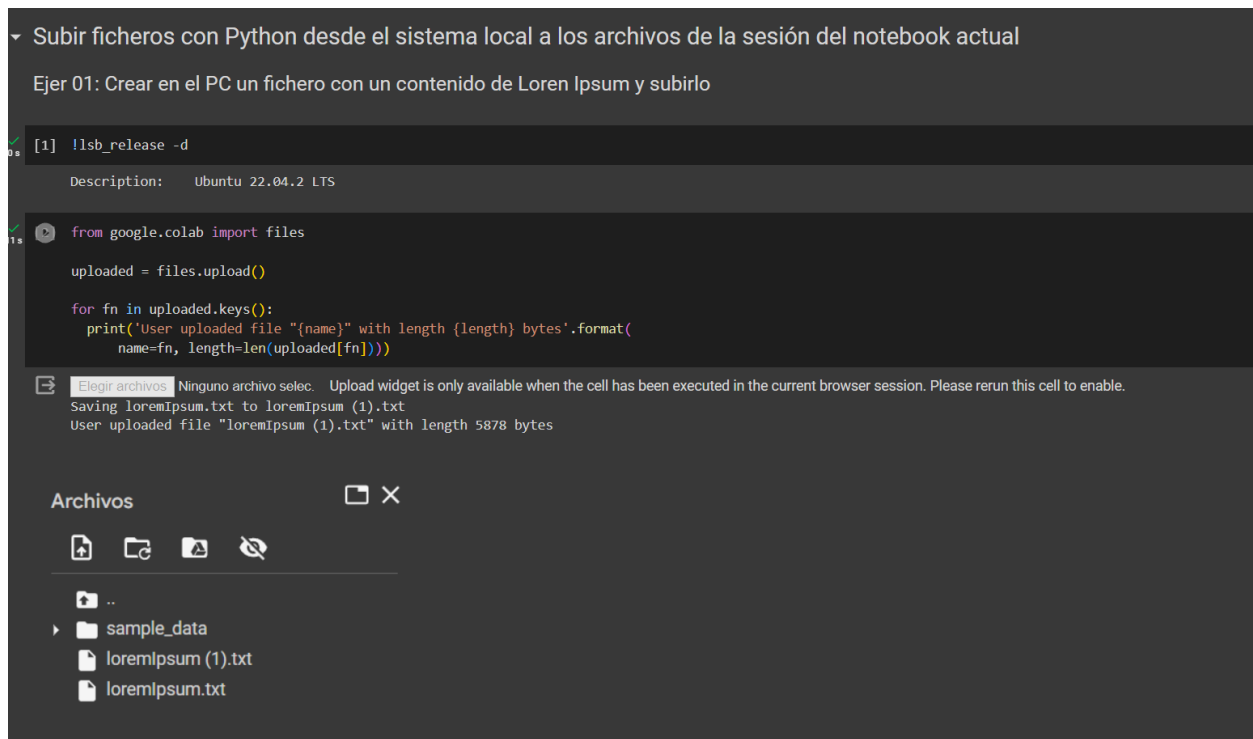


Actividad 1.1 – Gestión de ficheros en Google colab





▼ Descargar archivos en el sistema de archivos local

Ejer 02: Descargar el fichero subido en el apartado anterior

```
✓ 0 s from google.colab import files  
files.download('loremIpsum (1).txt')
```



 loremIpsum (1)	21/10/2023 11:34	Documento de tex...	6 KB
 loremIpsum	21/10/2023 11:05	Documento de tex...	6 KB

Google Drive

▼ Google Drive

▼ Activar (Montar) Google Drive de forma local

```
✓ 0 s # Muestra las unidades montadas  
!df -h
```



