Baby Names Part 1 - Files (100pts)
due: Apr 12, 2016 11:59pm

Important Notes

- mercurial repository hw6
- babynames files are in /rd/cs246s2016/shared/hw6/babynames
- **This assignment is to be done on your own.** If you need help, see the instructor or TA.
- Please start the assignment as soon as possible and get your questions answered early.
- Read through this specification completely before you start.
- Some aspects of this specification are subject to change, in response to issues detected by students or the course staff.

In this project, based on the names and corresponding genders from a given input file, your program will read and process 25 files in the following format:

```
rank Male_name Male_number Female_name Female_number
```

where

`rank`	The ranking of the names in this file
`Male_name`	A male name of this rank
`Male_number`	Number of males with this name
`Female_name`	A female name of this rank
`Female_number`	Number of females with this name

This is the format of database files obtained from the U.S. Social Secutiry Administration of the top 1000 registered baby names. Each line begins with the rank, followed by the male name at that rank, followed by the number of males with that name, etc. Here is an example file containing data from the year 2014:

```
1 Noah 19144 Emma 20799
2 Liam 18342 Olivia 19674
3 Mason 17092 Sophia 18490
4 Jacob 16712 Isabella 16950
5 William 16687 Ava 15586
6 Ethan 15619 Mia 13442
7 Michael 15323 Emily 12562
8 Alexander 15293 Abigail 11985
9 James 14301 Madison 10247
10 Daniel 13829 Charlotte 10048
......
991 Ayan 207 Samiyah 266
992 Boden 207 Yaritza 266
993 Foster 207 Cordelia 264
994 Jair 207 Micah 264
```

```
995 Reyansh 207 Nala 264
996 Tyree 207 Belen 263
997 Ean 206 Cambria 263
998 Leif 205 Natalya 263
999 Reagan 205 Kaelynn 262
1000 Rylen 205 Kai 262
```

You can see that, in 2014, there were 19144 male babies named Noah and 20799 female babies named Emma. Both names are the most popular names used in 2014.
Based on the given names and corresponding genders, your program will process each of these files containing baby names and output the following values for each given name :

- the rank, number of babies named this given name, and the percentage of the number of the given name out of the the total number of the babies with the same gender in each file;
- the rank, the number of babies named this given name, and the percentage of the total number of the babies with the same gender in **all** files.

Note that in the first item, the rank, number of the babies of the given name can be obtained directly from the file. You need to compute the total number of babies with the same gender to get the percentage. In the second item, all three items will have to be computed since you need to consider the data from all files.
For example, in 2014, line 166, `166 Bennett 2525 Elise 1959`, we can see that 2525 male babies were named Bennett, making Bennett the 166th most popular boy's name. Your output should then look like:

```
babynames/names1900
Emma(F): 29, 3312, 0.007612
Bennett(M): 712, 44, 0.000102
......
......
babynames/names2014
Emma(F): 1, 20799, 0.015810
Bennett(M): 166, 2525, 0.001576
......
Total:
Name(M/F): rank, total number, percentage
Emma(F): 2, 293856, 0.010817
Bennett(M): 378, 16916, 0.000520
......
```

The format of the given input file is as follows:

```
6
Emma F
Bennett M
Olivia F
Noah M
Liam M
Sophia F
```

The first line of the file contains the total number of names. From the second line, each line contains a name and its gender. When a mismatch between the number given in the first line and the actual total number of the given names happens, your program should report the error. Note that you need to use command line argument to accept the input file.

What to hand in:

- A header file to go with your .cpp file
- A Makefile
- A README file

Things to note:

- Before you start to program, sketch your design.
- Structure your program as modular as possible.
- Comment your program nicely.