



I predicted on 25th August a second-wave in London in 2-4 weeks time. It looks like the recent media's coverage and the government's policy reaction are consistent with my prediction.

The policy response of rule-of-six-in-social-gathering is mild. Hopefully though, it forces people to realise that they cannot afford to relax too much. Nevertheless, from talking to friends & colleagues, I don't get the sense that people are ready to scale back their activity and social events!

I have fitted an exponential curve to the recent data and I get a whopping 98% R2 and <1% p-value (OK – R2 is a bit inflated on overlapping data but still.). Here I added to my graph the fitted value and the model's predicted value for the next four weeks.

#covid19 #secondwave

The full code below. The code is essentially based on my previous post with the additional exponential model in the middle section.

```
library(data.table)

library(ggplot2)

data_url <- "https://c19downloads.azureedge.net/downloads/csv/coronavirus-cases_latest.csv"

raw_data <- fread(data_url, check.names = TRUE)

area_name <- "London"

area_type <- "region"

area_data <- raw_data[
  Area.name == area_name &
  Area.type == area_type,,
][,Specimen.date := as.Date(Specimen.date)
][,c("Specimen.date","Daily.lab.confirmed.cases")][
  order(Specimen.date)
]

area_data <- merge(area_data,
  data.table(Specimen.date = seq(
    min(area_data[,Specimen.date]),
    max(area_data[,Specimen.date]),
    by = "1 day"
  )), all = TRUE, by = "Specimen.date")

setkey(area_data, Specimen.date)

setnafill(area_data, type = "const", fill = 0,
  cols = c("Daily.lab.confirmed.cases"))

area_data[,roll_mean := frollmean(Daily.lab.confirmed.cases, n = 7, align = "right")]
```

```
#####
#####Exponential model#####
#####
area_data[,increasing := c(rep(NA,7), roll_mean[-(1:7)]- roll_mean[-(.N-6):.N])>0)]
end_date <- area_data[order(Specimen.date, decreasing = TRUE)][increasing==TRUE,,][,
  Specimen.date[1], by="increasing"]$V1
start_date <- area_data[order(Specimen.date, decreasing = TRUE)][
  increasing==FALSE & Specimen.date < end_date,,][,
  Specimen.date[1], by="increasing"]$V1
exp_lm_data <- area_data[Specimen.date > start_date & Specimen.date <= end_date,,]
exp_lm_data[, days := 1:.N]
exp_lm <- lm(log(roll_mean)~ days, data = exp_lm_data)
exp_lm_data[,fitted_numbers := exp(fitted.values(exp_lm))]
predicted_data <- data.table(days=max(exp_lm_data$days)+1:28)
predicted_data[,Specimen.date := min(exp_lm_data$Specimen.date)+ lubridate::days(days)]
predicted_data[,predicted_numbers := exp(predict.lm(exp_lm, predicted_data))]
#####
m_area_data <- melt(area_data, id.vars="Specimen.date",
  measure.vars = c("Daily.lab.confirmed.cases","roll_mean"))
exp_lm_data <- melt(dplyr::bind_rows(exp_lm_data, predicted_data),
  id.vars="Specimen.date",
  measure.vars = c("fitted_numbers","predicted_numbers"))
m_area_data <- rbind(m_area_data, exp_lm_data)
area_plot <- ggplot(m_area_data, aes(x = Specimen.date, y = value, fill = variable, color = variable))+
  geom_bar(data = subset(m_area_data, variable == "Daily.lab.confirmed.cases"),
    stat = "identity") +
  geom_line(data = subset(m_area_data, variable != "Daily.lab.confirmed.cases")) +
  labs(x="Specimen Date", y="Number of Confirmed Cases",
    fill = "", color = "") +
  scale_fill_manual(values = c("#ff0000","#05d153","#cad105","#000000"),
    labels = c(sprintf("%s # Daily Confirmed cases",area_name),
      "fitted","predicted","7 day average")) +
  scale_color_manual(values = c("#ff0000","#05d153","#cad105","#000000"),
    labels = c(sprintf("%s # Daily Confirmed cases",area_name),
      "fitted","predicted","7 day average")) +
  scale_x_date(date_breaks = "4 weeks", date_labels = "%Y-%m-%d") +
  theme_bw() %+replace% theme(legend.position = "top",
    legend.justification = "left")
```