## Data Science**: ePortfolio**

|  |  |
| --- | --- |
| First Name | Kevin |
| Last Name | Gemson |
| Student Number | EC2111447 |
| Email Address | ec2111447@edinburghcollege.ac.uk |

## Part 1: Title – **Data Capture & Modelling**

## Date Completed: **31/03**/2022

**Weekly summary, complete the following:**

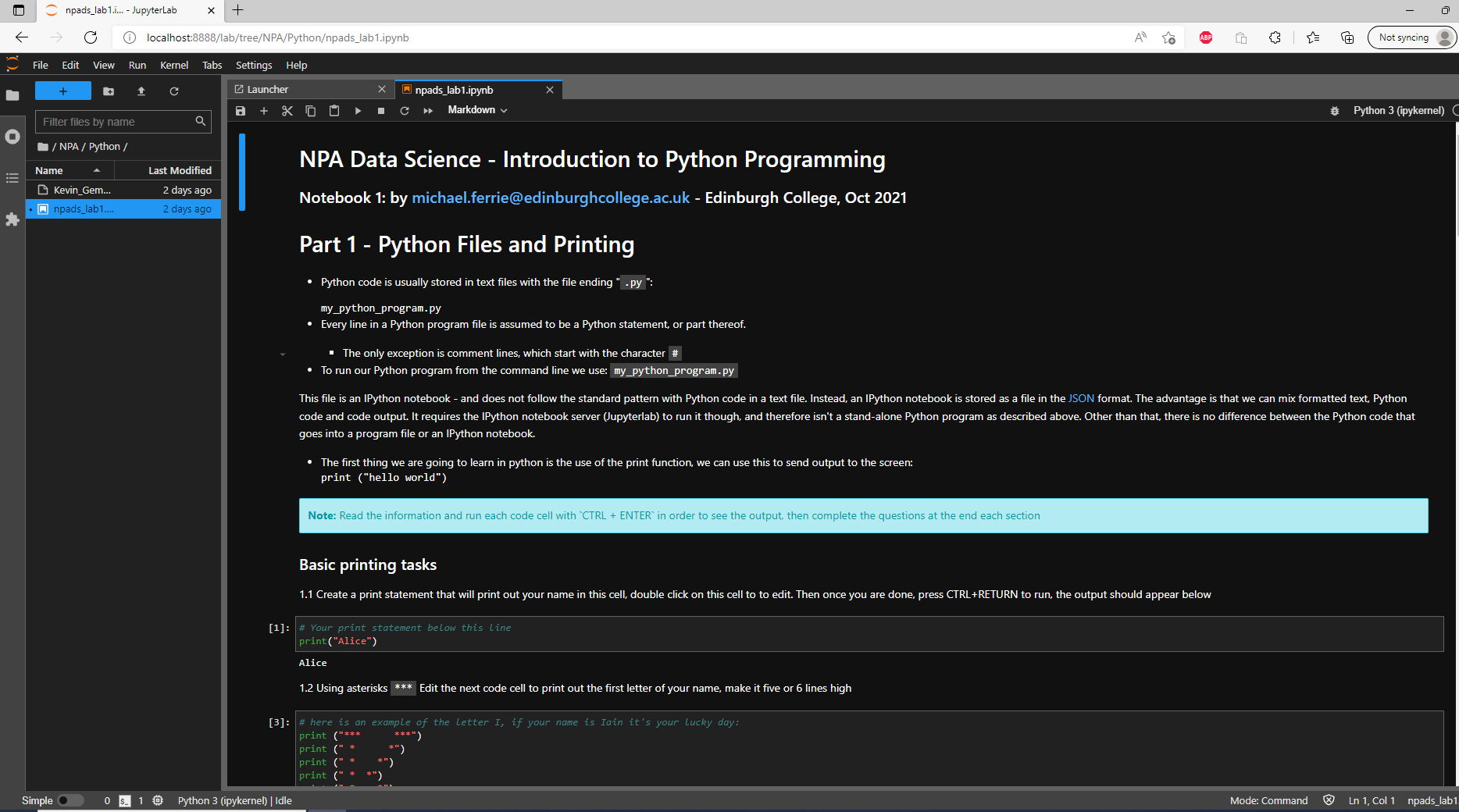
This week I learned … how to install Jupyterlab and get Python scripts running on it. From the PDF, I learned about different data capture techniques, as well as different ways of cleaning and validation. Also reviewed Python basics from the YouTube video

