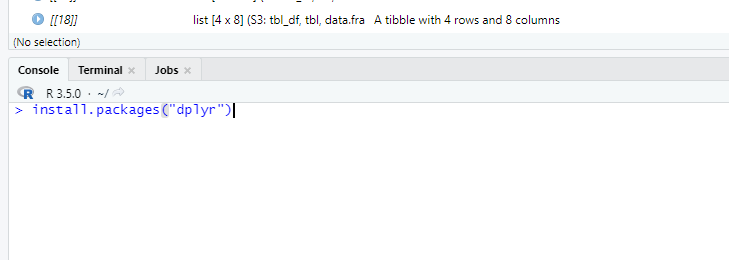
**Assignment 4: Copy outputs of all the examples/exercises provided in the Unit 2 Session 1 (Session 6) and Unit 2 Session 2 (Session 7) slides and interpret/discuss them carefully.**

Session 2 ( unit 2)

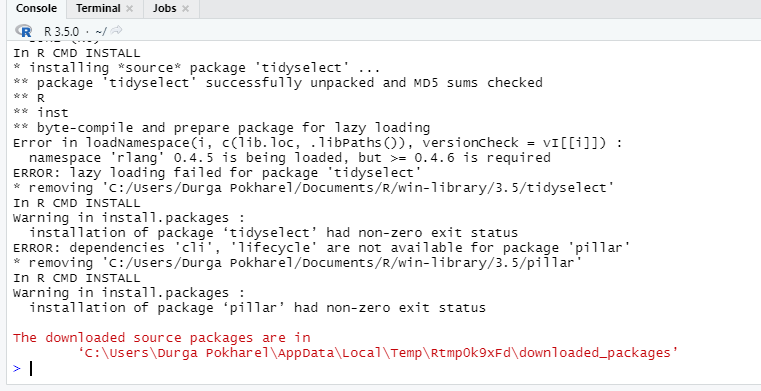
Solution:

1. Install a very popular R package in Rstudio.
2. dplyr b. Rcmdr

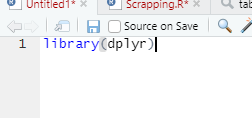
While installing these packages I used following process



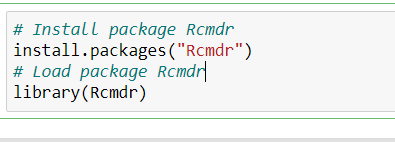
I typed this in a R Console after run this I got following



To load the package I type

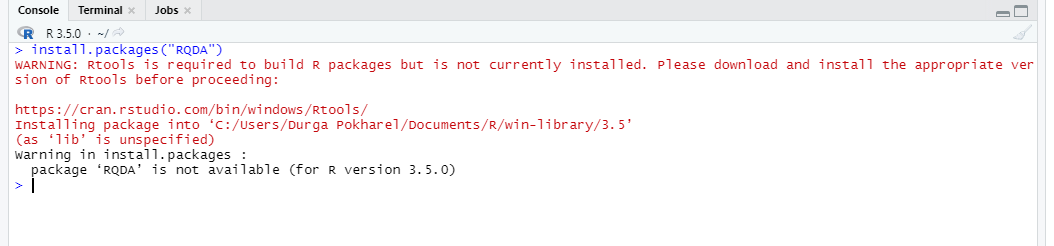


Similarly, for Rcmdr same process can follow

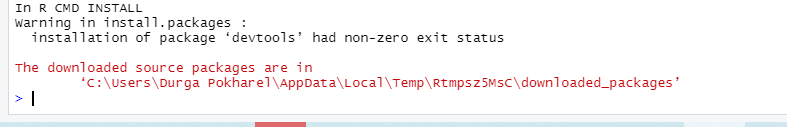


1. Install package RQDA

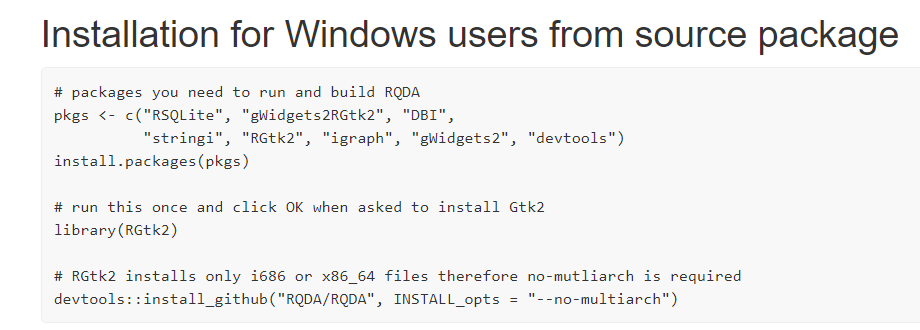
The packages RQDA is not directly install in Rstudio because it is not available for current version. So, we need to kept it from somewhere.

