REPORT DATA STRUCTURES ASSIGNMENT 1

PART 2

My selection of Hash function for part 2 was pretty simple. As we could see, in part one, our hash function would evaluate the index depending on the ASCII value of our characters and made sure it would fit into the array size.

My choice for the function in part two was to add a multiplying factor with a prime number. This would add more variability to the calculation of our hash index making it more difficult to get the same value when doing the process.

As we can see in my program, using this method we decreased the collisions of the code significantly. On part 1, we were having 54 collisions. After using this new method, the collisions went down to 22. This shows the effectivity of this simple change in our method when hashing.

PART 3

In part three of our assignment, we were asked to use double hashing as our new method to handle collisions, replacing linear probing. Double hashing can result in better outcomes that linear probing as it doesn't go from index to index summing one till we find one that is free. As this can result in a very high number of collisions if not done properly. Double hashing uses two hash functions to calculate our hash index, which makes it more efficient when calculating our index. With the second function it calculates the step it takes depending on the values of our characters. This makes the step more random and better resulting. As we can see, double hashing does work better that linear probing as the collisions went down from 22 (part 2) to 15.

PART 4

For part 4, which comparing my output to the sample output in the PDF, I do not have the same. In task 4, I have used an approach that when going through the lines of the csv file, it would skip all the information in lines except the ones of the surname and name. This I have done using tokens and a counter variable to know in which part of the line I am at. This way of doing it and my search function has worked successfully and I am able to allow the user to input a surname and he is able to see all the names with that surname.

It is true that I have had many issues when doing this part, especially when reading the file correctly. This is why I have added a few functions

to get rid of whitespaces, commas and also change all the letters to lower case. This has made it easier to work with so many information and has allowed me to make my program run successfully.