

**S2.A2 Práctica 3: Aplicar el proceso ETL para generar un Mapa de recomendaciones (7 pts.)**

**Objetivo:** La persona estudiante en formación deberá aplicar el proceso ETL utilizando el lenguaje de programación Python, con la finalidad de realizar un análisis con los mejores Museos y Galerías de arte ubicados en la zona urbana de la ciudad de New York; para ello, es necesario considerar un dataset proporcionado por la plataforma open data de new york, el cual contiene información real con fines de educativos y de promoción del turismo. Finalmente, se deberá desplegar una imagen de un mapa que permita visualizar los 3 museos y/o galería de arte mejor calificados ubicados en el distrito de Brooklyn.



## Cargaremos los datos del csv a un dataframe para poder realizar el ETL.

```
WORKSPACEPYTHON
env
practica3
data
new-york-city-art-galleries.csv
practica_etl.py 1

practica3 > practica_etl.py > ...
1 import pandas as pd
2
3 # Cargar el archivo CSV en un DataFrame
4 df = pd.read_csv('data/new-york-city-art-galleries.csv', encoding='latin1')
5
6 #Mostramos primero 5 filas del DataFrame Original
7 print("----- Data frame Original -----")
8 print(df.head())
9
10 #Ejercicio1: Eliminar Filas Duplicadas
11
12 #L1 to chat, MX to generate

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS
bash - practica3 + - - - - -
self_engine = self._make_engine(self_engine)
File ~/Users/pepe_gope/Desktop/workspacepython/env/lib/python3.12/site-packages/pandas/io/parsers/readers.py, line 1888, in _make_engine
self.handles = get_handle(
File ~/Users/pepe_gope/Desktop/workspacepython/env/lib/python3.12/site-packages/pandas/io/common.py, line 873, in get_handle
handle = open(
FileNotFoundError [Errno 2] No such file or directory: 'data/new-york-city-art-galleries.csv'
(env) PEPE-GOPEs-MacBook-Pro:workspacepython pepe_gope$ cd practica3/
(env) PEPE-GOPEs-MacBook-Pro:practica3 pepe_gope$ ls
data
practica_etl.py
(env) PEPE-GOPEs-MacBook-Pro:practica3 pepe_gope$ python3 practica_etl.py
0
1 O'reilly William & Co Ltd POINT (-73.96273874561996 40.773888871625926) (212) 396-1822 http://www.nyc.com/arts_attractions/oreilly_w... 52 E 76th St NaN New York 10021 4.1
2 Organization of Independent Artists - Gallery 482 POINT (-74.0893857886238 40.71646858931738) (212) 218-9213 http://www.nonprofitgallery.com/main/usa/ny/ol... 19 Hudson St NaN New York 10013 5.2
3 Own Gallery POINT (-73.96453217845256 40.714088522488) (212) 879-7415 http://www.owngallery.com/about-us 19 E 75th St NaN New York 10021 4.3
4 P P O W Gallerie POINT (-74.08389261844238 40.74953548848486) (212) 647-1844 http://www.ppowgallery.com/ 511 W 25th St NaN New York 10001 7.2
5

> ESQUEMA
> LÍNEA DE TIEMPO
MX to generate a command - MX to autocomplete
Ln 12, col 1 Espacios: 4 UTF-8 LF Python 3.12.2 64-bit Copilot++ Go Live Zero Two
```

