Algebra\_COVID\_plots

Kevin Little

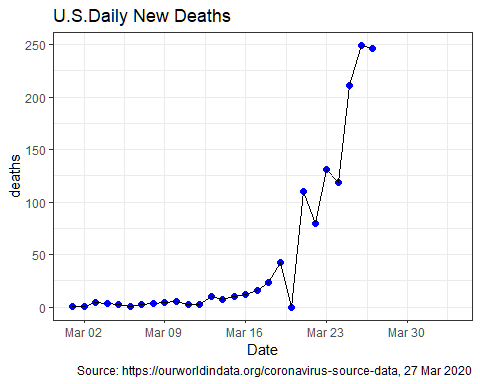
3/27/2020

## Table and Plot of U.S. Deaths

knitr::kable(data\_use)

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| dateRep | day | month | year | cases | deaths | countriesAndTerritories | geoId | countryterritoryCode | popData2018 | deaths\_nudge | log\_count\_deaths | serial\_day |
| 2020-03-01 | 1 | 3 | 2020 | 3 | 1 | United\_States\_of\_America | US | USA | 327167434 | 1 | 0.0000000 | 1 |
| 2020-03-02 | 2 | 3 | 2020 | 20 | 1 | United\_States\_of\_America | US | USA | 327167434 | 1 | 0.0000000 | 2 |
| 2020-03-03 | 3 | 3 | 2020 | 14 | 4 | United\_States\_of\_America | US | USA | 327167434 | 4 | 0.6020600 | 3 |
| 2020-03-04 | 4 | 3 | 2020 | 22 | 3 | United\_States\_of\_America | US | USA | 327167434 | 3 | 0.4771213 | 4 |
| 2020-03-05 | 5 | 3 | 2020 | 34 | 2 | United\_States\_of\_America | US | USA | 327167434 | 2 | 0.3010300 | 5 |
| 2020-03-06 | 6 | 3 | 2020 | 74 | 1 | United\_States\_of\_America | US | USA | 327167434 | 1 | 0.0000000 | 6 |
| 2020-03-07 | 7 | 3 | 2020 | 105 | 2 | United\_States\_of\_America | US | USA | 327167434 | 2 | 0.3010300 | 7 |
| 2020-03-08 | 8 | 3 | 2020 | 95 | 3 | United\_States\_of\_America | US | USA | 327167434 | 3 | 0.4771213 | 8 |
| 2020-03-09 | 9 | 3 | 2020 | 121 | 4 | United\_States\_of\_America | US | USA | 327167434 | 4 | 0.6020600 | 9 |
| 2020-03-10 | 10 | 3 | 2020 | 200 | 5 | United\_States\_of\_America | US | USA | 327167434 | 5 | 0.6989700 | 10 |
| 2020-03-11 | 11 | 3 | 2020 | 271 | 2 | United\_States\_of\_America | US | USA | 327167434 | 2 | 0.3010300 | 11 |
| 2020-03-12 | 12 | 3 | 2020 | 287 | 2 | United\_States\_of\_America | US | USA | 327167434 | 2 | 0.3010300 | 12 |
| 2020-03-13 | 13 | 3 | 2020 | 351 | 10 | United\_States\_of\_America | US | USA | 327167434 | 10 | 1.0000000 | 13 |
| 2020-03-14 | 14 | 3 | 2020 | 511 | 7 | United\_States\_of\_America | US | USA | 327167434 | 7 | 0.8450980 | 14 |
| 2020-03-15 | 15 | 3 | 2020 | 777 | 10 | United\_States\_of\_America | US | USA | 327167434 | 10 | 1.0000000 | 15 |
| 2020-03-16 | 16 | 3 | 2020 | 823 | 12 | United\_States\_of\_America | US | USA | 327167434 | 12 | 1.0791812 | 16 |
| 2020-03-17 | 17 | 3 | 2020 | 887 | 16 | United\_States\_of\_America | US | USA | 327167434 | 16 | 1.2041200 | 17 |
| 2020-03-18 | 18 | 3 | 2020 | 1766 | 23 | United\_States\_of\_America | US | USA | 327167434 | 23 | 1.3617278 | 18 |
| 2020-03-19 | 19 | 3 | 2020 | 2988 | 42 | United\_States\_of\_America | US | USA | 327167434 | 42 | 1.6232493 | 19 |
| 2020-03-20 | 20 | 3 | 2020 | 4835 | 0 | United\_States\_of\_America | US | USA | 327167434 | NA | NA | 20 |
| 2020-03-21 | 21 | 3 | 2020 | 5374 | 110 | United\_States\_of\_America | US | USA | 327167434 | 110 | 2.0413927 | 21 |
| 2020-03-22 | 22 | 3 | 2020 | 7123 | 80 | United\_States\_of\_America | US | USA | 327167434 | 80 | 1.9030900 | 22 |
| 2020-03-23 | 23 | 3 | 2020 | 8459 | 131 | United\_States\_of\_America | US | USA | 327167434 | 131 | 2.1172713 | 23 |
| 2020-03-24 | 24 | 3 | 2020 | 11236 | 119 | United\_States\_of\_America | US | USA | 327167434 | 119 | 2.0755470 | 24 |
| 2020-03-25 | 25 | 3 | 2020 | 8789 | 211 | United\_States\_of\_America | US | USA | 327167434 | 211 | 2.3242825 | 25 |
| 2020-03-26 | 26 | 3 | 2020 | 13963 | 249 | United\_States\_of\_America | US | USA | 327167434 | 249 | 2.3961993 | 26 |
| 2020-03-27 | 27 | 3 | 2020 | 16797 | 246 | United\_States\_of\_America | US | USA | 327167434 | 246 | 2.3909351 | 27 |

