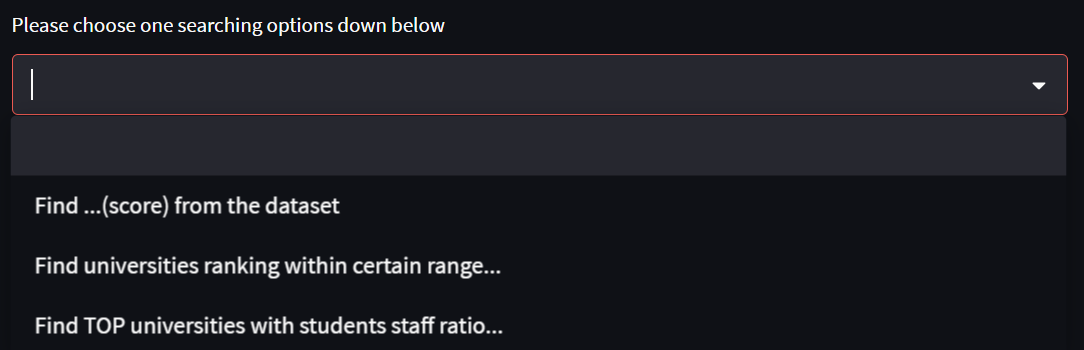
For the search functions on our EDFS, we implement partition-based map and reduce for different data. To better understanding how the search functions work and what are the results, the search page of the app include several query options you can choose. Within each query option, you can choose certain attributes to look for and thus user are able to navigate through the database and search.

Following pictures illustrates how the search functions look like in our app:

1. Firebase

There are three options.

The first one “Find ..(score) from the dataset” allows you to get specific scores for every university. Attributes selected are information to be included in the return dataframe structure. (See Figure 2.1.1)

