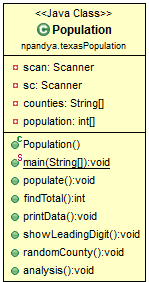
**Background**: Benford's law, also know as the first-digit law, states: given a list of numbers from a real world data set, the distribution of leading digits is often not uniform and instead skewed towards 1. Consider for example the populations of the counties in Texas from the 2010 census (file on moodle). Most people would guess there are roughly equal numbers of populations that start with 1, 2, 3, 4, 5, 6, 7, 8, and 9. (We don't consider 0 a leading digit for this assignment.) Given there are 254 Texas counties you might expect there to be about 28 counties (254 / 9 = 28) counties that have a leading digit of 1, 2, 3, and so forth. Travis County (home to UT) has a population of 1,024,266, a leading digit of 1.

Our intuition is often wrong. The breakdown of leading digits for populations of Texas counties according to the 2010 census are:

|  |  |  |
| --- | --- | --- |
| Leading Digit | Number of Counties | Percentage |
| 1 | 80 | 29.99 |
| 2 | 38 | 17.76 |
| 3 | 41 | 6.75 |
| 4 | 26 | 23.85 |
| 5 | 15 | 4.19 |
| 6 | 15 | 4.13 |
| 7 | 17 | 7.98 |
| 8 | 13 | 4.75 |
| 9 | 9 | 0.60 |

Not what you expected! Populations with a leading digit of 1 occur almost 1/3rd of the time! Not 1/9th as most people would guess. Benford's law does not hold for all data sets (for example height of humans in inches), but does hold for a surprisingly large number of real world measurements.

**Assignment:** Create a new Java project and then create a class that has the following specifications.

When the user runs the main method, it comes up with a menu as shown below:

Would you like to look at Texas population? yes/no:

yes

Please select one of the following:

1. Print the data from the file.

2. Find Texas total population

3. Randomly choose a county and display its population

4. Find counties population with a leading digit, entered by the user

5. Find total number of counties/percentage of population with each leading

digit formatted using printf

User will choose an option and the program will call an appropriate method to display the outcome on the terminal window.

Your program has multiple methods:

populate(): It reads the data from a file using Scanner and populates the arrays: one with a name of the county and the other one with the population for that county.

printData(): reads the data ie. all the counties and their population, from the arrays and prints it in a formatted output as shown below: (Use printf for formatting)

Reading data from a file into an array:

Anderson County 58458

Andrews County 14786

Angelina County 86771

Aransas County 23158

Archer County 9054

Armstrong County 1901

Atascosa County 44911

Austin County 28417

Bailey County 7165

Bandera County 20485

Bastrop County 74171

…

findTotal(): iterates through the arrays and returns the total population of Texas.

showLeadingDigit(): asks the user for a leading digit and then prints the name of the county and its population with the leading digit that matches the user's input:

Enter the leading digit 1 -9:

5

Showing data with a leading digit of: 5

Anderson County 58458

Cherokee County 50845

Delta County 5231

FortBend County 585375

Hansford County 5613

Hardin County 54635

Haskell County 5899

Hood County 51182

JimHogg County 5300

Lynn County 5915

Maverick County 54258

Rusk County 53330

VanZandt County 52579

Wheeler County 5410

Wise County 59127

randomCounty(): creates a random number between 0-253 (254 counties in all) and uses that number to randomly pick a county and its population to display on the terminal window:

Randomly chosen county at index 31 is: Camp County

and its population is: 12401

analysis(): This method will print out the total number of counties and the percentage of the population for that total for each leading digit as shown below: (Hint: use findTotal()method to get the total population and then find the percentage)

digit number percentage

1 80 29.99

2 38 17.76

3 41 6.75

4 26 23.85

5 15 4.19

6 15 4.13

7 17 7.98

8 13 4.75

9 9 0.60

In addition to the methods described above, you will have a main method for this class that presents the menu to the user as described above. Continue to document your code. Run the program for me when you are done. We will expand on this later so hang on to the program.

**I do want you to use arrays for this exercise, not ArrayList!**