Working with Known JSON Schemas - Lab

Introduction

In this lab you'll practice working with json files whose schema you know beforehand.

Objectives

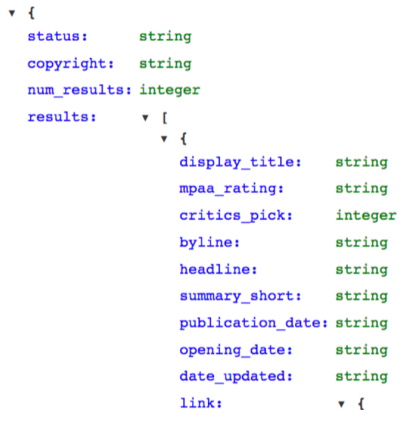
You will be able to:

- Read JSON Documentation Schemas and translate into code
- Extract data from known json schemas
- Write data to predefined JSON schemas

Reading a JSON Schema

Here's the JSON schema provided for a section of the NY Times API:

or a fully expanded view:



You can see this yourself here:
https://developer.nytimes.com/movie_reviews_v2.json#/Documentation/GET/critics/%7Bresource-type%7D.json

You can see that the master structure is a dictionary and has a key named 'response'. This is also a dictionary and has two keys: 'data' and 'meta'. As you continue to examine the schema hierarchy, you'll notice the vast majority in this case are dictionaries.

Loading the Data File

Start by importing the json file. The sample response from the api is stored in a file **ny_times_movies.json**

In [1]:
```
#Your code here
```

In [2]:
```
import json
```

In [3]:
```
f = open('ny_times_movies.json', 'r')
data = json.load(f)
```

Loading Specific Data

Create a DataFrame of the major data container within the json file, listed under the 'results' heading in the schema above.

In [4]:
```
#Your code here
import pandas as pd
df = pd.DataFrame(data['results'])
df.head()
```

Out[4]:

	byline	critics_pick	date_updated	display_title	headline	link	mpaa_rating	multim
0	A.O. SCOTT	1	2018-10-17 02:44:23	Can You Ever Forgive Me	Review: Melissa McCarthy Is Criminally Good in...	{'type': 'article', 'url': 'http://www.nytimes...	R	{' 'mediumThreeByTwo 'src': 'htt
1	BEN KENIGSBERG	1	2018-10-16 11:04:03	Charm City	Review: 'Charm City' Vividly Captures the Stre...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
2	GLENN KENNY	1	2018-10-16 11:04:04	Horn from the Heart: The Paul Butterfield Story	Review: Paul Butterfield's Story Is Told in 'H...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
3	A. O. SCOTT	0	2018-10-16 16:08:03	The Price of Everything	Review: 'The Price of Everything' Asks $56 Bil...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
4	BEN KENIGSBERG	0	2018-10-16 11:04:03	Impulso	Review: 'Impulso' Goes Backstage With a Flamen...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt

How many unique critics are there?

In [5]:
```
#Your code here
df.byline.nunique()
```

Out[5]: 8

Create a new column for the review's url. Title the column 'review_url'

In [6]:
```
#Your code here
df['review_url'] = df['link'].map(lambda x : x['url'])
df.head()
```

Out[6]:

	byline	critics_pick	date_updated	display_title	headline	link	mpaa_rating	multim
0	A.O. SCOTT	1	2018-10-17 02:44:23	Can You Ever Forgive Me	Review: Melissa McCarthy Is Criminally Good in...	{'type': 'article', 'url': 'http://www.nytimes...	R	{' 'mediumThreeByTwo 'src': 'htt
1	BEN KENIGSBERG	1	2018-10-16 11:04:03	Charm City	Review: 'Charm City' Vividly Captures the Stre...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
2	GLENN KENNY	1	2018-10-16 11:04:04	Horn from the Heart: The Paul Butterfield Story	Review: Paul Butterfield's Story Is Told in 'H...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
3	A. O. SCOTT	0	2018-10-16 16:08:03	The Price of Everything	Review: 'The Price of Everything' Asks $56 Bil...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt
4	BEN KENIGSBERG	0	2018-10-16 11:04:03	Impulso	Review: 'Impulso' Goes Backstage With a Flamen...	{'type': 'article', 'url': 'http://www.nytimes...		{' 'mediumThreeByTwo 'src': 'htt

How many results are in the file?

In [7]:
```
#Your code here
data['num_results']

#alternative solution:
#Len(df)
```

Out[7]: 20

Summary

Well done! Here you continued to gather practice extracting data from JSON files and transforming them into our standard tool of Pandas DataFrames.