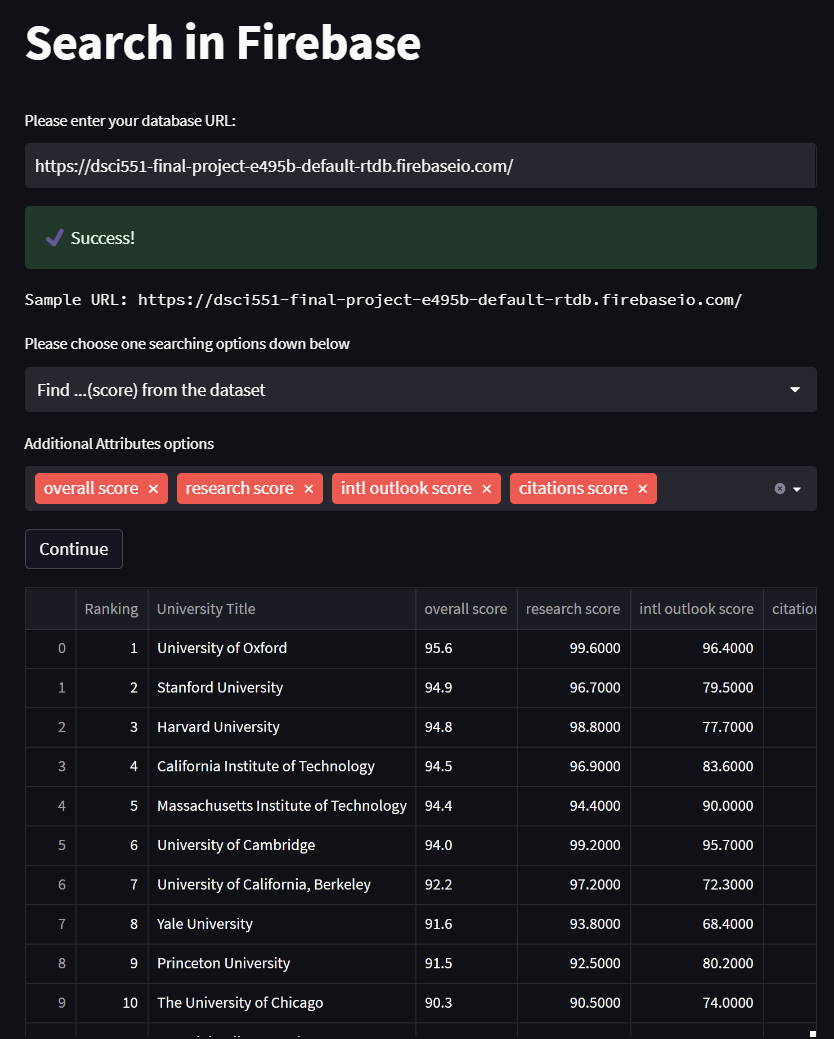


Figure 2.1.1

The second one “Find universities ranking within certain range” allows you to find universities within range you choose and it returns a dataframe contains rankings and school name. So the pseudo query if written in SQL would be “SELECT ‘rankings’ FROM dataset ORDER BY ‘rankings ‘”. The map after map-reduced is in the structure (key=ranking, value = UniversityName). (See Figure 2.1.2)

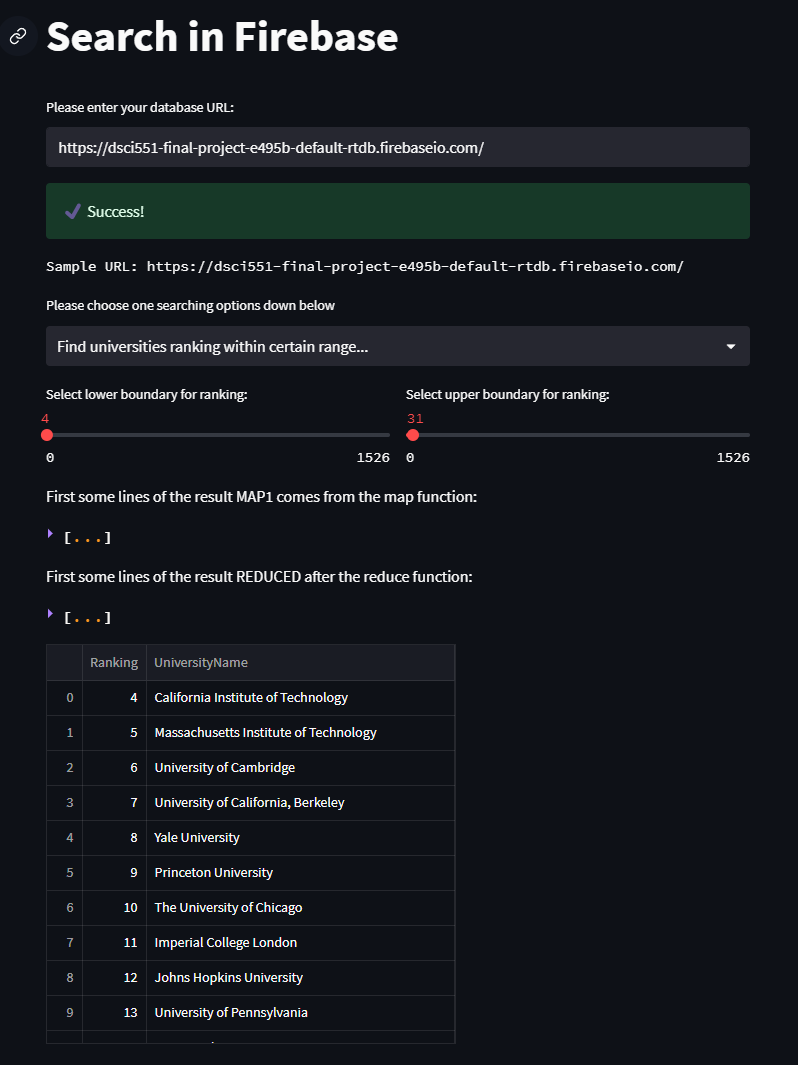


Figure 2.1.2

The third one allows you to find top... universities in student staff ratio, with some other attributes you can choose to have. For example you can add attributes “location”, “perc intl students” and “number students”. So the pseudo query if written in SQL would be “SELECT ‘students staff ratio’ FROM dataset WHERE location and perc intl students and number students ORDER BY ‘students staff ratio’ LIMIT number”. The map after map-reduced is in the structure (key=(ranking, universityName), value=(studens staff ratio, all other attributes selected)), Specifically we choose ranking and university name combined to be the key because we get unique information with them combined. (See Figure 2.1.3)

