

Programming Assignment #3: CAS CS460

NoSQL Databases – (BONUS!)

Due: 12/11 11:59 pm on gradescope.

1. Introduction

In this assignment, you will use a NoSQL database system called MongoDB. You will load all data from JSON file.

More about MongoDB and JSON you can find [here](#) and [here](#).

2. Setup the Database

1. Download and install MongoDB. Instructions for macOS users can be found [here](#). Instructions for Ubuntu users can be found [here](#).

2. Download the restaurants dataset at:

<https://raw.githubusercontent.com/mongodb/docs-assets/primer-dataset/primer-dataset.json>

3. Use [mongoimport](#) to import the restaurants dataset into MongoDB.

(use something like this):

```
> mongoimport --db test --collection restaurants --drop --file primer-dataset.json
```

You may need to specify the full directory of the file.

Now you are ready to start the project. Notice that the dataset contains information about restaurants in New York City.

3. Queries in MongoDB

After importing the data, write queries in MongoDB to answer the following questions (try to use [aggregation](#) pipelines when possible): To access the MongoDB environment, just open your terminal and type “mongo”.

(1) Find the restaurant name and ID of “Caffe Dante”.

(2) Find all restaurants whose name has “Ice Cream” in it, return only the restaurant’s ids and names.

(3) Find the names of all restaurants that have a score above 5 and are in the Brooklyn borough.

(4) Return the top 10 boroughs based on the number of American restaurants they have.

(5) Find the worst American restaurant in Manhattan in terms of total score.

Hint: You can use “\$unwind”.

(6) Consider a rectangle area on the location field, in which the vertices are $[-74, 40.5]$, $[-74, 40.7]$, $[-73.5, 40.5]$ and $[-73.5, 40.7]$. Find the number of restaurants in this area that have received a grade score (at least one) more than 50.

Hint: Use the \$geoWithin and \$box.

4. Submission instructions

Create a report with screenshots of each of the queries with their corresponding results, as we want to see the system running on your machine of choice. Additionally, turn in a text file of the exact same queries (just in case we want to easily copy and paste them!).

Make sure the report is in a PDF format.

Submit on gradescope!