

Atividade Prática – HBase

Em sua conta github, adicione uma pasta com nome hbase. Ali, você deve colocar os fontes e um README.md detalhando a atividade.

Inicializado e acessando o HBase

Por ser um banco mais complexo em termos de sistemas distribuídos, o HBase exige algumas configurações para ser inicializado. Como isso varia de SO, a melhor forma de utilizar esse banco é via Docker.

Baixe o Docker em seu Sistema Operacional e em seguida utilize a imagem do HBase do [Dajobe](#) da seguinte forma:

```
docker pull dajobe/hbase
```

Em seguida, inicie a imagem com o nome hbase-furb pois assim ficará fácil acessar a imagem depois. Caso contrário o Docker gera nomes aleatórios.

```
docker run --name hbase-furb dajobe/hbase
```

Através de outro terminal, acesse a VM da seguinte forma:

```
docker exec -it hbase-furb /bin/bash
```

Esse comando irá fazer você acessar a imagem Docker como se estivesse estabelecido uma conexão SSH. Para a execução dos comandos via terminal, é uma das formas mais simples.

Através do acesso ao ambiente do HBase, é possível acessar shell do banco com

```
hbase shell
```

Aquecendo com alguns dados

Importe o arquivo `italians.txt` do repositório [Downloads NoSQL FURB](#). Em seguida, carregue o mesmo no hbase. Para isso, você precisa enviar o arquivo para a imagem Docker primeiro.

```
docker cp italians.txt hbase-furb:/tmp
```

Agora, de dentro da imagem faça os seguintes procedimentos:

1. Crie a tabela com 2 famílias de colunas:
 - a. `personal-data`
 - b. `professional-data`
2. Importe o arquivo via linha de comando

Agora execute as seguintes operações:

1. Adicione mais 2 italianos mantendo adicionando informações como data de nascimento nas informações pessoais e um atributo de anos de experiência nas informações profissionais;
2. Adicione o controle de 5 versões na tabela de dados pessoais.
3. Faça 5 alterações em um dos italianos;
4. Com o operador `get`, verifique como o HBase armazenou o histórico.
5. Utilize o `scan` para mostrar apenas o nome e profissão dos italianos.
6. Apague os italianos com row id ímpar
7. Crie um contador de idade 55 para o italiano de row id 5
8. Incremente a idade do italiano em 1

Os italianos voltaram!

Com base no script de italianos utilizado no exercício de MongoDB, altere ele para uma versão do HBase. Você pode transformar ele para Ruby ou gerar os comandos para HBase a partir do JavaScript mesmo.

Modifique para inserir os dados com a seguinte estrutura:

Família de coluna: personal-data

firstname
surname
age
bloodType
city
father
mother

Família de colunas: app-data

id_num,
email
registerDate
ticketNumber
username

Família de Colunas: professional-data

jobs
salary

Família de Colunas: miscellaneous

favFruits
movies
cat
dog

Utilizando o MapReduce, faça um relatório ou mais de um indicando as seguintes informações:

1. Quantidade de gatos e cachorros na amostra
2. Média de gatos e cachorros na população
3. Quantidade de pais e mães
4. Média de pais e mães
5. Média de frutas e filmes por italiano
6. Salário médio dos italianos