## **Big Data Analytics | LAB 6**

Aim: Connecting to NoSQL database/s and querying to provide analysis using api like aggregation, etc. To be able to successfully import/export from/to CSV.

1. **Run JuPyter Notebook and using MongoDB kernel, do the same exercise shown in the screenshot by mongo client above.**

I have done this exercise in mongodb shell since I am not able to run imongo kernel in jupyter notebook.

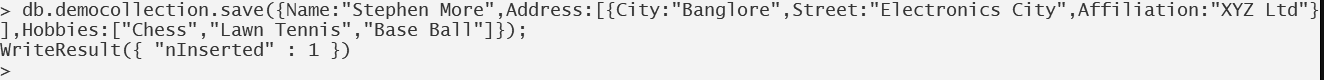


1. **Write the insert method to store the following document in MongoDB.**

Name : "Stephen More"

Address : { "City" : "Banglore", "Street" : "Electronics City", "Affiliation" : "XYZ Ltd" } Hobbies : Chess, Lawn Tennis, Base Ball

Step 1: In mongodb shell write following script.



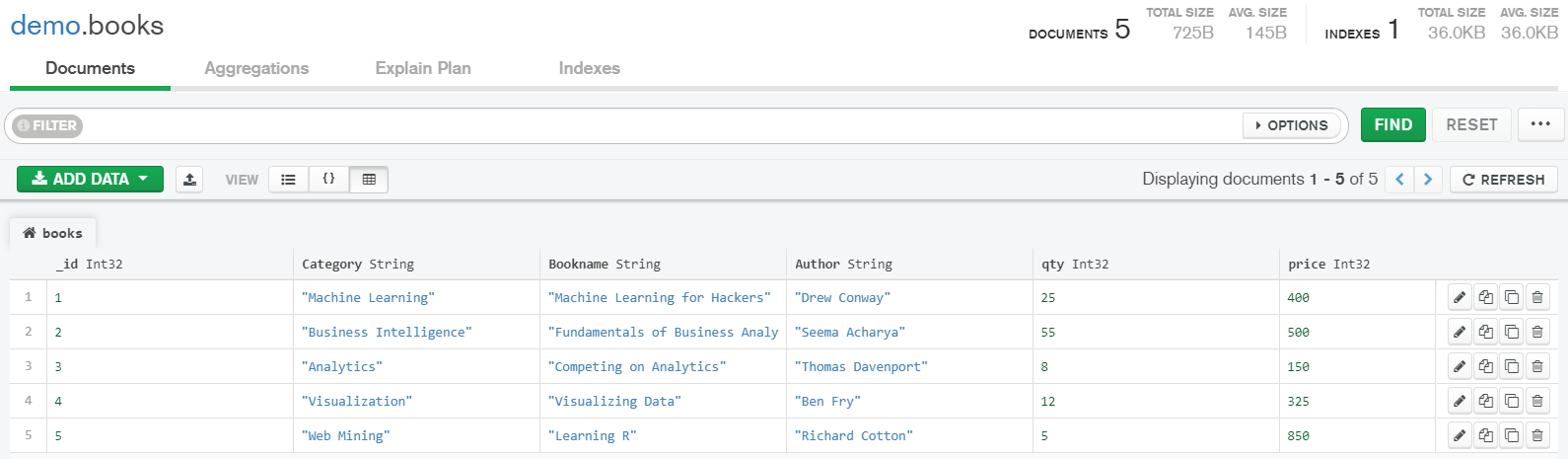
Step 2: Confirm inserted data using below command.