## Consideraciones:

### 1.- Eliminar filas o registros duplicados

```
WORKSPACEPYTHON
env
practica3
data
new-york-city-art-galleries.csv
practica_etl.py 1

practica3 > practica_etl.py > ...
1 import pandas as pd
2
3 # Cargar el archivo CSV en un DataFrame
4 df = pd.read_csv('data/new-york-city-art-galleries.csv', encoding='latin1')
5
6 #Mostramos primero 5 filas del DataFrame Original
7 print("----- Data frame Original -----")
8 print(df.head())
9
10 #Ejercicio1: Eliminar Filas Duplicadas
11 print("\n----- Ejercicio 1: eliminar filas duplicadas -----\n")
12
13 # Eliminar filas duplicadas
14 df_sin_duplicados = df.drop_duplicates()
15 # Mostrar el DataFrame sin filas duplicadas
16 print(df_sin_duplicados)
17
18

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS
bash - practica3 + - - - - -
P P O W Inc POINT (-74.08176389614576 40.72298742771331) (212) 941-8642 http://www.nyc.com/arts_attractions/p_p_o_w_i... 476 Broome St NaN New York 10013 5.1
----- Ejercicio 1: eliminar filas duplicadas -----
0
1 O'reilly William & Co Ltd POINT (-73.96273874561996 40.773888871625926) (212) 396-1822 http://www.nyc.com/arts_attractions/oreilly_w... 52 E 76th St NaN New York 10021 4.1
2 Organization of Independent Artists - Gallery 482 POINT (-74.0893857886238 40.71646858931738) (212) 218-9213 http://www.nonprofitgallery.com/main/usa/ny/ol... 19 Hudson St NaN New York 10013 5.2
3 Own Gallery POINT (-73.96453217845256 40.714088522488) (212) 879-7415 http://www.owngallery.com/about-us 19 E 75th St NaN New York 10021 4.3
4 P P O W Gallerie POINT (-74.08389261844238 40.74953548848486) (212) 647-1844 http://www.ppowgallery.com/ 511 W 25th St NaN New York 10001 7.2
5 P P O W Inc POINT (-74.08176389614576 40.72298742771331) (212) 941-8642 http://www.nyc.com/arts_attractions/p_p_o_w_i... 476 Broome St NaN New York 10013 5.1
6
7 Orensis Gallery POINT (-73.9629081361554 40.7618761712667) (212) 486-7238 http://www.livelihood.com/newyork/art/orensis... 389 E 61st St NaN New York 10021 6.2
8 Old Print Shop Inc POINT (-73.98194844848489 40.74387322872153) (212) 683-9998 http://www.oldprintshop.com/ 198 Lexington Ave NaN New York 10016 6.2
9 Orlingard Art Guarantee POINT (-74.814722258884 40.70838181103074) (212) 572-8888 http://www.citydata.com/business/68395189... 115 Broadway NaN New York 10006 6.1
10 One Twenty Eight POINT (-73.9889158385955 40.719886513486134) (212) 674-8284 http://www.galleryonetwentyeight.com/ 128 Livingston St NaN New York 10002 6.1
11 Opera Gallery POINT (-73.9999484559857 40.7238149375711) (212) 966-0675 http://www.ppopanda.com/events.htm 115 Spring St NaN New York 10012 6.0
12
13 (922 rows x 9 columns)
(env) PEPE-GOPEs-MacBook-Pro:practica3 pepe_gope$

> ESQUEMA
> LÍNEA DE TIEMPO
MX to generate a command - MX to autocomplete
Ln 16, col 25 Espacios: 4 UTF-8 LF Python 3.12.2 64-bit Copilot++ Go Live Zero Two
```

## 2.- Eliminar la columna Address 2

```
workspacePython
practica3 | practica_eli.py |
1 import pandas as pd
2
3 # Cargar el archivo CSV en un DataFrame
4 df = pd.read_csv('data/new-york-city-art-galleries.csv', encoding='latin1')
5
6 # Mostramos primera 5 filas del DataFrame Original
7 print("----- Data frame Original -----")
8 print(df.head())
9
10 #Ejercicio1: Eliminar Filas Duplicadas
11 print("\n----- Ejercicio 1: eliminar filas duplicadas -----")
12
13 # Eliminar filas duplicadas
14 df_sin_duplicados = df.drop_duplicates()
15 # Mostrar el DataFrame sin filas duplicadas
16 print(df_sin_duplicados)
17
18 #Ejercicio2: Eliminar Filas Duplicadas
19 print("\n----- Ejercicio 2: eliminar Columna ADDRESS2 -----")
20
21 # Eliminamos la columna "Address2"
22 df_sin_address = df.drop(columns=['ADDRESS2'])
23
24 # mostrar el DataFrame sin la columna "Address2"
25 print(df_sin_address)
```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS

[922 rows x 9 columns]

Ejercicio 2: eliminar Columna ADDRESS2

	NAME	the.geon	TEL	URL	ADDRESS1	CITY	ZIP	GRADING
0	O'Reilly William & Co Ltd	POINT (-73.96273874561996 40.773880871637578)	(212) 396-1822	http://www.nyc.gov/arts_attractions/oreilly.w...	52 E 76th St	New York	10021	4.1
1	Organization of Independent Artists - Gallery 402	POINT (-74.063057886208 40.7164030931789)	(212) 218-9223	http://www.independentartists.com/members/ryd...	19 Hudson St	New York	10013	3.2
2	Own Gallery	POINT (-73.96453178145859 40.7488882524689)	(212) 879-2435	http://www.owngallery.com/about-us	19 E 75th St	New York	10021	4.3
3	P P O W Gallery	POINT (-74.00302323846228 40.7408538080486)	(212) 647-1864	http://www.ppowgallery.com/	511 W 25th St	New York	10004	7.2
4	P P O W INC	POINT (-74.00176389614576 40.72290742771331)	(212) 941-8642	http://www.nyc.gov/arts_attractions/p_p_o_w_i...	476 Broome St	New York	10013	5.1
...	...	...	...	...	...	...	...	...
918	Odyssey Gallery	POINT (-73.96269813635554 40.7618747512869)	(212) 486-7338	http://www.livevillage.com/newyork/art/odyssey...	385 E 61st St	New York	10021	6.2
919	Old Print Shop Inc	POINT (-73.98194848484848 40.74367327071215)	(212) 683-3908	http://www.oldprintshop.com/	159 Lexington Ave	New York	10016	6.2
920	Omnigum Art Guarantee	POINT (-74.6114722358984 40.7893881165374)	(212) 577-8880	http://www.city-data.com/business/65395189...	115 Broadway	New York	10005	6.1
921	One Twenty Eight	POINT (-73.986913483955 40.71986513486134)	(212) 674-0244	http://www.galleryonewentyeight.org/	128 Livingston St	New York	10002	6.1
922	Opera Gallery	POINT (-73.999464339657 40.7233143379721)	(212) 966-6075	http://www.popsandco.com/events.htm	133 Spring St	New York	10012	6.0

