Capstone 2: Project Proposal

Context:
COVID-19 is a corona virus that originated in Wuhan, China, during the early fall of 2019. By spring of 2020 COVID-19 was a full-blown global pandemic and two years after its emergence, COVID-19 is still a threatening and global disease. From the very beginning of COVID-19's emergence, information and decisions surrounding the disease have been prone to politics, misinformation, and heavy media scrutiny.

Criteria for success:
The aim of this study is to examine the relations between COVID-19 deaths, patient demographics, and other variables. Another focus would be to examine the effect of lockdowns, mandates, and vaccinations on the count of new local cases. To be deemed successful, this study will be completed within a timely manner and provide new insight into COVID-19 cases in the US.

Scope of solution space:
The dataset contains information regarding patient demographics, geography, case/patient conditions, and other variables, but a large number of values are missing. A patient's characteristics (age/sex/ethnicity/health conditions) are correlated to their outcome following a positive COVID diagnosis. Their geographic location, vaccination status, and treatment are also likely correlative.

Constraints:
The available dataset has yes/no values for columns indicating case exposure and underlying conditions. It would be more useful if these parameters were more descriptive. As already mentioned a large portion of the dataset's values are missing. This foreseeably could constrain the study's findings.

Stakeholders:
This study has no true stakeholder but is of personal interest. However, this study is likely interesting to any person enduring through this global catastrophe.

Data sources:
The targeted data stems from the US Center for Disease Control and Prevention website.