Baby Names Part I - Files (100pts)
due: Nov 3, 2014 11:59pm

Important Notes

- Submission command: submit -c 246 -p **6** -d YourDirectory
- **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 23 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 2012:

```
1 Jacob 18899 Sophia 22158
2 Mason 18856 Emma 20791
3 Ethan 17547 Isabella 18931
4 Noah 17201 Olivia 17147
5 William 16726 Ava 15418
6 Liam 16687 Emily 13550
7 Jayden 16013 Abigail 12583
8 Michael 15996 Mia 11940
9 Alexander 15105 Madison 11319
10 Aiden 14779 Elizabeth 9596
......
991 Karsen 200 Devyn 253
992 Jarrett 199 Geraldine 253
993 Apollo 198 Analia 252
994 Denzel 198 Hayleigh 252
995 Foster 198 Landry 252
```

```
996 Gilbert 198 Sofie 252
997 Jaylon 198 Tess 252
998 Kylen 198 Ashtyn 251
999 Augustine 197 Jessa 251
1000 Dangelo 197 Katalina 251
```

You can see that, in 2012, there were 18899 male babies named Jacob and 22158 female babies named Sophia. Both names are the most popular names used in 2012.

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 2012, line 166, `166 George 2320 Aurora 1890`, we can see that 2320 male babies were named George, making George the 166th most popular boy's name. Your output should then look like:

```
babynames/names1900
Catherine(F): 19, 3717, 0.008543
George(M): 4, 13000, 0.030051
......
......
babynames/names2012
Catherine(F): 167, 1887, 0.001470
George(M): 166, 2320, 0.001468
......
Total:
Name(M/F): rank, total number, percentage
Catherine(F): 67, 68299, 0.002782
......
```

The format of the given input file is as follows:

```
7
Jake M
Catherine F
George F
George M
Jacob M
Mercedes F
Kenneth M
```

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 .c file
- 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.