COMP 306: Database Management Systems

Fall 2023 - Homework 4

In this homework assignment, you need to install MongoDB and complete the following tasks using MongoDB's command line interface (similar to what we did in the PS).

Prepare and submit a report that contains your answers to each question. For each question, you should include (i) the query/commands you used, (ii) screenshots of the results when you executed your queries/commands. Your screenshots should be legible.

Question 1) [5 pts] Switch to a new empty database and import the provided "zips.json" file to the new database. The database must be called "hw4" and the newly imported collection must be called "zipcodes". Show all the collections in the new database and show the first 15 documents in the zipcodes collection. Include all the commands you used and their outputs in your report.

Question 2) [10 pts] Write a query to find documents with population above 50,000 in the state of "CA" and latitude greater than 35. Limit the result to a maximum of 5 documents and sort them in descending order of population.

Longitude is the first index of the "loc" array, latitude is the second. Please read this documentation: https://www.mongodb.com/docs/manual/tutorial/query-arrays/.

Question 3) [10 pts] Write a query to find documents that satisfy the following two conditions:

- State is not "CA".
- Longitude is smaller than -120, or latitude is smaller than 40.

Limit the result to a maximum of 5 documents.

Question 4) [10 pts] Write an aggregation query to find the 5 most populated cities with their respective populations. For example: (Chicago, 2 million), (Los Angeles, 1.8 million), … The output must be ordered according to the descending order of populations.

Question 5) [15 pts] Write an aggregation query to find states that have higher than 300 and less than 500 zip codes associated with them. Display the state codes and their respective count of zip codes in ascending order of counts.

Question 6) [15 pts] We have to design a schema for a "customers" collection. The conditions of the schema are as follows:

- Each customer must have "name", "zipcode" and "avg_rating" fields.
- "zipcode" and "name" must be strings.
- "avg_rating" must be a double between 0.0 and 10.0

- “last_order” field must be an object containing an integer “year” and an array of strings called “tags” if it exists. If a document has this field, it must have the “year” field.

Include the command to create this collection with the specified schema validator and its output in your report.

Question 7) [10 pts] Now, we need to populate our new collection. Insert these 4 documents in one query:

- Name: <Your Name Surname> (Remove “<>”), Zip code: 99503, Avg. Rating: 8.3.
- Name: “Andrei T.”, Zip code: 90025, Avg. Rating: 3.5, Last Order: year 2009.
- Name: “Bela T.”, Zip code: 33126, Avg. Rating: 4.9, Last Order: year 2019, tags: art, melancholy.
- Name: “Nuri Bilge C.”, Zip code: 90010, Avg. Rating: 6.5, Last Order: year 3005.

After the insertion, show all documents in the collection.

Question 8) [5 pts] We have made a mistake. There are documents that contain an order after the year 2023, which is the current year. Correct this mistake by writing a deletion query to delete all documents containing an order after 2023. (You can’t look up the collection and use the matching name or id in your query, you have to specify the given condition.). After the deletion, show all documents in the collection.

Question 9) [5 pts] Let’s try to update the document with your name in the collection. Set its rating to 15. What is the query you execute? What is the response you receive from MongoDB? (Add a screenshot.) Why did you receive this response? Explain in 1-2 sentences.

Question 10) [10 pts] The new collection we have created (the “customers” collection) is related to the “zipcodes” collection since each customer has a zip code. Using the connection between the "zipcodes" and "customers" collections, write a query on the "customers" collection to display the Name, City, and State of each customer along with the corresponding zip code. Sort the results by the customer's name in ascending order.

Question 11) [5 pts] Export the “customers” collection to a JSON file called “customers.json”. (Use the “mongoexport” tool.) Add the command you used in your report. Include the customers.json file in your homework submission.

Submission

Your submission should consist of two files:

- A pdf report with the commands and screenshots
- The exported file in the last question