

# Unique Web Link Scraper

## Python:

A. When extracting the web links from the HTML code of the “Current Estimates” web link multiple parts of the code work together to accomplish the end result.

First, the needed libraries are imported:

```
1. from bs4 import BeautifulSoup, SoupStrainer
2. from urllib.parse import urljoin
3. from urllib.request import urlopen
```

Second, using the imported libraries the URL is opened, read and parsed:

```
1. url = "https://www.census.gov/data/tables/2016/demo/popest.html"
2.
3. """ Pulls unique links from specified web link and creates a
4.     csv file with the web links with the header web_link."""
5. def getalllinks(url):
6.     urlContent = urlopen(url)
7.     urlClient = urlContent.read()
8.     soup = BeautifulSoup(urlClient, 'html.parser')
9.     urlContent.close()
```

Third, a code is written to identify what we want to pull from web pages HTML that we just ran through BeautifulSoup:

```
1. """ Soups all 'a' tags where
2.     a href tag is present """
3. links = soup.find_all('a', href = True)
4.
5. url_set = set()
6.
7. """ Returns all hrefs and makes each relative
8.     link into an absolute link through urljoin.
9.     all links ending with a pdf, xlsx, doc, etc
10.    are not returned. Code also checks for
11.    duplicate links and returns only unique links """
12. for link in links:
13.     base_url = "https://www.census.gov/data/tables/2016/demo/popest/state-
total.html/"
14.     link = urljoin(base_url, link['href'])
15.     if link[-1] == '/': # Ensures that HTML link is a directory and not a file.
16.         link = link[:-1]
17.     ext = ["pdf", "xlsx", "doc", "docx", "mov", \
18.           "mp4", "mpg", "wmv", "javascript:window.print()"]
19.     if not link.endswith(tuple(ext)):
20.         if link and link not in url_set:
21.             f.write(str(link) + "\n")
```

Essentially, if we were to go to the selected web link

<https://www.census.gov/data/tables/2016/demo/popest/state-total.html> and right click and press inspect to view the raw HTML we could search through and find the correlating “a” tags that contain a href. However, this would-be time consuming and not practical if you had to scrape multiple web pages for certain information. Instead the code written above tells on computer when it wants to search for and return. Similarly, the code above tells your computer what it does not want to be returned in a roundabout way. In the website two types of links can be identified, file links and web links. The file link will open a file (i.e. pdf, xlsx, doc, etc.). Whereas, the web link will take you to another web page. To tell the program to only return the web links and to essentially ignore the file links the code refers to traditional HTML code structuring. In HTML language a web link ends with a forward slash “/”.

Conversely, a file link does not end with a forward slash. See code snippet below:

```
1. if link[-1] == '/': # Ensures that HTML link is a directory and not a file.
2.     link = link[:-1]
```

Essentially the code above tells the program to look for web links on the bases that they end in a forward slash.

**B.** To determine if a link is a locator to another HTML page the ‘a’ tag is requested where href equals True. It is import to note that out of all the tags in the HTML that could be selected the ‘a’ was selected because “The <a> tag defines a hyperlink, which is used to link from one page to another. The most important attribute of the <a> element is the href attribute, which indicates the link's destination. (w3schools.com, n.d.)” With both the <a> tag and the href element selected each unique link was extracted. Furthermore, as mentioned above in section “A” the code looks specifically for web links or links that are locators to another HTML page. This is done by looking for a forward slash “/” at the end of the link. Yet, as we most surly know all HTML code is not written perfectly. Consequently, a section of code is added to catch any files that may have come through despite our best efforts. This is done by looking for common file extensions at the end of a link. If the link ends with one of the specified extensions it is not returned. See code portion below.

```
1. links = soup.find_all('a', href = True)
2.
3. url_set = set()
4.
5. """ Retuns all hrefs and makes each relative
6.     link into an absolute link through urljoin.
7.     all links ending with a pdf, xlsx, doc, etc
8.     are not retuned. Code also checks for
9.     duplicate links and returns only unique links """
10. for link in links:
11.     base_url = "https://www.census.gov/data/tables/2016/demo/popest/state-
total.html/"
12.     link = urljoin(base_url, link['href'])
13.     if link[-1] == '/': # Ensures that HTML link is a directory and not a file.
14.         link = link[:-1]
15.     ext = ["pdf", "xlsx", "doc", "docx", "mov", \
16.           "mp4", "mpg", "wmv", "javascript:window.print()"]
17.     if not link.endswith(tuple(ext)):
18.         if link and link not in url_set:
19.             f.write(str(link) + "\n")
```

C. Absolute links were ensured by executing the following code:

```
1. for link in links:
2.     base_url = "https://www.census.gov/programs-surveys/popest.html/"
3.     link = urljoin(base_url, link['href'])
```

Through the URL Parse library, urljoin with a base URL was used to ensure relative links were converted to absolute links based on the href attribute. urljoin uses the base URL to check for missing pieces in a relative link and inputs the needed changes. Thus creating an absolute link.

