COVID-19 Canada Analysis

Liem Vu

2025-07-14

```
covid <- read_csv("covid_data.csv", show_col_types = FALSE)

canada <- covid %>%
  filter(country == "Canada", code == "CAN", continent == "North America") %>%
  select(date, total_cases, new_cases, total_deaths, new_deaths,
         people_vaccinated, people_fully_vaccinated, population)
glimpse(canada)
```

```
## Rows: 2,027
## Columns: 8
## $ date                   <date> 2020-01-01, 2020-01-02, 2020-01-03, 2020-01-0~
## $ total_cases            <dbl> NA, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ new_cases              <dbl> NA, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ total_deaths           <dbl> NA, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ new_deaths             <dbl> NA, NA, NA, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ people_vaccinated      <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ people_fully_vaccinated <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ population             <dbl> 38821209, 38821209, 38821209, 38821209, 388212~
```

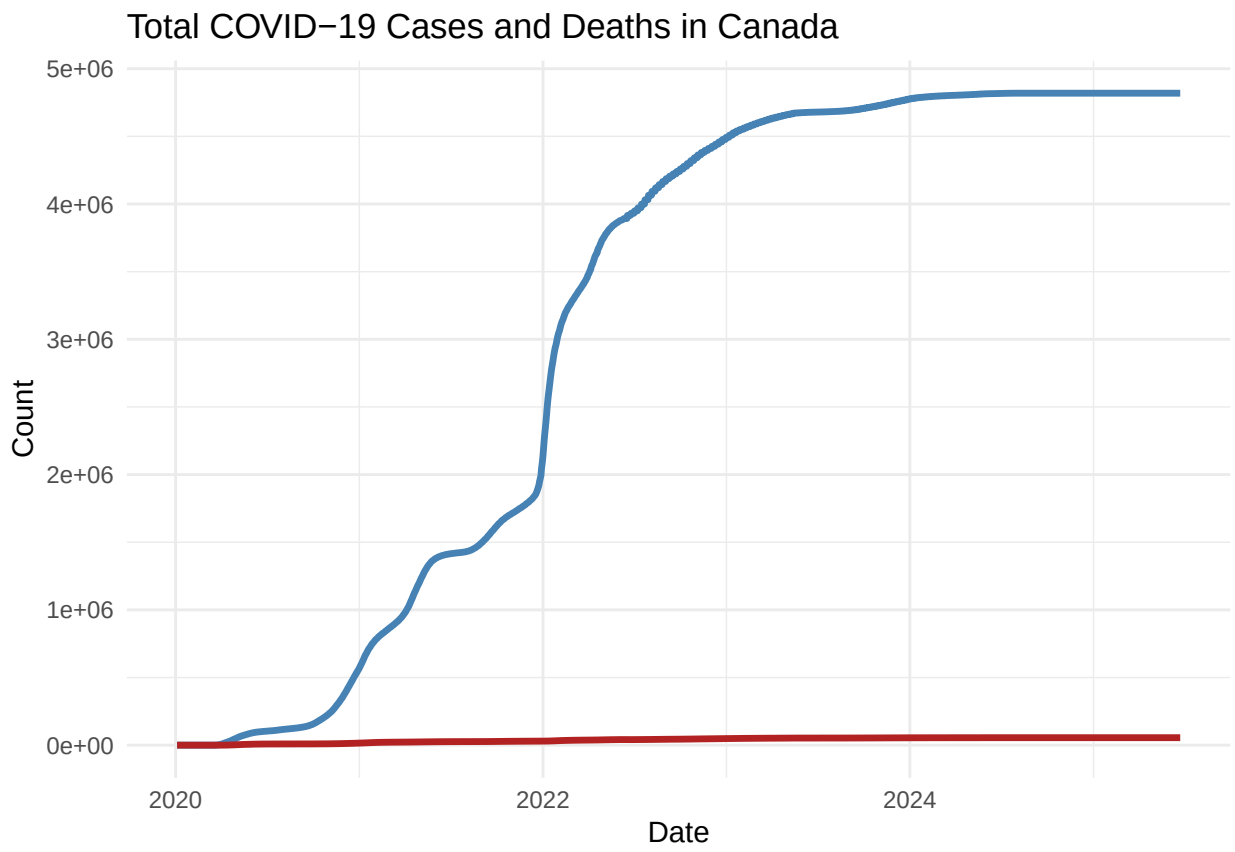
```
summary(canada)
```

```
##       date             total_cases        new_cases      total_deaths
##  Min.   :2020-01-01   Min.   :      0   Min.   :    0   Min.   :    0
##  1st Qu.:2021-05-21   1st Qu.:1324122   1st Qu.:    0   1st Qu.:24672
##  Median :2022-10-10   Median :4240286   Median :    0   Median :44333
##  Mean   :2022-10-11   Mean   :3119844   Mean   : 2414   Mean   :37667
##  3rd Qu.:2024-02-28   3rd Qu.:4793576   3rd Qu.: 2821   3rd Qu.:54801
##  Max.   :2025-12-28   Max.   :4819055   Max.   :49148   Max.   :55282
##                       NA's   :30        NA's   :31      NA's   :30
##    new_deaths     people_vaccinated  people_fully_vaccinated   population
##  Min.   :  0.00   Min.   :       5   Min.   :       0        Min.   :38821209
##  1st Qu.:  0.00   1st Qu.:23648904   1st Qu.: 7184406        1st Qu.:38821209
##  Median :  0.00   Median :30379510   Median :29390928        Median :38821209
##  Mean   : 27.86   Mean   :25002636   Mean   :22021141        Mean   :38821209
##  3rd Qu.: 33.00   3rd Qu.:32780732   3rd Qu.:31464510        3rd Qu.:38821209
##  Max.   :382.00   Max.   :34742936   Max.   :31758252        Max.   :38821209
##  NA's   :31       NA's   :1326       NA's   :1274
```

