Capstone Proposal:

In this pandemic, it could be a great help to understand the exact reason and why it is getting severe in some group and not in other groups. After Observing the reason and factors, causing death in COVID patients basing on weekly data from CDC.gov official website.

By analyzing, the time series data from CDC, which describes several different criteria’s Contributing to COVID-19 deaths and in those state and Age relevant dataset is one criteria, filtering needful columns and replacing null/Nan values with 0 and doing group by of columns. trends will be identified using Exploratory Data Analysis.

Linear Regression / Logistic Regression will be performed to understand the significance of relationship among dependent and independent variables. Hypothesis test will be performed and basing on the significance value it will be determined what variables will be considerable or not. For machine Learning modelling, cleansed data will be split in training dataset and testing dataset. Which will eventually be helpful to predict the trend based upon factors impacting death due to COVID-19.

Project report will be created after the analysis, evaluation, and verification of the model accuracy. The final deck will have all key information about how to identify a severe case of COVID-19.