COVID-19 Epidemic Information

Nick Lauerman

Contents

Set up

R Libraries

R is extended using packages or libraries. For this analysis the following packages are used.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------------------------------
```

```
## v ggplot2 3.3.0     v purrr   0.3.3
## v tibble  3.0.0     v dplyr   0.8.5
## v tidyr   1.0.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------------------------------- tidyve
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:dplyr':
##
##     intersect, setdiff, union
```

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

Note: the tidyverse package loads a set of packages that make up the tidy data universe. Use of the tidyverse package to load these packages is a convince instead of loading each individual backage.

Get data

Two R scripts are used to read in the data used. The source of this data is the WHO, for world wide data, and the CDC for US data. While the WHO data contains data for the US the CDC data is being used. The WHO data is a series of files scraped from their daily situtation reports and the CDC data is scraped from their web site for the CORANA virus. Due tothe WHO data being in 85 plus files the execution of these scripts is not shown to insure that the files is not overly large.

Create data set for Germany

The data for Germany is extracted from the WHO data for easy in displaying

```r
germany <- WHO %>%
      filter(`Country Territory area` == "Germany")
germany$`Week Number` <- week(germany$date)
```

Create data set for world wide data

This data set displays the daily total for world wide data.

```r
world <- WHO %>%
      group_by(`date`) %>%
      summarise(`Number of new cases` = sum(`Confirmed new cases`),
                cum = sum(`total Confirmed Cases`))
world$`Week Number` <- week(world$date)
```

Charts and graphs

World Wide

```r
epi <- ggplot(data = world)
```

```r
epi + geom_point(aes(x = date,
               y = `Number of new cases`))+
      geom_smooth(aes(x = date,
               y = `Number of new cases`),
               method = "loess",
               formula = y ~ x,
               color = "red",
               fill = "blue") +
      labs(y = "Cases",
           title = "Number of New Cases of COVID-19 Reported to the WHO")
```

```r
epi + geom_col(aes(x=date,
                 y=`Number of new cases`),
                fill= "blue") +
      geom_smooth(aes(x=date,
                 y=`Number of new cases`),
               method = "loess",
               formula = y ~ x,
                 color = "red") +
```
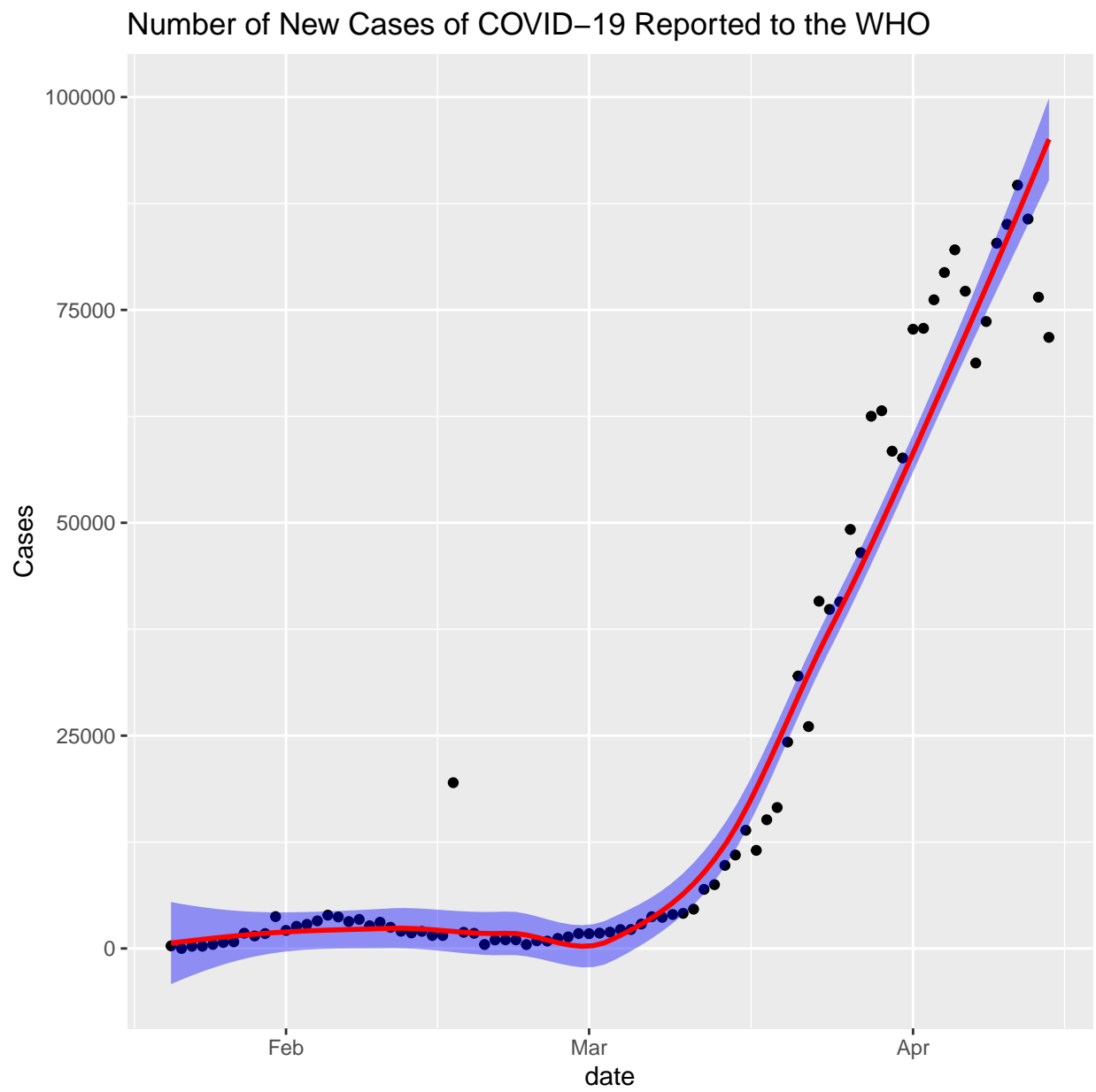
Figure 1: Epi curve 1

```
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported to the WHO")
```
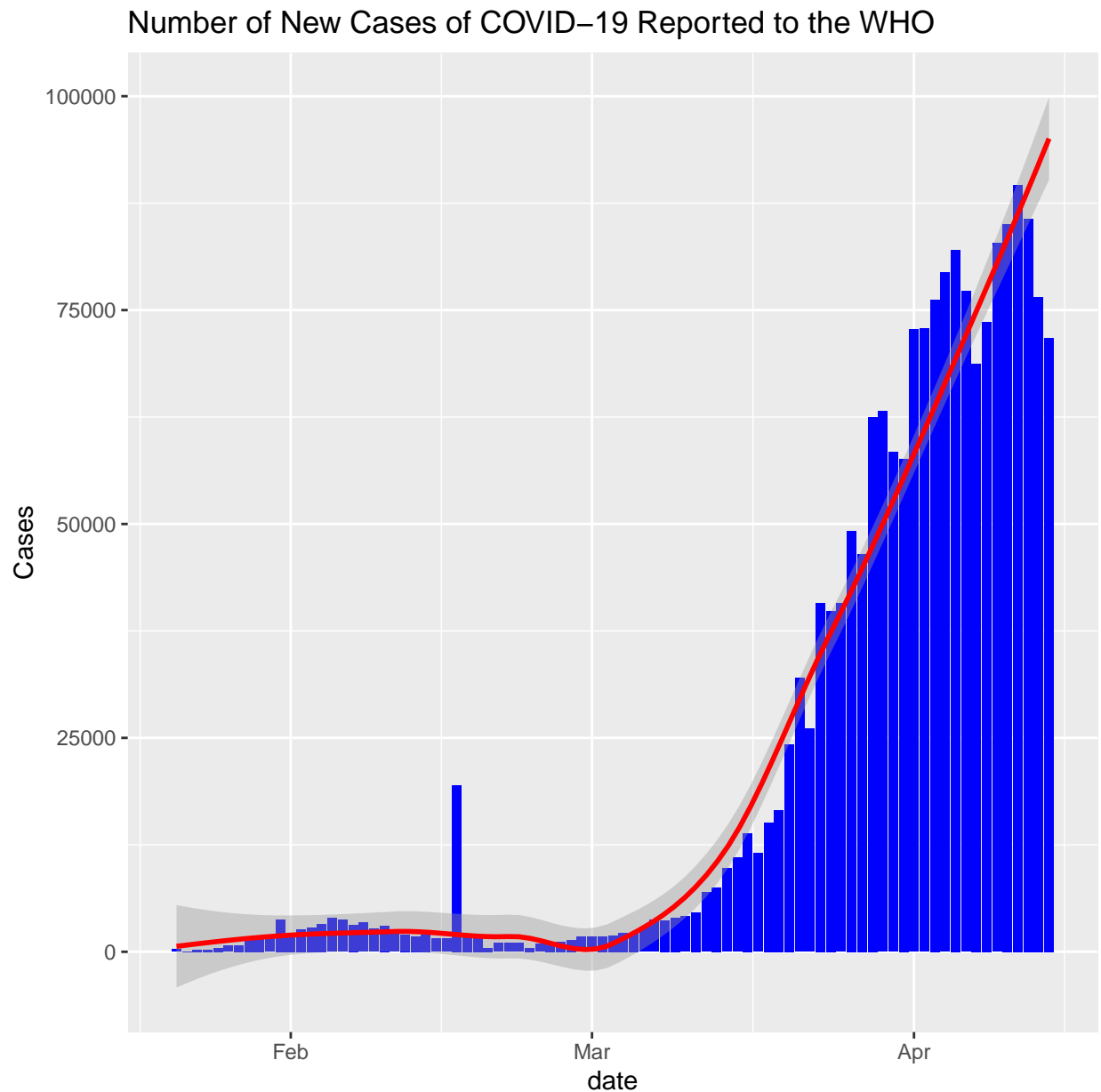
Number of New Cases of COVID−19 Reported to the WHO

Figure 2: Epi curve 2, traditional

```
epi + geom_point(aes(x = date,
           y = cum))+
    geom_smooth(aes(x = date,
           y = `cum`),
           method = "loess",
           formula = y ~ x,
           color = "red",
           fill = "blue") +
```

```
      labs(y = "Cumulative number of cases",
           title = "Cumulative Number of Cases of COVID-19 Reported to the WHO")
```


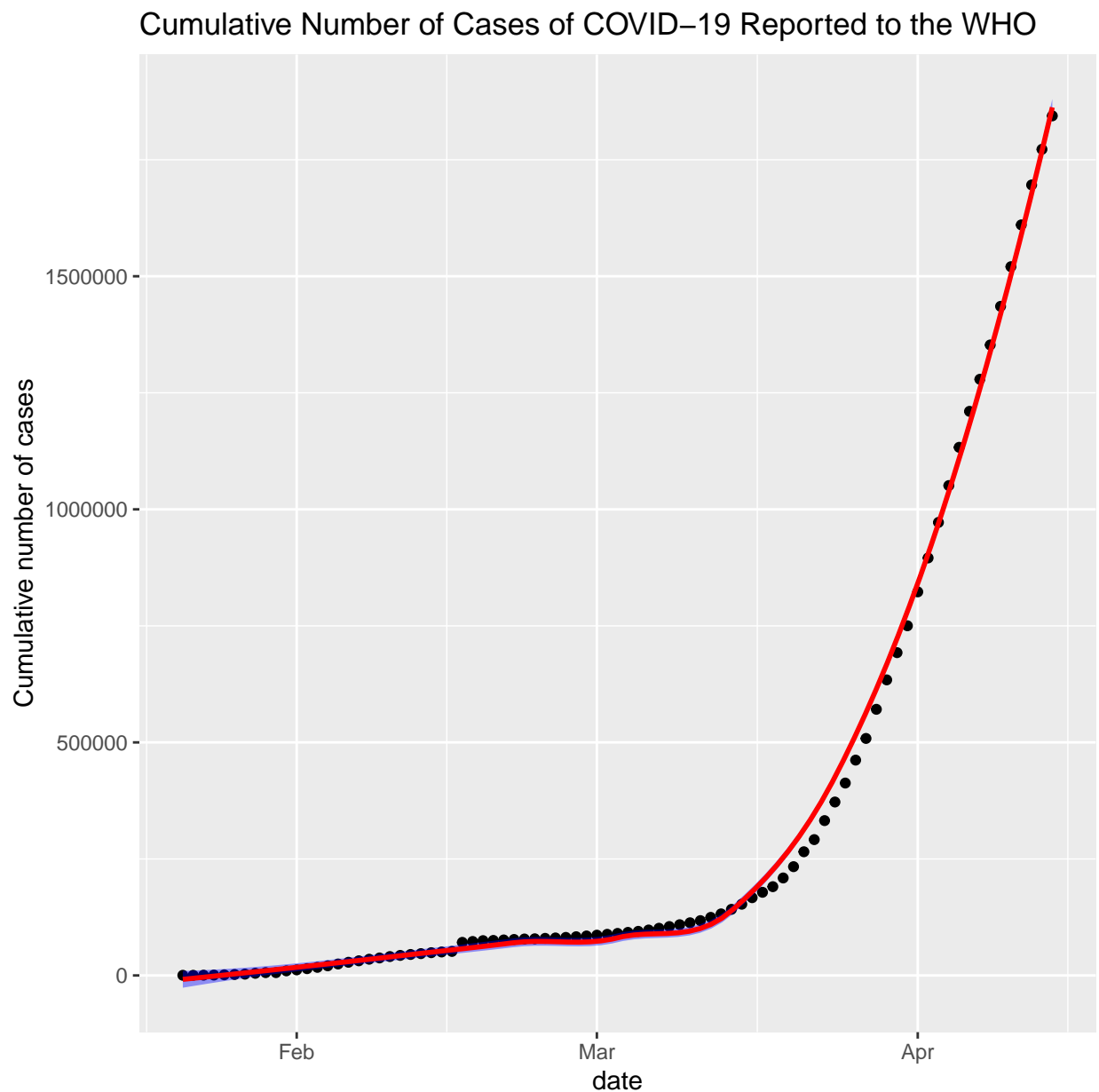Figure 3: Cumulative cases

```
gr <- world %>%
  filter(`Number of new cases` != 0)

ggplot(data = gr) +
  geom_line(aes(y = `Number of new cases`,
                x = cum)) +
  scale_x_log10() +
  scale_y_log10() +
```

```
geom_smooth(aes(y = `Number of new cases`,
                x = cum),
            method = "loess",
            formula = y ~ x) +
labs(title = "Daily Growth Rate",
     x = "Total Cases (log)",
     y = "New Cases (log)")
```

Daily Growth Rate
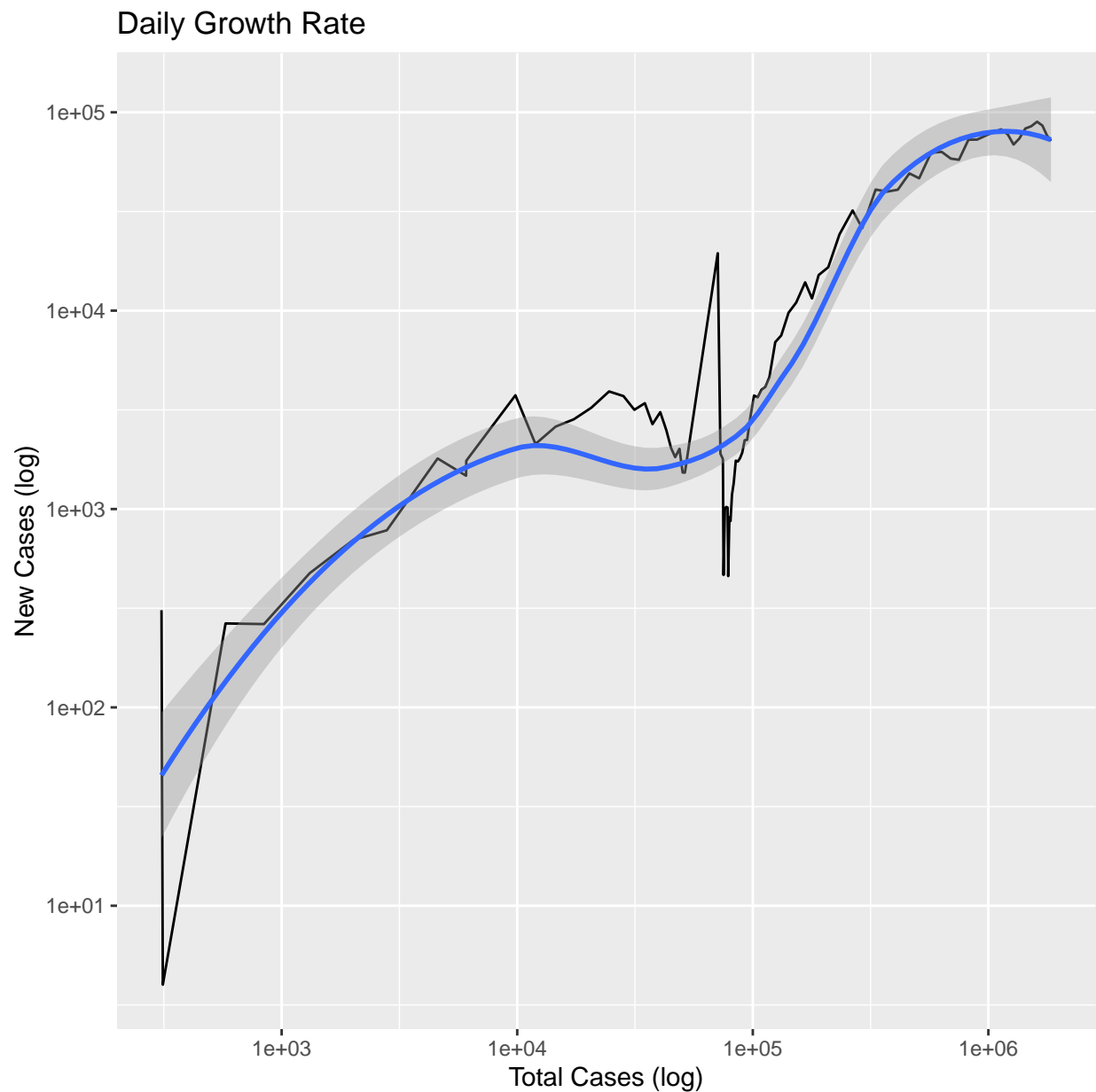

Figure 4: Growth Rate

