



University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

SEMESTER II EXAMINATION – 2015/2016

COMP 10020

Introduction to Programming 2

Prof. S. Dobson

Prof. P. Cunningham

Dr. B. Mac Namee *

Time Allowed: 1 Hour

Instructions for Candidates

Answer question 1 and two other questions. Total marks available 100.

Instructions for invigilators

1. (a) What is the output of `print str[2:5]` if `str = 'Hello World!'`?
(4 marks)

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

```
d = {"john":40, "peter":45}
```

What happens when the following code is executed:

```
d["lucy"] = 34
```

(4 marks)

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

- `__b`
- `2ndValue`
- `for`
- `my String`

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

(4 marks)

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

```
i = 0  
while i < 5:  
    print( i )  
    i += 1  
    if i == 3:  
        break  
print(0)
```

(4 marks)

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

```
j = i = 1  
i += j + j * 5  
print(i)
```

(4 marks)

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

(4 marks)

- (g) The following code snippet defines the start of a class to represent members of a club:

```
class Member:

    member_count = 0

    def __init__(self, name, number):
        self.name = name
        self.number = number
        member_count += 1
        testVal = 14
```

Name the *class variables* and the *instance variables* in the code given above.

(4 marks)

- (h) Write a **toString** method for the class defined in part (g) that returns a string containing the club member's name and number as follows:

Member: "Bob" [23456]

Where the club member's name is Bob and club membership number is 23456.

(4 marks)

- (i) Fill in the code to complete the following function for computing the *factorial* of an integer.

Note: The factorial of an integer is the product of the integer and all integers below it; e.g. factorial four (*4!*) is equal to $4*3*2*1 = 24$ and factorial six (*6!*) is equal to $6*5*4*3*2*1 = 720$

```
def factorial(n):
    if n == 0:
        return 1
    else:
        _____ # Complete this line
```

(4 marks)

- (j) The following code snippet defines a class in Python:

```
class JustCounter:  
    def __init__(self, count):  
        self.__secretCount = count  
    def count(self):  
        self.__secretCount += 1
```

Given the definition of the **JustCounter** class, what happens when the following code is executed?

```
counter = JustCounter(1)  
counter.count()  
print counter.__secretCount
```

(4 marks)

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

```
# A class to represent an athlete
class Athlete:

    def __init__(self, name, dob, country, qualified):

        # First 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 stored as a string
        self.country = country
        # A boolean variable to indicate whether the
        # athlete ever qualified for an olympic games
        self.qualified = qualified

    # A method to print the details of the athlete
    def show(self):

        print(name + " (" + dob + ") " + country )
```

- (a) Although the class definition code 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 marks)

- (b) Write python code to create two athletes with the following attribute values and to call the **show** method for each of them:

Name:	Helga	Name:	Patrick
Country:	Germany	Country:	Ireland
Dob:	01/01/1989	Dob:	11/08/1993
Qualified:	True	Qualified:	False

(10 marks)

- (c) Write a new Python class to represent runners that *inherits* from the **Athlete** class. The **Runner** class should have the following characteristics:
- Include an extra *instance variable* called **race_times** that stores a list of the runner's recent times in seconds.
 - To initialise a **Runner** object, the name, date of birth, country and qualified values should be initialised with values supplied as parameters and an empty list of recent race times should be created.
 - Include a method to add a race time (in seconds) to the list of race times
 - A **show** method that performs the same operation as the **Athlete** show method but also prints the list of race times

(15 marks)

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

(15 marks)

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

[54, 26, 93, 17, 77]

(10 marks)

- (c) A colleague has suggested to you that *insertion sort* is a faster approach to sorting than *bubble sort*. Is your colleague correct? Why or why not?

(5 marks)

4. The table below shows the first 10 rows from a dataset describing the players in a basketball league. In total the dataset contains 100 rows. This dataset will be used to build a predictive model that can be used to predict expected sponsorship earnings based on a player's other attributes.