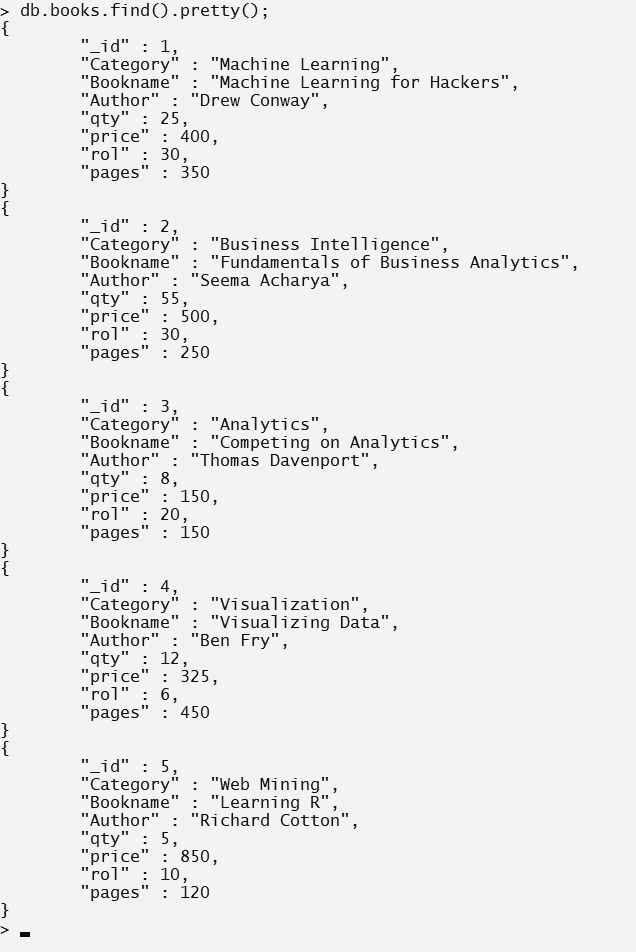
1. **To practice Map Reduce programming in MongoDB.**

Step 3.1: Insert 5 documents as shown below in a collection named 'books'.

Using mongodb compass insert the given data in books collection.

****

Step 3.2:Confirm presence of data using mongodb shell.



Step. 3.3: Write map and reduce functions to split the books into the following two categories:

(a) Big Books

(b) Small Books

Books which have more than 300 pages should be in the big book category. Books which have less than 300 pages should be in the small book category.

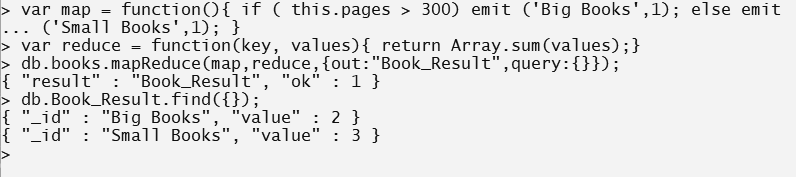
Step. 3.4: Count the number of books in each category.

Step 3.5: Store the output as follows as documents in a new collection, called “Book\_Result”. Book Category Count of the books.

“Big Books" 2

"Small Books" 3

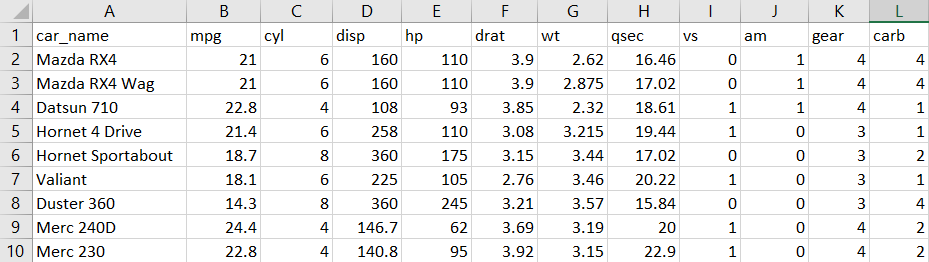
All Steps are shown in the below screenshot.



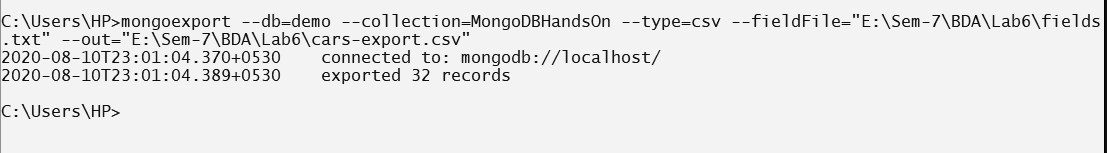
1. **To practice import, export and aggregation in MongoDB.**

