Mode  :character    Mode  :character
## Mean   :12.31      Mean   :15.86
## 3rd Qu.:18.00      3rd Qu.:30.00
## Max.   :23.00      Max.   :59.00
##
##      X              Y
## Min.   :      0      Min.   :      0
## 1st Qu.:490187      1st Qu.:5454209
## Median :491570      Median :5457169
## Mean   :448947      Mean   :4977180
## 3rd Qu.:493380      3rd Qu.:5458718
## Max.   :511303      Max.   :5512579
## NA's   :75          NA's   :75
```

```
head(crime_data)
```

```
##              TYPE YEAR MONTH DAY HOUR MINUTE  HUNDRED_BLOCK
## 1 Break and Enter Commercial 2012      12  14    8    52
## 2 Break and Enter Commercial 2019      3   7    2     6  10XX SITKA SQ
## 3 Break and Enter Commercial 2019      8  27    4    12  10XX ALBERNI ST
## 4 Break and Enter Commercial 2021      4  26    4    44  10XX ALBERNI ST
## 5 Break and Enter Commercial 2014      8   8    5    13  10XX ALBERNI ST
## 6 Break and Enter Commercial 2020      7  28   19    12  10XX ALBERNI ST
## NEIGHBOURHOOD      X      Y
## 1      Oakridge 491285.0 5453433
## 2      Fairview 490613.0 5457110
## 3      West End 491004.8 5459177
## 4      West End 491007.8 5459174
## 5      West End 491015.9 5459166
## 6      West End 491015.9 5459166
```

```
# Creating a basic dataframe.
crime_counts <- crime_data %>%
  filter(YEAR >= 2019) %>%
  group_by(YEAR, TYPE) %>%
  summarise(Count = n(), .groups = 'drop') %>%
  arrange(YEAR, desc(Count))

print(crime_counts)
```

```
## # A tibble: 55 x 3
##   YEAR TYPE      Count
##   <int> <chr>    <int>
```

```
## 1 2019 Theft from Vehicle 17005
## 2 2019 Other Theft 12323
## 3 2019 Mischief 5833
## 4 2019 Offence Against a Person 3473
## 5 2019 Break and Enter Commercial 2490
## 6 2019 Break and Enter Residential/Other 2269
## 7 2019 Theft of Bicycle 2221
## 8 2019 Theft of Vehicle 1378
## 9 2019 Vehicle Collision or Pedestrian Struck (with Injury) 1138
## 10 2019 Vehicle Collision or Pedestrian Struck (with Fatality) 14
## # i 45 more rows
```

```
# Creating a basic dataframe.
crime_counts <- crime_data %>%
  filter(YEAR >= 2020) %>%
  group_by(YEAR, TYPE) %>%
  summarise(Count = n(), .groups = 'drop') %>%
  arrange(YEAR, desc(Count))

print(crime_counts)
```

```
## # A tibble: 44 x 3
##   YEAR TYPE Count
##   <int> <chr> <int>
## 1 2020 Theft from Vehicle 10430
## 2 2020 Other Theft 8647
## 3 2020 Mischief 6109
## 4 2020 Offence Against a Person 3742
## 5 2020 Break and Enter Commercial 2788
## 6 2020 Break and Enter Residential/Other 2083
## 7 2020 Theft of Bicycle 1987
## 8 2020 Theft of Vehicle 853
## 9 2020 Vehicle Collision or Pedestrian Struck (with Injury) 853
## 10 2020 Homicide 19
## # i 34 more rows
```

```
# Creating a basic dataframe.
crime_counts <- crime_data %>%
  filter(YEAR >= 2021) %>%
  group_by(YEAR, TYPE) %>%
  summarise(Count = n(), .groups = 'drop') %>%
  arrange(YEAR, desc(Count))

print(crime_counts)
```

```
## # A tibble: 33 x 3
##   YEAR TYPE Count
##   <int> <chr> <int>
## 1 2021 Other Theft 8583
## 2 2021 Theft from Vehicle 7502
## 3 2021 Mischief 5415
## 4 2021 Offence Against a Person 3788
## 5 2021 Break and Enter Commercial 2079
```