D. To eliminate or avoid duplicate links in the output file an unordered collection of unique links was created through set(). Through set we created an unordered collection of unique elements that could be compared against itself. Each element was compared against each other or in essence link was compared against link. If one link already existed it was not returned to the final list. Below is the executed code:

```
1. url_set = set()
2.
3. """ Returns all hrefs and makes each relative
4.     link into an absolute link through urljoin.
5.     all links ending with a pdf, xlsx, doc, etc
6.     are not returned. Code also checks for
7.     duplicate links and returns only unique links """
8. for link in links:
9.     base_url = "https://www.census.gov/data/tables/2016/demo/popest/state-
total.html/"
10.    link = urljoin(base_url, link['href'])
11.    if link[-1] == '/': # Ensures that HTML link is a directory and not a file.
12.        link = link[:-1]
13.    ext = ["pdf", "xlsx", "doc", "docx", "mov", \
14.          "mp4", "mpg", "wmv", "javascript:window.print()"]
15.    if not link.endswith(tuple(ext)):
16.        if link and link not in url_set:
17.            f.write(str(link) + "\n")
18.    url_set.add(link)
```

E. Below is the full version of the code used to extract all the unique web links from the selected website:

```
1. from bs4 import BeautifulSoup, SoupStrainer
2. from urllib.parse import urljoin
3. from urllib.request import urlopen
4. from urllib.parse import urlparse
5.
6.
7. url = "https://www.census.gov/data/tables/2016/demo/popest/state-total.html"
8.
9. """ Pulls unique links from specified web link and creates a
10.    csv file with the web links with the header web_link."""
11. def getalllinks(url):
12.    urlContent = urlopen(url)
```

```

13.     urlClient = urlContent.read()
14.     soup = BeautifulSoup(urlClient, 'html.parser')
15.     urlContent.close()
16.
17.     filename = "web_links.csv"
18.     f = open(filename, "w")
19.     headers = "web_links\n"
20.     f.write(headers)
21.
22.     """ Soups all 'a' tags where
23.         a href tag is present """
24.     links = soup.find_all('a', href = True)
25.
26.     url_set = set()
27.
28.     """ Returns all hrefs and makes each relative
29.         link into an absolute link through urljoin.
30.         all links ending with a pdf, xlsx, doc, etc
31.         are not returned. Code also checks for
32.         duplicate links and returns only unique links """
33.     for link in links:
34.         base_url = "https://www.census.gov/data/tables/2016/demo/popest/state-
total.html/"
35.         link = urljoin(base_url, link['href'])
36.         if link[-1] == '/': # Ensures that HTML link is a directory and not a file.
37.             link = link[:-1]
38.         ext = ["pdf", "xlsx", "doc", "docx", "mov", \
39.             "mp4", "mpg", "wmv", "javascript:window.print()"]
40.         if not link.endswith(tuple(ext)):
41.             if link and link not in url_set:
42.                 f.write(str(link) + "\n")
43.                 url_set.add(link)
44.                 print(link) # used to view results of function
45.     f.close() #stops wrting file
46. print(getalllinks(url)) # prints results of the function

```

**F.** HTML code of the “Current Estimates” web page is provided in the `www_census_gov.txt` file

**G.** csv file created by script titled: `web_links.csv`

**H.** Results from the executed code + a screenshot:

```

1. https://www.census.gov/topics/population/ancestry.html
2. https://www.census.gov/data/tables.1931.html
3. https://www.census.gov/careers
4. https://www.census.gov/topics/public-sector/congressional-and-intergovernmental.html
5. https://www.census.gov/data/tables/time-series/econ/mhs/shipments.html
6. https://www.census.gov/programs-surveys/popest.html#NAV_30093459_0_accd
7. https://www.census.gov/data/training-workshops.html
8. https://www.census.gov/data/tables.1933.html
9. https://www.census.gov/data/data-tools/quickfacts.html
10. https://www.census.gov/2018censustest
11. https://www.census.gov/data/tables.1925.html
12. https://www.census.gov/topics/employment/data.html
13. https://www.census.gov/topics/population/international.html
14. https://www.census.gov/data/tables.1935.html
15. https://www.census.gov/programs-surveys/are-you-in-a-survey/fraudulent-activity-and-
scams.html
16. https://www.census.gov/data/tables.2008.html
17. https://www.census.gov/data/tables.1936.html

