WQD7007 Big Data Management
Project (20%)

Submission date: 09/13 Jan 2024 (Week 13)

Identify your group members in group assignment. Then, find another ONE (1) group and join together. The total group members should be between 6 to 8. Please assign a group leader.

Select one of the title based on the domain submitted in group assignment. Define one or more suitable problem statement(s) based on the title selected. Locate data that is (are) suitable to be used from the internet.

If one problem statement is defined, each group member must choose one or more Hadoop technology (e.g. Hive, MongoDB, and Pig) or any other suitable big data tools to solve the defined problem statement. (Comparison of different Hadoop technology in specific task)
OR
If multiple problem statements are defined, the group member can choose to use the same or different combination of Hadoop Technology (To solve different problems using the same tool to test the effectiveness of the tool in particular problem, or build a complete big data solution using different tools with a number of problem statements).

All the group leaders will need to send an **EMAIL** to the lecturer**,** LATEST BY WEEK 9 (12/16 Dec 2023), stating:
- Their group members
- Title chosen
- Problem statement(s) identified
- Tool(s) to be used
- Planned additional tools (you are allowed to add in any useful tools aside from Hadoop ecosystem to further enhance your report outcome, e.g. by using Tableau for interactive visualization, or extra analytics using R or Python)
- Choose a time slot to have a 10 minutes presentation between <u>week 13 to week 14</u>.

Each group is required submit a written report latest by WEEK 13, explaining the problem statement(s) and how the tools are utilized to solve the problem(s).

Page limit: 10 pages (front page and references excluded).

Plagiarism:
Your attention is drawn to the no-cheating policy. This covers cheating, attempts to cheat, plagiarism, collision and any other attempts to gain unfair advantage in assessment. The work you submit must conform to this policy. If cheating is discovered, any parties involved will take equal blame (get zero).