Archivos

- son estructura de datos persistentes (o sea que perdura más allá de la vida de quien las genera)
- son estructuras de datos empleados para compartir información entre usuarios, procesos, equipos
- debemos recordar que estamos haciendo programas en una arquitectura Von Neumann. Por lo tanto, para que los datos puedan ser procesados primero se deben alojar en la MI y para sacar los datos y los resultados también deben pasar por dicha MI
- generalmente los archivos tienen grandes cantidades de datos, por esto copiar un archivo completo en MI
 puede ser una tarea costosa y/o imposible. Una solución será copiar los datos por partes para que a medida
 que necesitemos información la vamos subiendo a la MI y a medida que la modifiquemos resguardaremos
 una copia en el archivo
- el SO reserva una zona de memoria y establece un canal de comunicación entre el archivo y el programa.
 Estos son conocidos como Buffers (mini espacios de intercambio entre un programa y un archivo). El SO implementa un sistema de tráfico de intercambio, lectura y escritura en bloque (si un programa necesita un dato específico, trae todo el bloque de información donde se encuentra ese dato, debido a que si trae solo el dato específico resulta un costo operativo alto. También es probable que se necesite los datos cercanos a este)
- abrir un archivo es decirle al SO que defina un buffer. Este va a tener el nombre del archivo, dónde se ubica en las estructura de directorios, qué haremos con él, etc
- un mismo archivo puede abrir simultáneamente los archivos necesitados. Para diferenciarlos se le asigna un alias (nombre interno) a cada archivo para referenciarlos dentro del programa

Primitivas Generales

 abrimos un archivo, en Python, con la función open(). Ésta solicita al SO que establezca el vínculo del trabajo y nos devuelva la dirección del inicio del buffer del archivo. Guardamos la dirección en una variable (nombre interno del archivo desde ese momento, dicho nombre interno es una etiqueta que apunta al inicio del buffer. Una vez abierto solo referenciamos al archivo por su nombre interno o alias)

SINTAXIS

alias = open(nombre_completo_archivo, modo)

EJEMPLO

archivo = open("texto.txt")

- nombre y extensión con el que fue guardado
- localización completa o relativa en los directorios (path)
- el modo es por default de lectura (o sea 'r') pero puede ser también w
- cuando terminamos de usar el archivo, ES IMPORTANTE realizar la operación de cierre (operación que libera el espacio de buffer y completa el proceso de grabado si hay algo pendiente), usando el método close().

El SO trae la información, por bloques sin preguntarnos y tampoco se molesta en salir a grabar un dato cuando lo solicitamos. Junta varios encargos de grabado antes de realizar la tarea. El cierre del archivo obliga a grabar todo lo que quedó pendiente

Modos de Apertura

Modo	Lee	Escribe	Existencia de Archivo	Posición Inicial de Lectura	Posición Inicial de Escritura
r	~	-	Si el archivo no existe, da error.	Inicio	-
w	-		Si el archivo no existe, lo crea. Si existe, le borra todo el contenido.	-	Inicio
а	-		Si el archivo no existe, lo crea.	-	Final
r+	▼	☑	Si el archivo no existe, da error.	Inicio	Depende
w+	3		Si el archivo no existe, lo crea. Si existe, le borra todo el contenido.	Inicio	Depende

1. Modo r

- r = read = leer
- se asigna por default, aunque también se puede aclarar en el código

- tres métodos de lectura
 - o **read** → lee todo el archivo junto y lo devuelve en un string
 - readlines → lee cada línea y las devuelve como una lista. Cada línea tiene \n que indica el cambio de línea
 - readline → lee línea por línea cada vez que se llama la función. Cada línea contiene el caracter especial \n pero no de forma explícita, sino implícita (o sea se imprime con un espacio de nueva línea pero no vemos el carácter)
- modo r+

```
archivo = open("archivo.txt", "r+")
```

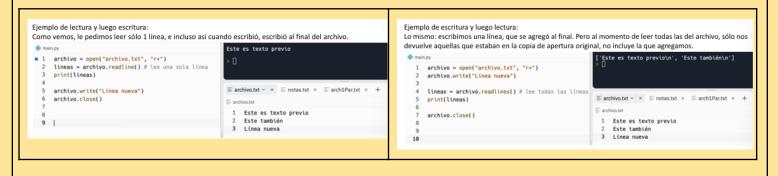
• tenemos tres comportamientos:





Comportamiento 3: Lectura y Escritura

- cuando leemos y escribimos usando este modo, va a pasar que independientemente de qué hagamos antes, todo lo que escribamos en el archivo siempre va a ir al final. Esto ocurre porque una vez que leemos el archivo, el puntero de posición con el que escribimos pasa a estar al final
- si le pedimos leer el archivo, siempre nos va a leer el contenido que teníamos antes de abrirlo, debido a
 que los write se guardan en el momento en el que se cierra el archivo



2. Modo w

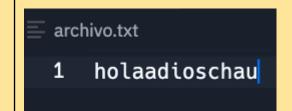
• w = write = escribir

```
archivo = open("archivo.txt", "w")
```

- dos métodos de lectura
 - write → recibe un string y los escribe en el archivo. Si no le agregamos el caracter especial
 (\n) imprime una palabra detrás de otra



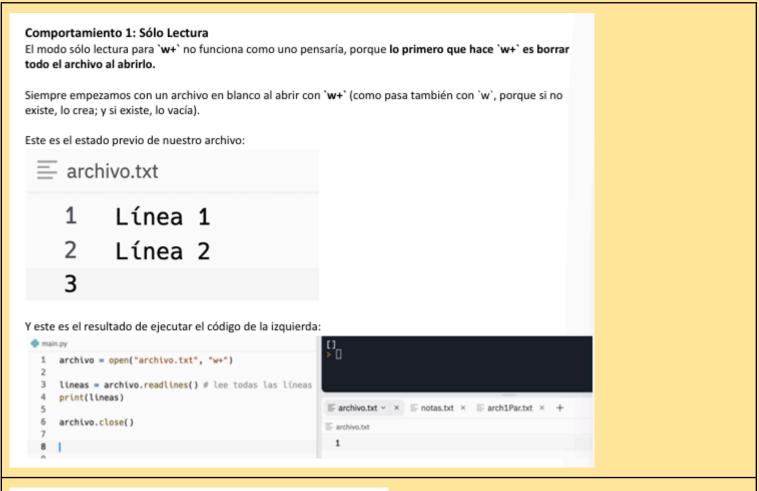
 o writelines → recibe un string o una lista de strings. Si no le agregamos el caracter especial imprime una palabra detrás de otra archivo = open("archivo.txt", "w")
archivo.writelines(["hola", "adios"])
archivo.writelines("chau")
archivo.close()



