



FACE FRONT,  
TRUE BELIEVERS!



*Stan Lee Quotes via Gecko&Fly*



# Web Scraping with Stan Lee

# What's our workflow?

- Import our libraries:  
requests, BeautifulSoup,  
and pandas
- Find our url: '[https://  
en.wikipedia.org/wiki/  
Stan Lee](https://en.wikipedia.org/wiki/Stan_Lee)'





Establish our connection to the web page using `requests.get(url)`.

Status code: 200 = True Believer, 403 = Doctor Doom, 404 = Savage Land

# soup = BeautifulSoup(html, 'lxml')

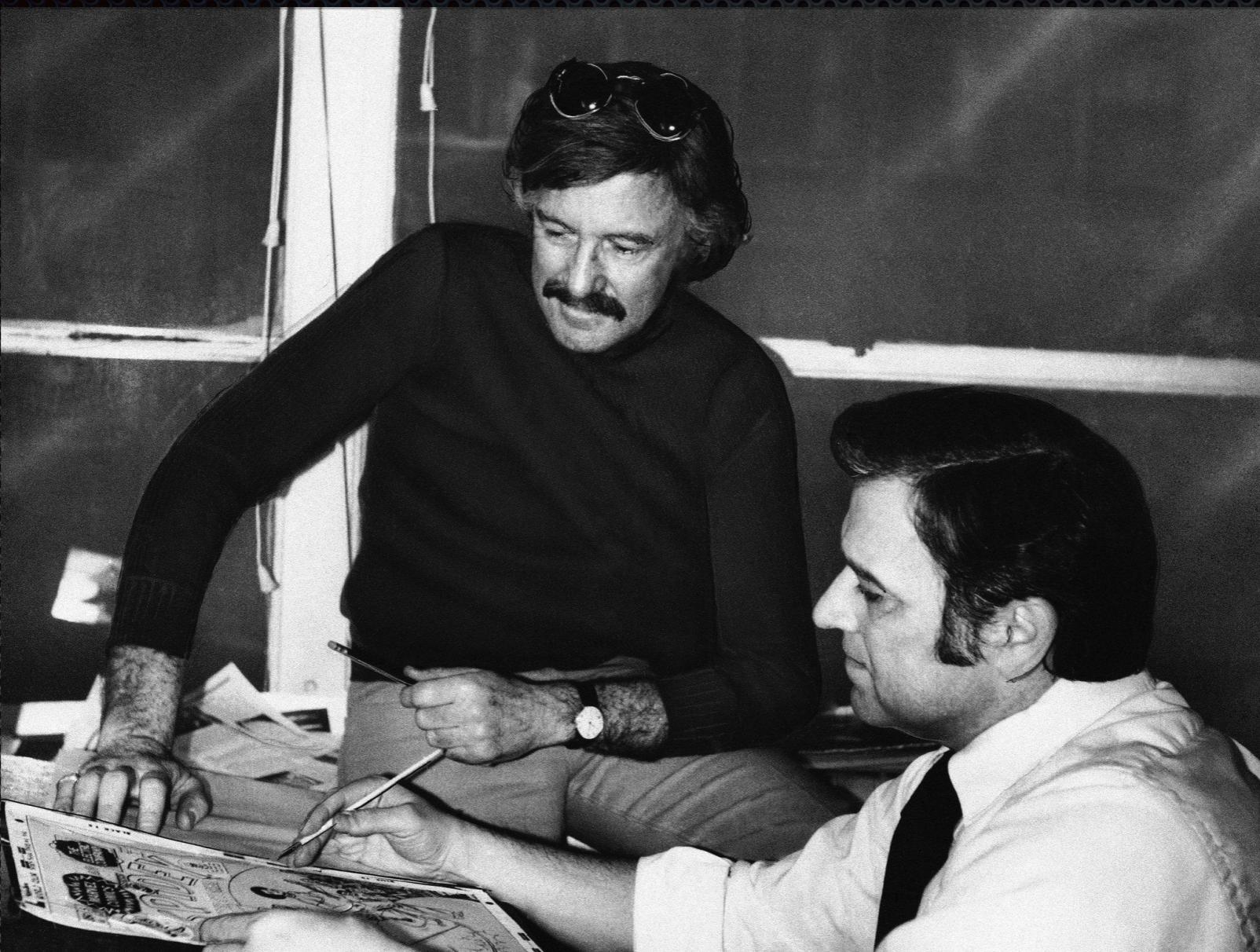
- Create a BeautifulSoup object
- Examine the website using either Mozilla or Google Chrome
- Inspect the html code and find the useful tags



