

1.)

- Wikipedia lists (actors & films)

Python script that I can run myself that would gather actors and/or movies from a wikipedia url. Here are some example lists:

Film lists:

https://en.wikipedia.org/wiki/Film_series
https://en.wikipedia.org/wiki/List_of_films_based_on_Marvel_Comics
https://en.wikipedia.org/wiki/List_of_films_based_on_DC_Comics
https://en.wikipedia.org/wiki/List_of_films_based_on_video_games
https://en.wikipedia.org/wiki/List_of_films_based_on_comics
https://en.wikipedia.org/wiki/List_of_films_based_on_television_programs
https://en.wikipedia.org/wiki/List_of_films_based_on_manga
[https://en.wikipedia.org/wiki/List_of_film_remakes_\(A%E2%80%93M\)](https://en.wikipedia.org/wiki/List_of_film_remakes_(A%E2%80%93M))
[https://en.wikipedia.org/wiki/List_of_film_remakes_\(N%E2%80%93Z\)](https://en.wikipedia.org/wiki/List_of_film_remakes_(N%E2%80%93Z))
https://en.wikipedia.org/wiki/List_of_films_considered_the_worst
https://en.wikipedia.org/wiki/Adam_Sandler_filmography
https://en.wikipedia.org/wiki/Category:Feminist_films
https://en.wikipedia.org/wiki/List_of_Holocaust_films

https://www.conservapedia.com/Essay:Greatest_Conservative_Movies
https://www.conservapedia.com/Essay:Worst_Liberal_Movies
https://www.conservapedia.com/Essay:20_Greatest_Conservative_Movies_of_the_Last_20_Years

Actors lists:

https://en.wikipedia.org/wiki/List_of_Jewish_actors
https://en.wikipedia.org/wiki/Category:Lists_of_actors_by_ethnicity
https://en.wikipedia.org/wiki/Category:Jewish_actors
https://en.wikipedia.org/wiki/List_of_African-American_actors
https://en.wikipedia.org/wiki/List_of_Italian-American_actors

https://en.wikipedia.org/wiki/List_of_Native_American_actors
https://en.wikipedia.org/wiki/List_of_American_actors_of_Irish_descent
https://en.wikipedia.org/wiki/List_of_Assyrian_actors
https://en.wikipedia.org/wiki/Category:Jewish_American_actresses
https://en.wikipedia.org/wiki/Category:Jewish_American_male_actors
https://en.wikipedia.org/wiki/Category:African-American_male_actors
https://en.wikipedia.org/wiki/Category:African-American_actresses

etc.

Instead of me manually finding all the lists, it makes more sense to have a Python crawler that finds all the different categories and subcategories so we can get every single wikipedia list related to film / actors.



The screenshot shows the Wikipedia page for the category "African-American_actors". The browser's address bar displays the URL: https://en.wikipedia.org/wiki/Category:African-American_actors. The page content includes a sidebar on the left with the word "pedia" visible. The main heading is "Subcategories". Below the heading, a text line states: "This category has the following 5 subcategories, out of 5 total." A plus sign (+) is shown, followed by a list of subcategories with their respective counts: "African-American actresses (1 C, 882 P)", "African-American child actresses (14 P)", "African-American male actors (1 C, 1,166 P)", and "African-American male child actors (37 P)".

https://en.wikipedia.org/wiki/Category:African-American_actors

Subcategories

This category has the following 5 subcategories, out of 5 total.

- +
 - ▼ African-American actresses (1 C, 882 P)
 - African-American child actresses (14 P)
 - ▼ African-American male actors (1 C, 1,166 P)
 - African-American male child actors (37 P)

2.)

- JewOrNotJew.com

full site




name, occupation category, results summary, verdict, rating1, rating2, rating3, link, etc.

Verdict CSS selector:

- **Businesspeople**
- **Comedians**
- **Directors**
- **Fictional Characters**
- **Historical Figures**
- **Media**
- **Musicians and Singers**
- **Politicians**
- **Religious Figures**
- **Scientists**
- **Sociopaths**
- **Writers**
- **[Uncategorized]**

Jew Score:

12

I5 O3 K4

two genders competed against each other. Men didn't not for Kronberger's reign. When the women got their national championship in 1924, the previous winners mixed competition were grandfathered (emphasis on as men's champions. Which, in Kronberger's case, was absurd.