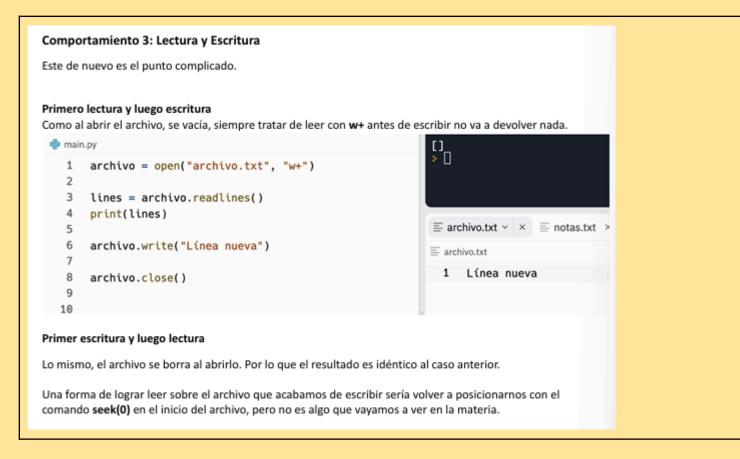
modo w+

```
archivo = open("archivo.txt", "w+")
```

tiene tres posibles comportamientos:







3. <u>Modo a</u>

a = append = agregar

```
archivo = open("archivo.txt", "a")
```

- el método de escritura es igual al de w, aunque se diferencia porque a SIEMPRE AGREGA EL CONTENIDO AL FINAL DEL ARCHIVO
- problema de escritura con append
 - o ocurre cuando nuestro archivo no tiene una línea vacía extra
 - nuestro indicador de posición en el archivo va a agregar el contenido a lado del anterior (o sea sin salto de línea ambos strings se van a juntar por así decirlo). Para esto tenemos dos soluciones: <u>la primera</u> es poner el carácter especial antes o dentro de escribir el string y <u>la segunda</u> es procurar que todos lo que se agrega al archivo tenga un carácter especial (\n) al final, de esta forma vamos a terminar las líneas con un punto y coma

```
archivo = open("archivo.txt", "a")
archivo.writelines(["\nhola", "adios"])
archivo.writelines("chau")
archivo.close()

1  Este es un texto previo
2  Este también
3  holaadioschau
```

4. Otra forma de abrir archivos

usando el with open as (para este modo no es necesario usar close())

```
main.py

1 v with open("archivo.txt", "r") as archivo:
2    lines = archivo.readlines()
3    print(lines)
4
```

el código dentro del bloque with mantiene el archivo abierto

Tipos de Archivo

Archivos Planos – Archivos de Texto (txt)

- archivo plano / archivo de texto plano / archivo de texto simple
- es un tipo de archivo informático que almacena datos estructurados pero sin usar ningún formato de codificación o estructura de datos compleja
- suelen llevar la extensión ".txt" al final (ejemplo: archivo.txt)

```
si ejecutamos el programa tendremos la siguiente
vaca_txt=open('arch1.txt','r+')
                                                                              salida
t = vaca_txt.readline()
print(t)
                                                                                  >_ Console ∨ × □ Output × +
t = input('Ingresá un texto con vaca: ')
while (t.lower()).count('vaca')==0:
                                                                                  La vaca LOCA
t = input('Ingresá un texto con vaca: ')
vaca_txt.write(t+'\n')
                                                                                  Ingresá un texto con vaca: señora vaca, señora vaca
vaca_txt.close()
                                                                                 La vaca LOCA
vaca_txt=open('arch1.txt')
                                                                                  VACAYENDO gente al baileee
todas = vaca_txt.readlines()
                                                                                  siga la vaca
for linea in todas:
    print(linea.strip('\n'))
                                                                                  vacaciones en faMIlia...
vaca_txt.close()
                                                                                  es lila la vaca de MILKA
Nota:
En este ejemplo tenemos en la misma carpeta del programa un archivo: arch1.txt que contiene al
                                                                                Y el archivo arch1.txt contendrá ahora lo siguiente:
momento de correr el programa lo siguiente:
                                                                                VACAYENDO gente al baileee
VACAYENDO gente al baileee
                                                                                 siga la vaca
siga la vaca
                                                                                 vacaciones en faMIlia...
vacaciones en faMIlia..
                                                                                 es lila la vaca de MILKA
                                                                                 señora vaca, señora vaca
es lila la vaca de MILKA
```

qué hicimos en el programa?

Código	Efecto
# pruebaarch1.py	
vaca_txt=open('arch1.txt','r+')	Abre arch1.txt para lectura y escritura, con la
	codificación por defecto del SO y define nombre
	interno o alias vaca_txt
	Lee desde el archivo la primer línea (todos los
t=vaca_txt.readline()	caracteres hasta encontrar un \n y se guarda en t
print(t)	Muestra la línea
t=input('Ingresá un texto con vaca: ')	Pide al usuario un nuevo texto con la palabra vaca
while (t.lower()).count('vaca')==0:	Valida que el texto incluya al menos vez
t=input('Ingresá un texto con vaca: ')	vaca
vaca_txt.write(t+'\n')	Graba en el archivo la línea leída
vaca_txt.close()	Cierra el archivo
vaca_txt=open('arch1.txt')	Reabre el archivo, pero ahora sólo lectura
	Carga la lista todas con las líneas del archivo. Cada
todas=vaca_txt.readlines()	línea es un elemento
for linea in todas:	Muestra el contenido de la lista quitando
	\n para evitar doble interlineado, uno sale
print(linea.strip('\n'))	naturalmente por el print()
vaca_txt.close()	Cierra el archivo

<u>Archivos Planos - Comma Separated Values (csv)</u>

- los datos se separan por comas o puntos y comas (este último lo usa excel)
- como al abrir un archivo se usa la función open(), se puede editar una string y automatizar la construcción del nombre del archivo

```
Integración con bucle de Carga
   completo=open('datosCompletos.csv','w')
b=[]
j in range(1,4):
    dat=open('datos'+str(j)+'.csv')
    a=dat.readlines()
    dat.close()
for i in range(len(a)):
     a[i]=a[i].strip('\n')
   a[i]=a[i].split(';')
   a[i][3]=int(a[i][3])
b.sort(reverse=True, key=lambda b:(b[3],b[2]))
for elemento in b:
   print(elemento)
for elemento in b:
   elemento[3]=str(elemento[3])
   ele=';'.join(elemento)+'\n'
   completo.write(ele)
completo.close()
```

Manejo de Errores

- mientras más tarde se identifica el error, más costos genera el proceso de corrección
- cuando hablamos de sistemas y programas
 - o el ingreso de datos en unos de los puntos con frecuencia de errores
 - o el ingreso de datos es el lugar más económico donde se pueden detectar y salvar errores
- para disminuir la tasa de errores
 - simplificamos el formato de los datos o los adaptamos a sistemas de lectura automática (código de barra, qr, etc)
 - validar los ingresos / lectura (proceso por el que se verifica y rechaza un ingreso de datos, buscando detectar posibles errores)
- siempre hay posibilidades de que el usuario se equivoque al ingresar los datos
- no todos los errores se podrán detectar en el ingreso. Los que si no todos tendrán una ecuación costo-beneficio que justifique su validación. Pero cuando validamos?. Cuando se pueda y cuando convenga.

se recomienda validar el ingreso siempre que se pueda y cuando la ecuación costo-beneficio sea positiva

- para validar un ingreso debemos escribir una o varias condiciones que nos permitan testear esos datos o identificar la posible presencia de errores
- si detectamos un ingreso erróneo, podemos hacernos las siguientes preguntas:
 - o rechazamos el ingreso y paramos la ejecución del programa?
 - continuamos la ejecución ignorando el ingreso?
 - o deja de tratarse el caso que contiene un error?

