Capítulo 3: Describir datos con estadísticas

### **Encontrar la media**

La *media* es una forma común e intuitiva de resumir un conjunto de números. Es lo que podríamos llamar simplemente la "media" en el uso cotidiano, aunque como veremos, también hay otros tipos de medias. Tomemos un conjunto de números de muestra y calculemos la media.

Supongamos que hay una escuela benéfica que ha estado recibiendo donativos durante un periodo de tiempo que abarca los últimos 12 días (lo denominaremos periodo A). En ese tiempo, los 12 números siguientes representan el importe total en dólares de los donativos recibidos cada día: 100, 60, 70, 900, 100, 200, 500, 500, 503, 600, 1000 y 1200. Podemos calcular la media sumando estos totales y dividiendo la suma por el número de días. En este caso, la suma de los números es 5733. Si dividimos este número por 12 (el número de días), obtenemos 477,75, que es la donación *media* por día. Este número nos da una idea general de cuánto dinero se donó en un día determinado.

Dentro de un momento, escribiremos un programa que calcule e imprima la media de una colección de números. Como acabamos de ver, para calcular la media, tendremos que tomar la suma de la lista de números y dividirla por el número de elementos de la lista. Veamos dos funciones de Python que facilitan estas dos operaciones: sum() y len().

Cuando utilizas la función sum() en una lista de números, suma todos los números de la lista y devuelve el resultado:

>>> shortlist = [1, 2, 3]  
>>> sum(shortlist)  
6

Podemos utilizar la función len() para obtener la longitud de una lista:

>>> len(shortlist)  
3

Cuando utilizamos la función len() en la lista, nos devuelve 3 porque hay tres elementos en shortlist. Ahora estamos preparados para escribir un programa que calcule la media de la lista de donaciones.

'''  
Calculating the mean  
'''  
  
def calculate\_mean(numbers):  
➊ s = sum(numbers)  
➋ N = len(numbers)  
# Calculate the mean  
➌ mean = s/N  
  
return mean  
  
if \_\_name\_\_ == '\_\_main\_\_':  
➍ donations = [100, 60, 70, 900, 100, 200, 500, 500, 503, 600, 1000, 1200]  
➎ mean = calculate\_mean(donations)  
N = len(donations)  
➏ print('Mean donation over the last {0} days is {1}'.format(N, mean))

Primero, definimos una función, calculate\_mean(), que acepta el argumento numbers, que es una lista de números. En ➊, utilizamos la función sum() para sumar los números de la lista y creamos una etiqueta, s, para referirnos al total. Del mismo modo, en ➋, utilizamos la función len() para obtener la longitud de la lista y creamos una etiqueta, N, para referirnos a ella. Después, como puedes ver en ➌, calculamos la media simplemente dividiendo la suma (s) por el número de miembros (N). En ➍, creamos una lista, donations, con los valores de las donaciones enumeradas anteriormente. A continuación, llamamos a la función calculate\_mean(), pasando esta lista como argumento en ➎. Por último, imprimimos la media calculada en ➏.

Cuando ejecutes el programa, deberías ver lo siguiente:

Mean donation over the last 12 days is 477.75

La función calculate\_mean() calculará la suma y la longitud de *cualquier* lista, así que también podemos reutilizarla para calcular la media de otros conjuntos de números.

Calculamos que la media de donaciones diarias es de 477,75. Cabe señalar que las donaciones de los primeros días fueron mucho menores que la donación media que calculamos y que las donaciones de los últimos días fueron mucho mayores. La media nos da una forma de resumir los datos, pero no nos da una imagen completa. Sin embargo, hay otras medidas estadísticas que pueden decirnos más sobre los datos cuando se comparan con la media.

[anterior](ch03_1.html)[Subtema 2 de 10: (Ver todo)](ch03.html)[siguiente](ch03_3.html)