

Ebola Forecasting Analysis

Andy Shen

8/13/2020

Data Input and Cleaning

```
rm(list=ls())
source("outbreak_vis.R") #script with functions

## -- Attaching packages -----
## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.1      v dplyr  1.0.0
## v tidyr   1.1.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0

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

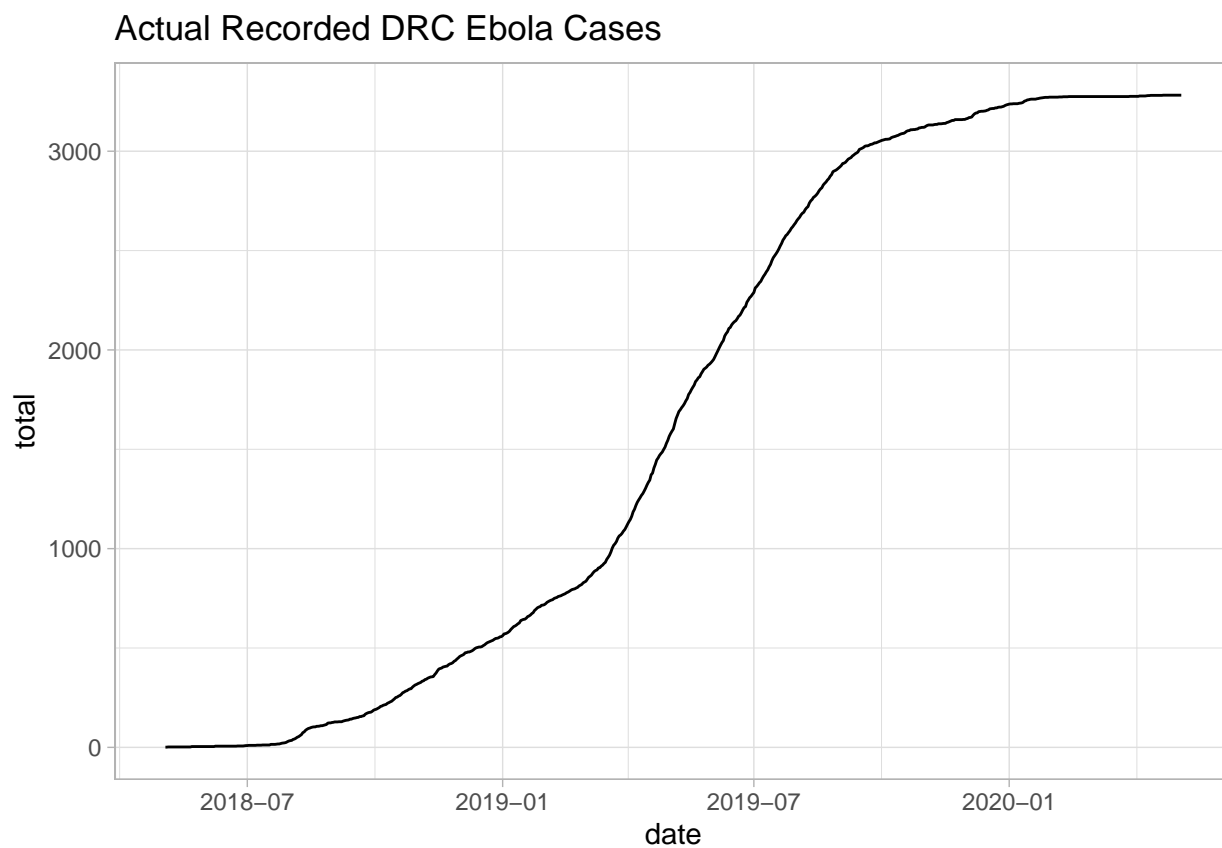
##
## Attaching package: 'lubridate'

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

rgx <- "\\d{1,2}\\./\\d{1,2}\\./\\d{4}" #date structure regex
true <- read.csv("/Volumes/GoogleDrive/.shortcut-targets-by-id/15UGkfREtfqH3LdfHmCsSpFJ5SrTnSeyt/ebola/")
true <- true[str_detect(true$Date, rgx),] #omits rows without a date
colnames(true) <- c("date", "cases")
true$date <- mdy(true$date)
true$cases[is.na(true$cases)] <- 0
true <- true %>% mutate(total = cumsum(cases))
last_date <- true$date[length(true$date)]
last_case <- true$total[length(true$total)]
```