Filesystem	Size	Used	Avail	Use%	Mounted on
overlay	108G	28G	81G	26%	/
tmpfs	64M	0	64M	0%	/dev
shm	5.8G	0	5.8G	0%	/dev/shm
/dev/root	2.0G	1.1G	885M	55%	/usr/sbin/docker-init
tmpfs	6.4G	220K	6.4G	1%	/var/colab
/dev/sda1	46G	28G	18G	62%	/etc/hosts
tmpfs	6.4G	0	6.4G	0%	/proc/acpi
tmpfs	6.4G	0	6.4G	0%	/proc/scsi
tmpfs	6.4G	0	6.4G	0%	/sys/firmware

```
✓ 35 s [5] from google.colab import drive  
drive.mount('/content/drive')  
# Observar que al montar la carpeta se puede navegar en Google Drive desde el explorar de carpetas que está a la  
# izquierda en Google Colab
```

Mounted at /content/drive

```
✓ 0 s [6] # Muestra las unidades montadas  
!df -h
```

Filesystem	Size	Used	Avail	Use%	Mounted on
overlay	108G	28G	81G	26%	/
tmpfs	64M	0	64M	0%	/dev
shm	5.8G	0	5.8G	0%	/dev/shm
/dev/root	2.0G	1.1G	885M	55%	/usr/sbin/docker-init
tmpfs	6.4G	224K	6.4G	1%	/var/colab
/dev/sda1	46G	28G	18G	62%	/etc/hosts
tmpfs	6.4G	0	6.4G	0%	/proc/acpi
tmpfs	6.4G	0	6.4G	0%	/proc/scsi
tmpfs	6.4G	0	6.4G	0%	/sys/firmware
drive	108G	32G	77G	29%	/content/drive

```
drive 108G 28G 77G 26% /content/drive

[7] # Aunque reiniciemos el entorno de ejecución la unidad continúa montada

from google.colab import drive
drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

▼ Crear un fichero en Drive desde Google Colab

[8] with open('/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/nestorPruebaFichero_01.txt', 'w') as f:
    f.write('Hello Google Drive 01!')

!cat '/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/nestorPruebaFichero_01.txt'

Hello Google Drive 01!

▼ SNS
  ▼ ACT1_1
    Actividad_1_1_Gestión de ficheros en Google Colab.pdf
    loremIpsum (1).txt
    loremIpsum.txt
    nestorPruebaFichero_01.txt
```

```
▼ Desmonta las unidades montadas

[9] drive.flush_and_unmount()

[10] # Muestra las unidades montadas
!df -h

Filesystem      Size  Used Avail Use% Mounted on
overlay         108G   28G   81G   26% /
tmpfs           64M     0   64M    0% /dev
shm            5.8G     0  5.8G    0% /dev/shm
/dev/root       2.0G   1.1G  885M   55% /usr/sbin/docker-init
tmpfs          6.4G   320K  6.4G    1% /var/colab
/dev/sda1       46G    28G   18G   62% /etc/hosts
tmpfs          6.4G     0  6.4G    0% /proc/acpi
tmpfs          6.4G     0  6.4G    0% /proc/scsi
tmpfs          6.4G     0  6.4G    0% /sys/firmware

[11] !ls -l

total 20
-rw-r--r-- 1 root root 5878 Oct 21 10:27 'loremIpsum (1).txt'
-rw-r--r-- 1 root root 5878 Oct 21 10:23 loremIpsum.txt
drwxr-xr-x 1 root root 4096 Oct 19 16:36 sample_data
```

▼ Crear un directorio en Google Drive

```
✓ [12] 2s from google.colab import drive
drive.flush_and_unmount # Lo utilizo para poder hacer el ejemplo con autonomía de otras celdas

drive.mount('/content/drive', force_remount=True)

!mkdir '/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/GoogleColab_test_2023_2024'
```

Mounted at /content/drive

- ▼ SNS
 - ▼ ACT1_1
 - ▶ GoogleColab_test_2023_2024
 - 📄 Actividad_1_1_Gestión de ficheros en Google Colab.pdf
 - 📄 loremIpsum (1).txt
 - 📄 loremIpsum.txt
 - 📄 nestorPruebaFichero_01.txt

