api-scrapin Documentation

Release 0.0.1

Benjamin M. Gross

May 19, 2015

Contents

Contents:

README.md

1.1 Introduction

Zillow aggregates some very interesting data, especially if you're interested aspects of home prices, demographics such as income, education, etc. – not to mention all of this information is provided with latitude and longitude coordinates to boot. This light-weight module takes advantage of that gives you the ability

1. Take the 'n' largest cities in the US (data scraped from Wikipedia)

2. Go to Zillow's API, & extract `regionid`'s (of which there are several hundred for any metropolitan city) along with some interesting Zillow Index Data

3. Join that data with demographic data, such as median house prices, cost per square foot, median income, etc. **all** provided at a "neighborhood", "city", & "state" level.

4. Put it all into Jesus' favorite data structure... pandas, to do some more interesting data analysis

1.2 Dependencies:

- Requests: Leveraged heavily to hit the Zillow API, as well as pass the API arguments

- BeautifulSoup: Specifically `bs4`, for parsing the horrific, sadder-than-baby-tears expunged `xml` from the Zillow API

1.3 Installation:

```
$ git clone git@github.com:benjaminmgross/api-scrapin.git #assuming ssh install
$ cd api-scrapin
$ python setup.py install
```

I know what you're thinking, "why can't I `pip install` it?" Stop whining! ... fine, I haven't figured out how to do that yet with packages, but I'm working on it...

1.4 Up and Running in 5 Steps

1.4.1 Step 1: Get Yourself a Zillow API Token

1. Go to Zillow's Registration Page where you will be prompted to create a login.

2. After you create a login, go to the Zillow API Overview Page

3. Click on the get a ZWSID

4. Fill out the information, click all of the check boxes of different APIs you might wannt, and then get ready to receive your Zillow API key in your inbox!

1.4.2 Step 2: Install the Package

See *installation instructions*

1.4.3 Step 3: Let 'er Rip

The crux of what makes this package special is the ability to merge what are called "region-id" and cities.

For instance, there are 267 `region-id`'s around the New York City area, and for each one of those `region-id`'s, there's extensive demographic information (such as income, commute times, etc), but this information is never provided "together" – as in, here's the city, all of it's `region-id`'s, and extensive demographic data about those `region-id`'s / cities.

You can try to figure out out how to join all that data from disparate Zillow API's... or you can just use this package.

1.4.4 Step 4: Do some cool analysis

You got this one covered...

1.4.5 Step 5: Write me an email and tell me you love me

1.5 To Do:

- [STRIKEOUT:Complete package installation so package can be installed]
- [STRIKEOUT:Finish `README.md`]
- Generate documentation with Sphynx

api-scrapin

Indices and tables

- genindex
- modindex
- search