[923 rows x 8 columns]

PEPE-GUPEs-MacBook-Pro:practica3 pepe\_gupe\$

## 3.- Seleccionar a los museos o galerías de arte del distrito Brooklyn

```
workspacePython
practica3 | practica_eli.py |
12 # Eliminar filas duplicadas
13 df_sin_duplicados = df.drop_duplicates()
14 # Mostrar el DataFrame sin filas duplicadas
15 print(df_sin_duplicados)
16
17 #Ejercicio2: Eliminar columna ADDRESS2
18 print("\n----- Ejercicio 2: eliminar Columna ADDRESS2 -----")
19
20 # Eliminamos la columna "Address2"
21 df_sin_address = df.drop(columns=['ADDRESS2'])
22
23 # mostrar el DataFrame sin la columna "Address2"
24 print(df_sin_address)
25
26
27
28 #Ejercicio3: eleccionar a los musos o galerias de arte del distrito Brooklyn
29 print("\n----- Ejercicio 3: eleccionar a los museos o galerias de arte del distrito Brooklyn -----")
30
31 # Seleccionar los musos o galerias de arte del distrito de Brooklyn
32 brooklyn_df = df[df['CITY'] == "Brooklyn"]
33
34 # Mostrar los musos o galerias de arte del distrito de Brooklyn
35 print(brooklyn_df)
36
```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS

[923 rows x 8 columns]

Ejercicio 3: eleccionar a los musos o galerias de arte del distrito Brooklyn

	NAME	the.geon	TEL	URL	ADDRESS1	ADDRESS2	CITY	ZIP	GRADING
19	Parker's Box	POINT (-73.964447802016 40.714380193203568)	(718) 386-2882	http://www.parkersbox.com/canibitions/current...	183 Grand St	NaN	Brooklyn	11211	8.2
39	Picture This Corporate Art	POINT (-73.99260358101613 40.6832413941588)	(718) 625-2722	http://www.ptcny.com/	249 Smith Street	# 182	Brooklyn	11231	7.5
40	Platop 2000	POINT (-73.9531816884418 40.71724935978865)	(718) 599-2144	http://www.platop1800.com/	117 N 9th St	NaN	Brooklyn	11211	9.5
43	Plaster Palace	POINT (-73.98315835383828 40.6257592774893)	(718) 621-3175	http://www.merchancircle.com/Business/Plaster...	5385 18th Ave	NaN	Brooklyn	11284	8.1
49	Polizzi Fine Art	POINT (-74.62866287478092 40.62876225412706)	(718) 688-0777	http://www.polizzifineart.com/	7833 3rd Ave	NaN	Brooklyn	11209	7.9
...	...	...	...	...	...	...	...	...	...
871	Master Drawing Assoc	POINT (-73.92892220452346 40.6880523394711)	(718) 576-1209	http://www.masterdrawing.org/	609 Monroe St	NaN	Brooklyn	11222	5.7
880	Moons Art Gallery	POINT (-73.96293555876488 40.7115840182591)	(718) 215-8858	http://find.mapeuse.com/details/art-gallery...	5482 7th Ave	NaN	Brooklyn	11229	5.6
893	Mooney Robert J Arts Management	POINT (-73.92816147524545 40.6888922225752)	(718) 852-8089	http://local.bovw.org/New_York/Brooklyn/Mooney...	355 Bedford Ave	NaN	Brooklyn	11211	5.7
916	Object-Image	POINT (-73.97789833482246 40.67984878697147)	(718) 623-2434	http://www.laurafantini.it/Documents/For_Immed...	174 Pacific St	NaN	Brooklyn	11201	5.6
					91 5th Ave	NaN	Brooklyn	11217	5.3

[62 rows x 9 columns]

