

Homework 1b

1) Britney Spears

A while back Google published a long list of the many ways to misspell the name Britney Spears to highlight the power of their spelling corrector. The correct spelling Britney Spears appeared 488,941 times.

The pattern I have drafted below matches almost all of the first column of Britney misspellings here: (<https://archive.google.com/jobs/britney.html>), aside from the really silly ones (Pritney, Britney's, etc.).

```
/b[yieaourt]{1,3}[ndt]{1,3}h?[eia]{0,2}n?[aei]?[nmr]{0,2}[eytu]{0,2}[s]? spears/
```

2) Porter Stemmer

“Engineers” into “engin”:

Step 1: Remove final “s”: Step 1ab, lines 174 - 180 (specifically 179-180). This gives us *engineer*.

Step 2: Remove “-er”: Step 4, lines 263 - 265. This gives us *engine*.

Step 3: Remove now-final “e.” Step 5, lines 310 - 313. This gives us *engin*.

“Continuously” into “continu”:

Step 1: Turn terminal “y” into “i”: Step 1c, lines 197-200. This gives us *continuousli*.

Step 2: Remove “li”: Step 2, line 222. This gives us *continuous*.

Step 3: Remove “-ous”: Step 4, lines 291 - 293. This gives us *continu*.

3) Date Regex: See ZIP file