- [Secret Files and Origins](#) (2002)
  - [Shazam!](#) (with Gary Frank) (2001)
- a** 158.73 × 16  
 Color #0645AD  
 Font 14px sans-serif  
 Contrast Aa 8.53

More

- [The Amazing Spider-Man](#) #1–100, 105–110, 116–118, 200, *Annual* #1–5, 18 (1962–84); (backup stories): #634–645 (2010–11)
- [The Amazing Spider-Man, strips](#) (1977–2018)<sup>[161]</sup>
- [The Avengers](#) #1–35 (1963–66)
- [Captain America](#) #100–141 (1968–71) (continues from [The Invaders](#))<sup>#99</sup>
- [Daredevil](#), #1–9, 11–50, 53, *Annual* #1 (1964–71)
- [Daredevil](#), vol. 2, #20 (backup stories)
- [Epic Illustrated](#) #1 (Silver Surfer)
- [Fantastic Four](#) #1–114, 120–121, 123–124, 126–127, 130–131, 133–134, 136–137, 140–141, 144–145, 148–149, 152–153, 156–157, 160–161, 164–165, 168–169, 172–173, 176–177, 180–181, 184–185, 188–189, 192–193, 196–197, 200–201, 204–205, 208–209, 212–213, 216–217, 220–221, 224–225, 228–229, 232–233, 236–237, 240–241, 244–245, 248–249, 252–253, 256–257, 260–261, 264–265, 268–269, 272–273, 276–277, 280–281, 284–285, 288–289, 292–293, 296–297, 300–301, 304–305, 308–309, 312–313, 316–317, 320–321, 324–325, 328–329, 332–333, 336–337, 340–341, 344–345, 348–349, 352–353, 356–357, 360–361, 364–365, 368–369, 372–373, 376–377, 380–381, 384–385, 388–389, 392–393, 396–397, 400–401, 404–405, 408–409, 412–413, 416–417, 420–421, 424–425, 428–429, 432–433, 436–437, 440–441, 444–445, 448–449, 452–453, 456–457, 460–461, 464–465, 468–469, 472–473, 476–477, 480–481, 484–485, 488–489, 492–493, 496–497, 500–501, 504–505, 508–509, 512–513, 516–517, 520–521, 524–525, 528–529, 532–533, 536–537, 540–541, 544–545, 548–549, 552–553, 556–557, 560–561, 564–565, 568–569, 572–573, 576–577, 580–581, 584–585, 588–589, 592–593, 596–597, 600–601, 604–605, 608–609, 612–613, 616–617, 620–621, 624–625, 628–629, 632–633, 636–637, 640–641, 644–645, 648–649, 652–653, 656–657, 660–661, 664–665, 668–669, 672–673, 676–677, 680–681, 684–685, 688–689, 692–693, 696–697, 700–701, 704–705, 708–709, 712–713, 716–717, 720–721, 724–725, 728–729, 732–733, 736–737, 740–741, 744–745, 748–749, 752–753, 756–757, 760–761, 764–765, 768–769, 772–773, 776–777, 780–781, 784–785, 788–789, 792–793, 796–797, 800–801, 804–805, 808–809, 812–813, 816–817, 820–821, 824–825, 828–829, 832–833, 836–837, 840–841, 844–845, 848–849, 852–853, 856–857, 860–861, 864–865, 868–869, 872–873, 876–877, 880–881, 884–885, 888–889, 892–893, 896–897, 900–901, 904–905, 908–909, 912–913, 916–917, 920–921, 924–925, 928–929, 932–933, 936–937, 940–941, 944–945, 948–949, 952–953, 956–957, 960–961, 964–965, 968–969, 972–973, 976–977, 980–981, 984–985, 988–989, 992–993, 996–997, 999–1000, 1004–1005, 1008–1009, 1012–1013, 1016–1017, 1020–1021, 1024–1025, 1028–1029, 1032–1033, 1036–1037, 1040–1041, 1044–1045, 1048–1049, 1052–1053, 1056–1057, 1060–1061, 1064–1065, 1068–1069, 1072–1073, 1076–1077, 1080–1081, 1084–1085, 1088–1089, 