



University College Dublin  
An Coláiste Ollscoile, Baile Átha Cliath

---

**SEMESTER II EXAMINATIONS**  
**ACADEMIC YEAR 2017/2018**

---

**COMP 10020**

**Introduction to Programming 2**

Prof. J. Pitt

Prof. P. Cunningham

Dr. B. Mac Namee \*

**Time Allowed: 1 Hour**

**Instructions for Candidates**

Answer Question 1 and any two other questions.

Total marks available **100**. The value of each part of each question is shown in brackets next to it.

This is a Closed Book/Notes exam. Students are **not** permitted to bring materials to the Exam Hall.

**Instructions for invigilators**

This is a Closed Book/Notes exam.

Students are **not** permitted to bring materials to the Exam Hall.

Students are **not** permitted to use a computational device or the Internet in any fashion.

1. (a) What is the output of `print(str[:6])` if `str = 'University College Dublin'`?

(4)

- (b) Only one of the following Python variable names is valid:

- `my.function`
- `while`
- `__value`
- `2nd_place`

Identify the valid variable name, and explain what is wrong with each of the others.

(4)

- (c) Suppose the following code is used to create a dictionary:

```
d = {"mary":87.5, "geoffrey":93.2}
```

What happens when the following code is executed:

```
d["brian"] = "A+"
```

(4)

- (d) What is output by the following Python code snippet?

```
counter = 0
while counter < 10:
    print(counter)
    if counter == 8:
        break
    counter += 2
print(10)
```

(4)

- (e) What is the difference between a *tuple* and a *list* in Python?

(4)

- (f) What value is printed after the following code is executed?

```
x = y = 2
y += x + x * 4
print(y)
```

(4)

- (g) Write a python function, `centimetres_to_feet_and_inches`, to convert a person's height given in centimetres to feet and inches.

**Note:** there are 2.54 centimetres in an inch, and 12 inches in a foot.

(4)

- (h) Write a Python function, **email\_to\_domain**, that extracts the domain part from an email address. For example, given the email address "**brian.macnamee@ucd.ie**" the function should return "**ucd.ie**".

(4)

- (i) Fill in the code to complete the following function for computing the sum of the *harmonic series* up to a specified limit  $n$ .

**Note:** The harmonic series is the series  $1/1, 1/2, 1/3, 1/4, 1/5, \dots$  continuing to infinity. To calculate the sum of the harmonic series up to a specified limit we simply add the values of the series together up to that limit. For example, the sum of the harmonic series up to the limit 5 is  $1/1 + 1/2 + 1/3 + 1/4 + 1/5 = 2.283$ .

```
def harmonic_sum(n):  
    if n < 2:  
        return 1  
    else:  
        _____ # Complete this line
```

(4)

- (j) Suppose that

```
s1 = {1, 2, 3, 4}
```

and

```
s2 = {6, 5, 4, 3}
```

What is the value of the expression **s1.difference(s2)** ?

(4)

2. The following Python code declares a class used to represent an Olympian object.

```
# A class to represent an Olympian
class Olympian:
    def __init__(self, name, dob, country):

        # Name stored as a string
        self.name = name

        # Date of birth stored as a string in the
        # form "dd/mm/yyyy"
        self.dob = dob

        # Country of citizenship
        self.country = country

        # Olympics and sports competed in
        self.competitions = dict()

    # Add competitive record
    def add_competition(year, disc, result):
        -

    # A method to print the details of the Olympian
    def show(self):
        print(name + " (" + dob + "), " + country)
```

- (a) Write Python code to create two Olympians with the following attribute values:

|          |            |          |            |
|----------|------------|----------|------------|
| Name:    | Usain      | Name:    | Sonia      |
| Dob:     | 21/08/1986 | Dob:     | 28/11/1969 |
| Country: | Jamaica    | Country: | Ireland    |

Also call the **show** method for each object created.

(5)

- (b) Although the class definition code written to answer part (a) will run, an error of type **NameError** will occur when the **show** method is called. Write a new version of the **show** method that corrects the error.

(5)

- (c) (i) Complete the code for the **add\_competition** method in the **Olympian** class. This should add the details of a competition to the **competitions** dictionary stored within an **Olympian** object.

The **competitions** dictionary is indexed by year and the values stored are a list of tuples describing a discipline and a result. Valid results are "gold", "silver", "bronze", "no medal", and "disqualified". For example, the following dictionary might be included for an **Olympian**:

```
{2012:[ ("100m", "gold"), ("200m", "gold")],  
 2016:[ ("100m", "gold"), ("4x100m", "no medal")]}
```

(10)

- (ii) Update the **print** method for the **Olympian** class to print the details of the **Olympian**'s competition record.

