Requirements

This project entails scraping real estate property data for houses. The script will need to take a parcel id as input and for each parcel id, scrap the property data.

The script should use the following software:

The scripts need be able run with Python 2.7.6 in a Linux env. The scraping should be done using Requests API version 2.13.0 with BeautifulSoup version 3.2.1. (No Selenium)

All output should be output to a csv file.

    i. For each property(parcel id), scrape property details from:
       http://qpublic9.qpublic.net/ga_search_dw.php?county=ga_fulton
   ii. Search steps:
        1. Select "search by Parcel Number"
        2. Input parcel number (sample parcel numbers below)
        3. Select parcel
  iii. Once the property page is returned, scrape the following elements and put in a property.csv file

        i.

Output Name	Website Element
ParcelID	Parcel Number
Location Address	Location Address
PropertyType	Property Class
Neighborhood	Neighborhood
Acreage	Acres
Exemption Codes	Exceptions
CurrentOwner	Owner Name (under Owner section)
OwnerAddress	Owner Address (under Owner section)
streetAddress	street (parse street from owner address)
city	city (parse city from owner address)
state	state (parse state from owner address)
zipcode	zipcode (parse zip from owner address)
If Commercial Improvement section available	
structure	Structure
Units	Units
Year Built	Year Built
Sq ft	Total Square Footage
If Residential Improvement section is available	
stories	Stories
construction	Exterior wall
style	style

Year built	Year built
Sq Ft	Res Sq Ft
basement	basement
Full Bath	Full Baths/Half Bath (text before slash)
Half Bath	Fll Baths/Half Bath (text after slash)
bedrooms	bedrooms
value	Total value (in first column) in Valuation section

    ii.

   iii. Create a seperate sales.csv from the sales section with each row being a separate row in the file with the following data

Output Name	Website Element
SalesDate – (output format should be mm/dd/yyyy)	Sale Date
SalesPrice	Sale Price
SalesType	Sales Validity (some data may have special chars, i.e 'Sale < = 1000')
Buyer	Grantee
Seller	Grantor

Sample input (parcel numbers)

14-0087-0003-023-6
14-0062- LL-193-0
17-0076-0001-003-6
09C-1100-0043-032-7
09F-3804-0178-025-0
13-0068- LL-178-7
14-0044-0002-135-1
17 -0224-0004-058-8