Homework Assignment 1 Overview

This program reads in a file with employee information such as first name, last name. Middle initial, ID number, and phone number, processes the text to be more standardized, and then creates an object for each person and outputs the information in an organized manner. The program first checks that a file path is in a sysarg and if so opens that file but if not the program ends. Once the file is opened, the program passes the file to a function to process the text inside. This processing function deletes the header line then creates a dictionary to hold the objects. It loops through each line and checks to see if the information provided matches the required format. If the information does match, the program moves onto the next line after creating an object but if something doesn't match the program prompts the user to input a correctly formatted answer. After the line is processed, the program checks that there are no duplicate IDs so if the ID is already present an error message is printed and it moves to the next line but if no duplicate is found the object is added to the dictionary. Once the dictionary is returned to the main function, it's saved as a pickle file, then opened and each object is printed.

One strength that python has when it comes to text processing is that python in general is a very natural and clean language to program in so it feels efficient when writing up the text processing sections. One weakness that python might have when it comes to text processing is that the runtime might become very big once more data is used. I'm not sure how python compares to other languages in text processing but that seems like it could become a problem. One thing I learned from this assignment was how to use regular expressions. Regular expressions were completely new to me before this assignment and seemed a little daunting but after spending some time researching them, I found them to be pretty interesting and very helpful for the text processing in this assignment.