1. Cases and Deaths

```r
canada_clean <- canada %>%
  mutate(date = ymd(date)) %>%
  drop_na(total_cases, total_deaths)

ggplot(canada_clean, aes(x = date)) +
  geom_line(aes(y = total_cases), color = "steelblue", linewidth = 1.2) +
  geom_line(aes(y = total_deaths), color = "firebrick", linewidth = 1.2) +
  labs(title = "Total COVID-19 Cases and Deaths in Canada",
       x = "Date", y = "Count") +
  theme_minimal()
```

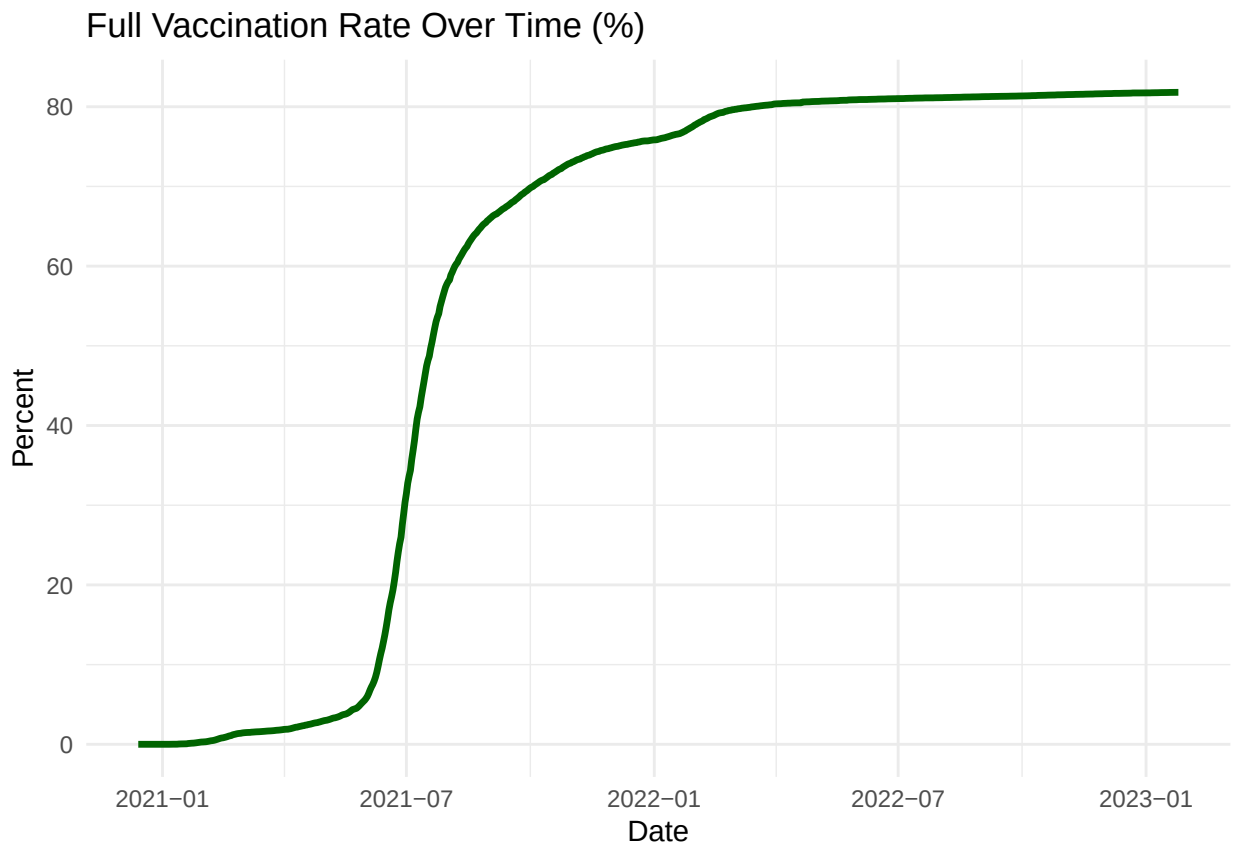


Interpretation: This line chart shows the growth of cumulative cases and deaths in Canada. Deaths increase more slowly than cases, likely due to vaccination and better treatment.