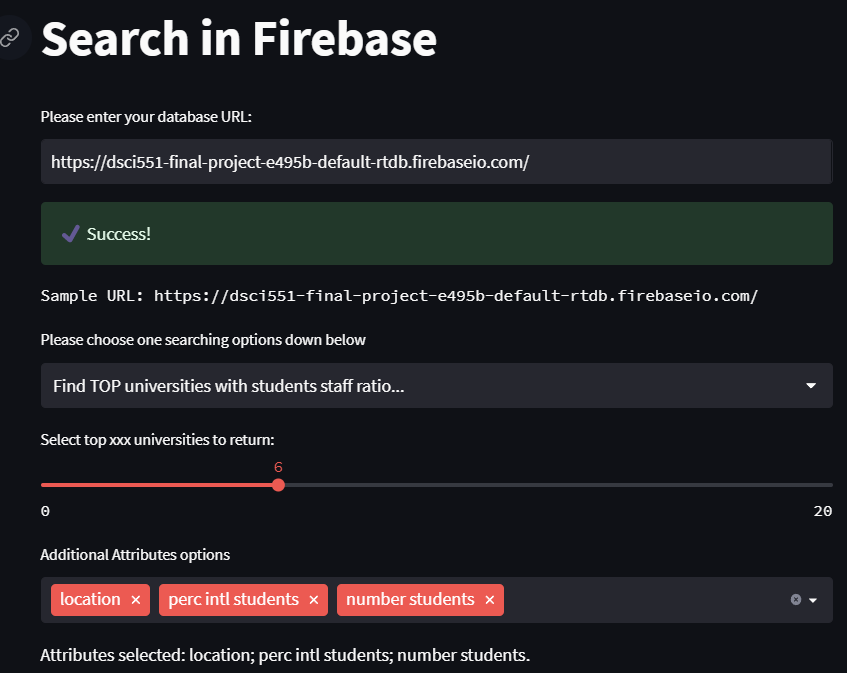


Figure 2.1.3 (Part 1)

