Coding Assignment

The task should be solved with Java/Groovy

Task definition

Create a program that takes as input a list of people names and finds duplicate names. There are several definitions of duplicate:

1. Order of name parts
   Bill Gates vs Gates Bill

2. Missing middle name
   Bill Gates vs Bill Henry Gates

3. Short or full variations of the first name
   Bill Gates vs William Gates

4. Misspelling of the last name
   Bill Gates vs Bill Gats

So all names from the above examples should be counted as duplicates. Note that there can be combinations of the above as well e.g. *Bill Gates vs William Henry Gatez*

The list of first names can be taken from here
The misspellings can be detected using Levenshtein distance. If it is not possible to detect which part of the name is a first name then the Levenshtein distance should be used on the full name.

Test input list may be found here. The output should be a list of tuples of duplicate names. The code should be able to process thousands of names in very short time.

What we will look at:
Code quality/elegancy/extendibility
Performance
Documentation
Test coverage