Analysing Covid Data From India

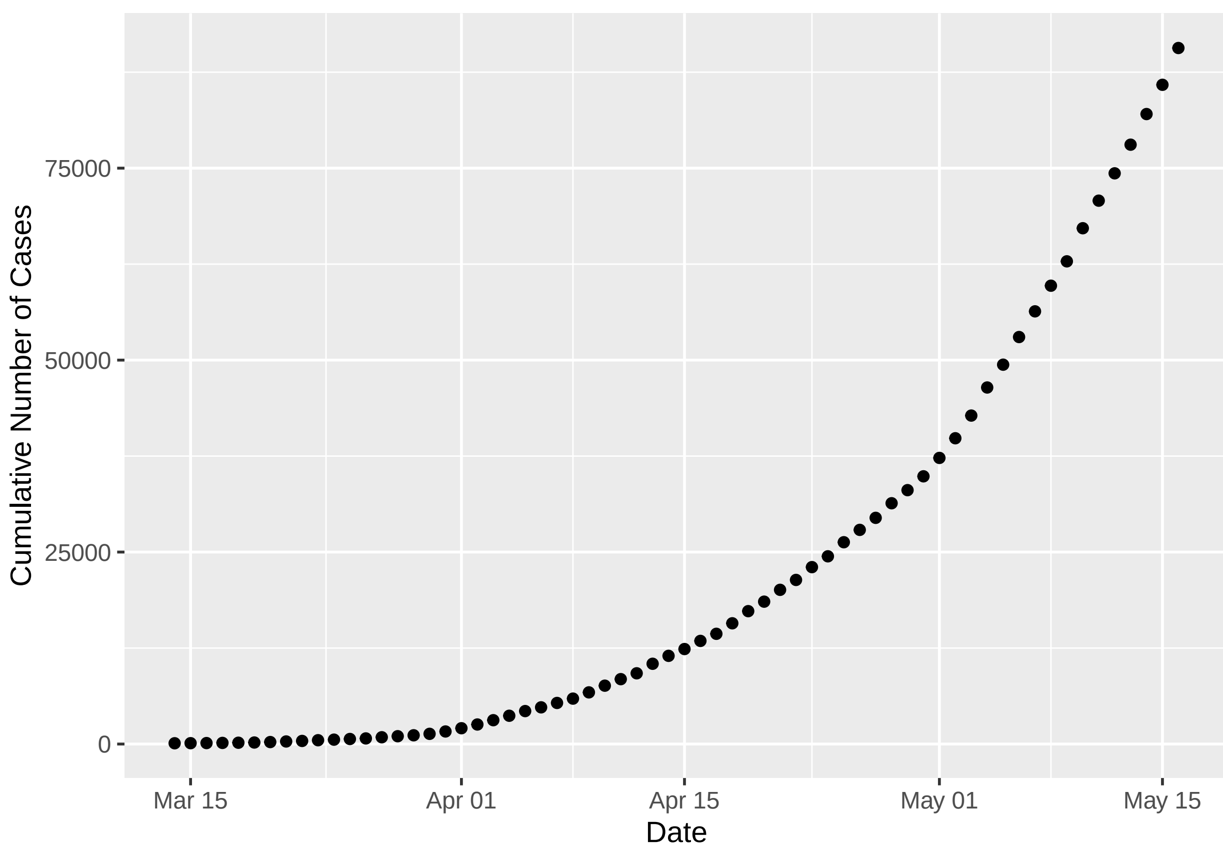
Vimal Simha

Overview

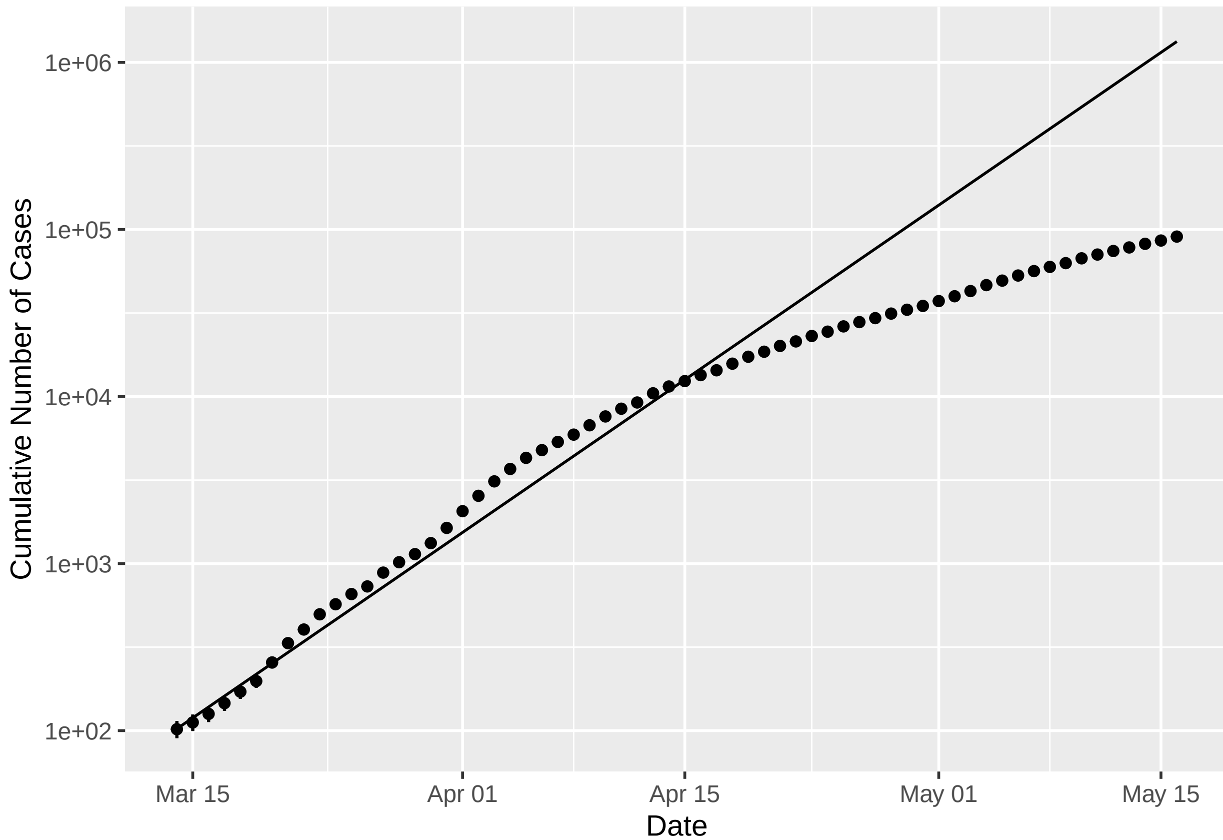
I examine some basic aspects of the Covid Data from India. My data are sourced from https://www.covid19india.org.