1092–1093, 1096–1097, 1100–1101, 1104–1105, 1108–1109, 1112–1113, 1116–1117, 1120–1121, 1124–1125, 1128–1129, 1132–1133, 1136–1137, 1140–1141, 1144–1145, 1148–1149, 1152–1153, 1156–1157, 1160–1161, 1164–1165, 1168–1169, 1172–1173, 1176–1177, 1180–1181, 1184–1185, 1188–1189, 1192–1193, 1196–1197, 1200–1201, 1204–1205, 1208–1209, 1212–1213, 1216–1217, 1220–1221, 1224–1225, 1228–1229, 1232–1233, 1236–1237, 1240–1241, 1244–1245, 1248–1249, 1252–1253, 1256–1257, 1260–1261, 1264–1265, 1268–1269, 1272–1273, 1276–1277, 1280–1281, 1284–1285, 1288–1289, 1292–1293, 1296–1297, 1299–1300, 1304–1305, 1308–1309, 1312–1313, 1316–1317, 1320–1321, 1324–1325, 1328–1329, 1332–1333, 1336–1337, 1340–1341, 1344–1345, 1348–1349, 1352–1353, 1356–1357, 1360–1361, 1364–1365, 1368–1369, 1372–1373, 1376–1377, 1380–1381, 1384–1385, 1388–1389, 1392–1393, 1396–1397, 1399–1400, 1404–1405, 1408–1409, 1412–1413, 1416–1417, 1420–1421, 1424–1425, 1428–1429, 1432–1433, 1436–1437, 1440–1441, 1444–1445, 1448–1449, 1452–1453, 1456–1457, 1460–1461, 1464–1465, 1468–1469, 1472–1473, 1476–1477, 1480–1481, 1484–1485, 1488–1489, 1492–1493, 1496–1497, 1499–1500, 1504–1505, 1508–1509, 1512–1513, 1516–1517, 1520–1521, 1524–1525, 1528–1529, 1532–1533, 1536–1537, 1540–1541, 1544–1545, 1548–1549, 1552–1553, 1556–1557, 1560–1561, 1564–1565, 1568–1569, 1572–1573, 1576–1577, 1580–1581, 1584–1585, 1588–1589, 1592–1593, 1596–1597, 1599–1600, 1604–1605, 1608–1609, 1612–1613, 1616–1617, 1620–1621, 1624–1625, 1628–1629, 1632–1633, 1636–1637, 1640–1641, 1644–1645, 1648–1649, 1652–1653, 1656–1657, 1660–1661, 1664–1665, 1668–1669, 1672–1673, 1676–1677, 1680–1681, 1684–1685, 1688–1689, 1692–1693, 1696–1697, 1699–1700, 1704–1705, 1708–1709, 1712–1713, 1716–1717, 1720–1721, 1724–1725, 1728–1729, 1732–1733, 1736–1737, 1740–1741, 1744–1745, 1748–1749, 1752–1753, 1756–1757, 1760–1761, 1764–1765, 1768–1769, 1772–1773, 1776–1777, 1780–1781, 1784–1785, 1788–1789, 1792–1793, 1796–1797, 1799–1800, 1804–1805, 1808–1809, 1812–1813, 1816–1817, 1820–1821, 1824–1825, 1828–1829, 1832–1833, 1836–1837, 1840–1841, 1844–1845, 1848–1849, 1852–1853, 1856–1857, 1860–1861, 1864–1865, 1868–1869, 1872–1873, 1876–1877, 1880–1881, 1884–1885, 1888–1889, 1892–1893, 1896–1897, 1899–1900, 1904–1905, 1908–1909, 1912–1913, 1916–1917, 1920–1921, 1924–1925, 1928–1929, 1932–1933, 1936–1937, 1940–1941, 1944–1945, 1948–1949, 1952–1953, 1956–1957, 1960–1961, 1964–1965, 1968–1969, 1972–1973, 1976–1977, 1980–1981, 1984–1985, 1988–1989, 1992–1993, 1996–1997, 1999–2000, 2004–2005, 2008–2009, 2012–2013, 2016–2017, 2020–2021, 2024–2025, 2028–2029, 2032–2033, 2036–2037, 2040–2041, 2044–2045, 2048–2049, 2052–2053, 2056–2057, 2060–2061, 2064–2065, 2068–2069, 2072–2073, 2076–2077, 2080–2081, 2084–2085, 2088–2089, 2092–2093, 2096–2097, 2099–2100, 2104–2105, 2108–2109, 2112–2113, 2116–2117, 2120–2121, 2124–2125, 2128–2129, 2132–2133, 2136–2137, 2140–2141, 2144–2145, 2148–2149, 