#extend data table by one week  
buffer <- 7  
  
p0 <- ggplot(data=data\_use,aes(x=dateRep,y=deaths))+  
 theme\_bw()+  
 geom\_point(size=rel(2.0),colour="blue")+  
 geom\_line()+  
 labs(title="U.S.Daily New Deaths", caption="Source: https://ourworldindata.org/coronavirus-source-data, 27 Mar 2020")+  
 xlab("Date")+  
 xlim(min(data\_use$dateRep),max(data\_use$dateRep)+buffer)  
  
print(p0)



## Linear Fit

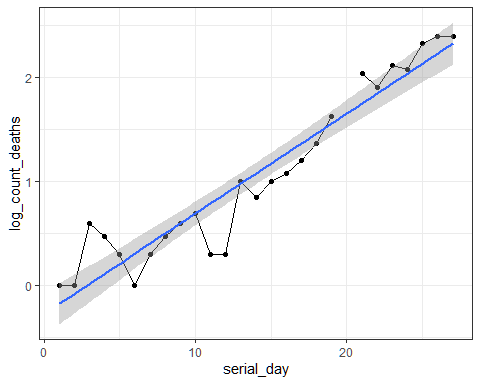
This is the linear fit

lm\_out <- lm(data=data\_use,data\_use$log\_count\_deaths ~ data\_use$serial\_day)  
#check plot  
p1<- ggplot(data=data\_use,aes(x=serial\_day,y=log\_count\_deaths))+  
 theme\_bw()+  
 geom\_point()+  
 geom\_line()  
   
p2 <- p1+ geom\_smooth(method=lm)  
print(p2)

## `geom\_smooth()` using formula 'y ~ x'

## Warning: Removed 1 rows containing non-finite values (stat\_smooth).

## Warning: Removed 1 rows containing missing values (geom\_point).



#intercept  
lm\_out$coefficients[1]

## (Intercept)   
## -0.2700875

#slope  
lm\_out$coefficients[2]

## data\_use$serial\_day   
## 0.09621738

## Exponential limits

cchart\_df <- data.frame(data\_use[!is.na(data\_use$log\_count\_deaths),c("dateRep","serial\_day")],  
 lm\_out$residuals,c(NA,diff(lm\_out$residuals)),lm\_out$fitted.values)  
  
  
  
AvgMR <- mean(abs(cchart\_df$lm\_out.residuals))  
cchart\_df$UCL <- lm\_out$fitted.values+2.66\*mean(AvgMR,na.rm=TRUE)  
cchart\_df$LCL <- lm\_out$fitted.values-2.66\*mean(AvgMR,na.rm=TRUE)  
  
#buffer with buffer days beyond max date  
buffer\_serial\_day <- seq(from=max(cchart\_df$serial\_day)+1,to=max(cchart\_df$serial\_day)+buffer,by=1)  
predicted\_value <- lm\_out$coefficients[1]+ lm\_out$coefficients[2]\*buffer\_serial\_day  
buffer\_dates <- seq.Date(from=max(cchart\_df$dateRep)+1,to=max(cchart\_df$dateRep)+buffer,by="day")  
buffer\_df <- cbind.data.frame(buffer\_dates,  
 buffer\_serial\_day,  
 rep(NA,buffer),  
 rep(NA,buffer),  
 predicted\_value,  
 predicted\_value +2.66\*mean(AvgMR,na.rm=TRUE),  
 predicted\_value - 2.66\*mean(AvgMR,na.rm=TRUE))  
  
names(buffer\_df) <- names(cchart\_df)  
  
  
df\_out <- rbind(cchart\_df,buffer\_df)  
  
df\_out$predict <- 10^df\_out$lm\_out.fitted.values  
df\_out$UCL\_anti\_log <- 10^df\_out$UCL  
df\_out$LCL\_anti\_log <- 10^df\_out$LCL  
  
p3 <- p0 + geom\_line(data=df\_out,aes(x=dateRep,y=predict),linetype="solid",colour="red")+  
 geom\_line(data=df\_out,aes(x=dateRep,y=UCL\_anti\_log),linetype="dotted")+  
 geom\_line(data=df\_out,aes(x=dateRep,y=LCL\_anti\_log),linetype="dotted")  
  
print(p3)