▼ Ejer 03: Crear en drive un directorio llamado GoogleColab_test_Ejer_23_24 y dentro de ella un fichero con un contenido de Loren Ipsum. Consultar/verificar su creación en Google Drive y en el navegador de Google Colab

```
✓ [13] 2s from google.colab import drive
drive.flush_and_unmount # Lo utilizo para poder hacer el ejemplo con autonomía de otras celdas

drive.mount('/content/drive', force_remount=True)

!mkdir '/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/GoogleColab_test_Ejer_23_24'

with open('/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/GoogleColab_test_Ejer_23_24/lorenIpsum.txt', 'w') as f:
    f.write('Lorem ipsum dolor sit amet consectetur adipiscing elit elementum, aptent a ridiculus vivamus inceptos sagittis tempor')

!cat '/content/drive/MyDrive/Master FP IA y BD/SNS/ACT1_1/GoogleColab_test_Ejer_23_24/lorenIpsum.txt'
```

Mounted at /content/drive

Lorem ipsum dolor sit amet consectetur adipiscing elit elementum, aptent a ridiculus vivamus inceptos sagittis tempor nisi vehic

Navegador Google Colab

- ▼ SNS
 - ▼ ACT1_1
 - ▶ GoogleColab_test_2023_2024
 - ▼ GoogleColab_test_Ejer_23_24
 - 📄 loremIpsum.txt
 - 📄 Actividad_1_1_Gestión de ficheros en Google Colab.pdf
 - 📄 loremIpsum (1).txt
 - 📄 loremIpsum.txt
 - 📄 nestorPruebaFichero_01.txt

Google Drive

...

>

SNS

>

ACT1_1

▼

Tipo

Personas

Modificado

Nombre	Propietario	Última modificación	Tamaño de a	
.ipynb_checkpoints	yo	11:43 yo	—	
GoogleColab_test_2023_2024	yo	11:47 yo	—	
GoogleColab_test_Ejer_23_24	yo	11:58 yo	—	
Actividad_1.1_Gestión de ficheros en Google Colab.pdf	yo	16 oct 2023 yo	469 kB	
lorem Ipsum (1).txt	yo	11:34 yo	6 kB	
lorem Ipsum.txt	yo	11:05 yo	6 kB	
nestorPruebaFichero_01.txt	yo	11:42 yo	22 bytes	

...

>

ACT1_1

>

GoogleColab_test_Ejer_...

▼

Tipo

Personas

Modificado

Nombre	Propietario	Última modificación	Tamaño de a	
lorem Ipsum.txt	yo	11:58 yo	672 bytes	

▼

Uso de PyDrive

```
[14] from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials

[15] auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

[16] uploaded = drive.CreateFile({'title': 'Sample upload 2324.txt'})
uploaded.SetContentString('Sample upload file content')
# En este caso lo crea y sube en el directorio raíz de Drive, en el caso de que deseemos utilizar un directorio en concreto se lo pasamos por parámetro
uploaded.Upload()
print('Uploaded file with ID {}'.format(uploaded.get('id')))

Uploaded file with ID 1MktDu0-SUJ-NsymBrl0lh5NFjZYAEnf

[22] file = drive.CreateFile({'id': uploaded.get('id')})
print('Downloaded content "{}".format(file.GetContentString()))

Downloaded content "Sample upload file content"

[18] # Para mostrar los directorios/ficheros y sus id's
# Fijarse que el nombre del fichero/directorio se identificaca con "title"
file_list = drive.ListFile({'q': "'root' in parents and trashed=false"}).GetList()
for file1 in file_list:
    print('title: {}, id: {}'.format(file1['title'], file1['id']))

title: Sample upload 2324.txt, id: 1MktDu0-SUJ-NsymBrl0lh5NFjZYAEnf
title: Economía 2023, id: 1uCdqRaiM07o2rn2x0z-1UzFvduGgd_kQFB4s5w6iv24
title: Master FP IA y BD, id: 1iKrxtl5NBX7judlqM_9cANwM7BeyQw4
title: Notas, id: 1csH3Ao6piPyEhJaNwhJHaBAewWeNoAitx-YdPf18oQ
```

