**Wiki Map**

In this project, you will use Python to parse Wiki pages and map the contents. Wikipedia has many 'List of" web pages that contain lists of points of interest (POI), such as museums or airports. E.g.,

<http://en.wikipedia.org/wiki/List_of_museums_in_North_Carolina>

<http://en.wikipedia.org/wiki/List_of_airports_in_North_Carolina>

<http://en.wikipedia.org/wiki/List_of_North_Carolina_state_parks>

The pages contain tables with POI attributes and links to the individual pages about each POI.  The individual pages have 'coordinates' (Lat/long) for each POI.  E.g., see the box on the right on this page:

<http://en.wikipedia.org/wiki/82nd_Airborne_Division_War_Memorial_Museum>

The code should be capable of taking one or more of these listing pages as input. Alternatively, the code should take a set of search terms such as ‘museums North Carolina’ and use this to find the url for the wiki page with a list of these (<http://en.wikipedia.org/wiki/List_of_museums_in_North_Carolina> in our example) or report that this page could not be found.

Next it should automatically parse all of the POIs in a page and bring them into a GIS dataset, using the linked subpage to extract the point coordinates (where available).   You would then need to figure out how to parse these pages, using BeautifulSoup.  Then you would map the result in ArcMap (with a basemap underneath to show context) and create an HTML report with a screenshot of the results and a summary of the output.    There will be a number of challenges in this, but it should be fun.   You could do this for specific specific pages or make it flexible so that the user could choose a search category (NC museums, NC airports, NC parks,...) and then it would map the data from the appropriate wiki page.

For this project, you can start looking at the data now and formulating your approach, but you may need to wait a few lessons until you can really start working on the code.  (We'll go over file I/O soon.  Later we will discuss fetching webpages and parsing the content with a module called BeautifulSoup,