WQD7005 (Data Mining)

2019/2010 Semester 1

This assignment should not only have learned in theory, but also have experience predictive analyzing and prescriptive analyzing data with the main objective of solving real-world problems. There are three key skills needed to succeed in data science, which we refer to as **c**reating, **c**onnecting, and **c**omputing.

It consists of **6 parts with 6 milestones:**

1. Acquisition of data in your familiar domain (Group : Max 2 students)

* Web crawling the real time data by using Python (WQD7004 Programming for Data Science)

Samples:

<https://drive.google.com/file/d/1CeguhXl6ZNIhC4YCQze90QwhAHEW4BGV/view>

https://github.com/GHEEJIAN/data-mining-project

Due Date : Week 3 (5 marks)

1. Management of data (Group : Max 2 students)

* Store data into hive data warehouse (WQD7007 Big Data Management)

Due Date : Week 5 (5 marks)

1. Processing of data (Group : Max 2 students)

Due Date : Week 8 (5 marks)

1. Interpretation of data (Individual)

Due Date : Week 9 (5 marks)

1. Communication of insights of data (Individual)

Due Date : Week 11 (5 marks)

1. Provide a user with recommendations around optimal actions to achieve investment objectives such as profits and return on investment (Individual)

Due Date: Week 14 (5 marks)