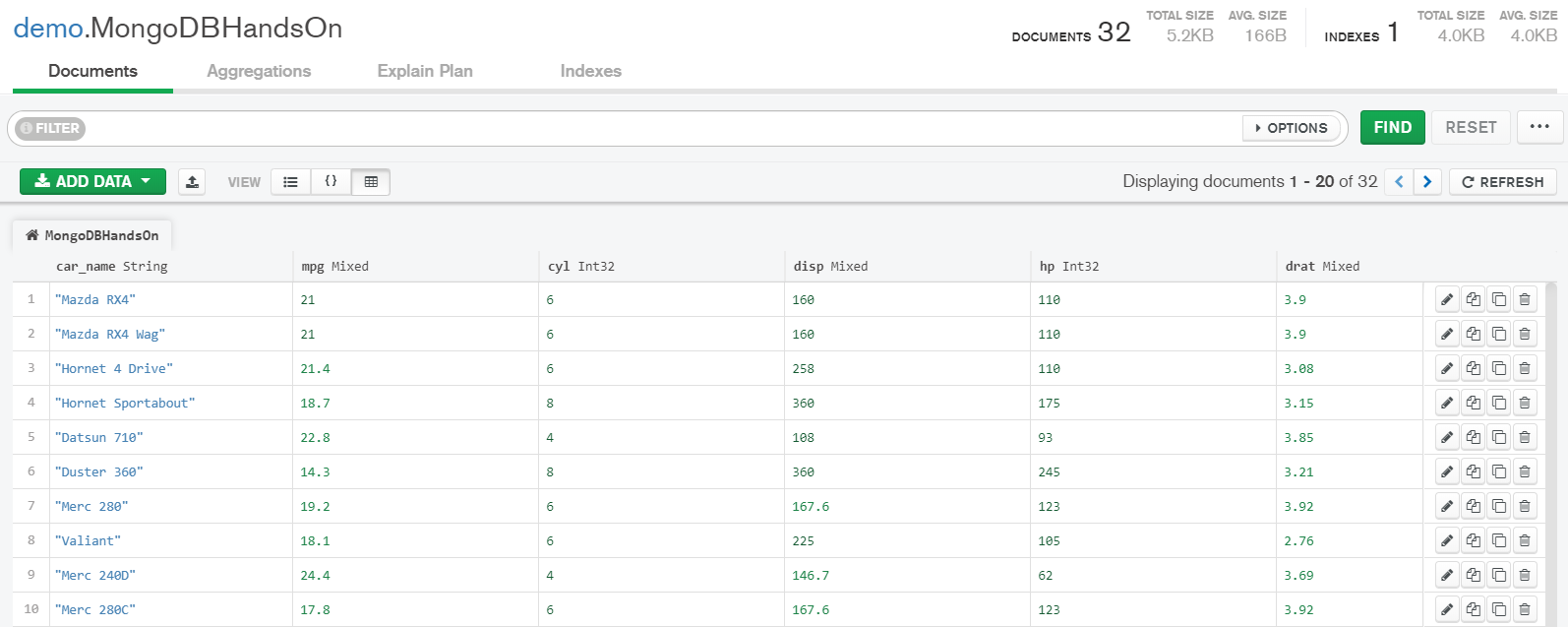
Step 4.1 Pick and public dataset from the site www.kdnuggets.com Convert it into CSV format. Make sure that you have at least two numeric columns.

I have downloaded cars.csv data from github.

****

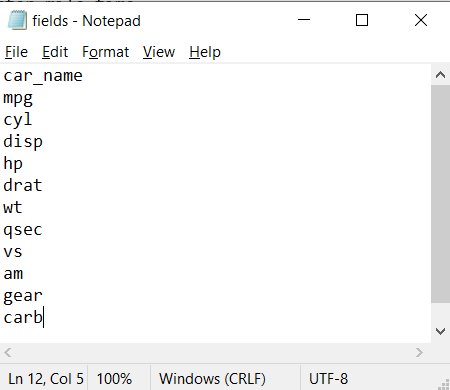
Step 4.2 Use MongoImport to import data from the CSV format file into MongoDB collection name “MongoDBHandsOn” in the test database.





