The site is open but it doesnt allow scraping for formats = rss and export = xls bring in the neccesary libraries from bs4 import BeautifulSoup as bs import requests import re import numpy as np import pandas as pd import matplotlib.pyplot as plt %matplotlib inline being able to scrap both pages use page !=2 page = 0 In [4]: while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" print(url) page = page + 1 https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page=0 https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page=1 Test the url to get a successful response On getting the HTTP response status codes...... Successful responses (200–299) #load html code from a url url = ('https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities') result = requests.get(url) result Out[11]: <Response [200]> **Authors** Method 1 In [6]: page = 0 # authors = [] while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') author = school.find all('a', class = "link person") # find all() method scans the entire document looking fun author =re.findall('rel="Person"....., str(author)) fun author = [item[19:]for item in fun author] #print(url) print(fun author) page = page + 1 ['Doha Abdelgawad<', 'Elizabeth Benjam', 'Joanne Callaghan', 'Tiziana Cervi-Wi', 'Yu-Hua Chen</spa', 'Zheci Che n</span', 'Thomas Cobb</spa', 'Ashley Cole</spa', 'Michael Cribb</s', 'Katie da Cunha L', 'James Dawson</sp', 'Mohammed Dayoub<', 'Miriam De Rosa</', 'Benjamin Dew</sp', 'Peter Finkenbusc', 'Zoe Gazeley-Eke<', 'Celia Geen </span', 'Mark Goodwin</sp', 'John Grima</span', 'James Henry</spa', 'Lewis Herrington', 'Begona Jauregui<', 'T im Kelly', 'Thomas Knowles</', 'Gavan Lennon</sp', 'Kristopher Lovel', 'Fiona Lowe</span', 'Pauline Male yrie', 'Marcus Maloney</', 'Emily Manns</pre>/spa', 'Simon May/span>', 'Lynsey McCulloch', 'Catherine Morgan', 'Joe Morrissey</s', 'Peta Murphy-Burk', 'Katie Nicholls</', 'Tim Nisbet</span', 'Masayo Nojiri Ge', 'Maria Norris</s p', 'Nick Oldham</spa', 'Monica Maria Ori', 'Marina Orsini-Jo', 'Katja Pilhuj</sp', 'Darren Reid</spa', 'Rebeca Reynolds<', 'Victoria Rios Ca', 'Felix Roesch</sp', 'Paul Ruscanu</sp', 'Ryan Scoats</spa', 'Emma Sheppard</s'] ['Chris Smith</spa', 'Simon Smith</spa', 'Kaori Stansall</', 'Sarah Turner</sp', 'Holly Vass</span', 'Liam Vern on</spa', 'Benet Vincent</s', 'Dimitra Vladimir', 'Thorsten Wojczew'] find the number of authors Method 2 In [7]: page = 0 authors = [] while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for a in school.find all('a', class ="link person"): authors.append(a.get text(strip=True)) #print(fun author) page = page + 1 authors Out[7]: ['Doha Abdelgawad', 'Elizabeth Benjamin', 'Joanne Callaghan', 'Tiziana Cervi-Wilson', 'Yu-Hua Chen', 'Zheci Chen', 'Thomas Cobb', 'Ashley Cole', 'Michael Cribb', 'Katie da Cunha Lewin', 'James Dawson', 'Mohammed Dayoub', 'Miriam De Rosa', 'Benjamin Dew', 'Peter Finkenbusch', 'Zoe Gazeley-Eke', 'Celia Geen', 'Mark Goodwin' 'John Grima', 'James Henry', 'Lewis Herrington', 'Begona Jauregui', 'Tim Kelly', 'Thomas Knowles', 'Gavan Lennon', 'Kristopher Lovell', 'Fiona Lowe', 'Pauline Maleyrie', 'Marcus Maloney', 'Emily Manns', 'Simon May', 'Lynsey McCulloch', 'Catherine Morgan', 'Joe Morrissey', 'Peta Murphy-Burke', 'Katie Nicholls', 'Tim Nisbet', 'Masayo Nojiri Gershon', 'Maria Norris', 'Nick Oldham', 'Monica Maria Oriani', 'Marina Orsini-Jones', 'Katja Pilhuj', 'Darren Reid', 'Rebeca Reynolds', 'Victoria Rios Castano', 'Felix Roesch', 'Paul Ruscanu', 'Ryan Scoats', 'Emma Sheppard', 'Chris Smith', 'Simon Smith', 'Kaori Stansall', 'Sarah Turner', 'Holly Vass', 'Liam Vernon', 'Benet Vincent', 'Dimitra Vladimirou', 'Thorsten Wojczewski'] len(authors) #number of authors Out[5]: 61 **Publications** In [6]: page = 0 publications = [] url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for a in school.find all('span', class ="minor dimmed"): publications.append(a.get text(strip=True)) #print(url) #print(fun author) page = page + 1 publications Out[6]: ['- Academic Marking Assistant', '- Lecturer in French', '- Lecturer in TEFL and CELTA', '- Assistant Professor Academic', '- Lecturer in Applied Linguistics', '- A251DEL, A150DEL, A251DEL/A351DEL, A150DEL/A250DEL/A350DEL, A351DEL, Academic Marking Assistant', '- Lecturer in International Relations', '- Academic Writing Tutor', '- Lecturer in Sociology and Criminology', '- Assistant Professor Academic', '- Lecturer in English Literature', '- Associate', '- Lecturer in Politics', '- A200DEL, Academic Marking Assistant', '- Project Supervisor', '- Project Supervisor', '- Associate', '- Associate Head of School - Recruitment and Marketing', '- Lecturer in International Relations', '- Assistant Professor (Academic)', '- Lecturer in Politics', '- Lecturer in Politics and History', '- Lecturer in English', '- Lecturer in International Relations', '- A281DEL, A381DEL, A181DEL, Academic Marking Assistant', '- Assistant Professor Academic', '- Curriculum Lead (Associate Professor - Academic)', '- Associate', '- Associate Head of School - Student Experience', '- Assistant Professor (Academic)', '- Academic Marking Assistant', '- 7136SOH, 7012SOH/7096SOH, 7139SOH/7148SOH, M06SOH/7009SOH, 7053SOH', '- Assistant Lecturer', '- Associate', '- Lecturer in Sociology', '- Lecturer in Interior Architecture and Design', '- Lecturer in International Relations', '- PHD Research Supervisor', '- PHD Research Supervisor', '- Curriculum Lead (Associate Professor - Academic)', '- Lecturer in English Literature', '- Associate', '- Associate Head of School - Enterprise and Innovation', '- CELTA', '- Assistant