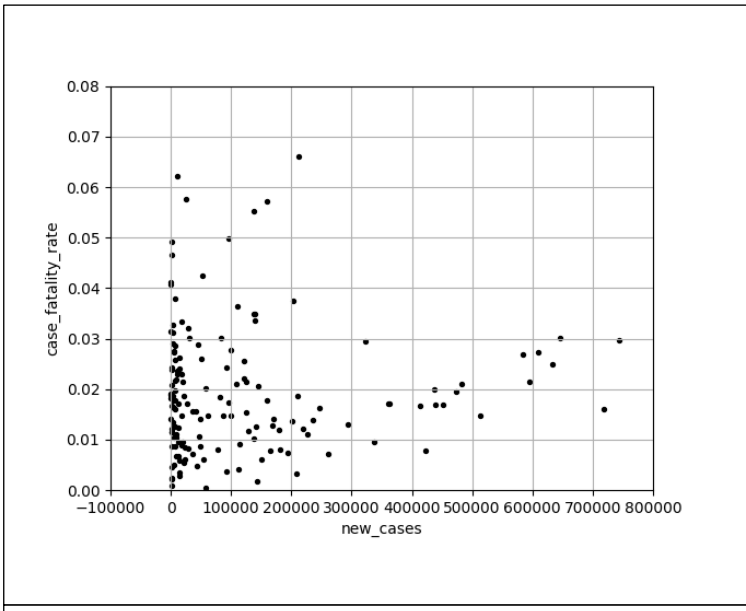
Investigation on COVID 2020

Introduction

This investigation is based on COVID cases in 2020, the raw data set is from covid.ourworldindata.org The data set shows the daily total cases, total deaths, new cases and new deaths from countries over the world. This data clearly shows the number of cases and deaths each day start from 2020 Feb, and all data is sorted based on location and date. However, the data set is also loss many information, for example death rate, positive rate, vaccinations etc. Therefore, in these investigation, we will analyze some of the data, and present it as a graph.

Graph A:

Graph below is data analyze from row data x-axis is new cases in year 2020, y-axis is case fatality rate in this year, each point represent a country. New cases which is equal to total cases in 2020, which can be taken from data set by selecting the max total cases in a country in 2020. Case fatality rate is calculated by the ratio of total deaths with total cases.

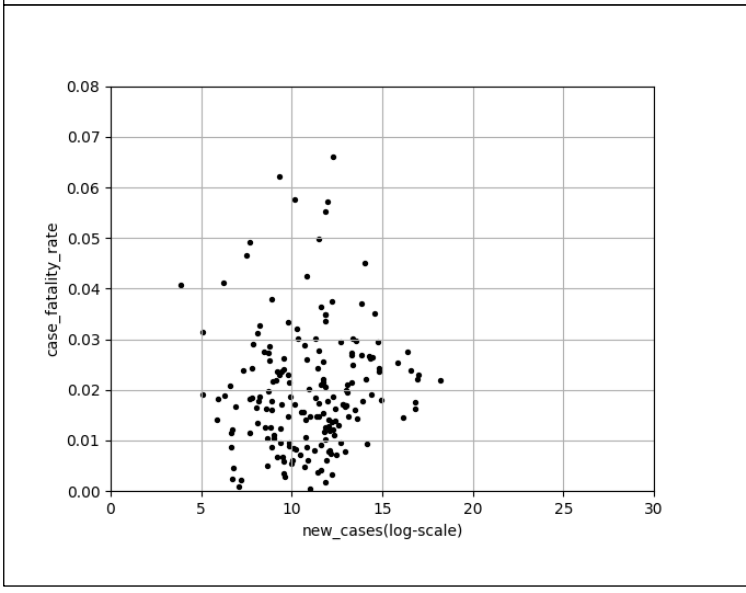


From this graph we can see most of the Country will have case fatality rate between 0% to 3%, which is not very high. And most Country have 0 to 50000 new cases in 2020.

Graph B:

Graph B have same data with only x-axis changed to log-scale, x-axis is just simply applying log to every new cases.

From this data, we can have a more clear view on the case fatality rate, which shows they all concentrated at 0 to 3%.



Evaluation:

From these two graph, we both can get same Information which is case fatality rate is at 0 to 3% since the point is most concentrated in this area. However, graph B can have a more clearly view on all data set, since the number of new cases in 2020 have a large range cross all countries, therefore, log-scale can let us more focus on case fatality rate and get a more clear view on it.