Module 7 – MongoDB Security

Case Study - 2

edureka!

edureka!

Business Requirement

You are working on Contract for a Grocery Store Company which is in middle east. Products are in millions and your task is to run various analytics operations on mongo.

In this case study, you will understand how you can integrate Mongo and Hadoop.

You have given Mongo scheme details along with mongo connection information. [In the previous case studies, you must have worked on product catalog schema]

There are 2 ways you can do operations on Mongo:

1. Mongo database
2. Mongo bson File

Both are important because when we don't need data in mongo, we achieve that data and you have to run Hadoop job on top of it.

Your task is to find out product count group by Brand and save back to Mongo database.

For example:

Output will be like this:

{ "_id" : "Advice", "value" : 1 }

{ "_id" : "Delux Look", "value" : 1 }

{ "_id" : "Granite", "value" : 1 }

As you can see it contains a brand and count of products for that brand.

You can perform below mentioned operations using Hadoop

1. Category wise product count
2. Seller code wise product count
3. Product have price greater then X

Products which are added in last 1 month.

Considerations

You have given various attributes of a Product.

1. PogId
2. Supc
3. Brand
4. Description
5. Size
6. Category
7. Sub Category
8. Price

9. Quantity
10. Country
11. Seller Code
12. Creation time
13. Stock

Data Volume
1 million products

Focus – Mongo Hadoop Integration
1. Set up Mongo
2. For Bson use mongorestore command (you have already created grocery database , you have to use that db bson file)
3. Write Mapper and Reducer Job