CS217 HW 8 – NoSQL using MongoDB

Overview

In this assignment, we will exercise some basic skills on MongoDB. MongoDB is complicated, and we only covered a very small fraction of it. Therefore, the main purpose of this assignment is to let you have a taste of it.

Before you start on this assignment:

1. **Make sure that you watched the lecture video on NoSQL. It's important that you understand every details of the setup and the demo.**
2. Learning to read the document is a major purpose of this assignment. Here are some helpful documentations:
   - https://pymongo.readthedocs.io/en/stable/tutorial.html
   - https://www.w3schools.com/python/python_mongodb_find.asp
   - MongoDB sample datasets: https://docs.atlas.mongodb.com/sample-data/available-sample-datasets/
3. DO NOT change any contents in the database. Do not terminate the cluster since we will need to run tests on your cluster.

First we look at the SampleAirBnB Listings Dataset, the collection is `sample_airbnb.listingsAndReviews`.

Q1 (1pt): Write a function `test_monogoDB` that returns the name of any entry in the collection `sample_airbnb.listingsAndReviews`. (this is mostly the same as in the demo)

```
>>> test_mongoDB()
'Ribeira Charming Duplex'
>>> type(test_mongoDB())
<class 'str'>
```

Q2 (2pt): Write a function `get_by_bd` that takes two integers (`a,b`) as input and returns an integer that represent the number of listings such that the number of bedrooms is between `a` and `b`, inclusive. (do not count those that do not specify the number of bedrooms)

Some code for verification:

```
>>> get_by_bd (7,10)
14
>>> type(get_by_bd (7,10))
<class 'int'>
```

The next three questions are about the Sample Supply Store Dataset, the collection is `sample_supplies.sales`. Please inspect the data format from MongoDB Atlas online portal first.

Q3 (1pt): Write a function `gender_dist` that takes no input and returns a tuple of integers indicating the number of transactions that are made by female and male respectively.

Q4 (1pt): We want to study the correlation between gender and their satisfaction of orders. Write a function `gender_stf` that takes no input and returns a tuple of floats, indicating the average `satisfaction` values made by female and male customers respectively.

Q5 (2pt): Write a function `num_prod` to query the total number of products sold on a specified day. The function takes three integers (`y, m, d`) as inputs indicating the year, month and day. It returns an integer of the total count of products.
Hint: check the data type for the field `saleDate`.
Documentation of `datetime` for reference: https://docs.python.org/3/library/datetime.html.

The next two questions are about Sample Restaurants Dataset. The collection is `sample_reataurants.reataurants`. Please inspect the data format from MongoDB Atlas online portal first.

Q6 (1pt): Write a function `num_cuisine` that takes no input and returns an integer indicating the total number of distinct cuisines in this database.

Q7 (2pt): Write a function `grades` that takes two strings (restaurant name and zipcode) as inputs and returns a list of all the letter grades this restaurant have ever received.
Some code for verification:

```
>>> grades("Lexler Deli","10174")
['A', 'A', 'A', 'A', 'A']
>>> grades ("Tov Kosher Kitchen","11374")
['Z', 'A', 'A', 'B']
```

Submission:

Submit the completed py file ONLY. DO NOT rename the file.
DO NOT call any function that you wrote in the py file. That is, please only keep the definition of the functions.
DO NOT print any information in your functions.
DO NOT change any function names!