```

18. <https://www.census.gov/about/contact-us.html>
19. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>
20. <https://www.census.gov/topics/employment/commuting.html>
21. <https://www.census.gov/data/tables.1950.html>
22. <https://www.census.gov/topics/public-sector/taxes.html>
23. <https://www.census.gov/library/visualizations.html>
24. <https://www.census.gov/topics/business/classification-codes.html>
25. <https://www.census.gov/EconomicCensus>
26. <https://www.census.gov/topics/income-poverty/income-inequality.html>
27. <https://www.census.gov/quality>
28. <https://www.census.gov/topics/health/data.html>
29. <https://www.census.gov/data/tables.1934.html>
30. <https://www.census.gov/popest>
31. <https://www.census.gov/data/tables.1956.html>
32. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_437121255\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_437121255_0_accd)
33. <https://www.census.gov/data/tables/2017/demo/soma/2017-qtly-metro-absorption.html>
34. <https://www.census.gov/topics/families/news.html>
35. <https://www.census.gov/topics/employment/industry-occupation.html>
36. <https://www.census.gov/topics/families.html>
37. <https://www.census.gov/data/tables.1977.html>
38. <https://www.census.gov/topics/education/school-enrollment.html>
39. <https://www.census.gov/topics/health/fertility.html>
40. <https://www.census.gov/data/tables.2014.html>
41. <https://www.census.gov/data/tables.1953.html>
42. <https://www.census.gov/data/tables.1978.html>
43. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_1207455997\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_1207455997_0_accd)
44. <https://www.census.gov/topics/public-sector.html>
45. <https://www.census.gov/topics/public-sector/apes.html>
46. <https://www.census.gov/data/tables.2017.html>
47. <https://www.census.gov/data/tables.1985.html>
48. <https://www.census.gov/topics/population/population-estimates.html>
49. <https://www.census.gov/data/tables.1907.html>
50. <https://www.census.gov/data/tables.1991.html>
51. <https://www.census.gov/privacy>
52. <https://www.census.gov/topics/business/construction.html>
53. <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-hinc.html>
54. <https://www.census.gov/topics/housing/guidance.html>
55. <https://www.census.gov/data/data-tools/my-congressional-district.html>
56. <https://www.census.gov/topics/population/migration.html>
57. <https://www.census.gov/geography/education.html>
58. <https://www.census.gov/data/tables.1895.html>
59. <https://www.census.gov/topics/health/health-insurance.html>
60. <https://www.census.gov/data/tables.1981.html>
61. <https://www.census.gov/topics/population/children.html>
62. <https://www.census.gov/topics/business/retail-trade.html>
63. <https://www.census.gov/topics/housing/homeownership.html>
64. <https://www.census.gov/data/tables.1999.html>
65. <https://www.census.gov/topics/economy/economic-indicators.html>
66. <https://www.census.gov/data/tables.1897.html>
67. <https://www.census.gov/topics/population/population-projections.html>
68. <https://www.census.gov/topics/families/same-sex-couples.html>
69. <https://www.census.gov/library/working-papers.html>
70. <https://www.census.gov/topics/population/hispanic-origin.html>
71. <https://www.census.gov/about/who.html>
72. <https://www.census.gov/data/tables.1945.html>
73. <https://www.census.gov/programs-surveys.html>
74. <https://www.census.gov/library/video.html>
75. <https://www.census.gov/topics/families/marriage-and-divorce.html>
76. <https://www.census.gov/geography/reference.html>
77. <https://www.census.gov/topics/housing/financial-characteristics.html>

78. <https://www.census.gov/topics/income-poverty/data.html>  
79. <https://www.census.gov/topics/business/services.html>  
80. <https://www.census.gov/data/tables.1918.html>  
81. <https://www.census.gov/data/tables.1947.html>  
82. <https://www.census.gov/topics/employment/labor-force.html>  
83. <https://www.census.gov/geography/metro.html>  
84. <https://www.census.gov/data/tables.1988.html>  
85. <https://www.census.gov/geography/interactive-maps.html>  
86. <https://www.census.gov/topics/population/race.html>  
87. <https://www.census.gov/topics/economy/isp.html>  
88. <https://www.census.gov/aian>  
89. <https://www.census.gov/AmericaCounts>  
90. <https://www.census.gov/topics/population/videos.html>  
91. <https://www.census.gov/data/tables/time-series/econ/mhs/latest-data.html>  
92. <https://www.census.gov/programs-surveys/popest.html#skiptoplinks>  
93. <https://www.census.gov/data/tables.2016.html>  
94. <https://www.census.gov/about/what/admin-data.html>  
95. <https://www.census.gov/topics/economy/stories.html>  
96. <https://www.census.gov/data/tables.1979.html>  
97. <https://www.census.gov/topics/income-poverty/saipe.html>  
98. <https://www.census.gov/topics/population/about.html>  
99. <https://www.census.gov/data/tables.1894.html>  
100. <https://www.census.gov/topics/employment/publications.html>  
101. <https://www.census.gov/data/tables.1989.html>  
102. <https://www.census.gov/data/tables.1927.html>  
103. <https://www.census.gov/data/data-tools.html>  
104. <https://www.census.gov/topics/education/public-school-finance.html>  
105. [https://www.census.gov/programs-surveys/popest.html#NAV\\_1713211272\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html#NAV_1713211272_0_accd)  
106. <https://www.census.gov/2020census>  
107. <https://www.census.gov/about/diversity-networks.html>  
108. <https://www.census.gov/topics/business/wholesale-trade.html>  
109. <https://www.census.gov/topics/population.html>  
110. <https://www.census.gov/topics/housing/rental-housing.html>  
111. <https://www.census.gov/data/tables.1928.html>  
112. <https://www.census.gov/data/tables.1901.html>  
113. <https://www.census.gov/data/tables.1974.html>  
114. <https://www.census.gov/geography.html>  
115. <https://www.census.gov/programs-surveys/popest.html#https://www.census.gov/newsroom.html>  
116. <https://www.census.gov/data/tables.1929.html>  
117. <https://www.census.gov/topics/health/news.html>  
118. <https://www.census.gov/topics/population/publications.html>  
119. <https://www.census.gov/topics/public-sector/stories.html>  
120. <https://www.census.gov/about.html>  
121. <https://www.census.gov/topics/health/social-assistance-industries.html>  
122. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-families.html>  
123. <https://www.census.gov/data/tables.1961.html>  
124. <https://www.census.gov/programs-surveys/popest.html#https://www.census.gov/programs-surveys.html>  
125. <https://www.census.gov/data/tables.1955.html>  
126. <https://www.census.gov/data/tables.1914.html>  
127. <https://www.census.gov/data/data-tools/census-explorer.html>  
128. <https://www.census.gov/data/tables.1905.html>  
129. <https://www.census.gov/topics/international-trade/definitions.html>  
130. <https://www.census.gov/data/tables.1959.html>  
131. <https://www.census.gov/data/tables.1930.html>  
132. <https://www.census.gov>  
133. <https://www.census.gov/topics/public-sector/government-finances.html>  
134. <https://www.census.gov/about/cong-gov-affairs.html>  
135. <https://www.census.gov/topics/families/child-care.html>

