INF 551 – MongoDB Lab

100 points
Due: 11/1/2019, Friday

Consider the New York Times best seller dataset (https://www.kaggle.com/cmenca/new-york-times-hardcover-fiction-best-sellers/). Import the dataset: nyt2.json, into your MongoDB database.

Write a MongoDB operation (find, aggregate, update, etc) for each of the following questions:

1. Find out the number of books whose title contains "history" (case insensitive).
2. Find out how many book were ranked numbered 1 (according to the overall "rank" value).
3. Find out the highest price of the books whose overall rank is top 20 (rank value is from 1 to 20). (Hint: use aggregate() with a match() for finding the qualified books and then $group, setting _id set to null, to find out the max value of price among all matched books).
4. Find out, for each publisher, the number of books it published. (list 20 of the query result in .txt)
5. Find out publishers who publish at least 10 books. (list 20 of the query result in .txt)
6. Find out the number of distinct publishers.
7. Find the titles of books published by "Harper" and appeared in the best-selling list at least 5 times. Output the titles in the descending order. (list first 10 of the query result in .txt)
8. Find out the average price of books published by "Harper".
9. Find the most productive authors (i.e., authors who published the largest number of books).
10. Change the price of book "Breathless" to 28.5. (paste the response after the update in .txt)

Submit a zip file that contains your queries and results.

Submission

1. Name your files as <firstname>_<lastname>_q1.txt, <firstname>_<lastname>_q2.txt, and so on.

2. Submit all .txt files in **a zip file** with the naming convention: <firstname>_<lastname>_lab3.zip.

(example content of .txt files)

```
[Example]your query for the question(no need to write this line)
db.<your collection>.find({'title':{$ne:''}},{title:1}).limit(10)

[Example]paste your result here(no need to write this line)
{ "_id" : ObjectId("5b4aa4ead3089013507db18b"), "title" : "ODD HOURS" }
{ "_id" : ObjectId("5b4aa4ead3089013507db18c"), "title" : "THE HOST" }
{ "_id" : ObjectId("5b4aa4ead3089013507db18d"), "title" : "LOVE THE ONE YOU'RE WITH" }
{ "_id" : ObjectId("5b4aa4ead3089013507db18e"), "title" : "THE FRONT" }
{ "_id" : ObjectId("5b4aa4ead3089013507db18f"), "title" : "SNUFF" }
{ "_id" : ObjectId("5b4aa4ead3089013507db190"), "title" : "SUNDAYS AT TIFFANY'S" }
{ "_id" : ObjectId("5b4aa4ead3089013507db191"), "title" : "PHANTOM PREY" }
{ "_id" : ObjectId("5b4aa4ead3089013507db192"), "title" : "SWINE NOT?" }
{ "_id" : ObjectId("5b4aa4ead3089013507db193"), "title" : "CARELESS IN RED" }
{ "_id" : ObjectId("5b4aa4ead3089013507db194"), "title" : "THE WHOLE TRUTH" }
```