```
ggplot(data = gr) +
    geom_line(aes(y = `cum`,
                  x = `date`)) +
```

```
    geom_point(aes(y = `cum`,
                   x = `date`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cum`,
                    x = `date`),
            method = "loess",
            formula = y ~ x) +
labs(title = "Growth Rate",
     x = "date",
     y = "Total Cases (log)")
```
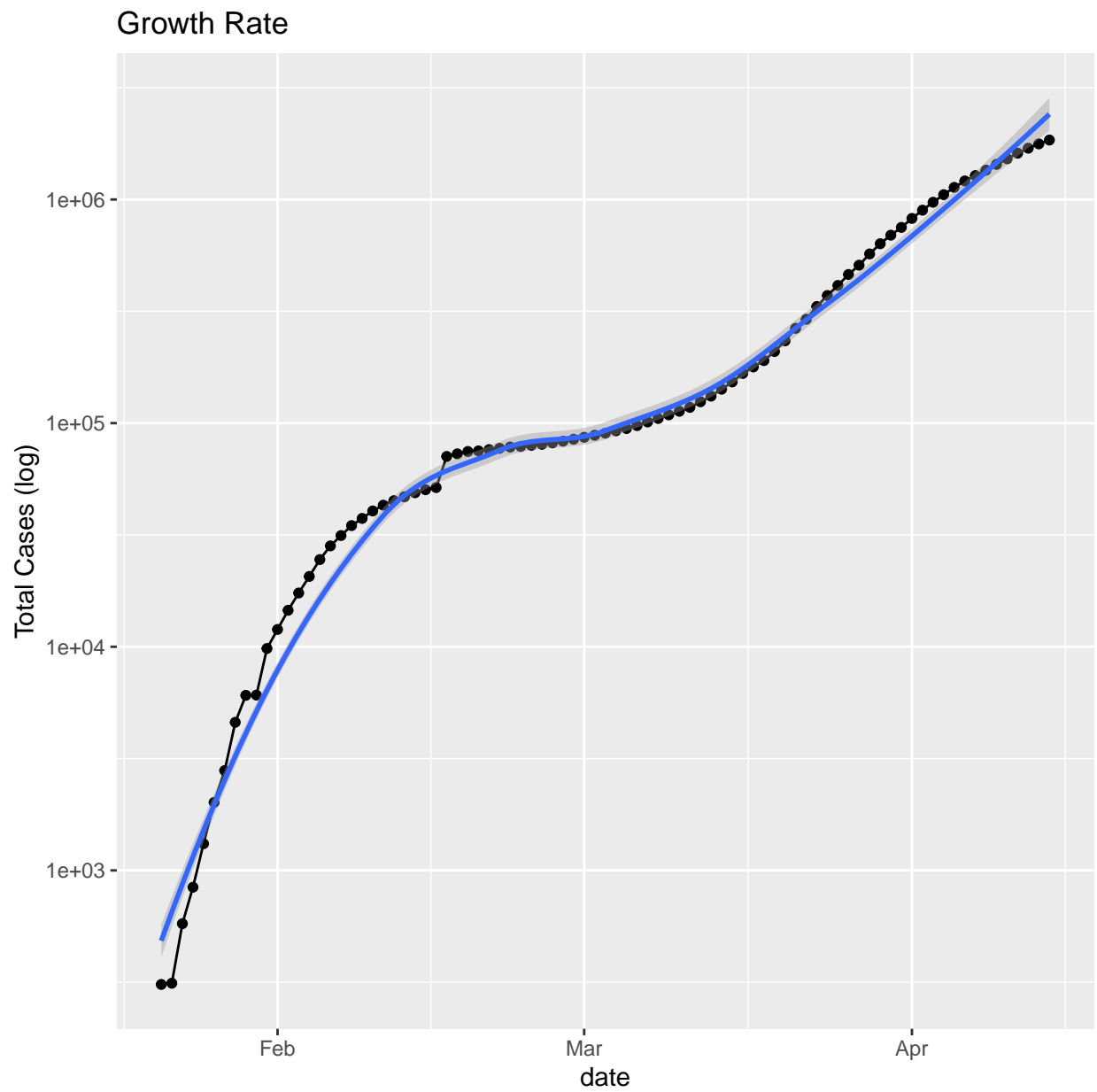
Growth Rate

Figure 5: Another look at growth

7

```
gr <- world %>%
    group_by(`Week Number`) %>%
    summarise(cases = sum(`Number of new cases`),
              cs = sum(cum),
              count = n())

ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = cs)) +
    geom_point(aes(y = `cases`,
                   x = cs)) +
    scale_x_log10() +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = cs),
             method = "loess",
             formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```
ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = `Week Number`)) +
    geom_point(aes(y = `cases`,
                   x = `Week Number`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = `Week Number`),
             method = "loess",
             formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Week of the year",
       y = "Cases (log)")
```

Germany

```
epi <- ggplot(data = germany)
```

```
epi + geom_point(aes(x = date,
             y = `Confirmed new cases`))+
    geom_smooth(aes(x = date,
             y = `Confirmed new cases`),
             method = "loess",
             formula = y ~ x,
             color = "red",
             fill = "blue") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19")
```

```
epi + geom_col(aes(x=date,
                   y=`Confirmed new cases`),
               fill= "blue") +
    geom_smooth(aes(x=date,
```
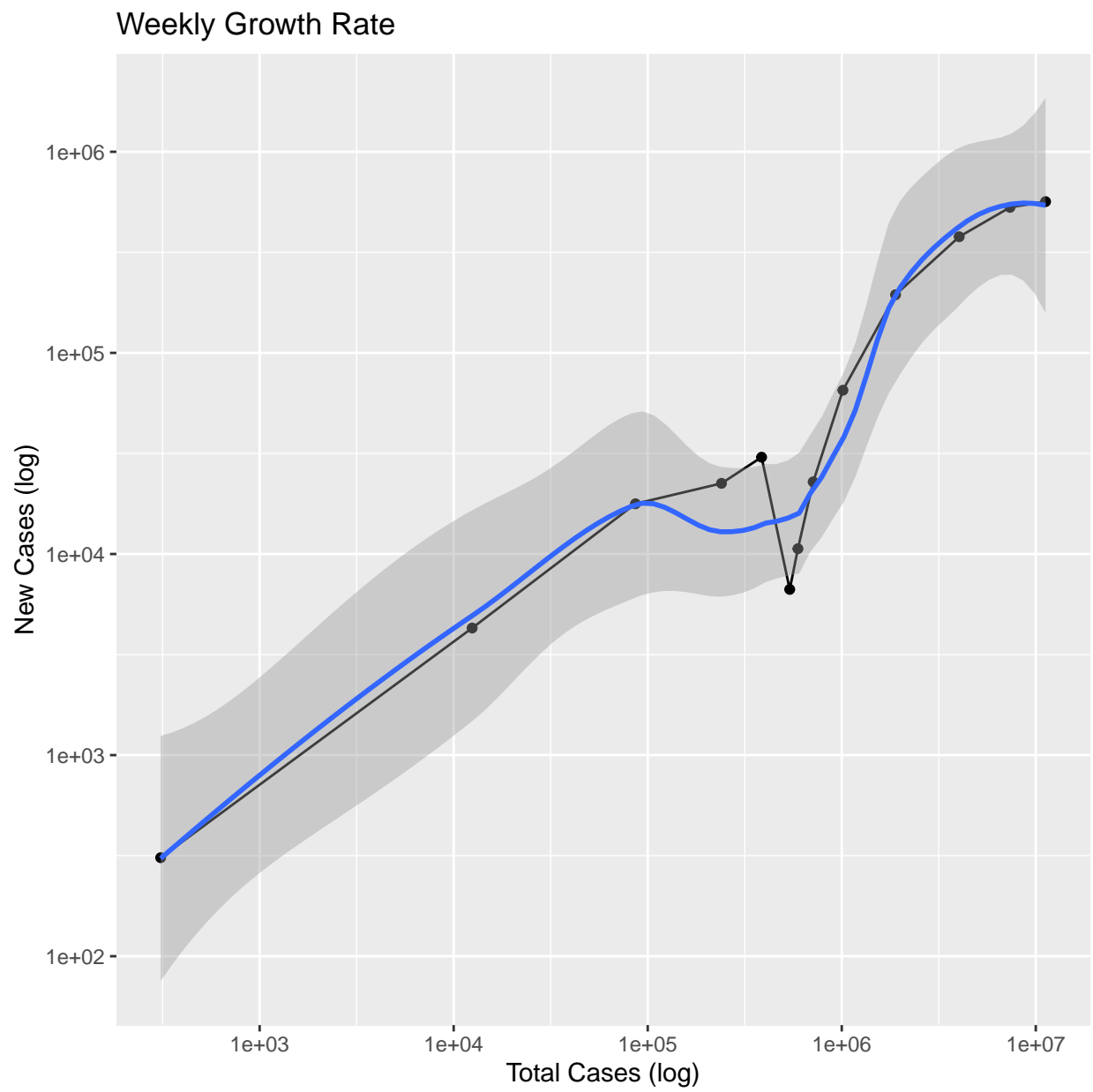
Figure 6: weekly growth rate

Figure 7: Another look at growth

Figure 8: Epi curve 1
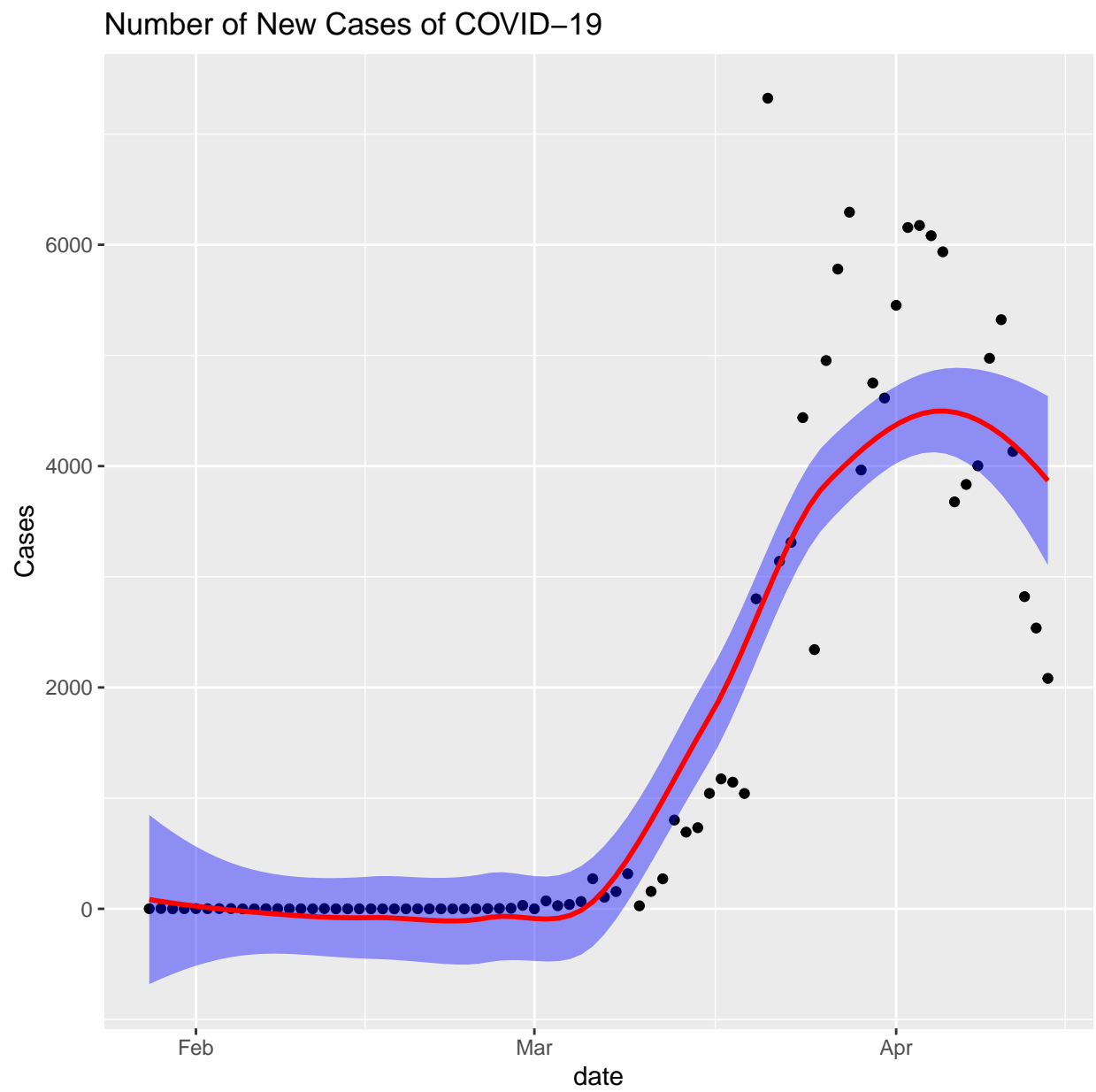
```
                    y=`Confirmed new cases`),
              method = "loess",
              formula = y ~ x,
                 color = "red") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported")
```
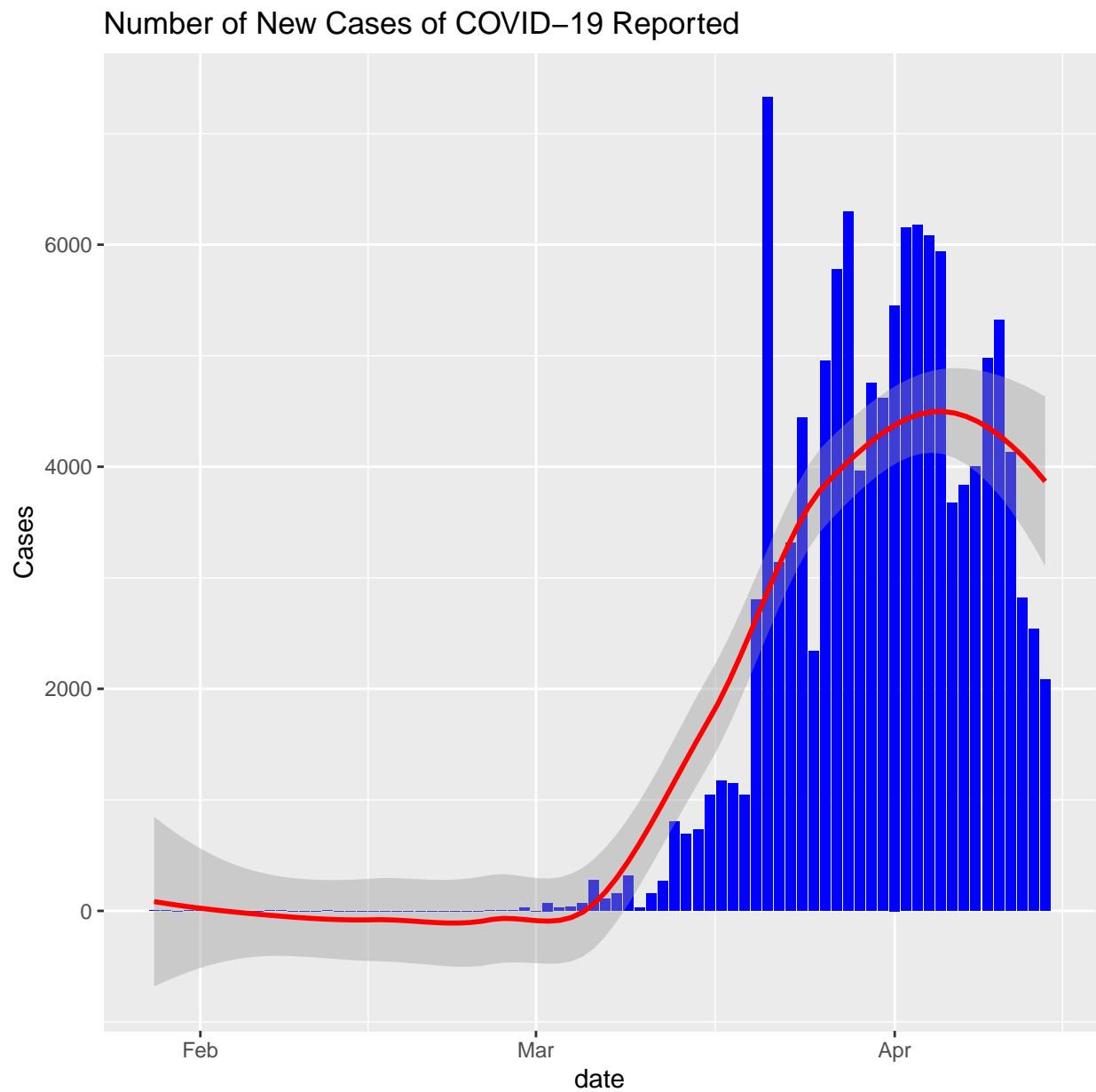
Number of New Cases of COVID−19 Reported

Figure 9: Epi curve 2, traditional

