

Manejo de estadísticas básicas

Cuando hablamos de estadísticas básicas nos referimos a contar elementos de una lista, cálculos de porcentajes, promedios, determinar el mayor o menor valor de una lista, encontrar el primer valor o el ultimo valor que cumple con una condición etc.

Definiciones Básicas:

Variable contadora: variable que se modifica así misma para llevar el computo de veces que ocurre un suceso, evento o situación, este tipo de variables su incremento es un valor constante:

$$A = A + \text{valor_constante} \text{ o } A += \text{valor_constante}.$$

Variable acumuladora: variable que se modifica así misma para acumular un valor variable:

$$B = B + \text{valor_variable} \text{ o } B += \text{valor_variable}.$$

Variable Bandera: variable que cambia su valor para indicar que algún suceso o evento ha tenido lugar, cambia de un valor a otro como de cero a uno o de falso a cierto.

Porcentaje: número de elementos de una lista que cumplen con una condición entre la cantidad de elementos a los cuales se les comprobó si cumplen o no con la condición por cien.

Promedio: sumatoria de un valor de una lista entre número de elementos sumados.

Para ilustrar esto desarrollaremos el siguiente proyecto:

En el archivo "Valore.txt" tenemos registrada la siguiente información de un grupo de estudiantes:

Nombre, Género (F: femenino, M: masculino), Edad y la Nota definitiva de un curso.

Para cada estudiante:

1. Mostrar sus datos en forma tabular formateada.

Para todos los estudiantes:

2. Porcentaje de estudiantes del género femenino que reprobaron el curso com respecto al total de estudiantes de ese género.
3. Nota promedio de los estudiantes del género masculino cuya definitiva está entre 14 y 20 puntos.
4. Cuantos estudiantes fueron procesados antes de encontrar el primer estudiante que obtuvo una calificación de 10 puntos.
5. Nombre del primer estudiante que obtuvo una calificación de 15 puntos.
6. Nombre del estudiante que obtuvo la mayor nota en el curso en caso de haber más de uno en la misma situación reporte el primero.

Para cada estudiante se nos pide simplemente mostrar el contenido del archivo en forma de tabla con los datos perfectamente alineados en columnas. De tal manera que la salida debería quedar así:

Nombre	Genero	Edad	Calificacion Final
Juan	M	18	14
Maria	F	16	16
Ana	F	15	9
Roberto	M	17	10
Jose	M	14	17
Aura	F	16	15
Pedro	M	19	10
Jesus	M	14	17
Frank	M	19	7
Antonia	F	20	5
Laura	F	19	15
Luis	M	17	17
Jerico	M	22	14
Mirian	F	17	4
Mercedes	F	22	17
Williams	M	18	8
Gertru	F	16	14
Victor	M	15	15
Paco	M	18	16

Aquí no se requiere realizar ningún tipo de cálculo para cada estudiante, si fuera el caso se harían los cálculos y se mostrarían los resultados dentro del ciclo de lectura del archivo.

Ahora cuando tengamos que hacer cálculos **para todos los estudiantes**, los datos requeridos para hacer los cálculos los recabamos dentro de ciclo y el cálculo se hace una vez terminado el ciclo de lectura, por ejemplo, se nos pide calcular la nota promedio del curso, que se requiere para calcular este promedio sumar (acumular) todas las notas de los estudiantes y contar los estudiantes cuyas notas fueron acumuladas, esto se hace dentro del ciclo de lectura, terminado el ciclo es cuando se procede a calcular el promedio dividiendo el acumulado entre el total de alumnos para luego proceder a mostrar este resultado.

Cuales son los requerimientos de memoria para este programa (Variables):

Para la lectura:

nom: para el nombre.

género: para el género del estudiante aquí se colocara F e el caso femenino y M en el caso masculino.

edad: para almacenar la edad del estudiante.

nota: para almacenar la nota final del estudiante.

Los cálculos a realizar para todos los estudiantes:

Punto 2:

- **contF**: cuenta la cantidad de estudiantes del género femenino que hay en el curso.
- **contFR**: cuenta la cantidad de estudiantes del género femenino que reprobaron.
- **porc**: almacena el cálculo del porcentaje.

Punto 3:

- **acumNota**: acumula la nota de los estudiantes del género masculino.
- **contM**: cuenta la cantidad de estudiantes del género masculino cuyo promedio calculamos.
- **prom**: almacena el cálculo del promedio.

Punto 4:

- **flag1**: bandera para controlar si lo encontramos o no.

Punto 5:

- **flag2**: bandera para controlar si lo encontramos o no.
- **primero**: para guardar el nombre del primer estudiante de la lista que obtuvo 15 puntos.

Punto 6:

- **mnom**: guarda el nombre del estudiante que obtuvo la mayor nota.
- **mnota**: guarda la mayor nota.
- **flag3**: bandera para controlar si el primer elemento de la lista fue procesado, cuando esto ocurra el valor de la bandera cambia de cero a 1, esto se hace con la finalidad de que el valor procesado se asume que es el mayor valor de la lista para luego compararlo con los demás valores de lista y determinar cuál es el mayor.