2152–2153, 2156–2157, 2160–2161, 2164–2165, 2168–2169, 2172–2173, 2176–2177, 2180–2181, 2184–2185, 2188–2189, 2192–2193, 2196–2197, 2199–2200, 2204–2205, 2208–2209, 2212–2213, 2216–2217, 2220–2221, 2224–2225, 2228–2229, 2232–2233, 2236–2237, 2240–2241, 2244–2245, 2248–2249, 2252–2253, 2256–2257, 2260–2261, 2264–2265, 2268–2269, 2272–2273, 2276–2277, 2280–2281, 2284–2285, 2288–2289, 2292–2293, 2296–2297, 2299–2300, 2304–2305, 2308–2309, 2312–2313, 2316–2317, 2320–2321, 2324–2325, 2328–2329, 2332–2333, 2336–2337, 2340–2341, 2344–2345, 2348–2349, 2352–2353, 2356–2357, 2360–2361, 2364–2365, 2368–2369, 2372–2373, 2376–2377, 2380–2381, 2384–2385, 2388–2389, 2392–2393, 2396–2397, 2399–2400, 2404–2405, 2408–2409, 2412–2413, 2416–2417, 2420–2421, 2424–2425, 2428–2429, 2432–2433, 2436–2437, 2440–2441, 2444–2445, 2448–2449, 2452–2453, 2456–2457, 2460–2461, 2464–2465, 2468–2469, 2472–2473, 2476–2477, 2480–2481, 2484–2485, 2488–2489, 2492–2493, 2496–2497, 2499–2500, 2504–2505, 2508–2509, 2512–2513, 2516–2517, 2520–2521, 2524–2525, 2528–2529, 2532–2533, 2536–2537, 2540–2541, 2544–2545, 2548–2549, 2552–2553, 2556–2557, 2560–2561, 2564–2565, 2568–2569, 2572–2573, 2576–2577, 2580–2581, 2584–2585, 2588–2589, 2592–2593, 2596–2597, 2599–2600, 2604–2605, 2608–2609, 2612–2613, 2616–2617, 2620–2621, 2624–2625, 2628–2629, 2632–2633, 2636–2637, 2640–2641, 2644–2645, 2648–2649, 2652–2653, 2656–2657, 2660–2661, 2664–2665, 2668–2669, 2672–2673, 2676–2677, 2680–2681, 2684–2685, 2688–2689, 2692–2693, 2696–2697, 2699–2700, 2704–2705, 2708–2709, 2712–2713, 2716–2717, 2720–2721, 2724–2725, 2728–2729, 2732–2733, 2736–2737, 2740–2741, 2744–2745, 2748–2749, 2752–2753, 2756–2757, 2760–2761, 2764–2765, 2768–2769, 2772–2773, 2776–2777, 2780–2781, 2784–2785, 2788–2789, 2792–2793, 2796–2797, 2799–2800, 2804–2805, 2808–2809, 2812–2813, 2816–2817, 2820–2821, 2824–2825, 2828–2829, 2832–2833, 2836–2837, 2840–2841, 2844–2845, 2848–2849, 2852–2853, 2856–2857, 2860–2861, 2864–2865, 2868–2869, 2872–2873, 2876–2877, 2880–2881, 2884–2885, 2888–2889, 2892–2893, 2896–2897, 2899–2900, 2904–2905, 2908–2909, 2912–2913, 2916–2917, 2920–2921, 2924–2925, 2928–2929, 2932–2933, 2936–2937, 2940–2941, 2944–2945, 2948–2949, 2952–2953, 2956–2957, 2960–2961, 2964–2965, 2968–2969, 2972–2973, 2976–2977, 2980–2981, 2984–2985, 2988–2989, 2992–2993, 2996–2997, 2999–3000, 3004–3005, 3008–3009, 3012–3013, 3016–3017, 3020–3021, 3024–3025, 3028–3029, 3032–3033, 3036–3037, 3040–3041, 3044–3045, 3048–3049, 3052–3053, 3056–3057, 3060–3061, 3064–3065, 3068–3069, 3072–3073, 3076–3077, 3080–3081, 3084–3085, 3088–3089, 3092–3093, 3096–3097, 3099–3100, 3104–3105, 3108–3109, 3112–3113, 3116–3117, 3120–3121, 3124–3125, 3128–3129, 3132–3133, 3136–3137, 3140–3141, 3144–3145, 3148–3149, 3152–3153, 3156–3157, 3160–3161, 3164–3165, 3168–3169, 3172–3173, 3176–3177, 3180–3181, 3184–3185, 3188–3189, 3192–3193, 3196–3197, 3199–3200, 3204–3205, 32