```
epi + geom_point(aes(x = date,
             y = `total Confirmed Cases`))+
     geom_smooth(aes(x = date,
             y = `total Confirmed Cases`),
```

```
            method = "loess",
            formula = y ~ x,
            color = "red",
            fill = "blue") +
      labs(y = "Cumulative number of cases",
           title = "Cumulative Number of Cases of COVID-19 Reported")
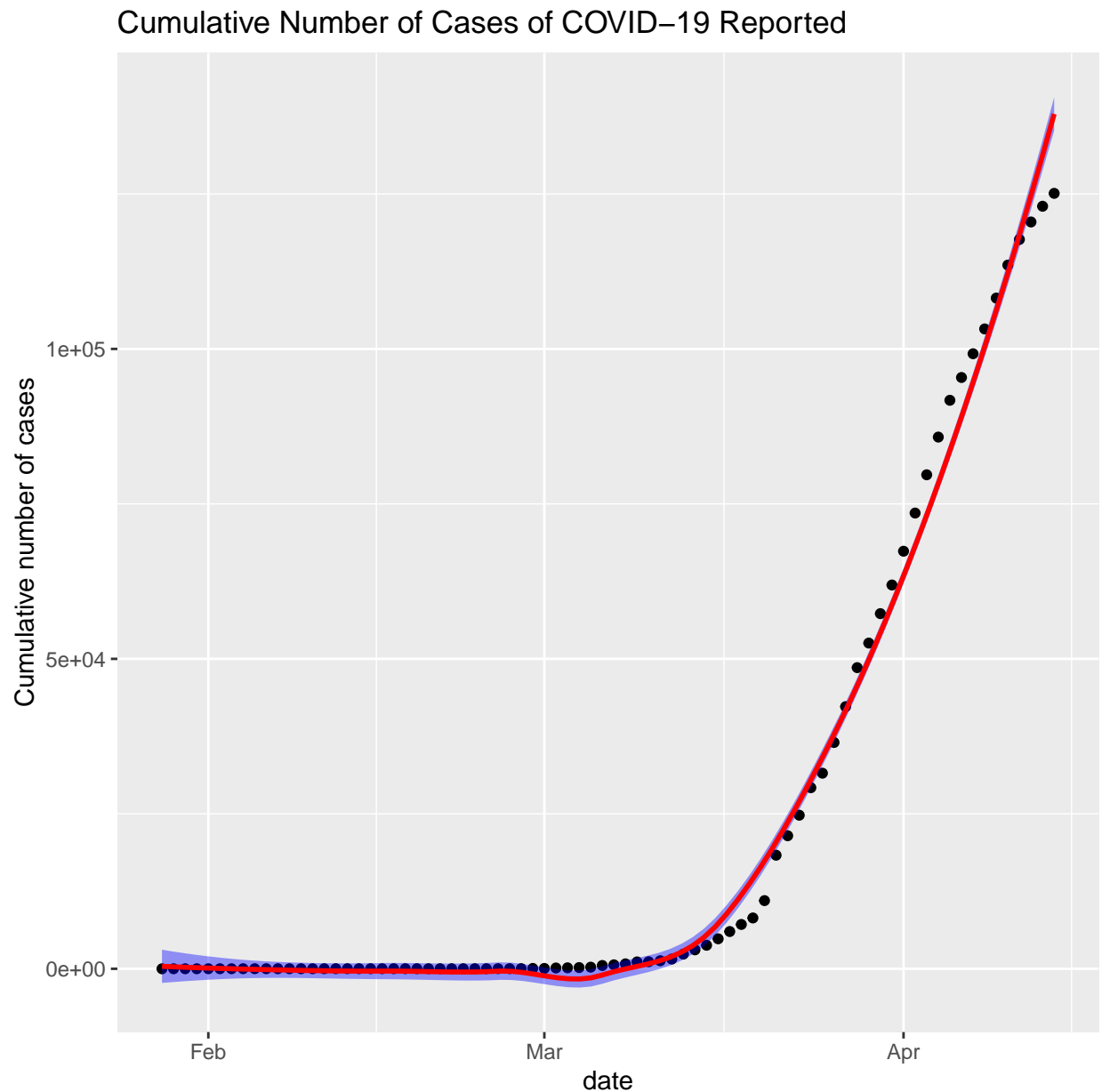```



Figure 10: Cumulative cases

```
gr <- germany %>%
  filter(`Confirmed new cases` != 0)

ggplot(data = gr) +
```

```
geom_line(aes(y = `Confirmed new cases`,
               x = `total Confirmed Cases`)) +
scale_x_log10() +
scale_y_log10() +
geom_smooth(aes(y = `Confirmed new cases`,
                x = `total Confirmed Cases`),
            method = "loess",
            formula = y ~ x) +
labs(title = "Daily Growth Rate",
     x = "Total Cases (log)",
     y = "New Cases (log)")
```


Figure 11: Growth Rate

14

```
ggplot(data = gr) +
  geom_line(aes(x = `date`,
                y = `total Confirmed Cases`)) +
  scale_y_log10() +
  geom_smooth(aes(x = `date`,
                  y = `total Confirmed Cases`),
              method = "loess",
              formula = y ~ x) +
  labs(title = "Daily Growth Rate",
       x = "Date",
       y = "Total Cases (log)")
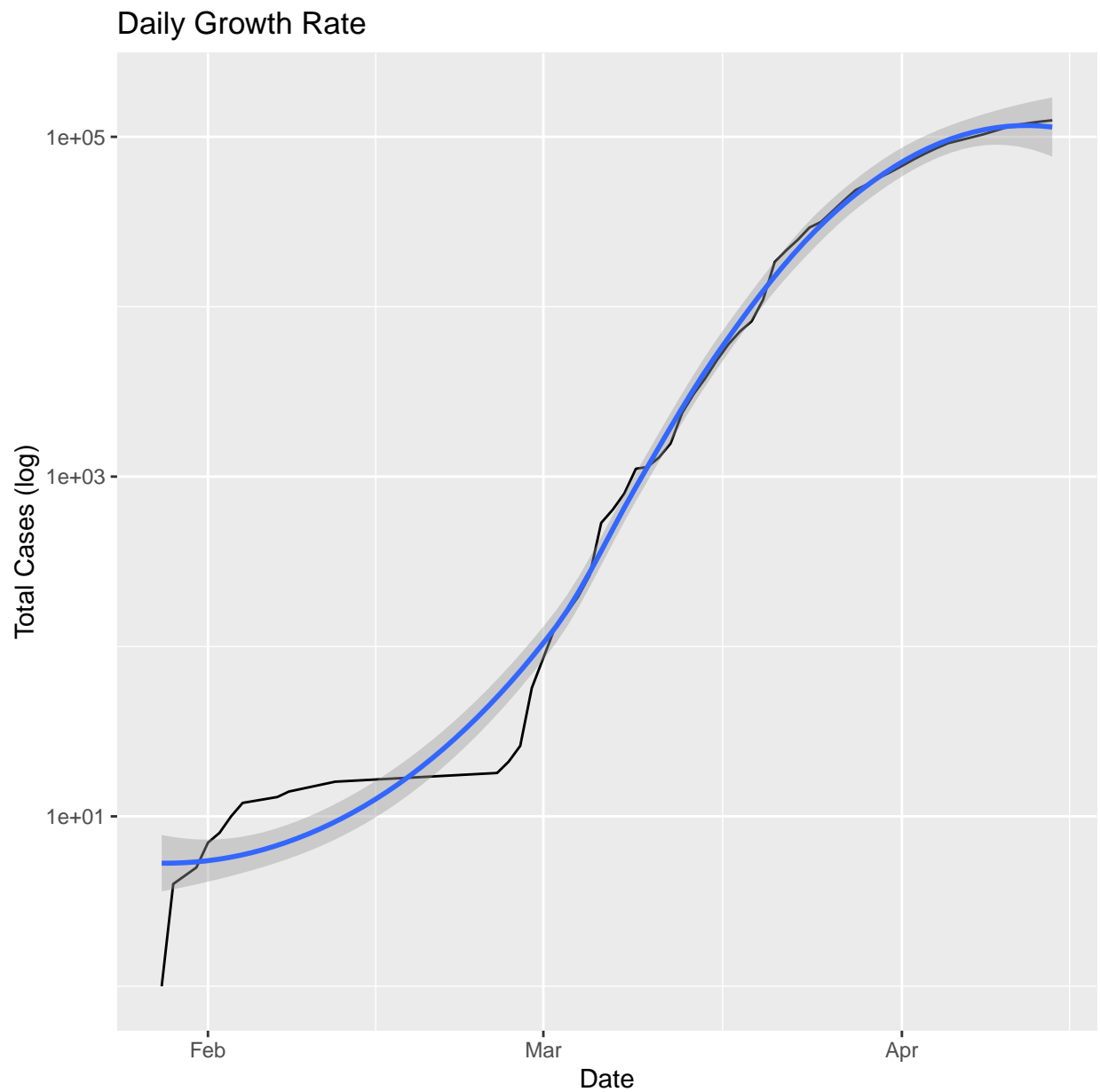```

Figure 12: Another look at growth

```
gr <- germany %>%
    group_by(`Week Number`) %>%
    summarise(cases = sum(`Confirmed new cases`),
              cs = sum(`total Confirmed Cases`),
              count = n()) %>%
    filter(cases != 0)

ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = cs)) +
    geom_point(aes(y = `cases`,
                   x = cs)) +
    scale_x_log10() +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = cs),
                method = "loess",
                formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```
ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = `Week Number`)) +
    geom_point(aes(y = `cases`,
                   x = `Week Number`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = `Week Number`),
                method = "loess",
                formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Week of the year",
       y = "New Cases (log)")
```

US

Complete data set

CDC designates a date to which they consider the data "good". Reported data after that date is considered incomplete due to delays in reporting. This intial set of plots uses the complete data set

```
epi <- ggplot(data = cdc)
```

```
epi + geom_point(aes(x = Date,
              y = `Number of new cases`))+
    geom_smooth(aes(x = Date,
              y = `Number of new cases`),
              method = "loess",
              formula = y ~ x,
              color = "red",
              fill = "blue") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported to the CDC")
```
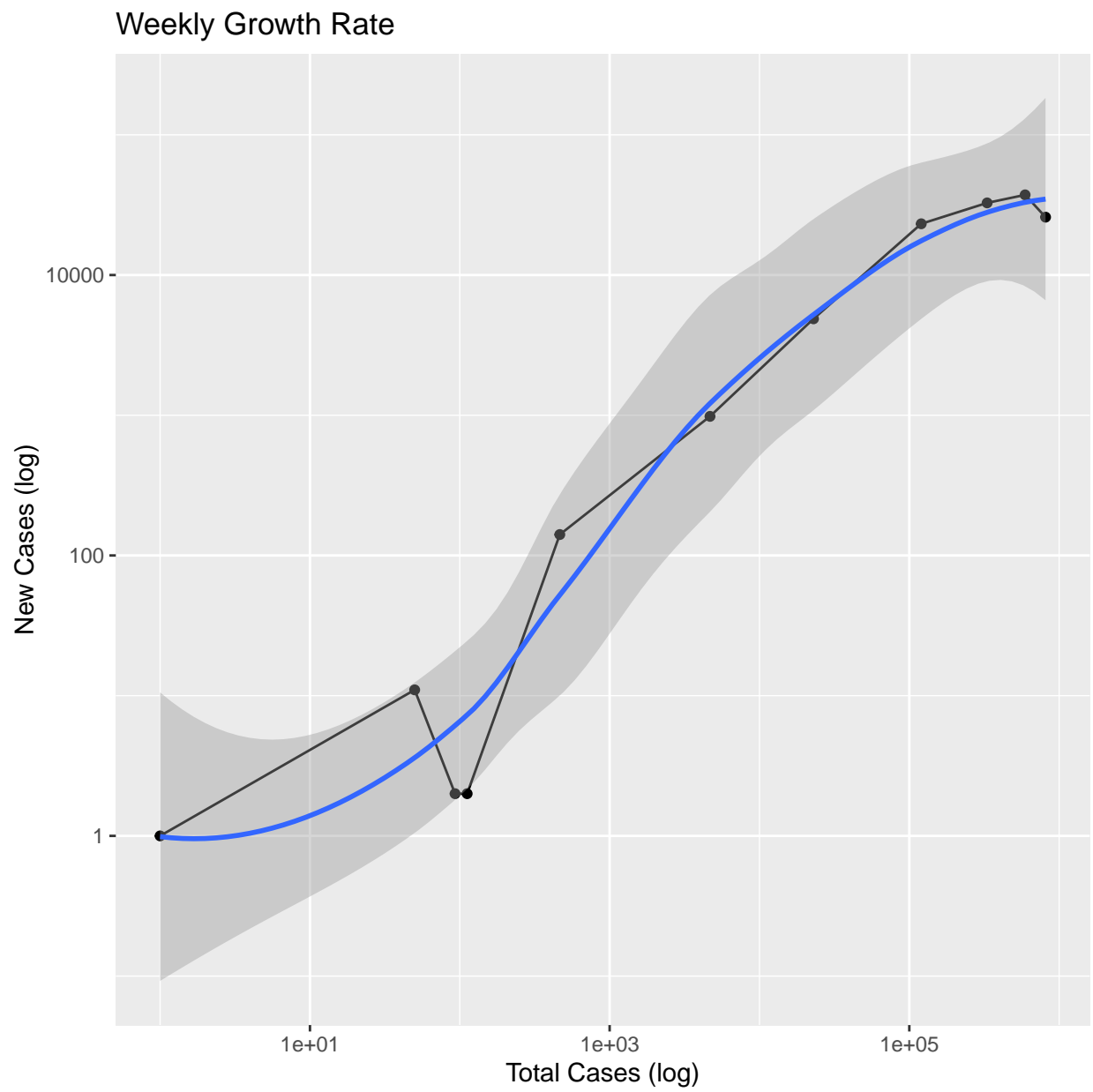
Figure 13: Weekly growth rate

Figure 14: Another look at growth
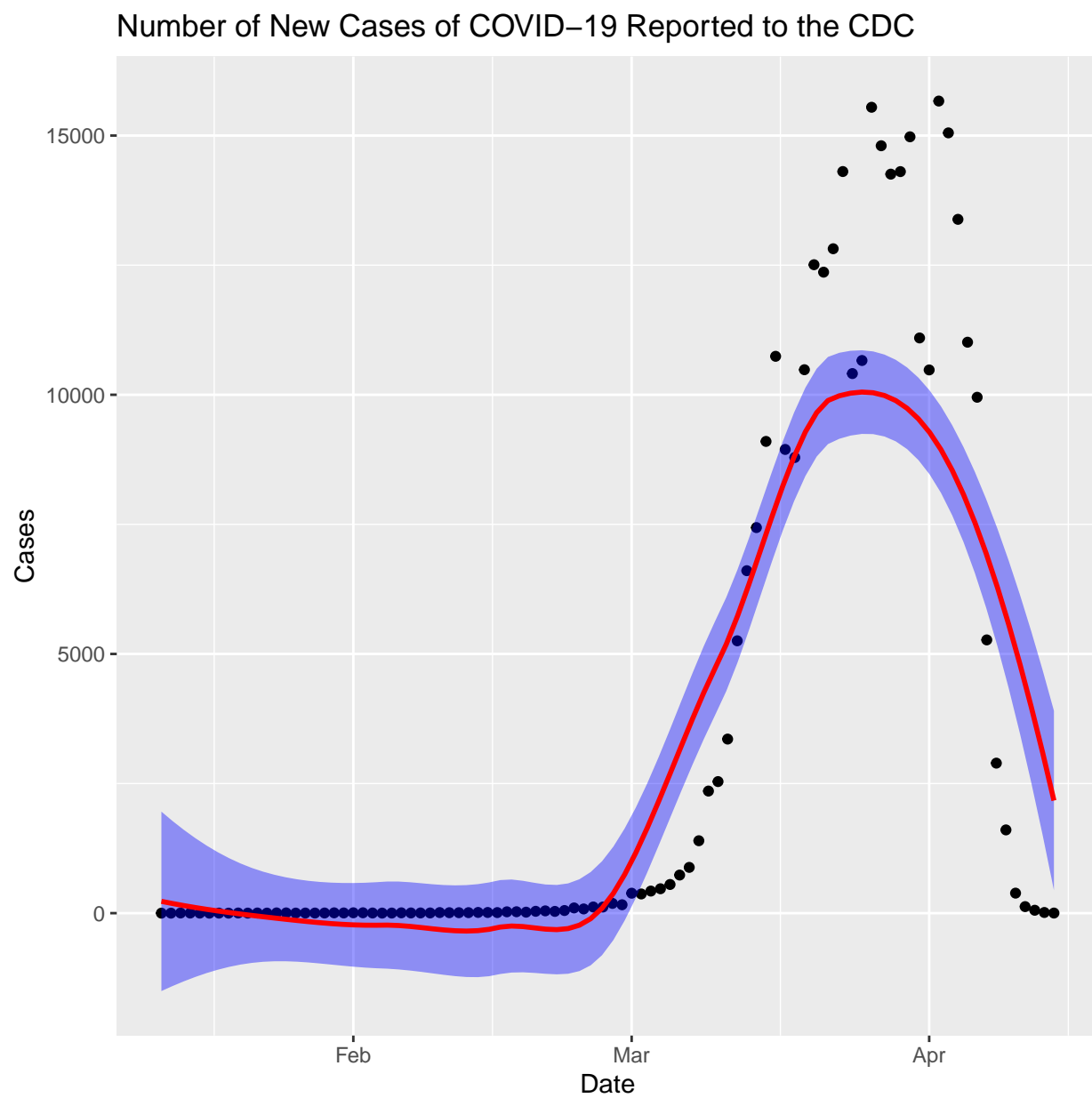
Figure 15: Epi curve 1

