

mcpp_taller8_Jhon_caro

October 24, 2016

1 Taller 8

Métodos Computacionales para Políticas Públicas - UROSARIO

Entrega: viernes 21-oct-2016 11:59 PM

[Jhon Alexander Caro Becerra] [Jhonalexbc@gmail.com]

1.1 Instrucciones:

- Guarde una copia de este *Jupyter Notebook* en su computador, idealmente en una carpeta destinada al material del curso.
- Modifique el nombre del archivo del *notebook*, agregando al final un guión inferior y su nombre y apellido, separados estos últimos por otro guión inferior. Por ejemplo, mi *notebook* se llamaría: mcpp_taller8_santiago_mataallana
- Marque el *notebook* con su nombre y e-mail en el bloque verde arriba. Reemplace el texto “[Su nombre acá]” con su nombre y apellido. Similar para su e-mail.
- Desarrolle la totalidad del taller sobre este *notebook*, insertando las celdas que sea necesario debajo de cada pregunta. Haga buen uso de las celdas para código y de las celdas tipo *markdown* según el caso.
- Recuerde salvar periódicamente sus avances.
- Cuando termine el taller:
 1. Descárguelo en PDF. Si tiene algún problema con la conversión, descárguelo en HTML.
 2. Suba todos los archivos a su repositorio en GitHub, en una carpeta destinada exclusivamente para este taller, antes de la fecha y hora límites.

Vamos a hacer “scraping” a esta página: <http://listas.20minutos.es/lista/500-mejores-peliculas-de-la-historia-segun-la-revista-empire-297804/>, que contiene un listado de las 500 mejores películas de la historia según la revista Empire.

Su tarea consiste en crear un “Pandas dataframe” que contenga 500 filas (una por película) y las siguientes columnas: - Ránking (#) - Título - Año - Director - Reparto - Sinopsis
Díviértase.

```
In [1]: from urllib.request import urlopen
        from bs4 import BeautifulSoup
        import re
        import pandas as pd
        import numpy as np
```

```

In [2]: base_url = "http://listas.20minutos.es/lista/500-mejores-peliculas-de-la-hi
In [3]: index_html = urlopen(base_url)
In [4]: index = BeautifulSoup(index_html, "lxml")
In [5]: info1 = index.find_all ("h3")
In [6]: mstring = "".join(map(str,info1))
In [7]: title = re.findall("</strong>(.*?)</h3>",mstring)
In [8]: ranking = re.findall("<strong>([^.].*?)</strong>",mstring)
In [9]: info2 = index.find_all("div", {"class": "descripcion_elemento"})
In [10]: mstring2 = "".join(map(str,info2))
In [11]: fecha =re.findall("[0-9]+\r|[0-9]+<",mstring2)
        fechas = []
        for i in range(len(fecha)):
            fecha2 = str(fecha[i])
            fechas.append(fecha2.rstrip("\r"))
        fecha = fechas

In [15]: info3 = index.find_all("div", {"class": "descripcion_elemento_larga"})
In [16]: mstring3 = "".join(map(str,info3))
In [17]: director =re.findall("Director:(.*?)\r|Directores:(.*?)\r|Director: (.*?)<
In [19]: reparto =re.findall("Reparto:(.*?)<br/>|Reparto: (.*?)<br/>|Reparto:\ (.*?)
In [41]: reparto.append("Åke Grönberg, Harriet Andersson, Hasse Ekman, Gudrun Brost
In [21]: sinopsis =re.findall("Sinopsis:(.*?)</div>|Sinopsis:(.*?) </div>|Sinopsis:
In [79]: movie = {"Ranking":ranking, "Título Pelicula":title, "Año":fecha, "Directo
pd.DataFrame(movie)

```

```

Out[79]:
      Año
0  1994\r
1  1997\r
2  2003\r
3  2000\r
4  1972\r
5  1994\r
6  2008\r
7  1994\r
8  1995\r
9  1993\r

Director Ranking \
0  ( Robert Zemeckis, , ) 1.
1  ( James Cameron, , ) 2.
2  ( Peter Jackson, , ) 3.
3  ( Ridley Scott, , ) 4.
4  ( Francis Ford Coppola, , ) 5.
5  (, Roger Allers y Rob Minkoff, ) 6.
6  ( Christopher Nolan, , ) 7.
7  ( Frank Darabont, , ) 8.
8  ( Mel Gibson, , ) 9.
9  ( Steven Spielberg, , ) 10.

