1. **Team Members**: Meagan Greenhill

Hesh Kuruppuge

Jacqueline X

Mike Murphy

1. **Project Description**:

The project uses John Hopkins University Covid-19 datasets sourced from Google Dataset Searches, and related datasets to extract the required data to create an integrated database that can be used to provide ongoing analysis of Covid\_19 and its global impact. The database will be updated daily and will facilitate analysis that currently requires disparate source datasets.

1. **Project Rationale**:

There is currently no single source that provides a comprehensive view of the global Covid pandemic. To get a comprehensive overview you have to use multiple data sources and combine their outputs to get a single view.

This project combines the main sources of global Covid data, global vaccination data and global population data to provide a single, integrated source of information in a relational database and to provide basic queries to retrieve views of the data. These queries can be copied and adapted to provide ad hoc query capability.

1. **Project Datasets**:

The datasets for the project can be found at the following links.

“JHU – Time Series Daily Reports”

<https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv>

“World population data”

<https://www.worldometers.info/world-population/population-by-country/>

“Vaccination rates per country”

<https://ourworldindata.org/covid-vaccinations>?

1. **Database Schema - ERD**

* ERD and description of data base to be inserted here

1. **ETL – Data Extract:**

Description of Data Extract to be inserted here

1. **ETL – Data Transform:**

Description of Data Transform to be inserted here

1. **ETL – Data Load:**

Description of Data Load to be inserted here