

Task 2&5 Supplementary notes

Task 2

Input specification:

Both starting year and ending year should be in range of 1880 – 2019 and starting year shouldn't be smaller than ending year.

Gender should be chosen from the choice box, which is to choose from Female and Male.

Ranking should be an integer from 1 to 1000.

Output explanation:

The percentage of each name is calculated by the occurrences of certain name at the input ranking divided by the sum of all the occurrences of names at the input ranking during the given period of time.

Task 5

Algorithm 2:

Input specification:

Input name should be found in the dataset of the user's born year and it is not case sensitive (both capital letter and small letter) is accepted.

Logic of the recommendation:

First, the system looks for the ranking of the user's name at his/her born year.

Second, according to the age preference of the user and gender of interest, the system reads data of certain years. If user looks for younger soulmates, data from two years after user born year to user born year is selected(inclusive). If user looks for older soulmates, data from two years before user born year to user born year is selected(inclusive).

Then it adds up the occurrence of names and sort them according to occurrence.

If the name at the ranking of user's name at born year is not user's own name, it is recommended, otherwise, if user's name is not the most popular name, the algorithm recommends the name with ranking higher than user's name by 1. If it is the most popular name, the algorithm recommends the second name.