Actual Recorded DRC Ebola Cases

```
p <- ggplot(data = true,  
            mapping = aes(x = date, y = total)) +  
  geom_line() + theme_light() # graph of runinng total of cases  
  
p + labs(title = "Actual Recorded DRC Ebola Cases")
```

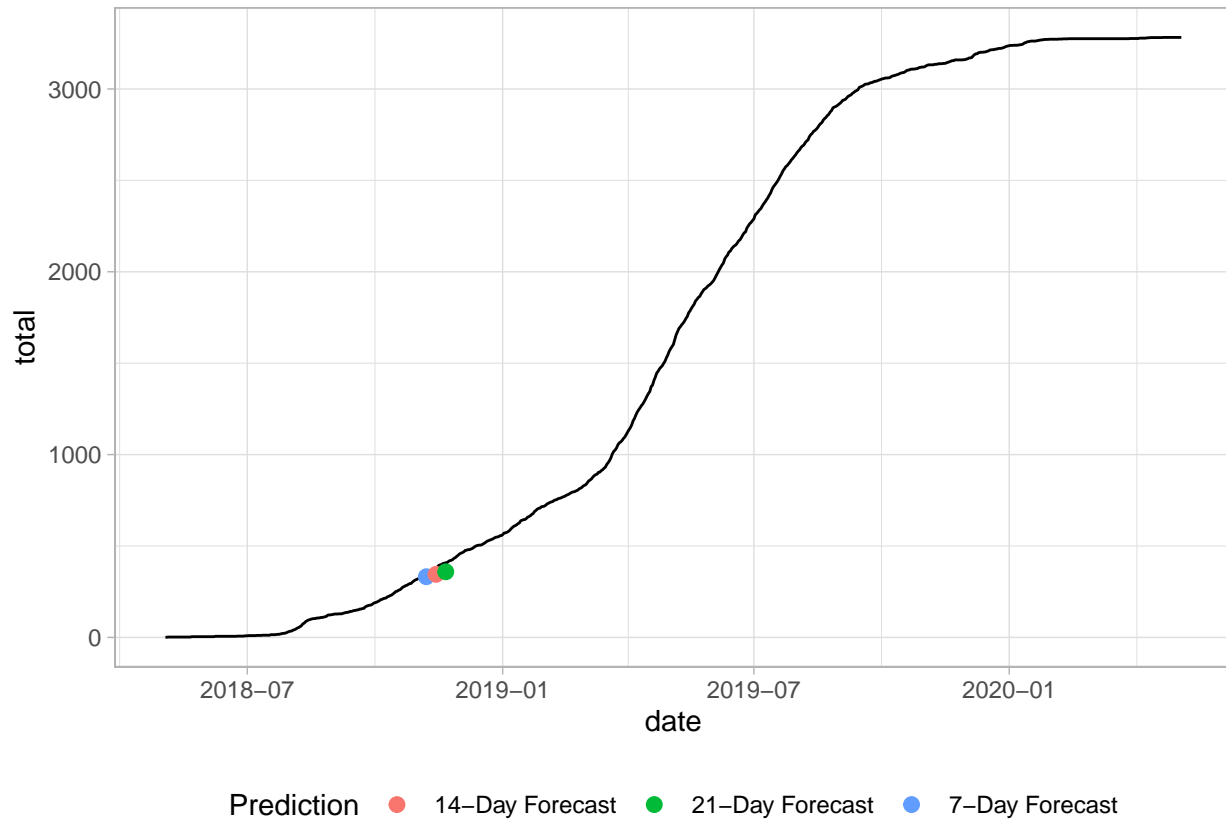


Accuracy of Hawkes Projections

Single Forecast Visualization for Entire Outbreak

Shows predicted vs actual for one forecast with respect to entire outbreak.

