

Ejercicios: Clase 5

Descargue el archivo `practice_netflix_titles.zip` desde el directorio con lo contenidos de la clase. Descomprimir el `.zip`, obteniendo el archivo en su interior, `netflix_titles.csv`.

Ejecute una notebook de Jupyter, donde pueda utilizar el kernel de Python de su entorno virtual de trabajo.

1. Cargue el contenido del archivo en un `DataFrame`.
2. Procese el `DataFrame` de manera tal de producir, a partir de la columna `country`, una nueva columna con un país por fila.
Deberá obtener un nuevo `DataFrame` de mayor longitud que el inicial. Para esto podrá utilizar los métodos `explode` y `apply`.
3. Vea cuántos títulos de cada país figuran en este. Mostrar en una `Series` cuáles son los cinco países con mayor cantidad de títulos.
4. Obtenga en un `DataFrame` el nombre de cada uno de los países que figuran, una columna mostrando cuándo se registró el título más viejo, y otra mostrando cuándo se registró el más reciente.
Para esto deberá utilizar la columna `date_added`, pudiendo convertir los valores de la misma a tipo `datetime` mediante la función `to_datetime`.
5. Genere un `DataFrame` que muestre una columna para el país, una columna para los ratings, y otra con la cantidad de títulos por rating y por país.
6. A partir de `DataFrame` anterior, derive otro donde pueda verse el rating con mayor cantidad de ocurrencias por país.
7. Cree dos archivos, uno en formato `.json` y otro en formato `.csv`, con los títulos de Francia estrenados en el año 2020.
8. Genere un `DataFrame` donde como columnas estén los cinco países con mayor cantidad de títulos registrados, como filas los últimos 10 años (que figuran en el dataset inicial), y mostrando como valores la cantidad de títulos por año y por país. Los valores deberán ser números enteros.
9. Investigue cómo puede generar gráficos a través del método `.plot`. Genere un gráfico dónde pueda verse la evolución de la cantidad de títulos por año de salida.