```
epi + geom_col(aes(x=Date,
                   y=`Number of new cases`),
               fill= "blue") +
    geom_smooth(aes(x=Date,
                    y=`Number of new cases`),
                method = "loess",
                formula = y ~ x,
                color = "red") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported to the CDC")
```
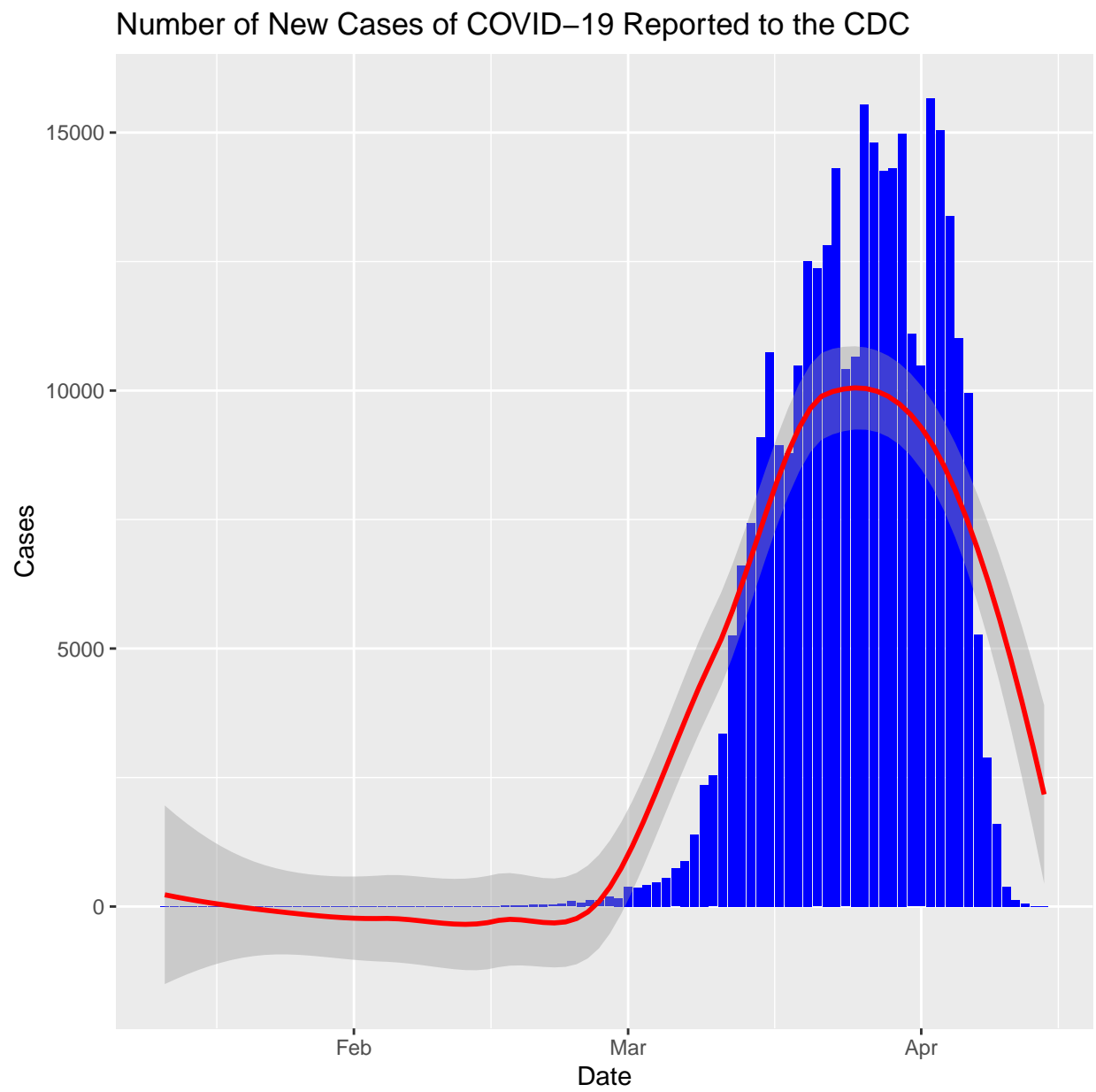
Figure 16: Epi curve 2, traditional

```r
epi + geom_point(aes(x = Date,
             y = cum))+
    # geom_line(aes(x = Date,
             #y = `cum`)) +
    geom_smooth(aes(x = Date,
             y = `cum`),
             method = "loess",
             formula = y ~ x,
             color = "red",
             fill = "blue") +
    labs(y = "Cumulative number of cases",
         title = "Cumulative Number of Cases of COVID-19 Reported to the CDC")
```

```r
gr <- cdc %>%
  filter(`Number of new cases` != 0)

ggplot(data = gr) +
  geom_line(aes(y = `Number of new cases`,
                x = cum)) +
  scale_x_log10() +
  scale_y_log10() +
  geom_smooth(aes(y = `Number of new cases`,
                x = cum),
             method = "loess",
             formula = y ~ x) +
  labs(title = "Daily Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```r
ggplot(data = gr) +
    geom_line(aes(y = `cum`,
                x = `Date`)) +
    geom_point(aes(y = `cum`,
                x = `Date`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cum`,
                x = `Date`),
             method = "loess",
             formula = y ~ x) +
  labs(title = "Growth Rate",
       x = "Date",
       y = "Total Cases (log)")
```

```r
gr <- cdc %>%
    group_by(`Week Number`) %>%
    summarise(cases = sum(`Number of new cases`),
              cs = sum(cum),
              count = n())

ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                x = cs)) +
    geom_point(aes(y = `cases`,
                x = cs)) +
```

Cumulative Number of Cases of COVID−19 Reported to the CDC
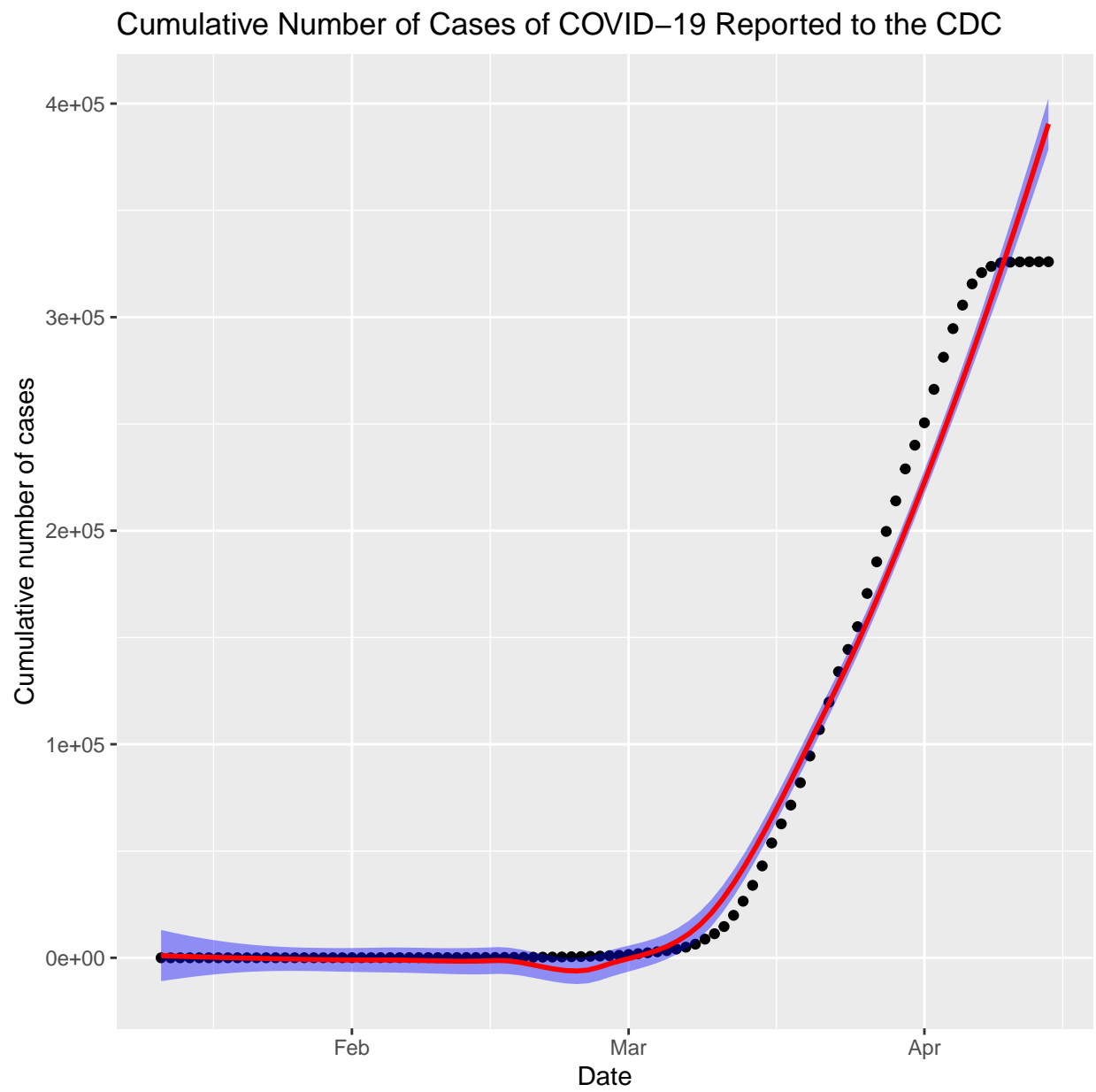
Figure 17: Cumulative cases

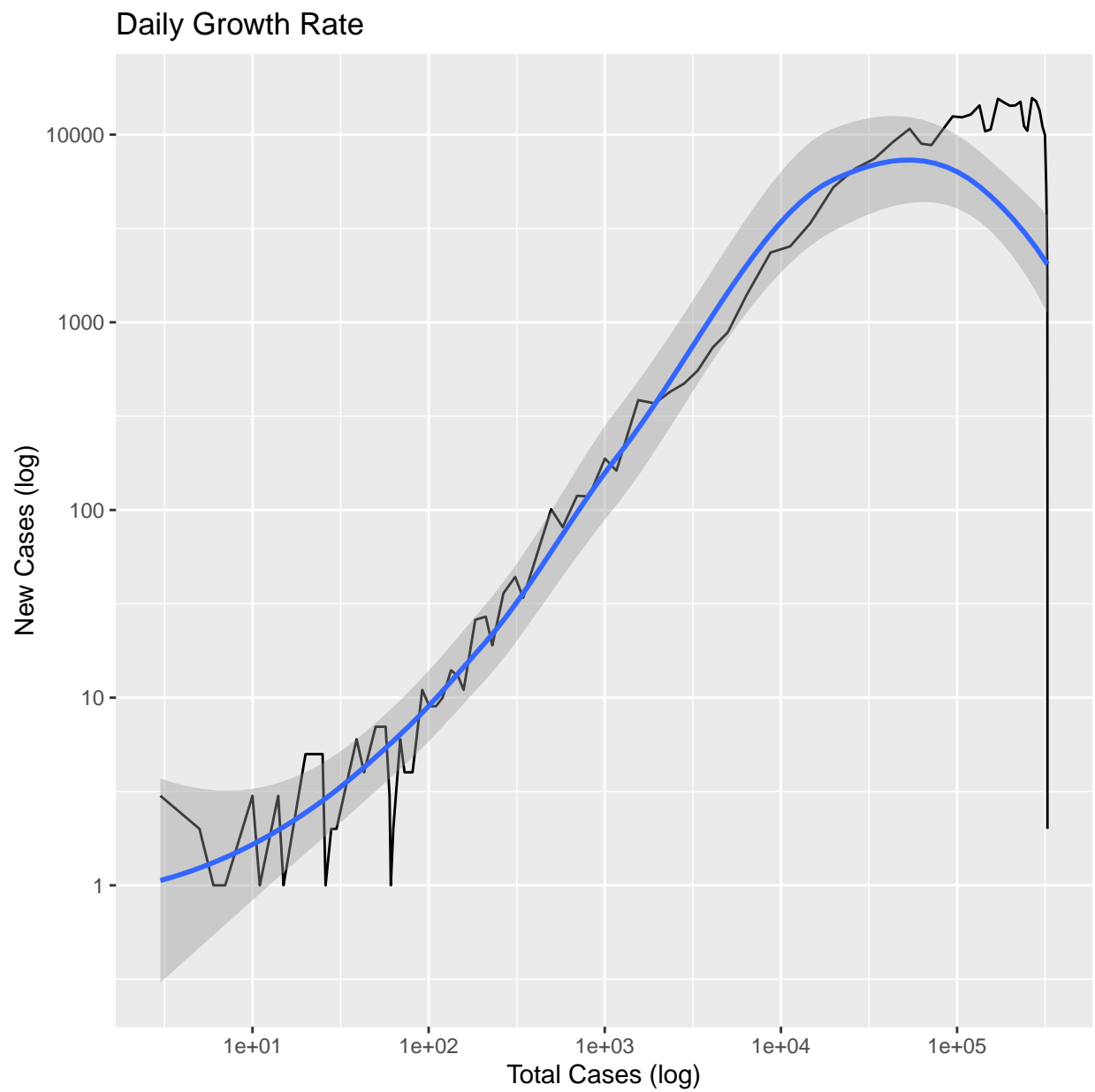Figure 18: Growth Rate
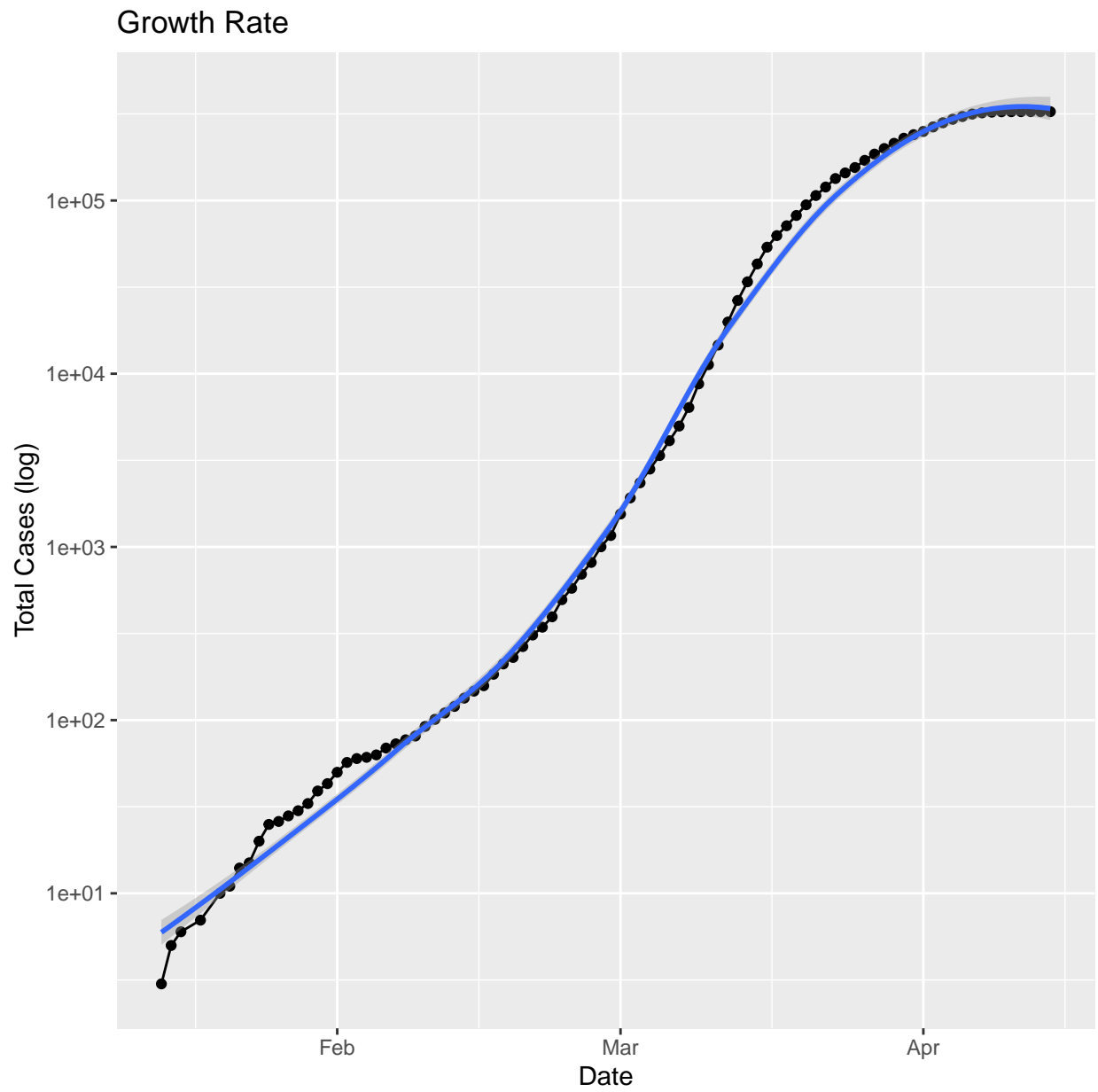
Figure 19: Another look at growth

