10/7/2020

*Submitted by:*

Sampreeth Amit Kumar

Vishal Gnanasekar 30382300

*Tutorial:* 02-P1

*Tutor Name:* Agnes and Ningran

GROUP ASSIGNMENT

FIT5137 Advanced Database Technology

Semester 2, 2020

**GROUP ASSIGNMENT COVER SHEET**

|  |  |  |
| --- | --- | --- |
| **Student ID Number** | **Surname** | **Given Names** |
|  |  |  |
|  |  |  |
| **\*** Please include the names of all other group members. | | |
| **Unit name and code** |  | |
| **Title of assignment** |  | |
| **Lecturer/tutor** |  | |
| **Tutorial day and time** |  | **Campus** |
| **Is this an authorised group assignment?  Yes  No** | | |
| **Has any part of this assignment been previously submitted as part of another unit/course?**  **Yes  No** | | |
| **Due Date** | | **Date submitted** |

All work must be submitted by the due date. If an extension of work is granted this must be specified with the signature of the lecturer/tutor.

**Extension granted until (date) ................................ Signature of lecturer/tutor .............................................**

Please note that it is your responsibility to retain copies of your assessments.

|  |
| --- |
| ***Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations***  **Plagiarism**: Plagiarism means taking and using another person’s ideas or manner of expressing them and passing them off as one’s own. For example, by failing to give appropriate acknowledgement. The material used can be from any source (staff, students or the internet, published and unpublished works).  **Collusion**: Collusion means unauthorised collaboration with another person on assessable written, oral or practical work and includes paying another person to complete all or part of the work.  Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or delegate, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing. |
| **Student Statement:**   * I have read the university’s Student Academic Integrity [Policy](http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-policy.html) and [Procedures](http://www.policy.monash.edu/policy-bank/academic/education/conduct/student-academic-integrity-managing-plagiarism-collusion-procedures.html). * I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) Regulations http://adm.monash.edu/legal/legislation/statutes * I have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied. * No part of this assignment has been previously submitted as part of another unit/course. * I acknowledge and agree that the assessor of this assignment may for the purposes of assessment, reproduce the assignment and:  1. provide to another member of faculty and any external marker; and/or 2. submit it to a text matching software; and/or 3. submit it to a text matching software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.  * I certify that I have not plagiarised the work of others or participated in unauthorised collaboration when preparing this assignment.   *Signature .............................................................................. Date……………………………………*  \* delete (iii) if not applicable |

Signature : Date : 7 October 2020

Signature : Vishal Gnanasekar Date : 7 October 2020

**Contribution Declaration Form**

**(to be completed by all team members)**

**Please fill in the form with the contribution from each student towards the assignment.**

**1 NAME AND CONTRIBUTION DETAILS**

|  |  |  |
| --- | --- | --- |
| Student ID | Student Name | Contribution Percentage |
|  |  | 50% |
| 30382300 | Vishal Gnanasekar | 50% |

**2 DECLARATION**

**We declare that:**

* The information we have supplied in or with this form is complete and correct.
* We understand that the information we have provided in this form will be used for individual assessment of the assignment.

**3 SIGNATURE**

**Signatures**

DayMonth Year

07 / 10 / 2020

**Date**

# C. Tasks

## C.1. Creating the Databases and Importing Data

### MongoDB (MongoDB Compass and Mongo Shell):

1. Using the Mongo Shell and appropriate Mongo Shell command(s), create a database called: FIT5137A1MRDB
2. Using the Mongo Shell and appropriate Mongo Shell command(s), create the following 2 collections: userProfiles placeProfiles
3. Using MongoDB Compass and appropriate MongoDB Data Types, import the data in

a. userProfile.json into the userProfiles collection.

b. placeProfiles.json into the placeProfiles collection.

1. In MongoDB, we understand that there are two data modelling methods, which are embedding and referencing.
   1. In your own words, explain what you understand by embedding and referencing data models.
   2. Import the data in openingHours.csv into your FIT5137A1MRDB database using appropriate data models.
   3. Include in your report, an explanation behind the selection of the data model.

### Cassandra (Cassandra Shell):

1. Create a keyspace called FIT5137A1\_MRDB for the Cassandra database, with SimpleStrategy and replication factor of 1.
2. Create the following column families using appropriate data types:

a. user\_ratings

b. place\_ratings

1. Using the cassandra COPY command import the following data into the tables:

a. user\_ratings.csv into the user\_ratings table.

b. place\_ratings.csv into the place\_ratings table.

## C.2. Modifying the Databases

1. MonR has gained some new information about a trendy new place. Therefore, without creating any new fields, insert all of the information provided in Table 1.
2. They have also realised that the user with user\_id 1108, no longer prefers Fast\_Food and also prefers to pay using debit\_cards instead of cash. Therefore, without looking up the existing values or adding any new fields, update user 1108’s favorite cuisines and favorite payment methods.
3. The management has realised that the user with user\_id 1063 was an error. Therefore delete the user 1063 from the database
4. To be consistent with the changes made in Task C.2 (1), (2), and (3), the management has also requested to update the reviews information to reflect the changes made to the users 1108 and remove the user 1063’s reviews. They have looked up the data in the reviews table and have provided the information in table 2.
5. It was also seen that user 1060 has reviewed Taco Jacks (ie. the new place with place id 70000), therefore using the information from table 3, insert the following data: (for this insert only you may look up the details of user 1060).

## C.3. Querying the Management.

1. How many users are there in the database?

2. How many places are there in the database?

3. How many reviews were made in the database?

4. How many reviews were created by places having public parking arrangements?

5. Display the user id, rating of places and the user’s personality for all users who are interested in technology, thrifty protector types of workers, with blue as their favorite colour and are casual drinkers. (Note that the number of rows returned should be greater than zero)

6. What are the place ids and ratings for food for places serving only Pizzeria cuisine? (Note the result should include cuisine).

7. Display all users who are students and prefer a medium budget restaurant.

8. Display all users who like Bakery cuisines and combine your output with all places having Bakery cuisines.

9. Display International restaurants that are open on Sunday.

10. Display the average place rating, average food rating, and average service rating for puesto de tacos. Show the average values in decimal points.

11. Display the average age according to each drinker level.

12. For each user whose favourite cuisine is Family, display the place ID, the place rating, the food rating and the user’s budget.

13. What are the top 3 most popular ambiences (friends/ family/ solitary) for a single when going to a Japanese restaurant?

14. List the names of unique cuisines in the database.

15. Display all of the restaurants and indicate using a separate field/column whether the restaurant includes mexican cuisines. For instance, you can display if the restaurant serves mexican cuisine then the result should show the restaurant name followed by “serves mexican food” in the next field/column, or if the restaurant does not serves mexican cuisine then the result should show the restaurant name followed by “doesn’t serves mexican food” in the next field/column.

Five additional queries

## C.4. Summary Reports

As a summary you are required to provide a comparison between the three databases i.e. Relational Database, Document-Oriented Database and Column-Oriented Database

C.5. Connecting to Drivers (Optional Bonus Section - up to 10 marks)