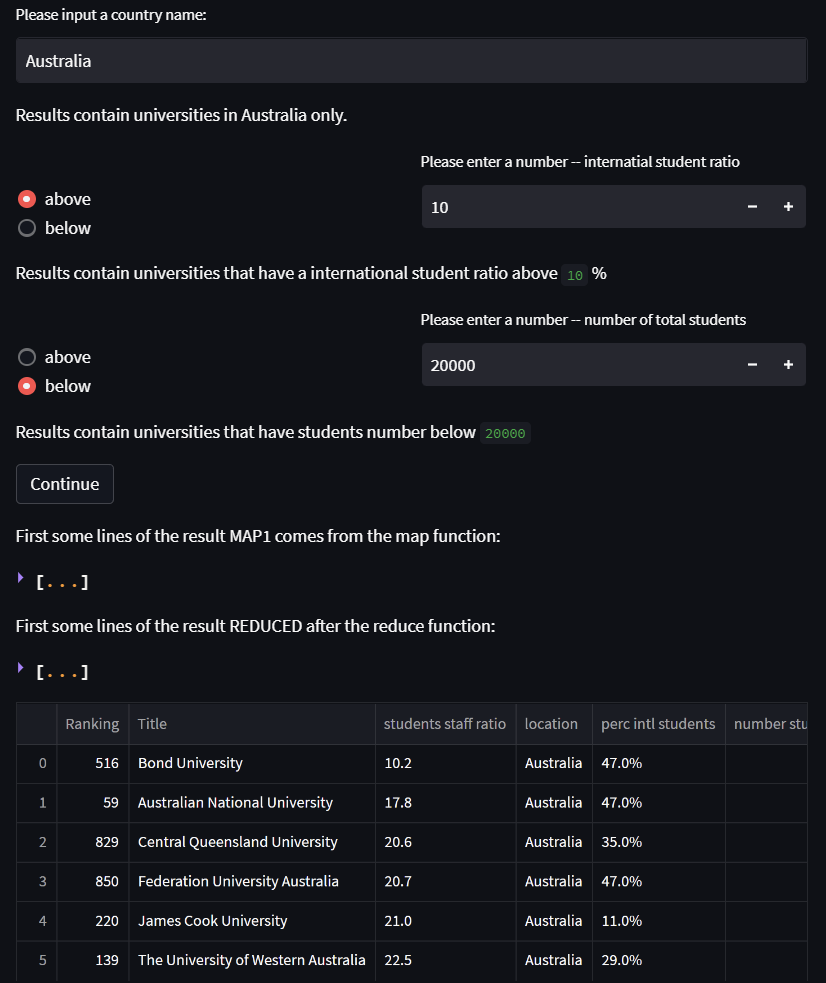


Figure 2.1.3 (Part2)

1. MongoDB

Search in MongoDB is implemented in a similar way.

The First query allows you to navigate through CountryDataP dataset, and find top ... countries in certain area. You can also select an additional attributes that you want to obtain its information at the same time. (See Figure 2.2.1)

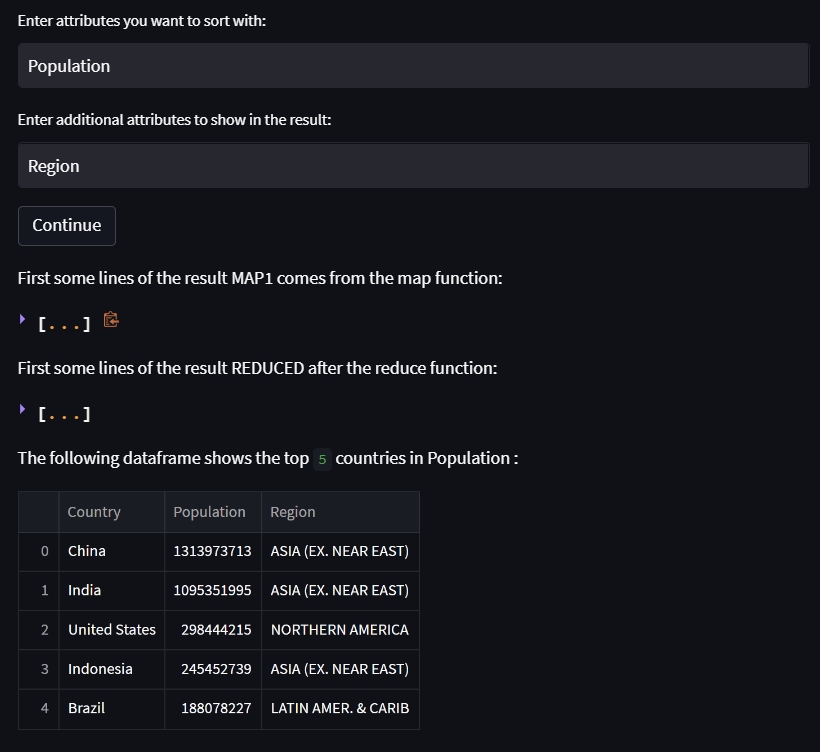
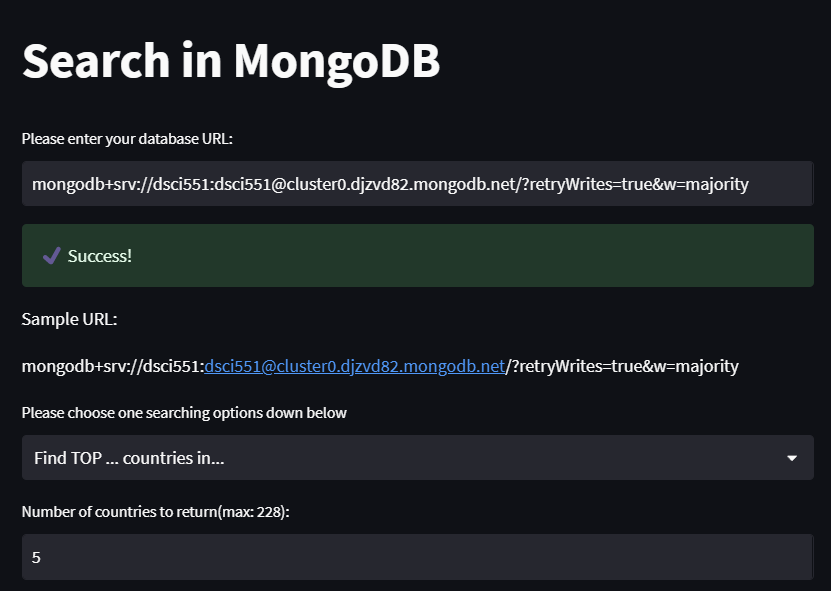