- se pide la colaboración del usuario y se le solicita un reingreso?
- Lo más lógico sería solicitar un reingreso para validarlo y ver si tiene errores, si los tiene se le vuelve a pedir hasta que no haya más errores.

```
Mostraremos un ejemplo de cómo validar números pares:

# Valida el ingreso de un número par
par = int(input('Ingresá un número par: '))

while par < 0 or par % 2 != 0: #también te serviría la condición par<0 or par%2
    par = int(input('Ingresá un número par: '))
print('Tu número es',par)

Una ejecución ejemplo podría ser:

Ingresá un número par: 11
Ingresá un número par: 8
Tu número es 8
>>>
```

Retenemos al usuario en el proceso de ingreso, la condición para continuar, o sea iterar dentro del bucle, es que se cumpla lo que no debe pasar. O sea no debemos escribir la condición de ingreso válido. Otra opción es crear la condición que queremos cumplir y agregarle el **operador not** adelante para negar al mismo.

 un ejemplo, cuando solicitamos el ingreso de un número entero y lo casteamos en la misma línea, si el usuario se equivoca e ingresa un dato erróneo Python enloquece. Para manejar este error tendríamos que validar que el valor ingresado sea compatible con el número solicitado y recién castear.

```
par=int(input('Ingresá un número par: '))
```

Ahora cuando ingresamos un número negativo lo toma como no válido, este problema también se da con los métodos de string como isdecimal() o isdigit()

```
# Solicita el ingreso de un número entero
numero = input('Ingresá un número entero: ')
while not numero.isnumeric():
   print('El valor ingresado no es válido.')
   numero = input('Ingresá un número entero: ')
print('Tu número entero es ',int(numero))
```

Manejo de Excepciones

- es un evento que ocurre durante la ejecución de programas que interrumpen el flujo normal de la ejecución. O sea es un objeto de Python que representa un error y contiene información del mismo
- Vemos que el error se da en la celda [1] de la línea 1
- tipo de error o nombre de la excepción → NameError
- breve descripción del error → el nombre 'Name' no está definido, Python no sabe a qué nos referimos cuando llamamos a Name

```
In [1]: name

NameError Traceback (most recent call last)
Cell In [1], line 1
----> 1 name

NameError: name 'name' is not defined
```

 Python viene con excepciones incorporadas y tiene la posibilidad de crear excepciones autodefinidas. Estas son las más usadas:

Excepción	Causa del error
AssertionError	Falla assert.
AttributeError	Falla la referencia de atributos.
EOFError	Se genera cuando la función input() alcanza la condición de fin de archivo (EOF).
FloatingPointError	Falla una operación de punto flotante.
GeneratorExit	Se genera cuando se llama al método close() de un generador.
ImportError	Se genera cuando no se encuentra el módulo importado.

н		
	IndexError	Se genera cuando el índice de una secuencia está fuera de rango.
1	KeyError	Se genera cuando una key no se encuentra en un diccionario.
1	KeyboardInterrupt	Se genera cuando el usuario presiona la tecla de interrupción (Ctrl+C o Eliminar)
1	MemoryError	Se genera cuando una operación se queda sin memoria.
1	NameError	Se genera cuando una variable no se encuentra en el ámbito local o global.
	OSError	Se genera cuando la operación del sistema causa un error relacionado con el sistema.
	ReferenceError	Se genera cuando se utiliza un proxy de referencia débil para

try / except

- sentencia estructurada usada para capturar y manejar excepciones
- Python ejecuta el código e intenta hacer una acción con la instrucción try, si tiene éxito el programa prosigue de lo contrario si surge un error, la sentencia lo captura y lo trata en la cláusula except
- try → cláusula que expresa lo que esperemos que falle
- except → cláusula donde planeamos manejar las fallas específicas en el código

```
try:
    # Corre este código, algo puede fallar.
except:
    # Se ejecuta este código cuando ocurre una excepción en el bloque
de try.

Mostraremos cómo resolver el problema del ingreso de un entero utilizando esta facilidad:

numero = None
while type(numero) is not int:
    try:
        numero = input('Ingresá un número entero: ')
        numero = int(numero)
        print('Tu número entero es ', int(numero))
        except ValueError:
        print('El valor ingresado no es válido.')
```

Una ejecución posible sería:
Ingresá un número entero: 1
El valor ingresado no es válido.
Ingresá un número entero: 0.6
El valor ingresado no es válido.
Ingresá un número entero: -7
Tu número entero es -7
>>>

- A tener en cuenta:
 - el manejo de excepciones dentro de un bucle es para darle al usuario los intentos que necesite. Sólo salimos del bucle cuando se puede ejecutar el bloque vinculado a la cláusula try
 - hay situaciones donde hay múltiples excepciones. Si decidimos usarlos debemos consignarlos, de lo contrario capturara todos los errores posibles

```
Caso 1: No se indica ninguna excepción en particular

try:
    # Corre este código, algo puede fallar.
except:
    # Se ejecuta este código cuando ocurre una excepción en el bloque de try.

Caso 2: Se indican distintas excepciones en un mismo bloque except.

try:
    # Corre este código, algo puede fallar.
except Excepción_1, Excepción_2:
    # Se ejecuta este código cuando ocurre cualquiera de las dos excepciones.
```

```
Caso 3: Se indican distintas excepciones en bloques diferentes de except.

try:
    # Corre este código, algo puede fallar.
except Excepción_1:
    # Se ejecuta este código cuando ocurre Excepcion_1.
except Excepción_2:
    # Se ejecuta este código cuando ocurre Excepcion_2.
```

```
▲ iAtención!
También podríamos querer "no hacer nada" frente a una excepción, pero evitando que el programa se
bloquee. Para ello, existe pass.
  # Corre este código, algo puede fallar.
except:
  pass # No hagas nada.
Quizá pass puede parecer poco útil en un primer momento, pero puede volverse tu mejor amigo. Por
ejemplo, mientras estás pensando la estructura del código, y querés ir probando algunas líneas cuando aún
no está terminado, y evitar que el programa se interrumpa.
if numero % 15 == 0:
  pass # Completar.
elif numero % 3 == 0:
  pass # Completar.
elif numero % 5 == 0:
  pass # Completar.
else:
  pass # Completar.
```

try con else

 else es una sentencia usada cuando deseamos ejecutar un bloque de código específico sólo si no ocurre ninguna excepción en el bloque try

```
try:
    # Corre este código, algo puede fallar.
except:
    # Se ejecuta este código cuando ocurre una excepción en el bloque de try.
else:
    # Se ejecuta este código cuando no existe una excepción.
```

- ¿Por qué es una mala idea poner más código en el bloque try?
 - hay posibilidad de encontrarnos con excepciones que no anticipamos
 - reduce la legibilidad. Cuanto más código está en la cláusula try, menos claridad tiene su estructura.
 Lo ideal sería que fuera lo más chico posible

try con finally

usado cuando queremos ejecutar una acción por más que haya ocurrido, o no, un error en un programa

```
try:
    # Corre este código, algo puede fallar.
except:
    # Se ejecuta este código cuando ocurre una excepción en el bloque de try.
else:
    # Se ejecuta este código cuando no existe una excepción.
finally:
    # Siempre se ejecuta y bloquea el código.
```