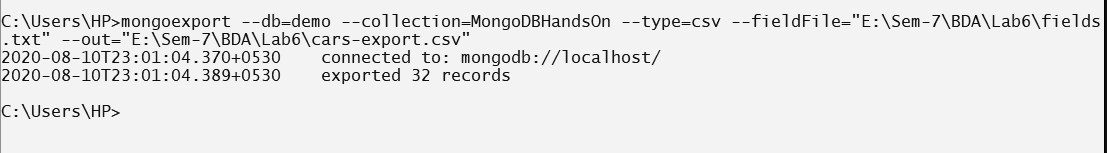
Step 4.3: Identifying a grouping column.

Create the fields.txt as shown below.

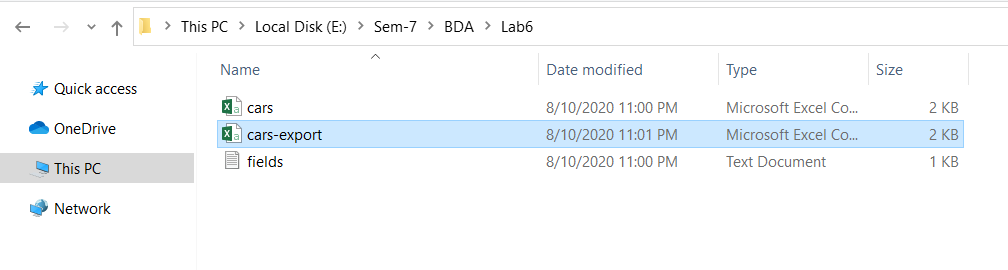


Type following command in cmd.

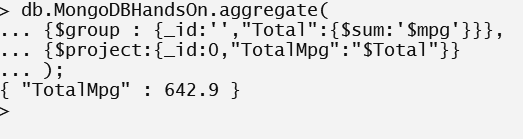
This will create cars-export.csv file using monogexport utility with fields.txt and data which was available in MongoDBHandsOn.



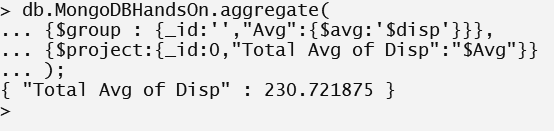
Confirm presence of exported file.



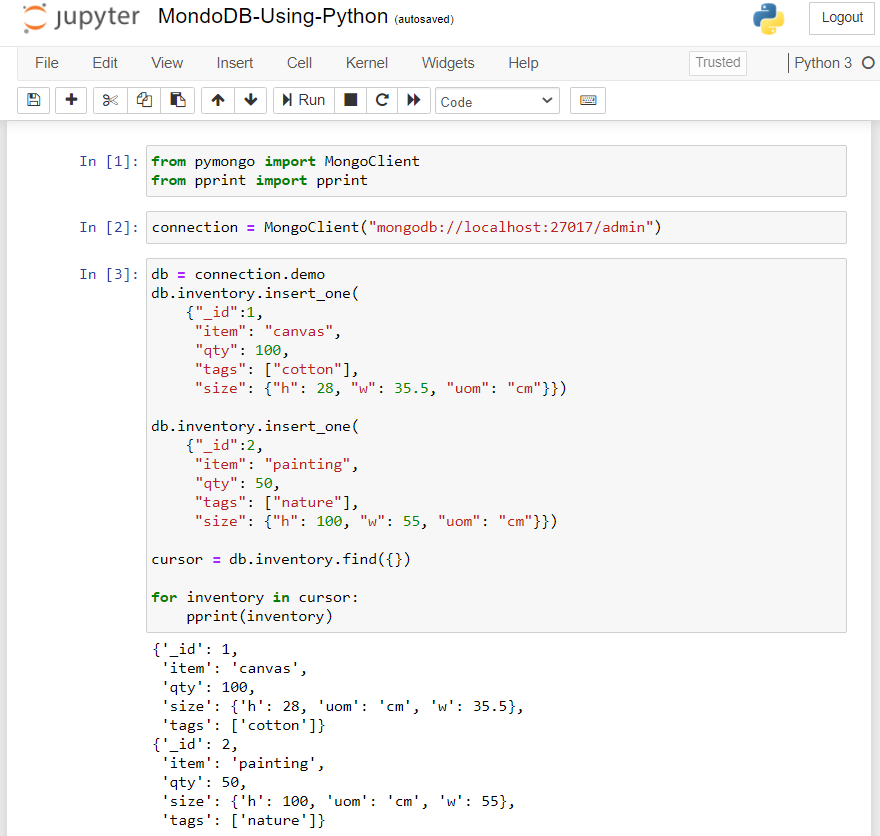
Step 4.4: Compute the sum of the values in the first numeric column.