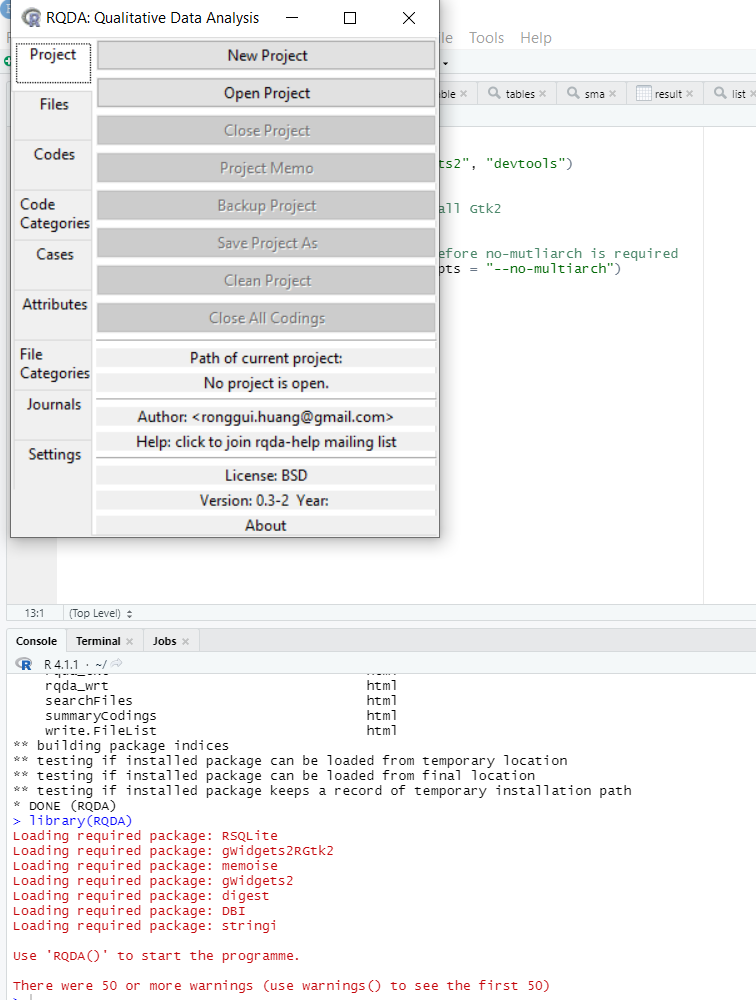


We pull it from git.

Before pull from git we should install ‘devtools’ packages by following same process as in above. 

