



MS5106 End-of-Module Exam Assignment 2022/23 – Part A

Module Code: MS5106

Module Name: Data Science and Big Data Analytics

Objective	<p>The objective of this assignment is to assess your understanding of the following learning outcome(s):</p> <ul style="list-style-type: none">• Demonstrate understanding of technologies for big data analysis, such as Hadoop, HDFS, MapReduce, Yarn, and Spark.• Demonstrate understanding of big data ecosystem technologies, such as HBase, Hive, Pig, and their role in big data analytics.
Lecturer	Anatoli Nachev
Marks Awarded	<p>This assignment carries 30% of the overall marks for the module.</p> <p>Guidelines for grading this assignment are included at the end of this document.</p>
Submission	<p>You should submit a MS Word document (.docx or .doc) containing cover page, provided. The text should not exceed 1,500 words, excluding the cover page, table of contents, bibliography, graphs, tables, appendices, if any.</p> <p>File name should contain your ID, module code, ExamAssignment, and PartA (e.g., 12345678_MS5106_ExamAssignment_PartA.docx). Please, follow strictly the file naming required.</p> <p>You should submit your completed assignment through the Blackboard electronic drop box provided by 4th May 2023 @23:59. The drop box is set to accept multiple submission attempts, the last one will be taken as valid.</p> <p>If you are unable to submit via the Blackboard, please email the file to your lecturer anatoli.nachev@universityofireland.ie</p> <p>Non-submission, or submission after the deadline will carry a mark of zero in determination of overall marks for this sitting.</p>

Referencing	<p>If you are using any references, please use the Harvard referencing style, instructions for which can be found on the NUI Galway Library Website, section Styles & Styles Guide: Citing & Referencing:</p> <p>http://libguides.library.nuigalway.ie/c.php?g=672922&p=4791378</p>
Deferral	<p>You may request a deferral of this assignment until Autumn, if you are unable to perform the work or submit it before the deadline. For details, see https://www.universityofgalway.ie/exams/timetable-advice/deferrals/ email business@nuigalway.ie,</p>
Assignment Details	<p>EU-Trade&Co is an e-commerce company with very successful online store. The company processes lots of data about their sales, compares prices of products across multiple shopping sites so that they come up with innovative promotions and good pricing strategy in order to beat the competitors. With ever increasing amount of data and demand for efficient data processing, EU-Trade&Co's management are seriously considering moving to big data technologies. They also believe that mining big data is the way forward.</p> <p>Based on your knowledge in big data technologies, you have been asked to consult the EU-Trade&Co's management by writing a report on the following issues:</p> <ul style="list-style-type: none"> • EU-Trade&Co don't understand how to build a platform for big data processing that can help to meet their goals. They have heard words, such as clusters, clouds, Hadoop, Spark, but do not really understand what to do and how to do it. They are also concerned about the cost and complexity of the big data processing. You are required to advise by outlining possible solutions and comparing alternatives. • Traditionally, the data they process are mostly structured and the technologies they used so far were mostly traditional databases and SQL. EU-Trade&Co don't have expertise in programming and they ask you to advise what would be the easiest and straightforward big data technologies they can adopt in order to rely on their current expertise, yet adding value to the company. Outline possible solutions and compare their features, strengths, and limitations. <p>Clearly state any assumptions that you have made in your report.</p> <p style="text-align: right;">(30 Marks)</p>

To help with expectations of standard of submission, following is a general set of Grade Descriptors as proposed by the SBE:

GRADE DESCRIPTORS

Percentage Translation

80%+	Exceptional
70 – 79%	Excellent
60 – 69%	Very good
50 – 59%	Good
40 – 49%	Satisfactory
35 – 39%	Fail* - Compensatable
<35%	Fail

First Class Honours 80%+

Exceptional performance, engaging deeply and systematically with the question set, with consistently impressive demonstration of:

- A comprehensive mastery of the subject matter; amply supported by evidence and citation;
- Reflecting deep and broad knowledge and critical insight as well as extensive reading;
- An exceptional ability to organise, analyse and present arguments fluently and lucidly with a high level of critical analysis;
- A highly developed capacity for original, creative and logical thinking.

First Class Honours 70-79%

Excellent performance, engaging closely and systematically with the question set, with consistently strong evidence of:

- A comprehensive mastery of the subject matter, ably supported by evidence and relevant citation;
- Excellent ability to organise, analyse and express arguments fluently and lucidly with a high level of critical analysis;
- A highly developed capacity for original, creative and logical thinking. Extensive knowledge and understanding of theory and concepts. Ability to demonstrate understanding through use of examples;
- An ability to draw on varied aspects of the course and previous studies in answering questions.

Second Class Honours 60-69%

Very Good performance, engaging substantially with the question set, demonstrating strong grasp of the subject matter, well supported by evidence and relevant citation, demonstrating:

- Well-developed capacity to analyse issues, organise material, present arguments clearly and cogently;
- Some original insights and capacity for creative and logical thinking.

Second Class Honours 50-59%

Good performance - intellectually competent answer (i.e., factually sound) with evidence of a reasonable familiarity with the relevant literature and techniques, demonstrating:

- Acceptable grasp of the subject material;
- Ideas stated rather than developed and insufficiently supported by evidence and relevant citation;
- Writing of sufficient quality to convey meaning but some lack of fluency and command of suitable vocabulary;

- Omission of parts of the subject in question or the appearance of several minor errors;
- Average critical awareness and analytical qualities;
- Limited evidence of capacity for original and logical thinking.

Pass 40-49%

Satisfactory performance – intellectually adequate answer with evidence of some familiarity with the relevant literature and techniques, demonstrating:

- Basic grasp of subject matter, but somewhat lacking in focus and structure;
- Main points covered in answer, but lacking detail;
- Some effort to engage, but only a basic understanding of the topic portrayed;
- Some development of argument;
- Only some critical awareness displayed;
- No evidence or relevant citation included in answer;
- Appearance of several minor errors or one major error;
- Lacking evidence of capacity for original and logical thinking.

Fail* 35-39% - Compensatable

Unacceptable performance, with either

- insufficient understanding of the question displayed
- failure to address the question resulting in a largely irrelevant answer
- a display of some knowledge of material relative to the question posed, but with very serious omissions / errors and/or major inaccuracies included in answer
- or answer left somewhat incomplete for lack of time

Also:

- limited understanding of question displayed
- a random layout / underdeveloped structure - not planned sufficiently
- poor analytical skills, with an absence of argument
- random and undisciplined development - limited structure
- lack of clarity, poor spelling
- material of marginal relevance predominating

Fail <35%

Wholly unacceptable performance, with

- deficient understanding of the question displayed
- complete failure to address the question resulting in an irrelevant answer
- inadequate knowledge displayed relative to the question posed
- or answer left incomplete for lack of time

Also:

- very poor analytical skills, with an absence of argument
- random and undisciplined development –poorly structured answer
- confused expression, poor spelling
- irrelevant material predominating