

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)

<terminated> YearHeightTree [Java Application] C:\Program Files\Java\jre1.8.0_151\bin\javaw.exe (8 janv. 2018 à 20:03:56)

```
Year : 1862 - Height : 22.0
Year : 1906 - Height : 16.0
Year : 1784 - Height : 30.0
Year : 1860 - Height : 45.0
Year : 1840 - Height : 40.0
Year : 1933 - Height : 16.0
Year : NA - Height : 30.0
Year : 1913 - Height : 33.0
Year : NA - Height : 30.0
Year : NA - Height : 35.0
Year : 1862 - Height : 35.0
Year : NA - Height : 12.0
Year : 1850 - Height : 30.0
Year : NA - Height : 10.0
Year : 1857 - Height : 10.0
Year : 1882 - Height : 27.0
Year : 1872 - Height : NA
Year : NA - Height : 30.0
```

2.8)

<terminated> DisplayStationData [Java Application] C:\Program Files\Java\jre1.8.0_151\bin\javaw.exe (8 janv. 2018 à 20:05:14)

```
Station : 999999 - Name : PORTLAND INTL JETPORT - Country : US - Elevation : +0019.2
Station : 999999 - Name : PROVIDENCE T F GREEN ARPT - Country : US - Elevation : +0018.9
Station : 999999 - Name : ROCHESTER MONROE CO ARPT - Country : US - Elevation : +0169.2
Station : 999999 - Name : SELINGSGROVE PENN VALLEY AP - Country : US - Elevation : +0137.2
Station : 999999 - Name : SYRACUSE MUNICIPL AP - Country : US - Elevation : +0127.1
Station : 999999 - Name : TICONDEROGA 3 NE - Country : US - Elevation : +0064.0
Station : 999999 - Name : UTICA CAA AP - Country : US - Elevation : +0163.1
Station : 999999 - Name : WESTFIELD BARNES MUNI AP - Country : US - Elevation : +0082.6
Station : 999999 - Name : LEBANON MUNICIPAL - Country : US - Elevation : +0173.7
Station : 999999 - Name : WILKES-BARRE SCRANTON AP - Country : US - Elevation : +0293.2
Station : 999999 - Name : WILLIAMSPORT-LYCOMING COUNTY - Country : US - Elevation : +0160.0
Station : 999999 - Name : LAKEHURST NAS - Country : US - Elevation : +0031.4
Station : 999999 - Name : AYER - Country : US - Elevation : +0082.0
Station : 999999 - Name : LEWISTON NAAF - Country : US - Elevation : +0082.9
Station : 999999 - Name : NEW YORK FLOYD BENNE - Country : US - Elevation : +0004.9
Station : 999999 - Name : QUONSET POINT NAS - Country : US - Elevation : +0009.1
Station : 999999 - Name : SQUANTUM NAS - Country : US - Elevation : +0004.9
Station : 999999 - Name : SOUTH WEYMOUTH NAS - Country : US - Elevation : +0049.4
Station : 999999 - Name : TRENTON MERCER COUNTY AP - Country : US - Elevation : +0064.9
Station : 999999 - Name : WILLOW GROVE NAS - Country : US - Elevation : +0110.3
Station : 999999 - Name : BEAR MOUNTAIN WB - Country : US - Elevation : +0396.8
Station : 999999 - Name : BLOCK ISLAND - Country : US - Elevation : +0013.1
Station : 999999 - Name : MOUNT CLEMENS SELEDTGE FLD - Country : US - Elevation : +0176.8
```

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 ».