```

10	2003\r	(Gore Verbinski, ,)	11.
11	1995\r	(John Lasseter, ,)	12.
12	2001\r	(Peter Jackson , ,)	13.
13	1990\r	(Tim Burton, ,)	14.
14	1998\r	(Steven Spielberg, ,)	15.
15	1985\r	(Robert Zemeckis, ,)	16.
16	2001\r	(Pete Docter, ,)	17.
17	2002\r	(Peter Jackson, ,)	18.
18	2003\r	(Andrew Stanton, ,)	19.
19	2004\r	(Alfonso Cuarón, ,)	20.
20	1998\r	(Tony Kaye, ,)	21.
21	2007\r	(Zack Snyder, ,)	22.
22	1999\r	(M. Night Shyamalan, ,)	23.
23	2005\r	(James McTeigue, ,)	24.
24	1994\r	(Quentin Tarantino, ,)	25.
25	1991\r	(Jonathan Demme, ,)	26.
26	1999\r	(David Fincher, ,)	27.
27	1976\r	(John G. Avildsen, ,)	28.
28	1999\r	(, Andy y Lana Wachowski,)	29.
29	1982\r	(Steven Spielberg, ,)	30.
..
469	1981\r	(Ermanno Olmi, ,)	470.
470	1978\r	(Hal Ashby, ,)	471.
471	1971\r	(Jean-Pierre Melville, ,)	472.
472	1962\r	(Sergio Corbucci, ,)	473.
473	1968\r	(Preston Sturges, ,)	474.
474	1948\r	(John Dahl, ,)	475.
475	1994\r	(Alexander Mackendrick, ,)	476.
476	1957\r	(Marcel Carné, ,)	477.
477	1938\r	(John Schlesinger, ,)	478.
478	1971\r	(Robert Bresson, ,)	479.
479	1966\r	(Todd Haynes, ,)	480.
480	1995\r	(Russ Meyer, ,)	481.
481	1970\r	(Jean Renoir, ,)	482.
482	1936\r	(, , Neil LaBute)	483.
483	1997<	(Robert Bresson, ,)	484.
484	1956\r	(Terence Davies, ,)	485.
485	1988\r	(Roy Andersson, ,)	486.
486	2000\r	(Bruce Robinson, ,)	487.
487	1987\r	(James William Guercio, ,)	488.
488	1973\r	(Roger Michell, ,)	489.
489	2004\r	(H.C. Potter, ,)	490.
490	1941\r	(James Foley, ,)	491.
491	1992\r	(Mario Bava, ,)	492.
492	1968\r	(Charles Burnett, ,)	493.
493	1977\r	(, Donald Cammell y Nic Roeg,)	494.
494	1970\r	(, Michael Powell y Emeric Pressburger,)	495.
495	1944\r	(Mu Fei, ,)	496.

496	1948\r	(Lukas Moodyson, ,)	497.
497	2000\r	(Ingmar Bergman, ,)	498.
498	1953\r	Peter Jackson	499.

