Week-9, Activity Questions

NOTE

- Fork this repl to for all problems related to files: https://replit.com/@pypod/week9files
- You can click on the `Fork` button on the right-top corner.
- Forking a repl creates your own local copy that you can play around with.
- All the files needed for working out the activity questions is given as a part of this repl. The files are named `file_i.txt` or `file_i.csv`.

Problem-1

Read the file `file_1.txt`.

- Print the contents of the file, one line at a time.
- Print the type of the variable used to store the contents of each line in the file. Is it `int`, `str`, `float` or something entirely different?
- Find the number of lines in the file.

Answer

Solution-1

```
f = open('file_1.txt', 'r')

for line in f:
    print(line.strip())

f.close()
```

Solution-2

```
f = open('file_1.txt', 'r')

line = f.readline()
while line:
    print(line.strip())
    line = f.readline()

f.close()
```

Solution-3

```
f = open('file_1.txt', 'r')

lines = f.readlines()
for line in lines:
    print(line.strip())

f.close()
```

Problem-2

The file `file_2.txt` is a collection of integers, one on each line. Read the file and print the sum of all the integers in it.

Answer

```
f = open('file_2.txt', 'r')

S = 0
line = f.readline()
while line:
    S += int(line)
    line = f.readline()

f.close()
print(S)
```

Problem-3

The file `file_3.txt` captures the maximum temperature recorded in Chennai on every day during the month of May, 2021. Find the average temperature for the month.

Note: Data is not accurate.

Answer

```
f = open('file_3.txt', 'r')

S, count = 0, 0
line = f.readline()
while line:
    S += float(line)
    count += 1
    line = f.readline()

print(f'{S / count:.2f}')
```

Problem-4

Consider the following string:

```
'this is a sentence that has ten words in it'
```

Split this sentence into a list of words. Call this list `words_1`.

```
'this is another sentence that has ten words in it'
```

Split the above sentence into a list of words. Call this list `words_2`.

Finally, create a list `words` that has `words_1` and `words_2` as two elements in it. What type of a list is `words`?

Note: Solve this problem before proceeding to the next one.

Answer

```
words_1 = 'this is a sentence that has ten words in it'.split(' ')
words_2 = 'this is another sentence that has ten words in it'.split(' ')
words = [words_1, words_2]
```

Problem-5

The file `file_4.txt` has a number of sentences on each line. Each sentence is made of certain number of words. There are no punctuation marks. All words are in lower case. With this information, answer the following questions:

- Find the number of sentences in the file.
- Find the number of words in each sentence.
- Find the sentence with the maximum number of words in it
- Find the frequency of occurrence of words in the file.
- Find the set of unique words used in the file.

Answer

Basic code to read lines of the file.

```
f = open('file_4.txt', 'r')

sentences = [ ]
line = f.readline()
while line:
    words = line.strip().split(' ')
    sentences.append(words)
    line = f.readline()

f.close()
```

Most of the other problems are straightforward. We will just work on finding the frequency of occurrence of words in the file:

```
P = { }
for sentence in sentences:
    for word in sentence:
        if word not in P:
            P[word] = 0
        P[word] += 1
```

Problem-6

`file_5.txt` is a text file that stores an $m \times n$ matrix. Each line represents one row of a matrix. Adjacent numbers are separated by a space. Write a code to extract the contents of the file and store it in a matrix `M`. What is the dimension of the matrix?

Answer

```
f = open('file_5.txt', 'r')

M = [ ]
line = f.readline()
while line:
    row = line.strip().split(' ')
    for i in range(len(row)):
        row[i] = int(row[i])
    M.append(row)
    line = f.readline()

print(M)
```

Problem-7

Open a file `file_6.txt` and write the first 100 perfect squares to it, one on each line. The first five lines should look like this:

```
1
4
9
16
25
```

Answer

```
f = open('file_6.txt', 'w')

for i in range(1, 101):
    line = f'{i ** 2}'
    if i != 100:
        line = line + '\n'
    f.write(line)

f.close()
```

Problem-8

Open a file `file_7.txt` and write the following pattern to it.

```
(a)
(b)
(c)
...
```

There should be 26 lines in the file. You get the idea.