In the lab we … installed Jupyterlab and ran some basic Python commands

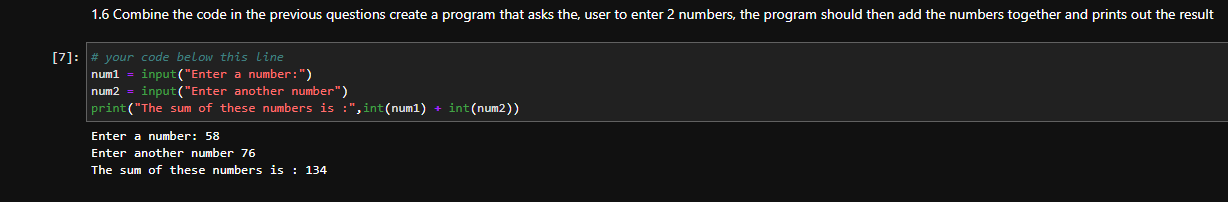
I enjoyed… learning about different data capture and cleaning techniques; having worked in a company that handles and processes large volumes of data, it is interesting to see the different techniques available, many of which I have not come across before. The company has tended to use traditional methods of handling data in the past, and it will be interesting to see which of these I may come across in the future.

I found it challenging when… I had some issues getting Jupyterlab to work; I had used Python before, running scripts through Visual Studio Code, but Jupyterlab works in a different way and took a bit of getting used to

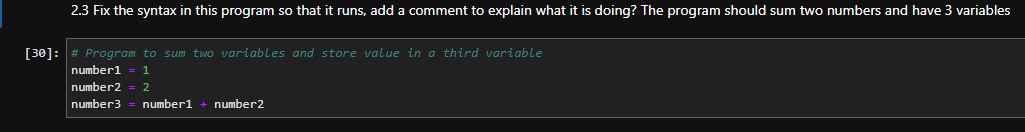
* A screenshot of Jupyterlab running on your computer



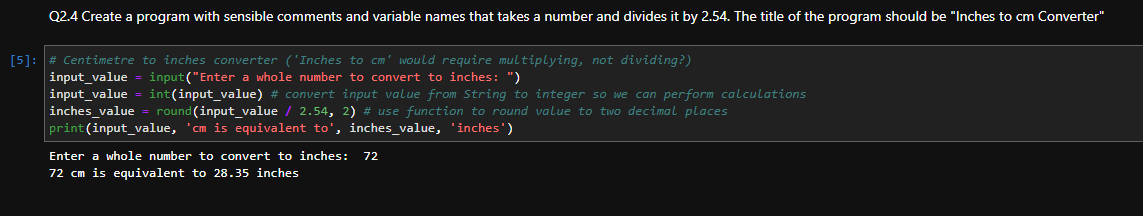
* A screenshot of your answer for question 1.6



* A screenshot of your answer for question 2.3



* A screenshot of your answer for question 2.4



## Lab 2: Title – **Data Analysis & Statistics**

## Date Completed: 03/04/2022

**Weekly summary, complete the following:**

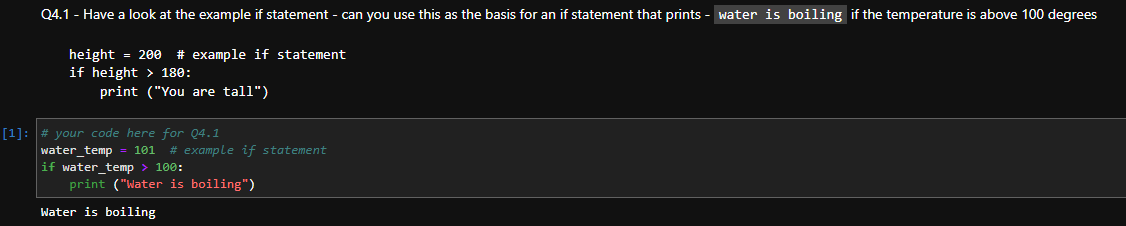
This week I learned … about the different steps in Data Analysis and how these can lead to allowing us to making predictions about future trends.