Professor Academic', '- A140DEL', '- Assistant Professor (Academic)', '- A311DEL, A111DEL, A221DEL, Academic Marking Assistant', '- A230DEL, A330DEL, A130DEL, Academic Marking Assistant', '- Associate', '- Professor in Education Practice', '- Lecturer in English Literature', '- Assistant Professor Academic', '- A353DEL', '- Associate', '- Lecturer in Spanish', '- Lecturer in Criminology', '- Associate Head of School - Research', '- Academic Marking Assistant', '- Lecturer in Sociology', '- Lecturer in Sociology', '- CELTA', '- Lecturer in History', '- Assistant Professor Academic', '- A340DEL, A240DEL, Academic Marking Assistant', '- Curriculum Lead Associate Professor - Academic', Assistant Professor (Academic)' '- Associate Head of School - Quality and Accreditation', '- Lecturer in International Relations', '- Associate', '- Assistant Professor in Applied Linguistics', '- Associate', '- Lecturer in English', '- Lecturer in International Relations'] len(publications) # number of the publications Out[7]: 74 **Titles** finding the titles In [8]: page = 0 title = [] while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for a in school.find_all('span', class_="minor dimmed"): title.append(a.get text(strip=True)) #print(url) #print(fun author) page = page + 1title Out[8]: ['- Academic Marking Assistant', '- Lecturer in French', Lecturer in TEFL and CELTA', '- Assistant Professor Academic' '- Lecturer in Applied Linguistics', '- A251DEL, A150DEL, A251DEL/A351DEL, A150DEL/A250DEL/A350DEL, A351DEL, Academic Marking Assistant', '- Lecturer in International Relations', '- Academic Writing Tutor', '- Lecturer in Sociology and Criminology', '- Assistant Professor Academic', '- Lecturer in English Literature', '- Associate', '- Lecturer in Politics', '- A200DEL, Academic Marking Assistant', '- Project Supervisor', '- Project Supervisor', '- Associate', '- Associate Head of School - Recruitment and Marketing', '- Lecturer in International Relations', '- Assistant Professor (Academic)', '- Lecturer in Politics', '- Lecturer in Politics and History', '- Lecturer in English', '- Lecturer in International Relations', '- A281DEL, A381DEL, A181DEL, Academic Marking Assistant', '- Assistant Professor Academic', '- Curriculum Lead (Associate Professor - Academic)', '- Associate', '- Associate Head of School - Student Experience', '- Assistant Professor (Academic)', '- Academic Marking Assistant', '- 7136SOH, 7012SOH/7096SOH, 7139SOH/7148SOH, М06SOH/7009SOH, 7053SOH', '- Assistant Lecturer', '- Associate', '- Lecturer in Sociology', '- Lecturer in Interior Architecture and Design', '- Lecturer in International Relations', '- PHD Research Supervisor', '- PHD Research Supervisor', '- Curriculum Lead (Associate Professor - Academic)', '- Lecturer in English Literature', '- Associate', '- Associate Head of School - Enterprise and Innovation', '- CELTA', '- Assistant Professor Academic', '- A140DEL', '- Assistant Professor (Academic) ' '- A311DEL, A111DEL, A221DEL, Academic Marking Assistant', '- A230DEL, A330DEL, A130DEL, Academic Marking Assistant', '- Associate', '- Professor in Education Practice', '- Lecturer in English Literature', '- Assistant Professor Academic', '- A353DEL', '- Associate', '- Lecturer in Spanish', '- Lecturer in Criminology', '- Associate Head of School - Research', '- Academic Marking Assistant', '- Lecturer in Sociology', '- Lecturer in Sociology' '- CELTA', '- Lecturer in History', '- Assistant Professor Academic', '- A340DEL, A240DEL, Academic Marking Assistant', '- Curriculum Lead Associate Professor - Academic', '- Assistant Professor (Academic)', '- Associate Head of School - Quality and Accreditation', '- Lecturer in International Relations', '- Associate', '- Assistant Professor in Applied Linguistics', '- Associate', '- Lecturer in English', '- Lecturer in International Relations'] **Years** In [11]: page = 0years = []while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for a in school.find all('span', class ="stacked-trend-graph-year"): years.append(a.get text(strip=True)) #print(url) #print(fun author) page = page + 1print([years]) [['2013', '2021', '2006', '2020', '2010', '2021', '2015', '2021', '2018', '2018', '2012', '2019', '2015', '202 0', '2005', '2020', '2016', '2021', '2015', '2020', '2019', '2019', '2014', '2019', '2010', '2021', '2012', '20 20', '2017', '2017', '2015', '2021', '2011', '2022', '2014', '2021', '2020', '2020', '1995', '2022', '2008', '2 021', '2005', '2022', '2016', '2021', '2010', '2022', '2015', '2021', '2017', '2021', '2012', '2022', '2001', '2021', '1995', '2016', '2013', '2022', '2004', '2017', '2013', '2021', '2010', '2021', '2010', '2021']] Finding the links In [21]: page = 0 links = []while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for link in soup.findAll('a'): links.append(link.get('href')) page = page + 1links Out[21]: ['#main-navigation', '#main-search', '#main-content' '/en/', 'https://share.coventry.ac.uk/staff/ps/research_office/pure/Pages/default.aspx', '/en/organisations/', '/en/persons/', '/en/publications/', '/en/projects/', '/en/studentTheses/', '/en/activities/', '/en/prizes/', '/en/datasets/', '/en/clippings/', '/en/equipments/', 'https://pureportal.coventry.ac.uk/en/organisations/coventry-university', 'https://pureportal.coventry.ac.uk/en/organisations/faculty-of-arts-humanities', '/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/fingerprints/', $\hbox{'/en/organisations/school-of-humanities/network/',}\\$ '/en/organisations/school-of-humanities/publications/', '/en/organisations/school-of-humanities/persons/', '/en/organisations/school-of-humanities/projects/' '/en/organisations/school-of-humanities/activities/', '/en/organisations/school-of-humanities/prizes/', '/en/organisations/school-of-humanities/studentTheses/', '/en/organisations/school-of-humanities/clippings/', "/en/organisations/school-of-humanities/persons/?page=\$7Bpage\$7D&ordering=lastNameFirstName&descending=true", and the substitution of the substi'https://pureportal.coventry.ac.uk/en/persons/doha-abdelgawad', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/elizabeth-benjamin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/joanne-callaghan', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tiziana-cervi-wilson', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/yu-hua-chen-chen', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zheci-chen-2', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-cobb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-academic-writing', 'https://pureportal.coventry.ac.uk/en/persons/ashley-cole', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/michael-cribb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-da-cunha-lewin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-dawson', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-trust-peace-and-social-relations-3', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mohammed-dayoub', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/miriam-de-rosa', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/benjamin-dew', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/peter-finkenbusch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zoe-gazeley-eke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mark-goodwin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/john-grima', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-henry', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lewis-herrington', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/begona-jauregui', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-kelly', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-knowles', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/gavan-lennon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/kristopher-lovell', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/fiona-lowe', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-nursing-midwifery-and-health', 'https://pureportal.coventry.ac.uk/en/persons/pauline-maleyrie', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marcus-maloney', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', https://pureportal.coventry.ac.uk/en/persons/emily-manns 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/simon-may', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lynsey-mcculloch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-media-and-performing-arts', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/catherine-morgan', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/joe-morrissey', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/peta-murphy-burke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-nicholls', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-nisbet', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/masayo-nojiri-gershon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/maria-norris', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/nick-oldham', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/monica-maria-oriani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marina-orsini-jones', 'https://pureportal.coventry.ac.uk/en/organisations/centre-global-learning', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katja-pilhuj', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/darren-reid', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/rebeca-reynolds', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/victoria-rios-castano', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/vladimir-rizov', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/felix-roesch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/paul-cosmin-ruscani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/ryan-scoats', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/emma-sheppard', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/persons/?page=1', '/en/organisations/school-of-humanities/persons/?page=1', 'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/pure', 'http://www.scopus.com/', 'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/elsevier-fingerprint-engin 'http://www.elsevier.com', '/en/about/cookies/', 'https://pure.coventry.ac.uk/admin/workspace.xhtml', '/en/web-accessibility/', 'mailto:enquiries.pure@coventry.ac.uk', '#main-navigation', '#main-search', '#main-content' '/en/', 'https://share.coventry.ac.uk/staff/ps/research office/pure/Pages/default.aspx', '/en/', '/en/organisations/', '/en/persons/', '/en/publications/', '/en/projects/', '/en/studentTheses/', '/en/activities/', '/en/prizes/', '/en/datasets/' '/en/clippings/', '/en/equipments/', 'https://pureportal.coventry.ac.uk/en/organisations/coventry-university', $\verb|'https://pureportal.coventry.ac.uk/en/organisations/faculty-of-arts-humanities'|,$ '/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/fingerprints/', /en/organisations/school-of-humanities/network/ '/en/organisations/school-of-humanities/publications/', '/en/organisations/school-of-humanities/persons/', '/en/organisations/school-of-humanities/projects/' '/en/organisations/school-of-humanities/activities/', '/en/organisations/school-of-humanities/prizes/', '/en/organisations/school-of-humanities/studentTheses/', '/en/organisations/school-of-humanities/clippings/', "/en/organisations/school-of-humanities/persons/?page=\$7Bpage\$7D&ordering=lastNameFirstName&descending=true", and