UNIVERSITY OF MALAYA

MIDDLE SEMESTER EXAMINATION FOR THE DEGREE OF MASTER OF DATA SCIENCE

ACADEMIC SESSION 2017/2018 : SEMESTER 2

WQD7001 : Principles of Data Science

March 2018 Time : 2 hours

INSTRUCTIONS TO CANDIDATES :

Answer **ALL** questions (50 marks).

**NAME**: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**MATRIC NO.**: **WQD1700\_\_\_\_\_\_**

(This question paper consists of 5 questions on 2 printed pages)

1. Based on your own understanding and using your own words, explain what data science is. You must use mind map or 5W 1H or drawing or any best and interesting way to assist your explanation, instead of just words.

[10 marks]

1. The FCSIT has loads of students data. The faculty needs to unlock some valuable insights from the data. Given a set of students’ data in the appendix, suggest what insights that could be derived from this dataset. Do state any assumptions, such as creating variable(s), that you make in this learning analytics task.

[10 marks]

1. (a) How datafication and analytics are related?

[4 marks]

(b) Missing data in general is one of the trickier issues that is dealt with when cleaning data. The simplest solution is to delete rows with missing values. What is your advice to a data scientist who chose this option?

[3 marks]

(c) You are finding some json files data for your project. What would you add to your Google query to look for files with a json extension?

[3 marks]

1. You are a CIO who is hiring a few data scientists to help shape the future of your company. You intend to look at their portfolios when making hiring decisions. Anticipate at least **FOUR** things that you expect to see in an impressive data science portfolio.

[10 marks]

1. Solve the cross-word puzzle.

[10 marks]

**END**