

▼ Actividad - Librería pandas

- **Nombre:** Ramona Nájera Fuentes
- **Matrícula:** A01423596

```
import pandas as pd
import numpy as np
```

Entregar: Archivo PDF de la actividad, así como el archivo .ipynb en tu repositorio. **Nota:** Recuerda habrá una penalización de 50 puntos si la actividad fue entregada fuera de la fecha límite.

1. Subir el archivo `netflix_titles.csv` que se encuentra en el repositorio de la clase y cargarlos en un `DataFrame`.

```
from google.colab import files

uploaded = files.upload()

for fn in uploaded.keys():
    print('User uploaded file "{name}" with length {length} bytes'.format(
        name=fn, length=len(uploaded[fn])))

DataFame = pd.read_csv('netflix_titles.csv')
```

2. Desplegar las primeras 5 líneas del `DataFrame`.

```
DataFrame.head(5)
```

	show_id	type	title	director	cast	country	date_added	release_year
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	2020
1	s2	Movie	7:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	2016
2	s3	Movie	23:59	Gilbert Chan	Tedd Chan, Stella Chung, Henley Hii, Lawrence	Singapore	December 20, 2018	2011

3. Implementa una función que despliegue el primer renglon de un `DataFrame`.

```
def the_first(df):
    return df.head(1)

the_first(DataFrame)
```

	show_id	type	title	director	cast	country	date_added	release_year
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato,	Brazil	August 14, 2020	2020

4. Implementa una función que regrese el número de registros en el DataFrame .

```
def row_number(df):
    return len(df)

row_number(DataFrame)

7787
```

5. Implementa una función que despliega las películas con un release_year determinado.

```
def released_year(year, df):
    return df[df.release_year == year]

year_query = 2020
released_year(year_query, DataFrame)
```

	show_id	type	title	director	cast	country	date_added	release_
0	s1	TV Show	3%	NaN	João Miguel, Bianca Comparato, Michel Gomes, R...	Brazil	August 14, 2020	
24	s25	TV Show	SAINT SEIYA: Knights of the Zodiac	NaN	Bryson Baugus, Emily Neves, Blake Shepard, Pat...	Japan	January 23, 2020	
26	s27	TV Show	(Un)Well	NaN	NaN	United States	August 12, 2020	
27	s28	Movie	#Alive	Cho Il	Yoo Ah-in, Park Shin-hye	South Korea	September 8, 2020	
29	s30	TV Show	#blackAF	NaN	Kenya Barris, Rashida Jones, Iman Benson, Genn...	United States	April 17, 2020	
...
7725	s7726	TV Show	You Cannot Hide	NaN	Blanca Soto, Eduardo Noriega, Iván Sánchez, Ma... Mauricio	Mexico	October 1, 2020	

6. Implementa una función que despliegue las películas que son de un país determinado.

```
def origin_country(country, df):
    return df[df.country == country]

country_query = 'Mexico'
origin_country(country_query, DataFrame)
```

	show_id	type	title	director	cast	country	date_added	release_year
1	s2	Movie	7:19	Jorge Michel Grau	Demián Bichir, Héctor Bonilla, Oscar Serrano, ...	Mexico	December 23, 2016	
12	s13	TV Show	1994	Diego Enrique Osorno	NaN	Mexico	May 17, 2019	
292	s293	Movie	Acapulco La vida va	Alfonso Serrano Maturino	Patricio Castillo, Sergio Bustamante, Alejandr...	Mexico	June 3, 2017	
370	s371	Movie	Alan Saldaña: Mi vida de pobre	Raúl Campos, Jan Suter	Alan Saldaña	Mexico	August 4, 2017	
378	s379	Movie	Alex Fernández: The Best Comedian in the World	Alex Díaz	Alex Fernández	Mexico	January 23, 2020	
...
7698	s7699	TV Show	Yankee	NaN	Pablo Lyle, Ana Layevska, Leonardo Daniel, Jav...	Mexico	June 14, 2019	
7725	s7726	TV Show	You Cannot Hide	NaN	Blanca Soto, Eduardo Noriega, Iván Sánchez, Ma... Mauricio	Mexico	October 1, 2020	

7. Crea un nuevo DataFrame que sólo contenga las columnas type, title, country y release_year.

```
dfAux = pd.read_csv('netflix_titles.csv', sep=',', usecols=['type', 'title', 'country', 'release_year'])
dfAux
```

	type	title	country	release_year
0	TV Show	3%	Brazil	2020
1	Movie	7:19	Mexico	2016
2	Movie	23:59	Singapore	2011
3	Movie	9	United States	2009
4	Movie	21	United States	2008
...
7782	Movie	Zozo	Sweden, Czech Republic, United Kingdom, Denmar...	2005
7783	Movie	Zubaan	India	2015
7784	Movie	Zulu Man in Japan	NaN	2019
7785	TV Show	Zumbo's Just Desserts	Australia	2019
...	...	ZZ TOP: THAT LITTLE OL'	United Kinodome. Canada.	...

8. Despliega, cuantos tipos de producciones (type) se hicieron por país determinado (country).

```
filtro = dfAux.groupby(['country', 'type']).count()
filtro
```

		title	release_year	
	country	type		
	Argentina	Movie	34	34
		TV Show	16	16
	Argentina, Brazil, France, Poland, Germany, Denmark	Movie	1	1
	Argentina, Chile	Movie	1	1
	Argentina, Chile, Peru	Movie	1	1

	Venezuela	Movie	1	1
	Venezuela, Colombia	Movie	1	1
	Vietnam	Movie	5	5
	West Germany	Movie	1	1
	Zimbabwe	Movie	1	1

773 rows × 2 columns

9. Implementa la función que genera el cuadrado de un número.

```
def square(x):
    return x*x
```

10. Usando el método `apply`, genera el cuadrado de todos los valores del siguiente `DataFrame`.

```
df2 = pd.DataFrame({'A': [1, 2, 3, 5, 6], 'B': [10, 20, 30, 40, 50]})
df2
```

	A	B
0	1	10
1	2	20
2	3	30
3	5	40
4	6	50

```
df2_2 = df2.apply(square, axis=0)
df2_2
```

	A	B
0	1	100
1	4	400
2	9	900
3	25	1600
4	36	2500