the substitution of the substi'https://pureportal.coventry.ac.uk/en/persons/doha-abdelgawad', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/elizabeth-benjamin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/joanne-callaghan', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tiziana-cervi-wilson', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/yu-hua-chen-chen', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zheci-chen-2', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-cobb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-academic-writing', 'https://pureportal.coventry.ac.uk/en/persons/ashley-cole', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/michael-cribb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-da-cunha-lewin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-dawson', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-trust-peace-and-social-relations-3', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mohammed-dayoub', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/miriam-de-rosa', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/benjamin-dew', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/peter-finkenbusch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zoe-gazeley-eke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mark-goodwin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/john-grima', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-henry', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lewis-herrington', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/begona-jauregui', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-kelly', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-knowles', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/gavan-lennon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/kristopher-lovell', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/fiona-lowe', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-nursing-midwifery-and-health', 'https://pureportal.coventry.ac.uk/en/persons/pauline-maleyrie', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marcus-maloney', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/emily-manns', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/simon-may', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lynsey-mcculloch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-media-and-performing-arts', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/catherine-morgan', https://pureportal.coventry.ac.uk/en/organisations/school-of-h 'https://pureportal.coventry.ac.uk/en/persons/joe-morrissey', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/peta-murphy-burke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-nicholls', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-nisbet', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/masayo-nojiri-gershon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/maria-norris', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/nick-oldham', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/monica-maria-oriani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marina-orsini-jones', 'https://pureportal.coventry.ac.uk/en/organisations/centre-global-learning', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katja-pilhuj', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/darren-reid', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/rebeca-reynolds', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/victoria-rios-castano', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/vladimir-rizov', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/felix-roesch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/paul-cosmin-ruscani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/ryan-scoats', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/emma-sheppard', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/persons/?page=1', '/en/organisations/school-of-humanities/persons/?page=1', 'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/pure', 'http://www.scopus.com/', 'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/elsevier-fingerprint-engin e', 'http://www.elsevier.com', '/en/about/cookies/', 'https://pure.coventry.ac.uk/admin/workspace.xhtml', '/en/web-accessibility/', 'mailto:enquiries.pure@coventry.ac.uk', from bs4 import BeautifulSoup from urllib.request import Request, urlopen import re req = Request("https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}") html page = urlopen(req) soup = BeautifulSoup(html page, "lxml") links = []for link in soup.findAll('a'): links.append(link.get('href')) links Out[16]: ['#main-navigation', '#main-search', '#main-content', '/en/', 'https://share.coventry.ac.uk/staff/ps/research office/pure/Pages/default.aspx', '/en/organisations/', '/en/persons/', '/en/publications/', '/en/projects/', '/en/studentTheses/', /en/activities/ '/en/prizes/', '/en/datasets/' '/en/clippings/', '/en/equipments/', 'https://pureportal.coventry.ac.uk/en/organisations/coventry-university', 