Exploratory Analysis

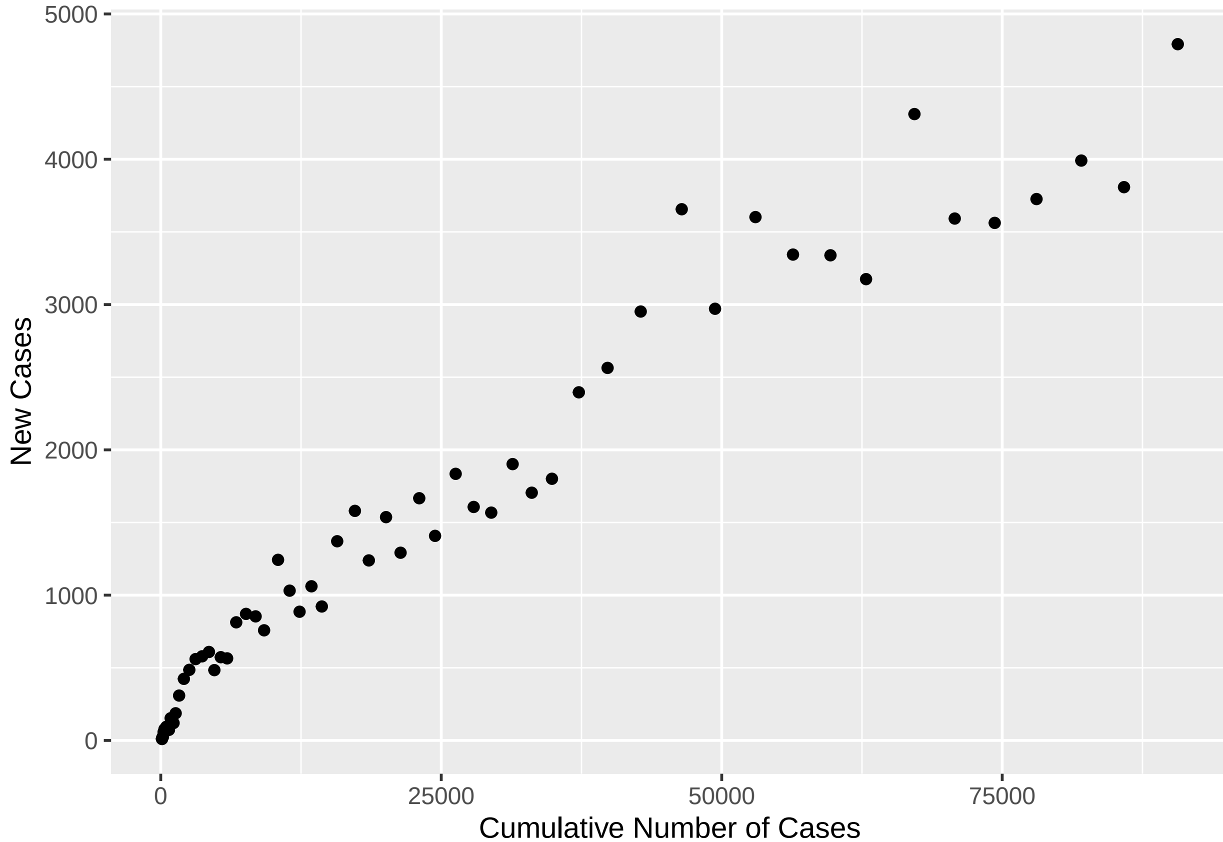
The figure shows the cumulative number of cases against date, starting from the day when there were 100 cases. It is the starting point of any analysis. It shows the number of cases rising roughly exponentially so far.



