



QQI

MSC INFORMATION SYSTEMS WITH COMPUTING

SUMMER 2014 EXAMINATIONS

Module Code: **B9IS106**

Module Description: **Data and Data Analytics**

Examiner: **Dr Shazia A Afzal**

Internal Moderator: **Mr Michael Gleeson**

External Examiner: **Dr Stephan Weibelzahl**

Date: **Friday, 9 May 2014**

Time: **14.00-17.00**

INSTRUCTIONS TO CANDIDATES

Time Allowed: 3 Hours

Question 1 is Compulsory. Answer any two out of the remaining three questions. Total: (100 marks).

Question 1: (40 marks); All other questions: (30 marks each).

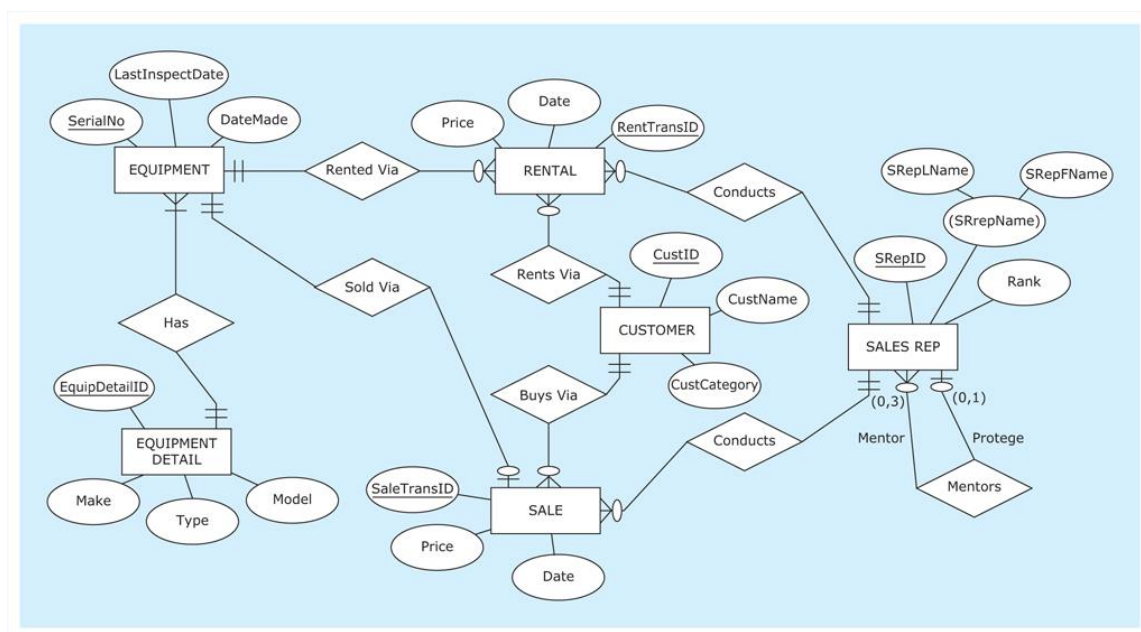
Question 1: (40 marks) – COMPULSORY

Answer ALL the following questions:

- i. Illustrate three characteristics of “Big Data”.
(6 marks)
 - ii. Explore two types of differences between analytical and operational data stores.
(6 marks)
 - iii. What is the difference between OLAP and OLTP?
(5 marks)
 - iv. Elucidate three basic analytical operations of OLAP tools.
(6 marks)
 - v. Discuss whether or not each of the following activities is a data mining task.
 - a. Predicting the outcomes of tossing a (fair) pair of dice.
 - b. Predicting the future stock price of a company using historical records.
 (5 marks)
 - vi. Distinguish between noise and outliers by giving three reasons.
(6 marks)
 - vii. Classify the following attributes as binary, discrete, or continuous. Also classify them as qualitative (nominal or ordinal) or quantitative (interval or ratio).
 - a. Angles as measured in degrees between 0° and 360°.
 - b. Bronze, Silver, and Gold medals as awarded at the Olympics.
 (6 marks)
- (Total: 40 marks)**

Question 2: (30 marks)

- i. Discuss the concept of star schema in dimensional modelling. In your discussion, include the analysis of several components of star schema.
(3+ 9 = 12 marks)
- ii. Jones Jozers wants to create an analytical data store to analyse its sales and rental revenue. The only available data source is the “Jones Dozers Sales and Rentals Database” depicted by the ER diagram given below:



Create a dimensional model for a data warehouse that enables analysis of sales and rentals revenue by:

- Date
- Type of revenue (sale or rental)
- Customer
- Equipment
- Sales rep

Each row in the fact table will represent the monetary amount of revenue taken in during one sale and rental transaction.

(18 marks)

Note: All dimensions and the fact table with proper attributes and relationships carry equal marks.

(Total: 30 marks)

Question 3: (30 marks)

- Examine in detail two data pre-processing approaches that can be used to make data suitable for data mining.

(7+7 = 14 marks)

- Discuss two data mining techniques: clustering and classification. Also illustrate the main differences between these two approaches?

(6+6+4 = 16 marks)

(Total: 30 marks)

Question 4: (30 marks)

Explain in detail your understanding of three out of the following by providing examples where appropriate:

- i. Extraction, Transform and Load (ETL)
- ii. Executive Dashboard
- iii. Record Data Sets
- iv. Map Reduce

(Total: 30 marks)