

## Übungsblatt 4

Abgabe für ESE: bis Dienstag, den 20. November um 16:00 Uhr

Abgabe für IEMS: bis Dienstag, den 4. Dezember um 16:00 Uhr

### Aufgabe 1 (10 Punkte)

Schreiben Sie eine Klasse *GeoNamesAnalyzer* mit folgenden Methoden:

1. Eine Methode *readInfoFromFile*, die alle für das Folgende relevanten Informationen aus der auf dem Wiki verlinkten GeoNames-Datei *allCountries.txt* einliest.
2. Eine Methode *computeMostFrequentCityNamesUsingMap*, die die weltweit am häufigsten vorkommenden Ortsnamen ermittelt, sowie für jeden dieser Ortsnamen die Information, in welchen Ländern wie oft. Verwenden Sie dabei ein assoziatives Array.
3. Eine Methode *computeMostFrequentCityNamesUsingSorting*, die genau das selbe berechnet, aber mittels Sortieren.

Als Orte gelten alle Zeilen in *allCountries.txt* mit einem *P* in Spalte 7. Entnehmen Sie den Ortsnamen der Spalte 2. Betrachten Sie nur Orte mit  $> 0$  Einwohnern (Spalte 15). Der Ländercode steht in Spalte 9.

Schreiben Sie für die beiden *compute...* Methoden jeweils einen Unit Test.

### Aufgabe 2 (10 Punkte)

Schreiben Sie ein Programm *GeoNamesAnalyzerMain*, welches die drei weltweit am häufigsten vorkommenden Ortsnamen berechnet + die oben beschriebene Zusatzinfo (in welchen Ländern wie oft). Vergleichen Sie die Laufzeiten Ihrer beiden *compute...* Methoden für diese Aufgabe.

Ändern Sie dann eine der beiden Methoden (oder beide) so ab, dass Sie die drei weltweit am häufigsten Ortsnamen berechnen (samt Zusatzinfo), die auch *mindestens einmal* in Deutschland (Ländercode DE) vorkommen.

Notieren Sie Ihre Ergebnisse, sowie eine kurze(!) Diskussion dazu in der Datei *erfahrungen.txt* (zusammen mit dem üblichen Feedback, siehe unten).

Committen Sie, wie gehabt, Ihren Code in das SVN, in einen neuen Unterordner *uebungsblatt\_04*, sowie, ebendort, Ihr Feedback in einer Textdatei *erfahrungen.txt*. Insbesondere: Wie lange haben Sie ungefähr gebraucht? An welchen Stellen gab es Probleme und wieviel Zeit hat Sie das gekostet?