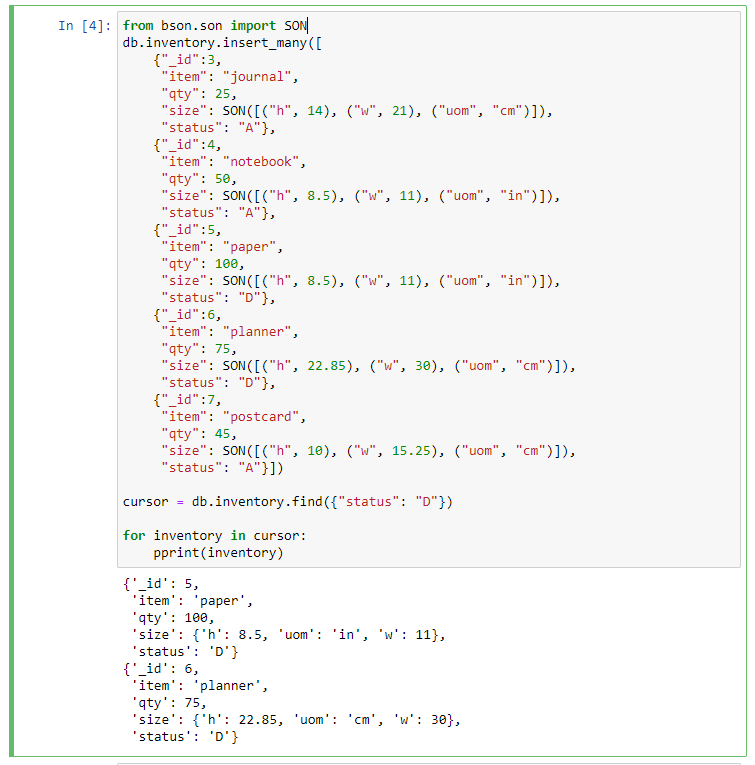


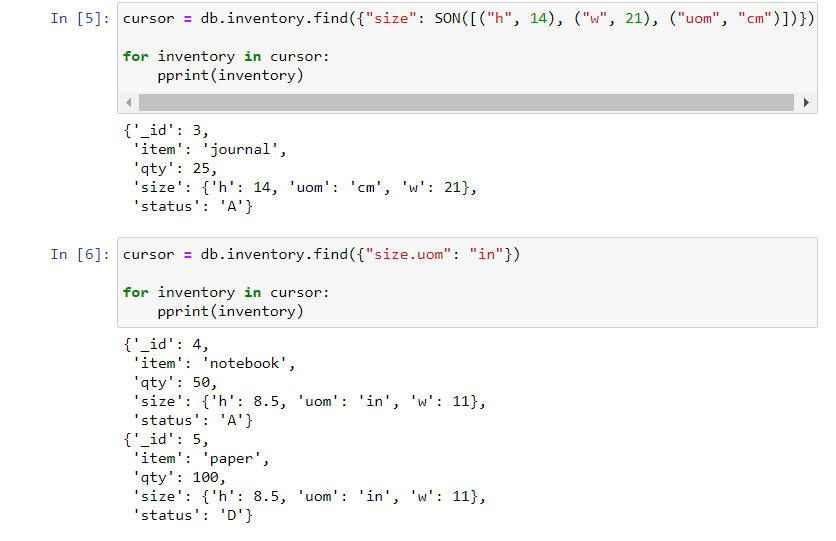
Step 4.5: Compute the average of the values in the second numeric column.



1. **Exercise Python to MongoDB Connectivity using JuPyter Notebook.**







**Document Link :**

[**https://docs.google.com/document/d/1u\_uErurBoF1xKtadyl6YSNZi7j-3RuB7b6w11daXhmQ/edit?usp=sharing**](https://docs.google.com/document/d/1u_uErurBoF1xKtadyl6YSNZi7j-3RuB7b6w11daXhmQ/edit?usp=sharing)