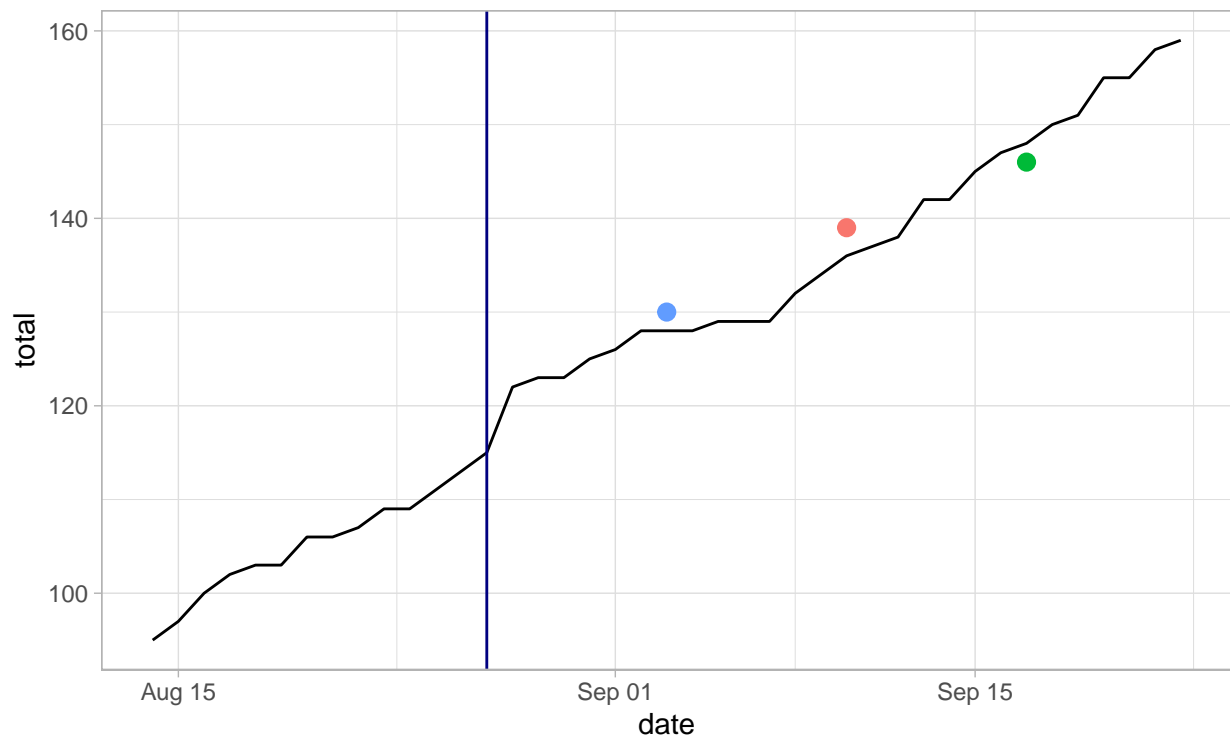
```
full_forecast("2018-10-31",c(15,28,42))
```



Single Forecast Visualization

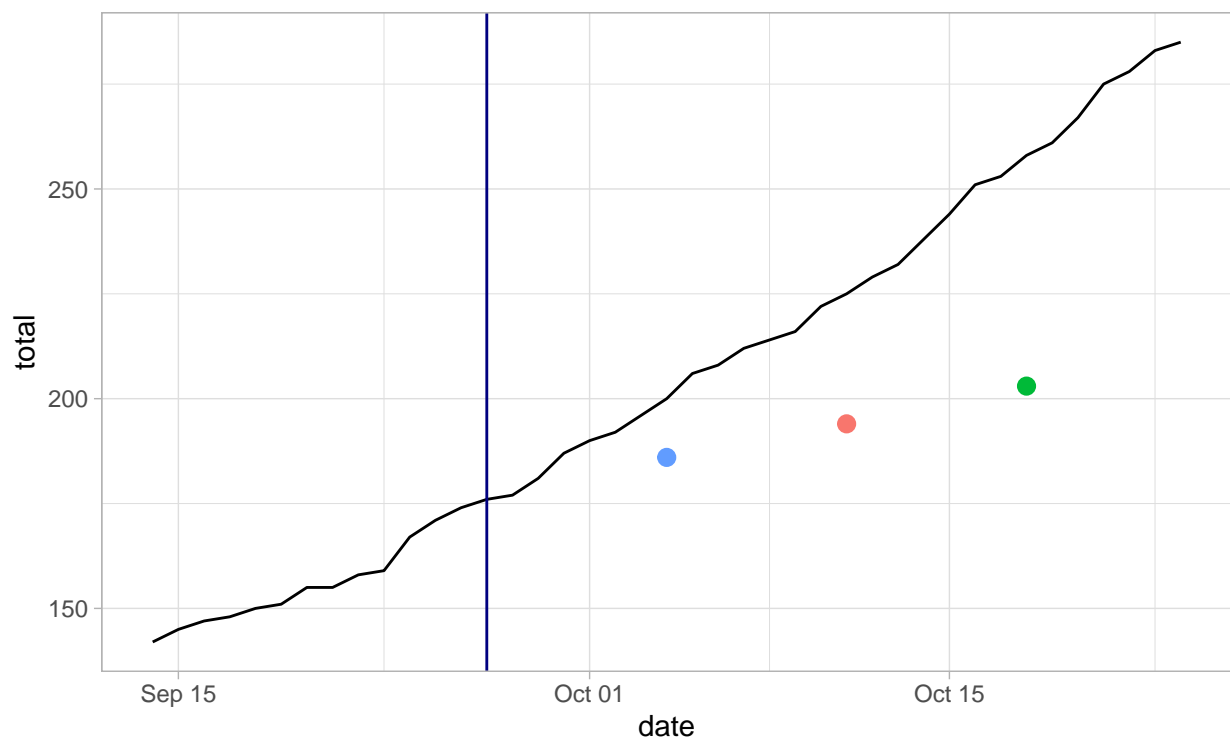
Shows predicted vs actual for one forecast with respect to that date range.

```
multi_forecast("2018-08-27",c(15,24,31)) #shows predicted vs actual for one forecast
```



