



School of Computer Science

COMP47470

Reflection Project
Data Management and Big Data

Teaching Assistant:	Cormac English
Coordinator:	Dr Anthony Ventresque
Date:	Friday 29 th October, 2021
Total Number of Pages:	2

General Instructions

- The report should not be longer than 6 pages.
- The report must be submitted as a PDF, but stylistically you may use whichever software you like to prepare it (e.g. Word, Google Docs, \LaTeX).
- Formal citations are not necessary, but if you are referring to something not explained in your text it would be preferable to provide some indication as to where the concept arises from.
- The breakdown of marks for the project will be as follows:
 - Exercise 1 (SQ/ vs. NoSQL): 20%
 - Exercise 2 (Hadoop/Spark): 30%
 - Exercise 3 (Paper reviews): 50%
- **Due date: 26/11/2021**

1. Compare relational and NoSQL database management models and in particular give your impressions on why one is better than the other and in which context. You can use some external content (papers, blogs, etc.) to support some of your ideas, but what we expect is a general discussion on the topic, with possibly some personal experience. This section should **not be more than a page**.
2. Write a short report (**max. 2 pages**) exploring the following topic: "Spark & Hadoop: Advantages and drawbacks of each technology, and why you might choose one over the other". Your discussion should address the following topics:
 - A brief overview of the two technologies
 - What the major technical differences are between Spark and Hadoop
 - In what circumstances would you choose Spark over Hadoop
 - In what circumstances would you choose Hadoop over Spark
3. Write a review (**no more than 3 pages**) on one of the research papers that are available on Brightspace. They are all different but they all contain some important information about the NoSQL and the Hadoop/Spark worlds. The following list of sections is an indication of how to write your paper. Some of the items might not be relevant for all papers, and you might want to add some sections in your report (e.g., evaluate the posterity of a solution etc.). We will have an open mind when reading your report and we just want to see how you analyse a research paper and are able to discuss it - in short there is no one single perfect report, everything that shows you made an effort to understand and focus on the important parts (research methodology, hypotheses, etc.) will be welcome.
 - identify the question/challenge the paper addresses. Explain in your own words what the motivation for the research is.
 - describe briefly the related work, i.e., the other (related) solutions that the authors compare themselves to. Show the limitations of these related solutions
 - Give an outline of the solution proposed by the authors (no need to go into details) showing the main components
 - describe their scientific method: what are the research questions they evaluate, how do they evaluate
 - describe briefly their results
 - give your impression on the idea, what you liked about the paper and whether you see any limitations etc.