```
## 6 2021 Break and Enter Residential/Other 1488
## 7 2021 Theft of Bicycle 1459
## 8 2021 Vehicle Collision or Pedestrian Struck (with Injury) 1021
## 9 2021 Theft of Vehicle 838
## 10 2021 Vehicle Collision or Pedestrian Struck (with Fatality) 22
## # i 23 more rows
```

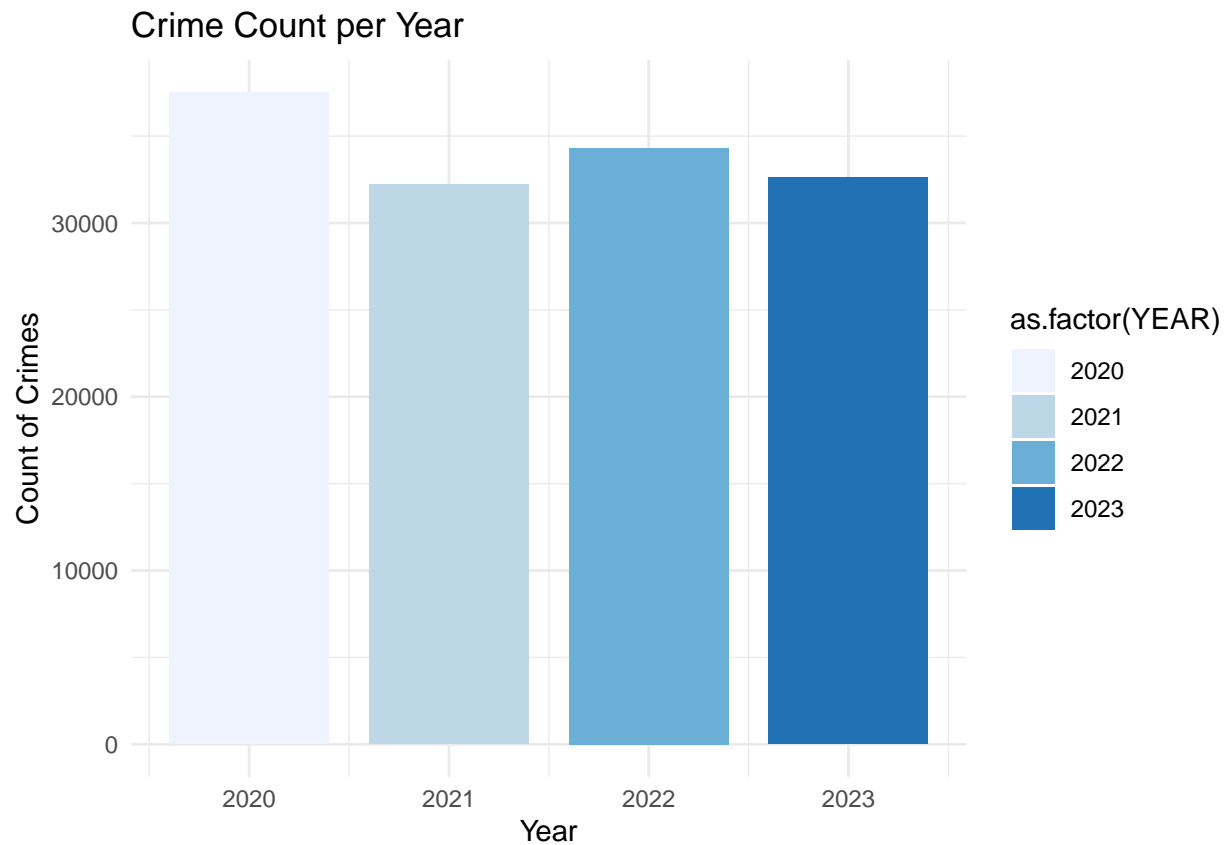
```
homicide_count <- crime_data %>%
  filter(YEAR %in% c(2019, 2020, 2021, 2022, 2023) & TYPE == "Homicide") %>%
  group_by(YEAR) %>%
  summarise(Count = n())