ID	Position	Height	Weight	Career Stage	Age	Sponsorship Earnings
1	forward	186	216	rookie	18	2166
2	guard	192	207	rookie	20	1473
3	forward	195	222	veteran	34	
4	guard			mid-career	30	86
5	guard	174	194	veteran	29	477
6	guard	175	195	rookie	18	2201
7	forward	185	217	mid-career	28	805
8	center			mid-career	27	598
9	guard	184	207	veteran	37	1128
10	forward	204	222	mid-career	33	-453
...

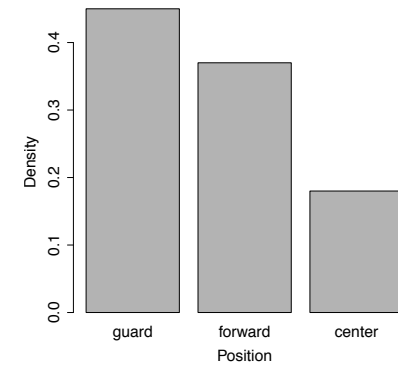
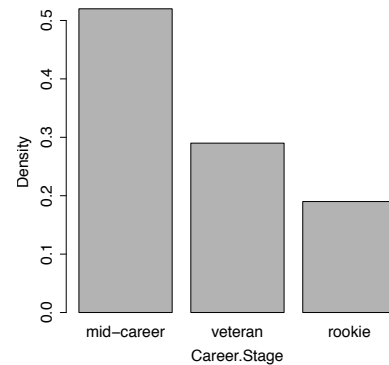
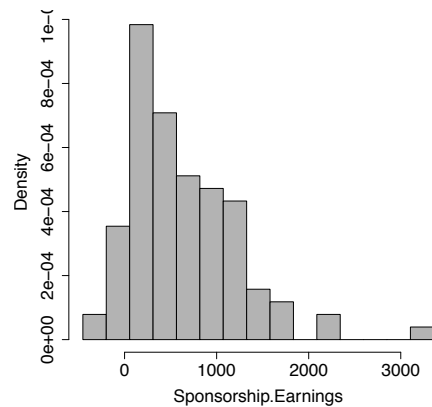
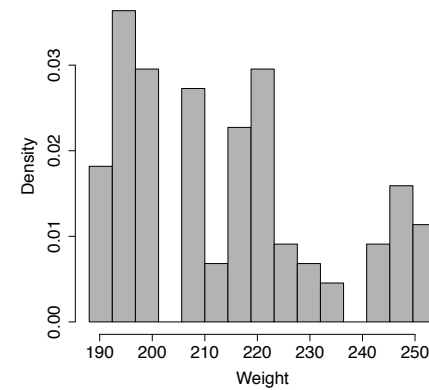
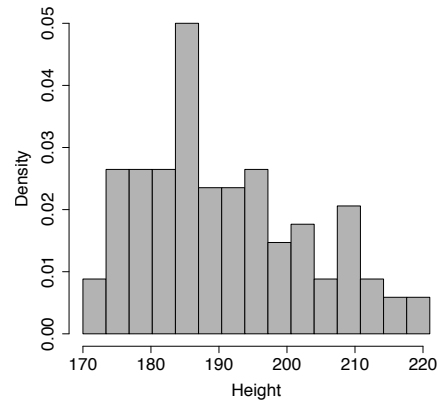
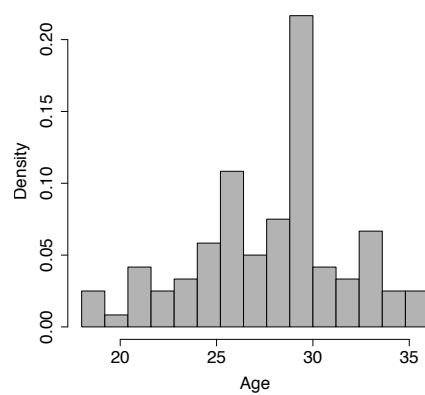
- (a) Page 8 shows a data quality report for this dataset. Perform an analysis of data quality issues that are apparent from this data quality report. **(15 marks)**
- (b) The **scatter plot matrix** (SPLOM) on page 9 visualises the relationships between the continuous features within the basketball player dataset. From a **data exploration** perspective, what can be learned from this visualisation? **(10 marks)**
- (c) The box plots on page 9 illustrate the relationships between the continuous target feature from the basketball dataset, *Sponsorship Earnings*, and the two categorical descriptive features, *Position* and *Career Stage*. From a **data exploration** perspective, what can be learned from these visualisations? **(5 marks)**

