

Web_Scrape

November 2, 2018

0.0.1 Web Scraping using Python and BeautifulSoup

```
In [3]: import requests
```

```
In [4]: page = requests.get("http://dataquestio.github.io/web-scraping-pages/simple.html")
        page1 = requests.get("https://github.com/mohitsharma44official/Python-Web-Scraping-/b
```

```
In [5]: page
```

```
Out[5]: <Response [200]>
```

```
In [6]: page.status_code
```

```
Out[6]: 200
```

```
In [7]: page.content
```

```
Out[7]: b'<!DOCTYPE html>\n<html>\n    <head>\n        <title>A simple example page</title>\n
```

0.0.2 BeautifulSoup

```
In [9]: from bs4 import BeautifulSoup
        soup = BeautifulSoup(page.content, 'html.parser')
```

```
In [10]: soup
```

```
Out[10]: <!DOCTYPE html>
```

```
    <html>
    <head>
    <title>A simple example page</title>
    </head>
    <body>
    <p>Here is some simple content for this page.</p>
    </body>
    </html>
```

```
In [11]: print(soup.prettify())
```

```

<!DOCTYPE html>
<html>
  <head>
    <title>
      A simple example page
    </title>
  </head>
  <body>
    <p>
      Here is some simple content for this page.
    </p>
  </body>
</html>

```

```
In [12]: soup.children
```

```
Out[12]: <list_iterator at 0x11527e438>
```

```
In [13]: list(soup.children)
```

```
Out[13]: ['html', '\n', <html>
  <head>
  <title>A simple example page</title>
</head>
<body>
<p>Here is some simple content for this page.</p>
</body>
</html>]
```

```
In [17]: list(soup.children)[2]
```

```
Out[17]: <html>
  <head>
  <title>A simple example page</title>
</head>
<body>
<p>Here is some simple content for this page.</p>
</body>
</html>
```

```
In [18]: # To see the type
         [type(i) for i in list(soup.children)]
```

```
Out[18]: [bs4.element.Doctype, bs4.element.NavigableString, bs4.element.Tag]
```

```
In [20]: html = list(soup.children)[2]
```

```
In [21]: list(html.children)
```

```
Out[21]: ['\n', <head>
          <title>A simple example page</title>
          </head>, '\n', <body>
          <p>Here is some simple content for this page.</p>
          </body>, '\n']
```

```
In [25]: body = list(html.children)[3]
```

```
In [26]: body
```

```
Out[26]: <body>
          <p>Here is some simple content for this page.</p>
          </body>
```

```
In [28]: list(body.children)[1]
```

```
Out[28]: <p>Here is some simple content for this page.</p>
```

```
In [29]: p = list(body.children)[1]
          p.get_text()
```

```
Out[29]: 'Here is some simple content for this page.'
```

0.0.3 Weather Scraping (SF)

```
In [34]: Ex_page = requests.get("http://forecast.weather.gov/MapClick.php?lat=37.7772&lon=-122
```

```
In [35]: Ex_page
```

```
Out[35]: <Response [200]>
```

```
In [36]: Ex_page.status_code
```

```
Out[36]: 200
```

```
In [37]: Ex_page.content
```

```
Out[37]: b'<!DOCTYPE html>\n<html class="no-js">\n      <head>\n          <!-- Meta -->\n          <
```

```
In [40]: from bs4 import BeautifulSoup
          soup = BeautifulSoup(Ex_page.content, 'html.parser')
```

```
In [41]: print(soup.children)
```

```
<list_iterator object at 0x115306c18>
```

```
In [42]: list(soup.children)
```

```

Out[42]: ['html', '\n', <html class="no-js">
  <head>
    <!-- Meta -->
    <meta content="width=device-width" name="viewport"/>
    <link href="http://purl.org/dc/elements/1.1/" rel="schema.DC"/><title>National Weather
    <!-- Icons -->
    <link href="./images/favicon.ico" rel="shortcut icon" type="image/x-icon"/>
    <!-- CSS -->
    <link href="css/bootstrap-3.2.0.min.css" rel="stylesheet"/>
    <link href="css/bootstrap-theme-3.2.0.min.css" rel="stylesheet"/>
    <link href="css/font-awesome-4.3.0.min.css" rel="stylesheet"/>
    <link href="css/ol-4.6.4.css" rel="stylesheet" type="text/css"/>
    <link href="css/mapclick.css" rel="stylesheet" type="text/css">
    <!--[if lte IE 7]><link rel="stylesheet" type="text/css" href="css/bootstrap-ie7.css"
    <!--[if lte IE 9]><link rel="stylesheet" type="text/css" href="css/mapclick-ie.css"
    <link href="css/print.css" rel="stylesheet" type="text/css">
    <link href="css/search.css" rel="stylesheet" type="text/css">
    <!-- Javascript -->
    <script src="js/lib/modernizr-2.8.3.js" type="text/javascript"></script>
    <script src="js/lib/json3-3.3.2.min.js" type="text/javascript"></script>
    <script src="js/lib/jquery-1.11.3.min.js" type="text/javascript"></script>
    <script src="js/lib/jquery.hoverIntent-1.8.1.min.js" type="text/javascript"></script>
    <script src="js/lib/bootstrap-3.2.0.min.js" type="text/javascript"></script>
    <script src="js/lib/ol-4.6.4.js" type="text/javascript"></script>
    <!--[if lte IE 8]><script type="text/javascript" src="js/respond.min.js"></script><!--
    <script src="js/jquery.autocomplete.min.js" type="text/javascript"></script>
    <script src="js/cfisurvey/cfi.js?v2" type="text/javascript"></script>
    <script src="js/forecast.esri.js" type="text/javascript"></script>
    <script src="js/forecast.search.js" type="text/javascript"></script>
    <script src="js/forecast.openlayers.js" type="text/javascript"></script>
    <script src="js/browserSniffer.js" type="text/javascript"></script>
    <script src="js/federated-analytics.js" type="text/javascript"></script>
    <script type="text/javascript">
      (function (i, s, o, g, r, a, m) {
        i['GoogleAnalyticsObject'] = r;
        i[r] = i[r] || function () {
          (i[r].q = i[r].q || []).push(arguments)
        }, i[r].l = 1 * new Date();
        a = s.createElement(o),
          m = s.getElementsByTagName(o)[0];
        a.async = 1;
        a.src = g;
        m.parentNode.insertBefore(a, m)
      })(window, document, 'script', '//www.google-analytics.com/analytics.js'

      ga('create', 'UA-40768555-1', 'weather.gov');
      ga('set', 'anonymizeIp', true);
      ga('require', 'linkid');