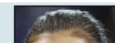

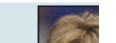
Yet not as absurd as her figure skating outfit...

Thanks, W **b#verdict** 28.58 x 16

Verdict: **Jew**.

February

See Also

v.com/profile.jsp?ID=361

Elements Console Sources Network Performance Memory Application Security Audits Adblock Plus

```

<td valign="top" class="normalText" width align="left">
  <span class="midText" id="subHead">...</span>
  <p>
    <div id="profileBody">...</div>
    <p>
      <span class="midText">
        "
        Verdict: "
        <b id="verdict">Jew</b> == $0
        "
        "
      </span>
    </p>
  </td>
  <td align="right" class="midText">...</td>

```

Styles

Filter

element.styl

```

}
strong, b {
  font-weight: bold;
}

```

Inherited from

```

.midText {
  font-family: sans-serif;
  font-size: 12px;
  color: black;
}

```

Inherited from

normalText


Ratings CSS Selectors:

Archive From

» Eddie Cantor

Categories




- Actors
- Actresses
- Artists
- Athletes and Coaches
- Businesspeople
- Comedians
- Directors
- Fictional Characters
- Historical Figures
- Media
- Musicians and Singers
- Politicians
- Religious Figures
- Scientists
- Sociopaths
- Writers
- [Uncategorized]



Jew Score:

td#ratbox2

40 × 71

I1


00

K4

If you are as confused as we Ryans, take the quiz below!

Reynolds!

Gosling!



Correct! That was Ryan Reynolds.

Still confused? OK, how about once married to Scarlett Johr point. **Gosling** was never ma

Console

Sources

Network

Performance

Memory

Application

Security

Audits

Adblock Plus

```

<br>
<table border="0" cellpadding="0" cellspacing="2">
  <tbody>
    <tr>...</tr>
    <tr>...</tr>
    <tr>
      <td width="5"></td>
      <td align="center" id="ratbox1" width="40">...</td>
      <td align="center" id="ratbox2" width="40">
         == $0
        <br>
        "
          0"
        <b>...</b>
      </td>
      <td align="center" id="ratbox3" width="40">...</td>
      <td width="5"></td>
    </tr>
    ...
  
```

Example pages:

<http://jewornotjew.com/profile.jsp?ID=349>

<http://jewornotjew.com/profile.jsp?ID=49>


3.)

- Ethnicelebs.com

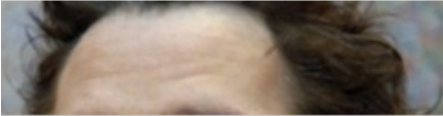
full site

name, occupation category, results, summary, source, link, etc.

No CSS selector, but a possible java regex pattern:



depo



Hills, CA, USA, 2014 – Keywords: Orientation: Portrait Face Cou
Credit: Janice Ogata / PRPhotos.com

Place of Birth: Detroit, Michigan, U.S.

Date of Birth: November 28, 1952

strong 211.75 × 23

Ethnicity: African-American

S. Epatha Merkerson is an American actress.

Her mother, Anna Bell (Bowen), was the only female

ents Console Sources Network Performance Memory Application Security Audits Adblock Plus

```
<div class="clear"></div>
<div class="entry themeform">
  <div class="entry-inner">
    <div class="code-block code-block-3 ai-viewport-3" style="margin: 8px auto 8px 0; text-align: left; display: block;
    both;"></div>
    <div id="attachment_214930" style="width: 410px" class="wp-caption alignnone"></div>
    <p>Place of Birth: Detroit, Michigan, U.S.</p>
    <p>Date of Birth: November 28, 1952</p>
    <p>
      <strong>Ethnicity: African-American</strong> == $0
    </p>
    <p>S. Epatha Merkerson is an American actress.</p>
    <div></div>
```

Example page:

<http://ethnicelebs.com/s-epatha-merkerson>

4.)

