

RUSH 2

Arthur DECAEN, Johan KOZAKEVICH

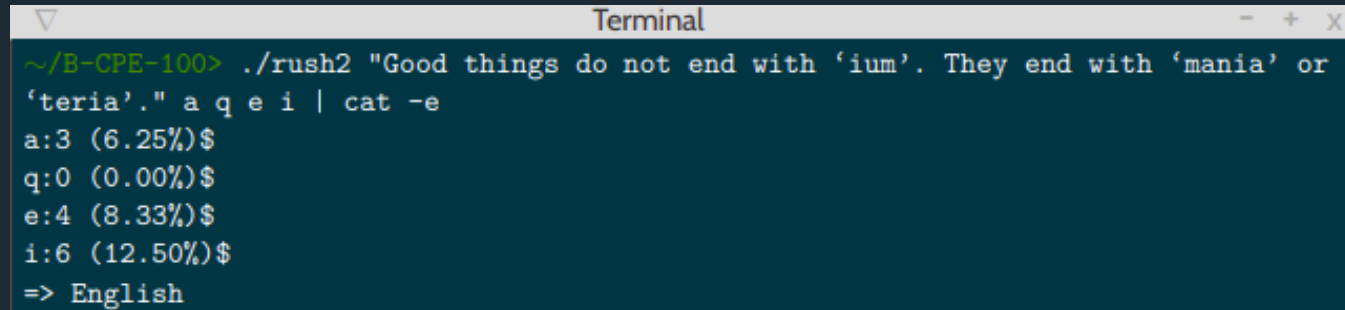
Consigne

WHAT LANGUAGE IS THIS?

The objective of this rush is to code a program that automatically guesses the language in which a text is written.

It is divided into 4 progressive steps.

Résultat attendu



```
~/B-CPE-100> ./rush2 "Good things do not end with 'ium'. They end with 'mania' or  
'teria'." a q e i | cat -e  
a:3 (6.25%)$  
q:0 (0.00%)$  
e:4 (8.33%)$  
i:6 (12.50%)$  
=> English
```

Algorithme (1/2)

Valeurs de chaque lettre (cf. Wikipedia)

{816, 149, 278, 425, 1270, 222, 201, 609, 696, 15, 77, 402, 240, 674, 750, 192, 9, 598, 632, 905, 275, 97, 236, 15, 197, 7}	English
{763, 90, 326, 366, 1471, 106, 86, 73, 752, 61, 7, 545, 296, 709, 579, 252, 136, 669, 794, 224, 631, 183, 4, 42, 12, 32}	French
{651, 188, 273, 507, 1639, 165, 300, 457, 655, 26, 141, 343, 253, 977, 259, 67, 1, 700, 727, 615, 416, 84, 192, 3, 3, 113}	German
{1152, 221, 401, 501, 1218, 69, 176, 70, 624, 49, 1, 496, 315, 671, 868, 251, 87, 687, 797, 463, 292, 113, 1, 21, 100, 46}	Spanish
{0, 0}	Liste vide
{1550, 0, 0, 120, 0, 0, 0, 0, 0, 0, 0, 0, 254, 0, 0, 0, 0, 0, 2900, 0, 0, 0, 0, 0, 0}	Liste pleine

Pourcentages calculés dans le programme

- * Somme de la différence de chaque lettre pour toutes les langues.
- * Le chiffre le plus bas correspond à la langue la plus probable.

Algorithme (2/2)

Somme de la slide précédente

English	French	German	Spanish
755	209	350	958

Bubble sort pour mettre le plus petit à gauche

French	German	English	Spanish
209	350	755	958

French
209

Langue la plus proche