Prediction ● 14-Day Forecast ● 21-Day Forecast ● 7-Day Forecast

`multi_forecast("2018-09-27",c(10,18,27))` *#shows predicted vs actual for one forecast*

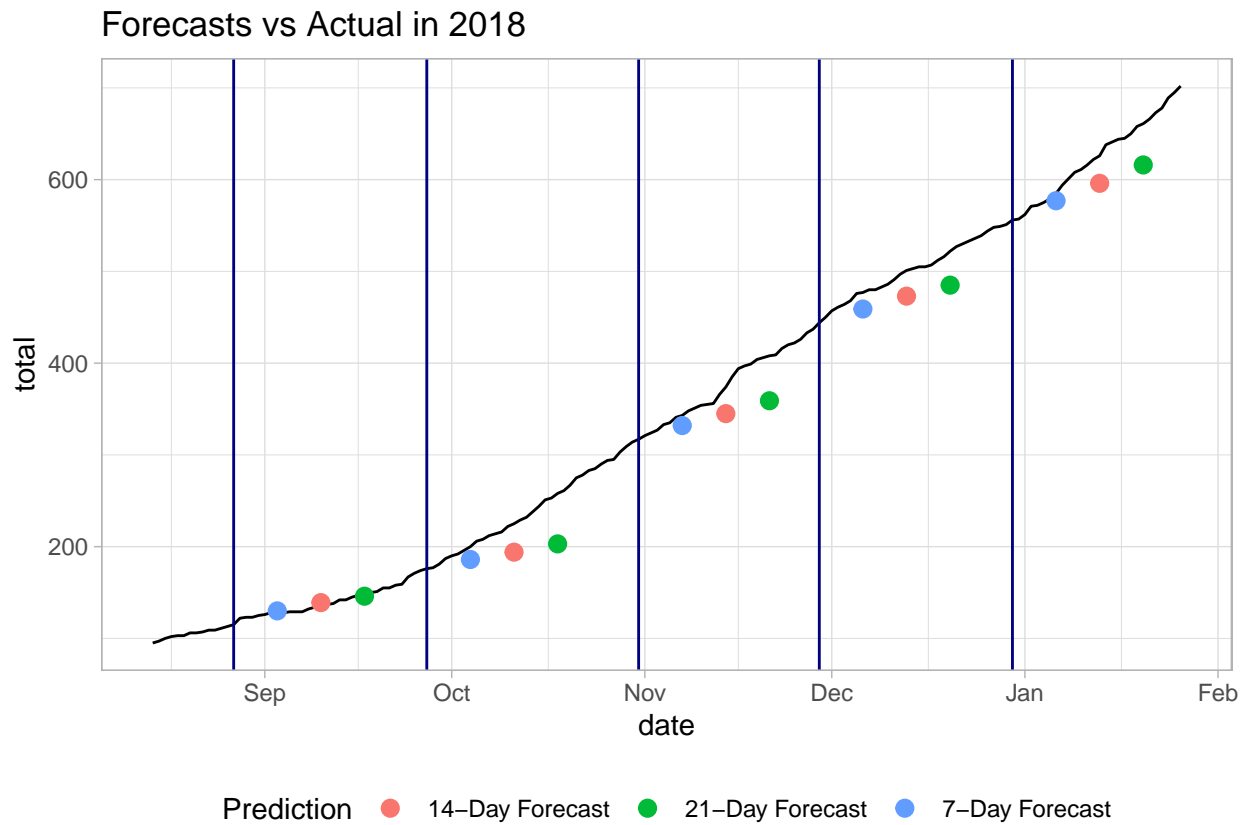


Prediction ● 14-Day Forecast ● 21-Day Forecast ● 7-Day Forecast

Forecasts vs Actual in 2018

Monthly forecasts during all available data in 2018.

```
title <- "Forecasts vs Actual in 2018"
