

scraping

December 30, 2018

```
In [1]: import urllib.request
        from bs4 import BeautifulSoup
        def get_soup(url):
            request = urllib.request.Request(url)
            response = urllib.request.urlopen(request)
            return BeautifulSoup(response, 'html.parser')

In [2]: from bs4 import BeautifulSoup
        url = 'https://en.wikipedia.org/wiki/William_Shakespeare'
        soup = get_soup(url)

In [3]: soup.title

Out[3]: <title>William Shakespeare - Wikipedia</title>

In [4]: def get_links(soup):
        try:
            section = soup.find_all(id='mw-normal-catlinks')[0]
        except (IndexError):
            return []

        links = []
        for catlink in section.find_all('a')[1:]:
            link = catlink.get('href')
            print(catlink.get('title'), '->', link)
            links.append(link)
        return links

In [7]: def get(base_links):
        soupes = map(lambda link: get_soup('https://en.wikipedia.org'+link), base_links)
        return map(get_links, soupes)
        base_links = get_links(soup)

        new_links = get(base_links)
        new_links2 = get(new_links)
        new_links3 = get(new_links2)
        print(list(new_links))
        print(list(new_links2))
        print(list(new_links3))
```

Category:William Shakespeare -> /wiki/Category:William_Shakespeare
 Category:1564 births -> /wiki/Category:1564_births
 Category:1616 deaths -> /wiki/Category:1616_deaths
 Category:16th-century English male actors -> /wiki/Category:16th-century_English_male_actors
 Category:English male stage actors -> /wiki/Category:English_male_stage_actors
 Category:16th-century dramatists and playwrights -> /wiki/Category:16th-century_dramatists_and.
 Category:17th-century English dramatists and playwrights -> /wiki/Category:17th-century_English
 Category:16th-century English poets -> /wiki/Category:16th-century_English_poets
 Category:Burials in Warwickshire -> /wiki/Category:Burials_in_Warwickshire
 Category:People from Warwickshire -> /wiki/Category:People_from_Warwickshire
 Category:17th-century English poets -> /wiki/Category:17th-century_English_poets
 Category:17th-century male writers -> /wiki/Category:17th-century_male_writers
 Category:English Renaissance dramatists -> /wiki/Category:English_Renaissance_dramatists
 Category:English male writers -> /wiki/Category:English_male_writers
 Category:People educated at King Edward VI School, Stratford-upon-Avon -> /wiki/Category:Peopl
 Category:People from Stratford-upon-Avon -> /wiki/Category:People_from_Stratford-upon-Avon
 Category:People of the Elizabethan era -> /wiki/Category:People_of_the_Elizabethan_era
 Category:People of the Stuart period -> /wiki/Category:People_of_the_Stuart_period
 Category:Shakespeare family -> /wiki/Category:Shakespeare_family
 Category:Sonneteers -> /wiki/Category:Sonneteers
 Category:King's Men (playing company) -> /wiki/Category:King%27s_Men_(playing_company)
 Category:17th-century English male actors -> /wiki/Category:17th-century_English_male_actors
 Category:English male dramatists and playwrights -> /wiki/Category:English_male_dramatists_and.
 Category:English male poets -> /wiki/Category:English_male_poets
 Category:1564 -> /wiki/Category:1564
 Category:1560s births -> /wiki/Category:1560s_births
 Category:1616 -> /wiki/Category:1616
 Category:1610s deaths -> /wiki/Category:1610s_deaths
 Category:16th-century English actors -> /wiki/Category:16th-century_English_actors
 Category:16th-century male actors -> /wiki/Category:16th-century_male_actors
 Category:English male actors by century -> /wiki/Category:English_male_actors_by_century
 Category:English male stage actors -> /wiki/Category:English_male_stage_actors
 Category:British male stage actors -> /wiki/Category:British_male_stage_actors
 Category:English male actors by medium -> /wiki/Category:English_male_actors_by_medium
 Category:English stage actors -> /wiki/Category:English_stage_actors
 Category:16th-century theatre -> /wiki/Category:16th-century_theatre
 Category:16th-century writers -> /wiki/Category:16th-century_writers
 Category:Dramatists and playwrights by century -> /wiki/Category:Dramatists_and_playwrights_by.
 Category:17th-century dramatists and playwrights -> /wiki/Category:17th-century_dramatists_and.
 Category:17th-century English writers -> /wiki/Category:17th-century_English_writers
 Category:English dramatists and playwrights by century -> /wiki/Category:English_dramatists_and.
 Category:English poets by century -> /wiki/Category:English_poets_by_century
 Category:16th-century English writers -> /wiki/Category:16th-century_English_writers
 Category:16th-century poets -> /wiki/Category:16th-century_poets
 Category:Burials in England by county -> /wiki/Category:Burials_in_England_by_county
 Category:Warwickshire -> /wiki/Category:Warwickshire
 Category:Burials in West Midlands (region) -> /wiki/Category:Burials_in_West_Midlands_(region)
 Category:People by county in England -> /wiki/Category:People_by_county_in_England