We shoud through this link 





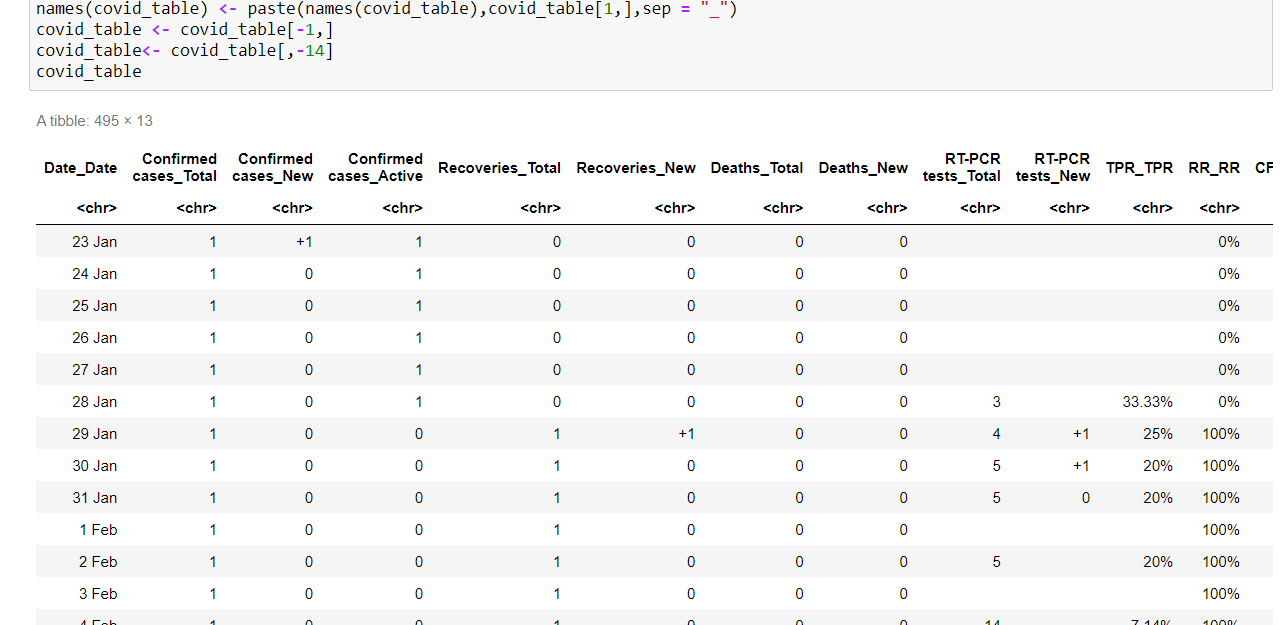
**Session 7**

**Web scrapping**

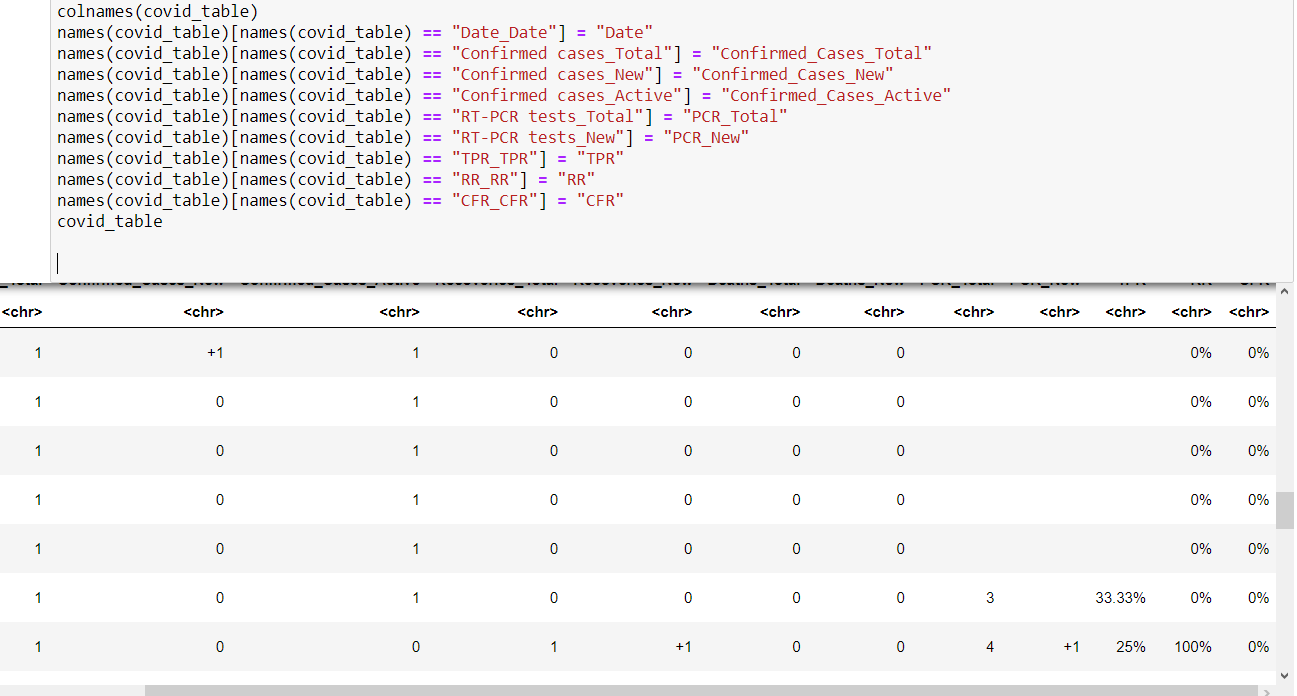
For this we load the packages `rvest` and `dplyr` .We provide wiki\_link . Read wiki link and then we gave instruction to read table from this wiki\_link . There are 40 tables out of this we gave instruction to pull 16 number’s first table. Which is given below.



