Homework Assignment #10

CS5004 – Object-Oriented Design
Northeastern University – Silicon Valley
Summer 2019

Due Sunday 08/11 at 11:00pm PDT

Grading: Each programming problem is graded as follows

- A submission which does not compile gets 0.

- A submission which compiles but does something completely irrelevant gets 0.

- A submission which works (partially) correctly, gets (up to) %80 of the total credit.

- %20 is reserved for the coding style. Follow the coding style described in the book.

Formatting: Each class must reside in its own file. If you need to instantiate objects of class `A` in your class `B`, your `B.java` file must import `A.java` first. The names you choose for your classes should match the ones specified in each problem. Also, for each problem i, you must have a `Problem_i.java` file containing (only) the class which has the `main` method. Finally, you must put all files related to problem i into a folder named `problem_i`. The files *must be placed in the right sub-folder*. Make sure that your code complies. See the note in Problem 1 for an example.

Submission: For each problem i submit one `problem_i.zip` file to Blackboard. The zip file should maintain the structure of the sub-folders. See the note in Problem 1 for an example.

Problem 1 [60pts]. Write a program that approximates the probability that 2 or more people in the same room have the same birthday, for 2 to 50 people in the room. The program should use simulation to approximate the answer. Over many trials (say, 5,000), randomly assign birthdays (i.e., a number from 1–365) to everyone in the room. Use a `HashSet` to store the birthdays. As the birthdays are randomly generated, use the contains method of a `HashSet` to see if someone with the same birthday is already in the room. If so, increment a counter that tracks how many times at least two people have the same birthday and then move on to the next trial. After the trials are over, divide the counter by the number of trials to get an estimated probability that two or more people share the same birthday for a given room size. Your output should look something like the following. It will not be exactly the same due to the random numbers:

```
For 2 people, the probability of two birthdays is about 0.002
For 3 people, the probability of two birthdays is about 0.0082
For 4 people, the probability of two birthdays is about 0.0163
...
For 49 people, the probability of two birthdays is about 0.9654
For 50 people, the probability of two birthdays is about 0.969
```

Problem 2 [60pts]. You are given two text files `boynames.txt` and `girlnames.txt`, which you can find on Piazza. These contain lists of the 1,000 most popular boy and girl names in the United States for the year 2005, as compiled by the Social Security Administration. These are blank-delimited files where the most popular name is listed first, the second most popular name is listed second, and so on to the 1,000th most popular name, which is listed last. Each line consists of the first name followed by a blank space followed by the number of registered births in the year using that name. For example, the `girlnames.txt` file begins with

```
Emily 25494
Emma 22532
```

This indicates that Emily is the most popular name with 25,494 registered namings, Emma is the second most popular with 22,532, and so on.

Write a program that determines how many names are on both the boys' and the girls' list. Use the following algorithm:

- Read each girl name as a String, ignoring the number of namings, and add it to a `HashSet` object.

- Read each boy name as a String, ignoring the number of namings, and add it to the same `HashSet` object. If the name is already in the `HashSet`, then the add method returns false. If you count the number of false returns, then this gives you the number of common namings.

- Add each common name to an `ArrayList` and output all of the common names from this list before the program exits.