

Tedi Dika
txd190008
CS 4395.001

Assignment 1

This program reads in a file containing data of Employees, including name, id, and phone number. It splits the data into a list and then sends it into a text processing function. If data is in an invalid format the program will prompt the user to enter valid data. If a duplicate id is found the program will output an error. The data is then standardized and an object is created for each Employee with the corrected data. These objects are put into a dictionary with the id as a key, pickled, and then each employee's information is printed out.

To run it simply make sure that 'data/data.csv' is input as a system argument as that is the file the program will read from.

The strength of python text processing is that it has a lot of useful built in functions. For example, to capitalize the first letter of a word and make the remaining letters all lowercase all I had to use was the capitalize() function. In Java I would have had to do an extra step and use substring to split the first letter. The 're' library made it very easy to process text as well. The weakness of text processing in Python is that it is slower than compiled languages like C++.

In this assignment I learned to process text with the 're' library. I also learned the concept of 'pickling' and how to implement it into a program. The object creation and file reading was straightforward and mostly just review. Overall I became more familiar with the Python language.