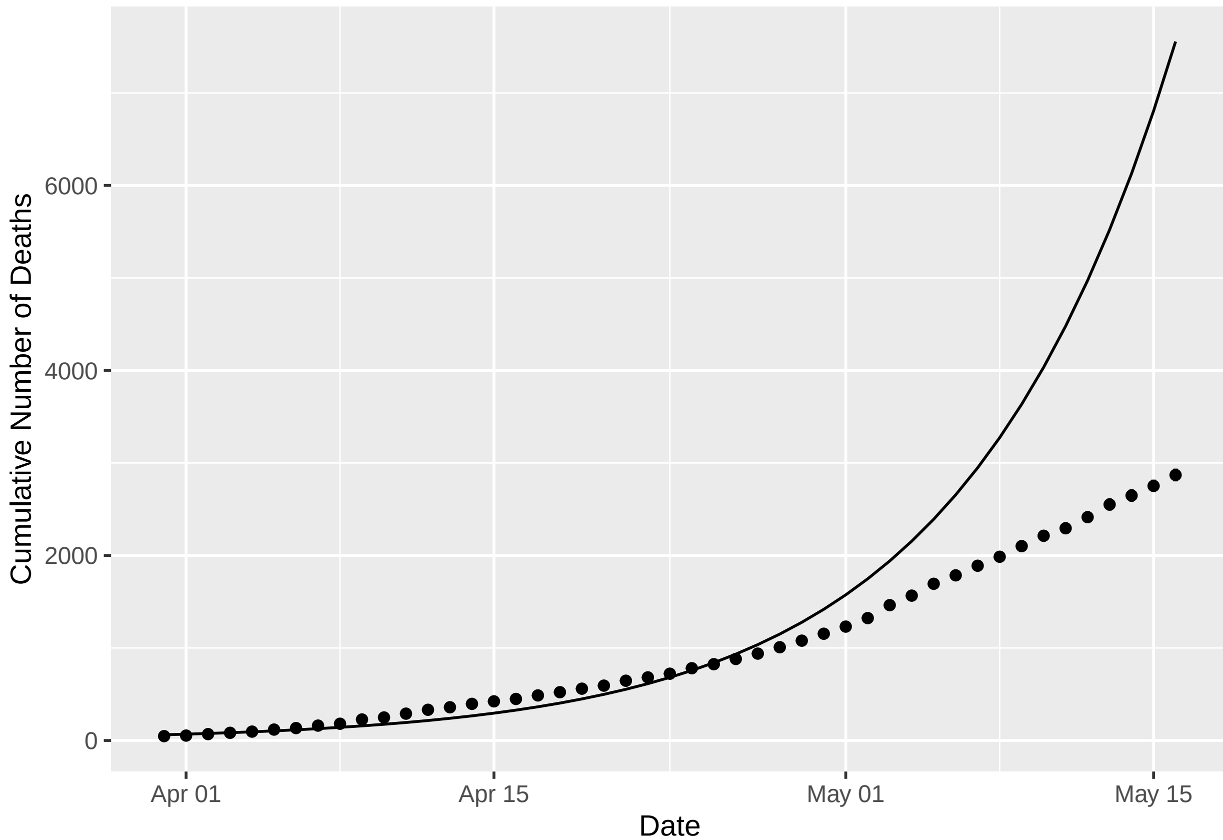
A key parameter of interest is the doubling period i.e. how many days does it take for the number of cases to double? There are a few equivalent ways of doing that. One is to fit an exponential curve to the figure above. To help visualise this better, the number of cases is plotted on a log axis which makes our exponential curve look like a straight line with its slope giving the reproduction rate (see appendix for details). The best-fit exponential curve is shown in the figure below and yields a doubling time of 4.1 days.