PEPE-GUPEs-MacBook-Pro:practica3 pepe\_gupe\$

#### 4.- Filtrar a los 3 mejor calificados (museo o galería)

Para mostrar los 3 mejores haremos uso del filtro anterior para poder ordenar los datos de las calificaciones de manera descendente para una mejor comprension

```
practica3 > python3 practica_etl.py > ...
27
28 #Ejercicio3: eleccionar a los museos o galerias de arte del distrito Brooklyn
29 print("\n----- Ejercicio 3: eleccionar a los museos o galerias de arte del distrito Brooklyn -----")
30 # Seleccionar los museos o galerias de arte del distrito de Brooklyn
31 brooklyn_df = df[df['CITY'] == 'Brooklyn']
32
33 # Mostrar los museos o galerias de arte del distrito de Brooklyn
34 print(brooklyn_df)
35
36
37 #Ejercicio4: Filtrar a los 3 mejor calificados (museo o galeria)
38 print("\n----- Ejercicio 4: Filtrar a los 3 mejor calificados (museo o galeria) de Brooklyn -----")
39
40 #ordenamos las calificaciones en orden descendente usando el filtro anterior
41 brooklyn_df_sorted = brooklyn_df.sort_values(by='GRADING', ascending=False)
42
43 # Filtrar los 3 mejores calificados
44 top_3_mejor_calificados = brooklyn_df_sorted.head(3)
45
46 # Mostrar los 3 mejores calificados
47 print(top_3_mejor_calificados)
```

PROBLEMAS	SALIDA	CONSOLA DE DEPURACIÓN	TERMINAL	PUESTOS
40	Pierogi 2000	POINT (-73.95573180884416 40.71875845378865)	(718) 599-2144	http://www.pierogi2000.com/ 177 N 9th St NaN Brooklyn 11211 9.5
43	Planter Palace	POINT (-73.98138533383629 40.62573597248058)	(718) 624-3175	http://www.merchantcircle.com/business/planter... 5385 18th Ave NaN Brooklyn 11204 8.1
49	Polizzi Fine Art	POINT (-74.02868202478092 40.62887623412286)	(718) 680-6777	http://www.polizzifineart.com/ 7813 3rd Ave NaN Brooklyn 11209 7.9
640	Master Drawing Assoc	POINT (-73.9292226435246 40.6889523738471)	(718) 573-1299	http://www.masterdrawings.org/ 669 Monroe St NaN Brooklyn 11221 5.7
671	Makay Art Gallery	POINT (-74.00855537930751 40.6483482223364)	(718) 686-9788	http://find-maps.com/details/art-galleries/... 5402 7th Ave NaN Brooklyn 11220 5.6
888	Moments Art Gallery	POINT (-73.96293535878489 40.7118841862991)	(718) 218-8058	http://momentart.org/ 359 Bedford Ave NaN Brooklyn 11211 5.7
916	Mooney Robert J Arts Management	POINT (-73.99361615424245 40.6988922223752)	(718) 826-9880	http://local.betw.org/New_York/Brooklyn/Mooney... 174 Pacific St NaN Brooklyn 11201 5.6
	Object-Image	POINT (-73.97789833482246 40.67984878697147)	(718) 623-2434	http://www.laursafantini.it/Documents/For_Imed... 91 5th Ave NaN Brooklyn 11217 5.3

```
[62 rows x 9 columns]
```

	NAME	the_geom	TEL	URL	ADDRESS1	ADDRESS2	CITY	ZIP	GRADING
242	Under Minerva	POINT (-73.99264054435559 40.66242444886783)	(718) 788-0170	http://blog.underminerva.com/	656 5th Ave	NaN	Brooklyn	11215	9.7
40	Pierogi 2000	POINT (-73.95573180884416 40.71875845378865)	(718) 599-2144	http://www.pierogi2000.com/	177 N 9th St	NaN	Brooklyn	11211	9.5
19	Parker's Box	POINT (-73.968447882816 40.714309136269266)	(718) 388-2882	http://www.parkersbox.com/caenhibitions/curren...	193 Grand St	NaN	Brooklyn	11211	8.2