```
    scale_x_log10() +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = cs),
              method = "loess",
              formula = y ~ x) +
labs(title = "Weekly Growth Rate",
     x = "Total Cases (log)",
     y = "New Cases (log)")
```
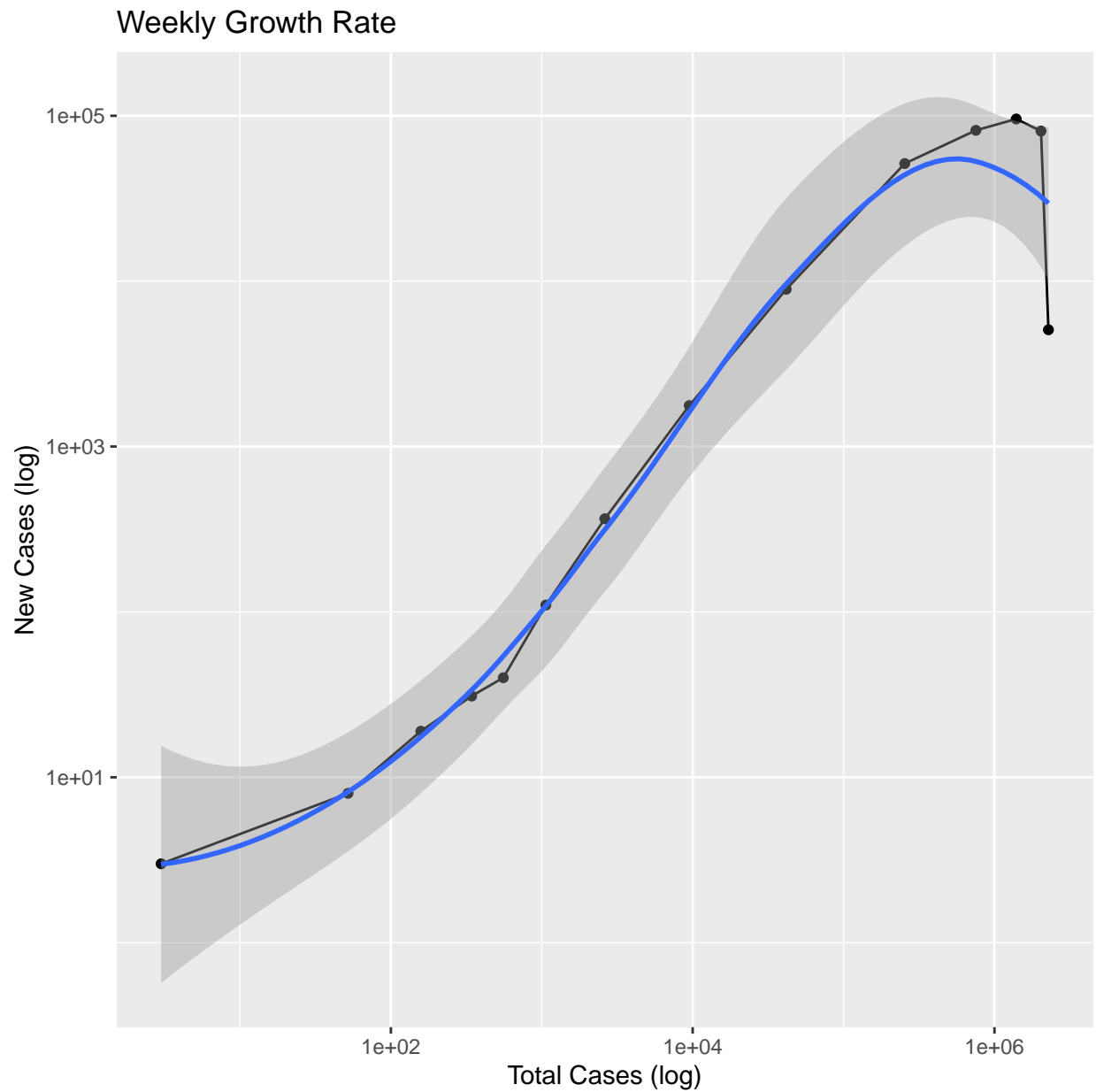
Weekly Growth Rate

Figure 20: weekly growth rate

```r
ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = `Week Number`)) +
    geom_point(aes(y = `cases`,
                   x = `Week Number`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = `Week Number`),
                method = "loess",
                formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Week of the year",
       y = "New Cases (log)")
```

Filtered to use only "complete data"

Remove dates on or after 4 April as this data may not be completely reported

```r
cdc <- cdc %>%
    filter(Date < as.Date("2020-04-04")) %>%
  filter(`Number of new cases` != 0)
```

```r
epi <- ggplot(data = cdc)
```

```r
epi + geom_point(aes(x = Date,
                 y = `Number of new cases`))+
    geom_smooth(aes(x = Date,
                y = `Number of new cases`),
                method = "loess",
                formula = y ~ x,
                color = "red",
                fill = "blue") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported to the CDC")
```

```r
epi + geom_col(aes(x=Date,
                   y=`Number of new cases`),
               fill= "blue") +
    geom_smooth(aes(x=Date,
                y=`Number of new cases`),
                method = "loess",
                formula = y ~ x,
                  color = "red") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported to the CDC")
```

```r
epi + geom_point(aes(x = Date,
                 y = cum))+
    # geom_line(aes(x = Date,
              #y = `cum`)) +
    geom_smooth(aes(x = Date,
                y = `cum`),
                method = "loess",
                formula = y ~ x,
                color = "red",
```
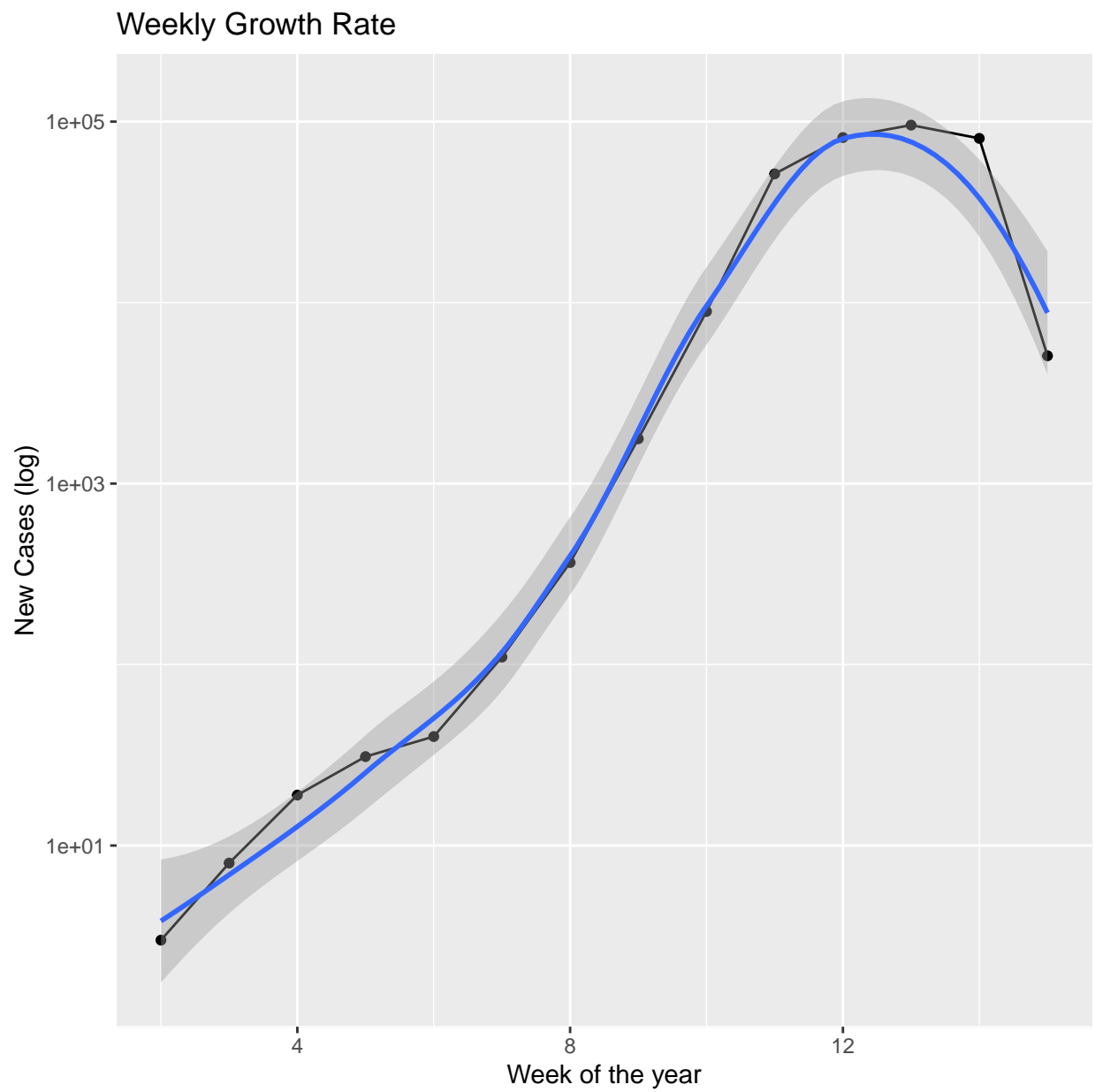
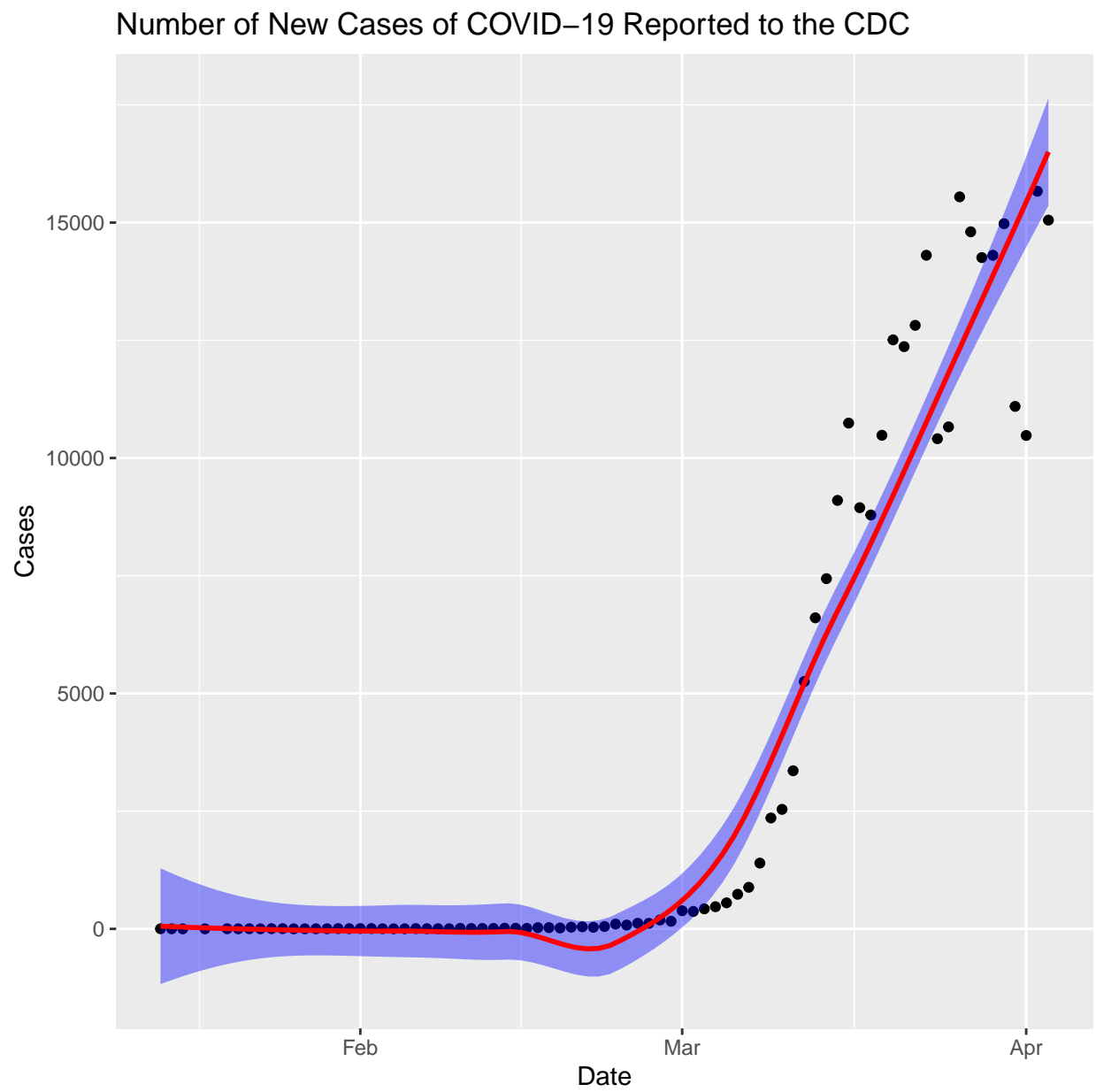Figure 21: Another look at growth

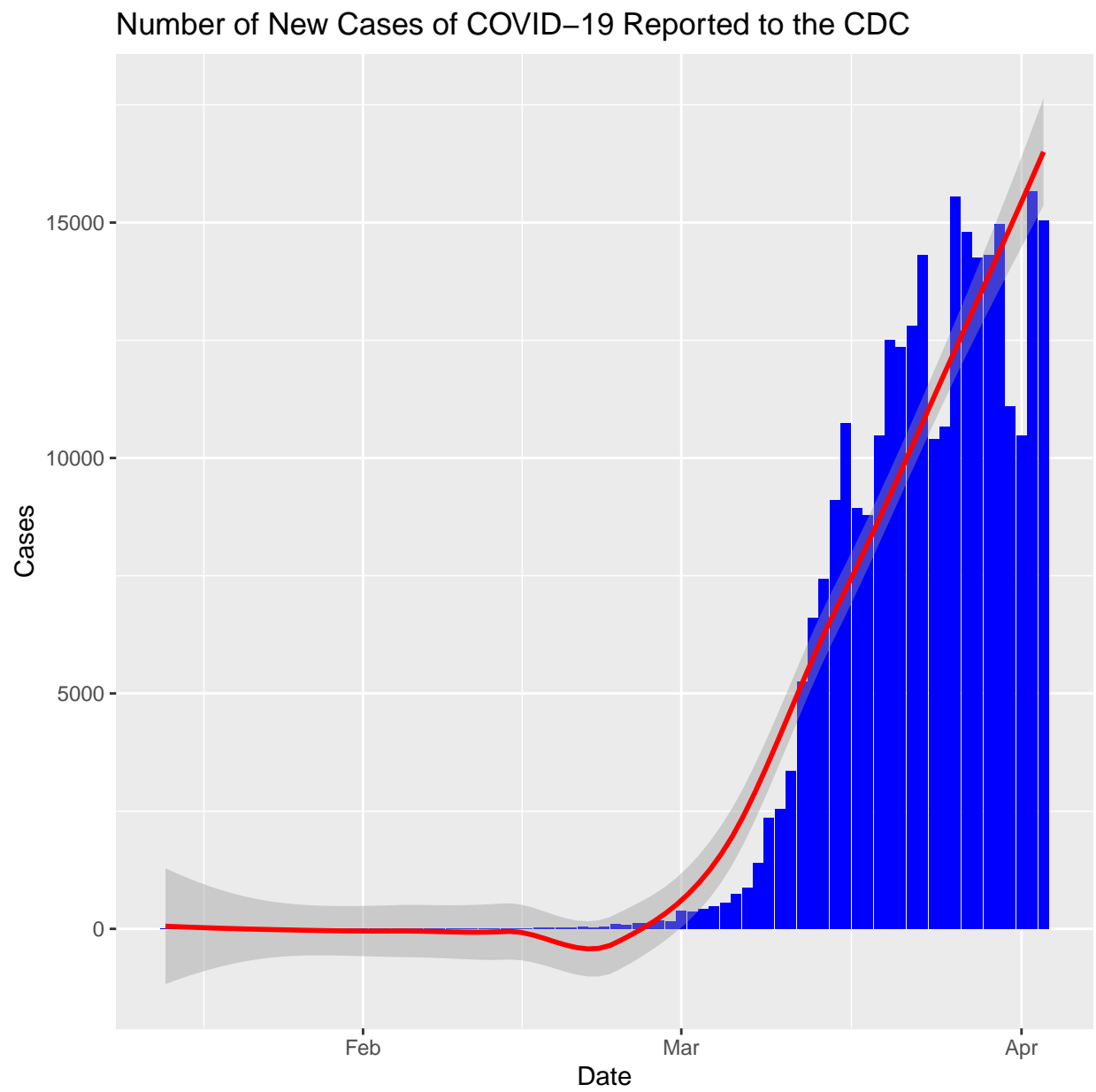Figure 22: Epi curve 1

Figure 23: Epi curve 2, traditional