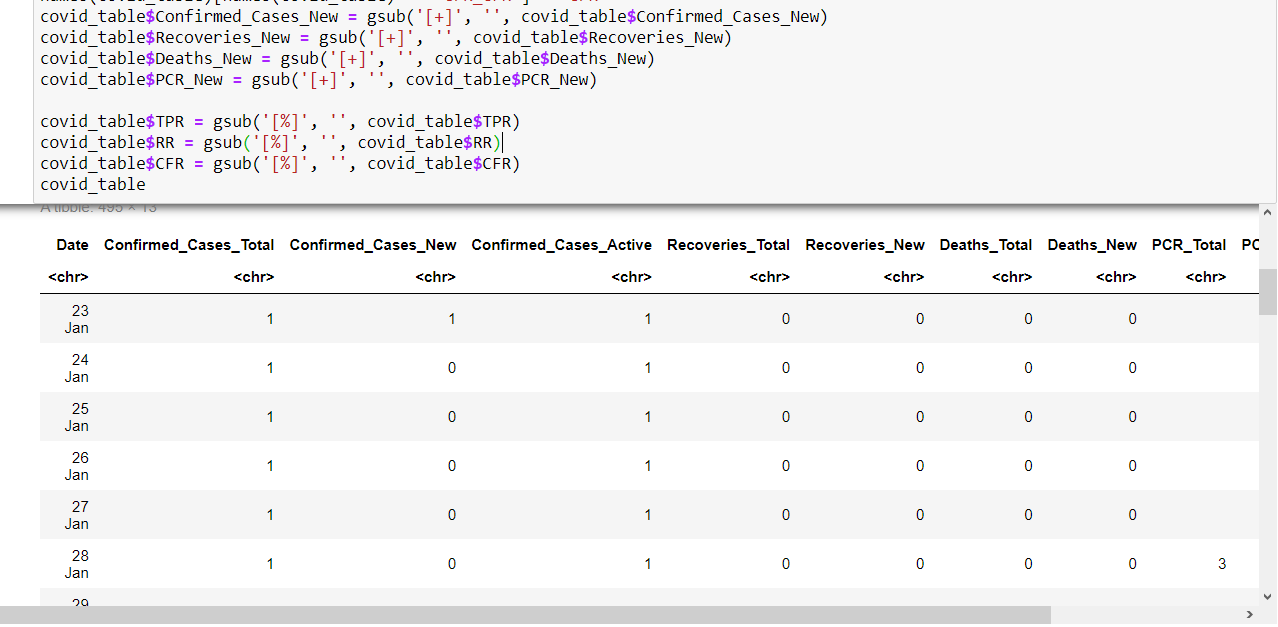
We first edit the table in our desire format. At first we add underscore in place of whitespace in first row. Also, we don’t need first row and 14 Column so we delete them.