- siempre se ejecutará
- finaliza el programa
- si se produce una excepción no controlada ejecutará el código antes que ocurra
- si tenemos un return en el bloque try, finally interrumpe ese retorno para ejecutar primero su propio código

cláusula raise

- permite lanzar una excepción si se produce cierta condición
- útil en situaciones donde necesitamos generar una excepción personalizada (ejemplo recibir datos incorrectos o cualquier error de validación

```
x = 10

if x > 5:
    raise Exception(f'No debe exceder 5. El valor es: {x}').

Miremos su ejecución:

Exception: No debe exceder 5. El valor es: 10
>>>
```

 tenemos que usarlo de forma razonable, no es correcto lanzar una excepción para atajar inmediatamente con un except. Lanzar una excepción sirve para elevar el error a una función previa y manejar el error desde ahí

Ejemplo de como usar declaraciones

Crear una función con parámetros x e que calcule el cociente entre ambos. Deberá usar:

- <u>try</u> para probar el cálculo
- except en caso de que ocurra una excepción con los números ingresados
- else para brindar la respuesta en un formato adecuado
- finally para mostrar al usuario que este bloque siempre se ejecuta.

```
realizamos un cálculo de división e identificamos
 def calcular_division(x, y):
                                                                        cualquier excepción que surja cuando el usuario
       try:
                                                                        proporciona valores no válidos (ejemplo de este caso
             cociente = x / y
                                                                        sería el denominador cero donde obtendremos
       except ZeroDivisionError:
                                                                        ZeroDivisionError, esta excepción que debe ser
             pass
                                                                        considerada con la cláusula except)
       else:
             pass
       finally:
             pass
                                                                        dentro la cláusula except escribimos el mensaje para
def calcular_division(x, y):
                                                                        informarle al usuario que los datos ingresados son
        cociente = x / y
                                                                        válidos. Devolvemos el cociente en la cláusula else y
    except ZeroDivisionError:
        print("No se puede dividir por cero.")
                                                                        agregamos el mensaje indicado en la consigna
        print("El cociente es: ", cociente)
     finally:
        print("El bloque finally siempre se ejecuta")
Caso 1: x = 15, y = 5
                                                                         No se puede dividir por cero.
                                                                         El bloque finally siempre se ejecuta
El cociente es: 5.0
El bloque finally siempre se ejecuta
                                                                         En este caso, Python intenta hacer la división 15 / 0 pero aparece el mensaje de error ZeroDivisionError, que es
En este caso, Python intenta hacer la división 15 / 5, y como es posible, ejecuta el bloque else. El bloque finally,
                                                                         capturado por la cláusula except que le indica cómo tratarlo: en este caso, mostrándole al usuario el mensaje
se ejecuta siempre.
                                                                         "No se puede dividir por cero". El bloque finally, se ejecuta siempre.
Caso 2: x = 15, y = 0
```

Aclaraciones

• independientemente cómo se ejecute el bloque de try-except y sus variaciones, al finalizar, el programa sigue con su flujo de control (siempre y cuando no explote)

```
def calcular_division(x, y):
    print(f"\nDividiendo {x} sobre {y}")
    try:
        cociente = x / y
    except ZeroDivisionError:
        print("No se puede dividir por cero.")
    else:
        print("El cociente es: ", cociente)
    finally:
        print("El bloque finally siempre se ejecuta")

print("Este texto siempre se va a imprimir después del bloque try-except")

calcular_division(15, 0)
calcular_division(15, 5)
```

```
Dividiendo 15 sobre 0
No se puede dividir por cero.
El bloque finally siempre se ejecuta
Este texto siempre se va a imprimir después del bloque try-except

Dividiendo 15 sobre 5
El cociente es: 3.0
El bloque finally siempre se ejecuta
Este texto siempre se va a imprimir después del bloque try-except
```

El manejo de errores con try-except es una forma donde nuestro programa puede ser más robusto y resistente a los errores. Al mismo tiempo, no todas las validaciones requieren de uno (hay validaciones que se pueden hacer con las herramientas de control que conocemos, if, elif, else)

Pandas

- es una biblioteca de Python usada para el análisis y la manipulación de los datos en bruto
- puede ayudar al usuario a limpiar, transformar y analizar datos rápido y eficientemente
- su estructura de datos principal es el DataFrame (df). Una tabla bidimensional compuesta por filas y
 columnas, semejantes a una hoja de cálculo de Excel.Usados para almacenar y manipular grandes
 cantidades de datos. Proporciona funcionalidades (selección y filtrado de datos, agregación de datos y
 relación de operaciones matemáticas y estadísticas)

<u>Pasos</u>

import pandas as pd	importamos pandas con su alias pd
pdversion Resultado: {"type":"string"}	vemos la versión
<pre>data = {'animal': ['cat', 'cat', 'snake', 'dog', 'dog', 'cat', 'snake', 'cat', 'dog', 'dog'],</pre>	tenemos el diccionario data y la lista labels

```
creamos un df a partir del diccionario y
df = pd.DataFrame(data, index=labels)
df # Esto lo ponemos para mostrarlo
                                                                           usando la lista como índice
Resultado:
  animal age visits priority
       2.5
    cat
                      yes
        3.0
                3
b
    cat
  snake
        0.5
                      nο
    dog
       NaN
                      yes
    dog
    cat
       2.0
                      no
g
  snake
        4.5
                1
                      no
    cat
       NaN
                      yes
    dog
        7.0
j
    dog
        3.0
                      no
                                                                          mostramos información
df.info()
                                                                           El método info() devuelve información del df
                                   df.describe() # Otra forma
Resultado:
                                                                          junto con sus índices y columnas, los valores
<class 'pandas.core.frame.DataFrame'>
                                                      visits
Index: 10 entries, a to j
                                                                          no nulos y el uso de memoria.
                                  count 8,000000
                                                   10.000000
Data columns (total 4 columns):
                                  mean
                                         3.437500
                                                    1.900000
                                                                          Otra forma es usando la función describe()
   Column
           Non-Null Count Dtype
                                  std
                                         2.007797
                                                    0.875595
                                                                          que devuelve una serie con resumen
0
   animal
            10 non-null
                                  min
                                         0.500000
                                                    1.000000
                         float64
            8 non-null
                                                                          descriptivo que incluye: número de datos, la
   age
                                  25%
                                         2.375000
                                                    1,000000
   visits
           10 non-null
                         int64
                                  50%
                                         3.000000
                                                    2.000000
                                                                          suma, el mínimo, la media, la desviación
   priority 10 non-null
                         object
                                  75%
                                         4,625000
                                                    2.750000
dtypes: float64(1), int64(1), object(2)
                                                                          típica y los cuartiles
                                         7.000000
                                                    3.000000
                                  max
memory usage: 400.0+ bytes
                                                                          El método head() devuelve las primeras filas
df.head(3)
                                     # Otra forma
                                     df.iloc[:3]
                                                                          especificadas del df, si no lo especificamos
Resultado:
                                                                          nos devuelve las primeros cinco.
                                     Resultado:
                                                                           El método iloc[] en este ejemplo nos
          age visits priority
  animal
                                       animal age
                                                   visits priority
     cat
           2.5
                     1
                                                                          devuelve las primeras tres filas
                             ves
а
                                         cat
                                              2.5
                                                        1
                                                               yes
b
     cat
           3.0
                     3
                             ves
                                                        3
                                         cat
                                              3.0
                                                               yes
          0.5
                     2
C
   snake
                              no
                                     C
                                       snake
                                              0.5
                                                                no
                                                                           Usado para mostrar columnas de un df.
 df.loc[:, ['animal', 'age']] # Otra manera
                                                                           El método loc[] devuelve un df con los
                                df[['animal', 'age']]
                                                                          elementos de las filas de la lista filas y las
 Resultado:
                                                                          columnas de la lista columnas. Si queremos
                                Resultado:
   animal
           age
                                                                          que muestre todas las columnas usamos :
                                  animal
                                         age
           2.5
      cat
                                                                           (dos puntos).
                                    cat
                                         2.5
 b
           3.0
      cat
                               b
                                    cat
                                         3.0
                                                                          También podemos usar doble corchete con el
    snake
           0.5
 C
                               C
                                  snake
                                         0.5
           NaN
                                                                          nombre de las columnas a especificar, por
 d
      dog
                               d
                                    dog
                                         NaN
 e
      dog
           5.0
                                                                          default nos va a traer todo
                               e
                                     dog
                                         5.0
 f
           2.0
      cat
                               f
                                         2.0
                                    cat
 g
    snake
           4.5
                                  snake
                                         4.5
      cat
           NaN
                                     cat
                                         NaN
                                                                          seleccionamos las filas [3, 4, 8] y las
df.loc[df.index[[3, 4, 8]], ['animal', 'age']]
                                                                          columnas 'animal' y 'age' usando el método
Resultado:
                                                                          index
  animal
         age
d
         NaN
    dog
    dog
         5.0
e
    dog
i
         7.0
                                                                           Ponemos un condición dentro de los
df[df['visits'] > 3] #No hay ninguno ja
                                                                          corchetes para filtrar los elementos de
<u>Resultado:</u>
                                                                           nuestra tabla.
                                                                          En el ejemplo seleccionamos las filas que
Empty DataFrame
Columns: [animal, age, visits, priority]
                                                                          tengan n° de visitas mayores que tres
Index: []
                                                                           La función isnull() devuelve los valores NaN.
df[df['age'].isnull()]
                              animal
                                             visits priority
                                       age
                                                                          ésta es usada como condición, donde
                             d
                                  dog
                                       NaN
                                                  3
                                                          yes
 Resultado:
                                                                          devuelve True si el valor de la columna es
                            h
                                  cat
                                       NaN
                                                          yes
                                                                          nulo o NaN
```