In the lab we … revised some basic operators and ran some scripts to demonstrate handling of different data types and conditionals

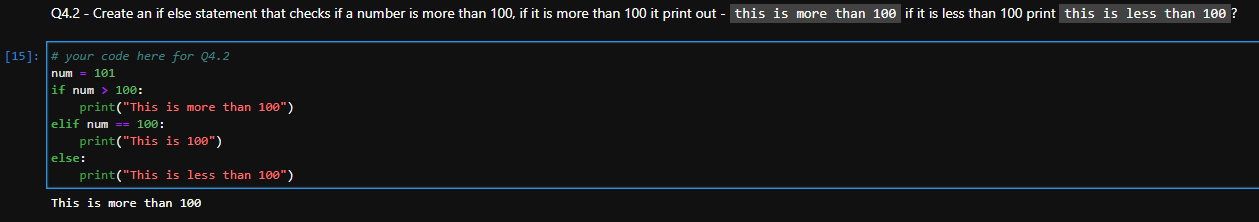
I enjoyed… revising some of the basic commands and taking on some of the programming challenges.

I found it challenging when … trying to get the appropriate data type for many of the scripts (e.g. converting string to int and performing division to get a float). Also considering the different options when using conditionals to ensure all outcomes are catered for.

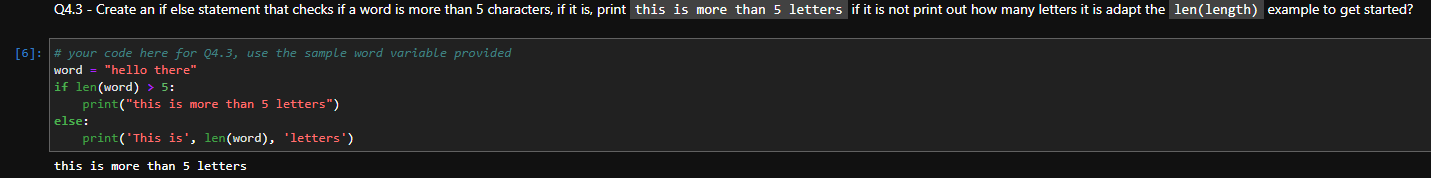
* A screenshot of the code in the code cell and the output for question 4.1



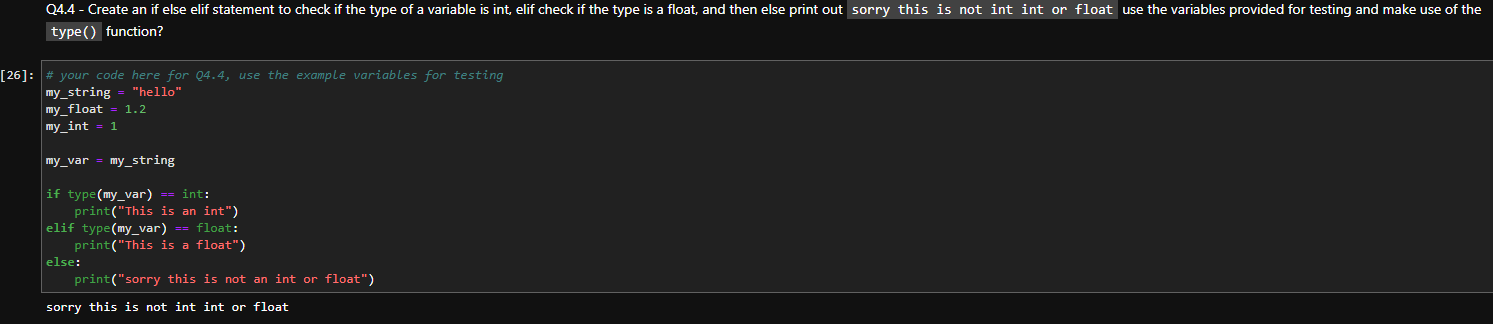
* A screenshot of the code in the code cell and the output for question 4.2



* A screenshot of the code in the code cell and the output for question 4.3



* A screenshot of the code in the code cell and the output for question 4.4



## Lab 3: Title – Data **Visualisation and Storytelling**

## Date Completed: **07/04/2022**

**Weekly summary, complete the following:**

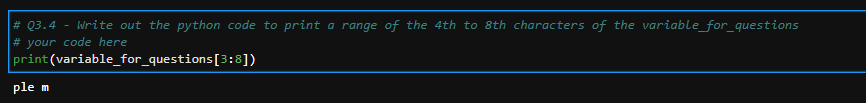
This week I learned… about data visualisation and the different options for presenting data based on what the data is trying to present. In Python, we learned about extracting individual elements from lists and also program control – for loops in conjunction with if/else statements.

In the lab we… ran a number of scripts to extract individual characters of ranges of characters from a string variable; also ran some scripts to take input data and process differently based on the values provided. Final challenge was to randomly generate a data value and handle the output based on the value.