dv <- c("2018-08-27", "2018-09-27", "2018-10-31", "2018-11-29", "2018-12-30")
mt <- cbind(c(15, 24, 31), c(10, 18, 27), c(15, 28, 42), c(15, 29, 41), c(21, 40, 60))
multi_forecast(dv, mt, title = title)
```

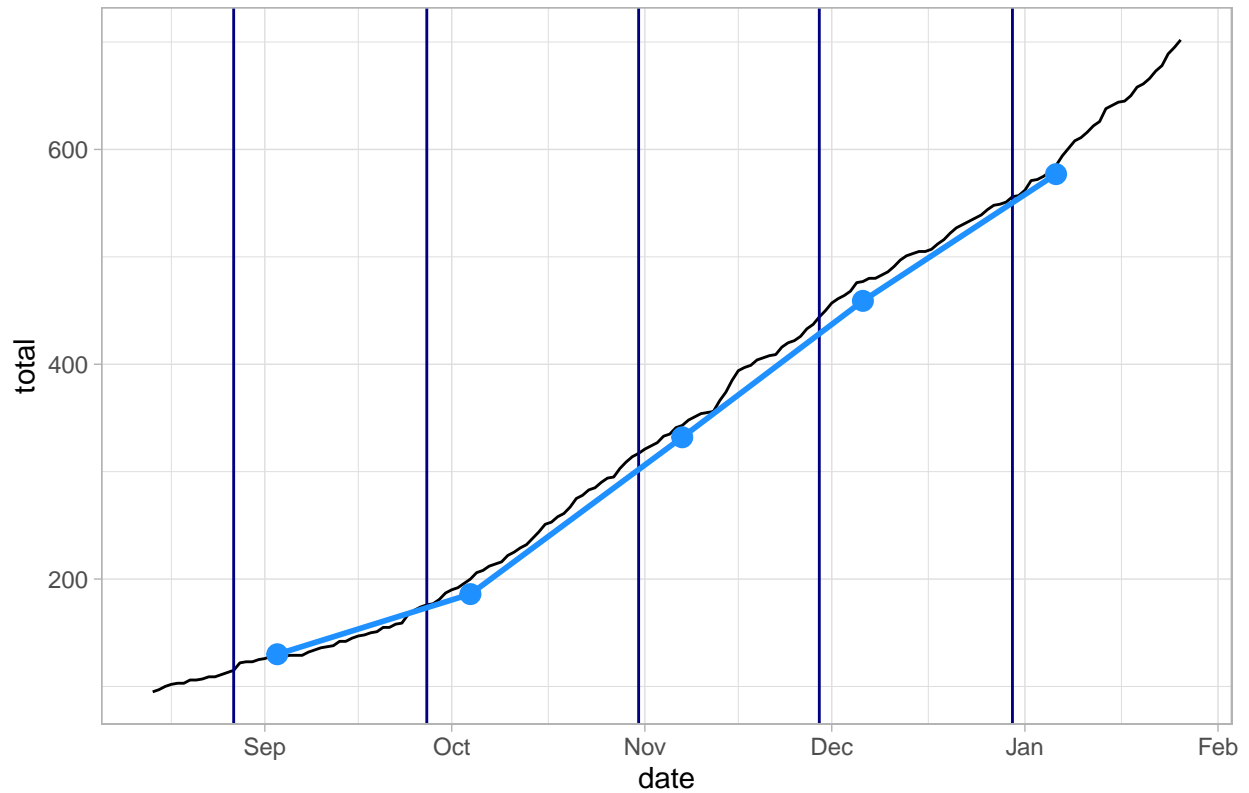


2018 7-day Forecasts

```
title <- "7-Day Forecasts vs Actual in 2018"
single_forecast(dv, mt, days = 7, title = title, res = TRUE)
```