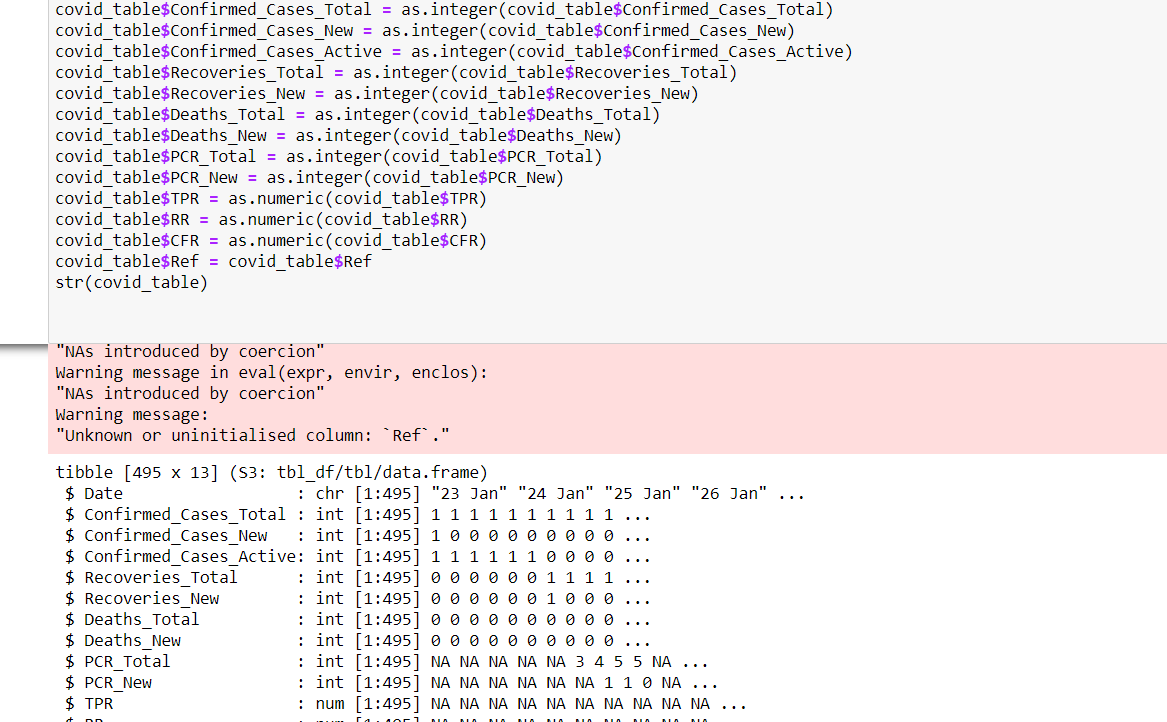


Then we rename our column by following process.

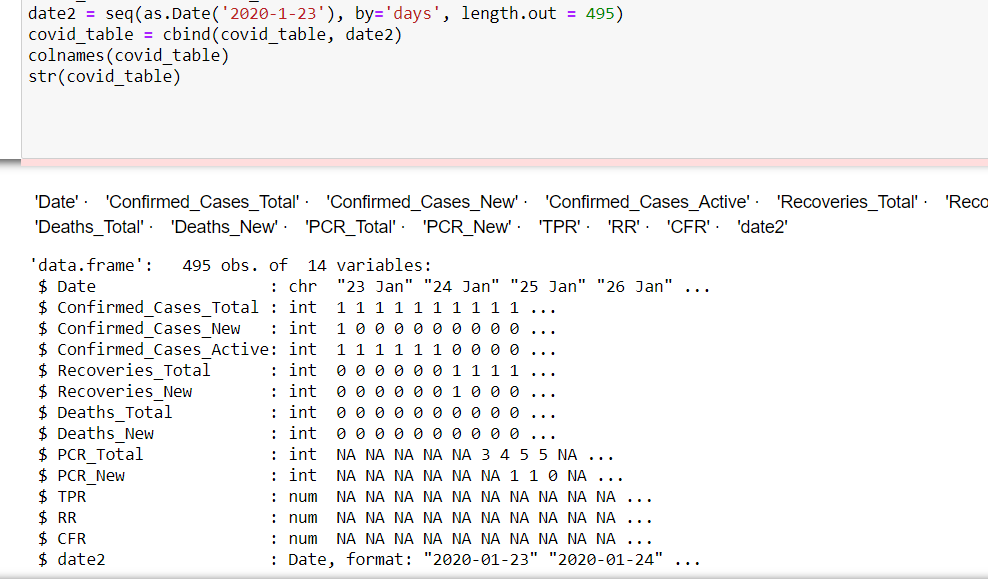


In our table in particular Columns there are not necessary symbol like +, % so we remove there by,