Néstor Batista Díaz

5/12

Ejer 04: En la carpeta creada en el ejercicio anterior (GoogleColab_test_Ejer_23_24), crear un fichero cuyo nombre sean tu nombre y apellidos, en su contenido incluir un texto LoremIpsum distinto al utilizado en el punto anterior (de 50 palabras y añadir tu nombre al principio) y al terminar mostrar su contenido. Similar al ejemplo. Validar su existencia consultando Gdrive.ListFile y el explorador de archivos de GoogleColab en este notebook

```
[30] folderName = 'GoogleColab_test_Ejer_23_24' # Please set the folder name.

folders = drive.ListFile(
    {'q': "title='" + folderName + "' and mimeType='application/vnd.google-apps.folder' and trashed=false"}).GetList()
for folder in folders:
    if folder['title'] == folderName:
        file2 = drive.CreateFile({'title': 'NestorBatistaDiaz.txt', 'parents': [{'id': folder['id']}]}))
        file2.SetContentString('Néstor Batista Díaz\nLorem ipsum dolor sit amet consectetur adipiscing, elit ut rhoncus nullam egestas lacinia, cursus morbi
        file2.Upload()

        file3 = drive.CreateFile({'id': file2.get('id')})
        print('Contenido del fichero ', file3['title'], ':')
        print(file3.GetContentString())
```

Contenido del fichero NestorBatistaDiaz.txt :
Néstor Batista Díaz

Lorem ipsum dolor sit amet consectetur adipiscing, elit ut rhoncus nullam egestas lacinia, cursus morbi quisque torquent mauris. Turpis sollicitudin tincidunt

▼ SNS

- ▼ ACT1_1
 - ▶ GoogleColab_test_2023_2024
 - ▼ GoogleColab_test_Ejer_23_24
 - NestorBatistaDiaz.txt
 - loremIpsum.txt

▼ Ejer 05: Mostrar una lista de los ficheros que existen en el directorio GoogleColab_test_Ejer_23_24

```
for folder in folders:
    if folder['title'] == folderName:
        file_list2 = drive.ListFile({'q': "'"+folder['id']+" in parents and trashed=false"}).GetList()
        for file2 in file_list2:
            print('title: %s, id: %s' % (file2['title'], file2['id']))
```

title: .ipynb_checkpoints, id: 1-LdeUbjhougOp88U-Tvn8nsFW0S8Rcx1
title: NestorBatistaDiaz.txt, id: 1FmniA0XKlbbOwmv-sYkb05XifhsDcExp
title: loremIpsum.txt, id: 1-DxLUJ7w3kRny3kLKMa17bxjMo0pTVaX