print(homicide_count)
```

```
## # A tibble: 5 x 2
##   YEAR Count
##   <int> <int>
## 1 2019    12
## 2 2020    19
## 3 2021    19
## 4 2022    11
## 5 2023    14
```

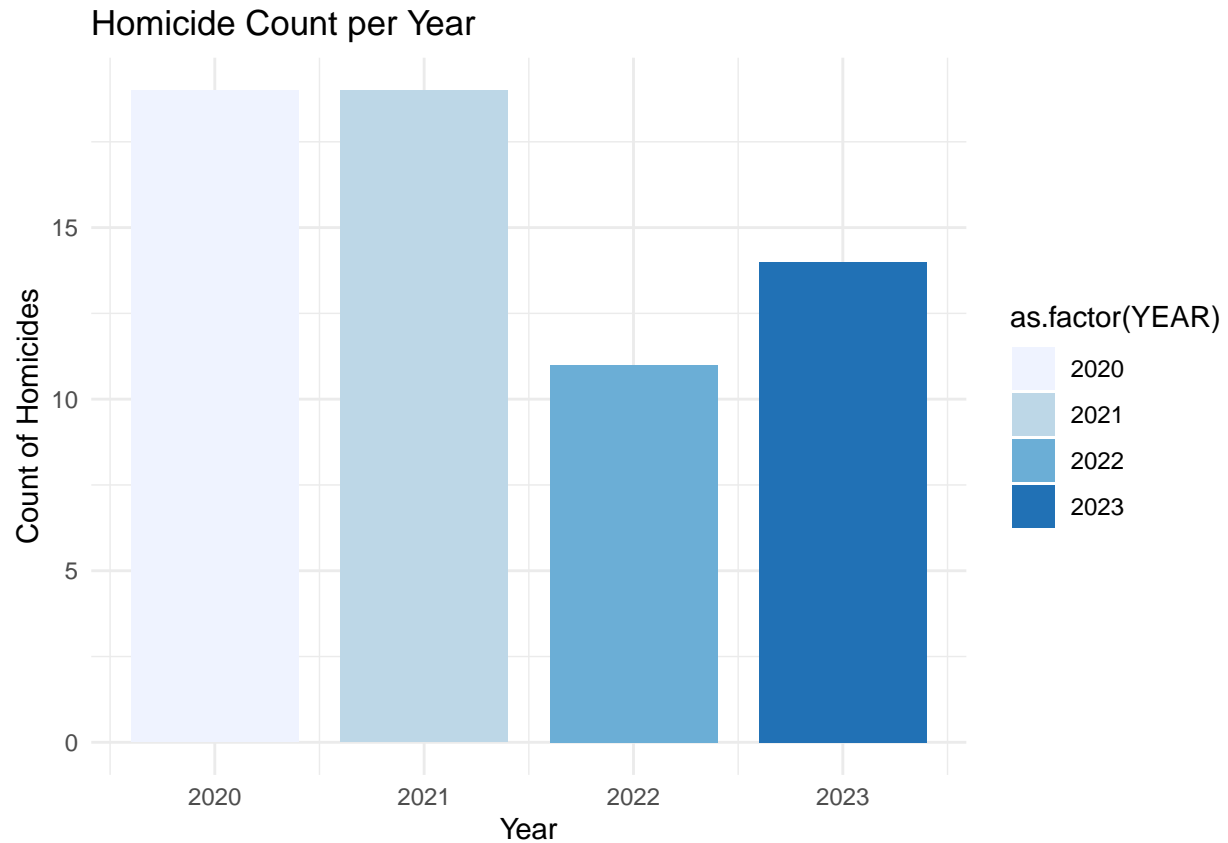
```
crime_data_selected_years <- crime_data %>%
  filter(YEAR %in% c(2020, 2021, 2022, 2023)) %>%
  count(YEAR)

ggplot(crime_data_selected_years, aes(x = YEAR, y = n, fill = as.factor(YEAR))) +
  geom_bar(stat = "identity", width = 0.8) +
  scale_fill_brewer(palette = "Blues") +
  labs(title = "Crime Count per Year", x = "Year", y = "Count of Crimes") +
  theme_minimal()
```



```
# Selecting for Homicides in Each Year Post Covid-19
homicide_data_years <- crime_data %>%
  filter(YEAR %in% c(2020, 2021, 2022, 2023) & TYPE == "Homicide") %>%
  count(YEAR)

