CS1117: Exercise Sheet Semester 2 Week 3

Question 1 (Files):

You have been given a file *dna.txt*. Open *dna.txt* in a text editor and have a look at it. It contains DNA sequences related to several people. It only makes sense to process this using the `read()` method (Why? If you don't understand ask a demonstrator).

Create a list that contains each person's DNA, using the read method to read from the file. Use a loop to print out the persons ID (starting at 0) and their associated DNA sequence. The output should look as follows:

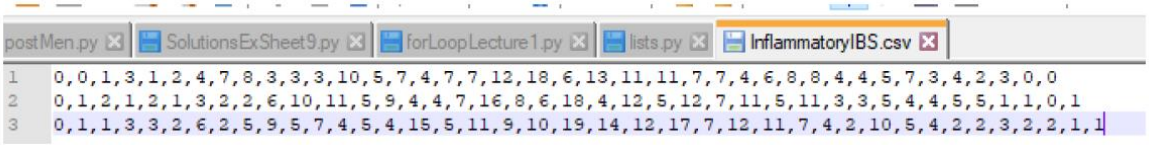
```
0 : CCACTGCACT….
1: GGCAGATTCC…
2: AAAAAAAAT…
etc
```

As you will need to pull out every person's DNA from the string you will need to split the string into parts (which are saved as a list). You can do this using the split method.

Question 2 (Files):

You have been given a file called InflammatoryIBS.csv. This contains data relating to a new medication for the inflammatory condition of Irritable Bowel Syndrome (IBS). The data is stored in comma-separate values (CSV) format.

A snippet of the file is shown below as an example.



Each number represents the number of inflammatory bouts that a particular patient experienced on a given day. The number "6" in row 3 (7th value from left) means that the third patient was experiencing inflammation six times on the seventh day of the clinical trial.

You want to:

1. Calculate the average number of inflammatory bouts per patient and print it to the screen along with their patient ID. This will occur in a function called *meanBoutsPerPatient*(). The average value should be rounded to a whole number. Complete this in AssignmentSolutionPart1.py
2. Calculate the average number of inflammatory bouts across all patients and print it to the screen. This will occur in a function called *meanBoutsAcrossAllPatients*(). This number should not be rounded. Complete this in AssignmentSolutionPart1.py

```
Patient 1 had 5 inflammatory bouts on average
Patient 2 had 5 inflammatory bouts on average
Patient 3 had 6 inflammatory bouts on average
The average number of inflammatory bouts on this trial medication is:  5.66
```

3. In AssignmentSolutionPart2.py, modify your code in *meanBoutsPerPatient*() so that instead of displaying this to the screen you are instead returning a list (meanPerPatient). This list contains the patient ID and mean bouts in individual patient lists e.g. patient 1 has a mean of 5.45, patient 2 has a mean of 5.425 etc.

```
[[1, 5.45], [2, 5.425], [3, 6.1]…]
```

Pass this list to a function called *writeToFile()* that writes the following lines to *meanBoutsPerPatient.txt* by processing the list.