Programming For Data Analytics
Lab Topic 02-Representing Data

CSV:

1. Create a CSV file with the following data, call it data.csv (you can copy and paste this data from here).

```
"id","age","name"
1,20,"Joe"
2,21,"Mary"
3,32,"Fred"
```

2. Write a program to read in the data and output each line as a list.

```
import csv

FILENAME= "data.csv"
DATADIR = “where did you put it”

with open (DATADIR + FILENAME, "rt") as fp:
    reader = csv.reader(fp, delimiter=",")
    for line in reader:
            print (line)
```

Note what is printed. What data type are these?

3. Modify the program to deal with the header line separately

```
# there is code not shown here
with open (DATADIR + FILENAME, "rt") as fp:
    reader = csv.reader(fp, delimiter=",")
    linecount = 0
    for line in reader:
        if not linecount: # first row ie header row
            print (f"{line}\n--------------------")
        else: # all subsequent rows
            print (line)
        linecount += 1
```

4. Modify the program to calculate the average age, there are a few ways to solve this;
    a. Convert the string that is read into an integer

```
with open (DATADIR + FILENAME, "rt") as fp:
    reader = csv.reader(fp, delimiter=",")
    linecount = 0
    total = 0
    for line in reader:
        if not linecount: # first row ie header row
            pass
        else: # all subsequent rows
            total += int(line[1]) # why 1
        linecount += 1
    print (f"average is {total/(linecount-1)}") # why -1 ?
```

    b. Use the quote parameter to read in the numbers as floats

```
with open (DATADIR + FILENAME, "rt") as fp:
    reader = csv.reader(fp, delimiter=",", quoting=csv.QUOTE_NONNUMERIC)
    linecount = 0
    total = 0
    for line in reader:
        if not linecount: # first row ie header row
            pass
        else: # all subsequent rows
            total += line[1] # why 1

        linecount += 1
    print (f"average is {total/(linecount-1)}") # why -1 ?
```

5. The CVS file could of course have been read in as a Dictionary object
   Using DictReader()

```
# some code is deleted from here
with open (DATADIR + FILENAME, "rt") as fp:
    reader = csv.DictReader(fp, delimiter="," , quoting=csv.QUOTE_NONNUMERIC)
    total = 0
    count = 0
    for line in reader:
        total += line['age']
        # print (line)
        count +=1
    print (f"average is {total/(count)}") # why is there no -1 this time?
```

Read JSON from internet

6. Copy this URL into browser and see the JSON it returns.
https://www.gov.uk/bank-holidays.json
7. Write a program to print this JSON to the console.

```
import requests

url =" https://www.gov.uk/bank-holidays.json"
response = requests.get(url)
data = response.json()
print(data)
```

   Is this JSON or a Dict object that is outputted.

8. Modify the program to only output the first holiday in northern ireland

```
print(data[' northern-ireland'][‘events’][0])
```