I have 2 datasets for 2 different questions.

I would like to set up a meeting during your office hours to see which would be a better question.

My first data set is, “Active DACA Recipients 9/3/21”. This is my preferred data set to conduct research on. I attained this data set from <https://www.uscis.gov/tools/reports-and-studies/immigration-and-citizenship-data?topic_id%5B%5D=33602&ddt_mon=&ddt_yr=2021&query=&items_per_page=10>

This is a government website that regularly posts datasets and information about immigration and citizenship. This is a reliable data source.

One issue I see with this dataset is that the layout of the excel sheet needs to be manipulated so that it can be read on R. The dataset also does not give some values when it is easy to pinpoint an exact person. For example, if there are 8 people from Alaska who have DACA, the dataset will put a value of D instead of 8.

Also there has been research done for this data set already. This dataset has been turned into a geospatial visualized website from someone on reddit. The person that made this has data from up to January 2021. I have data for SEP 2021.

There are also tons of research done for these DACA dataset online which I can use a peer review articles.

My back up dataset is Edu3\_week40. This is a dataset about education. I took this dataset from [Week 40 Household Pulse Survey: December 1 – December 13 (census.gov)](https://www.census.gov/data/tables/2021/demo/hhp/hhp40.html). This dataset is about the Impact of Coronavirus Pandemic on Post-Secondary Educational Plans, by Select Characteristics. The dataset will have to be manipulated so R can read it. I am sure I can find peer review articles for this type of information. Another dataset I can use from the same website is on the Impact of Coronavirus Pandemic on Post-Secondary Educational Plans, by Reasons for Changes