Figure 2.2.1

The second query allows you the find GPD expenditure greater than ... in percentage in a particular year. (See Figure 2.2.2)

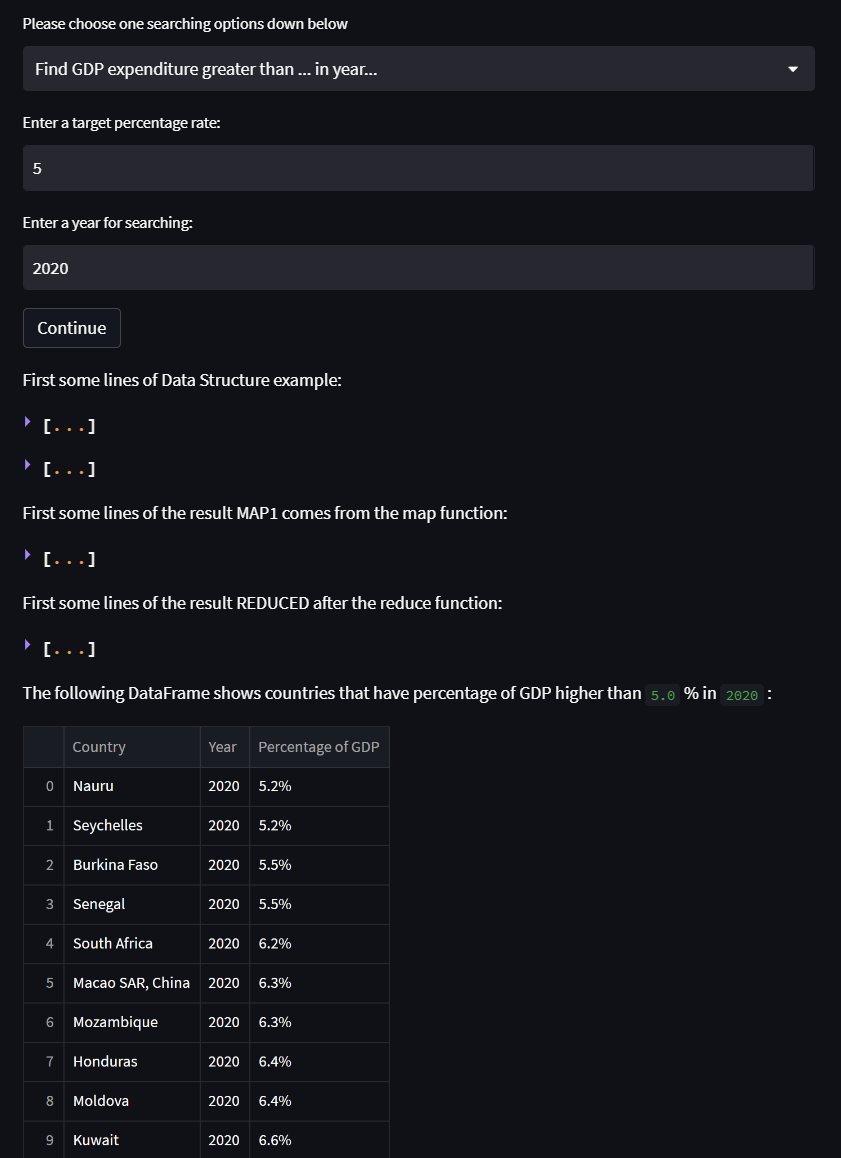


Figure 2.2.2