COVID 19 global dataset

1. Problem Statement - High level question/ prompt/ trigger for the project

* Choose 5 countries with complete data to analyze the increase/decrease in new cases through 12 months (Afghanistan, Australia, Morocco, Italy, USA?) randomly select 5 countries from each continent

1. Description/ Outline - 3-5 sentence summary of the project idea
2. Research Questions to Analyze - 3-4 at least
3. Dataset links

<https://www.kaggle.com/datasets/josephassaker/covid19-global-dataset>

Data breakdown:

“worldometer\_coronavirus\_daily\_d”

* cumulative\_total\_cases
* daily\_new\_cases
* active\_cases
* cumulative\_total\_deaths
* daily\_new\_deaths

“worldometer\_coronavirus\_summary”

* continent
* total\_confirmed
* total\_deaths
* total\_recovered
* active\_cases
* serious\_or\_critical
* total\_cases\_per\_1m\_population
* total\_deaths\_per\_1m\_population
* total\_tests
* total\_tests\_per\_1m\_population
* population

Skills we can use:

* merge datasets
* dataframe
* plot
* regression
* API?

Presentation preparation:

* Introduce datasets
* Walk people through the code
* Questions mentioned and answers
* Summary of our findings
* Limitations of the datasets