136. <https://www.census.gov/data/tables.1968.html>  
137. <https://www.census.gov/topics/education/surveys.html>  
138. <https://www.census.gov/topics/housing/stories.html>  
139. <https://www.census.gov/topics/population/computer-internet.html>  
140. <https://www.census.gov/data/related-sites.html>  
141. <https://www.census.gov/topics/health/publications.html>  
142. <https://www.census.gov/data/tables.1996.html>  
143. <https://www.census.gov/data/tables/2017/demo/soma/2017-qtly-absorptions.html>  
144. <https://www.census.gov/topics/population/age-and-sex.html>  
145. [https://www.census.gov/newsroom/releases/archives/embargoed\\_releases](https://www.census.gov/newsroom/releases/archives/embargoed_releases)  
146. [https://www.census.gov/about/contact-us/social\\_media.html](https://www.census.gov/about/contact-us/social_media.html)  
147. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_2133229556\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_2133229556_0_accd)  
148. <https://www.census.gov/data/tables.1972.html>  
149. <https://www.census.gov/topics/housing/physical-characteristics.html>  
150. <https://www.census.gov/data/tables.1896.html>  
151. <https://www.census.gov/topics/housing.html>  
152. <http://www.pinterest.com/uscensusbureau>  
153. <https://www.census.gov/data/tables.1892.html>  
154. <https://www.census.gov/topics/public-sector/criminal-justice.html>  
155. <https://www.census.gov/topics/education/news.html>  
156. <https://www.census.gov/topics/public-sector/bps.html>  
157. <https://www.census.gov/data/data-tools/cbb-redirect.html>  
158. <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov.html>  
159. <https://www.census.gov/topics/population/foreign-born.html>  
160. <https://www.census.gov/topics/employment/employers--public-sector.html>  
161. <https://www.census.gov/data/tables.1946.html>  
162. <https://www.census.gov/library/audio.html>  
163. <https://www.census.gov/topics/public-sector/voting.html>  
164. <https://www.census.gov/data/tables.1911.html>  
165. <https://www.census.gov/data/tables.2017.html/page/0.html>  
166. <https://www.census.gov/topics/income-poverty/library.html>  
167. <https://www.census.gov/topics/families/about.html>  
168. <https://www.census.gov/data/tables.1993.html>  
169. <https://www.census.gov/data/tables.1924.html>  
170. <https://www.census.gov/programs-surveys/sis.html>  
171. <https://www.census.gov/topics/international-trade/reference.html>  
172. <https://www.census.gov/data/tables.1958.html>  
173. <https://www.usa.gov>  
174. <https://www.census.gov/topics/public-sector/cog.html>  
175. <https://www.census.gov/programs-surveys/popest.html/#https://www.census.gov/library.html>  
176. <https://www.census.gov/topics/employment/work-from-home.html>  
177. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-inequality.html>  
178. <https://www.census.gov/regions>  
179. <https://www.census.gov/topics/economy/videos.html>  
180. <https://www.census.gov/programs-surveys/acs>  
181. <https://www.census.gov/foia>  
182. <https://www.census.gov/programs-surveys/ahs.html>  
183. <http://www.linkedin.com/company/us-census-bureau>  
184. <https://www.census.gov/topics/population/language-use.html>  
185. <https://www.census.gov/data/tables.1995.html>  
186. <https://www.census.gov/topics/economy/ces.html>  
187. <https://www.census.gov/programs-surveys/popest.html/#skipNav-CMSContent>  
188. <https://www.census.gov/topics/health/disability.html>  
189. <https://www.census.gov/topics/health/sahie.html>  
190. <https://www.census.gov/data/tables.2005.html>  
191. <http://twitter.com/uscensusbureau>  
192. <https://www.census.gov/data/tables.1990.html>  
193. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_2137395588\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_2137395588_0_accd)