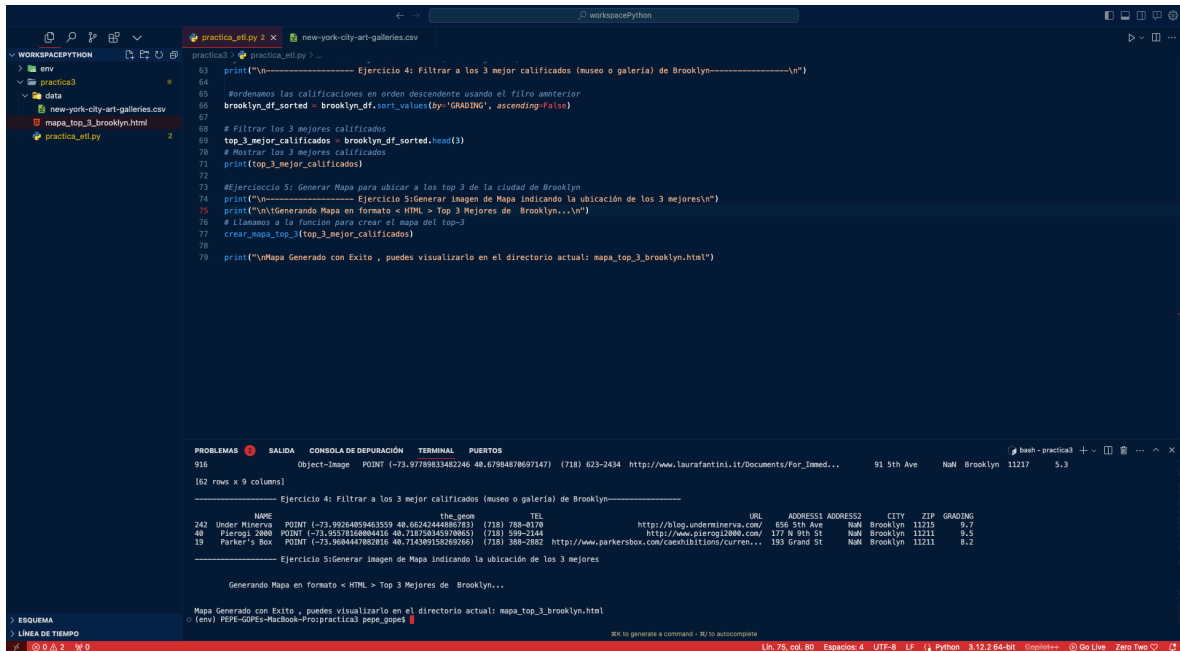
#### 5.- Generar imagen de Mapa indicando la ubicación de los 3 mejores

En este caso tendremos que hacer uso de la librería **folium** para la creacion de mapas creativos y la libreria **re** para el iuso de expresiones reulares .

Para hacer esto mas facil primero crearemos una funcion para crear el Mapa y este tenga como parametro undataframe de top\_3\_mejores y se le agregara un marcador acad uno para mejorar la visualizacion e interactividad.

```
practica3 > python3 practica_etl.py > ...
1 import pandas as pd
2 import folium
3 import re
4
5 #//////// Seccion de Funciones
6
7 # Función para crear el mapa con los marcadores de los 3 mejores museos o galerias de arte en Brooklyn usado en el ejercicio 5
8 def crear_mapa_top_3(top_3_mejor_calificados):
9     # Crear un mapa centrado en Brooklyn
10     mapa = folium.Map(location=[40.6782, -73.9442], zoom_start=12)
11
12     # Función para extraer coordenadas de la columna the_geom
13     def extraer_coordenadas(geom):
14         coords = re.findall(r"\d+\.\d+\.\d+", geom)
15         return [float(coords[1]), float(coords[0])]
16
17     # Agregar marcadores para los 3 mejores museos o galerias de arte en Brooklyn
18     for index, row in top_3_mejor_calificados.iterrows():
19         coords = extraer_coordenadas(row['the_geom'])
20         folium.Marker(coords, popup=row['NAME']).add_to(mapa)
21
22     # Guardar el mapa como un archivo HTML
23     mapa.save("mapa_top_3_brooklyn.html")
24
25 #//////// Fin de Seccion de Funciones
26
27 # Copiar el archivo CSV en un DataFrame
```

Ahora se procedera a llamar a la funcion **crear\_mapa\_top\_3()** y seejecutara el programa , esto hara quese cree el mapa en el directorio actual de la practica.



```
63 print("\n----- Ejercicio 4: Filtrar a los 3 mejor calificados (museo o galería) de Brooklyn-----\n")
64
65 #ordenamos las calificaciones en orden descendente usando el filtro anterior
66 brooklyn_df_sorted = brooklyn_df.sort_values(by="GRADING", ascending=False)
67
68 # Filtrar los 3 mejores calificados
69 top_3_mejor_calificados = brooklyn_df_sorted.head(3)
70 # Mostrar los 3 mejores calificados
71 print(top_3_mejor_calificados)
72
73 #Ejercicio 5: Generar Mapa para ubicar a los top 3 de la ciudad de Brooklyn
74 print("\n----- Ejercicio 5: Generar imagen de Mapa indicando la ubicación de los 3 mejores\n")
75 # Generar Mapa en formato < HTML > Top 3 Mejores de Brooklyn...
76 # llamamos a la función para crear el mapa del top-3
77 crear_mapa_top_3(top_3_mejor_calificados)
78
79 print("\nMapa Generado con Exito, puedes visualizarlo en el directorio actual: mapa_top_3_brooklyn.html")
```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS

916 Object-Image POINT (-73.97789833402246 40.67984878697247) (718) 623-2434 http://www.laurafantini.it/Documents/For\_Immed... 91 5th Ave NaN Brooklyn 11217 5-3

[62 rows x 9 columns]

