

08. Archivos Pandas Manipulacion Limpieza Visualizacion

August 16, 2024

1 Manejo de archivos

El manejo de archivos es una parte importante de cualquier aplicación web.

Python tiene varias funciones para crear, leer, actualizar y eliminar archivos.

1.1 Función open

La función clave para trabajar con archivos en Python es `open()`. La función `open()` toma dos parámetros: *nombre_archivo* y *modo*.

Hay cuatro métodos (modos) diferentes para abrir un archivo:

- “r”- Leer - Valor predeterminado. Abre un archivo para leerlo. Se produce un error si el archivo no existe.
- “a”- Anexar - Abre un archivo para anexarlo, crea el archivo si no existe.
- “w”- Escribir - Abre un archivo para escribir, crea el archivo si no existe.
- “x”- Crear - Crea el archivo especificado, devuelve un error si el archivo existe.

Además, puede especificar si el archivo debe manejarse en modo binario o de texto.

- “t”- Texto - Valor predeterminado. Modo texto.
- “b”- Binario - Modo binario (por ejemplo, imágenes).

1.2 Sintaxis

La función `open()` abre un archivo y lo devuelve como un objeto de archivo.

`open(file, mode)`

1.2.1 Valores predeterminados

Parámetro	Descripción
<i>file</i>	La ubicación y nombre del archivo
<i>mode</i>	Una cadena que define el modo como se abre el archivo "r" - Lectura - Valor por defecto. Abre un archivo para lectura, envía error si el archivo no existe. "a" - Añadir - Abre un archivo para añadir, crea el archivo si no existe.

Parámetro	Descripción
	<p>"w" - Escritura - Abre un archivo para escritura, crea el archivo si no existe.</p> <p>"x" - Creación - Crea un archivo, devuelve un error si el archivo ya existe.</p> <p>Adicionalmente se puede especificar si el archivo será manejado como texto o binario</p> <p>"t" - Texto - Modo texto, valor por defecto.</p> <p>"b" - Binario - Modo binario (i.e. imágenes)</p>

Para abrir un archivo para su lectura es suficiente especificar el nombre del archivo:

```
f = open("demofile.txt")
```

Este código es equivalente al siguiente:

```
f = open("demofile.txt", "rt")
```

Dado que 'r' para lectura y 't' para texto son los valores por defecto, no es necesario incluirlos.

1.3 Abrir un archivo localmente

Supongamos que tenemos el archivo `demofile.txt`, ubicado en la misma carpeta que el script de Python:

Archivo `demofile.txt`

```
Hello! Welcome to demofile.txt
This file is for testing purposes.
Good Luck!
```

Para abrir el archivo, se utiliza la función incorporada `open()`. Esta función devuelve un objeto tipo archivo, que tiene un método `read()` para leer el contenido del archivo:

```
[ ]: f = open("demofile.txt", 'r')
      print(f.read())
```

```
Hello! Welcome to demofile.txt
This file is for testing purposes.
Good Luck!
```

1.4 Abrir un archivo en Google Drive

Supongamos que tenemos el archivo `demofile.txt`, ubicado en Google Drive:

Archivo `demofile.txt`

```
Hello! Welcome to demofile.txt
This file is for testing purposes.
Good Luck!
```

Además un notebook de Python se encuentra en la misma carpeta. Para abrir un archivo es necesario utilizar la función `google.colab` dentro del paquete `drive`. Una vez que se haya incluido el paquete, ahora es necesario *montar* la unidad de almacenamiento de Google Drive en una ubicación dentro del sistema de archivos virtual: `/drive`. Cabe mencionar que esta dirección apuntará a la raíz de la unidad de almacenamiento en Google Drive.

```
from google.colab import drive
drive.mount("/drive")
```

Al ejecutar la función `drive.mount`, Google Drive preguntará por autorización para tener acceso a su unidad.

Posterior a esto, sólo es necesario utilizar la función `open` tal y como lo hicimos anteriormente.

```
f = open("/drive/My Drive/demofile.txt", 'r')
print(f.read())
```

Observe que se precede al nombre del archivo la ubicación `/drive/My Drive/`, que es donde se montó virtualmente la unidad de almacenamiento de Google Drive.

2 Pandas

3 Manipulación y Limpieza de Datos

4 Visualización de datos con Matplotlib y Seaborn

5 Referencias

- [Manejo de archivos en Python](#)
- Lutz M., Learning Python, O'Reilly. 2009