

TP 1 : MapReduce



Voici le schéma de la collection sur lequel vous allez travailler :

```
{
  "firstname": String,
  "lastname": String,
  "login": String,
  "password": String,
  "lastConnectionDate": Date,
  "website" : String,
  "articles": [
    {
      "title": String,
      "content": String,
      "date": String,
      "note": Number,
      "category" : String
    }
  ]
}
```

Il s'agit d'une collection d'utilisateurs avec la liste des articles qu'ils ont postés. Un utilisateur peut avoir entre 0 et 50 articles et la collection contient 20000 enregistrements.

Pour pouvoir travailler deans de bonnes conditions, je vous conseille d'avoir mongoDB 3.2+ d'installé sur vos machines et de passer par un gestionnaire graphique comme mongobooster ou robomongo.

Bon à savoir : Dans le cour, j'ai oublié de préciser que les fonctions de map et de reduce doivent absolument avoir un retour structuré de la même façon pour fonctionner (fields et type). Si vous voulez transformer le résultat final différemment de votre `emit()` , vous devrez utiliser la méthode `finalize()` .

Grâce à mapReduce, réalisez les requêtes suivantes :

Attention : Chaque résultat sera à mettre dans une collection différente à chaque fois.

1. Lister combien il y a d'utilisateur par site
2. Lister combien il y a d'articles par site
3. Lister la note moyenne par site (attention besoin de passer par finalize)
4. Lister le nombre d'articles par catégories
5. Lister le nombre d'articles par catégories pour le site `JV.com`
6. Lister la note moyenne par catégorie (finalize)
7. Lister le nombre d'articles de chaque site pour la catégorie `Vulgarisation`
8. Lister le nombre d'articles postés par année
9. Lister, par site, combien d'articles il y a dans chaque catégorie
10. Lister, par site, et par catégories, combien il y a d'articles chaque année
11. Lister pour chaque site quels sont les 5 articles les plus récents.
12. Lister, pour chaque catégorie quels sont les 5 articles les mieux notés de 2017 par ordre alphabétique.

13. Passage en mode `reduce` :

- Reprenez la requête de la question 5
- Effectuez la requête pour chaque site et faites un `reduce` sur la collection cible afin de retomber sur le même résultat que la question 4.