Answer

```
import string
f = open('file_7.txt', 'w')

count = 0
for char in string.ascii_lowercase:
    line = f'({char})'
    count += 1
    if count != 26:
        line += '\n'
    f.write(line)

f.close()
```

Problem-9

Open a file `file_8.txt` and write the following pattern to it. The first three lines are given for your reference:

```
1 2 3 4 5
6 7 8 9 10
11 12 13 14 15
```

The file should have 100 lines. That is, this pattern should continue for 97 more lines.

Answer

```
f = open('file_8.txt', 'w')

start, end = 1, 100 * 5 + 1
for num in range(start, end):
    line = f'{num}'
    if num != end - 1:
        if num % 5 == 0:
            line += '\n'
        else:
            line += ' '
    f.write(line)

f.close()
```

Problem-10

Consider the file `file_5.txt` that was discussed in problem-6. Create a file `file_9.txt`, where each line contains one number from the matrix given in `file_5.txt`. Proceed row-wise, left to right and top to bottom: fetch one number at a time from the first row of the matrix and write it to the file. Each line in the file corresponds to exactly one entry in the matrix. Then move to the second row of the matrix and so on.

Answer

```
inp_f = open('file_5.txt', 'r')
lines = [ ]
for line in inp_f:
    for num in line.strip().split(' '):
        lines.append(num)
inp_f.close()

out_f = open('file_9.txt', 'w')
start, end = 0, len(lines)
for i in range(start, end):
    line = lines[i]
    if i != end - 1:
        line += '\n'
    out_f.write(line)
out_f.close()
```

Problem-11

Consider the following file. It contains information about the runs scored by Kohli in a subset of the ODIs he has played. The first five lines are given below:

```
Opponent,Innings,Runs,Result
PAK,2,101,W
AUS,1,89,L
BAN,2,110,W
WI,1,65,W
```

- Find the number of matches in which Kohli has scored a century while chasing.
- How many centuries has Kohli made in a losing cause?

Answer

Basic code for extracting the contents from the file and storing it in a collection.

```
f = open('file_10.csv', 'r')

header = f.readline()   # ignore the header
mat = [ ]
line = f.readline()
while line:
    words = line.strip().split(',')
    opp, inn, runs, result = words
    inn = int(words[1])
    runs = int(words[2])
    mat.append((opp, inn, runs, result))
    line = f.readline()

f.close()
```

Problem-12

`students` is a list of dictionaries that has information on students in a class. The first two entries are given below:

```
scores = [ { 'Chemistry': 78,
    'CityTown': 'Erode',
    'DateOfBirth': '7 Nov',
    'Gender': 'M',
    'Mathematics': 68,
    'Name': 'Bhuvanesh',
    'Physics': 64,
    'SeqNo': 0,
    'Total': 210},
  { 'Chemistry': 91,
    'CityTown': 'Salem',
    'DateOfBirth': '3 Jun',
    'Gender': 'M',
    'Mathematics': 62,
    'Name': 'Harish',
    'Physics': 45,
    'SeqNo': 1,
    'Total': 198},
]
```

Write the student details into a comma separated file (csv). The file's header should be as follows:

SeqNo	Name	Mathematics	Physics	Chemistry

Answer

```
from scores import students # check out the repl
f = open('scores.csv', 'w')

header = 'SeqNo,Name,Mathematics,Physics,Chemistry'
f.write(header + '\n')

start, end = 0, len(students)
for i in range(start, end):
    student = students[i]
    seq_no = student['SeqNo']
    name = student['Name']
    math = student['Mathematics']
    phy = student['Physics']
    chem = student['Chemistry']
    line = f'{seq_no},{name},{math},{phy},{chem}'
    if i != end - 1:
        line += '\n'
    f.write(line)

f.close()
```