194. <https://www.census.gov/topics/housing/surveys.html>  
195. <https://www.census.gov/topics/families/grandparents-and-grandchildren.html>  
196. <https://www.census.gov/topics/population/news.html>  
197. <https://www.census.gov/topics/housing/data.html>  
198. <https://www.census.gov/topics/health/working-papers.html>  
199. <https://www.census.gov/topics/employment/news.html>  
200. <https://www.census.gov/data/product-catalog.html>  
201. <https://www.census.gov/sipp>  
202. <https://www.census.gov/data/tables.1938.html>  
203. <https://www.census.gov/data/tables.1997.html>  
204. <https://www.census.gov/data/tables.2013.html>  
205. <https://www.census.gov/topics/business/manufacturing.html>  
206. <https://www.census.gov/topics/employment.html>  
207. <https://www.census.gov/data/tables.1965.html>  
208. <https://www.census.gov/topics/employment/led.html>  
209. <https://www.census.gov/topics/families/visualizations.html>  
210. <https://www.census.gov/data/tables.1951.html>  
211. <https://www.census.gov/topics/health/expenses-investments.html>  
212. <https://www.census.gov/programs-surveys/are-you-in-a-survey.html>  
213. <https://www.census.gov/data/data-tools/interactive-population-map.html>  
214. <https://www.census.gov/topics/population/data.html>  
215. <https://www.census.gov/topics/employment/equal-employment-opportunity-tabulation.html>  
216. <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-hi.html>  
217. <https://www.census.gov/data/tables.2009.html>  
218. <https://www.census.gov/data/tables.1960.html>  
219. <https://www.census.gov/data/tables/time-series/econ/mhs/month-of-shipment.html>  
220. <https://www.census.gov/topics/international-trade/trade-regulations.html>  
221. <https://www.census.gov/data/tables.1893.html>  
222. <https://www.census.gov/topics/business/news.html>  
223. <https://www.census.gov/topics/international-trade/news.html>  
224. <http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml>  
225. <https://www.census.gov/data/tables.1891.html>  
226. <https://www.census.gov/data/tables.1932.html>  
227. <https://www.census.gov/programs-surveys/popest.html/#skipfooter>  
228. <https://www.census.gov/topics/employment/disability-employment-tabulation.html>  
229. <https://www.census.gov/topics/health/about.html>  
230. <https://www.census.gov/topics/families/families-and-households.html>  
231. <https://www.census.gov/data/tables.1900.html>  
232. <https://www.census.gov/topics/employment/about.html>  
233. <https://www.census.gov/topics/population/veterans.html>  
234. <https://www.census.gov/topics/families/data.html>  
235. <https://www.census.gov/newsroom/tip-sheets.html>  
236. <https://www.census.gov/topics/housing/library.html>  
237. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html>  
238. <https://www.census.gov/data/tables.1970.html>  
239. <https://www.census.gov/topics/education/related.html>  
240. <https://www.census.gov/topics/employment/cbp-redirect.html>  
241. <https://www.census.gov/topics/health/health-care-industries.html>  
242. <https://www.census.gov/data/tables.1973.html>  
243. <https://www.census.gov/newsroom/press-kits.html>  
244. <https://www.census.gov/data/tables.1943.html>  
245. <https://www.census.gov/data/tables.2015.html>  
246. <https://www.census.gov/data/tables/2017/demo/income-poverty/p60-259.html>  
247. <https://www.census.gov/about/leadership.html>  
248. <https://www.census.gov/topics/employment/stories.html>  
249. <https://www.census.gov/topics/housing/news.html>  
250. <https://www.calendarwiz.com/calendars/calendar.php?crd=cens1sample&cid=31793>  
251. <https://www.census.gov/newsroom/press-releases.html>  
252. <https://public.govdelivery.com/accounts/USCENSUS/subscriber/new>



