Sohaib Syed
CS422
9/26/21
Dr. Gurbani

Homework 3

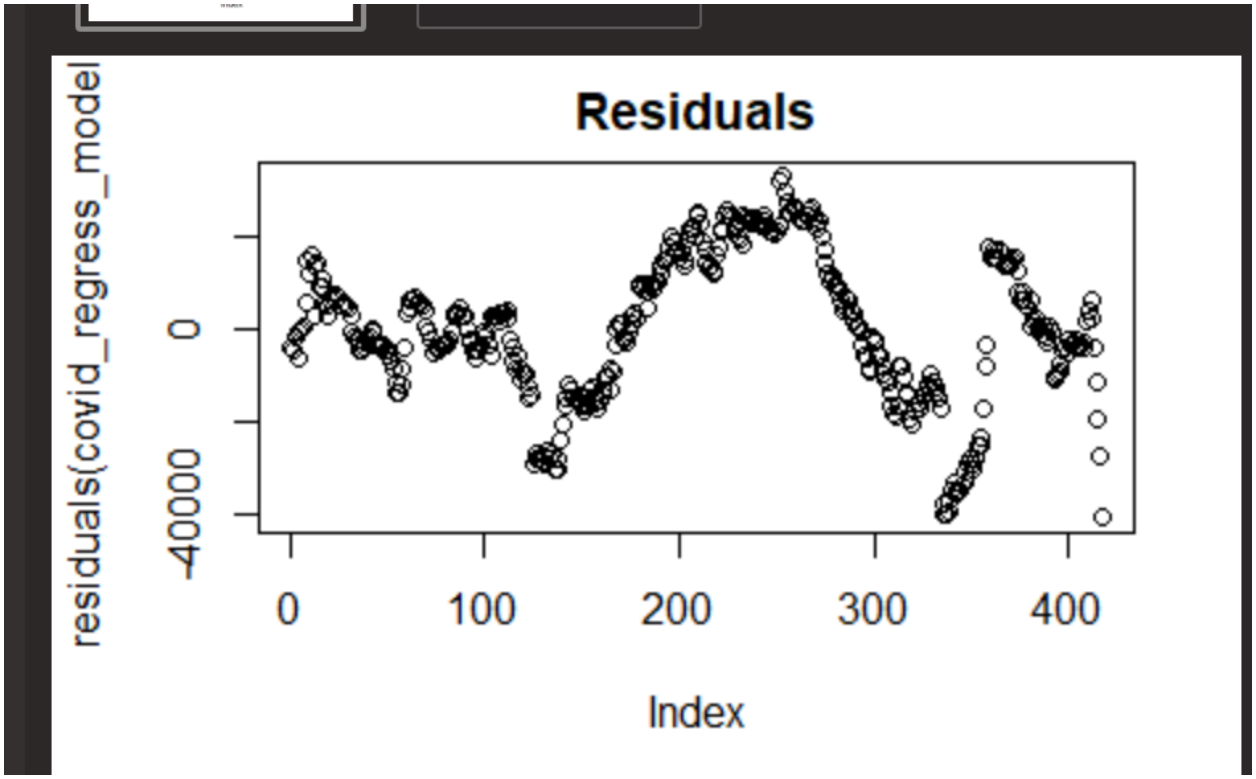
Exercises

1.1

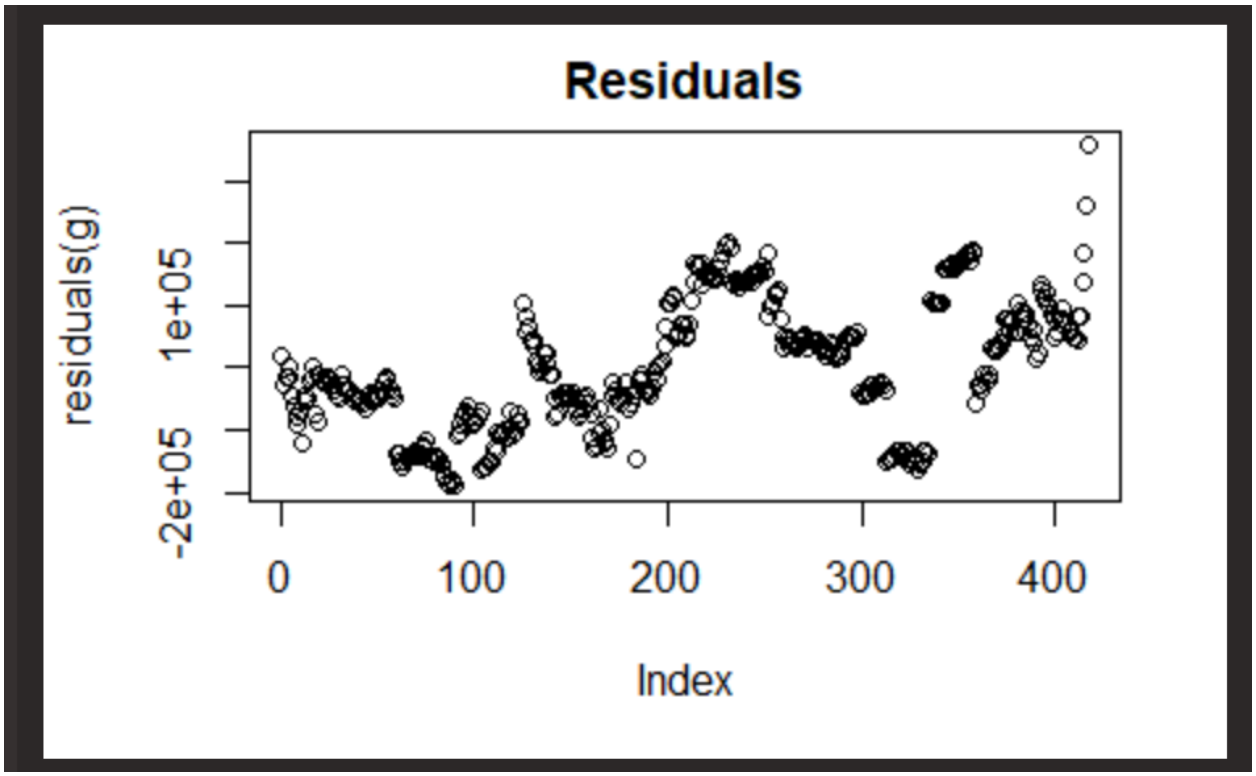
Given that the least squares line equation is: $y_i = \widehat{\beta_0} + \widehat{\beta_1} x_i$, what we can do is replace $\widehat{\beta_0}$ and end up with $y = \bar{y} - \beta_1(x - \bar{x})$. Then when we plug in $\bar{x}$, we get y=$\bar{y}$. This means that the least square line has to pass through ($\bar{x}$,$\bar{y}$).

1.2

I went back to the models and plotted the residuals and i got:



For part c



For part g

Both of these plots show that the residuals are not distributed well at all. They should be pretty much straight along the 0 line, but instead we get a large range in both models. I also got very large RSS values for both models. RSS for c:1 *10^11. RSS for g: 4.23*10^12. While c had a lower value, it still is ridiculously large and would not be a good prediction model