Ejercicio 4: Filtrar a los 3 mejor calificados (museo o galería) de Brooklyn-----

	NOME	LOC	TCL	URL	ADDRESS1	ADDRESS2	CITY	ZIP	GRADING
242	Under Minerva	POINT (-73.99264859463559 40.66242444867833)	(718) 788-0170	http://blog.underminerva.com/	656 5th Ave	NaN	Brooklyn	11215	9-7
48	Pierogi 2000	POINT (-73.9578118884416 40.718758345978865)	(718) 599-2144	http://www.pierogi2000.com/	177 N 9th St	NaN	Brooklyn	11211	9-5
19	Parker's Box	POINT (-73.984447983030 40.72489915292960)	(718) 388-2882	http://www.parkersbox.com/exhibitions/current...	103 Grand St	NaN	Brooklyn	11211	8-2

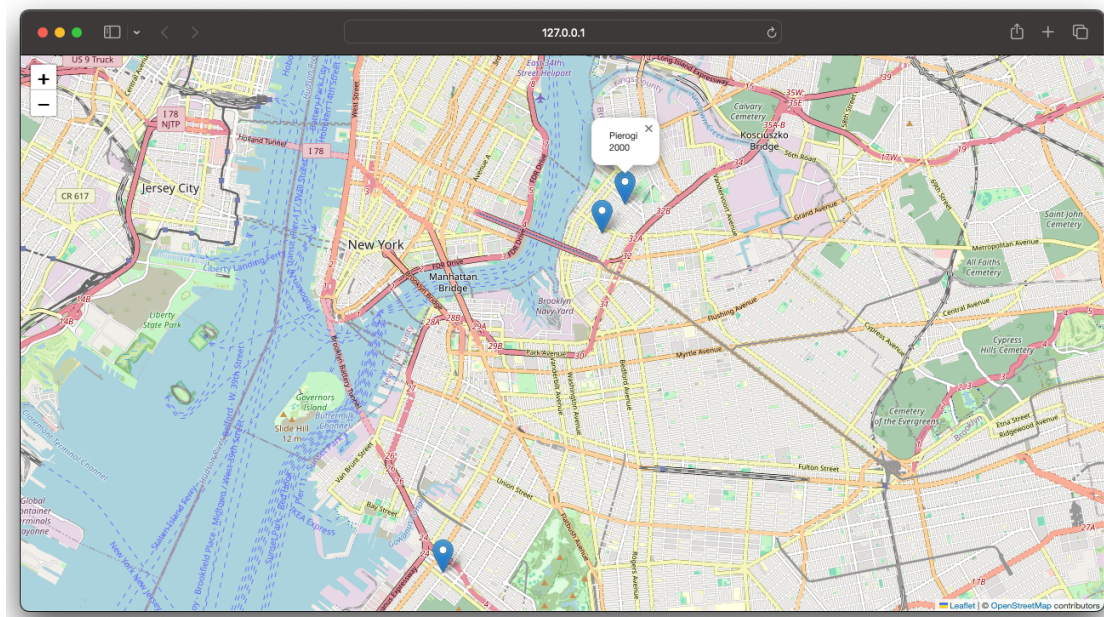
Ejercicio 5: Generar Mapa en formato < HTML > Top 3 Mejores de Brooklyn...

Generando Mapa en formato < HTML > Top 3 Mejores de Brooklyn...

Mapa Generado con Exito, puedes visualizarlo en el directorio actual: mapa\_top\_3\_brooklyn.html

envi PEPE-GOPES-MacBook-Pro[practica3] pepe\_gopes

Mapa ejecutado en un Servidor local:



6.- Generar el reporte final en un archivo de tipo PDF

En este ejercicios utilizaremos la librería de reportlab para la creacion de imágenes y el pdf que se utilizaran.

```
from reportlab.platypus import SimpleDocTemplate, Paragraph, Table,
TableStyle, Image
from reportlab.lib.pagesizes import letter
from reportlab.lib.styles import import getSampleStyleSheet
```