```

```

        ga('send', 'pageview');
    </script>
</link></link></link></meta></head>
<body>
<main class="container">
<header class="row clearfix" id="page-header">
<a class="pull-left" href="http://www.noaa.gov" id="header-noaa"><img alt="National O
<a class="pull-left" href="http://www.weather.gov" id="header-nws"><img alt="National
<a class="pull-right" href="http://www.commerce.gov" id="header-doc"><img alt="United
</header>
<nav class="navbar navbar-default row" role="navigation">
<div class="container-fluid">
<div class="navbar-header">
<button class="navbar-toggle collapsed" data-target="#top-nav" data-toggle="collapse">
<span class="sr-only">Toggle navigation</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="collapse navbar-collapse" id="top-nav">
<ul class="nav navbar-nav">
<li><a href="http://www.weather.gov">HOME</a></li>
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
<div id="site-search">
<form action="http://search.usa.gov/search" method="get" style="margin-bottom: 0; ma
<input name="v:project" type="hidden" value="firstgov"/>
<label for="query">Search For</label>
<input id="query" name="query" size="12" type="text"/>
<input type="submit" value="Go"/>
<p>
<input checked="checked" id="nws" name="affiliate" type="radio" value="nws.noaa.gov"/
<label class="search-scope" for="nws">NWS</label>
<input id="noaa" name="affiliate" type="radio" value="noaa.gov"/>
<label class="search-scope" for="noaa">All NOAA</label>
</p>
</form>
</div>
</li>
</ul> </li>
<li class="dropdown"><a class="dropdown-toggle" data-toggle="dropdown" href="http://v
</ul>

```

```

</div>
</div>
</nav>
<div class="contentArea">
<!-- Start Forecastsearch -->
<div class="" id="fcst-search">
<form action="https://forecast.weather.gov/zipcity.php" class="form-inline" id="getF
<div id="getfcst-body">
<input class="form-control" id="inputstring" name="inputstring" placeholder="" type=
<input class="btn btn-default" id="btnSearch" name="btnSearch" type="submit" value="
<div id="txtHelp"><a href="javascript:void(window.open('http://weather.gov/ForecastS
</div>
<div id="txtError">
<div id="errorNoResults" style="display:none;">Sorry, the location you searched for v
<div id="errorMultipleResults" style="display:none;">Multiple locations were found. P
<div id="errorChoices" style="display:none;"></div>
<input id="btnCloseError" style="display:none" type="button" value="Close"/>
</div>
<div id="getfcst-head">
<p>Your local forecast office is</p>
<h3 id="getfcst-headOffice"></h3>
</div>
</form>
</div>
<!-- end Forecastsearch -->
<link href="/css/topnews.css" rel="stylesheet" type="text/css"/>
<div id="news-items">
<div id="topnews">
<div class="icon"></div>
<div class="body">
<h1 style="font-size: 11pt;">Heavy Rain and Severe Storm Potential in the East; Rain
<p>
A strong cold front will bring strong to severe thunderstorms and heavy r
<a href="http://www.wpc.ncep.noaa.gov/discussions/hpcdiscussions.php?dis
</p>
</div>
</div>
</div>
</div>
<script type="text/javascript">(function ($) { var topnews = $("#topnews"); topnews.l
<script language="javascript">document.title = $('<div/>').html('7-Day Forecast for
<div id="quickLinks">
<span class="lang-spanish"><a href="http://forecast.weather.gov/MapClick.php?lat=37.
<div class="addthis_toolbox addthis_default_style addthis-forecast">
<a class="addthis_button_compact" href="http://www.addthis.com/bookmark.php?v=250&amp
<span class="addthis_separator">|</span>
<a class="addthis_button_preferred_1"></a>
<a class="addthis_button_preferred_2"></a>
<a class="addthis_button_preferred_3"></a>

