Phylogenetic Tree

The needlemen\_wensch function computes the Needleman-Wunsch distance between two strings given a dictionary scores for charactesr. The all\_pairs\_needleman\_wunsch function loops through all of the possible combinations of species and computes the Needleman-Wunsch distance between their DNA strings, saving all of the values to a new data dictionary. The Dendrogram\_Tree class has a construct dendrogram class that creates a leaf node for every species and keeps track of each nodes’ root in a dictionary. It uses the union find class to union species that are not already in the same component. The get\_clusters method takes in a threshold x value and returns a list of lists where each inner list has all of the species that are connected and passed the threshold value.

I worked on parts of this with Rachael Grossbauer and got help from Sam Gregory.