**Plateformes et Langages de Programmation : Partie Hadoop – Map Reduce**

Groupe : Adam TOUGHZA, Iliass CHEKOUR, Youssef EL HARIF

**Répartition des questions :**

2.7) Youssef El Harif

2.8) Youssef El Harif, Iliass Chekour

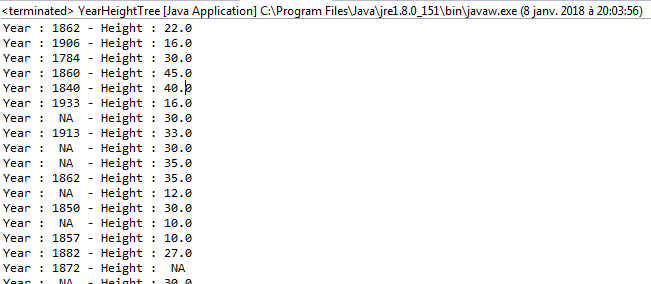
5.1) Adam Toughza

5.2) Adam Toughza

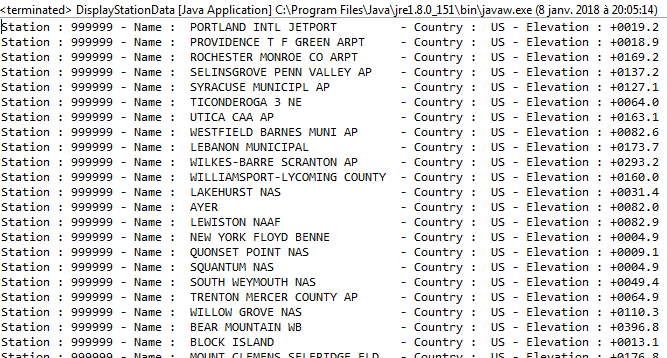
5.3) Iliass Chekour

**Résultats obtenus :**

**2.7)**

****

**2.8)**

****

**5.1)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Rank** | **Word** | **File** | **TFIDF** |
| 1 | buck | callwild.txt | 0,005259 |
| 2 | dogs | callwild.txt | 0,001724 |
| 3 | thornton | callwild.txt | 0,001490 |
| 4 | myself | defoe-robinson-103.txt | 0,001267 |
| 5 | spitz | callwild.txt | 0,000950 |
| 6 | sled | callwild.txt | 0,000920 |
| 7 | francois | callwild.txt | 0,000877 |
| 8 | friday | defoe-robinson-103.txt | 0,000803 |
| 9 | trail | callwild.txt | 0,000599 |
| 10 | john | callwild.txt | 0,000584 |
| 11 | perrault | callwild.txt | 0,000570 |
| 12 | hal | callwild.txt | 0,000541 |
| 13 | team | callwild.txt | 0,000497 |
| 14 | thoughts | defoe-robinson-103.txt | 0,000482 |
| 15 | sol | callwild.txt | 0,000424 |
| 16 | leks | callwild.txt | 0,000409 |
| 17 | traces | callwild.txt | 0,000409 |
| 18 | ice | callwild.txt | 0,000409 |
| 19 | around | callwild.txt | 0,000380 |
| 20 | dave | callwild.txt | 0,000365 |

**5.2)** Le calcul final du Page Rank n’était pas clair, ce qui nous a empêché de finaliser le code. On a bien codé la partie « map » qui renvoie les pairs <IDsource, IDtarget> mais on n’a pas finalisé la partie « reduce ».

**5.3)**

|  |  |  |
| --- | --- | --- |
| **Specie** | **Max Height** | **Count** |
| araucana | 9.0 | 1 |
| atlantica | 25.0 | 2 |
| australis | 16.0 | 1 |
| baccata | 13.0 | 2 |
| bignonioides | 15.0 | 1 |
| biloba | 33.0 | 5 |
| bungeana | 10.0 | 1 |
| cappadocicum | 16.0 | 1 |
| carpinifolia | 30.0 | 4 |
| colurna | 20.0 | 3 |
| coulteri | 14.0 | 1 |
| decurrens | 20.0 | 1 |
| dioicus | 10.0 | 1 |
| distichum | 35.0 | 3 |
| excelsior | 30.0 | 1 |
| fraxinifolia | 27.0 | 2 |
| giganteum | 35.0 | 5 |
| giraldii | 35.0 | 1 |
| glutinosa | 16.0 | 1 |
| grandiflora | 12.0 | 1 |
| hippocastanum | 30.0 | 3 |
| ilex | 15.0 | 1 |
| involucrata | 12.0 | 1 |
| japonicum | 10.0 | 1 |
| kaki | 14.0 | 2 |
| libanii | 30.0 | 2 |
| monspessulanum | 12.0 | 1 |
| nigra | 30.0 | 3 |
| nigra laricio | 30.0 | 1 |
| opalus | 15.0 | 1 |
| orientalis | 34.0 | 8 |
| papyrifera | 12.0 | 1 |
| petraea | 31.0 | 2 |
| pomifera | 13.0 | 1 |
| pseudoacacia | 11.0 | 1 |
| sempervirens | 30.0 | 1 |
| serrata | 18.0 | 1 |
| stenoptera | 30.0 | 1 |
| suber | 10.0 | 1 |
| sylvatica | 30.0 | 8 |
| tomentosa | 20.0 | 2 |
| tulipifera | 35.0 | 2 |
| ulmoides | 12.0 | 1 |
| virginiana | 14.0 | 2 |
| x acerifolia | 45.0 | 11 |