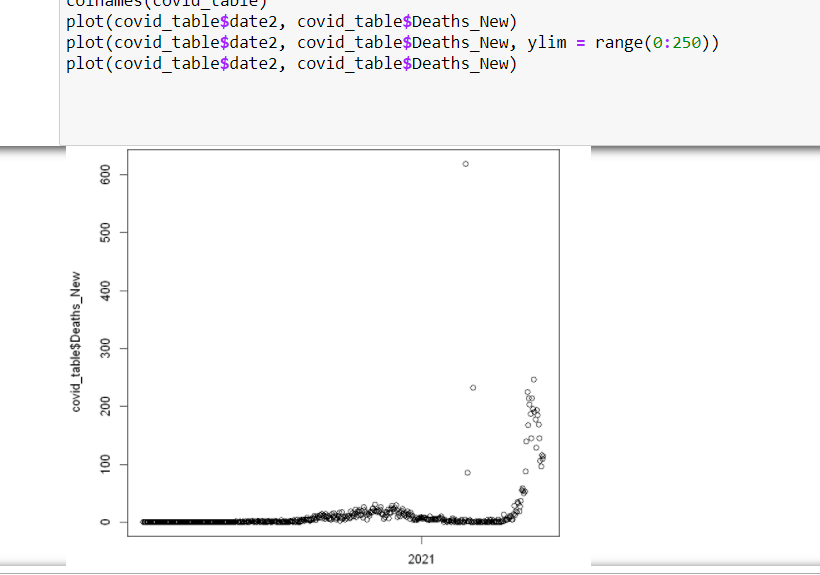


In our table all the data are in char format so we need to convert in required data format. We do following for this,

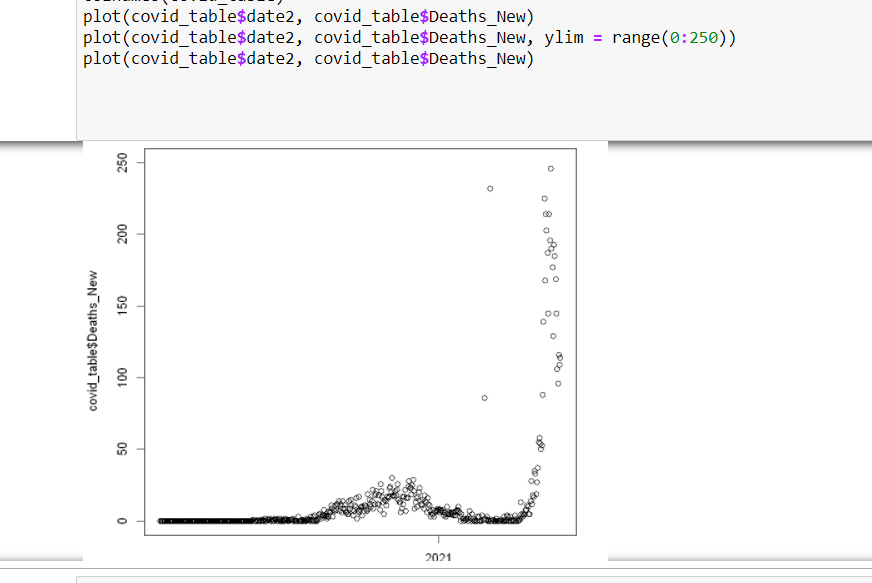
Here our date still not in date format to change in date format we do following step.



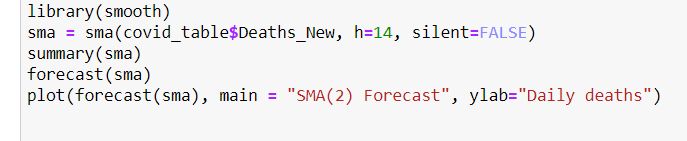
Now date2 in date format. We combine date2 with Covid\_table. Now, we plot date2 vs New\_ Deaths.



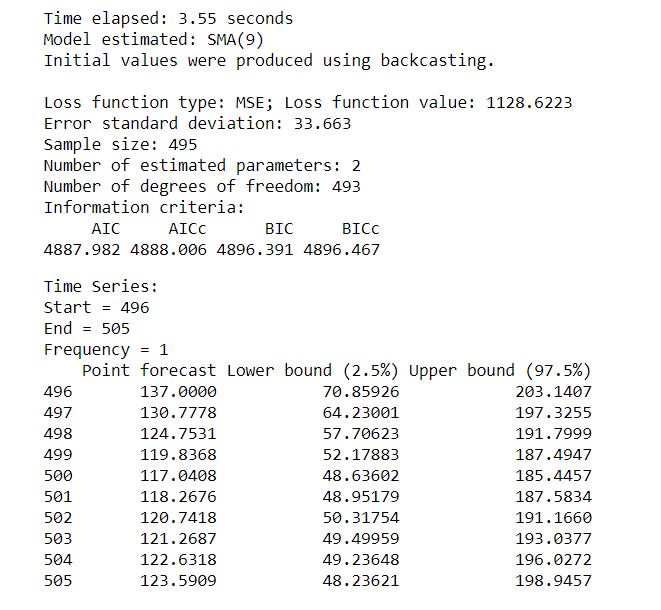
By typing range(0: 250) for y. It scale y from 0 to 250. From figure we can see that initially ratio of new deaths were constant for some months then slowly increase and then also decrease at last times new deaths increase rapidly. We see there are some outliers in figure because at the time of data collections some other day’s information also added at these particular days.