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

7-Day Forecasts vs Actual in 2018



```
##
```

```
## [[2]]
```

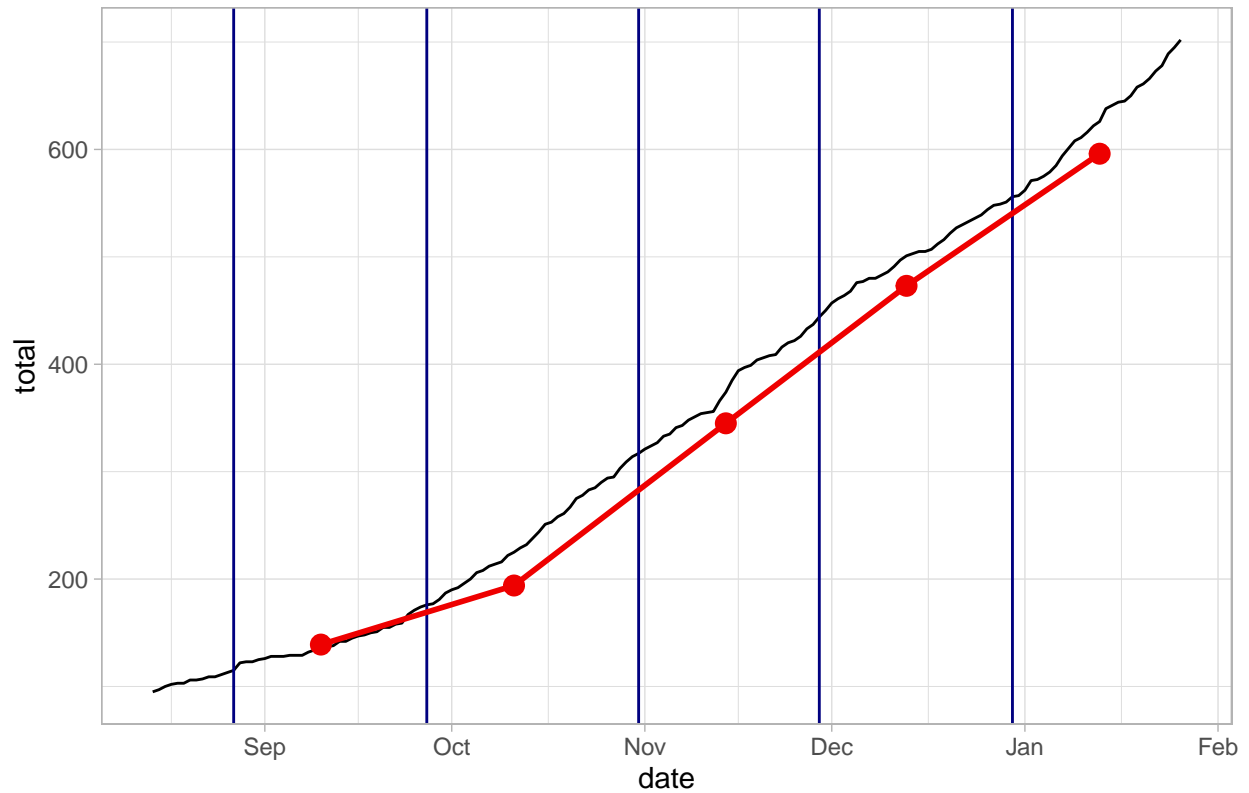
```
##   prior_date prior total forecast_date forecast.total.7 actual.total.7
## 1 2018-08-27    115    2018-09-03          130          128
## 2 2018-09-27    176    2018-10-04          186          200
## 3 2018-10-31    317    2018-11-07          332          343
## 4 2018-11-29    444    2018-12-06          459          477
## 5 2018-12-30    556    2019-01-06          577          585
```

2018 14-day Forecasts

```
title <- "14-Day Forecasts vs Actual in 2018"
single_forecast(dv, mt, days = 14, title = title, res = TRUE)
```