253. <https://www.census.gov/data/tables.1944.html>  
254. <https://www.census.gov/data.html>  
255. <https://www.census.gov/data/tables/time-series/econ/mhs/average-sales-price.html>  
256. <https://www.census.gov/data/tables.1916.html>  
257. <https://www.census.gov/programs-surveys/popest.html>  
258. <https://www.census.gov/topics/international-trade/schedule-b.html>  
259. <https://www.census.gov/data/tables/time-series/demo/health-insurance/acs-hi.html>  
260. <https://www.census.gov/topics/income-poverty/supplemental-poverty-measure.html>  
261. <https://www.census.gov/topics/business.html>  
262. <https://www.census.gov/topics/economy/news.html>  
263. <https://www.census.gov/programs-surveys/surveys-programs.html>  
264. <https://www.census.gov/data/tables.2004.html>  
265. <https://www.census.gov/data/tables.1954.html>  
266. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_820029714\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_820029714_0_accd)  
267. <https://www.census.gov/programs-surveys/popest.html/#https://www.census.gov/data.html>  
268. <https://www.census.gov/data/tables.1915.html>  
269. <https://www.census.gov/topics/public-sector/redistricting-data.html>  
270. <https://www.census.gov/topics/public-sector/congressional-apportionment.html>  
271. <https://www.census.gov/data/tables.1937.html>  
272. <https://www.census.gov/fieldjobs>  
273. <https://www.census.gov/data/tables.1919.html>  
274. <https://www.census.gov/data/tables.1994.html>  
275. <https://www.census.gov/data/tables.1952.html>  
276. <https://www.census.gov/data/tables.1898.html>  
277. <https://www.census.gov/2017censustests>  
278. <https://www.census.gov/data/tables.1982.html>  
279. <https://www.census.gov/topics/income-poverty.html>  
280. <https://www.census.gov/data/tables.2000.html>  
281. <https://www.census.gov/topics/employment/public-sector-employment-and-payroll.html>  
282. <https://www.census.gov/data/tables.1922.html>  
283. <https://www.census.gov/data/tables.1967.html>  
284. <https://www.census.gov/topics/income-poverty/news-updates.html>  
285. <https://www.census.gov/data/tables.1902.html>  
286. <https://www.census.gov/data/tables.1957.html>  
287. <https://www.census.gov/data/tables.2017.html/page/1.html>  
288. <https://www.census.gov/topics/education/educational-attainment.html>  
289. <https://www.census.gov/topics/education/teaching-about-statistics.html>  
290. <https://www.census.gov/topics/families/surveys.html>  
291. [https://www.census.gov/about/policies/quality/scientific\\_integrity.html](https://www.census.gov/about/policies/quality/scientific_integrity.html)  
292. <https://www.census.gov/programs-surveys/popest.html/#skipsideNav>  
293. <https://www.census.gov/data/tables.1986.html>  
294. <https://www.census.gov/data/tables.1923.html>  
295. <https://www.census.gov/topics/employment/surveys.html>  
296. <https://www.census.gov/topics/education/educational-services.html>  
297. <https://www.census.gov/programs-surveys/specialcensus.html>  
298. <https://www.census.gov/newsroom/facts-for-features.html>  
299. <https://www.census.gov/topics/education/visualizations.html>  
300. <https://www.census.gov/data/tables.1984.html>  
301. <https://www.census.gov/topics/business/business-characteristics.html>  
302. <https://www.census.gov/topics/housing/residential-financing.html>  
303. <https://www.census.gov/topics/population/stories.html>  
304. <https://www.census.gov/data/tables.2012.html>  
305. <https://www.census.gov/library.html>  
306. <https://www.census.gov/topics/health/hiv-aids.html>  
307. <https://www.census.gov/data/tables.1966.html>  
308. <https://www.census.gov/topics/health.html>  
309. <https://www.census.gov/topics/international-trade.html>

310. <https://www.census.gov/data/tables.1899.html>  
311. <https://www.census.gov/data/tables.1917.html>  
312. <https://www.census.gov/geography/maps.html>  
313. <https://www.census.gov/topics/population/visualizations.html>  
314. <https://www.census.gov/data/tables.1963.html>  
315. <https://www.census.gov/data/tables.1998.html>  
316. [https://www.census.gov/programs-surveys/popest.html/#NAV\\_1965438184\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html/#NAV_1965438184_0_accd)  
317. <https://www.census.gov/about/history.html>  
318. <https://www.census.gov/programs-surveys/popest.html/#skipNav>  
319. <https://www.census.gov/data/data-tools/my-tribal-area.html>  
320. <https://www.census.gov/data/tables.1939.html>  
321. <https://www.census.gov/topics/population/surveys.html>  
322. <https://www.census.gov/topics/preparedness.html>  
323. <https://www.census.gov/topics/education/about.html>  
324. <https://www.census.gov/library/photos.html>  
325. <https://www.census.gov/data/tables.1909.html>  
326. [https://www.census.gov/library/publications/time-series/statistical\\_abstracts.html](https://www.census.gov/library/publications/time-series/statistical_abstracts.html)  
327. <https://www.census.gov/data/tables.1890.html>  
328. <https://www.census.gov/programs-surveys/popest.html/#https://www.census.gov/about.html>  
329. <https://www.census.gov/programs-surveys/e-stats.html>  
330. <https://www.census.gov/about/what.html>  
331. <https://www.census.gov/about/business-opportunities.html>  
332. <https://www.census.gov/topics/families/children.html>  
333. <https://www.census.gov/data/tables.1926.html>  
334. <https://www.census.gov/en.html>  
335. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-people.html>  
336. <https://www.census.gov/glossary>  
337. <https://www.census.gov/data/tables.1903.html>  
338. <https://www.census.gov/blogs>  
339. <https://www.census.gov/data/tables.2001.html>  
340. <https://www.census.gov/data/tables.1908.html>  
341. <https://www.census.gov/data/tables.1980.html>  
342. <https://www.census.gov/data/tables.1942.html>  
343. <https://www.census.gov/data/tables.2006.html>  
344. <https://www.census.gov/data/tables.1913.html>  
345. <https://www.census.gov/topics/income-poverty/public-assistance.html>  
346. <https://www.census.gov/data/data-tools/population-clock.html>  
347. <https://www.census.gov/data/tables.All.html>  
348. <https://www.census.gov/topics/business/business-help.html>  
349. <https://www.census.gov/topics/income-poverty/well-being.html>  
350. <https://www.census.gov/data/tables.1983.html>  
351. <https://www.census.gov/topics/health/video.html>  
352. <https://www.census.gov/topics/education/publications.html>  
353. <https://www.census.gov/topics/population/genealogy.html>  
354. <https://www.census.gov/geography/about.html>  
355. <https://www.census.gov/data/tables.1992.html>  
356. <https://www.census.gov/data/tables.1912.html>  
357. <https://www.census.gov/about/partners.html>  
358. <https://www.census.gov/data/tables.1941.html>  
359. <https://www.census.gov/data/tables.1949.html>  
360. <https://www.census.gov/about/index.html>  
361. <https://www.census.gov/topics/employment/working-papers.html>  
362. <https://www.census.gov/newsroom/stories.html>  
363. <https://www.commerce.gov>  
364. <https://www.census.gov/topics/economy.html>  
365. <https://www.census.gov/topics/employment/visualizations.html>  
366. <https://www.census.gov/geography/research.html>  
367. <https://www.census.gov/topics/health/surveys.html>