```

```

<a class="addthis_button_preferred_4"></a>
<a class="addthis_button_preferred_5"></a>
</div>
<script type="text/javascript">
    var addthis_config = addthis_config || {data_track_addressbar:true,
    $(document).ready(function(){
        jQuery.ajax({
            url: "//s7.addthis.com/js/300/addthis_widget.js#asyn
            dataType: "script",
            cache: false

        });
    });
</script>
</div>
<!-- Current Conditions -->
<div class="panel panel-default" id="current-conditions">
<!-- Current Conditions header row -->
<div class="panel-heading">
<div>
<b>Current conditions at</b>
<h2 class="panel-title">SAN FRANCISCO DOWNTOWN (SFOC1)</h2>
<span class="smallTxt"><b>Lat:ä</b>37.77056řN<b>Lon:ä</b>122.42694řW<b>Elev:ä</b>150.
</div>
</div>
<div class="panel-body" id="current-conditions-body">
<!-- Graphic and temperatures -->
<div class="pull-left" id="current_conditions-summary">
<p class="myforecast-current">NA</p>
<p class="myforecast-current-lrg">67řF</p>
<p class="myforecast-current-sm">19řC</p>
</div>
<div class="pull-left" id="current_conditions_detail">
<table>
<tr>
<td class="text-right"><b>Humidity</b></td>
<td>63%</td>
</tr>
<tr>
<td class="text-right"><b>Wind Speed</b></td>
<td>NA</td>
</tr>
<tr>
<td class="text-right"><b>Barometer</b></td>
<td>NA</td>
</tr>
<tr>
<td class="text-right"><b>Dewpoint</b></td>
<td>54řF (12řC)</td>

```

```

</tr>
<tr>
<td class="text-right"><b>Visibility</b></td>
<td>NA</td>
</tr>
<tr>
<td class="text-right"><b>Last update</b></td>
<td>
02 Nov 1:43 pm PDT
</td>
</tr>
</table>
</div>
<div id="current_conditions_station">
<div class="current-conditions-extra">
<!-- Right hand section -->
<p class="moreInfo"><b>More Information:</b></p><p><a href="https://www.weather.gov/
<!-- /current_conditions_station -->
</div>
<!-- /current-conditions-body -->
</div>
<!-- /Current Conditions -->
</div>
<!-- 7-Day Forecast -->
<div class="panel panel-default" id="seven-day-forecast">
<div class="panel-heading">
<b>Extended Forecast for</b>
<h2 class="panel-title">
San Francisco CA
</h2>
</div>
<div class="panel-body" id="seven-day-forecast-body">
<div id="seven-day-forecast-container"><ul class="list-unstyled" id="seven-day-forec
<div class="tombstone-container">
<p class="period-name">This<br/>Afternoon</p>
<p><img alt="This Afternoon: Sunny, with a high near 71. West wind around 13 mph. "
<div class="tombstone-container">
<p class="period-name">Tonight<br/><br/></p>
<p><img alt="Tonight: Mostly clear, with a low around 56. West wind 5 to 11 mph. "
<div class="tombstone-container">
<p class="period-name">Saturday<br/><br/></p>
<p><img alt="Saturday: Sunny, with a high near 69. West wind 5 to 14 mph, with gusts
<div class="tombstone-container">
<p class="period-name">Saturday<br/>Night</p>
<p><img alt="Saturday Night: Increasing clouds, with a low around 55. West southwest
<div class="tombstone-container">
<p class="period-name">Sunday<br/><br/></p>
<p><img alt="Sunday: Mostly sunny, with a high near 68. West wind around 9 mph. "
<div class="tombstone-container">
<p class="period-name">Sunday<br/>Night</p>

```



```

<p>
<p class="period-name">Monday<br/><br/></p>
<p>
<p class="period-name">Monday<br/>Night</p>
<p>
<p class="period-name">Tuesday<br/><br/></p>
<p>
// equalize forecast heights
$(function () {
    var maxh = 0;
    $(".forecast-tombstone .short-desc").each(function () {
        var h = $(this).height();
        if (h > maxh) { maxh = h; }
    });
    $(".forecast-tombstone .short-desc").height(maxh);
});
</script> </div>
</div>
<!-- Everything between 7-Day Forecast and Footer goes in this row -->
<div class="row" id="floatingDivs">
<!-- Everything on the left-hand side -->
<div class="col-md-7 col-lg-8">
<!-- Detailed Forecast -->
<div class="panel panel-default" id="detailed-forecast">
<div class="panel-heading">
<h2 class="panel-title">Detailed Forecast</h2>
</div>
<div class="panel-body" id="detailed-forecast-body">
<div class="row row-odd row-forecast"><div class="col-sm-2 forecast-label"><b>This A
</div>
<!-- /Detailed Forecast -->
<!-- Additional Forecasts and Information -->
<div class="panel panel-default" id="additional_forecasts">
<div class="panel-heading">
<h2 class="panel-title">Additional Forecasts and Information</h2>
</div>
<div class="panel-body" id="additional-forecasts-body">
<p class="myforecast-location"><a href="MapClick.php?zoneid=CAZ006">Zone Area Forecas
<!-- First nine-ten links -->
<div id="linkBlockContainer">
<div class="linkBlock">
<ul class="list-unstyled">
<li><a href="http://forecast.weather.gov/product.php?site=MTR&amp;issuedby=MTR&amp;p
<li><a href="MapClick.php?lat=37.7772&amp;lon=-122.4168&amp;unit=0&amp;lg=english&amp;