<pre>df[(df['animal'] == 'cat') & (df['age'] < 3)] Resultado: animal age visits priority a cat 2.5 1 yes f cat 2.0 3 no</pre>	Ejemplo de una condición más compleja, donde decimos que queremos las filas donde el animal sea cat y la edad menor a tres
df[df['age'].between(2, 4)] Resultado: animal age visits priority a cat 2.5 1 yes b cat 3.0 3 yes f cat 2.0 3 no j dog 3.0 1 no	La función between() elige los elementos ubicados en el rango [a, b], sería equivalente a columna >= a & columna <=b . Como ejemplo le pedimos seleccionar las filas con edad entre dos y cuatro inclusive
<pre>df.loc['f', 'age'] = 1.5</pre>	Para cambiar el valor de una fila en específica, usamos el método loc[], especificamos las columnas y el = como si fuera una variable. En este ejemplo cambiamos la edad de la fila f a 1.5
<pre>df['visits'].sum() Resultado: 19</pre>	La función sum() es usada para sumar todos los valores de la columna seleccionada. En este ejemplo estamos sumando la cantidad de visitas de todas las filas del df, como resultado vemos que son 19
<pre>df.groupby('animal')['age'].mean() Resultado: animal cat 2.333333 dog 5.000000 snake 2.500000 Name: age, dtype: float64</pre>	En el ejemplo usamos dos métodos. El primero, groupby(), agrupa las filas según el animal y el segundo, mean(), calcula el promedio de la edad
<pre>df.loc['k'] = ['dog', 5.5, 2, 'no'] df Resultado: animal age visits priority a cat 2.5</pre>	Usamos loc para agregar una nueva fila al df junto a los valores que tiene
<pre>df = df.drop('k') df Resultado: animal age visits priority a cat 2.5</pre>	Usamos la función drop() para borrar la fila agregada recientemente

df['animal'].value_counts() Resultado: cat 4 snake dog 4 Name:	2 animal, dtype: int64	La función value_counts() devuelve la cantidad de veces que aparece un valor para una columna específica. En este ejemplo queremos contar la cantidad de animales en el df
df.sort_values(by=['age', 'visits'], ascending= Resultado: animal age visits priority i dog 7.0 2 no e dog 5.0 2 no g snake 4.5 1 no j dog 3.0 1 no b cat 3.0 3 yes a cat 2.5 1 yes f cat 1.5 3 no c snake 0.5 2 no h cat NaN 1 yes d dog NaN 3 yes	[False, True])	La función sort_values() ordena la tabla usando ciertos criterios como parámetros (by, ascending). En by ponemos una lista de las columnas y en ascending la lista de booleanos donde indicamos si queremos que el orden sea ascendente o no
<pre>df['priority'] = df['priority'].map({'yes': True, df</pre>	'no': False})	La función map() transforma los valores de la columna entera. Le pasamos un diccionario con elementos del tipo valor viejo - valor nuevo. En este ejemplo tenemos una columna llamada priority y le queremos cambiar los valores sí y no por true y false
<pre>df['animal'] = df['animal'].replace('snake' df Resultado: animal age visits priority a cat 2.5 1 True b cat 3.0 3 True c python 0.5 2 False d dog NaN 3 True e dog 5.0 2 False f cat 1.5 3 False g python 4.5 1 False h cat NaN 1 True i dog 7.0 2 False j dog 3.0 1 False</pre>	, 'python')	Usamos replace() para reemplazar el valor por uno nuevo (también se puede usar map()

<u>Funciones de Pandas</u>

Función	Definición			Ejemplo de Uso	
read_csv()	1	un archivo CSV y lo carga en un aFrame	,	df = pd.read_csv('data.csv')	
head() Data		Muestra las primeras filas del DataFrame. Se le puede pasar un número Dara pedirle una N cantidad de filas.		df.head() df.head(7)	
tail()	Mue	estra las últimas filas del	,	df.tail()	
dropna()		Elimina filas con valores DataFrame	faltar	ntes del	df.dropna()
fillna()		Rellena los valores faltantes del DataFrame con un valor dado			df.fillna(0)
groupby()	1 -	pa el DataFrame según una o varias mnas		<pre>grouped = df.groupby('columna')</pre>	
sum()		icula la suma de los valores en el taFrame		total = df['columna'].sum()	
mean()		cula la media de los valores en e taFrame	el average = df['columna'].mean()		['columna'].mean()
max()		cuentra el valor máximo en el taFrame	<pre>max_value = df['columna'].max(</pre>		df['columna'].max()
count()		ive el número de elementos no en el objeto	non_nu	ull_count = df['	columna'].count()
map()		una función o un diccionario a los ntos del objeto	<pre>mapped_values = df['columna'].map(func)</pre>		columna'].map(func)
value_counts()		lve una serie que contiene utos de valores únicos	<pre>value_counts = df['columna'].value_counts()</pre>		