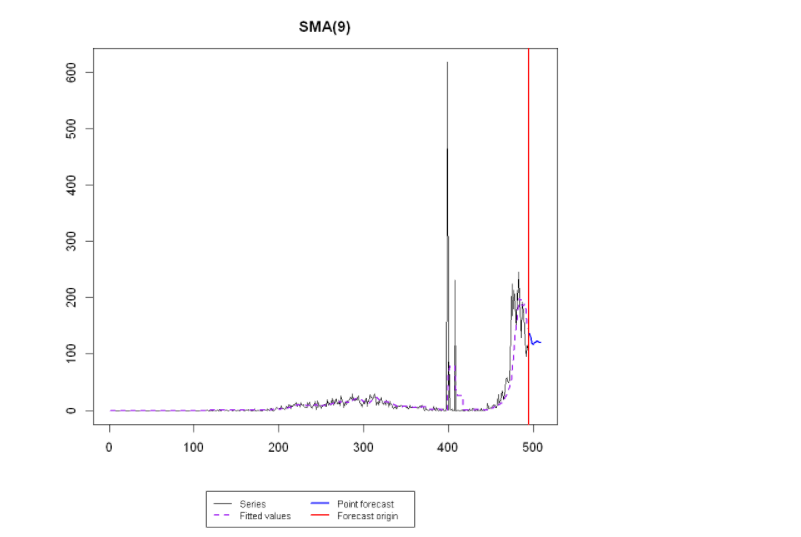


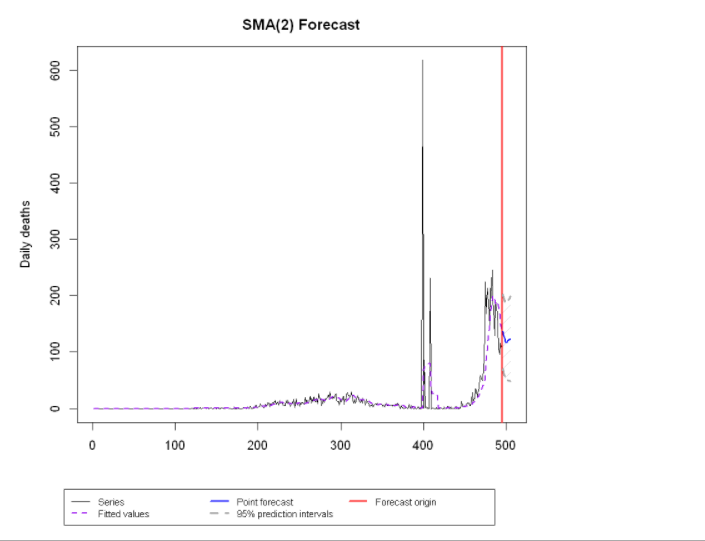
To calculate SMA, at first we need to install smooth package.



To calculate SMA at first load smooth package. And summary of SMA is given below. Here forecast value is in between lower bound and upper bound. So, accurate.