```
            fill = "blue") +
    labs(y = "Cumulative number of cases",
         title = "Cumulative Number of Cases of COVID-19 Reported to the CDC")
```
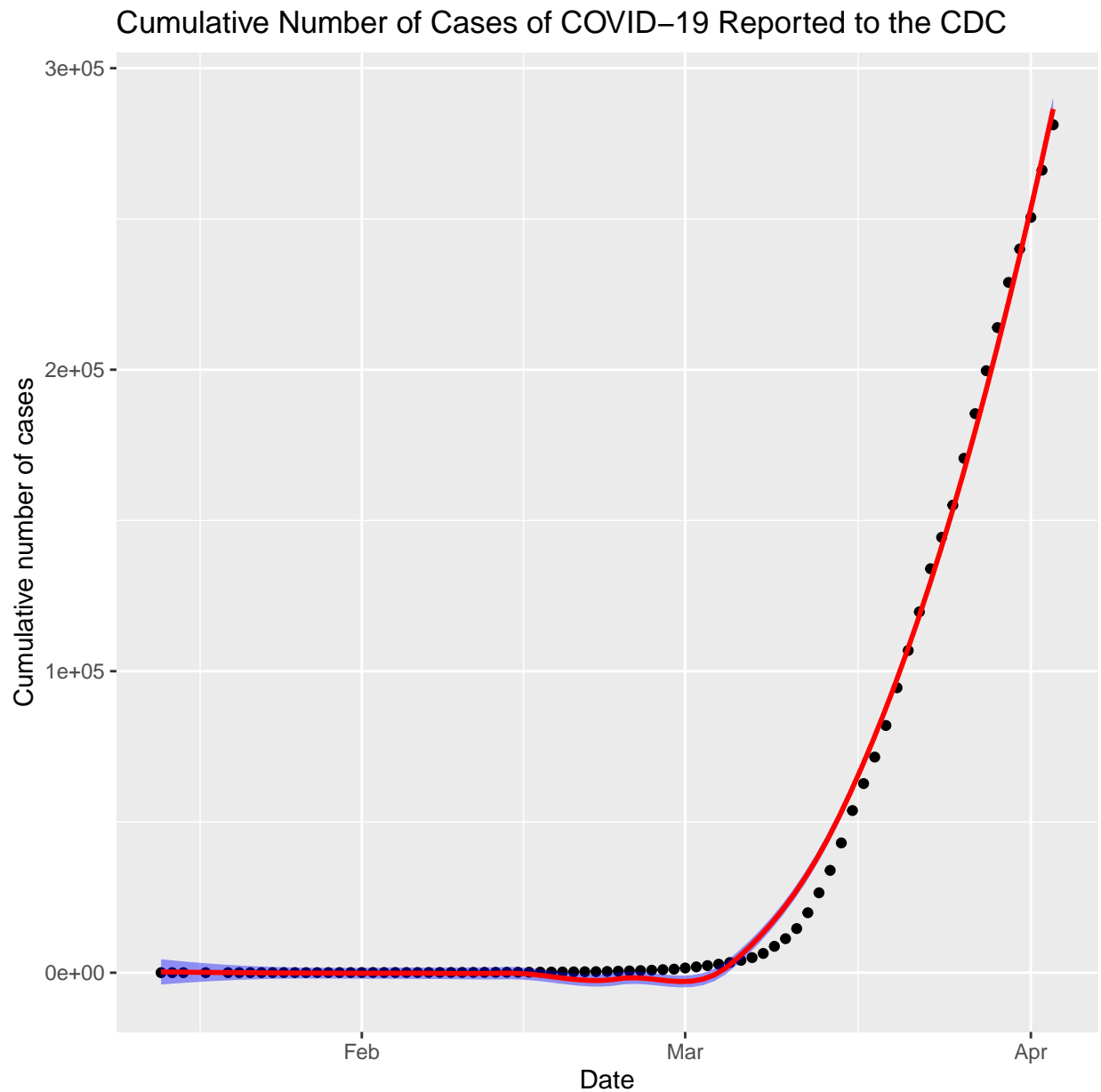

Figure 24: Cumulative cases

```
gr <- cdc %>%
  filter(`Number of new cases` != 0)

ggplot(data = gr) +
  geom_line(aes(y = `Number of new cases`,
                x = cum)) +
  scale_x_log10() +
```

```
scale_y_log10() +
geom_smooth(aes(y = `Number of new cases`,
                x = cum),
            method = "loess",
            formula = y ~ x) +
labs(title = "Daily Growth Rate",
     x = "Total Cases (log)",
     y = "New Cases (log)")
```
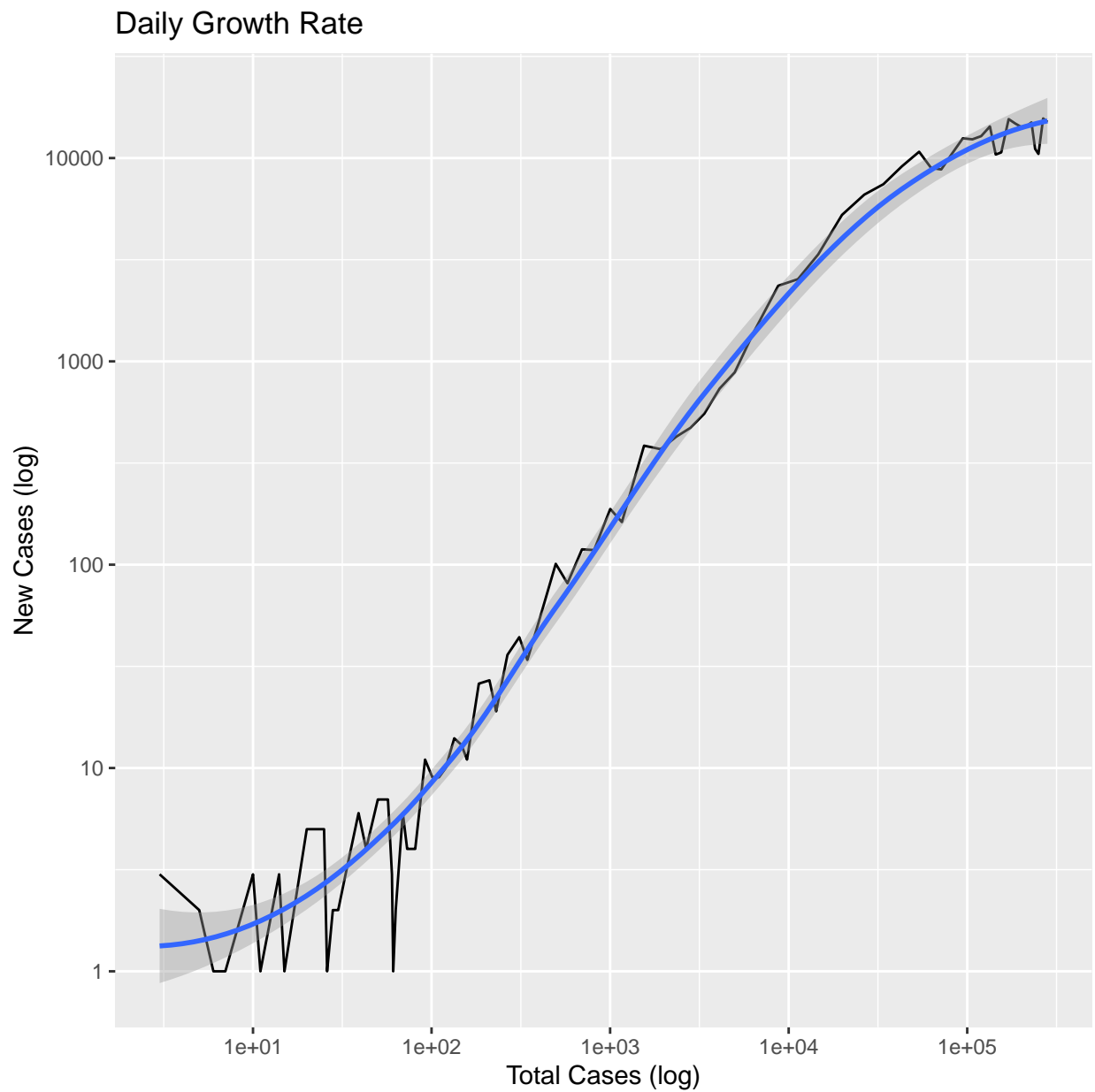

Figure 25: Growth Rate

```
ggplot(data = gr) +
    geom_line(aes(y = `cum`,
```

```
                    x = `Date`)) +
    geom_point(aes(y = `cum`,
                     x = `Date`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cum`,
                     x = `Date`),
            method = "loess",
            formula = y ~ x) +
  labs(title = "Growth Rate",
       x = "Date",
       y = "Total Cases (log)")
```
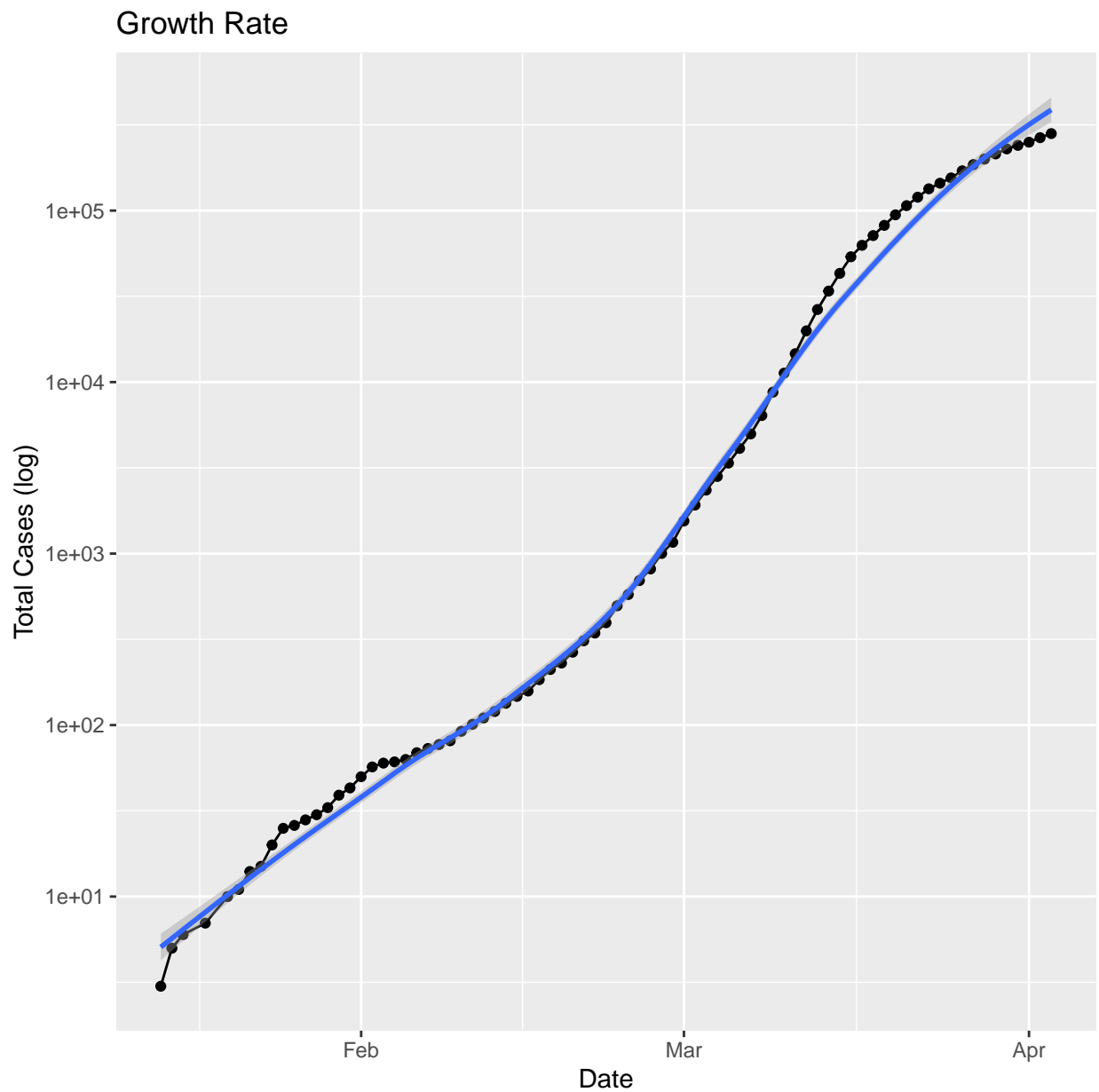
Growth Rate

Figure 26: Another look at growth

```r
gr <- cdc %>%
    group_by(`Week Number`) %>%
    summarise(cases = sum(`Number of new cases`),
              cs = sum(cum),
              count = n())

ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = cs)) +
    geom_point(aes(y = `cases`,
                   x = cs)) +
    scale_x_log10() +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = cs),
            method = "loess",
            formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```r
ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = `Week Number`)) +
    geom_point(aes(y = `cases`,
                   x = `Week Number`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = `Week Number`),
            method = "loess",
            formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Week of the year",
       y = "New Cases (log)")
```

Using WHO data for US

```r
US <- WHO %>%
  filter(WHO$`Country Territory area` == "United States of America")
```

