### Análisis de las cabeceras del archivo CSV

El módulo csv de la biblioteca estándar de Python analiza las líneas de un archivo CSV y nos permite extraer rápidamente los valores que nos interesan. Empecemos examinando la primera línea del archivo, que contiene una serie de cabeceras para los datos. Estas cabeceras nos indican qué tipo de información contienen los datos:

**sitka\_highs.py**

from pathlib import Path  
import csv  
  
❶ path = Path('weather\_data/sitka\_weather\_07-2021\_simple.csv')  
lines = path.read\_text().splitlines()  
  
❷ reader = csv.reader(lines)  
❸ header\_row = next(reader)  
print(header\_row)

Primero importamos Path y el módulo csv. A continuación, construimos un objeto Path que busca en la carpeta *weather\_data*, y apunta al archivo de datos meteorológicos concreto con el que queremos trabajar ❶. Leemos el archivo y encadenamos el método splitlines() para obtener una lista de todas las líneas del archivo, que asignamos a lines.

A continuación, construimos un objeto reader ❷. Éste es un objeto que puede utilizarse para analizar cada línea del archivo. Para crear un objeto lector, llama a la función csv.reader() y pásale la lista de líneas del archivo CSV.

Cuando se le da un objeto reader, la función next() devuelve la siguiente línea del archivo, empezando por el principio del archivo. Aquí llamamos a next() una sola vez, por lo que obtenemos la primera línea del archivo, que contiene las cabeceras del archivo ❸. Asignamos los datos devueltos a header\_row. Como puedes ver, header\_row contiene cabeceras significativas, relacionadas con el tiempo, que nos indican qué información contiene cada línea de datos:

['STATION', 'NAME', 'DATE', 'TAVG', 'TMAX', 'TMIN']

El objeto reader procesa la primera línea de valores separados por comas del archivo y almacena cada valor como un elemento de una lista. La cabecera STATION representa el código de la estación meteorológica que registró estos datos. La posición de esta cabecera nos indica que el primer valor de cada línea será el código de la estación meteorológica. La cabecera NAME indica que el segundo valor de cada línea es el nombre de la estación meteorológica que realizó la grabación. El resto de las cabeceras especifican qué tipo de información se registró en cada lectura. Los datos que más nos interesan por ahora son la fecha (DATE), la temperatura alta (TMAX) y la temperatura baja (TMIN). Se trata de un conjunto de datos sencillo que sólo contiene datos relacionados con la temperatura. Cuando descargues tus propios datos meteorológicos, puedes optar por incluir otra serie de mediciones relacionadas con la velocidad y dirección del viento, y datos sobre precipitaciones.

[anterior](c16_3.html)[Subtema 4 de 29: (Ver todo)](c16.html)[siguiente](c16_5.html)