Código:

```
1  # Inicializar variables.
2  nom=mnom=primero=segundo=ultimo=genero=""
3  edad=nota=mnota=0
4  flag1=flag2=flag3=0
5  prom=porc=0.0
6  contF=contFR=0
7  acumNota=contM=0
8  cont10=0
9  # Cabecera de las columnas de la lista de alumnos.
10 print("Nombre    Genero    Edad    Calificacion Final")
11 # Apertura del archivo Valores.txt
12 archivo=open("Valores.txt")
13 # Ciclo de lectura del archivo y procesar datos.
14 for registro in archivo:
15     campos=registro.split(",")
16     nom=campos[0]
17     genero=campos[1]
18     edad=int(campos[2])
19     nota=int(campos[3])
```

```

20 # Mostramos los valores por pantalla formateados.
21 print("{0:12} {1:2} {2:3d} {3:3d}".format(nom,genero,edad,nota))
22 if (genero=="F"): # Determinamos los valores para el calculo del porcentaje
23     contF+=1 # Contamos las alumnas del curso.
24     if (nota<9.5):
25         contFR+=1 # Contamos las alumnas reprobadas del curso
26 if (genero=="M" and (nota>=14 and nota<=20)): # Determinamos los valores para el calculo del promedio
27     contM+=1 # Contamos los alumnos
28     acumNota+=nota # Acumulamos las notas de los alumnos
29 if (flag1==0 and nota==10): # Contamos alumnos para encontrar 1er 10
30     flag1=1 # El cambio del flag1 de 0 a 1 indica que encuentre el alumno
31 elif (flag1==0):
32     cont10+=1 # Mientras flag1 = cero cuento los alumnos
33 if (flag2==0 and nota==15): # Determinamos el 1er alumno con 15
34     primero=nom
35     flag2=1 # Cuando se encuentra cambia flag2 de cero a 1.
36 if (flag3==0):
37     mnom=nom
38     mnota=nota
39     flag3=1
40 elif (mnota<nota):
41     mnom = nom
42     mnota = nota
43 # Calculos finales salida de resultados.
44 print("\n*****Salida de Resultados*****")
45 # calculo porcentaje.
46 porc=contFR/contF*100
47 print("Porcentaje de alumnas reprobadas: {0:5.2f}%".format(porc))
48 # Calculamos promedio
49 prom=acumNota/contM
50 print("Nota promedio: {0:5.2f} puntos".format(prom))
51 print("Para encontrar el primer estudiante con 10 pts contamos: ",cont10," estudiantes")
52 print("El primer estudiante con 15 en la lista fue: ",primero)
53 print("Estudiante con la mayor calificacion: ",mnom)

```

Análisis código:

- Líneas 01 a la 08: Inicializamos las variables de acuerdo a los requerimientos de memoria previamente indicados.
- Línea 10: Mostramos los encabezados de las columnas de los datos a mostrar.
- Línea 12: Abrimos el archivo "Valores.txt".
- Línea 14: Encabezado del ciclo **for**, que nos permitirá ejecutar el ciclo de lectura del archivo.
- Línea 15: Entramos al cuerpo del ciclo y fragmentamos el registro leído en cada uno de sus campos, almacenándolos en una lista.
- Líneas 16 a la 19: Asignamos los valores de lista a las variables de entrada correspondiente.
- Línea 21: Mostramos los datos de entrada en forma tabular formateada.
- Líneas 22 a la 25: Verificamos si el estudiante es del género femenino (22), de no ser así ignoramos, de ser así, contamos la alumna (23) y se verifica si la nota de la alumna es menor de 9.5 (24), de ser así, la contamos (25), de no ser así la ignoramos.
- Líneas 26 a la 28: Verificamos si el estudiante es del género masculino y si nota está entre 14 y 20 (26), de no ser así lo ignoramos, de ser así contamos el alumno (27) y acumulamos su nota (28).
- Líneas 29 a la 32: Verificamos si la bandera es igual a cero y si la nota es igual a 10 (29), de ser así esto nos indica que hemos encontrado el estudiante con la nota de 10 puntos procedemos a cambiar el valor de la bandera de cero a 1 (30), indicando con esto que el proceso de búsqueda se detiene, de no ser así verificamos si la bandera es cero (31) de ser así contamos el alumno procesado, de no ser así no se cuentan mas alumnos.
- Líneas 33 a la 35: Verificamos si la bandera es igual a cero y la nota es igual a 15 (33), de no ser así ignoramos, de ser así esto indica que hemos encontrado el primer alumno que obtuvo 15 puntos en la lista, asignamos su nombre a la variable **primero** (34) y cambiamos el valor

de la bandera a 1 (35), con esto evitamos que si hay más alumnos con 15 estos no será procesados.