```r
epi <- ggplot(data = US)
```

```r
epi + geom_point(aes(x = date,
               y = `Confirmed new cases`))+
    geom_smooth(aes(x = date,
               y = `Confirmed new cases`),
               method = "loess",
               formula = y ~ x,
               color = "red",
               fill = "blue") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19")
```

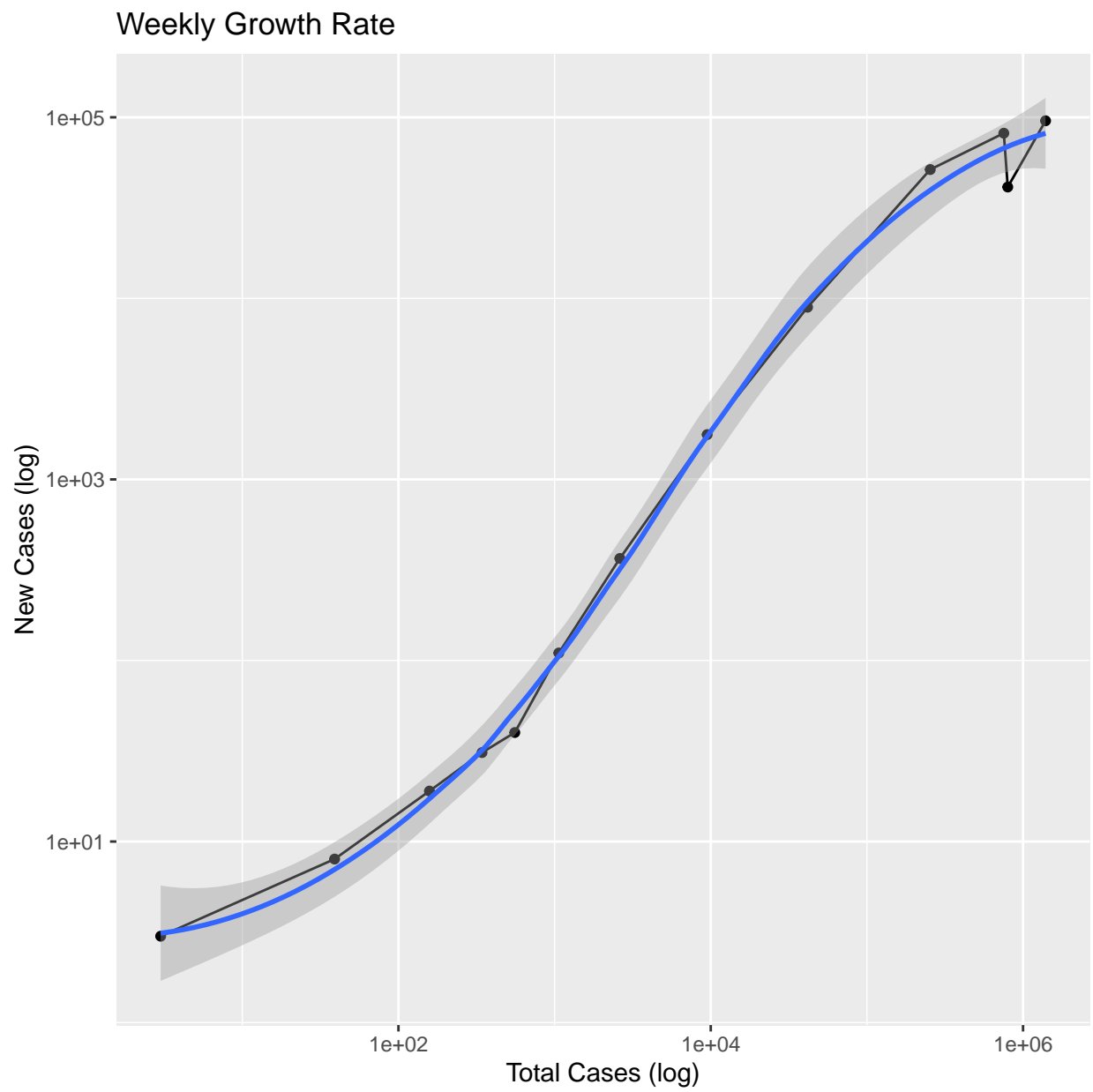# Weekly Growth Rate

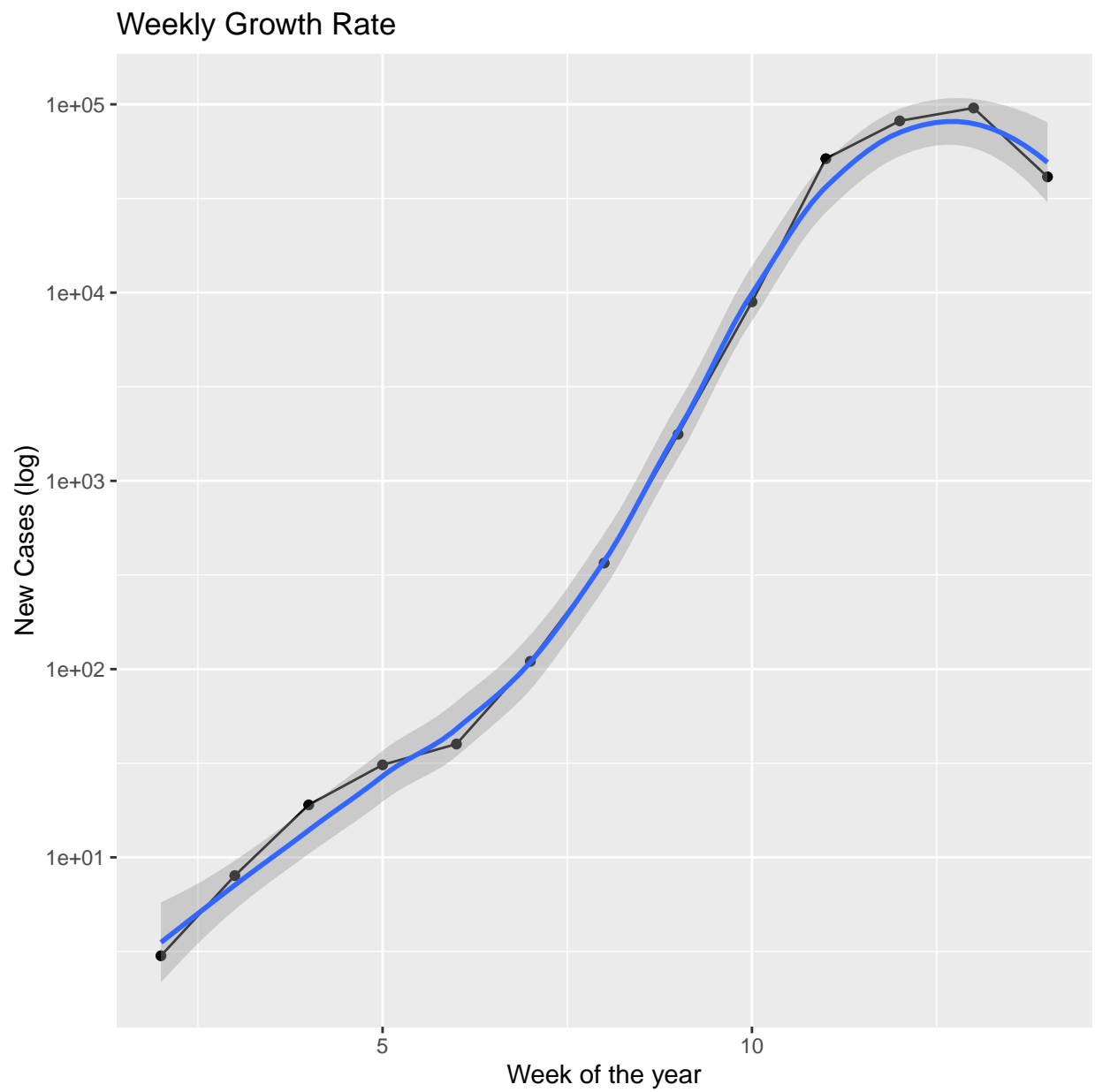Figure 27: weekly growth rate

Figure 28: Another look at growth
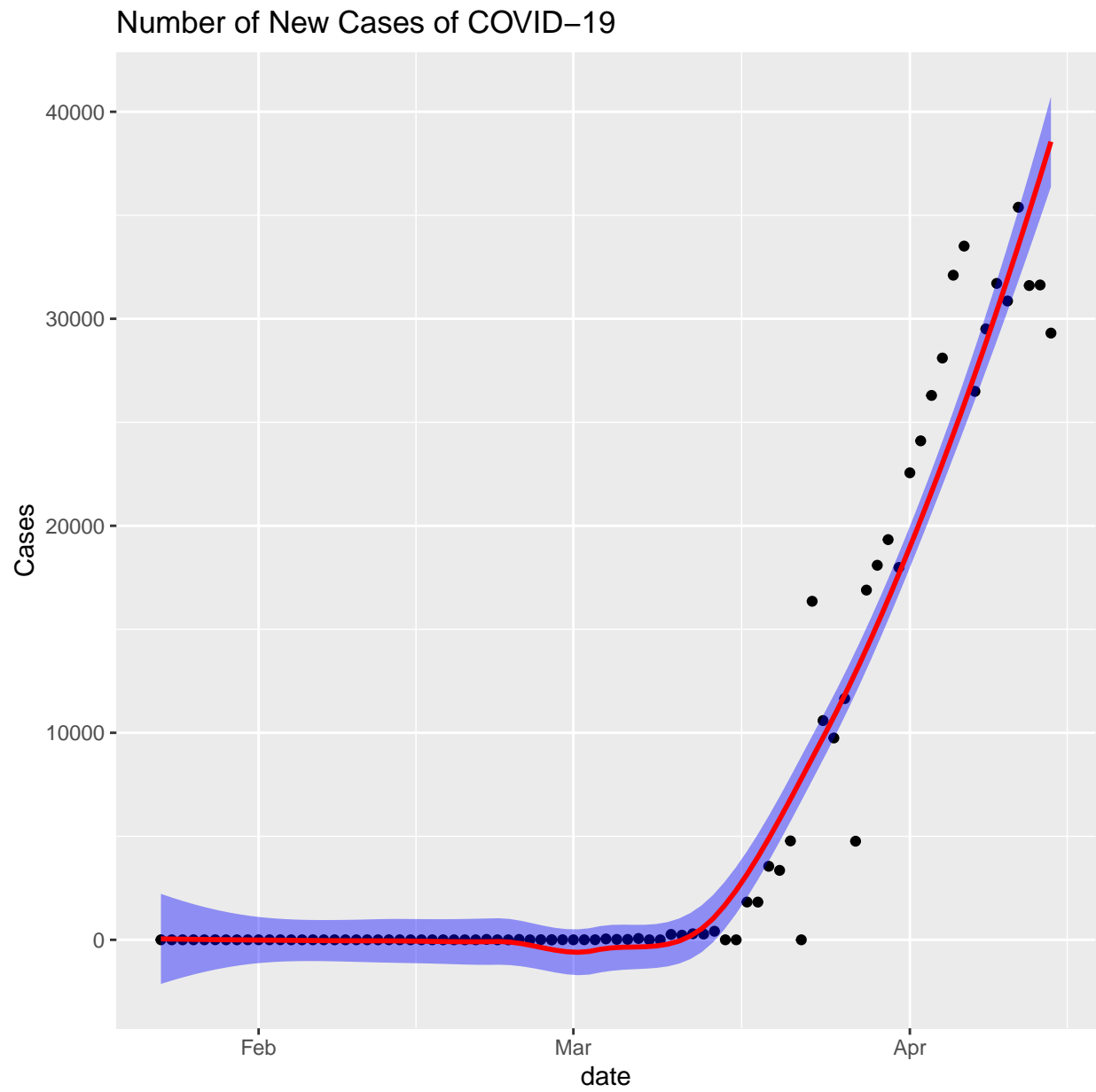
Figure 29: Epi curve 1

```
epi + geom_col(aes(x=date,
                   y=`Confirmed new cases`),
               fill= "blue") +
    geom_smooth(aes(x=date,
                   y=`Confirmed new cases`),
               method = "loess",
               formula = y ~ x,
               color = "red") +
    labs(y = "Cases",
         title = "Number of New Cases of COVID-19 Reported")
```
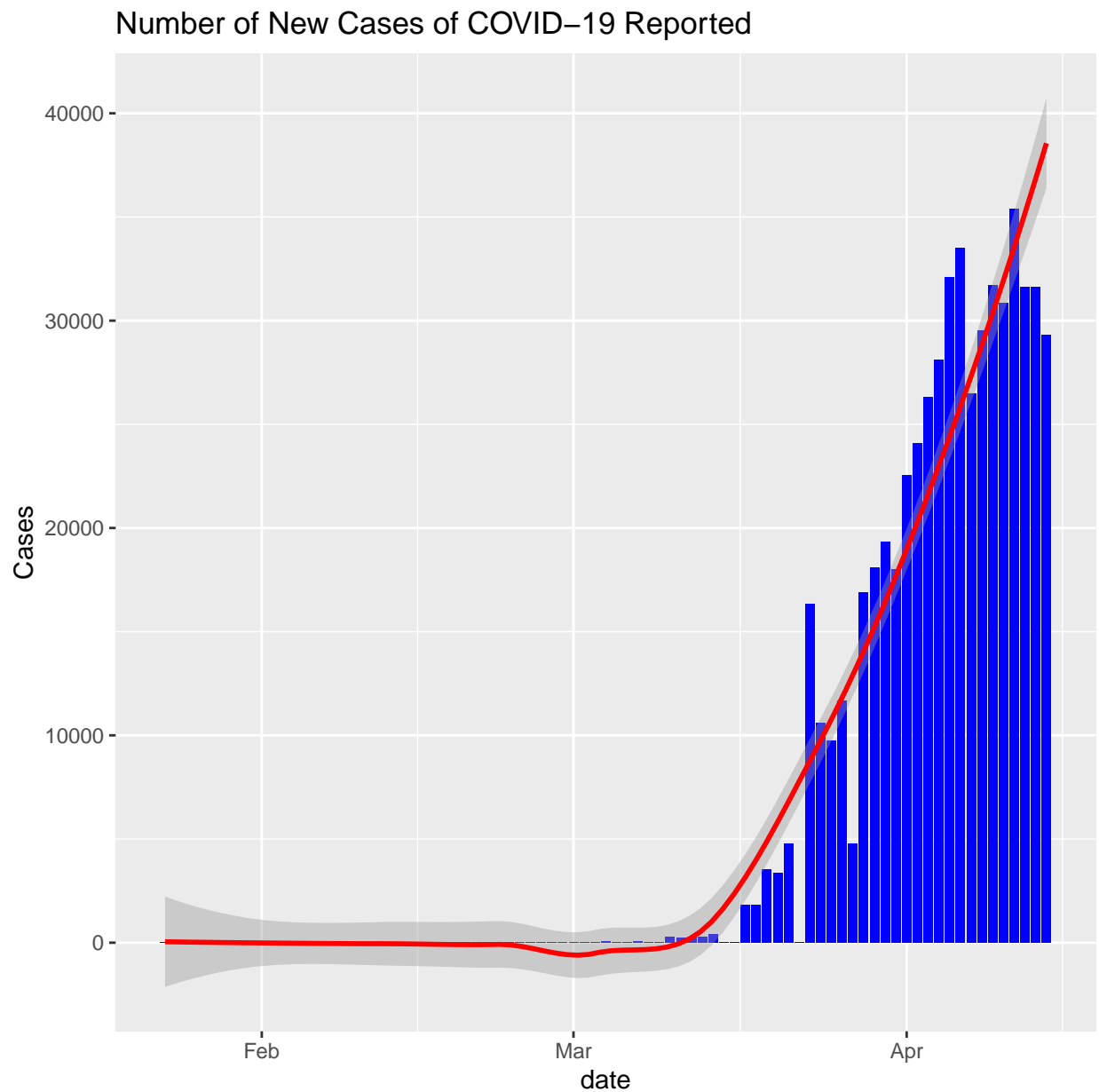

Figure 30: Epi curve 2, traditional