368. <https://www.census.gov/topics/business/mining.html>  
369. <https://www.census.gov/topics/public-sector/news.html>  
370. <https://www.census.gov/topics/income-poverty/wealth.html>  
371. <https://www.census.gov/programs-surveys/decennial-census/2010-census.html>  
372. [https://www.census.gov/programs-surveys/popest.html#NAV\\_893921976\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html#NAV_893921976_0_accd)  
373. <https://www.census.gov/data/tables.2002.html>  
374. <https://www.census.gov/topics/health/visualizations.html>  
375. <https://www.census.gov/data/tables.2011.html>  
376. <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-people.html>  
377. <https://www.census.gov/topics/education/data.html>  
378. <https://www.census.gov/topics/international-trade/exports-imports.html>  
379. <https://www.census.gov/data/tables.1921.html>  
380. <https://www.census.gov/topics/housing/new-housing-construction.html>  
381. <https://www.census.gov/newsroom.html>  
382. <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc.html>  
383. <https://www.census.gov/data/tables.1948.html>  
384. <http://www.instagram.com/uscensusbureau>  
385. <https://www.census.gov/topics/education/videos.html>  
386. <https://www.census.gov/data/visualizations.html>  
387. <https://www.census.gov/topics/population/working-papers.html>  
388. <https://www.census.gov/topics/families/publications.html>  
389. <https://www.census.gov/data/tables.1904.html>  
390. <https://www.census.gov/topics/economy/economic-census.html>  
391. <https://www.census.gov/topics/employment/related.html>  
392. <https://www.census.gov/topics/families/fertility.html>  
393. [https://www.census.gov/programs-surveys/popest.html#NAV\\_246571490\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html#NAV_246571490_0_accd)  
394. <https://www.census.gov/data/software.html>  
395. <https://www.census.gov/data/tables.1906.html>  
396. <https://www.census.gov/data/tables.1940.html>  
397. <https://www.census.gov/datalinkage>  
398. <https://www.census.gov/geography/partnerships.html>  
399. <https://www.census.gov/data/tables.1987.html>  
400. <https://www.census.gov/topics/housing/multifamily-housing.html>  
401. <https://www.census.gov/topics/income-poverty/income.html>  
402. <https://www.census.gov/topics/international-trade/aes.html>  
403. <https://www.census.gov/geography/gss-initiative.html>  
404. <https://www.census.gov/developers>  
405. <https://www.census.gov/topics/housing/about.html>  
406. <https://www.census.gov/topics/employment/survey-of-business-owners.html>  
407. <https://www.census.gov/topics/international-trade/videos.html>  
408. <https://www.census.gov/data/tables.1962.html>  
409. <https://www.census.gov/programs-surveys/sbo.html>  
410. [https://www.census.gov/programs-surveys/popest.html#NAV\\_1529793603\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html#NAV_1529793603_0_accd)  
411. <https://www.facebook.com/uscensusbureau>  
412. <https://www.census.gov/data/tables.2003.html>  
413. [https://www.census.gov/programs-surveys/popest.html#NAV\\_1082186784\\_0\\_accd](https://www.census.gov/programs-surveys/popest.html#NAV_1082186784_0_accd)  
414. <https://www.census.gov/topics/economy/classification-codes.html>  
415. <https://www.census.gov/topics/income-poverty/guidance.html>  
416. <http://www.youtube.com/user/uscensusbureau>  
417. <https://www.census.gov/topics/income-poverty/about.html>  
418. <https://www.census.gov/data/tables.1975.html>  
419. <https://www.census.gov/data/tables.1910.html>  
420. [https://www.census.gov/about/policies/privacy/privacy-policy.html#par\\_textimage\\_1](https://www.census.gov/about/policies/privacy/privacy-policy.html#par_textimage_1)  
421. <https://www.census.gov/topics/income-poverty/poverty.html>  
422. <https://www.census.gov/data/data-tools/american-factfinder.html>  
423. <https://www.census.gov/data/tables.2007.html>  
424. <https://www.census.gov/topics/education/school-districts.html>  
425. <https://www.census.gov/data/tables.1969.html>

```
426.      https://www.census.gov/topics/education.html
427.      https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-
finc.html
428.      https://www.census.gov/data/tables.2017.html/page/2.html
429.      https://www.census.gov/library/publications.html
430.      https://www.census.gov/topics/public-sector/government-organization.html
431.      https://www.census.gov/topics/housing/housing-vacancies.html
432.      https://www.census.gov/data/tables.2010.html
433.      https://www.census.gov/data/tables.1964.html
434.      https://www.census.gov/about/our-research.html
435.      https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-
series/hic.html
436.      https://www.census.gov/topics/business/videos.html
437.      https://ask.census.gov
438.      https://www.census.gov/programs-
surveys/popest.html/#https://www.census.gov/geography.html
439.      https://www.census.gov/topics/economy/other_econ.html
440.      https://www.census.gov/data/tables.1976.html
441.      https://www.census.gov/data/tables.1920.html
442.      https://www.census.gov/data/tables.1971.html
443.      https://www.census.gov/topics/families/stories.html
444.      None
445.      [Finished in 1.8s]
```

```
Camac@DESKTOP-AVBLRGO ~/Desktop/python_intro
$ python3 task1_redo2.py
```