Category:People from the West Midlands (region) -> /wiki/Category:People_from_the_West_Midlands
 Category:Warwickshire -> /wiki/Category:Warwickshire
 Category:English poets by century -> /wiki/Category:English_poets_by_century
 Category:17th-century English writers -> /wiki/Category:17th-century_English_writers
 Category:17th-century poets -> /wiki/Category:17th-century_poets
 Category:17th-century men -> /wiki/Category:17th-century_men
 Category:17th-century writers -> /wiki/Category:17th-century_writers
 Category:Male writers -> /wiki/Category:Male_writers
 Category:16th-century dramatists and playwrights -> /wiki/Category:16th-century_dramatists_and
 Category:17th-century dramatists and playwrights -> /wiki/Category:17th-century_dramatists_and
 Category:English dramatists and playwrights -> /wiki/Category:English_dramatists_and_playwright
 Category:Renaissance writers -> /wiki/Category:Renaissance_writers
 Category:16th-century English dramatists and playwrights -> /wiki/Category:16th-century_English
 Category:17th-century English dramatists and playwrights -> /wiki/Category:17th-century_English
 Category:British male writers -> /wiki/Category:British_male_writers
 Category:English men by occupation -> /wiki/Category:English_men_by_occupation
 Category:English writers -> /wiki/Category:English_writers
 Category:People educated by school in Warwickshire -> /wiki/Category:People_educated_by_school
 Category:People from Warwickshire -> /wiki/Category:People_from_Warwickshire
 Category:People by city or town in England -> /wiki/Category:People_by_city_or_town_in_England
 Category:Stratford-upon-Avon -> /wiki/Category:Stratford-upon-Avon
 Category:People from Stratford-on-Avon District -> /wiki/Category:People_from_Stratford-on-Avon
 Category:Elizabethan era -> /wiki/Category:Elizabethan_era
 Category:16th-century English people -> /wiki/Category:16th-century_English_people
 Category:People of the Tudor period -> /wiki/Category:People_of_the_Tudor_period
 Category:Stuart England -> /wiki/Category:Stuart_England
 Category:British people by period -> /wiki/Category:British_people_by_period
 Category:People of the Kingdom of England -> /wiki/Category:People_of_the_Kingdom_of_England
 Category:17th-century English people -> /wiki/Category:17th-century_English_people
 Category:People associated with Shakespeare -> /wiki/Category:People_associated_with_Shakespeare
 Category:English families -> /wiki/Category:English_families
 Category:Poets -> /wiki/Category:Poets
 Category:Sonnet studies -> /wiki/Category:Sonnet_studies
 Category:English early modern theatre companies -> /wiki/Category:English_early_modern_theatre
 Category:17th-century English male actors -> /wiki/Category:17th-century_English_male_actors
 Category:1603 establishments in England -> /wiki/Category:1603_establishments_in_England
 Category:1642 disestablishments -> /wiki/Category:1642_disestablishments
 Category:William Shakespeare -> /wiki/Category:William_Shakespeare
 Category:17th-century disestablishments in England -> /wiki/Category:17th-century_disestablishment
 Category:17th-century male actors -> /wiki/Category:17th-century_male_actors
 Category:17th-century English actors -> /wiki/Category:17th-century_English_actors
 Category:English male actors by century -> /wiki/Category:English_male_actors_by_century
 Category:English male stage actors -> /wiki/Category:English_male_stage_actors
 Category:English dramatists and playwrights -> /wiki/Category:English_dramatists_and_playwright
 Category:English male writers -> /wiki/Category:English_male_writers
 Category:British male dramatists and playwrights -> /wiki/Category:British_male_dramatists_and
 Category:English poets -> /wiki/Category:English_poets
 Category:English male writers -> /wiki/Category:English_male_writers

```
Category:British male poets -> /wiki/Category:British_male_poets
```

```
[[], ['/wiki/Category:1564', '/wiki/Category:1560s_births'], ['/wiki/Category:1616', '/wiki/Ca  
[]  
[]
```

```
In [ ]:
```