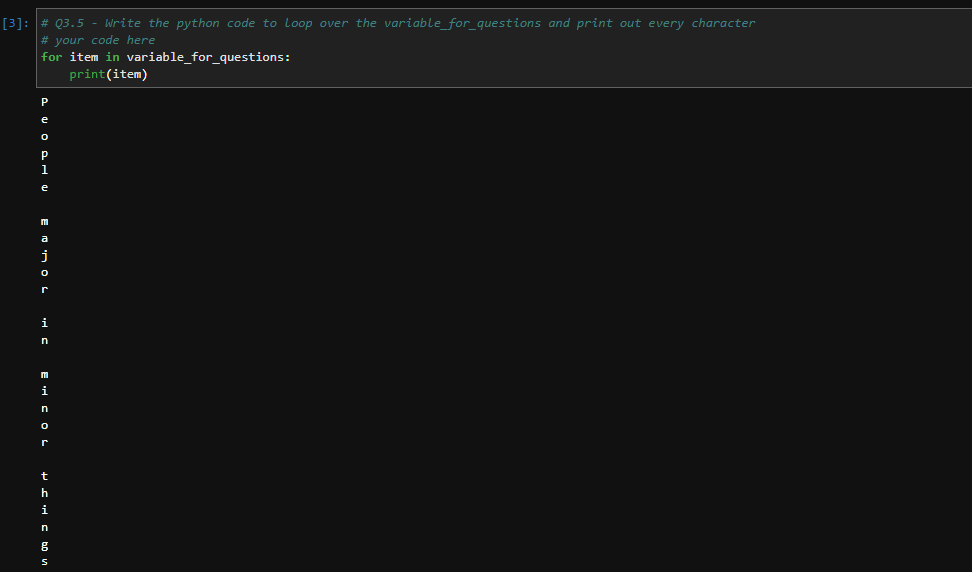
I enjoyed… the challenge of working through the different possible outcomes and writing code to handle these.

I found it challenging when… trying to work out the indexing values - I have used other programming languages that index starting at 1 and where ranges are inclusive, so ensuring I am extracting the correct values can take some thought.

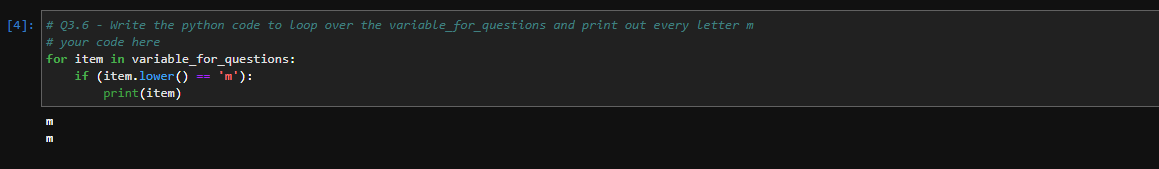
* A screenshot of the code in the code cell and the output for question 3.4



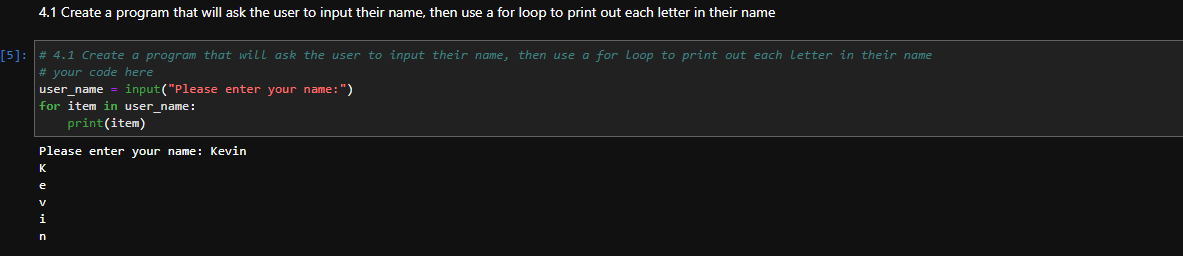
* A screenshot of the code in the code cell and the output for question 3.5



