Faculty Of Computing and Engineering Sciences

Assessment Cover Sheet and Feedback Form 2020-21

Module Code:	Module Title:	Module Team:
MS4S21	Big Data Engineering and Applications	Moizzah Asif
Assessment Title and Tasks: Set Tasks - not-time constrained 1		Assessment No. 2
Date Set: 28-Apr-2021 17:00	Submission Date: 14-May-2021 21:00	Return Date: 09-Jun-2021 21:00

IT IS YOUR RESPONSIBILITY TO KEEP RECORDS OF ALL WORK SUBMITTED

Marking and Assessment
This assignment will be marked out of 100% This assignment contributes to 40% of the total module marks.
Learning Outcomes to be assessed (as specified in the validated module descriptor https://icis.southwales.ac.uk/): 1) To appraise and contrast strategies for dealing with Big Data 2) To demonstrate an ability to apply Big Data concepts in non-trivial contexts
Provisional mark only: subject to change and / or confirmation by the Assessment Board

MS4S21 Coursework - I & II: 2020/21

Submission deadline: 14th May 2020 - 21:00

Your assessment is divided into two main parts: **Coursework I** and **Coursework II**

Coursework I is based on conducting experiments and providing the scripts along with brief overview and commentary in the sequence of execution in a single pdf file. It contributes 60% to this module's assessment.

Coursework II contributes 40% to the this module's assessment and is based on writing reports on specific big data technology research topics and tools.

You are required to:

1. submit a pdf file for both courseworks respectively,
2. submit any programming scripts and code used in Coursework I, and
3. attempt all the questions in both the courseworks.

You may attach appendices to each coursework's pdf file.

You should be able to complete these courseworks while staying well under your AWS classroom budgets. It will be your responsibility to not to exceed the budget and have contingency plans to evidence your work in the unlikely event that you loose your work due to reaching the budget limit. Furthermore, you are advised to plan the execution of your experiments for Coursework I in such a way that you are able to terminate any unused AWS services without affecting your coursework completion and progress.

The pdf and any accompanying files for both the courseworks should be submitted as separate submissions, under their own BB assignments provided in Blackboard on MS4S21 home page by the deadline.

Coursework II

Submission deadline: 14th May 2021 - 21:00

Contribution to module: 40%

This coursework is divided into two main sections: **Section A & Section B**.

Section A requires you to write brief reports on specific big data tool and technologies and their usability.

Section B requires you to write a short research report on a specific topic of research and its application in the context of big data.

You are required to attempt both sections.

Section A

20 marks

1. You are required to write a brief report which should:

   (a) provide an introduction to the database called MongoDB in the context of Big Data and technologies, and

   (b) draw comparisons between MongoDB and DynamoDB and their usability in the context of Big Data and technologies.

   The report should adhere to the word count of **600** (± *10%).*

2. What are data pipelines and how can they be utilised in cloud services such as AWS. You are required to pick up tasks or experiments from Coursework I to theoretically demonstrate the usability AWS data pipeline.
   The report should adhere to the word count of **600** (± *10%).*

Section B

20 marks

Write a formal academic style report of approximately **1200** words *(± 10%)* on Knolwedge graphs.

Your report should make references to the following:

1. introduction to knowledge graphs and their usability in general;
2. overview of RDF and ontologies;
3. the usability of knowledge graphs and a use-case or example of their application in either academia or industry;

4. and, an overview of a knowledge graph platform used either commercially or in academia with any type of license (open-source or paid).