milestone_1

June 12, 2022

1 Milestone 1

Joshua Burden DSC530 Week 2 Bellevue University Instructor: Catherine Williams Date: 06/12/2022

These data sets will have related data for covid 19 statistics for both the USA but also the world. We are looking for relationships found within the data as the data comes from the covid API project, a Kaggle report, and data from rapidAPI that is supported but is now sun-set by the authors. - https://rapidapi.com/axisbits-axisbits-default/api/covid-19-statistics/ > Contains an api that can be queried that shows total reports based on providences. - https://data.cdc.gov/resource/vbim-akqf.csv > contains sets of information that has demographics, exposure history, disease severity indicators and outcomes, presence of any underlying medical conditions and risk behaviors, and no geographic data. Comes directly from the CDC. - https://www.kaggle.com/datasets/imdevskp/corona-virus-report?select=country_wise_latest.csv#:~:text=calendar_view_week-,country_wise_latest,-.csv >Contains several datasets, linking day to day country wise number of cases, using the fully_grouped.csv file.

All of these datasets are related to case totals across the world and bounded by countries. I plan on looking this data deciphering whether this data has relationships that can be concluded based on geographical latitude and longitude or region. the data has a relationship based on publicly available patient level data sets that should have some commonality with each other. Some data has been sun set as of 2021. This might make the datasets difficult to work with as some data might not sync up correctly with the rest, and there would be no way of verifying if the data is correct or not.

The data itself might display causes of death or medical complications that are more of an anecdotal relationship rather than something that is completely conclusive. I think that if the data is either inconclusive, or anecdotal, you will not have a clear picture of the data once it is visualized, or the data could be misconstrued or misinterpreted to infer something that it is not. I will attempt to do my best not to allow the data to manipulate or create a narrative that could be swung in a direction for maleficence. The attempt of this project is to look at the data that is available, and to visualize the data as it presents itself, and to not infer a direct conclusive expectation or even question until after the data presents its patterns.

[ ]: