

## CET341 – ADVANCED DATA TECHNOLOGIES (2020/21)

### Assignment Two

---

The following learning outcomes will be assessed:

#### **Knowledge**

1. critical appreciation of current and new data models and database systems for traditional and Big Data systems;
2. appraisal of current and emerging trends in database systems and their application in the real world;

#### **Skills**

3. the ability to design and develop database systems using a range of different database development tools;
4. evidence of critical evaluation of the major developments and issues of databases within the database arena and their support in various application areas.

#### **Important Information**

You are required to submit your work within the bounds of the University Infringement of Assessment Regulations (see your Programme Guide). Plagiarism, paraphrasing and downloading large amounts of information from external sources, will not be tolerated and will be dealt with severely. You should make full use of any source material, which would normally be an occasional sentence and/or paragraph (referenced) followed by your own critical analysis/evaluation. You will receive no marks for work that is not your own. Your work may be subject to checks for originality which can include use of an electronic plagiarism detection service.

Where you are asked to submit an individual piece of work, the work must be entirely your own. The safety of your assessments is your responsibility. You must not permit another student access to your work.

Where referencing is required, unless otherwise stated, the Harvard referencing system must be used (see your Programme Guide).

Please ensure that you retain a duplicate of your assignment. We are required to send samples of student work to the external examiners for moderation purposes. It will also safeguard in the unlikely event of your work going astray.

<b>Submission Date and Time</b>	As specified on Canvas
<b>Submission Location</b>	Electronic submission via Canvas

This is the second assignment for this module, and is worth **65%** of the overall assignment mark.

All work must be done individually.

Your assignment is to **design, develop and query a database** using multiple DBMSs, i.e. Oracle and MongoDB, to provide a brief comparison of relational and NoSQL document store database technologies.

For a **case study of your choice**, perform the following tasks:

1. Produce one paragraph (300 words maximum) briefly describing your case study, focusing on describing the key functionality of the system. Subsequently provide a **UML class diagram** showing the main entities and attributes in your design and the relationships between them. The UML diagram **MUST** be data model independent, i.e. you are showing entities in the system, **NOT** tables. Ensure that you include complex relationships (e.g. inheritance hierarchies, aggregation, etc.) in your design.
2. Provide a complete SQL script file that will run in Oracle to implement the design from task 1 as a **relational** database. For each table, ensure that is fully normalised to BCNF, by specifying all functional dependencies that exist in each table, with a statement as to whether and why each table meets BCNF, and correctly resolving any normalisation issues. Ensure your SQL script file runs without errors, i.e. it drops then creates the tables and inserts sufficient sample data for your queries below.
3. Develop the following queries in **Oracle SQL** on your relational database from task 2 above. Ensure that you develop queries which include:
  - a. A join of three or more tables – you should consider various types of join in this query (e.g. inner join, outer joins, etc.);
  - b. A query which requires use of a sub-query;
  - c. A query using temporal features (e.g. timestamps, intervals, etc.) of Oracle SQL;
  - d. A query using OLAP (e.g. ROLLUP, CUBE, PARTITION) features of Oracle SQL.

Ensure that you give a natural language description of each query.

4. Develop a set of documents using MongoDB to develop the equivalent **document store** database for your design in task 1. Ensure you make use of appropriate document store features including, for example, nested documents, arrays, etc. Do not simply replicate the Oracle implementation but make full use of MongoDB document store features. Make sure you provide the complete script file for creating your MongoDB database, do not just provide a list of documents outputted from the interface.
5. Using MongoDB, implement queries which are equivalent (or as similar as possible) to those in task 3a-d above. Ensure you use the native MongoDB query language and **NOT** SQL.
6. Write a brief report (of no more than 500 words) which gives a **comparison between Oracle and MongoDB** for your database queries, using examples from your two database implementations to demonstrate your key points. You should ensure that you include discussion of problems or challenges that you encountered in how you implemented queries such as your temporal and OLAP SQL queries in MongoDB. Ensure that you cite appropriate academic references (i.e. research papers) in your discussion.

You need to ensure therefore that your case study is of sufficient detail to be able to complete tasks 1-6 above. As a suggestion, you should aim to have between four and six entities within your design which you would then represent in your two implementations using tables and document collections as appropriate. You may want to consider complex relationships between the entities such as inheritance and aggregation in your design. Marks given will take into account the challenge of your case study and the queries that you develop.

### **e-Portfolio**

Throughout the module you have also been required to create and maintain an e-Portfolio. You need to submit the private link to your e-Portfolio as part of the submission for this assignment. The e-portfolio is worth 10% of the module mark. Without submitting this link it will not be possible to mark your e-Portfolio and you will receive 0 marks for that element.

### **Submission Requirements**

Your submission for this assignment is a technical report which includes your solutions to tasks 1-6 above. All SQL, relational algebra, and MongoDB code must be included as well as screenshots to demonstrate that all of your database creation, inserts and queries work. Ensure each task is submitted in order within the one file, and the file must be converted to PDF for submission. The file must also, on the first page, contain the link to your e-Portfolio.

Staying within the bounds of University of Sunderland regulations, you should make full use of any source material available to you (particularly journals, conference papers or technical reports). Plagiarism, collusion and paraphrasing will not be tolerated, and will be dealt with under University Infringement regulations. If you are unclear about how to reference correctly please ask.

