

Devoir 1 : 8INF803 Bases de données réparties

Lien GitHub : https://github.com/AdrienRybarczyk/TP1_8INF803

Exercice 1

Il a fallu tout d'abord récupérer les éléments de la page web via Jsoup.

Pour le nom du sort, il est dans le heading de la page web.

Pour les autres éléments, ils sont tous dans la classe SPDet et il a fallu effectuer un traitement via la fonction split pour récupérer uniquement les informations nécessaires.

Pour les niveaux de sort, j'ai choisi de mettre à 5 si le sort n'est pas pour les sorciers. Ainsi, il ne sera pas compté dans la recherche des sorts pour Pito.

Pour les composants, il a fallu les remettre dans un tableau car il peut en avoir plusieurs.

Pour le spell resistance, il faut regarder le dernier élément de la classe SPDet. Si celui-ci contient un spell resistance on regarde la valeur sinon on la met à faux.

Enfin, pour récupérer uniquement les sorts utilisables par Pito, il a fallu filtrer sur les conditions de niveau et de composant.

Pour la recherche des sorts avec Spark DataFrame, il a fallu caster le RDD de Sort obtenu précédemment en RDD de Row. Il a fallu aussi créer un StructType qui correspondra aux différentes colonnes de la DataFrame. Pour obtenir le résultat, il a fallu faire une requête semblable à du SQL.

```

Blindness/Deafness Mythic Blindness/Deafness : 2 : ArrayBuffer(V) : true
Blur : 2 : ArrayBuffer(V) : true
Dimension Door : 4 : ArrayBuffer(V) : false
Feather Fall Mythic Feather Fall : 1 : ArrayBuffer(V) : true
Flare : 0 : ArrayBuffer(V) : true
Geas, Lesser : 4 : ArrayBuffer(V) : true
Hold Portal Mythic Hold Portal : 1 : ArrayBuffer(V) : false
Knock : 2 : ArrayBuffer(V) : false
Shout Mythic Shout : 4 : ArrayBuffer(V) : true
Dweomer Retaliation : 3 : ArrayBuffer(V) : true
Emergency Force Sphere : 4 : ArrayBuffer(V) : false
Liberating Comand : 1 : ArrayBuffer(V) : true
Chastise : 1 : ArrayBuffer(V) : false
Sotto Voce : 0 : ArrayBuffer(V) : true
Flare Burst : 1 : ArrayBuffer(V) : true
Steal Voice : 2 : ArrayBuffer(V) : true
Liberating Command : 1 : ArrayBuffer(V) : true
Buoyancy : 2 : ArrayBuffer(V) : true
Total resultat Crawler 18
*****

```

```

[Blindness/Deafness Mythic Blindness/Deafness]
[Blur]
[Dimension Door]
[Feather Fall Mythic Feather Fall]
[Flare]
[Geas, Lesser]
[Hold Portal Mythic Hold Portal]
[Knock]
[Shout Mythic Shout]
[Dweomer Retaliation]
[Emergency Force Sphere]
[Liberating Comand]
[Chastise]
[Sotto Voce]
[Flare Burst]
[Steal Voice]
[Liberating Command]
[Buoyancy]

```

```
Total resultat version Spark DataFrame 18
```

Exercice 2

Les explications du code sont en commentaires dans le code.

```
*****  
Itération numéro : 18  
(PageA,1.4902795873742496)  
(PageB,0.783235232500263)  
(PageC,1.5764851801254862)  
(PageD,0.15)  
*****  
Itération numéro : 19  
(PageA,1.490012403106663)  
(PageB,0.7833688246340561)  
(PageC,1.5766187722592795)  
(PageD,0.15)  
*****  
Itération numéro : 20  
(PageA,1.4901259564203875)  
(PageB,0.7832552713203318)  
(PageC,1.5766187722592795)  
(PageD,0.15)  
END
```