June 15, 2022

Chia yi quan

Student Number: 210220223

CM3010 Databases and Advanced Data Techniques

MID TERM REPORT

Table of Contents

[Stage 1. Find and critique a dataset 2](#_Toc106906529)

# Stage 1. Find and critique a dataset

## Choosing A Source Of Data And Explanation Of Interest

For this coursework, I have chosen an open data from [Our World In Data](https://ourworldindata.org/) and the dataset I used is the [coronavirus](https://ourworldindata.org/coronavirus-source-data) dataset. This is the [link](https://github.com/owid/covid-19-data/tree/master/public/data/) where the csv file is downloaded from. The reason why I picked this dataset is because coronavirus have been with us for the past 2+ years and there have been many mutations of the virus and what made this dataset interesting is that we can examine the number of cases, death rate, vaccination and other data by country. Because the data is in csv, we could use the database to help us organize the data better which will allow us to visualize the data easily. The main objective of this database is to allow us to better visualize the coronavirus cases, the relationship of the death rate and vaccination around the world by country, continent, and as a world.

## Assessing The Dataset