Database and Analytics Programming
CA1 (30%)

Masters of Science in Data Analytics

Instructions

- You have two hours, with five minutes grace. Submissions after this time will not be possible.
- You may consult any of the notes and code from the classes or the labs, as well as internet sources.
- You may not communicate with any other student by any means. Mobile phones must be left in your bag or on the floor. You may not have email or any social media page open on your computer at any time. Any violation of this rule will result in you being reported to the college's academic misconduct committee for disciplinary action.
- You must work on this CA in the classroom allocated. IP addresses can be checked. Anyone found to have submitted the CA from outside the college will receive 0 marks.
- All code must be **fully** commented.
- Create your code as a single Jupyter notebook and upload it to Moodle. Do **not** upload a Python (.py) file. Do **not** compress your submission using any means, including but not limited to Zip, RAR or 7z archives. Do **not** upload any datasets.

Question 1 (55 marks)

The Austin Animal Rescue dataset is a log of felines (cats) take into their care. The dataset provided on Moodle contains details for 1,000 such animals.

field	description
id	a number (1, . . . , 1000) identifying the animal taken into care
age_upon_outcome	age of animal at outcome date
breed	the breed of cat
color	the colour of the cat's coat
date_of_birth	the estimate date of birth of the cat
coat_pattern	the pattern of the cat's coat
name	the cat's name, if given
outcome_type	the type of outcome
sex_upon_outcome	whether the cat is intact or neutered at outcome date
sex	sex of cat

There are 1,000 records in total. You have been provided with an XML file containing this data (see the Moodle page).

a) Create a function to import this XML file. Your function should include appropriate exception handling clauses. **[15 marks]**

b) Use the print function to display the 'breed', 'color' and 'coat_pattern' of the first, third, fifth, seventh and ninth records in the XML dataset. (Hint: you may use the range() function). **[10 marks]**

c) Extract all the XML data and write it to a CSV file. Include appropriate exception handling. Your CSV file should also contain the column names. **[30 marks]**

Question 2 (25 marks)

a) Create a NumPy array filled with 2,000 numbers. Ensure that your array has 500 rows and 4 columns. **[5 marks]**

b) Using slicing, split this array into 5 separate arrays. The number of rows in each array should be equal, and there should still be 4 columns. **[10 marks]**

c) Reshape these 3 of these arrays into any dimensions of your choice. They should all have different dimensions. **[5 marks]**

d) Split 2 of these reshaped arrays horizontally **[5 marks]**

Question 3 (20 marks)

Given the following string:

```
"""
All I want is a proper cup of coffee
Made in a proper copper coffee pot
I may be off my dot
But I want a cup of coffee
From a proper coffee pot

Tin coffee pots and iron coffee pots
They're no use to me
If I can't have a proper cup of coffee
In a proper copper coffee pot
I'll have a cup of tea.
"""
```

a) Using **regular expressions**, write a single **function** to highlight the words "coffee","pot" or "pots" if they appear at the end of a line. **[15 marks]**

The text produced by your code look like this:

```
All I want is a proper cup of {coffee}
Made in a proper copper coffee {pot}
I may be off my dot
But I want a cup of {coffee}
From a proper coffee {pot}

Tin coffee pots and iron coffee {pots}
They're no use to me
If I can't have a proper cup of {coffee}
```

```
In a proper copper coffee {pot}
I'll have a cup of tea.
```

b) Using **regular expressions**, write a single **function** to highlight the words "proper" and "want" regardless of where they are found in a sentence. The words listed in part a) above should also be highlighted but only when they appear at the end of a sentence. **[5 marks]**

The text produced by your code should look like this:

```
All I {want} is a {proper} cup of {coffee}
Made in a {proper} copper coffee {pot}
I may be off my dot
But I {want} a cup of {coffee}
From a {proper} coffee pot.

Tin coffee pots and iron coffee {pots}
They're no use to me
If I can't have a {proper} cup of {coffee}
In a {proper} copper coffee {pot}
I'll have a cup of tea.
```