Creamos una funcion generar\_infrme\_pdf para crear el reporte que solicite como parametro un dataframe y mapa en este caso seran los antes creados en el ejerccion5.

```
practica5 > practica_01.py > new-york-city-art-galleries.csv

27 ##### Funcion para crear el Reporte PDF
28
29 # Función para generar el informe en PDF
30 def generar_informe_pdf(top_3_mejor_calificados, mapa_html):
31     # Crear un objeto PDF
32     doc = SimpleDocTemplate("informe_mejores_museos_galerias.pdf", pagesize=letter)
33
34     # Contenido del informe
35     contenido = []
36
37     # Estilo del párrafo
38     styles = getSampleStyleSheet()
39     estilo_titulo = styles["Title"]
40     estilo_normal = styles["Normal"]
41
42     # Título del informe
43     titulo = Paragraph("Mejores Museos y Galerías de Brooklyn", estilo_titulo)
44     contenido.append(titulo)
45
46     # Agregar el mapa
47     contenido.append(Image(mapa_html, width=500, height=500))
48
49     # Encabezados de la tabla
50     encabezados = [("Nombre", "Dirección", "Grados")]
51
52     # Datos para la tabla
53     datos_tabla = [(row['NAME'], row['ADDRESS1'], f'{row["GRADING"]}') for index, row in top_3_mejor_calificados.iterrows()]
54
55     # Crear la tabla
56     tabla = Table(encabezados + datos_tabla)
57
58     # Estilo de la tabla
59     estilo_tabla = TableStyle([('BACKGROUND', (0, 0), (-1, 0), (0.8, 0.8, 0.8)),
60                               ('TEXTCOLOR', (0, 0), (-1, 0), (0, 0, 0)),
61                               ('ALIGN', (0, 0), (-1, -1), 'CENTER'),
62                               ('FONTNAME', (0, 0), (-1, 0), 'Helvetica-Bold'),
63                               ('BOTTOMPADDING', (0, 0), (-1, 0), 12),
64                               ('BACKGROUND', (0, 1), (-1, -1), (0.8, 0.8, 0.8)),
65                               ('GRID', (0, 0), (-1, -1), 1, (0, 0, 0))])
66
67     tabla.setStyle(estilo_tabla)
68
69     # Agregar la tabla al contenido del informe
70     contenido.append(tabla)
71
72     # Generar el PDF
73     doc.build(contenido)
74
75     ##### Fin de Sección de Funciones
```

En el proceso se tomara el mapa en formato HTML para después convertirlo a una **PNG** o **JPG** para poder ingresarlo a la función `generar_informe_pdf()`, casi lo olvido, para poder realizar la conversión se tiene que usar la librería **imgKit**.

Ahora continuaremos con la llamada del método para poder generar el reporte pdf con las características solicitadas y se guardara en la ruta del directorio actual de la práctica.



```
131
132
133 #Ejercicio 6: Generar el reporte final en un archivo de tipo PDF con un Titulo , etiqueta nombre: Address y grados..
134 print("\n----- Ejercicio 6:Generar reporte final en formato pdf del mapa top_3 \n")
135 print("\n\tGenerando Reporte Final en PDF --> Espera...\n")
136
137 #Convertimos el mapa de top3.html en formato PNG
138 print("\t\tConvirtiendo ...")
139
140 # Convertir el archivo HTML a imagen usando imgkit
141 #mapa_html = 'mapa_top_3_brooklyn.html'
142 salida_mapa_png = 'mapa_top_3_brooklyn.png'
143 #imgkit.from_file (mapa_html, salida_mapa_png)
144
145 print("[mapa_top_3_brooklyn.png ] \n Completado...")
146
147 # Llamamos a la funcion para generar el Informe PDF
148 generar_informe_pdf(top_3_mejor_calificados, salida_mapa_png)
149
150 print("\n\tInforme Generado Exitosamente: Informe_mejores_museos_galerias.pdf\n")
151
152
```

PROBLEMAS SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS

Ejercicio 4: Filtrar a los 3 mejor calificados (museo o galeria) de Brooklyn

	NAME	POINT	the_geom	TEL	URL	ADDRESS1	ADDRESS2	CITY	ZIP	GRADING
242	Under Minerva	(-73.99264859463559 40.66242444898783)	(718) 788-0170	http://blog.underminerva.com/	656 5th Ave	NaN	Brooklyn	11215	9.7	
40	Pierogi 2000	(-73.95578160804416 40.718758345978063)	(718) 599-2144	http://www.pierogi2000.com/	177 N 9th St	NaN	Brooklyn	11211	9.5	
19	Parker's Box	(-73.9604447002816 40.714389150269266)	(718) 380-2882	http://www.parkersbox.com/caseinstitutions/curren...	193 Grand St	NaN	Brooklyn	11211	8.2	

Ejercicio 5:Generar imagen de Mapa indicando la ubicación de los 3 mejores

Generando Mapa en formato < HTML > Top 3 Mejores de Brooklyn...

Mapa Generado con Exito , puedes visualizarlo en el directorio actual: mapa\_top\_3\_brooklyn.html

Ejercicio 6:Generar reporte final en formato pdf del mapa top\_3

Generando Reporte Final en PDF --> Espera...

