

Actividad - Librería pandas

- **Nombre: Oscar Eduardo Nieto Espitia **
- **Matrícula: A01705090**

```
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])))
```

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

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

```
# El método `head` nos permite desplegar las primeras línea de un DataFrame.
df.head(5)
```

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

```
def the_first():  
    return df[0:1]  
  
the_first()
```

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

```
def row_number():  
    a = len(df)  
    return a  
  
row_number()  
  
7787
```

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

```
def released_year(year):  
  
    return df[df['release_year'].isin([year])]
```

```
year_query = 2020  
released_year(year_query)
```

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

```
def origin_country(country):
```

```
return df[df['country'].isin([country])]
```

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

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

```
df2 = pd.DataFrame({'type': pd.Categorical(["Movie", "TVShow", "test", "train"]),
                    'title': pd.Series("hola", index=list(range(4)), dtype='string'),
                    'country': pd.Series("adios", index=list(range(4)), dtype='string'),
                    'release_year': pd.Timestamp('20200101')
                    })
```

df2

	type	title	country	release_year	
0	Movie	hola	adios	2020-01-01	
1	TVShow	hola	adios	2020-01-01	
2	test	hola	adios	2020-01-01	
3	train	hola	adios	2020-01-01	

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

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

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

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
```

```
df2.applymap(np.square)
```

✓ 0 s completado a las 12:34