Líneas 36 a la 42: Verificamos si la bandera es igual a cero (36), de ser así esto indica que estamos procesando el primer elemento de lista (36), se asume en principio que este estudiante tiene la mayor nota del curso procediendo asignar su nombre a **mnom** (37) y su nota a **mnota** (38) que guardan el nombre y la nota del estudiante que obtuvo la mayor calificación respectivamente, y cambiamos el valor de la bandera de cero a 1 (39), indicando con esto que el primer valor de la lista fue leído, de no ser así comparamos la mayor nota encontrada (hasta ese momento) con la nota del estudiante que estamos procesando en ese momento (40), de ser así esto indica que hemos encontrado un estudiante cuya nota es mayor a la que veníamos considerando la mayor nota del curso, procediendo a cambiar los valores de **mnom** (41) y **mnota** (42), de o ser así ignoramos manteniendo los valores de **mnom** y **mnota**.

Línea 44: Cuando el programa ejecuta esta línea indica que el ciclo de lectura a terminado y se procesado todos los estudiantes, procediendo a mostrar por pantalla la salida de los resultados finales.

Línea 46: Calculamos el porcentaje usando los valores **contFR** y **contF** obtenidos en el cuerpo del ciclo.

Línea 47: Mostramos el porcentaje formateado a dos decimales.

Línea 49: Calculamos el promedio usando los valores **acumNota** y **contM** obtenidos el cuerpo del ciclo.

Línea 50: Mostramos el promedio formateado a dos decimales.

Línea 51: Mostramos la cantidad de estudiantes que hubo que procesar para encontrar el primer estudiante que obtuvo 10 puntos.

Línea 52: Mostramos el nombre del primer estudiante que obtuvo 15 puntos.

Línea 53: Mostramos el nombre de estudiante que obtuvo la mayor calificación del curso. Finalizando el programa.

Salida completa del programa:

Nombre	Genero	Edad	Calificacion Final
Juan	M	18	14
Maria	F	16	16
Ana	F	15	9
Roberto	M	17	10
Jose	M	14	17
Aura	F	16	15
Pedro	M	19	10
Jesus	M	14	17
Frank	M	19	7
Antonia	F	20	5
Laura	F	19	15
Luis	M	17	17
Jerico	M	22	14
Mirian	F	17	4
Mercedes	F	22	17
Williams	M	18	8
Gertru	F	16	14
Victor	M	15	15
Paco	M	18	16

*****Salida de Resultados*****
Porcentaje de alumnas reprobadas: 37.50%
Nota promedio: 15.71 puntos
Para encontrar el primer estudiante con 10 pts contamos: 3 estudiantes
El primer estudiante con 15 en la lista fue: Aura
Estudiante con la mayor calificacion: Jose

Variantes de este programa:

Determinar el valor mínimo, este caso solo tenemos cambiar en la línea 40 el operador relacional mayor que (>) por el operador menor que (<) y listo corra el programa de nuevo vera este resultado:

```
Estudiante con la menor calificacion: Mirian
```

Determinar el estudiante con la mayor calificación, en caso de haber más de un estudiante en la misma situación reporte el primero y cuente los demás las modificaciones en el código son las siguientes:

```
36         if (flag3==0):
37             mnom=nom
38             mnota=nota
39             flag3=1
40         elif (mnota<nota):
41             mnom = nom
42             mnota = nota
43             contOtros=0
44         elif (mnota==nota):
45             contOtros+=1
```

Solo hemos agregado tres líneas con respecto al código original, dos de ellas involucran el uso de una variable contadora, **contOtros**, que cuenta los estudiantes que tienen la misma mayor nota del curso y en la línea 44 agregamos un **elif** en el cual verificamos si la mayor nota del curso se repite en otros estudiantes, en caso de que el **elif** de la línea 40 no se cumpla. En caso de que la condición de la línea 44 se cumpla se incrementa en 1 el valor del contador, de o ser así no se producirá ningún incremento. Es de hacer notar que cuando se encumbra un estudiante cuya nota es mayor que la nota que se consideraba la mayor nota del curso hasta ese momento se debe inicializar la variable contadora en cero la razón tenemos otro estudiante tenemos otra cuenta.

La salida de este resultado también se altera veamos:

```
56     print("Estudiante con la mayor calificacion: ",mnom)
57     if (contOtros!=0):
58         print("Y aparte de el hay: ",contOtros," estudiantes mas con la misma nota")
```

Aparte de mostrar el nombre del estudiante con la mayor calificación agregamos un **if**, línea 57, que verifica si la variable contadora es diferente de cero, de ser así indicará que aparte del estudiante encontrado se encontraron otros mas con la misma mayor nota y se muestra (línea 58) cuantos mas, de no ser así, se ignora, indicando esto que solo se encontró un estudiante con esa nota. La salida sería esta:

```
Estudiante con la mayor calificacion: Jose
Y aparte de el hay: 3  estudiantes mas con la misma nota
```

Ejercicio: modificar el programa de tal manera que muestre el nombre del estudiante que obtuvo la mayor calificación, en caso de que esta situación se repita reporte el primero y el último encontrado.

Para el caso de esta lista la respuesta debería ser:

```
Estudiante con la mayor calificacion: Jose
el ultimo de la lista con la misma nota es: Mercedes
```

Problema: encuentre en la lista los estudiantes que ocuparon el primer y segundo lugar en cuanto a sus notas, reporte los primeros encontrados.