▼ API de REST de Drive

▼ Crear un archivo de Drive nuevo con datos de Python

```
✓ [46] from google.colab import auth
      auth.authenticate_user()
      from googleapiclient.discovery import build
      drive_service = build('drive', 'v3')
```

```
▼ tmp
  ▶ initgoogle_syslog_dir.0
  ▶ pyright-275-Kpwt5ayXpi2Z
  ▶ pyright-275-q0M4CEvB8EQm
  ▶ python-languageserver-cancellation
    dap_muxlexer.8a72cbd71623.root.log.INFO.20231021-102221.78
    dap_muxlexer.INFO
    directoryprefetcher_binary.8a72cbd71623.root.log.INFO.20231021-10385.
    directoryprefetcher_binary.8a72cbd71623.root.log.INFO.20231021-10473.
    directoryprefetcher_binary.8a72cbd71623.root.log.INFO.20231021-10583.
    directoryprefetcher_binary.INFO
    drive.8a72cbd71623.root.log.ERROR.20231021-103857.4215
    drive.8a72cbd71623.root.log.ERROR.20231021-104737.6425
    drive.8a72cbd71623.root.log.ERROR.20231021-105836.9177
    drive.8a72cbd71623.root.log.INFO.20231021-103852.4123
    drive.8a72cbd71623.root.log.INFO.20231021-103852.4130
    drive.8a72cbd71623.root.log.INFO.20231021-103857.4123
    drive.8a72cbd71623.root.log.INFO.20231021-103857.4215
    drive.8a72cbd71623.root.log.INFO.20231021-104439.5684
    drive.8a72cbd71623.root.log.INFO.20231021-104737.6422
    drive.8a72cbd71623.root.log.INFO.20231021-104737.6425
    drive.8a72cbd71623.root.log.INFO.20231021-105836.9173
    drive.8a72cbd71623.root.log.INFO.20231021-105836.9177
    drive.8a72cbd71623.root.log.WARNING.20231021-103857.4215
    drive.8a72cbd71623.root.log.WARNING.20231021-104737.6425
    drive.8a72cbd71623.root.log.WARNING.20231021-105836.9177
    drive.ERROR
    drive.INFO
    drive.WARNING
    language_service.8a72cbd71623.root.log.INFO.20231021-102255.270
    language_service.INFO
    to_upload_23_24.txt
```

```
✓ 0 s [48] with open('/tmp/to_upload_23_24.txt', 'w') as f:
      f.write('Este año tenemos nuevos retos.')

      print('/tmp/to_upload_23_24.txt contains:')
      !cat /tmp/to_upload_23_24.txt

/tmp/to_upload_23_24.txt contains:
Este año tenemos nuevos retos.

✓ 2 s [49] from google.colab import drive
      drive.mount('/content/drive')

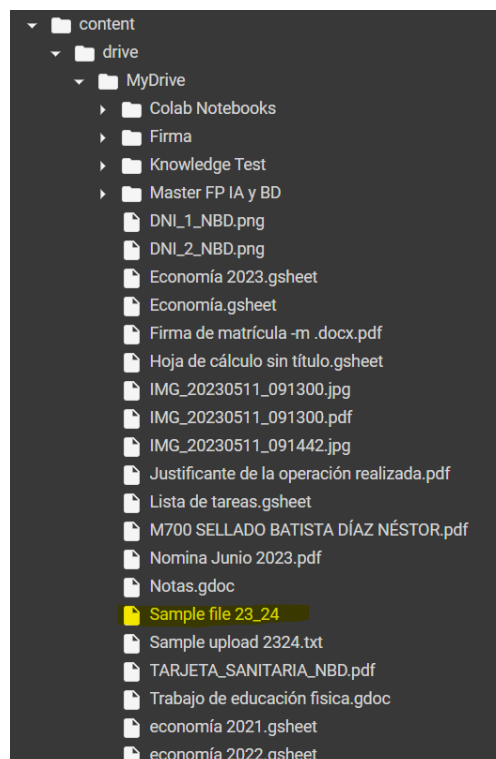
Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

✓ 1 s [50] from googleapiclient.http import MediaFileUpload

      file_metadata = {
          'name': 'Sample file 23_24',
          'mimeType': 'text/plain'
      }
      media = MediaFileUpload('/tmp/to_upload_23_24.txt',
                              mimetype='text/plain',
                              resumable=True)
      created = drive_service.files().create(body=file_metadata,
                                              media_body=media,
                                              fields='id').execute()

      print('File ID: {}'.format(created.get('id')))
```

File ID: 1fL-4cqpeetytKWccqIYp0baRkVmyCjRX



▼ Descargar datos de un archivo de Drive a un archivo lógico

✓
0 s

```
▶ file_id = created.get('id')

import io
from googleapiclient.http import MediaIoBaseDownload

request = drive_service.files().get_media(fileId=file_id)
downloaded = io.BytesIO()
downloader = MediaIoBaseDownload(downloaded, request)
done = False
while done is False:
    # _ is a placeholder for a progress object that we ignore.
    # (Our file is small, so we skip reporting progress.)
    _, done = downloader.next_chunk()

downloaded.seek(0)
print('Downloaded file contents are: {}'.format(downloaded.read()))
```

➞ Downloaded file contents are: b'Este a\x03\xb1o tenemos nuevos retos.'