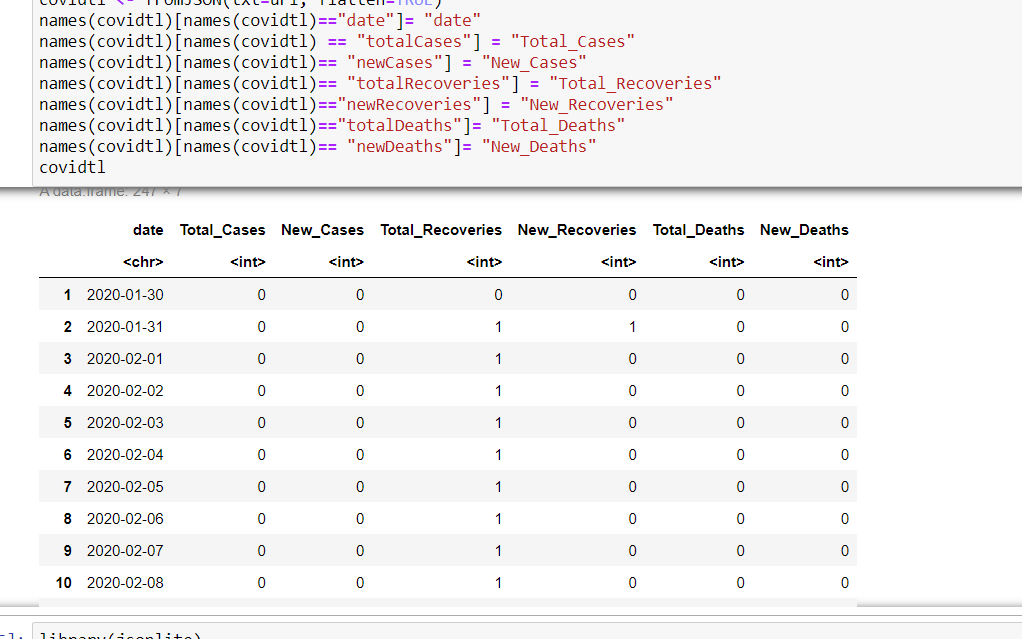


**Work from jsonlite**

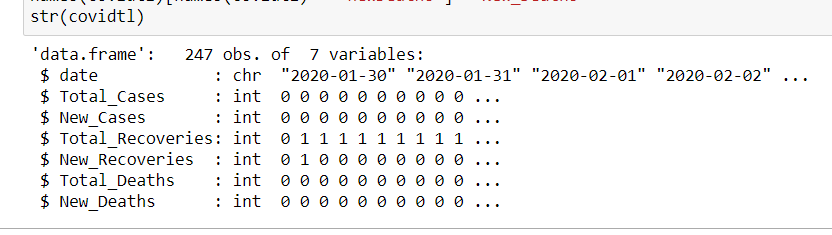
For this at first, we should install packages `jsonlite` and load the packages. To read the link we can see the following



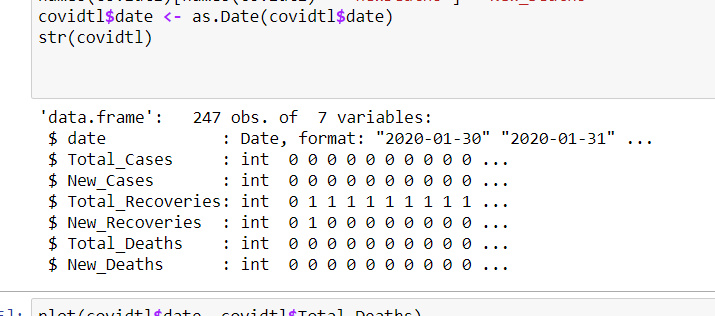
Here I rename the Columns name by,



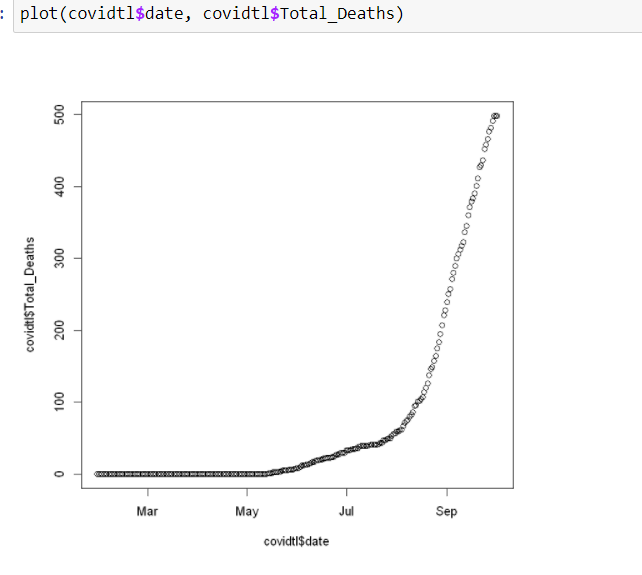
If we check the data type of each Columns, then we get



We require date column in date format we can change by,



Plot of date vs Total\_Deaths,



From graph we can see in beginning time less people were died later it increase slowly, and then at last time deaths rate increase rapidely.