

Python

# Archivos XLS



Rogelio Ferreira Escutia

Profesor / Investigador  
Tecnológico Nacional de México  
Campus Morelia



# Tipos de Archivos

# Tipos de Archivos

- **Básicamente hay 2 tipos de archivos:**

## Text File

- **Plain Text:** .txt, .csv
- **Source Code:** .py, .html, .css, .js
- **Data:** .json, .xml

## Binary File

- **Executable:** .exe, .dmg, .bin
- **Images:** .jpg, .png, .gif, .tiff, .ico
- **Video:** .mp4, .m4v, .mp4, .mov
- **Audio:** .aif, .mp3, .mpa, wav
- **Compressed:** .zip, .deb, .tar.gz
- **Font:** .woff, .otf, .ttf
- **Document:** .pdf, .docx, .xlsx

**XLS**

# XLS

- **Es el formato que usa para almacenar la información de las hojas de cálculo de Microsoft Excel.**
- **La primera versión salió para computadoras Mac en 1985 y actualmente hay versiones para varias plataformas y se ha convertido en uno de los formatos mas utilizados para el almacenamiento de información numérica en computadoras de escritorio.**

# XLS

- Ejemplo (hoja\_pruebas.xls):

	A	B	C	D	E	F
1	Examen 1	Examen 2	Examen 3	Examen 4	Examen 5	Promedio
2	1	2	3	4	5	3
3	6	7	8	9	10	8
4	11	12	13	14	15	13
5	16	17	18	19	20	18
6	21	22	23	24	25	23
7						

# XLS (1)

- **Ejemplo cargar “hoja\_pruebas.xls” e imprimirla en pantalla:**
- **Se requiere instalar “xlrd” (biblioteca para usar XLS).**
- **Se requiere instalar “Pandas”, el cual servirá para mejorar el procesamiento y cargaremos toda la hoja en Pandas.**
- **En la terminal:**
  - > **pip3 install pandas**
  - > **pip3 install xlrd**

# XLS (2)

- **Cargar bibliotecas, leer archivo “xls” y pasarlo a Pandas:**

```
# Cargar la biblioteca de "pandas"
import pandas as pd

# cargar la hoja de excel en un dataframe y seleccionar la Hoja1
df = pd.read_excel("hoja_pruebas.xls", "Hoja1")

print("\nDataFrame:\n", df)
```

# XLS (3)

- **Obtener el número de filas, de columnas y extraer elementos:**

```
# Obtener las filas
filas = df.shape[0]
print("\nFilas: " + str(filas))

# Obtener columnas
columnas = df.shape[1]
print("Columnas: " + str(columnas))

print("Extraer un elemento:\n")
print(df.loc[0][0])
```

# XLS (4)

- **Salida:**

```
DataFrame:  
   Examen 1  Examen 2  Examen 3  Examen 4  Examen 5  Promedio  
0         1         2         3         4         5         3  
1         6         7         8         9        10         8  
2        11        12        13        14        15        13  
3        16        17        18        19        20        18  
4        21        22        23        24        25        23
```

Filas: 5

Columnas: 6

Extraer un elemento:

1



## Rogelio Ferreira Escutia

Profesor / Investigador  
Tecnológico Nacional de México  
Campus Morelia



[rogelio.fe@morelia.tecnm.mx](mailto:rogelio.fe@morelia.tecnm.mx)



[rogeplus@gmail.com](mailto:rogeplus@gmail.com)



[xumarhu.net](http://xumarhu.net)



[@rogeplus](https://twitter.com/rogeplus)



[https://www.youtube.com/  
channel/UC0on88n3LwTKxJb8T09sGjg](https://www.youtube.com/channel/UC0on88n3LwTKxJb8T09sGjg)



[rogelioferreiraescutia](https://www.linkedin.com/in/rogelioferreiraescutia)