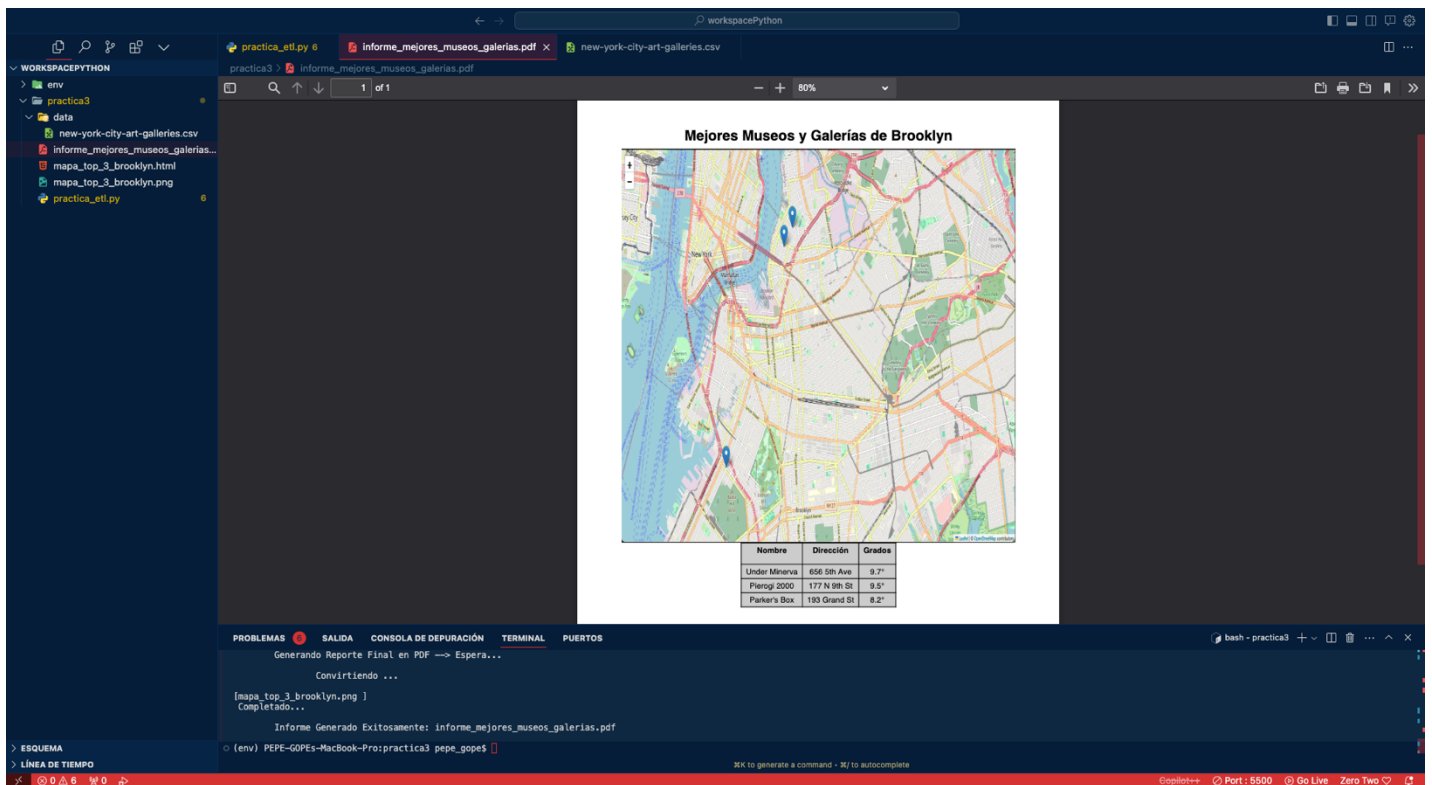
Convirtiendo ...

[mapa\_top\_3\_brooklyn.png ]  
Completado...

Informe Generado Exitosamente: Informe\_mejores\_museos\_galerias.pdf

(env) PEPE-GOPES-MacBook-Pro:practica3 pepe\_gope\$

Aqui se muestra como se ve el pdf generado :



En conclusion en esta práctica no3, hemos aprendido a manejar datos utilizando Python, realizando tareas de limpieza, transformación y visualización (ETL). Utilizamos pandas para manipular los datos, folium para crear mapas interactivos y reportlab para generar informes en PDF y la librería de imgkit para conversion para pfromatos PNG o JPG .Esto nos ha permitido obtener información significativa de los datos y presentarla de manera clara y efectiva mediante un dataframe, espero que sea de ayuda y comprsnible esta reporte.

El cododigo se encontrara en mi repositorio: por motivos de fallos tecnicos no podre subirloa en este momeonto , favor de esperar la actualizacion en la siguiente practica. Dejare el link del Repo para que puedan descargar el codigo:

<https://github.com/pepe1603/repo-BigData/tree/main/practica3>

Tutoriales y/o referencias de consulta:

- Folium Map : <https://medium.com/@sawsanyusuf/data-visualization-with-python-9-generating-maps-with-folium-e7563bc1a923>
- Static Map Project : <https://github.com/komoot/staticmap>
- Py StaticMap: <https://github.com/flopp/py-staticmaps>