

Ejercicios Procesamiento de texto y archivos

Miguel Solis

<http://www.innovacionyrobotica.com/~miguel.solis>

Facultad de Ingeniería - Escuela de Informática

Universidad Andrés Bello

Pregunta 1

Se han realizado 1600 millones de votaciones online para definir las nuevas 7 maravillas del mundo. Todas las votaciones se almacenan en un archivo llamado *maravillas.txt*. Esta información corresponde a un código interno, nombre de la maravilla, país en el que se encuentra y la votación obtenida en porcentaje. El archivo tiene la estructura *código/nombre:país:porcentaje_de_votos*.

El siguiente es un ejemplo del archivo.

maravillas.txt

```
103/Taj Mahal:India:7.7
219/Muralla China:China:11
312/Alhambra:España:2.83
403/Mahabalipuram:India:15
```

- a) Desarrolle la función *maravillas_seleccionadas(nombre,votacion)* que reciba el nombre del archivo y un porcentaje de votos, y que retorne un diccionario solo con los lugares que tengan porcentaje de votación mayor a la ingresada como parámetro, donde la llave del diccionario será el nombre de la maravilla y el valor, la cantidad de votos en millones.

Ejemplo:

```
maravillas_seleccionadas('maravillas.txt',7)
```

```
{'Taj Mahal': 123.2, 'Muralla China': 176.0, 'Mahabalipuram':240.0}
```

- b) Haga una función *suma(archivo,país)*, que reciba el nombre del archivo y un país y que sume los porcentajes para un país dado.

Ejemplo:

```
suma('maravillas.txt','India')
```

```
22.7
```

Pregunta 2

Considere el cine PythonMark, en el cual dependiendo de la edad del cliente, puede obtener diferentes descuentos sobre el precio normal.

Los precios se encuentran en el diccionario valores:

```
valores = {'menor': 1200, 'adulto': 2000, 'pensionado': 1500}
```

Considere que sobre los 65 años el cliente se toma como *pensionado*, y bajo 18 años el cliente se toma como *menor*.

Además, al comprar una entrada, cada cliente queda registrado en un archivo *clientes.txt* que contiene su nombre, apellido, edad, y fecha en que compró la entrada, en el formato **nombre:apellido:edad:fecha**, donde la fecha se encuentra en el formato **día-mes-año**.

clientes.txt

```
Juanito:Perez:14:13-6-2014
Juan:Perez:35:13-6-2014
John:Programation:60:28-7-2014
David:Python:70:18-9-2014
Juan:Perez:35:18-8-2014
```

- a) Implemente la función *ingresos(fecha)* que reciba la tupla *fecha = (día, mes, año)*, y retorne el total de ingresos acumulados por el cine durante ese día.

Ejemplo:

```
ingresos( (13,6,2014) )
```

3200

- b) Implemente la función *clientes_frecuentes(n)*, que retorna una lista con los nombres de los clientes que han ido *n* o más veces a este cine.

Ejemplo:

```
clientes_frecuentes(2)
```

```
['Juan Perez']
```

Pregunta 3

Considere el archivo *nombres.txt*, el cual contiene el nombre de una serie de archivos sospechosos, de los cuales no se tiene certeza si se encuentran infectados por un virus.

nombres.txt

```
ar2.txt
ar2.dat
ar4.txt
ar3.txt
ar1.txt
ar5.txt
```

Al abrir cada uno de estos archivos, se encontrará con una serie de caracteres, si el archivo se encuentra infectado, encontrar la palabra '**VIRUS**' en la secuencia de caracteres.

