

Covid-19 Project

Data Source

Covid-19 dataset from Our World in Data (OWID):

Website: <https://www.ourworldindata.org/coronavirus>

GitHub: <https://www.github.com/owid/covid-19-data/tree/master/public/data>

This is an external and open-source dataset. OWID collects data from governmental organisations around the world and compiles it in to one csv file. There are limitations to the collection of data, as different governments have different practises for collecting and sharing data.

I have chosen this data because it is a very current issue which is having a huge impact on peoples lives across the globe. There is also a large volume of data to analysis which makes it an interesting project.

Data Profile

72,096 records. 59 columns.

Cleaning Steps

- Checked missing values. Replaced NaN's with 0's.
- Checked duplicates.
- Checked mixed data types.

Understanding the Data

I have exported descriptive stats for the dataset into a separate csv file.

Limitations and Ethics

There are a wide range of variables recorded across a wide range of countries. As all countries have different data recording practises, there are missing records across the dataset. However, there is enough data present to conduct an interesting analysis. There may be some inaccuracy in recording of data, as some governments may not be as accurate or thorough as others. Our World in Data is considered a reliable source but naturally almost all of its data comes from other (mostly governmental) sources, so this should be considered when assessing data quality.

Questions to Explore

- How is the stringency of countries lockdowns related to cases of covid-19?
- How is the population density of countries related to cases of covid-19?
- How are HDI ratings related to covid-19 deaths?
- How does gdp per capita relate to covid-19 deaths?
- What is the effect of vaccinations on the reproduction rate of covid-19?
- How do rates of covid-19 compare amongst countries with a similar population?
- How do rates of cardiovascular disease effect hospitalisations due to covid-19?