'https://pureportal.coventry.ac.uk/en/organisations/faculty-of-arts-humanities', '/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/fingerprints/', '/en/organisations/school-of-humanities/network/', '/en/organisations/school-of-humanities/publications/', '/en/organisations/school-of-humanities/persons/', '/en/organisations/school-of-humanities/projects/', '/en/organisations/school-of-humanities/activities/', '/en/organisations/school-of-humanities/prizes/', '/en/organisations/school-of-humanities/studentTheses/', '/en/organisations/school-of-humanities/clippings/', '/en/organisations/school-of-humanities/persons/?page=%7Bpage%7D&ordering=lastNameFirstName&descending=true', 'https://pureportal.coventry.ac.uk/en/persons/doha-abdelgawad', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/elizabeth-benjamin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/joanne-callaghan', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tiziana-cervi-wilson', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/yu-hua-chen-chen', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zheci-chen-2', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-cobb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-academic-writing', 'https://pureportal.coventry.ac.uk/en/persons/ashley-cole', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/michael-cribb', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-da-cunha-lewin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-dawson', 'https://pureportal.coventry.ac.uk/en/organisations/centre-for-trust-peace-and-social-relations-3', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mohammed-dayoub', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/miriam-de-rosa', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/benjamin-dew', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/peter-finkenbusch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/zoe-gazeley-eke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/mark-goodwin', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/john-grima', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/james-henry', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lewis-herrington', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/begona-jauregui', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-kelly', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/thomas-knowles', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/gavan-lennon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/kristopher-lovell', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/fiona-lowe', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', $\verb|'https://pureportal.coventry.ac.uk/en/organisations/school-of-nursing-midwifery-and-health'|,$ 'https://pureportal.coventry.ac.uk/en/persons/pauline-maleyrie', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marcus-maloney', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-in-postdigital-cultures', https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities' 'https://pureportal.coventry.ac.uk/en/persons/emily-manns', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/simon-may', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/lynsey-mcculloch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-media-and-performing-arts', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/catherine-morgan', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/joe-morrissey', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/persons/peta-murphy-burke', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katie-nicholls', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/tim-nisbet', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/masayo-nojiri-gershon', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/maria-norris', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/nick-oldham', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/monica-maria-oriani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/marina-orsini-jones', 'https://pureportal.coventry.ac.uk/en/organisations/centre-global-learning', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/katja-pilhuj', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/darren-reid', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/rebeca-reynolds', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/victoria-rios-castano', 'https://pureportal.coventry.ac.uk/en/organisations/research-centre-for-arts-memory-and-communities', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/vladimir-rizov', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/felix-roesch', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/paul-cosmin-ruscani', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/ryan-scoats', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', 'https://pureportal.coventry.ac.uk/en/persons/emma-sheppard', 'https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities', '/en/organisations/school-of-humanities/persons/?page=1', '/en/organisations/school-of-humanities/persons/?page=1', $\verb|'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/pure'|,$ 'http://www.scopus.com/', 'http://www.elsevier.com/online-tools/research-intelligence/products-and-services/elsevier-fingerprint-engin e', 'http://www.elsevier.com', '/en/about/cookies/', 'https://pure.coventry.ac.uk/admin/workspace.xhtml', '/en/web-accessibility/', 'mailto:enquiries.pure@coventry.ac.uk', '#'] In [9]: page = 0 graphs = [] while page != 2: url = f"https://pureportal.coventry.ac.uk/en/organisations/school-of-humanities/persons/?page={page}" result = requests.get(url) school = bs(result.content, 'html.parser') for a in school.find all('span', class ="trend-bar-stacked "): years.append(a.get_text(strip=True)) #print(url) #print(fun author) page = page + 1 Useful resource to learn web scraping (https://www.youtube.com/watch?v=RvCBzhhydNk&list=WL&index=15) by Pythonology