Majadbi_Yousef_CSE511_NoSQL Project Report

Reflection

For this project, I implemented two functions: *FindBusinessBasedOnCity* and *FindBusinessBasedOnLocation* to query a NoSQL database using UnQLite and distance algorithm function *DistanceFunction*. The goal was to filter businesses based on City and Location then output the results to text files.

Implemented functions:

1. Distance algorithm function: The '*DistanceFunction' calculates the distance between two geographic coordinates using the Haversine formula.*

```
import math

def DistanceFunction(lat2, lon2, lat1, lon1):
        R = 3959
        φ1 = math.radians(lat1)
        φ2 = math.radians(lat2)
        Δφ = math.radians(lat2 - lat1)
        Δλ = math.radians(lon2 - lon1)

        a = math.sin(Δφ / 2) * math.sin(Δφ / 2) + math.cos(φ1) * math.cos(φ2) *
        math.sin(Δλ / 2) * math.sin(Δλ / 2)
        c = 2 * math.atan2(math.sqrt(a), math.sqrt(1 - a))

        d = R * c
        return d
```

2. FindBusinessBasedOnCity: This function filters businesses based on a given city name and writes the results to a text file.

```
def FindBusinessBasedOnCity(cityToSearch,saveLocation1,collection):
        results = collection.filter(lambda doc: doc['city'] == cityToSearch)
        with open(saveLocation1, 'w') as file:
                for business in results:
                        line =
        f"{business['name']}${business['full_address']}${business['city']}${business['st
        ate']}\n"
                        file.write(line)
```

3. FindBusinessBasedOnLocation: This function finds all businesses present in the
   'maxDistnace' from the given 'myLocation', using the distance algorithm function and
   save the results to text file.

```
def FindBusinessBasedOnLocation(categoriesToSearch, myLocation, maxDistance,
saveLocation2, collection):
        lat1, lon1 = myLocation

        results = collection.filter(lambda doc: any(cat in doc['categories'] for cat in
        categoriesToSearch))

        with open(saveLocation2, 'w') as f:
                for business in results:
                        lat2, lon2 = business['latitude'], business['longitude']
                        distance = DistanceFunction(lat2, lon2, lat1, lon1)
                        if distance <= maxDistance:
                                f.write(f"{business['name']}\n")
```

Lessons Learned

Through this project I learned the following:

1. Working with **NoSQL** databases, especially the **Document databases.**
2. The difference between NoSQL and SQL in the matter of storing and querying the data.
3. How to use **UnQLite** in **Python** to conduct several fundamental operations on NoSQL
   databases, including connecting to databases, accessing collections, and filtering
   documents.

Output

Below are screenshots from parts of my TXT output files:

Output_city.txt

```
1  VinciTorio's Restaurant$1835 E Elliot Rd, Ste C109, Tempe, AZ 85284$Tempe$AZ
2  Salt Creek Home$1725 W Ruby Dr, Tempe, AZ 85284$Tempe$AZ
3  P.croissants$7520 S Rural Rd, Tempe, AZ 85283$Tempe$AZ
4
```

Output_loc.txt

```
1  VinciTorio's Restaurant
2
```

Results

When running the test case to search for businesses in 'Tempe', I got the below result:

> Correct! You FindBusinessByCity function passes these test cases. This does not cover all possible test edge cases, however, so make sure that your function covers them before submitting!

And when running the test case to search for businesses in 'Buffets' category, in my location [33.3482589, -111.9088346] and maxdistance is 10, I got the below result:

> Correct! Your FindBusinessBasedOnLocation function passes these test cases. This does not cover all possible edge cases, so make sure your function does before submitting.