shape	Devuelve la forma (número de columnas) del DataFrame	shape = df.shape		
info()	Muestra información sobre el	df.info()		
describe()	Genera estadísticas descripti DataFrame	df.describe()		
columns	Devuelve los nombres de las del DataFrame	columns = df.columns		
dtypes	Devuelve los tipos de datos d columnas del DataFrame	dtypes = df.dtypes		
min()	incuentra el valor mínimo en el ataFrame	df['columna'].min()		
unique()	evuelve los valores únicos en una olumna del DataFrame	es = df['columna'].unique()		
nunique()	evuelve el número de valores únicos en na columna del DataFrame	<pre>num_unique = df['columna'].nunique()</pre>		
sort_values()	rdena el DataFrame por una o varias olumnas	df.sort_values('columna') df.sort_values(by=['columna1','columna2 ascending=[True,False])		
merge()	ombina dos DataFrames en función de na o varias columnas	<pre>merged_df = pd.merge(df1, df2, on='columna')</pre>		
to_csv() G	uarda el DataFrame en un archivo CSV	df.to_csv('out	iput.csv', index=False)	
ize()	evuelve el número total de elementos n el objeto	total_elements	s = df.size	

			loc[]	Permite la indexación y selección de	df.loc[2]
isnull()	Devuelve una máscara booleana que indica valores nulos	<pre>is_null_mask = df['columna'].isnull()</pre>	100[]	datos por etiquetas.	df.loc[1:3, 'columna']
					df.iloc[2]
notnull()	Devuelve una máscara booleana que indica valores no nulos	<pre>not_null_mask = df['columna'].notnull()</pre>	iloc[]	Permite la indexación y selección de datos por posición.	df.iloc[1:3, 0]
reset_index()	Restablece los índices del DataFrame	df_reset = df.reset_index()	get()	Devuelve el valor correspondiente a una clave dada	value = df.get('columna')

loc vs iloc

- son métodos de indexación usados para acceder a filas y columnas específicas
- loc → se basa en etiquetas, usa las etiquetas de filas y columnas para acceder a los datos
- iloc → se basa en números, usa las posiciones de los valores para acceder a los datos

	padron	nombre	apellido	edad	carrera	mail_fiuba	ingreso_en_pandemia
1	109234	Agustín	Ramírez	18	Química	AgustinRamirez	si
2	110456	María	Portillo	19	Civil	MariaPortillo	no
3	102890	Gustavo	Feliche	20	Mecánica	None	si
4	100450	Federico	Monte	35	Informática	FedericoMonte	si
5	100430	Felipe	Michael	25	Informática	FelipeMichael	no
6	100455	Manuela	Gomez	20	Informática	ManuelaGomez	no
7	105689	Laura	Ramírez	24	Sistemas	None	si

Para el dataframe de la foto, usando loc para obtener nombre y edad se ve así:

```
nombres_edades_v2 = df.loc[:, ['nombre', 'edad']]
nombres_edades_v2
```

Pero usando iloc se ve así:

```
nombres_edades_v2 = df.iloc[:, [1,3]]
```

nombres_edades_v2

Los dos puntos (:) marcan que queremos mostrar todas las filas, pero se puede modificar para mostrar menos. Por ejemplo:

:10 imprime del índice 0 al 10.

Matplotlib

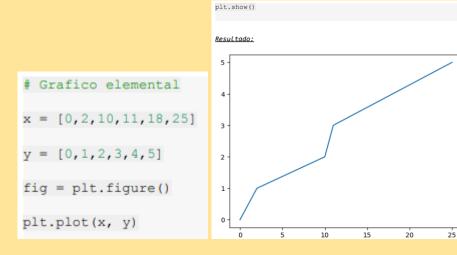
- es una biblioteca de Python usada para crear gráficos/plots
- es una forma rápida de graficar datos en diferentes formatos que pueden ser compartidos o publicados
- pyplot le proporciona un interfaz, está diseñada siguiendo el estilo matlab (tiene comandos y argumentos similares a éste)

import matplotlib.pyplot as plt

importamos la librería junto al alias plt

Función Plot y Show

- recibe dos vectores que tienen el mismo tamaño, estos representan una serie de puntos en el plano cartesiano. El primer arreglo son coordenadas X y el segundo son coordenadas Y
- tiene como función unir todos los puntos, con líneas, para que podamos ver el gráfico de la función
- plt.plot() → método usado para crear un gráfico para verificar resultados
- plt-subplots() → para gráficos más complejos con más finos
- show → función con la que creamos la imagen de todos los gráficos definidos



```
fig, ax = plt.subplots()
ax.plot(x, y)
plt.show()
```

- cuál se diferencia? → los outputs
 - plt.figure() crea una figura pero sin axes (por eso dice <Figure size 640x480 with 0 Axes>, se tiene una figura y 0 aces).
 - plt.subplots() permite crear ambos (por eso dice <Figure size 640x480 with 1 Axes>, <Axes: >, hay figure y axes).

FIGURA

- es el marco que delimita la zona donde se traza los gráficos
- una figura puede tener muchos axes

AXES

- son los gráficos, o sea las áreas donde los puntos se pueden especificar en términos de coordenadas

o también

 un axes determinado solo puede estar contenido en una única figura

NO CONFUNDIR AXES CON AXIS. Los axis son ejes cartesianos encargados de establecer los límites, la escala y las dimensiones del gráfico

un axes puede tener dos axis (si es un gráfico plano) o tres (si es un gráfico tridimensional)

```
# Grafico elemental explicado:

fig = plt.figure() # Se crea una figura vacía sin Axes

plt.plot(x, y) # Agregar ploteo (gráfico de linea)

plt.show() # Mostrar

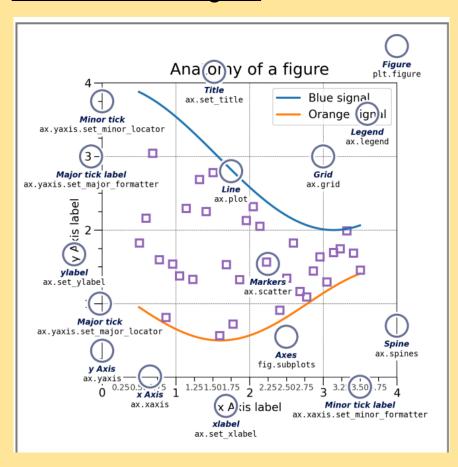
# Grafico elemental explicado:

fig, ax = plt.subplots() # Se crea una figura con un único Axes.

ax.plot(x, y) # Agregar los ploteos individuales (gráfico de linea)

plt.show() # Mostrar
```

Partes de una Figura



¿Cómo cambiar el aspecto de los gráficos?