```

```

<li><a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&am
</ul>
</div>
<div class="linkBlock">
<ul class="list-unstyled">
<li><a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&am
<li><a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&am
<!-- <li><a href="afm/PointClick.php?lat=37.7772&lon=-122.4168">Quick Forecast</a></
</ul>
</div>
<div class="linkBlock">
<ul class="list-unstyled">
<li><a href="http://weather.gov/aq/probe_aq_data.php?latitude=37.7772&longitude=
<li><a href="MapClick.php?lat=37.7772&lon=-122.4168&FcstType=text&unit=1
<li><a href="http://www.srh.weather.gov/srh/jetstream/webweather/pinpoint_max.htm">A
<li><a href="http://www.wrh.noaa.gov/forecast/wxtables/index.php?lat=37.7772&lon
</ul>
</div>
<!-- /First nine-ten links -->
<!-- Additional links -->
<div class="linkBlock"><ul class="list-unstyled"><li><a href="http://www.wrh.noaa.gov
</div> <!-- /linkBlockContainer -->
</div><!-- /additional-forecasts-body-->
</div> <!-- /additional_forecasts -->
</div> <!-- /Everything on the left-hand side -->
<!-- right-side-data -->
<div class="col-md-5 col-lg-4" id="right-side-data">
<div id="mapAndDescriptionArea">
<!-- openlayer map -->
<style>
#custom-search{
display: block;
position: relative;
z-index: 50;
top: 52px;
left: 60px;
}
#esri-geocoder-search{
display: block;
position: relative;
z-index: 50;
top: 52px;
left: 60px;
}
#emap{
margin-top:15px;
cursor:pointer;
height:370px;

```

```

width:100%;
border: 1px solid #ccc;
border-radius: 3px;
}
#switch-basemap-container{
}
#basemap-selection-form ul{
list-style: none;
margin: 0px;
}
#basemap-selection-form li{
float: left;
}
.disclaimer{
margin-top:350px;
margin-left: 5px;
z-index: 100;
position: absolute;
text-transform: none;
}
.esriAttributionLastItem{
text-transform: none;
}
.esriSimpleSlider div{
height:22px;
line-height:20px;
width:20px;
}
#point-forecast-map-label {
text-align:center;
font-weight:bold;
color:black;
}
@media (max-width: 767px) {
#emap{
margin-top:.5em;
height:270px;
}
.disclaimer{
margin-top:250px;
}
}
</style>
<!-- forecast-map -->
<div class="point-forecast-map">
<div class="point-forecast-map-header text-center">
<div id="toolbar">
<div id="switch-basemap-container">

```

```

<div id="basemap-selection-form" title="Choose a Basemap">
<div id="basemap-menu">
<select autocomplete="off" id="basemap-selected" name="basemap-selected" title="Basemap">
<option value="none">Select Basemap</option>
<option selected="" value="topo">Topographic</option>
<option value="streets">Streets</option>
<option value="satellite">Satellite</option>
<option value="ocean">Ocean</option>
</select>
</div>
</div>
<div id="point-forecast-map-label">
    Click Map For Forecast
</div>
</div><!-- //switch-basemap-container -->
<div style="clear:both;"></div>
</div><!-- //toolbar -->
</div><!-- //.point-forecast-map-header -->
<div id="emap">
<noscript><center><br/><br/><b>Map function requires Javascript and a compatible browser</b></noscript>
<div class="disclaimer"><a href="http://www.weather.gov/disclaimer#esri">Disclaimer</a></div><!-- //emap -->
<div class="point-forecast-map-footer">

</div><!-- //.point-forecast-map-footer -->
</div> <!-- //.point-forecast-map -->
<!-- //forecast-map -->
<!-- //openlayer map -->
<!-- About this Forecast -->
<div id="about_forecast">
<div class="fullRow">
<div class="left">Point Forecast:</div>
<div class="right">San Francisco CA<br/>37.77°N 122.41°W (Elev. 131 ft)</div>
</div>
<div class="fullRow">
<div class="left"><a href="http://www.weather.gov/glossary/index.php?word=Last+updated">Last updated</a></div>
<div class="right">1:46 pm PDT Nov 2, 2018</div>
</div>
<div class="fullRow">
<div class="left"><a href="http://www.weather.gov/glossary/index.php?word=forecast+validity">Forecast validity</a></div>
<div class="right">2pm PDT Nov 2, 2018-6pm PST Nov 8, 2018</div>
</div>
<div class="fullRow">
<div class="left"><a href="https://forecast.weather.gov/product.php?site=MTR&issued=201811021546">Forecast</a></div>
</div>
<div class="fullRow">
<div class="left"><a href="https://forecast.weather.gov/product.php?site=MTR&issued=201811021546">Forecast</a></div>
</div>

