

### JSON Files in Python

# JSON Files in Python

## Quizzes

Before you download the movie poster images to add to the word cloud as described at the end of the video above, let's first get comfortable with accessing *wptools* page object attributes. Let's inspect the *wptools* page object for the E.T. The Extra-Terrestial Wikipedia page. In the Jupyter Notebook below, you will access the *images* and *infobox* attributes and the data within them.



#### JSON Files in Python

```
In [2]: | page = wptools.page('E.T._the_Extra-Terrestrial').get()
en.wikipedia.org (query) E.T._the_Extra-Terrestrial
API warning: {'extracts': {'warnings': 'HTML may be malformed an
d/or unbalanced and may omit inline images. Use at your own risk.
Known problems are listed at https://www.mediawiki.org/wiki/Exten
sion:TextExtracts#Caveats.'}} (https://www.mediawiki.org/wiki/Ext
ension:TextExtracts#Caveats.'}})
en.wikipedia.org (parse) 73441
www.wikidata.org (wikidata) Q11621
www.wikidata.org (labels) Q1748409|Q102427|P3219|P3302|P3995|Q13
www.wikidata.org (labels) Q1044183|P3077|Q981030|Q586356|P135|P3
www.wikidata.org (labels) Q8877 | P905 | P2508 | Q952914 | P2638 | P3650 | Q
www.wikidata.org (labels) Q181508|P1562|P1970|P18|Q4376972|Q3936
en.wikipedia.org (restbase) /page/summary/E.T._the Extra-Terrestr
en.wikipedia.org (imageinfo) File:ET logo 3.svg File:E t the ext
E.T. the Extra-Terrestrial (en) data
  aliases: <list(2)> E.T., ET
  claims: <dict(84)> P1562, P57, P272, P345, P31, P161, P373, P48
  description: 1982 American science fiction film
  exhtml: <str(606)> <i><b>E.T. the Extra-Terrestrial</b></i>
  exrest: <str(575)> E.T. the Extra-Terrestrial is a 1982 America
  extext: <str(1803)> **E.T. the Extra-Terrestrial** is a 1982
  extract: <str(1899)> <i><b>E.T. the Extra-Terrestrial</b></i
  image: <list(4)> {'kind': 'parse-image', 'file': 'File:E t the
  infobox: <dict(19)> name, image, alt, caption, director, produc
e...
```

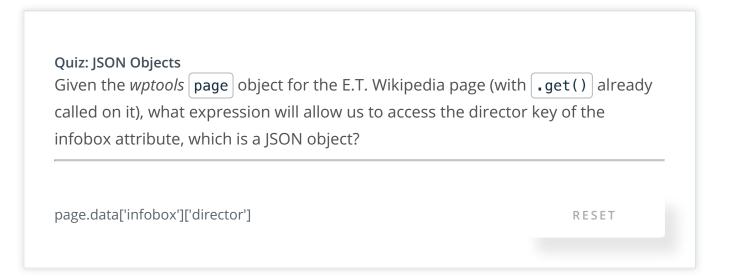
AMENU []EXPAND

#### **Quiz: JSON Arrays**

Given the *wptools* page object for the E.T. Wikipedia page (with get() already called on it), what expression will allow us to access the first image in the images attribute, which is a JSON array?



#### JSON Files in Python



## More JSON in Python

For the example in this lesson, JSON data was sourced from an API. That isn't always the case, though! Sometimes you're given a text file with human readable JSON within it. For this situation, the *json* library is indispensable. It can parse JSON *from* strings or files and it can parse JSON *into* a Python dictionary or list. It can also convert Python dictionaries or lists into JSON strings. The tutorial on the linked documentation page is handy. This Reading and Writing JSON to a File in Python article from Stack Abuse is also great, which outlines <code>json.dump</code>, <code>json.dumps</code>, <code>json.load</code>, and <code>json.loads</code> (four key *json* library methods) well.

pandas also has JSON functions (the read\_json function and the to\_json DataFrame method), but the hierarchical advantage of JSON is wasted in pandas' tabular DataFrame so the uses are limited.

NEXT