- definir los parámetros dentro de ax.plot()
- podemos establecer el tipo de línea y puntos, el grosor, color, etc
- color → para el color
- maker → forma de los puntos marcadores (^, o, v)
- linestyle → estilo de línea (solid/-, dashed/--, dotted/:)
- markersize, linewidth → establece el tamaño del marcador y el espesor de la línea (se usa números, si no le asignamos uno se establecen los predefinidos)

```
x = [0,2,10,11,18,25] # Tiempo (min)
y = [0,1,2,3,4,5] # Distancia (m)

fig, ax = plt.subplots()

ax.plot(x, y, color='green', marker='^', linestyle='--', markersize=8, linewidth=1.2)
plt.show()
```

Grilla o cuadrícula

- usado para leer cada punto
- se agrega usando ax.grid()
- podemos modificarla usando parámetros similares a los anteriores pero en la función ax.grid()

```
x = [0,2,10,11,18,25] # Tiempo (min)
y = [0,1,2,3,4,5] # Distancia (m)

fig, ax = plt.subplots()
ax.plot(x, y, color='green', marker='^', linestyle='--', markersize=8,
linewidth=1.2)
# Grilla preestablecida
ax.grid()
plt.show()
```

```
#Gráfica con la grilla preestablecida
x = [0,2,10,11,18,25] # Tiempo (min)
y = [0,1,2,3,4,5] # Distancia (m)

fig, ax = plt.subplots()

ax.plot(x, y, color='green', marker='^', linestyle='--', markersize=8, linewidth=1.2)

#Grilla modificada
ax.grid(axis = 'y', color = 'gray', linestyle = 'dashed')
plt.show()
```

Títulos

- es importante agregarle esto a los gráficos para que se pueda entender y explicar qué significa cada eje
- usamos las funciones ax.set_xlabel(), ax.set_ylabel(), ax.set_title(). Cada una recibe un string que se usará como etiqueta para los ejes x/y o el título

```
x = [0,2,10,11,18,25] # Tiempo (min)
y = [0,1,2,3,4,5] # Distancia (m)

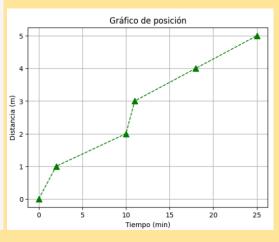
fig, ax = plt.subplots()

ax.plot(x, y, color='green', marker='^', linestyle='--', markersize=8, linewidth=1.2)

# Mostrar el título del gráfico
ax.set_title("Gráfico de posición")

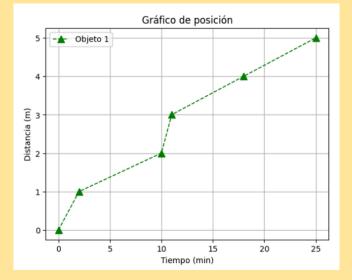
# Mostrar el título de los ejes
ax.set_xlabel('Tiempo (min)')
ax.set_ylabel('Distancia (m)')

# Grilla preestablecida
ax.grid()
plt.show()
```



Referencias

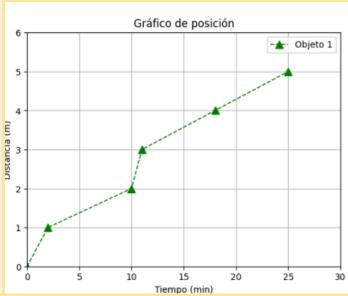
- usado cuando el gráfico tiene más de una línea
- lo ponemos dentro de ax.plot() donde definimos la referencia como label y luego colocamos ax.legend()



Características de los Ejes

podemos establecer los límites de los ejes usando ax.set_xlim() y ax.set_ylim()

```
x = [0,2,10,11,18,25] # Tiempo (min)
y = [0,1,2,3,4,5] # Distancia (m)
fig, ax = plt.subplots()
ax.plot(x, y, label='Objeto 1', color='green', marker='^', linestyle='--',
markersize=8, linewidth=1.2)
# Mostrar el título del gráfico
ax.set title("Gráfico de posición")
# Mostrar el título de los ejes
ax.set_xlabel('Tiempo (min)')
ax.set_ylabel('Distancia (m)')
# Establecer los límites de los ejes
ax.set_xlim(0, 30)
ax.set_ylim(0, 6)
# Agregar la refencia
ax.legend()
# Grilla preestablecida
ax.grid()
plt.show()
```



Tipos de Gráficos

Gráfico de Línea

- permite visualizar cambios en los valores a lo largo de un rango continuo (tendencias) (ejemplo: tiempo o distancia)
- lo creamos usando la función ax.plot()

```
x = [0,2,10,11,18,25]

y = [0,1,2,3,4,5]

fig, ax = plt.subplots()

ax.plot(x, y)

plt.show()
```

Gráfico de Dispersión / De Puntos

- permite visualizar la relación entre las variables
- la creamos usando la función ax.scatter()

```
x = [5,7,8,7,2,17,2,9,4,11,12,9,6]
y = [99,86,87,88,111,86,103,87,94,78,77,85,86]
fig, ax = plt.subplots()
ax.scatter(x, y)
```

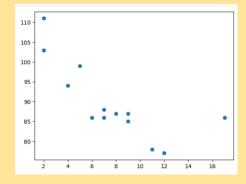


Gráfico de Barras

plt.show()

- permite visualizar proporciones, comparando dos o más valores entre sí
- usamos la función ax.bar() para crearlos
- primero recibe un arreglo con las etiquetas de las barras y otro con la altura de cada una de las barras

```
peso = [340, 115, 200, 200, 270]

ingredientes = ['chocolate', 'manteca', 'azúcar', 'huevo', 'harina']

fig, ax = plt.subplots()

ax.set_ylabel('Masa (g)')

ax.set_title("Receta")

ax.set_title("Receta")

plt.show()
```

- usamos la función ax.barh() si queremos las barras horizontales (no olvidarnos de cambiar los títulos de los ejes)

```
peso = [340, 115, 200, 200, 270]
ingredientes = ['chocolate', 'manteca', 'azúcar', 'huevo', 'harina']
```

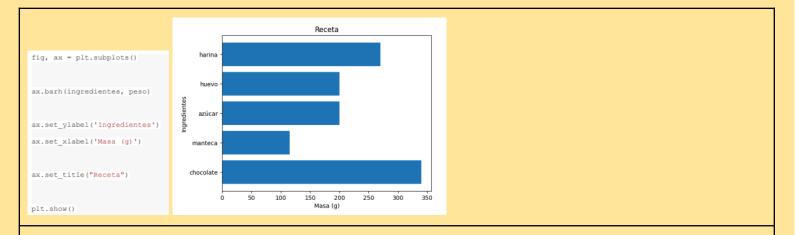
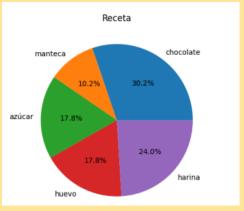


Gráfico de Tortas

- al igual que el de barras, este permite visualizar y comparar proporciones pero de manera circular y como parte de un todo
- se crea con la función ax.pie()
- usando el parámetro labels se le asigna una etiqueta a cada porción y el parámetro autopct para que indique como se mostrará el porcentaje

el parámetro autopot que indica cómo se mostrará el porcentaje: %1.1f% le indica que el porcentaje tendrá un decimal, mientras que %1.2f% tendrá 2 decimales.

```
peso = [340, 115, 200, 200, 270]
ingredientes = ['chocolate', 'manteca', 'azúcar', 'huevo', 'harina']
fig, ax = plt.subplots()
ax.pie(peso, labels= ingredientes, autopct='%1.1f%%')
ax.set_title("Receta")
```



Gráficos Múltiples

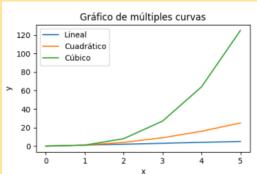
- sirve para graficar varias curvas en un mismo gráfico

```
# Valores que se desean graficar
x = [0, 1, 2, 3, 4, 5]
y_linear = [0, 1, 2, 3, 4, 5]
y_quadratic = [0, 1, 4, 9, 16, 25]
y_cubic = [0, 1, 8, 27, 64, 125]

fig, ax = plt.subplots(figsize=(5, 3))

ax.plot(x, y_linear, label='Lineal')
ax.plot(x, y_quadratic, label='Cuadrático')
ax.plot(x, y_cubic, label='Cúbico')
```

```
ax.set_title("Gráfico de múltiples curvas")
ax.set_xlabel('x')
ax.set_ylabel('y')
ax.legend()
```



Note que estamos agregando nuevos datos al mismo axes, por lo que siempre usamos ax.plot() pero con distintos valores de y. Además, se estableció un tamaño de la figura con figsize=(width, height).

