CORK INSTITUTE OF TECHNOLOGY
INSTITIÚID TEICNEOLAÍOCHTA CHORCAÍ

Computing@CIT

Assignment 2 Comp7037

Final Submission Due : Week 12 - 1st May

Submission location : Canvas

Submission Details : Single Document – PDF – which includes a screen shot of each query and its output, with appropriate titles

Worth : 35% of final Course Mark

This assignment work as it progresses will be reviewed in the lab with the student, please ensure that as you complete each of part of the assignment 1,2 and 3 you demonstrate your work to your lecturer in the lab.

The student's work will be reviewed in the lab if possible and the student must be able to explain logic and implementation of all work submitted. No mark will be given to student for work which has not been reviewed with the lecturer in the lab if possible. If it is not possible to review work in the lab an a review may be organised with the student using Skype.

Default Loan Data Set

The json objects contained in the default loan data set describe banking history for individuals. Most of the attributes are self-explanatory. The balance attribute indicates balance on current account.The outstandingloans attribute contains a list of the balances on outstanding loans.

Sample Json object


```
{
  "_id": "5e7b0e1c2ff059dbea2bf2d3",
  "isActive": true,
  "balance": 4471.96,
  "age": 59,
  "eyeColor": "brown",
  "name": "Phyllis Larson",
  "gender": "female",
  "company": "ZOUNDS",
  "email": "phyllislarson@zounds.com",
  "phone": "+1 (956) 532-3374",
  "address": "545 Bergen Street, Kirk, Virginia, 8745",
  "registered": "2017-10-23T04:27:38 -01:00",
  "outstandingloans": [
    {
      "id": 0,
      "balance": -1438
    },
    {
      "id": 1,
      "balance": -1989.31},
    {
      "id": 2,
      "balance": -482.08
    },
    {
      "id": 3,
      "balance": -1142.74
    }
  ]
}
```

Part 1 : Set up Mongdb Cluster on mongo db – use the restaurants data set provided in zip file - 40% - demo working version to lecturer in the lab

Or

Update the scripts which have been supplied for set up of the Mongo Cluster with compreshensive comments clearly explaining what the commands in the scripts are doing.

Submit a zip file of the updated scripts to Canvas as part of your submission.

If you do not have access to the lab to allow you set up the cluster you should comment the scripts.

Part 2 : Use Apoc to load the data from Mongo server into NEO 4j - 30%. Please consider the queries which need to be executed in Neo4j when creating the database and ensure that you are creating a graph that represents the data in an effective way.

Submit commands that were used to create your database in Neo4j and sample output from commands on a PDF.

I will be supplying video tutorials on Canvas on how to setup and use APOC

I will supply a csv file so that an alternative for those who have difficulties using APOC can use load csv – however the maximum marks given will be 20% of the 30% available marks

Part 3 : Do the following queries on both databases. Submit the queries and the output for Mongo and for Neo4j – 30%

1: Are males or females on average bigger defaulters on their loans ?

- The size of a person's loan default is calculated by summing the balances in their outstanding loans

2: What number of male loan defaulters who have 3 loans or more have defaulted on loans in total valued over 15,000 ?

Note : Standard CIT penalties apply for plagiarism and late submissions.