
# QQI

## MSc AI

## SUMMER 2022 EXAMINATIONS

*Module Code:* **B9AI108**

*Module Description:* **Programming for Data Analysis**

*Examiner:* **Paul Laird**

*Internal Moderator:* **Clive Gargan**

*External Examiner:* **Dr Svetlana Hensman**

*Date: 7th January 2022*
*Time: 9:30 - 12:30*

## INSTRUCTIONS TO CANDIDATES

**Time allowed is 3 hours**
**Answer All Questions**
**Run each question as python in Google Colab. You MUST submit the .ipynb source file. Failure to submit the source will result in a zero grade. Please also screen shot the program and its output and paste in the Answer Sheet.**

**Question 1**

**a)** Write a class in python, position, that represents food commodity orders or offers, with a type ('order'/'offer'), a product name - string, a quantity, price and traderID as private attributes, set in the constructor and accessed by getters. A method, trade, to reduce the outstanding quantity should be included, reducing the quantity remaining to trade by the parameter, amount, and a method to cancel, which sets the quantity to zero.

**(15 marks)**

**b)** Write a class market, which has an attribute product, and two orderbooks (order/offer) along with methods trade and checkDepth.

a. trade should accept a position, and if possible match it or partially match against opposite positions, otherwise storing it in the relevant orderbook

b. checkDepth should take a price and report the quantity on offer below that price, or (negative the quantity on order above that price.

**(25 Marks)**

**(Total: 40 Marks)**

**Question 2**

Data are provided on moodle in the file data.dat, which is in tab-separated format. Provide code snippets for the following:

a) Read the data into a dataframe and output a dictionary containing the mean of numerical columns, and the number of unique values for categorical columns, by column name.

b) Replace the values of feature B with a log-transformation of feature B.

c) Present a scatterplot with feature C on the y-axis against feature D on the x-axis for the records where the value of feature A is less than the cube of feature E

d) Create one hot encoded binary features for each unique value of feature F.

e) Create a test set containing a random 20% subset of the data, and create 5 folds of the remaining data for cross-validation.

**(5 * 5 = 25 Marks)**

**Question 3**

a) Write a program in python to implement a DBS asset registration system using OOP concepts where assets are assigned to departments. Create a separate department and asset class. Each asset should have an ID, value and department; and each department should have a named manager. Your program must be menu driven. The user should identify themselves at the main menu and return to the main menu when finished.

Please screenshot sample interaction with your application

**(35 Marks)**

**END OF EXAMINATION**