▼ Ejer 06: Haciendo uso de la api (ejemplo anterior) descargar/mostrar el contenido del fichero que creaste con contenido LoremIpsum en la carpeta GoogleColab_test_Ejer y con el nombre tuyo.

✓
0 s

```
[52] file_id = file3.get('id')

import io
from googleapiclient.http import MediaIoBaseDownload

request = drive_service.files().get_media(fileId=file_id)
downloaded = io.BytesIO()
downloader = MediaIoBaseDownload(downloaded, request)
done = False
while done is False:
    # _ is a placeholder for a progress object that we ignore.
    # (Our file is small, so we skip reporting progress.)
    _, done = downloader.next_chunk()

downloaded.seek(0)
print('Downloaded file contents are: {}'.format(downloaded.read()))
```

Downloaded file contents are: b'N\x03\xa9stor Batista D\x03\xadaz\n\nLorem ipsum dolor sit amet consectetur adipiscing, elit ut rhoncus

▼ Hojas de cálculo de Google

```
✓ [53] from google.colab import auth
      auth.authenticate_user()

      import gspread
      from google.auth import default
      creds, _ = default()

      gc = gspread.authorize(creds)
```

▼ Crear una hoja nueva con datos de Python

```
✓ [54] # Crea la hoja de cálculo en el directorio raíz de Google Drive
      sh = gc.create('My cool spreadsheet_23_24')
```

```
✓ [56] worksheet = gc.open('My cool spreadsheet_23_24').sheet1
1 s
      cell_list = worksheet.range('A1:C2')

      import random
      for cell in cell_list:
          cell.value = random.randint(1, 10)

      worksheet.update_cells(cell_list)
```

```
➡ {'spreadsheetId': '1qHu-A9hg28uaAvd2_LzJ2jkbTEf1iUroHSpZjvBhitI',
   'updatedRange': "'Hoja 1'!A1:C2",
   'updatedRows': 2,
   'updatedColumns': 3,
   'updatedCells': 6}
```

▼ Descargar datos de una hoja de cálculo en Python como un DataFrame de Pandas

```
✓ 1 s [57] worksheet = gc.open('My cool spreadsheet_23_24').sheet1

      # get_all_values gives a list of rows.
      rows = worksheet.get_all_values()
      print(rows)

      import pandas as pd
      df_Example01 = pd.DataFrame.from_records(rows)

      [['2', '9', '5'], ['9', '1', '3']]
```

▼ Ejer 07: Subir a un repositorio de GitHub una hoja de cálculo en formato csv y realizar la importación a un dataframe, utilizando como separador de campos un "punto y coma" ;

```
✓ 0 s [9] import pandas as pd
      # Importación de los datos

      repositorio = 'https://raw.githubusercontent.com/Nestorbd/Analisis-de-sentimientos-con-Python/master/code/data.csv'


      sentences = pd.read_csv(repositorio)

      sentences
```

	Sentence	Sentiment
0	The GeoSolutions technology will leverage Bene...	positive
1	\$ESI on lows, down \$1.50 to \$2.50 BK a real po...	negative
2	For the last quarter of 2010 , Componenta 's n...	positive
3	According to the Finnish-Russian Chamber of Co...	neutral
4	The Swedish buyout firm has sold its remaining...	neutral
...
5888	In a surprising turn of events, XYZ Pharma fac...	negative
5889	GreenTech Solutions announces plans for a \$200...	positive
5890	Despite economic challenges, XYZ Construction ...	positive
5891	After extensive research and development, XYZ ...	positive
5892	Market analysts speculate a potential merger b...	neutral

5893 rows x 2 columns

Referencias

-  SNS_ACT1_1_NestorBatistaDiaz.ipynb
- [GitHub](#)
- [DataSet](#)