- [The-Numbers.com](http://www.the-numbers.com)

movie, box office total, box office domestic, box office international, box office per country A, country B, country C, etc.

movie, ticket sales total, ticket sales domestic, ticket sales international, ticket sales per country A, country B, country C, etc.

movie, list of actors separated by commas in category Stars, rank, mpaa rating, production method, video release, movie release, production method, video sales, genre, keywords,

distributor, box office total, box office domestic, box office international, box office per country A, country B, country C, etc.

distributor, ticket sales total, ticket sales domestic, ticket sales international, ticket sales per country A, country B, country C, etc.

[https://www.the-numbers.com/movie/Frozen-\(2009\)#tab=video-sales](https://www.the-numbers.com/movie/Frozen-(2009)#tab=video-sales)

https://www.the-numbers.com/glossary#acting_roles

Frozen (2010)



Watch Now ▶

Theatrical Performance

Domestic Box Office	\$246,177	Details
International Box Office	\$48,286,888	Details
Worldwide Box Office	\$48,533,065	

Home Market Performance

Est. Domestic DVD Sales	\$3,362,513	Details
Est. Domestic Blu-ray Sales	\$531,392	Details
Total Est. Domestic Video Sales	\$3,893,905	

[Further financial details...](#)

Tweet

Share 0

Like 0

Share

Summary

News

Box Office

International

Video Sales

Full Financials

Cast & Crew

Trailer

Weekly US DVD Sales

Date	Rank	Units this Week	% Change	Total Units	Spending this Week	Total Spending	Weeks in Release
10/3/2010	16	41,868		41,868	\$439,195	\$439,195	1

Weekly US Blu-ray Sales

Date	Rank	Units this Week	% Change	Total Units	Spending this Week	Total Spending	Weeks in Release
10/3/2010	19	4,832		4,832	\$72,438	\$72,438	1
10/10/2010	16	9,043	+87%	13,876	\$135,560	\$207,998	2
10/17/2010	19	6,741	-25%	20,617	\$73,138	\$281,136	3

Our DVD and Blu-ray sales estimates are based on weekly retail surveys, which we use to build a weekly market share estimate for each title we are tracking. The market share is converted into a weekly sales estimate based on industry reports on the overall size of the market. For more information, see our [Home Media Metrics](#).

Categorization of Acting Roles

We use eight acting role types to define the different performances in a movie, TV show, or stage play.

Leading or Lead Ensemble Members

For movies, we generally use a simple rule to determine whether an actor or actress has a leading role: Do they appear in the movie's theatrical poster? If they do, they have a leading role. If they do not, they do not have a leading role. If there are more than four cast members on the poster, then we classify all actors in the poster as being "Lead Ensemble Members."

Note that movies with large ensemble casts will sometimes have several posters highlighting different members of the cast. See, for example, the posters of the Lord of the Rings trilogy. All characters who get a poster are considered "lead ensemble members."

If the poster shows no actors or actresses at all, then we look at the people credited at the top of the poster. These people are considered the lead actors in this case.

For TV Shows, an acting role is considered to be leading if the actor is credited at the beginning of the show. Lead actors in TV shows are generally considered "lead ensemble members", unless they are the main focus of the show. For example, Sarah Michelle Gellar is classed as having a leading acting role in Buffy the Vampire Slayer, while Anthony Stewart Head, Alison Hannigan, Nicholas Brendon etc. are all classed as lead ensemble members.

Supporting

With the exception of cameos and leading players, all other cast members listed in the credits are considered to have supporting roles.

Cameo

A cameo is a small role performed by a well-known actor or director, often limited to a single scene, with the casting done specifically because of who they are, not because of their acting talents. Examples of cameos are Alfred Hitchcock's appearances in his films and Stan Lee's brief appearances in Marvel films. Note that small parts played by actors before they became famous are not cameos. Rather, they are either supporting roles or roles as an "extra."

Extra

If someone is not credited, and does not have a cameo, we classify their role as being an "extra." In general, we do not include extras in the credits, although some early roles as extras by now-famous actors and actresses are included.

Interviewee, Documentary Subject, Narrator, Archival Footage

NOTE: They have an API which makes things much easier. I will register and do what's needed. Just let me know. But here are the links.

<https://www.the-numbers.com/research-analysis>

>

<http://www.opusdata.com/>