```

```

<div class="right">
<a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&Fc
<a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&Fc
</div>
</div>
</div>
<!-- /About this Forecast -->
</div>
<!--additionalForecast-->
<div class="panel panel-default" id="additionalForecast">
<div class="panel-heading">
<h2 class="panel-title">Additional Resources</h2>
</div>
<div class="panel-body">
<!-- Radar & Satellite Images -->
<div class="subItem" id="radar">
<h4>Radar & Satellite Image</h4>
<a href="http://radar.weather.gov/radar.php?rid=mux&product=NOR&overlay=1110
<!-- /Radar & Satellite Images -->
<!-- Hourly Weather Forecast -->
<div class="subItem" id="feature">
<h4>Hourly Weather Forecast</h4>
<a href="MapClick.php?lat=37.7772&lon=-122.4168&unit=0&lg=english&Fc
</div>
<!-- /Hourly Weather Forecast -->
<!-- NDFD -->
<div class="subItem" id="NDFD">
<h4>National Digital Forecast Database</h4>
<div class="one-sixth-first"><a href="http://graphical.weather.gov/sectors/pacsouthw
<p><a href="http://graphical.weather.gov/sectors/pacsouthwest.php?element=MaxT">High
<p><a href="http://graphical.weather.gov/sectors/pacsouthwest.php?element=Wx">Chance
<!-- /NDFD -->
</div>
</div>
<!-- /additionalForecast -->
</div>
<!-- /col-md-4 -->
<!-- /right-side-data -->
<script language="javascript">$( document ).ready(function() { load_openlayers_map('
<!-- /row -->
</img></div>
<!-- /PageFormat-Land -->
</main></body></html>, '\n', <footer>
<div class="sitemap-content row" id="sitemap">
<div class="col-xs-12">
<div class="sitemap-columns">
<div class="sitemap-section">
<div class="panel-heading">

```

```

<a class="sitemap-section-heading" href=" http://www.nws.noaa.gov/climate/">PAST WEATHER
<button class="menu-toggle pull-right" data-target="#sitemap-1" data-toggle="collapse"
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-1">
<ul class="list-unstyled">
<li><a href=" http://www.nws.noaa.gov/climate/">Past Weather </a></li>
<li><a href=" http://www.cpc.ncep.noaa.gov/products/MD_index.shtml">Climate Monitoring
<li><a href=" http://www.nws.noaa.gov/climate/">Heating/Cooling Days </a></li>
<li><a href=" http://www.nws.noaa.gov/climate/">Monthly Temps </a></li>
<li><a href=" http://www.nws.noaa.gov/climate/">Records </a></li>
<li><a href=" http://aa.usno.navy.mil/">Astronomical Data </a></li>
<li><a href="https://www.climate.gov/maps-data/dataset/past-weather-zip-code-data-table">
</li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://alerts.weather.gov">ACTIVE ALERTS</a>
<button class="menu-toggle pull-right" data-target="#sitemap-2" data-toggle="collapse"
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-2">
<ul class="list-unstyled">
<li><a href=" http://alerts.weather.gov">Warnings By State</a></li>
<li><a href=" http://www.wpc.ncep.noaa.gov/ww.shtml">Excessive Rainfall and Winter Weather
<li><a href="http://water.weather.gov/ahps/?current_color=flood&current_type=all">
<li><a href=" http://www.weather.gov">Latest Warnings</a></li>
<li><a href=" http://www.spc.noaa.gov/products/outlook/">Thunderstorm/Tornado Outlook
<li><a href=" http://www.nhc.noaa.gov/">Hurricanes </a></li>
<li><a href=" http://www.spc.noaa.gov/products/fire_wx/">Fire Weather Outlooks </a></li>
<li><a href=" http://www.cpc.ncep.noaa.gov/products/stratosphere/uv_index/uv_alert.shtml">
<li><a href=" http://www.drought.gov/">Drought </a></li>
<li><a href="http://www.swpc.noaa.gov/products/alerts-watches-and-warnings">Space Weather
<li><a href=" http://www.nws.noaa.gov/nwr/">NOAA Weather Radio </a></li>
<li><a href=" http://alerts.weather.gov/">NWS CAP Feeds </a></li>
</ul>
</div>
</div>
<div class="sitemap-section">

```

```

<div class="panel-heading">
<a class="sitemap-section-heading" href="http://www.weather.gov/current">CURRENT COND
<button class="menu-toggle pull-right" data-target="#sitemap-3" data-toggle="collapse"
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-3">
<ul class="list-unstyled">
<li><a href="http://www.weather.gov/Radar">Radar </a></li>
<li><a href="http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/">Climate Mon
<li><a href="http://water.weather.gov/ahps/">River Levels </a></li>
<li><a href="http://water.weather.gov/precip/">Observed Precipitation </a></li>
<li><a href="http://www.nws.noaa.gov/om/osd/portal.shtml">Surface Weather </a></li>
<li><a href="http://www.spc.noaa.gov/obswx/maps/">Upper Air </a></li>
<li><a href="http://www.ndbc.noaa.gov/">Marine and Buoy Reports </a></li>
<li><a href="http://www.nohrsc.noaa.gov/interactive/html/map.html">Snow Cover </a></li>
<li><a href="http://www.weather.gov/satellite">Satellite </a></li>
<li><a href="http://www.swpc.noaa.gov/">Space Weather </a></li>
<li><a href="http://www.weather.gov/pr">International Observations</a></li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://weather.gov/forecastmaps">FORECAST</a>
<button class="menu-toggle pull-right" data-target="#sitemap-4" data-toggle="collapse"
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-4">
<ul class="list-unstyled">
<li><a href="http://www.weather.gov/">Local Forecast </a></li>
<li><a href="http://www.weather.gov/pr">International Forecasts</a></li>
<li><a href="http://www.spc.noaa.gov/">Severe Weather </a></li>
<li><a href="http://www.wpc.ncep.noaa.gov/">Current Outlook Maps </a></li>
<li><a href="http://www.cpc.ncep.noaa.gov/products/Drought">Drought </a></li>
<li><a href="http://www.weather.gov/fire">Fire Weather </a></li>
<li><a href="http://www.wpc.ncep.noaa.gov/">Fronts/Precipitation Maps </a></li>
<li><a href="http://www.nws.noaa.gov/forecasts/graphical/">Current Graphical Forecas
<li><a href="http://water.weather.gov/ahps/forecasts.php">Rivers </a></li>
<li><a href="http://www.nws.noaa.gov/om/marine/home.htm">Marine </a></li>
<li><a href="http://www.opc.ncep.noaa.gov/marine_areas.php">Offshore and High Seas</a></li>