Reparto \

0	(Tom Hanks, Robin Wright Penn, Gary Sinise, My...
1	(Leonardo DiCaprio, Kate Winslet, Billy Zane, ...
2	(Elijah Wood, Viggo Mortensen, Ian McKellen, S...
3	(Russell Crowe, Joaquin Phoenix, Connie Nielse...
4	(Marlon Brando, Al Pacino, James Caan, Robert ...
5	(Animada, , ,)
6	(Christian Bale, Heath Ledger, Aaron Eckhart, ...
7	(Tim Robbins, Morgan Freeman, Bob Gunton, Jame...
8	(Mel Gibson, Sophie Marceau, Patrick McGoohan,...
9	(Liam Neeson, Ben Kingsley, Ralph Fiennes, Car...
10	(Johnny Depp, Geoffrey Rush, Orlando Bloom, Ke...
11	(Animada, , ,)
12	(Elijah Wood, Ian McKellen, Viggo Mortensen, S...
13	(Johnny Depp, Winona Ryder, Dianne Wiest, Anth...
14	(Tom Hanks, Tom Sizemore, Edward Burns, Barry ...
15	(Michael J. Fox, Christopher Lloyd, Lea Thomps...
16	(Animada, , ,)
17	(Elijah Wood, Viggo Mortensen, Ian McKellen, S...
18	(Animada, , ,)
19	(Daniel Radcliffe, Rupert Grint, Emma Watson, ...
20	(Edward Norton, Edward Furlong, Fairuza Balk, ...
21	(Gerard Butler, Lena Headey, David Wenham, Dom...
22	(Bruce Willis, Toni Collette, Haley Joel Osmen...
23	(Natalie Portman, Hugo Weaving, Stephen Rea, J...
24	(John Travolta, Samuel L. Jackson, Uma Thurman...
25	(Jodie Foster, Anthony Hopkins, Scott Glenn, T...
26	(Brad Pitt, Edward Norton, Helena Bonham Carte...
27	(Sylvester Stallone, Talia Shire, Burt Young, ...
28	(Keanu Reeves, Laurence Fishburne, Carrie-Anne...
29	(Henry Thomas, Dee Wallace, Robert MacNaughton...
..	...
469	(Jean Gabin, Michel Simon, Michèle Morgan, Pie...
470	(Glenda Jackson, Peter Finch, Murray Head, Peg...
471	(Anne Wiazemsky, Walter Green, François Lefarg...
472	(Julianne Moore, Xander Berkeley, Dean North, ...
473	(Dolly Read, Cynthia Myers, Marcia McBroom, Jo...
474	(Sylvia Bataille, Georges Darnoux, Jane Marken...
475	(François Leterrier, Charles Le Clainche, Maur...
476	(Freda Dowie, Pete Postlethwaite, Angela Walsh...
477	(Lars Nordh, Stefan Larsson, Bengt C.W. Carlss...
478	(Richard E. Grant, Paul McGann, Richard Griffi...
479	(Robert Blake, Billy Green Bush, Mitch Ryan, J...
480	(Daniel Craig, Samantha Morton, Rhys Ifans, Bi...

481 (Ole Olsen, Chic Johnson, Martha Raye, Hugh He...
 482 (Al Pacino, Ed Harris, Jack Lemmon, Alec Baldw...
 483 (John Phillip Law, Marisa Mell, Michel Piccoli...
 484 (Henry G. Sanders, Kaycee Moore, Charles Bracy...
 485 (James Fox, Mick Jagger, Anita Pallenberg, Mic...
 486 (Eric Portman, Sheila Sim, John Sweet, Dennis ...
 487 (Wei Wei, Li Wei, Shi Yu, Zhang Hongmei, Cui C...
 488 (Lisa Lindgren, Michael Nyqvist, Emma Samuelss...
 489 (Åke Grönberg, Harriet Andersson, Hasse Ekman,...
 490 Ole Olsen, Chic Johnson, Martha Raye, Hugh Her...
 491 Al Pacino, Ed Harris, Jack Lemmon, Alec Baldwi...
 492 John Phillip Law, Marisa Mell, Michel Piccoli,...
 493 Henry G. Sanders, Kaycee Moore, Charles Bracy,...
 494 James Fox, Mick Jagger, Anita Pallenberg, Mich...
 495 Eric Portman, Sheila Sim, John Sweet, Dennis P...
 496 Wei Wei, Li Wei, Shi Yu, Zhang Hongmei, Cui Ch...
 497 Lisa Lindgren, Michael Nyqvist, Emma Samuelss...
 498 Åke Grönberg, Harriet Andersson, Hasse Ekman, ...

Sinopsis \

0 (Forrest Gump (Tom Hanks) es un chico que sufr...
 1 (Jack (Leonardo DiCaprio) es un joven artista ...
 2 (Las fuerzas de Saruman han sido destruidas, y...
 3 (En el año 180, el Imperio Romano domina todo ...
 4 (Años 40. Don Vito Corleone es el jefe de una ...
 5 (La sabana africana es el escenario en el que ...
 6 (Batman (Christian Bale) regresa para continua...
 7 (Acusado del asesinato de su mujer, Andrew Duf...
 8 (En el siglo XIV, los escoceses viven oprimido...
 9 (Oskar Schindler (Liam Neeson), un hombre de e...
 10 (Mar Caribe, siglo XVIII. El aventurero capitá...
 11 (Los juguetes de Andy, un niño de 6 años, teme...
 12 (En la Tierra Media, el Señor Oscuro Saurón cr...
 13 (Durante una noche de Navidad, una anciana le ...
 14 (Durante la invasión de Normandía, en plena Se...
 15 (El adolescente Marty McFly es amigo de Doc, u...
 16 (Monsters Inc. es la mayor empresa de miedo de...
 17 (La Compañía del Anillo se ha disuelto. El por...
 18 (El pececillo Nemo, que es hijo único, es muy ...
 19 (Cuando Harry Potter y sus amigos vuelven a Ho...
 20 (Derek (Edward Norton) es un joven skin head c...
 21 (Adaptación del cómic de Frank Miller (autor d...
 22 (El Dr. Malcom Crowe es un conocido psicólogo ...
 23 (En un futuro no muy lejano, Gran Bretaña es u...
 24 (Jules y Vincent son dos asesinos a sueldo que...
 25 (El FBI busca a un asesino en serie, apodado B...
 26 (Jack es un personaje insomne, hastiado de su ...
 27 (Rocky Balboa es un desconocido boxeador a qui...