Ejemplo de archivo infectado

```
Tfdjbmqbyjcmqpxlpiwhcpfgo qwlallvesc
fjxrj
gannwlvVIRUSduynure
opitqumhtzzhmqvezcdomq swcretoyimghgeggpxdk mhciokyhwnvcqbfbizpglpt
nlhfulyrzylhqbqslgj
nhqhiauvskftkyhmrixzwwkygvvsrclhuujy
```

Ejemplo de archivo no infectado

```
TfdjbmqbVIR USyjcqpxlpiwhcpfgoqwlallvescVIR
USfjxrj
gannwlvduynurevirus
opitqumhtzzhmqvezcdomqswcretoyimghgeggpxdkmhciokyhwnvcqbfbizpglptnlhfulyrzylhqb
nhqhiauvskftkyhmrixzwwkygvvsrclhuujyVIR-US
```

- Implemente la función *analizar(archivos)* que reciba como entrada el archivo *nombres.txt* , y cree un archivo *infectados.txt* que contenga los nombres de todos los archivos infectados con virus.
- Implemente la función *ordenar(archivos)* que reciba como entrada algún archivo que contenga los nombres de varios otros archivos (como *nombres.txt* o *infectados.txt*), y cree un archivo *ordenados.txt* que contenga los mismos nombres que estaban en el archivo original que fue recibido como entrada, pero ordenados alfabeticamente.

ordenados.txt

```
ar1.txt
ar2.dat
ar2.txt
ar3.txt
ar4.txt
ar5.txt
```

Pregunta 4

Considere el archivo *personas.txt* que contiene el nombre (con apellido), el rut (sin puntos ni dígito verificador) y fecha de nacimiento de ciertas personas, en formato dd/mm/aaaa.

personas.txt

```
Juanito Perez:18785345:09/07/1991
David Python:14123321:27/10/1986
John Programation:20987234:29/09/1999
Juan Perez:10123654:18/05/1984
```

- a) Implemente la función *mayor(personas, anno)* que recibe el archivo de texto con la información de las personas, y además recibe el año actual. Luego, la función debe crear un archivo *mayores.txt* que tendrá el nombre y la edad de todas las personas que son mayores de edad hasta ese año. (Considere que la edad viene dada sólo por la resta del año actual y el año de nacimiento).

Ejemplo:

```
mayor('personas.txt', 2013)
```

mayores.txt

```
Juanito Perez - 22
David Python - 27
Juan Perez - 29
```

- b) Implemente la función *ordenar(personas)* que recibe el archivo de texto anterior, y crea el archivo *ordenados.txt* que contiene la misma información del archivo original, pero con las personas ordenadas alfabéticamente (según su apellido).

Ejemplo:

```
ordenar('personas.txt')
```

ordenados.txt

```
Juan Perez:10123654:18/05/84
Juanito Perez:18785345:09/07/91
John Programation:20987234:29/09/99
David Python:14123321:27/10/86
```

Pregunta 5

Considere el archivo *contactos.txt* que contiene los datos de los amigos de cierta persona, en el formato que se muestra a continuación, donde aparece el nombre, país y número de teléfono móvil que tienen en Chile.

contactos.txt

Jose Miguel Aranda, Chile, 9-7852341 Phillipe Fromage, Francia, 6-8794321 Mario Campos, Chile, 8-8331735 John Michael Smith, USA, 7-6753489

- a) Implemente la función *empresa(nombre)* que toma el archivo *contactos.txt* y retorna un diccionario que contiene como llaves el nombre de cada compañía telefónica, y valores asociados a cada llave tiene un conjunto con los números que pertenecen a dicha empresa.

Considere que hay 3 empresas telefónicas, la empresa dueña de cierto número vendrá dada por el primer dígito de cierto número: si el dígito es divisible **sólo** por 3, será de Empresa X, si es **sólo** divisible por 2 será de Empresa Y, y en otro caso será de Empresa Z.

Ejemplo:

empresa('contactos.txt')

{'Empresa Z' : {'Phillipe Fromage', 'John Michael Smith'}, 'Empresa X' : {'Jose Miguel Aranda'}, 'Empresa Y' : {'Mario Campos'} }

- b) Implemente la función *ordenar(nombre)* que toma el archivo *contactos.txt*, y genera un archivo distinto por cada país, almacenando el nombre (sin apellido) y el número de los contactos que pertenecen a este país.

Ejemplo:

ordenar('contactos.txt')

Chile.txt

Jose Miguel 9-7852341 Mario 8-8331735

Francia.txt

Phillipe 6-8794321