```

```

<li><a href=" http://www.nhc.noaa.gov/">Hurricanes </a></li>
<li><a href=" http://aviationweather.gov">Aviation Weather </a></li>
<li><a href="http://www.cpc.ncep.noaa.gov/products/OUTLOOKS_index.shtml">Climatic Out
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://www.weather.gov/informationcenter">I
<button class="menu-toggle pull-right" data-target="#sitemap-5" data-toggle="collaps
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-5">
<ul class="list-unstyled">
<li><a href=" http://www.spaceweather.gov">Space Weather </a></li>
<li><a href="http://www.weather.gov/briefing/">Daily Briefing </a></li>
<li><a href=" http://www.nws.noaa.gov/om/marine/home.htm">Marine </a></li>
<li><a href="http://www.nws.noaa.gov/climate">Climate </a></li>
<li><a href="http://www.weather.gov/fire">Fire Weather </a></li>
<li><a href=" http://www.aviationweather.gov/">Aviation </a></li>
<li><a href="http://mag.ncep.noaa.gov/">Forecast Models </a></li>
<li><a href="http://water.weather.gov/ahps/">Water </a></li>
<li><a href="https://www.weather.gov/gis/">GIS</a></li>
<li><a href=" http://www.nws.noaa.gov/om/coop/">Cooperative Observers </a></li>
<li><a href="https://www.weather.gov/skywarn/">Storm Spotters </a></li>
<li><a href="http://www.tsunami.gov">Tsunami Warning System</a></li>
<li><a href="http://water.noaa.gov/">National Water Center</a></li>
<li><a href="http://www.weather.gov/pr">International Weather</a></li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://weather.gov/safety">WEATHER SAFETY</a>
<button class="menu-toggle pull-right" data-target="#sitemap-6" data-toggle="collaps
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-6">
<ul class="list-unstyled">
<li><a href="http://www.weather.gov/nwr/">NOAA Weather Radio</a></li>

```



```

<li><a href="http://www.weather.gov/stormready/">StormReady</a></li>
<li><a href="http://www.nws.noaa.gov/om/heat/index.shtml">Heat </a></li>
<li><a href=" http://www.lightningsafety.noaa.gov/">Lightning </a></li>
<li><a href=" http://www.nhc.noaa.gov/prepare/">Hurricanes </a></li>
<li><a href="http://www.nws.noaa.gov/om/thunderstorm/">Thunderstorms </a></li>
<li><a href="https://www.weather.gov/safety/tornado">Tornadoes </a></li>
<li><a href=" http://www.ripcurrents.noaa.gov/">Rip Currents </a></li>
<li><a href="http://www.floodsafety.noaa.gov/">Floods </a></li>
<li><a href="https://www.weather.gov/safety/tsunami">Tsunamis</a></li>
<li><a href="https://www.weather.gov/tsunamiready/">TsunamiReady</a></li>
<li><a href=" http://www.weather.gov/om/winter/index.shtml">Winter Weather </a></li>
<li><a href="http://www.nws.noaa.gov/om/heat/uv.shtml">Ultra Violet Radiation </a></li>
<li><a href=" http://www.weather.gov/airquality/">Air Quality </a></li>
<li><a href=" http://www.weather.gov/om/hazstats.shtml">Damage/Fatality/Injury Statis</li>
<li><a href=" http://www.redcross.org/">Red Cross </a></li>
<li><a href=" http://www.fema.gov/">Federal Emergency Management Agency (FEMA) </a></li>
<li><a href=" http://www.weather.gov/om/brochures.shtml">Brochures </a></li>
<li><a href="http://www.nws.noaa.gov/os/marine/safeboating/">Safe Boating</a></li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://weather.gov/news">NEWS</a>
<button class="menu-toggle pull-right" data-target="#sitemap-7" data-toggle="collapse">
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-7">
<ul class="list-unstyled">
<li><a href=" http://weather.gov/news">Newsroom</a></li>
<li><a href=" http://weather.gov/socialmedia">Social Media </a></li>
<li><a href="http://www.nws.noaa.gov/com/weatherreadynation/calendar.html">Events</a></li>
<li><a href=" http://www.weather.gov/om/brochures.shtml">Pubs/Brochures/Booklets </a></li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://weather.gov/owlie">EDUCATION</a>
<button class="menu-toggle pull-right" data-target="#sitemap-8" data-toggle="collapse">
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>

