

Aaron Fox  
Final DS Project

For my final project it was my plan to take data from the UNWorld data collection. One data set had to do with education statistics based on country and continent, and the other had to do with population growth and fertility rates based on the same.

I started by creating structs to hold the data and function to parse through it, heavily editing the data files as it to make them easier to parse. Some of the municipalities held data from multiple years, so the code takes the most contemporary data into account.

It was then the goal to filter out municipalities that existed in one data set, but not the other. There were a few which did not have stats recorded for both.

I wanted to then cluster the data sets based on education rates, fertility, mortality, birth rates, etc. and compare the clusters to the existing world continents, countries that are a part of the UN or Nato, first world countries, etc.

I've spent over 30 hours trouble shooting and editing my code, starting from scratch 10 times, consulting peers as well as chatGPT as a means of troubleshooting, but have had no luck.