28 (Thomas Anderson es un brillante programador d...
 29 (Un pequeño ser de otro planeta se queda aband...

 469 (Mientras se producen una serie de hechos que ...
 470 (Producida por la productora de George Harris...
 471 (John Wintergreen es un policía de carretera ...
 472 (En un bonito día sin nubes una joven pareja e...
 473 (Dos estrellas del vodevil están rodando una p...
 474 (Narra las ajetradas vidas de unos agentes inm...
 475 (Diabolik (John Phillip Law), atractivo y educ...
 476 (Retrato de la vida de un hombre común, un tra...
 477 (Chas es un performer de la mafia, un matón a ...
 478 (Tres "peregrinos" se encuentran durante la gu...
 479 (En un pequeño pueblo de China, una mujer, Yuw...
 480 (Un grupo de jóvenes viven bajo el mismo techo...
 481 (Albert, el propietario de un circo, abandona ...
 482 Narra la historia de una familia que pasa un d...
 483 Cínico y despiadado retrato de dos machistas r...
 484 Resistencia francés André Devigny es encarcela...
 485 La acción se desarrolla en Liverpool en el sen...
 486 Mientras se producen una serie de hechos que p...
 487 Producida por la productora de George Harrison...
 488 ohn Wintergreen es un policía de carretera de ...
 489 En un bonito día sin nubes una joven pareja es...
 490 Dos estrellas del vodevil están rodando una pe...
 491 Narra las ajetradas vidas de unos agentes inmo...
 492 Diabolik (John Phillip Law), atractivo y educa...
 493 Retrato de la vida de un hombre común, un trab...
 494 Chas es un performer de la mafia, un matón a s...
 495 Tres peregrinos se encuentran durante la guerr...
 496 En un pequeño pueblo de China, una mujer, Yuwe...
 497 Un grupo de jóvenes viven bajo el mismo techo ...
 498 Albert, el propietario de un circo, abandona a...

	Título Pelicula
0	Forrest Gump
1	Titanic
2	El Señor de los anillos: El retorno del rey
3	Gladiator
4	El Padrino
5	El rey león
6	El caballero oscuro
7	Cadena perpetua
8	Braveheart
9	La lista de Schindler
10	Piratas del Caribe: La maldición de la Perla ...
11	Toy Story
12	El Señor de los anillos: La comunidad del anillo

13	Eduardo Manostijeras
14	Salvar al soldado Ryan
15	Regreso al futuro
16	Monstruos, S.A.
17	El Señor de los anillos: Las dos torres
18	Buscando a Nemo
19	Harry Potter y el Prisionero de Azkaban
20	American History X
21	300
22	El sexto sentido
23	V de Vendetta
24	Pulp Fiction
25	El silencio de los corderos
26	El club de la lucha
27	Rocky
28	Matrix
29	E.T.
..	...
469	Impacto
470	El árbol de los zuecos
471	Harold and Maude
472	El confidente
473	El gran silencio
474	Infielmente tuyo
475	La última seducción
476	Chantaje en Broadway
477	El muelle de las brumas
478	Domingo, maldito domingo
479	Al azar de Baltasar
480	Safe
481	El valle de los placeres
482	Una partida de campo
483	En compañía de hombres
484	A Man Escaped
485	Voces distantes
486	Canciones del segundo piso
487	Withnail And I
488	La piel en el asfalto
489	El intruso
490	Loquilandia
491	Éxito a cualquier precio
492	Danger: Diabolik
493	Killer of Sheep
494	Performance
495	Un cuento de Canterbury
496	Spring in a Small Town
497	Juntos
498	Noche de circo

[499 rows x 6 columns]