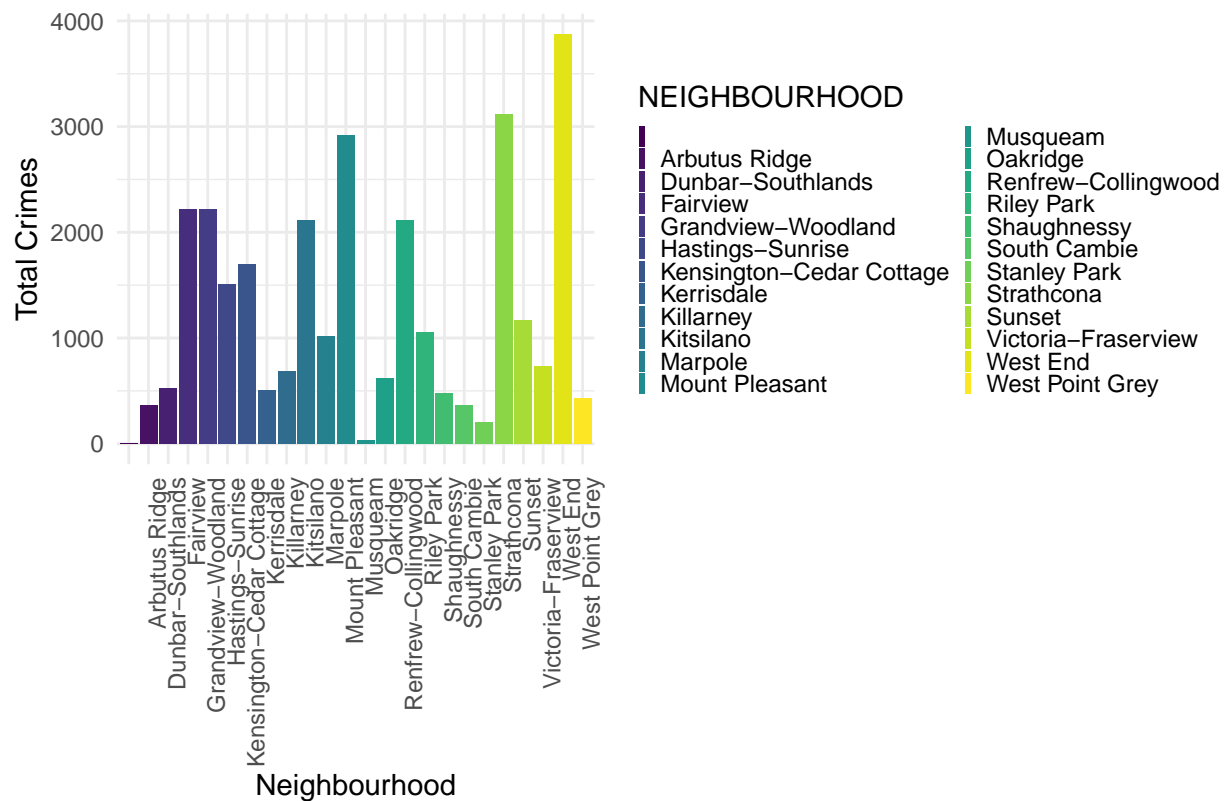
# Creating the Bar Chart
ggplot(homicide_data_years, aes(x = YEAR, y = n, fill = as.factor(YEAR))) +
  geom_bar(stat = "identity", width = 0.8) +
  scale_fill_brewer(palette = "Blues") +
  labs(title = "Homicide Count per Year", x = "Year", y = "Count of Homicides") +
  theme_minimal()
```



```
# Filter and aggregate data for 2018 excluding Central Business District.
crime_data_2018_total <- crime_data %>%
  filter(YEAR == 2018, NEIGHBOURHOOD != "Central Business District") %>%
  count(NEIGHBOURHOOD)