2. Vaccination Rate

```r
vaccine_data <- canada %>%
  mutate(date = ymd(date)) %>%
  drop_na(people_fully_vaccinated, population) %>%
  mutate(vaccine_rate = people_fully_vaccinated / population * 100)

ggplot(vaccine_data, aes(x = date, y = vaccine_rate)) +
  geom_line(color = "darkgreen", linewidth = 1.2) +
  labs(title = "Full Vaccination Rate Over Time (%)",
       x = "Date", y = "Percent") +
  theme_minimal()
```

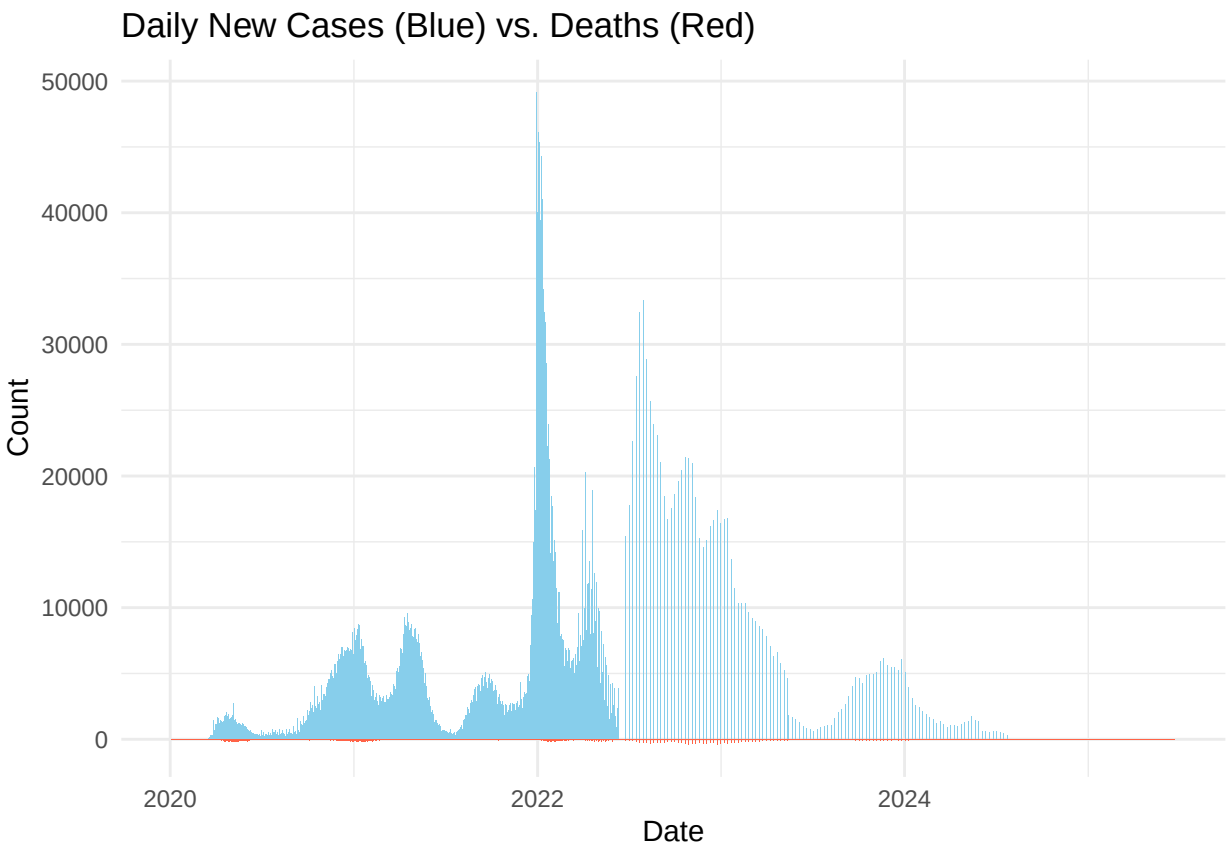


Interpretation: The percentage of fully vaccinated people rose sharply from early 2021, reaching over 70% in the end of the year.

3. Daily new cases and new deaths

```
daily <- canada %>%
  mutate(date = ymd(date)) %>%
  drop_na(new_cases, new_deaths)

ggplot(daily, aes(x = date)) +
  geom_col(aes(y = new_cases), fill = "skyblue") +
  geom_col(aes(y = -new_deaths), fill = "tomato") +
  labs(title = "Daily New Cases (Blue) vs. Deaths (Red)",
       x = "Date", y = "Count") +
  theme_minimal()
```



Interpretation: The gap between new cases and new deaths widened after vaccines were rolled out. Deaths lagged cases by 2–3 weeks.