Submit a 1 pager on your experience reflecting on the challenges, what was easier than you thought and what was harder. How would a utility like this be useful for other data projects you may encounter?

One thing that was easier than we thought was making changes to the files because we could reference previous codes we used for the class. It was nice to label the two different dataframes ("dfa" for the .csv file, and "dfb" for the .json file). We were able to make relevant changes to each one, such as adding a column regarding whether the crop needs watering or not for the .csv file (which was related to worldwide crop production). For the .json file, it was easy to add another column with the name of another relevant crop disease since this was a dataset related to plants either being healthy or having one of the listed diseases. Overall, this was likely the smoothest part of the project to run through.

One thing that was harder than we thought was converting file type, especially with user input. We ran into some problems due to not saving the right file into GoogleColab at the beginning of each run, since you need to reference the file in the code. The second part of the project related to converting the data types into different forms took us the longest time. We encountered problems with the code not running the second time because the file already existed from the first run. By adding the error handling lines, it helped to see where exactly the problem was coming from. Converting from .csv to .json and from .json to .csv was less confusing than converting from either to SQL, which is were we spent most of our time fixing.

Utilities like this could be useful if we need to convert one file type to another and if we need to make changes to the files. The input function also helps the user to see the data in the file type they want. It is important in a lot of industries to be able to read a database and have extra changes made to it, whether that be removing a column that is no longer relevant, or adding a column to capture extra information that was not previously recorded. Having the flexibility to start out with one type of database and being able to convert it to any other is really useful to have compatibility with various different software or programs. Some might only take in a certain input, so being able to change from one to the other can ensure that data can be used across multiple tools. Each type has its own strength, with .csv being used for spreadsheets and analytical tools, while .json is used for APIs and web development, and SQLite as a relational database to store structured data locally.