```
epi + geom_point(aes(x = date,
                     y = `total Confirmed Cases`))+
    geom_smooth(aes(x = date,
                    y = `total Confirmed Cases`),
                method = "loess",
                formula = y ~ x,
                color = "red",
                fill = "blue") +
    labs(y = "Cumulative number of cases",
         title = "Cumulative Number of Cases of COVID-19 Reported")
```
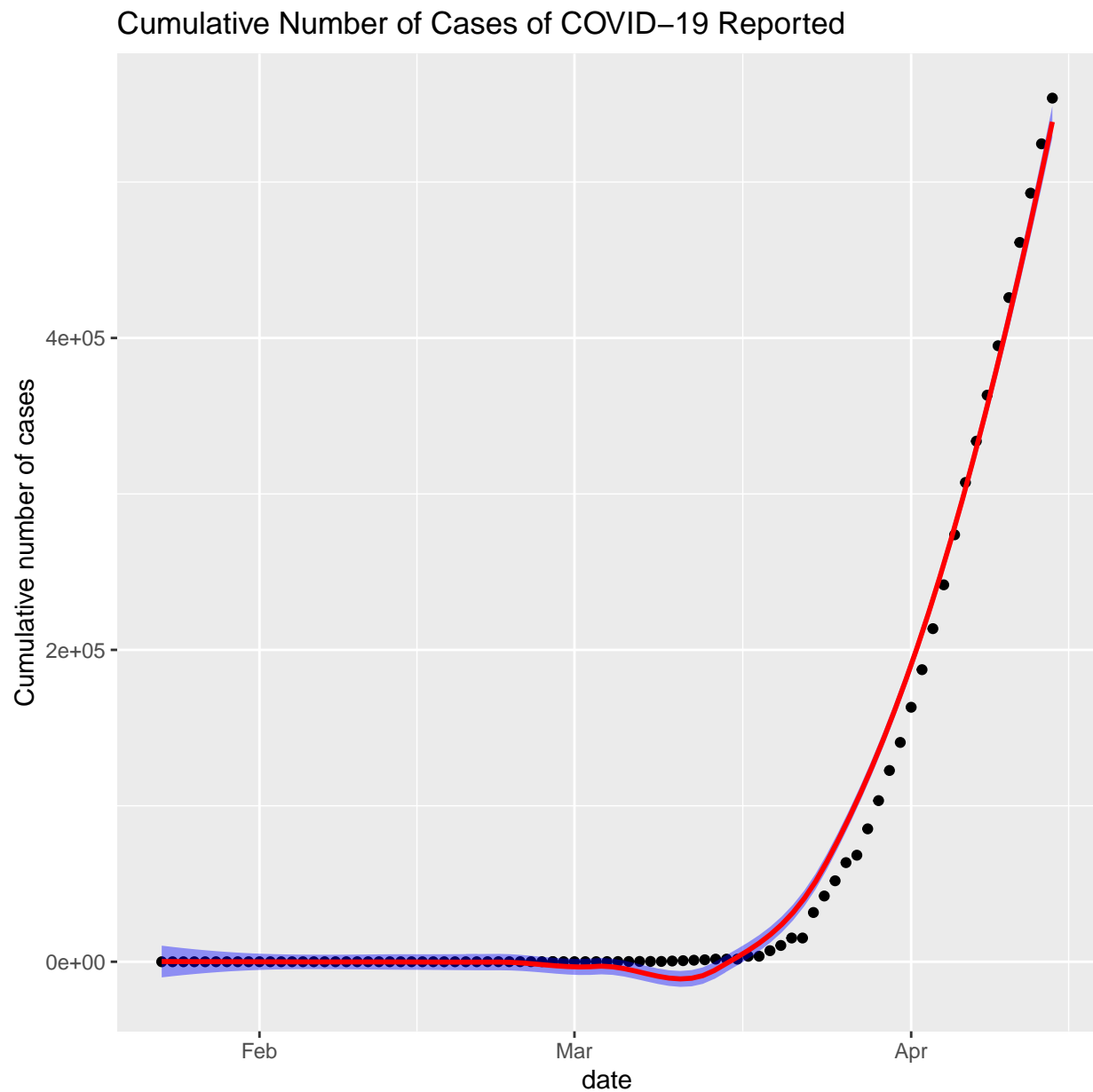

Figure 31: Cumulative cases

```r
gr <- US %>%
  filter(`Confirmed new cases` != 0)

ggplot(data = gr) +
  geom_line(aes(y = `Confirmed new cases`,
                x = `total Confirmed Cases`)) +
  scale_x_log10() +
  scale_y_log10() +
  geom_smooth(aes(y = `Confirmed new cases`,
                  x = `total Confirmed Cases`),
              method = "loess",
              formula = y ~ x) +
  labs(title = "Daily Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```r
ggplot(data = gr) +
  geom_line(aes(x = `date`,
                y = `total Confirmed Cases`)) +
  scale_y_log10() +
  geom_smooth(aes(x = `date`,
                  y = `total Confirmed Cases`),
              method = "loess",
              formula = y ~ x) +
  labs(title = "Daily Growth Rate",
       x = "Date",
       y = "Total Cases (log)")
```

```r
gr <- US %>%
    group_by(`Week Number`) %>%
    summarise(cases = sum(`Confirmed new cases`),
              cs = sum(`total Confirmed Cases`),
              count = n()) %>%
    filter(cases != 0)

ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = cs)) +
    geom_point(aes(y = `cases`,
                   x = cs)) +
    scale_x_log10() +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = cs),
              method = "loess",
              formula = y ~ x) +
  labs(title = "Weekly Growth Rate",
       x = "Total Cases (log)",
       y = "New Cases (log)")
```

```r
ggplot(data = gr) +
    geom_line(aes(y = `cases`,
                  x = `Week Number`)) +
    geom_point(aes(y = `cases`,
```
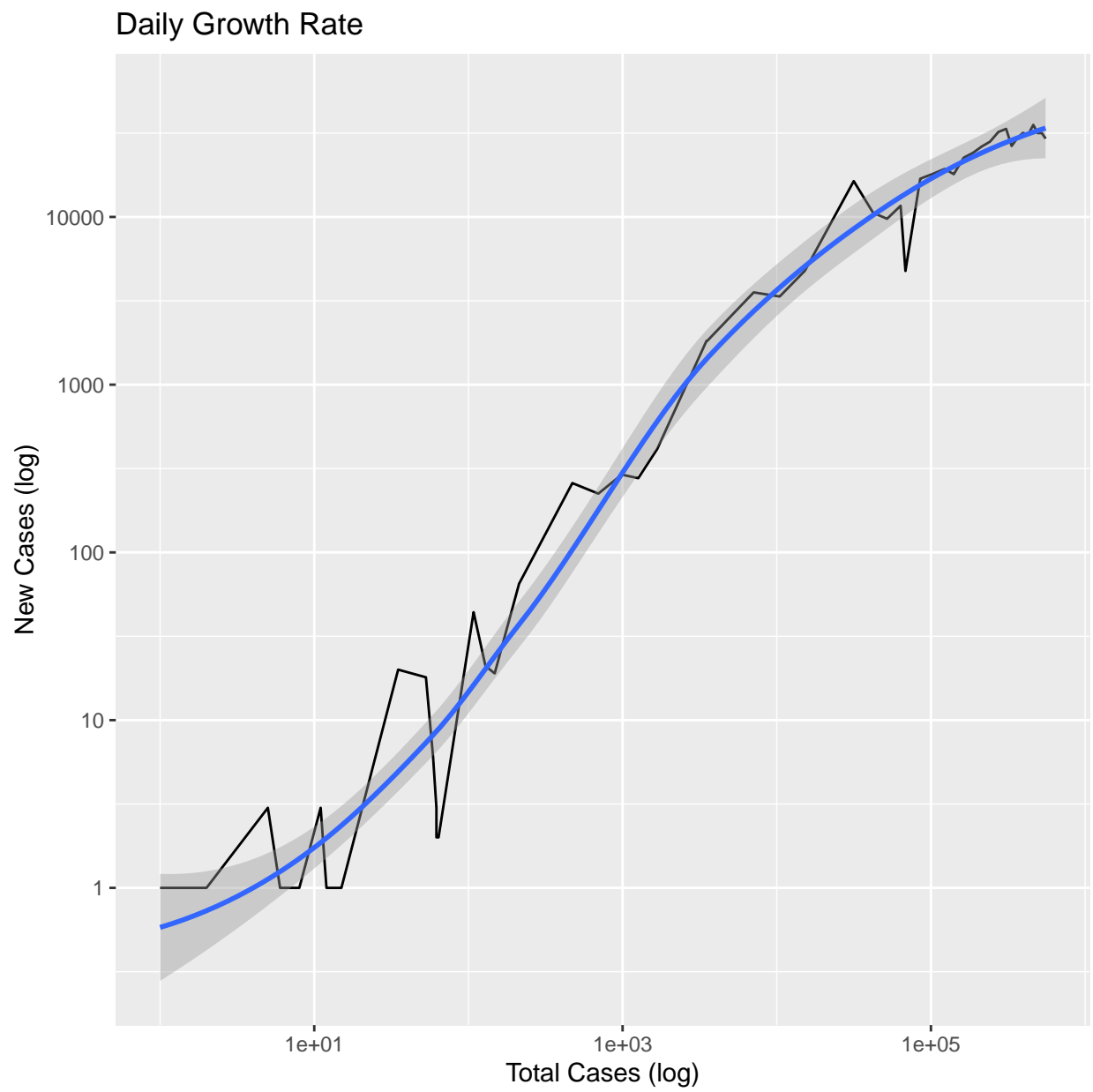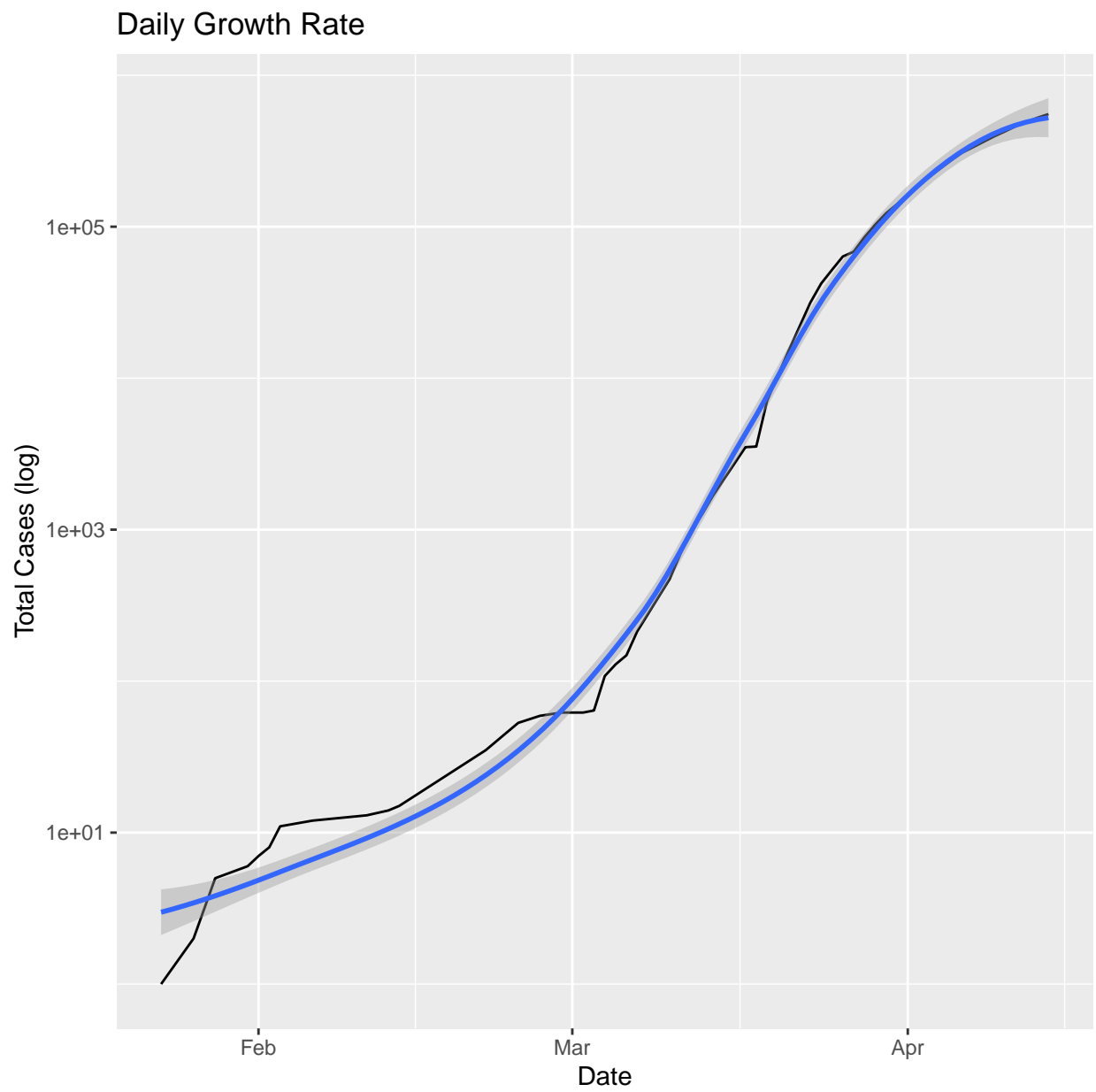
Figure 32: Growth Rate
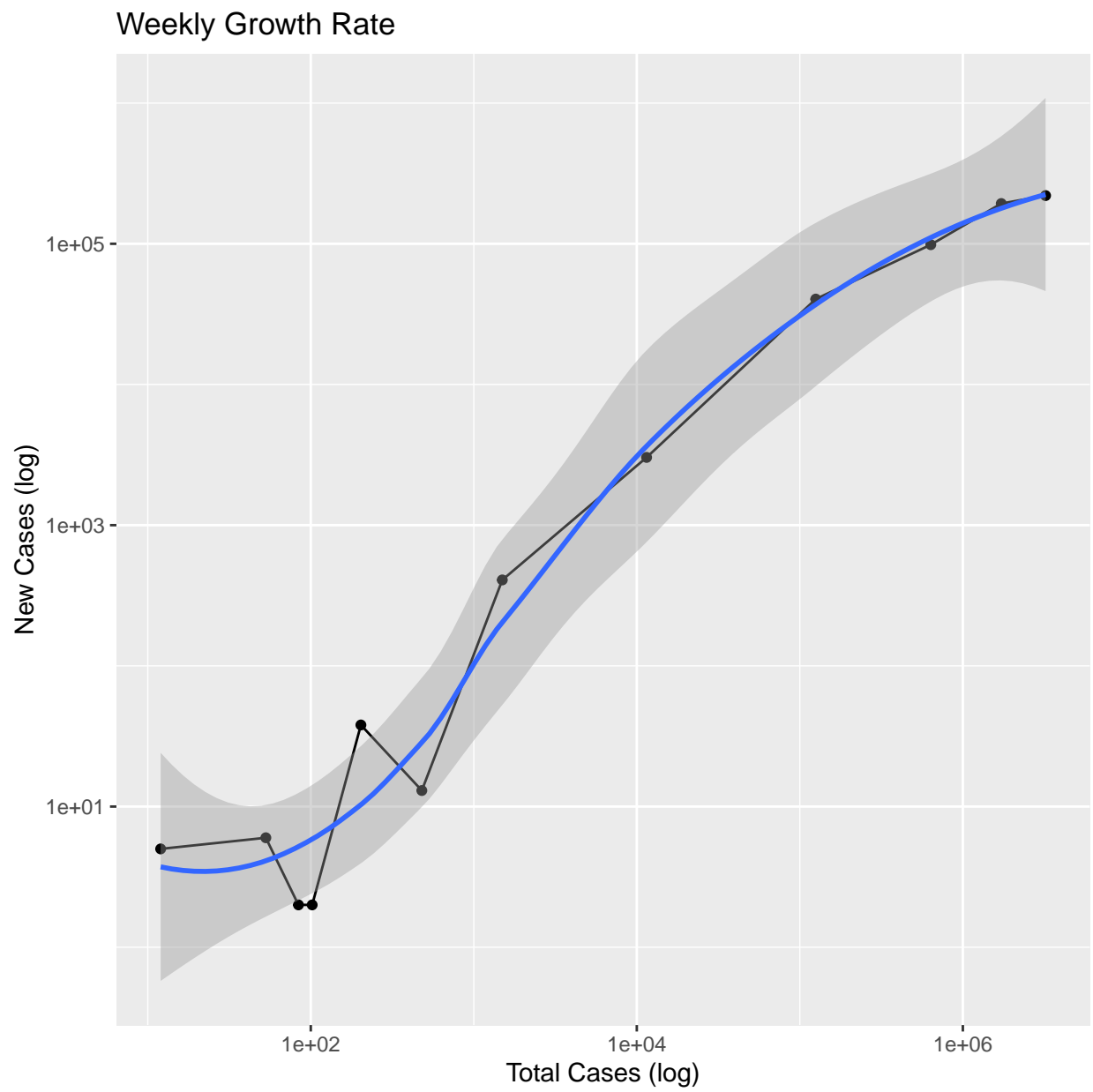
Figure 33: Another look at growth

Weekly Growth Rate

Figure 34: stable weekly growth rate

```
                  x = `Week Number`)) +
    scale_y_log10() +
    geom_smooth(aes(y = `cases`,
                    x = `Week Number`),
            method = "loess",
            formula = y ~ x) +
labs(title = "Weekly Growth Rate",
     x = "Week of the year",
     y = "New Cases (log)")
```

Weekly Growth Rate

Figure 35: Another look at growth