Grilla de Gráficos

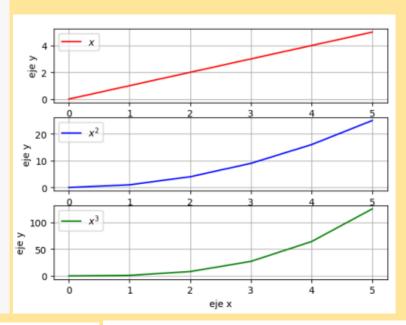
- sirve para ver varios axes en una misma figura
- necesitamos definir cuantas columnas (ncols) y filas (nrows) de gráficos necesitamos

```
fig, ax = plt.subplots(nrows=1, ncols=2) # o simplemente plt.subplots(1,2)
 1.0
                                        1.0
 0.8
                                        0.8
 0.6
                                        0.6
                                                                                          De manera análoga, podemos representar las 3 curvas de la figura titulada "Gráfico de múltiples curvas" pero viendo 3 filas en una única columna de gráficos:
 0.4
                                        0.4
                                                                                           x_{linear} = [0, 1, 2, 3, 4, 5]
 0.2
                                        0.2
                                                                                           x_quadratic = [0, 1, 4, 9, 16, 25]
                                                                                           x_cubic = [0, 1, 8, 27, 64, 125]
                                    --| 0.0 <del>|--</del>
1.0 0.0
 0.0
                                                                                          fig, ax = plt.subplots(nrows=3, ncols=1)
ax[0].plot(x, x_linear)
                                                                                     Lineal
ax[0].set_title('Lineal')
                                                                                   ,Cuadrático,
ax[1].plot(x, x_quadratic)
ax[1].set_title('Cuadrático')
                                                                                    Cúbico ;
                                                         100
ax[2].plot(x, x_cubic)
                                                          50
ax[2].set title('Cúbico')
                                                       Entonces, ahora en vez de graficar usando ax . plot() unicamente, tengo que indicar la
                                                       posición del axes con los números dentro del corchete, ¿pero si tengo varias columnas y filas?
plt.show()
                                                                                                                                                             0.0
                                                                                                                          0.0
fig, ax = plt.subplots(nrows=3, ncols=3)
                                                                                                                          1.0
fig.subplots_adjust(wspace=0.5, hspace=0.5) # Con esto indicamos el espacio
                                                                                                                          0.0
libre entre los subplots
                                                                                                                          1.0
ax[0, 1].plot(x, x_linear)
                                                                                                                                            0.0
plt.show()
                                                                                                                         Note que dentro del corchete, primero se indica la fila y luego la columna: ax[fila, columna].
```

Funciones de Gráficas

```
x = [0, 1, 2, 3, 4, 5]
```

```
x linear = [0, 1, 2, 3, 4, 5]
x_{quadratic} = [0, 1, 4, 9, 16, 25]
x cubic = [0, 1, 8, 27, 64, 125]
fig, ax=plt.subplots(3)
ax[0].plot(x,x linear,label="$x$",color="r")
ax[0].set_xlabel("eje x")
ax[0].set_ylabel("eje y")
ax[0].legend()
ax[0].grid()
ax[1].plot(x,x_quadratic,label="$x^2$",color="b")
ax[1].set xlabel("eje x")
ax[1].set_ylabel("eje y")
ax[1].legend()
ax[1].grid()
ax[2].plot(x,x_cubic,label="$x^3$",color="g")
ax[2].set_xlabel("eje x")
ax[2].set ylabel("eje y")
ax[2].legend()
ax[2].grid()
plt.show()
```



```
Para evitar lo anterior, definimos una función a la que le debemos entregar los valores a graficar:
```

```
def create_easy_graph(x, y, label, ax, xlabel, ylabel, title, color):
"""Crea un gráfico a partir de vectores con valores de los ejes x e y.
Recibe además:
- El texto para el label
- El subplot a donde graficar
- Un label para el eje x
- Un label para el eje y
- Un título para el gráfico
- Un color
El color y el eje pueden ser None. En ese caso toman valores por default"""
if color == None:
```

```
color = "blue"

# Si sólo haremos un gráfico, no necesito indicarle la posición
if ax == None:
fig, ax = plt.subplots()

# Definimos el gráfico
ax.plot(x, y, label=label, color=color)
ax.set_xlabel(xlabel)
ax.set_ylabel(ylabel)
ax.set_title(title)
```

Y por otro lado:

```
fig , ax = plt.subplots(3)

# En vez de copiar y pegar el código, llamo a la función easy_graph():

create_easy_graph(x, x_linear, "x", ax[0], "Eje x", "Eje y", "Lineal",

color="green")

create_easy_graph(x, x_quadratic, "$x^2$", ax[1], "Eje x", "Eje y",

"Cuadrática", color="red")

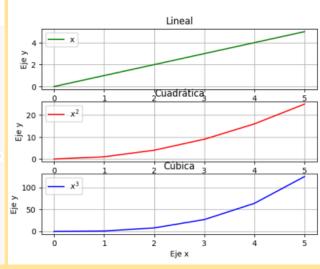
create_easy_graph(x, x_cubic, "$x^3$", ax[2], "Eje x", "Eje y", "Cúbica", None)

# Hacemos un for para agregar la cuadrícula y las referencias en cada axes:

for axes in fig.axes[:]:

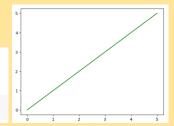
axes.grid()

axes.legend()
```



Como comentamos dentro de la función, también podemos usar easy_graph() para un único gráfico:

```
create_easy_graph(x, x_linear, "x", None, "", "", "", "green")
```



Gráficos usando DataFrames (df)

• usamos los datos contenidos en el df (no es necesario que provenga de vectores de valores)

```
import pandas as pd

data = ('animal': ['cat', 'snake', 'dog'],
   'age': [2.5, 3, 7],
   'visits': [1, 3, 2],
   'priority': ['yes', 'yes', 'no']}

df = pd.DataFrame(data)

df
```

	animal	age	visits	priority
0	cat	2.5	1	yes
1	snake	3.0	3	yes
2	dog	7.0	2	no

```
# Determino las columnas del DataFrame que queremos graficar
x_values = df['animal']
y_values = df['age']

fig, ax = plt.subplots()

ax.bar(x_values, y_values)

ax.set_xlabel('Animal')
ax.set_ylabel('Edad (años)')

ax.set_title("Mascotas")
```