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

14-Day Forecasts vs Actual in 2018



```
##
```

```
## [[2]]
```

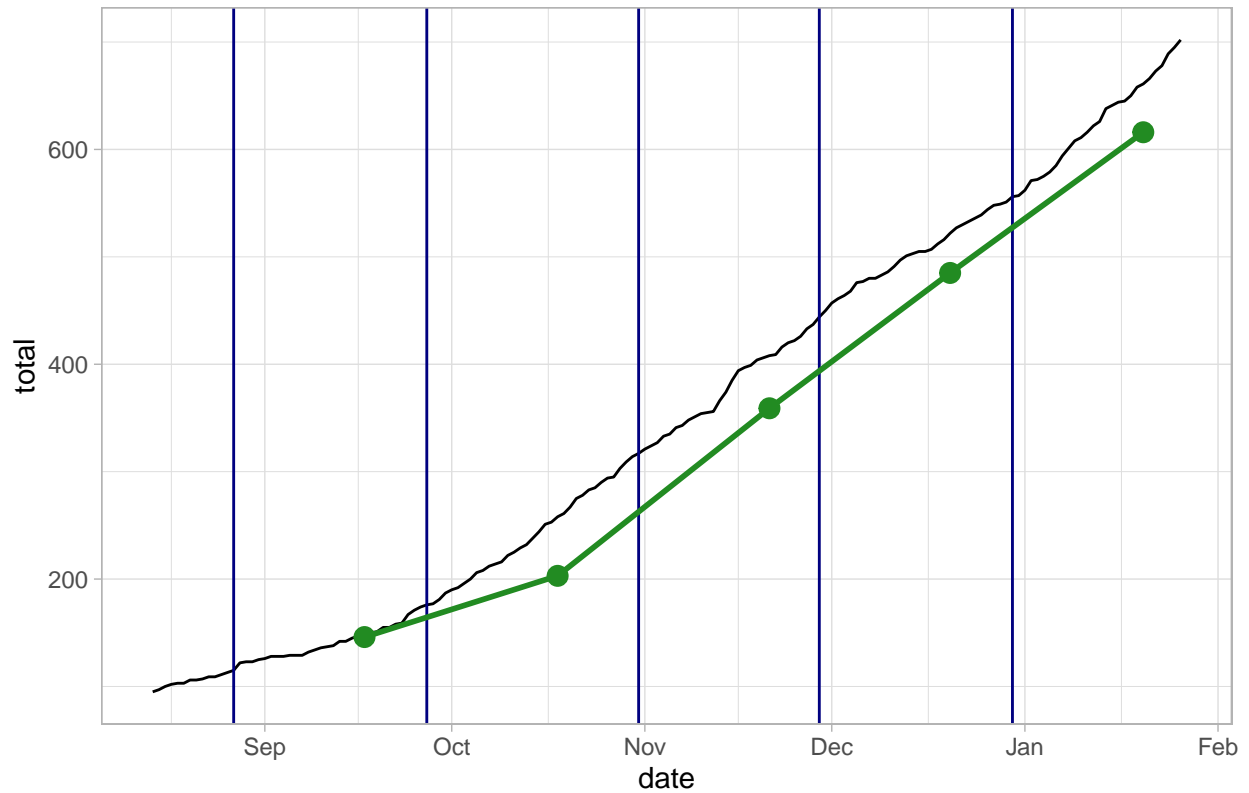
##	prior_date	prior	total	forecast_date	forecast.total.14	actual.total.14
## 1	2018-08-27	115		2018-09-10	139	136
## 2	2018-09-27	176		2018-10-11	194	225
## 3	2018-10-31	317		2018-11-14	345	374
## 4	2018-11-29	444		2018-12-13	473	501
## 5	2018-12-30	556		2019-01-13	596	626

2018 21-day Forecasts

```
title <- "21-Day Forecasts vs Actual in 2018"
single_forecast(dv, mt, days = 21, title = title, res = TRUE)
```

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

21-Day Forecasts vs Actual in 2018



```
##
```

```
## [[2]]
```

	prior_date	prior	total	forecast_date	forecast.total.21	actual.total.21
## 1	2018-08-27	115		2018-09-17	146	148
## 2	2018-09-27	176		2018-10-18	203	258
## 3	2018-10-31	317		2018-11-21	359	408
## 4	2018-11-29	444		2018-12-20	485	522
## 5	2018-12-30	556		2019-01-20	616	661