Grace Wang

Mr. Lin

Unit Project 7: Algorithms for Data Analysis

Feb 8, 2022

(1) problem description
The data file of the company employees is very unorganized. Therefore, I must:
Set 2 –
1. Create a sorted employee data file based on the descending order of their first names.
2. Print the data of top 5 highest salary managers.
3. Find the data of an employee whose first name is Elon and he is a manager.
4. Compare the average salaries of employees in Hawaii and Maryland.

(2) explanation of sorting/searching algorithms used in the program
1. In order to create a sorted employee data file based on the descending order of their first names, I had to compare the employees' names with each other. To do this, I assigned 2 variables i and j to the first two employees and compared the ASCII codes of the letters in their first names. If the first employee's name, assigned the variable "i" is greater than that of the second employee, they would swap positions, because the letter Z has less of a value than A since we want the first names to be in descending order. After the first two names have been compared, we assign the second employee to i and the third employee to j. We would continue going through this process with all the employees until j reaches 2000. The result of this algorithm is the first names organized by descending order. However, then I had to organize the first names of each employee with their corresponding information. Since I created a baseline, comparing the first names of each employee, the rest of the data such as last name, position and state would correspond to the first name when I compared them as well. Thus, I string compared and integer compared last names, positions, states, ages, salaries, education levels, and years of industrial experience and printed out the sorted employee data file based on descending order of their first names
2. In order to print the data of the top 5 highest salary managers, I had to organize the employee data file based on the ascending order of people's salaries. I did this by conducting the same procedure as that of sorting the employee data file based on the descending order of their first names. Then, I conducted a linear search to find the first 5 salary managers on the organized file. I did this by comparing every person's position to the word "Manager". If there was a match, I would print their name out, and add one count to the limit of 5, because we only want the first 5 salary managers. Once the count equaled 5, we would stop the search.
3. To find the data of an employee whose first name is Elon and he is a manager, I string compared the position of every employee to the word "Manager" and string compared the position of every employee to the word "Elon." When there was a match, I printed their data out.

4. To compare the average salaries of employees in Hawaii and Maryland, I went through every employee and searched if they came from Hawaii. If yes, I added +1 to the population count, and added their salary to a variable I created for the total salaries. When I went through everyone, I found the average salary of people from Hawaii by adding the total salaries added to the total people collected and printed the number out. I repeated this process with employees in Maryland.

(3) implementation (computer programs as attached compressed file) of your algorithms
https://replit.com/join/ogjukfkjvr-gracewang20

(4) the results (output) of your program.
1. The output of my program is the organized employee data file based on the descending order of their first names on a txt file
2. The output of my program is the top 5 highest salary managers
3. The output of my program is the data of an employee whose first name is Elon and he is a manager
4. The output of my program is the average salaries of employees in Hawaii and the average salaries of employees in Maryland.