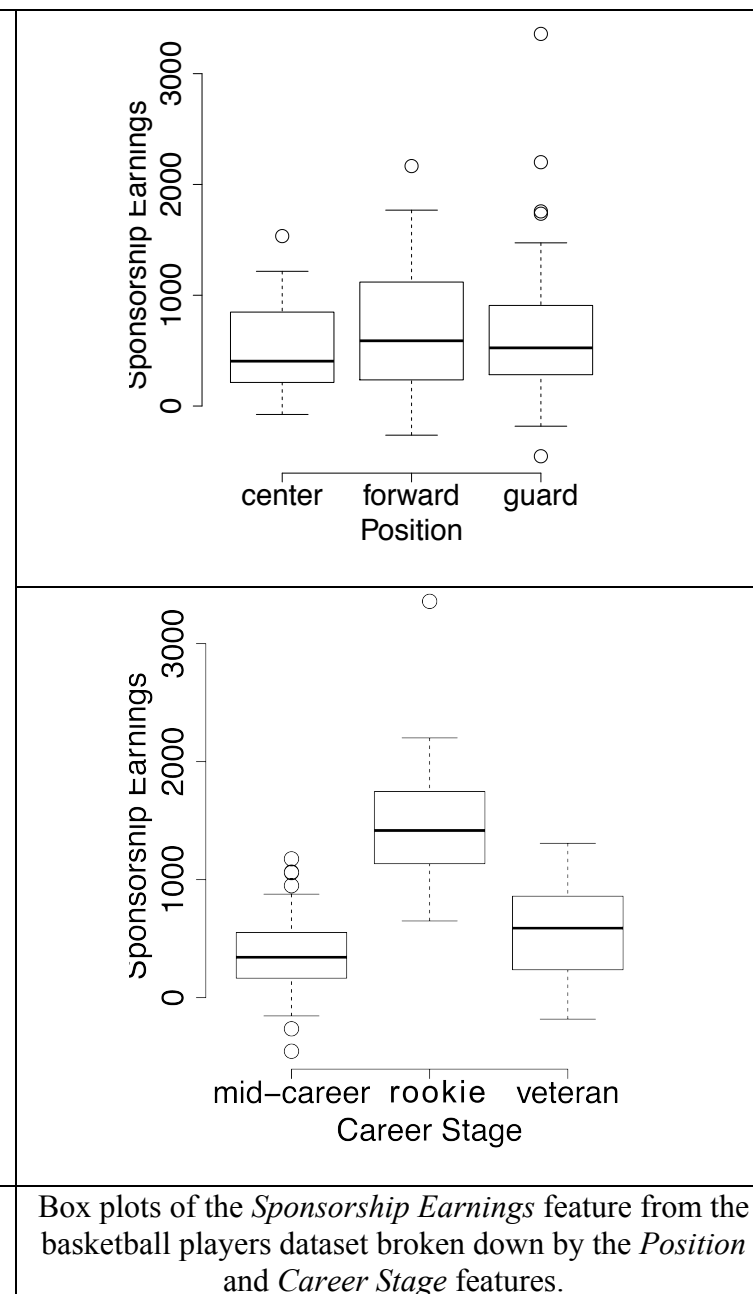
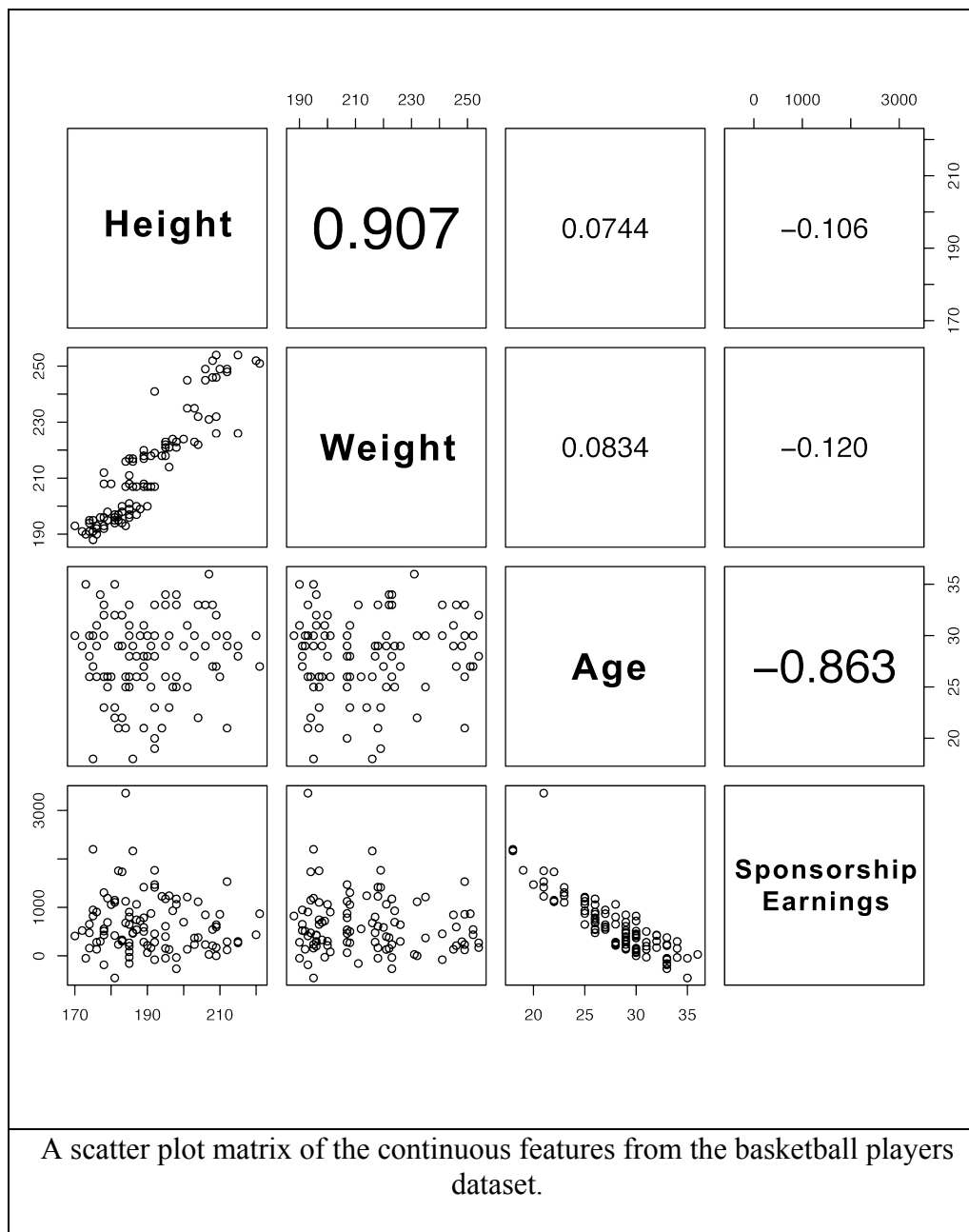
Continuous Features Data Quality

Feature	% Missing	Cardinality	Min	1st Quart.	Mean	Median	3rd Quart.	Max	Std. Dev.
Height	11	40	170	182	191	189	198	221	12
Weight	11	39	188	197	214	212	223	254	19
Age	0	18	18	26	28	28	30	36	4
Sponsorship Earnings	2	98	-453	236	642	519	937	3359	594

Categorical Features Data Quality

Feature	% Missing	Cardinality	Mode	Mode %	2nd Mode	2nd Mode %
Position	0	3	guard	45	forward	37
Career Stage	0	3	mid-career	52	veteran	29





Intentionally blank