# Create a colorful bar chart for 2018
ggplot(crime_data_2018_total, aes(x = NEIGHBOURHOOD, y = n, fill = NEIGHBOURHOOD)) +
  geom_bar(stat = "identity") +
  labs(title = "Total Crimes by Neighbourhood in 2019", x = "Neighbourhood", y = "Total Crimes") +
  theme_minimal() +
  theme(legend.key.size = unit(0.1, "cm"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  scale_fill_viridis_d()
```

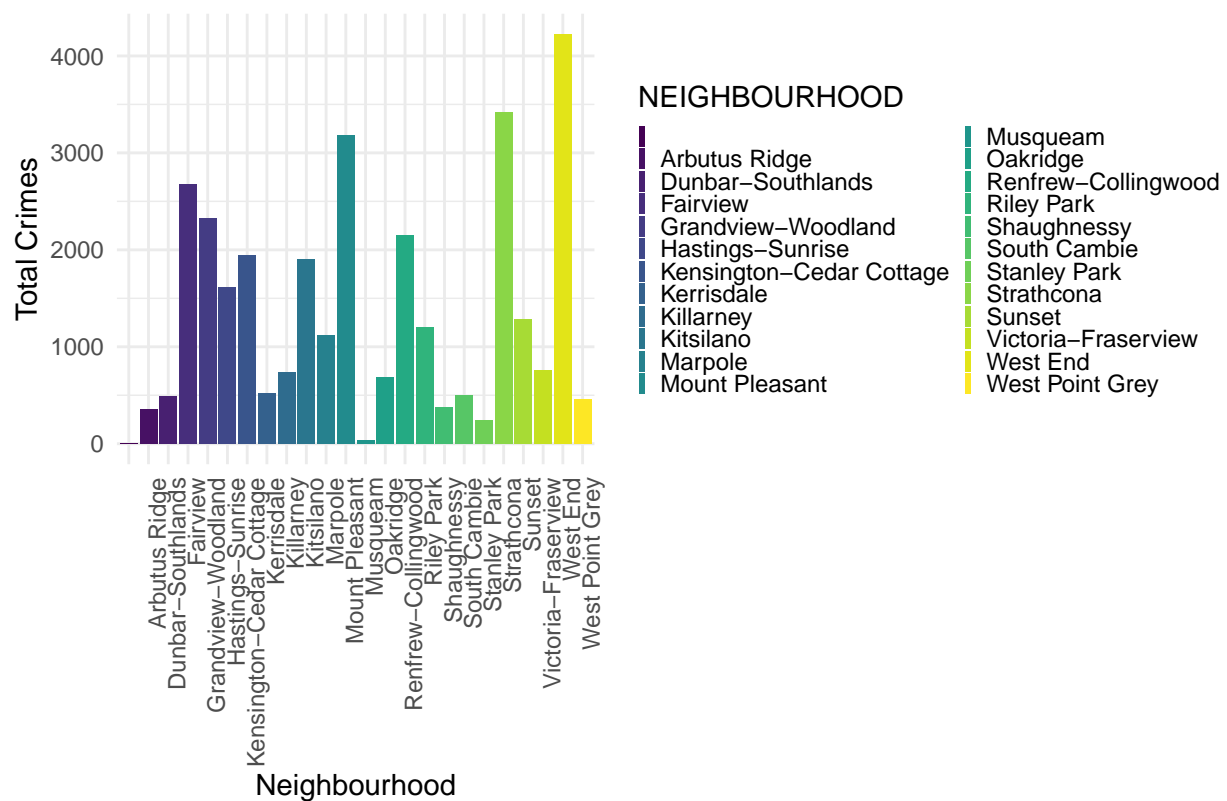
Total Crimes by Neighbourhood in 2019



```
# Filter and aggregate data for 2019 excluding Central Business District.
crime_data_2019_total <- crime_data %>%
  filter(YEAR == 2019, NEIGHBOURHOOD != "Central Business District") %>%
  count(NEIGHBOURHOOD)

# Create a colorful bar chart for 2019
ggplot(crime_data_2019_total, aes(x = NEIGHBOURHOOD, y = n, fill = NEIGHBOURHOOD)) +
  geom_bar(stat = "identity") +
  labs(title = "Total Crimes by Neighbourhood in 2019", x = "Neighbourhood", y = "Total Crimes") +
  theme_minimal() +
  theme(legend.key.size = unit(0.1, "cm"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  scale_fill_viridis_d()
```

Total Crimes by Neighbourhood in 2019



```
# Filter and aggregate data for 2023 excluding Central Business District.
crime_data_2023_total <- crime_data %>%
  filter(YEAR == 2023, NEIGHBOURHOOD != "Central Business District") %>%
  count(NEIGHBOURHOOD)

# Create a colorful bar chart for 2023
ggplot(crime_data_2023_total, aes(x = NEIGHBOURHOOD, y = n, fill = NEIGHBOURHOOD)) +
  geom_bar(stat = "identity") +
  labs(title = "Crime by Area 23", x = "Neighbourhood", y = "Total Crimes Committed") +
  theme_minimal() +
  theme(legend.key.size = unit(0.1, "cm"),
        axis.text.x = element_text(angle = 90, hjust = 1)) +
  scale_fill_viridis_d()
```



# Crime by Area 23