(10)

3. (a) Describe using pseudo code, and diagrams where appropriate, how the **bubble sort** algorithm works.

(15)

- (b) Show the steps the bubble sort algorithm would take to sort the following list:

[49, 17, 97, 6, 72]

(10)

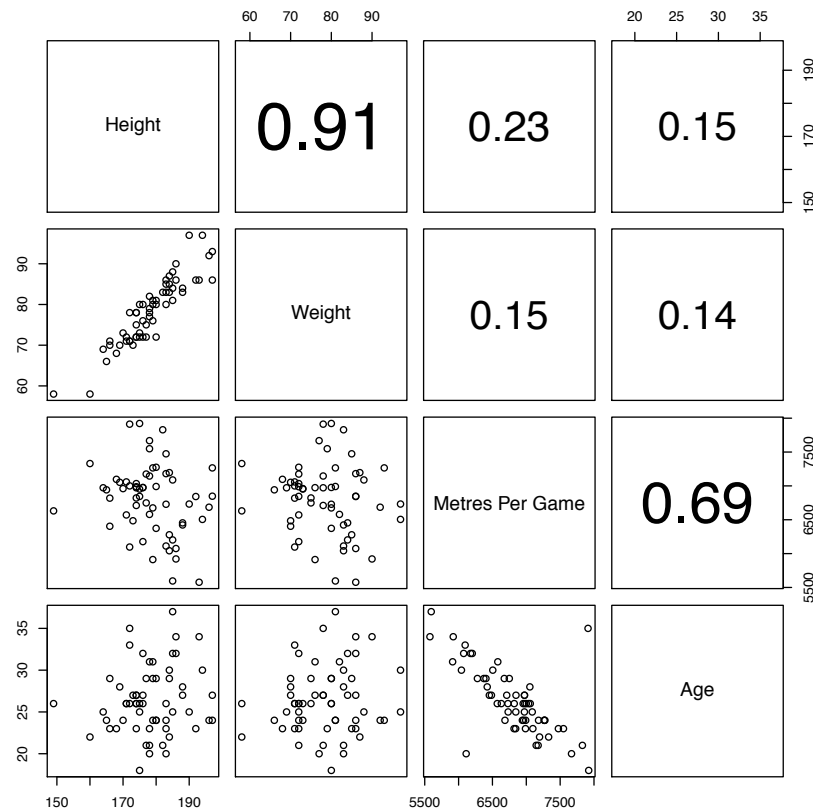
- (c) You and two colleagues have been asked by the UCD exams office to write a short Python script. The script should search a list of the student numbers of every student currently registered at UCD to check if it contains three particular student numbers. The list of student numbers that the exams office have is not sorted and contains over 30,000 student numbers. It is expected that the script you are being asked to write will only be used once.

One of your colleagues has suggested that a simple **linear search** would be the most efficient approach to this problem. Your other colleague has suggested that it would be more efficient to first use **insertion sort** to order the list and then perform a **binary search**.

Which colleague is correct?

(5)

4. (a) The image below shows a **scatter plot matrix** (SPLOM) of a dataset that describes the players on a soccer team. The variables included are players' height, weight, and age; and the average number of metres run by a player per game. Describe the relationships between these four variables that are apparent from this data visualisation.



(10)

- (b) The following table shows the first few rows from a **pandas data frame** referenced by the Python name **billionaires**. The full table contains over 1,000 entries. Net worth values are given in dollars.

| Name       | Age | Citizenship   | Net Worth      |
|------------|-----|---------------|----------------|
| Gates      | 62  | United States | 86,000,000,000 |
| Buffett    | 87  | United States | 75,600,000,000 |
| Bezos      | 54  | United States | 72,800,000,000 |
| Ortega     | 81  | Spain         | 71,300,000,000 |
| Zuckerberg | 33  | United States | 56,000,000,000 |
| Slim Helu  | 78  | Mexico        | 54,500,000,000 |
| ...        |     |               |                |

Write Python code (**using the pandas library**) to perform the following operations:

- (i) Select only the **Name** and **Age** columns from the data frame.  
(4)
- (ii) Calculate the **maximum** value for the **Age** and **Net Worth** columns of the data frame.  
(4)
- (iii) Select only the rows from the data frame that describe **people younger than 40**.  
(4)
- (iv) Determine the most common citizenship in the table.  
(4)
- (v) Create a new column in the data frame called **Years Per Billion** that contains the average number of years that it has taken each person to earn a billion dollars.  
(4)