## Marking Scheme

The marks breakdown is as follows:

Task		Weight	Not attempted	Fails to meet expected standard	Just below expected standard	Meets expected standard	Exceeds expected standard
1	Case study and UML class diagram	5%	Not attempted or very poor attempt.	Description of case study is not at an appropriate level and lacks challenge. UML class diagram provided but does not suitably describe the case study, and incorrect notation is used. Attributes and entities are not clearly shown and no complex relationships. Unsuitable number of entities is described.	Description of case study is almost at an appropriate level and challenge. UML class diagram correctly describes the case study, notation is mostly correct, but attributes and entities are not clearly shown with limited attempt to embed complex relationships. Suitable number of entities is described.	Description of case study is of an appropriate level and challenge. UML class diagram correctly describes the case study, notation is used correctly, clearly showing attributes and entities with some attempt to embed complex relationships. Suitable number of entities is described.	Very good description of case study at a very good level and challenge. UML class diagram correctly describes the case study, notation is used correctly, clearly showing attributes and entities with a very good attempt to embed complex relationships. A good number of entities is described to demonstrate the complexity of the scenario.
2	Oracle SQL database creation, normalisation and data insertion	10%	Not attempted or very poor attempt.	SQL script file developed but does not match scenario and contains errors. No, or limited, integrity rules specified. No, or very poor, attempt at normalisation to BCNF. No, or very limited, screenshots provided. No, or limited, sample data is provided.	SQL script file developed and mostly error free but does not fully match scenario. Not all database integrity rules correctly specified. Some attempt to normalise to either 3NF or BCNF, with some definition of functional dependencies. Some screenshots are provided but more could have been provided to clearly demonstrate that that system works. Some sample data is provided.	Correctly working SQL script file which can be run without error and matches the scenario described. Database integrity rules are correctly implemented. Functional dependencies are clearly stated and all tables are clearly normalised to BCNF. Any tables not in BCNF are attempted to be dealt with appropriately. Screenshots are provided to fully demonstrate that the code works. A good set of sample data is provided.	Correctly working SQL script file which can be run without error and completely matches the scenario described, utilising some advanced features of Oracle. Database integrity rules are correctly implemented including all types of integrity constraint. Functional dependencies are clearly stated and all tables are correctly normalised to BCNF. Any tables not in BCNF are dealt with correctly. An excellent set of screenshots are provided to fully demonstrate that the code works. An excellent set of sample data is provided.

3	Oracle queries	10%	Not attempted or very poor attempt.	Not all Oracle queries attempted and contain errors. Not at an appropriate level of challenge. No, or poor, natural language description is given. No, or limited, screenshots provided.	All queries attempted but may contain errors and not at an appropriate level of challenge. Natural language description is given but does not clearly describe the query. Screenshots are given which show data output.	All Oracle queries developed, work, and are at an appropriate level of challenge. Natural language description of each query is given. Screenshots are given which show the correct output from each query.	All Oracle queries developed, work and are at an excellent level of challenge. Natural language descriptions are given which clearly describe each query. Screenshots are given which show the correct output from each query.
4	MongoDB database creation and data insertion	10%	Not attempted or very poor attempt.	Poor attempt at the MongoDB database which shows lack of understanding of how to create a document store. Code contains many errors. Limited or no sample data provided. No, or limited, screenshots provided.	Good attempt at the MongoDB database but shows some issues in understanding of how to implement a document store. Code is mostly error free and a good matching data set has been provided. Some screenshots provided.	MongoDB database code provided which correctly implements the case study using appropriate document store facilities. MongoDB code will work without error. Sample data matches sample data provided in Oracle but correctly uses MongoDB concepts. Screenshots have been provided to demonstrate that code works.	MongoDB database code provided which correctly implements the case study using appropriate and some advanced document store facilities throughout. MongoDB code will work without error. Sample data matches sample data provided in Oracle but correctly uses some advanced MongoDB concepts. Screenshots have been provided to fully demonstrate that code works.
5	MongoDB queries	10%	Not attempted or very poor attempt.	Poor attempt at one of more of the MongoDB queries. Many errors in code. No, or few, screenshots have been provided.	Good attempt at 3 or 4 of the MongoDB queries, using MongoDB constructs where relevant. Some errors in code. Screenshots have been provided which attempt to show that the queries work with appropriate results.	MongoDB queries provided are a good attempt to match the Oracle queries provided, using appropriate MongoDB constructs. Screenshots are provided to fully demonstrate that the queries work, showing appropriate results.	Excellent set of MongoDB queries provided fully using appropriate and some advanced MongoDB constructs. Full set of screenshots provided which clearly demonstrate that the queries work with complete results shown.
6	Comparative report	10%	Not attempted or very poor attempt.	Limited or no critical comparison of MongoDB and Oracle, with no, or limited, use of examples. No findings are presented. No, or limited, use of references	Some good comparison of MongoDB and Oracle, some examples used but not necessarily from the scenario. Report is generic and does not focus entirely on database queries. Some use of academic referencing with a	Good critical comparison of MongoDB and Oracle with clear use and discussion of examples from the developed case study. Report focuses on database queries. Findings from discussion are	Excellent critical comparison of MongoDB and Oracle with excellent use throughout of examples from the case study. Report is focused and excellent findings are presented. Very good use of academic

					reference list provided but not necessarily in Harvard format.	presented and are relevant. Good use of academic referencing used and reference list correctly supplied in Harvard format.	referencing and a reference list of high quality sources is provided in Harvard format.
7	e-Portfolio	10%	Not attempted or very poor attempt.	Not all sections of e-Portfolio were submitted.	Most sections of e-Portfolio were submitted with some reflective discussion in each section.	All sections of e-Portfolio submitted with a good level of critical reflection in each section.	All sections of e-Portfolio submitted with an excellent level of critical reflection in each section.