```

```

</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-8">
<ul class="list-unstyled">
<li><a href="http://weather.gov/owlie">NWS Education Home</a></li>
<li><a href="http://www.nws.noaa.gov/com/weatherreadynation/force.html">Be A Force o
<li><a href=" http://www.education.noaa.gov/Weather_and_Atmosphere/">NOAA Education I
<li><a href=" http://www.weather.gov/glossary/">Glossary </a></li>
<li><a href="https://www.weather.gov/jetstream/">JetStream </a></li>
<li><a href=" http://www.weather.gov/training/">NWS Training Portal </a></li>
<li><a href="https://library.noaa.gov/">NOAA Library </a></li>
<li><a href="http://weather.gov/owlie">For Students, Parents and Teachers</a></li>
<li><a href="http://www.weather.gov/owlie/publication_brochures">Brochures </a></li>
</ul>
</div>
</div>
<div class="sitemap-section">
<div class="panel-heading">
<a class="sitemap-section-heading" href="http://weather.gov/about">ABOUT</a>
<button class="menu-toggle pull-right" data-target="#sitemap-9" data-toggle="collaps
<span class="sr-only">Toggle menu</span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
<span class="icon-bar"></span>
</button>
</div>
<div class="sitemap-section-body panel-body collapsable collapse" id="sitemap-9">
<ul class="list-unstyled">
<li><a href="http://weather.gov/organization">Organization </a></li>
<li><a href=" http://www.weather.gov/sp/">Strategic Plan </a></li>
<li><a href="https://sites.google.com/a/noaa.gov/nws-best-practices/">For NWS Employ
<li><a href="https://www.weather.gov/international/">International </a></li>
<li><a href="http://www.ncep.noaa.gov/">National Centers </a></li>
<li><a href=" http://www.weather.gov/tg/">Products and Services </a></li>
<li><a href="http://www.weather.gov/careers/">Careers</a></li>
<li><a href=" http://www.weather.gov/glossary/">Glossary </a></li>
<li><a href="http://weather.gov/contact">Contact Us </a></li>
</ul>
</div>
</div>
</div>
</div>
</div>
<!-- legal footer area -->
<div class="footer-legal">
<div class="col-xs-12 col-sm-2 col-md-2" id="footerLogo">
<a href="http://www.usa.gov">
<ul class="list-unstyled footer-legal-content">
<li><a href="http://www.commerce.gov">US Dept of Commerce</a></li>
<li><a href="http://www.noaa.gov">National Oceanic and Atmospheric Administration</a>
<li><a href="http://www.weather.gov">National Weather Service</a></li>
<li><a href="http://www.weather.gov/mtr">San Francisco Bay Area, CA</a></li><li><br>
</div>
<div class="col-xs-12 col-sm-3 col-md-3">
<ul class="list-unstyled">
<li><a href="http://www.weather.gov/disclaimer">Disclaimer</a></li>
<li><a href="http://www.cio.noaa.gov/services_programs/info_quality.html">Information
<li><a href="http://www.weather.gov/help">Help</a></li>
<li><a href="http://www.weather.gov/glossary">Glossary</a></li>
</ul>
</div>
<div class="col-xs-12 col-sm-3 col-md-3">
<ul class="list-unstyled">
<li><a href="http://www.weather.gov/privacy">Privacy Policy</a></li>
<li><a href="http://www.rdc.noaa.gov/~foia">Freedom of Information Act (FOIA)</a></li>
<li><a href="http://www.weather.gov/about">About Us</a></li>
<li><a href="http://www.weather.gov/careers">Career Opportunities</a></li>
</ul>
</div>
</div>
</footer>, '\n', '\n', '\n', '\n']