When print function is not enabled code runs faster than above 1.8s. See below result of 0.4s.

```
[Finished in 0.4s]
```

 web\_links.csv

10/26/2017 2:04 PM

1	web_links								
2	<a href="https://www.census.gov/topics/education/videos.html">https://www.census.gov/topics/education/videos.html</a>								
3	<a href="https://www.census.gov/topics/education/news.html">https://www.census.gov/topics/education/news.html</a>								
4	<a href="https://www.census.gov/about/our-research.html">https://www.census.gov/about/our-research.html</a>								
5	<a href="https://www.census.gov/topics/economy.html">https://www.census.gov/topics/economy.html</a>								
6	<a href="https://www.census.gov/data/tables.1897.html">https://www.census.gov/data/tables.1897.html</a>								
7	<a href="https://www.census.gov/topics/education/about.html">https://www.census.gov/topics/education/about.html</a>								
8	<a href="http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml">http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml</a>								
9	<a href="https://www.census.gov/programs-surveys/popest.html#https://www.census.gov/data.html">https://www.census.gov/programs-surveys/popest.html#https://www.census.gov/data.html</a>								
10	<a href="https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-series/hic.html">https://www.census.gov/data/tables/time-series/demo/health-insurance/historical-series/hic.html</a>								
11	<a href="https://www.census.gov/topics/education/school-enrollment.html">https://www.census.gov/topics/education/school-enrollment.html</a>								
12	<a href="https://www.census.gov/data/data-tools/my-tribal-area.html">https://www.census.gov/data/data-tools/my-tribal-area.html</a>								
13	<a href="https://www.census.gov/about/who.html">https://www.census.gov/about/who.html</a>								
14	<a href="https://www.census.gov/topics/income-poverty/supplemental-poverty-measure.html">https://www.census.gov/topics/income-poverty/supplemental-poverty-measure.html</a>								
15	<a href="https://www.census.gov/data/tables.1968.html">https://www.census.gov/data/tables.1968.html</a>								
16	<a href="https://www.census.gov/topics/population/age-and-sex.html">https://www.census.gov/topics/population/age-and-sex.html</a>								
17	<a href="https://www.census.gov/data/tables.2003.html">https://www.census.gov/data/tables.2003.html</a>								
18	<a href="https://www.census.gov/data/tables.1915.html">https://www.census.gov/data/tables.1915.html</a>								
19	<a href="https://www.census.gov/data/tables.1907.html">https://www.census.gov/data/tables.1907.html</a>								
20	<a href="https://www.census.gov/data/tables.1950.html">https://www.census.gov/data/tables.1950.html</a>								
21	<a href="https://www.census.gov/data/tables.1914.html">https://www.census.gov/data/tables.1914.html</a>								
22	<a href="https://www.census.gov/topics/income-poverty/news-updates.html">https://www.census.gov/topics/income-poverty/news-updates.html</a>								
23	<a href="https://www.census.gov/topics/international-trade/aes.html">https://www.census.gov/topics/international-trade/aes.html</a>								
24	<a href="https://www.census.gov/topics/employment/publications.html">https://www.census.gov/topics/employment/publications.html</a>								
25	<a href="https://www.census.gov/data/tables.1986.html">https://www.census.gov/data/tables.1986.html</a>								
26	<a href="https://www.census.gov/programs-surveys/specialcensus.html">https://www.census.gov/programs-surveys/specialcensus.html</a>								
27	<a href="https://www.census.gov/data/tables.1996.html">https://www.census.gov/data/tables.1996.html</a>								

## References

w3schools.com. (n.d.). *HTML <a> Tag*. Retrieved from [https://www.w3schools.com/TagS/tag\\_a.asp](https://www.w3schools.com/TagS/tag_a.asp)