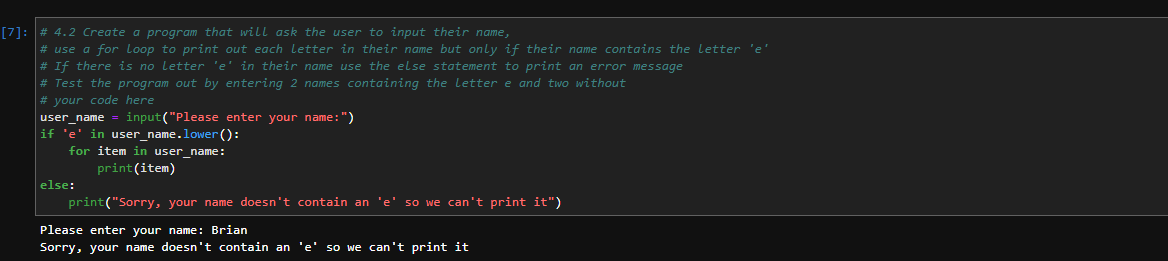
* A screenshot of the code in the code cell and the output for question 3.6



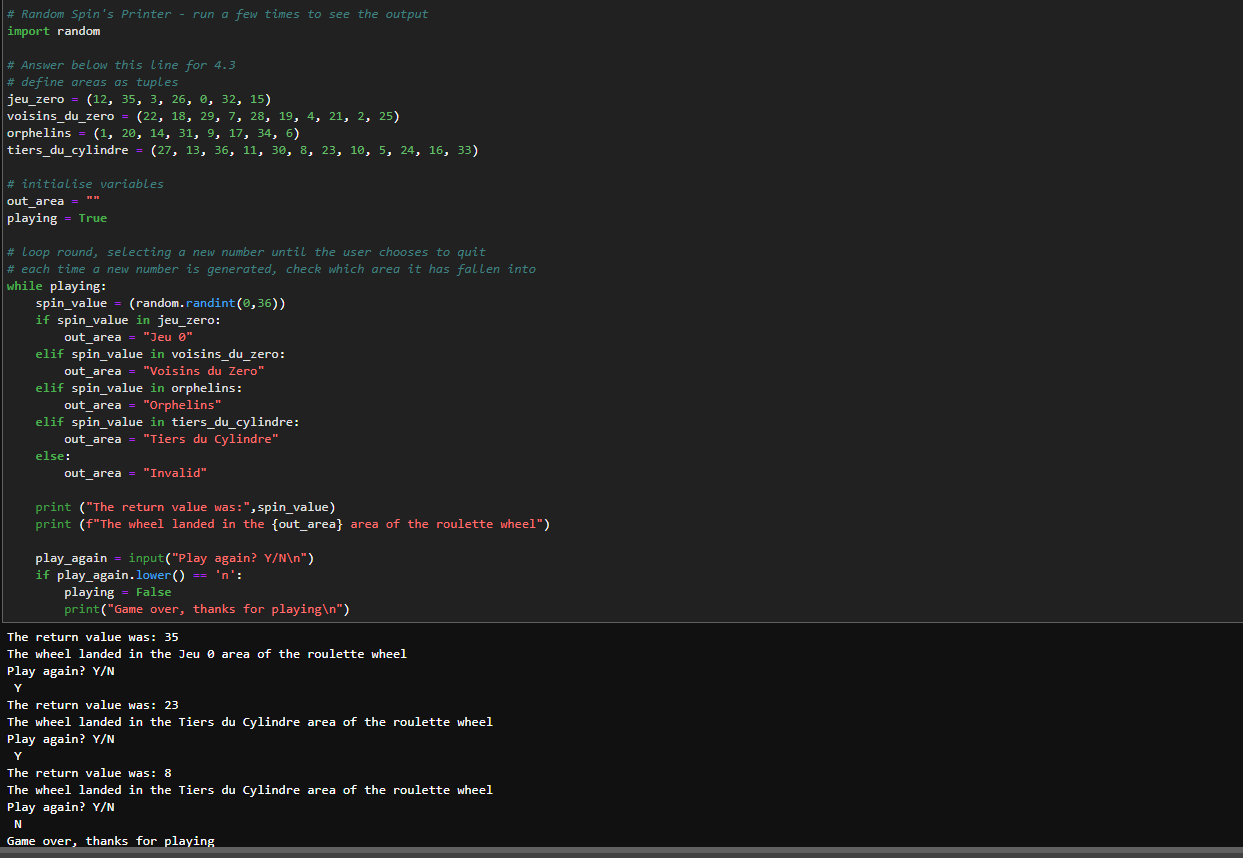
* A screenshot of the code in the code cell and the output for question 4.1



* A screenshot of the code in the code cell and the output for question 4.2



* A screenshot of the code in the code cell and the output for question 4.3



## Lab 4: Title – Computer Programming 1

## Date Completed: 12**/05/2022**

**Weekly summary, complete the following:**

An explanation of two different techniques for representing algorithms, how can we express an algorithm, why would we use these techniques?