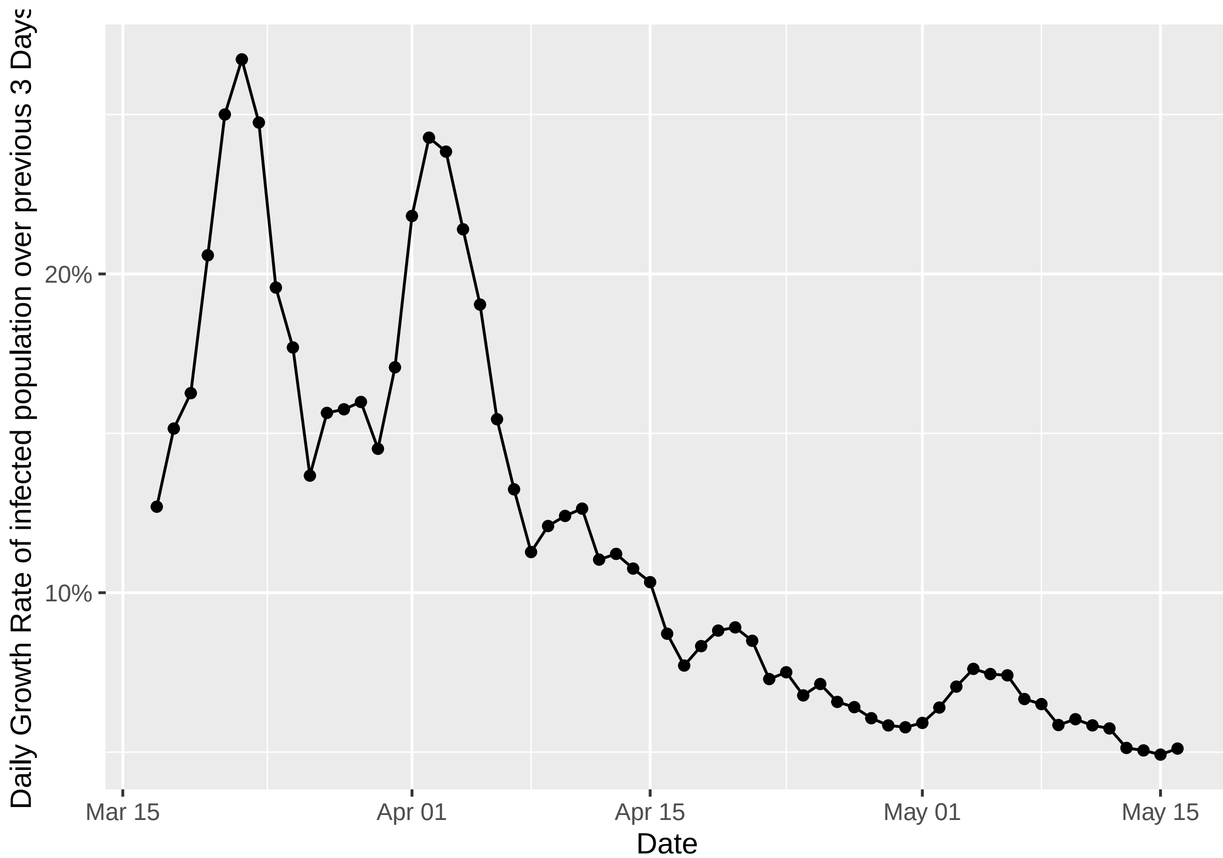
The figure below shows a plot of the number of new cases on a given day against the total number of cases on that day. It is roughly proportional, but with a lot of scatter.



We are also interested in how many people are dying, and how fast that is increasing. This is useful to know for its own sake, but besides that, it is also of interest because we might consider the data on deaths to be more accurate than the data on the number of cases. We plot the number of deaths against time starting from the time when we had at least 50 deaths. The data are well fit by an exponential with a doubling time of 4.4 days.



Another question of interest is whether the slope is changing i.e. is there evidence to suggest that the doubling rate is slowing down or speeding up? The figure below shows the daily growth rate over the last 3 day period i.e. we compute the number of new cases during the last 3 days and calculate the daily growth rate of the infected population that it implies.



The growth rate jumps around quite a bit, but there is no evidence of a transition from a high growth to a low growth regime or vice-versa. There is no reason to suppose that the growth rate of the infected population, or in other word the doubling rate, is changing.

Appendix: Computation of Doubling Rate from Reproduction Rate

Recall that, in the simplest models and during the initial stages in more complex models, the number of new cases ΔN is given by the reproduction rate, r multiplied by the Number of cases, N,

$$\Delta N = rN$$

and therefore the number of cases after t time periods, N is

$$N = N_0(1+r)^t$$

where N_0 is the initial number of cases.

If we take the logarithm of both sides

$$log(N) = log(N_0) + tlog(1+r).$$

The slope of the plot of $log(N)$ vs t is $log(1+r)$, and the doubling period d is given by:

$$d = log(2)/log(1+r).$$

Notice that these are the same formulae used in carbon dating, compound interest and many other things.