

One project involving data structures is data analysis using numerical linear algebra. A specific example was my summer research project in which large amount of infant heart beat data was analyzed. The data had to be stored in arrays in such a way that it could be easily accessed and run through other programs. Some of the data was corrupt or missing which meant large amounts of the data had to be sorted and the missing data (represented as NaN) had to be removed. Because these data files could each be multiple gigabytes, these sorting methods had to be efficient otherwise this step could take hours. After this was done, the actual numerical linear algebra could be performed which took the organized data and performed a decomposition, and returned more data in a new form. All of this data had to remain organized and easily accessible otherwise none of this would have been possible.

Another possible project would be an autocorrect program. Whenever somebody types a word, the computer has to check whether what they typed matches anything in a dictionary. If there is no match, then what was typed has to be compared to words that do exist and the closest one can be picked. However, the computer cannot simply look at the typed letters in order because the first letter may be the only mistyped letter. Therefore, all parts of what was typed has to be compared to the dictionary to find the most likely match.

A third project could be giving movie recommendations based off of what movies the user already like. Given a list of movies, the user could be asked to rate each one. Their ratings would then be used to choose which movies from another list should be recommended. Essentially, the first list would be sorted by their rating. Then this sorted list would determine how another list of movies is sorted. Finally, the second sorted list could be returned to give the recommendations to the user.