Use the `.find` and `.findall` functions to pull out our text from the website. In our case, all of the titles penned by Stan the Man!

# Iterate Using Fantastic For Loops

- Or by using list comprehensions!
- Here is a small sample of our titles from Wikipedia:

The Amazing Spider-Man #1–100, 105–110, 116–118, 200, Annual #1–5, 18 (1962–84); (backup stories): #634–645 (2010–11)  
The Amazing Spider-Man, strips (1977–2018)[161]  
The Avengers #1–35 (1963–66)  
Captain America #100–141 (1968–71) (continues from Tales of Suspense #99)  
Daredevil, #1–9, 11–50, 53, Annual #1 (1964–69)  
Daredevil, vol. 2, #20 (backup story) (2001)  
Epic Illustrated #1 (Silver Surfer) (1980)  
Fantastic Four #1–114, 120–125, Annual #1–6 (1961–72); #296 (1986)  
The Incredible Hulk #1–6 (continues to Tales to Astonish #59)  
The Incredible Hulk, vol. 2, #108–120 (1968–69)  
Journey into Mystery (Thor) plotter #83–96 (1962–63), writer #97–125, Annual #1 (1963–66) (continues to Thor #126)  
The Mighty Thor #126–192, 200, Annual #2 (1966–72), 385 (1987)





# Sources:

[https://en.wikipedia.org/wiki/  
Stan\\_Lee](https://en.wikipedia.org/wiki/Stan_Lee)

[https://medium.com/analytics-  
vidhya/web-scraping-wiki-tables-  
using-beautifulsoup-and-  
python-6b9ea26d8722](https://medium.com/analytics-vidhya/web-scraping-wiki-tables-using-beautifulsoup-and-python-6b9ea26d8722)

