**Short report of the data scraping project**

Bo Zhao, Zach Hansen, Collin Grant

1. **Data collection methods**

First, all HTML content in the specified url is captured by using **requests**.Second, the obtained HTML data is parsed by using **BeautifulSoup** and **lxml** and gives us a Beautifulsoup object.Then the Beautifulsoup object is used to navigate and search the desired entries using the tag name and attributes. In this way, we can collect and store the clean covid data and its corresponding name.

1. **Json file structure**

We are able to collect three consecutive days of covid data for 230 countries and territories. They are stored in three Json files named today.json, yesterday.json and two\_days\_ago.json. In each json file, there is a list of 230 dictionaries, each dictionary contains 22 key-value pairs that contain all the covid data of one country.