```

```
In [43]: seven_day = soup.find(id="seven-day-forecast")
```

```
In [46]: forecast_items = seven_day.find_all(class_="tombstone-container")
```

```
In [47]: print(forecast_items)
```

```

[<div class="tombstone-container">
<p class="period-name">This<br/>Afternoon</p>
<p><img alt="This Afternoon: Sunny, with a high near 71. West wind around 13 mph. " class="fore
<p class="period-name">Tonight<br/><br/></p>
<p><img alt="Tonight: Mostly clear, with a low around 56. West wind 5 to 11 mph. " class="fore
<p class="period-name">Saturday<br/><br/></p>
<p><img alt="Saturday: Sunny, with a high near 69. West wind 5 to 14 mph, with gusts as high as
<p class="period-name">Saturday<br/>Night</p>
<p><img alt="Saturday Night: Increasing clouds, with a low around 55. West southwest wind 8 to
<p class="period-name">Sunday<br/><br/></p>
<p><img alt="Sunday: Mostly sunny, with a high near 68. West wind around 9 mph. " class="forec
<p class="period-name">Sunday<br/>Night</p>
<p>Monday<br/><br/></p>
<p>Monday<br/>Night</p>
<p>Tuesday<br/><br/></p>
<p>
  <p class="period-name">
    This
  <br/>
  Afternoon
</p>
<p>
  <img alt="This Afternoon: Sunny, with a high near 71. West wind around 13 mph. " class="fore
</p>
  <p class="short-desc">
    Sunny
  </p>
  <p class="temp temp-high">
    High: 71 ̊F
  </p>
</div>
```

```
In [49]: period = tonight.find(class_ = "period-name").get_text()
        short_Desc = tonight.find(class_ = "short-desc").get_text()
        temp = tonight.find(class_ = "temp temp-high").get_text()

        print(period)
        print(short_Desc)
        print(temp)
```

```
ThisAfternoon
Sunny
High: 71 ̊F
```

```
In [54]: img = tonight.find('img')
        print(img.prettify())
```

```
<img alt="This Afternoon: Sunny, with a high near 71. West wind around 13 mph. " class="forecas
```

```
In [55]: desc =img['title']
        print(desc)
```

```
This Afternoon: Sunny, with a high near 71. West wind around 13 mph.
```

```
In [56]: period_tags = seven_day.select(".tombstone-container .period-name")
periods = [pt.get_text() for pt in period_tags]
periods
```

```
Out[56]: ['ThisAfternoon',
'Tonight',
'Saturday',
'SaturdayNight',
'Sunday',
'SundayNight',
'Monday',
'MondayNight',
'Tuesday']
```

```
In [57]: short_descs = [sd.get_text() for sd in seven_day.select(".tombstone-container .short-desc")]
temps = [t.get_text() for t in seven_day.select(".tombstone-container .temp")]
descs = [d["title"] for d in seven_day.select(".tombstone-container img")]
```

```
print(short_descs)
print(temps)
print(descs)
```

```
['Sunny', 'Mostly Clear', 'Sunny', 'IncreasingClouds', 'Mostly Sunny', 'Partly Cloudy', 'Mostly Sunny',
'High: 71 ̊F', 'Low: 56 ̊F', 'High: 69 ̊F', 'Low: 55 ̊F', 'High: 68 ̊F', 'Low: 56 ̊F', 'High: 67 ̊F',
'This Afternoon: Sunny, with a high near 71. West wind around 13 mph. ', 'Tonight: Mostly clear, with a low around 56. West wind around 13 mph. ', 'Saturday: Sunny, with a high near 69. West wind around 13 mph. ', 'Saturday Night: Increasing clouds, with a low around 55. West wind around 13 mph. ', 'Sunday: Mostly sunny, with a high near 68. West wind around 13 mph. ', 'Sunday Night: Partly cloudy, with a low around 56. West wind around 13 mph. ', 'Monday: Mostly sunny, with a high near 68. West wind around 13 mph. ', 'Monday Night: Mostly clear, with a low around 55. West wind around 13 mph. ', 'Tuesday: Sunny, with a high near 67. West wind around 13 mph. ']
```

```
In [58]: import pandas as pd
weather = pd.DataFrame({
    "period": periods,
    "short_desc": short_descs,
    "temp": temps,
    "desc":descs
})
weather
```

```
Out[58]:
```

	period	short_desc	temp \
0	ThisAfternoon	Sunny	High: 71 ̊F
1	Tonight	Mostly Clear	Low: 56 ̊F
2	Saturday	Sunny	High: 69 ̊F
3	SaturdayNight	IncreasingClouds	Low: 55 ̊F
4	Sunday	Mostly Sunny	High: 68 ̊F
5	SundayNight	Partly Cloudy	Low: 56 ̊F
6	Monday	Mostly Sunny	High: 68 ̊F
7	MondayNight	Mostly Clear	Low: 55 ̊F
8	Tuesday	Sunny	High: 67 ̊F

	desc
0	This Afternoon: Sunny, with a high near 71. We...

```

1 Tonight: Mostly clear, with a low around 56. W...
2 Saturday: Sunny, with a high near 69. West win...
3 Saturday Night: Increasing clouds, with a low ...
4 Sunday: Mostly sunny, with a high near 68. Wes...
5 Sunday Night: Partly cloudy, with a low around...
6 Monday: Mostly sunny, with a high near 68.
7 Monday Night: Mostly clear, with a low around 55.
8 Tuesday: Sunny, with a high near 67.

```

```

In [59]: temp_nums = weather["temp"].str.extract("(?P<temp_num>\d+)", expand=False)
weather["temp_num"] = temp_nums.astype('int')
temp_nums

```

```

Out[59]: 0    71
1     56
2     69
3     55
4     68
5     56
6     68
7     55
8     67
Name: temp_num, dtype: object

```

```

In [61]: is_night = weather["temp"].str.contains("Low")
weather["is_night"] = is_night
is_night

```

```

Out[61]: 0    False
1     True
2    False
3     True
4    False
5     True
6    False
7     True
8    False
Name: temp, dtype: bool

```

```

In [62]: weather[is_night]

```

```

Out[62]:
   period      short_desc  temp \
1  Tonight      Mostly Clear  Low: 56 ¨F
3 SaturdayNight IncreasingClouds  Low: 55 ¨F
5  SundayNight    Partly Cloudy  Low: 56 ¨F
7  MondayNight    Mostly Clear  Low: 55 ¨F

                                desc  temp_num  is_night
1  Tonight: Mostly clear, with a low around 56. W...      56      True

```

3	Saturday Night: Increasing clouds, with a low ...	55	True
5	Sunday Night: Partly cloudy, with a low around...	56	True
7	Monday Night: Mostly clear, with a low around 55.	55	True