An explanation of the following three concepts:

- algorithmic efficiency

- syntax

- semantics

The flow chart and decomposition for 2.2, 2.3 and 2.4

2.2 The next algorithm to create is one to check if a word contains a vowel ‘aeiou’:

* Write out your algorithm in decomposition notation.

> Read input word, set ‘match’ flag to ‘False’

> Check each letter to see if it is in ‘aeiou’

>> read next letter of word

>> if no more letters, exit loop

>> else, check to see if letter is in ‘aeiou’

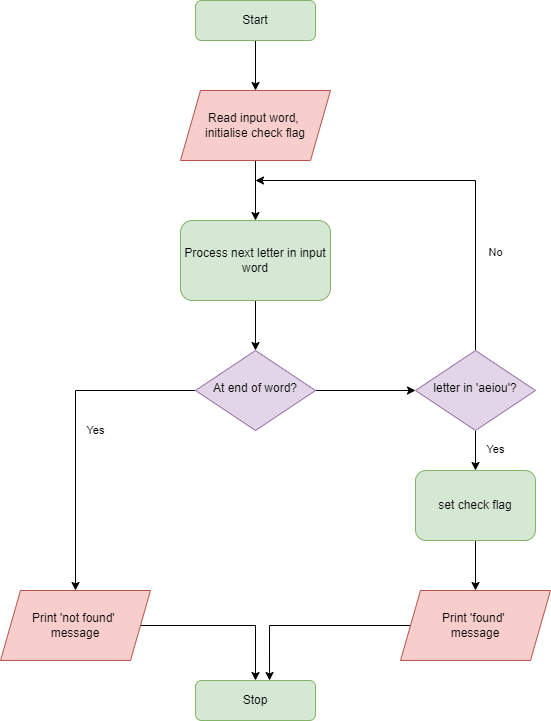
>>> if yes, set ‘match’ flag to ‘True’ and exit loop

> Print output message

>> if ‘match’ flag is ‘True’, print ‘match’ message

>> else print ‘no match’ message

* Draw your algorithm as a flowchart.



2.3 This algorithm should be an enhancement of the previous, in 2.2 you created an algorithm to determine if a word contained a vowel, can you improve this and say how many vowels - if any - a word contains?

* Write an algorithm that is an improvement of the algorithm in 2.2 in decomposition notation; this should now calculate the number of vowels in a word.

> Read input word, set counter to zero

> Check each letter to see if it is in ‘aeiou’

>> read next letter of word

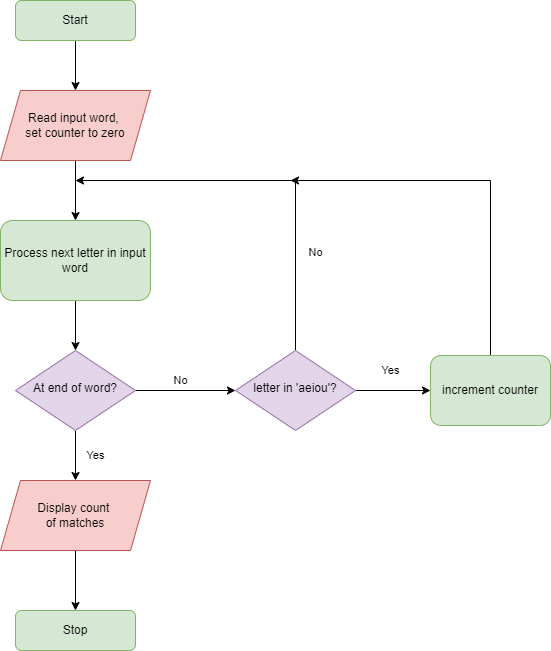
>> if no more letters, exit loop

>> else, check to see if letter is in ‘aeiou’

>>> if yes, increment counter

> Print output message containing count of matches

* Write a flowchart to represent your algorithm.



* 2.4 The next algorithm we want to create is an algorithm to check if a number can be divided by 3 without a remainder. The algorithm should then check if the number can be divided evenly by 4. Then the algorithm should check if the number can be divided evenly by both 3 and 4.
* Write out your algorithm in decomposition notation.
* Draw your algorithm as a flowchart.