Group: WH P2 lab
Group member: Ziming Wang zw2338
Jianzhi Wang jw6026

Since the outbreak of COVID-19, it has spread rapidly and spread to all over the world. Our group wants to find out the difference and similarities between COVID-19 and other viruses in the past, such as Ebola, SARS, and MERS. We want to know what caused the virus to spread so quickly, and we also want to understand how COVID-19 is different from other viruses. So when we encounter new viruses in the future, we hope that we can use more appropriate strategy to eliminate these viruses and save more lives

Through preliminary understanding, we can understand that many viral viruses have different outbreak areas and that their fatality rates are also different. By observing the current COVID-19 data, we can use the COVID-19 data to compare with other viruses, and we believe they must have a lot of information that can be compared. Our plan for this project is to analyze these data to find more details. By studying the experiences we had in the past, we will be able to know more about the viruses.

Our initial plan is to consider the speed of spread, the area of spread, the fatality rate and the duration of the past epidemics. Through further research, we might add some more specific factors, such as race and virus type. We are planning to use the spark SQL functions to help us analyze the data and get our deserved result.

datasets:
covid-19 compare to SARS, ARDS, Pneumocystis, Streptococcus…
https://github.com/ieee8023/covid-chestxray-dataset/blob/master/metadata.csv

covid-19 increasing rate (JHU)
https://github.com/datasets/covid-19/tree/master/data

Ebola dataset 2014-2016
https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset

Ebola, sars, mers, H1N1
https://www.kaggle.com/imdevskp/covid19-vs-sars-vs-mers-vs-ebola-vs-h1n1